

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«ПЕРМСКИЙ ГОСУДАРСТВЕННЫЙ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ»

А. С. Ажеганов, И. Л. Вольхин, Н. К. Шестакова

**КВАНТОВАЯ И ПОЛУПРОВОДНИКОВАЯ
ЭЛЕКТРОНИКА
КУРС ЛЕКЦИЙ**

*Допущено методическим советом
Пермского государственного
национального исследовательского университета
в качестве учебного пособия для студентов,
обучающихся по направлениям подготовки магистров
«Нанотехнологии и микросистемная техника»*



Пермь 2018

УДК 537.86+530.145+537.312 [075.8]

ББК 32.86

А342

Ажеганов А. С., Вольхин И. Л., Шестакова Н. К.

А342 Квантовая и полупроводниковая электроника: курс лекций: учеб. пособие / А. С. Ажеганов, И. Л. Вольхин, Н. К. Шестакова; Перм. гос. нац. исслед. ун-т. – Пермь, 2018.– 159 с.: ил.

ISBN 978-5-7944-3213-8

Рассмотрены принципы работы квантовых и полупроводниковых устройств: фотоэлектронных и оптоэлектронных приборов, приёмников теплового излучения, полупроводниковых лазеров, суперлюминесцентных диодов, волоконно-оптических усилителей, лазеров, интегральных электрооптических модуляторов.

Пособие предназначено для студентов, обучающихся по направлениям магистратуры «Нанотехнологии и микросистемная техника», «Радиофизика», а также смежных специальностей.

УДК 537.86+530.145+537.312 [075.8]

ББК 32.86

*Печатается по решению ученого совета физического факультета
Пермского государственного национального исследовательского университета*

Рецензенты: кафедра прикладной математики Пермского национального исследовательского политехнического университета (зав. каф. – д-р техн. наук, проф. **В. П. Первадчук**); гл. технолог волоконно-оптических технологий ПАО «Пермская научно-производственная приборостроительная компания» **О. Л. Кель**

ISBN 978-5-7944-3213-8

© ПГНИУ, 2018

© ПАО ПНППК, 2018

© Ажеганов А. С., Вольхин И. Л.,
Шестакова Н.К., 2018

ПРЕДИСЛОВИЕ

Учебное пособие “Квантовая и полупроводниковая электроника” включает в себя материал лекций, читаемых на физическом факультете ПГНИУ, и частично материалы, необходимые для самостоятельной работы студентов.

В пособии рассматриваются физические принципы работы и конструкции полупроводниковых, болометрических и пироэлектрических приёмников теплового излучения, суперлюминесцентных и лазерных диодов, оптоволоконных усилителей и лазеров, интегральных электрооптических модуляторов.

Полупроводниковые приёмники инфракрасного (ИК) излучения используются в оптических системах связи и автоматике. На основе приёмников излучения дальнего ИК диапазона строятся тепловизионные приборы, позволяющие получать изображение объектов по их собственному тепловому излучению.

Применение суперлюминесцентных и лазерных диодов накачки и внутриволоконных брэгговских решёток показателя преломления позволило существенно повысить выходную мощность, экономичность и надёжность волоконных усилителей и лазеров.

Спектральная область эрбиевых усилителей и лазеров совпадает с диапазоном длин волн волоконно-оптических линий связи. На их основе созданы полностью волоконные линии передачи информации.

Использование электрооптических модуляторов позволило существенно улучшить технические характеристики трактов передачи аналоговых и цифровых сигналов.

Закрепление теоретических знаний происходит в ходе самостоятельной работы студентов и при выполнении лабораторного практикума.

1. ПРИЁМНИКИ ОПТИЧЕСКИХ СИГНАЛОВ

1.1. Фоторезисторы, фотодиоды, фототранзисторы

1.1.1. Собственные полупроводники

Химически чистые полупроводники (т.е. не содержащие примесей других элементов или соединений) называют *собственными полупроводниками* или *полупроводниками i -типа* (intrinsic semiconductor). К ним относятся кремний Si, германий Ge, селен Se и ряд химических соединений, таких как арсенид галлия GaAs, фосфид галлия GaP, карбид кремния SiC и др.

Кристаллы элементов IV группы Периодической системы элементов – углерода, кремния и германия имеют одинаковую кристаллическую решётку, которую называют решёткой алмаза. Каждый атом в решётке имеет четыре ближайших соседа, с которыми он образует гомеополярные (ковалентные) связи. Ковалентные связи, образованные парами электронов на схеме (рис.1.1,*а*), изображены в виде двойных линий.

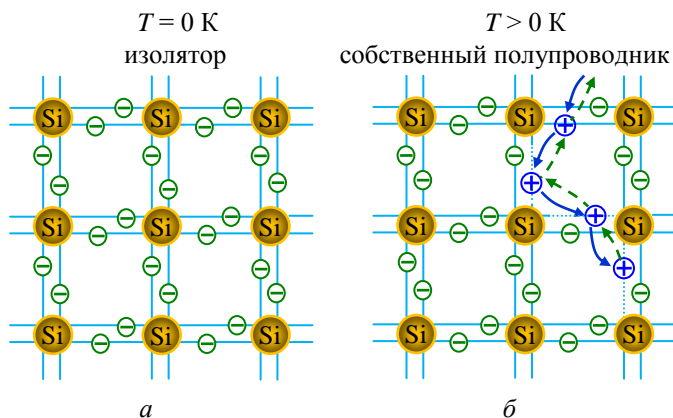


Рис.1.1. Схематичное изображение собственного полупроводника: *а* – при $T = 0\text{ K}$, свободных носителей нет, *б* – при $T > 0\text{ K}$, тепловая генерация пар электрон-дырка

Электронные конфигурации элементов IV группы показаны в таблице (заштрихованы оболочки атомных остовов).

Табл.1.1. Электронные конфигурации элементов главной подгруппы IV группы периодической системы элементов

Слой	K	L		M			N		
Оболочка	1s	2s	2p	3s	3p	3d	4s	4p	4d
Атом $_6\text{C}$	2	2	2	-					
Атом $_{14}\text{Si}$	2	2	6	2	2	-			
Атом $_{32}\text{Ge}$	2	2	6	2	6	10	2	2	-

Атомы IV группы имеют по 4 валентных электрона:

$(2s)^2$ и $(2p)^2$ – у углерода,

$(3s)^2$ и $(3p)^2$ – у кремния,

$(4s)^2$ и $(4p)^2$ – у германия.

При образовании парно-электронных связей каждый атом дополняет свою внешнюю оболочку до 8 электронов, что соответствует полностью заполненным s - и p -состояниям.

Энергетические зоны в кристалле полупроводника

В кристалле энергетические уровни s - и p -состояний расщепляются в единую энергетическую зону – *валентную зону* (рис.1.2). Валентные зоны кристаллов C, Si и Ge являются *заполненными*, так как все состояния заняты электронами.

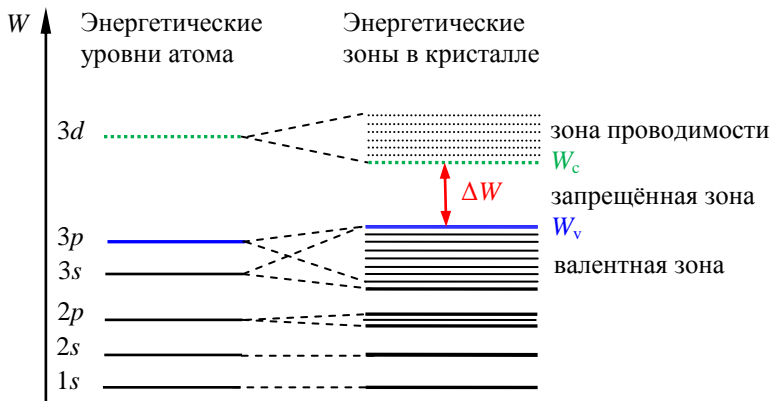


Рис.1.2. Образование энергетических зон в кристалле кремния из дискретных энергетических уровней отдельных атомов

Разрешённый энергетический уровень d -оболочки, свободный от электронов в невозбуждённом состоянии атома, расщепляясь, образует зону свободных уровней или *свободную зону* (рис.1.2), которую называют *зоной проводимости*. Между этой зоной и валентной находится *запрещённая зона*. Ширина запрещённой зоны $\Delta W = W_c - W_v$ в кристалле кремния равна 1,1 эВ, в кристалле германия – 0,7 эВ.

Носители заряда в собственном полупроводнике

При низких температурах ($T \rightarrow 0$ К) все валентные электроны атомов локализованы в ковалентных связях (рис.1.1,*а*). При повышении температуры возрастает число атомов N_T , у которых энергия тепловых колебаний превышает энергию ковалентной связи ΔW

$$\frac{N_T}{N} \sim \exp\left(-\frac{\Delta W}{kT}\right).$$

Атом может передать порцию энергии, большую или равную ΔW , одному из валентных электронов. Этот электрон выходит из валентной связи, и при этом у атомов, участвующих в связи оказывается некомпенсированным положительный заряд ядер. Такой положительный заряд называют *дыркой*. “Вакантное место” в незаполненной связи может быть занято электроном из соседней связи. Образовавшаяся вакансия, в свою очередь, может быть занята электроном из другой связи и т.д. Таким образом положительная дырка хаотически движется по кристаллу в направлении, противоположном перемещению электронов (синяя пунктирная линия на рис.1.1,*б*). Движение дырки можно рассматривать как тепловое движение квазичастицы, имеющей положительный заряд $+q$, численно равный заряду электрона.

На зонной диаграмме (рис.1.3,*б*) уход электрона из ковалентной связи изображается как переход электрона из валентной зоны в зону проводимости

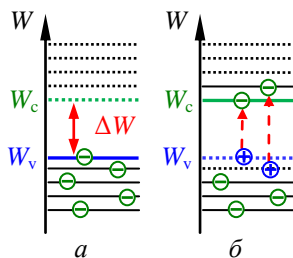


Рис.1.3. Схема зонной структуры собственного полупроводника: *а* – при $T = 0$ К, *б* – при $T > 0$ К

сти. В зоне проводимости появляются *свободные электроны*, а в валентной зоне – свободные подуровни (*дырки*).

Таким образом, нарушение валентной связи за счёт тепловой энергии приводит к появлению в кристалле полупроводника двух свободных носителей заряда – электрона в зоне проводимости и дырки в валентной зоне. Электропроводность полупроводника, возникающая в результате нарушения валентных связей, называется *собственной электропроводностью*.

Концентрация электронов и дырок в полупроводнике

Одновременно с генерацией носителей заряда происходит их рекомбинация. Электрон из зоны проводимости возвращается в валентную зону на свободный уровень энергии, и пара электрон-дырка исчезает (рекомбинирует). Возникшая при рекомбинации избыточная энергия выделяется в виде тепла или кванта света. В состоянии термодинамического равновесия процессы генерации и рекомбинации уравновешены. В собственном полупроводнике устанавливаются *равновесные концентрации* электронов и дырок:

$$n_i = N_c \exp\left(-\frac{\Delta W}{2kT}\right), \quad (1.1)$$

$$p_i = N_v \exp\left(-\frac{\Delta W}{2kT}\right), \quad (1.2)$$

где N_c и N_v – плотности расположения энергетических уровней (плотности квантовых состояний) в зоне проводимости и в валентной зоне соответственно.

В собственных полупроводниках концентрация электронов в зоне проводимости равна концентрации дырок в валентной зоне $n_i = p_i$, так как каждый электрон, переходя в зону проводимости, оставляет после своего ухода дырку. Произведение концентраций электронов n_i и дырок p_i для собственного полупроводника определяется шириной запрещённой зоны ΔW и температурой T :

$$n_i p_i = N_c N_v \exp\left(-\frac{\Delta W}{kT}\right). \quad (1.3)$$

$$n_i p_i = n_i^2 = p_i^2. \quad (1.4)$$

Расстояние между соседними подуровнями в валентной зоне и зоне проводимости настолько мало, что зоны можно считать практически непрерывными. При этом даже очень слабое электрическое поле может сообщить электронам дополнительную энергию, достаточную для их перехода на ближайшие свободные подуровни. В кристалле появится преимущественное движение электронов против направления поля, что эквивалентно движению положительных дырок по направлению поля, т.е. возникнет электрический ток.

Электропроводность собственных полупроводников

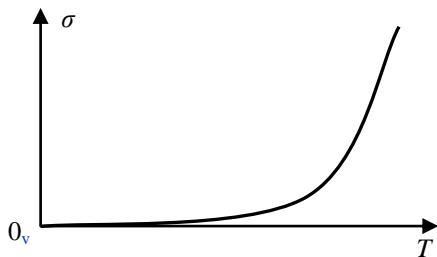
Удельная проводимость полупроводника σ (Ом·м)⁻¹ определяется формулой

$$\sigma = qn\mu_n + qp\mu_p, \quad (1.5)$$

где $q = 1,6 \cdot 10^{-19}$ Кл – элементарный заряд, n – концентрация электронов, p – концентрация дырок, μ_n и μ_p (м²/В·с) – величины подвижности соответственно электронов и дырок (т.е. средние скорости их перемещения вдоль направления электрического поля при напряжённости $E = 1$ В/м).

Подстановкой (1.1) и (1.2) в формулу (1.5) получим выражение для проводимости собственного полупроводника

$$\begin{aligned} \sigma &= q\mu_n N_c \exp\left(-\frac{\Delta W}{2kT}\right) + q\mu_p N_v \exp\left(-\frac{\Delta W}{2kT}\right), \\ \sigma &= \sigma_0 \exp\left(-\frac{\Delta W}{2kT}\right), \end{aligned} \quad (1.6)$$



где $\sigma_0 = q(\mu_n N_c + \mu_p N_v)$. Для собственных полупроводников характерен экспоненциальный рост проводимости при увеличении температуры (рис.1.4).

Рис.1.4. Температурная зависимость проводимости собственного полупроводника

1.1.2. Фоторезистивный эффект

Поглощение света в полупроводниках

Внутренний фотоэффект – это явление ионизации атомов полупроводника под действием света, приводящий к образованию добавочных, *неравновесных носителей заряда*. Добавочную проводимость, обусловленную внутренним фотоэффектом, называют *фотопроводимостью*.

На рис.1.5 изображены уровни энергии дна зоны проводимости W_c и верха валентной зоны W_v в собственном полупроводнике. Если энергия кванта света $h\nu \geq \Delta W$, то при её поглощении атомом полупроводника валентный электрон может перейти в зону проводимости, при этом образуется пара свободных носителей заряда – электрон и дырка.

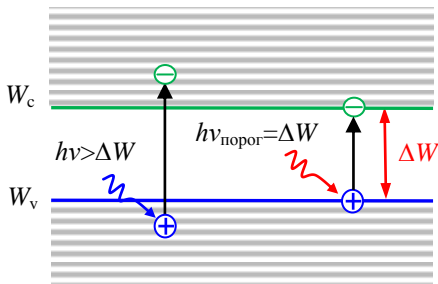


Рис.1.5. Энергетические зоны и собственное поглощение фотонов в полупроводнике

Энергия кванта, необходимая для перехода с верхнего уровня валентной зоны на нижний уровень зоны проводимости $h\nu_{\text{порог}} = \Delta W$, определяет *пороговую длину волны* (красную границу) фотоэффекта $\lambda_{\text{порог}} = c/\nu_{\text{порог}} = hc/\Delta W$. При поглощении фотонов с $\lambda < \lambda_{\text{порог}}$ возможен переход электронов с более низких уровней валентной зоны на более высокие уровни зоны проводимости.

При длинах волн, больших пороговой, энергии фотонов недостаточно для образования электронно-дырочных пар. Однако поглощение света все же происходит, поскольку фотоны могут поглощаться имеющимися в полупроводнике собственными носителями заряда – электронами и дырками.

Фотопроводимость полупроводников

Пусть на поверхность полупроводника падает монохроматический поток света Φ_λ ($\text{Дж/с}\cdot\text{м}^2$) с длиной волны $\lambda \leq hc/\Delta W$.

Число фотонов света, входящих в полупроводник в единицу времени, не учитывая отражение света от поверхности, определяется как $N = \Phi_\lambda / h\nu$.

Число пар свободных носителей заряда (электронов и дырок), возникающих (генерируемых) в единице объёма полупроводника под действием света в единицу времени равно:

$$\left(\frac{dn_\Phi}{dt} \right)_{\text{ген}} = \left(\frac{dp_\Phi}{dt} \right)_{\text{ген}} = \beta k N = \beta k \frac{\Phi_\lambda}{h\nu}, \quad (1.7)$$

где k – коэффициент поглощения света в полупроводнике. Коэффициент β называют *квантовым выходом*, который определяет число пар носителей заряда, образуемых одним поглощённым фотоном. Обычно $\beta \leq 1$.

Одновременно с генерацией *неравновесных* носителей заряда идёт процесс их рекомбинации. В собственном полупроводнике с малой концентрацией равновесных носителей заряда $n_0 = p_0 \ll n_\Phi = p_\Phi$ скорость рекомбинации пропорциональна квадрату концентрации неравновесных носителей (квадратичная релаксация):

$$\left(\frac{dn_\Phi}{dt} \right)_{\text{рек}} = \left(\frac{dp_\Phi}{dt} \right)_{\text{рек}} = -\mathcal{M}_\Phi p_\Phi = -\mathcal{M}_\Phi^2 = -\mathcal{P}_\Phi^2. \quad (1.8)$$

Изменение концентрации неравновесных носителей в единицу времени равно разности между скоростями генерации и рекомбинации:

$$\frac{dn_\Phi}{dt} = \beta k \frac{\Phi_\lambda}{h\nu} - \mathcal{M}_\Phi^2, \quad \frac{dp_\Phi}{dt} = \beta k \frac{\Phi_\lambda}{h\nu} - \mathcal{P}_\Phi^2. \quad (1.9)$$

Установившееся значение концентрации электронов и дырок ($dn_\Phi/dt = 0$ и $dp_\Phi/dt = 0$) пропорционально значению корня квадратного из интенсивности светового потока:

$$n_\Phi = p_\Phi = \sqrt{\frac{\beta k}{\gamma}} \cdot \sqrt{\frac{\Phi_\lambda}{h\nu}}. \quad (1.10)$$

Генерация неравновесных носителей заряда под действием света приводит к увеличению удельной электропроводности полупроводника

$$\sigma_c = \sigma_t + \sigma_\phi,$$

где $\sigma_t = q(n_0\mu_n + p_0\mu_p)$ – темновая проводимость (в отсутствие светового потока), $\sigma_\phi = q(n_\phi\mu_n + p_\phi\mu_p)$ – фотопроводимость полупроводника, q – элементарный заряд, μ_n и μ_p ($\text{м}^2/\text{В}\cdot\text{с}$) – величины подвижности соответственно электронов и дырок.

В общем случае можно считать, что фотопроводимость

$$\sigma_\phi = g\sqrt{\Phi_\lambda}, \quad (1.11)$$

где g – постоянный коэффициент, зависящий от материала полупроводника.

Спектральная зависимость фотопроводимости

Спектральная зависимость фотопроводимости определяется спектром поглощения полупроводника. Область собственного поглощения имеет длинноволновую (красную) границу, соответствующую длине волны света $\lambda_{\text{порог}} = hc/\Delta W$ (рис. 1.6).

С ростом энергии фотонов фотопроводимость быстро увеличивается, достигает максимума, а затем уменьшается (синяя

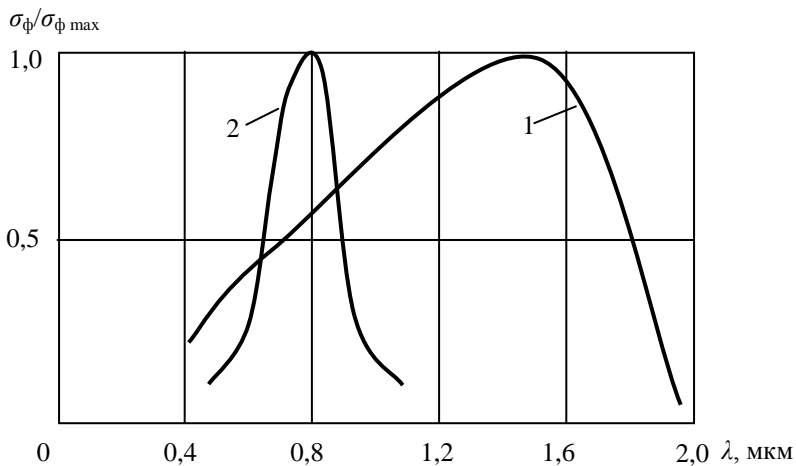


Рис.1.6. Нормированная спектральная характеристика фотопроводимости: 1 – германия, 2 – кремния

граница фотопроводимости). В районе синей границы весь световой поток поглощается в приповерхностном слое полупроводника, при этом за счёт поверхностной рекомбинации время жизни τ носителей заряда резко уменьшается.

Так как запрещённая зона различных полупроводниковых материалов имеет ширину от десятых долей электрон-вольта до двух-трёх электрон-вольт, фотопроводимость может обнаруживаться в ультрафиолетовой, видимой ($\lambda = 0,40 \div 0,76$ мкм) или инфракрасной частях спектра электромагнитных волн.

По длинноволновому краю фотопроводимости может быть определена ширина запрещённой зоны полупроводника. Из рис.1.6 видно, что пороговая длина волны для германия $\lambda_{\text{порог}} = 1,8$ мкм и находится в инфракрасной области спектра. Следовательно ширина запрещённой зоны для германия $\Delta W = hc/\lambda_{\text{порог}} = 0,72$ эВ. Для кремния пороговая длина волны $\lambda_{\text{порог}} = 1,15$ мкм также находится в инфракрасной области спектра и ширина запрещённой зоны кремния $\Delta W = 1,12$ эВ.

1.1.3. Фоторезисторы

Фоторезисторы – полупроводниковые приборы, работа которых основана на использовании явления внутреннего фотоэффекта. Конструкция фоторезистора и его включение в электрическую цепь схематично изображены на рис.1.7. Фоторезистор представляет собой тонкий слой поликристаллического полупроводникового материала, нанесённого на диэлектрическую подложку. В качестве фоточувствительного материала обычно используют сульфид кадмия CdS, селенид кадмия CdSe, сульфид свинца PbS или селенид свинца PbSe (рис.1.8). На по-

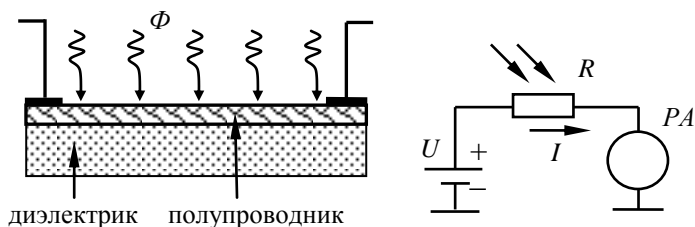


Рис.1.7. Конструкция фоторезистора и его включение во внешнюю цепь

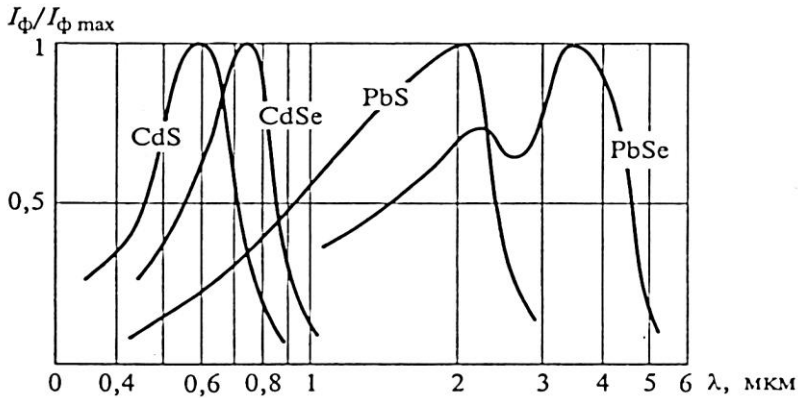


Рис.1.8. Нормированная спектральная характеристика для различных фоторезисторов

верхность фоточувствительного слоя нанесены металлические электроды.

Если фоторезистор R включен во внешнюю цепь последовательно с источником напряжения U , то при отсутствии освещения через него течёт *темновой ток*:

$$I_t = G_t U, \quad (1.12)$$

где $G_t = 1/R_t$ (Ом) $^{-1}$ – проводимость фоторезистора при отсутствии освещения определяется материалом и размерами фоточувствительного слоя. В простейшем случае $G_t = \sigma_t(a^2/b)$, где σ_t – удельная темновая проводимость, a^2 – сечение, b – длина фоточувствительного слоя.

При освещении поверхности фоторезистора в цепи течёт световой ток:

$$I_c = G_c U = G_t U + G_\phi U. \quad (1.13)$$

Разность между световым током и темновым током называется *фототоком*:

$$I_\phi = I_c - I_t = G_\phi U. \quad (1.14)$$

Вольт-амперными характеристиками (ВАХ) фоторезистора называются зависимости величин темнового тока, светового тока и фототока от приложенного к фоторезистору напряжения

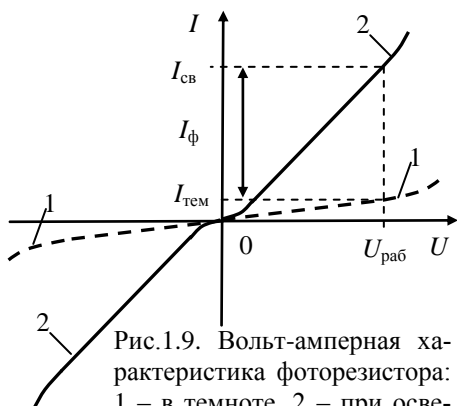


Рис.1.9. Вольт-амперная характеристика фоторезистора: 1 – в темноте, 2 – при освещении

при постоянной величине светового потока, падающего на фоторезистор (рис.1.9).

ВАХ имеют слабую нелинейность при малых напряжениях, что связано с туннелированием электронов через небольшие потенциальные барьеры между отдельными зёрнами или кристалликами полупроводника. При увеличении напряжения

энергия электронов начинает превышать высоту потенциальных барьеров между зёрнами, и ВАХ становится линейной. При больших напряжениях повышается температура фоторезистора из-за выделяющегося джоулева тепла $P = U \cdot I_c$. Темновой ток I_t возрастает, и ВАХ снова становится нелинейной.

Основные характеристики и параметры фоторезисторов

1. *Спектральные характеристики*, т.е. зависимости фототока I_ϕ , фотопроводимости G_ϕ или фотосопротивления $R_\phi = 1/G_\phi$ от длины волны падающего света. Максимумы этих характеристик находятся в видимой или инфракрасной частях спектра (рис.1.8).

2. *Световые характеристики фоторезистора*, т.е. зависимости фототока I_ϕ или фотопроводимости G_ϕ от величины освещённости E (см. прил. 1) при заданном постоянном напряжении, приложенном к фоторезистору (рис.1.10).

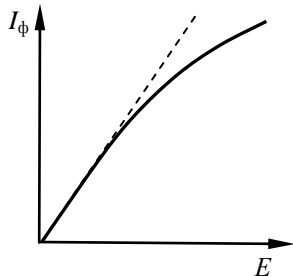


Рис.1.10. Световая характеристика фоторезистора

При малых освещённостях световая характеристика $I_\phi(E)$ близка к линейной. При больших освещённостях она начинает отличаться от линейной и изменение фототока (фото-

проводимости) может быть описано формулой (1.11):

$$I_{\phi} = G_{\phi} U = Y_{\phi} U \sqrt{E}, \quad (1.15)$$

где Y_{ϕ} – коэффициент, зависящий от материала и размеров фоточувствительного слоя.

На практике обычно используют интегральные световые (люкс-амперные) характеристики фоторезисторов, полученные при освещении фоторезистора светом лампы накаливания, вольфрамовая нить которой нагрета до температуры $T = 2840$ К.

3. Интегральная чувствительность $K_{\text{инт}}$ фоторезистора, т.е. отношение фотопроводимости к величине освещённости его поверхности:

$$K_{\text{инт}} = \frac{G_{\phi}}{E} = \frac{I_{\phi}}{U} \cdot \frac{1}{E}, \frac{\text{мА}}{\text{В} \cdot \text{лк}}. \quad (1.16)$$

При определении интегральной чувствительности освещение фоторезисторов производят также лампой накаливания, вольфрамовая нить которой нагрета до 2840 К.

4. Постоянная времени спада фототока τ , т.е. время, в течение которого фототок уменьшается в e раз после прекращения освещения фоторезистора. Постоянная времени τ характеризует инерционность фоторезистора и определяется временем жизни неосновных носителей заряда. Фоторезисторы различных типов имеют постоянные времени от 0,1 мс до 100 мс.

1.1.4. Примесные полупроводники, p-n-переход

Для придания полупроводнику необходимых электрических свойств в кристалл полупроводника вводят примесные атомы.

Электронные полупроводники

Пусть имеется кристалл кремния, в котором часть атомов кремния Si замещена атомами пентавалентного фосфора Р, примерно один атом фосфора на $10^4 \div 10^5$ атомов кремния.

Каждый атом кремния имеет 4 валентных электрона и образует ковалентные связи с четырьмя соседними атомами. Атом фосфора имеет 5 валентных электронов. Четыре из них войдут в ковалентные связи с соседними атомами кремния Si. Пятый ва-

лентный электрон в химической связи не участвует. Для того, чтобы оторвать его от атома фосфора и превратить в свободный

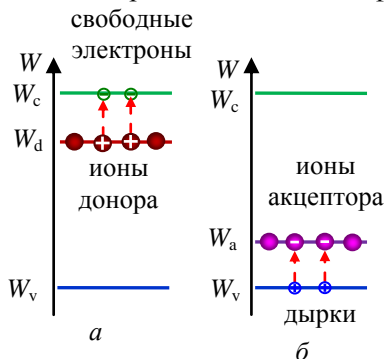


Рис.1.11. Зонная диаграмма активации примеси: а – донорной, б – акцепторной

электрон, требуется значительно меньшее количество энергии ($\Delta W_d < 0,01$ эВ), чем для высвобождения электрона из валентной связи ($\Delta W \sim 1$ эВ). При комнатной температуре ($kT \approx 0,025$ эВ) практически все атомы фосфора ионизированы.

Такие электроны при переходе в зону проводимости не оставляют вакансий (дырок) в валентной зоне. Концентрация электронов n_n превышает концентрацию

дырок p_n . Электроны становятся *основными носителями*, а дырки – *неосновными*.

Полупроводники, в которых основными носителями являются электроны, называются *электронными полупроводниками* или *полупроводниками n-типа* (от латинского negative – отрицательный). Атомы V группы Периодической системы элементов, являющиеся источником электронов проводимости, называются *донорными примесями*.

На зонной диаграмме полупроводника *n-типа* энергетический уровень пятого электрона атомов донорной примеси W_d располагается ниже дна зоны проводимости на $\Delta W_d \sim 0,01$ эВ (рис.1.11,а).

Дырочные полупроводники

Для создания полупроводника с дырочной проводимостью в кристаллическую решётку кремния Si вводят атомы III группы Периодической системы элементов, например, бора В. Внешняя электронная оболочка атома III группы содержит 3 валентных электрона. В кристаллической решётке кремния атом бора образует три заполненных валентных связи. Четвертая связь остается незаполненной. Вследствие теплового движе-

ния электрон одной из соседних валентных связей между атомами Si может перейти в эту связь. Образуются заполненные связи атома бора с четырьмя соседними атомами кремния. При этом в оболочке атома бора появляется лишний электрон, и атом бора превращается в отрицательный ион. В возникшей незаполненной связи между атомами кремния появляется положительный заряд – дырка.

Примеси, принимающие на свои валентные уровни валентные электроны полупроводника и являющиеся источником дырок, называются *акцепторными примесями*.

На зонной диаграмме полупроводника свободный уровень атома бора (*акцепторный уровень*) W_a расположен в запрещённой зоне недалеко от потолка валентной зоны кристалла кремния (рис.1.11,б). Один из валентных электронов покидает валентную зону и заполняет этот уровень, оставляя в валентной зоне свободный уровень – дырку (рис.1.11,б).

Введение акцепторной примеси в полупроводник повышает концентрацию дырок $p_p > n_p$. Дырки становятся *основными носителями*, а электроны *неосновными*. Полупроводник, в котором основными носителями являются дырки, называется дырочным полупроводником или *полупроводником p-типа* (от латинского positive – положительный).

Концентрация носителей заряда в примесных полупроводниках

При комнатной температуре практически все атомы доноров и акцепторов ионизированы и концентрация основных носителей заряда равна концентрации примесей:

$$n_n \approx N_d, \text{ и } p_p \approx N_a.$$

Произведение концентраций электронов и дырок в примесном полупроводнике равно произведению концентраций носителей в собственном полупроводнике при той же температуре:

$$\begin{aligned} \text{в } n\text{-полупроводнике:} \quad n_n p_n &= n_i p_i = n_i^2, \\ \text{в } p\text{-полупроводнике:} \quad n_p p_p &= n_i p_i = n_i^2. \end{aligned} \quad (1.17)$$

Так как $n_n \gg n_i$ и $p_p \gg p_i$, то концентрации неосновных носителей n_p и p_n в примесных полупроводниках оказываются много

ниже, чем концентрации носителей в собственном полупроводнике n_i и p_i при той же температуре.

Пример. Концентрации носителей заряда в собственном полупроводнике при $T = 300$ К $n_i = p_i \approx 10^{10} \frac{1}{\text{см}^3}$. Концентрации основных n_n и неосновных p_n носителей заряда в n - полупроводнике соответственно равны:

$$n_n = N_d \approx 10^{15} \frac{1}{\text{см}^3}, \quad p_n = n_i p_i / n_n \approx 10^5 \frac{1}{\text{см}^3}.$$

Проводимость примесных полупроводников определяем подстановкой $n_n \approx N_d$, $p_n \approx 0$ и $p_p \approx N_a$, $n_p \approx 0$ в формулу (1.5):

$$\begin{aligned} \sigma_n &= q\mu_n N_d. \\ \sigma_p &= q\mu_p N_a. \end{aligned} \quad (1.18)$$

Компенсированные полупроводники

Кристалл полупроводника может одновременно содержать и донорные, и акцепторные примеси. Если концентрация примесей одинакова, то примеси будут взаимно компенсировать друг друга. Донорные электроны будут занимать акцепторные уровни, и концентрация носителей заряда будет мало отличаться от концентрации носителей в собственном полупроводнике. Такой полупроводник называют *компенсированным*.

Компенсированные полупроводники могут обладать электронной или дырочной электропроводностью. Тип электропроводности определяется превышением концентрации одного типа примеси над другим.

Энергия Ферми

Распределение электронов по энергетическим уровням описывается статистикой Ферми-Дирака. Вероятность того, что разрешённое состояние с энергией W при данной температуре T занято электроном, выражается функцией Ферми

$$F_F(W, T) = \frac{1}{1 + \exp \frac{W - W_F}{kT}}. \quad (1.19)$$

Величина W_F называется *энергией Ферми* или *химическим потенциалом*. На зонной диаграмме *энергия Ферми* изображается в виде *уровня Ферми* (рис.1.12). Уровень Ферми в общем слу-

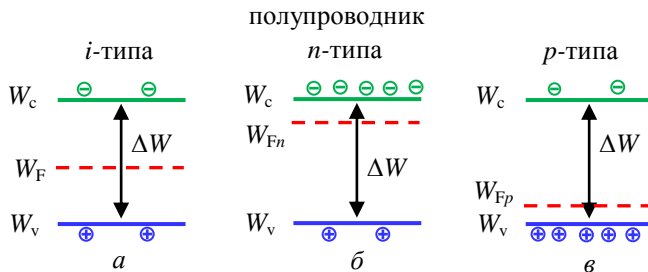


Рис.1.12. Положение уровня Ферми W_F на энергетических диаграммах собственного (а) и примесных (б, в) полупроводников

чае не совпадает ни с одним из разрешённых энергетических уровней электронов в кристалле. Поэтому на уровне W_F не может быть электронов, и величина W_F , введённая в функции распределения (1.19), имеет чисто статистический смысл.

Концентрация свободных носителей заряда в данной зоне определяется расстоянием этой зоны от уровня Ферми: чем меньше это расстояние, тем выше концентрация носителей. В собственных полупроводниках уровень Ферми располагается посередине запрещённой зоны (рис.1.12,а). В электронных полупроводниках уровень Ферми лежит вблизи уровня донорной примеси W_d (рис.1.12,б), а в дырочных – вблизи акцепторного уровня W_a (рис.1.12,в).

При заданной температуре T вероятность заполнения энергетического уровня электроном экспоненциально уменьшается с увеличением энергии W , соответствующей этому уровню (1.19). Таким образом, максимальная концентрация свободных электронов находится практически на уровне дна зоны проводимости. Аналогично, максимум концентрации других свободных носителей заряда – дырок находится вблизи верха валентной зоны.

p–n-переход

В результате технологического процесса в одном кристалле создаются контактирующие между собой области с p - и n -типом проводимости. На рис.1.13 слева изображена зонная диаграмма дырочного, справа – электронного полупроводника. Граница между полупроводниками p - и n -типа называется металлургиче-

ской границей, которая изображена на диаграмме сплошной чёрной линией, и соответствует точке 0 по оси x .

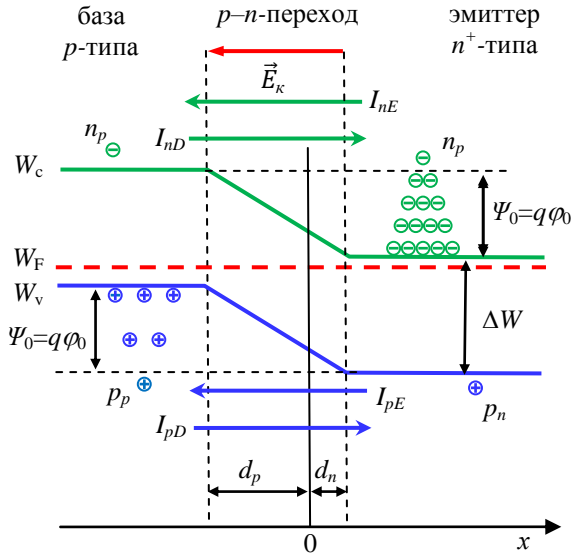


Рис.1.13. Зонная диаграмма несимметричного p - n -перехода с эмиттером электронов в состоянии термодинамического равновесия

При термодинамическом равновесии контактирующих полупроводников выравниваются их уровни Ферми

$$W_{Fp} = W_{Fn}.$$

Выполнение этого условия приводит к относительному смещению энергетических уровней p - и n -полупроводников и возникновению контактной разности потенциалов ϕ_0 :

$$\phi_0 = \phi_T \ln \frac{N_d N_a}{n_i p_i}. \quad (1.20)$$

Величину

$$\phi_T = \frac{kT}{q} \quad (1.21)$$

называют *температурным потенциалом*.

При $T = 300 \text{ К}$ $\phi_T = 1,38 \cdot 10^{-23} \cdot 300 / 1,6 \cdot 10^{-19} = 26 \cdot 10^{-3} \text{ В} = 26 \text{ мВ}$.

На границе раздела p - и n -областей образуется тонкий слой d , обеднённый основными носителями заряда, который называется p - n -переходом. Между p - и n -областями возникают электрическое поле \vec{E} и контактная разность потенциалов φ_0 , созданные нескомпенсированными зарядами донорных и акцепторных примесей в обеднённом слое. Потенциал n -области положительнее по отношению к потенциалу p -области.

Электрическое поле в слое объёмных зарядов

$$|E_k| = \frac{2\varphi_0}{d},$$

выталкивает основные носители (дырки в p -области и электроны в n -области) из p - n -перехода. Концентрация носителей заряда в p - n -переходе приближается к концентрации носителей в собственном полупроводнике.

По обе стороны p - n -перехода расположены области с различными концентрациями основных носителей заряда. Высоколегированная область полупроводника называется *эмиттером*, низколегированная область – *базой*. На рис.1.13 изображена зонная диаграмма p - n -перехода с *эмиттером электронов*. Концентрация донорной примеси в эмиттере обычно составляет $N_d = 10^{17} \div 10^{18} \frac{1}{\text{см}^3}$, что на $2 \div 3$ порядка больше концентрации акцепторной примеси в базе $N_a = 10^{15} \frac{1}{\text{см}^3}$. Полагая, что концентрации носителей заряда в собственном полупроводнике (кремнии) при $T = 300 \text{ К}$ $n_i = p_i \approx 10^{10} \frac{1}{\text{см}^3}$, по формуле (1.20) находим контактную разность потенциалов в переходе $\varphi_0 = 0,7 \div 0,8 \text{ В}$.

В *несимметричном* p - n -переходе *обеднённый носителями заряда* слой d в основном располагается в области низколегированного полупроводника (базы) d_6 . Толщиной обеднённого слоя в высоколегированном эмиттере d_3 , как правило, пренебрегают и считают, что толщина p - n -перехода приблизительно равна толщине обеднённого слоя в базе

$$d = d_6 + d_3 \approx d_6.$$

Движение носителей заряда через p – n -переход в состоянии термодинамического равновесия

Электрическое поле \vec{E} в p – n -переходе препятствует диффузии через переход основных носителей заряда из n - и p -областей. В то же время неосновные носители заряда (электроны в p -области и дырки в n -области), совершая тепловое движение, попадают в область p – n -перехода и перебрасываются электрическим полем (*дрейфуют* через переход) в области, где они являются основными носителями.

Потоки неосновных носителей через переход создают электронный I_{nE} и дырочный I_{pE} дрейфовые токи. Эти дрейфовые токи уравновешены диффузионными токами I_{nD} и I_{pD} основных носителей. Некоторое количество свободных электронов в n -области, обладающих достаточно высокой энергией теплового движения, могут преодолеть действие сил электрического поля и попасть в p -область. Аналогично, “горячие” дырки из p -области диффундируют в n -область. В состоянии термодинамического равновесия каждая из компонент тока и электронная $I_n = I_{nD} - I_{nE}$, и дырочная $I_p = I_{pD} - I_{pE}$ равны нулю и общий ток $I = I_n + I_p$ равен нулю.

Прямое смещение p – n -перехода

Зонная диаграмма прямосмещённого p – n -перехода приведена на рис.1.14. При прямом смещении к p – n -переходу подключается внешний источник напряжения U , плюсом к области p , минусом к области n .

Энергия электронов в области n увеличивается, и потенциальный барьер уменьшается до $\Psi = (\phi_0 - U)q$. При понижении потенциального барьера становится возможной диффузия основных носителей (электронов из n - в p -область и дырок из p - в n -область). Через переход протекает большой диффузионный ток

$$I_D = I_{nD} + I_{pD}.$$

Говорят, что прямосмещённый переход *открыт* для основных носителей.

Одновременно с инжекцией электронов в базу p -типа происходит инжекция дырок в эмиттер n -типа. В рассматриваемом

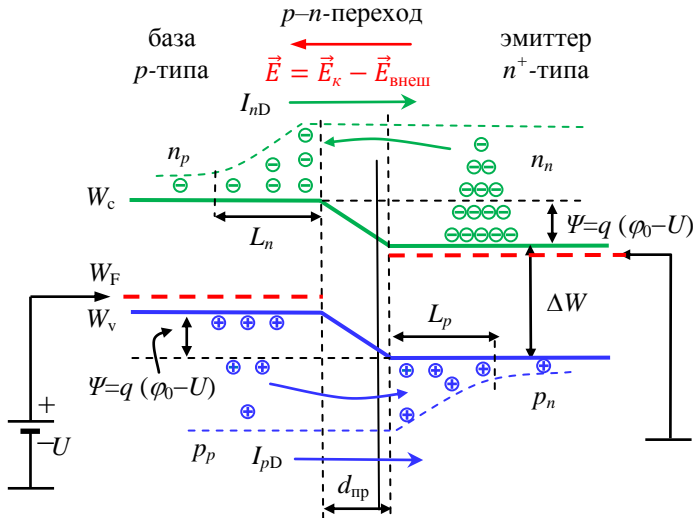


Рис.1.14. Зонная диаграмма прямосмещённого p - n -перехода

случае $n_n \gg p_p$ поэтому инжекция дырок пренебрежимо мала и учитывается только диффузионная составляющая электронного тока I_{nD} .

Обратное смещение p - n -перехода

Зонная диаграмма обратносмещённого p - n -перехода приведена на рис.1.15. При обратном смещении отрицательный полюс источника напряжения U подключается к p -области, а положительный – к n -области. В p - n -переходе внешнее электрическое поле $\vec{E}_{\text{внеш}}$ и внутреннее \vec{E}_k совпадают по направлению. В этом случае потенциальный барьер увеличивается до $\Psi = (\phi_0 + U)q$.

Диффузия электронов из области n в область p и дырок из области p в область n становится невозможной. Диффузионные токи основных носителей $I_{nD} = 0$ и $I_{pD} = 0$.

В то же время неосновные носители заряда (электроны в p -базе и дырки в n -эмиттере), совершая тепловое движение, попадают в область p - n -перехода и переносятся электрическим полем $(\vec{E}_{\text{внеш}} + \vec{E}_k)$ через переход: электроны n_p из базы в эмиттер и дырки p_n из эмиттера в базу. Это явление называют *экстракцией неосновных носителей*. Обратносмещённый p - n -переход

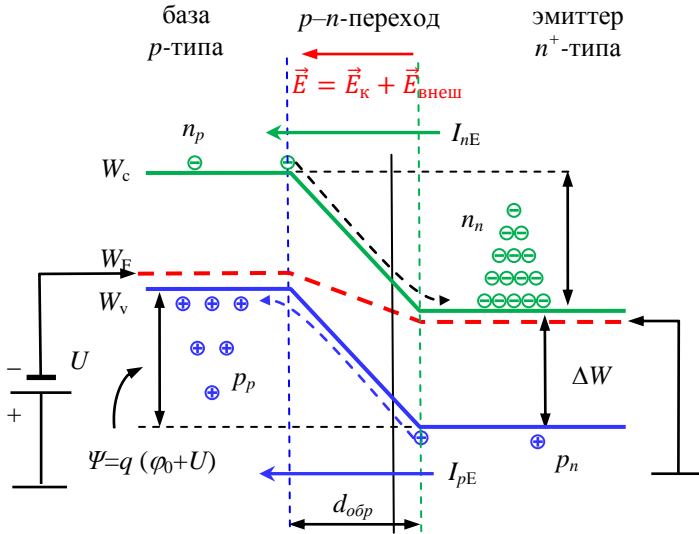


Рис.1.15. Зонная диаграмма обратносмещённого p - n -перехода

закрит для основных носителей и открыт для неосновных. Через него течёт малый обратный ток $I_0 = I_{nE} + I_{pE}$. Проводимость p - n -перехода близка к нулю.

Время жизни неосновных носителей τ_p или τ_n мало и до p - n -перехода могут диффундировать неосновные носители, возникшие в n - и p -областях на расстоянии, не превышающем соответствующей диффузионной длины L_p или L_n . Остальные неосновные носители, не успевшие дойти до перехода, рекомбинируют в объёме полупроводника. Поэтому обратный ток не зависит от напряжения смещения. Обратный ток через p - n -переход называют *током насыщения* или *тепловым током*

$$I_0 = j_0 S, \quad (1.22)$$

где

$$j_0 = q \left(p_n \frac{L_p}{\tau_p} + n_p \frac{L_n}{\tau_n} \right) \quad (1.23)$$

– плотность тока насыщения, S – площадь p - n -перехода.

При увеличении температуры полупроводника плотность тока насыщения увеличивается, так как с температурой экспоненциально растёт концентрация собственных носителей заряда (1.1, 1.2).

Вольт-амперная характеристика p - n -перехода

Сила тока через p - n -переход описывается выражением

$$I = jS = j_0 \left(\exp \frac{U}{\varphi_T} - 1 \right) S, \quad (1.24)$$

где j_0 – плотность тока насыщения (1.23), φ_T – температурный потенциал (1.21), S – площадь p - n -перехода, U – напряжение, приложенное к переходу.

ВАХ $I(U)$ p - n -перехода показана на рис.1.16. При $U = 0$ ток через переход $I = 0$. При приложении прямого напряжения (“плюсом” к p -базе, “минусом” к n -эмиттеру) уменьшается разность потенциалов в переходе $\varphi = \varphi_0 - U$ и возникает инжекция электронов в базу. В случае прямого напряжения $U > \varphi_T$ единиц в формуле (1.24) можно пренебречь и зависимость $I(U)$ будет иметь экспоненциальный характер. При дальнейшем увеличении U напряжение на p - n -переходе стремится к величине контактной разности потенциалов φ_0 , а остальное напряжение падает на металлических контактах и p - и n -областях (в основном на высокоомной области базы):

$$U \approx \varphi_0 + I r_{\bar{6}}. \quad (1.25)$$

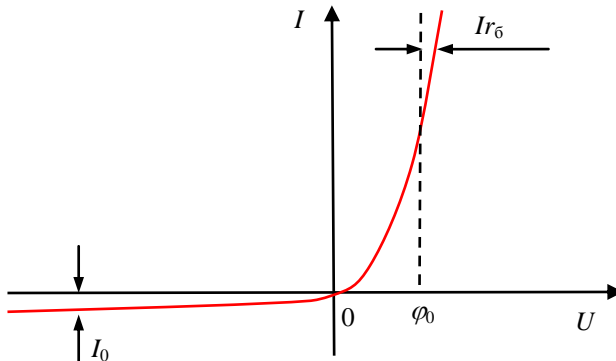


Рис.1.16. Вольт-амперная характеристика p - n -перехода

При достаточно большом значении обратного напряжения $|-U| > 3\varphi_T$ обратный ток через p - n -переход равен $I_0 = -j_0 S$ и практически не зависит от напряжения. Величина I_0 называется

током насыщения обратносмещённого перехода. При увеличении температуры обратный ток возрастает за счёт экспоненциального роста числа генерируемых пар электрон-дырка. Величина обратного тока I_0 меньше у полупроводников с широкой запрещённой зоной (1.1, 1.2).

Один из важных параметров прямой ветви ВАХ – дифференциальное сопротивление p – n -перехода. Дифференцированием формулы (1.24) можно получить

$$r_{p-n} = \frac{dU}{dI} \approx \frac{\varphi_T}{I}. \quad (1.26)$$

Типичная величина $r_{p-n} \approx 25$ Ом при токе $I = 1$ мА.

1.1.5. Фотодиоды, фототранзисторы

Воздействие света на p – n -переход

Зонная диаграмма освещённого p – n -перехода приведена на рис.1.17.

Пусть на поверхность полупроводника нормально к плоскости p – n -перехода падает монохроматический поток света Φ_λ (Дж/с·м²) с длиной волны $\lambda \leq hc/\Delta W$. Число квантов света, вхо-

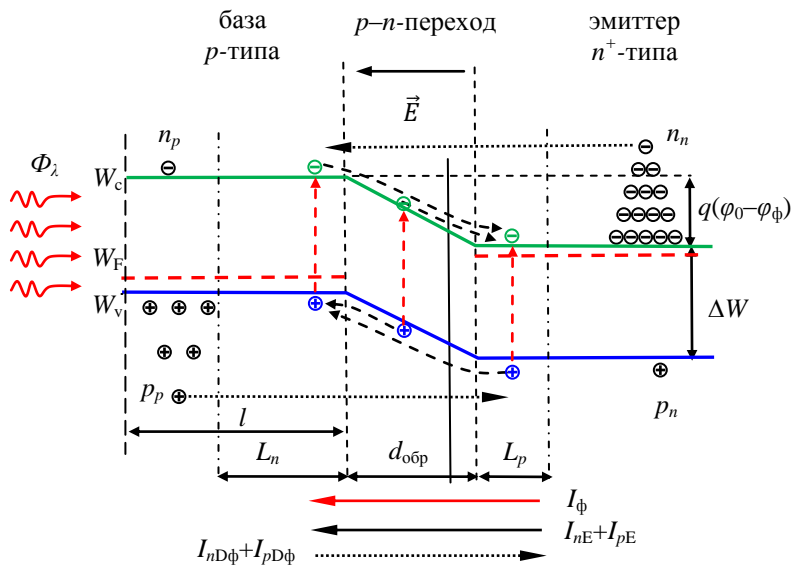


Рис.1.17. Зонная диаграмма освещённого p – n -перехода

дящего в полупроводник в единицу времени, $N = \Phi_\lambda/h\nu$.

При поглощении квантов света в полупроводнике образуются новые (неравновесные (*)) носители заряда – пары электрон–дырка. Неосновные неравновесные носители (свободные электроны в p -базе и дырки в n -эмиттере), возникшие в прилегающих к p - n -переходу областях, не превышающих средней диффузионной длины L_n для электронов и L_p для дырок (**), диффундируют к p - n -переходу и проходят через него под действием электрического поля \vec{E} . При этом возрастает дрейфовый ток через p - n -переход ($I_{nE} + I_{pE}$) на величину I_Φ . Величина фототока I_Φ пропорциональна числу квантов света, поглощённых в слое $L = L_n + d + L_p$ (рис.1.16):

$$I_\Phi = q\beta k\chi SLN = q\beta k\chi SL\Phi_\lambda/h\nu = K_\lambda\Phi_\lambda, \quad (1.27)$$

где β – квантовый выход, k – коэффициент поглощения света, χ – коэффициент переноса, в котором учитывается доля рекомбинировавших носителей заряда в общем количестве носителей, возникших под действием света, S – площадь p - n -перехода.

Таким образом, освещение полупроводника вызывает нарушение термодинамического равновесия. Фототок I_Φ , возникший в результате световой генерации пар электрон-дырка, приводит к увеличению числа основных носителей заряда в n - и p -областях полупроводника, что уменьшает общую разность потенциалов p - n -перехода $\varphi = \varphi_0 - \varphi_\Phi$ и величину потенциального барьера $q\varphi = q(\varphi_0 - \varphi_\Phi)$ для основных носителей. Вследствие этого возрастает диффузионный ток основных носителей $I_{nD\Phi}$ и $I_{pD\Phi}$ (рис.1.17, пунктирная стрелка).

В стационарном состоянии потоки зарядов через p - n -переход в обоих направлениях уравниваются друг друга:

$$-I_\Phi - I_{nE} - I_{pE} + I_{nD\Phi} + I_{pD\Phi} = 0. \quad (1.28)$$

Между p - и n -областями полупроводника устанавливается разность потенциалов – фото ЭДС φ_Φ .

(*) Равновесными называют носители, образовавшиеся в процессе тепловой генерации. Процессы тепловой генерации и рекомбинации уравнивают друг друга. Установившаяся концентрация носителей называется равновесной. Концентрация, превышающая равновесную, называется неравновесной, а дополнительно возникшие носители заряда – неравновесными носителями.

(**) Диффузия неосновных неравновесных носителей – распространение носителей из области с высокой концентрацией в области с пониженной концентрацией. Одновременно с диффузией неравновесных носителей идет процесс их рекомбинации с основными носителями. Диффузионная длина L определяется как расстояние, на котором концентрация неравновесных носителей уменьшается в $e = 2,72$ раз.

Фотодиоды

Фотоэлектрический полупроводниковый прибор с одним p – n -переходом называется *фотодиодом*. Структура фотодиода и его изображение на принципиальных схемах приведены на рис.1.18.

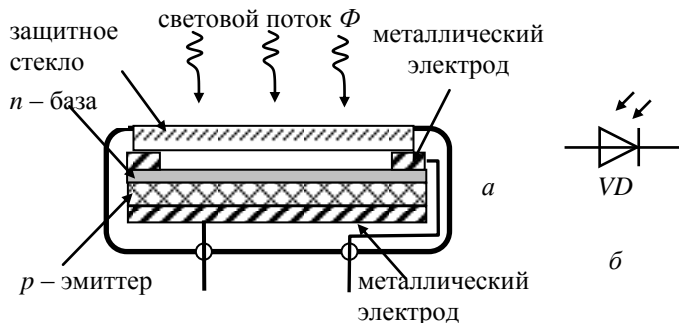


Рис.1.18. Фотодиод: *а* – структура, *б* – изображение на принципиальных схемах

Вольт-амперная характеристика p – n -перехода (рис.1.19) может быть записана в виде:

$$I = I_s \left[\exp \left(\frac{qU}{kT} \right) - 1 \right] - I_\Phi. \quad (1.29)$$

При обратном смещении p – n -перехода резко повышается потенциальный барьер для основных носителей заряда, и величина тока через переход определяется потоком неосновных носителей, которые проходят через него под действием электрического поля \vec{E} . Величина темнового обратного тока через фотодиод равна I_0 ; при освещении она возрастает на величину I_Φ за счёт появления добавочных неосновных носителей заряда:

$$I_{\text{обр}} = I_0 + I_\Phi. \quad (1.30)$$

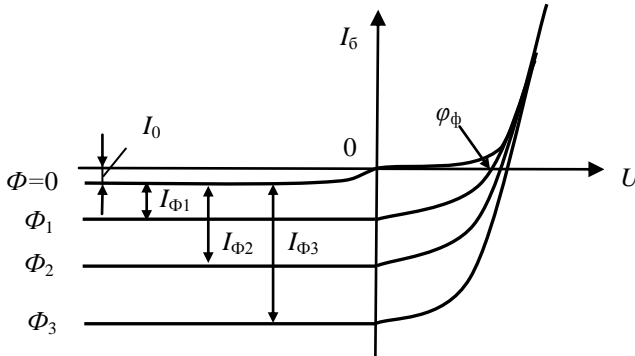


Рис.1.19. Вольт-амперная характеристика фотодиода: без освещения ($\Phi = 0$) и при освещении ($\Phi_1 < \Phi_2 < \Phi_3$)

Величина обратного тока почти не зависит от приложенного напряжения. Световая характеристика фотодиода $I_{\text{обр}} = I_{\text{обр}}(E)$ является линейной в широком диапазоне изменения освещённости E . Обычно толщина базы фотодиода существенно (рис.1.17) меньше средней длины диффузии неосновных носителей заряда $l < L_n$ (фотодиод с “тонкой” базой). Поэтому практически все возникшие в базе в результате световой генерации неосновные носители достигают p - n -перехода и принимают участие в образовании фототока.

При прямом смещении p - n -перехода понижается потенциальный барьер и через p - n -переход начинает протекать ток, созданный диффузией основных носителей заряда. При напряжении $U = \varphi_{\text{ф}}$ диффузионный ток полностью компенсирует дрейфовый ток, и результирующий ток через переход становится равным нулю (1.28).

При $U > \varphi_{\text{ф}}$ величина прямого тока быстро возрастает и вольт-амперные характеристики освещённого и затемнённого фотодиодов практически совпадают (рис.1.19).

В большинстве случаев фотодиоды используют в режиме обратного смещения p - n -перехода (рис.1.20). Если световой поток модулирован, то на сопротивлении на-

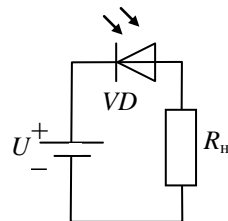


Рис.1.20. Схема включения фотодиода

грузки R_n появляется напряжение сигнала, пропорциональное изменению светового потока (1.27, 1.30):

$$U_c = I_{обр}R_n = (I_0 + I_\phi)R_n = I_0R_n + K_\lambda \Phi_\lambda R_n. \quad (1.31)$$

Преимуществом фотодиодов перед фоторезисторами является их малая инерционность. Инерционность диодов определяется в основном временем диффузии неосновных носителей заряда через базу и достигает порядка десятков наносекунд.

Одним из основных параметров фотодиода является *интегральная чувствительность*

$$K_{инт} = I_\phi / E, \quad (1.32)$$

где E – освещённость поверхности диода.

Спектральная характеристика фотодиодов (рис.1.6) при больших длинах волн ограничивается шириной запрещённой зоны $\lambda_{порог} = hc/\Delta W$, в области малых длин волн – увеличением показателя поглощения и скорости поверхностной рекомбинации носителей заряда.

Фотодиоды находят широкое применение в измерительной технике (фотометрия), автоматике, устройствах сигнализации, волоконно-оптических линиях связи, вычислительной технике (фотосчитывающие устройства CD, сканеры) и т.д.

Полупроводниковые фотоэлементы

В полупроводниковом фотоэлементе с p – n -переходом световая энергия преобразуется в электрическую. Он представляет собой плоскую пластину кремния с электропроводностью p -типа, на поверхности которой создан тонкий слой с проводимостью n^+ -типа. От фотодиода отличается значительно большей площадью p – n -перехода.

Вольт-амперная характеристика фотоэлемента аналогична ВАХ фотодиода (рис.1.19). Части кривых, расположенные в четвёртом квадранте, соответствуют режиму генерации фото-ЭДС.

Фотоэлемент работает без внешних источников питания и сам является источником электрической энергии (четвёртый квадрант на ВАХ p – n -перехода – рис.1.21,а). Если к фотоэлементу подключить внешнюю нагрузку (рис.1.21,б), то при освещении p – n -перехода по цепи потечёт ток. Точки пересечения

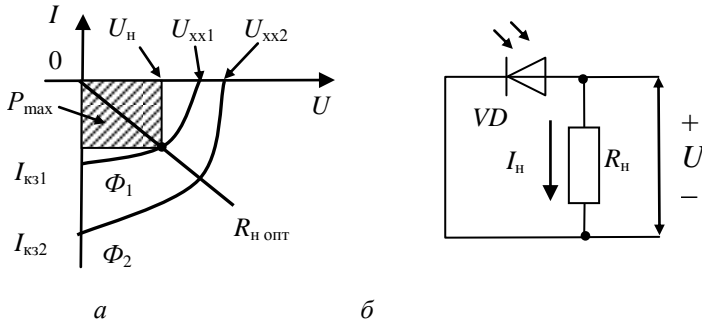


Рис.1.21. Фотозлемент: *a* – вольт-амперная характеристика при освещении ($\Phi_1 < \Phi_2$); *б* – схема включения

ВАХ с осью токов ($U = 0$) соответствуют значениям токов короткого замыкания ($R_n = 0$). Из уравнения (1.29) при $U = 0$ и уравнения (1.32) получим

$$I_{кз} = -I_{\phi} = -K_{инт}E. \quad (1.33)$$

Величина тока короткого замыкания зависит от площади p - n -перехода фотозлемента. У кремниевых фотозлементов плотность тока короткого замыкания $I_{кз}/S = 20 \div 25 \text{ мА/см}^2 = 200 \div 250 \text{ А/м}^2$ при средней освещённости солнечным светом.

Точки пересечения ВАХ с осью напряжений (при $R_n = \infty$) соответствуют значениям фото-ЭДС φ_{ϕ} или напряжениям холостого хода $U_{хх}$. Приняв в уравнении (1.29) $I = 0$, можно найти напряжение холостого хода:

$$U_{хх} = \varphi_{\phi} = \varphi_T \ln \left(\frac{I_{\phi}}{I_0} + 1 \right) = \varphi_T \ln \left(\frac{K_{инт}E}{I_0} + 1 \right), \quad (1.34)$$

где $I_0 = I_{nE} + I_{pE}$ – обратный ток p - n -перехода.

При большой интенсивности света максимальное значение фото-ЭДС стремится к величине контактной разности потенциалов p - n -перехода φ_0 . У кремниевых фотозлементов фото-ЭДС составляет $0,5 \div 0,55 \text{ В}$.

По ВАХ фотозлемента можно выбрать оптимальное сопротивление нагрузки $R_{н опт}$, при котором в нагрузке будет выделяться максимальная мощность электрического тока P_{max} . Для этого необходимо провести нагрузочную прямую из начала координат $I = -U/R_n$ (рис.1.21,*a*). Точка пересечения ВАХ с нагруз-

зочной прямой определяет напряжение U_n , ток нагрузки I_n и мощность в нагрузке $P_n = U_n \cdot |I_n|$.

Для кремниевых фотоэлементов при оптимальной нагрузке $U_n = 0,35 \div 0,4$ В и $j = 15 \div 20$ мА/см². Таким образом, при средней освещённости электрическая мощность, генерируемая одним квадратным сантиметром фотоэлемента, составляет $P = 5 \div 8$ мВт.

Световые характеристики фотоэлемента (рис.1.22) показывают зависимость напряжения холостого хода $U_{xx} = \varphi_{\phi}$ и тока короткого замыкания $I_{кз} = I_{\phi}$ от освещённости E , создаваемой лампой накаливания с температурой вольфрамовой нити $T = 2840$ К. Зависимость $I_{кз}$ от освещённости линейна (рис.1.22, сплошная линия), а напряжение холостого хода пропорционально логарифму освещённости (рис.1.22, штриховая линия).

Отношение $K_{\text{инт}} = I_{кз}/E$ называется *интегральной чувствительностью фотоэлемента*.

Спектральная характеристика фотоэлемента определяет зависимость величины тока короткого замыкания от длины волны падающего света: $I_{кз} = I_{кз}(\lambda)$ при $E = \text{const}$. Максимум спектральной характеристики кремниевых фотоэлементов (рис.1.6) почти совпадает с максимумом спектрального распределения энергии солнечного света. Поэтому кремниевые фотоэлементы широко используют для создания солнечных батарей.

Коэффициент полезного действия фотоэлемента $\eta = P_{\text{max}}/P_{\phi}$, где P_{ϕ} — мощность светового потока, падающего на

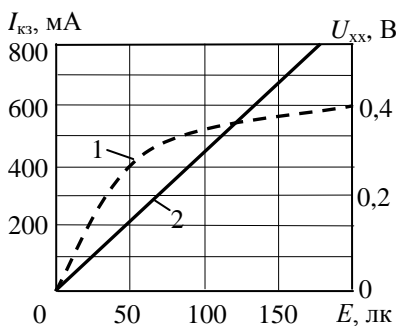


Рис.1.22. Световые характеристики фотоэлемента: 1 — $U_{xx} = U_{xx}(E)$, 2 — $I_{кз} = I_{кз}(E)$

рабочую поверхность фотоэлемента. Теоретический предел КПД кремниевых фотоэлементов составляет 23%.

Уменьшение КПД фотоэлементов происходит из-за частичного отражения излучения от поверхности полупроводника, поглощения квантов света без образования пар электрон — дырка, рекомбинации неравновесных носителей

до их разделения электрическим полем в p – n -переходе, потери мощности при прохождении тока через объёмное сопротивление базы фотозлемента. В результате КПД кремниевых фотозлементов, выпускаемых промышленностью, составляет порядка 12%.

Биполярный фототранзистор

Биполярный фототранзистор включают по схеме с общим эмиттером (рис.1.23). У фототранзистора обычно нет базового электрода. Падающий свет облучает область базы. При этом в базе происходит процесс световой генерации пар электрон – дырка, как показано на зонной диаграмме (рис.1.24). Дырки в базе оказываются сосредоточенными в потенциальной яме и не могут покинуть базу. Накопившийся дырочный заряд понижает потенциальный барьер перехода база-эмиттер, и электроны из эмиттера инжектируются в базу, где они являются неосновными носителями. Электроны диффундируют через область базы и

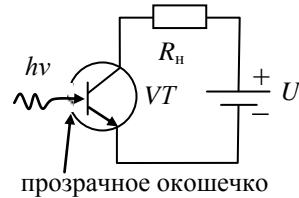


Рис.1.23. Схема включения фототранзистора

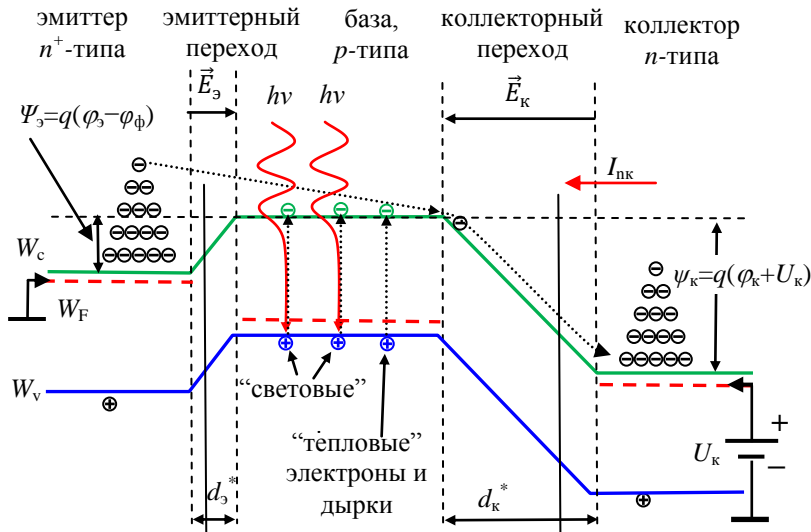


Рис.1.24. Зонная диаграмма фототранзистора n – p – n -типа включенного по схеме с общим эмиттером

втягиваются электрическим полем обратносмещённого коллекторного перехода. Попад в область коллектора, электроны становятся основными носителями и создают ток во внешней цепи. Таким образом, базовым током фототранзистора является фототок I_{ϕ} , который приводит к появлению электронной составляющей коллекторного тока $I_{\text{лк}} = \beta I_{\phi}$, где β – коэффициент усиления базового тока транзистора. Дырочная составляющая коллекторного тока $I_{\text{ко}}$ мала и на зонной диаграмме не показана.

Поскольку в фототранзисторе происходит усиление фототока, его интегральная чувствительность $K_{\text{инт}} = \beta I_{\phi}/E$ в β раз больше интегральной чувствительности фотодиода.

Недостатком фототранзистора без базового электрода является сильная температурная зависимость коллекторного тока, так как дырочный заряд в базе накапливается в результате не только световой, но и тепловой генерации пар электрон – дырка. И “тепловую” дырку нельзя отделить от “световой”. Некоторые типы фототранзисторов имеют электрический вывод базы. Их включают по схеме с температурной стабилизацией коллекторного тока, но при этом фоточувствительность транзисторов падает.

1.2. Фоточувствительные полупроводниковые приборы с зарядовой связью

1.2.1. Структура и принцип действия приборов с зарядовой связью

Прибор с зарядовой связью (ПЗС, англ. CCD – Charge-Coupled Device) – это полупроводниковый прибор, в котором электрический сигнал представлен пакетами подвижных носителей заряда, (а не током или напряжением, как в большинстве электронных приборов). Передача сигнала осуществляется управляемым переносом зарядовых пакетов между элементами прибора.

На рис.1.25 приведена конструкция ПЗС на пластине (подложке) низколегированного монокристалла кремния p -типа. Верхняя поверхность монокристалла кремния покрыта изолирующим слоем двуокиси кремния. На его поверхности размещён набор электродов (затворов), образующих линейки или

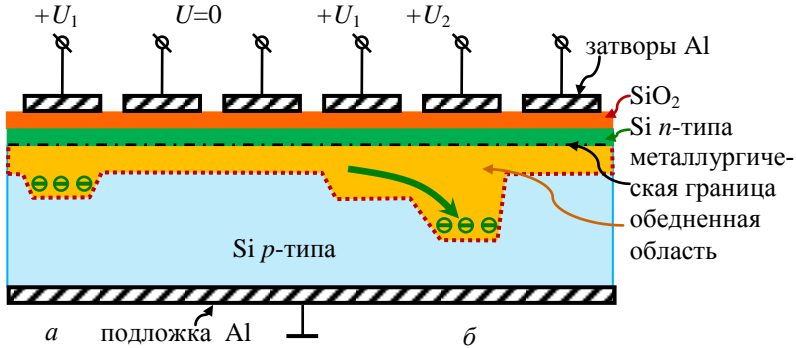


Рис.1.25. Процессы, происходящие в ПЗС: а – образование потенциальной ямы, б – перенос зарядового пакета

матрицы близко расположенных друг к другу МОП-конденсаторов – ячеек ПЗС. МОП или МДП-конденсатором называют структуру из металлической обкладки, слоя диэлектрика и обкладки в виде слоя полупроводника.

На p -подложке ионным легированием сформирован тонкий ($3\div 4$ мкм) n -слой с концентрацией донорной примеси, немного превышающей концентрацию акцепторов в подложке. Электроны проводимости n -слоя рекомбинировали с дырками. Электрическое поле объёмных зарядов ионов донорной примеси выталкивает дырки от поверхности вглубь подложки, поэтому у поверхности образуется слой (канал), обеднённый носителями заряда.

Если на затвор одной из ячеек подать большое ($U_1 = 10\div 30$ В относительно подложки) напряжение положительной полярности, то в полупроводнике под затвором возникает *потенциальная яма* для электронов. Попавшие в эту область электроны будут скапливаться вблизи металлургической границы между n -слоем и p -подложкой (рис.1.25,а).

Максимальная величина накопленного заряда (зарядового пакета)

$$Q_{n \max} = C_{\text{МОП}} U_1,$$

где $C_{\text{МОП}} = \epsilon_0 \epsilon S/d$ – ёмкость МОП-конденсатора, S – площадь затвора, d – глубина залегания металлургической границы.

Для большинства матриц максимальная величина зарядового пакета составляет 10^5 электронов.

Если теперь к затвору соседней ячейки приложить напряжение $U_2 > U_1$, то под ним образуется более глубокая потенциальная яма и электроны переместятся в неё (рис.1.25,б). Мелкая потенциальная яма под первым затвором “опустошается”. Последовательно прикладывая к цепочке затворов управляющие напряжения, можно переносить локализованный зарядовый пакет вдоль поверхности ПЗС. Канал переноса зарядовых пакетов ограничивают с боков *стоп-каналами* – узкими полосками p^+ -полупроводника с повышенной концентрацией акцепторной примеси (рис.1.26).

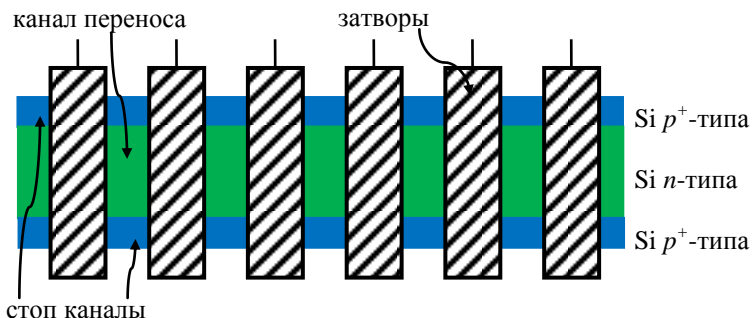


Рис.1.26. Конструкция канала переноса зарядовых пакетов

Сдвиговый регистр на ПЗС

Регистры сдвига на полупроводниковых приборах – триггерах широко используются для преобразования параллельного кода в последовательный и обратно. Регистры с последовательным вводом и выводом служат также в качестве элементов задержки сигнала, представленного в цифровой форме.

Ячейка ПЗС, как и триггер, может находиться в двух состояниях, определяемых наличием или отсутствием зарядового пакета. На рис.1.27,а показана структура двухрядного сдвигового регистра на ПЗС с последовательными вводом и выводом.

ПЗС-регистр имеет устройство ввода зарядовых пакетов, секцию переноса пакетов и устройство вывода. Устройство ввода (n^+ -исток и входной затвор) преобразует импульсы входного

напряжения в последовательность зарядовых пакетов. Обратное преобразование осуществляет устройство вывода, содержащее n^+ -сток и выходной усилитель на МДП-транзисторах, который создаётся на этой же подложке.

Секция переноса состоит из ряда ячеек ПЗС. Каждая три ячейки образуют один разряд регистра. Затворы ячеек соединены с тремя шинами синхронизации, на которые подаются импульсы напряжения, сдвинутые по фазе на $1/3$ длительности такта (рис.1.28).

Такты диаграммы работы сдвигового регистра показаны на рис.1.27,б-д. Во время первого синхроимпульса, $U_1 > 0$ под первым (и четвёртым) затвором возникает потенциальная яма для электронов. Если при этом на входной затвор подать импульс положительного напряжения $U_{вх}$ (логическая “1”), достаточного для образования проводящего канала, то электроны из истока будут проходить по каналу под входным затвором и накапливаться в потенциальной яме. Образование зарядового пакета под первым затвором – это запись логической “1” в первый разряд регистра (рис.1.27,б).

Для записи логического “нуля” в первый разряд регистра на входной затвор должно быть подано отрицательное напряжение. В этом случае не будет инжекции электронов из n^+ -истока в потенциальную яму под первым затвором (рис.1.27,д) и в ней может оказаться только небольшой отрицательный заряд, связанный с тепловой генерацией неосновных носителей в подложке.

Во время второго синхроимпульса U_2 зарядовый пакет перетекает в потенциальную яму под вторым затвором (рис.1.27,в). Потенциальная яма под первым затвором “опустошается”. Таким образом, при повышении потенциала на одной шине и понижении его на предыдущей происходит одновременная передача всех зарядовых пакетов под соседние затворы. За полный цикл (один такт на каждой шине) произойдёт сдвиг зарядовых пакетов на один разряд регистра.

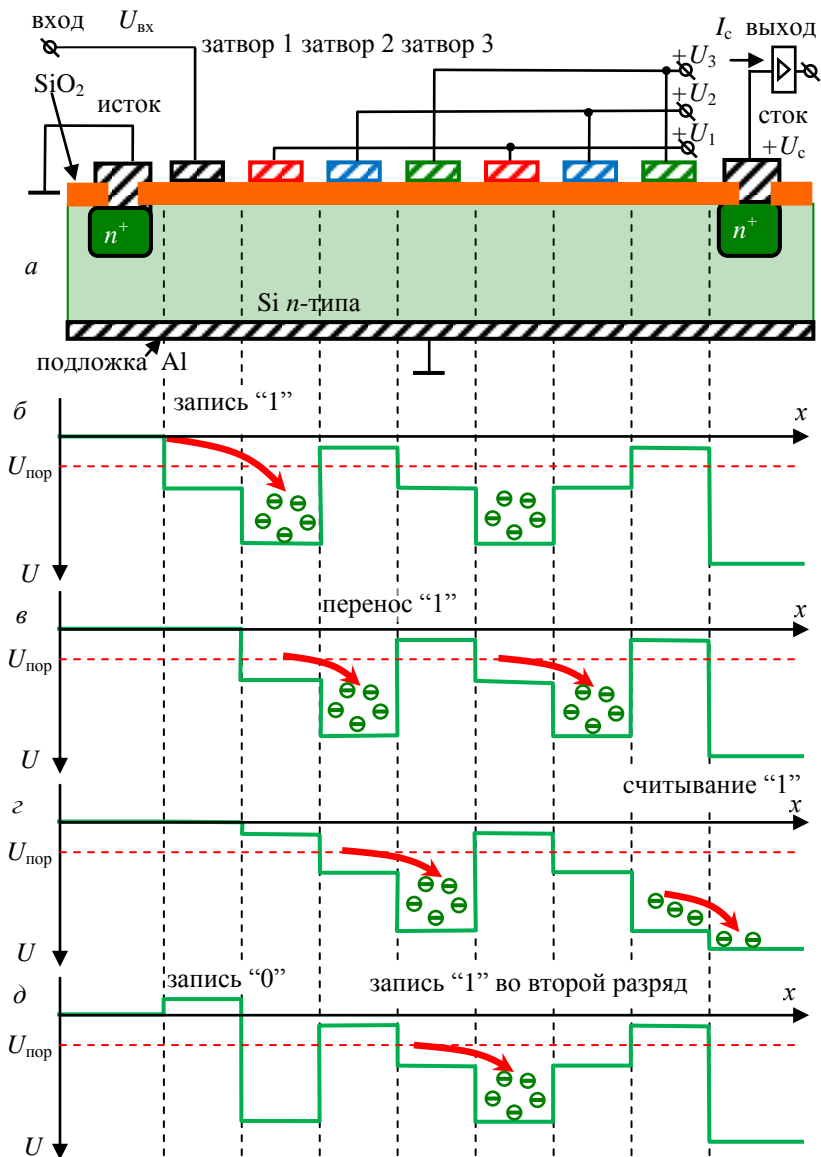


Рис.1.27. Структура сдвигового регистра – а и такты работы: б – запись "1", в – перенос "1", г – считывание "1", д – запись "0"

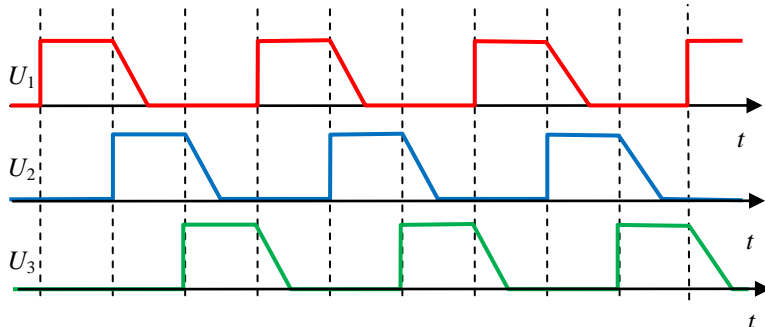


Рис.1.28. Тактовые диаграммы трёхтактового сдвигового регистра

Сток имеет более высокий потенциал по сравнению с затворами секции переноса. Если потенциальная яма, содержащая пакет электронов, во время третьего синхроимпульса U_3 подойдёт к p - n -переходу стока, произойдёт экстракция электронов из потенциальной ямы, и в цепи стока возникнет импульс тока (рис.1.27,з). Появление импульса тока стока означает считывание логической “1”. Если потенциальная яма под последним затвором окажется пустой (нет зарядового пакета), то отсутствие импульса тока стока во время третьего синхроимпульса означает считывание логического “нуля”.

Импульсы тока стока через усилитель-формирователь передаются на выход регистра.

Параметры приборов с зарядовой связью

ПЗС являются динамическими устройствами и могут работать при тактовых частотах, лежащих в определённом диапазоне.

Нижний предел тактовой частоты (единицы – десятки килогерц) связан с накоплением паразитных зарядов в потенциальных ямах. В основном это результат термогенерации электронно-дырочных пар в обеднённом слое. В зависимости от температуры и чистоты полупроводника накопление паразитного заряда, сравнимого по величине с информационным, может происходить за время от сотых долей до единиц секунд. Сдвиг потенциальных ям должен происходить на несколько порядков быстрее.

Максимальная тактовая частота ПЗС составляет десятки мегагерц и определяется временем перетекания заряда из одной потенциальной ямы в другую.

При прохождении зарядового пакета по цепочке потенциальных ям часть носителей захватывается ловушками. Уменьшение информационного заряда в заполненных ямах и накопление заряда в пустых приводят к ограничению максимальной длины ПЗС (до 10^4 затворов).

Благодаря непосредственной зарядовой связи между соседними элементами в ПЗС не нужны сигнальные проводники, как в микросхемах с транзисторными триггерами. На поверхности кристалла располагаются только затворы и управляющие шины, а сигнальные проводники используются только на входах и выходах ПЗС.

1.2.2. Фоточувствительные линейные ПЗС

Основное назначение фоточувствительных приборов с зарядовой связью ФПЗС – преобразование оптических изображений в последовательность электрических импульсов (формирование видеосигнала).

ФПЗС обычно содержат от нескольких тысяч до нескольких миллионов фоточувствительных элементов (пикселей). Размеры пикселей одинаковы и могут быть от единиц до десятков микрон. Пиксели могут быть выстроены в один ряд – тогда ФПЗС называется ФПЗС-линейкой (ЛФПЗС) или ровными рядами заполнять участок поверхности – тогда ФПЗС называют ФПЗС-матрицей (МФПЗС).

Основная область применения линейных фоточувствительных ПЗС – сканеры, а также спектро-анализаторы и другое научно-исследовательское оборудование.

Линия (строка) изображения выделяется щелевым экраном и фокусируется линзой на ФПЗС-линейке. Развёртка изображения по вертикали осуществляется механическим сканированием, в результате чего последовательно передаются все строки изображения.

В основе работы ФПЗС лежит явление внутреннего фотоэффекта. ФПЗС работают в трёх режимах: восприятие (интегрирование) изображения, т. е. преобразование светового потока в

зарядовые пакеты; хранение зарядовых пакетов; передача зарядовых пакетов на выход устройства.

Все три режима работы устройства могут последовательно выполняться одной линейкой ПЗС (рис.1.29,а). Каждый светочувствительный элемент ПЗС (пиксель) образован тремя ячейками ПЗС с затворами из поликристаллического кремния. Затворы из поликристаллического кремния прозрачны почти во всём видимом диапазоне.

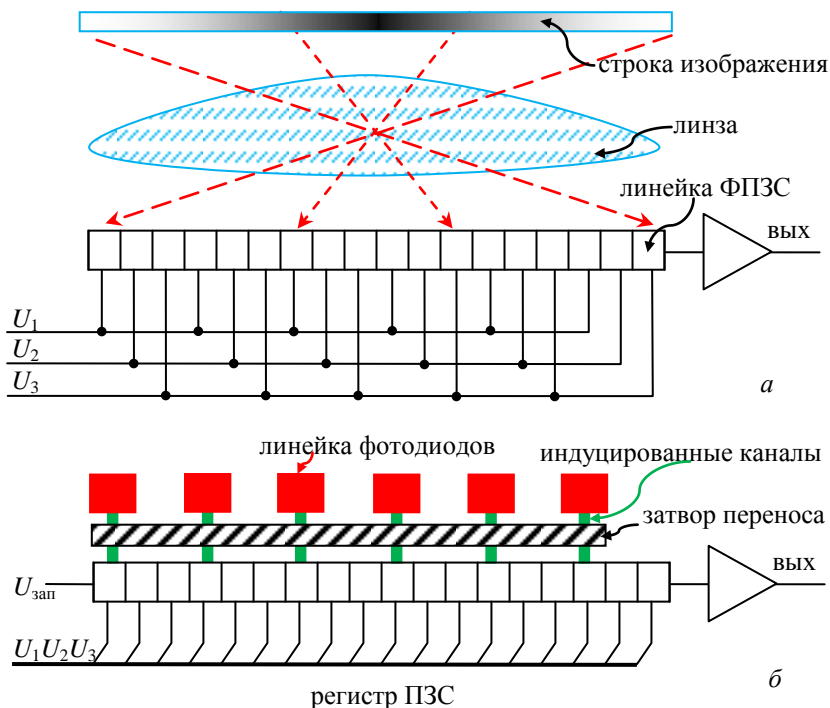


Рис.1.29. Линейные ФПЗС: а – состоящий из одного регистра ПЗС, б – с линейкой фотодиодов, отделенной от сдвигового регистра ПЗС

В режиме восприятия изображения положительный потенциал U_1 устанавливается на каждом третьем затворе. Все остальные затворы находятся при нулевом потенциале.

Если на ПЗС направить световой поток, то вблизи поверхности полупроводника начнется фотогенерация электроно-

дырочных пар. Электроны накапливаются в потенциальных ямах под затворами, к которым подведён положительный потенциал, а дырки вытесняются электрическим полем вглубь подложки. При этом величина накапливаемого заряда будет пропорциональна локальной освещённости, и по истечении некоторого времени (порядка нескольких миллисекунд) в ПЗС возникнет картина зарядовых пакетов, соответствующая изображению, спроектированному на ПЗС.

При включении тактовых импульсов U_1 , U_2 , U_3 ПЗС переходит в режим сдвигового регистра. Зарядовые пакеты перемещаются к выходному устройству считывания. В результате на выходе получается видеосигнал – последовательность электрических импульсов с разной амплитудой, огибающая которых соответствует распределению освещённости вдоль линейки ПЗС.

Основным недостатком ФПЗС, в котором линейка ПЗС попеременно выполняет функции накопления и сканирования, является “смазывание” изображения, возникающее при переносе зарядовых пакетов к выходу регистра. Во время переноса ячейки ФПЗС остаются открытыми для света и яркие участки изображения успевают дать вклад в чужой зарядовый пакет даже за то короткое время, когда он проходит через них.

Для исключения “смаза” линейку фоточувствительных элементов – фотодиодов выполняют отдельно от регистра сдвига (рис.1.29,б). На поверхности p -подложки создаются n^+ -области – светочувствительные площадки. При освещении фотодиода электрическое поле p - n -перехода выталкивает дырки в p -подложку, а электроны накапливаются в n -слое у поверхности полупроводника (рис.1.30).

Строка фотодиодов отделена затвором переноса от ПЗС-регистра сдвига. На затвор переноса подаётся короткий импульс напряжения $U_{\text{зат}}$ положительной полярности (аналогично записи “1” в сдвиговый регистр, показанный на рис.1.27). Под затвором возникают индуцированные n -каналы, и все заряды, накопленные в фоточувствительных элементах, параллельно переносятся в ячейки сдвигающего регистра (рис.1.30).

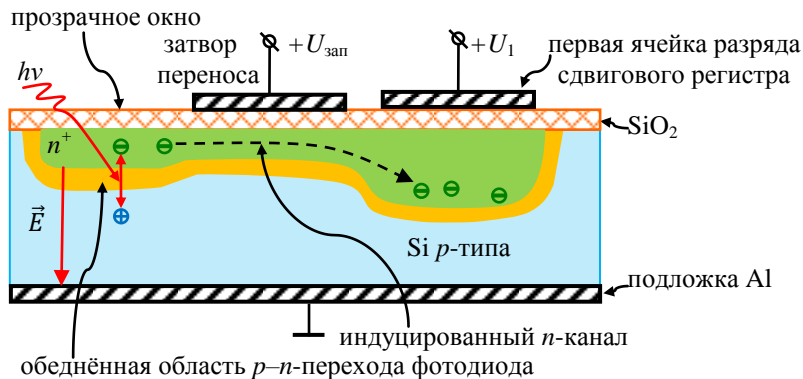


Рис.1.30. Конструкция фотодиодной ячейки ФПЗС

Сразу после запираания затвора в фоточувствительных элементах начинает накапливаться заряд следующей строки, а предыдущая строка изображения считывается с помощью регистра сдвига, который экранирован от падающего света дополнительным слоем металлизации поверхности.

Достоинства ФПЗС с отдельной строкой фотодиодов:

- отсутствие “смаза” изображения,
- длительность режима накопления равна 100% времени цикла.

Пиксели ФПЗС, воспринимающего цветное изображение, состоят из трёх фотодиодов, светочувствительные площадки которых закрыты красным, зелёным и синим светофильтрами.

1.2.3. Фоточувствительные ПЗС-матрицы

Используются три основных типа организации матричной структуры фоточувствительных приборов с зарядовой связью: полнокадровая (ПК), с кадровым переносом (КП) и со строчным переносом (СП). Эти организации отличаются способом считывания картины зарядовых пакетов.

Полнокадровая матричная структура

Полнокадровая матричная структура (full-frame, FF CCD-matrix) образована совокупностью регистров ФПЗС – вертикальных столбцов матрицы (рис.1.31). Регистры отделены друг от друга стоп-каналами – полосками p^+ -полупроводника (см.

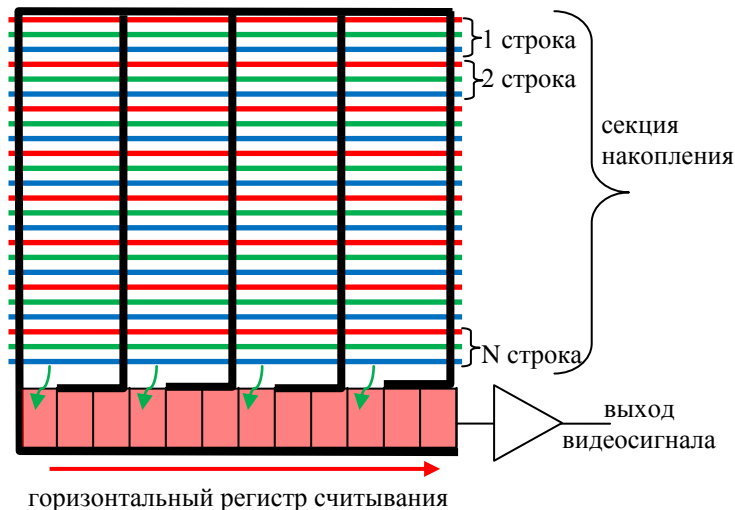


Рис.1.31. Полнокадровая ПЗС-матрица

рис.1.26). На рис.1.31 они изображены вертикальными жирными линиями.

Каждый разряд регистров (пиксель) образован тремя затворами. Одинаковые затворы соединены между собой (затворы конструктивно являются полосками из поликристаллического кремния, проходящими горизонтально над всеми регистрами). Совокупность разрядов регистров с тройками соединённых затворов образуют горизонтальные строки матрицы.

В течение времени накопления зарядовых пакетов на затворы матрицы подаются постоянные напряжения, формирующие потенциальные ямы только под первыми затворами во всех разрядах регистров. Изображение, проецируемое на поверхность матрицы, вызывает фотогенерацию электронно-дырочных пар. Фотогенерируемые электроны скапливаются в потенциальных ямах, а дырки уходят в подложку. Таким образом, под действием света в пикселях накапливается *зарядовый рельеф*, т.е. в каждом пикселе собирается заряд, пропорциональный его освещённости во время накопления.

По окончании накопления на затворы матрицы подаются тактовые импульсы, вызывающие синхронный перенос зарядов

во всех вертикальных регистрах. С каждой тройкой тактовых импульсов картина зарядовых пакетов сдвигается на одну строку вниз.

Из последней строки матрицы зарядовые пакеты передаются в горизонтальный регистр считывания. Каждый разряд регистра считывания принимает заряд из соответствующего столбца матрицы. Регистр считывания – это сдвиговый ПЗС регистр с параллельным вводом и последовательным выводом. Зарядовые пакеты последовательно передаются в выходной усилитель. По окончании вывода одной строки регистр считывания принимает зарядовые пакеты следующей строки и так построчно формируется видеосигнал, отображающий зарядовую картину в матрице.

Полнокадровая матричная структура применяется при регистрации неподвижных изображений. Преимуществом такой организации являются простота и однородность конструкции; фоточувствительная матрица занимает всю площадь кристалла. Основной недостаток – необходимость перекрывать (выключать) световой поток во время считывания накопленного кадра.

Матрица с кадровым переносом

Матрица с кадровым переносом, КП (frame transfer, FT CCD-matrix) содержит две секции – фоточувствительную секцию накопления и закрытую от света секцию хранения (рис.1.32). Обе секции имеют одинаковое число строк, а их столбцы образованы непрерывными регистрами сдвига. По окончании времени накопления кадра на обе секции подаются тактовые импульсы, вызывающие синхронный перенос зарядовых пакетов во всех столбцах матриц. После числа тактов, равного числу строк в каждой секции (каждая строка образована тремя полосками затворов), весь накопленный зарядовый рельеф целиком переместится в секцию хранения. Секция накопления полностью очищается от зарядов. Этот перенос из секции в секцию происходит достаточно быстро и составляет малую часть времени накопления.

Во время следующего цикла накопления секция накопления накапливает следующий кадр изображения.

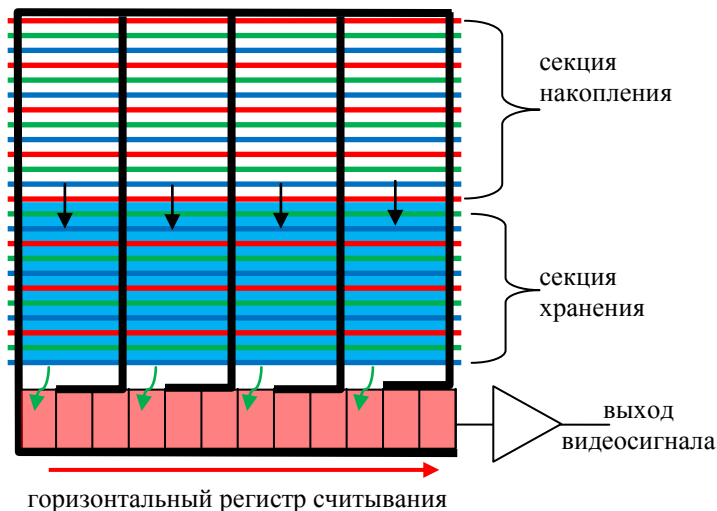


Рис.1.32. Полнокадровая ФПЗС-матрица с переносом кадра

Секция хранения закрыта от света. Во время накопления следующего кадра заряды из секции хранения построчно передаются в горизонтальный регистр считывания, формируя на выходе ФПЗС матрицы видеосигнал.

Достоинство матрицы с кадровым переносом – высокая светочувствительность, поскольку вся площадь секции накопления является светочувствительной. Видеосигнал формируется в виде непрерывной последовательности кадров, время накопления практически равно длительности кадра.

Матрицу отличают топологическая простота и регулярность структуры. Электроды затворов формируются в виде поперечных полосок из поликристаллического кремния, пересекающих весь кристалл. Каналы переноса зарядов в вертикальных регистрах отделяются друг от друга стоп-каналами. Секция хранения и регистр считывания защищаются от света с помощью дополнительной металлизации. Однако за счёт дополнительной секции хранения, содержащей такое же количество ПЗС, что и светочувствительная секция, общее количество элементов увеличивается вдвое.

Основной недостаток матрицы с КП – возникновение “смаза” изображения – появление вертикального следа от ярких де-

талей изображения. Во время кадрового переноса (переноса зарядовых пакетов из секции накопления в секцию хранения), занимающего доли миллисекунды, секция накопления остаётся освещённой и яркие участки изображения могут дать вклад в чужой зарядовый пакет за то короткое время, когда он проходит через них.

Матрицы со строчным переносом

Радикально проблема “смаза” решена в ФПЗС со строчным переносом, СП (interline transfer, IT CCD-matrix) (рис.1.33).

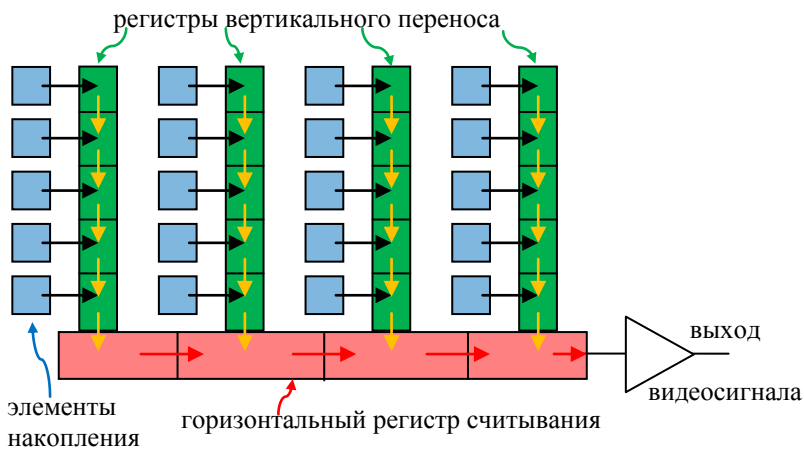


Рис.1.33. Архитектура ФПЗС-матрицы с переносом строк

Вертикальные столбцы матрицы образованы линейками фоточувствительных элементов (фотодиодов) и закрытыми от света регистрами сдвига (аналогично линейному ФПЗС, показанному на рис.1.29). Перенос зарядовых пакетов из всех фоточувствительных элементов в разряды регистров происходит за один такт, и “смаз”, связанный с переносом, не возникает.

Изображение следующего кадра (в виде картины зарядовых пакетов) накапливается в “очищенных от заряда” фотодиодах. В это же время предыдущий кадр сдвигается в экранированных от света вертикальных сдвиговых регистрах и построчно считывается выходным сдвиговым регистром.

Достоинством организации ФПЗС со строчным переносом является уменьшение количества переносов, так как для сдвига всей картины зарядовых пакетов в секцию хранения требуется только один перенос.

По сравнению с матрицами с КП фоточувствительность матриц с СП примерно вдвое меньше, поскольку около половины площади поверхности закрыто от света. “Смаз” изображения в матрице со строчным переносом пренебрежимо мал.

Пиксель цветной матрицы состоит из комбинации трёх пикселей, каждый из которых регистрирует свет либо в красной (Red), либо в зелёной (Green), либо в голубой (Blue) части оптического спектра. Таким образом, с каждого элемента цветной ПЗС-матрицы снимется три сигнала (RGB-сигнал).

1.2.4. Основные характеристики ФПЗС

Основными фотоприёмными характеристиками ФПЗС являются разрешающая способность, светочувствительность, спектральная область чувствительности, шумы, динамический диапазон.

Разрешающую способность характеризуют общим количеством элементов изображения, воспринимаемых всем ФПЗС. Подавляющее число моделей имеют стандартное число элементов, ориентированное на телевизионный стандарт: $(576 \text{ строк}) \times (768 \text{ элементов в строке}) = 442368$ пикселей. ФПЗС матрицы для профессиональной фотографии, космической и аэрофотосъёмки содержат 10^7 и более элементов.

Спектральная характеристика ФПЗС – зависимость выходного сигнала от длины волны падающего излучения. Спектральная характеристика ПЗС определяется двумя факторами – поглощением света при прохождении через электродную структуру и фотогенерацией в полупроводнике.

Поликристаллический кремний, из которого сделаны затворы, совершенно непрозрачен в области длин волн $430\div 450$ нм (синий и фиолетовый цвета). Использование фотодиодов значительно улучшает спектральные характеристики ФПЗС матриц, особенно в коротковолновой части спектра, поскольку не возникают проблемы, связанные с прозрачностью электродов.

Фотогенерацию вызывают только фотоны с энергией, превышающей ширину запрещённой зоны $\Delta W \approx 1,2$ эВ. Длина волны $\lambda = hc/\Delta W = 1,05$ мкм (это ближний ИК диапазон) является красной границей фотоэффекта в кремнии.

Светочувствительность ФПЗС определяют как минимальную освещённость объекта, при которой можно различить в видеосигнале переход от чёрного к белому. У цветных матриц чувствительность ниже из-за ослабления света в светофильтрах и уменьшения светочувствительной поверхности пикселей.

Неоднородность чувствительности. Пиксели ПЗС матрицы технологически имеют неодинаковую чувствительность к свету. Коррекция выполняется системой Auto Gain Control (AGC). В сканерах система AGC инициализируется при каждом включении прибора. Выполняются следующие действия:

- для каждого пикселя линейки ПЗС считывается значение при освещении его эталонным белым светом;
- значение сравнивается с опорным уровнем (например, средним);
- разница между выходным значением и опорным уровнем запоминается для каждого пикселя;
- при сканировании изображений эта разница компенсируется для каждого пикселя.

Для видеокамер и фотоаппаратов расчёт коррекции обычно даётся в заводских настройках.

Шумы ФПЗС

Фотонный шум обусловлен флуктуацией светового потока и случайным характером фотогенерации электронно-дырочных пар. Дискретный случайный процесс подчиняется статистике Пуассона: число фотоэлектронов, накопленных в ячейке, может быть определено с точностью до квадратного корня из их числа. Например, зарядовый пакет в 10^4 электронов будет флуктуировать от кадра к кадру со среднеквадратическим отклонением в 100 электронов. Отношение сигнал/шум для входного сигнала не более 100 и тем ниже, чем меньше зарядовый пакет (меньше площадь ячейки). Иначе говоря, чем больше у матрицы “мегапикселей” и чем меньше размеры кристалла, тем ниже соотношение сигнал/шум.

Шум темнового сигнала – один из основных источников шума, возникает вследствие случайности процесса тепловой генерации электронно-дырочных пар.

Шумы сканирования – возникают вследствие потери части зарядового пакета во время переноса по элементам ПЗС.

Шум выходного усилителя на МОП-транзисторах.

При соотношении сигнал/шум 50 дБ и более на мониторе видна чистая картинка, при 40 дБ – иногда заметны мелькающие точки, а при 30 дБ – “снег” по всему экрану.

Динамический диапазон ФПЗС определяется как отношение наибольшего зарядового пакета, который может быть накоплен в пикселе к шуму сканирования и усиления. Чем больше размер пикселя ПЗС, тем больше динамический диапазон. Для разных типов ПЗС эта величина составляет от 10^3 до 10^5 .

ФПЗС обладают высокой степенью линейности практически во всём динамическом диапазоне. Другими словами, число электронов, собираемых в пикселе, строго пропорционально числу попавших фотонов.

Оптическая пересветка (blooming) – расплывание ярких деталей изображения. С ростом накопленного заряда в потенциальной яме её потенциал уменьшается, и если он достигнет значения потенциала под соседним затвором, заряд начинает перетекать в соседний элемент.

Конструктивно-технологические достоинства ФПЗС:

- простота конфигурации и регулярность системы элементов – обязательное условие при создании многоэлементных (10^4 – 10^6 элементов) устройств;

- высокая степень интеграции (превышающая 10^5 элементов на одном кристалле) и высокая плотность упаковки (более 10^5 бит/см²);

- малое количество внешних выводов, что является определяющим при построении высоконадёжных систем.

Проблемы ФПЗС

Самые серьёзные из них – чрезвычайно жёсткие требования к однородности исходного кремния и степени совершенства технологического процесса изготовления ФПЗС. На кристалле формируется до 10^7 элементов. Создание таких сверхбольших

интегральных систем требует абсолютно бездефектной фотолитографии.

Если в матрице неисправен только один ПЗС-элемент, то будет потеряна информация от всего столбца, так как при сканировании через неисправный элемент не смогут быть переданы зарядовые пакеты других элементов столбца и на воспроизводимом изображении появится вертикальная полоса. Наличие одного дефектного элемента приводит к негодности всего кристалла.

1.3. Приёмники теплового излучения

Приёмниками излучений называют приборы, преобразующие энергию излучения в электрический или оптический сигналы. Приёмники теплового излучения можно разделить на датчики излучения и тепловизионные приборы.

Первые находят применение в системах автоматики и контроля технологических процессов, в охранной сигнализации, управлении освещением и т.д. Тепловизоры преобразуют тепловое инфракрасное (ИК) излучение от объектов и местности в видимое изображение. Основное отличие приёмников теплового излучения от приборов видимого диапазона состоит в том, что они регистрируют собственное тепловое излучение объекта, а не отражённый свет и не различия в отражательной способности элементов объекта.

В дальнем ИК-диапазоне длин волн $8\div 14$ мкм располагаются максимумы мощности излучения слабо нагретых тел. Согласно закону смещения Вина длина волны, соответствующая максимальному значению лучеиспускательной способности абсолютно чёрного тела, обратно пропорциональна его абсолютной температуре

$$\lambda_{\text{макс}} = \frac{b}{T},$$

где $b = 2898$ мкм·град – постоянная Вина.

Для $\lambda_{\text{макс}} = 8$ мкм – это тела с температурой плюс 90°C (363K), а для $\lambda_{\text{макс}} = 14$ мкм – соответственно минус 60°C (213K).

По принципу действия приёмники ИК-излучения подразделяются на термоэлектрические, фотоэлектронные, болометрические и пироэлектрические.

1.3.1. Термоэлектрические приёмники излучения

Термоэлектрические (термопарные) приёмники основаны на термоэлектрическом эффекте (эффекте Зеебека). Основным элементом данного устройства является контактная пара двух тонкоплёночных материалов А и В (рис.1.34) с различной рабо-

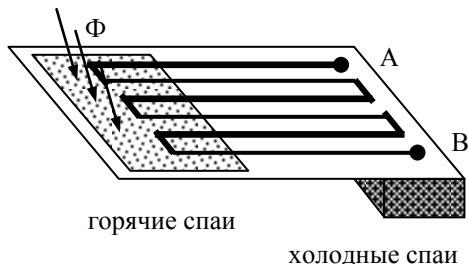


Рис.1.34. Структура плёночного термоэлектрического элемента

той выхода, в результате чего на их границе возникает разность потенциалов. Это напряжение пропорционально разнице температур между холодным и горячим контактами и называется термо-ЭДС:

$$U_{AB} = \alpha_{AB} \Delta T = (\alpha_A - \alpha_B)(T_{\text{гор}} - T_{\text{хол}}),$$

где α_A и α_B – коэффициенты термо-ЭДС материалов термопары, $T_{\text{гор}}$ и $T_{\text{хол}}$ – температуры горячего и холодного контактов.

Структура термопарного приёмника показана на рис.1.34. Термопары (поликристаллический кремний-алюминий) находятся на тонкой мембране, подвешенной на теплопроводящих стойках над подложкой. Чувствительным элементом является зачернённая область мембраны. Нагрев горячего контакта термопары происходит за счёт поглощения ИК излучения, а холодный контакт связан с подложкой, которая имеет более низкую температуру. Для повышения термо-ЭДС используют серию последовательно соединённых термопар.

Достоинства данных приёмников – широкий спектральный диапазон, совместимость с технологией кремниевой микроэлектроники и низкая стоимость.

Недостатки – чувствительность только к мощным потокам излучения и малое количество элементов в матрице (до 64×64).

1.3.2. Фотозлектрические приёмники теплового излучения

Приёмники на основе фотоэффекта, работающие в диапазоне $8 \div 14$ мкм, выпускаются в основном на базе кристалла твёрдого раствора CdHgTe . Это узкозонный полупроводник, ширина запрещённой зоны которого меньше энергии фотона с $\lambda = 14$ мкм:

$$\Delta W < h\nu = \frac{hc}{\lambda} = 0,09 \text{ эВ}.$$

Введением донорных или акцепторных примесей в кристалл CdHgTe получают узкозонные n - или p -полупроводники. На границе зон в кристалле образуется p - n -переход аналогично образованию перехода в кристалле кремния. Поглощение квантов ИК излучения вблизи перехода сопровождается генерацией пар электрон-дырка, и обратный ток диода возрастает на величину I_ϕ (1.30):

$$I_{\text{обр}} = I_0 + I_\phi.$$

Протекание через переход фототока I_ϕ приводит к увеличению числа основных носителей заряда в n - и p -областях полупроводника, что уменьшает общую разность потенциалов p - n -перехода $\varphi = \varphi_0 - \varphi_\phi$. Измеряя $I_{\text{обр}}$ или φ , можно определять величину потока ИК излучения Φ_λ , падающего на поверхность кристалла полупроводника, или её изменение $\Delta\Phi_\lambda$, см. формулы (1.27) и (1.30).

Концентрация собственных носителей заряда в узкозонном полупроводнике (1.1), (1.2):

$$n_i = p_i = N \exp\left(-\frac{\Delta W}{2kT}\right)$$

при комнатной температуре (300K) на несколько порядков больше, чем в кремнии. Во столько же раз будет больше и обратный (темновой) ток фотодиода I_0 (1.23). На его фоне слабый фототок будет неразличим. Чтобы уменьшить темновой ток I_0 и повысить чувствительность к ИК излучению, фотодиоды охлаждают до температуры жидкого азота (77 K) или гелия (4,2 K).

ИК-фотодиоды используют для регистрации излучений отдельных источников. Их конструкция подобна кремниевому фотодиоду (рис.1.18). Принципиальное отличие состоит в том, что окно закрыто не стеклом, а пластинкой германия, прозрачной для излучения в диапазоне $8 \div 12$ мкм.

Фотодиодные матрицы. Матичные приёмники теплового излучения

В настоящее время твёрдые растворы $\text{Cd}_x\text{Hg}_{1-x}\text{Te}$ являются основным материалом для изготовления фотоприёмников инфракрасного диапазона. Выбором x можно менять ширину запрещённой зоны от десятков до сотен мэВ. Кроме того, малое рассогласование постоянных решётки CdTe и HgTe позволяет выращивать эпитаксиальные структуры $\text{Cd}_x\text{Hg}_{1-x}\text{Te}$ на подложках CdTe большого размера. CdTe – широкозонный полупроводник, ширина запрещённой зоны которого составляет 1,6 эВ.

Используя достоинства этого материала, созданы фотодиодные матрицы, работающие в диапазоне теплового излучения. Объектив с германиевыми линзами фокусирует излучение наблюдаемого объекта на поверхности матрицы. Отклик каждого фотодиода в виде изменения разности потенциалов на p – n –переходе $\varphi_{\text{ф}}$ пропорционален его освещённости. Таким образом тепловая картина преобразуется в электрический рельеф. Для накопления и считывания сигналов фотодиодов на этой же подложке изготавливают схему мультиплексора по технологии приборов с зарядовой связью (ПЗС), см. рис.1.33.

В настоящее время выпускаются охлаждаемые матрицы с форматом 320×240 и 640×480 элементов (пикселей). Их спектральная чувствительность на 1–2 порядка выше болометрических. Постоянная времени порядка 10^{-5} с, т.е. их возможно применять для наблюдения за подвижными объектами.

Главным недостатком приёмников на основе фотоэффекта является необходимость охлаждения до низких температур (77 К или -196°C). Это значительно повышает стоимость самих приборов, их эксплуатации и создаёт ограничения в их использовании.

1.3.3. Болометрические приёмники теплового излучения

Термисторы

Термисторы – это полупроводниковые терморезисторы с отрицательным температурным коэффициентом сопротивления (ТКС). Их ещё называют NTC-термисторы (Negative temperature coefficient).

Температурная зависимость электропроводимости ковалентных полупроводников

Отрицательный ТКС термисторов, изготовленных из кристаллов ковалентных полупроводников (кремний, германий, соединения типа $A^{III}B^V$ и др.), вызван ростом концентрации собственных носителей заряда (электронов n_i и дырок p_i) с увеличением температуры (1.1), (1.2):

$$n_i = p_i = N \exp\left(-\frac{\Delta W}{2kT}\right)$$

где ΔW – ширина запрещённой зоны, N – плотность квантовых состояний у дна зоны проводимости и верха валентной зоны.

Проводимость собственного полупроводника определяется соотношением

$$\sigma = q_e \mu_n n_i + q_e \mu_p p_i = \sigma_0 \exp\left(-\frac{\Delta W}{2kT}\right), \quad (1.35)$$

где q_e – элементарный заряд, μ_n и μ_p – подвижности электронов и дырок, $\sigma_0 = q_e(\mu_n N_c + \mu_p N_v)$ – коэффициент, слабо зависящий от температуры по сравнению с экспоненциальным множителем.

Основной характеристикой термистора является температурная зависимость его сопротивления R_T (рис.1.35):

$$R_T \sim \rho = \frac{1}{\sigma}.$$

$$R_T = A \cdot \exp\left(\frac{B}{T}\right), \quad (1.36)$$

где R_T – сопротивление термистора при температуре T , A – величина, зависящая от материала и размеров термистора, $B = \Delta W/2k$ – коэффициент температурной чувствительности.

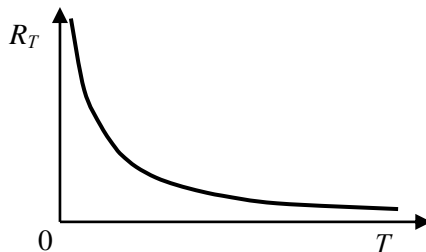


Рис.1.35. Зависимость сопротивления термистора от температуры

Температурная зависимость электропроводимости оксидных полупроводников

Основная часть термисторов, выпускаемых промышленностью, изготовлена из оксидных полупроводников – оксидов металлов переходной группы таблицы Д. И. Менделеева (от титана до цинка).

Электропроводность оксидных полупроводников с ионной связью отличается от электропроводности ковалентных полупроводников. В оксидных полупроводниках реализуется *перескоковый механизм проводимости*. Для металлов переходной группы характерны наличие незаполненных электронных оболочек и переменная валентность. При образовании оксида в одинаковых кристаллографических положениях оказываются ионы с разными зарядами. Электропроводность таких материалов связана с туннельными переходами электронов через потенциальные барьеры между соседними ионами. Например, из электронной оболочки иона Cu^{1+} валентный электрон может перейти в оболочку иона Cu^{2+} . При этом валентность первого иона увеличивается ($\text{Cu}^{1+} \rightarrow \text{Cu}^{2+}$), а второго – уменьшается ($\text{Cu}^{2+} \rightarrow \text{Cu}^{1+}$). Частота переходов возрастает с увеличением амплитуды тепловых колебаний решётки. Во внешнем электрическом поле наиболее вероятными будут переходы электронов в направлении против поля.

Подвижность носителей заряда в оксидном полупроводнике оказывается довольно низкой и экспоненциально возрастает с ростом температуры. В результате температурная зависимость сопротивления термистора из оксидного полупроводника имеет такой же характер, как и у термисторов, изготовленных из полупроводников с ковалентной связью, но она обусловлена изменением не концентрации свободных носителей заряда, а их подвижности.

Конструкция термисторов

Термисторы изготавливают из полупроводниковой керамики – смесей поликристаллических оксидов переходных металлов (CoO , MnO , CuO).

Термисторы изготавливают в виде стержней, дисков, шайб, бусинок и тонких пластинок методом керамической технологии: прессованием из смеси порошков и спеканием при высокой температуре. Их размеры могут варьироваться от 10 мкм до 1÷2 см.

Основные характеристики термисторов

Чувствительность термистора в качестве датчика температуры:

$$S_T = \frac{1}{R} \cdot \frac{dR}{dT} = \alpha_T = -\frac{B}{T^2}. \quad (1.37)$$

Интервал рабочих температур у металлооксидных термисторов составляет от -60°C до $+180^\circ\text{C}$.

Коэффициент теплового рассеяния H численно равен мощности, отводимой от термистора в окружающую среду при разности температур поверхности и окружающей среды в один градус.

Постоянная времени τ термистора – это время, в течение которого разность температур термистора и окружающей среды $(T - T_0)$ уменьшится в 2,71 раз (на 63 %) по отношению к начальной разности температур $(T_1 - T_0)$. Типичные значения постоянных времени находятся в диапазоне от десятых долей секунды до нескольких минут.

Применение термисторов

Термисторы применяются в приборах измерения, регулирования и стабилизации температуры, в датчиках пожарной сигнализации, в цепях температурной стабилизации режимов работы электронных устройств и др.

Цепь для измерения температуры (рис.1.36) состоит из источника тока, вольтметра и термистора. Измеряемой величиной является падение напряжения на термисторе $U_T = IR_T$. Величину тока выбирают таким образом, чтобы погрешность, вносимая саморазогревом термистора, на-

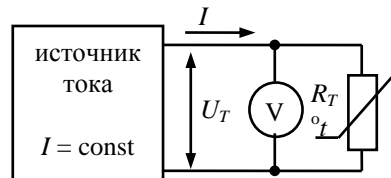


Рис.1.36. Схема включения термистора для измерения температуры

ходила в допустимых пределах. Перегрев термистора на 1 градус вызывается мощностью $I^2 R_T$, равной его коэффициенту теплового рассеяния H .

Термисторы имеют очень высокую температурную чувствительность (1.35) по сравнению с другими типами датчиков температуры. ТКС α_T термисторов составляет -5% на градус, в то время как для платинового терморезистора $+0,4\%$ на градус.

Болометр

Болометр (от др.-греч. βολή – луч и μέτρον – измерение) – прибор, предназначенный для регистрации и измерения мощности теплового излучения.

Полупроводниковый болометр состоит из двух одинаковых плёночных термисторов (рис.1.37,а). Температура первого термистора зависит от мощности падающего на него теплового потока. Его сопротивление R_{T1} изменяется по отношению ко второму термистору.

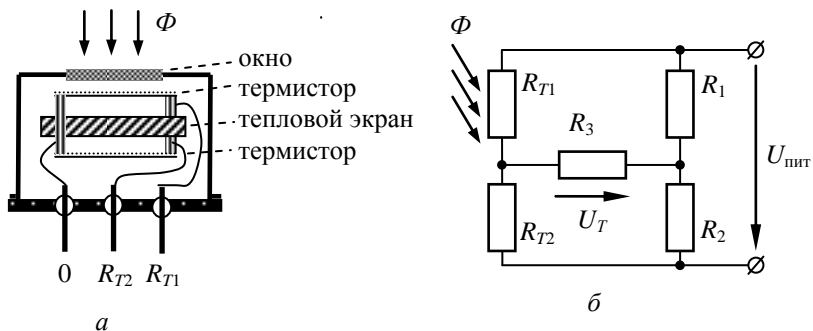


Рис.1.37. Болометр: а – конструкция; б – мостовая схема включения

Второй термистор R_{T2} экранирован от измеряемого излучения и служит для компенсации изменений температуры окружающей среды. Оба термистора помещаются в общий герметичный корпус. Плёнку первого термистора чернят для лучшего поглощения излучения.

Болометры обычно имеют три внешних вывода: от активного и компенсационного термисторов и от средней точки. Для измерения малых отклонений сопротивления термисторов их

включают в мостовую схему (рис.1.37,б), которую балансируют при отсутствии засветки

$$\frac{R_{T1}}{R_{T2}} = \frac{R_1}{R_2}.$$

Облучение вызывает разбалансировку моста ($R_{T1} < R_{T2}$) и появление сигнала на сопротивлении R_n . Одновременное изменение R_{T1} и R_{T2} из-за колебаний внешней температуры не нарушает балансировку моста. Мост питается переменным или постоянным напряжением.

Для изготовления термисторов используют плёнки окислов никеля, кобальта, марганца, меди, а также германия и кремния толщиной до 10 мкм. Поверхность чувствительного термистора чернят для увеличения коэффициента поглощения.

Основные параметры полупроводниковых болометров:

- сопротивление термисторов – от 0,1 до 10 МОм;
- интегральная чувствительность (1.37)

$$S = \frac{1}{R_T} \frac{dR_T}{dT} ;$$

– постоянная времени τ . Из-за малой теплоёмкости плёнок термисторов инерционность болометров лежит в пределах от 250 мкс до 40 мс;

– пороговая чувствительность (пороговый тепловой поток) болометров обычно лежит в интервале $10^{-9} \div 10^{-11}$ Вт.

Полупроводниковые болометры применяются в системах охранной сигнализации, автоматического управления освещением, для дистанционного измерения температуры объектов и в других областях.

Болометр чувствителен ко всему спектру излучения. Для ИК диапазона болометр – самый чувствительный датчик. Его применяют в астрономии. Источником излучения может быть свет звёзд или Солнца, прошедший через спектрометр и разложенный на множество спектральных линий, энергия в каждой из которых очень мала.

Матричные микроболометрические приёмники

Каждый элемент (пиксель) матрицы (рис.1.38,а) представляет собой микроболометр – полупроводниковый термистор. Микроболометр выполнен в виде тонкой прямоугольной пластины – мембраны (рис.1.38,б). Мембрана подвешена на двух консолях над кремниевой подложкой, на которой создано считывающее устройство (мультиплексор). Получается конструкция в виде “моста”, поднятого над подложкой на высоту ~2,5 мкм.

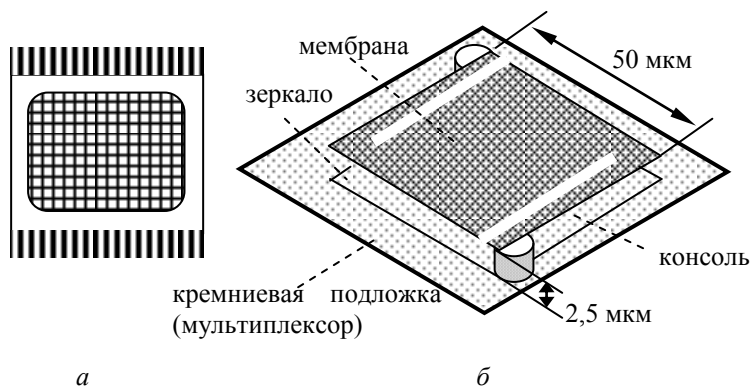


Рис.1.38. Микроболометр: а – матрица; б – конструкция элемента

Мембрана образована терморезистивной плёнкой окислов ванадия (VO_x) или аморфного кремния (Si) толщиной ~50 нм, находящейся между двух тонких слоёв нитрида кремния (Si_3N_4).

Консоли также изготовлены из нитрида кремния, материала с высокой прочностью и низкой теплопроводностью. На консоли напылён слой нихрома, который обеспечивает электрический контакт плёнки термистора со схемой мультиплексора.

При нагревании мембраны ИК-излучением сопротивление плёнки термистора уменьшается. Эти изменения регистрируются чувствительной схемой мультиплексора. Сопротивление плёнки термистора более 100 кОм, поэтому в схемах считывания используют планарные полевые транзисторы.

На подложку под мембраной наносится металлическое зеркало, зазор между ним и болометром составляет 2,5 мкм. Зеркало с мембраной болометра образуют четвертьволновый оптиче-

ский резонатор, увеличивающий поглощение излучения с длиной волны ~ 10 мкм.

Форматы микроболометрических матриц, выпускаемых промышленностью: 160×120 , 320×240 , 640×480 пикселей. Размер пикселя ~ 30 мкм.

Достоинствами микроболометров является их совместимость с технологией кремниевых интегральных схем. Микроболометрическая матрица создаётся на кремниевой пластине, на поверхности которой предварительно было изготовлено считывающее устройство (мультиплексор).

Последовательность основных операций при изготовлении матрицы показана на рис.1.39:

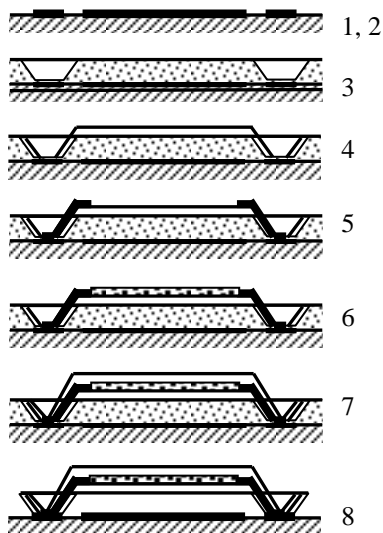


Рис.1.39. Последовательность операций при изготовлении микроболометрической матрицы

1) Формируются алюминиевые контактные площадки на кремниевой схеме считывания для подсоединения микроболометров;

2) Формируются алюминиевые зеркала оптических резонаторов, расположенные под микроболометрами;

3) Производится нанесение слоя полиимида (“жертвенного слоя”) толщиной 2,5 мкм. Формируются воронкообразные окна до алюминиевых контактных площадок. Полиимид – полимер-

ный материал, отличающийся прочностью, устойчивостью к воздействию химических веществ и высокой температуры;

4) Наносится нижний слой нитрида кремния. Используется метод плазмохимического нанесения в реакторе с индукционно-связанной плазмой. Толщина плёнки составляет 150 нм. Проводится плазмохимическое травление окон в слое нитрида кремния до алюминиевых контактных площадок;

5) Магнетронным распылением формируются нихромовые шины, идущие по нитридкремниевой консоли микроболометра, и контакты к оксиду ванадия. Толщина плёнки 50 нм;

6) Наносится термочувствительная плёнка оксида ванадия;

7) Наносится верхний слой нитрида кремния. Путём плазмохимического травления формируются нитридкремниевые консоли. Ширина консоли $1,5 \div 2,0$ мкм;

8) Удаление “жертвенного слоя”. Полиимид удаляется в кислородной плазме.

Для повышения чувствительности микроболометрических матриц нужно снижать теплопроводность между микроболометрами и их окружением, а также уменьшать их теплоёмкость. Это достигается уменьшением размера элементарного болометра. Однако имеется предел такого уменьшения. Он связан с длиной волны используемого излучения $8 \div 12$ мкм.

Уменьшение толщины мембраны снижает теплоёмкость микроболометра. Однако даже при использовании оптического резонатора толщина мембраны должна быть достаточной ($\sim 0,5$ мкм), чтобы обеспечить высокое поглощение падающего на неё излучения.

Микроболометрическая матрица помещается в вакуумированный корпус. Входное окно образует германиевая пластинка, прозрачная для излучения в рабочем диапазоне $8 \div 12$ мкм.

1.3.4. Пироэлектрические приёмники теплового излучения

Пироэлектрики (от греч. *pir* – огонь) – кристаллические диэлектрики, обладающие самопроизвольной (спонтанной) поляризацией.

Спонтанная поляризация возможна в кристаллах, имеющих одну ось симметрии (полярную ось). “Центры тяжести” положительных и отрицательных зарядов в элементарной ячейке могут

не совпадать. Такая ячейка имеет электрический дипольный момент p_i , который направлен вдоль оси симметрии. Сумма элементарных моментов p_i в единице объёма образует макроскопический дипольный момент (вектор поляризации) P' кристалла, также направленный вдоль его оси симметрии.

На поверхности кристалла образуется связанный заряд, поверхностная плотность которого равна нормальной компоненте вектора поляризации $\sigma' = P'$. Внутри кристалла возникает электрическое поле E'

$$E' = -\frac{\sigma'}{\varepsilon_0} = -\frac{P'}{\varepsilon_0}.$$

Пироэлектрический эффект состоит в изменении спонтанной поляризации диэлектриков при изменении температуры.

Для наблюдения пироэлектрического эффекта используют пластинку, вырезанную перпендикулярно полярной оси кристалла. На пластинку наносят металлические электроды. Образуется конденсатор с поляризованным диэлектриком. На обкладках конденсатора накапливаются заряды (за счёт адсорбции ионов воздуха, электропроводности самого кристалла, поверхностных токов утечки и др.), компенсирующие электрическое поле поляризованного диэлектрика. Разность потенциалов между обкладками устанавливается равной нулю, и кристалл внешне не обнаруживает электрической поляризации.

Пироэлектрические свойства можно наблюдать только при достаточно быстром изменении температуры кристалла. При этом изменяются амплитуды тепловых колебаний, размеры ячеек и их дипольные моменты.

Изменение температуры кристалла на величину ΔT вызывает изменения:

величины вектора поляризации

$$\Delta P' = \gamma \Delta T, \quad (1.38)$$

поверхностной плотности связанных зарядов

$$\Delta \sigma' = \Delta P',$$

величины заряда на электродах

$$\Delta Q = S \Delta \sigma' = \gamma S \Delta T, \quad (1.39)$$

и разности потенциалов между ними

$$\Delta U = \Delta E' d = \frac{\Delta P'}{\varepsilon_0} d = \left(\frac{\gamma d}{\varepsilon_0} \right) \Delta T, \quad (1.40)$$

где γ – пирозлектрический коэффициент, S – площадь, d – толщина пластинки пирозлектрика.

Таким образом, измеряя напряжение между электродами, можно следить за изменением температуры пластинки пирозлектрика.

В другом методе наблюдения измеряют ток заряда-разряда обкладок конденсатора, который пропорционален скорости изменения температуры пластинки

$$I = \frac{\Delta Q}{\Delta t} = \gamma S \frac{\Delta T}{\Delta t}. \quad (1.41)$$

Пирозлектрический эффект – нелинейное явление, но в диапазоне небольших изменений температуры ΔT используют показанное выше линейное приближение.

Одноэлементные пирозлектрические приёмники излучений

Вариант конструкции приёмника показан на рис.1.40,*а*. Пластина пирозлектрика с напылёнными металлическими электродами (пирозлектрический конденсатор) закреплена на тонкой термоизолирующей подложке. Под ними находится электронный усилитель, выполненный на полевом транзисторе

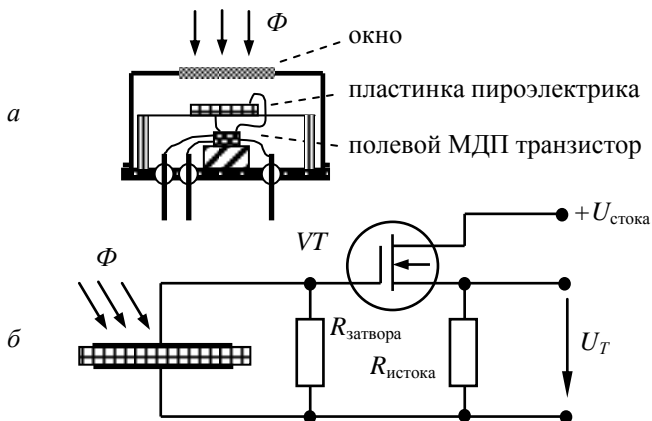


Рис.1.40. Пирозлектрический детектор ИК излучения:
а – конструкция; *б* – электрическая схема включения

по схеме истокового повторителя (рис.1.40,б). Транзистор имеет большое входное сопротивление, малую входную ёмкость и низкое выходное сопротивление. Сигнал с его выхода подается на электронные схемы для дальнейшего усиления и обработки. Устройство заключено в герметичный металлический корпус, из которого может быть откачан газ. В корпусе сделано окно, закрытое германиевой пластинкой, прозрачной для длинноволнового ИК излучения. Изменение температуры пластинки пироэлектрика падающим излучением Φ вызывает изменение его поляризации и возникновение напряжения U_T между электродами на пластинке.

Пусть на пластинку площадью A падает поток излучения $A\Phi$. Пренебрежём теплопроводностью элементов крепления и электрических выводов. Происходит только радиационный обмен энергией между пластинкой и окружением – корпусом прибора (рис.1.41.).

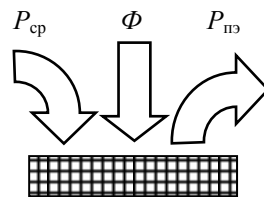


Рис.1.41. Энергетический баланс пластинки пироэлектрика

Уравнение энергетического баланса

$$C_p \left(\frac{dT_{пз}}{dt} \right) = \mu_1 A \Phi + P_{cp} - P_{пз}. \quad (1.42)$$

Здесь $C_p(dT_{пз}/dt)$ – изменение внутренней энергии пластинки; $P_{cp} = \mu_2 A \sigma T_{cp}^4$ – мощность излучения окружающей среды (корпуса детектора), поглощённая пироэлектрическим элементом; $P_{пз} = \mu_1 A \sigma T_{пз}^4$ – мощность излучения, испускаемого пироэлектрическим элементом и поглощённая корпусом; μ_1 и μ_2 – коэффициенты поглощения (зачернённой) пластинки и (белого) корпуса; $T_{пз}$ и T_{cp} – температуры пластинки и корпуса.

Уравнение энергетического баланса принимает вид

$$C_p \left(\frac{dT_{пз}}{dt} \right) = \mu_1 A \Phi + \mu_2 A \sigma T_{cp}^4 - \mu_1 A \sigma T_{пз}^4. \quad (1.43)$$

Полагая, что температура пластинки $T_{пз}$ немного отличается от температуры окружающей среды $(T_{пз} - T_{cp}) = \Delta T_{пз} \ll T_{cp}$, разложим $T_{пз}^4$ в ряд Тейлора, ограничиваясь первым членом:

$$T_{пз}^4 = (T_{cp} + \Delta T_{пз})^4 = T_{cp}^4 + 4T_{cp}^3 \Delta T_{пз}. \quad (1.44)$$

Подставив (1.44) в уравнение энергетического баланса (1.43), получим

$$C_p \left(\frac{dT_{пэ}}{dt} \right) = \mu_1 A \Phi - 4\mu_2 A \sigma T_{ср}^3 \Delta T_{пэ}. \quad (1.45)$$

При термодинамическом равновесии

$$C_p \left(\frac{dT_{пэ}}{dt} \right) = 0,$$

и превышение температуры пирозлектрической пластинки над температурой корпуса пропорционально потоку падающего на пластинку излучения

$$\Delta T_{пэ} = \frac{\Phi}{4\sigma(\mu_2/\mu_1)T_{ср}^3} = \frac{\Phi}{G_{пэ}}. \quad (1.46)$$

Для повышения чувствительности датчика нужно обеспечить теплоизоляцию пирозлектрического элемента от окружения ($\mu_2 \rightarrow 0$). Для сокращения времени установления температуры τ (рис.1.42) (повышения быстродействия) снижают тепло-

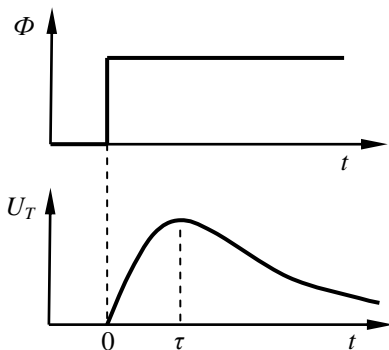


Рис. 1.42. Процесс установления разности потенциалов на обкладках пирозлектрической пластинки при воздействии потока ИК излучения

ёмкость C_p , уменьшая толщину пластинки d . Из формулы видно, что понижение температуры окружающей среды (корпуса прибора) $T_{ср}$ повышает чувствительность датчика. Поэтому в ряде случаев применяют термоэлектрические холодильники (-30°C – -100°C).

В качестве пластинок датчика используют тонкие срезы кристаллов танталата и ниобата лития LiTaO_3 и LiNbO_3 и более сложных соединений (цирконат титанат лантана-свинца и др.).

Во многих случаях вместо монокристаллов используют пластинки пирозлектрической керамики на основе титаната ба-

рия BaTiO_3 , титаната-цирконата свинца и др. Керамика представляет собой спрессованный и спечённый порошок из кристалликов сегнетоэлектрика. Керамику нагревают выше точки Кюри и охлаждают в сильном электрическом поле. Ориентации векторов поляризации в кристалликах приобретают преимущественное направление, совпадающее с направлением поля. Керамические пьезоэлектрики приближаются по своим свойствам к кристаллическим.

Наибольший объём применений пироприёмников связан с регистрацией ИК излучения. Пьезоэлектрические приёмники используются также в лазерной технике как измерители энергии излучения мощных лазеров. Пьезоэлектрические приёмники имеют ряд преимуществ по сравнению с полупроводниковыми фотодетекторами. Они могут работать в очень широком диапазоне длин волн и не нуждаются в охлаждении криогенными жидкостями. Работая при комнатной температуре, они способны регистрировать тепловое излучение в дальнем ИК диапазоне.

Пьезоэлектрический (пассивный) инфракрасный датчик движения (Pyroelectric (Passive) InfraRed sensor, PIR-sensor) применяется в охранной сигнализации, устройствах автоматического включения освещения, автоматического открывания раздвижных дверей при приближении человека. ПИР-датчик реагирует на слабое ИК излучение тела человека, температура которого в общем случае отличается от температуры окружающих предметов.

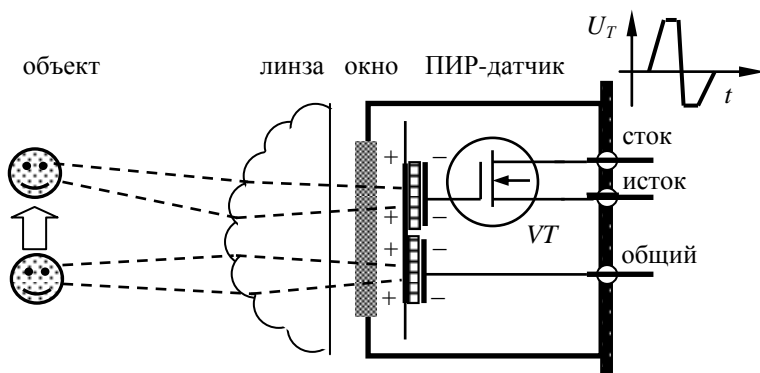


Рис. 1.43. Пьезоэлектрический датчик движения

Датчик содержит две пирозлектрические пластинки или плёнки (1×2 мм), расположенные рядом друг с другом. (рис.1.43). Пластмассовая линза волнистой формы фокусирует излучение теплого объекта (человека) на поверхности одной из пластинок. На её металлических обкладках возникает импульс напряжения.

При перемещении объекта излучение сместится на вторую пластинку, и она тоже сформирует импульс, но другой полярности. Импульсы усиливаются полевым транзистором. Таким образом, движущийся объект вызовет импульсы тока в электрической схеме, в которую включен датчик. На излучение неподвижных объектов датчик не реагирует.

Для расширения угла зрения линзу делают в виде полусферы, разбитой на несколько сегментов, подобно линзе Френеля. В этом случае излучение движущегося объекта будет много раз поочерёдно попадать на пирозлектрические пластинки. На выходе датчика сформируется последовательность импульсов.

Матричные пирозлектрические приёмники

Матричные пирозлектрические приёмники представляют собой прямоугольную матрицу (рис.1.44) точечных приёмников. Серийно выпускаются матрицы 160×120 , 320×240 элементов. На поверхность матрицы фокусируются потоки излучения от ис-

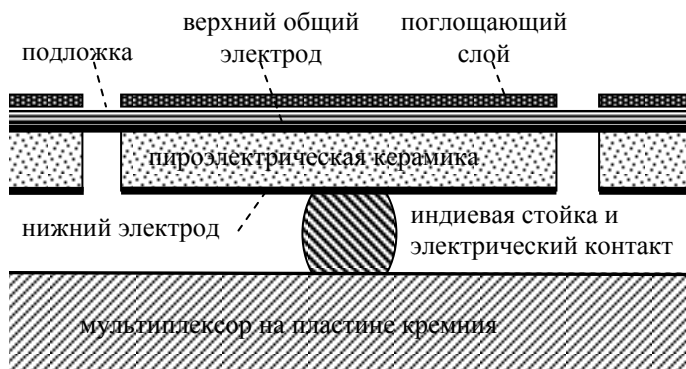


Рис.1.44. Структура пикселя пирозлектрической матрицы гибридной сборки на основе сегнетоэлектрической керамики титанат-бария стронция

следуемых источников. Матричные приёмники, помещённые в фокусной плоскости объектива, называют приёмниками излучения в фокальной плоскости (focal plane arrays – FPA или uncooled focal plane arrays – UFPA). Они дают возможность получать информацию о пространственном распределении и визуализировать различные потоки излучения.

Отдельные элементы (пикселы) матрицы имеют площадь от 30×30 до 50×50 мкм (минимальный размер 15 мкм для области спектра $\lambda = 8 \div 14$ мкм). Пирозлектрические элементы изготавливают методом напыления или эпитаксиального наращивания на подложке – диэлектрической плёнке, обеспечивающей электрическую и тепловую изоляцию элементов. Плёнка закреплена над поверхностью кремниевой пластины на металлических стойках, обеспечивающих механическую прочность и электрические контакты между электродами элементов и электронной схемой считывания и мультиплексирования сигналов (рис.1.44). Мультиплексоры (readout integrated circuit – ROIC) изготавливают на поверхности кремниевой пластины, используя технологии приборов с зарядовой связью (ПЗС) или комплиментарные металл-оксид-полупроводник (КМОП) технологии.

Основной проблемой при создании матричных пирозлектрических приёмников является несовместимость технологических процессов формирования пирозлектрических плёнок с технологическими процессами кремниевой микроэлектроники. В первую очередь – это высокая температура выращивания пирозлектрических плёнок ($\sim 700^\circ\text{C}$), более высокая, чем используется при изготовлении кремниевых микросхем. Поэтому производят, в основном, матрицы гибридной сборки (рис.1.44), в которых пирозлектрическая структура изготавливается отдельно и затем совмещается с кремниевым мультиплексором. Это напрямую отражается на количестве и стоимости производимых матриц. Параллельно разрабатываются технологии изготовления пирозлектрических матриц в виде микромостиковых структур, подобных микроболометрическим, в которых терморезистивная плёнка заменена на пирозлектрическую (рис.1.38).

1.3.5. Тепловизоры

Тепловидение – наблюдение объектов по их собственному излучению.

Тепловизор, работающий в диапазоне длин волн 8÷14 мкм, позволяет формировать тепловое изображение людей, животных, машин, оборудования, строительных конструкций, электросетей, трубопроводов, участков поверхности Земли и т.д. независимо от освещения, только за счёт собственного излучения объекта.

Для формирования теплового изображения необходим тепловой контраст, т. е. разница мощностей теплового излучения объектами, имеющими разные температуры или коэффициенты излучения. Современные тепловизионные приборы способны строить изображения с температурными контрастами в доли градуса.

Тепловизоры состоят из оптической системы (объектива), микроболометрической или пироэлектрической матрицы, устройств считывания и усиления сигналов с матрицы, блока обработки сигналов и дисплея.

Объективы тепловизоров изготавливают из германия, материала прозрачного для длинноволнового ИК излучения. Стеклые линзы для этой цели не пригодны. Объектив фокусирует тепловой поток от объекта на поверхность болометрической (или пироэлектрической) матрицы. Эти матрицы не требуют охлаждения. Их спектральная чувствительность равномерна в диапазоне от 8 до 12 мкм. Предел чувствительности для неохлаждаемых микроболометрических матриц около 0,08°C.

Мультиплексор считывает возникающие напряжения на микроболометрах или пироэлектрических элементах. Сигналы считываются одновременно или построчно со всей матрицы, затем усиливаются, оцифровываются и преобразуются в стандартные видеосигналы изображения.

Для представления изображения (термограммы) объекта в виде, удобном для анализа, используются специальные алгоритмы калибровки, коррекции сдвига и усиления, усреднения, накопления, замещения дефектных (“слепых”) элементов матрицы, компенсации шумов, а также регулировки яркости и кон-

траста в пределах выбранного динамического диапазона температур наблюдаемых объектов (от 2 до 60°C). Измерительные свойства тепловизора напрямую зависят от программных и технических решений конкретного производителя.

Пирозлектрические матрицы превосходят по чувствительности микроболометрические матрицы. Для регистрации температуры микроболометра необходимо пропустить через него электрический ток, вызывающий дополнительный нагрев. В то же время для регистрации сигнала от пироприёмника к нему не нужно прикладывать внешнее напряжение.

Электрические заряды на электродах пирозлектрического элемента возникают только при изменении его температуры. Поэтому приходится использовать прерыватели излучения, открывающие и закрывающие объектив с частотой 1÷50 Гц. В некоторых случаях используют качание камеры или её вращение для панорамных наблюдений.

Тепловизионные приборы применяются в промышленности, медицине, военной технике, строительстве, научных исследованиях и быту. Источниками инфракрасного излучения являются любые материальные объекты, имеющие температуру выше абсолютного нуля.

Примеры применения тепловизоров:

- обнаружение перегрева элементов электронных схем и электрических установок;
- контроль тепловой изоляции зданий и оборудования;
- обнаружение скрытых воспалительных процессов, вызывающих локальные повышения температуры;
- поиск пострадавших во время стихийных бедствий;
- в качестве приборов ночного видения.

шается, и наблюдается высокий уровень инжекции электронов в базу p -типа. Электроны в базе являются неосновными носителями и создают отрицательный заряд, сосредоточенный вблизи границы p – n -перехода, на расстоянии порядка средней длины диффузии электронов в базе L_n . Для сохранения электрической нейтральности из глубины базы подтягиваются дырки. Таким образом, вблизи границы p – n -перехода создаётся избыточная концентрация электронов и дырок, что приводит к их интенсивной рекомбинации с испусканием квантов света $h\nu$.

Переход электрона с нижних уровней зоны проводимости на верхние уровни валентной зоны называется межзонной рекомбинацией. Разность энергии $\Delta W = W_c - W_v$ выделяется в виде кванта электромагнитного излучения – фотона с длиной волны

$$\lambda = \frac{hc}{\Delta W}, \quad (2.1)$$

(излучательная рекомбинация), либо передаётся решётке в виде квантов тепловых колебаний – фононов (безызлучательная рекомбинация). Для увеличения вероятности прямой межзонной излучательной рекомбинации необходимо создать высокую неравновесную концентрацию электронов и дырок и уменьшить вероятность безызлучательной рекомбинации. В широко используемых полупроводниковых материалах на основе кремния и германия вероятность излучательной рекомбинации крайне мала. Вероятность излучательной рекомбинации выше в так называемых прямозонных полупроводниках, к которым относятся полупроводниковые соединения на основе элементов III и V групп, например GaAs, GaP, GaN и др.

Излучательные квантовые переходы в светодиодах происходят спонтанно, независимо друг от друга и в разные моменты времени, поэтому излучение является некогерентным. Ширина спектра излучения на уровне половины максимальной интенсивности составляет 20÷50 нм, что соответствует средней тепловой энергии электронов.

2.1.1. Рекомбинация неравновесных носителей заряда в полупроводниках

Прямая межзонная рекомбинация

Прямая межзонная рекомбинация происходит при непосредственном переходе электрона из зоны проводимости в валентную зону на один из свободных энергетических уровней (рис.2.2,а). Разность энергии $\Delta W = W_c - W_v$ может выделяться в

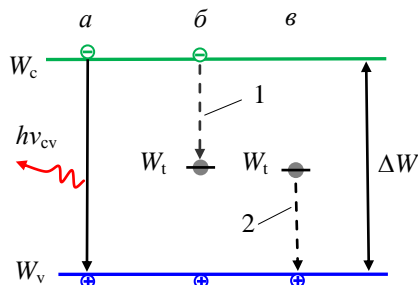


Рис.2.2. Схема рекомбинационных переходов в полупроводниках:

a – межзонный, *б* – из зоны проводимости на уровень ловушки, *в* – с уровня ловушки в валентную зону

виде кванта электромагнитного излучения – фотона $h\nu_{cv} = \Delta W$ (*излучательная рекомбинация*) либо передаваться решётке в виде квантов энергии тепловых колебаний – фононов (*безызлучательная рекомбинация*).

В большинстве полупроводников вероятность излучательной рекомбинации на несколько порядков меньше, чем безызлучательной. В то же время при безызлучательной межзонной рекомбинации должно происходить возбуждение большого количества фононов величиной порядка kT , т.е. порядка сотых долей электрон-вольта. Такие многофононные переходы также имеют малую вероятность.

Таким образом, прямая межзонная рекомбинация не является главным механизмом рекомбинации в полупроводниках.

Рекомбинация с участием ловушек

Основную роль играет рекомбинация на примесных центрах или дефектах кристаллической структуры, которые образуют *глубокие уровни* W_t , расположенные вблизи середины запрещённой зоны полупроводника. Их называют рекомбинационными ловушками (рис.2.2,в).

Рекомбинация с участием рекомбинационных ловушек протекает в два этапа. На первом этапе энергетический уровень рекомбинационной ловушки захватывает электрон из зоны проводимости (рис.2.2,б). В этом состоянии ловушка будет находиться до тех пор, пока в данном месте кристалла не окажется сво-

бодный энергетический уровень валентной зоны (дырка). При выполнении этих условий осуществляется второй этап рекомбинации – электрон переходит на свободный уровень валентной зоны (рис.2.2,в). На каждом этапе выделяется вдвое меньше энергии, чем при непосредственной рекомбинации. Это обстоятельство резко повышает вероятность передачи энергии фононам.

Для поверхностных слоёв полупроводника, где неизбежны нарушения периодичности решётки и разрывы ковалентных связей, характерна наиболее высокая скорость рекомбинации.

Переход электронов из зоны в зону может происходить и в обратном направлении (генерация пар электрон – дырка или внутренний фотоэффект) с поглощением фотонов той же или большей энергии, что и излучаемых при рекомбинации.

2.1.2. Прямозонные и непрямозонные полупроводники

В кристалле полупроводника электроны движутся в поле периодического электрического потенциала кристаллической решётки. Представим, что в некотором направлении x атомы, а следовательно, и периодический потенциал повторяются с частотой, соответствующей периоду решётки a . Электрон, движущийся в направлении x , может иметь различные значения волнового вектора k_x и импульса $p_x = \hbar k_x$ за исключением

$$\begin{aligned} k_x &= \pm(\pi/a)n, \\ p_x &= \pm(\hbar\pi/a)n, \end{aligned} \quad (2.2)$$

где $n = 1, 2, 3 \dots$, при которых длина волны де Бройля $\lambda = 2\pi/k$ оказывается кратной периоду решётки $n\lambda = 2a$. Это значит, что величины энергии, соответствующие значениям $k_x = \pm(\pi/a)n$, являются запрещёнными. Зона проводимости разделяется на ряд зон (зон Бриллюэна) разрешённых значений энергии.

В обычных условиях зона проводимости является частично заполненной зоной, в которой электроны занимают уровни, расположенные вблизи дна зоны. Эти уровни будут соответствовать самым нижним уровням первой зоны Бриллюэна ($|k_x| \ll \pi/a$). В этой области при малых значениях импульса электрона энергия электрона описывается зависимостью, аналогичной зависимости $W(p)$ свободного электрона:

$$W = \frac{p_x^2}{2m} = \frac{\hbar^2}{2m} k_x^2. \quad (2.3)$$

Кривая зависимости энергии электрона от волнового вектора $W(k)$ показана на рис.2.3. Аналогично может быть построена кривая зависимости энергии электрона от волнового вектора в валентной зоне. Как и в предыдущем случае, интерес представляют самые нижние энергетические уровни первой зоны Бриллюэна, т.е. уровни вблизи верха валентной зоны. Именно здесь находятся свободные уровни, которые интерпретируются как дырки (рис.2.3).

Из рис.2.3 можно видеть, что запрещённая зона полупроводника представляет собой зазор между экстремумами двух параболических кривых первой зоны Бриллюэна.

Мы рассмотрели одномерную задачу. В реальном трёхмерном кристалле электрон может двигаться в произвольном направлении, при этом картина зон Бриллюэна получается более

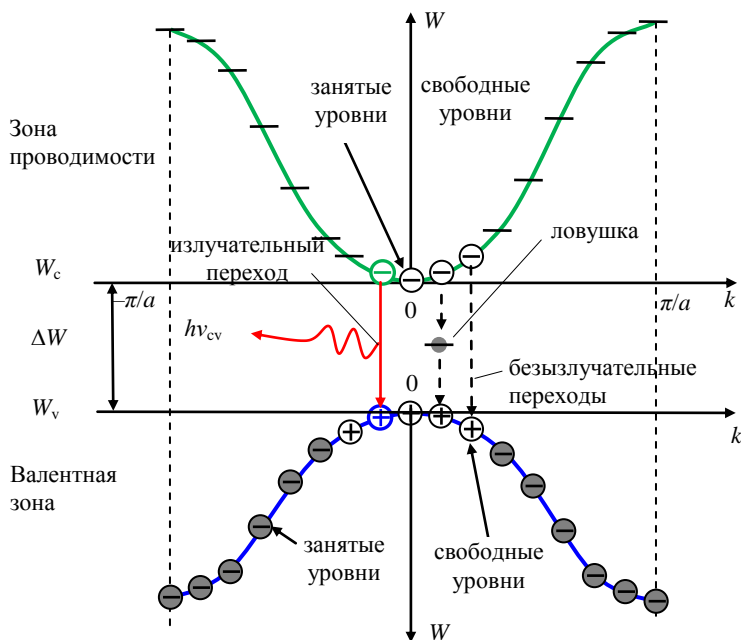


Рис.2.3. Зависимости энергии электрона от волнового вектора в прямозонном полупроводнике

сложной. Картина усложняется и за счет того, что зоны проводимости и валентная образуются расщеплением различных уровней электронных оболочек атомов. Сложная зависимость энергии от импульса приводит к тому, что положение экстремума (минимума или максимума) функции $W(k)$ может быть сдвинуто относительно нулевого значения k . Кроме того, оказывается возможным наличие нескольких экстремумов в зоне Бриллюэна, причём минимум, смещённый относительно $k = 0$, может быть более глубоким, т.е. соответствовать меньшему значению энергии по сравнению с минимумом, находящимся в точке $k = 0$ (рис.2.4).

Излучательная рекомбинация сопровождается испусканием фотона. Импульс, уносимый фотоном $p = h\nu/c$, пренебрежимо мал, и считают, что при излучательной рекомбинации импульс электрона практически не меняется. Изменяется лишь его потенциальная энергия. Поэтому излучательная рекомбинация

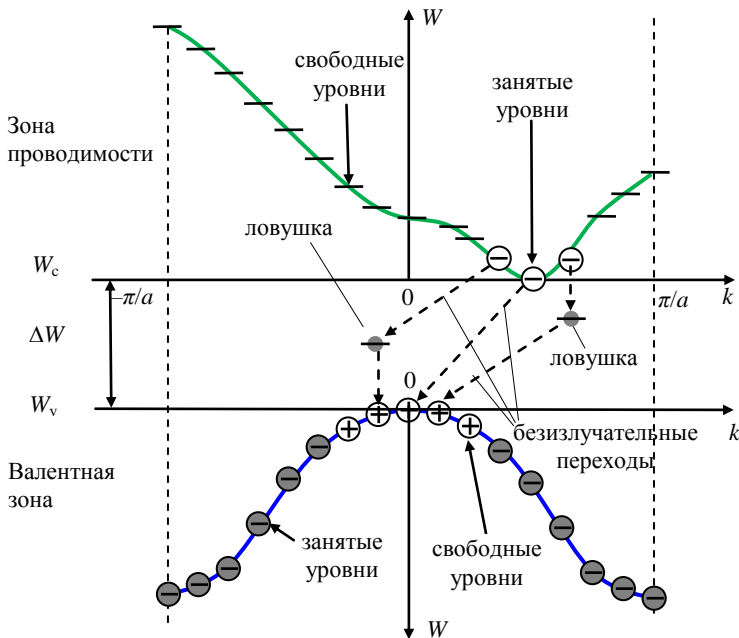


Рис.2.4. Зависимости энергии электрона от волнового вектора в непрямозонном полупроводнике

возможна только в *прямозонных полупроводниках* (рис. 2.3), у которых максимум валентной зоны и минимум зоны проводимости находятся в центре зоны Бриллюэна (в точке $k = 0$).

В *непрямозонных полупроводниках* (рис.2.4) переход электрона из зоны в зону осуществляется с изменением импульса в результате взаимодействия электрона и решётки. Непрямой переход всегда заканчивается выделением энергии в виде *фононов* (квантов тепловых колебаний кристаллической решётки). Импульс фононов определяется относительным положением минимумов первых зон Бриллюэна в зоне проводимости и валентной зоне. Рекомбинация без излучения фотона называется *безызлучательной рекомбинацией*.

В прямозонных и непрямозонных полупроводниках возможны также безызлучательные переходы с предварительным захватом электрона рекомбинационной ловушкой. Рекомбинационные ловушки являются своеобразным посредником, получающим от электрона и передающим решётке импульс фонона (рис.2.3 и 2.4).

2.1.3. Полупроводниковые материалы светоизлучающих диодов

Длина волны излучаемого света λ определяется разностью ΔE двух энергетических уровней, между которыми происходит переход электронов в процессе рекомбинации:

$$\lambda = hc/\Delta E,$$

где $h = 6,626 \cdot 10^{-34}$ Дж·с = $4,14 \cdot 10^{-15}$ эВ·с – постоянная Планка, $c = 2,998 \cdot 10^8$ м/с – скорость света.

Величина ΔE близка к энергетической ширине запрещённой зоны полупроводника ($\Delta E \approx \Delta W$). Чтобы кванты света – фотоны, освободившиеся при рекомбинации, соответствовали квантам видимого света ($0,4 \text{ мкм} < \lambda < 0,7 \text{ мкм}$), ширина запрещённой зоны должна быть относительно большой ($\Delta W > 1,8 \text{ эВ}$).

Кремний и германий являются непрямозонными полупроводниками. Для них характерна безызлучательная рекомбинация на примесных центрах.

Для изготовления светоизлучающих диодов используются кристаллы соединений типа $A_{III}B_V$ (элементов III и V групп таблицы Менделеева):

- арсенид галлия GaAs, $\Delta W = 1,42$ эВ, максимум излучения находится в инфракрасной области: $\lambda = 0,89$ мкм;
- фосфид галлия GaP, $\Delta W = 2,27$ эВ, максимум излучения расположен в зелёной области спектра: $\lambda = 0,55$ мкм;
- нитрид галлия GaN имеет наибольшую ширину запрещённой зоны: $\Delta W = 3,4$ эВ, что позволяет получать излучение в синей области, вплоть до фиолетового.

Кроме того, находят применение и другие широкозонные полупроводники, например:

- карбид кремния SiC ($A_{IV}B_{IV}$), свечение в диапазоне $\lambda = 0,56 \div 0,63$ мкм;
- сульфид цинка ZnS ($A_{II}B_{VI}$), $\Delta W = 3,74$ эВ, максимум спектрального распределения излучения на длине волны $\lambda = 0,38$ мкм (фиолетовая часть спектра).

Основными легирующими примесями в светодиодах на основе соединений $A_{III}B_V$ являются:

- элементы II группы Zn и Mg – акцепторы;
- элементы VI группы S, Se, Te – доноры.

2.1.4. Светодиод на основе гетероперехода

В p -базе светодиода вблизи границы p - n -перехода создаётся высокая концентрация электронов и дырок. Происходит их рекомбинация с испусканием фотонов. Однако через прямо-смещённый p - n -переход диффундируют не только электроны из эмиттера, но и дырки из базы. Это приводит к их рекомбинации как в p - n -переходе, так и в эмиттере и, следовательно, к снижению концентрации электронов в базе. Поэтому необходимо уменьшать дырочный ток через p - n -переход, используя эмиттеры с коэффициентом электронной инжекции близким к единице:

$$\gamma_n = I_n / (I_n + I_p) \approx 1,$$

где I_n – электронный, I_p – дырочный ток.

Идеальным эмиттером ($\gamma = 1$) является *гетеропереход*. *Гетеропереходом* называют контакт полупроводников, отличающихся шириной запрещённой зоны и величиной работы выхода. В то же время полупроводниковые материалы, образующие гетеропереход, должны иметь одну и ту же кристаллическую структуру, близкие периоды решёток (различие не более 0,5%) и коэффициенты теплового расширения. При создании элек-

трического контакта между низколегированным узкозонным ΔW_p полупроводником p -типа с большей работой выхода P_{cp} (рис.2.5,*а*) и высоколегированным широкозонным ΔW_n полупроводником n^+ -типа с меньшей работой выхода P_{cn} (рис.2.5,*б*) образуется выпрямляющий гетеропереход – эмиттер электронов (рис.2.5,*в*). При этом уровни энергии W_c и W_v имеют разрыв на металлургической границе (граница показана непрерывной вертикальной линией), а потенциальные барьеры для электронов Ψ_n и дырок Ψ_p оказываются разными. Разница в высоте потенциальных барьеров $\Delta\Psi = \Psi_p - \Psi_n = \Delta W_n - \Delta W_p$. Контактная разность потенциалов гетероперехода φ_0 равна разности термодинамических работ выхода p - и n -полупроводников, делённой на величину элементарного заряда: $\varphi_0 = [(W_0 - W_{Fp}) - (W_0 - W_{Fn})]/q$.

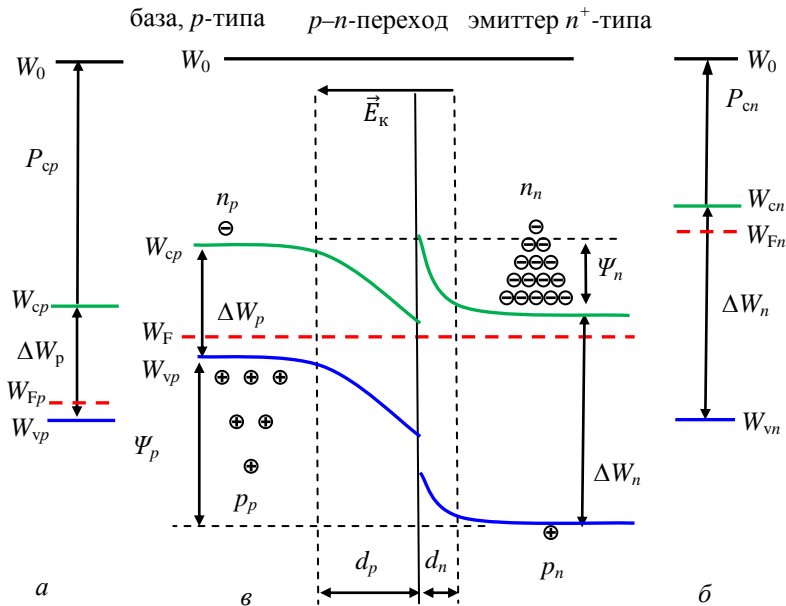


Рис. 2.5. Зонная диаграмма: *а* – узкозонного полупроводника p -типа; *б* – широкозонного полупроводника n -типа; *в* – несимметричного p - n -гетероперехода с эмиттером электронов в состоянии термодинамического равновесия

В прямосмещённом гетеропереходе (рис.2.6) при высоком уровне инжекции электронов дырочный ток практически равен нулю и $\gamma_n = 1$. В базе вблизи гетероперехода создается большая (по сравнению с гомопереходом) инверсная заселённость энергетических уровней, и квантовый выход инжекционной электролюминесценции увеличивается. Поэтому в современных светодиодах, как правило, используются гетеропереходы.

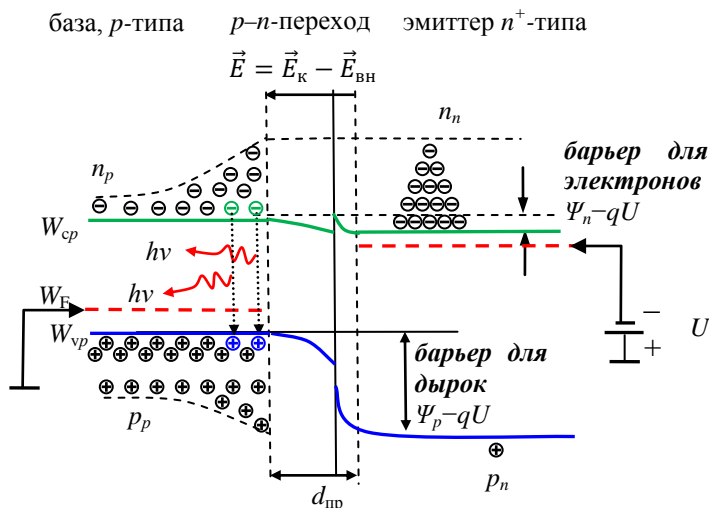


Рис. 2.6. Зонная диаграмма прямого p - n -гетероперехода

Материалы гетероструктур

Для создания гетероструктур используют кристаллы двойных, тройных и более сложных соединений. Нитриды элементов третьей группы: GaN, AlN, InN – и *твёрдые растворы* на их основе являются широкозонными полупроводниками с прямыми оптическими переходами. Твёрдым раствором называют кристалл, в котором атомы одного элемента частично замещены на атомы другого элемента, например $\text{In}_x\text{Ga}_{1-x}\text{N}$. Твёрдые растворы имеют идентичную кристаллическую структуру и близкие периоды решёток. Варьируя состав, можно получать полупроводники с различной шириной запрещённой зоны.

Гетероструктура $\text{In}_{0,06}\text{Ga}_{0,94}\text{N}/\text{GaN}$ используется в синих светодиодах.

Гетероструктура $\text{In}_{0,2}\text{Ga}_{0,8}\text{N}/\text{GaN}$ – в зелёных светодиодах.

На основе $\text{Al}_x\text{In}_y\text{Ga}_{1-x-y}\text{P}/\text{GaP}$ создают светодиоды жёлтого и красного света.

В инфракрасных светодиодах ($\lambda = 1,0 \div 1,3$ мкм) используется твёрдый раствор $\text{Ga}_x\text{In}_{1-x}\text{As}_y\text{P}_{1-y}$.

Особенности светодиодов на основе этих гетероструктур:

1) высокая интенсивность люминесценции, достигающая $2 \div 3$ кд;

2) высокое значение квантового выхода $\eta = 5,4 \%$;

3) перекрывание спектрами люминесценции всей области видимого спектра.

2.1.5. Инжекционный светодиод на основе многослойных гетероструктур

Современные инжекционные оптоэлектронные полупроводниковые приборы создаются на основе *многослойных гетероструктур*, в которых между *широкозонными* полупроводниками p - и n -типа имеется тонкий слой *узкозонного* собственного полупроводника (полупроводника i -типа). Область полупроводника i -типа является активной средой – потенциальной ямой, в которой накапливаются электроны и дырки при прямом смещении p – i – n -перехода. При высоком уровне инжекции их концентрация в i -слое становится на несколько порядков (четыре и более) выше равновесной, что приводит к их интенсивной рекомбинации.

Гетероструктуры создают методами эпитаксиального наращивания слоёв на кристаллических подложках. Типовая структура светодиода с генерацией излучения в *нелегированном активном слое* $\text{In}_x\text{Ga}_{1-x}\text{N}$ толщиной до $2 \div 3$ нм показана на рис.2.7. Активный слой InGaN имеет меньшую ширину запрещённой зоны, чем наружный слой GaN , поэтому верхний электрод является прозрачным для рекомбинационного излучения в активном слое.

Эмиттером n^+ -типа является монокристаллический слой нитрида галлия (GaN) с шириной запрещённой зоны $\Delta W_{\text{GaN}} = 3,4$ эВ. В слой GaN введены примесные атомы Si

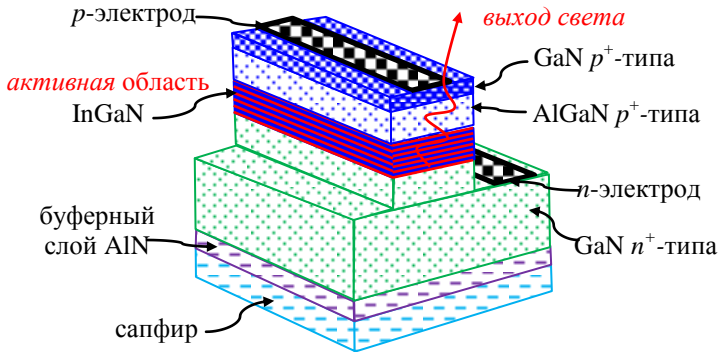


Рис. 2.7. Схема инжекционного светодиода на основе гетероструктуры AlGaN-InGaN-GaN

(IV группа), замещающие атомы Ga (III группа) и создающие высокую концентрацию электронов в зоне проводимости: $n_n = 10^{18} \div 10^{19} \text{ см}^{-3}$.

Активная область представляет собой тонкий слой ($2 \div 3 \text{ нм}$) тройного соединения $\text{In}_x\text{Ga}_{1-x}\text{N}$. При значениях параметра $x = 0,2 \div 0,4$ ширина запрещенной зоны лежит в диапазоне $\Delta W_{\text{InGaN}} = 2,8 \div 2,4 \text{ эВ}$. Примеси в активную область не вводятся.

Высоколегированной базой p^+ -типа является слой тройного соединения $\text{Al}_y\text{Ga}_{1-y}\text{N}$ с добавлением примеси магния (II группа). Ширина запрещенной зоны зависит от соотношения Al и Ga. Для $\text{Al}_{0,1}\text{Ga}_{0,9}\text{N}$ ширина $\Delta W_{\text{AlGaN}} \approx 4 \text{ эВ}$.

На легированную магнием p^+ -область GaN наносится металлический контакт никель – золото. Второй металлический контакт, титан – алюминий, наносится на нижний слой GaN n^+ -типа, как показано на рис.2.7.

На рис.2.8 представлена зонная диаграмма гетероструктуры типа $p\text{-GaN} / p\text{-AlGaN} / \text{InGaN} / n\text{-GaN}$ при прямом смещении напряжением U .

При прямом смещении гетероструктуры электроны из эмиттера и дырки из базы диффундируют в активную область (пунктирные стрелки на рис.2.8) и скапливаются в потенциальных ямах у дна зоны проводимости и верха валентной зоны соответственно. Барьеры для электронов и дырок, показанные на рис.2.8, препятствуют их дальнейшему перемещению.

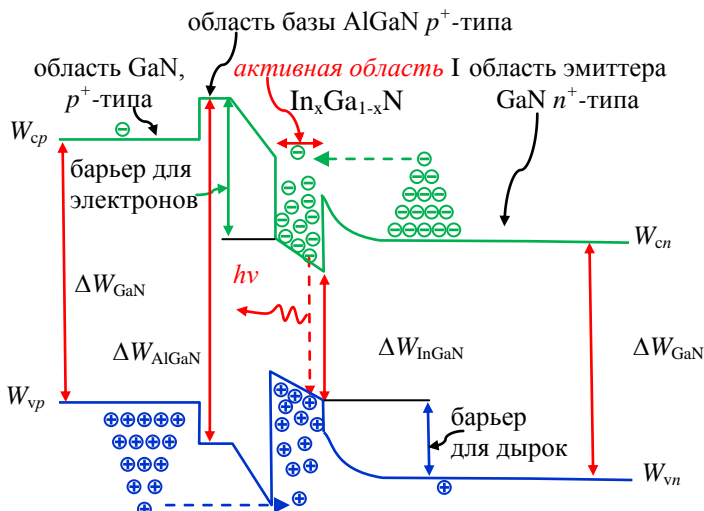


Рис. 2.8. Зонная диаграмма гетероструктуры типа p-GaN/p-AlGaN/InGaN/n-GaN при прямом смещении. Структура содержит активный слой собственного полупроводника $\text{In}_x\text{Ga}_{1-x}\text{N}$ (без легирующих добавок), в котором происходит межзонная рекомбинация с излучением фотонов

Активная область тонкая, примесей, которые могут быть центрами безызлучательной рекомбинации, в ней мало, а концентрация неравновесных носителей в потенциальных ямах велика. Это создаёт хорошие условия для прямого межзонного перехода электронов с уровня энергии дна зоны проводимости на уровень верха валентной зоны с испусканием квантов света.

Изменяя концентрацию индия и галлия в активном слое $\text{In}_x\text{Ga}_{1-x}\text{N}$, можно варьировать в нём ширину запрещённой зоны и длину волны светового излучения от 450 нм до 580 нм, что соответствует части видимого спектра от голубого до жёлтого света.

Гетеропереходы на основе соединений $\text{AlInGaP} / \text{GaP}$ покрывают длинноволновую часть видимого спектра от 580 нм до 700 нм, а на основе $\text{AlGaAs} / \text{GaAs}$ – ближнюю инфракрасную область от 700 нм до 900 нм.

Светодиод белого свечения

Светодиод белого свечения представляет собой голубой светодиод, покрытый слоем люминофора. Люминофоры – это вещества, способные преобразовывать поглощаемую ими энергию в световое излучение (люминесцировать). С гетероструктурами на основе GaN обычно используются люминофоры – смешанные с эпоксидной смолой порошки алюмоиттриевого граната $Y_3Al_5O_{12}$ с добавкой редкоземельных элементов. Такой люминофор испускает жёлто-зелёное свечение, которое, смешиваясь с голубым излучением кристалла, даёт белый свет.

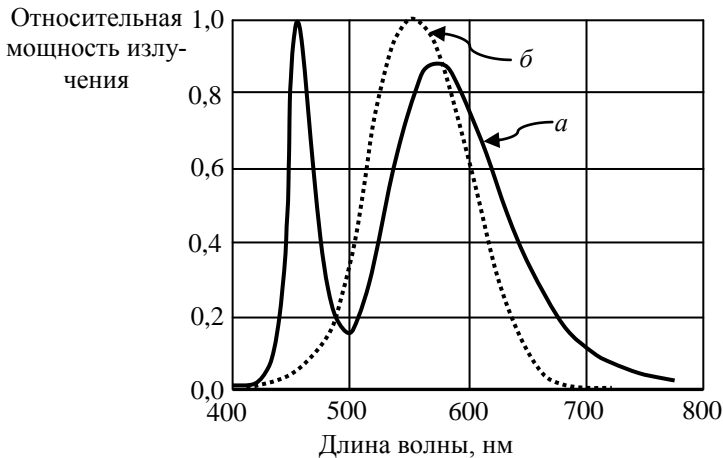


Рис.2.9. Спектр излучения “белого” светодиода: *a* – ССТ=4000 К, *б* – стандартная кривая чувствительности глаза

2.1.6. Основные характеристики и применение светодиодов

Внутренний и внешний квантовые выходы

В прямозонных полупроводниках могут происходить как излучательные, так и безызлучательные рекомбинации. Соотношение между излучательными и безызлучательными рекомбинациями характеризует внутренний квантовый выход, который является важнейшим показателем светоизлучающего диода:

$$\eta_{\text{внутр}} = N_{\text{изл}} / N_{\text{инж}}, \quad (2.4)$$

где $N_{\text{изл}}$ – число излучённых фотонов, а $N_{\text{инж}}$ – число инжектированных носителей в единицу времени.

Внутренний квантовый выход определяется в основном концентрацией центров безызлучательной рекомбинации. Поэтому в окрестности p – n -перехода стремятся снизить количество дефектов кристаллической решётки и нежелательных примесей, чтобы уменьшить скорость безызлучательной рекомбинации.

В видимой области спектра внутренний квантовый выход $\eta_{\text{внут}}$ у диодов с гомопереходом составляет единицы процентов. Поэтому в современных инжекционных светодиодах и лазерах используют гетероструктуры, в которых квантовый выход излучения значительно больше.

К числу важных параметров светодиода относится *внешний квантовый выход*

$$\eta_{\text{внеш}} = \eta_{\text{внут}} \cdot \eta_{\text{поверх}}, \quad (2.5)$$

где $\eta_{\text{поверх}}$ – коэффициент вывода света во внешнюю среду. На величину $\eta_{\text{поверх}}$ оказывают влияние ряд факторов:

Поглощение света полупроводником. В светодиодах с гомопереходом имеет место высокий коэффициент поглощения, поскольку полупроводник поглощает кванты с той же длиной волны, что и испускает. Чтобы снизить поглощение, глубину залегания p – n -перехода уменьшают до $1 \div 2$ мкм.

В светодиодах с гетеропереходом свет выходит через слои эмиттера или базы, в которых ширина запрещённой зоны больше, чем энергия кванта, испущенного в активном слое. Верхний электрод делают кольцевым, а нижний сплошной электрод выполняет функцию зеркала.

Отражение излучения от границы раздела полупроводник–воздух. Полупроводниковые материалы имеют высокий коэффициент преломления: $n = 3,3 \div 3,6$. Наружу может выйти только то излучение, которое падает на поверхность раздела под углом, меньшим критического: $\theta_{\text{крит}} = \arcsin n^{-1}$. Для фосфида галлия GaP этот угол $\theta_{\text{крит}} = 17,7^\circ$. С целью увеличения критического угла кристаллы полупроводника заливают полимерными компаундами с показателем преломления $n = 1,5 \div 2,0$ (рис.2.10).

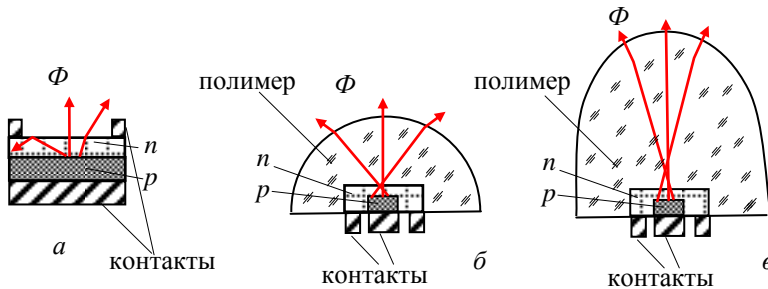


Рис.2.10. Конструкция светодиодов: *а* – плоская планарная, $\eta_{\text{вн}} = 13\%$; *б* – плоская планарная с полусферическим покрытием, $\eta_{\text{вн}} = 37\%$, *в* – плоская планарная с параболическим покрытием

Световая характеристика светодиода

Характеристикой диода как источника света является зависимость *силы света* I_v от прямого тока (световая характеристика). Сила света I_v – излучаемый диодом световой поток, приходящийся на единицу телесного угла в направлении, перпендикулярном плоскости излучающего кристалла. Измеряется в канделах (кд).

Сила света диода пропорциональна произведению квантового выхода (2.5) на число актов излучательной рекомбинации в единицу времени. На рис.2.11, *а* и *б* представлены зависимости внутреннего квантового выхода $\eta_{\text{внутр}}$ и силы света I_v от плотности прямого тока j . При малых значениях j происходит в основ-

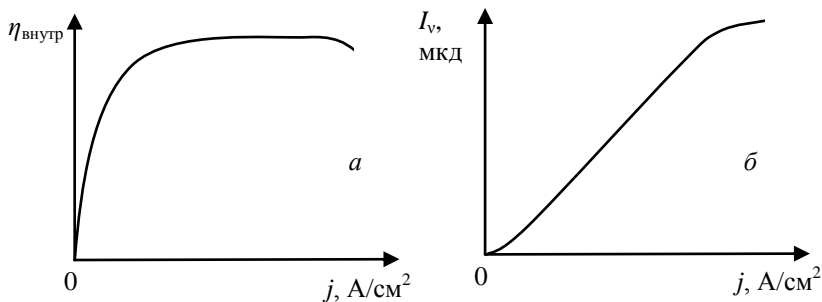


Рис. 2.11, Зависимости: *а* – внутреннего квантового выхода $\eta_{\text{внутр}}$ от плотности тока j ; *б* – силы света светодиода I_v от плотности тока j

ном безызлучательная рекомбинация носителей на примесных центрах в области p – n -перехода. Вследствие этого $\eta_{\text{внутр}}$ при малых значениях тока невелик и резко возрастает с увеличением инжекции носителей и их межзонной рекомбинации. При дальнейшем увеличении плотности тока имеется почти линейный участок световой характеристики $I_v(j)$, протяжённость которого определяется изменением силы света светодиода в пределах одного – двух порядков ($\eta_{\text{внутр}} \approx \text{const}$).

Светодиоды обладают высоким быстродействием. После подачи импульса прямого тока излучение нарастает за время менее 10^{-8} с.

Основные параметры светоизлучающих диодов

Основными параметрами светоизлучающих диодов являются следующие:

сила света I_v – излучаемый диодом световой поток, приходящийся на единицу телесного угла в направлении максимального излучения. Указывается при заданном значении прямого тока $I_{\text{пр}}$, мА и измеряется в милликанделах (мкд);

яркость L – величина, равная отношению силы света к площади светящейся поверхности (для мнемонических и много-сегментных индикаторов). Указывается при заданном значении прямого тока $I_{\text{пр}}$, мА и измеряется в $\text{кд}/\text{м}^2$;

мощность излучения (для излучающих диодов ИК-диапазона). Указывается при заданном значении $I_{\text{пр}}$, мА и измеряется в милливаттах (мВт);

постоянное прямое напряжение $U_{\text{пр}}$, В – значение напряжения на светодиоде при протекании заданного прямого тока;

максимум спектрального распределения λ_{max} , мкм – длина волны светового излучения, соответствующая максимуму спектральной характеристики светодиода;

спектральная характеристика светодиода – зависимость интенсивности излучения от длины волны излучаемого света;

диаграмма направленности – зависимость силы света от угла, под которым ведётся наблюдение (излучение может быть узконаправленным или рассеянным);

вольт-амперная характеристика – аналогична характеристике обычного диода;

зависимость силы света от температуры (в интервале рабочих температур I_v может меняться в 2÷3 раза, с повышением температуры сила света уменьшается);

предельно допустимые режимы эксплуатации – указываются максимальные значения прямого тока $I_{пр. \max}$, мА, обратного напряжения $U_{обр. \max}$, В и другие параметры, при которых обеспечивается заданная надёжность работы прибора.

Применение светодиодов

Условное обозначение светодиода приведено на рис.2.12.

Светодиоды в основном применяются как элементы индикации включения, готовности аппаратуры к работе, наличия напряжения питания в блоке, аварийной ситуации и других состояний.

Линейные шкалы (столбики из светодиодов) служат для отображения непрерывно изменяющейся информации – уровня звукового сигнала, уровня горючего в баке и т.д. Достоинство их – наглядность отображения информации.

Цифро-буквенные индикаторы используются в измери-

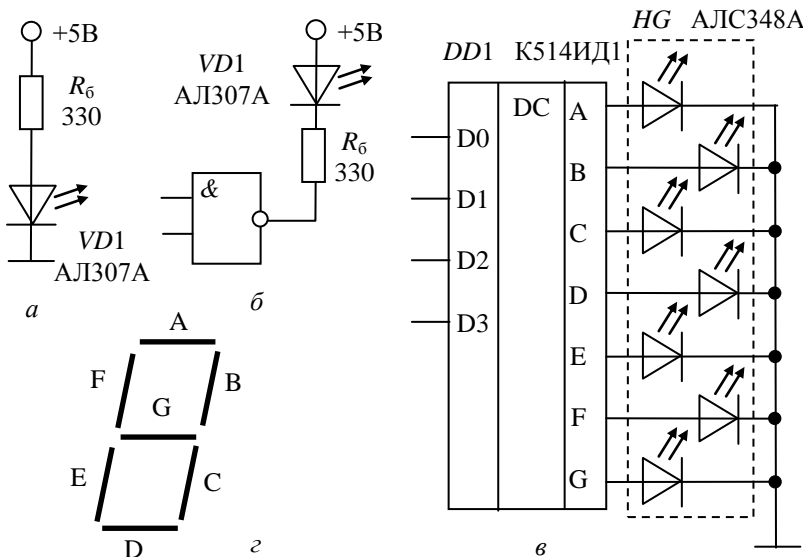


Рис. 2.12. Схемы включения светоизлучающих диодов

тельной аппаратуре, устройствах автоматики и вычислительной техники. (рис.2.12,з).

Инфракрасные излучающие диоды (ИК-диоды) применяются в оптических каналах передачи информации, датчиках систем автоматики, охранной сигнализации и т.д. Излучение ИК-диода не может быть воспринято человеческим глазом и регистрируется фотодиодом или фототранзистором.

Ввиду нелинейности вольт-амперной характеристики (резкого возрастания прямого тока при прямом напряжении выше $2\div 3$ В) светодиоды всегда подключают к источнику питания через балластное сопротивление (рис.2.12,а и б), величина которого может быть найдена из соотношения

$$R_б = (E - U_{пр}) / I_{пр}. \quad (2.6)$$

Обычно информация, предназначенная для отображения цифровым или шкальным индикатором, поступает в двоичном коде. Для преобразования её в натуральный (для шкал) или 7-сегментный код используются микросхемы – дешифраторы (рис.2.12,в). Некоторые типы дешифраторов имеют токовые выходы (создают заданную величину тока независимо от величины падения напряжения на диоде). В этом случае светодиоды индикатора можно подключать к выходам дешифратора без балластных резисторов (рис.2.12,г).

2.1.7. Оптопары

Комбинацией опто- и фотоэлектронных приборов является оптопара (оптрон). Структура оптопары приведена на рис.2.13. В ней входной электрический сигнал $I_{вх}$ преобразуется в оптический сигнал Φ , в качестве преобразователя обычно используют светодиод или полупроводниковый лазер (излучатель). Оптический сигнал Φ распространяется через оптический канал. Канал может быть открытым (вакуум, газ или жидкость) и закрытым (оптический световод). На выходе оптического канала имеется

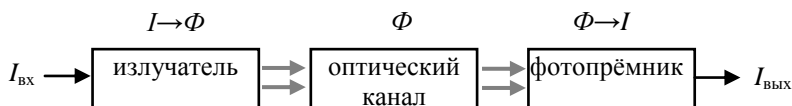


Рис.2.13. Структурная схема оптопары

фотоэлектронный преобразователь – фотоприёмник (фоторезистор, фотодиод, фототранзистор или фототиристор) в котором оптический сигнал Φ преобразуется в электрический $I_{\text{вых}}$. Двойное преобразование входного сигнала: $I \rightarrow \Phi$ и $\Phi \rightarrow I$ – позволяет получить гальваническую развязку входного и выходного сигналов, невосприимчивость оптических каналов к электромагнитным помехам, однонаправленность потока информации, отсутствие обратной связи с выхода на вход и широкую полосу пропускания. К недостаткам оптопар следует отнести нелинейность их проходных характеристик, низкий КПД и температурную зависимость параметров.

В диодной оптопаре (рис.2.14,*а* и 2.15,*б*) излучателем служит инфракрасный светодиод на основе арсенида галлия, фотоприёмный элемент – фотодиод на основе кремния. Максимум спектральной характеристики GaAs светодиода расположен около $\lambda_{\text{max}} \approx 1$ мкм. При облучении фотодиода инфракрасным светом с $\lambda \sim 1$ мкм в нём возникает генерация пар носителей заряда – электронов и дырок (внутренний фотоэффект). Интенсивность генерации пропорциональна силе света и, следовательно, входному току оптопары (току через светодиод).

Оптроны обычно используются в фотодиодном режиме (рис.2.14,*а*). К фотодиоду оптопары прикладывают обратное напряжение $U_{\text{обр}} > 0,5$ В. Электроны и дырки, генерируемые излучением, увеличивают обратный ток I_0 p – n -перехода:

$$I_{\text{вых}} = I_0 + I_{\text{фототок}} \approx I_{\text{фототок}}.$$

Передающая характеристика в фотодиодном режиме $I_{\text{вых}} = I_{\text{вых}}(I_{\text{вх}})$ практически линейна. Коэффициент передачи тока диодных оптопар:

$$K_I = I_{\text{вых}}/I_{\text{вх}} \quad (2.7)$$

составляет единицы процентов и примерно равен значению внешнего квантового выхода светодиода.

В транзисторной оптопаре (рис. 2.14,*б* и 2.15,*в*) светодиод конструктивно расположен так, что излучение направляется в базовую область фототранзистора. При отсутствии света в цепи коллектора фототранзистора протекает ток

$$I_{\text{тем}} \approx h_{21\beta} I_{\text{ко}},$$

где $I_{\text{ко}}$ – обратный ток p – n -перехода база – коллектор.

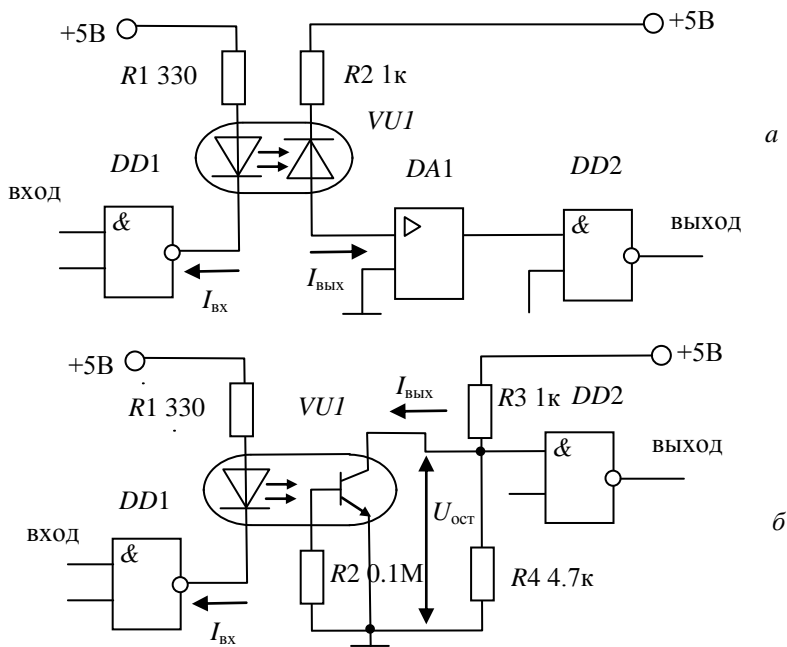


Рис.2.14. Схема включения оптопар: *а* – диодной, *б* – транзисторной. DD1, DD2 – интегральные ТТЛ схемы, DA1 – усилитель-формирователь импульсов

При облучении светом в базовой области генерируются пары электрон – дырка. Электроны втягиваются в область коллектора полем обратнсмещённого перехода база – коллектор. Дырки остаются в базе и увеличивают её потенциал, что приводит к инжекции электронов из эмиттера в базу. Соотношение фототоков коллектора $I_{\text{ВЫХ}}$ и базы

$$I_{\text{ВЫХ}} = h_{21\beta} I_{\text{ф.б}}, \quad (2.8)$$

где $h_{21\beta}$ – коэффициент передачи тока базы транзистора ($h_{21\beta} \approx 10^2$), $I_{\text{ф.б}}$ – генерируемый излучением фототок в базе транзистора (дырочная составляющая фототока).

Коэффициент передачи по току транзисторных оптопар $K_I = I_{\text{ВЫХ}}/I_{\text{ВХ}}$ имеет величину порядка единицы. Передаточная характеристика транзисторной оптопары $I_{\text{ВЫХ}} = I_{\text{ВЫХ}}(I_{\text{ВХ}})$ нелинейна. Транзисторные оптопары применяют обычно в ключевом режиме для передачи импульсных сигналов (рис.2.14,б).

Транзисторные оптопары используют для гальванической развязки логических цепей управления от высоковольтных цепей нагрузок большой мощности: для управления мощными тиристорами, в устройствах защиты вторичных источников питания и т.д.

Основные параметры оптопар

Условные обозначения микросхем светодиодных оптопар приведены на рис.2.15.

В системе параметров оптопар выделяют

входные параметры (параметры излучателя):

- входной номинальный ток оптопары $I_{\text{ВХ ном}}$, мА,
- входное напряжение $U_{\text{ВХ}}$, В при $I_{\text{ВХ ном}}$,
- максимальный входной ток $I_{\text{ВХ max}}$, мА,
- максимальное входное обратное напряжение $U_{\text{ВХ обр max}}$, В;

выходные параметры (параметры приёмника):

- максимально допустимое обратное выходное напряжение $U_{\text{ВЫХ обр max}}$, В,
- максимально допустимый выходной ток $I_{\text{ВЫХ max}}$, мА,
- выходной темновой ток $-I_{\text{ВЫХ т}}$ (или ток утечки на выходе $-I_{\text{ут}}$) при $I_{\text{ВХ}} = 0$ и заданном значении $U_{\text{ВЫХ}}$,
- выходное остаточное напряжение (на открытом фототранзисторе) $U_{\text{ост}}$, В;

передаточные характеристики:

- статический коэффициент передачи тока $K_I = I_{\text{ВЫХ}}/I_{\text{ВХ}}$,
- время нарастания импульсного выходного тока $t_{\text{нр}}$, с,
- время спада импульсного выходного тока $t_{\text{сп}}$, с.

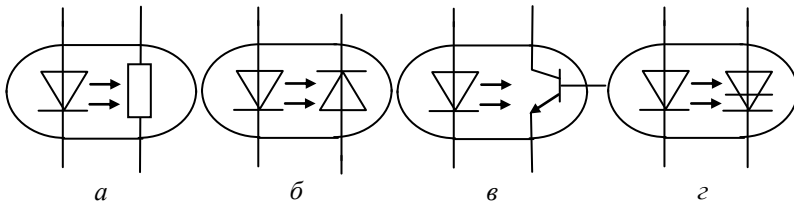


Рис.2.15. Условные обозначения светодиодных оптопар: а – фоторезисторная, б – фотодиодная, в – фототранзисторная, г – фототиристорная

2.2. Полупроводниковый лазер и суперлюминесцентный диод

2.2.1. Инжекционный полупроводниковый лазер

Лазер – это устройство, преобразующее энергию накачки (энергию внешнего источника) в энергию когерентного, монохроматического, поляризованного и узконаправленного потока излучения.

Слово “лазер” (laser) является аббревиатурой от полного английского названия “light amplification by stimulated emission of radiation”, что переводится на русский как “усиление света посредством вынужденного излучения”. Под термином “свет” подразумевают электромагнитное излучение в ультрафиолетовой ($0,1\div 0,4$ мкм), видимой ($0,4\div 0,8$ мкм) и инфракрасной ($0,8\div 50$ мкм) областях спектра. В отечественной литературе наряду с термином “лазер” используется название “*оптический квантовый генератор*” (ОКГ).

Необходимыми элементами любого лазера являются (рис.2.16):

- *активная среда* – рабочее вещество, в котором создаётся неравновесная инверсная заселённость энергетических уровней;
- *система накачки* – система в которой энергия внешнего воздействия преобразуется во внутреннюю энергию *активной среды*;
- *оптический резонатор* – элемент, в котором происходит интенсивное взаимодействие излучения с веществом активной среды, в результате которого внутренняя энергия активной сре-

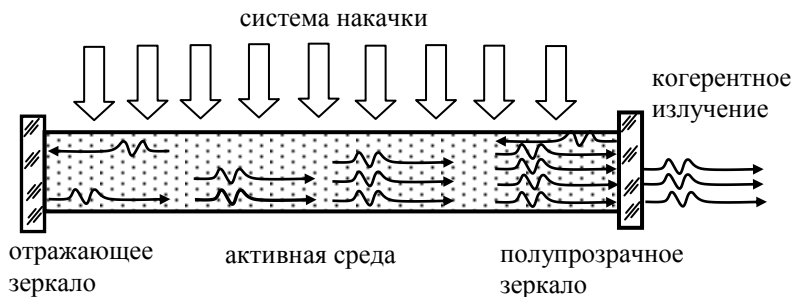


Рис. 2.16. Схема лазера

ды преобразуется в энергию когерентного светового излучения. Оптический резонатор – объём между отражающим и полупрозрачным зеркалами. Полупрозрачное зеркало обеспечивает вывод когерентного светового излучения во внешнюю среду.

Схема *инжекционного полупроводникового лазера* представлена на рис.2.17. Активной средой 1 является тонкий слой собственного полупроводника (*i*-типа). Роль системы накачки 2 выполняют слои полупроводников *n*- и *p*-типов, инжектирующие в *i*-слой электроны и дырки. Зеркалами 3 и 4 служат грани полупроводникового кристалла. Энергия, выделяющаяся при рекомбинации электронов и дырок, преобразуется в энергию светового излучения 5, которое выводится через грань кристалла 4. Излучение, вышедшее через противоположную грань 3, как правило, поглощается специальным поглотителем 6, например, зачернённой поверхностью корпуса.

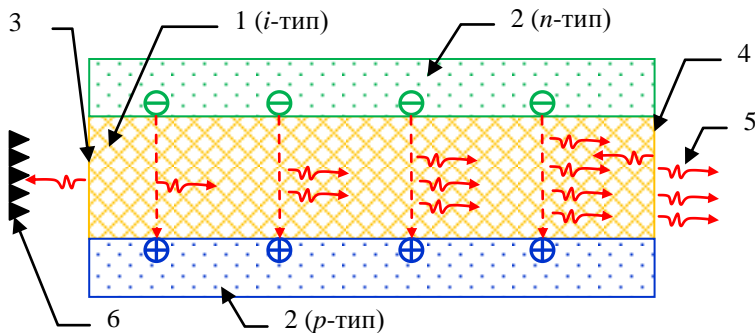


Рис.2.17. Схема инжекционного полупроводникового лазера:
1 – активная среда; 2 – система накачки; 3 и 4 – грани кристалла;
5 – когерентное излучение; 6 – поглотитель

2.2.2. Физические принципы работы инжекционного лазера

Люминесценция – нетепловое свечение вещества, атомы которого приведены в возбужденное состояние. В светодиодах люминесценция возникает в результате спонтанных излучательных переходов электронов с уровней энергии дна зоны проводимости на уровни энергии верха валентной зоны. Излучение является некогерентным и ненаправленным.

Суперлюминесценция – люминесценция веществ с инверсной заселённостью энергетических уровней, вызванная усилением спонтанного излучения за счёт индуцированного излучения. Это явление наблюдается в суперлюминесцентных диодах и инжекционных лазерах. Излучение лазеров является когерентным и направленным. В суперлюминесцентных диодах в отличие от лазеров нет оптического резонатора, поэтому спектр излучения у них более широкий, чем у лазера, но более узкий по сравнению со светодиодным.

Взаимодействие излучения с атомами активной среды инжекционного лазера

Активной средой в инжекционном лазере является тонкий слой собственного полупроводника i -типа. На рис.2.18 изображены энергетические уровни дна зоны проводимости W_c и верха валентной зоны W_v .

Электрон может самопроизвольно (спонтанно) перейти с уровня зоны проводимости в более низкое энергетическое состояние, на уровень валентной зоны (рекомбинировать с дыркой, рис.2.18,а), испустив при этом квант света (фотон):

$$h\nu \geq W_c - W_v.$$

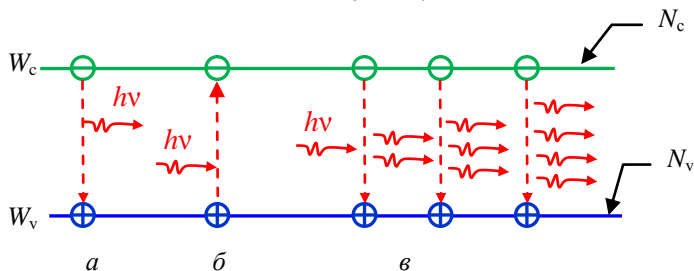


Рис.2.18. Процессы в активной среде инжекционного лазера: a – спонтанное излучение; b – поглощение; $в$ – стимулированное излучение

Спонтанные переходы являются случайными и не зависят от внешнего электромагнитного поля. Излученные фотоны имеют случайные частоту (в пределах полосы спектра испускания), фазу, поляризацию и направление распространения. Спонтанное излучение некогерентное. Такое излучение создают все

естественные источники света: Солнце, лампы накаливания и т.д.

Другим оптическим процессом является поглощение кванта света, сопровождающееся переходом электрона с уровня W_v на уровень W_c (рис.2.18,б).

В основе работы лазеров лежит явление *индуцированного излучения*. Фотон с энергией

$$h\nu \geq (W_c - W_v)$$

может инициировать процесс рекомбинации электрона и дырки. Испускаемый фотон идентичен инициирующему фотону (они имеют одинаковые частоту, фазу, поляризацию и направление распространения).

При распространении световой волны вдоль активной среды интенсивность индуцированного излучения экспоненциально возрастает:

$$I = I_0 \exp(\alpha x).$$

Коэффициент α называют коэффициентом усиления активной среды.

В реальной активной среде происходит также обратный процесс – снижение интенсивности излучения в результате поглощения, а также при рассеянии на неоднородностях или примесях, которые преобразуют энергию излучения в тепло. Эти потери учитывают в коэффициенте β . Тогда интенсивность когерентного излучения в среде определяется как:

$$I = I_0 \exp[(\alpha - \beta)x]. \quad (2.9)$$

Если $\alpha > \beta$, то интенсивность световой волны растёт при прохождении через активную среду. На рис.2.19 показано экспоненциальное увеличение амплитуды падающей затравочной световой волны при прохождении через активную среду:

$$E = E_0 \exp(\alpha' x),$$

где $\alpha' = \frac{\alpha - \beta}{2}$. Цифра “2” в знаменателе возникла потому, что интенсивность световой волны пропорциональна квадрату её амплитуды: $I \sim E^2$.

Как и в любом генераторе, для работы лазера необходимо выполнить два условия: условие баланса амплитуд и условие баланса фаз. Энергия, поступающая от источника питания в ак-

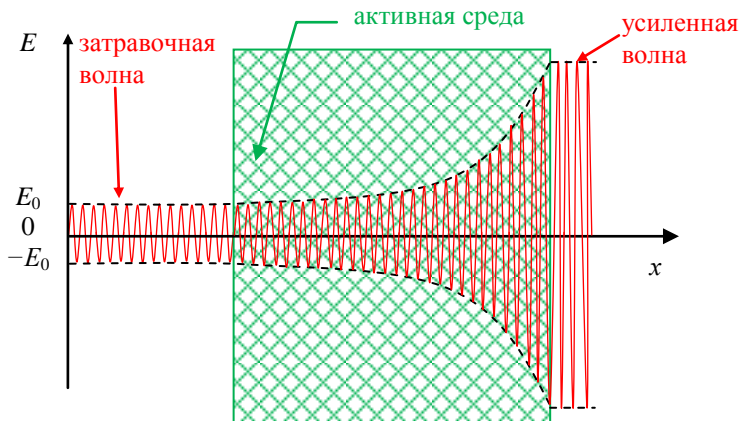


Рис. 2.19. Усиление амплитуды световой волны в активной среде

тивную среду оптического резонатора, равна энергии лазерного излучения плюс энергия потерь в самом резонаторе.

Условие баланса фаз обеспечивается подбором длины оптического резонатора l таким образом, чтобы на частоте индуцированного излучения внутри него возникала стоячая волна:

$$l = \frac{\lambda_q}{2} q, \quad (2.10)$$

где λ_q – длина волны лазерного излучения в резонаторе, $q = 1, 2, 3, \dots$ – число полуволн излучения, укладывающихся на длине резонатора.

На рис.2.20 показано распределение напряжённости электрического поля в активной среде оптического резонатора E_p

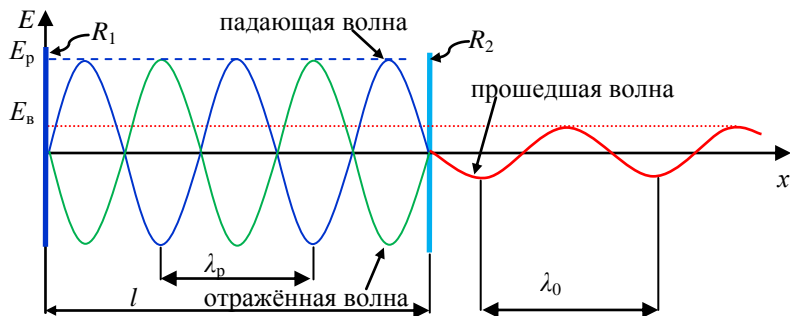


Рис.2.20. Распределение напряжённостей электрического поля E

(стоячая волна) и воздухе E_v (бегущая волна). Также указаны коэффициенты отражения R_1 и R_2 от граней полупроводникового кристалла. Напряжённость электрического поля прошедшей волны

$$E_v = (1 - R_2) E_p. \quad (2.11)$$

На рис.2.21 и 2.22 показаны изменения мощности и спектра излучения при переходе из светодиодного в лазерный режим с увеличением плотности инжекционного тока j . Видно, что при плотности тока выше пороговой $j_{\text{пор}}$ мощность излучения резко возрастает, а спектр излучения сужается.

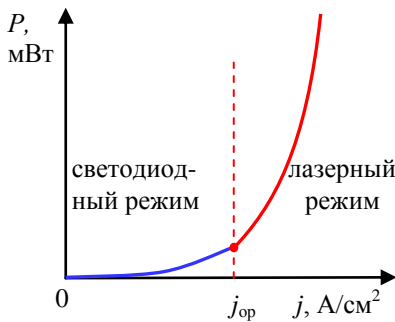


Рис.2.21. Зависимость мощности излучения P от плотности тока j

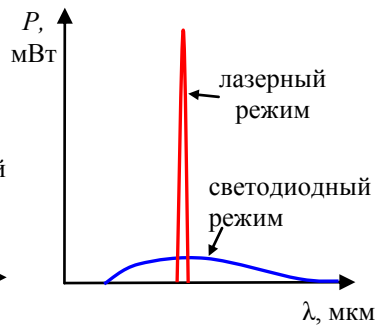


Рис.2.22. Изменение спектра при переходе из светодиодного в лазерный режим излучения

2.2.3. Инжекционный лазер на основе многослойных гетероструктур

Современные инжекционные лазеры изготавливают на основе многослойных гетероструктур, аналогичных структурам светодиодов. За работы по созданию полупроводниковых лазеров на гетеропереходах, способных работать при комнатной температуре и эффективно использоваться для накачки твёрдотельных лазеров, Ж.И. Алфёров удостоен Нобелевской премии по физике 2000 г.

На рис.2.23 показана структура инжекционного лазерного диода инфракрасного излучения, изготовленного на основе гетероперехода GaAs–AlGaAs. На подложку наращивается эпитаксиальный слой тройного соединения $\text{Al}_x\text{Ga}_{1-x}\text{As}$ с акцепторной примесью (полупроводник p -типа). Затем наращиваются

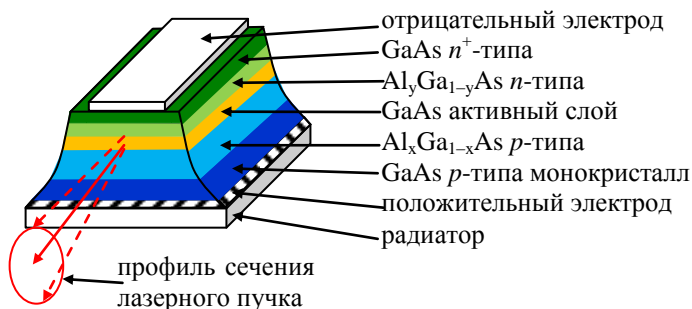


Рис.2.23. Структура инжекционного лазерного диода

тонкий слой собственного полупроводника GaAs и слой $\text{Al}_x\text{Ga}_{1-x}\text{As}$ с донорной примесью (полупроводник n -типа). Следующим наращивается высоколегированный слой GaAs n^+ -типа, на который напыляется отрицательный электрод.

Ширина запрещённой зоны ΔW в соединении $\text{Al}_x\text{Ga}_{1-x}\text{As}$ больше, чем в арсениде галлия GaAs. Таким образом, между двумя широкозонными полупроводниками $\text{Al}_x\text{Ga}_{1-x}\text{As}$ p -типа и $\text{Al}_y\text{Ga}_{1-y}\text{As}$ n -типа имеется тонкий слой (толщиной $1\div 3$ мкм) узкозонного полупроводника GaAs i -типа. Этот i -слой является активной средой – потенциальной ямой, в которой накапливаются электроны и дырки при прямом смещении p – i – n -перехода. При высоком уровне инжекции в i -слое происходит интенсивная рекомбинация электронов и дырок. Арсенид галлия GaAs – прямозонный полупроводник, ширина запрещённой зоны $\Delta W = 1,42$ эВ, максимум излучения лежит в инфракрасной области спектра $\lambda = 0,89$ мкм.

Зеркалами резонатора являются грани самого полупроводникового кристалла, получаемые обычно скалыванием его краёв перпендикулярно к плоскости активного слоя. Вынужденное излучение в активном слое генерируется параллельно p – i – n -переходу вдоль длинной стороны кристалла. Для предотвращения возникновения паразитного лазерного излучения в поперечном направлении боковые части кристалла стравливают под некоторым углом.

Типичные размеры лазерного кристалла: длина от 100 до 500 мкм, ширина от 200 до 400 мкм, высота от 80 до 100 мкм.

Толщина активного слоя на два порядка меньше его ширины и соизмерима с длиной волны лазерного излучения, поэтому вследствие явления дифракции при выходе излучения из активного слоя профиль сечения лазерного пучка получается эллиптическим, как показано на рис.2.23.

Инжекционные лазеры от других типов лазеров отличаются высоким КПД (до 80% в инфракрасной области спектра), простотой возбуждения, малыми размерами, низким напряжением источника питания. Перестройка лазера по частоте возможна в небольших пределах за счёт изменения плотности тока накачки и длины объёмного резонатора l при нагревании. Фотографии лазерных модулей приведены на рис.2.24, *а* и *б*.

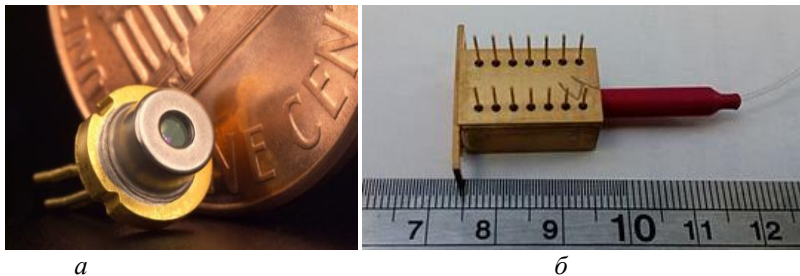


Рис.2.24. Лазерный диод: *а* – с открытым выходом; *б* – с оптоволоконным выходом

Лазерные модули, использующиеся в волоконной оптике, обычно включают в себя излучатель, термистор, элемент Пельтье, фотодиод, коллимирующую линзу и оптический изолятор. Для стабилизации частоты лазерного излучения температура гетероструктуры в процессе работы должна поддерживаться постоянной. Для этого используют внешний контроллер, датчиком температуры которого является термистор, а выходной ток, протекающий через элемент Пельтье, поддерживает постоянную температуру гетероструктуры. Коллимирующая линза служит для ввода лазерного пучка в оптическое волокно, а оптический изолятор защищает лазер от повреждения отраженным излучением.

2.2.4. Основные характеристики лазерного излучения

По структуре пучка излучения лазеры делятся на одномодовые и многомодовые. Лазером излучается плоская электромагнитная волна типа TEM (TEM-wave; transverse electromagnetic wave) или близкая к ней. Внутри оптического резонатора волны можно считать поперечными TEM_{mnq} , здесь индексы $m = 0, 1, 2, 3, \dots$, $n = 0, 1, 2, 3, \dots$ – горизонтальный и вертикальный индексы поперечной моды (m – число полуволн, укладываемых по ширине и n – число полуволн, укладываемых по толщине активной области), $q = 1, 2, 3, \dots$, – индекс продольный моды (см. рис. 2.20).

В одномодовом режиме генерации $m = n = 0$. Индекс q равен числу полуволн $\lambda_q/2$, уложившихся вдоль оси резонатора l :

$$q = \frac{l}{\lambda_q/2}.$$

Следует отметить, что вдоль реальных оптических резонаторов укладываются сотни – тысячи длин полуволн ($q \sim 10^3$).

Излучение инжекционного лазера в общем случае является многомодовым. Спектр излучения состоит из большого числа продольных мод:

$$\lambda_q = \frac{l}{q/2}, \quad q = q_0, (q_0 \pm 1), (q_0 \pm 2), (q_0 \pm 3) \dots, \quad (2.11)$$

находящихся в пределах ширины спектра излучения лазерного

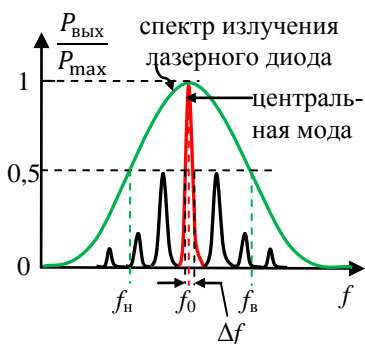


Рис.2.25. Режим генерации продольных мод лазерного диода

диода (рис.2.25). Наибольшую интенсивность имеет мода с максимальным коэффициентом усиления, находящаяся в центре спектра излучения диода.

Малая ширина лазерной линии $\Delta f = \frac{c}{\lambda^2} \Delta \lambda$ на уровне половинной мощности определяет *монохроматичность* лазерного излучения.

Одной из основных характеристик лазерного излучения яв-

ляется его *когерентность*. Различают временную и пространственную когерентности.

Время когерентности $\tau_{\text{ког}}$ определяется шириной полосы лазерного излучения Δf :

$$\tau_{\text{ког}} = \frac{1}{2\Delta f} = \frac{\lambda^2}{2c\Delta\lambda}. \quad (2.12)$$

Длина продольной когерентности $l_{\text{ког}} = c\tau_{\text{ког}}$ – это максимальная разность хода интерферирующих лучей, при которой ещё наблюдается их интерференция. Таким образом, чем уже спектр излучения, тем больше время и длина когерентности.

Максимальное расстояние между точками лазерного пучка, излучение в которых будет интерферировать, называется *длиной поперечной когерентности* (пространственной когерентности). Для многих лазеров длина поперечной когерентности равна диаметру пучка лазерного излучения.

При работе лазера в одномодовом режиме можно сфокусировать излучение в пятно малых размеров. На рис.2.26 показана диаграмма направленности излучения инжекционного лазерного диода, структура которого изображена на рис.2.23. Угловая ширина луча на уровне половинной мощности в горизонтальной плоскости составляет $\Delta\psi$, а в вертикальной – $\Delta\varphi$. Таким образом, пучок имеет эллиптическое сечение (рис.2.26,в).

Обычно величина $\Delta\psi \sim 1$ градус, $\Delta\varphi \sim 10$ градусов. Расходимость излучения является существенным недостатком инжек-

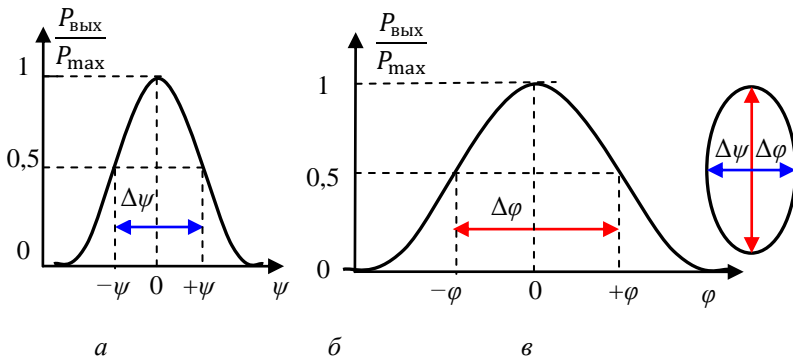


Рис.2.26. Диаграмма направленности излучения лазерного диода в плоскости: а – горизонтальной; б – вертикальной, в – в разрезе

ционных лазеров, поэтому для её уменьшения используют оптические линзы.

Инжекционные лазеры, изображенные на рис.2.23, генерируют излучение с вертикальной поляризацией.

Коэффициент полезного действия (КПД) лазера

$$\eta = \frac{P_{\text{изл}}}{P_{\text{нак}}} \cdot 100\%, \quad (2.13)$$

где $P_{\text{изл}}$ и $P_{\text{нак}}$ – мощности лазерного излучения и накачки соответственно. КПД инжекционных лазеров может достигать 80% в инфракрасном диапазоне.

2.2.5. Суперлюминесцентный диод

Конструкция и принцип действия суперлюминесцентного диода

Схема суперлюминесцентного диода (СЛД, англ. super luminescent diode, SLD) представлена на рис.2.27. Как и в полупроводниковом лазере (рис.2.17), в нём имеются активная среда 1 (активный слой собственного полупроводника) и система накачки, образованная полупроводниками n - и p -типов 2 и 3. Однако у СЛД отсутствуют отражающие зеркала. Для уменьшения отражения от задней поверхности 4 её скашивают под углом 10° либо на неё наносят слой поглощающего материала. Со стороны оптического выхода нанесено просветляющее покрытие 5. Такая схема обеспечивает условие однократного прохождения

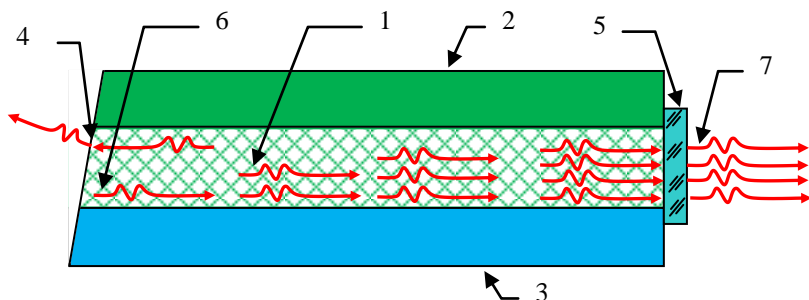


Рис.2.27. Схема суперлюминесцентного лазерного диода: 1 – активная среда; 2 – полупроводник n -типа; 3 – полупроводник p -типа; 4 – скошенная грань; 5 – просветляющее покрытие; 6 – затравочная волна; 7 – когерентное излучение

оптического излучения через активную среду. Затравочная волна 6, возникающая в результате спонтанного излучения, проходя через активную среду, усиливается, и на выходе СЛД возникает когерентное излучение 7. Поэтому мощность выходного излучения СЛД выше, чем у люминесцентного светодиода, и сравнима с мощностью аналогичного инжекционного полупроводникового лазера. Её типичная величина составляет десятки – сотни мВт, существуют и более мощные СЛД. Для увеличения выходной мощности длину активной среды СЛД увеличивают по сравнению с длиной инжекционных лазеров.

На рис.2.28,а,б и в приведены нормированные спектральные полосы излучения люминесцентного светодиода, суперлюминесцентного диода и инжекционного лазера соответственно. Типичная ширина спектральной полосы излучения суперлюминесцентного диода составляет $20 \div 50$ нм, что примерно на порядок меньше, чем у светодиода, и на один-два порядка больше, чем у лазера.

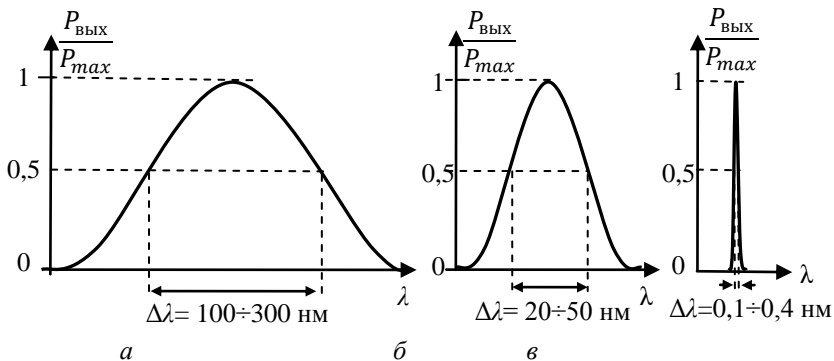


Рис.2.28. Нормированные спектральные характеристики: а – люминесцентного светодиода, б – суперлюминесцентного диода, в – одномодового инжекционного лазера

Для удобства эксплуатации СЛД выпускают в составе передающих оптоэлектронных модулей. Основными элементами модуля являются суперлюминесцентный диод, микрохолодильник, фотодиод обратной связи, оптический изолятор и одномодовый световод с цилиндрической линзой в начале световода.

Оптический изолятор в СЛД может заменяться схемой оптической развязки, представленной на рис.2.29.

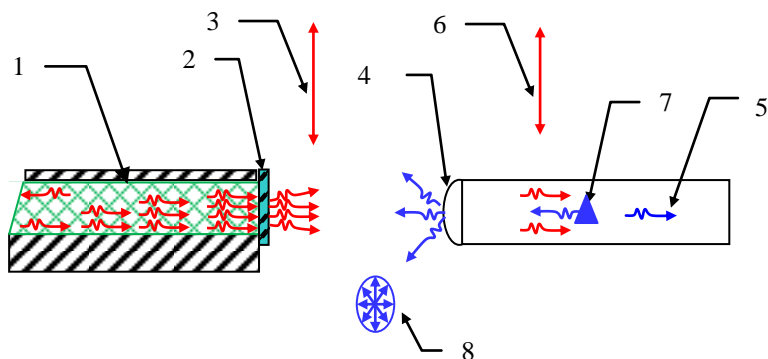


Рис.2.29. Оптическая развязка суперлюминесцентного диода

Оптическое излучение, возникшее в активной среде 1, направлено через просветляющее покрытие 2 в воздух в виде расходящегося в вертикальной плоскости пучка с вертикальной поляризацией 3. Цилиндрическая собирающая линза 4, приклеенная к торцу оптического волокна, фокусирует в его сердцевину 5 излучение с вертикальной поляризацией 6. Возникшее на оптической неоднородности 7 отражённое излучение имеет эллиптическую поляризацию 8 и при вторичном прохождении цилиндрической линзы 4 частично рассеивается, а частично возвращается в активную среду СЛД. При этом амплитуда вертикальной составляющей отражённого излучения, которая может нарушить режим генерации, оказывается ослабленной на несколько порядков.

Принцип действия просветляющего покрытия в суперлюминесцентном лазерном диоде

Принцип действия просветляющего покрытия поясняется на рис.2.30. Слева на нем изображена активная область с показателем преломления n_a , в середине – тонкая плёнка просветляющего покрытия с показателем преломления $n_{пл}$, а справа – воздух с $n_b = 1$. Толщина просветляющего покрытия $d = \frac{\lambda_{пл}}{4}$, где

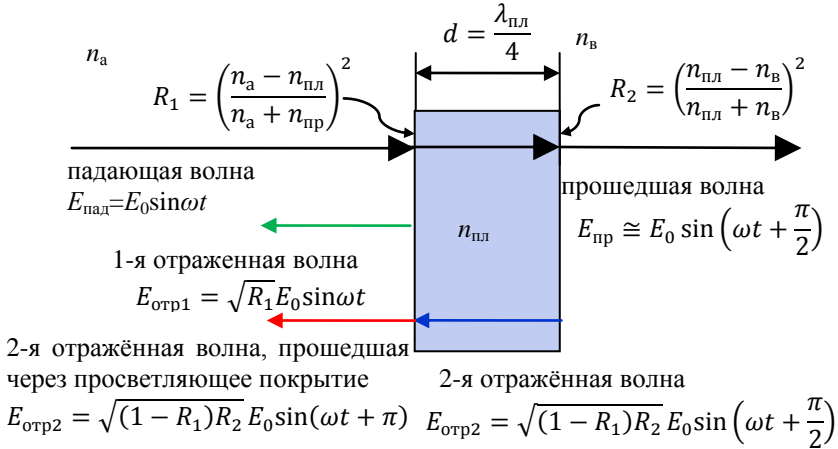


Рис.2.30. Принцип действия просветляющего покрытия

$\lambda_{пл}$ – длина световой волны в плёнке. Падающая перпендикулярно к границе раздела (активная среда – просветляющее покрытие) световая волна частично отражается, а частично проходит до границы раздела с воздухом, на которой образуется вторая отражённая волна. После вторичного прохождения через просветляющий слой фаза этой волны изменяется на π , и в активной среде обе отражённые волны складываются в противофазе.

При равенстве амплитуд первой и второй отражённых волн

$$\sqrt{R_1} E_0 = \sqrt{(1 - R_1) R_2} E_0$$

интерферирующие волны гасят друг друга, а волна, прошедшая в воздух, усиливается. Показатель преломления материала плёнки $n_{пл} = \sqrt{n_a n_b}$ подбирают так, чтобы получить коэффициенты отражения $R_1 = \left(\frac{n_a - n_{пл}}{n_a + n_{пр}} \right)^2$ и $R_2 = \left(\frac{n_{пл} - n_b}{n_{пл} + n_b} \right)^2$, обеспечивающие равенство амплитуд отражённых волн.

Использование просветляющего покрытия позволяет уменьшить коэффициент отражения от сколотой поверхности кристалла GaAs ($n_a = 6,1$) с 37% до 1%.

Основные характеристики передающего оптоэлектронного модуля с суперлюминесцентным диодом DL-CS5103A

Модуль состоит из корпуса с размещёнными в нем суперлюминесцентным диодом, микрохолодильником и терморезистором. Выход оптического излучения происходит через одномодовое оптическое волокно с диаметром сердцевины порядка 10 мкм. Диаметры оболочки, защитного покрытия и наружной изоляции составляют 125, 245 и 900 мкм соответственно. Длина оптического волокна – 1 м, диаметр его скрутки должен быть больше 40 мм. Тип оптического соединителя (коннектора) – FC/PC.

Основные характеристики суперлюминесцентного диода DL-CS5103A приведены в таблице. Ватт-амперная и спектральная характеристики изображены на рис.2.31 и 2.32.

Характеристики суперлюминесцентного диода DL-CS5103A

Параметр	Значение			Единица измерения
	минимальное	типичное	максимальное	
Напряжение	-	2	2,5	В
Сила тока	-	300	350	мА
Оптическая мощность	10	-	-	мВт
Центральная длина волны	1530	1550	1570	нм
Ширина спектра излучения	35	-	-	нм
Температура модуля	-40	-	+70	°С
Температура СЛД	0	-	+70	°С
Напряжение термоэлектрического охладителя	-	2,5	3,0	В
Сила тока термоэлектрического охладителя	-	1,1	1,8	А

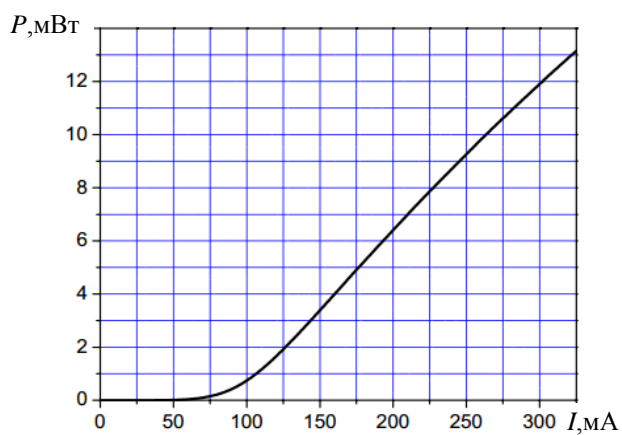
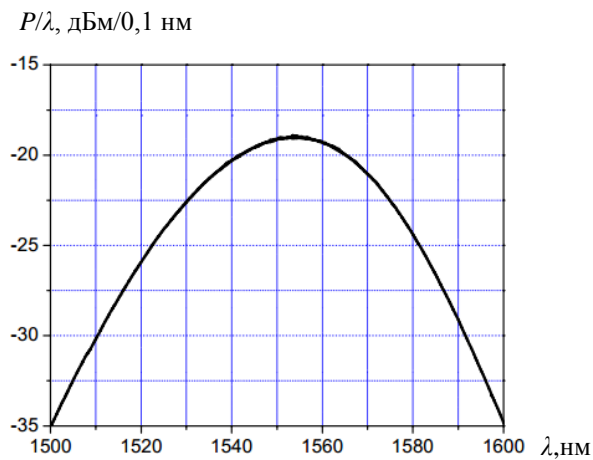
Рис.2.31. Ватт-амперная характеристика СЛД при $t = 25^\circ\text{C}$ 

Рис.2.32. Спектральная плотность мощности излучения СЛД

3. ГЕНЕРАЦИЯ И УСИЛЕНИЕ ИЗЛУЧЕНИЯ В ВОЛОКОННО-ОПТИЧЕСКИХ СИСТЕМАХ

3.1 Оптоволоконные усилители

Волоконно-оптические усилители на основе легированных эрбием оптических волокон (Erbium-Doped Fiber Amplifier – EDFA) применяются для повышения мощности оптического сигнала.

Волоконно-оптические усилители могут выполнять функции:

- усилителя мощности (бустера) для дополнительного усиления сигнала источника излучения на основе лазерного или суперлюминесцентного диода;
- линейного усилителя для восстановления уровня оптического сигнала в волоконно-оптических линиях передачи;
- предварительного усилителя на входе фотоприёмного устройства.

Преимущества эрбиевых усилителей:

- непосредственно усиливают проходящий оптический сигнал без преобразования в электрический сигнал;
- диапазон длин волн усиливаемых сигналов находится в области минимальных оптических потерь в кварцевом волокне (рис.3.1);

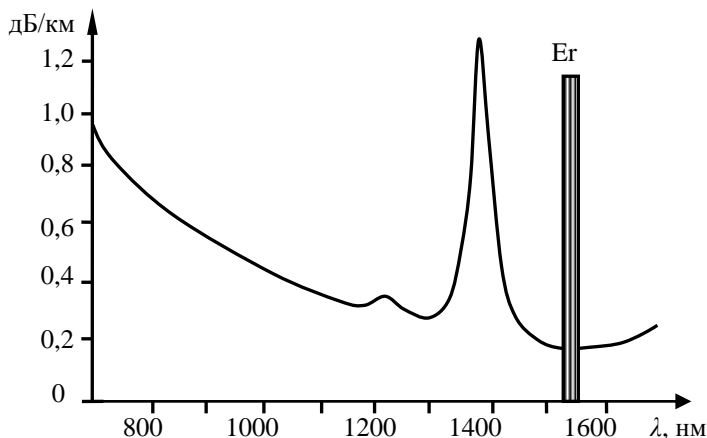


Рис.3.1. Оптические потери кварцевого волокна и диапазон работы эрбиевого усилителя

- обеспечивают возможность одновременного усиления сигналов с различными длинами волн;
- усиление оптических сигналов не зависит от метода их модуляции;
- достигается низкий уровень шумов и простота включения в волоконно-оптическую систему передачи.

Первые успешные эксперименты по усилению света в волоконном световоде, легированном ионами эрбия, были проведены в 1987 г., и уже в 1992 г. на рынке появились готовые для применения модули эрбиевых усилителей. Значительный вклад в их создание внесли научные группы из Университета Саутгемптона (Великобритания) под руководством Давида Пэйна (D. Payne) и из лаборатории Bell (США) под руководством Эммануила Десурвира (E. Desurvire). В нашей стране работы в этом направлении велись под руководством Е. М. Дианова в отделе волоконной оптики Института общей физики АН СССР (в настоящее время Научный центр волоконной оптики при Институте общей физики РАН).

Основная область применения волоконно-оптических усилителей – компенсация затухания сигнала в оптической кабеле волоконно-оптической линии связи. Линии длиной в сотни километров разделяют на участки (в несколько десятков километров), между которыми встраиваются волоконно-оптические усилители. Они, не изменяя формы сигнала, повышают его мощность. В одной волоконно-оптической линии связи может быть встроено несколько усилителей.

3.1.1. Оптическая схема эрбиевого волоконного усилителя

Упрощенные структурные схемы эрбиевых волоконных усилителей показаны на рис.3.2.

Основным элементом (*активной зоной*) волоконно-оптических усилителей является участок оптоволокна из кварцевого стекла ($\text{Si}^{4+}\text{O}_2^{2-}$), легированного ионами эрбия (Er^{3+}). Длина активного волокна составляет обычно $5 \div 8$ м. Эрбиевое волокно, как и обычное одномодовое оптическое волокно, имеет сердцевину диаметром $3 \div 6$ мкм, легированную германием или другим веществом, увеличивающим показатель преломления, и оболочку из чистого кварца диаметром 125 мкм. Центральная

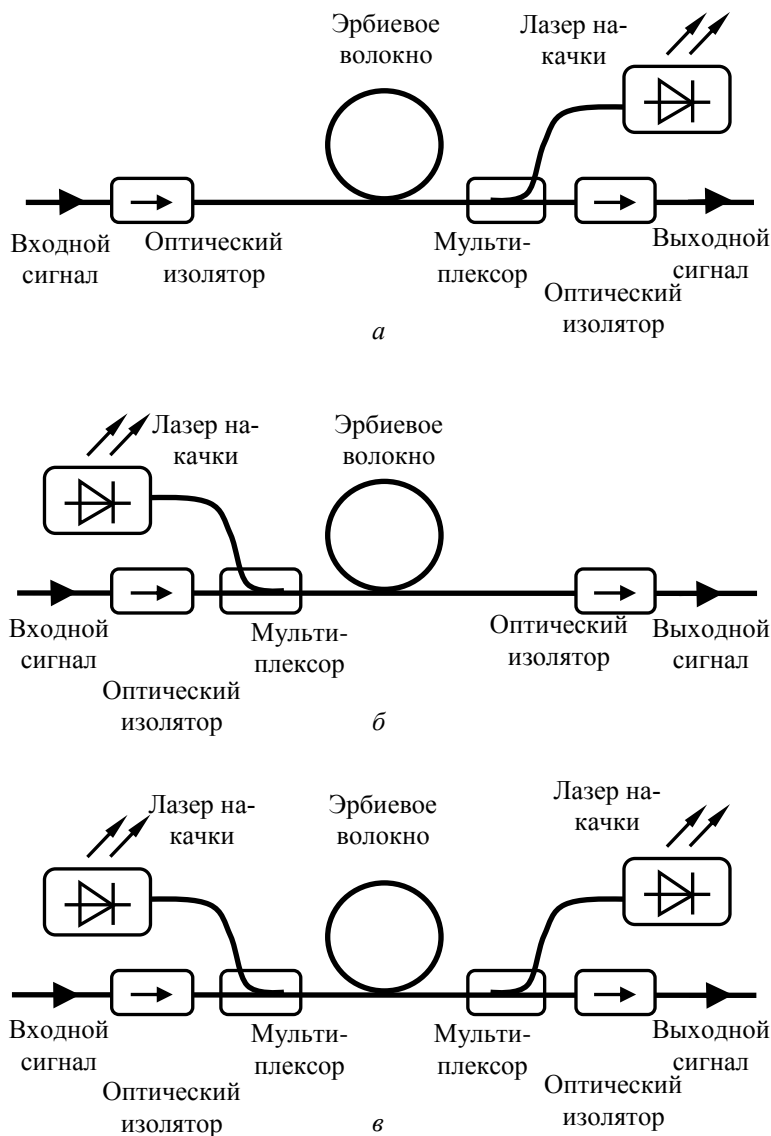


Рис.3.2. Структурные схемы эрбиевого волоконно-оптического усилителя (EDFA): *а* – накачка во встречном направлении; *б* – накачка в попутном направлении; *в* – накачка в двух направлениях

часть сердцевины диаметром примерно 2 мкм дополнительно легирована ионами эрбия. Концентрация ионов эрбия составляет $10^{17} \div 10^{19} \text{ см}^{-3}$ (менее 0,1%). Защитное покрытие волокна диаметром 250 мкм имеет большой коэффициент преломления и эффективно выводит из оболочки постороннее излучение.

В эрбиевом волокне сигналы с длиной волны около 1550 нм могут усиливаться за счёт использования энергии излучения *оптической накачки*, переводящей внешние (оптические) электроны ионов эрбия в возбуждённое состояние. Длина волны оптической накачки 980 или 1480 нм, что соответствует полосам поглощения ионов эрбия. Поток излучения накачки вводится в волокно с помощью *направленного ответвителя-мультиплексора*. В качестве источников накачки используются многомодовые полупроводниковые лазерные диоды с мощностью излучения от 100 до 250 мВт.

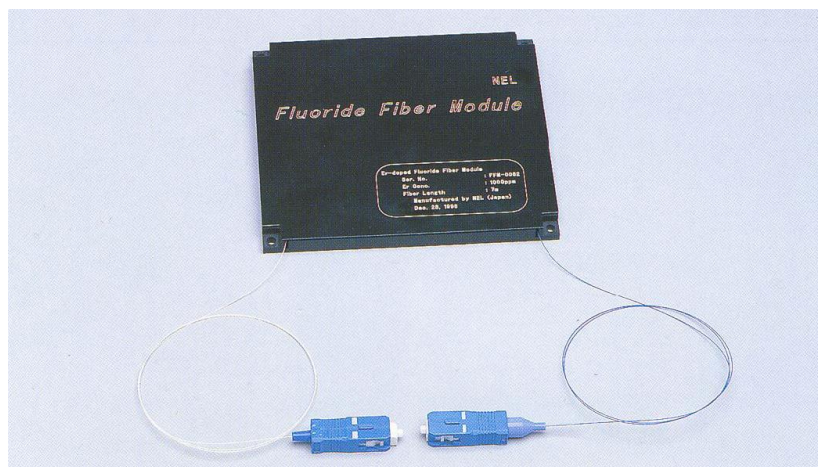
На входе и выходе оптического усилителя устанавливают *оптические изоляторы* – устройства, пропускающие световые сигналы только в одном направлении. Они предотвращают проникновение в усилитель паразитных сигналов, отражённых от неоднородностей линии связи. Отражённые сигналы являются помехами (шумами), ухудшающими работу усилителя.

На рис.3.2,*а* изображена схема усилителя со *встречной накачкой*, при которой оптический сигнал и излучение накачки проходят по волокну в противоположных направлениях. Существуют также схемы с попутной накачкой (рис.3.2,*б*) и накачкой в двух направлениях (рис.3.2,*в*).

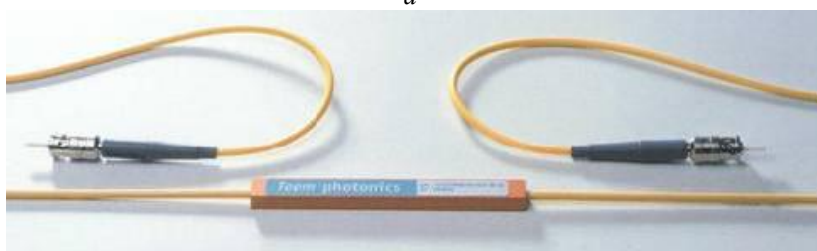
Эрбиевый усилитель может быть изготовлен в планарной волноводной геометрии (рис.3.3). Длина активного планарного волновода составляет $5 \div 8 \text{ см}$, а концентрация ионов эрбия в волноводе достигает $10^{20} \div 10^{21} \text{ ионов /см}^3$.

3.1.2. Принцип работы эрбиевого усилителя

Принцип работы усилителей EDFA основан на явлении усиления света при вынужденном (стимулированном) излучении ионов эрбия, находящихся в возбуждённом состоянии (Light Amplification by Stimulated Emission of Radiation, LASER). Это то же самое явление, которое обеспечивает возникновение генерации в лазерах.



а



б

Рис.3.3. Эрбиевые оптические усилители: а – волоконный, б – планарный

На рис.3.4,а приведена диаграмма энергетических уровней иона Er^{3+} , где W_1 – основное состояние иона, W_2 – метастабильное состояние и W_3 – возбуждённое состояние с малым временем жизни. В результате взаимодействия ионов эрбия (Er^{3+}) с электрическим полем соседних с ним ионов в кварцевом стекле каждый из уровней расщеплён на ряд подуровней, образующих энергетические зоны (эффект Штарка).

Число ионов в единице объёма, находящихся на некотором уровне энергии, называется заселённостью этого уровня. В отсутствие накачки заселённость основного уровня энергии W_1 максимальна, заселённость всех остальных уровней энергии пренебрежимо мала, поскольку они отстоят от основного уровня

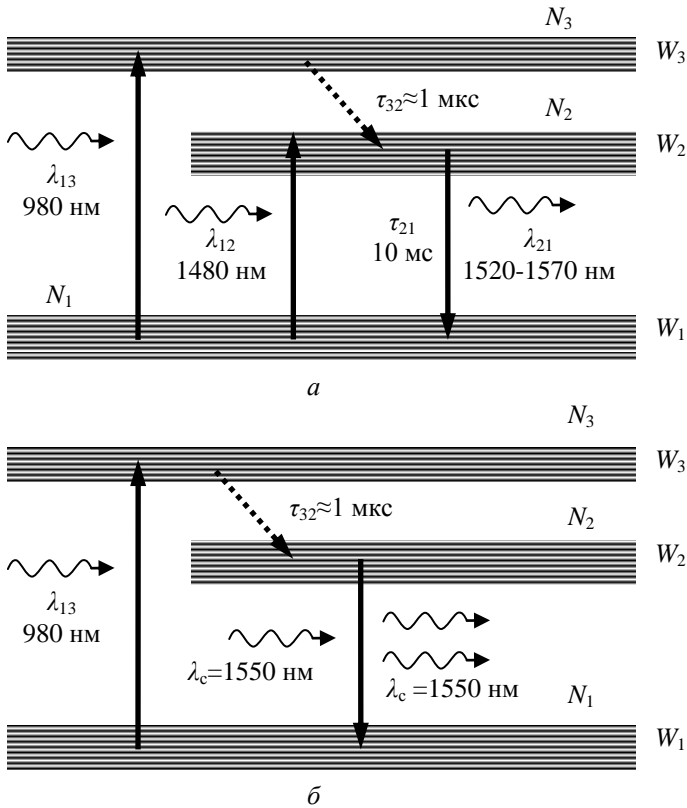


Рис.3.4. Упрощенная схема уровней энергии ионов эрбия (Er^{3+}) в кварцевом стекле

на расстояниях, значительно больших, чем средняя энергия кванта теплового движения kT :

$$W_3 - W_1 > W_2 - W_1 \gg kT.$$

При интенсивном облучении волокна мощным лазером накачки с длиной волны $\lambda_{13} = 980$ нм большое количество ионов эрбия переходит из основного состояния (с уровня W_1) в возбуждённое состояние (на уровень W_3):

$$W_3 - W_1 = \frac{hc}{\lambda_{13}}.$$

Состояние W_3 является короткоживущим (время жизни $\tau_3 \approx 1$ мкс).

С уровня W_3 ионы переходят на уровень W_2 с испусканием *фоонов* (квантов энергии тепловых колебаний атомов). Состояние W_2 является *метастабильным* с большим временем жизни: $\tau_2 \approx 10$ мс. Поэтому при соответствующей мощности накачки число ионов N_2 , находящихся на уровне W_2 , может превышать число ионов N_1 на уровне W_1 . Таким образом, в системе ионов эрбия происходит накопление энергии за счёт поглощения квантов излучения накачки.

Состояние среды, при котором заселённость более высокого уровня энергии превышает заселённость нижележащего уровня, получило название *состояния с инверсией заселённости*. В ряде схем усилителей состояние с инверсией заселённости уровней достигается при облучении волокна светом лазера с длиной волны 1480 нм (рис.3.4,а).

Пусть по волокну с инверсией заселённости уровней проходит полезный (усиливаемый) сигнал с энергией фотонов, совпадающей с энергией перехода из метастабильного состояния в основное

$$hc/\lambda_c = W_2 - W_1, \lambda_c \approx 1550 \text{ нм}.$$

Фотоны входного сигнала взаимодействуют с ионами эрбия и вынуждают их совершать переходы с метастабильного уровня W_2 на основной уровень W_1 с испусканием фотонов, когерентных с фотонами входного сигнала. Таким образом, в каждом элементарном акте вынужденного излучения возникают два фотона с идентичными свойствами: частотой, фазой, поляризацией и направлением распространения (рис.3.4,б). Так как все вновь рождённые фотоны когерентны, то их совокупность представляет собой электромагнитную волну, отличающуюся от электромагнитной волны входного сигнала только большей интенсивностью. При этом ионы эрбия, отдав запасённую энергию в ходе вынужденного излучения, возвращаются в основное, невозбуждённое, состояние.

Однако не все ионы эрбия находятся в метастабильном состоянии W_2 и обеспечивают усиление сигнала. Часть ионов находится на уровне W_1 . Эти ионы, взаимодействуя с фотонами, энергия которых совпадает с энергией перехода ($\lambda_c \approx 1550$ нм),

эффективно их поглощают, переходя на уровень W_2 . При этом спектр поглощения практически совпадает со спектром усиления.

Для изменения мощности сигнала dP_c при прохождении сигналом участка волокна длиной dx можно написать:

$$dP_c = [N_2\sigma_{21}(\lambda) - N_1\sigma_{12}(\lambda)]P_c dx, \quad (3.1)$$

где P_c – мощность излучения сигнала, $\sigma_{12}(\lambda)$ и $\sigma_{21}(\lambda)$ – сечения поглощения и вынужденного излучения на длине волны сигнала, N_1 и N_2 – показатели заселённости основного и метастабильного уровней.

Для двухуровневой схемы (W_1 и W_2) справедливо соотношение, связывающее сечение вынужденного излучения с сечением поглощения света:

$$g_2\sigma_{21}(\lambda) = g_1\sigma_{12}(\lambda), \quad (3.2)$$

где g_1 и g_2 – показатели кратности вырождения основного и возбуждённого состояний. Для иона Er^{3+} отношение $g_1/g_2 \approx 1$.

Если число ионов, находящихся на уровне W_2 , меньше числа ионов, находящихся на основном уровне W_1 ($N_2 < N_1$), то будет наблюдаться поглощение светового сигнала:

$$dP_s < 0.$$

Именно поэтому необходимым условием усиления света является создание инверсии заселённости ($N_2 > N_1$) между двумя рабочими уровнями энергии W_2 и W_1 . Для создания инверсии заселённости нужно перевести больше половины ионов эрбия на метастабильный уровень W_2 .

Мощность накачки $P_{\text{нак}}$ эрбиевого усилителя, при которой показатели заселённости уровней W_1 и W_2 равны, называется *пороговой мощностью*. Поглощение прекращается, и сигнал проходит по волокну, не изменяясь по величине ($dP_c = 0$).

При мощности накачки больше пороговой первое слагаемое в (3.1) превышает второе, и всё выражение становится положительным ($dP_s > 0$). Сигнал при прохождении волокна усиливается. Проинтегрировав (3.1), получим мощность сигнала после прохождения отрезка эрбиевого волокна длиной x :

$$P_c(x) = P_c(0)\exp\{[N_2\sigma_{21}(\lambda) - N_1\sigma_{12}(\lambda)]x\}, \quad (3.3)$$

здесь $P_c(0)$ – мощность сигнала на входе усилителя. Отношение мощностей сигнала на выходе и входе усилителя определяет его коэффициент усиления:

$$K = P_c(x)/P_c(0). \quad (3.4)$$

Зависимость коэффициента усиления усилителя от мощности источника накачки $P_{\text{нак}}$ показана на рис.3.5.

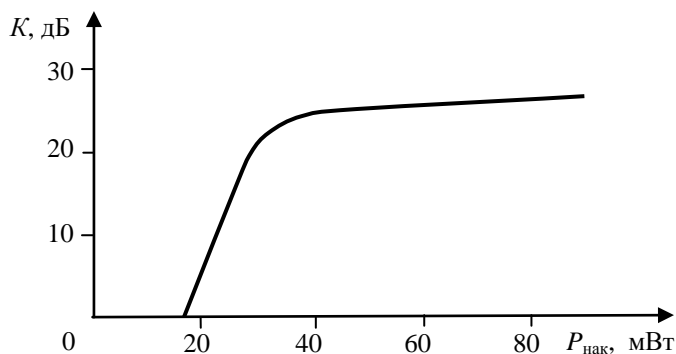


Рис.3.5. Зависимость коэффициента усиления усилителя от мощности источника накачки ($P_{\text{нак}}$)

Выражение (3.3) получено в предположении, что заселённости уровней одинаковы по всей длине волокна. Однако мощность излучения накачки уменьшается вдоль волокна и, как следствие, снижается заселённость метастабильного состояния N_2 .

Сечение поглощения σ_{12} и связанное с ним (3.2) сечение вынужденного излучения σ_{21} можно найти экспериментально, если измерить коэффициент поглощения света в отсутствие накачки. В этом случае заселённость основного уровня практически равна полной концентрации ионов эрбия:

$$N_1 \approx N, \text{ а } N_2 \approx 0.$$

Проинтегрировав (3.1), получим

$$P_c(x) = P_c(0)\exp[-N\sigma_{12}(\lambda)x] = P_c(0)\exp[-ax], \quad (3.5)$$

где $a = N\sigma_{12}$ – коэффициент поглощения, измерение которого позволяет определить сечение поглощения σ_{12} .

Часть возбуждённых ионов эрбия переходит с метастабильного уровня в основное состояние спонтанно (рис.3.4,а), излучая фотоны с длиной волны около 1550 нм:

$$\hbar c/\lambda = W_2 - W_1.$$

При этом энергия фотонов, начальные фазы волновых пакетов, поляризация и направления их распространения случайны. Излучение носит шумовой (некогерентный) характер. Усиленное спонтанное излучение является основным источником шумов усилителя.

Среднее время жизни метастабильного состояния W_2 (время релаксации для спонтанного излучения) составляет около 10 мс. Средним временем жизни называют время, в течение которого первоначальная заселённость уровня уменьшается в $e \approx 2,72$ раз.

3.1.3. Основные характеристики эрбиевого усилителя

Диапазон длин волн усиливаемых оптических сигналов $1,53 \div 1,58$ мкм. Ширина спектра усиливаемых сигналов равна ширине спектра излучения ионов эрбия и определяется шириной энергетических зон ионов (рис.3.4). Наиболее широкий спектр излучения, до 50 нм, имеют ионы эрбия в кварцевом стекле, легированном алюминием (рис.3.6).

Коэффициент усиления определяется как отношение мощности сигнала на выходе усилителя $P_{с, \text{ВЫХ}}$ к мощности входящего сигнала $P_{с, \text{ВХ}}$ (измеряется в децибелах):

$$K = 10 \lg \left(\frac{P_{с, \text{ВЫХ}}}{P_{с, \text{ВХ}}} \right), \text{ дБ.} \quad (3.6)$$

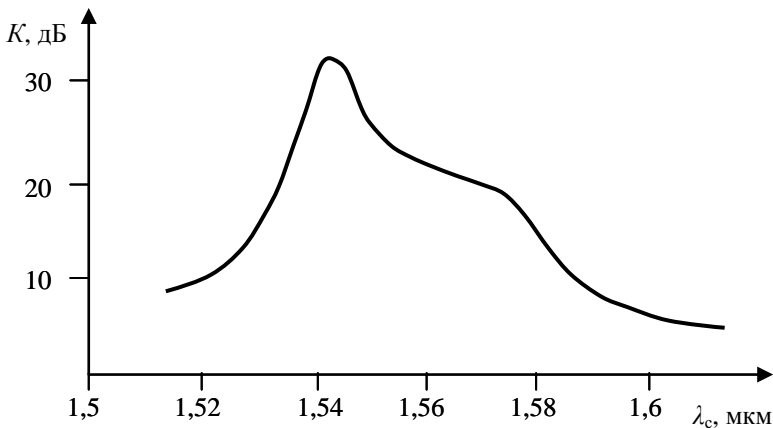


Рис.3.6. Спектр усиления эрбиевого усилителя

Если уровни мощности входного и выходного сигналов выражены в логарифмических единицах (дБм), то коэффициент усиления равен разности уровней мощности выходного и входного сигналов:

$$K [\text{дБ}] = P_{\text{с.вых}} [\text{дБм}] - P_{\text{с.вх}} [\text{дБм}]. \quad (3.7)$$

Величина, показывающая, на сколько децибел мощность сигнала P больше (или меньше) 1 мВт, называется уровнем мощности:

$$P = 10 \lg \frac{P, \text{ мВт}}{1 \text{ мВт}}, \text{ дБм.}$$

Например, запись $P = 30$ дБм означает, что уровень мощности составляет +30 дБ относительно 1 мВт, т. е.

$$P = 10^{30/10} \cdot 1 \text{ мВт} = 1 \text{ Вт.}$$

Типичное значение коэффициента усиления эрбиевого усилителя составляет 20÷30 дБ для слабых сигналов. Для слабых сигналов коэффициент усиления приблизительно постоянен и мощность на выходе усилителя изменяется пропорционально входной (рис.3.7, область А).

Высокий уровень входного сигнала вызывает интенсивное вынужденное излучение, при котором существенно уменьшается заселённость метастабильного уровня W_2 и тем самым понижается коэффициент усиления (рис.3.7, область В).

Мощность насыщения $P_{\text{вых,макс}}$ определяет максимальную мощность сигнала, которая может быть получена на выходе усилителя (рис.3.7, область В). Этот параметр зависит от схемы построения ВОУ (рис.3.2) и мощности лазера накачки. У мощных эрбиевых усилителей $P_{\text{вых,макс}}$ может превосходить 36 дБм (4 Вт).

Неравномерность коэффициента усиления определяется как разность между максимальным и минимальным значениями коэффициента усиления в пределах рабочего спектрального диапазона. Как видно из рис.3.6, изменение коэффициента усиления слабого сигнала в пределах полосы 1,53÷1,58 мкм может превышать 10 дБ. Обычно усилители мощности работают в режиме насыщения по отношению к сигналу на выходе. При этом выравниваются уровни сигналов с различными длинами волн.

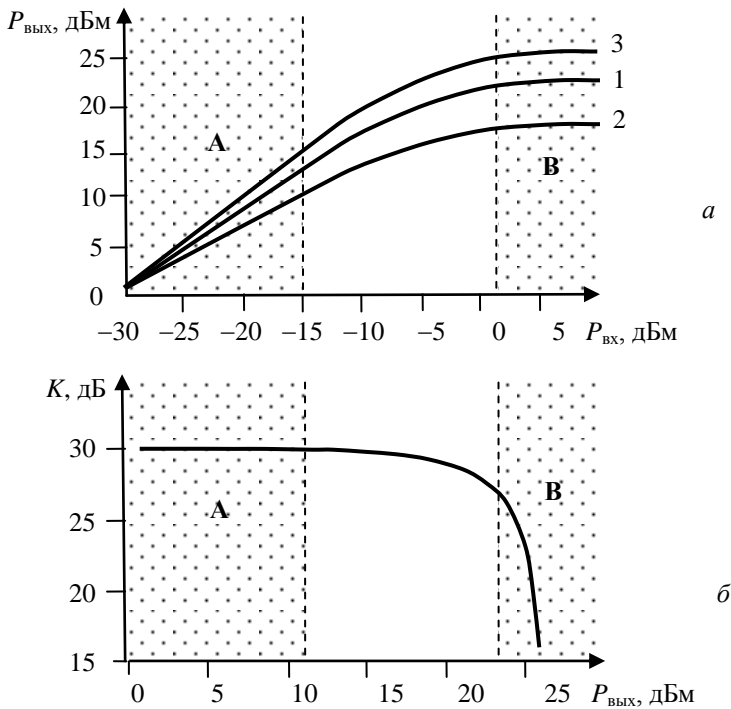


Рис.3.7. Зависимости: *a* – уровней выходной оптической мощности от входной (усилитель с обратной накачкой – кривая 1, с прямой накачкой – 2, с двунаправленной накачкой – 3, при одинаковой мощности источников накачки); *б* – коэффициента усиления от выходной мощности (А – область работы линейного усилителя; В – область работы усилителя мощности)

В случае усиления слабых сигналов необходимо провести *сглаживание спектральной характеристики*. Для этого в оптическую цепь усилителя вводят спектрально-селективные поглощающие фильтры. Основным видом фильтра является фотоиндуцированная длиннопериодная брегговская решётка, полученная путём пространственно-периодического облучения сердцевины оптоволокна ультрафиолетовым излучением через его поверхность. Спектр и интенсивность поглощения задаются периодом решётки и временем облучения. Применение сглаживающих фильтров, изготовленных по этой технологии, позволя-

ет уменьшить неравномерность коэффициента усиления до десятых долей дБ в пределах рабочего диапазона.

Энергетическая эффективность определяется отношением приращения мощности сигнала к величине мощности накачки:

$$P_{\text{ef}} = \frac{P_{\text{с.вых}} - P_{\text{с.вх}}}{P_{\text{н}}}. \quad (3.8.)$$

Отношение числа фотонов сигнала $N_{\text{с}}$, появившихся в процессе усиления сигнала, к числу поглощённых фотонов накачки $N_{\text{н}}$ называется *квантовой эффективностью* накачки:

$$Q_{\text{ef}} = N_{\text{с}} / N_{\text{н}}. \quad (3.9)$$

Накачка на длине волны 980 нм позволяет получить энергетическую эффективность 55% при квантовой эффективности 86%.

Шум усилителя – некогерентное излучение, возникающее в эрбиевом волокне. Ионы эрбия могут самопроизвольно (спонтанно) переходить из метастабильного состояния в основное с испусканием фотона:

$$hc/\lambda = W_2 - W_1.$$

Спонтанное излучение усиливается так же, как полезный сигнал, распространяющийся по усилителю. Поэтому, даже если нет входного сигнала, на выходе усилителя всегда имеется усиленное спонтанное излучение (amplified spontaneous emission – ASE). Ширина спектра шума на выходе усилителя (рис.3.8) приблизительно такая же, как и ширина полосы усиливаемых сигналов (рис.3.6).

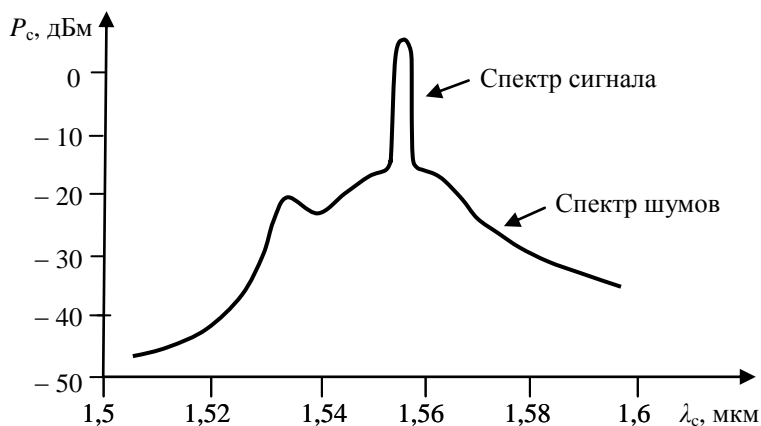


Рис.3.8. Спектры полезного сигнала и шумов на выходе усилителя

Оценим величину мощности шумов на выходе усилителя. Мощность спонтанного излучения ионов эрбия в метастабильном состоянии пропорциональна средней энергии кванта и полюсе частот излучения:

$$P_{\text{сп}} \sim h\nu\Delta\nu,$$

где $h = 6,6 \cdot 10^{-34}$ Вт·с² – постоянная Планка, $\nu \approx 2 \cdot 10^{14}$ Гц – средняя частота и $\Delta\nu \approx 0,04 \cdot 10^{14}$ Гц – ширина спектральной полосы усиливаемых сигналов, соответствующая диапазону длин волн 1530÷1560 нм.

Мощность усиленного спонтанного излучения на выходе усилителя

$$P_{\text{ш вых}} \approx P_{\text{сп}} \cdot K,$$

где $K \approx 10^3$ (30 дБ) – коэффициент усиления усилителя.

$$P_{\text{ш вых}} \approx h\nu \cdot \Delta\nu \cdot K \approx 0,5 \cdot 10^{-6} \text{ Вт} = 0,5 \cdot 10^{-3} \text{ мВт},$$

$$P_{\text{ш вых}} \approx 10\lg(0,5 \cdot 10^{-3}) = -33 \text{ дБм}.$$

Шум-фактор (коэффициент шума) NF (noise figure) характеризует ухудшение отношения сигнал/шум после прохождения сигнала через оптический усилитель и определяется как отношение сигнал/шум на входе ($P_{\text{с вх}}/P_{\text{ш вх}}$) к отношению сигнал/шум на выходе ($P_{\text{с вых}}/P_{\text{ш вых}}$):

$$NF = \frac{P_{\text{с вх}}/P_{\text{ш вх}}}{P_{\text{с вых}}/P_{\text{ш вых}}}. \quad (3.10)$$

Мощность шума на выходе состоит из суммы мощности усиленного спонтанного излучения и усиленной мощности шума во входном сигнале:

$$P_{\text{ш вых}} = P_{\text{сп}}K + P_{\text{ш вх}}K.$$

$$NF = \frac{P_{\text{с вх}}/P_{\text{ш вх}}}{P_{\text{с вх}}K/(P_{\text{сп}}K + P_{\text{ш вх}}K)} = 1 + \frac{P_{\text{сп}}}{P_{\text{ш вх}}}.$$

Часто значение шум-фактора указывается в дБ:

$$NF = 10\lg\left(1 + \frac{P_{\text{сп}}}{P_{\text{ш вх}}}\right), \text{ дБ}. \quad (3.11)$$

Шум-фактор идеального усилителя, в котором отсутствует спонтанное излучение, равен 1 (0 дБ). На практике типичные значения шум-фактора эрбиевых усилителей составляют 5÷6 дБ.

3.2. Оптоволоконные лазеры

3.2.1. Принцип работы лазера с оптической накачкой

Лазер с оптической накачкой – это устройство, преобразующее энергию некогерентного излучения внешнего источника в энергию когерентного, монохроматического, поляризованного и узконаправленного светового луча. В отечественной литературе синонимом термина “лазер” является *оптический квантовый генератор* (ОКГ).

Схема лазера с оптической накачкой показана на рис.3.9. В основе его работы лежит явление индуцированного (вынужден-

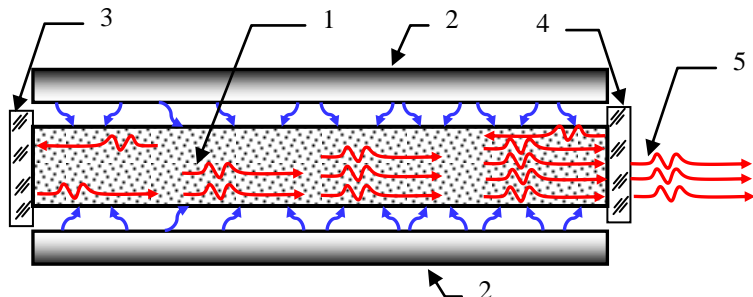


Рис. 3.9. Схема лазера с оптической накачкой : 1 – активная среда; 2 – система оптической накачки; 3 – отражающее зеркало; 4 – полупрозрачное зеркало; 5 – когерентное излучение

ного) излучения. Основными элементами лазера являются:

1. *Активная среда* – рабочее вещество, в котором имеется неравновесная инверсная заселённость энергетических уровней.

2. *Система накачки* – устройство, с помощью которого энергия внешнего источника излучения преобразуется во внутреннюю энергию активной среды, создавая инверсию заселённости её энергетических уровней.

3. *Оптический резонатор* – элемент, определяющий длину волны генерируемого излучения. Он состоит из двух зеркал – отражающего зеркала и полупрозрачного зеркала. В резонаторе происходит интенсивное взаимодействие генерируемого излучения с рабочим веществом, в результате которого внутренняя энергия активной среды преобразуется в энергию когерентного светового излучения.

4. Устройство, обеспечивающее вывод энергии в виде узконаправленного светового луча. Устройство состоит из полупрозрачного зеркала и системы оптических линз, формирующих луч нужного сечения и угла расходимости.

5. Дополнительные элементы – приборы управления лучом, модуляторы и т.д.

Взаимодействие излучения с атомами активной среды лазера

Основное состояние атома – состояние с минимальной энергией. В возбуждённом состоянии атома один из его внешних электронов находится на более высоком разрешенном уровне энергии. Возбуждённый атом может самопроизвольно (спонтанно) перейти в более низкое энергетическое состояние (рис.3.10,*a*), испустив при этом квант света (фотон) с энергией:

$$h\nu = W_2 - W_1 .$$

Скорость спонтанного перехода с уровня W_2 на уровень W_1 :

$$\frac{dN_2}{dt} = A_{21}N_2 , \quad (3.12)$$

где N_2 – концентрация возбуждённых атомов (заселённость уровня W_2). Коэффициент Эйнштейна для спонтанного излучения A_{21} определяет вероятность перехода и имеет размерность $[с^{-1}]$. Обратная величина называется временем жизни атома на уровне W_2

$$\tau_2 = \frac{1}{A_{21}} . \quad (3.13)$$

Значение коэффициента Эйнштейна для спонтанных переходов в оптическом диапазоне может изменяться от $10^8 с^{-1}$ для *разрешённых переходов* до $1 с^{-1}$ для *запрещённых переходов*. Переход называют разрешённым, если он удовлетворяет правилам отбора, накладывающим ограничения на изменения квантовых чисел атома:

- изменение квантового числа суммарного орбитального момента электронов в оболочке $\Delta L = \pm 1$;
- изменение квантового числа суммарного спинового момента электронов в оболочке $\Delta S = 0$;
- изменение квантового числа полного момента импульса электронов в оболочке $\Delta J = 0, \pm 1$.

Переходы, не удовлетворяющие правилам отбора (запрещённые переходы), также происходят, но со значительно меньшей вероятностью.

Спонтанные переходы (рис.3.10,*а*) случайны и не зависят от внешнего электромагнитного поля. Излучённые фотоны имеют случайную частоту (в пределах ширины спектральной линии), фазу, поляризацию и направление распространения. Спонтанное излучение некогерентное. Такое излучение создают естественные источники света.

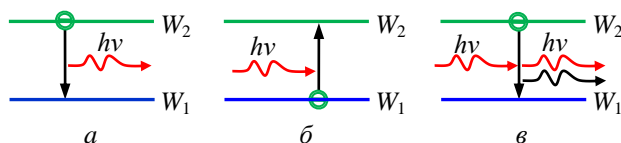


Рис.3.10. Три вида взаимодействия квантов света (фотонов) с атомами, ионами или молекулами вещества: *а* – спонтанное излучение, *б* – поглощение и *в* – индуцированное излучение

Другим оптическим процессом является поглощение кванта света атомом с переходом его в возбуждённое состояние (рис.3.10,*б*). Скорость поглощения

$$\frac{dN_1}{dt} = -B_{12}\rho_\nu N_1, \quad (3.14)$$

где B_{12} – коэффициент Эйнштейна для вынужденного поглощения; N_1 – концентрация атомов, находящихся в основном состоянии, или заселённость нижнего энергетического уровня; ρ_ν – плотность энергии электромагнитного поля с частотой $\nu = (W_2 - W_1)/h$ в единичном спектральном интервале [Дж/(м³·Гц)].

При взаимодействии возбуждённого атома с фотоном возможен переход атома в состояние с меньшей энергией и с излучением дополнительного фотона (рис.3.10,*в*). Скорость вынужденного или индуцированного излучения

$$\frac{dN_2}{dt} = -B_{21}\rho_\nu N_2, \quad (3.15)$$

где B_{21} – коэффициент Эйнштейна для вынужденного излучения; N_2 – концентрация атомов в возбуждённом состоянии или

заселённость верхнего энергетического уровня; ρ_ν – плотность энергии электромагнитного поля с частотой $\nu = (W_2 - W_1)/h$.

В отличие от спонтанного излучения вероятность квантовых переходов с поглощением или вынужденным излучением пропорциональна плотности энергии электромагнитного поля ρ_ν с частотой перехода $\nu = (W_2 - W_1)/h$.

Коэффициенты Эйнштейна для поглощения и вынужденного излучения связаны соотношениями

$$g_2 B_{21} = g_1 B_{12}, \quad (3.16)$$

где g_1 и g_2 – кратности вырождения уровней атома с энергиями W_2 и W_1 . Для квантовых переходов вероятности поглощения и вынужденного излучения равны $B_{12}=B_{21}$ или отличаются на постоянный множитель.

Индукцированное излучение является процессом, составляющим физическую основу работы ОКГ. Основное отличие индуцированного излучения от спонтанного – вторичный фотон не отличим от первичного, индуцирующего фотона. Оба фотона имеют одинаковую частоту, фазу и поляризацию.

Усиление и поглощение света в активной среде

Рассмотрим среду, в которой одновременно имеются атомы в основном W_1 и возбуждённом W_2 состояниях. Заселённости состояний в единице объёма – N_1 и N_2 .

Пусть через среду проходит электромагнитная волна с частотой $\nu = (W_2 - W_1)/h$ и плотностью энергии ρ_ν . При этом единицей объёма среды поглощается мощность

$$P_{\text{погл}} = h\nu B_{12} \rho_\nu N_1. \quad (3.17)$$

Одновременно в процессе вынужденного излучения выделяется мощность

$$P_{\text{изл}} = h\nu B_{21} \rho_\nu N_2. \quad (3.18)$$

Разность этих мощностей представляет собой изменение плотности энергии поля в единицу времени (в предположении, что $B_{12}=B_{21}$):

$$\frac{d\rho_\nu}{dt} = P_{\text{изл}} - P_{\text{погл}} = h\nu B_{21} \rho_\nu (N_2 - N_1). \quad (3.19)$$

При термодинамическом равновесии в среде

$$\frac{N_2}{N_1} = \exp\left(-\frac{W_2 - W_1}{kT}\right) \quad (3.20)$$

заселённость верхнего уровня пренебрежимо мала $N_2 \ll N_1$, поскольку расстояние между уровнями в оптическом диапазоне много больше kT . Поэтому согласно (3.19) будет происходить поглощение энергии внешнего поля $d\rho_v/dt < 0$, которая в результате релаксационных процессов переходит в тепловую.

Для излучения, распространяющегося в направлении z со скоростью c , коэффициент поглощения определяется как

$$\alpha = -\frac{1}{I} \frac{dI}{dz}, \quad (3.21)$$

где I – интенсивность (плотность мощности) излучения [Вт/м²]. Так как $I \sim \rho$ и $dz = cdt$, то

$$\alpha = -\frac{1}{c\rho} \frac{d\rho}{dt},$$

$$\alpha = -\frac{h\nu}{c} B_{21}(N_2 - N_1) > 0. \quad (3.22)$$

Если же в квантовой системе каким-либо образом создать инверсную заселённость уровней $N_2 > N_1$, то это будет система с отрицательным поглощением $\alpha < 0$. При распространении излучения в среде с отрицательным поглощением его интенсивность нарастает

$$I(z) = I_0 \exp\left[\frac{h\nu}{c} B_{21}(N_2 - N_1)z\right]. \quad (3.23)$$

Таким образом, при наличии в среде инверсной заселённости уровней $N_1 > N_2$ можно получить квантовый усилитель света.

Совокупность квантовых частиц с инверсией заселённости энергетических уровней называют *активной средой*. Иногда такую инвертированную систему называют системой с отрицательной температурой (см. формулу (3.20): если $N_2 > N_1$, то $kT < 0$). Коэффициент $|\alpha|$ называют коэффициентом усиления активной среды.

В реальной активной среде происходит уменьшение интенсивности излучения при рассеянии на неоднородностях, примесях, поглощении, которые переводят энергию излучения в тепло. Эти потери учитываются коэффициентом β , тогда

$$I = I_0 \exp[(|\alpha| - \beta)x]. \quad (3.24)$$

Для усиления необходимо выполнить условие $|\alpha_{\min}| > \beta$, следовательно, усиление интенсивности световой волны будет происходить, если разность заселённостей будет больше некоторого порогового значения.

Кроме того, при выводе формул не учитывалось, что энергетические уровни и спектральная линия имеют конечные ширины:

$$\Delta\nu \geq \frac{1}{2\pi\tau},$$

где τ – время жизни атома в возбуждённом состоянии, определяется скоростями спонтанного излучения и безызлучательных релаксационных переходов. Учёт этих явлений приводит к уменьшению рассчитанного значения коэффициента усиления.

Создание инверсии заселённостей энергетических уровней при оптической накачке

Для создания инверсии заселённостей необходимо дополнительное внешнее воздействие. Это воздействие должно преодолеть релаксационные процессы, направленные на восстановление равновесной разности заселённостей. Противостоять процессам восстановления равновесной заселённости можно, только затрачивая энергию – энергию накачки, поступающую от внешнего источника.

При *оптической накачке* активная среда облучается светом (инфракрасным, видимым или ультрафиолетовым), который при поглощении переводит активные атомы или молекулы рабочего вещества из основного в возбуждённое состояние. В качестве источника света используются: ксеноновые лампы-вспышки, криптоновые и ртутные лампы высокого давления, светодиоды и лазеры.

Двухуровневая система

Рассмотрим двухуровневую систему, в которой атомы активной среды в основном состоянии находятся на уровне W_1 и в возбуждённом состоянии на уровне W_2 . Среда подвергается действию излучения накачки с плотностью ρ_n . В среде проходят три основных процесса: возбуждение с переходом на уровень W_2 в

результате накачки, а также и спонтанный, и вынужденный излучательные переходы в основное состояние (рис.3.11,а).

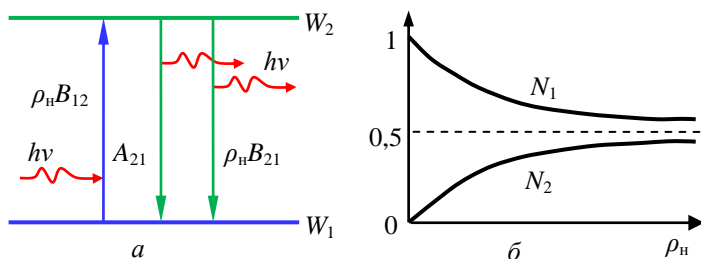


Рис. 3.11. Двухуровневая квантовая система: а – схема энергетических состояний; б – зависимость относительной заселённости уровней от плотности излучения накачки

В установившемся состоянии скорости “заселения” и “опустошения” уровня W_2 равны:

$$\rho_n B_{12} N_1 = (\rho_n B_{21} + A_{21}) N_2,$$

отсюда

$$\frac{N_2}{N_1} = \frac{\rho_n B_{12}}{\rho_n B_{21} + A_{21}}. \quad (3.25)$$

При отсутствии возбуждения $\rho_n = 0$ все атомы находятся на нижнем уровне W_1 . С увеличением плотности энергии накачки заселённость уровня W_1 уменьшается, а уровня W_2 растёт (рис.3.11,б). В пределе при бесконечно большой плотности энергии накачки заселённости обоих уровней станут равными ($N_2/N_1 = 1$). В двухуровневой системе нельзя достичь инверсии при использовании оптической накачки.

Трёхуровневая система активной среды лазера

Энергетическая диаграмма трёхуровневой активной среды лазера показана на рис.3.12. Активная среда в основном состоянии находится на уровне W_1 . В оптическом диапазоне

$$(W_3 - W_1) > (W_2 - W_1) \gg kT$$

и заселённостью вышележащих уровней W_2 и W_3 можно пренебречь. Пусть излучение накачки $h\nu_n = W_3 - W_1$ индуцирует переходы между уровнями W_1 и W_3 с вероятностью

$$w_{13} = w_{31} = \rho_n B_{13} = \rho_n B_{31}. \quad (3.26)$$

Состояние W_3 является короткоживущим. С уровня W_3 происходят безызлучательные релаксационные переходы на уровень W_2 с вероятностью w_{32} . Уровень W_2 должен быть метастабильным, т.е. вероятность релаксационного перехода с этого уровня w_{21} должна быть малой $w_{21} \ll w_{32}$ и $w_{21} \ll w_{13}$. Выполнение этих неравенств приводит к тому, что метастабильный уровень W_2 заселяется быстрее, чем опустошается.

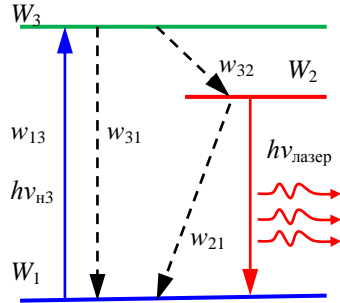


Рис.3.12. Энергетическая диаграмма трёхуровневой системы

В этих условиях разность заселённостей уровней составляет

$$N_2 - N_1 = N \frac{w_{13} - w_{21}}{w_{13} + w_{21}}, \quad (3.27)$$

$N_2 > N_1$, т.е. устанавливается инверсная заселённость энергетических уровней W_2 и W_1 .

Использование трёхуровневых систем в качестве активных сред для квантовых усилителей и генераторов было предложено в 1955 г. Н.Г.Басовым, А.М.Прохоровым и Ч.Таунсом (Нобелевская премия по физике 1964 г.). Первым твёрдотельным лазером был рубиновый лазер, реализованный в 1960 г. Т.Мейсманом. Его активным веществом является рубин — кристалл α -модификации корунда (Al_2O_3) с примесью ионов хрома Cr^{3+} в количестве около 0,05%, придающих ему бледно-розовую окраску.

Ионы Cr^{3+} в кристалле рубина находятся в сильном внутрикристаллическом электрическом поле, создаваемым соседними ионами O^{2-} и Al^{3+} . Вследствие эффекта Штарка вырождение верхнего уровня W_3 частично снимается, и этот уровень расщепляется на две группы подуровней, дающих полосы поглощения с длинами волн 410 нм и 550 нм (рис.3.13).

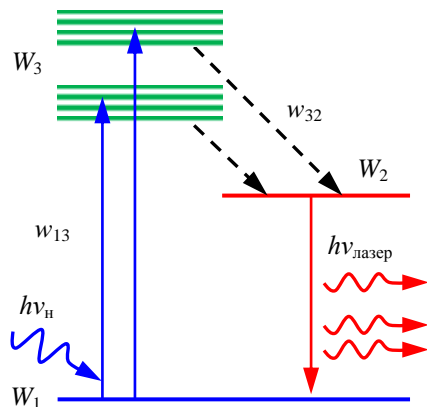


Рис. 3.13. Уровни энергии иона хрома в рубине

При поглощении квантов излучения накачки с длинами волн $\lambda_n = 560$ нм или $\lambda_n = 440$ нм происходит переход ионов Cr^{3+} в возбуждённое состояние. На уровне W_3 ионы не накапливаются, так как они сразу же переходят на метастабильный уровень W_2 . Время безызлучательной передачи энергии из полос поглощения на уровень W_2 составляет около 100 нс.

В трёхуровневой системе для достижения инверсии необходимо перевести из основного состояния W_1 через уровень резонансного поглощения W_3 на метастабильный уровень W_2 не менее половины всех ионов.

Усиление луча света при однократном прохождении через активную среду получается небольшим. Чтобы достичь высокой энергии излучения, создают условия для многократного прохождения луча. Для этого активная среда помещается в оптический резонатор.

Простейший оптический резонатор открытого типа (отражающие стенки не создают замкнутого объёма) образуют два плоских зеркала, расположенные параллельно друг другу. Электромагнитные волны, распространяющиеся вдоль оптической оси резонатора, создают в резонаторе плоскую стоячую волну (рис.3.14). Частоты или длины волн резонансных колебаний называются *модами*. Вдоль оси резонатора укладывается большое число q полуволн: если длина резонатора 10 см, а длина волны 1 мкм, то $q = 2 \cdot 10^5$. Расстояние между соседними модами много меньше ширины линии излучения активной среды. Поэтому в резонаторе могут одновременно возбуждаться несколько мод. Такой резонатор называют многомодовым.

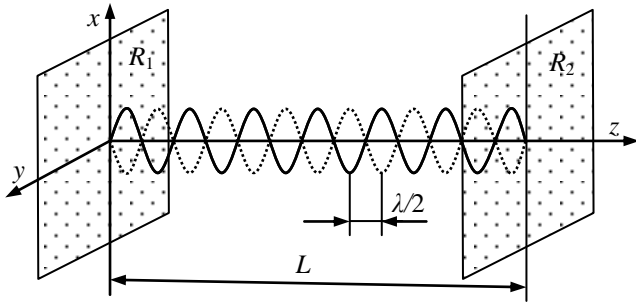


Рис.3.14. Оптический резонатор открытого типа:
 L – длина резонатора, R_1 и R_2 – плоские зеркала

Для вывода излучения из резонатора одно из зеркал делается полупрозрачным. Оно определяет основные потери энергии в резонаторе – потери на излучение.

Важной характеристикой является *добротность* резонатора – отношение энергии, запасенной в резонаторе, к величине энергии, теряемой за период колебаний:

$$Q = \frac{2\pi L}{\lambda(1 - R_2)}, \quad (3.28)$$

где R_2 – коэффициент отражения полупрозрачного зеркала, L – длина резонатора.

Для генерации излучения необходимо, чтобы увеличение энергии волны $\delta\rho_{\text{усил}}$ за один проход активной среды (3.19)

$$\delta\rho_{\text{усил}} = h\nu B_{21}\rho_\nu(N_2 - N_1)\delta t,$$

где $\delta t = L/c$ – время прохождения волной расстояния L между зеркалами резонатора превышало её потери:

$$\delta\rho_{\text{потерь}} = \frac{\rho_\nu}{Q} \cdot \frac{q}{2},$$

где $q = 2L/\lambda = 2Lv/c$ – количество полувольт, укладываемых на длине резонатора.

Отсюда получим условие генерации колебаний:

$$\begin{aligned} \delta\rho_{\text{усил}} &> \delta\rho_{\text{потерь}}, \\ hB_{21}(N_2 - N_1) &> \frac{1}{Q}. \end{aligned} \quad (3.29)$$

Инверсия заселённости уровней достигается при возбуждении 50% ионов Cr^{3+} . Минимальное (пороговое) значение объём-

ной плотности энергии накачки, необходимое для возникновения инверсного состояния, составляет около 3 Дж/см^3 .

Рубиновые лазеры работают в импульсно-периодическом режиме, с длительностью импульсов порядка 1 мкс. Накачка обычно осуществляется излучением мощных ксеноновых ламп-вспышек. При длительности импульсов накачки 1 мс они должны создавать излучение мощностью

$$\frac{3 \text{ Дж/см}^3}{0,001 \text{ с}} = 3 \text{ кВт/см}^3.$$

В лазерах используются рубиновые круглые стержни диаметром $1 \div 2 \text{ см}$ и длиной $10 \div 30 \text{ см}$. Торцевые поверхности полируют и на них напыляют зеркальные покрытия. Высокая теплопроводность рубина позволяет получать последовательности импульсов большой мощности с высокой частотой повторения.

3.2.2. Оптоволоконные лазеры

Принцип работы оптоволоконного лазера

Волоконный лазер состоит из источника накачки (лазерных диодов), волоконного усилителя (активного световода), глухого R_1 и полупрозрачного R_2 зеркал (рис.3.15).

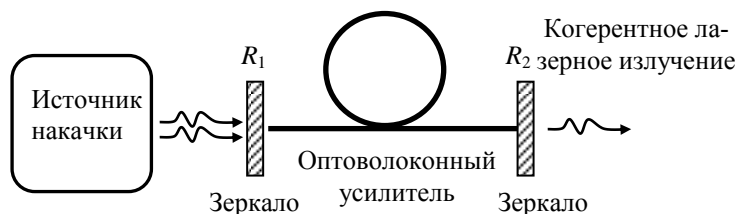
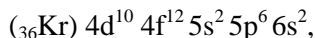


Рис.3.15. Схема волоконного лазера

В качестве активного световода используется кварцевое оптоволокно, легированное ионами одного из редкоземельных элементов: неодима ${}_{60}\text{Nd}^{3+}$, гольмия ${}_{67}\text{Ho}^{3+}$, эрбия ${}_{68}\text{Er}^{3+}$, тулия ${}_{69}\text{Tm}^{3+}$ или иттербия ${}_{70}\text{Yb}$. Для атомов этих элементов характерна дестройка внутренней 4f-оболочки при заполненных внешних оболочках. Например, электронная конфигурация атома эрбия



где (36Kr) обозначает электронную конфигурацию атома криптона, цифрой обозначено главное квантовое число $n = 4, 5, 6$, буква обозначает орбитальное квантовое число $l = 0, 1, 2, 3, \dots = (s, p, d, f, \dots)$. Электронные оболочки $5s^2 5p^6 6s^2$ заполнены, оболочка $4f$ заполнена частично. Заполненная f -оболочка должна содержать $2(2l + 1) = 2(2 \cdot 3 + 1) = 14$ электронов.

Наличие незаполненной электронной оболочки приводит к появлению выраженной дискретной структуры возбуждённых состояний атомов и ионов. Диаграмма энергетических уровней иона Er^{3+} приведена на рис.3.4.

Спектральные полосы люминесценции ионов редкоземельных элементов лежат в близкой ИК области:

$\text{Nd}^{3+} - 0,93; 1,1; 1,34 \text{ мкм};$

$\text{Ho}^{3+} - 2,1 \text{ мкм};$

$\text{Er}^{3+} - 1,53 \text{ мкм};$

$\text{Tm}^{3+} - 1,8 \text{ мкм};$

$\text{Yb}^{3+} - 1,1 \text{ мкм}.$

Для волоконных лазеров были разработаны кварцевые *волоконные световоды с двойной оболочкой* (рис.3.16). Активной средой лазера является одномодовая сердцевина волокна, легированная ионами редкоземельного элемента, а также примесями, повышающими показатель преломления.

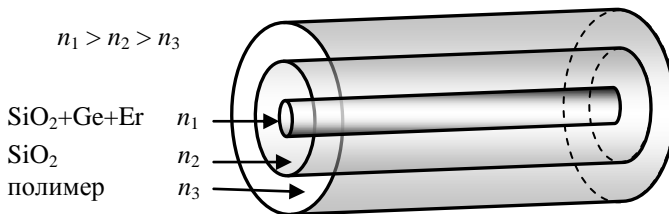


Рис.3.16. Волоконные световоды с двойной оболочкой

Сердцевина окружена внутренней оболочкой из кварцевого стекла. Внутренняя оболочка окружена внешней оболочкой из полимерного материала с показателем преломления более низким, чем у кварцевого стекла.

Внутренняя оболочка вместе с внешней образуют многомодовый световод, по которому распространяется излучение накачки (рис.3.17). Распространяющееся по внутренней оболочке излучение накачки проникает в сердцевину оптоволокна и поглощается ионами редкоземельного элемента, переводя их в возбуждённое состояние.

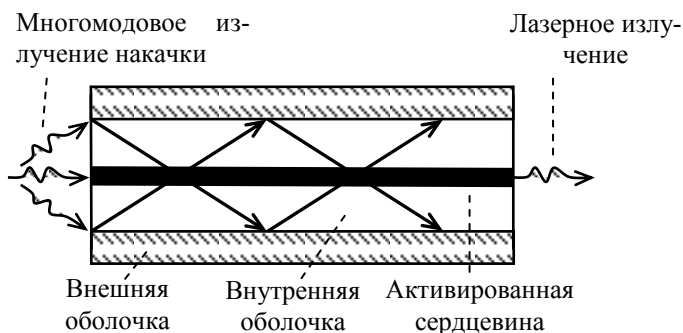


Рис. 3.17. Схема накачки волоконного лазера

В качестве источников накачки волоконных лазеров используются как отдельные лазерные диоды, так и группы лазерных диодов с волоконным выходом. Ряд фирм производит лазерные модули накачки с выходной мощностью до нескольких сотен ватт. Выходной волоконный световод таких модулей имеет сердцевину диаметром 200 мкм и больше.

Зеркала резонатора (рис.3.15) волоконного лазера обеспечивают многократное прохождение излучения лазера через усиливающую среду (сердцевину оптоволокна), и если достигается порог генерации (коэффициент усиления превышает суммарные потери при двойном проходе между зеркалами), мощность излучения резко возрастает. Максимальная мощность излучения ограничена как мощностью источника накачки, так и эффектом насыщения, при котором усиление излучения становится равным его потерям в резонаторе.

Резонатор на основе волоконных брэгговских решёток

В качестве зеркал резонаторов в волоконных лазерах используются волоконные брэгговские решётки, сформированные

непосредственно в активном световоде. На рис.3.18 показана схема волоконного лазера с брэгговскими зеркалами.

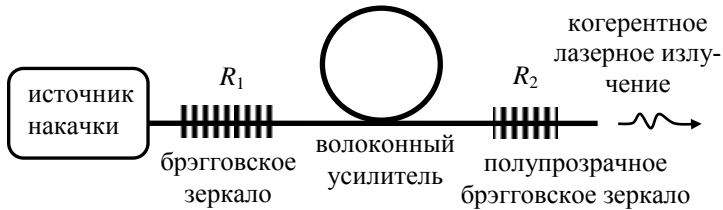


Рис. 3.18. Схема волоконного лазера с зеркалами на основе внутриволоконных брэгговских решеток

Волоконная брэгговская решетка (ВБР) представляет собой участок оптического волокна (ОВ), в сердцевине которого создано периодическое изменение показателя преломления (ПП) в продольном направлении (рис. 3.19).

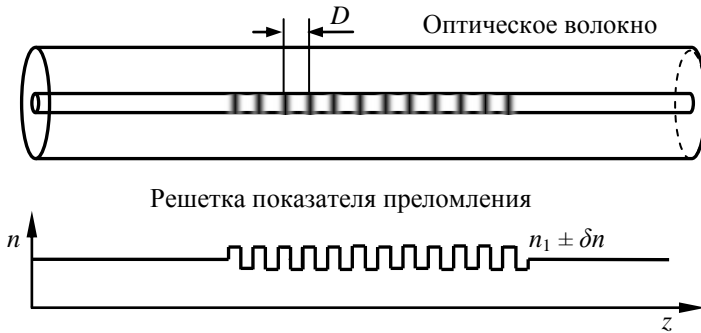


Рис.3.19. Волоконная решётка Брэгга

Свет, распространяющийся в сердцевине ОВ, отражается от каждой границы изменения показателя преломления, от каждого штриха решётки. Если длина волны удовлетворяет условию Брэгга (3.30), то волны, отражённые от штрихов решётки ПП имеют одинаковую фазу, складываются и образуют результирующую отражённую волну, распространяющуюся по волокну в обратном направлении.

Условие Брэгговского отражения

$$\lambda_B = 2n_{\text{эфф}}D, \quad (3.30)$$

где λ_B – длина волны брэгговского резонанса, $n_{\text{эфф}}$ – эффективный ПП сердцевины волокна, D – период брэгговской решётки.

Если условия Брэгга не выполняются, то волны, отражённые от штрихов решётки, не синфазны и, складываясь, образуют рассеянный свет.

Основными характеристиками внутриволоконных брэгговских решёток являются:

- период модуляции показателя преломления D ;
- амплитуда изменения показателя преломления $\Delta n \sim 10^{-5} \div 10^{-3}$;
- число штрихов решётки N_p и её длина L .

Изменяя параметры внутриволоконных брэгговских решёток, можно создавать волоконные зеркала с заданными коэффициентами отражения ($1 \div 99,9\%$), тем самым формировать полностью волоконный резонатор, рассчитанный на нужную длину волны и нужную добротность.

Селективность по частоте брэгговских решёток позволяет получить лазер, работающий на одной продольной моде с узкой частотной полосой генерации.

Метод записи брэгговской решётки в германо-силикатную сердцевину оптоволокна

Разработаны и применяются несколько методов создания периодического изменения показателя преломления кварцевого волокна, легированного германием. Все они основаны на изменении ПП под действием ультрафиолетового излучения. На рис.3.20 показана схема записи ВБР пошаговым методом. УФ свет фокусируется на сердцевине волокна. Свет периодически включается и освещённая точка перемещается вдоль волокна.

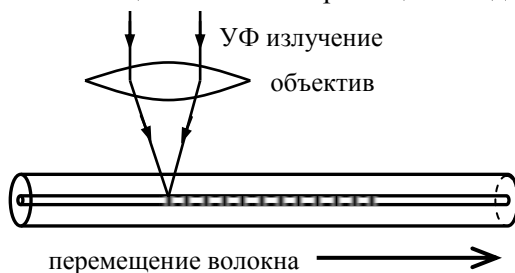


Рис. 3.20. Схема записи волоконной решётки Брегга пошаговым методом

Модель фоточувствительности германо-силикатного стекла

Был предложен ряд моделей, описывающих свойство германо-силикатного стекла изменять ПП под действием УФ излучения. Мы рассмотрим электрострикционную модель формирования ВБР (рис.3.21).

Германий имеет два стабильных состояния окисления Ge^{+4} и Ge^{+2} .

Поглощение германо-силикатных стекол имеет три максимума в УФ диапазоне спектра, расположенных вблизи 185, 242 и 325 нм. В результате внутреннего фотоэффекта в стекле воз-

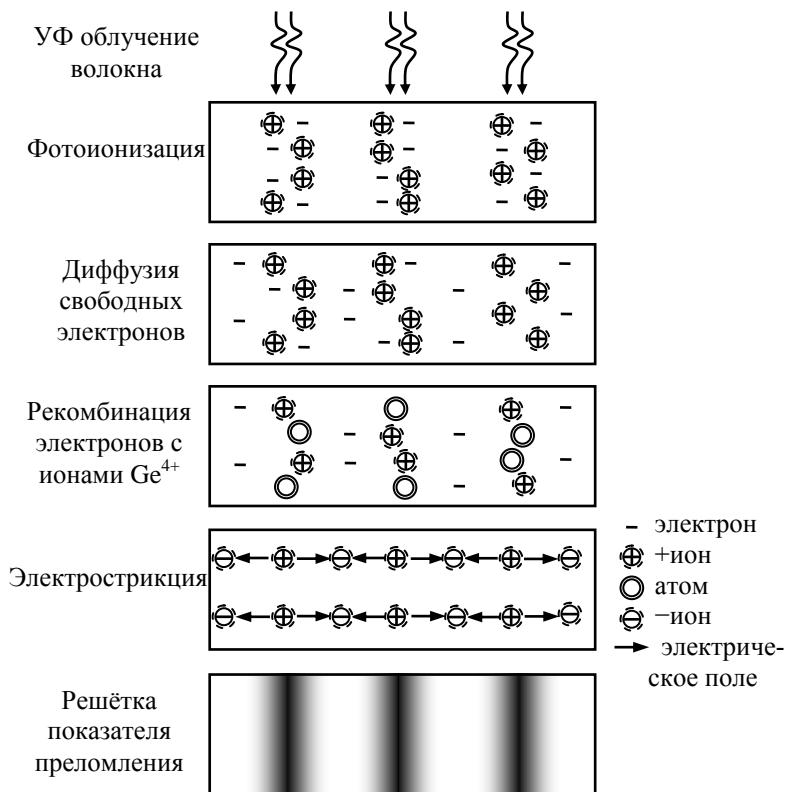


Рис.3.21. Электрострикционная модель формирования волоконной решётки Брэгга

никают положительно заряженные центры и свободные электроны. Свободные электроны диффундируют из области с большей освещённостью в область с меньшей освещённостью и захватываются ловушками электронов – ионами Ge^{+4} . В результате часть свободных электронов не возвращается на исходные доноры.

В максимуме облучения образуется избыточный положительный заряд. Избыточный отрицательный заряд создает симметричные максимумы между положительными объёмными зарядами. В решётке зарядов возникает электрическое поле, периодически меняющееся вдоль оси световода. В результате электрострикции в области максимального электрического поля возникает сжатие сердцевины ОВ, тогда как в области минимального электрического поля по законам упругости – растяжение. Таким образом, решётка зарядов создает периодическую деформацию объёма сердцевины, модуляцию её плотности и показателя преломления.

Непрерывные волоконные лазеры на основе активных световодов, легированных ионами редкоземельных элементов

Непрерывные волоконные лазеры находят применение в различных областях науки и техники, поскольку они оказываются более компактными, дешёвыми и надёжными по сравнению с другими твёрдотельными лазерами. Использование многооболочечных одномодовых активных световодов позволяет достичь высокой эффективности преобразования излучения накачки в лазерное излучение (до 40÷80%).

Образующийся при оптической накачке активного волокна тепловой поток распределён вдоль большой длины световода. Это позволяет эффективно рассеивать тепло в окружающую среду, что особенно важно при непрерывной генерации излучения большой мощности.

В настоящее время созданы волоконные лазеры на основе активных световодов, легированных Er-Yb и Tm , с мощностью одномодового излучения более 100 Вт и высоким к.п.д. (до 20%) преобразования электрической энергии в энергию излучения.

4. УПРАВЛЕНИЕ ИЗЛУЧЕНИЕМ В ВОЛНОВОДАХ

4.1. Интегральный электрооптический модулятор фазы

Для преобразования электрического сигнала в оптический сигнал, передаваемый по волоконному световоду, используют электрооптические модуляторы. Цифровому сигналу, представляющему последовательность импульсов напряжения или тока, ставится в соответствие последовательность импульсов интенсивности света.

При частотах импульсного сигнала до 5 ГГц обычно применяют модуляцию тока накачки лазера. Длительность фронтов импульсов определяется переходными процессами в оптическом резонаторе лазера – временами нарастания и затухания колебаний.

При частотах 20 ГГц, 40 ГГц и выше используют электрооптические модуляторы, принцип действия которых основан на использовании эффекта Поккельса. Источник излучения – лазер работает в непрерывном режиме. Модуляторы фазы световой волны, модулятор интенсивности, группы электродов и система волноводов, образующих интерферометр Маха-Цендера, объединены (интегрированы) на поверхности оптически анизотропного кристалла ниобата лития.

Интегральные электрооптические модуляторы имеют ряд достоинств, определяющих их практическое применение. Среди них – совместимость с оптоволоконными линиями, малая нелинейность, высокий динамический диапазон и широкий диапазон модулирующих частот.

Интегральный электрооптический модулятор Маха-Цендера (electro-optical modulators of the Mach-Zehnder) изготавливают на поверхности пластины, вырезанной из монокристалла ниобата лития. В основе его работы лежит *линейный электрооптический эффект* (эффект Поккельса).

4.1.1. Электрооптические свойства кристалла ниобата лития

Ниобат лития (LiNbO_3) – смешанный оксид ниобия и лития, бесцветный кристалл тригональной (или ромбоэдрической)

кристаллографической системы. Его кристаллическая решётка имеет одну ось симметрии третьего порядка ($L3$).

Направление (ось z), совпадающее с кристаллографической осью $L3$, называется *оптической осью кристалла*. *Главным сечением кристалла* называется плоскость, в которой лежат оптическая ось кристалла и волновой вектор \vec{k} светового луча, проходящего через кристалл (рис.4.1).

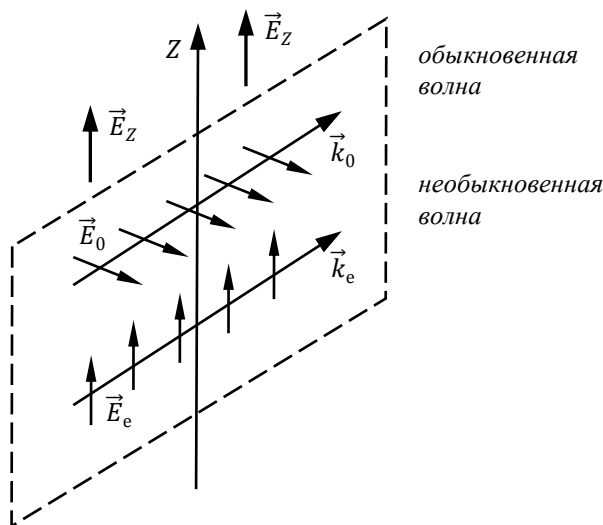


Рис.4.1. Главное сечение кристалла и направления поляризации обыкновенной и необыкновенной волн

Одноосные кристаллы являются *оптически анизотропными*, т. е. показатель преломления и фазовая скорость световой волны зависят от направления её распространения и от ориентации плоскости поляризации. Диэлектрическая проницаемость ϵ и показатель преломления ($n = \sqrt{\epsilon}$) в направлении оптической оси и в направлениях, перпендикулярных к ней, имеют различные значения. Следовательно, электромагнитным волнам с различными направлениями колебания вектора напряжённости электрического поля \vec{E} соответствуют разные значения фазовой скорости $v = c/n$ и показателя преломления. Это явление носит название *двойное лучепреломление*.

Колебания вектора напряжённости электрического поля \vec{E}_0 в “обыкновенной” волне совершаются в направлении, перпендикулярном главному сечению кристалла для этой волны (рис.4.2). В “необыкновенной” волне колебания вектора напряжённости \vec{E}_e совершаются в плоскости, совпадающей с главным сечением для данной волны (рис.4.1). Для С-диапазона длин волн ($\lambda = 1530 \div 1565$ нм), преимущественно используемого в волоконно-оптических системах, “обыкновенный” n_o и “необыкновенный” n_e показатели преломления имеют величины $n_o = 2,225$ и $n_e = 2,135$.

Электрооптический эффект в ниобате лития состоит в изменении показателя преломления кристалла под действием электрического поля. В случае, когда световая волна, поляризованная вдоль оси Z , распространяется в направлении, перпендикулярном к оси Z (“необыкновенная” волна \vec{k}_e , см. рис.4.1), показатель преломления линейно зависит от напряжённости электрического поля:

$$n_z = n_e - \frac{n_e^3}{2} r_{33} E_z, \quad (4.1)$$

где: n_e – показатель преломления для “необыкновенной” волны в отсутствие электрического поля, E_z – величина напряжённости внешнего электрического поля, приложенного вдоль кристаллографической оси Z ; $r_{33} = 30,9 \cdot 10^{-12}$ м/В – *электрооптический коэффициент* ниобата лития.

Электрооптический эффект практически безынерционен (быстродействие порядка $10^{-10} \div 10^{-11}$ с) благодаря чему верхний предел частоты колебаний напряжённости электрического поля в принципе может достигать 100 ГГц. На практике, у современных модуляторов ширина полосы по уровню -3 дБ составляет $40 \div 50$ ГГц.

Ниобат лития прозрачен для длин волн от 0,4 до 5 мкм, что делает его почти идеальным материалом для применения в оптоволоконных системах и интегрально-оптических устройствах.

4.1.2. Конструкция интегрального электрооптического модулятора Маха-Цендера

Интегральный электрооптический модулятор представляет собой интерферометр Маха-Цендера, изготовленный на X -срезе монокристалла ниобата лития. X -срезом называют тонкую пластину, вырезанную из монокристалла таким образом, что кристаллографическая ось X перпендикулярна, а кристаллографические оси Z и Y параллельны поверхности среза (рис.4.2).

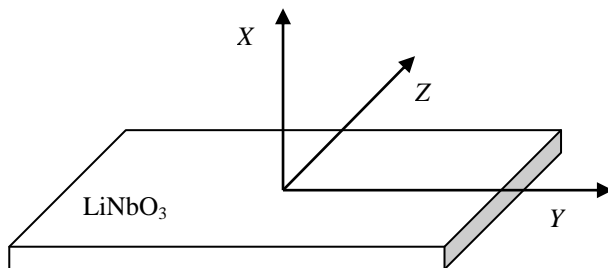


Рис.4.2. X -срез кристалла ниобата лития

Интерферометр состоит из двух Y -разветвителей, соединённых прямолинейными участками канальных волноводов. Волноводы направлены вдоль кристаллографической оси Y и перпендикулярно оси Z . Рядом с волноводами напылены металлические электроды (рис.4.3).

Для изготовления модулятора на поверхность пластины методом фотолитографии наносят защитные маски, через которые производят формирование волноводов и напыление электродов.

Канальные световоды толщиной до 1 мкм создаются методом протонного обмена в расплаве бензойной кислоты C_6H_5COOH при $t \approx 200^\circ C$. В ходе реакции ионы лития в поверхностном слое частично замещаются на ионы водорода: $H_xLi_{1-x}NbO_3$, $x = (0,4 \div 0,6)$. Показатель преломления канального световода оказывается больше, чем показатель преломления в объёме кристалла. Разность показателей преломления волновода и подложки или контрастность волновода составляет около 0,1.

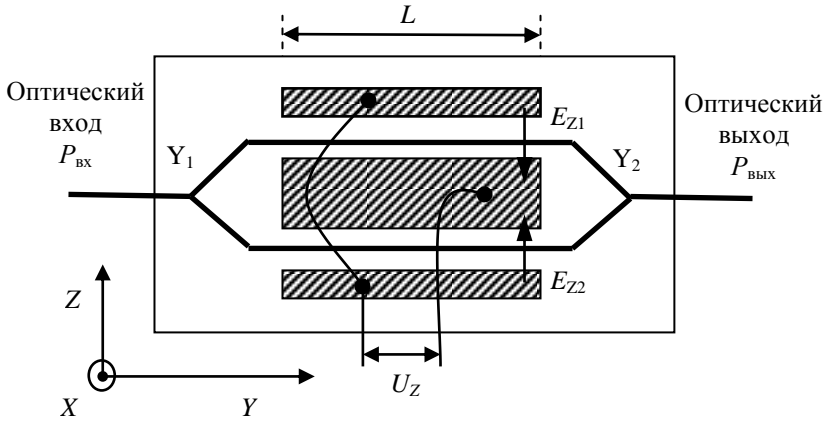


Рис.4.3. Структурная схема интегрального электрооптического модулятора Маха-Цендера

Оптическое лазерное излучение $P_{\text{вх}}$ подаётся на вход интерферометра Маха-Цендера через волоконные световоды типа “Panda” сохраняющие поляризацию излучения. В Y_1 -разветвителе излучение разделяется на два равных по мощности ($50 \pm 2\%$) потока и поступает в два плеча интерферометра. Оптическое излучение в волноводах должно быть поляризовано в направлении оси Z кристалла. Вдоль оси Z направлен и вектор напряжённости электрического поля \vec{E} между электродами (рис.4.3).

X -срез является главным сечением кристалла для волн, распространяющихся по волноводам. Волна, распространяющаяся по волноводу в направлении Y и поляризованная вдоль оси Z , является необыкновенной волной в кристалле (см. рис.4.1).

Рядом с волноводами на поверхности пластины напылены металлические электроды. При подаче на них напряжения U в поверхностном слое кристалла, по которому проходят волноводы, возникают электрические поля

$$E_{z1} \approx E_{z2} \approx \frac{U}{\epsilon_e d}.$$

Электрические поля в плечах интерферометра направлены в противоположные стороны (рис.4.4), что вызывает равные по величине и противоположные по знаку изменения показателя преломления (4.1) в волноводах за счёт электрооптического эффекта.

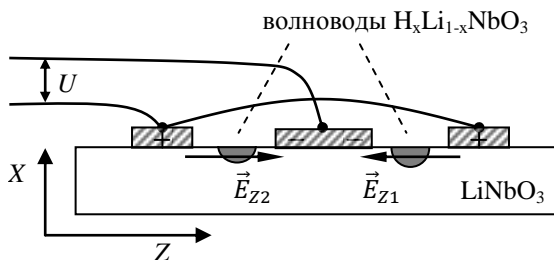


Рис.4.4. Поперечное сечение модулятора Маха-Цендера

4.1.3. Принцип действия интегрального электрооптического модулятора фазы

При распространении волн вдоль плеч интерферометра между ними возникает *разность фаз*, величина которой зависит от оптических длин волноводов. Оптическая длина определяется как физической длиной волновода, так и его показателем преломления, который изменяется под действием электрического поля.

Интенсивность света $P_{\text{вых}}$ на выходе интерферометра определяется разностью фаз волн, интерферирующих в Y₂-разветвителе, которая, в свою очередь, определяется величиной напряжения, приложенного к электродам.

После первого Y₁-разветвителя входящая световая волна e_0 разделяется на две волны с равными амплитудами $e_0/2$ в верхнем e_1 и нижнем e_2 плечах интерферометра Маха-Цендера:

$$e_1 = \frac{e_0}{2} \cos \left(2\pi vt - \frac{2\pi}{\lambda_1} y \right), \quad (4.2)$$

$$e_2 = \frac{e_0}{2} \cos \left(2\pi vt - \frac{2\pi}{\lambda_2} y \right), \quad (4.3)$$

здесь λ_1 , λ_2 и λ_0 – длины волн оптического излучения с частотой ν в волноводах верхнего и нижнего плеч интерферометра и в вакууме:

$$\lambda_1 = \lambda_0/n_{z1}, \quad \lambda_2 = \lambda_0/n_{z2}.$$

В электрических полях E_{z1} и E_{z2} , создаваемых электродами и направленных вдоль или против оси Z , показатели преломления в верхнем n_{z1} и нижнем n_{z2} волноводах вследствие электрооптического эффекта будут различны (4.1):

$$n_{z1} = n_e + \frac{n_e^3}{2} r_{33} E_{z1}, \quad (4.4)$$

$$n_{z2} = n_e - \frac{n_e^3}{2} r_{33} E_{z2}. \quad (4.5)$$

Разность фаз световых волн e_1 и e_2 после прохождения области электрооптического взаимодействия длиной L (рис.4.3):

$$2\Delta\varphi = \left(\frac{2\pi}{\lambda_0} n_{z1} L - \frac{2\pi}{\lambda_0} n_{z2} L \right). \quad (4.6)$$

Подставив (4.4) и (4.5) в формулу (4.6), получим, что наличие поля $E_z = E_{z1} = E_{z2}$ приводит к возникновению разности фаз в плечах интерферометра:

$$2\Delta\varphi = \frac{2\pi}{\lambda_0} n_e^3 r_{33} E_z L. \quad (4.7)$$

4.2. Интегральный электрооптический модулятор амплитуды

4.2.1. Принцип действия интегрального электрооптического модулятора амплитуды

В результате интерференции волн e_1 и e_2 на выходе Y_2 -разветвителя (объединителя) образуется световая волна

$$e_{\text{вых}} = \frac{e_0}{2} \cos(2\pi\nu t + \Delta\varphi) + \frac{e_0}{2} \cos(2\pi\nu t - \Delta\varphi),$$

$$e_{\text{вых}} = [e_0 \cos(\Delta\varphi)] \cos(2\pi\nu t). \quad (4.8)$$

Выражение в квадратных скобках – амплитуда волны.

Интенсивность (мощность) света на выходе интерферометра

$$P_{\text{вых}} = P_{\text{вх}} \cos^2(\Delta\varphi) = \frac{1}{2} P_{\text{вх}} [1 + \cos(2\Delta\varphi)], \quad (4.9)$$

где $P_{\text{вх}}$ – мощность оптического излучения на входе.

Для количественной характеристики чувствительности интерферометра используют величину E_π – величину Z-компоненты напряжённости внешнего электрического поля, необходимую для изменения разности фаз (4.7) в плечах интерферометра на угол $2\Delta\varphi = \pi$:

$$\pi = \frac{2\pi}{\lambda_0} n_e^3 r_{33} E_\pi L, \\ E_\pi = \frac{\lambda_0}{2n_e^3 r_{33} L}. \quad (4.10)$$

При таком поле световые волны гасят друг друга, и интенсивность света на выходе интерферометра равна нулю. С использованием этого параметра (4.10) уравнение (4.9) принимает вид

$$\frac{P_{\text{ВЫХ}}}{P_{\text{ВХ}}} = \frac{1}{2} \left[1 + \cos \left(\pi \frac{E_z}{E_\pi} \right) \right], \quad (4.11)$$

или

$$\frac{P_{\text{ВЫХ}}}{P_{\text{ВХ}}} = \frac{1}{2} \left[1 + \cos \left(\pi \frac{U_z}{U_\pi} \right) \right], \quad (4.12)$$

где U_z – напряжение, подаваемое на электроды модулятора, U_π – величина “полуволнового” напряжения.

Передаточной характеристикой электрооптического модулятора называется зависимость относительной интенсивности света $P_{\text{ВЫХ}}/P_{\text{ВХ}}$ на выходе модулятора от нормированной величины модулирующего напряжения U_z/U_π (рис.4.5).

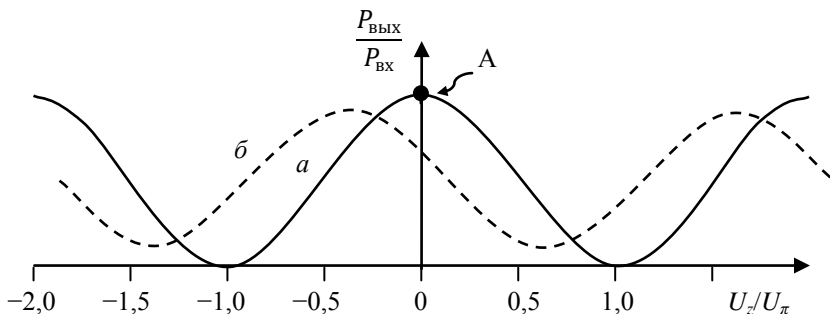


Рис.4.5. Передаточные характеристики: *a* – идеального модулятора с симметричной топологией волноводов, *б* – реального модулятора

4.2.2. Выбор рабочей точки модулятора

Рабочей точкой (РТ) модулятора называется точка на передаточной характеристике, которая соответствует имеющейся в интерферометре постоянной разности фаз. При равной оптической длине плеч интерферометра и соответственно нулевой разности фаз рабочая точка А находится в максимуме передаточной характеристики (характеристика “а” на рис.4.5). Такое положение РТ соответствует участку с максимальной кривизной передаточной характеристики, что приводит к искажению передаваемого сигнала.

Для того, чтобы модулятор работал на линейном участке передаточной характеристики, необходимо обеспечить начальный сдвиг фаз в каналах интерферометра $\varphi_0 = \pi/2$:

$$\frac{P_{\text{вых}}}{P_{\text{вх}}} = \frac{1}{2} \left[1 + \cos \left(\pi \frac{U_{\text{сигн}}}{U_{\pi}} + \frac{\pi}{2} \right) \right]. \quad (4.13)$$

При этом рабочая точка В будет находиться в середине линейного участка передаточной характеристики модулятора (рис.4.6). Этого можно достичь, если приложить постоянное электрическое $E_{\text{кв}}$ поле путём подачи на электроды модулятора

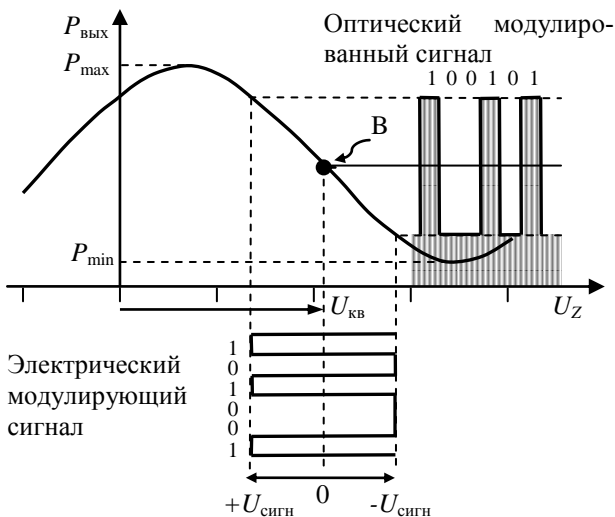


Рис.4.6. Положение рабочей точки интерферометра в амплитудном модуляторе

постоянного напряжения $U_{\text{кв}}$:

$$U_z = U_{\text{кв}} + U_{\text{сигн}}. \quad (4.14)$$

Это смещение РТ называется *смещением в квадратуру*. Как видно из графика на рис.4.6, при смещении РТ в квадратуру средняя выходная интенсивность излучения равна половине входной интенсивности, что соответствует оптическим потерям в 3 дБ. Часть энергии излучения, не прошедшая через Y_2 -разветвитель в волноводный канал, рассеивается в окружающее пространство. Глубина модуляции оптического сигнала (максимальное значение):

$$M[\text{дБ}] = P_{\text{max}}[\text{дБм}] - P_{\text{min}}[\text{дБм}]. \quad (4.15)$$

Сделав преобразование тригонометрической функции косинуса в формуле (4.12), получим

$$P_{\text{вых}} = \frac{1}{2} P_{\text{вх}} \left[1 - \sin \left(\pi \frac{U_z}{U_{\pi}} \right) \right], \quad (4.16)$$

$$\frac{P_{\text{вых}}}{P_{\text{вх}}} = \frac{1}{2} \left[1 - \sin \left(\pi \frac{U_z}{U_{\pi}} \right) \right]. \quad (4.17)$$

На линейном участке передаточной характеристики (вблизи рабочей точки, $U_z \ll U_{\pi}$):

$$\begin{aligned} \sin \left(\pi \frac{U_z}{U_{\pi}} \right) &\approx \pi \frac{U_z}{U_{\pi}}, \\ \frac{P_{\text{вых}}}{P_{\text{вх}}} &= \frac{1}{2} \left[1 - \pi \frac{U_z}{U_{\pi}} \right]. \end{aligned} \quad (4.18)$$

При подаче на модулятор переменного (СВЧ) напряжения $U_z(t)$ интенсивность светового сигнала на выходе интерферометра будет изменяться с той же частотой.

Использование электрооптических модуляторов позволяет существенно улучшить технические характеристики трактов передачи аналоговых и цифровых сигналов, обеспечить электромагнитную совместимость линий передачи информации.

На основе компонентов интегральной и волоконной оптики возможно также создание датчиков для контроля таких физических величин, как напряжённости электрического и магнитного полей, температура, давление, газовый состав. Интегральные электрооптические модуляторы Маха-Цендера применяются в системах связи, а также являются одним из основных узлов волоконно-оптических гироскопов.

Приложение

Единицы измерения световых величин

Потоком энергии Φ электромагнитной волны называется количество энергии, переносимое электромагнитной волной через некоторую поверхность за единицу времени. Единица измерения потока энергии равна 1 Дж/с или 1 Вт.

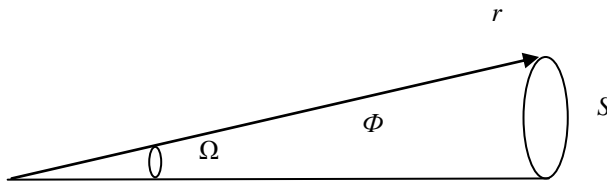
Для измерения энергии, переносимой электромагнитной волной оптического диапазона ($\lambda = 0,40 \div 0,76$ мкм), используются фотометрические (световые) единицы. В фотометрии световым потоком Φ называют мощность светового излучения, оцениваемую по его воздействию на человеческий глаз. Световой поток измеряют в люменах (лм). Для монохроматического излучения с длиной волны $\lambda = 0,555$ мкм, при которой чувствительность глаза максимальна, световой поток равен 683 лм, если мощность излучения (поток энергии) равна 1 Вт.

Силой света I точечного источника называется величина светового потока, излучаемого внутри единичного телесного угла

$$I = \Phi / \Omega.$$

Телесный угол Ω определяется как отношение площади S поверхности шарового сегмента к квадрату радиуса сферы

$$\Omega = S / r^2.$$



Единицей силы света является кандела (кд).
1 кд = 1 лм/стерадиан.

Источник излучения с $\lambda = 0,555$ мкм и силой света $I = 1$ кд излучает внутри телесного угла в 1 стерадиан поток энергии $\Phi = 1/683$ Вт.

Освещённостью E поверхности называется световой поток, приходящийся на единицу площади освещаемой поверхности

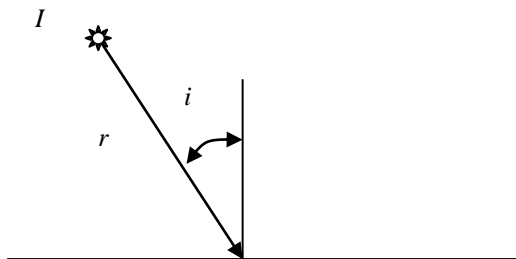
$$E = \Phi/S.$$

Световой поток в 1 лм, равномерно распределенный по площади 1 м^2 , создает освещённость в 1 люкс (лк).

Освещённость, которую создает точечный источник с силой света I на поверхности, удаленной на расстояние r от источника

$$E = \frac{I \cos i}{r^2},$$

где i – угол падения.



Яркость светящейся поверхности S численно равна силе света, излучаемой единицей поверхности в направлении, перпендикулярном этой поверхности

$$B = I/S.$$

За единицу яркости (1 кд/м^2) принимают яркость такого источника, который излучает с 1 м^2 светящейся поверхности свет силой 1 кд.

Список литературы

1. Азанова И. С., Шевцов Д. И. Физические свойства и структура волоконно-оптических систем: учеб.-метод. пособие. / Пермь: изд-во Перм. гос. ун-та, 2007. С. 42
2. Азанова И.С., Волынцев А.Б., Тайсин И.Ф., Шевцов Д.И. Метастабильные фазы в протонообменных волноводах на х-срезе ниобата лития // Физика твердого тела. 2006. Т. 48. № 6. С. 1059-1063.
3. Андропова И.А., Малыкин Г.Б. Физические проблемы волоконной гироскопии на эффекте Саньяка // УФН. 2002. Т. 172, вып.8. Р.849-873.
4. Афанасьев В.М. Электрооптический модулятор по схеме интерферометра Маха–Цендера // Прикладная фотоника. 2016. Т. 3, № 4. С. 341-369.
5. Бурков В. Д., Губин В. П., Сазонов А. И. Волоконно-оптические гироскопы с эрбиевыми волоконными источниками излучения // Лесной вестник. М., 2007. №2. С. 8–13. URL: http://les-vest.msfu.ru/les_vest/2007.pdf (дата обращения: 04.10.2017).
6. Буш А. А. Пироэлектрический эффект и его применения: учеб. пособие / Государственное образовательное учреждение высшего профессионального образования «Московский государственный институт радиотехники, электроники и автоматики (технический университет)». М., 2005. 212 с.
7. Варжель С.В. Волоконные брэгговские решетки. СПб: изд-во Университета ИТМО, 2015. 65 с.
8. Васильев С.А., Дианов Е.М., Курков А.С., Медведков О.И., Протопопов В.Н. Фотоиндуцированные внутриволоконные решетки показателя преломления для связи мод сердцевины-оболочка // Квантовая электроника. 1997. т. 24. С.151-154.
9. Гермогенов В.П. Материалы, структуры и приборы полупроводниковой оптоэлектроники : учеб. пособие. Томск: Изд-во. Том. гос. ун-та, 2015. 272 с.
10. Гуртов В. Твердотельная электроника. М., 2005. 408 с.
11. Демьяненко М. А., Есаев Д. Г, Овсяк В. Н., Фомин Б. И., Асеев А. Л., Князев Б. А., Кулипанов Г. Н., Винокуров Н. А. Матричные микроболометрические приемники для инфра-

красного и терагерцового диапазонов // Оптический журнал. 2009. Т.76, №12. С. 5-11.

12. Дианов Е. М. Волоконные лазеры // УФН. 2004. Т. 174, №10. С. 1139-1142

13. Дианов Е.М., Карпов В.И., Курков А.С., Протопопов В.Н. Методы сглаживания спектра усиления эрбиевых волоконных усилителей // Квантовая электроника. 1996. Т. 23. С. 1059-1064.

14. Зорин С. М. Разработка основ технологии изготовления тонкоплёночной мишени пироэлектрического электронно-оптического преобразователя инфракрасного изображения на основе дифенил-2,2',4,4'-тетраамина: дис. на соис. учен. степ. канд. техн. наук. М., 2016.

15. Игумнов Д.В., Костюнина Г.П. Основы полупроводниковой электроники. М., 2005. 392 с.

16. Квантовая и полупроводниковая электроника: лабораторный практикум: учеб.-метод. пособие / сост. И. Л. Вольхин, А. С. Ажеганов; Перм. гос. нац. исслед. ун-т. Пермь, 2018. 168 с.

17. Котов Л.В. Волоконные лазеры с высокой пиковой и средней мощностью на основе легированных эрбием световодов с двойной отражающей оболочкой: дис. на соис. учен. степ. канд. физ.-мат. М., 2015.

18. Котов О.И., Ликумович Л.Б., Медведев А.В. Интерференционный метод измерения коэффициента экстинкции двулучепреломляющих волоконных световодов // Журнал технической физики. 2007. Т.77, вып.9. С. 102-107.

19. Кузьминов Ю.С. Электрооптический и нелинейно-оптический кристалл ниобата лития. М.: Наука, 1987.

20. Кульчин Ю.Н., Змеу С.Б., Субботин Е.П., Никитин А.И. Волоконные лазеры // Вестник ДВО РАН. 2015. № 3. С.67-78.

21. Курков А. С., Дианов Е. М. Непрерывные волоконные лазеры средней мощности // Квантовая электроника. 2004. Т.34, №10. С. 881–899.

22. Курков А. С., Наний О. Е. Эрбиевые волоконно-оптические усилители // Lightwave Russian Edition. 2003. №1. р. 14–19. URL: <http://t8.ru/wp-content/uploads/2012/01/28.pdf> (дата обращения: 04.10.2017).

23. Никоноров Н. В., Шандаров С. М. Волноводная фотоника: учеб. пособие (курс лекций). СПб., 2008. 143 с.
24. Оптическая и квантовая электроника: учеб. пособие / В.И. Светцов; Иван. гос. хим.-техн. ун-т. Иваново, 2004. 122 с.
25. Паранин В. Д., Матюнин С. А. Элементы и устройства интегральной оптоэлектроники [Электронный ресурс] : курс лекций : электрон. учеб. пособие / Минобрнауки России, Самар. гос. аэрокосм. ун-т им. С. П. Королева (нац. исслед. ун-т). Самара, 2012. Режим доступа: <http://rtfmoodle.ssau.ru/course/view.php?id=66> (дата обращения: 04.10.2017).
26. Пасынков В.В., Чиркин Л.К. Полупроводниковые приборы. СПб., 2006. 480 с.
27. Пихтин Н.А., Ильин Ю.В., Лешко А.Ю., Лютецкий А.В., Станкевич А.Л., Тарасов И.С., Фетисова Н.В. Мощный широкополосный одномодовый InGaAsP/InP суперлюминесцентный диод // Письма ЖТФ. 1999. Т. 25, вып. 15. С. 16-22.
28. Полупроводниковая электроника: лабораторный практикум: учеб.-метод. пособие / сост. И.Л.Вольхин, А.С.Ажеганов; Перм. гос. нац. исслед. ун-т. Пермь, 2012. 172 с.
29. Полупроводниковая электроника: учеб.-метод. пособие / сост. И.Л.Вольхин, А.С.Ажеганов; Перм. гос. нац. исслед. ун-т. Пермь, 2012. 172 с.
30. Пономарев Р. С. Структурная модель дрейфовых явлений в интегрально-оптических схемах на основе $H_xLi_{1-x}NbO_3$ канальных волноводов: Дис. на соис. учен. степ. канд. физ.-мат. наук / Перм. гос. нац. исслед. ун-т. Пермь, 2014.
31. Румянцев В.В. Фотопроводимость и фотолюминесценция эпитаксиальных пленок и структур с квантовыми ямами на основе HgCdTe в среднем и дальнем инфракрасном диапазоне: Дис. на соис. учен. степ. канд. физ.-мат. наук / Институт физики микроструктур Российской академии наук. Нижний Новгород, 2014.
32. Физика лазеров: учеб. пособие / В.С. Айрапетян, О.К. Ушаков. Новосибирск: СГГА, 2012. 134 с.
33. Шевцов Д.И., Азанова И.С., Тайсин И.Ф., Волынцев А.Б. Метастабильные фазы в протоннообменных волноводах на X-срезе ниобата лития // Физика твердого тела. 2006. Т.48, вып. 6. С. 996-1000.

34. Шуберт Ф. Светодиоды / пер. с англ; под ред. А.Э.Юновича. М.: ФИЗМАТЛИТ, 2008. 496 с.

35. Щербаков В. В., Солодков А. Ф., Заdernовский А. А. Передача сигналов модуляции интенсивности света в аналоговых волоконно-оптических линиях связи // Радиоэлектроника, наносистемы, информационные технологии. М., 2016. Т. 8, № 1. С. 9-24.

Оглавление

1. ПРИЁМНИКИ ОПТИЧЕСКИХ СИГНАЛОВ	4
1.1.1. Собственные полупроводники	4
1.1.2. Фоторезистивный эффект	9
1.1.3. Фоторезисторы	12
1.1.4. Примесные полупроводники, р–п-переход	15
1.1.5. Фотодиоды, фототранзисторы	26
1.2. Фоточувствительные полупроводниковые приборы с зарядовой связью	34
1.2.1. Структура и принцип действия приборов с зарядовой связью	34
1.2.2. Фоточувствительные линейные ПЗС	40
1.2.3. Фоточувствительные ПЗС-матрицы	43
1.2.4. Основные характеристики ФПЗС	48
1.3. Приёмники теплового излучения	51
1.3.1. Термоэлектрические приёмники излучения	52
1.3.2. Фотоэлектрические приёмники теплового излучения	53
1.3.3. Болометрические приёмники теплового излучения	54
1.3.4. Пироэлектрические приёмники теплового излучения	62
1.3.5. Тепловизоры	70
2. СВЕТОИЗЛУЧАЮЩИЕ ПРИБОРЫ	72
2.1. Светоизлучающие диоды	72
2.1.1. Рекомбинация неравновесных носителей заряда в полупроводниках	73
2.1.2. Прямозонные и непрямозонные полупроводники	75
2.1.3. Полупроводниковые материалы светоизлучающих диодов	78
2.1.4. Светодиод на основе гетероперехода	79
2.1.5. Инжекционный светодиод на основе многослойных гетероструктур	82
2.1.6. Основные характеристики и применение светодиодов	85
2.1.7. Оптопары	90
2.2. Полупроводниковый лазер и суперлюминесцентный диод	94
2.2.1. Инжекционный полупроводниковый лазер	94
2.2.2. Физические принципы работы инжекционного лазера	95
2.2.3. Инжекционный лазер на основе многослойных гетероструктур	99
2.2.4. Основные характеристики лазерного излучения	102
2.2.5. Суперлюминесцентный диод	104
3. ГЕНЕРАЦИЯ И УСИЛЕНИЕ ИЗЛУЧЕНИЯ В ВОЛОКОННО-ОПТИЧЕСКИХ СИСТЕМАХ	110
3.1. Оптоволоконные усилители	110

3.1.1. Оптическая схема эрбиевого волоконного усилителя	111
3.1.2. Принцип работы эрбиевого усилителя	113
3.1.3. Основные характеристики эрбиевого усилителя	119
3.2. Оптоволоконные лазеры	124
3.2.1. Принцип работы лазера с оптической накачкой	124
3.2.2. Оптоволоконные лазеры	134
4. УПРАВЛЕНИЕ ИЗЛУЧЕНИЕМ В ВОЛНОВОДАХ	141
4.1. Интегральный электрооптический модулятор фазы	141
4.1.1. Электрооптические свойства кристалла ниобата лития	141
4.1.2. Конструкция интегрального электрооптического модулятора Маха-Цендера	144
4.1.3. Принцип действия интегрального электрооптического модулятора фазы	146
4.2. Интегральный электрооптический модулятор амплитуды	147
4.2.1. Принцип действия интегрального электрооптического модулятора амплитуды	147
4.2.2. Выбор рабочей точки модулятора	149
Приложение. Единицы измерения световых величин	151

Ажеганов Александр Сергеевич
Вольхин Игорь Львович
Шестакова Наталия Константиновна

КВАНТОВАЯ И ПОЛУПРОВОДНИКОВАЯ ЭЛЕКТРОНИКА КУРС ЛЕКЦИЙ

Учебное пособие

Редактор *Н. И. Стреколовская*
Корректор *А. В. Цветкова*
Компьютерная вёрстка: *И. Л. Вольхин*

Подписано в печать 03.12.2018. Формат 60х84/16.
Усл. печ. л. 9,24. Тираж 100 экз. Заказ 240

Издательский центр
Пермского государственного национального
исследовательского университета
614990, Пермь, ул. Букирева, 15

Типография ПГНИУ
614990, Пермь, ул. Букирева, 15