

Носов Ю.Р.

Оптоэлектроника. - М.: Радио и связь, 1989.

1.2. РАСПРОСТРАНЕНИЕ ВОЛНЫ ИЗЛУЧЕНИЯ В РЕАЛЬНОЙ СРЕДЕ

Выясним, к каким изменениям решения уравнений Максвелла приведет учет неидеальностей, имеющих место в реальных условиях распространения световых волн. Прежде всего очень часто приходится учитывать то, что оптическая среда небезгранична. Типичные примеры этого – распространение излучения по волоконно-оптическим световодам, имеющим цилиндрическую симметрию, или по плоским интегрально-оптическим диэлектрическим волноводам. В обоих случаях поперечные размеры оптических сред соизмеримы с длиной волны света λ , протяженность же среды вдоль направления распространения можно считать бесконечной. Другой пример – распространение световой волны внутри лазера. Здесь необходимо учитывать наличие границ по всем трем координатам, т. е. мы встречаемся с задачей возбуждения объемного резонатора.

Рассмотрим однородный бесконечно протяженный волновод (рис. 1.4), у которого обозначим через S контур, ограничивающий это сечение. Решать уравнения Максвелла удобно, пользуясь поляризационным потенциалом (или вектором Герца) Π_g , связанным с вектором-потенциалом A [см. формулу (1.6)] простым соотношением

$$A = \frac{1}{c} \cdot \frac{\partial \Pi_g}{\partial t}. \quad (1.18)$$

Если положить, что поля в волноводе меняются по периодическому закону $\exp(-i\omega t)$, омических потерь нет ($s = 0$) и вектор Π_g имеет лишь один компонент

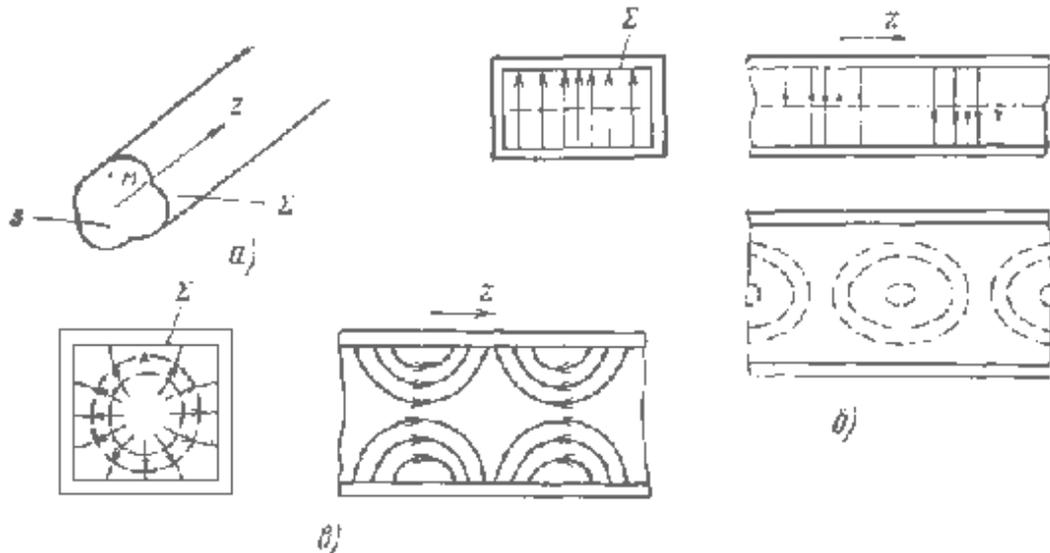


Рис. 14. К распределению излучения в оптическом волноводе:
 а — обобщенная модель волновода; б, в — структуры полей в прямоугольных и цилиндрических поперечных сечениях (непрерывные линии — силовые линии электрического поля E ; штриховые — силовые линии магнитного поля H .)

направленный вдоль оси z , то решение может быть представлено в виде

$$\text{Pr}(M, z, t) = y(M) f(z) \exp(-i\omega t), \quad (1.19)$$

где M — произвольная точка, лежащая в поперечном сечении s , имеющая координаты x и y ; $y(M)$ — собственная функция задачи о колебаниях мембраны, закрепленной по контуру, т. е.

$$D^2y + a^2y = 0;$$

$$yS = 0, \quad (1.20)$$

имеющей бесконечную систему собственных значений $\{a_m^2\}$ и собственных функций $\{y_m\}$. Частные решения z -составляющей вектора Герца $f_m(z)$ определяются из уравнения

$$\frac{\partial^2 f_m}{\partial z^2} + (k^2 - a_m^2) f_m = 0. \quad (1.21)$$

Для волны, распространяющейся в положительном направлении оси z ,

$$f_m = A_m \exp(i\sqrt{k^2 - a_m^2} z) \quad (1.22).$$

Общее решение уравнения Максвелла для волновода на рис. 1-4,а, представляющее собой суперпозицию всех частных решений, имеет вид:

$$H_z(M, z, t) = \sum_{m=1}^{\infty} A_m \psi_m(M) \exp\left[-i(\omega t - \sqrt{k^2 - \alpha_m^2} z)\right] \quad (1.23).$$

Подобные выражения могут быть получены и для волн $H_z(M, z, t)$, у которых имеется z -составляющая не электрического [как при выводе (1.23)], а магнитного вектора. В этом случае амплитуда волны представляет собой решение для колебаний мембраны, не закрепленной (свободной) по контуру S .

Итак, волна, распространяющаяся вдоль оси волновода (см. рис. 1.4,а) представляет собой дискретную совокупность плоских волн, амплитуды и фазы которых в каждой точке сечения s определяются решением задачи для колебание закрепленной (или свободной) мембраны, а волновые числа равны $\sqrt{k^2 - \alpha_m^2}$. Каждая из составляющих в (1.23) представляет тип волны, или вол-новую моду, характеризующуюся специфической структурой электромагнитного поля. Все моды волновода относятся к двум группам:

волны $H_z(M, z, t)$, у которых $E_z = 0$, $H_z \neq 0$, называемые электрическими. Е- или ТМ-волнами;

волны $H_z(M, z, t)$, у которых $E_z \neq 0$, $H_z = 0$ (т. е. поле E поляризовано в плоскости $z = \text{const}$), называемые магнитными Н- или ТЕ-волнами¹.

Любое поле в волноводе представляет суперпозицию полей ТЕ и ТМ.

Уравнение (1.23) описывает незатухающие волны, распространяющиеся в направлении z , при условии $k > \alpha_{\text{min}}$. Таким образом, с учетом (1.15) для каждого волновода получаем критические значения частоты и длины волны

$$\omega_{\text{кр}} = \alpha_{\text{min}} v_{\text{ф}} = \alpha_{\text{min}} c/n; \quad l_{\text{кр}} = 2\pi/\alpha_{\text{min}}. \quad (1.24)$$

По волноводу могут распространяться лишь волны с $\omega > \omega_{\text{кр}}$ и $l < l_{\text{кр}}$. Вследствие этого число распространяющихся мод всегда конечно, причем их тем больше, чем значительнее различие между частотой вводимого в волновод излучения ω и критическим значением собственной частоты волновода $\omega_{\text{кр}}$ (или l и $l_{\text{кр}}$).

Волновое число каждой составляющей в (1.23) $k_z = \sqrt{k^2 - \alpha_m^2}$. отсюда

$$v_{z,e} = (c/n)\sqrt{1 - (a_m/k)^2}; \quad \lambda_e = \lambda/\sqrt{1 - (a_m/k)^2}; \quad (1.25)$$

1 В аббревиатурах ТМ, ТЕ буква Т (от англ. transverse – поперечный) означает, что ТМ и ТЕ – поперечные по Н и по Е волны.

Отсюда следует, что разные составляющие электромагнитной волны волновода имеют разные групповые скорости, при приближении к критическому режиму ($1 \rightarrow 1_{кр}$) скорость распространения волны $v_{г,в} \rightarrow 0$, а длина волны $1_{в} \rightarrow \infty$, т. е. режим из волнового становится колебательным.

Проиллюстрируем общие положения на примерах. Для волновода с сечением s и виде прямоугольника (b_1 – толщина, b_2 – ширина) собственные значения уравнения (1.20) зависят от двух целочисленных аргументов m и j :

$$a_m^2 = a_{m,j}^2 = (m \cdot \pi/b_1)^2 + (j \cdot \pi/b_2)^2. \quad (1.26)$$

Простейшая электрическая волна – ТМ₁₁ (здесь $m = 1, 0$ и $j = 1, 0$, так как силовые линии магнитного поля не могут упираться в стенки и замыкаются сами на себя), а простейшая магнитная – ТЕ₁₀. Структура этих волн (расположение силовых линий электрического и магнитного полей для фиксированного момента времени) представлена на рис. 1.4,б,в. Если принять $b_2 \gg b_1$ (что типично для интегрально-оптических волноводов), то, полагая в (1.26) $m = 1$, получаем, что при $1 \gg 2b_1$ в волноводе может существовать лишь одна мода ТЕ₁₀.

Для круглого волновода радиуса r_0 собственные значения

$$a_{m,j}^2 = (\chi_j^{(m)} / r_0)^2, \quad (1.27)$$

где $\chi_j^{(m)}$ – j -и корень уравнения $J_m(\chi) = 0$, где J_m – функция Бесселя m -го порядка. Поскольку $\chi(0) = 2,4$, то с учетом (1.24) получим условие существования лишь одной волны для круглого волновода: $1 \gg 2,6r_0$.

Наглядное представление о волновых модах дает лучевая теория (см. § 1.4), основанная на концепции зигзагообразного распространения света в волноводе. Соответственно этому каждая мода волнового поля представляет собой луч, падающий на границу волновода – окружающая среда под определенным углом φ и распространяющийся по волноводу вследствие полного внутреннего отражения. Фактически тождественность волноводных мод и «косых» световых лучей вытекает и из решений уравнения Максвелла для волн $\Pi(M, z, t)$ и $\Pi_{\square}(M, z, t)$. Действительно, для всех мод, кроме ТЕ₁₀, наряду с радиальными составляющими полей E и H имеются еще и осевые (вдоль оси z) хотя бы одного из этих полей. Геометрически это и означает плоскую волну, направление распространения которой отлично от z , причем это направление изменяется во времени (отражение от границ в лучевой теории). Несмотря на зигзагообразный путь такого луча свет (точнее, энергия излучения) распространяется лишь в направлении z (см. рис. 1.4), образуя в перпендикулярном направлении стоячую волну. Предельное значение угла падения, при котором могут существовать волноводные моды, определяется условием полного внутреннего отражения [см. формулу (1.36)] с заменой ап.в.о. на qp.в.о. Для некоторых лучей внутри этого угла выполняются такие фазовые соотношения, при которых многократно отраженные от различных границ плоские волны интерферируют в фазе, т. е. не подвергаются «самогашению». Именно эти лучи (или волны) представляют собой

волноводные моды, прочие типы колебаний по мере распространения постепенно затухают. Волноводные моды определяются набором углов $Q_{m,j}$, (целочисленные аргументы m и j отражают дискретность Q в направлениях x и y), представляющих собой частные решения трансцендентного уравнения, определяющего изменение фазы волны при распространении по волноводу. Итак, волноводные моды в лучевой теории тождественны световым лучам, направленным под углами $Q_{m,j}$ к оси Oz , причем число разрешенных значений $Q_{m,j}$ конечно.

Другой практически важный случай отступления от идеализации (см. § 1.1) — это падение светового луча на границу раздела двух сред. Рассмотрим простейшую задачу. Пусть граница раздела двух изотропных полубесконечных сред с показателями преломления n_1 и n_2 представляет собой плоскость xu , на которую под углом a_1 падает плоская электромагнитная волна (рис. 1.5). Уравнения Максвелла решаются по той же схеме, что и раньше, при этом граничные условия в плоскости xu выражаются в непрерывности тангенциальных составляющих векторов E и H . Решение приводит к появлению двух составляющих распространяющейся волны — преломленной (луч 2) и отраженной (луч 3). Существуют простые взаимосвязи направлений лучей 1, 2 и 3 (законы отражения и преломления):

$$a_1 = a_3; \quad n_1 \times \sin a_1 = n_2 \sin a_2. \quad (1.28)$$

При этом все три луча лежат в одной плоскости — плоскости падения. Если разложить электрический вектор падающей волны на составляющие E_{1p} , лежащую в плоскости падения, и E_{1s} , перпендикулярную ей, то для аналогичных составляющих преломленной и отраженной волн решение уравнений Максвелла дает известные формулы Френеля:

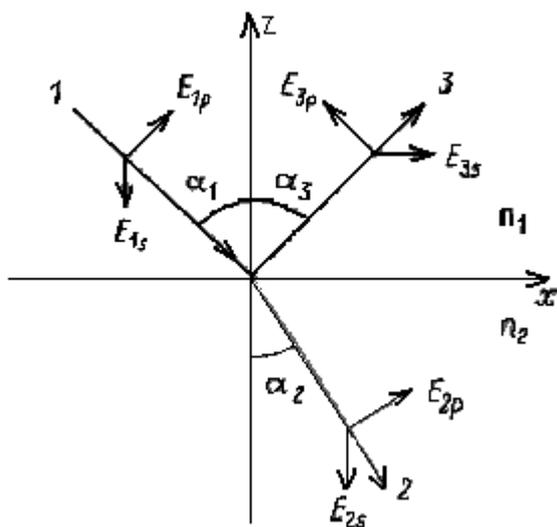


Рис. 1.5. Преломление и отражение света на границе двух сред

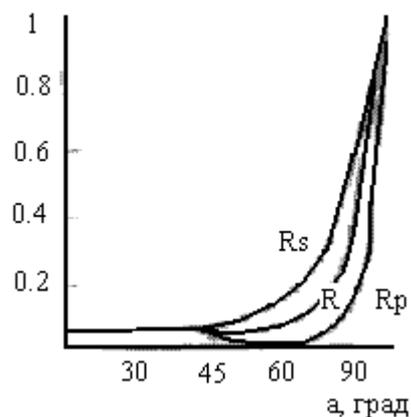


Рис.1.6. Зависимость коэффициента отражения R и его составляющих от угла падения для системы воздух-кварц (R_p соответствует поляризации света в плоскости p проходящей через лучи 1 и 3, R_s — поляризации света в плоскости s , проходящей через

луч 3 и перпендикулярной плоскости p)

$$\begin{aligned}
 E_{2p} &= \frac{2\sin\alpha_2 \cos\alpha_1}{\sin(\alpha_1 + \alpha_2) \cos(\alpha_1 - \alpha_2)} E_{1p}; \\
 E_{2s} &= \frac{2\sin\alpha_2 \cos\alpha_1}{\sin(\alpha_1 + \alpha_2)} E_{1s}; \\
 E_{3p} &= \frac{\operatorname{tg}(\alpha_1 - \alpha_2)}{\operatorname{tg}(\alpha_1 + \alpha_2)} E_{1p}; \\
 E_{3s} &= \frac{-\sin(\alpha_1 - \alpha_2)}{\sin(\alpha_1 + \alpha_2)} E_{1s}.
 \end{aligned}
 \tag{1.29}$$

Интересно, что эти формулы получены за 40 лет до того, как были предложены уравнения электродинамики. Заложенная при их выводе неверная физически, но верная математически модель — поперечные колебания «эфира» — привела к правильным результатам.

Из анализа (1.28) и (1.29) можно сделать ряд важных выводов:

1. Фаза преломленной волны всегда совпадает с фазой падающей волны (E_{2p} и E_{1p} , E_{2s} и E_{1s} имеют одинаковые знаки при любых соотношениях α_1 и α_2); фазы составляющих отраженной волны могут отличаться от фаз таких же составляющих падающей волны на $0 \dots \pi$ в зависимости от угла падения.
2. Во всех экспериментах с отраженными и преломленными волнами измеряются их интенсивности, которые согласно (1-13) пропорциональны квадратам соответствующих амплитуд. Вводимые соответственно этому коэффициенты пропускания T и отражения R (см. рис. 1.5).

$$T = I_2 / I_1; \quad R = I_3 / I_1; \quad R + T = 1,
 \tag{1.30}$$

Для любого конкретного случая могут быть определены на основе формул Френеля. Дифференциальная форма этих параметров определяется как $t = dI_2/dI_1$ и $r = dI_3/dI_1$. Для неполяризованного (естественного) света получаем

$$R = \frac{1}{2} \left[\frac{\sin^2(\alpha_1 - \alpha_2)}{\sin^2(\alpha_1 + \alpha_2)} + \frac{\operatorname{tg}^2(\alpha_1 - \alpha_2)}{\operatorname{tg}^2(\alpha_1 + \alpha_2)} \right]
 \tag{1.31}$$

Из расчетных зависимостей коэффициентов отражения от угла падения для разных случаев поляризации (рис. 1.6) видно, что до некоторого угла α_1 величина R и ее составляющие почти постоянны, а затем резко возрастают до $R = 1$.

3. При нормальном падении луча света на границу раздела

$$R^{\perp} = (n_2 - n_1)/(n_2 + n_1), \quad (1-32)$$

Независимо от вида поляризации и от того, с какой стороны падает волна. Для некоторых практически важных в оптоэлектронике границ раздела значения R^{\perp} приведены в табл. 1.1. Коэффициент отражения изменяется от единиц до нескольких десятков процентов. Характерно, что и для лучей, отличных от нормального, но близких к нему (до $a_1 \gg a_B$, где a_B — угол Брюстера, см. ниже), $R \gg R^{\perp}$. Из табл. 1.1 видно, в частности, что если между средами 1 и 2 поместить промежуточный слой с показателем преломления $n_1 < n_{np} < n_2$, то энергия отраженной волны уменьшится. Наиболее эффективно этот промежуточный просветляющий слой действует,

$$n_{np} = \sqrt{n_1 n_2}. \quad (1-33)$$

Таблица 1.1

Оптические характеристики некоторых границ раздела

Граница раздела	Коэффициент отражения при нормальном падении луча $R_{\perp}, \%$	Угол Брюстера α_B , град	Степень поляризации при $a_1 = \alpha_B$ $P, \%$	Угол полного внутреннего отражения $\alpha_{вн}$, град
Воздух (1) – кварц (1.45)	3.4	55	6.6	44
Воздух (1) – стекло (1.52)	4.3	56	8.5	40
Оптический клей (1.52) – кремний (3.44)	15	66	62	26
Оптический клей (1.52) – арсенид галлия (3.5)	16	66	62	25
Воздух (1) – кремний (3.44)	30	74	55	17
Воздух (1) – арсенид галлия (3.5)	31	74	56	16.5

4. Из (1.29) следует, что при $(a_1 + a_2) \gg p/2$ $E_3 \approx 0$ и вектор E_3 совпадает с E_3 , т. е. отраженный свет линейно поляризован в плоскости падения луча. Физически этот вывод, составляющий содержание закона Брюстера, основан на том, что поле волны 1, проникая во вторую среду, вызывает в ней колебания электронов, параллельные вектору E_2 , и излучение ими вторичных волн в направлении, перпендикулярном их колебаниям. Вторичные волны и составляют отраженную волну 3. Понятно, что если угол между прошедшим 2 и отраженным 3 лучами составит $p/2$, то отраженный луч не будет содержать p-компонент. Условие $a_1 + a_2 = p/2$ и определяет угол Брюстера

$$a_B = \arctg(n_2/n_1). \quad (1.34)$$

Степень поляризации естественного света при выполнении условия $a_1 = a_B$

$$L = \frac{(n_1^2 + n_2^2) - 4n_2^2}{(n_1^2 + n_2^2) + 4n_2^2}. \quad (1.35)$$

Из табл. 1.1 видно, что отражение от сред с большим показателем преломления может быть использовано для эффективной поляризации естественного света. Если рассчитанная по (1.35) степень поляризации недостаточна, применяют набор плоскопараллельных пластин, при этом L приближается к 100%.

5. Если на рис. 1.5 поменять направления лучей 2 и 1 (луч 2 станет падающим, а луч 1 — преломленным), то при условии

$$a_2 = a_{п.в.о.} \cdot \arcsin(n_1/n_2), \quad (1.36)$$

наступает полное внутреннее отражение; энергия волны не переходит из среды 2 в среду 1. Детальное рассмотрение этого процесса показывает, что некоторое проникновение волны в среду 1 все же имеет место, но глубина этого проникновения (оцениваемая по уменьшению интенсивности в e раз), как правило, не превышает 0,5 λ .

Из условия (1.36) $a_{п.в.о.}$ имеет действительные значения лишь при $n_2 > n_1$, т. е. при распространении света в оптически более плотной среде. При большом различии n_2 и n_1 выйти из второй среды в первую могут лишь лучи, близкие к нормальным (см табл. 1.1).

Соотношения (1.28) – (1.36) широко используются в опто-электронике при расчете чувствительности фотоприемников, излучения светодиодов и лазеров, распространения волн по двухслойным оптическим волокнам, при изготовлении поляризаторов и модуляторов света.

Вернемся к рассмотрению безграничной изотропной среды, но учтем, что реально $\sigma \neq 0$. В этом случае в отличие от (1.7) получим

$$\Delta u = \frac{\epsilon \cdot \mu}{c^2} \frac{\partial^2 u}{\partial t^2} + \frac{4 \cdot \pi \cdot \sigma \cdot \mu}{c^2} \frac{\partial u}{\partial t}. \quad (1.37)$$

где, как и прежде, под u подразумеваются \vec{j} , A , E , H или любая их составляющая. Уравнение (1.37) описывает процесс распространения плоской электромагнитной волны с затуханием, вызываемым диссипацией энергии вследствие проводимости. Формально (1.37) может быть сведено к (1.7), если ввести комплексный волновой вектор или (что согласно (1.15) эквивалентно) комплексный показатель преломления

$$\tilde{n} = n - ig. \quad (1.38)$$

Здесь p — действительная часть, равная введенному по (1.14) показателю преломления, а g — коэффициент экстинкции. Подставляя (1.38) в (1.15) и далее в формулу плоской волны (1.8), ограничиваясь при этом одномерным случаем (1.9), получаем

$$\mathbf{E} = \mathbf{e}_y F_0 \exp\left(-\frac{\omega \cdot \gamma}{c} x\right) \cos\left(\omega t - \frac{n \cdot \omega}{c} x\right). \quad (1.39)$$

Это формула плоской волны с амплитудой, экспоненциально убывающей по мере проникновения излучения в поглощающую среду. Феноменологическое рассмотрение

ослабления интенсивности света, распространяющегося в поглощающей среде, приводит к закону Бугера – Ламберта:

$$I(x) = I_0 \exp(-C \times x), \quad (1.40)$$

где C — экспериментально измеряемый линейный показатель поглощения. Учитывая квадратичную связь I_0 с E_0 [см. формулу (1.13)], получаем

$$\chi = 2 \frac{\omega}{c} \gamma = \frac{4\pi}{n \cdot \lambda} \gamma. \quad (1.41)$$

Поскольку в эксперименте измеряется длина волны света в вакууме l_v , а не в среде l , то обычно связь линейного показателя поглощения с коэффициентом экстинкции задается в виде

$$c = 4\pi g / l_v. \quad (1.42)$$

Если комплексное значение показателя преломления (1.38) ввести в формулы, определяющие коэффициенты отражения и пропускания, то получим их обобщение для случая поглощающей среды. В частности, для нормального падения луча из вакуума ($n_1 = 1$) на поверхность поглощающей среды ($n^2 = n - i g$) имеем

$$R_{\perp} = \left| \frac{(n - i \cdot \gamma) - I}{(n - i \cdot \gamma) + I} \right|^2 = \frac{(n - I)^2 + \gamma^2}{(n + I)^2 + \gamma^2}. \quad (1.43)$$

При $g \rightarrow 0$ (1.43) переходит в (1.32), а при $g \rightarrow \infty$ $R_{\perp} \rightarrow 1$, т. е. сильно поглощающее вещество ведет себя как металлическое зеркало. Закон сохранения энергии [третье уравнение в (1.30)] в случае поглощающей среды принимает вид:

$$R + T + A = 1, \quad (1.44)$$

или в дифференциальной форме

$$r + t + a = 1, \quad (1.44a)$$

где $A(a)$ и $T(t)$ — доли энергии волны, поглощенной в слое вещества и прошедшей через этот слой.

Проанализируем роль еще одного отличия реальных потоков излучения от идеализации § 1.1 — немонохроматичность световых волн. Для этого необходимо учитывать дисперсию, т. е. зависимость фазовой скорости волны от частоты (или частотную зависимость показателя преломления). Согласно классической теории замедление и поглощение световой волны в веществе обусловлено ее взаимодействием с элементарными атомными осцилляторами этого вещества. В простейшем приближении, когда все осцилляторы одинаковы, не взаимодействуют друг с другом и имеют собственные частоты колебаний ω_0 , решение уравнения вынужденных колебаний, возбуждаемых внешней силой, изменяющейся с частотой ω , дает

$$n^2 - \gamma^2 = 1 + \frac{4\pi \cdot N \cdot q^2}{m_e} \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \omega^2 f^2}; \quad (1.45)$$

$$2n \cdot \gamma = \frac{4\pi \cdot N \cdot q^2}{m_e} \frac{\omega \cdot f}{(\omega_0^2 - \omega^2)^2 + \omega^2 f^2};$$

где N — концентрация осцилляторов, m_e и q — масса и заряд электрона, f — константа, определяемая природой осциллятора.

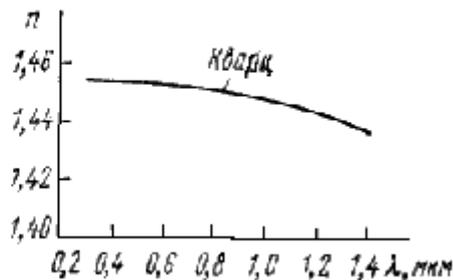
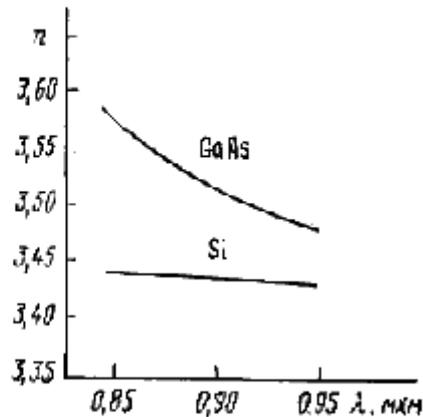
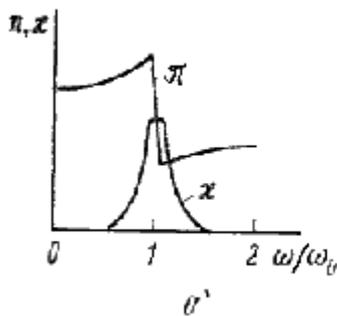


Рис. 1.7. Дисперсия оптических сред:

а — теоретические частотные зависимости показателей преломления и поглощения; **б** — экспериментальные зависимости $n(\lambda)$ для некоторых сред

Из формул (1.45) видно, что и n , и γ должны менять ход своих частотных зависимостей вблизи резонансной частоты $\omega = \omega_0$, что, как правило, и наблюдается экспериментально (рис. 1.7). Вдали от резонанса ($\omega < \omega_0$ или $\omega > \omega_0$) наблюдается нормальная дисперсия, для которой характерен рост n при увеличении ω , около $\omega = \omega_0$ — аномальная дисперсия. Для этой же области характерен и максимум поглощения. Обычно даваемые в справочниках значения n соответствуют $\omega \rightarrow 0$. Количественной мерой дисперсии служит величина $DN/D\omega$ или в пределе $dn/d\omega$, иногда ее и называют дисперсией D . Если распространяющаяся волна включает в себя группу монохроматических волн, занимающих полосу $D\omega$, то согласно (1.16) и с учетом (1.15) ее групповая скорость

$$v_g = v_\phi(1 - Dn/n) = v_\phi n g/n, \quad (1.46)$$

где $v_\phi = c/n$ — фазовая скорость, соответствующая частотному «центру» группы волн, а $ng = n - Dn$ — групповой показатель преломления.

Важнейшее практическое следствие дисперсии состоит в том, что по мере распространения в среде оптический сигнал «расплывается», т. е. происходит то же самое, что и с электрическим сигналом, распространяющимся по цепям, содержащим реактивности: искажение фронта и среза и увеличение общей длительности импульса. Оценки показывают, что уширение оптического импульса оказывается пропорциональным $l^{1/2}$, где l — расстояние, пройденное им в диспергирующей среде.

Понятие дисперсии в более общем смысле как эффекта зависимости фазовой скорости волны от волнового вектора уже встречалось при анализе модового состава излучения, распространяющегося в волноводах [см. формулу (1.25)]; отличие состоит в том, что там дисперсия обуславливалась геометрическими факторами, а не тем, что $D = Dn/D\omega \neq 0$, поэтому она имеет место и для вакуума ($n = 1$). Эта модовая дисперсия важна, в частности, в системах волоконно-оптической связи.

В заключение отметим, что имеются и другие формы неидеальности оптической среды (шероховатость границ раздела, мутность, наличие статистически распределенных неоднородностей и т. п.), но их анализ (математически более сложный) основывается, как правило, на материале, рассмотренном в данном параграфе.