

С.ЗИ
ФИЗИКА
ПОЛУПРОВОДНИКОВЫХ
ПРИБОРОВ

2

**Physics of
Semiconductor
Devices**

Second Edition

S. M. Sze

Bell Laboratories,
Incorporated
Murray Hill, New Jersey

A Wiley-Interscience
publication
John Wiley & Sons
New York · Chichester · Brisbane
Toronto · Singapore
1981

С. Зи Физика полупроводниковых приборов

В 2-х книгах

2

Перевод с английского
канд. физ.-мат. наук В. А. Гергеля,
канд. физ.-мат. наук Н. В. Зыкова,
д-ра физ.-мат. наук Б. И. Фукса
и канд. физ.-мат. наук Р. З. Хафизова
под редакцией
д-ра физ.-мат. наук Р. А. Суриса

ББК 32.852
3-59
УДК 621.382

Зи С.

3-59 Физика полупроводниковых приборов: В 2-х книгах.
Кн. 2. Пер. с англ. — 2-е перераб. и доп. изд. — М.: Мир,
1984. — 456 с., ил.

Монография написана известным американским специалистом в области полупроводниковой электроники. В переводе на русский язык выходит в двух книгах. Книга 2 посвящена физике приборов на туннельном эффекте и оптоэлектронных устройств (светодиодов, лазеров, фотодетекторов и солнечных батарей).

Для научных работников и инженеров, работающих в области электроники и вычислительной техники, а также студентов старших курсов вузов.

3 $\frac{2403000000-431}{041(01)-84}$ 157—85, ч. 1

ББК 32.852
6Ф 0.32

Редакция литературы по новой технике

- © 1981, by John Wiley and Sons, Inc.
All rights reserved. Authorized translation
from English language edition published
by John Wiley and Sons, Inc.
- © Перевод на русский язык, «Мир», 1984

8.1. ВВЕДЕНИЕ

Полевой транзистор со структурой металл—окисел—полупроводник (МОП-транзистор)¹ в настоящее время является основным элементом сверхбольших интегральных схем (СБИС), таких, как микропроцессоры и полупроводниковые запоминающие устройства. В последнее время они находят широкое применение и в мощных переключающих схемах. Принцип работы поверхностного полевого транзистора впервые предложен Лилиенфельдом [1] и Хейлом [2] в начале 30-х годов. Позднее, в конце 40-х годов, теорию работы этих приборов развивали Шокли и Пирсон [3]. В 1960 г. Канг и Аталла [4] изготовили первый кремниевый МОП-транзистор, используя термическое окисление. Основные характеристики МОП-транзисторов затем исследовали Ихантола и Молл [5, 6], Са [7], Хофштейн и Хейман [8]. Для ознакомления с современным состоянием технологии, применений и физики МОП-транзисторов можно рекомендовать обзорные статьи Уолмарка и Джонсона [9], Ричмана [10] и Брюса [11].

Поскольку ток в МОП-транзисторах переносится в основном носителями только одного типа (например, электронами в *n*-канальных приборах), МОП-транзисторы обычно относят к классу униполярных приборов. Работа других приборов этого класса — полевых транзисторов с *p*—*n*-переходом в качестве затвора (JFET) и полевых транзисторов с затвором Шоттки (MESFET) — была рассмотрена в гл. 6. Хотя к настоящему времени известны МОП-транзисторы, изготовленные на различных полупроводниковых материалах, таких, как Ge [12], Si и GaAs [13], с использованием различных диэлектрических слоев (SiO_2 , Si_3N_4 и Al_2O_3), наибольшее распространение получили МОП-транзисторы на Si—SiO₂.

Сначала мы рассмотрим основные характеристики обычных, так называемых длинноканальных МОП-транзисторов, в которых

¹) Приборы этого типа в зарубежной литературе носят разные названия, в том числе MOSFET (metal — oxide — semiconductor field-effect transistor, т. е. полевой транзистор металл — окисел — полупроводник), IGFET (insulated-gate field-effect transistor, т. е. полевой транзистор с изолированным затвором), MISFET (metal — insulator — semiconductor field-effect transistor, т. е. полевой транзистор металл — изолятор — полупроводник), MOST (metal — oxide — semiconductor transistor, т. е. транзистор металл — окисел — полупроводник). В отечественной литературе наиболее часто используются термины «МОП-транзистор» и «МДП-транзистор». — Прим. перев.

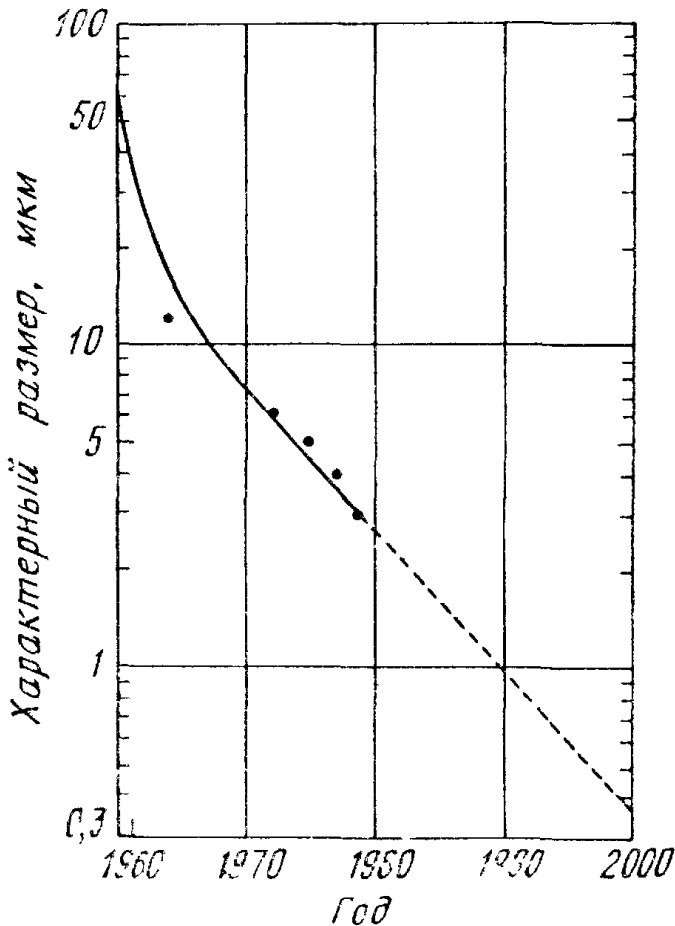


Рис. 1. Зависимость характерного размера отдельного прибора в интегральных схемах коммерческого назначения от времени [14].

длина канала L значительно превышает сумму толщин обедненных слоев стока и истока $(W_S + W_D)^1$. Это поможет также понять специфику работы так называемых короткоканальных приборов, в которых $L \leq (W_S + W_D)$. Данные, представленные на рис. 1, показывают, что с 1959 г. — «начала эры

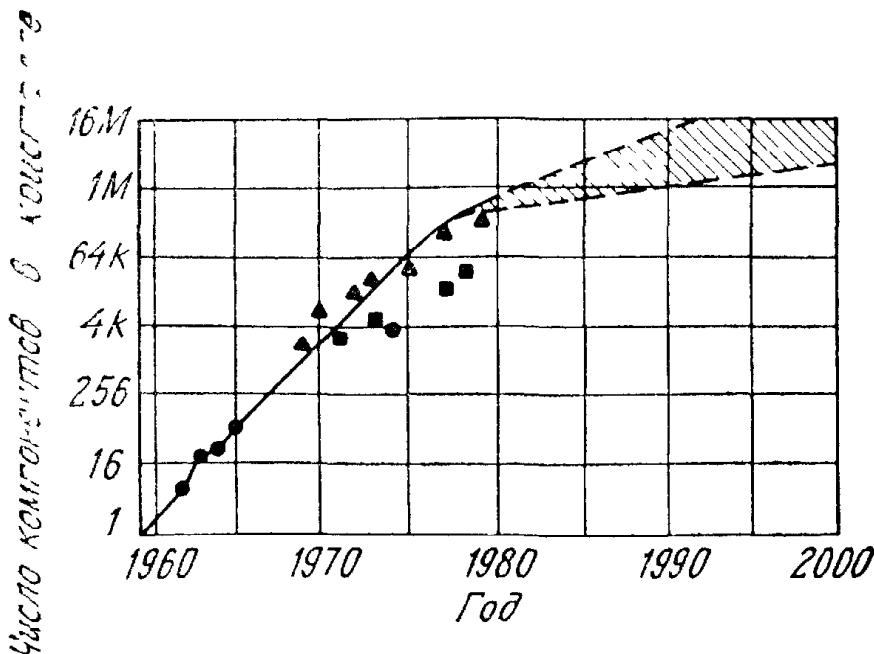


Рис. 2. Зависимость сложности интегральных схем от времени [15].

▲ — МОП-память; ■ — МОП-логика; ● — биполярная логика.

¹ Эти термины определены в разд. 8.2.

интегральных схем» — и до настоящего времени характерный размер отдельного элемента уменьшался, а число компонентов в одном кристалле соответственно увеличивалось экспоненциально во времени [14]. Ожидается, что «скорость» этих процессов в ближайшее время должна существенно снизиться из-за возрастающих трудностей конструирования и производства интегральных схем такого большого объема. Однако есть основания полагать, что к 1990 г. сложность интегральных схем будет составлять 10^6 и более компонентов на кристалл при характерном размере отдельного элемента ≤ 1 мкм (рис. 2) [15]. На рабочие характеристики МОП-транзисторов таких малых размеров могут существенно влиять так называемые короткоканальные эффекты: двумерный характер распределения потенциала под затвором, насыщение дрейфовой скорости носителей в сильных электрических полях, зарядка окисла и т. д.

В настоящее время для улучшения тех или иных рабочих характеристик предложено довольно большое число различных структурных модификаций МОП-транзисторов. Особенности работы некоторых, наиболее представительных из них описаны ниже. Мы рассмотрим также работу специальных МОП-транзисторных структур с многослойным затвором, которые используются в настоящее время в качестве основного элемента энергонезависимых полупроводниковых запоминающих устройств.

8.2. ОСНОВНЫЕ ХАРАКТЕРИСТИКИ

Основная структура полевого транзистора металл—окисел—полупроводник (МОП-транзистора) показана на рис. 3. Этот четырехполюсный прибор состоит из полупроводниковой подложки

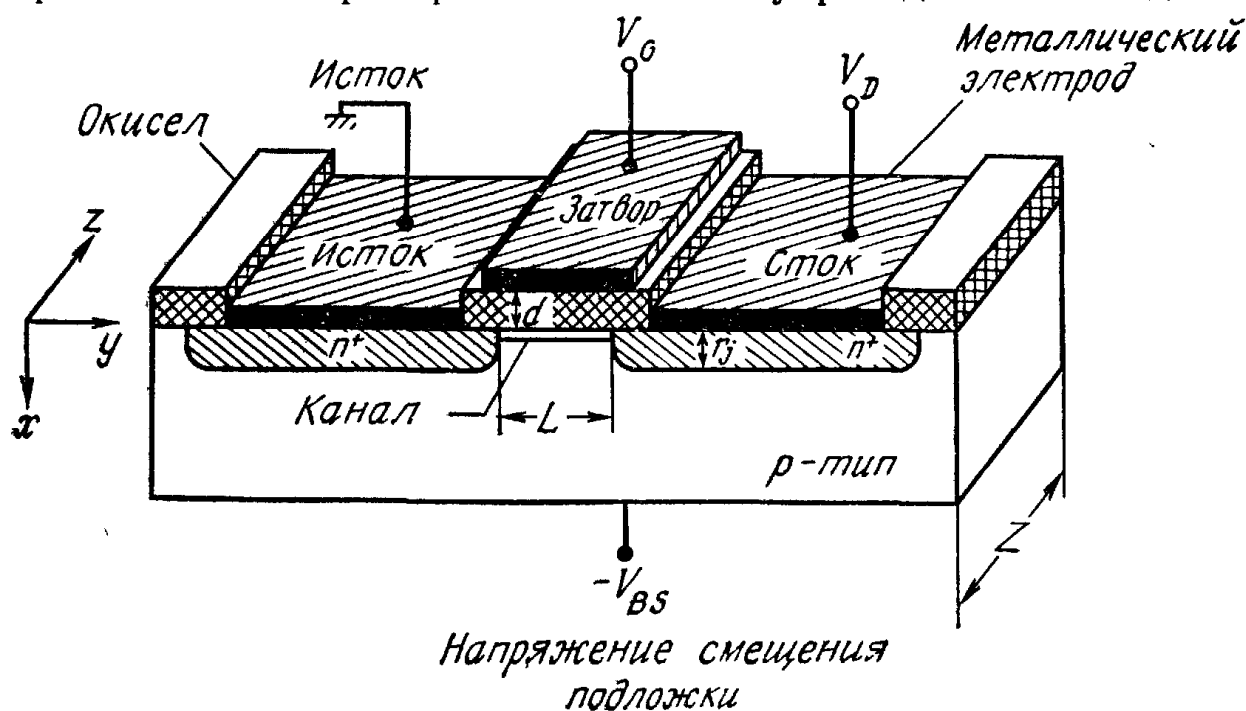


Рис. 3. Схема структуры МОП-транзистора [4].

p -типа, в которой сформированы (например, с помощью ионной имплантации) две высоколегированные n^+ -области — сток и исток¹. Металлический электрод, отделенный от подложки слоем окисла, называется затвором. В последнее время в качестве затвора используются высоколегированный поликремний, а также некоторые комбинации поликремния и силицидов. Основными параметрами структуры являются длина канала L — расстояние между металлургическими границами n^+ — p -переходов стока и истока, ширина канала Z , толщина слоя изолятора d , глубина переходов r_j и уровень легирования подложки N_A . В кремниевых интегральных схемах отдельный МОП-транзистор окружен в целях изоляции областью с толстым слоем окисла, который называется пассивирующим или полевым (в отличие от тонкого слоя подзатворного окисла).

Рассматривая работу прибора, мы будем все напряжения отсчитывать от потенциала истока, т. е. считать его заземленным. Когда напряжение на затворе отсутствует, электрическая цепь исток — сток представляет собой два n^+ — p -перехода, включенных навстречу друг другу. При этом в ней может протекать очень малый ток, равный току утечки обратносмещенного перехода². Если же к затвору приложено достаточно большое положительное напряжение, у границы с диэлектриком образуется инверсионный слой или канал, соединяющий n^+ -области стока и истока. Проводимость этого инверсионного канала модулируется при изменении напряжения на затворе. Тыловой контакт к подложке может находиться либо под тем же опорным потенциалом, что и исток, либо под напряжением, соответствующим обратному смещению n^+ — p -перехода исток—подложка. Напряжение обратного смещения подложки также влияет на проводимость инверсионного канала

8.2.1. Неравновесные условия

Когда между стоком и истоком транзистора приложено напряжение, МОП-структура оказывается в неравновесных условиях. При этом квазиуровень Ферми неосновных носителей (в данном случае электронов) оказывается ниже равновесного уровня Ферми. На рис. 4 приведены двумерные зонные диаграммы, соответствующие различным режимам работы МОП-транзистора, который для удобства показан развернутым на 90° (рис. 4, а) [16]. Зонная

¹ Это так называемый n -канальный прибор. Для анализа работы p -канального МОП-транзистора следует просто поменять местами индексы n и p , а также изменить полярность приложенных напряжений.

² Здесь рассматривается работа нормально закрытого n -канального МОП-транзистора (обогащенного типа). Работа МОП-транзисторов других типов описана в последующих разделах.

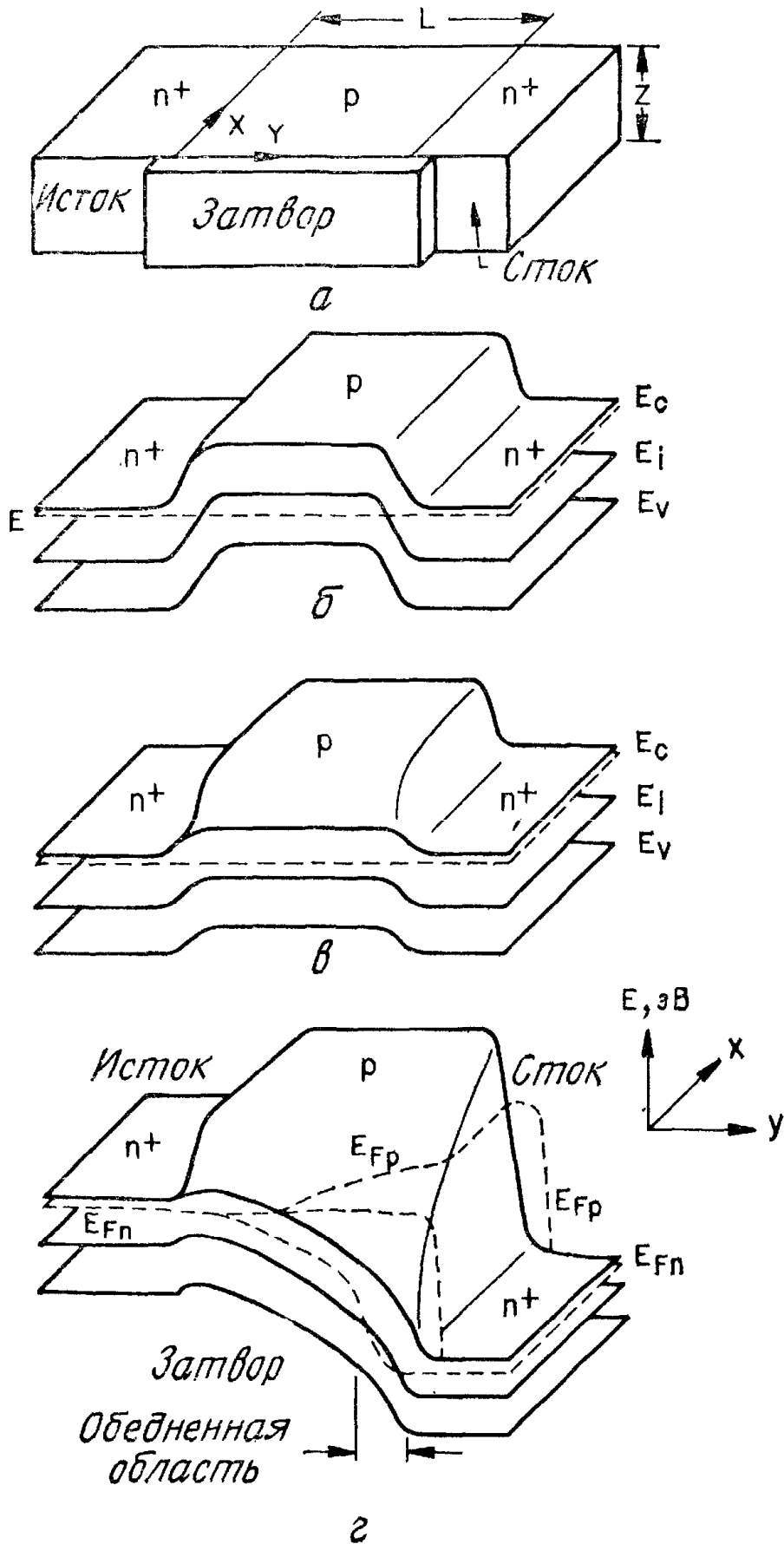


Рис. 4. Двумерная зонная диаграмма n -канального МОП-транзистора [16].
 а — конфигурация прибора; б — равновесная зонная диаграмма (плоские зоны, нулевое смещение подложки); в — равновесная зонная диаграмма при инвертирующем смещении на затворе; г — неравновесная ситуация (напряжения поданы на затвор и сток структуры).

схема на рис. 4, б соответствует равновесным условиям ($V_D = 0$), отсутствию обратного смещения на подложке ($V_{BS} = 0$) и нулевому напряжению на затворе ($V_G = 0$, плоские зоны). На рис. 4, в показана также равновесная ситуация, но уже с инвертированной границей раздела ($V_G > 0$). Неравновесная ситуация, когда соответствующие положительные напряжения приложены и к затвору, и к стоку, показана на рис. 4, г. Штриховыми линиями показаны дырочный и электронный квазиуровни Ферми. В то время как дырочный квазиуровень Ферми E_{FP} не зависит от координат (остается равным равновесному уровню Ферми в объеме подложки), электронный квазиуровень Ферми E_{FN} уменьшается по мере приближения к стоку. Из рис. 4, г видно также, что в неравновесных условиях, для того чтобы инвертировать поверхность вблизи стока, требуется большее напряжение на затворе, чем в равновесном случае, когда $\psi_s(\text{inv}) \simeq 2\psi_B$. Это обусловлено тем, что напряжение на стоке понижает электронный квазиуровень Ферми, а инверсионный слой на поверхности может образоваться только в том случае, когда поверхностный потенциал, отсчитанный от собственного уровня энергии E_i , пересекает квазиуровень Ферми неосновных носителей.

На рис. 5 приведены для сравнения пространственные распределения зарядов и зонные диаграммы в примыкающей к стоку части p -области в равновесных и неравновесных условиях [17]. В первом случае, который подробно рассмотрен в гл. 7, при инверсии толщина обедненной области достигает своего максимального равновесного значения W_m . В неравновесном случае толщина обедненного слоя зависит от напряжения на стоке V_D , а поверхностный потенциал, соответствующий началу сильной инверсии, с достаточной степенью точности можно считать равным

$$\psi_s(\text{inv}) \simeq V_D + 2\psi_B. \quad (1)$$

Соотношения, устанавливающие связь между зарядом инверсионного слоя, зарядом в обедненной области и напряжением на затвор для неравновесного случая выводятся аналогично тому, как это сделано в гл. 7 для равновесного случая. При этом, однако, используются еще два дополнительных предположения: 1) квазиуровни Ферми основных и неосновных носителей не зависят от расстояния до границы раздела; 2) разность между этими квазиуровнями равна напряжению на стоковом переходе, т. е. $E_{FP} = E_{FN} + qV_D$ (для p -подложки). Первое предположение в условиях инверсии приводит к несущественным ошибкам, поскольку в этом случае заряд основных носителей составляет пренебрежимо малую часть полного пространственного заряда в приповерхностном слое. Второе предположение справедливо только в условиях сильной инверсии, когда заряд инверсионного слоя составляет

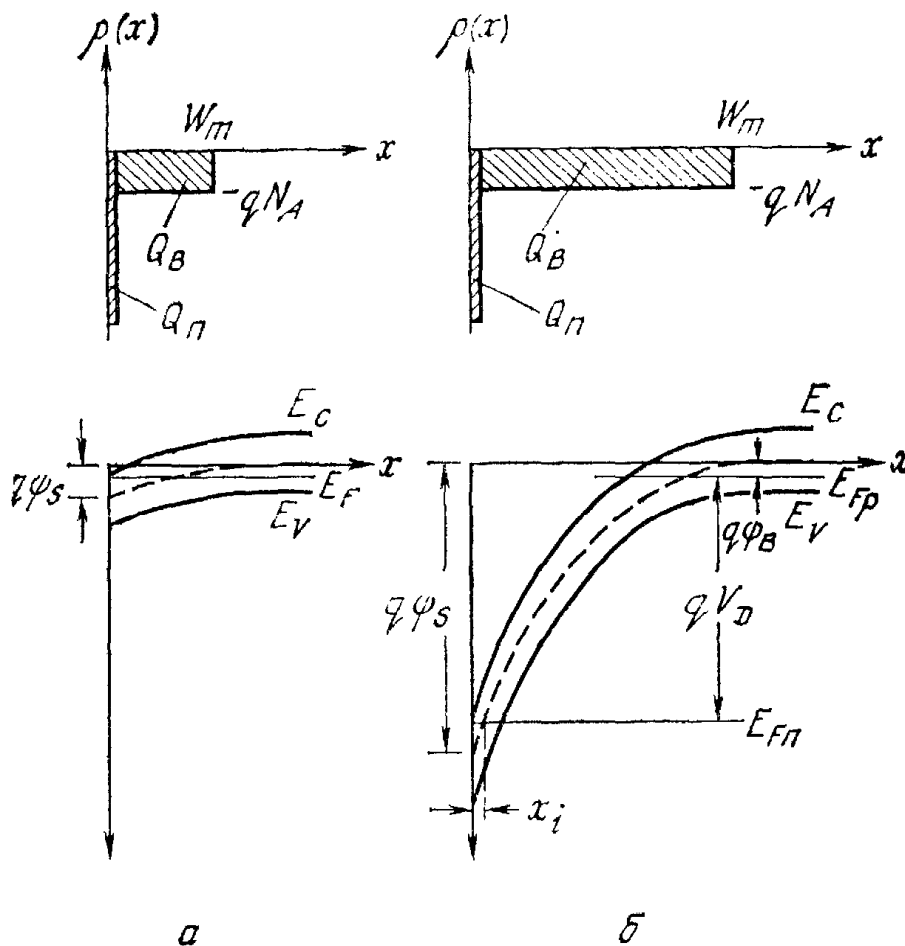


Рис. 5. Распределение зарядов и зонные диаграммы в инвертированной части p -подложки [17].

a — при равновесных условиях; b — неравновесные условия вблизи стока.

существенную часть полного заряда приповерхностной области полупроводниковой подложки.

Основываясь на этих предположениях, запишем одномерное уравнение Пуассона для области пространственного заряда в прилегающей к стоку части p -подложки в виде

$$\frac{\partial^2 \psi}{\partial x^2} = -\frac{q}{\epsilon_s} (N_D^+ - N_A^- + p - n), \quad (2)$$

где, как и в гл. 7,

$$N_D^+ - N_A^- = n_{p0} - p_{p0}, \quad p_{p0} \simeq N_A, \quad p = p_{p0} e^{-\beta \psi},$$

$$\text{но } n = n_{p0} e^{\beta \psi - \beta V_D}, \quad \beta \equiv q/kT. \quad (3)$$

Тем же способом, что и в гл. 7, получим отсюда

$$\mathcal{E} = -\frac{\partial \psi}{\partial x} = \pm \frac{\sqrt{2kT}}{qL_D} F\left(\beta \psi, V_D, \frac{n_{p0}}{p_{p0}}\right) \quad (4)$$

и

$$Q_s = -\epsilon_s \mathcal{E}_s = \mp \frac{\sqrt{2\epsilon_s kT}}{qL_D} F\left(\beta \psi_s, V_D, \frac{n_{p0}}{p_{p0}}\right), \quad (5)$$

где

$$F\left(\beta\psi, V_D, \frac{n_{p0}}{p_{p0}}\right) \equiv \left[e^{-\beta\psi} + \beta\psi - 1 + \frac{n_{p0}}{p_{p0}} e^{-\beta V_D} (e^{\beta\psi} - \beta\psi e^{\beta V_D} - 1) \right]^{1/2}, \quad (6)$$

а

$$L_D \equiv \left(\frac{kT\epsilon_s}{p_{p0}q^2} \right)^{1/2}. \quad (7)$$

В режиме сильной инверсии полный поверхностный заряд (5) можно представить в виде

$$Q_s = Q_n + Q_B, \quad (8)$$

где

$$Q_B = -qN_A W_m = -\sqrt{2qN_A\epsilon_s(V_D + 2\psi_B)} \quad (9)$$

и Q_n заряда, обусловленный неосновными носителями в инверсионном слое:

$$|Q_n| \equiv q \int_0^{x_i} n(x) dx = q \int_{\psi_s}^{\psi_B} \frac{n(\psi) d\psi}{d\psi/dx}, \quad (10)$$

$$|Q_n| = q \int_{\psi_s}^{\psi_B} \frac{n_{p0} e^{(\beta\psi - \beta V_D)} d\psi}{(\sqrt{2} kT/qL_D) F(\beta\psi, V_D, n_{p0}/p_{p0})}. \quad (11)$$

Здесь x_i — точка, в которой собственный уровень Ферми пересекает квазиуровень Ферми неосновных носителей — электронов. Для используемых на практике уровней легирования кремния величина x_i довольно мала (30—300 Å). Выражение (11) является одним из основных соотношений теории МОП-транзистора с длинным каналом, и в конкретных ситуациях соответствующая зависимость $Q_n(\psi_s)$ рассчитывается численными методами. Более простое и удобное выражение для Q_n , справедливое при достаточно сильной инверсии, дает так называемая модель заряженных слоев [18]:

$$|Q_n| = \sqrt{2}qN_AL_D \left\{ \left[\beta\psi_s + \left(\frac{n_{p0}}{p_{p0}} \right) e^{(\beta\psi_s - \beta V_D)} \right]^{1/2} - (\beta\psi_s)^{1/2} \right\}. \quad (12)$$

Выражения (11) и (12) для Q_n выведены для случая, когда обратное смещение на подложке отсутствует ($V_{BS} = 0$). При наличии такого смещения при расчете Q_n следует подставить $\beta(V_D + V_{BS})$ в формулу (12) вместо βV_D .

8.2.2. Линейная область характеристик. Область насыщения

Рассмотрим сначала качественно работу МОП-транзистора. Будем считать, что напряжение на затворе структуры G достаточно для сильной инверсии на границе с окислом (рис. 6, а). Тогда, если напряжение на стоке V_D не слишком велико, инвер-

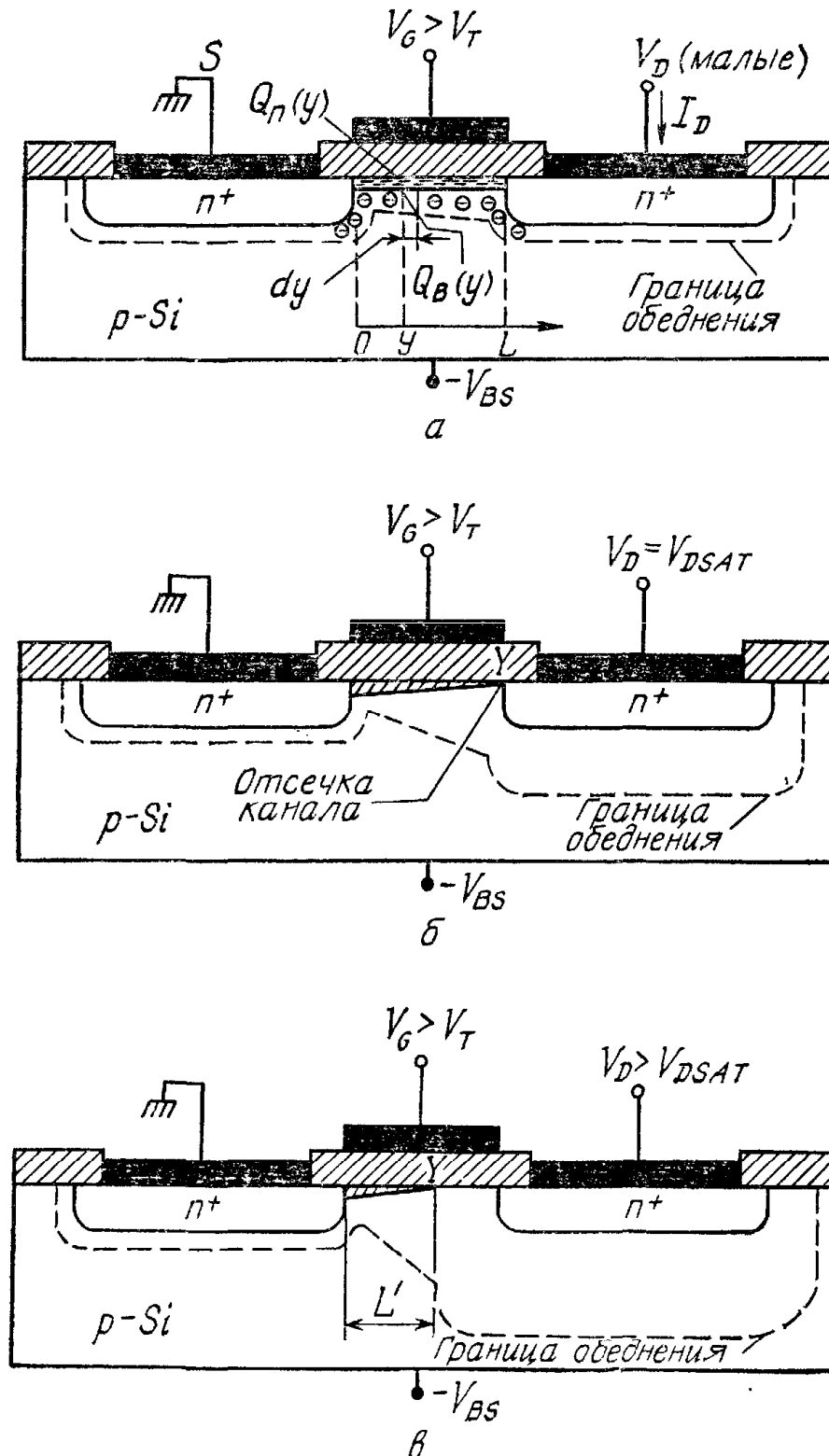


Рис. 6. МОП-транзистор в линейном режиме (малые напряжения на стоке) (а), в начале насыщения (отсечка канала на границе со стоком $Y = L$) (б) и в режиме насыщения (показано уменьшение эффективной длины канала) (в).

сионный слой действует как обычное сопротивление, и ток через проводящий инверсионный канал I_D будет увеличиваться пропорционально напряжению стока V_D . Эта область называется линейной областью работы прибора. С увеличением напряжения на стоке в конце концов достигается момент, когда ширина канала x_i , а следовательно, и заряд инверсионного слоя Q_n в точке $y = L$ (прямо на границе стока) становится равным нулю. Эти условия соответствуют началу режима отсечки канала (рис. 6, б). Напряжение стока в данном режиме обозначим $V_{D\text{sat}}$. При больших напряжениях ($V_D > V_{D\text{sat}}$) точка отсечки Y сдвигается к истоку, а ток стока при этом увеличивается лишь незначительно (рис. 6, в), поскольку напряжение между истоком и точкой отсечки канала Y по-прежнему равно $V_{D\text{sat}}$. Инжекция неосновных носителей из точки Y в обедненную область стока подобна процессу инжекции носителей переходом эмиттер—база биполярного транзистора в обедненную область его коллекторного перехода.

Теперь выведем аналитические соотношения, определяющие основные характеристики МОП-транзисторов, используя при этом следующие допущения: 1) подзатворная область прибора представляет собой идеальную МОП-структуру, что в соответствии с определением этого понятия, данным в гл. 7, означает отсутствие поверхностных состояний, фиксированного заряда в окисле, разности работ выхода между полупроводником и металлом и т. д.; 2) учитывается только дрейфовая компонента тока; 3) подвижность носителей в инверсионном слое постоянна (не зависит от электрического поля); 4) концентрация легирующей примеси в канале однородна; 5) обратные токи утечки p — n -переходов пренебрежимо малы; 6) поперечное электрическое поле \mathcal{E}_x , направленное перпендикулярно границе раздела (по оси x), значительно превышает продольное электрическое поле \mathcal{E}_y , направленное вдоль поверхности (по оси y). Последнее допущение называется приближением плавного канала.

В соответствии с принятыми допущениями поверхностная плотность полного заряда в полупроводнике на расстоянии y от истока и поверхностный потенциал $\psi_s(y)$ в той же точке связаны следующим соотношением:

$$Q_s(y) = [-V_G + \psi_s(y)] C_i, \quad (13)$$

где $C_i \equiv \epsilon_i/d$ — удельная емкость окисла. Используя выражение (8) для заряда инверсионного слоя, получим

$$Q_n(y) = Q_s(y) - Q_B(y) = -[V_G - \psi_s(y)] C_i - Q_B(y). \quad (14)$$

В режиме сильной инверсии поверхностный потенциал ψ_s равен $2\psi_B + V(y)$, где $V(y)$ — напряжение обратного смещения

в точке y по отношению к потенциалу истока (который мы считаем заземленным). Как уже говорилось выше, заряд обедненного слоя

$$Q_B(y) = -qN_A W_m = -\sqrt{2\epsilon_s q N_A [V(y) + 2\psi_B]}. \quad (15)$$

Подставив это выражение в выражение (14), получим

$$Q_n(y) = -[V_G - V(y) - 2\psi_B] C_i + \sqrt{2\epsilon_s q N_A [V(y) + 2\psi_B]}. \quad (16)$$

Проводимость канала представляет собой интеграл

$$g = \frac{Z}{L} \int_0^{x_i} \sigma(x) dx \quad (17)$$

от объемной удельной проводимости

$$\sigma(x) = qn(x) \mu_n(x). \quad (18)$$

Полагая, что подвижность μ не зависит от x , получим

$$g = \frac{qZ\mu_n}{L} \int_0^{x_i} n(x) dx = qZ\mu_n |Q_n|/L. \quad (19)$$

Сопротивление элементарного участка канала длиной dy есть

$$dR = \frac{dy}{gL} = \frac{dy}{Z\mu_n |Q_n(y)|} \quad (20)$$

и, следовательно, падение напряжения на этом участке

$$dV = I_D dR = \frac{I_D dy}{Z\mu_n |Q_n(y)|}, \quad (21)$$

где ток стока I_D — константа, не зависящая от y . Умножив выражение (16) на выражение (21) и проинтегрировав получившееся выражение от истока ($y = 0$, $V = 0$) до стока ($y = L$, $V = V_D$), для рассматриваемого идеализированного случая найдем

$$I_D = \frac{Z}{L} \mu_n C_i \left\{ \left(V_G - 2\psi_B - \frac{V_D}{2} \right) V_D - \frac{2}{3} \frac{\sqrt{2\epsilon_s q N_A}}{C_i} \left[(V_D + 2\psi_B)^{3/2} - (2\psi_B)^{3/2} \right] \right\}. \quad (22)$$

Из выражения (22) следует, что при фиксированном напряжении на затворе V_G ток стока I_D сначала линейно увеличивается с напряжением стока V_D (линейная область), затем крутизна характеристики плавно уменьшается до нуля, после чего начинается область насыщения. Семейство выходных ($I_D - V_D$) характеристик идеального МОП-транзистора приведено на рис. 7.

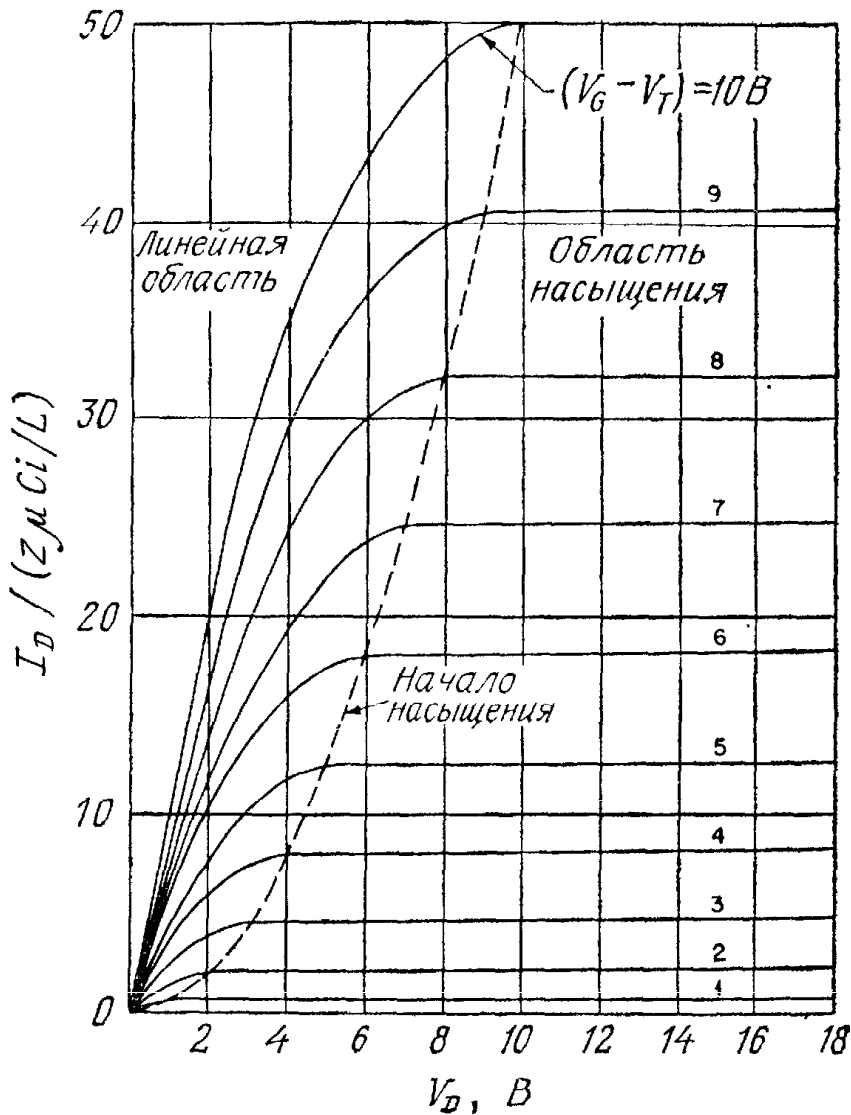


Рис. 7. Выходные (стоковые) характеристики идеального МОП-транзистора (I_D в зависимости от V_D). Штриховая линия соединяет точки начала насыщения на характеристиках, соответствующих различным смещениям на затворе транзистора ($J_{D \text{ sat}}$, $V_{D \text{ sat}}$). При $V_D > V_{D \text{ sat}}$ ток стока остается постоянным.

Рассмотрим названные выше две области более подробно. Для малых напряжений стока, разлагая выражение (22), получим

$$I_D \approx \frac{z}{L} \mu_n C_i \left[(V_G - V_T) V_D - \left(\frac{1}{2} + \frac{\sqrt{\varepsilon_s q N_A / \psi_B}}{4 C_i} \right) V_D^2 \right], \quad (23)$$

или еще проще (при $V_D \ll (V_G - V_T)$)

$$I_D \approx \frac{z}{L} \mu_n C_i (V_G - V_T) V_D, \quad (23a)$$

где

$$V_T = 2\psi_B + \frac{\sqrt{2\varepsilon_s q N_A (2\psi_B)}}{C_i} \quad (24)$$

— пороговое напряжение. Теоретические зависимости порогового напряжения от концентрации примеси и толщины изолятора для Si—SiO₂-систем были приведены в гл. 7 (рис. 14). Эксперимен-

тально пороговое напряжение определяют, линейно продолжая измеренную при малых V_D сток-затворную характеристику ($I_D - V_G$) до пересечения с осью напряжений.

Дифференцируя выражение (23а), получим для проводимости g_D и крутизны g_m идеального МОП-транзистора в линейной области

$$g_D \equiv \left. \frac{\partial I_D}{\partial V_D} \right|_{V_G = \text{const}} = \frac{Z}{L} \mu_n C_i (V_G - V_T), \quad (25)$$

$$g_m \equiv \left. \frac{\partial I_D}{\partial V_G} \right|_{V_D = \text{const}} = \frac{Z}{L} \mu_n C_i V_D. \quad (26)$$

По мере увеличения напряжения стока V_D заряд инверсионного слоя Q_n в точке $y = L$ (у края стока) уменьшается и, наконец, становится равным нулю при $V_D = V_{D \text{ sat}}$. Это явление по аналогии с подобным эффектом в полевых транзисторах с $p-n$ -переходом в качестве затвора называется отсечкой канала. При $V_D = V_{D \text{ sat}}$ ток стока достигает максимального значения $I_{D \text{ sat}}$. При больших напряжениях ($V_D > V_{D \text{ sat}}$) наступает насыщение ($I_D = I_{D \text{ sat}}$). Величину напряжения отсечки $V_{D \text{ sat}}$ получим, приравняв нулю выражение (16) с $V(L) = V_{D \text{ sat}}$:

$$V_{D \text{ sat}} = V_G - 2\psi_B + K^2 (1 - \sqrt{1 + 2V_G/K^2}), \quad (27)$$

где $K \equiv \sqrt{\epsilon_s q N_A / C_i}$. Ток насыщения $I_{D \text{ sat}}$ можно найти, подставив выражение (27) в выражения (22) и (23):

$$I_{D \text{ sat}} \simeq \frac{mZ}{L} \mu_n C_i (V_G - V_T)^2, \quad (28)$$

где коэффициент m в общем случае зависит от уровня легирования подложки. При сравнительно малых концентрациях примеси $m = 1/2$ [11]. Пороговое напряжение V_T в этом выражении в режиме насыщения для МОП-транзисторов с низколегированными подложками определяется также уравнением (24). В сильнолегированных структурах V_T оказывается зависящим от V_G . Крутизна МОП-транзистора в области насыщения в соответствии с выражением (28) равна

$$g_m = \left. \frac{\partial I_D}{\partial V_G} \right|_{V_D = \text{const}} = \frac{2mZ}{L} \mu_n C_i (V_G - V_T). \quad (29)$$

При выводе соотношений, определяющих основные характеристики МОП-транзисторов, мы использовали ряд упрощающих предположений. Попробуем теперь отказаться от первых двух предположений, т. е. проанализируем, какое влияние на характеристики МОП-транзистора оказывают неидеальность МОП-структуры и диффузионная компонента тока в канале. Как уже говорилось в гл. 7, фиксированный заряд окисла и разность

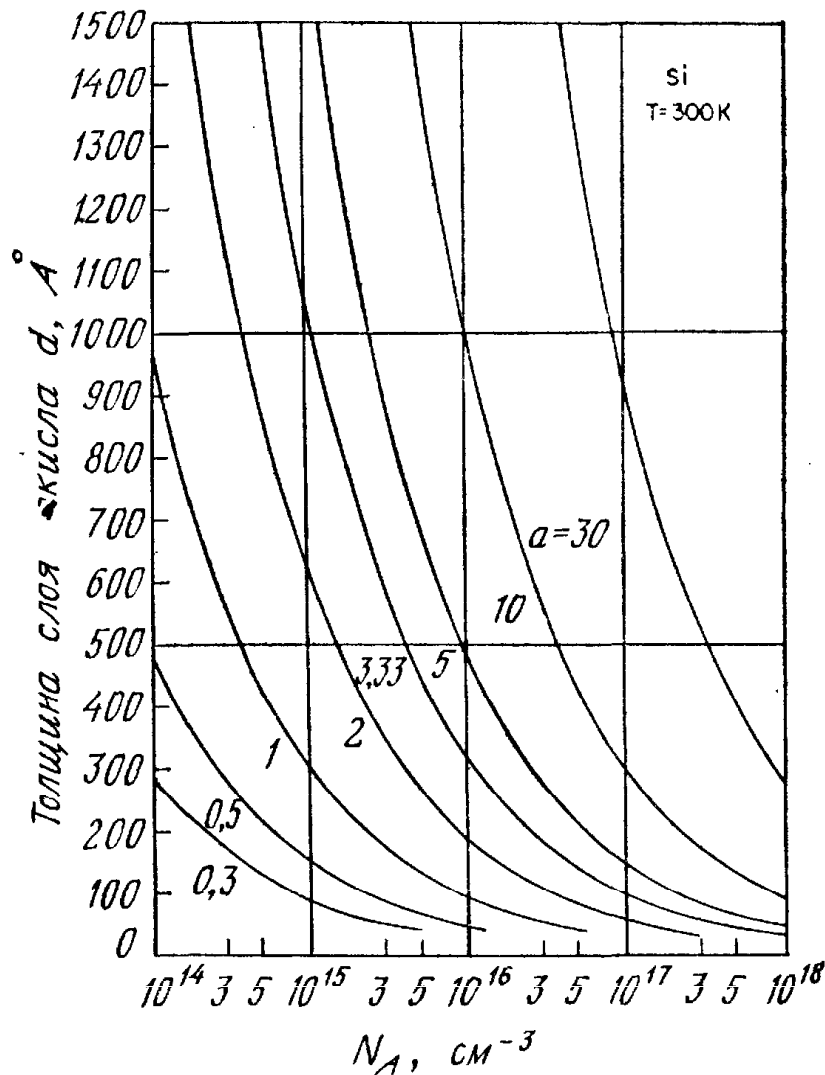


Рис. 8. Номограмма для определения значений безразмерного параметра a [19].

работ выхода электрона из металла и полупроводника обуславливают сдвиг напряжения плоских зон МОП-структуры V_{FB} . Ясно, что на эту же величину должно сдвинуться и пороговое напряжение МОП-транзистора:

$$\begin{aligned}
 V_T &= V_{FB} + 2\psi_B + \frac{\sqrt{2\varepsilon_s q N_A (2\psi_B)}}{C_i} = \\
 &= \left(\varphi_{ms} - \frac{Q_f}{C_i} \right) + 2\psi_B + \frac{\sqrt{4\varepsilon_s q N_A \psi_B}}{C_i}. \quad (30)
 \end{aligned}$$

На пороговое напряжение МОП-транзистора влияет также напряжение обратного смещения подложки V_{BS} :

$$V_T = V_{FB} + 2\psi_B + \sqrt{2\varepsilon_s q N_A (2\psi_B + V_{BS})} / C_i, \quad (31)$$

или

$$\begin{aligned}
 \Delta V_T &= V_T(V_{BS}) - V_T(V_{BS} = 0) = \\
 &= \frac{\sqrt{2\varepsilon_s q N_A}}{C_i} (\sqrt{2\psi_B + V_{BS}} - \sqrt{2\psi_B}) =
 \end{aligned}$$

$$= \frac{a}{\beta} (\sqrt{2\beta\psi_B + \beta V_{BS}} - \sqrt{2\beta\psi_B}), \quad (32)$$

где

$$a \equiv \sqrt{2} (\epsilon_s/L_D)/C_i = \sqrt{2} (\epsilon_s/\epsilon_i) (d/L_D). \quad (33)$$

При выполнении конкретных расчетов для нахождения a удобно использовать номограмму, приведенную на рис. 8 [19]. Семейство зависимостей сдвига порогового напряжения ΔV_T от смещения подложки V_{BS} для некоторых типичных значений параметра a приведено на рис. 9. Легко видеть, что чем больше a , тем больше ΔV_T .

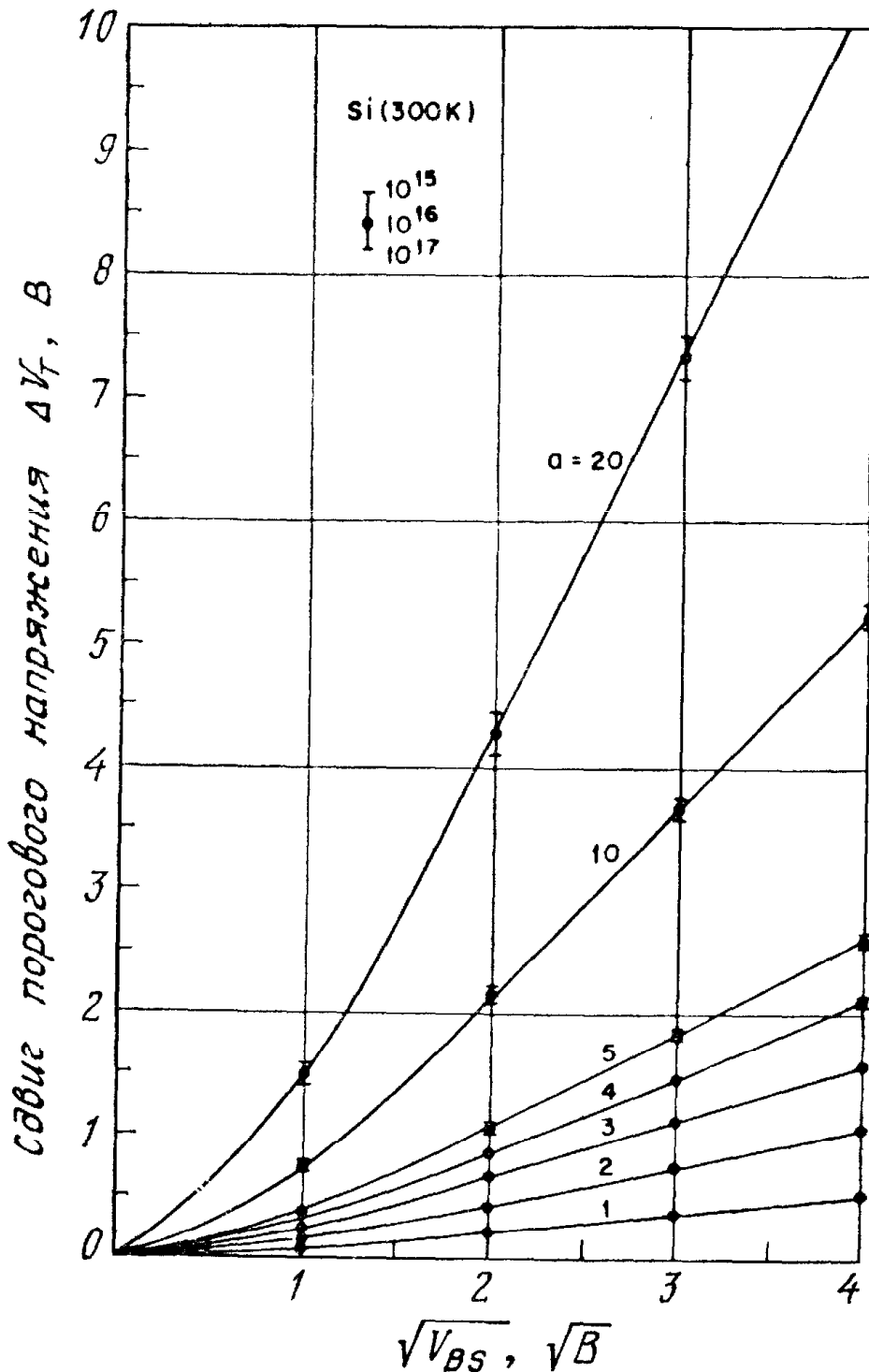


Рис. 9. Сдвиг порогового напряжения за счет обратного смещения подложки при различных значениях параметра a .

Для исследования влияния диффузионной компоненты тока канала рассмотрим двумерную зонную схему МОП-транзистора в неравновесных условиях (рис. 4, *з*), где, в частности, показано изменение положения электронного квазиуровня Ферми по длине канала. С помощью этого квазиуровня ψ_{Fn} плотность полного тока канала, включая обе его компоненты (диффузионную и дрейфовую), можно записать в виде

$$J_D(x, y) = q\mu_n n \mathcal{E}_y + qD_n \nabla n = -qD_n n(x, y) \nabla \psi_{Fn}. \quad (34)$$

Снова используя приближение плавного канала, из выражения (34) получим для полного тока

$$\begin{aligned} I_D &= \int_0^{x_i} J_D(x, y) Z dx = \\ &= \frac{1}{L} \int_0^L D_n q Z \left(\frac{\partial \psi_{Fn}}{\partial y} \right) \int_0^{x_i} n(x, y) dx dy = \\ &= \frac{Z}{L} \frac{\varepsilon_s \mu_n}{L_D} \int_0^{V_D} \int_{\psi_B}^{\psi_s} \frac{e^{\beta\psi - \beta V}}{F(\beta\psi, V, n_{p0}/p_{p0})} d\psi dV. \end{aligned} \quad (35)$$

Напряжение на затворе связано с поверхностным потенциалом ψ_s соотношением

$$\begin{aligned} V_G &= V_G - V_{FB} = -\frac{Q_s}{C_i} + \psi_s = \\ &= \frac{2\varepsilon_s kT}{C_i q L_D} F\left(\beta\psi_s, V, \frac{n_{p0}}{p_{p0}}\right) + \psi_s. \end{aligned} \quad (36)$$

Когда напряжение на затворе V_G существенно превышает пороговое напряжение V_T , выражение (35) переходит в приближенную формулу (22). Однако в окрестности порогового напряжения, а также вблизи отсечки канала точность формулы (22) недостаточна, и здесь следует использовать точное выражение (35). Для конкретных структур с определенными значениями физических параметров (размеров, концентрации примеси, подвижности) уравнения (35) и (36) решаются численными методами. Получаемые при этом зависимости справедливы при любых напряжениях на стоке (и в линейной области, и в области насыщения). В качестве примера на рис. 10 приведены результаты одного из таких расчетов [16], которые отчетливо демонстрируют эффект насыщения тока — типичное свойство выходных характеристик МОП-транзисторов с длинным каналом.

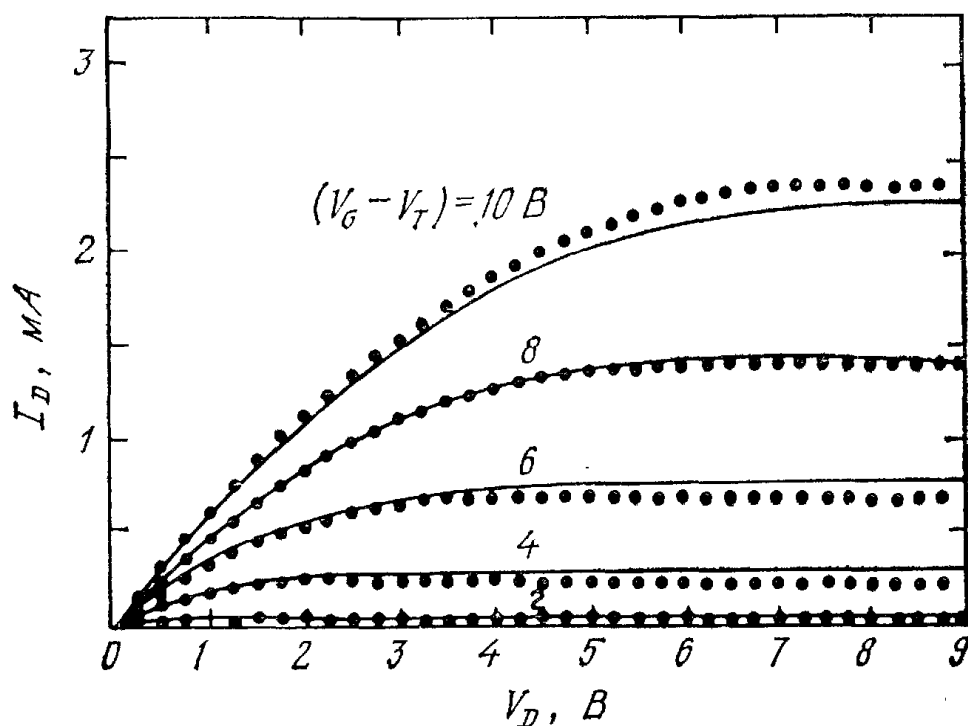


Рис. 10. Теоретические (●) и экспериментальные (—) стоковые характеристики p -канального МОП-транзистора с параметрами $d = 2000 \text{ \AA}$, $N_D = 4,6 \cdot 10^{14} \text{ см}^{-3}$ и $\mu_p = 256 \text{ см}^2 \cdot \text{В}^{-1} \cdot \text{с}^{-1}$ [16].

8.2.3. Подпороговая область

Если напряжение на затворе МОП-транзистора ниже порогового и граница с окислом находится в условиях слабой инверсии, соответствующий ток стока называют подпороговым током [20, 21]. Подпороговая область характеристик особенно важна для МОП-транзисторов, предназначенных для работы в низковольтных полупроводниковых устройствах с малым энергопотреблением (цифровые логические схемы, устройства памяти), поскольку именно этот режим соответствует закрытому состоянию МОП-транзистора, а также описывает процесс переключения из закрытого состояния в открытое.

В условиях слабой инверсии основной компонентой тока канала является диффузионная составляющая. Формула для тока стока выводится так же, как и выражение для коллекторного тока в биполярном транзисторе с однородно легированной базой. Рассматривая МОП-транзистор как $n-p-n$ - (исток—подложка—сток)-биполярный транзистор, можно написать

$$I_D = -qAD_n \frac{dn}{dy} = qAD_n \frac{n(0) - n(L)}{L}, \quad (37)$$

где A — эффективная поперечная площадь канала, а $n(0)$ и $n(L)$ — объемные концентрации электронов у истока и стока соответственно (рис. 6, а):

$$n(0) = n_{p0} e^{\beta \psi_s}, \quad (38a)$$

$$n(L) = n_{p0} e^{\beta \psi_s - \beta V_D}. \quad (38b)$$

Здесь ψ_s — поверхностный потенциал у истока. Эффективная поперечная площадь канала представляет собой произведение ширины канала Z на эффективную толщину инверсионного слоя в направлении, перпендикулярном поверхности. Поскольку концентрация электронов экспоненциально зависит от потенциала ψ , эффективная толщина канала равна расстоянию от поверхности, на котором потенциал уменьшается на величину порядка kT/q . Следовательно, эффективная толщина канала равна $kT/q\mathcal{E}_s$, где \mathcal{E}_s — поверхностное электрическое поле, которое в режиме слабой инверсии определяется выражением

$$\mathcal{E}_s = -Q_B/\epsilon_s = \sqrt{2qN_A\psi_s/\epsilon_s}. \quad (39)$$

Подставив выражения (38) и (39) в выражение (37), получим [18, 22]

$$I_D = \mu_n \left(\frac{Z}{L}\right) \frac{aC_i}{2\beta^2} \left(\frac{n_i}{N_A}\right)^2 \left(1 - e^{-\beta V_D}\right) e^{\beta\psi_s} (\beta\psi_s)^{-1/2}. \quad (40)$$

В последнем выражении были использованы соотношение Эйнштейна $D_n = \mu_n kT/q$ и безразмерный параметр a , характеризующий МОП-структуру. Поверхностный потенциал ψ_s у истока в этих обозначениях записывается через напряжение на затворе V_G следующим образом [18, 19]:

$$\psi_s = (V_G - V_{FB}) - \frac{a^2}{2\beta} \left\{ \left[1 + \frac{4}{a^2} (\beta V_G - \beta V_{FB} - 1) \right]^{1/2} - 1 \right\}. \quad (41)$$

Из выражения (40) следует, что в подпороговой области ток стока экспоненциально увеличивается с ростом напряжения на затворе V_G и перестает зависеть от напряжения стока при $V_D > 3kT/q$. С помощью формул (40) и (41) можно определить характерное изменение напряжения на затворе S , требуемое для изменения тока стока на порядок. По определению

$$\begin{aligned} S &\equiv \ln 10 dV_G/d(\ln I_D) = \\ &= (kT/q) \ln 10 d(\beta V_G)/d(\ln I_D) = \\ &= (kT/q) \ln 10 \left[1 + C_D(\psi_s)/C_i \right] \left\{ 1 - \left(\frac{2}{a^2}\right) [C_D(\psi_s)/C_i]^2 \right\}. \end{aligned} \quad (42)$$

В обычных ситуациях, когда $a \gg C_D/C_i$,

$$S \simeq \frac{kT}{q} \ln 10 (1 + C_D/C_i). \quad (43)$$

Слагаемое в скобках здесь представляет собой отношение емкостного делителя $(C_i + C_D)/C_i$. В структурах с высокой плотностью поверхностных состояний D_{it} эффективную емкость поверхностных состояний $C_{it} = qD_{it}$ нужно «включить» параллельно емкости обед-

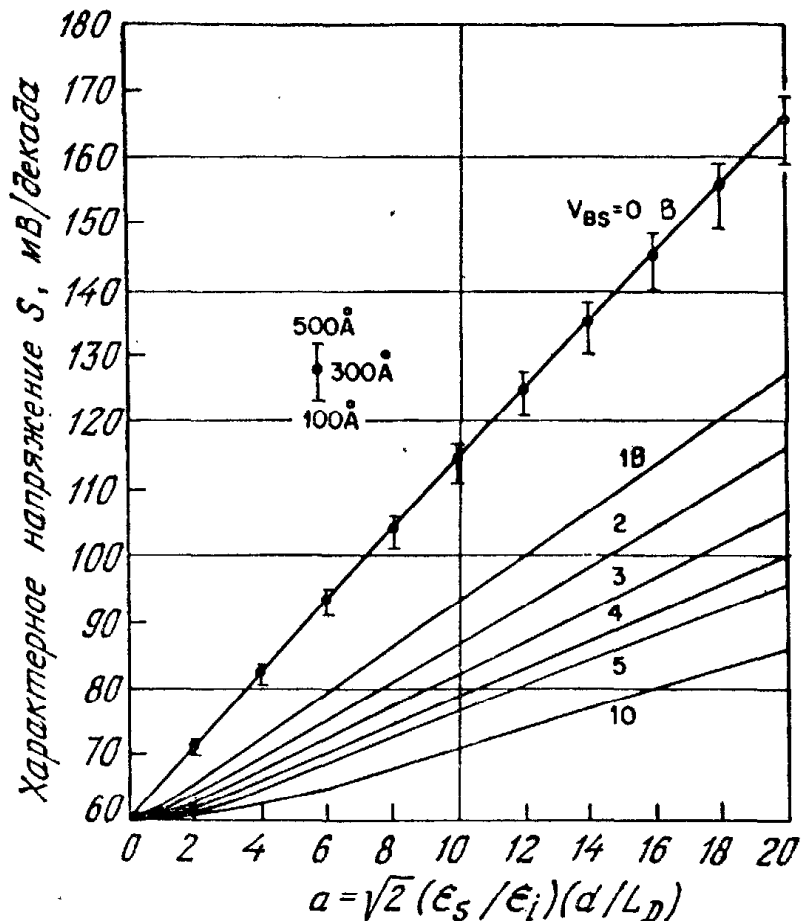


Рис. 11. Характерный сдвиг затворного напряжения S , обеспечивающий в подпороговой области изменение тока на порядок, при различных обратных смещениях подложки [19].

ненного слоя C_D , т. е. заменить в выражении (43) C_D на $C_D + C_{it}$. При этом

$$S_{Dit} \approx \frac{kT}{q} \ln 10 \left(1 + \frac{C_D + C_{it}}{C_i} \right). \quad (44)$$

Когда к подложке МОП-транзистора приложено обратное смещение, поверхностный потенциал ψ_s увеличивается и соответственно уменьшается емкость обедненного слоя C_D , а следовательно, и характерное изменение затворного напряжения S . На рис. 11 приведены теоретические зависимости S от величины параметра a при различных обратных смещениях на подложке МОП-транзистора [19]. Следует отметить, что существенное уменьшение S вызывает уже «первый вольт» обратного смещения подложки.

Экспериментальные подпороговые характеристики МОП-транзистора с длинным каналом (15,5 мкм) для трех значений напряжения обратного смещения V_{BS} приведены на рис. 12 [23]. Как и следовало ожидать, ниже порогового напряжения V_T , отмеченного на каждой из кривых стрелочкой, ток стока МОП-транзистора экспоненциально зависит от напряжения на затворе.

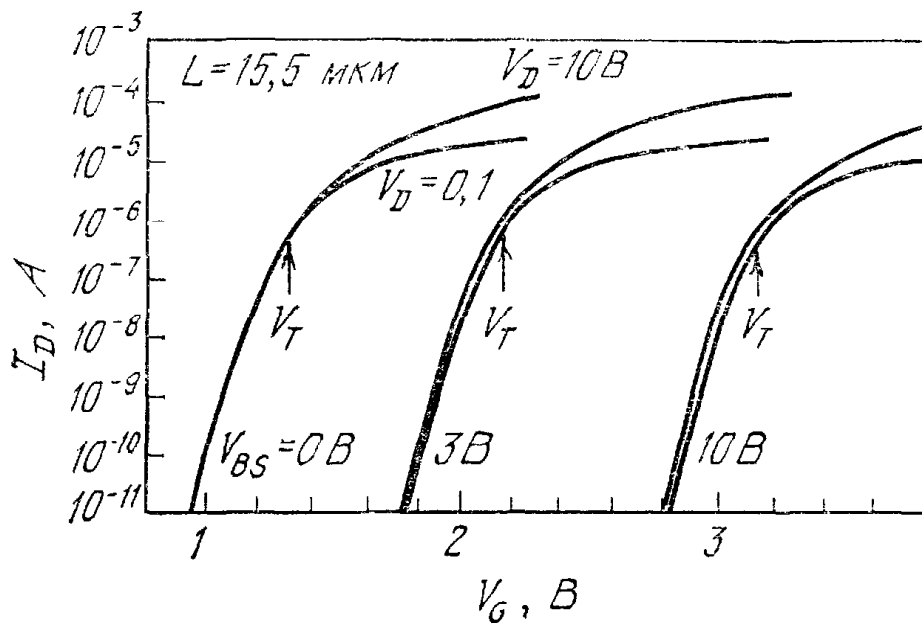


Рис. 12. Экспериментальные подпороговые характеристики МОП-транзистора с длинным каналом ($L = 15,5$ мкм) [23].

Отметим также, что подпороговый участок экспериментальных кривых практически не изменяется при изменении напряжения стока от 0,1 до 10 В. Независимость I_D от V_D на подпороговом участке характерна для длинноканального поведения МОП-транзистора. Толщина слоя окисла и уровень легирования подложки прибора, характеристики которого приведены на рис. 12, составляли 570 \AA и $5,6 \cdot 10^{15} \text{ см}^{-3}$. (Параметр a равен 4.) При этом теоретические значения $S = 83$ мВ на декаду для $V_{BS} = 0$, $S = 67$ мВ на декаду для $V_{BS} = 3$ В и $S = 63$ мВ на декаду при $V_{BS} = 10$ В (рис. 11), а $\Delta V_T = 0,75$ В для $V_{BS} = 3$ В и $\Delta V_T = 1,7$ В для $V_{BS} = 10$ В (рис. 9). Эти результаты близки к экспериментальным значениям соответствующих величин, найденным по характеристикам прибора (рис. 12).

8.2.4. Подвижность носителей

При малых продольных электрических полях \mathcal{E}_y дрейфовая скорость носителей инверсионного слоя пропорциональна приложенному полю $v_d = \mu_n \mathcal{E}_y$ (\mathcal{E}_y — параллельно направлению тока). Из результатов измерений следует, что дрейфовая поверхностная подвижность μ_n зависит от поперечного электрического поля \mathcal{E}_x (\mathcal{E}_x перпендикулярно направлению тока). На рис. 13 приведены зависимости $\mu_n(\mathcal{E}_x)$, измеренные [24] при трех различных температурах в кремниевых МОП-структурах на p -подложке с ориентацией $\langle 100 \rangle$. Отметим, что, согласно результатам работы [24], подвижность μ_n не зависит от обработки поверхности и уровня легирования подложки при $N_A < 10^{17} \text{ см}^{-3}$. При постоянной

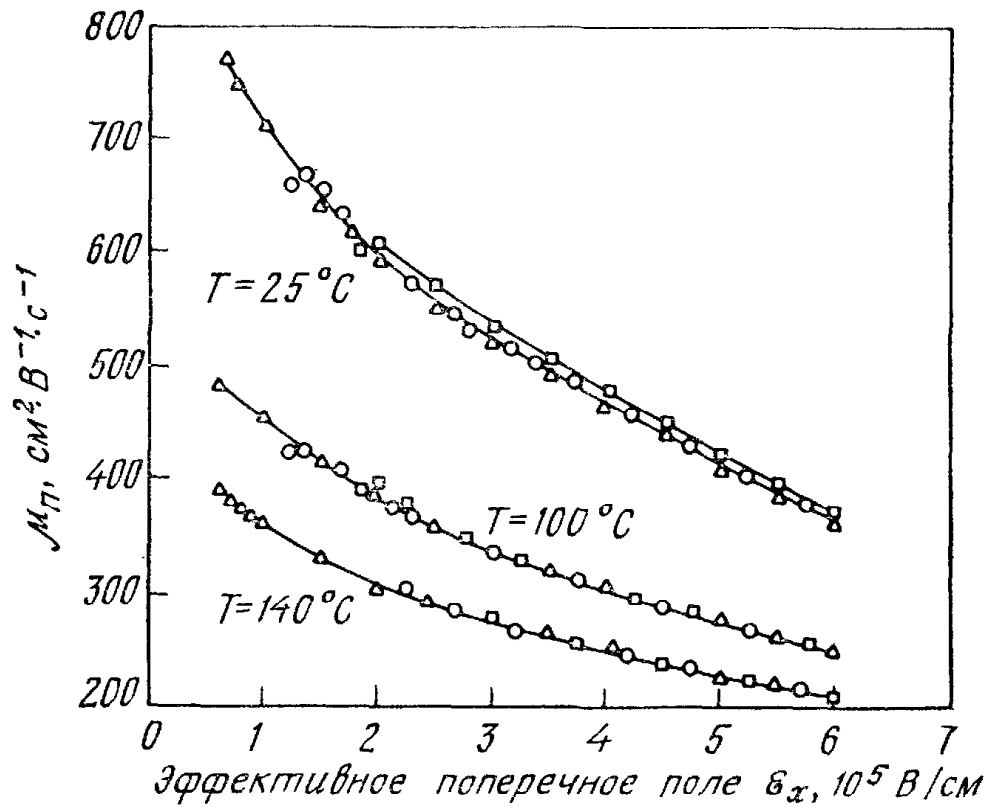


Рис. 13. Зависимость подвижности инверсионного слоя μ_n от эффективного поперечного поля \mathcal{E}_x при трех температурах ($\mathcal{E}_y < 10^3$ В/см) [24].

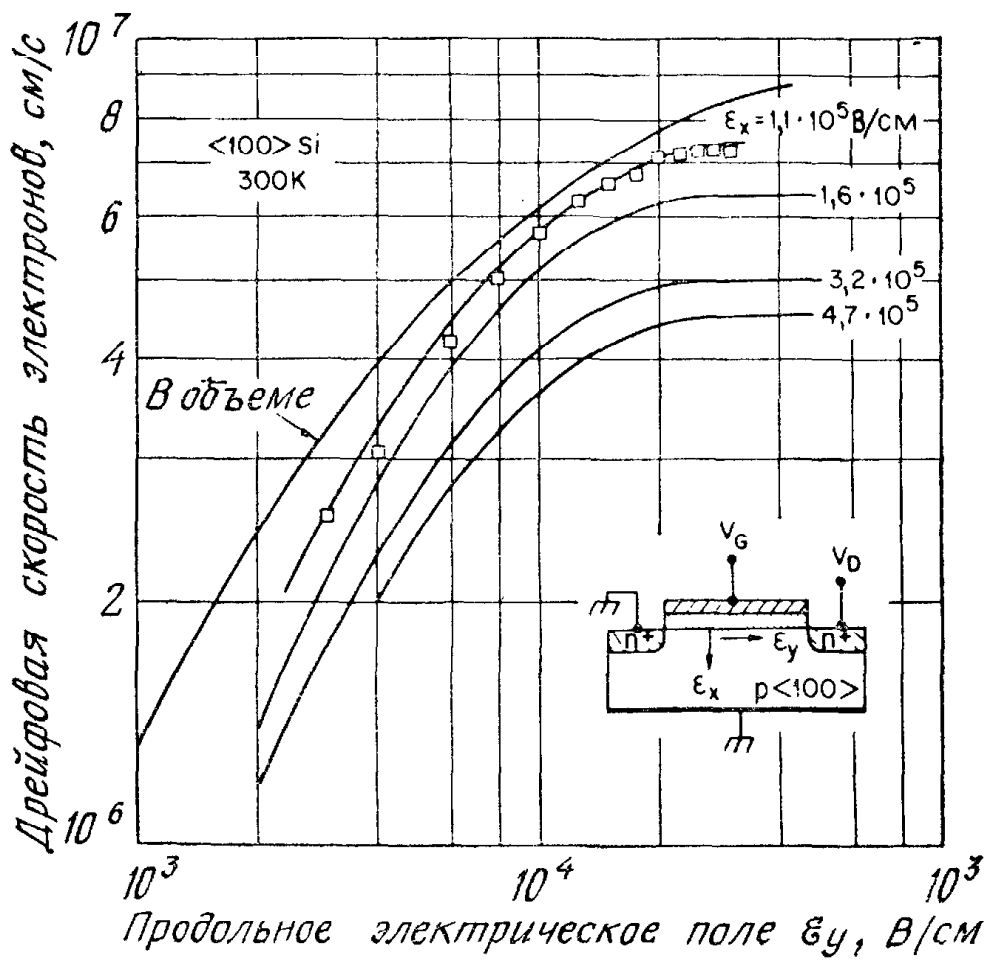


Рис. 14. Зависимость дрейфовой скорости электронов от продольного поля при некоторых значениях поперечного поля [25, 26].

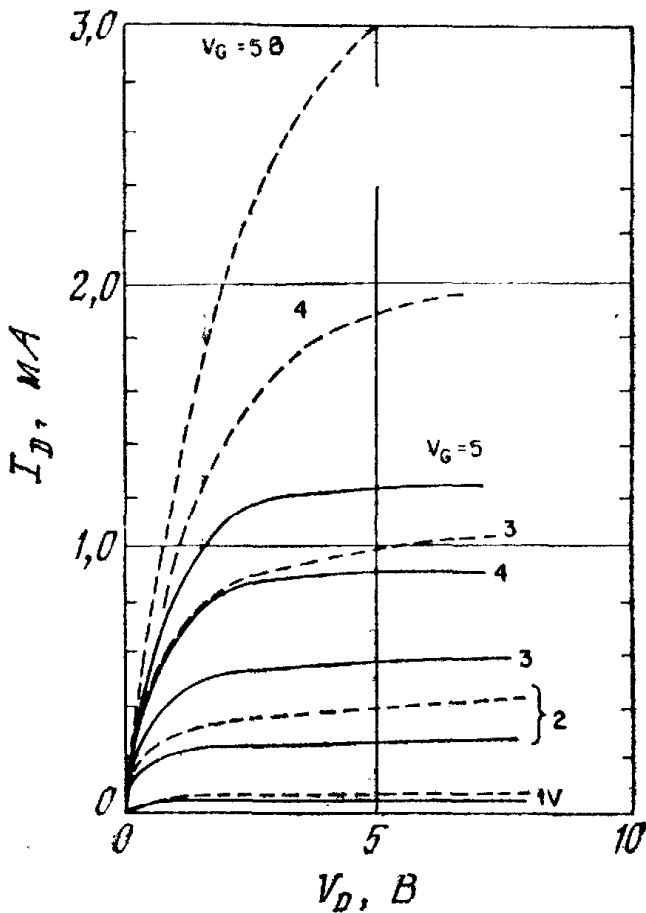


Рис. 15. Стоковые характеристики, рассчитанные в предположении постоянной подвижности (---) и с учетом полевой зависимости подвижности (—) [28].

температуре поверхностная подвижность уменьшается с увеличением эффективного поперечного электрического поля

$$(\mathcal{E}_x)_{\text{eff}} = \frac{1}{\epsilon_s} \left(Q_B + \frac{1}{2} Q_n \right), \quad (45)$$

где $(\mathcal{E}_x)_{\text{eff}}$ — поперечное электрическое поле, усредненное по распределению электронов в инверсионном слое.

При увеличении продольного электрического поля \mathcal{E}_y наступает насыщение дрейфовой скорости (аналогично тому, что наблюдается в объеме кремния). Экспериментальные зависимости дрейфовой скорости инверсионных электронов от величины продольного электрического поля приведены на рис. 14 [25, 26]. При фиксированном значении поперечного электрического поля \mathcal{E}_x дрейфовая скорость при малых значениях \mathcal{E}_y пропорциональна продольному полю \mathcal{E}_y , причем коэффициент пропорциональности равен подвижности μ_n . Зависимость μ_n от \mathcal{E}_x приведена на рис. 13. При увеличении поля \mathcal{E}_y дрейфовая скорость стремится к определенному пределу. Ее зависимость от поля \mathcal{E}_y можно аппроксимировать довольно общим выражением [27]

$$v_d = v_0 \left[1 + \left(\frac{v_0}{v_c} \right)^2 \left(\frac{v_0}{v_c} + G \right)^{-1} + \left(\frac{v_0}{v_s} \right) \right]^{-1/2}, \quad (46)$$

где v_c , v_s и G — подгоночные параметры, а

$$v_0 \equiv \mu_n (\mathcal{E}_x) \mathcal{E}_y. \quad (47)$$

Отметим, что подвижность $\mu_n(\mathcal{E}_x)$ — однозначная функция поперечного поля \mathcal{E}_x . Если $\mathcal{E}_y \rightarrow 0$, зависимость $v_d(\mathcal{E}_y)$ стремится к $v_d = v_0 = \mu_n(\mathcal{E}_x) \mathcal{E}_y$. Если же $\mu_n \mathcal{E}_y > v_s$, дрейфовая скорость примерно равна величине v_s , которая также зависит от поперечного электрического поля \mathcal{E}_x .

В гл. 6 рассматривалось влияние насыщения дрейфовой скорости на характеристики полевых транзисторов с p — n -переходом в качестве затвора. Аналогичные результаты для МОП-транзистора приведены на рис. 15, где сравниваются расчетные характеристики, полученные в предположении постоянной подвижности, и экспериментальные зависимости в том же приборе с учетом насыщения дрейфовой скорости [28]. Эффект насыщения дрейфовой скорости приводит, во-первых, к существенному уменьшению тока насыщения МОП-транзистора и, во-вторых, к тому, что ток насыщения примерно пропорционален напряжению на затворе, а не квадратичен, как предсказывает теория идеального МОП-транзистора [28]. В условиях насыщения дрейфовой скорости ток насыщения МОП-транзистора

$$I_{D \text{ sat}} = ZC_i (V_G - V_T) v_s, \quad (48)$$

и, следовательно, его крутизна в этом режиме постоянна:

$$g_m = (\partial I_{D \text{ sat}}) / \partial V_G = ZC_i v_s. \quad (49)$$

8.2.5. Температурная зависимость

Температура влияет на многие параметры и характеристики МОП-транзисторов, в особенности на подвижность, пороговое напряжение и форму подпорогового участка характеристик. При напряжениях на затворе, обеспечивающих сильную инверсию, подвижность инверсионных носителей при $T > 300$ К уменьшается с температурой по закону T^{-2} [24].

Рассмотрим температурную зависимость порогового напряжения (выражение (30)) в линейной области работы МОП-транзистора

$$V_T = \varphi_{ms} - \frac{Q_f}{C_i} + 2\psi_B + \frac{\sqrt{4\varepsilon_s q N_A \psi_B}}{C_i}. \quad (50)$$

Поскольку разность работ выхода φ_{ms} и фиксированный заряд окисла не зависят от температуры, дифференцируя выражение (50) по температуре, получим [29]

$$\frac{dV_T}{dT} = \frac{d\psi_B}{dT} \left(2 + \frac{1}{C_i} \sqrt{\frac{\varepsilon_s q N_A}{\psi_B}} \right), \quad (51)$$

где

$$\frac{d\psi_B}{dT} \simeq \pm \frac{1}{T} \left[\frac{E_g(T=0)}{2q} - |\psi_B(T)| \right]. \quad (51a)$$

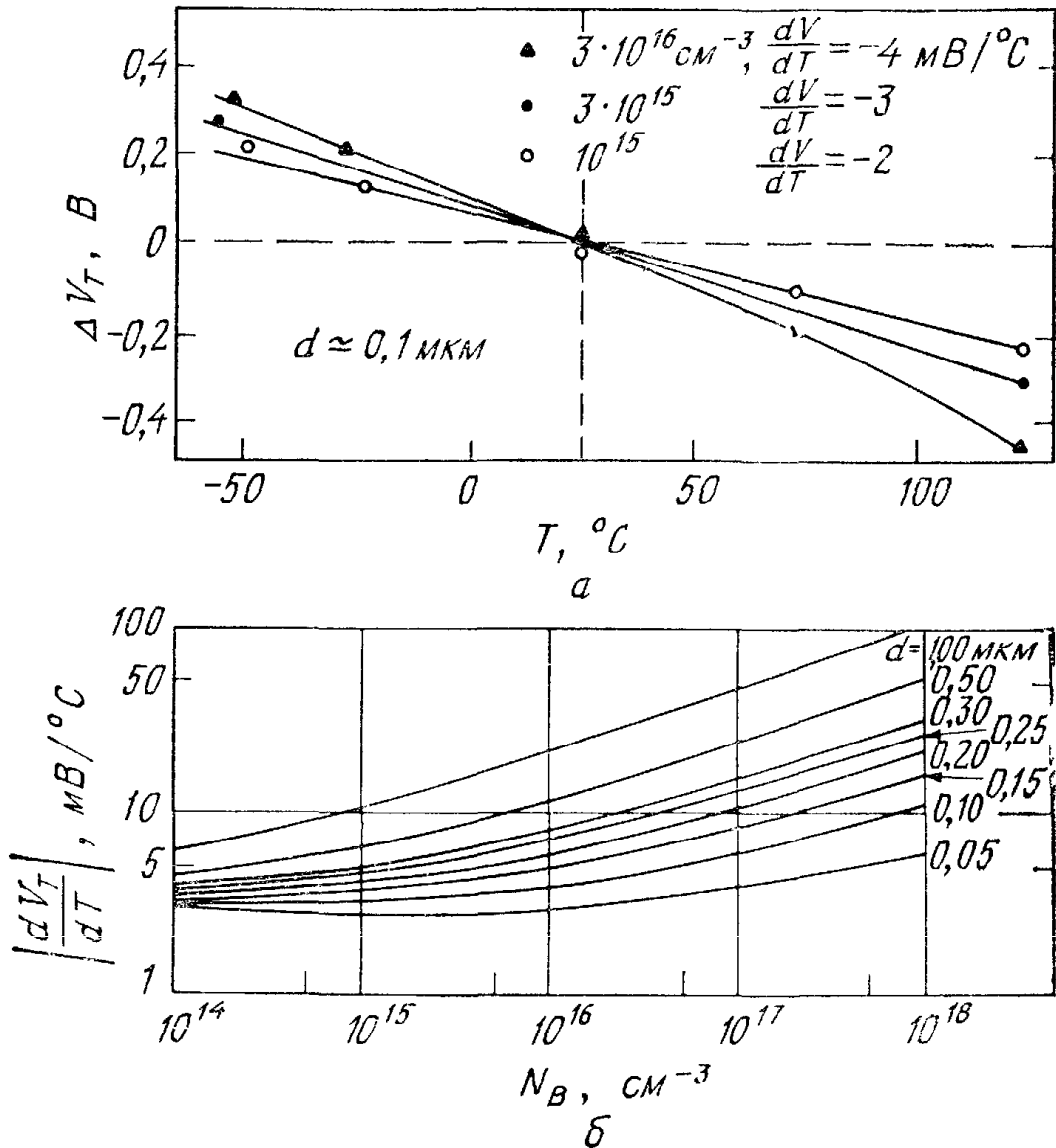


Рис. 16. Зависимость порогового напряжения от температуры (а) и зависимость dV_T/dT от уровня легирования в Si — SiO₂-структурах с различной толщиной окисного слоя (б) [29, 30].

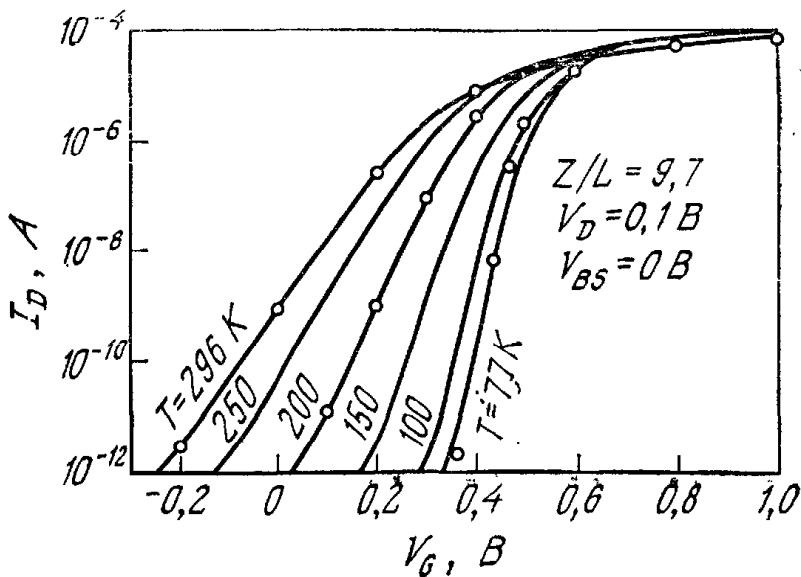


Рис. 17. Передаточные характеристики длинноканального прибора ($L = 9 \text{ мкм}$) при различных температурах [31].
 — теоретические, —○— экспериментальные.

Типичные экспериментальные значения порогового напряжения в Si—SiO₂-системах в окрестности комнатной температуры приведены на рис. 16, а [29, 30]. Видно, что экспериментальные точки в этом температурном диапазоне лежат на почти прямых линиях. Значения соответствующих производных можно рассчитать, воспользовавшись формулой (51) при комнатной температуре. Результаты такого расчета для МОП-структур с различной толщиной слоя окисла в зависимости от уровня легирования подложки приведены на рис. 16, б. Отметим, что при заданной толщине слоя окисла величина производной dV_T/dT в основном увеличивается с ростом уровня легирования.

С понижением температуры характеристики МОП-транзисторов улучшаются (особенно в подпороговой области). На рис. 17 приведены передаточные (сток-затворные) характеристики длинноканального МОП-транзистора ($L = 9$ мкм), параметром которых является температура [31]. Отметим прежде всего, что при уменьшении температуры от 296 до 77 К пороговое напряжение возрастает от $V_T = 0,25$ до 0,5 В. Такое увеличение V_T подобно показанному на рис. 16, а. Однако наиболее важным улучшением является значительное уменьшение характерного перепада затворного напряжения S , требуемого для изменения подпорогового тока на порядок (от 80 мВ на декаду при 296 К до 22 мВ на декаду при 77 К). Такое (примерно четырехкратное) уменьшение S в основном обусловлено температурным множителем kT/q в формуле (42). Другими преимуществами работы МОП-транзисторов при 77 К являются повышенная подвижность, более высокие крутизна и проводимость в окрестности порога, пониженное рассеяние мощности, меньшие токи утечки p — n -переходов и меньшие сопротивления металлических шин. Основным недостатком при этом является необходимость погружения приборов в специальную инертную охлаждающую среду (например, в жидкий азот) и связанные с этим дополнительные, чисто технические трудности.

8.2.6. Типы МОП-транзисторов

Идеальный МОП-транзистор представляет собой усилитель с бесконечным входным сопротивлением и генератором тока на выходе. Электрические свойства реальных приборов представляются более сложной эквивалентной схемой. На рис. 18 приведена одна из таких схем (для схемы включения с общим истоком) [32]. О дифференциальной крутизне g_m мы уже говорили выше. Входная проводимость G_{in} определяется током утечки через тонкий слой подзатворного окисла. Для термически выращенных слоев двуокиси кремния ток утечки между затвором и каналом очень мал (его плотность составляет 10^{-10} А·см⁻²), и

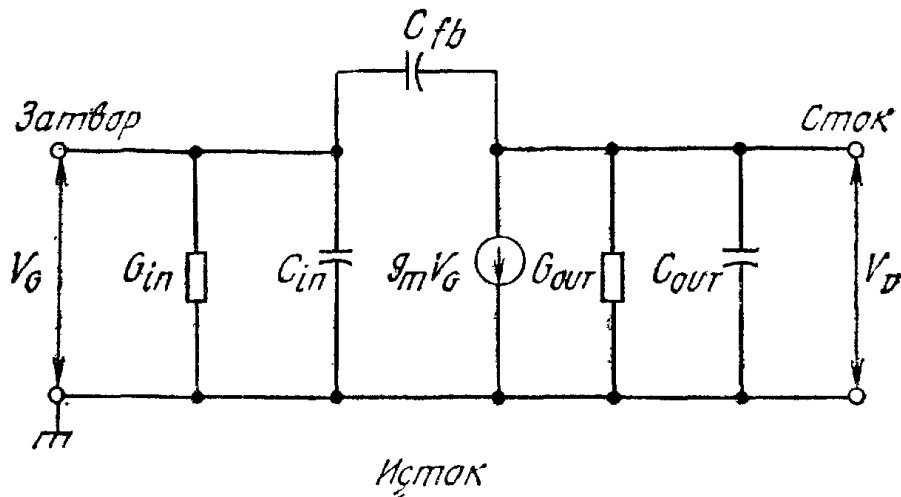


Рис. 18. Эквивалентная электрическая схема МОП-транзистора для включения с общим истоком [6].

обычно входной проводимостью пренебрегают. Входная емкость C_{in} равна $\partial Q_M / \partial V_G$, где Q_M — полный заряд на затворе [16]. В реальных приборах тонкий слой окисла и затвор частично перекрывают области стока и истока. Этот краевой эффект дает основной вклад в проходную емкость C_{fb} . Выходная проводимость G_{out} равна проводимости стокового $p-n$ -перехода. Выходная емкость C_{out} представляет главным образом емкости $p-n$ -переходов стока и истока, последовательно соединенные через объем полупроводниковой подложки. В линейном режиме МОП-транзистора максимальная рабочая частота прибора

$$f_m = \frac{\omega_m}{2\pi} = \frac{g_m}{2\pi C_{in}} \approx \frac{\mu_n V_D}{2\pi L^2}. \quad (52)$$

Этот результат следует из выражения (26) и того факта, что $C_{in} \approx \approx ZLC_i$. В области насыщения (см. выражение (49))

$$f_m \approx \frac{v_s}{2\pi L}, \quad (53)$$

что соответствует времени пролета канала со скоростью насыщения

$$\tau = \frac{L}{v_s}. \quad (54)$$

Для $L = 1$ мкм и $v_s = 10^7$ см·с⁻¹ время пролета τ составляет всего лишь 10 пс. Однако в типичных кольцевых осцилляторах на МОП-транзисторах с длиной канала 1 мкм измеряемые времена задержки обычно на порядок больше. Следовательно, такие задержки обусловлены главным образом различными паразитными емкостями и сопротивлениями в окружении прибора.

В зависимости от типа инверсионного канала различают четыре основных типа МОП-транзисторов. Если при нулевом

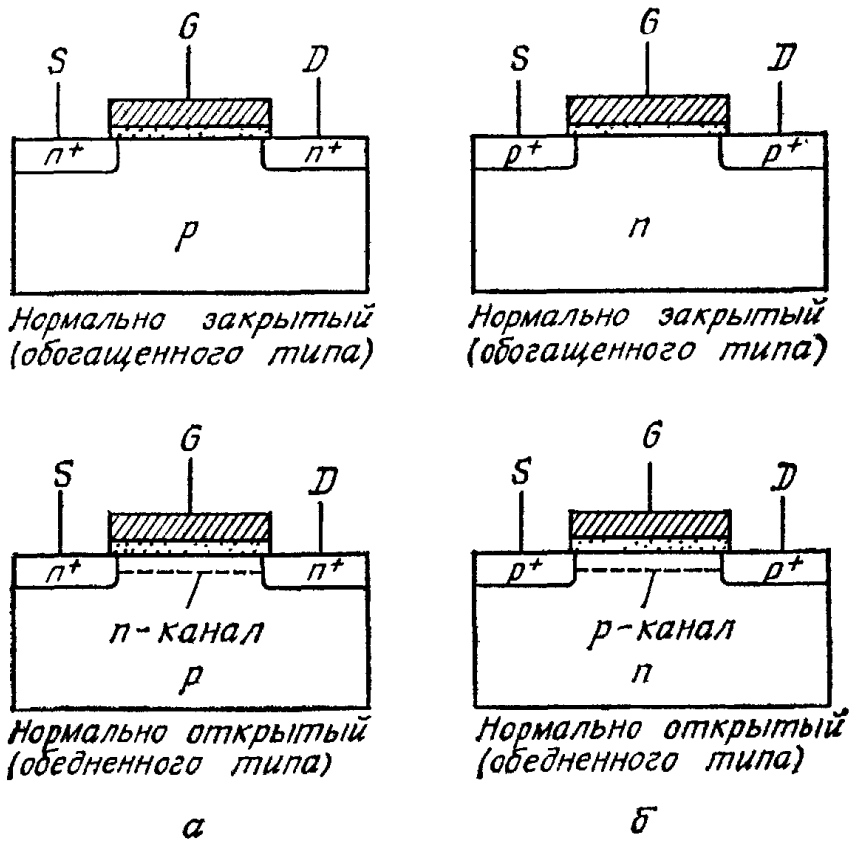


Рис. 19. Основные типы МОП-транзисторов.
 а — n-канальные; б — p-канальные.

Тип	Электрический символ	Выходные характеристики	Передаточные характеристики
n-канал Нормально закрытый (обогащенного типа)			
n-канал Нормально открытый (обедненного типа)			
p-канал Нормально закрытый (обогащенного типа)			
p-канал Нормально открытый (обедненного типа)			

Рис. 20. Электрические символы, передаточные и выходные характеристики четырех типов МОП-транзисторов [33].

напряжении на затворе проводимость канала очень мала и для образования инверсионного n -канала к затвору нужно приложить положительное напряжение, такой прибор называется нормально закрытым (обогащенного типа) n -канальным МОП-транзистором. Если же n -канал существует уже при нулевом смещении на затворе и, чтобы уменьшить его проводимость, к затвору следует приложить отрицательное обедняющее напряжение, такой прибор называют нормально открытым (обедненного типа) n -канальным МОП-транзистором. Оба этих типа n -канальных приборов показаны на рис. 19, а. Аналогичным образом классифицируются p -канальные МОП-транзисторы: нормально закрытые (обогащенного типа) и нормально открытые (обедненного типа) (рис. 19, б).

Электрические символы, передаточные и выходные характеристики этих четырех типов МОП-транзисторов приведены на рис. 20 [33]. Отметим, что для нормально закрытого n -канального прибора, чтобы получить заметный ток стока, нужно приложить к затвору достаточно большое положительное смещение, превышающее пороговое напряжение V_T . В нормально открытых приборах значительные токи могут протекать уже при $V_G = 0$, а смещение затвора увеличивает или уменьшает их величину. Сказанное выше можно легко распространить и на p -канальные приборы, соответствующим образом изменив полярность напряжений

8.3. НЕОДНОРОДНОЕ ЛЕГИРОВАНИЕ И ПРИБОРЫ СО СКРЫТЫМ КАНАЛОМ

В разд. 8.2 концентрация легирующей примеси в подложке предполагалась постоянной. В реальных же приборах распределение примеси практически всегда неоднородно. Определенная неоднородность возникает даже в приборах, изготавливаемых на начально однородно легированных подложках, за счет перераспределения примесей в процессе термического окисления. Более того, в современной МОП-технологии специальное неоднородное легирование подложки, главным образом с помощью ионной имплантации, широко используется для улучшения рабочих параметров приборов. Так, например, в настоящее время ионная имплантация используется: во-первых, для самосовмещения стока и истока с затвором, чтобы уменьшить емкости перекрытия; во-вторых, для подлегирования границы раздела $\text{Si}-\text{SiO}_2$ для корректировки пороговых напряжений; в-третьих, для создания имплантированных каналов на слаболегированных подложках, что предотвращает прокол (смыкание) стока с истоком; в четвертых, для создания приборов со скрытым каналом путем введения в приповерхностную область примесей противоположного типа.

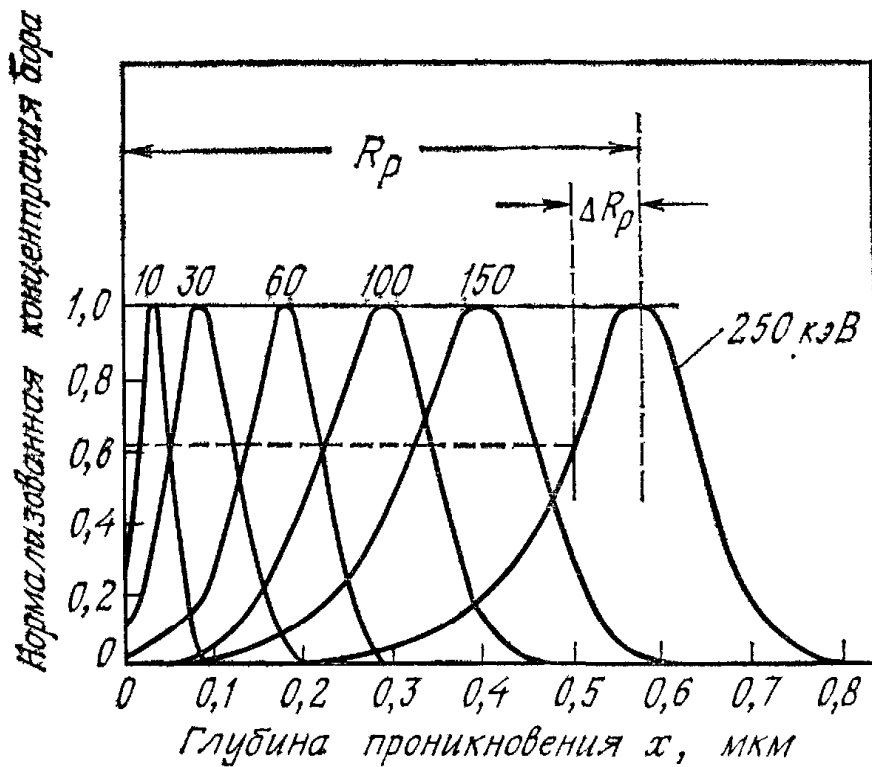


Рис. 21. Нормализованные распределения концентрации имплантированного бора по глубине кремниевой подложки при различных энергиях имплантации [34].

Профили концентрации примеси при ионной имплантации близки к форме гауссова распределения с максимумом на средней длине пробега R_p и стандартным отклонением ΔR_p :

$$N(x) = \frac{D}{\sqrt{2\pi}\Delta R_p} \exp \left[-\frac{(x - R_p)^2}{2(\Delta R_p)^2} \right]. \quad (55)$$

Здесь D — доза ионов на единицу площади (рис. 21) [34]. Средняя длина пробега и стандартное отклонение увеличиваются с ростом энергии имплантируемых ионов [34]. Значения R_p и ΔR_p для бора, мышьяка и фосфора в кремнии приведены в гл. 2.

Ниже мы рассмотрим влияние неоднородного легирования канала на характеристики, в особенности на пороговые напряжения и характерный наклон подпорогового участка характеристик. Мы также рассмотрим МОП-транзисторы со скрытым каналом, которые можно использовать в качестве нагрузки (нормально открытый) в схеме инвертора и как нормально закрытые быстродействующие приборы.

8.3.1. Сдвиг порога

Чтобы рассчитать сдвиг порогового напряжения, обусловленный ионной имплантацией, рассмотрим сначала идеализированный ступенчатый профиль концентрации (рис. 22) [35]. Отметим, что первоначальный профиль концентрации имплантированных ионов заметно изменяется в процессе термического отжига (рис. 22).

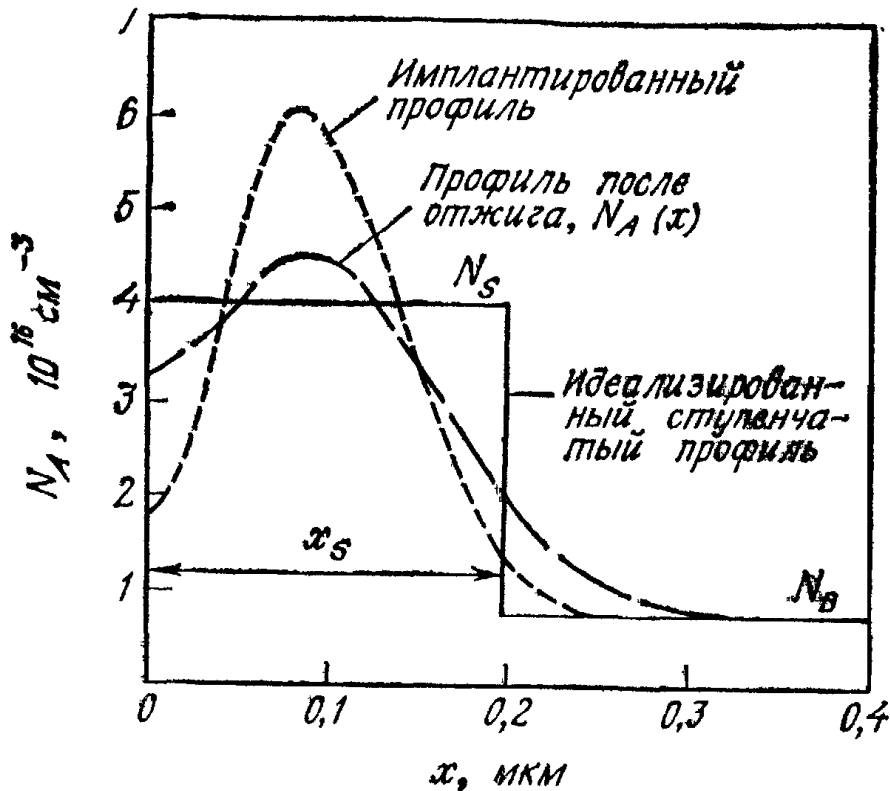


Рис. 22. Профиль легирования в области имплантации под затворным окислом [35].

Профиль после отжига можно аппроксимировать ступенчатой функцией с глубиной x_s , равной сумме средней длины пробега и стандартного отклонения первоначального профиля. Высота ступеньки $N_S - N_B$ определяется соотношением

$$(N_S - N_B) x_s = \int_0^{\infty} [N_A(x) - N_B] dx = D_I. \quad (56)$$

При $x_s \rightarrow 0$ (распределение в виде δ -функции на границе раздела) влияние имплантированных акцепторов (например, отрицательно заряженных ионизированных атомов бора) сводится к уменьшению положительного фиксированного заряда окисла на величину qD_I . Следовательно, в этом случае

$$V_T = V_{FB} + 2\psi_B + \frac{\sqrt{2\varepsilon_s q N_B (2\psi_B + V_{BS})}}{C_i} + \frac{qD_I}{C_i}, \quad (57)$$

а соответствующий сдвиг порога

$$\Delta V_T = V_T(D_I) - V_T(D_I = 0) = \frac{qD_I}{C_i}. \quad (58)$$

Для широких распределений, таких, что $x_s > W_m$ (W_m — максимальная толщина обедненного слоя при сильной инверсии), приповерхностный слой полупроводника можно рассматривать как однородно легированную область с концентрацией N_S . В этом случае пороговое напряжение определяется выражением (31),

в котором N_A следует заменить на N_S и соответствующим образом скорректировать величину ψ_B .

В случае $x_s < W_m$ для определения толщины обедненного слоя и падения напряжения в окисле необходимо решать уравнение Пуассона для ступенчатого профиля концентрации примеси:

$$W_m = \sqrt{\frac{2\epsilon_s}{qN_B} \left[\psi_s + V_{BS} - \frac{qx_s^2}{2\epsilon_s} (N_S - N_B) \right]^{1/2}} \quad (59)$$

$$\begin{aligned} V_i &= \frac{q}{C_i} \int_0^{W_m} N_A(x) dx = \\ &= \frac{q}{C_i} [N_B W_m + (N_S - N_B) x_s] = \frac{q}{C_i} (N_B W_m + D_I), \end{aligned} \quad (60)$$

где ψ_s — поверхностный изгиб зон при сильной инверсии, а V_i — падение напряжения в слое окисла. Отсюда пороговое напряжение

$$V_T = V_{FB} + \psi_s + \frac{\sqrt{2q\epsilon_s N_B}}{C_i} \left(\psi_s + V_{BS} - \frac{qx_s}{2\epsilon_s} D_I \right)^{1/2} + \frac{qD_I}{C_i}. \quad (61)$$

Легко видеть, что, когда $V_{BS} + \psi_s$ становится много больше $qx_s D_I / 2\epsilon_s$, выражение (61) стремится к приближенной формуле (57), соответствующей импульсному (типа δ -функции) профилю примеси на границе с окислом.

Вычисленные с помощью формул (57)—(61) зависимости порогового напряжения от смещения подложки V_{BS} приведены на рис. 23. Кривые *a* и *б* соответствуют случаю однородного легирования с $N_B = 7,5 \cdot 10^{15}$ и $4 \cdot 10^{16}$ см⁻³ соответственно. Эти зависимости аналогичны зависимостям, приведенным на рис. 9. Кривая *в* на рис. 23 получена в приближении импульсного профиля имплантированной примеси на границе раздела. В соответствии с формулой (57) она просто сдвинута относительно кривой *a* на величину $qD_I/C_i = 1,1$ В. Кривая *г* представляет точное решение уравнения Пуассона для ступенчатого профиля. Легко видеть, что при малых V_{BS} , когда $W_m < x_s = 0,2$ мкм, она близка к кривой *б*, соответствующей однородному распределению с высокой концентрацией примеси ($4 \cdot 10^{16}$ см⁻³). При больших V_{BS} , когда $W > x_s$, кривая *a* приближается к кривой *в*, соответствующей импульсной аппроксимации профиля концентрации имплантированных ионов.

Со схемотехнической точки зрения желательно иметь как можно меньшую чувствительность порогового напряжения МОП-транзистора к смещению подложки. Из рис. 23 видно, что для корректировки V_T следует использовать мелкую (с малыми x_s) имплантацию. Это позволяет при изменениях пороговых на-

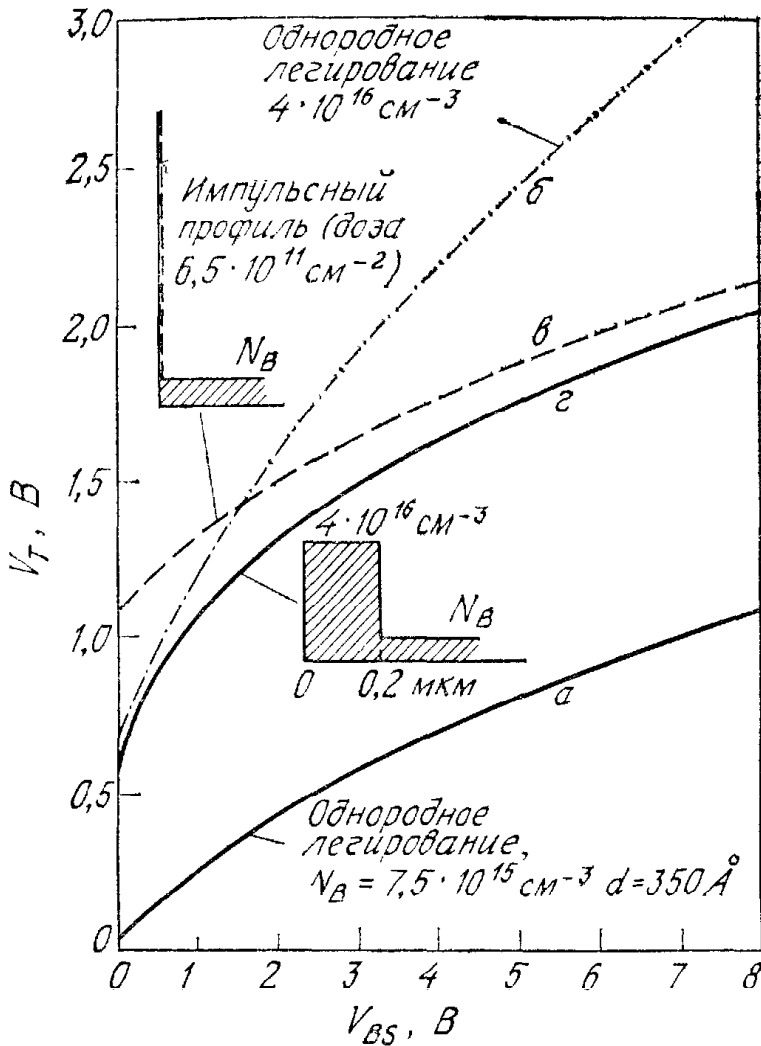


Рис. 23. Теоретическая зависимость чувствительности порога к смещению на подложке для различных профилей легирования [35].

пряжений сохранить исходную малую чувствительность V_T к смещению подложки.

Описанное выше приближение ступенчатого профиля представляет собой оценку первого порядка для сдвига порогового напряжения при ионной имплантации. Чтобы получить точный результат, необходимо решать уравнение Пуассона с точным профилем концентрации имплантированной примеси. Оказывается, что для сравнительно глубоких и узких распределений концентрации имплантированных ионов (рис. 24, а) лучшей, чем ступенчатая, является аппроксимация реального распределения импульсным профилем (типа δ -функции), центрированным на

$$x_c = \int_0^{x_1} [N_A(x) - N_B] x dx / D_I \quad (62)$$

и с дозой

$$D_I = \int_0^{x_1} [N_A(x) - N_B] dx. \quad (63)$$

Соответствующее этому импульсному профилю распределение электрического поля показано на рис. 24, б [36]. Площадь под

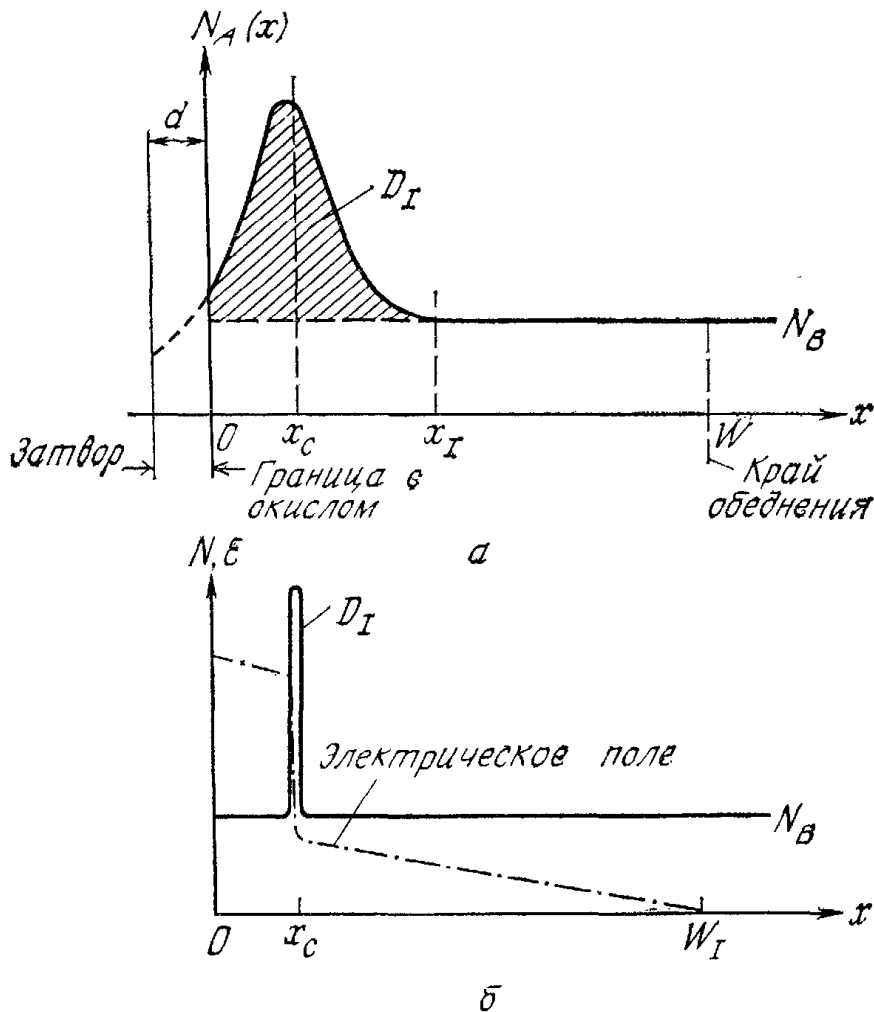


Рис. 24. Реальный профиль концентрации имплантированной примеси (а) и его импульсная аппроксимация (б) [36].

кривой равна поверхностному потенциалу ψ_s . Из уравнения Пуассона можно найти глубину обедненного слоя:

$$W_I = \sqrt{2L_D} [\beta\psi_s - D_I x_c / (N_B L_D^2) - 1]^{1/2}. \quad (64)$$

Напомним, что в однородно легированной подложке глубина обедненного слоя

$$W_U = \sqrt{2L_D} [\beta\psi_s - 1]^{1/2}. \quad (65)$$

Из сравнения выражений (64) и (65) следует, что ионная имплантация уменьшает глубину обедненного слоя.

Теперь с помощью выражения (64) найдем обусловленный имплантацией сдвиг порогового напряжения ΔV_T , определив его [11, 36] как разность между напряжениями на затворе, требуемыми для создания одного и того же инверсионного заряда Q_n в МОП-структуре с имплантированной примесью, и в аналогичной структуре на однородно легированной подложке. В однородно легированной МОП-структуре с напряжением на затворе V_{GU} и поверхностным потенциалом ψ_{SU} плотность инверсионного заряда равна

$$Q_n = Q_s - Q_B = C_i (V_{GU} - \psi_{SU}) - qN_B W_U. \quad (66)$$

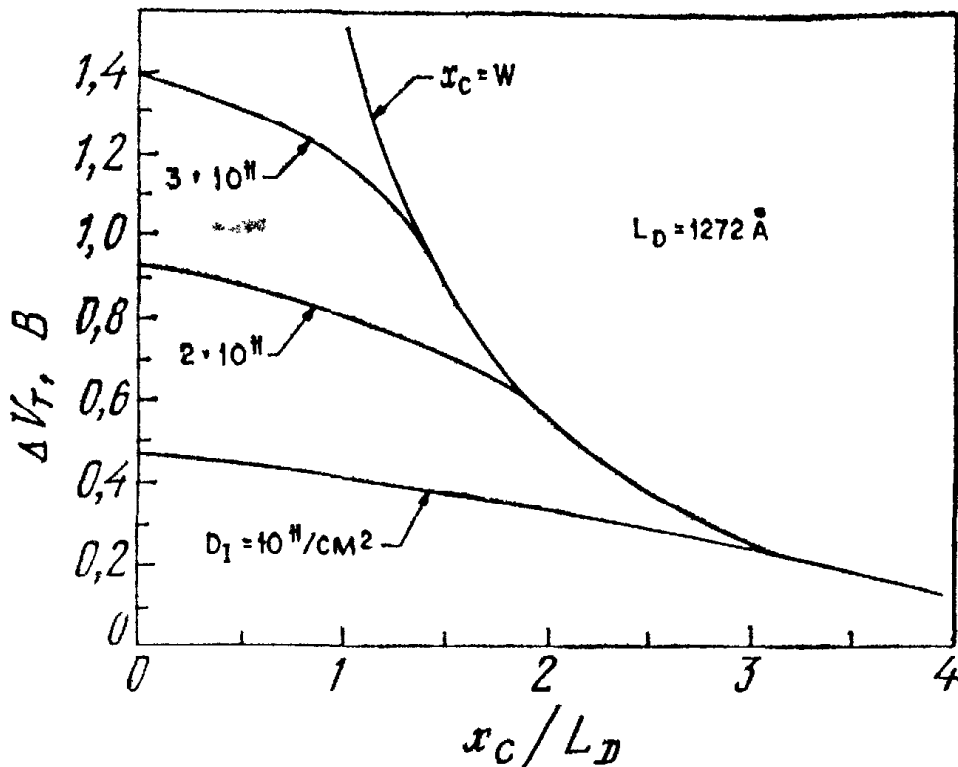


Рис. 25. Зависимость сдвига порогового напряжения от центрады распределения имплантированной примеси при различных дозах ионов [36].

В случае имплантации при напряжении на затворе V_{GI} и изгибе зон ψ_{SI}

$$Q_n = C_i (V_{GI} - \psi_{SI}) - qN_B W_I - qD_I. \quad (67)$$

Приравняв выражения (66) и (67), для сдвига порогового напряжения ΔV_T получим

$$\Delta V_T = V_{GI} - V_{GU} = \psi_{SI} - \psi_{SU} - qN_B (W_U - W_I)/C_i + qD_I/C_i. \quad (68)$$

На практике при достаточно большом Q_n (в режиме сильной инверсии) поверхностные потенциалы ψ_{SI} и ψ_{SU} различаются на величину, которая меньше kT/q , так что в выражении (68) можно пренебречь этой разностью. Подставив в выражение (68) выражения (64) и (65), найдем

$$\Delta V_T = \frac{qD_I}{C_i} - \frac{V\sqrt{2}qL_D N_B}{C_i} \left[(\beta\psi_{SU} - 1)^{1/2} - \left(\beta\psi_{SU} - \frac{D_I x_c}{N_B L_D^2} - 1 \right)^{1/2} \right]. \quad (69)$$

Как уже говорилось выше, изгиб зон ψ_{SU} при сильной инверсии примерно равен $2\psi_B + V_{BS}$, и для расчета зависимости сдвига порога от дозы D_I и центрады x_c имплантированной примеси остается лишь подставить $\psi_{SU} = 2\psi_B + V_{BS}$ в выражение (69).

Полученные таким образом зависимости ΔV_T от центрады x_c при различных дозах [36] приведены на рис. 25. При $x_c = 0$, т. е. когда δ -функция имплантации расположена на границе

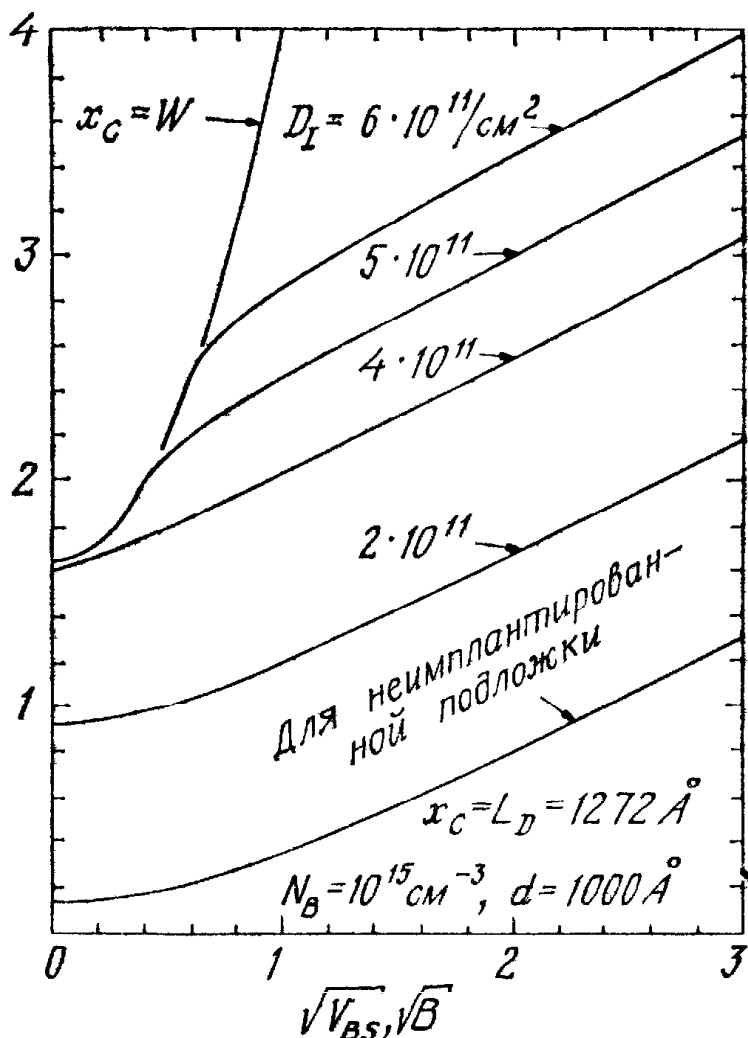


Рис. 26. Зависимость порогового напряжения от смещения на подложке при различных дозах имплантированных ионов [36].

раздела Si—SiO₂, $\Delta V_T = qD_I/C_i$, поскольку слагаемое в квадратных скобках в формуле (69) равно нулю. С ростом x_c величина ΔV_T уменьшается. Наконец, x_c становится равным толщине слоя обеднения W , и край обедненной области совпадает с имплантированным слоем $W = x_c$. Значение дозы, при которой $W = x_c$, можно получить из уравнения (64):

$$D_I(x_c = W) = N_B (W_U^2 - x_c^2)/2x_c. \quad (70)$$

На рис. 26 приведена зависимость порогового напряжения от смещения на подложке для импульсного профиля имплантации с $x_c = L_D = 1272 \text{ \AA}$. При $D_I < 4 \cdot 10^{11} \text{ см}^{-2}$ имплантация приводит к параллельному сдвигу зависимости $V_T(V_{BS})$ относительно исходной кривой, соответствующей однородно легированной подложке (неимплантированной). При таких малых D_I главную роль в уравнении (69) играет первое слагаемое qD_I/C_i . При $D > 4 \cdot 10^{11} \text{ см}^{-2}$ край обедненного слоя фиксируется в плоскости имплантации. В этом случае с ростом смещения V_{BS} изменения порогового напряжения сначала описываются кривой, соответ-

ствующей условию $x_c = W$. Однако при больших смещениях на подложке, таких, что $(2\psi_B + V_{BS}) > D_I x_c / \beta N_B L_D^2$, край обедненного слоя «отрывается» от плоскости имплантации x_c , первое слагаемое в выражении (69) опять становится доминирующим, а зависимость $V_T(V_{BS})$ становится параллельной исходной («неимплантированной») кривой.

8.3.2. Наклон подпороговых характеристик

Как уже говорилось выше, в подпороговой области прибора с однородно легированной подложкой логарифм тока стока $\ln I_D$ увеличивается почти линейно с ростом напряжения на затворе. Обратный наклон этой прямой определяет, на сколько следует уменьшить напряжение на затворе МОП-транзистора, чтобы уменьшить ток прибора до требуемого уровня. Поэтому подпороговую область, соответствующую закрытому состоянию МОП-транзистора, обычно характеризуют напряжением S , необходимым для уменьшения тока стока на порядок величины. В случае однородного легирования [19]

$$S \simeq \frac{kT}{q} \ln 10 (1 + C_D/C_i) \left/ \left[1 - \left(\frac{2}{a^2} \right) \left(\frac{C_D}{C_i} \right)^2 \right] \right., \quad (71)$$

где C_D — удельная дифференциальная емкость обедненного слоя:

$$C_D = \varepsilon_s / W_I. \quad (72)$$

Влияние имплантации на подпороговое напряжение S определяется главным образом изменением (в соответствии с выражением [64]) дифференциальной емкости обедненного слоя. На рис. 27 приведены зависимости S от центриды имплантированного профиля примеси для различных доз D_I . При $x_c = 0$ исходное значение S , как и в однородно легированной подложке, определяется только концентрацией примеси и толщиной слоя окисла. При фиксированной дозе D_I характерное напряжение S сначала увеличивается с ростом x_c , поскольку при этом край обедненного слоя приближается к границе раздела Si—SiO₂ и дифференциальная емкость C_D растет. Затем край слоя обеднения попадает в плоскость имплантации, и теперь при росте x_c он уже удаляется от границы с окислом вместе с плоскостью имплантации. При этом, естественно, C_D и S уменьшаются до исходных значений, соответствующих неимплантированной подложке. Излом кривых на рис. 27 соответствует ситуации, когда $x_c = W$. В реальных условиях острота пика, естественно, сглаживается за счет ненулевой толщины слоя имплантированной примеси. Из данных, приведенных на рис. 27, следует, что с точки зрения поддержания высокой крутизны участка выключения МОП-транзистора (малые значения S) предпочтительна мелкая (с малыми x_c) имплантация.

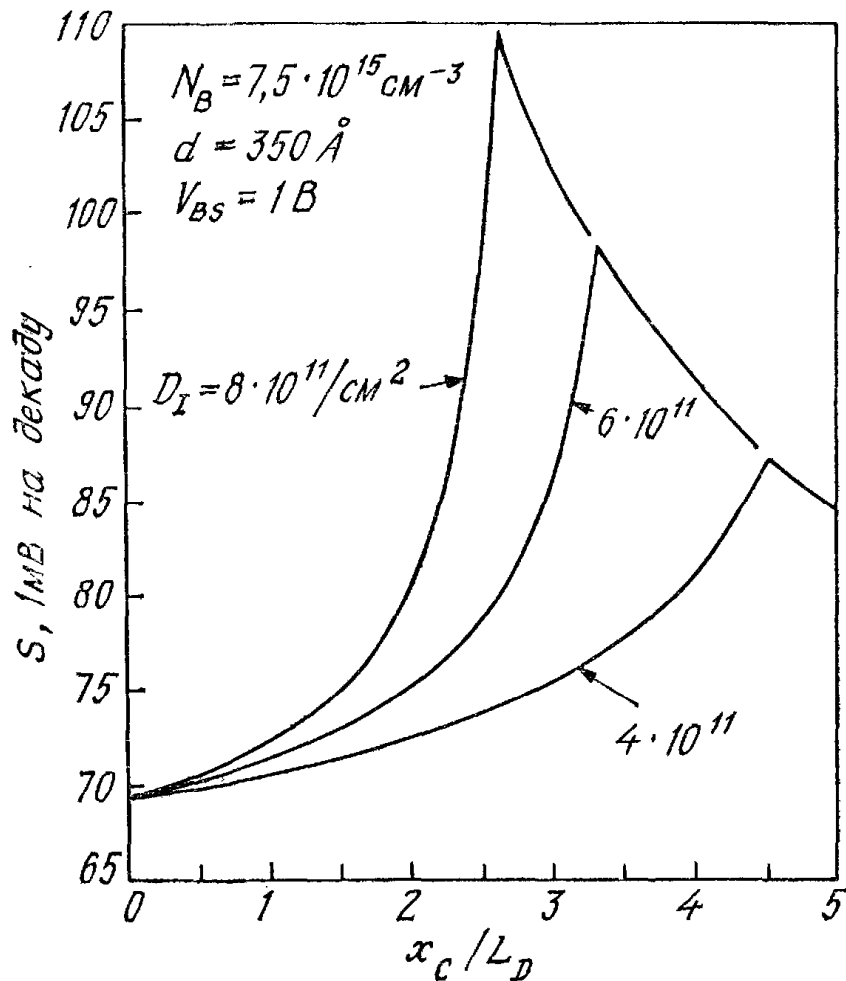


Рис. 27. Зависимость характерного подпорогового напряжения S от центроида распределения имплантированных ионов при различных дозах [19].

8.3.3. Приборы со скрытым каналом

Когда тип имплантированной примеси противоположен типу примесных атомов в исходной подложке (например, при имплантации мышьяка в p -кремний), в приповерхностной части подложки может сформироваться так называемый скрытый канал. На рис. 28, а показан условный поперечный разрез такого n -канального МОП-транзистора со скрытым каналом, где в отличие от обычного МОП-транзистора проводящий канал заглублен в объем полупроводника от границы раздела $\text{Si}-\text{SiO}_2$. Истинный профиль легирования приповерхностной области подложки такого прибора и его ступенчатая аппроксимация приведены на рис. 28, б. Ясно, что в окрестности металлургического $p-n$ -перехода ($x = x_j$) образуется обедненная область, локальная толщина которой определяется величиной напряжений, приложенных к структуре. Напряжение на затворе, кроме того, модулирует локальную толщину приповерхностной подзатворной области обеднения.

МОП-транзисторы со скрытым каналом могут быть нормально открытыми и нормально закрытыми приборами в зависимости от глубины и уровня легирования канала. Рассмотрим сначала

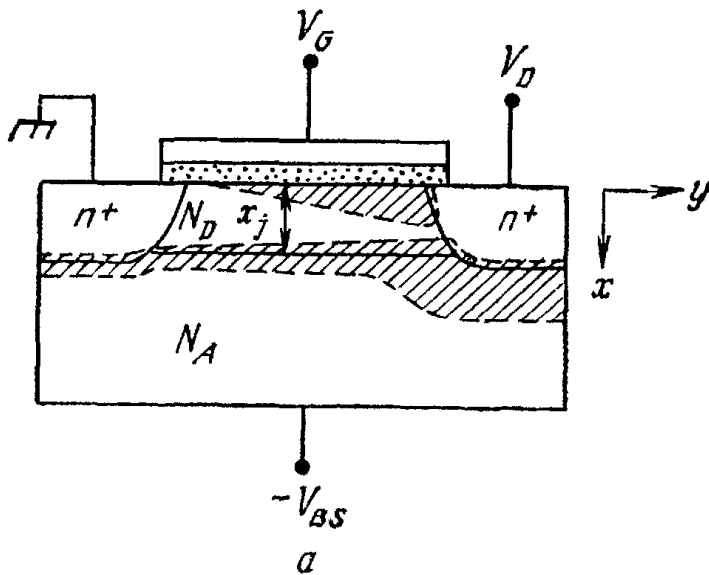
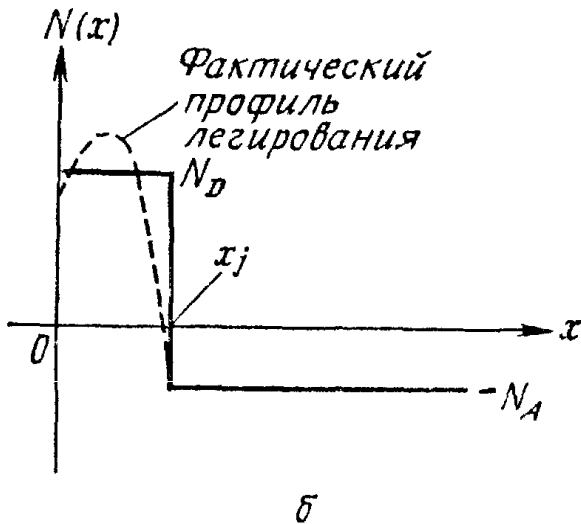


Рис. 28. МОП-транзистор со скрытым каналом (а) и профиль легирования канала (б).



работу нормально открытого прибора. Его условное поперечное сечение в области вблизи истока [37] показано на рис. 29 (вставка). Когда напряжение на затворе увеличивается (в рассматриваемом случае оно отрицательно), ширина скрытого канала уменьшается. Наконец, при $V_G = V_P$ величина x_b становится равной нулю и происходит отсечка канала. Напряжение отсечки канала V_P рассчитывается с помощью уравнения Пуассона и закона сохранения электрической индукции (Гаусса) на границе Si—SiO₂ в качестве граничного условия. Для ступенчатого профиля (рис. 28, б) в предположении $N_D \gg N_A$ получаем

$$V_P - V_{KB} = -\frac{qN_D x_j}{C_i} \left(1 + \frac{C_i}{2C_D}\right) + \frac{\sqrt{2\epsilon_s q N_A}}{C_i} \left(1 + \frac{C_i}{C_D}\right) (V_{bi} + V_{BS})^{1/2}, \quad (73)$$

где $C_D = \epsilon_s/x_j$, а V_{bi} — контактная разность (встроенный потенциал) p — n -перехода. На рис. 29 приведены рассчитанные по этой формуле зависимости напряжения отсечки канала V_P от

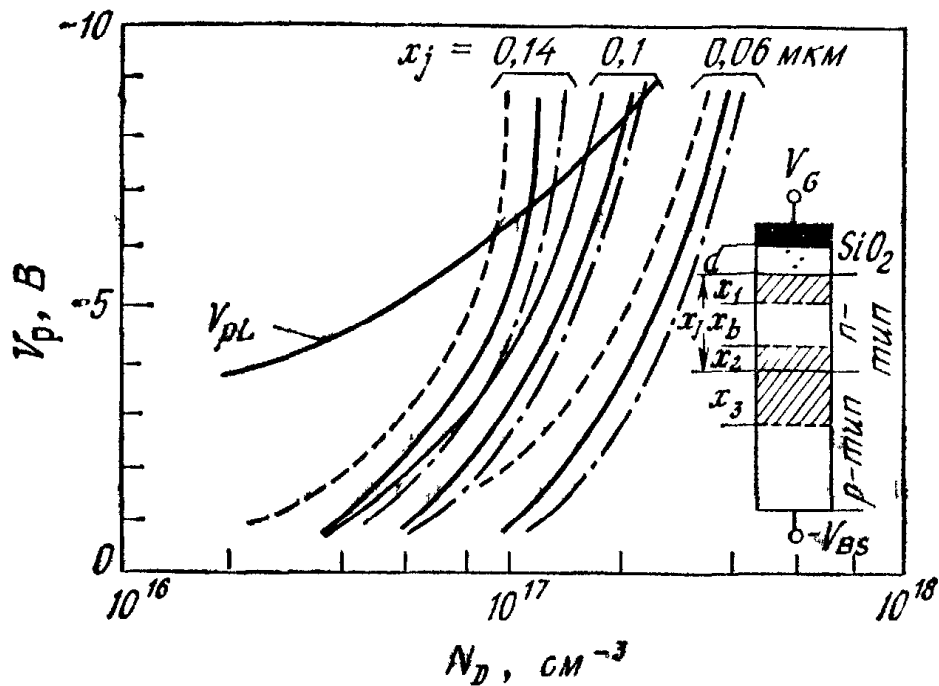


Рис. 29. Зависимость напряжения отсечки от концентрации доноров при различных значениях глубины канала и уровня легирования подложки. На вставке показана толщина обедненных слоев в области у истока [37].

--- $N_A = 10^{16} \text{ см}^{-3}$, — $N_A = 3 \cdot 10^{16} \text{ см}^{-3}$, - · - $N_A = 5 \cdot 10^{16} \text{ см}^{-3}$.

концентрации доноров в приповерхностном слое при нескольких значениях глубины перехода x_1 для случая $V_{BS} = 0$. Видно, что при фиксированных x_j напряжение отсечки канала увеличивается с ростом концентрации. Однако при слишком больших N_D (либо x_1) полностью перекрыть проводящий канал напряжением на затворе невозможно, поскольку в конечном счете увеличение x_1 с ростом V_G будет стабилизировано (на уровне W_m) образованием дырочного инверсионного слоя на границе раздела $\text{Si}-\text{SiO}_2$. Соответствующее предельное значение напряжения отсечки канала равно

$$V_{pL} = V_{FB} + 2\psi_B + \sqrt{2\epsilon_s q N_D (2\psi_B) / C_i}. \quad (74)$$

Предельное значение глубины перехода x_{jL} , при превышении которого канал уже не перекрывается полностью ($x_b \neq 0$), легко вычисляется по формуле

$$x_{jL} = \frac{1}{N_D} \sqrt{\frac{2\epsilon_s N_A}{q}} (\sqrt{2V_{bi} N_D / N_A} + \sqrt{V_{bi} + V_{BS}}). \quad (75)$$

Выходные (стоковые) характеристики МОП-транзистора со скрытым каналом аналогичны характеристикам полевого транзистора с $p-n$ -переходом в качестве затвора (гл. 6). Они также линейны при малых V_D и насыщаются при больших стоковых напряжениях. Согласно работе [38], ток стока транзистора со скрытым каналом

$$I_D = Z (\mu D_i - \mu Q_s - \mu Q_B) \frac{dV}{dy}, \quad (76)$$

где

$$D_f = \int_0^{x_j} [N(x) - N_A] dx \simeq (N_D - N_A) x_j \quad (77)$$

— плотность имплантированных центров на единицу площади, а

$$Q_s = -\bar{C} [V_G - V_{FB} - V(y)] \quad (78)$$

— плотность заряда (приповерхностного слоя обеднения либо аккумуляции) на единицу площади границы раздела, \bar{C} — усредненная емкость затвора,

$$Q_B = \sqrt{2\epsilon_s q N'_A} [V_{bi} + V_{BS} + V(y)]^{1/2} \quad (79)$$

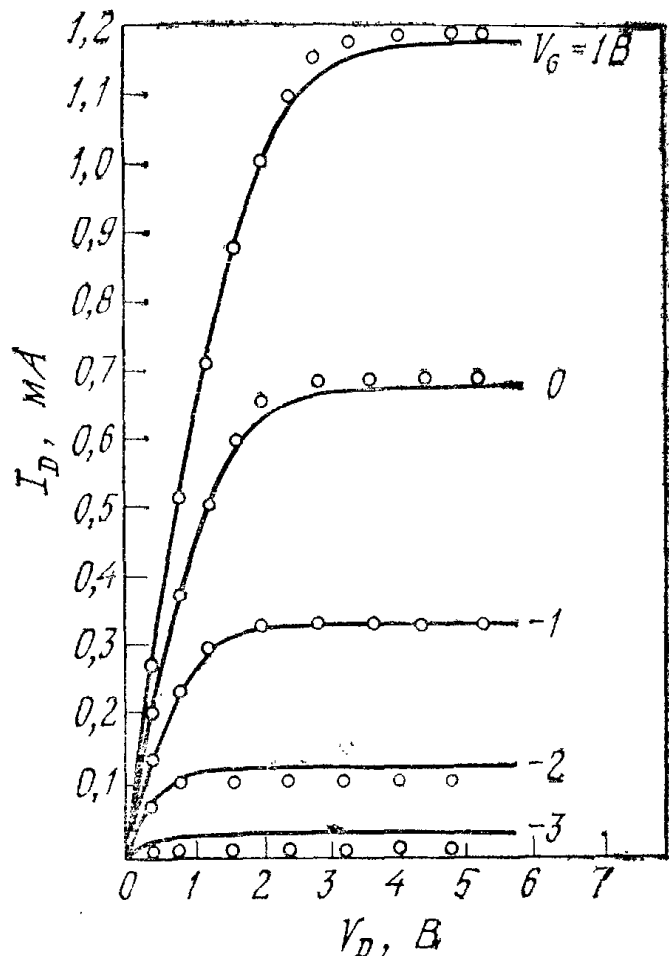
— заряд в обедненном слое p — n -перехода, а $N'_A = N_D N_A / (N_D + N_A)$. В нормально обедненном режиме $[(V_G - V_{FB}) \leq 0]$ ток стока протекает только в скрытом объемном канале, поэтому в выражение (76) следует подставлять значение объемной подвижности μ_B . Подставив выражения (77)—(79) в выражение (76) и проинтегрировав от 0 до L , получим

$$I_D = \frac{\mu_B Z}{L} \left\{ D_f V_D + \bar{C} \left[(V_G - V_{FB}) V_D - \frac{1}{2} V_D^2 \right] - \right. \\ \left. - \frac{2}{3} \sqrt{2\epsilon_s q N'_A} [(V_{bi} + V_{BS} + V_D)^{3/2} - (V_{bi} + V_{BS})^{3/2}] \right\}. \quad (80)$$

Рассчитанные по этой формуле выходные характеристики и экспериментальные зависимости $I_D(V_D)$ хорошо соответствуют друг другу (рис. 30). Отметим, что, поскольку рассмотренный МОП-транзистор является нормально открытым прибором, в нем уже при $V_G = 0$ протекают довольно большие токи. При изменении напряжения на затворе ток стока также изменяется.

На рис. 31, а показаны профиль легирования и толщины обедненных слоев нормально открытого n -канального МОП-транзистора со скрытым каналом в режиме, предшествующем отсечке канала. На рис. 31, б и в приведены распределения зарядов и подвижных электронов в момент отсечки ($x_1 + x_2 = x_j$, $V_G = V_P$). Закрытое состояние прибора при $V_G > V_P$ полностью соответствует подпороговому режиму обычного МОП-транзистора обогащенного типа. Главной компонентой тока при этом является диффузионная компонента, а сам подпороговый ток зависит экспоненциально от напряжения затвора. Характеризующий наклон подпорогового (экспоненциального) участка параметр S

Рис. 30. Экспериментальные (—) и рассчитанные (○) стоковые характеристики нормально открытого МОП-транзистора [38].
 $V_{FB} = -1,25$ В; $V_{BS} = 0$ В.



определяется соотношением емкостей, показанных на рис. 31, а [39]:

$$S = \frac{kT}{q} \ln 10 [(C_1 + C_2)/C_1] = \frac{kT}{q} \ln 10 \left[1 + \frac{\epsilon_i x_1 + \epsilon_s d}{\epsilon_i (x_2 + x_3)} \right]. \quad (81)$$

Очевидно, что рассматриваемый МОП-транзистор со скрытым каналом будет нормально закрытым прибором, когда обедненные слои (приповерхностный и $p-n$ -перехода) будут сомкнуты уже при нулевом напряжении на затворе $V_G = 0$. Для этого необходимо, чтобы напряжение отсечки V_P , которое теперь можно по аналогии назвать пороговым напряжением V_T :

$$V_{FB} - V_T = \frac{qN_D(x_j - x_1)}{C_i} + \frac{qN_D(x_j - x_1)^2}{2\epsilon_s}, \quad (82)$$

было бы больше нуля, что возможно в структурах с большими напряжениями плоских зон V_{FB} . На практике для этого применяются затворы из сильнолегированного бором поликремния или из металлов с большой работой выхода.

По сравнению с поверхностными МОП-транзисторами в приборах со скрытым каналом можно ожидать более высоких значений подвижности, поскольку движущиеся в объеме носители не

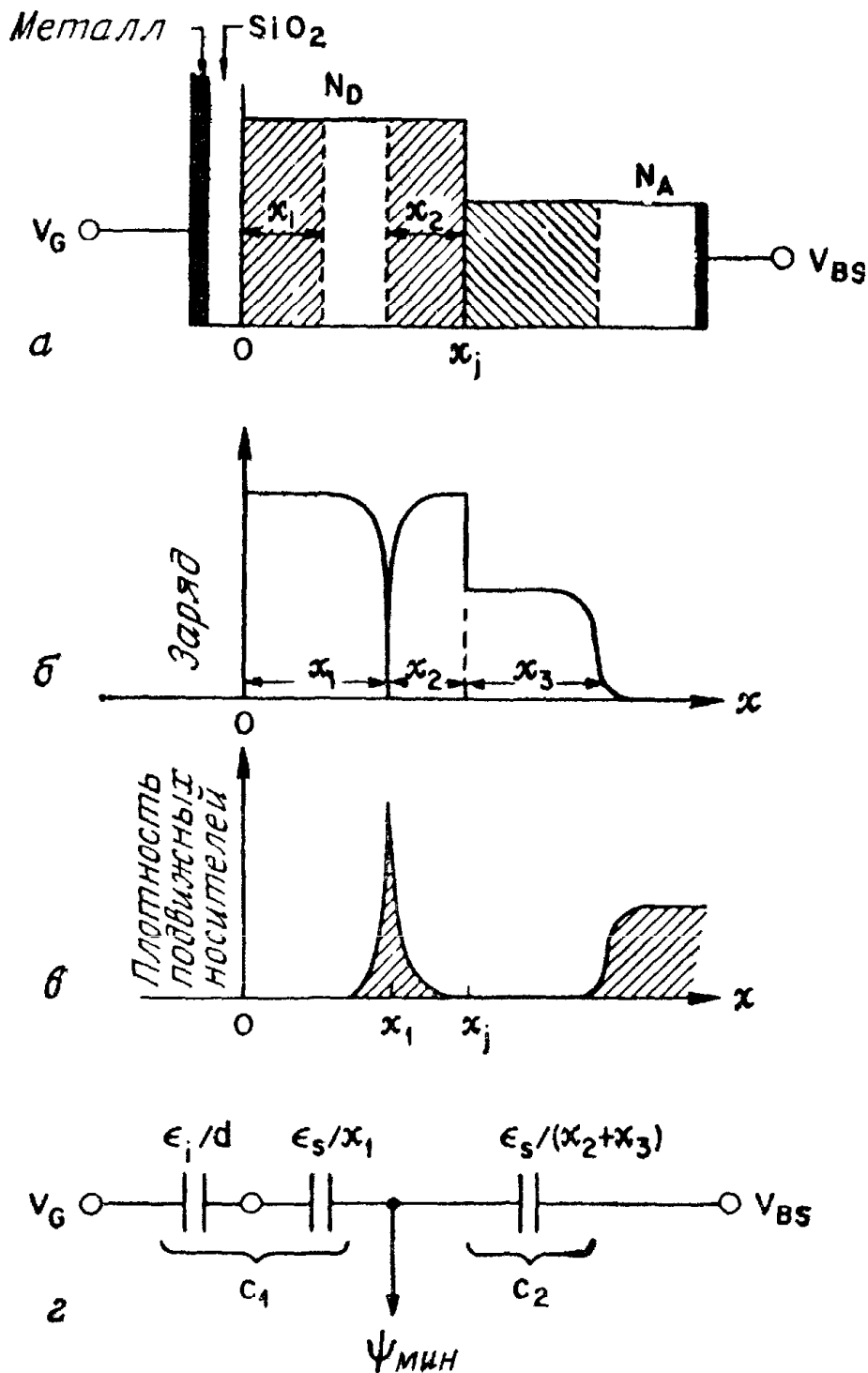


Рис. 31. МОП-транзистор со скрытым каналом в режиме, предшествующем отсечке канала [39]

а — профиль легирования; б — плотность заряда; в — плотность подвижных носителей; г — эквивалентные емкости.

рассеиваются на неоднородностях границы раздела с диэлектриком. Эти соображения подтверждаются результатами соответствующих измерений (рис. 32, а), из которых видно, что в приборах со скрытым каналом подвижность примерно на 50 % выше, чем в обычных МОП-транзисторах. На рис. 32, б для сравнения приведены зависимости порогового напряжения соответствующих приборов от длины канала L . В приборах с длинным каналом напряжение V_T не зависит от длины L . При малых L , однако,

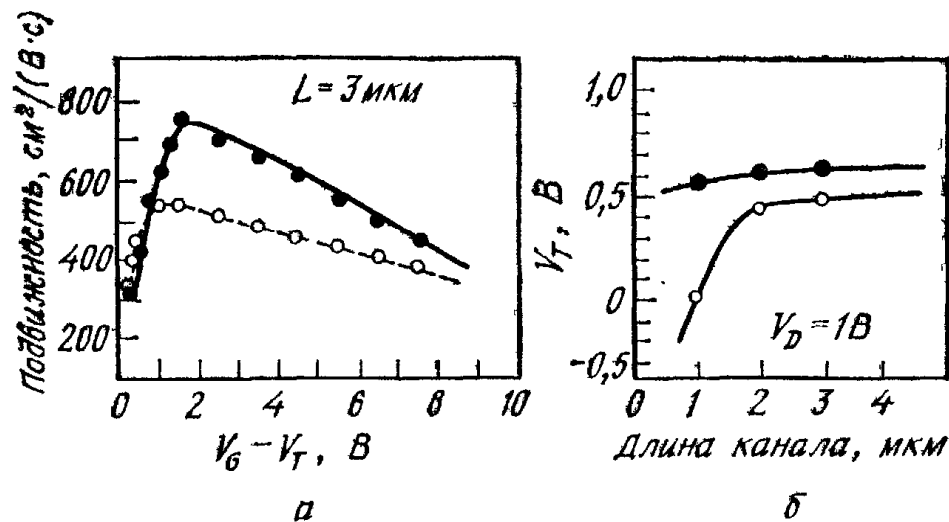


Рис. 32. Характеристики МОП-транзисторов с поверхностным (\circ) и скрытым (\bullet) каналами [40].

a — зависимость подвижности от напряжения затвора; *б* — зависимость порогового напряжения от длины канала.

происходит нежелательное уменьшение V_T с длиной канала, обусловленное так называемыми короткоканальными эффектами, которые мы рассмотрим в разд. 8.4. Однако уже теперь можно отметить большую «устойчивость» МОП-транзисторов со скрытым каналом к влиянию короткоканальных эффектов по сравнению с обычными приборами. Поэтому нормально закрытые МОП-транзисторы со скрытым каналом весьма полезны для высокочастотных применений.

8.4. КОРОТКОКАНАЛЬНЫЕ ЭФФЕКТЫ

С 1959 г., который можно считать «началом эры интегральных схем», до настоящего времени характерный размер отдельного полупроводникового прибора уменьшился по крайней мере на два порядка. Можно полагать, что такое сокращение будет продолжаться и в обозримом будущем (рис. 1). Однако по мере уменьшения длины канала МОП-транзистора свойства последнего начинают резко отличаться от свойств обычных длинноканальных приборов, рассмотренных в разд. 8.2. Эти отклонения — так называемые короткоканальные эффекты — обусловлены существенно двумерным характером распределения электрических полей в активной области и сравнительно высокими абсолютными значениями напряженности полей.

Если при неизменной концентрации легирующей примеси в подложке сокращать длину канала прибора, в конце концов она станет величиной порядка толщины обедненных слоев p — n -переходов стока и истока. При этом распределение потенциала в канале будет равным образом определяться поперечным полем \mathcal{E}_x , обусловленным напряжениями на затворе и подложке, и продольным полем \mathcal{E}_y , инициированным напряжением смещения стока

транзистора. Иными словами, распределение потенциала в таком короткоканальном приборе имеет двумерный характер, и для его описания уже нельзя использовать приближение плавного канала, которое предполагает, что $\mathcal{E}_x \gg \mathcal{E}_y$. Двумерный характер распределения потенциала существенно изменяет подпороговый участок характеристики прибора, обуславливает нежелательную зависимость порогового напряжения от длины канала и напряжений смещения на электродах, уменьшает выходное сопротивление, препятствуя отсечке канала.

При повышенных значениях электрических полей, характерных для короткоканальных приборов, становится важной полевая зависимость подвижности, которая в конечном итоге приводит к насыщению дрейфовой скорости (разд. 8.2). При еще больших полях в окрестности стокового перехода начинается ударная ионизация, становится существенной дополнительная проводимость по подложке и происходит так называемое включение паразитного биполярного транзистора. Высокие электрические поля приводят также к разогреву носителей и соответствующей инжекции горячих носителей в окисел. Такая зарядка окисла обуславливает сдвиг порога, дрейф характеристик и ухудшение крутизны прибора.

Все перечисленные короткоканальные эффекты усложняют работу прибора и ухудшают его рабочие характеристики. Следовательно, в практике конструирования следует стремиться к устранению или минимизации этих эффектов, чтобы «физически» короткоканальный прибор был «электрически» подобен длинноканальному.

8.4.1. Подпороговые токи

На рис. 33 приведены экспериментальные передаточные (т. е. $\ln I_D$ от V_G) характеристики нескольких МОП-транзисторов, изготовленных по стандартной n -канальной технологии на пластинах p -кремния с ориентацией $\langle 100 \rangle$. После выращивания подзатворного окисла заданной толщины с помощью рентгеновской литографии на пластинах формировались поликремниевые затворы длиной от 1 до 10 мкм и одинаковой шириной 70 мкм. Переходы стока и истока были получены имплантацией ионов мышьяка с последующим отжигом. Глубина переходов варьировалась от 0,25 до 1,5 мкм в зависимости от энергии имплантации и режима отжига. Для контактной металлизации был использован алюминий.

В МОП-транзисторах, характеристики которых приведены на рис. 33, a , толщина окисла составляла 130 Å, концентрация легирующей примеси в подложке 10^{15} см⁻³, а глубина переходов 0,33 мкм. Прибор с каналом длиной 7 мкм демонстрирует типичное

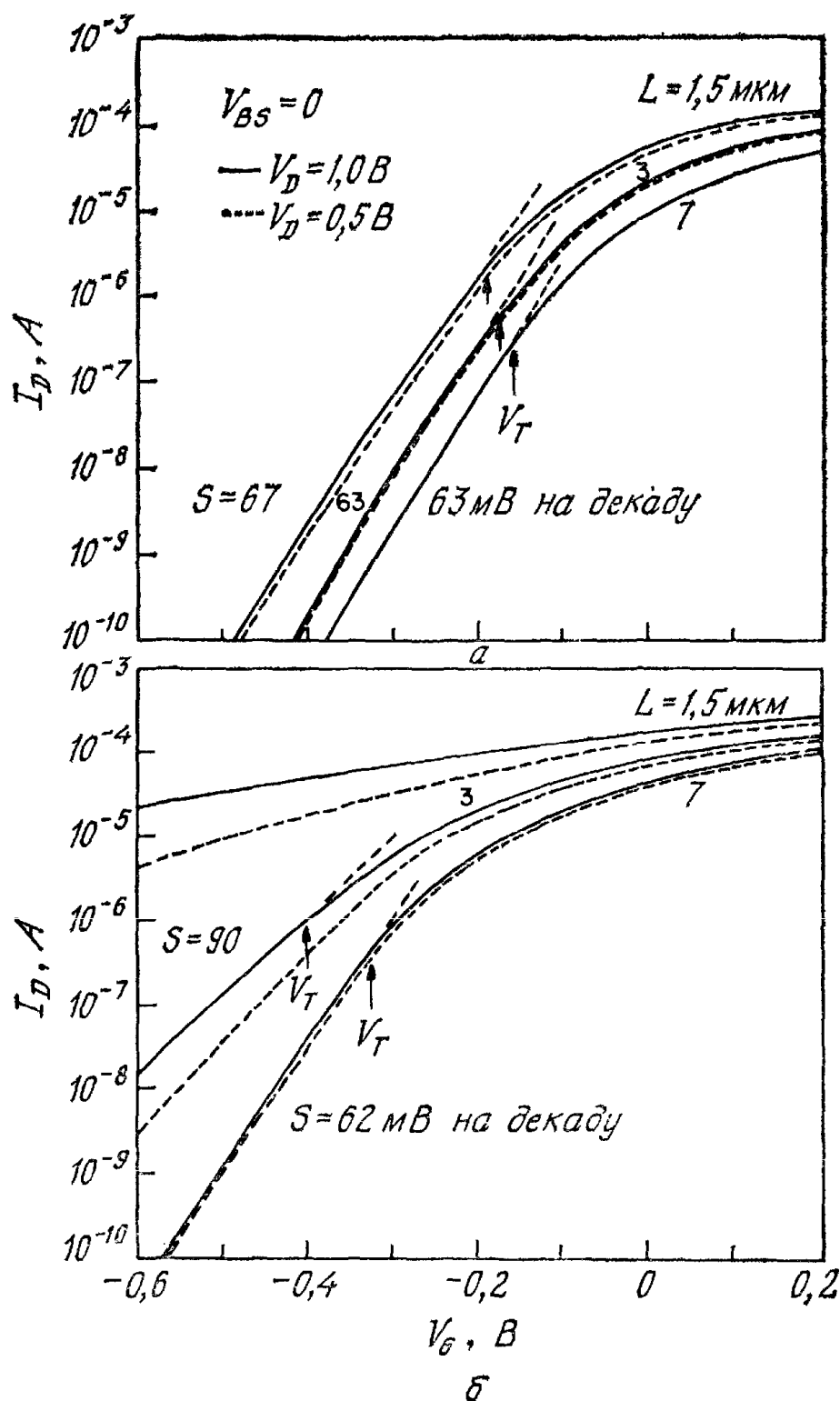


Рис. 33. Подпороговые характеристики МОП-транзисторов с различной длиной канала.

$a - N_B = 10^{16} \text{ см}^{-3}$; $b - N_B = 10^{14} \text{ см}^{-3}$.

длинноканальное поведение. Его подпороговая характеристика не зависит от напряжения стока, как и должно быть при $V_D > > 3kT/q$ согласно формуле (40). В приборе с $L = 3 \mu\text{m}$ изменение V_D немного сдвигает подпороговый участок. Несколько больший сдвиг при изменении напряжения стока от $V_D = 0,5$ до $1,0 \text{ В}$ наблюдается в транзисторе с самым коротким ($L = 1,5 \mu\text{m}$) каналом. Соответственно изменяются и пороговые напряжения V_T ,

определенные здесь как точка (показанная на рис. 33 стрелками), в которой $I-V$ -зависимость начинает отклоняться от прямой линии. При увеличении V_D уменьшается также крутизна подпорогового участка (S увеличивается).

Иные характеристики имеют приборы с идентичными (за исключением концентрации примеси в подложке $N_B = 10^{14} \text{ см}^{-3}$) параметрами (рис. 33, б). Здесь небольшой сдвиг характеристик при изменении V_D наблюдается уже в приборе с $L = 7$ мкм. В МОП-транзисторе с каналом длиной 3 мкм подпороговые токи и соответствующие S существенно увеличены. И наконец, характеристики прибора с самым коротким каналом ($L = 1,5$ мкм) кардинально отличаются от характеристик обычных длинноканальных приборов (прибор «не выключается»). Критерий, разграничивающий длинно- и короткоканальные приборы, можно попытаться определить, используя одно из двух характерных свойств обычного длинноканального МОП-транзистора: 1) обратно пропорциональную зависимость тока стока от длины канала $I_D \sim 1/L$; 2) независимость подпорогового тока обычного прибора от напряжения стока при $V_D > 3kT/q$. На рис. 34 приведены зависимости I_D и $\Delta I_D/I_D$ от $1/L$, где I_D — ток стока транзисторов при напряжении на затворе V_G , равном пороговому напряже-

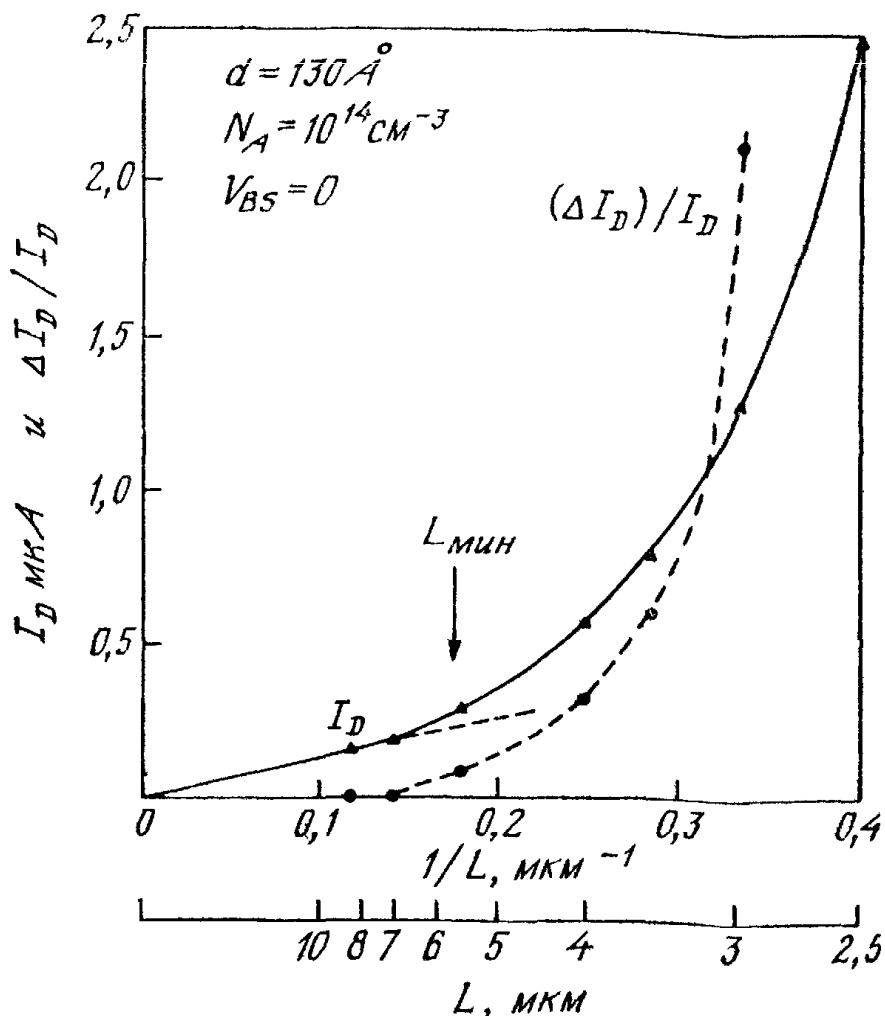


Рис. 34. Зависимость тока стока и отношения $\Delta I_D/I_D$ от величины $1/L$.

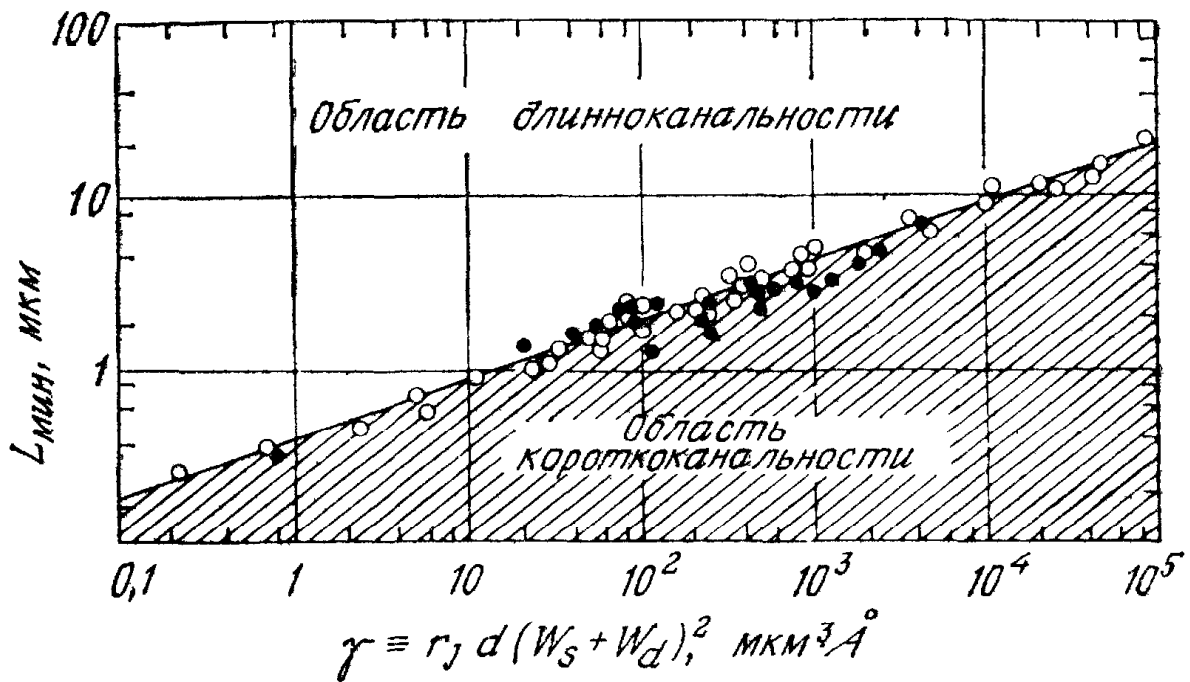


Рис. 35. Зависимость минимальной длинноканальной длины $L_{\text{мин}}$ от параметра $\gamma \equiv r_j d (W_S + W_D)^2$ [41].

нию V_T , а $\Delta I_D / I_D$ — относительная разность токов при том же $V_G = V_T$ и двух различных напряжениях стока V_D . Началом короткоканального поведения будем считать момент, когда отклонение тока стока от длинноканальной зависимости $I_D \sim 1/L$ составляет 10 % или когда относительная разность $\Delta I_D / I_D$ равна 0,1.

Результаты большого количества измерений, выполненных в МОП-транзисторах, параметры которых варьировались в широких пределах (толщина подзатворного окисла 100—1000 Å, концентрация примеси в подложке 10^{14} — 10^{17} см $^{-3}$, глубина переходов 0,18—1,5 мкм при напряжениях на стоке до 5 В), обобщаются довольно простым эмпирическим соотношением [41]

$$L_{\text{мин}} = 0,4 [r_j d (W_S + W_D)^2]^{1/3} \equiv 0,4 (\gamma)^{1/3}. \quad (83)$$

Здесь $L_{\text{мин}}$ — минимальная длина канала, при которой подпороговый участок еще сохраняет длинноканальный характер, мкм; r_j — глубина переходов, мкм; d — толщина слоя окисла, Å; $W_S + W_D$ — сумма толщин обедненных слоев стока и истока, вычисленная в приближении одномерного резкого p — n -перехода:

$$W_D = \sqrt{\frac{2\epsilon_s}{qN_A} (V_D + V_{bi} + V_{BS})} \text{ [мкм]}, \quad (84)$$

где V_{bi} — контактная разность перехода, а V_{BS} — смещение на подложке. При нулевом смещении стока ($V_D = 0$) толщина W_D равна W_S .

На рис. 35 приведены для сравнения результаты соответствующих экспериментальных измерений (—●—) и вычислений по

эмпирическому соотношению (83). Здесь также приведены результаты двумерных машинных расчетов зависимости $L_{\text{мин}}$ от γ (—○—), полученные на основе сформулированных выше критериев короткоканального поведения. Отметим, что в самых худших случаях экспериментальные и расчетные результаты отличаются от эмпирического соотношения (83) не более чем на 20 %, а для основного числа точек наблюдается значительно лучшее соответствие. Следовательно, эмпирическую формулу (83) можно использовать в качестве главного ориентира в практике миниатюризации МОП-транзисторов. Все МОП-транзисторы с параметрами, которым соответствует точка в заштрихованной области рис. 35, по своим электрическим свойствам будут короткоканальными приборами. МОП-транзисторы с параметрами, попадающими в незаштрихованную область рис. 35, в электрическом смысле являются длинноканальными. Так, например, свойства прибора с $L = 10$ мкм, у которого $\gamma = 10^5$ мкм³·Å, будут короткоканальными, в то время как МОП-транзистор с $L = 0,5$ мкм, но с $\gamma = 1$ мкм³·Å будет вести себя как длинноканальный прибор.

В первом приближении отклонения от длинноканального поведения можно проанализировать, воспользовавшись принципом электронейтральности (рис. 36, а) [42]

$$Q'_M + Q'_O + Q'_n + Q'_B = 0, \quad (85)$$

где Q'_M — полный заряд на затворе, Q'_O — полный эффективный (отнесенный к границе раздела Si—SiO₂) заряд окисла, Q'_n — полный заряд инверсионного слоя, а Q'_B — полный эффективный заряд ионизированных примесей в области обеднения под затвором. Поделив на $C_i A$, это соотношение можно записать в виде

$$V_G = V_{FB} + \psi_s + Q'_B / C_i A, \quad (86)$$

где V_{FB} — напряжение плоских зон, ψ_s — поверхностный потенциал и A — площадь. Пороговое напряжение V_T получим, подставив в выражение (86) $\psi_s = 2\psi_B$:

$$V_T = V_{FB} + 2\psi_B + Q'_B / C_i A. \quad (87)$$

Для длинноканального прибора $Q'_B = qN_A W A$, где W — толщина обедненного слоя. В приборах с коротким каналом влияние Q'_B на пороговое напряжение уменьшается, поскольку в краевых областях канала вблизи стока и истока часть силовых линий с заряженных примесных атомов «уходит в сторону», закорачиваясь на высоколегированных n^+ -областях, а не на затворе, как это показано на рис. 36, а для $V_D = 0$. Естественно, что при $V_D > 0$ обедненная область вблизи стока шире, чем вблизи истока (рис. 36, б). Отметим, что с учетом влияния поперечных электрических полей на распределение потенциала в при-

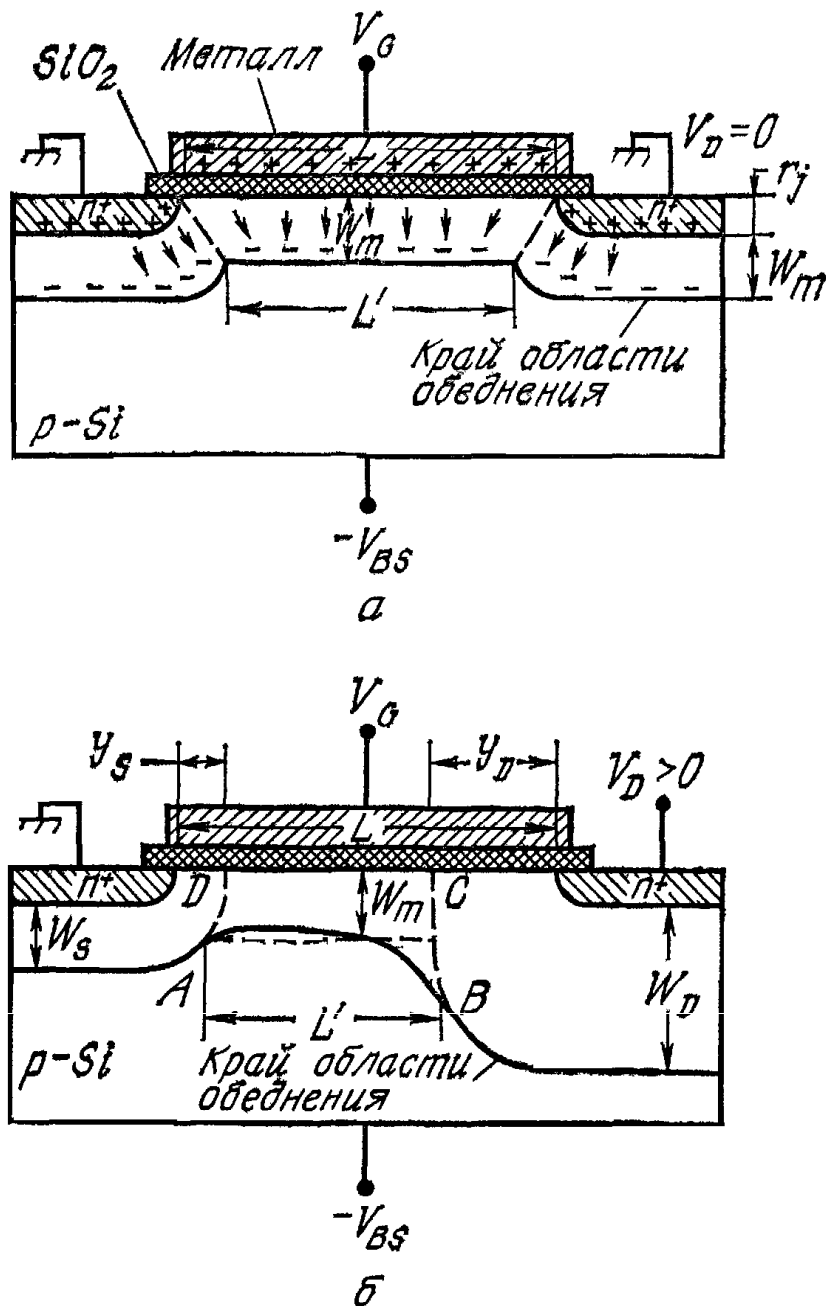


Рис. 36. Модель принципа электронейтральности при рассмотрении короткоканальных эффектов [42].

а) $V_D = 0$; б) $V_D > 0$.

поверхностном слое «горизонтальные» толщины обеднения y_S и y_D оказываются несколько меньше «вертикальных» глубин обеднения W_S и W_D соответственно.

В результате уменьшения эффективного заряда обедненного слоя Q'_B поверхностный потенциал ψ_s для данного значения напряжения на затворе увеличивается, что приводит к соответствующему увеличению подпороговых токов (выражения (37) и (38)). Поверхностный потенциал можно записать следующим образом [43]:

$$V_G - V_{FB} = \psi_s + \frac{1}{C_i} \sqrt{q\epsilon_s N_A (\psi_s + V_{BS})/2} \left(1 + \frac{L - W_D - W_S}{L - y_D - y_S} \right), \quad (88)$$

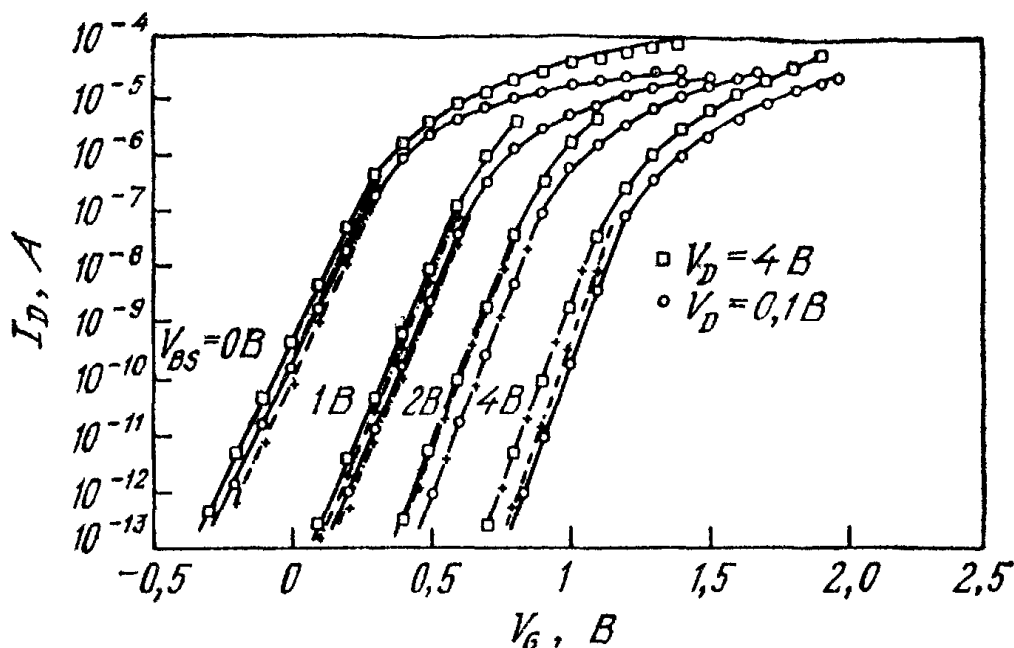


Рис. 37. Подпороговые токи в короткоканальном МОП-транзисторе с $d = 530 \text{ \AA}$, $N_A = 1,2 \cdot 10^{16} \text{ см}^{-3}$, $L = 2,56 \text{ мкм}$ и $Z = 21,5 \text{ мкм}$ [22].

—+— по формуле (90); — двумерное моделирование.

где W_D и W_S определяются выражением (84), а

$$y_S \approx \sqrt{\frac{2\epsilon_s}{qN_A} (V_{bi} - \psi_s)} \quad (89a)$$

и

$$y_D \approx \sqrt{\frac{2\epsilon_s}{qN_A} (V_{bi} - \psi_s + V_D)}. \quad (89b)$$

Для подпорогового тока будем иметь

$$I_D = \mu_n \left(\frac{Z}{L - y_S - y_D} \right) \frac{aC_i}{2\beta^2} \left(\frac{n_i}{N_A} \right)^2 (1 - e^{-\beta V_D}) e^{\beta \psi_s} (\beta \psi_s)^{1/2}. \quad (90)$$

Это выражение отличается от выражения (40) только тем, что в нем фигурирует меньшая длина канала:

$$L_{\text{eff}} = L - y_S - y_D. \quad (90a)$$

Для сравнения на рис. 37 приведены экспериментальные подпороговые кривые, характеристики, рассчитанные по формуле (90), и соответствующие результаты, полученные двумерным численным моделированием. Все они находятся в достаточно близком соответствии друг с другом.

8.4.2. Пороговые напряжения

В первом приближении величину порогового напряжения короткоканального прибора можно оценить, используя простые геометрические соображения [42]. Будем считать, что эффектив-

ный полный заряд Q'_B равен заряду ионизированных примесей в трапецеидальной области под затвором (рис. 36, а):

$$\frac{Q'_B}{Z} = qN_A W_m \left(\frac{L+L'}{2} \right). \quad (91)$$

Из обычных тригонометрических соотношений получим для $L+L'$ следующее выражение:

$$\frac{L+L'}{2L} = 1 - \frac{r_j}{L} \left(\sqrt{1 + \frac{2W_m}{r_j}} - 1 \right). \quad (92)$$

При этом сдвиг порогового напряжения

$$\begin{aligned} \Delta V_T &= \frac{1}{C_i} \left(\frac{Q'_B}{ZL} - qN_A W_m \right) = - \frac{qN_A W_m}{C_i} \left(1 - \frac{L+L'}{2L} \right) = \\ &= - \frac{qN_A W_m r_j}{C_i L} \left(\sqrt{1 + \frac{2W_m}{r_j}} - 1 \right). \end{aligned} \quad (93)$$

Для учета влияния напряжения стока и смещения на подложке выражение (93) запишем в модифицированной форме [22]:

$$\Delta V_T = - \frac{qN_A W_m r_j}{2C_i L} \left[\left(\sqrt{1 + \frac{2y_S}{r_j}} - 1 \right) + \left(\sqrt{1 + \frac{2y_D}{r_j}} - 1 \right) \right], \quad (94)$$

где y_S и y_D определяются выражениями (89), а

$$W_m = \sqrt{2\epsilon_s (2\psi_B + V_{BS}) / qN_A}. \quad (95)$$

Как следует из рис. 38, выражение (94) достаточно хорошо описывает результаты соответствующих измерений.

Определенное влияние на пороговое напряжение оказывает также краевой эффект на боковой границе затвора. Соответствующий сдвиг порога при этом обусловлен тем фактом, что область обеднения выходит (в направлении z , перпендикулярном току I_D) за боковой край затвора (вставка на рис. 39). Предполагая эту краевую часть области пространственного заряда цилиндрической [44], для полного заряда в области обеднения будем иметь

$$Q_{BT} = qN_A ZLW \left(1 + \frac{\pi}{2} \frac{W}{Z} \right). \quad (96)$$

Из этого выражения видно, что боковое расширение области пространственного заряда описывается соответствующим нагрузочным коэффициентом $1 + \pi W/2Z$. В результате наблюдается соответствующее увеличение порогового напряжения:

$$V_T = V_{FB} + 2\psi_B + \frac{\sqrt{2\epsilon_s qN_A (2\psi_B + V_{BS})}}{C_i} \left(1 + \frac{\pi}{2} \frac{W}{Z} \right). \quad (97)$$

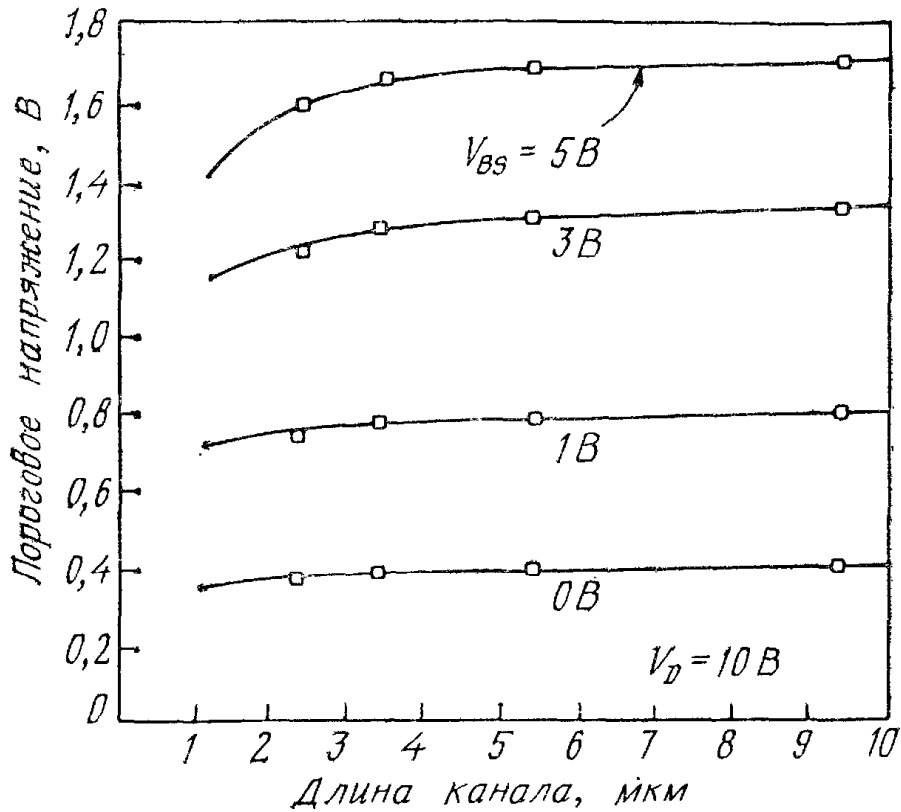


Рис. 38. Зависимость порогового напряжения от длины канала для МОП-транзисторов с $d = 360 \text{ \AA}$, $N_A = 2,5 \cdot 10^{16} \text{ см}^{-3}$ и $L = 2,4-9,4 \text{ мкм}$ [22].

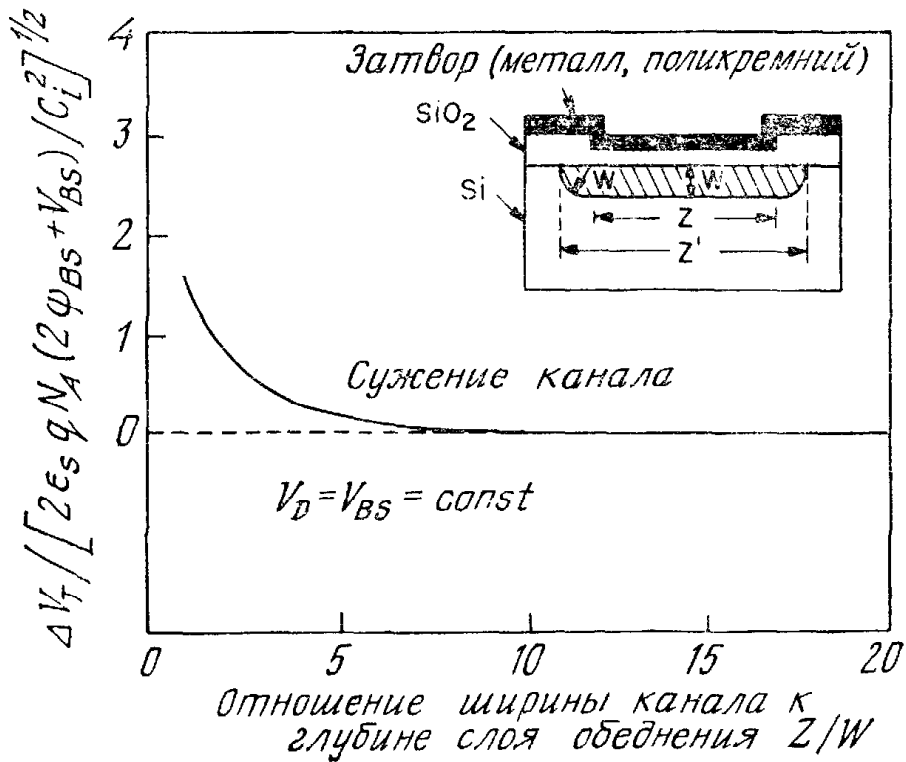


Рис. 39. Влияние бокового сужения канала на пороговое напряжение [44].

Влияние бокового сужения канала проиллюстрировано на рис. 39. Ясно, что заметного эффекта в этом случае можно ожидать, когда ширина канала Z становится сравнимой с характерной глубиной слоя обеднения W . Так, например, для прибора с шириной затвора $L = 1$ мкм соответствующее увеличение V_T становится существенным при уровнях легирования подложки, меньших 10^{16} см $^{-3}$.

8.4.3. Линейная область и режим насыщения

На рис. 40 приведены выходные вольт-амперные характеристики МОП-транзистора с длиной канала 0,73 мкм и толщиной слоя окисла 258 Å [45]. В область канала прибора сквозь окисел была имплантирована доза ионов бора $2 \cdot 10^{12}$ см $^{-2}$ с энергией 150 кэВ. Результирующий профиль концентрации можно было считать ступенчатым с глубиной $x_s = 0,3$ мкм и поверхностной концентрацией $7 \cdot 10^{16}$ см $^{-3}$. При напряжениях стока 0—5 В параметр γ (выражение (83)) изменялся в пределах 3—10 мкм $^3 \cdot \text{Å}$, что соответствует $L_{\text{мин}} = 0,6—0,9$ мкм. Следовательно, рассматриваемый прибор (в соответствии с классификацией, данной на рис. 35) занимает промежуточное положение у края области, соответствующей длинноканальным приборам. В первом приближении токи насыщения транзистора можно вычислить, используя выражение (28) и подставив в него вместо L редуцированную эффективную длину канала (выражение (90a)). С ростом V_D

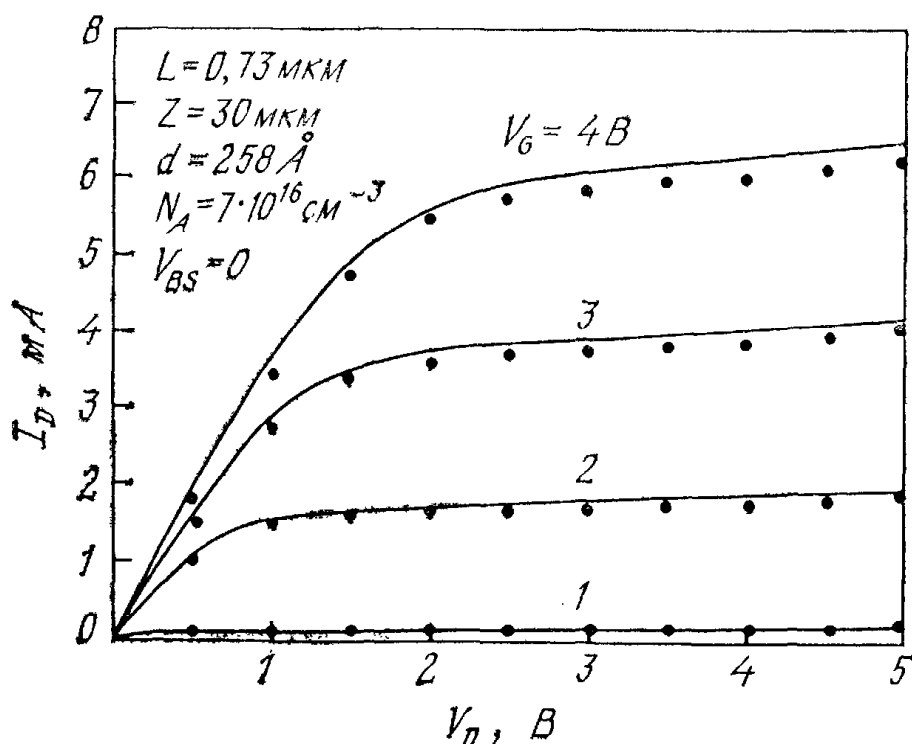


Рис. 40. Стоковые характеристики МОП-транзистора с длиной канала 0,73 мкм [45].

— экспериментальные; ●●● теоретические.

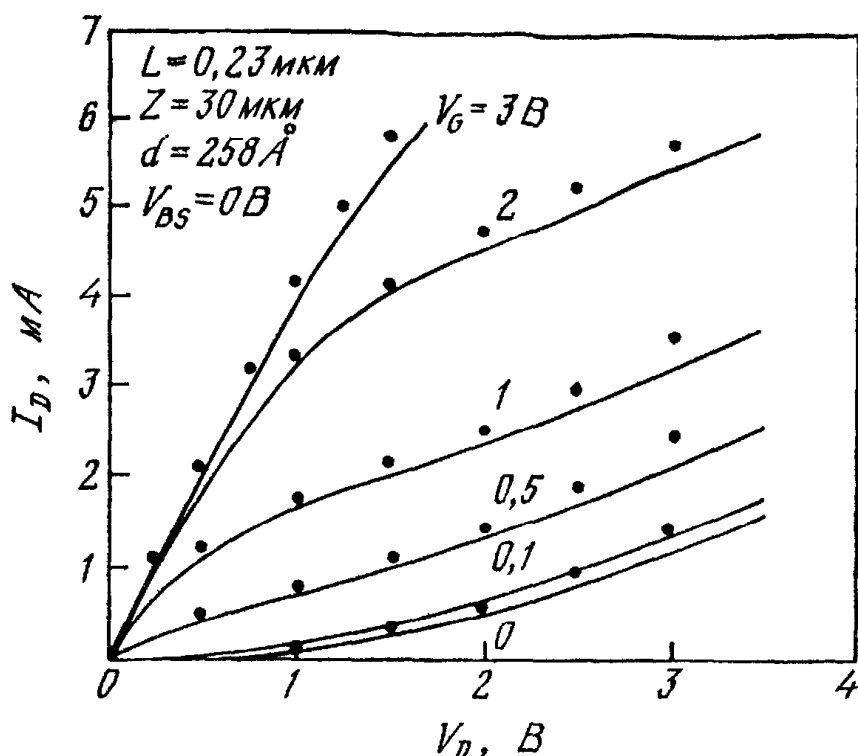


Рис. 41. Стоковые характеристики МОП-транзистора с длиной канала 0,23 мкм [45].

— экспериментальные; ●●● теоретические.

длина L_{eff} уменьшается, что обуславливает соответствующее увеличение тока стока I_D .

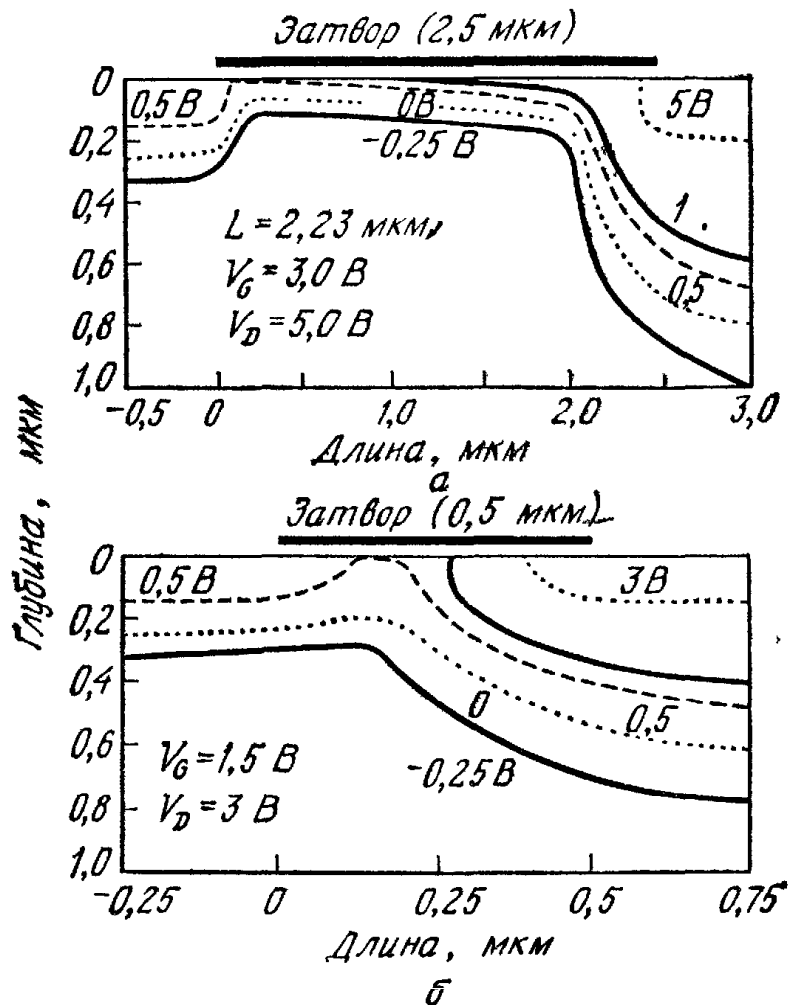
На рис. 41 приведены $I-V$ -характеристики прибора, все параметры которого идентичны параметрам МОП-транзистора с характеристиками, приведенными на рис. 40, за исключением длины канала, которая в данном случае составляла 0,23 мкм. Характеристики этого прибора демонстрируют ярко выраженное короткоканальное поведение. В нем уже при $V_D = 0$ происходит смыкание обедненных областей стока и истока, поскольку соответствующая сумма $y_S + y_D = 0,26$ мкм больше длины канала L . Следовательно, и во всем диапазоне стоковых напряжений (рис. 41) рассматриваемый прибор работает в режиме смыкания (прокола). В этом случае электроны из истока будут инжектироваться непосредственно в объем обедненной области канала, где они подхватываются электрическим полем и сразу же коллектируются стоком. Напряжение на стоке, требуемое для такого смыкания [10], равно

$$V_{pt} \simeq \frac{qN_A(L - y_S)^2}{2\epsilon_s} - V_{bt}. \quad (98)$$

Если V_D существенно превышает эту величину, рассматриваемый ток инжекции ограничивает накопление заряда подвижных носителей в области пролета. В этом так называемом режиме ограничения тока пространственным зарядом

$$I_D \simeq 9\epsilon_s\mu_nAV_D^2/8L^3, \quad (99)$$

Рис. 42. Распределение потенциала в структурах, характеристики которых приведены на рис. 40 и 41 [45].
 а — длинноканальная структура с $L = 2,23$ мкм; б — короткоканальная структура с $L = 0,23$ мкм.



где параметр A по порядку величины равен площади $n^+ - p$ -перехода истока. Ток инжекции I_D , увеличивающийся пропорционально V_D^2 , течет в объеме полупроводника параллельно току инверсионного канала, пропорциональному напряжению на затворе V_G .

Теоретические точки на рис. 40 и 41 получены в результате сложного двумерного численного моделирования при учете эффекта смыкания и полевой зависимости подвижности без использования каких-либо подгоночных параметров. Поэтому следует особо отметить хорошее соответствие этих результатов экспериментальным данным. На рис. 42, а и б приведены двумерные распределения потенциала в структурах длинноканального МОП-транзистора и рассмотренного выше короткоканального прибора с $L = 0,23$ мкм. Все параметры обоих приборов идентичны, за исключением длины канала, которая в длинноканальном приборе ($L = 2,3$ мкм) была в 10 раз больше, чем в короткоканальном. Отметим, что длина канала в обоих приборах примерно на 0,27 мкм короче длины соответствующих затворов в результате боковой диффузии имплантированных ионов под край затвора. В длинноканальном приборе (рис. 42, а) эквипотенциали в канале параллельны друг другу и сосредоточены

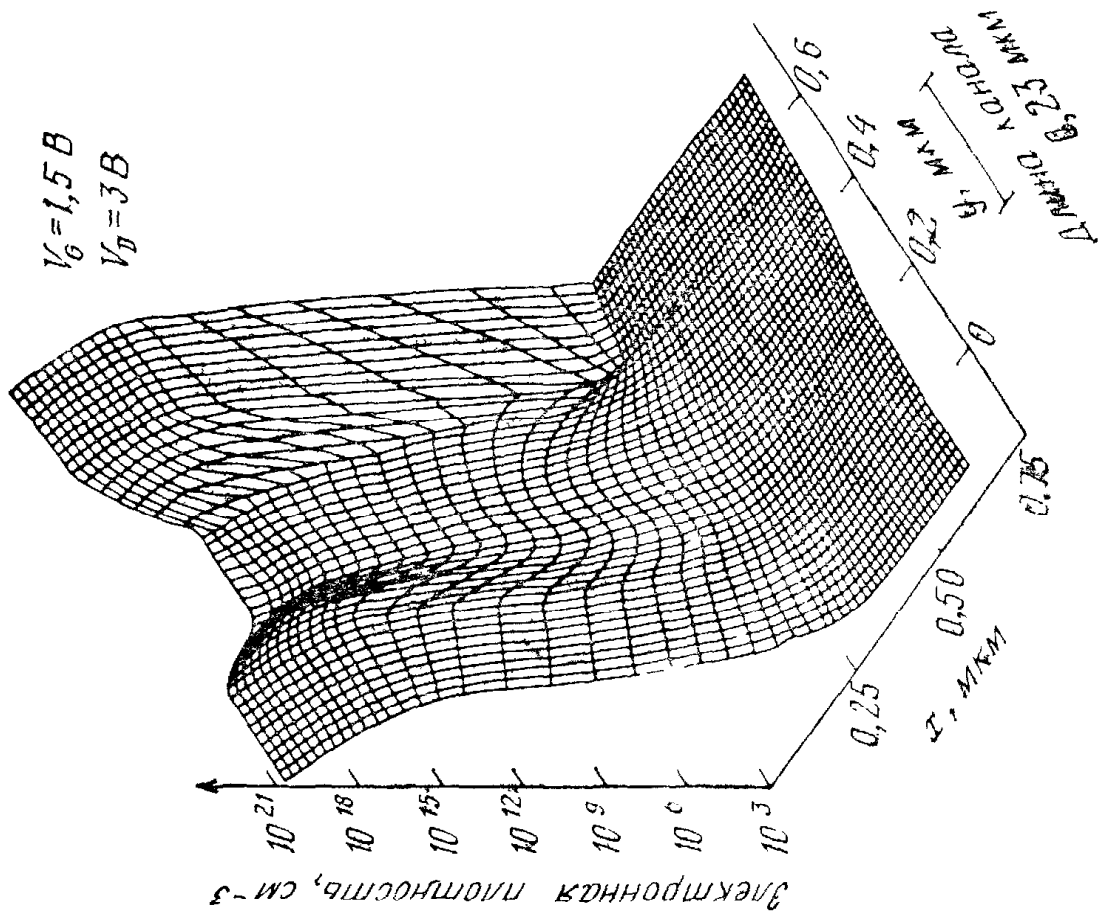


Рис. 44. Двумерное распределение электронной плотности в МОП-транзисторе с $L = 0,23$ мкм [45].

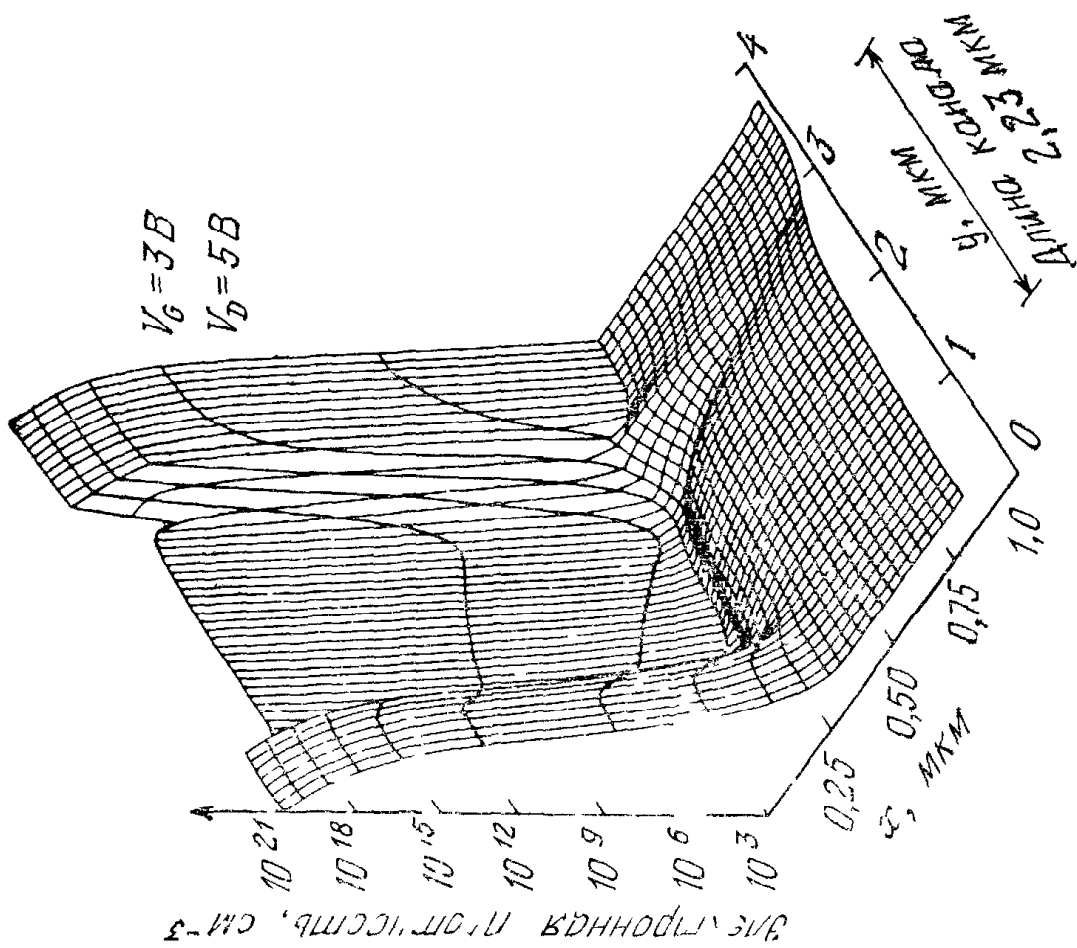


Рис. 43. Двумерное распределение электронной плотности в МОП-транзисторе с $L = 2,23$ мкм [45].

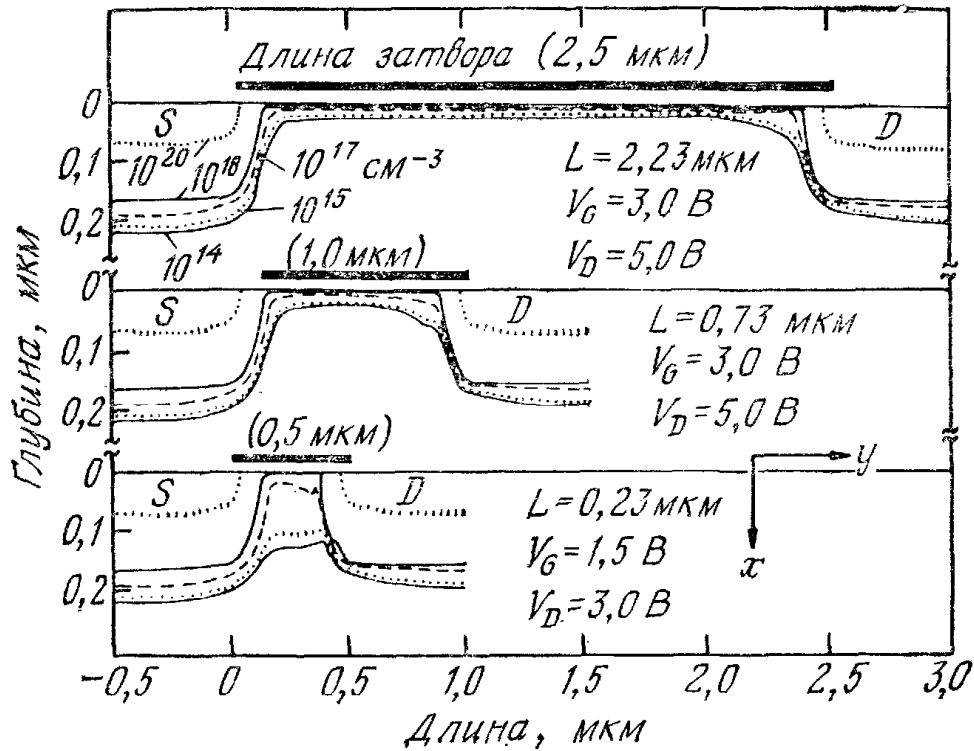


Рис. 45. Линии постоянной электронной плотности в МОП-транзисторах с длиной канала 2,23; 0,73 и 0,23 мкм [45].

в непосредственной близости от границы раздела с окислом. Напротив, в короткоканальном приборе (рис. 42, б) эквипотенциали уходят глубоко в объем подложки, что указывает на эффект смыкания.

На рис. 43 и 44 приведено распределение электронной плотности в рассматриваемых приборах. Из рис. 43, соответствующего режиму насыщения длинноканального прибора, ясно видно характерное уменьшение электронной плотности у края стока (отсечка канала). В длинноканальном приборе (рис. 43) электроны локализованы у границы с окислом, в то время как в короткоканальном МОП-транзисторе (рис. 44) электронная плотность довольно глубоко простирается в толщу подложки.

На рис. 45 показаны линии одинаковой электронной плотности в плоскости $x-y$, рассчитанные для рассматриваемых МОП-транзисторов с длиной канала 2,23; 0,73 и 0,23 мкм. Видно, что, как уже говорилось выше, в длинноканальном МОП-транзисторе с длиной затвора 2,5 мкм электронный ток сосредоточен в слое толщиной $\sim 200 \text{ \AA}$ у границы раздела. Отметим также, что у края стока, где происходит отсечка канала, часть линий тока уходит от поверхности в объем подложки. В приборе с длиной затвора 1 мкм ток все еще остается сосредоточенным у границы с окислом, однако область отсечки уже захватывает значительно большую часть длины канала. При уменьшении длины канала до 0,23 мкм (длина затвора 0,5 мкм) линии тока уходят глубоко в подложку, а свойства прибора теряют длинноканальный характер.

В заключение следует констатировать, что из-за сложной природы двумерного распределения потенциала выразить в простой аналитической форме вольт-амперные характеристики короткоканальных приборов невозможно. Подходящие приближенные выражения можно получить, определенным образом модифицируя основное выражение (22), тем или иным способом «выделяя» действующие заряды (например так, как это показано на рис. 36). Обычно в таких упрощенных аналитических моделях удается путем подбора подходящих значений подгоночных параметров обеспечить неплохое соответствие с экспериментальными данными [46—48].

8.4.4. Умножение носителей и зарядка окисла

В длинноканальных приборах при достаточно высоком напряжении на стоке в области отсечки у границы канала со стоком происходит слабый лавинный пробой. Генерированные при этом электроны сразу же уходят в сток, а генерированные дырки коллектируются подложкой, создавая ток подложки I_{BS} . Зависимость этого тока от напряжения на затворе в длинноканальном приборе с $L = 10$ мкм приведена на рис. 46 [49]. Здесь показаны также полный ток стока I_D транзистора и генерационно-рекомбинационный ток в области обеднения. Диапазон изменения тока на рис. 46 перекрывает все характерные режимы

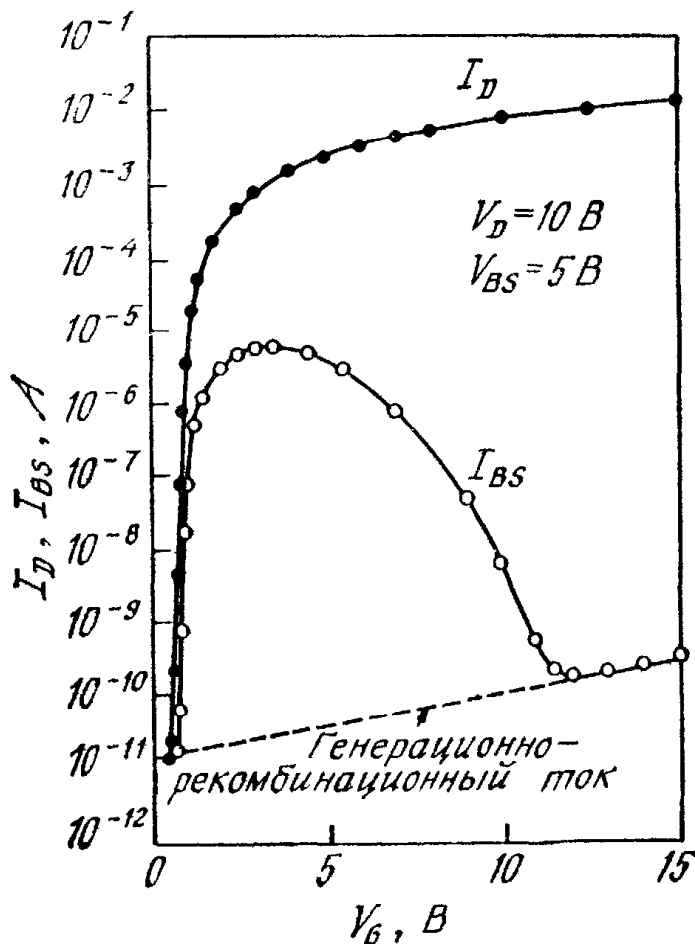


Рис. 46. Зависимость тока стока и тока подложки в длинноканальном МОП-транзисторе от напряжения на затворе [49].

работы прибора (подпороговый участок, линейную область и насыщение). При увеличении напряжения на затворе ток подложки I_{BS} сначала увеличивается, а затем, проходя через максимум, снова уменьшается. Такое поведение тока I_{BS} можно объяснить следующим образом. Предположим для простоты, что ударная ионизация происходит однородно по длине области отсечки канала ΔL . Тогда для тока подложки I_{BS} можно написать

$$I_{BS} = I_D \alpha \Delta L, \quad (100)$$

где α — коэффициент ударной ионизации, т. е. число электронно-дырочных пар, генерируемых электроном на единице длины. При заданном напряжении на стоке и при увеличении V_G одновременно увеличиваются ток I_D и напряжение $V_{D\text{sat}}$. Однако с ростом $V_{D\text{sat}}$ уменьшается поверхностное поле в области отсечки $(V_D - V_{D\text{sat}})/L$, что, естественно, приводит к уменьшению α . Следовательно, имеются два противоположных фактора: 1) рост I_D обуславливает увеличение I_{BS} при малых V_G и 2) уменьшение α приводит к снижению I_{BS} при больших напряжениях на затворе. Максимуму I_{BS} соответствует напряжение на затворе, при котором действие обоих факторов сбалансировано.

В короткоканальных приборах возникает дополнительный эффект, обусловленный током дырок, генерированных в области пробоя. Дело в том, что при малом расстоянии между стоком и истоком и при достаточном напряжении стока V_D значительная часть этих дырок увлекается электрическим полем в исток, а не в подложку (рис. 47) [50]. Встречая на пути энергетический барьер $n^+ - p$ -перехода истока, дырки накапливаются у его внешней границы и «открывают» переход. В результате исток начинает инжектировать в объем подложки электроны, которые коллектируются стоком. Все это представляет собой механизм «включения» паразитного биполярного $n - p - n$ (исток—подложка—сток)-транзистора в короткоканальных приборах. Ясно, что развитие этого процесса приводит к электрическому пробую структуры, начало которого определяется условием

$$\alpha_{npn} M = 1. \quad (101)$$

Здесь α_{npn} — коэффициент усиления тока биполярного транзистора в схеме с общей базой:

$$\alpha_{npn} = \text{sech}(L_G/L_{\text{diff}}) \simeq 1 - \frac{L_G^2}{2L_{\text{diff}}^2}, \quad (102)$$

где L_G — эффективная толщина базы, равная по порядку величины длине затвора, а L_{diff} — диффузионная длина в подложке. Коэффициент умножения M можно записать в виде

$$M = \left[1 - \left(\frac{BV_{CE0}}{V_{D\text{sub}}} \right)^n \right]^{-1}, \quad (103)$$

где BV_{CE0} — напряжение пробоя биполярного транзистора в схеме с общим эмиттером и разомкнутой базой, а $V_{D\text{sub}}$ — напряжение пробоя перехода сток—подложка. Из выражений (101)—(103) для результирующего напряжения пробоя короткоканального МОП-транзистора получим

$$V_B = BV_{CE0} \approx \frac{V_{D\text{sub}}}{2^{1/n}} \left(\frac{L_G}{L_{\text{diff}}} \right)^{2/n}. \quad (104)$$

На рис. 48 приведены результаты измерений напряжения пробоя в структурах с различной длиной канала L . Удовлетворительное согласие этих данных с результатами расчета по формуле (104) получается при $n = 5,43$. Различие пробивных напряжений структур с различной глубиной стокового перехода можно объяснить зависимостью $V_{D\text{sub}}$ от кривизны перехода (гл. 2).

Другим эффектом, существенным для работы короткоканальных приборов, который возникает при больших полях, является зарядка окисла [51]. Оказывается, что при высоких продольных электрических полях часть электронов, двигаясь в канале, разогревается до энергии, достаточной для преодоления энергетического барьера на границе Si—SiO₂ (3,1 эВ), и инжектируется в окисел (рис. 49 (вставка)). Горячие электроны могут инжектироваться в окисел также из электронно-дырочной плазмы лавинного пробоя в области отсечки канала у границы со стоком. Кроме того, в SiO₂ могут инжектироваться электроны, термически гене-

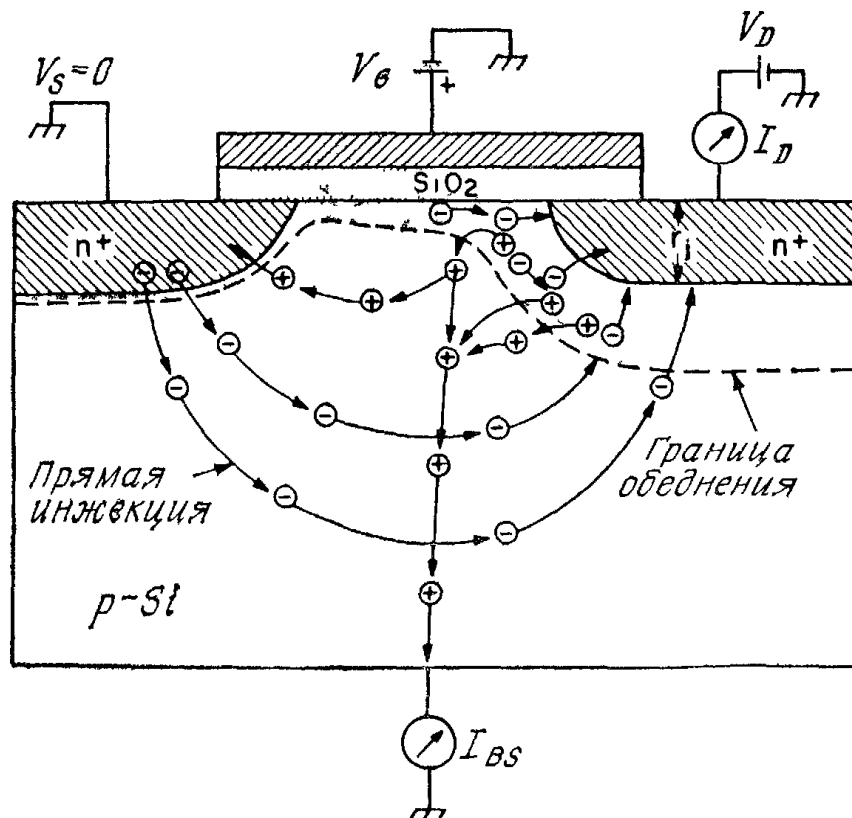


Рис. 47. Эффект паразитного биполярного транзистора [50].

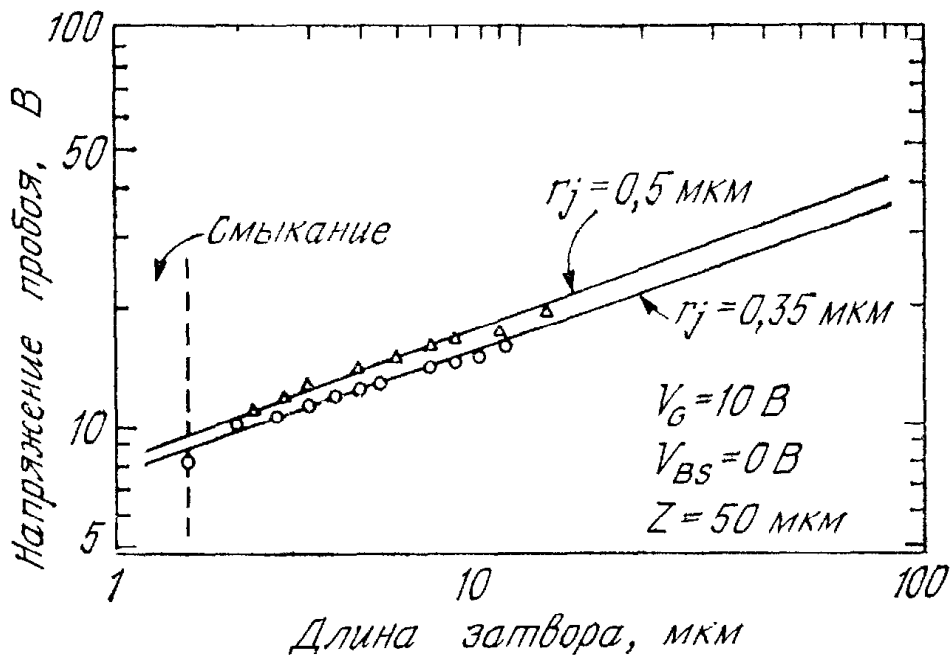


Рис. 48. Зависимость напряжения пробоя от длины затвора, обусловленная эффектом паразитного биполярного транзистора [50].

рированные в объеме подложки и разогретые большим поперечным электрическим полем на пути к границе раздела. Влияние инжекции горячих электронов в окисел на свойства МОП-транзистора показано на рис. 49. Отметим, что, во-первых, пороговое напряжение при зарядке окисла становится более положительным и, во-вторых, снижается крутизна МОП-транзистора (меньший наклон $I-V$ -зависимости) вследствие уменьшения подвижно-

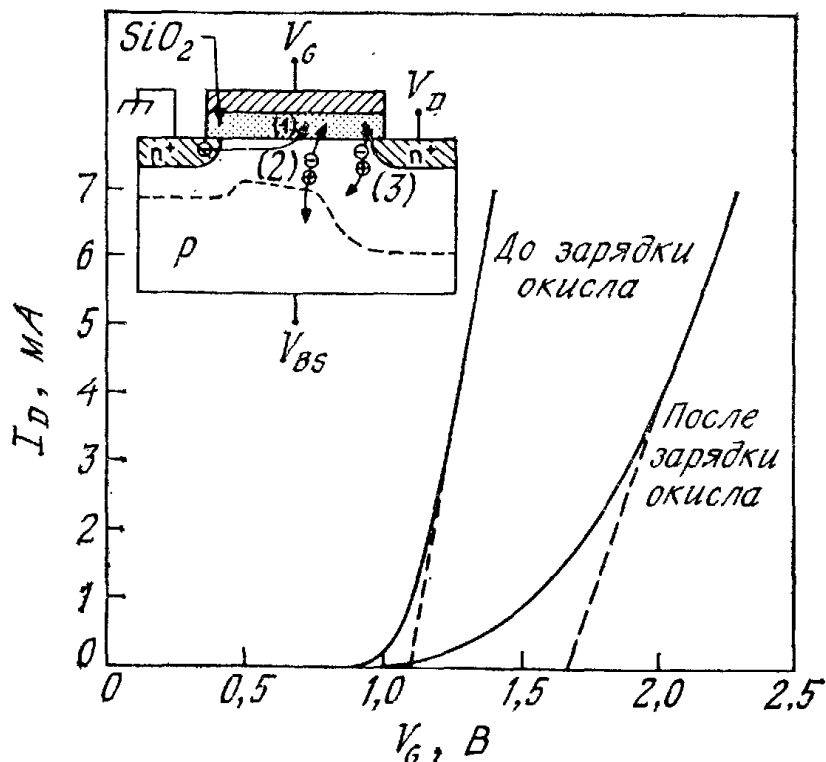


Рис. 49. Процессы, обуславливающие зарядку окисла и результирующее изменение характеристики МОП-транзистора [51].

(1) — ток в канале; (2) — горячие электроны; (3) — электронно-дырочная плазма лавинного пробоя.

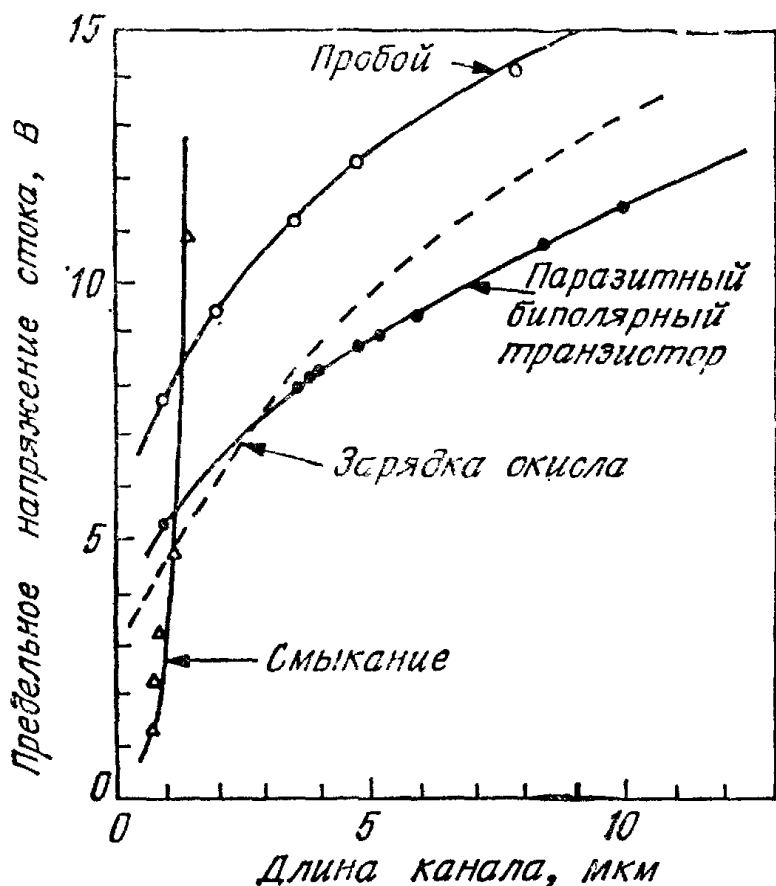


Рис. 50. Зависимость предельных напряжений стока от длины канала при различных эффектах [52].

сти в канале. Кроме того, возрастают подпороговые токи прибора вследствие увеличения плотности поверхностных состояний.

Зарядка окисла оказывает существенное влияние на стабильность работы МОП-транзистора и в конечном счете определяет срок его службы, поскольку этот процесс приводит к непрерывной деградации рабочих характеристик прибора. Для замедления деградации и тем самым увеличения срока службы прибора придется ограничивать максимальную амплитуду рабочих напряжений МОП-транзистора. Из графиков, приведенных на рис. 50, можно выявить влияние рассмотренных выше механизмов на предельные напряжения стока [52]. Эти зависимости получены для некоторых определенных значений параметров прибора (d , r_f , N_A). При ином выборе этих параметров относительная роль рассмотренных механизмов может, естественно, изменяться.

Для устранения эффекта паразитного транзистора следует уменьшать сопротивление подложки R_{sub} , так чтобы произведение тока подложки на это сопротивление не превышало напряжения 0,6 В, требуемого для открывания истока, когда напряжение на стоке равно или превышает соответствующее напряжение BV_{CEO} . При этом напряжение пробоя короткоканального МОП-транзистора уже не ограничивается величиной BV_{CEO} и прибор может работать при более высоких напряжениях, что обеспечивает большую надежность [50]. Чтобы уменьшить зарядку окисла, следует уменьшать плотность электронных ловушек в SiO_2 [53].

Для увеличения напряжения смыкания применяют одно- или даже двукратно ионно-импланированные структуры с повышенным легированием приповерхностной области подложки. Некоторые из них рассмотрены в разд. 8.5.

8.5. МОП-ТРАНЗИСТОРНЫЕ СТРУКТУРЫ

Для улучшения рабочих характеристик МОП-транзисторов (повышения быстродействия, уменьшения энергопотребления, повышения надежности и уровня переключаемых мощностей) в настоящее время предложено большое число разнообразных структурных решений. Ниже рассмотрены некоторые наиболее типичные и употребительные из этих структур.

8.5.1. Масштабное уменьшение размеров

Как уже говорилось выше, короткоканальные эффекты ухудшают рабочие характеристики МОП-транзисторов и поэтому нежелательны. Одним из способов устранения этих эффектов является пропорциональное уменьшение всех характерных размеров прибора. При этом следует во столько же раз уменьшить и характерные значения рабочих напряжений, с тем чтобы внешние электрические поля в приборе оставались на прежнем уровне. Такое масштабное уменьшение размеров представляет собой простейший подход к проблеме миниатюризации МОП-транзисторов.

На рис. 51, а показан обычный прибор, а на рис. 51, б — прибор, все характерные размеры которого (толщина окисла, длина и ширина канала, размеры n^+ — p -переходов) [54] уменьшены в k раз (k — масштабный фактор). Уровень легирования подложки во втором приборе в k раз больше, чем в обычном приборе, и, поскольку питающие напряжения также уменьшены в k раз, толщина соответствующих обедненных областей оказывается в k раз меньше. На рис. 51, в приведены сток-затворные характеристики обоих приборов. Отметим, что пороговое напряжение масштабно уменьшенной структуры также уменьшилось в k раз. Такое пропорциональное уменьшение характерных размеров отдельного прибора увеличивает плотность компоновки приборов на кристалле в k^2 раз, уменьшает характерные времена задержки (увеличивает быстродействие) в k раз и уменьшает мощность, рассеиваемую в отдельном приборе, в k^2 раз.

Обратим внимание на то, что область подпороговых токов на характеристиках обоих приборов (рис. 51, в) практически одинакова. Это неудивительно, поскольку при таком масштабном преобразовании характерное напряжение $S \sim (1 + C_D/C_i)$ остается неизменным (емкости C_D и C_i увеличились в k раз). Отметим также, что контактная разность переходов V_{ib} и характерный

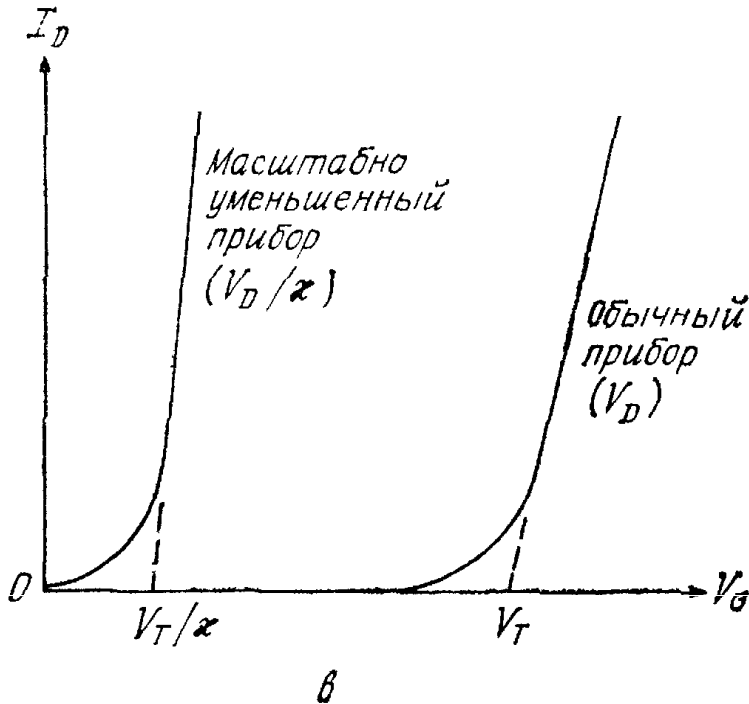
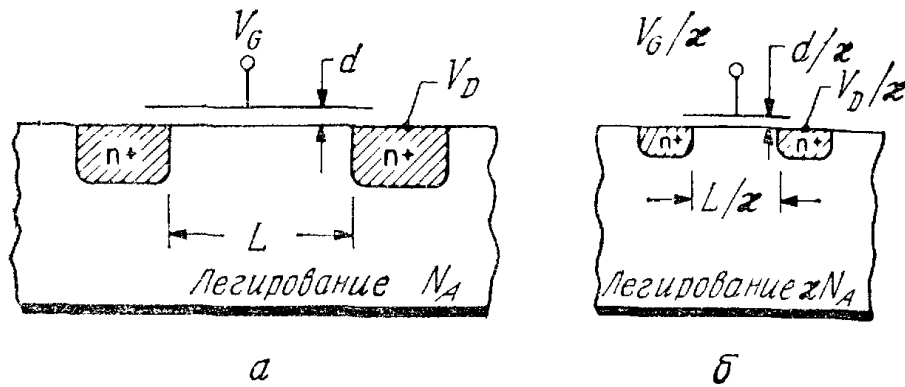


Рис. 51. Пропорциональное уменьшение размеров как принцип миниатюризации [54].
 а — длинноканальный прибор;
 б — прибор с пропорционально уменьшенными размерами; в — стоковые характеристики.

поверхностный потенциал ψ_B не уменьшаются при таком «масштабировании», а даже увеличиваются (на $\sim 10\%$ при увеличении уровня легирования в 10 раз). Поэтому не изменяется и характерная разность затворных напряжений ($\sim 0,5$ В), соответствующих началу сильной инверсии и режиму обеднения. Паразитные емкости при таком масштабном уменьшении размеров могут и не уменьшиться. Сопротивление шин разводки обычно увеличивается с уменьшением размеров.

Большую «гибкость» в решении вопросов миниатюризации допускает подход, основанный на использовании эмпирического соотношения (83) [41]. В этом случае при обеспечении заданного минимального размера L_{\min} можно в определенных пределах варьировать различные параметры структуры, поддерживая при этом неизменной требуемую величину синтетического параметра γ . Следовательно, при таком подходе не обязательно уменьшать все характерные размеры структуры в одинаковое число раз, а в определенных пределах можно даже варьировать геометрию структуры. Последнее можно использовать в целях упрощения технологии либо для оптимизации других аспектов функционирования прибора.

8.5.2. Высококачественные МОП-структуры (НМОС)

Так называют МОП-транзисторы, изготовленные с использованием ионной имплантации. Структура, показанная на рис. 52, а, содержит один имплантированный слой [55]. Он контролирует пороговое напряжение МОП-транзистора и в определенной мере препятствует смыканию обедненных областей стока и истока. Отметим, что имплантированный слой здесь довольно узкий, и обедненная область стока выходит за его пределы в низколегированную подложку. Это обеспечивает достаточно малую выходную емкость МОП-транзистора (емкость перехода сток — подложка). На рис. 52, б показана такая структура с двойной ионной имплантацией [56]. Параметры первого имплантированного слоя p_1 выбираются для обеспечения требуемого порогового напряжения МОП-транзистора, а характеристики второго слоя p_2 — для предотвращения смыкания. Использование двойной ионной имплантации позволяет существенно минимизировать

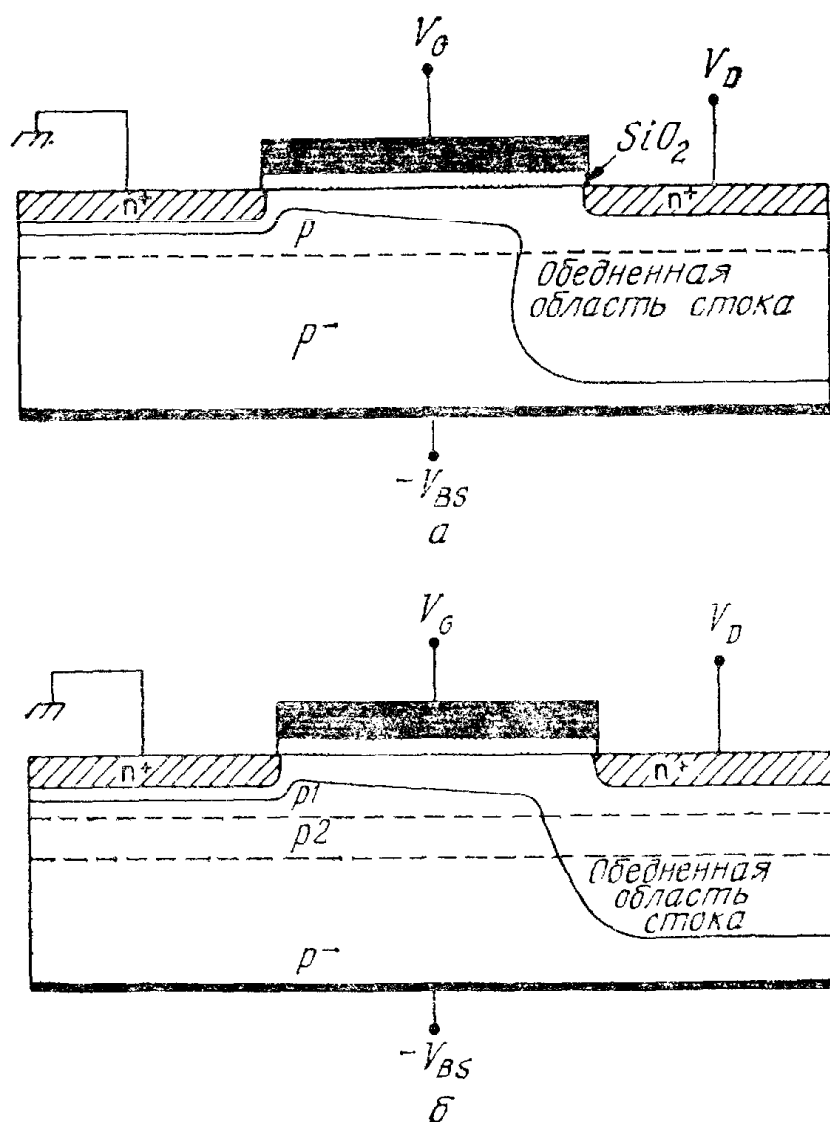


Рис. 52. Высококачественная МОП-структура (НМОС) [55, 56].

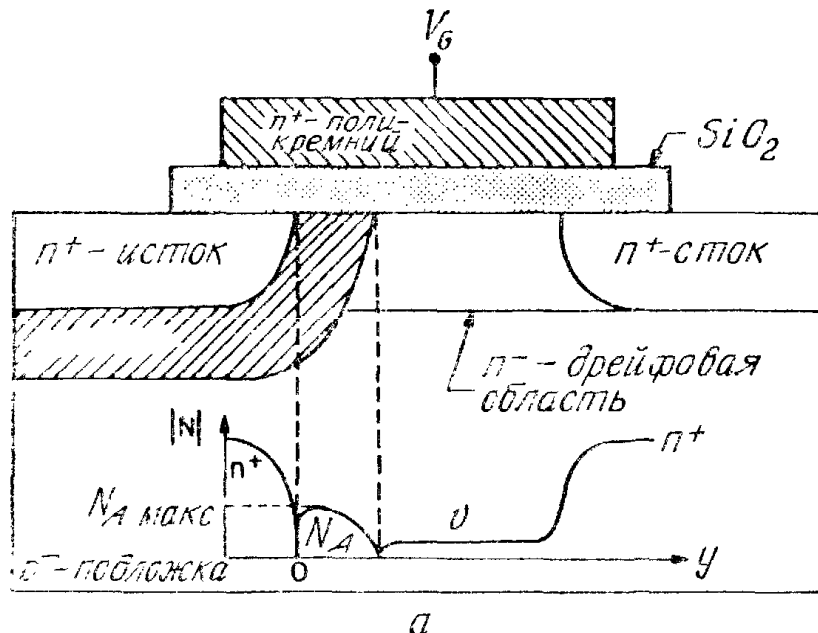
а — с одним имплантированным слоем; б — двукратно имплантированная структура.

короткоканальные эффекты в таких МОП-структурах с физически малыми длинами каналов.

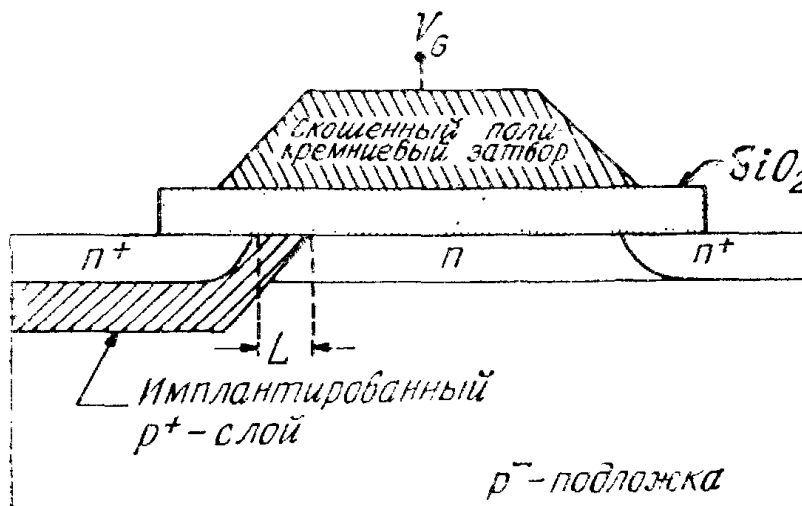
Имплантация, однако, ухудшает подпороговый участок характеристик [19] (увеличенные значения характерного напряжения S) и повышает чувствительность порога к смещению подложки V_{BS} . Эти нежелательные эффекты необходимо учитывать при оптимизации параметров структуры.

8.5.3. МОП-структуры с двойной диффузией (DMOS)

МОП-структура, схема которой приведена на рис. 53, а, изготовлена с использованием так называемой двойной диффузии [57]. Этот процесс, широко применяемый в биполярной технологии, основан на том, что скорость диффузии p -примеси (например, бора) в кремнии выше, чем для n -примеси (например, фосфора).



а



б

Рис. 53. МОП-структуры с двойной диффузией (а) и двукратной ионной имплантацией (б) [57, 58].

В рассматриваемом случае он позволяет получать структуры со сравнительно короткими каналами. На рис. 53, а также показан профиль легирования в такой структуре вдоль границы с окислом. Отметим, что сразу за каналом следует низколегированная дрейфовая область. В аналогичной структуре, изготовленной с использованием двукратной ионной имплантации [58] (рис. 53, б), в качестве маски используется поликремниевый затвор транзистора, скошенные края которого обеспечивают выход имплантированной p -области на границу с окислом.

Такие структуры могут иметь очень короткие каналы, и при их изготовлении используется обычный литографический процесс. Обе структуры (рис. 53) весьма устойчивы по отношению к смыканию из-за высокой концентрации примеси в p -слое. Сравнительно протяженная низколегированная дрейфовая область служит как бы нагрузкой, стабилизируя электрическое поле в канале на уровне $\approx 10^4$ В·см⁻¹, достаточном для достижения носителями скорости насыщения [59]. Такое ограничение максимальных электрических полей в структуре препятствует развитию нежелательных процессов лавинного умножения носителей и зарядки окисла. В этом отношении структуры с двойной диффузией имеют значительные преимущества перед МОП-транзисторами, изготовленными по стандартной и NMOS-технологии [11].

Однако в этих структурах труднее контролировать пороговое напряжение МОП-транзистора, которое (рис. 53, а) определяется максимальным значением концентрации $N_{A \text{ макс}}$ на границе с окислом [60]. Кроме того, сравнительно высокая концентрация примеси в узком p -слое, необходимая для предотвращения смыкания стока с истоком, ухудшает подпороговый участок характеристик, соответствующий закрытому состоянию МОП-транзисторов.

8.5.4. МОП-транзисторы с заглубленным каналом

Конструкция этого прибора [61], показанная на рис. 54 (вставка), служит примером использования концепции миниатюризации, основанной на эмпирическом соотношении (83): минимальная длина канала $L_{\text{мин}}$, при которой еще сохраняется «длинноканальность» электрических характеристик МОП-транзистора и уменьшается с уменьшением глубины переходов по закону $r_j^{1/3}$. Представленные на рис. 54 зависимости порогового напряжения от длины канала демонстрируют такую минимизацию короткоканальных эффектов с уменьшением r_j . Легко видеть, что рассматриваемая конструкция с заглубленным каналом позволяет снизить r_j до нуля и даже до отрицательных значений.

Очевидным недостатком структур с заглубленным каналом, особенно существенным для приборов с субмикронными разме-

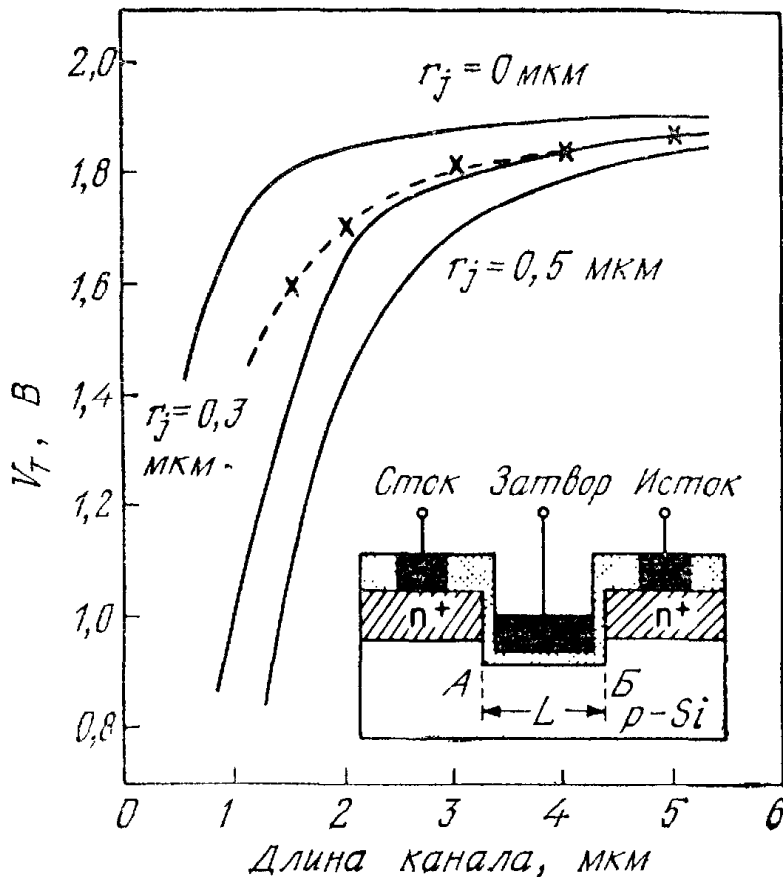


Рис. 54. Расчетные (—) и экспериментальные (---x---) зависимости порогового напряжения от длины канала при различных значениях глубины перехода. На вставке показан МОП-транзистор в заглубленном канале [61].

рами, являются технологические трудности, возникающие при формировании требуемой «канавки». Кроме того, в такой геометрии плохо контролируется толщина окисла в «углах» А и Б, которая определяет пороговое напряжение транзистора. Отметим также, что в структурах с заглубленным каналом можно ожидать повышенную скорость зарядки окисла в результате более интенсивной инжекции горячих электронов.

8.5.5. Барьеры Шоттки в качестве стока и истока

Использование барьеров Шоттки в качестве стока и истока МОП-транзисторов имеет ряд достоинств с точки зрения рабочих характеристик приборов и технологии их изготовления. Схематически такая МОП-структура показана на рис. 55, а [62]. Барьеры Шоттки имеют сравнительно малую толщину перехода, что весьма благоприятно для минимизации короткоканальных эффектов. Высокая проводимость контактов Шоттки также обеспечивает достаточно малую величину эффективного последовательного сопротивления истока. С точки зрения технологии исключение ряда высокотемпературных операций способствует повышению качества окисного слоя и облегчает воспроизведение геометрии структуры. Кроме того, МОП-транзисторы с барьерами Шоттки могут быть изготовлены на полупроводниках (таких, как, например, CdS), в которых не удастся получить достаточно качественных $p-n$ -переходов.

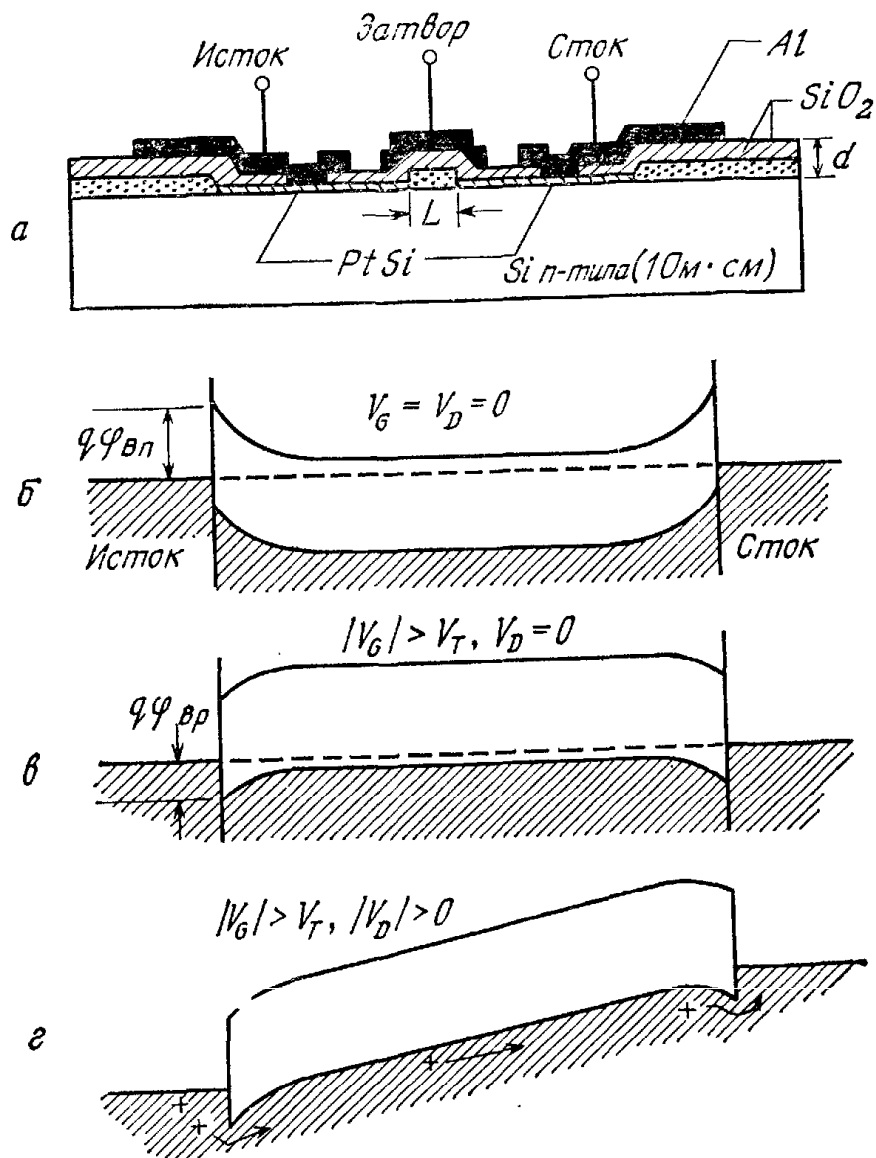


Рис. 55. МОП-транзистор с барьерами Шоттки в качестве стока и истока [62].
 а — схема поперечного разреза прибора; б, в и г — зонные диаграммы при различных комбинациях напряжений на электродах.

На рис. 55, б приведена зонная диаграмма структуры в условиях термодинамического равновесия ($V_G = V_D = 0$). Здесь $q\varphi_{Вп}$ — высота энергетического барьера между металлом и n -подложкой (например, для контакта PtSi—Si $q\varphi_{Вп} = 0,85$ эВ). Когда напряжение на затворе достаточно велико, чтобы инвертировать приповерхностную область полупроводника от n -типа к p -типу (рис. 55, в), энергетический барьер между истоком и p -инверсионным слоем равен уже $q\varphi_{Вр} = 0,25$ эВ (для того же контакта PtSi—Si). Термоионная эмиссия через такой малый барьер поставляет достаточное число дырок в инверсионный слой структуры, где они увлекаются электрическим полем к контакту Шоттки стока (рис. 55, г). Отметим, что в этом рабочем режиме прибора барьер Шоттки истока оказывается обратномещенным. При комнатных температурах плотность термоионного тока через барьер высотой 0,25 эВ составляет $\sim 10^3$ А·см⁻². При необходи-

мости ее увеличения следует использовать контакты Шоттки с большей высотой энергетического барьера по отношению к n -кремнию. Изготовление аналогичных МОП-транзисторов на p -подложке (с n -каналом) в настоящее время представляет большие трудности, поскольку еще не найдены контакты Шоттки с достаточно большой высотой энергетического барьера по отношению к p -кремнию.

8.5.6. Тонкопленочный транзистор

Тонкопленочный транзистор (ТПТ) схематически изображен на рис. 56 (вставка). Его структуру составляют последовательно нанесенные на изолирующую подложку слои полупроводника (например, CdS), диэлектрика и металла [63]. Вольт-амперные характеристики ТПТ обогащенного типа (рис. 56) подобны соответствующим характеристикам обычных МОП-транзисторов. Поскольку полупроводниковая пленка в ТПТ обычно наносится простым распылением, в ней содержится значительно большее число дефектов и нарушений кристаллической решетки, чем в соответствующих монокристаллических образцах. Поэтому транспортные процессы в таких пленках более сложны, чем в моно-

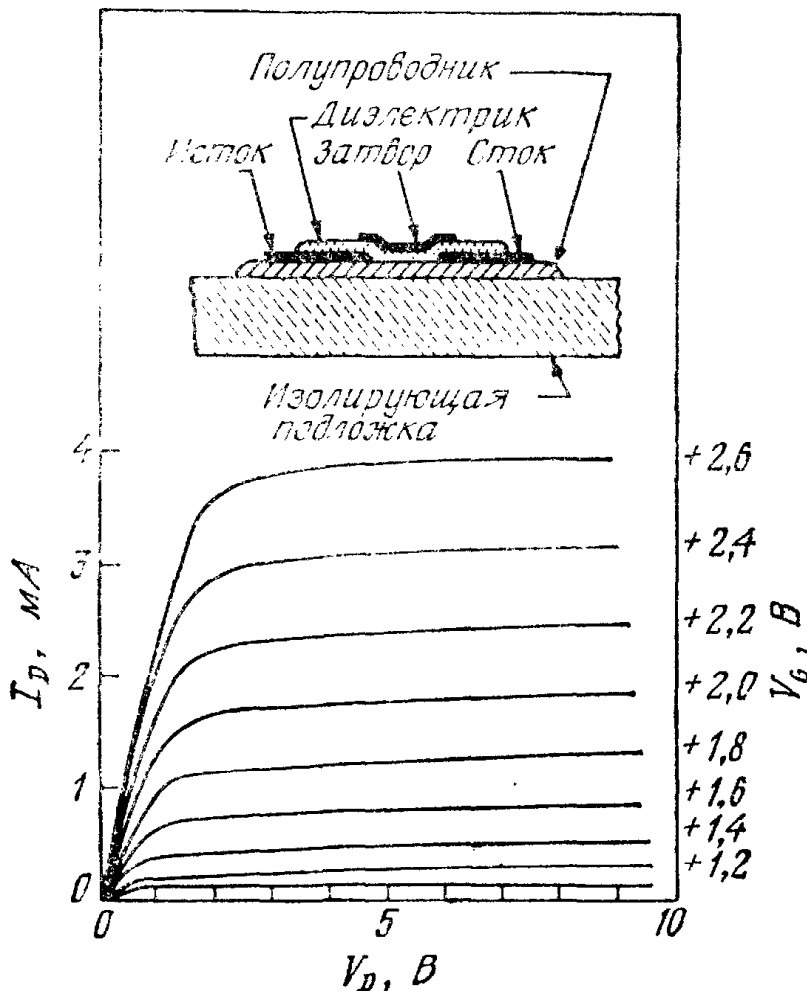


Рис. 56. Тонкопленочный транзистор и его стоковые характеристики [63].

кристаллах. Для улучшения приборных характеристик ТПТ, надежности и воспроизводимости следует снижать плотность объемных и поверхностных ловушек в полупроводниковой пленке.

8.5.7. Структуры типа «кремний на изоляторе» (КНИ)

К настоящему времени предложено довольно большое число структур типа «кремний на изоляторе» (КНИ): кремний на сапфире (КНС), кремний на шпинели, кремний на нитриде и кремний на окисле. КНИ-приборы формируются по стандартной МОП-технологии на монокристаллической кремниевой пленке, эпитаксиально выращенной на изолирующей подложке (например, на Al_2O_3 в случае КНС) [64]. Схема такой структуры приведена на рис. 57, а. Диэлектрическая подложка обеспечивает электрическую изоляцию соседних приборов. КНИ-структуры имеют низкие паразитные емкости, что обеспечивает высокое быстродействие этих приборов. Типичные характеристики n -канального КНС-прибора показаны на рис. 57, б. Перегибы в области насыщения характеристик можно объяснить следующим образом. Когда напряжение стока увеличивается, в области вблизи стокового контакта начинается слабый процесс лавинного умножения. Генерированные при этом электроны сразу же поступают в сток, а генерированные дырки двигаются к истоку, составляя ток подложки. Однако, поскольку КНС-прибор не имеет контакта к подложке, эти дырки накапливаются у границы истока, смещая его в прямом направлении. При этом прибор фактически как бы находится в режиме с положительно смещенной подложкой $V_{BS} > 0$, что эквивалентно уменьшению порогового напряжения. Этим объясняются наблюдаемые скачки тока. Амплитуда скачков порогового напряжения сильно зависит от степени легирования под-

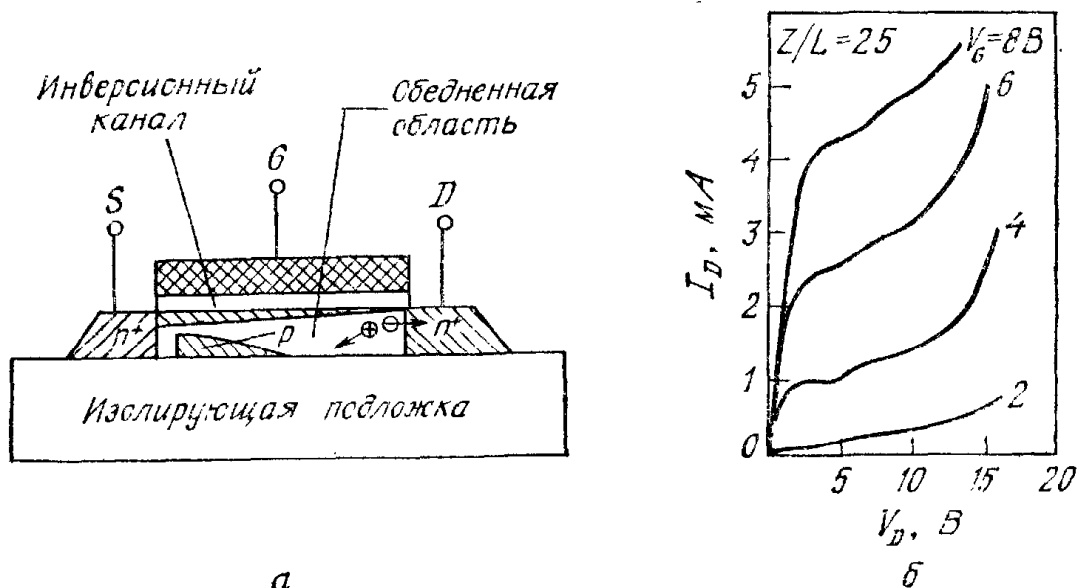


Рис. 57. МОП-транзистор типа «кремний на изоляторе» (а) и стоковые характеристики КНС-прибора (б) [64].

ложки: чем больше концентрация, тем больше скачок V_T и тем заметнее перегибы вольт-амперных характеристик.

В работах [65, 66] для изготовления КНИ-приборов предложено использовать рекристаллизацию лазерным отжигом пленок аморфного кремния на изолирующих подложках, таких, как Si_3N_4 и SiO_2 . В настоящее время рабочие характеристики КНИ-структур значительно уступают соответствующим характеристикам приборов на монокристаллических подложках. Это обусловлено высокой плотностью дефектов в объеме таких эпитаксиальных кремниевых пленок и большой плотностью поверхностных состояний на границах раздела кремний—изолятор. При улучшении качества пленок КНИ-приборы нашли бы широкое применение в сверхбыстродействующих схемах. Кроме того, на основе КНИ-структур можно было бы попытаться освоить трехмерный уровень интеграции микросхем, «накладывая» друг на друга соответствующие плоские базовые КНИ-структуры [80].

8.5.8. МОП-транзистор с V-канавкой

Структура МОП-транзистора с V-канавкой [67] схематически показана на рис. 58, а, а структура МОП-транзистора с U-канавкой — на рис. 58, б [68]. Эти приборы изготавливаются на кремниевых подложках с ориентацией $\langle 100 \rangle$ на основе эффекта анизотропного травления. Вытравливаемые V-канавки имеют характерный угол наклона на $54,7^\circ$ по отношению к горизонтальной плоскости. Как видно из рис. 58, а, распределение легирующих примесей здесь аналогично распределению примесей в МОП-транзисторах с двойной диффузией. В этих приборах два параллельных друг другу канала длиной L расположены на обеих сторонах вытравленной канавки. Общим стоком служит подложка. В такой конфигурации легко соединить параллельно требуемое число отдельных приборов, что позволяет создавать таким способом мощные переключатели, выдерживающие большие токи [69]. В тех ситуациях, когда требуется планарная организация схемы, т. е. нужно, чтобы все контакты выходили на внешнюю поверхность кристалла, можно исключить из структуры n^+ -область и увеличить степень легирования p -слоя (заменить p на p^+). Левая p^+ -область тогда будет стоком, правая — истоком (или наоборот), а собственно канал будет расположен по границе ABC .

8.5.9. Гексагональный МОП-транзистор

На рис. 59, а показана полупроводниковая структура, представляющая собой плотно упакованный массив так называемых гексагональных МОП-транзисторов [70]. Отдельная ячейка структуры представляет собой МОП-транзистор с двойной диффузией

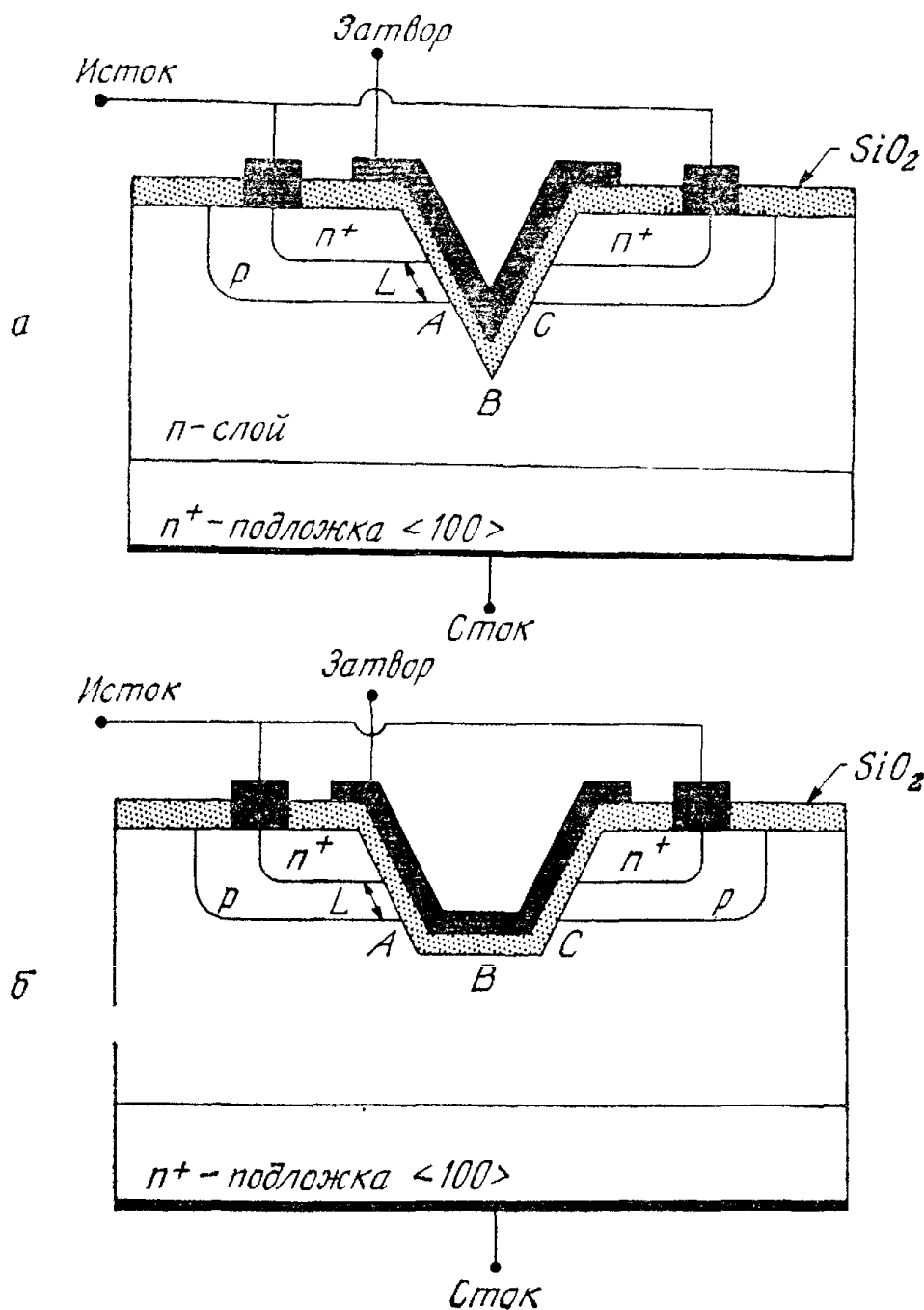


Рис. 58. МОП-транзистор с V-образной канавкой (а) и U-образной канавкой (б) [67, 68].

и шестиугольным истоком. Узкий диффузионный канал, расположенный по его периферии, перекрыт поликремниевым затвором. Общим стоком служит подложка структуры. Гексагональная структура имеет высокую плотность размещения элементов на кристалле ($\sim 10^5 \text{ см}^{-2}$). Большие значения отношения ширины к длине канала ($6ZN/L$) обеспечивают высокую крутизну и малое выходное сопротивление во включенном состоянии.

На рис. 59, б приведены выходные характеристики мощной гексагональной структуры, рассчитанной на переключение токов до 100 А. Напряжения, соответствующие участку насыщения характеристик, имеют те же значения, что и в аналогичных биполярных приборах таких же размеров. Выходное сопротивление

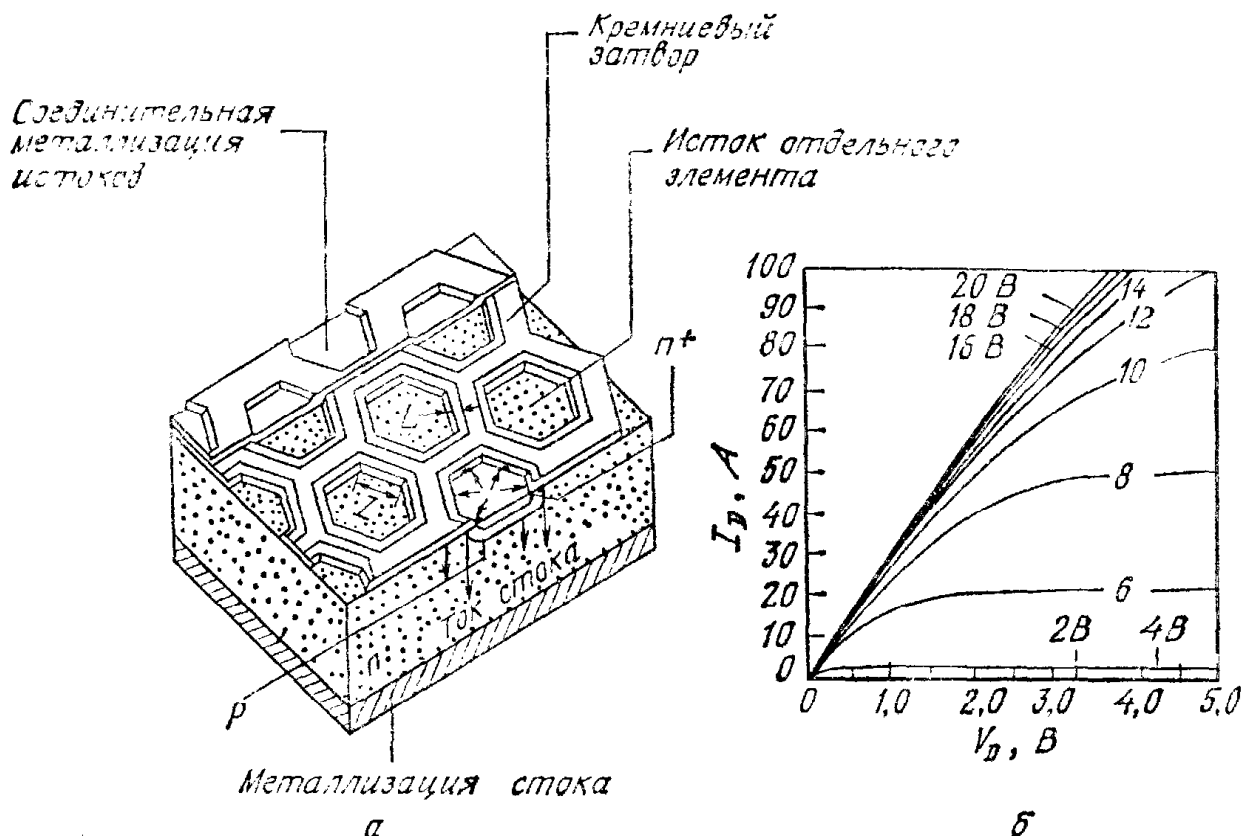


Рис. 59. Гексагональный МОП-транзистор (а) и его выходные характеристики (б) [70].

во включенном состоянии составляет всего лишь 0,05 Ом. При необходимости можно получить меньшие выходные сопротивления, увеличив отношение $6NZ/L$.

8.6. ЭНЕРГОНЕЗАВИСИМЫЕ ЭЛЕМЕНТЫ ПАМЯТИ

Если модифицировать затвор обычного МОП-транзистора таким образом, чтобы в подзатворном диэлектрике мог храниться электрический заряд, то мы получим новый прибор — энергонезависимый элемент памяти. С тех пор как в 1967 г. был предложен первый такой прибор [71], энергонезависимые элементы памяти получили широкое развитие и распространение. На их основе в настоящее время разрабатывается и изготавливается широкий класс полупроводниковых интегральных запоминающих устройств (ЗУ), таких, как электрически репрограммируемые постоянные ЗУ (ПЗУ с электрической перезаписью), программируемые ПЗУ со стиранием, энергонезависимые ЗУ с произвольной выборкой и т. п.

Энергонезависимые элементы памяти можно разделить на два класса: приборы с плавающим затвором и структуры с двухслойным диэлектриком — МДОП (металл—диэлектрик—окисел—полупроводник). «Запись» заряда в них осуществляется инжекцией носителей из кремния через слой окисла. Заряд хранится

либо в плавающем затворе, либо на границе окисла со вторым диэлектриком (в МДОП-структурах). Этот заряд, изменяя пороговое напряжение МОП-транзистора, «сдвигает» прибор в состояние с высоким порогом. В хорошо сконструированной ячейке такое состояние сохраняется более 100 лет. «Стирание» хранящегося заряда (возврат прибора в исходное состояние с низким порогом) осуществляется либо электрически (с помощью соответствующих импульсов напряжения на затворе), либо некоторыми другими способами, например ультрафиолетовым облучением структуры.

8.6.1. Приборы с плавающим затвором

В структуре с плавающим затвором (рис. 60) на первом, тонком слое окисла $I(1)$ расположен металлический электрод $M(1)$ — плавающий затвор, отделенный вторым, толстым слоем диэлектрика $I(2)$ от внешнего металлического затвора $M(2)$. Зонные диаграммы структуры с плавающим затвором, соответствующие режимам записи, хранения и стирания, приведены на рис. 61.

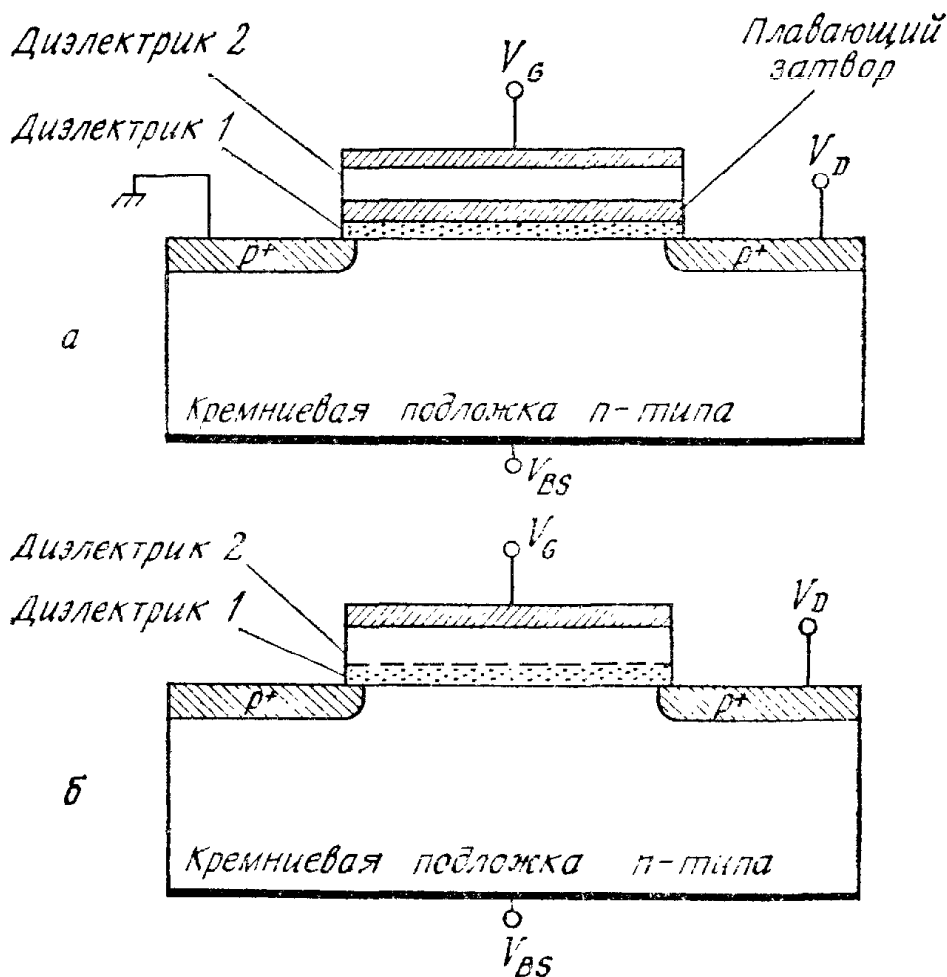


Рис. 60. Энергонезависимый элемент памяти с плавающим затвором (а) и энергонезависимый элемент памяти типа МДОП (б).

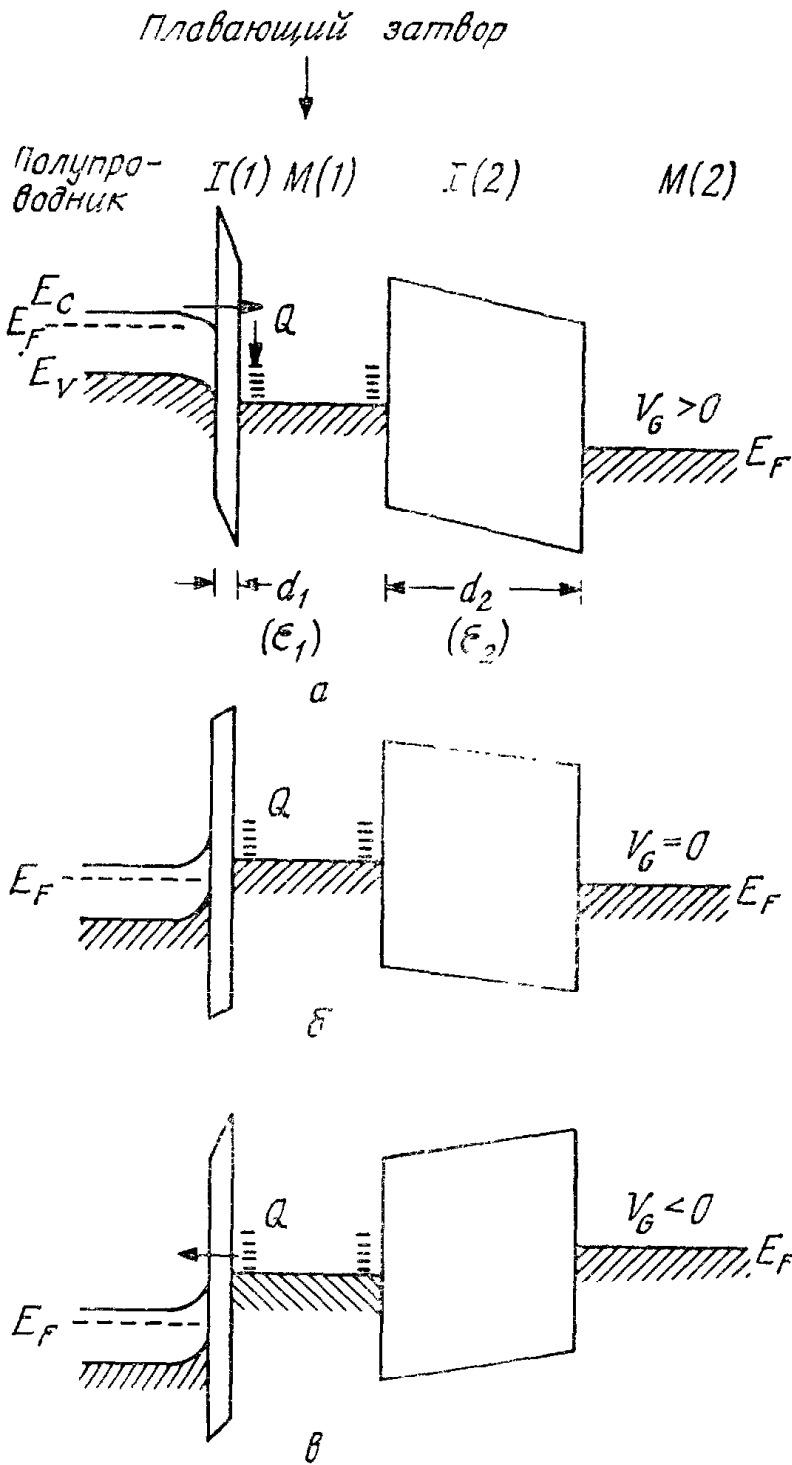


Рис. 61. Зонные диаграммы элемента памяти с плавающим затвором [71].

a — зарядка (операция записи);
б — хранение заряда; *в* — разрядка плавающего затвора (операция стирания).

Напряжение V_G , приложенное к внешнему затвору структуры $M(2)$, делится между обоими диэлектрическими слоями (рис. 61, *a*):

$$V_G = V_1 + V_2 = d_1 \mathcal{E}_1 + d_2 \mathcal{E}_2, \quad (105)$$

где \mathcal{E}_1 и \mathcal{E}_2 — напряженность электрического поля в этих слоях. Поскольку, согласно закону Гаусса,

$$\varepsilon_1 \mathcal{E}_1 = \varepsilon_2 \mathcal{E}_2 + Q, \quad (106)$$

электрическое поле в каждом из слоев (например, в первом) можно записать в виде

$$\mathcal{E}_1 = \frac{V_G}{d_1 + d_2(\varepsilon_1/\varepsilon_2)} + \frac{Q}{\varepsilon_1 + \varepsilon_2(d_1/d_2)}, \quad (107)$$

где Q — заряд в плавающем затворе, а ϵ_1 и ϵ_2 — диэлектрические проницаемости первого и второго диэлектрических слоев.

Если токи в обоих диэлектрических слоях не равны друг другу, заряд плавающего затвора Q будет изменяться во времени:

$$Q(t) = \int_0^t [J_1(\mathcal{E}_1) - J_2(\mathcal{E}_2)] dt \quad [\text{Кл} \cdot \text{см}^{-2}], \quad (108)$$

где $J_1(\mathcal{E}_1)$ и $J_2(\mathcal{E}_2)$ — плотности тока в диэлектриках 1 и 2.

Обычно токи в диэлектриках сильно зависят от напряженности электрического поля. Так, например, при туннелировании по механизму Фаулера—Нордгейма

$$J = C_1 \mathcal{E}^2 \exp(-\mathcal{E}_0/\mathcal{E}), \quad (109)$$

где \mathcal{E} — электрическое поле, а C_1 и \mathcal{E}_0 — константы, зависящие от эффективной массы носителей и высоты барьера. Как уже говорилось в гл. 7, данный механизм проводимости играет основную роль в тонких слоях SiO_2 и Al_2O_3 . Когда транспорт заряда в диэлектрике осуществляется по механизму Пула—Френкеля, как, например, в Si_3N_4 , тогда

$$J = C_2 \mathcal{E} \exp[-q(\varphi_B - \sqrt{q\mathcal{E}/\pi\epsilon_i})/kT], \quad (110)$$

где C_2 — константа, пропорциональная плотности ловушек в диэлектрике; φ_B — глубина ловушки, а ϵ_i — динамическая диэлектрическая проницаемость.

В результате разбаланса токов в диэлектриках 1 и 2 за время действия большого положительного напряжения V_G , приложенного к внешнему затвору (импульса записи), в плавающем затворе накапливается заряд Q (рис. 61, б), который сдвигает пороговое напряжение структуры на величину

$$\Delta V_T = - \frac{d_2}{\epsilon_2} Q. \quad (111)$$

Стирание хранящегося заряда осуществляется импульсом противоположной полярности $V_G < 0$ (рис. 61, в).

На рис. 62 приведены результаты теоретического расчета динамики записи заряда по формулам (105)—(109) для следующих значений параметров: $d_1 = 50 \text{ \AA}$, $\epsilon_1 = 3,85\epsilon_0$ (для SiO_2); $d_2 = 1000 \text{ \AA}$, $\epsilon_2 = 30\epsilon_0$ (для ZrO_2), при напряжении $V_G = 50 \text{ В}$ и в предположении $J_2 = 0$. Как видно из рис. 62, в начальной стадии записи заряд Q линейно увеличивается со временем, но затем эта зависимость насыщается. В свою очередь ток инжекции сначала слабо меняется, а затем начинает резко уменьшаться. Такое поведение можно объяснить следующим образом. В начале записи при $t = 0$ заряд Q также равен нулю, а электрическое поле в слое 1 максимально ($\mathcal{E}_1 = V_G/[d_1 + (\epsilon_1/\epsilon_2)d_2]$). Пока Q

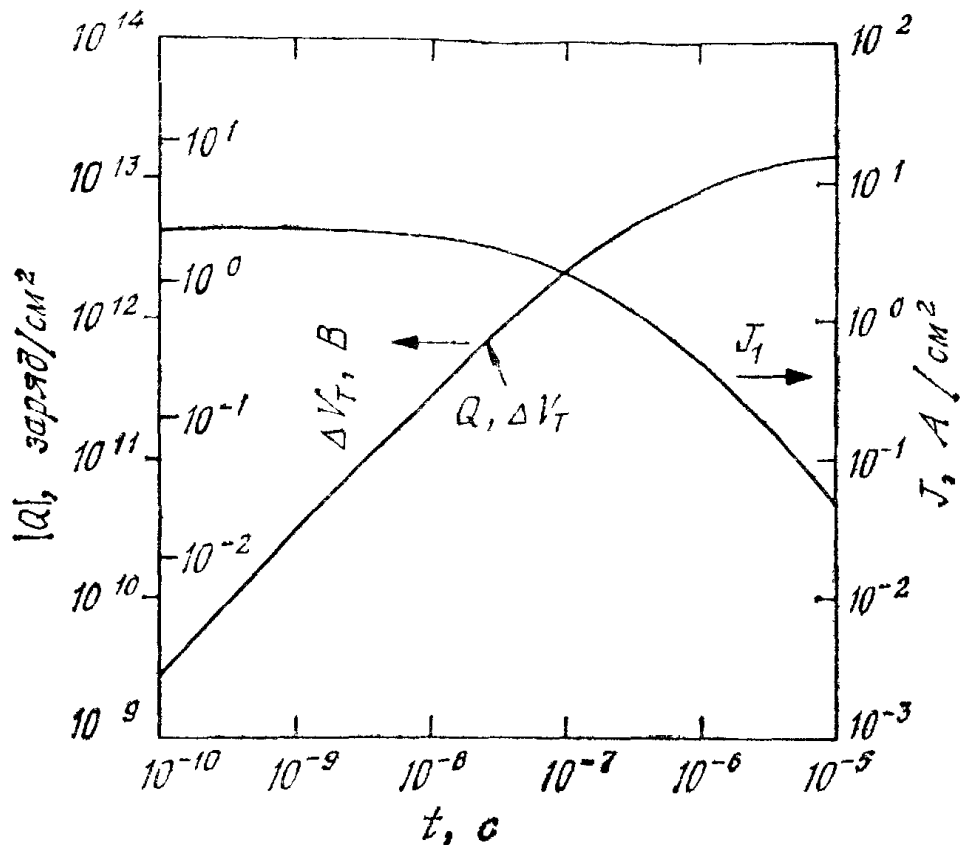


Рис. 62. Расчетные зависимости тока зарядки и накопленного заряда от времени [71].

мало и соответствующее уменьшение \mathcal{E}_1 незначительно, ток J_1 уменьшается слабо. Поэтому Q в начале процесса увеличивается линейно во времени. Однако в дальнейшем увеличение Q становится существенным и приводит к уменьшению поля \mathcal{E}_1 , в результате чего ток J_1 начинает заметно снижаться, а его интеграл (заряд Q) насыщаться.

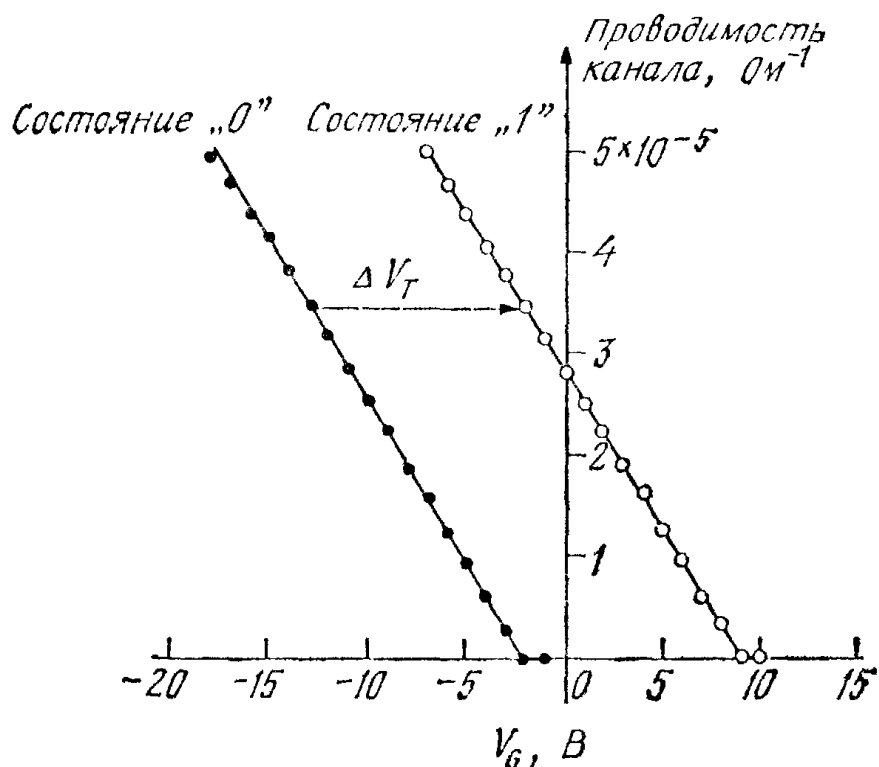
На рис. 62 приведен также соответствующий заряду Q сдвиг порогового напряжения ΔV_T (выражение (111)). В реальных приборах время зарядки (время записи), необходимое для сдвига порогового напряжения на 1 В, не превышает 10^{-7} с. Экспериментально сдвиг порогового напряжения можно определить, измеряя проводимость канала g_D . При малых напряжениях на стоке (для p -канального транзистора)

$$g_D = -\frac{Z}{L} \mu_p C_i (V_G - V_T), \quad V_G > V_T. \quad (112)$$

После изменений заряда на плавающем затворе Q (Q — электронный заряд и поэтому отрицателен) график зависимости $g_D(V_G)$ сдвигается вправо на величину ΔV_T (рис. 63) [73].

Мы подробно проанализировали работу предельно упрощенной запоминающей ячейки (рис. 60, а), чтобы проиллюстрировать основные принципы функционирования реальных энергонезависимых элементов памяти, о которых речь пойдет ниже: среди них

Рис. 63. Проводимость канала до (состояние «0») и после (состояние «1») выполнения операции записи [73].



принцип туннельной инжекции, используемый в энергонезависимых запоминающих ячейках типа МДОП, и принцип хранения заряда в плавающем затворе, на котором основана работа реальных ЗУ с плавающим поликремниевым затвором.

В первых программируемых ПЗУ со стиранием в качестве материала плавающего затвора использовался сильнолегированный поликремний (рис. 64, а). В этих приборах, получивших название МОП-ЗУ с плавающим затвором и лавинной инжекцией [74], поликремниевый затвор со всех сторон окружен слоем окисла. Довольно большая толщина слоя окисла ($d_1 \approx 1000 \text{ \AA}$) выбрана для того, чтобы исключить в окисле слабые места и коротки, через которые поликремниевый плавающий затвор мог бы быть электрически связан с подложкой. Инжекция заряда осуществляется горячими электронами, генерированными в области лавинного пробоя у стока (рис. 64, а), для чего сток смещается до напряжений, требуемых для пробоя. Горячие электроны при этом втягиваются в плавающий затвор электрическим полем, возникающим вследствие его емкостной связи с электродами стока и истока. Для стирания информационного заряда в такой структуре используется ультрафиолетовое или рентгеновское облучение. Электрическое стирание здесь невозможно, поскольку в структуре отсутствует внешний затвор.

При необходимости электрического стирания используется аналогичная ячейка памяти с двумя поликремниевыми затворами (рис. 64, б) [75]. Внешний затвор, кроме того, улучшает эффективность записи. Иногда в приборах этого типа изоляция между поликремниевыми затворами осуществляется слоем SiO_2 , обога-

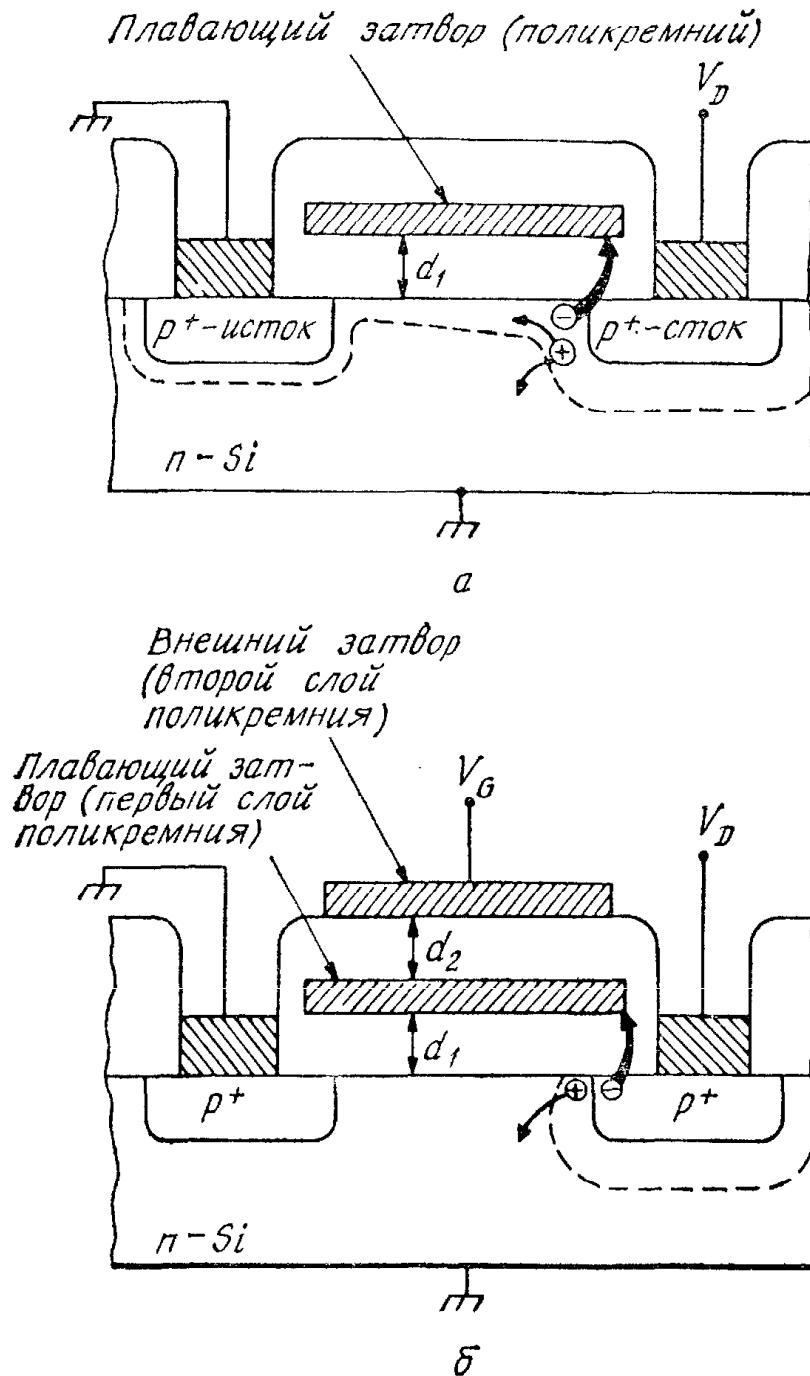
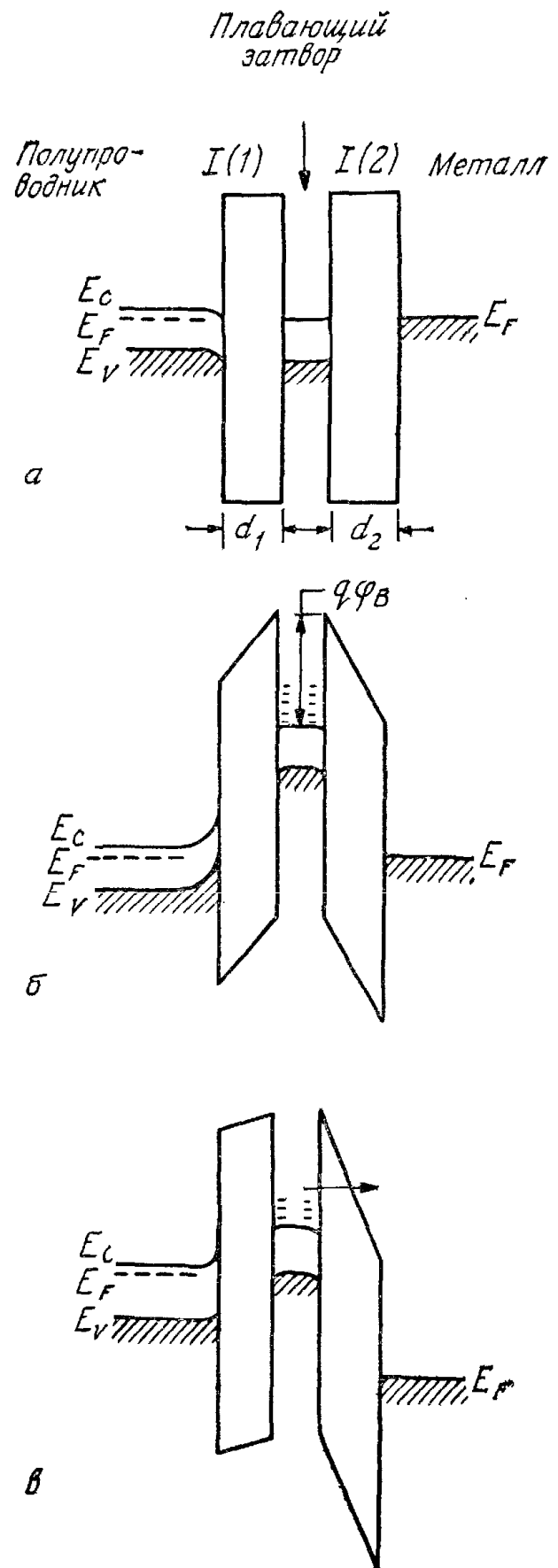


Рис. 64. Ячейка МОП-ЗУ с плавающим затвором и лавинной инжекцией (а) [74] и аналогичная ячейка с двумя затворами, обеспечивающая возможность электрического стирания (б) [75].

щенным кремнием [81]. На рис. 65 приведены зонные диаграммы, поясняющие работу такой ячейки с двумя затворами. Равновесное состояние структуры с $Q = 0$ (рис. 65, а) соответствует хранению информационного нуля «0». Состояние с накопленным в процессе лавинной инжекции зарядом на плавающем затворе (рис. 65, б) соответствует хранению информационной единицы «1». При больших положительных напряжениях на внешнем затворе накопленный заряд удаляется путем его вытягивания во внешний затвор (рис. 65, в), и элемент памяти возвращается в исходное состояние «0».

Рис. 65. Зонные диаграммы двухзатворного элемента памяти в состоянии «0» (а), состоянии «1» (б) и в режиме разрядки (стирания) (в).



На рис. 66 приведены теоретические кривые и соответствующие экспериментальные результаты, показывающие влияние напряжения на внешнем затворе на эффективность записи в такой структуре [75]. При фиксированном напряжении на стоке V_D напряжение на внешнем затворе V_G сначала способствует записи

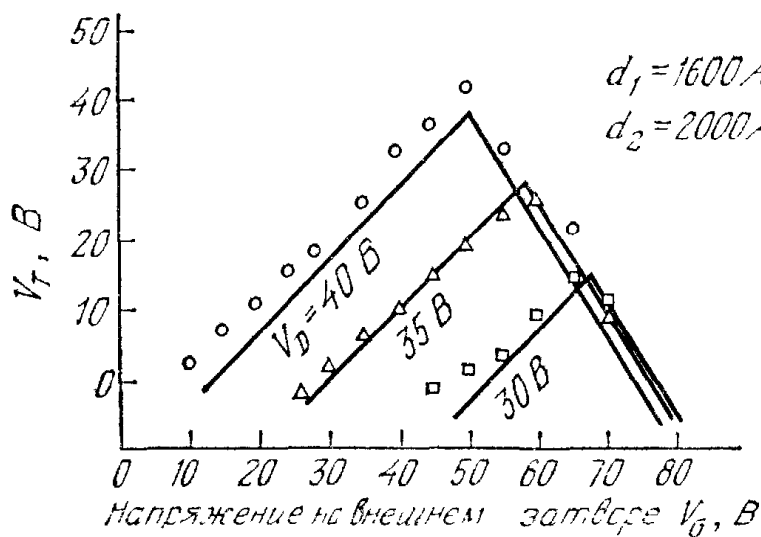


Рис. 66. Зависимость сдвига порогового напряжения при записи в двухзатворном элементе памяти от напряжения на внешнем затворе [75].

— теоретические данные;
○, △, □ — экспериментальные данные.

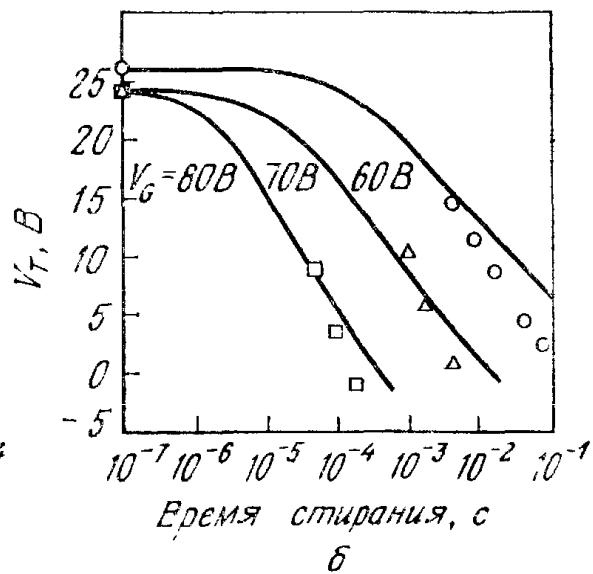
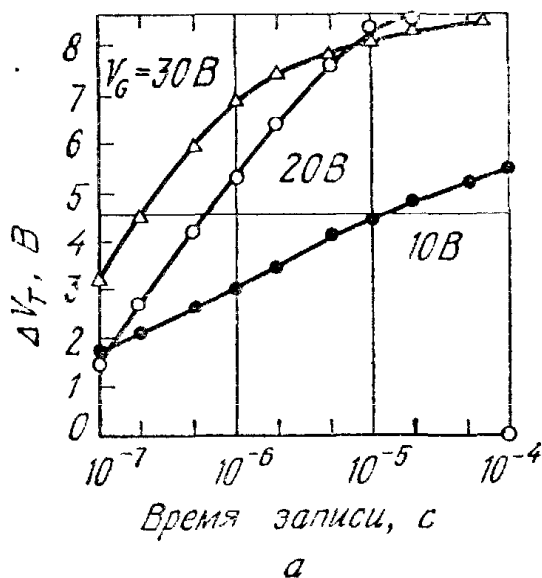


Рис. 67. Зависимость сдвига порогового напряжения от длительности импульсов записи и стирания в двухзатворном элементе памяти [75].

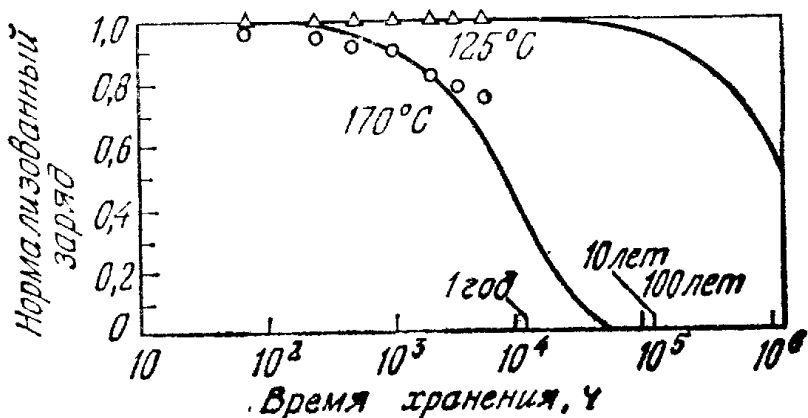
— теоретические данные; ○, △, □ экспериментальные данные.

(V_T увеличивается), усиливая дрейф горячих электронов через первый слой SiO_2 на плавающий затвор. Однако при достаточно больших V_G начинает преобладать конкурирующий процесс полевого вытягивания заряда из плавающего затвора во внешний. Поэтому с определенного значения V_G эффективность записи падает (V_T уменьшается). На рис. 67 приведены зависимости изменения порогового напряжения от времени записи и стирания для той же, что и на рис. 66, структуры. Отметим, что при $V_G = 20 \text{ В}$ для сдвига порога на 5 В требуется время $\sim 10^{-6} \text{ с}$, в то время как в аналогичных структурах без внешнего затвора (рис. 64, а) для этого необходимы времена, на два порядка большие. Кроме того, скорость операции стирания, естественно, увеличивается с ростом напряжения затвора V_G .

Важной характеристикой энергонезависимых элементов памяти является время хранения информации. Оно определяется

Рис. 68. Кривые стекания хранящегося заряда для двух температур [72].

— теоретические данные;
 Δ, ○ экспериментальные данные.



как время, за которое записанный заряд уменьшается в результате стекания на 50 %:

$$t_R = \ln 2 / [v \exp(-q\phi_B/kT)]. \quad (113)$$

Здесь v — частота диэлектрической релаксации, а $q\phi_B$ — высота энергетического барьера (рис. 65, б). Кривые стекания заряда, рассчитанные при $q\phi_B = 1,7$ эВ для температуры 125 и 170 °С, приведены на рис. 68 вместе с соответствующими экспериментальными данными. Как следует из рис. 68, время хранения информации при $T = 125$ и 170 °С составляет примерно 100 лет и 8000 ч соответственно.

8.6.2. МДОП-структуры

Среди различных запоминающих устройств со структурой металл—диэлектрик— SiO_2 — Si (МДОП) наиболее популярны МНОП-структуры (металл— Si_3N_4 — SiO_2 — Si). Кроме того, известны МДОП-структуры, в которых вместо нитрида кремния используются другие диэлектрические пленки, такие, как окись алюминия, окись тантала и окись титана. К этому типу приборов относится также структура, содержащая только один слой двуокиси кремния, во внешнюю часть которого имплантированы ионы золота (или других металлов) [76]. Такая имплантация изменяет проводимость в соответствующей части окисла, а также приводит к возникновению пограничных центров, способных захватывать и хранить заряд.

На рис. 69 приведены зонные диаграммы МНОП-структуры, соответствующие операциям записи и стирания информационного заряда [77]. Ток в окисле J_0 обусловлен туннелированием Фаулера—Нордгейма, в то время как ток через нитрид J_N соответствует эмиссии Пула—Френкеля. Уравнения, описывающие процесс накопления заряда в МНОП-структурах, близки к системе уравнений (105)—(108). Различие состоит в том, что электронные ловушки, захватывающие заряд в нитриде у границы с окислом, вероятно, непрерывно распределены по энергии связи, а также

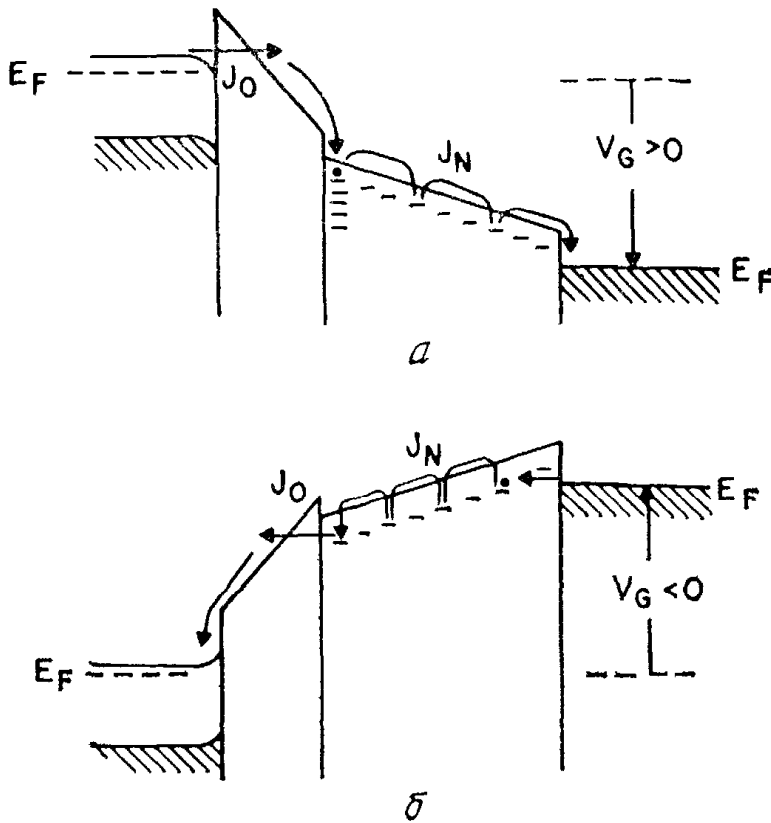


Рис. 69. Операции записи (а) и стирания (б) в МНОП-структуре [77].

по толщине приграничной области нитрида кремния. Этот факт тем или иным способом учитывается в различных теоретических моделях, предложенных в настоящее время для объяснения экспериментальных характеристик записи в МНОП-структурах. Данные, приведенные на рис. 70, представляют собой пример такого сопоставления, демонстрирующий неплохое соответствие теоретических зависимостей и результатов эксперимента [78].

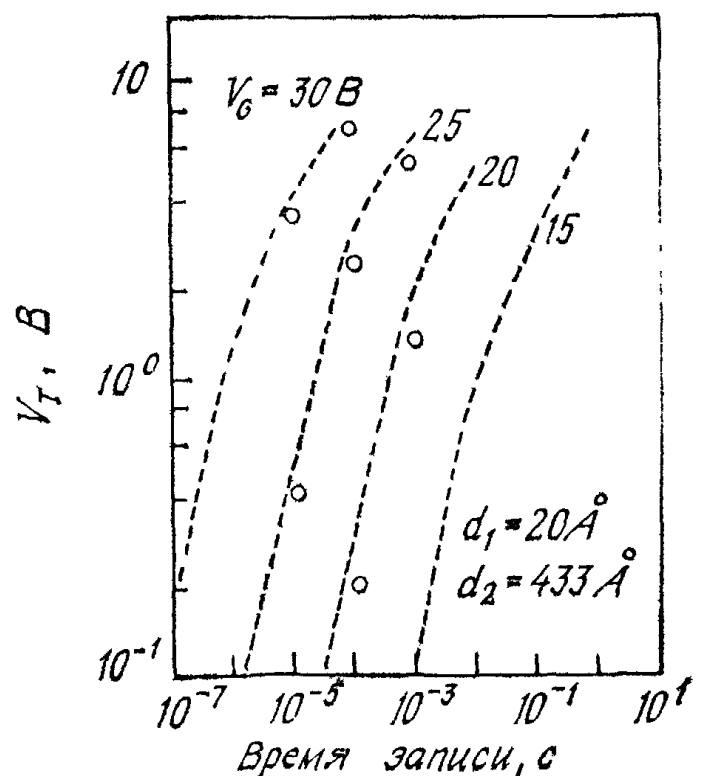


Рис. 70. Рассчитанные (— — —) и измеренные (\diamond) сдвиги порога в МНОП-структурах [78].

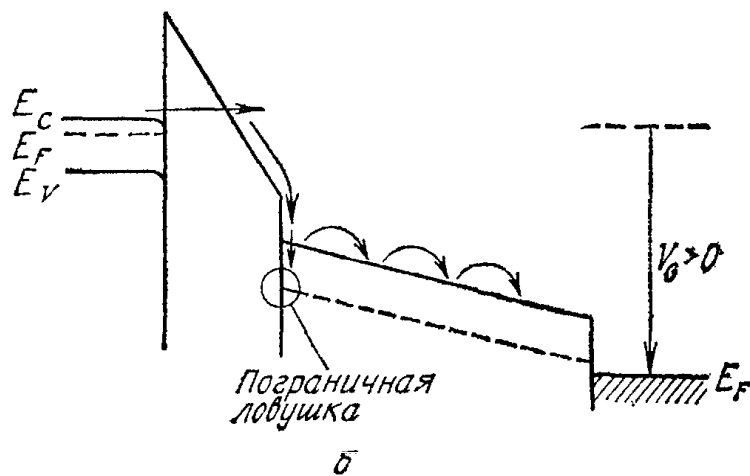
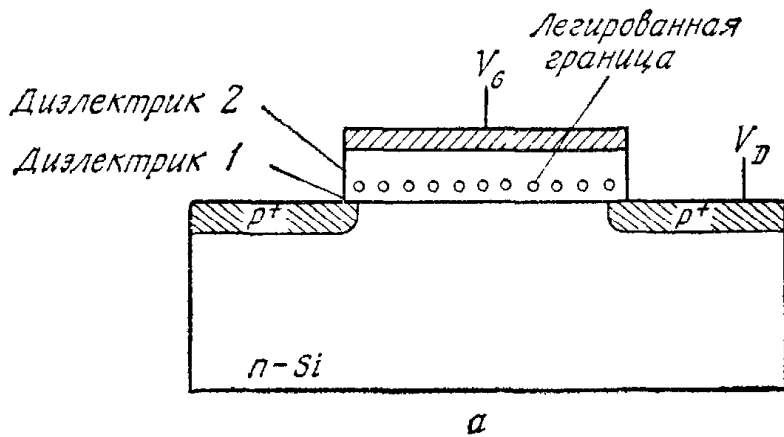


Рис. 71. Элемент памяти с легированной границей раздела диэлектриков [79].
 а — поперечное сечение прибора; б — зонная диаграмма при наличии напряжения на ватворе.

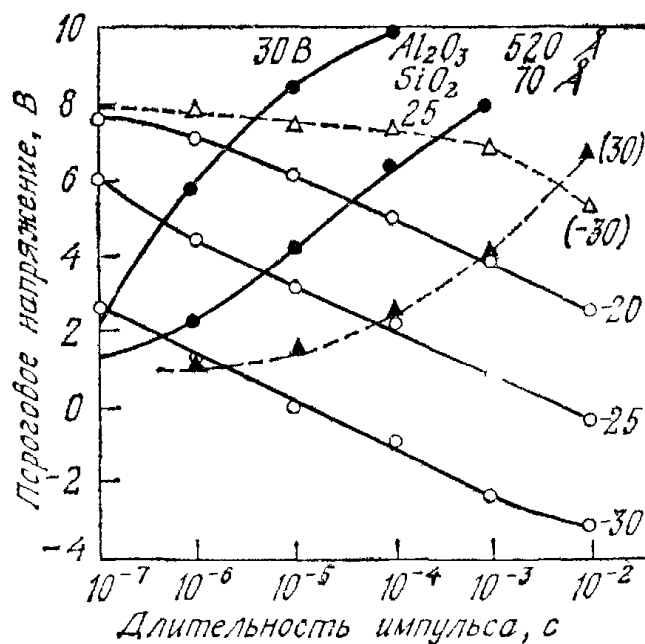


Рис. 72. Зависимости сдвига порогового напряжения от длительности импульсов записи и стирания в структуре с двойным диэлектриком [79].
 ▲ записи и Δ стирание в нелегированной структуре; ● записи и ○ стирание в структуре с легированной границей раздела ($1,5 \cdot 10^{15} \text{ см}^{-2}$).

Как видно из рис. 70, время записи, необходимое для обеспечения заданного сдвига порога, быстро уменьшается с ростом V_G . В свою очередь при фиксированном V_G сдвиг порога ΔV_T при больших временах записи стремится к насыщению (так же, как на рис. 62).

В качестве одного из способов улучшения рабочих характеристик МДОП-элементов памяти было предложено легировать границу раздела между диэлектриками (например, вольфрамом) для создания дополнительных центров захвата заряда (рис. 71,а) [79]. Оказывается, что соответствующее значительное увеличение плотности поверхностных ловушек позволяет существенно снизить времена записи и стирания (рис. 71, б). Из сравнения экспериментальных результатов, приведенных на рис. 72, следует, что времена записи ($V_G = 30$ В) и времена стирания ($V_G = -30$ В) в структурах с легированной границей раздела диэлектриков примерно на три порядка меньше соответствующих времен в обычных МНОП-структурах.

ЛИТЕРАТУРА

1. Lilienfeld J. E., U. S. Patent 1745175 (1930).
2. Heil O. British Patent 439457 (1935).
3. Shockley W., Pearson G. L. Modulation of Conductance of Thin Films of Semiconductors by Surface Charges, *Phys. Rev.*, **74**, 232 (1948).
4. Kahng D., Atalla M. M. Silicon — Silicon Dioxide Field Induced Surface Devices, IRE Solid-State Device Res. Conf., Carnegie Institute of Technology, Pittsburgh, Pa., 1960; Kahng D. A Historical Perspective on the Development of MOS Transistors and Related Devices, *IEEE Trans. Electron Devices*, **ED-23**, 655 (1976).
5. Ihantola H. K. Design Theory of a Surface Field-Effect Transistor, Stanford Electron. Lab. Tech. Rep. No. 1661-1 (1961).
6. Ihantola H. K. J., Moll J. L. Design Theory of a Surface Field-Effect Transistor, *Solid State Electron.*, **7**, 423 (1964).
7. Sah C. T. Characteristics of the Metal — Oxide — Semiconductor Transistors, *IEEE Trans. Electron Devices*, **ED-11**, 324 (1964).
8. Hofstein S. R., Heiman F. P. The Silicon Insulated-Gate Field-Effect Transistor, *Proc. IEEE*, **51**, 1190 (1963).
9. Wallmark J. T., Johnson H. Field-Effect Transistors, Physics, Technology, and Applications, Prentice-Hall, Englewood Cliffs, N. J., 1966.
10. Richman P. MOSFET's and Integrated Circuits, Wiley, N. Y., 1973.
11. Brews J. R. Physics of the MOS Transistor, in Kahng D., Ed., Applied Solid State Science, Suppl. 2A, Academic, N. Y., 1981.
12. Chang L. L., Yu H. N. The Germanium Insulated-Gate Field-Effect Transistor (FET), *Proc. IEEE*, **53**, 316 (1965).
13. Wilnsen C. W., Szpak S. MOS Processing for III—V Compound Semiconductors: Overview and Bibliography, *Thin Solid Films*, **46**, 17 (1977).
14. Looking Ahead to the Year 2000, Technology, *Electronics*, **53**, (9) 530 (1980).
15. Moore G. VLSI: Some Fundamental Challenges, *IEEE Spectrum*, **16** (4) 30 (1979).
16. Pao H. C., Sah C. T. Effects of Diffusion Current on Characteristics of Metal — Oxide (Insulator) — Semiconductor Transistors (MOST), *IEEE Trans. Electron Devices*, **ED-12**, 139 (1965).

17. Grove A. S., Fitzgerald D. J. Surface Effects on $p - n$ Junctions: Characteristics of Surface Space-Charge Regions under Nonequilibrium Conditions, *Solid State Electron.*, **9**, 783 (1966).
18. Brews J. R. A Charge-Sheet Model of the MOSFET, *Solid State Electron.*, **21**, 345 (1978).
19. Brews J. R. Subthreshold Behavior of Uniformly and Nonuniformly Doped Long-Channel MOSFET, *IEEE Trans. Electron Devices*, **ED-26**, 1282 (1979).
20. Barron M. B. Low Level Currents in Insulated Gate Field Effect Transistors, *Solid State Electron.*, **15**, 293 (1972).
21. Gosney W. M. Subthreshold Drain Leakage Current in MOS Field-Effect Transistors, *IEEE Trans. Electron Devices*, **ED-19**, 213 (1972).
22. Fichtner W., Potzl H. W. MOS Modeling by Analytical Approximations. I. Subthreshold Current and Threshold Voltage, *Int. J. Electron.*, **46**, 33 (1979).
23. Troutman R. R. Subthreshold Design Considerations for IGFET's, *IEEE J. Solid State Circuits*, **SC-9**, 55 (1974).
24. Sabnis A. G., Clemens J. T. Characterization of the Electron Mobility on the Inverted (100) Si Surface, IEEE Tech. Dig. Int. Electron Device Meet., 1979, p. 18.
25. Coen R. W., Muller R. S. Velocity of Surface Carriers in Inversion Layers on Silicon, *Solid State Electron.*, **23**, 35 (1980).
26. Cooper J. A., Nelson D. F. Observation of High Field Transport in Si Inversion Layer by a Time-of-Flight Technique, IEEE Device Res. Conf., June 23, 1980.
27. Thornber K. K. Relation of Drift Velocity to Low-Field Mobility and High Field Saturation Velocity, *J. Appl. Phys.*, **51**, 2127 (1980).
28. Yamaguchi K. Field-Dependent Mobility Model for Two-Dimensional Numerical Analysis of MOSFETs, *IEEE Trans. Electron Devices*, **ED-26**, 1068 (1979).
29. Vadasz L., Grove A. S. Temperature Dependence of MOS Transistor Characteristics below Saturation, *IEEE Trans. Electron Devices*, **ED-13**, 863 (1966).
30. Wang R., Dunkley J., DeMassa T. A., Jelsma L. F. Threshold Voltage Variations with Temperature in MOS Transistors, *IEEE Trans. Electron Devices*, **ED-18**, 386 (1971).
31. Gaensslen F., Rideout V. L., Walker E. J., Walker J. J. Very Small MOSFET's for Low-Temperature Operation, *IEEE Trans. Electron Devices*, **ED-24**, 218 (1977).
32. Fischer W. Equivalent Circuit and Gain of MOS Field-Effect Transistors, *Solid State Electron.*, **9**, 71 (1966).
33. Gallagher R. C., Corak W. S. A Metal — Oxide — Semiconductor (MOS) Hall Element, *Solid State Electron.*, **9**, 571 (1966).
34. Wittmack K., Maul J., Schulz F. Ion Implantation in Semiconductor and Other Materials, Plenum, N. Y., 1973.
35. Rideout V. L., Gaensslen F. H., LeBlanc A. Device Design Consideration for Ion Implanted n -channel MOSFETs, IBM J. Res. Dev., p. 50 (Jan. 1975).
36. Brews J. R. Threshold Shifts Due to Nonuniform Doping Profiles in Surface Channel MOSFETs, *IEEE Trans. Electron Devices*, **ED-26**, 1696 (1979).
37. Merckel G. Ion Implanted MOS Transistors — Depletion Mode Devices, in Van de Wiele F., Engle W. L., Jaspers P. G., Eds., Process and Device Modeling for IC Design, Noordhoff, Leyden, 1977.
38. Huang J. S. T., Taylor G. W. Modeling of an Ion-Implanted Silicon-Gate Depletion-Mode IGFET, *IEEE Trans. Electron Devices*, **ED-22**, 995 (1975).
39. Hendrikson T. E. A Simplified Model for Subpinchoff Condition in Depletion Mode IGFET's, *IEEE Trans. Electron Devices*, **ED-25**, 435 (1978).
40. Nishiuchi K., Oka H., Nakamura T., Ishikawa H., Shinoda M. A Normally-Off Type Buried Channel MOSFET for VLSI Circuits, IEEE Tech. Dig., Int. Electron Device Meet., 1978, p. 26.

41. Brews J. R., Fichtner W., Nicollian E. H., Sze S. M. Generalized Guide for MOSFET Miniaturization, *IEEE Electron Devices Lett.*, EDL-1, 2 (1980).
42. Yau L. D., A Simple Theory to Predict the Threshold Voltage of Short-Channel IGFET's, *Solid State Electron.*, 17, 1059 (1974).
43. Taylor G. W. Subthreshold Conduction in MOSFET's, *IEEE Trans. Electron Devices*, ED-25, 337 (1978).
44. Merckel G. Short Channels — Scaled Down MOSFET's, in Ref. 37, p. 705.
45. Fichtner W. Scaling Calculation for MOSFET's, IEEE Solid State Circuits and Technology Workshop on Scaling and Microlithography, N. Y., Apr. 22, 1980.
46. Taylor G. W. The Effects of Two-Dimensional Charge Sharing on the Above Threshold Characteristics of Short-Channel IGFET's, *Solid State Electron.*, 22, 701 (1979).
47. Yau L. D. Simple I—V Model for Short-Channel IGFET's in the Triode Region, *Electron. Lett.*, 11, 44 (1975).
48. Fukuma M., Matsumura M. A Simple Model for Short Channel MOSFET's, *Proc. IEEE*, 65, 1212 (1977)
49. Kamata T., Tanabashi K., Kobayashi K. Substrate Current Due to Impact Ionization in MOSFET, *Jpn. Appl. Phys.*, 15, 1127 (1976).
50. Sun E., Moll J., Berger J., Alders B. Breakdown Mechanism in Short-Channel MOS Transistors, IEEE Tech. Dig. Int. Electron Device Meet., Washington D. C., 1978, p. 478.
51. Ning T. H., Osburn C. M., Yu H. N. Effect of Electron Trapping on IGFET Characteristics, *J. Electron. Mater.*, 6, 65 (1977).
52. Matsunaga J., Konaka M., Kohyama S., Iizuku H. 11th Int. Conf. Solid State Devices, Aug. 27, 1979, p. 45.
53. Nicollian E. H., Berglund C. N. Avalanche Injection of Electrons into Insulating SiO₂ Using MOS Structures, *J. Appl. Phys.*, 41, 3052 (1970)
54. Dennard R. H., Gaensslen F. H., Yu H., Rideout V. L., Bassons E., LeBlanc A. R. Design of Ion-Implanted MOSFET's with Very Small Physical Dimensions, *IEEE J. Solid State Circuits*, SC-9, 256 (1974).
55. Shannon J. M., Stephen J., Freeman J. H. MOS Frequency Soars with Ion-Implanted Layers, *Electronics*, p. 96 (Feb. 3, 1969).
56. Nihira H., Konaka M., Iwai H., Nishi Y. Anomalous Drain Current in *n*-MOSFET's and Its Suppression by Deep Ion Implantation, IEEE Tech. Dig., Int. Electron Device Meet., 1978, p. 487.
57. Tarui Y., Hayashi Y., Sekigawa T. Diffusion Self-Aligned Enhance-Depletion MOS-IC, Proc. 2nd Conf. Solid State Devices, *Suppl. J Jpn. Soc. Appl. Phys.*, 40, 193 (1971).
58. Tihanyi J., Widmann D. DIMOS — A Novel IC Technology with Submicron Effective Channel MOSFET's, IEEE Tech. Dig., Int. Electron Device Meet., 1977, p. 399.
59. Masuhara T., Muller R. S. Analytical Technique for the Design of DMOS Transistors, *Jpn. J. Appl. Phys.*, 16, 173 (1976).
60. Pocha M. D., Gonzalez A. G., Dutton R. W. Threshold Voltage Controllability in Double-Diffused MOS Transistors, *IEEE Trans. Electron Devices*, ED-21, 778 (1974).
61. Nishimatsu S., Kawamoto Y., Masuda H., Hori R., Minato O. Grooved Gate MOSFET, *Jpn. J. Appl. Phys.*, 16; *Suppl. 16-1*, 179 (1977).
62. Lepselter M. P., Sze S. M. SB-IGFET: An Insulated-Gate Field-Effect Transistor Using Schottky Barrier Contacts as Source and Drain, *Proc. IEEE*, 56, 1088 (1968).
63. Weimer P. K. The TFT — A New Thin-Film Transistor, *Proc. IRE*, 50, 1462 (1962).
64. Tihanyi J., Schlotterer H. Influence of the Floating Substrate Potential on the Characteristics of ESFI MOS Transistors, *Solid State Electron.*, 18, 309 (1975).

65. Lam H. W., Tasch A. F., Holloway T. C., Lee K. F., Gibbons J. F. Ring Oscillators Fabricated in Laser-Annealed Silicon-on-Insulator, *IEEE Electron Devices Lett.*, **EDL-1**, 99 (1980).
66. Geis M. W., Flanders D. C., Antoniadis D. A., Smith H. I. Crystalline Silicon on Insulators by Graphoepitaxy, IEEE Tech. Dig., Int. Electron Device Meet., 1979, p. 210.
67. Holmes F. E., Salama C. A. T. VMOS — A New MOS Integrated Circuit Technology, *Solid State Electron.*, **17**, 791 (1974).
68. Salama C. A. T. A New Short Channel MOSFET Structure (UMOST), *Solid State Electron.*, **20**, 1003 (1977).
69. Lidow A., Herman T., Collins H. W. Power MOSFET Technology, IEEE Tech. Dig., Int. Electron Device Meet., 1979, p. 79.
70. Collins H. W., Pelly B. HEXFET, A New Power Technology, Cutts On-Resistance, Boosts Ratings, *Electron. Des.*, **17** (12), 36 (1979).
71. Kahng D., Sze S. M. A Floating Gate and Its Application to Memory Devices, *Bell Syst. Tech. J.*, **46**, 1283 (1967).
72. Nishi Y., Iizuka H. Nonvolatile Memories, in Kahng D., Ed., Applied Solid State Science, Suppl. 2A, Academic, N. Y., 1981.
73. Card H. C., Worrall A. G. Reversible Floating-Gate Memory, *J. Appl. Phys.*, **44**, 2326 (1973).
74. Frohman-Bentchkowsky D. FAMOS — A New Semiconductor Charge Storage Device, *Solid State Electron.*, **17**, 517 (1974).
75. Iizuka H., Masuoka F., Sato T., Ishikawa M. Electrically Alterable Avalanche-Injection-Type MOS Read-Only Memory with Stacked-Gate Structures, *IEEE Trans. Electron Devices*, **ED-23**, 379 (1976).
76. Chen L. I., Pickar K. A., Sze S. M. Carrier Transport and Storage Effects in Au Ion Implanted SiO₂ Structures, *Solid State Electron.*, **15**, 979 (1972).
77. Frohman-Bentchkowsky D. The Metal — Nitride — Oxide — Silicon (MNOS) Transistor Characteristics and Applications, *Proc. IEEE*, **58**, 1207 (1970).
78. Card H. C., Elmasry M. I. Functional Modeling of Nonvolatile MOS Memory Devices, *Solid State Electron.*, **19**, 863 (1976).
79. Kahng D., Sundburg W. J., Boulin D. M., Ligenza J. R. Interfacial Dopants for Dual-Dielectric Charge-Storage Cells, *Bell Syst. Tech. J.*, **53**, 1723 (1974).
80. Gibbons J. J., Lee K. F. One-Gate-Wide CMOS Inverter on Laser-Recrystallized Polysilicon, *IEEE Electron Devices Lett.*, **EDL-1**, 117 (1980).
81. DiMaria D. J., DeMeyer K. M., Dong D. W. Electrically-Alterable Memory Using a Dual Electron Injector Structure, *IEEE Electron Devices Lett.*, **EDL-1**, 179 (1980).

Часть IV

Полупроводниковые приборы СВЧ-диапазона

Глава 9

ТУННЕЛЬНЫЕ ПРИБОРЫ

9.1. ВВЕДЕНИЕ

В этой главе, а также в гл. 10 и 11 рассмотрены некоторые полупроводниковые приборы СВЧ-диапазона, перекрывающего область частот $\sim 0,1$ ГГц (10^8 Гц) — 1000 ГГц (300 см — 0,3 мм). Этот диапазон в свою очередь подразделяется на отдельные полосы частот. Названия полос и их границы приведены в табл. 1. Полоса частот 30—300 ГГц называется также диапазоном миллиметровых волн, поскольку длины волн, соответствующие этим частотам, лежат между 10 и 1 мм. Полоса более высоких частот называется полосой субмиллиметровых волн. В 1970 г. были приняты новые обозначения полос частот СВЧ-диапазона (табл. 2). Во избежание недоразумений рекомендуется одновременно указывать как название полосы частот, так и ее границы. В СВЧ-диапазоне могут работать многие полупроводниковые приборы, рассмотренные в предыдущих главах. В табл. 3 указаны основные приборы СВЧ-диапазона и принципы их работы.

В данной главе описаны приборы, работа которых основана на явлении квантовомеханического туннелирования. Первая работа, подтверждающая реальность создания туннельных приборов, была посвящена туннельному диоду, называемому также диодом Есаки, и опубликована Л. Есаки в 1958 г. [1]. Есаки в процессе изучения внутренней полевой эмиссии в вырожденном германиевом $p-n$ -переходе обнаружил «аномальную» вольт-амперную характеристику: дифференциальное сопротивление на одном из участков характеристики было отрицательным. Этот эффект он объяснил с помощью концепции квантовомеханического туннелирования и при этом получил приемлемое согласие между теоретическими и экспериментальными результатами.

В явлении туннелирования главную роль играют основные носители. Время туннелирования носителей через потенциальный барьер не описывается на привычном языке времени пролета ($\tau = W/v$, где W — ширина барьера, v — скорость носителей); оно определяется с помощью вероятности квантовомеханического

Таблица 1. Названия диапазонов длин волн и полос частот [63]

Название диапазона (полосы)	Размер волновода, см	Полоса частот, ГГц
L	—	1,0—2,6
S	7,6×3,8	2,60—3,95
G	5×2,5	3,95—5,85
C	4,4×2,3	4,90—7,05
J	3,8×1,9	5,30—8,20
H	3,2×1,3	7,05—10,00
X	2,5×1,25	8,20—12,40
M	2,1×1,2	10,00—15,00
P	1,8×1,0	12,40—18,00
N	1,5×0,85	15,00—22,00
Ku	—	15,30—18,00
K	1,2×0,65	18,00—26,50
R	0,9×0,56	26,50—40,00
Миллиметровый		>30—300
Субмиллиметровый		>300

Таблица 2. Новые названия диапазонов и полос частот [64]

Название диапазона	Полоса частот, ГГц
A	0,100—0,250
B	0,250—0,500
C	0,500—1,000
D	1,000—2,000
E	2,000—3,000
F	3,000—4,000
G	4,000—6,000
H	6,000—8,000
I	8,000—10,000
J	10,000—20,000
K	20,000—40,000
L	40,000—60,000
M	60,000—100,00
Миллиметровый	>30—300
Субмиллиметровый	>300

перехода в единицу времени. Эта вероятность пропорциональна $\exp[-2\bar{k}(0)W]$, где $\bar{k}(0)$ — среднее значение волнового вектора в процессе туннелирования, приходящееся на один носитель с нулевым поперечным импульсом и энергией, равной энергии Ферми [2]. Отсюда следует, что время туннелирования пропор-

Таблица 3. Полупроводниковые приборы СВЧ-диапазона

Прибор	Принцип работы
Варактор (гл. 2) $p-i-n$ -диод (гл. 2)	Зависимость емкости от напряжения смещения Емкость практически постоянна, высокое напряжение пробоя
Биполярный транзистор (гл. 3)	Совместное участие электронов и дырок в транспортных процессах
Диод с точечным контактом (гл. 5)	Малая площадь, малая емкость
Диод Шоттки (гл. 5)	Ток основных носителей, термоэлектронная эмиссия
Полевой транзистор с $p-n$ -переходом (гл. 6)	Основные носители; ток, модулируемый напряжением на затворе
Полевой транзистор с барьером Шоттки (гл. 6)	Основные носители; ток, модулируемый смещением на барьере Шоттки
МОП-транзистор (гл. 8)	Неосновные носители движутся в инверсионном канале у поверхности
Туннельный диод (гл. 9)	Туннелирование в прямосмещенном p^+-n^+ -переходе, отрицательное дифференциальное сопротивление
Обращенный диод (гл. 9)	Туннелирование в p^+-n^+ -переходе при обратном или почти нулевом смещении
ЛПД (гл. 10)	Лавинное умножение и пролетные эффекты для генерации мощного излучения
ИПД (гл. 10)	Эмиссия через барьер и пролетные эффекты
Пролетный диод с захваченным объемным зарядом лавины (гл. 10)	Захват плазмы в лавинно-пролетном диоде
Прибор на эффекте междолинного перехода электронов (гл. 11)	Электроны переходят из низкоэнергетических состояний с высокой подвижностью в высокоэнергетические состояния с низкой подвижностью

ционально $\exp [2k(0)W]$. Оно очень мало, и поэтому туннельные приборы можно использовать в диапазоне миллиметровых волн.

9.2. ТУННЕЛЬНЫЙ ДИОД

Благодаря высокой надежности и совершенству технологии изготовления туннельные диоды используются в специальных СВЧ-приборах с низким уровнем мощности, таких, как гетеродин и схемы для синхронизации частоты. В данном разделе описаны характеристики туннельных диодов на постоянном токе и в СВЧ-диапазоне.

Туннельный диод представляет собой простой $p-n$ -переход, обе стороны которого вырождены (т. е. очень сильно легированы примесями). На рис. 1 приведена энергетическая диаграмма туннельного диода, находящегося в состоянии термического

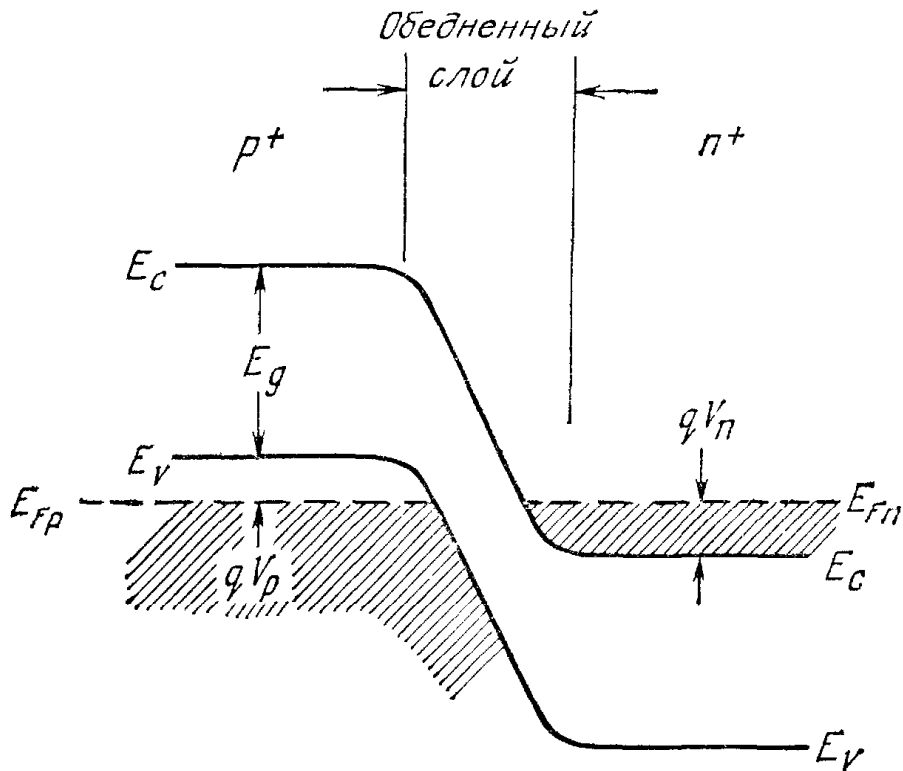


Рис. 1. Энергетическая диаграмма туннельного диода в состоянии термического равновесия. V_p и V_n — степени вырождения p -области и n -области соответственно.

равновесия. В результате сильного легирования уровень Ферми проходит внутри разрешенных зон. Степени вырождения V_p и V_n обычно составляют несколько kT/q , а ширина обедненного слоя $\sim 100 \text{ \AA}$ и меньше, т. е. намного меньше, чем в обычном p — n -переходе.

На рис. 2, а приведена типичная статическая вольт-амперная характеристика туннельного диода, из которой видно, что ток в обратном направлении (потенциал p -области отрицателен по отношению к потенциалу n -области) монотонно увеличивается. В прямом направлении ток сначала возрастает до максимального значения (пикового значения I_p) при напряжении V_p , а затем уменьшается до минимальной величины I_v при напряжении V_v . При напряжениях, превышающих V_v , ток возрастает с ростом напряжения по экспоненциальному закону. Полный статический ток диода представляет собой сумму тока туннелирования из зоны в зону, избыточного и диффузионного тока (рис. 2, б).

Сначала проведем качественный анализ туннельных процессов при температуре абсолютного нуля для упрощенной зонной структуры (рис. 3) [3]. Отметим, что уровни Ферми проходят внутри разрешенных зон полупроводника, и в состоянии термического равновесия (рис. 3, б) уровень Ферми постоянен по всему полупроводнику. Выше уровня Ферми все состояния по обеим сторонам перехода оказываются пустыми, а ниже уровня Ферми все разрешенные состояния по обеим сторонам перехода запол-

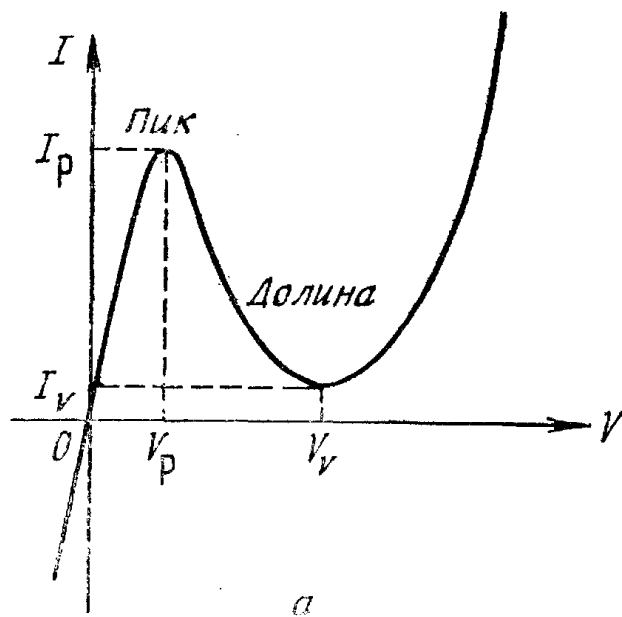
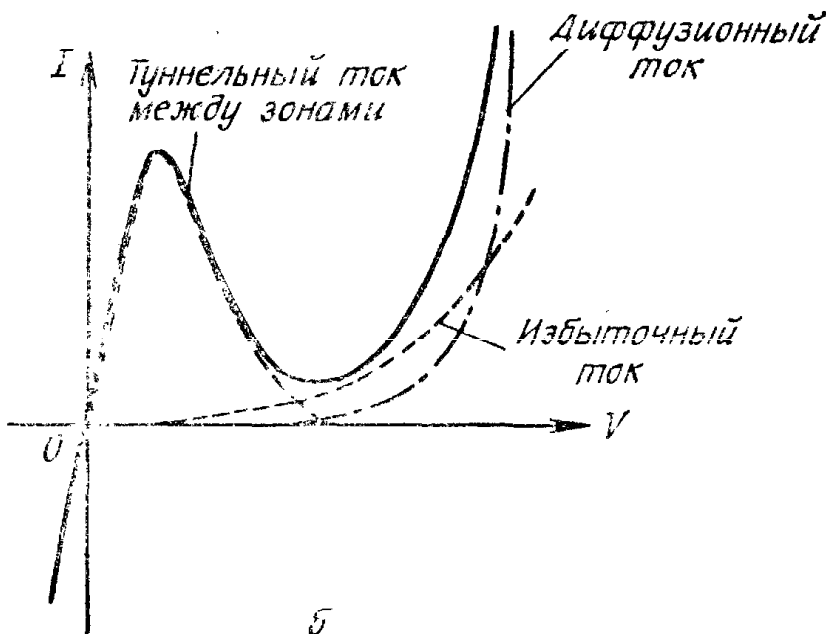


Рис. 2. Типичная статическая вольт-амперная характеристика туннельного диода (а) и три компоненты полного тока в туннельных диодах (б).



нены электронами. Поэтому в отсутствие приложенного напряжения туннельный ток не протекает.

При подаче напряжения на переход электроны могут туннелировать из валентной зоны в зону проводимости или наоборот. Для протекания туннельного тока необходимо выполнение следующих условий: 1) энергетические состояния на той стороне перехода, откуда туннелируют электроны, должны быть заполнены; 2) на другой стороне перехода энергетические состояния с той же энергией должны быть пустыми; 3) высота и ширина потенциального барьера должны быть достаточно малыми, чтобы существовала конечная вероятность туннелирования; 4) должен сохраняться квазиимпульс.

На рис. 3, а показано, как туннелируют электроны из валентной зоны в зону проводимости при обратном напряжении на диоде. Соответствующая величина тока отмечена точкой на вольт-ам-

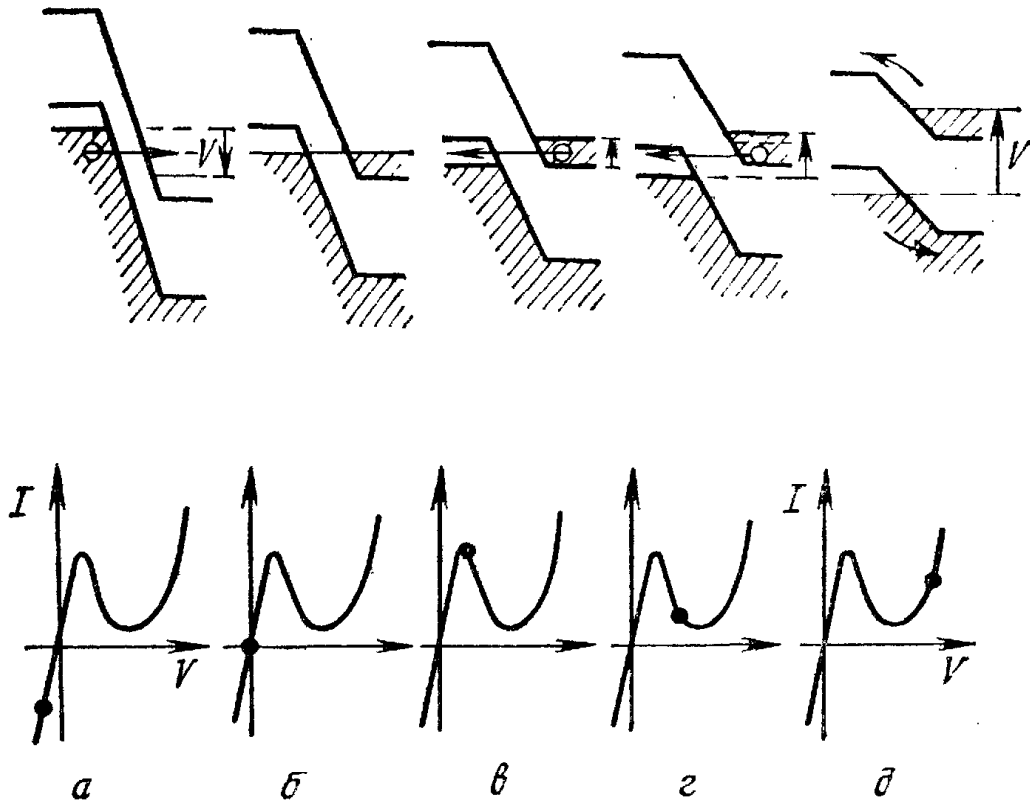


Рис. 3. Упрощенные энергетические диаграммы туннельного диода [3].

a — при обратном смещении; *б* — в тепловом равновесии, при нулевом смещении; *в* — при прямом смещении, обеспечивающем максимальное значение тока; *г* — при прямом смещении, соответствующем протеканию тока, близкого к долинному; *д* — при прямом смещении, соответствующем диффузионному току.

перной характеристике. При прямом напряжении (рис. 3, *в*) существует диапазон энергий, при которых состояния в *n*-области заполнены, а разрешенные состояния в *p*-области пусты. Естественно, что при этом электроны могут туннелировать из *n*-области в *p*-область. При увеличении прямого напряжения число разрешенных пустых состояний в *p*-области, в которые могут туннелировать электроны из *n*-области, уменьшается (рис. 3, *г*). Если же прямое напряжение имеет такое значение, что зоны «не перекрываются», т. е. энергия дна зоны проводимости точно совпадает с энергией потолка валентной зоны, то неразрешенные пустые состояния, соответствующие заполненным состояниям, отсутствуют. Следовательно, в этой точке туннельный ток должен исчезать. При дальнейшем увеличении напряжения будет протекать обычный диффузионный ток (рис. 3, *д*), который экспоненциально возрастает с ростом напряжения. Таким образом, следует ожидать, что при увеличении прямого напряжения туннельный ток сначала возрастает от нуля до максимального значения I_p , а затем уменьшается до нуля, когда приложенное прямое напряжение $V = V_n + V_p$ (рис. 1), где V_n — степень вырождения *n*-области ($V_n \equiv (E_{Fn} - E_C)/q$), а V_p — степень вырождения *p*-области ($V_p \equiv (E_V - E_{Fp})/q$). Падающий участок

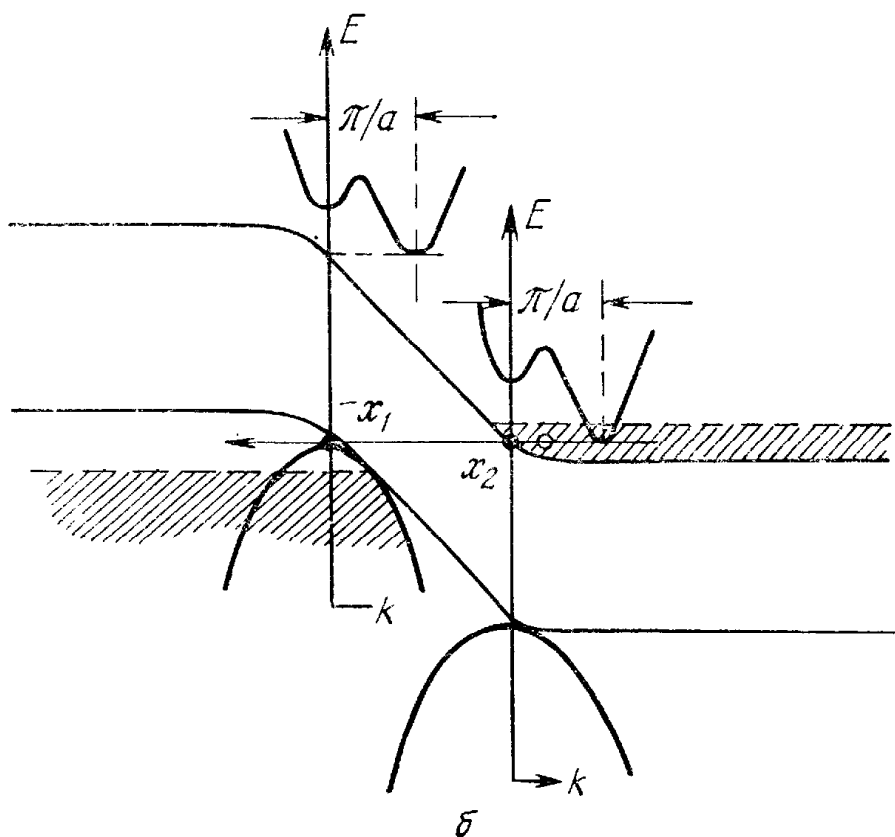
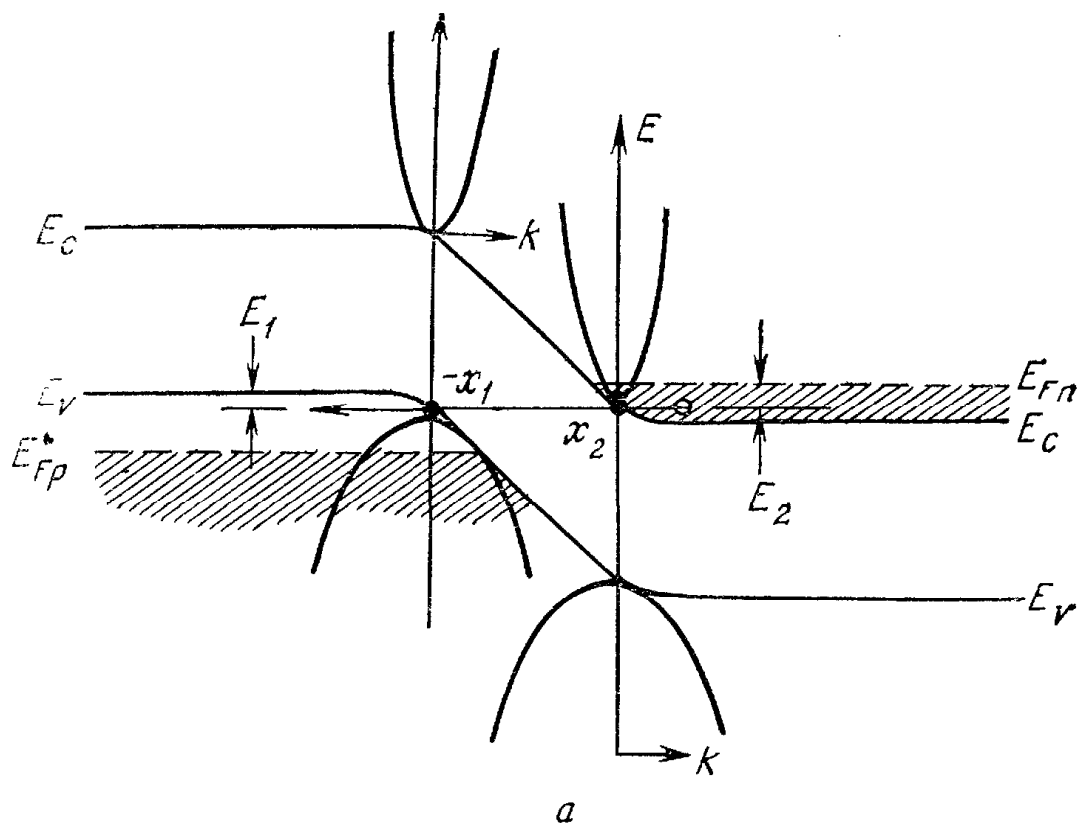


Рис. 4. Суперпозиция в классических точках поворота ($-x_1$ и x_2) структуры зон в импульсном пространстве $E - k$ и энергетической диаграммы туннельного перехода в координатном пространстве $E - x$.
 а — случай прямого туннелирования ($k_{\text{мин}} = k_{\text{макс}}$); б — случай непрямого туннелирования ($k_{\text{мин}} \neq k_{\text{макс}}$).

вольт-амперной характеристики соответствует области отрицательного дифференциального сопротивления.

Процесс туннелирования может быть прямым и непрямым. Случай прямого туннелирования показан на рис. 4, а, где структура зон в импульсном пространстве $E - k$ в классических точках поворота наложена на энергетическую диаграмму туннельного перехода в координатном пространстве $E - x$. При такой структуре зон электроны могут туннелировать из окрестности минимума зоны проводимости в окрестность максимума валентной зоны, сохраняя значение импульса. Таким образом, для того чтобы происходило прямое туннелирование, положения дна зоны проводимости и потолка валентной зоны в пространстве импульсов должны совпадать. Это условие выполняется в полупроводниках с прямой запрещенной зоной (в таких, как GaAs и GaSb). Оно может выполняться также в полупроводниках с непрямой запрещенной зоной (например, в Ge) при достаточно больших приложенных напряжениях, таких, что максимум валентной зоны (точка Г) [4] находится на одном уровне с непрямым минимумом зоны проводимости (точка Г).

Непрямое туннелирование происходит в том случае, когда минимум зоны проводимости и максимум валентной зоны расположены при различных значениях импульса (рис. 4, б). Для выполнения закона сохранения импульса разность между импульсами, соответствующими минимуму зоны проводимости и потолку валентной зоны, должна передаваться рассеивателям, таким, как фононы или примесные центры. Законы сохранения энергии и импульса для туннелирования с участием фононов формулируются следующим образом: сумма энергии фонона и начальной энергии электрона, туннелирующего из n -области в p -область, равна конечной энергии электрона, протуннелировавшего в p -область; сумма начального импульса электрона и импульса фонона ($\hbar k_p$) равна конечному импульсу протуннелировавшего электрона. В общем случае вероятность непрямого туннелирования гораздо меньше, чем вероятность прямого. Кроме того, вероятность непрямого туннелирования с участием нескольких фононов гораздо меньше вероятности аналогичного процесса с участием лишь одного фонона.

9.2.1. Вероятность туннелирования и туннельный ток

Если электрическое поле в полупроводнике достаточно велико ($\sim 10^6$ В/см), вероятность квантовомеханического туннелирования, т. е. прямого возбуждения электронов из валентной зоны в зону проводимости, становится существенной. Использование приближения ВКБ (метода Вентцеля—Крамерса—Брил-

люэна) приводит к следующему выражению для вероятности туннелирования T_t [5]:

$$T_t \simeq \exp \left[-2 \int_{-x_1}^{x_2} |k(x)| dx \right], \quad (1)$$

где $|k(x)|$ — модуль волнового вектора частицы в области барьера, а $-x_1$ и x_2 — классические точки поворота (рис. 4).

Туннелирование электрона через запрещенную зону формально идентично туннелированию частицы через барьер. Проанализируем ситуацию в случае потенциальных барьеров двух видов: треугольного и параболического (рис. 5, *a* и *б*). Внутри треугольного барьера модуль волнового вектора равен

$$k(x) = \sqrt{\frac{2m^*}{\hbar^2} (PE - E)} = \sqrt{\frac{2m^*}{\hbar^2} \left(\frac{E_g}{2} - q\mathcal{E}x \right)}, \quad (2)$$

где PE — потенциальная энергия, E — энергия электрона, E_g — ширина запрещенной зоны полупроводника, \mathcal{E} — напряженность электрического поля. Подставляя выражение (2) в формулу (1), получаем

$$\begin{aligned} T_t &\simeq \exp \left[-2 \int_{-x_1}^{x_2} \sqrt{\frac{2m^*}{\hbar^2} \left(\frac{E_g}{2} - q\mathcal{E}x \right)} dx \right] = \\ &= \exp \left[+ \frac{4}{3} \frac{\sqrt{2m^*}}{q\mathcal{E}\hbar} \left(\frac{E_g}{2} - q\mathcal{E}x \right)^{3/2} \right]_{-x_1}^{x_2}. \end{aligned} \quad (3)$$

Поскольку

$$\frac{E_g}{2} - q\mathcal{E}x = 0 \quad \text{при } x = x_2,$$

$$\frac{E_g}{2} - q\mathcal{E}x = E_g \quad \text{при } x = -x_1,$$

выражение (3) принимает вид

$$T_t \simeq \exp \left(- \frac{4 \sqrt{2m^*} E_g^{3/2}}{3g\hbar\mathcal{E}} \right). \quad (4)$$

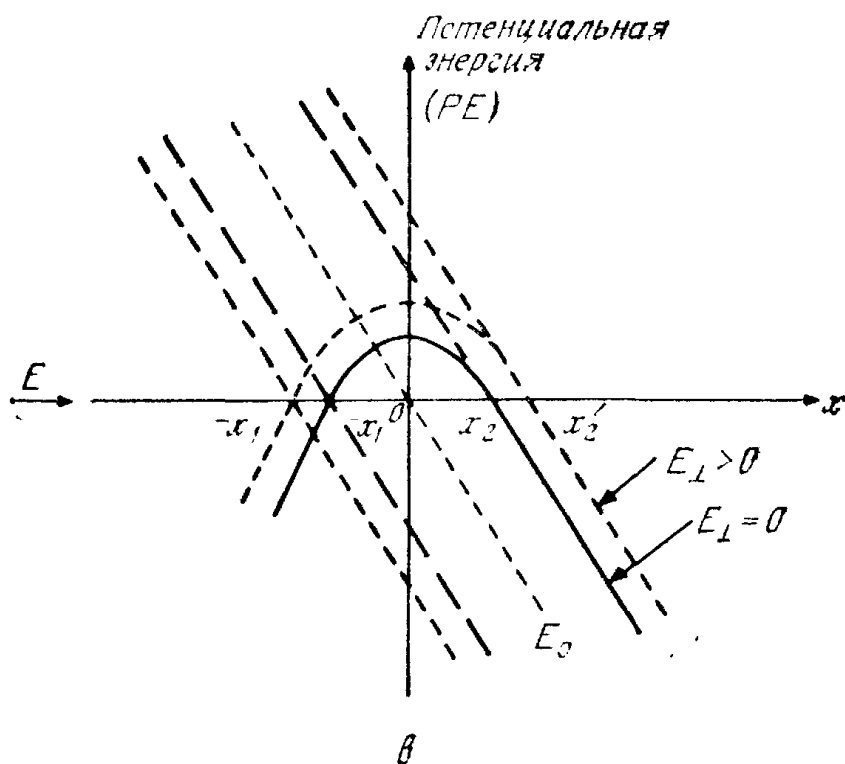
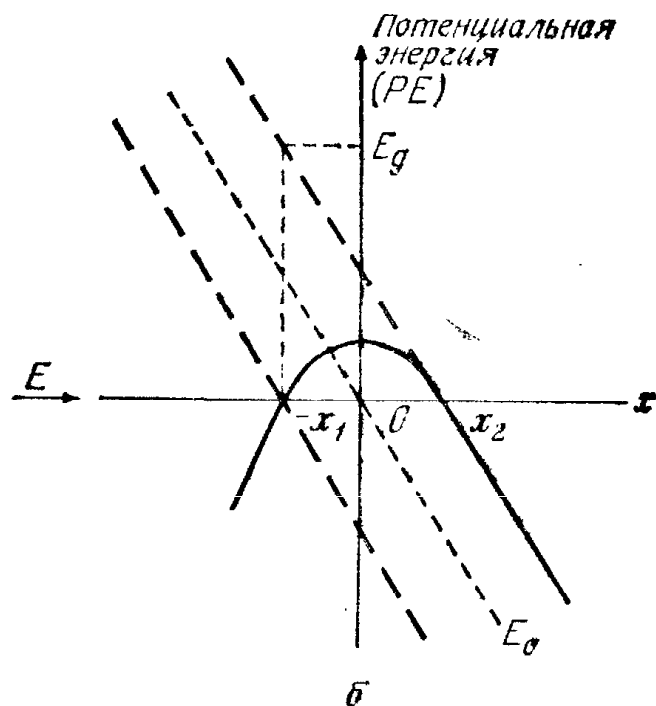
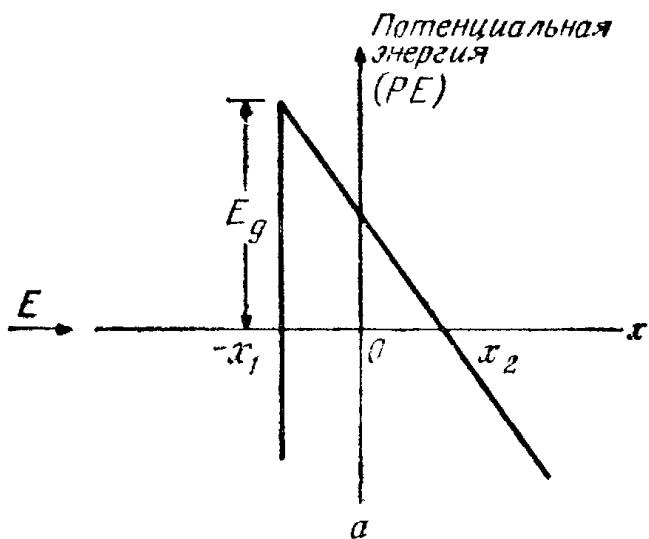
Для параболического барьера

$$PE - E = \frac{(E_g/2)^2 - E_0^2}{E_g} = \frac{(E_g/2)^2 - (q\mathcal{E}x)^2}{E_g}, \quad (5)$$

где E_0 — энергия электрона в электрическом поле. Это выражение представляет собой простейшую алгебраическую функцию,

Рис. 5. Потенциальные барьеры [10].

a — треугольный; *б* — параболический; θ — параболический при $E_{\perp} > 0$.



правильно описывающую туннелирование вблизи краев запрещенной зоны [6]. При этом вероятность туннелирования равна

$$\begin{aligned}
 T_t &\simeq \exp \left[-2 \int_{-x_1}^{x_2} \sqrt{\frac{2m^*}{\hbar^2} \left(\frac{E_g^2/4 - q^2 \mathcal{E}^2 x^2}{E_g} \right)} dx \right] = \\
 &= \exp \left[-\frac{m^*{}^{1/2} E_g^{3/2}}{2\sqrt{2} q \hbar \mathcal{E}} \int_{-1}^1 (1-y^2)^{1/2} dy \right] \Bigg|_{y=2q\mathcal{E}x/E_g} = \\
 &= \exp \left(-\frac{\pi m^*{}^{1/2} E_g^{3/2}}{2\sqrt{2} q \hbar \mathcal{E}} \right). \quad (6)
 \end{aligned}$$

Выражение (6) совпадает с выражением (4) с точностью до числового множителя.

Поскольку при туннелировании полный импульс должен сохраняться, при вычислении вероятности туннелирования необходимо учитывать роль поперечной компоненты импульса. В связи с этим полную энергию будем представлять в виде двух компонент E_x и E_\perp , где E_\perp — энергия, связанная с движением в плоскости, перпендикулярной направлению туннелирования (или с поперечной компонентой импульса), а E_x — энергия движения в направлении туннелирования. Тогда при $E_\perp > 0$ (рис. 5, в)

$$pE - E_x = \frac{E_g^2/4 - E_0^2}{E_g} + E_\perp, \quad (7)$$

и координаты классических точек поворота равны

$$-x_1, x_2 = \mp \frac{1}{q\mathcal{E}} \sqrt{E_g^2/4 + E_g/E_\perp}. \quad (8)$$

Из соотношений (1), (7) и (8) получаем вероятность туннелирования

$$T_t \simeq \exp \left(-\frac{\pi m^*{}^{1/2} E_g^{3/2}}{2\sqrt{2} q \hbar \mathcal{E}} \right) \exp \left(-\frac{2E_\perp}{\bar{E}} \right), \quad (9)$$

где первая экспонента та же, что и в выражении (6), и соответствует вероятности туннелирования с нулевым поперечным импульсом. Характерная энергия при туннелировании с отличным от нуля поперечным импульсом \bar{E} равна

$$\bar{E} \equiv \frac{4\sqrt{2} q \hbar \mathcal{E}}{3\pi m^*{}^{1/2} E_g^{1/2}}. \quad (10)$$

Если эта величина очень мала, могут туннелировать лишь электроны с малым поперечным импульсом. Другими словами, нали-

чие поперечного импульса у частицы приводит к дополнительному уменьшению вероятности туннелирования в $\exp(-2E_{\perp}/\bar{E})$ раз. При прямом туннелировании поперечный импульс должен сохраняться. Из приведенных выше результатов следует, что вероятность туннелирования тем больше, чем меньше эффективная масса и ширина запрещенной зоны и чем больше электрическое поле.

Теперь найдем величину туннельного тока. Сначала сделаем это в первом приближении, используя плотности состояний в зоне проводимости и валентной зоне [1], а затем обсудим более строгую теорию прямого туннелирования [6].

В условиях теплового равновесия туннельный ток $I_{V \rightarrow C}$ с заполненных состояний валентной зоны на пустые состояния в зоне проводимости равен току $I_{C \rightarrow V}$ с заполненных состояний зоны проводимости на пустые состояния валентной зоны в соответствии с принципом детального равновесия. Выражения для $I_{C \rightarrow V}$ и $I_{V \rightarrow C}$ имеют следующий вид:

$$I_{C \rightarrow V} = A \int_{E_C}^{E_V} F_C(E) n_C(E) T_t [1 - F_V(E)] n_V(E) dE, \quad (11a)$$

$$I_{V \rightarrow C} = A \int_E^{E_V} F_V(E) n_V(E) T_t [1 - F_C(E)] n_C(E) dE, \quad (11b)$$

где A — константа; T_t — вероятность туннелирования, которая считается равной для обоих направлений; $F_C(E)$ и $F_V(E)$ — функции распределения Ферми—Дирака; $n_C(E)$ и $n_V(E)$ — плотности состояний в зоне проводимости и валентной зоне соответственно. При смещении перехода наблюдаемый ток I_t равен

$$I_t = I_{C \rightarrow V} - I_{V \rightarrow C} = A \int_{E_C}^{E_V} [F_C(E) - F_V(E)] T_t n_C(E) n_V(E) dE. \quad (12)$$

В результате последующих вычислений формула (12) приводится к виду [7]

$$I_t = I_P (V/V_P) \exp(1 - V/V_P), \quad (13)$$

где I_P и V_P — ток и напряжение, определенные в соответствии с рис. 2. Напряжение V_P можно оценить из условия максимума произведения числа заполненных состояний в n -области на число пустых состояний с той же энергией в p -области. Результат можно записать в виде [7]

$$V_P \approx (V_n + V_p)/3. \quad (14)$$

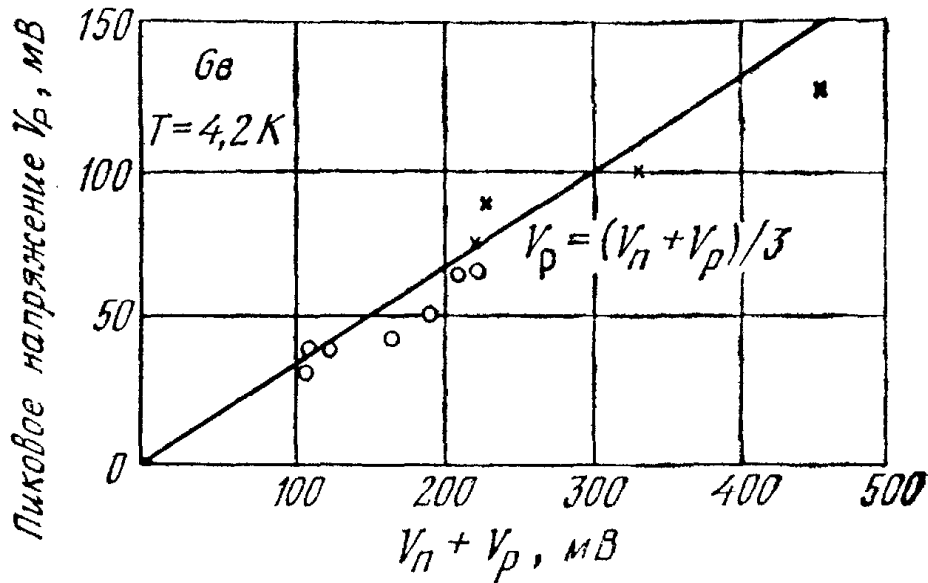


Рис. 6. Зависимость напряжения V_p для германиевых туннельных диодов от суммы V_n и V_p [7, 9].

x при изменении V_p ; O при изменении V_n .

Степень вырождения n -области V_n можно оценить из интеграла Ферми—Дирака [8]:

$$V_n \approx \frac{kT}{q} \left[\ln \left(\frac{N_D}{N_G} \right) + 0,35 \left(\frac{N_D}{N_G} \right) \right], \quad (15)$$

где N_D — концентрация доноров, N_G — эффективная плотность состояний в зоне проводимости. Аналогичное выражение можно написать для степени вырождения p -области V_p , если только в соотношении (15) заменить N_D на N_A , а N_G на N_V .

На рис. 6 приведена зависимость напряжения V_p от суммы степеней вырождения V_n и V_p , измеренная на германиевых туннельных диодах [9]. Видно, что сдвиг напряжения V_p к большим значениям при увеличении легирования и экспериментальные значения V_p достаточно хорошо согласуются с формулой (14).

При вычислении тока I_p мы сначала рассмотрим случай прямого туннелирования, т. е. будем считать импульс сохраняющимся. Предельная плотность тока в интервале энергий $dE_x dE_{\perp}$ равна [6, 10]

$$dJ_x = \frac{q (m_y^* m_z^*)^{1/2}}{2\pi^2 \hbar^3} dE_x dE_{\perp}, \quad (16)$$

где

$$E = E_x + E_{\perp}, \quad E_x = \hbar^2 k_x^2 / 2m_x^*. \quad (17)$$

Полагая m^* изотропной и равной для электронов и дырок и используя соотношения (9) и (17), получим выражение для плотности туннельного тока J_t :

$$J_t = \frac{qm^*}{2\pi^2\hbar^3} \exp\left(-\frac{\pi m^{*1/2} E_g^{3/2}}{2\sqrt{2} q\hbar \mathcal{E}}\right) \int [F_C(E) - F_V(E)] \times \\ \times \exp(-2E_\perp/\bar{E}) dE dE_\perp. \quad (18)$$

Для того чтобы в качестве переменных интегрирования фигурировали E и E_\perp , мы использовали выражение (16). Пределы интегрирования по E_\perp определяются условиями $0 \leq E_\perp \leq E_1$ и $0 \leq E_\perp \leq E_2$, где E_1 и E_2 — энергии электрона, отсчитанные от краев зоны проводимости и валентной зоны соответственно. Пределы интегрирования по E определяются краями зон.

После интегрирования по E_\perp получаем

$$J_t = \frac{qm^*}{2\pi^2\hbar^3} \exp\left(-\frac{\pi m^{*1/2} E_g^{3/2}}{2\sqrt{2} \hbar q \mathcal{E}}\right) \left(\frac{\bar{E}}{2}\right) D, \quad (19)$$

$$D \equiv \int [F_C(E) - F_V(E)] [1 - \exp(-2E_S/\bar{E})] dE, \quad (20)$$

$$\mathcal{E} \equiv (qV_{bi} N^*/2\epsilon_s)^{1/2}, \quad (21)$$

где в уравнение (20) введена величина E_S , обозначающая меньшую из величин E_1 и E_2 , а величина \bar{E} определяется выражением (10). Величина \mathcal{E} — средняя напряженность электрического поля резкого перехода; qV_{bi} — разность потенциальных энергий на краях перехода, примерно равная ширине запрещенной зоны; N^* — эффективная концентрация легирования:

$$N^* \equiv N_A N_D / (N_A + N_D). \quad (22)$$

Для германиевого туннельного диода в качестве эффективной массы в соотношении (19) следует [11] подставлять величину

$$m^* = 2 \left(\frac{1}{m_e^*} + \frac{1}{m_{lh}^*} \right)^{-1},$$

если речь идет о туннелировании из зоны легких дырок в точки $\langle 000 \rangle$ зоны проводимости Ge. Здесь $m_{lh}^* = 0,04m_0$ — масса легких дырок, а $m_e^* = 0,036m_0$ — масса электрона в зоне проводимости в точке $\langle 000 \rangle$. При туннелировании в направлении $\langle 100 \rangle$ в минимумы $\langle 111 \rangle$ эффективная масса равна

$$m^* = 2 \left[\left(\frac{1}{3m_l^*} + \frac{2}{3m_l^*} \right) + \frac{1}{m_{lh}^*} \right]^{-1},$$

где $m_l^* = 1,6m_0$ и $m_{lh}^* = 0,082m_0$ — соответственно продольная и поперечная массы минимумов $\langle 111 \rangle$. Однако показатели экспоненты в уравнении (19) отличаются для этих двух случаев лишь

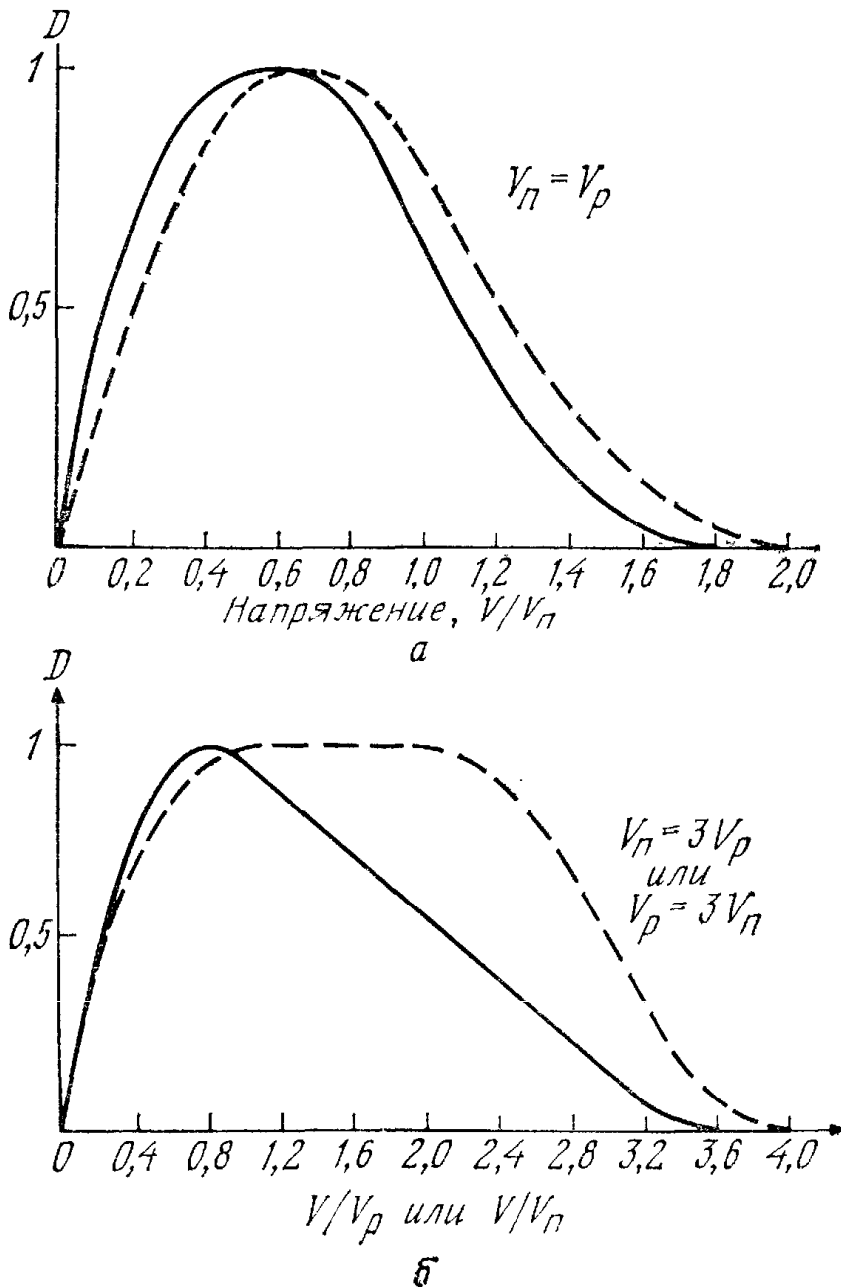


Рис. 7. Зависимость эффективной плотности состояний D от прямого смещения перехода для случаев прямого (штриховые кривые) и непрямого (сплошные кривые) туннелирования (величина \bar{E} полагается очень большой) [6].

на 5%. Величина D , определяемая уравнением (20), представляет собой интеграл перекрытия, от которого зависит форма вольт-амперной характеристики. Она имеет размерность энергии (эВ) и зависит от температуры и степеней вырождения qV_n и qV_p . При $T = 0$ К F_C и F_V имеют вид ступенек. На рис. 7 приведены зависимости величины D от прямого смещения (штриховые кривые) для случаев $V_n = V_p$ и $V_n = 3V_p$. Видно, что максимум D достигается при $V = (V_n + V_p)/3$, т. е. в полном соответствии с формулой (14).

Аналогичные зависимости приведены на рис. 7 и для случая непрямого туннелирования (сплошные кривые). Для непрямого туннелирования с участием фононов вероятность туннелирования равна [6, 12]

$$T_t \simeq \exp \left[\frac{-4\sqrt{2} m_{rx}^*{}^{1/2} (E_g - E_p)^{3/2}}{3q\hbar\mathcal{E}} \right], \quad (23)$$

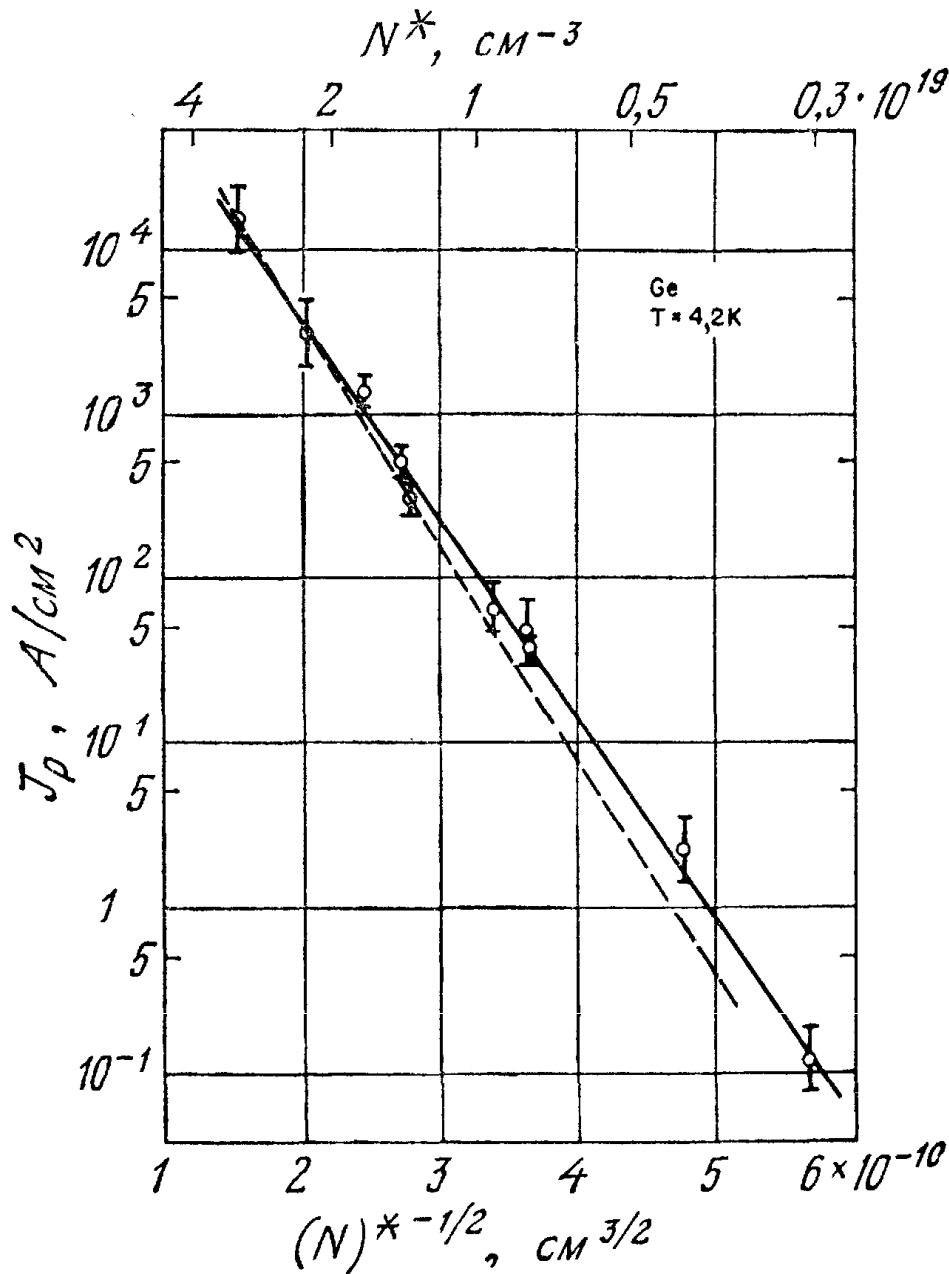


Рис. 8. Зависимость плотности пикового тока от эффективной концентрации легирования для германиевых диодов [7, 9].

где m_{rx}^* — приведенная эффективная масса в направлении туннелирования, а E_p — энергия фонона. Выражение для туннельного тока в этом случае аналогично соотношению (19). Максимум тока на прямой ветви вольт-амперной характеристики достигается при $V = V_n + V_p - (V_n^2 + V_p^2)^{1/2}$. На рис. 8 приведены экспериментальная (сплошная линия) и теоретическая, вычисленная по формуле (19), (штриховая линия) зависимости пикового тока для нескольких германиевых диодов [7].

9.2.2. Вольт-амперные характеристики

Как следует из рис. 2, б, вид статической вольт-амперной характеристики определяется тремя компонентами тока: туннельным, избыточным и диффузионным токами. В идеальном туннель-

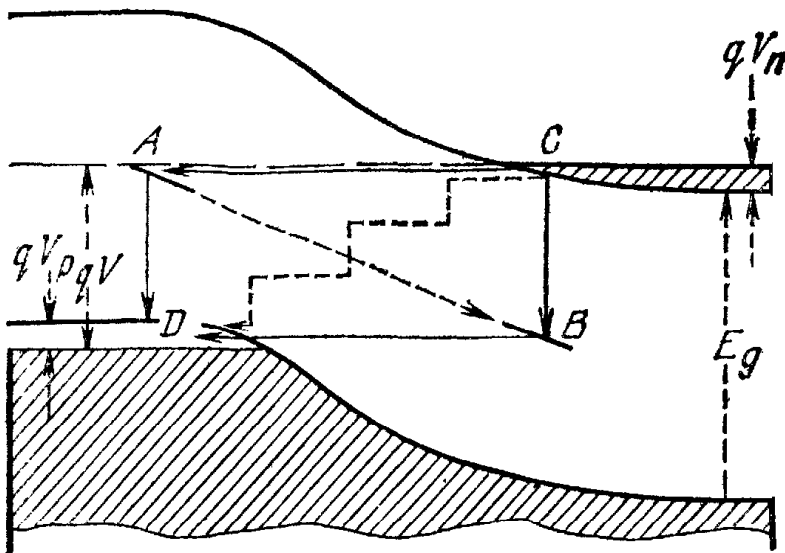


Рис. 9. Зонная диаграмма, иллюстрирующая различные механизмы туннелирования через состояния в запрещенной зоне [13].

ном диоде туннельный ток должен уменьшаться до нуля при напряжениях $V \geq V_n + V_p$; при больших напряжениях должен течь только обычный диодный ток, вызванный инжекцией неосновных носителей. Однако на практике ток туннельного диода при таких напряжениях существенно превышает обычный диодный ток, в связи с чем он получил название «избыточный ток». Основная причина избыточного тока — туннелирование носителей через энергетические состояния, расположенные в запрещенной зоне.

Для определения величины избыточного тока воспользуемся рис. 9 [13], на котором приведены несколько возможных путей туннелирования. На первом пути электрон, выходящий из точки C в зоне проводимости, может протуннелировать на подходящий локализованный уровень в точку A , из которой, отдав энергию, может попасть затем в точку D , расположенную в валентной зоне. По второму пути электрон «падает» из точки C на пустой уровень в точку B , из которой затем туннелирует в точку D . По третьему пути (маршрут $CABD$) электрон отдает свою избыточную энергию в процессе, который можно назвать проводимостью по примесной зоне между точками A и B . И наконец, четвертый путь — лестница из точки C в точку D , которая состоит из целого ряда туннельных переходов между различными локальными уровнями, сопровождаемых вертикальными скачками — потерями энергии при переходах электрона с одного уровня на другой; подобный процесс возможен при достаточно большой концентрации промежуточных уровней. Основным можно считать маршрут CBD ; остальные пути фактически являются лишь более сложными его модификациями. Рассмотрим электрон, туннелирующий из точки B в точку D при напряжении смещения на переходе V . Высота барьера, под которым он должен туннелировать, равна

$$E_x \approx E_g - qV + q(V_n + V_p) \approx q(V_{bi} - V), \quad (24)$$

где V_{bi} — контактная разность потенциалов (здесь полагается, что электрон появляется в валентной зоне вблизи ее потолка). Вероятность туннелирования T_t для электрона, движущегося на уровне точки B , равна

$$T_t \simeq \exp\left(\frac{-4\sqrt{2}m_x^*{}^{1/2}E_x^{3/2}}{3q\hbar\mathcal{E}}\right) = \exp(-\alpha_x E_x^{3/2}/\mathcal{E}). \quad (25)$$

Это выражение аналогично формуле (4), в которой величина E_g заменена величиной E_x . Максимальная напряженность электрического поля в резком переходе равна

$$\mathcal{E} = 2(V_{bi} - V)/W, \quad (26)$$

где W — ширина обедненного слоя, описываемая выражением

$$W = \left[\frac{2\epsilon_s}{q} \left(\frac{N_A + N_D}{N_A N_D} \right) (V_{bi} - V) \right]^{1/2}. \quad (27)$$

Если плотность заполненных состояний, расположенных выше потолка валентной зоны на величину E_x , равна D_x , то плотность избыточного тока определяется выражением

$$J_x \simeq A_1 D_x T_t, \quad (28)$$

где A_1 — константа, а зависимость избыточного тока от различных параметров определяется в основном их влиянием на вероятность туннелирования T_t , а не на плотность состояний D_x . Тогда, подставляя выражения (24)—(27) в формулу (28), получим следующее выражение для избыточного тока [13]:

$$J_x \simeq A_1 D_x \exp\{-\alpha'_x [E_g - qV + 0,6q(V_n + V_p)]\}, \quad (29)$$

где α'_x — константа. Из выражения (29) следует, что избыточный ток возрастает при увеличении плотности локализованных состояний в запрещенной зоне (благодаря множителю D_x), а также экспоненциально зависит от приложенного напряжения V (при условии, что $E_g \gg qV$). Выражение (29) можно также записать в виде [14]

$$\begin{aligned} J_x &= J_V \exp\left[\frac{4}{3} \left(\frac{m_x^* \epsilon_s}{N^*}\right)^{1/2} (V - V_V)\right] = \\ &= J_V \exp[A_2 (V - V_V)], \end{aligned} \quad (30)$$

где J_V — плотность тока в минимуме (в долине) при напряжении V_V , а A_2 — множитель в экспоненте. Экспериментальные зависимости $\ln J_x$ от V для туннельных диодов из кремния и арсенида галлия в соответствии с формулой (30) оказываются линейными.

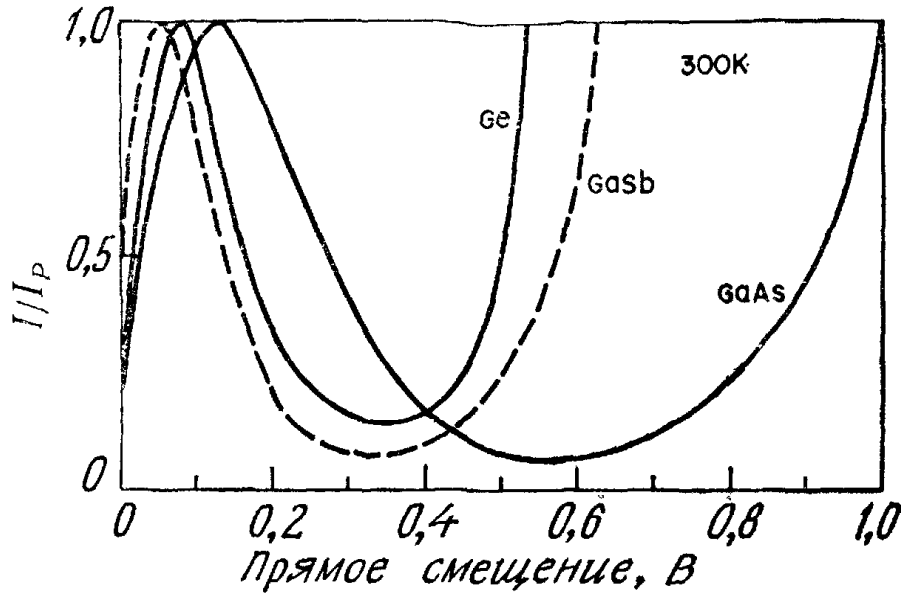


Рис. 10. Типичные вольт-амперные характеристики туннельных диодов из Ge, GaSb и GaAs при 300 К.

Диффузионный ток представляет собой ток инжекции неосновных носителей в $p-n$ -переходах

$$J_{th} = J_0 (e^{qV/kT} - 1), \quad (31)$$

где J_0 — плотность тока насыщения (гл. 2).

Полная статическая вольт-амперная характеристика определяется суммой этих трех компонент тока:

$$\begin{aligned} J &= J_t + J_x + J_{th} = \\ &= J_P \left(\frac{V}{V_P} \right) \exp \left(1 - \frac{V}{V_P} \right) + J_V \exp [A_2 (V - V_V)] + J_0 \exp \left(\frac{qV}{kT} \right). \end{aligned} \quad (32)$$

Вклад туннельного тока в полный ток является определяющим при $V < V_V$, избыточный ток наиболее существен при $V \approx V_V$, а диффузионный ток преобладает при $V > V_V$.

На рис. 10 для сравнения одновременно представлены типичные вольт-амперные характеристики туннельных диодов из Ge, GaSb [15] и GaAs [16, 17] при температуре 300 К. Величина отношения I_P/I_V равна 8 для Ge и 12 для GaSb и GaAs. Туннельные диоды были изготовлены из многих полупроводников. В диодах из Si отношение I_P/I_V равнялось примерно 4 [18], из InAs оно составляло 2 при комнатной температуре и возрастало до 10 при 4 К [19] (вследствие малой ширины запрещенной зоны), из InP отношение I_P/I_V составляло 5 [20], а в диодах на тройном соединении $\text{Ga}_{0,7}\text{Al}_{0,3}\text{As}$ достигало 12 [21]. В общем случае эта величина для данного полупроводника возрастает с увеличением уровня легирования как n -, так и p -областей. Предельное значение этого соотношения зависит, во-первых, от пикового тока,

который определяется эффективной массой туннелирования и шириной запрещенной зоны, и, во-вторых, от минимального тока, который определяется распределением и концентрацией локализованных состояний в запрещенной зоне.

Рассмотрим кратко влияние на вольт-амперную характеристику температуры, облучения электронами и давления. Температурная зависимость пикового значения тока объясняется зависимостью от температуры величин D и E_g , входящих в формулу (19). При высоких концентрациях температурная зависимость слабая и изменения вероятности туннелирования определяются в первую очередь отрицательной величиной dE_g/dT , поэтому пиковый ток растет с увеличением температуры. В слаболегированных туннельных диодах важнейшим фактором является уменьшение D с ростом температуры. Естественно, что при этом пиковый ток уменьшается с ростом температуры. Для типичных германиевых туннельных диодов изменения пикового тока в диапазоне температур -50 — $+100$ °C составляют примерно ± 10 % [22]. Ток в минимуме в общем случае растет с увеличением температуры, так как ширина запрещенной зоны уменьшается с ростом температуры (см. выражение (29)).

Основное влияние электронного облучения состоит в возрастании избыточного тока, что связано с увеличением концентрации энергетических уровней в запрещенной зоне [23]. Возросший избыточный ток можно постепенно снизить с помощью отжига. Подобные результаты получают и для других видов облучения, например γ -лучей. Влияние давления на вольт-амперную характеристику сводится к увеличению избыточного тока германиевых и кремниевых туннельных диодов [24], что обусловлено глубокими уровнями, связанными с возникающими под давлением дефектами решетки в обедненном слое. Исследования показали обратимость этих изменений. Однако для GaSb I_P и I_V уменьшаются с увеличением гидростатического давления [25], что может быть объяснено расширением запрещенной зоны и уменьшением степеней вырождения V_n и V_p при возрастании давления.

9.2.3. Характеристики прибора

Большинство туннельных диодов изготавливаются одним из следующих способов. 1) *Вплавление таблетки*. Маленькая таблетка металлического сплава, в котором в высокой концентрации растворена легирующая примесь одного типа, вплавляется в поверхность полупроводниковой подложки с сильнолегированной примесью противоположного типа при точном соблюдении режима вплавления в атмосфере инертного газа или водорода (например, мышьяк из легированной мышьяком тонкой оловянной

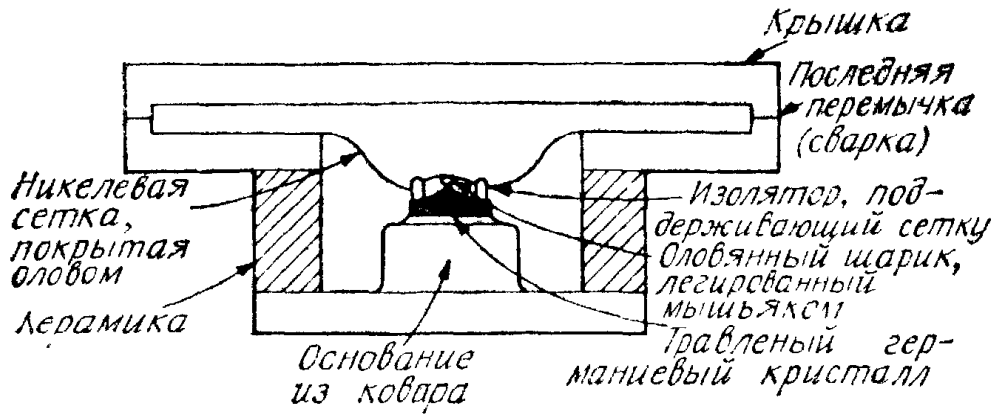


Рис. 11. Корпус, используемый в СВЧ-технике, вместе с помещенным в него германиевым туннельным диодом [27].

таблетки создает области n^+ -типа на поверхности германиевой подложки p^+ -типа). Необходимая величина пикового тока достигается после проведения процесса травления. 2) *Импульсная сварка*. Контакт и переход формируются одновременно при приложении импульсного напряжения к полупроводниковой подложке и помещенному на нее сплаву, содержащему компенсирующую примесь. 3) *Планарная технология* [26]. Планарные туннельные диоды создаются с использованием таких процессов, как эпитаксия, диффузия и контролируемое вплавление. На рис. 11 показан корпус, используемый в СВЧ-технике, с помещенным в него германиевым туннельным диодом [27]. Верхним контактом к оловянной таблетке служит луженая проволочная сетка. Диаметр перехода очень мал: он составляет $\sim 2,5$ мкм для прибора, работающего на частоте 6 ГГц, и меньше для более высокочастотных диодов.

На рис. 12, а показано символическое обозначение туннельного диода [28], а на рис. 12, б — его основная эквивалентная схема, в которую входят четыре элемента: последовательно соединенные индуктивность L_S и сопротивление R_S и параллельно соединенные емкость диода C и отрицательное сопротивление $-R$.

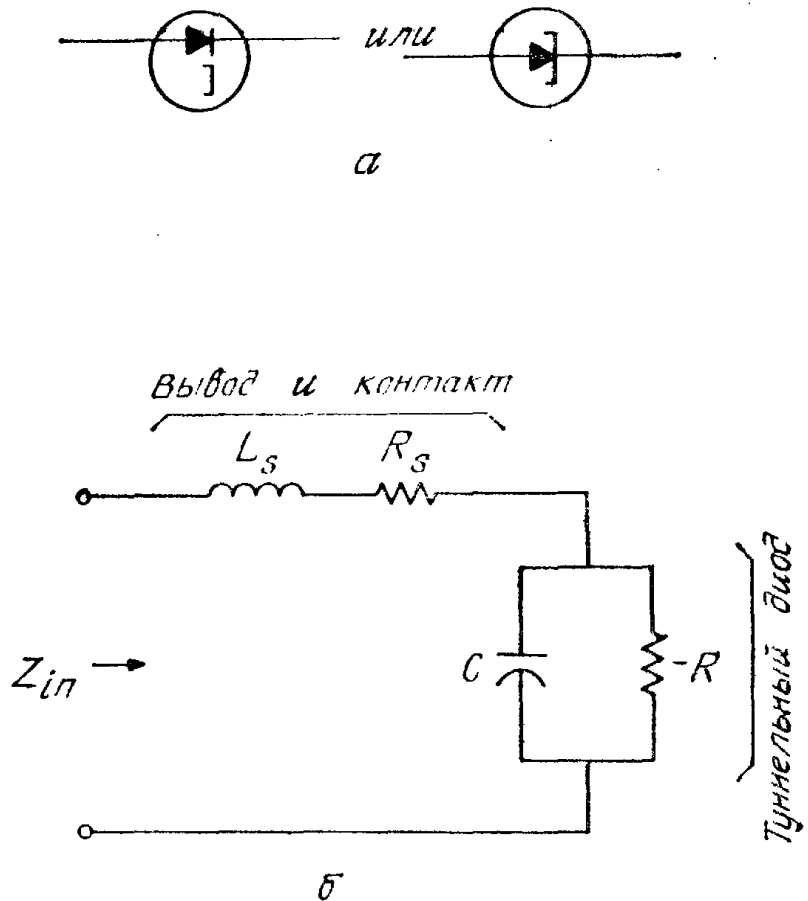
Последовательное сопротивление R состоит из сопротивлений вывода, омических контактов и сопротивления растекания подложки, равного $\rho/2d$, где ρ — удельное сопротивление полупроводника, а d — диаметр площади диода. Последовательная индуктивность L_S для коаксиальной линии равна [29]

$$L_S = \frac{2,303\mu_0 l}{2\pi} \ln\left(\frac{r_2}{r_1}\right), \quad (33)$$

где μ_0 — магнитная проницаемость среды, l — длина, а r_1 и r_2 — соответственно внутренний и внешний радиусы коаксиальной линии. Ниже показано, что эти паразитные элементы существенно ограничивают характеристики туннельного диода.

Для определения величины емкости и отрицательного сопротивления диода воспользуемся его типичной вольт-амперной ха-

Рис. 12. Символическое обозначение туннельного диода (а) и его эквивалентная схема (б) [28].



рактической характеристикой (рис. 13, а). Получаемая отсюда зависимость дифференциального сопротивления, определяемого как $(dI/dV)^{-1}$, приведена на рис. 13, б. Величина отрицательного сопротивления в точке перегиба вольт-амперной характеристики, где отрицательное сопротивление достигает минимального значения, обозначена через $R_{\text{мин}}$. Эту величину можно аппроксимировать выражением

$$R_{\text{мин}} \approx 2V_P/I_P, \quad (34)$$

где I_P и V_P — пиковый ток и соответствующее напряжение. На рис. 13, в приведена зависимость проводимости dI/dV от напряжения. При напряжениях, соответствующих пиковому току и току долины, проводимость оказывается равной нулю; емкость диода обычно измеряется в минимуме вольт-амперной характеристики и обозначается C_j .

Импеданс эквивалентной цепи Z_{in} , изображенной на рис. 12, равен

$$Z_{in} = \left[R_s + \frac{-R}{1 + (\omega RC)^2} \right] + j \left[\omega L_s + \frac{-\omega CR^2}{1 + (\omega RC)^2} \right]. \quad (35)$$

Из этого соотношения видно, что активная (действительная) составляющая импеданса при некоторой частоте может оказаться равной нулю, в то время как для реактивной (мнимой) составляющей импеданса подобное соотношение может выполняться при другом значении частоты. Мы будем называть эти частоты актив-

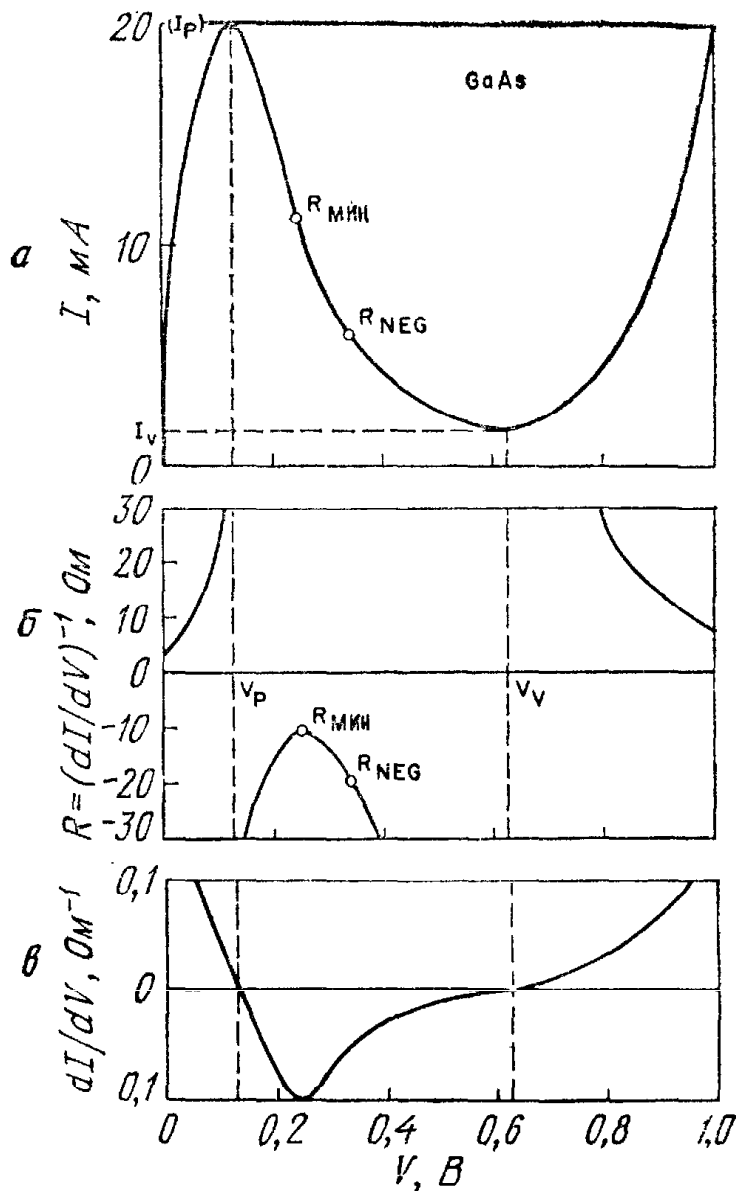


Рис. 13 Вольт-амперные характеристики туннельного диода из GaAs при 300 К (а), зависимость дифференциального сопротивления $(dI/dV)^{-1}$ от напряжения (б) ($R_{\text{мин}}$ — минимальное отрицательное сопротивление R_{NEG} — сопротивление в области минимального шум-фактора) и зависимость дифференциальной проводимости $G \equiv dI/dV$ от напряжения (в) (при пиковом и долинном токе $G = 0$).

ной частотой отсечки f_r и реактивной частотой отсечки f_x соответственно. Они равны

$$f_r \equiv \frac{1}{2\pi RC} \sqrt{\frac{R}{R_S} - 1}, \quad (36)$$

$$f_x \equiv \frac{1}{2\pi} \sqrt{\frac{1}{L_S C} - \frac{1}{(RC)^2}}. \quad (37)$$

Для частот отсечки, определенных при минимальном сопротивлении и емкости, соответствующей точке минимума вольт-амперной характеристики, имеем

$$f_{r0} \equiv \frac{1}{2R_{\text{мин}}C_j} \sqrt{\frac{R_{\text{мин}}}{R_S} - 1} \geq f_r, \quad (38)$$

$$f_{x0} \equiv \frac{1}{2\pi} \sqrt{\frac{1}{L_S C_j} - \frac{1}{(R_{\text{мин}}C_j)^2}} \leq f_x, \quad (39)$$

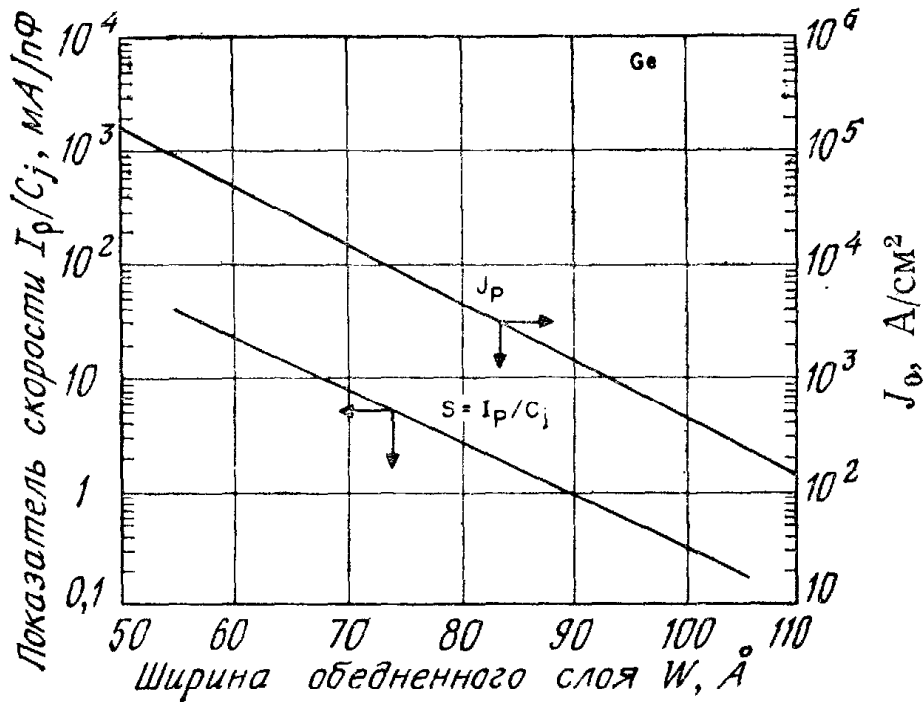


Рис. 14. Зависимости средней величины плотности пикового тока и показателя скорости переключения I_p/C_j от ширины обедненного слоя для германиевых туннельных диодов при 300 К [26].

где f_{r0} — максимальная активная частота отсечки, выше которой сопротивление диода всегда положительно; f_{x0} — минимальная реактивная частота отсечки либо резонансная частота диода, при которой реактивность диода равна нулю (диод служит генератором колебаний, если $f_{r0} > f_{x0}$). Для большинства приложений, в которых используется работа диода в области отрицательного сопротивления, требуется выполнение соотношений $f_{x0} > f_{r0}$ и $f_{r0} \gg f_0$, где f_0 — рабочая частота. Из формул (38) и (39) следует, что для выполнения требования $f_{x0} > f_{r0}$ следует уменьшать величину последовательной индуктивности L_S . Характеристикой качества туннельных диодов служит показатель скорости переключения, который определяется как отношение пикового тока к величине емкости I_p/C_j при напряжении V_p . Скорость переключения туннельного диода определяется током, необходимым для зарядки емкости перехода, и поэтому зависит от величины тока источника питания и средней величины произведения RC . Поскольку отрицательное сопротивление R обратно пропорционально пиковому току, для быстрого переключения требуется большой показатель скорости переключения (или малое произведение RC). На рис. 14 приведены зависимости пикового тока и показателя скорости переключения [26] от ширины обедненного слоя в германиевых туннельных диодах при 300 К. Мы видим, что для получения большого показателя скорости переключения необходима малая ширина слоя, т. е. более сильное легирование.

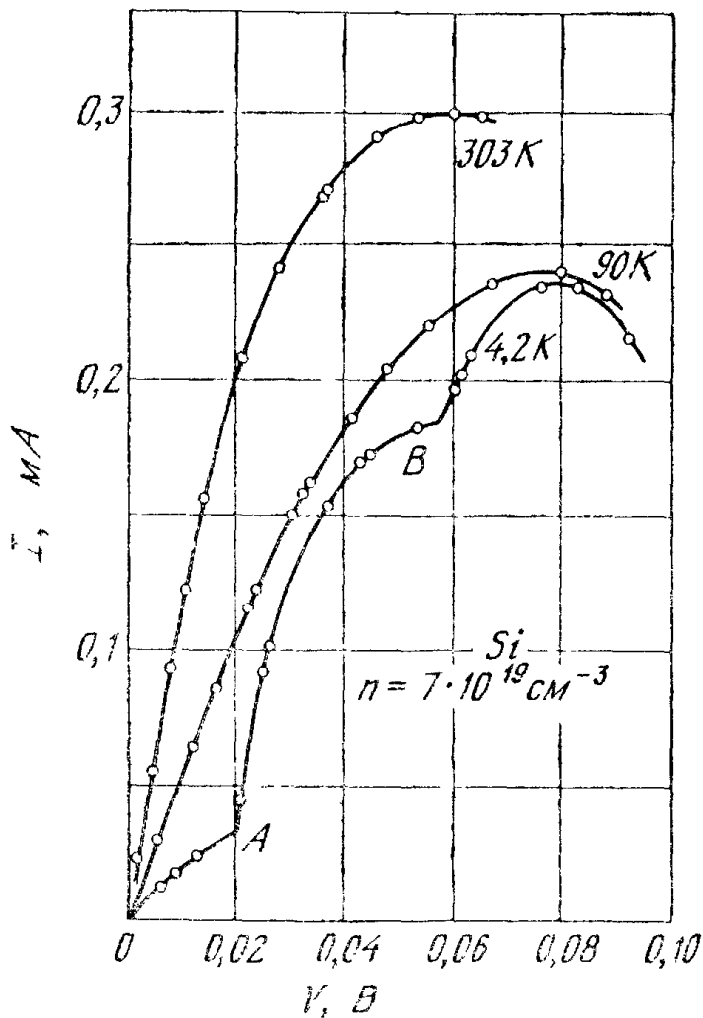


Рис. 15. Вольт-амперные характеристики кремниевого туннельного диода при трех температурах [30].

С эквивалентной цепью связана другая важная величина — шум-фактор, определяемый выражением

$$NF \equiv 1 + \frac{q}{2kT} |RI|_{\text{мин}}, \quad (40)$$

где $|RI|_{\text{мин}}$ — минимальное для данной вольт-амперной характеристики значение произведения тока на отрицательное сопротивление. Соответствующая величина R (обозначенная как R_{NFG}) приведена на рис. 13. Произведение $q|RI|_{\text{мин}}/2kT$ называется шумовой постоянной K ; эта величина является константой для данного материала. Типичные значения K при комнатной температуре равны 1,2 для Ge, 2,4 для GaAs, 0,9 для GaSb. Шум-фактор для германиевых туннельных диодов равен ~ 5 дБ при 6 ГГц и ~ 6 дБ при 14 ГГц.

Помимо использования в СВЧ- и вычислительной технике туннельный диод полезен для изучения фундаментальных физических параметров. Такой диод можно использовать для туннельной спектроскопии — метода определения энергетического спектра полупроводника, в котором зондирование спектра осуществляется при помощи туннелирования электронов с известным распределением по энергиям вместо применяемого в оптической спектроскопии поглощения фотонов фиксированной частоты. Туннельная спек-

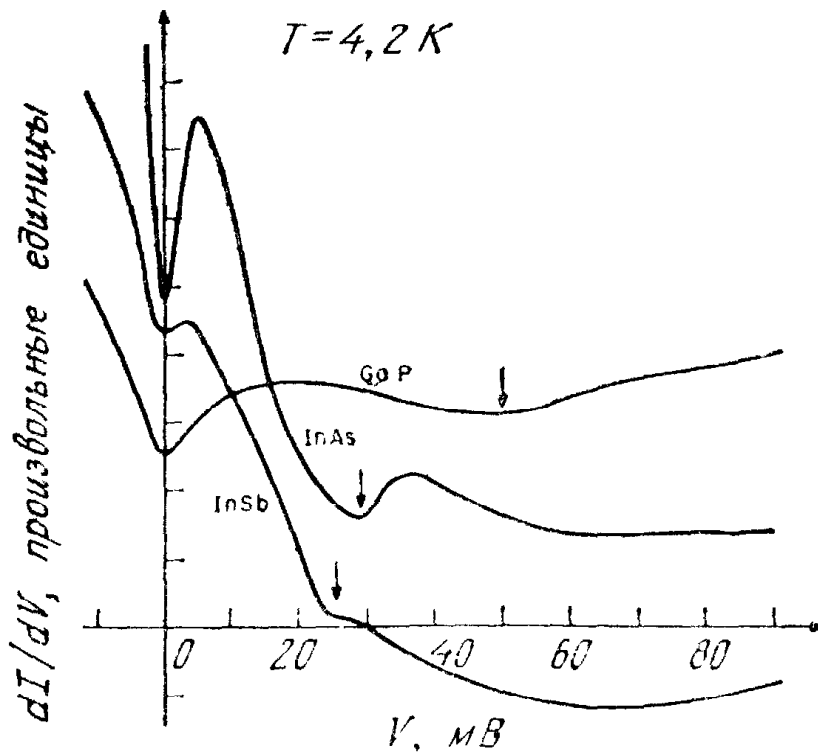


Рис. 16. Зависимость проводимости dI/dV от напряжения для туннельных диодов, созданных на полупроводниках группы $A^{III}B^V$ [31].

троскопия использовалась для изучения электронных энергетических состояний в твердых телах и для наблюдения мод возбуждения в таких приборах с $p-n$ -переходами. На рис. 15 приведены вольт-амперные характеристики кремниевого туннельного диода [30]. При уменьшении температуры до 4,2 К на кривой появляются две точки перегиба A и B . Эти точки перегиба соответствуют началу туннелирования с испусканием фононов, а энергии (или напряжения) в точках A и B равны энергиям акустических и оптических фононов соответственно. Подобные наблюдения проведены на туннельных переходах в сложных полупроводниках типа $A^{III}B^V$. На рис. 16 приведены зависимости проводимости dI/dV от напряжения V для GaP, InAs и InSb при 4,2 К [31]. Стрелками указаны энергии, соответствующие энергиям оптических фононов в этих полупроводниках. Более детальное изучение непрямого туннелирования может быть проведено при анализе второй производной d^2I/dV^2 [32].

9.3. ОБРАЩЕННЫЙ ДИОД

Если концентрации легирующих примесей таковы, что p - и n -области $p-n$ -перехода близки к вырождению либо слабо вырождены, то при малых смещениях ток в «обратном» направлении (рис. 17) превышает ток в «прямом» направлении. Отсюда происходит название такого туннельного диода — «обращенный диод». В состоянии теплового равновесия уровень Ферми в обра-

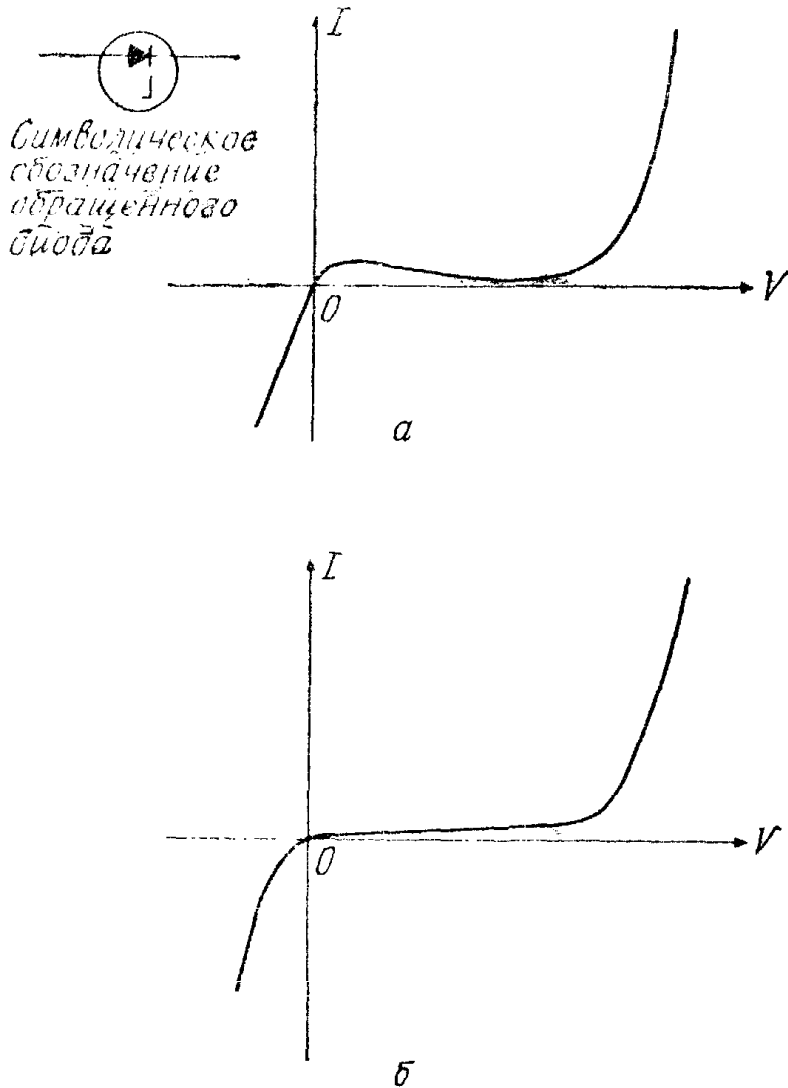


Рис. 17. Символическое обозначение обращенного диода и его вольт-амперные характеристики при наличии отрицательного сопротивления (а) и без отрицательного сопротивления (б).

щенном диоде очень близок к краям зон. При малых обратных смещениях (при отрицательном потенциале p -области по отношению к потенциалу n -области) диаграмма энергетических зон обращенного диода подобна диаграмме, приведенной на рис. 3, а, с одним отличием — обе стороны перехода теперь не вырождены. При обратном смещении электроны могут легко туннелировать из валентной зоны в зону проводимости, что по формуле (19), которую можно записать в виде

$$J \simeq B_1 \exp(+|V|/B_2), \quad (41)$$

будет приводить к возрастанию туннельного тока с ростом приложенного напряжения V . Здесь B_1 и B_2 — положительные величины, слабо зависящие от V . Из формулы (41) следует, что обратный ток зависит от напряжения примерно по экспоненциальному закону.

Обращенный диод можно использовать как выпрямитель малых сигналов, детектор СВЧ-излучения и смеситель. Подобно туннельному диоду, обращенный диод имеет хорошие частотные характеристики, поскольку его работа не связана с накоплением неосновных носителей. Кроме того, вольт-амперная характери-

стика обращенного диода нечувствительна к влиянию температуры и облучения, и в таких диодах очень мал $1/f$ -шум [33—36].

При работе в нелинейном режиме, например в качестве высокоскоростного переключателя, характеристикой совершенства прибора служит параметр γ , определяемый как отношение второй и первой производных вольт-амперной характеристики. Этот параметр также называют коэффициентом кривизны [37]:

$$\gamma \equiv \frac{d^2 I / dV^2}{dI / dV}. \quad (42)$$

Величина γ служит мерой степени нелинейности, нормированной на величину рабочей полной проводимости. При прямом смещении p — n -перехода или барьера Шоттки (гл. 5) величина γ равна q/nkT , т. е. обратно пропорциональна T . При комнатной температуре для идеального p — n -перехода ($n = 1$) $\gamma \simeq 40 \text{ В}^{-1}$ независимо от смещения. Однако при обратном смещении p — n -перехода величина γ очень мала при малых смещениях и растет пропорционально фактору лавинного умножения вблизи напряжения пробоя [38]. Хотя идеальная обратная вольт-амперная характеристика дает вблизи напряжения пробоя величину γ , большую чем 40 В^{-1} , однако из-за статистического распределения примесей и ограничения тока объемным зарядом реальные значения γ могут оказаться гораздо меньше.

Величину γ для обращенного диода можно получить на основании формул (4), (13), (14), (26). Она равна [39]

$$\gamma(\text{при } V = 0) = \frac{4}{V_n + V_p} + \frac{2}{\hbar} \sqrt{\frac{2\epsilon_s m^*}{N^*}}, \quad (43)$$

где m^* — средняя эффективная масса носителей:

$$m^* \simeq m_e^* m_h^* / (m_e^* + m_h^*),$$

а N^* — эффективная концентрация легирования, определяемая формулой (22). Ясно, что коэффициент кривизны γ зависит от концентраций примесей в обеих областях перехода и от эффективных масс. В отличие от барьеров Шоттки в обращенном диоде величина γ относительно слабо зависит от изменений температуры, поскольку параметры, фигурирующие в выражении (43), являются медленно меняющимися функциями температуры.

На рис. 18, а и б приведены теоретические и экспериментальные значения γ для германиевых обращенных диодов. Сплошными кривыми представлены результаты вычислений по формуле (43) с использованием следующих значений эффективных масс: $m_e^* = 0,22m_0$ и $m_h^* = 0,39m_0$. Совпадение результатов во всем рассмотренном диапазоне концентраций легирующих примесей в об-

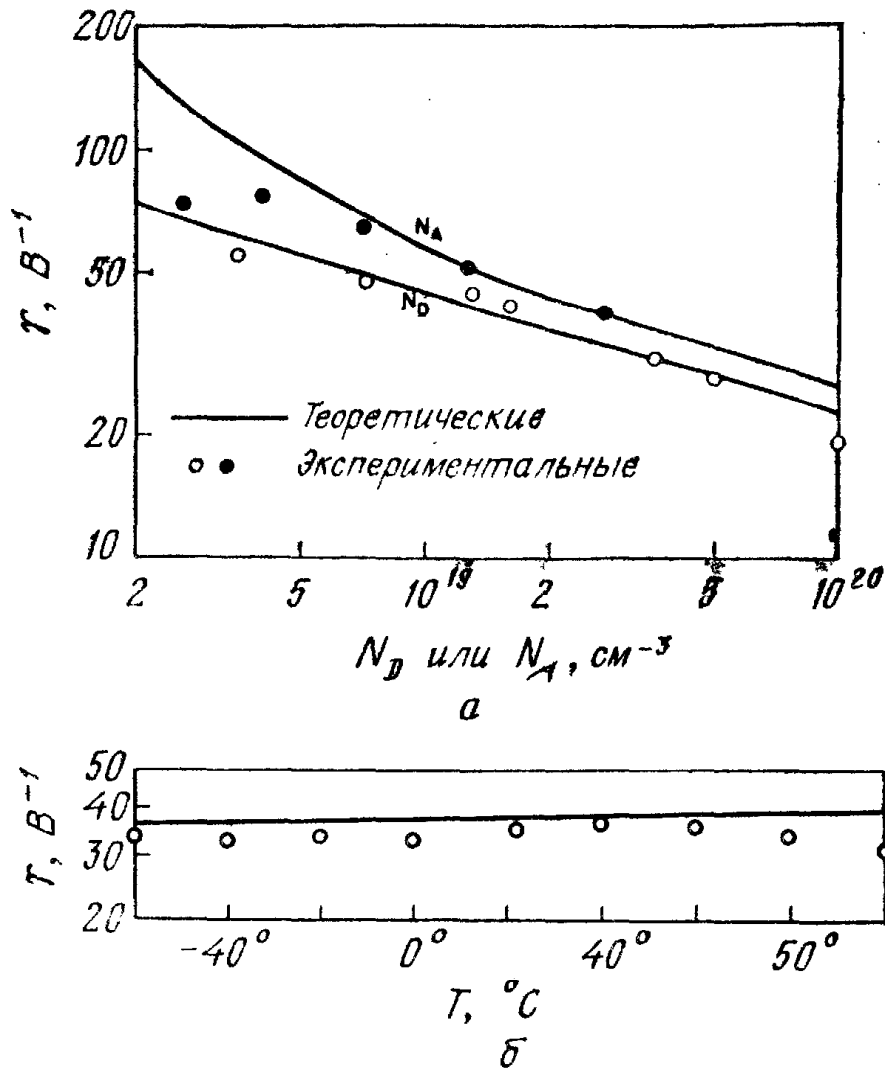


Рис. 18. Зависимость γ при 300 К и $V \simeq 0$ в германиевых диодах от концентрации акцепторов (при фиксированном значении $N_D = 2 \cdot 10^{19} \text{ см}^{-3}$) и доноров (а) (при фиксированном значении $N_A = 10^{19} \text{ см}^{-3}$ [39]) и от температуры (б) [39].

щем довольно хорошее. Отметим также две интересные особенности поведения γ для обращенных диодов: 1) γ может быть больше 40 V^{-1} , 2) нечувствительность γ к изменениям температуры.

9.4. туннельный мдп-диод

Для диода со структурой металл—диэлектрик—полупроводник (МДП) вольт-амперные характеристики критическим образом зависят от толщины диэлектрика. Если диэлектрический слой достаточно толстый (например, для системы $\text{Si}-\text{SiO}_2$ больше 50 \AA), то переходом носителей сквозь этот слой можно пренебречь, и МДП-диод представляет собой обычную МДП-емкость (гл. 7). Если же, наоборот, диэлектрический слой очень тонкий (меньше 10 \AA), то он практически не мешает движению носителей между металлом и полупроводником, и тогда структура представляет собой барьер Шоттки (гл. 5). Если же диэлектрический

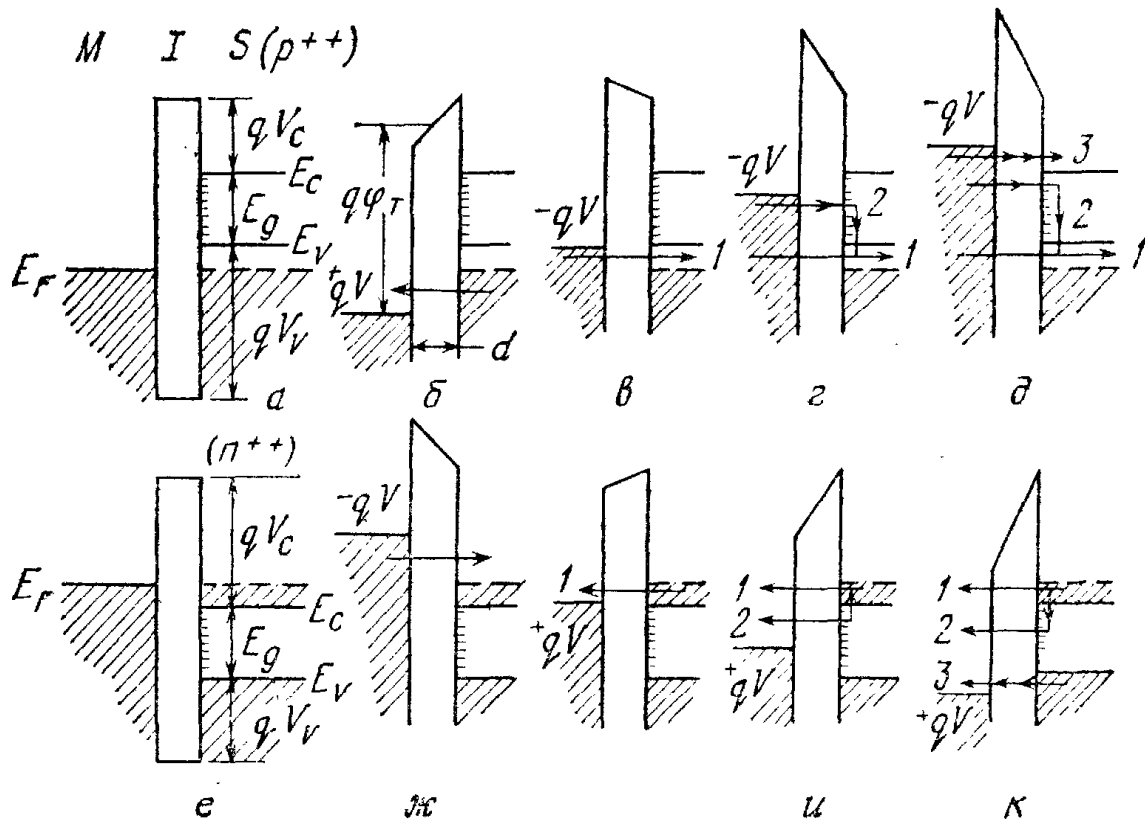


Рис. 19. Упрощенные энергетические диаграммы (учитывающие влияние поверхностных состояний) туннельных МДП-диодов на вырожденных подложках [40].

слой имеет промежуточную толщину ($10 \text{ \AA} < d < 50 \text{ \AA}$), то МДП-диод представляет собой прибор третьего класса — туннельный МДП-диод. В этом разделе рассмотрены сначала свойства туннельных МДП-диодов на вырожденных полупроводниковых подложках. Такие диоды могут обладать отрицательным сопротивлением такой же природы, что и простой туннельный диод. Затем рассмотрены свойства туннельных МДП-диодов на невырожденных подложках.

9.4.1. Вырожденный полупроводник

На рис. 19 приведены упрощенные энергетические диаграммы, учитывающие роль поверхностных состояний, для туннельных МДП-диодов на p^{++} - и n^{++} -полупроводниковых подложках [40]. Для простоты здесь не учитываются изгиб зон под действием силы изображения и падение потенциала на слое окисла в условиях равновесия. Сначала рассмотрим случай подложки p^{++} -типа. Приложение положительного напряжения к металлу (рис. 19, б) вызывает туннелирование электронов из валентной зоны в металл. Величина туннельного тока определяется выражением, аналогичным формуле (12). Используя приближение ВКБ и законы сохранения энергии и поперечного импульса k_{\perp} , выражение для плотности туннельного тока, протекающего в направлениях x

между двумя проводящими областями через запрещенную область, можно записать в виде [41]

$$J = \frac{q}{4\pi^2\hbar} \iint T_t [F_1(E) - F_2(E)] dk_{\perp}^2 dE, \quad (44)$$

где F_1 и F_2 — распределения Ферми в обеих проводящих областях, а T_t — вероятность туннелирования. В рассматриваемом нами здесь случае МДП-диода следует иметь в виду, что поверхности Ферми для электронов в полупроводнике значительно меньше, чем в металле. Вследствие этого можно полагать, что туннелирование электронов из полупроводника в металл всегда разрешено. Если же кроме этого предположить, что энергетические зоны имеют параболическую форму с изотропной электронной массой m^* , то формулу (44) можно записать в более простом виде

$$J = \frac{m^*q}{2\pi^2\hbar^3} \iint T_t dE_{\perp} dE, \quad (45)$$

где E_{\perp} и E — поперечная и полная кинетические энергии электронов в полупроводнике. Пределами интегрирования по E_{\perp} являются нуль и E , а пределами интегрирования по E служат два уровня Ферми. Вероятность туннелирования через прямоугольный барьер с эффективной высотой $q\phi_T$ и шириной d (рис. 19, б) может быть получена из формулы (1) [42]:

$$T_t \simeq \exp \left[- 2 (2m^*/\hbar^2)^{1/2} (q\phi_T)^{1/2} d \right] \simeq \quad (46)$$

$$\simeq \exp \left(- \alpha_T \phi_T^{1/2} d \right), \quad (46a)$$

где постоянная $\alpha_T \simeq 1$, если эффективная масса в диэлектрике равна массе свободного электрона и если значения ϕ_T и d измеряются соответственно в вольтах и ангстремах.

Этот туннельный ток (рис. 19, б) монотонно возрастает с увеличением разности энергий между положениями уровней Ферми; кроме того, он увеличивается с уменьшением высоты барьера. Приложение к металлу малого отрицательного напряжения (рис. 19, в) приводит к туннелированию электронов из металла на свободные состояния в валентной зоне полупроводника. Как видно из рис. 19, г, увеличение напряжения $-V$ приводит к увеличению эффективной высоты барьера, через который электроны туннелируют из металла на пустые состояния в валентной зоне, а следовательно, к появлению падающего участка на вольт-амперной характеристике (при условии, что $qV_c < qV$ (рис. 19, а)). Однако вклад в ток может вносить и компонента другой природы, обусловленная туннелированием электронов с большими энергиями из металла на пустые поверхностные состояния, где эти электроны будут рекомбинировать с дырками из валентной зоны. Поскольку высота барьера для такого туннелирования умень-

шается с ростом смещения, вольт-амперная характеристика этой компоненты тока будет монотонно возрастающей. Наконец, при дальнейшем увеличении смещения появляется третья, очень быстро нарастающая компонента туннельного тока, обусловленная переходом носителей из металла в зону проводимости полупроводника (рис. 19, *д*).

Теперь рассмотрим туннелирование в структуре с полупроводником n^{++} -типа. Эффективная высота диэлектрического барьера в случае полупроводника n^{++} -типа (рис. 19, *е*) должна быть меньше, чем в случае полупроводника p^{++} -типа, а следовательно, должна быть больше и величина туннельного тока. При отрицательном напряжении на металле электроны туннелируют из металла на пустые состояния в зоне проводимости полупроводника, что приводит к большому, быстро нарастающему току (рис. 19, *ж*). Малое положительное напряжение на металле приводит к усилению туннелирования электронов из зоны проводимости полупроводника в металл (рис. 19, *з*). Если поверхностные состояния заполнены электронами, захваченными из зоны проводимости, дальнейшее увеличение напряжения приведет к росту второй компоненты тока, обусловленной туннелированием электронов с поверхностных состояний в металл. Эта компонента тока растет с увеличением напряжения вследствие уменьшения эффективного диэлектрического барьера (рис. 19, *и*). При больших напряжениях, кроме того, оказывается возможным туннелирование из валентной зоны в металл. Однако вклад этой компоненты тока в полную вольт-амперную характеристику невелик из-за относительно большой высоты барьера (рис. 19, *к*). Таким образом, влияние зонной структуры полупроводника на туннельные характеристики МДП-структур с подложками n^{++} -типа оказывается гораздо более слабым, чем в случае структур p^{++} -типа.

На рис. 20 приведены вольт-амперные характеристики трех образцов, изготовленных на кремниевых p^{++} -подложках со слоями окисла толщиной 20 Å (0,0006 Ом·см) и подвергнутых различным обработкам. Окислы выращивались в атмосфере сухого кислорода (кривая *а*), во влажной атмосфере (кривая *б*) и во влажной атмосфере с последующим 30-минутным отжигом при 350 °С в атмосфере H_2 (кривая *в*). Зонная структура полупроводника лучше проявляется на вольт-амперных характеристиках, снятых при отрицательных напряжениях. Видно, что ток относительно медленно нарастает при увеличении напряжения до значения $V \approx \approx -1$ В, после чего скорость возрастания тока увеличивается. Это напряжение соответствует ширине запрещенной зоны кремния при сильном легировании, приводящем к образованию хвостов плотности состояний вблизи краев зон.

Очевидно, что отрицательное сопротивление, предсказываемое при малых отрицательных напряжениях, как это говорилось

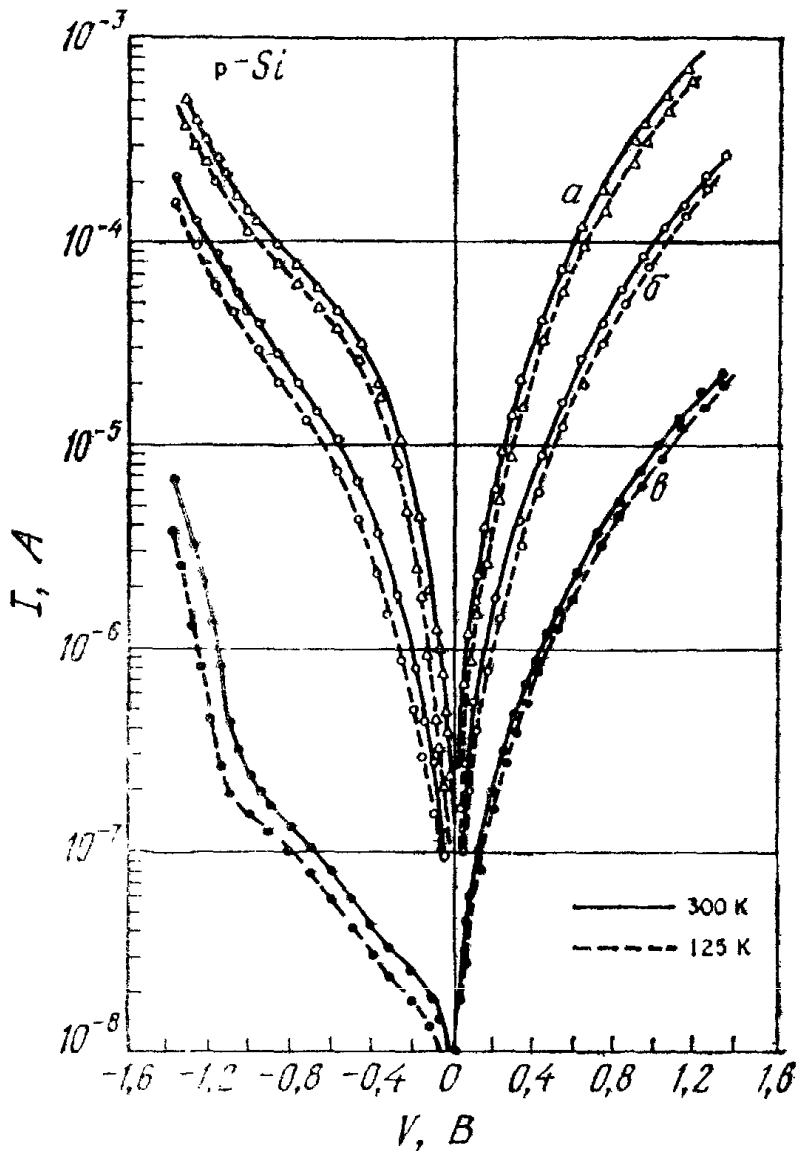
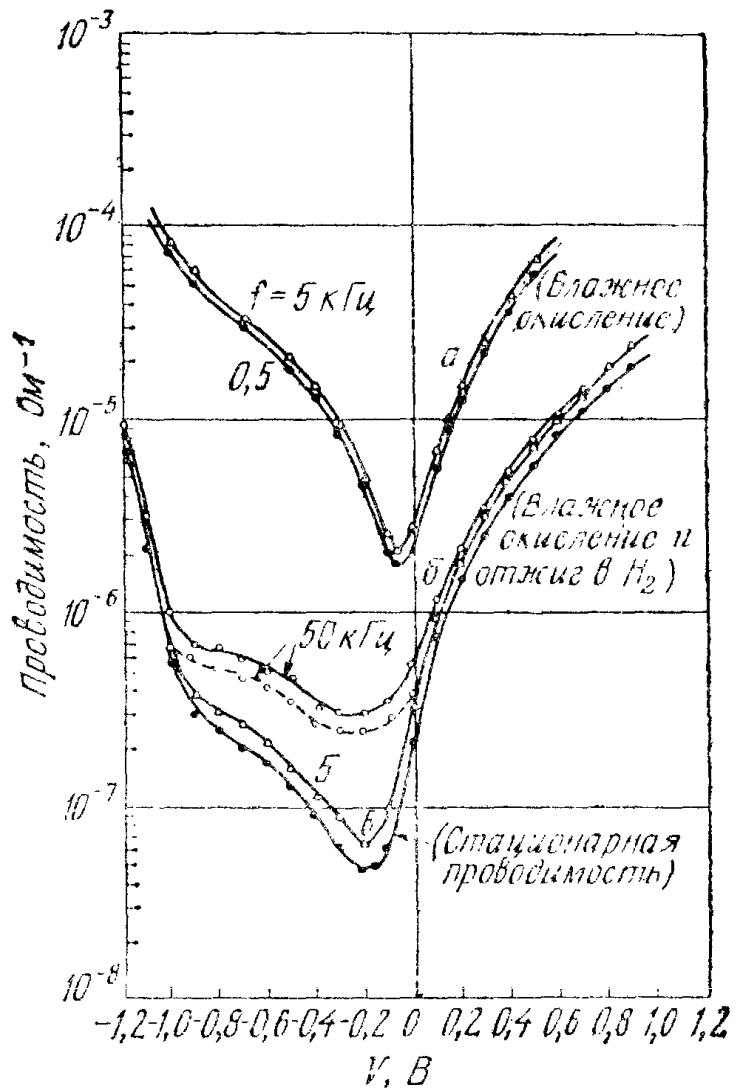


Рис. 20. Вольт-амперные характеристики трех образцов на кремниевой p^{++} -подложке со слоем окисла толщиной 20 Å, созданным различными способами. Характеристики, измеренные при комнатной температуре (300 К), указаны сплошными линиями, а измеренные при температуре жидкого азота (125 К) — штриховыми линиями [40].

выше, маскируется туннелированием электронов из металла на поверхностные состояния. Кривые a и b принципиально не отличаются от кривой b , однако для них характерны существенно высокие значения тока при напряжениях, соответствующих запрещенной зоне ($-1,1 \text{ В} < V < 0$). Если предположить, что величина тока в этом диапазоне напряжений пропорциональна плотности поверхностных состояний, то отсюда следует вывод о возрастании плотности поверхностных состояний на один-два порядка величины при отказе от водородного отжига в случае влажного окисления или при выращивании окисла в атмосфере сухого кислорода. Этот вывод качественно согласуется с экспериментальными данными, полученными при изучении плотности поверхностных состояний, возникающих при аналогичных способах выращивания окислов большой толщины ($d \approx 1000 \text{ Å}$) [43]. Влияние зонной структуры полупроводника и плотности поверхностных состояний на характеристики туннелирования в стационарных условиях оказывается еще более ярко выраженным на зависимости проводимости от напряжения (рис. 21, кри-

Рис. 21. Зависимости проводимости от напряжения, измеренные при различных частотах. Кривая для стационарной проводимости получена с помощью дифференцирования кривых, приведенных на рис. 20 [40].

— 300 К; --- 125 К.



вая в), которая получена дифференцированием кривых, приведенных на рис. 20. Левые ветви кривых при $V < -1,1$ В соответствуют туннелированию электронов из металла в зону проводимости; правые ветви при $V > 0$ соответствуют туннелированию из валентной зоны в металл. Минимум проводимости при малых отрицательных напряжениях есть результат суперпозиции двух компонент тока (рис. 19, г). Очевидно, что ожидаемая отрицательная проводимость, связанная с первой компонентой тока, компенсируется большей положительной проводимостью, связанной с протеканием тока через поверхностные состояния.

На рис. 21 и 22 приведены соответственно зависимости проводимости, измеренной на переменном токе, от приложенного напряжения и частоты. Как и следовало ожидать, кривые для емкости неотожженных образцов расположены выше, чем кривые для емкости образцов, подвергнутых отжигу, поскольку в первом случае плотность поверхностных состояний выше. Частотная зависимость проводимости оказывается более сильной. При частотах, меньших 5 кГц, зависимость проводимости от напряжения практически совпадает со статической. На рис. 22 (вставка) показана основная эквивалентная схема, с помощью которой учиты-

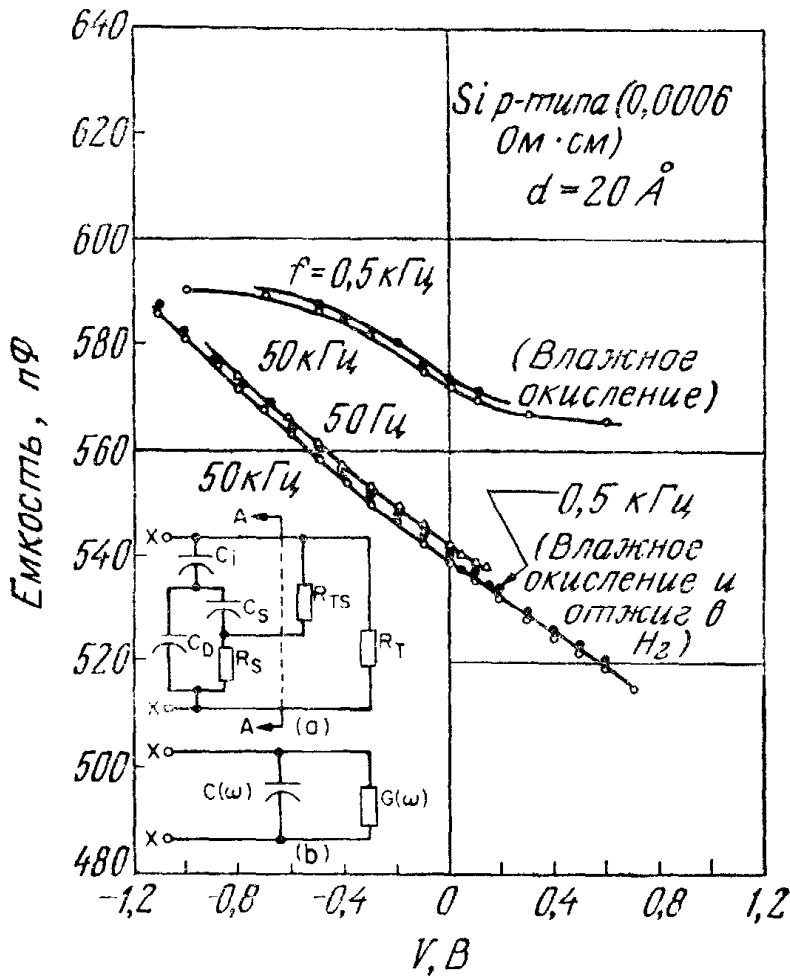


Рис. 22. Зависимости емкости от напряжения, измеренные при различных частотах. На вставке показана эквивалентная схема туннельного МДП-диода [40].

вается влияние тунnelирования. RC -цепь, расположенная левее линии AA , идентична RC -цепи для более толстого слоя окисла. R_T — эквивалентное сопротивление, соответствующее тунnelированию электронов в валентную зону и (или) в зону проводимости p^{++} -полупроводника (ему соответствуют первая и третья компоненты тока на рис. 19, в и д). R_{TS} — эквивалентное сопротивление, соответствующее тунnelированию электронов на поверхностные состояния и последующей их рекомбинации с дырками из валентной зоны (чему соответствует вторая компонента тока на рис. 19, г или д). Величины R_T и R_{TS} являются функциями приложенного напряжения. Полную цепь можно упростить, заменив ее параллельно соединенными емкостью $C(\omega)$ и проводимостью $G(\omega)$, зависящими от частоты.

Можно показать, что $d[C(\omega)]/d\omega \leq 0$ и $d[G(\omega)]/d\omega \geq 0$. Для тонких окисных слоев и (или) сильнолегированных полупроводниковых подложек емкость увеличивается по сравнению с высокочастотным значением [40]

$$C(\infty) = \frac{C_i C_D}{C_i + C_D} \quad (47a)$$

вплоть до величины

$$C(0) \approx C(\infty) + AC_s, \quad (47б)$$

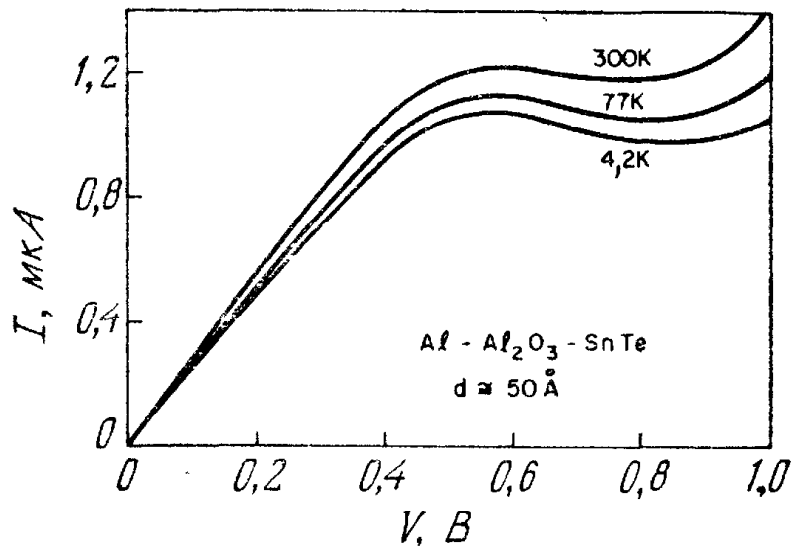


Рис. 23. Вольт-амперные характеристики туннельного МДП-диода (Al — Al₂O₃ — SnTe) при трех температурах [44].

если $\omega \rightarrow 0$. Проводимость увеличивается от низкочастотного значения

$$G(0) = \frac{1}{R_T} + \frac{1}{R_{TS} + R_s} \quad (48a)$$

вплоть до величины

$$G(\infty) = G(0) + \frac{A(R_{TS} + R_s)}{R_{TS}R_s} \quad (48b)$$

при $\omega \rightarrow \infty$. Величина

$$A \equiv \frac{(C_i R_{TS} - C_D R_s)^2}{(C_i + C_D)^2 (R_{TS} + R_s)^2} \quad (49)$$

Экспериментальные результаты находятся в хорошем согласии с приведенными выше качественными соображениями (рис. 21 и 22).

Отрицательное сопротивление наблюдалось в туннельных МДП-диодах из Al—Al₂O₃—SnTe [44]. Материал SnTe представлял собой сильнолегированный полупроводник *p*-типа с концентрацией носителей $8 \cdot 10^{20} \text{ см}^{-3}$; толщина слоя Al₂O₃ составляла $\sim 50 \text{ \AA}$. На рис. 23 приведены измеренные вольт-амперные характеристики такой структуры при трех различных температурах. Участок отрицательного сопротивления существует в диапазоне 0,6—0,8 В. Эти результаты хорошо согласуются с теоретическими результатами [41], полученными по формуле (45).

9.4.2. Невырожденный полупроводник

Туннельные МДП-диоды с очень тонким слоем окисла, выращенным на невырожденной полупроводниковой подложке, не обладают отрицательным сопротивлением. Этот диод используется

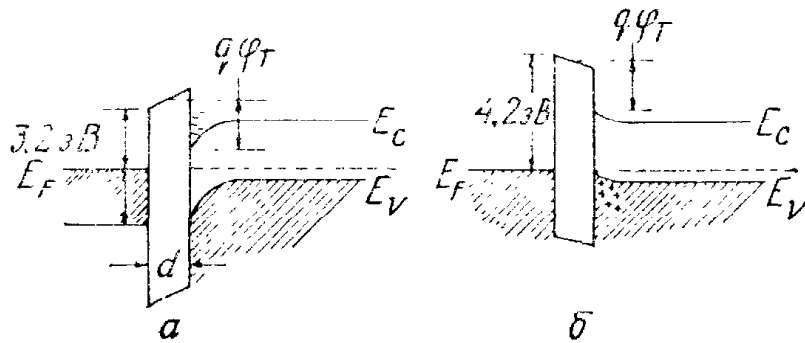


Рис. 24 Энергетические диаграммы туннельных МДП-диодов на невырожденных подложках [45].

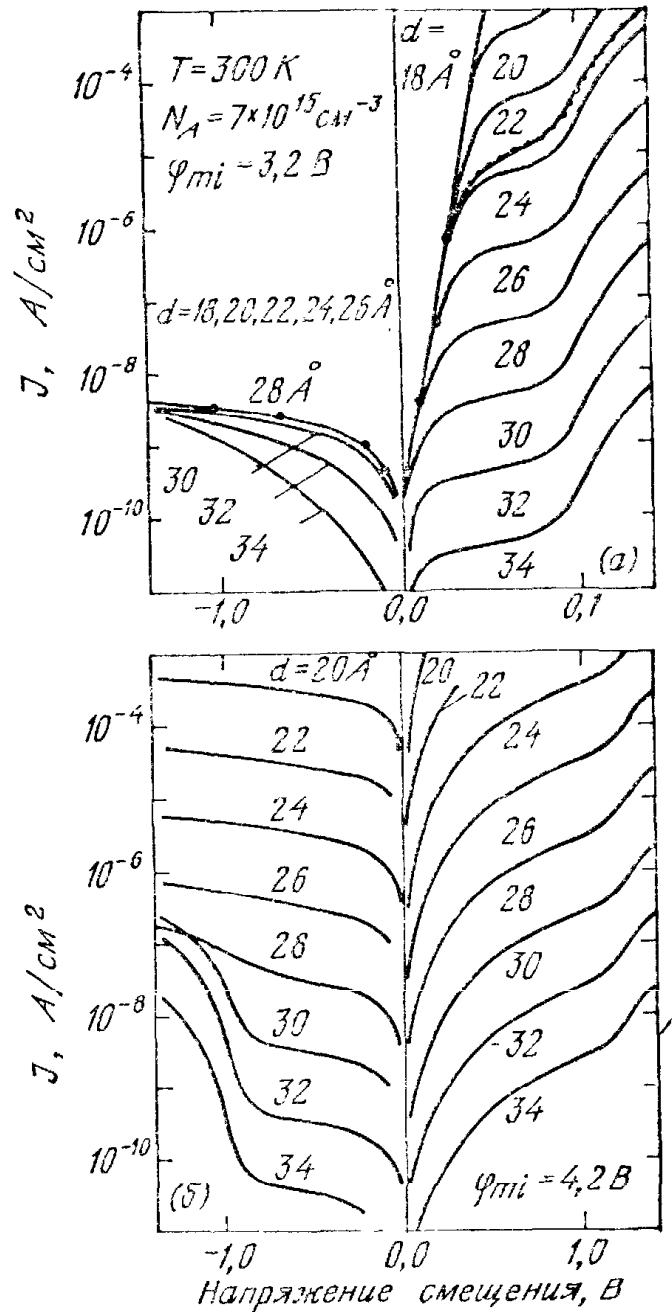
a — низкий барьер в системе металл — диэлектрик; *b* — высокий барьер в системе металл — диэлектрик.

в элементах памяти (например, МНОП-приборы), в преобразователях энергии (например, солнечные батареи) и в СВЧ-приборах в качестве перехода эмиттер—база туннельного транзистора.

Одним из важнейших параметров такого диода является высота барьера на границе металл—диэлектрик, которая существенно влияет на вольт-амперные характеристики [45, 46]. На рис. 24 приведены диаграммы энергетических зон туннельных МДП-диодов, находящихся в состоянии теплового равновесия. Диаграммы соответствуют подложке *p*-типа и двум разным высотам барьера в системе металл—диэлектрик. При малой высоте барьера ($\varphi_{mi} = 3,2$ В, что соответствует системе Al—SiO₂) поверхность дырочного кремния инвертирована. При большой высоте барьера ($\varphi_{mi} = 4,2$ В, что соответствует системе Au—SiO₂) на поверхности кремния накапливаются дырки. Туннельный ток состоит из двух основных компонент: тока из зоны проводимости в металл (J_{ct}) и тока из валентной зоны в металл (J_{vt}). Обе эти компоненты определяются выражениями, подобными формуле (44). На рис. 25 приведены теоретические вольт-амперные характеристики диодов с малой и большой высотой барьера. Если высота барьера невелика (рис. 25, *a*), то при малых прямых и обратных смещениях главным является ток неосновных носителей (электронов) J_{ct} , поскольку уровень Ферми проходит ближе к краю зоны проводимости. При увеличении прямого смещения (положительного напряжения на полупроводнике) ток монотонно возрастает. При фиксированном смещении ток резко возрастает с уменьшением толщины диэлектрика в результате того, что ограничивается вероятностью туннелирования (выражение (46)), которая экспоненциально зависит от толщины диэлектрика. При обратном смещении ток практически не зависит от толщины диэлектрика до тех пор, пока $d < < 30$ Å, поскольку в этом случае ток ограничивается скоростью генерации неосновных носителей (электронов) в объеме полупроводника, т. е. он аналогичен току насыщения обратносмещенного *p*—*n*-перехода. На рис. 25, *a* приведены также экспериментальные

Рис. 25. Вольт-амперные характеристики туннельных МДП-диодов [45].

a — с низким барьером; *б* — с высоким барьером.



результаты (они представлены черными точками), полученные при $d = 23,5 \text{ \AA}$. Отметим хорошее соответствие теоретических данных экспериментальным, а также то, что полученные вольт-амперные характеристики подобны вольт-амперным характеристикам $p-n$ -перехода.

Если высота барьера большая (рис. 25, б), то при прямых смещениях главным является ток основных носителей (дырок), туннелирующих из валентной зоны в металл, вследствие чего величина тока экспоненциально возрастает с уменьшением толщины диэлектрика. Однако в этом случае и при обратных смещениях не должно быть такой независимости тока от толщины диэлектрика, которая проиллюстрирована рис. 25, а. Вместо этого ток резко возрастает с уменьшением толщины диэлектрика, поскольку ток основных носителей в обоих направлениях ограни-

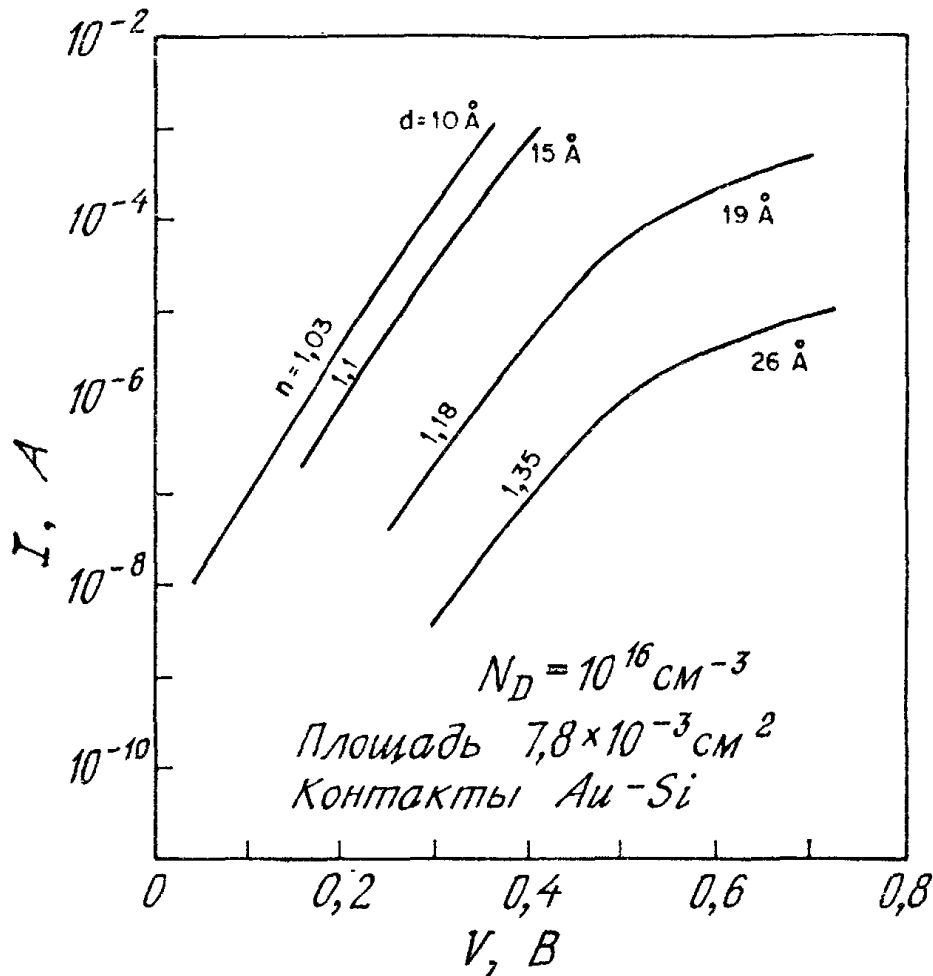


Рис. 26. Экспериментальные вольт-амперные характеристики туннельных МДП-диодов с различной толщиной слоя окисла [47].

чивается вероятностью туннелирования, а не скоростью генерации носителей.

Подставив формулу (46а) в выражение (45) и оценив интеграл по энергиям, получим соотношение [42, 47]

$$J = A^* T^2 e^{(-\alpha_T \Phi_T^{1/2} d)} e^{-q\Phi_B/kT} (e^{qV/kT} - 1), \quad (50)$$

где $A^* = 4\pi m_i^* q k^2 h^{-3}$ — эффективная постоянная Ричардсона, Φ_B — высота барьера. Выражение (50) идентично обычному выражению для термоэлектронной эмиссии в случае барьера Шоттки, за исключением множителя $\exp(-\alpha_T \Phi_T^{1/2} d)$, который представляет собой вероятность туннелирования. Отсюда следует, что при $\Phi_T \simeq 1$ В и $d > 50$ Å вероятность туннелирования составляет $\sim e^{-50} \approx 10^{-22}$ и туннельным током действительно можно пренебречь. При уменьшении d и (или) Φ_T ток быстро возрастает вплоть до величины тока термоэлектронной эмиссии. На рис. 26 приведены прямые ветви вольт-амперных характеристик четырех туннельных диодов типа Au—SiO₂—Si. При $d = 10$ Å зависимость тока от напряжения совпадает с обычной зависимостью для диода

Шоттки с фактором идеальности $n \approx 1$. При увеличении толщины диэлектрика ток резко уменьшается, а фактор идеальности начинает отличаться от 1.

9.5. ПЕРЕКЛЮЧАЮЩИЙ МДП-ДИОД

Переключающий МДП-диод представляет собой четырехслойную структуру (рис. 27, а). Этот диод, так же как и диод Шокли (гл. 4), был предложен для реализации S-образной вольт-амперной характеристики (рис. 27, б [48]). Когда к прибору прикладывается отрицательное напряжение (p^+ -подложка при этом заземлена), реализуется вольт-амперная характеристика с большим значением импеданса (прибор находится в состоянии «выключено»). При достаточно большом напряжении (напряжении переключения V_S) прибор резко переключается в состояние с большим током и низким напряжением (в состоянии «включено»). Переключение

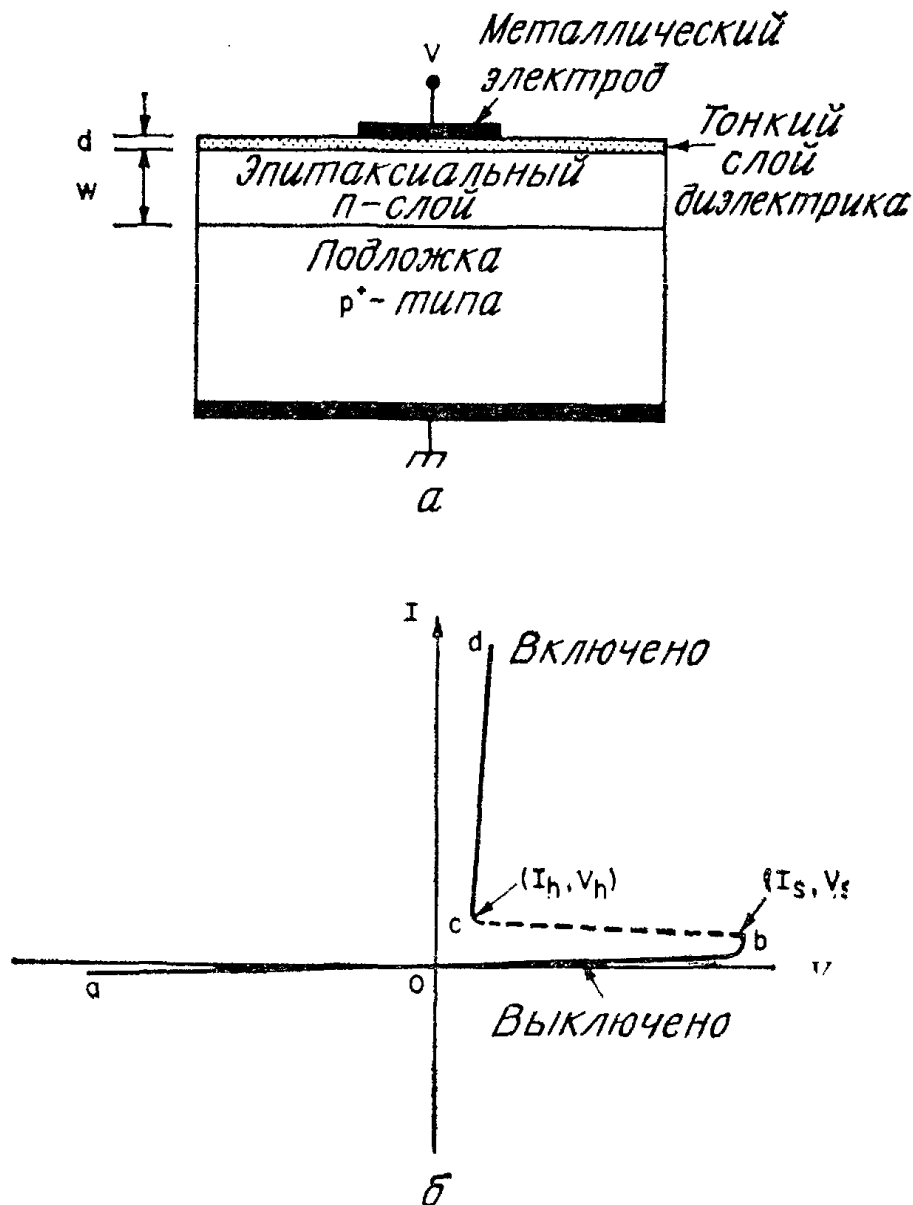


Рис. 27. Переключающий МДП-диод (а) и его S-образная вольт-амперная характеристика (б).

происходит либо благодаря расширению приповерхностного обедненного слоя вплоть до границы $p^+ - n$ -перехода (прокол n -слоя), либо благодаря лавинному пробоею вблизи поверхности n -слоя [49]. К достоинствам переключающего МДП-диода относятся совпадение уровня токов и напряжений с уровнем, который используется в цепях цифровой техники, высокая скорость переключения (1 нс и меньше) и высокая чувствительность напряжения переключения V_S к оптической или электрической инжекции. Впервые такой диод был изготовлен из кремния с использованием SiO_2 в качестве туннельного диэлектрика. Эффект переключения наблюдался также при использовании других диэлектриков (например, Si_3N_4) и толстых слоев поликристаллического кремния. Было показано, что, так же как в тиристоре (гл. 4), инжекция носителей в n -слой через третий контакт контролирует величину напряжения переключения.

Для того чтобы понять эффект переключения, рассмотрим энергетические диаграммы, приведенные на рис. 28 [50]. На рис. 28, а показана диаграмма идеального диода в состоянии теплового равновесия, а на рис. 28, б — аналогичная диаграмма для случая, когда на металлический электрод подано положительное напряжение. Если проводимость туннельного диэлектрика гораздо выше, чем проводимость обратносмещенного $p^+ - n$ -перехода, то большая часть приложенного напряжения падает на $p^+ - n$ -переходе. При этом вольт-амперная характеристика прибора практически та же, что и для обратносмещенного $p^+ - n$ -перехода, и определяется генерацией носителей в обедненном слое:

$$J_s = \frac{qn_i W_j}{2\tau_g} \simeq \frac{n_i}{\tau_g} \left(\frac{q\varphi_s}{2N_D} \right)^{1/2} (V + V_{bi})^{1/2}, \quad (51)$$

где n_i — собственная концентрация носителей, τ_g — время жизни неосновных носителей, W_j — ширина обедненного слоя, N_D — концентрация доноров в n -слое, V_{bi} — контактная разность потенциалов, V — приложенное напряжение. Поведение этого тока иллюстрируется участком кривой 0а вольт-амперной характеристики, приведенной на рис. 27, б.

При увеличении отрицательного напряжения на электроде (рис. 28, в) свободные электроны вытягиваются из эпитаксиального слоя через $p^+ - n$ -переход и обедненный слой под электродом расширяется. При генерации электронно-дырочных пар в этом слое электроны вытягиваются из прибора через $p^+ - n$ -переход, а дырки подтягиваются к границе раздела $\text{SiO}_2 - \text{Si}$. В случае толстого слоя окисла ($d \geq 50 \text{ \AA}$) дырки накапливаются у границы раздела и инвертируют поверхность. Если же слой окисла достаточно тонкий, дырки туннелируют через него. Тогда при увеличении напряжения происходит обеднение поверхности n -слоя

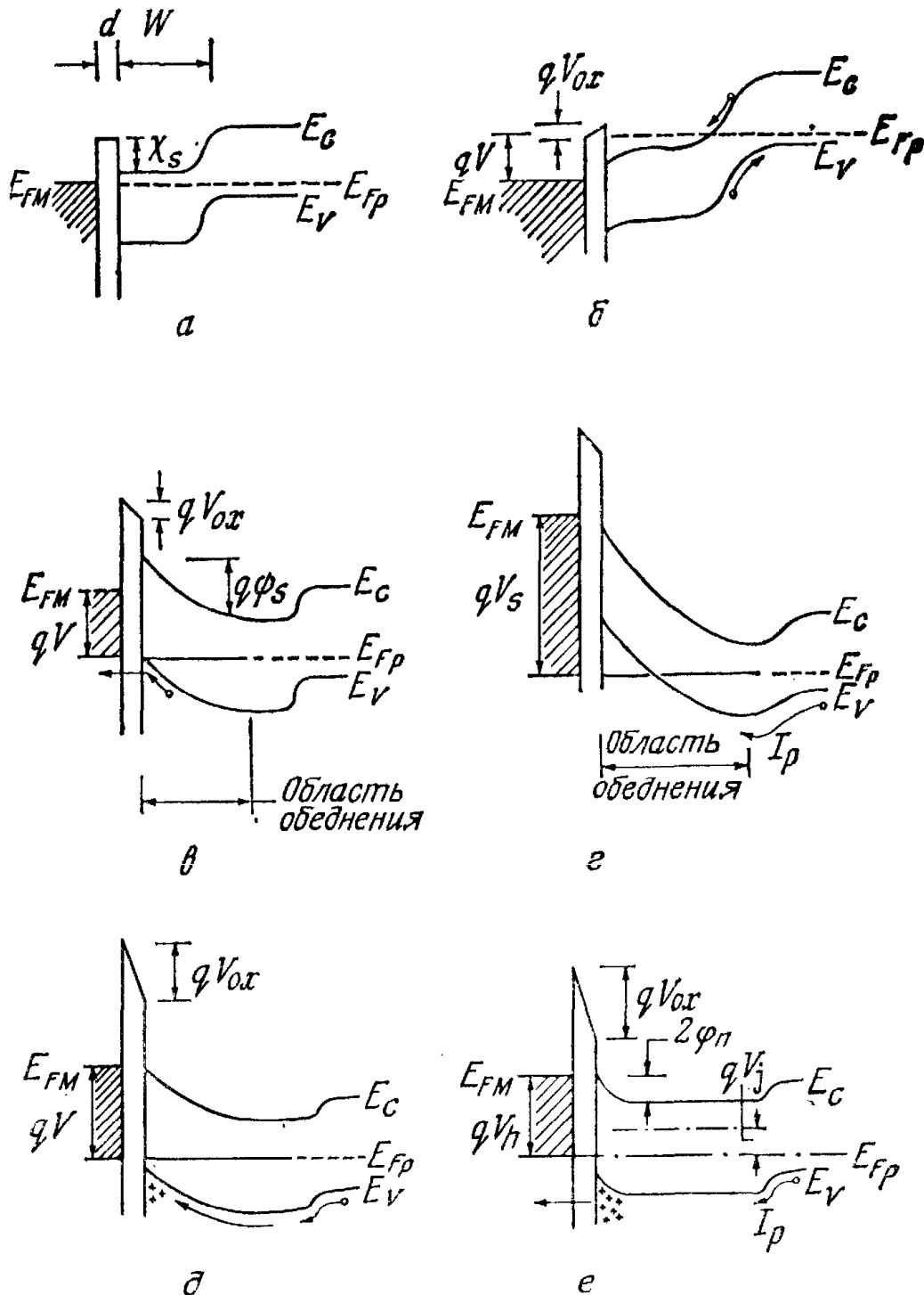


Рис. 28. Энергетические диаграммы переключающего МДП-диода при различных напряжениях [50].

вместо инверсии. При этом генерационный ток определяется выражением, аналогичным выражению (51), но в последнем множитель $(V + V_{bi})^{1/2}$ заменяется на $(\psi_s)^{1/2}$, где ψ_s — величина поверхностного потенциала. Этот ток также слабо зависит от напряжения (участок кривой Ob на рис. 27, б).

Если приложено достаточно большое напряжение, например $V = V_s$ (напряжение переключения), обедненный слой под окис-

лом расширяется вплоть до $p^+—n$ -перехода (рис. 28, *з*), и происходит прокол перехода. Напряжение прокола равно

$$V_s \approx \frac{qN_D (W - W_{J_0})^2}{2\epsilon_s}, \quad (52)$$

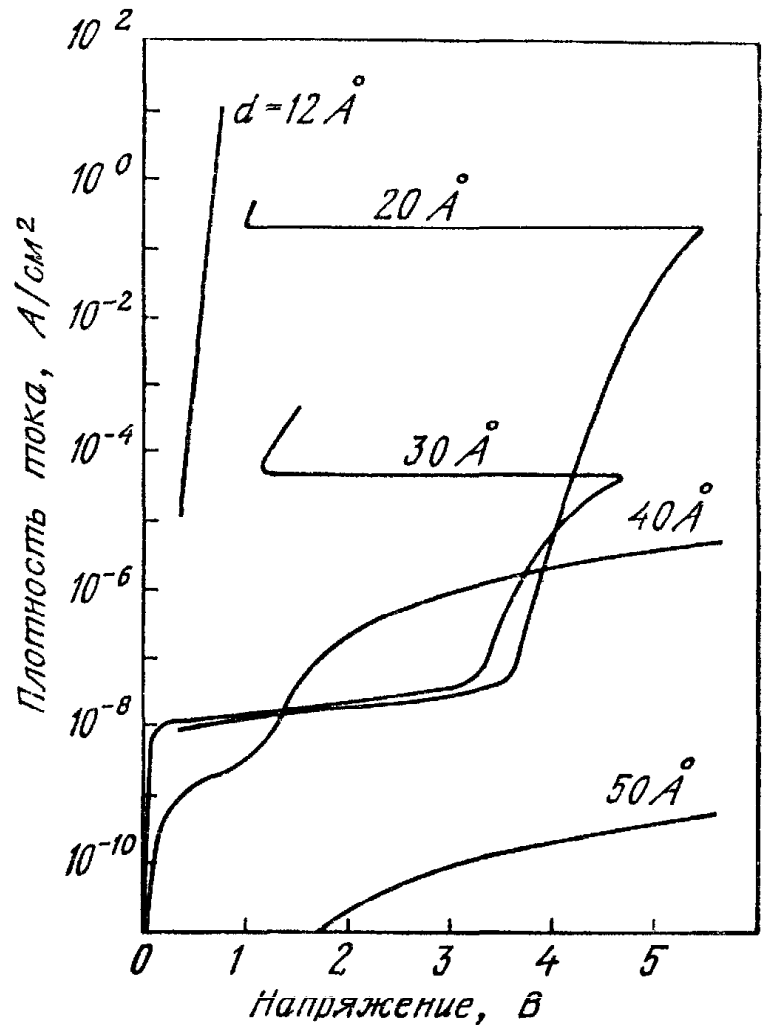
где W — толщина эпитаксиального слоя n -типа, W_{J_0} — толщина обедненного слоя $p^+—n$ -перехода при нулевом смещении. При проколе уровень инжекции дырок из перехода возрастает экспоненциально с увеличением смещения перехода. В первый момент после прокола величина поля в окисле оказывается недостаточной для того, чтобы большой поток инжектированных дырок мог туннелировать через диэлектрик, поскольку при этом ток ограничен туннелированием. Следовательно, большое увеличение напряжения приведет к накоплению дырок на границе раздела $Si—SiO_2$. Это накопление переведет поверхность из глубокого обеднения в область инверсии (рис. 28, *д*); при этом ψ_s и напряжение на приборе уменьшаются. Падение напряжения на окисле увеличится, вследствие чего возрастет ток электронов, туннелирующих из металла. Этот ток увеличит смещение на $p^+—n$ -переходе. Данный механизм обеспечивает положительную обратную связь и приводит к появлению на вольт-амперной характеристике участка с отрицательным дифференциальным сопротивлением (участок кривой *bc* на рис. 27, *б*). Из энергетической диаграммы прибора при остаточном напряжении (рис. 28, *е*) видно, что остаточное напряжение равно

$$V_h = V_{ox} + \psi_s + V_J, \quad (53)$$

где V_{ox} — падение напряжения на окисле, ψ_s — поверхностный потенциал, V_J — напряжение на $p^+—n$ -переходе. Приведенный анализ касается случая переключения в условиях прокола n -слоя. Тот же механизм положительной обратной связи ответствен за переключение в условиях лавинного пробоя. При этом напряжение переключения, определяемое выражением (52), следует заменить на напряжение лавинного пробоя n -слоя вблизи границы раздела $Si—SiO_2$.

В соответствии с проведенным анализом основным параметром для эффекта переключения служит толщина слоя окисла. Этот вывод подтверждают вольт-амперные характеристики, приведенные на рис. 29 [49]. Для слоев окисла большой толщины ($d \geq 50 \text{ \AA}$) сопротивление туннелирования слишком высоко для переключения диода при низких напряжениях ($V \ll 5 \text{ В}$). В случае очень тонкого слоя окисла ($d < 15 \text{ \AA}$) переход может полностью сместиться в прямом направлении еще до того, как произойдет глубокое обеднение поверхности; следовательно, вольт-амперная

Рис. 29. Рассчитанные вольт-амперные характеристики переключающего МДП-диода при различной толщине слоя окисла. Параметры прибора: $W = 10$ мкм, $N_D = 10^{14}$ см $^{-3}$ и $\tau_g = 3,5 \cdot 10^{-5}$ с [49].



характеристика прибора будет совпадать с характеристикой $p-n$ -перехода. Эффект переключения наблюдался только на слоях окисла промежуточной толщины ($15 \text{ \AA} \leq d \leq 40 \text{ \AA}$).

9.6. ТУННЕЛЬНЫЙ МДМ-ДИОД

Туннельный диод на структуре металл—диэлектрик—металл (МДМ) представляет собой тонкопленочный прибор, в котором электроны из первого металла могут туннелировать через пленку диэлектрика и коллектироваться вторым металлом. На рис. 30 (вставка) приведены основные энергетические диаграммы симметричного МДМ-диода. Поскольку все приложенное напряжение падает на диэлектрике, плотность туннельного тока через него в соответствии с выражением (44) равна

$$J = \frac{4\pi q m^*}{h^3} \iint T_t [F(E) - F(E + qV)] dE_{\perp} dE. \quad (54)$$

При абсолютном нуле температуры выражение (54) принимает более простой вид [51]

$$J = J_0 [\bar{\phi} \exp(-A\sqrt{\bar{\phi}}) - (\bar{\phi} + V) \exp(-A\sqrt{\bar{\phi} + V})], \quad (55)$$

где

$$J_0 \equiv q^2 / [2\pi h (\Delta d)^2], \quad A \equiv 4\pi (\Delta d) \sqrt{2mq/h},$$

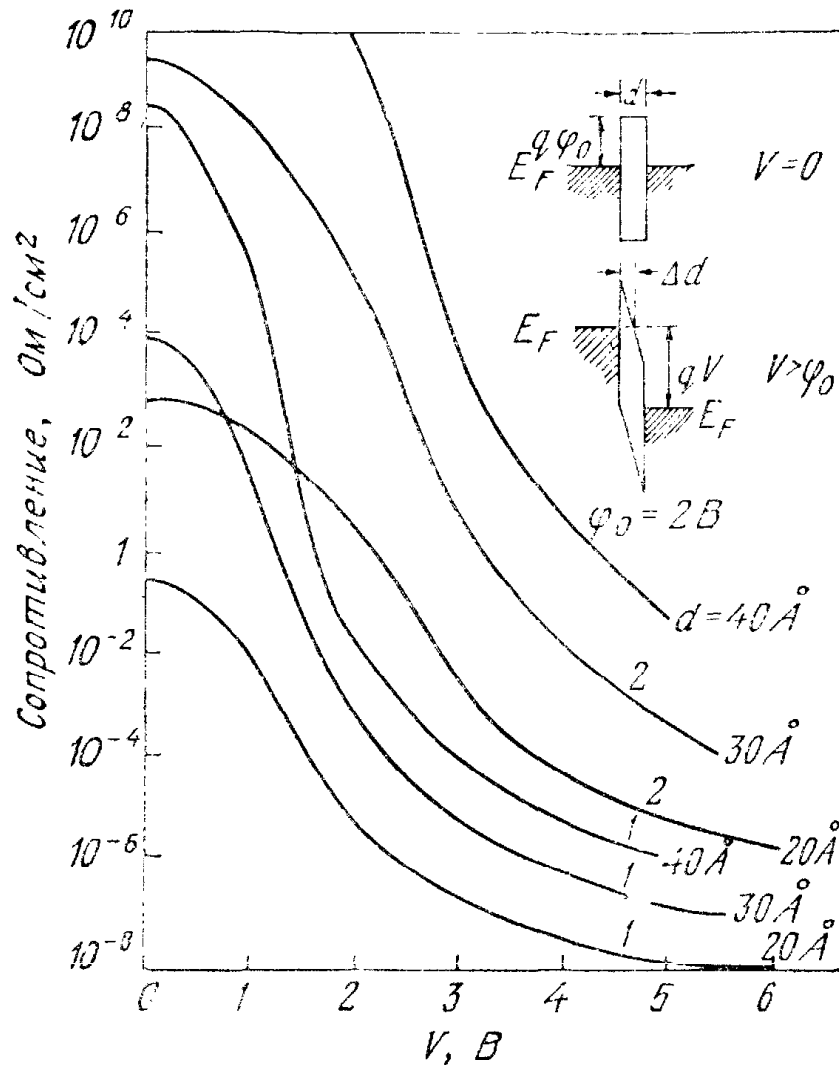


Рис. 30. Зависимости туннельного сопротивления симметричной МДМ-структуры от напряжения. На вставке показаны зонные диаграммы при $V = 0$ и при $V > \varphi_0$ [51].

а $\bar{\varphi}$ — средняя высота барьера над уровнем Ферми. Выражение для плотности тока (55) можно интерпретировать как результат суперпозиции тока с плотностью $J_0 \bar{\varphi} \exp(-A \sqrt{\bar{\varphi}})$, текущего от электрода 1 к электроду 2, и тока с плотностью $J_0 (\bar{\varphi} + V) \exp(-A \sqrt{\bar{\varphi} + V})$, текущего от электрода 2 к электроду 1.

Ниже мы используем выражение (55) для описания свойств идеальной МДМ-структуры (рис. 30). Под идеальностью здесь понимается пренебрежение влиянием температуры, сил изображения и проникновения поля в глубь металлических электродов. При $0 \leq V \leq \varphi_0$, $\Delta d = d$ и $\bar{\varphi} = \varphi_0 - V/2$ плотность тока равна

$$J = J_0 [(\varphi_0 - V/2) \exp(-A \sqrt{\varphi_0 - V/2}) - (\varphi_0 + V/2) \exp(-A \sqrt{\varphi_0 + V/2})]. \quad (56)$$

При большем напряжении ($V > \varphi_0$) имеем $\Delta d = d\varphi_0/V$ и $\bar{\varphi} = \varphi_0/2$. Тогда плотность тока равна

$$J = \frac{q^2 \mathcal{E}^2}{4\pi h \varphi_0} [\exp(-\mathcal{E}_0/\mathcal{E}) - (1 + 2V/\varphi_0) \exp(-\mathcal{E}_0 \sqrt{1 + 2V/\varphi_0}/\mathcal{E})], \quad (57)$$

где $\mathcal{E} = V/d$ — поле в диэлектрике, а $\mathcal{E}_0 \equiv \frac{8}{3} V \sqrt{\pi q} (\varphi_0)^{3/2}$. При очень больших напряжениях, таких, что $V > (\varphi_0 + E_F/q)$, вторым слагаемым в квадратных скобках в выражении (57) можно пренебречь, и тогда мы получим хорошо известную формулу Фаулера—Нордгейма. На рис. 30 приведены результаты расчета туннельного сопротивления V/J при разных высотах барьера и толщинах слоя диэлектрика. Отметим, что туннельное сопротивление быстро уменьшается с увеличением приложенного напряжения.

Для идеальной асимметричной МДМ-структуры (рис. 31, вставка) в диапазоне низких напряжений $0 < V < \varphi_1$ величины Δd и $\bar{\varphi}$ равны d и $(\varphi_1 + \varphi_2 - V)/2$ и не зависят от полярности напряжения. При этом не зависит от полярности и вольт-амперная характеристика. При больших напряжениях ($V > \varphi_2$) средняя высота барьера $\bar{\varphi}$ и эффективная длина туннелирования Δd начинают зависеть от полярности. Поэтому значения тока при различных полярностях оказываются разными. На рис. 31 приведены зависимости туннельного сопротивления от напряжения при $d = 20, 30, 40 \text{ \AA}$, $\varphi_1 = 1 \text{ В}$ и $\varphi_2 = 2 \text{ В}$. Экспериментальные результаты, полученные на структурах $\text{Al—Al}_2\text{O}_3\text{—Al}$ при $d = 30 \text{ \AA}$, $\varphi_1 = 1,6 \text{ В}$ и $\varphi_2 = 2,5 \text{ В}$, хорошо согласуются с предсказаниями теории, основанными на выражениях (56) и (57). Однако эффективная площадь, использованная для сравнения с экспериментальными данными, была принята равной 1 % площади электрода [52]. Совсем иная ситуация наблюдается для туннельных МДМ-диодов на структурах $\text{Al—SiO}_2\text{—SiO}_2$, где эффективная площадь обычно равна площади алюминиевого электрода. Относительно малые значения площади туннелирования в МДМ-диоде могут быть объяснены статистической природой образования диэлектрических пленок на металлической подложке [53]. Туннельный ток определяется только самыми тонкими участками диэлектрической пленки. Из-за статистических флуктуаций толщины пленки емкость МДМ-структуры всегда оказывается больше той, которая вычисляется исходя из средней толщины слоя диэлектрика [54]. Кроме того, на величину емкости может влиять распределение потенциала на металлических электродах. В результате проникновения электрического поля в глубь металла

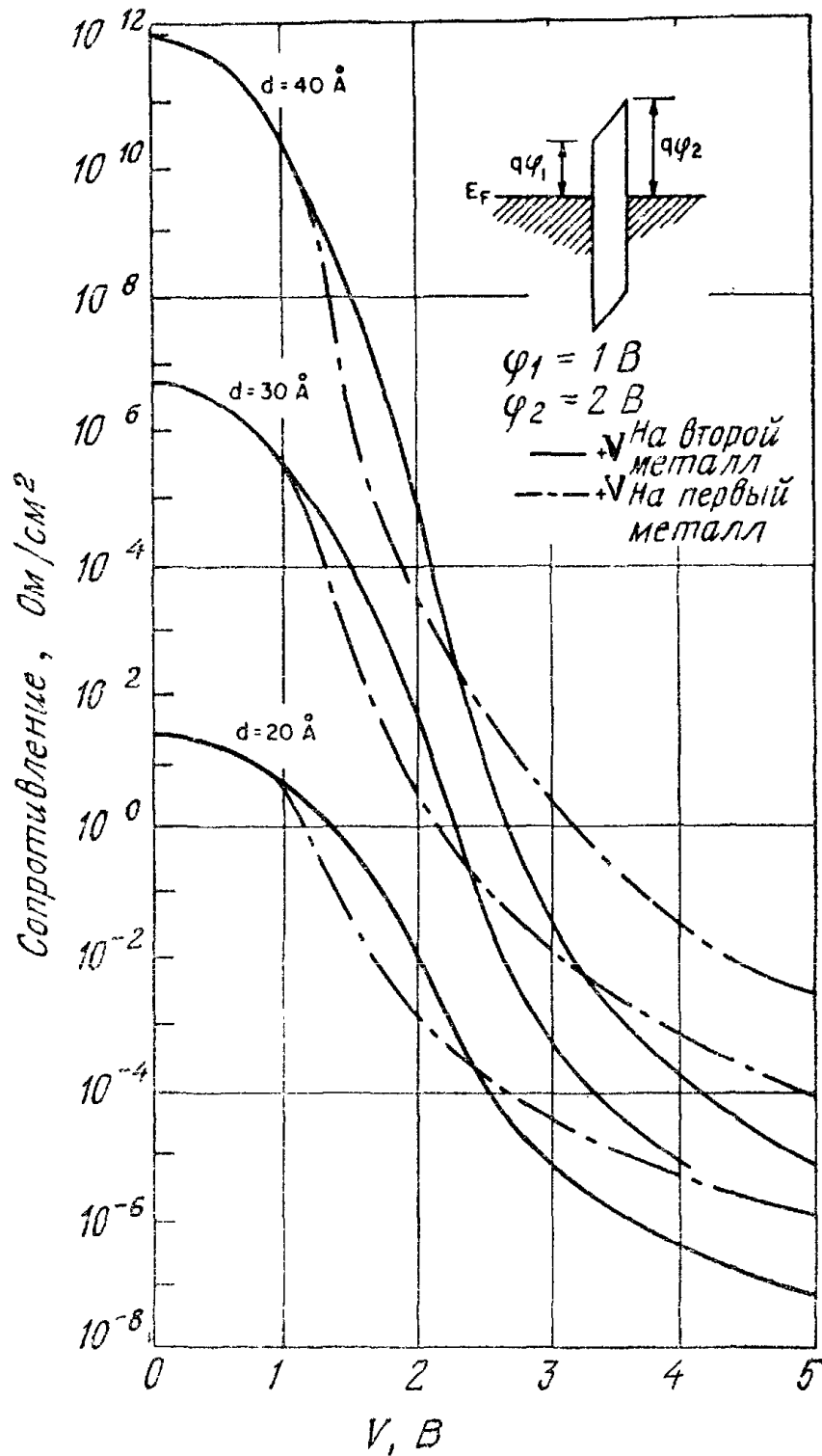


Рис. 31. Зависимости туннельного сопротивления асимметричной МДМ-структуры от напряжения. На вставке показана зонная диаграмма при $V = 0$ [51].

полная емкость фактически равняется двум последовательно соединенным емкостям [55]:

$$1/C = d/\epsilon_i + 2,3\lambda_m/\epsilon_m, \quad (58)$$

где ϵ_i и ϵ_m — диэлектрические проницаемости диэлектрика и металла соответственно, а λ_m — характерная глубина проникновения электрического поля в металл ($\sim 0,5 \text{ \AA}$ для благородных металлов). Из выражения (58) видно, что емкость, обусловленная

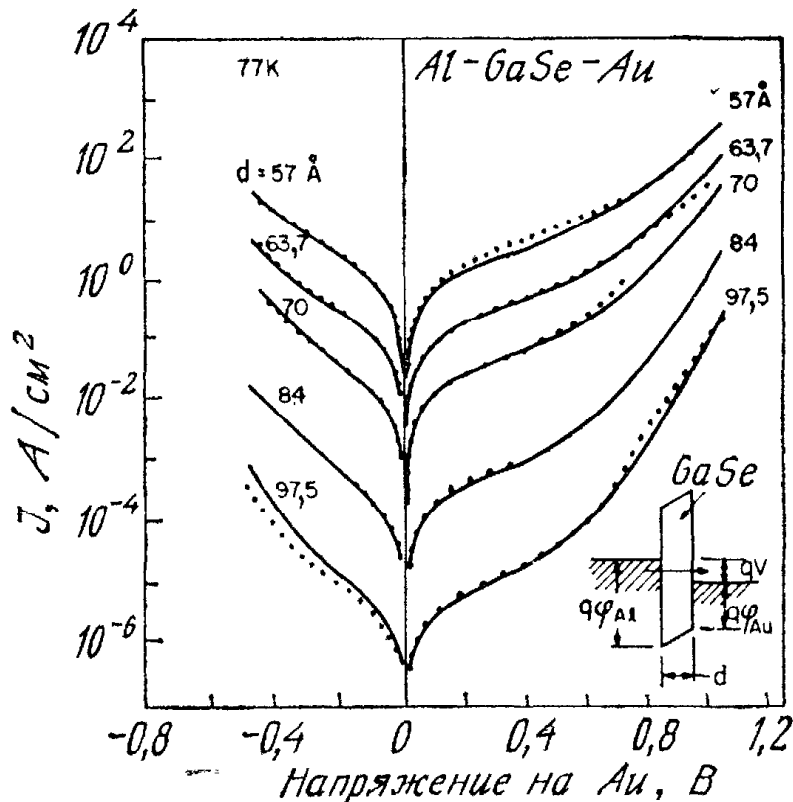


Рис. 32. Зависимость плотности тока от приложенного напряжения для нескольких структур Al — GaSe — Au. Точками представлены экспериментальные данные, кривыми — результаты вычислений. На вставке показана зонная диаграмма при приложенном напряжении [56].

вторым слагаемым, составляет 10^6 пФ/см². Полная величина емкости единицы площади при $d \rightarrow 5$ Å окажется гораздо меньше ϵ_i/d .

Туннельные МДМ-диоды использовались для изучения связи между энергией и импульсом в запрещенной зоне монокристаллических диэлектриков, например GaSe ($E_g = 2,0$ эВ, $\epsilon_s/\epsilon_0 = 8$ и плотность носителей $p \approx 3 \cdot 10^{14}$ см⁻³ при 300 К). GaSe имеет слоистую кристаллическую структуру, что упрощает изготовление монокристаллических чешуек требуемой толщины (< 100 Å). Туннельную МДМ-структуру создают путем размещения монокристаллического образца между двумя металлическими электродами, например Al—GaSe—Au (барьер на границе Al—GaSe равен 1,08 В, а на границе Au—GaSe 0,52 В). Зонная диаграмма такой структуры приведена на рис. 32 (вставка). Здесь приведены также результаты измерений на структурах Al—GaSe—Au различной толщины. Используя одну из $J(V)$ -кривых, можно с помощью формул (44) и (54) вычислить зависимость импульса от энергии $k(E)$ для GaSe. Определение зависимости $k(E)$ позволяет без использования подгоночных параметров вычислить значения туннельного тока при любой толщине слоя диэлектрика. Таким способом были рассчитаны теоретические вольт-амперные характеристики, приведенные на рис. 32, из которого видно идеальное

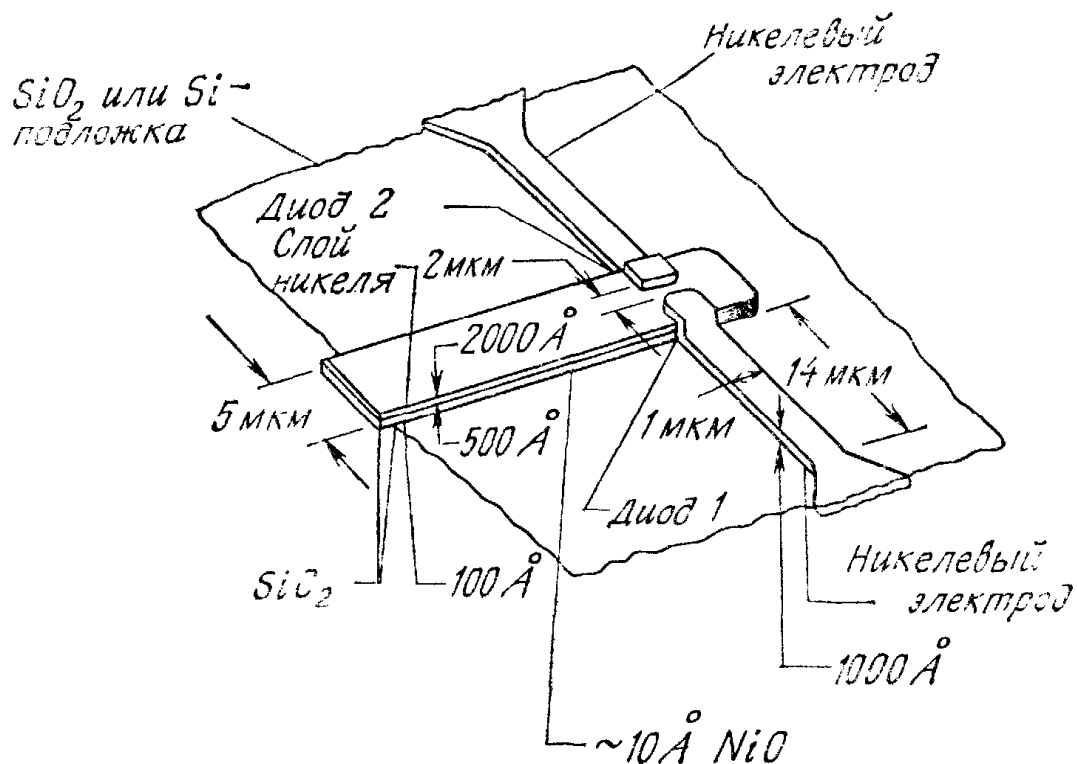


Рис. 33. Схема краевого МОМ-диода [57].

совпадение между теоретическими и экспериментальными данными для всех значений напряжения и толщины.

Стабильным и воспроизводимым МДМ-диодом является краевой МОМ-диод (металл—окисел—металл), который отличается строго фиксированной площадью туннелирования [57]. На рис. 33 приведено схематическое изображение такого прибора. Перекрытие между металлической (Ni) полоской шириной 1 мкм и оксидированным краем металлической (Ni полоски, толщина которой близка к 100 \AA , приводит к площади туннелирования $\sim 10^{-10} \text{ см}^2$. Толщина слоя окисла (NiO) составляет 10 \AA . Коэффициент кривизны γ (уравнение (42)) для краевого МОМ-диода со структурой Ni—NiO—Ni лежит в диапазоне 1—10. Этот прибор можно использовать в качестве широкополосного детектора и смесителя, поскольку нелинейность его вольт-амперной характеристики постоянна в диапазоне от звуковых частот и вплоть до частот ближнего инфракрасного излучения.

9.7. ТУННЕЛЬНЫЙ ТРАНЗИСТОР

За последние годы было предпринято много попыток изобрести или сконструировать новые твердотельные приборы с лучшими рабочими характеристиками по сравнению с биполярными или полевыми транзисторами. К числу наиболее интересных предложений относятся туннельные транзисторы. Первый туннельный транзистор был создан Мидом в 1960 г. [58]. На рис. 34, а по-

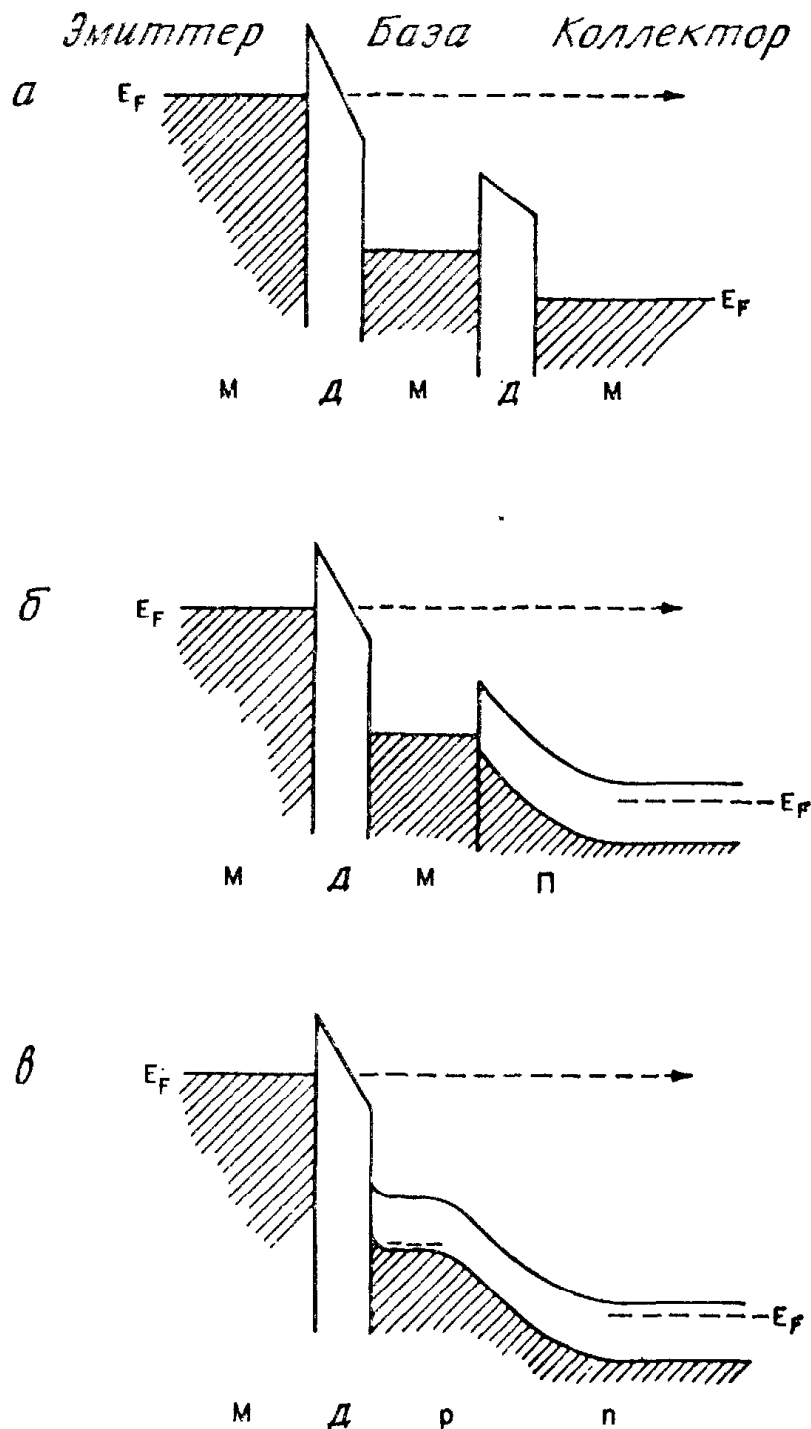


Рис. 34. Туннельный МДМДМ-транзистор (а), туннельный МДМП-транзистор (б) и туннельный МД (р — n)-транзистор (в) [58, 60, 61].

казана энергетическая диаграмма прибора со структурой металл—диэлектрик—металл—диэлектрик—металл (МДМДМ), в котором ток течет через первый диэлектрический слой благодаря туннелированию. После туннелирования из эмиттера в базу энергия электронов не менее чем на несколько kT превышает энергию Ферми в базе; эти электроны называются горячими, поскольку они не находятся в тепловом равновесии с решеткой. Такой транзистор называют также транзистором на горячих электронах [59]. Спратт, Шварц и Кейн установили, что коэффициент усиления по току такой структуры можно существенно

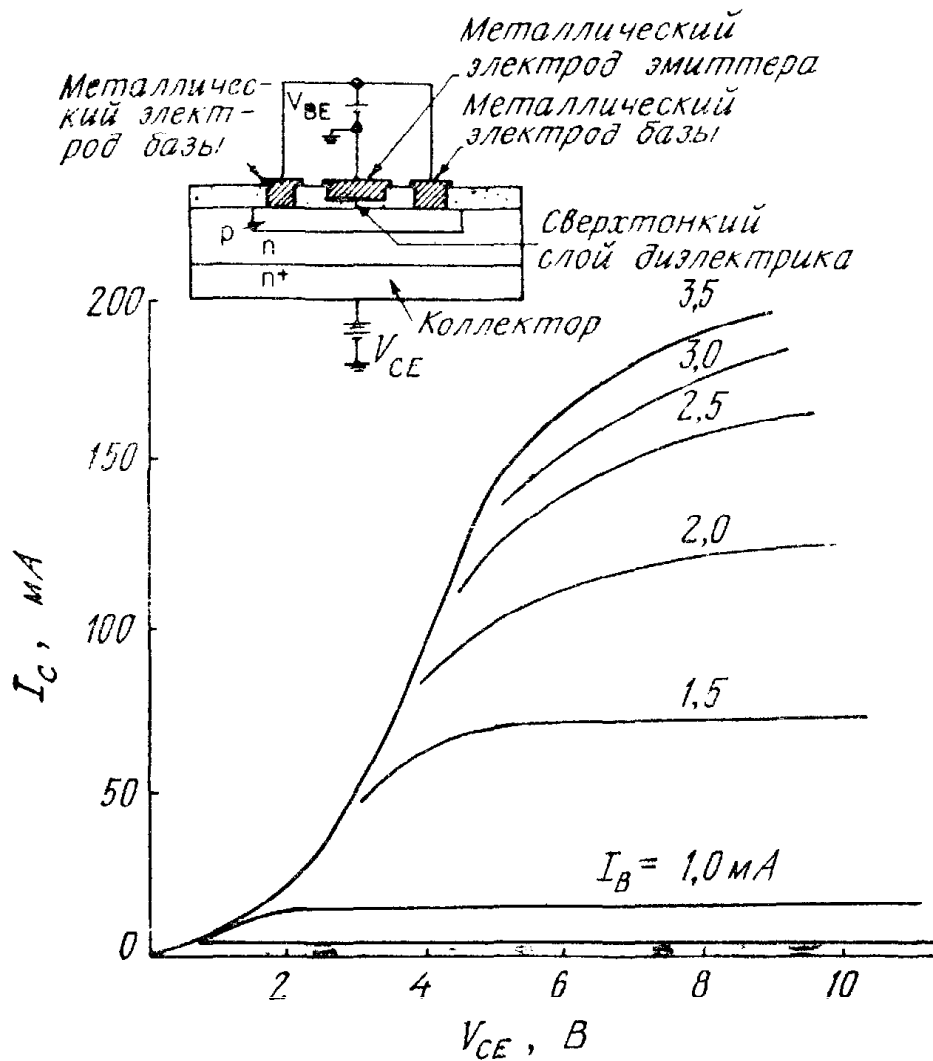


Рис. 35. Вольт-амперные характеристики коллектора туннельного МД ($p-n$)-транзистора. На вставке показана диаграмма поперечного сечения прибора [61].

улучшить, заменив коллектирующий диэлектрик полупроводником, образующим барьер Шоттки со вторым металлом (рис. 34, б) [60]. Максимальная рабочая частота для МДМП-структуры меньше, чем для биполярного транзистора в основном из-за того, что в такой структуре время зарядки эмиттера больше (что обусловлено большой емкостью эмиттера), а коэффициент усиления по току в схеме с общей базой меньше (что обусловлено интенсивным рассеянием горячих электронов в области базы).

На рис. 34, в приведена энергетическая диаграмма туннельного транзистора [61], представляющего собой МДП-структуру, в которой электроны могут туннелировать из эмиттирующего металлического электрода в базу (слой p -типа, полученный диффузией). На рис. 35 (вставка) показана диаграмма поперечного сечения такой структуры. В качестве металлического электрода используется Al , нанесенный на тонкий слой SiO_2 толщиной около 20 \AA , термически выращенный в атмосфере сухого кислорода ($550^\circ C$ в течение 30 мин). Толщина базы p -типа составляет $\sim 1500 \text{ \AA}$, а концентрация акцепторов в ней $\sim 10^{18} \text{ см}^{-3}$. Концентрация до-

норов в эмиттерном слое n -типа равна $8 \cdot 10^{15} \text{ см}^{-3}$. На рис. 35 приведены также вольт-амперные характеристики коллектора. Коэффициент усиления по току в схеме с общим эмиттером при низких уровнях тока в базе может превышать 100.

В работе [62] предложен транзистор с туннельно-тонкой базой, состоящий из двух гетеропереходов. На рис. 36, а приведена зонная диаграмма такого транзистора в состоянии теплового равновесия. Для нормального функционирования этого транзистора необходимо выполнение трех условий: 1) разрывы зон ΔE_c и ΔE_v должны происходить в таком направлении и иметь такую величину, чтобы образовались яма для электронов и барьер для дырок (рис. 36, а); 2) n -область при всех условиях должна оставаться проводящей; 3) толщина базы и конфигурация энергетических зон должны быть такими, чтобы основным током

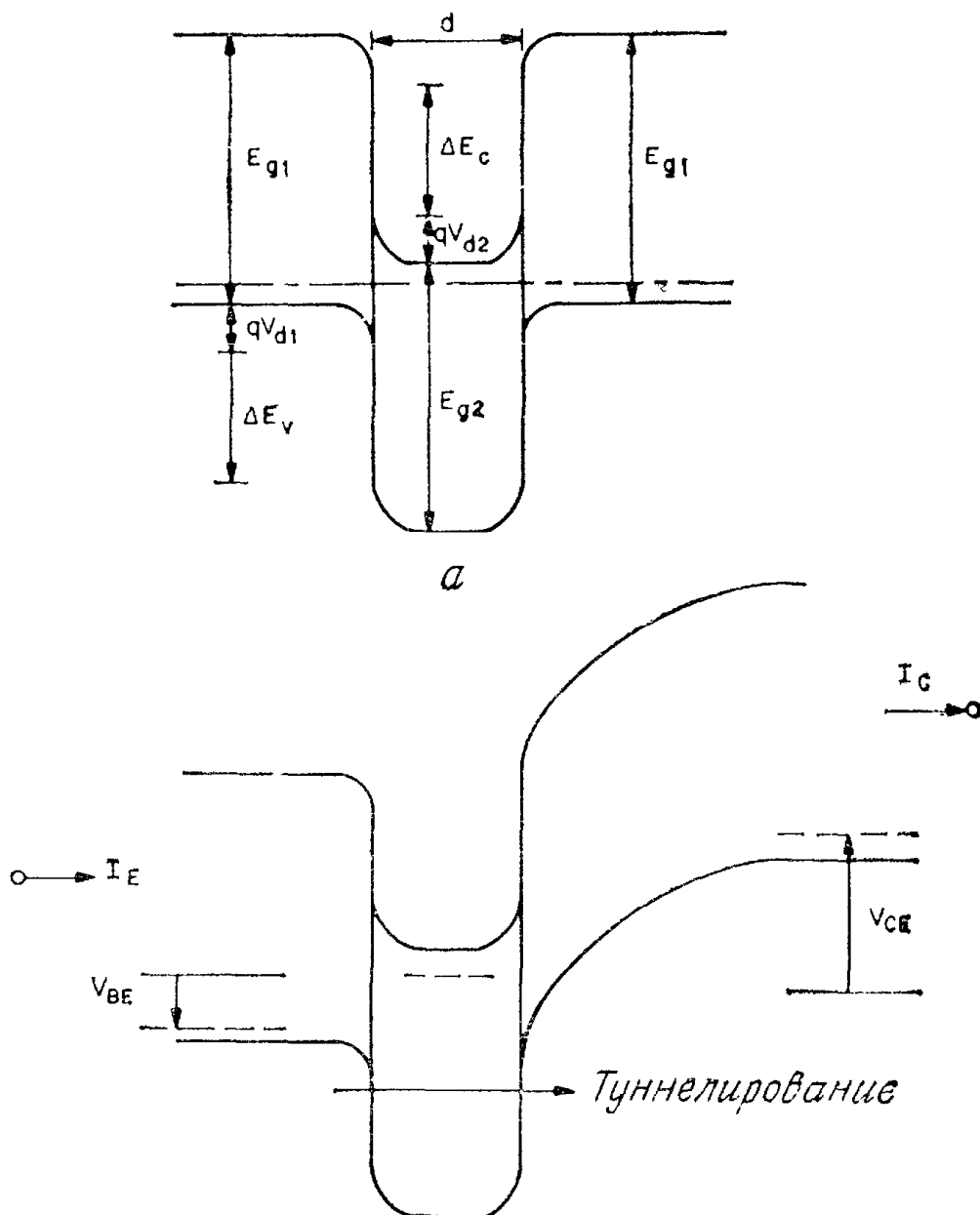


Рис. 36. Энергетические диаграммы транзистора с туннельно-тонкой базой в состоянии теплового равновесия (а) и при напряжении (б) [62].

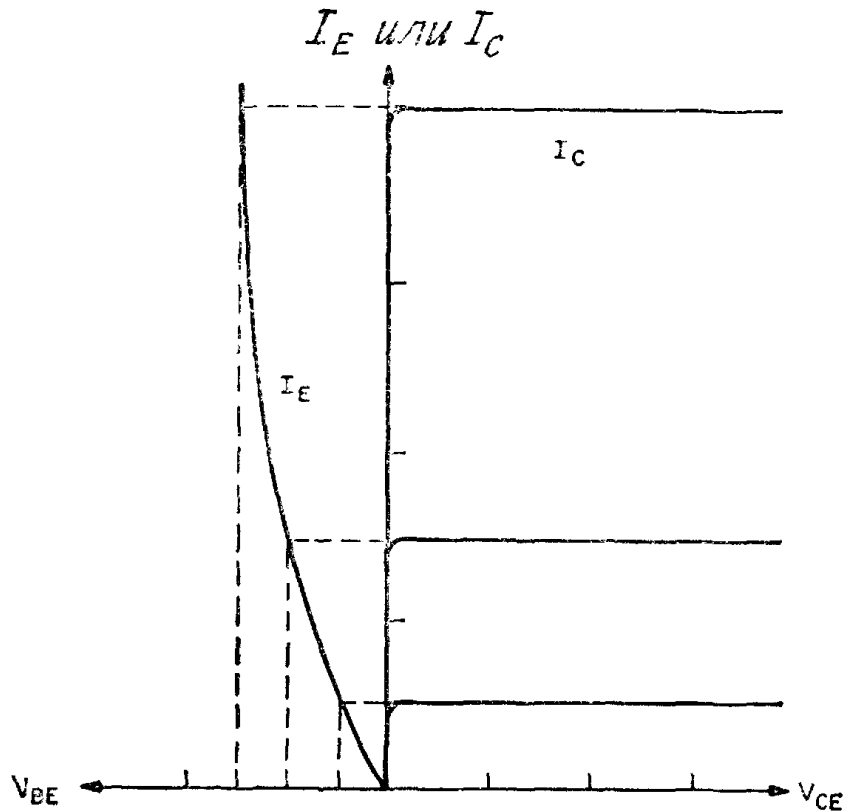


Рис. 37. Зависимость тока эмиттера и тока коллектора от напряжения на базе и коллекторе [62].

был туннельный. Этим условиям удовлетворяют, например, $\text{GaAs}_{0,5}\text{Sb}_{0,5}$ p -типа с $N_A \approx 10^{16} \text{ см}^{-3}$ в качестве материала эмиттера и коллектора и $\text{Ga}_{0,5}\text{In}_{0,5}\text{As}$ n -типа с $N_D \approx 10^{19} \text{ см}^{-3}$ при $d \approx 50 \text{ \AA}$ в качестве материала базы. Для этой пары $E_{g1} \approx \approx E_{g2} = 0,8 \text{ эВ}$ и $\Delta E_C = \Delta E_V = 0,5 \text{ эВ}$. Выращивание этих слоев можно осуществить методом молекулярно-лучевой эпитаксии.

На рис. 36, б приведена зонная диаграмма такого туннельного транзистора при подаче напряжения V_{CE} . Основным током здесь будет туннельный ток через базу, для которого можно записать $J_t \sim \exp(qV_{BE}/kT)$. На рис. 37 приведены зависимости тока эмиттера I_E и тока коллектора I_C от напряжения. Близость величин I_E и I_C означает, что эффективность эмиттера и коэффициент переноса по базе равны 1, а следовательно, транзистор имеет большой коэффициент усиления по току. Этот туннельный транзистор обладает тремя преимуществами перед обычным биполярным транзистором: 1) пренебрежимо малым временем пролета через базу, поскольку времена туннелирования всегда малы; 2) малой емкостью эмиттера, что обусловлено его слабым легированием; 3) низким сопротивлением базы, обусловленным ее сильным легированием. Можно ожидать, что благодаря этим достоинствам транзистор с туннельно-тонкой базой будет иметь лучшие характеристики в СВЧ-диапазоне.

ЛИТЕРАТУРА

1. Esaki L. New Phenomenon in Narrow Germanium $p - n$ Junctions, *Phys. Rev.*, **109**, 603 (1958); Long Journey into Tunneling, *Proc. IEEE*, **62**, 825 (1974); Discovery of the Tunnel Diode, *IEEE Trans. Electron Devices*, **ED-23**, 644 (1976).
2. Thornber K. K., McGill Thomas C., Mead C. A. The Tunneling Time of an Electron, *J. Appl. Phys.*, **38**, 2384 (1967).
3. Hall R. N. Tunnel Diodes, *IRE Trans. Electron Devices*, **ED-7**, 1 (1960).
4. Morgan J. V., Kane E. O. Observation of Direct Tunneling in Germanium, *Phys. Rev. Lett.*, **3**, 466 (1959).
5. Ландау Л. Д., Лифшиц Е. М. Квантовая механика, ГИФМЛ, Москва, 1963, стр. 211.
6. Kane E. O. Theory of Tunneling. *J. Appl. Phys.*, **32**, 83 (1961); Tunneling in InSb, *Phys. Chem. Solids*, **2**, 181 (1960).
7. Demassa T. A., Knott D. P. The Prediction of Tunnel Diode Voltage — Current Characteristics, *Solid State Electron.*, **13**, 131 (1970).
8. Kroemer H. The Einstein Relation for Degenerate Carrier Concentration, *IEEE Trans. Electron Devices*, **ED-25**, 850 (1978).
9. Meyerhofer D., Brown G. A., Sommers H. S., Jr. Degenerate Germanium I, Tunnel, Excess, and Thermal Current in Tunnel Diodes, *Phys. Rev.*, **126**, 1329 (1962).
10. Moll J. L. Physics of Semiconductors, McGraw-Hill, N. Y., 1964, p. 252.
11. Butcher P. N., Hulme K. F., Morgan J. R. Dependence of Peak Current Density on Acceptor Concentration in Germanium Tunnel Diodes, *Solid State Electron.*, **3**, 358 (1962).
12. Келдыш Л. В. Поведение неметаллических кристаллов в сильных электрических полях, *ЖЭТФ*, **6**, 763 (1958).
13. Chynoweth A. G., Feldmann W. L., Logan R. A. Excess Tunnel Current in Silicon Esaki Junctions, *Phys. Rev.*, **121**, 684 (1961).
14. Roy D. K. On the Prediction of Tunnel Diode I—V Characteristics, *Solid State Electron.*, **14**, 520 (1971).
15. Carr W. N. Reversible Degradation Effects in GaSb Tunnel Diodes, *Solid State Electron.*, **5**, 261 (1962).
16. Nanavati R. P., Morato De Andrade C. A. Excess Current in Gallium Arsenide Tunnel Diodes, *Proc. IEEE*, **52**, 869 (1964).
17. Holonyak N., Jr. Evidence of States in the Forbidden Gap of Degenerate GaAs and InP — Secondary Tunnel Current and Negative Resistance, *J. Appl. Phys.*, **31**, 130 (1960).
18. Franks V. M., Hulme K. F., Morgan J. R. An Alloy Process for Making High Current Density Silicon Tunnel Junction, *Solid State Electron.*, **8**, 343 (1965).
19. Kleinknecht H. P. Indium Arsenide Tunnel Diodes, *Solid State Electron.*, **2**, 133 (1961).
20. Burrus C. A. Indium Phosphide Esaki Diodes, *Solid State Electron.*, **3**, 357 (1962).
21. Yoshihito A., Konagai M., Sakai Y. Mixed Crystal Tunnel Diode, *Jpn. J. Appl. Phys.*, **12**, 480 (1973).
22. Minton R. M., Glicksman R. Theoretical and Experimental Analysis of Germanium Tunnel Diode Characteristics, *Solid State Electron.*, **7**, 491 (1964).
23. Logan R. A., Augustyniak W. M., Gilber J. F. Electron Bombardment Damage in Silicon Esaki Diodes, *J. Appl. Phys.*, **32**, 1201 (1961).
24. Bernard W., Rindner W., Roth H. Anisotropic Stress Effect on the Excess Current in Tunnel Diodes., *J. Appl. Phys.*, **35**, 1860 (1964).
25. Голованов В. В., Панахов А. З. Влияние гидростатического давления на туннельный ток в GaAs-диодах. *ФТП*, **6**, (1972).

26. Davis R. E., Gibbons G. Design Principles and Construction of Planar Ge Esaki Diodes, *Solid State Electron.* 10, 461 (1967).
27. Hindin H. J. Tunnel Diode Flying High, *Electronics*. 52, 81 (1979).
28. Standards on Definitions, Symbols, and Methods of Test for Semiconductor Tunnel (Esaki) Diodes and Backward Diodes, *IEEE Trans. Electron Devices*, 12, 374 (1965).
29. Hauer W. B. Definition and Determination of the Series Inductance of Tunnel Diodes, *IRE Trans. Electron Devices*, 8, 470 (1961).
30. Esaki L., Miyahara Y. A New Device Using the Tunneling Process in Narrow $p - n$ Junctions, *Solid State Electron.*, 1, 13 (1960).
31. Hall R. N., Racette J. H., Ehrenreich H. Direct Observation of Polarons and Phonons During Tunneling in Group 3—5 Semiconductor Junctions, *Phys. Rev. Lett.*, 4, 456 (1960).
32. Chynoweth A. G., Logan R. A., Thomas D. E. Thion-Assisted Tunneling in Silicon and Germanium Esaki Junctions, *Phys. Rev.*, 125, 877 (1962).
33. Shurmer H. V. Backward Diodes as Microwave Detectors, *Proc. Inst. Electr. Eng., Lond.*, 111, 15111 (1964).
34. Eng. S. T. Low-Noise Properties of Microwave Backward Diodes, *IRE Trans. Microwave Theory Tech.*, MTT 8, 419 (1961).
35. Hopkins J. B. Microwave Backward Diodes in InAs, *Solid State Electron.* 13, 697 (1970).
36. Bhattacharyya A. B., Sarrot S. L. Switching Time Analysis of Backward Diodes, *Proc. IEEE* 58, 513 (1970).
37. Torrey H. C., Whitmer C. A., *Crystal Rectifiers*, McGraw-Hill, N. Y., 1948, Ch. 8.
38. Sze S. M., Ryder R. M. The Nonlinearity of the Reverse Current — Voltage Characteristics of a $p - n$ Junction near Avalanche Breakdown, *Bell Syst. Tech. J.*, 46, 1135 (1967).
39. Karlovsky J. The Curvature Coefficient of Germanium Tunnel and Backward Diodes, *Solid State Electron.*, 10, 1109 (1967).
40. Dahlke W. E., Sze S. M. Tunneling in Metal — Oxide — Silicon Structures, *Solid State Electron.*, 10, 865 (1967).
41. Chang L. L., Stiles P. J., Esaki L. Electron Tunneling between a Metal and a Semiconductor: Characteristics of Al — Al₂O₃ — SnTe and GeTe Junctions, *J. Appl. Phys.*, 38, 4440 (1967).
42. Kumar V., Dahlke W. E. Characteristics of Cr — SiO₂ — nSi Tunnel Diodes, *Solid State Electron.*, 20, 143 (1977).
43. Nicollian E. H., Goetzberger A. The Si — SiO₂ Interface Electrical Properties as Determined by the MIS Conductance Technique, *Bell Syst. Tech. J.*, 46, 1055 (1967).
44. Esaki L., Stiles P. J. New Type of Negative Resistance in Barrier Tunneling, *Phys. Rev. Lett.*, 16, 1108 (1966).
45. Green M. A., King F. D., Shewchun J. Minority Carrier MIS Tunnel Diodes and Their Application to Electron and Photovoltaic Energy Conversion: I. Theory, *Solid State Electron.*, 17, 551 (1974). II Experiment, *Solid State Electron.*, 17, 563 (1974).
46. Temple V. A. K., Green M. A., Shewchun J. Equilibrium-to-Nonequilibrium Transition in MOS Tunnel Diodes, *J. Appl. Phys.*, 45, 4934 (1974).
47. Card H. C., Rhoderick E. H. Studies of Tunnel MOS Diodes I. Interface Effects in Silicon Schottky Diodes, *J. Phys. D Appl. Phys.*, 4, 1589 (1971).
48. Yamamoto T., Morimoto M. Thin-MIS-Structure Si Negative Resistance Diode, *Appl. Phys. Lett.*, 20, 269 (1972).
49. Habib S. E.-D., Simmons J. G. Theory of Switching in $p - n$ Insulator (Tunnel) — Metal Devices, *Solid State Electron.*, 22, 181 (1979).
50. Simmons J. G., El-Badry A. Theory of Switching Phenomena in Metal/Semi-Insulator/ $n - p^+$ Silicon Devices, *Solid State Electron.*, 20, 955 (1977); A. El-

- Badry and Simmons J. G., Experimental Studies of Switching in Metal Semi-Insulating np^+ Silicon Devices, *Solid State Electron.*, **20**, 963 (1977).
51. Simmons J. G. Generalized Formula for the Electric Effect between Similar Electrodes Separated by a Thin Insulating Film. *J. Appl. Phys.*, **34**, 1793 (1963).
 52. Pollack S. R., Morris C. E. Tunneling through Gaseous Oxidized Films of Al_2O_3 , *Trans. AIME*, **233**, 497 (1965).
 53. Hurych Z. Influence of Nonuniform Thickness of Dielectric Layers on Capacitance and Tunnel Current, *Solid State Electron.*, **9**, 967 (1966).
 54. Pochobradsky J. On the Capacitance of Metal — Insulator — Metal Structures with Nonuniform Thickness, *Solid State Electron.*, **10**, 973 (1967).
 55. Ku H. Y., Ullman F. G. Capacitance of Thin Dielectric Structures, *J. Appl. Phys.* **35**, 265 (1964).
 56. Kurtin S., McGill T. C., Mead C. A. Tunneling Currents and $E - k$ Relation, *Phys. Rev. Lett.*, **25**, 756 (1970); Direct Interelectrode Tunneling in GaAs., *Phys. Rev.* **B3**, 3368 (1971).
 57. Heiblum M., Wang S. Y., Whinnery J. R., Gustafson T. K. Characteristics of Integrated MOM Junction at DC and at Optical Frequencies, *IEEE J. Quantum Electron.*, **QE-14**, 159 (1978).
 58. Mead C. A. Tunnel-Emission Amplifiers, *Proc. IRE.* **48**, 359 (1960).
 59. Crowell C. R., Sze S. M. Hot Electron Transport and Electron Tunneling in Thin-Film Structures, in Haas G., Thun R. E., Eds. *Physics of Thin Films*, Vol. 4, Academic, N. Y., 1967.
 60. Spratt J. P., Schwartz R. F., Kane W. M. Hot Electrons in Metal Films: Injection and Collection, *Phys. Rev., Lett.*, **6**, 341 (1961).
 61. Kisaki H. Tunnel Transistor, *Proc IEEE*, **61**, 1053 (1978).
 62. Chang L. L., Esaki L. Tunnel Triode — A Tunneling Base Transistor, *Appl. Phys. Lett.*, **31**, 687 (1977).
 63. Hewlett-Packard Electronic Test Instruments, Hewlett-Packard, Palo Alto, Calif., 1961.
 64. Liao S. Y. *Microwave Devices and Circuits*, Prentice-Hall, Englewood Cliffs, N. J., 1980.

ЛАВИННО-ПРОЛЕТНЫЕ ДИОДЫ И ДРУГИЕ ПРИБОРЫ НА ПРОЛЕТНЫХ ЭФФЕКТАХ

10.1. ВВЕДЕНИЕ

Принцип работы лавинно-пролетных диодов¹ (ЛПД) основан на возникновении отрицательного сопротивления в диапазоне сверхвысоких частот, которое обусловлено процессами лавинного умножения носителей и их пролета через полупроводниковую структуру. Появление отрицательного сопротивления связано с временным запаздыванием этих двух процессов, приводящим к фазовому сдвигу между током и напряжением. «Лавинное запаздывание» появляется за счет конечного времени нарастания лавинного тока, а «пролетное запаздывание» — за счет конечного времени прохождения носителями области дрейфа. Сопротивление диода отрицательно на некоторой частоте, если сумма этих времен равна полупериоду колебания. Впервые экспериментальное наблюдение осцилляций было осуществлено² в 1965 г. Джонстоном, Делочем и Коэном в смещенном в область обратного лавинного пробоя кремниевом диоде, который был помещен в СВЧ-резонатор [1, 2].

Возникающее за счет пролетных эффектов в полупроводниковых диодах отрицательное сопротивление впервые изучил в 1954 г. Шокли, который считал, что двухконтактные приборы из-за структурной простоты обладают потенциальными преимуществами по сравнению с трехконтактными транзисторами [3]. В 1958 г. Рид предложил конструкцию высокочастотного полупроводникового диода, в котором область лавинного умножения расположена на одном конце относительно высокоомного слоя, служащего дрейфовым пролетным пространством для генерируемых носителей заряда (т. е. $p^+—n—i—n^+$ или $n^+—p—i—p^+$) [4]. Осцилляции в диоде Рида впервые наблюдали Ли и др. [5]. Результаты малосигнальной теории, развитой Мисавой [6], Джилденом и Хайнсом [7], подтвердили, что отрицательное сопротивление, обусловленное лавинно-пролетным эффектом, может возникать

¹ В английской литературе эти приборы сокращенно называются IMPATT (impact ionization avalanche transit time) — *Прим. перев.*

² Лавинно-пролетный диод был впервые создан в СССР на основе обнаруженного в 1959 г. эффекта генерации когерентных СВЧ-колебаний при лавинном пробое германиевых дифференциальных диодов (Тагер А. С., Мельников А. И., Кобельков Г. П., Цебиев А. М. Генерация и усиление радиоволн сантиметрового и миллиметрового диапазонов с помощью полупроводникового диода в области положительного наклона его статической вольт-амперной характеристики. Диплом на открытие № 24, приоритет 27.X.1959 г.). — *Прим. ред.*

в обычном диоде с $p-n$ -переходом или контакте металл—полупроводник с произвольным профилем распределения примеси.

В настоящее время лавинно-пролетный диод является одним из самых мощных твердотельных источников СВЧ-излучения [8, 9]. ЛПД могут генерировать в непрерывном режиме наибольшую мощность в диапазоне частот, соответствующих миллиметровым длинам волн (т. е. свыше 30 ГГц) [10, 10а]. Однако следует сказать о трудностях, с которыми сталкиваются при работе ЛПД во внешней цепи: 1) высокий уровень шума; 2) необходимость тщательного расчета цепей (чтобы избежать расстройки или даже перегорания диода, поскольку реактивность большая и сильно зависит от амплитуды осцилляций) [11].

Другим прибором на пролетном эффекте является инжекционно-пролетный диод¹ (ИПД), название которого указывает на механизмы, лежащие в основе его работы. Колмен и Зи в 1971 г. сообщили о работе структуры металл—полупроводник—металл в режиме с проколом в качестве ИПД [12]. Аналогичные структуры были предложены в 1968 г. Ругом [13] (им был проведен анализ работы в режиме большого сигнала) и Райтом [14], который рассмотрел механизм переноса, ограниченного объемным зарядом. Поскольку в ИПД нет запаздывания, связанного с конечным временем нарастания лавины, предполагается, что эти приборы будут генерировать меньшие мощности при более низких к. п. д. по сравнению с лавинно-пролетными диодами. С другой стороны, шум, связанный с инжекцией носителей заряда через барьер, меньше лавинного шума ЛПД. Низкий уровень шума и хорошая стабильность позволяют применять ИПД в качестве гетеродина. Инжекционно-пролетные диоды обладают большей мощностью и меньшей величиной минимального обнаружимого сигнала, чем ЛПД или приборы на эффекте междолинного перехода электронов при их использовании в качестве доплеровских детекторов [15, 16]. Родственным инжекционно-пролетному диоду является двухскоростной пролетный диод² (ДСПД), в котором скорость носителей у одного контакта значительно меньше, чем у другого. До сих пор не было сообщений об экспериментальном наблюдении СВЧ-генерации в ДСПД. Тем не менее, согласно теоретическим оценкам, эти приборы будут иметь более высокие к. п. д. по сравнению с инжекционно-пролетными диодами, поскольку в ДСПД есть дополнительное запаздывание на время пролета области, в которой скорость носителей относительно низка.

¹ В английской литературе принято сокращение BARITT (barrier injection transit time). — Прим. перев.

² В английской литературе используется сокращение DOVETT (double velocity transit time). — Прим. перев.

Родственным ЛПД прибором является также пролетный диод с захваченным объемным зарядом лавины¹. Впервые такой режим работы был обнаружен Прагером, Чангом и Вайсбродом в 1967 г. в кремниевых лавинных диодах [17]. Рабочая частота этого прибора намного меньше пролетной, а к. п. д. значительно выше, чем у ЛПД. В результате теоретических исследований было установлено, что при работе в режиме большого сигнала лавинный процесс начинается в области высокого поля, а затем быстро распространяется на весь образец, в результате чего последний заполняется высокопроводящей электронно-дырочной плазмой, пространственный заряд которой снижает напряжение на диоде до очень малых величин [18—20]. Так как плазма не может быть быстро выведена из прибора, этот режим работы назван режимом с захваченным объемным зарядом лавины. Пролетные диоды с захваченным объемным зарядом лавины применяются в импульсных передатчиках и в радарх с фазными решетками [15, 21].

10.2. СТАТИЧЕСКИЕ ХАРАКТЕРИСТИКИ

Основными представителями семейства лавинно-пролетных диодов являются диод Рида, асимметричный резкий $p-n$ -переход, симметричный $p-n$ -переход (диод с двумя дрейфовыми областями), диод с двухслойной базой, диод с трехслойной базой (модифицированный диод Рида) и $p-i-n$ -диод. Ниже рассмотрены статические характеристики этих диодов (распределение электрического поля, напряжение пробоя и эффекты пространственного заряда).

На рис. 1 приведены распределения примесей, электрического поля и эффективного коэффициента ударной ионизации в идеализированном диоде Рида ($p^+-n-i-n^+$ - или $n^+-p-i-p^+$ -диоде) при пробое. Эффективный коэффициент ударной ионизации равен

$$\langle \alpha \rangle \equiv \alpha_n \exp \left[- \int_x^W (\alpha_n - \alpha_p) dx' \right], \quad (1)$$

где α_n и α_p — коэффициенты ударной ионизации электронов и дырок соответственно и W — ширина обедненной области.

Как показано в гл. 2, условие лавинного пробоя можно записать в виде

$$\int_0^W \langle \alpha \rangle dx = 1. \quad (2)$$

¹ В английской литературе этот прибор сокращенно называется TRAPATT (trapped plasma avalanche triggered transit). — Прим. перев.

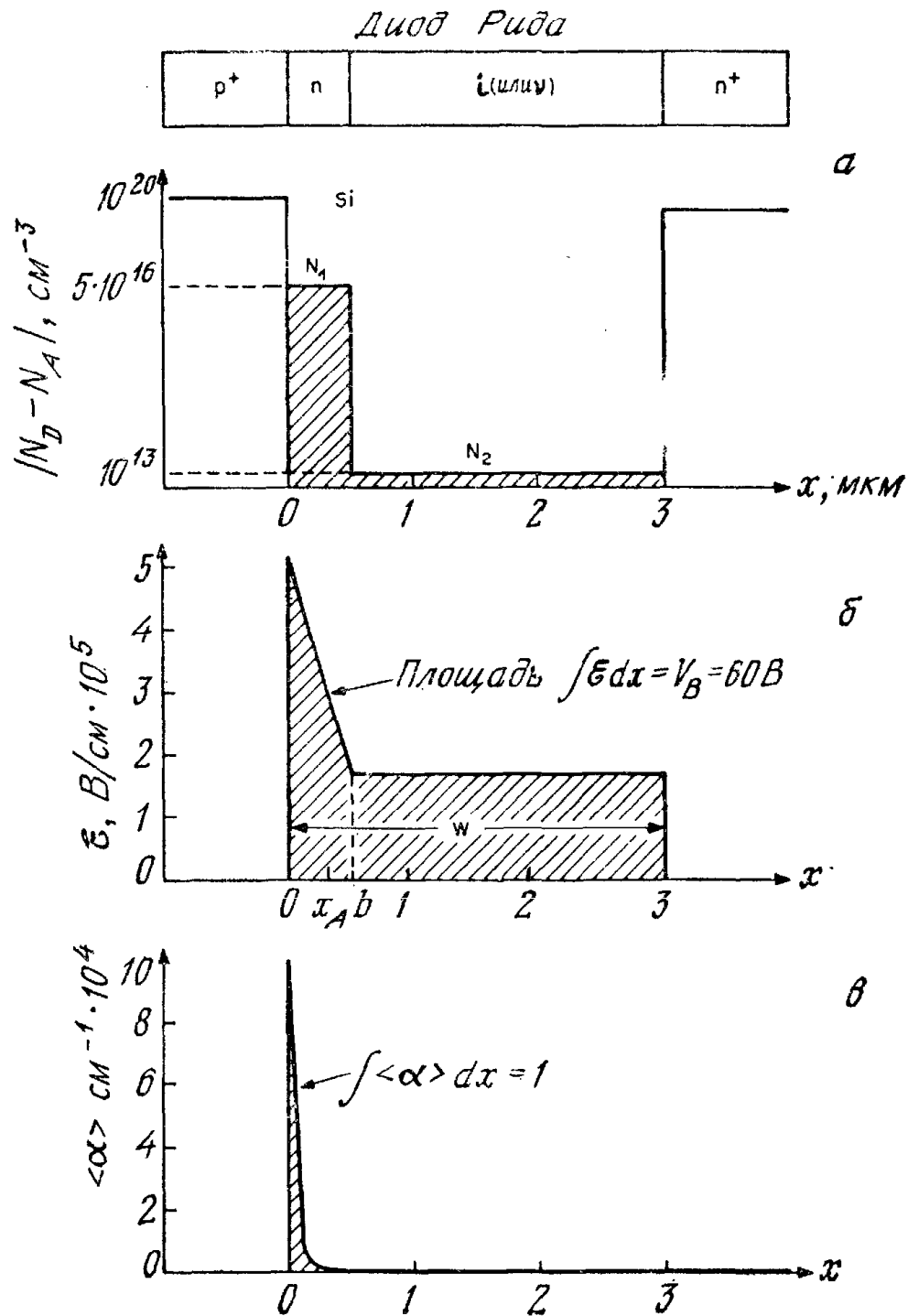


Рис. 1. Распределения примесей (а), напряженности электрического поля (б) и эффективного коэффициента ионизации (в) при пробое в $p^+ - n - i - n^+$ -диоде Рида [4].

Из-за сильной зависимости коэффициентов ударной ионизации от напряженности электрического поля область лавинного умножения сильно локализована, т. е. практически весь процесс умножения носителей происходит в узком слое высокого поля от 0 до x_A , где x_A — ширина области умножения. Падение напряжения на области умножения обозначим V_A . Ниже мы покажем, что оптимальная плотность тока и максимальный к. п. д. лавинно-пролетных диодов сильно зависят как от x_A , так и от V_A . Слой вне области умножения ($x_A \leq x \leq W$) называется областью дрейфа.

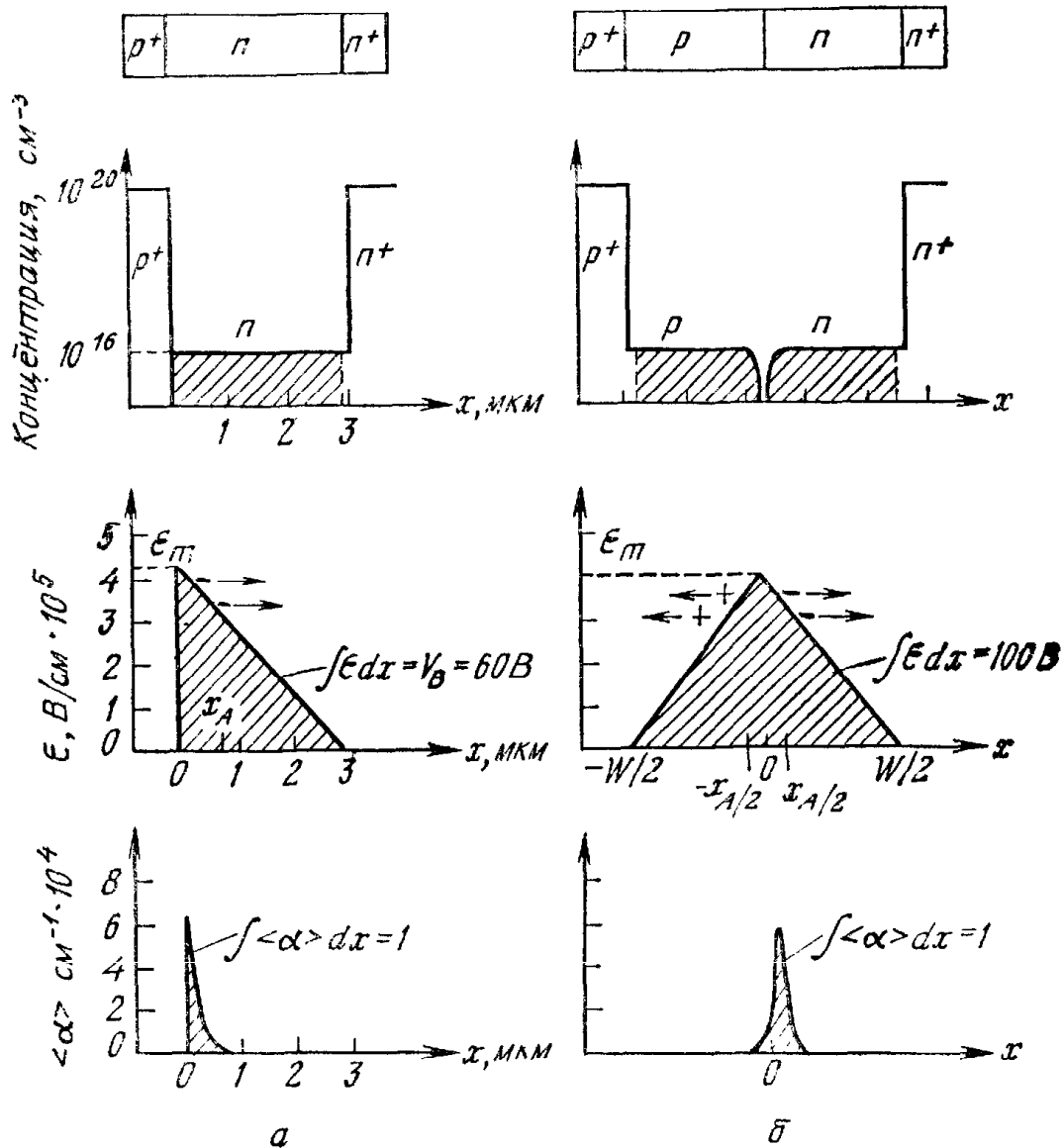


Рис. 2. Распределения примесей, напряженности электрического поля и эффективного коэффициента ионизации в диоде с асимметричным резким p^+ — n -переходом (диод с одной дрейфовой областью) (а) и в p^+ — p — n — n^+ -диоде с симметричным резким p — n -переходом (диод с двумя дрейфовыми областями) (б).

Существуют два предельных случая распределения примеси в диодах Рида. Если отсутствует область с концентрацией примеси N_2 , то мы имеем резкий p^+ — n -переход. Если же отсутствует область с концентрацией N_1 , то диод Рида вырождается в p — i — n -диод, который иногда называют диодом Мисавы [6]. На рис. 2, а приведены статические характеристики диода с асимметричным резким p — n -переходом. Область лавинного умножения сильно локализована. На рис. 2, б приведены характеристики кремниевого диода с симметричным резким p — n -переходом. Область умножения расположена вблизи центра обедненной области. Небольшая асимметрия эффективного коэффициента ударной ионизации по отношению к точке, в которой напряженность электрического поля максимальна, связана с тем, что в кремнии α_n и α_p

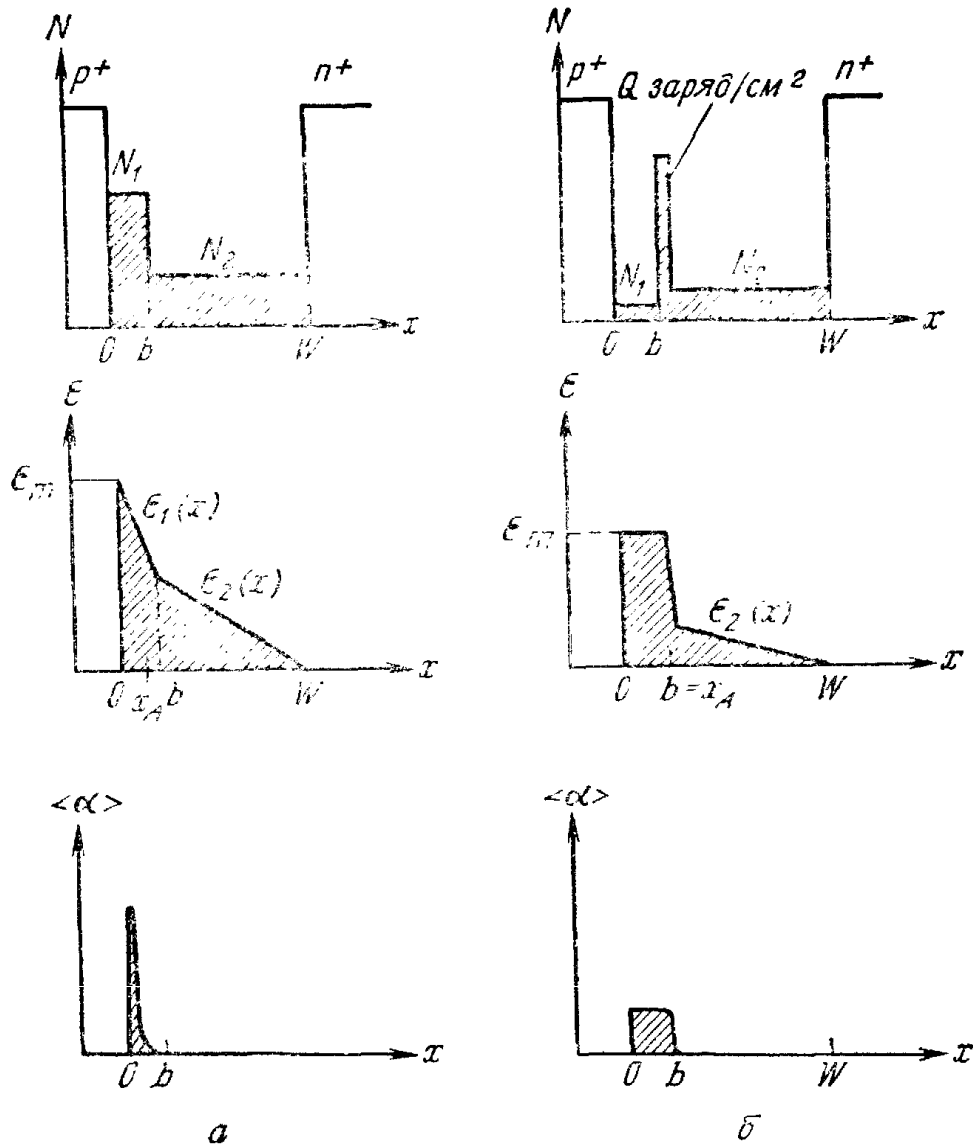


Рис. 3. Распределения примесей, напряженности электрического поля и эффективного коэффициента ионизации в модифицированных диодах Рида: с двухслойной базой (а) и с трехслойной базой (б).

сильно различаются. Если $\alpha_n \simeq \alpha_p$, как это имеет место в GaP, эффективный коэффициент ионизации

$$\langle \alpha \rangle = \alpha_n = \alpha_p, \quad (3)$$

и область лавинного умножения симметрична по отношению к точке $x = 0$.

На рис. 3, а приведены статические характеристики модифицированной структуры Рида — диода с двухслойной базой, в котором концентрация N_2 существенно превышает соответствующую величину для диода Рида [22]. Статические характеристики другой модификации — диода с трехслойной базой приведены на рис. 3, б. В этом диоде «сгусток» заряда Q расположен в точке $x = b$. Поскольку достаточно высокая напряженность электрического поля остается постоянной между точками $x = 0$ и $x = b$, то максимальное поле может быть меньше, чем в диоде

в двухслойной базой. В $p-i-n$ -диоде электрическое поле однородно во всей i -области при малых токах. В этом случае область лавинного умножения соответствует всей ширине слоя собственной концентрации.

10.2.1. Напряжение пробоя

Напряжение пробоя для асимметричных резких $p-n$ -переходов было рассмотрено в гл. 2. Мы можем воспользоваться развитым в этой главе методом для того, чтобы найти напряжения пробоя для других типов диодов. Для симметричных резких переходов (рис. 2, б) напряжение пробоя определяется равенством

$$V_B = \frac{1}{2} \mathcal{E}_m W = \frac{\epsilon_s \mathcal{E}_m^2}{q N_B}, \quad (4)$$

где \mathcal{E}_m — величина максимальной напряженности электрического поля при $x = 0$. На рис. 4 приведены максимальные напряженности электрического поля при пробое в симметричных и асимметричных $p-n$ -переходах из кремния и арсенида галлия с ориентацией $\langle 100 \rangle$ [23, 24]. Если известна концентрация примеси, то напряжение пробоя можно найти из равенства (4), пользуясь графиками, приведенными на рис. 4. Обратное напряжение при пробое равно $(V_B - V_{bi})$, где V_{bi} — контактная разность потенциалов, причем $V_{bi} = (2kT/q) \ln(N_B/n_i)$ в случае симметричных резких переходов (в работающих ЛПД величиной V_{bi} можно пренебрегать).

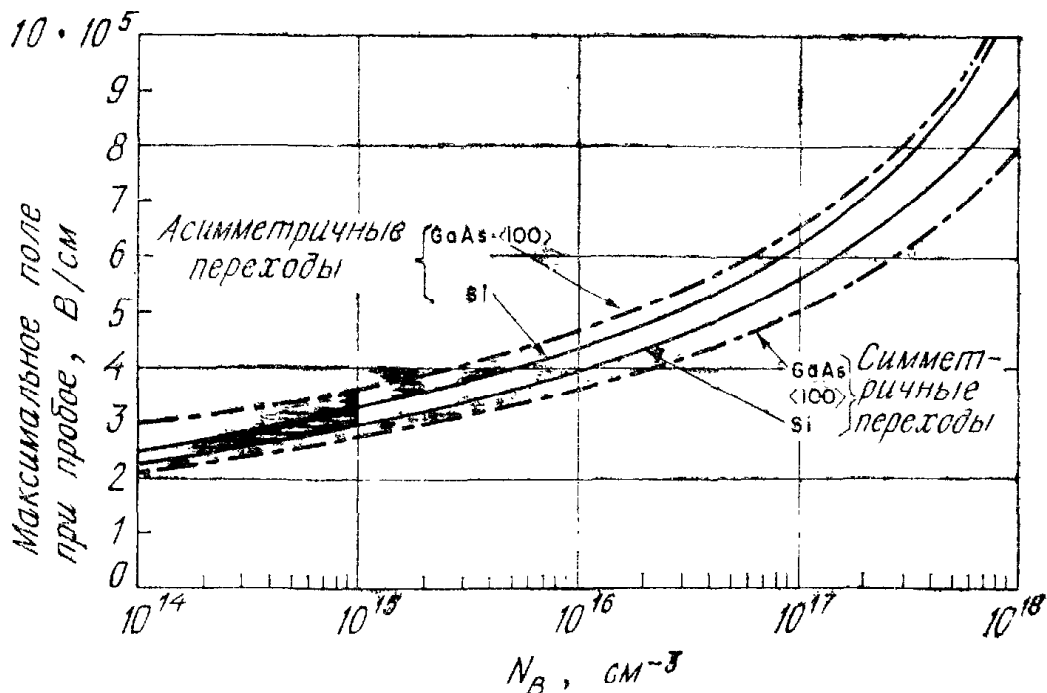


Рис. 4. Максимальная напряженность электрического поля при пробое как функция концентрации для асимметричных и симметричных резких переходов из Si и GaAs [23, 24].

Напряжение пробоя и ширина обедненной области в диоде Рида и диоде с двухслойной базой определяются следующими равенствами:

$$V_B = \left(\mathcal{E}_m - \frac{qN_1b}{2\epsilon_s} \right) b - \frac{1}{2} \left(\mathcal{E}_m - \frac{qN_1b}{\epsilon_s} \right) (W - b), \quad (5)$$

$$W = \frac{\epsilon_s \mathcal{E}_m}{qN_2} - b \left(\frac{N_1}{N_2} - 1 \right). \quad (6)$$

Ширина обедненной области в диоде Рида ограничена толщиной эпитаксиального слоя, которая меньше величины, вычисленной с помощью равенства (6). Максимальная напряженность электрического поля при пробое в диоде Рида или диоде с двухслойной базой с точностью до 1 % совпадает с соответствующей величиной для асимметричного резкого $p^+ - n$ -перехода с такой же концентрацией примеси N_1 при условии, что ширина области лавинного умножения $x_A < b$ [25]. Следовательно, напряжение пробоя можно найти из равенства (5) и (6), используя величину максимальной напряженности электрического поля из графиков, приведенных на рис. 4.

В диоде с трехслойной базой при очень узком слое объемного заряда с высокой концентрацией примеси напряжение пробоя и полная ширина обедненной области определяются выражениями

$$V_B = \mathcal{E}_m b + \frac{1}{2} \left(\mathcal{E}_m - \frac{qQ}{\epsilon_s} \right) (W - b), \quad (7)$$

$$W = \frac{\epsilon_s}{qN_2} \left(\mathcal{E}_m - \frac{qQ}{\epsilon_s} \right) + b, \quad (8)$$

где Q — число атомов примеси на 1 см^2 в слое объемного заряда. Поскольку напряженность электрического поля приблизительно постоянна при $0 \leq x \leq b$, эффективный коэффициент ударной ионизации $\langle \alpha \rangle = 1/b$. Максимальную напряженность электрического поля можно найти, если известна зависимость коэффициента ионизации от поля.

10.2.2. Область лавинного умножения и область дрейфа

В идеальном $p-i-n$ -диоде область лавинного умножения занимает всю ширину слоя собственной концентрации. Однако в диодах Рида и $p-n$ -переходах область умножения носителей заряда очень узкая и расположена вблизи металлургической границы перехода. Вклад в интеграл в выражении (2) быстро уменьшается по мере удаления x от металлургической границы. Поэтому имеет смысл определить ширину области умножения x_A таким

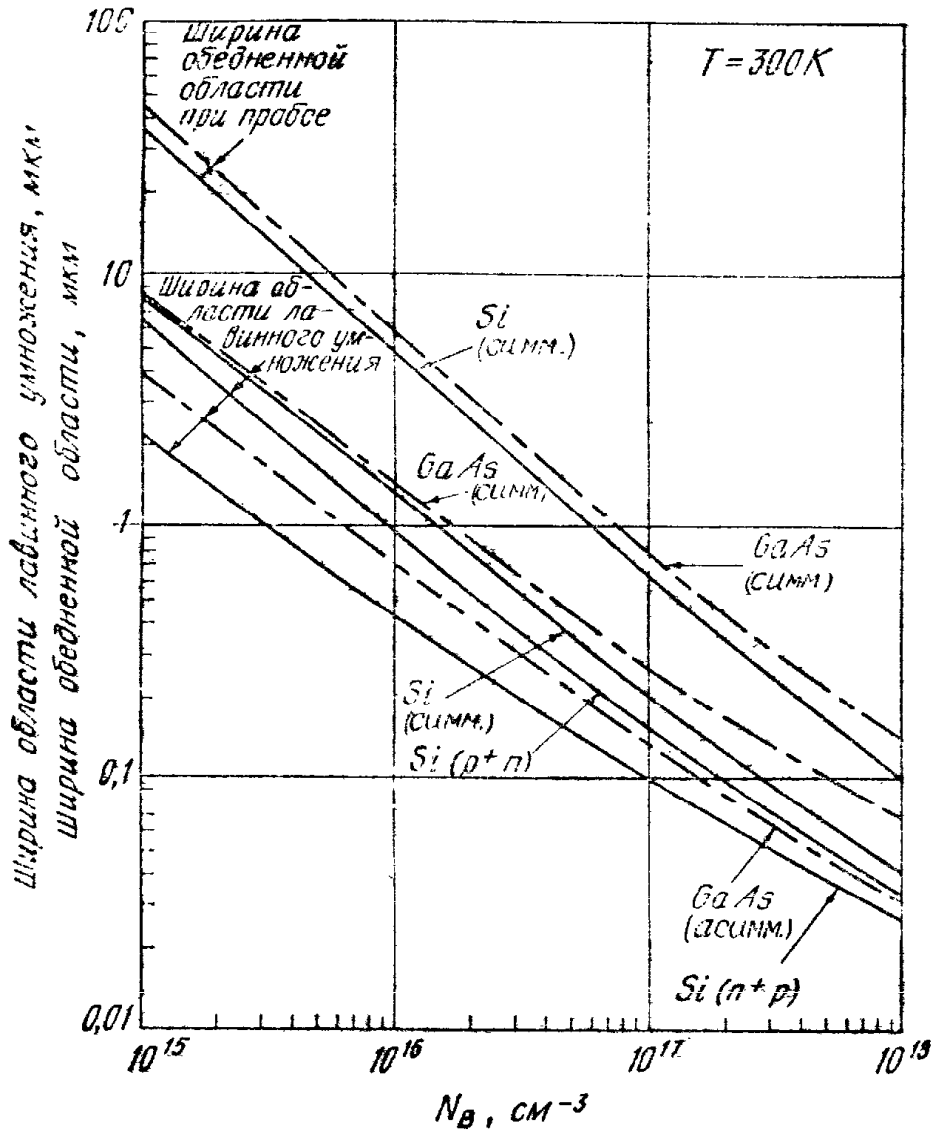


Рис. 5. Ширина области лавинного умножения в переходах из кремния и арсенида галлия и ширина области обеднения в зависимости от концентрации примеси в симметричных переходах из Si и GaAs [24].

образом, чтобы при интегрировании эффективного коэффициента ионизации от 0 до x_A получить 95 % полной величины:

$$\int_0^{x_A} \langle \alpha \rangle dx = 0,95, \quad (9)$$

а для диода Рида, диода с двухслойной базой, асимметричного резкого $p^+ - n$ -перехода и симметричных $p - n$ -переходов при пробое

$$\int_{-x_A/2}^{x_A/2} \langle \alpha \rangle dx = 0,95. \quad (10)$$

На рис. 5 приведены зависимости ширины области лавинного умножения от концентрации примеси в диодах из кремния и арсенида галлия [24], а также зависимости ширины обедненной области

в симметричных переходах. Из-за того что в кремнии коэффициенты ионизации различны ($\alpha_n > \alpha_p$), при одинаковой концентрации примеси ширина области умножения в $n^+—p$ -переходе меньше, чем в $p^+—n$ -переходе. В диоде Рида или в диоде с двухслойной базой ширина области умножения такая же, как и в асимметричном резком $p^+—n$ -переходе, если концентрация примеси N_1 одинакова. В диоде с трехслойной базой область умножения занимает пространство от металлургической границы до слоя заряда высокой концентрации при $x_A = b$.

Область дрейфа совпадает с обедненной областью, за исключением слоя лавинного умножения носителей, или $x_A < x < W$. Важнейшей характеристикой области дрейфа является скорость, с которой носители заряда пролетают через эту область. Время пролета области дрейфа минимально при достаточно большой напряженности электрического поля, когда носители движутся со скоростью насыщения v_s . В кремнии напряженность электрического поля должна превышать 10^4 В/см. В арсениде галлия поле может быть значительно меньше ($\sim 10^3$ В/см).

В $p—i—n$ -диодах это требование выполняется автоматически, так как электрическое поле при пробое приблизительно одинаково на всей длине области собственной концентрации и существенно превышает поле, необходимое для насыщения скорости. В диоде Рида минимальная напряженность электрического поля в области дрейфа равна $\mathcal{E}_{\text{мин}} = \mathcal{E}_m - q [N_1 b + N_2 (W - b)] / \epsilon_s$.

Из сказанного выше следует, что диод Рида нужно рассчитать так, чтобы $\mathcal{E}_{\text{мин}}$ было достаточно велико. В резких $p—n$ -переходах всегда существуют области, в которых напряженность электрического поля меньше необходимой для начала пробоя. Однако эти области занимают лишь небольшую часть всего обедненного слоя. Например, в кремниевом $p^+—n$ -переходе с концентрацией примеси 10^{16} см $^{-3}$ максимальное поле при пробое равно 4×10^5 В/см. Отношение ширины области, в которой напряженность электрического поля относительно низка (меньше 10^4 В/см), к полной ширине обедненной области составляет $10^4 / 4 \cdot 10^5 = 2,5$ %. В $p^+—n$ -переходе из арсенида галлия при такой же концентрации примеси ширина области слабого поля составляет менее 0,2 % ширины обедненной области. Поэтому существование области слабого поля практически не сказывается на времени пролета носителей через обедненную область.

10.2.8. Влияние температуры и объемного заряда носителей

Напряжение пробоя и максимальные напряженности электрического поля, о которых речь шла выше, вычисляются при комнатных температурах без учета эффектов объемного заряда носителей, разогрева и в отсутствие генерации. Однако в рабочих

условиях ЛПД смещен в область лавинного пробоя, и обычно плотность тока очень велика. Это приводит к заметному увеличению температуры в переходе и появлению эффектов, связанных с объемным зарядом носителей.

Коэффициенты ионизации электронов и дырок уменьшаются при увеличении температуры [26]. Поэтому при заданной концентрации примеси напряжение пробоя ЛПД будет возрастать по мере того, как температура повышается. С ростом потребляемой на постоянном токе мощности (произведение напряжения на ток при обратном смещении) возрастают как температура в переходе, так и напряжение пробоя. В конце концов диод выходит из строя в основном из-за сильного разогрева, вызывающего локальные разрушения. Поэтому увеличение температуры в $p-n$ -переходе накладывает сильные ограничения на режим работы диода. Для того чтобы предотвратить увеличение температуры в приборе, необходим хороший теплоотвод. Этот вопрос мы обсудим в разд 10.4.

Объемный заряд генерируемых носителей приводит к изменению напряженности электрического поля в обедненной области. Этот эффект вызывает появление дополнительного статического сопротивления, которое положительно для резких переходов и отрицательно для $p-i-n$ -диодов [27].

Сначала рассмотрим p^+-n-n^+ -диод с асимметричным резким переходом (рис. 6, а). Когда приложенное напряжение V равно напряжению пробоя V_B , абсолютная величина напряженности электрического поля $|\mathcal{E}_x|$ максимальна в точке $x = 0$ и равна \mathcal{E}_m . Если предположить, что электроны движутся в обедненной области со скоростью насыщения v_s , то ток I равен

$$I = v_s \rho A, \quad (11)$$

где ρ — плотность заряда носителей и A — площадь. Изменение напряженности электрического поля $\Delta\mathcal{E}(x)$, связанное с объемным зарядом носителей, можно найти из равенства (11) и уравнения Пуассона

$$\Delta\mathcal{E}(x) \approx \frac{Ix}{A\epsilon_s v_s}. \quad (12)$$

Если предположить, что все носители генерируются в области умножения шириной x_A , то изменение напряжения, связанное с объемным зарядом в области дрейфа $W - x_A$, можно найти, проинтегрировав $\Delta\mathcal{E}(x)$:

$$\Delta V_B \approx \int_0^{W-x_A} \frac{Ix}{A\epsilon_s v_s} dx = I \frac{(W-x_A)^2}{2A\epsilon_s v_s}. \quad (13)$$

Полная величина приложенного напряжения равна

$$V = V_B + \Delta V_B = V_B + IR_{SC}, \quad (14)$$

где R_{SC} — сопротивление области пространственного заряда [28], которое определяется из выражений (13) и (14):

$$R_{SC} \equiv \frac{\Delta V_B}{I} \approx \frac{(W - x_A)^2}{2A\epsilon_s v_s}. \quad (15)$$

Для значений параметров, приведенных на рис. 6, а, сопротивление области пространственного заряда приблизительно равно 20 Ом при $A = 10^{-4} \text{ см}^2$.

В $p-i-n$ - или $p-v-n$ -диодах ситуация отличается от случая p^+-n -перехода. Когда приложенное напряжение достаточно для начала лавинного пробоя, обратный ток мал. Влиянием объемного

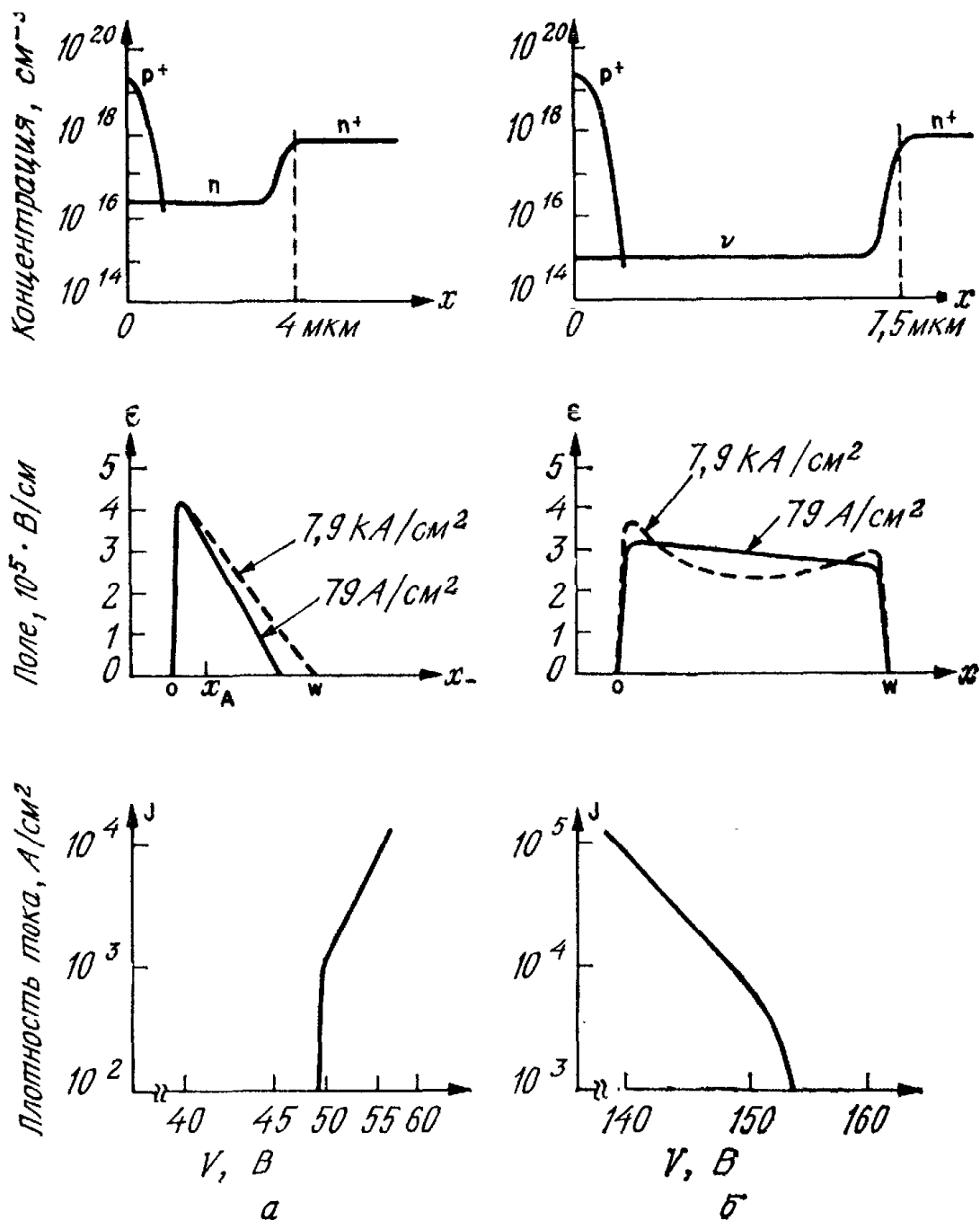


Рис. 6. Распределения примесей электрического поля и вольт-амперные характеристики $p^+ - n - n^+$ -диода (а) и $p^+ - v - n^+$ -диода (б) [27].

заряда можно пренебречь, а поле считать практически однородным в обедненной области. По мере увеличения тока все большее количество электронов генерируется за счет ударной ионизации вблизи $p-i$ -перехода, а дырок — вблизи $n-i$ перехода (так как поле имеет два максимума, см. рис. 6, б). Возникающий пространственный заряд снижает электрическое поле в центре v -области.

Поэтому уменьшается и напряжение, которое равно $\int_0^W \mathcal{E} dx$. Это приводит к уменьшению сопротивления $p-v-n$ -диода по постоянному току (рис. 6, б).

10.3. ДИНАМИЧЕСКИЕ ХАРАКТЕРИСТИКИ

10.3.1. Запаздывание, связанное с инжекцией.

Пролетный эффект

Сначала рассмотрим запаздывание фазы, связанное с инжекцией, и пролетный эффект в идеализированном приборе [29]. Предположим, что импульс тока проводимости инжектируется при $x = 0$ (рис. 7) с фазовым углом φ по отношению к полному току, а приложенное к диоду постоянное смещение таково, что инжектированные носители заряда пролетают область дрейфа со скоростью насыщения v_s . Плотность переменного тока проводимости \tilde{J}_c при $x = 0$ равна полной плотности тока \tilde{J} со сдвигом по фазе:

$$\tilde{J}_c(x=0) = \tilde{J} \exp(-j\varphi). \quad (16)$$

В произвольной точке области дрейфа полный переменный ток равен сумме тока проводимости и тока смещения:

$$\tilde{J}(x) = \tilde{J}_c(x) + \tilde{J}_d(x) = \tilde{J}_c(x=0) e^{-j\omega x/v_s} + j\omega \epsilon_s \tilde{\mathcal{E}}(x), \quad (17)$$

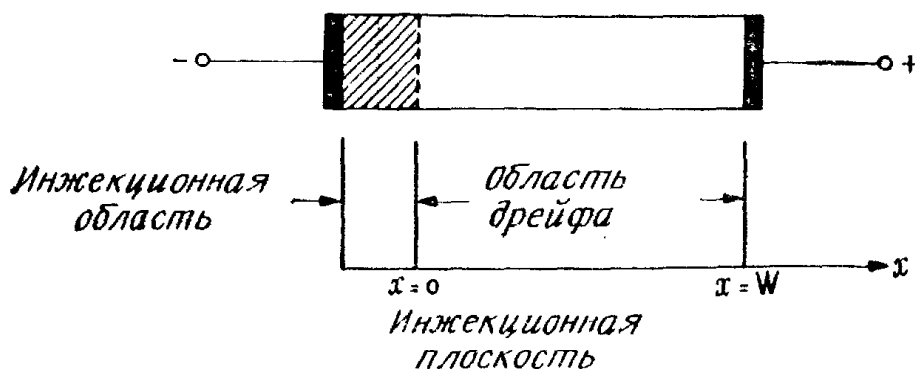


Рис. 7. Идеализированный диод, в котором носители инжектируются при $x = 0$ и движутся в области дрейфа со скоростью насыщения.

где $\tilde{\mathcal{E}}(x)$ — переменная компонента напряженности электрического поля. Из выражений (16) и (17) получаем

$$\tilde{\mathcal{E}}(x) = \frac{\tilde{J} (1 - e^{-j\omega x/v_s - j\varphi})}{j\omega\epsilon_s}. \quad (18)$$

Проинтегрировав выражение (18), найдем импеданс диода

$$Z \equiv \frac{\int_0^W \tilde{\mathcal{E}}(x) dx}{\tilde{J}} = \frac{1}{j\omega C} \left[1 - \frac{e^{-j\varphi} (1 - e^{-j\theta})}{j\theta} \right], \quad (19)$$

где $C = \epsilon_s/W$ — емкость на единицу площади и $\theta = \omega W/v_s$ — пролетный угол. Выделив действительную и мнимую части выражения (19), получим

$$R = \frac{\cos \varphi - \cos (\varphi + \theta)}{\omega C \theta}, \quad (20)$$

$$X = -\frac{1}{\omega C} + \frac{\sin (\theta + \varphi) - \sin \varphi}{\omega C \theta}. \quad (21)$$

Ниже мы изучим влияние инжекционного фазового угла φ на сопротивление по переменному току R . Когда угол φ равен нулю (нет фазового запаздывания), сопротивление пропорционально величине $(1 - \cos \theta)/\theta$, которая всегда больше или равна нулю (рис. 8, а), и поэтому сопротивление положительно. Следовательно, только пролетный эффект не может приводить к возникновению отрицательного сопротивления. Однако если $\varphi \neq 0$, то сопротивление отрицательно при некоторых значениях пролетных углов. Например, при $\varphi = \pi/2$ наибольшее отрицательное сопротивление соответствует углу пролета $\theta \approx 3\pi/2$ (рис. 8, б). На этом принципе основана работа инжекционно-пролетного диода: инжекция неосновных носителей через барьер вносит фазовое запаздывание 90° , а пролетный угол, равный 270° , оптимизирует характеристику прибора. Детальный анализ работы ИПД проведен в разд. 10.7. Если $\varphi = \pi$, сопротивление максимально при $\theta = \pi$ (рис. 8, в). Эта ситуация соответствует лавинно-пролетному режиму, в котором фазовый сдвиг 180° достигается за счет конечного времени нарастания лавинного тока, а пролетный эффект приводит к добавочному запаздыванию на 180° .

Изложенное выше подтверждает важность инжекционного запаздывания. Проблема создания прибора, использующего при работе пролетные эффекты, сводится, таким образом, к поиску способа, с помощью которого можно вызвать запаздывание инжекции носителей в область дрейфа. Из рис. 8 видно, что сумма инжекционной фазы и оптимального угла пролета $\varphi + \theta_{opt}$ приблизительно равна 2π , а отрицательное сопротивление рас-

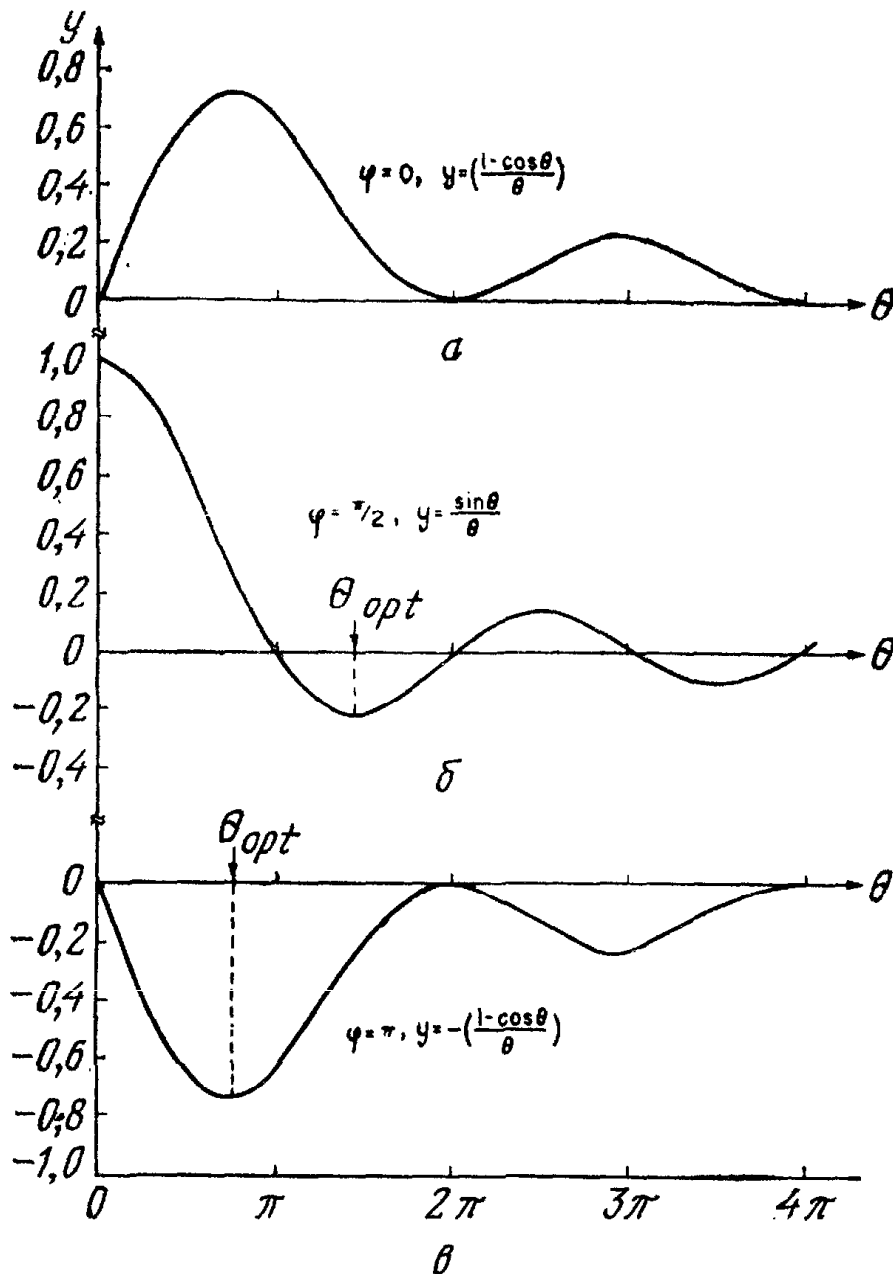


Рис. 8. Зависимость сопротивления от пролетного угла при трех значениях инжекционного запаздывания фазы.

тет по мере того, как φ становится все более положительным. Из-за более высокого сопротивления лавинно-пролетные диоды имеют лучшие к. п. д. и генерируют большие мощности, чем инжекционно-пролетные диоды.

10.3.2. Малосигнальный анализ

Анализ малосигнальных процессов впервые проведен Ридом [4] и затем развит Джилденом и Хайнсом [7]. Для простоты предположим, что коэффициенты ионизации $\alpha_n = \alpha_p = \alpha$ и скорости насыщения электронов и дырок равны. На рис. 9, а приведена схема диода Рида. Воспользовавшись соображениями, изложенными в разд. 10.2, разобьем диод на три области: 1) область

лавинного умножения, которая предполагается настолько тонкой, что в ней можно не учитывать объемный заряд и запаздывание сигнала; 2) область дрейфа, в которой отсутствует генерация носителей, а имеющиеся носители движутся со скоростями насыщения; 3) пассивную область, которая вносит нежелательное паразитное сопротивление.

Две активные области взаимодействуют друг с другом, так как переменное электрическое поле непрерывно на их границе. Будем использовать нижний индекс «0» для обозначения постоянных величин, а тильду (\sim) для малосигнальных переменных. Величины, содержащие постоянную и переменную составляющие, обозначим буквами без индексов. Плотность лавинного тока в области умножения \tilde{J}_A равна плотности переменного тока проводимости в этой области, а полная плотность переменного тока равна \tilde{J} . Поскольку мы предположили, что область умножения узка, то ток \tilde{J}_A достигает границы области дрейфа без запаздывания. Если носители движутся со скоростью насыщения v_s , плотность переменного тока проводимости $\tilde{J}_c(x)$ в области дрейфа представляет собой незатухающую волну (в которой изменяется лишь фаза), распространяющуюся со скоростью насыщения:

$$\tilde{J}_c(x) = \tilde{J}_A e^{-j\omega x/v_s} \equiv \gamma \tilde{J} e^{-j\omega x/v_s}, \quad (22)$$

где $\gamma \equiv \tilde{J}_A/\tilde{J}$ — комплексная величина, равная отношению плотности лавинного тока к плотности полного тока.

В любом поперечном сечении плотность полного тока \tilde{J} равна сумме плотностей тока проводимости \tilde{J}_c и тока смещения \tilde{J}_d . Эта сумма не зависит от x :

$$\tilde{J} = \tilde{J}_c(x) + \tilde{J}_d(x) \neq f(x). \quad (23)$$

Плотность тока смещения связана с напряженностью переменного электрического поля $\tilde{\mathcal{E}}(x)$ следующим соотношением:

$$\tilde{J}_d = j\omega\epsilon_s \tilde{\mathcal{E}}(x). \quad (24)$$

Из выражений (22)—(24) можно получить выражение для переменной компоненты напряженности электрического поля в области дрейфа как функции x и \tilde{J} :

$$\tilde{\mathcal{E}}(x) = \tilde{J} \frac{1 - \gamma e^{-j\omega x/v_s}}{j\omega\epsilon_s}. \quad (25)$$

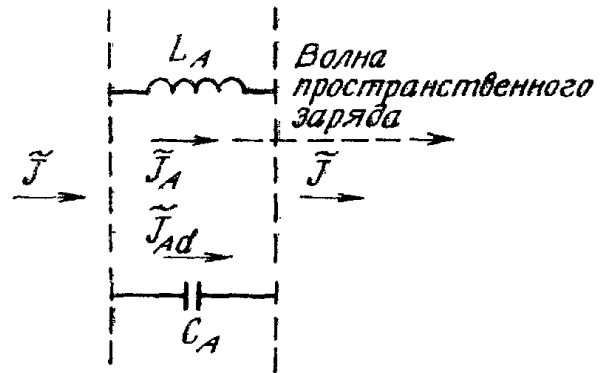
В результате интегрирования $\tilde{\mathcal{E}}(x)$ получим зависимость падения напряжения на этой области. Выражение для коэффициента γ мы получим ниже.

Область лавинного умножения. Сначала рассмотрим область умножения. Плотность постоянного тока $J_0 = J_{p0} + J_{n0}$ связана с плотностью обратного тока насыщения, обусловленного термической генерацией $J_s = J_{ps} + J_{ns}$, соотношением

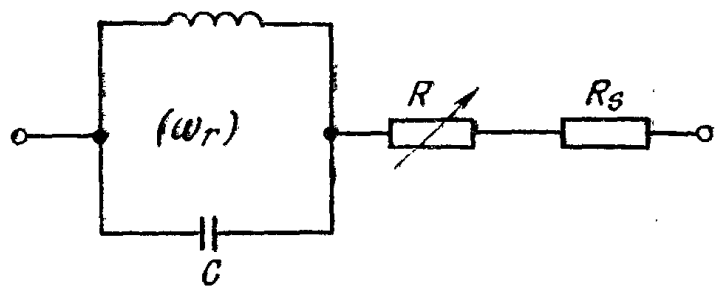
$$\frac{J_s}{J_0} = 1 - \int_0^W \langle \alpha \rangle dx. \quad (26)$$



а



б



в

Рис. 9. Модель диода Рида с областью лавинного умножения, областью дрейфа (б), эквивалентная схема диода Рида при малых пролетных углах (в) и за

При пробое J_0 стремится к бесконечности, а $\int_0^w \langle \alpha \rangle dx = 1$.

В стационарном случае интеграл от эффективного коэффициента ионизации не превышает 1. Это не обязательно имеет место в случае быстро изменяющегося поля. Ниже мы получим дифференциальное уравнение для тока, зависящего от времени. Если предположить, что коэффициенты ионизации и скорости насыщения электронов и дырок равны и дрейфовые токи существенно превышают диффузионные, то основные уравнения, описывающие работу прибора, в одномерном приближении имеют вид

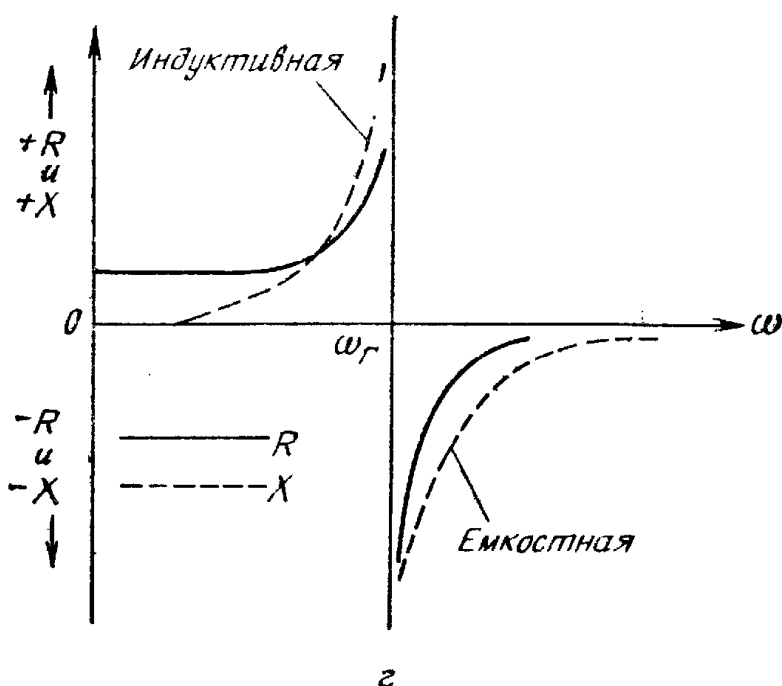
$$\frac{\partial \mathcal{E}}{\partial x} = \frac{q}{\epsilon_s} (N_D^+ - N_A^- + p - n) \quad (\text{уравнение Пуассона}), \quad (27)$$

$$\left. \begin{aligned} J_n &= qv_s n, \\ J_p &= qv_s p, \\ J &= J_n + J_p \end{aligned} \right\} \quad (\text{равенства для плотности тока}), \quad (28)$$

$$\left. \begin{aligned} \frac{\partial n}{\partial t} &= \frac{1}{q} \frac{\partial J_n}{\partial x} + \alpha v_s (n + p), \\ \frac{\partial p}{\partial t} &= -\frac{1}{q} \frac{\partial J_p}{\partial x} + \alpha v_s (n + p) \end{aligned} \right\} \quad (\text{уравнения непрерывности}). \quad (29a)$$

$$(29b)$$

Вторые слагаемые в правых частях выражений (29a) и (29b) равны скорости генерации электронно-дырочных пар за счет лавинного умножения. Эта величина значительно превышает скорость термической генерации, которой можно пренебречь. Про-



и пассивной областью (а), эквивалентная схема для области лавинного умножения (б), зависимости активных и реактивных составляющих импеданса от частоты (в) [7].

суммировав выражения (29а) и (29б) с учетом выражения (28) и проинтегрировав от $x = 0$ до $x = x_A$, получим

$$\tau_A \frac{dJ}{dt} = - (J_p - J_n)_0^{x_A} + 2J \int_0^{x_A} \alpha dx, \quad (30)$$

где $\tau_A = x_A/v_s$ — время пролета носителями области умножения.

Граничным условием служит равенство тока электронов при $x = 0$ обратному току насыщения J_{ns} . Поэтому в плоскости $x = 0$ $J_p - J_n = -2J_n + J = -2J_{ns} + J$. В плоскости $x = x_A$ дырочный ток равен обратному току насыщения J_{ps} генерируемых в области пространственного заряда дырок. Следовательно, $J_p - J_n = 2J_p - J = 2J_{ps} - J$. С учетом этих граничных условий уравнение (30) принимает следующий вид:

$$\frac{dJ}{dt} = \frac{2J}{\tau_A} \left[\int_0^{x_A} \alpha dx - 1 \right] + \frac{2J_s}{\tau_A}. \quad (31)$$

В стационарных условиях J совпадает с постоянным током J_0 , и уравнение (31) переходит в равенство (26).

Упростим уравнение (31), подставив вместо α средний коэффициент ионизации $\bar{\alpha}$, полученный интегрированием по области умножения. Если пренебречь слагаемым с J_s , получим

$$\frac{dJ}{dt} = \frac{2J}{\tau_A} (\bar{\alpha} x_A - 1). \quad (32)$$

Сделаем предположение о малости переменного сигнала:

$$\begin{aligned} \bar{\alpha} &= \bar{\alpha}_0 + \bar{\alpha} e^{j\omega t} \simeq \bar{\alpha}_0 + \bar{\alpha}' \tilde{\mathcal{E}}_A e^{j\omega t}, \\ \bar{\alpha} x_A &= 1 + x_A \alpha' \tilde{\mathcal{E}}_A e^{j\omega t}, \\ J &= J_0 + \tilde{J}_A e^{j\omega t}, \quad \mathcal{E} = \mathcal{E}_0 + \tilde{\mathcal{E}}_A e^{j\omega t}, \end{aligned} \quad (33)$$

где $\alpha' \equiv \partial \alpha / \partial \mathcal{E}$ и $\bar{\alpha} = \alpha' \tilde{\mathcal{E}}_A$. Подставляя приведенные выше соотношения в уравнение (32) и пренебрегая членами более высокого порядка малости, получим выражение для переменной компоненты плотности лавинного тока проводимости

$$\tilde{J}_A = \frac{2\alpha' x_A J_0 \tilde{\mathcal{E}}_A}{j\omega \tau_A}. \quad (34)$$

Ток смещения в области умножения равен

$$\tilde{J}_{Ad} = j\omega \epsilon_s \tilde{\mathcal{E}}_A. \quad (35)$$

Это две компоненты полного тока, протекающего в цепи. Для заданной величины напряженности электрического поля лавинный

ток \tilde{J}_A имеет реактивный характер и изменяется обратно пропорционально ω , как в индуктивности. Другая составляющая тока \tilde{J}_{Ad} также имеет реактивный характер и изменяется прямо пропорционально частоте, как в емкости. Таким образом, область умножения эквивалентна цепи, в которой индуктивность и емкость соединены параллельно (рис. 9, б), причем их величины определяются следующими равенствами (A — площадь поперечного сечения диода):

$$\begin{aligned} L_A &= \tau_A / 2J_0 \alpha' A, \\ C_A &= \epsilon_s A / x_A. \end{aligned} \quad (36)$$

Резонансная частота такой цепи равна

$$f_r \equiv \frac{\omega_r}{2\pi} = \frac{1}{2\pi} \sqrt{\frac{2\alpha' v_s J_0}{\epsilon_s}}. \quad (37)$$

Импеданс области умножения имеет простой вид

$$z_A = \frac{x_A}{j\omega\epsilon_s A} \left(\frac{1}{1 - \omega_r^2/\omega^2} \right) = \frac{1}{j\omega C_A} \left(\frac{1}{1 - \omega_r^2/\omega^2} \right). \quad (38)$$

Множитель γ равен

$$\gamma \equiv \frac{\tilde{J}_A}{\tilde{J}} = \frac{1}{1 - \omega^2/\omega_r^2}. \quad (39)$$

Таким образом, тонкая область умножения ведет себя как параллельная резонансная цепочка с собственной частотой, пропорциональной квадратному корню из плотности постоянного тока (выражение (37)).

Область дрейфа и полный импеданс. Подставляя выражение (39) и выражение (25) и интегрируя поле по длине дрейфовой области ($W - x_A$), получим выражение для падения переменного напряжения

$$\tilde{V}_d = \frac{(W - x_A) \tilde{J}}{j\omega\epsilon_s} \left[1 - \frac{1}{1 - \omega^2/\omega_r^2} \left(\frac{1 - e^{-j\theta_d}}{j\theta_d} \right) \right], \quad (40)$$

где θ_d — угол пролета в области дрейфа:

$$\theta_d \equiv \frac{\omega(W - x_A)}{v_s} \equiv \omega\tau_d \quad (41)$$

и

$$\tau_d = \frac{W - x_A}{v_s}. \quad (41a)$$

Определим $C_d = A\varepsilon_s/(W - x_A)$ как емкость области дрейфа. Тогда из выражения (40) мы получим импеданс этой области

$$Z_d = \frac{\tilde{V}_d}{\tilde{I}_A} = \frac{1}{\omega C_d} \left[\frac{1}{1 - \omega^2/\omega_r^2} \left(\frac{1 - \cos \theta_d}{\theta_d} \right) \right] + \\ + \frac{j}{\omega C_d} \left[-1 + \frac{1}{1 - \omega^2/\omega_r^2} \left(\frac{\sin \theta_d}{\theta_d} \right) \right] = R + jX, \quad (42)$$

где R и X — активная и реактивная компоненты импеданса соответственно. Очевидно, что реальная часть импеданса (активное сопротивление) будет отрицательной для всех частот, больших ω_r (рис. 9, z), за исключением нулей при $\theta_d = 2\pi n$, где n — целое. Активное сопротивление положительно при $\omega < \omega_r$, а при уменьшении частоты стремится к

$$R(\omega \rightarrow 0) = \frac{\tau_d}{2C_d} = \frac{(W - x_A)^2}{2A\varepsilon_s v_s}.$$

Низкочастотное малосигнальное сопротивление соответствует конечной величине объемного заряда в области дрейфа, а приведенное выше выражение совпадает с равенством (15).

Полный импеданс диода равен сумме импедансов области лавинного умножения, области дрейфа и сопротивления пассивной области R_s :

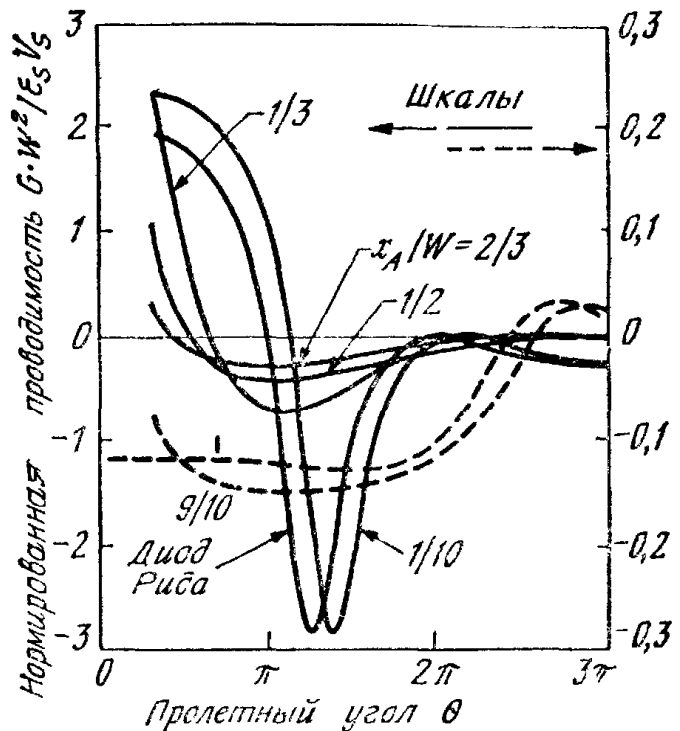
$$Z = \frac{(W - x_A)^2}{2A\varepsilon_s v_s} \left(\frac{1}{1 - \omega^2/\omega_r^2} \right) \frac{1 - \cos \theta_d}{\theta_d^2/2} + R_s + \\ + \frac{j}{\omega C_d} \left[\left(\frac{\sin \theta_d}{\theta_d} - 1 \right) - \left(\frac{\frac{\sin \theta_d}{\theta_d} + \frac{x_A}{W - x_A}}{1 - \omega_r^2/\omega^2} \right) \right]. \quad (43)$$

Равенство (43) записано в форме, удобной для упрощения в случае малых углов пролета θ_d . При $\theta_d < \pi/4$ выражение (43) имеет вид

$$Z = \frac{(W - x_A)^2}{2Av_s\varepsilon_s (1 - \omega^2/\omega_r^2)} + R_s + \frac{j}{\omega C} \left[\frac{1}{(\omega_r^2/\omega^2) - 1} \right], \quad (44)$$

где $C \equiv \varepsilon_s A/W$ — емкость всей области обеднения. Отметим, что первое слагаемое в выражении (44) представляет собой активное сопротивление, которое становится отрицательным при $\omega > \omega_r$. Третье слагаемое — реактивная компонента сопротивления параллельной резонансной цепочки, в которой емкость диода зашунтирована индуктивностью. Реактивное сопротивление носит индуктивный характер при $\omega < \omega_r$ и емкостной при $\omega > \omega_r$. Другими словами, сопротивление становится отрицательным на частотах, соответствующих смене знака реактивной компоненты. Эквивалентная схема и частотные зависимости активной и ре-

Рис. 10. Зависимости нормированной величины проводимости от пролетного угла для шести лавинно-пролетных диодов [30].



активной составляющих импеданса приведены на рис. 9, в и г соответственно.

Импеданс диода Z связан с адмиттансом Y следующим соотношением:

$$Z \equiv \frac{\tilde{V}}{\tilde{I}_A} = R + jX = \frac{1}{G + jB} = \frac{1}{Y}, \quad (45)$$

где G и B — активная и реактивная компоненты адмиттанса соответственно. Теоретические зависимости малосигнальной активной проводимости от пролетного угла для шести ЛПД приведены на рис. 10 [30]. Эти диоды имеют одинаковые области обеднения, но различные по ширине области лавинного умножения ($x_A/W = 1/10, 1/3, 1/2, 2/3, 9/10, 1$). Отметим, что при увеличении толщины слоя лавинного умножения расширяется диапазон частот, в котором сопротивление диода отрицательно. Диод Рида обладает большой отрицательной проводимостью в узком диапазоне частот. Напротив, отрицательная проводимость диода Мисавы ($p-i-n$ -диода с $x_A/W = 1$) мала, но диапазон частот шире и содержит низкие частоты.

10.4. МОЩНОСТЬ И КОЭФФИЦИЕНТ ПОЛЕЗНОГО ДЕЙСТВИЯ

10.4.1. Анализ работы в режиме большого сигнала

На рис. 11 приведены характеристики диода Рида при работе в режиме большого сигнала. Область лавинного умножения с высоким электрическим полем, которое генерирует электронно-дырочные пары, расположена вблизи $p^+ - n$ -перехода (рис. 11, б),

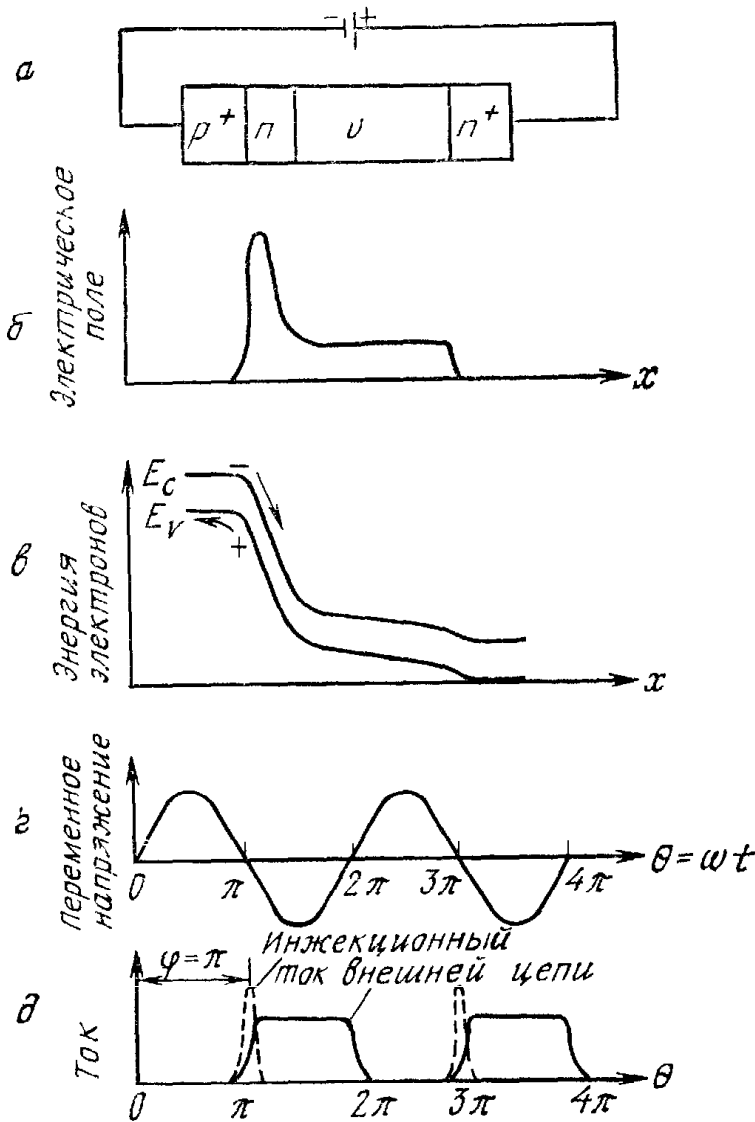


Рис. 11. Характеристики диода при работе в режиме большого сигнала [4].

а — $p^+ - n - v - n^+$ -структура; *б* — распределение поля при пробое; *в* — энергетическая зонная диаграмма; *г* — переменное напряжение; *д* — инжекционный ток и ток внешней цепи для диода Рида.

а слаболегированный v -слой образует область дрейфа с практически однородным полем. Генерируемые дырки быстро попадают в p^+ -область, а генерируемые электроны инжектируются в область дрейфа (рис. 11, б), где совершают работу, равную выделяемой СВЧ-мощности. При периодических колебаниях электрического поля относительно среднего значения (рис. 11, г) коэффициент ионизации, отнесенный к числу носителей заряда, практически мгновенно отслеживает колебания поля. Однако концентрация носителей не повторяет изменения поля, так как генерация зависит от числа уже имеющихся носителей. Даже после того, как поле достигает максимального значения, концентрация носителей продолжает возрастать, поскольку скорость генерации все еще превышает среднее значение. Концентрация носителей максимальна приблизительно в тот момент, когда поле уменьшилось до среднего значения. Таким образом, изменение концентрации инжектированных носителей запаздывает по отношению к переменному напряжению приблизительно на 90° (это так называемое лавинное запаздывание) даже тогда, когда коэффициент ионизации изменяется в фазе с полем. Описанная выше ситуация проиллю-

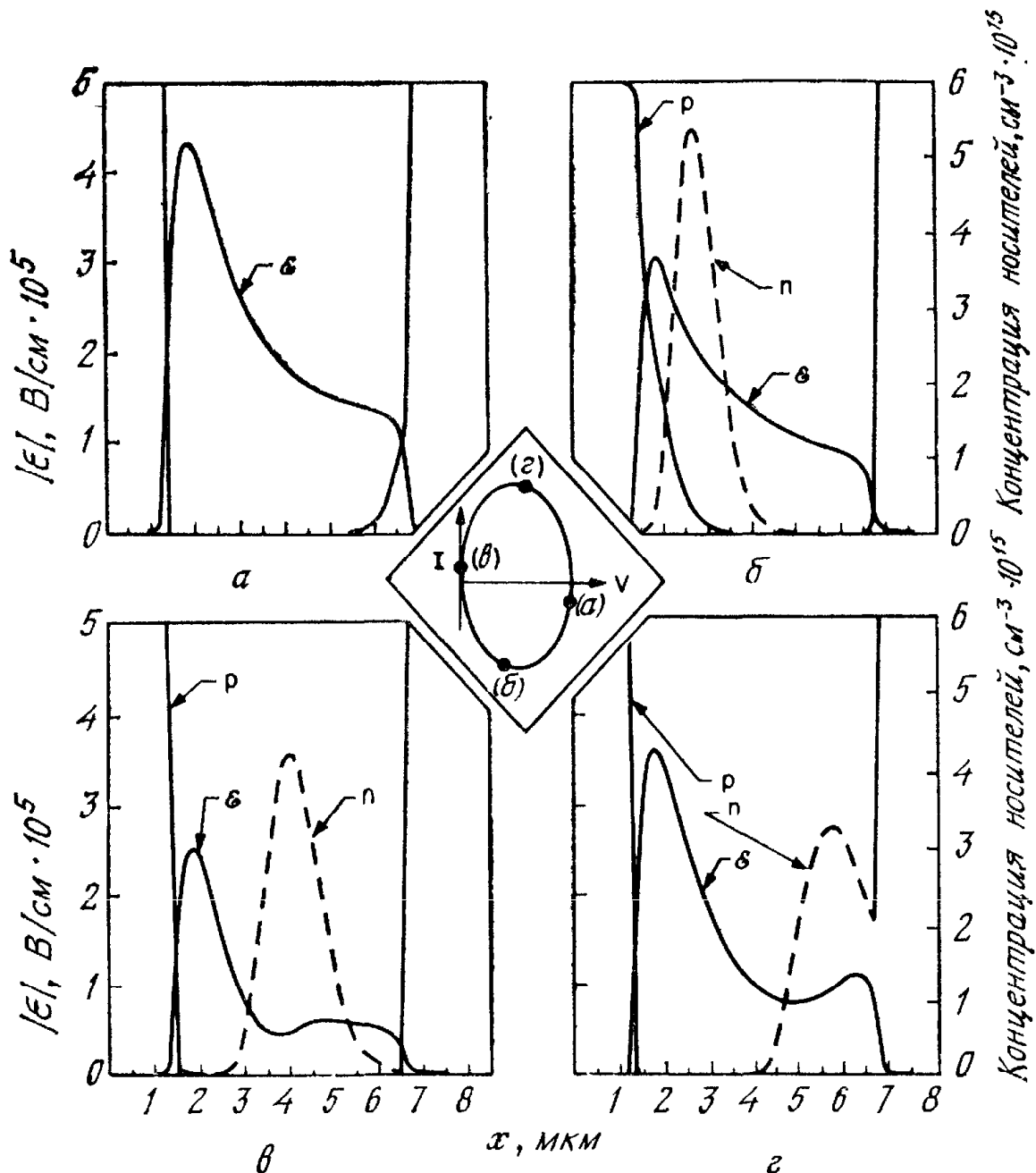


Рис. 12. Результаты расчета на ЭВМ основной моды колебаний ЛПД [31] для четырех случаев, разделенных интервалом в четверть периода.

На вставке показана фазовая траектория в координатах ток — напряжение на зажимах.

стрирована на рис. 11, δ изменением инжекционного тока. Максимум переменного электрического поля (или напряжения) соответствует фазе $\pi/2$, а наибольшая концентрация инжектированных носителей — фазе π (т. е. $\varphi = \pi$, как показано на рис. 11). Инжектированные носители затем попадают в область дрейфа, которую они пролетают со скоростью насыщения, внося «пролетное запаздывание». Форма электрического тока, возбуждаемого во внешней цепи, также приведена на рис. 11, δ . Из сравнения временных зависимостей тока и напряжения следует, что сопротивление диода отрицательно.

Детализированные характеристики диода при работе в режиме большого сигнала можно получить с помощью соотношений (27)—

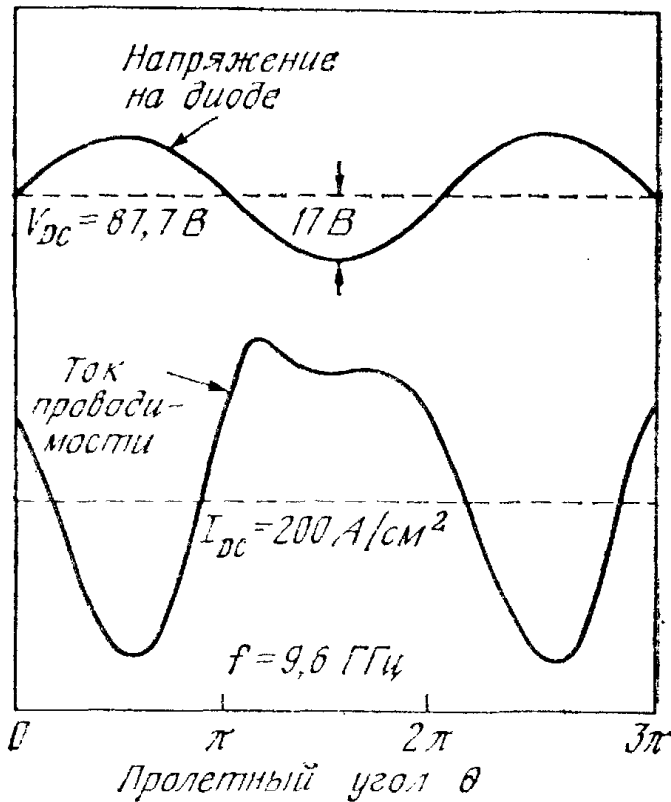


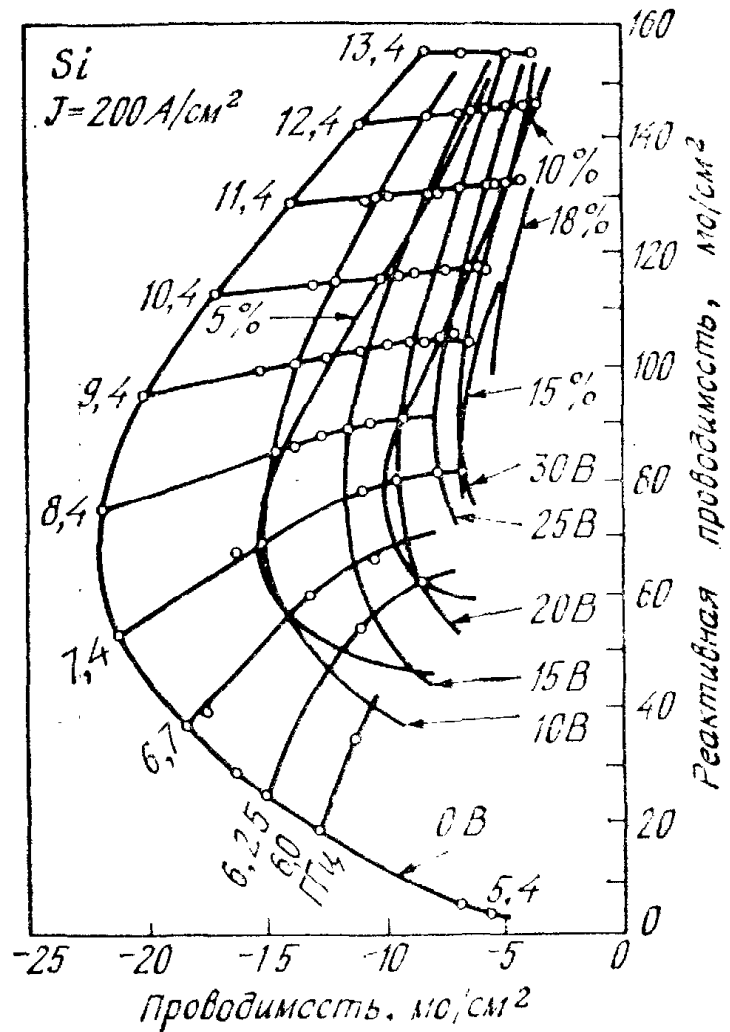
Рис. 13. Зависимости напряжения на зажимах и тока проводимости от пролетного угла для диода с характеристиками, приведенными на рис. 12, при 9,6 ГГц [31].

(29) и приближенных граничных условий. Рассчитанные на ЭВМ зависимости электрического поля и концентраций электронов и дырок от расстояния x в обедненной области кремниевого $p^+ - n - v - n^+$ -диода (при $N_1 = 10^{16} \text{ см}^{-3}$, $b = 1 \text{ мкм}$, $N_2 = 10^{15} \text{ см}^{-3}$ и $W = 6 \text{ мкм}$) в моменты времени, разделенные промежутками в четверть периода, приведены на рис. 12 [31]. Отметим, что 1) генерация электронов и дырок начинается при максимальном напряжении, а через четверть периода импульсы заряда уже полностью сформированы и движутся через соответствующие области дрейфа; 2) дырки быстро покидают активную область, в то время как электроны дрейфуют приблизительно в течение полупериода, индуцируя положительный ток проводимости при отрицательном напряжении; 3) за последующую четверть периода оставшиеся электроны выносятся из активной области, так как напряжение снова приближается к максимальному значению; 4) ток смещения достаточно большой и положительный (соответствует положительному смещению p^+ -контакта), в то время как напряжение на зажимах диода отрицательно; 5) амплитуда модуляции напряжения и тока велика, а эффективность высокая, если область лавинного умножения достаточно широкая.

Зависимости напряжения на зажимах диода и тока проводимости (рис. 12) от пролетного угла приведены на рис. 13. Соотношения между фазами колебаний напряжения на зажимах и тока проводимости близки к идеальным, показанным на рис. 11.

Реактивная проводимость и отрицательная активная проводимость диода в режиме большого сигнала при различных значениях

Рис. 14. Результаты расчетов реактивной проводимости и отрицательной активной проводимости [31].



переменного напряжения приведены на рис. 14. Там же указаны значения к. п. д. Отметим, что реактивная проводимость обычно возрастает с увеличением амплитуды напряжения, а отрицательная активная проводимость убывает. Результаты расчетов показывают, что к. п. д. может достигать 18 %. Эта величина близка к максимальной экспериментально полученной величине к. п. д. кремниевого лавинного-пролетного диода. Тем не менее в некоторых модифицированных диодах Рида из арсенида галлия к. п. д. достигал 40 %. Ниже мы рассмотрим ограничения на величину генерируемой мощности и коэффициент полезного действия.

10.4.2. Температурные ограничения

При низких частотах характеристики ЛДП в непрерывном режиме генерации ограничены в первую очередь температурным фактором, т. е. мощностью, которая может быть рассеяна в виде тепла в полупроводниковом кристалле. Типичный способ монтажа прибора показан на рис. 15, а. Упрощенная структура диода и теплоотвода изображена на рис. 15, б. Полное тепловое сопротив-

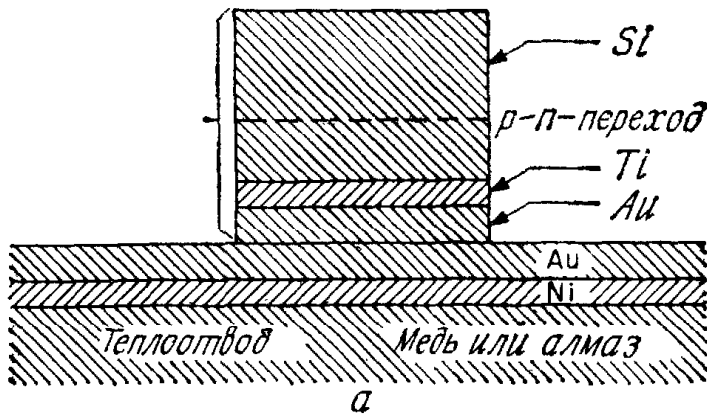
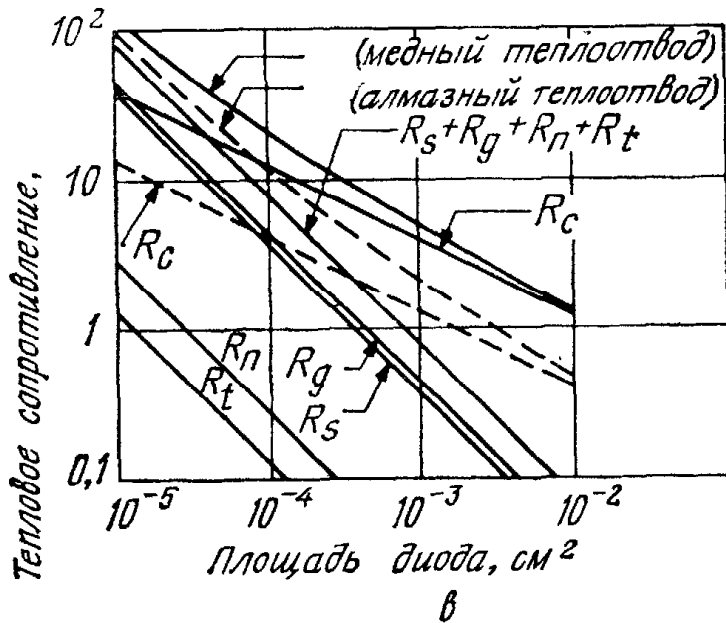
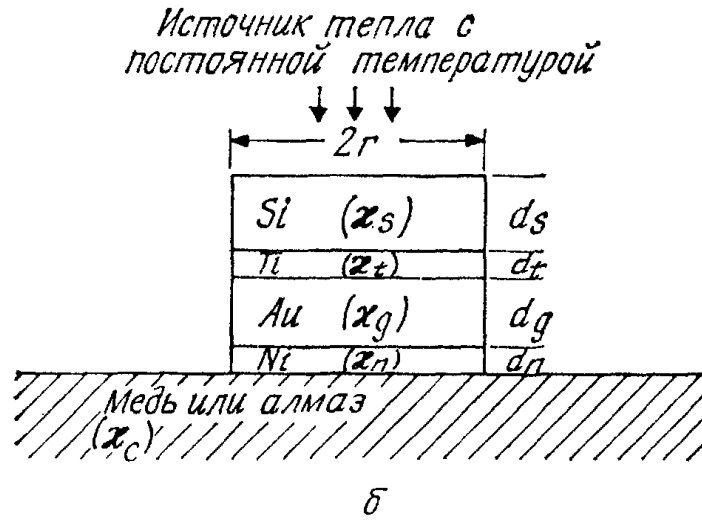


Рис. 15. Диод, смонтированный на медном теплоотводе, в металлических слоях, используемых для термокомпрессированного соединения золото — золото (а), упрощенная структура диода и теплоотвода (б) и зависимости полного теплового сопротивления R_T ($^{\circ}\text{C}/\text{Вт}$) и его компонент от площади диода (в) [32].



ление для кругового источника тепла радиусом r на глубине d_s в кремнии равно [32]

$$R_T = R_s + R_t + R_g + R_n + R_c = \frac{1}{A} \left(\frac{d_s}{\kappa_s} + \frac{d_t}{\kappa_t} + \frac{d_g}{\kappa_g} + \frac{d_n}{\kappa_n} \right) + \frac{1}{4\pi\kappa_c}. \quad (46)$$

Символы поясняются рис. 15, б. Последнее слагаемое равно тепло-вому сопротивлению теплоотвода, занимающего бесконечное полу-

Таблица 1. Теплопроводность и типичные значения толщин слоев различных материалов для диода, работающего в диапазоне частот 15—18 ГГц при 300 К

Материал	Теплопроводность κ , Вт/(см·К)	Толщина d , мкм	d/κ 10^{-4} см ² ·К/Вт
Кремний ¹	0,80	3,00	3,80
Титан	0,16	0,02	0,13
Золото	3,00	12,50	4,20
Никель	0,71	0,20	0,28
Медь	3,90	—	—
Алмаз	20,00	—	—

¹ Теплопроводность для кремния дана при 500 К (по предположению при максимальной рабочей температуре).

пространство. Зависимости различных компонент теплового сопротивления от площади диода приведены на рис. 15, в. Штриховыми линиями показаны составляющая теплового сопротивления алмазного (тип II) теплоотвода и соответствующее полное тепловое сопротивление R_T . Предполагается, что теплопроводность алмаза при 300 К в 3 раза превышает соответствующую величину для меди, а значение теплопроводности кремния соответствует максимальной рабочей температуре 500 К (табл. 1). Из рисунка видно, что алмазный теплоотвод уменьшает тепловое сопротивление приблизительно в 2 раза, а R_T убывает с ростом площади диода.

Мощность P , которая может быть рассеяна в виде тепла в диоде, должна быть равна мощности, которая может быть передана в теплоотвод. Поэтому P равно $\Delta T/R_T$, где ΔT — разность температур p — n -перехода и теплоотвода. Если реактивная проводимость $2\pi fC$ (где $C = A\epsilon_s/W$) остается постоянной, а основной вклад в тепловое сопротивление вносит полупроводник (в предположении, что $d_s \approx W$), то для заданной величины ΔT

$$Pf = \left(\frac{\Delta T}{R_T} \right) f \sim \left[\frac{\Delta T}{(1/A)(W/\kappa_s)} \right] \left(\frac{W}{A\epsilon_s} \right) = \frac{\kappa_s \Delta T}{\epsilon_s} = \text{const.} \quad (47)$$

В этих условиях величина генерируемой в непрерывном режиме мощности уменьшается пропорционально $1/f$.

10.4.3. Ограничения, связанные со свойствами носителей

Максимальная выходная мощность диода на заданной частоте зависит от предельных характеристик¹ полупроводниковых материалов, а также от достижимого в СВЧ-цепях уровня импеданса.

Ограничения, вытекающие из свойств полупроводниковых материалов, связаны со следующими параметрами: 1) величиной напряженности электрического поля \mathcal{E}_m , при которой происходит лавинный пробой; 2) максимально достижимой скоростью дрейфа носителей в полупроводнике v_s — скоростью насыщения.

Максимальное приложенное к полупроводниковому образцу напряжение ограничено напряжением пробоя, которое в случае однородной лавины равно $V_m = \mathcal{E}_m W$, где W — ширина обедненной области. Поскольку ток увеличивает напряженность электрического поля в области пространственного заряда (это следует из уравнения Пуассона), максимальная величина протекающего в образце тока также ограничена лавинными процессами. Если максимальное электрическое поле равно \mathcal{E}_m , максимальный ток определяется выражением $I_m = \mathcal{E}_m \epsilon_s v_s A / W$. Следовательно, верхний предел потребляемой от источника питания мощности

$$P_m = V_m I_m = \mathcal{E}_m^2 \epsilon_s v_s A. \quad (48)$$

Частота, соответствующая углу пролета $\theta = \pi$, равна

$$f = \frac{v_s}{2(W - x_A)} \quad (49)$$

$$\text{и} \quad f \simeq \frac{v_s}{2W} \quad \text{для} \quad W \gg x_A. \quad (49a)$$

Равенство (48) можно записать в виде

$$P_m f^2 \simeq \mathcal{E}_m^2 v_s^2 / 8\pi X_c, \quad (50)$$

где $X_c = (2\pi f C)^{-1}$ — реактивная составляющая импеданса.

Если предположить, что величина импеданса цепи не может быть меньше некоторого значения, то из выражения (50) следует, что максимальная мощность, которая может быть передана подвижным носителям, уменьшается с частотой по закону $1/f^2$. Можно ожидать, что это электронное ограничение будет основным на частотах выше 30 ГГц в кремнии и арсениде галлия.

Следовательно, в режиме непрерывной генерации на относительно низких частотах существенны температурные ограничения ($P \sim 1/f$), а на более высоких частотах — ограничения, связанные со свойствами носителей ($P \sim 1/f^2$). Граничная частота, разделяющая эти интервалы, для данного полупроводника зависит от максимально допустимого увеличения температуры, минимально достижимого уровня импеданса цепи и произведения \mathcal{E}_m на v_s .

Для рабочих температур в диапазоне 150—200 °С максимальная напряженность электрического поля \mathcal{E}_m в кремнии почти на 10 % меньше, чем в арсениде галлия, а скорость насыщения v_s в кремнии почти в 2 раза больше. Поэтому на частотах, больших 30 ГГц, выходная мощность кремниевого ЛПД должна по рас-

четам приблизительно в 3 раза превышать выходную мощность ЛПД из арсенида галлия, работающего на тех же частотах [33]. Можно полагать, что в области субмиллиметровых длин волны выгоднее использовать диоды Мисавы, в которых электрическое поле однородно, так как они обладают широким диапазоном частот, в котором сопротивление отрицательно, а пролетные эффекты не играют такой важной роли в появлении отрицательного сопротивления, как в диодах Рида [34].

При работе в импульсном режиме, когда температурные эффекты можно не учитывать (т. е. при очень коротких импульсах), предельно достижимая мощность на любых частотах $P \sim 1/f^2$.

10.4.4. Ограничения на коэффициент полезного действия

Для эффективной работы ЛПД необходима генерация в области лавинного умножения максимально допустимой величины заряда Q_m при условии, что электрическое поле в области дрейфа не упадет настолько, что скорость носителей станет меньше скорости насыщения. Движение заряда Q_m через область дрейфа приводит к появлению переменной составляющей напряжения на диоде mV_D , где m — коэффициент модуляции ($m \leq 1$) и V_D — среднее падение напряжения на области дрейфа. На оптимальной частоте $\sim v_s/2W$ движение заряда Q_m также приводит к возникновению переменного тока проводимости, который отстает по фазе на угол φ от переменного напряжения на диоде. Средняя величина тока проводимости равна постоянному току J_0 . Следовательно, ток проводимости может изменяться лишь в пределах $0-2J_0$. В случае прямоугольного импульса тока проводимости и синусоидального изменения напряжения на диоде по амплитуде и фазе к. п. д. генерации СВЧ-мощности η равен [31, 35]

$$\eta = \frac{\text{Мощность генерируемого сигнала}}{\text{Мощность, потребляемая от источника питания}} = \frac{(2J_0/\pi)(mV_D)}{J_0(V_A + V_D)} |\cos \varphi| = \frac{2m}{\pi} \frac{1}{1 + V_A/V_D} |\cos \varphi|, \quad (51)$$

где V_A — падение напряжения на области лавинного умножения, а сумма V_A и V_D равна полной величине приложенного постоянного напряжения. Угол φ равен запаздыванию фазы тока проводимости. При идеальных условиях $\varphi = 180^\circ$, а $|\cos \varphi| = 1$. Для диодов с двумя дрейфовыми областями напряжение V_D в выражении (51) следует заменить на $2V_D$. Здесь мы не учли вклад области умножения, поскольку напряжение на этом слое носит индуктивный характер и сдвинуто по фазе относительно тока проводимости на $+\pi/2$. Ток смещения сдвинут по фазе относительно напряжения на $+\pi/2$ и поэтому тоже не даст вклада в среднюю генерируемую мощность.

Из выражения (51) следует, что для повышения к. п. д. диода необходимо увеличивать коэффициент модуляции, оптимизировать угол запаздывания фазы так, чтобы он был близок к 180° , и уменьшать отношение V_A/V_D . Тем не менее для того, чтобы лавинный процесс начинался быстро, напряжение V_A должно быть достаточно велико. При уменьшении отношения V_A/V_D ниже некоторого оптимального значения происходит резкое снижение к. п. д. диода практически до нуля [31].

Если зависимость скорости от поля такая, что насыщение дрейфовой скорости происходит при очень малых значениях напряженности электрического поля, то коэффициент модуляции $m \approx 1$ и скорость носителей все время равна скорости насыщения. В арсениде галлия n -типа эффективное насыщение дрейфовой скорости происходит при полях $\sim 10^3$ В/см, а для кремния n -типа — при $\sim 2 \cdot 10^4$ В/см. Следовательно, в диодах из арсенида галлия n -типа можно получить значительно бóльшие амплитуды переменного напряжения и более высокие к. п. д. [36—38].

Для того чтобы оценить оптимальное значение отношения V_A/V_D , предположим, что пролетная частота, определяемая равенством (49), приблизительно на 20 % превышает резонансную частоту (выражение (37)):

$$f = \frac{v_s}{2(W - x_A)} = 1,2f_r = \frac{1,2}{2\pi} \sqrt{\frac{2\alpha' v_s J_0}{\epsilon_s}}. \quad (52)$$

Напряжение V_D равно

$$V_D = \langle \mathcal{E}_D \rangle (W - x_A) = \langle \mathcal{E}_D \rangle v_s / 2f, \quad (53)$$

где $\langle \mathcal{E}_D \rangle$ — средняя напряженность электрического поля в области дрейфа. При 100 %-ной модуляции тока ($J_0 = J_{dc} = J_{ac}$) максимальная величина заряда $Q_m = m\epsilon_s \langle \mathcal{E}_D \rangle$ определяет плотность тока:

$$J_0 = Q_m f = m\epsilon_s \langle \mathcal{E}_D \rangle f. \quad (54)$$

Если коэффициент ионизации $\alpha \sim \mathcal{E}^\zeta$, величина α' может быть получена следующим образом:

$$\alpha' \equiv \frac{d\alpha}{d\mathcal{E}} = \frac{\zeta\alpha}{\mathcal{E}} \simeq \frac{\zeta(W - x_A)\alpha}{V_D}. \quad (55)$$

Комбинируя выражения (52)—(55), получим соотношение [35]

$$(V_A/V_D)_{\text{opt}} \simeq 4m \left(\frac{1,2}{2\pi} \right)^2 \zeta \alpha x_A. \quad (56)$$

При частоте $f \simeq 10$ ГГц для арсенида галлия оптимальное значение $V_A/V_D = 0,65$ и $m \simeq 1$, в то время как для кремния оно равно $\sim 1,1$ при $m \simeq 1/2$.

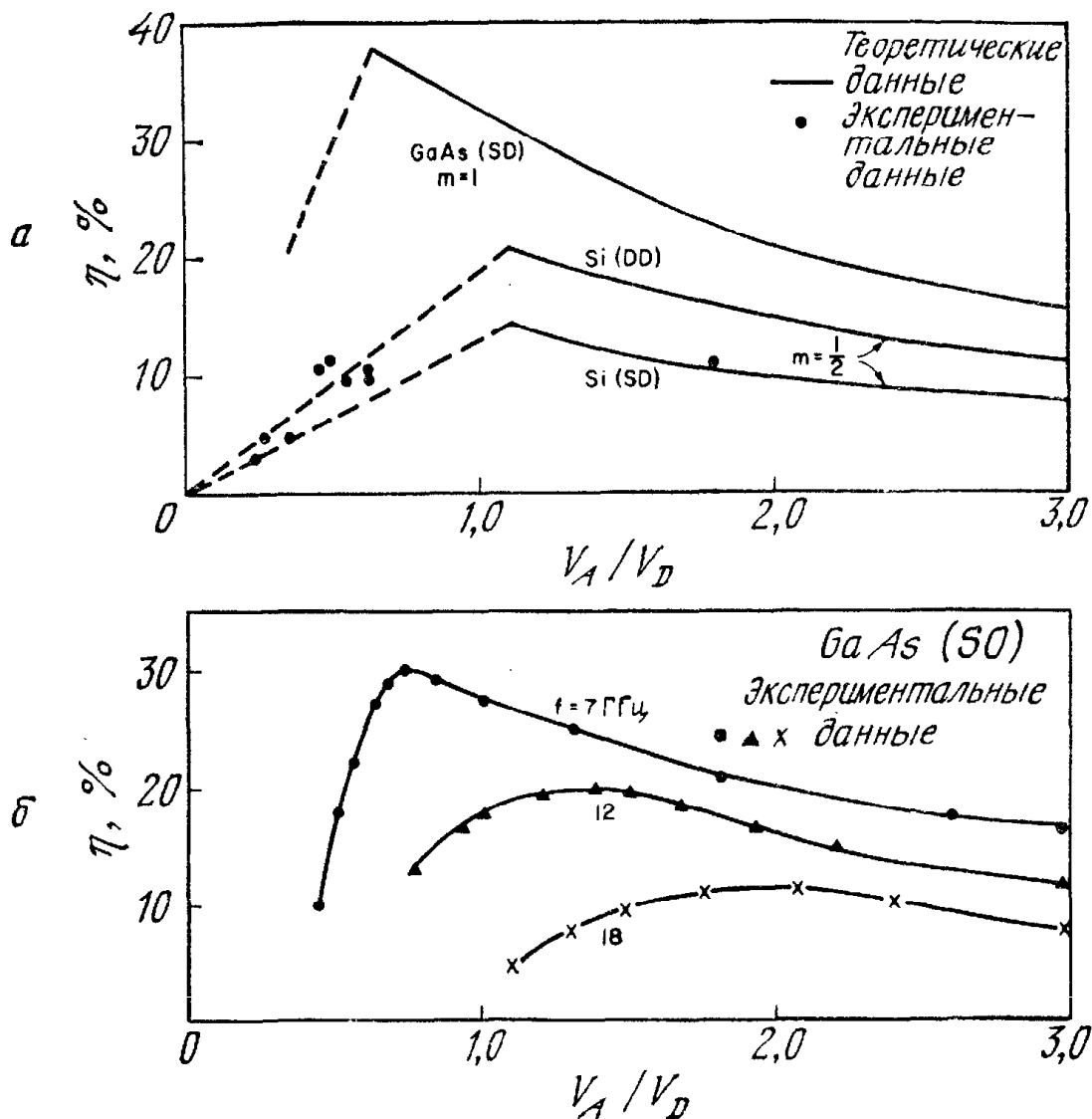


Рис. 16. Зависимость к. п. д. от V_A/V_D для диодов из кремния и арсенида галлия (а) [35] и экспериментально полученные значения к. п. д. в зависимости от V_A/V_D в диодах с двухслойной базой из арсенида галлия (б) [39].

SD — одна область дрейфа; DD — две области дрейфа.

На рис. 16, а приведены зависимости к. п. д. диодов из кремния и арсенида галлия от V_A/V_D . При вычислении максимальных к. п. д. использованы оптимальные значения величин, которые мы получили выше. Ожидаемое значение максимального к. п. д. для кремниевого диода с одной дрейфовой областью составляет около 15 %, с двумя дрейфовыми областями 21 %, а для диодов из арсенида галлия с одной дрейфовой областью — около 38 %. Предыдущие оценки совпадают с экспериментально полученными результатами. На высоких частотах происходит увеличение оптимального значения V_A/V_D , что приводит к уменьшению максимального к. п. д. Приведенные на рис. 16, б экспериментальные результаты для диода из арсенида галлия с одной областью дрейфа вполне согласуются с полученными нами выше оценками [39].

На практике множество различных факторов служит причиной уменьшения к. п. д. лавинно-пролетных диодов. К ним можно отнести влияние пространственного заряда носителей, обратного

тока насыщения, последовательного сопротивления, скин-эффект, насыщение коэффициента ионизации, туннельный эффект, конечность «времени лавинного отклика» и эффекты, связанные с накоплением неосновных носителей.

Влияние пространственного заряда [40] иллюстрируется рис. 17. Генерируемые дырки уменьшают напряженность электрического поля (рис. 17, а), что может привести к преждевременному прекращению процесса лавинного умножения и уменьшению угла запаздывания фазы, который в оптимальных условиях равен 180° . По мере того как дырки дрейфуют вправо (рис. 17, б), пространственный заряд может снизить электрическое поле на левом краю зарядового пакета до значений, меньших, чем это необходимо для насыщения скорости. Такое уменьшение в свою очередь приведет к изменению формы импульса тока и снижению уровня мощности, генерируемой на пролетных частотах.

Влияние обратного тока насыщения [41] иллюстрируется рис. 18. Большой ток насыщения вызывает быстрое нарастание

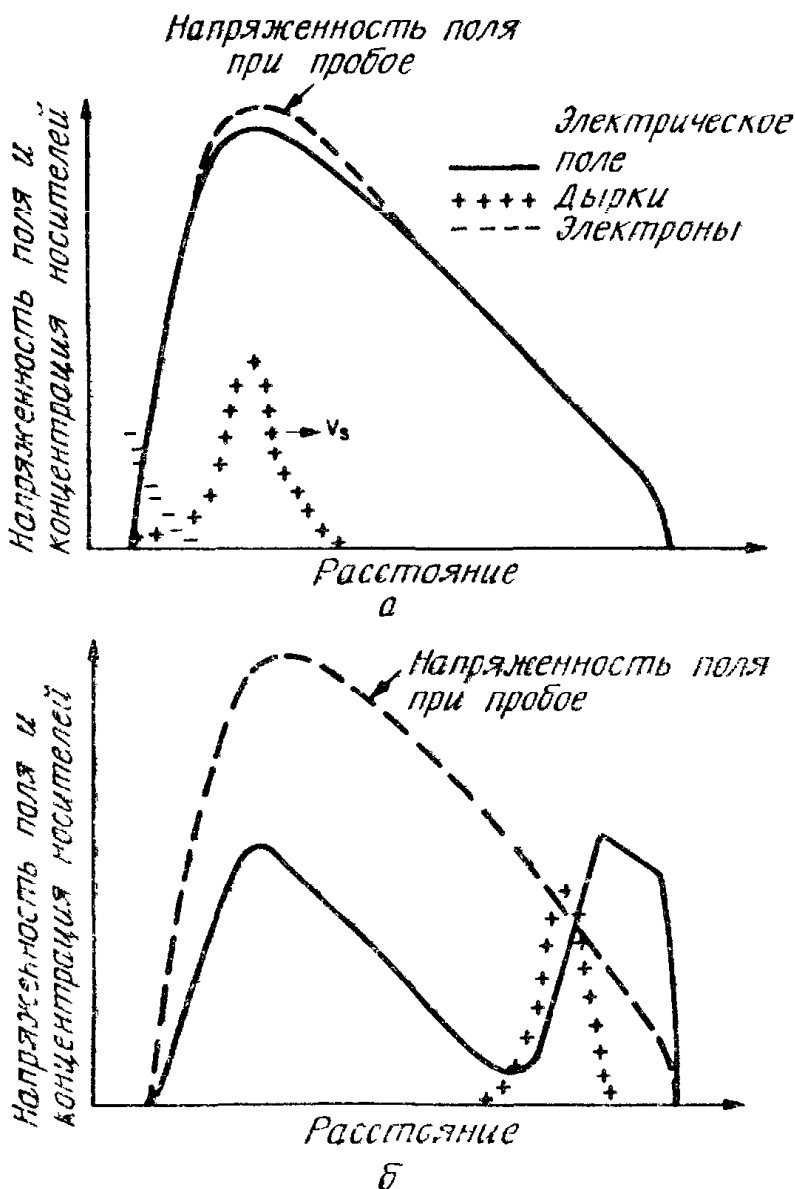


Рис. 17. Мгновенные распределения электрического поля и носителей заряда в диоде Рида [40]
 а — процесс лавинного умножения завершен, заряд начинает двигаться через диод;
 б — пролет заряда почти закончен. Видно, что пространственный заряд носителей сильно снижает электрическое поле

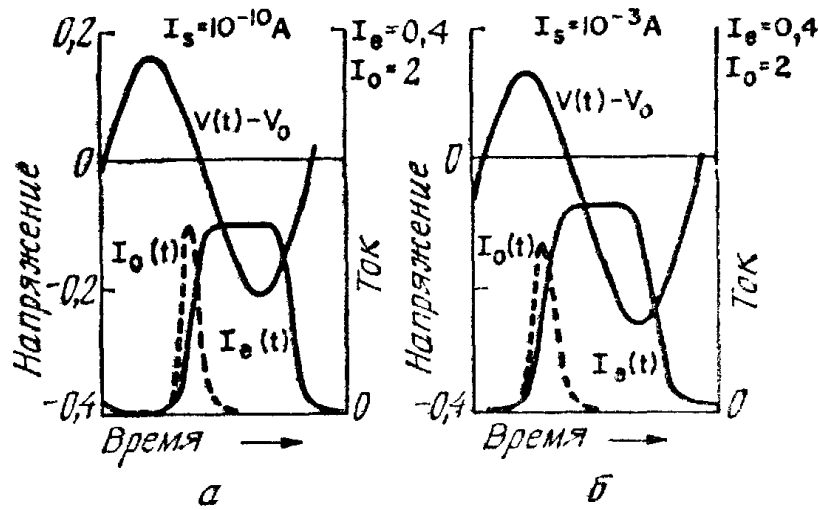


Рис. 18. Фазовые сдвиги для инжекционного тока $I_0(t)$ и тока внешней цепи $I_e(t)$ по отношению к переменному напряжению на диоде. Результаты расчета на ЭВМ для тока насыщения 10^{-10} А и 10^{-3} А [41].

лавины, приводя к уменьшению соответствующего запаздывания фазы. Для $I_s = 10^{-10}$ А $|\cos \varphi|$ близок к 1 (рис. 18, а), но для $I_s = 10^{-3}$ А $|\cos \varphi|$ равен лишь 0,43 (рис. 18, б), а к. п. д. меньше примерно в два раза. Инжекция неосновных носителей из «плотных» омических контактов также приводит к увеличению обратного тока насыщения и уменьшению к. п. д.

Эффект неполного «перекрывания» эпитаксиального слоя подложки обедненной областью [42, 43] иллюстрируется рис. 19. Наличие области несмыкания приводит к дополнительному последовательному сопротивлению и снижает отрицательное сопротивление диода. Отметим, что этот эффект в арсениде галлия вы-

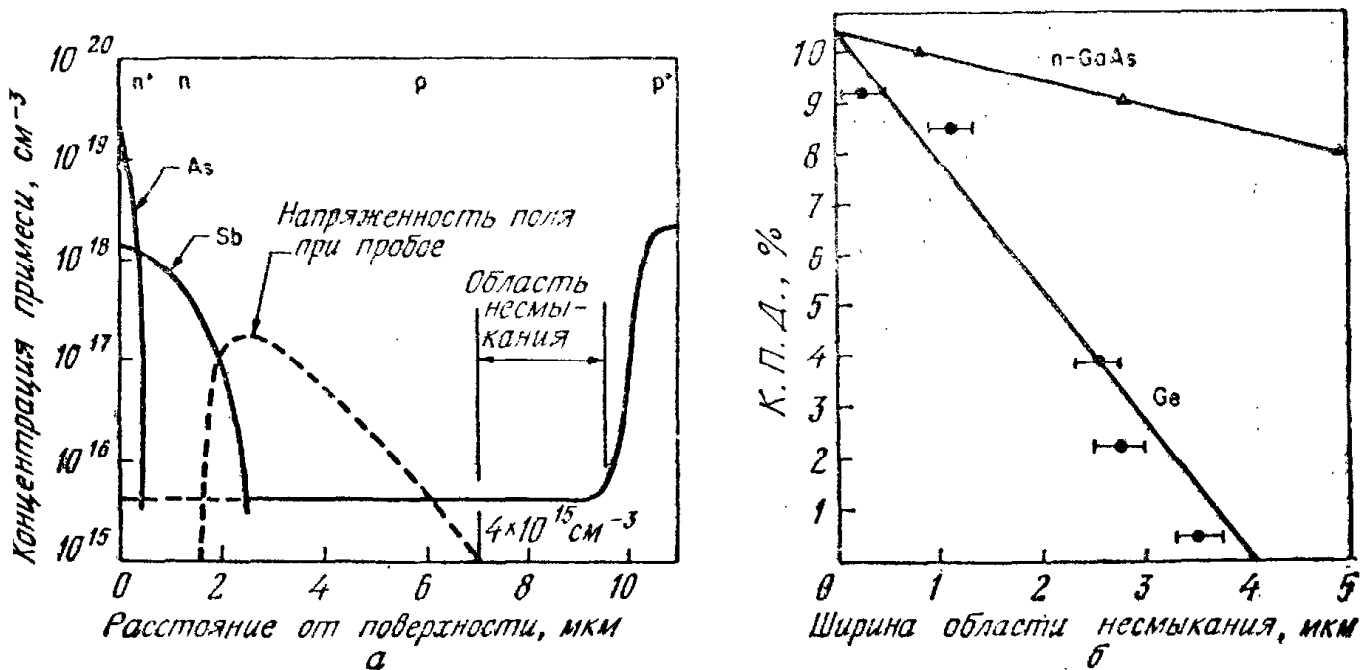


Рис. 19. Распределение электрического поля при пробое в германиевом ЛПД (а) и к. п. д. в зависимости от ширины области несмыкания в эпитаксиальном слое для n -GaAs и Ge-диодов (б) [42, 43]

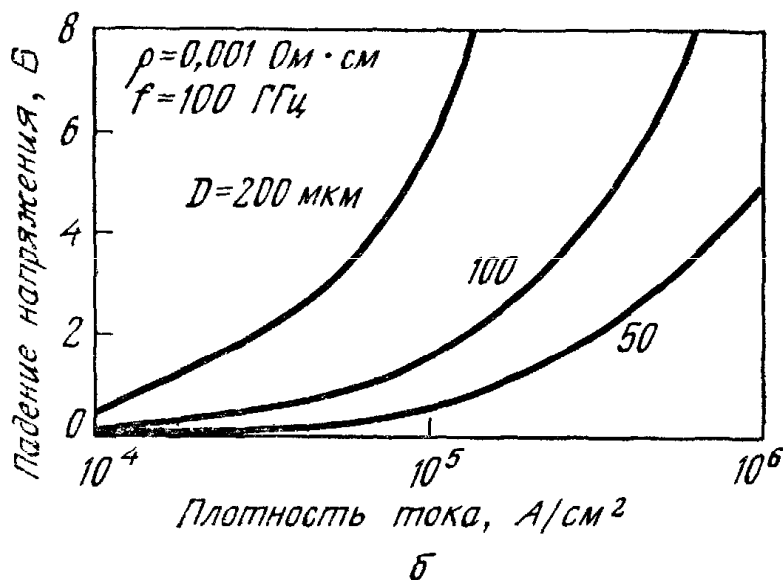
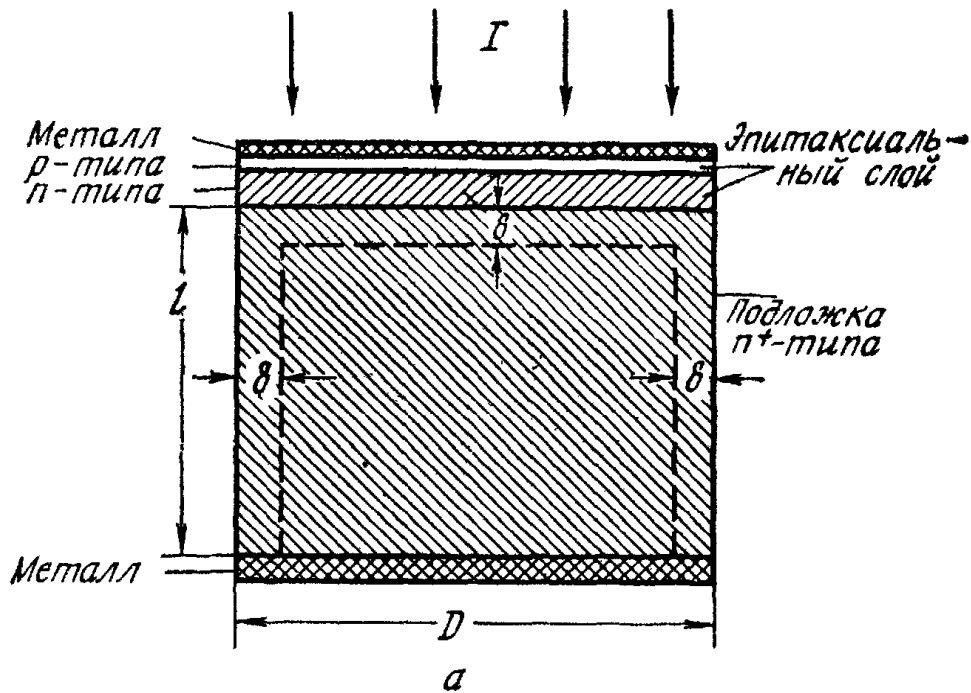


Рис. 20. Скин-эффект [44].

a — протекание тока ограничено тонким поверхностным слоем толщиной δ , что приводит к неоднородности и резистивным потерям; *б* — рассчитанные зависимости падения напряжения на подложке при 100 ГГц для диодов с различными диаметрами.

ражен слабее, поскольку при малых электрических полях в GaAs подвижность носителей существенно выше.

Рассмотрим теперь влияние скин-эффекта [44] (рис. 20). Если рабочая частота ЛПД лежит в диапазоне миллиметровых длин волн, область протекания тока ограничена скин-слоем глубиной δ на поверхности подложки. Поэтому эффективное сопротивление подложки возрастает и возникает падение напряжения в плоскости, перпендикулярной направлению тока (рис. 20, б). Это в свою очередь приводит к неоднородному распределению плотности тока в данной плоскости и дополнительному последовательному сопротивлению, что также снижает к. п. д.

При достаточно высоких рабочих частотах ширина обедненной области очень мала, а необходимое для начала ударной ионизации электрическое поле велико. При сильных полях имеют место два основных эффекта. Во-первых, коэффициент ионизации медленно изменяется, что приводит к уширению импульса инжекционного тока [45], изменению формы тока на выходе и уменьшению к. п. д. Во-вторых, туннельный ток может стать доминирующим. Поскольку этот ток находится в фазе с приложенным полем, запаздывание в 180° , характерное для лавинного умножения, в этом случае не может быть достигнуто. Туннельные механизмы изучались в связи с туннельно-пролетным режимом работы [4, 46, 47]. Предполагается, что шум туннельно-пролетного диода будет меньше, чем шум ЛПД, однако его выходная мощность и к. п. д. также будут значительно меньше.

Другим ограничивающим фактором при работе в субмиллиметровом диапазоне длин волны является конечное запаздывание отклика коэффициента ионизации на изменение электрического поля. Для кремния «время лавинного отклика» $\tau_i < 10^{-13}$ с. Предполагается, что кремниевые ЛПД будут эффективно работать на частотах 300 ГГц и выше, так как τ_i мало по сравнению с временем пролета в диапазоне субмиллиметровых волн. Однако в арсениде галлия τ_i на порядок больше, чем в кремнии [48]. Такие большие значения τ_i могут ограничить рабочий диапазон GaAs-ЛПД частотами, меньшими 100 ГГц.

Эффекты накопления неосновных носителей в $p^+ - n$ - или $n^+ - p$ -диодах, возникающие из-за диффузии генерируемых в активном слое электронов или дырок в нейтральную p^+ - или n^+ -область, также приводят к снижению к. п. д. В то время как основные носители пролетают область дрейфа, неосновные носители накапливаются в нейтральной области, а в следующую фазу цикла диффундируют обратно в активную область, вызывая преждевременную лавину, которая искажает соотношение фаз между током и напряжением.

10.4.5. Необратимый термический пробой диода вследствие образования шнура тока

Причиной перегорания диода может быть не только перегрев, но и неоднородное распределение тока по площади при образовании шнуров с высокой плотностью тока. Такая ситуация возможна в диоде в режиме с отрицательным сопротивлением по постоянному току, когда локальные области максимальной плотности тока характеризуются минимальным напряжением пробоя. Это причина выхода из строя $p-i-n$ -диодов. Объемный заряд движущихся в области дрейфа носителей препятствует возникновению режима с низкочастотным отрицательным сопротивлением

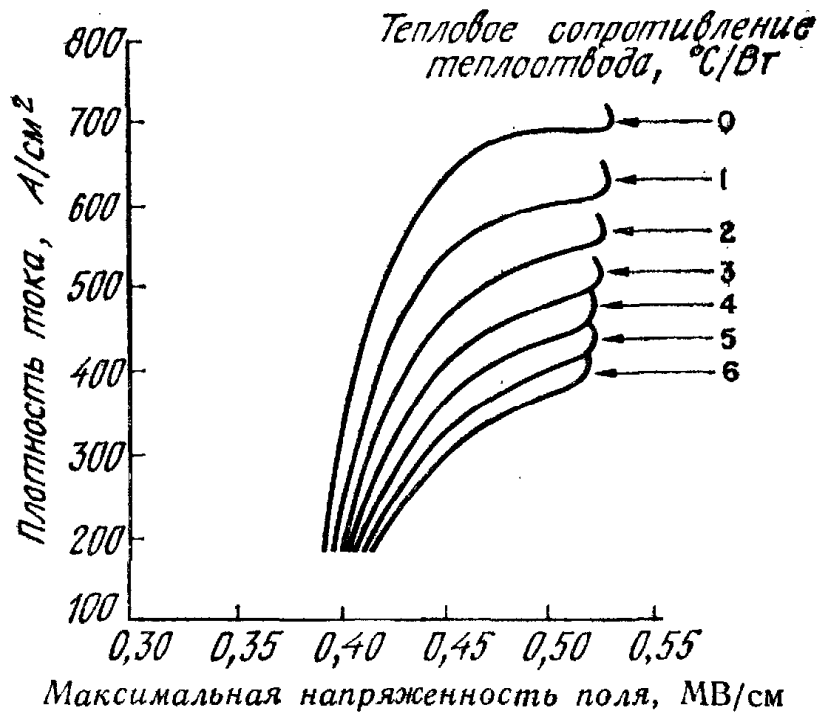


Рис. 21. Зависимость плотности тока от максимальной напряженности электрического поля в кремниевых барьерах Шоттки с $N_D = 6,2 \cdot 10^{15} \text{ см}^{-3}$; $\Phi_{Вр} = 0,79 \text{ В}$ при температуре 22 °С [49].

и, следовательно, перегоранию из-за образования шнура тока. Однако диоды, обладающие положительным статическим сопротивлением при малых токах, могут иметь отрицательное динамическое сопротивление и перегорать при больших токах.

На рис. 21 приведена зависимость плотности тока от максимальной напряженности электрического поля в GaAs-диоде с барьером Шоттки [49] при концентрации примеси $6,2 \cdot 10^{15} \text{ см}^{-3}$ и высоте барьера $0,79 \text{ В}$. Если в режиме непрерывной генерации увеличивается ток прибора, то возрастает и температура в переходе. В конце концов после достижения поворотной точки зависимости тока от поля, за которой начинается область отрицательного дифференциального сопротивления, диод выходит из строя из-за образования шнура тока с высокой плотностью. Отметим, что диод может работать при больших плотностях тока, предшествующих точке поворота, если использовать теплоотвод с малым тепловым сопротивлением.

10.5. ШУМЫ ЛПД

Основной причиной шумов ЛПД является статистическая природа генерации электронно-дырочных пар в области лавинного умножения. Поскольку шумы определяют минимальный уровень СВЧ-сигнала, который может быть усилен, теория шумов в ЛПД очень важна.

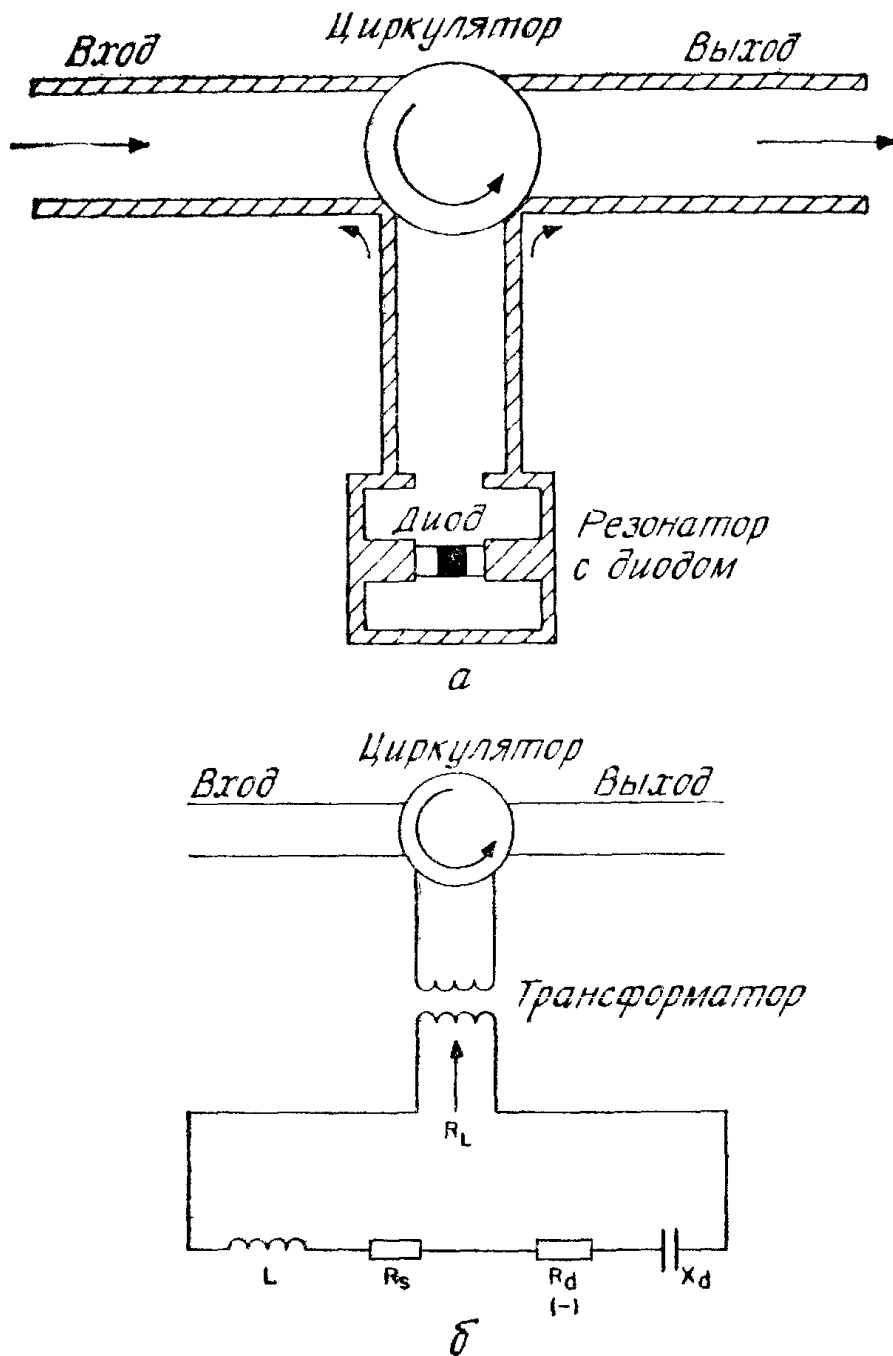


Рис. 22. ЛПД, помещенный в резонатор (а), и эквивалентная схема (б) [50].

Для усиления сигнала лавинно-пролетный диод можно поместить в резонатор, связанный с передающей линией [50], которая в свою очередь связана с входным и выходным трактами с помощью циркулятора (рис. 22, а). На рис. 22, б приведена эквивалентная схема, используемая при малосигнальном анализе. Ниже мы введем две характеристики: коэффициент шума и шумовое отношение. Коэффициент шума определим следующим образом:

$$\begin{aligned}
 F &= 1 + \frac{\text{Выходная мощность шума, возникающего в усилителе}}{\text{Коэффициент усиления мощности} \times (kT_0 B_1)} = \\
 &= 1 + \frac{\bar{I}_\pi^2 R_L}{P_G k T_0 B_1}, \quad (57)
 \end{aligned}$$

где P_G — коэффициент усиления мощности, R_L — сопротивление нагрузки, k — постоянная Больцмана, $T_0 = 290$ К, B_1 — шумовая полоса, \bar{I}_n^2 — средний квадрат шумового тока диода, индуцируемого в замкнутой цепи (рис. 22, б). Шумовое отношение определим как

$$M \equiv \frac{\bar{I}_n^2}{4kT_0GB_1} \quad (58)$$

или

$$M \equiv \frac{\bar{V}_n^2}{4kT_0(-Z_{\text{real}})B_1}, \quad (59)$$

где G — отрицательная действительная компонента адмиттанса, $-Z_{\text{real}}$ — действительная часть импеданса диода и \bar{V}_n^2 — средний квадрат шумового напряжения. Отметим, что коэффициент шума и шумовое отношение зависят от среднеквадратичного шумового тока (или среднеквадратичного шумового напряжения). Ниже мы покажем, что при частотах, больших резонансной f_r , шум диода уменьшается, однако при этом уменьшается также отрицательное сопротивление. Поэтому приемлемой величиной, характеризующей работу диода в качестве усилителя, является шумовое отношение, и мы заинтересованы в том, чтобы оно было минимальным.

Коэффициент шума в усилителе с большим коэффициентом усиления равен [50]

$$NF = 1 + \frac{qV_A/kT_0}{4\zeta\tau_A^2(\omega^2 - \omega_r^2)}, \quad (60)$$

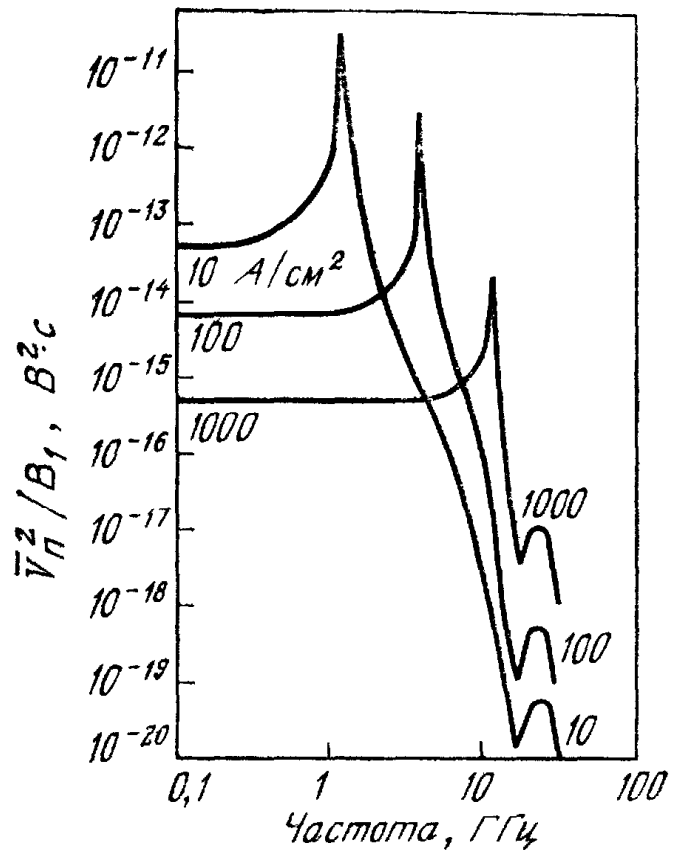
где ζ — показатель степени в зависимости $\alpha \sim \mathcal{E}^\zeta$, τ_A и V_A — время пролета области умножения и падение напряжения на ней соответственно, f_r — резонансная частота (разд. 10.3). Соотношение (60) получено в приближении тонкой области умножения и равенства коэффициентов ионизации электронов и дырок. При $\zeta = 6$ (для кремния), $\omega = 2\omega_r$ и $V_A = 3$ В коэффициент шума на частоте $f = 10$ ГГц по расчетам равен 11 000, или 40,5 дБ.

При реальных значениях коэффициентов ионизации ($\alpha_n \neq \alpha_p$ для кремния) и произвольном профиле распределения примеси выражение для среднего квадрата шумового напряжения на низких частотах имеет вид [51]

$$\bar{V}_n^2 = \frac{2qB_1}{J_0A} \left(\frac{1 + W/x_A}{\alpha'} \right)^2 \sim \frac{1}{J_0}, \quad (61)$$

где $\alpha' = d\alpha/d\mathcal{E}$. На рис. 23 приведена зависимость \bar{V}_n^2/B_1 от частоты для кремниевого ЛПД с $A = 10^{-4}$ см², $W = 5$ мкм и $x_A = 1$ мкм. Отметим, что, как следует из выражения (61), при низких частотах средний квадрат шумового напряжения \bar{V}_n^2

Рис. 23. Средний квадрат шумового напряжения, отнесенный к ширине полосы, в зависимости от частоты для кремниевого ЛПД [51].



обратно пропорционален плотности постоянного тока. Вблизи резонансной частоты (которая пропорциональна $\sqrt{J_0}$) \bar{V}_n^2 достигает максимального значения, а затем быстро убывает приблизительно как четвертая степень частоты. Таким образом, уровень шума может быть снижен, если работать при малых токах и на частотах, превышающих лавинную. Эти условия противоположны режиму большой мощности и высокого к. п. д., и поэтому необходимо выбирать промежуточный режим, оптимальный для конкретных применений.

На рис. 24 приведены типичные теоретические и экспериментально полученные [52] зависимости шумового отношения для ЛПД из арсенида галлия. На пролетной частоте 6 ГГц шумовое отношение приблизительно равно 32 дБ. Минимальное шумовое отношение, равное 22 дБ, соответствует частотам, превышающим приблизительно в 2 раза пролетную частоту. Шумовое отношение для ЛПД из арсенида галлия значительно ниже, чем для кремниевых диодов. В табл. 2 приведены шумовые отношения для диодов из германия, кремния и арсенида галлия. Шумы в режимах усиления и генерации даны для идеальной цепи без потерь при частотах, соответствующих максимальной эффективности генерации без гармонической настройки.

Основная причина низкого уровня шумов в GaAs-ЛПД состоит в том, что при заданной напряженности электрического поля коэффициенты ионизации электронов и дырок в арсениде галлия одинаковы, в то время как в кремнии они существенно различа-

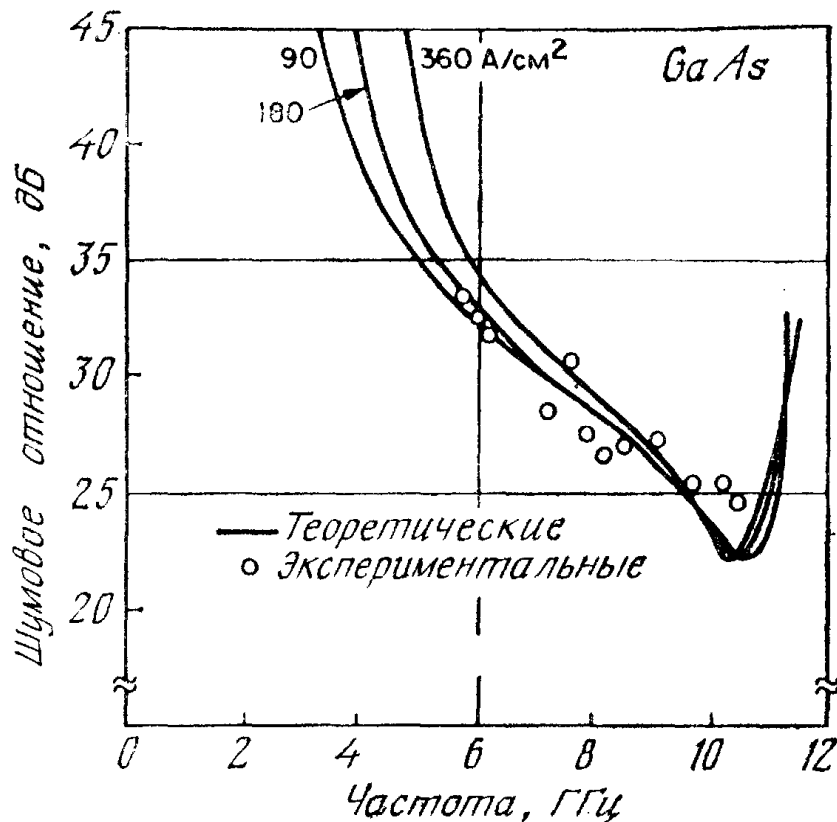


Рис. 24. Шумовое отношение для GaAs-ЛПД с пролетной частотой 6 ГГц [52].

ются. Используя интеграл лавинного умножения, можно показать, что для получения больших коэффициентов умножения M средняя длина ионизации $1/\langle\alpha\rangle$ должна приблизительно равняться ширине слоя умножения x_A , если $\alpha_n = \alpha_p$, и $x_A/\ln M$, если $\alpha_n \gg \alpha_p$. Так, при заданной ширине x_A в кремнии будет происходить значительно больше актов ионизации, и, следовательно, шумы ЛПД будут больше.

На рис. 25 приведены зависимости выходной мощности от шумового отношения для кремниевых и GaAs-ЛПД на частоте 6 ГГц (в случае частотно-модулированного шума) [53]. Мощность нормирована на 1 мВт, т. е. равна $10 \log (P \cdot 10^3)$ дБ, где мощность P выражена в ваттах. Диоды помещались в настроенный на определенную частоту коаксиальный резонатор, причем сопротивле-

Таблица 2. Шумовые отношения для лавинно-пролетных диодов

Полупроводник	Ge	Si	GaAs
Шумовое отношение в режиме малого сигнала, дБ	30	40	25
Шумовое отношение в режиме большого сигнала, дБ	40	55	35

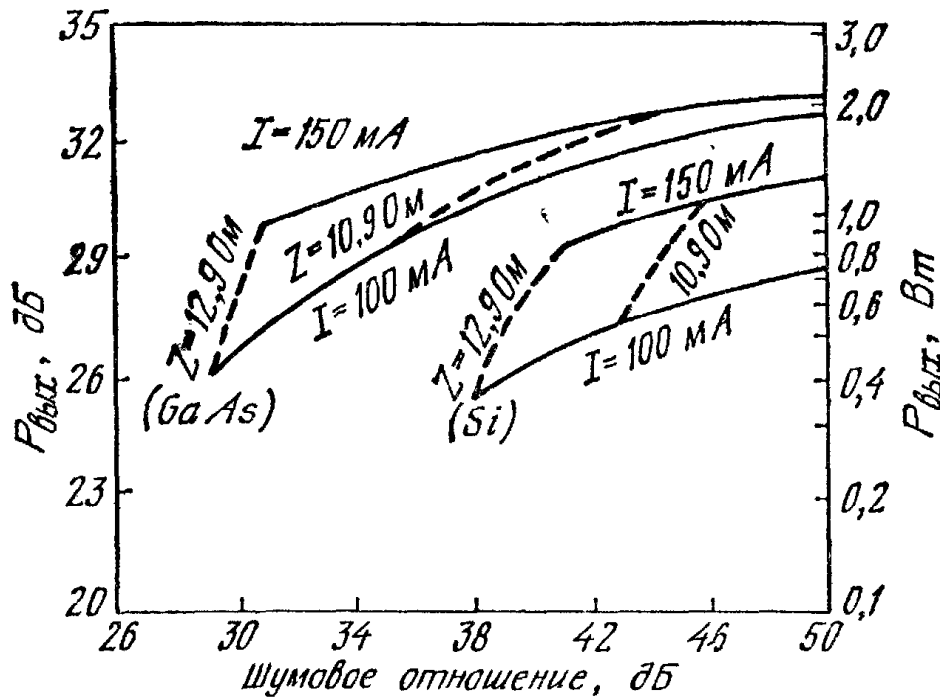


Рис. 25. Зависимости выходной мощности от шумового отношения для генератора с синхронизованной фазой. Величина синхронизирующей мощности поддерживалась постоянной на уровне 4 дБ. Показаны линии, соответствующие постоянной нагрузке Z и постоянному току диода I [53].

ние нагрузки изменялось постепенно при помощи импедансного трансформатора Z . При максимальной выходной мощности шумовое отношение относительно мало. Шумовое отношение можно уменьшить за счет небольшого понижения уровня выходной мощности. Еще раз отметим, что при заданном уровне мощности (например, 1 Вт или 30 дБ) шум ЛПД из арсенида галлия приблизительно на 10 дБ меньше шума кремниевого ЛПД.

10.6. КОНСТРУКЦИЯ ПРИБОРА И ХАРАКТЕРИСТИКИ

10.6.1. Изготовление диода

Основываясь на соображениях, изложенных в разд. 10.4, можно сделать вывод, что лавинно-пролетные диоды обычно конструируются так, чтобы выходная мощность и к. п. д. были максимальными. На рис. 26 показаны некоторые типичные структуры мощных ЛПД.

Диодная структура, изображенная на рис. 26, а, получена с помощью двойной эпитаксии или диффузии в эпитаксиальный слой [54]. Для снижения последовательного сопротивления используют n^+ -подложку. Толщину эпитаксиального слоя следует контролировать, чтобы при пробое отсутствовала область несмыкания. При работе на высоких частотах даже подложка n^+ -типа должна быть достаточно тонкой (порядка нескольких микромет-

ров), чтобы снизить потери и влияние неоднородностей из-за скин-эффекта.

На рис. 26, б показан ЛПД с барьером Шоттки, который представляет собой выпрямляющий контакт металл—полупроводник (гл. 5) [55—57]. Несмотря на то что распределения электрического поля в структурах, показанных на рис. 26, а и б, идентичны, диод с барьером Шоттки обладает рядом преимуществ. Во-первых, напряженность электрического поля максимальна на металлургической границе раздела, и поэтому выделяющееся тепло легко отводится через металлический контакт. Диод может быть изготовлен в форме усеченного конуса (рис. 26, б) с целью снижения влияния краевых эффектов, связанных с большой напряженностью электрического поля, и получения однородной области лавинного пробоя. Во-вторых, диод с барьером Шоттки можно создать при относительно низких температурах, при которых не происходит нарушения высококачественной структуры эпитаксиального слоя. Однако такие диоды обладают существенным недостатком. Дело в том, что в присутствии электронов и дырок высоких энергий атомы полупроводника могут химически воздействовать на металл, что приводит к ухудшению характеристик контакта.

Барьер Шоттки можно также использовать в модифицированном диоде Рида, если вместо p^+ -слоя нанести металлический контакт (рис. 26, в). Поскольку в барьерах Шоттки главную роль играют основные носители, эффект накопления неосновных носителей [58], который имеет место в обычных структурах Рида, здесь несуществен, а к. п. д. может быть даже выше.

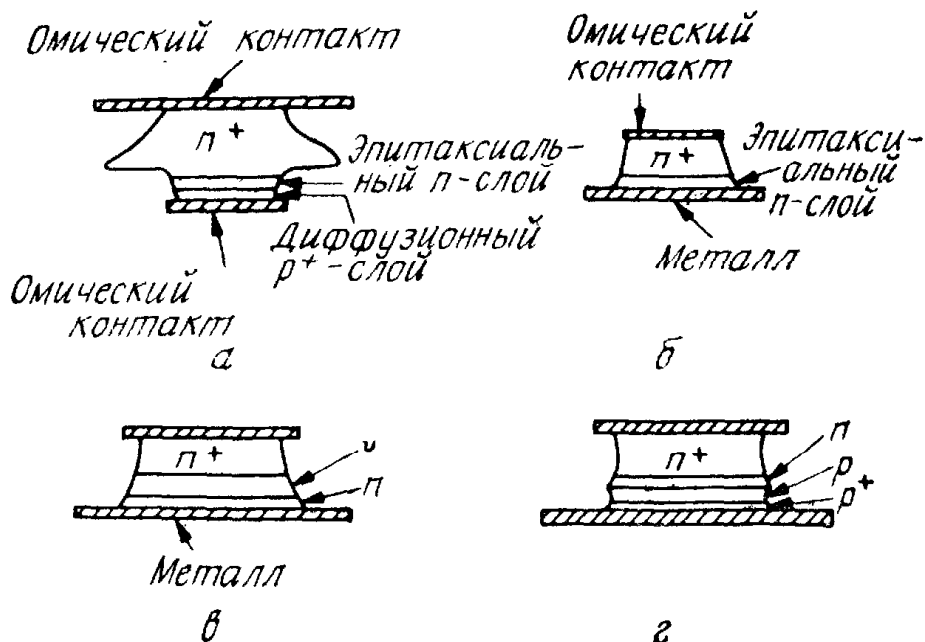


Рис. 26. Структуры некоторых ЛПД.

а — диод, созданный с помощью диффузии или двойной эпитаксии; б — барьер Шоттки; в — диод с двухслойной базой; г — диод с двумя областями дрейфа, созданный с помощью ионной имплантации.

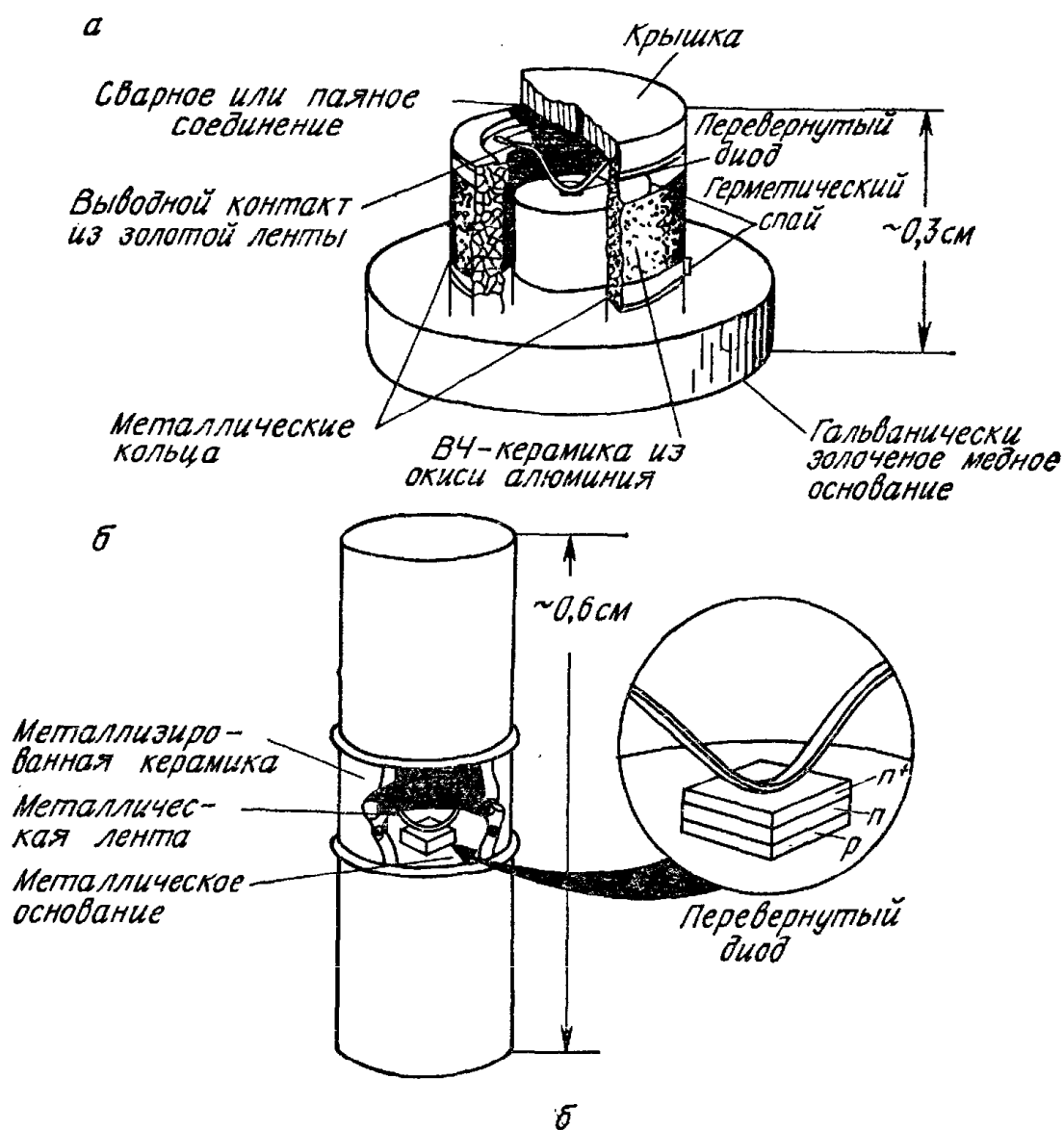


Рис. 27. Два СВЧ-корпуса с помещенными в них ЛПД.

Используя модифицированные структуры Рида, можно получить значительно большие к. п. д. по сравнению с обычными диодами с постоянным профилем примеси. Однако при изготовлении модифицированных диодов Рида требуется более строгий контроль профиля примеси, чтобы получить прибор с определенной частотной характеристикой. Метод самоограниченного анодного травления используется для уменьшения толщины высоколегированного слоя (в диоде с двухслойной базой) и низколегированного слоя (в диоде с трехслойной базой) с целью получения определенного напряжения пробоя (т. е. рабочей частоты) [59]. В большинстве барьеров Шоттки на GaAs высота барьера большая, а обратный ток насыщения мал. Однако платина реагирует с арсенидом галлия при рабочих температурах, образуя $PtAs_2$ и вызывая смещение поверхности барьера. Это «въедание» платины в арсенид галлия изменяет напряжение пробоя и приводит к деградации прибора. Эффект можно контролировать, если нанести на подложку слой платины толщиной 200—500 Å с последующим

нанесением слоя ртути или тантала для ограничения реакции платины с арсенидом галлия [60, 61].

На рис. 26, *г* показана структура диода, созданного с помощью метода ионной имплантации [62]. В случае бора или фосфора, которые наиболее часто используются при легировании полупроводников, глубина проникновения ионов составляет 0,5 мкм/100 кэВ. Поэтому слой шириной 1 мкм можно легко получить на установке для ионной имплантации с энергией ионов, равной нескольким сотням килоэлектронвольт. Такие источники используются при изготовлении ЛПД, работающих в диапазоне миллиметровых длин волн. К достоинствам этого метода следует отнести возможность изготовления структур с двумя дрейфовыми областями (рис. 26, *г*), в которых выходная мощность и импеданс, отнесенные к единичной площади, приблизительно в 2 раза больше. Поэтому ожидается, что такие структуры будут генерировать большие выходные мощности с более высокими к. п. д. [62].

ЛПД могут быть изготовлены методом молекулярно-лучевой эпитаксии [63]. Поскольку при этом можно контролировать уровень легирования и толщину вплоть до атомных размеров, предполагается, что данный метод окажется весьма полезным при создании ЛПД, работающих в миллиметровом и субмиллиметровом диапазонах длин волн.

Изготовленный диод обычно монтируют в СВЧ-корпус. Два типичных корпуса показаны на рис. 27. В обоих случаях диод крепится диффузионной областью или металлическим электродом на медный или алмазный теплоотвод для обеспечения эффективного охлаждения перехода во время работы.

10.6.2. Связь параметров диода с рабочей частотой и СВЧ-характеристики

Из малосигнальной теории можно получить приближенные соотношения между различными параметрами диода и рабочей частотой. Выражение (43) для сопротивления можно записать в следующем виде:

$$-R \sim \frac{W^2}{2A\epsilon_s v_s} \left(\frac{1}{\omega^2/\omega_r^2 - 1} \right) \left(\frac{1 - \cos \theta_d}{\theta_d^2/2} \right), \quad (62)$$

где θ_d — угол пролета, равный $\omega\tau$, и $\tau \simeq W/v_s$. Для того чтобы сопротивление $-R$ не зависело от ω/ω_r , необходимо выполнить следующие равенства (выражение (62)):

$$W^2/2A\epsilon_s v_s = \text{const}, \quad \omega_r \tau = \text{const}. \quad (63a), (63b)$$

Поскольку ширина обедненной области W обратно пропорциональна рабочей частоте (выражение (49)), площадь прибора, пропорциональная W^2 , зависит по закону ω^{-2} . Из условия лавинного

Таблица 3. Приближенные соотношения между параметрами ЛПД и рабочей частотой

Параметр	Зависимость от частоты
Площадь поперечного сечения перехода A	f^{-2}
Плотность постоянного тока J	f
Ширина обедненной области W	f^{-1}
Напряжение пробоя V_B	f^{-1}
Выходная мощность $P_{\text{вых}}$	f^{-1}
температурные ограничения	f^{-1}
ограничения, связанные со свойствами полупроводников	f^{-2}
К. п. д. η	Не зависит

пробоя (равенство (2)) можно показать, что коэффициент ионизации α и его производная по напряженности электрического поля α' обратно пропорциональны ширине обедненной области. Комбинируя выражения (37) и (63б) с соотношением $\alpha' \sim 1/W$, получим выражение для плотности постоянного тока

$$J_0 \sim \frac{\omega_f^2}{\alpha'} \sim \frac{\omega^2}{1/W} \sim \omega. \quad (64)$$

Соотношения между параметрами диода и рабочей частотой приведены в табл. 3. Они используются при экстраполяции характеристик и проектировании приборов, работающих в других частотных диапазонах.

Ограничения на выходную мощность рассмотрены в разд. 10.4. Предполагается, что при низких частотах коэффициент полезного действия будет очень слабо зависеть от частоты. Однако в диапазоне миллиметровых длин волн плотности рабочих токов велики ($\sim f$), а площадь мала ($\sim f^{-2}$), поэтому рабочая температура диода высокая. Повышение температуры в свою очередь приводит к увеличению обратного тока насыщения и снижению к. п. д. Кроме того, скин-эффект, туннельный и другие эффекты, которые связаны с высокими частотами и сильными электрическими полями, также ухудшают к. п. д. Следовательно, по мере увеличения частоты следует ожидать уменьшения к. п. д.

На рис. 28 показана СВЧ-цепь для измерений на частотах в диапазонах Н (6 — 8 ГГц) и I (8—10 ГГц) [64]. Наиболее важными компонентами схемы служат: 1) корпус диода, который обеспечивает необходимый теплоотвод, а также снабжен подвижными втулками, используемыми при настройке на резонанс в цепи; 2) измеритель мощности, который служит источником СВЧ-мощности; 3) генератор развертки, с помощью которого детекти-

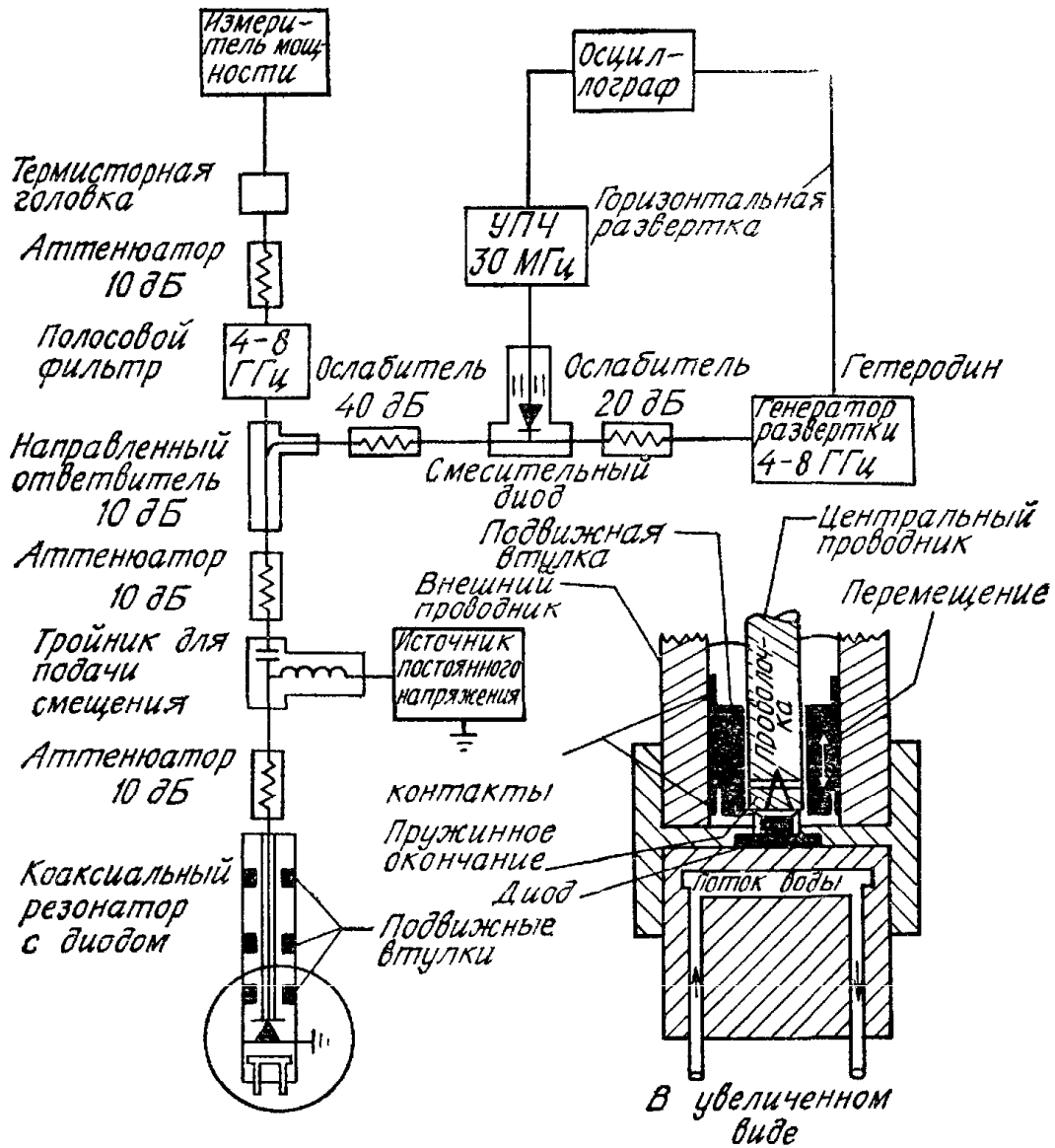


Рис. 28. Блок-схема широкодиапазонной измерительной установки для ЛПД [65].

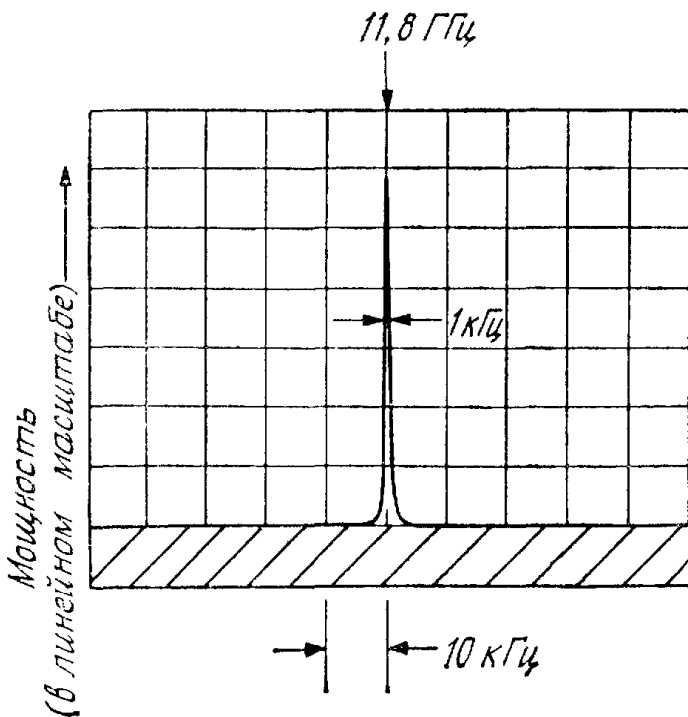


Рис. 29. Спектр генерируемых $p-n$ -диодом колебаний ($V_B = 54$ В) [66].

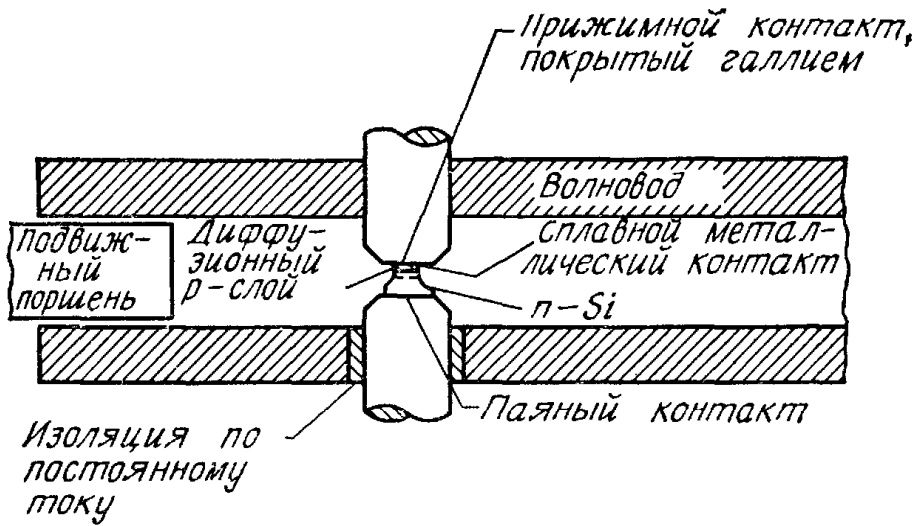


Рис. 30. СВЧ-резонатор для ЛПД миллиметрового диапазона длин волн [67].

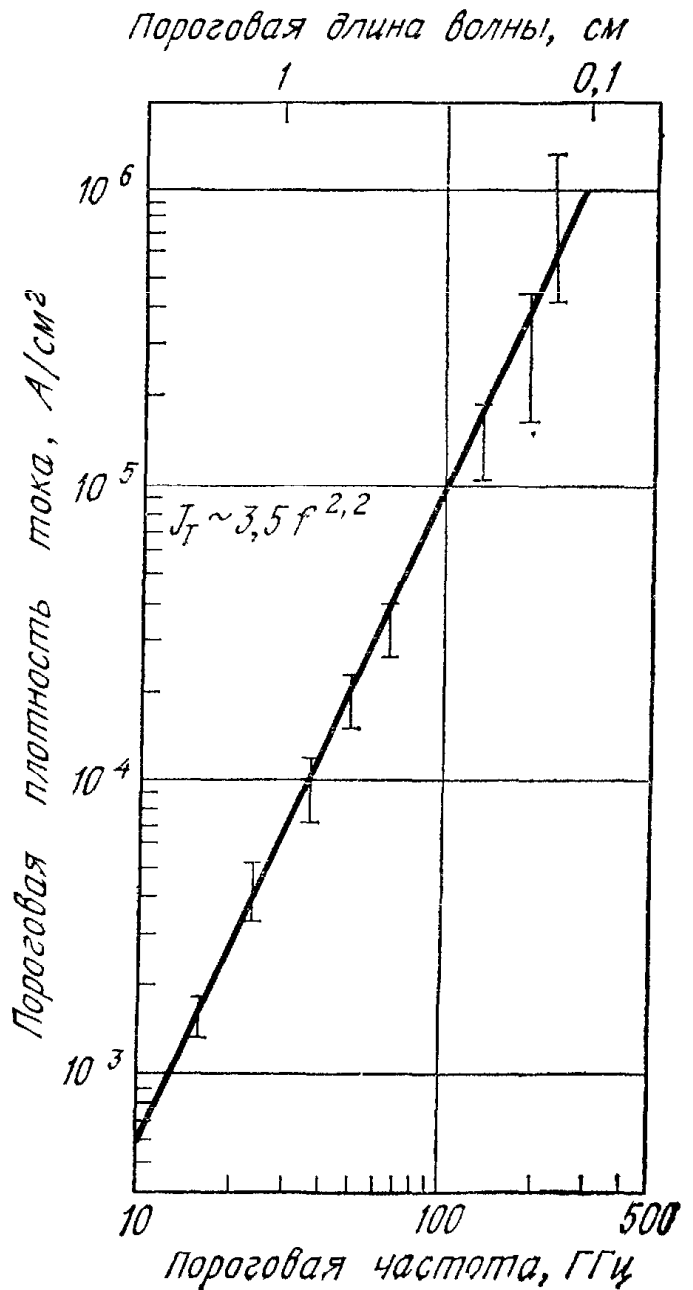


Рис. 31. Зависимость пороговой частоты от плотности тока [67].

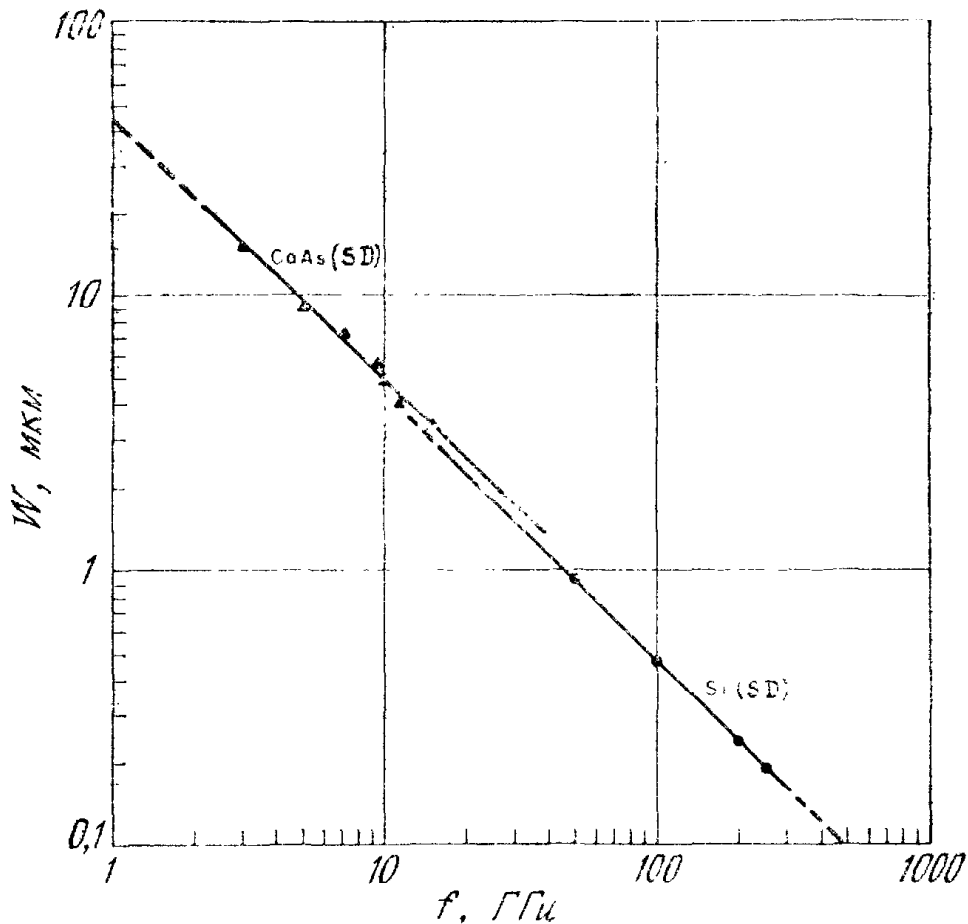


Рис. 32. Зависимость ширины обедненной области от частоты в Si- и GaAs-ЛПД [68, 69] (SD — одна область дрейфа).

руются СВЧ-сигналы; 4) осциллограф для наблюдения выходного сигнала [65]. Типичный спектр колебаний (генерируемых кремниевым $p-n$ -диодом), полученный с экрана осциллографа, показан на рис. 29 [66]. Рабочая частота равна 11,8 ГГц, а ширина полосы меньше 1 кГц. Для диапазона миллиметровых длин волн следует использовать другую измерительную схему и другой способ крепления диода; типичный СВЧ-резонатор с диодом показан на рис. 30 [67].

На рис. 31 приведена зависимость пороговой плотности тока, т. е. минимальной плотности тока, при которой начинается генерация, от частоты [67]. Отметим, что пороговая плотность тока изменяется приблизительно как квадрат частоты в соответствии с общими выводами, сделанными нами выше. Зависимость оптимальной ширины обедненной области от частоты для ЛПД из кремния и арсенида галлия [68, 69] приведена на рис. 32. Ширина обедненной области изменяется, как и ожидалось, обратно пропорционально частоте. Интересно, что при частотах выше 100 ГГц ширина обедненной области менее 0,5 мкм. Такая малая ширина слоя может дать представление о трудностях, с которыми приходится сталкиваться при изготовлении модифицированных

диодов Рида и диодов с двумя дрейфовыми областями, работающих в этом диапазоне частот.

В настоящее время наибольшая величина произведения мощности на квадрат частоты получена для диодов с двумя дрейфовыми областями. На рис. 33 приведены характеристики [70] диодов с одной и двумя дрейфовыми областями, работающих на частоте 50 ГГц. Созданный методами ионной имплантации кремниевый ЛПД с двумя дрейфовыми областями в непрерывном режиме генерирует мощность выше 1 Вт на частоте 50 ГГц при максимальном к. п. д. 14%. Эти данные можно сравнить с характеристиками аналогичного диода с одной дрейфовой областью, который генерирует мощность 0,5 Вт при к. п. д. 10%. Преимущества диодов с двумя дрейфовыми областями обусловлены тем, что в этих диодах электроны и дырки, генерируемые в лавине, пролетая соответствующие дрейфовые области, совершают работу против сил высокочастотного электрического поля. В диодах с одной дрейфовой областью используется лишь один тип носителей.

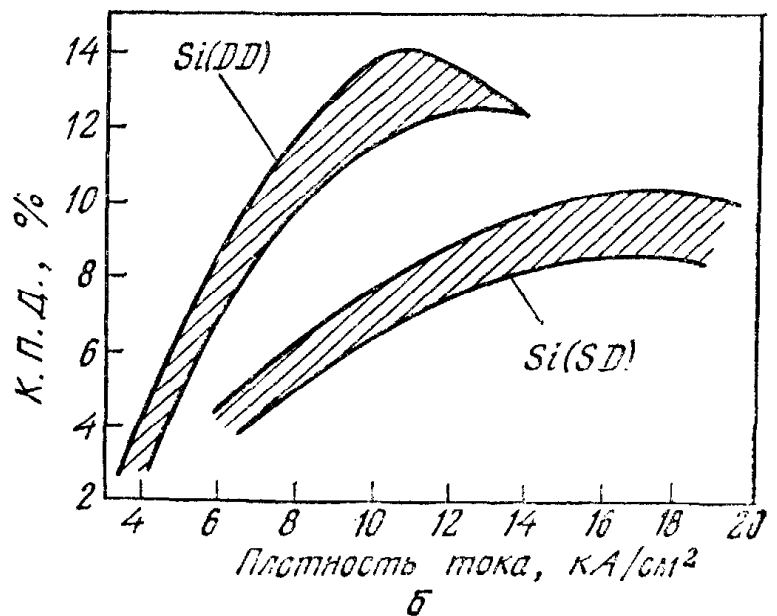
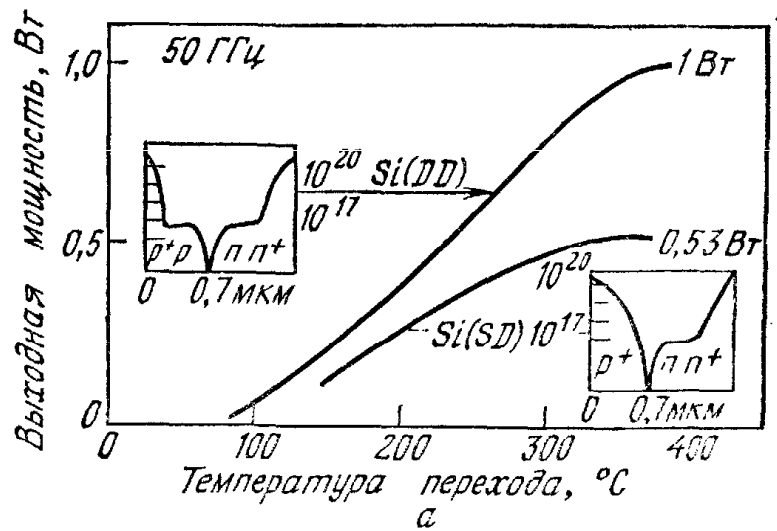


Рис. 33. Выходная мощность (а) и к. п. д. (б) для Si-ЛПД с одной (SD) и двумя (DD) областями дрейфа при 50 ГГц. Значения к. п. д. приведены для четырех диодов каждого типа [70].

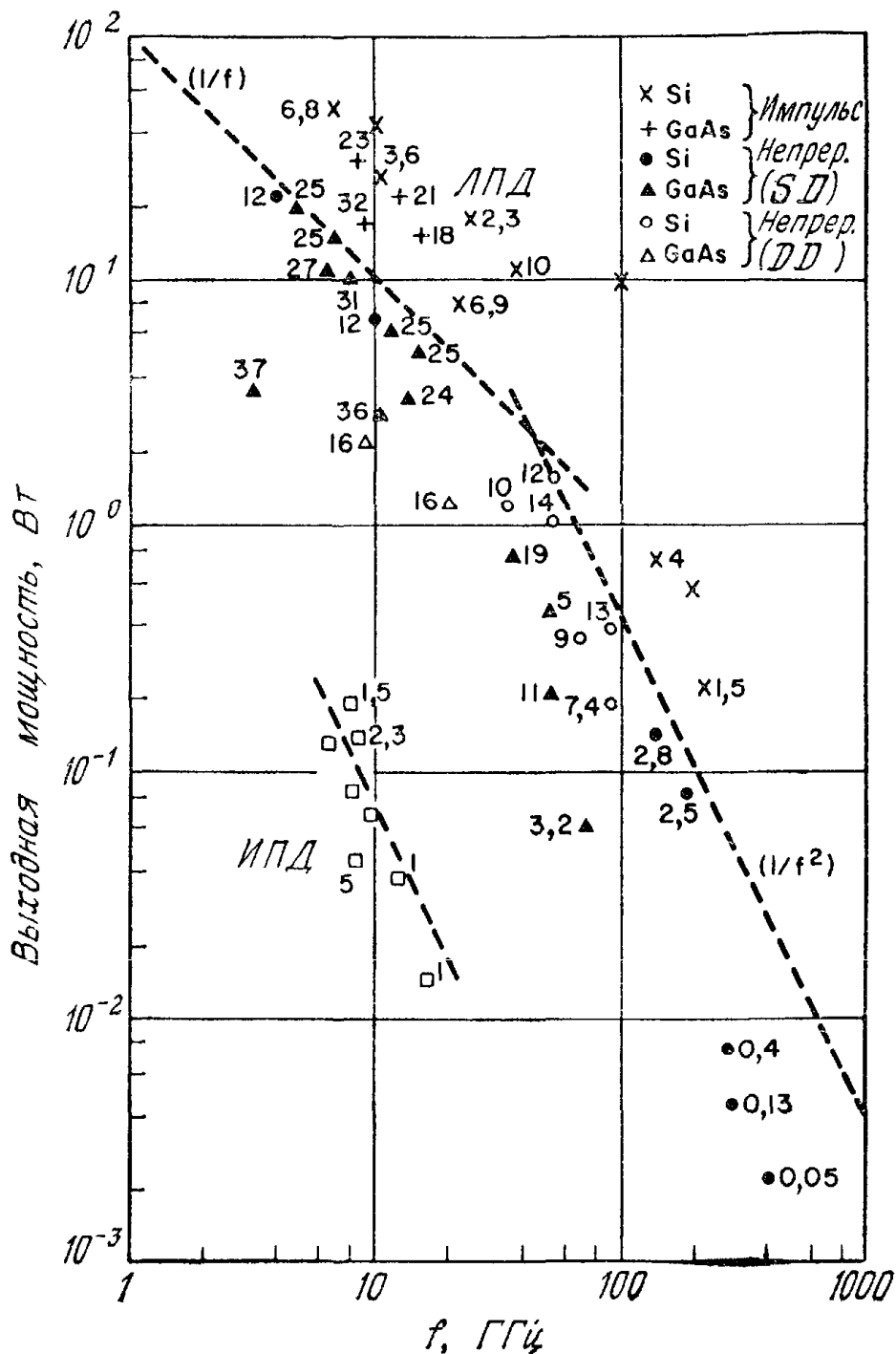


Рис 34. Характеристики современных ЛПД и ИПД. Рядом с экспериментальными точками указаны значения к. п. д. в процентах [10, 85, 97].

SD — одна область дрейфа; DD — две области дрейфа.

Характеристики современных ЛПД [10, 71, 72] приведены на рис. 34 вместе с характеристиками инжекционно-пролетных диодов (ИПД), которые изучены в разд. 10.7. При низких частотах выходная мощность ограничена тепловыми эффектами и зависит от частоты как $1/f$, а при более высоких частотах (>50 ГГц) мощность ограничена характеристиками носителей и пропорциональна $1/f^2$. На частотах, близких к 10 ГГц, для модифицированных диодов с трехслойной базой и контактом Pt—GaAs получены к. п. д., близкие к 40 % [73]. Из рис. 34 видно, что ЛПД являются наиболее мощными твердотельными источниками СВЧ-

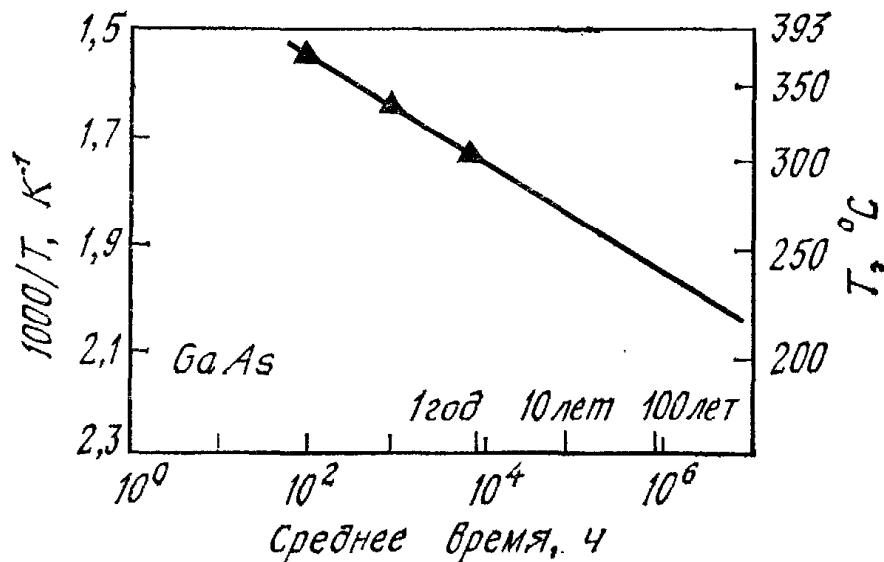


Рис. 35. Зависимость среднего времени работы GaAs-ЛПД до выхода из строя от температуры [74].

колебаний. Из всех твердотельных приборов ЛПД могут генерировать максимальную мощность в непрерывном режиме в миллиметровом диапазоне длин волн.

ЛПД очень надежны при работе в высокотемпературных условиях, если при их изготовлении применяется улучшенная технология. На рис. 35 приведена зависимость среднего времени работы прибора до выхода из строя от температуры для ЛПД из арсенида галлия [74]. Величина энергии активации, стоящая в показателе экспоненты этой зависимости, равна 1,9 эВ. Из экстраполяции этой зависимости следует, что при рабочих температурах в переходе $\sim 200^\circ\text{C}$ срок службы такого диода составит $\sim 10^7$ ч. Аналогичные результаты получены для кремниевых ЛПД [75].

10.7. ИНЖЕКЦИОННО-ПРОЛЕТНЫЕ ДИОДЫ

И ДВУХСКОРОСТНЫЕ ПРОЛЕТНЫЕ ДИОДЫ

Инжекционно-пролетный диод принадлежит к семейству пролетных СВЧ-диодов [12]. В основе его работы лежат два механизма: термоэмиссия (инжекция) и диффузия неосновных носителей через прямосмещенный барьер и пролет носителей через область дрейфа (что приводит к запаздыванию фазы на угол пролета $3\pi/2$). Обычно ИПД маломощны и имеют низкий к. п. д., однако отношение сигнал/шум лучше, чем у других диодов. ИПД используются чаще всего в качестве гетеродинов в СВЧ-приемниках.

Инжекционно-пролетный диод представляет собой два выпрямляющих контакта с общей базой, которая в рабочих условиях полностью обеднена носителями. Сначала мы рассмотрим протекание тока в таком диоде. Затем проведем анализ работы в режимах

малого и большого сигналов. В разд. 10.7.4 мы рассмотрим работу родственного прибора — двухскоростного пролетного диода (ДСПД).

10.7.1. Протекание тока

Сначала рассмотрим протекание электрического тока в симметричной структуре металл—полупроводник—металл (МПМ) с однородно легированным слоем полупроводника n -типа (рис. 36, *а*) [76]. На рис. 36, *б—г* приведены распределения объемного заряда, электрического поля и энергетическая диаграмма

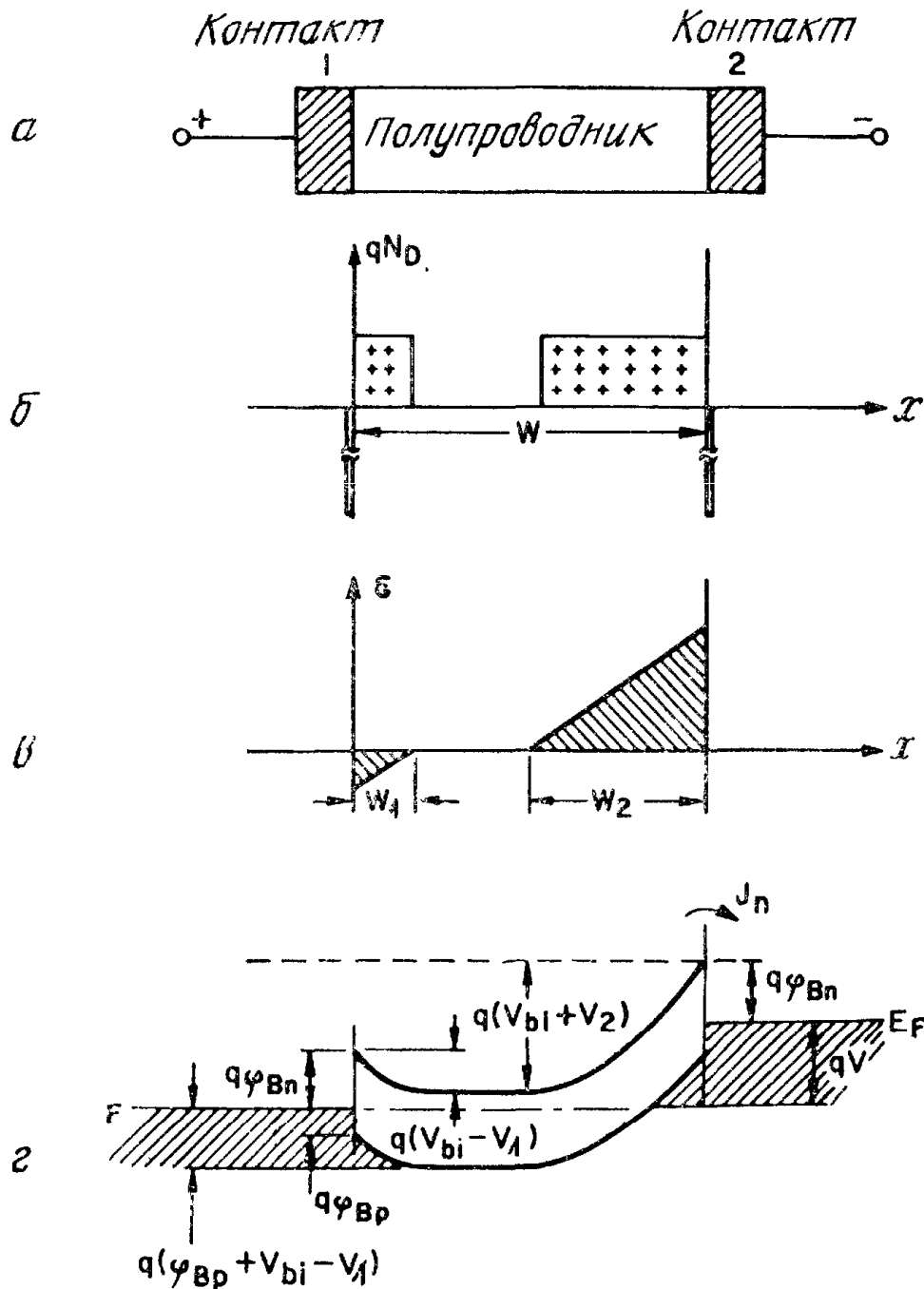


Рис. 36. Структура металл — полупроводник — металл (МПМ-структура) [76]. *а* — МПМ-структура с однородно легированным полупроводником n -типа; *б* — распределение пространственного заряда при малых смещениях диода; *в* — распределение поля; *г* — энергетическая диаграмма.

смещенного диода соответственно (меньшее положительное напряжение приложено к контакту 1, причем контакт 1 смещен, таким образом, в прямом направлении, а контакт 2 — в обратном). Ширина обедненных областей определяется следующими равенствами:

$$W_1 = \sqrt{\frac{2\epsilon_s}{qN_D} (V_{bi} - V_1)} \quad (65)$$

и

$$W_2 = \sqrt{\frac{2\epsilon_s}{qN_D} (V_{bi} + V_2)}, \quad (66)$$

где W_1 и W_2 — ширина обедненных областей прямо- и обратно-смещенных барьеров соответственно, N_D — концентрация ионизированной примеси и V_{bi} — контактная разность потенциалов. В этих условиях полный ток равен сумме обратного тока насыщения диода Шоттки с высотой барьера ϕ_{Bn} , генерационно-рекомбинационного тока и тока поверхностных утечек.

По мере увеличения напряжения в конце концов произойдет смыкание обедненной области контакта, смещенного в обратном направлении, с обедненной областью прямосмещенного контакта (рис. 37, а). Соответствующее напряжение называется напряжением смыкания (или напряжением прокола) V_{RT} . Это напряжение можно найти, приравняв $W_1 + W_2 = W$ (где W — ширина n -области):

$$\begin{aligned} V_{RT} &= \frac{qN_D}{2\epsilon_s} W^2 - W \left[\frac{2qN_D}{\epsilon_s} (V_{bi} - V_1) \right]^{1/2} \simeq \\ &\simeq \frac{qN_D}{2\epsilon_s} W^2 - W \left(\frac{2qN_D V_{bi}}{\epsilon_s} \right)^{1/2}. \end{aligned} \quad (67)$$

При дальнейшем увеличении напряжения энергетические зоны на контакте 1 становятся плоскими. В этом случае электрическое поле при $x = 0$ равно нулю, а падение напряжения на первом контакте $V_1 = V_{bi}$ (рис. 37, б). Соответствующее напряжение называется напряжением плоских зон V_{FB} :

$$V_{FB} \equiv \frac{qN_D W^2}{2\epsilon_s}. \quad (68)$$

Зависимость напряжения плоских зон от концентрации примеси в кремниевых диодах с различной шириной базы приведена на рис. 38. Для заданной ширины базы максимальная величина V_{FB} ограничена напряжением лавинного пробоя.

Величина постоянного смещения ИПД в условиях генерации СВЧ-мощности обычно лежит между V_{RT} и V_{FB} . В этом случае ($V_{RT} < V < V_{FB}$) приложенное напряжение V связано с высотой барьера, смещенного в прямом направлении, следующим соотношением:

$$V_{bi} - V_1 = (V_{FB} - V)^2 / 4V_{FB}. \quad (69)$$

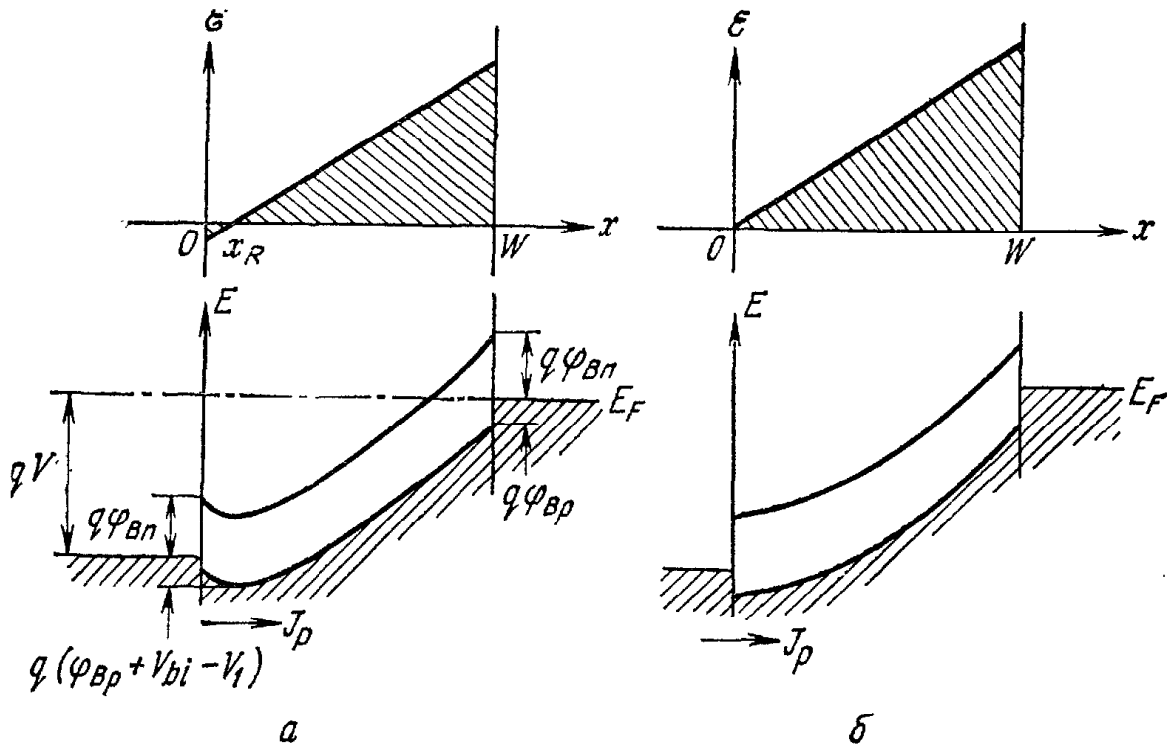


Рис. 37. Распределение электрического поля и энергетическая диаграмма МПМ-структуры при проколе (а) и в условиях плоских зон (б) [76].

Положение точки x_R , в которой происходит смыкание обедненных областей, определяется равенством

$$x_R/W = (V_{FB} - V)/2V_{FB}. \tag{70}$$

После смыкания обедненных областей ток термически эмиттированных через барьер ϕ_{Bp} дырок становится доминирующим:

$$J_p = A_p^* T^2 e^{-q(\phi_{Bp} + V_{bi})/kT} (e^{qV_1/kT} - 1), \tag{71}$$

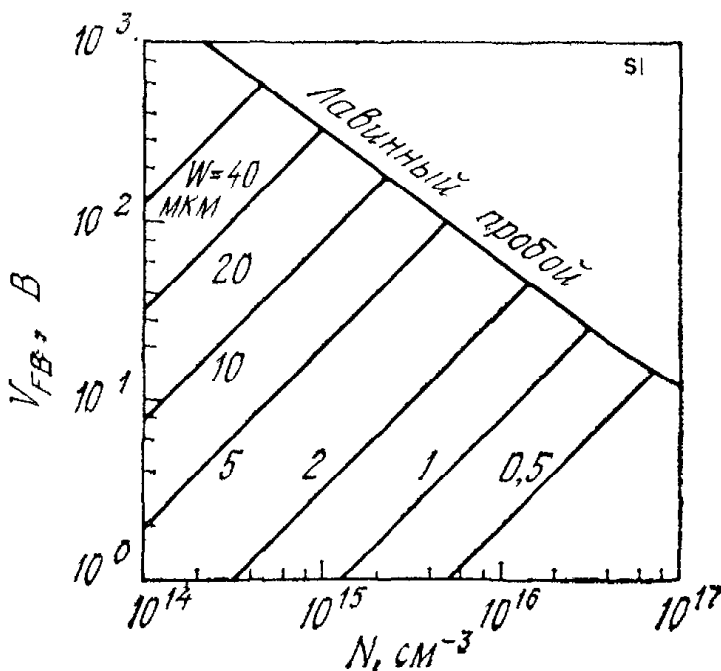


Рис. 38. Зависимость напряжения плоских зон от концентрации примеси в кремниевых диодах с различной шириной базы. Для заданной ширины базы максимальная величина напряжения плоских зон ограничена напряжением лавинного пробоя [76].

где A_p^* — эффективная постоянная Ричардсона (гл. 5). Из равенства (69) для $V \geq V_{RT}$ получаем

$$J_p = A_p^* T^2 e^{-q\phi_{BP}} \exp \left[-\frac{q(V_{FB} - V)^2}{4kTV_{FB}} \right]. \quad (72)$$

Таким образом, после смыкания обедненных носителями областей зависимость тока от напряжения будет экспоненциальной.

Если ток настолько велик, что концентрация инжектированных носителей сравнивается с концентрацией ионизованной примеси, объемный заряд подвижных носителей влияет на распределение электрического поля в области дрейфа. Это явление называется эффектом ограничения объемным зарядом. Если дырки пересекают n -область со скоростью насыщения v_s , а ток $J \gg qv_s N_D$, то уравнение Пуассона принимает вид

$$\frac{d\mathcal{E}}{dx} = \frac{\rho}{\epsilon_s} = \frac{q}{\epsilon_s} \left(N_D + \frac{J}{qv_s} \right) \simeq \frac{J}{\epsilon_s v_s}. \quad (73)$$

Дважды проинтегрировав это уравнение с граничными условиями $\mathcal{E} = 0$, $V = 0$ при $x = 0$, получим [77]

$$J = \left(\frac{2\epsilon_s v_s}{W^2} \right) V = qv_s N_D \left(\frac{V}{V_{FB}} \right). \quad (74)$$

Изложенные выше соображения можно применить для других структур, таких, как $p^+ - n - p^+$ - и $p^+ - i - n - \pi - p^+$ -диоды (рис. 39). Выражения для напряжений смыкания и плоских зон в $p^+ - n - p^+$ -диоде имеют такой же вид, как и для МПМ-структуры. Для усложненной $p^+ - i - n - \pi - p^+$ -структуры из аналогичных рассуждений можно получить величины V_{RT} , V_{FB} , x_R и соотношения между приложенным напряжением и падением напряжения на смещенном в прямом направлении переходе [78].

Протекание тока через $p^+ - n - p^+$ -диод, в котором произошел прокол, определяется теми же механизмами, что и в МПМ-структуре. Единственным отличием является отсутствие в равенствах (71) и (72) множителя $\exp(-q\phi_{BP}/kT)$ в случае инжекции носителей через смещенный в прямом направлении $p^+ - n$ -переход [77], т. е.

$$J = A^* T^2 \exp[-q(V_{FB} - V)^2/4kTV_{FB}] = J_{FB} \exp \left[-\frac{q(V_{FB} - V)^2}{4kTV_{FB}} \right]. \quad (75)$$

Высота барьера в структуре PtSi—Si равна 0,2 эВ. Следовательно, при температуре 300 К и заданном напряжении, превышающем напряжение смыкания, ток $p^+ - n - p^+$ -диола будет приблизительно в 3000 раз больше тока МПМ-структуры. При комнатной температуре J_{FB} ($\equiv A^* T^2$) составляет $\sim 10^7$ А/см². Поэтому в обычных рабочих условиях эффект ограничения объемным зарядом становится существенным при токах, значительно меньших J_{FB} .

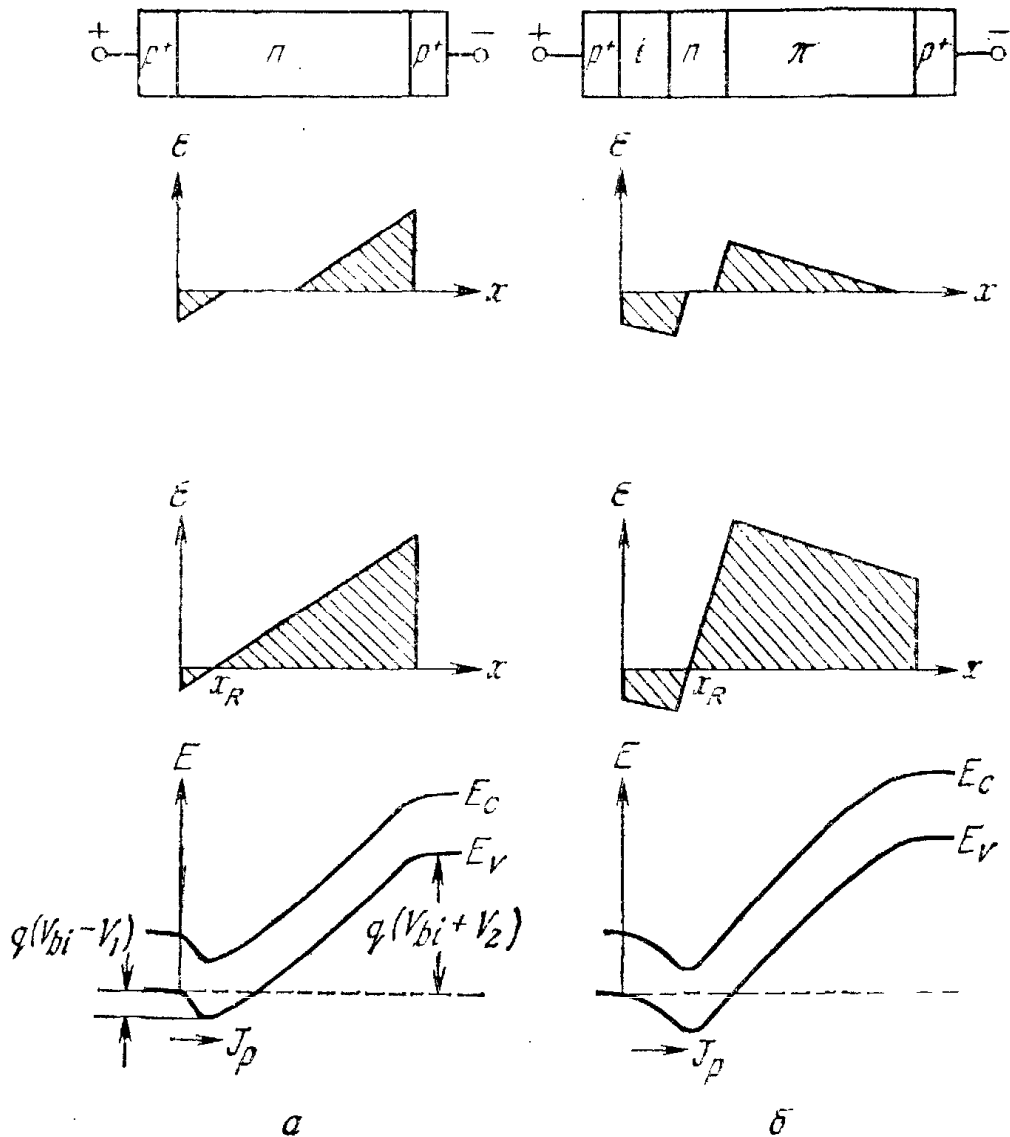


Рис. 39. Распределение электрического поля при малых смещениях и проколе и энергетические диаграммы при проколе для $p^+ - n - p^+$ -структуры (а) и $p^+ - i - n - p^+$ -структуры (б).

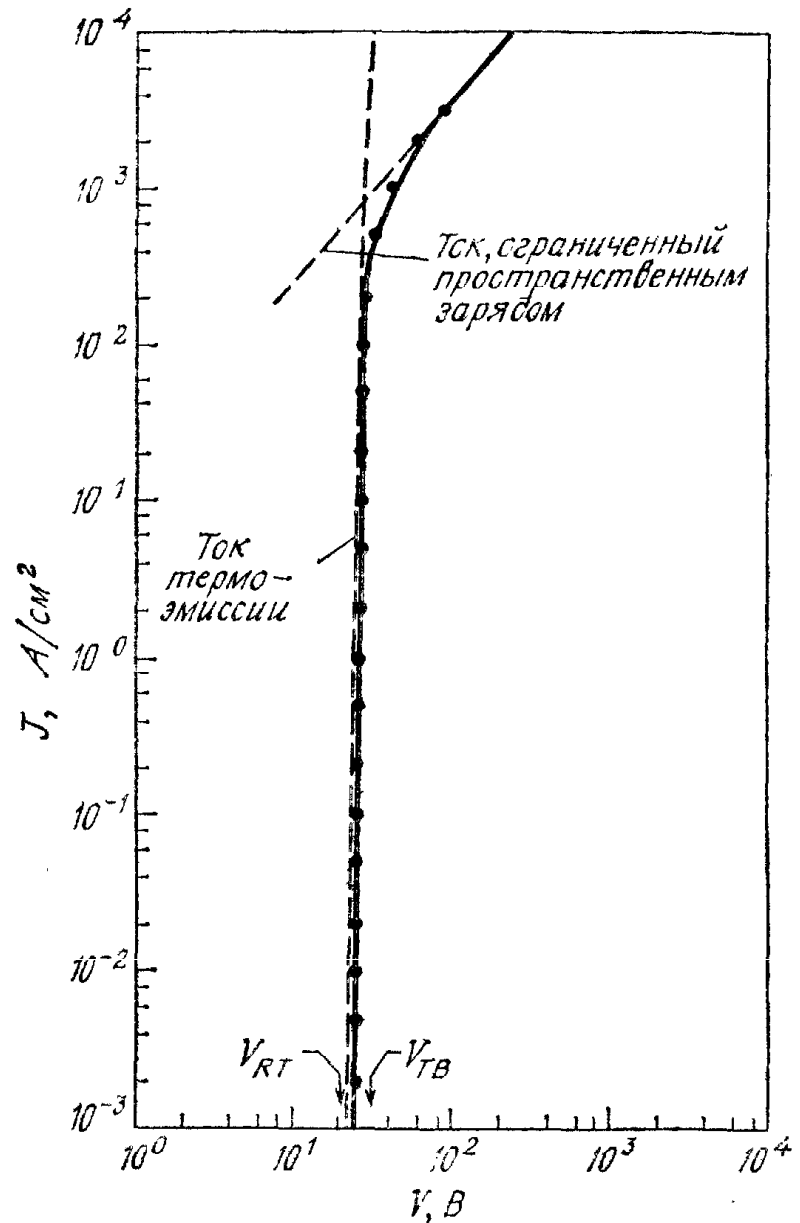
Типичная вольт-амперная характеристика кремниевого $p^+ - n - p^+$ -диода с концентрацией примеси $5 \cdot 10^{14} \text{ см}^{-3}$ и шириной базы 8,5 мкм приведена на рис. 40. Напряжение плоских зон равно 29 В, а напряжение смыкания составляет ~ 21 В. Отметим, что сначала ток возрастает экспоненциально, а затем линейно с напряжением. Экспериментальные результаты находятся в хорошем соответствии с рассчитанными с помощью равенств (74) и (75) зависимостями.

Для эффективной работы ИПД необходимо контролировать величину инжектированного заряда. Поэтому ток должен резко возрастать с увеличением напряжения. Линейная вольт-амперная зависимость, связанная с эффектом ограничения объемным зарядом, будет ухудшать характеристики прибора. Обычно оптимальная плотность тока значительно меньше $J = qv_s N_D$.

Описанный выше диод со смыканием обедненных областей (который также называют диодом с проколом) используется в каче-

Рис. 40. Вольт-амперная характеристика кремниевого $p^+ - n - p^+$ -диода в условиях работы с проколом [77].

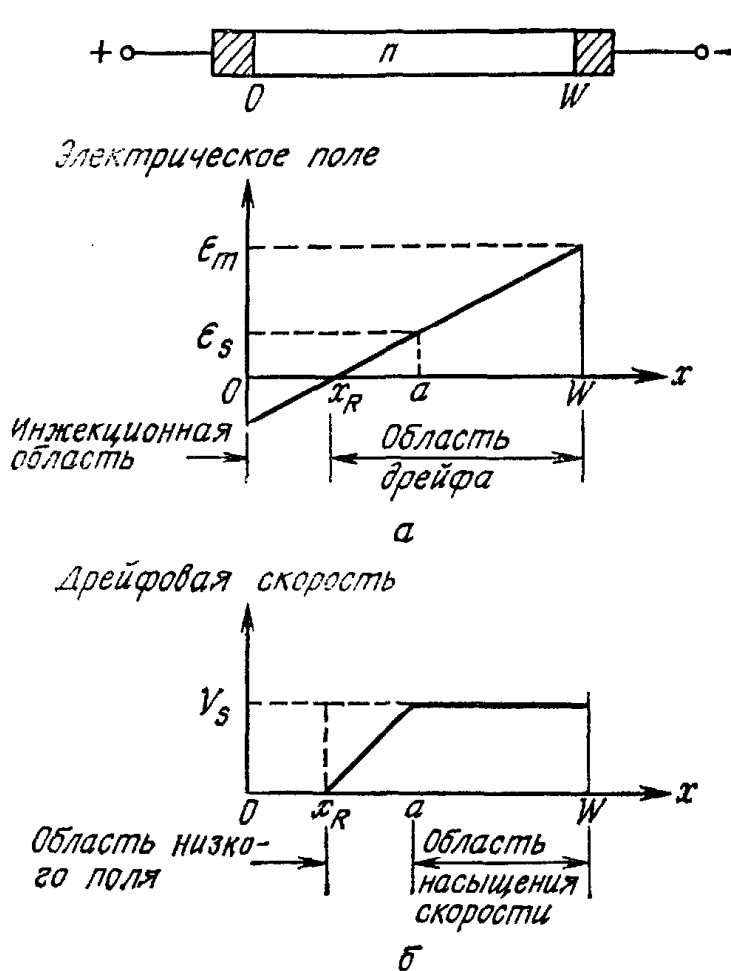
— — — теоретическая; —●— экспериментальная.



стве быстродействующего ограничителя напряжения, поскольку ток в таком диоде резко возрастает, как только напряжение превысит напряжение смыкания. Эффект накопления заряда практически отсутствует; кроме того, диод обладает хорошей температурной стабильностью. Были созданы диоды со смыканием с рабочим напряжением $\sim 1,5$ В, причем характеристики были сравнимы и даже лучше, чем характеристики зинеровского диода, в котором используются явления лавинного пробоя или туннелирования (при больших напряжениях — лавинный пробой, при малых — туннелирование) [79].

10.7.2. Малосигнальные характеристики

Ниже показано, что малосигнальное сопротивление инжекционно-пролетного диода отрицательно и, следовательно, в нем самопроизвольно возбуждаются колебания. Рассмотрим $p^+ - n - p^+$ -



структуру. Распределение электрического поля при напряжении смещения, большем напряжения смыкания, приведено на рис. 41, а. Точка x_R соответствует максимальной потенциальной энергии инжектированных дырок; ее положение определяется выражением (70). Точка a отделяет область относительно слабых электрических полей от области, в которой дрейфовая скорость носителей равна скорости насыщения, т. е. $\mathcal{E} \geq \mathcal{E}_s$, а $v = v_s$ (рис. 41, б). Если концентрация инжектированных носителей невелика, то

$$a \approx \frac{\epsilon_s \mathcal{E}_s}{qN_D} + x_R. \quad (76)$$

Время пролета носителями области дрейфа ($x_R < x < W$) определяется следующим выражением [80]:

$$\tau_d = \int_{x_R}^a \frac{dx}{\mu_n \mathcal{E}(x)} + \int_a^W \frac{dx}{v_s} = \int_{x_R}^a \frac{dx}{\mu_n q N_D x / \epsilon_s} + \frac{W - a}{v_s} \approx \frac{3,75 \epsilon_s}{q \mu_n N_D} + \frac{W - a}{v_s}. \quad (77)$$

Для того чтобы получить выражение для малосигнального импеданса диода, используем подход, аналогичный подходу, примененному в разд. 10.3.2, и представим зависящие от времени вели-

чины в виде сумм, не зависящих от времени и малосигнальных переменных компонент:

$$\begin{aligned} J(t) &= J_0 + \tilde{J}e^{j\omega t}, \\ V(t) &= V_0 + W\tilde{\mathcal{E}}e^{j\omega t}. \end{aligned} \quad (78)$$

Подставив эти соотношения в выражение (75), получим выражение для линеаризованной переменной составляющей инжекционного тока дырок

$$\tilde{J} = \sigma\tilde{\mathcal{E}}, \quad (79)$$

где σ — инжекционная проводимость на единичную площадь, равная

$$\sigma = J_0 \frac{\epsilon_s (V_{FB} - V_0)}{N_D W kT}, \quad (80)$$

причем J_0 — плотность тока, определяемая выражением (75), в котором V следует заменить на V_0 . Инжекционная проводимость возрастает при увеличении напряжения, достигает максимума, а затем быстро убывает, когда V_0 приближается к V_{FB} . Из выражений (75) и (80) можно найти смещение диода при максимальной проводимости σ :

$$V_{0\sigma \text{ макс}} = V_{FB} - \sqrt{2kTV_{FB}/q}. \quad (81)$$

Поскольку переменное электрическое поле непрерывно на границе инжекционной и дрейфовой областей, эти области взаимодействуют друг с другом. Определим \tilde{J} как плотность полного переменного тока, а \tilde{J}_1 — плотность инжекционного тока. Предположим, что инжекционная область настолько тонкая, что носители заряда без запаздывания поступают в область дрейфа. Переменная компонента плотности тока проводимости определяется следующим равенством:

$$\tilde{J}_c(x) = \tilde{J}_1 e^{-j\omega\tau(x)} \equiv \gamma \tilde{J} e^{-j\omega\tau(x)} \quad (82)$$

и представляет собой незатухающую волну, распространяющуюся по направлению к точке $x = W$ с пролетным запаздыванием по фазе, равным $\omega\tau(x)$. Комплексная величина $\gamma \equiv \tilde{J}_1/\tilde{J}$ равна отношению плотностей переменных инжекционного и полного токов.

В заданной точке области дрейфа плотность полного переменного тока \tilde{J} равна сумме тока проводимости \tilde{J}_c и тока смещения \tilde{J}_d :

$$\tilde{J} = \tilde{J}_c(x) + J_d(x) \neq f(x). \quad (83)$$

Ток смещения связан с напряженностью переменного электрического поля $\tilde{\mathcal{E}}(x)$ следующим соотношением:

$$\tilde{J}_d(x) = j\omega\epsilon_s\tilde{\mathcal{E}}(x). \quad (84)$$

Комбинируя выражения (80), (82) и (84), получим переменную составляющую напряженности электрического поля в области дрейфа как функцию x и \tilde{J} :

$$\mathcal{E}(x) = \frac{\tilde{J}}{j\omega\epsilon_s} [1 - \gamma e^{-j\omega\tau(x)}]. \quad (85)$$

Проинтегрировав $\mathcal{E}(x)$, получим зависимость переменной компоненты падения напряжения на области дрейфа от плотности тока \tilde{J} . Коэффициент γ определяется выражением

$$\gamma = \frac{\tilde{J}_1}{\tilde{J}_1 + \tilde{J}_d} = \frac{\sigma}{\sigma + j\omega\epsilon_s}. \quad (86)$$

Подставив γ в выражение (85) и проинтегрировав по длине области дрейфа ($W - x_R$) с учетом граничных условий $\tau = 0$ при $x = x_R$ и $\tau = \tau_d$ при $x = W$, получим переменную составляющую падения напряжения на области дрейфа

$$V_d = \frac{\tilde{J}(W - x_R)}{j\omega\epsilon_s} \left[1 - \frac{\sigma}{\sigma + j\omega\epsilon_s} \frac{1 - e^{j\theta_d}}{j\theta_1} \right], \quad (87)$$

где θ_d — угол пролета в области дрейфа:

$$\theta_d = \frac{\omega[(W - x_R) + (a - x_R)]}{v_s} = \omega\tau_d, \quad (88)$$

а θ_1 — постоянная, равная

$$\theta_1 \equiv \omega \left(\frac{W - x_R}{v_s} \right). \quad (89)$$

Определим $C_d = \epsilon_s/(W - x_R)$ как емкость области дрейфа. Из выражения (87) получим малосигнальный импеданс структуры

$$Z \equiv \frac{\tilde{V}_d}{\tilde{J}} = R_d - jX_d, \quad (90)$$

где R_d и X_d — малосигнальные действительная и мнимая составляющие импеданса соответственно:

$$R_d = \frac{1}{\omega C_d} \left(\frac{\sigma}{\sigma^2 + \omega^2 \epsilon_s^2} \right) \left[\frac{\sigma(1 - \cos \theta_d) + \omega \epsilon_s \sin \theta_d}{\theta_1} \right],$$

$$X_d = \frac{1}{\omega C_d} - \frac{1}{\omega C_d} \left(\frac{\sigma}{\sigma^2 + \omega^2 \epsilon_s^2} \right) \left[\frac{\sigma \sin \theta_d - \omega \epsilon_s (1 - \cos \theta_d)}{\theta_1} \right].$$

Отметим, что действительная составляющая (сопротивление) отрицательна при условии, что угол пролета θ_d больше π , но меньше 2π и если $|(1 - \cos \theta_d)/\sin \theta_d|$ меньше $\omega \epsilon_s/\sigma$.

На рис. 42 приведены рассчитанные значения (штриховые линии) и экспериментально полученные значения (сплошные

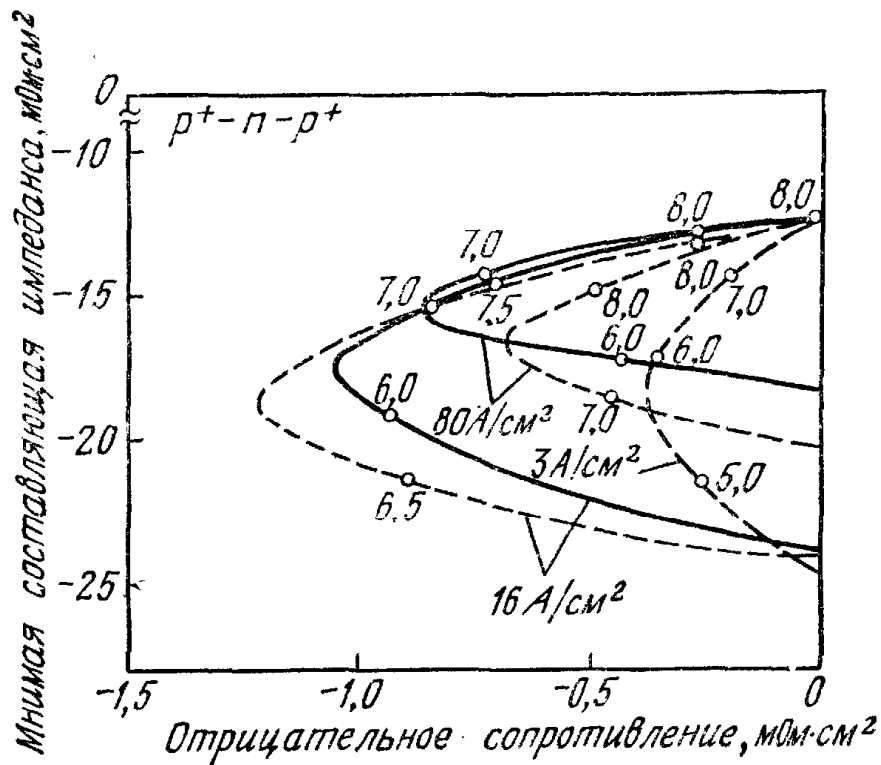


Рис. 42. Малосигнальный импеданс $p^+ - n - p^+$ -диода в области отрицательного сопротивления при трех различных плотностях тока [81]:
 --- теоретические значения; — экспериментальные значения.

линии) действительной и мнимой составляющих малосигнального импеданса $p^+ - n - p^+$ -диода с $L = 8,5$ мкм и $N = 5 \cdot 10^{14}$ см $^{-3}$ [81]. Для плотности тока 3 А/см 2 отрицательное сопротивление $-R_d$ максимально на частоте $f = 5,7$ ГГц. При увеличении плотности тока до 16 А/см 2 растут амплитуда максимума $-R_d$ и соответствующая частота. При увеличении плотности тока до 80 А/см 2 максимальное отрицательное сопротивление $-R_d$ начинает убывать в результате уменьшения инжекционной проводимости (выражение (80)).

Теоретические и экспериментальные результаты совпадают в общих чертах. Малосигнальная теория также применима к инжекционно-пролетным диодам с МПМ-структурой, для которых получено хорошее соответствие с экспериментальными данными.

Из полученных выше результатов следует, что 1) инжекционно-пролетные диоды имеют отрицательное малосигнальное сопротивление и, следовательно, могут самовозбуждаться; 2) инжекция через смещенный в прямом направлении $p^+ - n$ -переход или барьер Шоттки служит источником носителей заряда; 3) время пролета носителями области дрейфа является важной величиной для частотных характеристик.

Инжекционно-пролетный диод представляет собой прибор с низким уровнем шума, основными источниками которого являются дробовой шум инжектированных носителей (инжекцион-

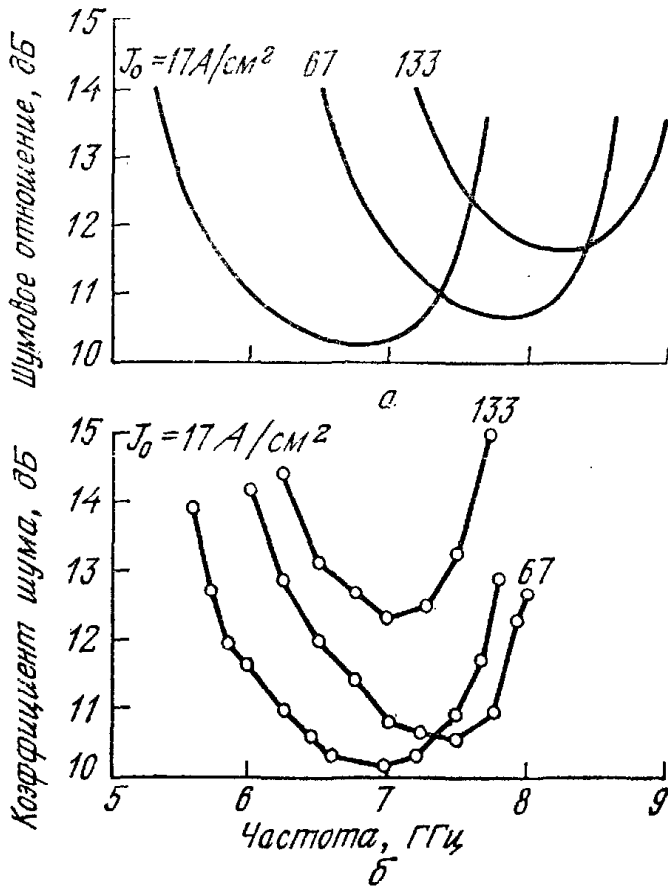


Рис. 43. Рассчитанные шумовые отношения (а) и экспериментально полученные коэффициенты шума (б) Si-ИПД с $W = 7,9$ мкм и $N_D = 1,2 \cdot 10^{15}$ см $^{-3}$ [82].

ный шум) и случайные флуктуации скорости в области дрейфа (диффузный шум).

Рассчитанные и экспериментально полученные результаты [82] находятся в хорошем соответствии при малых плотностях токов (рис. 43). Для $p^+ - n - p^+$ -диода с $N_D = 1,2 \cdot 10^{15}$ см $^{-3}$ и $W = 7,9$ мкм теоретическая величина шумового отношения составляет ~ 10 дБ при плотности тока 17 А/см 2 . При увеличении плотности тока кривые сдвигаются в сторону более высоких частот. Из сравнения этих результатов с шумами ЛПД видно, что уровень шума в инжекционно-пролетных диодах существенно ниже.

10.7.3. Характеристики при работе в режиме большого сигнала [83]

Основные характеристики ИПД при работе в режиме большого сигнала приведены на рис. 44. Носители инжектируются в виде δ -импульса, когда величина переменного напряжения максимальна ($\varphi = \pi/2$). В течение трех четвертей периода эти носители движутся в направлении отрицательно смещенного контакта:

$$\theta_d = \omega \tau_d = 3\pi/2 \quad (91)$$

или

$$f = \frac{3}{4\tau_d}, \quad (91a)$$

где θ_d — угол пролета и τ_d — время пролета. В первом приближении τ_d равно W/v_s . Следовательно, частота осцилляций

$$f \approx \frac{3v_s}{4W}. \quad (92)$$

Более точное выражение для оптимальной частоты можно получить, подставив выражение (77) в выражение (92).

По оценкам максимальный к. п. д. ИПД равен $\sim 10\%$ при условии, что носители инжектируются при $\varphi = \pi/2$ (рис. 44). Однако можно достигнуть больших к. п. д., если инжектировать носители в более поздний момент цикла, т. е. при $\pi/2 < \varphi \leq \pi$. С этой целью был создан многослойный $n^+ - i - p - v - n^+$ -ИПД [84] (на рис. 39, б показана комплементарная $p^+ - i - n - \pi - p^+$ -структура). Тормозящее поле $n^+ - i - p$ -области увеличивает время инжекционного запаздывания. Измеренные значения выходной мощности и к. п. д. приведены на рис. 45. Максимальное значение к. п. д. составляет более 5% при токе $J \approx 50$ А/см². Соответствующее значение выходной мощности равно 40 мВт

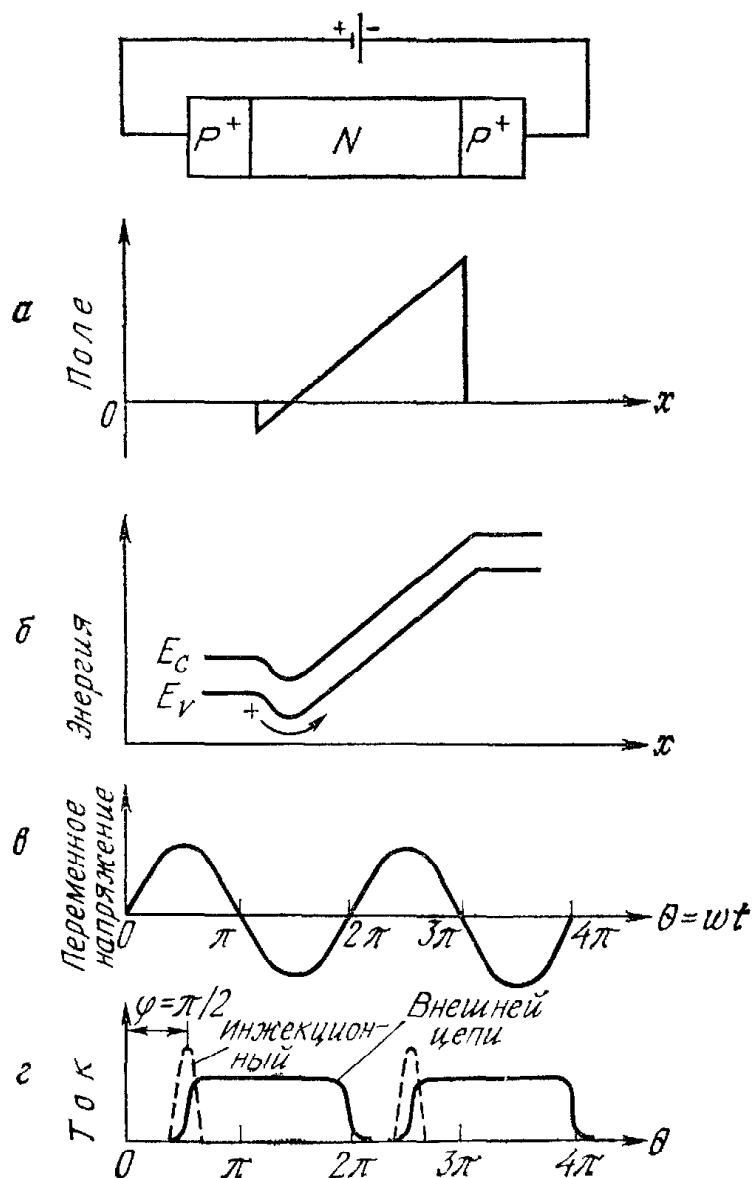


Рис. 44. Работа ИПД в режиме большого сигнала:

а — $p^+ - n - p^+$ -структура; б — распределение поля после прокола; в — энергетическая диаграмма; г — переменное напряжение; д — инжекционный ток и ток внешней цепи.

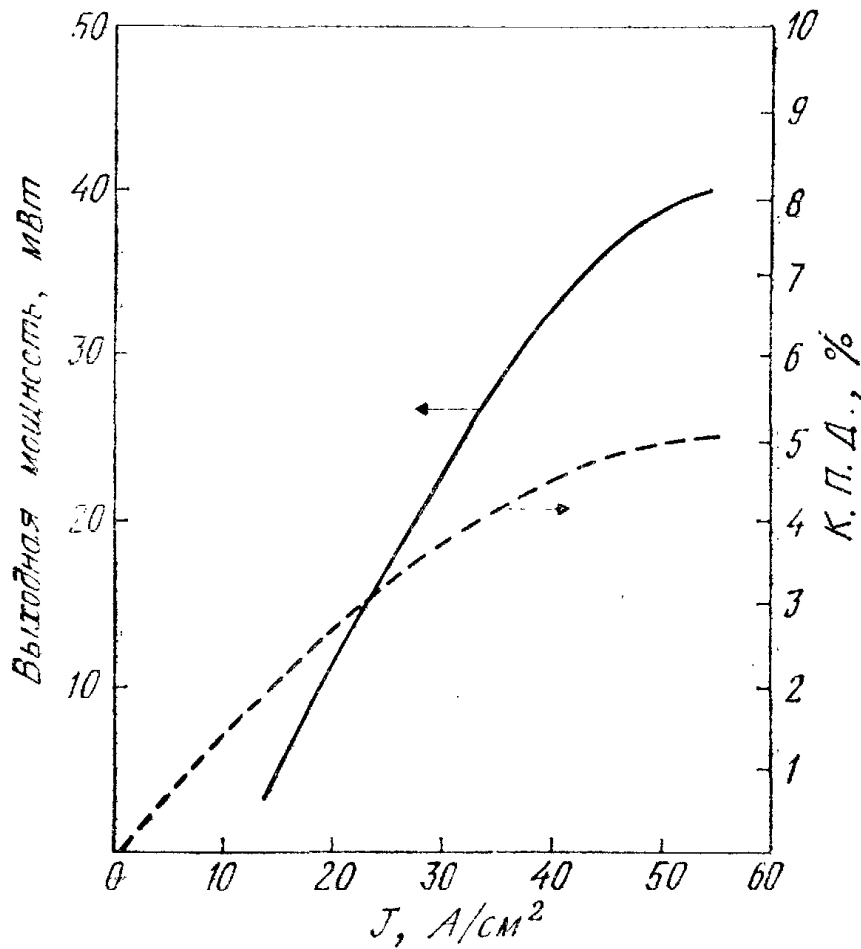


Рис. 45. Зависимости выходной мощности и к. п. д. $n^+ - i - p - v - n^+$ -ИПД от плотности тока [84].

на частоте 7,6 ГГц. После нанесения платинового контакта Шоттки на эпитаксиальный $n-p^+$ -слой с концентрацией $N_D = 2,5 \times 10^{15} \text{ см}^{-3}$ и $W = 6,5 \text{ мкм}$ была получена выходная мощность 152 мВт на частоте 8,6 ГГц при максимальном к. п. д. 2,3 % [85]. Характеристики современных ИПД приведены на рис. 34 [84, 85, 97]. Хотя выходная мощность на частоте ~ 10 ГГц приблизительно на два порядка меньше, чем в ЛПД, шумовое отношение ИПД меньше приблизительно во столько же раз. Предполагается, что путем оптимизации времени инжекционного запаздывания можно реализовать потенциальные преимущества ИПД как малошумящего прибора с приемлемыми выходной мощностью и к. п. д.

10.7.4. Двухскоростной пролетный диод

Диод с двухскоростным режимом пролета, или двухскоростной пролетный диод (ДСПД), подобен ИПД за одним исключением: скорость носителей вблизи инжектирующего контакта в нем значительно ниже, чем вблизи другого контакта [86, 87]. Для $n^+ - p - n - n^+$ -ДСПД предложена гетероструктура с $n^+ - p$ -переходом из $\text{Al}_x\text{Ga}_{1-x}\text{As}$ и $n - n$ -переходом из арсенида галлия

(вставка на рис. 46). На этой вставке также приведено распределение дрейфовых скоростей носителей в диоде, причем скорость насыщения v_1 в области 1 меньше скорости насыщения v_2 в области 2. Постоянное смещение диода таково, что область p — n -гетеро-перехода полностью «проколота» электрическим полем, а инжектируемый через смещенный в прямом направлении n^+ — p -переход ток очень большой.

Из-за дополнительного запаздывания фазы при пролете носителями области с малой скоростью насыщения ожидается, что отрицательное сопротивление ДСПД будет больше, чем сопротивление ИПД (который является предельным случаем ДСПД с равными скоростями насыщения $v_1 = v_2$). Максимальное малосигнальное отрицательное сопротивление определяется выражением [87]

$$R_{\text{макс}} = \frac{[v_1(1 + \cos \varphi) - 2v_2] \cos \varphi}{A\omega^2\epsilon_s}, \quad (93)$$

где φ — инжекционное запаздывание фазы, A — площадь прибора. На рис. 46 приведены теоретические значения $R_{\text{макс}}$ в зависимости от инжекционного запаздывания фазы φ при различных соотношениях между скоростями насыщения. Отметим, что

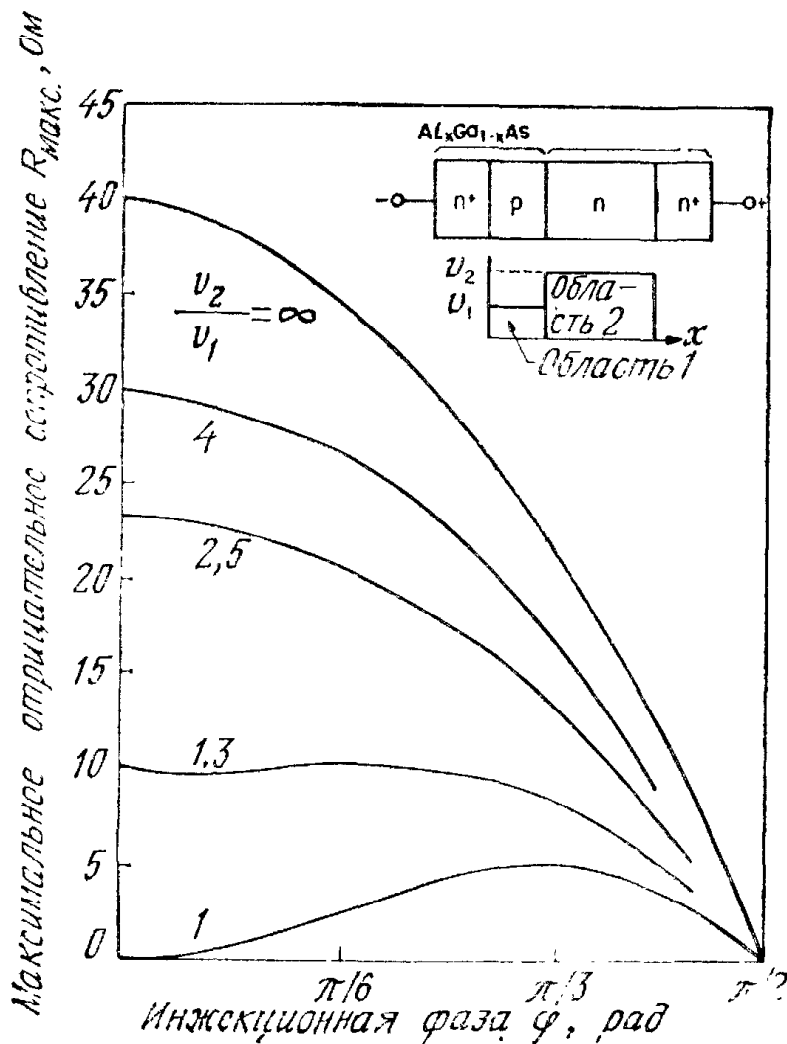


Рис. 46. Зависимость максимального отрицательного сопротивления от инжекционного запаздывания фазы в ДСПД при 10 ГГц с $v_2 = 8 \cdot 10^6$ см/с и $A = 10^{-4}$ см². На вставке показаны структура прибора и распределение дрейфовых скоростей [87].

в ИПД ($v_2/v_1 = 1$) для получения максимального отрицательного сопротивления необходимо инжекционное запаздывание фазы, равное $\sim \pi/3$ рад. Из рис. 46 ясно видны преимущества ДСПД с большим отношением скоростей насыщения. Отрицательное сопротивление сравнительно велико, а прибор может работать даже при условии отсутствия инжекционного запаздывания ($\varphi = 0$). Поэтому токи в ДСПД могут быть больше, чем в ИПД, в котором с ростом плотности тока происходит уменьшение инжекционного запаздывания до тех пор, пока не исчезнет отрицательное сопротивление. Для ДСПД теоретическое значение к. п. д. превышает 25 %.

10.8. ПРОЛЕТНЫЙ ДИОД С ЗАХВАЧЕННЫМ ОБЪЕМНЫМ ЗАРЯДОМ ЛАВИНЫ

Этот диод является мощным прибором с высоким коэффициентом полезного действия. В настоящее время максимальная экспериментально полученная выходная мощность диода в импульсном режиме генерации равна 1,2 кВт на частоте 1,1 ГГц (пять последовательно соединенных диодов [88]), а максимальный к. п. д. равен 75 % на частоте 0,6 ГГц [89]. Однако способ генерации сигнала в таком диоде весьма сложен, и, кроме того, работа в этом режиме требует точного контроля параметров цепи и самого диода. Шумовое отношение этого диода существенно больше, чем шумовое отношение ЛПД, а максимально достижимая рабочая частота, по-видимому, меньше частот диапазона миллиметровых длин волн.

Для того чтобы понять, как возникает пролетный режим с захваченным объемным зарядом лавины, рассмотрим $n^+ - p - p^+$ -диод, в котором ток резко возрастает в момент $t = 0$ (рис. 47, а). Если при $t = 0$ максимальная напряженность поля в диоде меньше критической напряженности \mathcal{E}_m , которую необходимо достичь для того, чтобы начался лавинный пробой, то со временем будет происходить увеличение напряженности электрического поля (рис. 47, б). Причина этого в том, что ток связан с полем следующим соотношением [20, 90]:

$$J = \epsilon_s \frac{d\mathcal{E}}{dt}. \quad (94)$$

Следовательно, напряженность электрического поля равна

$$\mathcal{E}(x, t) = \mathcal{E}_m - \frac{qN_A}{\epsilon_s} x + \frac{Jt}{\epsilon_s}, \quad (95)$$

где N_A — концентрация примеси в p -области. Поэтому промежуток времени, в течение которого электрическое поле достигает критического значения \mathcal{E}_m на заданном расстоянии x от контакта

в обедненной области, можно найти, приравняв $\mathcal{E}(x, t) = \mathcal{E}_m$, т. е.

$$\frac{qN_A x}{\epsilon_s} = Jt. \quad (96)$$

Продифференцировав соотношение (96), найдем скорость распространения границы области лавинного умножения (или скорость лавинной зоны) v_z (рис. 47, б):

$$v_z \equiv \frac{dx}{dt} = \frac{J}{qN_A}. \quad (97)$$

Так, например, для $N_A = 10^{15} \text{ см}^{-3}$ и $J = 10^4 \text{ А/см}^2$ скорость v_z значительно превышает скорость насыщения и равна $6 \cdot 10^7 \text{ см/с}$. То есть лавинная зона [19] (или ударный фронт лавины) быстро пересекает практически всю базовую область, которая заполняется высокопроводящей электронно-дырочной плазмой, причем пространственный заряд плазмы сильно снижает напряжение на диоде [18, 91].

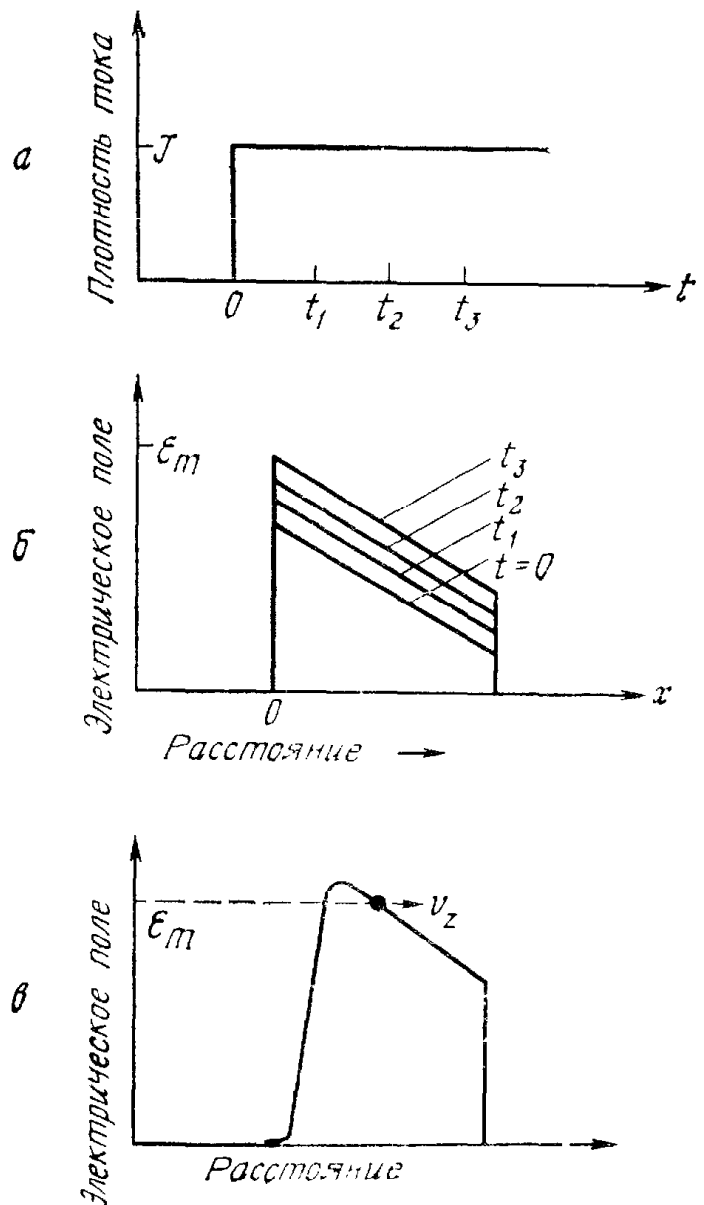


Рис. 47. Характеристики пролетного диода с захваченным объемным зарядом лавины [20, 94]:

а — плотность тока в зависимости от времени; б — идеальное распределение электрического поля, возрастающего со временем; в — ударный фронт лавины, распространяющийся в диоде со скоростью v_z , большей скорости насыщения v_g .

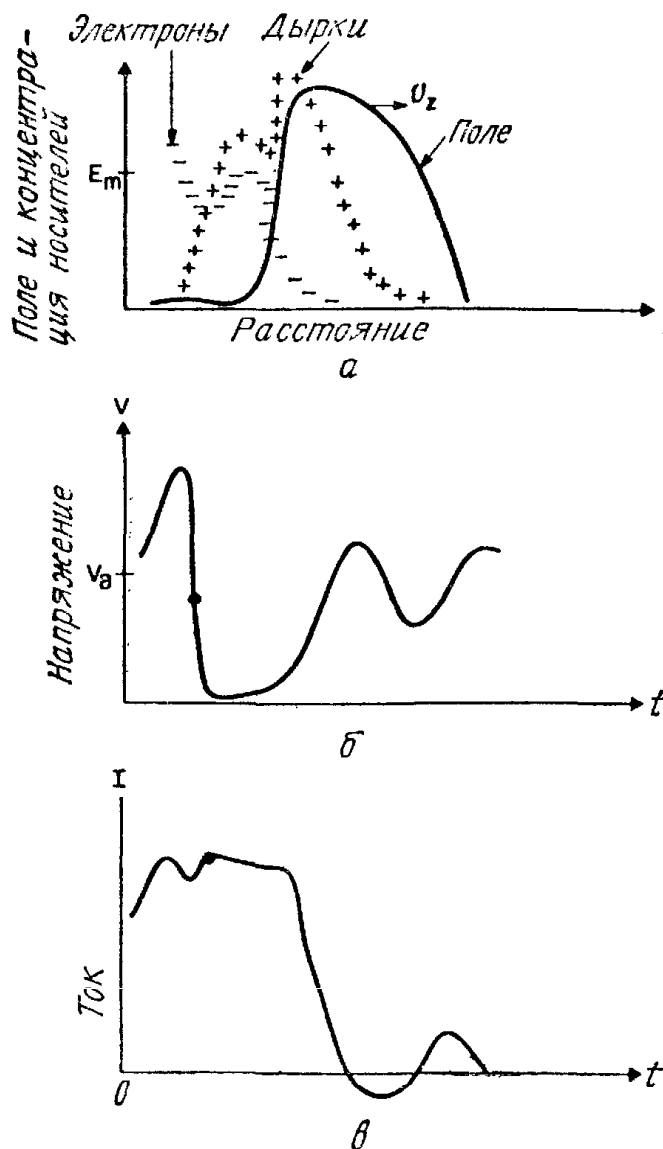
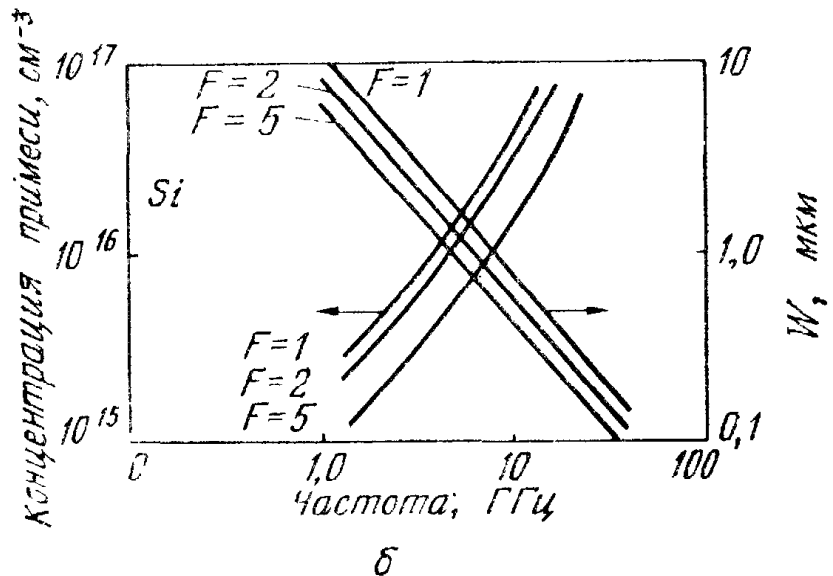
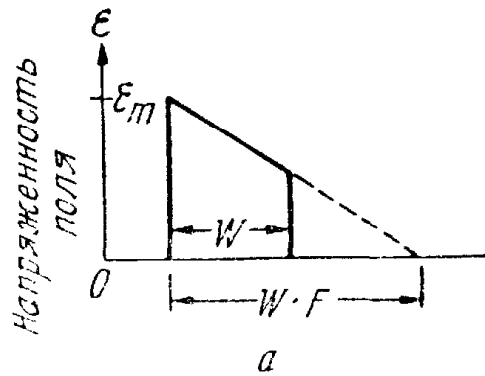


Рис. 48. Характеристики работы диода в режиме с захваченным объемным зарядом лавины [92]: а — результаты расчета на ЭВМ для мгновенного распределения поля и пространственного заряда; б и в — зависимости напряжения и тока от времени соответственно. Точкой на зависимостях на рис. б и в отмечен момент времени, для которого распределения приведены на рис. а.

Рассчитанные на ЭВМ распределения плотности заряда и электрического поля в диоде [92] в некоторый момент времени приведены на рис. 48, а. Зависимости напряжения и тока диода от времени приведены на рис. 48, б и в соответственно. Падение напряжения на диоде в начале периода может существенно превышать стационарное напряжение пробоя. Оно уменьшается после того, как лавинная зона пройдет через базу прибора, однако вскоре опять увеличится, поскольку электрическое поле поляризует объемный заряд подвижных носителей. Если скорость дрейфа носителей не слишком велика, то она пропорциональна напряженности электрического поля (т. е. $v = \mu \mathcal{E}$), а время пролета электронов и дырок через область дрейфа определяется подвижностями и будет значительно больше W/v_s .

Таким образом, в этом диоде возможна генерация колебаний значительно меньших частот, поскольку время рассасывания объемного заряда плазмы, равное отношению заряда Q к току I , может сильно превышать обычное пролетное время при больших напряженностях электрического поля. Этот диод является пролетным прибором в том смысле, что время запаздывания носите-

Рис. 49. Параметры прибора, используемые при расчете. Коэффициент перекрытия F определяется, как показано на рис. 49, а [40].



лей, возникающее из-за конечной скорости пересечения области дрейфа (т. е. время между инжекцией на одном контакте и сбором носителей заряда на другом контакте), используется здесь для получения соответствующего сдвига по фазе тока относительно приложенного напряжения.

Для работы диода необходима внешняя цепь, которая хорошо пропускает гармоники основной частоты при больших амплитудах напряжения. Кроме того, гармоник должно быть много, чтобы обеспечить требуемую задержку тока по фазе при таких низких частотах. Диод не должен выходить из строя при максимальном напряжении основной частоты. Необходимо, чтобы диод выдерживал смещение приблизительно в течение четверти периода до тех пор, пока пик гармонического напряжения не «запустит» процесс лавинного умножения. Так как напряжение на зажимах такого диода может быть малым при больших токах, и наоборот, к. п. д. довольно высок. Для некоторых кремниевых диодов были предсказаны к. п. д. [93, 94], превышающие 50 %.

Конструкция и характеристики таких диодов более сложные, поскольку сильная взаимосвязь прибора и внешней цепи предъявляет особые требования к большинству параметров. Для того чтобы уменьшить напряжение, необходимое для начала движения лавинной зоны, диод обычно изготавливают с большим коэффициентом

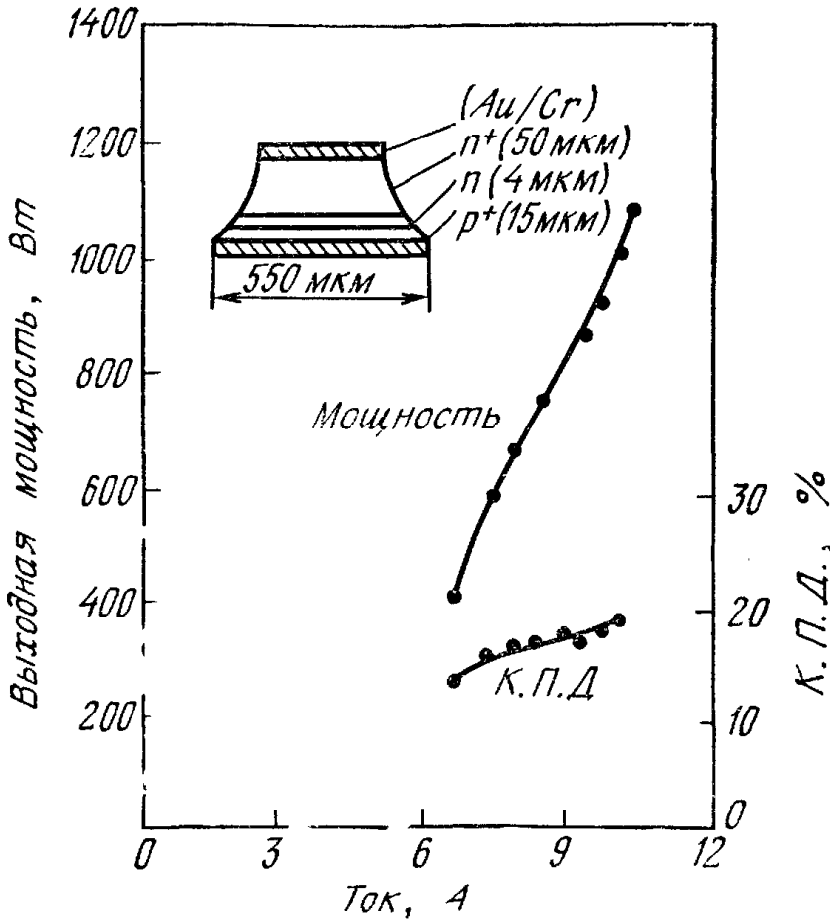


Рис. 50. Зависимость выходной мощности и к. п. д. от тока для пяти соединенных последовательно диодов при 1,1 ГГц [88].

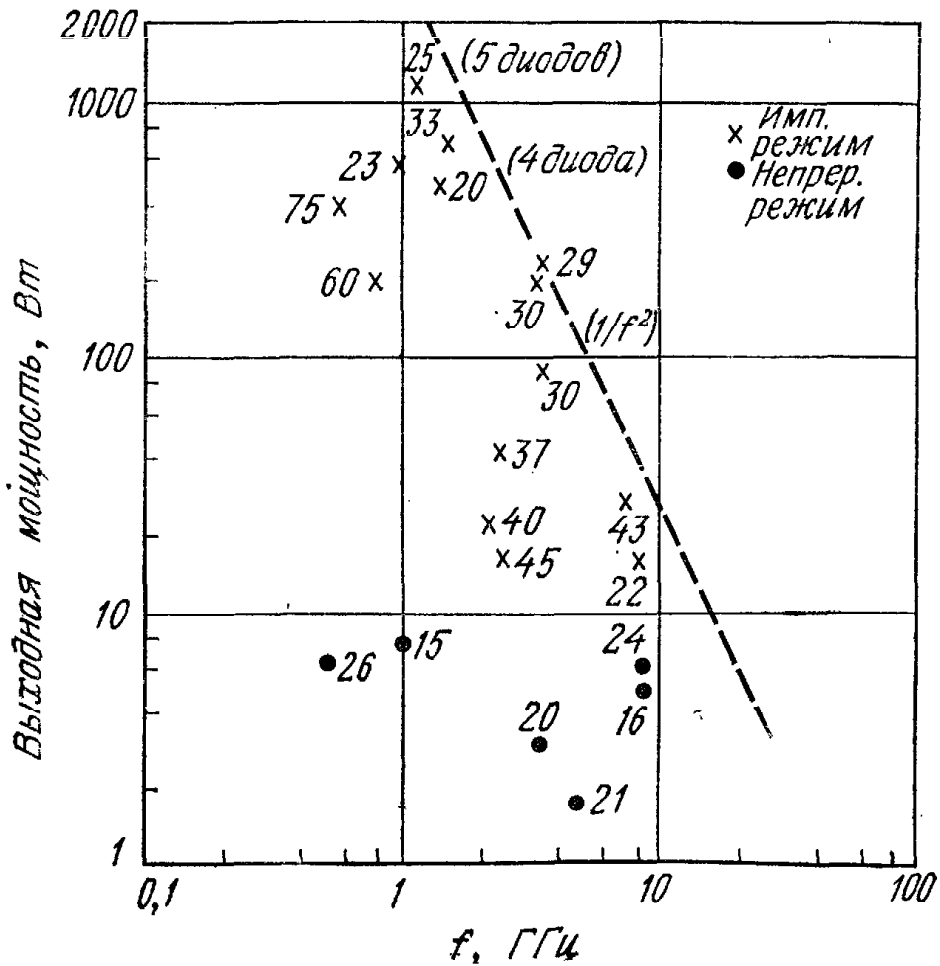


Рис. 51. Характеристики современных диодов с захваченным объемным зарядом лавины,

том перекрытия F (рис. 49, а). Чем больше коэффициент перекрытия F , тем больше прибор напоминает $p-i-n$ -диод. Зависимости ширины обедненной области и концентрации примеси от рабочей частоты для трех значений коэффициента перекрытия приведены на рис. 49, б.

Характеристики наиболее мощных современных диодов с захваченным объемным зарядом приведены на рис. 50 и 51 [88, 8, 95, 96]. Выходная мощность пяти соединенных последовательно диодов в импульсном режиме генерации достигает 1,2 кВт при к. п. д. 25 %. Отметим, что верхний предел рабочих частот близок к 10 ГГц, а наибольшие выходные мощности получены в диодах, работающих в импульсном режиме. Такие высокие значения мощности, генерируемой в импульсном режиме, существенно превышают выходные мощности других полупроводниковых СВЧ-приборов. Однако большая чувствительность к малым изменениям параметров цепи, рабочих условий или температуры затрудняет создание приборов со стабильными характеристиками при работе в изменяющихся внешних условиях.

ЛИТЕРАТУРА

1. Johnson R. L., DeLoach B. C., Jr., Cohen B. G. A Silicon Diode Oscillator, *Bell Syst. Tech. J.*, **44**, 369 (1965).
2. DeLoach B. C., Jr. The IMPATT Story, *IEEE Trans. Electron Devices*, **ED-23**, 57 (1976).
3. Shockley W. Negative Resistance Arising from Transit Time in Semiconductor Diodes, *Bell Syst. Tech. J.*, **33**, 799 (1954).
4. Read W. T. A Proposed High-Frequency Negative Resistance Diode, *Bell Syst. Tech. J.*, **37**, 401 (1958).
5. Lee C. A., Batdorf R. L., Wiegman W., Kaminsky G. The Read Diode and Avalanche, Transit-Time, Negative-Resistance Oscillator, *Appl. Phys. Lett.*, **6**, 89 (1965).
6. Misawa T. Negative Resistance on $p-n$ Junction under Avalanche Breakdown Conditions, Parts I and II, *IEEE Trans. Electron Devices*, **ED-13**, 137 (1966).
7. Gilden M., Hines M. F. Electronic Tuning Effects in the Read Microwave Avalanche Diode, *IEEE Trans. Electron Devices*, **ED-13**, 169 (1966).
8. Sze S. M., Ryder R. M. Microwave Avalanche Diodes, *Proc. IEEE*, **59**, 1140 (1971).
9. Schroeder W. E., Haddad G. I. Nonlinear Properties of IMPATT Devices, *Proc. IEEE*, **61**, 153 (1973).
10. Special Issue on Solid-State Microwave Millimeter-Wave Power Generation, Amplification, and Control, *IEEE Trans. Microwave Theory Tech.*, **MTT-27** (May 1979).
- 10a. Eastman L. F. Microwave Semiconductor Devices: State-of-the-Art and Limiting Effects, *IEEE Tech. Dig., Int. Electron Devices Meet.*, 1978, p. 364.
11. Brackett C. A. The Elimination of Tuning Induced Burnout and Bias Circuit Oscillation in IMPATT Oscillators, *Bell Syst. Tech. J.*, **52**, 271 (1973).
12. Coleman D. J., Jr., Sze S. M. The Baritt Diode — A New Low Noise Microwave Oscillator, *IEEE Device Res. Conf.*, Ann Arbor, Mich., June 28, 1971; A Low-Noise Metal — Semiconductor — Metal (MSM) Microwave Oscillator, *Bell Syst. Tech. J.*, **50**, 1695 (1971).

13. Ruegg H. W. A Proposed Punch-Through Microwave Negative Resistance Diode, *IEEE Trans. Electron Devices*, **ED-15**, 577 (1968).
14. Wright G. T. Punch-Through Transit-Time Oscillator, *Electron. Lett.*, **4**, 543 (1968).
15. Sobol H., Sterzer F. Microwave Power Sources, *IEEE Spectrum*, **9**, 20 (1972).
16. East J. R., Nguyen-Ba H., Haddad G. I. Microwave and mm Wave Baritt Doppler Detector, *Microwave J.*, **19**, 51 (1976).
17. Prager H. J., Chang K. K. N., Weisbrod S. High Power, High Efficiency Silicon Avalanche Diodes at Ultrahigh Frequencies, *Proc. IEEE*, **55**, 586 (1967). [Имеется перевод: Прагер, Чанг, Вайсброд. Мощные кремниевые лавинные диоды с высоким к. п. д. для ультравысоких частот. — ТИИЭР, 1967, 55, № 4, 119—121.]
18. Scharfetter D. L., Bartelink D. J., Gummel H. K., Johnston R. L. Computer Simulation of Low-Frequency High-Efficiency Oscillation in Germanium IMPATT Diodes, *IEEE Trans. Electron Devices*, **ED-15**, 691 (1968).
19. Bartelink D. J., Scharfetter D. L. Avalanche Shock Fronts in $p - n$ Junctions, *Appl. Phys. Lett.*, **14**, 420 (1969).
20. Clorfeine A. S., Ikola R. J., Napoli L. S. A Theory for the High-Efficiency Mode of Oscillation in Avalanche Diodes, *RCA Rev.*, **30**, 397 (1969).
21. Chang K. K. N., Kawamoto H., Prager H. J., Reynolds J. F., Rosen A., Mikenas V. A. TRAPATT Amplifiers for Phased-Array Radar Systems, *Microwave J.*, **16**, 27 (1973).
22. Salmer G., Pribetich H., Farrayre A., Kramer B. Theoretical and Experimental Study of GaAs Impatt Oscillator Efficiency, *J. Appl. Phys.*, **44**, 314 (1973).
23. Sze S. M., Gibbons G. Avalanche Breakdown Voltages of Abrupt and Linearly Graded $p - n$ Junctions in Ge, Si, GaAs, and GaP, *Appl. Phys. Lett.*, **8**, 111 (1966).
24. Schroeder W. E., Haddad G. I. Avalanche Region Width in Various Structures of IMPATT Diodes, *Proc. IEEE*, **59**, 1245 (1971). [Имеется перевод: Шредер, Хаддад, Ширина слоя умножения в IMPATT-диодах различных структур. — ТИИЭР, 1971, 59, № 8, 129—132.]
25. Gibbons G., Sze S. M. Avalanche Breakdown in Read and $p - i - n$ Diodes, *Solid State Electron.*, **11**, 225 (1968).
26. Crowell C. R., Sze S. M. Temperature Dependence of Avalanche Multiplication in Semiconductors, *Appl. Phys. Lett.*, **9**, 242 (1966).
27. Bowers H. C. Space-Charge-Limited Negative Resistance in Avalanche Diodes, *IEEE Trans. Electron Devices*, **ED-15**, 343 (1968).
28. Sze S. M., Shockley W. Unit-Cube Expression for Space-Charge Resistance, *Bell Syst. Tech. J.*, **46**, 837 (1967).
29. Weissglas P. Avalanche and Barrier Injection Devices, in Howes M. J., Morgan D. V., Eds., *Microwave Devices — Device Circuit Interactions*, Wiley, N. Y., 1976, Chap. 3.
30. Misawa T. Multiple Uniform Layer Approximation in Analysis of Negative Resistance in $p - n$ Junctions in Breakdown, *IEEE Trans. Electron Devices*, **ED-14**, 795 (1967).
31. Scharfetter D. L., Gummel H. K. Large-Signal Analysis of a Silicon Read Diode Oscillator, *IEEE Trans. Electron Devices*, **ED-16**, 64 (1969).
32. Swan C. B., Misawa T., Marinaccio L. Composite Avalanche Diode Structures for Increased Power Capability, *IEEE Trans. Electron Devices*, **ED-14**, 684 (1967).
33. Scharfetter D. L. Power-Impedance-Frequency Limitation of IMPATT Oscillators Calculated from a Scaling Approximation, *IEEE Trans. Electron Devices*, **ED-18**, 536 (1971).
34. Thim H. W., Poetze H. W. Search for Higher Frequencies in Microwave Semiconductor Devices, 6th Eur. Solid State Device Res. Conf., *Inst. Phys. Conf. Ser.*, **32**, 73 (1977).

- 35 Seidel T. E., Niehaus W. C., Iglestas D. E. Double-Drift Silicon IMPATTs at X Band, *IEEE Trans. Electron Devices*, **ED-21**, 523 (1974).
- 36 Kuvvas R. L. Carrier Transport in the Drift Region of Read-Type Diodes, *IEEE Trans. Electron Devices*, **ED-25**, 660 (1978).
- 37 Hirachi Y., Kobayashi K., Ogasawara K., Toyama Y. A New Concept for High Efficiency Operation of Hi — Lo-Type GaAs IMPATT Diodes, *IEEE Trans. Electron Devices*, **ED-25**, 666 (1978).
- 38 Blakey P. A., Culshaw B., Giblin R. A. Comprehensive Models for the Analysis of High Efficiency GaAs IMPATTs, *IEEE Trans. Electron Devices*, **ED-25**, 674 (1978).
- 39 Nishitani K., Sawano H., Ishihara O., Ishii T., Mitsui S. Optimum Design for High-Power and High Efficiency GaAs Hi — Lo IMPATT Diodes, *IEEE Trans. Electron Devices*, **ED-26**, 210 (1979).
- 40 Evans W. J. Avalanche Diode Oscillators, in Hershberger W. D., Ed. *Solid State and Quantum Electronics*, Wiley, N. Y., 1971.
- 41 Misawa T. Saturation Current and Large Signal Operation of a Read Diode, *Solid State Electron.*, **13**, 1363 (1970).
- 42 Kovel S. R., Gibbons G. The Effect of Unswept Epitaxial Material on the Microwave Efficiency of IMPATT Diodes, *Proc. IEEE Lett.*, **55**, 2066 (1967).
- 43 Aono Y., Okuto Y. Effect of Undepleted High Resistivity Region on Microwave Efficiency of GaAs IMPATT Diodes, *Proc. IEEE*, **63**, 724 (1975). [Имеется перевод: Аоно, Окуто. Влияние нейтрального участка высокоомной области на к. п. д. IMPATT-диодов из арсенида галлия. — ТИИЭР, 1975, 63, № 3, 56—57.]
- 44 DeLoach B. C., Jr. Thin Skin IMPATTs, *IEEE Trans. Microwave Theory Tech.*, **MTT-18**, 72 (1970).
- 45 Misawa T. High Frequency Fall-Off of IMPATT Diode Efficiency, *Solid State Electron.*, **15**, 457 (1972).
- 46 Elta M. E., Haddad G. I. Mixed Tunneling and Avalanche Mechanisms in $p - n$ Junctions and Their Effects on Microwave Transit-Time Devices, *IEEE Trans. Electron Devices*, **ED-25**, 694 (1978).
- 47 Nishizawa J. I., Motoya K., Okuno Y. 200 GHz TUNNETT Diodes, *Jpn. J. Appl. Phys.*, **17**, Suppl. 17—1, 167 (1977).
- 48 Berenz J. J., Kinoshita J., Hierl T. L., Lee C. A. Orientation Dependence of n -type GaAs Intrinsic Avalanche Response Time, *Electron. Lett.*, **15**, 150 (1979).
- 49 Olson H. M. A Mechanism for Catastrophic Failure of Avalanche Diodes, *IEEE Trans. Electron Devices*, **ED-22**, 842 (1975).
- 50 Hines M. F. Noise Theory of Read Type Avalanche Diode, *IEEE Trans. Electron Devices*, **ED-13**, 158 (1966).
- 51 Gummel H. K., Blue J. L. A Small-Signal Theory of Avalanche Noise of IMPATT Diodes, *IEEE Trans. Electron Devices*, **ED-14**, 569 (1967).
- 52 Blue J. L. Preliminary Theoretical Results on Low Noise GaAs IMPATT Diodes, *IEEE Device Res. Conf.*, Seattle, Wash., June 1970.
- 53 Irvin J. C., Coleman D. J., Johnson W. A., Tatsuguchi I., Decker D. R., Dunn C. N. Fabrication and Noise Performance of High-Power GaAs IMPATTs, *Proc. IEEE*, **59**, 1212 (1971). [Имеется перевод: Ирвин, Коулмен, Джонсон и др. Изготовление и шумовые свойства мощных IMPATT-диодов из арсенида галлия. — ТИИЭР, 1971, 59, № 8, 95—99.]
- 54 Misawa T. Microwave Si Avalanche Diode With Nearly-Abrupt-Type Junction, *IEEE Trans. Electron Devices*, **ED-14**, 580 (1967).
- 55 Sze S. M., Lepselter M. P., MacDonald R. W. Metal — Semiconductor IMPATT Diode, *Solid State Electron.*, **12**, 107 (1969).
- 56 De Nobel D., Kock H. G. A Silicon Schottky Barrier Avalanche Transit Time Diode, *Proc. IEEE*, **57**, 2088 (1969). [Имеется перевод: де Нобел, Кок. Кремниевый лавинно-пролетный диод с барьером Шоттки. — ТИИЭР, 1969, 57, № 10, 200.]

57. Kim C. K., Armstrong L. D. GaAs Schottky-Barrier Avalanche Diode, *Solid State Electron.*, **13**, 53 (1970).
58. Misawa T. Minority Carrier Storage and Oscillation Efficiency in Read Diodes, *Solid State Electron.*, **13**, 1369 (1970)
59. Niehaus W. C., Schwartz B. A Self-Limiting Anodic Etch-to-Voltage Technique for Fabrication of Modified Read IMPATTs, *Solid State Electron.*, **19**, 175 (1976).
60. Mahoney G. E. Retardation of IMPATT Diode Aging by Use of Tungsten in the Electrodes, *Appl. Phys. Lett.*, **27**, 613 (1975).
61. Heaton J. L., Walline R. E., Carroll J. F. Low — High — Low Profile GaAs IMPATT Reliability, *IEEE Trans. Electron Devices*, **ED-26**, 96 (1979).
62. Scharfetter D. L., Evans W. J., Johnston R. L. Double-Drift-Region (p^+pnn^+) Avalanche Diode Oscillators, *Proc. IEEE*, **58**, 1131 (1970). [Имеется перевод: Шарфеттер, Эванс, Джонстон ГЛПД на диодах с двумя пролетными пространствами (структуры $p^+ - p - n - n^+$). — ТИИЭР 1970, 58, № 8, 113—114.]
63. Cho A. Y., Dunn C. N., Kuvass R. L., Schroeder W. E. GaAs IMPATT Diodes, Prepared by Molecular Beam Epitaxy, *Appl. Phys. Lett.*, **25**, 224 (1974).
64. Liao S. Y. *Microwave Devices and Circuits*, Prentice-Hall, Englewood Cliffs, N. J., 1980.
65. Iglesias D. E. Circuit for Testing High Efficiency IMPATT Diodes, *Proc. IEEE*, **55**, 2065 (1967). [Имеется перевод: Иглесиас. Схема для проверки лавинно-пролетных диодов с большим к. п. д. — ТИИЭР, 1967, 55, № 11, 319—320.]
66. Misawa T., Marinaccio L. P. A 1/4 Watt Si $p - v - n$ X-Band IMPATT Diode, Int. *Electron Device Meet.*, Washington, D. C., Oct. 1966.
67. Bowman L. S., Burrus C. A., Jr., Pulse-Driven Silicon $p - n$ Junction Avalanche Oscillators for the 0,9 to 20 mm Band *IEEE Trans. Electron Devices*, **ED-14**, 411 (1967).
68. Ino M., Ishibashi T., Ohmori M. Submillimeter Wave Si p^+pn^+ IMPATT Diodes, *Jpn. J. Appl. Phys.*, **16** Suppl. 16-1, 89 (1977).
69. Pribetich J., Chive M., Constant E., Farrayre A. Design and Performance, of Maximum-Efficiency Single and Double-Drift-Region GaAs IMPATT Diodes in the 3—18 HGz Frequency Range, *J. Appl. Phys.*, **49**, 5584 (1978).
70. Seidel T. E., Davis R. E., Iglesias D. E. Double-Drift-Region Ion-Implanted Millimeter-Wave IMPATT Diodes, *Proc. IEEE*, **59**, 1222 (1971). [Имеется перевод: Зейдел, Дэвис, Иглесиас. Двухпролетные IMPATT-диоды миллиметрового диапазона, изготовленные методом ионного внедрения. — ТИИЭР, 1971, 59, № 8, 107—114.]
71. Cohen E. D. Trapatts and Impatts — State of the Art and Application, *Microwave J.*, **20**, 22 (1977).
72. Grexy K. W. Recent Advances in Microwave Devices, *Jpn. J. Appl. Phys.*, **16**, Suppl. 16-1, 81 (1974).
73. Bozler C. O., Donnelly J. P., Murphy R. A., Laton R. W., Subduny R. W. High-Efficiency Ion-Implanted Lo — Hi — Lo GaAs IMPATT Diodes, *Appl. Phys. Lett.*, **29**, 125 (1976).
74. Hierl T. L., Berenz J. J., Kinoshita J., Zubeck I. U. High Efficiency Pulsed GaAs Read IMPATT Diodes, *Electron. Lett.*, **14**, 155 (1978).
75. Potteiger D. C., private communication.
76. Sze S. M., Coleman D. J., Loya A. Current Transport in Metal — Semiconductor — Metal (MSM) Structures, *Solid State Electron.*, **14**, 1209 (1971).
77. Chu J. L., Persky G., Sze S. M. Thermionic Injection and Space-Charge-Limited Current in Reach-Through p^+np^+ Structures, *J. Appl. Phys.*, **43**, 3510 (1972).
78. Eknayan O., Yang E. S., Sze S. M. Multilayered Ion-Implanted Baritt Diodes with Improved Efficiency, *Solid State Electron.*, **20**, 291 (1977).

79. De Cogan D. The Punch-Through Diode, *Microelectronics*, 8, 20 (1976).
80. Nguyen-Ba H., Haddad G. I. Effects of Doping Profile on the Performance of Baritt Devices, *IEEE Trans. Electron Devices*, ED-24, 1154 (1977).
81. Chu J. L., Sze S. M. Microwave Oscillation in *pnp* Reach-Through Baritt Diodes, *Solid State Electron.*, 16, 85 (1973).
82. Sjolund A. Small-Signal Analysis of Punch-Through Injection Microwave Devices, *Solid State Electron.*, 16, 559 (1973).
83. Kwok S. P., Haddad G. I. Power Limitation in Baritt Devices, *Solid State Electron.*, 19, 795 (1976).
84. Eknayan O., Sze S. M., Yang E. S. Microwave Baritt Diode with Retarding Field — An Investigation, *Solid State Electron.*, 20, 285 (1977).
85. Ahmad S., Freyer J. High-Power Pt Schottky Baritt Diodes, *Electron. Lett.*, 12, 238 (1976).
86. Sitch J. E., Majerfeld A., Robson P. N., Hasegawa F. Transit — Time — Induced Microwave Negative Resistance in GaAlAs — GaAs Heterostructure Diodes, *Electron. Lett.*, 11, 457 (1975).
87. Sitch J. E., Robson P. N. Efficiency of BARITT and DOVETT Oscillators, *Solid State Electron Devices*, 1, 31 (1976).
88. Liu S. G., Risko J. J. Fabrication and Performance of Kilowatt L-Band Avalanche Diodes, *RCA Rev.*, 31, 3 (1970).
89. Kostichack D. F. UHF Avalanche Diode Oscillator Providing 400 Watts Peak Power and 75 Percent Efficiency, *Proc. IEEE Let.*, 58, 1282 (1970).
90. DeLoach B. C., Scharfetter Jr. and D. L. Device Physics of TRAPATT Oscillators, *IEEE Trans. Electron Devices*, ED-17, 9 (1970).
91. Johnston R. L., Scharfetter D. L., Bartelink D. J. High-Efficiency Oscillations in Germanium Avalanche Diodes Below the Transit-Time Frequency, *Proc. IEEE*, 56, 1611 (1968). [Имеется перевод: Джонстон, Шарфеттер, Бартелинк. Генерация колебаний германиевыми лавинно-пролетными диодами с высоким к. п. д. на частотах ниже номинальной. — ТИИЭР, 1968, № 9, 222—225.]
92. Evans W. J. Circuits for High-Efficiency Avalanche-Diode Oscillators, *IEEE Trans. Microwave Theory Tech.*, MTT-17, 1060 (1969).
93. Scharfetter D. L. Power-Frequency Characteristics of the PRAPATT Diode Mode of High Efficiency Power Generation in Germanium and Silicon Avalanche Diodes, *Bell Syst. Tech. J.*, 49, 799 (1970).
94. Trew R. J., Haddad G. I., Masnari N. A. A Simplified Model of a TRAPATT Diode, *IEEE Trans. Electron Devices*, ED-23, 28 (1976).
95. Obah C. O. G., Benko E., Midford T. A., Bowers H. C., Chao P. Y. Single-Diode 0,5 kW TRAPATT Oscillators, *Electron. Lett.*, 10, 430 (1974).
96. Ying R. S., Fong T. T. C-Band Complementary TRAPATT Diodes, *Proc. IEEE*, 62, 287 (1974). [Имеется перевод: Ин, Фын. TRAPATT-диоды С-диапазона с дополнительной структурой. — ТИИЭР, 1974, 62, № 3, 147—148.]
97. Ahmad S., Freyer J. Design and Development of High-Power Microwave Silicon BARITT Diodes, *IEEE Trans. Electron Devices*, ED-26, 1370 (1979).

ПРИБОРЫ НА ЭФФЕКТЕ МЕЖДОЛИННОГО ПЕРЕХОДА ЭЛЕКТРОНОВ

11.1. ВВЕДЕНИЕ

Приборы на эффекте междолинного перехода электронов¹⁾ (одни из важнейших СВЧ-приборов) широко используются в качестве гетеродина и усилителя мощности в диапазоне частот 1—100 ГГц. Твердотельные генераторы на этих приборах применяются в радарх, системах раннего оповещения и контрольно-измерительной аппаратуре.

В 1963 г. Ганн [1] обнаружил генерацию когерентных СВЧ-колебаний при наложении на произвольно ориентированный короткий образец из арсенида галлия или фосфида индия *n*-типа постоянного электрического поля, превышающего пороговое значение, равное нескольким киловольтам на 1 см. Частота колебаний приблизительно равнялась величине, обратной времени пролета носителей через образец. Кремер [2] указал на то, что все наблюдаемые свойства этого процесса согласуются с результатами теории отрицательного дифференциального удельного сопротивления, предложенной независимо Ридли и Уоткинсом [3, 3а], а также Хилсумом [4, 4а]. Причиной возникновения отрицательного дифференциального удельного сопротивления служит стимулированный электрическим полем переход электронов зоны проводимости из низкоэнергетической долины, в которой их подвижность относительно велика, в более высокоэнергетические долины с меньшими подвижностями. Хатсон и др. [5] обнаружили влияние давления на характеристики прибора из арсенида галлия, а Аллен и др. [6], исследуя генерацию в тройных соединениях $GaAs_{1-x}P_x$, показали, что величина порогового электрического поля уменьшается при уменьшении энергетического зазора между основным и побочным минимумами зоны проводимости. Все это убедительно доказывало, что причиной ганновской генерации является эффект междолинного перехода электронов.

Этот эффект также называют механизмом Ридли—Уоткинса—Хилсума или эффектом Ганна. Приборы на эффекте междолинного перехода электронов подробно описаны в работах [7—9]. В работе [10] дан обзор твердотельных СВЧ-генераторов.

¹⁾ В зарубежной литературе эти приборы называются TED (Transferred-Electron Devices). — *Прим. перев.*

11.2. ЭФФЕКТ МЕЖДОЛИННОГО ПЕРЕХОДА ЭЛЕКТРОНОВ

11.2.1. Объемное отрицательное дифференциальное удельное сопротивление

Физические механизмы возникновения объемного отрицательного дифференциального удельного сопротивления весьма различны. Одним из важнейших служит эффект междолинного перехода электронов. Поскольку случайная флюктуация концентрации свободных носителей в любой точке образца с объемным отрицательным удельным сопротивлением приводит к моментальному возникновению пространственного заряда, величина которого возрастает со временем по экспоненциальному закону, такой полупроводник является по своей природе нестабильным. Запишем уравнение непрерывности в одномерном случае ¹⁾

$$\frac{\partial n}{\partial t} + \frac{1}{q} \frac{\partial J}{\partial x} = 0. \quad (1)$$

Малое локальное отклонение концентрации основных носителей от однородной равновесной n_0 приводит к появлению пространственного заряда плотностью $n - n_0$. Уравнение Пуассона и выражение для плотности тока имеют вид

$$\frac{\partial \mathcal{E}}{\partial x} = \frac{q(n - n_0)}{\epsilon_s}, \quad (2)$$

$$J = \mathcal{E}/\rho - qD \frac{\partial n}{\partial x}, \quad (3)$$

где ϵ_s — диэлектрическая проницаемость, ρ — удельное сопротивление и D — коэффициент диффузии. Продифференцировав выражение (3) по x и учитывая уравнение Пуассона, получим

$$\frac{1}{q} \frac{\partial J}{\partial x} = \frac{n - n_0}{\rho \epsilon_s} - D \frac{\partial^2 n}{\partial x^2}. \quad (4)$$

Подстановка этого выражения в уравнение (1) приводит к уравнению

$$\frac{\partial n}{\partial t} + \frac{n - n_0}{\rho \epsilon_s} - D \frac{\partial^2 n}{\partial x^2} = 0. \quad (5)$$

Уравнение (5) можно решить методом разделения переменных. Если концентрация зависит только от координаты, то решение уравнения (5) имеет вид

$$n - n_0 = (n - n_0)_{x=0} \exp(-x/L_D), \quad (6)$$

¹⁾ Во избежание появления многочисленных отрицательных знаков мы будем считать в этой главе заряд электронов положительным.

где L_D — дебаевская длина, равная

$$L_D \equiv \sqrt{\frac{kT\epsilon_s}{q^2 n_0}} \quad (7)$$

и определяющая расстояние, на котором затухают малые флуктуации заряда. Если концентрация зависит только от времени, то решение уравнения (5) имеет вид

$$n - n_0 = (n - n_0)_{t=0} \exp(-t/\tau_R), \quad (8)$$

где τ_R — время диэлектрической релаксации:

$$\tau_R \equiv \rho\epsilon_s = \epsilon_s/q\mu n \approx \epsilon_s/q\mu n_0, \quad (9)$$

за которое происходит рассасывание пространственного заряда, если дифференциальное удельное сопротивление и дифференциальная подвижность μ положительны. Однако в полупроводнике с отрицательным дифференциальным сопротивлением (ОДС) любая флуктуация концентрации будет нарастать с постоянной времени, равной $|\tau_R|$, а не затухать.

Приборы с объемным ОДС можно разделить на две группы: приборы, управляемые напряжением (с N-образными характеристиками), и приборы, управляемые током (с S-образными характеристиками). Типичные зависимости плотности тока от напряженности электрического поля, характерные для этих двух типов приборов, приведены на рис. 1, а и б [11]. Также показано удельное дифференциальное сопротивление ($\partial\mathcal{E}/\partial J$) при изменении напряжения смещения от нуля до порогового, при котором начинается участок отрицательного сопротивления (рис. 1, в и г). В приборах с N-образными характеристиками зависимость напряженности электрического поля от плотности тока (рис. 1, а) неоднозначна, а в приборах с S-образными J — \mathcal{E} -характеристиками наоборот (рис. 1, б).

Характеристики на рис. 1 подобны приведенным в главах, посвященных контактному явлению. Так, например, туннельный диод является прибором с N-образной J — \mathcal{E} -характеристикой, а тиристор — прибором с S-образной характеристикой. Объемное ОДС обусловлено микроскопическими свойствами полупроводника: стимулированным электрическим полем захватом носителей на глубокие уровни [12], ударной ионизацией мелких примесных уровней в компенсированных полупроводниках [13] и переходом электронов из основного минимума зоны проводимости в побочные (эффект междолинного перехода электронов). Мы рассмотрим основные характеристики приборов с ОДС независимо от конкретных физических механизмов возникновения отрицательного сопротивления.

В однородном полупроводнике с ОДС распределение электрического поля, соответствующее устойчивому состоянию, существ-

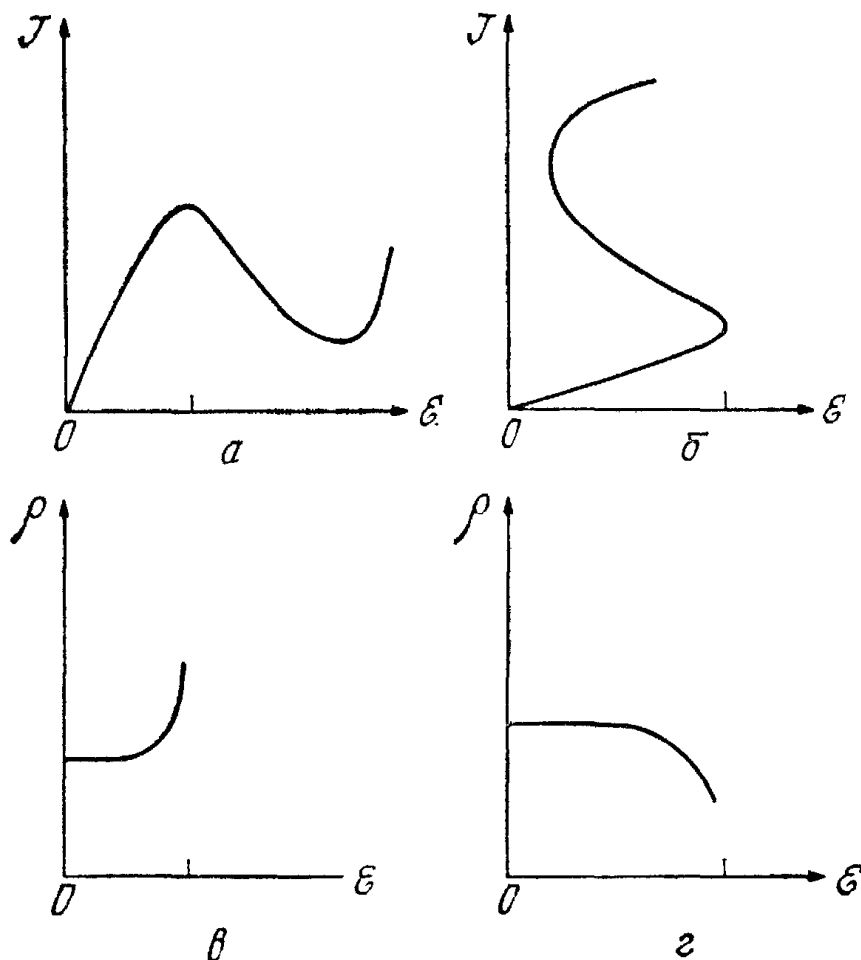


Рис. 1. Зависимости плотности тока и дифференциального удельного сопротивления от напряженности электрического поля.

a и v — для приборов с N-образными $J - \mathcal{E}$ -характеристиками; b и z — для приборов с S-образными $J - \mathcal{E}$ -характеристиками.

венно неоднородно. Ниже мы приведем простые соображения, из которых следует, что в приборах с N-образными $J - \mathcal{E}$ -характеристиками будут образовываться домены высокого электрического поля (или обогащенные слои), а в приборах с S-образными характеристиками — шнуры тока с высокой плотностью [11]. При описании различных режимов работы приборов в последующих разделах мы лишь видоизменим нашу простую модель.

Как следует из рис. 1, v , положительное дифференциальное удельное сопротивление увеличивается с ростом электрического поля, т. е. $d\rho/d\mathcal{E} > 0$. Если напряженность электрического поля в некоторой области полупроводника немного больше средней (рис. 2, a), то ее удельное сопротивление также больше. Поэтому плотность тока будет меньше. Это приведет к увеличению размеров области и образованию домена (области высокого электрического поля), вне которого поле относительно мало. Поверхности, ограничивающие домен высокого поля, эквипотенциальны и поэтому перпендикулярны направлению электрического тока.

Для прибора с S-образной характеристикой первоначальная величина положительного дифференциального удельного сопро-

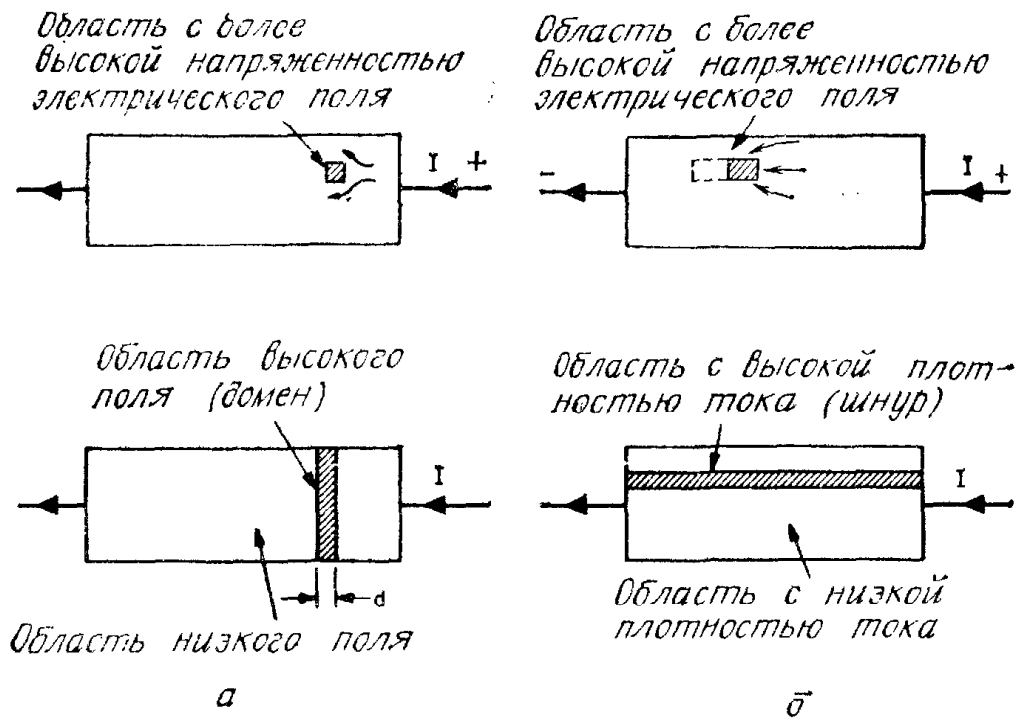


Рис. 2. Формирование домена высокого электрического поля (а) в образце с отрицательным дифференциальным удельным сопротивлением, управляемым напряжением (N-образной характеристикой), а также формирование шнура тока с высокой плотностью (б) в образце с отрицательным дифференциальным удельным сопротивлением, управляемым током (S-образной характеристикой).

тивления уменьшается при увеличении напряженности электрического поля, т. е. $dp/d\mathcal{E} < 0$. Если поле в некоторой области полупроводника немного больше среднего, то удельное сопротивление этой области меньше. Поэтому ток будет втекать в эту область, что приведет к ее удлинению в направлении электрического поля и образованию шнура тока с высокой плотностью.

Для более подробного изучения нестабильностей объемного заряда [14] в приборах с N-образными $J-\mathcal{E}$ -характеристиками обратимся к рис. 3. На рис. 3, а приведена типичная мгновенная зависимость J от \mathcal{E} , а на рис. 3, б — схема прибора. Предположим, что в точке А образовался избыток отрицательного заряда (см. рис. 3, б). Причиной этого могут быть случайные шумы или неоднородное легирование (рис. 3, в). Проинтегрировав один раз уравнение Пуассона, мы получим распределение электрического поля в образце, приведенное на рис. 3, г, причем поле слева от точки А меньше чем справа. Если постоянное напряжение на приборе соответствует точке \mathcal{E}_A на $J-\mathcal{E}$ -кривой, то втекающий в точку А ток будет больше вытекающего, что приведет к накоплению отрицательного пространственного заряда в точке А. При этом электрическое поле слева от точки А становится еще меньше, что способствует дальнейшему накоплению заряда. Этот процесс продолжается до тех пор, пока

напряженности электрического поля в области вне точки A не будут соответствовать точкам 1 и 2 J — \mathcal{E} -кривой на рис. 3, a , в которых токи равны, а дифференциальное удельное сопротивление положительно. В результате образуется движущийся слой, обогащенный носителями. Этот процесс, безусловно, зависит от числа электронов в кристалле, которое должно быть достаточ-

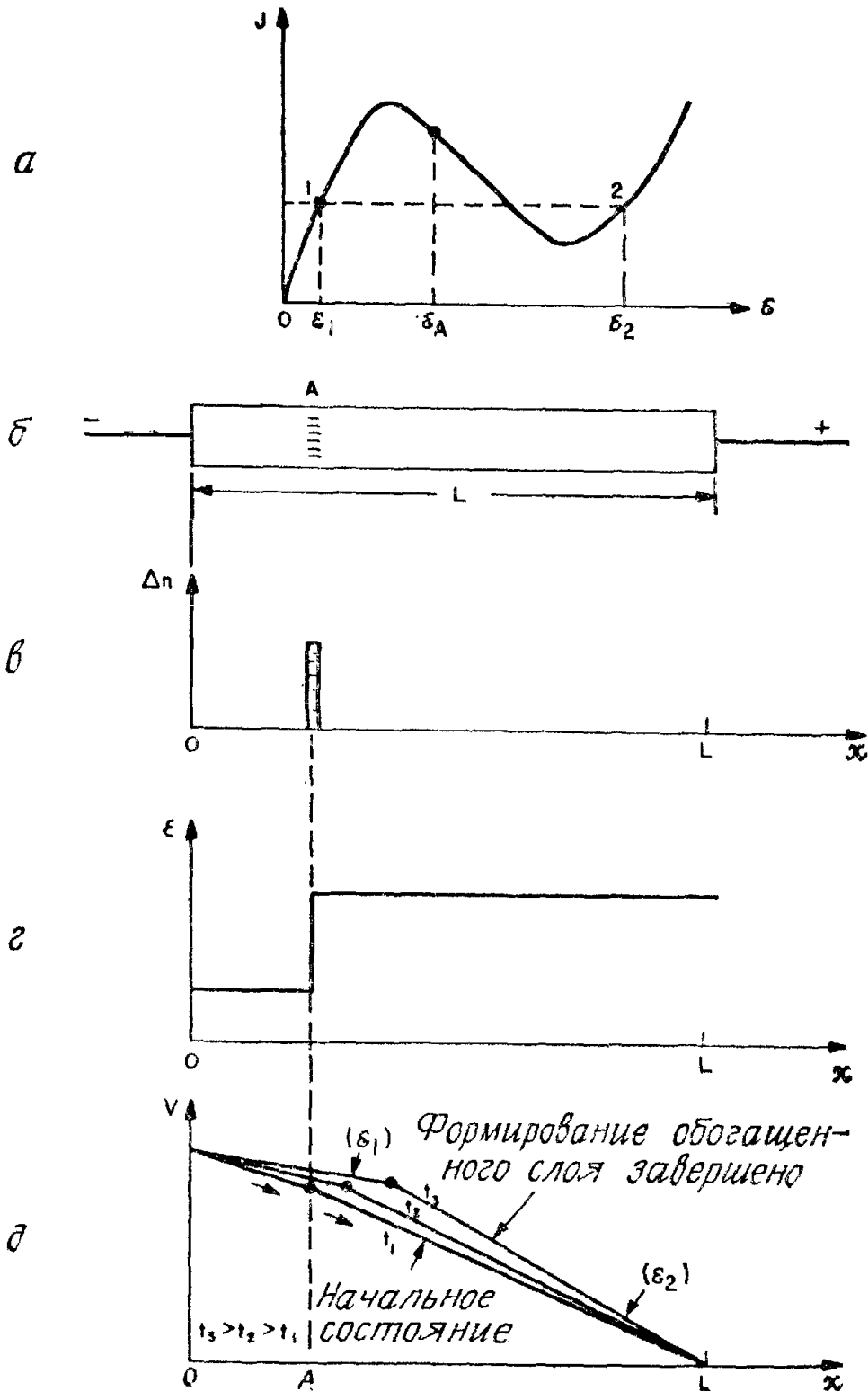


Рис. 3. Формирование обогащенного электронами слоя в возмущенной среде с ОДС [14].

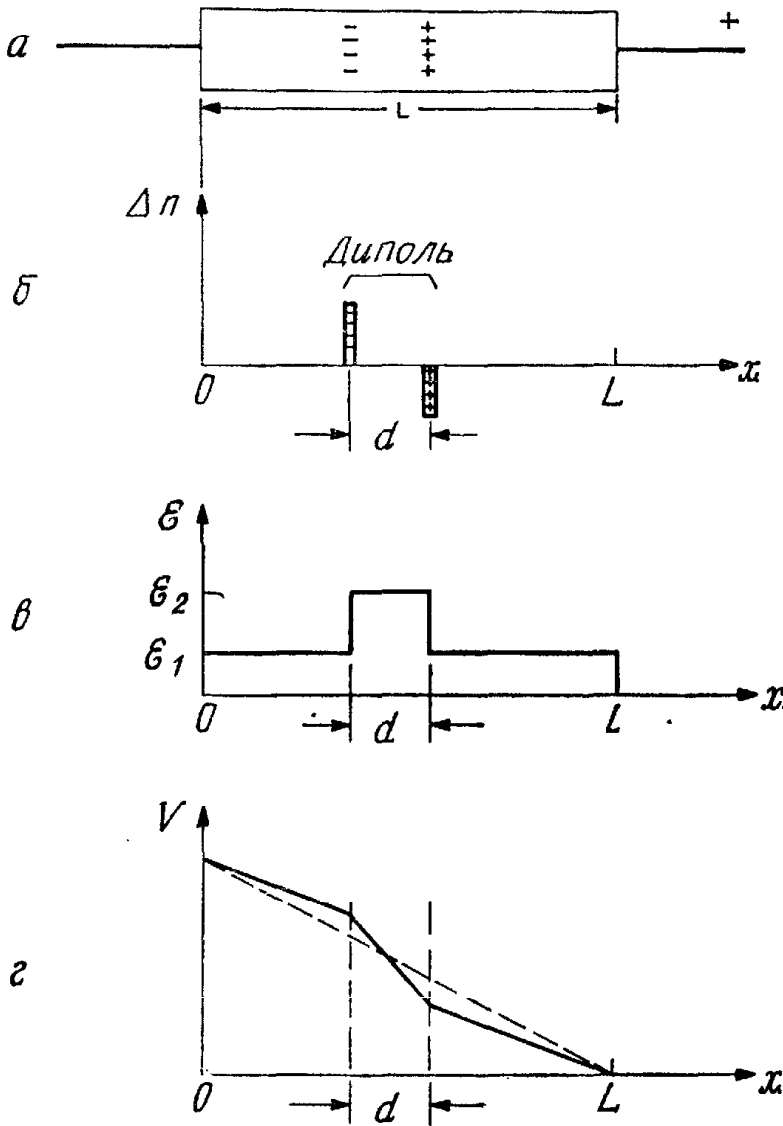


Рис. 4. Формирование электронного дипольного слоя в возмущенной среде в ОДС [14].

ным для формирования слоя пространственного заряда за время его пролета через образец [14].

Идеальный обогащенный слой, о котором до сих пор говорилось, представляет собой простейший пример неустойчивости объемного заряда. Если положительные и отрицательные заряды разделены небольшим пространственным промежутком (рис. 4), то мы имеем дело с образованием дипольного слоя (домена). Поле внутри домена больше, чем вне его (рис. 4, б). Из-за ОДС ток в области низкого электрического поля будет больше, чем внутри домена. Значения напряженности электрического поля вне и внутри домена будут стремиться к стационарным, соответствующим лежащим вне участка ОДС точкам зависимости J от \mathcal{E} , в которых электрические токи равны (рис. 3, а). (При этом мы не учитывали толщину доменных стенок.) Таким образом домен достигает стабильной конфигурации. Дипольный слой движется через кристалл и исчезает на аноде. В этот момент электрическое поле начинает возрастать, оставаясь однородным в полупроводнике, до тех пор, пока не превысит пороговое $\mathcal{E} \gg \mathcal{E}_T$ (рис. 5, а), что приведет к образованию нового домена и повторению процесса.

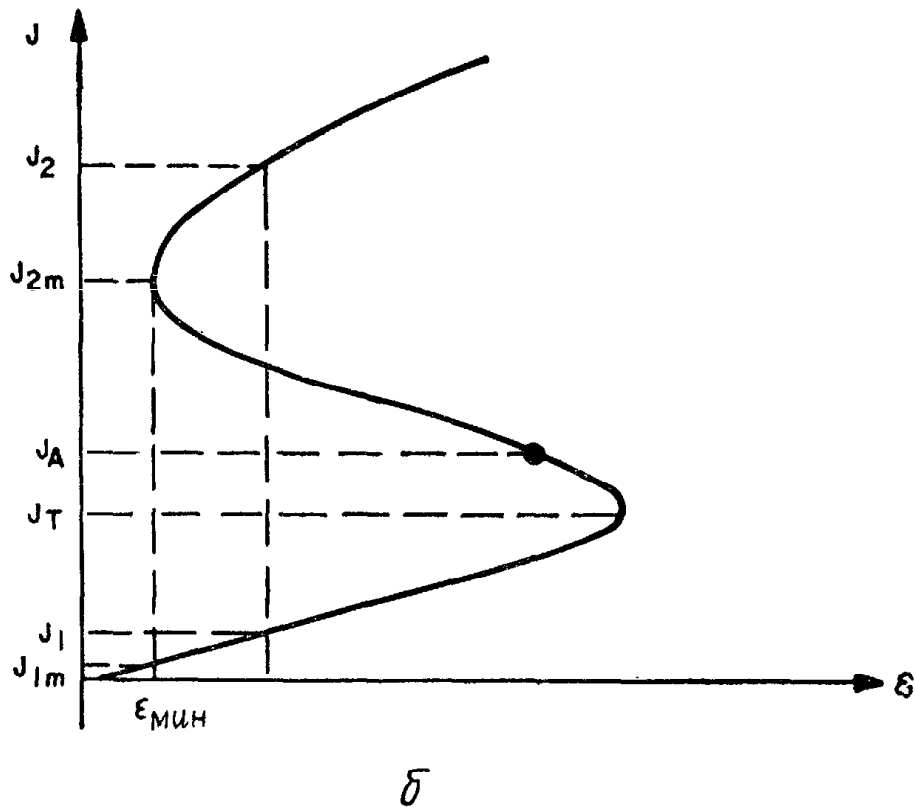
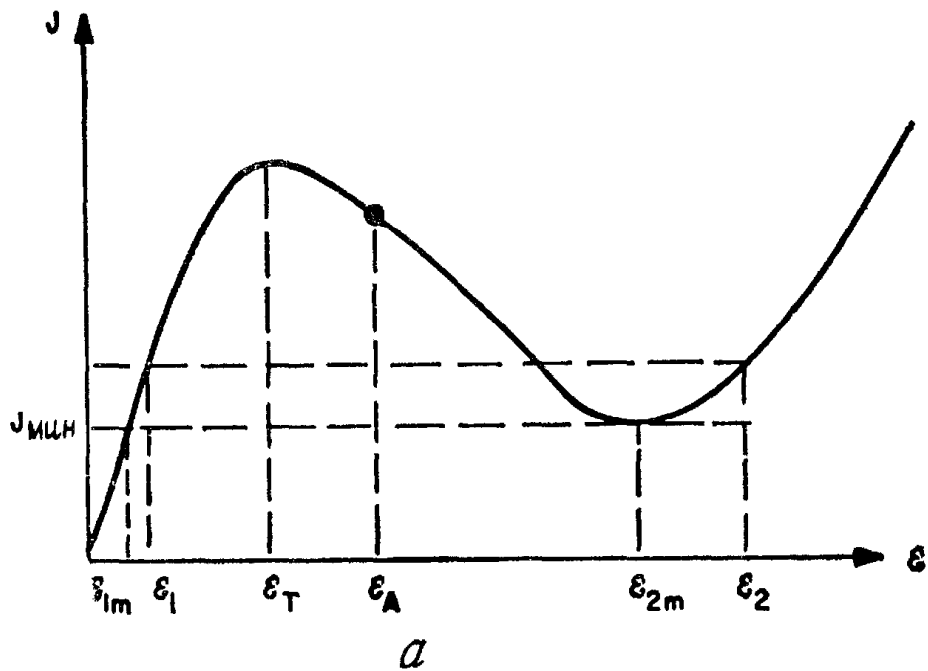


Рис. 5. Минимальная плотность тока и соответствующая напряженность электрического поля для приборов с N-образными $J - \mathcal{E}$ -характеристиками, управляемых напряжением (а), и для приборов с S-образными $J - \mathcal{E}$ -характеристиками, управляемых током (б) [11].

Для того чтобы оценить толщину дипольного слоя d , используем следующие соотношения:

$$V = \mathcal{E}_A L = \mathcal{E}_2 d + (L - d) \mathcal{E}_1 \quad (10)$$

и

$$d = L \left(\frac{\mathcal{E}_A - \mathcal{E}_1}{\mathcal{E}_2 - \mathcal{E}_1} \right), \quad (11)$$

где L — длина образца. Предположим, что наиболее стабильное состояние соответствует минимальной мощности, потребляемой прибором от источника питания. Это означает, что при постоянном напряжении ток должен быть минимальным. Как следует из рис. 5, а, толщина дипольного слоя d определяется выражением (11), если заменить в нем \mathcal{E}_1 и \mathcal{E}_2 на \mathcal{E}_{1m} и \mathcal{E}_{2m} соответственно.

Аналогичное рассмотрение можно провести и для прибора с S-образной J — \mathcal{E} -характеристикой. Вместо домена мы рассмотрим шнур с площадью поперечного сечения a . В стационарных условиях для заданной величины тока I (рис. 5, б) получим

$$I = J_A A = J_2 a + (A - a) J_1 \quad (12)$$

и

$$a = A \left(\frac{J_A - J_1}{J_2 - J_1} \right), \quad (13)$$

где A — площадь поперечного сечения прибора. Если потребляемая мощность минимальна, напряжение также должно быть минимальным, а площадь поперечного сечения шнура можно найти из выражения (13), заменив в нем J_2 и J_1 на J_{2m} и J_{1m} соответственно.

11.2.2. Междолинный переход электронов

Переход электронов из основного минимума зоны проводимости с относительно большой подвижностью в побочные более высокоэнергетические минимумы с меньшими подвижностями называется эффектом междолинного перехода электронов. Для того чтобы понять, как этот эффект приводит к появлению отрицательного дифференциального удельного сопротивления, рассмотрим зависимости энергии электрона от квазиимпульса в арсениде галлия и фосфиде индия (рис. 6) — двух наиболее важных для таких приборов полупроводниках [15, 16]. Видно, что зонные структуры этих полупроводников почти одинаковы. Зона проводимости состоит из нескольких подзон. Дно зоны проводимости соответствует $k = 0$ (точка Γ). Первая более высокоэнергетическая подзона расположена в направлении $\langle 111 \rangle$ (L), а следующая — вдоль оси $\langle 100 \rangle$ (X). Таким образом, подзоны в этих полупроводниках имеют следующую последовательность: $\Gamma - L - X$. Обычно предполагалось, что первая подзона в арсениде галлия отделена при комнатной температуре энергетическим зазором $\sim 0,36$ эВ от основного минимума (X -подзона). Однако в 1976 г. Аспнес [15] измерил электроотражение синхротронного излучения на барьере Шоттки и установил правильную последовательность подзон в арсениде галлия $\Gamma - L - X$ (которая идентична фосфиду индия (рис. 6)).

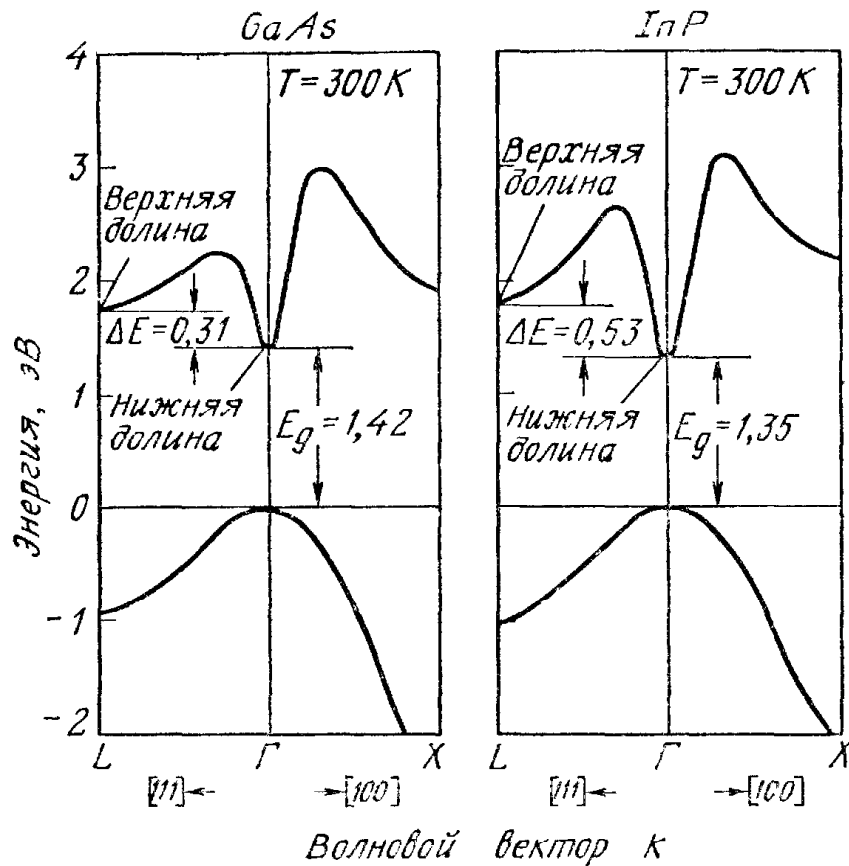


Рис. 6. Структура энергетических зон арсенида галлия и фосфида индия [15, 16].

Теперь мы получим приближенное соотношение между дрейфовой скоростью и напряженностью электрического поля, опираясь на предположение о равенстве электронных температур T_e в нижней (Γ) и верхней (L) долинах [4, 17]. Величина энергетического зазора между минимумами зоны проводимости $\Delta E \approx 0,31$ эВ для арсенида галлия и $\sim 0,53$ эВ для фосфида индия. Введем обозначения: m_1^* и m_2^* — эффективные массы, μ_1 и μ_2 — подвижности, n_1 и n_2 — концентрации электронов в нижней и верхней долинах соответственно, причем полная концентрация носителей заряда равна $n = n_1 + n_2$. Плотность стационарного тока в полупроводнике можно представить следующим образом:

$$J = q (\mu_1 n_1 + \mu_2 n_2) \mathcal{E} = q n v, \quad (14)$$

где v — средняя дрейфовая скорость и

$$v = \left(\frac{\mu_1 n_1 + \mu_2 n_2}{n_1 + n_2} \right) \mathcal{E} \approx \frac{\mu_1 \mathcal{E}}{1 + (n_2/n_1)}, \quad (15)$$

так как $\mu_1 \gg \mu_2$. Отношение заселенностей верхней и нижней долины, разделенных энергетическим зазором ΔE , равно

$$\frac{n_2}{n_1} = R \exp \left(- \frac{\Delta E}{kT_e} \right), \quad (16)$$

где R — отношение плотностей состояний и

$$R = \frac{\text{Число состояний в верхней долине}}{\text{Число состояний в нижней долине}} = \frac{M_2}{M_1} \left(\frac{m_2^*}{m_1^*} \right)^{3/2}, \quad (17)$$

а M_1 и M_2 — число верхних и нижних долин соответственно. Для арсенида галлия $M_1 = 1$, а число верхних долин вдоль оси L равно 8, но они расположены у края первой зоны Бриллюэна, и поэтому $M_2 = 4$. Используя значения эффективных масс электронов в арсениде галлия $m_1^* = 0,067m_0$ и $m_2^* = 0,55m_0$, получим $R = 94$.

Поскольку электрическое поле ускоряет электроны и увеличивает их кинетическую энергию, электронная температура T_e превышает температуру решетки T . Электронная температура определяется с помощью времени релаксации энергии:

$$q\mathcal{E}v = \frac{3}{2} k (T_e - T)/\tau_e, \quad (18)$$

где время релаксации энергии τ_e предполагается равным $\sim 10^{-12}$ с. Подставив v из выражения (15) и n_2/n_1 из выражения (16) в выражение (18), получим

$$T_e = T + \frac{2q\tau_e\mu_1}{3k} \mathcal{E}^2 \left[1 + R \exp \left(-\frac{\Delta E}{kT_e} \right) \right]^{-1}. \quad (19)$$

Используя это равенство, можно рассчитать зависимость T_e от напряженности электрического поля при заданной величине T . Из выражений (15) и (16) следует соотношение между дрейфовой скоростью и полем:

$$v = \mu_1 \mathcal{E} \left[1 + R \exp \left(-\frac{\Delta E}{kT_e} \right) \right]^{-1}. \quad (20)$$

Рассчитанные с помощью выражений (19) и (20) типичные зависимости v от \mathcal{E} для арсенида галлия при трех температурах решетки приведены на рис. 7 (сплошные кривые). Показана также населенность верхней долины как функция напряженности электрического поля (штриховая кривая). Из рассмотренной выше простой модели можно сделать следующие выводы: 1) существует определенное пороговое значение напряженности электрического поля \mathcal{E}_T , при котором возникает участок ОДС (или отрицательной дифференциальной подвижности); 2) пороговое значение напряженности электрического поля увеличивается с ростом температуры решетки; 3) участок отрицательной дифференциальной подвижности может отсутствовать, если температура решетки достаточно высока или энергетический зазор ΔE между минимумами зоны проводимости мал. Поэтому для того, чтобы междолинный переход электронов приводил к возникновению объемного ОДС, необходимо выполнение следующих условий: 1) температура решетки должна быть настолько малой, чтобы в отсут-

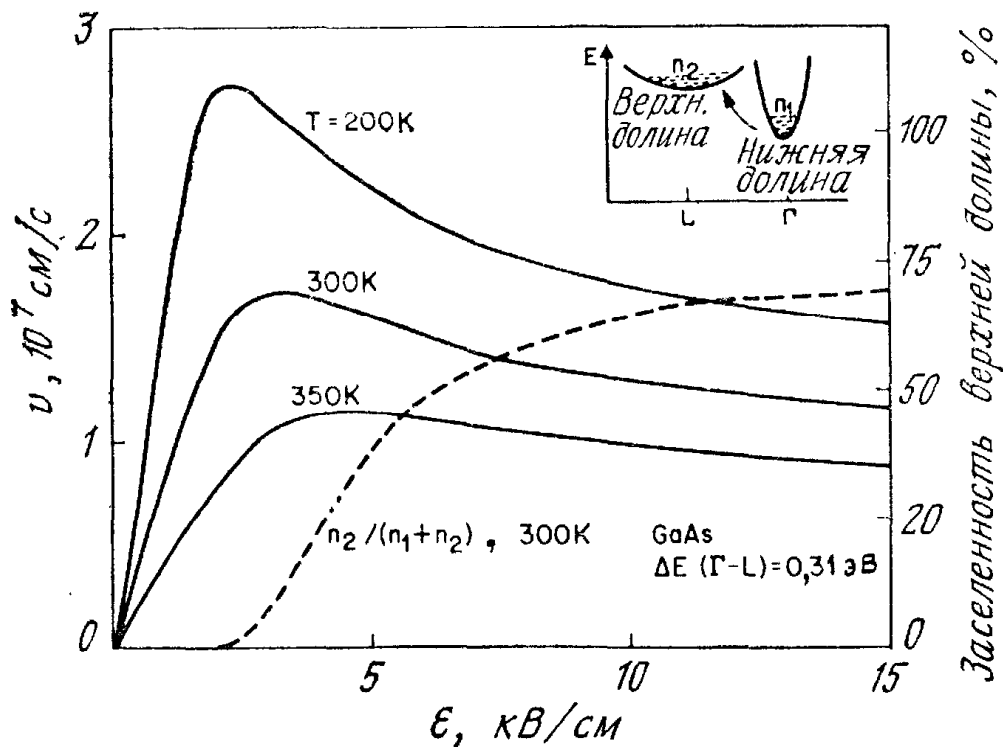


Рис. 7. Зависимости дрейфовой скорости от напряженности электрического поля в GaAs при трех температурах решетки (двухдолинная модель в предположении равенства электронных температур в обоих минимумах).

ствие поля большинство электронов находилось в основном минимуме зоны проводимости (т. е. $kT < \Delta E$); 2) в основном минимуме зоны проводимости электроны должны иметь высокую подвижность, малую эффективную массу и малую плотность состояний, в то время как в побочных минимумах электроны должны иметь низкую подвижность, большую эффективную массу, а плотность состояний должна быть высокой; 3) энергетический зазор между минимумами должен быть меньше ширины запрещенной зоны, чтобы лавинный пробой не произошел раньше междолинных переходов электронов.

Из всех полупроводников, удовлетворяющих этим условиям, арсенид галлия и фосфид индия n -типа наиболее исследованы и широко используются. Тем не менее эффект междолинного перехода электронов наблюдался во многих других полупроводниках, в том числе в германии, бинарных, тройных и четверных соединениях (табл. 1) [9, 18, 19]. Эффект междолинного перехода электронов наблюдался также в InAs и InSb при воздействии гидростатического давления, которое прикладывается для уменьшения энергетического зазора ΔE , при обычных условиях превышающего ширину запрещенной зоны. Особый интерес для возможного использования в маломощных, но быстродействующих приборах представляют тройные соединения элементов III—V групп GaInSb, так как в них пороговые поля малы, а дрейфовые скорости велики. В полупроводниках с большим энергетическим за-

Таблица 1. Полупроводниковые материалы, в которых наблюдается эффект междолинного перехода электронов при 300 К

Полупроводник	E_g , эВ	Энергетический зазор		\mathcal{E}_T , кВ/см	v_p , 10^7 см/с
		между минимумами	ΔE , эВ		
GaAs	1,42	$\Gamma-L$	0,31	3,2	2,2
InP	1,35	—	0,53	10,5	2,5
Ge ^{а)}	0,74	$L-\Gamma$	0,18	2,3	1,4
CdTe	1,50	$\Gamma-L$	0,51	11,0	1,5
InAs ^{б)}	0,36	$\Gamma-L$	1,28	1,6	3,6
InSb ^{в)}	0,28	$\Gamma-L$	0,41	0,6	5,0
ZnSe	2,60	$\Gamma-L$	—	38,0	1,5
Ga _{0,5} In _{0,5} Sb	0,36	$\Gamma-L$	0,36	0,6	2,5
Ga _{0,3} In _{0,7} Sb	0,24	$\Gamma-L$	—	0,6	2,9
InAs _{0,2} P _{0,8}	1,10	$\Gamma-L$	0,95	5,7	2,7
Ga _{0,13} In _{0,87} As _{0,37} P _{0,63}	1,05	—	—	5,5—8,6	1,2

а) При 77 К и ориентации $\langle 100 \rangle$ или $\langle 110 \rangle$.

б) Давление 14 кбар.

в) При 77 К и давлении 8 кбар.

зором между долинами (например, в $Al_{0,25}In_{0,75}As$ $\Delta E = 1,12$ эВ, а в $Ga_{0,6}In_{0,4}As$ $\Delta E = 0,72$ эВ) причиной возникновения ОДС могут быть электроны, находящиеся в основном (Γ) минимуме зоны проводимости [59]. Численные расчеты на ЭВМ по методу Монте-Карло показали, что в таких полупроводниках наличие побочных минимумов не является необходимым условием существования ОДС. Причиной появления максимума в зависимости дрейфовой скорости от поля и возникновения ОДС может служить эффект рассеяния электронов на полярных оптических фононах в непараболическом основном минимуме.

Экспериментальные зависимости дрейфовой скорости от напряженности электрического поля в арсениде галлия и фосфиде индия при комнатной температуре приведены на рис. 8 [16, 20]. Результаты теоретического анализа процессов переноса электронов в сильных электрических полях находятся в хорошем соответствии с экспериментальными данными [21, 22]. Пороговая величина напряженности электрического поля \mathcal{E}_T , определяющая начало участка ОДС, равна $\sim 3,2$ кВ/см для GaAs и 10,5 кВ/см для InP. Максимальная величина дрейфовой скорости v_p составляет $\sim 2,2 \cdot 10^7$ см/с для сверхчистых кристаллов арсенида галлия и $\sim 2,5 \cdot 10^7$ см/с для фосфида индия. Максимальная величина отрицательной дифференциальной подвижности приблизительно

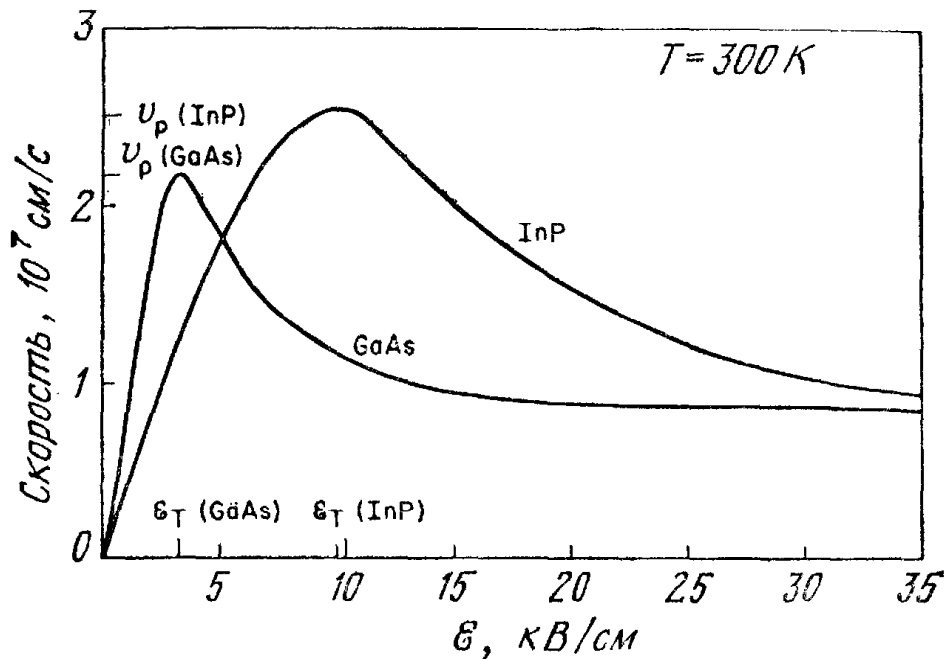


Рис. 8. Экспериментальные зависимости дрейфовой скорости от напряженности электрического поля в GaAs и InP [20, 16].

равна $-2400 \text{ см}^2/\text{В} \cdot \text{с}$ для GaAs и $-2000 \text{ см}^2/\text{В} \cdot \text{с}$ для InP. Экспериментальные зависимости нормированных величин напряженности порогового поля $\mathcal{E}_T/\mathcal{E}_T$ (300K) и максимальной дрейфовой скорости v_p/v_p (300K) от температуры решетки арсенида галлия показаны на рис. 9 [23]. Полученные с помощью простой теоретической модели результаты (рис. 7) качественно согласуются с экспериментальными. Если известна зависимость дрейфовой скорости v от напряженности электрического поля \mathcal{E} , то равенство для плотности тока (3) можно записать в следующем виде:

$$J = qnv(\mathcal{E}) - qD \frac{\partial n}{\partial x}, \quad (21)$$

или

$$J = q(n_1\mu_1 + n_2\mu_2)\mathcal{E} - q\left(D_1 \frac{\partial n_1}{\partial x} + D_2 \frac{\partial n_2}{\partial x}\right), \quad (21a)$$

где первое слагаемое в правой части обусловлено дрейфом, а второе — диффузией носителей заряда, причем предполагается, что коэффициент диффузии не зависит от поля.

В стационарных условиях при малых градиентах концентрации носителей выражение (21) принимает вид $J = qnv(\mathcal{E})$. Подставив этот результат в уравнение Пуассона (уравнение (2)), получим

$$\frac{\partial \mathcal{E}}{\partial x} = -\frac{qn_0}{\epsilon_s} \left[1 - \frac{J/qn_0}{v(\mathcal{E})}\right]. \quad (22)$$

Граничное условие для этого нелинейного дифференциального уравнения первого порядка формулируется на основании утверждения о непрерывности напряженности электрического поля

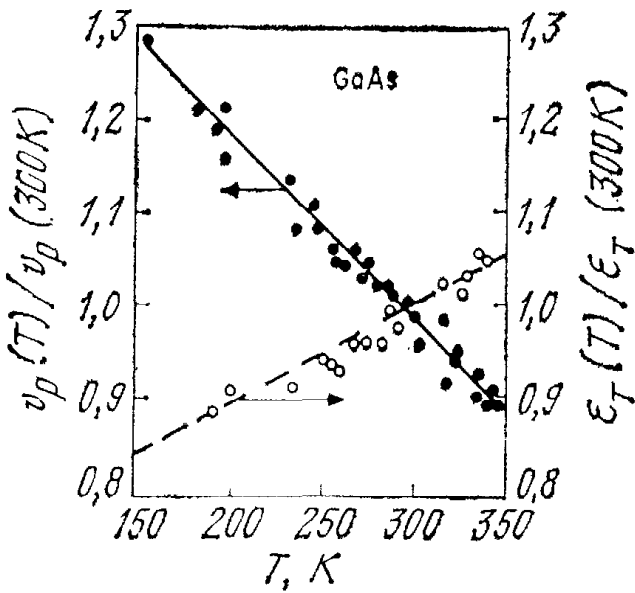


Рис. 9. Экспериментальные зависимости нормированных величин максимальной дрейфовой скорости и пороговой напряженности электрического поля от температуры в GaAs [23].

$\mathcal{E}(x)$ в полупроводнике. Результаты численного решения уравнения (22) для образца из арсенида галлия приведены на рис. 10, где $J_T = qn_0v_p$ [17]. Из рисунка видно, что в любой точке x напряженность электрического поля является монотонно возрастающей функцией плотности тока. Поэтому, если правильно сформулировать граничные условия, стационарное решение не дает отрицательного сопротивления. Это не удивительно, поскольку еще Шокли указал на то, что приборы с объемным ОДС должны

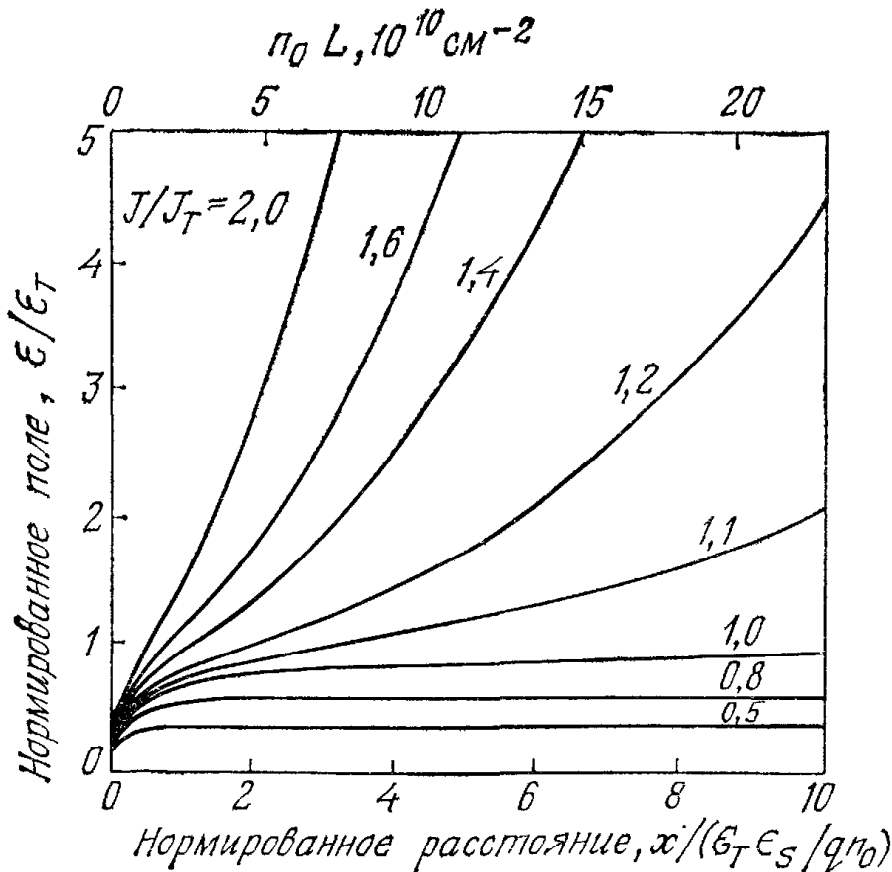


Рис. 10. Зависимости напряженности электрического поля от расстояния x ($\mathcal{E}(0) = 0$ при $x = 0$) [17].

быть стабильными на постоянном токе [24]. Действительно, в полупроводнике имеется внутренний пространственный заряд (избыточные электроны, инжектированные катодом), который возрастает при увеличении напряжения настолько, что ток также увеличивается, несмотря на уменьшение дрейфовой скорости электронов из-за отрицательной дифференциальной подвижности. Однако эти стационарные решения не обязательно являются устойчивыми по отношению к малым флюктуациям.

11.3. РЕЖИМЫ РАБОТЫ

С тех пор как в 1963 г. Гани впервые наблюдал СВЧ-колебания в приборах на эффекте междолинного перехода электронов из арсенида галлия и фосфида индия, были изучены различные режимы работы этих приборов. Пять основных факторов определяют тот или иной режим ¹⁾: 1) величина концентрации примеси и однородность ее распределения в образце; 2) ширина активной области; 3) свойства катодного контакта; 4) тип используемой внешней цепи; 5) величина напряжения смещения на приборе.

Как уже говорилось выше, в нестационарном состоянии в приборах на эффекте междолинного перехода электронов формируются обогащенные слои, или домены. Процесс образования домена высокого поля зависит от числа носителей заряда в полупроводнике и длины прибора, которые должны быть достаточно большими, чтобы за пролетное время происходило формирование пространственного заряда требуемой величины. Исходя из этого, устанавливается критерий для того или иного режима работы прибора. Как следует из выражения (8), увеличение со временем пространственного заряда в приборе с ОДС описывается на начальной стадии следующей зависимостью:

$$(n - n_0) = (n - n_0)_{t=0} \exp(t/|\tau_R|),$$

где $|\tau_R| = \epsilon/qn_0|\mu_-|$ и μ_- — отрицательная дифференциальная подвижность. Если это соотношение остается справедливым на протяжении времени пролета слоя пространственного заряда через образец, то максимальный фактор роста будет равен $\exp(L/v|\tau_R|)$, где v — средняя скорость дрейфа слоя. Для существенного увеличения заряда этот фактор должен превышать 1, и, следовательно, $L/v|\tau_R| > 1$ или

$$n_0L > \epsilon_s v/q|\mu_-|. \quad (23)$$

¹⁾ В отдельных работах возникает ряд вопросов в связи с определением граничных условий на катоде. Поэтому представляется необходимым дальнейшее изучение этого вопроса с целью установления справедливости некоторых утверждений, которые включены в эту главу.

Для арсенида галлия и фосфида индия n -типа правая часть неравенства составляет $\sim 10^{12}$ см $^{-2}$. Если произведение $n_0L < 10^{12}$ см $^{-2}$, то распределение поля в таком приборе устойчиво (рис. 10). Поэтому важной границей, разделяющей режимы работы, является величина произведения концентрации носителей и длины прибора, равная $n_0L = 10^{12}$ см $^{-2}$. Ниже мы рассмотрим некоторые важные режимы работы, причем начнем с идеального режима однородного электрического поля.

11.3.1. Идеальный режим однородного электрического поля

В идеализированных условиях, когда в исходном образце отсутствует пространственный заряд и электрическое поле однородно, вольт-амперная характеристика прибора может быть получена путем пересчета зависимости скорости от поля. Простейшей формой импульсов напряжения является прямоугольная (рис. 11) [25]. Введем два безразмерных параметра $\alpha \equiv I_V/I_T$ и $\beta \equiv V_0/V_T$. Из предполагаемой формы импульсов следует, что средняя величина тока I_0 равна

$$I_0 = (1 + \alpha) I_T/2. \quad (24)$$

Мощность, потребляемая прибором от источника питания,

$$P_0 = V_0 I_0 = \frac{1}{2} \beta (1 + \alpha) V_T I_T. \quad (25)$$

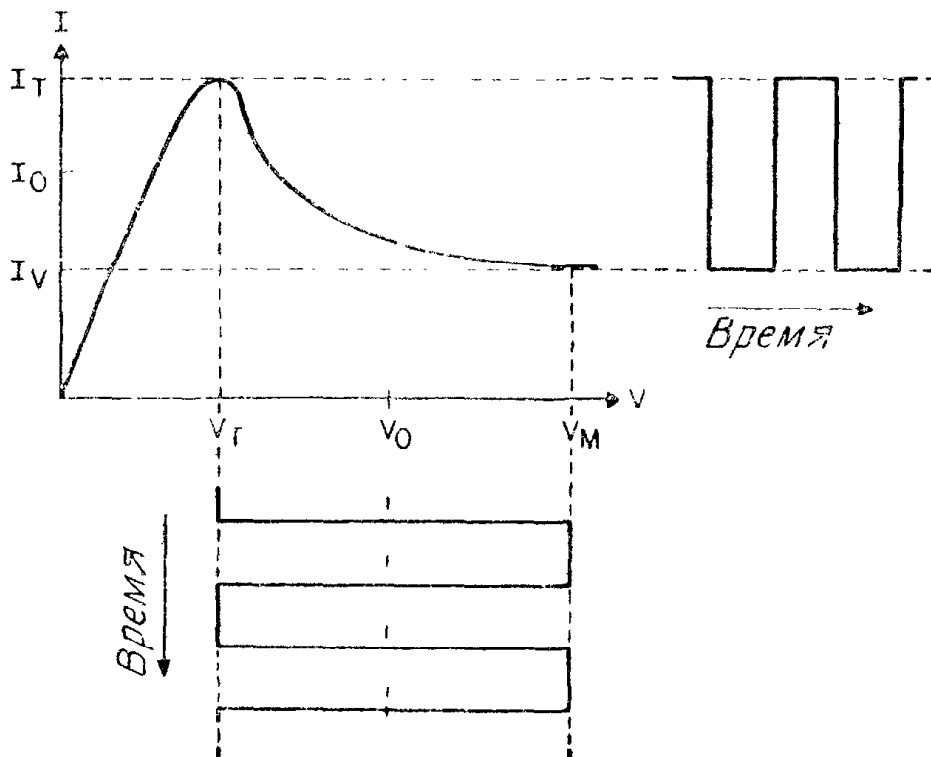


Рис. 11. Идеализированные прямоугольные формы импульсов тока и напряжения в режиме однородного электрического поля [25].

Полная величина СВЧ-мощности, выделяемой на нагрузке, равна

$$P_{rf} = \frac{V_M - V_T}{2} \frac{I_T - I_V}{2} = \frac{(\beta - 1)(1 - \alpha)}{2} V_1 I_T. \quad (26)$$

Поэтому к. п. д. прибора

$$\eta = \frac{(1 - \alpha)(\beta - 1)}{(1 + \alpha)\beta}, \quad (27)$$

а для основной частоты

$$\eta_1 = \frac{8}{\pi^2} \eta. \quad (27a)$$

Из выражения (27) следует, что наибольший к. п. д. получается при максимально возможном напряжении постоянного смещения ($\beta \rightarrow \infty$) и отношении $1/\alpha$. Теоретически достижимая величина к. п. д. прибора из арсенида галлия равна 30 % ($1/\alpha = 2,2$) и 45 % для прибора из фосфида индия ($1/\alpha = 3,5$). Эти значения не должны зависеть от частоты до тех пор, пока последняя меньше величин, обратной времени релаксации энергии, и величин, обратной времени междолинного рассеяния.

В экспериментальных условиях такие высокие значения к. п. д. не были получены, а рабочая частота обычно равнялась пролетной. Это объясняется следующими причинами: 1) величина постоянного смещения ограничена напряжением лавинного пробоя; 2) формирование слоя пространственного заряда в обычных условиях приводит к неоднородному распределению поля в образце; 3) в резонансной цепи трудно получить идеальную форму импульсов тока и напряжения.

11.3.2. Режим с обогащенным слоем

Распределение электрического поля в коротком или слаболегированном образце ($n_0 L < 10^{12} \text{ см}^{-2}$) устойчиво, а его сопротивление по постоянному току положительно. На рис. 10 приведены стационарные распределения напряженности электрического поля в таком приборе при различных плотностях тока. Если к образцу внезапно прикладывается однородное электрическое поле, то динамику обогащенного слоя можно понять из упрощенных зависимостей, приведенных на рис. 12 [26]. В момент времени t обогащенный слой (т. е. избыточные электроны) инжектируется из катода, что приводит к «расщеплению» распределения электрического поля на два участка (момент времени 2). При этом скорости носителей вне обогащенного слоя изменяются, как показано на рис. 12, а. Поскольку напряжение на приборе предполагается постоянным, площади под кривыми на рис. 12, б должны быть равны. Из зависимости дрейфовой скорости от напряженности электрического поля следует, что по мере продвижения обога-

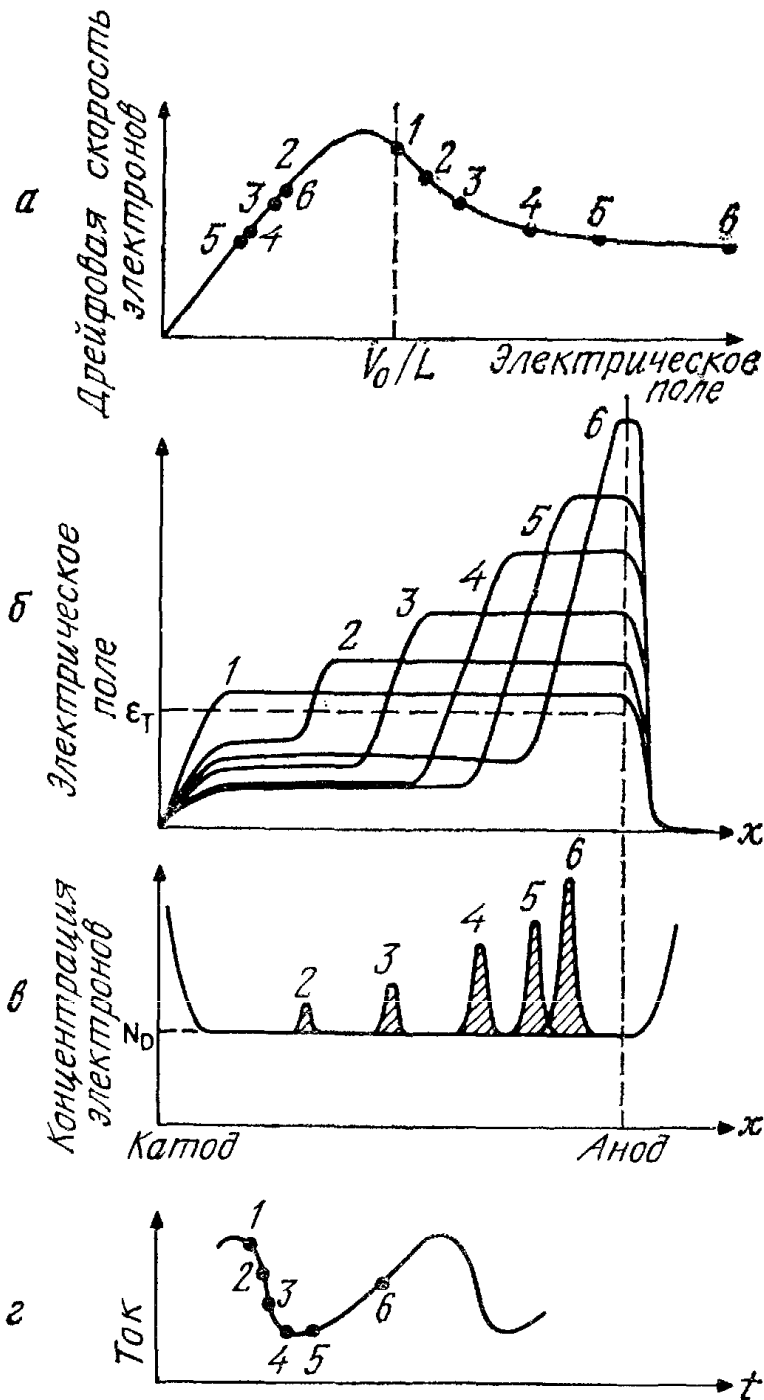


Рис. 12. Характеристики прибора в режиме работы с пролетом обогащенного слоя при постоянном напряжении на приборе [26].

щенного слоя к аноду это равенство может иметь место только при уменьшении скорости носителей вне слоя (моменты времени 3—5). В конце концов в момент времени 6 обогащенный слой достигает анода и исчезает. Поле вблизи катода возрастает до порогового, инжектируется другой обогащенный слой, и процесс повторяется. Как видно из рис. 12, г, зависимость тока от времени достаточно гладкая.

Если n_0L меньше критического значения ($n_0L < 10^{12}$ см⁻²), то сопротивление прибора может быть отрицательным на частотах, близких к пролетной частоте или ее гармоникам. Такой прибор может работать как усилитель [27]. На рис. 13 приведены экспериментальные частотные зависимости действительной и мнимой составляющих проводимости образца из арсенида галлия *n*-типа

с $n_0 = 3 \cdot 10^{13} \text{ см}^{-3}$ и $L = 70 \text{ мкм}$ ($n_0 L = 2,1 \cdot 10^{11} \text{ см}^{-2}$) [28]. Максимумы отрицательной действительной составляющей проводимости приблизительно соответствуют гармоникам пролетной частоты. Экспериментально полученная зависимость коэффициента усиления от частоты приведена на рис. 14 для прибора с такой же длиной, но меньшей концентрацией ($n_0 = 1,5 \cdot 10^{13} \text{ см}^{-3}$). При увеличении напряженности электрического поля пик коэффициента усиления сдвигается в область более высоких частот, поскольку растет максимум отрицательной действительной составляющей проводимости и сдвигается в ту же область.

Прибор будет генерировать сигнал в режиме с пролетом обогащенного слоя, если величина $n_0 L$ меньше критического значения, а прибор включен в параллельную резонансную цепь с сопротивлением нагрузки $\sim 10R_0$, где R_0 — сопротивление прибора при малых полях. На рис. 15 приведены распределения напряженности электрического поля в образце в четыре различных момента одного СВЧ-периода [29]. Показан также профиль легирования ($n_0 = 2 \cdot 10^{14} \text{ см}^{-3}$) и зависимости напряжения и тока от времени. Напряжение всегда превышает пороговое ($V > V_T = \mathcal{E}_T L$). Форма генерируемых колебаний далека от идеальной, а к. п. д. составляет лишь 5 %. При последовательном соединении прибора с нагрузочным сопротивлением и катушкой индуктивности можно получить более приемлемую форму колебаний и к. п. д. $\sim 10 \%$.

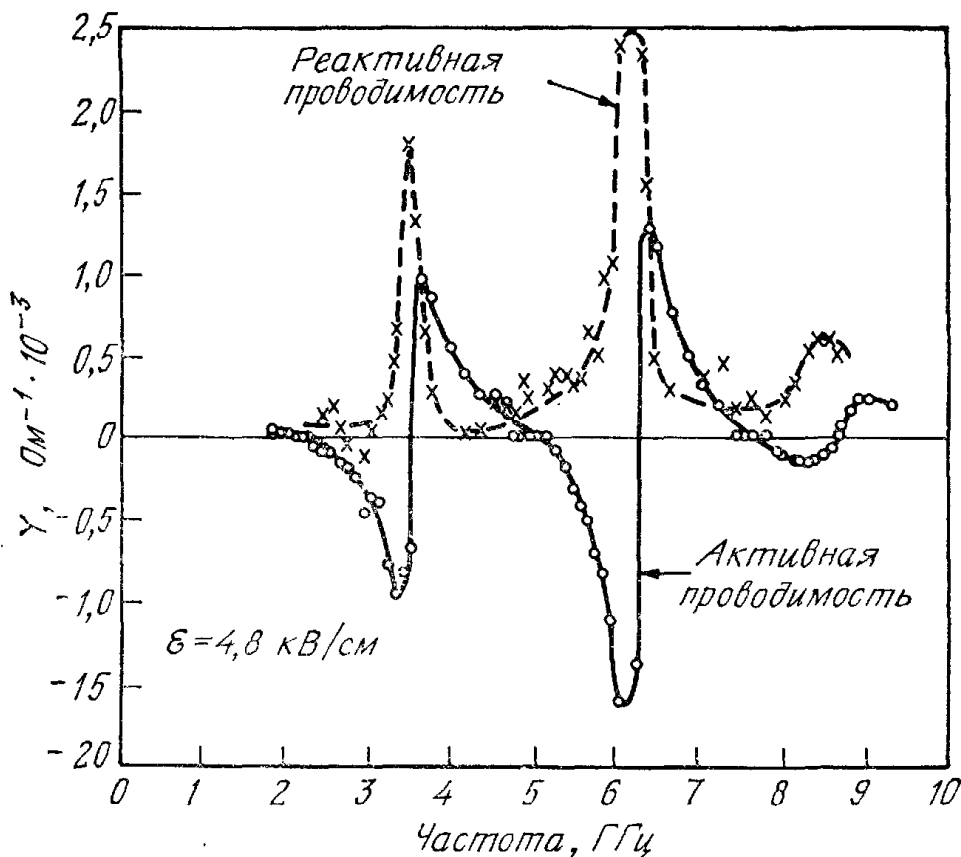


Рис. 13. Экспериментальная зависимость проводимости образца из арсенида галлия ($n_0 = 3 \cdot 10^{13} \text{ см}^{-3}$ и $L = 70 \text{ мкм}$) от частоты [28].

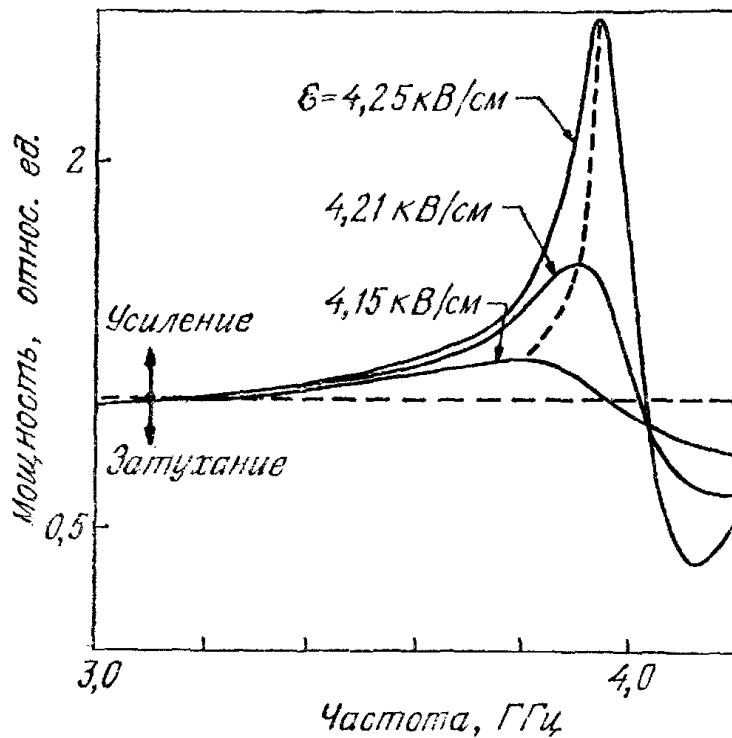


Рис. 14. Экспериментально полученный коэффициент усиления как функция частоты при различных напряжениях смещения на образце из арсенида галлия с $n_0 = 1,5 \cdot 10^{13} \text{ см}^{-3}$ и $L = 70 \text{ мкм}$ [28].

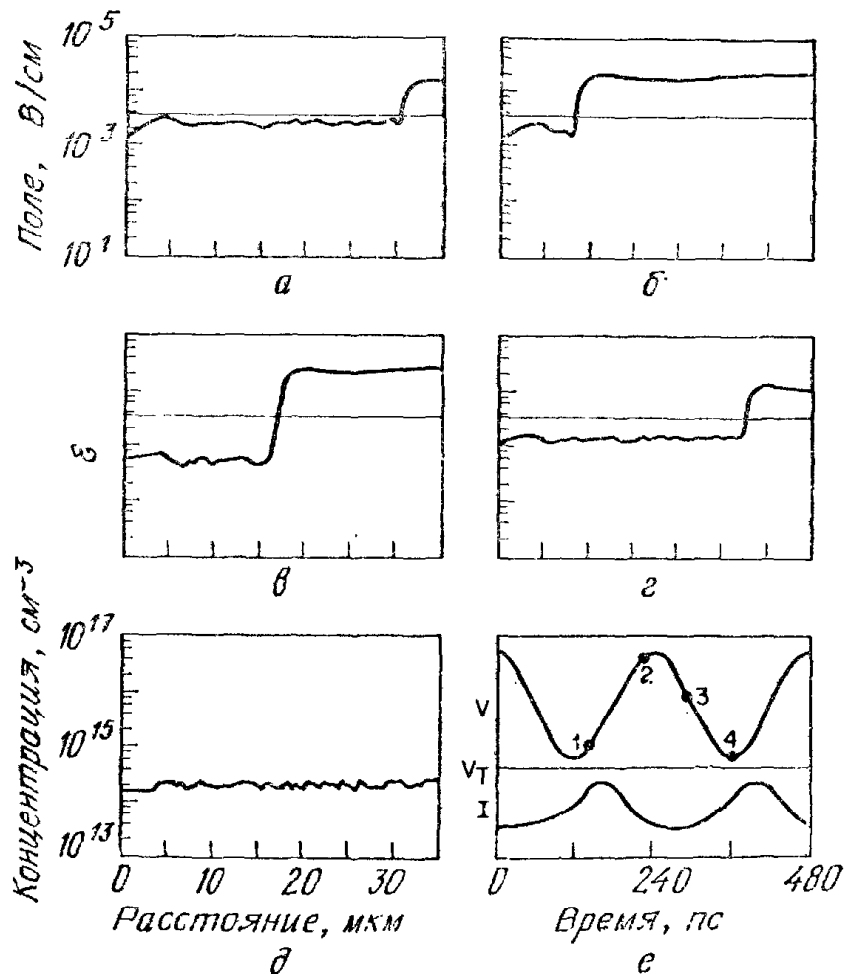


Рис. 15. Распределения электрического поля в четыре различных момента одного СВЧ-периода (а — г), профиль легирования примесью (д) и зависимости тока и напряжения от времени (е) прибора из GaAs при работе в резонансной цепи в режиме с обогащенным слоем для $fL = 1,4 \cdot 10^7 \text{ см/с}$ и $n_0/f = 5 \cdot 10^4 \text{ с/см}^3$ [29].

11.3.3. Режим пролета домена

Если произведение n_0L превышает 10^{12} см⁻², то флуктуации пространственного заряда в полупроводнике возрастают экспоненциально со временем и расстоянием и формируются домены, которые движутся к аноду. Обычно домен образуется вблизи катодного контакта, поскольку флуктуации концентрации примеси и объемного заряда здесь максимальны. Причиной экспериментально наблюдаемой ганновской генерации служат периодическое образование и последующее исчезновение на аноде доменов, которые успевают полностью сформироваться. В разд. 11.2 мы качественно рассмотрели неустойчивости пространственного заряда и формирование доменов. В этом разделе мы изучим диполи более подробно. Дипольные слои являются стабильными в том смысле, что движутся с определенной скоростью и не изменяются со временем. Мы будем считать, что скорость дрейфа электронов подчиняется статической зависимости, приведенной на рис. 16 (сплошная кривая). Система уравнений, описывающих поведение электронов, состоит из уравнения Пуассона (уравнение (2)) и уравнения для плотности полного тока

$$J = qnv(\mathcal{E}) - q \frac{\partial D(\mathcal{E})n}{\partial x} + \epsilon_s \frac{\partial \mathcal{E}}{\partial t}. \quad (28)$$

За исключением третьего слагаемого, которое равно току смещения, это уравнение идентично выражению (21).

Решение должно иметь вид домена высокого электрического поля, который без изменения формы движется со скоростью $v_{\text{дом}}$. Вне домена концентрация носителей и электрическое поле постоянны и равны $n = n_0$ и $\mathcal{E} = \mathcal{E}_r$ соответственно. Для решения такого вида напряженность электрического поля \mathcal{E} и концентра-

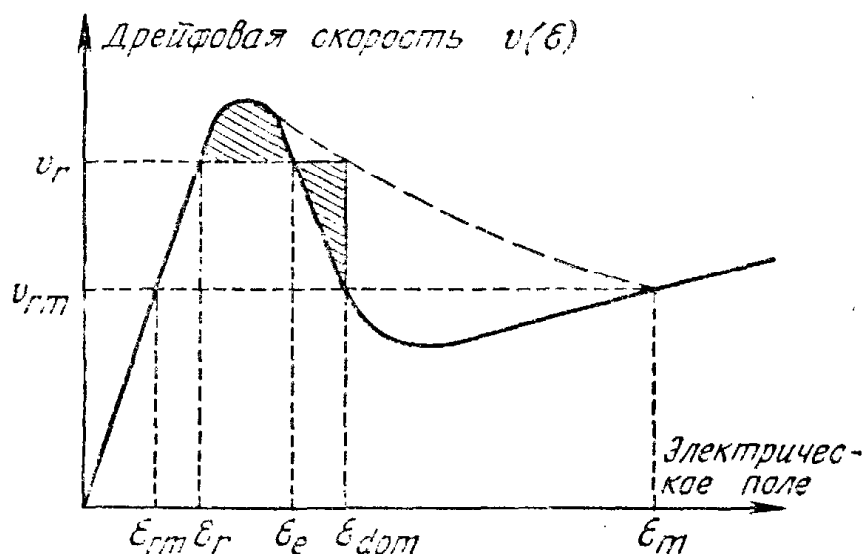


Рис. 16. Зависимости скорости от напряженности электрического поля (сплошная кривая) и максимальной напряженности электрического поля в домене от дрейфовой скорости вне его (штриховая кривая) [30].

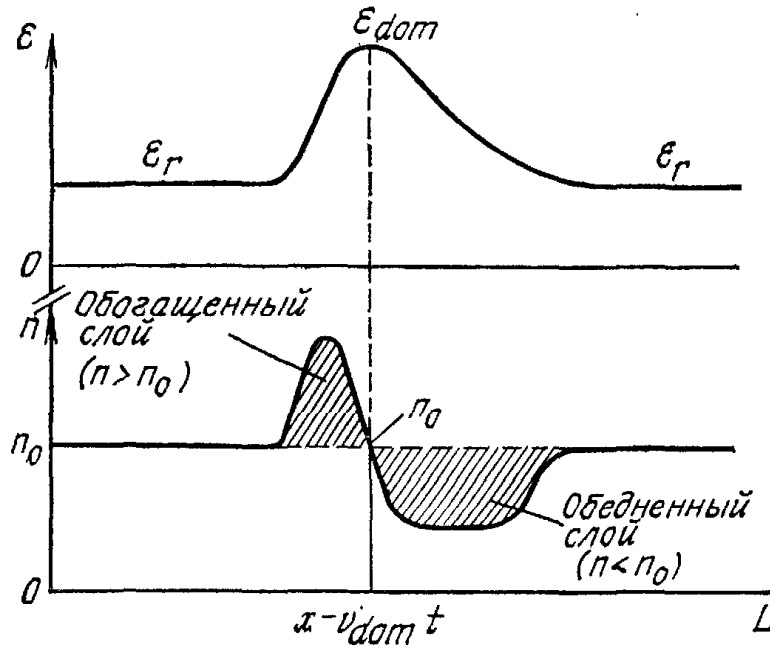


Рис. 17. Распределения электрического поля и концентрация электронов для стабильного домена высокого поля, движущегося со скоростью v_r [31].

ция n зависят от одной переменной $x' = x - v_{\text{dom}}t$, а их распределения приведены на рис. 17. Отметим, что концентрация является двузначной функцией электрического поля. Домен состоит из обогащенного носителями слоя, в котором $n > n_0$, и обедненного слоя ($n < n_0$). Концентрация носителей n равна n_0 при двух значениях напряженности электрического поля: при $\mathcal{E} = \mathcal{E}_r$ вне домена и при $\mathcal{E} = \mathcal{E}_{\text{dom}}$ в максимуме поля в домене.

Предположим, что величина поля вне домена \mathcal{E}_r известна. (Ниже мы покажем, что величину \mathcal{E}_r можно легко найти.) Ток вне домена состоит только из тока проводимости и определяется выражением $J = qn_0v_r$, в котором $v_r = v(\mathcal{E}_r)$. Так как

$$\frac{\partial \mathcal{E}}{\partial x} = \frac{\partial \mathcal{E}}{\partial x'} \quad \text{и} \quad \frac{\partial \mathcal{E}}{\partial t} = -v_{\text{dom}} \frac{\partial \mathcal{E}}{\partial x'},$$

выражения (2) и (28) можно записать следующим образом:

$$\frac{\partial \mathcal{E}}{\partial x'} = \frac{q}{\epsilon_s} (n - n_0) \quad (29)$$

и

$$\frac{d}{dx'} [D(\mathcal{E})n] = n[v(\mathcal{E}) - v_{\text{dom}}] - n_0(v_r - v_{\text{dom}}). \quad (30)$$

Разделив выражение (30) на выражение (29), мы можем исключить переменную x' и получить дифференциальное уравнение для $[D(\mathcal{E})n]$ как функцию электрического поля:

$$\frac{q}{\epsilon_s} \frac{d}{d\mathcal{E}} [D(\mathcal{E})n] = \{n[v(\mathcal{E}) - v_{\text{dom}}] - n_0(v_r - v_{\text{dom}})\} / (n - n_0). \quad (31)$$

В общем виде уравнение (31) может быть решено только численными методами [30—32]. Однако задача существенно упрощается,

если предположить, что коэффициент диффузии не зависит от электрического поля, т. е. $D(\mathcal{E}) = D$. С учетом этого приближения решение уравнения (31) будет иметь вид

$$\frac{n}{n_0} - \ln\left(\frac{n}{n_0}\right) - 1 = \frac{\varepsilon_s}{qn_0D} \int_{\mathcal{E}_r}^{\mathcal{E}} \left\{ [v(\mathcal{E}') - v_{\text{dom}}] - \frac{n_0}{n} (v_r - v_{\text{dom}}) \right\} d\mathcal{E}', \quad (32)$$

в чем можно убедиться, продифференцировав его по \mathcal{E} .

Отметим, что если $\mathcal{E} = \mathcal{E}_r$ или $\mathcal{E} = \mathcal{E}_{\text{dom}}$, то концентрация $n = n_0$ (рис. 17) и левая часть выражения (32) равна 0. Поэтому интеграл в правой части выражения (32) должен также равняться 0 при $\mathcal{E} = \mathcal{E}_{\text{dom}}$. Однако проинтегрировать от \mathcal{E}_r до \mathcal{E}_{dom} можно либо по обедненному слою, в котором $n < n_0$, либо по обогащенному слою ($n > n_0$). Поскольку первое слагаемое под интегралом в выражении (32) не зависит от n , а результаты интегрирования второго слагаемого различаются в двух указанных выше случаях, для равенства интеграла нулю при интегрировании как по обедненному слою, так и по обогащенному необходимо, чтобы $v_r = v_{\text{dom}}$. Тогда при $\mathcal{E} = \mathcal{E}_{\text{dom}}$ уравнение (32) принимает вид

$$\int_{\mathcal{E}_r}^{\mathcal{E}_{\text{dom}}} [v(\mathcal{E}') - v_r] d\mathcal{E}' = 0. \quad (33)$$

Это условие выполняется, если равны площади заштрихованных областей на рис. 16. Воспользовавшись «правилом равных площадей» [30], можно определить максимальную величину поля в домене \mathcal{E}_{dom} , если известна величина \mathcal{E}_r . Штриховая кривая на рис. 16 представляет собой зависимость \mathcal{E}_{dom} от v_r , найденную с помощью этого правила. Кривая начинается в максимуме зависимости скорости от поля, в котором напряженность равна пороговой, и заканчивается в точке (\mathcal{E}_m, v_{rm}) . Если напряжение смещения прибора такое, что скорость $v(\mathcal{E}_r)$ меньше v_{rm} , то правило равных площадей не может быть выполнено и стабильный домен не существует [31].

Для решения уравнения (31) с учетом полевой зависимости коэффициента диффузии применяются численные методы. Из полученных решений следует, что при заданной величине напряженности электрического поля вне домена \mathcal{E}_r существует не более одного значения превышения скорости доменом $(v_{\text{dom}} - v_r)$, при котором это решение существует. Другими словами, для любой \mathcal{E}_r существует единственная устойчивая доменная конфигурация. На рис. 18 приведены зависимости превышения скорости доменом от дрейфовой скорости носителей вне его v_r при двух значениях концентрации n_0 [32]. Отметим, что для $n_0 = 10^{15} \text{ см}^{-3}$ выпол-

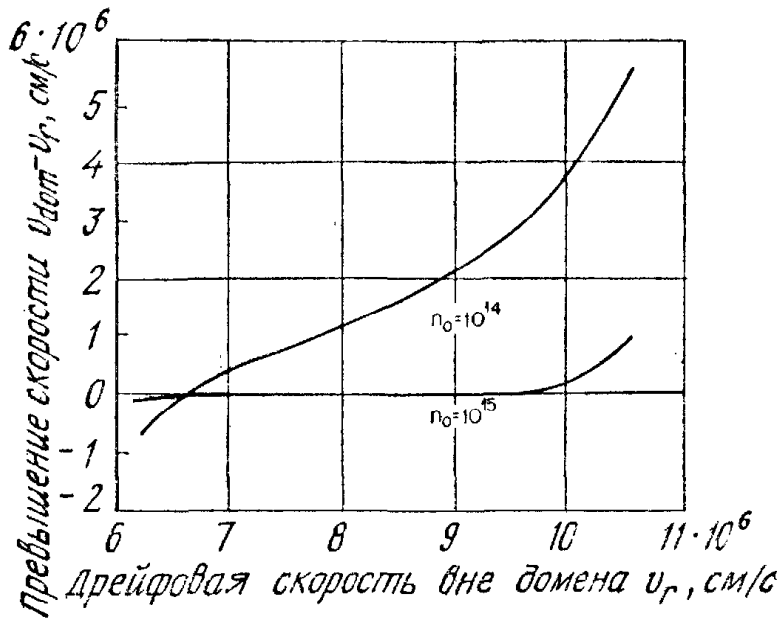


Рис. 18. Зависимость преышения скорости доменом высокого поля от дрейфовой скорости вне его [32].

няется равенство $v_{\text{dom}} = v_r$ в широком диапазоне v_r , в котором применимо правило равных площадей.

Теперь рассмотрим некоторые характеристики домена высокого поля. Если диполь не находится в контакте с каким-нибудь электродом, ток прибора зависит от напряженности электрического поля вне домена \mathcal{E}_r :

$$J = qn_0v(\mathcal{E}_r). \quad (34)$$

Поэтому при заданной концентрации n_0 поле вне домена определяет плотность тока J . Избыточное напряжение на диполе можно найти следующим образом:

$$V_{ex} = \int_{-\infty}^{\infty} [\mathcal{E}(x) - \mathcal{E}_r] dx. \quad (35)$$

Результаты вычисления интеграла (35) на ЭВМ для различных значений концентрации носителей и напряженности поля вне домена приведены на рис. 19 [32]. Эти кривые можно использовать для определения поля вне домена \mathcal{E}_r в образце длиной L с концентрацией примеси n_0 при напряжении смещения V , если учесть, что одновременно с уравнением (35) должно выполняться равенство

$$V_{ex} = V - L\mathcal{E}_r. \quad (36)$$

Прямая линия, определяемая этим уравнением, называется приборной прямой. Для частного случая $L = 25$ мкм и $V = 10$ В она нанесена штрихами на рис. 19. Если $V/L > \mathcal{E}_T$, то пересечение приборной прямой с зависимостью (35) однозначно определяет \mathcal{E}_r и, следовательно, ток прибора. Наклон приборной прямой задается величиной L , а точка пересечения, соответствующая \mathcal{E}_r , может смещаться при изменении постоянного напряжения V .

Когда домен достигает анода, ток во внешней цепи возрастает, а электрическое поле перераспределяется в полупроводнике так, что образуется новый домен. Частота колебаний тока зависит от скорости движения домена в образце v_{dom} . Если возрастает v_{dom} , то увеличивается частота, и наоборот. Зависимость v_{dom} от напряжения смещения может быть легко найдена.

На рис. 20 приведена зависимость ширины домена от избыточного напряжения [32]. Отметим, что при заданной величине V_{ex} ширина диполя уменьшается с увеличением концентрации. При бесконечно малом коэффициенте диффузии домен будет иметь треугольную форму, так как при $\mathcal{E}_r < \mathcal{E} < \mathcal{E}_{dom}$ правая часть выражения (32) стремится к бесконечности, если $D \rightarrow 0$. Поэтому левая часть выражения (32) также должна неограниченно возрастать. Это означает, что $n \rightarrow 0$ в обедненном слое и $n \rightarrow \infty$ в обогащенном слое. Напряженность электрического поля линейно уменьшается от \mathcal{E}_{dom} до \mathcal{E}_r с расстоянием, а ширина домена равна

$$d = \frac{\epsilon_s}{qn_0} (\mathcal{E}_{dom} - \mathcal{E}_r). \quad (37)$$

Избыточное напряжение на домене

$$V_{ex} = (\mathcal{E}_{dom} - \mathcal{E}_r) d/2 = \frac{\epsilon_s (\mathcal{E}_{dom} - \mathcal{E}_r)^2}{2qn_0}. \quad (38)$$

На практике в приборах из арсенида галлия и фосфида индия генерировались только домены треугольной формы.

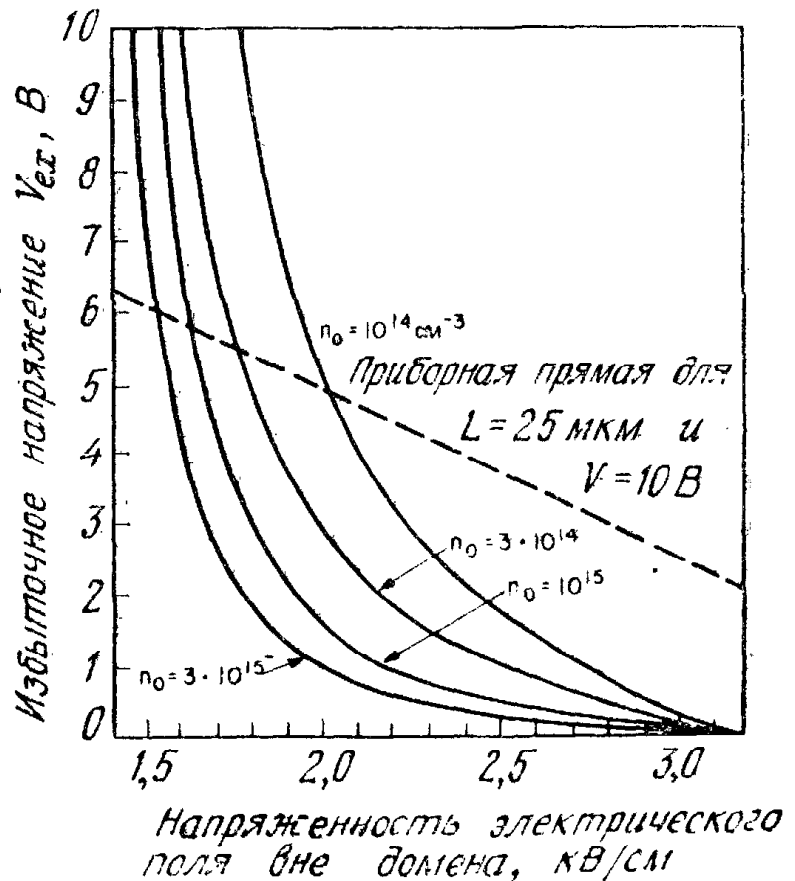


Рис. 19. Зависимость избыточного напряжения на домене от напряженности электрического поля вне его при различных концентрациях примеси [32].

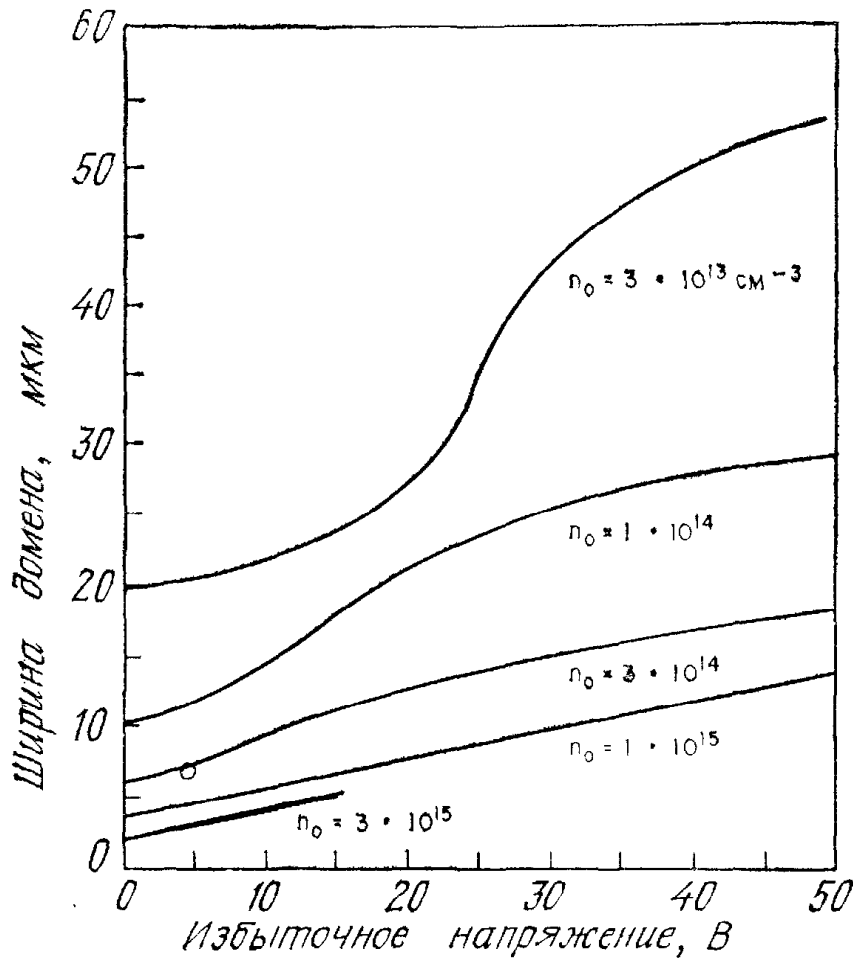
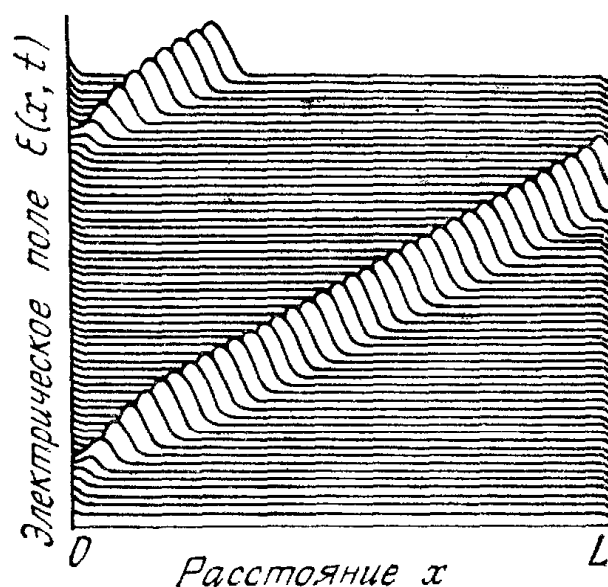


Рис. 20. Зависимость ширины домена от избыточного напряжения при различных концентрациях примеси [32].

Режим пролета домена может быть реализован, если прибор с величиной произведения n_0L , превышающей критическое значение, включить в параллельную резонансную цепь (например, в СВЧ-резонатор с высокой добротностью). В этом режиме домен высокого поля формируется у катода и движется через образец к аноду. Когда домен достигает анода, напряженность электрического поля в образце увеличивается до тех пор, пока не превысит пороговую, и новый диполь образуется у катода. На рис. 21 показаны результаты численного моделирования процесса распространения домена в приборах из арсенида галлия длиной 100 мкм при концентрации примеси $5 \cdot 10^{14}$ см $^{-3}$ ($n_0L = 5 \cdot 10^{12}$ см $^{-2}$) [33]. Две пространственные зависимости $\mathcal{E}(x, t)$, изображенные соседними кривыми, разделены промежутком $16\tau_R$, где τ_R — время диэлектрической релаксации (выражение (9)), равное для данного прибора 1,5 пс. В момент исчезновения домена на аноде происходит увеличение тока во внешней цепи; поэтому зависимость тока от времени в образце, длина которого значительно превышает ширину домена, будет иметь резкие «всплески», а не плавную синусоидальную форму. На рис. 22 показана экспериментальная зависимость тока от времени в образце длиной 100

Рис. 21. Численное моделирование процесса распространения возникающего у катода дипольного слоя в режиме пролета домена. Распределения поля, изображенные соседними кривыми, соответствуют моментам времени, разделенным интервалом 24 пс [33].



мкм и с $n_0L = 3 \cdot 10^{13} \text{ см}^{-2}$ [34]. Очевидно, что для получения формы колебаний, близкой к синусоидальной, следует уменьшать длину образца (что приведет к увеличению частоты) или увеличивать ширину домена. Из рис. 20 видно, что ширина домена растет при уменьшении концентрации примеси n_0 . В общем случае, уменьшая величину произведения n_0L , можно получить форму колебаний, более близкую к синусоидальной. На рис. 23 приведены распределения напряженности электрического поля в образце длиной 35 мкм в четыре различных момента в течение одного СВЧ-периода, а также профиль легирования и зависимости напряжения и тока от времени [29]. Для этого прибора форма тока больше похожа на синусоидальную. Теоретический анализ показывает, что к. п. д. пролетного режима максимален, когда произведение $n_0L \approx 10^{12} \text{ см}^{-2}$, а домен занимает приблизительно половину длины образца. В этом случае форма колебаний тока близка к синусоидальной. Максимально достижимая величина к. п. д. для этого режима составляет 10 %. Если форма импульсов тока близка к прямоугольной, то к. п. д. будет выше. Такую фор-

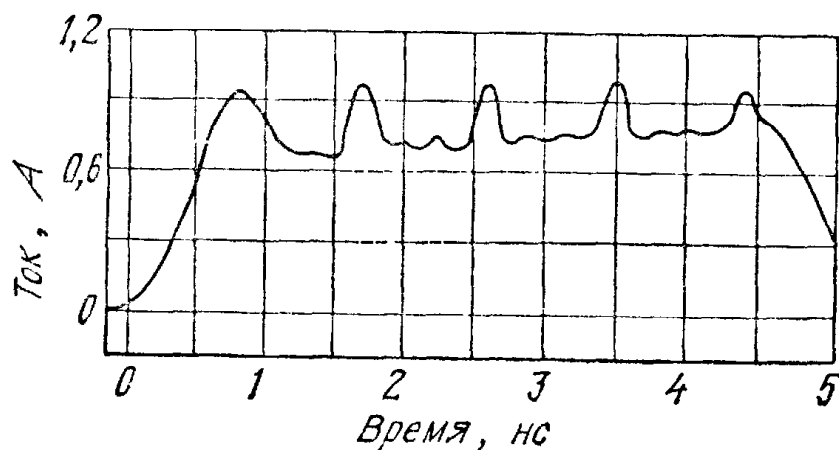


Рис. 22. Экспериментальная зависимость тока от времени в образце из арсенида галлия длиной 100 мкм при концентрации примеси $3 \cdot 10^{15} \text{ см}^{-3}$.

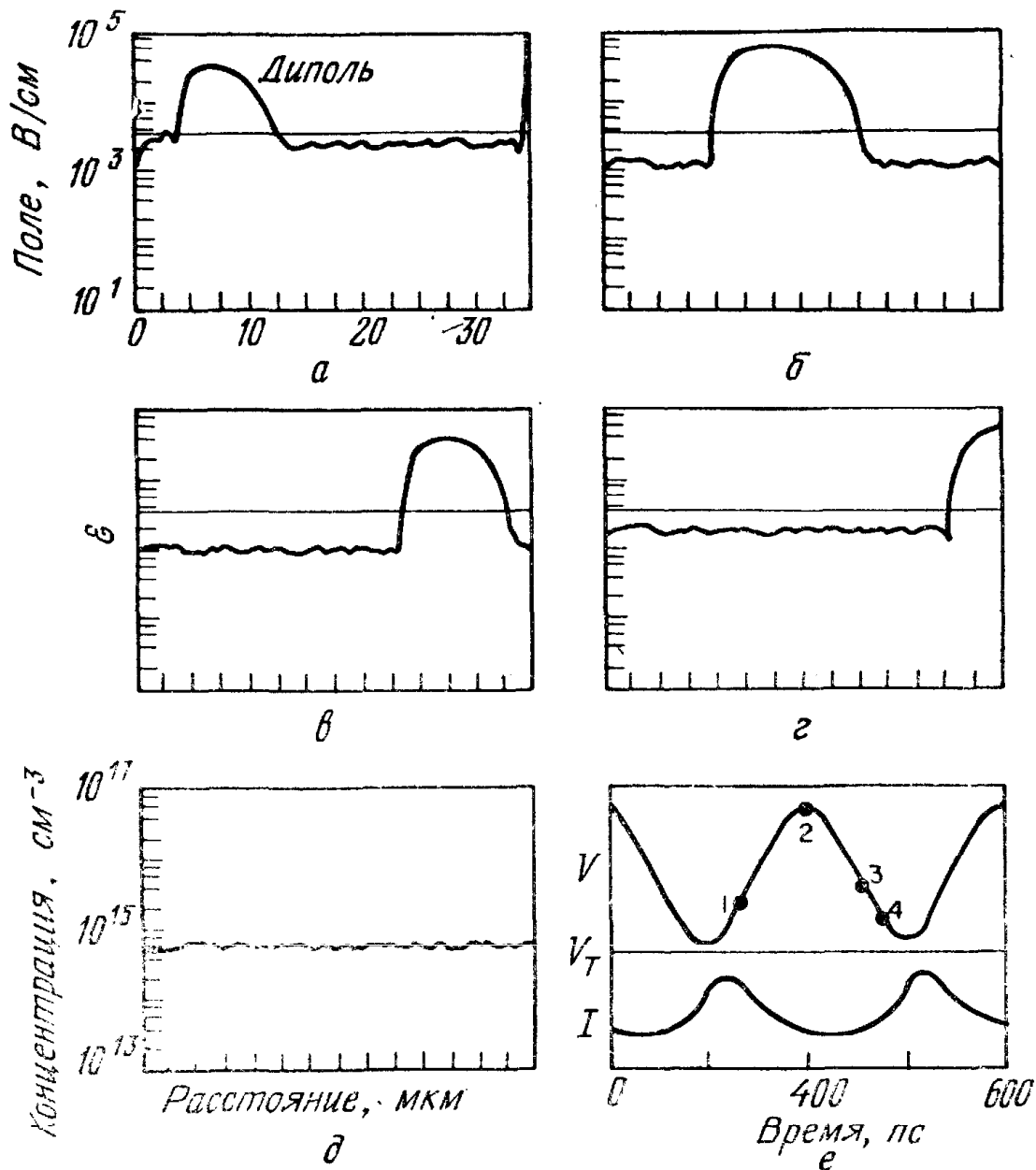


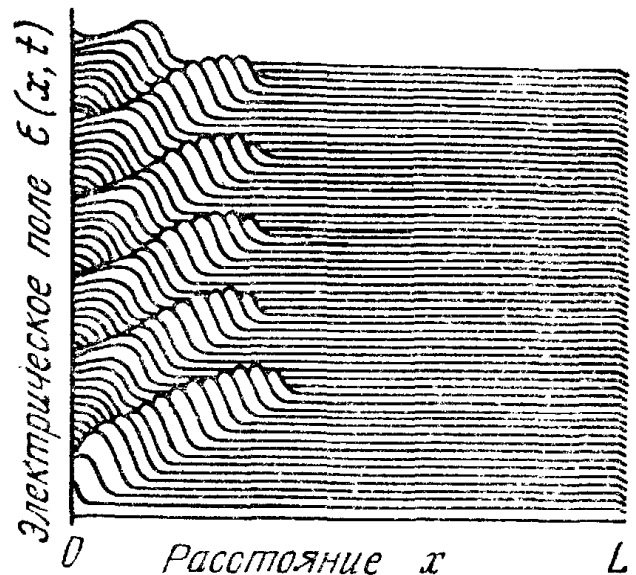
Рис. 23. Распределение поля в четыре различных момента одного СВЧ-периода для режима с пролетом домена (а—г), а также профиль легирования примесью (д) и зависимости тока и напряжения от времени (е) для образца из GaAs при $n_0L = 2,1 \cdot 10^{12} \text{ см}^{-2}$ и $fL = 0,9 \cdot 10^7 \text{ см/с}$ [29].

му генерируемых импульсов можно получить, если в момент исчезновения домена на аноде напряжение будет меньше порогового. Домен сформируется только после того, как напряжение превысит пороговое. Однако при реализации режима с запаздывающим формированием домена возникают большие трудности, связанные с настройкой.

11.3.4. Режим с разрушением домена

Включенный в резонансную цепь прибор на эффекте междолинного перехода электронов может работать на частотах, больших пролетной, при условии, что домен высокого поля рассеивается до того, как достигнет анода. В режиме пролета диполь-

Рис. 24. Численное моделирование режима с разрушением домена [33].



ного слоя большая часть падения напряжения на приборе приходится на домен. Поэтому ширина диполя уменьшается при понижении напряжения на диоде за счет изменения тока в резонансной цепи (рис. 20) до тех пор, пока в некоторой точке обедненный и обогащенный слой не нейтрализуют друг друга. Напряжение в этот момент равно V_s . Рассасывание домена происходит при уменьшении напряжения на приборе ниже V_s . Когда напряжение вновь превысит пороговое, возникнет новый домен, и процесс повторится. Следовательно, генерация будет происходить на частоте, определяемой параметрами резонансной цепи, а не на пролетной частоте.

На рис. 24 показан режим с разрушением домена [33]. Параметры прибора такие же, как и для кривых на рис. 21. Домен рассасывается на расстоянии $\sim L/3$ от катода, а рабочая частота примерно в три раза превышает пролетную (см. рис. 21).

Экспериментально [35] и теоретически [29] показано, что в режиме с разрушением домена в образцах, для которых резонансная частота цепи в несколько раз превышает пролетную (т. е. $fL > 2 \cdot 10^7$ см/с), а рабочая частота равна величине, обратной времени диэлектрической релаксации (т. е. $n_0/f \approx \epsilon_s/q|\mu|$), обычно образуются несколько диполей, поскольку электрическое поле не успевает перераспределиться в один домен. На рис. 25 показано формирование нескольких дипольных слоев в режиме работы с разрушением домена. Для этого образца $n_0L = 4,2 \times 10^{12}$ см⁻², $fL = 4,2 \cdot 10^7$ см/с и $n_0/f = 10^5$ с/см³.

Верхний предел рабочих частот в данном режиме определяется скоростью рассасывания домена, которая зависит от двух постоянных времени: времени диэлектрической релаксации и времени RC -цепи, где R равно сопротивлению той части полупроводника, в которой нет диполей, а C — емкость следующих друг за другом доменов. Первая постоянная времени определяет минимальную величину отношения $n_0/f \approx 10^4$ с/см³ для арсенида гал-

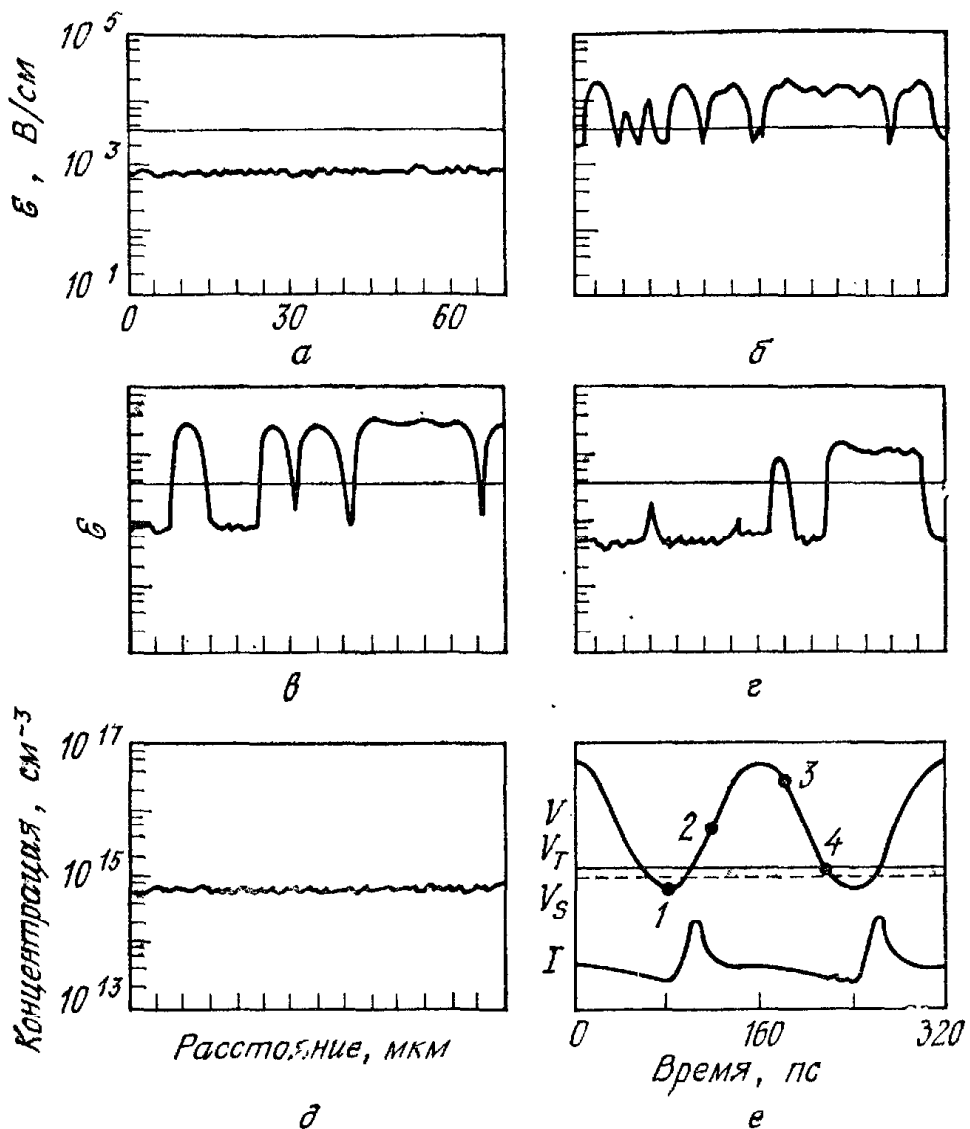


Рис. 25. Образование нескольких диполей в режиме с разрушением домена [29].

лия и фосфида индия n -типа [37, 38]. Вторая постоянная времени зависит от числа доменов и длины образца. Теоретическое максимальное значение к. п. д. генератора, работающего в режиме с разрушением домена, равно 13 % [36].

11.3.5. Режим ограниченного накопления объемного заряда [37]

В модели этого режима предполагается, что напряженность электрического поля увеличивается до пороговой и затем уменьшается так быстро, что распределение пространственного заряда, приводящее к образованию домена, не успевает сформироваться. У катода образуется только первичный обогащенный слой, а распределение электрического поля в остальной части прибора почти однородно, если только флуктуации концентрации примеси малы и их величина недостаточна для возникновения доменов. В этих условиях напряженность электрического поля в большей части прибора одинакова, что приводит к эффективной генерации мощности на частоте, определяемой параметрами внешней цепи. Чем

выше частота, тем меньше длина пути, проходимого слоем пространственного заряда, причем в большей части полупроводника дифференциальная подвижность остается отрицательной. Для работы прибора в этом режиме необходимо, чтобы в течение одного СВЧ-периода объемный заряд не успевал заметно возрасти, а обогащенный слой, наоборот, успел полностью разрушиться. Поэтому τ_R (выражение (9)), соответствующее отрицательной дифференциальной подвижности, должно быть больше по абсолютной величине, чем СВЧ-период, а τ_R , соответствующее положительной дифференциальной подвижности, должно быть меньше. Эти требования можно записать в виде неравенств [38]

$$\frac{\epsilon_s}{q\mu_+} \ll \frac{n_0}{f} < \frac{\epsilon_s}{q|\mu_-|}, \quad (39)$$

где μ_+ — положительная дифференциальная подвижность при малых напряженностях электрического поля, а μ_- — средняя отрицательная дифференциальная подвижность при полях, больших порогового. Для GaAs и InP получаем

$$10^4 < n_0/f < 10^5 \text{ с/см}^3. \quad (40)$$

Интересно отметить, что режим с разрушением нескольких доменов можно реализовать и в образце с флюктуациями примесной концентрации в некотором диапазоне значений n_0/f . Работающие в этом режиме приборы пригодны для генерации коротких импульсов большой мощности, поскольку можно использовать диоды большой длины (непролетная мода), для которых затруднен теплоотвод. Однако максимальная рабочая частота приборов, работающих в данном режиме, существенно ниже, чем приборов, работающих в пролетном режиме. Это связано с тем, что релаксация энергии электронов в основном минимуме зоны проводимости происходит достаточно медленно, что приводит к большим временам рассасывания домена. Из результатов численного моделирования на ЭВМ следует, что при генерации в непрерывном режиме максимальное время, в течение которого прибор на основе GaAs должен находиться при допороговом смещении, составляет ~ 20 пс; это ограничивает верхний предел рабочих частот величиной ~ 20 ГГц [39, 40]. Ожидается, что приборы из фосфида индия будут иметь более высокий верхний предел.

11.4. ХАРАКТЕРИСТИКИ ПРИБОРОВ НА ЭФФЕКТЕ МЕЖДОЛИННОГО ПЕРЕХОДА ЭЛЕКТРОНОВ

11.4.1. Катодные контакты

Для изготовления приборов на эффекте междолинного перехода электронов необходимы чрезвычайно чистые и однородные полупроводниковые материалы с низкой концентрацией глубоких

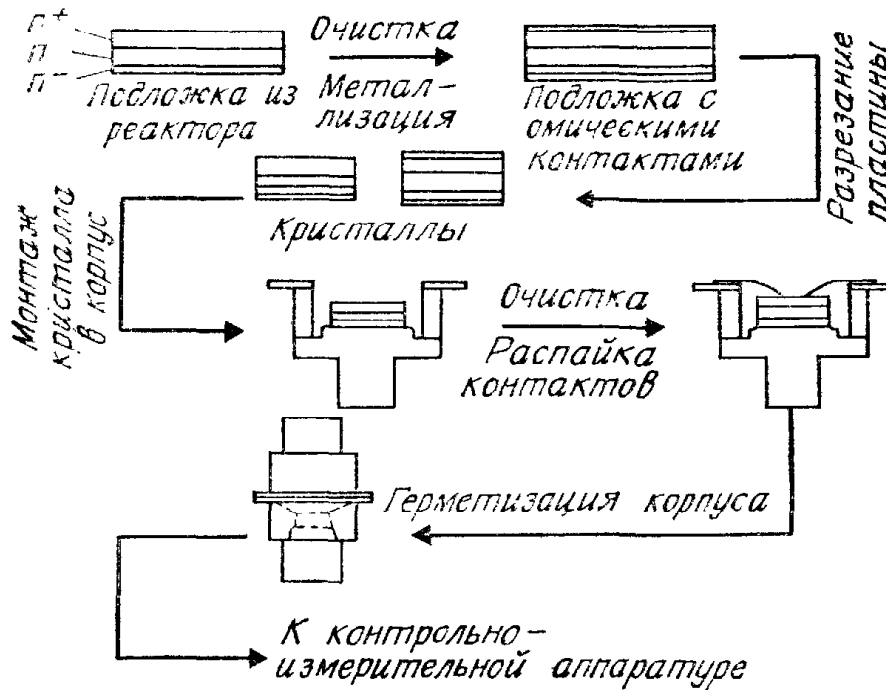


Рис. 26. Основные стадии процесса изготовления приборов на эффекте междолинного перехода электронов [41].

донорных уровней и ловушек, особенно для приборов, работающих в режиме с рассасыванием объемного заряда. Первые приборы изготавливались из арсенида галлия и фосфида индия со сплавными омическими контактами. В настоящее время для этих целей используются эпитаксиальные слои, наносимые на n^+ -подложку методами эпитаксии из газовой и жидкой фаз или методом молекулярно-лучевой эпитаксии. Типичный диапазон донорных концентраций составляет 10^{14} — 10^{16} см^{-3} , а длин прибора — от нескольких микрометров до нескольких сотен микрометров. Основные стадии процесса изготовления такого прибора на эпитаксиальном арсениде галлия $n^+—n—n^+$ -типа показаны на рис. 26 [41]. Кристаллы с приборами на эффекте междолинного перехода электронов монтируются в корпуса, которые, как и теплоотводы, аналогичны корпусам для ЛПД, описанным в гл. 10. Некоторые высокоомощные приборы изготавливаются при помощи селективной металлизации или травления мезоструктур.

Для улучшения характеристик прибора вместо омических n^+ -контактов применяют контакты с ограниченной инжекцией [42—44]. При использовании таких контактов можно реализовать ситуации, когда пороговое поле для тока приблизительно равно пороговой напряженности \mathcal{E}_T , определяющей начало участка ОДС. Поэтому поле в полупроводнике может быть однородным. В приборе с омическими контактами формирование обогащенного слоя или домена происходит на некотором расстоянии от катода из-за конечного времени разогрева электронов в основном минимуме зоны проводимости. Ширина «мертвой зоны» может достигать 1 мкм, что накладывает ограничения на минимальную

длину диода и, следовательно, максимальную рабочую частоту. В контактах с ограниченной инжекцией горячие электроны эмитируются из катода, что приводит к уменьшению ширины «мертвой зоны». Поскольку влияние пролетных эффектов можно свести к минимуму, отрицательное сопротивление прибора не будет зависеть от частоты и будет шунтироваться геометрической емкостью диода. Если к прибору подсоединить индуктивность и достаточно малое сопротивление, то прибор может генерировать колебания резонансной частоты в режиме однородного электрического поля. Теоретическая величина к. п. д. для этого случая найдена в разд. 11.3.1.

Были изучены два класса контактов с ограниченной инжекцией: контакт Шоттки с малой высотой барьера и двухслойная катодная структура. На рис. 27 приведены характеристики трех катодных контактов. В приборе с омическими контактами (рис. 27, а) всегда существует область низкой напряженности электрического поля у катода, а поле в образце неоднородно. В обратнорасположенном барьере Шоттки [45] можно получить достаточно однородное распределение электрического поля (рис. 27, б). Обратный ток равен (гл. 5)

$$J_R = A^{**} T^2 \exp(-q\phi_B/kT), \quad (41)$$

где A^{**} — эффективная константа Ричардсона, а $q\phi_B$ — высота барьера. Плотностям тока в диапазоне 10^2 — 10^4 А/см² соответствуют

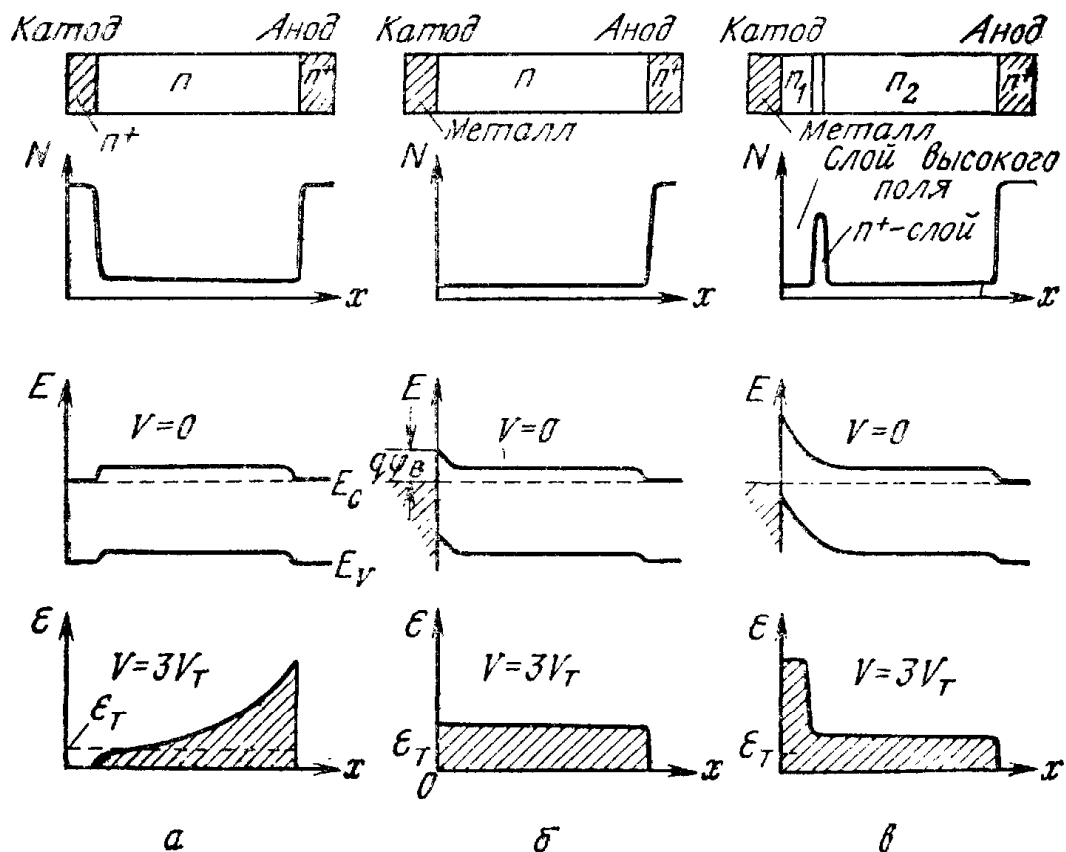


Рис. 27. Характеристики трех катодных контактов. а — омический; б — барьер Шоттки; в — двухслойный контакт с барьером Шоттки.

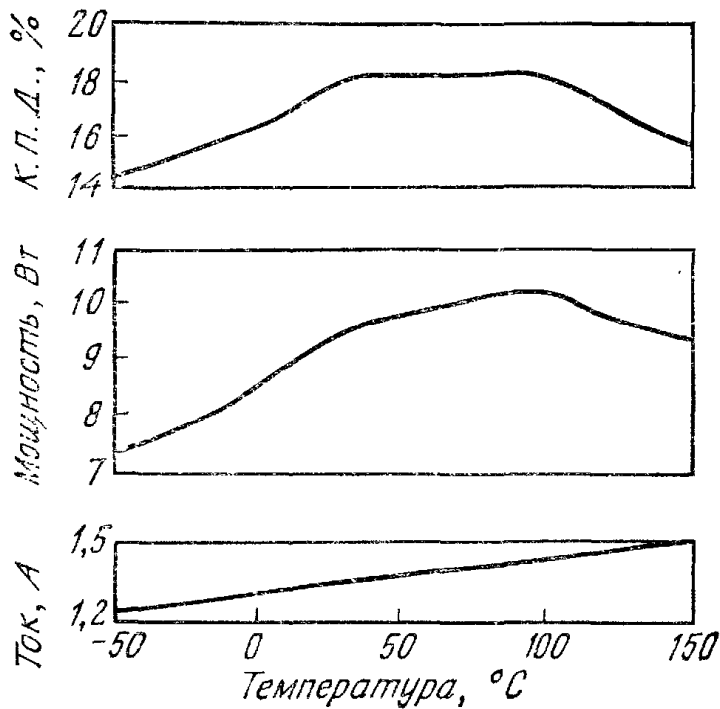


Рис. 28. К. п. д., генерируемая СВЧ-мощность и ток в приборе из фосфида индия с двухслойным катодным контактом, работающим в импульсном режиме при $n_0 = 2 \cdot 10^{15} \text{ см}^{-3}$ и $L = 8 \text{ мкм}$ и в области температур -50 — 150 °C [46].

значения высот барьера в пределах $0,15$ — $0,3 \text{ эВ}$. Однако нелегко получить барьер Шоттки небольшой высоты в полупроводниковых соединениях типа $A^{III}B^V$; кроме того, область рабочих температур ограничена, поскольку ток изменяется экспоненциально с температурой (выражение (41)).

Двухслойный катодный контакт состоит из слоя высокой напряженности электрического поля и n^+ -слоя (рис. 27, в) [46]. Его конфигурация аналогична конфигурации ЛПД с трехслойной базой (гл. 10). Электроны разогреваются в слое высокого электрического поля, а затем инжектируются в активную область, где поле однородно. Такие структуры успешно используются в широком диапазоне температур. Коэффициент полезного действия, генерируемая в импульсном режиме мощность и ток прибора из фосфида индия в области температур -50 — 150 °C приведены на рис. 28. Отметим, что в этом температурном интервале к. п. д. и мощность изменяются приблизительно всего лишь на $\pm 10 \%$. Наибольший к. п. д., полученный у приборов из фосфида индия с двухслойным катодным контактом, равен 24% . Из-за захвата уровня Ферми на поверхностные состояния катодные контакты с ограниченной инжекцией в приборах из GaAs до сих пор не созданы.

11.4.2. Зависимость мощности от частоты и шум

Процесс передачи энергии электрического поля электронам и переход последних из основного минимума зоны проводимости в побочные занимают конечный промежуток времени. Поэтому рабочие частоты приборов ограничены сверху величинами, обратными времени релаксации энергии и времени междолинного

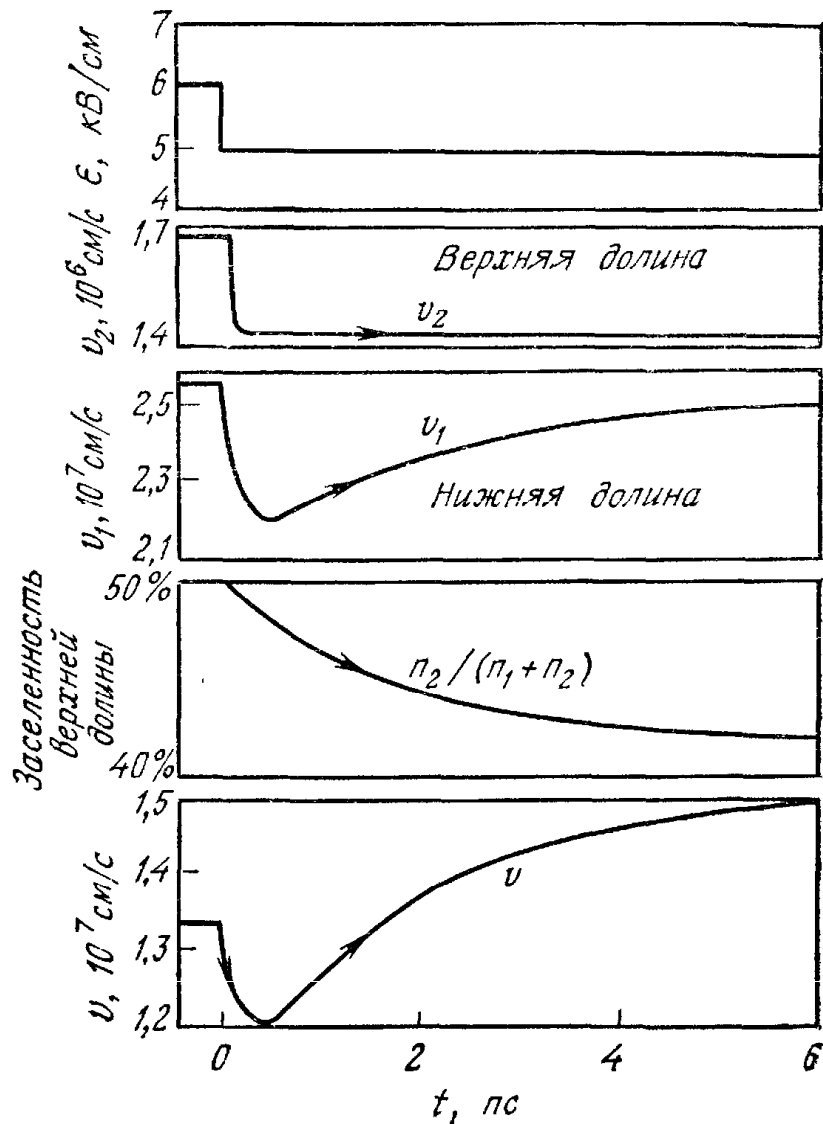


Рис. 29. Временные зависимости при внезапном изменении напряженности электрического поля от 6 до 5 кВ/см в момент $t = 0$ [47].

рассеяния. На рис. 29 приведены временные зависимости дрейфовой скорости носителей в верхней и нижней долинах, средней скорости и заселенности верхней долины при внезапном уменьшении напряженности электрического поля от 6 до 5 кВ/см [47]. Отметим, что скорость электронов в верхней долине v_2 практически мгновенно реагирует на изменение поля. Однако скорость в нижней долине v_1 изменяется медленно (характерное время отклика составляет ~ 5 пс), поскольку горячие электроны в основном минимуме слабо рассеиваются. Кроме того, причиной медленного изменения концентрации n_2 является большая величина времени междолинного рассеяния электронов. Поэтому отклик средней скорости v связан частично с изменением скорости v_1 и частично с междолинным рассеянием. Оценка верхнего предела рабочей частоты прибора из-за конечного времени отклика составляет ~ 150 ГГц.

В пролетном режиме рабочая частота обратно пропорциональна длине диода, т. е. $f = v/L$. Соотношение между генерируемой мощностью и частотой имеет вид

$$P_{rf} = V_{rf}^2/R = \mathcal{E}_{rf}^2 L^2/R = \frac{\mathcal{E}_{rf}^2 v^2}{Rf^2} \sim \frac{1}{f^2}, \quad (42)$$

где V_{rf} и \mathcal{E}_{rf} — СВЧ-напряжение и напряженность электрического поля соответственно, а R — импеданс. Поэтому ожидаемое изменение мощности с частотой пропорционально $1/f^2$. Зависимость генерируемой в непрерывном и импульсном режимах СВЧ-мощности от частоты для современных приборов на эффекте междолинного перехода электронов из арсенида галлия и фосфида индия показана на рис. 30 [48—53, 57, 58]. В скобках рядом с экспериментальными точками приведены значения к. п. д. в процентах. Как следует из формулы (42), изменение мощности пропорционально $1/f^2$. В импульсном режиме в приборах из GaAs с достаточно длинной базой получена СВЧ-мощность до 6 кВт на частоте, близкой к 2 ГГц. Генерируемая в непрерывном режиме мощность приблизительно равна 2 Вт на 10 ГГц, что примерно в пять раз меньше, чем для лавинно-пролетного диода (ЛПД). Однако при фиксированной частоте напряжение смещения в приборе на эффекте междолинного перехода электронов меньше (в 2—5 раз), чем в лавинно-пролетном диоде.

Характеристики приборов из фосфида индия в непрерывном режиме сравнимы с характеристиками приборов из арсенида галлия. Однако в импульсном режиме они хуже, несмотря на то что по теоретическим оценкам к. п. д. приборов из фосфида индия выше. Причиной этого является более развитая технология арсенида галлия.

В приборах на эффекте междолинного переноса обнаруживаются шумы двух видов: амплитудно-модулированный (АМ) и частотно-модулированный (ЧМ), причем эти шумы возникают из-за флюктуаций тепловой скорости электронов. Обычно АМ-шум мал, поскольку амплитуда относительно стабильна из-за сильной нелинейности зависимости скорости от напряженности электрического поля. Величина среднего отклонения частоты в случае ЧМ-шума равна [54]

$$f_{rms} = \frac{f_0}{Q_{ex}} \sqrt{\frac{kT_{eq}(f_m)B}{P}}, \quad (43)$$

где f_0 — пролетная частота, Q_{ex} — добротность внешней цепи, P_0 — генерируемая мощность, B — ширина полосы. Эквивалентная ЧМ-шумовая температура T_{eq} равна

$$T_{eq}(f_m) = qD/k|\mu_-|, \quad (44)$$

где величина средней отрицательной дифференциальной подвижности μ_- зависит от амплитуды напряжения. Поскольку отноше-

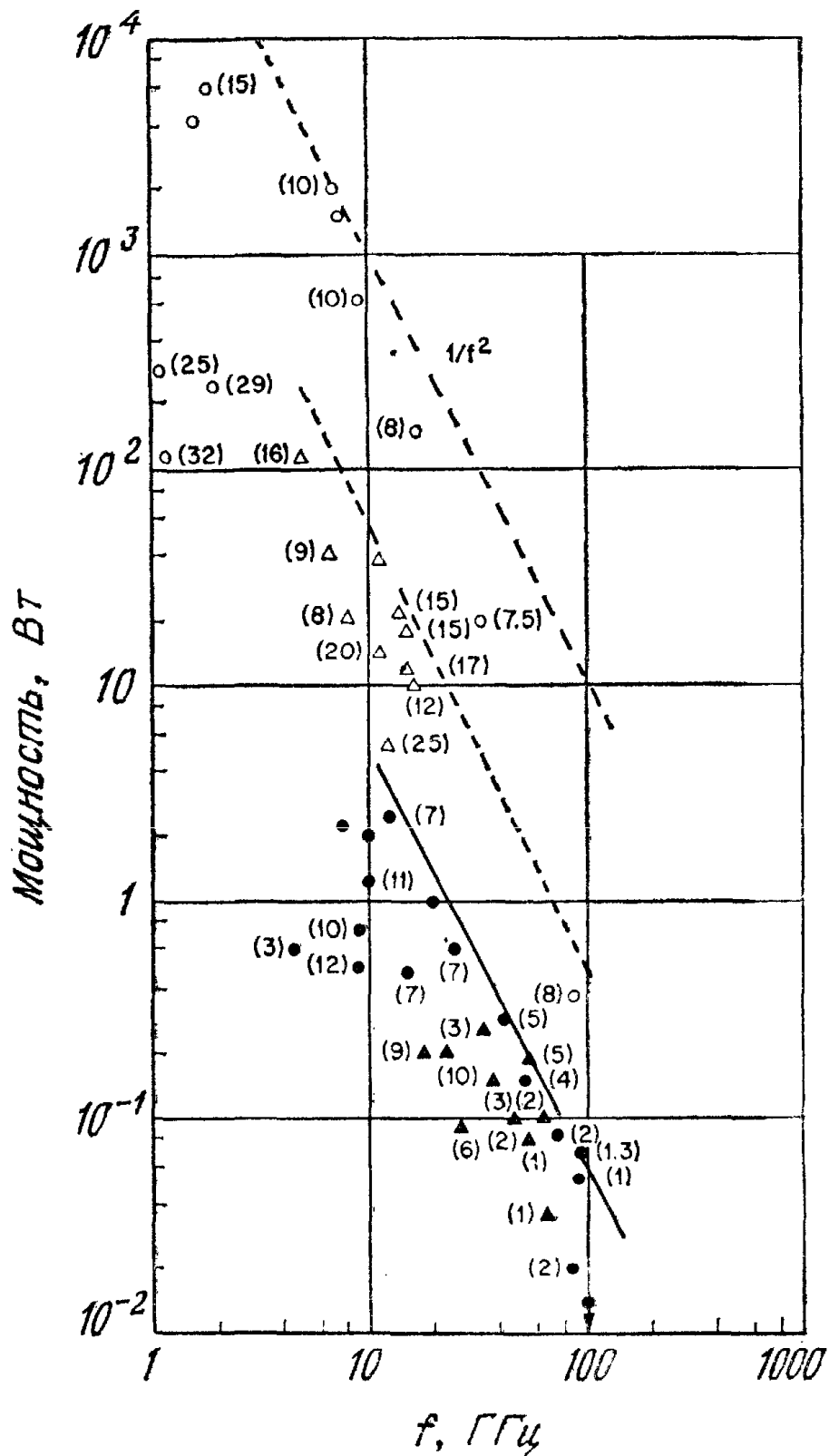
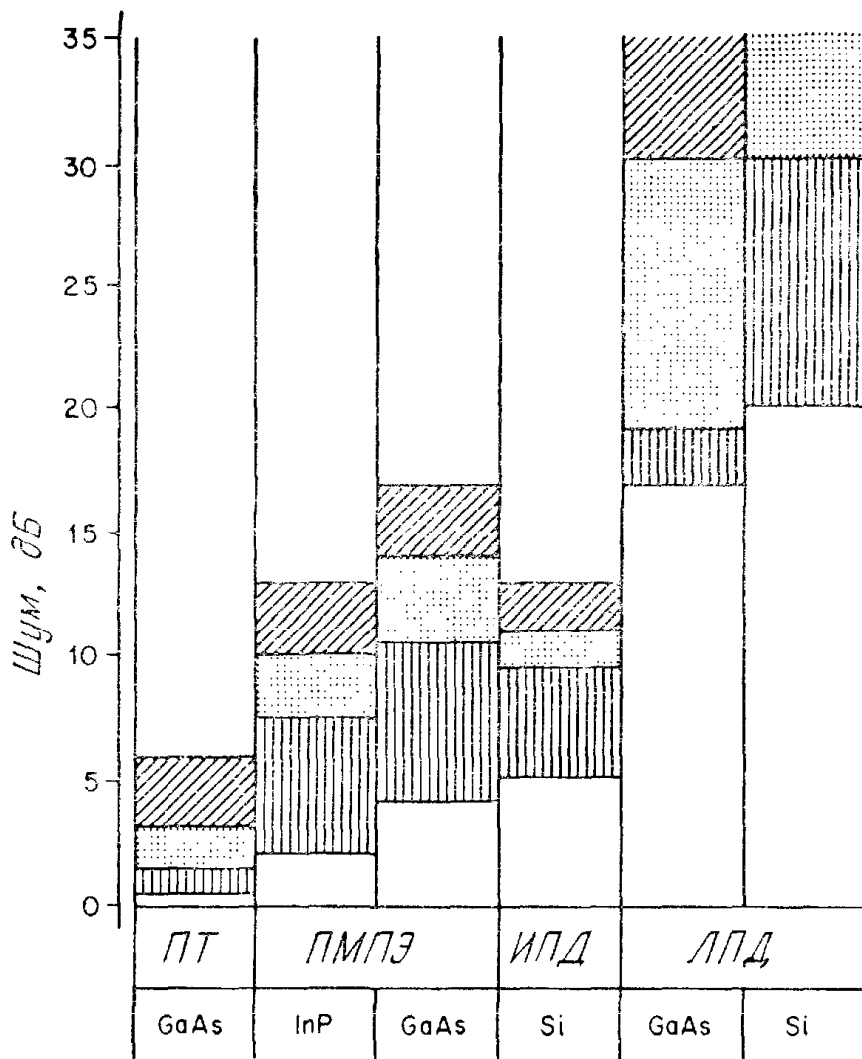


Рис. 30. Зависимость генерируемой СВЧ-мощности от частоты для приборов на эффекте междолинного перехода электронов.

○ импульсный режим, прибор из GaAs; Δ непрерывный режим, прибор из GaAs; ● импульсный режим, прибор из InP; ▲ непрерывный режим, прибор из InP.



□ Минимальная теоретическая величина
 ▒ Экспериментальные значения
 ▨ Типичные результаты $M=(F-1)/(1-1/\text{усиление})$

Рис. 31. Малосигнальный шум в полупроводниковых СВЧ-приборах в диапазоне частот 8—12 ГГц [55].

ПТ — полевой транзистор с затвором Шоттки; ПМПЭ — приборы на междолинном переходе электронов.

ние $D/|\mu_-|$ в фосфиде индия меньше, чем в арсениде галлия, ожидается, что уровень шума в приборах из InP будет также ниже.

Результаты измерения малосигнального шума в СВЧ-приборах в диапазоне частот 8—12 ГГц приведены на рис. 31 [55]. Наименьший шум наблюдается у полевого транзистора с затвором Шоттки из арсенида галлия (гл. 6). Шумы кремниевого ИПД сравнимы с шумами приборов на эффекте междолинного перехода электронов. Причиной того, что шумы в ЛПД больше, чем в других приборах, являются лавинные процессы.

11.4.3. Другие приборы

Мы рассмотрели эффект междолинного перехода электронов и его использование в СВЧ-генераторах и усилителях. Приборы на этом эффекте применяются в быстродействующих цифровых и ана-

логовых устройствах. Ниже мы рассмотрим приборы с изменяющейся вдоль направления тока площадью поперечного сечения и приборы с тремя контактами.

При анализе процессов в генераторе с переменной площадью поперечного сечения можно воспользоваться результатами одномерной теории доменов высокого поля, если предположить, что толщина диполя мала, и рассматривать явления в практически однородной области полупроводника вблизи домена. Эти допущения справедливы, если $n_0L \gg 10^{12} \text{ см}^{-2}$, а площадь поперечного сечения диода изменяется плавно при удалении от контакта. Воспользовавшись теоретическими результатами, полученными в предыдущем разделе, можно показать, что существует такое избыточное напряжение на домене $V_{ex} = V_{rm}$, при превышении которого электрическое поле вне домена \mathcal{E}_r не изменяется. Соответствующая величина напряженности электрического поля вне домена равна \mathcal{E}_{rm} (рис. 16). Когда избыточное напряжение на домене $V_{ex} = V_{rm}$, плотность тока равна

$$J_{rm} = qn_0v_{rm}. \quad (45)$$

Такие «насыщенные» домены движутся в генераторе с постоянной скоростью. Пусть толщина объемного генератора равна s , а изме-

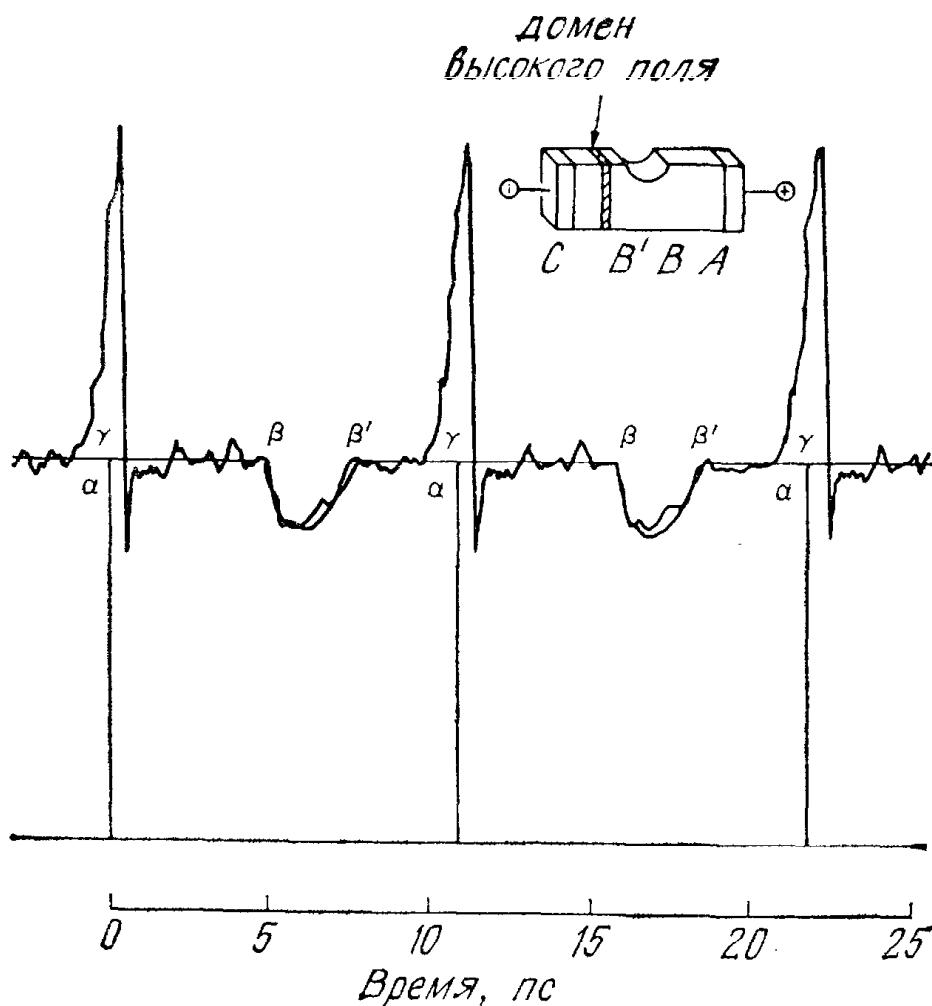


Рис. 32. Форма тока прибора, показанного в верхней части рисунка [56].

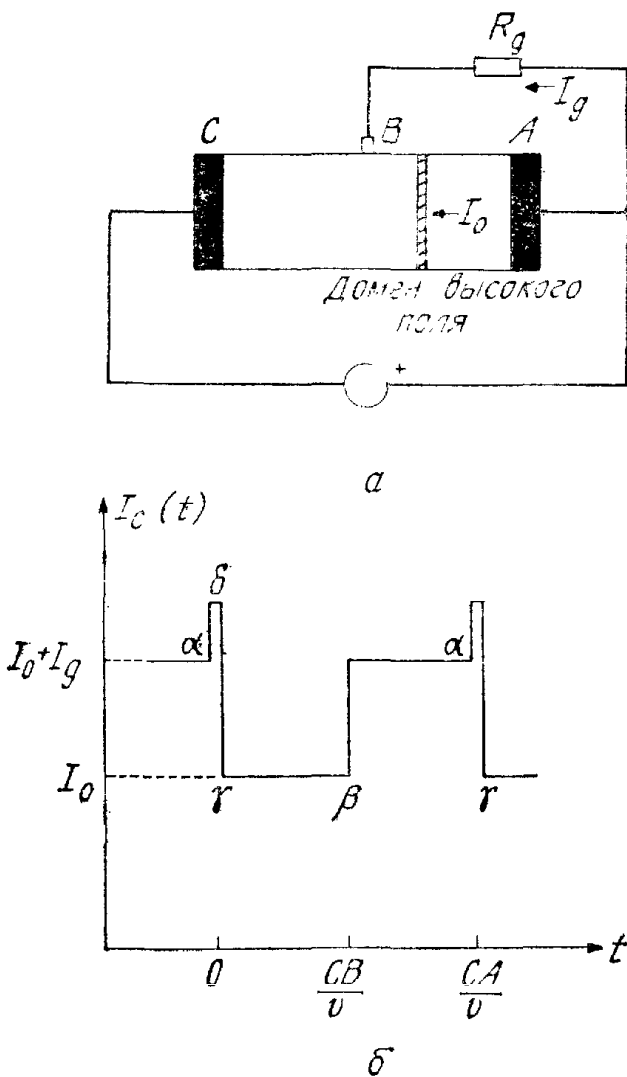


Рис. 33. Генератор управляемых прямоугольных импульсов тока (а) и их форма (б).

няющаяся с расстоянием от катода x ширина — $b(x)$. Если домен высокого поля зарождается у катода при $t = 0$, то в момент времени t диполь находится на расстоянии $x(t) = v_{rm}t$. В любом поперечном сечении прибора ток должен быть одинаков, а вблизи домена равен

$$I(t) = J_{rm}sb(v_{rm}t) + I_g(t), \quad (46)$$

где $I_g(t)$ — связанный с исчезновением домена ток, который отличен от нуля только в конце периода. Поэтому из равенства (46) следует, что ток пропорционален $b(v_{rm}t)$.

На рис. 32 приведена форма импульсов тока прибора, показанного в верхней части рисунка [56]. Тонкими сплошными линиями показана ожидаемая для такой конфигурации прибора форма тока. Обозначения α , β , β' и γ соответствуют моментам времени, в которые домен находится в точках А, В, В' и С.

До сих пор мы рассматривали только приборы с двумя контактами. Формой генерируемых импульсов можно управлять с помощью одного или нескольких электродов, расположенных вдоль длины прибора. На рис. 33, а показана конфигурация такого прибора с электродом в точке В [56]. Ожидаемая форма тока приве-

дена на рис. 33, б. Здесь снова используется теория насыщенного домена, описанная выше. После отрыва домена от катода при $t = 0$ ток катодного контакта равен току насыщенного домена (Aqn_0v_{rm}) до тех пор, пока диполь не достигнет электрода, расположенного в точке B . В этот момент ток становится равным сумме тока насыщенного домена и протекающего через сопротивление R_g тока I_g . Ток I_g равен напряжению между точками A и B при наличии между ними домена, поделенному на R_g . После этого ток катода остается равным

$$I_c(t) = Aqn_0v_{rm} + I_g \quad (47)$$

до тех пор, пока домен не достигнет анода, вызвав короткий «всплеск» тока.

ЛИТЕРАТУРА

1. Gunn J. B. Microwave Oscillation of Current in III—V Semiconductors, *Solid State Commun.*, 1, 88 (1963); Instabilities of Current in III—V Semiconductors, *IBM J. Res. Dev.*, 8, 141 (1964).
2. Kroemer H. Theory of the Gunn Effect, *Proc. IEEE*, 52, 1736 (1964).
3. Ridley B. K., Watkins T. B. The Possibility of Negative Resistance Effects in Semiconductors, *Proc. Phys. Soc. Lond.*, 78, 293 (1961).
- 3a. Ridley B. K. Anatomy of the Transferred-Electron Effect in III—V Semiconductors, *J. Appl. Phys.*, 48, 754 (1977).
4. Hilsum C. Transferred Electron Amplifiers and Oscillators, *Proc. IRE*, 50, 185 (1962).
- 4a. Hilsum C. Historical Background of Hot Electron Physics, *Solid State Electron.*, 21, 5 (1978).
5. Hutson A. R., Jayaraman A., Chynoweth A. G., Coriel A. S., Feldmann W. L. Mechanism of the Gunn Effect from a Pressure Experiment, *Phys. Rev. Lett.*, 14, 639 (1965).
6. Allen J. W., Shyam M., Chen Y. S., Pearson G. L. Microwave Oscillations in $\text{GaAs}_{1-x}\text{P}_x$ Alloys, *Appl. Phys. Lett.*, 7, 78 (1965).
7. Carroll J. E. Hot Electron Microwave Generator, Edward Arnold, London, 1970.
8. Bulman P. J., Hobson G. S., Taylor B. S. Transferred Electron Devices, Academic, N. Y., 1972.
9. Bosch B. G., Engelmann R. W. H. Gunn-Effect Electronics, Wiley, N. Y., 1975.
10. Thim H. W. Solid State Microwave Sources, in Hilsum C., Ed., Handbook on Semiconductors, Vol. 4, Device Physics, North-Holland, Amsterdam, 1980.
11. Ridley B. K. Specific Negative Resistance in Solids, *Proc. Phys. Soc. Lond.*, 82, 954 (1963).
12. Ridley B. K., Pratt R. G. A Bulk Differential Negative Resistance Due to Electron Tunneling through an Impurity Potential Barrier, *Phys. Lett.*, 4, 300 (1963).
13. McWhorter A. L., Rediker R. H. The Cryosar — A New Low-Temperature Computer Component, *Proc. IRE*, 47, 1207 (1959).
14. Kroemer H. Negative Conductance in Semiconductors, *IEEE Spectrum*, 5, 47 (1968).
15. Aspnes D. E. GaAs Lower Conduction Band Minimum: Ordering and Properties, *Phys. Rev.*, 14, 5331 (1976).
16. Rees H. D., Gray K. W. Indium Phosphide: A Semiconductor for Microwave Devices, *Solid State Electron Devices*, 1, 1 (1976).

17. McCumber D. E., Chynoweth A. G. Theory of Negative Conductance Application and Gunn Instabilities in 'Two-Valley' Semiconductors, *IEEE Trans. Electron Devices*, **ED-13**, 4 (1966).
18. Sakai K., Ikoma T., Adachi Y. Velocity-Field Characteristics of $\text{Ga}_x\text{In}_{1-x}\text{Sb}$ Calculated by the Monte Carlo Method, *Electron Lett.*, **10**, 402 (1974).
19. Hayes R. E., Raymond R. M. Observation of the Transferred-Electron Effect in GaInAsP , *Appl. Phys. Lett.*, **31**, 300 (1977).
20. Ruch J. G., Kino G. S. Measurement of the Velocity-Field Characteristics of Gallium Arsenide, *Appl. Phys. Lett.*, **10**, 40 (1967).
21. Butcher P. N., Fawcett W. Calculation of the Velocity-Field Characteristics for Gallium Arsenide, *Phys. Lett.*, **21**, 489 (1966).
22. Littlejohn M. A., Hauser J. R., Glisson T. H. Velocity — Field Characteristics of GaAs with $\Gamma - L - X$ Conduction-Band Ordering, *J. Appl. Phys.*, **48**, 4587 (1977).
23. Mojzes I., Podor B., Balogh I. On the Temperature Dependence of Peak Electron Velocity and Threshold Field Measured on GaAs Gunn Diodes, *Phys. Status Solidi*, **39**, K123 (1977).
24. Shockley W. Negative Resistance Arising from Transit Time in Semiconductor Diodes, *Bell Syst. Tech. J.*, **33**, 799 (1954).
25. Kino G. S., Kuru I. High-Efficiency Operation of a Gunn Oscillator in the Domain Mode, *IEEE Trans. Electron Devices*, **ED-16**, 735 (1969).
26. Hobson G. S. The Gunn Effect, Clarendon, Oxford, 1974.
27. Thim H. W., Haydl W. Microwave Amplifier Circuit Consideration, in Howes M. J., Morgan D. V., Eds., *Microwave Devices*, Wiley, N. Y., 1976, Chap. 6.
28. Hakki B. W. Amplification in Two-Valley Semiconductors, *J. Appl. Phys.*, **38**, 808 (1967).
29. Thim H. W. Computer Study of Bulk GaAs Devices with Random One-Dimensional Doping Fluctuations, *J. Appl. Phys.*, **39**, 3897 (1968).
30. Butcher P. N. Theory of Stable Domain Propagation in the Gunn Effect, *Phys. Lett.*, **19**, 546 (1965).
31. Butcher P. N., Fawcett W., Hilsum C. A Simple Analysis of Stable Domain Propagation in the Gunn Effect, *Br. J. Appl. Phys.*, **17**, 841 (1966).
32. Copeland J. A. Electrostatic Domains in Two-Valley Semiconductors, *IEEE Trans. Electron Devices* **ED-13**, 187 (1966).
33. Shaw M., Grubin H. L., Solomon P. R. The Gunn — Hilsum Effect, Academic, N. Y. 1979.
34. Fukui H. New Method of Observing Current Waveforms in Bulk GaAs. *Proc. IEEE*, **54**, 792 (1966).
35. Thim H. W., Barber M. R. Observation of Multiple High-Field Domains in n -GaAs, *Proc. IEEE*, **56**, 110 (1968).
36. Barber M. R. High Power Quenched Gunn Oscillators, *Proc. IEEE*, **56**, 752 (1968).
37. Copeland J. A. A New Mode of Operation for Bulk Negative Resistance Oscillators, *Proc. IEEE*, **54**, 1479 (1966).
38. Copeland J. A. LSA Oscillator Diode Theory, *J. Appl. Phys.*, **38**, 3096 (1967).
39. Jones D., Rees H. D. Electron-Relaxation Effects in Transferred-Electron Devices Revealed by New Simulation Method, *Electron. Lett.*, **8**, 363 (1972).
40. Kroemer H. Hot Electron Relaxation Effects in Devices, *Solid State Electron.*, **21**, 61 (1978).
41. Sterzer F. Transferred Electron Amplifiers and Oscillators for Microwave Application, *Proc. IEEE*, **59**, 1155 (1971).
42. Kroemer H. The Gunn Effect under Imperfect Cathode Boundary Condition, *IEEE Trans. Electron Devices*, **ED-15**, 819 (1968).
43. Atalla M. M., Moll J. L. Emitter Controlled Negative Resistance in GaAs, *Solid State Electron.*, **12**, 619 (1969).

44. Yu S. P., Tantraporn W., Young J. D. Transit-Time Negative Conductance in GaAs Bulk — Effect Diodes, *IEEE Trans. Electron. Devices*, ED-18, 88 (1971).
45. Colliver D. J., Irving L. D., Pattison J. E., Rees H. D. High-Efficiency InP Transferred-Electron Oscillators, *Electron. Lett.*, 10, 221 (1974).
46. Gray K. W., Pattison J. E., Rees J. E., Prew B. A., Clarke R. C., Irving L. D. InP Microwave Oscillator with 2-Zone Cathode, *Electron. Lett.*, 11, 402 (1975).
47. Rees H. D. Time Response of the High-Field Electron Distribution Function in GaAs. *IBM J. Res. Dev.*, 13, 537 (1969).
48. Hobson G. S. Recent Development in Transferred Devices, *J. Phys. E*, 7, 229 (1974).
49. Irving L. D., Pattison J. E., Braddock P. W., Gray K. W. Improved Mean Power and Long Pulse-Width Operation of InP TEDs in J Band, *Electron. Lett.*, 14, 116 (1978).
50. Hamilton R. J., Fairman R. D., Long S. I., Omori M., Fank F. B. InP Gunn-Effect Devices for Millimeter-Wave Amplifiers and Oscillators, *IEEE Trans. Microwave Theory Tech.*, MTT-24, 775 (1976).
51. Day W. R. Gunn Oscillators, *Microwave Syst. News*, 8, 56 (1978).
52. Suzuki T., Ito M., Ishii T., Mitsui S. Design and Fabrication for High Efficiency and High Output Power Gunn Diodes, *Trans. Inst. Electron. Commun. Eng., Jpn.*, E61, 932 (1978).
53. Mun J. High-Efficiency and High-Peak-Power InP Transferred-Electron Oscillators, *Electron. Lett.*, 13, 275 (1977).
54. Ataman A., Harth W. Intrinsic FM Noise of Gunn Oscillators, *IEEE Trans. Electron Devices*, ED-20, 12 (1973).
55. Constant E. Noise in Microwave, Injection, Transit Time and Transferred-Electron Devices, *Physica*, 83B, 24 (1976).
56. Shoji M. Functional Bulk Semiconductor Oscillators, *IEEE Trans. Electron Devices*, DED-14, 535 (1967).
57. Ondria J. Wide-Band Mechanically Tunable W-Band (75—110GHz) CW GaAs Diode Oscillators, Proc. 7th Bienn. Conf. Active Microwave Semicond. Devices Circuits, Cornell University, Ithaca, N. Y., 1979, p. 309.
58. Crowley J. D., Fank F. B., Hyder S. B., Sowers J. J., Tringali D. Millimeter Wave InP Transferred Electron Devices, Proc. 7th Bienn. Conf. Active Microwave Semicond. Devices Circuits, Cornell University, Ithaca, N. Y., 1979, p. 331.
59. Hauser J. R., Glisson T. H., Littlejohn M. A. Negative Resistance and Peak Velocity in the Central (000) Valley of III—V Semiconductors, *Solid State Electron.*, 22, 487 (1979).

Часть V

Оптоэлектронные приборы

Глава 12

СВЕТОДИОДЫ И ПОЛУПРОВОДНИКОВЫЕ ЛАЗЕРЫ

12.1. ВВЕДЕНИЕ

К оптоэлектронным приборам относятся приборы, в которых основные процессы протекают с участием квантов света — фотонов. Оптоэлектронные приборы можно разделить на три группы: 1) приборы, преобразующие электрическую энергию в оптическое излучение (светодиоды, полупроводниковые лазеры); 2) приборы, детектирующие оптические сигналы за счет протекающих под действием света электронных процессов (фотодетекторы); 3) приборы, осуществляющие преобразование оптического излучения в электрическую энергию (фотовольтаические приборы, солнечные батареи).

В данной главе рассматривается первая группа приборов, а гл. 13 и 14 посвящены соответственно фотодетекторам и солнечным батареям.

Явление электролюминесценции было открыто в 1907 г. [1]. Оно состоит в генерации света в материале, сквозь который протекает электрический ток, вызванный электрическим полем. Излучение электролюминесценции в отличие от теплового характеризуется сравнительно узким спектром (для светодиодов спектральная ширина линии лежит обычно в диапазоне длин волн 100—500 Å). Излучение может быть практически монохроматичным; например, для лазерного диода спектральная ширина излучения лежит в пределах 0,1—1 Å. Подробное изучение светодиодов проведено в работе [2], а вопросам практического использования светодиодов в качестве источников света, дисплеев и оптронов посвящена работа [3].

После открытия в 1954 г. мазера и последующей разработки оптических и рубиновых лазеров [5, 6] было предложено использовать полупроводники в качестве материала для лазеров [7—10]. В 1961 г. на основе теоретических расчетов [11] были найдены условия, необходимые для осуществления режима генерации лазерного излучения с использованием концепции квазиуровней Ферми. В 1962 г. в работе [12] было показано, что генерация

может быть практически реализована в прямозонных полупроводниках, и были найдены важнейшие критерии для такого режима. В том же 1962 г. три группы авторов почти одновременно сообщили об осуществлении режима лазерной генерации в полупроводниках [13—15]. Было получено импульсное излучение с длиной волны 0,84 мкм в $p-n$ -переходах на основе GaAs, смещенных в прямом направлении и охлажденных до температуры жидкого азота. Вскоре после этого появилось сообщение о получении генерации излучения на длине волны 0,71 мкм в $p-n$ -переходе на основе тройного соединения $\text{GaAs}_{1-x}\text{P}_x$ [16]. В 1970 г. авторы работы [17] реализовали режим непрерывной генерации при комнатной температуре в лазерах на основе двойных гетеропереходов. Впервые структуры такого типа были предложены в работах [18, 19].

Со времени этих первых открытий было найдено множество новых материалов для лазеров. Диапазон длин волн когерентного излучения включает близкую ультрафиолетовую, видимую и дальнюю инфракрасную области спектра (от $\sim 0,3$ до ~ 30 мкм). Полупроводниковые лазеры рассматриваются как основные источники света для волоконно-оптических систем связи. Они могут найти также широкое применение во многих областях фундаментальных исследований, таких, как газовая спектроскопия высокого разрешения и контроль атмосферных загрязнений. Свойства лазеров на основе гетероструктур подробно рассмотрены в работах [20, 21].

12.2. ИЗЛУЧАТЕЛЬНЫЕ ПЕРЕХОДЫ

Светодиоды и полупроводниковые лазеры принадлежат к классу люминесцентных приборов. Люминесценцией называется оптическое излучение (в ультрафиолетовой, видимой или инфракрасной областях спектра), возникающее в результате электронного возбуждения материала, в отличие от теплового излучения, обусловленного только тепловым нагревом материала. На рис. 1 приведена диаграмма электромагнитного спектра. Несмотря на то что возбуждение излучения на разных длинах волн осуществляется разными методами, природа излучения во всех случаях одинакова. Излучение, видимое человеческим глазом, лежит в узкой области длин волн примерно от 0,4 до 0,7 мкм. На рис. 1 показаны основные цветовые зоны видимого излучения от фиолетовой до красной. Инфракрасная область простирается от 0,7 до ~ 1000 мкм, а ультрафиолетовая область включает длины волн от 0,4 до $\sim 0,01$ мкм (т. е. 100 Å, или 10 нм). В этой и последующих главах нас в основном будут интересовать области длин волн от близкой ультрафиолетовой ($\sim 0,3$ мкм) до близкой инфракрасной ($\sim 1,5$ мкм).

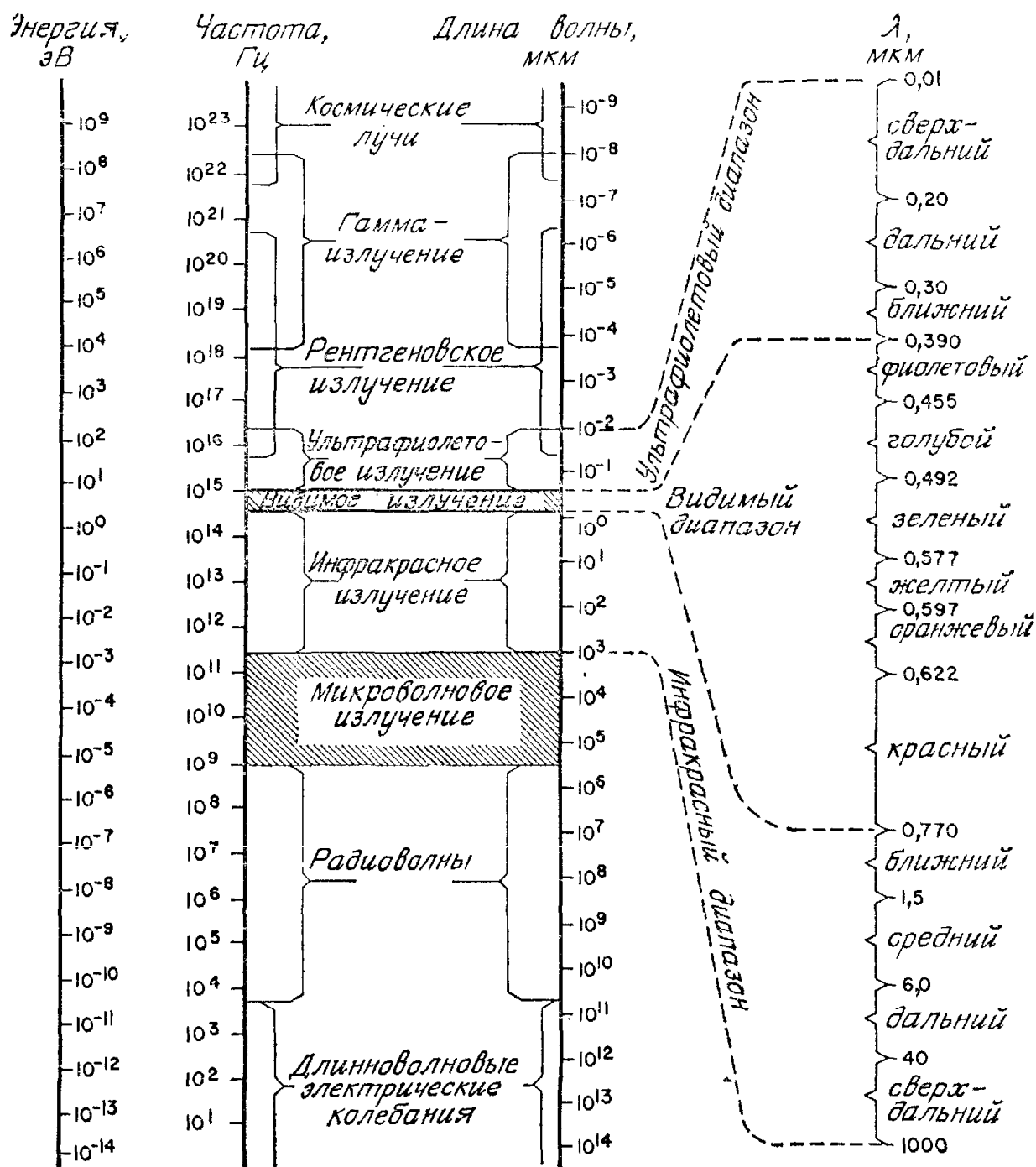


Рис. 1. Диаграмма электромагнитного спектра.

В зависимости от источника энергии возбуждения люминесценцию можно разделить [22] на: 1) фотолюминесценцию (возбуждение оптическим излучением); катодолюминесценцию (возбуждение под действием пучка электронов, или катодных лучей); 3) радиационную люминесценцию (возбуждение другими быстрыми частицами или излучением высокой энергии) и 4) электролюминесценцию (возбуждение электрическим полем или током). Мы будем рассматривать в основном электролюминесценцию, и прежде всего инжекционную электролюминесценцию, которая представляет собой оптическое излучение, возникающее при инжекции неосновных носителей в область полупроводникового $p - n$ -перехода, где имеются излучательные переходы.

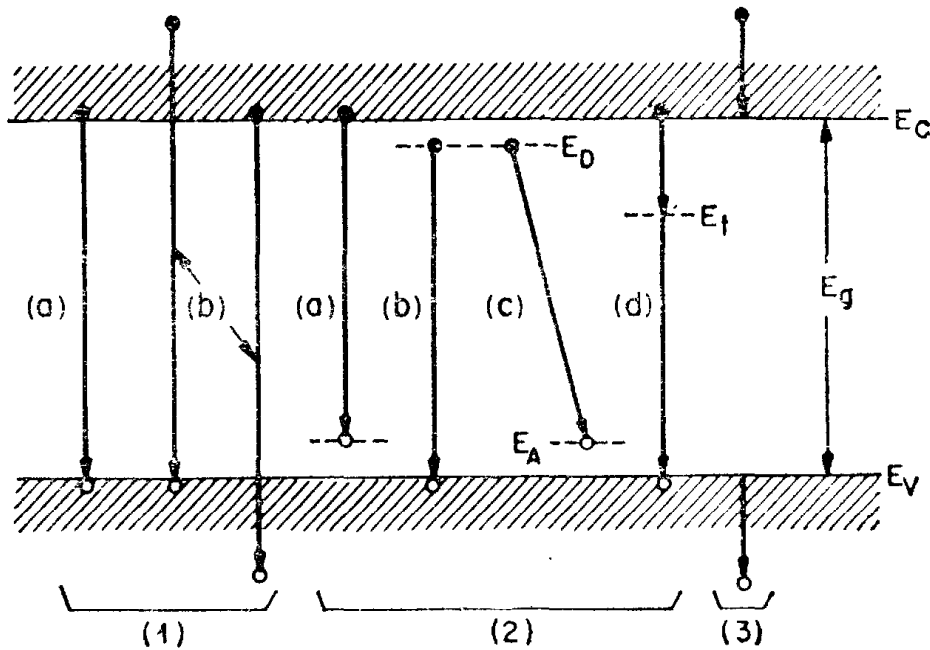


Рис. 2. Основные переходы в полупроводнике [23].

● электрон; ○ дырка.

На рис. 2 схематически представлены основные переходы в полупроводнике [23]. Эти переходы можно классифицировать следующим образом. Во-первых, это междузонные переходы, вызывающие: а) собственное излучение с энергией, очень близкой к ширине запрещенной зоны, которое может сопровождаться возбуждением фононов или экситонов; б) излучение с более высокой энергией с участием энергичных или «горячих» носителей, которое иногда может быть связано с лавинным пробоем. Во-вторых, это переходы с участием химических примесей или физических дефектов: а) между зоной проводимости и акцепторным уровнем; б) между донорным уровнем и валентной зоной; в) между донорным и акцепторным уровнями (межпримесное излучение); г) через глубокие уровни. В-третьих, это внутрizonные переходы, которые вызывают излучение, называемое иногда тормозным, и которые протекают с участием «горячих» носителей.

Не все переходы могут возникать в одном и том же материале или при одних и тех же условиях и не все из них являются излучательными. Эффективным с точки зрения люминесценции является такой материал, в котором излучательные переходы преобладают над безизлучательными (к последним относятся, например, переходы при оже-рекомбинации) [2].

12.2.1. Спектры излучения

Существуют три типа взаимодействия между фотонами и электронами в твердом теле. Фотон может поглотиться в результате перехода электрона из заполненного состояния валентной зоны в свободное состояние зоны проводимости. Кроме того, он может

стимулировать излучение подобного себе фотона, вызывая переход электрона из заполненного состояния в зоне проводимости в свободное состояние в валентной зоне. Могут также возникать спонтанные обратные переходы электронов из зоны проводимости на свободные состояния в валентной зоне, вызывающие испускание фотонов.

Интенсивность спонтанного излучения зависит от плотности заполненных состояний в зоне проводимости и плотности свободных состояний в валентной зоне [20]:

$$I(h\nu) \sim \nu \langle M \rangle^2 N_c N_v F_c(E) F_v(E), \quad (1)$$

где $\langle M \rangle$ — матричный элемент перехода; N_c — плотность состояний в зоне проводимости; N_v — плотность состояний в валентной зоне; $F_c(E)$ и $F_v(E)$ — функции распределения Ферми — Дирака для электронов и дырок соответственно.

Спектр спонтанного излучения обычно имеет вид [24]

$$I(h\nu) \sim \nu^2 (h\nu - E_g)^{1/2} \exp[-(h\nu - E_g)/kT], \quad (2)$$

где E_g — ширина запрещенной зоны.

Следует отметить, что при характерных для светодиодов и полупроводниковых лазеров концентрациях примеси представление локализованных примесных уровней, отстоящих от краев запрещенной зоны, уже неприменимо. В этом случае случайное распределение заряженных примесей в кристалле приводит к флуктуациям потенциала, которые создают на краях запрещенной зоны хвосты плотности состояний валентной зоны и зоны проводимости (разд. 12.4.3).

По мере увеличения концентрации примеси энергия ионизации донорных и акцепторных уровней уменьшается до нуля, уровни ионизируются (предполагается, что они не образуют комплексов), и концентрация свободных носителей уже не зависит от температуры. Поэтому в примесном полупроводнике с низкой степенью компенсации концентрация свободных электронов n_0 или дырок p_0 будет равна концентрации доноров или акцепторов соответственно. Значение концентрации, при которой энергия ионизации приближается к нулю, находится из условия

$$r/a^* \approx 3,0. \quad (3)$$

В уравнении (3) величина r составляет среднее расстояние между примесными атомами, равное

$$r \equiv \left(\frac{3}{4\pi n} \right)^{1/3}. \quad (4)$$

Здесь n — концентрация примеси и a^* — радиус водородоподобного примесного центра, равный

$$a^* = \frac{1}{2} \left(\frac{\epsilon_s}{\epsilon_0} \right) \left(\frac{m_0}{m^*} \right), \quad (5)$$

где m^* — эффективная масса электрона или дырки соответственно для донорного или акцепторного полупроводника.

В основе стандартной теории оптических междузонных переходов лежит так называемое правило k -отбора. Волновой вектор k_1 , соответствующий волновой функции валентной зоны, и волновой вектор k_2 , соответствующий волновой функции зоны проводимости, должны различаться на волновой вектор фотона, т. е. матричный элемент $\langle M \rangle$ равен нулю. Поскольку волновой вектор электрона существенно превосходит волновой вектор фотона, правило k -отбора обычно записывается в виде равенства

$$k_1 = k_2. \quad (6)$$

Разрешенными являются переходы, при которых начальное и конечное состояния характеризуются одинаковыми волновыми векторами; такие переходы называются «прямыми» или «вертикальными». Если минимум зоны проводимости и максимум валентной зоны не соответствуют одному и тому же значению вектора k , то для сохранения квазиимпульса при переходах необходимо участие фонона; эти переходы называются «непрямыми». Введение в полупроводник примесей приводит к возмущению волновых функций и матричных элементов перехода, вследствие чего правило k -отбора строго не выполняется. Поэтому в обычно рас-

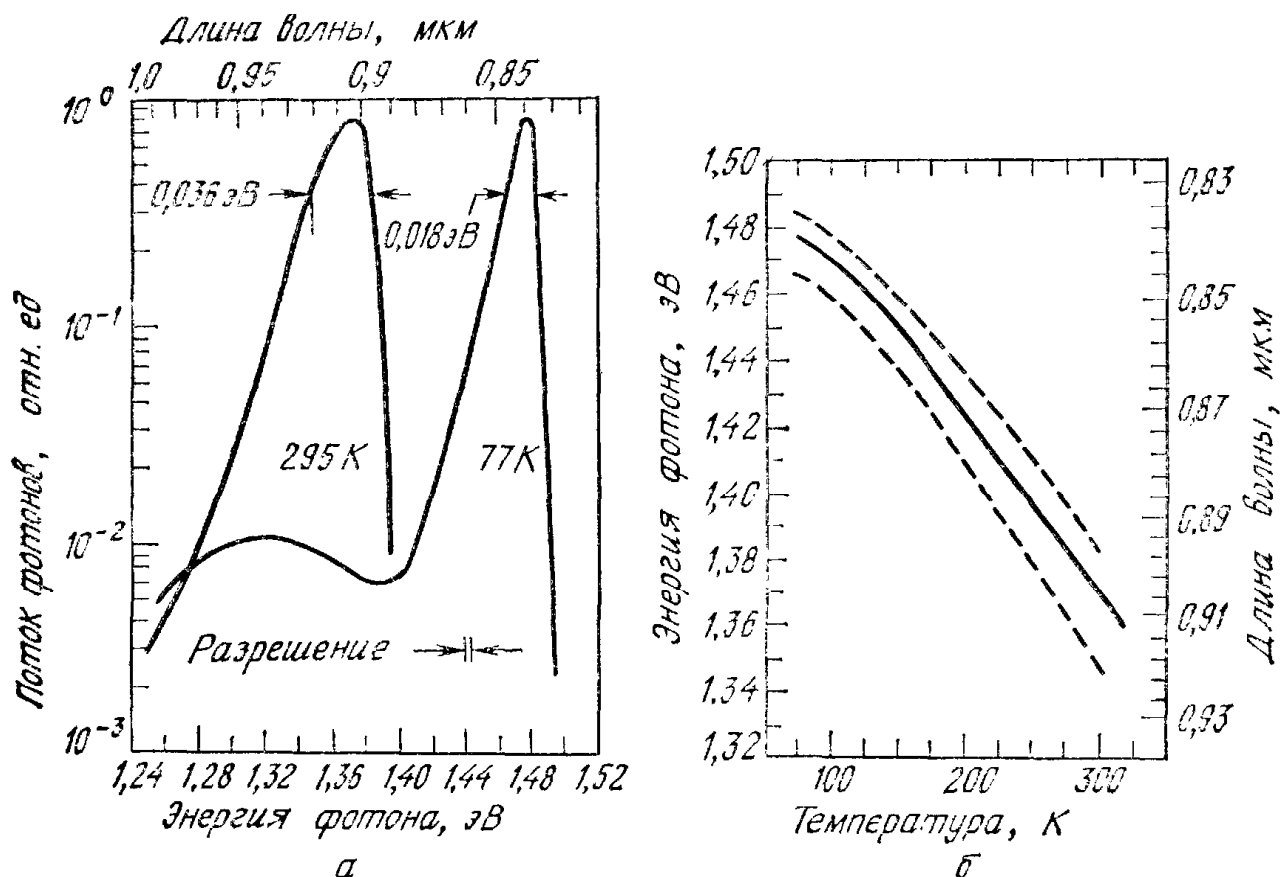


Рис. 3. Спектры излучения диода на GaAs (а) при 295 и 77 К и температурная зависимость энергии максимума (сплошная кривая) и уровней полуширины спектра излучения (штриховые кривые) (б) [25].

сма триваемых случаях наблюдаются «запрещенные» переходы, и матричный элемент перехода зависит от энергии фотона.

На рис. 3, а представлены спектры излучения $p-n$ -перехода на основе GaAs, наблюдаемые при температурах 77 и 295 К [25]. Уменьшение энергии фотона, соответствующей максимуму спектра, при возрастании температуры обусловлено в основном сужением запрещенной зоны. На рис. 3, б приведена температурная зависимость энергии максимума и уровней полуширины спектра излучения диода. Полуширина спектра слабо увеличивается с ростом температуры, как это и следует ожидать из уравнения (2).

12.2.2. Эффективность люминесценции

При заданной энергии возбуждения наряду с излучательной рекомбинацией протекают конкурирующие безызлучательные процессы (рис. 4) [26]. Квантовая эффективность люминесценции определяется как отношение числа возбужденных носителей, дающих вклад в излучение, к полному числу носителей, участвующих в рекомбинации, и может быть выражена через времена жизни следующим образом:

$$\eta_q = R_r/R = \tau_{nr}/(\tau_{nr} + \tau_r), \quad (7)$$

где τ_r и τ_{nr} — времена жизни излучательной и безызлучательной рекомбинации соответственно, а R_r и R — скорости излучательной и полной рекомбинации. Для полупроводниковых слоев p -типа проводимости скорость рекомбинации и время жизни связаны соотношением

$$R = (n - n_0)/\tau. \quad (8)$$

Аналогично для слоев n -типа

$$R = (p - p_0)/\tau, \quad (9)$$

где n_0 и p_0 — концентрации электронов и дырок в состоянии теплового равновесия, а n и p — электронная и дырочная кон-

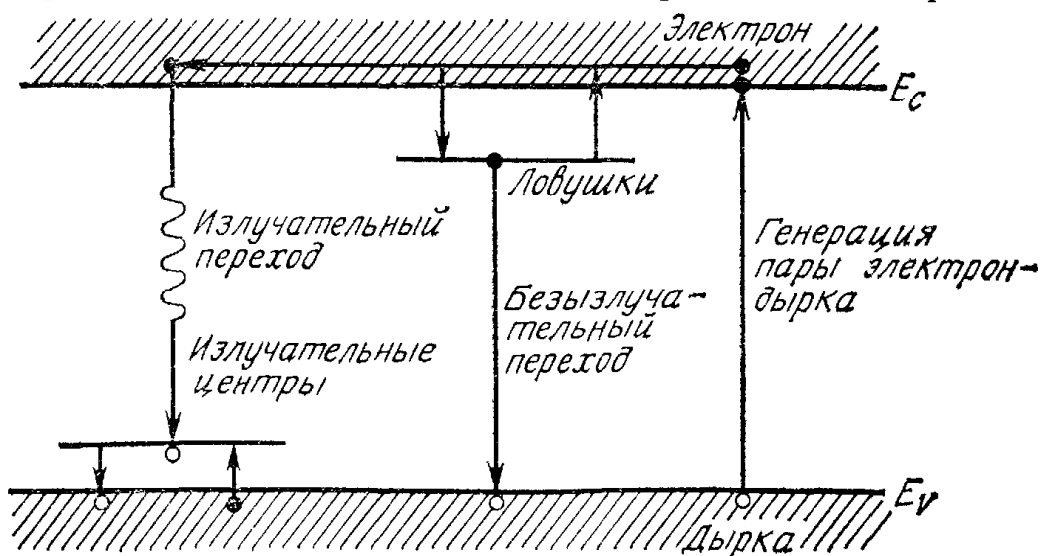


Рис. 4. Процессы излучательной и безызлучательной рекомбинации [26].

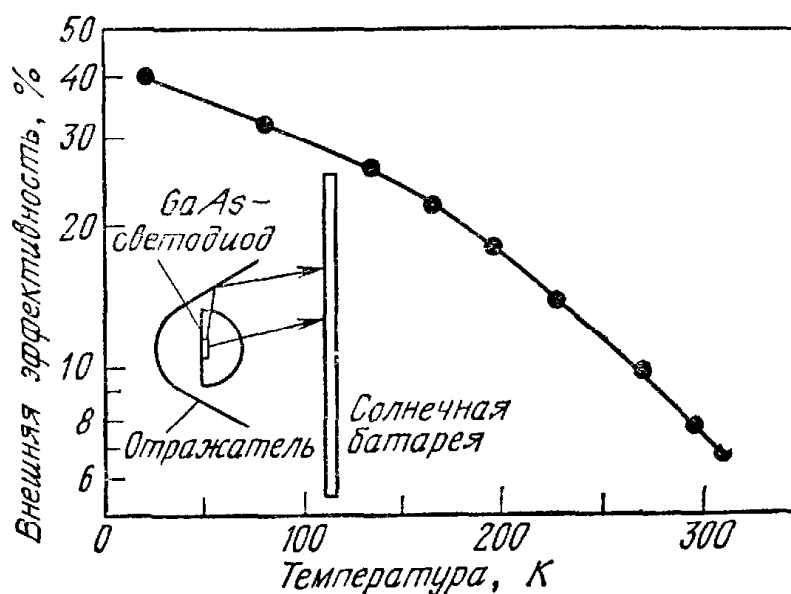


Рис. 5. Зависимость квантовой эффективности GaAs-светодиода от температуры. На вставке показано поперечное сечение устройства, используемого для измерения эффективности [25].

центрации при оптическом возбуждении. Время жизни τ неосновных носителей дается выражением

$$\tau = \tau_r \tau_{nr} / (\tau_{nr} + \tau_r). \quad (10)$$

Из уравнения (7) видно, что для обеспечения высокой квантовой эффективности время жизни τ_r должно быть мало.

Вставка на рис. 5 иллюстрирует схему измерения эффективности. Измерение излучения производится непосредственно по току короткого замыкания солнечной батареи. Внешняя эффективность, измеренная как функция от температуры, представлена на рис. 5 [25]. С ростом температуры эффективность уменьшается и имеет значения около 40, 32 и 7 % при температурах 20, 77 и 300 К соответственно. Уменьшение эффективности обусловлено уменьшением отношения τ/τ_r с ростом температуры.

12.2.3. Методы возбуждения

К способам, которыми можно возбуждать электролюминесценцию, относятся междузонное (собственное) возбуждение, лавинное возбуждение и возбуждение при туннелировании и инжекции. Междузонное возбуждение обеспечивается при воздействии переменного электрического поля на порошкообразный полупроводник (например, ZnS), впрессованный в диэлектрик (пластмассу или стекло). Электролюминесценция обычно имеет место при частотах, лежащих в звуковом диапазоне. Ее эффективность обычно низка ($\leq 1\%$). Основными механизмами возбуждения при этом являются ударная ионизация ускоренных электронов или полевая эмиссия электронов с ловушек [23, 27].

Для того чтобы создать лавинное возбуждение, p — n -переход или контакт металл — полупроводник смещают в обратном направлении в область лавинного пробоя. Электронно-дырочные пары, возникающие при ударной ионизации, могут вызвать излучение за счет междузонных (лавинное излучение) или внутризонных (тормозное излучение) переходов. Электролюминесценция может быть также результатом туннелирования в область прямо-смещенного или обратосмещенного перехода. При достаточно большом обратном смещении на контакте металл — полупроводник (вырожденный, p -типа) дырки с уровня Ферми в металле могут туннелировать в валентную зону и вызывать излучательную рекомбинацию с электронами, туннелирующими из валентной зоны в зону проводимости [28].

Наиболее важным методом возбуждения электролюминесценции является инжекционный [29]. При прямом смещении на p — n -переходе инжекция неосновных носителей через переход может привести к весьма эффективной излучательной рекомбинации, так как в этом случае электрическая энергия непосредственно преобразуется в фотоны. В последующих разделах мы в первую очередь будем рассматривать инжекционные электролюминесцентные приборы.

12.3. СВЕТОДИОДЫ

Светодиоды — это p — n -переходы, которые при прямом смещении могут испускать спонтанное излучение в ультрафиолетовой, видимой и инфракрасной областях электромагнитного спектра. Мы рассмотрим сначала светодиоды видимого участка спектра, которые широко применяются в информационных каналах, обеспечивающих связь электронной аппаратуры с ее пользователями, а затем инфракрасные светодиоды, которые эффективно используются в оптронах и представляют собой перспективные источники света для волоконно-оптических линий связи.

12.3.1. Светодиоды видимого диапазона

Эффективность воздействия света на глаз человека определяется функцией относительной видности глаза $V(\lambda)$, зависящей от длины волны. На рис. 6 приведена функция относительной видности глаза при угле зрения 2° , определенная Международной комиссией по освещению (МКО) [2]. Чувствительность глаза максимальна на длине волны 0,555 мкм ($V(0,555) = 1,0$) и спадает до нуля при длинах волн 0,39 и 0,77 мкм, соответствующих краям видимого диапазона спектра. Эффективность использования лучистой энергии при зрительном восприятии характеризуется яркостью лучистой энергии. Для стандартного фотометрического

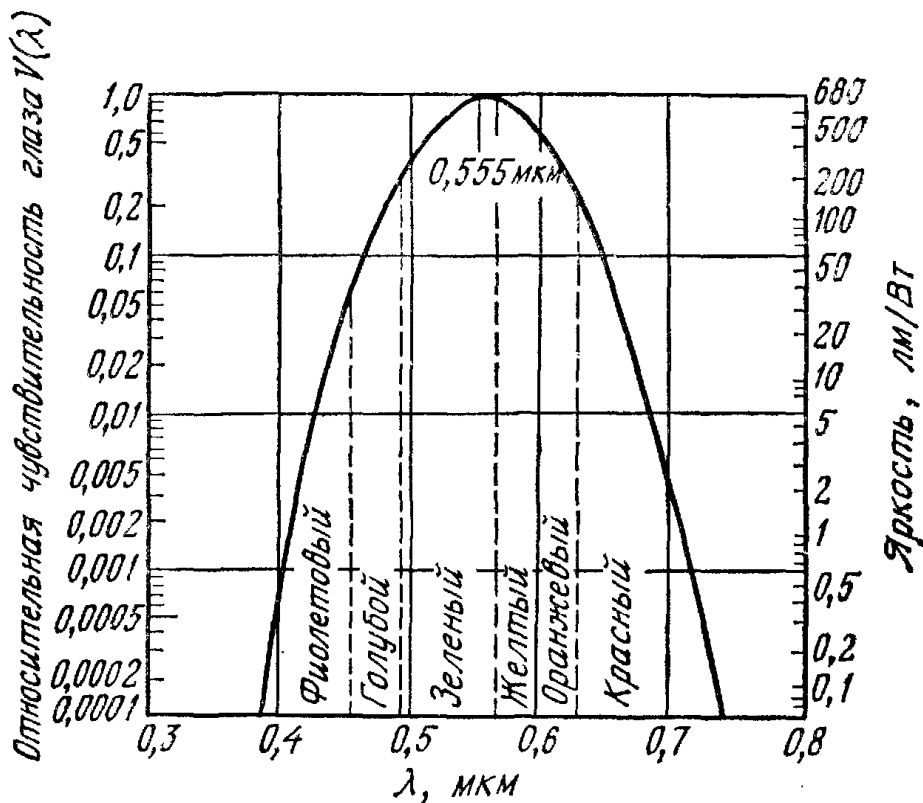


Рис. 6. Функция относительной видности глаза, определенная Международной комиссией по освещению (МКО) для стандартного фотометрического наблюдателя. Обозначены основные цветовые зоны видимого излучения.

наблюдателя 1 Вт лучистой энергии в максимуме чувствительности глаза (0,555 мкм) соответствует 680 лм. На рис. 6 показаны шесть основных цветовых участков видимого диапазона от фиолетового до красного.

Материалы для светодиодов видимого диапазона. Поскольку глаз чувствителен только к свету с энергией $h\nu \geq 1,8$ эВ ($\sim 0,7$ мкм), то полупроводники, которые могут быть использованы для создания светодиодов видимого диапазона, должны иметь ширину запрещенной зоны больше этого значения. На рис. 7 приведен перечень полупроводников, представляющих интерес с этой точки зрения. Заштрихованные прямозонные полупроводники особенно важны для электролюминесцентных приборов, так как излучательная рекомбинация в них является процессом первого порядка (протекает без участия фононов) и ожидаемая квантовая эффективность должна быть значительно более высокой, чем для полупроводников с непрямой зоной, в которых при переходах возбуждаются фононы. Среди всех приведенных на рис. 7 полупроводников наибольший интерес представляет $\text{GaAs}_{1-x}\text{P}_x$, относящийся к соединениям типа $A^{\text{III}} - B^{\text{V}}$.

На рис. 8, а приведена зависимость ширины запрещенной зоны $\text{GaAs}_{1-x}\text{P}_x$ от молярного содержания компонент x [20]. В диапазоне $0 < x < 0,45$ запрещенная зона прямая, а ее ширина увеличивается от $E_g = 1,424$ эВ при $x = 0$ до $E_g = 1,977$ эВ при

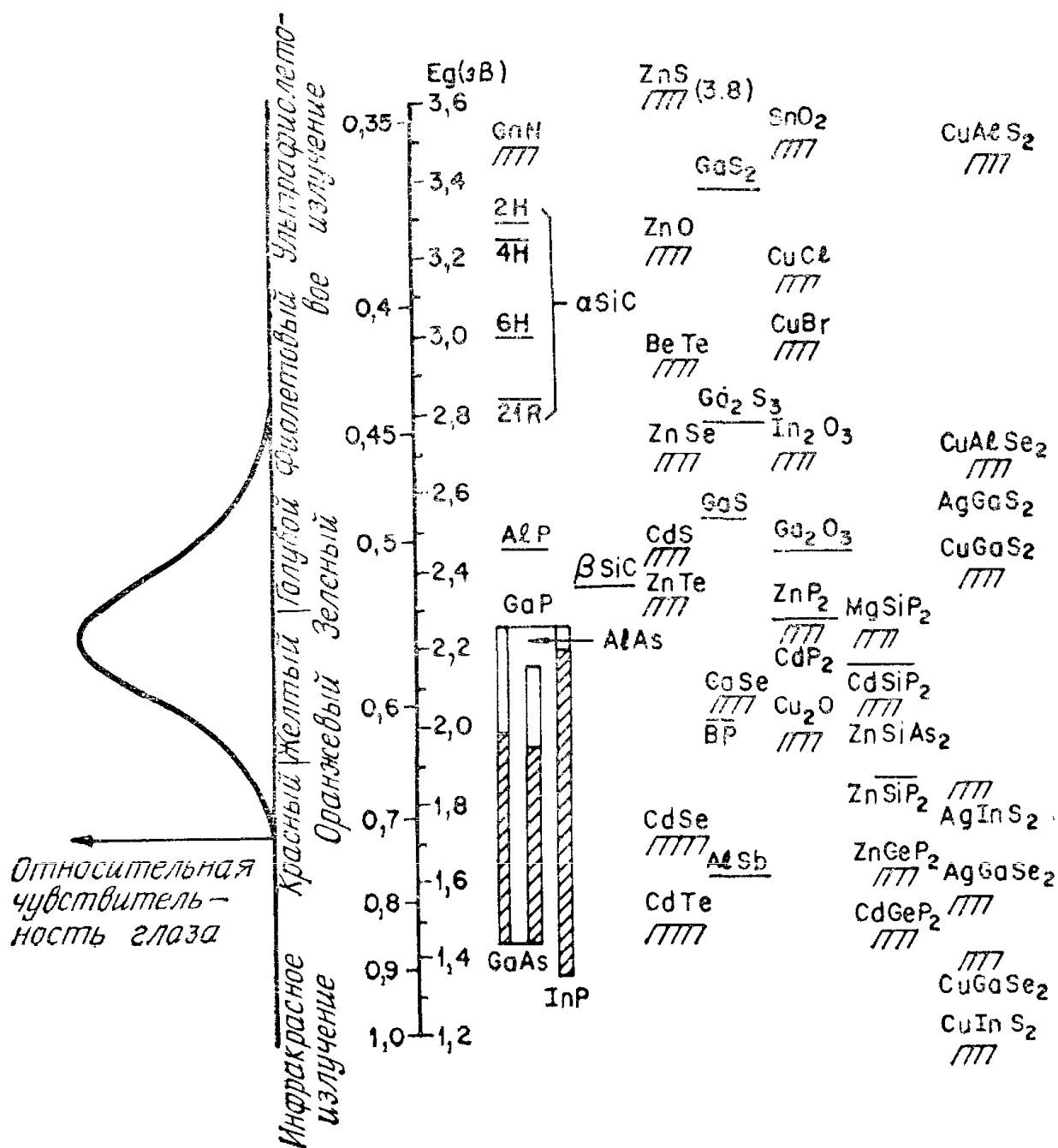


Рис. 7. Полупроводники, на основе которых могут быть созданы светодиоды со спектром излучения, включающим диапазон видимого света [2].

$x = 0,45$. При $x > 0,45$ полупроводник становится непрямозонным. На рис. 8, б приведена зависимость энергии от волнового вектора для нескольких значений x [30], из которой следует, что зона проводимости имеет два минимума. Те минимумы, которые расположены в точке Γ , являются прямыми, тогда как другие, расположенные на оси x , не прямые. Электроны в прямом минимуме зоны проводимости и дырки в максимуме валентной зоны обладают одинаковыми квазиимпульсами; электроны в непрямом минимуме имеют другое значение квазиимпульса. Для таких прямозонных полупроводников, как GaAs и $\text{GaAs}_{1-x}\text{P}_x$ ($x \leq 0,45$), квазиимпульс при междузонных переходах сохраняется, поэтому эти переходы характеризуются высокой степенью вероятности.

При этом энергия фотона приблизительно равна ширине запрещенной зоны полупроводника. В прямозонных материалах процесс излучательной рекомбинации является доминирующим. В то же время для $\text{GaAs}_{1-x}\text{P}_x$ при $x > 0,45$ и GaP , у которых запрещенная зона непрямая, вероятность междузонных переходов чрезвычайно мала, поскольку в этом случае для преобразования квазиимпульса при переходе требуется участие фононов или других факторов рассеяния. Поэтому для усиления излучательных процессов в непрямозонных полупроводниках, таких, например, как GaP , специально создаются рекомбинационные центры.

Эффективные центры излучательной рекомбинации в $\text{GaAs}_{1-x}\text{P}_x$ могут быть созданы путем внедрения специальных примесей, например азота [31]. Азот, введенный в полупроводник, замещает атомы фосфора в узлах решетки. Азот и фосфор имеют одинаковую внешнюю электронную структуру (оба относятся к V группе элементов периодической системы), а структуры их внутренних оболочек сильно различаются. Это приводит к возникновению вблизи зоны проводимости электронного уровня захвата. Полученный таким образом рекомбинационный центр называется изоэлектронным центром. В нормальном состоянии изоэлектронные центры нейтральны. В материале p -типа инжектированный электрон сначала захватывается на центр. Заряженный отрица-

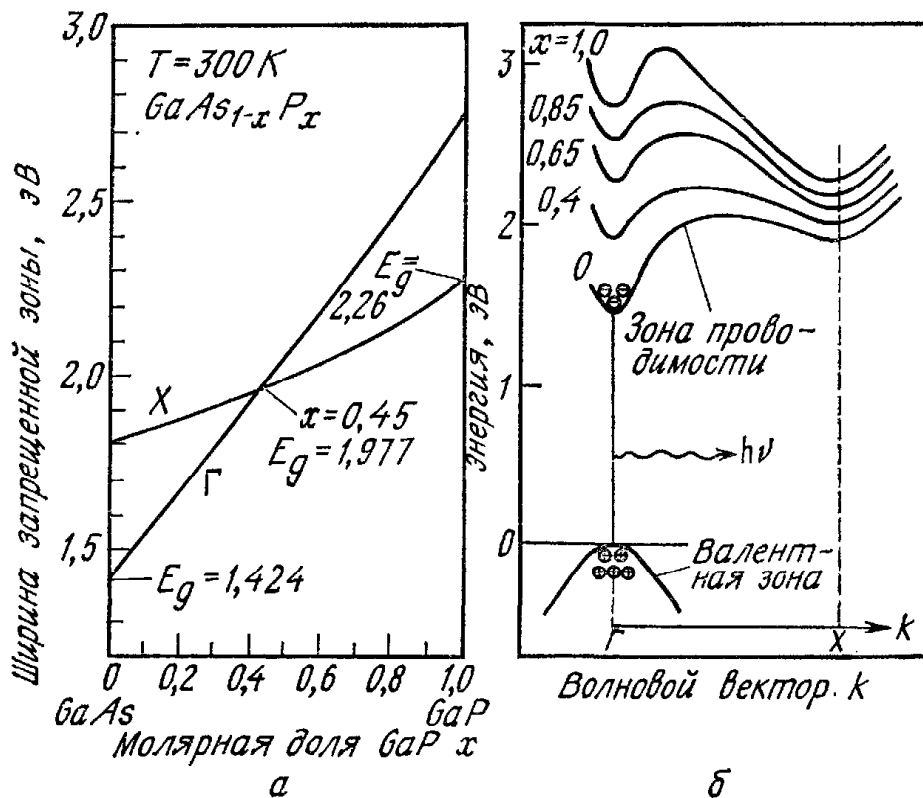


Рис. 8. Зависимость ширины энергетических зазоров (а) для прямого (Γ) и непрямого (X) минимумов зоны проводимости от состава $\text{GaAs}_{1-x}\text{P}_x$ [20] и энергетическая зонная структура $\text{GaAs}_{1-x}\text{P}_x$ (б). Значения состава соответствуют красному ($x = 0,4$), оранжевому (0,65), желтому (0,85) и зеленому свету (1,0) [30].

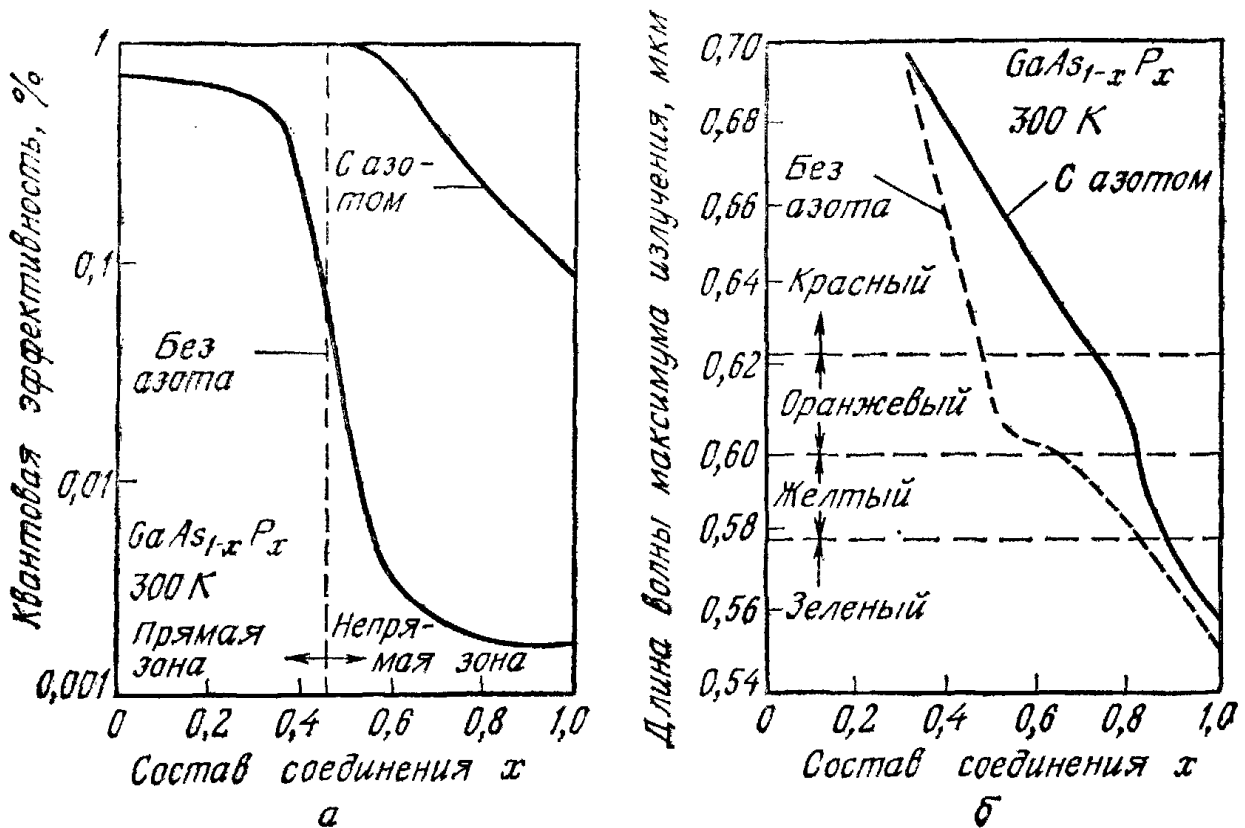


Рис. 9. Зависимость квантовой эффективности (а) и длины волны, соответствующей максимуму излучения (б), от состава соединения при наличии и в отсутствие изоэлектронной примеси азота [31].

тельно центр затем захватывает дырку из валентной зоны, формируя связанный экситон. Последующая аннигиляция этой электронно-дырочной пары приводит к рождению фотона с энергией, примерно равной разности между шириной запрещенной зоны и энергией связи центра. Так как захваченный электрон сильно локализован на центре, его импульс рассеивается. Таким образом обеспечивается преобразование квазиимпульса, вследствие чего вероятность прямого перехода существенно возрастает. В непрямозонных материалах, таких, как GaP, описанный механизм излучательной рекомбинации является преобладающим.

На рис. 9, а приведена зависимость квантовой эффективности от состава $GaAs_{1-x}P_x$ при наличии и в отсутствие изоэлектронной примеси азота [31]. Резкое уменьшение эффективности излучения $GaAs_{1-x}P_x$ без примеси азота в диапазоне $0,4 < x < 0,5$ обусловлено изменением механизма рекомбинации: при $x = 0,45$ преобладающими становятся непрямые переходы. Наличие азота обеспечивает более высокие значения эффективности при $x > 0,45$, однако последняя все же постепенно уменьшается с ростом x вследствие увеличения разности глубин прямой и непрямой запрещенных зон (рис. 8, б). Кроме того, длина волны максимума излучения для соединений, легированных азотом, смещается под влиянием энергии связи изоэлектронного центра (рис. 9, б).

Конструкции и характеристики светодиодов. Среди светодиодных структур основной является структура с плоской геометрией (рис. 10). Обычно прямозонные светодиоды (красное излучение) формируются на подложках GaAs (рис. 10, а), тогда как непрямозонные (оранжевое, желтое и зеленое излучения) — на подложках GaP (рис. 10, б). При использовании подложки GaAs на нее эпитаксиально наращивается переходный слой $\text{GaAs}_{1-x}\text{P}_x$ переменного состава с x , изменяющимся в пределах 0—0,4, а затем слой $\text{GaAs}_{1-x}\text{P}_x$ с постоянным составом. Переходная область ограничивает образование безызлучательных центров, обусловленных различием решеток. Фотоны, генерируемые в области перехода, испускаются во всех направлениях, однако наблюдателя достигает лишь та их часть, которая проходит через поверхность.

Уменьшение количества излучаемых светодиодом фотонов обусловлено поглощением в материале светодиода, потерями за счет отражения и потерями за счет полного внутреннего отражения. Потери, связанные с поглощением, весьма существенны в светодиодах на подложках GaAs (рис. 10, а), так как в этом случае

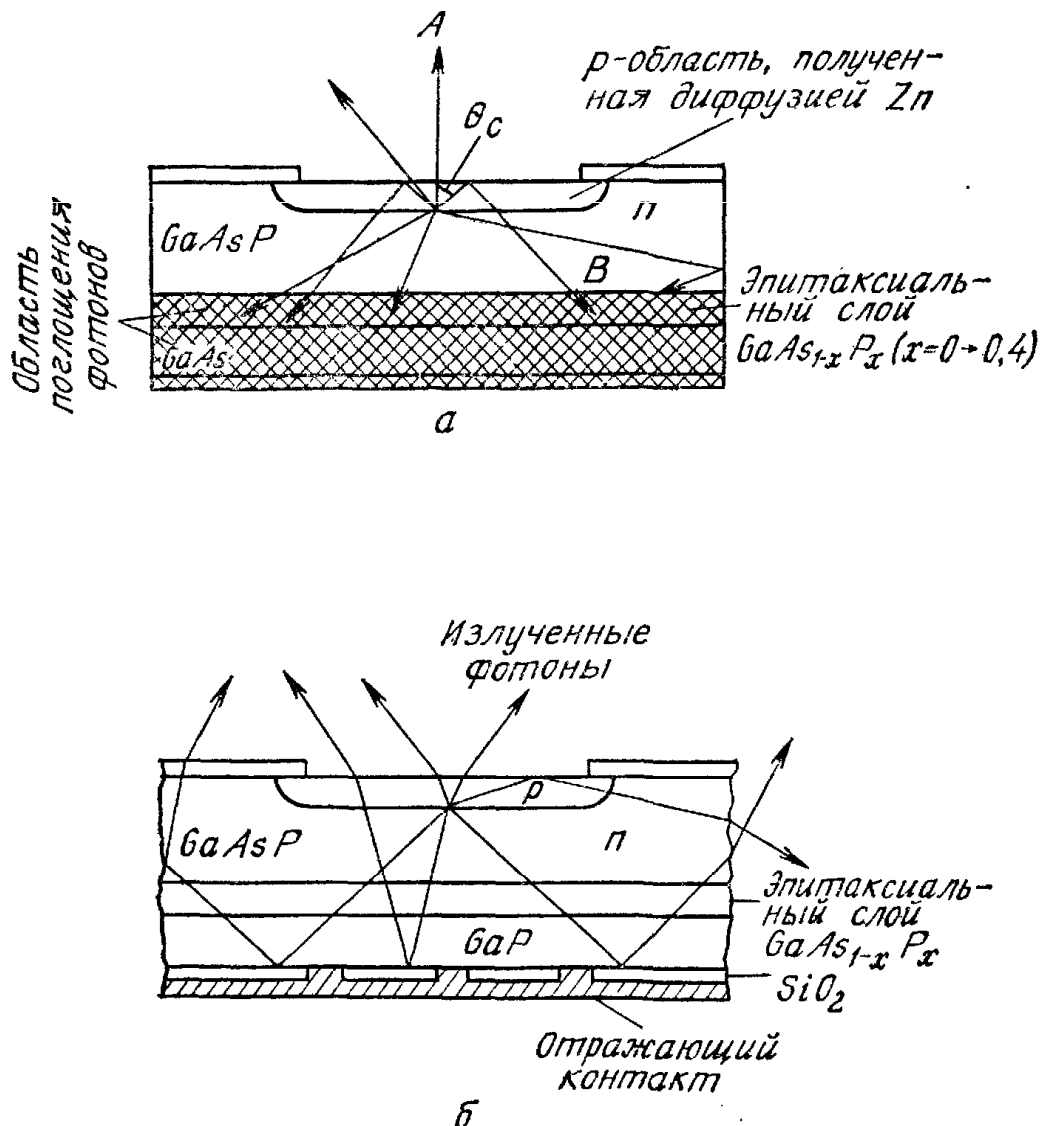


Рис. 10. Прохождение света, излучаемого переходом в структурах с непрозрачной (а) и прозрачной подложками (б) [3].

подложка непрозрачна для света и поглощает примерно 85 % фотонов, излучаемых переходом. В светодиодах на подложках GaP (рис. 10, б) фотоны, испускаемые в направлении тылового контакта, могут отражаться от него, при этом поглощение составляет $\sim 25\%$, и эффективность излучения может быть существенно увеличена. При переходе фотонов из среды с показателем преломления \bar{n}_2 (например, из GaAs с $\bar{n}_2 = 3,66$) в среду с показателем преломления \bar{n}_1 (например, в воздух с $\bar{n}_1 = 0,1$) часть света претерпевает отражение от границы раздела сред в обратном направлении. Эти потери называются потерями отражения. При нормальном падении света коэффициент отражения равен

$$R = \left(\frac{\bar{n}_2 - \bar{n}_1}{\bar{n}_2 + \bar{n}_1} \right)^2. \quad (11)$$

Третий фактор обусловлен полным внутренним отражением света, падающего на границу раздела под углом, больше критического угла θ_c , определяемого выражением

$$\theta_c = \sin^{-1}(\bar{n}_1/\bar{n}_2). \quad (12)$$

Для GaAs критический угол составляет $\sim 16^\circ$, а для GaP с $\bar{n}_2 = 3,45$ он равен $\sim 17^\circ$.

При плоской геометрии (рис. 10, а) луч А ослабляется вследствие поглощения, а также частично из-за потерь на отражение. Лучи В, которые приходят к границе раздела полупроводник — воздух под углами $\theta \geq \theta_c$, претерпевают полное внутреннее отражение. Полная эффективность преобразования электрического сигнала в оптический дается выражением [32]

$$\eta_F = \frac{q}{P} (1 - R) (1 - \cos \theta_c) \frac{\int \Phi(\lambda) (1 + R_1 e^{-2\alpha_1(\lambda)x_1}) e^{-\alpha_2(\lambda)x_2} d\lambda}{\int \Phi(\lambda) d\lambda}, \quad (13)$$

$$\eta_F \approx \frac{4\bar{n}_2\bar{n}_1}{(\bar{n}_2 + \bar{n}_1)^2} (1 - \cos \theta_c), \quad (14)$$

где P — мощность на входе; $4\bar{n}_2\bar{n}_1/(\bar{n}_2 + \bar{n}_1)^2$ — коэффициент передачи, равный $4\bar{n}_2/(\bar{n}_2 + 1)^2$ для границы раздела полупроводник — воздух; $(1 - \cos \theta_c)$ — телесный угол; $\Phi(\lambda)$ — скорость генерации фотонов в единицах фотон/(с·см²); R_1 — коэффициент отражения от тылового контакта; $\alpha(\lambda)$ и x — соответственно коэффициенты поглощения и толщина p - и n -областей прибора.

На рис. 11 показаны поперечные разрезы других светодиодов, которые имеют полусферическую, усеченно сферическую и параболическую геометрию [33]. Аналогичные выражения для эффективности могут быть получены также и для этих конфигураций. Основное отличие этих трех структур от структуры с плоской

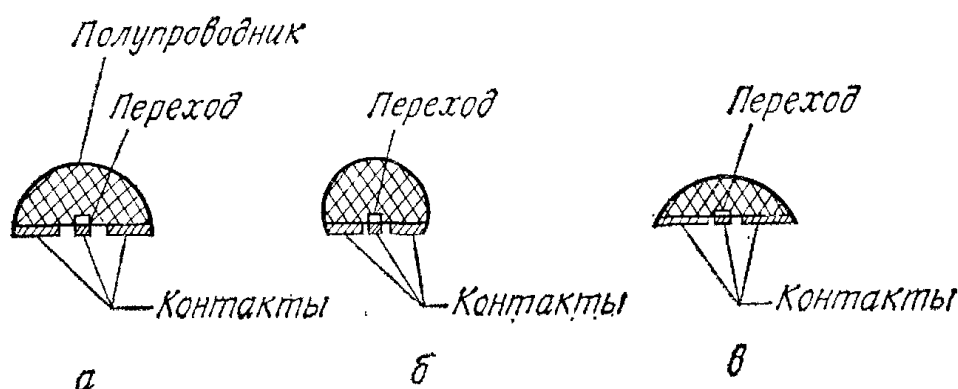


Рис. 11. Разрез трех светодиодов [33].

a — полусфера; *б* — усеченная сфера; *в* — параболоид.

геометрией (рис. 10, *a*) состоит в том, что телесный угол для них равен 1. Таким образом, отношение эффективностей равно

$$\frac{\eta}{\eta_F} \approx \frac{1}{1 - \cos \theta_c} = \frac{1}{1 - \sqrt{1 - 1/\bar{n}_2^2}}, \quad (15)$$

$$\eta/\eta_F = 2\bar{n}_2^2 \text{ при } \bar{n}_2 \gg 1. \quad (16)$$

Это означает, что для структур на GaP с $\bar{n}_2 = 3,45$ при данной геометрии можно ожидать увеличения эффективности на порядок величины (рис. 11). На рис. 12 приведены типичные диаграммы направленности для плоского, полусферического и параболического диодов [34]. Очевидно, что требуемая диаграмма направленности может быть получена за счет изменения геометрии прибора.

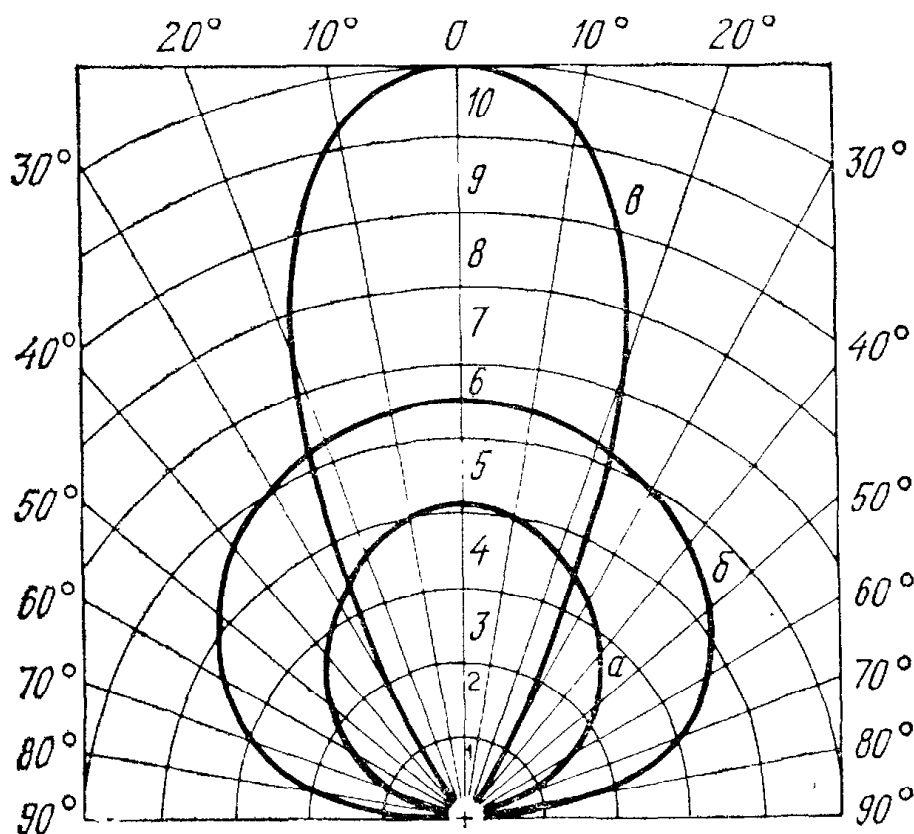


Рис. 12. Диаграмма направленности излучения светодиодов плоской (*a*), полусферической (*б*) и параболической (*в*) геометрии [34].

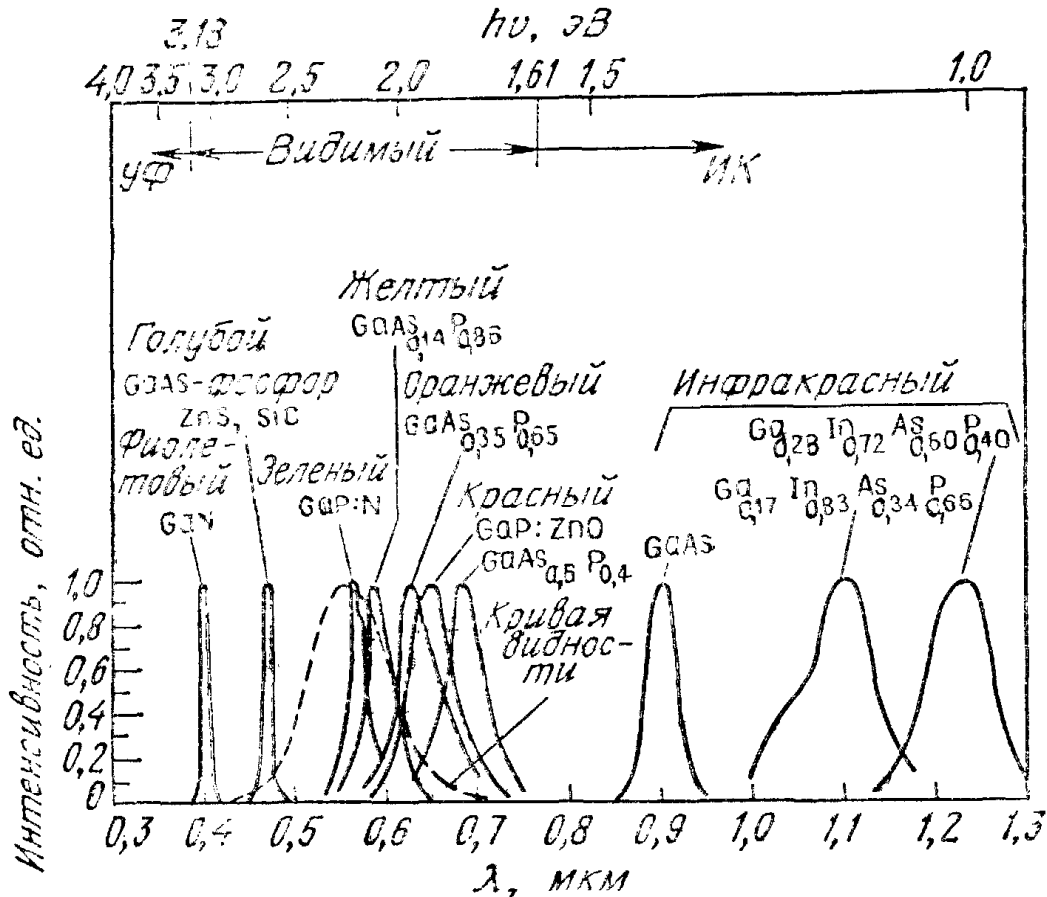


Рис. 13. Относительные спектры излучения различных светодиодов видимого и инфракрасного диапазонов.

На рис. 13 представлены относительные спектры излучения некоторых типичных светодиодов при комнатной температуре. Инфракрасные светодиоды из GaAs и соединений типа $A^{III}B^V$ будут рассмотрены в разделе 12.3.2. Красные светодиоды могут быть изготовлены из прямозонного соединения $GaAs_{0.6}P_{0.4}$, а также на основе GaP с примесью ZnO, которая, обладая таким же числом общих валентных электронов, как и пары соседних атомов галлий — фосфор, замещает их в решетке полупроводника [2]. ZnO является одной из примесей, формирующих изоэлектронные центры. Уровень рекомбинации, обусловленный этой примесью, расположен примерно на 0,3 эВ ниже края зоны проводимости GaP, в результате чего испускаются кванты с энергией 1,95 эВ. Светодиоды, излучающие оранжевый, желтый и зеленый свет [35], создаются на основе непрямозонных соединений $GaAs_{1-x}P_x$ или GaP, легированных изоэлектронной примесью азота. Светодиоды, излучающие голубой свет, могут быть изготовлены из соединений ZnS, SiC [36], а также в виде преобразователей инфракрасного излучения в видимое [37]. Рис. 14, а иллюстрирует основную структуру преобразователя, в котором инфракрасное излучение светодиода GaAs поглощается слоем фосфора с добавками ионов редкоземельных элементов, например иттербия (Yb^{3+}) или эрбия (Er^{3+}). Характеристики преобразования зависят от

процесса последовательного поглощения двух фотонов в инфракрасной области спектра, сопровождающегося испусканием одного фотона в видимой области спектра (рис. 14, б). На рис. 13 в фиолетовой части спектра расположен спектр излучения светодиода с $p-n$ -переходом на основе GaN. Отметим, что с увеличением длины волны λ_0 , соответствующей максимуму спектра излучения, увеличивается также и полуширина спектра. Это обусловлено тем, что ширина спектра спонтанного излучения пропорциональна λ_0^2 , как и следует из уравнения (2).

На рис. 13 приведена также характеристика чувствительности глаза (штриховая кривая) с максимумом на длине волны 0,555 мкм (зеленый свет). По мере смещения длины волны в красную область видимого спектра чувствительность глаза быстро падает. Спектральная характеристика чувствительности глаза должна учитываться при оценках эффективности визуального восприятия излучения от различных светодиодов с известными значениями энергии излучения. Для светодиодов с высокой эффективностью и широким спектром яркостный эквивалент излучения дается выражением [29]:

$$\text{Яркость} = L_0 \int V(\lambda) P(\lambda) d\lambda \quad [\text{лм/Вт}], \quad (17)$$

где L_0 — максимальное значение яркости, равное 680 лм/Вт; $V(\lambda)$ — функция относительной видности глаза; $P(\lambda)$ — спектр излучения. Например, как показано на рис. 15, максимум красного излучения диода на основе GaP — ZnO лежит на длине волны 0,69 мкм. Максимум кривой, соответствующей суперпозиции спектра излучения с кривой видности, смещается к значению $\lambda = 0,65$ мкм и соответствует яркостному эквиваленту, равному 15 лм/Вт.

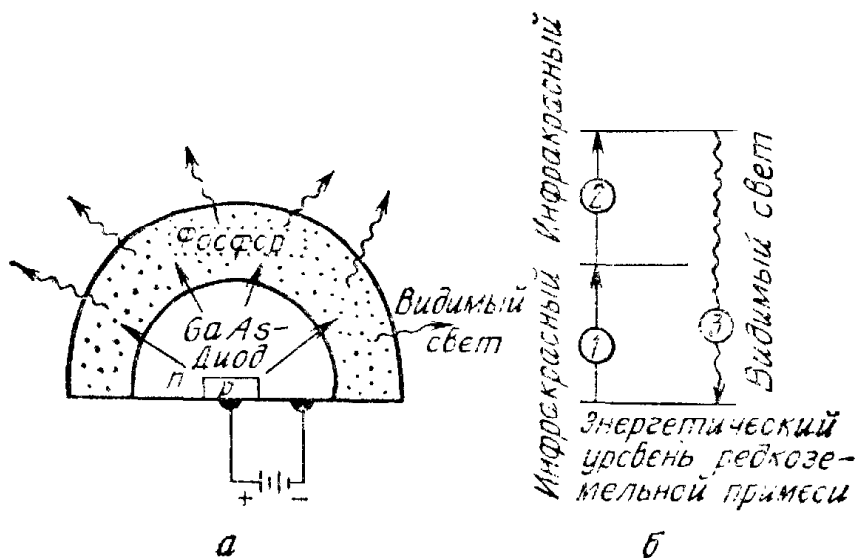


Рис. 14. Преобразователь, в котором излучение GaAs-диода возбуждает испускание света фосфором (а) и энергетические уровни преобразования (б) [37].

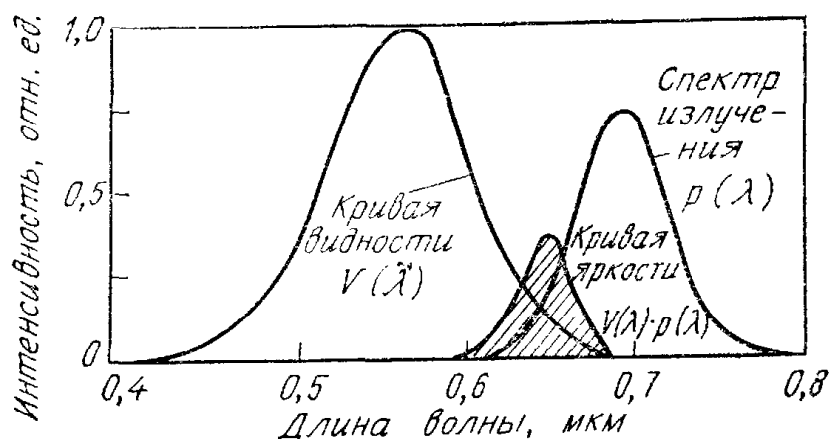


Рис. 15. Оценка яркостного эквивалента излучения [29].

Светодиоды, излучающие в видимой области спектра, могут использоваться в качестве индикаторов и дисплеев, а также в оптронах [3]. На рис. 16 приведены схемы конструкций различных светодиодных источников [2]. Источник содержит кристалл со светодиодом и пластмассовую линзу, которая обычно окрашивается с целью усиления контраста и фильтрации излучения. Источники, показанные на рис. 16, а и б, выполнены на базе обычных транзисторных и диодных корпусов. В источнике, изображенном на рис. 16, в, используется базовый держатель в виде металлических проводников, на которых размещены светодиод и последовательное сопротивление. Пластмассовая линза задает диаграмму направленности и угол зрения прибора. Конструкция корпуса на рис. 16, г удовлетворяет требованиям для размещения прозрачных светодиодных кристаллов, например, таких, как GaP, которые излучают через все пять своих граней.

На рис. 17 приведены основные форматы светодиодных дисплеев. Для отображения цифр от 0 до 9 обычно применяются 7-сегментные дисплеи или светодиодные матрицы с числом элементов 3×5 . В алфавитно-цифровых дисплеях (0 — 9 и A — Z) используются форматы, содержащие 14 сегментов или 5×7 элементов. Дисплеи можно изготавливать на базе интегральной технологии так же, как кремниевые монолитные схемы, или используя гибридную сборку, в ходе которой на отражателе формируются сегменты из отдельных светодиодных чипов.

Светодиоды могут использоваться в оптронах, которые обеспечивают электрическую развязку между входным или управляющим сигналом и выходом [2, 3]. На рис. 18, а показана схема оптрона, содержащего светодиод в качестве источника света и фотодиод в качестве детектора. Излучение, генерируемое светодиодом при подаче на него входного электрического сигнала, детектируется фотодиодом. Затем излучение вновь преобразуется в электрический сигнал в виде тока, который протекает через сопротивление нагрузки. Типичное значение коэффициента пре-

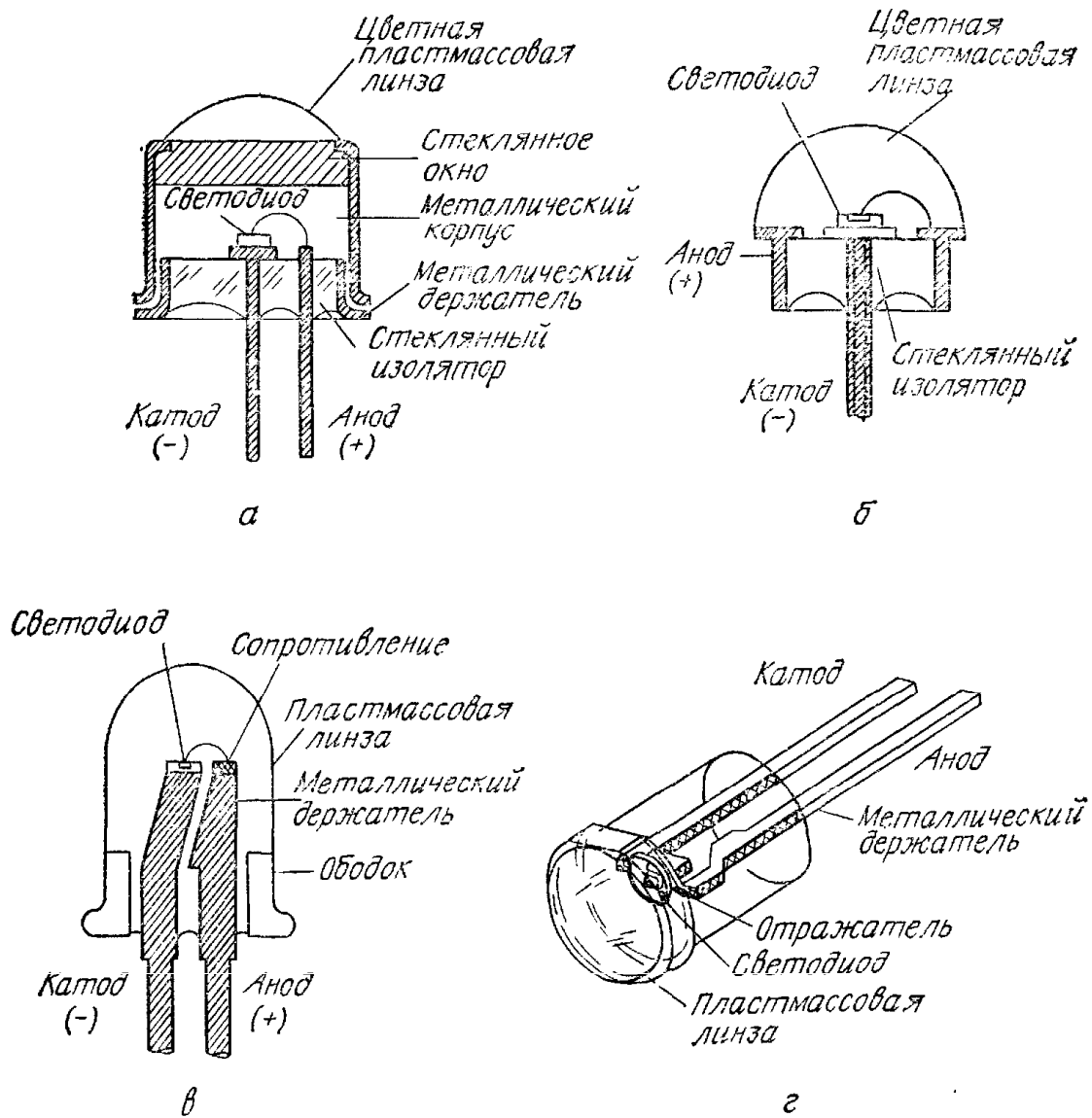


Рис. 16, Схемы конструкций некоторых светодиодных излучателей [2].

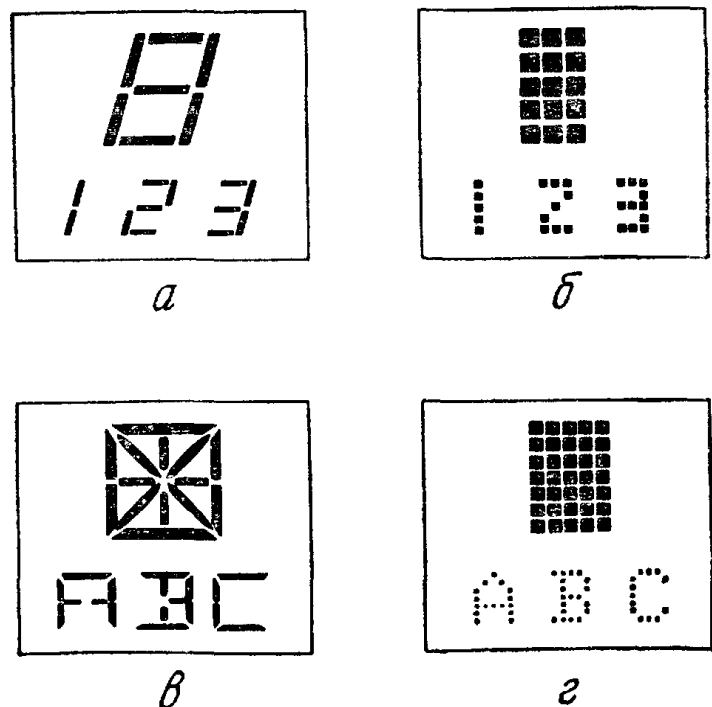


Рис. 17. Форматы цифровых и буквенных светодиодных дисплеев.

а — 7-сегментный (цифровой); б — матричный 3 × 5 (цифровой); в — 14-сегментный (алфавитно-цифровой); г — матричный (алфавитно-цифровой).

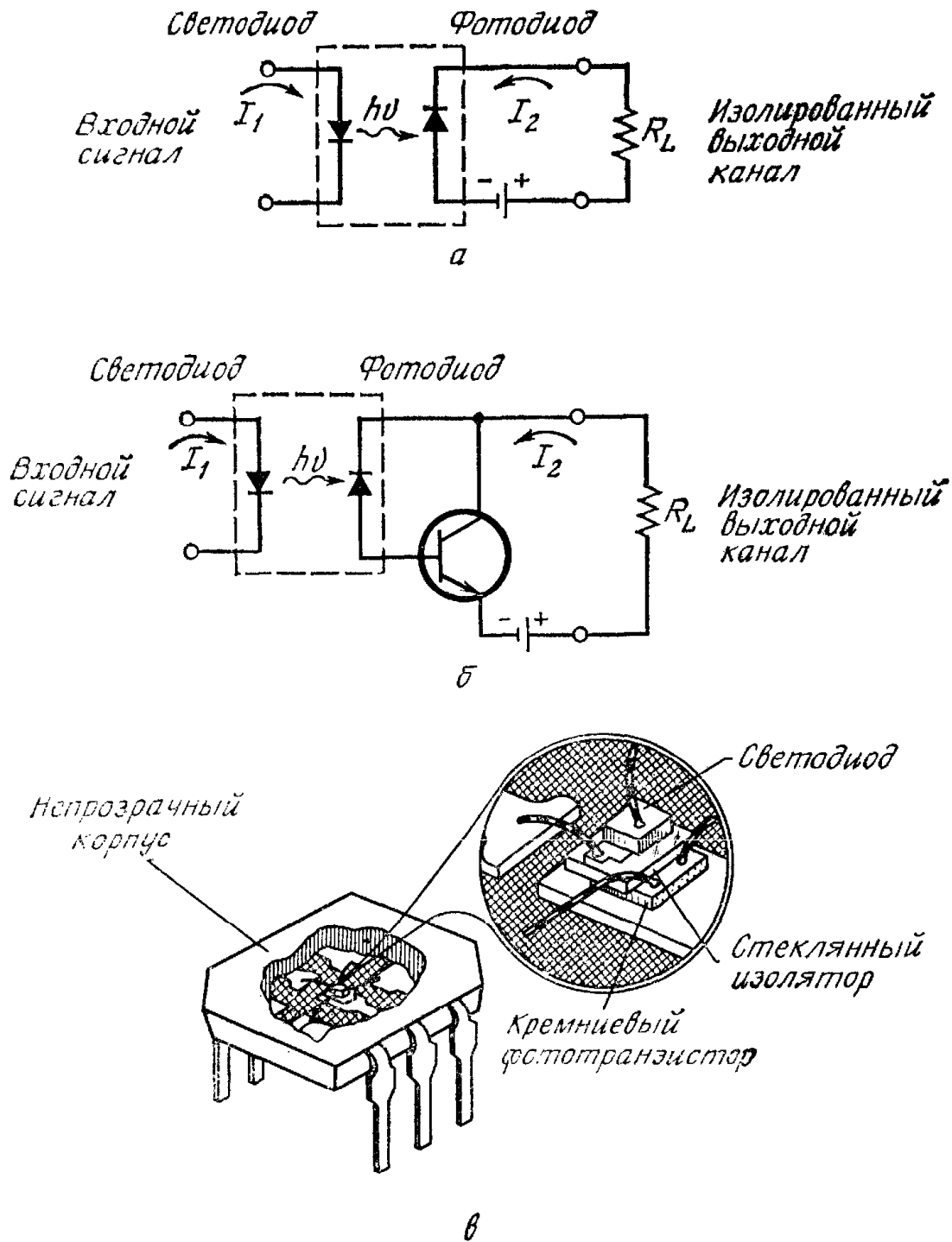


Рис. 18. Оптроны [2].

a — схема; *б* — оптрон с высоким коэффициентом преобразования; *в* — оптрон, расположенный на одной контактной площадке.

образования по току I_2/I_1 составляет $\sim 10^{-3}$. Для усиления коэффициента преобразования в оптронах в качестве детекторов используются фототранзисторы, при этом отношение токов может быть повышено до 0,1 : 1. На рис. 18, *в* показана оптронная структура, размещенная в одном корпусе. В таких приборах осуществляется оптическая передача сигналов со скоростью света и обеспечивается электрическая изоляция, так как отсутствует обратная связь с выхода на вход.

12.3.2. Инфракрасные светодиоды

Инфракрасные светодиоды, как и светодиоды видимого излучения, имеют различную конфигурацию (рис. 10 и 11). На рис. 13 среди других спектров приведен также спектр излучения светодиода на основе GaAs. Этот инфракрасный источник в настоящее время обладает наибольшей эффективностью электролюминесценции в основном благодаря тому, что среди всех прямозонных полупроводников GaAs является технологически наиболее освоенным. Одним из важных применений диодов на основе GaAs являются источники излучения в оптронах. Для изготовления инфракрасных светодиодов используются многие другие полупроводники, имеющие запрещенную зону шириной менее 1,5 эВ. К ним относятся твердые растворы, в состав которых входят три или четыре элемента III и V групп периодической системы. На рис. 13 приведены спектры излучения инфракрасных светодиодов из четырехэлементных полупроводников $Ga_xIn_{1-x}As_{1-y}P_y$ [38]. Структура запрещенной зоны и постоянные решетки этих полупроводников рассмотрены в разд. 12.4.

Инфракрасные светодиоды являются перспективными источниками для волоконно-оптических линий связи. Светодиоды по сравнению с лазерами имеют как недостатки, так и преимущества. К последним относятся высокая рабочая температура, меньшее влияние температуры на мощность излучения, а также простота конструкции и схемы питания. К недостаткам относятся меньшая яркость, более низкие частоты модуляции и большая спектральная ширина линии излучения, типичное значение которой составляет 100—500 Å, тогда как лазеры характеризуются шириной линии 0,1—1 Å.

Различают два основных типа светодиодов, обеспечивающих ввод излучения в оптические волокна малого диаметра: светодиоды с излучающей поверхностью и светодиоды с излучающей гранью [39, 40]. В поверхностных излучателях (рис. 19, а) излучающая область перехода ограничена слоем окисла, а диаметр контакта обычно составляет 15—100 мкм. Для уменьшения потерь поглощения и обеспечения плотного контакта излучающей поверхности с торцом волокна слои полупроводника, через которые проходит излучение, должны быть очень тонкими (10—15 мкм). В гетеропереходах (например, GaAs — $Al_xGa_{1-x}As$) мощность излучения на выходе может быть повышена за счет эффекта ограничения носителей, который возникает при окружении слоя излучательной рекомбинации (например, GaAs) областью полупроводника с более широкой запрещенной зоной (например, $Al_xGa_{1-x}As$). Подробно этот эффект рассмотрен в последующих разделах. Гетеропереход может также служить в качестве окна,

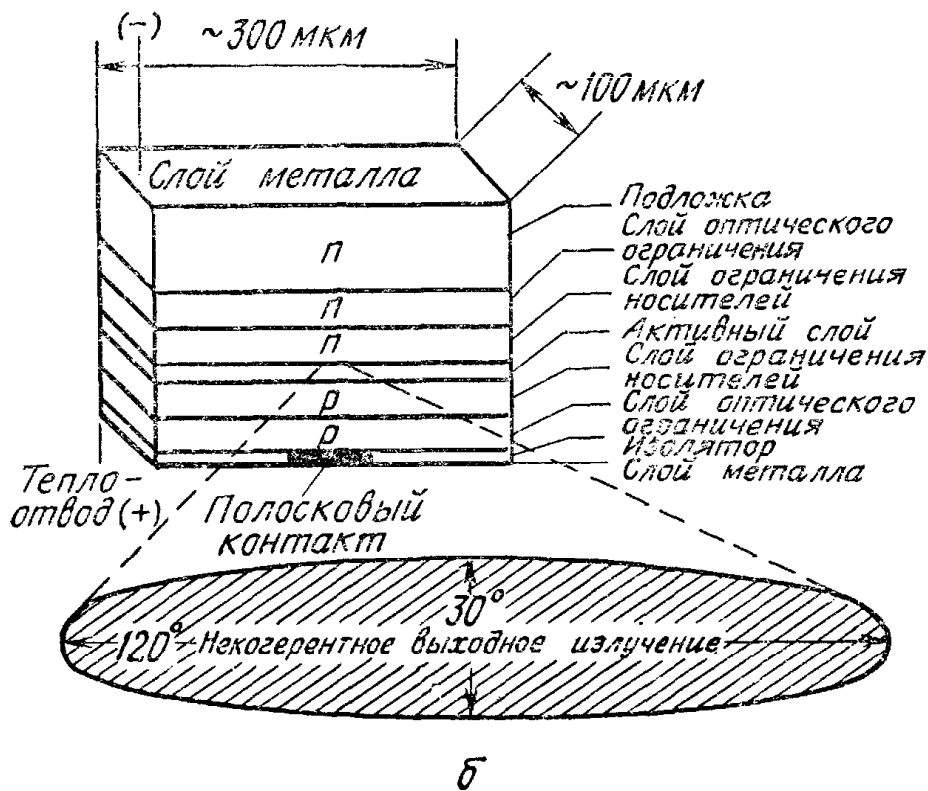
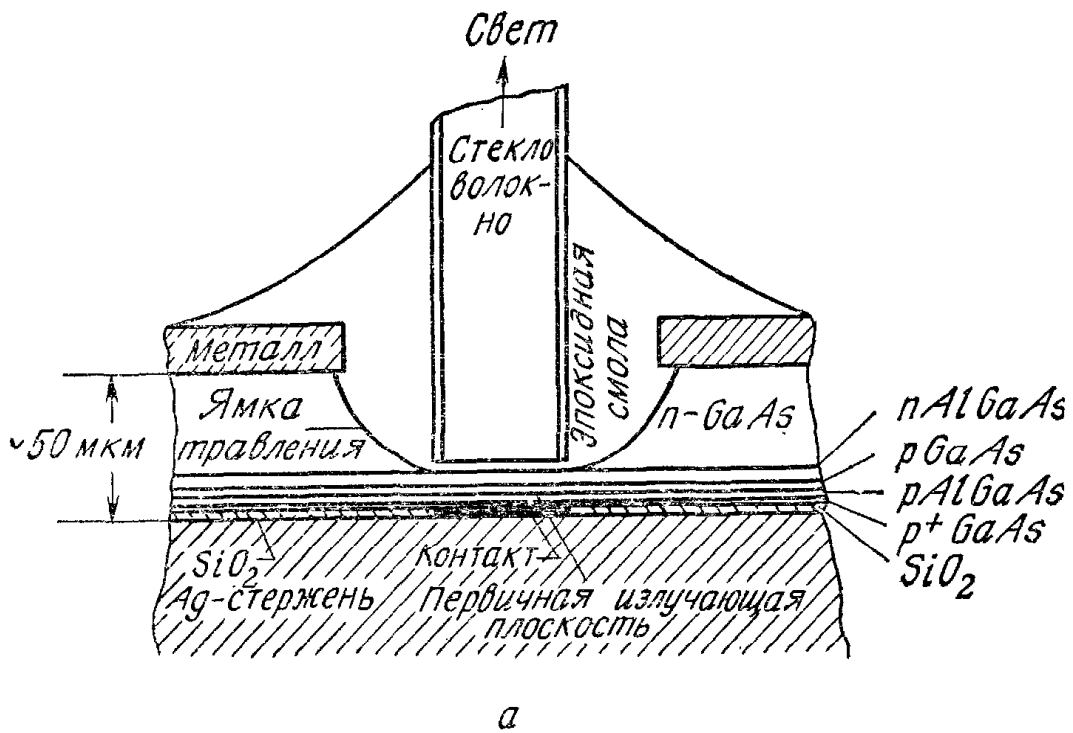


Рис. 19. Поверхностно-излучающий светодиод (а) на основе двойной гетероструктуры AlGaAs, имеющий малую площадь и высокую излучающую способность и совмещенный со стекловолокном [39], и светодиод с излучающей гранью на основе двойной гетероструктуры (б) [40].

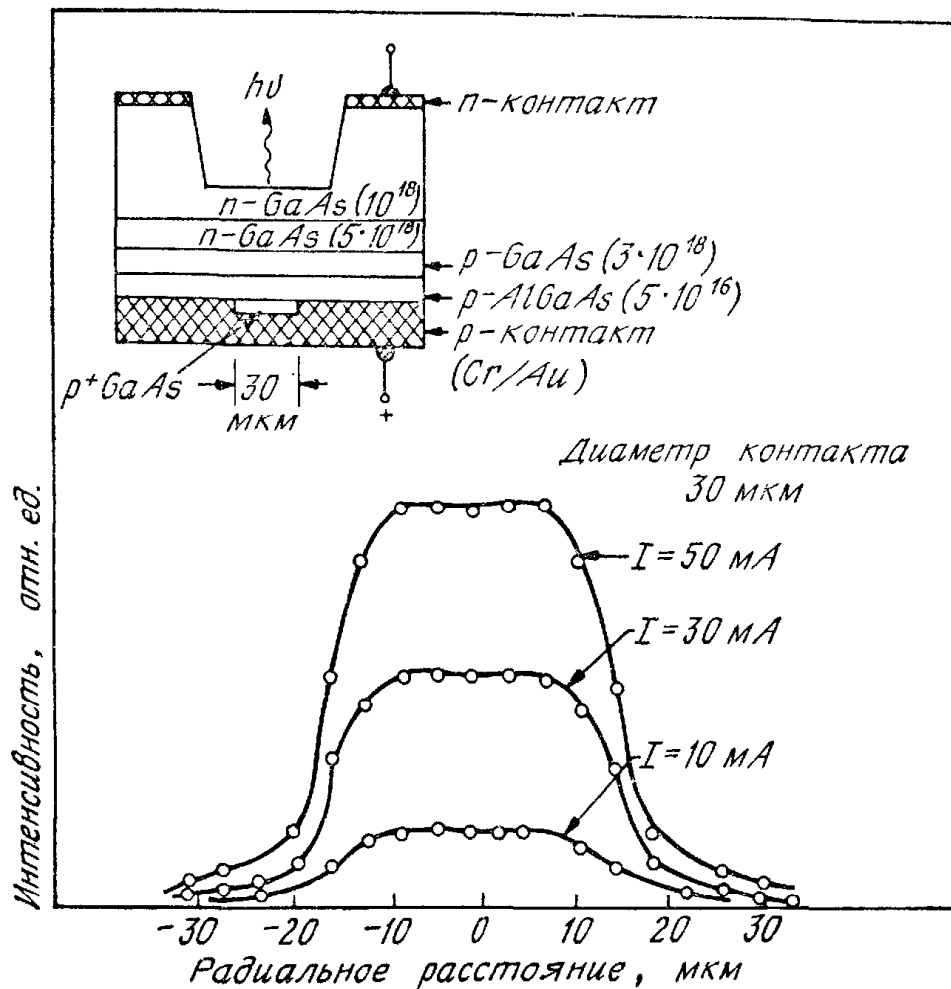


Рис. 20. Распределение интенсивности света вблизи излучающей поверхности светодиода. На вставке показана структура светодиода [41].

которое не поглощает излучение, возникающее в полупроводнике с меньшей запрещенной зоной.

При выводе через грань (рис. 19, б) излучение может быть сосредоточено в относительно направленном луче, что повышает эффективность связи светодиода с волокнами, имеющими малый приемный угол. Пространственное распределение излучения при этом аналогично распределению для лазеров на гетеропереходах, которые рассматриваются в разд. 12.4.

Вставка на рис. 20 иллюстрирует структуру поверхностного излучателя, в котором область излучения ограничена областью с высоким контактным сопротивлением [41]. Диаметр омического контакта к слою $p^+\text{-GaAs}$ составляет 30 мкм. Слой металлизации Cr/Au вокруг области омического контакта имеет высокое контактное сопротивление со слоем $p\text{-AlGaAs}$ (концентрация легирования $5 \cdot 10^{16} \text{ см}^{-3}$). Поэтому основная часть тока протекает через область омического контакта. На рис. 20 приведено распределение интенсивности света вблизи излучающей поверхности. Видно, что диаметр области излучения ограничен размером 30 мкм, задаваемым конфигурацией контакта.

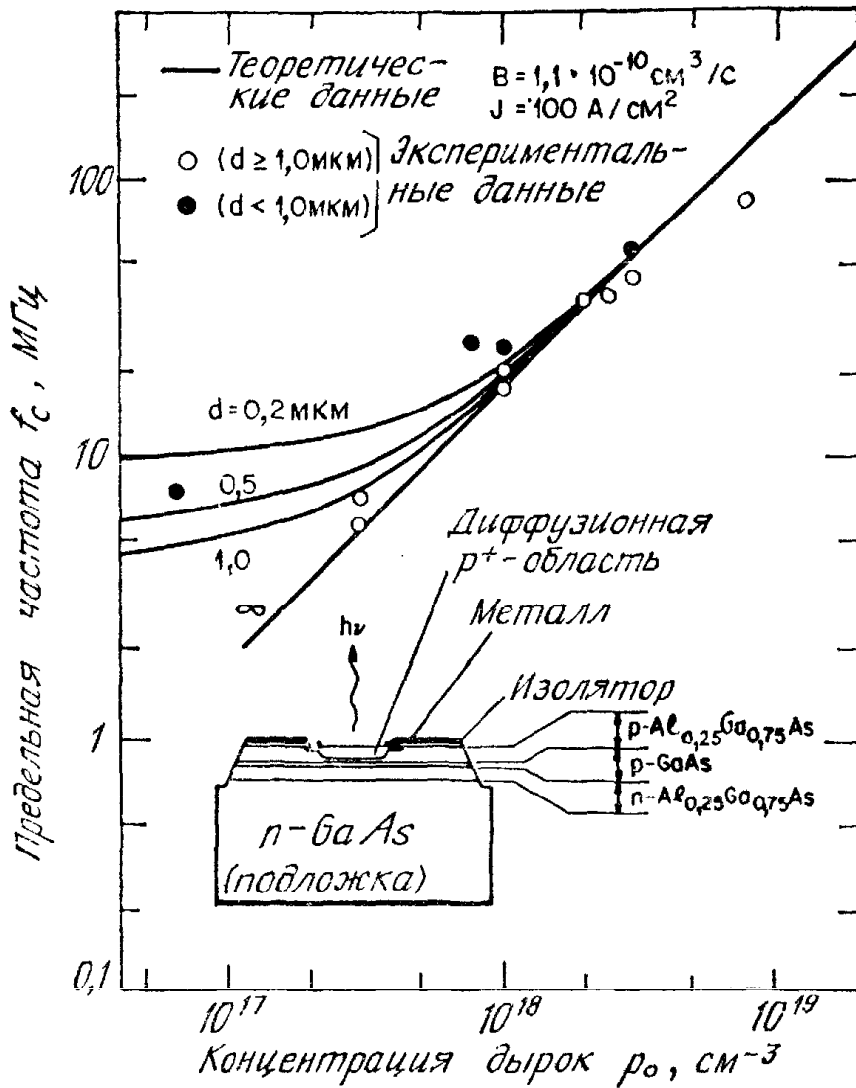


Рис. 21. Зависимость предельной частоты светодиода от концентрации дырок в активном слое [42].

Другим важным параметром, который должен учитываться при конструировании светодиодов для оптических систем связи, является диапазон рабочих частот. Как следует из уравнений (7) и (10), квантовая эффективность люминесценции и полное время жизни связаны с временем жизни излучательной рекомбинации. При внешнем возбуждении скорость полной излучательной рекомбинации определяется выражением [41а]

$$R_r = Gn_p/n_0p_0, \tag{18}$$

где G — скорость полной термической генерации. Обычно это соотношение записывается следующим образом:

$$R_r = Bnp, \tag{19}$$

где B — константа излучательной рекомбинации, равная G/n_0p_0 . При достаточно низких уровнях возбуждения, таких, что в материале p -типа $p \approx p_0$, время жизни излучательной рекомбинации становится равным

$$\tau_r = (n - n_0)/R_r \approx 1/Bp_0, \tag{20}$$

а для материала n -типа, когда $n \approx n_0$,

$$\tau_r = (\rho - \rho_0)/R_r \approx 1/Bn_0. \quad (21)$$

Предельная частота светодиода

$$f_c = \frac{1}{2\pi\tau}. \quad (22)$$

Если $\tau_r \leq \tau_{nr}$, то τ в уравнении (22) становится близким к τ_r . Таким образом, как следует из уравнений (20) и (21), при увеличении концентрации τ_r уменьшается, а f_c растет. Результаты теоретического расчета и экспериментальные данные измерения зависимости предельной частоты от концентрации дырок для светодиода на основе гетероструктуры с выводом излучения через поверхность приведены на рис. 21 [42]. Видно, что экспериментально измеренная предельная частота возрастает с концентрацией p_0 , что находится в соответствии с уравнениями (10) и (20). Для достижения высоких значений f_c необходимо уменьшать толщину рекомбинационной области и увеличивать концентрацию носителей.

12.4. ФИЗИКА ПОЛУПРОВОДНИКОВЫХ ЛАЗЕРОВ

12.4.1. Общие замечания

Полупроводниковые лазеры, подобно другим лазерам (таким, как рубиновый лазер или же лазер на смеси He — Ne), испускают излучение, когерентное в пространстве и во времени. Это означает, что излучение лазера высокомонохроматично (имеет узкую полосу спектра) и создает строго направленный луч света. Вместе с тем по ряду важных характеристик полупроводниковые лазеры существенно отличаются от лазеров других типов:

1. В обычных лазерах квантовые переходы происходят между дискретными энергетическими уровнями, тогда как в полупроводниковых лазерах переходы обусловлены зонной структурой материала.

2. Полупроводниковые лазеры имеют очень малые размеры ($\sim 0,1$ мм в длину), и, так как активная область у них очень узкая (~ 1 мкм и меньше), расхождение лазерного луча значительно больше, чем у обычного лазера.

3. Пространственные и спектральные характеристики излучения полупроводникового лазера сильно зависят от свойств материала, из которого сделан переход (таких свойств, как структура запрещенной зоны и коэффициент преломления).

4. В лазере с p — n -переходом лазерное излучение возникает непосредственно под действием тока, протекающего через прямо-смещенный диод. В результате система в целом очень эффективна, поскольку позволяет легко осуществлять модуляцию излучения



Рис. 22. Зависимость потерь от длины волны излучения в кварцевом волокне. Обозначены три значения длины волны, представляющие наибольший интерес [43].

за счет модуляции тока. Так как полупроводниковые лазеры характеризуются очень малыми временами стимулированного излучения, модуляция может проводиться на высоких частотах.

Благодаря малым размерам и возможности высокочастотной модуляции полупроводниковый лазер представляет собой один из перспективных источников излучения для волоконно-оптических систем связи. На рис. 22 приведена зависимость потерь от длины волны для экспериментальных оптических волокон [43]. На графике выделены три значения длины волны, которые представляют практический интерес. В интервале длин волн вблизи 0,9 мкм в качестве источников излучения используются гетеролазеры на основе $\text{GaAs} - \text{Al}_x\text{Ga}_{1-x}\text{As}$, а в качестве фотодетекторов — Si-фотодиоды или Si-фотодиоды с лавинным умножением. Вблизи длины волны 1,3 мкм волокно имеет низкие потери (0,6 дБ/км) и слабую дисперсию, а в окрестности длины волны 1,55 мкм потери достигают минимального значения 0,2 дБ/км. Для этих двух длин волн в качестве источников излучения могут использоваться лазеры на основе четверных соединений типа $A^{\text{III}}B^{\text{V}}$, таких, как $\text{Ga}_x\text{In}_{1-x}\text{As}_y\text{P}_{1-y} - \text{InP}$, а в качестве фотодетекторов — германиевые лавинные фотодиоды, а также фотодиоды на тройных и четверных соединениях [44]. Волоконно-оптические линии связи для больших значений длин волн рассмотрены в работе [45]. Для создания таких систем необходимы волокна с исключительно низкими потерями, а также эффективные источники излучения (например, светодиоды $\text{PbS}_{0,1}\text{Se}_{0,9}$ [77]) и чувствительные фотодетекторы.

Полупроводниковые материалы. Перечень полупроводниковых материалов, проявляющих лазерные свойства, постоянно растет.

В настоящее время фактически все лазерные полупроводники обладают прямой запрещенной зоной. Это обусловлено тем, что излучательные переходы в прямозонных полупроводниках представляют собой процесс первого порядка (импульс сохраняется автоматически) и вероятность переходов высока. В полупроводниках с непрямой зоной излучательная рекомбинация выступает как процесс второго порядка (т. е. сохранение импульса и энергии при переходах обеспечивается участием фононов или другими факторами рассеяния), так что вероятность излучательных переходов существенно ниже. Кроме того, в непрямозонных полупроводниках при увеличении степени возбуждения потери, связанные с поглощением излучения на инжектируемых свободных носителях, возрастают быстрее, чем усиление [12].

На рис. 23 представлен диапазон длин волн лазерного излучения, включающий спектральные интервалы излучения различных полупроводников и охватывающий области спектра от ультрафиолетовой до инфракрасной [20]. В связи с тем, что GaAs стал первым лазерным материалом, наиболее интенсивное исследование и развитие получили родственные ему твердые растворы элементов III и V групп. Такие соединения типа $A^{IV}B^V$, как PbS, PbTe, PbSe и родственные им твердые растворы, также обладают

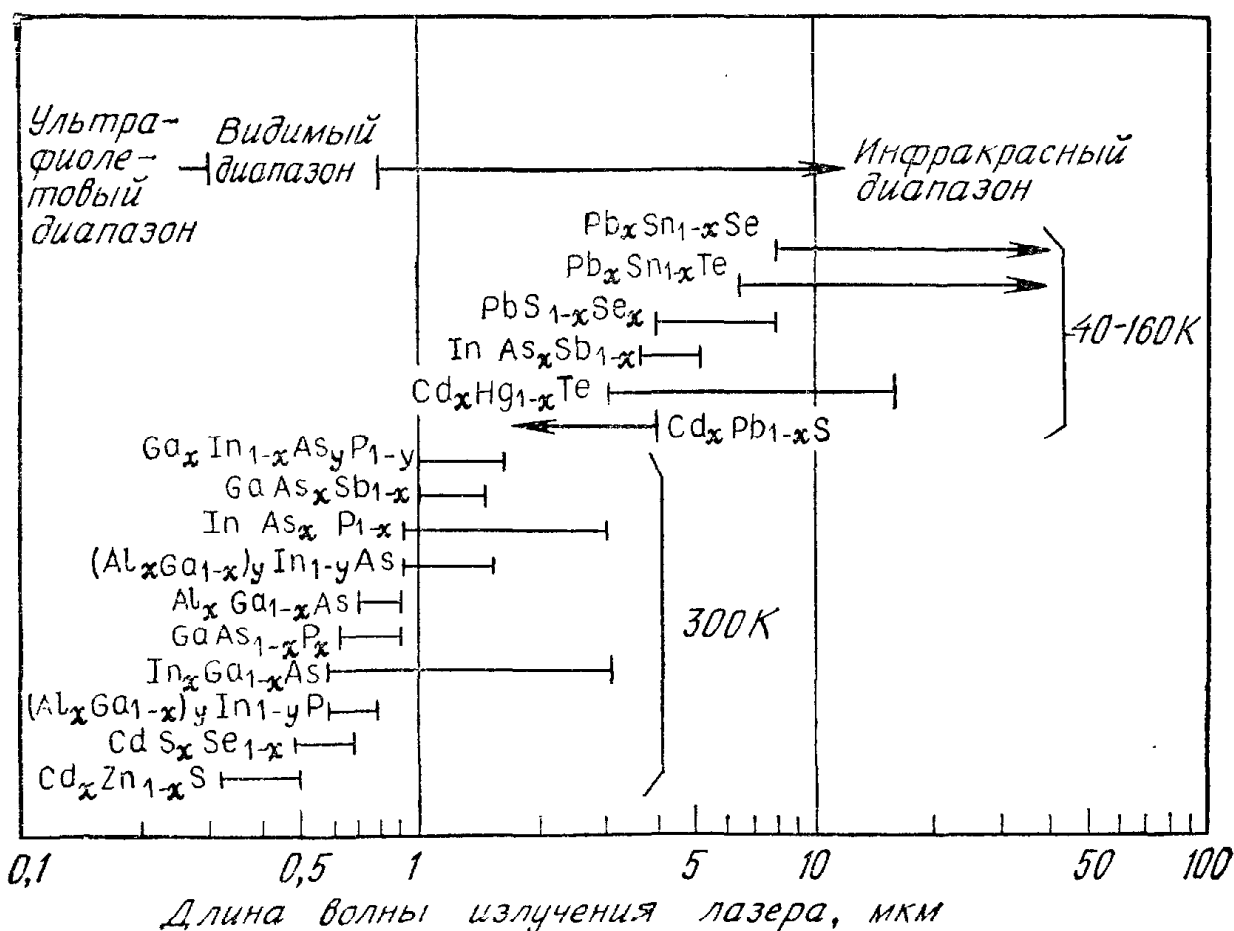


Рис. 23. Интервалы длин волн, для которых имеются или могут быть созданы гетеролазеры на соединениях $A^{III}B^{IV}$ и $A^{IV}B^{VI}$ [20].

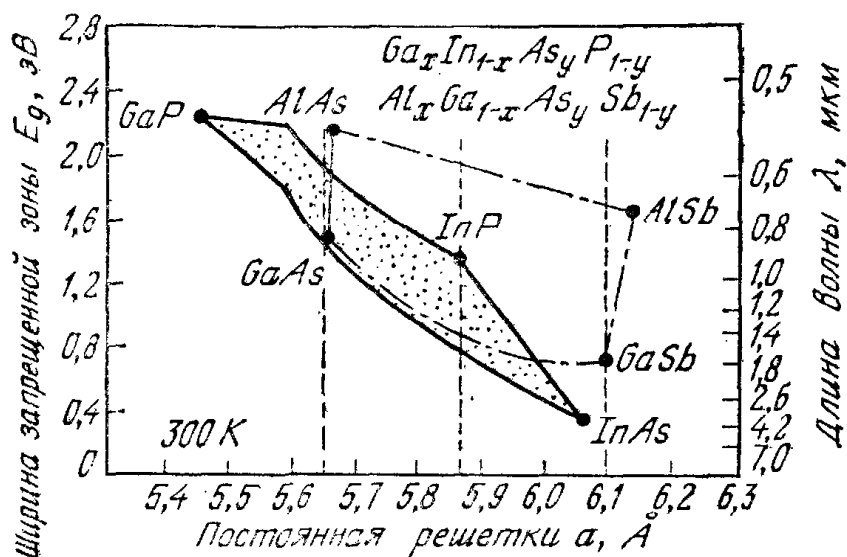


Рис. 24. Ширина запрещенной зоны и постоянная решетки для твердых растворов типа $A^{III}B^{IV}$ [20].

лазерными свойствами. В этих соединениях, относящихся к прямозонным материалам, энергетические экстремумы в зоне Бриллюэна локализованы вдоль направления $\langle 111 \rangle$ [46] в отличие от GaAs, у которого экстремумы расположены в центре зоны.

В этом разделе мы рассмотрим основы физики лазеров, причем в качестве иллюстраций будем пользоваться в основном примерами, относящимися к приборам на основе полупроводниковых соединений типа $A^{III}B^V$. Если узлы решетки, соответствующие элементам III группы, случайным образом заняты различными атомами элементов этой группы или подобная ситуация реализуется в узлах решетки элемента V группы, то такие соединения относятся к кристаллическим твердым растворам. При их описании часто используются обозначения типа $A_x B_{1-x} C$ и $A_x B_{1-x} C_y D_{1-y}$ соответственно для тройных и четверных соединений, где A и B — элементы III группы, а C и D — элементы V группы. В пределах каждой группы элементы могут быть расположены в алфавитном порядке или в соответствии с их атомным номером. В этой главе мы будем пользоваться алфавитным порядком расположения элементов внутри групп.

К наиболее важным из сложных соединений относятся твердые растворы $Al_x Ga_{1-x} As_y Sb_{1-y}$ и $Ga_x In_{1-x} As_y P_{1-y}$ [20]. На рис. 24 приведена диаграмма значений ширины запрещенной зоны и постоянной решетки для бинарных полупроводников типа $A^{III}B^V$ и их промежуточных тройных и четверных соединений. Для создания гетероструктур с пренебрежимо малой плотностью состояний на границе раздела необходимы пары полупроводников с близкими постоянными решетки. При использовании GaAs ($a = 5,6533 \text{ \AA}$) в качестве подложки для гетероструктуры с тройным соединением $Al_x Ga_{1-x} As$ можно достичь

рассогласования решеток менее 0,1 %. Аналогично подложка из InP ($a = 5,8686 \text{ \AA}$) может иметь почти идеальное согласование постоянных решетки с четверным соединением $\text{Ga}_x\text{In}_{1-x}\text{As}_y\text{P}_{1-y}$, что обозначено вертикальными линиями на рис. 24.

Стимулированное излучение. Как говорилось выше, работа лазера связана с тремя основными процессами, обусловленными переходом носителей: поглощением, спонтанной эмиссией и стимулированным излучением. Для описания этих процессов мы будем использовать простую модель [47]. Рассмотрим два энергетических уровня E_1 и E_2 , один из которых E_1 характеризует основное, а другой E_2 — возбужденное состояние атома (рис. 25). Любой переход между этими состояниями сопровождается испусканием или поглощением фотона с частотой ν_{12} , определяемой из соотношения $h\nu_{12} = E_2 - E_1$, где h — постоянная Планка. При обычных температурах большинство атомов находится в основном состоянии. Эта ситуация нарушается в результате воздействия на систему фотона с энергией, равной $h\nu_{12}$. Атом в состоянии E_1 поглощает фотон и переходит в возбужденное состояние E_2 . Это и составляет процесс поглощения излучения (рис. 25, а). Возбужденное состояние является нестабильным и через короткий промежуток времени без какого-либо внешнего воздействия атом переходит в основное состояние, испуская фотон с энергией $h\nu_{12}$. Этот процесс называется спонтанной эмиссией (рис. 26, б). Время жизни, связанное со спонтанной эмиссией (т. е. среднее время возбужденного состояния), может изменяться в широком диапазоне, обычно в пределах 10^{-9} — 10^{-3} с, в зависимости от параметров полупроводника, таких, как структура зон (прямая или непрямая) и плотность рекомбинационных центров. Столкновение фотона, обладающего энергией $h\nu_{12}$, с атомом, находящимся в возбужденном состоянии, стимулирует мгновенный переход атома в основное состояние с испусканием фотона с энергией $h\nu_{12}$ и фазой, соответствующей фазе падающего излучения. Этот процесс называется стимулированным излучением (рис. 25, в).

Лазеры на основе двойных гетероструктур. На рис. 26 показана базовая структура лазера с p — n -переходом. Две боковые грани структуры скальваются или полируются перпендикулярно плоскости перехода. Две другие грани делаются шероховатыми для того, чтобы исключить излучение в направлениях, не совпадающих с главным. Такая структура называется резонатором Фабри — Перо. Смещение лазерного диода в прямом направлении вызывает протекание тока. Вначале, при низких значениях тока, возникает спонтанное излучение, распространяющееся во всех направлениях. При увеличении смещения ток достигает порогового значения, при котором создаются условия для стимулированного излучения, и p — n -переход испускает монохроматичный и строго направленный луч света (рис. 26).

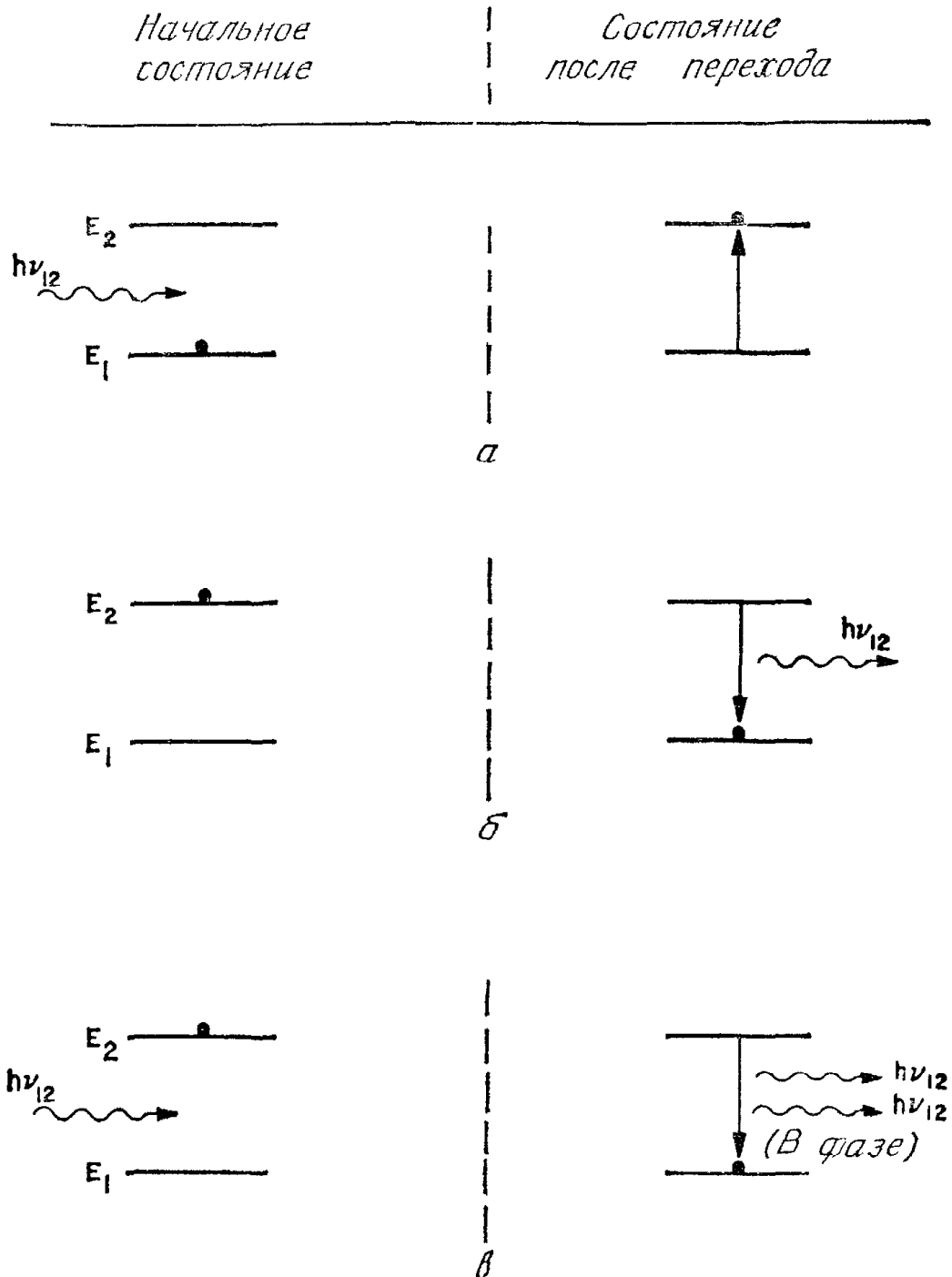


Рис. 25. Три основных типа переходов между двумя энергетическими уровнями E_1 и E_2 . Черными кружками обозначены состояния атома [47].
 а — поглощение; б — спонтанное излучение; в — стимулированное излучение.

Для гомоструктур (например, $p-n$ -переходов на основе GaAs) пороговая плотность тока быстро увеличивается с ростом температуры. При комнатной температуре типичное значение J_{th} (полученное при измерении в импульсном режиме) составляет $\sim 5,0 \cdot 10^4$ А/см². Такая большая плотность тока создает серьезные трудности для реализации режима непрерывной генерации при 300 К.

С целью уменьшения пороговой плотности тока были предложены, а затем с помощью эпитаксиальной технологии реализованы

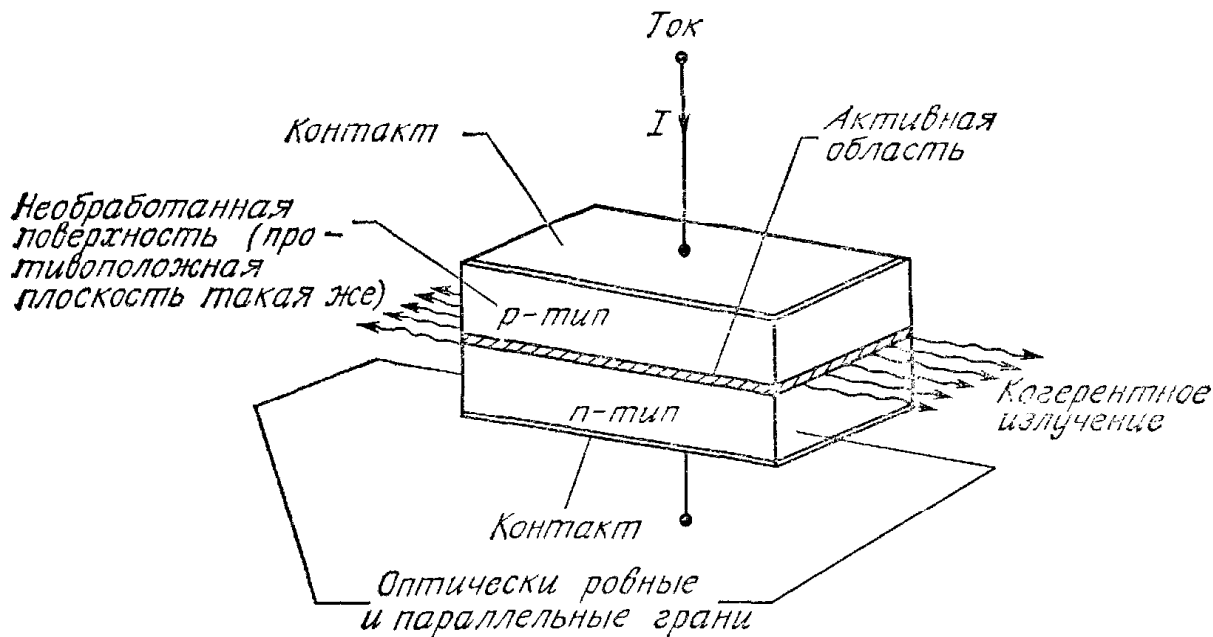


Рис. 26. Основная структура лазера с $p-n$ -переходом в виде резонатора Фабри — Перо.

лазеры на гетероструктурах. На рис. 27 схематически изображены гомоструктура, структуры с одним и двумя гетеропереходами, их зонные диаграммы в условиях прямого смещения [48], а также профили распределения показателя преломления и интенсивности генерируемого в $p-n$ -переходе света. Как видно из рис. 27, в структуре с двумя гетеропереходами носители сосредоточены внутри активной области d , ограниченной с обеих сторон потенциальными барьерами; излучение также ограничено этой

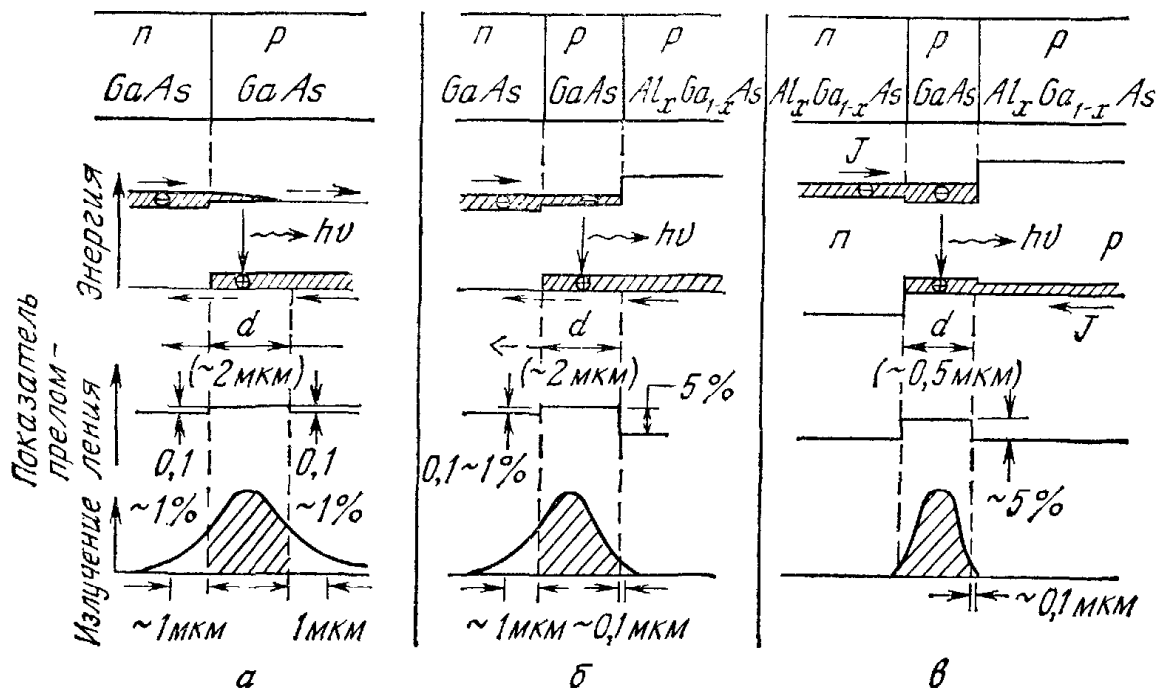


Рис. 27. Сравнительные характеристики некоторых лазерных структур [48].
 а — гомоструктура; б — структура с одним гетеропереходом; в — структура с двумя гетеропереходами.

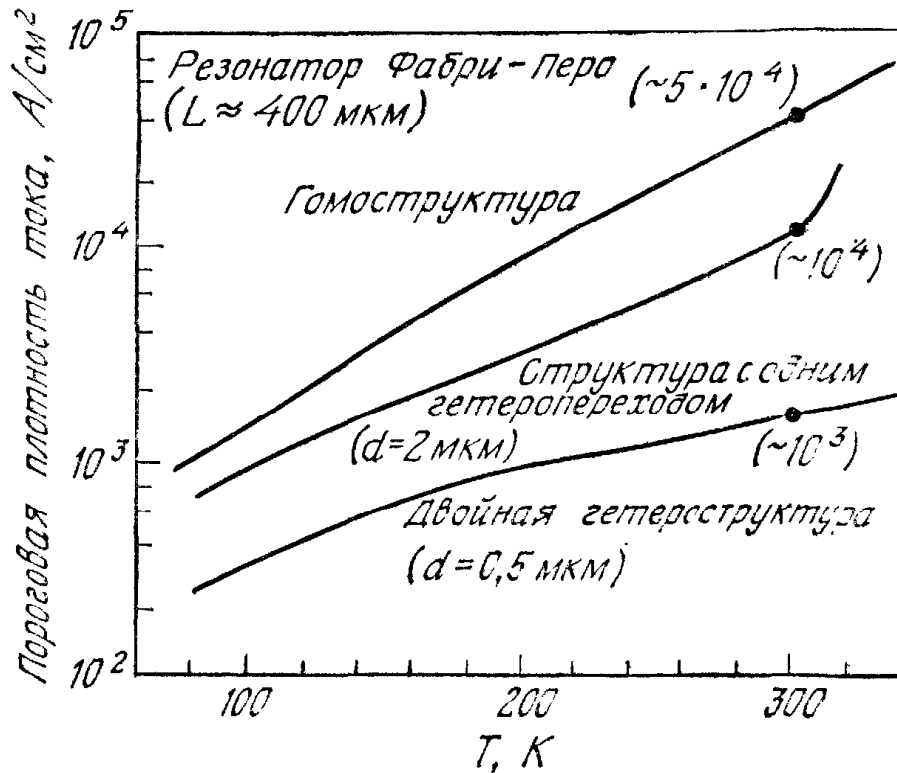


Рис. 28. Температурная зависимость пороговой плотности тока для трех лазерных структур [48].

областью вследствие скачкообразного уменьшения показателя преломления за ее пределами. Эти ограничения способствуют усилению стимулированного излучения и соответственно уменьшению пороговой плотности тока.

На рис. 28 для сравнения приведены зависимости J_{th} от рабочей температуры для трех структур, изображенных на рис. 27. Следует отметить, что самая слабая температурная зависимость наблюдается для лазеров на двойных гетероструктурах (ДГ-лазеры). Поскольку J_{th} в ДГ-лазерах при 300 К может достигать значений $\sim 10^3$ А/см² и менее, оказывается возможным реализовать режим непрерывной генерации при комнатной температуре. Благодаря этому лазеры нашли широкое применение в науке и промышленности, в частности в волоконно-оптических системах связи. Ниже рассмотрены в основном ДГ-лазеры.

12.4.2. Волноводный эффект

Коэффициент оптического ограничения. В ДГ-лазере свет удерживается и направляется диэлектрическим волноводом. На рис. 29, а показан диэлектрический волновод на основе трехслойной структуры, в которой активный слой расположен между двумя неактивными слоями. Если показатели преломления слоев \bar{n}_1 , \bar{n}_2 и \bar{n}_3 удовлетворяют условию

$$\bar{n}_2 > \bar{n}_1 \geq \bar{n}_3, \quad (23)$$

то угол падения луча θ_{12} на границу раздела между слоями 1 и 2 (рис. 29, б) будет больше критического угла, определяемого уравнением (12). Аналогичная ситуация наблюдается на границе раздела слоев 2 и 3. Таким образом, если показатель преломления активной области больше показателей преломления окружающих ее слоев (неравенство (23)), то электромагнитное излучение распространяется в направлении, параллельном границам раздела слоев.

В лазерах с гомоструктурой разность между показателем преломления центрального слоя волновода и показателями преломления соседних слоев составляет лишь 0,1—1 %. В гетеролазерах можно задать более существенное изменение показателей преломления на каждом гетеропереходе (до $\sim 10\%$) и тем самым обеспечить четко выраженное ограничение излучения.

Для более подробного анализа свойств волноводов обратимся к рис. 30, на котором схематически изображен лазер в декартовой системе координат. Плоскости $z = 0$ и $z = -L$ совпадают с двумя отражающими гранями или «зеркалами» лазера, которые представляют собой либо сколотые, либо отполированные поверх-

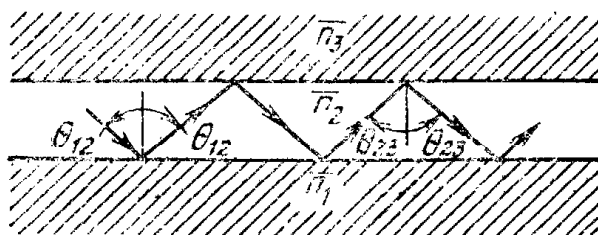
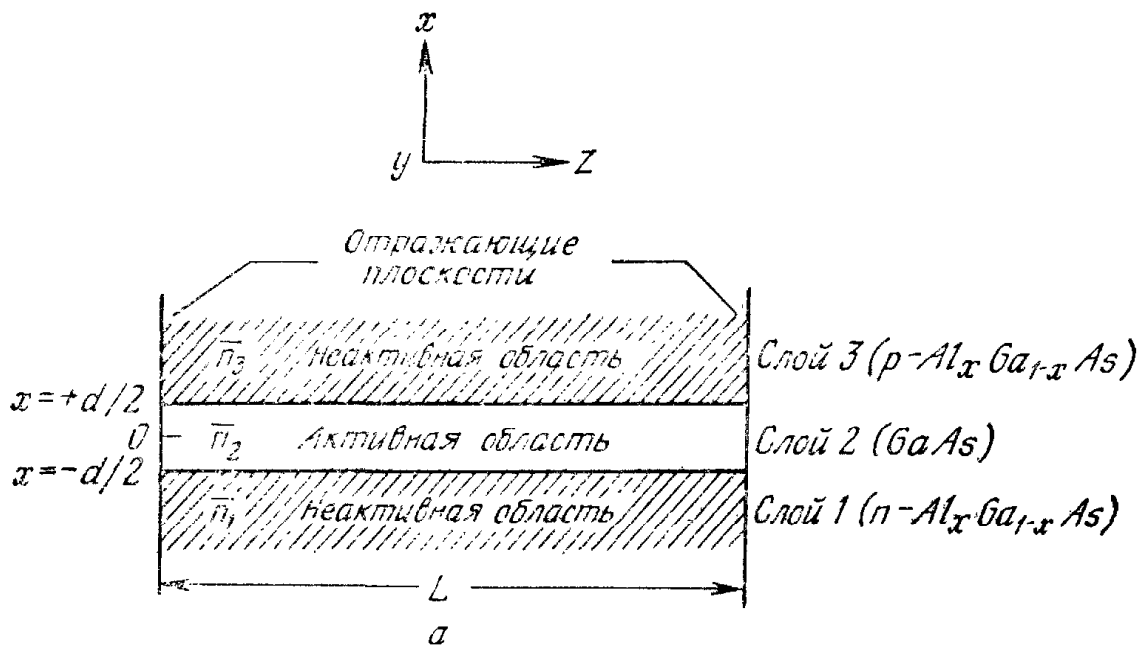


Рис. 29. Трехслойный диэлектрический волновод (а) и ход лучей в нем (б).

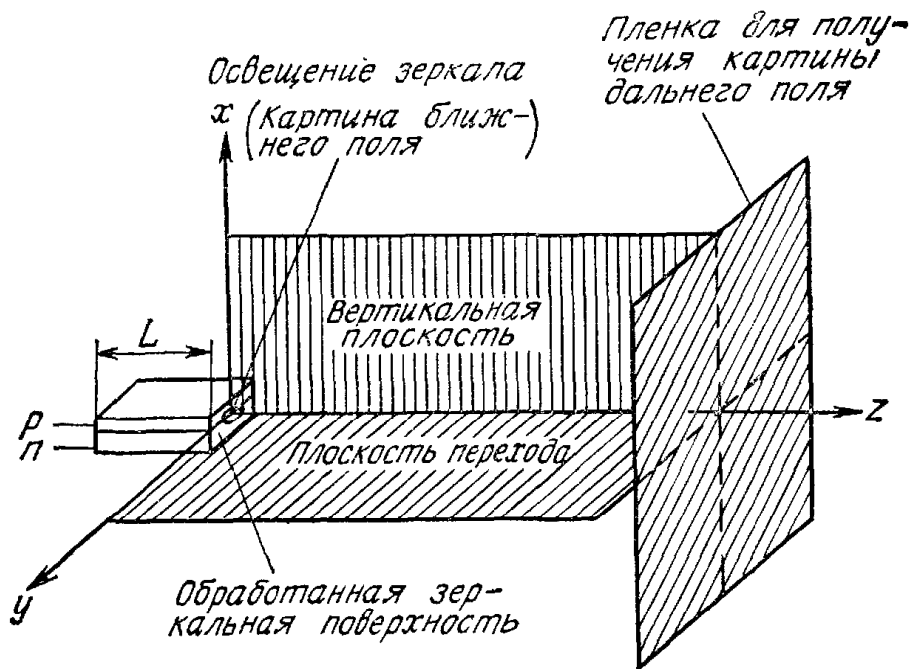


Рис. 30. Лазер с $p-n$ -переходом в декартовой системе координат.

ности. Переднее зеркало может излучать в полупространство $z > 0$. Пересекающиеся координатные оси x и y направлены соответственно перпендикулярно и параллельно плоскости перехода.

Рассмотрим симметричный трехслойный диэлектрический волновод с показателями преломления $\bar{n}_2 > \bar{n}_1 = \bar{n}_3$ (рис. 29). В электромагнитной волне с поперечной поляризацией электрического поля (ТЕ-поляризацией) компонента \mathcal{E}_z равна 0. Используя предположение о неограниченности волновода в направлении y , получим (вследствие $\partial/\partial y = 0$) упрощенное волновое уравнение

$$\partial^2 \mathcal{E}_y / \partial x^2 + \partial^2 \mathcal{E}_y / \partial z^2 = \mu_0 \epsilon \partial^2 \mathcal{E}_y / \partial t^2, \quad (24)$$

где μ_0 — магнитная проницаемость, а ϵ — диэлектрическая постоянная. Решая (методом разделения переменных) уравнение (24), получим для волны с ТЕ-поляризацией в пределах $-d/2 \leq x \leq d/2$ активного слоя

$$\mathcal{E}_y(x, z, t) = A_e \cos(\kappa x) \exp[j(\omega t - \beta z)]. \quad (25)$$

В выражении (25)

$$\kappa^2 = \bar{n}_2^2 k_0^2 - \beta^2, \quad (26)$$

где $k_0 \equiv (\omega/\bar{n}_2) \sqrt{\mu_0 \epsilon}$, а β — постоянная распространения. Магнитное поле в направлении z дается выражением

$$\begin{aligned} \mathcal{H}_z(x, z, t) &= (j/\omega\mu_0) (\partial \mathcal{E}_y / \partial x) = \\ &= (-j\kappa/\omega\mu_0) A_e \sin(\kappa x) \exp[j(\omega t - \beta z)]. \end{aligned} \quad (27)$$

Амплитуда удерживаемой волноводом электромагнитной волны спадает за пределами активного слоя. При $|x| > d/2$ решение для поперечного электрического и продольного магнитного полей имеет вид

$$\mathcal{E}_y(x, z, t) = A_e \cos(\kappa d/2) \exp[-\gamma(|x| - d/2)] \exp[j(\omega t - \beta z)], \quad (28)$$

$$\mathcal{H}_z(x, z, t) = (-x/|x|) (j\gamma/\omega\mu_0) A_e \cos(\kappa d/2) \times \\ \times \exp[-\gamma(|x| - d/2)] \exp[j(\omega t - \beta z)], \quad (29)$$

где

$$\gamma^2 \equiv \beta^2 - \bar{n}_1^2 k_0^2. \quad (30)$$

Из уравнений (26) и (30) видно, что, поскольку κ и γ должны быть положительными действительными величинами, условие существования волновых мод представляет собой неравенства $\bar{n}_2 k_0^2 > \beta^2$ и $\beta^2 > \bar{n}_1^2 k_0^2$ или

$$\bar{n}_2 > \bar{n}_1. \quad (31)$$

Этот результат идентичен соотношению (23).

Для определения постоянной распространения используем граничное условие, состоящее в том, что на границе раздела диэлектриков тангенциальная компонента магнитного поля \mathcal{H}_z должна быть непрерывной. Дисперсионное уравнение, полученное из уравнений (27) и (29) с использованием этого условия, имеет вид

$$\operatorname{tg}(\kappa d/2) = \gamma/\kappa = [(\beta^2 - \bar{n}_1^2 k_0^2)/(\bar{n}_2^2 k_0^2 - \beta^2)]^{1/2}. \quad (32)$$

Решение уравнения (32) зависит от аргумента тангенса, который определен с точностью до слагаемого $2\pi m$ (m — целое число). Моде низшего порядка или основной моде соответствует $m = 0$. При $m = 1$ реализуется мода первого порядка, при $m = 2$ — второго и т. д. При заданном числе m уравнение (32) может быть решено численно или графически. Полученное решение можно затем использовать в уравнениях (25) и (29) для определения электрического и магнитного полей.

Найдем теперь выражение для коэффициента оптического ограничения Γ , который представляет собой отношение интенсивности света в активном слое к полной интенсивности падающего света. Поскольку интенсивность света представляется вектором Пойнтинга $\mathcal{E} \times \mathcal{H}$, который пропорционален $|\mathcal{E}_y|^2$, то выражение для коэффициента оптического ограничения симметричного трехслойного диэлектрического волновода может быть

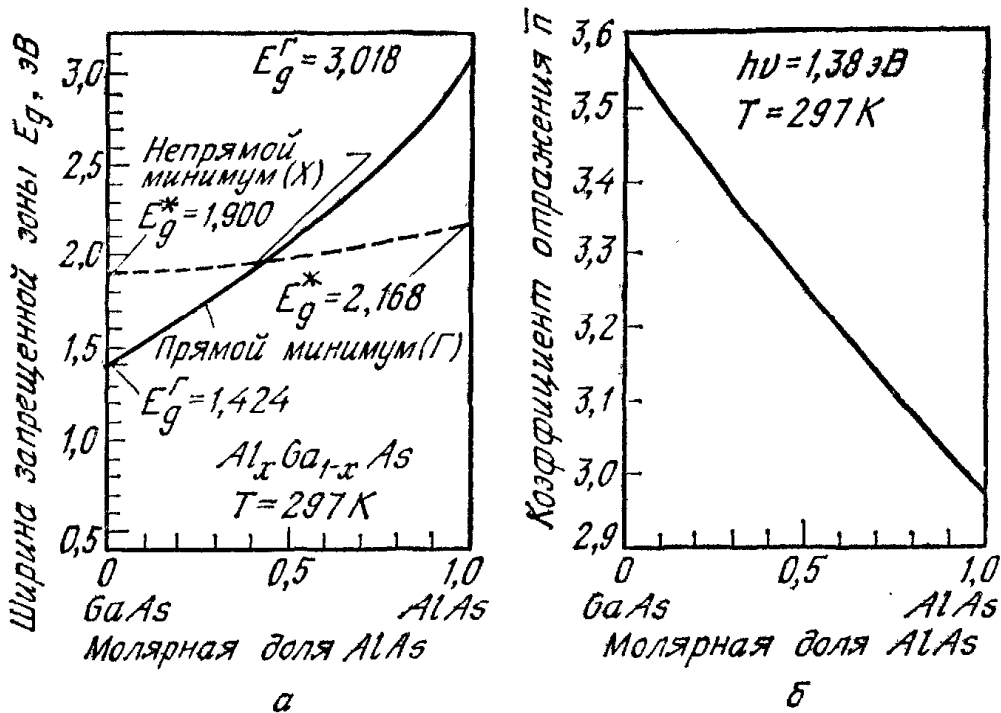


Рис. 31. Зависимость ширины запрещенной зоны (а) от состава $\text{Al}_x\text{Ga}_{1-x}\text{As}$ и показателя преломления (б) от состава для $h\nu = 1,38$ эВ [20].

получено с помощью уравнений (25) и (28). Для четных ТЕ-мод это выражение имеет вид

$$\Gamma = \frac{\int_0^{d/2} \cos^2(\kappa x) dx}{\int_0^{d/2} \cos^2(\kappa x) dx + \int_{d/2}^{\infty} \cos^2(\kappa d/2) \exp[-2\gamma(x - d/2)] dx} = \left\{ 1 + \frac{\cos^2(\kappa d/2)}{\gamma [(d/2) + (1/\kappa) \sin(\kappa d/2) \cos(\kappa d/2)]} \right\}^{-1}. \quad (33)$$

Аналогичные выражения можно получить для нечетных ТЕ-мод, а также для мод с поперечной поляризацией магнитного поля (ТМ-поляризация). Коэффициент оптического ограничения широко используется потому, что он дает представление о доле энергии, которая переносится волной, распространяющейся внутри активного слоя.

Среди лазерных гетероструктур в настоящее время наиболее интенсивно изучается система $\text{GaAs} - \text{Al}_x\text{Ga}_{1-x}\text{As}$. На рис. 31, а приведена зависимость ширины запрещенной зоны $\text{Al}_x\text{Ga}_{1-x}\text{As}$ от молярной доли AlAs. Для лазеров на основе гетероструктур наибольший интерес представляет диапазон $0 < x < 0,35$, в котором ширина прямой запрещенной зоны может быть представлена [20] в виде

$$E_g(x) = 1,424 + 1,247x \text{ [эВ]}. \quad (34)$$

На рис. 31, б приведена зависимость показателя преломления от состава соединения, которая может быть аппроксимирована выражением

$$\bar{n}(x) = 3,590 - 0,710x + 0,091x^2. \quad (35)$$

Например, при $x = 0,3$ ширина запрещенной зоны $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$ составляет 1,798 эВ, что на 0,374 эВ больше, чем у GaAs; при этом показатель преломления $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$ примерно на 6 % меньше, чем у GaAs.

Зависимости, приведенные на рис. 32, а, иллюстрируют влияние состава на распределение интенсивности $|\mathcal{E}_y|^2$ в трехслойном диэлектрическом волноводе на основе структуры $\text{Al}_x\text{Ga}_{1-x}\text{As} - \text{GaAs} - \text{Al}_x\text{Ga}_{1-x}\text{As}$ в направлении, перпендикулярном плоскости переходов. Расчет кривой проведен с помощью уравнений (25) и (32) для основной моды излучения ($m = 0$) с длиной волны 0,90 мкм (1,38 эВ). При заданной толщине активного слоя d в качестве переменного параметра выступает состав соединения. Значительное улучшение ограничения наблюдается при увеличении x от 0,1 до 0,2. На рис. 32, б показано изменение ограничения в зависимости от d при $x = 0,3$. По мере уменьшения толщины активного слоя свет все глубже проникает в $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$ и все меньшая часть общей интенсивности сосредоточивается внутри активного слоя. При увеличении d оказываются разрешенными моды более высокого порядка. Из данных, приведенных на рис. 32, в, следует, что с ростом порядка моды все большая часть света выходит за пределы активной области. Таким образом, для достижения эффективного оптического ограничения предпочтительны моды более низкого порядка.

На рис. 33 приведена зависимость коэффициента оптического ограничения Γ для основной моды от толщины активного слоя d и состава соединения. При $d < \lambda/\bar{n}_2$ ($\approx 0,5$ мкм) наблюдается резкое уменьшение Γ вследствие того, что толщина активного слоя становится меньше длины волны излучения. Представление той доли интенсивности волноводной моды, которая распространяется внутри активного слоя, с помощью коэффициента оптического ограничения Γ является важной концепцией для понимания взаимосвязи между толщиной активного слоя и пороговой плотностью тока.

Картина дальнего поля¹. Картиной дальнего поля называется распределение интенсивности излучения в свободном пространстве. На рис. 34 схематически представлена картина дальнего поля для излучения ДГ-лазера. Полные углы θ_{\parallel} и θ_{\perp} , измерен-

¹ В литературе употребляется также термин «картина излучения в дальней зоне». — Прим. перев.

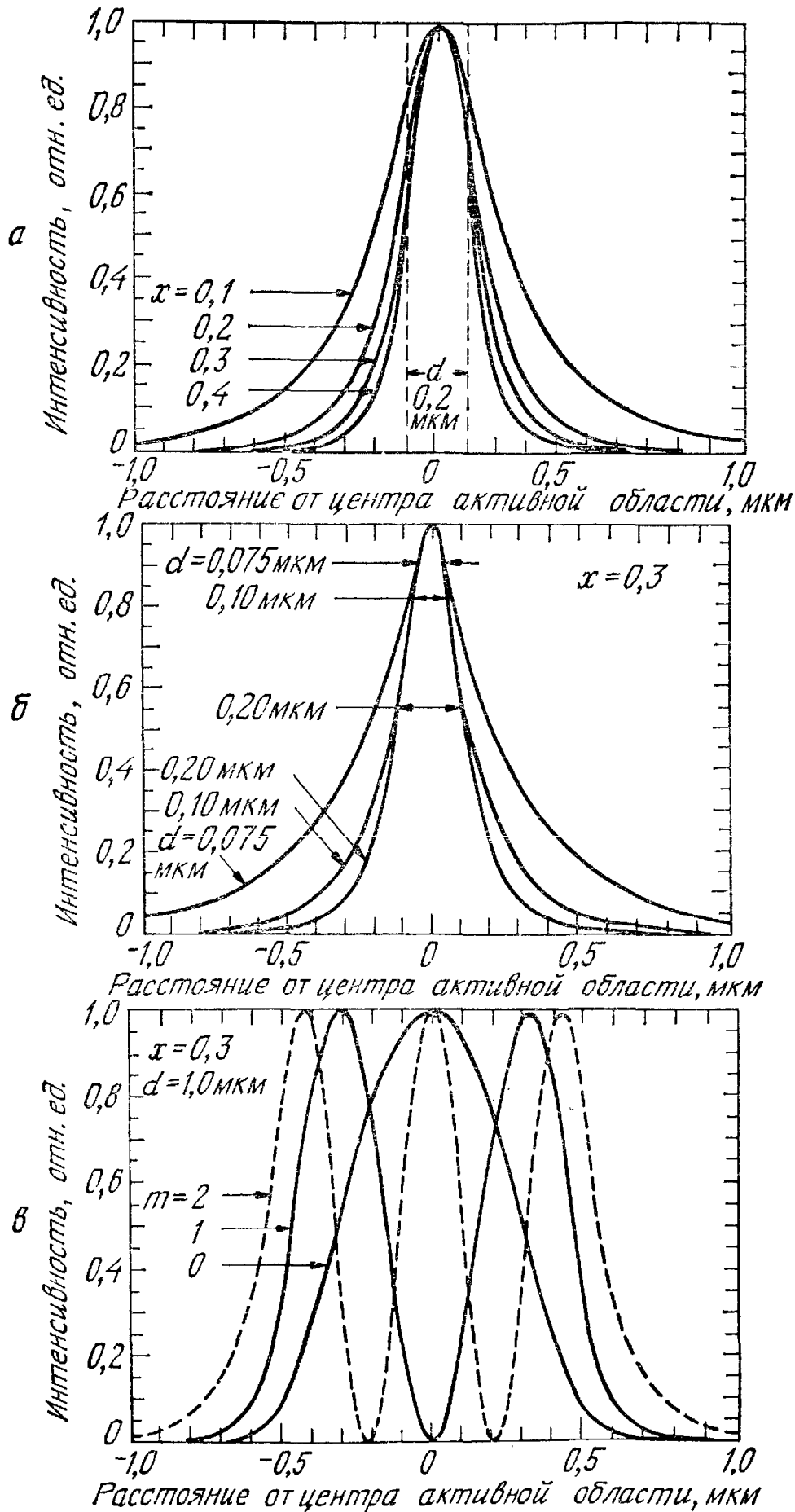


Рис. 32. Распределение интенсивности излучения в волноводе на основе двойной гетероструктуры [20].

а — изменение распределения в зависимости от молярной доли AlAs при $d = 0,2$ мкм; б — изменение распределения в зависимости от d при $x = 0,3$; в — изменение распределения в зависимости от порядка моды при постоянных x и d .

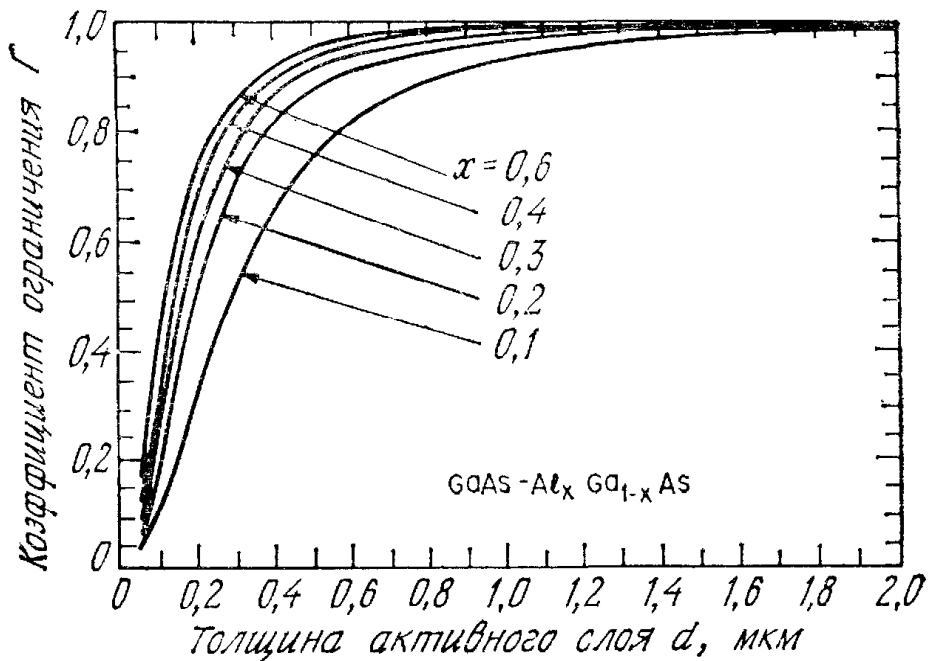


Рис. 33. Зависимость коэффициента оптического ограничения для фундаментальной моды от толщины активного слоя при различных значениях x в симметричном трехслойном диэлектрическом $\text{GaAs}-\text{Al}_x\text{Ga}_{1-x}\text{As}$ -волноводе [20].

ные на уровне половины максимальной интенсивности, характеризуют расходимость луча соответственно в плоскости перехода и перпендикулярно ей. Обычно θ_{\parallel} составляет $\sim 10^\circ$, тогда как θ_{\perp} принимает значительно большие значения ($35-60^\circ$) и зависит от толщины активного слоя и состава соединения.

Картину дальнего поля можно получить, в первую очередь, исходя из расчета пространственного распределения волн с ТЕ-поляризацией при $z > 0$. Для этого используется волновое уравне-

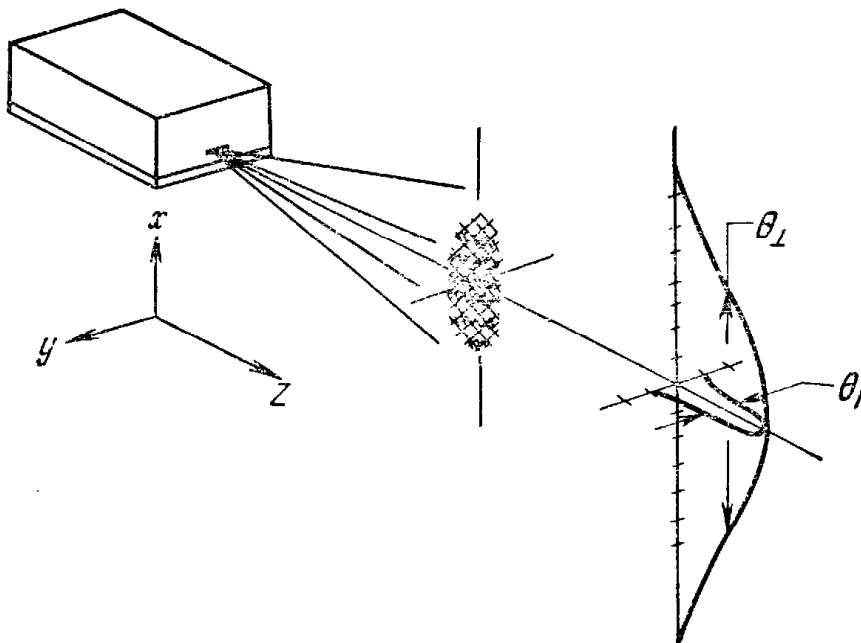


Рис. 34. Картина дальнего поля полоскового ДГ-лазера. Показаны полные углы на уровне половины максимальной интенсивности в плоскости перехода и перпендикулярно ей.

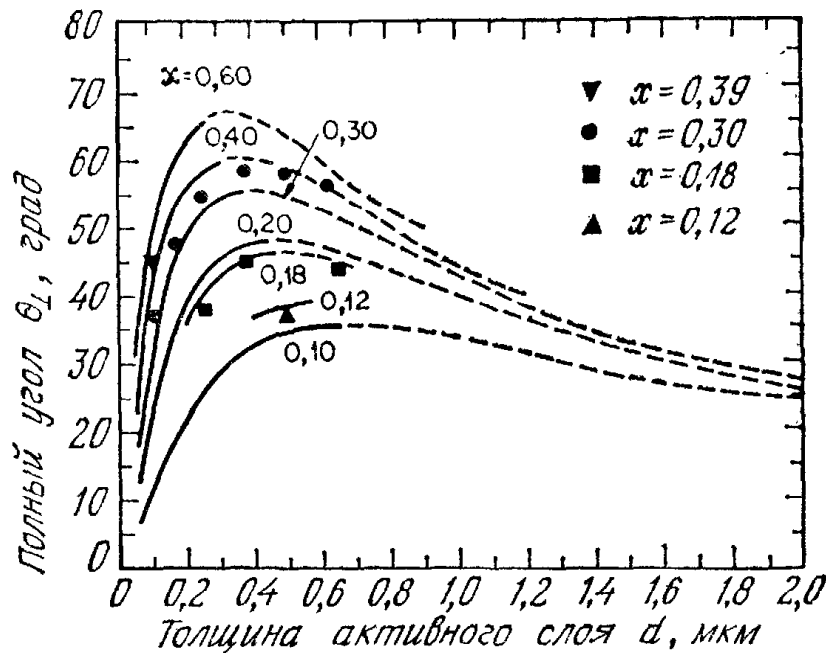


Рис. 35. Зависимость полного угла на уровне половины максимальной интенсивности от толщины активного слоя и состава соединения в GaAs — AlGaAs-лазерах [49].

ние (24), в котором ϵ заменено на ϵ_0 для свободного пространства:

$$\frac{I(\theta_{\perp})}{I(0)} = \frac{\cos^2 \theta_{\perp} \left| \int_{-\infty}^{\infty} \mathcal{E}_y(x, 0) \exp(j \sin \theta_{\perp} k_0 x) dx \right|^2}{\left| \int_{-\infty}^{\infty} \mathcal{E}_y(x, 0) dx \right|^2}. \quad (36)$$

В случае симметричного трехслойного волновода (ДГ-лазер) в уравнение (36) может быть подставлено выражение для электрического поля из уравнений (25) и (28). На рис. 35 приведены данные расчета и экспериментальные результаты для угла θ_{\perp} , полученные при исследовании картины дальнего поля GaAs — AlGaAs-лазера [49]. Сплошными и штриховыми линиями представлены кривые расходимости излучения основной моды, вычисленные из уравнения (36). Часть кривых, изображенных штриховой линией, соответствует толщине активного слоя, при которой возможны моды более высокого порядка. Экспериментальные результаты и данные расчета находятся в хорошем согласии. Для типичного значения толщины активного слоя 0,2 мкм ДГ-лазера на основе GaAs — Al_{0,3}Ga_{0,7}As угол θ_{\perp} составляет 50°.

На рис. 36 изображена базовая структура лазера с плоским волноводом, в котором излучающей является вся площадь перехода. Следует, однако, отметить, что гетеролазеры в основном имеют полосковую геометрию; на рис. 36 показаны две структуры

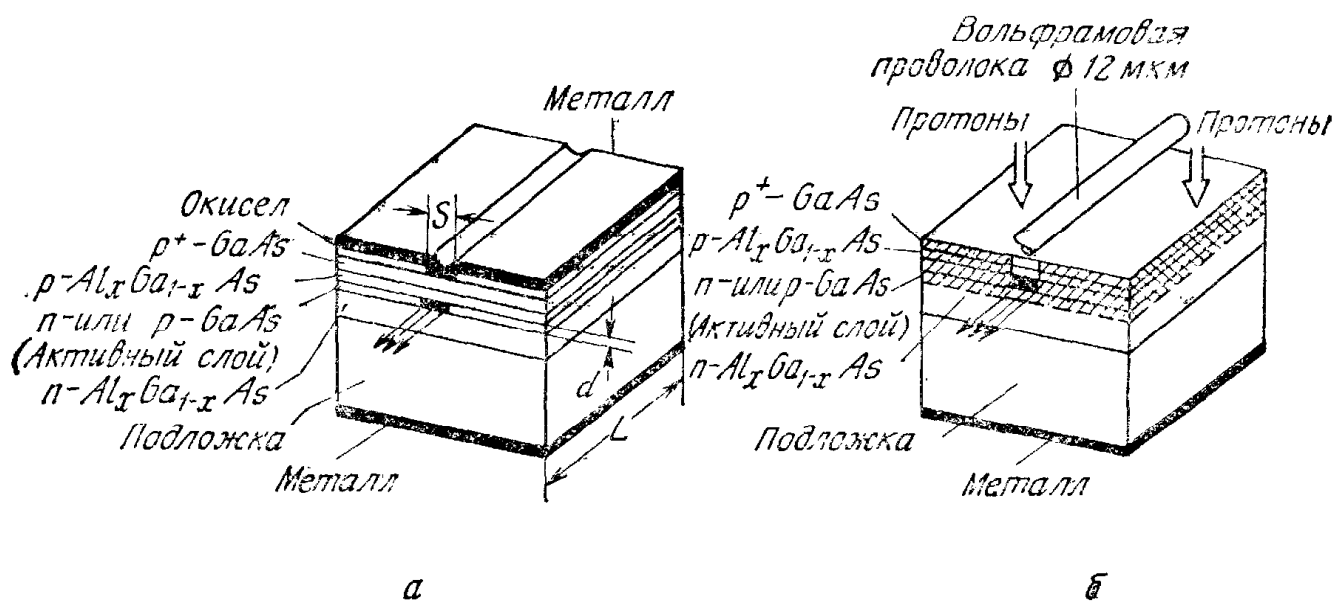


Рис. 36. ДГ-лазеры с полосковой геометрией [50].

а — с использованием окисной изоляции; б — с использованием протонной бомбардировки.

с такой геометрией [50]. В структуре на рис. 36, а слой окисла изолирует всю площадь перехода, кроме узкого контакта, который ограничивает излучающую поверхность, расположенную под ним. В структуре на рис. 36, б полосковая геометрия формируется в процессе протонной бомбардировки, с помощью которой создаются высокоомные области. Излучающая поверхность сосредоточена в центре области, не подвергнутой бомбардировке. Ширина канала обычно составляет 5—~30 мкм. За счет применения полосковой геометрии удается реализовать ряд преимуществ: 1) уменьшить площадь поперечного сечения излучающей области, что приводит к снижению рабочего тока; 2) устранить возможность возникновения нескольких областей с высокой интенсивностью излучения (так называемых шнуров) при увеличении тока накачки (при ширине канала менее 15 мкм в лазере с полосковой геометрией можно обеспечить одношнуровой режим с излучением основной моды вдоль плоскости перехода); 3) улучшить надежность за счет отдаления большей части периметра перехода от граней кристалла; 4) уменьшить инерционность за счет снижения емкости перехода.

В лазере с полосковой геометрией напряженность электрического поля вдоль направления, параллельного плоскости перехода (ось y), сильно зависит от пространственного изменения диэлектрической постоянной. Для структуры на рис. 37 волновое уравнение для электрического поля, имеющего временную зависимость вида $e^{j\omega t}$, записывается следующим образом [51]:

$$\nabla^2 \mathcal{E}_y + (k_0^2 \epsilon / \epsilon_0) \mathcal{E}_y = 0. \quad (37)$$

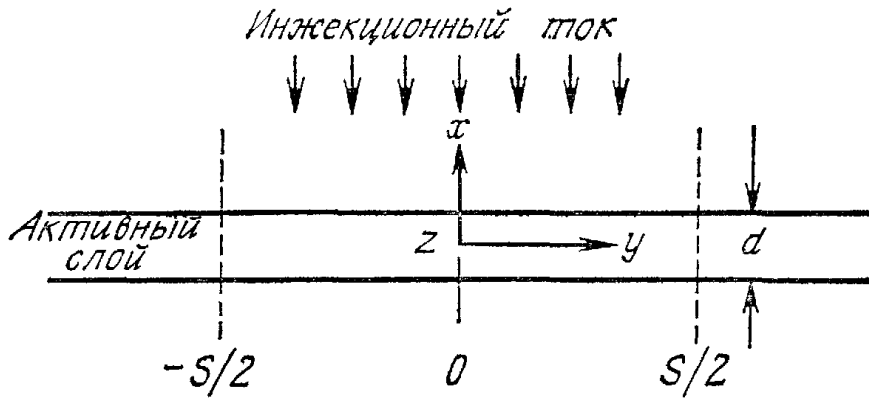


Рис. 37. Система координат для полоскового лазера, полученного протонной бомбардировкой, с шириной полосы S .

В этом уравнении $k_0 = 2\pi/\lambda$, а отношение $\varepsilon/\varepsilon_0$ представляется для области внутри активного слоя в двумерной форме

$$\varepsilon(x, y)/\varepsilon_0 = [\varepsilon(0) - a^2 y^2]/\varepsilon_0. \quad (38)$$

В прилегающих слоях оно имеет вид

$$\varepsilon(x, y)/\varepsilon_0 = \varepsilon_1/\varepsilon_0. \quad (39)$$

В уравнении (38) $\varepsilon(0) = \varepsilon_r(0) + j\varepsilon_i(0)$ представляет собой комплексную диэлектрическую постоянную активного слоя при $y = 0$, а коэффициент a — комплексную константу $a_r + ja_i$. Приближенное решение уравнения (37) при использовании диэлектрических постоянных из уравнений (38) и (39) имеет вид

$$\mathcal{E}_x(x, y, z) = \mathcal{E}_y(x) \mathcal{E}_y(y) \exp(-j\beta_z z). \quad (40)$$

Поскольку $\varepsilon(x, y)$ слабо зависит от y в плоскости перехода, то компонента \mathcal{E}_y почти не ограничивается вдоль направления y и поэтому может быть представлена полученными выше выражениями (25) и (28). Из уравнения (37) с помощью разделения переменных можно получить

$$[\partial^2 \mathcal{E}_y(x)/\partial x^2] + \beta_x^2 \mathcal{E}_y(x) = 0. \quad (41)$$

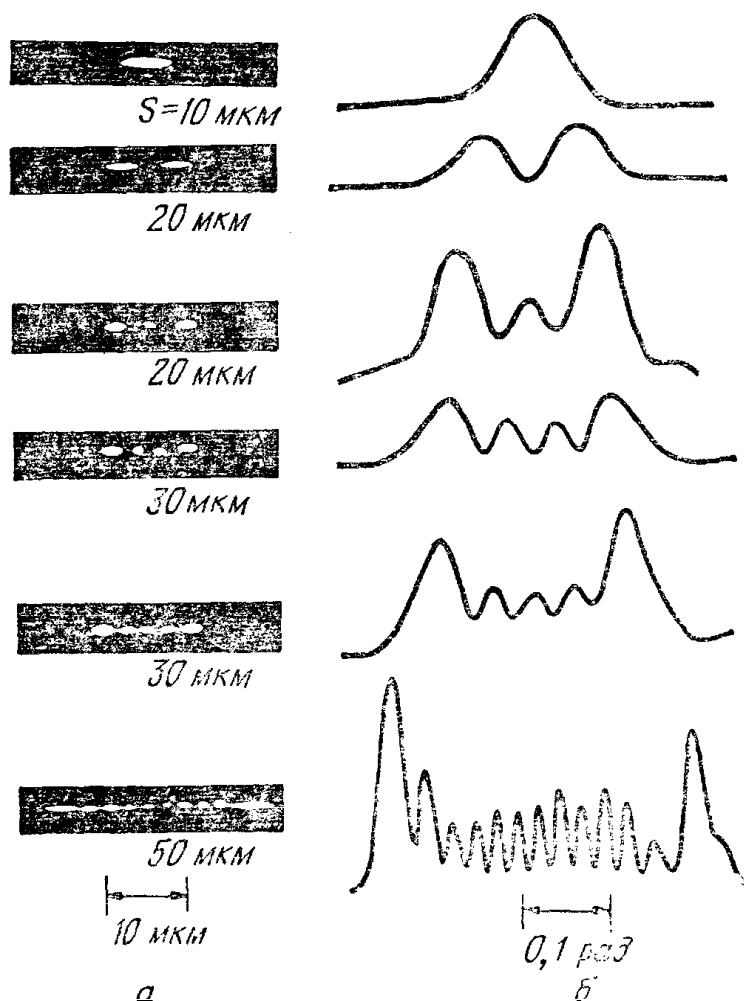
Подстановка уравнений (40) и (41) в уравнение (37) с последующим исключением из него компоненты $\mathcal{E}_y(x)$ путем умножения на сопряженное этой компоненте комплексное выражение и интегрирования по x приводит к дифференциальному уравнению для $\mathcal{E}_y(y)$:

$$[\partial^2 \mathcal{E}_y(y)/\partial y^2] + \{k_0^2 [\Gamma \varepsilon(0)/\varepsilon_0 + (1 - \Gamma) \varepsilon_1/\varepsilon_0] - \beta_x^2 - \beta_z^2 - (\Gamma k_0^2 a^2 y^2/\varepsilon_0)\} \mathcal{E}_y(y) = 0, \quad (42)$$

в котором Γ — коэффициент оптического ограничения.

Рис. 38. Структура мод вдоль плоскости $p-n$ -перехода в зависимости от ширины полоски S в планарном полосковом ДГ-лазере [52].

a — картины ближнего поля; b — картины дальнего поля.



Распределение поля $\mathcal{E}_y(y)$, полученное при решении уравнения (42), представляет собой функцию Эрмита — Гаусса

$$\mathcal{E}_y(y) = H_p \left[\left(\Gamma^{1/2} a k_0 / \varepsilon_0^{1/2} \right)^{1/2} y \right] \exp \left[-\frac{1}{2} \left(\Gamma / \varepsilon_0 \right)^{1/2} a k_0 y^2 \right], \quad (43)$$

где H_p — полином Эрмита порядка p , равный

$$H_p(\xi) \equiv (-1)^p \exp(\xi^2) \partial^p \exp(-\xi^2) / \partial \xi^p. \quad (44)$$

Приведем три первых полинома Эрмита: $H_0(\xi) = 1$, $H_1(\xi) = 2\xi$ и $H_2(\xi) = 4\xi^2 - 2$. Таким образом, распределение интенсивности основной моды дается выражением в виде функции Гаусса

$$|\mathcal{E}_y(y)|^2 = \exp \left[-\left(\Gamma / \varepsilon_0 \right)^{1/2} a_r k_0 y^2 \right], \quad (45)$$

из которого следует, что распределение интенсивности вдоль плоскости перехода зависит от a_r .

На рис. 38 показаны картины ближнего и дальнего поля вдоль плоскости перехода для лазера с полосковой геометрией [52]. При ширине канала 10 мкм наблюдается основная мода с гауссовым распределением интенсивности. По мере увеличения ширины канала вдоль плоскости перехода возникают моды более высокого порядка. Эти моды характеризуются распределением интенсивности в виде функций Эрмита — Гаусса, определяемых уравнением (43).

12.4.3. Пороговая плотность тока

В условиях теплового равновесия в основном состоянии находится больше атомов, чем в возбужденном. Если возникает обратная ситуация, то говорят об инверсной населенности. При взаимодействии фотонов, обладающих энергией $h\nu_{12}$, с простой системой (рис. 25), в которой уровень E_2 инверсно заселен по отношению к уровню E_1 , стимулированная эмиссия будет преобладать над поглощением, и в результате систему будет покидать больше фотонов с энергией $h\nu_{12}$, чем входит в нее. Такое явление называют квантовым усилением.

Для того чтобы рассмотреть условия инверсной населенности в полупроводниковых лазерах, обратимся к рис. 39, на котором приведены зависимости энергии от плотности состояний в полупроводнике с прямой зоной [53]. Ситуация, изображенная на рис. 39, а, соответствует условию равновесия при $T = 0$ К для собственного полупроводника, в котором все состояния в валентной зоне заняты (заштрихованная площадь). На рис. 39, б показан случай инверсной населенности при $T = 0$ К, которая возникает, например, при фотовозбуждении электронов фотонами с энергией, превышающей ширину запрещенной зоны E_g . Валентная зона выше уровня E_{FV} свободна от электронов, а зона проводимости заполнена ими до уровня E_{FC} . В такой ситуации фотоны с энергией $h\nu$ в диапазоне $E_g < h\nu < (E_{FC} - E_{FV})$ будут вызывать переходы электронов на нижние уровни, что в свою очередь приведет к возникновению стимулированной эмиссии.

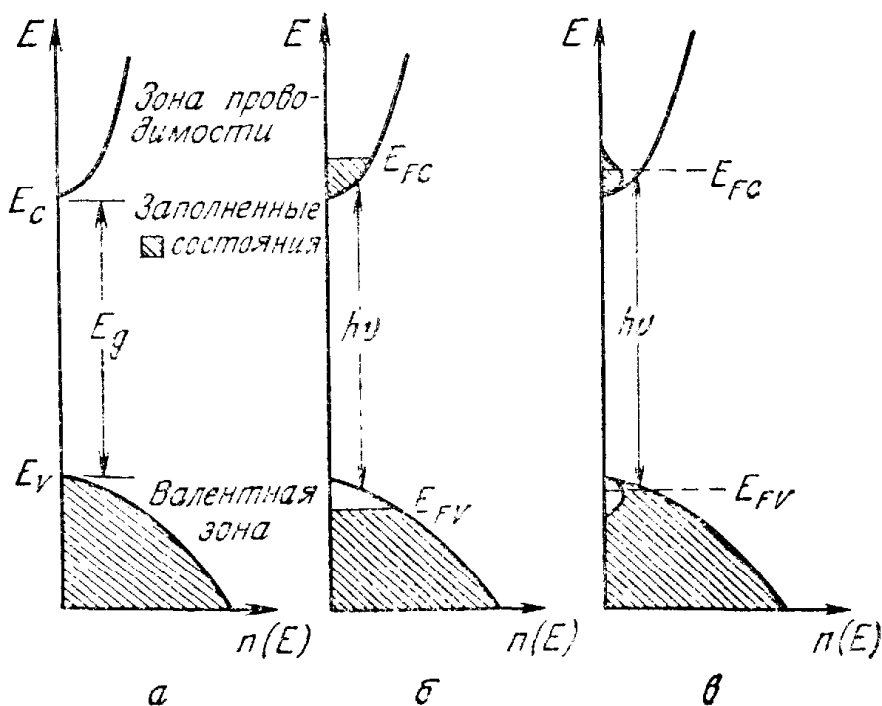


Рис. 39. Зависимость плотности состояний от энергии в полупроводнике [53]. а — состояние равновесия, $T = 0$ К; б — инверсия, $T = 0$ К; в — инверсия, $T > 0$ К.

При конечных температурах распределение носителей «размазано» по энергетическому спектру (рис. 39, в). Несмотря на отсутствие общего теплового равновесия, носители в данной энергетической зоне находятся в тепловом равновесии друг с другом. Вероятность заполнения состояния в зоне проводимости дается распределением Ферми — Дирака

$$F_C(E) = \frac{1}{1 + \exp\left(\frac{E - E_{FC}}{kT}\right)}, \quad (46)$$

где E_{FC} — квазиуровень Ферми для электронов в зоне проводимости. Аналогичное выражение справедливо и для валентной зоны.

Рассмотрим эмиссию фотонов с энергией $h\nu$, возникающую в результате переходов электронов с верхних состояний, которым соответствуют энергии вблизи уровня E в зоне проводимости, на нижние состояния, которым соответствуют энергии вблизи уровня $(E - h\nu)$ в валентной зоне. Скорость такой эмиссии пропорциональна произведению плотности занятых верхних состояний $n_C(E) F_C(E)$ и плотности свободных нижних состояний $n_V(E - h\nu) [1 - F_V(E - h\nu)]$. Полная скорость эмиссии определяется в результате интегрирования по всем энергиям:

$$W_{\text{спонт}}(h\nu) = B \int n_C(E) n_V(E - h\nu) F_C(E) [1 - F_V(E - h\nu)] |\langle M \rangle|^2 dE. \quad (47)$$

Таким же образом можно получить выражение для скорости поглощения:

$$W_{\text{погл}}(h\nu) = B \int n_V(E - h\nu) n_C(E) F_V(E - h\nu) [1 - F_C(E)] |\langle M \rangle|^2 dE. \quad (48)$$

Коэффициент B в этих выражениях определяется соотношением

$$B = (4\pi\bar{n}q^2 h\nu / m^2 \epsilon_0 h^2 c^3) \text{Vol}, \quad (49)$$

где $\langle M \rangle$ — матричный элемент, а Vol — объем кристалла. Усиление возникает при условии $W_{\text{спонт}} \triangleright W_{\text{погл}}$. Из уравнений (46) и (48) с соответствующими квазиуровнями Ферми E_{FC} и E_{FV} получим [11]

$$(E_{FC} - E_{FV}) \triangleright h\nu. \quad (50)$$

Это неравенство является необходимым условием для того, чтобы стимулированное излучение преобладало над поглощением.

В полупроводниковом лазере коэффициент усиления g (т. е. приращение потока энергии на единицу длины) определяется структурой энергетических зон и сложным образом зависит от степени легирования, плотности тока, температуры и частоты.

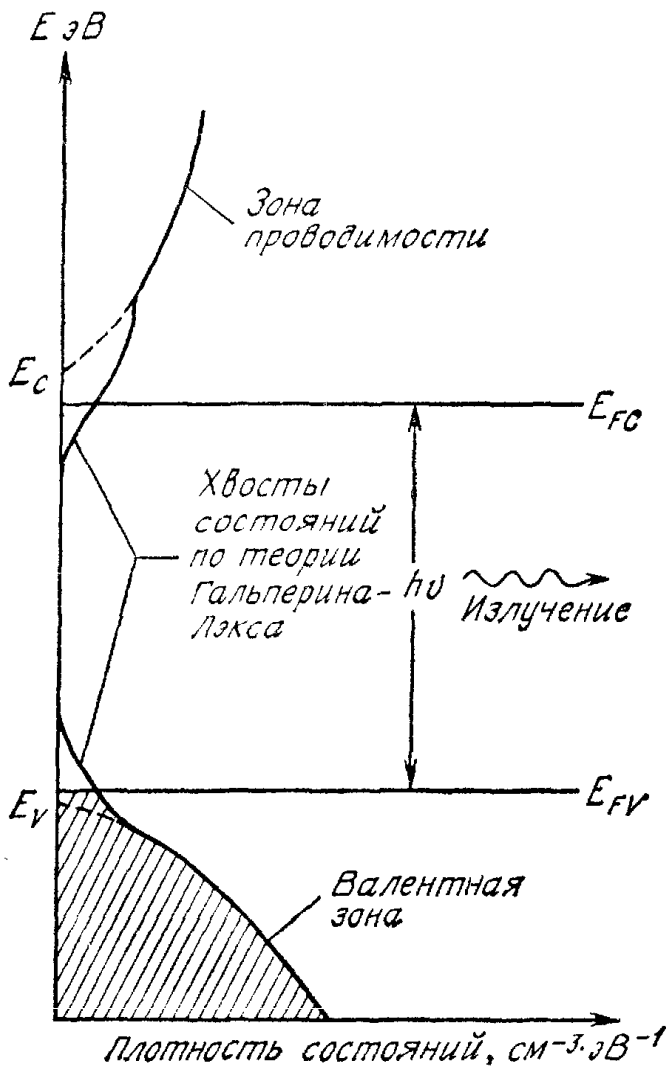


Рис. 40. Распределение плотности состояний при наличии хвостов состояний валентной зоны и зоны проводимости [55].

При возрастании интенсивности возбуждения функции распределения $E_c(E)$ и $F_v(E)$ изменяются таким образом, что E_{FC} увеличивается, а E_{FV} уменьшается. Изменяется также и форма кривой, описывающей зависимость коэффициента усиления от энергии фотонов.

Коэффициент усиления может быть рассчитан для специального распределения плотности состояний [55], при котором как валентная зона, так и зона проводимости имеют хвосты плотности состояний (рис. 40). Такое распределение характерно для большинства ДГ-лазеров на GaAs. Если распределение плотности состояний найдено, то для заданных значений $\Delta E \equiv (E_{FC} - E_{FV})$ и температуры с помощью уравнения (47) можно рассчитать функцию спонтанного излучения и получить зависимость коэффициента усиления $g = -\alpha$ от ΔE . Полученные результаты дают возможность установить также зависимость $g(h\nu)$ от номинальной плотности тока [56] $J_{\text{ном}}$, которая определяется как плотность тока, необходимая для поддержания постоянного возбуждения в активном слое толщиной 1 мкм при квантовой эффективности $\eta = 1$. Действительное значение плотности тока при этом определяется выражением

$$J \text{ (A/cm}^2\text{)} = J_{\text{ном}} d / \eta. \quad (51)$$

На рис. 41 приведена зависимость коэффициента усиления от номинальной плотности тока, рассчитанная для GaAs с концентрацией ионизованных акцепторов $5 \cdot 10^{17} \text{ см}^{-3}$ и концентрацией ионизованных доноров $1 \cdot 10^{17} \text{ см}^{-3}$ [56]. Эта зависимость сверхлинейна при малых значениях коэффициента усиления и линейна при $50 \leq g \leq 400 \text{ см}^{-1}$. Штриховая линия соответствует зависимости, которая может быть описана уравнением

$$g = (g_0/J_0) (J_{\text{ном}} - J_0), \quad (52)$$

где $g_0/J_0 = 5 \cdot 10^{-2} \text{ см} \cdot \text{мкм}/\text{А}$ и $J_0 = 4,5 \cdot 10^3 \text{ А} (\text{см}^2 \cdot \text{мкм})$.

Как уже говорилось выше, при малых токах спонтанное излучение распространяется во всех направлениях. По мере увеличения тока усиление растет (рис. 41) до тех пор, пока не достигается порог лазерной генерации, т. е. не создаются условия, при которых световая волна проходит через резонатор без затухания:

$$R \exp [(\Gamma g - \alpha) L] = 1, \quad (53)$$

$$\text{или } \Gamma g = \alpha + \frac{1}{L} \ln \left(\frac{1}{R} \right), \quad (54)$$

где g — пороговый коэффициент усиления; Γ — коэффициент оптического ограничения; α — потери на излучения на единице длины¹, обусловленные поглощением на свободных носителях и рассеянием на дефектах; L — длина резонатора; R — коэффициент отражения зеркал резонатора (при разных коэффициентах отражения зеркал резонатора $R = \sqrt{R_1 R_2}$). Из урав-

¹ То есть коэффициент поглощения. — Прим. перев.

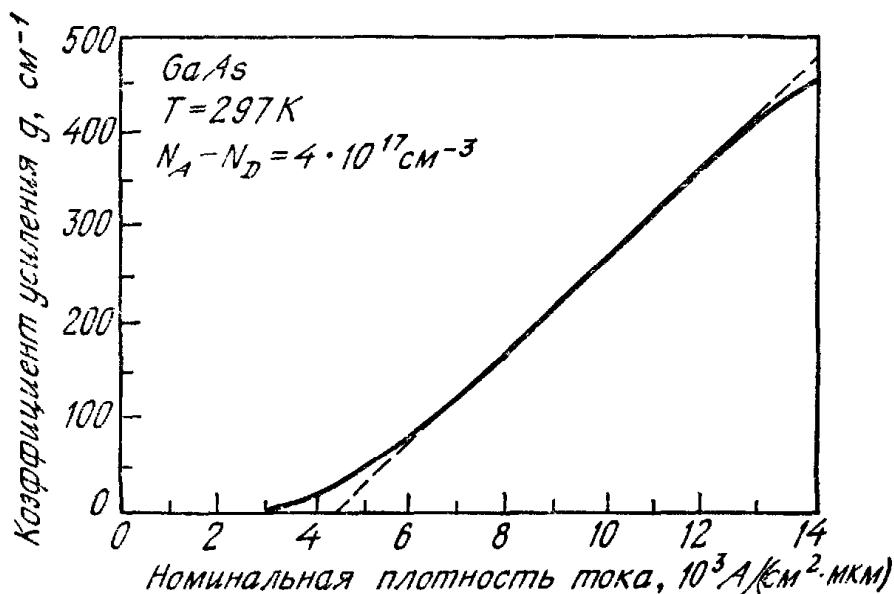


Рис. 41. Зависимость коэффициента усиления от номинальной плотности тока. Штриховой линией представлена линейная зависимость [56].

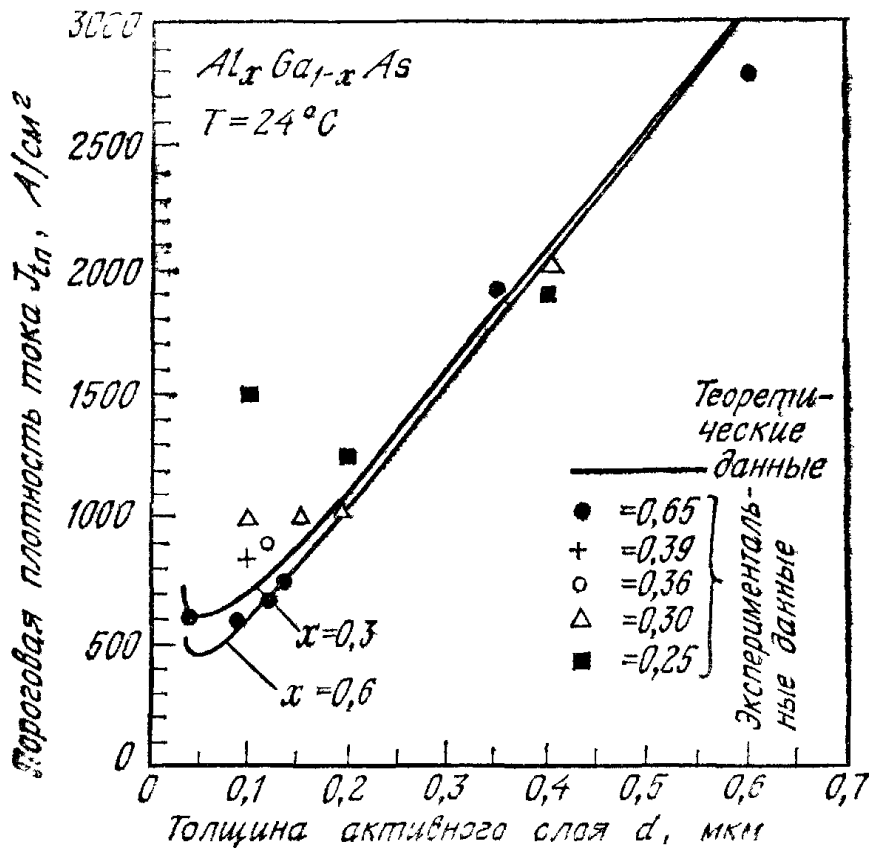


Рис. 42. Сравнение экспериментальной и рассчитанной по уравнению [55] зависимостей J_{th} от толщины активного слоя [57].

нений (51), (52) и (54) можно получить выражение для пороговой плотности тока [57]

$$J_{th} (\text{A/cm}^2) = \frac{J_0 d}{\eta} + \frac{J_0 d}{g_0 \eta \Gamma} \left[\alpha + \frac{1}{L} \ln \left(\frac{1}{R} \right) \right]. \quad (55)$$

Для снижения пороговой плотности тока можно увеличивать η , Γ , L и R , а также уменьшать d и α . На рис. 42 приведены для сравнения экспериментальные результаты [57] и теоретические кривые, рассчитанные по уравнению (55). С уменьшением d пороговая плотность тока уменьшается, достигает минимального значения, а затем снова возрастает. Увеличение J_{th} при очень малых толщинах активного слоя обусловлено ухудшением оптического ограничения. При заданных d уменьшение J_{th} с увеличением x (молярной доли AlAs) происходит вследствие улучшения оптического ограничения. Аналогичные результаты были получены для ДГ-лазеров на основе $\text{InP} - \text{Ga}_x\text{In}_{1-x}\text{As}_y\text{P}_{1-y} - \text{InP}$ [58, 59]. При этом в лазерах с активным слоем из $\text{Ga}_{0,25}\text{In}_{0,75}\text{As}_{0,54}\text{P}_{0,46}$ ($\lambda = 1,23$ мкм) толщиной $\sim 0,2$ мкм минимальное значение J_{th} при комнатной температуре составляет $\sim 1,5$ кА/см².

12.5. РАБОЧИЕ ХАРАКТЕРИСТИКИ ЛАЗЕРОВ

12.5.1. Лазерные структуры

Низкая пороговая плотность в гетеролазерах обеспечивается двумя факторами: 1) ограничением носителей энергетическими барьерами между активной областью и слоем полупроводника с более широкой запрещенной зоной и 2) оптическим ограничением за счет резкого уменьшения показателя преломления за пределами активной области. В разд. 12.4 были рассмотрены ДГ-лазеры, в которых ограничение обоих типов происходит внутри активного слоя на одной и той же толщине.

Для создания гетеролазеров используются и некоторые другие гетероструктуры [60]. На рис. 43, *a* показана последовательность слоев в гетеролазере с отдельным ограничением и приведены зависимости ширины запрещенной зоны, показателя преломления и распределение интенсивности света в направлении, перпендикулярном плоскости перехода. Скачок E_g на границе между GaAs и $\text{Al}_{0,1}\text{Ga}_{0,9}\text{As}$ достаточен для того, чтобы ограничить число носителей внутри слоя толщиной d , в то же время скачок \bar{n} на этой

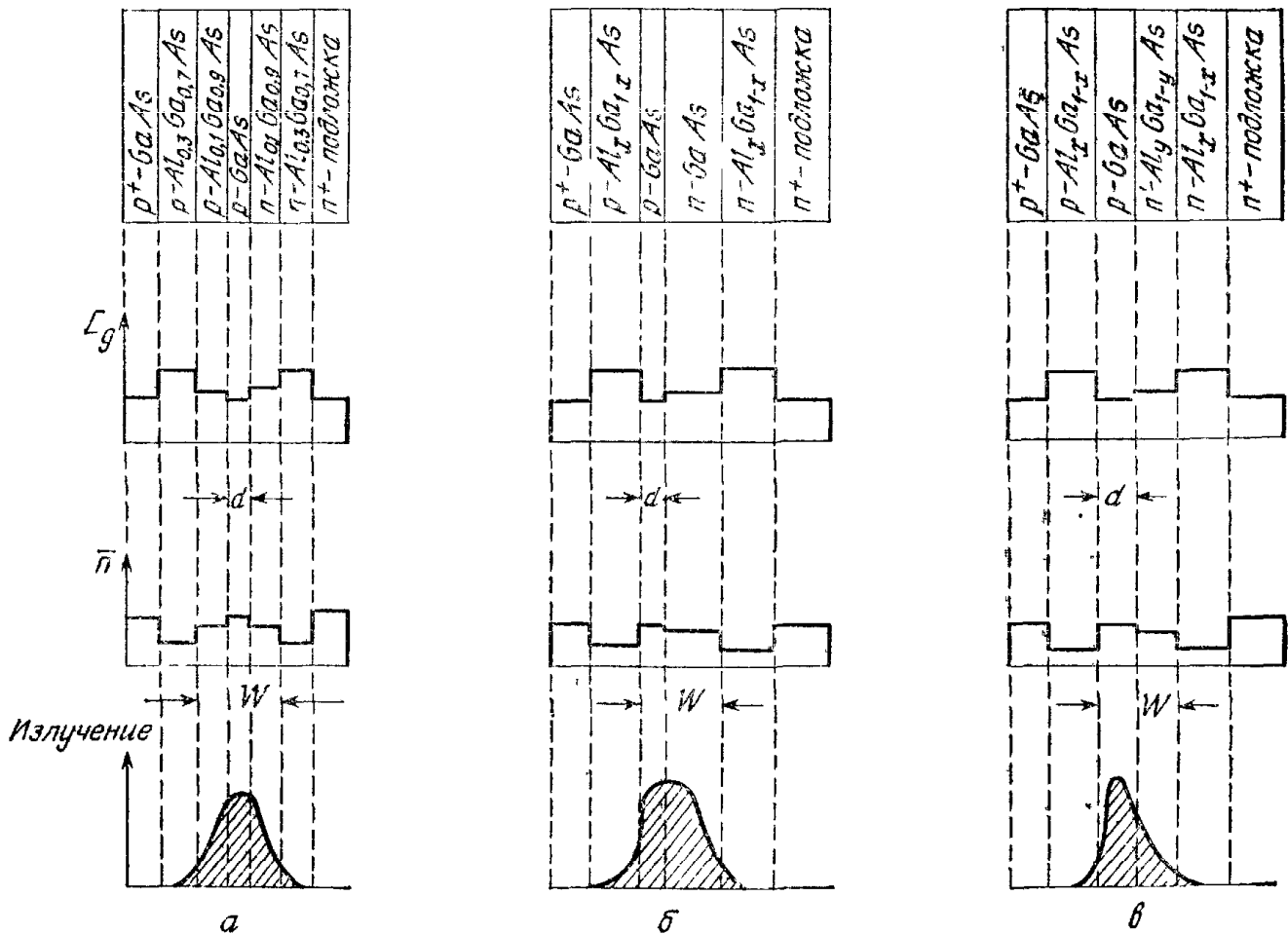


Рис. 43. Энергетическая диаграмма, показатель преломления и интенсивность света в трех типах гетеролазеров [60].

a — ДГ-лазер с отдельным ограничением; *б* — ДГ-лазер с расширенным волноводом; *в* — четырехслойный гетеролазер.

границе не обеспечивает оптического ограничения. Большая величина скачка показателя преломления между $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$ и $\text{Al}_{0,1}\text{Ga}_{0,9}\text{As}$ служит для оптического ограничения и формирует оптический волновод толщиной W . В такой структуре была получена низкая пороговая плотность тока $J_{th} \approx 500 \text{ А/см}^2$.

Четырехслойный гетеролазер с расширенным волноводом (рис. 43, б) имеет структуру, аналогичную структуре ДГ-лазера с отдельным ограничением, за исключением того, что в нем в качестве волновода используется p — n -гомопереход, который помещен между двумя гетеропереходами. Большая часть тока, протекающего через переход, обусловлена инжекцией электронов в p -слой, который играет роль активной области. Гетеропереход $p\text{-GaAs}/p\text{-Al}_x\text{Ga}_{1-x}\text{As}$ обеспечивает оптическое и электронное ограничения. Преимущество таких лазеров состоит в том, что в них в отличие от ДГ-лазеров увеличение толщины оптического волновода W достигается без существенного возрастания J_{th} . Поскольку в этих лазерах сечение оптического волновода может быть увеличено, то по сравнению с ДГ-лазерами они могут излучать большую мощность.

На рис. 43, в показан четырехслойный гетеролазер, в котором слой $n\text{-GaAs}$ заменен слоем $n'\text{-Al}_y\text{Ga}_{1-y}\text{As}$ с $y = 0,1—0,2$. Эта структура может работать в режиме генерации основной поперечной моды при сравнительно низкой плотности J_{th} , имеет высокую выходную мощность и относительно небольшую расходимость луча.

Большинство гетеролазеров имеет полосковую геометрию, аналогичную той, которая показана на рис. 36. На рис. 44 приведены поперечные сечения нескольких других полосковых лазерных структур. Для создания структур используются подложки из высококачественных бинарных полупроводников (например, GaAs или InP). Последующие эпитаксиальные слои из различных соединений и с различными легирующими примесями выращиваются методами жидкофазной, газофазной или молекулярно-лучевой эпитаксии. Для каждой структуры на рисунке указано, какой слой является активным, и обозначено сечение излучения.

На рис. 44, а показан полосковый лазер, полученный простым методом [61]. Мезополосковая структура лазера создается путем травления. После нанесения металлических контактов осуществляется завершающая операция скалывания граней, в результате которой формируется резонатор Фабри — Перо. В такой структуре не надо создавать окисную изоляцию или проводить протонную бомбардировку. Боковое ограничение тока достигается за счет высокого контактного сопротивления между металлом и слаболегированным $p\text{-AlGaAs}$. Этот лазер имеет низкую пороговую плотность тока, линейную ватт-амперную характеристику и работает в режиме генерации основной поперечной моды с одной

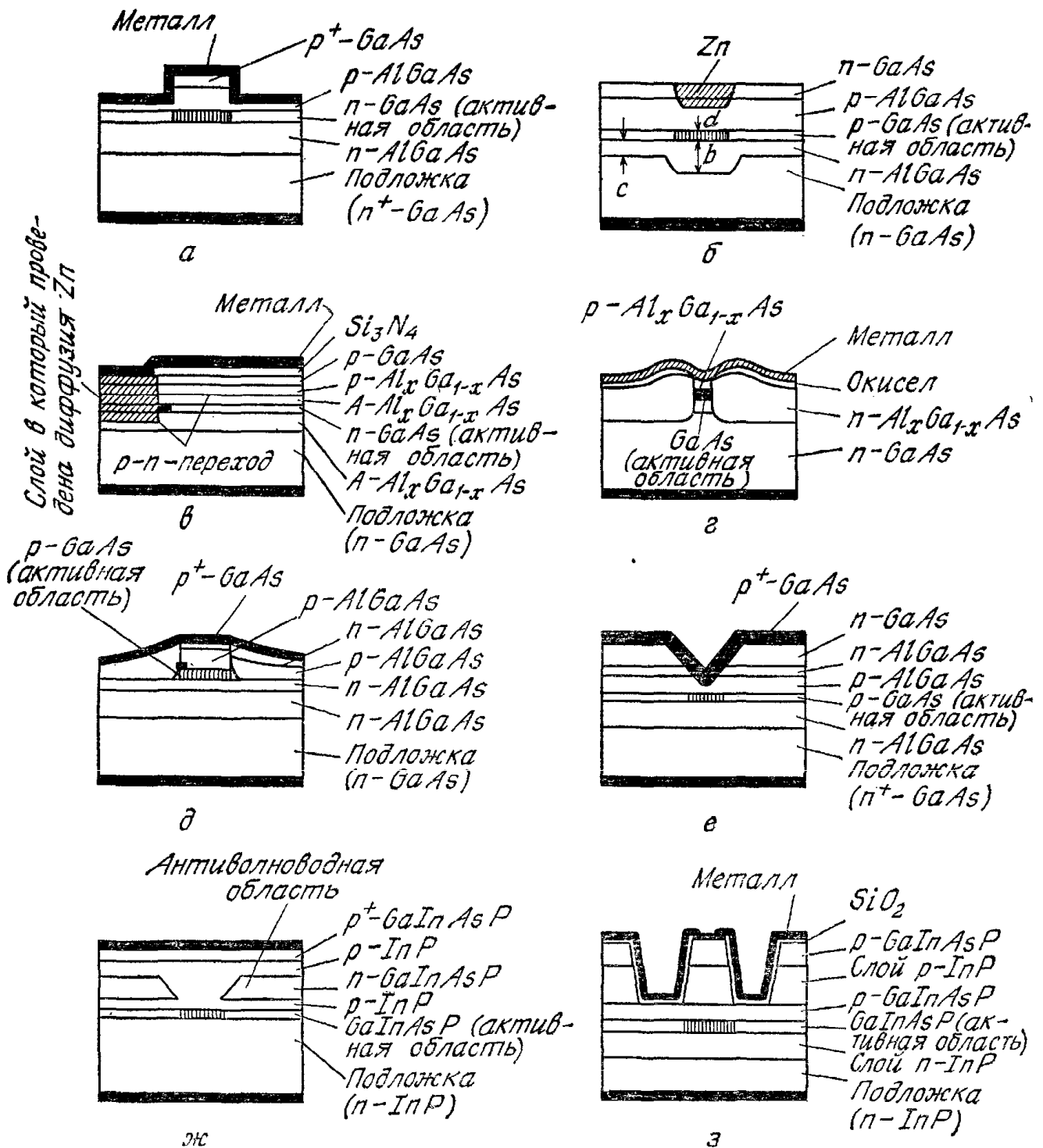


Рис. 44. Поперечные сечения различных гетеролазеров.

продольной модой на длине волны 0,861 мкм (разд. 12.5.2). На рис. 44, б показана структура гетеролазера с каналом в подложке [62]. Формирование канала в подложке проводят перед выращиванием эпитаксиальных слоев. Эта структура также обладает хорошими лазерными характеристиками. На рис. 44, в показана структура, в которой проведена диффузия Zn [63]. Она называется полосковым лазером с поперечным $p-n$ -переходом. Поскольку в этом случае инжекция идет из слоя $n-GaAs$ в слой $p-GaAs$, полученный диффузией Zn, то такой прибор фактически

представляет собой очень узкий гомолазер. В нем были получены низкие пороговые токи и генерация с одной продольной модой.

В зарощенном гетеролазере [64] активная область полностью окружена слоями из материалов с более широкой зоной и с меньшим показателем преломления (рис. 44, *г*). Этот лазер характеризуется малым сечением активной области, достигающим 1 мкм². В нем были получены очень низкие пороговые токи (до 15 мА) и почти симметричное распределение интенсивности в дальнем поле. Для создания зарощенного полоскового лазера [65] используется регулярная пятислойная эпитаксиальная структура с соответствующим содержанием Al и легирующих примесей (рис. 44, *д*). На слое n' -AlGaAs фотолитографическим способом формируется полоска шириной 10 мкм. Затем вместо стравленных областей наращиваются слои p - и n -AlGaAs. Ток ограничивается в центре полоски, так как восстановленный p — n -переход, окружающий центральную область, смещен в обратном направлении. Структура представляет собой четырехслойный лазер, который имеет линейную ватт-амперную характеристику и два симметричных выходных зеркала.

На рис. 44, *е* показана новая структура полоскового лазера, в котором для ограничения ширины полоски используется V-образная канавка [66]. V-канавка формируется травлением поверхности системы GaAs — AlGaAs; с помощью диффузии Zn прилегающий слой n -AlGaAs преобразуют под V-канавкой в слой p -типа. Глубина диффузии определяется профилем V-канавки. Была получена линейная ватт-амперная характеристика в сочетании с расходимостью светового луча в диапазоне 30—40°.

На рис. 44, *ж* показан самосовмещенный GaInAsP — InP-лазер [67]. Электронное и оптическое ограничения обеспечиваются в нем с помощью антиволноводных областей n -GaInAsP, которые расположены за пределами полосковой области и зарощены слоем p -InP, покрывающим активный слой. Носители инжектируются в узкую полосковую область, ограниченную обратносмещенными p — n — p — n -переходами. Прибор стабильно функционирует в режиме генерации основной моды с длиной волны $\sim 1,3$ мкм, причем вплоть до 72° генерация осуществляется в непрерывном режиме. На рис. 44, *з* показан ДГ-лазер с гребенчатым волноводом [68], излучающий на длине волны 1,55 мкм. Ширина полосковой области задается шириной гребенки, которая формируется методом ионного травления. Преимущество этой структуры состоит в том, что в процессе ее изготовления можно исключить повреждение активной области. Прибор также функционирует в режиме фундаментальной поперечной моды. Так как он излучает на длине волны 1,55 мкм, его можно использовать в качестве источника в оптических системах связи, работающих в одномодовом режиме и обладающих низкими потерями на больших расстояниях.

Во всех рассмотренных выше структурах обратная связь, необходимая для генерации лазерного излучения, обеспечивается за счет отражения на торцевых гранях резонатора, сформированных скалыванием или полировкой. Обратная связь может быть также получена путем периодического изменения показателя преломления внутри оптического волновода, достигаемого обычно гофрированием границы раздела между двумя диэлектрическими слоями. Пример такой структуры приведен на рис. 45 (вставка).

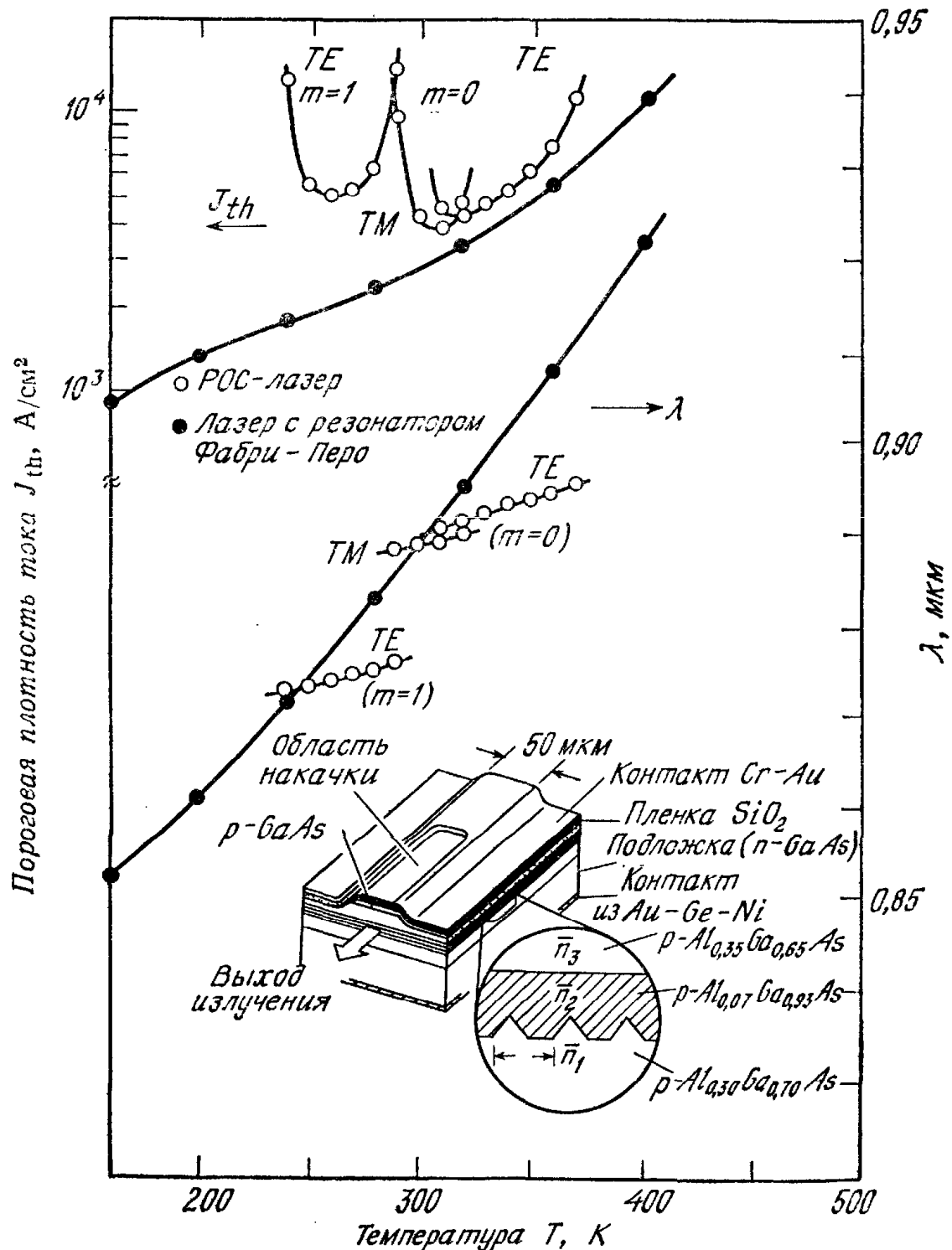


Рис. 45. Зависимость пороговой плотности тока и длины волны генерации от температуры p — n -перехода. На вставке показан лазер с распределенной обратной связью (РОС-лазер) [70].

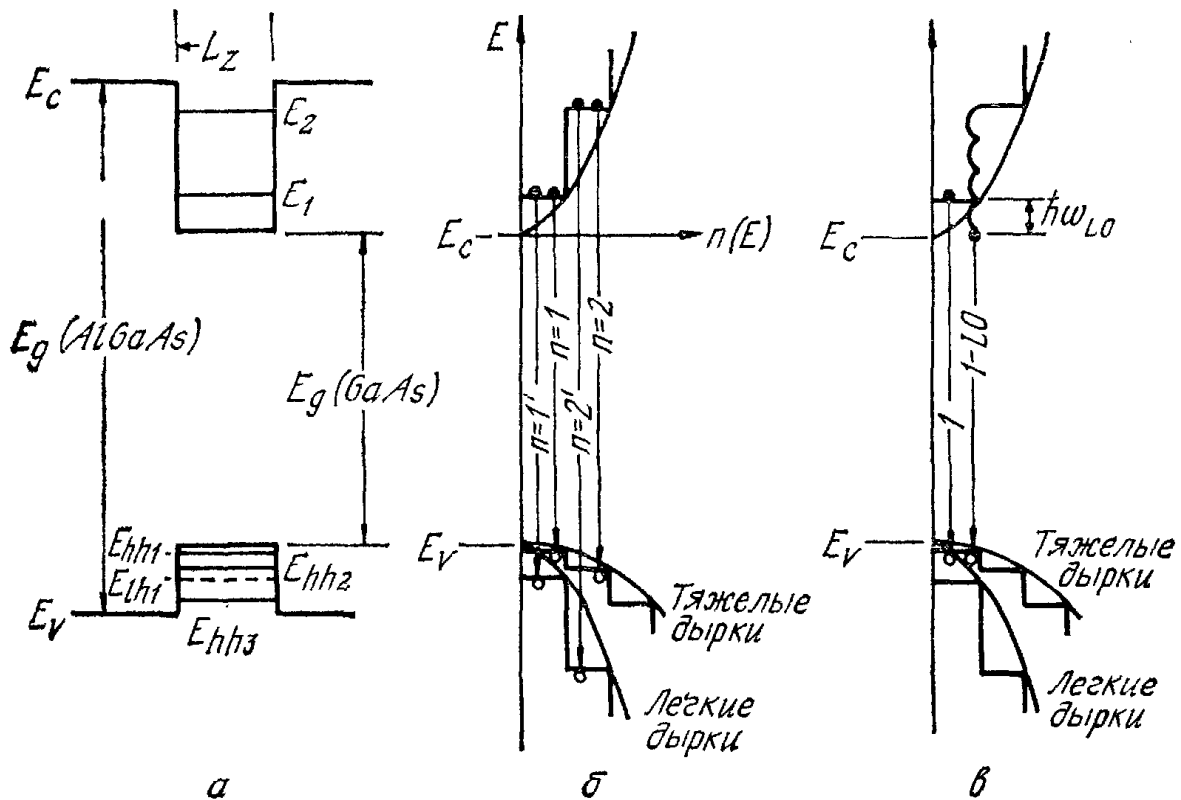


Рис. 46. Зонная диаграмма в гетероструктуре с квантовой ямой (а), диаграмма плотности состояний (б) и фоновое рассеяние в гетеролазере с квантовой ямой (в) [78].

Периодическое изменение \bar{n} , заданное конструкцией прибора, вызывает интерференцию. Лазеры, в которых используются такие гофрированные структуры, называются лазерами с распределенной обратной связью (РОС) или с распределенным брэгговским отражателем (РБО) [69]. Эти гетеролазеры могут с успехом применяться в интегральной оптике в качестве источников излучения. На рис. 45 приведены характеристики лазера с распределенной обратной связью и (для сравнения) лазера с резонатором Фабри — Перо [70]. В температурном диапазоне 150—400 К для РОС-лазера наблюдались две ТЕ-моды и одна ТМ-мода. Излучение лазера с резонатором Фабри — Перо связано с температурной зависимостью ширины запрещенной зоны, а изменение длины волны генерации РОС-лазера — с более слабой температурной зависимостью показателя преломления.

В настоящее время созданы полупроводниковые лазеры на основе структур со сверхрешеткой (гл. 2). Такие приборы называются гетеролазерами с квантовыми ямами [78]. Уменьшение толщины активного слоя ДГ-лазера до величины, сравнимой с длиной волны де Бройля для носителей ($\lambda = h/p$, где h — постоянная Планка, а p — импульс), приводит к двумерному квантованию и возникновению дискретных уровней. Энергия уровней зависит от величины потенциальных барьеров на границах конечной

прямоугольной ямы. На рис. 46, а приведена энергетическая диаграмма гетероструктуры $\text{Al}_x\text{Ga}_{1-x}\text{As} - \text{GaAs}$, в которой $L_z \approx 100 \text{ \AA}$. Уровни E_1 , E_2 представляют энергетические состояния электронов, E_{hh1} , E_{hh2} , E_{hh3} — состояния тяжелых дырок, E_{lh1} , E_{lh2} — легких дырок.

На рис. 46, б представлена соответствующая диаграмма плотности состояний. Полупараболы, представляющие собой края зоны проводимости E_C и валентной зоны E_V , соответствуют плотностям состояний объемного полупроводника. Ступенчатое изменение плотности состояний характеризует структуру с квантовыми ямами. Междузонные рекомбинационные переходы (правило отбора $\Delta n = 0$) происходят из основного состояния в зоне проводимости (например, E_1) в основное состояние в валентной зоне (например, E_{hh1}). Энергия перехода дается выражением

$$\hbar\omega = E_g(\text{GaAs}) + E_1 + E_{hh1}. \quad (56)$$

Такая рекомбинация в отличие от рекомбинации в объемном полупроводнике с непрерывным энергетическим спектром носителей, обусловленным параболическим изменением плотности состояний, осуществляется при наличии хорошо локализованных энергетических уровней. Рис. 46, в иллюстрирует другую важную особенность гетероструктур с квантовыми ямами, состоящую в том, что носители, инжектированные с большими энергиями, могут возбуждать фононы и рассеивать свою энергию, переходя в конечном счете на состояния с меньшей плотностью. В объемном полупроводнике процесс возбуждения фононов ограничивается уменьшением плотности состояний (особенно у края зоны), тогда как в системе с квантовыми ямами в пределах области с постоянной плотностью состояний такое ограничение отсутствует. В результате этого процесса энергия электрона может оказаться ниже уровня, обусловленного квантовым ограничением, например ниже уровня E_1 (рис. 46, в), что может привести к генерации лазерного излучения с энергией фотонов $\hbar\omega < E_g$ вместо энергии $\hbar\omega \geq E_g$, которая характеризует рекомбинацию без участия фононов.

Гетеролазеры с квантовыми ямами были изготовлены на основе структуры $\text{Al}_x\text{Ga}_{1-x}\text{As} - \text{GaAs}$ методом химического вакуумного осаждения из металлоорганики. Изготовленные лазеры имели полосковую геометрию, в качестве ограничительных областей применялись слои $\text{Al}_{0,4}\text{Ga}_{0,6}\text{As}$, а нелегированная активная область содержала шесть квантовых ям из слоев GaAs толщиной $L_z \approx 120 \text{ \AA}$, разделенных пятью барьерными слоями $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As}$ толщиной $\sim 120 \text{ \AA}$. Потенциальные преимущества лазеров с квантовыми ямами состоят в высокой квантовой эффективности, низком пороговом токе ($\sim 1 \text{ мА}$ и менее) и слабой чувствительности к изменениям температуры.

12.5.2. Выходная мощность и спектры излучения

На рис. 47 приведена типичная зависимость мощности ДГ-лазера при возрастании тока от низких значений, характерных для спонтанной эмиссии, до значений, превышающих порог лазерной генерации. Вставка на рисунке поясняет схему детектирования, которая использовалась для измерения интенсивности излучения лазерного диода. На начальном участке, соответствующем спонтанной эмиссии, интенсивность излучения (пропорциональная току детектора) медленно растет с увеличением тока, протекающего через диод, а после возбуждения лазерной генерации резко возрастает. Экстраполяция этой кривой к нулевому значению тока I_D , показанная штриховой линией, дает значение порогового тока. Плотность порогового тока J_{th} равна I_{th}/A ,

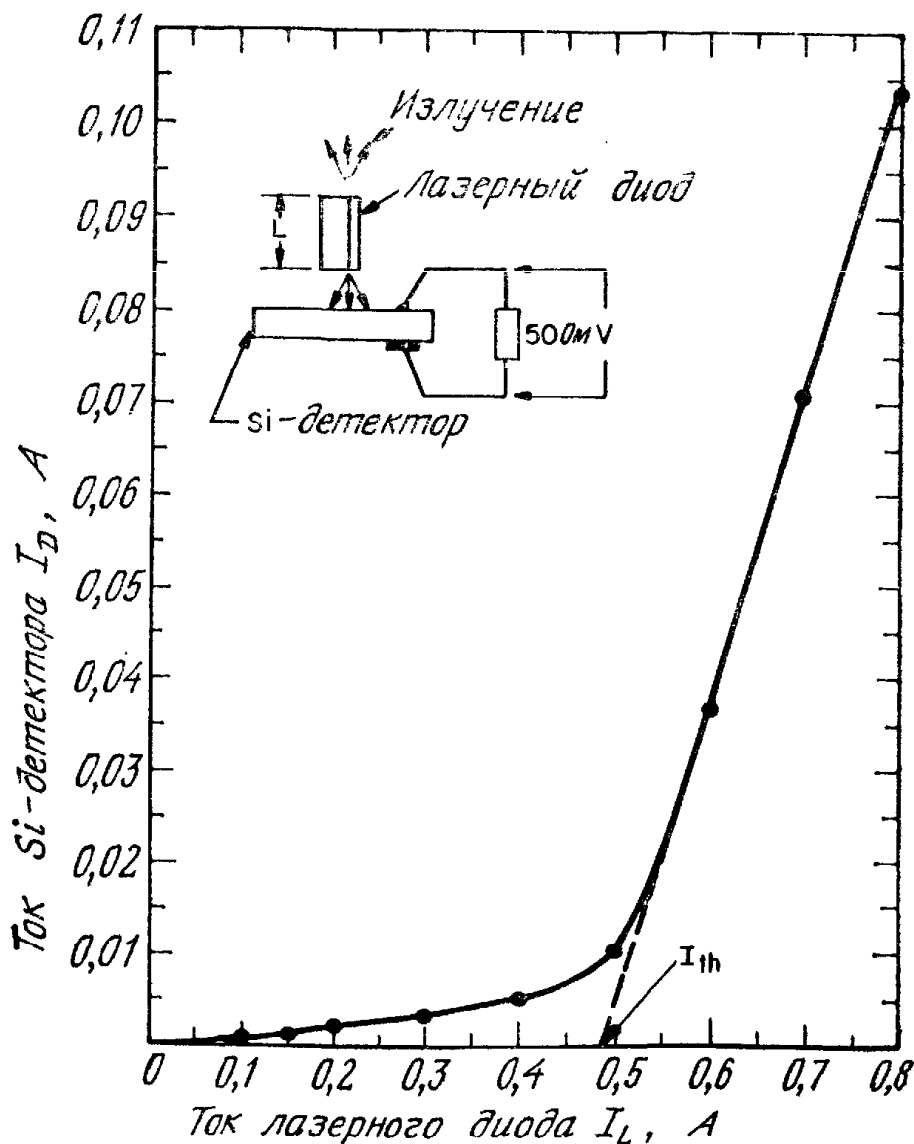


Рис. 47. Зависимость мощности излучения от тока накачки ДГ-лазера на GaAs — Al_xGa_{1-x}As при комнатной температуре. На вставке приведена схема измерения [20].

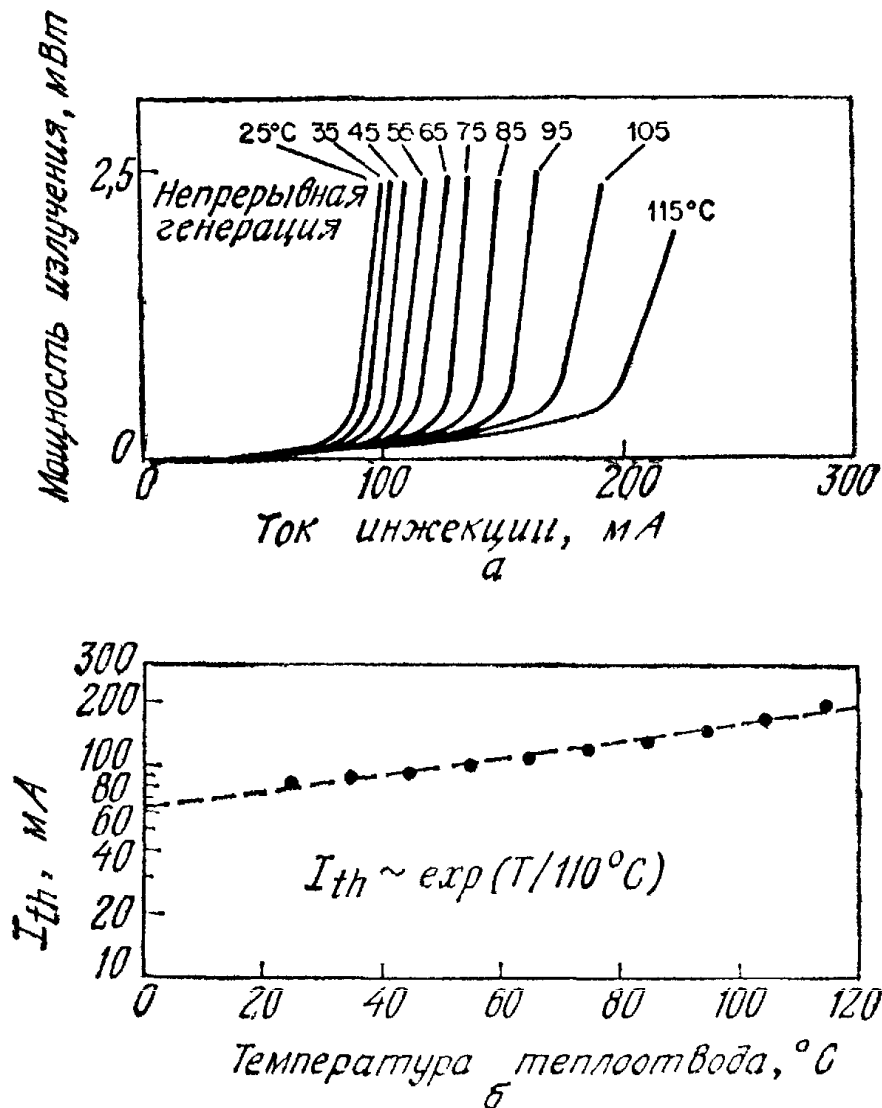


Рис. 48. Зависимость мощности излучения от тока накачки полоскового GaAs — $\text{Al}_x\text{Ga}_{1-x}\text{As}$ -гетеролазера с зарощенным активным слоем (а) и зависимость порогового тока в режиме непрерывной генерации от температуры (б) [65].

где A — активная площадь. Из графика на рис. 47 можно получить также дифференциальную квантовую эффективность

$$\eta_D \equiv \Delta I_D / \Delta I_L. \quad (57)$$

Для приведенного примера $\eta_D \approx 30\%$. Температурная зависимость порогового тока полоскового лазера с зарощенным активным слоем для режима непрерывной генерации (рис. 44, д) приведена на рис. 48. Зависимость мощности излучения от тока инжекции при непрерывной генерации (рис. 48, а) для различных температур теплоотвода в диапазоне 25—115°C почти идеально линейна. Зависимость порогового тока от температуры приведена на рис. 48, б. Пороговый ток экспоненциально возрастает с увеличением температуры:

$$I_{th} \sim \exp(T/T_0), \quad (58)$$

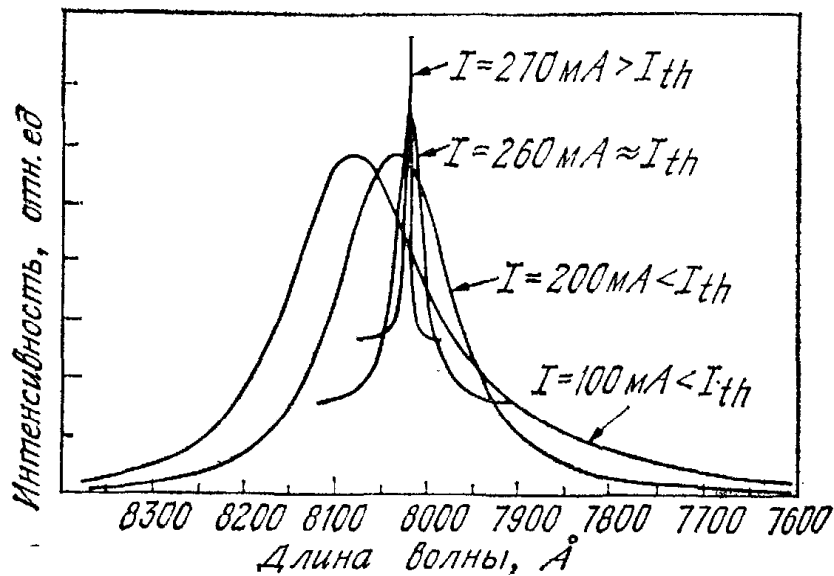


Рис. 49. Спектры излучения диодного лазера при токах накачки ниже порогового, вблизи порогового и выше порогового значений, иллюстрирующие эффект сужения полосы излучения при переходе к режиму лазерной генерации [21].

где T — температура теплоотвода, °С, а показатель T_0 , равный 110 °С, сравним по величине с аналогичным показателем для обычных ДГ-лазеров.

Режим спонтанной эмиссии, наблюдающийся при низких токах, характеризуется широким спектром излучения (полуширина спектра обычно составляет 100—500 Å). При возрастании тока до значений, близких к пороговому, спектр излучения становится уже. На рис. 49 приведены спектры для различных значений тока, иллюстрирующие эффект сужения полосы излучения при переходе к режиму лазерной генерации.

На рис. 50 показан спектр излучения полоскового ДГ-лазера, полученного протонной бомбардировкой, на основе структуры InP — GaInAsP. При токах, незначительно превышающих пороговый ток (145 мА), возникает довольно много линий излучения, которые разделены практически равными интервалами $\Delta\lambda \simeq 7,5$ Å. При более высоких значениях тока ($I = 155$ мА) наблюдается тенденция к преобладанию в спектре единственной линии вблизи 1,285 мкм, характерной для одномодовой генерации. Эти линии излучения соответствуют продольным модам. Выведем основные соотношения для этих мод.

Для структуры, изображенной на рис. 29, а, получим основное условие отбора мод в продольном направлении (z -направление), вытекающее из того, что на длине полупроводника L между отражающими плоскостями должно укладываться целое число полуволен:

$$m \left(\frac{\lambda}{2\bar{n}} \right) = L, \quad (59)$$

$$\text{или } m\lambda = 2L\bar{n}, \quad (60)$$

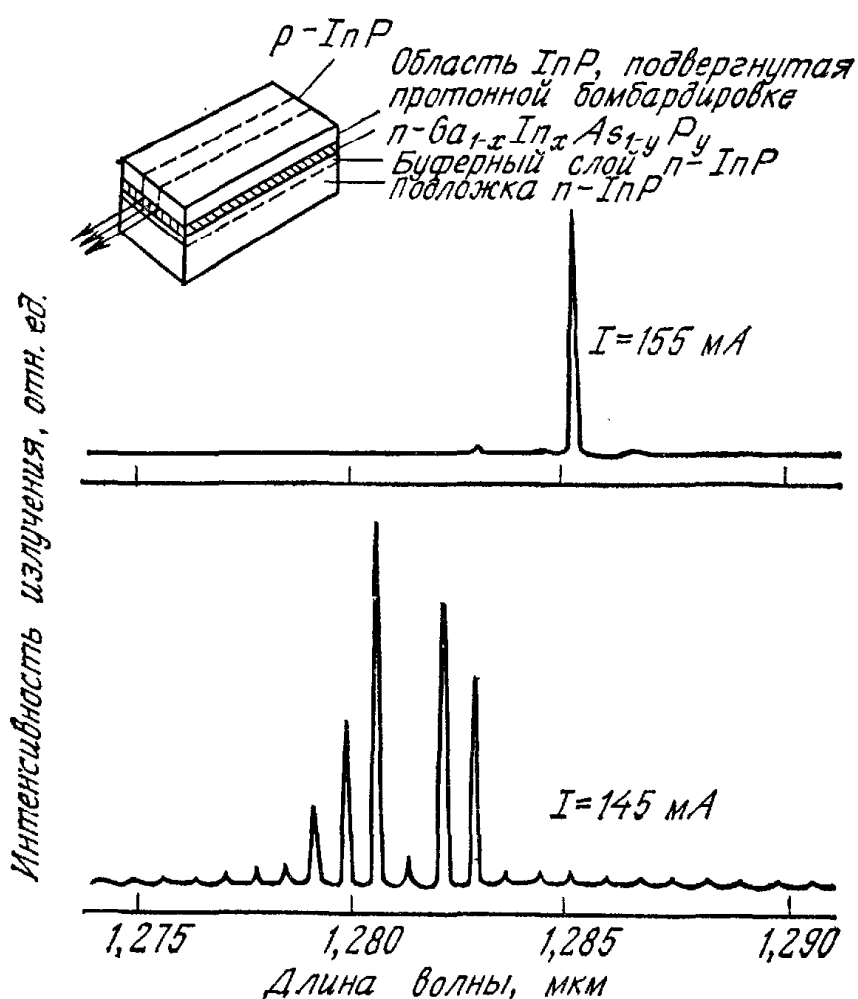


Рис. 50. Спектры излучения ДГ-лазера на $\text{InP} - \text{GaInAsP}$, измеренные с высоким разрешением [44].

где \bar{n} — показатель преломления среды, соответствующий длине волны λ . Интервал $\Delta\lambda$ между этими разрешенными продольными модами представляет собой разность в длинах волн, соответствующих целым числам m и $m + 1$. Дифференцируя уравнение (60) по λ , получим для больших m

$$\Delta\lambda = \frac{\lambda^2 \Delta m}{2\bar{n}L [1 - (\lambda/\bar{n})(d\bar{n}/d\lambda)]}. \quad (61)$$

Член в квадратных скобках связан с дисперсией. В газовых лазерах \bar{n} почти не зависит от λ и дисперсионный член дает лишь малую поправку. Спектральный интервал между продольными модами обратно пропорционален длине L . У полупроводниковых лазеров вследствие их малых размеров интервал $\Delta\lambda$ значительно больше, чем у газовых лазеров.

На рис. 51, а приведено относительное распределение интенсивности излучения гетеролазера с квантовыми ямами на основе $\text{Al}_{0,3}\text{Ga}_{0,7}\text{As} - \text{GaAs}$ (шесть квантовых и пять разделительных слоев, каждый толщиной $L_z = 120 \text{ \AA}$), работающего в режиме непрерывной генерации при $T = 300 \text{ K}$. Пик линии излучения на длине волны $0,8732 \text{ мкм}$ соответствует переходу с энергией

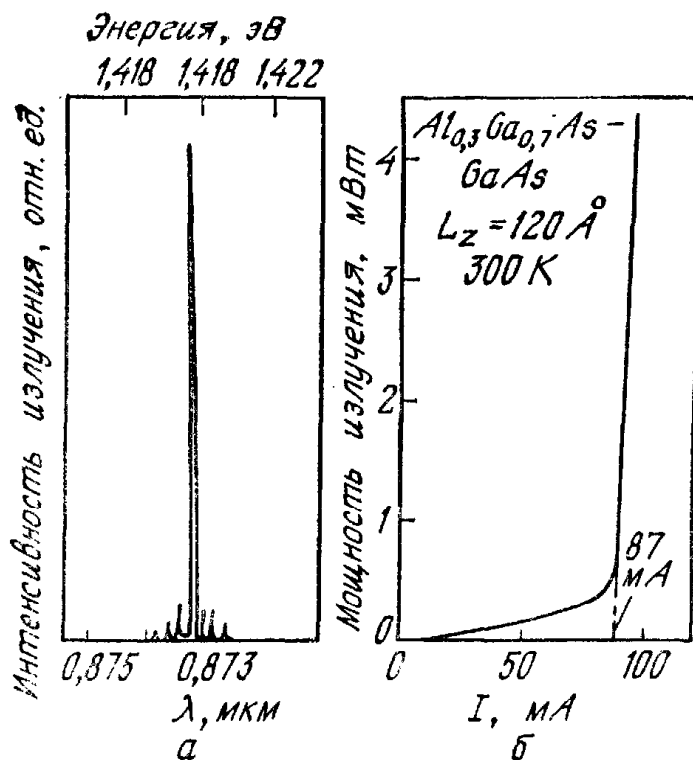


Рис. 51. Спектр излучения AlGaAs — GaAs-гетеролазера с квантовыми ямами, работающего в режиме непрерывной генерации при температуре 300 К (а), и зависимость мощности излучения от тока накачки (б) [78].

$\hbar\omega = (E_g + E_1 + E_{hh1} - \hbar\omega_{L0})$, где $\hbar\omega_{L0}$ — энергия продольного оптического фонона (рис. 46, в). На рис. 51, б приведена соответствующая ватт-амперная характеристика. Дифференциальная квантовая эффективность (уравнение (57)) равна 85 %. Температурная зависимость порогового тока описывается уравнением (58) с $T_0 = 220$ °С. Таким образом, лазеры с квантовыми ямами менее чувствительны к температурным изменениям, чем обычные ДГ-лазеры.

12.5.3. Задержка включения и частота модуляции

При пропускании через лазер импульса тока с амплитудой, достаточной для получения стимулированного излучения, генерация излучения обычно начинается с задержкой, равной нескольким наносекундам (вставка на рис. 52). Время задержки t_d связано с временем жизни неосновных носителей.

Для вычисления времени задержки рассмотрим уравнение непрерывности для электронов в полупроводнике p -типа. Предположим, что через активный слой d протекает ток I постоянной величины и концентрация инжектированных электронов n много больше равновесной концентрации. Тогда уравнение непрерывности примет вид

$$dn/dt = I/qAd - n/\tau_e, \quad (62)$$

где A — площадь, а τ_e — время жизни носителей. Первый член уравнения представляет собой постоянную скорость генерации, а второй — скорость рекомбинации. Аналогичное выражение

можно записать для дырок в активном слое n -типа. Решение этого уравнения при $n(0) = 0$ имеет вид

$$n(t) = (\tau_e I / qAd) [1 - \exp(-t/\tau_e)] \quad (63)$$

или

$$t = \tau_e \ln \left[\frac{I}{I - qn(t) Ad / \tau_e} \right]. \quad (64)$$

Когда $n(t)$ достигает порогового значения, $n(t) = n_{th}$ и $I_{th} = qn_{th} Ad / \tau_e$. Так как при $n(t) = n_{th}$, $t = t_d$, то выражение для времени задержки имеет вид

$$t_d = \tau_e \ln \left(\frac{I}{I - I_{th}} \right). \quad (65)$$

Если перед включением импульса через лазер пропускается ток, то решение уравнения (60) при начальном условии $n(0) = I_0 \tau_e / qAd$ будет иметь вид

$$t_d = \tau_e \ln [(I - I_0) / (I - I_{th})]. \quad (66)$$

Экспериментальные результаты, иллюстрирующие изменение задержки включения лазера при увеличении тока сверх его порогового значения при различных концентрациях акцепторной примеси в активном слое, приведены на рис. 52 [71]. В соответствии

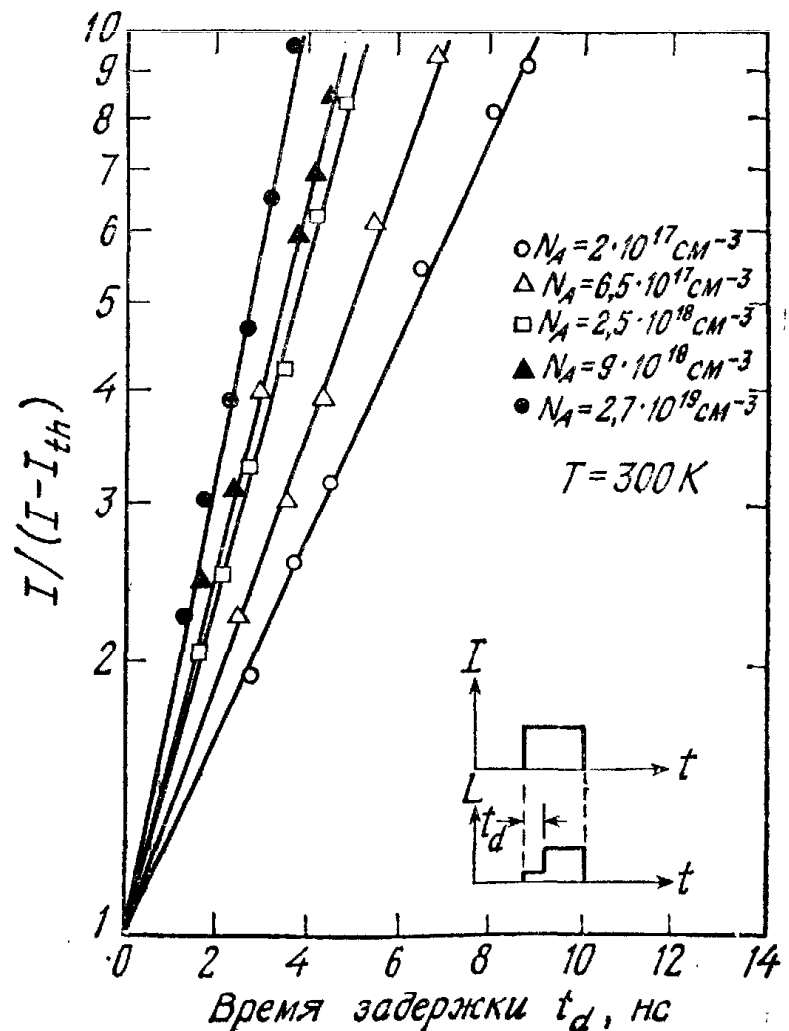


Рис. 52. Зависимость времени задержки от тока. Время задержки t_d показано на вставке [71].

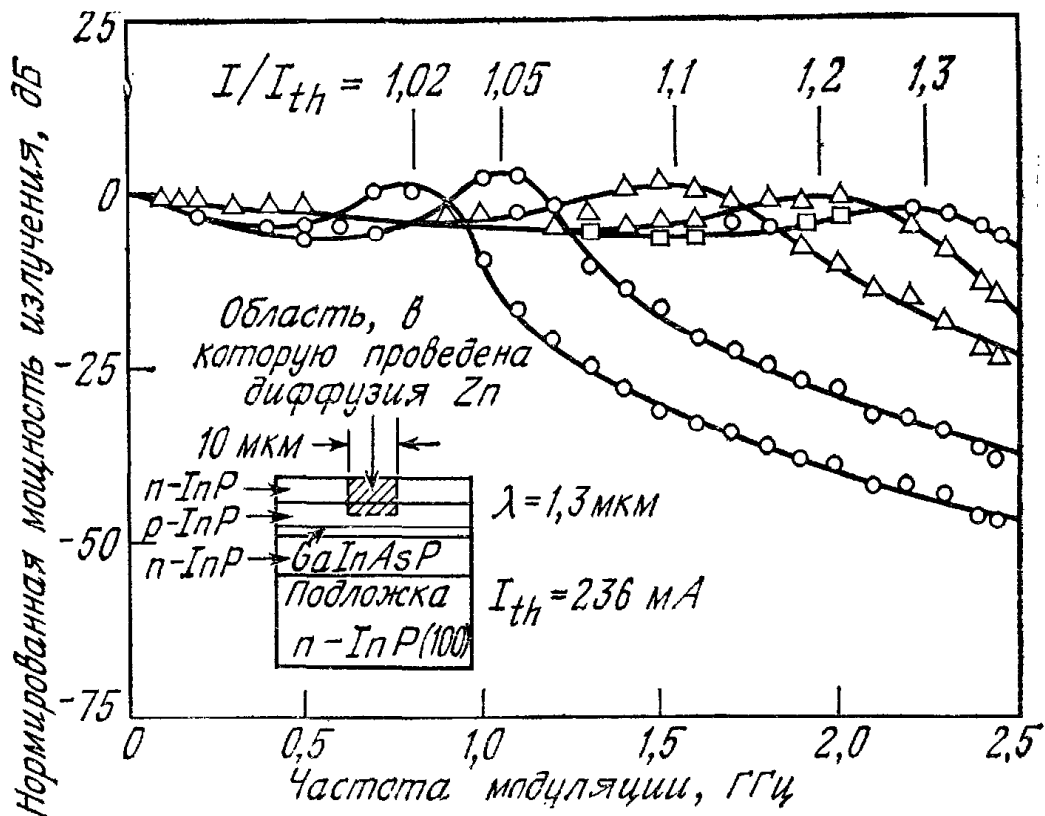


Рис. 53. Зависимость нормированной мощности излучения от частоты модуляции. На вставке показана структура лазера [72].

с уравнением (65) наблюдается логарифмическая зависимость времени задержки t_d от тока; с увеличением N_A время задержки уменьшается (поскольку уменьшается время жизни неосновных носителей).

Источники излучения, предназначенные для волоконно-оптических линий связи, должны обеспечивать возможность осуществления высокочастотной модуляции. ДГ-лазеры на основе GaAs и GaInP обладают хорошими модуляционными характеристиками в пределах гигагерцевого диапазона частот. На рис. 53 приведена зависимость нормированной выходной мощности модулированного излучения от частоты модуляции для ДГ-лазерного диода на основе GaInAsP—InP [72]. Диод излучает на длине волны 1,3 мкм и непосредственно модулируется синусоидальным током, наложенным на постоянный ток накачки. При токе накачки, превышающем значение порогового тока в 1,3 раза, мощность излучения на выходе практически не изменяется.

12.5.4. Лазерная перестройка [73]

Длину волны излучения полупроводникового лазера можно регулировать путем изменения тока диода или температуры теплоотвода, а также с помощью магнитного поля или давления. Такие достоинства полупроводниковых лазеров, как возможность перестройки длины волны узкой линии излучения, высокая стабиль-

ность, низкая потребляемая мощность, простота конструкции, открывают широкие перспективы их применения в промышленности и фундаментальных исследованиях, таких, как молекулярная и атомная спектроскопия, газовая спектроскопия высокого разрешения и контроль загрязнений атмосферы.

На рис. 23 показан диапазон длин волн, который перекрывают лазеры на основе полупроводниковых соединений. Выбирая соответствующий состав соединения для лазера, можно получить излучение с любой требуемой длиной волны из широкого спектрального диапазона 0,3—30 мкм.

Рис. 54 иллюстрирует температурную перестройку излучения ДГ-лазера на основе $\text{PbTe}-\text{Pb}_{1-x}\text{Sn}_x\text{Te}$ [74]. Регулируя температуру теплоотвода в пределах 10—120 К, можно изменять длину волны излучения примерно от 9 до 16 мкм. Это изменение в основном обусловлено эффектом расширения запрещенной зоны в соединениях типа $\text{A}^{\text{III}}\text{B}^{\text{VI}}$ при увеличении температуры.

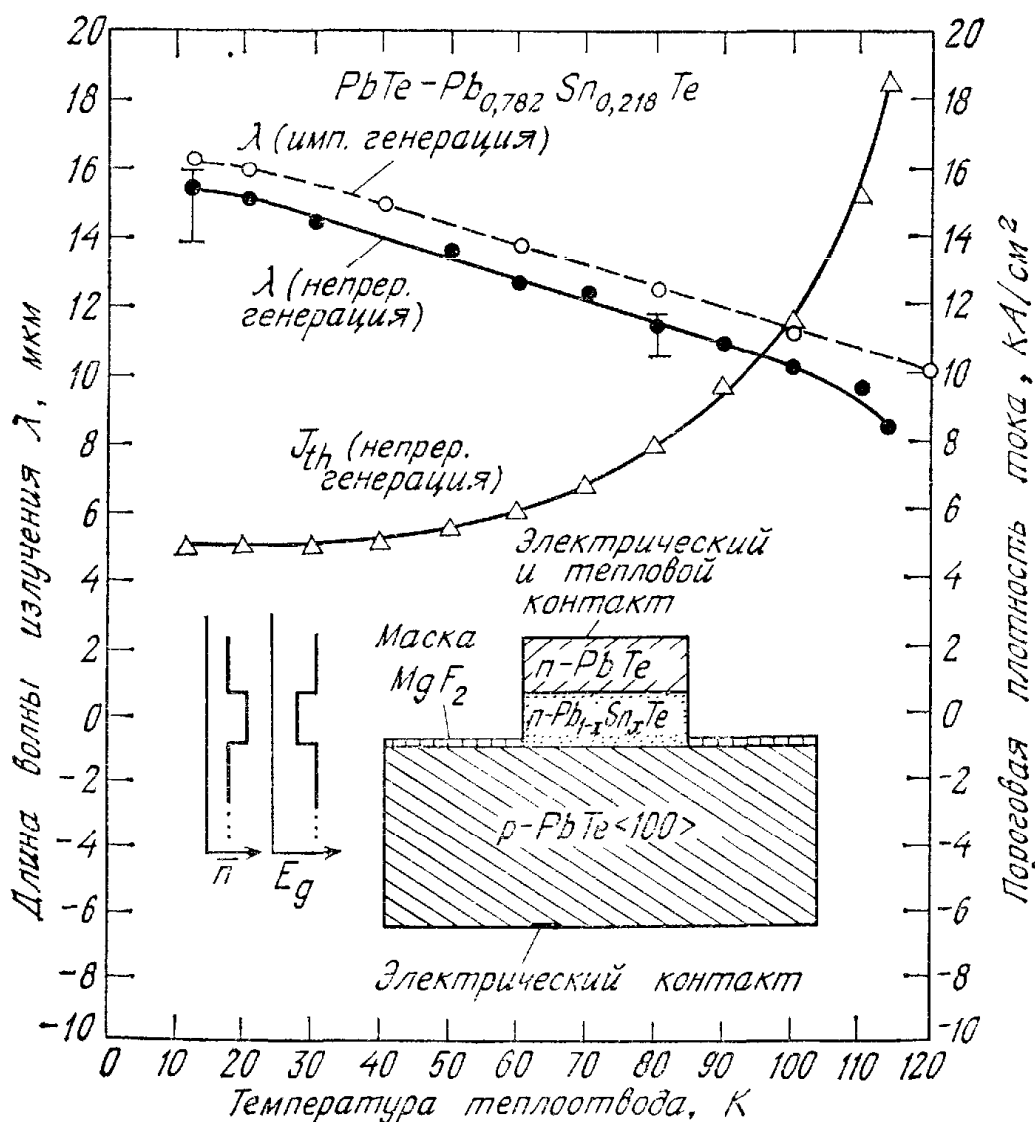


Рис. 54. Изменение длины волны излучения и порогового тока в зависимости от температуры. На вставке показана структура ДГ-лазера на $\text{PbTe}-\text{PbSnTe}$ [74].

Подвергая лазерный диод действию гидростатического давления, можно достичь смещения линии излучения в очень широком диапазоне. Изменение ширины запрещенной зоны в некоторых бинарных соединениях (например, InSb, PbS, PbSe) при изменении гидростатического давления происходит по линейному закону. За счет увеличения гидростатического давления до 14 кбар можно обеспечить перестройку линии излучения лазера на PbSe, работающего при 77 К, в диапазоне 7,5—22 мкм [73].

Диодные лазеры могут также перестраиваться под действием магнитного поля. Для полупроводников с сильной анизотропией эффективных масс положение магнитных энергетических уровней зависит от ориентации приложенного магнитного поля по отношению к оси кристалла. Как зона проводимости, так и валентная зона имеют свои энергии квантования уровней Ландау. При увеличении магнитного поля происходит увеличение энергетических интервалов между разрешенными переходами, вызывающее изменение длины волны излучения. Под действием магнитного поля 10 кГс, приложенного в направлении $\langle 100 \rangle$, лазер на основе $\text{Pb}_{0,79}\text{Sn}_{0,21}\text{Te}$ при $T = 7$ К перестраивается по длине волны излучения от 15 до 14 мкм.

12.5.5. Деградация лазеров [20]

Деградация инжекционных лазеров обусловлена целым рядом механизмов. Выделяют три основных типа деградации: 1) катастрофическое разрушение; 2) образование дефектов темных линий; 3) постепенная деградация.

Катастрофическая деградация происходит под действием больших мощностей излучения, приводящих к непрерывному повреждению зеркал лазера вследствие образования на их поверхности ямок и канавок. В результате модификаций структуры прибора, уменьшающих поверхностную рекомбинацию и поглощение, мощность излучения может увеличиться до предельного значения, при котором происходит разрушение лазера [75].

Дефекты темных линий представляют собой сетку дислокаций, которая может формироваться в процессе работы лазера и внедряться внутрь резонатора. Появившись, она может широко разрастись в течение нескольких часов и вызвать увеличение плотности порогового тока. Для уменьшения вероятности образования дефектов темных линий следует использовать высококачественные эпитаксиальные слои, выращенные на подложках с низкой плотностью дислокаций, а также обеспечивать тщательное соединение лазера с теплоотводом для минимизации напряжений.

ДГ-лазеры могут работать в течение длительного времени с относительно медленной деградацией, если устранить условия,

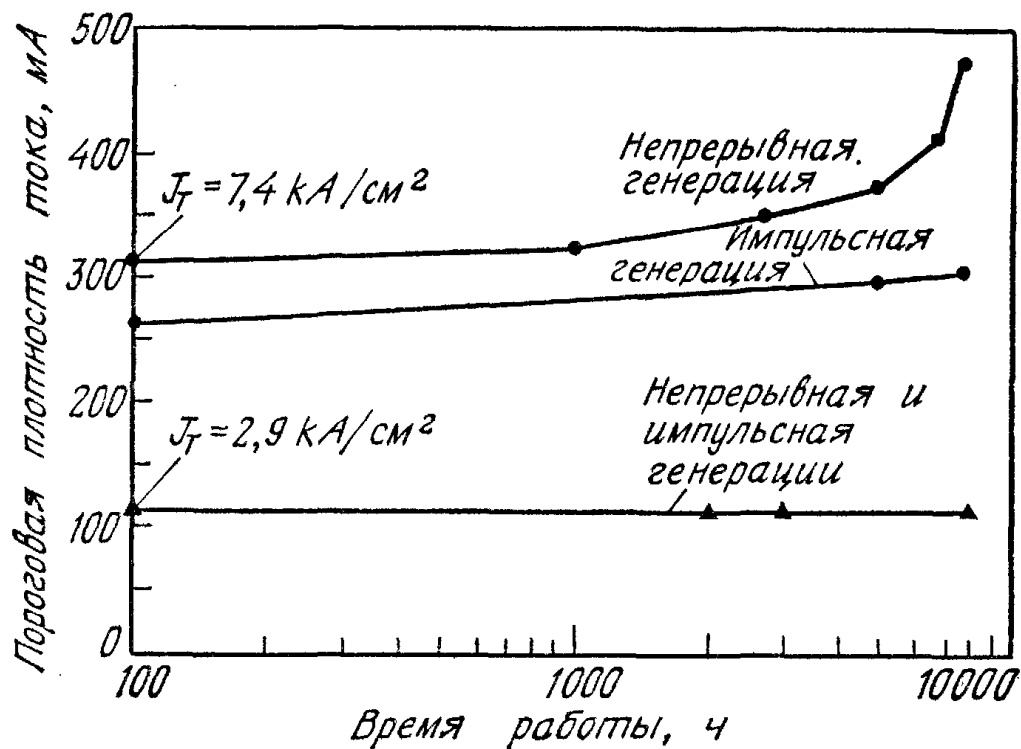


Рис. 55. Зависимость порогового тока от времени работы ДГ-лазера на InP — GaInAsP в режиме непрерывной генерации [44].

которые вызывают мгновенное катастрофическое разрушение и быструю деградацию, обусловленную образованием дефектов темных линий. Так, например, в ДГ-лазерах на основе GaAs—AlGaAs, работающих в режиме непрерывной генерации в течение более $2,6 \cdot 10^4$ ч (3 года) при температуре 30°C , не обнаруживалось признаков деградации [76]. Экстраполяция этих данных на температуру теплоотвода 22°C дает срок службы более 10^6 ч (около 100 лет). На рис. 55 приведена зависимость порогового тока от времени работы в режиме непрерывной генерации для ДГ-лазеров на основе GaIn—AsP—InP [44]. После 8000 ч работы характеристики лазера с низким значением порогового тока ($J_{th} = 2,9 \text{ kA/cm}^2$) остались прежними, а в лазере с высоким пороговым током существенно увеличилось тепловое сопротивление, но признаки внутренней деградации отсутствовали. Имеются все основания полагать, что срок службы, достигнутый для ДГ-лазеров на GaAs, может быть также получен и для лазеров, излучающих в более длинноволновой области. Лазеры с длительным сроком службы необходимы для волоконно-оптических систем дальней связи, а также для ряда других систем.

ЛИТЕРАТУРА

1. Round H. J. A Note on Carborundum, *Electron World*, 19, 309 (1907).
2. Bergh A. A., Dean P. J. Light-Emitting Diodes, Clarendon, Oxford, 1976. [имеется русский перевод: Берг А., Дин П. Светодиоды, М.: «Мир», 1979.]
3. Gage S., Evans D., Hodapp M., Sorenson H. Optoelectronics Applications Manual McGraw-Hill, N. Y., 1977.

4. Gordon J. P., Zeiger H. J., Townes C. H., Molecular Microwave Oscillator and New Hyperfine Structure in the Microwave Spectrum of NH_3 , *Phys. Rev.*, 95, 282 (1954).
5. Schawlow A. L., Townes C. H. Infrared and Optical Masers, *Phys. Rev.*, 112, 1940 (1958).
6. Maiman T. H. Stimulated Optical Radiation in Ruby Masers, *Nature (Lond.)*, 187, 493 (1960).
7. Aigrain P. (1958), as reported in Proc. Conf. Quantum Electron., Paris, 1963, p. 1762.
8. Басов Н. Г., Вул Б. М., Попов Ю. М. Квантовомеханические полупроводниковые генераторы и усилители электромагнитных колебаний, ЖЭТФ (1960).
9. Nishizawa J. I., Watanabe Y. Japanese Patent (Apr. 1957).
10. Boyle W. S., Thomas D. G., U. S. Patent 3059117 (Oct. 16, 1962).
11. Bernard M. G. A., Duraffourg G. Laser Conditions in Semiconductors, *Phys. Status Solidi*, 1, 699 (1961).
12. Dumke W. P. Interband Transitions and Maser Action, *Phys. Rev.*, 127, 1559 (1962).
13. Hall R. N., Genner G. E., Kingsley J. D., Soltys T. J., Carlson R. O. Coherent Light Emission from GaAs Junctions, *Phys. Rev. Lett.*, 9, 366 (1962).
14. Nathan M. I., Dumke W. P., Burns G., Dill F. J., Jr., Lasher G. J. Stimulated Emission of Radiation from GaAs $p-n$ Junction, *Appl. Phys. Lett.*, 1, 62 (1962).
15. Quist T. M., Rediker R. H., Keyers R. J., Krag W. E., Lax B., McWhorter A. L., Zeigler H. J. Semiconductor Maser of GaAs, *Appl. Phys. Lett.*, 1, 91 (1962).
16. Holonyak N., Jr., Bevacqua S. F. Coherent (Visible) Light Emission from $\text{Ga}(\text{As}_{1-x}\text{P}_x)$ Junction, *Appl. Phys. Lett.*, 1, 82 (1962).
17. Hayashi I., Panish M. B., Foy P. W., Sumski S. Junction Lasers which Operate Continuously at Room Temperature, *Appl. Phys. Lett.*, 17, 109 (1970).
18. Kroemer H. A Proposed Class of Heterojunction Injection Lasers, *Proc. IEEE*, 51, 1782 (1963).
19. Алферов Ж. И., Казаринов Р. Ф. Авторское свидетельство № 181737 (1963), Алферов Ж. И. ФТП, т. 1, стр. 436 (1967).
20. Casey H. C., Jr., Panish M. B. Heterostructure Lasers, Academic, N. Y., 1978. [имеется русский перевод: Кейси Х., Панши М. Лазеры на гетероструктурах, М., Мир, 1981.]
21. Kressel H., Butler J. K. Semiconductor Lasers and Heterojunction LEDs, Academic, N. Y., 1977.
22. Goldberg P. Ed., Luminiscence of Inorganic Solids, Academic, N. Y., 1966.
23. Ivey H. F. Electroluminescence and Semiconductor Lasers, *IEEE J. Quantum Electron.*, QE-2, 713 (1966).
24. Mooradian A., Fan H. Y. Recombination Emission in InSb, *Phys. Rev.*, 148, 873 (1966).
25. Carr W. N. Characteristics of a GaAs Spontaneous Infrared Source with 40 Percent Efficiency, *IEEE Trans. Electron Devices*, ED-12, 531 (1965).
26. Ivey H. F. Electroluminescence and Related Effects, Suppl. 1 to Marton L., Ed., Advances in Electronics and Electron Physics, Academic, N. Y., 1963, p. 205.
27. Wang S. Solid-State Electronics, McGraw-Hill, N. Y., 1966.
28. Eastman P. C., Haering R. R., Barnes P. A. Injection Electroluminescence in Metal — Semiconductor Tunnel Diodes, *Solid State Electron.*, 7, 879 (1964).
29. Gooch C. N. Injection Electroluminescent Devices, Wiley, N. Y., 1973.
30. Craford M. G. Recent Development in LED Technology, *IEEE Trans. Electron Devices*, ED-24, 935 (1977).
31. Groves W. O., Herzog A. H., Craford M. G. The Effect of Nitrogen Doping on GaAsP Electroluminescent Diodes, *Appl. Phys. Lett.*, 19, 184 (1971).

32. Carr W. N., Puttman G. E. One Watt GaAs $p - n$ Junction Infrared Source, *Appl. Phys. Lett.*, **3**, 173 (1963).
33. Carr W. N. Photometric Figures of Merit for Semiconductor Luminescent Sources Operating in Spontaneous Mode, *Infrared Phys.*, **6**, 1 (1966).
34. Galginaitis S. V. Improving the External Efficiency of Electroluminescent Diodes, *J. Appl. Phys.*, **36**, 460 (1965).
35. Wight D. R. Green Luminescence Efficiency in GaP, *J. Phys. D*, **10**, 431 (1977).
36. Von Munch W. Silicon Carbide Technology for Blue Emitting Diodes, *J. Electron. Mater.*, **6**, 449 (1977).
37. Geusic J. E., Ostermayer F. W., Marcos H. M., Van Uitert L. G., Van Der Ziel J. P. Efficiency of Red, Green and Blue Infrared-to-Visible Conversion Sources, *J. Appl. Phys.*, **42**, 1958 (1971).
38. Dentai A. G., Lee T. P., Burrus C. A. Small-Area High-Radiance cw InGaAsP LEDs Emitting at 1.2 to 1.3 μ m, *Electron Lett.*, **13**, 484 (1977).
39. Burrus C. A., Miller B. I. Small-Area, Double Heterostructure AlGaAs Electroluminescent Diode Source for Optical-Fiber Transmission Lines, *Opt. Commun.*, **4**, 307 (1971).
40. Ettenberg M., Kressel H., Wittke J. P. Very High Radiance Edge-Emitting LED, *IEEE J. Quantum Electron.*, **QE-12**, 360 (1976).
41. Amann M. C., Proebster W. Small-Area GaAs — GaAlAs Heterostructure LED with Improved Current Confinement, *Electron. Lett.*, **15**, 599 (1979).
- 41a. Hall R. N., *Proc. Inst. Electr. Eng. Lond.*, **B106**, 923 (1959).
42. Ikeda K., Horiuchi S., Tanaka T., Susaki W. Design Parameters of Frequency Response of GaAs — AlGaAs DH LED's for Optical Communications, *IEEE Trans. Electron Devices*, **ED-24**, 10001 (1977).
43. Miya T., Terunuma Y., Hosaka T., Miyashita T. Ultimate Low-Loss Single Mode Fiber at 1.55 μ m, *Electron. Lett.*, **15**, 108 (1979).
44. Foyt A. G. 1.0—1.6 μ m Sources and Detectors for Fiber Optics Applications, *IEEE Device Res. Conf.*, Boulder, Colo., June 25, 1979.
45. Goodman C. H. L. Devices and Materials for 4 μ m-band Fibre-Optical Communication, *Solid State Electron Devices*, **2**, 129 (1978).
46. Lin P. J., Kleinman L. Energy Bands of PbTe, PbSe and PbS, *Phys. Rev.*, **142**, 478 (1966).
47. Levine A. K. Lasers, *Am. Sci.*, **51**, 14 (1963).
48. Panish M. B., Hayashi I., Sumski S. Double-Heterostructure Injection Lasers with Room Temperature Threshold as Low as 2300 A/cm², *Appl. Phys. Lett.*, **16**, 326 (1970).
49. Casey H. C., Jr., Panish M. B., Merz J. L. Beam Divergence of the Emission from Double-Heterostructure Injection Lasers, *J. Appl. Phys.*, **44**, 5470 (1973).
50. D'Asaro L. A. Advances in GaAs Junction Lasers with Stripe Geometry, *J. Lumin.*, **7**, 310 (1973).
51. Paoli T. L. Waveguiding in a Stripe-Geometry Junction Laser, *IEEE J. Quantum Electron.*, **QE-13**, 662 (1977).
52. Yonezu H., Sakuma I., Kobayashi K., Kamejima T., Ueno M., Nannichi Y. A GaAs — Al_xGa_{1-x}As Double Heterostructure Planar Stripe Laser, *Jpn. J. Appl. Phys.*, **12**, 1585 (1973).
53. Nathan M. I. Semiconductor Lasers, *Proc. IEEE*, **54**, 1276 (1966).
54. Lasher G., Stern F. Spontaneous and Stimulated Recombination Radiation in Semiconductors, *Phys. Rev.*, **133**, A553 (1964).
55. Halperin B. I., Lax M. Impurity-Band Tails in the High-Density Limit. I. Minimum Counting Methods, *Phys. Rev.*, **148**, 722 (1966).
56. Stern F. Calculated Spectral Dependence of Gain in Excited GaAs, *J. Appl. Phys.*, **47**, 5382 (1976).
57. Casey H. C., Jr., Room Temperature Threshold-Current Dependence of GaAs — Al_xGa_{1-x}As Double Heterostructure Lasers on x and Active-Layer Thickness, *J. Appl. Phys.*, **49**, 3684 (1978).

58. Nahory R. E., Pollack M. A. Threshold Dependence on Active-Layer Thickness in InGaAsP/InP DH Lasers, *Electron. Lett.*, **14**, 727 (1978).
59. Yana M., Nishi H., Takusagawa M. Theoretical and Experimental Study of Threshold Characteristics in InGaAsP/InP DH Lasers, *IEEE J. Quantum Electron.*, **QE-15**, 571 (1979).
60. Burrus C. A., Casey H. C., Jr., Li T. Y. Optical Sources, in Miller S. E., Chynoweth A. G., Eds., *Optical Fiber Communication*, Academic, N. Y., 1979.
61. Amann M. C. New Stripe-Geometry Laser with Simplified Fabrication Process, *Electron. Lett.*, **15**, 441 (1979).
62. Aiki K., Nakamura M., Kuroda T., Umeda J. Channeled-Substrate Planar Structure AlGaAs Injection Laser, *Appl. Phys. Lett.*, **30**, 649 (1977).
63. Namizaki H. Transverse-Junction-Stripe Lasers with a GaAs $p-n$ Homojunction, *IEEE J. Quantum Electron.*, **QE-11**, 427 (1975).
64. Tsukada T. GaAs — AlGaAs Buried-Heterostructure Injection Lasers. *J. Appl. Phys.*, **45**, 4899 (1974).
65. Tsang W. T., Logan R. A., Van der Ziel J. P. Low-Current-Threshold Stripe-Buried-Heterostructure Lasers with Self-Aligned Current Injection Stripes, *Appl. Phys. Lett.*, **34**, 644 (1979).
66. Gosch J. Simple Technology Irons Out Kinks in Laser Diode Output, *Electronics* **52**, (26), 59 (1979).
67. Nishi H., Yano M., Nishitani Y., Akita Y., Takusagawa M. Self-Aligned Structure InGaAsP/InP Lasers, *Appl. Phys. Lett.*, **35**, 232 (1979).
68. Kaminow I. P., Nahory R. E., Pollack M. A., Stulz L. W., Dewinter J. C. Single-Mode cw Ridge-Waveguide Laser Emitting at 1.55 μ m, *Electron. Lett.*, **15**, 763 (1979).
69. Kogelnik H., Shank C. V. Stimulated Emission in a Periodic Structure, *Appl. Phys. Lett.*, **18**, 152 (1971); Coupled-Wave Theory of Distributed Feedback Lasers, *J. Appl. Phys.*, **43**, 2327 (1973).
70. Aiki K., Nakamura M., Umeda J. Lasing Characteristics of Distributed-Feedback GaAs — GaAlAs Diode Lasers with Separate Optical and Carrier Confinement, *IEEE J. Quantum Electron.*, **QE-12**, 597 (1976).
71. Hwang C. J., Dymont J. C. Dependence of Threshold and Electron Lifetime on Acceptor Concentration in GaAs — Ga_{1-x}Al_xAs Lasers, *J. Appl. Phys.*, **44**, 3240 (1973).
72. Akiba S., Sakai K., Yamamoto T. Direct Modulation of InGaAsP/InP Double Heterostructure Lasers, *Electron. Lett.*, **14**, 197 (1978).
73. Melngailis I., Mooradian A. Tunable Semiconductor Diode Lasers and Applications, in Jacobs S., Sargent M., Scott J. F., Seully M. O., Eds., *Laser Applications to Optics and Spectroscopy*. Addison-Wesley Reading, Mass., 1975.
74. Walpole J. N., Calawa A. R., Harman T. C., Groves S. H. Double-Heterostructure PbSnTe Lasers Grown by Molecular-Beam Epitaxy with CW Operation up to 114 K, *Appl. Phys. Lett.*, **28**, 552 (1976).
75. Yonezu H., Sakuma I., Kamojima T., Ueno M., Iwamoto K., Hino I., Hayashi I. High Optical Power Density Emission from a Window Stripe AlGaAs DH Laser, *Appl. Phys. Lett.*, **34**, 637 (1979).
76. Hartman R. L., Schumaker N. E., Dixon R. W. Continuously Operated AlGaAs DH Lasers with 70 °C Lifetimes as Long as Two Years, *Appl. Phys. Lett.*, **31**, 756 (1977).
77. Lo W., Swets D. E. Room Temperature 4.5 μ m Light Emitting Diodes, *Appl. Phys. Lett.*, **36**, 450 (1980).
78. Holonyak N., Jr., Kolbas R. M., Dupuis R. D., Dapkus P. D. Quantum-Well Heterostructure Lasers, *IEEE J. Quantum Electron.*, **QE-16**, 170 (1980).

13.1. ВВЕДЕНИЕ

Фотодетекторы представляют собой полупроводниковые приборы, которые могут детектировать оптические сигналы посредством протекающих в них электронных процессов. Расширение спектрального диапазона излучения когерентных и некогерентных источников света в дальнюю инфракрасную область, с одной стороны, и в ультрафиолетовую — с другой, привело к увеличению потребности в быстродействующих чувствительных фотодетекторах. В общем случае в фотодетекторе протекают три основных процесса: 1) генерация носителей под действием внешнего излучения; 2) перенос носителей и их умножение за счет того или иного механизма усиления тока, характерного для данного полупроводника; 3) взаимодействие тока с внешней цепью, обеспечивающее получение выходного сигнала.

Фотодетекторы играют важную роль в волоконно-оптических системах связи ближнего ИК-диапазона (0,8—1,6 мкм). Они детектируют оптические сигналы, т. е. преобразуют изменения интенсивности излучения в электрические колебания, которые соответствующим образом усиливаются и подвергаются последующей обработке. Фотодетекторы, предназначенные для таких систем, должны обладать высокой чувствительностью в рабочем диапазоне длин волн, высоким быстродействием и низким уровнем шумов. Кроме того, они должны иметь малые размеры, низкие управляющие напряжения или токи, а также высокую надежность в работе.

Быстродействующие фотодетекторы рассмотрены в работах [1—3]. В работе [4] имеется исчерпывающий обзор инфракрасных фотодетекторов, в котором подробно рассмотрены лавинные фотодиоды [5]. Вопросам использования фотодетекторов в составе волоконно-оптических систем посвящена работа [6].

13.2. ФОТОРЕЗИСТОР

Фоторезистор представляет собой пластину полупроводника (монокристаллическую или в виде тонкой пленки), на противоположных концах которой созданы омические контакты (рис. 1). При падении излучения на поверхность фоторезистора в последнем генерируются носители вследствие возбуждения либо междузонных переходов (собственное возбуждение), либо переходов с участием

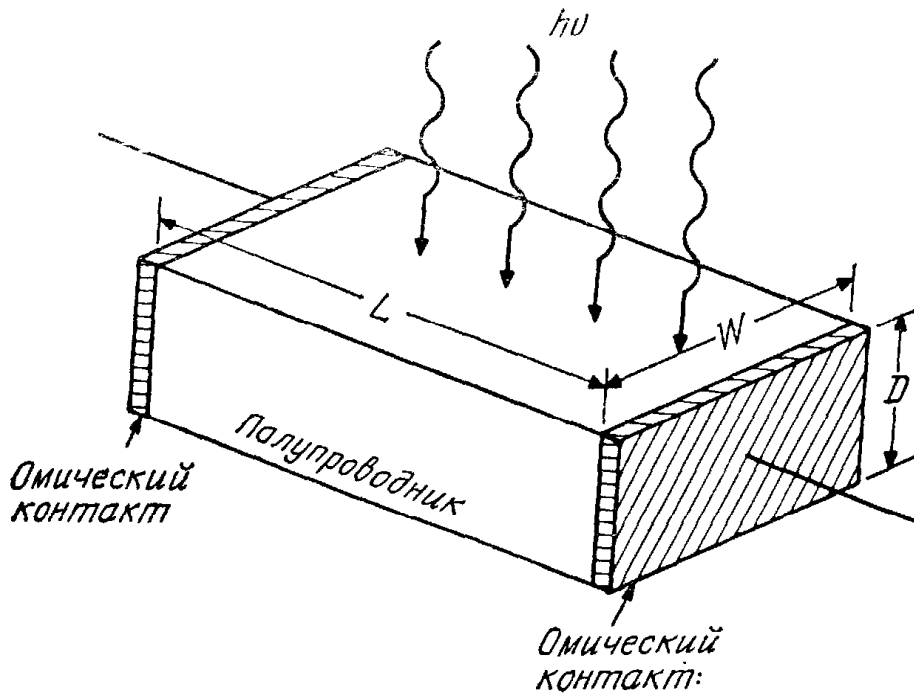


Рис. 1. Фоторезистор, состоящий из пластины полупроводника с омическими контактами.

энергетических уровней в запрещенной зоне (примесное возбуждение), что приводит к увеличению проводимости. Процессы собственного и примесного фотовозбуждения носителей схематически показаны на рис. 2.

Проводимость собственных фоторезисторов описывается формулой $\sigma = q(\mu_n n + \mu_p p)$, и увеличение проводимости под действием освещения в основном связано с увеличением числа носителей. Длинноволновая граница фотопроводимости в этом случае определяется из соотношения

$$\lambda_c = \frac{hc}{E_g} = \frac{1,24}{E_g (\text{эВ})}, \quad (1)$$

где λ_c — длина волны, соответствующая ширине запрещенной зоны E_g . Излучение с длинами волн, меньшими λ_c , поглощается в полупроводнике с образованием электронно-дырочных пар.

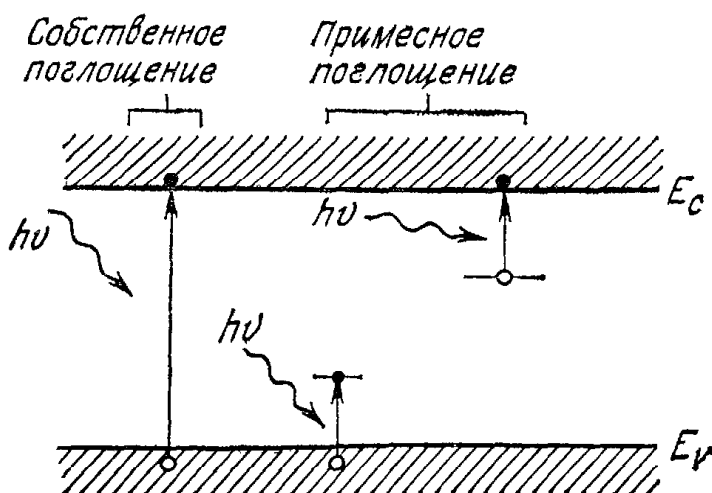


Рис. 2. Процессы собственного (зона — зона) и примесного фотовозбуждения.

В примесном фоторезисторе фотовозбуждение может происходить между краем зоны и энергетическим уровнем в запрещенной зоне. Фотопроводимость может возникать в результате поглощения фотонов с энергией, которая равна или превышает энергию, отделяющую примесный уровень от зоны проводимости или валентной зоны. При этом длинноволновая граница фотопроводимости определяется глубиной залегания примесного уровня в запрещенной зоне.

Работа фотодетекторов вообще и фоторезисторов в частности характеризуется тремя параметрами: 1) квантовой эффективностью или усилением, 2) временем фотоответа и 3) чувствительностью (обнаружительной способностью). Рассмотрим сначала процессы, протекающие в фоторезисторе при воздействии излучения (рис. 1). Пусть в какой-то начальный момент времени $t = 0$ число носителей, возникших в единице объема за счет генерации потоком фотонов, равно n_0 . В последующие моменты времени t число носителей в том же объеме уменьшается за счет рекомбинации по закону $n = n_0 \exp(-t/\tau)$, где τ — время жизни носителей. Другими словами, скорость рекомбинации равна $1/\tau$. Если поток фотонов постоянен и распределен равномерно по поверхности фоторезистора с площадью $A = WL$, то общее число фотонов, падающих на поверхность в единицу времени, равно $P_{\text{opt}}/h\nu$, где P_{opt} — мощность падающего излучения и $h\nu$ — энергия фотона.

В стационарном состоянии скорость генерации носителей должна быть равна скорости рекомбинации. Если толщина прибора D значительно больше, чем глубина проникновения света $1/\alpha$, то общая стационарная скорость генерации носителей в единице объема равна

$$G = \frac{n}{\tau} = \frac{\eta (P_{\text{opt}}/h\nu)}{WLD}, \quad (2)$$

где η — квантовая эффективность (т. е. число фотогенерированных носителей, отнесенное к числу падающих фотонов) и n — число носителей в единице объема (плотность носителей). Фототок, протекающий между контактами, равен

$$I_p = (q\mathcal{E})WD = (q\mu_n n \mathcal{E})WD = (qn v_d)WD, \quad (3)$$

где \mathcal{E} — электрическое поле внутри фоторезистора и v_d — дрейфовая скорость. Подставляя n из уравнения (2) в уравнение (3), получим

$$I_p = q \left(\eta \frac{P_{\text{opt}}}{h\nu} \right) \left(\frac{\mu_n \tau \mathcal{E}}{L} \right). \quad (3a)$$

Определяя исходный фототок как

$$I_{ph} \equiv q \left(\eta \frac{P_{\text{opt}}}{h\nu} \right),$$

Таблица 1. Типичные значения коэффициента усиления и времени фотоответа

Фотодетектор	Усиление	Время фотоответа, с	Рабочая температура, К
Фоторезистор	1—10 ⁶	10 ⁻³ —10 ⁻⁸	4,2—300
p—n-переход	1	10 ⁻¹¹	300
p—i—n-переход	1	10 ⁻⁸ —10 ⁻¹⁰	300
Диод со структурой металл—полупроводник	1	10 ⁻¹¹	300
Лавинный фотодиод	10 ² —10 ⁴	10 ⁻¹⁰	300
Биполярный фототранзистор	10 ²	10 ⁻⁸	300
Полевой транзистор	10 ²	10 ⁻⁷	300

из уравнения (3а) получим коэффициент усиления фототока

$$\frac{I_p}{I_{ph}} = \frac{\mu_n \tau \mathcal{E}}{L} = \frac{\tau}{t_r}, \quad (4)$$

где $t_r = L/v_d$ — время пролета носителей. Коэффициент усиления, зависящий от отношения времени жизни и времени пролета, является исключительно важным параметром фоторезисторов. Для образцов с большим временем жизни и малым расстоянием между контактами коэффициент усиления может быть существенно больше единицы. Время фотоответа фоторезистора определяется временем пролета t_r . Поскольку для фоторезисторов характерны большие расстояния между контактами и слабые электрические поля, то их время фотоответа обычно больше аналогичного параметра для фотодиодов. Некоторые типичные значения коэффициента усиления и времени фотоответа приведены в табл. 1 [2, 8].

Рассмотрим теперь оптический сигнал с интенсивностью, модулированной по закону

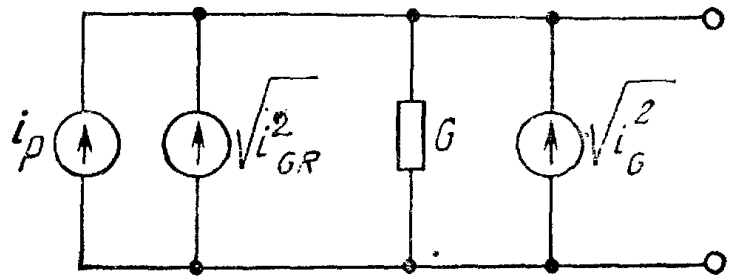
$$P(\omega) = P_{opt} (1 + m e^{j\omega t}), \quad (5)$$

где P_{opt} — средняя мощность оптического сигнала, m — коэффициент модуляции и ω — частота модуляции. Средний ток I_p , обусловленный оптическим сигналом, определяется выражением (3а). Среднеквадратичная мощность модулированного оптического сигнала равна $mP_{opt}/\sqrt{2}$, а среднеквадратичный токовый сигнал может быть представлен в виде [9]

$$I_p \approx \frac{q\eta m P_{opt}}{\sqrt{2} h\nu} \left(\frac{\tau}{t_r} \right) \frac{1}{(1 + \omega^2 \tau^2)^{1/2}}. \quad (6)$$

На рис. 3 показана эквивалентная схема фоторезистора. Проводимость G включает проводимость, обусловленную темновым током,

Рис. 3. Эквивалентная схема фоторезистора [9].



и проводимость, вызываемую фототоком и фоновым током. Тепловой шум, обусловленный проводимостью G , равен

$$i_G^2 = 4kTGB, \quad (7)$$

где k — константа Больцмана, T — абсолютная температура и B — ширина полосы. Генерационно-рекомбинационный шум (дробовой шум) определяется выражением [10]

$$\overline{i_{GR}^2} = \frac{\tau}{t_r} \frac{4qI_0B}{1 + \omega^2\tau^2}, \quad (8)$$

где I_0 — стационарный выходной фототок, равный $I_p (= \tau I_{ph}/t_r)$. Используя уравнения (6)–(8), можно получить выражение для отношения сигнал/шум:

$$(S/N)_{power} = \frac{i_p^2}{i_{GR}^2 + i_G^2} = \frac{\eta m^2 (P_{opt}/hv)}{8B} \left[1 + \frac{kT}{q} \frac{t_r}{\tau} (1 + \omega^2\tau^2) \frac{G}{I_0} \right]^{-1}. \quad (9)$$

Относительным критерием качества фотодетекторов [11] является мощность, эквивалентная шуму (NEP)¹, которая определяется как среднеквадратичная мощность падающего излучения, необходимая для получения отношения сигнал/шум, равного 1, в полосе частот 1 Гц. Выражение для NEP (т. е. $mP_{opt}/V\bar{2}$) можно получить из уравнения (9), подставив в него $S/N = 1$ и $B = 1$. Для инфракрасных детекторов наиболее употребительным критерием качества является удельная обнаружительная способность D^* [12]:

$$D^* = \frac{A^{1/2}B^{1/2}}{NEP} \text{ [см (Гц)}^{1/2}\text{/Вт)].} \quad (10)$$

Для того чтобы исключить неопределенность в D^* , необходимо указывать, является ли источником излучения абсолютно черное тело или монохроматический источник и при какой частоте осуществляется модуляция. Рекомендуется выразить D^* как $D^*(\lambda, f, 1)$ или $D^*(T, f, 1)$, где λ — длина волны (мкм), f — частота модуляции (Гц), T — температура (К). При этом ширина полосы всегда равна 1 Гц.

Для фоторезистора, чувствительность которого ограничена фоновым излучением, идеальная обнаружительная способность D^*

¹ Сокращенное обозначение словосочетания noise equivalent power (NEP). — Прим. перев.

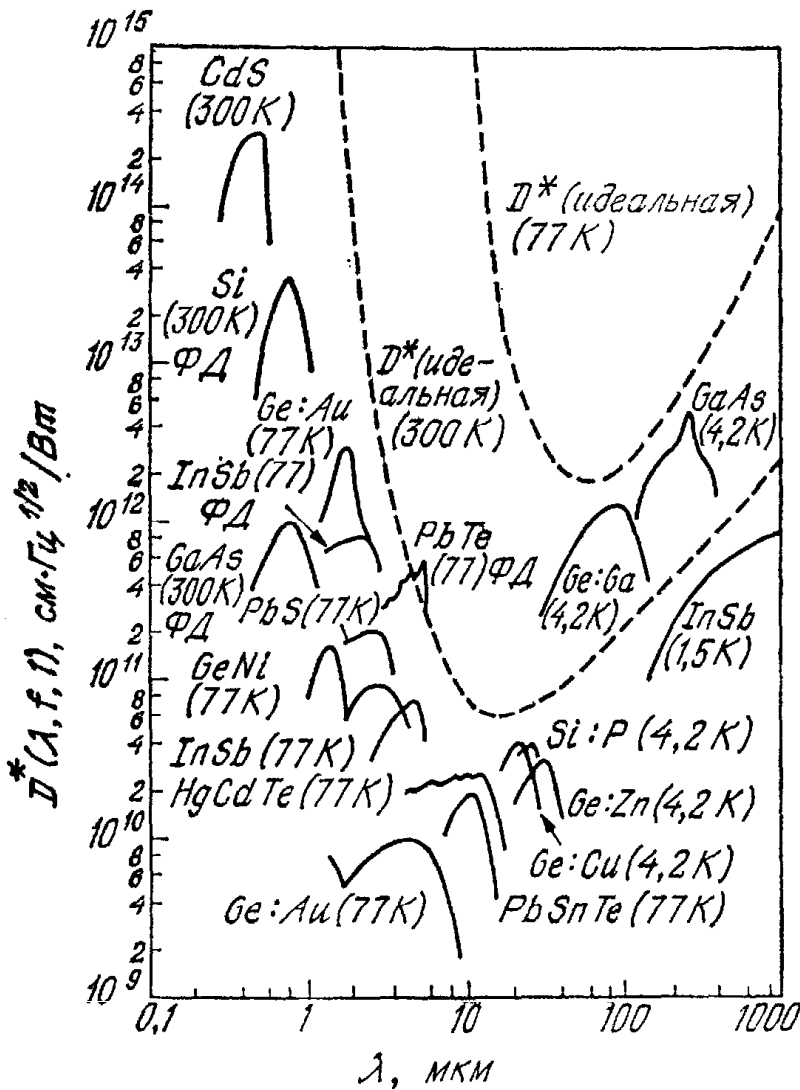


Рис. 4. Зависимость обнаружительной способности D^* от длины волны излучения для различных фоторезисторов и фотодиодов (фотодиоды обозначены ФД). Штриховыми кривыми представлена теоретическая идеальная D^* при 77 и 300 К и угле 2π ср [2, 3].

(рис. 4) при квантовой эффективности, равной 1, определяется выражением [12]

$$D^*(\lambda, f, 1) = \frac{c \exp(\zeta)}{2 \sqrt{\pi h k T} v^2 (1 + 2\zeta + 2\zeta^2)^{1/2}}, \quad (11)$$

где c — скорость света и $\zeta \equiv h\nu/kT$. На рис. 4 представлены также некоторые типичные значения D^* для таких фоторезисторов, как CdS, PbS, InSb, германий, легированный никелем (Ge—Ni), кремний, легированный фосфором Si—P [4, 7] (энергетические уровни примесей см. в гл. 1). Отметим, что для детектирования излучения в среднем, дальнем и сверхдальнем ИК-диапазонах фоторезисторы охлаждаются до низких температур (77 и 4,2 К). При таких температурах уменьшаются тепловые эффекты, вызывающие термическую ионизацию и опустошение энергетических уровней, и увеличиваются усиление и эффективность детектирования. На длинах волн вблизи 0,5 мкм наиболее высокой чувствительностью обладает фоторезистор на CdS, в то время как при $\lambda = 10$ мкм предпочтительны фоторезисторы на HgCdTe [13]. Для детектирования излучения в диапазоне длин волн 100—400 мкм лучше всего выбрать фоторезистор на GaAs [14], так как он

имеет большие значения D^* . Данный фоторезистор обладает широким динамическим диапазоном и вследствие этого может обеспечить хорошие характеристики детектирования сигналов высокой интенсивности. Однако для детектирования слабых сигналов в микроволновом диапазоне частот фотодиоды обеспечивают большее быстродействие и сравнительно более высокое отношение сигнал/шум. Таким образом, применение фоторезисторов в качестве высокочастотных оптических демодуляторов, например смесителей мощных оптических сигналов, ограничено. Вместе с тем они широко используются для детектирования в инфракрасной области спектра, особенно на длинах волн выше нескольких микрометров, где, несмотря на интенсивную работу, пока еще отсутствуют удовлетворительные альтернативные варианты методов детектирования.

13.3. ФОТОДИОД

13.3.1. Общий анализ

В фотодиоде имеется обедненная область полупроводника с сильным электрическим полем, в котором происходит разделение электронно-дырочных пар, возбужденных под действием света. Для работы фотодиода на высоких частотах необходимо обеспечить малые времена пролета, поэтому обедненная область должна быть тонкой. С другой стороны, для увеличения квантовой эффективности (число фотогенерированных электронно-дырочных пар, отнесенное к числу падающих фотонов) обедненный слой должен быть достаточно толстым, чтобы обеспечить поглощение большей части падающего излучения. Таким образом, существует взаимосвязь между быстродействием и квантовой эффективностью.

Фотодиод может работать в фотовольтаическом режиме, характерном для солнечных батарей (гл. 14) — несмещенный диод подсоединяется к нагрузочному сопротивлению. Однако конструктивно фотодиод существенно отличается от солнечной батареи. В работе фотодиода определяющую роль играет лишь излучение, сосредоточенное в узком интервале длин волн в центре оптического диапазона, а солнечные батареи должны обладать высокой чувствительностью в широком диапазоне длин волн солнечного излучения. Фотодиоды имеют малые размеры, необходимые для минимизации емкости перехода, тогда как солнечные батареи относятся к приборам с большой площадью. Одной из наиболее важных характеристик фотодиодов является квантовая эффективность, тогда как качество солнечных батарей в основном оценивается по эффективности преобразования мощности (мощность, выделяемая на нагрузке, отнесенная к мощности падающего солнечного излучения).

При работе в видимом и ближнем инфракрасном диапазонах диоды обычно смещаются в обратном направлении с помощью сравнительно большого напряжения, чтобы уменьшить время пролета носителей и снизить емкость перехода. Напряжение смещения, однако, не настолько велико, чтобы вызвать лавинный пробой. Этим ограничением на величину смещения отличаются рассматриваемые фотодиоды от лавинных фотодиодов, в которых внутреннее усиление достигается как раз за счет ударной ионизации в условиях лавинного пробоя. В семейство фотодиодов входят диоды с $p-n$ -переходом, $p-i-n$ -диоды, диоды со структурой металл—полупроводник (с барьером Шоттки) и диоды с гетеропереходом.

В этом разделе мы кратко рассмотрим общие характеристики фотодиодов: квантовую эффективность, скорость фотоответа и приборные шумы.

Квантовая эффективность, как уже говорилось выше, представляет собой число фотогенерированных электронно-дырочных пар, отнесенное к числу падающих фотонов:

$$\eta = (I_p/q)/(P_{\text{opt}}/h\nu), \quad (12)$$

где I_p — фототок, обусловленный поглощением падающего оптического излучения с мощностью P_{opt} и длиной волны λ (соответствующей энергии фотона $h\nu$).

Сравнительным критерием качества является чувствительность, которая определяется как отношение фототока к оптической мощности:

$$\mathcal{R} = \frac{I_p}{P_{\text{opt}}} = \frac{\eta q}{h\nu} = \frac{\eta \lambda}{1,24}. \quad (13)$$

Таким образом, для данной квантовой эффективности чувствительность линейно растет с ростом λ . Для идеального фотодиода ($\eta = 1$) $\mathcal{R} = (\lambda/1,24)$ (А/Вт), где длина λ выражена в микрометрах.

Одним из основных факторов, определяющих квантовую эффективность, является коэффициент поглощения. На рис. 5 и 6 приведены зависимости коэффициентов поглощения α от длины волны, измеренные для различных фоточувствительных материалов [2]. Для Ge, Si и полупроводниковых соединений типа $A^{III}B^V$ при увеличении температуры кривые сдвигаются в область больших длин волн. Для некоторых соединений типа $A^{IV}B^{VI}$ (например, PbSe) имеет место обратная ситуация, что обусловлено расширением запрещенной зоны при увеличении температуры. На рисунках указаны также длины волн, соответствующие излучению некоторых важных лазеров.

Вследствие того что α сильно зависит от длины волны, для данного полупроводника область длин волн, в которой можно получить заметный фототок, ограничена. Длинноволновая гра-

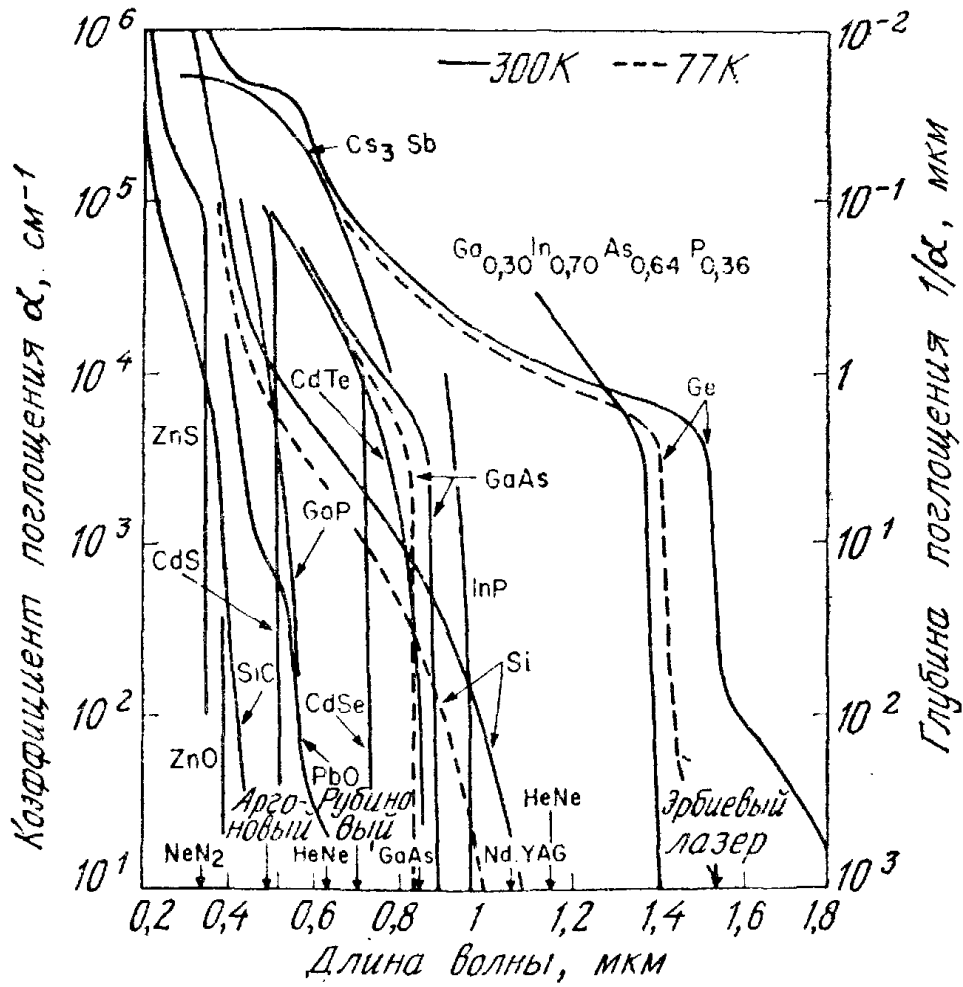


Рис. 5. Зависимость коэффициента поглощения от длины волны излучения для различных фоточувствительных материалов [2].

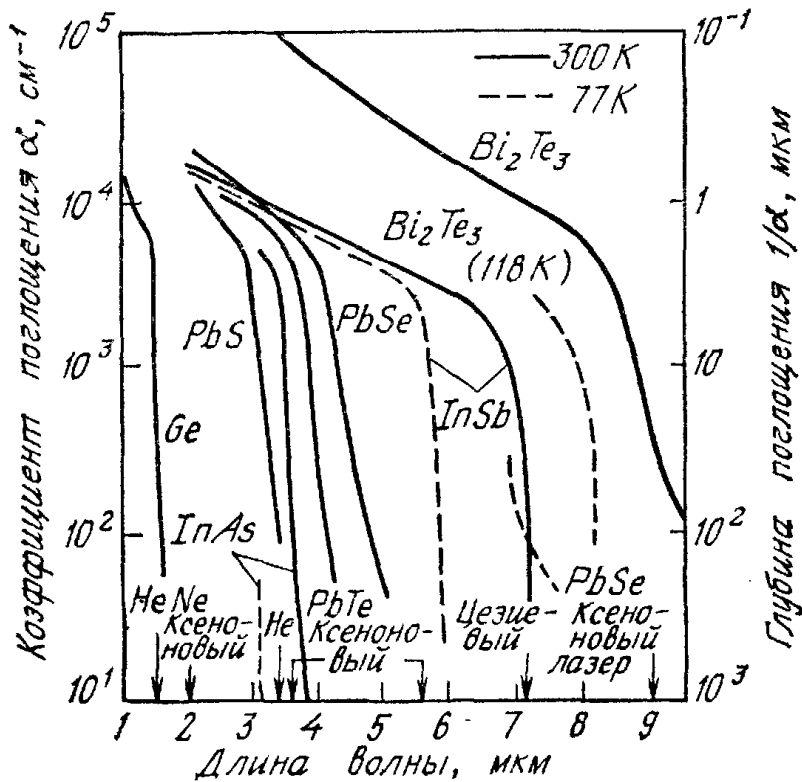


Рис. 6. Зависимость коэффициента поглощения от длины волны излучения для материалов, чувствительных в ИК-области спектра [2].

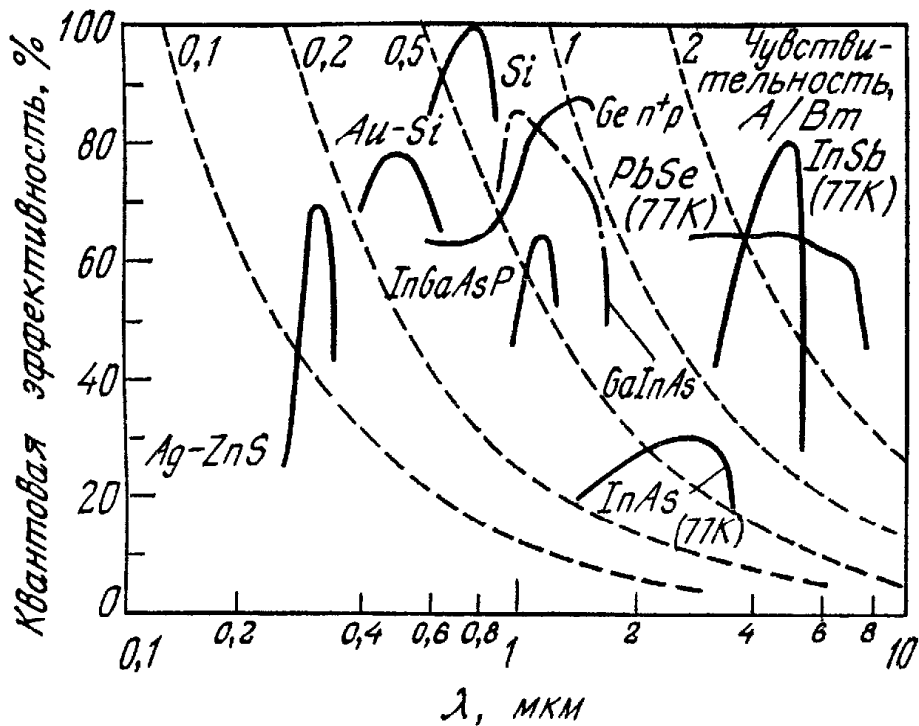


Рис. 7. Зависимости квантовой эффективности (сплошная кривая) и чувствительности (штриховая кривая) различных фотодетекторов от длины волны излучения.

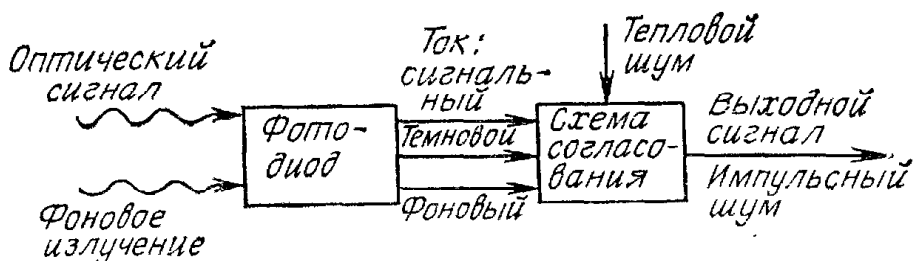
ница λ_c определяется шириной запрещенной зоны полупроводника в соответствии с уравнением (1) и равна $\sim 1,7$ мкм для Ge и $\sim 1,1$ мкм для Si. Для длин волн, больших λ_c , значение α слишком мало, чтобы обеспечить заметное поглощение. Коротковолновая граница фотоответа обусловлена тем, что в области коротких волн коэффициент α очень велик ($\geq 10^5$ см $^{-1}$) и излучение поглощается в непосредственной близости поверхности, где времена рекомбинации малы. Вследствие этого фотоносители рекомбинируют раньше, чем уходят из поверхностного слоя. На рис. 7 приведены типичные зависимости квантовой эффективности от длины волны для некоторых высокочастотных фотодиодов и кривые квантовой эффективности, соответствующие заданным значениям чувствительности. В ультрафиолетовой и видимой областях спектра высокой квантовой эффективностью обладают фотодиоды со структурой металл—полупроводник. В ближнем ИК-диапазоне (на длинах волн $\sim 0,8$ — $0,9$ мкм) в кремниевых фотодиодах с просветляющим покрытием может быть достигнута квантовая эффективность 100%. В области длин волн 1,0—1,6 мкм высокой квантовой эффективностью характеризуются фотодиоды на Ge, а также на основе тройных (например, GaInAs) и четверных (например, GaInAsP) соединений элементов III и V групп. Для обеспечения высокоэффективной работы в длинноволновой области фотодиоды охлаждаются до низких температур (например, до 77 К).

Скорость фотоответа ограничивается тремя факторами: диффузией носителей, временем дрейфа через обедненную область и емкостью обедненной области. Носители, генерированные за пре-

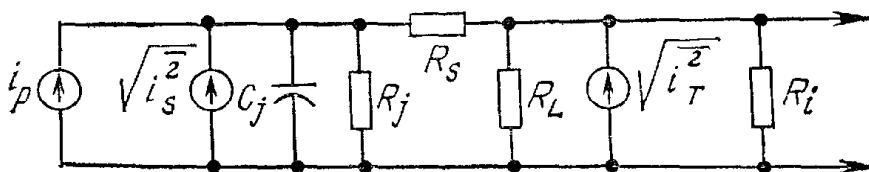
делами обедненной области, должны диффундировать к переходу, что существенно увеличивает время задержки. Чтобы уменьшить влияние диффузии, необходимо формировать переход вблизи поверхности. Основная часть излучения поглощается в том случае, когда обедненная область достаточно широка (порядка $1/\alpha$); при достаточно высоких обратных смещениях дрейф носителей осуществляется со скоростью насыщения. Обедненный слой, однако, не должен быть слишком широким, иначе время переноса будет ограничивать скорость фотоответа. Кроме того, он не должен быть и слишком узким, иначе за счет большой емкости C будет велика постоянная времени RC (R — сопротивление нагрузки). Оптимальной шириной обедненного слоя считается такая ширина, при которой время переноса равно примерно половине периода модуляции излучения. Например, для частоты модуляции, равной 10 ГГц, оптимальная ширина обедненного слоя в кремнии (скорость насыщения $\sim 10^7$ см/с) составляет ~ 5 мкм.

Для изучения шумовых свойств фотодиода рассмотрим обобщенный процесс детектирования излучения, схематически представленный на рис. 8, а [5]. При поглощении оптического сигнала и фонового излучения в фотодиоде генерируются электронно-дырочные пары. Электроны и дырки разделяются электрическим полем и дрейфуют в противоположные стороны от p - n -перехода. При этом через внешнее нагрузочное сопротивление протекает ток смещения. Для определения генерационного тока, возникающего в результате этого фотоэлектронного процесса, рассмотрим модулированный по интенсивности оптический сигнал, заданный уравнением (5). Средний фототок, обусловленный оптическим сигналом, равен

$$I_P = q\eta P_{\text{opt}}/h\nu. \quad (14)$$



а



б

Рис. 8. Процесс детектирования излучения (а) и эквивалентная схема фотодиода (б) [5].

Среднеквадратичная мощность модулированного оптического сигнала равна $mP_{\text{opt}}/\sqrt{2}$, а среднеквадратичный ток сигнала можно записать в виде

$$i_p = q\eta mP_{\text{opt}}/\sqrt{2} h\nu. \quad (15)$$

Ток, обусловленный фоновым излучением, равен I_B , а темновой ток термической генерации электронно-дырочных пар в обедненной области I_D . Вследствие случайного характера процесса генерации этих токов дробовой шум определяется выражением

$$\langle i_s^2 \rangle = 2q(I_P + I_B + I_D)B, \quad (16)$$

где B — полоса частот.

Эквивалентная схема фотодиода [5] показана на рис. 8, б, где i_p , $\sqrt{i_s^2}$, C_j , R_j и R_s определяются параметрами фотодиода. Компонента C_j представляет собой емкость перехода, R_j — сопротивление перехода, а R_s — последовательное сопротивление. Изменяемое сопротивление R_L соответствует внешнему нагрузочному сопротивлению, а R_i — входному сопротивлению усилителя [55]. Все сопротивления вносят в систему дополнительные тепловые шумы. Последовательное сопротивление R_s обычно намного меньше, чем другие сопротивления, поэтому его можно не учитывать. Тепловой шум равен

$$\langle i_T^2 \rangle = 4kT(1/R_{eq})B, \quad (17)$$

где $1/R_{eq} = (1/R_j) + (1/R_L) + (1/R_i)$.

При 100 %-ной модуляции оптического сигнала со средней мощностью P_{opt} выражение для отношения сигнал/шум можно записать в виде

$$(S/N)_{\text{power}} = \frac{i_p^2 R_{eq}}{(\langle i_s^2 \rangle + \langle i_T^2 \rangle) R_{eq}} = \frac{1/2 (q\eta P_{\text{opt}}/h\nu)^2}{2q(I_P + I_B + I_D)B + 4kTB/R_{eq}}. \quad (18)$$

Из этого уравнения следует, что минимальная оптическая мощность, необходимая для получения заданного отношения сигнал/шум, равна

$$(P_{\text{opt}})_{\text{мин}} = \frac{2h\nu B}{\eta} \left(\frac{S}{N} \right) \left\{ 1 + \left[1 + \frac{I_{eq}}{qB(S/N)} \right]^{1/2} \right\}, \quad (19)$$

где $I_{eq} = I_B + I_D + 2kT/qR_{eq}$. В пределе, когда $I_{eq}/qB(S/N) \ll \ll 1$, минимальная оптическая мощность определяется квантовыми шумами, связанными с самим оптическим сигналом. В другом предельном случае, когда $I_{eq}/qB(S/N) \gg 1$, преобладающими становятся шум фонового излучения и тепловой шум эквивалентного

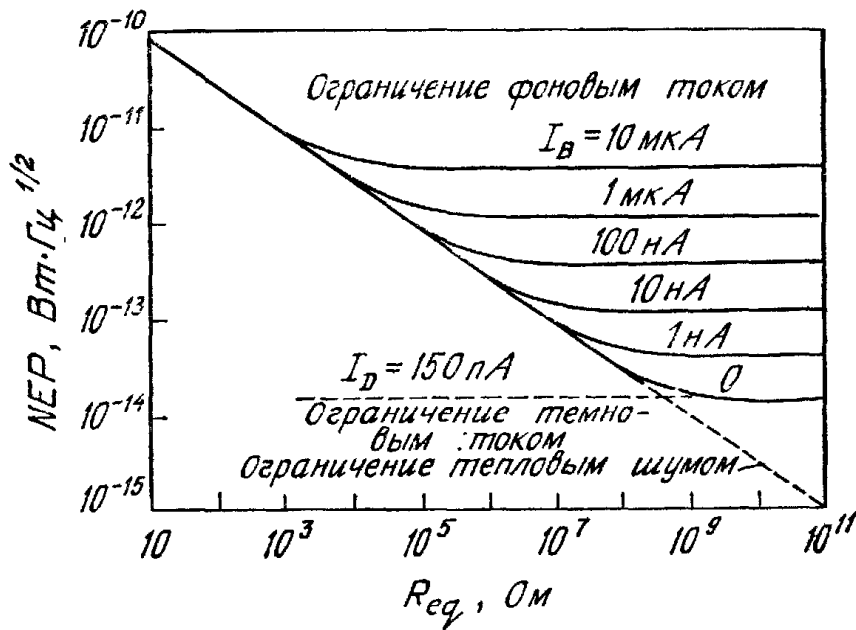


Рис. 9. Зависимость NEP $p-i-n$ -диода от величины нагрузочного сопротивления при темновом токе 150 пА и различных значениях фонового тока.

сопротивления. В этих условиях мощность, эквивалентная шуму (NEP) и равная среднеквадратичной оптической мощности при $S/N = 1$ и $B = 1$ Гц, описывается выражением

$$NEP = \sqrt{2} (h\nu/\eta) (I_{eq}/q)^{1/2} [\text{Вт/см}^2 \text{ Гц}^{1/2}]. \quad (20)$$

Для того чтобы улучшить чувствительность фотодиода, необходимо η и R_{eq} увеличивать, а I_B и I_D уменьшать. Зависимость NEP от величины нагрузочного сопротивления для типичного кремниевого фотодиода, вычисленная при $\eta = 75\%$, $\lambda = 0,77$ мкм и $I_D = 1,5 \cdot 10^{-10}$ А [5], приведена на рис. 9. Из графика следует, что NEP при наличии дробового шума темнового тока или фонового излучения может быть достигнута лишь при достаточно высоких значениях R_{eq} .

13.3.2. $p-i-n$ -Фотодиод

$p-i-n$ -Фотодиод является наиболее распространенным фотодетектором, поскольку толщина обедненной области (слоя собственной проводимости) может быть создана такой, чтобы обеспечить оптимальную квантовую эффективность и быстродействие. На рис. 10 схематически представлены структура $p-i-n$ -диода, его энергетическая диаграмма в условиях обратного смещения и распределение интенсивности излучения [2]. С помощью этой схемы мы проведем анализ некоторых особенностей работы $p-i-n$ -фотодиода. Этот анализ будет справедливым также и для фотодиода с $p-n$ -переходом. Поглощение излучения в полупровод-

нике приводит к генерации электронно-дырочных пар. Пары, создаваемые в обедненной области или на расстоянии диффузионной длины от нее, разделяются электрическим полем, и во внешней цепи во время дрейфа носителей через обедненную область появляется ток.

В стационарных условиях плотность полного тока, протекающего через обратносмещенный обедненный слой, равна [15]

$$J_{\text{tot}} = J_{\text{dr}} + J_{\text{diff}}, \quad (21)$$

где J_{dr} — плотность дрейфового тока, обусловленного генерацией носителей в обедненной области, и J_{diff} — плотность диффузионного тока, обусловленного генерацией носителей в объеме полупроводника за пределами обедненного слоя и их диффузией к обратносмещенному переходу. Выведем выражение для полного тока при условии, что током термической генерации можно пренебречь и что приповерхностный слой p -типа существенно меньше $1/\alpha$. В соответствии с рис. 10, a выражение для скорости генерации электронно-дырочных пар имеет вид

$$G(x) = \Phi_0 \alpha e^{-\alpha x}. \quad (22)$$

Здесь Φ_0 — поток падающих фотонов, приходящийся на единицу площади и равный $P_{\text{opt}}(1 - R)/A h\nu$, где R — коэффициент

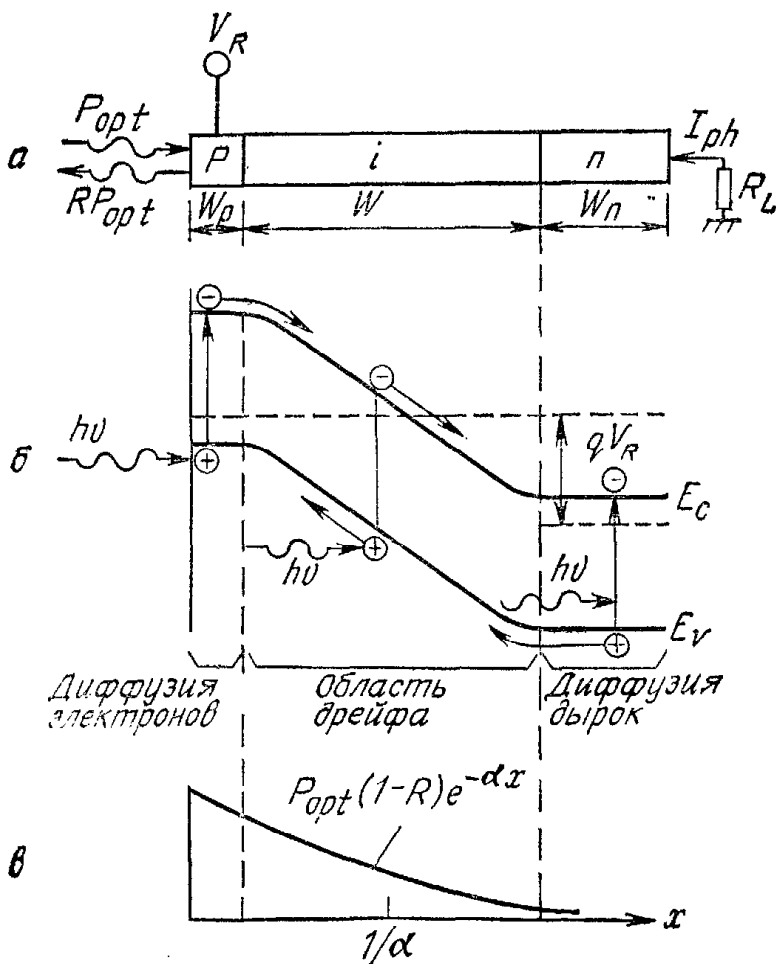


Рис. 10. Принцип работы фотодиода [2].

a — поперечный разрез $p-i-n$ -диода; b — зонная диаграмма в условиях обратного смещения; c — распределение интенсивности излучения.

отражения и A — площадь прибора. Таким образом, дрейфовый ток

$$J_{dr} = -q \int_0^W G(x) dx = q\Phi_0 (1 - e^{-\alpha W}), \quad (23)$$

где W — толщина обедненного слоя. При $x > W$ плотность неосновных носителей (дырок) в объеме полупроводника определяется из одномерного диффузионного уравнения

$$D_p \frac{\partial^2 p_n}{\partial x^2} - \frac{p_n - p_{n0}}{\tau_p} + G(x) = 0, \quad (24)$$

где D_p — коэффициент диффузии дырок, τ_p — время жизни неравновесных носителей и p_{n0} — равновесная плотность дырок. Решение уравнения (24) при граничных условиях $p_n = p_{n0}$ при $x = \infty$ и $p_n = 0$ при $x = W$ имеет вид

$$p_n = p_{n0} - (p_{n0} + C_1 e^{-\alpha W}) e^{(W-x)/L_p} + C_1 e^{-\alpha x}, \quad (25)$$

где $L_p = \sqrt{D_p \tau_p}$ и

$$C_1 \equiv \left(\frac{\Phi_0}{D_p} \right) \frac{\alpha L_p^2}{1 - \alpha^2 L_p^2}. \quad (26)$$

Плотность диффузионного тока, определяемая как $J_{diff} = -qD_p (\partial p_n / \partial x)_{x=W}$, равна

$$J_{diff} = q\Phi_0 \frac{\alpha L_p}{1 + \alpha L_p} e^{-\alpha W} + qp_{n0} \frac{D_p}{L_p}. \quad (27)$$

Плотность полного тока описывается выражением

$$J_{tot} = q\Phi_0 \left(1 - \frac{e^{-\alpha W}}{1 + \alpha L_p} \right) + qp_{n0} \frac{D_p}{L_p}. \quad (28)$$

При нормальных условиях работы член, содержащий p_{n0} , значительно меньше первого члена и полный фототок пропорционален потоку фотонов. Из уравнений (12) и (28) можно вывести выражение для квантовой эффективности

$$\eta = \frac{J_{tot}/q}{P_{opt}/Ah\nu} = (1 - R) \left(1 - \frac{e^{-\alpha W}}{1 + \alpha L_p} \right). \quad (29)$$

Для получения высокой квантовой эффективности желательно, чтобы коэффициент отражения был мал и выполнялось условие $\alpha W \gg 1$. Однако при $W \gg 1/\alpha$ может оказаться значительным время пролета. Рассмотрим теперь влияние времени пролета.

Так как для прохождения носителей через обедненный слой требуется конечное время, то при высокочастотной модуляции интенсивности падающего излучения появляется разность фаз между потоком фотонов и фототоком. Для получения количественного выражения, описывающего этот эффект, обратимся к рис.

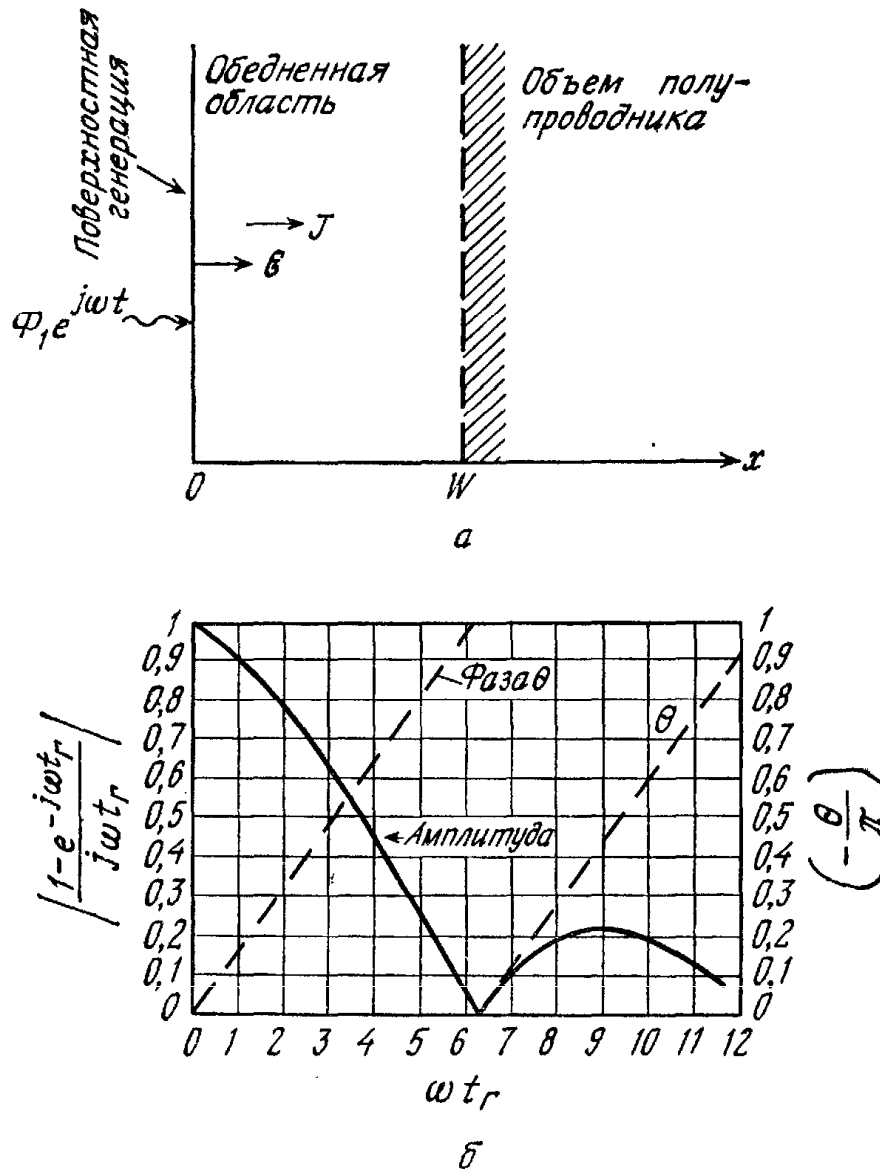


Рис. 11. Представление диода (а), принятое при рассмотрении эффектов, связанных с влиянием времени пролета, и фотоответ (нормированный ток или напряжение) (б) как функция от нормализованной частоты модулятора падающего потока фотонов (здесь $\theta \equiv \omega t_r/2$) [15].

11, а, на котором схематически изображен наиболее простой случай. Предполагается, что приложенное напряжение достаточно велико для того, чтобы область собственной проводимости была приведена в состояние полного обеднения и скорость дрейфа носителей достигла значения скорости насыщения. При плотности потока фотонов, равной $\Phi_1 e^{j\omega t}$ (фотон/с см²), плотность тока проводимости J_{cond} в точке x оказывается равной

$$J_{\text{cond}}(x) = q\Phi_1 e^{j\omega(t-x/v_s)}. \quad (30)$$

Поскольку $\nabla \cdot J_{\text{tot}} = 0$, то можно написать

$$J_{\text{tot}} = \frac{1}{W} \int_0^W \left(J_{\text{cond}} + \epsilon_s \frac{\partial \mathcal{E}}{\partial t} \right) dx, \quad (31)$$

где второй член в скобках представляет собой плотность тока смещения, а ϵ_s и \mathcal{E} — диэлектрическая постоянная и электрическое поле соответственно. Подставляя уравнение (30) в уравнение (31), получим

$$J_{\text{tot}} = \left(\frac{j\omega\epsilon_s V}{W} + q\Phi_1 \frac{1 - e^{-j\omega t_r}}{j\omega t_r} \right) e^{j\omega t}, \quad (32)$$

где V — сумма приложенного и встроенного напряжений и $t_r \equiv \equiv W/v_s$ — время пролета носителей через обедненную область.

Из уравнения (32) следует, что плотность тока короткого замыкания ($V = 0$) равна

$$J_{sc} = \frac{q\Phi_1 (1 - e^{j\omega t_r})}{j\omega t_r} e^{j\omega t}. \quad (33)$$

Эффекты, проявляющиеся на высоких частотах и связанные с влиянием времени пролета, иллюстрируются на рис. 11, б, на котором приведены зависимости амплитуды и фазового угла нормализованного тока от нормализованной частоты модуляции. Отметим, что амплитуда переменной составляющей фототока резко уменьшается, когда ωt_r становится больше 1. При $\omega t_r = 2,4$ уменьшение амплитуды в $\sqrt{2}$ раз сопровождается изменением фазового угла на величину $2\pi/5$. Быстродействие фотодетектора, таким образом, ограничивается временем пролета носителей через обедненный слой. Разумный компромисс между быстродействием и квантовой эффективностью достигается при значениях ширины области поглощения от $1/\alpha$ до $2/\alpha$.

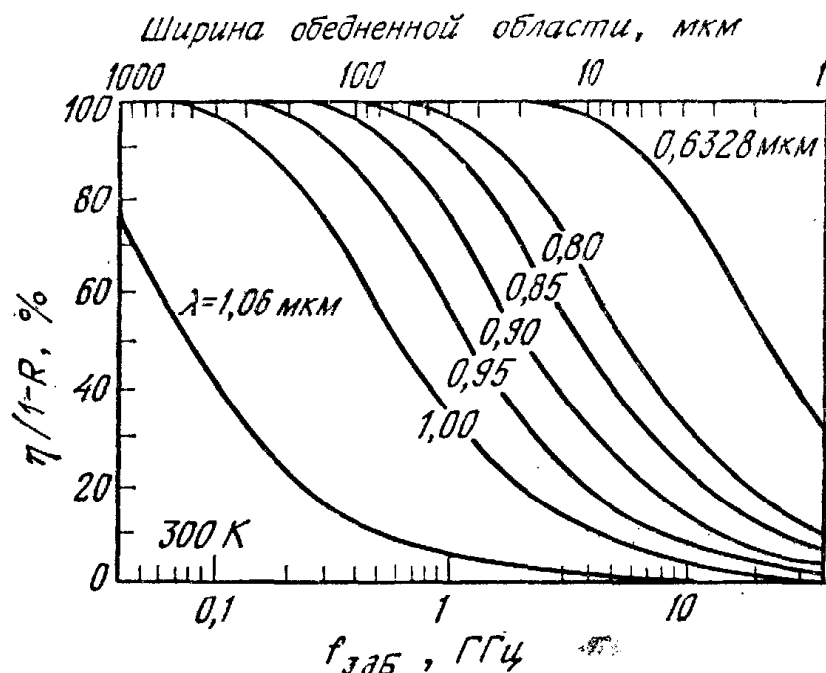


Рис. 12. Зависимость квантовой эффективности кремниевого $p-i-n$ -фотодиода от ширины обедненной области и граничной частоты (на уровне 3 дБ), определяемой временем пролета для различных длин волн. Скорость насыщения равна 10^7 см/с.

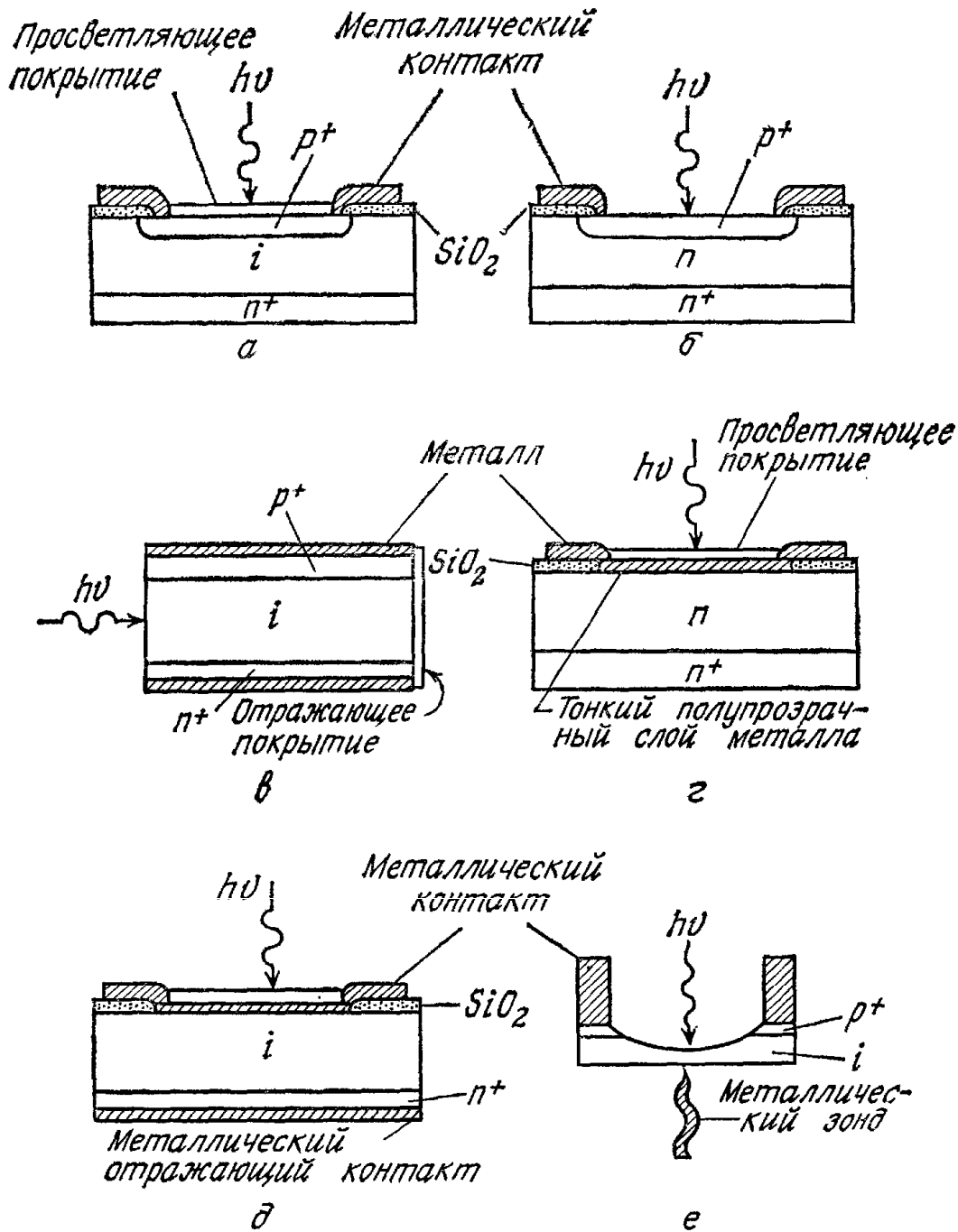


Рис. 13. Конструкции некоторых высокочастотных фотодиодов [2].

a — *p-i-n*-диод; *б* — диод с *p-n*-переходом; *в* — *p-i-n*-диод в освещении параллельно переходу; *г* — диод со структурой металл — полупроводник; *д* — диод со структурой металл — *i-n*; *е* — полупроводниковый точечно-контактный диод.

В *p-i-n*-диоде ширина *i*-области выбирается равной $1/\alpha$. Время пролета носителей равно времени их дрейфа через *i*-область. Из уравнения (33) следует, что ослабление фототока, соответствующее 3 дБ, происходит на частоте

$$f_{3 \text{ дБ}} = \frac{2,4}{2\pi l_s} \approx \frac{0,4v_s}{W} \approx 0,4\alpha v_s. \quad (34)$$

На рис. 12 приведена зависимость внутренней квантовой эффективности, т. е. $\eta/(1 - R)$, кремниевого *p-i-n*-фотодиода от $f_{3 \text{ дБ}}$ и ширины обедненной области, вычисленная из уравнения (34)

с учетом данных, представленных на рис. 5. Кривые иллюстрируют связь между быстродействием (частота $f_{з дБ}$) и квантовой эффективностью при различных длинах волн.

Несколько конструкций высокочастотных фотодиодов показаны на рис. 13 [2]. $p-i-n$ -Фотодиод (рис. 13, а) снабжен просветляющим покрытием для увеличения квантовой эффективности. Толщина области собственной проводимости (на практике используется высокоомный слой n -типа, т. е. v -область, или высокоомный слой p -типа, т. е. π -область) выбирается оптимальной для заданной длины волны оптического сигнала и частоты модуляции. В $p-n$ -фотодиоде n -слой имеет более высокую концентрацию легирующей примеси и в результате этого обеднен не полностью (рис. 13, б). Вследствие большего вклада диффузионного тока $p-n$ -фотодиоды обычно обладают меньшим быстродействием, чем $p-i-n$ -фотодиоды. При длинах волн вблизи красной границы чувствительности требуемая толщина поглощения становится очень большой (при $\alpha = 10 \text{ см}^{-1}$ или $1/\alpha = 1000 \text{ мкм}$). В этом случае согласование между квантовой эффективностью и быстродействием может быть достигнуто, если свет входит в полупроводник через грань параллельно переходу (рис. 13, в). Можно также выбрать такие углы падения света, которые обеспечивают его многократное отражение внутри прибора и тем самым приводят к увеличению эффективности толщины поглощения, сохраняя при этом малую длину дрейфа носителей [16, 17]. Три других прибора, представляющие собой фотодиоды с барьером Шоттки, рассмотрены в следующем разделе.

13.3.3. Фотодиод с барьером Шоттки

Фотодиоды с барьером Шоттки могут быть использованы в качестве высокоэффективных фотодетекторов [18, 19]. Зонная энергетическая диаграмма и протекание тока в диоде с барьером Шоттки рассмотрены в гл. 5, а типичная структура изображена на рис. 13, г. С целью исключения больших потерь на отражение и поглощение при освещении диода через металлический контакт следует использовать очень тонкие слои металла ($\sim 100 \text{ \AA}$) и применять просветляющие покрытия. Диод может работать в различных режимах в зависимости от энергии фотонов и прикладываемого смещения:

1. При $E_g \gg h\nu \gg q\phi_{Вп}$ и $V < V_B$ (рис. 14, а) (где V_B — напряжение лавинного пробоя) электроны, возбужденные в металле под действием излучения и преодолевшие барьер, будут входить в полупроводник. Этот эффект часто используется для определения высоты барьера Шоттки и изучения процессов переноса горячих электронов в металлических пленках [20].

2. При $h\nu > E_g$ и $V < V_B$ (рис. 14, б) излучение вызывает генерацию электронно-дырочных пар, и основные характеристики

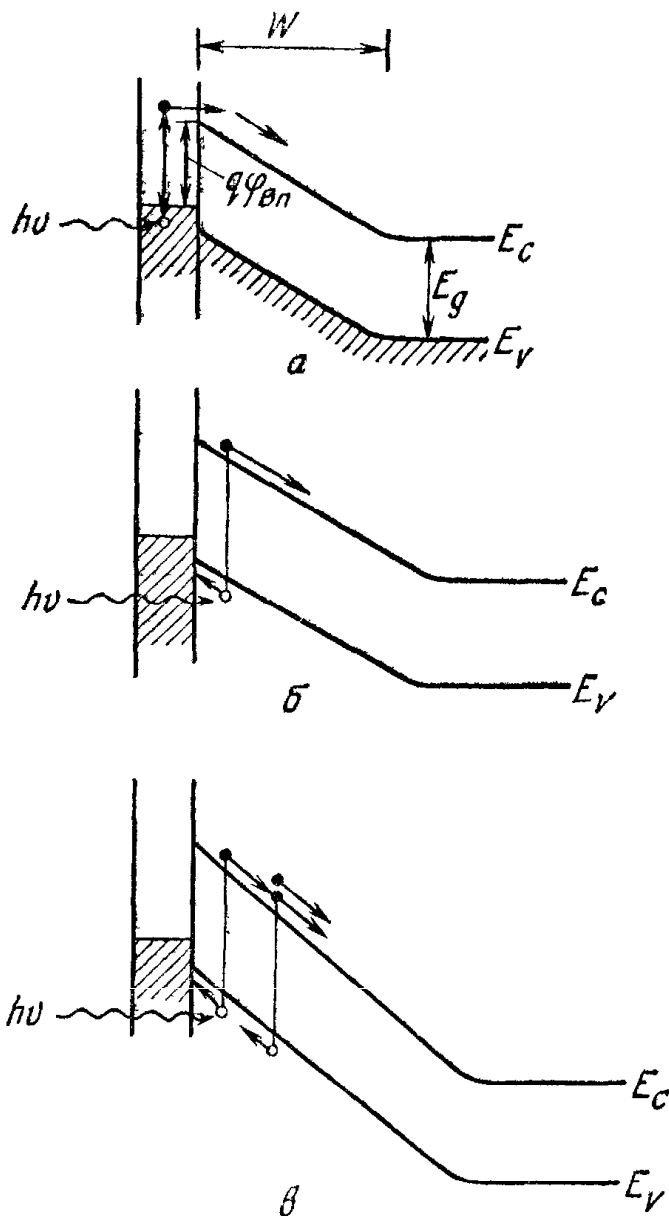


Рис. 14. Фотоэмиссия электронов из металла в полупроводник (а) ($E_g > h\nu > q\Phi_{Bn}$), собственное возбуждение (зона — зона) электронно-дырочной пары (б) ($h\nu > E_g$), а также генерация электронно-дырочной пары и лавинное умножение в условиях большого обратного смещения (в) ($h\nu > E_g$ и $V \simeq V_B$).

диода очень похожи на характеристики $p-i-n$ -фотодиода. Квантовая эффективность определяется выражением, идентичным уравнению (29).

3. При $h\nu > E_g$ и $V \simeq V_B$ (большое обратное смещение) (рис. 14, в) диод может работать как лавинный фотодиод (последний рассматривается в разд. 13.4).

Фотодиоды с барьером Шоттки особенно эффективны в видимой и ультрафиолетовой областях спектра. Для этих спектральных областей коэффициент поглощения α в наиболее распространенных полупроводниках очень велик ($\sim 10^5 \text{ см}^{-1}$ и более) и эффективная глубина поглощения мала ($1/\alpha \simeq 0,1 \text{ мкм}$ и менее). Поэтому оказывается возможным подобрать металл и просветляющее покрытие так, чтобы бóльшая часть падающего излучения поглощалась вблизи поверхности полупроводника.

Структура, показанная на рис. 15, а, имеет тщательно подобранное просветляющее покрытие из ZnS толщиной 500 \AA и коэффициентом преломления 2,30. Пленка из золота имеет комплексный

коэффициент преломления $\bar{n} = 0,28 + j0,301$ на длине волны $\lambda = 6328 \text{ \AA}$ (излучение He—Ne-лазера), в то время как на этой длине волны кремниевая подложка имеет коэффициент преломления, равный $3,75 + j0,018$. На рис. 15, б представлены зависимости пропускания, коэффициента отражения и потерь в пленке Au от ее толщины. Видно, что при толщине пленки около 100 \AA более 95 % падающего излучения проникает в кремниевую подложку. Экспериментальные результаты, полученные на диоде, подобном тому, который показан на рис. 13, г, показывают, что полная квантовая эффективность диода на длине волны $\lambda = 6328 \text{ \AA}$ равна 70 % и время нарастания фотоотклика достигает 0,1 нс ($1 \text{ нс} = 10^{-9} \text{ с}$) [17].

Квантовая эффективность и коэффициент пропускания Au—n—Si-фотодиода [21] (толщина пленки Au равна 107 \AA) для энергии фотона в диапазоне 1—6 эВ представлены на рис. 16. В видимой области спектра (вплоть до энергии $h\nu = 3,1 \text{ эВ}$) значения η и коэффициента пропускания практически одинаковы. В ультрафиолетовой области они различаются менее чем на 25 %. Найдено, что квантовая эффективность для всех значений энергии фотона не зависит от концентрации легирования при ее изменении от $2 \cdot 10^{13}$ до $2 \cdot 10^{17} \text{ см}^{-3}$. Это означает, что потери, обусловленные

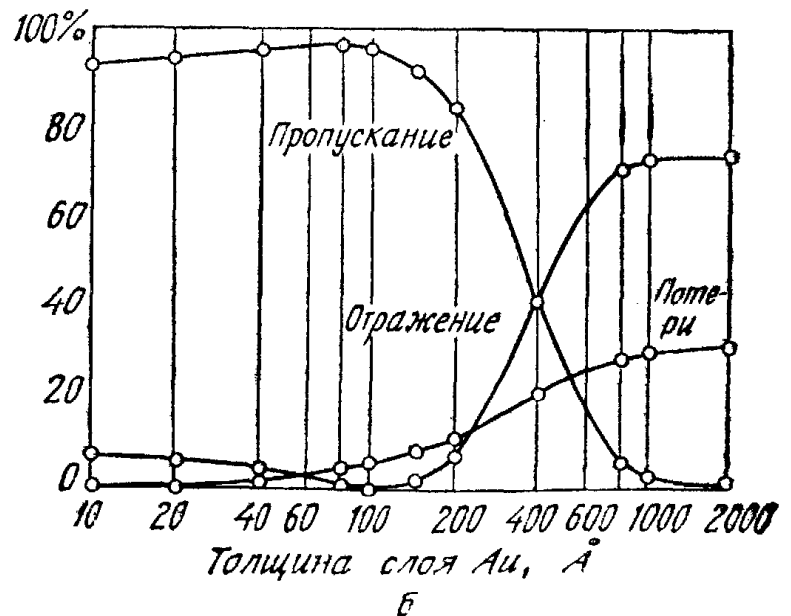
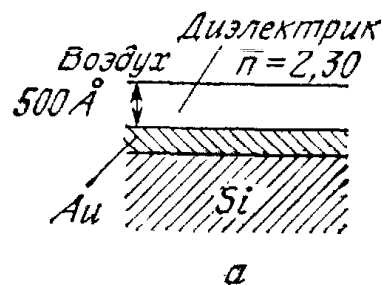


Рис. 15. Структура с просветляющим покрытием из пленки ZnS толщиной 500 \AA (а) и зависимость пропускания, отражения и потерь излучения ($\lambda = 0,6328 \text{ мкм}$) в пленке Au от ее толщины (б) [19].

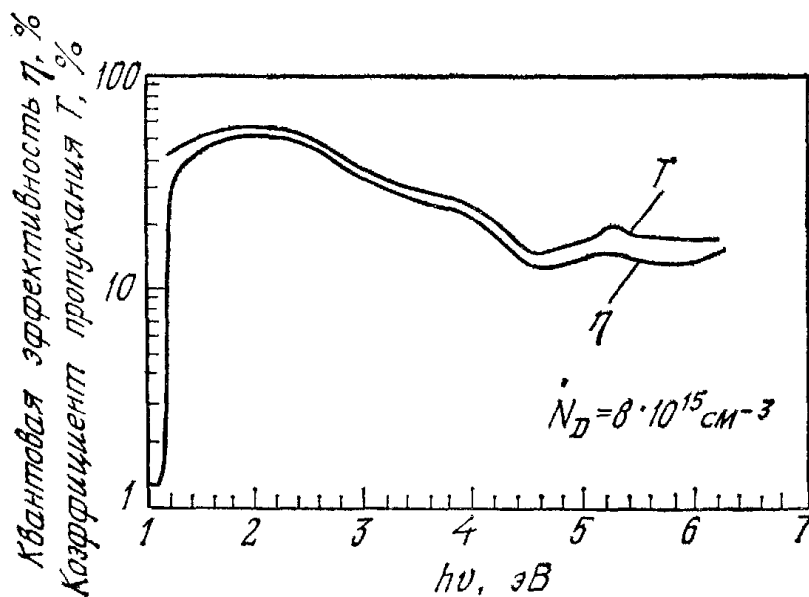


Рис. 16. Квантовая эффективность и коэффициент пропускания в Au — n — Si фотодиоде [21].

диффузией электронов в обедненной области в направлении, противоположном их дрейфу в электрическом поле, или рекомбинацией на ловушках на границе металл—полупроводник, пренебрежимо малы [22].

Используя слаболегированный *i*-слой, можно создать фотодиод со структурой металл—*i*—*n*, аналогичный *p*—*i*—*n*-диоду (рис. 13, *д*). Специальный диод, изображенный на рис. 13, *е*, представляет собой точечно-контактный фотодиод [23]. Активный объем в таких диодах очень мал, и соответственно чрезвычайно малы время пролета и емкость. Поэтому они пригодны для детектирования излучения при высоких частотах модуляции. Применение этих диодов, однако, ограничено лишь теми ситуациями, в которых излучение может быть сфокусировано в пятно диаметром несколько микрон.

13.3.4. Фотодиод с гетеропереходом

Фотодиод можно создать на основе гетероперехода, в котором переход сформирован между двумя полупроводниками с различной шириной запрещенной зоны (гл. 2). Одно из преимуществ такого фотодиода состоит в том, что его квантовая эффективность не зависит критическим образом от расстояния, на котором расположен переход от поверхности, поскольку материал с более широкой запрещенной зоной можно использовать в качестве «окна», пропускающего оптическое излучение. Кроме того, гетеропереход дает уникальную возможность выбора такого сочетания материалов, которое обеспечивает оптимальные для заданной длины волны оптического сигнала значения квантовой эффективности и быстродействия.

Для того чтобы получить гетеропереход с низкими токами утечки, необходимо обеспечить хорошее согласование постоянных решеток обоих полупроводников. Тройные полупроводниковые соединения $Al_xGa_{1-x}As$ (с прямой зоной при $x < 0,4$), выращенные в виде эпитаксиальных слоев на подложке GaAs, формируют гетеропереходы с совершенным согласованием решеток [24] (с постоянной решетки, равной $5,653 \text{ \AA}$). Эти гетеропереходы играют важную роль в оптоэлектронных приборах, работающих в спектральном диапазоне $0,65-0,85 \text{ мкм}$. Фотодетектор с двойным гетеропереходом ($n-Al_{0,24}Ga_{0,76}As/p-GaAs/p-Al_{0,24}Ga_{0,76}As$), снабженный просветляющим покрытием и работающий на длине волны $\lambda = 0,8075 \text{ мкм}$, имеет квантовую эффективность, равную 92% [25]. Каскадная структура [26], состоящая из пары $p-i-n$ -диодов с двойным гетеропереходом из AlGaAs, соединенных последовательно с туннельным переходом, также характеризуется высокой эффективностью на длине волны $\lambda = 0,815 \text{ мкм}$ и имеет напряжение холостого хода, равное $1,78 \text{ В}$.

Для больших длин волн ($1-1,6 \text{ мкм}$) могут быть использованы тройные соединения, такие, как $Ga_{0,47}In_{0,53}As$ ($E_g = 0,73 \text{ эВ}$) [59] и четверные соединения, такие, как $Ga_xIn_{1-x}As_yP_{1-y}$ (например, $Ga_{0,27}In_{0,73}As_{0,63}P_{0,37}$ с $E_g = 0,95 \text{ эВ}$) [27], которые имеют совершенное согласование решеток с подложкой InP (постоянная

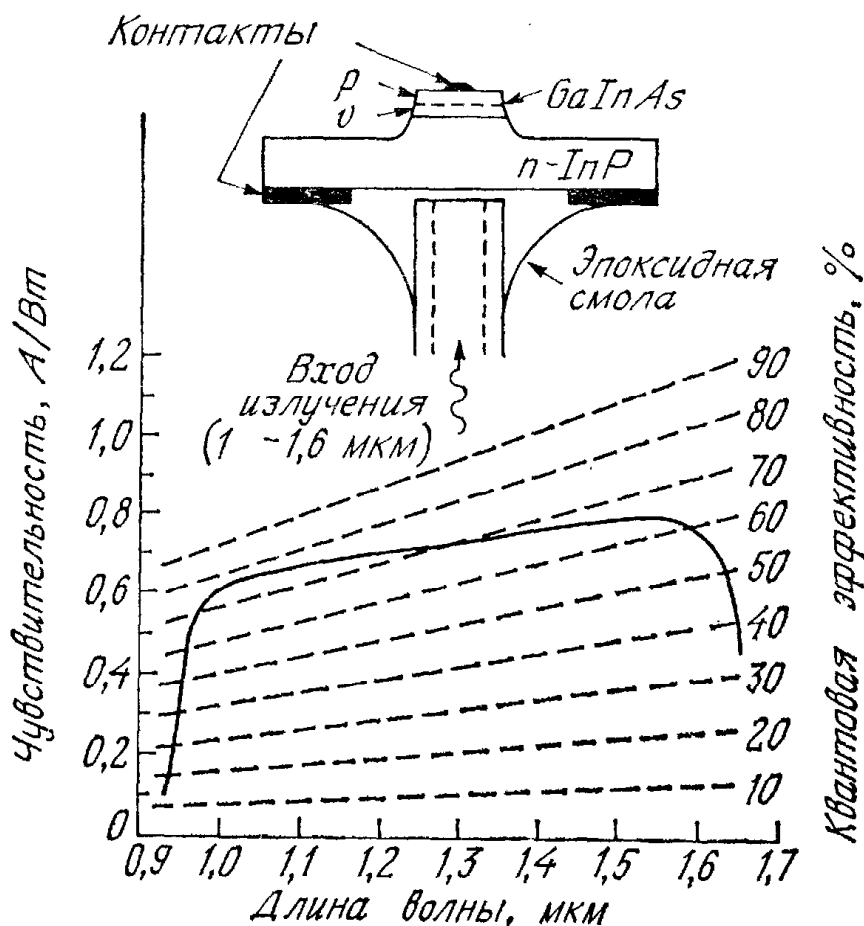


Рис. 17. Зависимость чувствительности и квантовой эффективности GaInAs — $p-i-n$ -фотодиода от длины волны [28].

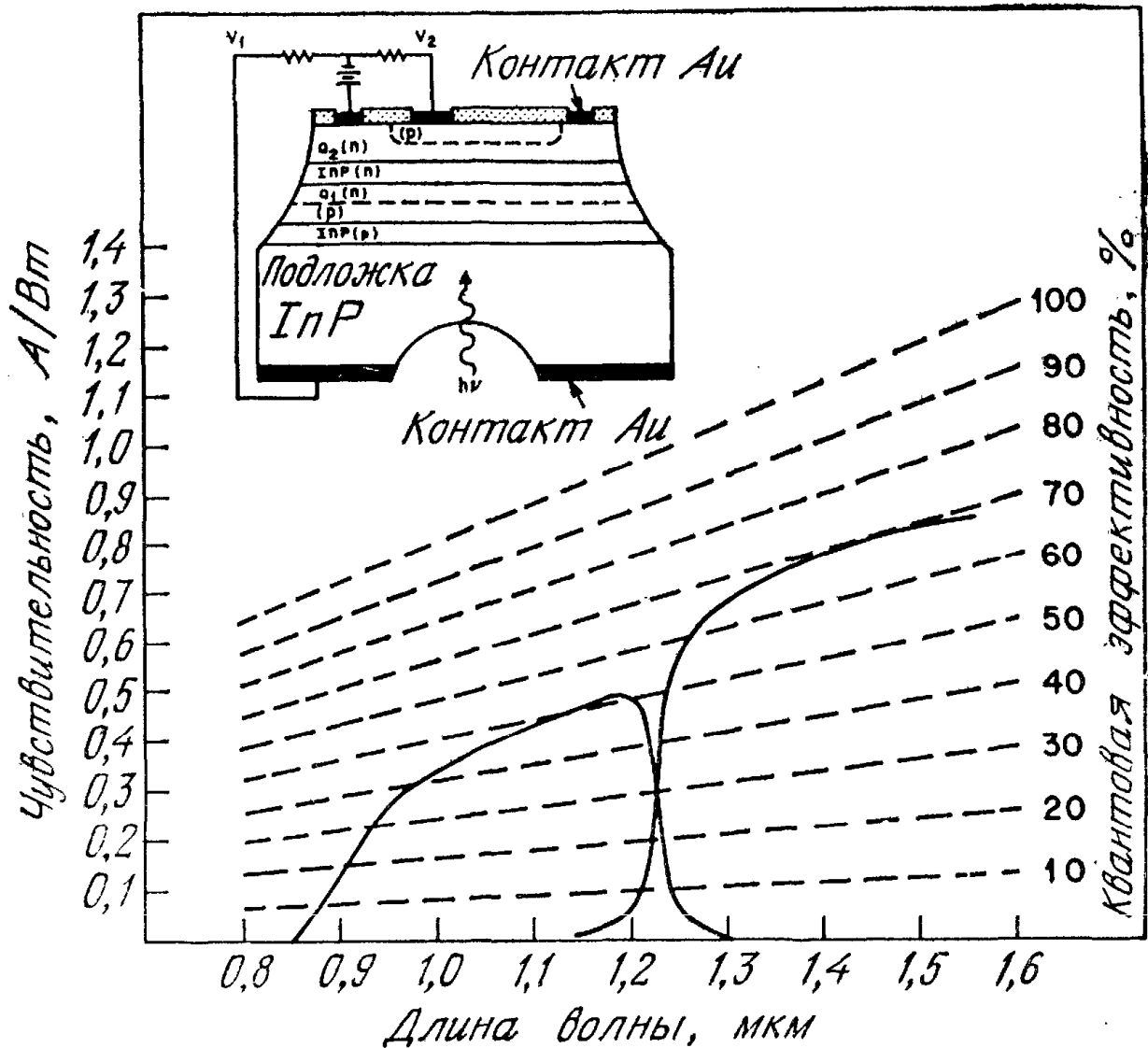


Рис. 18. Зависимость чувствительности и квантовой эффективности спектрального фотодиода от длины волны [56].

решетки $5,8686 \text{ \AA}$, см. рис. 24 гл. 12). На рис. 17 (вставка) изображен фотодиод, освещаемый с обратной стороны, на основе мезоструктуры $p\text{-Ga}_{0,47}\text{In}_{0,53}\text{As}/v\text{-Ga}_{0,47}\text{In}_{0,53}\text{As}/n^+\text{-InP}$. Фотодиод имеет чувствительность более $0,6 \text{ А/Вт}$, которая практически постоянна во всем спектральном диапазоне от $1,0$ до $1,5 \text{ мкм}$ (рис. 17) [28]. Наибольшее значение внешней квантовой эффективности, равное 70% , достигается при длине волны $1,3 \text{ мкм}$; типичные значения квантовой эффективности составляют $55\text{--}60 \%$. Этот прибор, по-видимому, должен иметь более высокие характеристики, чем Ge-фотодиод. Поскольку он обладает прямой запрещенной зоной и, соответственно, большим коэффициентом поглощения вблизи края собственного поглощения, то за счет уменьшения толщины обедненной области можно достичь более высокого быстродействия [29].

Многослойная структура фотодиода с гетеропереходами, в которой одновременно осуществляется детектирование излучения и его разложение на два спектральных интервала, показана на рис. 18

(вставка) [56]. Спектральное разложение обеспечивают два слоя четверного соединения $\text{Ga}_x\text{In}_{1-x}\text{As}_y\text{P}_{1-y}$. Решетки этих слоев согласованы с решеткой InP ; они имеют различный кристаллический состав и, следовательно, характеризуются различными значениями ширины запрещенной зоны. Слой Q_1 имеет более широкую запрещенную зону (1,0 эВ), чем слой Q_2 (0,74 эВ), а подложка прозрачна для излучения падающего с обратной стороны структуры. Перекрытие зон fotocувствительности является следствием оптических перекрестных помех между двумя каналами (рис. 18). Необходимо выбирать такие области применения прибора, в которых спектральный интервал, соответствующий этому перекрытию, не используется. Из графика на рис. 18 следует, например, что на коротковолновой границе интервала перекрытия перекрестные помехи составляют -19 дБ при $\lambda = 1,2$ мкм и уменьшаются до -30 дБ при $\lambda = 1,15$ мкм, а на длинноволновой границе они равны -43 дБ при $\lambda = 1,3$ мкм.

13.4. ЛАВИННЫЙ ФОТОДИОД

Лавинные фотодиоды работают при высоких обратных смещениях, вызывающих лавинное умножение [30]. В результате умножения возникает внутреннее усиление тока. Произведение коэффициента усиления на ширину полосы лавинного фотодиода может достигать более 100 Гц, за счет чего прибор обеспечивает детектирование излучения в микроволновом диапазоне частот модуляции [2, 31]. Квантовая эффективность и быстродействие лавинных фотодиодов определяются так же, как и для обычных фотодиодов. Однако их шумовые свойства и эффект лавинного умножения требуют особого рассмотрения.

13.4.1. Отношение сигнал/шум

Процесс детектирования излучения в лавинном фотодиоде и его эквивалентная схема представлены на рис. 19 [5]. Усиление тока приводит к умножению сигнального, фонового и темнового токов. Среднеквадратичное значение фототока определяется выражением (15), в котором учитывается наличие коэффициента умножения (или коэффициента лавинного умножения) M :

$$i_p = q\eta n P_{\text{opt}} M / \sqrt{2} h\nu. \quad (35)$$

Другие элементы эквивалентной схемы на рис. 19, б полностью совпадают с элементами эквивалентной схемы $p-i-n$ -фотодиода. Среднеквадратичное значение дробового шума после умножения определяется выражением

$$\langle i_s^2 \rangle = 2q (I_p + I_B + I_D) \langle M^2 \rangle B, \quad (36)$$

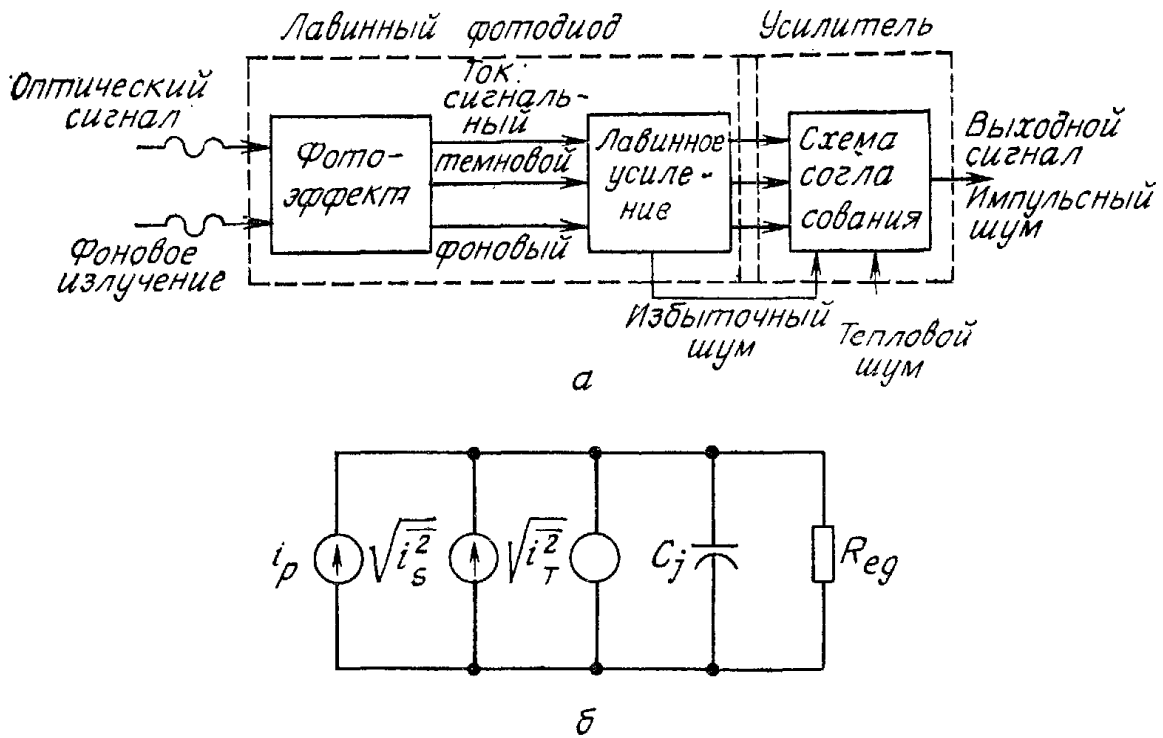


Рис. 19. Процесс детектирования излучения в лавинном фотодиоде (а) и эквивалентная схема лавинного фотодиода (б) [5].

или

$$\langle i_s^2 \rangle = 2q(I_p + I_B + I_D) M^2 F(M) B, \quad (37)$$

где $\langle M^2 \rangle$ — среднеквадратичное значение внутреннего коэффициента усиления и $F(M) = \langle M^2 \rangle / M^2$ — шум-фактор, который представляет собой меру увеличения дробового шума по отношению к идеальному нешумящему усилителю. Тепловой шум, так же как и для $p-i-n$ -фотодиода, определяется уравнением (17).

Отношение сигнал/шум лавинного фотодиода при 100 %-ной модуляции излучения со средней мощностью P_{opt} имеет вид

$$\begin{aligned} S/N &= \frac{1/2 (q\eta P_{opt}/h\nu)^2 M^2}{2q(I_p + I_B + I_D) F(M) M^2 B + 4kTB/R_{eq}} = \\ &= \frac{1/2 (q\eta P_{opt}/h\nu)^2}{2q(I_p + I_B + I_D) F(M) B + 4kTB/R_{eq} M^2}. \end{aligned} \quad (38)$$

Из этого уравнения видно, что лавинное умножение, обеспечивающее уменьшение вклада второго члена в знаменателе, может привести к увеличению отношения сигнал/шум. Ниже показано, что шум-фактор $F(M)$ всегда больше или равен единице и монотонно увеличивается с ростом коэффициента умножения, за исключением тех случаев, когда процесс умножения не вносит дополнительных шумов. Таким образом, существует оптимальное значение коэффициента умножения M , которое обеспечивает максимальное отношение сигнал/шум при заданной оптической мощности. Такое оптимальное умножение имеет место, когда первый и второй члены знаменателя примерно равны.

Используя уравнение (38), можно найти выражение для минимальной оптической мощности P_{opt} , которая необходима для

получения заданного отношения S/N при наличии лавинного умножения. Это выражение имеет вид

$$(P_{\text{opt}})_{\text{мин}} = \frac{2h\nu}{\eta} \left(\frac{S}{N} \right) \left\{ 1 + \left[1 + \frac{I_{eq}}{qBF(M)^2(S/N)} \right]^{1/2} \right\}, \quad (39)$$

где

$$I_{eq} \equiv (I_B + I_D) F(M) + 2kT/qR_{eq}M^2. \quad (40)$$

В условиях, когда $I_{eq}/qBF^2(S/N)$ пренебрежимо мало, превышение минимально обнаружимой мощности над собственным квантовым шумом сигнала при бесшумовом умножении ($F = 2$) составляет 3 дБ. При детектировании сигналов высокой частоты и с широкой полосой пропускания минимально обнаружимая мощность ограничивается тепловым шумом нагрузочного сопротивления и шум-фактором последующей стадии усиления. В этих условиях отношение $I_{eq}/qBF(M)^2(S/N)$ больше единицы, и мощность, эквивалентная шуму, определяется выражением

$$\text{NEP} = \sqrt{2} (h\nu/\eta) [I_{eq}/qF(M)^2]^{1/2}. \quad (41)$$

Поскольку лавинное усиление позволяет значительно уменьшить NEP, лавинные фотодиоды имеют существенное преимущество перед обычными фотодиодами.

13.4.2. Коэффициент умножения

Коэффициент лавинного усиления, называемый также коэффициентом умножения, был введен в гл. 2. Выражение для коэффициента умножения электронов при низких частотах имеет вид

$$M = \left\{ 1 - \int_0^W \alpha_n \exp \left[- \int_0^x (\alpha_n - \alpha_p) dx' \right] dx \right\}^{-1}, \quad (42)$$

где W — ширина обедненного слоя, α_n и α_p — коэффициенты ионизации электронов и дырок соответственно. При независимых от координаты коэффициентах ионизации, что имеет место, например, в $p-i-n$ -диоде, коэффициент умножения электронов, инжектированных в область сильного поля, в точке $x = 0$ равен

$$M = \frac{(1 - \alpha_p/\alpha_n) \exp [\alpha_n W (1 - \alpha_p/\alpha_n)]}{1 - (\alpha_p/\alpha_n) \exp [\alpha_n W (1 - \alpha_p/\alpha_n)]}. \quad (43)$$

Если $\alpha_n = \alpha_p$, выражение для M принимает простую форму

$$M = 1/(1 - \alpha_n W). \quad (44)$$

Если $\alpha_n W = 1$, возникает пробой полупроводника. Восстанавливающийся лавинный процесс создает большое число носителей в области сильного поля в течение длительного времени после прохождения через эту область первичных электронов. Чем выше коэффициент умножения, тем дольше продолжается лавинный

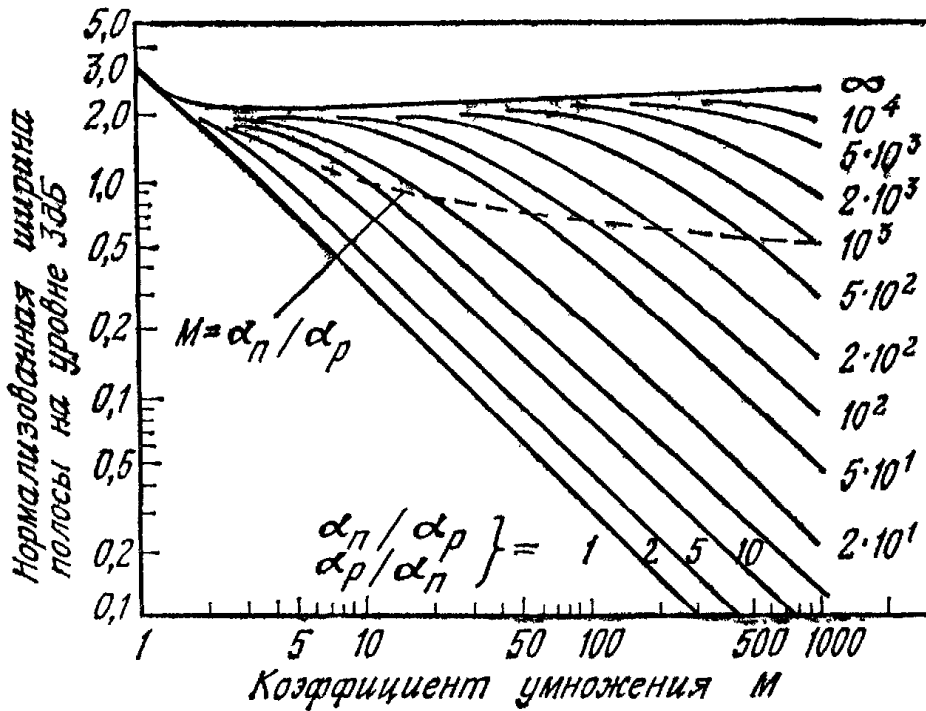


Рис. 20. Теоретическая зависимость граничной частоты (на уровне 3 дБ), умноженной на $2\pi\tau_{av}$ (τ_{av} — среднее время пролета носителей), от низкочастотного коэффициента умножения при различных значениях α_p/α_n (инжекция электронов) и α_n/α_p (инжекция дырок) для лавинного фотодиода [32].

процесс. Эта взаимосвязь устанавливается произведением усиления на ширину полосы. Когда коэффициенты ионизации равны, а $M \rightarrow \infty$, выражение для произведения коэффициента усиления тока на ширину полосы имеет вид

$$\text{Усиление} \times \text{Ширина полосы} = \frac{3}{2\pi\tau_{av}}, \quad (45)$$

где $\tau_{av} = (t_n + t_p)/2$; t_n — время пролета электронов, равное W/v_n (v_n — скорость насыщения электронов), и t_p — время пролета дырок, которое можно записать аналогичным образом.

На рис. 20 приведены значения ширины полосы [32], рассчитанной для идеализированного лавинного $p-i-n$ -фотодиода с постоянной величиной электрического поля в области лавинного умножения. Расчет проведен для ширины полосы B , соответствующей ослаблению 3 дБ и нормализованной на $2\pi\tau_{av}$, в зависимости от низкочастотного коэффициента умножения, причем в качестве параметра взято отношение коэффициентов лавинного умножения. Выше кривой, соответствующей $M = \alpha_n/\alpha_p$, ширина полосы в значительной степени определяется временем пролета носителей и практически не зависит от усиления. Ниже этой кривой наблюдается почти прямолинейная зависимость, соответствующая постоянным значениям произведения усиления на ширину полосы:

$$M(\omega)\omega \approx \frac{1}{(W/v_n)(\alpha_p/\alpha_n)}. \quad (46)$$

Таким образом, для получения больших значений произведения усиления на ширину полосы необходимо, чтобы α_p/α_n было велико, а W — мало.

В реальных приборах максимально достижимое усиление по постоянному току при высоких интенсивностях излучения ограничено последовательным сопротивлением и влиянием пространственного заряда. Эти факторы можно свести к одному эффективному последовательному сопротивлению R . Коэффициент умножения для фотогенерированных носителей может быть описан эмпирическим соотношением [33]:

$$M_{ph} = \frac{I - I_{MD}}{I_P - I_D} = \frac{1}{\left[1 - \left(\frac{V_R - IR}{V_B}\right)^n\right]}, \quad (47)$$

где I — общий усиленный ток, I_P — общий первичный ток (до усиления), I_D и I_{MD} — первичный и усиленный темновые токи соответственно, V_R — напряжение обратного смещения, V_B — напряжение пробоя и n — показатель, зависящий от полупроводникового материала, профиля легирования и длины волны излучения. При высокой интенсивности излучения ($I_P \gg I_D$) и $IR \ll V_B$ максимальная величина коэффициента умножения фототока определяется выражением

$$(M_{ph})_{\max} \simeq \frac{I}{I_P} = \frac{1}{1 - \left(\frac{V_R - IR}{V_B}\right)^n} \Bigg|_{V_R \rightarrow V_B} \approx \frac{1}{nIR/V_B}, \quad (48)$$

или

$$(M_{ph})_{\max} = \sqrt{V_B/nI_P R}. \quad (49)$$

При малых фототоках максимальное умножение ограничено темновым током и в уравнении (49) необходимо вместо I_P подставить I_D . Таким образом, весьма важно, чтобы темновой ток был как можно меньше и не ограничивал ни $(M_{ph})_{\max}$, ни минимально обнаружимую мощность (уравнение (39)).

13.4.3. Шум лавинного умножения

Лавинный процесс является по природе статистическим процессом, так как электронно-дырочные пары, генерируемые в данной точке обедненной области, претерпевают неодинаковое умножение. Поскольку коэффициент умножения флюктуирует, среднеквадратичное значение усиления превышает квадрат его среднего значения. Избыточный шум характеризуется шум-фактором $F(M) = \langle M^2 \rangle / M^2$, который зависит от отношения коэффициентов ионизации α_p/α_n и низкочастотного коэффициента умножения M . Если $\alpha_n = \alpha_p$, на каждый фотоноситель, возбужденный падающим излучением, в области умножения приходится три носителя (пер-

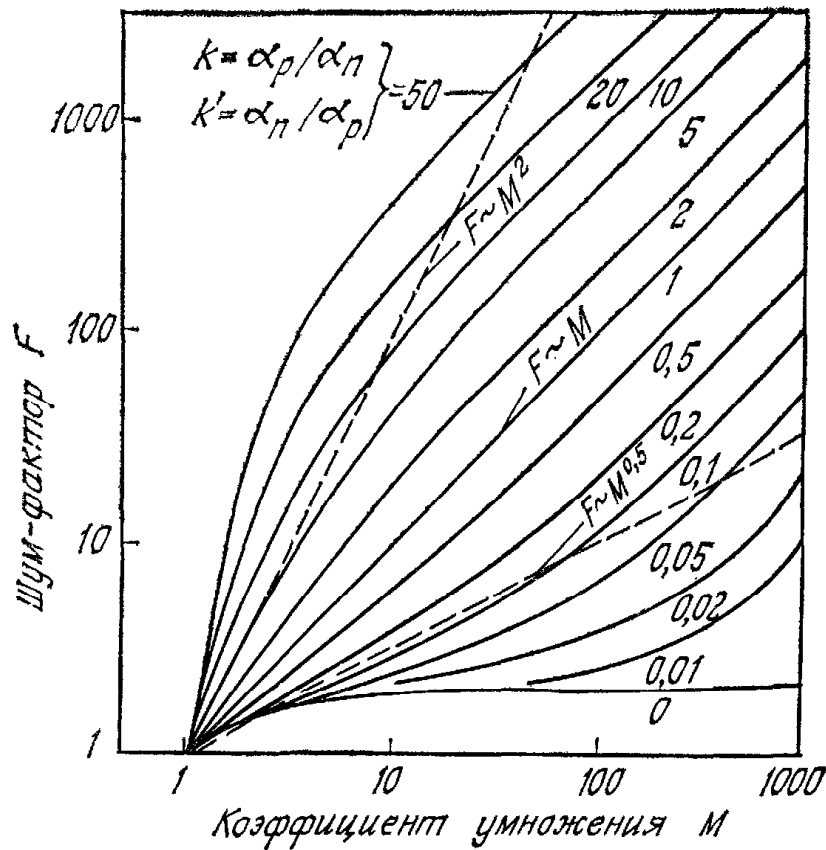


Рис. 21. Теоретическая зависимость шум-фактора от коэффициента умножения для различных значений α_n/α_p и α_p/α_n [34].

вичный носитель и вторичные электрон и дырка). Поэтому флуктуации, изменяющей число носителей на один, соответствует существенное относительное изменение их числа и шум-фактор оказывается большим. Если же один из коэффициентов ударной ионизации пренебрежимо мал (например, $\alpha_p \rightarrow 0$), то на каждый фотоноситель в области умножения приходится около M носителей. В этом случае флуктуация, соответствующая изменению этого числа на один носитель, вызовет лишь незначительное относительное возмущение. Таким образом, следует ожидать, что при существенно различных α_n и α_p значение шум-фактора будет мало.

Если инжектируются только электроны, то шум-фактор можно записать в виде [34]

$$F = kM + (2 - 1/M)(1 - k), \quad (50)$$

где $k \equiv \alpha_p/\alpha_n$ предполагается постоянным во всей области лавинного умножения. Это выражение применимо и для случая инжекции дырок, если k заменить на $k' \equiv \alpha_n/\alpha_p$. Для двух специальных случаев, а именно $\alpha_p = \alpha_n$ (т. е. $k = 1$) и $\alpha_p \rightarrow 0$ (т. е. $k = 0$), из уравнения (50) получаем значения $F = M$ и $F = 2$ (при больших M) соответственно. Зависимость шум-фактора от коэффициента умножения при различных отношениях коэффициентов ионизации

приведена на рис. 21. Видно, что снижение уровня дополнительных шумов достигается при малых значениях k и k' .

На рис. 22 представлены экспериментальные результаты [35], полученные для лавинного Si-фотодиода на частоте 600 кГц и при токе первичной инжекции, равном 0,1 мкА. Светлыми кружками обозначен шум первичного дырочного фототока, возбуждаемого коротковолновым излучением (см. вставку). Темными кружками представлен шум для случая, когда имеет место первичная инжекция электронов. Электронная инжекция характеризуется существенно меньшим шум-фактором, чем дырочная, поскольку в кремнии α_n значительно больше α_p .

Результаты, представленные на рис. 21, могут быть использованы применительно к лавинным $p-i-n$ -фотодиодам и фотодиодам с трехслойной базой, которые имеют постоянное поле в области лавинного умножения. Для обычных лавинных фотодиодов с изменяющимся электрическим полем в области лавинного умножения коэффициенты ионизации в уравнении (50) должны

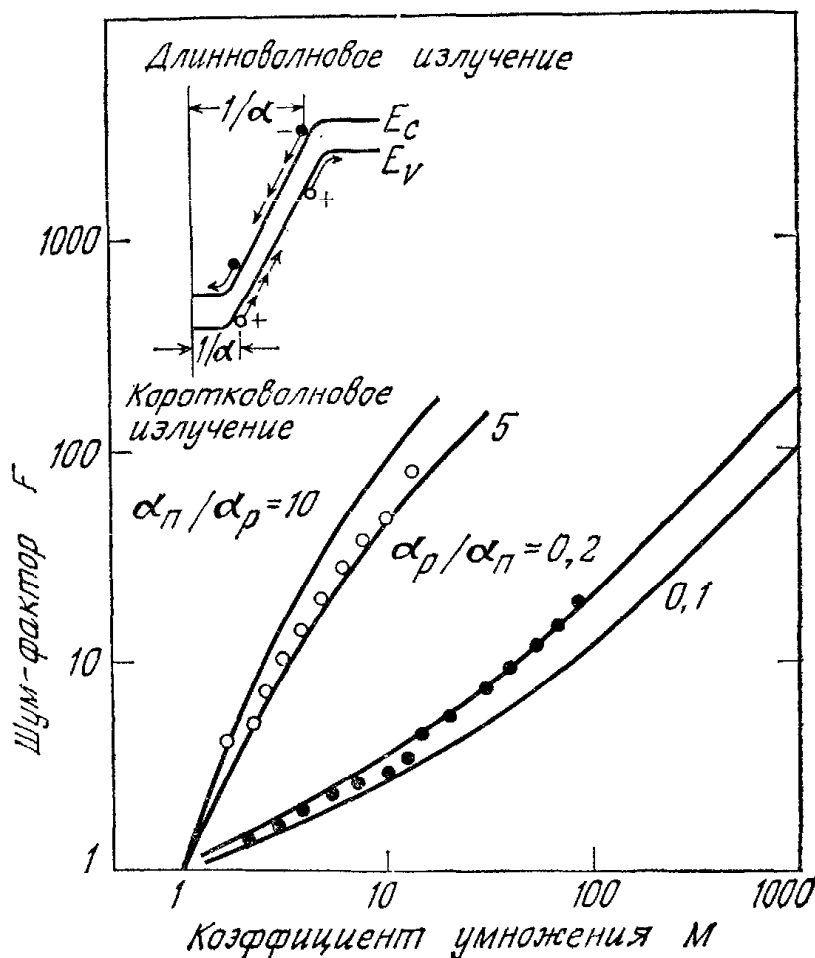


Рис. 22. Результаты экспериментального исследования шум-фактора кремниевого лавинного фотодиода с током первичной инжекции, равным 0,1 мкА. На вставке показана зонная диаграмма лавинного фотодиода в условиях электронной и дырочной инжекции, зависящей от длины волны падающего излучения [35].

— теоретические данные; ● для инжекции электронов; ○ для инжекции дырок.

быть усложнены [36]; k необходимо заменить на k_{eff} , а k' — на k'_{eff} , где

$$k_{\text{eff}} = \int_0^W \alpha_p(x) M^2(x) dx \bigg/ \int_0^W \alpha_n(x) M^2(x) dx, \quad (51)$$

$$k'_{\text{eff}} = k_{\text{eff}} \left[\int_0^W \alpha_p(x) M(x) dx \bigg/ \int_0^W \alpha_n(x) M(x) dx \right]^{-2}. \quad (52)$$

Избыточный шум возникает в том случае, когда излучение поглощается как в n -, так и p -областях, прилегающих к переходам, вызывая инжекцию в область лавинного умножения как электронов, так и дырок. Например, для $k_{\text{eff}} = 0,005$ и $M = 10$ значение шум-фактора возрастает примерно от 2 при чисто электронной инжекции до 20 при электронной инжекции, составляющей 10 % [37]. Таким образом, для достижения в лавинных фотодиодах низкого уровня шума и широкой полосы частот необходимо, чтобы коэффициенты ионизации различались как можно более сильно, а лавинный процесс стимулировался носителями, обладающими более высоким коэффициентом ионизации.

13.4.4. Рабочие характеристики

В лавинных фотодиодах необходимо обеспечить пространственно однородное умножение по всей фоточувствительной площади [38]. Должно быть исключено образование микроплазм, представляющих собой локализованные участки, в которых пробивное напряжение меньше, чем во всем переходе. Количество микроплазм в активной области уменьшается при использовании материалов с низкой плотностью дислокаций и конструкций приборов с активной областью, не превышающей по своим размерам область, которая необходима для сбора падающего светового пучка (обычно диаметром от нескольких микрометров до 100 мкм). Избыточные токи утечки по периферии перехода, обусловленные краевым эффектом [39] или сгущением линий электрического поля, устраняются за счет применения охранного кольца или поверхностно-конических структур.

На рис. 23 показаны некоторые конструкции приборов [2]. Охранное кольцо в структуре на рис. 23, а должно иметь низкий градиент примеси на границе p — n -перехода и достаточно большой радиус кривизны, чтобы центральная область резкого p — n^+ -перехода пробивалась раньше, чем область охранного кольца. Охранное кольцо можно использовать и в более сложной n^+ — p — n — p^+ -структуре (рис. 23, б). Оно может быть применено также для устранения высокого поля на периферии контакта в лавинном фотодиоде Шоттки (рис. 23, в). Мезаструктура (рис. 23, г) и коническая

структура (рис. 23, д) характеризуются низким поверхностным полем по всей площади перехода и однородным лавинным пробоем. Для детектирования излучения с длинами волн вблизи края собственного поглощения может быть использован лавинный фотодиод с боковым освещением, с помощью которого можно улучшить квантовую эффективность и отношение сигнал/шум (рис. 23, е).

Лавинные фотодиоды изготавливаются на основе различных полупроводников, в том числе Ge, Si, соединений типа $A^{III}B^V$ и других. Основными факторами при выборе того или иного полупроводника являются квантовая эффективность в заданном спектральном диапазоне, быстродействие и шумы. Мы рассмотрим несколько типичных примеров, иллюстрирующих характеристики приборов.

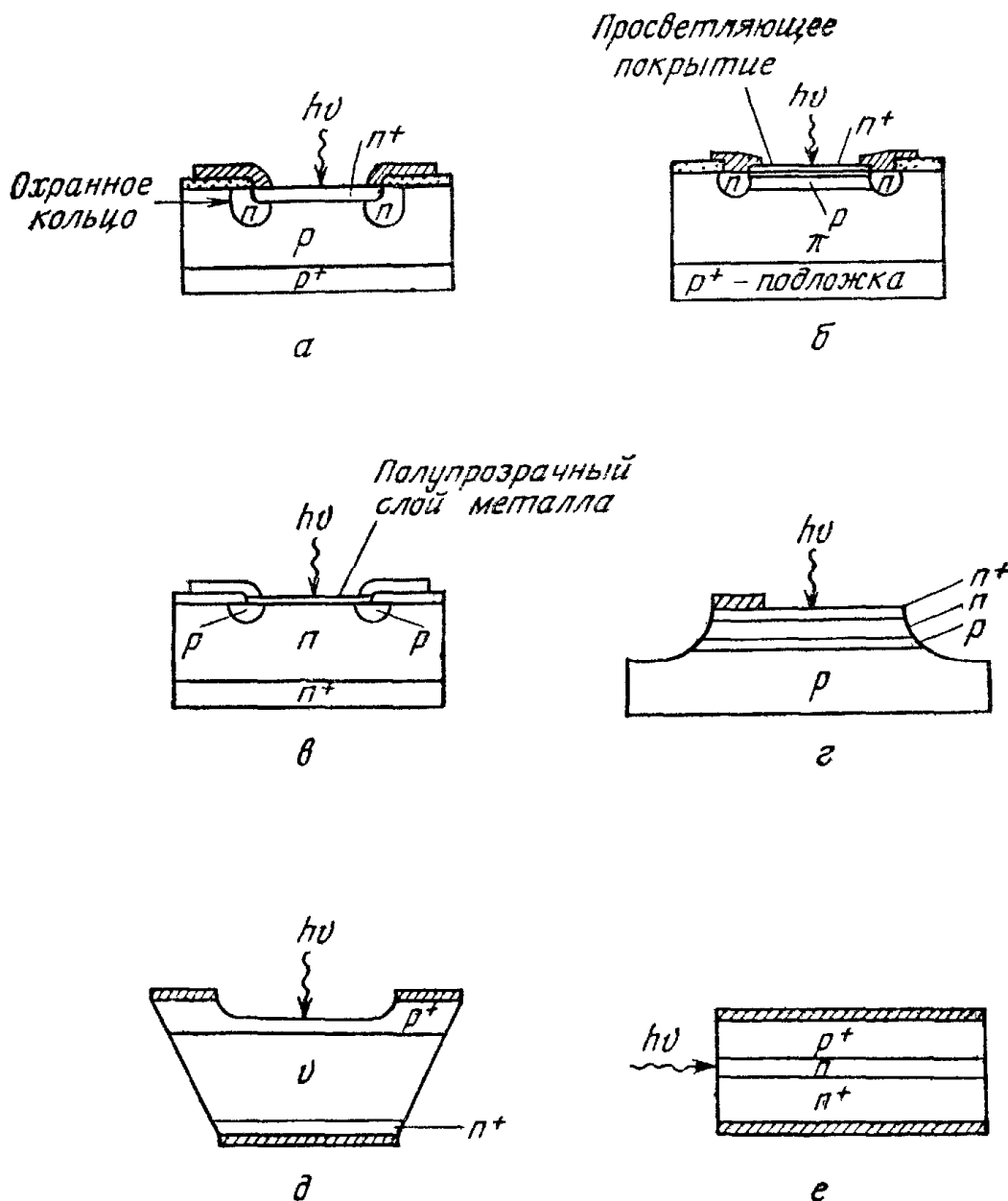


Рис. 23. Структуры некоторых лавинных фотодиодов [2].
 а — структура с охранным кольцом; б — $n^+ - p - n - p^+$ -структура с охранным кольцом; в — диод со структурой металл — полупроводник; г — мез-структура; е — $p^+ - n - n^+$ -структура с боковым освещением.

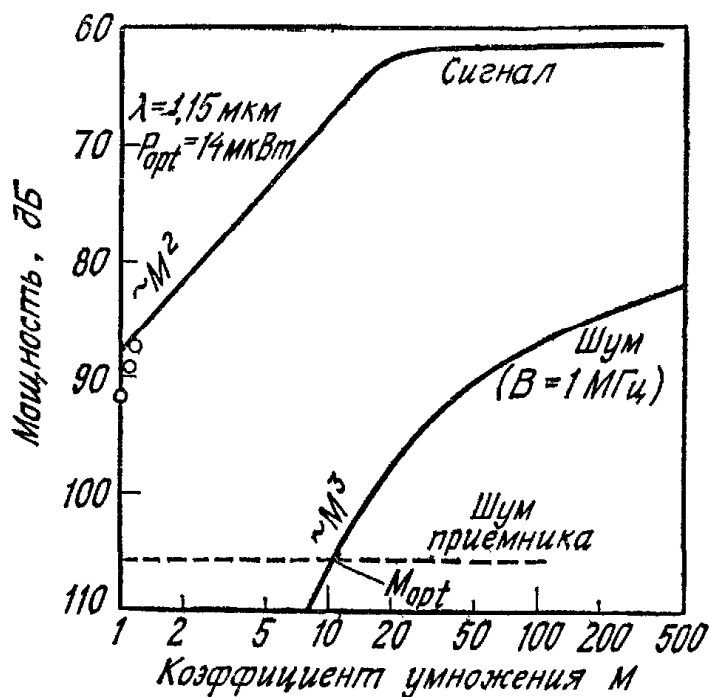


Рис. 24. Мощность сигнала и шума относительно уровня 1 мВт на выходе лавинного Ge-фотодиода при частоте 8 ГГц в полосе 1 МГц. На графике обозначена оптимальная рабочая точка M_{opt} , в которой достигается максимальное значение отношения сигнал/шум и чувствительности [2].

Германиевые лавинные фотодиоды обеспечивают высокую квантовую эффективность в спектральном диапазоне 1—1,6 мкм. Поскольку коэффициенты ионизации электронов и дырок в Ge сравнимы, то шум-фактор близок по величине к $F = M$ (уравнение (50)) и среднеквадратичный дробовой шумовой ток изменяется по закону M^3 (уравнение (37)). Структура, показанная на рис. 23, а, была применена для изготовления высокочастотного Ge-фотодиода [33, 40]. Удельное сопротивление p -слоя составляло 0,5 Ом·см, а его толщина 150 мкм. После диффузии сурьмы для создания охранного кольца глубиной 7,5 мкм проводилась диффузия мышьяка на глубину 0,4 мкм для формирования n^+ -слоя. Прибор имел квантовую эффективность около 80 % во всем спектральном диапазоне от 1 до 1,55 мкм (рис. 7). Плотность тока при комнатной температуре и напряжении смещения, равном $\frac{1}{2} V_B$ (где V_B — пробивное напряжение), составила $\sim 3 \cdot 10^{-4}$ А/см² [33]. Время нарастания импульса фотоотклика при воздействии АИГ-лазера ($\lambda = 1,06$ мкм) равнялось 100 пс.

На рис. 24 представлены результаты исследования [2] выходного сигнала и шума Ge-лавинного фотодиода при частоте модуляции 3 гГц и первичном токе 5,3 мкА. Отметим, что при $M < 30$ мощность сигнала растет как M^2 , а мощность шума — как M^3 . Это хорошо согласуется с теоретическими предположениями. Отметим также, что наибольшее отношение сигнал/шум (~ 40 дБ) получено при $M \simeq 10$, т. е. когда шумы диода примерно равны шумам приемника. При более высоких значениях M отношение S/N уменьшается, так как шум лавины растет быстрее, чем умноженный сигнал.

Кремниевые лавинные фотодиоды особенно эффективны в спектральном диапазоне 0,6—1,0 мкм, в котором для приборов с про-

светляющим покрытием достигнут квантовый выход, близкий к 100 %. Отношение коэффициентов ионизации электронов и дырок ($k = \alpha_p/\alpha_n$) в кремнии сильно зависит от электрического поля; оно изменяется примерно от 0,1 при поле $3 \cdot 10^5$ В/см до 0,5 при поле 6×10^5 В/см. Поэтому для получения минимальных шумов необходимо, чтобы электрическое поле лавинного пробоя было низким и чтобы лавинное умножение стимулировалось электронами.

Идеализированный профиль легирования в $p-\pi-p-\pi-n^+$ -структуре показан на рис. 25, а. Этот профиль аналогичен профилю трехслойной базы в структуре лавинно-пролетного диода (гл. 10). Структура $p-\pi-p-\pi-n^+$ называется структурой со «сквозным проколом» [41], поскольку электрическое поле в ней распространяется на все пространство от n^+ - π -области умножения до p^+ -слоя (рис. 25, б). В области дрейфа с низким электрическим полем носители могут переноситься со скоростью, достигающей их скорости насыщения (10^7 см/с при $\mathcal{E}_D > 10^4$ В/см²). Максимальное поле \mathcal{E}_m в области умножения можно регулировать за счет изменения толщины b . Условие пробоя можно записать в виде [42]

$$\alpha_n b = \frac{\ln k}{k-1} \quad (k \equiv \alpha_p/\alpha_n), \quad (53)$$

а пробивное напряжение определяется выражением

$$V_B \simeq \mathcal{E}_m b + \mathcal{E}_D (W - b). \quad (54)$$

С целью оптимизации характеристик прибора для заданной длины волны можно выбрать сначала значение W (из условия $W = 1/\alpha$), а затем независимо подобрать величину b . Как видно из рис. 25, в, большая часть излучения поглощается в π -области; при этом лавинный процесс стимулируется электронами, поступающими в область умножения (рис. 25, г). В приборах с $p^+-\pi-p-\pi-p^+$ -структурой можно получить высокие значения квантовой эффективности, быстродействия и отношения сигнал/шум.

Формирование узкой p^+ -области с целью создания профиля, показанного на рис. 25, а, может быть сопряжено с большими трудностями, поэтому на практике используют приборы со структурой $p^+-\pi-p-n^+$, в которой p -область занимает все пространство до n -подложки. Такой профиль аналогичен профилю двухслойной базы в структуре лавинно-пролетного диода. В кремниевом лавинном фотодиоде со «сквозным проколом» [37], в котором полная ширина обедненного слоя составляла 200 мкм, а π -область имела удельное сопротивление 5000 Ом·см, была получена квантовая эффективность ~ 90 % при $\lambda = 0,9$ мкм. При $M = 100$ шум-фактор принимает значение, равное всего лишь 4, а при $M = 1000$ он возрастает до 20, что находится в соответствии с низкой величиной $k_{\text{eff}} = 0,016$.

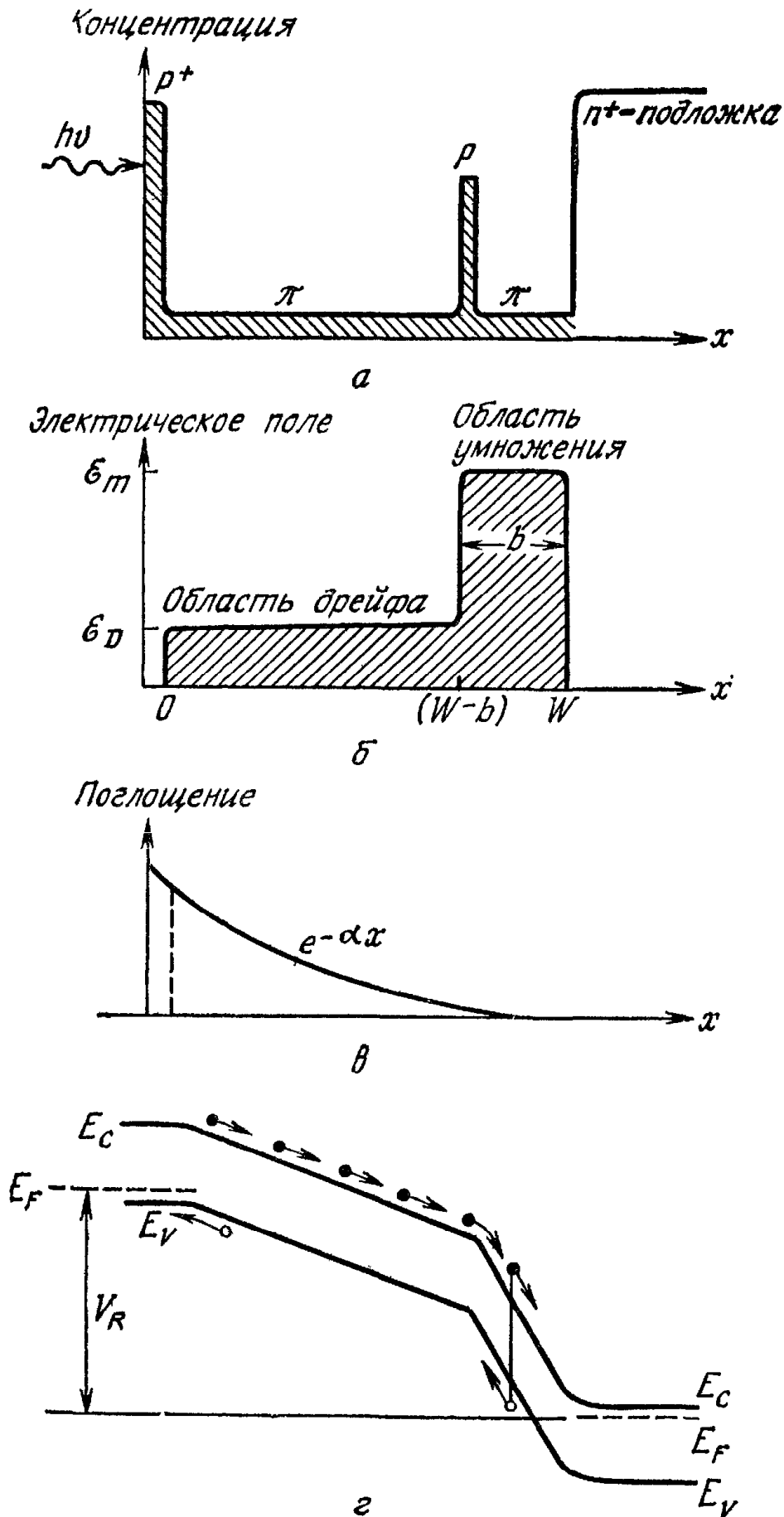


Рис. 25. Лавинный фотодиод со «сквозным проколом» на основе структуры с профилем трехслойной базы.

a — профиль легирования; *b* — распределение поля; *в* — поглощение падающего излучения; *г* — зонная диаграмма, которая иллюстрирует процесс умножения, стимулируемый электронами.

Если сигнальный фототок существенно больше фонового и темнового токов, то доминирующий дробовой шум в выражении (38) описывается членом $2qI_p M^2 FB$, в котором F определяется выражением (50). Оптимальный коэффициент умножения M_{opt} находится из условия $d(S/N)/dM = 0$. Подставляя M_{opt} в выражение (38), получим максимальное соотношение сигнал/шум, соответствующее условию большого сигнального фототока [43]:

$$(S/N)_{\max} \sim (\eta/\sqrt{k_{\text{eff}}})^{4/3}. \quad (55)$$

Таким образом, для получения предельно большого отношения S/N необходимо увеличивать $\eta/\sqrt{k_{\text{eff}}}$. На рис. 26 приведена зависимость $\eta/\sqrt{k_{\text{eff}}}$ от b при различных пробивных напряжениях в кремниевом лавинном фотодиоде с профилем трехслойной базы. Длина волны 0,85 мкм ($\alpha = 600 \text{ см}^{-1}$) соответствует излучению GaAs-лазера. Отметим, что при заданном V_B зависимость имеет максимум, который увеличивается с ростом V_B . Проведенная оценка справедлива и для других примесных профилей.

На рис. 27, а показана структура, позволяющая создавать приборы на кремниевых пластинах большого диаметра и обеспечи-

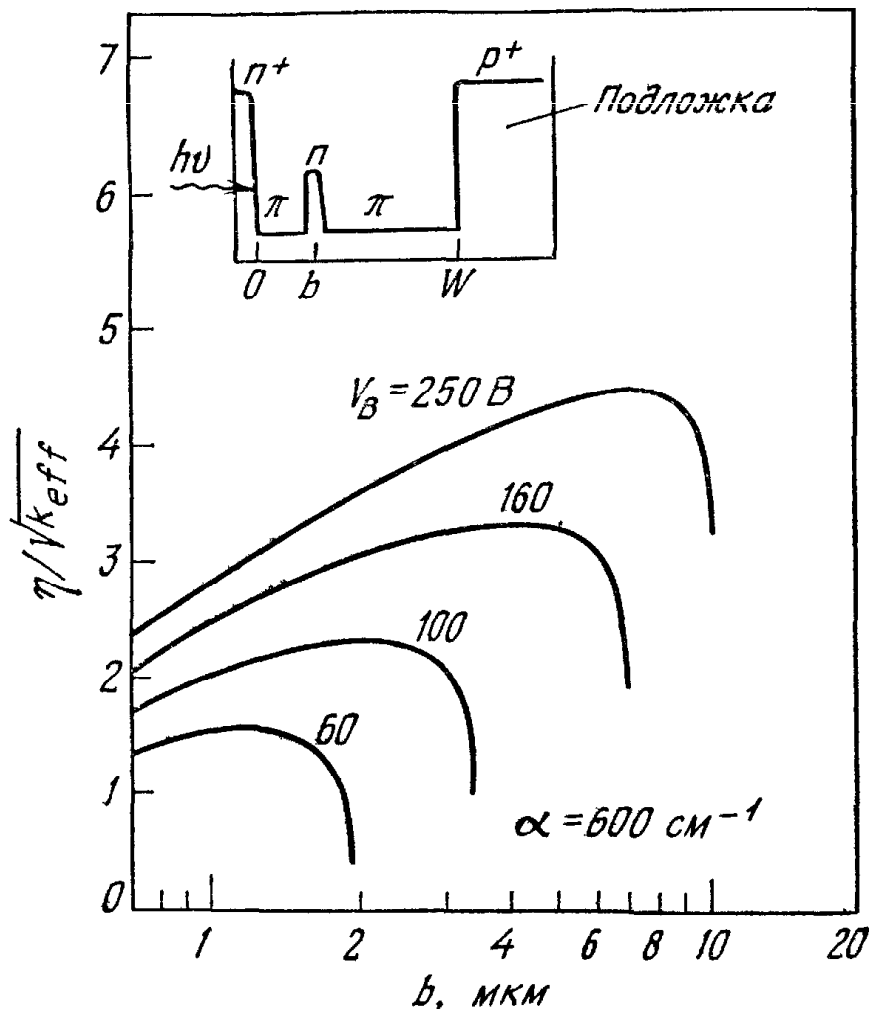


Рис. 26. Зависимость отношения $\eta/\sqrt{k_{\text{eff}}}$ от толщины области умножения b для кремниевого лавинного фотодиода с трехслойной базой [43].

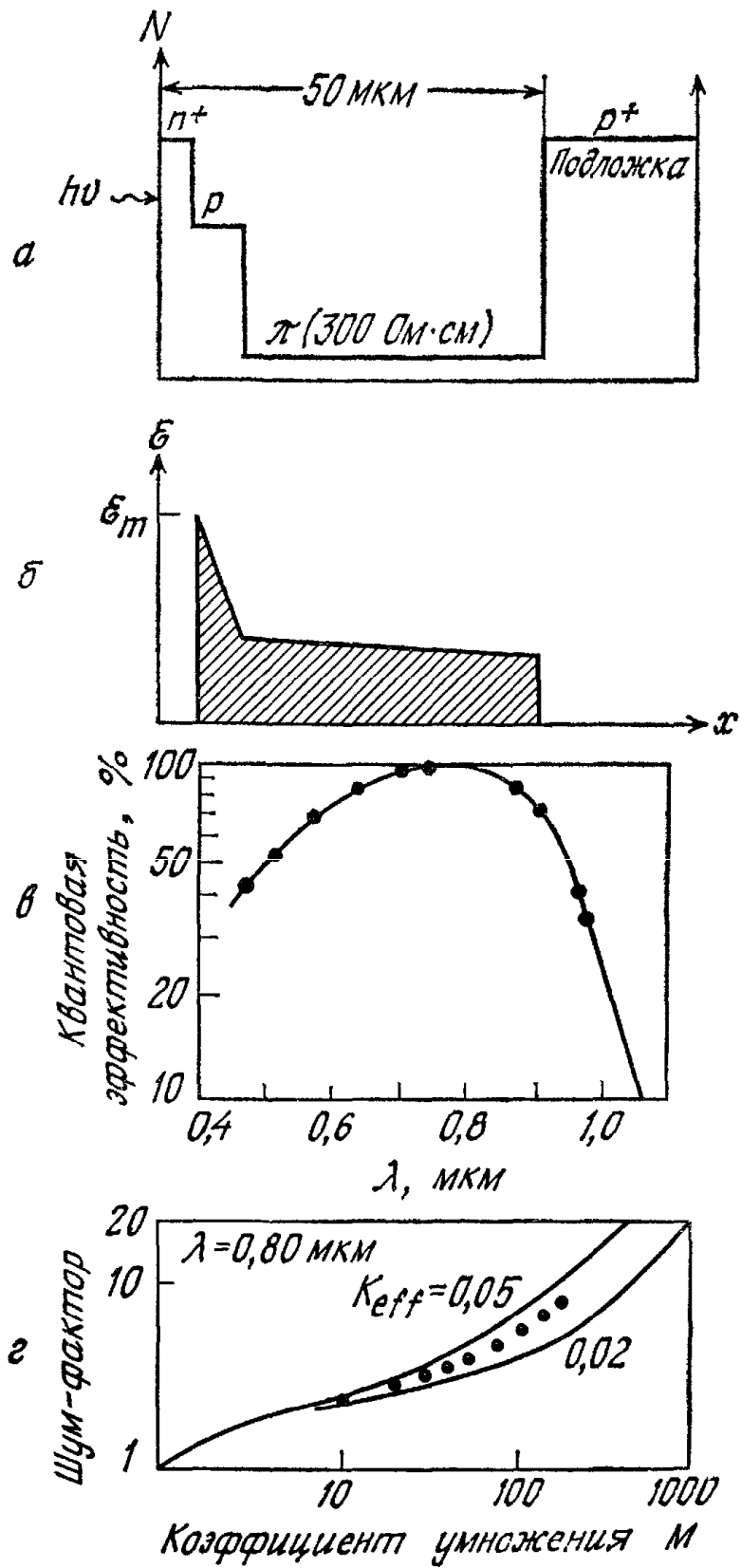


Рис. 27. Кремниевый лавинный фотодиод с двухслойной базой [44].
 а — профиль легирования; б — распределение поля; в — квантовая эффективность; г — шум-фактор [44].

вающая хорошее управление профилем легирования с помощью процессов ионной имплантации и диффузии [44]. Поперечный разрез структуры представлен на рис. 23, б. На рис. 27, б изображено распределение электрического поля, для которого характерно наличие узкой области лавинного умножения и протяженной области дрейфа. В приборе с просветляющим покрытием из пленок $\text{SiO}_2\text{—Si}_3\text{N}_4$ квантовая эффективность близка к 100 % на длине волны 0,8 мкм (рис. 27, в). Поскольку вклад дырок в стимулирование процесса умножения незначителен, то прибор обладает более высоким шум-фактором ($k_{\text{eff}} \simeq 0,04$), чем структура, показанная на рис. 25.

Лавинные фотодиоды со структурой металл—полупроводник (с барьером Шоттки) эффективны в видимой и ультрафиолетовой областях спектра. Характеристики лавинных фотодиодов с барьером Шоттки подобны характеристикам фотодиодов с $p\text{—}n$ -переходом [45]. Фотодиод с барьером Шоттки был изготовлен на кремниевой подложке n -типа с удельным сопротивлением 0,5 Ом·см, с тонкой пленкой PtSi ($\sim 100 \text{ \AA}$) на поверхности и охранным кольцом [46]. Поскольку токи утечки по периферии перехода устранены благодаря применению охрannого кольца, то в такой структуре может быть получен идеальный обратный ток насыщения с плотностью $2 \cdot 10^{-7} \text{ А/см}^2$. В фотодиодах с барьером Шоттки в режиме лавинного умножения может быть получено 35-кратное усиление амплитуды высокочастотного импульса фототока длительностью 0,8 нс (рис. 28). Разность между фототоком в отсутствие умножения и умноженным фототоком показывает, какое большое усиление по току можно получить для заданной длительности импульса (в рассматриваемом случае 0,8 нс) при определенном уровне освещения. Средняя мощность излучения лазера на длине волны 6328 \AA равна 0,8 мВт, а пиковая мощность в импульсе, соответствующая 0 дБ на рис. 28, примерно равна 7 мВт. Измерения шума лавинного умножения PtSi—Si-диодов показали, что шум умноженного фототока растет по закону $\sim M^3$ при возбуждении светом в видимой области спектра. При уменьшении длины волны преобладает первичный инжекционный фототок электронов и шумы уменьшаются, что находится в хорошем согласии с теорией.

Лавинные фотодиоды с барьером Шоттки, изготовленные на кремниевых подложках n -типа, особенно перспективны для применения в качестве высокочастотных фотодетекторов ультрафиолетового излучения. Ультрафиолетовое излучение, проходящее через тонкие металлические электроды, поглощается в приповерхностном слое кремния толщиной 100 \AA . Вследствие этого умножение носителей в основном стимулируется электронами, что обеспечивает низкий шум и большое произведение коэффициента усиления на ширину полосы. Лавинное умножение фототока повышает чувстви-

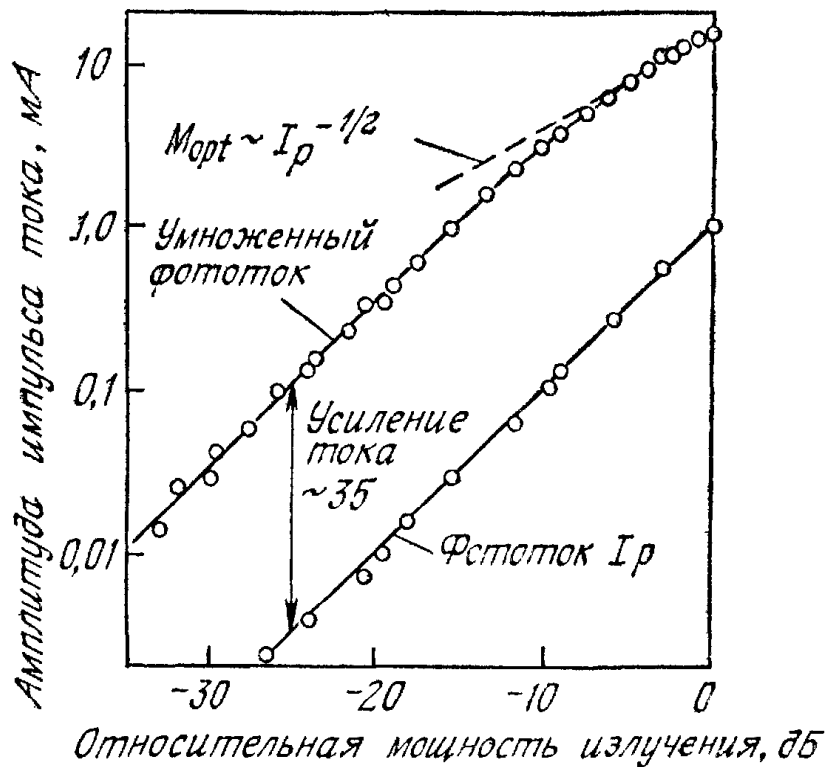


Рис. 28. Зависимость амплитуды импульсов фототока в лавинном фотодиоде с барьером Шоттки от интенсивности излучения при $\lambda = 0,6328$ мкм и длительности световых импульсов 0,8 нс. Нижняя кривая соответствует фототоку при низком обратном смещении. На верхней кривой представлены максимальные значения амплитуды импульсов умноженного фототока, которые достигаются при смещении, равном напряжению пробоя [45].

тельность детектирующих систем, состоящих из фотодиодов с барьером Шоттки и приемника. Кроме того, имеется возможность усиления высокочастотных импульсов фототока.

Лавинные фотодиоды с гетеропереходами, особенно из соединений типа $A^{III}B^V$, обладают рядом потенциальных преимуществ по сравнению с приборами на основе Ge и Si. Изменяя состав соединений, можно регулировать спектральную характеристику чувствительности прибора. Поскольку прямозонные соединения типа $A^{III}B^V$ имеют высокий коэффициент поглощения, то в приборах на их основе может быть получена высокая квантовая эффективность даже в тех случаях, когда с целью достижения высокого быстродействия в них используются тонкие обедненные слои. Кроме того, прозрачный слой гетероструктуры (поверхностный слой с более широкой запрещенной зоной) может быть выращен по технологии, которая обеспечивает высокочастотные характеристики и минимальные потери, связанные с рекомбинацией фотоносителей.

Лавинные фотодиоды были изготовлены [47, 48] с использованием различных систем из соединений типа $A^{III}B^V$, в том числе AlGaAs/GaAs, AlGaSb/GaSb, GaInAs/InP и GaInAsP/InP. На начальной стадии развития таких фотодиодов был получен ряд обнадеживающих результатов по быстродействию и квантовой эффективности. Последние могут превосходить аналогичные харак-

теристики существующих лавинных Ge-фотодиодов, работающих в диапазоне 1—1,6 мкм. Вместе с тем в этом направлении предстоит еще широкие исследования с целью изучения таких параметров материалов, как темновые токи, коэффициенты поглощения, отношения коэффициентов ионизации, а также с целью совершенствования материалов, методов изготовления, конструкций приборов и их надежности.

Последние исследования отношения коэффициентов ионизации в $Al_xGa_{1-x}Sb$ -переходах показали, что, когда спин-орбитальное расщепление валентной зоны Δ приближается к значению, равному ширине запрещенной зоны (рис. 29, вставка), это отношение может быть очень малым [57]. График на рис. 29 демонстрирует резкое уменьшение k' при $\Delta/E_g \approx 1$. Были получены значения k' менее 0,04, что соответствует значениям шум-фактора менее 5 при $M = 100$.

Ряд лавинных фотодиодов был создан с использованием конструкции, изображенной на рис. 23, г. В качестве подложек использовались полупроводники III—V групп (например, GaAs и InP). Слои тройных или четверных соединений с близкими параметрами решеток выращивались на подложках, например, с помощью методов жидкостной, газофазной или молекулярно-лучевой эпитаксии. На рис. 30 (вставка) показана структура $Al_{0,85}Ga_{0,15}As/GaAs$ [48]. Верхний слой структуры играет роль «окна», прозрачного для падающего излучения в диапазоне 0,5—0,9 мкм. Квантовая эффективность при $\lambda = 0,53$ мкм равна $\sim 70\%$ и может быть увеличена до 95% при использовании просветляю-

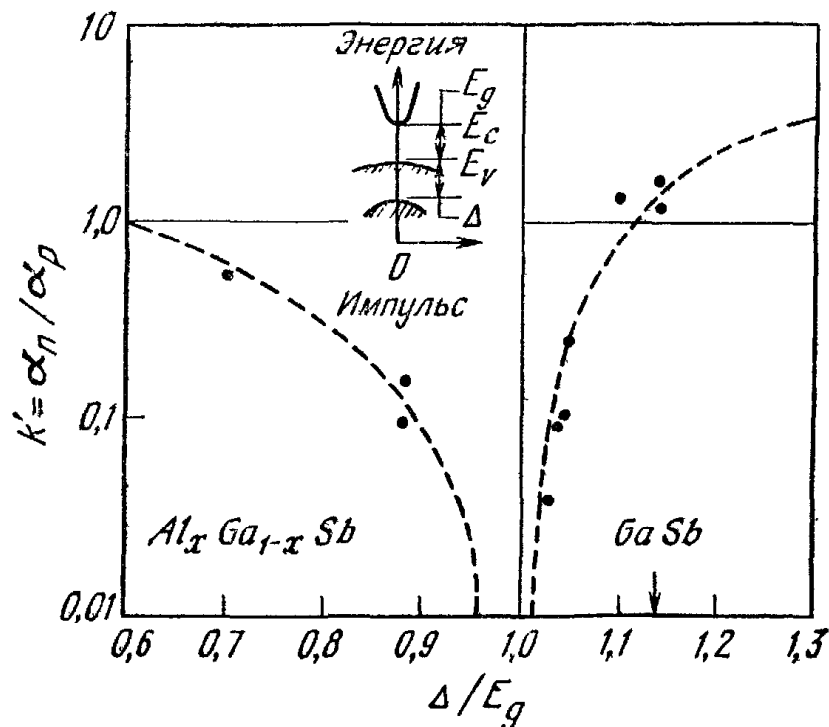


Рис. 29. Зависимость отношения коэффициентов ионизации в $Al_xGa_{1-x}Sb$ -диоде от Δ/E_g , где Δ — спин-орбитальное расщепление валентной зоны (вставка) [57].

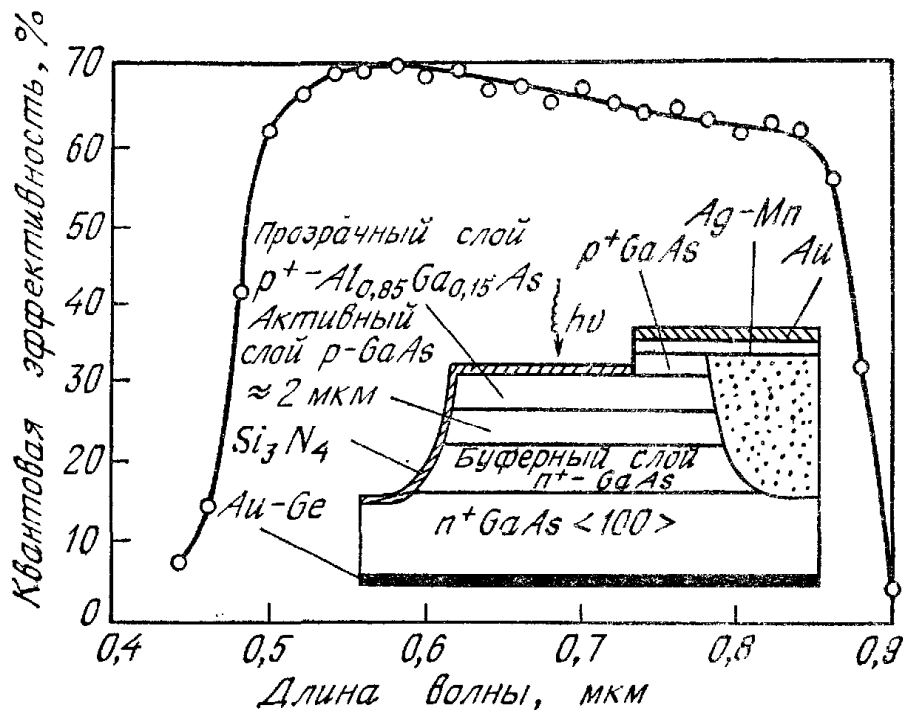


Рис. 30. Зависимость квантовой эффективности от длины волны излучения для лавинного фотодиода с гетеропереходом AlGaAs/GaAs. На вставке показана структура прибора (лавинного Al_{0,85}Ga_{0,15}As/GaAs-фотодиода) [48].

щего покрытия из Si₃N₄ (рис. 30) Темновой ток при напряжении, равном половине пробивного напряжения, не превышал $\sim 10^{-8}$ А/см². Время нарастания импульса составляет 35 пс, а коэффициент умножения на частоте 273 МГц равен 100. Отношение $k = \alpha_p/\alpha_n$ составило 0,83, что, по-видимому, является следствием сравнимых значений коэффициентов ионизации электронов и дырок в GaAs с ориентацией <100>. Необходимо отметить, что в GaAs с ориентацией <111> коэффициент ионизации дырок существенно больше, чем коэффициент ионизации электронов (см. гл. 1). Следовательно, для достижения минимального уровня шума необходимо использовать GaAs с ориентацией <111>, в котором процесс лавинного умножения стимулируется дырками.

На рис. 31 (вставка) представлена гетероструктура [49], в которой планарный p—n-переход сформирован в прозрачном слое InP, отделенном от поглощающего слоя GaInAsP с более узкой запрещенной зоной. В этом диоде обедненная область распространяется из n-InP в слой n-GaInAsP от плоской части перехода, оставаясь внутри InP на периферии перехода. Так как пробивное напряжение V_B пропорционально $E_g^{3/2}$, то его значение в InP примерно на 60 % выше, чем в GaInAsP (при условии, что оба слоя имеют одинаковые уровни легирования). В результате этого в рассматриваемой диодной структуре предотвращается краевой пробой. Уменьшение токов утечки достигается также за счет подавления туннельного пробоя в широкозонном материале (т. е. в InP), соседствующего со слоем с более узкой запрещенной

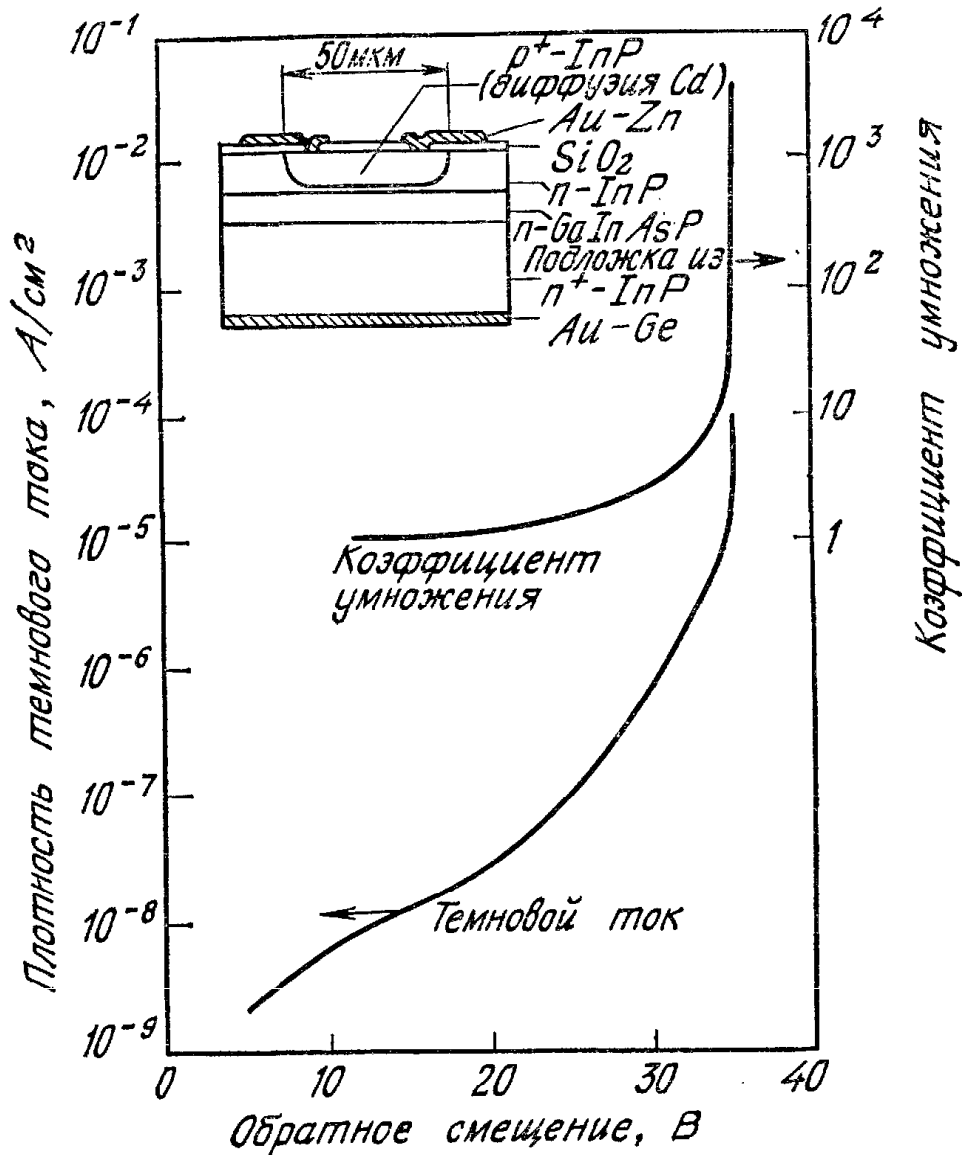


Рис. 31. Зависимость коэффициента умножения и темнового тока от величины обратного смещения для лавинного фотодиода с гетероструктурой GaInAsP [49].

зоной (т. е. GaInAsP). Прибор имеет низкую плотность темнового тока $\sim 10^{-6}$ А/см² при $\frac{1}{2} V_B$ и коэффициент умножения 3000.

На рис. 32, а показана зонная диаграмма лавинного фотодиода в состоянии равновесия, имеющего отдельные области поглощения и умножения [58]. Переход $p^+ - n$ сформирован в слое n -InP (область умножения). Слой GaInAs, выращенный на n -InP, используется в качестве области поглощения. Поскольку в InP коэффициент ионизации дырок в 2—3 раза больше, чем коэффициент ионизации электронов ($k' = 0,4$), то процесс лавинного умножения должен стимулироваться дырками. Для обеспечения инжекции дырок в область умножения через барьер на границе слоев n -InP и n -GaInAs необходимо подобрать уровни легирования и толщины этих слоев таким образом, чтобы обедненная область в n -InP в условиях лавинного пробоя достигала слоя GaInAs (рис. 32, б) Квантовая эффективность прибора составляет 40 %

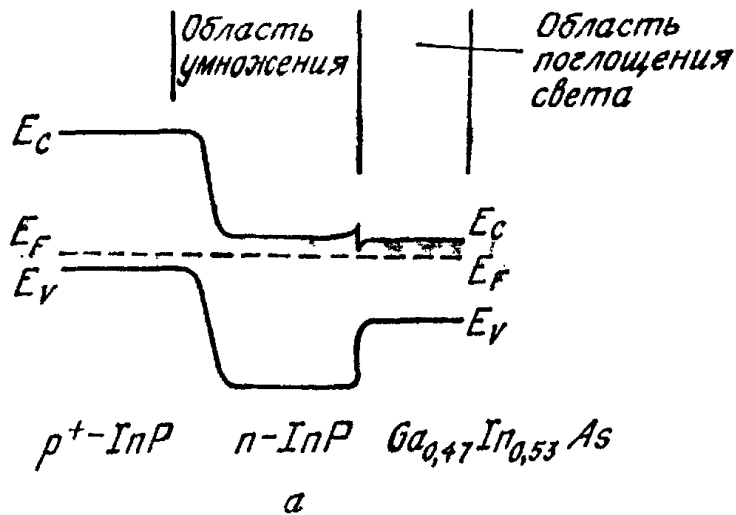
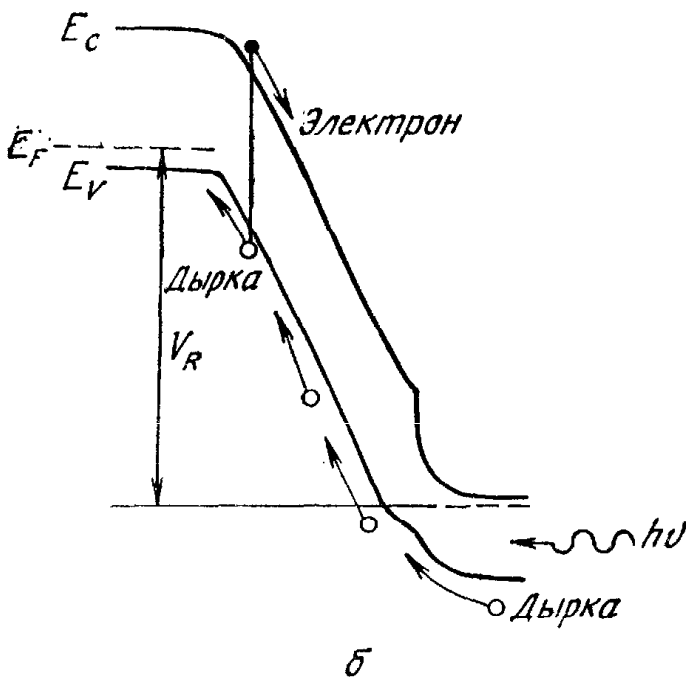


Рис. 32. Зонная диаграмма гетероперехода $\text{InP} - \text{GaInAs}$ в условиях теплового равновесия (а) и зонная диаграмма при лавинном пробое (б) [58].



при $\lambda = 1,3$ мкм и 50 % при $\lambda = 1,6$ мкм, а шум-фактор на 3 дБ ниже, чем в лавинном Ge-фотодиоде, работающем на длине волны 1,15 мкм.

13.5. ФОТОТРАНЗИСТОР

Все рассмотренные в предыдущих главах биполярные и униполярные транзисторы могут выполнять функции фотодетекторов и, следовательно, называться фототранзисторами. Присущий этим приборам транзисторный эффект обеспечивает высокое усиление. Вместе с тем по сравнению с фотодиодом фототранзистор более сложен в изготовлении и уступает ему в быстродействии вследствие большей площади.

Структура биполярного фототранзистора [51] и его эквивалентная схема приведены на рис. 33, а. Он отличается от обычного биполярного транзистора тем, что переход база — коллектор,

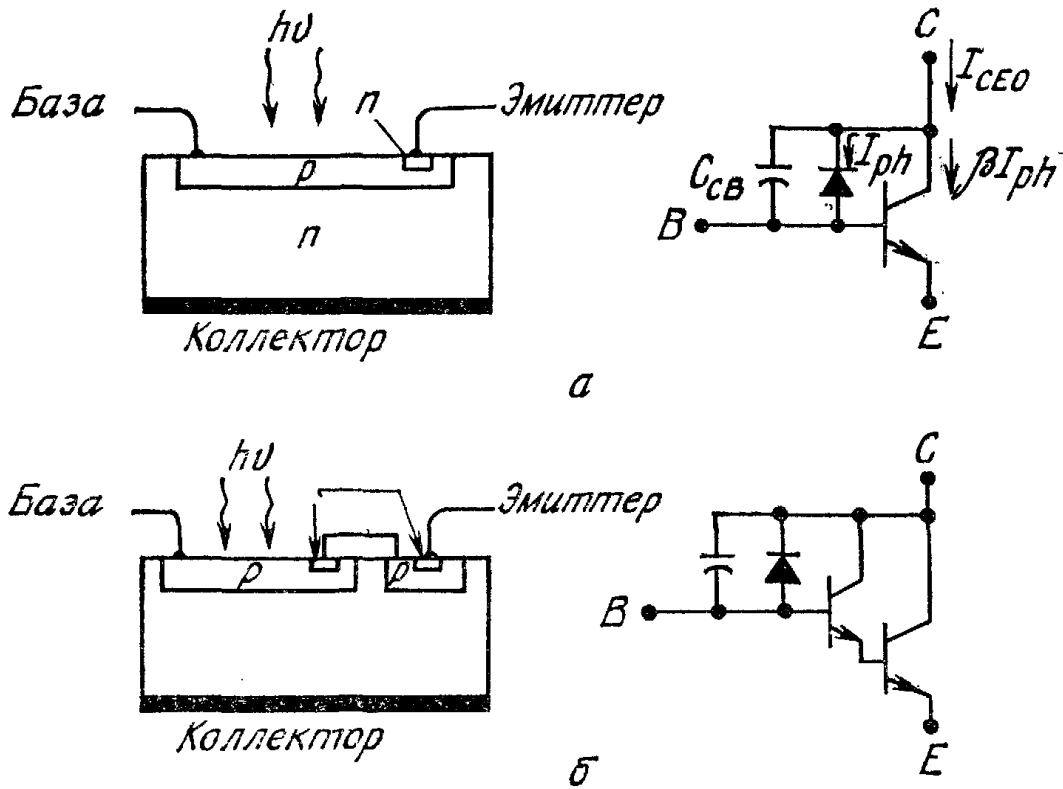


Рис. 33. Биполярный фототранзистор (а) и составной фототранзистор (б) [51].

выполняющий в нем роль чувствительного элемента (на рис. 33, а показан в виде диода с параллельно включенной емкостью), имеет большую площадь. Фототранзисторы особенно эффективны в составе оптронов, поскольку обеспечивают высокий коэффициент преобразования по току (т. е. отношение выходного тока фотодетектора к входному току светодиода), достигающий 50 % и более, тогда как при использовании фотодиодов этот коэффициент не превышает 0,2 %. В режиме работы с плавающей базой фотоносители дают вклад в ток коллектора в виде фототока I_{ph} . Кроме того, дырки, фотогенерированные в базе, а также дырки, приходящие в базу из коллектора, понижают разность потенциалов между собой и эмиттером, что приводит к инжекции электронов через базу в коллектор. Общий коллекторный ток равен

$$I_{CEO} = I_{ph} + h_{FE} I_{ph} = (1 + h_{FE}) I_{ph}, \quad (56)$$

где h_{FE} — статический коэффициент усиления транзистора в схеме с общим эмиттером, который может быть существенно больше единицы. Вследствие этого эффективный квантовый выход фототранзистора в $(1 + h_{FE})$ раз больше, чем квантовый выход фотодиода, которым является переход база — коллектор.

Мощность, эквивалентная шуму, определяется выражением (20), в котором [52]

$$I_{eq} = I_{CEO} (1 + 2h_{je}^2 / h_{FE}), \quad (57)$$

где h_{fe} — дифференциальный коэффициент усиления транзистора в схеме с общим эмиттером. Таким образом, уровень шума прибора и усиление взаимосвязаны.

Биполярный фототранзистор может быть интегрально совмещен с другими приборами. Например, используя дополнительный транзистор, можно сформировать составной фототранзистор с существенно более высоким коэффициентом усиления (рис. 33, б). Быстродействие рассмотренных структур ограничивается большой емкостью перехода база — коллектор и снижается при увеличении усиления за счет эффекта обратной связи. Так, типичное время отклика для фотодиода составляет $\sim 0,01$ мкс, тогда как для биполярного фототранзистора оно равно ~ 5 мкс, а для составного фототранзистора ~ 50 мкс.

Фототранзистор может быть изготовлен на основе гетероструктуры. На рис. 34 (вставка) показана зонная энергетическая диаграмма симметричной структуры $n\text{-AlGaAs}/p\text{-GaAs}/n\text{-AlGaAs}$ [53]. Этот прибор характеризуется высокими значениями напряжения запирающего и коэффициента усиления для обеих полярностей смещения и линейной зависимостью тока от напряжения, проходящей через начало координат (рис. 34). Эффективная инжекционная способность обеспечивается за счет того, что слой базы имеет более узкую запрещенную зону, чем слои коллектора и эмиттера. Таким путем устраняется присущее обычным транзисторам ограничение, состоящее в необходимости создания слаболегированного базового слоя, и можно реализовать условия, при которых обе-

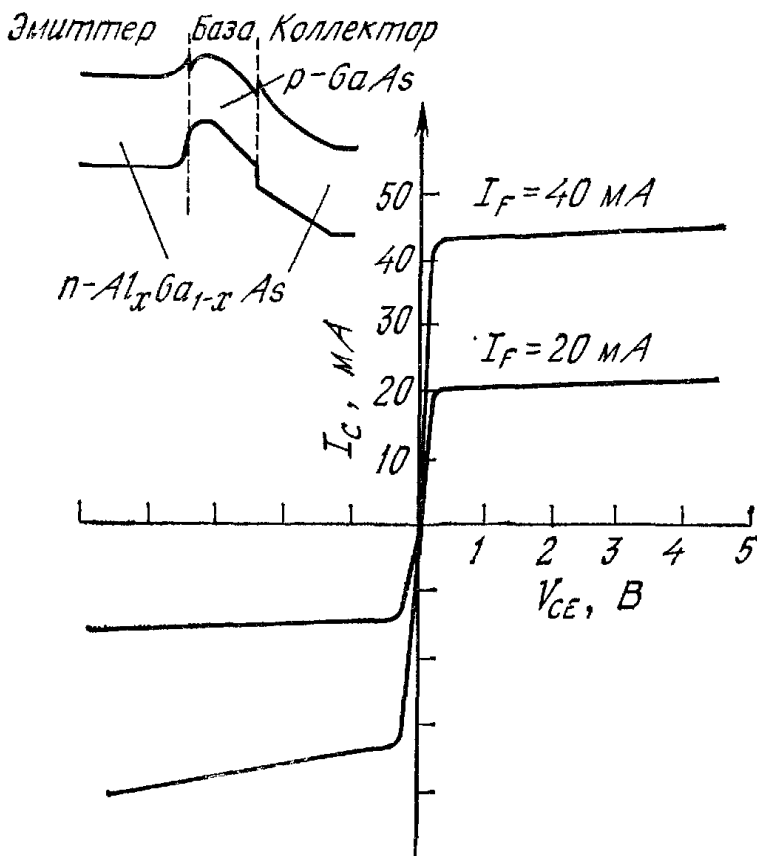
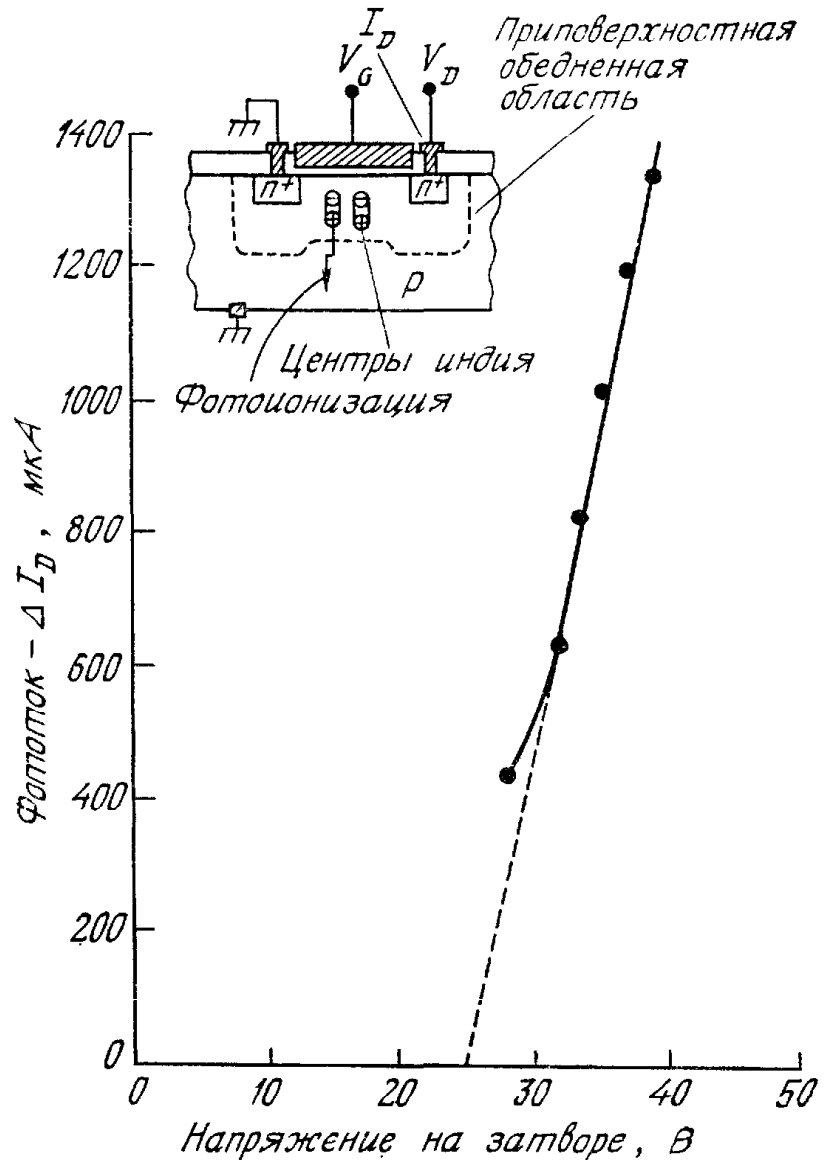


Рис. 34. Вольт-амперная характеристика фототранзистора с двусторонним гетеропереходом [53].

Рис. 35. Зависимость фототока в области насыщения от напряжения на полевом электроде в МОП-транзисторе с подложкой, легированной индием [54].



ненная область запирающего перехода в основном сосредоточена в слое коллектора. Это позволяет получить в одном приборе высокое усиление и высокое напряжение запирающего. Прибор обладает симметричным коэффициентом усиления, который при напряжении запирающего ± 10 В, соответствующем толщине базы GaAs 2,1 мкм, составляет 180, а при напряжении запирающего $\pm 2,6$ В, соответствующем базе толщиной 0,3 мкм, равен 3000.

Структура кремниевого МДП-транзистора с изолированным затвором, чувствительного к ИК-излучению [54], схематически представлена на рис. 35 (вставка). (Характеристики МДП-транзисторов с изолированным затвором рассмотрены в гл. 8.) Подложка p -типа легирована бором, дающим мелкий акцепторный уровень, и индием, дающим глубокий акцепторный уровень (глубина залегания 0,16 эВ от края валентной зоны). В заполненном дырками состоянии центры индия нейтральны, а при возбуждении дырок в валентную зону становятся отрицательно заряженными. Прибор работает следующим образом. Отрицательное напряжение на полевом электроде вызывает обогащение канала; при этом все

центры индия захватывают дырки и переходят в нейтральное состояние. Под действием инфракрасного излучения в диапазоне 2—7 мкм, обеспечивающего энергию фотона, достаточную для фотоионизации центров индия, происходит возбуждение дырок в валентную зону. Этот процесс приводит к увеличению суммарной плотности заряженных акцепторов в области обеднения, что в свою очередь модулирует проводимость поверхностного канала. Выражение для изменения порогового напряжения имеет вид

$$\Delta V_T = 2 \sqrt{q \epsilon_s \psi_B} (\sqrt{N_A + N_{In}} - \sqrt{N_A}) / C_i, \quad (58)$$

где ψ_B — разность потенциалов между уровнем Ферми и уровнем Ферми собственного полупроводника, C_i — удельная емкость окисла, N_A — концентрация бора и N_{In} — концентрация фотоионизованных центров индия. Ток канала в области насыщения пропорционален $(V_G - V_T)^2$, и, таким образом, изменение тока канала может быть представлено в виде

$$\Delta I_D \sim \Delta V_T (V_G - V_T - \Delta V_T / 2). \quad (59)$$

На рис. 35 приведена зависимость фототока в области насыщения от напряжения на полевом электроде. Полученная зависимость хорошо согласуется с уравнением (59).

Используя вместо индия другие акцепторные примеси, можно обеспечить оптимальное согласование области чувствительности прибора с требуемым спектральным диапазоном. Рассмотренный ИК-детектор может быть интегрально совмещен с элементами накопления и памяти и выходным усилителем, что открывает широкие перспективы для его использования в многоэлементных ИК-фотоприемниках большой степени интеграции.

ЛИТЕРАТУРА

1. Anderson L. K., Didomenico M., Jr., Fisher M. B. High-Speed Photodetector for Microwave Demodulation of Light, in Young L., Ed., *Advances in Microwaves*, Vol. 5, Academic, N. Y., 1970, pp. 1—122.
2. Melchior H. Demodulation and Photodetection Techniques, in Arecchi F. T., Schulz-Dubois E. O., Eds., *Laser Handbook*, Vol. 1, North-Holland, Amsterdam, 1972, pp. 725—835.
3. Melchior H. Detector for Lightwave Communication, *Phys. Today*, p. 32 (Nov. 1977).
4. Willardson R. K., Bear A. C., Eds. *Semiconductors and Semimetals*, Vol. 12, Infrared Detector II, Academic, N. Y., 1977.
5. Stillman G. E., Wolfe C. M. Avalanche Photodiode, in Ref. 4, pp. 291—394.
6. Lee T. P., Li T. Y. Photodetectors, in Miller S. E., Chynoweth A. G., Eds. *Optical Fiber Communications*, Academic, N. Y., 1979, Chap. 18.
7. Bratt P. R. Impurity Ge and Si Infrared Detectors, in Ref. 4, pp. 39—142.
8. Bube R. H. Comparison of Solid State Photoelectronic Radiation Detectors, *Trans. AIME*, 239, 291 (1967).
9. DiDomenico M., Jr., Svelto O. Solid State Photodetection Comparison between Photodiodes and Photoconductors, *Proc. IEEE*, 52, 136 (1964).

10. Van der Ziel A. Fluctuation Phenomena in Semiconductors, Academic, N. Y., 1959, Chap. 9. [Имеется русский перевод: Ван-дер-Зил А., Флуктуационные явления в полупроводниках, М., 1961.]
11. Ross M. Laser Receivers-Devices, Techniques, Systems, Wiley, N. Y., 1966. [Имеется русский перевод: Росс М. Лазерные приемники, М.: Мир, 1969.]
12. Kruse P. W., McGlauchlin L. D., McQuistan R. B. Elements of Infrared Technology, Wiley, N. Y., 1962.
13. Eisenman W. L., Merriam J. D., Potter R. F. Operational Characteristics of Infrared Photodiode, in Ref. 4, pp. 1—38.
14. Stillman G. E., Wolfe C. M., Dimmock J. O. Far-Infrared Photoconductivity in High Purity GaAs, in Ref. 4, pp. 169—290.
15. Gartner W. W. Depletion-Layer Photoeffects in Semiconductors, *Phys. Rev.*, **116**, 84 (1959).
16. Lee H. S., Sze S. M. Silicon $p-i-n$ Photodetector Using Internal Reflection Method, *IEEE Trans. Electron Devices*, **ED-17**, 342 (1970).
17. Muller J. Thin Silicon Film $p-i-n$ Photodiodes with Internal Reflection, *IEEE Trans. Electron Devices*, **ED-25**, 247 (1978).
18. Ahlstrom E., Gartner W. W. Silicon Surface-Barrier Photocells, *J. Appl. Phys.*, **33**, 2602 (1962).
19. Schneider M. V. Schottky Barrier Photodiodes with Antireflection Coating, *Bell Syst. Tech. J.*, **45**, 1611 (1966).
20. Crowell C. R., Sze S. M. Hot Electron Transport and Electron Tunneling in Thin Film Structures, in Thun R. E., Ed., Physics of Thin Films, Vol. 4, Academic, N. Y., 1967, pp. 325—371.
21. Гуткин А. А., Дмитриев В. М., Хайт В. М., *ФТП*, **11** (1977).
22. Lavagna M., Pique J. P., Marfaing Y. Theoretical Analysis of Quantum Yield in Schottky Diodes, *Solid State Electron.*, **20**, 235 (1977).
23. Sharpless W. M. Cartridge-Type Point Contact Photodiode, *Proc. IEEE*, **52**, 207 (1964).
24. Casey H. C., Jr., Panish M. B. Heterostructure Lasers, Part B, Academic, N. Y., 1978, p. 3. [Имеется русский перевод: Кейси Х., Паниш М. Лазеры на гетероструктурах, М.: Мир, 1981.]
25. Miller R. C., Schwartz B., Koszi L. A., Wagner W. R. A High-Efficiency GaAlAs Double-Heterostructure Photovoltaic Detector, *Appl. Phys. Lett.*, **33**, 721 (1978).
26. Ilegems M., Schwartz B., Koszi L. A., Miller R. C. Integrated Multijunction GaAs Photodetector with High Output Voltage, *Appl. Phys. Lett.*, **33**, 629 (1978).
27. Burrus C. A., Dentai A. G., Lee T. P. InGaAsP $p-i-n$ Photodiodes with Low Dark Current and Small Capacitance, *Electron. Lett.*, **15**, 655 (1979).
28. Lee T. P., Burrus C. A., Dentai A. G. InGaAs/InP $p-i-n$ Photodiodes for Lightwave Communications at 0,95 to 1,65 μm Wavelengths, *IEEE J. Quantum Electron.*, **QE-17**, 232 (1981).
29. Ahmad K., Mabbitt A. W. GaInAs Photodiodes, *Solid State Electron.*, **22**, 327 (1979).
30. Johnson K. M. High-Speed Photodiode Signal Enhancement at Avalanche Breakdown Voltage, *IEEE Trans. Electron Devices*, **ED-12**, 55 (1965).
31. Anderson L. K., McMurtry B. J. High Speed Photodetectors, *Proc. IEEE*, **54**, 1353 (1966).
32. Emmons R. B. Avalanche Photodiode Frequency Response, *J. Appl. Phys.*, **38**, 3705 (1967).
33. Melchior H., Lynch W. T. Signal and Noise Response of High Speed Germanium Avalanche Photodiodes, *IEEE Trans. Electron Devices*, **ED-13**, 829 (1966).
34. McIntyre R. J. Multiplication Noise in Uniform Avalanche Diodes, *IEEE Trans. Electron Devices*, **ED-13**, 164 (1966).
35. Baertsch R. D. Noise and Ionization Rate Measurements in Silicon Photodiodes, *IEEE Trans. Electron Devices*, **ED-13**, 987 (1966).

36. McIntyre R. J. The Distribution of Gains in Uniformly Multiplying Avalanche Photodiodes: Theory, *IEEE Trans. Electron Devices*, **ED-19**, 703 (1972).
37. Webb R. P., McIntyre R. J., Conradi J. Properties of Avalanche Photodiodes, *RCA Rec.*, **35**, 234 (1974).
38. Anderson L. K., McMullin P. G., D'Asaro L. A., Goetzberger A., Microwave Photodiodes Exhibiting Microplasma — Free Carrier Multiplication, *Appl. Phys. Lett.*, **6**, 62 (1965).
39. Sze S. M., Gibbons G., Effect of Junction Curvature of Breakdown Voltage in Semiconductors, *Solid State Electron.*, **9**, 831 (1966).
40. Ando H., Kanbe H., Kimura T., Yamaoka T., Kaneda T. Characteristics of Ge Avalanche Photodiodes in the Wavelength Region of 1—1,6 μm , *IEEE J. Quantum Electron.*, **QE-14**, 804 (1978).
41. Ruegg H. W. An Optimized Avalanche Photodiode, *IEEE Trans. Electron Devices*, **ED-14**, 239 (1967).
42. Moll J. Physics of Semiconductors, McGraw-Hill, N. Y., 1964.
43. Kanbe H., Kimura T. Figure of Merit for Avalanche Photodiodes, *Electron. Lett.*, **13**, 262 (1977).
44. Melchior H., Hartman A. R., Schinke A. R., Seidel T. E. Planar Epitaxial Silicon Avalanche Photodiode, *Bell Syst. Tech. J.*, **57**, 1791 (1978).
45. Melchior H., Lepselter M. P., Sze S. M. Metal — Semiconductor Avalanche Photodiode, IEEE Solid-State Device Res. Conf., Boulder, Colo, June 17—19, 1968.
46. Lepselter M. P., Sze S. M. Silicon Schottky Barrier Diode with Near-Ideal I—V Characteristics, *Bell Syst. Tech. J.*, **47**, 195 (1968).
47. Hurwitz C. E., Hsieh J. J. GaInAsP/InP Avalanche Photodiodes, *Appl. Phys. Lett.*, **32**, 487 (1978).
48. Law H. D., Nakano K., Tomasetta L. R. III—V Alloy Heterostructure High Speed Avalanche Photodiodes, *IEEE, J. Quantum Electron.*, **QE-15**, 549 (1979).
49. Nishida K., Taguchi K., Matsumoto Y. InGaAsP Heterostructure Avalanche Photodiodes with High Avalanche Gain, *Appl. Phys. Lett.*, **35**, 251 (1979).
50. Shockley W., Sparks M., Teal G. K. $p - n$ Junction Transistor, *Phys. Rev.*, **83**, 151 (1951).
51. Jayson J. S., Knight S. Opto-Isolator, in Wolfe R., Ed., Applied Solid State Science, Vol. 6, Academic, N. Y., 1976, pp. 119—168.
52. DeLaMoneda F. H., Chenette E. R., Van Der Ziel A. Noise in Phototransistors, *IEEE Trans. Electron Devices*, **ED-18**, 340 (1971).
53. Knight S., Dawson L. R., Keramidis U. G., Spencer M. G. An Optically Triggered Double Heterostructure Linear Bilateral Phototransistors, IEEE Tech. Dig., Int. Electron Device Meet., 1977, p. 472.
54. Forbes L., Wittmer L. L., Loh K. W. Characteristic of the Indium-Doped Infrared Sensing MOSFET, *IEEE Trans. Electron Devices*, **ED-23**, 1272 (1976).
55. Smith R. G., Personick S. D. Receiver Design for Optical Communications Systems, Kressel H., Ed., Semiconductor Devices for Optical Communication, Springer-Verlag, N. Y., 1979, Chap. 4.
56. Campbell J. C., Dentai A. G., Lee T. P., Burrus C. A. Improved Two-Wavelength Demultiplexing InGaAsP Photodetector, *IEEE J. Quantum Electron.*, **QE-16**, 601 (1980).
57. Hildebrand O., Kuebart W., Pikuhn M. H. Resonant Enhancement of Impact Ionization in $\text{Al}_x\text{Ga}_{1-x}\text{Sb}$ $i - i - n$ Avalanche Photodiodes, IEEE Device Res. Conf., June 23—25, 1980.
58. Susa N., Nakagome H., Mikami O., Ando H., Kanbe H. New InGaAs/InP Avalanche Photodiode Structure for the 1—1,6 μm Wavelength Region *IEEE J. Quantum Electron.*, **QE-16**, 864 (1980).
59. Forrest S. R., Leheny R. F., Nahory R. E., Pollack M. A. $\text{In}_{0,53}\text{Ga}_{0,47}\text{As}$ Photodiodes with Dark Current Limited by Generation-Recombination and Tunneling, *Appl. Phys. Lett.*, **37**, 322 (1980).

14.1. ВВЕДЕНИЕ

В настоящее время солнечные батареи являются наиболее важным источником энергии для спутников и космических кораблей при длительных полетах. Они успешно применяются также и на Земле. В связи с увеличением потребления энергии во всем мире запасы обычных источников энергии (различных видов ископаемого топлива) должны истощиться в не слишком отдаленном будущем. Поэтому необходимо разрабатывать и использовать альтернативные источники энергии, и особенно наш единственный практически неисчерпаемый естественный источник энергии — Солнце. Считается, что основным поставщиком солнечной энергии будут солнечные батареи, поскольку они преобразуют солнечный свет непосредственно в электричество с высоким коэффициентом преобразования, создают почти постоянную мощность при низких эксплуатационных расходах и фактически не загрязняют окружающую среду. В последнее время наблюдается расширение исследований и разработок дешевых плоскочелюстных, а также тонкопленочных солнечных батарей, систем концентраторов и многих новых идей. Следует ожидать, что в ближайшем будущем стоимость отдельного солнечного элемента и комплектуемых на его основе больших солнечных батарей снизится настолько, что окажется экономически выгодным использование солнечной энергии в больших масштабах.

Первый солнечный элемент был создан Чапеном, Фуллером и Пирсоном в 1954 г. на основе диффузионного кремниевого $p-n$ -перехода [1]. Впоследствии Рейнольдс и др. разработали солнечный элемент на сульфиде кадмия [2]. Затем солнечные элементы были созданы на многих других полупроводниках с использованием различных конструкций прибора и применением монокристаллических и поликристаллических материалов и аморфных тонкопленочных структур. Исчерпывающий анализ основных характеристик солнечных элементов проведен Ховелом [3]. Обзор основных важнейших работ по солнечным батареям за период до 1974 г. составлен Бекусом [4]. Палфри [5] и Джонстон [6] составили обзоры по фотовольтаическим элементам, а в работе Бечмана [7] обсуждаются проблемы выбора материала для солнечных батарей. Результаты последних исследований можно найти в сборниках конференций для специалистов по фотовольтаическим эффектам [8, 9].

14.2. СОЛНЕЧНОЕ ИЗЛУЧЕНИЕ И ИДЕАЛЬНАЯ ЭФФЕКТИВНОСТЬ ПРЕОБРАЗОВАНИЯ

14.2.1. Солнечное излучение

Источником энергии солнечного излучения служит термоядерная реакция. Каждую секунду примерно $6 \cdot 10^{11}$ кг H_2 превращаются в He. Дефект массы при этом составляет $4 \cdot 10^3$ кг, что приводит в соответствии с соотношением Эйнштейна $E = mc^2$ к выделению энергии, равной $4 \cdot 10^{20}$ Дж. Основная часть этой энергии испускается в виде электромагнитного излучения в диапазоне от ультрафиолетового до инфракрасного (0,2—3 мкм). Полная масса Солнца в настоящее время составляет $\sim 2 \cdot 10^{30}$ кг, что должно обеспечивать его достаточно стабильное существование примерно с постоянным выделением энергии в течение свыше 10 млрд. (10^{10}) лет.

Интенсивность солнечного излучения в свободном пространстве на расстоянии, равном среднему расстоянию между Землей и Солнцем, называется солнечной постоянной. Ее величина равна 1353 Вт/м^2 [4, 10]. При прохождении через атмосферу солнечный свет ослабляется в основном благодаря поглощению инфракрасного излучения парами воды, поглощению ультрафиолетового излучения озоном и рассеянию излучения находящимися в воздухе частицами пыли и аэрозолями. Показатель атмосферного влияния на интенсивность солнечного излучения, доходящего до земной поверхности, определяется «воздушной массой» (AM). Величина последней равна секансу угла между Солнцем и зенитом ($\sec \theta$).

На рис. 1 приведены четыре кривые, иллюстрирующие спектральное распределение интенсивности солнечного излучения [10] (мощность на единицу площади в единичном интервале длин волн). Верхняя кривая соответствует солнечному спектру за пределами земной атмосферы, т. е. при нулевой воздушной массе (AM0). Это распределение можно аппроксимировать распределением интенсивности черного тела при температуре 5800 К. Спектр AM0 определяет работу солнечных батарей на спутниках и космических кораблях. Спектр AM1 соответствует распределению интенсивности солнечного излучения на поверхности Земли, когда Солнце стоит в зените; при этом полная мощность излучения составляет $\sim 925 \text{ Вт/м}^2$. Спектр AM2 реализуется при угле $\theta = 60^\circ$. В этом случае полная мощность излучения равна 691 Вт/м^2 .

Средняя интенсивность излучения на Земле примерно совпадает с интенсивностью излучения, прошедшего через воздушную массу, равную 1,5, что соответствует положению Солнца под углом 45° к горизонту. На рис. 2 приведено распределение числа фотонов, приходящихся на единичный энергетический интервал

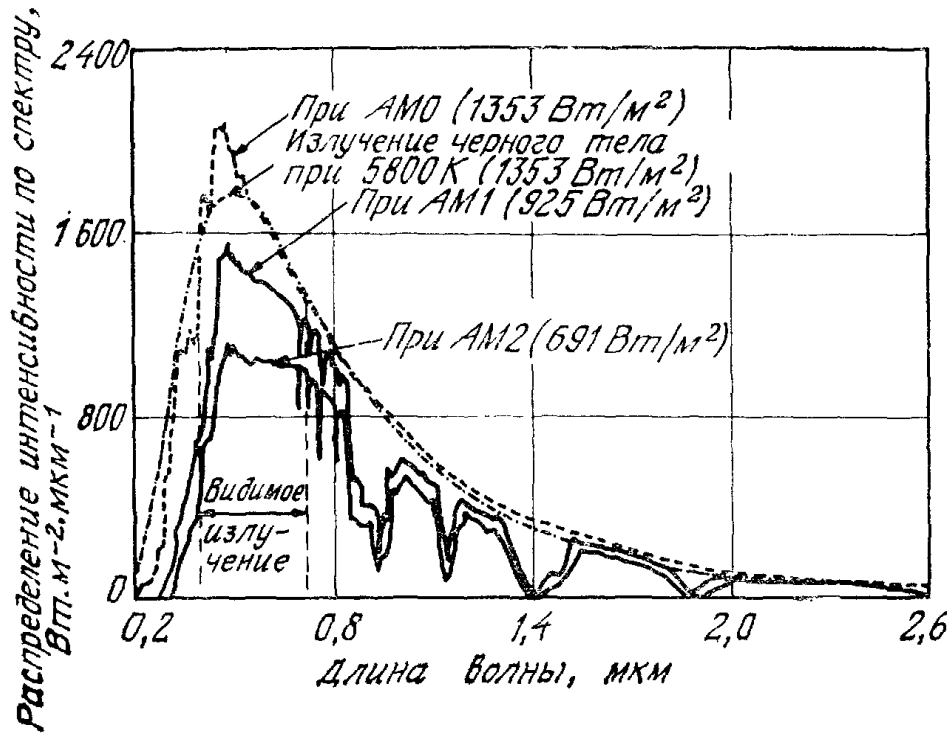


Рис. 1. Четыре кривые, иллюстрирующие распределение солнечной энергии по спектру [10].

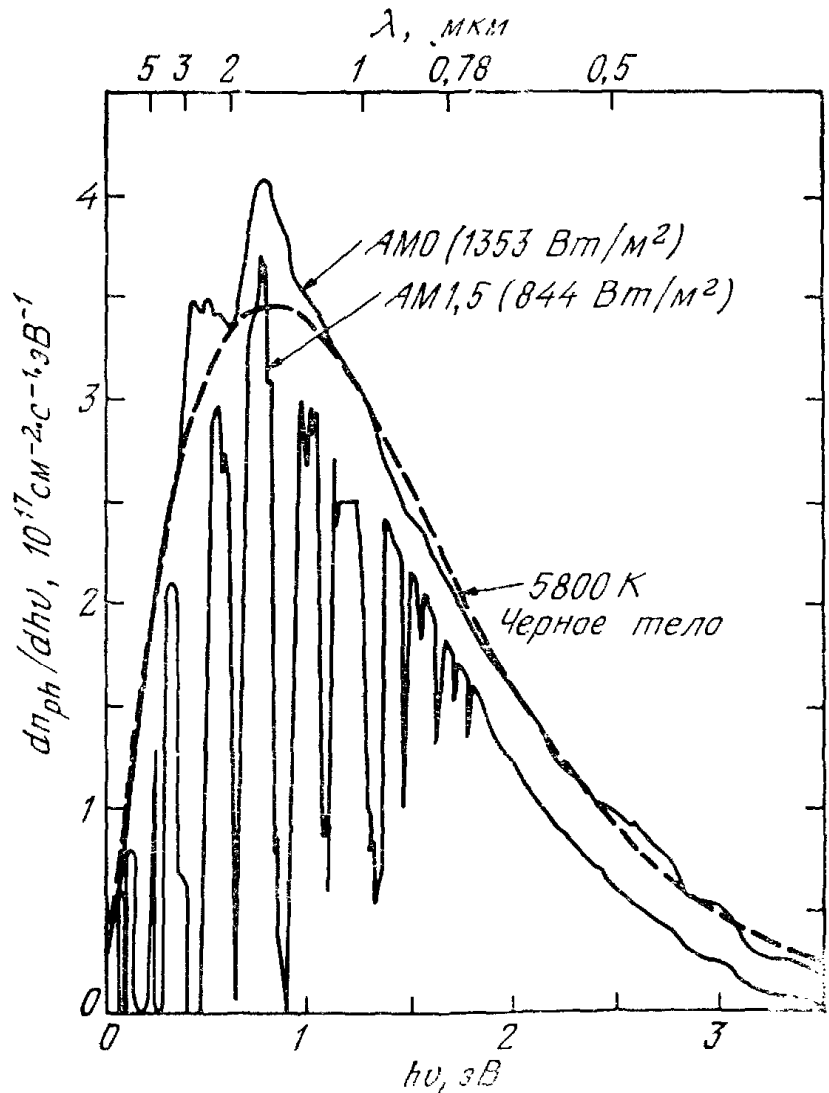


Рис. 2. Спектральное распределение потока фотонов, поступающих от Солнца в условиях AM0 и AM1,5 [11].

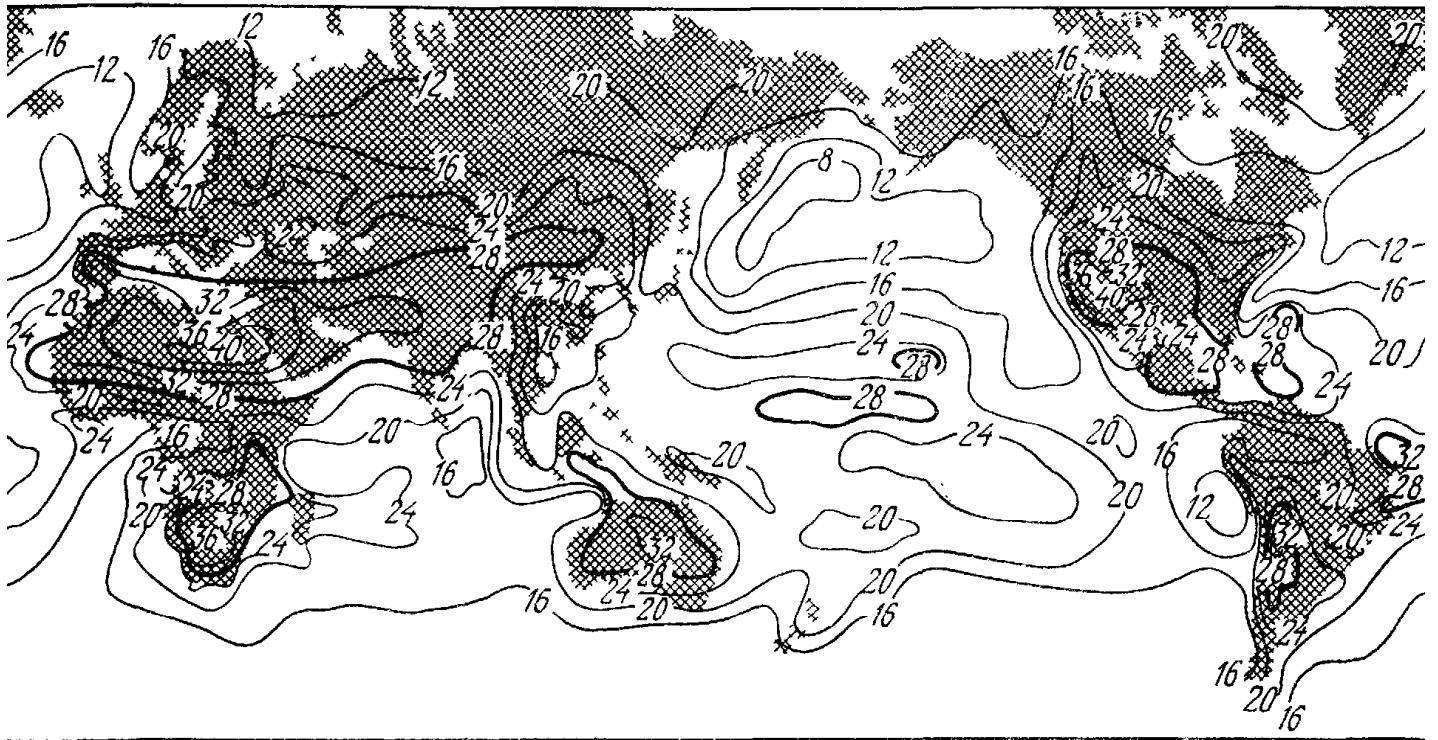


Рис. 3. Карта распределения солнечной энергии по земной поверхности. Контурные линии проходят через области равной облученности. Цифрами обозначено годовое количество в сотнях световых часов [5].

на 1 см^2 за 1 с в условиях АМ0 и АМ1,5 [11]. Для перевода длины волны (мкм) в энергию фотона (эВ) мы использовали соотношение

$$\lambda = \frac{c}{\nu} = \frac{1,24}{h\nu} . \quad (1)$$

Полная мощность солнечного излучения при АМ1,5 составляет 844 Вт/м^2 .

Для наиболее эффективного использования солнечных элементов необходимо знать примерное количество солнечной энергии, поступающей в течение года в различных регионах. На рис. 3 [5] приведена карта распределения солнечной энергии по земной поверхности с указанием средней продолжительности солнечного освещения. Жирными контурными линиями очерчены регионы, которые освещаются в среднем по 8 ч и более ежедневно. Поскольку на всех континентах имеются довольно большие площади с высоким средним уровнем солнечного освещения, можно ожидать, что в ближайшем будущем солнечная энергия будет широко использоваться во всем мире.

14.2.2. Идеальная эффективность преобразования

Обычный солнечный элемент (например, $p-n$ -переход) имеет лишь одну характерную энергию — ширину запрещенной зоны E_g . Когда на элемент попадает солнечный свет, фотоны с энергией,

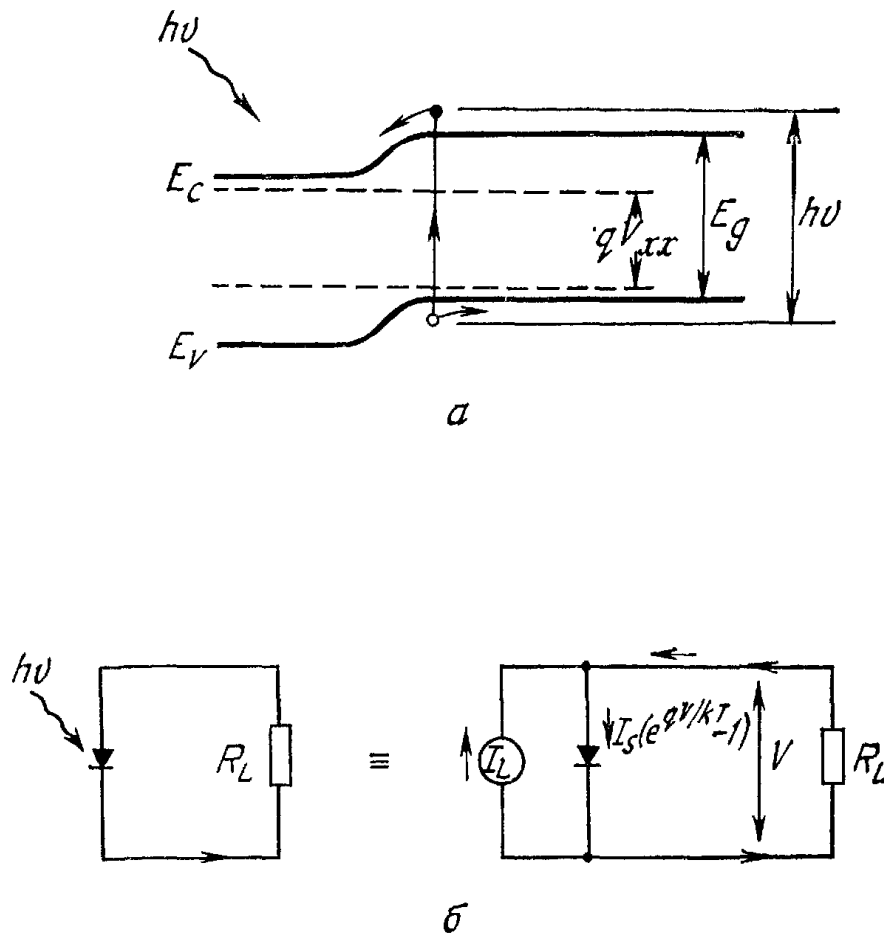


Рис. 4. Энергетическая диаграмма солнечного элемента с $p-n$ -переходом при солнечном освещении (а) и идеализированная эквивалентная схема солнечного элемента (б).

меньшей E_g , не дают вклада в выходную мощность элемента (в пренебрежении поглощением света, сопровождаемым возбуждением фононов). Каждый фотон с энергией, большей E_g , дает вклад, равный E_g , в выходную мощность, а оставшаяся часть энергии фотона переходит в тепло. Для того чтобы определить эффективность (или идеальный к. п. д.) преобразования, рассмотрим диаграмму энергетических зон освещаемого $p-n$ -перехода (рис. 4,а). Будем полагать, что солнечный элемент имеет идеальную вольт-амперную характеристику. Соответствующая эквивалентная цепь показана на рис. 4, б, где параллельно переходу введен источник постоянного тока I_L , описывающий возбуждение неравновесных носителей солнечным излучением. Ток насыщения диода I_s определен в гл. 2, R_L — нагрузочное сопротивление.

Вольт-амперная характеристика такого прибора определяется выражениями

$$I = I_s (e^{qV/kT} - 1) - I_L \quad (2)$$

и

$$J_s = I_s/A = qN_c N_v \left(\frac{1}{N_A} \sqrt{\frac{D_n}{\tau_n}} + \frac{1}{N_D} \sqrt{\frac{D_p}{\tau_p}} \right) e^{-E_g/kT}, \quad (3)$$

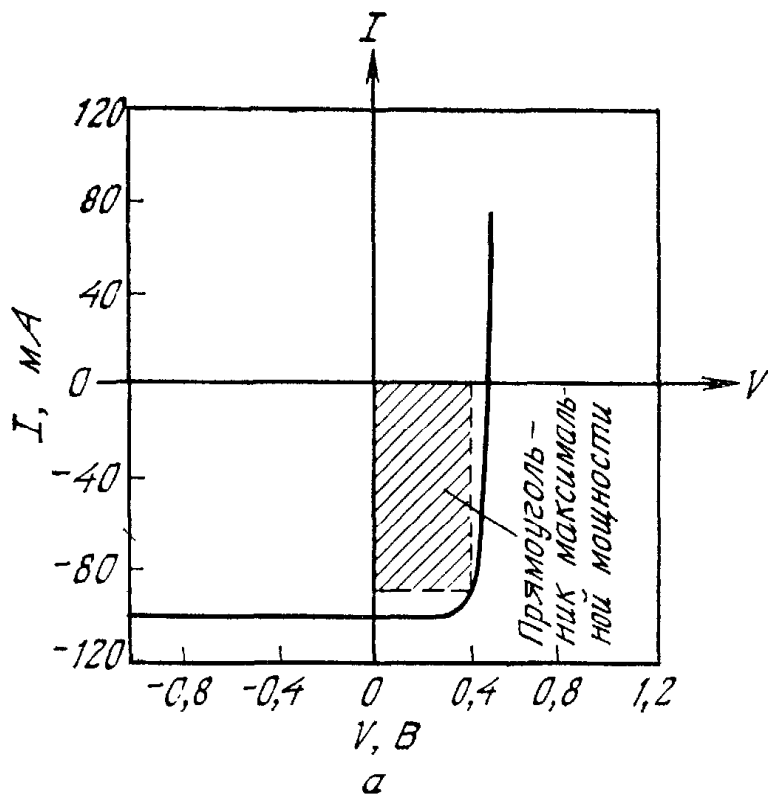
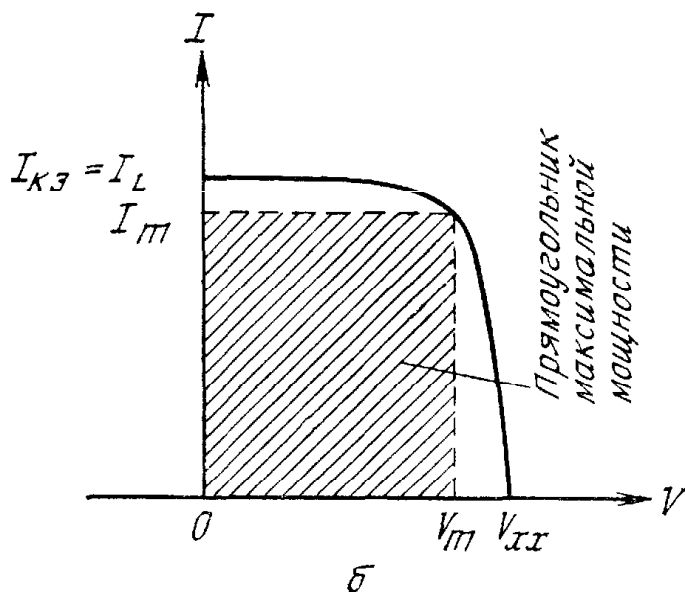


Рис. 5. Вольт-амперная характеристика освещенного солнечного элемента [12].



где A — площадь прибора. График вольт-амперной характеристики, определяемой формулой (2), приведен на рис. 5, а для следующих параметров: $I_L = 100$ мА, $I_s = 1$ нА, $A = 4$ см² и $T = 300$ К [12]. Поскольку вольт-амперная характеристика проходит через четвертый квадрант, это означает, что прибор служит источником энергии. При соответствующем подборе нагрузочного сопротивления вырабатываемая энергия может достигать 80 % произведения $I_{кз} V_{хх}$ ($I_{кз}$ — ток короткого замыкания, $V_{хх}$ — напряжение холостого хода элементов). Чаще вольт-амперную характеристику элемента изображают так, как показано на рис. 5, б. На рис. 5, б также определены величины I_m и V_m — значения тока и напряжения, при которых реализуется максимальная выходная мощность P_m ($P_m = I_m V_m$).

Из уравнения (2) получаем, что напряжение холостого хода (при $I = 0$) равно

$$V_{\text{xx}} = \frac{kT}{q} \ln \left(\frac{I_L}{I_s} + 1 \right) \approx \frac{kT}{q} \ln \left(\frac{I_L}{I_s} \right). \quad (4)$$

Следовательно, при заданном токе I_L напряжение холостого хода логарифмически возрастает при уменьшении тока насыщения I_s . Выходная мощность равна

$$P = IV = I_s V (e^{qV/kT} - 1) - I_L V. \quad (5)$$

Условие максимума мощности получаем, положив $dP/dV = 0$, откуда имеем

$$I_m = I_s \beta V_m e^{\beta V_m} \simeq I_L \left(1 - \frac{1}{\beta V_m} \right), \quad (6)$$

$$V_m = \frac{1}{\beta} \ln \left(\frac{I_L/I_s + 1}{1 + \beta V_m} \right) \simeq V_{\text{xx}} - \frac{1}{\beta} \ln(1 + \beta V_m), \quad (7)$$

где $\beta = q/kT$. Максимальная выходная мощность определяется выражением

$$P_m = I_m V_m \simeq I_L \left[V_{\text{xx}} - \frac{1}{\beta} \ln(1 + \beta V_m) - \frac{1}{\beta} \right] = I_L (E_m/q), \quad (8)$$

где

$$E_m \equiv q \left[V_{\text{xx}} - \frac{1}{\beta} \ln(1 + \beta V_m) - \frac{1}{\beta} \right]. \quad (9)$$

Величина E_m соответствует той максимальной энергии, которая выделяется в нагрузке при поглощении одного фотона и при оптимальном согласовании элемента с внешней цепью.

Для данного полупроводника плотность тока насыщения J_s может быть получена из формулы (3). Минимальное значение J_s для Si при 300 К составляет $\sim 10^{-15}$ А/см². При AM1,5 плотность тока короткого замыкания $J_{\text{кз}}$, равную J_L , можно получить графически из рис. 2:

$$J_{\text{кз}} = J_L(E_g) = q \int_{h\nu=E_g}^{\infty} (dn_{ph'} d h\nu) d(h\nu). \quad (10)$$

Результат такого интегрирования показан на рис. 6 (кривая 1) [11]. Если значения J_s и J_L известны, величину E_m можно получить с помощью численного решения трансцендентных уравнений (4), (7) и (9). Поскольку величина E_m зависит от J_s , она зависит также от параметров материала (таких, как τ , D и уровни легирования). Идеальная эффективность преобразования реализуется при оптимальном выборе параметров материала, когда величина J_s минимальна. Поведение E_m для полупроводников с различной шириной запрещенной зоны показано на рис. 6 (кривая 2). Идеальная эффективность преобразования равна от-

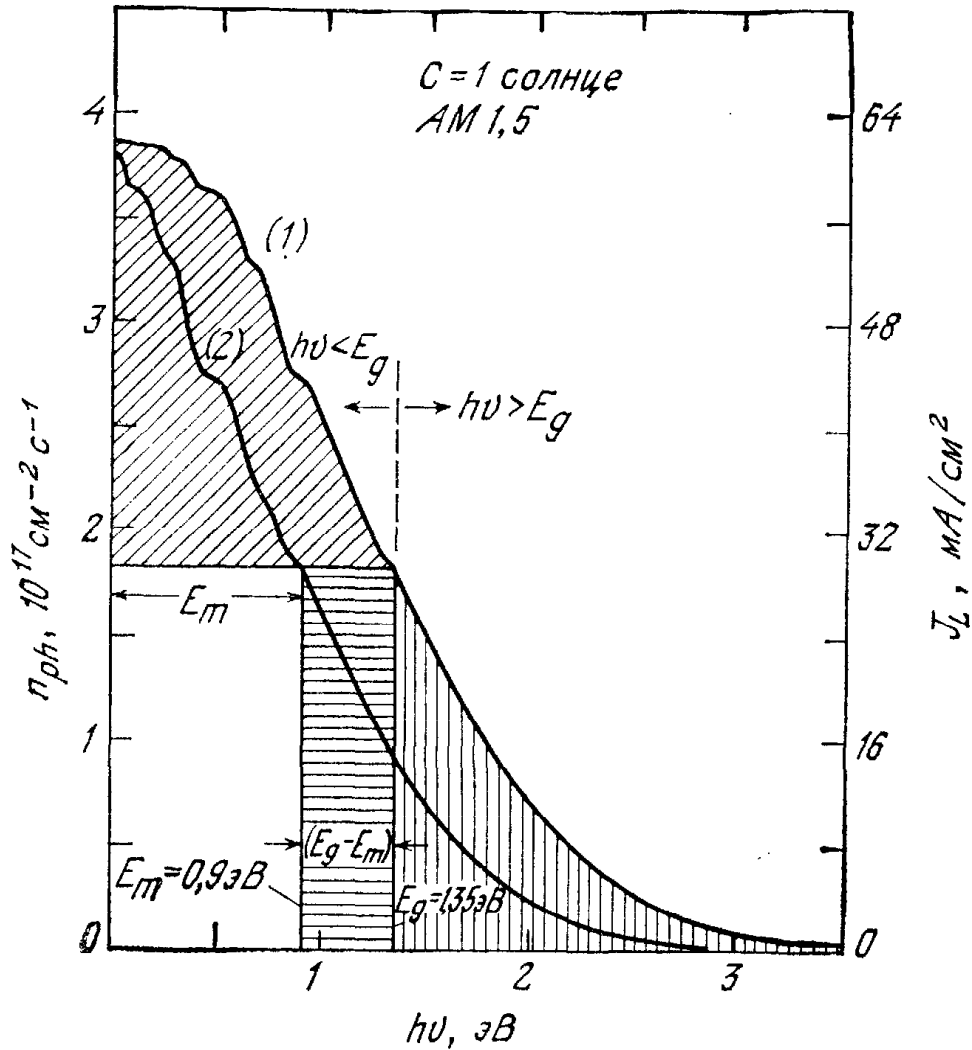


Рис. 6. Зависимость плотности потока фотонов в солнечном спектре от энергии фотона и графический метод определения эффективности преобразования [11].

ношению максимальной выходной мощности к внешней мощности (мощности падающего излучения) P_{in} и может быть определена графически из рис. 6:

$$\eta = \frac{P_m}{P_{in}} = \frac{I_L (E_m/q)}{P_{in}} = [V_m^2 I_s (q/kT) e^{qV_m/kT}] / P_{in}, \quad (11)$$

или

$$\eta = \frac{\text{Прямоугольник со сторонами } E_m \text{ и } n_{ph}}{\text{Площадь под кривой } I}, \quad (11a)$$

где площадь под кривой I равна $5,2 \cdot 10^{17}$ эВ/(см²·с). Максимальная эффективность оказывается равной 31 % и достигается при $E_g = 1,35$ эВ, если использовать параметры материала, характерные для полупроводников типа $A^{III}B^V$.

На рис. 7 приведена зависимость идеальной эффективности солнечного элемента при 300 К от ширины запрещенной зоны [13]. Кривая с отметкой $C = 1$ соответствует AM1,5. Небольшие осцилляции на ней связаны с характером поглощения излучения в атмосфере. Предельное значение идеальной эффективности можно также вычислить на основании принципа детального равно-

веса [14] или в предположении, что потери обусловлены лишь излучательной рекомбинацией [11]. Эти методы также приводят к результатам, близким к получаемым с помощью графического построения и приведенным на рис. 7. Отметим, что кривая эффективности имеет широкий максимум, в пределах которого к. п. д. слабо зависит от E_g . Поэтому все полупроводники, которые имеют ширину запрещенной зоны от 1 до 2 эВ, относятся к материалам, пригодным для создания солнечных элементов. Существует ряд факторов, снижающих идеальную эффективность, вследствие чего реальные значения эффективности преобразования ниже идеальных. О реальных солнечных элементах речь пойдет в следующих разделах.

На рис. 7 приведена также кривая идеальной эффективности для случая оптически сконцентрированного излучения интенсивностью 1000 солнц (т. е. 844 кВт/м^2). Детальное обсуждение проблем, связанных с оптической концентрацией излучения, проведено в разд. 14.5. Максимум идеальной эффективности преобразования возрастает с 31 % (при $C = 1$) до 37 % (при $C = 1000$). Это возрастание связано в основном с увеличением $V_{\text{хх}}$, которое приводит к увеличению E_m в соответствии с формулой (9). Поведение плотности тока короткого замыкания J_L иллюстрируется рис. 8 [11]. Отметим, что расстояние по горизонтали между кривыми 1 и 2 на этом рисунке меньше, чем на рис. 6. На рис. 8 приведена также зонная диаграмма полупроводниковой

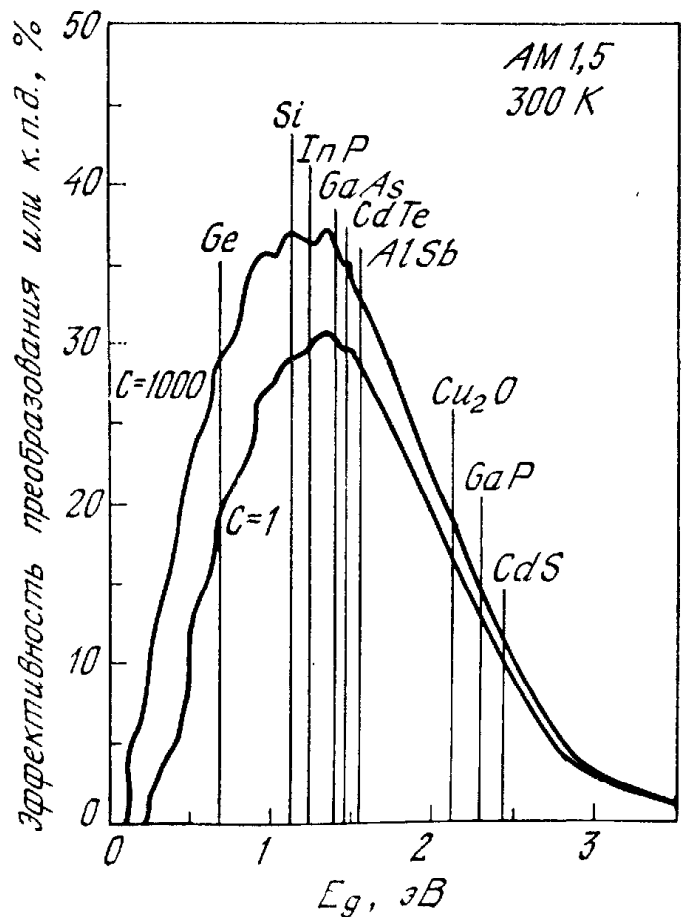


Рис. 7. Зависимость идеального к. п. д. солнечного элемента от E_g при солнечном освещении и при 1000-кратной концентрации солнечного излучения ($T = 300 \text{ K}$) [13].

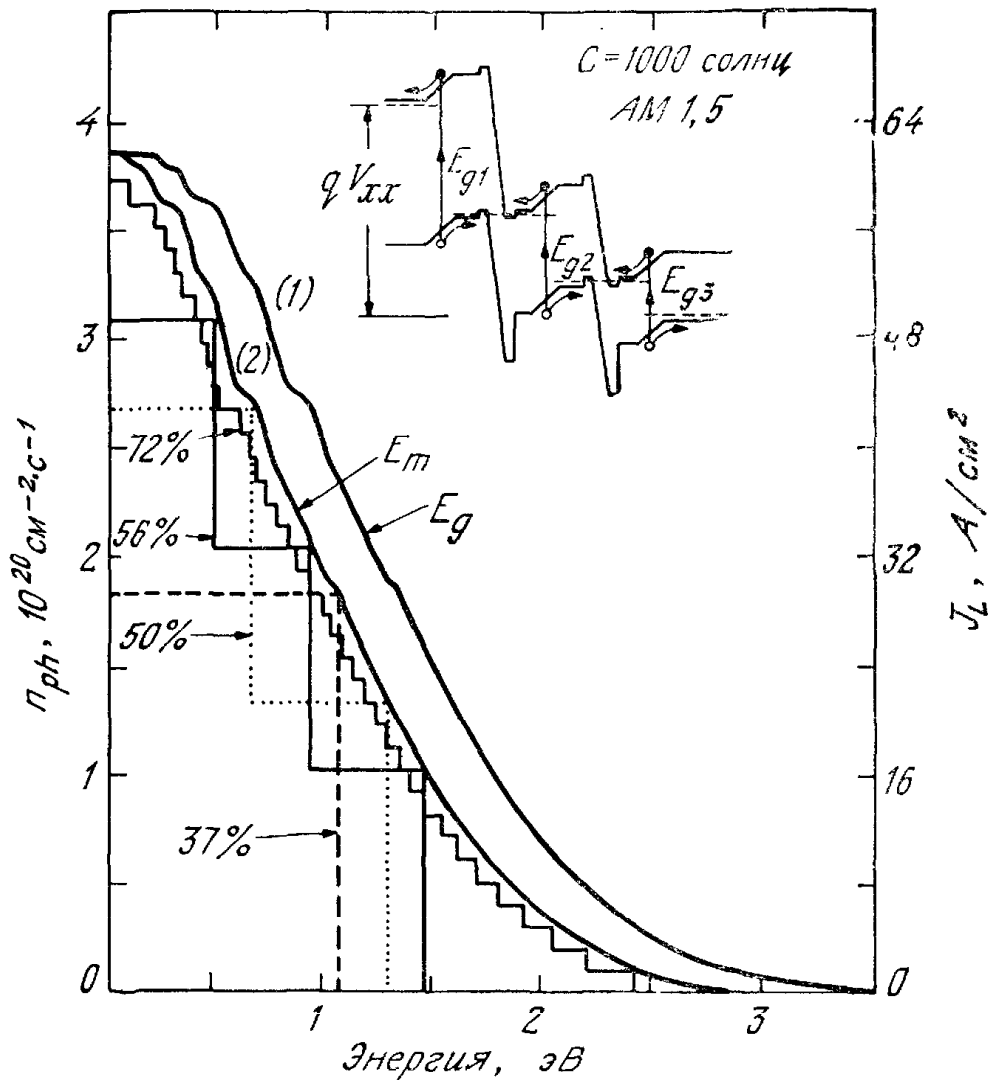


Рис. 8. Зависимость плотности потока фотонов от их энергии в спектре концентрированного солнечного излучения и графический метод определения к. п. д. в каскадных солнечных элементах с последовательно соединенными $p-n$ -переходами [11].

структуры, которую можно использовать для высокоэффективного преобразования концентрированного солнечного излучения. Для нахождения максимальной эффективности можно использовать графический метод, положив плотности тока во всех $p-n$ -переходах одинаковыми. Для двух последовательных $p-n$ -переходов максимально возможный к. п. д. достигает 50 % и реализуется при $E_{g1} = 1,56$ эВ и $E_{g2} = 0,94$ эВ. Для трех переходов к. п. д. достигает 56 % при $E_{g1} = 1,75$ эВ, $E_{g2} = 1,18$ эВ и $E_{g3} = 0,75$ эВ. При дальнейшем увеличении числа переходов к. п. д. растет очень медленно: так, при 36 переходах максимум к. п. д. равен 72 %.

14.3. СОЛНЕЧНЫЕ ЭЛЕМЕНТЫ НА $p-n$ -ПЕРЕХОДАХ

В этом разделе рассмотрены прежде всего кремниевые солнечные элементы с $p-n$ -переходами, поскольку они служат как бы эталонным прибором для всех солнечных батарей. На спутниках и космических кораблях плоские солнечные батареи слу-

жат основным источником питания при длительных полетах. При этом главными проблемами являются эффективность преобразования и надежность элемента, поскольку под влиянием частиц с высокой энергией на внешних орбитах характеристики элемента ухудшаются со временем. В наземных условиях можно применять как плоские солнечные батареи, так и системы солнечных батарей с концентраторами. При этом наряду с проблемами повышения эффективности преобразования и надежности элемента на первом плане оказывается проблема снижения его стоимости, поскольку необходимо, чтобы наземные солнечные системы оказались в конце концов конкурентоспособными с другими источниками энергии.

Работы по наземным плоскочелюпным системам направлены в основном на максимально возможное снижение стоимости батарей при одновременном сохранении к. п. д. преобразования на уровне не менее 10 %. Для этой цели используются такие методы создания подложек, как выращивание ленты через фильеру или способом «пластина к пластине» и выращивание дендритных лент [7]. Для снижения стоимости подложек также применяется выращивание поликристаллического кремния на керамике либо на металлургическом кремнии. К наиболее вероятным кандидатам для использования в плоскочелюпных системах следует отнести солнечные элементы на аморфном кремнии и на тонких пленках CdS. Последние описаны в разд. 14.4. Работы по созданию концентрирующих систем направлены в основном на увеличение к. п. д. преобразования в условиях высокой концентрации солнечной энергии и на минимизацию стоимости всей системы как целого. Солнечные элементы для концентраторов рассмотрены в разд. 14.5. В разд. 14.3 мы анализируем основные рабочие характеристики солнечных элементов с $p-n$ -переходами.

14.3.1. Спектральный отклик

Схематически типичный солнечный элемент изображен на рис. 9. Он состоит из мелкого $p-n$ -перехода, созданного у поверхности (например, с помощью диффузии), лицевого омического полоскового контакта, тылового омического сплошного контакта и просветляющего покрытия на лицевой поверхности.

Если на лицевую поверхность падает монохроматический свет с длиной волны λ , фототок и спектральный отклик (число коллектируемых электронов, приходящихся на один падающий фотон с данной длиной волны) можно определить следующим образом. Скорость генерации электронно-дырочных пар на расстоянии x от поверхности полупроводника (рис. 10, а) определяется выражением

$$G(\lambda, x) = \alpha(\lambda) F(\lambda) [1 - R(\lambda)] \exp[-\alpha(\lambda)x], \quad (12)$$

где $\alpha(\lambda)$ — коэффициент поглощения, $F(\lambda)$ — плотность потока падающих фотонов в единичном спектральном интервале, $R(\lambda)$ — доля фотонов, отражающихся от поверхности [3].

При низком уровне инжекции одномерные стационарные уравнения непрерывности имеют вид

$$G_n - \frac{n_p - n_{p0}}{\tau_n} + \frac{1}{q} \frac{dJ_n}{dx} = 0 \quad (13)$$

для электронов в полупроводнике p -типа и

$$G_p - \frac{p_n - p_{n0}}{\tau_p} - \frac{1}{q} \frac{dJ_p}{dx} = 0 \quad (14)$$

для дырок в полупроводнике n -типа. При этом плотности тока электронов и дырок соответственно определяются выражениями

$$J_n = q\mu_n n_p \mathcal{E} + qD_n \frac{dn_p}{dx}, \quad (15)$$

$$J_p = q\mu_p p_n \mathcal{E} - qD_p \frac{dp_n}{dx}. \quad (16)$$

В солнечном элементе с резким p — n -переходом и постоянными уровнями легирования по обеим сторонам перехода (рис. 10, б и в) вне обедненной области электрическое поле отсутствует. Если рассматривается переход с подложкой p -типа и n -слоем у лицевой поверхности, с помощью уравнений (12), (14) и (16) можно получить уравнение, описывающее распределение неосновных носителей в верхней части перехода:

$$D_p \frac{d^2 p_n}{dx^2} - \frac{p_n - p_{n0}}{\tau_p} + \alpha F (1 - R) \exp(-\alpha x) = 0. \quad (17)$$

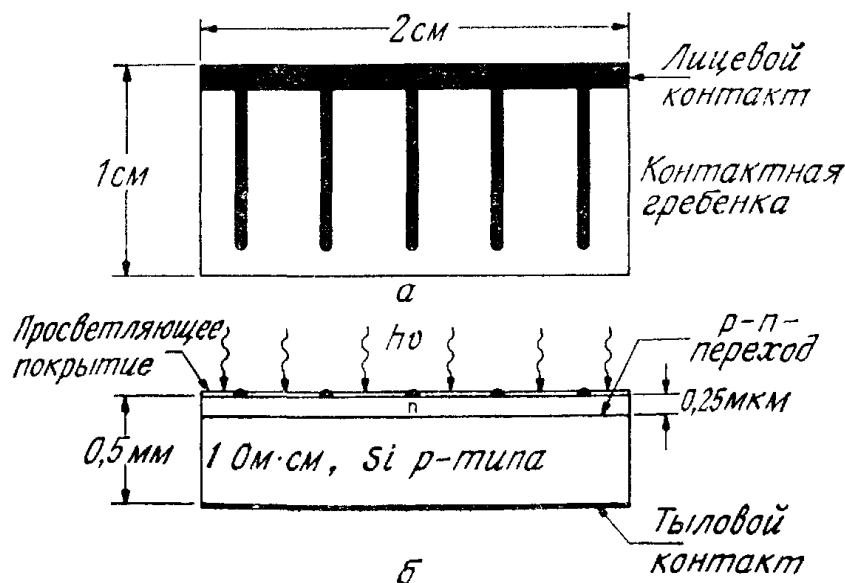


Рис. 9. Схематическое изображение кремниевого солнечного элемента с p — n -переходом.

а — вид сверху; б — вид сбоку.

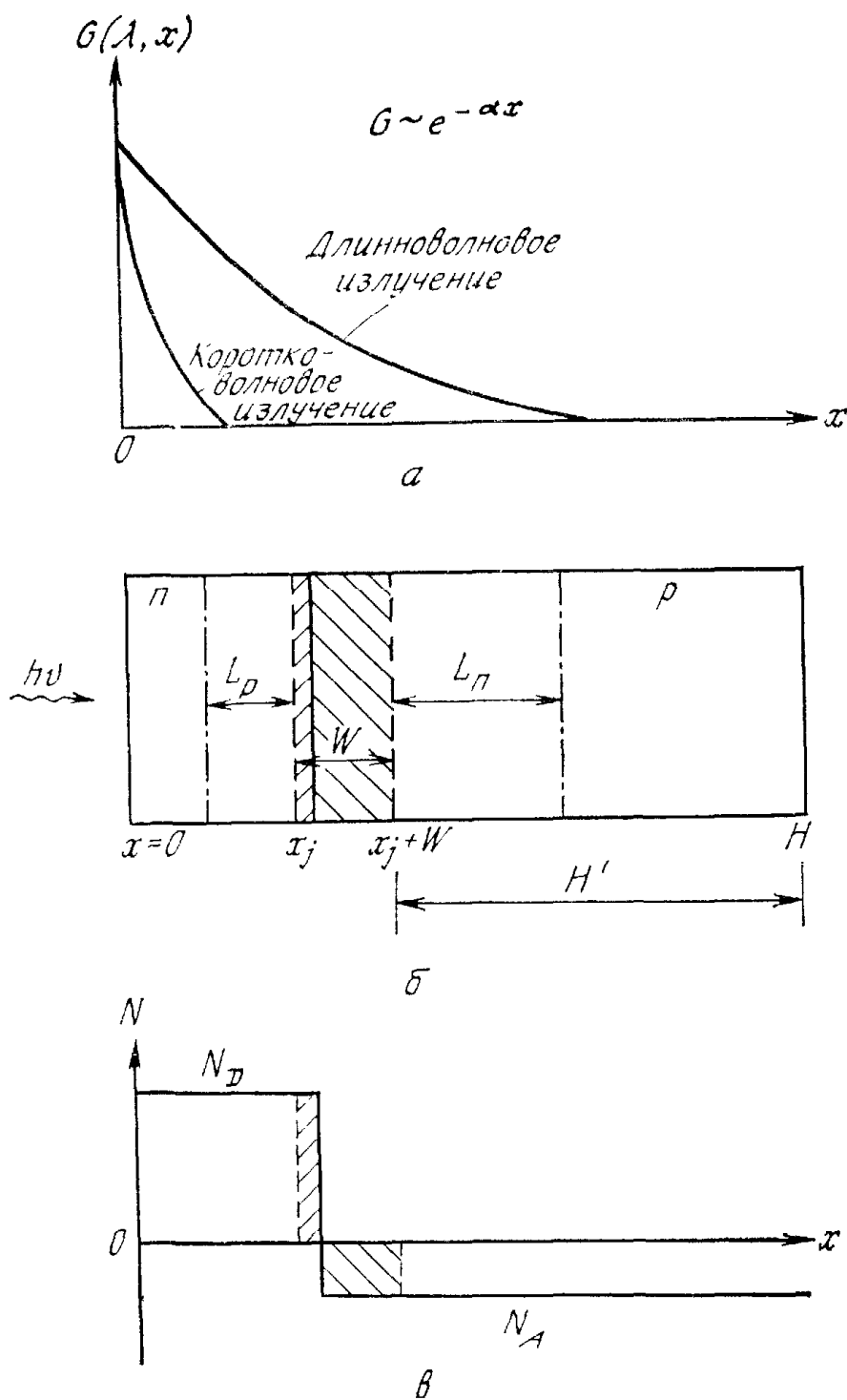


Рис. 10. Зависимость скорости генерации электронно-дырочных пар от расстояния от поверхности для длинноволнового и коротковолнового света (а), размеры солнечного элемента и характерные длины диффузии неосновных носителей (б), а также предполагаемый профиль легирования солнечного элемента (в).

Общее решение этого уравнения имеет вид

$$p_n - p_{no} = A \operatorname{ch}(x/L_p) + B \operatorname{sh}(x/L_p) - \frac{\alpha F (1-R) \tau_p}{\alpha^2 L_p^2 - 1} \exp(-\alpha x), \quad (18)$$

где $L_p = (D_p \tau_p)^{1/2}$ — диффузионная длина. Это решение должно удовлетворять двум граничным условиям: на поверхности полу-

проводника происходит поверхностная рекомбинация носителей, характеризующаяся скоростью S_p ; поэтому

$$D_p \frac{d(p_n - p_{no})}{dx} = S_p (p_n - p_{no}) \quad \text{при } x = 0; \quad (19)$$

на краю области обеднения избыточная концентрация неосновных носителей относительно мала (это связано с существованием тянущего поля в обедненном слое), поэтому

$$p_n - p_{no} \simeq 0 \quad \text{при } x = x_j. \quad (20)$$

Подставляя эти граничные условия в формулу (18), получаем выражение для концентрации дырок

$$p_n - p_{no} = [\alpha F (1 - R) \tau_p / (\alpha^2 L_p^2 - 1)] \times \\ \times \frac{\left(\frac{S_p L_p}{D_p} + \alpha L_p \right) \operatorname{sh} \left(\frac{x_j - x}{L_p} \right) + e^{-\alpha x_j} \left(\frac{S_p L_p}{D_p} \operatorname{sh} \frac{x}{L_p} + \operatorname{ch} \frac{x}{L_p} \right)}{(S_p L_p / D_p) \operatorname{sh} (x_j / L_p) + \operatorname{ch} (x_j / L_p)} - e^{-\alpha x}. \quad (21)$$

Отсюда вытекает, что результирующая плотность дырочного фототока на краю обедненного слоя равна

$$J_p = -q D_p \left(\frac{dp_n}{dx} \right)_{x_j} = [q F (1 - R) \alpha L_p / (\alpha^2 L_p^2 - 1)] \times \\ \times \left[\frac{\left(\frac{S_p L_p}{D_p} + \alpha L_p \right) - e^{-\alpha x_j} \left(\frac{S_p L_p}{D_p} \operatorname{ch} \frac{x_j}{L_p} + \operatorname{sh} \frac{x_j}{L_p} \right)}{(S_p L_p / D_p) \operatorname{sh} (x_j / L_p) + \operatorname{ch} (x_j / L_p)} - \alpha L_p e^{-\alpha x_j} \right]. \quad (22)$$

Этот фототок будет коллектироваться с лицевой стороны n - p -перехода солнечного элемента с p -базой при заданной длине волны падающего излучения и в предположении, что время жизни носителей, их подвижность и уровень легирования в этой области постоянны.

Чтобы найти электронный фототок, коллектируемый из подложки, следует использовать уравнения (12), (13) и (15) при следующих граничных условиях:

$$n_p - n_{po} \cong 0 \quad \text{при } x = x_j + W, \quad (23)$$

$$S_n (n_p - n_{po}) = -D_n dn_p / dx \quad \text{при } x = H, \quad (24)$$

где W — ширина обедненного слоя, H — полная толщина элемента. Условие (23) означает, что избыточная концентрация неосновных носителей вблизи края обедненного слоя равна нулю, а условие (24) определяет скорость поверхностной рекомбинации вблизи омического контакта на тыловой поверхности.

При использовании этих граничных условий распределение концентрации электронов в однородно легированной базе p -типа имеет вид

$$n_p - n_{p0} = \frac{\alpha F (1 - R) \tau_n}{\alpha^2 L_n^2 - 1} \exp[-\alpha(x_j + W)] \cdot \left\{ \operatorname{ch}\left(\frac{x - x_j - W}{L_n}\right) - \right. \\ \left. - e^{-\alpha(x - x_j - W)} - \frac{(S_n L_n / D_n) [\operatorname{ch}(H'/L_n) - \exp(-\alpha H')] + \operatorname{sh}(H'/L_n) + \alpha L_n e^{-\alpha H'}}{(S_n L_n / D_n) \operatorname{sh}(H'/L_n) + \operatorname{ch}(H'/L_n)} \times \right. \\ \left. \times \operatorname{sh}\left(\frac{x - x_j - W}{L_n}\right) \right\}, \quad (25)$$

а электронный фототок, коллектируемый на краю обедненного слоя (при $x = x_j + W$), равен

$$J_n = q D_n \left(\frac{dn_p}{dx} \right)_{x_j + W} = \frac{q F (1 - R) \alpha L_n}{\alpha^2 L_n^2 - 1} \exp[-\alpha(x_j + W)] \times \\ \times \left\{ \alpha L_n - \frac{(S_n L_n / D_n) [\operatorname{ch}(H'/L_n) - \exp(-\alpha H')] + \operatorname{sh}(H'/L_n) + \alpha L_n \exp(-\alpha H')}{(S_n L_n / D_n) \operatorname{sh}(H'/L_n) + \operatorname{ch}(H'/L_n)} \right\}, \quad (26)$$

где H — толщина квазинейтральной области p -базы (рис. 10, б). Выражение (26) выведено в предположении, что время жизни носителей, их подвижность и уровень легирования полупроводника постоянны по всей толщине базы. Если же эти величины изменяются с координатой, ток можно найти лишь с помощью численных методов решения.

Фототок возникает вследствие генерации носителей внутри обедненного слоя. В этой области электрическое поле настолько велико, что электроны и дырки, появляющиеся под действием света, выносятся из обедненного слоя еще до того, как успеют рекомбинировать между собой. Поэтому фототок обедненного слоя в единичном спектральном интервале равен числу фотонов, поглощаемых в этом слое в единицу времени:

$$J_{dr} = q F (1 - R) \exp(-\alpha x_j) [1 - \exp(-\alpha W)]. \quad (27)$$

Полный фототок, возникающий при поглощении света с заданной длиной волны, равен сумме выражений (22), (26) и (27):

$$J(\lambda) = J_p(\lambda) + J_n(\lambda) + J_{dr}(\lambda). \quad (28)$$

Спектральный отклик SP равен этой сумме, деленной на величину qF , если речь идет о наблюдаемом (внешнем) отклике, либо деленной на величину $qF(1 - R)$, если речь идет о внутреннем спектральном отклике:

$$SR(\lambda) = \frac{1}{qF(\lambda) [1 - R(\lambda)]} [J_p(\lambda) + J_n(\lambda) + J_{dr}(\lambda)]. \quad (29)$$

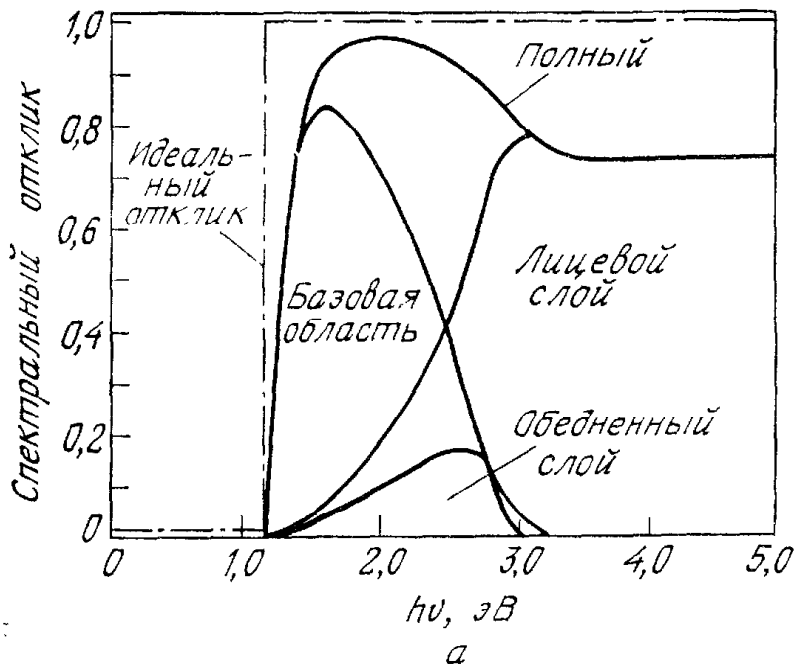
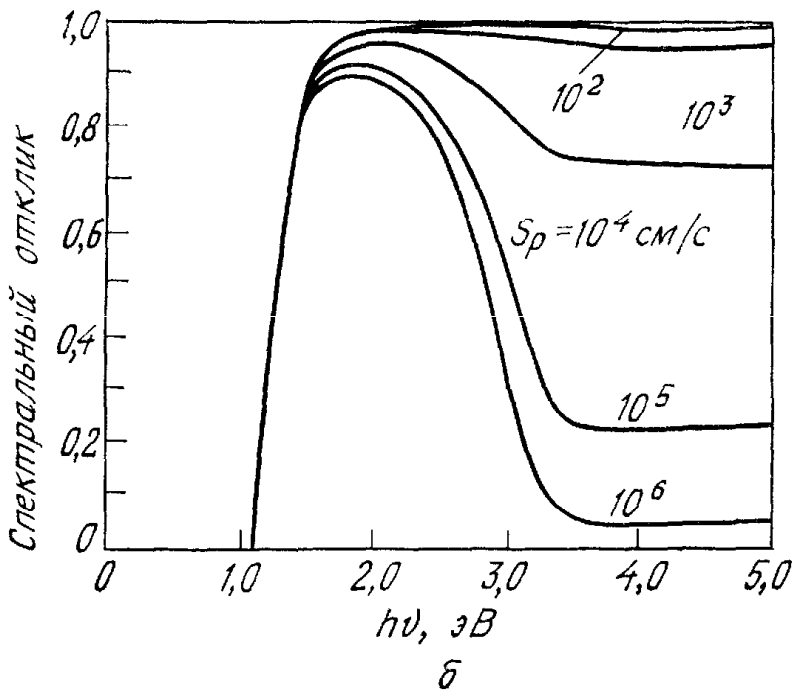


Рис. 11. Рассчитанный внутренний спектральный отклик кремниевого элемента с p -базой (а) (отдельно приведены вклады каждой из трех областей) и тот же рассчитанный отклик с p -базой при разных скоростях поверхностной рекомбинации (б) [3].



Идеальный внутренний спектральный отклик для полупроводника с шириной запрещенной зоны E_g представляет собой ступеньку: он равен 0 при $h\nu < E_g$ и равен 1 при $h\nu \geq E_g$ (штрихпунктирная кривая на рис. 11, а). На рис. 11, а показан реальный внутренний спектральный отклик для кремниевого солнечного элемента с p -базой. Этот отклик существенно отличается от идеальной ступеньки при больших энергиях фотонов [3]. На рис. 11, а проведены также спектральные зависимости вклада каждой из трех областей (трех компонент тока). Здесь использовались следующие параметры прибора: $N_D = 5 \cdot 10^{19} \text{ см}^{-3}$, $N_A = 1,5 \cdot 10^{16} \text{ см}^{-3}$, $\tau_p = 0,4 \text{ мкс}$, $\tau_n = 10 \text{ мкс}$, $x_j = 0,5 \text{ мкм}$, $H = 450 \text{ мкм}$, S_p (лицевая поверхность) $= 10^4 \text{ см/с}$ и S_n (тыловая поверхность) $= \infty$. При поглощении фотонов с низкими энергиями основная доля носителей генерируется в базовой об-

ласти, поскольку коэффициент поглощения в Si при этом мал. Если же энергия фотонов превышает 2,5 эВ, основной вклад в фототок вносит лицевой слой. При энергиях фотонов выше 3,5 эВ α превышает 10^6 см^{-1} , и спектральный отклик полностью определяется лицевым слоем. Поскольку величина S_p предполагается достаточно большой, поверхностная рекомбинация на лицевой поверхности элемента приводит к значительному уменьшению спектрального отклика по сравнению с идеальным значением. При $\alpha L_p \gg 1$ и $\alpha x_j \gg 1$ спектральный отклик приближается асимптотически к величине (которая определяется фототоком лицевой стороны перехода), равной

$$SR = \frac{1 + S_p/\alpha D_p}{(S_p L_p/D_p) \text{sh}(x_j/L_p) + \text{ch}(x_j/L_p)}. \quad (30)$$

Скорость поверхностной рекомбинации S_p сильно влияет на спектральный отклик, особенно при высоких энергиях фотонов. На рис. 11, б приведены кривые спектрального отклика для приборов с теми же параметрами, что и на рис. 11, а (с одним отличием: здесь S_p меняется в диапазоне $10^2 - 10^6 \text{ см/с}$). Следует отметить сильное уменьшение спектрального отклика при увеличении S_p . Из выражения (30) следует, что при заданной величине S_p спектральный отклик улучшается при увеличении диффузионной длины L_p . В общем случае для увеличения спектрального отклика в диапазоне используемых длин волн следует уменьшать S_n и S_p и увеличивать L_n и L_p .

Если спектральный отклик известен, полная плотность фототока, протекающего при облучении элемента солнечным светом со спектральным распределением $F(\lambda)$ (рис. 1), равна

$$J_L = q \int_0^{\lambda_m} F(\lambda) [1 - R(\lambda)] SR(\lambda) d\lambda, \quad (31)$$

где λ_m — красная граница поглощения, определяемая шириной запрещенной зоны полупроводника. Для получения максимального значения J_L следует уменьшить $R(\lambda)$ и увеличивать SR во всем диапазоне длин волн $0 < \lambda < \lambda_m$. Идеальная плотность фототока J_L , показанная на рис. 6, получена при $R(\lambda) = 0$ и $SR(\lambda) = 1$.

14.3.2. Вольт-амперные характеристики

Для реальных солнечных элементов следует изменить идеальную эквивалентную схему, приведенную на рис. 4, б, включив в нее последовательное сопротивление, обусловленное омическим падением напряжения на лицевом контакте, и шунтирующее параллельное сопротивление, обусловленное токами утечки (рис. 12,

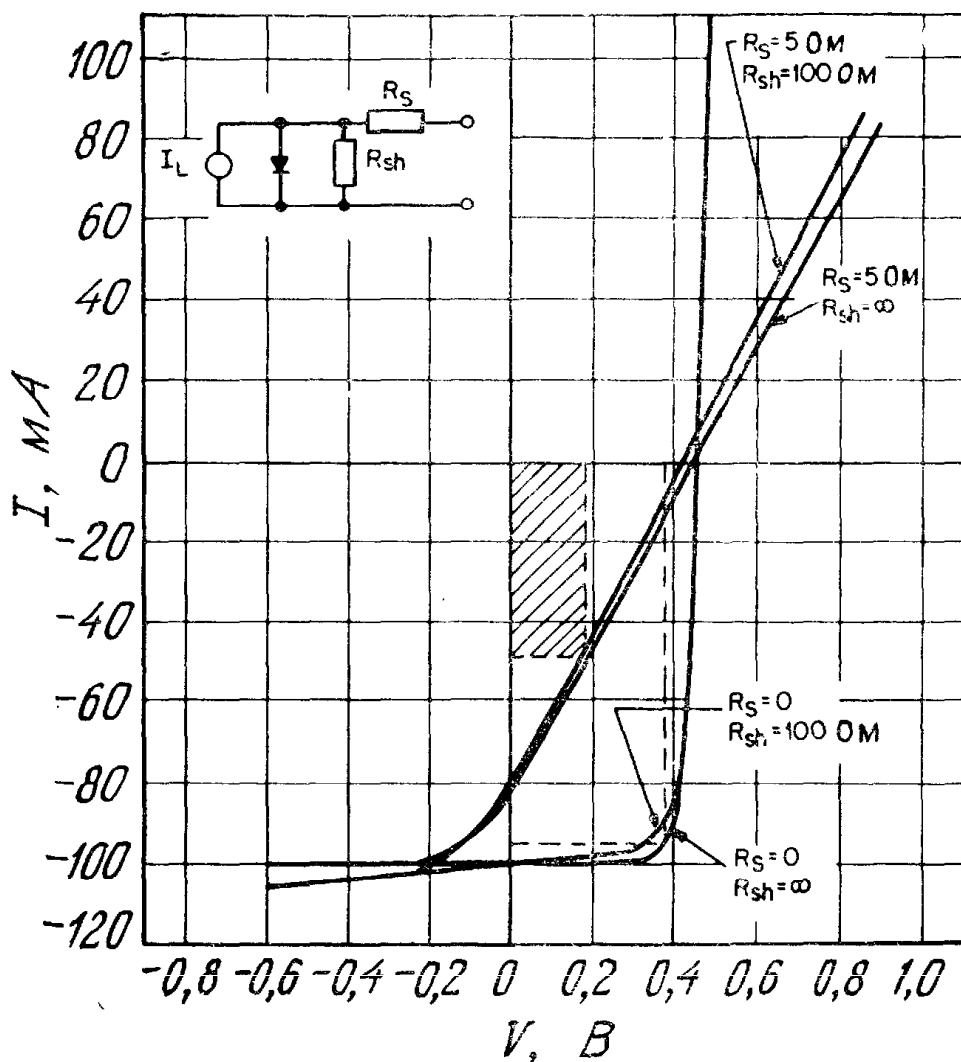


Рис. 12. Теоретические вольт-амперные характеристики солнечного элемента при различных значениях последовательного и параллельного сопротивлений. На вставке показана эквивалентная схема. Параметры элемента те же, что использовались на рис. 5 [12].

вставка). Если ток диода определяется формулой (2), то вольт-амперная характеристика удовлетворяет уравнению [12]

$$\ln \left(\frac{I + I_L}{I_s} - \frac{V - IR_s}{I_s R_{sh}} + 1 \right) = \frac{q}{kT} (V - IR_s). \quad (32)$$

Решения этого уравнения при различных комбинациях параметров R_s и R_{sh} приведены на рис. 12. Значения I_s , I_L и T здесь те же, что на рис. 5. Видно, что шунтирующее сопротивление, даже такое малое, как 100 Ом, не влияет заметно на выходную мощность прибора, тогда как даже весьма небольшое последовательное сопротивление 5 Ом снижает выходную мощность более чем на 30 % по сравнению с оптимальной величиной, достигаемой при $R_s = 0$. Поэтому в дальнейшем мы будем пренебрегать вли-

янием R_{sh} . Тогда ток во внешней цепи и выходная мощность определяются соответственно выражениями

$$I = I_s \left\{ \exp \left[\frac{q(V - IR_s)}{kT} \right] - 1 \right\} - I_L, \quad (33)$$

$$P = |IV| = I \left[\frac{kT}{q} \ln \left(\frac{I + I_L}{I_s} + 1 \right) + IR_s \right]. \quad (34)$$

Относительный максимум выходной мощности составляет 1; 0,77; 0,57; 0,27 и 0,14 при R_s , равном соответственно 0; 1; 2; 5 и 10 Ом. Последовательное сопротивление солнечного элемента зависит от глубины залегания перехода, концентраций примесей в n - и p -областях и от конструкции лицевого омического контакта. Для типичного кремниевого солнечного элемента с геометрией, показанной на рис. 9, последовательное сопротивление составляет $\sim 0,7$ Ом для элементов с p -базой и $\sim 0,4$ Ом для элементов с n -базой [15]. Это различие в величинах сопротивления обусловлено в основном относительно низким сопротивлением подложек n -типа.

С помощью рис. 5 и 12 можно определить фактор заполнения FF :

$$FF \equiv \frac{I_m V_m}{I_L V_{xx}}. \quad (35)$$

Эффективность преобразования (к. п. д.) элемента равна

$$\eta = \frac{I_m V_m}{P_{in}} = \frac{FF \cdot I_L \cdot V_{xx}}{P_{in}}. \quad (36)$$

Для получения максимального к. п. д. следует максимизировать все три сомножителя в числителе в правой части выражения (36).

В реальных солнечных элементах прямой ток может определяться рекомбинацией носителей в обедненном слое. При этом эффективность преобразования, вообще говоря, ниже, чем в идеальном диоде (разд. 14.2.2). В случае одноуровневых рекомбинационных центров рекомбинационный ток можно записать в виде [16, 17]

$$I_{rec} = I'_s \left[\exp \left(\frac{qV}{2kT} \right) - 1 \right], \quad (37)$$

где

$$\frac{I'_s}{A} = \frac{qn_i W}{V \tau_p \tau_n}. \quad (37a)$$

И в этом случае выражение для эффективности преобразования легко получить с помощью простых уравнений, аналогичных формулам (4)–(9), в которых величину I_s следует заменить на I'_s , а показатель экспоненты разделить на 2. Эффективность преобразования в том случае, когда прямой ток определяется реком-

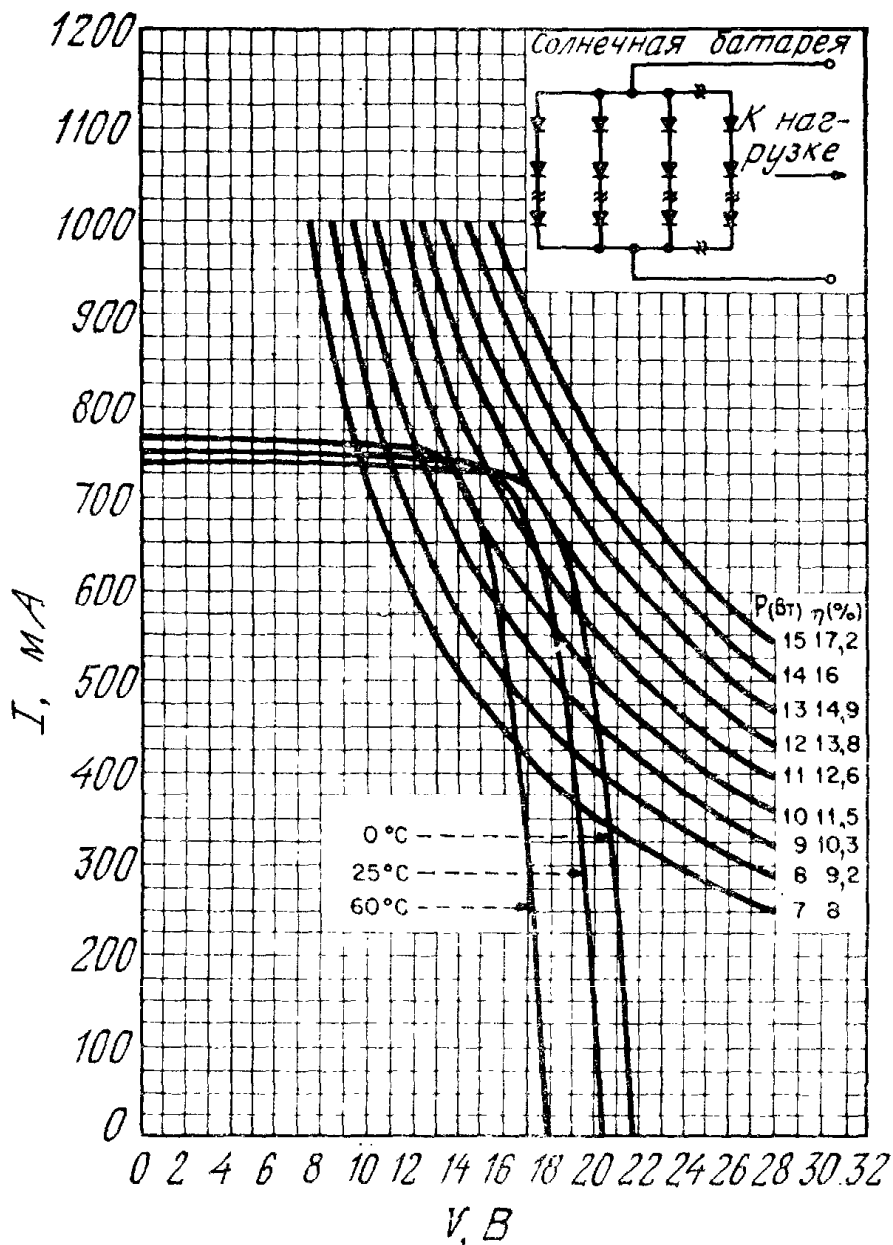


Рис. 13. Вольт-амперные характеристики солнечной батареи и кривые постоянной мощности [5].

бинацией в обедненном слое, оказывается значительно меньше, чем в идеальном диоде, что обусловлено уменьшением как V_{xx} , так и фактора заполнения. В кремниевых солнечных элементах при 300 К рекомбинационный ток приводит к снижению эффективности преобразования на 25 % [18].

В солнечных элементах, в которых диффузионный и рекомбинационный токи сравнимы по величине или содержится много дефектов, экспоненциальная зависимость прямого тока от смещения перехода может иметь вид $\exp(qV/nkT)$, где величина n называется фактором идеальности. Обычно эффективность преобразования уменьшается с увеличением n .

Хотя отдельный кремниевый солнечный элемент площадью 2 см^2 имеет напряжение холостого хода лишь 0,5—0,6 В и ток короткого замыкания от 30 до 60 мА, последовательно-параллель-

ное соединение таких элементов в большую батарею позволяет подводить к нагрузке существенно бóльшие напряжения и токи. Эквивалентная схема такой солнечной батареи показана на рис. 13 (вставка) [5]. Эта батарея при температуре 60 °С вырабатывает мощность 10 Вт с к. п. д. = 11,5 % при AM1 (~100 мВт/см²); вольт-амперная характеристика касается соответствующей кривой постоянной мощности. Значения V_m и I_m , соответствующие максимальному выделению мощности батареей, равны 14 В и 720 мА соответственно. Работа при пониженных температурах позволяет повысить выходную мощность и к. п. д. батареи.

14.3.3. Влияние температуры и радиации

При увеличении температуры диффузионные длины в Si и GaAs возрастают, поскольку коэффициент диффузии не изменяется либо растет, а время жизни неосновных носителей возрастает при повышении температуры. Увеличение диффузионной длины неосновных носителей приводит к возрастанию J_L . Однако величина V_{xx} при этом уменьшается быстрее за счет экспоненциальной зависимости тока насыщения от температуры. Кроме того, более плавная форма вольт-амперной характеристики при повышенных температурах приводит к уменьшению фактора заполнения. Поэтому в целом повышение температуры приводит к уменьшению эффективности преобразования.

Нормированные значения к. п. д. для солнечных батарей на Si и GaAs приведены на рис. 14 [18]. В условиях протекания идеального тока эффективность линейно уменьшается с возрастанием температуры вплоть до ~200 °С для Si и до ~300 °С для GaAs. При протекании рекомбинационного тока к. п. д. при 20 °С на ~25 % ниже. И в этом случае к. п. д. также уменьшается с увеличением температуры примерно по линейному закону. Из рис. 14 видно, что повышение рабочей температуры элемента всегда приводит к снижению к. п. д.

В космических условиях выходная мощность солнечных элементов понижается в связи с тем, что облучение частицами высоких энергий на удаленных орбитах приводит к образованию дефектов в полупроводнике. Поэтому важно уметь оценивать продолжительность периода работоспособности космических солнечных батарей. Эта продолжительность определяется временем, в течение которого выходная мощность солнечной батареи достаточна для функционирования аппаратуры спутника.

Из выражений (22) и (26) видно, что фототок уменьшается с уменьшением диффузионных длин L_n и L_p . Время жизни избыточных неосновных носителей при облучении полупроводника частицами высоких энергий меняется по закону

$$\frac{1}{\tau} = \frac{1}{\tau_0} + K'\Phi, \quad (38)$$

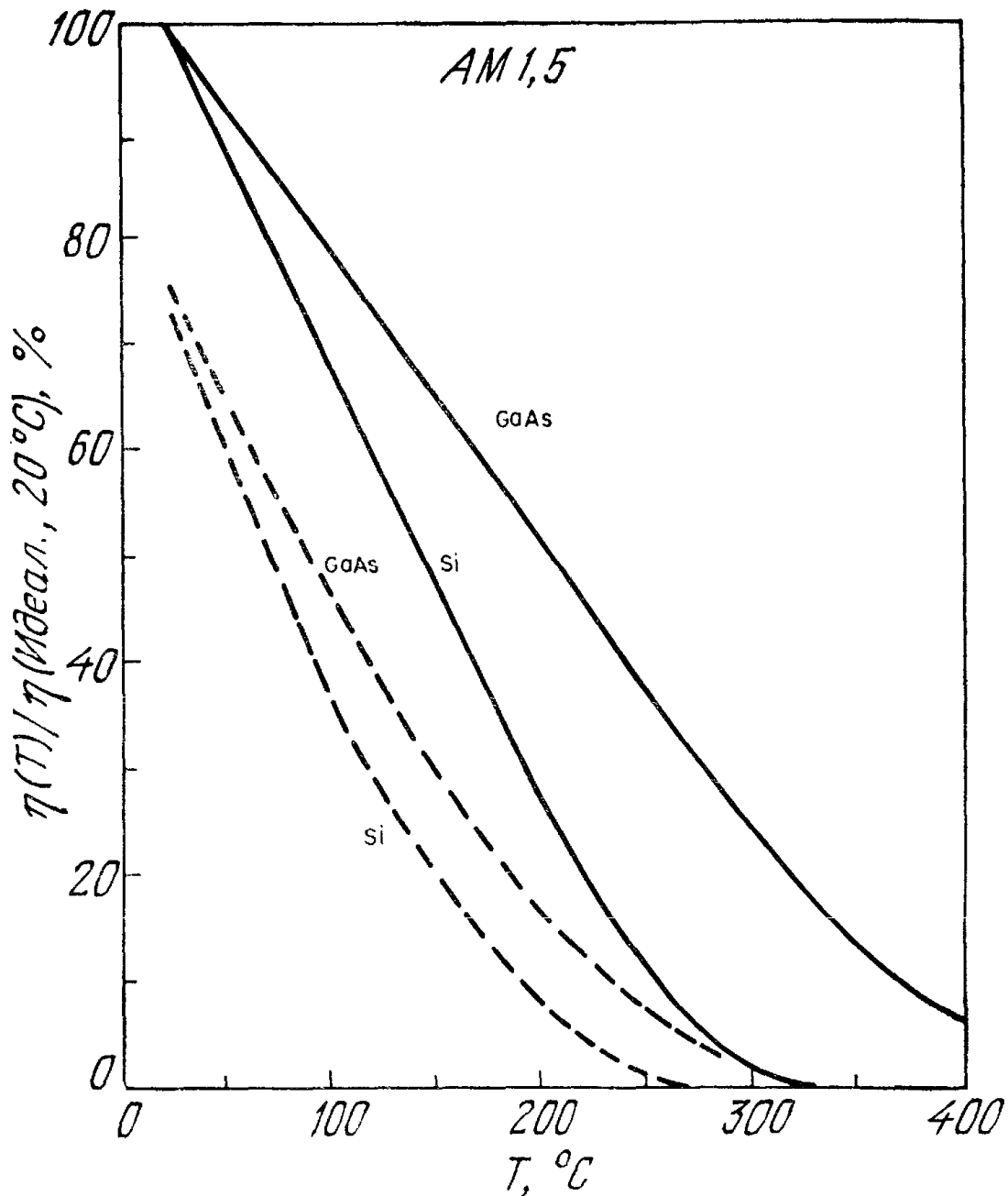


Рис. 14. Нормированная эффективность преобразования для солнечных элементов с $p-n$ -переходами на Si и GaAs при идеальном (сплошные линии) и рекомбинационном (штриховые линии) токах перехода [18].

где τ_0 — исходное время жизни, K' — постоянная и Φ — доза радиации [19]. Из этого выражения следует, что скорость рекомбинации неосновных носителей пропорциональна исходной концентрации рекомбинационных центров и их концентрации, вводимой в полупроводник в процессе облучения, причем последняя концентрация пропорциональна дозе попадающих на полупроводник частиц. Поскольку диффузионная длина равна $\sqrt{D\tau}$, а D слабо зависит от облучения (или от уровня легирования), соотношение (38) можно записать в виде

$$\frac{1}{L^2} = \frac{1}{L_0^2} + K\Phi, \quad (39)$$

где L_0 — исходная диффузионная длина и $K = K'/D$. На рис. 15 приведены измеренные зависимости диффузионной длины для трех кремниевых солнечных элементов от дозы облучения электронами с энергией 1 МэВ. Элемент с p -базой, чувствительный в голубой области спектра, имеет диффузионный слой n -типа и просветляющее покрытие. Диффузионная длина в этом элементе выбрана такой, чтобы обеспечить максимальный спектральный отклик в голубой области спектра (0,45—0,5 мкм), на которую приходится максимум энергии солнечного излучения. Чувствительный в голубой области спектра элемент с n -базой аналогичен рассмотренному выше элементу с p -базой, за исключением того, что области n - и p -типа здесь меняются местами. Мы видим, что экспериментальные результаты для этих элементов неплохо описываются уравнением (39). Кривая, проходящая через экспериментальные точки элемента с p -базой, рассчитана с помощью формулы (39) при $L_0 = 119$ мкм и $K = 1,7 \cdot 10^{-10}$. Экспериментальные точки для чувствительного в голубой области элемента с n -базой и обычного элемента с n -базой можно аппроксимировать с помощью формулы (39) при $L_0 = 146$ мкм и $K = 1,22 \cdot 10^{-8}$. Из рис. 15 видно, что радиационная стойкость элементов с p -базой существенно выше, чем элементов с n -базой.

Для повышения радиационной стойкости в солнечные элементы вводится литий [20], который легко диффундирует и образует

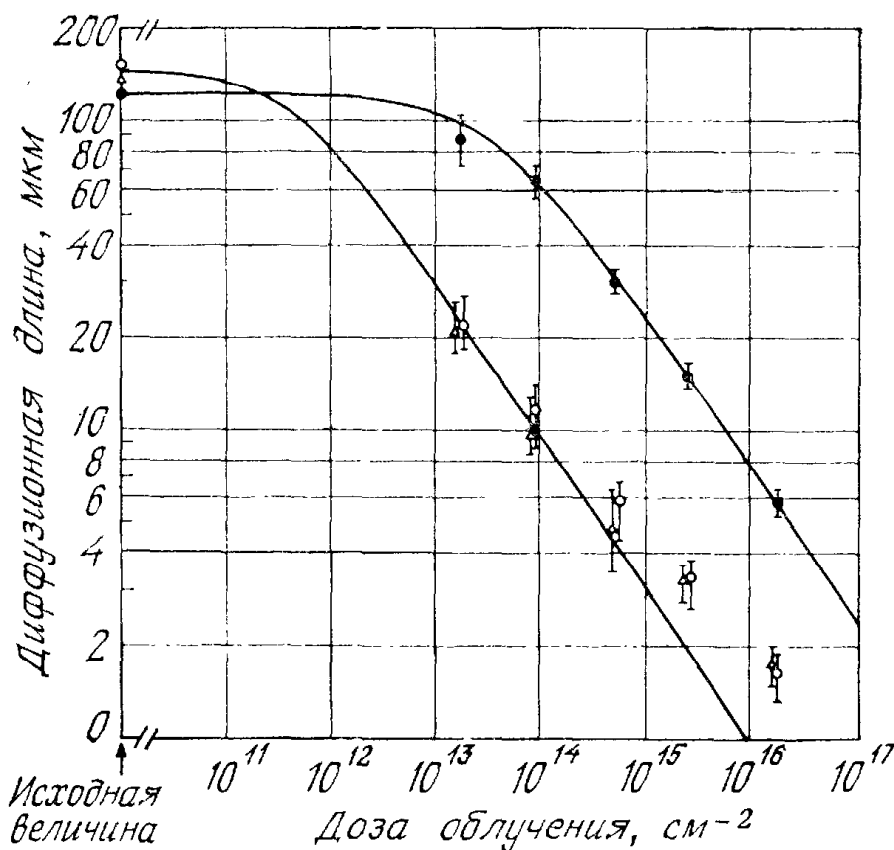


Рис. 15. Зависимость диффузионной длины неосновных носителей от дозы облучения электронами с энергией 1 МэВ [19].

● чувствительный в голубом свете с p -базой; Δ обычный с n -базой; ○ чувствительный в голубом свете с n -базой.

комплексы с радиационными точечными дефектами. Очевидно, Li нейтрализует дефекты и препятствует деградации времени жизни. Для снижения числа частиц высоких энергий, достигающих элемента в космическом пространстве, перед лицевой поверхностью элемента всегда должно помещаться защитное покрытие (например, содержащая церий тончайшая бумага).

14.3.4. Конструкции прибора

Для достижения более высокой эффективности преобразования предлагалось большее число различных конструкций солнечных элементов. Мы рассмотрим некоторые из них.

Солнечный элемент с барьером вблизи тыловой поверхности (БТП) имеет существенно большее выходное напряжение, чем обычные элементы [21]. Зонная диаграмма такого элемента показана на рис. 16. Лицевая поверхность этого элемента изготавливается обычным способом, а вблизи тыловой поверхности перед металлическим омическим контактом создается сильнолегированный полупроводниковый слой. Потенциальный барьер $q\phi_p$, возникающий между двумя базовыми областями, препятствует выходу неосновных носителей из менее легированной области. Элемент с БТП эквивалентен обычному элементу с толщиной полупроводника $x_j + W + W_p$ и с очень малой скоростью поверхностной рекомбинации вблизи тыловой поверхности ($S_n < 100$ см/с). Уменьшение S_n приводит к увеличению спектрального отклика для фотонов с малыми энергиями, а следовательно, и к возрастанию плотности тока короткого замыкания. Кроме того, повышается напряжение холостого хода, что обусловлено возрастанием тока короткого замыкания, уменьшением рекомбинационного тока тылового контакта и, наконец, появлением дополнительной потенциальной энергии $q\phi_p$ между p - и p^+ -областями.

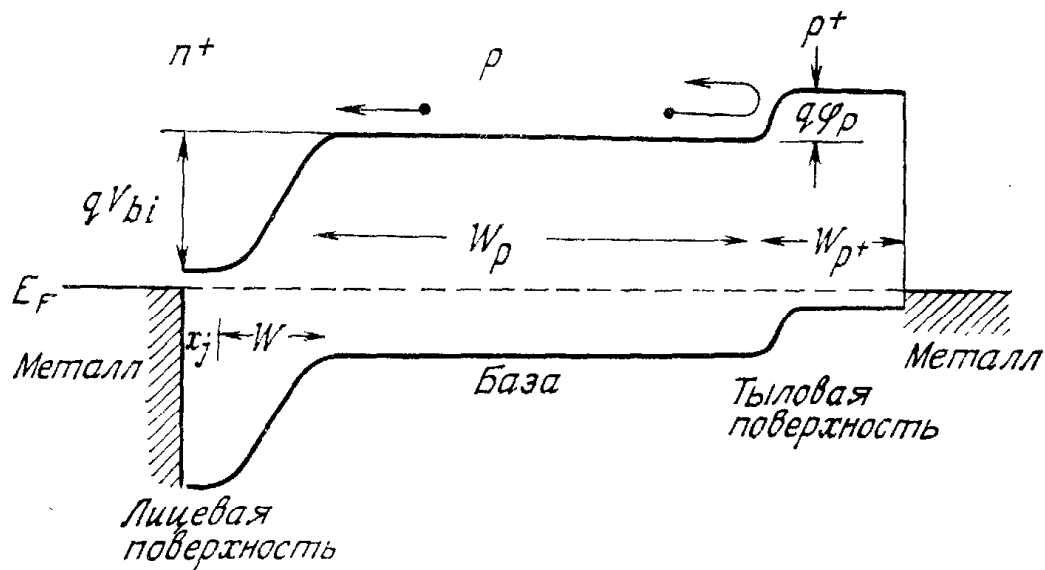


Рис. 16. Зонная диаграмма солнечного $n^+ - p - p^+$ -элемента с барьером вблизи тыловой поверхности [21].

Для создания «фиолетовых» элементов используют пониженную концентрацию легирующей примеси вблизи лицевой поверхности и меньшую глубину залегания перехода [22]. При этом благодаря повышению времени жизни неосновных носителей вблизи лицевой поверхности и относительно узкому лицевому легированному слою спектральный отклик для высокоэнергетичных фотонов сильно возрастает. На рис. 17 для сравнения представлены вместе результаты измерений спектрального отклика обычного элемента с $x_j = 0,4$ мкм и $N_D = 5 \cdot 10^{19}$ см⁻³ и «фиолетового» элемента с $x_j = 0,2$ мкм и поверхностной концентрацией $5 \cdot 10^{18}$ см⁻³. Видно, что «фиолетовый» элемент имеет существенно больший спектральный отклик в фиолетовой области спектра ($h\nu > 2,75$ эВ); отсюда происходит название этого элемента. Форма его спектрального отклика напоминает представленную на рис. 11, б теоретическую кривую, относящуюся к случаю, когда скорость поверхностной рекомбинации на лицевой поверхности $S_p \cong 10^4$ см/с.

На рис. 18, а показан текстурированный элемент, лицевая поверхность которого покрыта пирамидами, созданными с помощью анизотропного травления поверхности кремния, ориентированной в направлении $\langle 100 \rangle$ [23]. Свет, падающий на боковую поверхность одной из пирамид, отражается на другую пирамиду, что приводит к уменьшению оптических потерь. Коэффициент отражения чистого кремния уменьшается с ~ 35 % для плоской поверхности до ~ 20 % для текстурированной поверхности. Нанесение дополнительного просветляющего покрытия снижает потери на отражение вплоть до нескольких процентов (рис. 18, б). Уменьшение отражения приводит к возрастанию как тока короткого замыкания, так и напряжения холостого хода, что в свою очередь повышает эффективность преобразования. Для текстурированных элементов в условиях АМО получен к. п. д. выше 15 %.

Был предложен еще один новый тип солнечных элементов, использующий анизотропное травление поверхности кремния

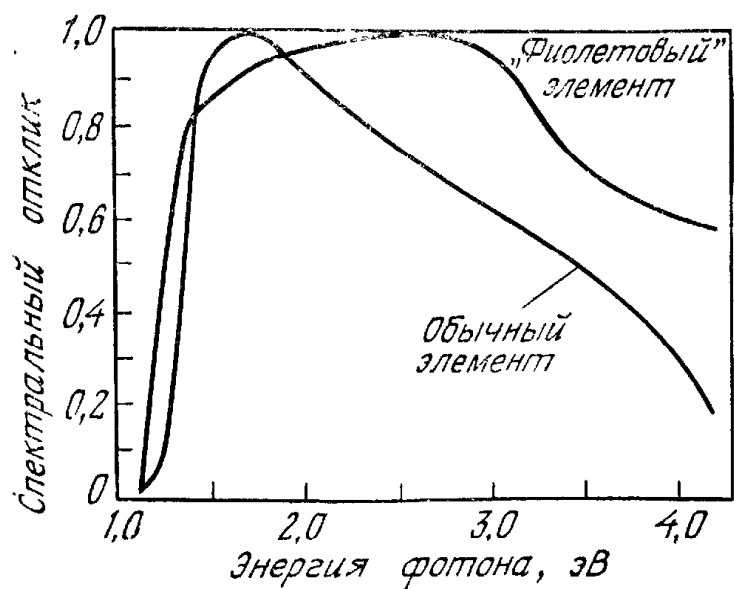


Рис. 17. Измеренные спектральные отклики кремниевых солнечных элементов с p -базой. В фиолетовой области спектра отклик «фиолетового» элемента значительно превышает отклик обычного элемента [22].

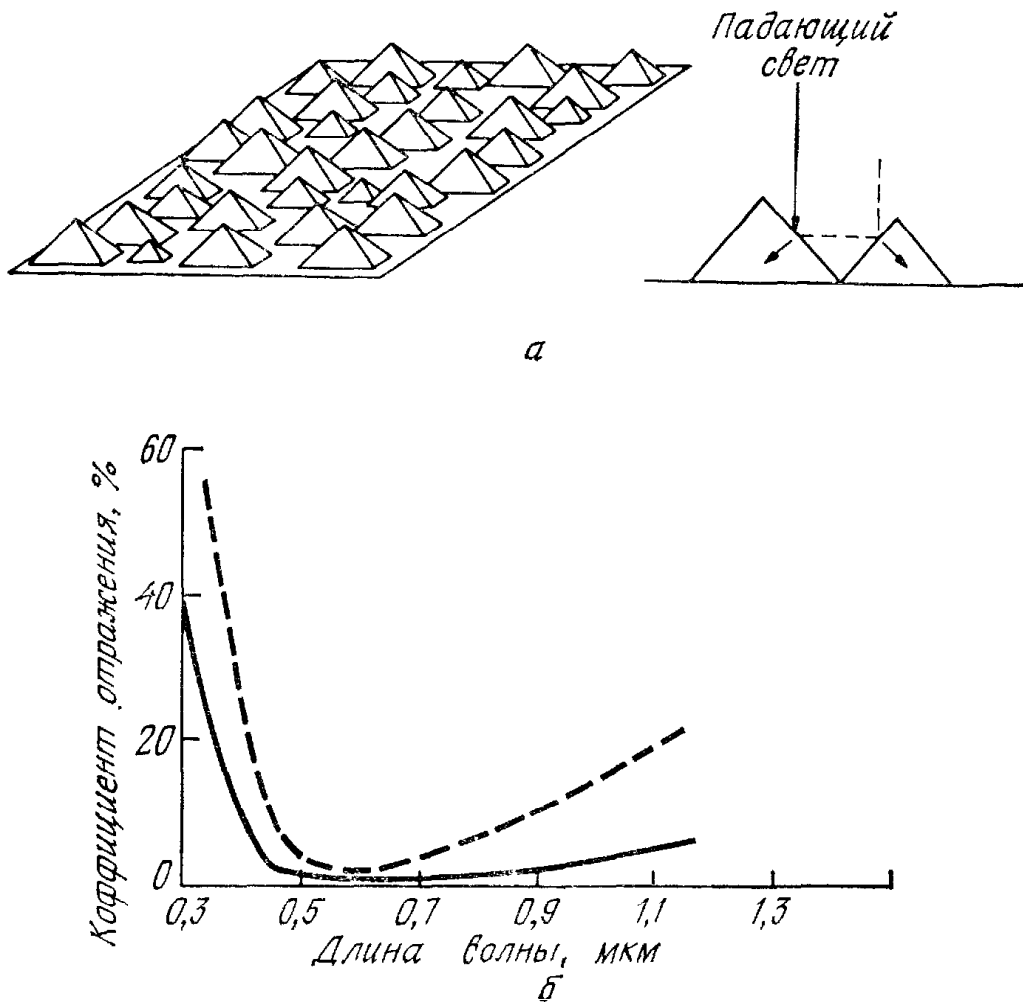


Рис. 18. Текстурированный элемент с пирамидальной поверхностью (а) и зависимость от длины волны коэффициентов отражения солнечного элемента с плоской поверхностью и просветляющим покрытием (штриховая линия) и текстурированного солнечного элемента с просветляющим покрытием (сплошная кривая) (б) [23].

с ориентацией $\langle 100 \rangle$, — многопереходные солнечные элементы на V-канавках [24]. Такой элемент состоит из ряда отдельных $p^+ - p - n^+$ (или $p^+ - p - n^+$)-диодов, соединенных последовательно (рис. 19, а). Трапецеидальная форма отдельных диодов получается за счет анизотропного травления поверхности кремния $\langle 100 \rangle$ через маску термически выращенной двуокиси кремния. Эффективная оптическая толщина такого диода показана на рис. 19, б. При толщине диода 50 мкм эффективная оптическая длина, усредненная по различным траекториям света, превышает 250 мкм. Вследствие этого эффективность фундаментального поглощения (т. е. отношение числа поглощенных фотонов к полному числу фотонов с энергией $h\nu \geq E_g$, попадающих в полупроводник) превышает 93 %. В результате повышенной эффективности поглощения, относительно низкого последовательного сопротивления и отсутствия затенения лицевой поверхности (в этой конструкции нет коллектирующей гребенки) можно ожидать, что данный элемент будет иметь к. п. д. выше 20 %.

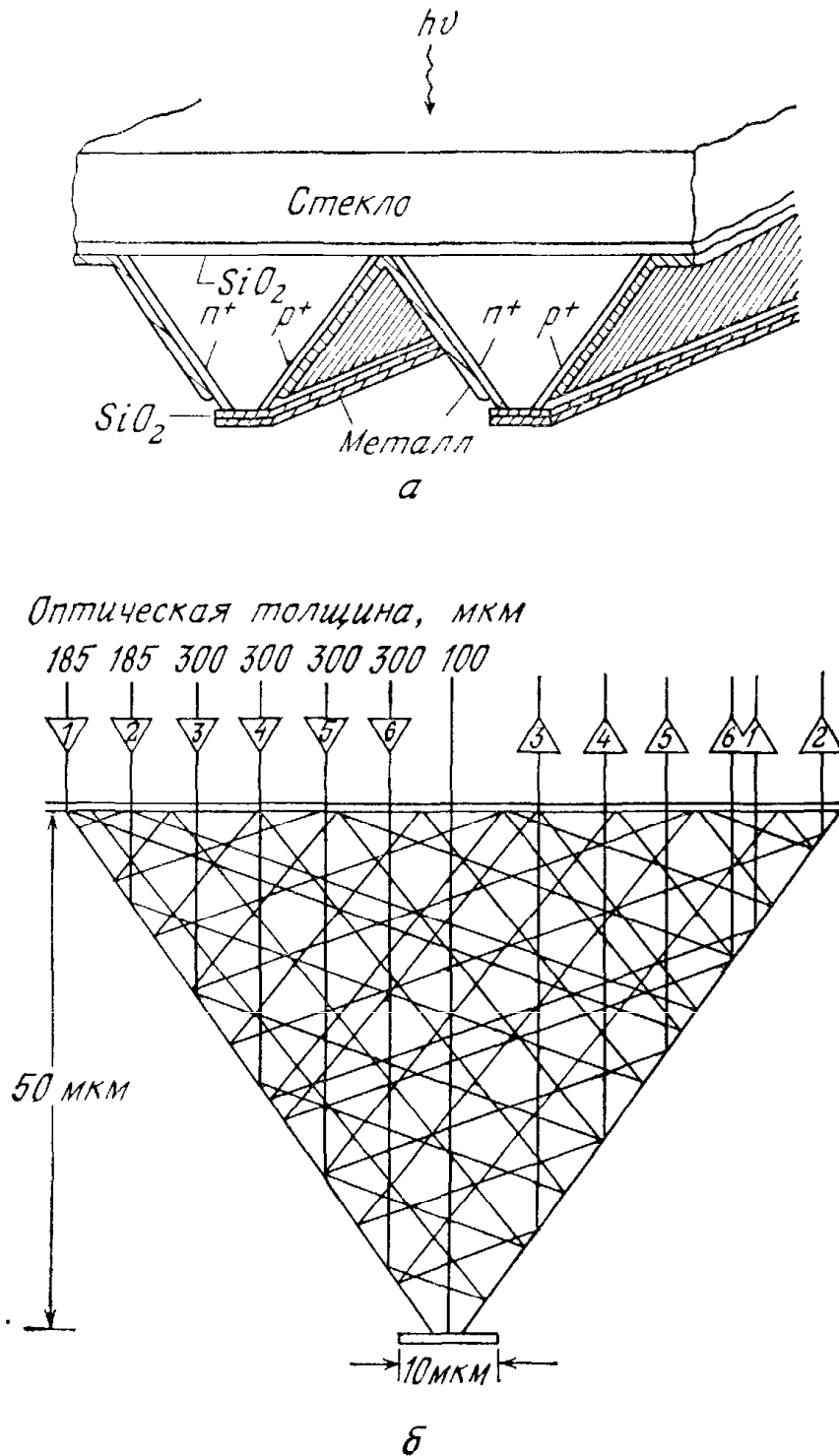


Рис. 19. Многопереходный солнечный элемент на V-канавках (а) и оптические пути различных лучей в полупроводнике (б). Средняя эффективная оптическая толщина полупроводника составляет 250 мкм [24].

На рис. 20, а показан солнечный элемент со спаренными переходами [25], в котором сочетаются концепции элемента с БТП и текстурированного элемента. В этом элементе нет тени от металла и облегчаются проблемы межсоединений благодаря тому, что n⁺- и p⁺- контакты создаются на тыловой поверхности. Прибор работает как биполярный транзистор с изолированным n⁺-эмиттером на лицевой поверхности (рис. 20, б). Электроны, генерируемые светом в эмиттере или в базе, собираются n⁺-коллектором, подобно тому как это имеет место в транзисторе. В та-

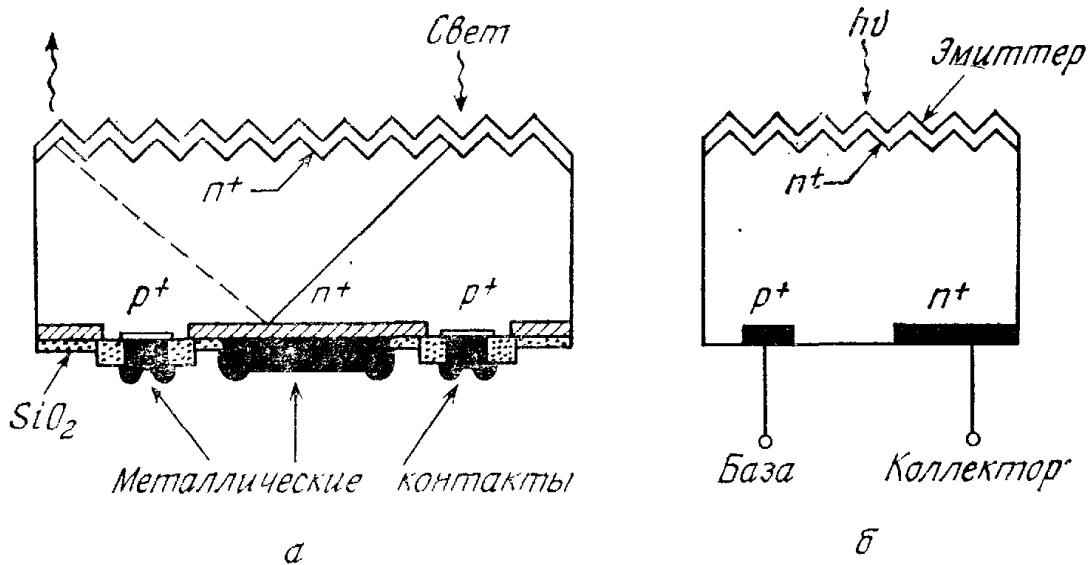


Рис. 20. Поперечное сечение солнечного элемента со спаренными переходами (ЭСП) (а), структура эквивалентного транзистора (б) и зависимость тока короткого замыкания от толщины прибора для ЭСП и обычного элемента в условиях АМ0 (в) [25].

ком элементе зависимость тока короткого замыкания от толщины прибора совершенно непохожа на соответствующие зависимости для обычных солнечных элементов. Для сравнения такие зависимости приведены на рис. 20, в. В солнечных элементах со спаренными переходами используются гораздо более тонкие базовые области, и к. п. д. этих элементов также может превышать 20 %.

Еще одной новой конструкцией являются солнечные элементы с вертикальными переходами, в которых поверхности переходов и металлические поверхности расположены перпендикулярно поверхности элемента [26]. Прибор с вертикальными переходами

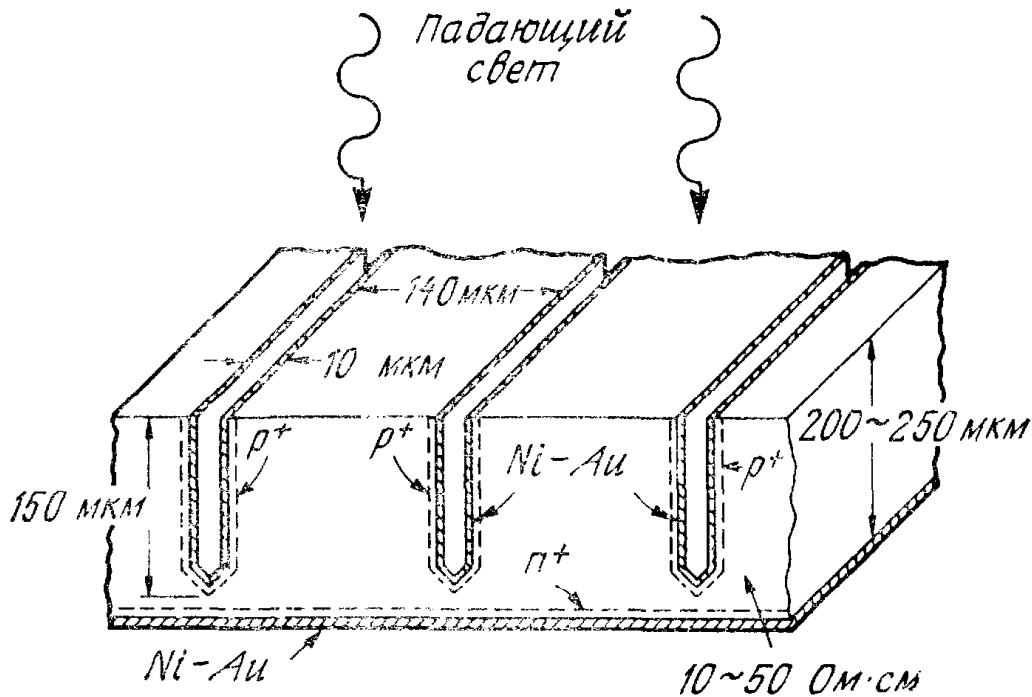


Рис. 21. Солнечный элемент с вертикальными переходами, имеющий низкое последовательное сопротивление [26].

схематически представлен на рис. 21. Диффузионные и металлические контакты расположены в глубоко вытравленных канавках, перпендикулярных поверхности и созданных благодаря анизотропному травлению поверхности кремния $\langle 110 \rangle$. Расстояние между канавками равно 140 мкм, что составляет величину того же порядка, что и диффузионная длина неосновных носителей. В этом случае диффузия с лицевой поверхности не проводится, а доля затененной поверхности благодаря небольшой ширине канавок (10 мкм) не превышает 7 %. Контактные полосы к каждому выводу и просветляющее покрытие на этом рисунке не показаны.

В солнечном элементе с вертикальными переходами ток протекает непосредственно через диффузионные p^+ -области и металлический слой, расположенные на стенках канавок, и выводится через контактные полосы, лежащие на боковых поверхностях элемента. Эти полосы соединяют все элементы параллельно, вследствие чего последовательное сопротивление прибора оказывается очень малым. На таких элементах в условиях АМ1 экспериментально были получены следующие результаты: $V_{\text{хх}} = 0,59$ В, $J_{\text{кз}} = 33$ мА/см², фактор заполнения $FF = 0,80$, к. п. д. = 15,6 %.

Многие из названных выше конструкций были предложены специально для использования в условиях высокой интенсивности излучения. Поэтому их рабочие характеристики мы обсудим в разд. 14.5.

14.4. СОЛНЕЧНЫЕ ЭЛЕМЕНТЫ С ГЕТЕРОПЕРЕХОДАМИ. ПОВЕРХНОСТНЫЕ И ТОНКОПЛЕНОЧНЫЕ СОЛНЕЧНЫЕ ЭЛЕМЕНТЫ

14.4.1. Солнечные элементы с гетеропереходами

Гетеропереходы представляют собой переходы, образующиеся при контакте двух полупроводников с различными энергетическими положениями запрещенной зоны (гл. 2). Типичная энергетическая диаграмма n — p -гетероперехода, находящегося в состоянии термодинамического равновесия, показана на рис. 22. Фотоны с энергией, меньшей E_{g1} , но большей E_{g2} , будут проходить через слой первого полупроводника, который играет роль оптического окна, и поглощаться во втором полупроводнике. Носители, генерируемые излучением внутри обедненного слоя и в электронейтральном объеме полупроводника в пределах диффузионной длины от перехода, будут коллектироваться переходом подобно тому, как это имеет место в солнечных элементах с n — p -гомопереходами. Фотоны с энергией, большей E_{g1} , поглощаются в первом полупроводнике, и переход будет коллектировать носители, генерируемые этим излучением на расстоянии от перехода, не превышающем диффузионную длину, либо непосредственно в области пространственного заряда перехода.

Преимущества солнечных элементов с гетеропереходами перед обычными солнечными элементами с p — n -переходами состоят в следующем [27]: 1) в увеличении спектрального отклика в коротковолновом диапазоне при условии, что энергия E_{g1} достаточно велика и фотоны с высокой энергией поглощаются в обедненном слое второго полупроводника; 2) в понижении последовательного сопротивления при условии, что первый полупроводник можно сильно легировать, не ухудшая при этом условия прохождения света через него; 3) в высокой радиационной стой-

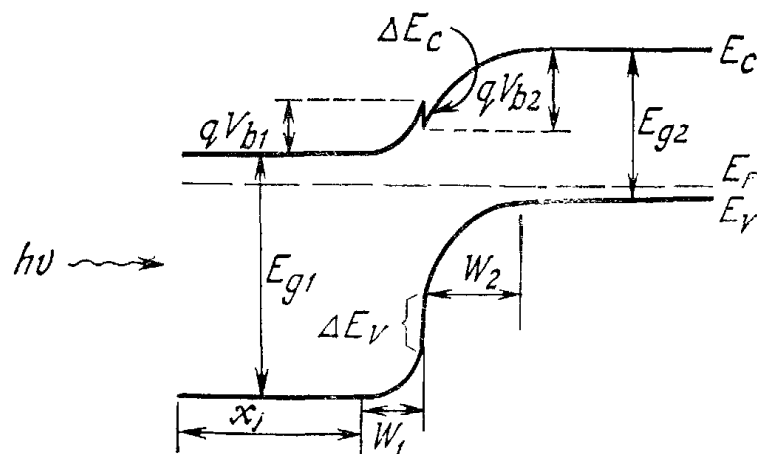


Рис. 22. Диаграмма энергетических зон n — p -гетеропереходов в состоянии термодинамического равновесия.

кости, если первый слой полупроводника достаточно толстый и полупроводник имеет широкую запрещенную зону.

Выражения для фототока в гетеропереходах весьма близки к выражениям для фототока в гомопереходах [3]. В случае n — p -гетероперехода плотность дырочного фототока в первом полупроводнике определяется выражением (22), в котором α следует заменить на α_1 и L_p на L_{p1} , где α_1 и L_{p1} — соответственно коэффициент поглощения и диффузионная длина в первом полупроводнике. Плотность электронного тока определяется формулой (26), в которой следует заменить α на α_2 , L_n на L_{n2} и $\alpha(x_i + W)$ на $[\alpha_1(x_j + W_1) + \alpha_2 W_2]$, где α_2 и L_{n2} — соответствующие параметры второго полупроводника, а W_1 и W_2 — толщины обедненного слоя в этих двух полупроводниках. Фототок обедненных слоев равен

$$J_{dr} = qF(1 - R) [e^{-\alpha_1 x_j} (1 - e^{-\alpha_1 W_1}) + e^{-\alpha_1 (W_1 + x_j)} (1 - e^{-\alpha_2 W_2})]. \quad (40)$$

Рассмотренные выше выражения для фототока получены при следующих предположениях: 1) скачок зон проводимости ΔE_C мал (если речь идет о p — n -гетеропереходах, должна быть мала и величина ΔE_V) и не мешает протеканию неосновных носителей из второго полупроводника через переход; 2) решетки обоих полупроводников хорошо согласуются между собой, вследствие чего плотность поверхностных состояний в гетеропереходе оказывается малой и не снижает время жизни как внутри, так и около обедненного слоя. Спектральный отклик гетеропереходов описывается выражением, подобным формуле (29). Его длинноволновая граница определяется величиной E_{g2} , а отклик в коротковолновой части спектра зависит от ширины запрещенной зоны и толщины первого полупроводника. Если E_{g1} существенно больше E_{g2} и рекомбинация на лицевой поверхности и поверхности гетероперехода пренебрежимо мала, то форма спектрального отклика близка к кривым, приведенным на рис. 11, б, при $S_p < 10^4$ см/с. И наоборот, увеличение ΔE_C и плотности состояний на поверхности гетероперехода приведет к снижению спектрального отклика.

В качестве производного от слова «гетеропереход» используется термин «солнечный элемент с лицевой гетероповерхностью» для обозначения элементов с p — n -гомопереходом и добавленным к нему слоем полупроводника с большей шириной запрещенной зоны. На рис. 23 (вставка) показана зонная диаграмма элемента p -Ga_{1-x}Al_xAs/ p -GaAs/ n -GaAs. Широкозонный полупроводник здесь используется в качестве оптического окна, пропускающего фотоны с энергией, меньшей E_{g1} . Фотоны с энергиями в диапазоне $E_{g2} - E_{g1}$ создают носители в гомопереходе. Если коэффициент поглощения света в узкозонном полупроводнике большой, генерация носителей будет происходить в обедненном

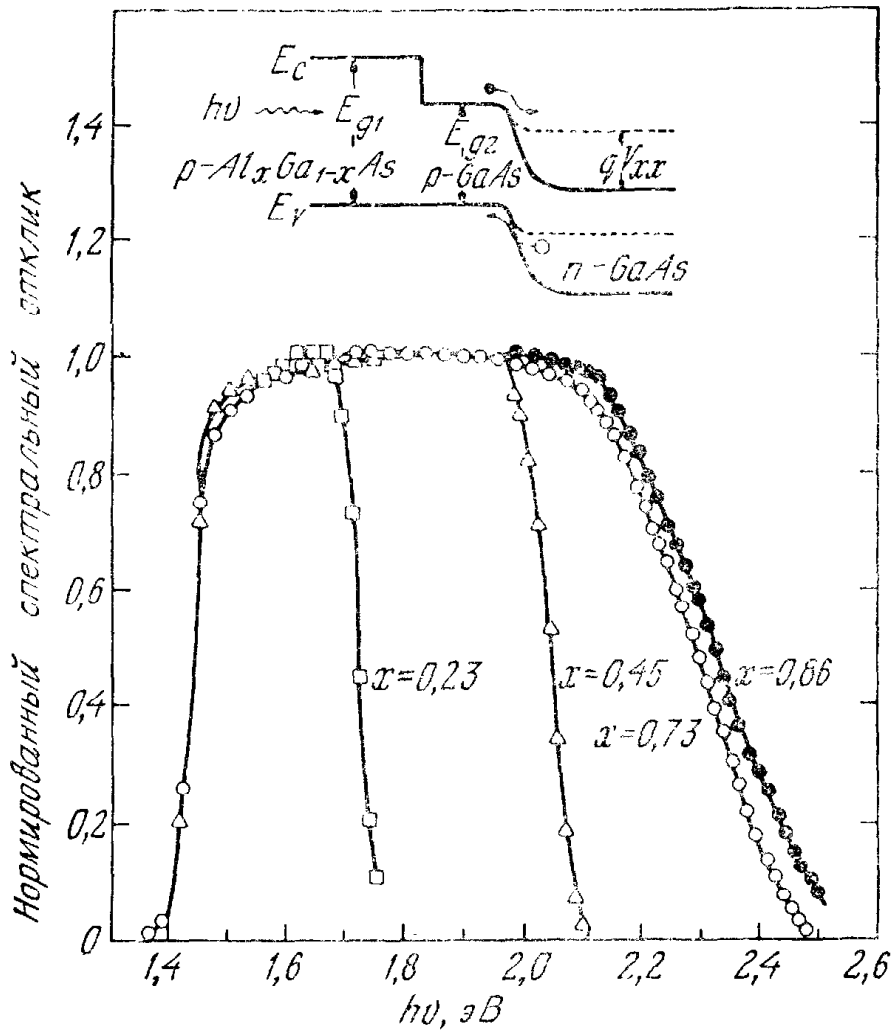


Рис. 23. Нормированный спектральный отклик нескольких солнечных элементов AlGaAs/GaAs различного состава. На вставке показана зонная диаграмма солнечного элемента с лицевой гетеропереходной поверхностью [28].

слое или в близлежащей области и эффективность собирания носителей окажется высокой. Если полупроводниковый материал лицевой гетеропереходной поверхности имеет непрямую запрещенную зону (как, например, $\text{Ga}_{1-x}\text{Al}_x\text{As}$ при $x > 0,4$) и его слой достаточно тонкий, то многие фотоны с энергией $h\nu > E_{g1}$ будут проходить через этот слой и генерировать носители непосредственно в узкозонном полупроводнике. На рис. 23 приведены нормированные зависимости спектрального отклика для нескольких солнечных элементов из $\text{Ga}_{1-x}\text{Al}_x\text{As}-\text{GaAs}$, имеющих одинаковые уровни легирования и одинаковую глубину перехода. При увеличении x ширина запрещенной зоны E_{g1} увеличивается, поэтому возрастает спектральный отклик в коротковолновой области спектра.

Интересным примером солнечного элемента с гетеропереходом является гетеропереход проводящее стекло — полупроводник. К проводящим стеклам относятся окисные полупроводники, такие, как окисел индия (In_2O_3 с $E_g = 3,5$ эВ и электронным сродством $\chi = 4,45$ эВ), окисел олова (SnO_2 с $E_g = 3,5$ эВ и электронным сродством $\chi = 4,8$ эВ) и окисел сплава индия и

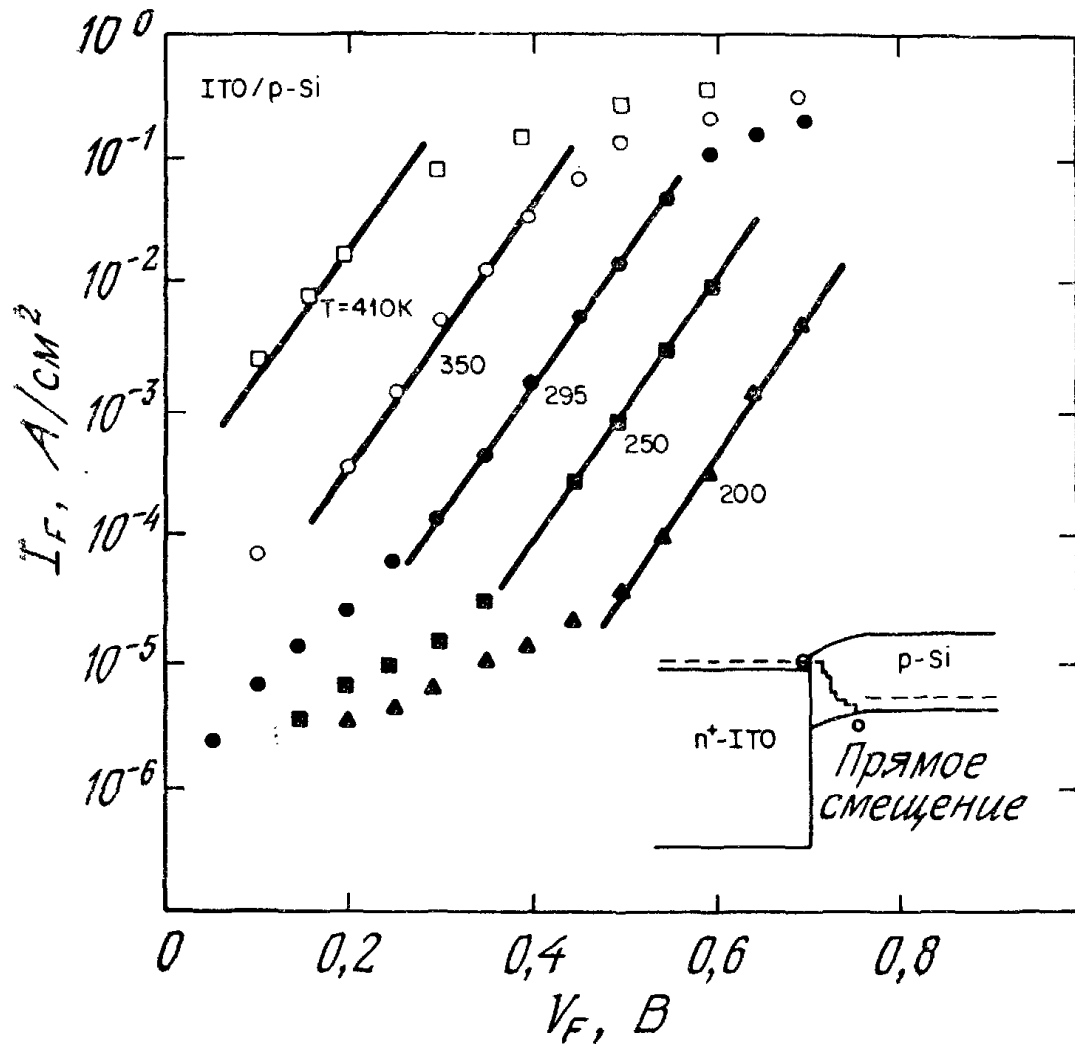


Рис. 24. Вольт-амперные характеристики гетероперехода ITO — Si. На вставке показана зонная диаграмма при прямом смещении [29].

олова (ITO, смесь In_2O_3 и SnO_2 с $E_g = 3,7$ эВ и электронным сродством χ от 4,2 до 4,5 эВ). Эти окисные полупроводники в виде тонких пленок обладают замечательным сочетанием хорошей электрической проводимости и высокой оптической прозрачности. Поэтому они играют роль части гетероперехода и просветляющего покрытия.

На рис. 24 (вставка) показана зонная диаграмма солнечного элемента ITO—Si [29]. Верхним слоем служит окисел ITO n -типа. Толщина слоя равна 4000 \AA , а его удельное сопротивление составляет $5 \cdot 10^{-4} \text{ Ом}\cdot\text{см}$. В качестве подложки выбран кремний p -типа с удельным сопротивлением $2 \text{ Ом}\cdot\text{см}$. Все кривые на рис. 24 при плотностях тока порядка $1 \text{ мА}/\text{см}^2$ параллельны друг другу. Наклон $d(\ln J)/dV$ составляет $\sim 24 \text{ В}^{-1}$ и не зависит от температуры. Такой наклон соответствует многоступенчатому туннелированию внутри гетероперехода. Значения к. п. д. лежат в диапазоне 12—15%. Для солнечных элементов n -ITO/ p -InP в условиях AM2 к. п. д. = 14% [30].

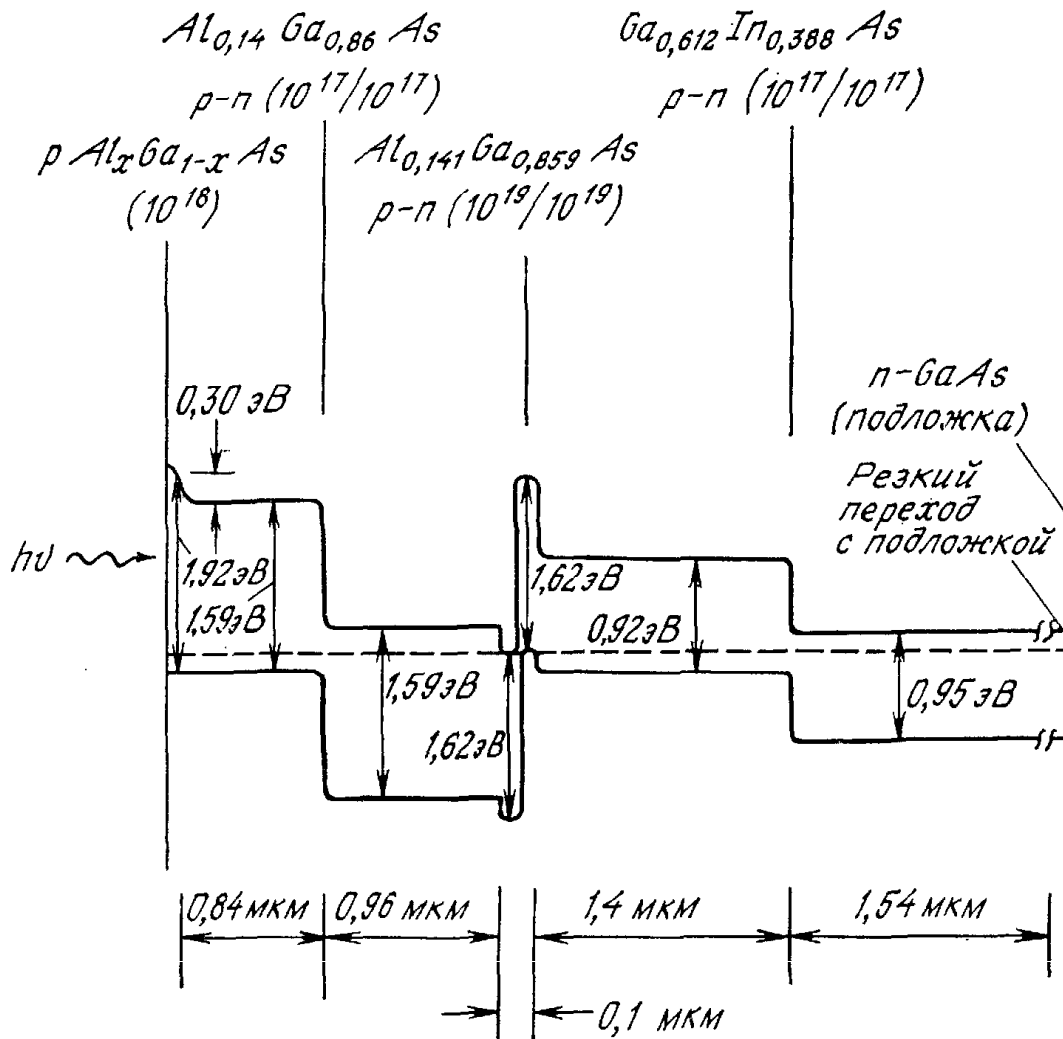


Рис. 25. Идеализированная диаграмма энергетических зон каскадного солнечного элемента с двумя переходами [31].

Для получения высокой эффективности преобразования был предложен каскадный солнечный элемент с гетеропереходами [31]. Идеализированная зонная диаграмма этого элемента показана на рис. 25. Прибор состоит из широкозонного ($E_g = 1,59$) и узкозонного ($E_g = 0,95$ эВ) элементов, последовательно соединенных туннельным диодом на гетеропереходе. Туннельный диод создается во время единого цикла изготовления такой монокристаллической структуры. На рисунке также показана лицевая гетероповерхность, которая служит оптическим окном и позволяет снизить потери от поверхностной рекомбинации. Свет, который проходит первый элемент, не поглощаясь в нем, не поглотится также в сверхтонком туннельном диоде, а приведет к генерации и коллектированию носителей в узкозонном элементе. Оптимизировав соотношение между запрещенными зонами этих двух элементов, можно уравнивать значения их рабочих токов. При этом предельное теоретическое значение к. п. д. в условиях AM1,5 при комнатной температуре оказывается выше 30 %.

14.4.2. Солнечные элементы на барьерах Шоттки и МДП-структурах

Основные характеристики диодов с барьерами Шоттки рассматривались в гл. 5. На рис. 26 представлена диаграмма энергетических зон освещенного солнечного элемента с барьером Шоттки. При этом слой металла должен быть достаточно тонким, чтобы основная доля света достигала полупроводника. Можно выделить три компонента фототока. Одна из них обусловлена поглощением в металле фотонов с энергией $h\nu \rightarrow q\phi_B$ ($q\phi_B$ — высота барьера), что вызывает возбуждение дырок через барьер в полупроводник (эта компонента на рис. 26 обозначена цифрой 1). Как уже говорилось в гл. 5, этот процесс используется для фотоэлектрических измерений высоты барьера. Попадающий в полупроводник коротковолновый свет поглощается главным образом в обедненном слое (соответствующий фототок на рис. 26 обозначен цифрой 2). Длинноволновый свет, поглощаясь в нейтральном объеме полупроводника, создает электронно-дырочные пары; затем электроны, так же как и в случае обычного p — n -перехода, диффундируют к краю обедненного слоя, где происходит их коллектирование (этот фототок на рис. 26 обозначен цифрой 3). В условиях, типичных для работы солнечных элементов, возбуждение светом носителей из металла в полупроводник составляет менее 1 % полного фототока, и поэтому этим процессом можно пренебречь.

К преимуществам солнечных элементов с барьерами Шоттки относится: 1) изготовление таких элементов при низких температурах, поскольку отпадает необходимость в проведении высокотемпературной операции — диффузии; 2) применение данной технологии при создании поликристаллических и тонкопленочных солнечных элементов; 3) высокая радиационная стойкость элементов, поскольку вблизи их поверхности существует сильное

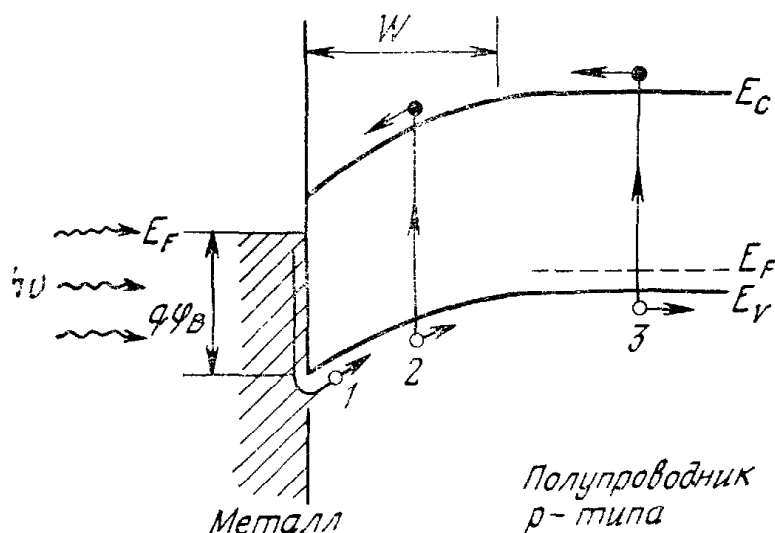


Рис. 26. Диаграмма энергетических зон освещенного солнечного элемента с барьером Шоттки.

электрическое поле; 4) большой выходной ток и хороший спектральный отклик, что также обусловлено непосредственным примыканием обедненного слоя к поверхности полупроводника, вследствие чего ослабляется негативное влияние малых времен жизни и высокой скорости поверхностной рекомбинации.

Две основные компоненты спектрального отклика и фототока связаны с генерацией носителей в обедненном слое и в электронной базовой области. Коллектирование носителей в обедненном слое происходит так же, как и в p — n -переходе. Сильное поле в обедненном слое выносит из него генерируемые светом носители еще до того, как они успевают рекомбинировать, вследствие чего фототок оказывается равным

$$J_{ar} = qT(\lambda)F(\lambda)\{1 - \exp(-\alpha W)\}, \quad (41)$$

где $T(\lambda)$ — коэффициент пропускания металлом монохроматического света с длиной волны λ . Фототок, создаваемый генерацией носителей в базовой области, описывается выражением, идентичным формуле (26), в которой величина $(1 - R)$ должна быть заменена на $T(\lambda)$, а $\alpha(x_i + W)$ — на αW . Если тыловой контакт является омическим и толщина прибора гораздо больше диффузионной длины $H' \gg L_p$, выражение для фототока базовой области можно упростить и представить в виде

$$J_n = qT(\lambda)F(\lambda)\{\alpha L_n/(\alpha L_n + 1)\} \exp(-\alpha W). \quad (42)$$

Полный фототок равен сумме выражений (41) и (42). Видно, что для увеличения фототока следует повышать коэффициент пропускания T и диффузионную длину L_n . Спектральный отклик элемента с идеальным барьером Шоттки аналогичен кривым, приведенным на рис. 11, б при $S_n < 10^4$ см/с. Однако его величина при любой заданной энергии фотона оказывается несколько меньше за счет отражения и поглощения света металлической пленкой. Коэффициент пропускания света золотыми пленками (толщиной 10—100 Å) с просветляющим покрытием может достигать 90—95 %.

Вольт-амперные характеристики освещенного солнечного элемента с барьером Шоттки имеют вид

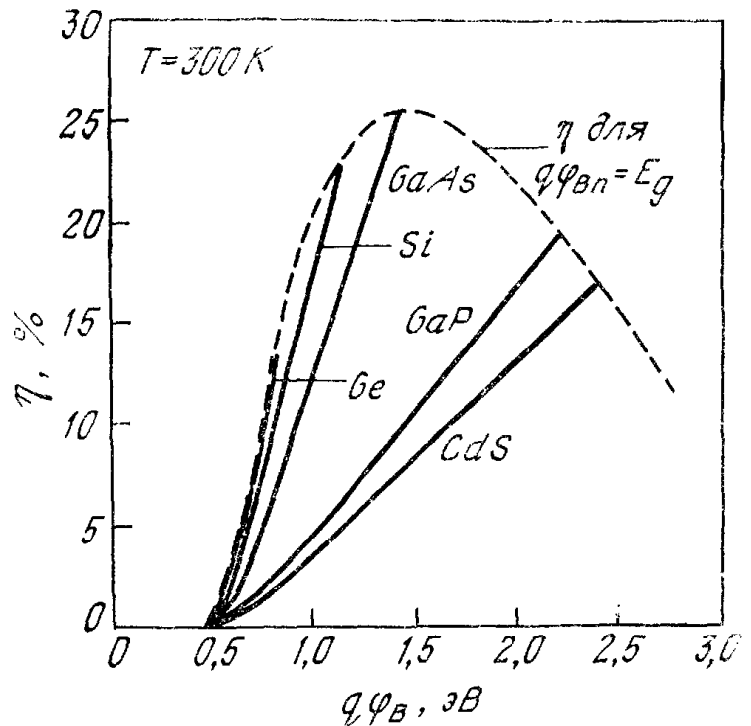
$$I = I_s(e^{qV/nkT} - 1) - I_L, \quad (43)$$

где

$$I_s = AA^{**}T^2 \exp(-q\varphi_B/kT), \quad (44)$$

n — фактор идеальности, A — площадь, A^{**} — эффективная постоянная Ричардсона (гл. 5) и $q\varphi_B$ — высота барьера. Используя выражение для эффективности преобразования (11) и формулы (43) и (44), можно вычислить зависимость к. п. д. от высоты барьера для данного полупроводника. Такие идеальные зависимости (не учитывается отражение света и резистивные потери и предполагается, что фактор идеальности равен 1) приведены

Рис. 27. Зависимость эффективности преобразования от высоты барьера. Огибающая соответствует максимальным значениям к. п. д., вычисленным при $q\varphi_B = E_g$ [32].



на рис. 27 [32]. Видно, что к. п. д. возрастает с увеличением высоты барьера и в предельном случае $q\varphi_B$ (макс) = E_g его максимальное значение составляет $\sim 25\%$. Показанные на рис. 27 предельные значения к. п. д. для различных полупроводников сравнимы с предельными значениями к. п. д. в случае гомопереходов.

Чтобы получить большую высоту барьера Шоттки, обычно для полупроводников n -типа используют металлы с высокой работой выхода, а для полупроводников p -типа — металлы с низкой работой выхода. Например, для Si p -типа используется двухслойный электрод (слой меди толщиной 50 \AA и слой хрома толщиной 50 \AA), в котором хром служит для создания барьера Шоттки, а медь используется в качестве защитного покрытия [33]. Для большинства систем металл — полупроводник, создаваемых на однородных подложках, максимальная высота барьера составляет $\sim 2/3E_g$. Однако отношение высоты барьера к ширине запрещенной зоны может быть увеличено, если между металлом и полупроводниковой подложкой ввести сильнолегированный тонкий полупроводниковый слой (толщиной 100 \AA и с противоположным по отношению к подложке типом легирования). На рис. 28 (вставка) показана зонная диаграмма барьера Шоттки в системе металл — p^+ — n^- -полупроводник [34]. Тогда эффективная высота барьера равна

$$q\varphi_B = q\varphi_m - q\chi + qV_{pn}, \quad (45)$$

где $q\varphi_m$ — работа выхода из металла, χ — электронное сродство полупроводника, V_{pn} — высота максимума потенциала в полупроводнике, равная

$$V_{pn} = \left(\frac{q}{2\varepsilon_s N_A} \right) (N_A W_p - N_D W_n)^2. \quad (46)$$

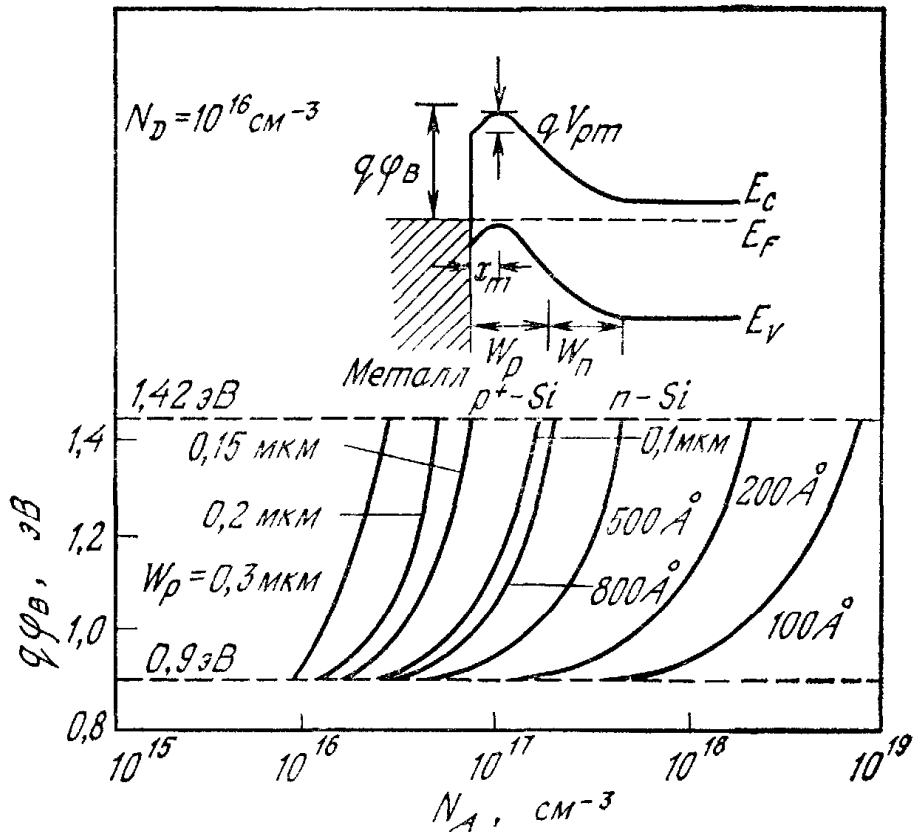


Рис. 28. Зависимость высоты барьера от уровня легирования и толщины p -слоя в солнечных элементах с барьерами Шоттки Au — GaAs. На вставке показана диаграмма энергетических зон элемента со структурой металл — p^+ — n -полупроводник [34].

Эти выражения получены в предположении $N_A W_p \gg N_D W_n$, что соответствует полной ионизации тонкого p -слоя и существованию максимума потенциала внутри этого слоя. При перестановке n - и p -слоев получается комплементарный прибор со структурой металл — n^+ — p -полупроводник. На рис. 28 приведены также вычисленные зависимости высоты барьера от N_A и W_p в GaAs при $N_D = 10^{16}$ см⁻³. Сильнолегированный приповерхностный слой можно создавать с помощью ионной имплантации и молекулярно-лучевой эпитаксии. Другим методом получения такого слоя будет металлургический. Высота барьера в контакте Al—Si (n -типа) при термообработке (< 580 °C) возрастает от 0,68 до 0,9 эВ. При этом образуется тонкий рекристаллизованный p^+ -слой (~ 100 Å), содержащий $\approx 10^{19}$ см⁻³ акцепторов (Al) [35], и зонная диаграмма приобретает вид, показанный на рис. 28.

В солнечных элементах на МДП-структурах между металлом и полупроводниковой подложкой создается тонкий изолирующий слой. Преимущество солнечных элементов на МДП-структурах связано с тем, что в них имеется сильное электрическое поле вблизи поверхности полупроводника, которое направлено так, что помогает коллектированию носителей, создаваемых коротковолновым светом; кроме того, в активной области таких элемен-

тов нет кристаллических дефектов, возникающих при диффузии примесей и присущих солнечным элементам с диффузионными $p-n$ -переходами. В туннельных МДП-структурах выражение для плотности тока насыщения отличается от соответствующего выражения для барьеров Шоттки дополнительным туннельным множителем (гл. 9)

$$J_s = A^{**} T^2 \exp(-q\phi_B/kT) \exp[-(q\phi_T)^{1/2} \delta], \quad (47)$$

где $q\phi_T$ — средняя высота барьера в изолирующем слое (мкм) и δ — толщина изолирующего слоя (Å). Подставляя в уравнение (43) $V = V_{xx}$ и $J = 0$, получаем

$$V_{xx} = \frac{nkT}{q} \left[\ln \left(\frac{J_L}{A^{**} T^2} \right) + \frac{q\phi_B}{kT} + (q\phi_T)^{1/2} \delta \right]. \quad (48)$$

Из формулы (48) видно, что в солнечных элементах на МДП-структурах V_{xx} растет с увеличением δ . В то же время при увеличении толщины диэлектрика δ уменьшается ток короткого замыкания, что может приводить к уменьшению эффективности преобразования. Показано [36], что оптимальная толщина диэлектрика в системе металл — SiO_2 — Si составляет $\sim 20 \text{ \AA}$.

В рассмотренных выше солнечных элементах на МДП-структурах вся поверхность элемента покрыта очень тонким слоем металла. На рис. 29 показана новая конструкция элемента [37], в которой используется гребенка в виде МДП-структуры с толстым металлом. Между отдельными полосками полупроводник покрыт прозрачным диэлектрическим слоем. Слой SiO_2 толщиной 1000 \AA служит одновременно и прозрачным диэлектрическим слоем, и просветляющим покрытием. Эта конструкция отличается от гетероперехода ITO-Si , поскольку SiO_2 — диэлектрик и весь ток протекает через зубцы гребенчатой туннельной МДП-структуры. Если в окисле имеется положительный встроенный заряд, то вблизи поверхности полупроводника будет образовываться помимо обедненного слоя также и инверсионный слой. Генерируемые светом неосновные носители (электроны в случае подложки p -типа) будут коллектироваться из обедненного

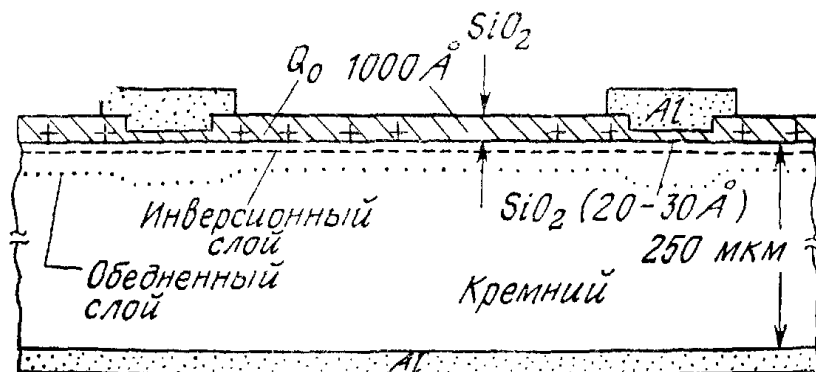


Рис. 29. Поперечное сечение солнечного элемента с гребенчатой МДП-структурой [37].

и инверсионного слоев с помощью гребенчатой туннельной МДП-структуры со слоем SiO_2 толщиной 20—30 Å. Кроме того, инверсионный слой будет экранировать центры поверхностной рекомбинации. На таких структурах в условиях АМІ к. п. д. достигает 18 % [49, 50]. Поскольку при создании таких структур не надо проводить диффузию, а выращивание окисла возможно и при низких температурах, предполагается, что именно такие структуры, изготавливаемые на поликристаллических и аморфных подложках, окажутся экономически наиболее выгодными источниками энергии.

14.4.3. Тонкопленочные солнечные элементы

В тонкопленочных солнечных элементах активными полупроводниковыми слоями являются поликристаллические или неупорядоченные пленки, которые наносятся или выращиваются на электрически активных или пассивных подложках (таких, как стекло, пластмасса, керамика, металл, графит или металлургический кремний). Тонкие пленки CdS , Si , GaAs , InP , CdTe и т. д. наносятся на подложку с помощью различных методов, таких, как газовая эпитаксия, плазменное распыление и осаждение. Если толщина слоя полупроводника превышает обратный коэффициент поглощения, большая часть света будет поглощаться в полупроводнике; если диффузионная длина превышает толщину пленки, большая часть генерируемых светом носителей может быть коллектирована.

Основное достоинство тонкопленочных солнечных элементов заключается в том, что их стоимость может быть низкой, поскольку технология их изготовления дешевая и при этом используются сравнительно дешевые материалы. К основным недостаткам относятся низкий к. п. д. и постепенная деградация характеристик. Низкий к. п. д. обусловлен частично влиянием границ зерен, а частично низким качеством полупроводникового материала, выращиваемого на различных подложках. Проблема деградации связана с химическими реакциями, протекающими между полупроводником и окружающей атмосферой (кислородом, парами воды).

Тонкопленочный солнечный элемент на CdS схематически показан на рис. 30. Элементы создаются с использованием подложек, покрытых электролитически осажденной медью, поверх которой наносится слой цинка толщиной 0,5 мкм. На подложку, подогретую до 220 °С, осаждается слой CdS толщиной 20 мкм. Реакция между пленкой CdS и ионами меди приводит к образованию слоя Cu_2S толщиной 1000 Å. Поверх слоя наносятся прозрачная контактная гребенка и слой просветляющего покрытия. На рис. 31 показана энергетическая диаграмма элемента Cu_2S —

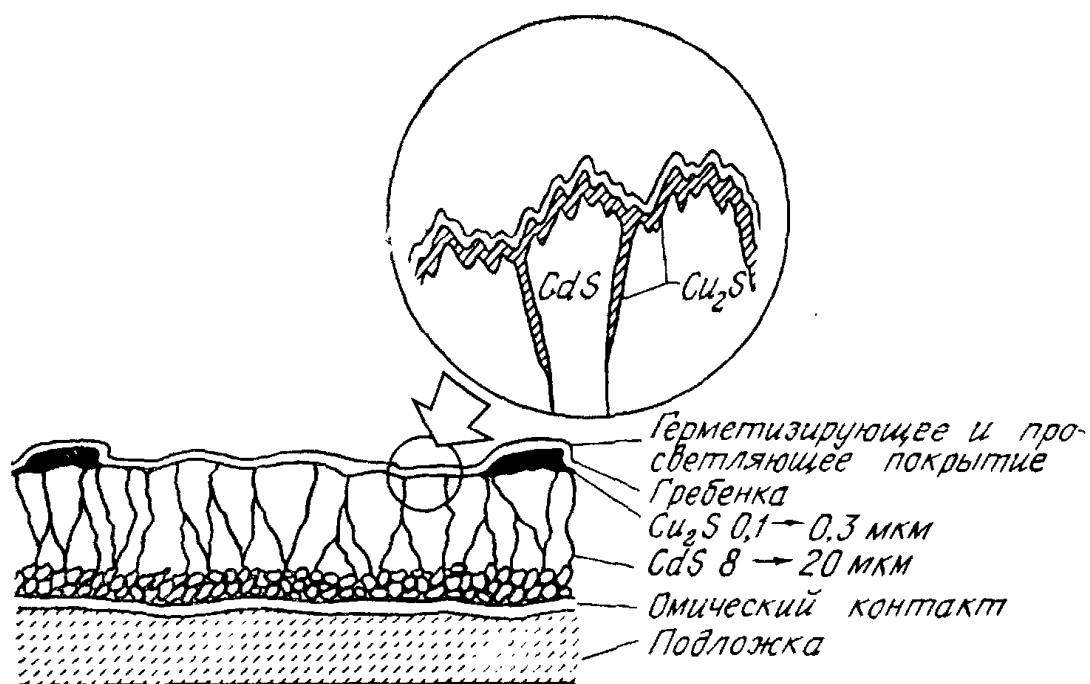


Рис. 30. Тонкопленочный солнечный элемент на CdS [38].

CdS [39]. По существу этот элемент представляет собой гетеропереход с большой плотностью ловушек на границе раздела. При освещении с лицевой стороны излучение поглощается в основном в Cu_2S . Спектральный отклик и фототок ограничиваются высокой скоростью поверхностной рекомбинации, малой диффузионной длиной и интенсивной рекомбинацией на границе раздела. Несмотря на это, к. п. д. такого элемента близок к 10 % [38]. Кроме того, здесь возможен дальнейший прогресс: ожидается, что при замене от 15 до 25 % цинка кадмием вырастет выходное напряжение и к. п. д. окажется выше 14 %.

Основой другого тонкопленочного элемента служит гетеропереход $\text{CuInSe}_2/\text{CdS}$. Его зонная диаграмма показана на рис. 32 (вставка). Ток короткого замыкания и к. п. д. зависят от размера зерен. При увеличении радиуса зерна от 0,3 до 1 мкм и выше к. п. д. возрастает от 4 до 6,6 % [40]. В качестве материалов для создания дешевых солнечных элементов можно рассматривать целый ряд тройных соединений. На рис. 33 представлены некоторые полупроводники групп I—III—VI₂ и II—IV—V₂ [[41], которые пригодны для эффективного фотовольтаического преобразования солнечной энергии.

В качестве материала для тонкопленочных солнечных элементов используется также аморфный кремний (a-Si). Слои толщиной 1—3 мкм выращиваются на стеклянных подложках, покрытых слоем металла или ITO, с помощью разложения силана в высокочастотном разряде. Кристаллический и аморфный кремний сильно различаются: первый имеет непрямую запрещенную зону шириной 1,1 эВ, в то время как характеристики оптического поглощения в гидроденизированной a-Si напоминают характерис-

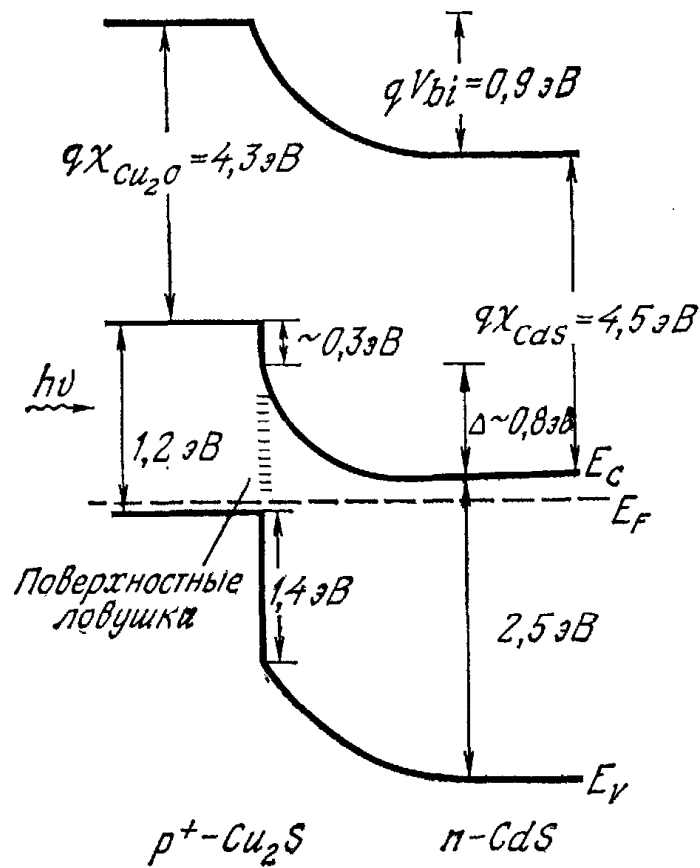


Рис. 31. Диаграмма энергетических зон солнечного элемента $\text{Cu}_2\text{S} - \text{CdS}$ [39].

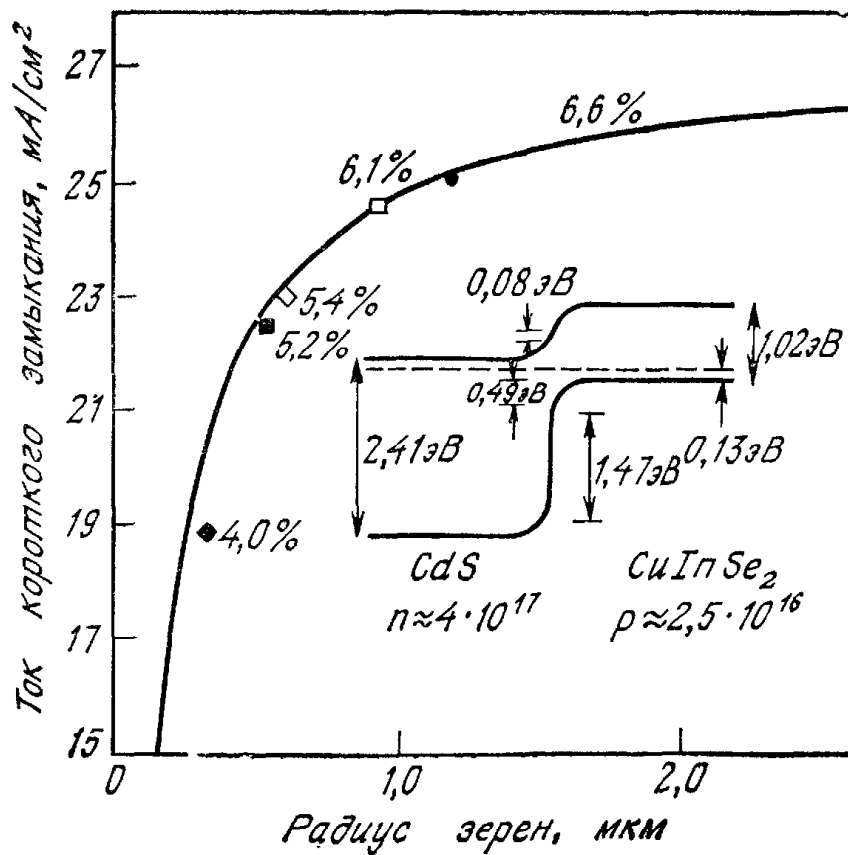
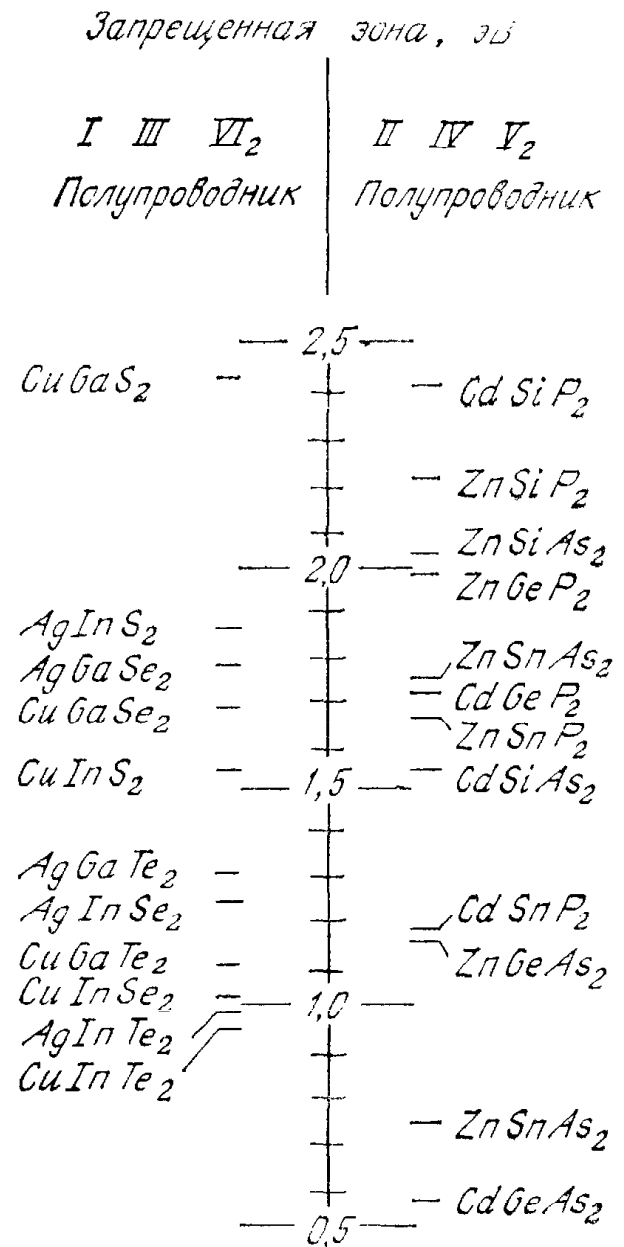


Рис. 32. Зависимость тока короткого замыкания и эффективности преобразования от радиуса зерен в солнечных элементах $\text{CdS}/\text{CuInSe}_2$. На вставке показана диаграмма энергетических зон [40].

Рис. 33. Значение ширины запрещенных зон полупроводников групп I—III—VI₂ и II—IV—V₂, представляющих интерес для фотовольтаического преобразования солнечной энергии [41].



тики кристалла с прямой запрещенной зоной шириной 1,6 эВ (рис. 34). На тонких пленках гидрогенизированного a-Si были созданы солнечные элементы как с *p-n*-переходами, так и с барьерами Шоттки. На рис. 35 схематически представлено несколько диаграмм различных структур солнечных элементов на a-Si [43]. Поскольку в видимой части солнечного спектра коэффициент поглощения в a-Si составляет 10^4 — 10^5 см⁻¹, большинство носителей генерируется на расстоянии, не превышающем 1 мкм от освещенной поверхности. Диффузионная длина не превышает ~1 мкм, а темновое удельное сопротивление оказывается порядка 100 МОм·см. Отсюда получаем, что в *p-i-n*-приборе с толщиной *i*-слоя 1 мкм последовательное сопротивление составляет 10^4 Ом·см. Поэтому даже наилучшие результаты для элемента с барьером Шоттки (прозрачный металлический слой) относительно невысоки (к. п. д. ~6 %).

Для проведения оценки предельных возможностей приборов на неупорядоченных тонких пленках рассмотрим идеализирован-

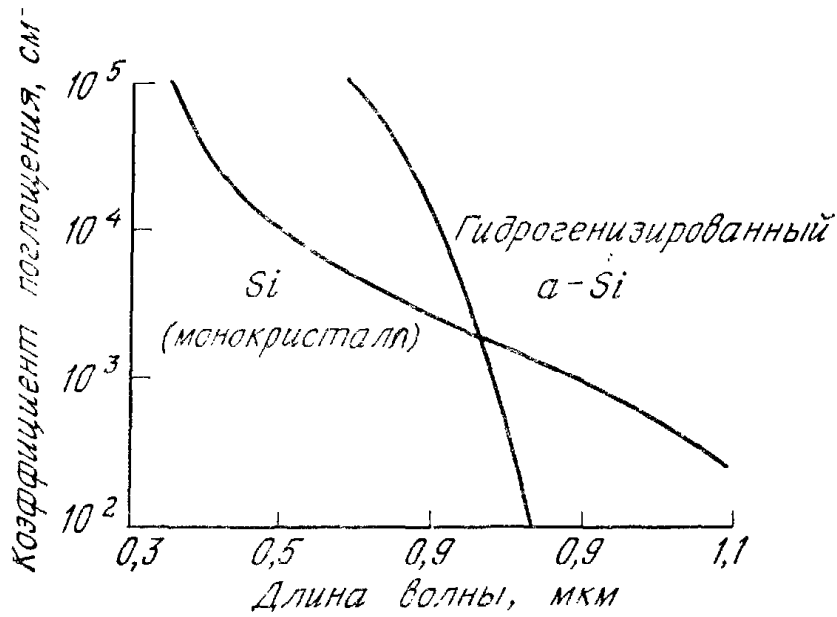


Рис. 34. Зависимость коэффициента поглощения от длины волны для кристаллического и гидрогенизированного аморфного Si [42].

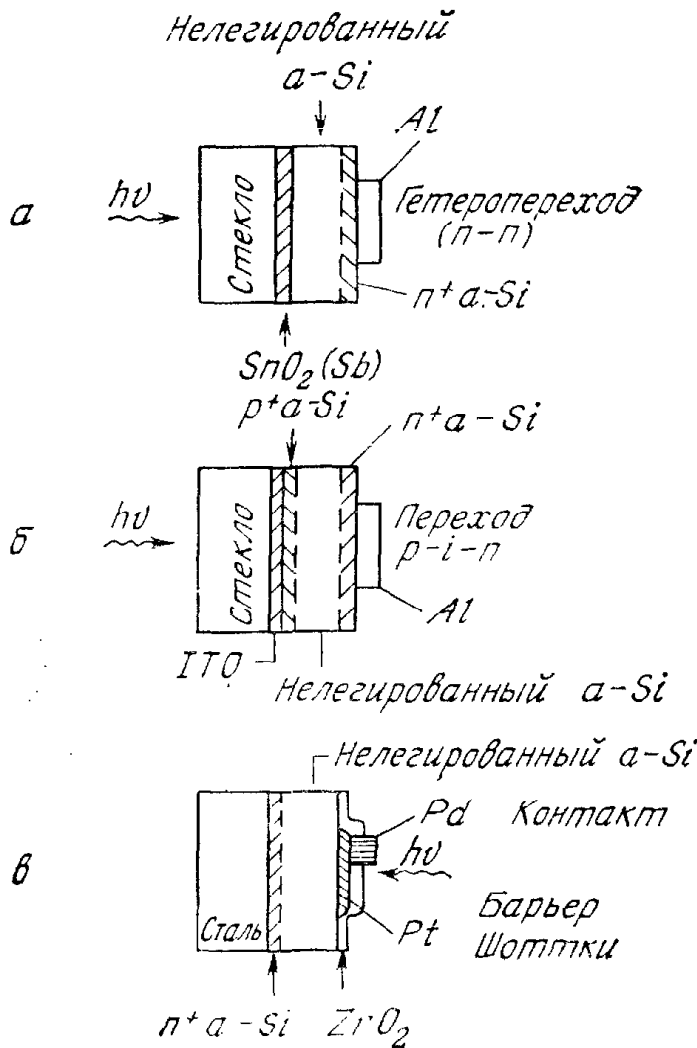


Рис. 35. Схематические диаграммы структур различных солнечных элементов на a-Si [43].

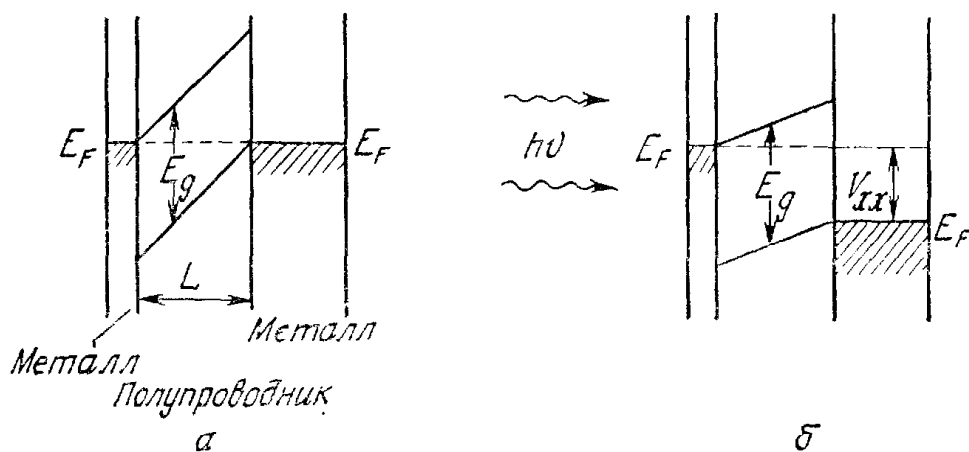


Рис. 36. Идеализированная диаграмма энергетических зон тонкопленочного солнечного элемента в равновесном состоянии (а) и при освещении (б) [13].

ный тонкопленочный солнечный элемент (рис. 36). Зонная диаграмма этого элемента в состоянии термодинамического равновесия показана на рис. 36, а. Элемент состоит из слоя полупроводника толщиной L , заключенного между двумя металлическими электродами с различными работами выхода. Металлы выбраны так, что уровень Ферми в одном из них проходит вблизи края зоны проводимости, а в другом — вблизи края валентной зоны. Один из электродов сделан настолько тонким, что практически не мешает прохождению солнечного излучения. Предполагается, что в прямозонном полупроводнике для фотонов с энергией, превышающей на несколько десятых эВ энергию края поглощения, $\alpha > 10^5 \text{ см}^{-1}$. Диаграмма на рис. 36, б построена в предположении, что свет поглощается в области сильного электрического поля, вследствие чего происходит эффективное разделение электронов и дырок.

Поскольку в таких полупроводниках концентрация ловушек довольно высокая, следует оценить, при какой концентрации ловушек характеристики прибора начнут ухудшаться. В отсутствие заряженных центров электрическое поле, как показано на рис. 36, б, оказывается однородным и равным $\mathcal{E} = E_g/qL$. При толщине $1/\alpha$ ($\sim 0,1 \text{ мкм}$) и $E_g = 1,5 \text{ эВ}$ напряженность поля равна $1,5 \cdot 10^5 \text{ В/см}$. Если в полупроводнике имеются ловушки с концентрацией n_t , плотность ловушечного пространственного заряда будет равна n_c , где $n_c < n_t$. Эти заряженные дефекты изменяют поле на величину $\Delta\mathcal{E} = qn_cL/\epsilon_s$. Положив диэлектрическую постоянную равной 4, получим, что $\Delta\mathcal{E} \ll \mathcal{E}$, если $n_c < 10^{16} \text{ см}^{-3}$. Это означает, что даже при такой высокой концентрации ловушек, как 10^{17} см^{-3} , их заполнение не приводит к существенным изменениям распределения электрического поля в полупроводнике. Следующее условие, которое также должно выполняться, состоит в том, чтобы плотность тока, ограниченного пространственным зарядом, превышала 100 мА/см^2 , т. е. была существенно выше

плотности тока короткого замыкания при солнечном освещении. Из этого условия следует, что при толщине полупроводника 0,1 мкм допустима концентрация ловушек также $\sim 10^{17}$ см $^{-3}$.

Электрическое поле должно быть настолько большим, чтобы время пролета $L/\mu\mathcal{E}$, за которое электроны и дырки уходят из полупроводника, было меньше рекомбинационного времени жизни $(n_t\nu\sigma)^{-1}$, где σ — сечение захвата $\sim 10^{-14}$ см 2 и ν — тепловая скорость $\sim 10^7$ см/с. Это условие удовлетворяется, если

$$\mu > \frac{n_t\nu\sigma L}{\mathcal{E}} = n_t\nu\sigma qL^2/E_g \simeq 1 \text{ см}^2/\text{В}\cdot\text{с}. \quad (49)$$

Выполнение соотношения (49) вполне реально.

Таким образом, из приведенного обсуждения следует, что высокоэффективные солнечные элементы можно создавать из полупроводников с очень высокой плотностью дефектов, если полупроводниковые пленки достаточно тонкие, имеют высокий коэффициент поглощения вблизи края фундаментального поглощения и в них достигаются требуемые значения подвижности.

14.5. КОНЦЕНТРИРОВАНИЕ СВЕТА

Используя зеркала и линзы, можно фокусировать солнечный свет. Оптическое концентрирование излучения представляет собой весьма привлекательный и удобный способ снижения высокой стоимости солнечных элементов, поскольку при этом большая часть площади солнечного элемента заменяется поверхностью концентратора. Кроме того, имеются и другие достоинства: 1) возрастает к. п. д. элемента (см. рис. 7); 2) появляется возможность создания гибридных систем, которые одновременно вырабатывают электрическую энергию и тепло; 3) ослабляется зависимость эффективности работы элемента от температуры окружающей среды.

На рис. 37, а показана стандартная плоско-выпуклая линза, а на рис. 37, б — эквивалентная ей линза Френеля. В концентрирующем модуле, обеспечивающем 500—1000-кратную концентрацию солнечного света (рис. 37, в), солнечный свет, отраженный первым зеркалом, попадает на второе зеркало, которое фокусирует его на солнечном элементе, расположенном в блоке с водяным охлаждением. Экспериментальные результаты, полученные с помощью кремниевого солнечного элемента с вертикальными переходами (рис. 21), смонтированного в концентраторе (рис. 37, в), приведены на рис. 38 [26, 44]. Видно, что рабочие характеристики прибора улучшаются при увеличении интенсивности излучения от эквивалентной 1 солнцу до интенсивности, эквивалентной 1000 солнц. Ток короткого замыкания возрастает линейно с увеличением концентрации излучения. Напряжение холостого хода

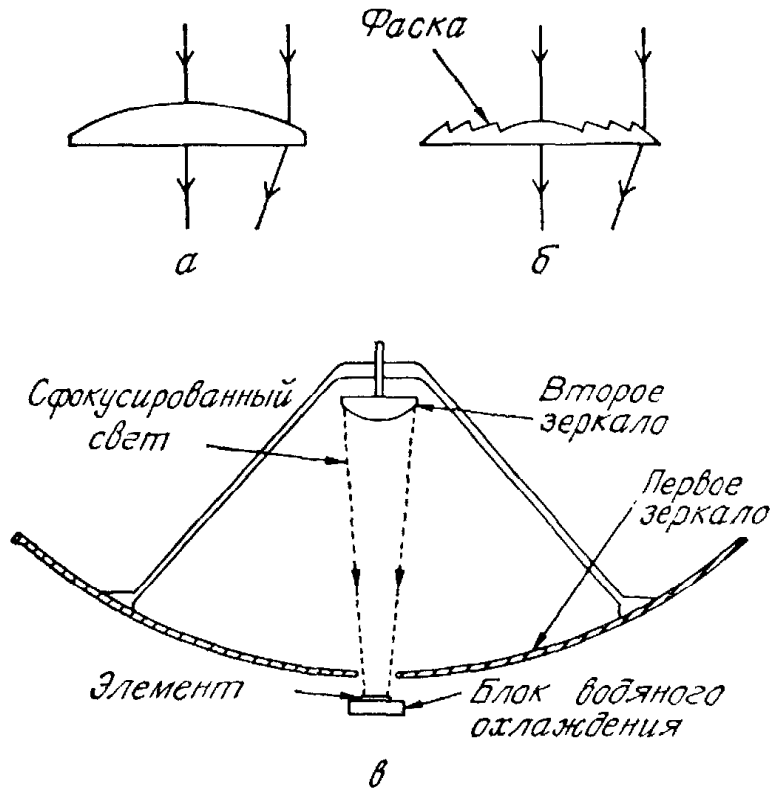


Рис. 37. Концентраторы солнечного света.

а — обычная линза; б — линза Френеля; в — типичный концентрирующий модуль.

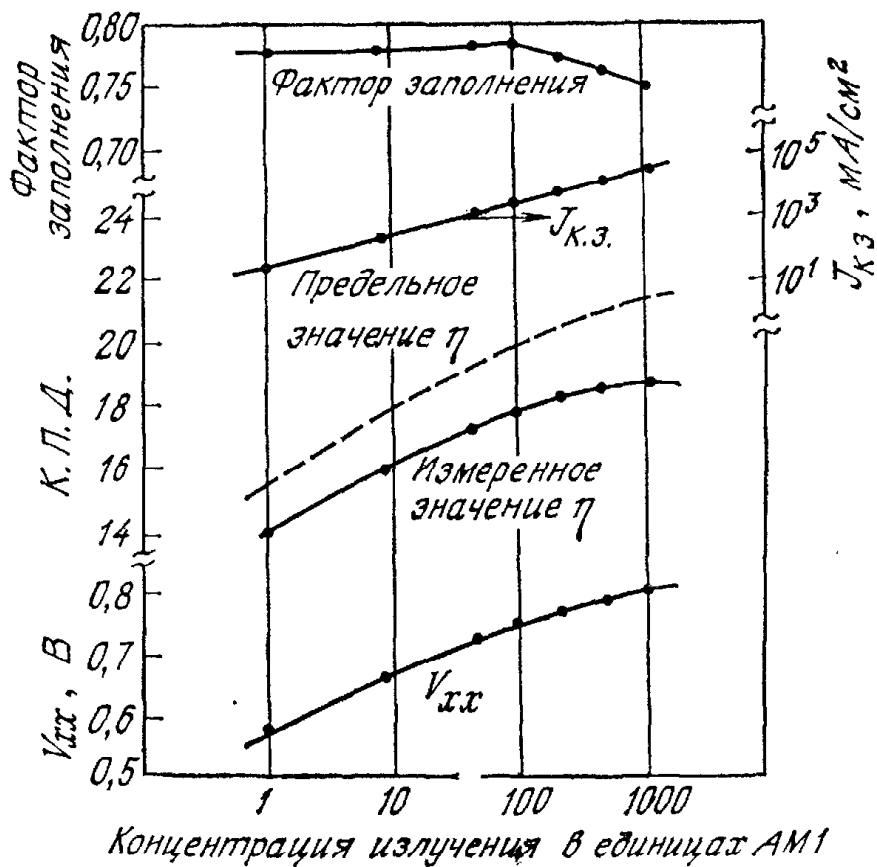


Рис. 38. Зависимость к. п. д., V_{xx} , $J_{кз}$ и фактора заполнения от кратности концентрации солнечного излучения при АМ1 в элементе с вертикальными переходами [44].

возрастает на 0,1 В при увеличении интенсивности излучения в 10 раз, в то время как фактор заполнения изменяется слабо. Коэффициент полезного действия элемента, который является произведением трех указанных выше величин, деленным на мощность излучения, падающего на концентратор, возрастает на $\sim 2\%$ при увеличении концентрации излучения в 10 раз. Считается, что при использовании соответствующих просветляющих покрытий к. п. д. может составлять $\sim 22\%$ при интенсивности излучения, равной интенсивности 1000 солнц. Поэтому один элемент, освещаемый 1000-кратной концентрацией солнечного излучения, обеспечивает выходную мощность, равную мощности 1300 элементов, освещаемых светом одного солнца. Следовательно, метод концентрации света обеспечивает возможность замены дорогостоящих солнечных элементов сравнительно дешевыми материалами концентратора, и его можно реализовать с помощью относительно простого устройства, что позволяет минимизировать стоимость всей системы.

При сильном концентрировании излучения концентрация неравновесных носителей становится сравнимой по порядку величины с концентрацией легирующей примеси в подложке, и таким образом реализуется условие высокого уровня инжекции. При этом плотность тока пропорциональна $\exp(qV/nkT)$ с $n = 2$. Напряжение холостого хода равняется

$$V_{xx} = \frac{2kT}{q} \ln \left(\frac{J_L}{J_s} + 1 \right). \quad (50)$$

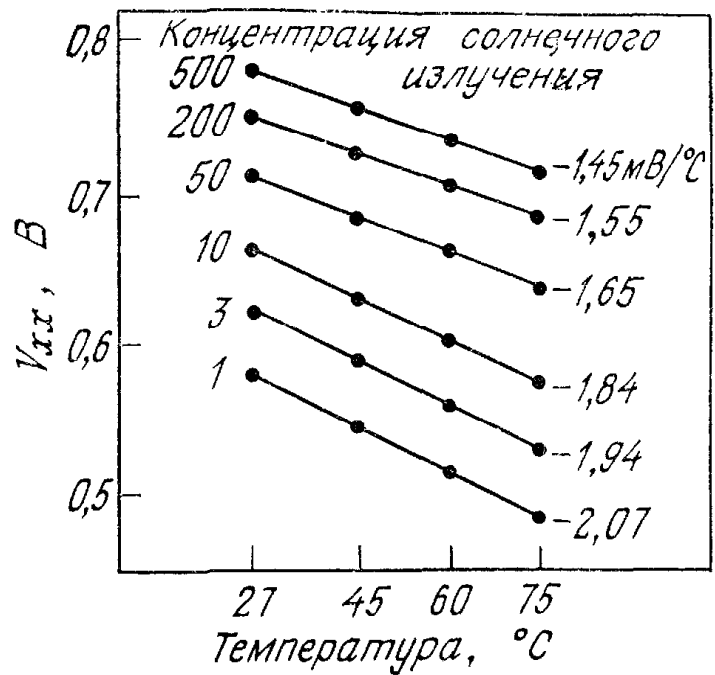
Величину J_s можно представить в виде

$$J_s = \text{const} \left(\frac{T}{T_0} \right)^{3/2} \exp \left[-\frac{E_g(T)}{2kT} \right], \quad (51)$$

где T — рабочая температура, $T_0 = 300$ К. Температурная зависимость V_{xx} при различных уровнях концентрирования приведена на рис. 39 [26]. Температурный коэффициент изменения V_{xx} уменьшается по абсолютной величине от $-2,07$ мВ/°С при солнечном излучении до $-1,45$ мВ/°С при 500-кратной концентрации солнечного излучения. Таким образом, для кремниевых солнечных элементов высокие уровни концентрирования солнечного излучения снижают эффективность потерь, обусловленных работой элемента при повышенных температурах.

При концентрировании излучения хорошие рабочие характеристики могут иметь и солнечные элементы на GaAs с лицевой гетероповерхностью. Поперечное сечение такого элемента схематически показано на рис. 40 (вставка) [45], где также приведена вольт-амперная характеристика этого элемента при интенсивности излучения, эквивалентной излучению 945 солнц в условиях AM1,5. Фактор заполнения этого прибора равен 0,79, напряжение

Рис. 39. Зависимость напряжения холостого хода от температуры прибора при различной кратности концентрации солнечного излучения [26].



холостого хода составляет 1,13 В, максимальная выходная мощность 10,7 Вт и к. п. д. 23,3 % при 50 °С.

Для повышения к. п. д. и выходной мощности можно использовать многокаскадные солнечные элементы (рис. 8 и 25) либо устройства спектрального разложения света. В последнем случае солнечное излучение разлагается на много узких спектральных полос и излучение из каждой полосы преобразуется с помощью элемента, ширина запрещенной зоны которого выбрана наиболее

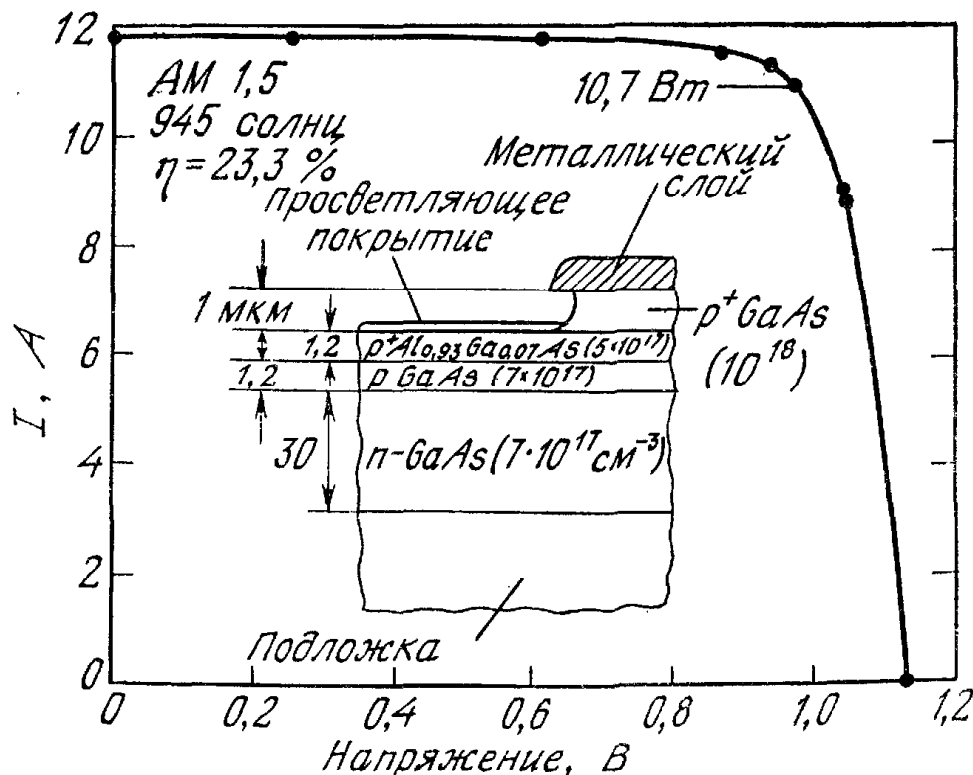


Рис. 40. Измеренная вольт-амперная характеристика солнечного элемента AlGaAs/GaAs при 945-кратной концентрации солнечного излучения. На вставке показано поперечное сечение прибора [45].

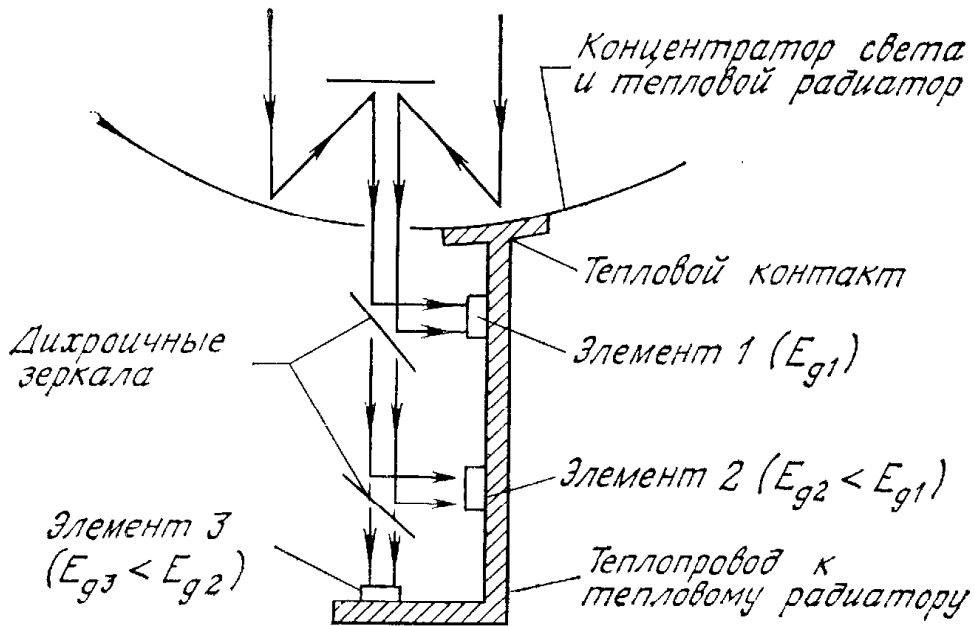


Рис. 41. Возможная конструкция модуля спектрального разложения света, обеспечивающего высокую эффективность преобразования [46].

оптимальной по отношению к спектральному составу данной полосы. Возможная конструкция устройства спектрального разложения света показана на рис. 41 [46]. Дихроичные зеркала разлагают падающий концентрированный свет, отражая фотоны с высокой энергией в элемент 1 и пропуская фотоны с низкой энергией к элементу 2 и далее к элементу 3. При 1000-кратном концентрировании солнечного излучения максимальное значение к. п. д. при делении света на два спектральных диапазона можно получить с помощью описанного выше (рис. 8) графического метода. Это значение составляет $\sim 60\%$. При делении излучения на 3 спектральные полосы максимум к. п. д. близок к 70% , а при делении на 10 полос он составляет $\sim 85\%$. Идеальное значение к. п. д. при разложении света на n спектральных полос выше, чем при последовательном соединении n каскадов с различными ширинами запрещенной зоны. Дело в том, что при разложении спектра плотность тока в каждом элементе оптимизируется независимо. Это обеспечивает максимум к. п. д. всей системы.

На рис. 42 (вставка) показано экспериментально реализованное устройство, состоящее из солнечного элемента $\text{Al}_x\text{Ga}_{1-x}\text{As}$ с $x < 0,24$, кремниевого солнечного элемента и фильтра, осуществляющего разложение спектра. На этом же рисунке приведены вольт-амперные характеристики обоих элементов при 165-кратной концентрации солнечного излучения (AM1,2) [47]. Максимальная выходная мощность составляет 1,44 Вт и 0,95 Вт для элементов на AlGaAs ($E_g = 1,61$ эВ) и Si ($E_g = 1,1$ эВ) соответственно, а значения к. п. д. равны 17,4 и 11,1%. Полный к. п. д. системы равен 28,5% при 30°C .

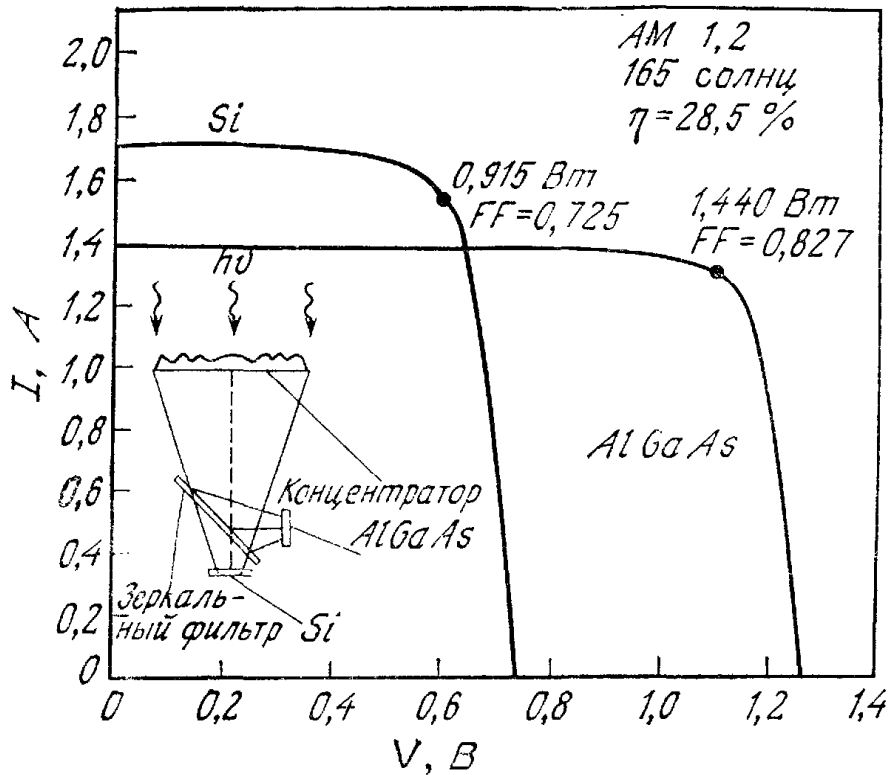


Рис. 42. Вольт-амперные характеристики солнечных элементов на Si и AlGaAs, работающих в устройстве спектрального разложения [47].

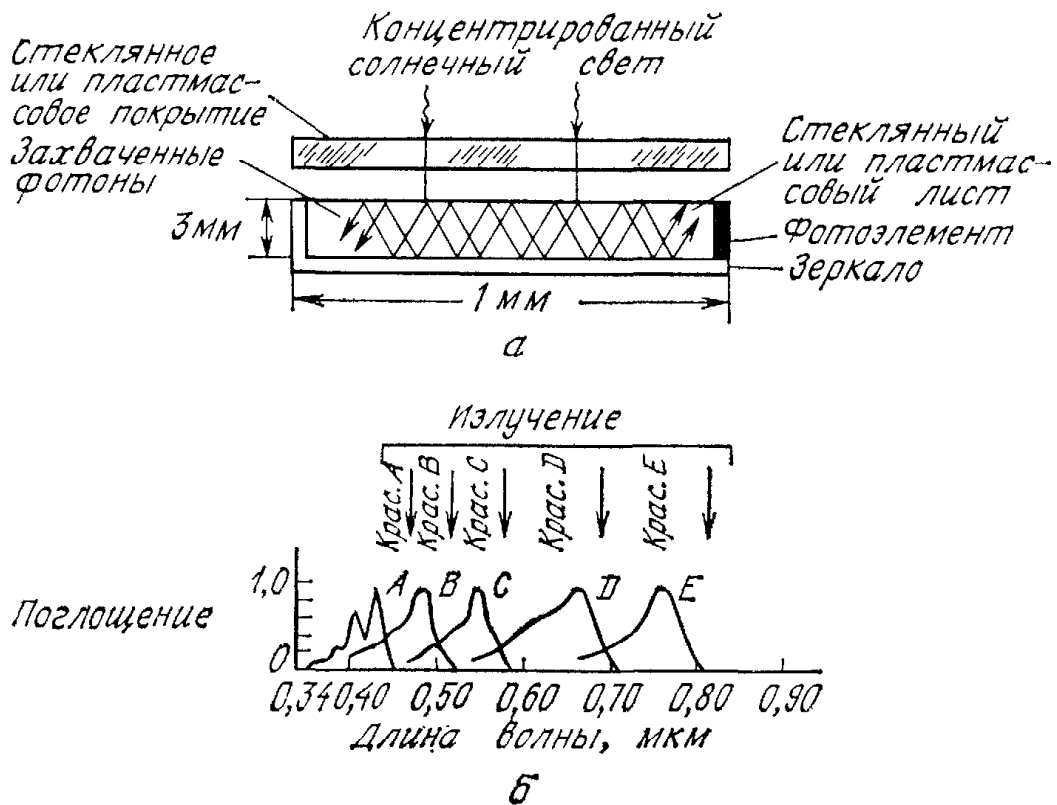


Рис. 43. Накопитель солнечного излучения, в котором используются люминесцентные красители, поглощающие узкие спектральные полосы солнечного света, а затем испускающие фотоны с различными длинами волн [48].

Другой перспективный метод преобразования солнечной энергии состоит в использовании листа стекла или пластмассы, покрытого люминесцирующими красителями, которые поглощают солнечный свет в узких спектральных диапазонах, а затем испускают фотоны разных энергий во многих направлениях (рис. 43). Свет, отражаемый от границ, оказывается «захваченным» внутри листа, поскольку красители уже не могут поглощать его и он в конце концов попадает к краю листа, где расположен солнечный элемент. Для такой системы не нужно устройство слежения за Солнцем, поскольку красители поглощают свет, падающий под любым углом. Теоретическое значение к. п. д. здесь превышает 50 %.

ЛИТЕРАТУРА

1. Chapin D. M., Fuller C. S., Pearson G. L. A New Silicon $p-n$ Junction Photocell for Converting Solar Radiation into Electrical Power, *J. Appl. Phys.*, **25**, 676 (1954).
2. Reynolds D. C., Leies G., Antes L. L., Marbutger R. E. Photovoltaic Effect in Cadmium Sulfide, *Phys. Rev.*, **96**, 533 (1954).
3. Hovel H. J. Solar Cells, in Willardson R. K., Beer A. C. Eds., *Semiconductors and Semimetals*, Vol. 11, Academic, N. Y., 1975; *Photovoltaic Materials and Devices for Terrestrial Applications*, *IEEE Tech. Dig. Int. Electron Device Meet.*, 1979, p. 3.
4. Backus C. E. Ed., *Solar Cells*. IEEE Press, N. Y., 1976.
5. Pulfrey D. L. *Photovoltaic Power Generation*, Van Nostrand Reinhold, N. Y., 1978.
6. Johnston W. D. Jr. *Solar Voltaic Cells*, Dekker N. Y., 1980.
7. Bachmann K. J. Material Aspects of Solar Cells, in Kaldis E., Ed., *Current Topics in Material Science*, Vol. 3, North-Holland, Amsterdam, 1979.
8. Conf. Rec. 13th IEEE Photovoltaic Spec. Conf., IEEE, N. Y., 1978.
9. Conf. Rec. 14th IEEE Photovoltaic Spec. Conf., IEEE, N. Y., 1980.
10. Thekaekara M. P. Data on Incident Solar Energy, *Suppl. Proc. 20th Ann. Meet. Inst. Environ. Sci.*, 1974, p. 21.
11. Henry C. H. Limiting Efficiency of Ideal Single and Multiple Energy Gap Terrestrial Solar Cells, *J. Appl. Phys.*, **51**, 4494 (1980).
12. Prince M. B. Silicon Solar Energy Converters, *J. Appl. Phys.*, **26**, 534 (1955).
13. Principal Conclusions of the American Physical Society Study Group on Solar Photovoltaic Energy Conversion, American Physical Society, N. Y., 1979.
14. Shockley W., Queisser H. J. Detailed Balance Limit of Efficiency of pn Junction Solar Cells, *J. Appl. Phys.*, **32**, 510 (1961).
15. Handy R. J. Theoretical Analysis of the Series Resistance of a Solar Cell, *Solid State Electron.*, **10**, 765 (1967).
16. Hall R. N. Electron-Hole Recombination in Germanium, *Phys. Rev.* **87**, 367 (1952).
17. Shockley W., Read W. T. Statistics of the Recombination of Holes and Electrons, *Phys. Rev.*, **87**, 835 (1952).
18. Wysocki J. J., Rappaport P. Effect of Temperature on Photovoltaic Solar Energy Conversion, *J. Appl. Phys.*, **31**, 571 (1960).
19. Rosenzweig W., Gummel H. K., Smits F. M. Solar Cell Degradation under 1 MeV Electron Bombardment, *Bell Syst. Tech. J.*, **42**, 399 (1963).
20. Wysocki J. J., Rappaport P., Davison E., Hand R., Loferski J. J. Lit-

- hium-Doped Radiation-Resistant Silicon Solar Cells, *Appl. Phys. Lett.*, **3**, 44 (1966).
21. Mandelkorn J., Lamneck J. H. Jr., Simplified Fabrication of Back Surface Electric Field Silicon Cells and Novel Characteristic of Such Cells, Conf. Rec. 9th IEEE Photovoltaic Spec. Conf., IEEE, N. Y., 1972, p. 66.
 22. Lindmayer J., Allison J. F. The Violet Cell: An Improved Silicon Solar Cell, Conf. Rec. 9th IEEE Photovoltaic Spec. Conf., IEEE, N. Y., 1972, p. 83.
 23. Arndt R. A., Allison J. F., Haunos J. G., Meulenberg A. Jr., Optical Properties of the COMSAT Non-Reflective Cell, Conf. Rec. 11th IEEE Photovoltaic Spec. Conf., IEEE, N. Y., 1975, p. 40.
 24. Chappel T. I. The V-Groove Multijunction Solar Cell, *IEEE Trans. Electron Devices*, **ED-26**, 1091 (1979).
 25. Chiang S. Y., Carbajal B. G., Wakefield G. F. Improved Performance Thin Solar Cells, *IEEE Trans. Electron Devices*, **ED-25**, 1405 (1978); Matzen W. T., Chiang S. Y., Carbajal B. G. A Device Model for the Tandem Junction Solar Cell, *IEEE Trans. Electron Devices*, **ED-26**, 1365 (1979).
 26. Frank R. I., Goodrich J. L., Kaplow R. A Low Series Resistance Silicon Photovoltaic Cell for High Intensity Applications, Conf. Rec. 14th IEEE Photovoltaic Spec. Conf., IEEE, N. Y., 1980, p. 1350.
 27. Feucht D. L. Heterojunction in Photovoltaic Devices, *J. Vac. Sci. Technol.*, **14**, 57 (1977).
 28. Hovel H. J., Woodall J. M. $\text{Ga}_{1-x}\text{Al}_x\text{GaAs } p-p-n$ Heterojunction Solar Cells, *J. Electrochem. Soc.*, **120**, 1246 (1973).
 29. Sites J. R. Current Mechanisms and Barrier Height in ITO/Si Heterojunctions, *Inst. Phys. Conf. Ser.* **43**, Chap. 22 (1979).
 30. Harsha K. S. S., Bachmann K. J., Schmidt P. H., Spencer E. G., Thiel F. A. n -Indium Tin Oxide/ p -InP Solar Cell, *Appl. Phys. Lett.*, **30**, 645 (1977).
 31. Lamorte M. F., Abbott D. Analysis of AlGaAs — GaInAs Cascade Solar Cell under AM0-AM5 Spectra, *Solid State Electron.*, **22**, 467 (1979).
 32. Pulfrey D. L., McQuat R. F. Schottky-Barrier Solar-Cell Calculations, *Appl. Phys. Lett.*, **24**, 167 (1974).
 33. Anderson W. A., Vernon S. M., Mathe P., Lalevic B. Schottky Solar Cells on Thin Epitaxial Silicon, *Solid State Electron.*, **19**, 973 (1976).
 34. Li S. S. A Proposed Novel MPN GaAs Schottky Barrier Solar Cell, *Jpn. J. Appl. Phys.*, **17**, Suppl. 17-1, 291 (1978).
 35. Card H. C., Yang E. S., Panayotatos P. Peaked Schottky-Barrier Solar Cells by Al — Si Metallurgical Reactions, *Appl. Phys. Lett.*, **30**, 643 (1977).
 36. Card H. C., Yang E. S. MIS-Schottky Theory under Conditions of Optical Carrier Generation in Solar Cells, *Appl. Phys. Lett.*, **29**, 51 (1976).
 37. Van Halen P., Mertens R. P., Van Overstraeten R. J., Thomas R. E., Van Meerbergen J. New TiO_x -MIS and SiO_2 -MIS Silicon Solar Cells, *IEEE Trans. Electron Devices*, **ED-25**, 507 (1978).
 38. Barnett A. M., Bragagnolo J. A., Hall R. B., Phillips J. E., Meakin J. D. Achievement of 9.15 % Efficiency in Thin Film CdS/ Cu_2S Solar Cells, Conf. Rec. 13th IEEE Photovoltaic Spec. Conf. IEEE, N. Y., 1978, p. 419.
 39. Burton L. C., Baron B., Devaney W., Hench T. L., Lorenz S., Meakin J. Studies Related to $\text{Zn}_x\text{Cd}_{1-x}\text{S} - \text{Cu}_2\text{S}$ Solar Cells, Conf. Rec. 12th IEEE Photovoltaic Spec. Conf., IEEE, N. Y., 1977, p. 526.
 40. Kazmerski L. L., Ireland P. J., White F. R., Cooper R. B. The Performance of Copper-Ternary Based Thin-Film Solar Cells, Conf. Rec. 13th IEEE Photovoltaic Spec. Conf., IEEE, N. Y., 1978, p. 185.
 41. Wagner S., Bridenbaugh P. M. Multicomponent Tetrahedral Compounds for Solar Cells, *J. Cryst. Growth*, **39**, 151 (1977).
 42. Gibson R. A., LeComber P. G., Spear W. E. Doped Amorphous Silicon and Its Application in Photovoltaic Devices, *IEEE J. Solid State Electron Devices*, **2**, 83 (1978).

43. Carlson D. E. Amorphous Silicon Solar Cells, *IEEE Trans. Electron Devices*, **ED-24**, 449 (1977).
44. Frank R. I., Goodrich J. L., Kaplow R. A Novel Silicon High-Intensity Photovoltaic Cell, GOMAC Conference, Houston, Nov. 1980.
45. Vander Plas H. A., James L. W., Moon R. L., Nelson N. J. Performance of AlGaAs/GaAs Terrestrial Concentrator Solar Cells, Conf. Rec. 13th IEEE Photovoltaic Spec. Conf., IEEE, N. Y., 1978, p. 934.
46. Blocker W. High Efficiency Solar Energy Conversion through Flux Concentration and Spectrum Splitting, *Proc. IEEE*, **66**, 104 (1978).
47. Moon R. L., James L. W., Vander Plas H. A., Yep T. O., Antypas G. A., Chai Y. Multigap Solar Cell Requirements and the Performance of AlGaAs and Si Cells in Concentrated Sunlight, Conf. Rec. 13th IEEE Photovoltaic Spec. Conf., IEEE, N. Y., 1978, p. 859.
48. Javetski J. A Burst of Energy in Photovoltaics, *Electronics*, **52**, 105 (1979).
49. Green M. A., Godfrey R. B., Willison M. R., Blakers A. W. High Efficiency Silicon minMIS Solar Cells, Conf. Rec. 14th IEEE Photovoltaic Spec. Conf., IEEE, N. Y., 1980, p. 684.
50. Thomas R. E., Norman C. F., North R. B. High Efficiency MIS/Inversion Layer Silicon Solar Cells, Conf. Rec. 14th IEEE Photovoltaic Spec. Conf., IEEE, N. Y., 1980. p. 1350.

Приложения

Приложение А. Перечень обозначений

Символ	Физическая величина	Единица измерения
a	Постоянная решетки	Å
\mathcal{B}	Магнитная индукция	$\text{В} \cdot \text{м}^{-2}$
c	Скорость света в вакууме	$\text{см} \cdot \text{с}^{-1}$
C	Емкость	Ф
\mathcal{D}	Электрическая индукция	$\text{Кл} \cdot \text{см}^{-2}$
D	Коэффициент диффузии	$\text{см}^2 \cdot \text{с}^{-1}$
E	Энергия	эВ
E_C	Энергия дна зоны проводимости	эВ
E_F	Энергия Ферми	эВ
E_g	Ширина запрещенной зоны	эВ
E_V	Энергия потолка валентной зоны	эВ
\mathcal{E}	Электрическое поле	$\text{В} \cdot \text{см}^{-1}$
\mathcal{E}_c	Критическое поле	$\text{В} \cdot \text{см}^{-1}$
\mathcal{E}_m	Максимальное поле	$\text{В} \cdot \text{см}^{-1}$
f	Частота	Гц
$F(E)$	Функция распределения Ферми—Дирака	
h	Постоянная Планка	Дж·с
$h\nu$	Энергия фотона	эВ
I	Ток	А
I_C	Коллекторный ток	А
J	Плотность тока	$\text{А} \cdot \text{см}^{-2}$
J_t	Пороговая плотность тока	$\text{А} \cdot \text{см}^{-2}$
k	Постоянная Больцмана	Дж·К ⁻¹
kT	Тепловая энергия	эВ
L	Длина	см или мкм
m_0	Масса свободного электрона	кг
m^*	Эффективная масса	кг
\bar{n}	Показатель преломления	
n	Концентрация свободных электронов	см^{-3}
n_i	Собственная концентрация	см^{-3}
N	Концентрация легирующей примеси	см^{-3}
N_A	Концентрация акцепторов	см^{-3}
N_C	Эффективная плотность состояния в зоне проводимости	см^{-3}
N_D	Концентрация доноров	см^{-3}
N_V	Эффективная плотность состояния в валентной зоне	см^{-3}
p	Концентрация свободных дырок	см^{-3}
P	Давление	$\text{Н} \cdot \text{м}^{-2}$
q	Заряд электрона	Кл
Q_{it}	Плотность поверхностных состояний	см^{-2}
R	Сопротивление	Ом
t	Время	с
T	Абсолютная температура	К
v	Скорость носителей	$\text{см} \cdot \text{с}^{-1}$
v_s	Скорость насыщения	$\text{см} \cdot \text{с}^{-1}$
v_{th}	Тепловая скорость	$\text{см} \cdot \text{с}^{-1}$
V	Напряжение	В
V_{bi}	Встроенный потенциал	В
V_{EB}	Напряжение эмиттер—база	В
V_B	Напряжение пробоя	В
W	Толщина	см или мкм

Продолжение прилож.

Символ	Физическая величина	Единица измерения
W_B	Толщина базы	см или мкм
x	x -направление	
∇	Дифференциальный оператор	
∇T	Температурный градиент	$K \cdot cm^{-1}$
ϵ_0	Диэлектрическая постоянная вакуума	$\Phi \cdot cm^{-1}$
ϵ_s	Диэлектрическая проницаемость полупроводника	$\Phi \cdot cm^{-1}$
ϵ_i	Диэлектрическая проницаемость изолятора	$\Phi \cdot cm^{-1}$
ϵ_s/ϵ_0	Относительная диэлектрическая проницаемость	
τ	Время жизни или время затухания	с
θ	Угол	рад
λ	Длина волны	мкм или Å
ν	Частота света	Гц
μ_0	Магнитная проницаемость вакуума	$Gn \cdot cm^{-1}$
μ_n	Подвижность электронов	$cm^2 \cdot V^{-1} \cdot s^{-1}$
μ_p	Подвижность дырок	$cm^2 \cdot V^{-1} \cdot s^{-1}$
ρ	Удельное сопротивление	Ом·см
ϕ	Высота барьера или квазиуровень Ферми	В
ϕ_{Bn}	Высота барьера Шоттки к полупроводнику n -типа	В
ϕ_{Bp}	Высота барьера Шоттки к полупроводнику p -типа	В
ϕ_m	Работа выхода металла	В
ω	Круговая частота	cm^{-1}
Ω	Сопротивление	Ом

Приложение Б. Международная система единиц

Величина	Единица	Символ	Размерность
Длина	Метр	м	
Масса	Килограмм	кг	
Время	Секунда	с	
Температура	Кельвин	К	
Ток	Ампер	А	
Частота	Герц	Гц	s^{-1}
Сила	Ньютон	Н	$kg \cdot m \cdot s^{-2}$
Давление	Паскаль	Па	$N \cdot m^{-2}$
Энергия	Джоуль	Дж	$N \cdot m$
Мощность	Ватт	Вт	$Дж \cdot с^{-1}$
Электрический заряд	Кулон	Кл	$A \cdot c$
Потенциал	Вольт	В	$Дж \cdot Кл^{-1}$
Проводимость	Сименс	См	$A \cdot V^{-1}$
Сопротивление	Ом	Ом	$V \cdot A^{-1}$
Емкость	Фарада	Ф	$Кл \cdot В^{-1}$
Магнитный поток	Вебер	Вб	$V \cdot c$
Магнитная индукция	Тесла	Тл	$Вб \cdot м^{-2}$
Индуктивность	Генри	Гн	$Вб \cdot A^{-1}$

Приложение В. Физические константы

Величина	Символ	Значение
Ангстрем	Å	$1 \text{ Å} = 10^{-4} \text{ мкм} = 10^{-8} \text{ см}$
Число Авогадро	N_{AVO}	$6,02204 \cdot 10^{23} \text{ моль}^{-1}$
Боровский радиус	a_B	$0,52917 \text{ Å}$
Постоянная Больцмана	k	$1,38066 \cdot 10^{-23} \text{ Дж} \cdot \text{К}^{-1}$
Элементарный заряд	q	$1,60218 \cdot 10^{-19} \text{ Кл}$
Масса свободного электрона	m_0	$0,91095 \cdot 10^{-30} \text{ кг}$

Продолжение прилож. В

Величина	Символ	Значение
Электронвольт	эВ	1 эВ = 1,60218 · 10 ⁻¹⁹ Дж = = 23,053 ккал · моль ⁻¹
Газовая постоянная	R	1,98719 кал · моль ⁻¹ К ⁻¹
Магнитная проницаемость вакуума	μ ₀	1,25663 · 10 ⁻⁸ Гн · см ⁻¹ (4π · 10 ⁻⁹)
Диэлектрическая проницаемость вакуума	ε ₀	8,85418 · 10 ⁻¹⁴ Ф · см ⁻¹ (1/μ ₀ c ²)
Постоянная Планка	h	6,62617 · 10 ⁻³⁴ Дж · с
Редуцированная постоянная Планка	ħ	1,05458 · 10 ⁻³⁴ Дж · с (ħ = $\frac{h}{2\pi}$)
Масса протона	M _p	1,67264 · 10 ⁻²⁷ кг
Скорость света в вакууме	c	2,99792 · 10 ¹⁰ см · с ⁻¹
Стандартная атмосфера		1,01325 · 10 ⁵ Н · м ⁻²
Напряжение, соответствующее температуре 300 К	kT/q	0,0259 В
Длина волны фотона с энергией 1 эВ	λ	1,23977 мкм

Приложение Г. Постоянные решеток

Элемент или соединение	Название	Кристаллическая структура ¹	Постоянная решетки при 300 К (Å)
Элемент			
C	Углерод (алмаз)	D	3,56683
Ge	Германий	D	5,64613
Si	Кремний	D	5,43095
Sn	Серое олово	D	6,48920
Соединения типа			
A ^{IV} B ^{IV}	Карбид кремния	W	a = 3,086, c = 15,117
A ^{III} B ^V	Арсенид алюминия	Z	5,6605
	Фосфид алюминия	Z	5,4510
	Антимонид алюминия	Z	6,1355
	Нитрид бора	Z	3,6150
	Фосфид бора	Z	4,5380
	Арсенид галлия	Z	5,6533
	Нитрид галлия	W	a = 3,189, c = 5,185
	Фосфид галлия	Z	5,4512
	Антимонид галлия	Z	6,0959
	Арсенид индия	Z	6,0584
	Фосфид индия	Z	5,8686
	Антимонид индия	Z	6,4794
A ^{II} B ^{VI}	Сульфид кадмия	Z	5,8320
	Сульфид кадмия	W	a = 4,16, c = 6,756
	Селенид кадмия	Z	6,050
	Теллурид кадмия	Z	6,482
	Окись цинка	R	4,580
	Сульфид цинка	Z	5,420
	Сульфид цинка	W	a = 3,82, c = 6,26
A ^{IV} B ^{VI}	Сульфид свинца	R	5,9362
	Теллурид свинца	R	6,4620

¹ D — структура алмаза, W — структура вюрцита, Z — структура цинковой обманки, R — структура каменной соли.

Приложение Д. Физические параметры важнейших полупроводников

Полупроводники	Ширина запрещенной зоны, Эв		Подвижность ¹ при 300 К, см ² ·В ⁻¹ ·с ⁻¹		Зона ²	Эффективная масса m^*/m_0		ϵ_s/ϵ_0
	300 К	0 К	электронов	дырок		электронов	дырок	
Элементарные								
C	5,47	5,48	1 800	1200	I	0,2	0,25	5,7
Ge	0,66	0,74	3 900	1900	I	1,64 ³⁾ 0,082 ⁴⁾	0,04 ⁵⁾ 0,28 ⁶⁾	16,0
Si	1,12	1,17	1 500	450	I	0,98 ³⁾ 0,19 ⁴⁾	0,16 ⁵⁾ 0,49 ⁶⁾	11,9
Sn		0,082	1 400	1200	D			
Соединения типа A ^{IV} B ^{IV}								
α-SiC	2,996	3,03	400	50	I	0,60	1,00	10,0
AlSb	1,58	1,68	200	420	I	0,12	0,98	14,4
BN	7,5				I			7,1
BP	2,0							
GaN	3,36	3,50	380			0,19	0,60	12,2
GaSb	0,72	0,81	5 000	850	D	0,042	0,40	15,7
GaAs	1,42	1,52	8 500	400	D	0,067	0,082	13,1
GaP	2,26	2,34	110	75	I	0,82	0,60	11,6
InSb	0,17	0,23	80 000	1250	D	0,0145	0,40	17,7
InAs	0,36	0,42	33 000	460	D	0,023	0,40	14,6
InP	1,35	1,42	4 600	150	D	0,077	0,64	12,4
Типа A ^{III} B ^{VI}								
CdS	2,42	2,56	340	50	D	0,21	0,80	5,4
CdSe	1,70	1,85	800		D	0,13	0,45	10,0
CdTe	1,56		1 050	100	D			10,2
ZnO	3,35	3,42	200	180	D	0,27		9,0
ZnS	3,68	3,84	165	5	D	0,40		5,2
PbS	0,41	0,286	600	700	I	0,25	0,25	17,0
PbTe	0,31	0,19	6 000	4000	I	0,17	0,20	30,0

¹⁾ Здесь приведены значения подвижности, измеренные в наиболее чистых и совершенных на сегодняшний день материалах.

²⁾ I — непрямая зона; D — прямая зона.
³⁾ Продольная эффективная масса.
⁴⁾ Поперечная эффективная масса.
⁵⁾ Эффективная масса легких дырок.
⁶⁾ Эффективная масса тяжелых дырок.

Приложение E. Свойства Ge, Si и GaAs при температуре 300 К

	Ge	Si	GaAs
Число атомов в 1 см ³	4,42 · 10 ²²	5,0 · 10 ²²	4,42 · 10 ²²
Атомный вес	72,60	28,09	144,63
Поле пробоя, В · см ⁻¹	~10 ⁵	~3 · 10 ⁵	~4 · 10 ⁵
Кристаллическая структура	Структура алмаза	Структура алмаза	Структура цинковой обманки
Плотность, г · см ⁻³	5,3267	2,328	5,32
Диэлектрическая постоянная	16,0	11,9	13,1
Эффективная плотность состояний в зоне проводимости N_C , см ⁻³	1,04 · 10 ¹⁹	2,8 · 10 ¹⁹	4,7 · 10 ¹⁷
Эффективная плотность состояний в валентной зоне N_V , см ⁻³	6,0 · 10 ¹⁸	1,04 · 10 ¹⁹	7,0 · 10 ¹⁷
Эффективная масса m^*/m_0 электронов	$m_i^* = 1,64$	$m_i^* = 0,98$	0,067
дырок	$m_i^* = 0,082$	$m_i^* = 0,19$	
	$m_{lh}^* = 0,044$	$m_{lh}^* = 0,16$	$m_{lh}^* = 0,082$
	$m_{hh}^* = 0,28$	$m_{hh}^* = 0,49$	$m_{hh}^* = 0,45$
Электронное сродство	4,0	4,05	4,07
Ширина запрещенной зоны, эВ	0,66	1,12	1,424
Собственная концентрация носителей, см ⁻³	2,4 · 10 ¹³	1,45 · 10 ¹⁰	1,79 · 10 ⁶
Собственная дебаевская длина, мкм	0,68	24	2250
Собственное удельное сопротивление, Ом · см	47	2,3 · 10 ⁵	10 ⁸
Постоянная решетки, Å	5,64613	5,43095	5,6533
Коэффициент линейного теплового расширения $\Delta L/L \Delta T$, °C ⁻¹	5,8 · 10 ⁻⁶	2,6 · 10 ⁻⁶	6,86 · 10 ⁻⁶
Температура плавления, °C	937	1415	1238
Время жизни носителей, с	10 ⁻³	2,5 · 10 ⁻³	10 ⁻⁸
Дрейфовая подвижность, см ² · В ⁻¹ · с ⁻¹	{ 3900 1900	1500 450	8500 400
Энергия оптического фонона, эВ	0,037	0,063	0,035
Средняя длина свободного пробега носителя до испускания оптического фонона λ_0 , Å	105	76 (для элек. тронов) 55 (для дырок)	58
Теплоемкость, Дж/г · °C	0,31	0,7	0,35
Коэффициент теплопроводности при 300 К, Вт/см · °C	0,6	1,5	0,46
Коэффициент тепловой диффузии, см ² · с ⁻¹	0,36	0,9	0,14
Давление насыщенных паров, Па	1 при 1330 °C 10 ⁻⁶ при 760 °C	1 при 1650 °C 10 ⁻⁶ при 900 °C	100 при 1050 °C 1 при 900 °C

Приложение Ж. Свойства SiO_2 и Si_3N_4
при температуре 300 К

Структура	SiO_2 (аморфный)	Si_3N_4 (аморфный)
Температура плавления, °С	~1600	—
Плотность, г·см ⁻³	2,2	3,1
Показатель преломления	1,46	2,05
Диэлектрическая постоянная	3,9	7,5
Электрическая прочность, В·см	10 ⁷	10 ⁷
Линия ИК-поглощения, мкм	9,3	11,5—12,0
Ширина запрещенной зоны, эВ	9,0	~5,0
Коэффициент теплового расширения, °С ⁻¹	5·10 ⁻⁷	—
Коэффициент теплопроводности, Вт·см ⁻¹ К ⁻¹	0,014	—
Удельное сопротивление по постоянному току при 25 °С, Ом·см	10 ¹⁴ —10 ¹⁶	10 ¹⁴
при 500 °С, Ом·см	—	2·10 ¹³
Скорость травления в буферном HF ¹⁾ , Å/мин	1000	5—10

¹⁾ Буферный HF: 34,6 % NH_4F , 6,8 % HF, 58,6 % H_2O .

Предметный указатель

Цифры, набранные полужирным шрифтом, соответствуют страницам кн. 2.

- Акцепторные поверхностные состояния 396
Акцепторный уровень 25
Акцепторы 21
«Альфа-частота» отсечки 169
Амбиполярное время жизни 94, 129
- Балластный эмиттерный транзистор 181
Барьер *Мотта* 257, 314
— *Шоттки* 257
— — в качестве стока и истока 72
Барьерная емкость 86
Биполярный транзистор 142, 166
— — статические характеристики 143
Блоховские функции 16
Больцмановская статистика 22
- Валентная зона 18
Варакторы 123
— чувствительность 124
Варисторы 122
Вероятность туннелирования 101, 132
«Воздушная масса» (AM) 390
Время включения вторичного пробоя 183
— — тиристора 225
— — выключения тиристора 225
— — жизни неосновных носителей 131
— — носителей 60
— — пролета 356
Встроенный потенциал 42
Вторичный пробой 182, 183
Выключающий тиристор 240
Выпрямитель 121
Высокочастотные характеристики транзистора 171
Высота барьера 283
- Геометрический коэффициент 37
Гетеролазеры с квантовыми ямами 324
Гетеропереход 132
«Горячие» точки 184
Горячие электроны 143
Градиентное напряжение 89
Граничная частота полевых канальных транзисторов 357
- Двухскоростной пролетный диод (ДСПД) 151, 214
Дебаевская длина дрейфа 84, 383
Деградация лазеров 334
Диак 243
Диод *Мисавы* 154
— *Рида* 152
— с насыщенном дрейфовой скорости 367
— — проколом 206
— *Шоттки* 126
Диодный тиристор 221
Диоды с быстрым восстановлением 125
— — накоплением заряда 126
Диффузионная емкость 101, 103
Диффузионный метод 71
Диффузия носителей 63
Длина амбиполярной диффузии 130
Длинноканальный МОП-транзистор 5
Донорные поверхностные состояния 396
Донорный уровень 25
Доноры 21
Дополнительная функция ошибок 74
Дрейфовая подвижность 36, 40
— скорость носителей 33, 51
Дрейфовый транзистор 146
Дробовый шум 120
Дырка 21
 dV/dt -переключение транзистора 225
- Закон *Фика* 73
Закороченный катод в тиристоре 228
Запрещенные переходы 276
Затвор 8
- Гамма-функция 40
Гексагональный МОП-транзистор 76
Генерационно-рекомбинационный шум 343

- Зона Бриллюэна 17, 19
 — легких дырок 18
 — проводимости 18
 — тяжелых дырок 18
 Зонная структура идеального резкого гетероперехода 132
 Зонные диаграммы МДП-структуры 419
 — энергетические спектры твердых тел 17
- Идеальная МПД-структура 378
 — — характеристики 385
 Излучательная рекомбинация 41
 Излучательные переходы 271
 Изменения высоты барьера 292
 Изоэлектронный центр 281
 Инверсная населенность 314
 Индексы Миллера 14, 15
 Инжекционная электролюминесценция 272
 Инжекционно-пролетный диод (ИПД) 151
 Инжекция неосновных носителей 280
 Ионная имплантация 72
- Картина дальнего поля 307
 Квантовая эффективность люминесценции 276
 Квантовое усиление 314
 КНИ-прибор 75
 Конструкции высокочастотных фотодиодов 356
 — мощных МП-транзисторов для СВЧ-диапазона 362
 Контактная разность потенциалов 257
 Короткоканальные МОП-транзисторы 6
 — эффекты 47, 67
 Коэффициент амбиполярной диффузии 94, 131, 219
 — диффузии 74
 — — электронов 35
 — кривизны 131
 — оптического ограничения 305, 309
 — отражения 284
 — умножения 365
 — усиления тока 149
 — устойчивости 172
 — Холла 40
 — шума 187
 Краевое поле 438
 Краевой МОМ-диод 142
 Кремниевый выключатель 251
- СВЧ-транзистор 166
 Кристаллическая структура полупроводника 11, 12
 Крутизна канала 332
 — транзистора 335
- Лавинное запаздывание 150, 172
 — умножение 106
 Лавинно-пролетный диод (ЛПД) 150
 — — шумы 186—191
 Лавинный пробой 61
 — фотодиод 363
 — — структуры 371, 373
 Лазерная перестройка 332
 Лазерные структуры 319
 Лазеры на двойных гетероструктурах (ДГ-лазеры) 302
 — — основе гетероструктур 299
 — с распределенной обратной связью 324
 Литографические методы 71
 Люминесценция 271
 — квантовая эффективность 276
- Магниторезистивный эффект 41
 Малосигнальный анализ 164
 Максимальная диэлектрическая прочность 427
 — рабочая частота тиристора 232
 — частота генерации 174
 Масштабирование транзисторов 192
 Масштабный фактор 67
 Матрица рассеяния 171
 МДОП-структуры 87
 МДП-структуры 377
 Междолинное рассеяние 34
 Междозонное возбуждение 277
 Междозонное поглощение 47
 Метод вольт-амперных характеристик 292
 — вольт-фарадных характеристик 299
 — проводимости 400
 — сплавления 70
 — энергии активации 297
 Методы возбуждения электролюминесценции 277
 — изготовления полупроводниковых приборов 70
 Микроплазма 370
 Многоканальный полевой транзистор 371
 МНОП-структуры 87
 Модель двух областей 343
 — Гуммеля — Гуна 163
 — с полностью насыщенной скоростью 346

- Эберса — Молла 161, 187
- МОП-структуры с двойной диффузией 70
- МОП-транзистор 5
 - гексагональный 76
 - линейная область работы 14, 57
 - неравновесные условия 8
 - подпороговая область 21
 - подпороговое напряжение 16
 - подпороговые токи 48
 - приближение главного канала 14
 - прокол 58
 - режим насыщения 57
 - — отсечки канала 14, 61
 - — смыкания 58
 - с заглубленным каналом 71
 - — V-канавкой 76
 - типы 29
 - эффект насыщения тока 20
- Мощность, эквивалентная шуму (NEP) 343
- Мощные транзисторы 180
- МП-транзисторы с двойной гетеро-структурой 364

- Напряжение включения транзистора 217
 - плоских зон 203
 - пробоя 156
 - — тиристора 205
 - — прокола 136
 - отсечки канала 333
 - смыкания 203
 - Эрли 159
- Нейтронное легирование 207
- Непрямые переходы 275
- «Непустой нуль» 443
- Нестационарная диффузия носителей 63
- Нормально закрытый полевой транзистор 336

- Обедненный слой 260
- Область безопасной работы транзистора 185
 - дрейфа 153
- Обобщенное соотношение Эйнштейна 35
- Обратная решетка 14
- Обращенный диод 119
- Однопереходные транзисторы 248, 250
- Оже-рекомбинация 41, 153
- Омический контакт 318
- Оптические свойства полупроводников 47
- Оптическое концентрирование излучения 434
- Основные носители 33
- Отрицательный температурный коэффициент 106
- Отсечка канала 17, 42

- Пассивирующий слой 8
- Переключающий МДП-диод 123
 - транзистор 186
 - — эквивалентная схема 189
- Перенос в пленках изоляторов 422
- Переходные процессы в $p-n$ -переходе 117
- Первичный пробой 183
- ПЗС (прибор с зарядовой связью) 377
 - со скрытым каналом 446
 - хранение заряда 429
- $p-i-n$ -диоды 126
- Плавный линейный переход 88
- Планарный процесс 71
- Плоскостные $p-n$ -переходы 69
- Поверхностная рекомбинация 64
- Подвижность дырок 36
 - носителей 24, 33, 51
 - — полевая зависимость 338
- Поглощение излучения 299
- Полевой диод 366
 - слой 8
 - транзистор 325, 333, 368
 - тиристор 252
- Полупроводниковые лазеры 295
 - материалы для лазеров 296—298
- Пороговая энергия ударной ионизации 53
- Постоянная решетки 14
- Потенциальная яма 429
- Правило k -отбора 275
 - равных площадей 249
- Приближение ВКБ (Вентцеля — Крамерса — Бриллюэна) 101
 - плавного канала 330
- Приборы с плавающим затвором 73
 - со скрытым каналом 41
- Примесное возбуждение 340
- Пробой $p-n$ -перехода 103
- Проводимость 36
 - канала 332, 335
 - пленок 425
- Пролетное запаздывание 150, 173
- Пролетный диод с захваченным объемным зарядом лавины 216
 - угол 163
 - эффект 162, 163

- Процессы генерации-рекомбинации носителей 97
 Прямые переходы 275
- Работа выхода 259, 262
 Радиус водородоподобного примесного центра 274
 Разность работ выхода 413
 Рассеяние на акустических фонах 33
 — — ионизированных примесных атомах 33
 Режим инверсии 380
 — обеднения 379
 — обратного запираания тиристора 205
 — прямого запираания тиристора 211
 Резкий переход 80
 Резонатор *Фабри — Перо* 299
 Рекомбинационные процессы 41, 42
 Релаксация фотовозбужденных носителей 60
- Сверхрезкий переход 124
 Сверхрешетки 136
 Светодиоды видимого диапазона 278
 — инфракрасные 291
 — конструкции 283
 Сдвиг напряжения плоских зон 413
 Сдвоенный диод *Шоттки* 317
 Сила *Лоренца* 38
 Скин-эффект 184
 Скорость поверхностной рекомбинации 64
 Скрытый канал 41
 Собственная температура полупроводника 25
 Собственное возбуждение 339
 Собственный полупроводник 21
 Солнечная постоянная 390
 Солнечные элементы, вольт-амперные характеристики 405—409
 — — конструкции 412—417
 — — на барьерах *Шоттки* 423—426
 — — — МДП-структурах 426—428
 — — — $p-n$ -переходе 398
 — — спектральный отклик 399
 — — с гетеропереходами 418—419
 — — тонкопленочные 428
 Соотношения *Больцмана* 91
 Спектры излучения 273
 Спонтанная эмиссия 299
 Сродство к электрону *см.* электронное сродство 259, 378
 Стабилизаторы тока 366
 Стабилитроны 122
- Стимулированное излучение 299
 Структура с диффузионным охранным кольцом 316
 — со сквозным проколом 373
 — тиристора с электродом в форме эвольвенты 231
 Структуры с фаской 208
- Температура «горячих» точек 183
 Теорема *Блоха* 16
 Теория термоэлектронной эмиссии 267
 Тепловая неустойчивость 105
 Тепловые свойства полупроводников 49
 Тепловой шум 119
 Тиристоры 203, 232, 235
 Ток отсечки канала 331
 Тонкопленочный транзистор 74
 Тормозное излучение 273
 Точечно-контактный выпрямитель 310
 Транзистор с проникаемой базой 196, 198
 — со ступенчатым электродом 178
 Транзисторы на горячих электронах 195
 — с гетеропереходом 195
 Трехфазный ПЗС 433, 442
 Триаки 245
 Триодный тиристор 221
 Туннелирование носителей 94
 Туннельная спектроскопия 119
 Туннельный диод 96
 — — вольт-амперные характеристики 109
 — МДП-диод 122—129
 — МДМ-диод 137
 — — скорость переключения 117
 — — частота отсечки 116
 — транзистор 142
 — механизм проводимости 424
 — ток 276
 — эффект 105
- Ударная ионизация 51, 53, 106
 Удельная обнаружительная способность 343
 Удельное сопротивление 36
 Униполярная структура выпрямителя 131
 Уравнение *Пуассона* 116, 133, 330, 382
 — *Шредингера* 16
 Уравнения *Максвелла* 58
 Уровень инжекции 43
 — *Ферми* 23

- Условие лавинного пробоя 152
— электронейтральности 29, 382
- Фактор вырождения донорного уровня 29
— неидеальности 308, 356
- Фликер-шум 120
- Фононные спектры 45
- Формула Шокли 95, 120
- Фотодетекторы 339
- Фотодиод 345
— с барьером Шоттки 357
— — гетеропереходом 360—361
- Фоторезистор 339
- Фототиристор 237
- Фототранзистор 382
- Фотоэлектрический метод 300
- Функция Гаусса 74
— распределения Ферми — Дирака 22
— Эрмита — Гаусса 313
- Хвосты плотности состояний 316
- Холловская подвижность 36, 40
- Частота отсечки 168
- Число Гуммеля 148
- Ширина запрещенной зоны 18
- Шкала электроотрицательности 290
- Шум лавинного умножения 367
— генерационно-рекомбинационный 343
— дробовой 120
— тепловой 119
- Шумовая постоянная 118
- Шумовое отношение 188
- Шум-фактор 175, 361, 118, 364
- Шумовые характеристики полевых транзисторов 361
- Шумы 119
- Экситон 48
- Электроды переноса 434
- Электролюминесценция 270
— методы возбуждения 277
- Электронное сродство 259, 378
- Электроотрицательность 289
- Эмиссия Пула — Френкеля 423
- Энергетические зоны 16
- Энергонезависимые элементы памяти 78
— — — время хранения информации 86
- Энергия ионизации 27
- Эпитаксиальное выращивание 72
- Эпитаксиальные структуры 71
- Эффект Ганна 226—227
— dV/dt 229
— Кирка 155
— междолинного перехода электронов см. эффект Ганна 226—227
— ограничения объемным зарядом 205
— постоянной Ричардсона 274
— Холла 36, 38
— Шоттки 262
- Эффективная постоянная Ричардсона 135
- Эффективное время жизни носителей 97
- Эффективность переноса 439
— преобразования 393
— — идеальная 395
- Яркостный эквивалент излучения 287
- Ячейка Вигнера — Зейтца 15

Оглавление

Глава 8. МОП-транзисторы	5
8.1. Введение	5
8.2. Основные характеристики	7
8.3. Неоднородное легирование и приборы со скрытым каналом	32
8.4. Короткоканальные эффекты	47
8.5. МОП-транзисторные структуры	67
8.6. Энергонезависимые элементы памяти	78
Литература	90
ЧАСТЬ IV. ПОЛУПРОВОДНИКОВЫЕ ПРИБОРЫ СВЧ-ДИАПАЗОНА	94
Глава 9. Туннельные приборы	94
9.1. Введение	94
9.2. Туннельный диод	96
9.3. Обращенный диод	119
9.4. Туннельный МДП-диод	122
9.5. Переключающий МДП-диод	133
9.6. Туннельный МДМ-диод	137
9.7. Туннельный транзистор	142
Литература	147
Глава 10. Лавинно-пролетные диоды и другие приборы на пролетных эффектах	150
10.1. Введение	150
10.2. Статические характеристики	152
10.3. Динамические характеристики	162
10.4. Мощность и коэффициент полезного действия	171
10.5. Шумы ЛПД	186
10.6. Конструкция прибора и характеристики	191
10.7. Инжекционно-пролетные диоды и двухскоростные пролетные диоды	201
10.8. Пролетный диод с захваченным объемным зарядом лавины	216
Литература	221
Глава 11. Приборы на эффекте междолинного перехода электронов	226
11.1. Введение	226
11.2. Эффект междолинного перехода электронов	227
11.3. Режимы работы	341
11.4. Характеристики приборов на эффекте междолинного перехода электронов	257
Литература	267

ЧАСТЬ V. ОПТОЭЛЕКТРОННЫЕ ПРИБОРЫ	270
Глава 12. Светодиоды и полупроводниковые лазеры	270
12.1. Введение	270
12.2. Излучательные переходы	271
12.3. Светодиоды	278
12.4. Физика полупроводниковых лазеров	295
12.5. Рабочие характеристики лазеров	319
Литература	335
Глава 13. Фотодетекторы	339
13.1. Введение	339
13.2. Фоторезистор	339
13.3. Фотодиод	345
13.4. Лавинный фотодиод	363
13.5. Фототранзистор	382
Литература	386
Глава 14. Солнечные батареи	389
14.1. Введение	389
14.2. Солнечное излучение и идеальная эффективность преобразования	390
14.3. Солнечные элементы на $p - n$ -переходах	398
14.4. Солнечные элементы с гетеропереходами. Поверхностные и тонкопленочные солнечные элементы	418
14.5. Концентрирование света	434
Литература	440
Приложения	443
Предметный указатель	449

УВАЖАЕМЫЙ ЧИТАТЕЛЫ

Ваши замечания о содержании книги, ее оформлении, качестве перевода и другие просим присылать по адресу: 129820, Москва, И-110, ГСП, 1-й Рижский пер., д. 2, издательство «Мир».

С. Зи

ФИЗИКА ПОЛУПРОВОДНИКОВЫХ ПРИБОРОВ

В 2-х книгах

Книга 2

Старший научный редактор И. М. Андреева
Младший научный редактор Л. С. Сысоева
Художник Д. А. Аникеев
Художественный редактор В. Б. Прищепа
Технический редактор В. П. Сизова
Корректор В. С. Соколов

ИБ № 3811

Сдано в набор 10.04.84.
Подписано к печати 06.11.84.
Формат 60×90^{1/16}
Бумага типографская № 1. Объем 14,25.
Гарнитура литературная. Печать высокая.
Усл. печ. л. 28,50. Усл. кр. отт. 28,50.
Уч.-изд. л. 25,95. Изд. № 20/2919.
Тираж 16 000 экз. Зак. 110. Цена 2 р. 20 к.

ИЗДАТЕЛЬСТВО «МИР»

129820, Москва, И-110, ГСП, 1-й Рижский пер., 2.

Ленинградская типография № 6 ордена Трудового Красного Знамени
Ленинградского объединения «Техническая книга» им. Евгении Соколовой
Союзполиграфпрома при Государственном комитете СССР
по делам издательств, полиграфии и книжной торговли.
193144, г. Ленинград, ул. Моисеенко, 10.