



Light Transmission Optics

Dietrich Marcuse

Member of Technical Staff
Bell Telephone Laboratories
Holmdel, New Jersey

Van Nostrand Reinhold Company
New York • Cincinnati • Toronto • London • Melbourne
1972

Д. МАРКУЗЕ

ОПТИЧЕСКИЕ ВОЛНОВОДЫ

Перевод с английского

Под редакцией
В. В. ШЕВЧЕНКО

ИЗДАТЕЛЬСТВО «МИР»
МОСКВА 1974

Оптические волноводы исследуются с точки зрения их применения в качестве средств сверхширокополосной дальней связи и элементов миниатюрных быстродействующих функциональных узлов.

Данная книга является первой монографией, в которой изложены вопросы теории почти всех известных видов оптических волноводов. В ней рассмотрены линейные, непрерывные газовые, фибровые и плоские (пленочные) волноводы. Приводятся результаты как для регулярных, так и для нерегулярных (в том числе статистически нерегулярных) волноводов. В книгу включены также наиболее важные вопросы теории геометрической и волновой оптики и теории дифракции света, необходимые для понимания работы оптических волноводов.

Книга предназначена для специалистов по оптике, радиофизике и радиотехнике, занимающихся вопросами разработки новых средств передачи сигнала, а также для специалистов в области микроэлектроники и вычислительной техники, интересующихся вопросами миниатюризации и быстродействия радиоэлектронной и вычислительной аппаратуры.

Редакция литературы по новой технике

© Перевод на русский язык, «Мир», 1974

Д. Маркузе

ОПТИЧЕСКИЕ ВОЛНОВОДЫ

Редактор Л. П. Якименко

Художник Ю. С. Урманцев. Художественный редактор И. А. Шаврова
Технический редактор Т. А. Максимова. Корректор И. С. Соколова

Сдано в набор 29/I 1974 г. Подписано к печати 27/VI 1974 г. Бумага № 1
84×108¹/₃₂=9 бум. л. 30,24 усл. печ. л. Уч.-изд. л. 28,21. Изд. № 20/7179
Цена 3 р. 03 к. Зак. 087

ИЗДАТЕЛЬСТВО «МИР»

Москва, 1-й Рижский пер., 2

Ордена Трудового Красного знамени Московская типография № 7 «Искра революции» Союзполиграфпрома при Государственном комитете Совета Министров СССР по делам издательств, полиграфии и книжной торговли.
Москва, К-1, Трехпрудный пер., 9

ПРЕДИСЛОВИЕ РЕДАКТОРА РУССКОГО ИЗДАНИЯ

Книга Д. Маркузе посвящена теории волноводов оптического диапазона. В настоящее время оптические волноводы исследуются в связи с проблемой дальней световой связи и возможным их применением в так называемой интегральной оптике (создание сверхбыстродействующих миниатюрных функциональных устройств на оптических микроволноводах). Работами данной тематики занимаются ведущие фирмы ряда капиталистических стран (США, Англии, Японии, ФРГ); такие работы проводятся и в СССР.

В книге рассмотрены новые проблемы оптики, возникающие при конструировании световых линий связи, в которых для передачи сигнала без искажения необходимо сохранять пространственную структуру поля несущей волны. Подобные физические задачи не встречались при разработке обычных оптических устройств. В то же время они являются привычными для специалистов по сверхвысокочастотной радиотехнике. Поэтому и за рубежом и у нас данной тематикой занимаются в основном специалисты по сверхвысокочастотной электродинамике и волноводной технике.

Д. Маркузе является как раз таким специалистом. Он работает в одной из ведущих фирм США (Bell Telephone). Поэтому основу книги составили работы, выполненные в этой фирме.

Книга состоит из 10 глав. Первые три главы автор посвятил общей теории волновой оптики, дифракции света и геометрической оптике. Следует отметить оригинальное изложение некоторых вопросов. Так, в первой главе автор

вводит понятия направляющей среды и направляемых волн, что не типично для оптики, а в третий главе проводит сравнение теории оптических лучей с квантовой теорией. Здесь же вводится понятие аналога фазового пространства в оптике и формулируется теорема Лиувилля о постоянстве элементарного объема в фазовом пространстве, принадлежащего связке лучей. Эта теорема, как отмечает автор, является важной и полезной, когда простая физическая интуиция может привести к серьезным ошибкам. Теорема Лиувилля используется затем при рассмотрении теории тонкой линзы.

Следующие две главы посвящены теории диэлектрических и газовых линз и построенных на их основе лучевых (пучковых) волноводов. Формула тонкой линзы и ее частный вид для параксиального приближения лучей выводятся из общей формулы, вытекающей из теоремы Лиувилля. При анализе же волновых фронтов линза рассматривается как фазовый корректор. Обычно при изложении оптических свойств линзы эти подходы не применяются.

Газовые линзы, которые осуществляются путем создания в газе подходящего температурного градиента, являются новым оптическим инструментом для формирования изображения. Исходное уравнение термодинамики решается в книге для трубчатой газовой линзы в предположении осевой симметрии. На основе полученного температурного распределения рассматриваются оптические свойства газовой линзы, в частности ее фокусирующие свойства в зависимости от скорости газового потока.

Изложение теории линзового волновода начинается с рассмотрения лучевой оптики идеального линзового волновода. При этом проводится сравнение линзовых волноводов с открытыми лазерными резонаторами с фокусирующими зеркалами. Такое сравнение позволяет рассмотреть вопросы устойчивости колебаний в волноводе и резонаторе с произвольной фокальностью. В книге приводятся также результаты исследования нерегулярных линзовых волноводов с изогнутой осью и со случайными смещениями линз. Для подтверждения результатов лучевой теории линзового волновода проводится их сравнение с волновой (модовой) теорией, которая наиболее подробно изложена для конфокального волновода.

В гл. 6 рассмотрены законы распространения гауссовых пучков в свободном пространстве, а в гл. 7 — вопросы, связанные с формированием таких пучков в среде с квадратичным изменением параметров. Приводятся результаты по модовой структуре поля в среде с квадратичным изменением параметров.

Последние три главы посвящены теории оптических плоских и круглых (фибровых) диэлектрических волноводов. Рассматриваются вопросы модовой структуры направляемых волн в диэлектрических волноводах как дискретной, так и непрерывной частей спектра. Эти вопросы имеют важное значение при исследовании волноводов с нерегулярностями: переменная толщина, неровная поверхность стенки. Особое внимание уделяется в книге периодической неровности поверхности стенок вдоль оси волновода и статистической шероховатости стенок. Важным результатом является вычисление потерь волн на преобразование и излучение.

В конце книги рассмотрены связанные плоские и круглые диэлектрические волноводы и проблема перекрестной связи.

Монография адресована, как указывает автор в предисловии к английскому изданию книги, инженерам и физикам, которые интересуются средствами передачи света и пытаются с помощью методов математической физики решать технические задачи.

Данная книга является первой монографией по теории оптических волноводов, охватывающей практически все их виды. Ранее и за рубежом и в СССР публиковались книги по отдельным вопросам или определенным видам оптических волноводов. В сборнике статей [1], а также в монографиях Б. З. Каценеленбаума [2] и Л. А. Вайнштейна [3] дано изложение ряда вопросов теории линзовых волноводов. Книга В. Ф. Взятышева [4] посвящена теории диэлектрических волноводов. В монографиях Н. С. Капани [5], В. Б. Вейнберга и Д. К. Саттарова [6] и Л. М. Кучикяна [7] рассмотрены диэлектрические волноводы, поперечные размеры которых велики по сравнению с длиной волны. В таких волноводах почти не проявляется модовая структура поля направляемых волн и они не используются для средств связи. Теория и применения толстых

фибров, описанные в монографиях [5—7], существенно отличны от изложенных в данной книге. В сборнике [8] рассмотрены вопросы технологии оптических диэлектрических волноводов, в монографии [9] изложена теория конвективных газовых линз.

В заключение следует отметить, что в книге Маркузе слабо отражены результаты советских исследований по оптическим волноводам. По этой причине некоторые вопросы оказались недостаточно полно раскрытыми. Этот недостаток в настоящем издании в некоторой степени компенсируется списком дополнительной литературы по советским работам. Список не претендует на полноту, но составлен таким образом, что либо в самих работах, либо в имеющихся в них ссылках можно найти все основные результаты, относящиеся к теории оптических, а также квазиоптических волноводов.

Перевод книги выполнен Г. П. Веселковым (гл. 1, 3, 4), Ю. М. Александровым (гл. 2), В. А. Поветкиным (гл. 5—7), В. В. Черным и О. И. Черной (гл. 8—10). Список дополнительной литературы составлен Р. Б. Вагановым и В. В. Черным, предметный указатель подготовила О. И. Черная.

В. В. Шевченко

ЛИТЕРАТУРА

1. Квазиоптика, сб. статей, под ред. Б. З. Каценеленбаума и В. В. Шевченко, изд-во «Мир», 1966.
2. Каценеленбаум Б. З. Высокочастотная электродинамика, изд-во «Наука», 1966.
3. Вайнштейн Л. А. Открытые резонаторы и открытые волноводы, изд-во «Советское радио», 1966.
4. Взятых В. Ф. Диэлектрические волноводы, изд-во «Советское радио», 1970.
5. Капани Н. С. Волоконная оптика, изд-во «Мир», 1969.
6. Вейнберг В. Б., Саттаров Д. К. Оптика световодов, изд-во «Машиностроение», 1969.
7. Кучикян Л. М. Световоды, изд-во «Энергия», 1973.
8. Оптические волокна и волоконные элементы, сб. статей, под ред. К. И. Блох, изд-во «Химия», 1972.
9. Мартыненко О. Г., Колесников П. М., Колпащиков В. Л. Введение в теорию конвективных газовых линз, изд-во «Наука и техника», 1972.

ПРЕДИСЛОВИЕ АВТОРА К РУССКОМУ ИЗДАНИЮ

С большим удовольствием я принял предложение В. В. Шевченко написать предисловие к русскому изданию моей книги. Как я уже отметил в предисловии к английскому изданию, в книге описываются физические принципы и аналитические методы решения задач распространения света по оптическим волноводам разного типа. Первые три главы имеют вводный характер и содержат элементы электромагнитной теории и в первую очередь приближенные методы, являющиеся весьма полезными в области малых длин волн, которые мы называем оптическими волнами. Большая часть этого вводного материала является традиционной. Например, уравнения лучевой оптики выводятся из скалярного волнового уравнения обычным способом. Однако в книге рассматривается и обратная задача: волновое уравнение выводится из уравнений геометрической оптики с применением методов квантовой механики к лучевой теории. Другой необычной чертой моего изложения лучевой оптики является использование теоремы Лиувилля, взятой из статистической механики.

Рассмотрение оптических волноводов концентрируется на вопросах распространения света вдоль линзовых систем, диэлектрической среды с параболическим законом изменения показателя преломления и по диэлектрическим волноводам. Исследуются как идеальные структуры, так и волноводы с нерегулярностями.

Поскольку цель книги изложена в предисловии к английскому изданию, то здесь я дополнительно хотел бы сделать только несколько замечаний относительно истории появления этой книги. Английское издание опубликовано в серии Bell Laboratories, издаваемой Van Nost-

rand Reinhold Book Company. Материал книги основывался главным образом на работах автора и его коллег, проведенных в фирме Bell Laboratories. Следовательно, в ней не нашло отражения большое число работ других исследователей, выполненных как в Соединенных Штатах, так и в Советском Союзе. Я писал книгу в свободное время. Сначала в первом варианте книга была написана с намерением издать учебник для студентов университета. Позже она была переработана в соответствии с пожеланием американского издательства опубликовать книгу для инженеров.

При подготовке русского издания я использовал возможность исправить некоторые опечатки (те, о которых мне стало известно к моменту написания этого предисловия), оставшиеся в книге даже после тщательной правки корректуры.

В заключение хотел бы выразить мою искреннюю благодарность редактору перевода В. В. Шевченко, переводчикам и издательству «Мир» за их интерес к моей книге, за быструю и квалифицированную работу, сделавшую возможной столь скорую публикацию книги на русском языке.

Д. Маркузе

ПРЕДИСЛОВИЕ

Изучение света и особенностей его распространения не является новой областью исследования. Но с изобретением лазера резко возрос интерес к свету как носителю информации. Цель настоящей книги — изложение физических основ, необходимых для понимания прохождения света через различные направляющие среды, и методов передачи света применительно к системам оптической связи.

Проблема передачи света настолько обширна и разнообразна, что пришлось произвести определенный отбор материала и опустить некоторые детали, которые могли бы, возможно, оказаться существенными для некоторых читателей. Отбор материала обычно с неизбежностью отражает личные интересы и опыт автора. Настоящая книга в этом смысле не является исключением. На отборе материала сильно сказалось то, что автор многие годы работал в области передачи света по различным оптическим волноводам. В то же время я полагаю, что приведенный материал отражает современное состояние предмета и послужит для читателя хорошим введением в теорию направленной передачи и распространения света. Рассмотрение материала проводится аналитическими методами. Особое внимание уделяется математическим методам, используемым для решения задач, связанных с направляемыми волнами. В качестве математического аппарата, мне кажется, лучше использовать метод нормальных волн, хотя некоторые авторы предпочитают методы, основанные на функции Грина.

Книга предназначена для инженеров и физиков, которые интересуются методами передачи света и хотят освоить математические методы решения задач, связанных с передачей света. Можно надеяться также, что читатель найдет для себя полезные сведения о свойствах световодов,

оптических волокон и методах передачи света с помощью гауссовых пучков. Эту книгу я рассматриваю как учебное пособие по курсу оптики. В настоящее время оптика не является дисциплиной, изучаемой на всех технических факультетах. Но быстро растущий интерес к оптической связи и блестящие перспективы использования света в качестве носителя информации позволяют надеяться, что курс оптики войдет в программу всех технических факультетов. Эта книга должна восполнить пробел в учебниках по технической оптике.

Книга начинается с введения в волновую оптику, являющуюся основой при изучении передачи света. Введение довольно краткое, так как этот вопрос рассматривается во многих книгах по оптике и электромагнетизму. Во второй главе излагается скалярная теория дифракции, необходимая для понимания работы линзовых волноводов. Особое внимание уделяется приложению этой теории к анализу прохождения света через апертуры, к дифракционным решеткам и дифракции Брэгга.

Геометрическая оптика рассматривается как предельный случай волновой оптики, поэтому мы начинаем ее рассмотрение с вывода уравнения луча из уравнения эйконала, которое в свою очередь получается из скалярного волнового уравнения. Затем строится лучевая теория из принципа Ферма в формулировке Гамильтона. Такой подход особенно удобен для выявления связи между лучевой оптикой и волновой оптикой при использовании квантования в лучевой теории. Оригинальным является использование при рассмотрении геометрической оптики теоремы Лиувилля, взятой из статистической механики.

На основании свойств тонких линз проводится анализ работы линзового волновода и обсуждаются его свойства с точки зрения волновой и лучевой оптики. Рассматриваются статистические смещения линзы, а также аналогия между линзовыми волноводами и резонаторами лазеров.

Анализ линзовых волноводов естественным образом связан с гауссовыми пучками. Распространение этих пучков в свободном пространстве и трансформация их при прохождении через линзы исследуются в шестой главе. Седьмая глава посвящена вопросам распространения света в среде с квадратичным изменением показателя пре-

ломления и опять-таки относится к гауссовым пучкам и линзовым волноводам.

В остальных трех главах книги рассматриваются диэлектрические волноводы и оптические волокна. Основное внимание уделяется нормальным модам в диэлектрических волноводах. Изучаются как направляемые моды, так и моды излучения. Развитая теория используется для исследования излучения и преобразования типов волн из-за несовершенства диэлектрического волновода. Рассматривается влияние систематических и случайных искажений диаметра сердцевины оптического волокна, а также радиационные потери, обусловленные изгибом.

В последней главе излагается теория связанных диэлектрических волноводов. Выводится уравнение связанных волн. Выражения для коэффициентов связи являются достаточно общими и справедливы для любых диэлектрических волноводов, выполненных из материала с потерями. Эта теория используется для анализа перекрестной связи между двумя параллельными оптическими волокнами.

Материал книги в значительной мере новый и в других книгах ранее не рассматривался. Однако данная книга не исчерпывает всех сторон рассматриваемой проблемы: для этого потребовался бы больший объем изложения. Как отмечалось выше, на отбор материала и глубину изложения оказали влияние интересы и опыт автора, а также ограничения, связанные с объемом книги.

Д. Маркузе

ВОЛНОВАЯ ОПТИКА

1.1. ВВЕДЕНИЕ

Свет представляет собой электромагнитное явление. Поэтому оптика должна быть просто разделом электродинамики. По историческим причинам все же оптика обычно рассматривается как самостоятельный предмет, поскольку световые явления изучались задолго до того, как была установлена электромагнитная природа света. Существенным обстоятельством является то, что мы ощущаем свет глазами, тогда как электромагнитные колебания частот, отличных от оптических, могут обнаруживаться только с помощью специальных устройств.

Другой особенностью света является очень короткая длина волны, что позволяет использовать методы приближенного анализа. На более длинных волнах это сделать нельзя. В связи с этим в оптике используются два различных метода приближения. Волновая оптика непосредственно основывается на уравнениях Максвелла, тогда как лучевая оптика использует малость длины волны света для упрощения многих задач, связанных с распространением света. Лучевая оптика во многих отношениях подобна механике материальной точки, а волновая оптика близка к квантовой теории световых лучей. Этот вопрос будет подробно рассмотрен в главе, посвященной геометрической оптике.

В настоящей главе запишем уравнения Максвелла и выведем из них волновое уравнение. Затем применим эти уравнения для решения некоторых характерных и важных задач. Эти задачи волновой оптики выбраны так, чтобы облегчить понимание последующих глав. Волновая оптика является обширным и хорошо разработанным разделом оптики и излагается во многих учебниках [1]. Поэтому мы не будем стремиться к полному изло-

жению волновой оптики, а ограничимся рассмотрением здесь основных свойств отражения и преломления света на границе между двумя диэлектрическими средами и в следующей главе обсудим основные задачи дифракции.

1.2. УРАВНЕНИЯ МАКСВЕЛЛА

Уравнения Максвелла подробно рассмотрены в ряде учебников по электромагнетизму (см., например, [2] и [3]), поэтому здесь достаточно лишь сформулировать их. Вектор напряженности электрического поля \mathbf{E} и вектор электрической индукции \mathbf{D} связаны с вектором напряженности магнитного поля \mathbf{H} и вектором индукции \mathbf{B} следующими уравнениями ¹⁾:

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} \quad (1.2.1)$$

и

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}. \quad (1.2.2)$$

Член с током в первом уравнении Максвелла (1.2.1) опущен, так как возбуждение света чрезвычайно редко описывают посредством токов.

Электрические векторы связаны друг с другом. В общем случае эта связь может оказаться весьма сложной, например тензорного или даже нелинейного характера. Однако для многих практически интересных случаев можно предполагать простое линейное соотношение

$$\mathbf{D} = \varepsilon \mathbf{E}, \quad (1.2.3)$$

которое справедливо для линейной изотропной среды. Константа ε есть диэлектрическая проницаемость. Отношение $\varepsilon/\varepsilon_0$, где ε_0 — проницаемость вакуума, называют относительной диэлектрической проницаемостью. Связь между \mathbf{H} и \mathbf{B} задается аналогичным соотношением

$$\mathbf{B} = \mu \mathbf{H}. \quad (1.2.4)$$

Величина μ известна как магнитная проницаемость; у немагнитных материалов ее величина очень близка к проницаемости вакуума μ_0 .

¹⁾ Оператор ∇ является вектором с компонентами $\partial/\partial x$, $\partial/\partial y$, $\partial/\partial z$.

При отсутствии электрических зарядов вектор \mathbf{D} удовлетворяет уравнению

$$\nabla \cdot \mathbf{D} = 0. \quad (1.2.5)$$

Вектор же магнитной индукции всегда удовлетворяет уравнению

$$\nabla \cdot \mathbf{B} = 0 \quad (1.2.6)$$

независимо от наличия или отсутствия электрических зарядов.

Уравнения (1.2.1)–(1.2.6) полностью описывают электромагнитное поле в линейной изотропной среде при отсутствии токов и свободных зарядов.

Очень важным является вектор плотности потока мощности

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}, \quad (1.2.7)$$

который называется также вектором Пойнтинга. Он описывает поток электромагнитной мощности в пространстве. Для того чтобы получить мощность, проходящую через поверхность A с внешним единичным вектором нормали \mathbf{n} в каждой точке, нужно вычислить поверхностный интеграл

$$P = \int_A \mathbf{S} \cdot \mathbf{n} dA. \quad (1.2.8)$$

Под световыми лучами во многих случаях можно понимать линии в пространстве, вдоль которых узким пучком распространяется электромагнитная энергия света.

Мы в основном будем иметь дело с монохроматическими полями, совершающими колебания с одной определенной частотой f . В этом случае удобно воспользоваться обозначениями, при которых компоненты электрического и магнитного полей могут быть выражены соотношениями следующего вида:

$$F(x, y, z, t) = \operatorname{Re} [G(x, y, z) e^{i\omega t}], \quad (1.2.9)$$

где круговая частота ω определена как

$$\omega = 2\pi f, \quad (1.2.10)$$

а $G(x, y, z)$ — комплексная функция вещественных переменных x, y, z . Символ $\operatorname{Re} [\]$ означает, что берется ве-

вещественная часть от выражения в скобках. Мы всегда будем опускать символ $\text{Re}[\]$ в уравнениях, даже если они должны быть интерпретированы как (1.2.9), подразумевая при этом, что только вещественная часть величины имеет физический смысл. Таким образом, будем просто писать

$$F(x, y, z, t) = G(x, y, z) e^{i\omega t}. \quad (1.2.11)$$

С помощью такой комплексной записи вектор Пойнтинга (1.2.7) представится в виде

$$\bar{\mathbf{S}} = \frac{1}{2} (\mathbf{E} \times \mathbf{H}^*). \quad (1.2.12)$$

Звездочка означает комплексно сопряженную величину. Множитель $1/2$ появляется в (1.2.12) при вычислении среднего по времени значения вектора Пойнтинга, на что указывает черточка сверху. Вещественная часть от выражения (1.2.12) представляет собой физический средний во времени вектор потока мощности.

1.3. ВОЛНОВОЕ УРАВНЕНИЕ

Из уравнений Максвелла многими способами можно получать уравнения, удобные для тех или иных конкретных приложений. Например, подставим величину (1.2.4) в уравнение (1.2.2) и возьмем ротор от этого уравнения. В результате получим

$$\nabla \times (\nabla \times \mathbf{E}) = -\mu \frac{\partial}{\partial t} (\nabla \times \mathbf{H}). \quad (1.3.1)$$

При выводе уравнения (1.3.1) мы предположили, что величина μ не зависит от координат. Подстановка (1.2.1) и (1.2.3) в (1.3.1) приводит к уравнению, содержащему только вектор \mathbf{E} :

$$\nabla \times (\nabla \times \mathbf{E}) + \epsilon \mu \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0. \quad (1.3.2)$$

Это уравнение справедливо и в том случае, если ϵ изменяется в пространстве. Ввиду того что оператор $\nabla \times \nabla \times$ не очень удобен для применения, целесообразно воспользоваться векторным тождеством

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla (\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}, \quad (1.3.3)$$

которое справедливо в декартовой системе координат. Использование соотношений (1.2.3) и (1.2.5) дает возможность переписать уравнение (1.3.2) в следующем виде:

$$\nabla^2 \mathbf{E} + \nabla \left(\mathbf{E} \cdot \frac{\nabla \varepsilon}{\varepsilon} \right) = \varepsilon \mu \frac{\partial^2 \mathbf{E}}{\partial t^2}. \quad (1.3.4)$$

В частном случае, когда величина ε постоянна в пространстве, градиент ε обращается в нуль и уравнение (1.3.4) принимает вид волнового уравнения

$$\nabla^2 \mathbf{E} = \varepsilon \mu \frac{\partial^2 \mathbf{E}}{\partial t^2}. \quad (1.3.5)$$

Волновое уравнение (1.3.5) справедливо для каждой декартовой компоненты вектора электрического поля, т. е. каждая его декартова компонента удовлетворяет скалярному волновому уравнению

$$\nabla^2 \psi = \frac{1}{v^2} \frac{\partial^2 \psi}{\partial t^2}, \quad (1.3.6)$$

где величина

$$v = (\varepsilon \mu)^{-1/2} \quad (1.3.7)$$

имеет физический смысл скорости света в среде с диэлектрической проницаемостью $\varepsilon/\varepsilon_0$.

Волновое уравнение (1.3.6) приближенно удовлетворяется для каждой компоненты вектора электрического поля даже в случае, когда ε изменяется в пространстве, при условии, что ее изменения незначительны на расстоянии порядка длины волны света. К этому мы еще вернемся несколько позже.

Смысл волнового уравнения легко понять, если учесть, что любая функция вида

$$\psi = f \left(t - \frac{1}{v} \mathbf{n} \cdot \mathbf{r} \right) \quad (1.3.8)$$

есть решение этого уравнения при условии, что существует вторая производная от f . Компоненты вектора \mathbf{r} являются координатами точки наблюдения поля; \mathbf{n} — единичный вектор. В решении (1.3.8) волнового уравнения (1.3.6) скорость v не должна зависеть от частоты.

Решение (1.3.8) волнового уравнения представляет собой плоские волны, распространяющиеся в простран-

стве, однородно заполненном средой с диэлектрической проницаемостью $\varepsilon/\varepsilon_0$. Для того чтобы убедиться, что функция (1.3.8) описывает плоские волны, нужно рассмотреть некоторое фиксированное значение аргумента

$$u = t - \frac{1}{v} \mathbf{n} \cdot \mathbf{r}. \quad (1.3.9)$$

При любом данном значении u функция f имеет соответствующее фиксированное значение $f(u)$. Величина $u = \text{const}$ при фиксированном значении времени t реализуется в плоскости, определяемой соотношением $\mathbf{n} \cdot \mathbf{r} = \text{const}$. Вектор \mathbf{n} направлен перпендикулярно плоскости. Поэтому функция имеет одно и то же значение в пространстве именно на этой бесконечной плоскости. Для того чтобы посмотреть, каким образом это фиксированное значение функции ведет себя с течением времени, предположим, что t изменилось на Δt , вектор \mathbf{r} — на $\Delta \mathbf{r}$, но так, что величина u осталась неизменной. Связь между приращением времени и изменением положения вектора при фиксированном значении u задается соотношением

$$\mathbf{n} \cdot \Delta \mathbf{r} = v \Delta t. \quad (1.3.10)$$

Конец вектора $\Delta \mathbf{r}$ снова лежит на плоскости. Вектор \mathbf{n} , очевидно, является нормалью как к первоначальной, так и к смещенной плоскостям. Плоскость, описываемая формулой (1.3.9), перемещается в направлении \mathbf{n} на расстояние $v \Delta t$ за время Δt . Отсюда следует, что плоскость движется в пространстве со скоростью v . Уравнение (1.3.8), таким образом, описывает плоское волновое возмущение, движущееся со скоростью v . Вид функции $f(u)$ произволен. Изменяя знак v в (1.3.8), получаем другое решение волнового уравнения. Легко видеть, что функция

$$\psi = f\left(t + \frac{1}{v} \mathbf{n} \cdot \mathbf{r}\right)$$

представляет плоскую волну, распространяющуюся в направлении — \mathbf{n} .

Весьма важным частным случаем являются такие решения волнового уравнения, которые в каждой точке пространства изменяются во времени по синусоидальному

закону. Такую плоскую волну можно представить в виде

$$g = A \cos \left(\omega t - \frac{\omega}{v} \mathbf{n} \cdot \mathbf{r} \right).$$

Частота колебаний

$$\nu = \frac{\omega}{2\pi},$$

где ω — так называемая круговая частота. Иногда будем называть ω просто частотой, подразумевая при этом, что она является умноженной на 2π истинной частотой колебаний. Удобно ввести вектор

$$\mathbf{k} = \frac{\omega}{v} \mathbf{n} \quad (1.3.11)$$

и записать

$$g = A \cos (\omega t - \mathbf{k} \cdot \mathbf{r}). \quad (1.3.12)$$

Назовем \mathbf{k} волновым вектором. Если допустить, что \mathbf{r} получает приращение

$$\Delta \mathbf{r} = \lambda \mathbf{n},$$

то из требования, чтобы функция (1.3.12) совершила полный цикл своего изменения при изменении \mathbf{r} до величины $\mathbf{r} + \Delta \mathbf{r}$, найдем соотношение

$$k\lambda = 2\pi,$$

или

$$k = \frac{2\pi}{\lambda} = \frac{\omega}{v} = \omega \sqrt{\epsilon\mu}, \quad (1.3.13)$$

где

$$k = |\mathbf{k}| \quad (1.3.14)$$

есть модуль волнового вектора. Правая часть уравнения (1.3.13) следует из (1.3.11) и (1.3.7).

Большое физическое значение волн вида (1.3.12) обусловлено тем фактом, что в большинстве сред, за исключением вакуума, v не является константой, а зависит от частоты:

$$v = v(\omega).$$

Синусоидальные волны разных частот распространяются с разными фазовыми скоростями. Это явление называется дисперсией. Выражение общего вида (1.3.8) для плоской

волны произвольной формы справедливо поэтому только в вакууме. В другой среде это выражение можно использовать лишь как разумное приближение, если дисперсия не слишком явно выражена. Легко можно предсказать путь, по которому распространяется общее возмущение в диспергирующей среде. Допустим, что на некоторой плоскости существует возмущение вида $f(t)$. Для того чтобы определить, как это возмущение будет распространяться в диспергирующей среде, нужно представить произвольную функцию f в виде суперпозиции синусоидальных колебаний. Это осуществляется с помощью интегрального преобразования Фурье функции $f(t)$. Каждое гармоническое колебание, согласно (1.3.12), распространяется через диспергирующую среду как плоская волна (поскольку с самого начала возмущение предполагалось плоским). Располагая систему координат ради удобства таким образом, чтобы волна распространялась в направлении оси z , мы выразим форму плоской волны общего вида, имеющей вид $f(t)$ при $z = 0$, с помощью интеграла Фурье

$$f(z, t) = \frac{1}{\pi} \int_0^{\infty} h(\omega) \cos(\omega t - kz + \theta(\omega)) d\omega. \quad (1.3.15)$$

Вводя комплексную функцию

$$\varphi(\omega) = h(\omega) e^{i\theta(\omega)} \quad (1.3.16)$$

и распространяя определение фазы и амплитуды на отрицательные частоты как

$$\theta(-\omega) = -\theta(\omega) \quad (1.3.17)$$

и

$$h(-\omega) = h(\omega) \quad (1.3.18)$$

(k также изменяет знак), можно переписать выражение (1.3.15) в виде комплексного интеграла Фурье

$$f(z, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi(\omega) e^{i(\omega t - kz)} d\omega. \quad (1.3.19)$$

Амплитудная функция $\varphi(\omega)$ определяется через известную форму волны при $z = 0$:

$$\varphi(\omega) = \int_{-\infty}^{\infty} f(0, t) e^{-i\omega t} dt. \quad (1.3.20)$$

Выражение (1.3.19) не является решением волнового уравнения, если v зависит от частоты, поскольку волновое уравнение имеет физический смысл, лишь когда нет дисперсии (т. е. когда v не зависит от частоты) или когда функция $f(z, t)$ описывается очень узким спектром частот. Фазовая скорость v в формуле (1.3.6) не имеет смысла для функции типа (1.3.19). Однако (1.3.19) правильно описывает распространение плоской волны общего вида в диспергирующей среде.

Нетрудно распространить концепцию полей, распространяющихся в диспергирующей среде, на волновые возмущения неплоской структуры. Чтобы сделать это, введем вновь плоскую синусоидальную волну

$$g(x, y, z, t) = \varphi(k_x, k_y, \omega) \exp[i(\omega t - \mathbf{k} \cdot \mathbf{r})], \quad (1.3.21)$$

распространяющуюся в направлении вектора \mathbf{k} . Здесь использованы комплексные обозначения. Вещественная часть от (1.3.21) описывает физическую плоскую волну. Амплитудный множитель φ может зависеть от независимых переменных k_x, k_y и ω . Поскольку абсолютная величина k вектора \mathbf{k} должна удовлетворять соотношению (1.3.13), можно выразить k_z через независимые переменные:

$$k_z = \sqrt{\left(\frac{\omega}{v}\right)^2 - k_x^2 - k_y^2}. \quad (1.3.22)$$

Из этой формулы видно, что z -составляющая (или любая другая компонента k_x или k_y , если предпочтительнее считать z -составляющую независимой переменной) вектора \mathbf{k} будет мнимой, когда подкоренное выражение становится отрицательным. В этом случае вместо плоской волны мы имеем дело с нераспространяющейся (локальной) волной. Нераспространяющиеся волны, таким образом, также являются решениями волнового уравнения. Используя суперпозицию волн вида (1.3.21) со всевозможными частотами, бегущих во всевозможных направлениях,

можно составить общее выражение для волны, распространяющейся в среде с дисперсией:

$$f(x, y, z, t) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} dk_x \int_{-\infty}^{\infty} dk_y \varphi(k_x, k_y, \omega) \times \\ \times \exp[i(\omega t - k_x x - k_y y - k_z z)]. \quad (1.3.23)$$

Здесь z -составляющая волнового вектора определяется формулой (1.3.22). Используя вещественные функции, можно также написать

$$f(x, y, z, t) = \frac{1}{4\pi^3} \int_0^{\infty} d\omega \int_{-\infty}^{\infty} dk_x \int_{-\infty}^{\infty} dk_y |\varphi(k_x, k_y, \omega)| \times \\ \times \cos(\omega t - k_x x - k_y y - k_z z + \theta). \quad (1.3.24)$$

Это интегральное представление более общей волны состоит не только из синусоидальных плоских волн всевозможных частот, распространяющихся по всевозможным направлениям, но также и из нераспространяющихся волн. При гармонической (синусоидальной) зависимости решения волнового уравнения от времени интегрирование по ω можно из формулы (1.3.24) исключить.

Существование нераспространяющихся волн в свободном пространстве может показаться неожиданным, поскольку обычно под нераспространяющимися волнами понимают нечто присущее волноводам, когда их частота ниже критической. Нераспространяющиеся волны, встречающиеся в наших рассуждениях, полностью обусловлены явлением полного внутреннего отражения волны, пытающейся проникнуть из среды с высокой диэлектрической постоянной в другую среду, обладающую более низкой диэлектрической постоянной. Чтобы понять это явление, рассмотрим гармоническую плоскую волну (1.3.21). Она обладает длиной волны λ , определяемой формулой (1.3.13), и распространяется в пространстве в направлении, задаваемом компонентами вектора \mathbf{k} . Для простоты предположим, что волна распространяется в плоскости x, z , так что $k_y = 0$. Теперь начнем поворачивать вектор \mathbf{k} по направлению к оси x . В пределе при $k_z = 0$ волна пойдет параллельно оси x . Ее синусоидальное изменение в пространстве по направлению x имеет пространственный пери-

од λ . Однако математический аппарат допускает выбор $k_x > 2\pi/\lambda$. С точки зрения физики это означает, что мы вынуждаем поле изменяться с пространственным периодом, меньшим λ . Это действительно можно сделать, но поле противодействует нашим усилиям путем сжатия в направлении оси z . Невозможно вынудить поле к пространственным колебаниям, более быстрым, чем колебания с длиной волны в свободном пространстве на рабочей частоте. Это рассмотрение показывает, что нераспространяющиеся волны встречаются всякий раз, когда мы накладываем на поле пространственные вариации, которые являются более быстрыми, чем это совместимо с распространением в свободном пространстве гармонической плоской волны. Как известно, именно это имеет место при полном внутреннем отражении на границе двух сред.

До сих пор обсуждение касалось волнового уравнения (1.3.6). Однако волновое уравнение было получено как частный случай уравнения (1.3.4). Необходимо исследовать, при каких условиях волновое уравнение является достаточно хорошим приближением к уравнению (1.3.4), так как с последним уравнением гораздо труднее обращаться и оно почти не применяется при практических вычислениях. Весьма удачным является то обстоятельство, что в большинстве приложений, встречающихся в оптике, можно использовать простое волновое уравнение, даже если условие его справедливости [равенство нулю второго члена в уравнении (1.3.4)] не вполне выполняется.

Среди членов уравнения (1.3.4) доминирующими являются первый член в левой части и член в правой части, порядок величин которых одинаковый. Нижеследующий анализ относится лишь к оценке порядка величин и не претендует на строгость. Член в правой части уравнения (1.3.4) по порядку величины равен

$$\epsilon\mu \frac{\partial^2 \mathbf{E}}{\partial t^2} = \omega^2 \epsilon\mu \mathbf{E} = \frac{\omega^2}{v^2} \mathbf{E} = \left(\frac{2\pi}{\lambda}\right)^2 \mathbf{E}. \quad (1.3.25)$$

Заменяя оператор ∇ производной по некоторому направлению S в пространстве, можно записать

$$\nabla \left(\mathbf{E} \cdot \frac{\nabla \epsilon}{\epsilon} \right) \approx \frac{\partial}{\partial S} \left(\mathbf{E} \cdot \frac{\nabla \epsilon}{\epsilon} \right) \approx \frac{2\pi}{\lambda} \mathbf{E} \cdot \frac{\nabla \epsilon}{\epsilon} + \mathbf{E} \cdot \frac{\partial}{\partial S} \left(\frac{\nabla \epsilon}{\epsilon} \right). \quad (1.3.26)$$

Интересным является случай, когда второй член в правой части (1.3.26) много меньше первого. Сравним два рассматриваемых члена уравнения (1.3.4):

$$R = \frac{\left[\nabla \left(\mathbf{E} \cdot \frac{\nabla \varepsilon}{\varepsilon} \right) \right]}{\left[\varepsilon_{\mu} \frac{\partial^2 E}{\partial t^2} \right]} \approx \frac{\frac{2\pi}{\lambda} \frac{\nabla \varepsilon}{\varepsilon}}{\left(\frac{2\pi}{\lambda} \right)^2} = \\ = \frac{1}{2\pi} \lambda \frac{\nabla \varepsilon}{\varepsilon} \approx \frac{1}{2\pi} \lambda \frac{\varepsilon_2 - \varepsilon_1}{\varepsilon \Delta S}. \quad (1.3.27)$$

Здесь квадратные скобки указывают на то, что рассматривается только порядок соответствующей величины. Считается, что градиент ε по порядку величины равен отношению разности $\varepsilon_2 - \varepsilon_1$ диэлектрических проницаемостей в двух близко расположенных точках к расстоянию ΔS между ними.

В качестве последнего шага выберем $\Delta S = \lambda$ и получим

$$R = \frac{1}{2\pi} \frac{\varepsilon_2 - \varepsilon_1}{\varepsilon}. \quad (1.3.28)$$

Чтобы можно было пренебречь вторым членом левой части уравнения (1.3.4), нужно потребовать выполнения условия $R \ll 1$. Как видно из соотношения (1.3.28), это означает, что относительное изменение ε на расстоянии одной длины волны должно быть много меньше единицы. Это условие часто выполняется в оптически неоднородных средах. Мы убедимся ниже в том, что условие $R \ll 1$ выполняется для случая оптических волноводов, у которых диэлектрическая проницаемость непрерывно изменяется, так что можно ограничиться решением волнового уравнения вместо гораздо более сложного уравнения (1.3.4). Только на границе раздела двух областей с различными диэлектрическими проницаемостями величина (1.3.28) почти равна или даже больше единицы. Большинство оптических приборов включает области, такие, как воздух, и материалы, диэлектрическая проницаемость которых отлична от диэлектрической проницаемости воздуха, но опять-таки постоянна. Например, на границе между стеклянными линзами и воздухом величина (1.3.28) большая. Однако даже в этих случаях требуется решать лишь волновое уравнение, так как оно справедливо всюду, кроме

границы раздела сред. При этом решают волновое уравнение в различных однородных областях и сшивают эти решения посредством граничных условий. Граничные условия будут рассматриваться ниже.

Проверим, насколько обоснованным было пренебрежение вторым членом в правой части (1.3.26). Порядок величины пренебрегаемого члена относительно оставляемого члена есть

$$\frac{\left[\mathbf{E} \cdot \frac{\partial}{\partial S} \left(\frac{\nabla \varepsilon}{\varepsilon} \right) \right]}{\left[\frac{2\pi}{\lambda} \mathbf{E} \cdot \frac{\nabla \varepsilon}{\varepsilon} \right]} \approx \frac{1}{2\pi} \lambda \frac{\left(\frac{|\nabla \varepsilon|}{\varepsilon} \right)_2 - \left(\frac{|\nabla \varepsilon|}{\varepsilon} \right)_1}{\frac{|\nabla \varepsilon|}{\varepsilon} \nabla S}. \quad (1.3.29)$$

Нет необходимости требовать, чтобы отношение (1.3.29) на самом деле было много меньше единицы. Вторым членом в (1.3.4) можно пренебречь и в том случае, когда оба слагаемых в (1.3.26) имеют одинаковый порядок величины. Для этого необходимо потребовать, чтобы изменение градиента ε на расстоянии одной длины волны примерно равнялось или было меньше самого градиента. Это требование выполняется при условиях, когда $R \ll 1$.

Приведенные аргументы показывают, что волновое уравнение можно использовать даже тогда, когда величина ε не остается постоянной, а меняется в пространстве, но ее изменения должны быть незначительными на расстоянии в одну длину волны света. За исключением граничных поверхностей между двумя различными диэлектрическими средами, это условие почти всегда выполняется. Таким образом, распространение света в неоднородной среде можно исследовать путем решения волнового уравнения. Разница между волновым уравнением и более точным уравнением (1.3.4) в большинстве практически интересных случаев незначительна.

1.4. ОПТИЧЕСКИЕ СИСТЕМЫ

С ЦИЛИНДРИЧЕСКОЙ СИММЕТРИЕЙ

Общее решение (1.3.24) волнового уравнения было построено для случая, когда ε и μ постоянны в пространстве. Везде ниже опять будем предполагать, что величина μ постоянна. Но если ε является функцией пространствен-

ных координат, то суперпозиция плоских волн (1.3.24) уже не будет решением волнового уравнения. Оказывается, что все-таки и в этом случае можно построить общее решение из более простых решений. Такой подход известен как метод нормальных мод. Каждая плоская волна в предыдущем разделе может рассматриваться как мода некоторой структуры. Термин «мода» не всегда удается легко определить. Можно определить моду как собственное решение уравнений Максвелла, соответствующее некоторому собственному значению и удовлетворяющее всем граничным условиям задачи. Плоские волны, рассмотренные в предыдущем разделе, удовлетворяют всем этим требованиям. Векторы

$$\mathbf{E} = A\mathbf{e} \exp [i (\omega t - \mathbf{k} \cdot \mathbf{r})], \quad (1.4.1)$$

$$\mathbf{H} = B\mathbf{h} \exp [i (\omega t - \mathbf{k} \cdot \mathbf{r})] \quad (1.4.2)$$

с комплексными коэффициентами A и B и единичными векторами \mathbf{e} и \mathbf{h} удовлетворяют уравнениям Максвелла, если удовлетворяются уравнения

$$-i (\mathbf{k} \times \mathbf{h}) B = i\omega\epsilon A\mathbf{e}, \quad (1.4.3)$$

$$-i (\mathbf{k} \times \mathbf{e}) A = -i\omega\mu B\mathbf{h}, \quad (1.4.4)$$

полученные после подстановки выражений (1.4.1) и (1.4.2) в соотношения (1.2.1) и (1.2.2). Вводя единичный вектор \mathbf{n} по (1.3.11), находим, что равенства (1.4.3) и (1.4.4) удовлетворяются, если имеют место следующие соотношения:

$$\mathbf{n} \cdot \mathbf{e} = 0, \quad (1.4.5)$$

$$\mathbf{h} = \mathbf{n} \times \mathbf{e}, \quad (1.4.6)$$

$$B = \sqrt{\frac{\epsilon}{\mu}} A. \quad (1.4.7)$$

Из уравнения (1.4.5) видно, что мы имеем дело с поперечными волнами. Поскольку в этом случае нет граничных условий, плоские волны (1.4.1) и (1.4.2) можно рассматривать как моды в соответствии с нашим определением. Между прочим, каждая компонента векторов \mathbf{E} и \mathbf{H} удовлетворяет волновому уравнению. Составляя для каждой компоненты суперпозицию волн, получим общее решение уравнений Максвелла в однородной среде. Если же,

однако, ε не остается постоянной в пространстве, то (1.4.1) и (1.4.2) не будут решениями ни волнового уравнения, ни уравнений Максвелла.

Рассмотрим решение уравнений Максвелла в неоднородной среде для частного случая, представляющего для нас особый интерес. Часто ε не зависит от одной пространственной координаты. Без ограничения общности можно считать, что диэлектрическая проницаемость не зависит от координаты z .

Будем искать модовые решения в виде

$$\mathbf{E} = \mathbf{E}_0(x, y) \exp[i(\omega t - \beta z)], \quad (1.4.8)$$

$$\mathbf{H} = \mathbf{H}_0(x, y) \exp[i(\omega t - \beta z)]. \quad (1.4.9)$$

После подстановки этих выражений в (1.2.1) и (1.2.2) и использования соотношений (1.2.3) и (1.2.4) приходим к следующим уравнениям:

$$\frac{\partial H_z}{\partial y} + i\beta H_y = i\omega\varepsilon E_x, \quad (1.4.10)$$

$$-i\beta H_x - \frac{\partial H_z}{\partial x} = i\omega\varepsilon E_y, \quad (1.4.11)$$

$$\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} = i\omega\varepsilon E_z, \quad (1.4.12)$$

$$\frac{\partial E_z}{\partial y} + i\beta E_y = -i\omega\mu H_x, \quad (1.4.13)$$

$$i\beta E_x + \frac{\partial E_z}{\partial x} = i\omega\mu H_y, \quad (1.4.14)$$

$$\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -i\omega\mu H_z. \quad (1.4.15)$$

С помощью соотношений (1.4.10), (1.4.11), (1.4.13) и (1.4.14) можно выразить поперечные компоненты поля через E_z и H_z :

$$E_x = -\frac{i}{\kappa^2} \left(\beta \frac{\partial E_z}{\partial x} + \omega\mu \frac{\partial H_z}{\partial y} \right), \quad (1.4.16)$$

$$E_y = -\frac{i}{\kappa^2} \left(\beta \frac{\partial E_z}{\partial y} - \omega\mu \frac{\partial H_z}{\partial x} \right), \quad (1.4.17)$$

$$H_x = -\frac{i}{\kappa^2} \left(\beta \frac{\partial H_z}{\partial x} - \omega\varepsilon \frac{\partial E_z}{\partial y} \right), \quad (1.4.18)$$

$$H_y = -\frac{i}{\kappa^2} \left(\beta \frac{\partial H_z}{\partial y} + \omega\varepsilon \frac{\partial E_z}{\partial x} \right), \quad (1.4.19)$$

где

$$\kappa^2 = k^2 - \beta^2, \quad (1.4.20)$$

$$k^2 = \omega^2 \varepsilon \mu. \quad (1.4.21)$$

Соответствующие выражения в цилиндрических координатах будут приведены в разд. 8.2. Уравнения (1.4.16)—(1.4.19) являются точными. Замена H_x и H_y в (1.4.12) выражениями (1.4.18) и (1.4.19) приводит к

$$\frac{\partial^2 E_z}{\partial x^2} + \frac{\partial^2 E_z}{\partial y^2} + \kappa^2 E_z = 0. \quad (1.4.22)$$

Из уравнения (1.4.15) аналогичным образом получаем

$$\frac{\partial^2 H_z}{\partial x^2} + \frac{\partial^2 H_z}{\partial y^2} + \kappa^2 H_z = 0. \quad (1.4.23)$$

Уравнения (1.4.22) и (1.4.23) выполняются строго лишь в случае $\varepsilon = \text{const}$. Примечательным здесь является то, что каждое из этих уравнений содержит либо только E_z , либо только H_z . Продольные компоненты векторов \mathbf{E} и \mathbf{H} не связаны друг с другом и могут быть выбраны произвольным образом, лишь бы они удовлетворяли уравнениям (1.4.22) и (1.4.23). Как правило, обе продольные компоненты поля связываются граничными условиями. Если же из граничных условий не вытекает связь этих компонент, то можно получить модовые решения при $E_z = 0$ или $H_z = 0$. Их называют соответственно поперечными электрическими или ТЕ-модами и поперечными магнитными или ТМ-модами.

В общем случае, когда ε зависит от x и y , уравнения (1.4.22) и (1.4.23) становятся приближенными. Однако они дают хорошие приближения, если относительные изменения ε много меньше единицы на расстояниях порядка длины волны. В случае оптических полей с их весьма короткими длинами волн эти приближения обычно достаточно точны и существенно упрощают вычисления.

До сих пор осталась не определенной постоянная пространства β . Она является собственным значением рассматриваемой граничной задачи. Ее величина или, точнее, ее возможные значения определяются граничными условиями задачи. Общие решения уравнений Максвелла получаются путем суперпозиции мод.

1.5. ГРАНИЧНЫЕ УСЛОВИЯ

Как неоднократно отмечалось, уравнения Максвелла не определяют электромагнитное поле полностью. Среди неограниченного числа возможных решений уравнений Максвелла нужно выделить те, которые удовлетворяют также граничным условиям рассматриваемой частной задачи.

В задачах о неоднородных структурах без скачкообразного изменения диэлектрической проницаемости граничным условием обычно является требование исчезновения поля в бесконечности и его ограниченность внутри любой конечной области пространства. Задачи подобного типа будут здесь встречаться. Требование исчезновения поля в бесконечности приводит к направляемым модам. Поле направляемых волн ограничено направляющей структурой и не теряет мощность на излучение.

Наиболее общий тип граничных условий соответствует кусочно-однородному распределению диэлектрической проницаемости. Этот случай имеет место, например, когда на пути светового луча стоит стеклянная линза. Для нас представляют интерес только граничные условия для переменных во времени полей (которые отличаются от граничных условий статических полей). Искомые граничные условия получаются из уравнений Максвелла в интегральной форме. Эти уравнения можно вывести интегрированием (1.2.1) и (1.2.2) по произвольному объему. Применение интегральной теоремы Стокса приводит к следующим интегральным соотношениям:

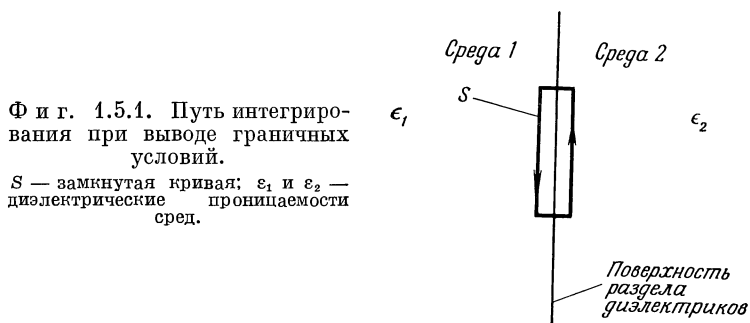
$$\oint_S \mathbf{H} d\mathbf{S} = \int_A \frac{\partial \mathbf{D}}{\partial t} \mathbf{n} dA, \quad (1.5.1)$$

$$\oint_S \mathbf{E} d\mathbf{S} = - \int_A \frac{\partial \mathbf{B}}{\partial t} \mathbf{n} dA. \quad (1.5.2)$$

Единичный вектор \mathbf{n} направлен перпендикулярно элементу поверхности dA , тогда как $d\mathbf{S}$ указывает направление, касательное к кривой S , ограничивающей поверхность A .

Поскольку форма и положение замкнутой кривой S произвольны, будем считать, что она расположена так, как показано на фиг. 1.5.1. Замкнутый путь интегрирования находится бесконечно близко к границе между двумя

диэлектрическими средами, обладающими различными значениями ϵ . Одна часть кривой расположена в среде 1, а другая — в среде 2. Участки пути, идущие параллельно границе, значительно короче длины волны излучения, но все же много длиннее участков пути интегрирования поперек границы. Фактически мы сокращаем длину этих последних участков пути до нуля. Область внутри кривой



становится исчезающе малой и потому правые части уравнений (1.5.1) и (1.5.2) можно принять равными нулю. Величины \mathbf{E} и \mathbf{H} на коротких, но конечных участках пути, расположенных целиком в одной или другой среде, можно считать постоянными, так что

$$[(H_t)_1 - (H_t)_2] \Delta S = 0,$$

$$[(E_t)_1 - (E_t)_2] \Delta S = 0.$$

Индекс t указывает, что речь идет о компонентах поля, тангенциальных к границе. Длина бесконечно малого участка пути интегрирования в каждой среде обозначена через ΔS . Отсюда сразу получаем искомые граничные условия для тангенциальных компонент поля:

$$(H_t)_1 = (H_t)_2, \quad (1.5.3)$$

$$(E_t)_1 = (E_t)_2. \quad (1.5.4)$$

Физический смысл соотношений (1.5.3) и (1.5.4) состоит в том, что тангенциальные компоненты полей \mathbf{E} и \mathbf{H} непрерывны на границе сред. Эти граничные условия совместно с условиями на бесконечности, рассмотренными

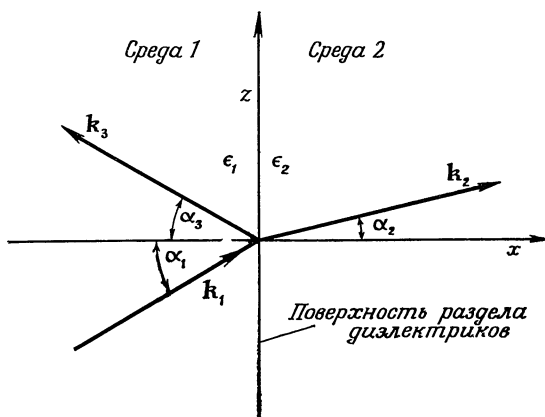
ранее, используются для выделения тех решений уравнений Максвелла, которые соответствуют данной конкретной физической задаче. Существуют граничные условия для нормальных компонент векторов \mathbf{D} и \mathbf{B} , которые вытекают из уравнений для дивергенций (1.2.5) и (1.2.6). Они показывают, что компоненты этих векторов, нормальные к границе раздела сред, также должны быть непрерывными. Однако уравнения для дивергенций могут быть получены из уравнений Максвелла, если взять дивергенцию от (1.2.1) и (1.2.2). Поскольку дивергенция ротора равна нулю, сразу же получаем, что должны быть равны нулю временные производные от дивергенций \mathbf{D} и \mathbf{B} . Но временные производные от величин, зависящих от времени, могут равняться нулю лишь в том случае, когда сами эти величины равны нулю. В электродинамике при наличии зависимости от времени уравнения для дивергенций вытекают из уравнений Максвелла и не являются независимыми, т. е. они не являются условиями, которые можно налагать отдельно. Граничные условия для нормальных компонент векторов \mathbf{D} и \mathbf{B} , вытекающие из уравнений для дивергенций, также не являются независимыми и уже содержатся в уравнениях Максвелла. Поэтому граничные условия для нормальных компонент полей \mathbf{D} и \mathbf{B} автоматически выполняются, если справедливы соотношения (1.5.3) и (1.5.4). Лишь в случае статических магнитных и электрических полей граничные условия для нормальных компонент независимы от граничных условий для тангенциальных компонент. Соотношения (1.5.3) и (1.5.4) являются граничными условиями только в случае разрыва однородности внутри диэлектрической среды.

Известно несколько иное граничное условие, часто используемое в случае идеального металлического проводника. Это условие требует равенства нулю компоненты электрического поля, касательной к поверхности идеального проводника. В этом случае не требуется дополнительного условия для компоненты поля \mathbf{H} . В оптике понятие идеального проводника почти не применяется, так как его нельзя реализовать со сколько-нибудь удовлетворительной степенью точности с помощью существующих металлов. Металл следует рассматривать как диэлектрический материал с комплексным значением диэлектриче-

ской проницаемости. Тогда граничные условия, установленные выше, пригодны и полностью описывают рассматриваемую ситуацию.

1.6. ОТРАЖЕНИЕ И ПРЕЛОМЛЕНИЕ НА ГРАНИЦЕ ДИЭЛЕКТРИКОВ

Мы теперь подготовлены для решения простейшей, но весьма важной задачи распространения волн. Пусть плоская волна с волновым вектором \mathbf{k}_1 падает из однородной среды с диэлектрической проницаемостью ϵ_1/ϵ_0 на плоскую границу раздела между этой средой и другой



Ф и г. 1.6.1. Отражение и преломление плоской волны на границе раздела двух диэлектрических сред.

Волновые векторы падающей, прошедшей и отраженной волн обозначены соответственно через \mathbf{k}_1 , \mathbf{k}_2 и \mathbf{k}_3 .

средой с диэлектрической проницаемостью, равной ϵ_2/ϵ_0 . Наша цель состоит в изучении явления отражения и преломления волны при ее падении на подобную границу.

Геометрия задачи схематически дана на фиг. 1.6.1. Система координат выбрана так, чтобы $k_y = 0$. Рассмотрим два различных случая. Предположим сначала, что электрическое поле поляризовано по направлению оси y , т. е. вектор электрического поля параллелен плоской

границе раздела. Позже мы положим, что вектор \mathbf{E} лежит в плоскости x, z . В любом случае будем рассматривать падающую и отраженную волны в среде 1 и прошедшую волну в среде 2. Уравнения для плоской волны, распространяющейся в произвольном направлении \mathbf{k} , уже были даны в разд. 1.4. Осталось сложить две плоские волны в среде 1, учесть наличие плоской волны в среде 2 и подобрать направления волновых векторов и амплитуды волн таким образом, чтобы удовлетворялись граничные условия (1.5.3) и (1.5.4).

Рассмотрим сначала случай, когда имеется только составляющая E_y электрического поля, так что в среде 1

$$E_y = Ae^{-ik_1 r} + Be^{-ik_3 r}, \quad (1.6.1)$$

а в среде 2

$$E_y = Ce^{-ik_2 r}. \quad (1.6.2)$$

Множитель $\exp(i\omega t)$ для простоты опущен. Поляризация поля \mathbf{H} следует из (1.4.6), так что в среде 1 (напомним, что имеется только y -компонента вектора \mathbf{E})

$$H_x = -An_{1z} \sqrt{\frac{\epsilon_1}{\mu}} e^{-ik_1 r} - Bn_{3z} \sqrt{\frac{\epsilon_1}{\mu}} e^{-ik_3 r}, \quad (1.6.3)$$

$$H_z = An_{1x} \sqrt{\frac{\epsilon_1}{\mu}} e^{-ik_1 r} + Bn_{3x} \sqrt{\frac{\epsilon_1}{\mu}} e^{-ik_3 r}, \quad (1.6.4)$$

а в среде 2

$$H_x = -Cn_{2z} \sqrt{\frac{\epsilon_2}{\mu}} e^{-ik_2 r}, \quad (1.6.5)$$

$$H_z = Cn_{2x} \sqrt{\frac{\epsilon_2}{\mu}} e^{-ik_2 r}. \quad (1.6.6)$$

Связь между векторами \mathbf{n} и \mathbf{k} дается соотношением (1.3.11). Теперь нужно потребовать, чтобы тангенциальные компоненты поля E_y и H_z удовлетворяли граничным условиям (1.5.3) и (1.5.4). При $x = 0$ получаем

$$Ae^{-ik_{1z}z} + Be^{-ik_{3z}z} = Ce^{-ik_{2z}z}, \quad (1.6.7)$$

$$An_{1x} \sqrt{\epsilon_1} e^{-ik_{1z}z} + Bn_{3x} \sqrt{\epsilon_1} e^{-ik_{3z}z} = Cn_{2x} \sqrt{\epsilon_2} e^{-ik_{2z}z}. \quad (1.6.8)$$

Эти уравнения должны выполняться для всех значений z ; иными словами, зависимость от z должна здесь отсут-

воватъ. А это возможно только, когда

$$k_{1z} = k_{2z} = k_{3z}. \quad (1.6.9)$$

Выражения для составляющих k_x и, следовательно, компонент n_x вектора \mathbf{n} получаются из (1.3.13) и (1.3.14):

$$k_{1x} = \sqrt{\omega^2 \varepsilon_1 \mu - k_{1z}^2} = n_{1x} k_1, \quad (1.6.10)$$

$$k_{2x} = \sqrt{\omega^2 \varepsilon_2 \mu - k_{2z}^2} = n_{2x} k_2, \quad (1.6.11)$$

$$k_{3x} = -k_{1x} = n_{3x} k_1. \quad (1.6.12)$$

Проекция k_{1z} , определяющая направление падающей плоской волны, может быть выбрана произвольно. Знак минус в (1.6.12) не следует из (1.3.13), но он необходим, чтобы избежать противоречия между (1.6.7) и (1.6.8). Если $k_{3x} = \omega \sqrt{\varepsilon_1 \mu} n_{3x}$ имеет тот же знак, что и k_{1x} , то левые части (1.6.7) и (1.6.8) будут пропорциональны и правые части будут противоречить друг другу. Уравнения (1.6.7) и (1.6.8) можно теперь использовать для определения амплитуд B и C , отнесенных к произвольной амплитуде A падающей плоской волны. Имеем систему уравнений

$$-B + C = A,$$

$$B + \frac{k_{2x}}{k_{1x}} C = A.$$

Ее решение имеет вид

$$C = \frac{2A}{1 + \frac{k_{2x}}{k_{1x}}}, \quad (1.6.13)$$

$$B = \frac{\left(1 - \frac{k_{2x}}{k_{1x}}\right) A}{1 + \frac{k_{2x}}{k_{1x}}}, \quad (1.6.14)$$

причем

$$\frac{k_{2x}}{k_{1x}} = \frac{n_{2x}}{n_{1x}} \sqrt{\frac{\varepsilon_2}{\varepsilon_1}} = \frac{\sqrt{\omega^2 \varepsilon_2 \mu - k_{1z}^2}}{\sqrt{\omega^2 \varepsilon_1 \mu - k_{1z}^2}}. \quad (1.6.15)$$

Соотношения (1.6.10)–(1.6.15) полностью определяют решение задачи. Остается исследовать это решение и определить физические свойства отраженной и прошедшей волн.

Так как модули векторов \mathbf{k}_1 и \mathbf{k}_3 , как и их компоненты k_z , одинаковы, из формулы (1.6.12) и фиг. 1.6.1

очевидно, что

$$\alpha_1 = \alpha_3. \quad (1.6.16)$$

Падающая и отраженная волны составляют одинаковые углы с нормалью к отражающей поверхности. Это и есть закон отражения. Обозначая абсолютную величину векторов \mathbf{k}_1 и \mathbf{k}_2 через k_1 и k_2 , получим из (1.3.13) следующее соотношение:

$$\frac{k_2}{k_1} = \sqrt{\frac{\varepsilon_2}{\varepsilon_1}}. \quad (1.6.17)$$

Согласно фиг. 1.6.1, синус угла α_i равен

$$\sin \alpha_i = \frac{k_{iz}}{k_i}, \quad i=1, 2 \text{ или } 3. \quad (1.6.18)$$

Используя соотношения (1.6.9), (1.6.17) и (1.6.18), можно записать

$$\frac{\sin \alpha_1}{\sin \alpha_2} = \frac{n_2}{n_1}. \quad (1.6.19)$$

В правой части стоит отношение показателей преломления двух сред. Показатель преломления определяется соотношением

$$n_i = \sqrt{\frac{\varepsilon_i}{\varepsilon_0}}. \quad (1.6.20)$$

Уравнение (1.6.19) является очень важным и известно под названием закона Снеллиуса. Этот простой закон позволяет проследить за ходом световых лучей через последовательность однородных сред. Поэтому оказывается возможным проанализировать способность линзовых систем формировать изображение путем проведения луча в соответствии с законом Снеллиуса. В соответствии с законом преломления световых лучей луч представляет собой ломаную линию, приближающуюся к направлению нормали к поверхности раздела при переходе из одной диэлектрической среды в другую, обладающую более высокой диэлектрической проницаемостью.

Фиг. 1.6.1 относится к случаю, когда показатель преломления среды 2 выше, чем у среды 1. Допустим, однако, на время, что $\varepsilon_1 > \varepsilon_2$ и, следовательно, $n_1 > n_2$. Тогда из формулы (1.6.19) имеем

$$\sin \alpha_2 = \frac{n_1}{n_2} \sin \alpha_1 > \sin \alpha_1. \quad (1.6.21)$$

По мере роста угла α_1 функция $\sin \alpha_2$ достигает единицы до того, как α_1 станет равным 90° . В этом предельном случае плоская волна входит в среду 2 параллельно границе между двумя средами. Бесконечно малое возрастание α_1 приводит к $\sin \alpha_2 > 1$, что невозможно удовлетворить ни при каких вещественных значениях α_2 . Физическая интерпретация этого явления заключается в том, что прошедшая волна больше не пересекает границу, т. е. остается лишь отраженная волна. Это явление называют полным внутренним отражением. Анализ на этом предельном случае не заканчивается, и мы можем изучить поведение плоских волн более детально. Из формул (1.6.9) и (1.6.18) следует, что

$$k_{2z} = k_{1z} > k_2 = \omega \sqrt{\epsilon_2 \mu}, \quad (1.6.22)$$

так что, согласно (1.6.11), величина k_{2x} становится мнимой. Зависимость амплитуды волны от x в среде 2 теперь задается функцией ¹⁾

$$e^{-|k_{2x}|x},$$

указывающей, что волна больше не распространяется в среду 2, а является спадающей (нераспространяющейся) волной с амплитудой поля, убывающей экспоненциально с ростом x .

Полное внутреннее отражение представляет собой пример существования в свободном пространстве нераспространяющихся волн, которые были рассмотрены в разд. 1.3. Полное внутреннее отражение наблюдается всякий раз, когда угол падения волны, приходящей к границе раздела двух диэлектрических сред со стороны более плотной среды (среды с большим показателем преломления), превышает некоторое критическое значение, определяемое уравнением

$$\sin \alpha_{ic} = \frac{n_2}{n_1}. \quad (1.6.23)$$

¹⁾ Квадратный корень в (1.6.11) может иметь знак либо плюс, либо минус. Выбираем $k_{2x} = -i |k_{2x}|$, так как в противном случае получается нарастающая волна, которая нарушает граничное условие конечности (или равенства нулю) интенсивности поля в бесконечности.

Уравнение (1.6.23) определяет критический угол для полного внутреннего отражения. Тот факт, что волна в среде 2 экспоненциально убывает, когда α_1 превышает критический угол, говорит о том, что поле в этой среде возбуждается со слишком короткой длиной волны, чтобы быть распространяющейся волной в среде с более низким показателем преломления (см. обсуждение данного вопроса в разд. 1.3).

Амплитуды полей прошедшей и отраженной волн даются выражениями (1.6.13) и (1.6.14). Из формулы (1.6.15) видно, что отношение k_{2x}/k_{1x} всегда положительно¹⁾. Как следствие этого величина C/A всегда положительна независимо от того, поступают волны из плотной среды в среду с низким показателем преломления или наоборот. Знак амплитуды B отраженной волны зависит от характера двух сред вблизи границы. Если волна отражается от среды с более высоким показателем преломления, то, как следует из формулы (1.6.15), величина k_{2x}/k_{1x} больше единицы и, следовательно, амплитуда B отрицательна. Отражение от более плотной среды изменяет знак напряженности электрического поля. В противном случае, когда световая волна отражается от среды с более низким показателем преломления по сравнению со средой, откуда она пришла, амплитуда B положительна, т. е. вектор E не меняет знака при отражении.

Представляет интерес рассмотреть отраженную и прошедшую через границу мощности. При использовании метода комплексных амплитуд исходным является соотношение (1.2.12)

$$S = \frac{1}{2} \operatorname{Re} (\mathbf{E} \times \mathbf{H}^*). \quad (1.6.24)$$

В частном случае, когда отлична от нуля только компонента E_y вектора E , для x -проекции вектора S получаем

$$P = |S_x| = \frac{1}{2} |\operatorname{Re} (E_y H_z^*)|. \quad (1.6.25)$$

x -проекция вектора Пойнтинга представляет собой мощность, проходящую перпендикулярно поверхности. Мощ-

¹⁾ В данном случае исключаем из рассмотрения явление полного внутреннего отражения.

ность S_z , протекающая параллельно границе, на поверхность раздела не падает. Рассмотрим мощность падающей, отраженной и прошедшей волн. Мощность падающей волны получим путем подстановки слагаемого с множителем A выражений (1.6.1) и (1.6.4) в (1.6.25). В итоге имеем

$$P_i = \frac{1}{2} n_{1x} \sqrt{\frac{\varepsilon_1}{\mu}} |A|^2. \quad (1.6.26)$$

Мощность отраженной волны содержит множитель B :

$$P_r = \frac{1}{2} n_{1x} \sqrt{\frac{\varepsilon_1}{\mu}} |B|^2. \quad (1.6.27)$$

Наконец, мощность прошедшей волны дается выражением

$$P_t = \frac{1}{2} n_{2x} \sqrt{\frac{\varepsilon_2}{\mu}} |C|^2. \quad (1.6.28)$$

Коэффициент отражения и коэффициент прохождения определяются следующими соотношениями:

$$R = \frac{P_r}{P_i} = \frac{|B|^2}{|A|^2}, \quad (1.6.29)$$

$$T = \frac{P_t}{P_i} = \frac{k_{2x}}{k_{1x}} \frac{|C|^2}{|A|^2}. \quad (1.6.30)$$

Для того чтобы выразить коэффициенты отражения и прохождения через угол падения волны, запишем выражение (1.6.15) с учетом (1.6.18) в виде

$$\frac{k_{2x}}{k_{1x}} = \frac{n_2}{n_1} \frac{\cos \alpha_2}{\cos \alpha_1}.$$

Подставив вместо α_2 выражение, полученное из закона Снеллиуса (1.6.19), имеем

$$\frac{k_{2x}}{k_{1x}} = \frac{\sqrt{n_2^2 - n_1^2 \sin^2 \alpha_1}}{n_1 \cos \alpha_1}. \quad (1.6.31)$$

Отношение прошедшей мощности к падающей можно теперь получить из формул (1.6.13), (1.6.30) и (1.6.31):

$$T_E = \frac{4n_1 \cos \alpha_1 \sqrt{n_2^2 - n_1^2 \sin^2 \alpha_1}}{(n_1 \cos \alpha_1 + \sqrt{n_2^2 - n_1^2 \sin^2 \alpha_1})^2}. \quad (1.6.32)$$

Аналогичным образом получается отношение отраженной мощности к падающей:

$$R_E = \frac{(n_1 \cos \alpha_1 - \sqrt{n_2^2 - n_1^2 \sin^2 \alpha_1})^2}{(n_1 \cos \alpha_1 + \sqrt{n_2^2 - n_1^2 \sin^2 \alpha_1})^2}. \quad (1.6.33)$$

Эти соотношения справедливы лишь в тех случаях, когда квадратный корень в (1.6.32) и (1.6.33) является вещественным. Случай мнимого корня требует более детального рассмотрения. Квадратный корень становится мнимым, когда угол α_1 превышает свое критическое значение, определяемое формулой (1.6.23). В результате получается полное внутреннее отражение; распространение волны в среде 2 отсутствует. Компонента k_{2x} волнового вектора в среде 2 становится мнимой, так что компонента n_{2x} единичного вектора, которая входит в формулу (1.6.6), также мнимая, в результате чего выражение (1.6.25) обращается в нуль, если его применить к прошедшей волне. В случае полного внутреннего отражения коэффициент прохождения

$$T = 0,$$

а коэффициент отражения

$$R = 1.$$

Это становится очевидным, если учесть, что в этом случае отношение k_{2x}/k_{1x} является чисто мнимой величиной и квадрат абсолютной величины B , как следует из формулы (1.6.14), равен $|A|^2$. В общем случае, когда присутствуют как отраженная, так и прошедшая волны, из формул (1.6.32) и (1.6.33) вытекает равенство

$$T + R = 1, \quad (1.6.34)$$

что указывает на сохранение полной энергии в рассматриваемом процессе.

Выражения для коэффициентов прохождения и отражения упрощаются, когда волна падает на граничную поверхность между двумя диэлектрическими средами под прямым углом. Полагая $\alpha_1 = 0$, из формулы (1.6.32) получаем

$$T_E = \frac{4n_1 n_2}{(n_1 + n_2)^2}, \quad (1.6.35)$$

а из (1.6.33) следует, что

$$R_E = \frac{(n_1 - n_2)^2}{(n_1 + n_2)^2}. \quad (1.6.36)$$

Другой крайний случай — случай скользящего падения — должен быть рассмотрен отдельно для каждого из двух возможных вариантов. При $n_1 > n_2$ имеем полное внутреннее отражение и вся мощность отражается при углах падения, больших критического. В случае $n_1 < n_2$ полное внутреннее отражение отсутствует, но в пределе, когда волна падает параллельно стенке ($\alpha_1 = 90^\circ$), получаем из формул (1.6.32) и (1.6.33)

$$T_E = 0$$

и

$$R_E = 1.$$

Таким образом, при скользящем падении вся мощность отражается от границы раздела диэлектриков независимо от того, падает ли волна на оптически более плотную или менее плотную среду.

До сих пор речь шла о частном случае, когда вектор \mathbf{E} падающей, а также отраженной и прошедшей волн параллелен граничной поверхности. То обстоятельство, что векторы \mathbf{E} отраженной и прошедшей волн направлены параллельно вектору \mathbf{E} падающей волны, вытекает из уравнений Максвелла и граничных условий. Решения граничных задач единственны. Поэтому если получено некоторое частное решение, то никакое другое решение не может существовать.

Рассмотрим, наконец, случай падения плоской волны на плоскую границу раздела двух диэлектрических сред, когда вектор \mathbf{E} волны лежит в плоскости, определяемой волновым вектором и нормалью к поверхности. Свойства плоских волн проходить и отражаться от границы раздела сред зависят от их поляризации, так что необходимо исследовать волну, поляризованную перпендикулярно поляризации, рассмотренной выше. Пусть в среде 1 имеет место суперпозиция падающей и отраженной волн:

$$E_x = A e_{1x} e^{-ik_1 r} + B e_{3x} e^{-ik_3 r}, \quad (1.6.37)$$

$$E_z = A e_{1z} e^{-k_1 r} + B e_{3z} e^{-ik_3 r}, \quad (1.6.38)$$

$$H_y = A \sqrt{\frac{\epsilon_1}{\mu}} e^{-ik_1 r} + B \sqrt{\frac{\epsilon_1}{\mu}} e^{-ik_3 r}. \quad (1.6.39)$$

В среде 2 прошедшая волна в общем случае имеет вид

$$E_x = C e_{2x} e^{-ik_2 r}, \quad (1.6.40)$$

$$E_z = C e_{2z} e^{-ik_2 r}, \quad (1.6.41)$$

$$H_y = C \sqrt{\frac{\varepsilon_2}{\mu}} e^{-ik_2 r}. \quad (1.6.42)$$

Выражения (1.6.37) — (1.6.42) являются решениями уравнений Максвелла. Коэффициенты A , B , C , а также направления волновых векторов должны определяться из граничных условий (1.5.3) и (1.5.4). Граничные условия при $x = 0$ приводят к уравнениям (напомним, что $k_y = 0$)

$$A e_{1z} e^{-ik_{1zz}} + B e_{3z} e^{-ik_{3zz}} = C e_{2z} e^{-ik_{2zz}}, \quad (1.6.43)$$

$$A e^{-ik_{1zz}} + B e^{-ik_{3zz}} = \sqrt{\frac{\varepsilon_2}{\varepsilon_1}} C e^{-ik_{2zz}}. \quad (1.6.44)$$

Если потребовать, чтобы эти уравнения не содержали зависимость от z , то вновь приходим к условию (1.6.9). Ввиду того что соотношения между компонентами волновых векторов такие же, как в рассмотренном выше случае, можно утверждать, что закон Снеллиуса (1.6.19) справедлив и для этой поляризации. Направления отраженной и прошедшей волн в точности такие же, как и в предыдущем случае, т. е. они не зависят от поляризации. Также наблюдается явление полного внутреннего отражения, если имеет место переход из более плотной в менее плотную среду.

Осталось определить отношения C/A и B/A из уравнений

$$A e_{1z} + B e_{3z} = C e_{2z}, \quad (1.6.45)$$

$$A + B = \frac{n_2}{n_1} C. \quad (1.6.46)$$

Их решения

$$C = \frac{e_{1z} - e_{3z}}{e_{2z} - \frac{n_2}{n_1} e_{3z}} A, \quad (1.6.47)$$

$$B = \frac{\frac{n_2}{n_1} e_{1z} - e_{2z}}{e_{2z} - \frac{n_2}{n_1} e_{3z}} A. \quad (1.6.48)$$

Составляющие векторов \mathbf{e} можно выразить через компоненты векторов \mathbf{k} , если воспользоваться соотношением (1.4.5), которое может быть записано также в форме

$$\mathbf{k}_i \cdot \mathbf{e}_i = 0, \quad i = 1, 2, 3. \quad (1.6.49)$$

Поскольку для рассматриваемой поляризации $e_y = 0$, то

$$\frac{e_{iz}}{e_{ix}} = -\frac{k_{ix}}{k_{iz}}. \quad (1.6.50)$$

Это уравнение совместно с условием нормировки

$$e_{ix}^2 + e_{iz}^2 = 1 \quad (1.6.51)$$

дает возможность определить составляющие e_z :

$$e_{1z} = \frac{k_{1x}}{k_1}, \quad (1.6.52)$$

$$e_{2z} = \frac{k_{2x}}{k_2}, \quad (1.6.53)$$

$$e_{3z} = -e_{1z}. \quad (1.6.54)$$

Знак минус в последнем уравнении необходим, чтобы избежать противоречия между (1.6.45) и (1.6.46). Окончательно, используя формулы (1.6.17) и (1.6.20), получаем решения

$$C = \frac{2A}{\frac{n_2}{n_1} + \frac{n_1}{n_2} \frac{k_{2x}}{k_{1x}}}, \quad (1.6.55)$$

$$B = \frac{\frac{n_2}{n_1} - \frac{n_1}{n_2} \frac{k_{2x}}{k_{1x}}}{\frac{n_2}{n_1} + \frac{n_1}{n_2} \frac{k_{2x}}{k_{1x}}} A. \quad (1.6.56)$$

Выражения для коэффициента прохождения и коэффициента отражения получаются из соотношений (1.6.29) и (1.6.30) с учетом формулы (1.6.31) ¹⁾:

$$T_H = \frac{4n_1 \cos \alpha_1 \sqrt{n_2^2 - n_1^2 \sin^2 \alpha_1}}{\left(n_2 \cos \alpha_1 + \frac{n_1}{n_2} \sqrt{n_2^2 - n_1^2 \sin^2 \alpha_1} \right)^2}, \quad (1.6.57)$$

$$R_H = \frac{\left(n_2 \cos \alpha_1 - \frac{n_1}{n_2} \sqrt{n_2^2 - n_1^2 \sin^2 \alpha_1} \right)^2}{\left(n_2 \cos \alpha_1 + \frac{n_1}{n_2} \sqrt{n_2^2 - n_1^2 \sin^2 \alpha_1} \right)^2}. \quad (1.6.58)$$

¹⁾ Легко показать, что равенства (1.6.29) и (1.6.30) в этом случае также справедливы.

Эти выражения справедливы только для вещественных значений квадратных корней. Для мнимых значений вновь получается $T = 0$ и $R = 1$. Легко видеть, что и для R_H и T_H имеет место сохранение энергии (1.6.34).

Коэффициенты отражения и прохождения в обоих рассмотренных случаях имеют одинаковый вид. Однако имеется одно существенное отличие. Нетрудно показать, что коэффициент R_E , определяемый формулой (1.6.33), никогда не обращается в нуль при $n_1 \neq n_2$. Коэффициент же отражения R_H волны с электрическим вектором, лежащим в одной плоскости с вектором \mathbf{k} и нормалью к поверхности, обращается в нуль при

$$\sin \alpha_1 = \frac{n_1}{\sqrt{n_1^2 + n_2^2}} \quad (1.6.59)$$

[см. формулу (1.6.58)]. Угол, определяемый выражением (1.6.59), известен как угол Брюстера. Волна, поляризованная в плоскости падения (плоскость падения определяется вектором \mathbf{k} и нормалью к отражающей поверхности), проходит через границу без отражения, если она падает под углом Брюстера. Формулу для угла Брюстера можно записать более изящно:

$$\operatorname{tg} \alpha_{1B} = \frac{n_2}{n_1}. \quad (1.6.60)$$

Используя закон Снеллиуса (1.6.19), можно записать угол Брюстера преломленной волны в следующем виде:

$$\operatorname{tg} \alpha_{2B} = \frac{n_1}{n_2}. \quad (1.6.61)$$

Сравнение выражений (1.6.60) и (1.6.61) показывает, что если волну направить под углом α_2 и при этом заменить показатель преломления среды 1 на n_2 , а в среде 2 взять n_1 , то волна снова будет падать под углом Брюстера. Иными словами, если направление волны, падающей под углом Брюстера, изменить на обратное, то, возвращаясь по пройденному пути, она снова пройдет без отражения. Волна, падающая на стеклянный предмет под углом Брюстера, проходит стекло без отражения, если обе поверхности стекла параллельны.

Направления отраженной (если она существует) и прошедшей волн находятся в замечательном соотношении друг с другом, если волна падает под углом Брюстера. Чтобы получить это соотношение, перемножим выражения (1.6.60) и (1.6.61):

$$\operatorname{tg} \alpha_{1B} \operatorname{tg} \alpha_{2B} = 1.$$

Это соотношение можно представить следующим образом:

$$\cos \alpha_{1B} \cos \alpha_{2B} - \sin \alpha_{1B} \sin \alpha_{2B} = \cos (\alpha_{1B} + \alpha_{2B}) = 0. \quad (1.6.62)$$

Отсюда имеем, что сумма углов падения и прохождения при условии, что волна падает под углом Брюстера, равна 90° :

$$\alpha_{1B} + \alpha_{2B} = 90^\circ. \quad (1.6.63)$$

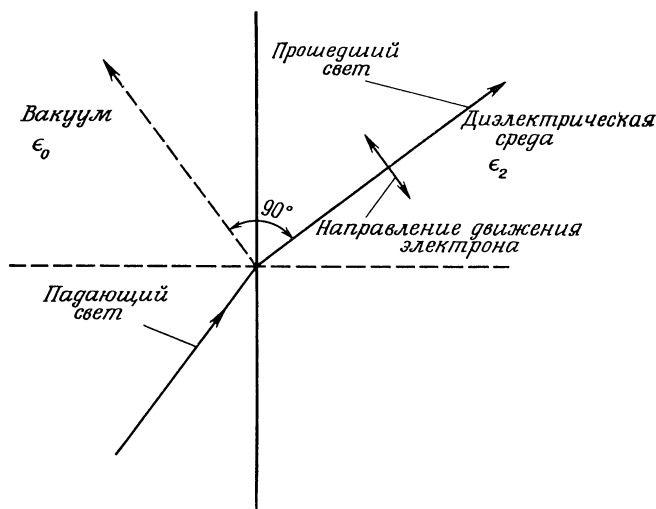
Характеристикой угла Брюстера является отсутствие отраженной волны. Однако закон отражения формально имеет место, так что отраженная волна, если бы она существовала, шла бы под углом $\alpha_{3B} = \alpha_{1B}$, т. е.

$$\alpha_{3B} + \alpha_{2B} = 90^\circ. \quad (1.6.64)$$

Из фиг. 1.6.1 видно, что при условии выполнения соотношения (1.6.64) направления прошедшей и отраженной волн составляют прямой угол.

Направления падающей, прошедшей и отраженной (если бы она имела место) волн показаны на фиг. 1.6.2. С ее помощью можно объяснить существование угла Брюстера. Для этого нужно знать, что явления отражения и преломления обусловлены изменением фазовой скорости вследствие взаимодействия поля со связанными электронами в материале среды. Электроны осциллируют с частотой световой волны. Осциллирующие электроны в свою очередь излучают вторичные световые волны. Суперпозиция падающей световой волны и волны, вызванной осциллирующими электронами, является причиной появления отраженной и преломленной волн. Внутри материала проходит лишь преломленная волна, которая возникает при совместном действии падающей волны и вынужденных колебаний электронов. В связи с тем что плоское электромагнитное поле содержит лишь поперечные волны [см. формулу (1.4.5)], вынужденные колебания

электронов под влиянием преломленной волны происходят перпендикулярно распространению этой волны. Падающая волна, электрический вектор которой поляризован в плоскости вектора \mathbf{k} и нормали к поверхности диэлектрика, заставляет электроны внутри материала совершать колебания в той же плоскости. На фиг. 1.6.2

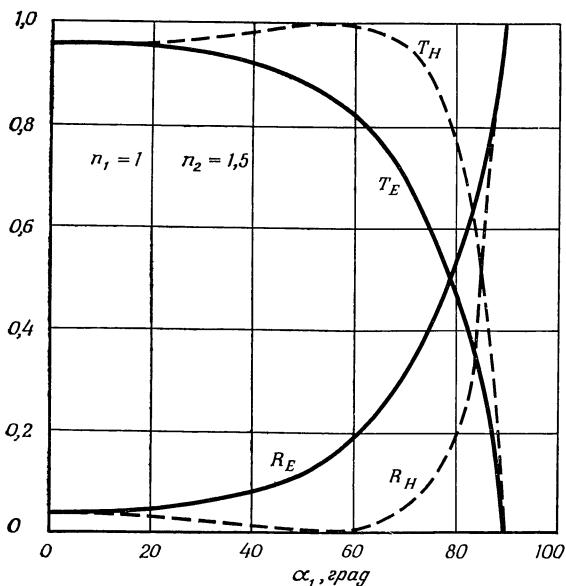


Ф и г. 1.6.2. К объяснению существования угла Брюстера.

маленькой двойной стрелкой указано направление вынужденных колебаний электронов. Из формулы (1.6.64) видно, что электроны движутся параллельно направлению, в котором должна была распространяться отраженная волна. По теории излучения электрического диполя [2] электроны не излучают в направлении своего движения. Следовательно, отраженная волна, обусловленная переизлучением осциллирующих электронов, не может существовать.

Явления обычного отражения и преломления, угол Брюстера и полное внутреннее отражение отображены на фиг. 1.6.3 и 1.6.4, где представлены графики коэффициентов прохождения и отражения, определяемых формулами (1.6.32), (1.6.33), (1.6.57) и (1.6.58). На фиг. 1.6.3

изображены графики коэффициентов отражения и пропускания для обеих поляризаций, когда волна входит из среды 1 в более плотную среду 2 ($n_1 < n_2$). В этом случае явление полного внутреннего отражения отсутствует.

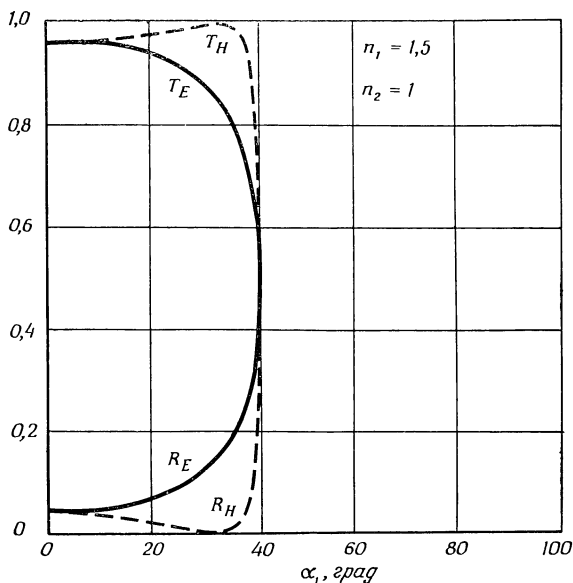


Ф и г. 1.6.3. Зависимость коэффициентов отражения R_E и R_H и коэффициентов прохождения T_E и T_H от угла падения α_1 в среде 1. Точка $R_H = 0$ определяет угол Брюстера. Среда 2 ($n_2 = 1,5$) более плотная, чем среда 1 ($n_1 = 1$).

Из кривой R_H видно исчезновение отраженной волны при угле Брюстера. При отношении $n_2/n_1 = 1,5$ (фиг. 1.6.3) угол Брюстера равен 56° . На фиг. 1.6.3 видно также, как растет коэффициент отражения, когда волна падает под все более пологими углами. Величина R приближается к единице, когда угол падения близок к 90° .

Фиг. 1.6.4 относится к случаю, когда волна входит в среду с меньшим показателем преломления. В этом случае возможно полное внутреннее отражение. Для $n_1/n_2 = 1,5$ угол полного внутреннего отражения равен $41,8^\circ$.

Таким образом, простой пример плоской волны, проходящей через поверхность раздела диэлектрических сред,



Ф и г. 1.6.4. Зависимость коэффициентов отражения и прохождения от угла падения α_1 для света, падающего из более плотной среды 1 ($n_1 = 1,5$; $n_2 = 1$). Отчетливо различимы угол Брюстера ($R_H = 0$) и полное внутреннее отражение.

содержит много важного и полезного для физиков. Другие примеры распространения волн в диэлектрической среде будут рассмотрены в последующих главах.

2

ТЕОРИЯ ДИФРАКЦИИ

2.1. ВВЕДЕНИЕ

Дифракция является характерной чертой явления распространения волн. Если рассматривать свет с точки зрения геометрической оптики, то следует ожидать появления за краем непрозрачного препятствия резких теней. Однако из волновой оптики следует, что резких теней не может быть, потому что часть света проникает в область геометрооптической тени. Важным применением теории дифракции является изучение поля в области, соответствующей большим углам расходимости коллимированного пучка света. Если плоская волна падает на отверстие в непрозрачном экране, то за экраном световой луч начинает расходиться и вдали от отверстия заходит в область, определяемую углом расходимости прошедшего светового пучка. Этот угол зависит от соотношения между размером отверстия и длиной волны света.

Знание свойств дифракции существенно для понимания работы фокусирующих световодов, рассматриваемых в главе 5. В настоящей главе математический аппарат описания дифракции использован при рассмотрении характерных примеров. Наиболее элементарными примерами являются случаи дифракции на щели и на круглом отверстии в непрозрачном экране. В этой главе рассматривается также дифракция Брэгга. Часто это явление называют отражением Брэгга. Мы выбрали термин «дифракция Брэгга» потому, что это явление имеет место в трехмерной дифракционной решетке. Дифракция Брэгга может рассматриваться как особый случай рассеяния света. Рассмотрение дифракции Брэгга с помощью метода «связанных» волн, проведенное в последнем разделе этой главы, представляет собой пример анализа явления взаимодействия волн, которое будет обсуждаться также и в последующих главах.

Понятия дифракционной решетки и дифракции Брэгга чрезвычайно важны также для понимания голографии, рассмотрение которой выходит за рамки этой книги.

2.2. ДИФРАКЦИОННЫЙ ИНТЕГРАЛ КИРХГОФА — ГЮЙГЕНСА

Наиболее замечательными явлениями, сопровождающими распространение волн (в противоположность движению частиц), являются интерференция и дифракция. Дифракцией объясняется появление света в области тени, чего нельзя понять с точки зрения геометрической оптики. Поляризация световой волны играет важную роль при отражении и преломлении световых волн; при рассмотрении дифракции ее влияние можно не учитывать. Особенно это справедливо, если длина волны света существенно короче любого линейного размера предмета, на который падает свет. По этой причине при описании дифракции достаточно использовать скалярное волновое уравнение (1.3.6), игнорируя векторный характер световой волны. Каждая компонента электромагнитного поля является решением волнового уравнения, так что скалярная функция ψ может представлять любую (электрическую или магнитную) компоненту волны.

Мы ограничим наше рассмотрение монохроматическим излучением, т. е. излучением определенной частоты. Более общие формы излучения могут быть описаны как суперпозиции его синусоидальных компонент. В этом и следующем разделах зависимость от времени выражается в виде экспоненциального множителя

$$e^{i\omega t}, \quad (2.2.1)$$

который во всех уравнениях будет опущен. Использование зависимости от времени в форме (2.2.1) приводит к тому, что дифференцирование по времени может быть заменено умножением на $i\omega$. Тогда волновое уравнение (1.3.6) принимает вид

$$\nabla^2 \psi + k^2 \psi = 0, \quad (2.2.2)$$

где k определяется из формулы (1.3.13). Следуя примеру Куранта [4], назовем соотношение (2.2.2) приведенным

волновым уравнением. Волновое уравнение является уравнением в частных производных, содержащим вторые производные как по времени, так и по пространственным координатам. Поскольку производная по времени исключена из уравнения (2.2.2), название «приведенное волновое уравнение» кажется уместным. Уравнение (2.2.2) известно также как уравнение Гельмгольца.

Решение приведенного волнового уравнения существенно упрощается при использовании δ -функции Дирака [5]. Предполагается, что читатель знаком с ее свойствами. Математики не любят работать с δ -функцией из-за ее сингулярного поведения. Однако она очень полезна при решении инженерных и физических задач и часто упрощает вычисления. Использование этой функции обычно приводит к правильным результатам, что подтверждает целесообразность ее применения ¹⁾.

Поэтому наряду с уравнением (2.2.2) рассмотрим уравнение

$$\nabla^2 G + k^2 G = \delta(\mathbf{r} - \mathbf{r}'). \quad (2.2.3)$$

Дельта-функция векторного аргумента может быть раскрыта в виде

$$\delta(\mathbf{r} - \mathbf{r}') = \delta(x - x') \delta(y - y') \delta(z - z'). \quad (2.2.4)$$

Функция G , известная как функция Грина [6], зависит от двух наборов переменных:

$$G = G(x, y, z, x', y', z'). \quad (2.2.5)$$

Решение приведенного волнового уравнения (2.2.2) может быть теперь получено следующим образом. Умножим обе части уравнения (2.2.2) на G , а обе части уравнения (2.2.3) — на ψ и вычтем одно уравнение из другого. После интегрирования по некоторому объему получим

$$\psi(x', y', z') = \int_V (\psi \nabla^2 G - G \nabla^2 \psi) dV. \quad (2.2.6)$$

¹⁾ Последовательное применение результатов теории обобщенных функций делает оперирование с δ -функциями и другими сингулярными функциями строгим в математическом смысле (см., например, В. С. Владимиров, Уравнения математической физики, изд-во «Наука», 1971). — Прим. ред.

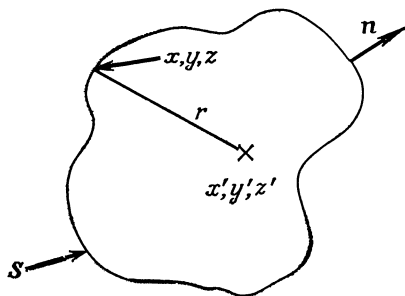
Функция ψ с аргументами x', y', z' получена в результате интегрирования произведения δ -функции на ψ . Подынтегральное выражение может быть переписано в виде

$$\psi \nabla^2 G - G \nabla^2 \psi = \nabla \cdot (\psi \nabla G - G \nabla \psi). \quad (2.2.7)$$

Используя теорему Стокса [3], преобразуя интеграл по объему в интеграл по поверхности, получим решение приведенного уравнения в виде

$$\psi(x', y', z') = \int_S \left(\psi \frac{\partial G}{\partial n} - G \frac{\partial \psi}{\partial n} \right) dS. \quad (2.2.8)$$

Интегрирование проводится по замкнутой поверхности S , а n — направление внешней нормали к этой поверхности



Фиг. 2.2.1. Геометрическая иллюстрация представления интеграла (2.2.8). Поверхность S окружает точку наблюдения. Единичный вектор n нормален к поверхности.

(фиг. 2.2.1). Точка x', y', z' , в которой вычисляется функция ψ , расположена внутри замкнутой поверхности S . Может показаться, что, получив формулу (2.2.8), мы не многого достигли. Однако в однородных средах функцию Грина найти не трудно, а функция ψ на поверхности S в типичных дифракционных задачах по крайней мере частично известна.

Функция Грина в свободном пространстве дается выражением

$$G = -\frac{1}{4\pi} \frac{e^{-ikr}}{r}, \quad (2.2.9)$$

где

$$r = [(x' - x)^2 + (y' - y)^2 + (z' - z)^2]^{1/2}. \quad (2.2.10)$$

Легко показать, что выражение (2.2.9) является решением уравнения (2.2.3). Вычислим первую частную производ-

ную G по x :

$$-4\pi \frac{\partial G}{\partial x} = -\left(ik + \frac{1}{r}\right) \frac{x-x'}{r} \frac{e^{-ikr}}{r}. \quad (2.2.11)$$

Дифференцирование этого выражения второй раз дает

$$-4\pi \frac{\partial^2 G}{\partial x^2} = \left[\left(\frac{3ik}{r} + \frac{3}{r^2} - k^2 \right) \frac{(x-x')^2}{r^3} - \frac{ik}{r^2} - \frac{1}{r^3} \right] e^{-ikr}. \quad (2.2.12)$$

Вторые производные по y и z имеют точно такой же вид, с тем лишь исключением, что $x - x'$ заменяется на $y - y'$ или $z - z'$ соответственно. Тогда сразу же получаем

$$\nabla^2 G + k^2 G = 0 \quad \text{для } r \neq 0. \quad (2.2.13)$$

При $r = 0$ функция G имеет сингулярность так что соотношения (2.2.11) и (2.2.12) не справедливы в этой точке. Для изучения природы сингулярности производных запишем интеграл

$$\int_V (\nabla^2 G + k^2 G) dV = I. \quad (2.2.14)$$

Согласно уравнению (2.2.13), подынтегральное выражение равно нулю везде, за исключением точки $r = 0$. Определим значение интеграла по бесконечно малому объему v , включающему точку $r = 0$, используя следующее соотношение:

$$I = \lim_{v \rightarrow 0} \int_v (\nabla^2 G + k^2 G) dV = \lim_{\rho \rightarrow 0} \int_s \nabla G \cdot \mathbf{n} ds + \lim_{v \rightarrow 0} k^2 \int_v G dV. \quad (2.2.15)$$

В качестве поверхности s выберем бесконечно малую сферу радиуса ρ , окружающую точку $r = 0$; вектор \mathbf{n} является внешней нормалью к этой сфере. Второй член в правой части можно вычислить, используя выражение (2.2.9)

$$\lim_{v \rightarrow 0} \int_v G dV = - \lim_{\rho \rightarrow 0} \int_0^\rho r dr = 0. \quad (2.2.16)$$

Первый член правой части в (2.2.15) может быть вычислен также легко:

$$\begin{aligned} \lim_{\rho \rightarrow 0} \int_s \nabla G \cdot \mathbf{n} \, ds &= \lim_{\rho \rightarrow 0} \rho^2 \int_0^\pi d\theta \int_0^{2\pi} \left(\frac{\partial G}{\partial r} \right)_{r=\rho} \sin \theta \, d\varphi = \\ &= \lim_{\rho \rightarrow 0} \left\{ -\frac{\rho^2}{4\pi} \left[4\pi \left(-ik - \frac{1}{\rho} \right) \frac{e^{-ik\rho}}{\rho} \right] \right\} = 1. \end{aligned} \quad (2.2.17)$$

Используя формулы (2.2.16) и (2.2.17), получим

$$I = 1. \quad (2.2.18)$$

Это как раз тот результат, который мы должны получить из определения функции Грина. Проведенные здесь вычисления показывают, что функция Грина (2.2.9) действительно удовлетворяет уравнению (2.2.3). Левая часть уравнения (2.2.3) обращается в нуль везде [см. формулу (2.2.13)], за исключением точки $r = 0$. Сингулярность в этой точке должна описываться δ -функцией, потому что интеграл по бесконечно малому объему, включающему сингулярность, равен единице.

Подставляя выражение (2.2.9) в (2.2.8), получаем искомого решение волнового уравнения

$$\psi(x', y', z') = \frac{1}{4\pi} \int_s \left[\frac{\partial \psi}{\partial n} \frac{e^{-ikr}}{r} - \psi \frac{\partial}{\partial n} \left(\frac{e^{-ikr}}{r} \right) \right] dS. \quad (2.2.19)$$

Здесь вектор \mathbf{n} задает направление внешней нормали, т. е. этот вектор направлен наружу от объема, содержащего точку x', y', z' .

Дифракционный интеграл Кирхгофа — Гюйгенса (2.2.19) очень полезен для решения большого числа дифракционных задач. Однако следует иметь в виду, что существуют и другие способы решения задач дифракции. Например, в выражение (2.2.8) можно подставить и другие функции Грина. Можно выбрать функцию Грина, являющуюся решением уравнения (2.2.3) и удовлетворяющую, кроме того, определенным граничным условиям на поверхности S [7]. Функция Грина может быть выбрана так, чтобы она обращалась, например, в нуль на поверхности S . Такая специальная функция Грина в выражении (2.2.8) обращает в нуль член, содержащий $\partial\psi/\partial n$,

Это упрощает задачу, потому что теперь достаточно знать на поверхности S лишь функцию ψ . Однако это «упрощение» достигается за счет использования вместо простой функции (2.2.9) гораздо более сложной функции Грина. Фактически функция Грина, которая удовлетворяет граничному условию $G = 0$ на поверхности S , может быть найдена только для задач с простейшей геометрией. Другая возможность заключается в том, чтобы использовать функцию Грина, которая удовлетворяет граничному условию $\partial G / \partial n = 0$ на S . В этом случае, чтобы оценить интеграл (2.2.8), надо знать на поверхности S величину $\partial \psi / \partial n$. Нахождение такой функции Грина конечно не легче, чем в предыдущем случае.

Вместо того чтобы искать решения волнового уравнения в форме выражения (2.2.8) с подходящей функцией Грина, можно использовать метод разложения поля по плоским волнам, в общих чертах описанный в разд. 1.3. Для примера в разд. 2.3 мы решим простую дифракционную задачу двумя методами: используя интеграл Кирхгофа — Гюйгенса (2.2.19) и находя решение в виде суперпозиции плоских волн в соответствии с формулой (1.3.23).

Для практических расчетов часто полезно использовать некоторые приближения дифракционного интеграла (2.2.19). Первое приближение, которое почти всегда справедливо для оптических задач, основано на малости длины волны света. В большинстве задач, представляющих практический интерес, точка x' , y' , z' , в которой вычисляется функция ψ , удалена от S на много длин волн, так что условие

$$k \gg \frac{1}{r} \quad (2.2.20)$$

справедливо для всех r , входящих в интеграл (2.2.19). Тогда можно пренебречь производной от $1/r$ по сравнению с производной от $\exp(-ikr)$ и записать приближенное выражение

$$\psi(x', y', z') = \frac{1}{4\pi} \int_S \left(\frac{\partial \psi}{\partial n} + ik \frac{\partial r}{\partial n} \psi \right) \frac{e^{-ikr}}{r} dS. \quad (2.2.21)$$

Следующее приближение связано с получением для r более простого, чем (2.2.10), выражения. Аппроксимация для r

возможна в тех случаях, когда апертура отверстий дифракционных экранов мала по сравнению с расстояниями от них до экранов, на которых исследуются поля. В этих случаях удовлетворяются условия

$$|x' - x| \ll |z' - z| \quad \text{и} \quad |y' - y| \ll |z' - z|. \quad (2.2.22)$$

Тогда удобно для величины r использовать разложение в ряд Тейлора

$$r = z' - z + \frac{1}{2} \frac{(x' - x)^2 + (y' - y)^2}{z' - z} + \dots \quad (2.2.23)$$

Важно отметить, что выполнения только условия (2.2.22) еще недостаточно для использования в выражении (2.2.21) приближения (2.2.23). Поскольку r входит в экспоненту в (2.2.19) в комбинации kr (а это большое число), то именно произведение k на отброшенные малые члены выражения (2.2.23) должно быть мало по сравнению с единицей. Если это обстоятельство не учесть, могут возникнуть существенные ошибки.

Наконец, возьмем в качестве поверхности S непрозрачный плоский экран с отверстием. На этот экран слева падает плоская волна. Это простейшая дифракционная задача, но именно благодаря своей простоте она поможет глубоко понять свойства дифракции света. Физические аспекты этой задачи мы обсудим в следующих разделах. В этом разделе рассмотрим лишь различные приближенные формы дифракционного интеграла. Геометрия дифракции на отверстии в экране показана на фиг. 2.2.2.

В соответствии с выбранной геометрией

$$\frac{\partial}{\partial n} = -\frac{\partial}{\partial z}, \quad (2.2.24)$$

так что, используя формулу (1.3.21) и положив $g = \psi$, имеем

$$\frac{\partial \psi}{\partial n} = -\frac{\partial \psi}{\partial z} = ik_z \psi = ik \psi \cos \gamma. \quad (2.2.25)$$

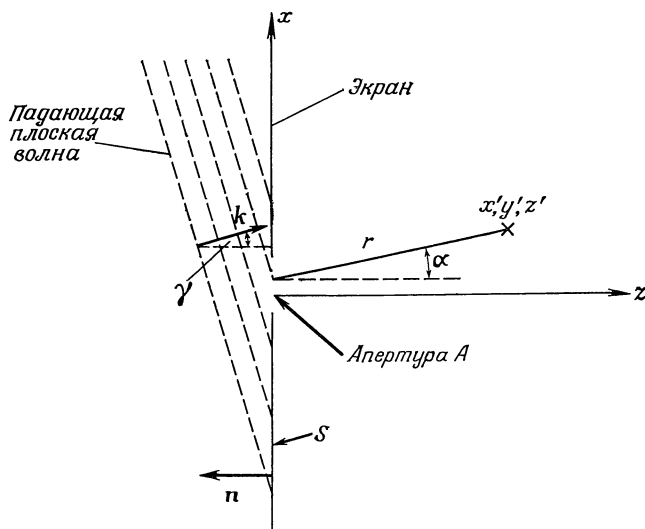
Производная от r равна

$$\frac{\partial r}{\partial n} = -\frac{\partial r}{\partial z} = \frac{z' - z}{r} = \cos \alpha, \quad (2.2.26)$$

Используя формулы (2.2.25) и (2.2.26), получаем из (2.2.21)

$$\psi(x', y', z') = \frac{ik}{4\pi} \int_A (\cos \gamma + \cos \alpha) \psi(x, y, z) \frac{e^{-ikr}}{r} dS. \quad (2.2.27)$$

Поверхность S , по которой производится интегрирование, включает экран, показанный на фиг. 2.2.2, и бесконечную



Ф и г. 2.2.2. Дифракция плоской волны на щели в бесконечном непрозрачном экране.

полусфере, ограничивающую пространство справа от экрана. Мы рассматриваем падающую волну как часть очень длинного, но конечного цуга волн. Конечный цуг волн не может достичь бесконечной полусферы за конечное время, так что интеграл по поверхности этой полусферы обращается в нуль. По-видимому, выражение (2.2.27) столь же точно, как и (2.2.21), из которого оно получено. Однако мы заменили область интегрирования S на A , т. е. интегрирование проводится теперь только по апертуре отверстия A в экране. При переходе от (2.2.21) к (2.2.27) подразумевается, что поле ψ и его производная $\partial\psi/\partial n$

обращаются в нуль на непрозрачном экране. С первого взгляда это предположение может показаться логичным. Однако можно проверить, что решение волнового уравнения обращается в нуль везде, если оно и его первая производная одновременно обращаются в нуль на любом конечном интервале [4]. Это означает, что, строго говоря, наше предположение, состоящее в том, что поле и его первая производная обращаются в нуль на непрозрачном экране, противоречит условию, что функция ψ должна быть решением волнового уравнения. Другими словами, это предположение несовместимо с тем фактом, что формула (2.2.19) получена как решение волнового уравнения. Действительно, если мы вычислим значения ψ из (2.2.19), предполагая, что функция ψ и ее нормальная производная равны нулю на непрозрачном экране, то придем к противоречию, не получив ψ , обращающуюся в нуль вместе с производной на экране. Это обстоятельство представляет серьезную математическую проблему. Чтобы можно было использовать дифракционный интеграл, необходимо знать значения ψ и ее нормальной производной не только на отверстии, но и на экране. Итак, чтобы получить желаемое решение, необходимо знать функцию в области, где мы ее не можем знать. К счастью, несмотря на эту математическую дилемму, соотношение (2.2.27) приводит к результату, хорошо совпадающему с экспериментальными данными. Нарушение математических принципов, по-видимому, не настолько серьезно, чтобы быть преградой при получении полезных результатов. Значения ψ и ее нормальной производной очень малы на экране и могут быть измерены только в непосредственной близости от отверстия. Малая ошибка, которая возникла из-за предположения о равенстве функции нулю, не столь существенна, чтобы привести к заметным неприятностям. Дифракционный интеграл в виде (2.2.27) широко и успешно используется при решении задач оптической дифракции и при расчетах СВЧ-антенн. Строго говоря, несправедливо и другое предположение, которое подразумевается в (2.2.27), о том, что поле в отверстии такое же, как и поле падающей волны при отсутствии экрана. Однако это возражение той же природы, что и касающееся равенства ψ нулю на экране.

Во многих случаях апертура отверстия в экране настолько мала ¹⁾ и точка x', y', z' настолько близка к оси, что можно получить

$$\cos \alpha \approx 1. \quad (2.2.28)$$

Если, наконец, плоская волна падает перпендикулярно или по крайней мере почти перпендикулярно к поверхности отверстия, так что можно считать

$$\cos \gamma = 1, \quad (2.2.29)$$

то

$$\psi(x', y', z') = \frac{i}{\lambda} \int_A \psi(x, y, z) \frac{e^{-ikr}}{r} dS. \quad (2.2.30)$$

Приближение (2.2.30) интересно не только с практической, но и с исторической точки зрения. Гюйгенс в 1690 г. высказал догадку, что распространение световых волн можно описать интегралом такого вида. Чтобы понять интерпретацию выражения (2.2.30), использующую принцип Гюйгенса, рассмотрим волну, описываемую формулой (2.2.9), более внимательно. В разд. 1.3 показано, что волна, описываемая равенством (1.3.21), является синусоидальной плоской волной, распространяющейся со скоростью $v = \omega/k$ в направлении вектора \mathbf{k} . Подобные аргументы показывают, что формула (2.2.9) описывает сферическую волну, распространяющуюся со скоростью $v = \omega/k$ в радиальном направлении от исходной точки $r = 0$. В точке расположения источника ($r = 0$), генерирующего сферическую волну, амплитуда волны бесконечна. Амплитуда сферической волны не постоянна, а равномерно убывает как $1/r$.

Физический смысл выражения (2.2.30) теперь ясен. Каждая точка отверстия может рассматриваться как источник сферической волны. Сила источника определяется амплитудой падающей волны в этой точке отверстия. Поле в точке x', y', z' образуется в результате суперпозиции всех сферических волн, достигающих этой точки. Интуитивно ясно, что это объяснение обладает внутренней логикой; оно было предложено Гюйгенсом без реше-

¹⁾ Мала по сравнению с расстоянием от экрана, но велика по отношению к длине волны.— *Прим. ред.*

ния волнового уравнения. Наш вывод показывает, что принцип Гюйгенса является лишь аппроксимацией. Точная форма дифракционного интеграла (2.2.19) в действительности совсем другая. Но наиболее важная часть утверждения, содержащаяся в этом принципе — возмущения, возникающие в отверстии, распространяются как сферические волны из каждого элемента отверстия — действительно справедлива. Соотношение (2.2.27) тоже является аппроксимацией, но оно отлично от принципа Гюйгенса. Это соотношение показывает, что еще недостаточно просто предположить, что падающая волна действует как источник сферических волн с силой, пропорциональной амплитуде падающей волны в каждой точке отверстия. Необходимо приписать силе источника весовые коэффициенты, чтобы обеспечить направленность (фиктивных) источников.

Возвращаясь к выражению (2.2.27), окончательно получим полезную аппроксимацию дифракционного интеграла путем подстановки выражения (2.2.23). Еще раз предостережем: чтобы использовать это приближение, необходимо проверить, действительно ли произведение k на член третьего порядка [которым пренебрегли в (2.2.23)] много меньше единицы. Если это условие выполняется, получаем приближение

$$\psi(x', y', z') = \frac{i}{2\lambda} e^{-ikz'} \int_A (\cos \gamma + \cos \alpha) \frac{\psi}{z' - z} \times \\ \times \exp \left[ik \left(z - \frac{(x' - x)^2 + (y' - y)^2}{2(z' - z)} \right) \right] dS. \quad (2.2.31)$$

В соответствии с приближением (2.2.23) r в знаменателе заменено на $z' - z$. Дифракционный интеграл такого вида действительно наиболее часто используется на практике. Обычно различают две области его применения. Если рассматриваемое поле настолько близко к отверстию, что член $x^2 + y^2$, появляющийся в показателе степени экспоненты, следует принимать во внимание, имеет место дифракция Френеля. Если этот член пренебрежимо мал, наблюдается дифракция Фраунгофера. Дифракция Фраунгофера имеет место вдали от источника дифракции. Дифракция Френеля наблюдается также в областях, достаточно удаленных от отверстия, чтобы приближение (2.2.31) было справедливо, но не настолько далеких,

чтобы можно было говорить о дифракции Фраунгофера. Указанная терминология отражает историю развития оптики. Дальнейшее обсуждение приближения Френеля можно найти в разд. 4.6.

Часто бывает достаточно рассмотреть задачу в двух измерениях, а не трехмерную задачу, как мы делали до сих пор. Преимуществом двумерных задач является то, что их проще рассматривать, и исследование таких задач часто оказывается достаточным для полного понимания проблем дифракции. Исследования дифракционных интегралов, проведенные до сих пор, и в частности важная аппроксимация (2.2.31), основаны на разложении падающего поля на непрерывный набор источников сферических волн. Сферические волны естественно появляются в трехмерных задачах. Для двумерных задач эквивалентом сферических волн являются цилиндрические волны. Следующий вывод дифракционных интегралов обычно не приводится в учебниках по оптике.

Для двумерных задач характерно, что все характеристики поля, так же как и геометрия пространства, в котором распространяются поля, не зависят от одной из пространственных переменных. Схема дифракционного опыта, изображенная на фиг. 2.2.2, становится двумерной, если предположить, что ни одна из интересующих нас переменных не зависит от координаты y . Символически это можно выразить с помощью соотношения

$$\frac{\partial}{\partial y} = 0. \quad (2.2.32)$$

Будем называть задачу двумерной, если справедливо соотношение (2.2.32). В двух измерениях волновое уравнение принимает вид

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial z^2} + k^2 \psi = 0. \quad (2.2.33)$$

Продолжая рассуждения точно так же, как и прежде, рассмотрим в дополнение к волновому уравнению (2.2.33) неоднородное уравнение

$$\frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial z^2} + k^2 G = \delta(z - z') \delta(x - x'). \quad (2.2.34)$$

Умножим обе части уравнения (2.2.33) на G , а (2.2.34) — на ψ и вычтем одно из другого. Интегрируя полученное уравнение по области в плоскости x, z , получим

$$\psi(x', z') = \int_C \left(\psi \frac{\partial G}{\partial n} - G \frac{\partial \psi}{\partial n} \right) dC. \quad (2.2.35)$$

Здесь была использована двумерная интегральная теорема для преобразования интеграла по поверхности в интеграл по замкнутой кривой, окружающей эту поверхность. Направление n совпадает с направлением внешней нормали к кривой C .

Функция Грина наиболее легко может быть найдена в полярных координатах. Направим полярную ось вдоль оси y , расстояние от оси обозначим через r , а угол между радиусом r и осью z — через α . Тогда уравнение (2.2.34) может быть записано в виде

$$\frac{\partial^2 G}{\partial r^2} + \frac{1}{r} \frac{\partial G}{\partial r} + \frac{1}{r^2} \frac{\partial^2 G}{\partial \alpha^2} + k^2 G = \delta(z - z') \delta(x - x'). \quad (2.2.36)$$

Нас интересуют решения, симметричные в плоскости x, z , другими словами, решения, не зависящие от α . Причина этого ограничения станет ясна, когда мы рассмотрим более детально общие решения с зависимостью от α . Такие решения могут быть записаны как произведение функции, которая зависит только от α , на функцию только r . Последняя представляет собой либо $\cos(n\alpha)$, либо $\sin(n\alpha)$. Постоянная n должна быть целой, так как мы требуем, чтобы функция G принимала то же самое значение при изменении α от 0 до 2π . Если теперь найти особенность, которая появляется после подстановки решения в (2.2.36), то получим, что интеграл от левой части равен нулю для всех функций, кроме тех, для которых $n = 0$. Другими словами, только функции, не зависящие от α , обладают нужной особенностью в точке $r = 0$.

Дополнительно мы должны потребовать, чтобы функция Грина представляла волны, исходящие из точки x', z' . Имеются также решения, представляющие волны, которые приходят из бесконечности и собираются в точке x', z' . Эти волны должны быть исключены, так как из физических соображений очевидно, что поле должно образовываться волнами, исходящими из области источника конеч-

ной протяженности. В наших дифракционных задачах следует ожидать, что поле в точке x' , z' генерируется (фиктивными) источниками, расположенными в отверстии. Волны, собирающиеся из бесконечности на точку в отверстии, не удовлетворяют этому требованию. Это ограничение на допустимые функции Грина было впервые сформулировано Зоммерфельдом и называется условием Зоммерфельда.

Возвращаясь к уравнению (2.2.36), теперь можно записать просто

$$\frac{\partial^2 G}{\partial r^2} + \frac{1}{r} \frac{\partial G}{\partial r} + k^2 G = \delta(x - x') \delta(z - z'). \quad (2.2.37)$$

Левая часть этого уравнения имеет типичный вид дифференциального уравнения Бесселя. Следовательно, решениями уравнения (2.2.37) должны быть цилиндрические функции. Существуют цилиндрические функции различного вида. Так как (2.2.37) является дифференциальным уравнением второго порядка, то, как известно, существуют два независимых решения. Функции Бесселя и Неймана являются независимыми решениями этого вида [6, 8]. Однако ни одно из этих решений не пригодно здесь. Функции Бесселя не имеют особенностей, которые требуются согласно правой части уравнения (2.2.37). Функция Неймана обладает необходимым сингулярным поведением в источнике, но она представляет стоячие волны и таким образом нарушает условие Зоммерфельда. Суперпозиция этих двух функций могла бы удовлетворить обоим требованиям. Такая суперпозиция существует в форме функции Ханкеля двух родов. Функция Ханкеля первого рода представляет волны, приходящие из бесконечности [при выборе временного множителя $\exp(i\omega t)$]. Функция Ханкеля второго рода как раз соответствует нашим целям. Она имеет требуемую сингулярность и представляет волны, исходящие из начала системы координат, в которой описана функция. Следовательно, функция Грина имеет вид

$$G(x, x', z, z') = \frac{i}{4} H_0^{(2)}(kr). \quad (2.2.38)$$

Оставляем читателю проверить, что G удовлетворяет уравнению (2.2.37), в частности, что сингулярность приводит к нужному виду правой части уравнения.

Решение двумерной дифракционной задачи может теперь быть записано в виде

$$\psi(x', z') = \frac{i}{4} \int_C \left(\psi \frac{\partial H_0^{(2)}}{\partial n} - H_0^{(2)} \frac{\partial \psi}{\partial n} \right) dC. \quad (2.2.39)$$

Так же как и в трехмерном случае, для конкретных расчетов удобно использовать подходящие приближения этого точного выражения. Для первого приближения воспользуемся тем фактом, что вследствие малости длины волны света аргумент функции Ханкеля всегда большой. Это позволяет использовать асимптотическое выражение для функций Ханкеля при большом аргументе [8]:

$$H_0^{(2)}(kr) = \sqrt{\frac{2}{\pi kr}} e^{-i(kr - \pi/4)}. \quad (2.2.40)$$

С помощью выражения (2.2.40) решение (2.2.39) приобретает форму, очень похожую на ту, которая была получена в трехмерном случае [формула (2.2.19)]:

$$\psi(x', z') = \frac{1}{\sqrt{8\pi k}} e^{-i\pi/4} \int_C \left[\frac{\partial \psi}{\partial n} \frac{e^{-ikr}}{\sqrt{r}} - \psi \frac{\partial}{\partial n} \left(\frac{e^{-ikr}}{\sqrt{r}} \right) \right] dC. \quad (2.2.41)$$

Здесь

$$r = \sqrt{(z' - z)^2 + (x' - x)^2}. \quad (2.2.42)$$

Разница между выражениями (2.2.19) и (2.2.41) заключается в появлении цилиндрических волн в двумерном решении. Амплитуды таких волн спадают как $r^{-1/2}$ вместо r^{-1} , как в случае сферических волн.

Далее можно перейти к приближению выражения (2.2.41), соответствующему приближению (2.2.31). При условиях, подобных указанным выше, получим приближенное выражение

$$\begin{aligned} \psi(x', z') = & \frac{1}{2\sqrt{\lambda}} e^{i\pi/4} e^{-ikz'} \int_C \frac{\cos \gamma + \cos \alpha}{\sqrt{z' - z}} \psi(x, z) \times \\ & \times \exp \left[ik \left(z - \frac{(x' - x)^2}{2(z' - z)} \right) \right] dC, \end{aligned} \quad (2.2.43)$$

более простое по сравнению с выражением (2.2.31). Упрощение состоит в том, что двумерное выражение включает

только интеграл по кривой C . В случае плоской апертуры интегрирование проводится просто по x . Интеграл в трехмерной задаче берется по поверхности отверстия, что для плоской апертуры приводит к двойному интегралу по x и y . Это очень облегчает расчеты, если существуют аналитические решения, и особенно полезным оказывается, когда можно получить лишь численное решение на ЭВМ. Одно интегрирование вместо двух приводит к большой экономии времени. Мы будем иметь возможность рассмотреть численное интегрирование выражения типа (2.2.43) при исследовании сложной задачи прохождения волн через систему апертур (см. разд. 5.8). Эта задача может быть решена только численно. Хорошее понимание поведения волн можно получить, сведя задачу к двум измерениям. Это упрощение делает машинное решение экономически возможным.

2.3. ДИФРАКЦИЯ НА ЩЕЛИ В НЕПРОЗРАЧНОМ ЭКРАНЕ

В качестве первого приложения дифракционного интеграла рассмотрим решение задачи о дифракции волн, когда плоская волна падает на щель в непрозрачном экране. Задача с прямоугольным отверстием в непрозрачном экране ненамного сложнее задачи с щелью, но, так как из решения задачи с прямоугольным отверстием можно извлечь не так много дополнительной информации, мы остановимся на задаче с щелью.

Поучительно рассмотреть два способа решения этой задачи. Как отмечалось выше, решение с использованием дифракционного интеграла не является единственным способом решения дифракционных задач. Для иллюстрации этого сначала получим решение задачи с использованием решения волнового уравнения в виде (1.3.23). Для данной конкретной задачи это выражение можно упростить. Поскольку нас интересует задача рассеяния плоской монохроматической волны, то можно произвести интегрирование по ω . Монохроматическая волна обладает одной частотой. Ее спектр, следовательно, имеет одну δ -образную линию, которую можно выразить математически, взяв

$$\varphi(k_x, k_y, \omega) = 2\pi \bar{\varphi}(k_x, k_y) \delta(\omega - \omega_0). \quad (2.3.1)$$

Подстановка выражения (2.3.1) позволяет исключить интегрирование по ω в (1.3.23).

Нашу задачу со щелью можно проиллюстрировать на фиг. 2.2.2. Щель имеет ширину d и расположена так, что края щели параллельны направлению y . Предположим, что падающая волна имеет волновой вектор, лежащий в плоскости x, z , так что $k_y = 0$. Геометрия задачи такова, что если компонента k_y отсутствует, на щели не будут создаваться волны, распространяющиеся в направлении y . Следовательно, рассеянное поле не зависит от y . Таким образом, наша задача является двумерной в смысле определений, приведенных в предыдущем разделе. Эти предварительные соображения позволяют упростить интеграл (1.3.23) с помощью следующего соотношения:

$$\varphi(k_x, k_y, \omega) = (2\pi)^2 \varphi'(k_x) \delta(\omega - \omega_0) \delta(k_y). \quad (2.3.2)$$

Такое представление амплитуды преобразует выражение (1.3.23) к следующему виду:

$$\psi(x', z', t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi'(k_x) e^{-i(k_x x' + k_z z')} dk_x. \quad (2.3.3)$$

Временной множитель $\exp(i\omega_0 t)$ здесь опущен для простоты. Чтобы еще упростить задачу, предположим, что волна падает на щель перпендикулярно. Следовательно, падающая волна может быть описана выражением

$$\psi(x, z) = \psi_0 e^{-ikz}. \quad (2.3.4)$$

Предположим, что поле в щели с достаточной точностью может быть описано просто путем усечения функции (2.3.4). При расположении щели в плоскости $z = 0$ усеченная функция имеет вид

$$\psi(x, z) = \begin{cases} \psi_0, & |x| < d/2, \\ 0, & |x| > d/2. \end{cases} \quad (2.3.5)$$

Амплитуда φ' может быть получена просто как преобразование Фурье функции (2.3.5)

$$\varphi'(k_x) = \int_{-d/2}^{d/2} \psi_0 e^{ik_x x} dx = 2\psi_0 \frac{\sin k_x \frac{d}{2}}{k_x}. \quad (2.3.6)$$

Поле дифракции за экраном можно описать интегралом

$$\psi(x', z') = \frac{\psi_0}{\pi} \int_{-\infty}^{\infty} \frac{\sin k_x \frac{d}{2}}{k_x} \exp[-i(k_x x' + \sqrt{k^2 - k_x^2} z')] dk_x. \quad (2.3.7)$$

Компонента k_z была получена из формул (1.3.22) и (1.3.13) при $k_y = 0$.

Интеграл (2.3.7) нельзя вычислить точно. Необходимо применить некоторые разумные приближения, чтобы найти аналитическое решение нашей задачи. Только при $d = \infty$ решение очевидно. Дельта-функция Дирака может быть представлена в следующем виде [12]:

$$\delta(k_x) = \lim_{d \rightarrow \infty} \frac{1}{\pi} \frac{\sin k_x \frac{d}{2}}{k_x}. \quad (2.3.8)$$

Выражение (2.3.8) позволяет сразу вычислить интеграл (2.3.7) в пределе при $d \rightarrow \infty$. Решение

$$\psi(x', z') = \psi_0 e^{-ikz'} \quad (2.3.9)$$

не является сюрпризом. Оно показывает, что падающая плоская волна проходит через бесконечно широкую щель невозмущенной.

Приближение для случая конечной ширины щели можно получить, если ограничиться изучением дифрагированного поля на больших расстояниях за экраном. Чтобы получить приближение для дальнего поля, используем метод стационарной фазы [9]. Метод стационарной фазы чрезвычайно полезен для получения приближенных решений задач дифракции на больших расстояниях, которые представляют интерес в большом числе приложений. Мы будем иметь возможность использовать такое приближение в данной книге неоднократно. Этот метод основан на следующих соображениях. Пусть значения x' и z' (или по крайней мере одной из этих переменных) очень велики. Вследствие этого аргумент экспоненциальной функции в выражении (2.3.7) имеет большую величину и экспоненциальная функция изменяется очень быстро с изменением k_x . В подынтегральном выражении быстро

меняющаяся осциллирующая функция стремится уничтожить вклад от функции, которая умножается на нее. Следовательно, если функция, которая умножена на экспоненциальный множитель, изменяется медленно по сравнению с экспоненциальной функцией, результат интегрирования будет исчезающе мал. Теперь предположим, что на оси k_x имеется точка, где производная аргумента экспоненты обращается в нуль. Такая точка максимума или минимума аргумента определяет малую область, в которой аргумент стационарен, другими словами, он не меняется в первом порядке при изменении k_x . Такая область дает большой вклад в интеграл, так как быстрые осцилляции прекращаются в этой области изменения k_x и никакого взаимного уничтожения не происходит. Метод стационарной фазы состоит в нахождении точки или точек, где аргумент быстро осциллирующей функции имеет стационарное значение. Аргумент разлагается в ряд Тейлора в окрестности этой точки: разложение проводится до членов второго порядка. Член первого порядка, конечно, обращается в нуль по определению стационарной точки. Функция, умножаемая на быстро осциллирующую функцию, взятая в стационарной точке, выносится из-под интеграла. Метод стационарной фазы аналогичен действию δ -функции, так как интеграл пропорционален величине функции, умноженной на осциллирующую функцию со значением аргумента, взятым в стационарной точке.

Прежде чем использовать этот метод применительно к интегралу (2.3.7), перепишем его таким образом, чтобы экспоненциальные функции, содержащиеся в его подынтегральном выражении, были представлены в виде

$$\psi(x', z') = \frac{\psi_0}{2i\pi} \int_{-\infty}^{\infty} \left(\frac{1}{k_x} e^{-i\theta_-} - \frac{1}{k_x} e^{-i\theta_+} \right) dk_x, \quad (2.3.10)$$

где

$$\theta_- = \left(x' - \frac{d}{2} \right) k_x + \sqrt{k^2 - k_x^2} z', \quad (2.3.11)$$

а

$$\theta_+ = \left(x' + \frac{d}{2} \right) k_x + \sqrt{k^2 - k_x^2} z'. \quad (2.3.12)$$

Чтобы найти стационарные точки фазы θ_- , возьмем ее производную и приравняем нулю

$$\frac{\partial \theta_-}{\partial k_x} = x' - \frac{d}{2} - \frac{k_x}{\sqrt{k^2 - k_x^2}} z' = 0. \quad (2.3.13)$$

Следовательно, стационарной точкой является

$$k_x^{(-)} = \frac{x' - \frac{d}{2}}{\sqrt{z'^2 + \left(x' - \frac{d}{2}\right)^2}} k. \quad (2.3.14)$$

Знак минус перед квадратным корнем должен быть отброшен, так как при этом корень k_x не удовлетворяет уравнению (2.3.13). Стационарная точка для θ_+ получается из формулы (2.3.14) изменением знака перед d . Чтобы разложить фазы в ряд Тейлора до второй степени $(k_x - k_x^{(-)})$, нужно найти вторую производную от θ . Дифференцируя (2.3.13) еще раз, получим

$$\frac{\partial^2 \theta_-}{\partial k_x^2} = - \frac{k^2}{(k^2 - k_x^2)^{3/2}} z'. \quad (2.3.15)$$

Разложение θ_- в ряд Тейлора

$$\theta_- = (\theta_-)_{k_x = k_x^{(-)}} + \frac{1}{2} \left(\frac{\partial^2 \theta_-}{\partial k_x^2} \right)_{k_x = k_x^{(-)}} (k_x - k_x^{(-)})^2 \quad (2.3.16)$$

с учетом выражения (2.3.14) принимает вид

$$\sqrt{z'^2 + \left(x' - \frac{d}{2}\right)^2} k - \frac{\left[z'^2 + \left(x' - \frac{d}{2}\right)^2\right]^{3/2}}{2kz'^2} u^2, \quad (2.3.17)$$

где

$$u = k_x - k_x^{(-)}. \quad (2.3.18)$$

Наша аппроксимация справедлива только для далеких полей. Следовательно, разумно использовать тот факт, что

$$z' \gg d, \quad (2.3.19)$$

и получить следующие приближения:

$$\sqrt{z'^2 + \left(x' - \frac{d}{2}\right)^2} = r \left(1 - \frac{x'd}{2r^2}\right) \quad (2.3.20)$$

и

$$\left[z'^2 + \left(x' - \frac{d}{2} \right)^2 \right]^{3/2} = r^3 \left(1 - \frac{3}{2} \frac{x'd}{r^2} \right), \quad (2.3.21)$$

где

$$r^2 = x'^2 + z'^2. \quad (2.3.22)$$

При использовании метода стационарной фазы функция, которая умножается на быстро осциллирующую экспоненциальную функцию, выносится из-под интеграла. В рассматриваемом случае множитель $1/k_x$ в выражении (2.3.10) может быть вынесен из-под интеграла и заменен на $1/k_x^{(-)}$ или $1/k_x^{(+)}$ соответственно. Интегрируется только функция $\exp(iau^2)$, в результате чего имеем

$$\int_{-\infty}^{\infty} e^{iau^2} du = (1+i) \sqrt{\frac{\pi}{2a}} = \sqrt{\frac{\pi}{a}} e^{i\pi/4}. \quad (2.3.23)$$

Таким образом, мы получили все необходимые результаты для вычисления интеграла (2.3.10). После проведения необходимых подстановок с учетом того, что члены, соответствующие θ_+ , получаются путем замены d на $-d$ во всех уравнениях, имеем ¹⁾

$$\psi(x', z') = \sqrt{\frac{2}{\pi}} \psi_0 e^{i\pi/4} \frac{e^{-ikr}}{\sqrt{kr}} \frac{\sin \pi \frac{d}{\lambda} \alpha}{\alpha}. \quad (2.3.24)$$

Малый угол

$$\alpha \approx \operatorname{tg} \alpha = \frac{x'}{z'} \approx \frac{x'}{r} \quad (2.3.25)$$

определяет направление, в котором точка x', z' оказывается видимой из щели.

Метод стационарной фазы является не просто удобной математической аппроксимацией, но имеет и определенный физический смысл. Точка $k_x^{(\mp)}$ определяет ту область на оси k_x , которая дает максимальный вклад в интеграл. Так как интеграл представляет собой суперпозицию плоских волн, выражение (2.3.14) показывает, какая часть

¹⁾ Все члены, содержащие d , кроме входящих в экспоненту, опущены как несущественные.

этих волн вносит наиболее существенный вклад в процесс дифракции. Отношение $k_x^{(-)}/k$ определяет синус угла между направлением распространения волн и осью z . С учетом нашего приближения дальнего поля можно записать выражение (2.3.14) в виде

$$\frac{k_x^{(-)}}{k} \approx \frac{x'}{\sqrt{z'^2 + x'^2}} = \sin \alpha \approx \alpha. \quad (2.3.26)$$

Правая часть выражения (2.3.26) вычислена с помощью формулы (2.3.25). Таким образом, получено, что направление плоских волн, вносящих наибольший вклад в процесс дифракции, совпадает с направлением, под которым точка поля x' , z' видна из щели. Это интересный результат, так как плоские волны, движущиеся в других направлениях, можно не учитывать в точке наблюдения. В действительности дифракционная картина, конечно, получается как суперпозиция бесконечного числа плоских волн. Метод стационарной фазы раскрывает тот факт, что только волны, которые находятся в непосредственной близости (в пространстве k) к волнам, направленным на точку наблюдения, дают существенный вклад в поле в этой точке.

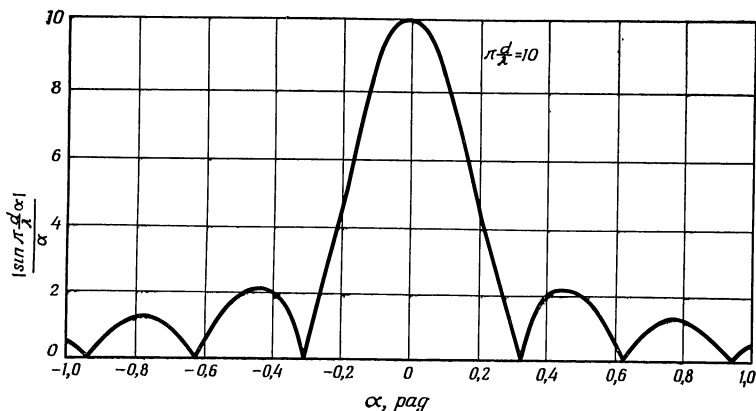
Дифракционное поле (2.3.24) имеет форму цилиндрической волны, спадающей с расстоянием как $r^{-1/2}$. Эта волна промодулирована по углу множителем

$$\frac{\sin \pi \frac{d}{\lambda} \alpha}{\alpha}. \quad (2.3.27)$$

Множитель (2.3.27), как показано на фиг. 2.3.1, имеет знакомый вид функции $\sin x/x$. Он имеет максимум при $\alpha = 0$. Функция симметрична относительно α и убывает при возрастании $|\alpha|$. Ширина главного лепестка выражения (2.3.27) определяется величиной $\pi d/\lambda$. Определим угловую ширину лепестка излучения как угол $\alpha = \Delta\alpha$, при котором функция (2.3.27) первый раз обращается в нуль. Этой точке соответствует значение аргумента синуса, равное π . Отсюда находим

$$\Delta\alpha = \frac{\lambda}{d}. \quad (2.3.28)$$

Это важная формула. Она показывает, что полуширина $\Delta\alpha$ главного лепестка излучения уменьшается при уменьшении длины волны λ излучения, а также при увеличении ширины щели d . Проходя через узкую щель, свет рассеивается в стороны тем больше, чем уже щель и чем больше длина волны. Соотношение (2.3.28) имеет большое значение для расчета направленных антенн, для которых наши результаты непосредственно применимы. В разд. 3.6 мы



Ф и г. 2.3.1. Дифракционная картина, создаваемая плоской волной, проходящей через щель в непрозрачном экране.

увидим, что выражение (2.3.28) может быть интерпретировано как следствие знаменитого принципа неопределенности Гейзенберга [10].

Осталось показать, что решение (2.3.24) также можно получить с помощью дифракционного интеграла (2.2.43). Поскольку мы имеем дело с далекой частью дифрагированного излучения, вызванного волной, которая падает нормально на поверхность щели, то, прежде чем использовать выражение (2.2.43), можно сделать ряд упрощающих приближений.

Положим

$$\cos \gamma = \cos \alpha = 1, \quad (2.3.29)$$

так как волна падает нормально ($\gamma = 0$) и мы предполагаем, что в приближении (2.3.25) величина α мала. Нако-

неп, пренебрежем членом x^2 в показателе экспоненты под интегралом (2.2.43). С учетом этих дополнительных приближений интеграл (2.2.43) можно записать в виде

$$\psi(x', z') = \frac{e^{i\pi/4}}{\sqrt{\lambda}} e^{-ikz'} \frac{\psi_0}{\sqrt{z'}} e^{-ik(x'^2/2z')} \int_{-d/2}^{d/2} e^{ik(x'/z')x} dx. \quad (2.3.30)$$

Пренебрежение в показателе экспоненты под интегралом членом x^2 приводит к результату (2.3.30), соответствующему дифракции Фраунгофера. В приближении (2.3.25) подразумевается, что $x' \ll z'$, так что можно заменить z' на r везде, кроме показателя экспоненты, куда входит произведение малой величины x'/z' и большого числа k . Используя приближение

$$r = z' + \frac{x'^2}{2z'} \quad (2.3.31)$$

для членов в показателе экспонент перед интегралом и проводя интегрирование в (2.3.30), получим

$$\psi(x', z') = \frac{2e^{i\pi/4}}{k\sqrt{\lambda}} \psi_0 \frac{e^{-ikr}}{\sqrt{r}} \frac{\sin \pi \frac{d}{\lambda} \alpha}{\alpha}. \quad (2.3.32)$$

Поскольку $k = 2\pi/\lambda$, выражения (2.3.32) и (2.3.24) равны. Оба подхода, применяемые для получения поля дифракции на узкой щели, с учетом принятых приближений приводят к одному и тому же результату. Если использовать выражение (2.2.43) в качестве исходного, то метод дифракционного интеграла оказывается много проще метода плоских волн.

2.4. ДИФРАКЦИЯ НА КРУГЛОМ ОТВЕРСТИИ

Дифракционную задачу, рассмотренную в предыдущем разделе, оказалось возможным решить с помощью двумерных уравнений. В этом разделе рассмотрим задачу, которая может быть решена лишь с помощью трехмерного дифракционного интеграла (2.2.31)

Пусть поле плоской волны падает перпендикулярно поверхности круглого отверстия в непрозрачном экране.

Используя приближение (2.3.29) и пренебрегая членами x^2 и y^2 в показателе экспоненты, запишем выражение (2.2.31) в следующем приближенном виде:

$$\psi(x', y', z') = \frac{i}{\lambda z'} \psi_0 e^{-ik[z' + (x'^2 + y'^2)/2z']} \int_A e^{ik(x'x + y'y)/z'} dS. \quad (2.4.1)$$

Здесь предполагается, что экран расположен при $z = 0$. Чтобы оценить интеграл, введем полярную систему координат с полярной осью, параллельной оси z' . Тогда координаты x' , y' и x , y можно выразить следующим образом:

$$x' = \rho \cos \beta', \quad y' = \rho \sin \beta' \quad (2.4.2)$$

и

$$x = \sigma \cos \beta, \quad y = \sigma \sin \beta. \quad (2.4.3)$$

Комбинация, в которой эти величины входят в подынтегральное выражение функции (2.4.1), принимает вид

$$x'x + y'y = \rho\sigma \cos(\beta - \beta').$$

Располагая ось, от которой отсчитываются углы, таким образом, чтобы $\beta' = 0$, получим

$$x'x + y'y = \rho\sigma \cos \beta. \quad (2.4.4)$$

Чтобы упростить обозначения, введем величину r_0 , равную радиусу, получаемому из выражения (2.2.23) при $z = x = y = 0$. Таким образом, из выражения (2.4.1) имеем

$$\psi(x', y', z') = \frac{i}{\lambda z'} \psi_0 e^{-ikr_0} \int_0^a \sigma d\sigma \int_0^{2\pi} e^{ik(\rho/z')\sigma \cos \beta} d\beta. \quad (2.4.5)$$

Интеграл по β может быть выражен через функцию Бесселя нулевого порядка, которая имеет следующее интегральное представление [11]:

$$J_0(z) = \frac{1}{2\pi} \int_0^{2\pi} e^{iz \cos \beta} d\beta. \quad (2.4.6)$$

Используя хорошо известное соотношение [11] между J_0 и функцией Бесселя первого порядка J_1

$$\int z J_0(z) dz = z J_1(z), \quad (2.4.7)$$

из соотношения (2.4.5) получим

$$\psi(x', y', z') = i \frac{a}{\rho} \psi_0 e^{-ikr_0} J_1 \left(k \frac{\rho}{z'} a \right). \quad (2.4.8)$$

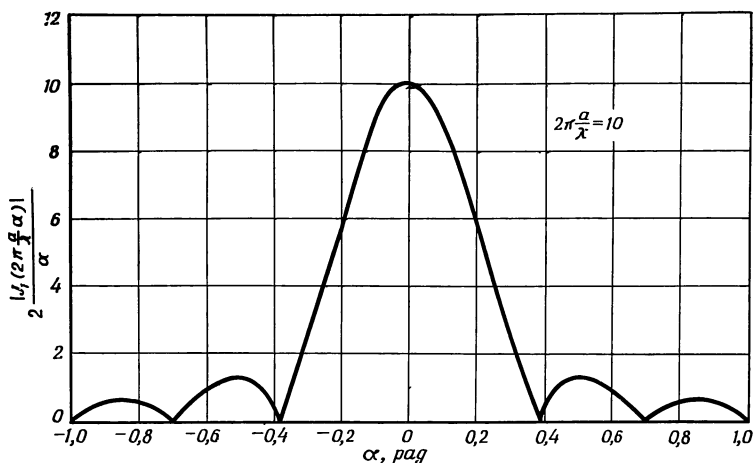
По аналогии с формулой (2.3.25) введем угол α , под которым точка поля x', y', z' видна из отверстия,

$$\alpha \approx \frac{\rho}{z'} \approx \frac{\rho}{r_0} \quad (2.4.9)$$

и получим окончательно

$$\psi(x', y', z') = ia\psi_0 \frac{e^{-ikr_0}}{r_0} \frac{1}{\alpha} J_1 \left(2\pi \frac{a}{\lambda} \alpha \right). \quad (2.4.10)$$

Дифракционная картина на больших расстояниях для поля за круглым отверстием (фиг. 2.4.1) очень похожа



Ф и г. 2.4.1. Дифракционная картина, создаваемая плоской волной, проходящей через круглое отверстие в непрозрачном экране.

на картину в случае щели (фиг. 2.3.1), если в обоих случаях рассматривается зависимость от угла α . Конечно, поле за щелью распространяется параллельно щели, образуя дифракционную картину лишь в направлении x . Дифракционная картина за круглым отверстием состоит из concentрических кругов. Полуширина главного лепест-

ка опять определяется углом, при котором поле обращается в нуль. Первый нуль функции Бесселя $J_1(z)$ имеет место при $z = 3,832$ [11], так что по аналогии с (2.3.28) получаем

$$\Delta\alpha = \frac{3,832}{2\pi} \frac{\lambda}{a} = 1,22 \frac{\lambda}{2a}. \quad (2.4.11)$$

Эта формула очень похожа на формулу (2.3.28). Если считать, что отверстие диаметром $2a$ соответствует щели шириной d , то получим, что половина углового размера $\Delta\alpha$ главного лепестка излучения дифракционного поля за круглым отверстием мало отличается от полуширины главного лепестка в случае щели.

2.5. ДИФРАКЦИОННАЯ РЕШЕТКА

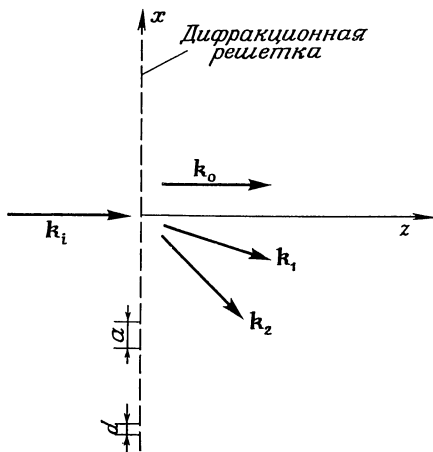
Дифракционные задачи, рассмотренные в двух предыдущих разделах, дают ценную вводную информацию о физике дифракции света и помогают научиться оперировать с полезными приближениями для нахождения дифракционных полей на больших расстояниях. Однако ни щель, ни круглая диафрагма сами по себе еще не являются устройствами, применяемыми на практике. Задача, рассматриваемая в этом разделе, гораздо важнее для практики.

Дифракционные решетки являются приборами, которые применяются для анализа частот, составляющих некогерентный свет. Простой пример дифракционной решетки приведен на фиг. 2.5.1. Непрозрачный экран имеет большое число щелей шириной d . Эти щели параллельны друг другу и расположены точно периодически на расстоянии a одна от другой. Полное число щелей $N + 1$ заполняет область экрана шириной D так, что имеет место соотношение

$$D = a(N + 1). \quad (2.5.1)$$

Чтобы упростить последующие выкладки, предположим, что $N/2$ — целое. Векторы \mathbf{k} , показанные на фиг. 2.5.1, представляют собой векторы распространения. Предполагается, что вектор \mathbf{k}_i направлен перпендикулярно поверхности экрана. Вектор \mathbf{k}_0 соответствует части падаю-

щей волны, которая за экраном продолжает распространяться в первоначальном направлении. Векторы \mathbf{k}_1 и \mathbf{k}_2 добавлены, чтобы изобразить рассеянный свет. Единственной (но существенной) разницей между этой задачей



Ф и г. 2.5.1. Иллюстрация дифракционной решетки, состоящей из многих тонких щелей в непрозрачном экране.

и задачей о щели, разобранный в разд. 2.3, является число существующих щелей. Чтобы решить задачу о дифракционной решетке и получить полное поле за решеткой, надо сложить рассеянный свет от каждой щели.

Используя формулу (2.3.31), заменим расстояние от центра каждой щели до точки x' , z' выражением

$$r_n = z' + \frac{(x' - x_n)^2}{2z'} = r_0 - \frac{x'x_n}{z'} + \frac{x_n^2}{2z'}, \quad (2.5.2)$$

где

$$r_0 = z' + \frac{x'^2}{2z'}. \quad (2.5.3)$$

Решим задачу о дифракционной решетке в приближении Фраунгофера точно так же, как решались дифракционные задачи в предыдущих двух разделах. Имеется в виду, что мы пренебрежем членом x_n^2 в выражении (2.5.2). Координата $x_n = na$ определяет центр n -й щели решетки.

Предполагается, что, хотя решетка содержит очень много щелей, ее полный размер мал по сравнению с расстоянием, на котором наблюдается дифракционное поле. Исходя из выражений (2.3.32) и (2.5.2), запишем поле на больших расстояниях от дифракционной решетки в приближении Фраунгофера:

$$\psi(x', z') = \frac{2e^{i\pi/4}}{k \sqrt{\lambda}} \psi_0 \frac{e^{-ikr_0}}{\sqrt{r_0}} \frac{\sin \frac{\pi d}{\lambda} \alpha}{\alpha} \sum_{n=-N/2}^{N/2} e^{ik\alpha na}. \quad (2.5.4)$$

Здесь было использовано соотношение (2.3.25), чтобы заменить отношение x' к z' углом α , под которым точка поля видна из центра дифракционной решетки.

Сумму в формуле (2.5.4) легко можно оценить, представив в виде

$$\begin{aligned} \sum_{n=-N/2}^{N/2} e^{ik\alpha na} &= e^{-ik\alpha(N/2)a} \frac{e^{ik\alpha(N+1)a} - 1}{e^{ik\alpha a} - 1} = \\ &= \frac{\sin \frac{1}{2} k\alpha (N+1) a}{\sin \frac{1}{2} k\alpha a}. \end{aligned} \quad (2.5.5)$$

Используя соотношения $k = 2\pi/\lambda$ и (2.5.1), преобразуем выражение (2.5.4) к виду

$$\psi(x', z') = \frac{\sqrt{\lambda}}{\pi} \psi_0 e^{i\pi/4} \frac{e^{-ikr_0}}{\sqrt{r_0}} \frac{\sin \frac{\pi d}{\lambda} \alpha}{\alpha} \frac{\sin \pi \frac{D}{\lambda} \alpha}{\sin \pi \frac{a}{\lambda} \alpha}. \quad (2.5.6)$$

Выражение для поля на больших расстояниях от дифракционной решетки содержит два существенных множителя. Первый соответствует действию каждой отдельной щели, тогда как второй член описывает влияние собственно решетки. Этот член имеет вид

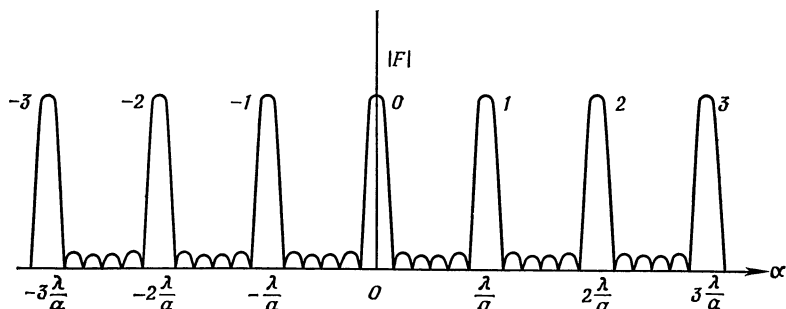
$$F(\alpha) = \frac{\sin \pi \frac{D}{\lambda} \alpha}{\sin \pi \frac{a}{\lambda} \alpha}; \quad (2.5.7)$$

график этой функции приведен на фиг. 2.5.2. В тех интервалах, где величина синуса в знаменателе велика,

функция $F(\alpha)$ очень быстро осциллирует с частотой $D/2\lambda$ (по переменной α). Из формулы (2.5.5) видно, что всякий раз, когда синус в знаменателе обращается в нуль при

$$\pi \frac{a}{\lambda} \alpha = m\pi \quad (2.5.8)$$

(m — целое), синус в числителе выражения (2.5.7) также



Ф и г. 2.5.2. Графическое представление функции $|F(\alpha)|$, задаваемой формулой (2.5.7), которая определяет дифракционную картину дифракционной решетки.

обращается в нуль. В точках (2.5.8) функция принимает значения

$$F_{\max} = \frac{D}{a} = N + 1. \quad (2.5.9)$$

Величина пика, характеризующего действие дифракционной решетки, пропорциональна числу щелей в решетке. Чтобы получить представление о полной картине действия решетки как функции α , график на фиг. 2.5.2 надо мысленно представить умноженным на функцию, соответствующую одной щели, которая показана на фиг. 2.3.4. Угловая полуширина лепестка для решетки может быть определена, так же как и в случае одной щели, точкой, в которой функция F обращается в нуль первый раз после прохождения через максимум. Таким образом, угловая полуширина лепестка для решетки равна

$$\Delta\alpha = \frac{\lambda}{D} = \frac{1}{N+1} \frac{\lambda}{a}. \quad (2.5.10)$$

Следовательно, полуширина лепестка для решетки обратно пропорциональна числу щелей в решетке.

Лепесток с максимумом при $\alpha = 0$ называется лепестком нулевого порядка. Он соответствует полю, которое по существу проходит через решетку, не дифрагируя. Лепесток m -го порядка, полученный из формулы (2.5.8), соответствует волне, выходящей из решетки под углом

$$\alpha_m = m \frac{\lambda}{a}. \quad (2.5.11)$$

Относительная полуширина лепестка m -го порядка равна

$$\frac{\Delta\alpha}{\alpha_m} = \frac{1}{m(N+1)}. \quad (2.5.12)$$

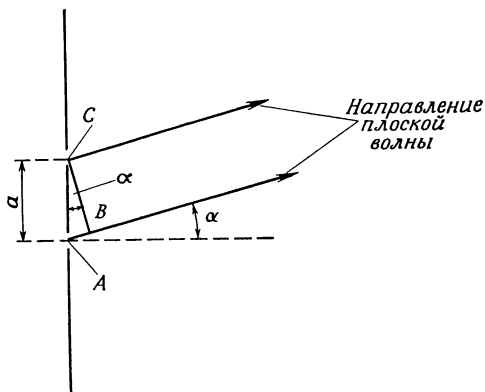
Поле на больших расстояниях от дифракционной решетки состоит из ряда почти плоских волн, распространяющихся в разных направлениях. Лепестки дифракционной картины различных порядков можно получить в виде линий на экране, сфокусировав на него картину дальнего поля с помощью линзы. Из формулы (2.5.11) видно, что направление каждого лепестка зависит от длины волны излучения. Следовательно, линза фокусирует свет каждой длины волны, содержащийся в падающей плоской волне, на свое место на экране (в свое пятно). Это свойство дифракционных решеток делает их очень полезными для спектрального анализа света. Спектральное разрешение дифракционной решетки может быть легко получено из соотношений (2.5.10) и (2.5.11). Будем считать две спектральные линии еще различимыми, если пик одной из них попадает на первый провал, ближайший к пику, другой линии. Это условие называется критерием Рэлея для разрешения оптических приборов. Выбранное правило означает, что изменение α_m для определенного изменения длины волны $\Delta\lambda$ должно быть равно $\Delta\alpha$ согласно (2.5.10). Такое требование сразу может быть выражено через относительное изменение длины волны, которое дифракционная решетка может разрешить,

$$\frac{\Delta\lambda}{\lambda} = \frac{1}{mN}. \quad (2.5.13)$$

В этом выражении мы пренебрегли единицей по сравнению с большим числом N . Очевидно, что разрешающая способ-

ность решетки улучшается при увеличении не только числа щелей, но и используемого порядка дифракции. Разрешающая способность прибора возрастает при работе в области более высоких порядков.

Соотношение (2.5.11) для угла, под которым появляется данный порядок, может быть понято с помощью очень



Ф и г. 2.5.3. Иллюстрация работы дифракционной решетки.

простых рассуждений. На фиг. 2.5.3 изображены две щели дифракционной решетки; показано направление распространения двух плоских волн, исходящих из каждой щели. Концепция плоской волны в дифракционной картине заимствована из разложения по плоским волнам произвольного поля [см. (1.3.23)]. Дифракционный лепесток может получиться только в том случае, если две волны, исходящие из соседних щелей, интерферируют, усиливаясь. Это означает, что точки *B* и *C*, показанные на фиг. 2.5.3, должны соответствовать одинаковым фазам. Это требование удовлетворяется, если отрезок *AB* равен целому числу длин волн

$$AB = m\lambda, \quad (2.5.14)$$

или

$$a \sin \alpha = m\lambda. \quad (2.5.15)$$

В приближении малых углов, используемом при выводе, соотношения (2.5.11) и (2.5.15) совпадают.

Дифракционную решетку, изображенную на фиг. 2.5.1, можно описать, введя функцию пропускания T следующего вида:

$$T = \begin{cases} 1, & x_n - \frac{d}{2} < x < x_n + \frac{d}{2}, \\ 0, & x_n - \left(a - \frac{d}{2}\right) < x < x_n - \frac{d}{2}. \end{cases} \quad (2.5.16)$$

Соотношение (2.5.16) означает, что $T = 0$ на непрозрачной части решетки и $T = 1$ на щелях. Хотя мы и не следовали этим путем, выражение (2.5.4) может быть получено подстановкой

$$\psi(x, z) = T\psi_0 \quad (2.5.17)$$

в подынтегральное выражение интеграла (2.2.43).

Дифракционную решетку можно получить, придав функции пропускания не только форму, описываемую соотношением (2.5.16), но и в других случаях, когда функция пропускания периодична. Особый интерес для голографии имеет решетка, функция пропускания которой описывается синусоидальным законом

$$T = 1 + t \cos 2\pi \frac{x}{a}. \quad (2.5.18)$$

Подставив выражения (2.3.29), (2.5.17) и (2.5.18) в (2.2.43), получим

$$\begin{aligned} \psi(x', z') = & \frac{\psi_0}{V\lambda} \frac{e^{-ikr_0}}{Vr_0} e^{i\pi/4} \times \\ & \times \int_{-(N/2)a}^{(N/2)a} \left(1 + t \cos 2\pi \frac{x}{a}\right) e^{ikh\alpha x} dx. \end{aligned} \quad (2.5.19)$$

Можно сделать обычные приближения, применимые к случаю дифракции Фраунгофера, и использовать выражение (2.3.25) в показателе экспоненты под знаком интеграла. Легче всего вычислить интеграл, выразив косинус

через экспоненты. В результате имеем

$$\psi(x', z') = \frac{\psi_0}{\sqrt{\lambda}} e^{i\pi/4} \frac{e^{-ikr_0}}{\sqrt{r_0}} \times \\ \times \left[2 \frac{\sin k\alpha a \frac{N}{2}}{k\alpha} + t \frac{\sin(k\alpha a + 2\pi) \frac{N}{2}}{k\alpha + \frac{2\pi}{a}} + t \frac{\sin(k\alpha a - 2\pi) \frac{N}{2}}{k\alpha - \frac{2\pi}{a}} \right]. \quad (2.5.20)$$

Действие такой синусоидальной решетки существенно отличается от действия щелевой решетки (2.5.6). В то время как щелевая решетка имеет бесконечно много лепестков, дифракционная картина (2.5.20) имеет только три лепестка. Лепесток нулевого порядка соответствует первому члену в скобках. Этот член просто представляет дифракционную картину щели, размер которой равен размеру всей решетки. Этот лепесток очень узкий и имеет максимум при $\alpha = 0$. Каждый из других членов в скобках соответствует одному боковому лепестку. Существенной разницей между двумя решетками является появление множителя $\sin \pi (a/\lambda) \alpha$ в знаменателе выражения для щелевой решетки. Этот множитель отсутствует в выражении для синусоидальной решетки. Знаменатели в случае синусоидальной решетки обращаются в нуль только в одной точке, в которой и возникает лепесток. Но они не обращаются в нуль периодически, как в случае щелевой решетки. Этот пример показывает, что можно сконструировать решетку, имеющую только по одному боковому лепестку с каждой стороны лепестка нулевого порядка.

Завершим рассмотрение дифракционных решеток анализом фазовых решеток. Две рассмотренные выше решетки модулировали амплитуду падающей волны по закону, заданному периодической функцией пропускания. Можно получить дифракционную решетку, модулируя не амплитуду, а фазу падающей плоской волны. Это можно сделать, помещая на пути волны совершенно прозрачный предмет. Фаза прошедшей волны зависит от длины оптического пути света, проходящего через определенную часть прозрачного предмета. Изготовив стеклянную пластинку периодически меняющейся толщины, получим периодиче-

скую фазовую модуляцию падающей плоской волны. Фазовую решетку можно также осуществить с помощью звуковых волн в жидкости. В комплексных обозначениях фазовую модуляцию можно представить просто, введя фазовый множитель для функции пропускания:

$$T = e^{i[1+t \cos 2\pi(x/a)]}. \quad (2.5.21)$$

Подставив это выражение в соотношения (2.5.17) и (2.2.43), получим

$$\begin{aligned} \psi(x', z') = & \frac{\psi_0}{\sqrt{\lambda}} e^{i\pi/4} \frac{e^{-ikr_0}}{\sqrt{r_0}} \times \\ & \times \int_{-D/2}^{D/2} \exp \left[i \left(1 + t \cos 2\pi \frac{x}{a} + k\alpha x \right) \right] dx. \end{aligned} \quad (2.5.22)$$

Интеграл в этом выражении оценить гораздо труднее, чем интегралы, вычисленные ранее. Чтобы получить представление о поведении фазовой решетки, начнем с приближенного решения. Положим $t \ll 1$ и пренебрежем членами второго и более высокого порядка по t . Это приближение позволяет получить следующее выражение:

$$\begin{aligned} \psi(x', z') = & \frac{\psi_0}{\sqrt{\lambda}} e^{i\pi/4} \frac{e^{-ikr_0}}{\sqrt{r_0}} e^i \times \\ & \times \int_{-D/2}^{D/2} \left(1 + it \cos 2\pi \frac{x}{a} \right) e^{ik\alpha x} dx. \end{aligned} \quad (2.5.23)$$

За исключением фазового множителя перед интегралом и того факта, что вместо t в выражении появилось t , умноженное на i , выражение (2.5.23) идентично формуле (2.5.19) для синусоидальной амплитудной решетки. Выражение (2.5.20) дает сразу решение задачи о фазовой решетке после замены t на it и умножения всего выражения на несущественный фазовый множитель e^i . Таким образом, видно, что при малых значениях t имеется только три лепестка: лепесток нулевого порядка, соответствующий щели шириной, равной полной ширине решетки, и два боковых лепестка первого порядка по обе стороны щели. Однако этот результат не точный. Если в выражении функции пропускания (2.5.21) сохранить члены более

высокого порядка, то получим лепестки дополнительных дифракционных порядков. Согласно формуле (2.5.20), максимуму лепестка первого порядка соответствует равенство

$$\frac{ak\alpha}{2\pi} = \pm 1,$$

которое через длину волны выражается следующим образом:

$$\alpha \frac{a}{\lambda} = \pm 1. \quad (2.5.24)$$

Главные максимумы произвольных порядков соответствуют

$$\alpha \frac{a}{\lambda} = m, \quad (2.5.25)$$

где m — любое положительное или отрицательное целое число или нуль. Чтобы проверить это утверждение и получить выражение, которое позволит определить амплитуды максимумов в каждом порядке, а также ширину лепестка, положим

$$v = \alpha \frac{a}{\lambda} = -m + \varepsilon, \quad (2.5.26)$$

где m — целое, а

$$\varepsilon \ll 1. \quad (2.5.27)$$

Вводя новую переменную

$$2\pi \frac{x}{a} = \theta, \quad (2.5.28)$$

получим из формулы (2.5.22)

$$\begin{aligned} \psi(x', z') = a \frac{\psi_0}{\sqrt{\lambda}} e^{i\pi/4} \frac{e^{-ikr_0}}{\sqrt{r_0}} \frac{e^i}{2\pi} \times \\ \times \int_{-\pi(D/a)}^{\pi(D/a)} e^{i(v\theta + t \cos \theta)} d\theta. \end{aligned} \quad (2.5.29)$$

Разделим пределы интегрирования на участки длиной 2π . Используя выражение (2.5.1), получим

$$\begin{aligned} I = \int_{-\pi(D/a)}^{\pi(D/a)} e^{i(v\theta + t \cos \theta)} d\theta = \\ = \sum_{n=-N/2}^{N/2} \int_{(2n-1)\pi}^{(2n+1)\pi} e^{i(v\theta + t \cos \theta)} d\theta. \end{aligned} \quad (2.5.30)$$

Используя соотношения (2.5.26) и (2.5.27), можно записать приближенное равенство

$$I = \sum_{n=-N/2}^{N/2} e^{i2n\pi\varepsilon} \int_{(2n-1)\pi}^{(2n+1)\pi} e^{i(-m\theta+t \cos \theta)} d\theta. \quad (2.5.31)$$

Из-за малости величины ε функция $e^{i\varepsilon\theta}$ почти постоянна в пределах интегрирования и вынесена из-под знака интеграла. Оставшееся в (2.5.31) подынтегральное выражение является периодической функцией с периодом 2π . Пределы интегрирования интеграла по полному периоду периодической функции могут быть сдвинуты без изменения величины интеграла. Фактически надо только заменить θ на $\theta' + 2n\pi$, чтобы получить

$$I = \sum_{n=-N/2}^{N/2} e^{i2n\pi\varepsilon} \int_{-\pi}^{\pi} e^{i(-m\theta'+t \cos \theta')} d\theta'. \quad (2.5.32)$$

Интеграл в этом выражении является функцией Бесселя. Одно из интегральных представлений функции Бесселя [11] имеет вид

$$J_n(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(-n\theta+z \sin \theta)} d\theta. \quad (2.5.33)$$

Вводя

$$\theta = \theta' + \frac{\pi}{2}$$

и сдвигая область интегрирования периодической функции, чтобы снова ее сделать симметричной, получим

$$J_n(z) = \frac{e^{-in(\pi/2)}}{2\pi} \int_{-\pi}^{\pi} e^{i(-n\theta+z \cos \theta)} d\theta. \quad (2.5.34)$$

Выражение (2.5.34) позволяет переписать соотношение (2.5.32) в виде

$$I = 2\pi e^{im(\pi/2)} J_m(t) \sum_{n=-N/2}^{N/2} e^{i2n\pi\varepsilon}. \quad (2.5.35)$$

Суммирование уже было проведено в (2.5.5), так что из соотношения (2.5.29) получаем

$$\psi(x', z') = a \frac{\psi_0}{\sqrt{\lambda}} e^{i[(m+1/2)\pi/2+1]} \frac{e^{-ikr_0}}{\sqrt{r_0}} J_m(t) \times \\ \times \frac{\sin \pi \varepsilon (N+1)}{\pi \varepsilon}. \quad (2.5.36)$$

Здесь $\sin(\pi \varepsilon)$ заменен на $\pi \varepsilon$.

Из соотношения (2.5.36) видно, что амплитуды рассеянного излучения становятся большими только при $\varepsilon = 0$. В этой точке

$$\lim_{\varepsilon \rightarrow 0} \frac{\sin \pi \varepsilon (N+1)}{\pi \varepsilon} = N+1. \quad (2.5.37)$$

Как было указано выше, главные дифракционные максимумы приходятся на целые значения α (a/λ) в соответствии с формулой (2.5.26). Фазовая решетка имеет бесконечно много максимумов, амплитуда которых спадает как функция Бесселя m -го порядка

$$J_m(t). \quad (2.5.38)$$

В табл. 2.5.1 приведены некоторые значения $J_m(t)$.

Таблица 2.5.1

Выборочные значения функции Бесселя $J_m(t)$ для различных значений порядка m и аргумента t

m	t				
	0,2	0,4	0,6	0,8	1,0
0	$9,90 \cdot 10^{-1}$	$9,60 \cdot 10^{-1}$	$9,12 \cdot 10^{-1}$	$8,46 \cdot 10^{-1}$	$7,65 \cdot 10^{-1}$
1	$9,95 \cdot 10^{-2}$	$1,96 \cdot 10^{-1}$	$2,87 \cdot 10^{-1}$	$3,69 \cdot 10^{-1}$	$4,40 \cdot 10^{-1}$
2	$4,98 \cdot 10^{-3}$	$1,97 \cdot 10^{-2}$	$4,37 \cdot 10^{-2}$	$7,58 \cdot 10^{-2}$	$1,15 \cdot 10^{-2}$
3	$1,66 \cdot 10^{-4}$	$1,32 \cdot 10^{-3}$	$4,40 \cdot 10^{-3}$	$1,02 \cdot 10^{-2}$	$1,96 \cdot 10^{-2}$
4	$4,16 \cdot 10^{-6}$	$6,61 \cdot 10^{-5}$	$3,31 \cdot 10^{-4}$	$1,03 \cdot 10^{-3}$	$2,48 \cdot 10^{-3}$
5	$8,32 \cdot 10^{-8}$	$2,65 \cdot 10^{-6}$	$1,99 \cdot 10^{-5}$	$8,31 \cdot 10^{-5}$	$2,50 \cdot 10^{-4}$
6	$1,39 \cdot 10^{-9}$	$8,84 \cdot 10^{-8}$	$1,00 \cdot 10^{-6}$	$5,56 \cdot 10^{-6}$	$2,09 \cdot 10^{-5}$

Амплитуды дифракционных максимумов не только уменьшаются при возрастании порядка m , но они также

зависят от глубины модуляции t . Для $t \ll 1$ приближение (2.5.23) справедливо при $m = \pm 1$.

Читатель, вероятно, заметил очень тесную связь между угловой характеристикой дифракционной решетки и частотным спектром модулированных радиоволн. Случай простой синусоидальной амплитудной модуляции соответствует случаю решетки с синусоидальной функцией пропускания. Случай фазовой решетки соответствует синусоидальной фазовой (или частотной) модуляции в радиотехнике. Наконец, простая щелевая решетка соответствует импульсной амплитудной модуляции несущей высокой частоты. Это соответствие угла отклонения частотному спектру в процессах, зависящих от времени, не ограничивается дифракционными решетками. Оно находит еще более полезное применение при рассмотрении свойств преобразования в линзах, которые будут рассматриваться в одной из последующих глав. Соответствие угла отклонения частотному спектру в процессах, зависящих от времени, обусловлено тем фактом, что в условиях дифракции Фраунгофера дифракционный интеграл принимает форму интеграла Фурье, как это легко можно видеть из соотношения (2.5.19) или (2.5.22). Свойства преобразований Фурье для дифракционных процессов, в частности при установке линз на пути светового пучка, будут рассмотрены в главе о линзах.

2.6. ДИФРАКЦИЯ БРЭГГА. ТЕОРИЯ ВОЗМУЩЕНИЙ

В предыдущем разделе рассматривались двумерные дифракционные решетки нулевой толщины. В этом разделе будут исследованы толстые фазовые решетки.

Пусть показатель преломления плоской пластины диэлектрического вещества меняется по закону

$$n = n_0 + \eta \cos(\beta \cdot \mathbf{r} + \phi). \quad (2.6.1)$$

Показатель преломления синусоидально модулирован в пространстве. Области постоянного показателя преломления лежат в плоскостях, перпендикулярных вектору β . Расстояние D между двумя последовательными максимумами n (или любыми другими фиксированными значениями)

дается формулой

$$D = \frac{2\pi}{\beta}, \quad (2.6.2)$$

где β — модуль вектора β .

Распределение (2.6.1) может быть реализовано с помощью звуковых волн в диэлектрической среде. Сжатия и разрежения, возникающие в звуковом поле, вызывают такого рода изменения n . Изменения показателя преломления в звуковой волне проходили бы через вещество со скоростью звука. Это движение волн не нашло отражения в формуле (2.6.1).

Можно рассматривать это выражение как моментальный снимок трехмерного звукового поля. Однако распределение показателя преломления обсуждаемого здесь типа может быть получено путем фотографирования трехмерной интерференционной картины на фотоэмульсию. После того как эмульсия будет проявлена и атомы серебра обесцвечены, в желатине останется распределение показателя преломления, подобное (2.6.1). Такое распределение, конечно, стационарно и не перемещается в веществе. Оно хорошо соответствует нашей математической модели.

Распределение показателя n (2.6.1) соответствует трехмерной фазовой решетке в диэлектрических материалах. Эта решетка обладает свойствами, подобными свойствам двумерной фазовой решетки, рассмотренной в предыдущем разделе. Действительно, если слой вещества, из которого состоит решетка, имеет не слишком большую толщину и если падающая волна входит в вещество перпендикулярно, а вектор β направлен параллельно поверхности вещества, можно в хорошем приближении говорить, что решетка двумерная.

Геометрия общей задачи представлена на фиг. 2.6.1, где, кроме вектора падающей плоской волны \mathbf{k}_i и вектора рассеянной плоской волны \mathbf{k}_s , приведен также вектор β . Линии постоянного показателя преломления показаны наклонной штриховкой.

Решим эту задачу рассеяния методом теории возмущений. Будем считать, что

$$\eta \ll 1 \quad (2.6.3)$$

и запишем выражение для квадрата показателя преломления

$$n^2 \approx n_0^2 + 2\eta n_0 \cos(\beta \cdot \mathbf{r} + \varphi) = n_0^2 + \Delta\epsilon. \quad (2.6.4)$$

Приведенное волновое уравнение (2.2.2) имеет вид ¹⁾

$$\nabla^2 \psi + n^2 k_0^2 \psi = 0. \quad (2.6.5)$$

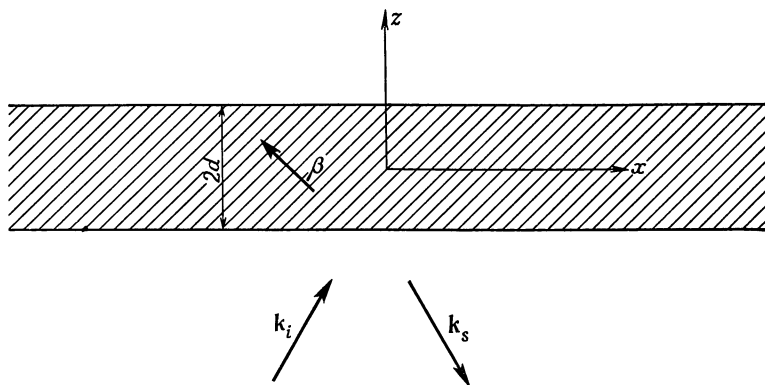
Коэффициент k_0 — постоянная распространения свободного пространства

$$k_0 = \omega \sqrt{\epsilon_0 \mu_0}. \quad (2.6.6)$$

Пусть падает плоская волна вида

$$\psi_i = A e^{-i\mathbf{k}_i \cdot \mathbf{r}} \quad (2.6.7)$$

при $k_i = n_0 k_0$. Предполагается, что область вне объема с синусоидально модулированным показателем преломле-



Ф и г. 2.6.1. Схема дифракции Брэгга.

Заштрихованная полоса показывает область с периодически меняющимся показателем преломления. \mathbf{k}_i и \mathbf{k}_s — векторы распространения падающей и рассеянной волн, вектор β нормален плоскостям постоянных значений показателя преломления.

ния имеет среднее значение показателя преломления n_0 . Это позволяет избежать ненужных усложнений, возникаю-

¹⁾ Скалярная волновая теория, которая используется здесь, применима непосредственно к волновому полю, которое удовлетворяет условию $\mathbf{E} \cdot \beta = 0$. Для других поляризаций скалярная теория не обязательно приведет к тому же результату, что и векторная теория.

щих из-за отражения и преломления волн на границе раздела двух сред. Эти явления на плоской границе раздела диэлектриков уже были рассмотрены в разд. 1.6.

Кроме падающей волны, имеется рассеянная волна, которую представим в виде интеграла Фурье

$$\psi_s = \int B(\mathbf{k}_s) e^{-i\mathbf{k}_s \cdot \mathbf{r}} d^3\mathbf{k}_s. \quad (2.6.8)$$

Здесь $d^3\mathbf{k}_s$ является сокращенной записью произведения:

$$d^3\mathbf{k}_s = dk_{sx} dk_{sy} dk_{sz}. \quad (2.6.9)$$

Аргумент \mathbf{k}_s функции B представляет в сокращенном виде ее зависимость от трех переменных k_{sx} , k_{sy} , k_{sz} . Интеграл берется по всем трем переменным от $-\infty$ до $+\infty$. Выражение (2.6.8) отличается от решения волнового уравнения (1.3.23). Здесь мы считаем, что все три компоненты вектора \mathbf{k}_s изменяются независимо во всей области интегрирования. Все ограничения, которые волновое уравнение накладывает на функцию, должны быть выражены с помощью функциональной зависимости B .

Полное поле является суммой падающей и отраженной волн:

$$\psi = \psi_i + \psi_s. \quad (2.6.10)$$

Подставим выражения (2.6.4), (2.6.7), (2.6.8) и (2.6.10) в приведенное волновое уравнение (2.6.5)

$$\begin{aligned} & (n_0^2 k_0^2 - k_i^2) A e^{-i\mathbf{k}_i \cdot \mathbf{r}} + \Delta \epsilon k_0^2 A e^{-i\mathbf{k}_i \cdot \mathbf{r}} + \\ & + \int (n_0^2 k_0^2 - k_s^2) B(\mathbf{k}_s) e^{-i\mathbf{k}_s \cdot \mathbf{r}} d^3\mathbf{k}_s + \\ & + \int \Delta \epsilon k_0^2 B(\mathbf{k}_s) e^{-i\mathbf{k}_s \cdot \mathbf{r}} d^3\mathbf{k}_s = 0. \end{aligned} \quad (2.6.11)$$

Падающая волна является решением невозмущенного волнового уравнения. Следовательно, первый член в уравнении (2.6.11) исключается. Пока уравнение (2.6.11) является точным. Однако мы хотим решить эту задачу методом теории возмущений и ограничиться приближением первого порядка. Поскольку $\Delta \epsilon$ и B являются величинами, малыми в первом порядке, можно пренебречь их произведением и опустить последний член в уравнении (2.6.11). Умножим оставшиеся слагаемые на $\exp(i\mathbf{k}_s \cdot \mathbf{r})$

и проинтегрируем по всему пространству. Интеграл от экспоненциальной функции сводится в результате к δ -функции

$$\begin{aligned} \int e^{i(\mathbf{k}'_s - \mathbf{k}_s) \cdot \mathbf{r}} d^3\mathbf{r} &= \\ &= (2\pi)^3 \delta(k'_{sx} - k_{sx}) \delta(k'_{sy} - k_{sy}) \delta(k'_{sz} - k_{sz}) = \\ &= (2\pi)^3 \delta(\mathbf{k}'_s - \mathbf{k}_s). \end{aligned} \quad (2.6.12)$$

Здесь $\delta(\mathbf{k}' - \mathbf{k}_s)$ является сокращенным обозначением трехмерной δ -функции.

Используя указанную выше процедуру, получим из уравнения (2.6.11)

$$\begin{aligned} B(\mathbf{k}_s) &= \frac{\eta k_0^2 A n_0}{(2\pi)^3 (k_s^2 - n_0^2 k_0^2)} \times \\ &\times \int_V (e^{i[(\mathbf{k}_s + \boldsymbol{\beta} - \mathbf{k}_i) \cdot \mathbf{r} + \phi]} + e^{i[(\mathbf{k}_s - \boldsymbol{\beta} - \mathbf{k}_i) \cdot \mathbf{r} - \phi]}) d^3\mathbf{r}, \end{aligned} \quad (2.6.13)$$

куда подставлено выражение для $\Delta \varepsilon$ из (2.6.4). Поскольку $\Delta \varepsilon$ обращается в нуль вне заштрихованного на фиг. 2.6.1 объема, то интегрирование в (2.6.13) распространяется только на область взаимодействия, имеющую объем V . Под знаком интеграла стоят осциллирующие функции, которые дают существенный вклад в интеграл, если их аргументы обращаются в нуль. Легко видеть, что это имеет место при

$$\mathbf{k}_s = \mathbf{k}_i \pm \boldsymbol{\beta}. \quad (2.6.14)$$

Это важное соотношение известно как условие Брэгга. Предполагается, что объем взаимодействия простирается до бесконечности в направлениях x и y . Следовательно, интегрирование по x и y приводит к δ -функциям, так что выражение (2.6.13) можно записать в виде

$$\begin{aligned} B(\mathbf{k}_s) &= \frac{\eta n_0 k_0^2 A}{\pi (k_s^2 - n_0^2 k_0^2)} \left[\delta(k_{sx} + \beta_x - k_{ix}) \delta(k_{sy} + \beta_y - k_{iy}) \times \right. \\ &\times \frac{\sin(k_{sz} + \beta_z - k_{iz}) d}{k_{sz} + \beta_z - k_{iz}} e^{i\phi} + \delta(k_{sx} - \beta_x - k_{ix}) \delta(k_{sy} - \beta_y - k_{iy}) \times \\ &\times \left. \frac{-\sin(k_{sz} - \beta_z - k_{iz}) d}{k_{sz} - \beta_z - k_{iz}} e^{-i\phi} \right]. \end{aligned} \quad (2.6.15)$$

Рассеянную волну можно получить из выражений (2.6.8) и (2.6.15)

$$\begin{aligned} \psi = & \eta n_0 k_0^2 A \frac{1}{\pi} \left[e^{i\phi} e^{-i(k_{sx}^{(-)}x + k_{sy}^{(-)}y)} \times \right. \\ & \times \int_{-\infty}^{+\infty} \frac{e^{-ik_{sz}z}}{k_s^{(-)2} - n_0^2 k_0^2} \frac{\sin(k_{sz} + \beta_z - k_{iz})d}{k_{sz} + \beta_z - k_{iz}} dk_{sz} + \\ & + e^{-i\phi} e^{-i(k_{sx}^{(+)}x + k_{sy}^{(+)}y)} \int_{-\infty}^{+\infty} \frac{e^{-ik_{sz}z}}{k_s^{(+2} - n_0^2 k_0^2} \times \\ & \left. \times \frac{\sin(k_{sz} - \beta_z - k_{iz})d}{k_{sz} - \beta_z - k_{iz}} dk_{sz} \right]. \quad (2.6.16) \end{aligned}$$

Верхние индексы «+» и «—» у компонент \mathbf{k}_s указывают, что должны быть подставлены значения k_s , соответствующие знакам плюс и минус в выражении (2.6.14). Таким образом, компоненты x и y вектора \mathbf{k}_s определяются выражением (2.6.14), тогда как его z -компонента произвольна.

Интегралы в (2.6.16) легко могут быть сведены к интегралам по контуру. Путь интегрирования проходит вдоль реальной оси k_{sz} от $-\infty$ до $+\infty$. Каждый интеграл имеет два полюса, лежащие на пути интегрирования. Однако эти полюсы можно сместить с реальной оси на комплексную плоскость, используя физические аргументы. Мы все время предполагали, что показатель преломления — величина вещественная. Однако это идеализация. Все реальные материалы ослабляют электромагнитные волны, распространяющиеся в них. Можно не вводить предположения об идеальном диэлектрике без потерь. Пусть плоская волна распространяется в идеальном диэлектрике вдоль оси z . При выборе зависимости поля от времени в виде (2.2.1) можно записать плоскую волну в виде

$$\psi = \psi_0 e^{-i n_0 k_0 z}. \quad (2.6.17)$$

В диэлектрическом материале с потерями волна будет ослабляться. Можно описать затухающую волну, введя комплексную диэлектрическую проницаемость

$$n_0 = n_r - i n_i. \quad (2.6.18)$$

Для положительных значений n_r и n_i формула (2.6.17) описывает плоскую волну, распространяющуюся в поло-

жительном направлении оси z и экспоненциально затухающую по закону

$$\psi = e^{-n_i k_0 z} \psi_0 e^{-i n_r k_0 z}. \quad (2.6.19)$$

Эти рассуждения показывают, как модифицировать выражение (2.6.16), чтобы учесть небольшие потери в диэлектрике. В конце вычислений мы опять положим мнимую часть диэлектрической проницаемости равной нулю.

Подстановка выражения (2.6.18) для n_0 в (2.6.16) смещает полюсы подынтегрального выражения с реальной оси. Путь интегрирования не проходит теперь ни через один полюс, так что интегралы не имеют особенностей и легко могут быть вычислены. Опуская верхние индексы у k_s , можно вычислить оба интеграла одновременно. Рассмотрим знаменатель

$$\begin{aligned} D &= k_s^2 - n_0^2 k_0^2 = k_{sz}^2 - (n_0^2 k_0^2 - k_{sx}^2 - k_{sy}^2) = \\ &= (k_{sz} - k'_{sz})(k_{sz} + k'_{sz}), \end{aligned} \quad (2.6.20)$$

где

$$k'_{sz} = \sqrt{n_0^2 k_0^2 - k_{sx}^2 - k_{sy}^2}. \quad (2.6.21)$$

Мы считаем, что $n_i \ll 1$, и можем записать в хорошем приближении

$$k'_{sz} = \sqrt{n_r^2 k_0^2 - k_{sx}^2 - k_{sy}^2} - \frac{i n_i n_r k_0^2}{\sqrt{n_r^2 k_0^2 - k_{sx}^2 - k_{sy}^2}}. \quad (2.6.22)$$

Величину n_i можно взять как угодно малой. В пределе при $n_i \rightarrow 0$ выражение (2.6.22) становится точным. Путь интегрирования и положение полюсов схематически показаны на фиг. 2.6.2. Пунктирные кривые изображают две бесконечно большие полуокружности, которые замыкают путь интегрирования вдоль реальной оси.

Нас интересуют лишь значения z вне трехмерной дифракционной решетки. Следовательно, достаточно ограничиться случаем

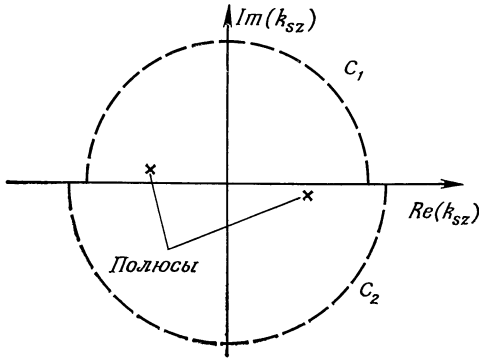
$$|z| > d. \quad (2.6.23)$$

В области, удовлетворяющей соотношению (2.6.23), экспоненциальный множитель $\exp(-ik_{sz}z)$ определяет сходящийся либо расходящийся характер подынтегрального выражения. Необходимо разграничить эти два случая.

Сначала предположим, что z положительно:

$$z > d. \quad (2.6.24)$$

В этом случае экспоненциальный множитель обращается в нуль на кривой C_2 , показанной на фиг. 2.6.2. Добавление интегрирования вдоль C_2 не изменит величины интег-



Фиг. 2.6.2. Путь интегрирования в комплексной плоскости k_{sz} .

рала. Так как нет других полюсов, кроме указанного, можно использовать интегральную теорему Коши, чтобы стянуть путь интегрирования в бесконечно малую окружность. Тогда значение интеграла равно вычету. Пусть величина z отрицательна:

$$z < -d. \quad (2.6.25)$$

Путь интегрирования вдоль C_1 не дает вклада по аналогичным причинам. Замена интегралов в (2.6.16) контурными приводит к следующему результату:

$$\psi_s = -i\eta n_0 k_0^2 A \left[\frac{e^{i\phi}}{k_{sz}} e^{-i\mathbf{k}_s^{(-)} \cdot \mathbf{r}} \frac{\sin(k_{sz} + \beta_z - k_{iz})d}{k_{sz} + \beta_z - k_{iz}} + \frac{e^{-i\phi}}{k_{sz}} e^{-i\mathbf{k}_s^{(+)} \cdot \mathbf{r}} \frac{\sin(k_{sz} - \beta_z - k_{iz})d}{k_{sz} - \beta_z - k_{iz}} \right]. \quad (2.6.26)$$

Величина k_{sz} , входящая в выражение (2.6.26), задается формулой (2.6.21). Штрих в индексе отброшен для случая сходимости. Выражение (2.6.26) является решением

дифракционной задачи Брэгга в первом порядке приближения теории возмущений.

Выражение (2.6.26) указывает на несколько интересных свойств рассеяния Брэгга, к рассмотрению которых мы перейдем.

Условие Брэгга, сформулированное соотношением (2.6.14), не удовлетворяется точно для z -компоненты этого векторного уравнения. Этот факт становится очевидным, если рассмотреть синусоидальные множители в (2.6.26). Соотношение (2.6.14) для x - и y -компонент является точным, потому что по предположению решетка бесконечна в направлениях x и y . Устремляя $d \rightarrow \infty$, получим δ -функции вместо синусоидальных множителей [см. (2.3.8)], и z -компонента в условии Брэгга также становится точной. Для конечной величины d имеется небольшое отклонение от точного выполнения условия (2.6.14) для z -компоненты. Члены с синусоидальными функциями оказываются уменьшенными, как показано на фиг. 2.3.1. Пик увеличивается при возрастании d и достигает бесконечности в пределе при $d \rightarrow \infty$. Ширина пика уменьшается по мере увеличения его высоты. В большинстве случаев $d \gg \lambda$ и для z -компоненты условие Брэгга удовлетворяется с хорошей точностью.

Однако мы получили, что, кроме условия Брэгга, должно удовлетворяться условие (2.6.21). Это соотношение может быть переписано в виде

$$n_0^2 k_0^2 = k_s^2. \quad (2.6.27)$$

Это важное соотношение устанавливает, что рассеянная волна распространяется с той же скоростью, что и падающая волна, и что обе они распространяются с фазовой скоростью плоских волн в среде с показателем преломления n_0 . Соотношение (2.6.27) и условие (2.6.14) для x - и y -компонент определяют все компоненты вектора \mathbf{k}_s .

Рассеянная волна является плоской. Ее амплитуда определяется точностью, с которой удовлетворяется по компонентам соотношение (2.6.14), и выражается через синусоидальные функции, входящие в выражение (2.6.26). Если условие Брэгга (2.6.14) выполняется даже для z -компоненты, амплитуда рассеянной волны достигает макси-

муна:

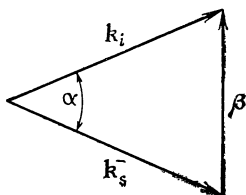
$$\psi_s = -i\eta n_0 \frac{k_0^2 d}{k_{sz}} A e^{-ik(\pm)\cdot r} e^{\mp i\phi}. \quad (2.6.28)$$

Амплитуда рассеянной волны, конечно, пропорциональна амплитуде падающей волны, а также изменению показателя преломления η и толщине d трехмерной решетки. Интересным обстоятельством является зависимость амплитуды от обратной величины k_{sz} . При уменьшении z -компоненты вектора распространения амплитуда возрастает. Это можно понять, если учесть, что волна проходит большее расстояние внутри решетки, когда вектор распространения параллелен границе решетки. Чем дальше рассеянная волна остается внутри решетки, тем больше она взаимодействует с падающей волной и тем большая мощность может быть передана от одной волны другой. В предельном случае $k_{sz} = 0$ теория возмущений неприменима. Эта теория не учитывает уменьшения амплитуды падающей волны. Точная теория, развитая в разд. 2.7, учитывает потери мощности падающей волны, которые приводят к уменьшению ее амплитуды.

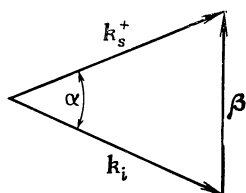
Условие Брэгга (2.6.14) и соотношение (2.6.27) показывают, как связаны направления падающей и рассеянной волн. Оба знака в соотношении (2.6.14) допустимы, так что имеем две диаграммы, изображенные на фиг. 2.6.3 и 2.6.4. Обе диаграммы представляют собой равнобедренные треугольники. Вектор β направлен перпендикулярно плоскостям постоянного показателя преломления. Падающая волна рассеивается этими плоскостями. Две диаграммы получаются потому, что волна может падать на решетку сверху или снизу. Существенным моментом в дифракции Брэгга является тот факт, что рассеяние происходит как отражение света от плоских диэлектрических поверхностей. Если рассеяние имеет место, то оно происходит зеркально в соответствии с законом отражения света от зеркал. Однако условия зеркального отражения недостаточно, чтобы появилась рассеянная волна. Рассеяние имеет место, если волна падает на отражающие плоскости постоянного показателя преломления под углом $\alpha/2$. Этот угол легко получить из диаграмм:

$$\sin \frac{\alpha}{2} = \frac{\beta}{2k_i} = \frac{\lambda}{2D}. \quad (2.6.29)$$

Это соотношение также известно как условие Брэгга. Для получения правой части было использовано соотношение (2.6.2). Уже отмечалось, что условие Брэгга выполняется точно, только когда толщина дифракционной решетки становится бесконечно большой. Для решетки

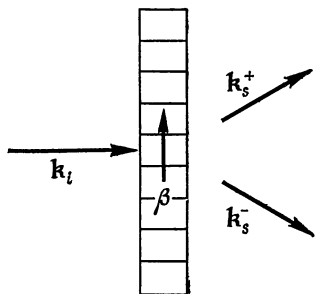


Ф и г. 2.6.3. Иллюстрация условия Брэгга для случая, когда в соотношении (2.6.14) используется знак минус.



Ф и г. 2.6.4. Иллюстрация условия Брэгга для случая, когда в соотношении (2.6.14) используется знак плюс.

конечной толщины условие Брэгга выполняется лишь приближенно. Этот факт можно использовать при определении положения дифракционных максимумов двумерной



Ф и г. 2.6.5. Тонкая дифракционная решетка Брэгга при нормальном падении плоской волны.

фазовой решетки из законов рассеяния Брэгга. Предположим, что толщина d сравнительно мала. Тогда условие Брэгга выполняется неточно. Пусть вектор β направлен параллельно стороне дифракционной решетки, а волна падает на решетку перпендикулярно, как показано на фиг. 2.6.5. Условие Брэгга справедливо в этом случае только для x - и y -компонент (фиг. 2.6.6). Конечно, длина векторов k_s и k_i та же, но векторная диаграмма не согла-

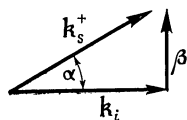
сована. Для малых углов соотношение (2.6.29) справедливо приближенно, так что получаем

$$\alpha = \frac{\lambda}{D}. \quad (2.6.30)$$

Это выражение идентично выражению (2.5.25) для дифракционного максимума первого порядка, полученному в разд. 2.5. Величина D в разд. 2.5 обозначалась как a .

Сравнение формул (2.6.30) и (2.5.25) выявляет важный общий момент. Вывод законов дифракции Брэгга, основанный на теории возмущений в первом порядке, позволил

Ф и г. 2.6.6. Для тонкой дифракционной решетки условие Брэгга точно не выполняется.



получить только первый порядок интерференции трехмерной дифракционной решетки. Эта ситуация полностью аналогична приближению первого порядка теории двумерной фазовой решетки, использование которого дало максимумы первого порядка (2.5.24). Точная теория дифракции Брэгга показывает, что имеются также высшие порядки интерференции, поэтому точное условие Брэгга должно иметь вид

$$\mathbf{k}_s = \mathbf{k}_i \pm m\boldsymbol{\beta}. \quad (2.6.31)$$

Целое m указывает порядок интерференции дифракционной решетки. Требование зеркального отражения сохраняется. Единственная разница заключается в том, что имеется большое число возможных направлений для падающей волны, при которых возникает дифракционная картина. Однако для большинства практических приложений, таких, как голография или рассеяние света звуковыми волнами в жидкости, изменение показателя преломления η столь незначительно, что надежно можно наблюдать только первый порядок дифракции.

Так как теория дифракции Брэгга применима к рассеянию света звуковыми волнами в жидкостях и твердых телах, не удивительно, что наши результаты могут быть получены из квантовой теории рассеяния света фононами.

Этот процесс рассеяния известен также как рассеяние Бриллюэна [10]. Квантовая теория приводит к соотношению (2.6.31) при удовлетворении требования, что импульсы взаимодействующих фотонов и фононов сохраняются. Рассеяние Бриллюэна является частным случаем комбинационного рассеяния света [10]¹⁾: либо падающий световой фотон испускает фонон и излучает оставшуюся энергию в виде рассеянного фотона, либо падающий фотон поглощает фонон и преобразуется в рассеянный фотон. Закон сохранения импульса для рассеяния, включающего испускание или поглощение m фононов, приводит к соотношению (2.6.31).

Требование сохранения энергии дает

$$\omega_s = \omega_i \pm m\omega_{ph}, \quad (2.6.32)$$

где ω_s , ω_i и ω_{ph} — круговые частоты рассеянного фотона, падающего фотона и фонона. Сдвиг частоты, который возникает при рассеянии света на звуковых волнах, является результатом эффекта Доплера. С точки зрения квантовой механики это является следствием сохранения энергии частиц, участвующих в процессе.

Наше рассмотрение дифракции Брэгга до сих пор ограничивалось случаем фазовой трехмерной решетки. Однако легко можно распространить область применимости нашего анализа на решетки, в которых вместо показателя преломления в пространстве варьируется коэффициент поглощения. Выше было показано, что комплексный показатель преломления в виде (2.6.18) описывает среду с потерями. Используя вместо (2.6.1) выражение

$$n = n_0 + i\xi \cos \boldsymbol{\beta} \cdot \mathbf{r}, \quad (2.6.33)$$

получим трехмерную решетку с периодически меняющимся коэффициентом поглощения. То обстоятельство, что, как может показаться, соотношение (2.6.33) описывает среду с перемежающимися областями, ослабляющими и усиливающими свет, легко устраняется добавлением постоянных фиксированных потерь путем введения соответствующим образом выбранной малой мнимой части в n_0 .

¹⁾ Явление комбинационного рассеяния света, как известно, было открыто Мандельштамом и Ландсбергом в 1920 г. и независимо от них Раманом. — *Прим. перев.*

Все полученные выше соотношения остаются справедливыми для решетки с переменными потерями, если произвести замену

$$\eta = i\xi. \quad (2.6.34)$$

Другое полезное развитие данной теории получается путем обобщения формы распределения показателя преломления переходом от простой синусоиды к произвольной форме с малой амплитудой. Используем вместо формул (2.6.1) или (2.6.33) выражение

$$n = n_0 + \sum_i \eta_i \cos(\beta_i \cdot \mathbf{r} + \phi_i). \quad (2.6.35)$$

Из-за предположения

$$\eta_i \ll 1 \quad (2.6.36)$$

при вычислении n^2 перекрестные члены выпадут. Рассеянная волна, образующаяся в случае распределения показателя преломления общего вида, находится сразу, если приписать индексы величинам η и ϕ в формуле (2.6.26) и просуммировать полученные выражения.

Однако следует помнить, что условие Брэгга (2.6.14) для x - и y -компонент должно быть удовлетворено, так что суммирование надо ограничить лишь теми членами, для которых оно выполняется.

Можно применить развитую теорию со всеми обобщениями к многослойным диэлектрикам, которые часто используются в зеркалах и фильтрах. Для этого нужно разложить данное распределение показателя преломления в ряд Фурье. Для любой падающей волны условию Брэгга может удовлетворить не больше чем один член ряда. Из формулы (2.6.14) получаем сразу требуемую толщину слоя для любого желательного угла падения. Однако амплитуда волны может быть получена с хорошей точностью только для зеркал с низкой отражательной способностью из-за ограничений, свойственных теории возмущений. Общее рассмотрение можно провести для случая, когда свет падает нормально на слоистое вещество. Наиболее эффективным способом анализа для этого случая является использование лишь основного члена в разложении (2.6.35). При нормальном падении в формуле (2.6.14) сохраняются только z -компоненты векторов и они равны

(с точностью до знака) модулям этих векторов. Таким образом, из (2.6.14) получаем условие оптимума для периодического многослойного зеркала

$$-k_s = k_i - \beta. \quad (2.6.37)$$

Используя формулу (2.6.2) и соотношение $k_s = k_i = 2\pi/\lambda$, получим длину периода D , требуемую для максимального отражения от многослойной периодической структуры,

$$D = \frac{\lambda}{2}. \quad (2.6.38)$$

2.7. ДИФРАКЦИЯ БРЭГГА. ТЕОРИЯ СВЯЗАННЫХ ВОЛН [13]

Рассмотрение дифракции Брэгга в предыдущем разделе привело нас к важнейшим результатам. Оно позволило получить дифракционную картину тонкой дифракционной решетки, а также найти важное условие Брэгга в общей векторной форме (2.6.14) и в более привычной форме (2.6.29). Единственным недостатком теории возмущений является невозможность использования ее для предсказания амплитуд падающей и рассеянной волн в случае сильной связи или большой длины взаимодействия.

В этом разделе завершается обсуждение дифракции Брэгга рассмотрением теории связанных волн применительно к взаимодействию падающей волны с диффракгированной.

В качестве исходных примем выражение (2.6.4) и введенное волновое уравнение. Будем использовать систему координат, показанную на фиг. 2.6.1, так что плоскости $z = \pm d$ совпадают с границами пластины, имеющей периодически меняющийся показатель преломления. Используя сведения, полученные из теории возмущений, о том, что имеется определенная диффракгированная плоская волна, предположим, что полное поле можно представить в виде

$$\psi = A(z) e^{-ik_i \cdot \mathbf{r}} + B(z) e^{ik_s \cdot \mathbf{r}}. \quad (2.7.1)$$

Амплитуда A падающей волны и амплитуда B диффракгированной волны считаются функциями z . Взаимодействие волн не столь сильно, чтобы существенно изменить амплитуды волн на пути в одну длину волны. Это позволяет

пренебречь вторыми производными амплитуд по сравнению с остальными членами. Такой прием в книге используется неоднократно. Полученное важное математическое приближение приводит к связанным дифференциальным уравнениям первого порядка для амплитуд волн.

Чтобы можно было рассмотреть дифракцию Брэгга в среде с периодической вариацией коэффициента поглощения, необходимо, чтобы среднее значение n_0 показателя преломления содержало член, соответствующий потерям. Таким образом, мы воспользуемся соотношением (2.6.18) при дополнительном предположении

$$n_i \ll n_r. \quad (2.7.2)$$

Пусть также

$$|\eta| \ll n_r. \quad (2.7.3)$$

Тогда соотношение (2.6.4) можно аппроксимировать следующим образом:

$$n^2 \approx n_r^2 - 2in_r n_i + 2\eta n_r \cos \mathbf{b} \cdot \mathbf{r}. \quad (2.7.4)$$

Предположение (2.7.2) означает, что потери в среде настолько малы, что амплитуда волны не изменяется существенно на пути, равном длине волны. Аналогично из соотношения (2.7.3) следует, что модуляция коэффициента поглощения или изменение показателя преломления слабые. Эти предположения подтверждаются во всех случаях, представляющих практический интерес. Если потери в веществе столь велики, что приближения (2.7.2) и (2.7.3) не применимы, падающая волна поглощается так сильно, что дифракция Брэгга уже наблюдаться не может.

Пренебрегая вторыми производными амплитуды, получим

$$\nabla^2 (A e^{-i\mathbf{k}_i \cdot \mathbf{r}}) = - \left(k_i^2 A + 2ik_{iz} \frac{\partial A}{\partial z} \right) e^{-i\mathbf{k}_i \cdot \mathbf{r}}, \quad (2.7.5)$$

где k_{iz} — z -компонента вектора \mathbf{k}_i . Модуль постоянных распространения определяется соотношением

$$k_i^2 = k_s^2 = n_r^2 k_0^2. \quad (2.7.6)$$

Используя это соотношение и условия, рассмотренные выше, а также подставляя выражение (2.7.1) в приведен-

ное волновое уравнение (2.6.5), получим

$$\begin{aligned} & \left[-2ik_{iz} \frac{\partial A}{\partial z} - 2in_r n_i k_0^2 A + \eta n_r k_0^2 B e^{i(k_i - k_s - \beta) \cdot r} \right] e^{-ik_i \cdot r} + \\ & + \left[-2ik_{sz} \frac{\partial B}{\partial z} - 2in_r n_i k_0^2 B + \eta n_r k_0^2 A e^{i(k_s - k_i + \beta) \cdot r} \right] e^{-ik_s \cdot r} + \\ & + \eta n_r k_0^2 [A e^{-i(k_i + \beta) \cdot r} + B e^{-i(k_s - \beta) \cdot r}] = 0. \quad (2.7.7) \end{aligned}$$

Используем то обстоятельство, что векторы распространения рассеянной и падающей волн должны удовлетворять условию Брэгга (2.6.14). Тогда экспоненциальные множители в первых двух скобках равны единице. Теперь можно повторить процедуру, которая применялась ранее для решения уравнения (2.6.11). Умножим обе части равенства (2.7.7) на $\exp(i\mathbf{k}_i \cdot \mathbf{r})$ и проинтегрируем полученное уравнение по всей плоскости x, y . Пусть в направлении z интегрирование распространяется на много периодов осцилляций, а интервал интегрирования остается коротким по сравнению с расстоянием, на котором $A(z)$ и $B(z)$ изменяются значительно. Все члены, кроме первого, в равенстве (2.7.7) приближенно равны нулю, и мы получаем

$$\frac{\partial A}{\partial z} + \frac{n_r n_i k_0^2}{k_{iz}} A = \frac{\eta n_r k_0^2}{2ik_{iz}} B. \quad (2.7.8)$$

Аналогично, умножая обе части равенства (2.7.7) на $\exp(i\mathbf{k}_s \cdot \mathbf{r})$ и интегрируя, находим

$$\frac{\partial B}{\partial z} + \frac{n_r n_i k_0^2}{k_{sz}} B = \frac{\eta n_r k_0^2}{2ik_{sz}} A. \quad (2.7.9)$$

При отсутствии периодической вариации показателя преломления ($\eta = 0$) уравнения (2.7.8) и (2.7.9) независимы. Уравнение (2.7.8) имеет решение

$$A = A_0 e^{-\alpha_i(z+d)}, \quad (2.7.10)$$

где

$$\alpha_i = \frac{n_r n_i k_0^2}{k_{iz}}, \quad (2.7.11)$$

а уравнение (2.7.9) при $\eta = 0$ имеет решение

$$B = B_0 e^{-\alpha_s(z+d)}, \quad (2.7.12)$$

где

$$\alpha_s = \frac{n_r n_i k_0^2}{k_{sz}}. \quad (2.7.13)$$

При отсутствии связи обе волны в среде с потерями распространяются с затухающими амплитудами. Интересно заметить, что амплитуда уменьшается равномерно не в направлении волнового вектора \mathbf{k} , а вдоль плоскостей $z = \text{const}$. Если плоская волна распространяется внутри однородной среды с потерями, ее амплитуда уменьшается как $\exp[-\alpha(|\mathbf{k}|/k) \cdot \mathbf{r}]$. Однако в данном случае ситуация другая. Мы предположили, что плоская волна падает из среды без потерь и входит в среду с потерями в плоскости $z = -d$. Амплитуда A_0 этой волны постоянна вне среды с потерями и уменьшается как функция z в этой среде. Граничное условие при $z = -d$: амплитуда $A(z)$ должна быть равна амплитуде падающей волны A_0 . Амплитуда волны (2.7.10) удовлетворяет этому граничному условию, и, следовательно, это решение задачи правильное. Отражение и преломление на границе раздела отсутствуют, так как, согласно предположению (2.7.2), которое позволяет пренебречь вторыми производными от $A(z)$ по z , потери достаточно малы.

Введя сокращение

$$\kappa = \frac{1}{2} \eta n_r k_0, \quad (2.7.14)$$

можно записать систему связанных уравнений в виде

$$\frac{\partial A}{\partial z} + \alpha_i A = -i \frac{k_0}{k_{iz}} \kappa B \quad (2.7.15)$$

и

$$\frac{\partial B}{\partial z} + \alpha_s B = -i \frac{k_0}{k_{sz}} \kappa A. \quad (2.7.16)$$

Совместное решение уравнений (2.7.15) и (2.7.16) позволяет исключить B :

$$\frac{\partial^2 A}{\partial z^2} + (\alpha_i + \alpha_s) \frac{\partial A}{\partial z} + \left(\alpha_i \alpha_s + \frac{k_0^2 \kappa^2}{k_{iz} k_{sz}} \right) A = 0. \quad (2.7.17)$$

Так как α_i , α_s и κ малы, мы не можем теперь пренебречь второй производной A . Будем искать решение в виде

$$A = e^{\sigma z}. \quad (2.7.18)$$

Тогда из уравнения (2.7.17) получим

$$\sigma^2 + (\alpha_i + \alpha_s)\sigma + \left(\alpha_i \alpha_s + \frac{k_0^2 \kappa^2}{k_{iz} k_{sz}} \right) = 0. \quad (2.7.19)$$

Это квадратное уравнение имеет два решения:

$$\sigma_{\pm} = -\frac{\alpha_i + \alpha_s}{2} \pm \frac{1}{2} \sqrt{(\alpha_i - \alpha_s)^2 - 4 \frac{k_0^2 \kappa^2}{k_{iz} k_{sz}}}. \quad (2.7.20)$$

Тогда общее решение дифференциального уравнения (2.7.17) имеет вид

$$A = a e^{\sigma_+(z+d)} + b e^{\sigma_-(z+d)}. \quad (2.7.21)$$

Из формулы (2.7.15) получаем

$$B = i \frac{k_{iz}}{k_0 \kappa} [(\sigma_+ + \alpha_i) a e^{\sigma_+(z+d)} + (\sigma_- + \alpha_i) b e^{\sigma_-(z+d)}]. \quad (2.7.22)$$

Таким образом, мы нашли общее решение задачи о дифракции Брэгга. Остается изучить это решение, чтобы получить сведения о некоторых интересных частных случаях.

Существует два возможных способа наблюдать дифракцию Брэгга. В зависимости от ориентации вектора β условие Брэгга (2.6.14) выполняется или для отраженной рассеянной волны, или для прошедшей рассеянной волны. Таким образом, мы должны рассмотреть случай, когда рассеянная волна возникает на той же стороне рассеивающего материала, с которой первичная волна падает, либо случай, когда она возникает с противоположной стороны.

Прежде всего рассмотрим случай отраженной рассеянной волны. При этом на противоположной стороне рассеивающего слоя толщиной $2d$ рассеянная волна отсутствует (фиг. 2.6.1). Следовательно, начальными условиями являются $B = 0$ при $z = d$ и $A = A_0$ при $z = -d$. Эти два условия позволяют найти коэффициенты a и b , определяющие амплитуды,

$$a = \frac{(\sigma_- + \alpha_i) e^{-(\sigma_+ - \sigma_-)d}}{\sigma_- e^{-(\sigma_+ - \sigma_-)d} - \sigma_+ e^{(\sigma_+ - \sigma_-)d} - 2\alpha_i \operatorname{sh}(\sigma_+ - \sigma_-)d} A_0 \quad (2.7.23)$$

и

$$b = \frac{-(\sigma_+ + \alpha_i) e^{(\sigma_+ - \sigma_-)d}}{\sigma_- e^{-(\sigma_+ - \sigma_-)d} - \sigma_+ e^{(\sigma_+ - \sigma_-)d} - 2\alpha_i \operatorname{sh}(\sigma_+ - \sigma_-)d} A_0. \quad (2.7.24)$$

В случае проходящей волны, когда рассеянная волна возникает на противоположной стороне рассеивающего мате-

риала по отношению к падающей волне, налагаются условия $B = 0$ и $A = A_0$ при $z = -d$. Амплитудные коэффициенты в этом случае имеют вид

$$a = \frac{\sigma_- + \alpha_i}{\sigma_- - \sigma_+} A_0 \quad (2.7.25)$$

и

$$b = -\frac{\sigma_+ + \alpha_i}{\sigma_- - \sigma_+} A_0. \quad (2.7.26)$$

Дальнейшее рассмотрение дифракции Брэгга проведем для реальной фазовой решетки, а также для решетки, образованной периодическими слоями с потерями. В обоих случаях необходимо различать рассеяние вперед и назад.

Если периодические изменения показателя преломления имеют вещественные значения (фазовая решетка), то потери пренебрежимо малы и можно положить

$$\alpha_i = \alpha_s = 0. \quad (2.7.27)$$

Из формулы (2.7.20) в этом случае получим

$$\sigma_{\pm} = \pm i\rho \text{ (случай прохождения),} \quad (2.7.28)$$

где

$$\rho = \frac{k_0 \kappa}{\sqrt{k_{iz} k_{sz}}}. \quad (2.7.29)$$

Для случая прохождения (рассеянная волна возникает на противоположной стороне слоя рассеивающего материала) из соотношений (2.7.24), (2.7.22), (2.7.25) и (2.7.26) имеем

$$A(z) = A_0 \cos \rho(z+d) \text{ для } -d \leq z \leq d, \quad (2.7.30)$$

$$B(z) = -i \sqrt{\frac{k_{iz}}{k_{sz}}} A_0 \sin \rho(z+d) \text{ для } -d \leq z \leq d. \quad (2.7.31)$$

Этот результат ясно показывает, как периодически происходит обмен мощностью между двумя волновыми процессами. Если

$$2\rho d = (2n+1) \frac{\pi}{2}, \quad n=0, 1, 2, \dots, \quad (2.7.32)$$

то вся мощность падающей волны переходит к рассеянной волне. Таким образом, мы получили результат, который не могли вывести из теории возмущений: дифракция Брэг-

га может иметь 100 %-ную эффективность. Количество переданной мощности зависит от силы связи и толщины слоя материала. При $2\rho d \ll 1$ получаем

$$B(d) = -2id \sqrt{\frac{k_{iz}}{k_{sz}}} \rho A_0 = -i \frac{n_r \eta k_0^2}{k_{sz}} A_0 d. \quad (2.7.33)$$

Тогда амплитуда при $z = d$ согласуется с результатом (2.6.28), полученным методом теории возмущений.

В случае отражения (рассеянная волна возникает на той же стороне, что и падающая волна), падающая и рассеянная волны движутся в противоположных направлениях. Это означает, что z -компоненты их волновых векторов имеют противоположные знаки, т. е.

$$k_{iz} k_{sz} < 0. \quad (2.7.34)$$

Параметр ρ в (2.7.29) теперь мнимый и удобнее написать

$$\sigma_{\pm} = \pm |\rho|. \quad (2.7.35)$$

Амплитуды поля находятся снова из соотношений (2.7.21) и (2.7.22). В этом случае с помощью формул (2.7.23) и (2.7.24) находим

$$A(z) = \frac{\text{ch} |\rho| (z-d)}{\text{ch} 2 |\rho| d} A_0 \quad \text{при} \quad -d \leq z \leq d \quad (2.7.36)$$

и

$$B(z) = i \sqrt{\left| \frac{k_{iz}}{k_{sz}} \right|} \frac{\text{sh} |\rho| (z-d)}{\text{ch} 2 |\rho| d} A_0 \quad \text{при} \quad -d \leq z \leq d. \quad (2.7.37)$$

Для случая $2 |\rho| d \ll 1$ из формулы (2.7.37) при $B(-d)$ получаем тот же результат (2.6.28), что и из теории возмущений. Следует отметить интересное различие между случаями отражения и прохождения волн. В случае прохождения мощность может осциллировать, т. е. переходить от одной волны к другой периодически. Полная передача мощности возможна при надлежащем выборе толщины рассеивающего материала. В случае отражения осцилляции мощности между волнами отсутствуют. Полная передача мощности от падающей волны к рассеянной возможна, если рассеивающее вещество имеет достаточно большую толщину. Дальнейшее увеличение толщины не приводит к заметным дополнительным эффектам. Точный выбор толщины рассеивающего вещества в этом случае не крити-

чен. Если материал имеет достаточную толщину, всегда будет иметь место полная передача мощности. Примером такого случая являются многослойные диэлектрические зеркала. В случае прохождения волны необходимо очень точно подбирать толщину материала, чтобы получить полное прохождение энергии.

Далее рассмотрим случай рассеивающего вещества с периодическими вариациями потерь (амплитудную решетку). Мы не можем теперь считать, что постоянная часть диэлектрических потерь обращается в нуль, поскольку, для того чтобы в веществе не было усиления волны, мнимая часть показателя преломления (2.6.1) не должна становиться положительной.

Необходимо потребовать, чтобы

$$n_i \geq |\eta|, \quad (2.7.38)$$

где η — мнимая величина, так что

$$\kappa^2 < 0. \quad (2.7.39)$$

В случае прохождения волн имеем

$$k_{iz}k_{sz} > 0, \quad (2.7.40)$$

так что

$$\sigma_+ = -\alpha + \gamma \quad (2.7.41)$$

и

$$\sigma_- = -\alpha - \gamma, \quad (2.7.42)$$

где

$$\alpha = \frac{1}{2} (\alpha_i + \alpha_s) \quad (2.7.43)$$

и

$$\gamma = \frac{1}{2} \sqrt{(\alpha_i - \alpha_s)^2 + 4 \left| \frac{k_0^2 \kappa^2}{k_{iz} k_{sz}} \right|}. \quad (2.7.44)$$

Теперь получим амплитуды волн для случая прохождения через среду с периодическими вариациями потерь (для $-d \leq z \leq d$)

$$A(z) = e^{-\alpha(z+d)} \left[\operatorname{ch} \gamma(z+d) - \frac{\alpha_i - \alpha_s}{2\gamma} \operatorname{sh} \gamma(z+d) \right] A_0, \quad (2.7.45)$$

$$B(z) = \frac{ik_0 \kappa^* A_0}{\gamma k_{sz}} e^{-\alpha(z+d)} \operatorname{sh} \gamma(z+d). \quad (2.7.46)$$

Этот результат существенно сложнее, чем в случае фазовой решетки без потерь. Полный обмен мощностью, конечно, не возможен, так как обе волны испытывают потери при прохождении через рассеивающую среду. В случае рассеивания вперед при нормальном падении, $k_{sz} = k_{iz}$, также имеем $\alpha_i = \alpha_s = \alpha$, так что из соотношений (2.7.44), (2.7.11) и (2.7.14) получаем

$$\gamma = \frac{|\eta|}{2n_i} \alpha, \quad (2.7.47)$$

и выражения для амплитуд волн упрощаются:

$$A(z) = A_0 e^{-\alpha(z+d)} \operatorname{ch} \left[\frac{1}{2} \frac{|\eta|}{n_i} \alpha (z+d) \right] \quad (2.7.48)$$

и

$$B(z) = A_0 e^{-\alpha(z+d)} \operatorname{sh} \left[\frac{1}{2} \frac{|\eta|}{n_i} \alpha (z+d) \right], \quad (2.7.49)$$

где $-d \leq z \leq d$.

Интересно заметить, что осцилляции обмена мощностью, которые имеют место при прохождении волн в фазовой решетке, не происходят в случае решетки с потерями (амплитудной решетки). Амплитуда рассеянной волны при $z = d$ равна

$$B(d) = A_0 e^{-2\alpha d} \operatorname{sh} \left(\frac{|\eta|}{n_i} \alpha d \right). \quad (2.7.50)$$

Постоянные η и n_i должны удовлетворять соотношению (2.7.38). Следовательно, наилучшим условием для такой решетки является равенство

$$|\eta| = n_i. \quad (2.7.51)$$

Тогда амплитуда волны, рассеянной вперед, равна

$$B(d) = \frac{1}{2} A_0 (e^{-\alpha d} - e^{-3\alpha d}). \quad (2.7.52)$$

При

$$2\alpha d = \ln 3 \quad (2.7.53)$$

амплитуда достигает максимума:

$$B_{\max}(d) = \frac{A_0}{3^{3/2}} = 0,192 A_0. \quad (2.7.54)$$

Мощность, переносимая рассеянной волной, пропорциональна $|B|^2$. Решетка с периодическими вариациями поглощения, удовлетворяющая условию оптимума (2.7.51), может передать не более 3,7% мощности падающей волны волне, рассеянной вперед. В связи с этим, конечно, бессмысленно использовать объемные решетки при рассеянии вперед, так как падающая и рассеянная волны трудно различимы. Наши расчеты преследуют цель установить оптимальные величины для рассеяния почти вперед при работе с амплитудными решетками. Интересно заметить, что для той же величины $\alpha_i = \alpha$ и толщины материала, определяемой формулой (2.7.53), амплитуда падающей волны после прохождения материала с тем же значением n_i , но при $\eta = 0$ равна

$$A(d) = A_0 e^{-2\alpha d} = \frac{1}{3} A_0. \quad (2.7.55)$$

Далее рассмотрим случай отражения от амплитудной решетки. В этом случае выполняется неравенство (2.7.34), так что имеем

$$\sigma_+ = -\alpha + \delta \quad (2.7.56)$$

и

$$\sigma_- = -\alpha - \delta, \quad (2.7.57)$$

где

$$\delta = \frac{1}{2} \sqrt{(\alpha_i - \alpha_s)^2 - 4 \left| \frac{k_0^2 \kappa^2}{k_{iz} k_{sz}} \right|}, \quad (2.7.58)$$

а α дается соотношением (2.7.43). Амплитуды волн получаются из формул (2.7.21) и (2.7.22) с помощью соотношений (2.7.23) и (2.7.24) (для $-d \leq z \leq d$)

$$A(z) = A_0 e^{-\alpha(z+d)} \frac{\delta \operatorname{ch} \delta(z-d) - \frac{1}{2} (\alpha_i - \alpha_s) \operatorname{sh} \delta(z-d)}{\delta \operatorname{ch} 2\delta d + \frac{1}{2} (\alpha_i - \alpha_s) \operatorname{sh} 2\delta d}, \quad (2.7.59)$$

$$B(z) = - \left| \frac{k_0 \kappa}{k_{sz}} \right| A_0 e^{-\alpha(z+d)} \frac{\operatorname{sh} \delta(z-d)}{\delta \operatorname{ch} 2\delta d + \frac{1}{2} (\alpha_i - \alpha_s) \operatorname{sh} 2\delta d}. \quad (2.7.60)$$

Рассеянная волна возникает теперь при $z = -d$. Предполагая, что рассеяние происходит назад при нормальном

падении, $k_{iz} = -k_{sz}$, имеем $\alpha_s = -\alpha_i$ и, следовательно, $\alpha = 0$. Используя еще и условие оптимума (2.7.51), получаем из формулы (2.7.58)

$$\delta = \frac{\sqrt{3}}{2} \alpha_i. \quad (2.7.61)$$

Учитывая равенство

$$\frac{1}{2}(\alpha_i - \alpha_s) = \alpha_i, \quad (2.7.62)$$

из формулы (2.7.60) имеем

$$B(-d) = \frac{iA_0}{2 + \sqrt{3} \operatorname{cth}(\sqrt{3} \alpha_i d)}. \quad (2.7.63)$$

Оптимума по толщине теперь нет. Максимальное значение $B(-d)$ достигается, когда $2d$ устремляется к бесконечности:

$$B_{\max}(-d) = \frac{iA_0}{2 + \sqrt{3}} = 0,268iA_0. \quad (2.7.64)$$

Оптимальная эффективность передачи мощности в случае отражения достигает 7,2%. С амплитудной решеткой можно получить оптимальную передачу мощности для случая отражения назад, в 2 раза превышающую мощность, рассеянную вперед. Таким образом, фазовая решетка без потерь существенно эффективнее любой решетки с периодическими вариациями коэффициента поглощения, так как с фазовой решеткой можно достичь эффективности передачи мощности 100%.

При рассеянии вперед (или почти вперед) вектор β должен быть направлен параллельно плоскостям $z = \text{const}$. Это означает, что плоскости постоянных потерь или постоянного показателя преломления должны быть ориентированы почти перпендикулярно плоскостям $z = \text{const}$. При рассеянии назад вектор β должен быть направлен почти перпендикулярно плоскостям $z = \text{const}$, так что слои постоянного показателя преломления параллельны плоскостям $z = \text{const}$.

ГЕОМЕТРИЧЕСКАЯ ОПТИКА

3.1. ВВЕДЕНИЕ

Замечательной особенностью видимого света является малость длины его волны. Существует электромагнитное излучение с более короткой волной. Это ультрафиолетовое излучение, рентгеновские лучи и гамма-излучение. Области видимого света соответствуют длинам волн примерно от 0,4 до 0,7 мкм. Более длинные волны называются инфракрасными волнами, волнами диапазона СВЧ и радиоволнами. Часто области ультрафиолетового и инфракрасного излучений также относят к оптике. Однако не так важно, как именно мы определим область оптического излучения, если имеем дело с излучением, длина волны которого существенно меньше любых размеров, встречающихся в нашей повседневной жизни. Линейные размеры элементов во многих оптических приборах гораздо больше длины волны света, проходящего через эти приборы. В таких случаях для описания процессов распространения света можно пользоваться приближенными методами геометрической оптики. При этом нет необходимости в решении уравнений Максвелла или волнового уравнения. Геометрическая оптика применима в тех случаях, когда дифракцией света можно пренебречь.

Лучевая оптика, как иногда называют геометрическую оптику, может быть использована при рассмотрении явлений, описываемых волновым уравнением, при условии, что длина волны мала по сравнению с размерами аппаратуры. Лучевая оптика применима поэтому для решения многих задач, связанных с распространением звука. Она близка к классической механике материальной точки. Действительно, соотношение между лучевой оптикой и волновой оптикой такое же, как между классической механикой и волновой механикой.

Начнем с вывода уравнений геометрической оптики из волнового уравнения. Затем получим эти же уравнения из принципа Ферма и рассмотрим принцип Гамильтона применительно к геометрической оптике. Покажем, что теорема Лиувилля, обычно рассматриваемая в статистической механике, имеет важные и интересные приложения в оптике. В данную главу включена также квантовая теория лучевой оптики, показывающая, что квантование уравнений геометрической оптики также приводит к волновому уравнению. Сопоставление релятивистской механики материальной точки и релятивистской квантовой механики показывает, что так называемое параксиальное приближение в лучевой оптике эквивалентно нерелятивистской механике материальной точки.

Принцип Гамильтона в лучевой оптике и квантовая теория световых лучей не являются совершенно необходимыми для понимания остальной части книги. Поэтому читатель при желании может опустить разд. 3.5 и 3.6.

3.2. ВЫВОД УРАВНЕНИЙ ЛУЧЕВОЙ ОПТИКИ ¹⁾ ИЗ ВОЛНОВОГО УРАВНЕНИЯ

В гл. 1 было выведено волновое уравнение (1.3.6) из уравнений Максвелла. Отмечалось, что волновому уравнению удовлетворяет каждая составляющая векторов электрического и магнитного полей в однородной среде. Однако даже в случае неоднородной среды можно использовать волновое уравнение с хорошим приближением, если показатель преломления среды изменяется очень мало на расстояниях, сравнимых с длиной волны излучения. Волновое уравнение описывает не только электромагнитные явления, но и множество других задач, связанных с распространением волн. В линейном приближении волновым уравнением может быть описано распространение звука в жидкостях, газах и твердых телах. Тот факт, что волновое уравнение берется в качестве отправной точки геометрической оптики, свидетельствует о том, что мы собираемся получить приближенную теорию, которая может быть использована при исследовании распространения не только

¹⁾ Лучевая оптика хорошо изложена в [1, 15, 16].

света, но и звука. Векторная природа света теряется в представлениях геометрической оптики. Свет рассматривается как скалярный процесс. В терминах фотонной теории это означает, что мы пренебрегаем свойствами спина фотона.

Ограничимся гармоническими во времени процессами, предполагая, что все переменные поля имеют зависимость от времени, которая в комплексных обозначениях может быть выражена через $\exp(i\omega t)$, так что волновое уравнение можно заменить приведенным волновым уравнением (2.2.2) ¹⁾

$$\nabla^2 \psi + k^2 \psi = 0, \quad (3.2.1)$$

в котором k может быть выражено в любой из следующих форм:

$$k = \frac{2\pi}{\lambda} = \omega \sqrt{\varepsilon \mu_0}. \quad (3.2.2)$$

Следует сразу же оговориться, что приведенное волновое уравнение в действительности не накладывает больших ограничений по сравнению с волновым уравнением (1.3.6), как это может показаться, потому что, как мы убедились в разд. 1.3, в дисперсионной среде волновое уравнение имеет физический смысл только в узком диапазоне частот, для которого можно определить фазовую скорость излучения. Волновое уравнение справедливо для всех частот только в свободном пространстве. Однако случай свободного пространства представляет наименьший интерес для геометрической оптики.

Чтобы получить желаемое приближение, предположим, что волновая функция ψ имеет следующий вид:

$$\psi = \psi_0(x, y, z) e^{-ik_0 S(x, y, z)}, \quad \text{где } k_0 = \omega \sqrt{\varepsilon_0 \mu_0}. \quad (3.2.3)$$

Преимуществом этой формы волновой функции является то, что она отделяет быстрые фазовые изменения, вызванные малостью длины волны поля излучения в свободном пространстве, от гораздо более медленных изменений амплитуды ψ_0 . Как ψ_0 , так и S являются функциями, кото-

¹⁾ Как уже отмечалось, это уравнение также называется уравнением Гельмгольца. — *Прим. ред.*

рые можно считать медленно изменяющимися по отношению к длине волны излучения. Подстановка в приведенное волновое уравнение дает

$$k_0^2 \left(\frac{k^2}{k_0^2} - \nabla S \cdot \nabla S \right) \psi_0 - ik_0 (2\nabla S \cdot \nabla \psi_0 + \psi_0 \nabla^2 S) + \nabla^2 \psi_0 = 0. \quad (3.2.4)$$

Члены в (3.2.4) сгруппированы в порядке уменьшения степени величины k_0 . Из-за малости λ величину k_0 следует рассматривать как очень большое число. Далее можно предположить, что ψ_0 и S являются действительными величинами. Это означает, что член с множителем ik_0 должен равняться нулю независимо от двух других членов. Некоторые авторы [14] используют комплексные ψ_0 и S , что приводит к представлению о комплексных лучах. В данной книге этот формализм не используется, но вместо этого предполагается, что выражение (3.2.3) представляет собой разложение комплексного числа ψ на амплитуду ψ_0 и фазу $k_0 S$. Для коротких длин волн первый член в (3.2.4) много больше последнего, так что с хорошим приближением можно записать

$$(\nabla S)^2 = n^2. \quad (3.2.5)$$

Показатель преломления, как обычно, определяется соотношением

$$n^2 = \frac{k^2}{k_0^2} = \frac{\varepsilon}{\varepsilon_0}. \quad (3.2.6)$$

Уравнение (3.2.5) известно как уравнение эйконала. Оно задает функцию S , которая позволяет определить поверхности постоянных фаз уравнением

$$S(x, y, z) = \text{const.} \quad (3.2.7)$$

Поверхности постоянных фаз определяют форму поля излучения, поэтому уравнение эйконала описывает распространение волны в приближении геометрической оптики.

В соответствии с названием «лучевая оптика» геометрическая оптика использует понятие световых лучей для описания распространения поля излучения. Понятие о световом луче возникает естественным образом, когда мы наблюдаем, как тонкий световой пучок распространяется

в пространстве. В большинстве сред свет постепенно рассеивается на пути его распространения, что делает видимой траекторию светового пучка. Каждому знаком вид тонкого светового пучка в накуренной комнате. Интенсивные световые лучи видимы даже в относительно чистом воздухе, поскольку частицами пыли рассеивается энергия, достаточная, чтобы проследить траекторию луча. Однако понятие светового луча было несколько обобщено. Интуитивно ясно, что два световых пучка пойдут почти по одному и тому же пути, если они вышли параллельно и близко друг от друга. Таким образом, можно представить себе пучок (связку) световых лучей, определяющих направление распространения более широкого светового поля. В терминах волновой оптики можно говорить о световом пучке как об участке плоской волны, который заполняет только весьма узкую область пространства. С некоторым приближением такой пучок можно получить пропусканием плоской волны через отверстие в непрозрачном экране. Как следует из разд. 2.3 и 2.4, плоская волна, прошедшая через очень узкое отверстие, не сохраняет своей формы. Однако, если отверстие велико по сравнению с длиной волны (это предположение существенно для геометрической оптики), световой луч, прошедший через отверстие, будет расширяться благодаря дифракции сравнительно медленно, а понятие об ограниченной плоской волне будет оставаться приблизительно справедливым на некотором расстоянии за отверстием. Известно, что поверхности постоянных фаз перпендикулярны направлению распространения света плоской волны. Теперь можно определить световые лучи как траектории, ортогональные фазовым фронтам световой волны. Если известны поверхности постоянных фаз, можно получить световые лучи, построив линии, перпендикулярные фазовым фронтам. Когда фазовые фронты изгибаются в пространстве из-за изменений показателя преломления, то соответственно изгибаются и световые лучи.

Желательно иметь возможность определять поведение световых лучей непосредственно без построения фазовых фронтов из уравнения эйконала.

Выберем фиксированную начальную точку и проведем из нее вектор ко всем точкам светового луча. Знать вектор

\mathbf{r} для всех точек вдоль луча — это все равно, что иметь математическое описание светового луча. Определив s как расстояние, измеренное вдоль светового луча, получим единичный вектор

$$\mathbf{u} = \frac{d\mathbf{r}}{ds}. \quad (3.2.8)$$

По определению единичный вектор \mathbf{u} , касательный к световому лучу, перпендикулярен фазовым фронтам. Из формулы (3.2.7) вытекает, что вектор, перпендикулярный фазовым фронтам, является градиентом от S :

$$\mathbf{v} = \nabla S. \quad (3.2.9)$$

Векторы \mathbf{v} и \mathbf{u} должны быть параллельны. Величину вектора \mathbf{v} можно найти из уравнения эйконала (3.2.5)

$$|\mathbf{v}| = n. \quad (3.2.10)$$

Отсюда сразу же получается соотношение

$$\mathbf{u} = \frac{\mathbf{v}}{n}, \quad (3.2.11)$$

или, подробнее,

$$n \frac{d\mathbf{r}}{ds} = \nabla S. \quad (3.2.12)$$

Дифференцирование по s можно представить как умножение единичного вектора (3.2.8) на оператор ∇ :

$$\frac{d}{ds} = \sum_i \frac{dx_i}{ds} \frac{\partial}{\partial x_i} = \frac{d\mathbf{r}}{ds} \cdot \nabla. \quad (3.2.13)$$

Чтобы получить уравнение луча, найдем градиент от обеих частей уравнения эйконала (3.2.5):

$$2\nabla S \cdot \nabla \nabla S = 2n \nabla n. \quad (3.2.14)$$

Произведение $\nabla \nabla$ является тензорным оператором. Соотношение (3.2.14) может быть проверено по составляющим векторов. Используя формулы (3.2.12) и (3.2.13), можно переписать (3.2.14) и получить следующее соотношение:

$$\frac{d}{ds} \nabla S = \nabla n. \quad (3.2.15)$$

Взяв производную от (3.2.12) по s и используя (3.2.15), получаем уравнение луча

$$\frac{d}{ds} \left(n \frac{d\mathbf{r}}{ds} \right) = \nabla n. \quad (3.2.16)$$

Уравнение луча и уравнение эйконала являются двумя альтернативными описаниями геометрической оптики. Уравнение луча более удобно для определения траектории световых лучей в однородной среде. Однако решение точного (в пределах лучевой оптики) уравнения (3.2.16) трудно получить. Для многих практических применений представляет интерес приближенное уравнение луча. В некоторых оптических задачах рассматриваются световые лучи, которые всегда распространяются почти параллельно оптической оси системы. Это не значит, что лучи не могут значительно отклоняться от оптической оси. Здесь имеется в виду, что угол α , который световые лучи образуют с этой осью, остается достаточно малым и поэтому можно пользоваться приближением

$$\operatorname{tg} \alpha \approx \sin \alpha \approx \alpha, \quad \cos \alpha \approx 1. \quad (3.2.17)$$

Если принять оптическую ось за ось z , то

$$ds \approx dz \quad (3.2.18)$$

с точностью, предполагаемой допущением (3.2.17). Упрощение, следующее из (3.2.17), известно как параксиальное приближение. В этом приближении уравнение (3.2.16) упрощается до уравнения параксиальных лучей

$$\frac{d}{dz} \left(n \frac{d\mathbf{r}}{dz} \right) = \nabla n. \quad (3.2.19)$$

Уравнение параксиальных лучей будет использовано в гл. 7 для решения задач распространения лучей в оптических волноводах с параболическим распределением показателя преломления.

Уравнение луча можно получить также из имеющего более общий смысл вариационного принципа. Прежде чем перейти к этому выводу, исследуем граничные условия, которым подчиняются световые лучи на границе сред с различными диэлектрическими проницаемостями.

3.3. ГРАНИЧНЫЕ УСЛОВИЯ ДЛЯ СВЕТОВЫХ ЛУЧЕЙ

Граничное условие для световых лучей можно получить из формулы (3.2.12). Используя результат теории линейных интегралов [17], сразу же находим

$$\oint \nabla S \cdot d\mathbf{l} = 0. \quad (3.3.1)$$

Интеграл берется по замкнутой кривой ¹⁾. Вектор $d\mathbf{l}$ направлен по касательной к этой кривой. Используя формулы (3.3.1) и (3.2.12), получаем важное соотношение

$$\oint n \frac{d\mathbf{r}}{ds} \cdot d\mathbf{l} = 0. \quad (3.3.2)$$

В слегка измененном виде уравнение (3.3.2) выражается следующим образом:

$$\int_{P_1}^{P_2} n \frac{d\mathbf{r}}{ds} \cdot d\mathbf{l}_1 = \int_{P_1}^{P_2} n \frac{d\mathbf{r}}{ds} \cdot d\mathbf{l}_2. \quad (3.3.3)$$

Из уравнения (3.3.3) следует, что результат интегрирования $n (d\mathbf{r}/ds) \cdot \mathbf{t}$ (\mathbf{t} — вектор, касательный к кривой и параллельный $d\mathbf{l}$) в пределах от P_1 до P_2 не зависит от пути интегрирования. Индексы 1 и 2 у $d\mathbf{l}$ означают интегрирование вдоль разных кривых. Этот результат можно интерпретировать двумя различными способами. При одной интерпретации рассматривается поле лучей в пространстве. В этом случае из уравнения (3.3.3) следует, что если проекция вектора $n (d\mathbf{r}/ds)$ на направление кривой (вдоль которой берется интеграл) интегрируется от точки P_1 до точки P_2 , то результат интегрирования не зависит от выбора пути интегрирования. Конечно, вектор $d\mathbf{r}/ds$ направлен по касательной к направлению каждого из световых лучей, пересекаемых кривой интегрирования. Это интересный и удивительный результат. Другая интерпретация уравнения (3.3.3) приводит к результату, который еще более важен для лучевой оптики. В данном случае предположим, что две точки P_1 и P_2 соединяются двумя различными световыми лучами. Тогда можно взять две

¹⁾ Соотношение (3.3.1) следует из того, что $\text{rot } \nabla S = 0$. — *Прим. ред.*

кривые интегрирования, каждая из которых совпадает с одним из световых лучей. Векторы $d\mathbf{r}/ds$ и $d\mathbf{l}$ параллельны в этом случае, и их произведение является просто линейным элементом ds светового луча. Соотношение (3.3.3) принимает вид

$$\int_{P_1}^{P_2} n ds_1 = \int_{P_1}^{P_2} n ds_2. \quad (3.3.4)$$

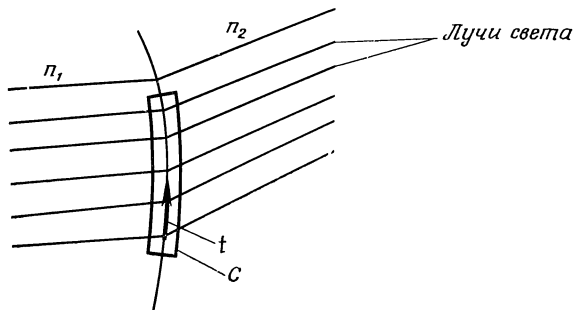
Интеграл от n вдоль светового луча известен как оптическая длина пути. Тогда уравнение (3.3.4) можно интерпретировать следующим образом. Две точки P_1 и P_2 можно соединить двумя различными световыми лучами только в том случае, если оптические длины путей одинаковы для обоих лучей. Этот результат важен для систем формирования изображения. Изображение формируется тогда, когда каждый луч, выходящий из произвольной точки объекта, встречается со всеми остальными лучами, выходящими из той же точки объекта, в некоторой другой точке пространства. Точка, в которой все лучи, исходящие из данной точки объекта, сходятся, называется точкой изображения. Из равенства (3.3.4) следует, что оптические длины путей всех лучей, соединяющих точку объекта с точкой изображения, одинаковы.

После такого отступления вернемся к нашей исходной цели — получению граничных условий для световых лучей. Предположим, что поле световых лучей проходит через границу между двумя средами с различными показателями преломления, как показано на фиг. 3.3.1. Здесь изображена также кривая C , которая проходит параллельно границе в среде 1, пересекает ее в направлении от среды 1 к среде 2 и возвращается к исходной точке, проходя параллельно границе в среде 2. Выбирая кривую C в качестве пути интегрирования в (3.3.2), получаем

$$n_1 \left(\frac{d\mathbf{r}}{ds} \right)_1 \cdot \mathbf{t} = n_2 \left(\frac{d\mathbf{r}}{ds} \right)_2 \cdot \mathbf{t}. \quad (3.3.5)$$

Единичный вектор \mathbf{t} направлен по касательной к границе. Уравнение (3.3.5) получено для случая, когда путь вдоль границы достаточно короткий, чтобы при этом векторы $d\mathbf{r}/ds$ можно было рассматривать как постоянные вдоль

этих участков интегрирования. Отрезки пути интегрирования, пересекающие границу между двумя средами, считаются исчезающе малыми. Векторы dr/ds и t имеют еди-



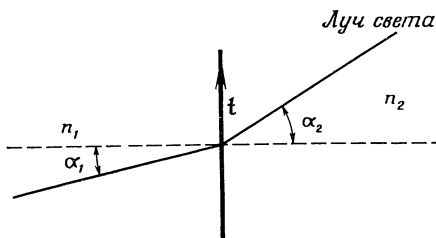
Ф и г. 3.3.1. Путь интегрирования при выводе граничных условий для световых лучей.

C — замкнутая кривая, идущая параллельно поверхности раздела двух сред и пересекающая ее; t — единичный вектор, тангенциальный к поверхности.

ничную длину. С учетом углов, показанных на фиг. 3.3.2, сразу же получаем необходимое граничное условие

$$n_1 \sin \alpha_1 = n_2 \sin \alpha_2. \quad (3.3.6)$$

Это соотношение уже было выведено в разд. 1.6 [уравнение (1.6.19)]. Таким образом, снова получен закон Снеллиуса методом геометрической оптики.



Ф и г. 3.3.2. Расположение углов α_1 и α_2 , входящих в закон Снеллиуса.

Этот результат тем более удивителен, что была сделана оговорка о приближенности результатов лучевой оптики на границе двух диэлектрических сред. Эта оговорка основывалась на том, что уравнения лучевой оптики следуют из волнового уравнения, которое само является лишь

приближением в случае неоднородной среды. Приближенный характер волнового уравнения является следствием его вывода из уравнений Максвелла. На границе сред показатель преломления (или диэлектрическая проницаемость) изменяется очень значительно, что приводит к нарушению условия $(\epsilon_2 - \epsilon_1)/\epsilon \ll 1$ (полученного в разд. 1.3), при котором справедливо волновое уравнение. Чтобы получить равенство (3.3.6) из уравнений лучевой оптики, необходимо использовать волновое уравнение именно в той области, где его применимость становится сомнительной. Однако результат (3.3.6) подтверждается строгим выводом равенства (1.6.19), который был основан на уравнениях Максвелла. Более того, можно вывести уравнение эйконала непосредственно из уравнений Максвелла [18].

Справедливость изложенного метода можно обосновать следующим аргументом. Примем, что граница между двумя средами является не линией, а размазанной областью, в которой показатель преломления постепенно изменяется от n_1 до n_2 . Пусть это изменение происходит достаточно быстро, так что оно завершается на расстоянии, малом по сравнению с другими, интересующими нас размерами, но большом по сравнению с длиной волны света. В этом случае в переходной области можно использовать волновое уравнение. В итоге получаем результат (3.3.6), как и раньше в предположении, что отрезки пути, параллельные границе, много больше отрезков интегрирования, соединяющих обе области пространства. Поскольку лучевая оптика справедлива при $\lambda \rightarrow 0$, гладкая переходная область может быть сколь угодно малой и в то же время достаточно широкой по сравнению с длиной волны.

3.4. ПРИНЦИП ФЕРМА

В разд. 3.2 из приведенного волнового уравнения были получены уравнение эйконала и уравнение луча. В этом разделе будет использован другой подход. Уравнения геометрической оптики можно получить из вариационного принципа. Этот подход имеет то преимущество, что вскрывает тесную связь между лучевой оптикой и классической механикой.

Большинство законов физики может быть получено из вариационных принципов. Наиболее известным из вариационных принципов физики является принцип наименьшего действия Гамильтона [19]. Принцип Ферма очень похож по форме на знаменитый закон Гамильтона. Есть, однако, одно существенное отличие: принцип Гамильтона основан на минимизации функций от времени, а принцип Ферма — на минимизации функций от пространственных координат. Сходство между классической механикой и оптикой требует поэтому замены времени пространственной координатой. За исключением этого, сходство между лучевой оптикой и механикой очень близкое.

Принцип Ферма основывается на понятии об оптической длине пути, которое введено в (3.3.4). Он утверждает, что световой луч всегда выбирает траекторию с минимальной оптической длиной пути. Как указал Льюнеберг [16], в редких случаях этот минимум фактически может быть максимумом. Кроме того, важно понять, что эта минимизация длины пути не приводит к абсолютному минимуму. Принцип Ферма требует, чтобы любой путь в непосредственной близости от минимальной траектории был длиннее ее. В математических обозначениях принцип Ферма принимает вид

$$\int_{P_1}^{P_2} n(x, y, z) ds = \min. \quad (3.4.1)$$

Вместо длины пути можно ввести понятие о времени прохождения путем деления уравнения (3.4.1) на постоянную c , т. е. скорость света в вакууме. Величина c/n является скоростью света в среде с показателем преломления n . Тогда выражение под знаком интеграла есть время, которое необходимо свету для прохождения расстояния ds . Поэтому принцип Ферма можно записать в виде

$$\int_{P_1}^{P_2} dt = \min. \quad (3.4.2)$$

P_1 и P_2 — это две фиксированные точки в пространстве. Световой луч должен пройти от P_1 к P_2 за минимальное время, т. е. он должен найти траекторию, которая приводит

его из P_1 в P_2 по минимальному оптическому пути. Решение этой задачи в свободном пространстве тривиально, так как прямая линия является кратчайшим путем между двумя точками, и, следовательно, свет будет распространяться наиболее быстро в случае выбора прямолинейного пути. Если точки P_1 и P_2 расположены в двух различных однородных диэлектрических средах с плоской границей раздела, принцип Ферма непосредственно приводит к закону преломления Снеллиуса. Этот результат можно получить прямым вычислением. Справедливость его следует из того, что уравнения эйконала и луча удовлетворяют принципу Ферма и из того, что закон Снеллиуса является следствием уравнений лучевой оптики.

Решение вариационной задачи (3.4.1) оказывается более легким, если перейти к новой переменной интегрирования. Мы используем определение элемента длины

$$ds = \sqrt{dx^2 + dy^2 + dz^2} = \sqrt{1 + x'^2 + y'^2} dz, \quad (3.4.3)$$

где

$$x' = \frac{dx}{dz}, \quad y' = \frac{dy}{dz}, \quad (3.4.4)$$

чтобы выразить (3.4.1) в виде

$$\int_{P_1}^{P_2} L(x, y, x', y', z) dz = \min. \quad (3.4.5)$$

Функция L задается соотношением

$$L(x, y, x', y', z) = n(x, y, z) \sqrt{1 + x'^2 + y'^2}. \quad (3.4.6)$$

Уравнение (3.4.5) имеет точно ту же форму, что и принцип наименьшего действия Гамильтона [19]. Единственное отличие принципа Ферма от принципа Гамильтона, как уже отмечалось, заключается в том, что в принципе Ферма вместо временной координаты t используется пространственная координата z . В классической механике функция L называется лагранжианом. Координата z обычно выбирается совпадающей с предпочтительным направлением оптической системы, известным как оптическая ось. Большинство оптических систем имеет ось симметрии, которая является также осью вращения.

Решение задачи (3.4.5) хорошо известно в вариационном исчислении и нет необходимости приводить его здесь [20]. Оно дается уравнениями Эйлера для вариационной задачи

$$\frac{d}{dz} \frac{\partial L}{\partial x'} - \frac{\partial L}{\partial x} = 0, \quad (3.4.7)$$

$$\frac{d}{dz} \frac{\partial L}{\partial y'} - \frac{\partial L}{\partial y} = 0. \quad (3.4.8)$$

Подстановка выражения (3.4.6) приводит к

$$\frac{d}{dz} \frac{nx'}{\sqrt{1+x'^2+y'^2}} = \sqrt{1+x'^2+y'^2} \frac{\partial n}{\partial x}, \quad (3.4.9)$$

$$\frac{d}{dz} \frac{ny'}{\sqrt{1+x'^2+y'^2}} = \sqrt{1+x'^2+y'^2} \frac{\partial n}{\partial y}. \quad (3.4.10)$$

С помощью формулы (3.4.3) эти уравнения можно записать в более сжатой форме:

$$\frac{d}{ds} \left(n \frac{dx}{ds} \right) = \frac{\partial n}{\partial x}, \quad (3.4.11)$$

$$\frac{d}{ds} \left(n \frac{dy}{ds} \right) = \frac{\partial n}{\partial y}. \quad (3.4.12)$$

Сравнение формул (3.4.11) и (3.4.12) с (3.2.16) показывает их идентичность. В самом деле, из принципа Ферма следует, что двух уравнений (3.4.11) и (3.4.12) должно быть достаточно для определения траектории луча. Это свидетельствует о том, что z -компонента уравнения луча является лишней. Действительно, можно показать, что соответствующее уравнение для z можно получить из уравнений для x и y . Для облегчения этого доказательства преобразуем соотношение (3.4.3) к виду

$$\begin{aligned} \sqrt{1+x'^2+y'^2} &= \frac{ds}{dz} = \frac{ds}{\sqrt{(ds)^2 - dx^2 - dy^2}} = \\ &= \frac{1}{\sqrt{1 - \dot{x}^2 - \dot{y}^2}}. \end{aligned} \quad (3.4.13)$$

Здесь

$$\dot{x} = \frac{dx}{ds}, \quad \dot{y} = \frac{dy}{ds}. \quad (3.4.14)$$

Далее

$$\begin{aligned} \frac{d}{ds} \left(n \frac{dz}{ds} \right) &= \frac{d}{ds} \left(n \sqrt{1 - \dot{x}^2 - \dot{y}^2} \right) = \\ &= \frac{dn}{ds} \sqrt{1 - \dot{x}^2 - \dot{y}^2} - n \frac{\ddot{x}\dot{x} + \ddot{y}\dot{y}}{\sqrt{1 - \dot{x}^2 - \dot{y}^2}} = \\ &= \frac{(1 - \dot{x}^2 - \dot{y}^2) \frac{dn}{ds} - n(\ddot{x}\dot{x} + \ddot{y}\dot{y})}{\sqrt{1 - \dot{x}^2 - \dot{y}^2}}. \end{aligned} \quad (3.4.15)$$

Из формулы (3.4.11) путем умножения на \dot{x} получаем

$$\dot{n} \dot{x} \dot{x} = \dot{x} \frac{\partial n}{\partial x} - \dot{x}^2 \frac{dn}{ds},$$

и из (3.4.12) аналогично следует

$$n \dot{y} \ddot{y} = \dot{y} \frac{\partial n}{\partial y} - \dot{y}^2 \frac{dn}{ds}.$$

Используя эти уравнения, упростим (3.4.15):

$$\frac{d}{ds} \left(n \frac{dz}{ds} \right) = \frac{\frac{dn}{ds} - \dot{x} \frac{\partial n}{\partial x} - \dot{y} \frac{\partial n}{\partial y}}{\sqrt{1 - \dot{x}^2 - \dot{y}^2}}.$$

С помощью соотношения

$$\frac{dn}{ds} = \frac{\partial n}{\partial z} \frac{dz}{ds} + \frac{\partial n}{\partial x} \frac{dx}{ds} + \frac{\partial n}{\partial y} \frac{dy}{ds}$$

и используя (3.4.13), получаем окончательно

$$\frac{d}{ds} \left(n \frac{dz}{ds} \right) = \frac{\partial n}{\partial z}. \quad (3.4.16)$$

Таким образом, получена z -составляющая векторного уравнения луча (3.2.16) при условии, что не все три компоненты этого уравнения являются независимыми. Уравнений (3.4.11) и (3.4.12) достаточно для описания траекторий лучей.

3.5. ГАМИЛЬТОНОВА ФОРМУЛИРОВКА ЛУЧЕВОЙ ОПТИКИ

Аналогия между лучевой оптикой и механикой становится наиболее явной, если выразить уравнения лучевой оптики в гамильтоновой форме. Предполагается, что чита-

тель знаком с уравнениями движения Гамильтона [19].

Переход от уравнений Эйлера (в механике они обычно называются уравнениями Лагранжа) к уравнениям Гамильтона достигается следующим образом. Начнем с введения обобщенных импульсов p_x и p_y , которые рассматриваются как переменные, канонически сопряженные с x и y . Обобщенные импульсы задаются уравнениями

$$p_x = \frac{\partial L}{\partial x'}, \quad (3.5.1)$$

$$p_y = \frac{\partial L}{\partial y'}. \quad (3.5.2)$$

Далее определим гамильтониан соотношением

$$H(x, y, p_x, p_y) = p_x x' + p_y y' - L. \quad (3.5.3)$$

Лагранжиан L дается формулой (3.4.6). Ключевым моментом в этом определении является выбор независимых переменных для гамильтониана H . Предполагается, что лагранжиан зависит от x, y, x' и y' , как это следует из (3.4.6). По определению независимыми переменными гамильтониана являются x, y, p_x и p_y . Замена переменных достигается, если заменить x' и y' на p_x и p_y с помощью формул (3.5.1) и (3.5.2). Наличие различных независимых переменных указывает на то, что p_x, p_y, x' и y' считаются здесь независимыми от x и y , тогда как x' и y' зависят от p_x и p_y . Воспользовавшись взаимной зависимостью этих переменных, получаем

$$\frac{\partial H}{\partial p_x} = x' + p_x \frac{\partial x'}{\partial p_x} + p_y \frac{\partial y'}{\partial p_x} - \frac{\partial L}{\partial x'} \frac{\partial x'}{\partial p_x} - \frac{\partial L}{\partial y'} \frac{\partial y'}{\partial p_x}.$$

Использование соотношений (3.5.1) и (3.5.2) с учетом формул (3.4.4) дает одно из уравнений Гамильтона:

$$\frac{dx}{dz} = \frac{\partial H}{\partial p_x}. \quad (3.5.4)$$

С помощью аналогичного вычисления получаем также, что

$$\frac{dy}{dz} = \frac{\partial H}{\partial p_y}. \quad (3.5.5)$$

Остальные два уравнения находятся из (3.5.3).

$$\frac{\partial H}{\partial x} = -\frac{\partial L}{\partial x}. \quad (3.5.6)$$

Уравнения (3.4.7) и (3.5.1) позволяют записать

$$\frac{dp_x}{dz} = -\frac{\partial H}{\partial x}. \quad (3.5.7)$$

Аналогично получаем последнее из уравнений Гамильтона

$$\frac{dp_y}{dz} = -\frac{\partial H}{\partial y}. \quad (3.5.8)$$

Чтобы завершить формулировку лучевой оптики в гамильтоновой форме, необходимо записать гамильтониан через соответствующие переменные. Это достигается выражением x' и y' через p_x и p_y на основании формул (3.5.1) и (3.5.2). Из формулы (3.5.1) с учетом соотношения (3.4.6) имеем

$$p_x = \frac{nx'}{\sqrt{1+x'^2+y'^2}}, \quad (3.5.9)$$

а из (3.5.2) находим, что

$$p_y = \frac{ny'}{\sqrt{1+x'^2+y'^2}}. \quad (3.5.10)$$

Эти уравнения можно решить относительно x' и y' , что дает

$$x' = \frac{p_x}{\sqrt{n^2 - p_x^2 - p_y^2}} \quad (3.5.11)$$

и

$$y' = \frac{p_y}{\sqrt{n^2 - p_x^2 - p_y^2}}. \quad (3.5.12)$$

Используем эти соотношения, чтобы получить гамильтониан в его обычном виде:

$$H = -\sqrt{n^2 - p_x^2 - p_y^2}. \quad (3.5.13)$$

В лучевой оптике гамильтониан имеет некоторое сходство с релятивистской энергией точечной частицы с массой покоя m_0 [21]

$$E = c\sqrt{m_0^2 c^2 + p_x^2 + p_y^2 + p_z^2}. \quad (3.5.14)$$

Однако сходство между гамильтонианом лучевой оптики и гамильтонианом точечной частицы наиболее ярко выражено в «нерелятивистском» приближении. Аналогом механического нерелятивистского случая является параксиальное приближение, с которым мы уже встречались

в разд. 3.2 [см. уравнения (3.2.18) и (3.2.19)]. Производные x' и y' описывают наклон луча по отношению к оси z . В параксиальном приближении принимается, что наклон луча очень мал и поэтому можно принять

$$x' \ll 1, \quad y' \ll 1. \quad (3.5.15)$$

Соотношения (3.5.9) и (3.5.10) требуют, чтобы выполнялись следующие условия:

$$p_x \ll n, \quad p_y \ll n. \quad (3.5.16)$$

Это позволяет разложить квадратный корень в (3.5.13) и получить гамильтониан лучевой оптики в параксиальном приближении:

$$H = \frac{p_x^2 + p_y^2}{2n_0} - n. \quad (3.5.17)$$

Необходимо также выразить показатель преломления в виде суммы постоянной части n_0 и небольшой, изменяющейся в пространстве части Δn :

$$n = n_0 - \Delta n. \quad (3.5.18)$$

Это соотношение является необходимым для параксиального приближения, так как если показатель преломления резко изменяется в пространстве, то, вероятно, встретимся с лучами, имеющими большие наклоны. Условие

$$\Delta n \ll n_0 \quad (3.5.19)$$

вместе с (3.5.16) позволяет заменить n на n_0 в зависящей от p части гамильтониана. Дифференциальные уравнения лучевой оптики неприменимы при разрывах непрерывности диэлектрической проницаемости, когда Δn не является малой. Однако разрывы непрерывности могут быть учтены путем сшивания решений для непрерывных участков с помощью граничных условий.

Параксиальный гамильтониан геометрической оптики тесно связан с нерелятивистским гамильтонианом механики точечных частиц [19]:

$$H = \frac{p_x^2 + p_y^2 + p_z^2}{2m} + V. \quad (3.5.20)$$

Задачи лучевой оптики имеют на одно измерение меньше по сравнению с соответствующими задачами механики

точечных частиц. Потенциал частицы V заменяется на показатель преломления оптической среды. Разница в знаке двух потенциальных членов не является существенной. Действительно, можно использовать Δn из (3.5.18) в качестве потенциала лучевой оптики и достичь полного соответствия даже в знаке потенциального члена. Аддитивная постоянная n_0 не имеет физического смысла, так как любой потенциал определяется с точностью до произвольной постоянной. Уравнение в частных производных Гамильтона — Якоби [19]

$$\frac{\partial S}{\partial t} + H\left(x, y, z, \frac{\partial S}{\partial x}, \frac{\partial S}{\partial y}, \frac{\partial S}{\partial z}\right) = 0 \quad (3.5.21)$$

играет важную роль в гамильтоновой механике точечных частиц. Функция S введена как решение уравнения в частных производных Гамильтона — Якоби (3.5.21). Производные S по координатам заменяют импульсы в функции Гамильтона. В соответствии с общей связью между механикой и лучевой оптикой нужно заменить временную координату t пространственной переменной z и понизить размерность задачи на единицу. Таким образом получаем уравнение Гамильтона — Якоби для лучевой оптики

$$\frac{\partial S}{\partial z} = -H\left(x, y, \frac{\partial S}{\partial x}, \frac{\partial S}{\partial y}\right). \quad (3.5.22)$$

Возводя в квадрат обе части этого уравнения и учитывая точное выражение (3.5.13) для гамильтониана, получаем

$$\left(\frac{\partial S}{\partial z}\right)^2 = n^2 - \left(\frac{\partial S}{\partial x}\right)^2 - \left(\frac{\partial S}{\partial y}\right)^2. \quad (3.5.23)$$

Как видим, вновь пришли к уравнению эйконала (3.2.5)

$$(\nabla S)^2 = n^2. \quad (3.5.24)$$

Таким образом, мы описали полный круг, получив уравнение эйконала, которое являлось исходным при выводе уравнения луча в разд. 3.2. Поэтому вывод уравнений геометрической оптики из принципа Ферма полностью эквивалентен предыдущему выводу из приведенного волнового уравнения. Различные выводы законов механики или лучевой оптики из вариационного принципа типичны для вариационных задач такого рода. Эта взаимосвязь между вариационным принципом, уравнениями дви-

жения в форме лагранжиана или гамильтониана и уравнением в частных производных Гамильтона — Якоби хорошо прослежена в классической книге Куранта [4].

Как увидим в следующем разделе, можно пойти еще дальше и вывести волновое уравнение как уравнение Клейна — Гордона релятивистской механики для лучевой оптики.

Возвращаясь к дифференциальным уравнениям Гамильтона (3.5.4), (3.5.5), (3.5.7) и (3.5.8), покажем, что они тоже эквивалентны уравнению луча. Используя формулы (3.5.4) и (3.5.5), из гамильтониана (3.5.13) получаем

$$\frac{dx}{dz} = \frac{p_x}{\sqrt{n^2 - p_x^2 - p_y^2}} \quad (3.5.25)$$

и

$$\frac{dy}{dz} = \frac{p_y}{\sqrt{n^2 - p_x^2 - p_y^2}}. \quad (3.5.26)$$

Эти уравнения уже были получены [см. (3.5.11) и (3.5.12)]. Из формул (3.5.7) и (3.5.8) имеем

$$\frac{dp_x}{dz} = \frac{n}{\sqrt{n^2 - p_x^2 - p_y^2}} \frac{\partial n}{\partial x}, \quad (3.5.27)$$

$$\frac{\partial p_y}{\partial z} = \frac{n}{\sqrt{n^2 - p_x^2 - p_y^2}} \frac{\partial n}{\partial y}. \quad (3.5.28)$$

Из выражений (3.5.25) и (3.5.26) находим

$$\frac{ds}{dz} = \sqrt{1 + x'^2 + y'^2} = \frac{n}{\sqrt{n^2 - p_x^2 - p_y^2}}. \quad (3.5.29)$$

Эта формула позволяет упростить (3.5.27) и (3.5.28):

$$\frac{dp_x}{ds} = \frac{\partial n}{\partial x}, \quad (3.5.30)$$

$$\frac{dp_y}{ds} = \frac{\partial n}{\partial y}. \quad (3.5.31)$$

Используя еще раз формулу (3.5.29), из равенств (3.5.25) и (3.5.26) получаем обобщенные импульсы в простом виде:

$$p_x = n \frac{dx}{ds} \quad (3.5.32)$$

и

$$p_y = n \frac{dy}{ds}. \quad (3.5.33)$$

Подстановка выражения (3.5.32) в равенство (3.5.30) дает составляющую x уравнения луча (3.2.16) или (3.4.14). Составляющая y вычисляется из двух других уравнений.

Параксиальное приближение лучевого уравнения получается сразу же, если воспользоваться параксиальным приближением гамильтониана в уравнениях Гамильтона. Из формул (3.5.4) и (3.5.17) имеем

$$\frac{dx}{dz} = \frac{1}{n_0} p_x. \quad (3.5.34)$$

Уравнения (3.5.7) и (3.5.17) приводят к соотношению

$$\frac{dp_x}{dz} = \frac{\partial n}{\partial x}. \quad (3.5.35)$$

Объединение равенств (3.5.34) и (3.5.35) дает x -составляющую параксиального лучевого уравнения

$$n_0 \frac{d^2x}{dz^2} = \frac{\partial n}{\partial x}. \quad (3.5.36)$$

Соответствующая y -составляющая равна

$$n_0 \frac{d^2y}{dz^2} = \frac{\partial n}{\partial y}. \quad (3.5.37)$$

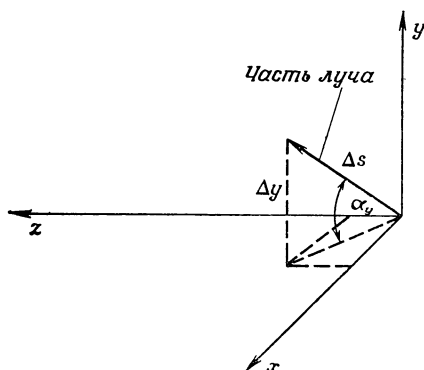
Эти уравнения имеют здесь более подходящий вид, чем форма (3.2.19), так как показатель преломления в левой части заменен его средним значением n_0 . Это необходимо в случае параксиального приближения, так как произведения Δn из (3.5.18) и производных от x и y по z являются членами второго порядка малости, которыми следует пренебречь в случае последовательного приближения первого порядка.

На фиг. 3.5.1 и 3.5.2 показаны участок светового луча и углы, которые он образует с плоскостями x , z и y , z . Обобщенные моменты (3.5.32) и (3.5.33) можно выразить через углы α_x и α_y :

$$p_x = n \sin \alpha_x, \quad (3.5.38)$$

$$p_y = n \sin \alpha_y. \quad (3.5.39)$$

Если считать плоскость x , y границей между двумя средами с различными диэлектрическими проницаемостями и расположить систему координат так, что луч, проходящий через границу из одной среды в другую, распростра-

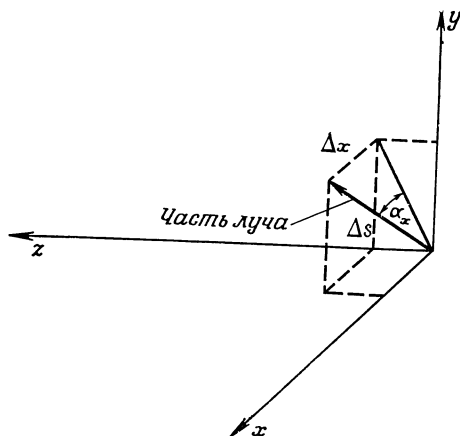


Ф и г. 3.5.1. Определение угла α_y .

няется в плоскости y, z , то можно выразить закон Снеллиуса (3.3.6) через обобщенные импульсы следующим образом:

$$\begin{aligned} (p_x)_1 &= (p_x)_2 = 0, \\ (p_y)_1 &= (p_y)_2. \end{aligned} \quad (3.5.40)$$

Обобщенные импульсы, таким образом, сохраняются при прохождении луча из одной среды в другую. Этот резуль-



Ф и г. 3.5.2. Определение угла α_x .

тат следует из соотношений (3.5.30) и (3.5.31), потому что при нашем выборе системы координат везде (включая границу) $\partial n/\partial x = 0$, $\partial n/\partial y = 0$, и, следовательно, p_x и p_y должны быть постоянными.

Обобщенный импульс луча нельзя путать с импульсом фотона. Они не имеют ничего общего. Импульс фотона это действительный механический импульс, которым обладает фотон, тогда как обобщенный импульс светового луча был получен чисто формально из гамильтонова формализма лучевой оптики. Обобщенный импульс светового луча описывает его наклон по отношению к фиксированным координатным осям.

3.6. КВАНТОВАЯ ТЕОРИЯ СВЕТОВЫХ ЛУЧЕЙ¹⁾

Принцип Ферма позволил получить основные уравнения геометрической оптики, ранее выведенные из приведенного волнового уравнения. Однако чисто алгебраические выкладки не могут привести к успеху при получении волнового уравнения из принципа Ферма, так как мы пренебрегли некоторыми членами при выводе уравнений лучевой оптики из волновой оптики. Эти отброшенные члены лишний раз напоминают о том, что геометрическая оптика является лишь приближением волновой оптики и не может описывать оптические явления так же полно, как волновая оптика. В этом разделе покажем, что волновое уравнение все же может быть получено из лучевой оптики, но не просто с помощью алгебраических преобразований, а скорее при использовании дополнительного условия — квантования. Применение правил квантовой механики к лучевой оптике приводит к квантовой теории световых лучей [42]. Оказывается, однако, что такая теория идентична скалярной волновой теории, использующей приведенное волновое уравнение. В предыдущем разделе было указано, что точные (в рамках лучевой оптики) уравнения лучевой оптики эквивалентны уравнениям релятивистской механики и что уравнения параксиальной оптики эквивалентны уравнениям нерелятивистской механики. Эта

¹⁾ Этот раздел не является существенным для понимания дальнейшего изложения. Читатели, не знакомые с квантовой механикой, могут его пропустить.

аналогия приводит нас к квантовой теории световых лучей. Приведенное волновое уравнение получается из эквивалента релятивистского уравнения Клейна — Гордона [22] волновой механики, тогда как геометрооптический эквивалент обычного нерелятивистского уравнения Шредингера следует из параксиального приближения. Эйхман [66] сумел сделать еще один шаг. Он построил теорию лучевой оптики в духе Дирака и показал, что уравнение Дирака для световых лучей эквивалентно уравнениям Максвелла, не зависящим от времени.

Квантование в физической теории осуществляется путем замены физических величин операторами. В волновой механике координаты по-прежнему остаются числами, но канонически сопряженные переменные — импульсы — становятся дифференциальными операторами [5, 10]:

$$p_x = -i\kappa \frac{\partial}{\partial x} \quad (3.6.1)$$

и

$$p_y = -i\kappa \frac{\partial}{\partial y}. \quad (3.6.2)$$

Мы воздержались от обозначения постоянной в уравнениях (3.6.1) и (3.6.2) обычным символом \hbar , так как наша квантовая теория несколько отличается от обычной квантовой механики дискретных частиц. Временная координата механики заменяется на координату z . Поэтому временная координата в постоянной Планка также должна быть заменена координатой длины, и \hbar уже нельзя считать имеющей размерность энергии, умноженной на время. Вместо этого постоянная κ , введенная вместо \hbar , должна иметь размерность гамильтониана, умноженного на длину. Гамильтониан (3.5.13) является безразмерным, поэтому κ имеет размерность длины.

В классической квантовой механике энергия оператора Гамильтона выражается с помощью производной по времени. Поэтому мы применяем соотношение

$$H = i\kappa \frac{\partial}{\partial z}, \quad (3.6.3)$$

опять используя соответствие между переменными «время» и «длина».

В релятивистской квантовой механике принято возводить в квадрат соотношение (3.6.3), т. е. не использовать для операторов выражения через квадратные корни. Поэтому запишем

$$H^2 = -\kappa^2 \frac{\partial^2}{\partial z^2}. \quad (3.6.4)$$

Применение этого операторного соотношения к волновой функции дает геометрикооптический эквивалент релятивистского волнового уравнения. Это уравнение называют уравнением Клейна — Гордона [22], чтобы отличить его от обычного нерелятивистского уравнения Шредингера. Используя формулы (3.5.13), (3.6.1) и (3.6.2) и применяя (3.6.4) к волновой функции, получаем уравнение Клейна — Гордона квантовой теории лучевой оптики:

$$n^2 \psi + \kappa^2 \frac{\partial^2 \psi}{\partial x^2} + \kappa^2 \frac{\partial^2 \psi}{\partial y^2} = -\kappa^2 \frac{\partial^2 \psi}{\partial z^2}. \quad (3.6.5)$$

Перегруппировка членов уравнения дает

$$\nabla^2 \psi + \frac{n^2}{\kappa^2} \psi = 0. \quad (3.6.6)$$

Очевидно, что это волновое уравнение квантовой теории геометрической оптики идентично приведенному волновому уравнению (3.2.1), которое для данного, интересующего нас случая запишем в виде

$$\nabla^2 \psi + \left(\frac{2\pi n}{\lambda_0} \right)^2 \psi = 0 \quad (3.6.7)$$

с тем, чтобы использовать понятие длины волны в свободном пространстве λ_0 . Сравнение двух уравнений позволяет определить постоянную κ :

$$\kappa = \frac{\lambda_0}{2\pi}. \quad (3.6.8)$$

Это более чем удовлетворительный результат. Постоянная κ эквивалентна квантовомеханической постоянной \hbar . Для квантовой теории характерно, что ее результаты совпадают с результатами классической механики при $\hbar \rightarrow 0$. Можно ожидать, что результаты квантовой теории лучевой оптики совпадают с результатами геометрической оптики при $\kappa \rightarrow 0$. Однако, как видим теперь, этот предел эквивалентен пределу $\lambda_0 \rightarrow 0$. Из предыдущего вывода уравнений

лучевой оптики из волновой оптики известно, что уравнения лучевой оптики становятся точными при $\lambda_0 \rightarrow 0$. Очень удобно, что постоянная Планка квантовой теории лучевой оптики оказывается длиной волны света в вакууме.

Можно непосредственно использовать все хорошо известные результаты квантовой механики и применить их к квантовой теории лучевой оптики, которая, как мы показали, идентична скалярной волновой теории света или любого другого явления, описываемого волновым уравнением. На практике установление эквивалентности волновой оптики и квантовой теории лучей приносит больше пользы лучевой оптике, чем волновой. Волновая оптика является сама по себе полной теорией. Однако граница применимости лучевой оптики — вопрос открытый, так как лучевая оптика является всего лишь приближенной теорией. Поскольку лучевая оптика — это «классическая механика» волновой оптики, то можно использовать наши знания классической и волновой механики в качестве руководства для оценки справедливости и применимости лучевой оптики. Квантовая механика не заменяет классическую механику. Последняя используется в тех областях, где она считается применимой. Известно, например, что движение электронов в электрическом и магнитном полях с большой точностью описывается законами классической механики. Только в случае очень сильных полей, например вблизи ядер, мы вынуждены использовать квантовую механику для правильного описания движения электронов. В этом случае классическая механика бессильна. Известно также, что классическая механика неприменима, когда электрон проявляет свою волновую природу. Эти аналогии могут быть использованы в качестве руководства при выборе лучевой или волновой оптики для конкретной оптической задачи, а теорема Эренфеста [22] помогает установить пределы классической механики. Позднее в этом разделе будут рассмотрены эти вопросы и сделаны соответствующие выводы о применимости лучевой оптики.

Первым промежуточным результатом развитой здесь квантовой теории лучевой оптики является параксиальное приближение для приведенного волнового уравнения. Мы неоднократно указывали, что параксиальное приближение соответствует нерелятивистской механике. Подста-

новка параксиального гамильтониана (3.5.17) в операторное уравнение (3.6.3) позволяет получить геометрооптический эквивалент нерелятивистского уравнения Шредингера. Это уравнение также является параксиальным приближением приведенного волнового уравнения. Оно имеет вид

$$-\frac{\lambda_0^2}{8\pi^2 n_0} \left(\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} \right) - n\psi = i \frac{\lambda_0}{2\pi} \frac{\partial \psi}{\partial z}. \quad (3.6.9)$$

Перегрупуировав члены, получим параксиальное волновое уравнение

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + i \frac{4\pi}{\lambda_0} n_0 \frac{\partial \psi}{\partial z} + \frac{8\pi^2}{\lambda_0^2} n_0 n \psi = 0. \quad (3.6.10)$$

Уравнение (3.6.9) имеет вид нерелятивистского уравнения Шредингера.

Другим важным уравнением квантовой механики является уравнение собственных значений энергии

$$H\psi = E\psi, \quad (3.6.11)$$

которое можно переписать в виде, более удобном для релятивистского гамильтониана:

$$H^2\psi = E^2\psi. \quad (3.6.12)$$

Из формул (3.5.17) и (3.6.11) получаем уравнение собственных значений энергии для параксиального случая

$$-\frac{\lambda_0^2}{8\pi^2 n_0} \left(\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} \right) - n\psi = E\psi. \quad (3.6.13)$$

Используя волновую функцию вида [знак показателя экспоненты сравните с (3.6.27)]

$$\psi = \psi_0(x, y) e^{i\beta z} \quad (3.6.14)$$

с постоянной распространения β и сравнивая уравнение (3.6.9) с уравнением собственных значений энергии (3.6.13), находим, что собственное значение E пропорционально постоянной распространения

$$E = -\frac{\lambda_0}{2\pi} \beta. \quad (3.6.15)$$

Тот же результат можно получить из равенства (3.6.12). Выражение (3.6.15) является эквивалентом квантовоме-

ханического выражения

$$E = \frac{\hbar}{2\pi} \omega. \quad (3.6.16)$$

Разница в знаке между (3.6.15) и (3.6.16) возникает только из-за того, что функция (3.6.14) описывает волну, распространяющуюся в отрицательном направлении z . Постоянная Планка \hbar превращается в λ_0 ; как видно, частота ω классической квантовой механики в квантовой теории лучей заменяется постоянной распространения — β в направлении оси z . Уравнения собственных значений энергии (3.6.11) и (3.6.12) эквивалентны граничной задаче нахождения мод в оптическом волноводе. Подобная задача будет решена в другой главе для оптического волновода с параболическим распределением показателя преломления.

Волновая функция ψ квантовой механики лучей является обычной скалярной волновой функцией волновой оптики. В данной теории она получает дополнительную интерпретацию как амплитуда вероятности. Волновая функция описывает состояние статистического ансамбля лучей. Квадрат ее абсолютной величины

$$\bar{P} = |\psi(x, y, z)|^2 \quad (3.6.17)$$

приобретает значение плотности вероятностей прохождения светового луча через единичную площадку в плоскости x, y с координатой z . Следовательно, полная вероятность прохождения световым лучом данной площади A определяется как

$$P = \int_A |\psi|^2 dA. \quad (3.6.18)$$

Требование того, что каждый луч света должен пересекать любую плоскость, перпендикулярную оси z , дает условие нормировки

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\psi(x, y, z)|^2 dx dy = 1. \quad (3.6.19)$$

Волновая функция описывает вероятность нахождения луча в области световых лучей, статистическое состояние которых характеризуется функцией ψ .

Можно также записать волновое уравнение в пространстве импульсов [10]. Для этого преобразования нужно знать собственные функции и собственные значения импульсов. Уравнение собственных значений для оператора импульса p_x , определяемого формулой (3.6.1), имеет вид

$$-i\kappa \frac{\partial \psi_{p_x}}{\partial x} = p'_x \psi_{p_x}. \quad (3.6.20)$$

Решением этого уравнения является собственная функция импульса, соответствующая к собственному значению импульса p'_x :

$$\psi_{p'_x} = \frac{1}{\sqrt{2\pi}} e^{i(p'_x/\kappa)x} e^{i\phi(z)}. \quad (3.6.21)$$

Множитель $(2\pi)^{-1/2}$ необходим для правильной нормировки собственной функции импульса; фазовая функция ϕ произвольна. Поскольку спектр собственных значений собственной функции непрерывен, эта функция должна быть нормирована на дельта-функцию:

$$\int_{-\infty}^{\infty} \psi_{p'_x}^* \psi_{p''_x} dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i(1/\kappa)(p''_x - p'_x)x} dx = \delta(p''_x - p'_x). \quad (3.6.22)$$

Чтобы можно было дать физическую интерпретацию собственной функции импульса, нужно выбрать ее фазу ϕ такой, что ψ_{p_x} также становится решением уравнения Клейна — Гордона (3.6.5). Пробное решение

$$\psi_{p_x} = \frac{1}{\sqrt{2\pi}} e^{i[(p'_x/\kappa)x + (p'_z/\kappa)z]} \quad (3.6.23)$$

является решением релятивистского уравнения Шредингера (3.6.5), если имеет место следующее соотношение:

$$p_x^2 + p_z^2 = n^2, \quad (3.6.24)$$

где показатель преломления есть постоянная ($n = n_0$). Это означает, что собственная функция импульса может быть физическим состоянием, т. е. удовлетворять уравнению Шредингера, только в том случае, когда показатель преломления n постоянен в пространстве. Если это требование удовлетворено, собственное состояние импульса существует и представляет собой плоскую волну, распро-

страняющуюся в направлении вектора

$$\mathbf{p} = p_x \mathbf{e}_x + p_z \mathbf{e}_z \quad (3.6.25)$$

(\mathbf{e}_x и \mathbf{e}_z — единичные векторы в направлении осей x и z). Этот результат согласуется с нашей физической интерпретацией лучевой оптики. Собственное состояние импульса может существовать как физическое состояние только тогда, когда каждое измерение (проведенное для определения величины «импульса» лучей) дает один и тот же точный результат. В соответствии с формулой (3.5.38) «импульс» лучей определяет их наклон. Световой луч может иметь определенный наклон в рассматриваемой системе координат, если он соответствует плоской волне. Лучи являются траекториями, ортогональными фазовым фронтам сопровождающей волны. Если фазовые фронты изогнуты, наклоны лучей — их «импульсы» — различны в разных частях пространства. Поэтому собственное «импульсное» состояние должно быть плоской волной. При выбранных нами постоянных выражение (3.6.23) описывает плоскую волну, распространяющуюся в отрицательном направлении вектора \mathbf{p} , если используем для него временную зависимость в виде множителя

$$e^{i\omega t}. \quad (3.6.26)$$

Волновая функция, являющаяся одновременно собственной функцией операторов p_x и p_y , представляет собой плоскую волну более общего вида

$$\psi = \frac{1}{2\pi} e^{i(p_x x + p_y y + p_z z)/\hbar}. \quad (3.6.27)$$

Эта волна является решением релятивистского уравнения Шредингера (3.6.5) в случае постоянного $n = n_0$, если выполняется условие

$$p_x^2 + p_y^2 + p_z^2 = n_0^2. \quad (3.6.28)$$

Если бы мы потребовали, чтобы собственная функция импульса была решением нерелятивистского уравнения Шредингера (3.6.9), то получили бы соотношение

$$\frac{1}{2n_0} (p_x^2 + p_y^2) + p_z = n_0, \quad (3.6.29)$$

которое, конечно, соответствует параксиальному приближению уравнения (3.6.28) при $p_z \approx n_0$, т. е. своему параксиальному значению.

Теперь можно выполнить преобразование произвольной волновой функции ψ в пространство импульсов. Это преобразование имеет вид

$$\phi(p_x, p_y, z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi(x, y, z) e^{-i(\kappa)(p_x x + p_y y)} dx dy. \quad (3.6.30)$$

Вероятность нахождения луча с составляющими «импульса» p_x, p_y в интервале $dp_x dp_y$ дается формулой

$$dP = |\phi(p_x, p_y)|^2 dp_x dp_y. \quad (3.6.31)$$

Обратное преобразование (3.6.30) представляет трудную задачу, так как диапазон физически возможных значений p_x и p_y заключен между $-n$ и n . Однако если κ — очень малая величина, так что диапазон значений p_x/κ заключен между $-2\pi n/\lambda_0$ и $+2\pi n/\lambda_0$, то при очень малых значениях λ_0 интервал интегрирования весьма широк. Можно предположить, что функция $\phi(p_x, p_y)$ становится исчезающе малой вблизи границ интервала интегрирования, так как вероятность прохождения лучей перпендикулярно оси в большинстве практических случаев должна быть очень близкой к нулю. Заметной ошибки не будет допущено, если положить $\phi = 0$ вне физического интервала интегрирования и распространить этот интервал от $-\infty$ до $+\infty$. В таком случае преобразование (3.6.30) можно рассматривать как интегральное преобразование Фурье с обратным преобразованием

$$\begin{aligned} \psi(x, y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(p_x, p_y) e^{i(\kappa)(p_x x + p_y y)} \times \\ \times d\left(\frac{p_x}{\kappa}\right) d\left(\frac{p_y}{\kappa}\right). \end{aligned} \quad (3.6.32)$$

Трактовка волновой функции как амплитуды вероятности позволяет сразу же определить математическое ожидание [10, 22] всех операторов, появляющихся в квантовой теории лучей. Для любого оператора A можно определить

математическое ожидание уравнением

$$\langle A \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^*(x, y) A \psi(x, y) dx dy. \quad (3.6.33)$$

Математическое ожидание можно выразить в пространстве импульсов соотношением

$$\langle A \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi^*(p_x, p_y) A_p \phi(p_x, p_y) d\left(\frac{p_x}{\kappa}\right) d\left(\frac{p_y}{\kappa}\right). \quad (3.6.34)$$

Оператор A_p является эквивалентом A в пространстве импульсов [10]. Теперь мы подготовлены к тому, чтобы рассмотреть теорему Эренфеста. Ранее было отмечено, что эта теорема служит мостом между классической и квантовой теориями. Теорема Эренфеста в релятивистской квантовой механике почти не применяется. Чаще всего она используется в нерелятивистской квантовой механике. Поэтому ограничимся параксиальным приближением и выведем теорему Эренфеста из нерелятивистского уравнения Шредингера (3.6.9), которое в операторной форме имеет вид

$$H\psi = i\kappa \frac{\partial \psi}{\partial z}, \quad (3.6.35)$$

где H определяется формулой (3.5.17). В теорему Эренфеста входят производные величин математического ожидания координат x и y и «импульсов» p_x и p_y . Поэтому возьмем производную от величины математического ожидания координаты x :

$$\begin{aligned} \frac{d\langle x \rangle}{dz} &= \frac{d}{dz} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^* x \psi dx dy = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{d\psi^*}{dz} x \psi + \psi^* x \frac{d\psi}{dz} \right) dx dy. \end{aligned} \quad (3.6.36)$$

Производные от волновой функции можно исключить с помощью уравнения Шредингера (3.6.35)

$$\frac{d\langle x \rangle}{dz} = \frac{i}{\kappa} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [(H\psi^*) x \psi - \psi^* x H\psi] dx dy.$$

Для любого гамильтониана можно произвести следующую перегруппировку членов [22]:

$$\frac{d\langle x \rangle}{dz} = \frac{i}{\kappa} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^* (Hx - xH) \psi dx dy. \quad (3.6.37)$$

В нашем частном случае это можно проверить выполнением частичного интегрирования с использованием специального вида гамильтониана, который получается подстановкой выражений (3.6.1) и (3.6.2) в (3.5.17). Можно легко оценить операторное выражение в круглых скобках. Начнем с коммутаторов

$$xp_x - p_x x = i\kappa \quad (3.6.38)$$

и

$$xp_y - p_y x = 0. \quad (3.6.39)$$

Эти фундаментальные соотношения квантовой механики [40] хорошо известны. Их можно подтвердить, применяя к произвольной волновой функции и используя точные выражения (3.6.1) и (3.6.2) для p_x и p_y . Повторное применение тех же коммутаторов дает

$$xp_x^2 - p_x^2 x = p_x x p_x + i\kappa p_x - p_x x p_x + i\kappa p_x = 2i\kappa p_x, \quad (3.6.40)$$

$$xp_y^2 - p_y^2 x = 0. \quad (3.6.41)$$

Используя эти коммутационные соотношения, сразу же получаем желаемый результат:

$$\begin{aligned} Hx - xH &= - \left[\frac{1}{2n_0} (xp_x^2 - p_x^2 x + xp_y^2 - p_y^2 x) + xn - nx \right] = \\ &= - \frac{i\kappa}{n_0} p_x. \end{aligned} \quad (3.6.42)$$

Уравнение (3.6.37) теперь принимает вид одного из уравнений теоремы Эренфеста

$$\frac{d\langle x \rangle}{dz} = \frac{1}{n_0} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^* p_x \psi dx dy; \quad (3.6.43)$$

используя определение математического ожидания операторов (3.6.33), можно записать его в виде

$$\frac{d\langle x \rangle}{dz} = \frac{1}{n_0} \langle p_x \rangle. \quad (3.6.44)$$

Аналогичное соотношение можно получить для y -компоненты:

$$\frac{d\langle y \rangle}{dt} = \frac{1}{n_0} \langle p_y \rangle. \quad (3.6.45)$$

Необходимо также определить производную от математического ожидания p_x :

$$\begin{aligned} \frac{d\langle p_x \rangle}{dz} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{d\psi^*}{dz} p_x \psi + \psi^* p_x \frac{d\psi}{dz} \right) dx dy = \\ &= \frac{i}{\hbar} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \psi^* (H p_x - p_x H) \psi dx dy. \end{aligned} \quad (3.6.46)$$

Для определения этого коммутатора необходимо использовать коммутационное соотношение

$$p_x f(x) - f(x) p_x = -i\hbar \frac{\partial f(x)}{\partial x}. \quad (3.6.47)$$

Это соотношение непосредственно проверяется применением его к произвольной волновой функции с учетом (3.6.1). Используя это коммутационное соотношение и учитывая то обстоятельство, что p_x коммутирует сам с собой и p_x , из формулы (3.6.46) получаем

$$\frac{d\langle p_x \rangle}{dz} = \left\langle \frac{\partial n}{\partial x} \right\rangle. \quad (3.6.48)$$

Аналогичное соотношение имеет место и для y -компоненты:

$$\frac{d\langle p_y \rangle}{dz} = \left\langle \frac{\partial n}{\partial y} \right\rangle. \quad (3.6.49)$$

Уравнения (3.6.44), (3.6.45), (3.6.48) и (3.6.49) составляют теорему Эренфеста. Эти уравнения очень похожи на уравнения параксиальных лучей (3.5.34) и (3.5.35). Действительно, можно объединить уравнения (3.6.44) и (3.6.48) в одно, которое имеет почти такой же вид, как уравнение параксиальных лучей (3.5.36):

$$n_0 \frac{d^2 \langle x \rangle}{dz^2} = \left\langle \frac{\partial n(x, y)}{\partial x} \right\rangle. \quad (3.6.50)$$

Аналогичное соотношение, конечно, справедливо и для y -компоненты. Если бы можно было представить уравне-

ние (3.6.50) как

$$n_0 \frac{d^2 \langle x \rangle}{dz^2} = \frac{\partial}{\partial \langle x \rangle} n(\langle x \rangle, \langle y \rangle), \quad (3.6.51)$$

то отсюда следовало бы, что величина математического ожидания координаты x светового луча сама «движется» подобно лучу. Однако уравнение (3.6.50) заметно отличается от (3.6.51) и такое утверждение не может быть сделано в общем случае. Можно лишь сказать, что если производная $\partial n(\langle x \rangle, \langle y \rangle) / \partial \langle x \rangle$ достаточно близка к $\langle \partial n(x, y) / \partial x \rangle$, то величина математического ожидания луча «движется» подобно реальному лучу. Это очень важная теорема, так как она утверждает, что центр тяжести светового поля (который совпадает с величиной математического ожидания $\langle x \rangle$) движется примерно так же, как световой луч. Такое сравнение лучей геометрической оптики и центра тяжести светового поля является ключом к интерпретации лучевой оптики. Мы используем лучевую оптику только для получения упрощенного и приближенного описания траекторий световых лучей. Было бы полезно иметь уверенность в том, что уравнение луча описывает траекторию центра тяжести действительного светового луча. В этом случае лучевая оптика точно указывала бы, где проходит большая часть светового поля. Однако сравнение полученного в действительности уравнения (3.6.50) и желаемого уравнения (3.6.51) показывает, что лучевая теория не обязательно предсказывает движение центра тяжести световых лучей. Чтобы более подробно изучить условия, при которых уравнение луча предсказывает движение центра тяжести светового поля, разложим показатель преломления в степенной ряд:

$$n = n_0 + n_1 x + n_2 y + n_3 x^2 + n_4 xy + n_5 y^2 + n_6 x^3 + n_7 x^2 y + \\ + n_8 xy^2 + n_9 y^3 + \dots \quad (3.6.52)$$

Производная этого разложения показателя преломления равна

$$\frac{\partial n}{\partial x} = n_1 + 2n_3 x + n_4 y + 3n_6 x^2 + 2n_7 xy + n_8 y^2 + \dots \quad (3.6.53)$$

Если бы в разложении (3.6.52) существовали только члены первого и второго порядка, то мы бы точно имели

$$\left\langle \frac{\partial}{\partial x} n(x, y) \right\rangle = \frac{\partial}{\partial \langle x \rangle} n(\langle x \rangle, \langle y \rangle). \quad (3.6.54)$$

Только в этом случае можно точно утверждать, что центр тяжести светового луча движется в соответствии с лучевой оптикой. Распределения показателя преломления, описываемые разложением вида

$$n = n_0 + n_1x + n_2y + n_3x^2 + n_4xy + n_5y^2, \quad (3.6.55)$$

конечно, возможны. Действительно, уже получены оптические волокна с параболическим распределением показателя преломления [23, 112]. Такая «квадратичная среда», которая может быть использована в качестве световых волноводов, будет рассмотрена в гл. 7. Среда с квадратичным распределением представляет большой интерес, так как известно, что решения уравнений лучевой оптики точно описывают движение центра тяжести светового луча, распространяющегося в ней. В случае среды более общего характера уравнения лучей также разрешимы, но не ясно, как эти математические лучи связаны с действительным движением светового поля. В среде с медленно и слабо изменяющимся показателем преломления распределение показателя преломления может быть приблизительно описано разложением, содержащим только члены первого и второго порядка. Можно утверждать, что луч описывает движение центра тяжести светового поля только в приближении, допускаемом разложением (3.6.55). В случае самого общего распределения коэффициента преломления мало что можно сказать о связи между решениями лучевой оптики и действительной траекторией светового луча.

В заключение рассмотрим принцип неопределенности применительно к данной квантовой теории световых лучей. В книгах по квантовой теории [10, 22] показывается, что коммутационное соотношение (3.6.38) и соответствующее соотношение для y -компонент приводят к следующим соотношениям неопределенности:

$$\Delta x \Delta p_x \geq \frac{1}{2} \kappa = \frac{\lambda_0}{4\pi} \quad (3.6.56)$$

и

$$\Delta y \Delta p_y \geq \frac{\lambda_0}{4\pi}. \quad (3.6.57)$$

Неопределенности x и p_x определяются выражениями вида

$$\Delta x = [\langle (x - \langle x \rangle)^2 \rangle]^{1/2} \quad (3.6.58)$$

и

$$\Delta p_x = [\langle (p_x - \langle p_x \rangle)^2 \rangle]^{1/2}. \quad (3.6.59)$$

Математические ожидания определяются по формуле (3.6.33). Согласно полученным соотношениям, если световой луч сформирован таким образом, что x известен с некоторой точностью, то p_x может быть измерен только с точностью, определяемой формулой (3.6.56). Рассмотрим два примера. Сначала предположим, что состояние луча описывается плоской волной вида (3.6.23). Это состояние является собственным «импульсом» и показывает, что каждое измерение наклона лучей должно давать величину p'_x . В этом случае положение луча совершенно неизвестно, так как бесконечная плоская волна занимает все пространство и нельзя определить какое-либо конкретное положение луча. Далее рассмотрим световой луч, прошедший через очень узкую щель. Положение луча, проходящего через щель, известно с точностью до ширины щели. Однако наклон выходящего луча становится все менее и менее определенным при сужении щели, так как световая волна дифрагирует в стороны по мере удаления поля от щели. Если предположить, что можно принять $\Delta x = d$, т. е. ширине щели, то из формулы (3.6.56) получим

$$\Delta p_x \geq \frac{\lambda_0}{4\pi d} \quad (3.6.60)$$

или, используя соотношение (3.5.38) для малых значений α , получим неопределенность угла луча для $n = 1$:

$$\Delta \alpha \geq \frac{\lambda_0}{4\pi d}. \quad (3.6.61)$$

Сравнение с формулой (2.3.28) показывает, что развитая ранее дифракционная теория приводила к угловому расширению примерно того же порядка величины. Полученный результат представляет собой всего лишь неравенство и не противоречит равенству (2.3.28). Однако естественно ожидать, что можно выбрать более благоприятные распределения поля, которые приведут к меньшему расширению луча по сравнению с диафрагмированной плоской волной с однородным распределением амплитуды в щели. Такими волновыми пакетами с минимальной неопределенностью являются гауссовы пучки (см. гл. 6),

3.7. ТЕОРЕМА ЛИУВИЛЛЯ [24]

Теорема Лиувилля играет важную роль в статистической механике. Приведенные выше рассуждения, подчеркивавшие тесное сходство лучевой оптики с механикой, подготовили нас к тому, чтобы не удивляться при обнаружении применимости теорем механики в оптике. Световые лучи обычно не рассматриваются статистическими теориями. Однако квантовая механика является статистической теорией, поэтому волновая оптика, будучи квантовой теорией лучевой оптики, сводится к статистической трактовке световых лучей. Статистические положения в строгом классическом смысле также с успехом могут быть применены для описания коллективного движения пучков лучей. Большая важность теоремы Лиувилля в оптике обусловлена ее способностью служить руководством в ситуациях, когда рассуждения, основанные на интуиции, могут привести к грубым ошибкам. Пример такого рода будет рассмотрен в гл. 4. Он включает обобщение концепции тонких линз как фазового корректора. Теорема Лиувилля может быть применена для проверки физической реализуемости математических предположений, которые выглядят интуитивно правдоподобными, но неумовным образом нарушают физические принципы. Теорема Лиувилля не часто обсуждается в книгах по оптике. Но она является мощным инструментом и содержит как частные случаи несколько хорошо известных оптических теорем.

Вывод теоремы Лиувилля можно найти в книгах по статистической механике [25]. Выведем ее здесь, чтобы не заставлять читателя обращаться к литературе по статистической механике. Отправная точка для формулирования и понимания теоремы — это понятие о фазовом пространстве. Ранее было показано, что в гамильтоновой формулировке лучевой оптики используются два рода координат: положение луча x и y и «импульсы» луча p_x и p_y . Вместе эти переменные образуют пространство четырех измерений (шести измерений в классической механике точечной частицы). Физическое состояние светового луча может быть представлено точкой в фазовом пространстве. Каждая точка в фазовом пространстве дает информацию о положении и угле наклона («импульсе») луча. Траектория в фазо-

вом пространстве дает нам такое же количество информации, какое мы получили бы, изучая положение и наклон луча в зависимости от координаты z . Переход от одного луча к многим лучам и, наконец, к статистике лучей естественным образом приводит к рассмотрению элементов объема в фазовом пространстве. Рассмотрим пучок лучей, заполняющих некоторую площадь в реальном пространстве и имеющих некоторый разброс по направлению. Каждый луч пучка в фазовом пространстве представлен точкой, слегка отличающейся по положению от любого другого луча данного пучка. Если положение лучей в реальном пространстве и углы их наклона ограничены, то представляющие их точки в фазовом пространстве заполнят некоторый конечный объем. Проследив за изменением светового пучка при его распространении через оптическую систему, можно заметить, что объем, который первоначально занимали лучи, изменяет свою форму и движется в фазовом пространстве. Вместо того чтобы изучать историю каждого отдельного луча в пучке, легче наблюдать движение объема в фазовом пространстве. Теорема Лиувилля дает сведения об объеме пучка лучей в фазовом пространстве. Ее можно выразить в нескольких эквивалентных формах. Через объем фазового пространства она формулируется следующим образом: *объем в фазовом пространстве, заполненный точками, соответствующими пучку лучей, остается постоянным*. Этот результат основан на уравнениях лучей. Сами лучи могут сильно отклоняться, вызывая деформацию и вытягивание объема фазового пространства. Однако объем, который первоначально был занят лучами в их исходном состоянии, не меняется даже при изменении его первоначальной формы.

Теорема Лиувилля основана на плотности ρ точек в фазовом пространстве. Плотность — это число точек в единице объема фазового пространства. Плотность удовлетворяет уравнению непрерывности

$$\frac{\partial \rho}{\partial z} + \operatorname{div}(\rho \mathbf{v}) = 0. \quad (3.7.1)$$

Заметим, что временная координата механики снова заменена на z . Согласно уравнению непрерывности, общее число точек должно оставаться постоянным. Справедли-

вость этого уравнения подтверждается тем, что изменение общего числа точек, заключенных внутри фиксированного объема V_f фазового пространства, должно объясняться движением точек как внутрь этого объема, так и наружу из него. Так как движение точек происходит вдоль оси z , то

$$\frac{\partial}{\partial z} \int_{V_f} \rho dV = - \int_S \rho \mathbf{v} \cdot \mathbf{n} dS. \quad (3.7.2)$$

Интеграл слева есть общее число точек в объеме V_f , тогда как интеграл справа представляет поток частиц с плотностью тока $\rho \mathbf{v}$ (\mathbf{v} — скорость частиц), который течет через замкнутую поверхность, окружающую V_f . Единичный вектор \mathbf{n} направлен по внешней нормали. Поэтому отрицательный поток означает, что в объеме накапливаются частицы. Это учтено отрицательным знаком правой части. Применение теоремы о дивергенции к интегралу в правой части приводит к уравнению (3.7.1), если учесть, что объем V_f произволен.

Мы рассматривали задачу о потоке так, как если бы он имел место в обычном трехмерном пространстве с координатой z , представляющей временную координату. Однако математические операции, которые были применены, не ограничены случаем трехмерного пространства, так что каждое утверждение применимо и к четырехмерному пространству. Дивергенция, появляющаяся в уравнении (3.7.1), является обобщением обычной трехмерной дивергенции на четырехмерное фазовое пространство. Можно записать уравнение (3.7.1) через составляющие следующим образом:

$$\frac{\partial \rho}{\partial z} + \sum_{i=1}^2 \left[\frac{\partial}{\partial x_i} \left(\rho \frac{dx_i}{dz} \right) - \frac{\partial}{\partial p_i} \left(\rho \frac{dp_i}{dz} \right) \right] = 0. \quad (3.7.3)$$

Здесь использованы обозначения $x_1 = x$, $x_2 = y$, $p_1 = p_x$ и $p_2 = p_y$. Координаты и импульсы представляют действительные лучи, которые подчиняются законам лучевого распространения. Использование соотношений (3.5.4) — (3.5.8) позволяет записать

$$\frac{\partial \rho}{\partial z} + \sum_{i=1}^2 \left[\frac{\partial}{\partial x_i} \left(\rho \frac{\partial H}{\partial p_i} \right) - \frac{\partial}{\partial p_i} \left(\rho \frac{\partial H}{\partial x_i} \right) \right] = 0. \quad (3.7.4)$$

Полезно помнить, что законы лучевой оптики использованы именно здесь. Вычисление производных от произведений позволяет сократить некоторые члены, так что в результате получаем

$$\frac{\partial \rho}{\partial z} + \sum_{i=1}^2 \left(\frac{\partial \rho}{\partial x_i} \frac{\partial H}{\partial p_i} - \frac{\partial \rho}{\partial p_i} \frac{\partial H}{\partial x_i} \right) = 0. \quad (3.7.5)$$

Повторное применение уравнений Гамильтона приводит к равенству

$$\frac{\partial \rho}{\partial z} + \sum_{i=1}^2 \left(\frac{\partial \rho}{\partial x_i} \frac{dx_i}{dz} + \frac{\partial \rho}{\partial p_i} \frac{dp_i}{dz} \right) = 0. \quad (3.7.6)$$

Плотность ρ является функцией x , y , p_x и p_y . Левая часть является полной производной ρ по z , так что получаем теорему Лиувилля в следующем виде:

$$\frac{d\rho}{dz} = 0. \quad (3.7.7)$$

Плотность точек в фазовом пространстве не зависит от координаты z . Можно использовать этот результат для получения второй формы теоремы Лиувилля. Заключим определенное число точек в объем V и будем изменять границы этого объема таким образом, чтобы общее число точек, содержащихся внутри, не изменялось. Тогда по определению

$$\frac{d}{dz} \int_V \rho dV = 0. \quad (3.7.8)$$

Можно расположить точки любым произвольным образом. Расположим их так, чтобы плотность ρ была постоянна внутри объема V . Это позволяет вынести ρ из-под знака интеграла и, используя равенство (3.7.7), поставить ее перед знаком производной. В результате получается уравнение

$$\frac{dV}{dz} = 0. \quad (3.7.9)$$

Это теорема Лиувилля в ее второй форме. Она утверждает, что объем, содержащий фиксированное число точек фазового пространства, не зависит от z . Этот объем переносится

вместе с точками, но численная величина объема не изменяется, хотя его форма может существенно изменяться.

Можно выразить теорему Лиувилля в третьей форме, которая наиболее полезна для наших целей. Получение этой последней формы требует небольшого введения. Возьмем луч во входной плоскости $z = z_i$ с координатами x_i, y_i, p_{xi}, p_{yi} . При прохождении луча от начальной точки $z = z_i$ до произвольной точки $z = z_f$ вдоль его траектории изменяются все значения координат. Обозначим переменные в произвольной точке луча через x_f, y_f, p_{xf}, p_{yf} . Рассматриваемый луч является одним из лучей пучка, который заполняет некоторый объем в фазовом пространстве. Соседние лучи имеют координаты, начальные и конечные значения которых несколько отличаются от координат выбранного луча. Следует учесть, что если изменять начальные координаты луча в фиксированной плоскости $z = z_f$, его конечные координаты в этой плоскости $z = z_f$ изменятся тоже. Действительно, конечные координаты являются функциями начальных координат:

$$x_f = x_f(x_i, y_i, p_{xi}, p_{yi}), \quad (3.7.10)$$

$$p_{xf} = p_{xf}(x_i, y_i, p_{xi}, p_{yi}). \quad (3.7.11)$$

Аналогичные функциональные соотношения справедливы и для двух других координат. Координата z считается постоянной и нет необходимости включать ее в эти функциональные соотношения. Далее нам необходим один результат из теории объемных интегралов. Если интеграл, выраженный через ряд переменных, необходимо записать через новый ряд переменных, то преобразование выполняется с помощью якобиана преобразования [26]. Для случая объемных интегралов это можно представить уравнением

$$dV_f = \frac{\partial(x_f, y_f, p_{xf}, p_{yf})}{\partial(x_i, y_i, p_{xi}, p_{yi})} dV_i. \quad (3.7.12)$$

Применим эту теорему к нашему ряду переменных. Уравнения (3.7.10) и (3.7.11) и остальные уравнения не выражают явно формы преобразования от одного ряда переменных к другому, поэтому применяется теория преобразования объемных интегралов. Теорема Лиувилля в ее второй форме (3.7.8) утверждает, что два элемента объема

должны быть одинаковыми, так как объем совокупности точек в фазовом пространстве не изменяется. Должно выполняться равенство

$$dV_f = dV_i. \quad (3.7.13)$$

Следовательно, якобиан, входящий в формулу (3.7.12), должен быть равен единице. Из определения якобиана получаем третью форму теоремы Лиувилля

$$\frac{\partial (x_f, y_f, p_{xf}, p_{yf})}{\partial (x_i, y_i, p_{xi}, p_{yi})} = \begin{vmatrix} \frac{\partial x_f}{\partial x_i} & \frac{\partial y_f}{\partial x_i} & \frac{\partial p_{xf}}{\partial x_i} & \frac{\partial p_{yf}}{\partial x_i} \\ \frac{\partial x_f}{\partial y_i} & \frac{\partial y_f}{\partial y_i} & \frac{\partial p_{xf}}{\partial y_i} & \frac{\partial p_{yf}}{\partial y_i} \\ \frac{\partial x_f}{\partial p_{xi}} & \frac{\partial y_f}{\partial p_{xi}} & \frac{\partial p_{xf}}{\partial p_{xi}} & \frac{\partial p_{yf}}{\partial p_{xi}} \\ \frac{\partial x_f}{\partial p_{yi}} & \frac{\partial y_f}{\partial p_{yi}} & \frac{\partial p_{xf}}{\partial p_{yi}} & \frac{\partial p_{yf}}{\partial p_{yi}} \end{vmatrix} = 1. \quad (3.7.14)$$

Теорема Лиувилля справедлива не только для лучей в средах с непрерывным распределением показателя преломления, но также и для лучей, пересекающих границу между средами с разными показателями преломления, и при отражении света от изогнутых зеркал. Доказательство справедливости теоремы при отражении света представляется читателю в качестве упражнения. Справедливость теоремы Лиувилля для диэлектрических границ может быть подтверждена следующим образом. Предположим, что скачкообразное изменение показателя преломления заменено произвольным плавным распределением, как показано на фиг. 3.7.1. Ось z координатной системы выбрана совпадающей с нормалью к границе. Показатель преломления постоянен в обеих областях и является непрерывной функцией $n(z)$ внутри переходной области. При таком выборе системы координат и распределения показателя преломления из формул (3.5.30) и (3.5.31) получаем

$$p_x = \text{const}, \quad (3.7.15)$$

$$p_y = \text{const} \quad (3.7.16)$$

во всем пространстве. Этот результат совершенно не зависит от вида функции $n(z)$. Уравнения (3.5.32) и (3.5.33)

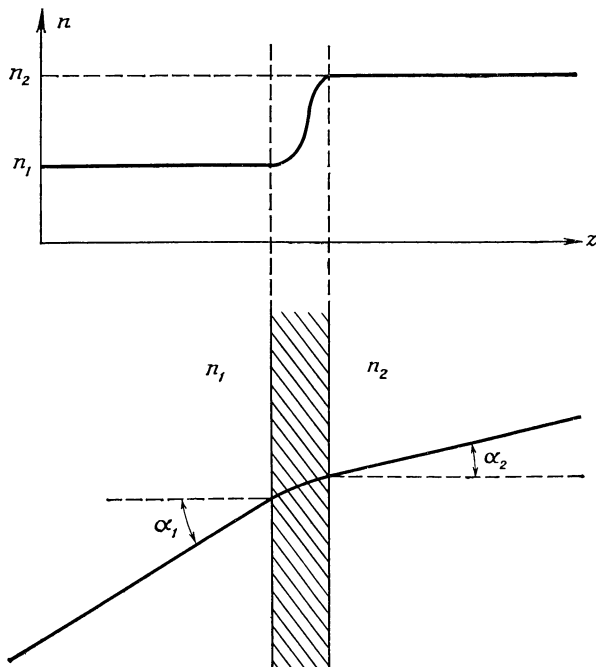
могут быть дополнены соотношением

$$p_z = n \frac{dz}{ds}. \quad (3.7.17)$$

Все три соотношения приводят к квадрату длины вектора «импульса»:

$$p_x^2 + p_y^2 + p_z^2 = n^2 \left[\left(\frac{dx}{ds} \right)^2 + \left(\frac{dy}{ds} \right)^2 + \left(\frac{dz}{ds} \right)^2 \right] = n^2. \quad (3.7.18)$$

Используя угол α_1 , определяющий наклон луча в среде 1, и угол α_2 для наклона луча в среде 2 (см. определение этих



Ф и г. 3.7.1. Замена резкого скачка показателя преломления на границе раздела диэлектриков плавным изменением.

углов на фиг. 3.7.1), из формулы (3.7.15) сразу же получаем закон Снеллиуса (считается, что система координат расположена так, что $p_y = 0$):

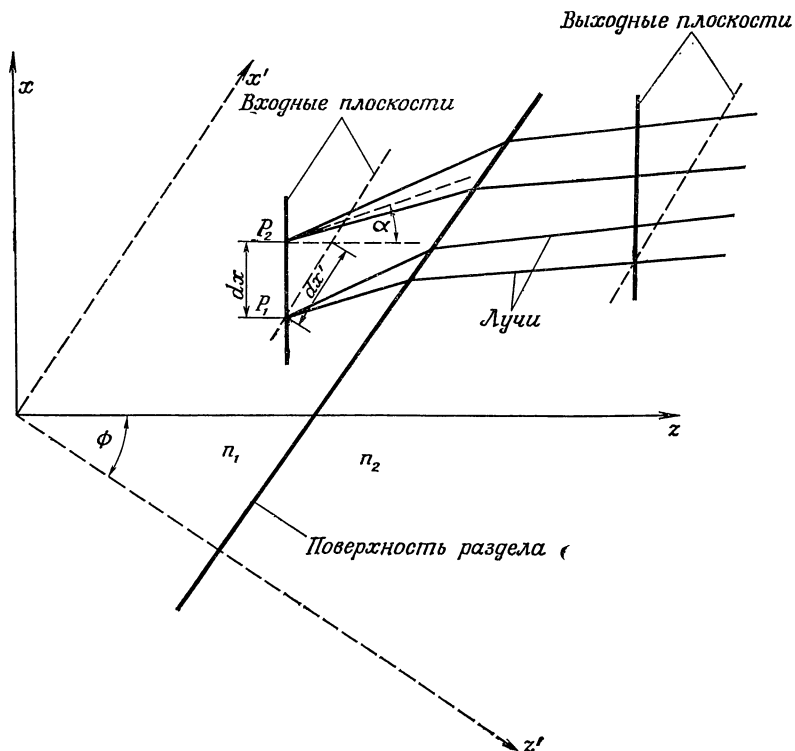
$$(p_x)_1 = n_1 \sin \alpha_1 = (p_x)_2 = n_2 \sin \alpha_2. \quad (3.7.19)$$

Эти рассуждения показывают, что закон Снеллиуса справедлив не только в случае резкого изменения показателя преломления на границе, но даже в случае произвольного его распределения и произвольной ширины переходной области. Поскольку показатель преломления здесь вообще не используется, можно принять его изменяющимся произвольным образом и получить, в частности, резкую границу в результате предельного перехода. Таким образом, любое резкое изменение показателя преломления можно считать сглаженным, так что уравнивания лучей и, следовательно, теорема Лиувилля являются справедливыми и в случае резкого изменения показателя преломления. Траектории лучей в среде со сглаженным распределением показателя преломления асимптотически идентичны траекториям в разрывной среде, если считать, что изменение показателя преломления становится все более и более резким.

Второе доказательство справедливости теоремы Лиувилля использует с самого начала резкую границу. Разумеется, справедливость (3.7.14) можно подтвердить прямым вычислением в любом конкретном случае. Однако в случае произвольной границы такое непосредственное вычисление весьма трудоемко. Второй подход показывает, как такое прямое вычисление может быть упрощено вращением системы координат.

Рассмотрим задачу лучевой оптики, схематически изображенную на фиг. 3.7.2. Плоская граница расположена под произвольным углом к координатным осям. Задача для простоты сведена к двумерной. Гораздо легче проводить лучи через границу, расположенную перпендикулярно оси z системы координат. Поэтому введем координаты, показанные на фиг. 3.7.2 пунктирными линиями. Однако поворот системы координат вызывает определенные сложности, напоминающие те, которые возникают в галилеевском варианте преобразования Лоренца [21]. Теорема Лиувилля принципиально основывается на том факте, что все координаты луча x , y , p_x и p_y берутся при одном и том же значении координаты z . Лучи в точках P_1 и P_2 , показанных на фиг. 3.7.2, удовлетворяют этому требованию в первоначальной системе координат x , y . Однако в повернутой системе каждая точка имеет другое значение z' . Чтобы

можно было использовать теорему Лиувилля, нужно не просто отнести положение луча к координатам x' и y' , а рассматривать их как новые координаты, смещенные



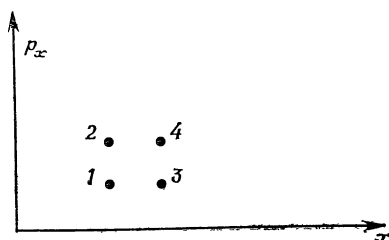
Ф и г. 3.7.2. Пояснение к доказательству теоремы Лиувилля.

Лучи пересекают границу раздела двух сред. Ось x' новой системы координат параллельна граничной поверхности.

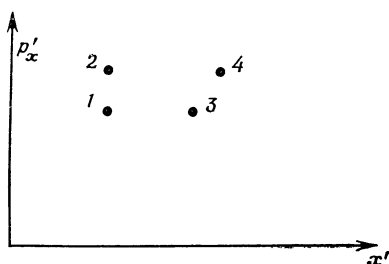
по лучу до точки его пересечения с плоскостью, перпендикулярной оси z' . Мы не имеем права просто геометрически поворачивать систему координат, а должны изменить координаты каждой точки так, чтобы быть уверенными, что каждый луч снова соотнесен с той же самой точкой z' . Эта задача аналогична задаче одновременности, встречающейся в теории относительности. Здесь обнаруживается

еще одна связь лучевой оптики с релятивистской механикой. Только в параксиальном (нерелятивистском) приближении, когда все углы (включая углы лучей с нормалью к границе) малы, простой поворот системы координат будет приблизительно правильным. Представление в фазовом пространстве четырех лучей в точках P_1 и P_2 , выраженное

Ф и г. 3.7.3. Фазовое пространство для лучей, проходящих через точки P_1 и P_2 на фиг. 3.7.2.



в старой системе координат, показано на фиг. 3.7.3. Точки 1 и 2 на фиг. 3.7.3 соответствуют точке P_1 на фиг. 3.7.2, а точки 3 и 4 соответствуют P_2 . В новой системе координат положения четырех лучей должны быть представлены



Ф и г. 3.7.4. Фазовое пространство для лучей, пересекающих пунктирную линию, которая проходит через точку P_1 на фиг. 3.7.2.

пересечениями пунктирной линии (проходящей через P_1) с лучами. Соответствующее представление в фазовом пространстве показано на фиг. 3.7.4. Все точки были сдвинуты вдоль p'_x , и первоначальный квадрат приобрел слегка трапециевидальную форму. Используя угол наклона α , средний между углами наклона лучей 3 и 4, можно выразить связь между приращением длины dx' в новой системе координат и приращением dx в старой системе в виде

$$dx = (\cos \phi - \sin \phi \operatorname{tg} \alpha) dx'. \quad (3.7.20)$$

Импульс в старой системе координат получается из соотношений (3.5.38) или (3.7.19):

$$p_x = n_1 \sin \alpha. \quad (3.7.21)$$

Его приращение дается формулой

$$dp_x = n_1 \cos \alpha d\alpha. \quad (3.7.22)$$

Импульс в новой системе координат равен

$$p'_x = n_1 \sin (\alpha + \phi), \quad (3.7.23)$$

а приращение

$$dp'_x = n_1 \cos (\alpha + \phi) d\alpha. \quad (3.7.24)$$

Теперь можно выразить приращение в старой системе координат через приращение в новой системе:

$$dp_x = \frac{\cos \alpha}{\cos (\alpha + \phi)} dp'_x. \quad (3.7.25)$$

Элементы объема в фазовом пространстве, выраженные в двухкоординатной системе, связаны соотношением

$$dx dp_x = \frac{\cos \alpha (\cos \phi - \operatorname{tg} \alpha \sin \phi)}{\cos (\alpha + \phi)} dx' dp'_x, \quad (3.7.26)$$

или, используя равенство $dV = dx dp_x$ и теорему дополнения тригонометрических функций, получаем

$$dV = dV'. \quad (3.7.27)$$

Таким образом, при преобразовании к новой системе координат объем в фазовом пространстве сохраняется. Это полезный результат, так как он показывает, что можно поворачивать систему координат до любого удобного положения, когда мы пытаемся подтвердить теорему Лиувилля для лучей, которые или отражаются зеркалом, или, как в этом случае, преломляются диэлектрической границей. Вывод об инвариантности объема фазового пространства не зависит от величины показателя преломления n , конечно, справедлив также и в случае, когда система координат возвращается обратно в первоначальное положение после прохождения лучей через границу. Этот результат справедлив также и для реального трехмерного случая. Выбор двумерной задачи мотивировался лишь удобством изложения.

Чтобы доказать теорему Лиувилля для лучей, проходящих через границу между двумя средами с различными показателями преломления, используем повернутую систему координат с осью z' , перпендикулярной границе в точке, где лучи ее пересекают. Граничное условие для лучей (3.3.5) с помощью формул (3.5.32) и (3.5.33) может быть выражено в виде

$$(\mathbf{r}_t)_1 = (\mathbf{r}_t)_2. \quad (3.7.28)$$

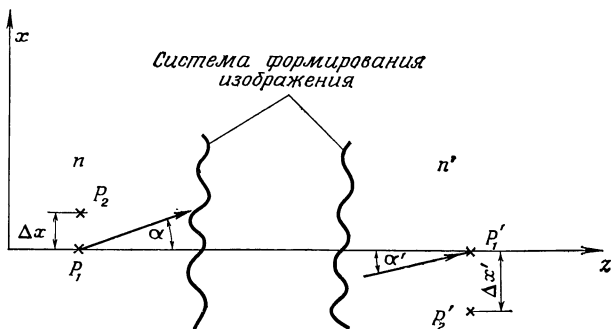
Тангенциальные составляющие вектора \mathbf{r} при пересечении границы остаются непрерывными. При выборе повернутой системы координат тангенциальными компонентами являются r'_x и r'_y . Теорема Лиувилля справедлива вдоль траектории луча в свободном пространстве, поэтому при доказательстве справедливости (3.7.14) можно использовать положения луча в непосредственной близости к границе. Пересечение границы не изменяет положения луча. Так как тангенциальные x' - и y' -компоненты \mathbf{r} также одинаковы на обеих сторонах границы, имеем

$$dV_1 = dV_2, \quad (3.7.29)$$

т. е. теорема Лиувилля справедлива для лучей на разрывах непрерывности показателя преломления. Поскольку систему координат можно поворачивать произвольным образом, доказательство того, что теорема Лиувилля справедлива на разрывной диэлектрической границе, становится тривиальным. Этот метод прямой проверки теоремы Лиувилля является наиболее подходящим для установления ее справедливости в случае отражения от произвольной поверхности. С помощью предельного перехода, рассмотренного ранее, мы убедились, что теорема должна быть справедливой на диэлектрических границах раздела. Однако случай отражения не может быть рассмотрен таким же способом, поэтому желательным является прямое доказательство. Это доказательство предоставляется читателю в качестве упражнения. Непосредственное вычисление становится совершенно простым, если использовать тот факт, что можно поворачивать систему координат и относить положение лучей к плоскости постоянных значений z' .

Теперь воспользуемся теоремой Лиувилля, чтобы доказать теорему синусов Аббе [1] для оптических изображе-

ний. Эта теорема не является существенной для содержания настоящей книги. Ее вывод включен только в качестве иллюстрации полезности теоремы Лиувилля. Условие синусов применяется для систем формирования изображения. Точнее, оно применяется к изображениям без аберраций. Пусть необходимо определить связь между углами лучей и расстояниями между соседними точками изображения и объекта. На фиг. 3.7.5 показана геометрия задачи.



Ф и г. 3.7.5. Схематическое изображение системы формирования оптического изображения. Показаны точки объекта P_1 и P_2 и точки изображения P'_1 и P'_2 , а также направления двух типичных лучей, используемых в теореме синусов Аббе.

Точка объекта P_1 расположена на оптической оси системы. Соответствующая точка изображения P'_1 также расположена на оптической оси. Фактически ими определяется оптическая ось. Вследствие нашего предположения о правильности оптического изображения каждый луч, выходящий из P_1 под произвольным углом, должен пройти через P'_1 . Разумеется, то же самое справедливо для каждого луча, выходящего из соседней точки объекта P_2 . Они все должны пройти через ее изображение P'_2 . Теорема Лиувилля в двумерной форме дает

$$\left| \begin{array}{cc} \frac{\partial x'}{\partial x} & \frac{\partial p'_x}{\partial x} \\ \frac{\partial x'}{\partial p_x} & \frac{\partial p'_x}{\partial p_x} \end{array} \right| = 1. \quad (3.7.30)$$

Предполагаем, что Δx и $\Delta x'$ являются расстояниями по вертикали между двумя соседними точками объекта и изображения. Для стигматичности (правильности) изображения необходимо, чтобы положение луча x' в плоскости изображения не зависело от угла наклона луча (его «импульса»). Это означает [с использованием производных в (3.7.30)], что

$$\frac{\partial x'}{\partial p_x} = 0. \quad (3.7.31)$$

Теорема Лиувилля (3.7.30) требует, чтобы выполнялось следующее соотношение между соседними точками объекта и изображения и «импульсами» лучей:

$$\Delta x \Delta p_x = \Delta x' \Delta p'_x. \quad (3.7.32)$$

Так как положение каждой пары точек объекта и изображения не зависит от угла, под которым луч выходит из точки объекта, можно проинтегрировать соотношение по p_x , что дает

$$\Delta x p_x = \Delta x' p'_x. \quad (3.7.33)$$

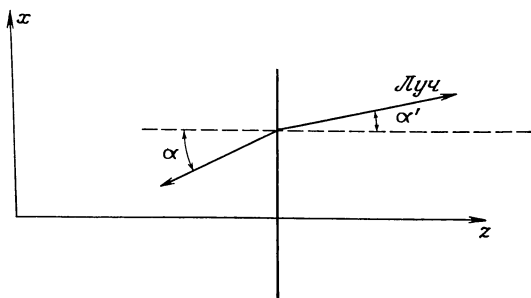
Наконец, используя выражение (3.5.38), получаем условие синусов Аббе

$$n \Delta x \sin \alpha = n' \Delta x' \sin \alpha'. \quad (3.7.34)$$

Это условие обычно выводится без использования теоремы Лиувилля [1]. Наш вывод показывает, что условие синусов является лишь частным случаем теоремы Лиувилля. Отношение $\Delta x' / \Delta x$ есть увеличение изображения. Теорема синусов устанавливает связь между увеличением изображения и углами выходящих и входящих лучей. Она основывается на предположении о стигматичности изображения.

Наконец, выведем другое соотношение, которое будет использовано при рассмотрении распространения света через оптические передающие системы довольно общего вида. Удобно использовать представление о тонком оптическом корректоре. Это прибор, имеющий (в идеале) нулевую толщину, но изменяющий угол наклона каждого луча, проходящего через него, в зависимости от положения луча. Другими словами, можно сказать, что тонкий оптический корректор производит в общем случае переменный

по фронту волны фазовый сдвиг. Схематически тонкий оптический корректор показан на фиг. 3.7.6. Тонким оптическим корректором может быть тонкая оптическая линза. То, что она показана расположенной перпендикулярно оси z , не является ограничением, так как мы убедились, что можем поворачивать систему координат, не нарушая справедливости теоремы Лиувилля. Снова достаточно применить теорему Лиувилля в ее двумерной форме (3.7.30).



Ф и г. 3.7.6. Тонкий оптический корректор. Луч, падающий под углом α , выходит из корректора под углом α' .

Мы выбираем две точки x и x' на траектории луча непосредственно перед и после корректора. Это означает, что $x = x'$ и что x' не зависит от p_x . Теорема Лиувилля (3.7.30) дает

$$\frac{\partial p'_x}{\partial p_x} = 1. \quad (3.7.35)$$

Умножим обе части равенства (3.7.35) на dp_x и проинтегрируем. Интегрирование дает произвольную постоянную, которая может быть функцией от x :

$$p'_x = p_x + F(x). \quad (3.7.36)$$

Через углы это уравнение выражается следующим образом:

$$\sin \alpha' = \sin \alpha + F(x). \quad (3.7.37)$$

Предполагается, что показатель преломления среды перед и после корректора имеет одно и то же значение. Функция $F(x)$ произвольна и дает значительную свободу в выборе тонких оптических корректоров. Однако изменение угла падения луча, которое может внести корректор, сильно

ограничено теоремой Лиувилля. Используя теорему дополнения синусоидальных функций, можно выразить (3.7.37) через разность углов $\alpha' - \alpha$:

$$\sin \frac{\alpha' - \alpha}{2} = \frac{F(x)}{2 \cos \frac{\alpha' + \alpha}{2}}. \quad (3.7.38)$$

В параксиальном приближении предполагается, что α и α' очень малы, поэтому можно использовать приближенную форму (3.7.38) для получения хорошо известной формулы тонкой линзы

$$\alpha' - \alpha = F(x). \quad (3.7.39)$$

В параксиальном приближении тонкий оптический корректор (тонкую линзу) можно рассматривать как прибор, который изменяет угол луча в зависимости от положения луча, но вне зависимости от угла падения. Однако изменение угла должно зависеть и от угла падения, если этот угол велик. Мы не имеем права предположить, что можем создать тонкий оптический прибор, который просто изменяет угол луча независимо от угла падения. Это важный и неожиданный результат. Работая с тонкими линзами и с углами падения, близкими к нормальным, мы привыкли к мысли о том, что тонкая линза лишь отклоняет каждый луч на фиксированную величину, зависящую только от точки, через которую луч проходит через линзу, но не зависящую от угла падения. Если это простое правило применить к тонким линзам, которые наклонены по отношению к оси или, возможно, даже искажены более общим образом, можно получить серьезные ошибки, если не принять во внимание следствия теоремы Лиувилля. Траектория луча, проходящего через тонкую линзу, описывается формулой (3.7.38). Единственная свобода, предоставленная экспериментатору, заключается в выборе функции $F(x)$. Следствия этих рассуждений будут рассмотрены в другой главе при исследовании прохождения световых лучей через системы неидеальных линз.

Применение теоремы Лиувилля к задаче сжатия пучка световых лучей сужающимися образцами диэлектрического материала и линзами можно найти в работе [102].

4.1. ВВЕДЕНИЕ

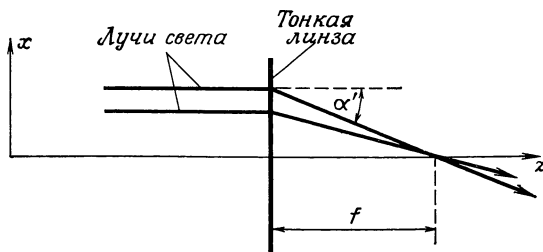
Нет более важных оптических элементов, чем линзы. Линза является основным элементом любой системы, формирующей изображение, в том числе и глаза животных. Линзы в искусственных оптических системах обычно изготавливаются из стекла. Назначение линз состоит в изменении пути световых лучей таким образом, чтобы все лучи, исходящие из некоторой точки объекта, снова собрались в соответствующей точке изображения. Не существует линз, которые выполняли бы эту операцию идеально во всех случаях. Для устранения недостатков, присущих одиночной линзе, используются сложные комбинированные системы, позволяющие достичь компромисса между противоречивыми требованиями.

Принципы конструирования одиночных и сложных линз изложены в литературе по оптическим приборам [1, 15]. По этой причине здесь не будут детально описываться способы коррекции линз. Ограничимся рассмотрением таких свойств линз, которые являются важными при передаче света в линзовых волноводах, и фурье-преобразующих свойств, существенных при оптической обработке информации и в оптических аналоговых вычислительных устройствах. Принцип работы этих устройств может быть объяснен с помощью понятия тонкой линзы. Линза называется тонкой, если траекторию света внутри ее можно не рассматривать в деталях. Тонкая линза работает как устройство, которое просто изменяет направление проходящих через него лучей. На языке волновой оптики тонкую линзу можно рассматривать как фазовый корректор, изменяющий фазу проходящей волны. Фазовый сдвиг зависит от того, на какую часть линзы падает свет.

Рассмотрим свойства тонких линз как с позиции геометрической оптики, так и с позиции волновой оптики.

4.2. ЛУЧЕВАЯ ОПТИКА ТОНКИХ ЛИНЗ

Определим тонкую линзу как устройство, которое отклоняет любой световой луч, падающий параллельно оптической оси, так, что он пересекает оптическую ось



Фиг. 4.2.1. Схема идеальной тонкой линзы. Все горизонтальные входные лучи проходят за линзой через фокальную точку.

на фиксированном расстоянии f после прохождения линзы. На фиг. 4.2.1 показаны тонкая собирающая линза и ее действие на два световых луча в двумерном случае. Тонкая линза должна подчиняться общей формуле (3.7.37). Для удовлетворения условия, чтобы каждый луч, идущий параллельно оптической оси, пересекал ее на расстоянии f от линзы, функция $F(x)$ должна иметь следующий вид (при $\alpha = 0$):

$$F(x) = \sin \alpha' = \frac{x}{\sqrt{x^2 + f^2}}. \quad (4.2.1)$$

Для линзы с осевой симметрией расстояние x заменяется расстоянием вдоль радиуса r , на котором луч проходит через линзу:

$$F(r) = \frac{r}{\sqrt{r^2 + f^2}}. \quad (4.2.2)$$

Общая формула тонкой линзы для луча, входящего в линзу под углом α по отношению к нормали к плоскости тонкой линзы и выходящего из нее под углом α' , вытекает

из формул (3.7.37) и (4.2.2):

$$\sin \alpha' = \sin \alpha + \frac{r}{\sqrt{r^2 + f^2}}. \quad (4.2.3)$$

В параксиальном приближении уравнение тонкой линзы принимает хорошо известную форму

$$\alpha' - \alpha = \frac{r}{f}. \quad (4.2.4)$$

Если фокусное расстояние f также является функцией r , то оказывается, что тонкая линза допускает искажения. Реальные физические линзы всегда вносят некоторые искажения. Даже если воспользоваться приближением тонкой линзы, то можно ожидать, что f есть функция не только r , но также зависит от частоты. До тех пор пока величина r остается малой, обычно можно пренебречь зависимостью f от r и применять параксиальное приближение с постоянным f . Мы воспользуемся параксиальной формулой (4.2.4) для описания способности линзы формировать изображение. Из формулы (4.2.3), впрочем как и из (4.2.4), видно, что луч, проходящий через линзу по оптической оси ($r = 0$), не меняет своего направления. То, что луч, падающий на линзу параллельно оптической оси, пересекает ось в фокальной точке f , а луч, проходящий через точку f , выходя из линзы, распространяется параллельно оптической оси и, наконец, что луч, который пересекает линзу при $r = 0$, не изменяет своего направления, оказывается достаточной информацией для нахождения изображения любого объекта, помещенного перед линзой. Обозначим расстояние от линзы до объекта через a , а расстояние от линзы до изображения — через b . На фиг. 4.2.2 приведено построение, позволяющее найти изображение; показаны три упомянутых луча, причем для получения изображения точки достаточно двух из них. Для луча, проходящего через фокус справа от линзы, можно записать соотношение

$$\frac{h_0}{f} = \frac{h_0 + h_i}{b}. \quad (4.2.5)$$

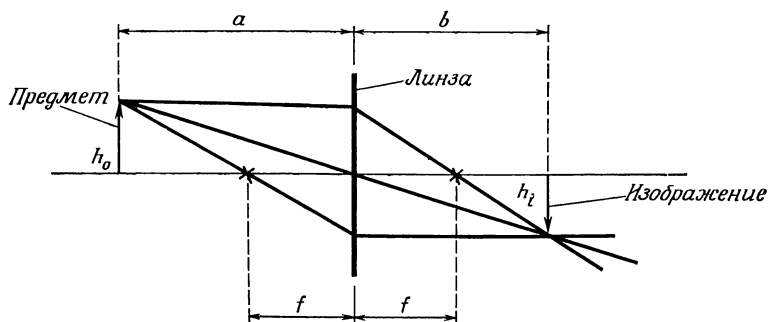
Аналогичное соотношение получается для луча, проходящего через левую фокальную точку:

$$\frac{h_i}{f} = \frac{h_0 + h_i}{a}. \quad (4.2.6)$$

Сложение обоих соотношений дает в результате важное уравнение, которое связывает расстояние до объекта a и расстояние до изображения b с фокусным расстоянием линзы:

$$\frac{1}{f} = \frac{1}{a} + \frac{1}{b}. \quad (4.2.7)$$

Из этого равенства видно, что для бесконечно удаленного объекта $b = f$. По мере приближения объекта к левому



Ф и г. 4.2.2. Формирование изображения с помощью идеальной тонкой линзы.

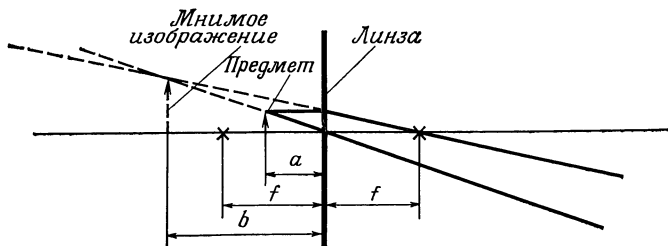
фокусу величина b возрастает. При $a = f$ имеем $b = \infty$. При $a < f$ величина b оказывается отрицательной. Для того чтобы не рассматривать отрицательное расстояние до изображения, будем считать, как это обычно делается, что отрицательная величина b соответствует так называемому мнимому изображению. Реальный объект и его мнимое изображение показаны на фиг. 4.2.3.

Наблюдатель, рассматривающий объект сквозь линзу, интерпретирует лучи, достигающие его справа от линзы, как приходящие от мнимого изображения (см. фиг. 4.2.3). Фиг. 4.2.2 поясняет принцип работы фотографической камеры. Если светочувствительную пленку поместить в положение b , то на ней зарегистрируется изображение реального объекта. Фиг. 4.2.3 объясняет действие увеличительного стекла. Наблюдатель держит линзу так, чтобы мнимое изображение формировалось на удобном для рассматривания расстоянии b . Объект должен быть помещен

между линзой и левым фокусом. Беря отношение (4.2.5) и (4.2.6), получаем для величины изображения соотношение

$$\frac{h_i}{h_0} = \frac{b}{a}. \quad (4.2.8)$$

В заключение настоящего раздела приведем краткое описание разрешающей способности оптических приборов,

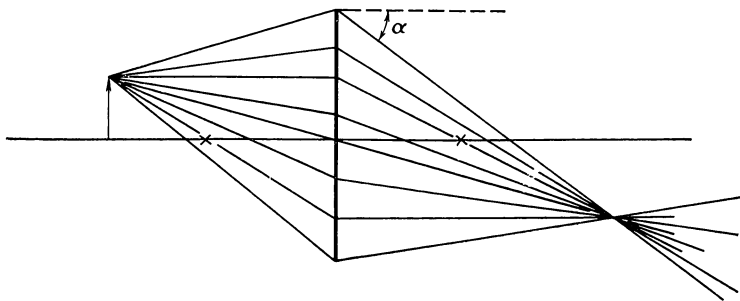


Ф и г. 4.2.3. Мнимое изображение объекта, расположенного между линзой и фокусом.

не вдаваясь глубоко в рассмотрение свойств оптических систем, связанных с формированием изображения. Точность, с которой можно получить изображение точки, зависит от предельной точности определения положения луча. Лучи в оптике служат лишь для приближенного описания оптических явлений. Точное описание должно быть дано в рамках волновой оптики. Поскольку волновая оптика определена ранее в разд. 3.6 как квантовая теория лучевой оптики, то мы уже знаем, что волновая функция есть величина вероятности нахождения луча в данной точке. Положение и «импульс» световых лучей представляют собой две величины, которые не могут быть определены с произвольной точностью. Для того чтобы определить положение светового луча с точностью Δx в поперечном направлении, следует допустить разброс «импульса» луча. Из соотношения неопределенности (3.6.56) получим ограничение на точность определения положения светового луча или, что то же самое, на разрешение изображения точки

$$\Delta x \geq \frac{\lambda}{4\pi\Delta p}. \quad (4.2.9)$$

Импульс луча связан с углом его наклона соотношением (3.5.38). Все лучи, выходящие из какой-либо точки объекта, сходятся в соответствующей точке изображения, как показано на фиг. 4.2.4. Неопределенность лучевого импульса должна быть связана с максимальным углом наклона



Фиг. 4.2.4. Все лучи, исходящие из точки объекта, сходятся в точке изображения за идеальной тонкой линзой.

луча, который допускается апертурой линзы. Принимая в качестве грубой оценки величины неопределенности лучевого импульса

$$\Delta p = n \sin \alpha, \quad (4.2.10)$$

получаем оценку для предельной разрешающей способности оптической системы

$$\Delta x \geq \frac{\lambda}{4\pi n \sin \alpha}. \quad (4.2.11)$$

Эта формула совпадает с общеизвестной предельной разрешающей способностью, полученной Рэлеем, по крайней мере по порядку величины [1]. Мы еще вернемся к оптической разрешающей способности в связи с описанием линз в рамках волновой оптики.

4.3. ВОЛНОВАЯ ОПТИКА ТОНКИХ ЛИНЗ

Для понимания свойств тонких линз в рамках волновой оптики не требуется использования теории Максвелла. Поляризация света оказывает влияние лишь на отражение

от передней и задней поверхностей линз. Если же игнорировать это явление, то для описания волновой оптики линз будет достаточно использовать лишь скалярное волновое уравнение. Мы ограничимся рассмотрением только тонких линз. Дополнительные трудности, обусловленные толщиной линз, лучше трактуются с помощью лучевой оптики. Такую трактовку можно найти в большинстве книг по оптике [1, 15].

Тонкая линза является фазовым корректором. Если рассмотреть волновое поле до и непосредственно после линзы, то обнаружится, что фаза волны изменяется, тогда как ее амплитуда в большинстве практических случаев остается без изменения. Обозначим входное поле непосредственно перед линзой через ψ_i , а выходное поле сразу за линзой через ψ_0 . Предполагая для простоты круговую симметрию поля, действие линзы можно описать уравнением

$$\psi_0(r) = e^{i\gamma(r)}\psi_i(r), \quad (4.3.1)$$

где r — расстояние от оптической оси. Потребуем, чтобы линза имела такие же свойства, что и линза, описываемая уравнением (4.2.4). Тот факт, что лучи представляют собой траектории, ортогональные волновым фронтам, позволяет перевести картину, представленную на фиг. 4.2.1, на язык волновой оптики. Результат показан на фиг. 4.3.1.

Пунктирными линиями изображены волновые фронты. Падающая плоская волна превращается в сферическую волну, сходящуюся в фокальной точке на расстоянии f позади линзы. Фаза волны непосредственно после прохождения волны через линзу определяется соотношением

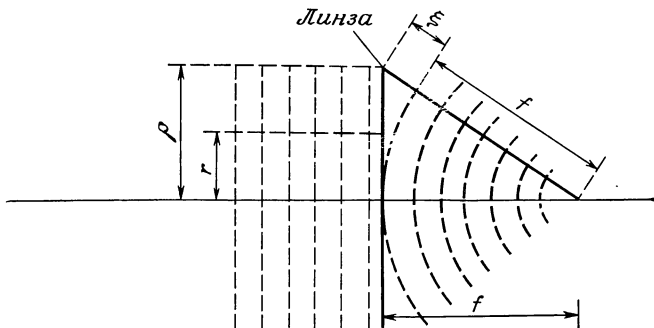
$$\gamma(r) = -(\rho^2 - r^2)K. \quad (4.3.2)$$

Знак минус связан с тем, что распространяющаяся волна зависит от времени по закону (2.2.1). Уравнение (4.3.2) можно обосновать следующим образом. При r , равном ρ , т. е. радиусу линзы, фазовый сдвиг равен нулю. Он возрастает при уменьшении r и достигает максимума при $r = 0$. Такое поведение становится очевидным, если вспомнить, что собирающие линзы толще в центральной своей части вблизи оптической оси и становятся все тоньше при удалении от центра по направлению к периферии. Более толстый участок в центре соответствует большей оптиче-

ской длине пути для светового поля, чем тонкий участок у края. Более длинный оптический путь означает больший фазовый сдвиг. Предположение о квадратичной зависимости является приближением, справедливым для линз, для которых удовлетворяется соотношение

$$\rho \ll f. \quad (4.3.3)$$

Квадратичная закономерность становится понятной, если учесть, что фазовый сдвиг должен быть одинаков для r и



Фиг. 4.3.1. Тонкая линза как фазовый корректор. Падающая плоская волна превращается за линзой в сферическую волну, сходящуюся в фокальной точке.

— r . Квадрат есть наинизшая степень, удовлетворяющая этому требованию. Наличие членов четвертого порядка указывало бы на то, что линза обладает аберрациями. Эти простые соображения позволяют определить фазовый сдвиг с точностью до неизвестного параметра K . Эта константа должна описывать оптическую силу линзы. Мы определим эту константу в параксиальном приближении. Поскольку каждый волновой фронт является поверхностью постоянной фазы, то, следовательно, волновой фронт на расстоянии ξ от края линзы (см. фиг. 4.3.1) имеет фазу $\gamma(0)$. Получаем соотношение

$$\gamma(0) = -\rho^2 K = -k\xi. \quad (4.3.4)$$

Фазовый сдвиг $k\xi$ соответствует прохождению волны на расстояние ξ от края линзы до волнового фронта, касаю-

щегося тонкой линзы в точке $r = 0$. Величина k есть постоянная распространения (4.3.13). Расстояние ξ можно получить из теоремы Пифагора, примененной к треугольнику, показанному на фиг. 4.3.1:

$$(f + \xi)^2 = f^2 + \rho^2. \quad (4.3.5)$$

Пренебрегая ξ^2 , приходим к решению

$$\xi = \frac{\rho^2}{2f}. \quad (4.3.6)$$

Константа K теперь получается из формул (4.3.4) и (4.3.6):

$$K = \frac{k}{2f}. \quad (4.3.7)$$

Фазовый сдвиг, обусловленный прохождением волны через тонкую линзу, получим окончательно из равенств (4.3.2) и (4.3.7):

$$\gamma(r) = -\frac{k}{2f}(\rho^2 - r^2). \quad (4.3.8)$$

Фазовый сдвиг позволяет рассчитать прохождение волн через тонкую линзу. Для простоты рассмотрим двумерный случай. Следуя обычной нашей практике, назовем задачу двумерной, если отсутствует зависимость от одной декартовой координаты. Предполагая, что нет зависимости от y , рассмотрим прохождение волнового поля от плоскости, находящейся на расстоянии a перед линзой, до плоскости на расстоянии b позади линзы. В нашей двумерной задаче мы имеем дело с тонкой линзой, эквивалентной цилиндрической линзе. Геометрия задачи показана на фиг. 4.3.2.

Прохождение входного поля ψ_i из сечения $z_i = \text{const}$ в положение непосредственно перед линзой описывается дифракционным интегралом (2.2.43)

$$\begin{aligned} \psi_1 = & \frac{e^{i\pi/4}}{\sqrt{\lambda(z' - z_i)}} e^{-ik(z' - z_i)} \times \\ & \times \int_{-\infty}^{\infty} \psi_i(x_i, z_i) \exp \left[-ik \frac{(x' - x_i)^2}{2(z' - z_i)} \right] dx_i. \end{aligned} \quad (4.3.9)$$

В соответствии с характером параксиального приближения мы заменили косинусы, стоящие в (2.2.43), на единицу.

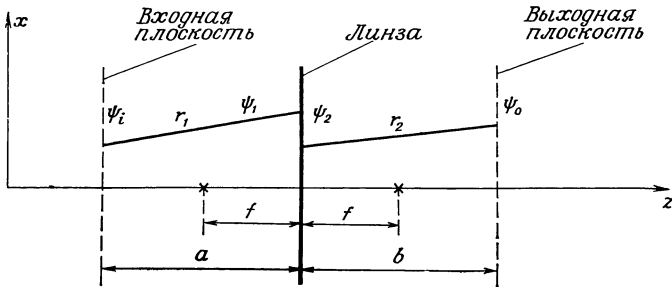
Это — поле, описывающее волну непосредственно слева от линзы. После прохождения линзы оно преобразуется следующим образом [см. (4.3.1) и (4.3.8)]:

$$\psi_2 = e^{-i(k/2f)(\rho^2 - x'^2)} \psi_1. \quad (4.3.10)$$

Функция ψ_2 описывает волну непосредственно за линзой. Поле в выходной плоскости имеет вид

$$\begin{aligned} \psi_0 = & \frac{e^{i\pi/4}}{\sqrt{\lambda(z_0 - z')}} e^{-ik(z_0 - z')} \int_{-\infty}^{\infty} \psi_2(x', z') \times \\ & \times \exp \left[-ik \frac{(x_0 - x')^2}{2(z_0 - z')} \right] dx'. \end{aligned} \quad (4.3.11)$$

После подстановки выражения (4.3.9) в (4.3.10) и использования полученного соотношения для исключения ψ_2



Ф и г. 4.3.2. К определению некоторых величин, используемых в волновой оптике для описания формирования изображения с помощью тонкой линзы.

из (4.3.11) получаем двойной интеграл. Интегрирование, не затрагивающее входное поле ψ_i , приводит к интегралу

$$\begin{aligned} I = & \int_{-\infty}^{\infty} \exp \left[-\frac{ik}{2} \left(\frac{1}{a} + \frac{1}{b} - \frac{1}{f} \right) x'^2 \right] \times \\ & \times \exp \left[ik \left(\frac{x_i}{a} + \frac{x_0}{b} \right) x' \right] dx'. \end{aligned} \quad (4.3.12)$$

Для упрощения записи в выражении (4.3.12) были использованы соотношения

$$z' - z_i = a \quad (4.3.13)$$

и

$$z_0 - z' = b; \quad (4.3.14)$$

имеем интеграл вида

$$I = \int_{-\infty}^{\infty} e^{i(\alpha x^2 + \beta x)} dx = \sqrt{\frac{i\pi}{\alpha}} e^{-i(\beta^2/4\alpha)}, \quad (4.3.15)$$

где

$$\alpha = -\frac{k}{2} \left(\frac{1}{a} + \frac{1}{b} - \frac{1}{f} \right) \quad (4.3.16)$$

и

$$\beta = k \left(\frac{x_i}{a} + \frac{x_0}{b} \right). \quad (4.3.17)$$

В результате получаем соотношение между входным полем и полем на выходе

$$\begin{aligned} \psi_0 = & \frac{\sqrt{\pi} e^{i(3/4)\pi}}{\lambda \sqrt{ab}} e^{-ik(z_0 - z_i)} e^{-i(k/2f)\rho^2} \frac{1}{\sqrt{\alpha}} \int_{-\infty}^{\infty} \psi_i(x_i, z_i) \times \\ & \times \exp \left\{ -\frac{ik}{2} \left[\frac{x_i^2}{a} + \frac{x_0^2}{b} + \frac{k}{2\alpha} \left(\frac{x_i}{a} + \frac{x_0}{b} \right)^2 \right] \right\} dx_i. \end{aligned} \quad (4.3.18)$$

В общем виде это соотношение оказывается довольно сложным. Имеются, однако, весьма интересные частные случаи, заслуживающие специального исследования. Начнем с рассмотрения случая

$$\alpha = 0. \quad (4.3.19)$$

Поскольку α стоит в знаменателе выражения (4.3.18), то очевидно, что мы сталкиваемся здесь с сингулярностью. Однако с данной сингулярностью легко справиться, если обратиться к (4.3.15). При $\alpha = 0$ член с x^2 исчезает из подынтегрального выражения, и мы имеем дело с представлением δ -функции:

$$I = \int_{-\infty}^{\infty} e^{i\beta x} dx = 2\pi\delta(\beta). \quad (4.3.20)$$

Используя формулы (4.3.15) и (4.3.20), получаем результат для предельного перехода

$$\lim_{\alpha \rightarrow 0} \sqrt{\frac{i\pi}{\alpha}} e^{-i(\beta^2/4\alpha)} = 2\pi\delta(\beta). \quad (4.3.21)$$

Таким образом, мы получили другое представление δ -функции. Этот результат дает возможность записать

$$\psi_0 = \frac{2\pi i}{\lambda \sqrt{ab}} e^{-ik(a+b)} e^{-i(k/2f)\rho^2} \int_{-\infty}^{\infty} \psi_i(x_i, z_i) \times \\ \times e^{-i(k/2)(x_i^2/a + x_0^2/b)} \delta\left[k\left(\frac{x_i}{a} + \frac{x_0}{b}\right)\right] dx_i. \quad (4.3.22)$$

Интеграл можно теперь вычислить с помощью δ -функции, входящей в явном виде в подынтегральное выражение:

$$\psi_0(x_0, z_0) = i \sqrt{\frac{a}{b}} e^{-ik(a+b)} e^{-i(k/2f)[\rho^2 + (a/b)x_0^2]} \times \\ \times \psi_i\left(-\frac{a}{b}x_0, z_i\right). \quad (4.3.23)$$

Для упрощения здесь использованы соотношение

$$k = \frac{2\pi}{\lambda} \quad (4.3.24)$$

и равенство

$$\frac{1}{a} + \frac{1}{b} = \frac{1}{f}, \quad (4.3.25)$$

вытекающее из (4.3.16) при $\alpha = 0$. Уравнение (4.3.23) представляет собой закон формирования изображения, выраженный в терминах волновой оптики. Условие $\alpha = 0$ привело к соотношению (4.3.25), которое связывает расстояние до объекта и расстояние до изображения с фокусным расстоянием линзы. Это соотношение уже появлялось ранее как уравнение (4.2.7) при лучевой трактовке формирования изображения линзами. Формула (4.3.23) выражает тот факт, что поле изображения и поле объекта идентичны, за исключением противоположного знака в координатах изображения, что выражается знаком минус у аргумента x_0 . Кроме того, появляется скалярный множитель a/b , отражающий тот факт, что изображение отличается по размеру от объекта. Коэффициент увеличения изображения b/a уже встречался нам в лучевой трактовке формирования изображения (4.2.8). Фазовые множители в (4.3.23) не влияют на способность линзы формировать изображение. Как фотографическая пленка, так и сетчатка глаза нечувствительны к фазе светового поля. Они реагируют лишь на $|\psi|^2$, а не на ψ . Множитель $\sqrt{a/b}$ показы-

вает связь световой интенсивности изображения с интенсивностью объекта (если изображение увеличено, $b/a > 1$).

Читатель, несомненно, обратил внимание на то, что размеры линзы здесь предполагаются настолько большими, что ее истинными границами можно пренебречь и распространить интегрирование от $-\infty$ до $+\infty$. Очевидно, что это приближение оправдано, если размер светового пятна мал по сравнению с размером линзы. Однако пренебрегаемый нами конечный размер линзы в случае его учета приводит к дифракционным эффектам, которые ограничивают достижимое разрешение изображения (см. разд. 4.6).

Вернемся к выражению (4.3.18) и продолжим обсуждение волновой оптики линз, рассмотрев еще один частный случай. Предположим, что

$$a = b = f. \quad (4.3.26)$$

Иными словами, поле объекта и поле изображения расположены в передней и задней фокальных плоскостях линзы. Подстановка выражений (4.3.16) и (4.3.26) в (4.3.18) приводит к следующему соотношению:

$$\psi_0(x_0, z_0) = \frac{e^{i\pi/4}}{\sqrt{\lambda f}} e^{-2ikf} e^{-i(k/2f)\rho^2} \int_{-\infty}^{\infty} \psi_i(x_i, z_i) \times \\ \times e^{i(k/f)x_0 x_i} dx_i. \quad (4.3.27)$$

Исключая из рассмотрения несущественный фазовый множитель, мы видим, что выходное поле в задней фокальной плоскости линзы представляет собой преобразование Фурье входного поля, расположенного в передней фокальной плоскости. Линза, следовательно, способна осуществлять преобразование Фурье. Этот результат является очень важным при оптической обработке данных и в случае пространственной фильтрации. Поскольку преобразование Фурье какого-либо входного поля, расположенного в передней фокальной плоскости, осуществляется в задней фокальной плоскости, то оказывается возможным воздействовать на полученный «спектр» с помощью пространственной фильтрации, фазового сдвига или перемножения с другими функциями для того, чтобы придать преобразо-

ванному полю любой желаемый вид. Можно записать преобразование Фурье в более привычном виде, если ввести новые координаты

$$u = \sqrt{\frac{k}{f}} x_i = \sqrt{\frac{2\pi}{\lambda f}} x_i \quad (4.3.28)$$

и

$$v = \sqrt{\frac{k}{f}} x_0 \quad (4.3.29)$$

и, кроме того, функции

$$\psi_i(x_i, z_i) = f(u) \quad (4.3.30)$$

и

$$\psi_0(x_0, z_0) = F(v). \quad (4.3.31)$$

Это позволяет записать выражение (4.3.27) в виде

$$F(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u) e^{iuv} du. \quad (4.3.32)$$

Фазовый множитель опущен, поскольку его всегда можно устранить смещением входной или выходной плоскости на расстояние, не превышающее одну длину световой волны. Такая незначительная «дефокусировка» не влияет на качество оптических изображений, но позволяет подправить фазу таким образом, чтобы сделать ее в точности кратной 2π . Единственное ограничение для идеальной реализации преобразования Фурье с помощью тонких линз состоит в наличии конечной апертуры линзы. В приведенном здесь выводе соотношения (4.3.32) не учитывался размер апертуры линзы. Интеграл (4.3.12), строго говоря, должен быть взят в конечных пределах. Однако до тех пор пока размер светового пятна, падающего на линзовую апертуру, мал, приближение (4.3.32) справедливо.

При прохождении света от передней в заднюю фокальную плоскость достигается почти идеальное преобразование Фурье. Однако мы получим формулу преобразования Фурье также путем рассмотрения входного поля непосредственно перед линзой. Формулу (4.3.18) нельзя использовать в этом случае, так как она предполагает выполнимость параксиальных приближений, которые здесь

не выполняются. Однако результат для этого случая вытекает из формул (4.3.10) и (4.3.11):

$$\psi_0 = \frac{e^{i\pi/4}}{\sqrt{\lambda b}} e^{-ikb} e^{-(i/2)k(\rho^2/f + x_0^2/b)} \times \\ \times \int_{-\infty}^{\infty} \psi_i(x', z') e^{(i/2)k[(1/f - 1/b)x'^2 + (2/b)x_0 x']} dx'. \quad (4.3.33)$$

Располагая выходное поле вновь в задней фокальной плоскости

$$b = f \quad (4.3.34)$$

и используя формулы (4.3.28) — (4.3.31), после замены x_i на x' получим следующий результат:

$$F(v) = \frac{1}{\sqrt{2\pi}} e^{-ik(x_0^2/2f + z_0)} \int_{-\infty}^{\infty} f(u) e^{iuv} du. \quad (4.3.35)$$

Не интересующие нас постоянные фазы снова опущены по тем же соображениям, которые были приведены ранее. Преобразование Фурье (4.3.35) отличается наличием фазового множителя, который не является постоянным, а зависит от выходной координаты x_0 . Это отличие, однако, можно исключить, если рассматривать выходное поле не в плоскости $z_0 = f$, а на криволинейной поверхности

$$\frac{x_0^2}{2f} + z_0 = f + z'.$$

Анализ выходного поля на этой параболической поверхности вместо плоскости позволяет рассматривать результат прохождения поля через линзу как преобразование Фурье.

До сих пор мы ограничивались обсуждением двумерного случая. Нетрудно перейти к более общему случаю. При рассмотрении трехмерного поля нужно исходить из выражения (2.2.31) вместо (2.2.43). Различие между этими двумя выражениями незначительно и может быть легко учтено. Наиболее существенное отличие состоит в том, что под знаком интеграла стоит произведение фазовых множителей, зависящих от x и y . Это приводит к тому, что вместо интеграла (4.3.12) появляется произведение этого интеграла

на другой, полученный путем замены всех x -составляющих на соответствующие y -составляющие. Принимая это во внимание и рассматривая выражения (2.2.31) и (2.2.43), можно получить результат, аналогичный двумерному уравнению (4.3.18):

$$\begin{aligned} \psi_0 = & \frac{\pi e^{i(3/2)\pi}}{\lambda^2 ab} e^{-ik(z_0 - z_i)} e^{-i(k/2f)\rho^2} \times \\ & \times \frac{1}{\alpha} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left\{ \psi_i(x_i, y_i, z_i) \times \right. \\ & \times \exp \left[-\frac{i}{2} k \left(\frac{x_i^2 + y_i^2}{a} + \frac{x_0^2 + y_0^2}{b} + \frac{k}{2\alpha} \left(\frac{x_i}{a} + \frac{x_0}{b} \right)^2 + \right. \right. \\ & \left. \left. + \frac{k}{2\alpha} \left(\frac{y_i}{a} + \frac{y_0}{b} \right)^2 \right) \right] \right\} dx_i dy_i. \quad (4.3.36) \end{aligned}$$

Выводы, которые можно получить из этого выражения, почти идентичны тем, которые были сделаны для двумерного случая. Если расстояния до объекта и до изображения выбраны так, чтобы $\alpha = 0$, то получается перевернутое увеличенное изображение входного поля. Если входная плоскость находится при $a = f$, а выходная при $b = f$, то вновь получаем, что выходное поле представляет собой преобразование Фурье (теперь двумерное) входного поля. Используя формулу (4.3.26) и обозначения

$$u_x = \sqrt{\frac{k}{f}} x_i, \quad (4.3.37)$$

$$u_y = \sqrt{\frac{k}{f}} y_i, \quad (4.3.38)$$

$$v_x = \sqrt{\frac{k}{f}} x_0, \quad (4.3.39)$$

$$v_y = \sqrt{\frac{k}{f}} y_0, \quad (4.3.40)$$

находим из (4.3.36)

$$\begin{aligned} F(v_x, v_y) = & \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u_x, u_y) \times \\ & \times e^{i(u_x v_x + u_y v_y)} du_x du_y. \quad (4.3.41) \end{aligned}$$

Мы воспользовались формулами (4.3.30) и (4.3.31) и очевидным здесь обобщением на две переменные. Постоянный фазовый множитель опущен по изложенным ранее соображениям.

Сравнение выражений для преобразования Фурье между входным и выходным полями (входным полем является поле в передней фокальной плоскости или в плоскости непосредственно перед линзой), трансформируемыми линзой, с уравнениями (2.2.31) и (2.2.43) показывает, что поле в задней фокальной плоскости линзы соответствует дальнему дифракционному полю в области Фраунгофера. Напомним, что область Фраунгофера определяется тем, что квадратичные члены в аргументе экспоненциальных функций пренебрежимо малы. Дальнее поле у оси системы также пропорционально преобразованию Фурье входного поля без применения линзы.

4.4. ОПТИЧЕСКОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ И ПРОСТРАНСТВЕННАЯ ФИЛЬТРАЦИЯ

То обстоятельство, что линзы можно применять для получения преобразования Фурье двумерных распределений света, является очень важным для оптической обработки данных. Из выражений (4.3.32) и (4.3.41) видно, что преобразование Фурье объекта, расположенного в передней фокальной плоскости линзы, появляется в ее задней фокальной плоскости. Преобразование Фурье распределения входного поля, таким образом, легкодоступно и может быть использовано для изменения распределения поля известными методами. Оптический аналог частотной фильтрации электрических сигналов может быть осуществлен без сложных электрических схем просто путем размещения апертур на пути световой волны. Более сложная техника пространственной фильтрации [7] включает аттенюаторы или фазовращатели в задней фокальной плоскости линзы для изменения фурье-спектра входного распределения поля.

Допустим, что мы хотим получить преобразование Фурье данной функции. Начнем с приготовления оптического транспаранта, у которого амплитуда коэффициента прохождения представляет желаемую функцию. В прин-

цпе возможно рассмотреть транспаранты с заданной амплитудной и фазовой характеристиками. Однако для контроля фазы транспаранта требуется чрезвычайно точный контроль оптической длины пути в транспаранте в пределах малой доли длины волны. Такая точность обычно не может быть реализована. Однако воспроизведение оптического поля с желаемыми амплитудой и фазой возможно путем использования методов голографии [7].

Ограничимся рассмотрением простого двумерного случая. Предположим, что мы приготовили транспарант с амплитудной передаточной функцией $t(x)$. Освещение этого транспаранта в передней фокальной плоскости цилиндрической линзы плоской световой волной сразу же приводит к следующему распределению поля справа от транспаранта:

$$\psi_i(x_i) = At(x_i) = f(u). \quad (4.4.1)$$

Связь между x_i и u дается формулой (4.3.28). Константа A есть амплитуда падающей плоской волны. После прохождения транспаранта амплитуда этой плоской волны модулируется амплитудной передаточной функцией $t(x)$ транспаранта. Согласно выражению (4.3.32), находим преобразование Фурье входного поля (4.4.1) в задней фокальной плоскости цилиндрической линзы

$$\psi_0(x_0) = F(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u) e^{iuv} du. \quad (4.4.2)$$

Специалисты в области радиотехники привыкли к фурье-преобразованиям временных функций. Преобразование Фурье временной функции есть функция частоты. Здесь же временная координата заменена пространственной координатой

$$u = \sqrt{\frac{2\pi}{\lambda f}} x_i; \quad (4.4.3)$$

круговая частота ω заменена другой пространственной координатой

$$v = \sqrt{\frac{2\pi}{\lambda f}} x_0. \quad (4.4.4)$$

Координата x_0 в выходной (задней фокальной) плоскости цилиндрической линзы пропорциональна частотному интервалу обычного преобразования Фурье временной функции.

Это простое соотношение между функцией и ее преобразованием Фурье идеально подходит для осуществления операции фильтрации преобразования Фурье в частотной области. Поскольку частотная переменная заменена здесь пространственной переменной, подобная операция фильтрации известна как пространственная фильтрация.

Допустим, что мы хотим моделировать действие фильтра нижних частот. В случае электрического сигнала с временной зависимостью нужно передать напряжение или ток через соответствующую электрическую цепь, которая позволяет пропустить все частоты вплоть до граничной частоты ω_c , тогда как все частоты выше ω_c задерживаются. Соответствующий пространственный фильтр представляет собой просто щель в светонепроницаемом экране. Щель простирается от $x_0 = -x_{0c}$ до $x_0 = x_{0c}$, где

$$x_{0c} = \sqrt{\frac{\lambda f}{2\pi}} \omega_c. \quad (4.4.5)$$

Свет, достигающий задней фокальной плоскости при $|x_0| > |x_{0c}|$, встречает непрозрачный экран. Можно поставить вторую линзу для осуществления обратного преобразования Фурье и воспроизведения функции

$$\psi_t(x_t) = G(w) = \frac{1}{\sqrt{2\pi}} \int_{-\omega_c}^{\omega_c} \psi_0(x_0) e^{i w w} dv, \quad (4.4.6)$$

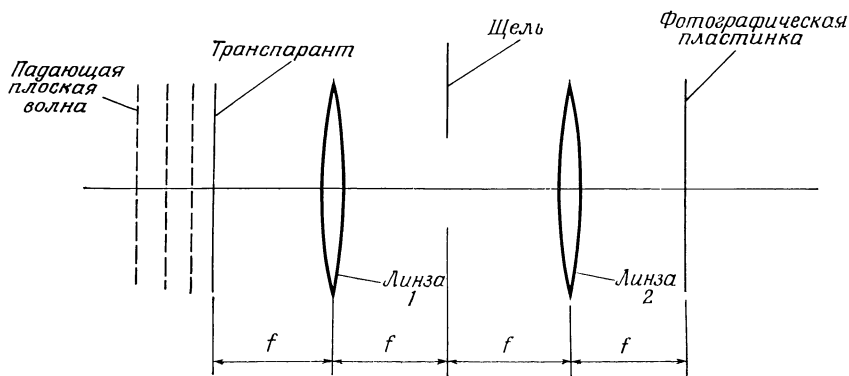
где

$$w = \sqrt{\frac{k}{f}} x_t. \quad (4.4.7)$$

Точная математическая форма преобразования Фурье получается при замене w (и соответственно x_t) на $-w$. Эта замена приводит к появлению отрицательного знака в показателе экспоненциальной функции. Изменение знака, необходимое для получения привычного преобразования Фурье и соответствующего ему обращения, показывает, что пространственное изображение, получаемое при использовании двух линз (в отсутствии фильтра нижних

частот), оказывается перевернутым относительно входного поля. Описанное действие пространственной фильтрации показано на фиг. 4.4.1.

Простое оптическое устройство, изображенное на фиг. 4.4.1, является аналоговым вычислительным устройством.



Ф и г. 4.4.1. Оптическое аналоговое вычислительное устройство, действующее как фильтр нижних частот. Если экран с щелью заменить транспарантом, то можно осуществить более общие операции фильтрации.

Оно позволяет получить результат фильтрации нижних частот по функции $t(x)$. Результирующая функция записывается на фотографической пластинке, помещенной в задней фокальной плоскости второй линзы. Однако необходимо помнить, что фотографическая пластинка чувствительна к энергии поля, а не к его амплитуде. В случае использования пластинки в ее линейной области регистрируется квадрат абсолютной величины амплитуды распределения поля, но не распределение самого поля. Результат аналогового вычисления, выполняемого оптической системой, получается, если взять квадратный корень от оптической плотности, записанной на пленке. Оптическая плотность входного транспаранта представляет квадрат функции $t(x)$: коэффициент прохождения $t(x)$ представляет амплитуду, а не мощность.

Не следует забывать, что вследствие конечного размера линз преобразование Фурье, осуществляемое линзами, является не точным, а лишь приближенным. Интеграл в (4.4.2), строго говоря, должен распространяться не на бесконечную область, а только на область, занятую линзой. Приближение, однако, может быть довольно хорошим, если использовать достаточно большие линзы, чтобы иметь гарантию, что большая часть света проходит через их апертуру.

Если поместить фотографическую пластинку в задней фокальной плоскости первой линзы, то будет зарегистрировано преобразование Фурье (точнее, его квадрат) входной функции $t(x)$. В такой конфигурации оптическое аналоговое вычислительное устройство представляет собой анализатор Фурье.

Пространственная фильтрация, конечно, не ограничивается только нижними частотами. Так же легко можно осуществить фильтрацию верхних частот, если установить экран, который не пропускает света в области $|x_0| < |x_{0c}|$, но пропускает его при $|x_0| > |x_{0c}|$. При размещении на пути светового пучка в задней фокальной плоскости первой линзы вместо светонепроницаемых апертур оптических транспарантов осуществляются более сложные операции фильтрации. Наиболее интересным примером проведения вычислений такого рода является получение свертки двух функций.

Пусть нужно получить свертку функций $t(x)$ и $g(x)$. Этого можно достичь (по крайней мере в принципе) при использовании транспаранта, амплитудный коэффициент прохождения которого является преобразованием Фурье от $g(x)$. Изготовление такого транспаранта представляет в общем случае весьма трудную задачу, и поэтому только очень простые функции могут быть представлены таким способом. Трудности возникают из-за того, что невозможно точно проконтролировать фазу транспаранта вследствие хаотических изменений толщины пленки и показателя преломления материала. Проблема хаотического изменения фазы может быть упрощена, если поместить транспарант внутрь контейнера, стенки которого выполнены плоскими с оптической точностью. Оптическую длину пути через транспарант можно до некоторой степени регу-

лизовать, заполняя стеклянный сосуд, содержащий транспарант, маслом со специально подобранным для согласования показателем преломления. Однако из-за непостоянства показателя преломления транспаранта невозможно полностью выравнять хаотические изменения оптической длины пути через транспарант, так что описанный метод выравнивания длины пути не решает нашу проблему. Преобразование Фурье G функции $g(x)$ характеризуется как фазовой, так и амплитудной функцией. Невозможно приготовить транспарант с фазовым сдвигом, который изменялся бы заданным образом. Эта задача может быть решена с помощью голографической техники, которая позволяет сохранять на пленке или пластинке информацию как об амплитуде, так и о фазе. Описание пространственных фильтров, изготовленных с помощью техники голографии, можно найти в книге Гудмана [7].

Здесь мы рассмотрим транспарант с амплитудной передаточной характеристикой

$$G(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g\left(\sqrt{\frac{f}{k}} u\right) e^{iuv} du. \quad (4.4.8)$$

Размещение этого транспаранта в задней фокальной плоскости первой линзы, показанной на фиг. 4.4.1, приводит к тому, что распределение поля $F(v)$, являющегося преобразованием Фурье от $t(x)$, согласно (4.4.1) и (4.4.2) умножается на амплитудную передаточную функцию $G(v)$. Распределение поля непосредственно справа от транспаранта в задней фокальной плоскости первой линзы будет теперь иметь вид $F(v) G(v)$.

Вторая линза осуществляет второе преобразование Фурье от этой функции, так что поле в задней фокальной плоскости второй линзы равно

$$H(w) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(v) G(v) e^{i vw} dv \quad (4.4.9)$$

или после подстановки выражения (4.4.8)

$$H(w) = \frac{1}{2\pi} \int_{-\infty}^{\infty} du g\left(\sqrt{\frac{f}{k}} u\right) \int_{-\infty}^{\infty} F(v) e^{i(u+w)v} dv. \quad (4.4.10)$$

Обратное преобразование Фурье $F(v)$ определяется формулой

$$f(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(v) e^{-iuv} dv. \quad (4.4.11)$$

Используя этот результат, можно переписать выражение (4.4.10) в следующем виде:

$$H(w) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g\left(\sqrt{\frac{f}{k}} u\right) f(-u-w) du. \quad (4.4.12)$$

Возвращаясь к исходной пространственной координате x_i (4.4.3) и вводя соответствующее выражение для w

$$w = \sqrt{\frac{k}{f}} x_t, \quad (4.4.13)$$

запишем выражение (4.4.12) как

$$H\left(\sqrt{\frac{k}{f}} x_t\right) = \frac{A}{\sqrt{\lambda f}} \int_{-\infty}^{\infty} g(x_i) t(-x_i - x_t) dx_i. \quad (4.4.14)$$

Для того чтобы выразить f через исходную функцию t , была использована формула (4.4.1). Наконец, вводя функцию

$$\varphi(-x_t) = H\left(\sqrt{\frac{k}{f}} x_t\right), \quad (4.4.15)$$

получаем для распределения поля в задней фокальной плоскости второй линзы выражение

$$\varphi(x_t) = \frac{A}{\sqrt{\lambda f}} \int_{-\infty}^{\infty} g(x_i) t(x_t - x_i) dx_i. \quad (4.4.16)$$

Заметим, что определение (4.4.15) функции $\varphi(x)$ включает обратное преобразование Фурье координатного направления в выходной плоскости.

Интеграл в выражении (4.4.16) известен как свертка функций g и t . Отфильтрованное распределение поля $t(x)$ с фурье-преобразованием функции $g(x)$ получается при вычислении свертки двух функций с помощью аналого-

вого вычислительного устройства, схематически изображенного на фиг. 4.4.1 (цель, показанная на фиг. 4.4.1, заменена транспарантом с амплитудной передаточной функцией G). Фотографическая пластинка в выходной плоскости регистрирует скорее квадрат модуля функции $\varphi(x)$, нежели саму функцию.

Трехмерный вариант двумерного оптического устройства, рассматривавшегося до сих пор, приводит к двумерному преобразованию Фурье (4.3.44). Пространственная фильтрация с использованием обычных линз вместо цилиндрических может быть осуществлена по аналогии с двумерным случаем. Применение обычных линз и транспаранта, являющегося двумерным преобразованием Фурье функций двух переменных, ведет к двумерной свертке функций двух переменных. Пусть входной транспарант определяется функцией $t(x, y)$, а пространственный фильтр — преобразованием Фурье функции $g(x, y)$

$$(v_x, v_y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g\left(\sqrt{\frac{f}{k}} u_x, \sqrt{\frac{f}{k}} u_y\right) \times \\ \times e^{i(u_x v_x + u_y v_y)} du_x du_y. \quad (4.4.17)$$

Распределение выходного поля в задней фокальной плоскости второй (обычной) линзы будет

$$\varphi(x_t, y_t) = \frac{A}{\lambda f} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x_i, y_i) t(x_t - x_i, y_t - y_i) dx_i dy_i. \quad (4.4.18)$$

Свертка двух функций или двух одинаковых функций (автокорреляция) является весьма ценной операцией. Так, например, процессом автокорреляции можно выделить замаскированный шумом сигнал.

Для рассмотренного здесь оптического аналогового вычислительного устройства требуется когерентный свет. Когерентным с большой степенью является свет лазера, тогда как свет, исходящий из теплового источника, не является когерентным. Свет называется когерентным, если амплитуды поля в разных точках пространства меняются при вполне определенном фазовом соотношении между ними. Отсутствие определенного соотношения фаз характеризует некогерентный свет. Если имеет место частичная кор-

реляция между фазами поля в разных точках пространства, то говорят о частично когерентном свете. В дополнение к когерентности необходимо также потребовать монохроматичность света (т. е. наличие лишь одной частоты или очень узкого спектра частот), если он используется для оптического вычисления функций, рассмотренного в данном разделе. Свет лазеров идеально соответствует этим целям. Необходимо, однако, обеспечить работу лазера на одной из многих возможных частот и генерирование одной определенной поперечной моды [см. обсуждение формулы (6.6.26)] для уверенности в том, что его выход достаточно монохроматичен и пространственно когерентен. Обычный свет теплового источника или газового разряда можно сделать монохроматичным и когерентным путем фильтрации узкой полосы частот и пропускания ее через маленькое отверстие. Однако при этом теряется столь много световой энергии, что остающегося частично когерентного света обычно оказывается недостаточно для проведения экспериментов по пространственной фильтрации.

4.5. ГАЗОВЫЕ ЛИНЗЫ

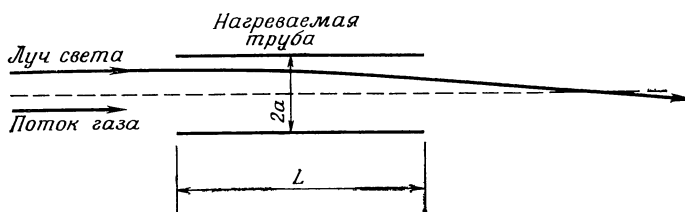
Стеклянные линзы идеально приспособлены для формирования изображения в оптических приборах. Если простая линза, сделанная из стекла лишь одного сорта, имеет значительное число нежелательных аберраций, то можно изготовить хорошие линзы путем комбинирования линз, выполненных из разных стекол [1, 15]. Однако, помимо формирования изображения, имеются иные приложения для линз, и это требует наличия у них разнообразных свойств. Вопросы, связанные с передачей света на большие расстояния, детально будут рассмотрены в гл. 5. Весьма перспективная световодная линия передачи состоит из последовательности линз, используемых для нейтрализации дифракционного расширения пучка света. Если можно пренебречь потерями, обусловленными рассеянием света внутри стекла и на поверхности линзы, то такая линзовая линия может быть сделана с экстремально низкими потерями. Действительно, можно сконструировать линзовую линию с произвольно малыми потерями, если

взять достаточно большие линзы и сделать так, чтобы пучок света не отклонялся значительно от оси линии. Однако линзы не идеальны, т. е. нельзя игнорировать потери из-за рассеяния внутри их и на поверхности. Даже в том случае, когда линза имеет неотражающие покрытия, потери на каждой поверхности линзы трудно сделать менее 1%. Бóльшая часть этих потерь обусловлена остаточным отражением, а также хаотическим рассеянием от пыли и других неоднородностей на поверхности линзы. Таким образом, основная доля потерь происходит на границе воздух — линза. Если бы удалось осуществить линзы без заметных изменений показателя преломления на поверхностях, то проблема потерь на рассеяние и отражение была бы устранена.

Такие линзы с малыми потерями в действительности осуществимы. Впервые подобную линзу предложил Берреман [27], который использовал подогреваемую металлическую спираль, помещенную внутри охлаждаемой металлической трубы, чтобы согреть газ внутри трубы и обеспечить его конвективное циркулирование. Он показал, что изменение плотности неравномерно нагретого газа вызывает такое изменение показателя преломления, которое приводит к фокусированию светового пучка, проходящего вблизи оси трубы. Линза Берремана достаточно эффективна, но она очень трудно поддается теоретическому анализу. Поэтому рассмотрим другой тип газовых линз, которые при эквивалентной эффективности можно подвергнуть анализу, с тем чтобы предсказать их свойства.

Поскольку имеется возможность сконструировать линзы, действие которых основано на использовании изменений показателя преломления газа, то можно исключить из рассмотрения проблему потерь, столь существенную для стеклянных линз. В газе с переменным показателем преломления, обусловленным неравномерным подогревом, между кусочно-однородными средами нет граничных поверхностей. Поэтому здесь полностью исключены потери на отражение и на поверхностное рассеяние. Однако необходимо предостеречь читателя от мысли, будто газовые линзы идеальны во всех отношениях. Газовые линзы имеют целый ряд своих собственных недостатков. Обычно они обладают довольно сильными аберрациями, т. е. их фокус-

ное расстояние зависит от положения светового пучка относительно центра линзы [28]. Нельзя выполнить газовые линзы с большими апертурами, так как газовый поток в широкой трубе имеет тенденцию становиться турбулентным. Наконец, работа с газовыми линзами обходится дороже, чем со стеклянными, поскольку требуется затрачивать энергию для поддержания необходимого температурного градиента [29]. Однако малые потери газовых линз являются довольно существенным их преимуществом и весьма вероятно, что эта отличительная черта может



Ф и г. 4.5.1. Схема газовой линзы.

оказаться достаточно привлекательной для того, чтобы обеспечить их использование для определенных целей. Малые потери газовых линз были наглядно продемонстрированы Бекком [30]. Хорошо известно, что в резонаторах гелий-неоновых лазеров недопустимы значительные потери. Обычно достаточно внести внутрь резонатора на пути лазерного пучка микроскопическое постороннее включение, как колебания сразу пропадают. По этой причине использование стеклянных линз в качестве фокусирующих устройств внутри резонаторов гелий-неонового лазера встречает большие трудности. Бекк продемонстрировал работу гелий-неонового лазера, который имел внутри резонатора не одну, а 78 газовых линз. Длина этого резонатора со всеми линзами составляла 78 м, тогда как работал этот лазер с обычной лазерной трубкой длиной 30 см.

Схематически рассматриваемая газовая линза изображена на фиг. 4.5.1.

Газовая линза состоит из нагреваемой металлической трубы, в которой поддерживается ламинарное течение газа.

Газ поступает в трубу при комнатной температуре слева и покидает ее при более высокой температуре справа. Поскольку газ нагревается стенками трубы, то тепло, радиально проникая внутрь газа, образует некоторый температурный градиент. В каждом поперечном сечении наиболее охлажденные участки находятся на оси трубы, и чем ближе к стенке, тем более нагретым является газ [31].

Объем V , давление p и температура T идеального газа связаны уравнением

$$pV = \mu RT. \quad (4.5.1)$$

Газовая постоянная R равна

$$R = 8,315 \text{ Дж} \cdot \text{град}^{-1} \cdot \text{моль}^{-1}. \quad (4.5.2)$$

Постоянная μ равна числу молей в объеме V . Обозначим через M вес одного моля газа (один моль газа равен его молекулярному весу в граммах). Например, кислород O_2 имеет молекулярный вес 32. Один моль кислорода составляет, следовательно, 32 г этого газа. Тогда

$$m = \mu M \quad (4.5.3)$$

есть истинная масса газа, содержащегося в объеме V . Плотность ρ газа определяется как его масса в единице объема:

$$\rho = \frac{m}{V}. \quad (4.5.4)$$

С помощью приведенных определений уравнение состояния идеального газа можно переписать в виде

$$\frac{p}{\rho} = \frac{R}{M} T. \quad (4.5.5)$$

Экспериментально установлено, что показатель преломления газа зависит от его плотности следующим образом:

$$n = 1 + (n_0 - 1) \frac{\rho}{\rho_0}, \quad (4.5.6)$$

где ρ_0 — некоторая постоянная средняя плотность газа, а n_0 — показатель преломления, соответствующий этой плотности.

Газовая линза работает при постоянном (атмосферном) давлении. Это значит, что плотность газа обратно пропор-

циональна его температуре. Уравнение (4.5.6) можно, следовательно, записать как

$$n = 1 + (n_0 - 1) \frac{T_0}{T}. \quad (4.5.7)$$

Отсюда видно, что показатель преломления газа уменьшается с ростом температуры. В центре трубы газ имеет наибольший показатель преломления, так как там он менее нагрет, и величина n тем меньше, чем ближе к стенке трубы. Световой пучок, проходящий внутри газовой линзы, отклоняется по направлению к области с более высоким значением показателя преломления. Пучок света изгибается, как показано на фиг. 4.5.1, и покидает газовую линзу под более крутым углом к оптической оси, чем тот, который он имеет на входе.

РЕШЕНИЕ ТЕРМОДИНАМИЧЕСКОЙ ЗАДАЧИ

Детальный анализ газовой линзы требует некоторого отступления в область термодинамики. Поскольку соответствующие уравнения легко получить, их вывод будет здесь приведен.

Рассмотрим некоторый объем, который неподвижен относительно трубы и расположен в движущейся среде с неоднородным распределением температуры. Количество тепла в малом объеме дается выражением

$$c_p \int_V \rho T dV. \quad (4.5.8)$$

Константа c_p называется теплоемкостью при постоянном давлении. Произведение температуры T и массы $dm = \rho dV$ дает количество тепла, содержащегося в элементе объема с массой dm ; путем интегрирования получается количество тепла внутри объема V . Если количество тепла в объеме V изменяется, то это происходит вследствие проникновения тепла в объем из-за теплопроводности и из-за течения газа в объеме. Количество тепла, поступающего в объем за 1 с благодаря теплопроводности, пропорционально градиенту температуры T :

$$k \int_A \nabla T \cdot dA. \quad (4.5.9)$$

Вектор $d\mathbf{A}$ является внешним по отношению к замкнутой поверхности A . Это выражение записано для притока тепла через поверхность A , ограничивающую объем V . Множитель k называется теплопроводностью. Наконец, тепло может поступать путем притока нагретого газа внутрь объема. Полный приток тепла в соответствии с этим механизмом равен

$$-c_p \int_A \rho T \mathbf{v} \cdot d\mathbf{A}. \quad (4.5.10)$$

Знак минус стоит потому, что вектор $d\mathbf{A}$ направлен наружу из объема. Комбинация $-\mathbf{v} \cdot d\mathbf{A}$ определяет объем вещества, входящего внутрь объема V за единицу времени. Три полученных выражения позволяют составить уравнение энергетического баланса

$$c_p \frac{\partial}{\partial t} \int_V \rho T dV = k \int_A \nabla T \cdot d\mathbf{A} - c_p \int_A \rho T \mathbf{v} \cdot d\mathbf{A}. \quad (4.5.11)$$

Изменение энергии в объеме за единицу времени (левая часть уравнения) вызывается притоком тепла (первый член справа) и втеканием газа (второй член справа). Применение теоремы о дивергенции к обоим поверхностным интегралам дает

$$\int_V \left\{ c_p \left[\frac{\partial}{\partial t} (\rho T) + \operatorname{div} (\rho T \mathbf{v}) \right] - k \nabla^2 T \right\} dV = 0. \quad (4.5.12)$$

Поскольку объем совершенно произволен, то подынтегральная функция должна быть равна нулю. Запишем это уравнение в слегка измененной форме

$$c_p \left\{ T \left[\frac{\partial \rho}{\partial t} + \operatorname{div} (\rho \mathbf{v}) \right] + \rho \frac{\partial T}{\partial t} + \rho \mathbf{v} \cdot \nabla T \right\} = k \nabla^2 T. \quad (4.5.13)$$

Теперь воспользуемся уравнением непрерывности (3.7.1), выведенным в разд. 3.7. Применим его здесь для трехмерного случая и заменим в (3.7.1) координату z на время t . Это позволит опустить первый член в левой части уравнения (4.5.13), так что окончательно получим дифференциальное уравнение в частных производных, описывающее перенос тепла в движущейся среде:

$$\frac{k}{c_p} \nabla^2 T = \rho \frac{\partial T}{\partial t} + \rho \mathbf{v} \cdot \nabla T. \quad (4.5.14)$$

Для теории газовой линзы интерес представляет только стационарный процесс, который устанавливается после протекания газа в трубе в течение некоторого времени. Для стационарного случая

$$\frac{\partial T}{\partial t} = 0. \quad (4.5.15)$$

Получающееся при этом уравнение все еще не поддается решению, если считать плотность ρ и скорость v переменными величинами. Точное решение задачи требует рассмотрения динамики движущейся среды при наличии неоднородного распределения тепла. Эта задача слишком трудна. Поэтому для приближенного решения предполагают, что плотность ρ почти постоянна внутри трубы, скорость соответствует ламинарному течению, вязкий поток на протяжении трубы находится при постоянной температуре. Распределение скорости вязкой среды в ламинарном потоке при постоянной температуре определяется компонентами

$$\begin{aligned} v_x &= 0, \\ v_y &= 0, \end{aligned} \quad (4.5.16)$$

$$v_z = v_0 \left[1 - \left(\frac{r}{a} \right)^2 \right], \quad (4.5.17)$$

где a — радиус трубы. Задача о потоке в такой упрощенной форме известна как задача Гратца. Она рассмотрена в книге Якоба [32], посвященной переносу тепла. Задача о стационарном процессе в этом приближении следует из (4.5.14)—(4.5.17):

$$\alpha \left(\frac{\partial^2 T}{\partial r^2} + \frac{1}{r} \frac{\partial T}{\partial r} + \frac{\partial^2 T}{\partial z^2} \right) = v_0 \left[1 - \left(\frac{r}{a} \right)^2 \right] \frac{\partial T}{\partial z}, \quad (4.5.18)$$

где

$$\alpha = \frac{k}{\rho c_p}. \quad (4.5.19)$$

Здесь предполагается, что распределение температуры не зависит от азимутального угла ϕ . Это предположение является еще одним приближением, поскольку гравитационное поле Земли искажает газовый поток, если газовая линза ориентирована горизонтально. Гравитационный

эффе́кт приводит к искажению линзы, но ради упрощения будем игнорировать этот эффе́кт, ибо учет его значительно усложняет задачу [33, 70*, 71*].

Начнем решение задачи о распределении тепла в газовой линзе с введения новой переменной

$$\Theta = \frac{T - T_w}{T_w - T_0}. \quad (4.5.20)$$

Здесь T_w — температура стенки трубы; T_0 — температура газа, поступающего в трубу. Функция Θ является решением такого же дифференциального уравнения, как и (4.5.18), но с более простыми граничными условиями. Потребуем, чтобы $T = T_w$ при $r = a$. Если воспользоваться относительной независимой переменной

$$u = \frac{r}{a}, \quad (4.5.21)$$

то это требование запишется как

$$\Theta(u) = 0 \quad \text{при} \quad u = 1. \quad (4.5.22)$$

Из физических соображений можно ожидать, что решение симметрично по координате u и поэтому функция Θ симметрична относительно аргумента u . Будем искать решение дифференциального уравнения в частных производных (4.5.18) в виде

$$\Theta = AR(u) e^{-\beta^2(\alpha/a^2 v_0)z}. \quad (4.5.23)$$

Подставляя выражение (4.5.23) в (4.5.18) (с заменой T на Θ), получаем обыкновенное дифференциальное уравнение

$$\frac{d^2 R}{du^2} + \frac{1}{u} \frac{dR}{du} + \beta^2(1 - u^2)R = 0. \quad (4.5.24)$$

Членом $d^2 T/dz^2$ в (4.5.18) мы пренебрегли. Это приближение справедливо для достаточно больших значений v_0 .

Решения дифференциального уравнения (4.5.24) относятся к так называемым функциям Уиттекера. Однако в связи с тем, что они недостаточно широко табулированы, необходимо получить для них разложение в степенной ряд. Поскольку R является четной функцией u , можно ограничиться степенным рядом из четных степеней u

$$R = \sum_{\nu=0}^{\infty} C_{2\nu} u^{2\nu}. \quad (4.5.25)$$

Тогда

$$R(0) = 1, \quad (4.5.26)$$

$$C_0 = 1. \quad (4.5.27)$$

Подстановка выражения (4.5.25) в (4.5.24) с учетом (4.5.27) дает

$$C_2 = -\frac{1}{4} \beta^2 \quad (4.5.28)$$

при $u = 0$. Сравнение коэффициентов при u^2 в уравнении, полученном после подстановки выражения (4.5.25) в (4.5.24), приводит к следующему рекуррентному соотношению для коэффициентов:

$$C_{2v} = \frac{\beta^2}{(2v)^2} (C_{2v-4} - C_{2v-2}) \quad \text{при } v \geq 2. \quad (4.5.29)$$

Если β известно, то задача решена. Но собственное значение β должно быть определено из уравнения

$$R(1) = 0, \quad (4.5.30)$$

которое следует из граничного условия (4.5.22).

Решить эту задачу о собственном значении можно лишь численно с помощью вычислительной машины. Имеется бесчисленное множество решений уравнения (4.5.30). Их отличают друг от друга при помощи индексов, приписываемых β и R . Решение физической задачи получают как суперпозицию всех собственных решений. Функцию Θ можно, таким образом, переписать для наиболее общего случая в виде

$$\Theta = \sum_{\mu=0}^{\infty} A_{\mu} R_{\mu}(u) e^{-(\alpha/a^2 v_0) \beta_{\mu}^2 z}. \quad (4.5.31)$$

Вычисление собственных значений β_{μ} высокого порядка представляет довольно сложную задачу. Ряд (4.5.25) сходится не очень быстро. Коэффициенты C_{2v} возрастают до огромной величины, прежде чем начнут уменьшаться. Оказалось возможным получить решения задачи о собственных значениях только до β_8 даже с использованием двойной точности, так как коэффициенты росли до величин порядка 10^{20} , тогда как величина $R(u)$ не превосходила единицы. По этой причине возникла необходимость применить разложение в ряд $R(u)$. Разложение (4.5.25)

можно использовать при $0 \leq u \leq 0,5$. Произведения $C_{2\nu} u^{2\nu}$ остаются при этом в разумных пределах. Если $u > 0,5$, то нужно ввести новую переменную

$$w = 1 - u \quad (4.5.32)$$

и воспользоваться разложением

$$R(u) = \sum_{\nu=0}^{\infty} D_{\nu} w^{\nu}. \quad (4.5.33)$$

Дифференциальное уравнение для R , выраженное через переменную w , имеет вид

$$(1-w) \frac{d^2 R}{dw^2} - \frac{dR}{dw} + \beta^2 (2w - 3w^2 + w^3) R = 0. \quad (4.5.34)$$

Уравнение (4.5.30) теперь записывается как

$$R = 0 \text{ при } w = 0 \quad (4.5.35)$$

и приводит к условию

$$D_0 = 0. \quad (4.5.36)$$

Подставляя (4.5.33) в (4.5.34), получаем рекуррентные соотношения

$$\begin{aligned} D_2 &= \frac{1}{2} D_1, \\ D_3 &= \frac{1}{3} D_1, \\ D_4 &= \left(\frac{1}{4} - \frac{1}{6} \beta^2 \right) D_1, \end{aligned} \quad (4.5.37)$$

$$D_{\nu} = \frac{1}{\nu(\nu-1)} [(\nu-1)^2 D_{\nu-1} - \beta^2 (2D_{\nu-3} - 3D_{\nu-4} + D_{\nu-5})]. \quad (4.5.38)$$

Собственное значение β и коэффициент D_1 должны теперь выбираться так, чтобы функция, представленная рядом (4.5.25), и ее первая производная были непрерывны. Эта трудная задача может быть решена при помощи электронной вычислительной машины [34]. Результаты вычислений собственных значений, первой производной R' функции $R(u)$ при $u = 1$, а также производной $\partial R / \partial \beta$ при $u = 1$ и $\beta = \beta_{\mu}$ приведены в табл. 4.5.1.

Таблица 4.5.1

Решения задачи о собственных значениях,
производные функции $R_\mu(u)$ при $u=1$, а также ее
производные по собственному значению β ,
вычисленные при $u=1$ и $\beta=\beta_\mu$

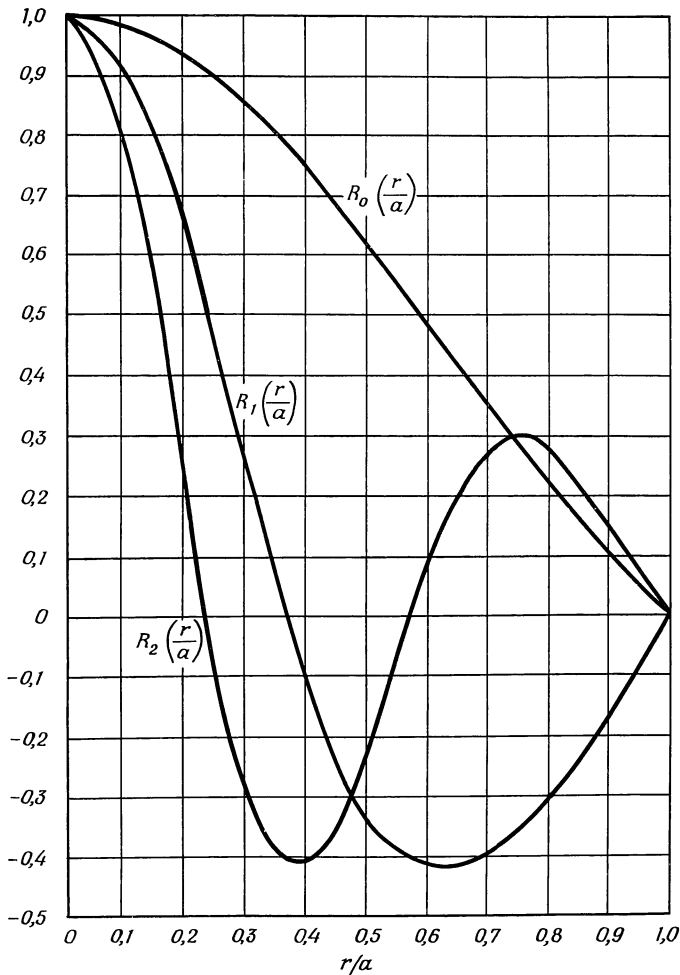
μ	β_μ	$R'_\mu(1)$	$\frac{\delta R_\mu}{\delta \beta} (\beta = \beta_\mu, u = 1)$
0	2,70436	-1,01430	-0,50090
1	6,67903	1,34924	0,37146
2	10,67338	-1,57232	-0,31826
3	14,6711	1,74600	0,28648
4	18,6699	-1,89090	-0,26449
5	22,6691	2,01647	0,24799
6	26,6686	-2,12814	-0,23491
7	30,6682	2,22038	0,22485
8	34,6679	-2,32214	-0,21548
9	38,6676	2,40274	0,20779
10	42,6667	-2,48992	-0,20108
11	46,6667	2,56223	0,19516
12	50,6668	-2,64962	-0,18988
13	54,6668	2,70216	0,18513
14	58,6668	-2,76421	-0,18083

Функции $R_\mu(u)$ трех первых порядков и изображены на фиг. 4.5.2. Для определения коэффициентов A_μ разложения (4.5.31) необходимо иметь соотношение ортогональности функций $R_\nu(u)$. С помощью дифференциального уравнения (4.5.24) можно доказать, что имеет место следующее условие ортогональности:

$$\int_0^1 u(1-u^2) R_\mu(u) R_\nu(u) du = 0 \quad \text{при} \quad \mu \neq \nu. \quad (4.5.39)$$

Условие нормировки функций R следует из интеграла

$$\int_0^1 u(1-u^2) R_\mu^2(u) du = \frac{1}{2\beta_\mu} \left(\frac{\partial R_\mu}{\partial \beta} \frac{\partial R_\mu}{\partial u} \right)_{u=1, \beta=\beta_\mu}. \quad (4.5.40)$$



Ф и г. 4.5.2. Графики функций $R_\mu(r/a)$ трех первых порядков [31].

Коэффициенты разложения в (4.5.31) могут быть получены из условия $T = T_0$ при $z = 0$. Температура газа T_0 на входе не зависит от u . Это условие на входе дает

$$\Theta = -1 \quad \text{при} \quad z=0. \quad (4.5.41)$$

Из формул (4.5.41) и (4.5.31) получаем коэффициенты разложения с помощью соотношений (4.5.39) и (4.5.40):

$$\begin{aligned} \frac{1}{2\beta_\mu} \left(\frac{\partial R_\mu}{\partial \beta} \frac{\partial R_\mu}{\partial u} \right)_{\substack{u=1 \\ \beta=\beta_\mu}} A_\mu &= - \int_0^1 u (1-u^2) R_\mu(u) du = \\ &= \frac{1}{\beta_\mu^2} \left(\frac{\partial R_\mu}{\partial u} \right)_{u=1}. \end{aligned} \quad (4.5.42)$$

Вычисление интеграла было проведено с помощью дифференциального уравнения (4.5.24). Коэффициенты разложения равны

$$A_\mu = \frac{2}{\beta_\mu \left(\frac{\partial R_\mu}{\partial \beta} \right)_{\substack{u=1 \\ \beta=\beta_\mu}}}. \quad (4.5.43)$$

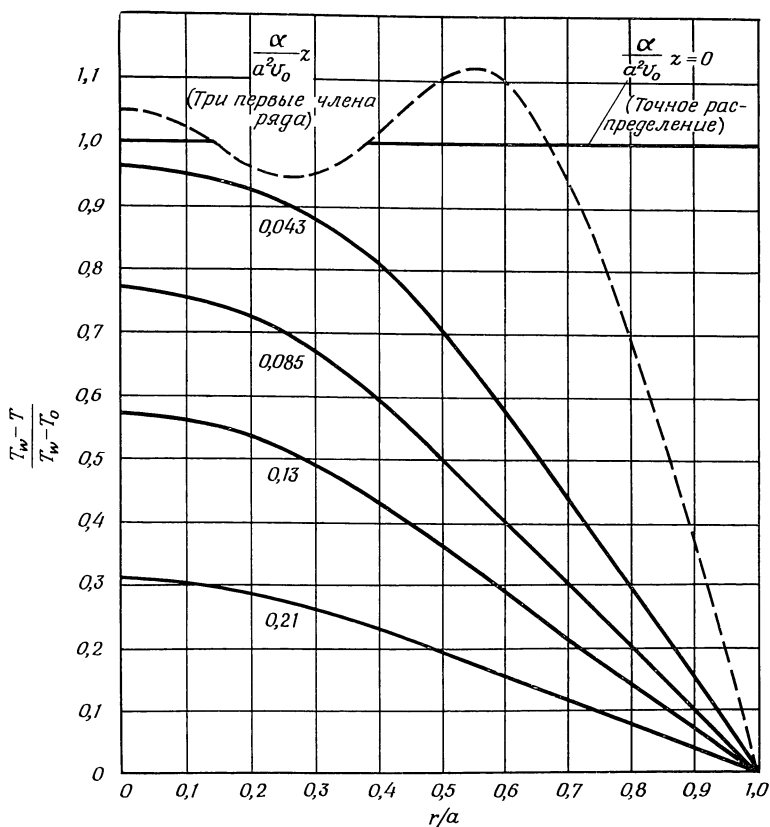
Производные от функции R по собственному значению приведены в табл. 4.5.1. Распределение температуры внутри газовой линзы получаем теперь из формул (4.5.20), (4.5.31) и (4.5.43):

$$\begin{aligned} T &= T_w + 2(T_w - T_0) \sum_{\mu=0}^{\infty} R_\mu(u) \times \\ &\times [\beta_\mu (\partial R_\mu / \partial \beta)]_{\substack{u=1 \\ \beta=\beta_\mu}}^{-1} e^{-(\alpha/a^2 v_0) \beta_\mu^2 z}. \end{aligned} \quad (4.5.44)$$

Этот ряд хорошо сходится при положительных значениях z . Из табл. 4.5.1 видно, что значения β_μ очень быстро возрастают с ростом μ . Экспоненциальная функция в (4.5.44) быстро становится исчезающе малой и ею можно пренебречь везде, за исключением нескольких первых значений μ . Ряд медленно сходится только при $z = 0$.

Данное рассмотрение дает возможность получить приближенное решение задачи о тепловом потоке в газовой линзе. Распределение температуры внутри газовой линзы показано на фиг. 4.5.3 для нескольких значений нормированной длины вдоль трубы. Хорошая сходимость ряда (4.5.44) становится очевидной, если учесть, что при $\alpha z/a^2 v_0 = 0,043$ экспоненциальные множители в разложении равны 0,73; 0,147; 0,0075; 0,000096 и т. д. Это говорит

о том, что уже первые три члена ряда дают очень хорошее приближение к истинному температурному распределению.



Ф и г. 4.5.3. Распределение температуры в газовой линзе для пяти различных значений нормированного параметра длины $\alpha z / a^2 \nu_0$. Пунктирной линией дана приближенная кривая, построенная по трем членам ряда для $z = 0$ [31].

Нормированный параметр длины кривых на фиг. 4.5.3 соответствует множителю, стоящему в показателе экспоненциальной функции в (4.5.44).

ТРАЕКТОРИЯ ЛУЧА

После того как установлено температурное распределение внутри газовой линзы, можно перейти к определению ее оптических свойств. Основываться будем на уравнении параксиальных лучей (3.2.19). Поскольку показатель преломления газа весьма близок к единице (для воздуха $n = 1,000293$ при 20°C), то положим в левой части уравнения (3.2.19) $n = 1$. Используя цилиндрические координаты r , ϕ , z и принимая, что для траекторий лучей, падающих параллельно оптической оси, отсутствует зависимость от ϕ , получаем с хорошим приближением

$$\frac{d^2 r}{dz^2} = \frac{\partial n}{\partial r}. \quad (4.5.45)$$

Показатель преломления дается выражением (4.5.7). Для случая малых температурных перепадов можно написать

$$\frac{\partial n}{\partial r} = -(n_0 - 1) \frac{T_0}{T^2} \frac{\partial T}{\partial r} \approx -(n_0 - 1) \frac{1}{T_0} \frac{\partial T}{\partial r}. \quad (4.5.46)$$

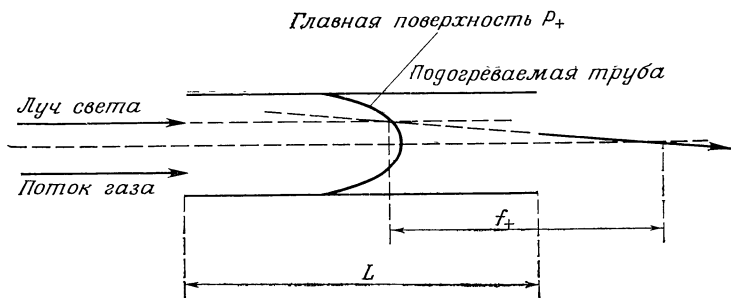
Уравнение параксиальных лучей принимает вид

$$\frac{d^2 r}{dz^2} = -(n_0 - 1) \frac{1}{T_0} \frac{\partial T}{\partial r}. \quad (4.5.47)$$

Формулы (4.5.47) и (4.5.44) могут быть использованы для получения траекторий параксиальных лучей, проходящих через газовую линзу. Вследствие сложного вида температурного распределения аналитическое решение уравнения (4.5.47) получить невозможно. Для того чтобы определить свойства линз исследуемого устройства, необходимо обратиться к численным решениям, полученным с помощью вычислительной машины [34].

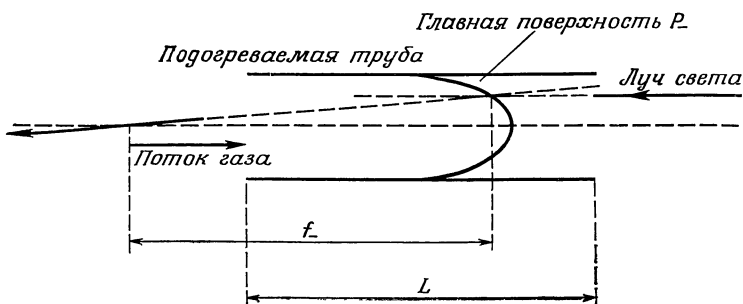
Свойства газовой линзы определяются двумя величинами. Как и в случае обычной линзы, необходимо знать фокусное расстояние. Однако в связи с протяженной физической длиной газовой линзы необходимо добавить еще один параметр, для того чтобы полностью охарактеризовать ее свойства. Оказывается, что газовую линзу можно приближенно считать тонкой, хотя она имеет значительную длину. Однако эта эквивалентная тонкая линза не будет плоской и должна иметь кривую поверхность

с круговой симметрией. В теории толстых линз используется концепция главных плоскостей для описания свойств



Ф и г. 4.5.4. Определение главной поверхности в газовой линзе в случае, когда свет распространяется вдоль газового потока.

линз такого рода [1, 15]. Мы воспользуемся этой концепцией для описания газовой линзы, но вместо главных плоскостей будем рассматривать главные поверхности. Определение главных поверхностей дано на фиг. 4.5.4 и 4.5.5.



Ф и г. 4.5.5. Определение второй главной поверхности в газовой линзе в случае, когда свет распространяется навстречу газовому потоку.

Газовый поток направлен слева направо. На фиг. 4.5.4 изображен луч света, падающий слева параллельно оптической оси. Его истинная траектория внутри линзы не показана. Траектория входного луча продлена внутрь линзы прямой пунктирной линией. Выходной луч внутри линзы

изображен другой прямой пунктирной линией. Пересечение этих двух линий определяет одну из точек главной поверхности P_+ . Совокупность точек, определяемых с помощью всех возможных входных лучей, параллельных оптической оси, не лежит в одной плоскости. Нахождение этой главной поверхности является одной из задач при численном решении уравнения (4.5.47). Другая главная поверхность P_- показана на фиг. 4.5.5. Ее определение идентично определению P_+ , но падающий луч входит в линзу справа навстречу газовому потоку параллельно оптической оси. Вовсе не обязательно, чтобы обе поверхности P_+ и P_- совпадали. Их несовпадение фактически указывает на наличие оптической толщины у газовой линзы. Если же обе поверхности совпадают, то для того чтобы проследить за ходом луча через линзу, оказывается достаточным одной главной поверхности. В этом случае линза является оптически тонкой. Тонкая линза в разд. 4.2 рассматривалась как плоская. Траектории лучей находились путем преломления каждого луча в точке пересечения его с плоскостью на величину, которая зависит от фокусирующей способности линзы. Можно аналогичным образом описать действие газовой линзы, преломляя каждый луч в точке пересечения его с главной поверхностью. Если у линзы две главные поверхности, то ход лучей, падающих параллельно оптической оси, можно проследить, преломляя их на соответствующих главных поверхностях. Если у линзы только одна главная поверхность, то именно на ней должен быть преломлен каждый луч; найти траекторию луча, прошедшего через газовую линзу, не труднее, чем определить ее в случае обычной тонкой линзы при условии, что известна форма главной поверхности. Действие газовой линзы можно, с другой стороны, описать с использованием фокусного расстояния. Как показано на фиг. 4.5.4 и 4.5.5, фокусное расстояние измеряется от той точки, в которой лучи пересекают главную поверхность. Если имеются две различные главные поверхности, то не удивительно, что для описания свойств линзы нужно ввести два разных фокусных расстояния f_+ и f_- . К счастью, газовая линза практически является тонкой линзой. Обе ее главные поверхности почти совпадают, и поэтому фокусные расстояния f_+ и f_- близки друг

к другу. Численное решение уравнения (4.5.47) дает положение и наклон лучей в том сечении, где они покидают линзу. Геометрическое место точек главной поверхности P_+ определяется соотношением

$$z = P_+(r), \quad (4.5.48)$$

а для главной поверхности P_-

$$z = P_-(r). \quad (4.5.49)$$

Положение луча при $z = 0$ описывается функцией $r_{\pm}(0)$, а положение при $z = L$ — функцией $r_{\pm}(L)$. Аналогичным образом наклоны лучей dr/dz описываются функциями $r'_{\pm}(0)$ и $r'_{\pm}(L)$. Таким образом, фокусные расстояния равны

$$f_+ = -\frac{r_+(0)}{r'_+(L)} \quad (4.5.50)$$

и

$$f_- = \frac{r_-(L)}{r'_-(0)}. \quad (4.5.51)$$

Главные поверхности определяются теперь соотношениями

$$P_+ = L + \frac{r_+(0) - r_+(L)}{r'_+(L)} \quad (4.5.52)$$

и

$$P_- = \frac{r_-(L) - r_-(0)}{r'_-(0)}. \quad (4.5.53)$$

Индексами «+» и «-» отмечены лучи, идущие соответственно вправо и влево. Плоскость $z = 0$ совпадает с входным сечением газовой линзы (левый конец трубы). Выбор знака диктуется тем обстоятельством, что $r'_+(L) < 0$, а $r'_-(0) > 0$.

Для того чтобы иметь возможность оперировать с безразмерными величинами, введем некоторые нормированные параметры. Для этой цели полезным является параметр ¹⁾

$$V = \frac{\alpha L}{a^2} = \frac{kL}{\rho c a_p^2}. \quad (4.5.54)$$

Поскольку здесь показатель степени экспоненциальной функции должен быть безразмерной величиной, то ясно,

¹⁾ Для воздуха при 20 °C имеем: $\rho = 1,21 \cdot 10^{-3}$ г/см³, $k = 6,28 \cdot 10^{-5}$ кал/см · °град, $c_p = 0,24$ кал/г · °град, $n - 1 = 2,93 \cdot 10^{-4}$.

что V имеет размерность скорости. Важно помнить, что величина V пропорциональна длине газовой линзы. Кроме того, этот параметр определяется также теплопроводностью k и удельной теплоемкостью c_p применяемого газа, а также размером поперечного сечения линзы. В связи с тем, что координата z входит в формулу (4.5.44) в виде комбинации

$$s = \frac{\alpha}{a^2 v_0} z = \frac{V}{v_0} \frac{z}{L}, \quad (4.5.55)$$

удобно выразить уравнение лучей через эту координату. Учитывая выражение (4.5.21), получаем из (4.5.47)

$$\frac{d^2 u}{ds^2} = - \left(\frac{v_0}{V} \right)^2 \left(\frac{L}{a} \right)^2 (n_0 - 1) \frac{1}{T_0} \frac{\partial T}{\partial u}. \quad (4.5.56)$$

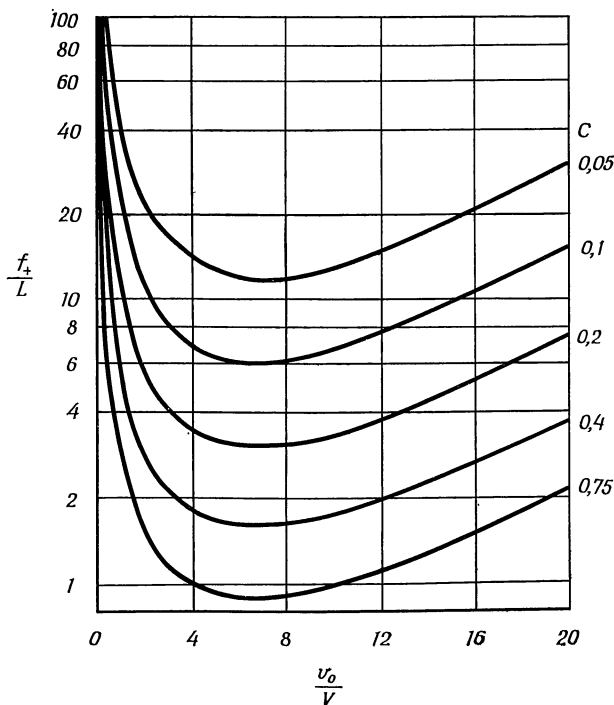
Это уравнение содержит только безразмерные величины и поэтому может быть использовано для численного расчета. Из формул (4.5.44) и (4.5.56) видно, что константа

$$C = (n_0 - 1) \frac{T_w - T_0}{T_0} \left(\frac{L}{a} \right)^2 \quad (4.5.57)$$

является еще одним безразмерным параметром, используемым для расчета линзы. Он содержит информацию о длине и поперечном сечении линзы, а также о показателе преломления газа и температуре нагреваемой трубы.

Результаты численного решения рассматриваемой задачи о газовой линзе приведены на нескольких фигурах. На фиг. 4.5.6 представлены графики зависимости нормированного фокусного расстояния газовой линзы для лучей, весьма близких к оси линзы ($r_+(0)/a = 0,1$), от нормированной скорости газа. Наиболее замечательным свойством этих кривых является наличие минимума. Тот факт, что существует оптимальная скорость газа, при которой фокусное расстояние короче, чем при любой другой скорости, можно легко объяснить. Рассмотрим сначала газовую линзу, в которой газ неподвижен ($v_0 = 0$). В этом случае газ в трубе прогрет до температуры трубы. При одинаковой температуре газа линза не фокусирует ($f = \infty$). При протекании газа через линзу устанавливается некоторое температурное распределение между началом и концом трубы. Однако в случае очень малых скоростей газ успевает полностью прогреться, прежде чем достигнет конца трубы,

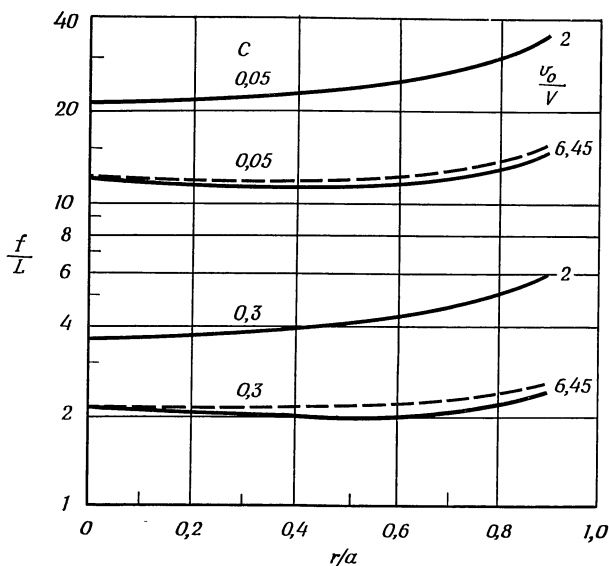
и в результате фокусирующее действие оказывается слабым. По мере возрастания скорости фокусирующая способность линзы становится все сильнее. Этим и объясняется



Ф и г. 4.5.6. Нормированное фокусное расстояние f_+/L как функция нормированной скорости газового потока v_0/V [34].

уменьшение фокусного расстояния слева от минимума на фиг. 4.5.6. При возрастании скорости наступает такой момент, когда газ покидает трубу, прежде чем успеет прогреться в достаточной мере. Фокусирующая способность линзы начинает ослабевать, и фокусное расстояние опять возрастает. При бесконечно большой скорости (если это возможно для ламинарного потока) газ проходит через линзу, не изменяя своей входной температуры, и поэтому фокусирующее действие становится невозмож-

ным. Таким образом, фокусное расстояние обращается в бесконечность в двух крайних случаях: для нулевой и бесконечно большой скорости газа. Параметр C соответствует разным температурам трубы или газам с разными показателями преломления или комбинации этих двух случаев. В случае продувки воздуха через трубу длиной 20 см и радиусом 0,3 см, нагретую на 50 °С больше входной

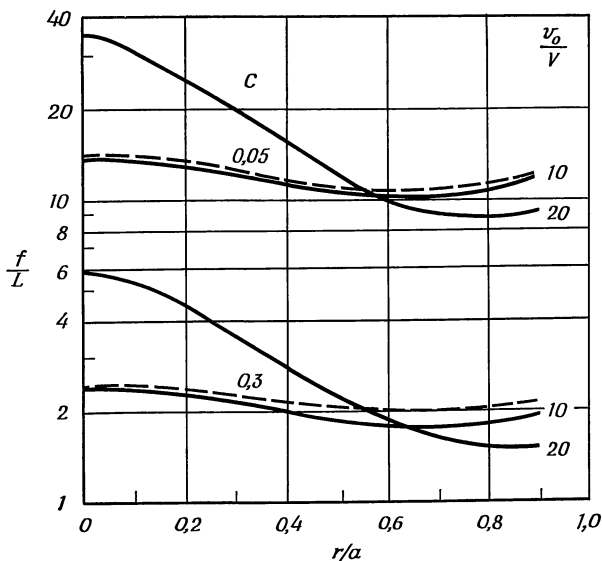


Ф и г. 4.5.7. Фокусное расстояние как функция расстояния от оси входного луча при $v_0/V = 2$ и 6,45 [34].

— f_+/L , — — — f_-/L .

температуры газа, получаем $C = 0,2$. Оптимальная скорость течения $v_0/V = 6,4$ соответствует истинной осевой скорости $v_0 = 272$ см/с (эта осевая скорость получается при скорости течения газа 0,04 л/с). Оптимальное фокусное расстояние $f = 3L = 60$ см. Можно получить еще более короткое фокусное расстояние, если выбрать более длинную газовую трубу, газ с более высоким показателем преломления и более высокой рабочей температурой. Фокусное расстояние порядка 20 см можно реализовать даже с воздухом [35].

Фокусирующая способность газовой линзы довольно значительна. Однако, чтобы получить полную картину работы линзы этого типа, нужно изучить вносимые ею искажения. На фиг. 4.5.7 и 4.5.8 показана зависимость фокусного расстояния от положения луча на входе трубы.



Ф и г. 4.5.8. Фокусное расстояние как функция расстояния от оси входного луча при $v_0/V = 10$ и 20 [34].
 ————— f_+/L , — — — — f_-/L .

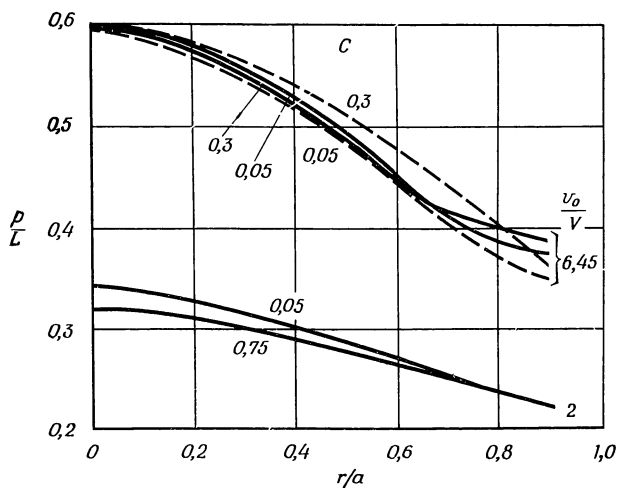
Видно, что при оптимальном значении скорости течения $v_0/V = 6,4$ зависимость фокусного расстояния от r незначительна. При более низких скоростях ($v_0/V = 2$) абберации несколько возрастают. Искажение фокусного расстояния становится весьма большим при высоких скоростях потока. Это видно из кривых на фиг. 4.5.8; фокусные расстояния f_+ (в случае луча, входящего в линзу в направлении течения газа) и f_- (для лучей, проходящих в противоположном направлении) почти не отличаются друг от друга.

Однако по одним только кривым фокусного расстояния нельзя сделать окончательного вывода о всех искажениях

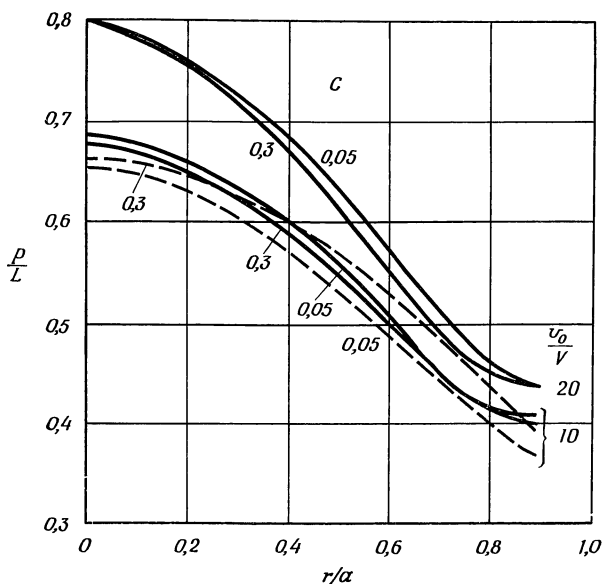
линзы. Нужно помнить, что фокусное расстояние линзы измерялось относительно главных поверхностей. Общие сведения об искажениях линзы можно получить только после исследования формы главных поверхностей.

Форма главных поверхностей показана на фиг. 4.5.9 и 4.5.10, из которых видно, что главные поверхности P_+ и P_- почти совпадают. Это подтверждает то, что газовая линза подобна тонкой линзе. Однако искривление главной поверхности довольно значительное. Для скорости потока, соответствующей оптимуму по фиг. 4.5.6, главная поверхность размещается вблизи середины линзы, а в продольном направлении она занимает более трети длины газовой линзы. Для коротких линз и больших фокусных расстояний это искривление можно сделать малым. Из сравнения фиг. 4.5.7 и фиг. 4.5.9 видно, что искажение фокусного расстояния и искривление главной плоскости имеют тенденцию компенсировать друг друга. Из фиг. 4.5.7 следует, что фокусное расстояние длиннее для лучей, входящих в линзу ближе к стенке, тогда как из фиг. 4.5.9 вытекает, что главная поверхность размещена ближе к входному концу трубы. Если разместить фокус относительно входного конца трубы при $z = 0$, то возрастание фокусного расстояния и отклонение главной поверхности стремятся компенсировать друг друга, так что расположение фокуса относительно газовой трубы почти не изменяется. Такая компенсация с успехом может быть использована при создании газовой линзы с минимальными искажениями.

Для исследования справедливости предположения о тонкой линзе было проведено следующее численное моделирование с помощью ЭВМ [36]. Лучи проходили через последовательность 100 газовых линз. Ход прохождения луча в одном случае был промоделирован путем численного решения лучевого уравнения, для того чтобы найти положение луча у каждой линзы. Второй способ заключался в использовании эквивалентной тонкой линзы, являющейся моделью газовой линзы (фиг. 4.5.11). Форма этой эквивалентной линзы была выбрана в соответствии с главной поверхностью газовой линзы, а фокусное расстояние для каждой точки искривленной тонкой линзы взято соответствующим фокусному расстоянию f_+ , рассчитанному путем проведения параллельных входных лучей через

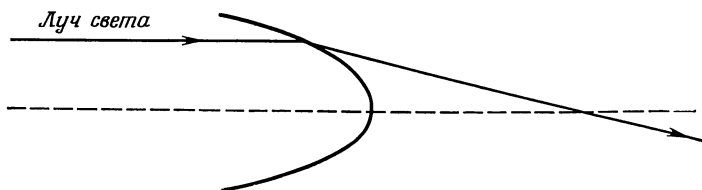


Ф и г. 4.5.9. Форма главных поверхностей при $v_0/V = 2$ и $6,45$ [34]
 ——— P_+/L , — — — — P_-/L .



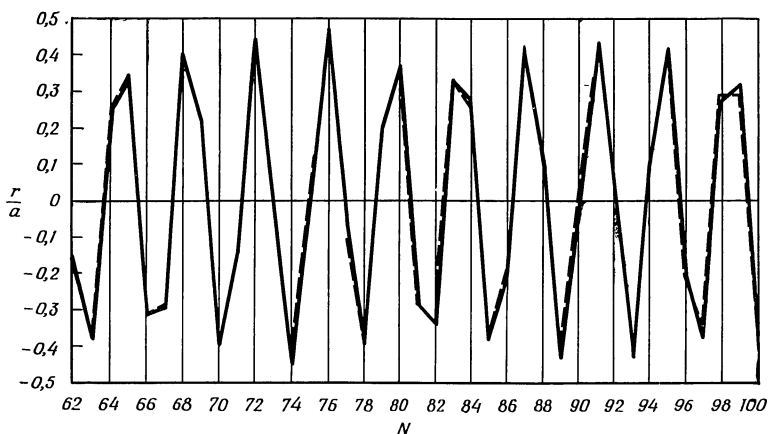
Ф и г. 4.5.10. Форма главных поверхностей при $v_0/V = 10$ и 20 [34].
 ——— P_+/L , — — — — P_-/L ,

линзу. Истинная траектория луча была найдена с помощью формулы (3.7.37). Фиг. 4.5.12 и 4.5.13 демонстрируют



Ф и г. 4.5.11. Форма главной поверхности тонкой линзы, эквивалентной газовой линзе.

хорошее совпадение между результатами, полученными обоими способами. В обоих случаях линзы располагались

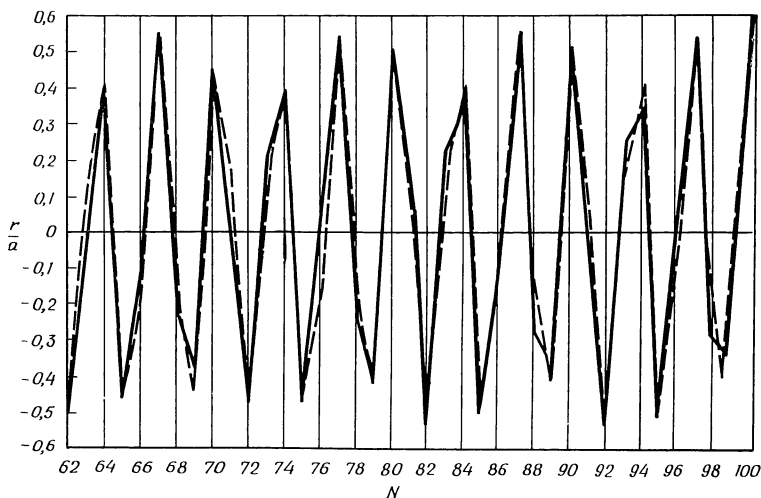


Ф и г. 4.5.12. Сравнение траекторий лучей в линии из газовых линз и эквивалентных им тонких линз при $D/a = 1200$, $D/f_0 = 2,16$, $L/a = 50$ (N — число линз, a — радиус газовой линзы, D — расстояние между линзами, f_0 — фокусное расстояние, L — длина газовой линзы) [36].

друг от друга на расстоянии, примерно равном 0,8 от фокусного расстояния. Линзы, соответствующие фиг. 4.5.12, примерно в 4 раза слабее линз, соответствующих фиг. 4.5.13. Согласие между обоими способами лучше для

слабых линз, но согласие неплохое и в случае сильных линз ¹⁾).

Газовые линзы вполне эффективны по своим фокусирующим свойствам и могут быть выполнены с достаточно



Ф и г. 4.5.13. Траектории лучей в линзах из газовых линз и эквивалентных им тонких линз при $D/a = 330$, $D/f_0 = 2,74$, $L/a = 50$ [36].

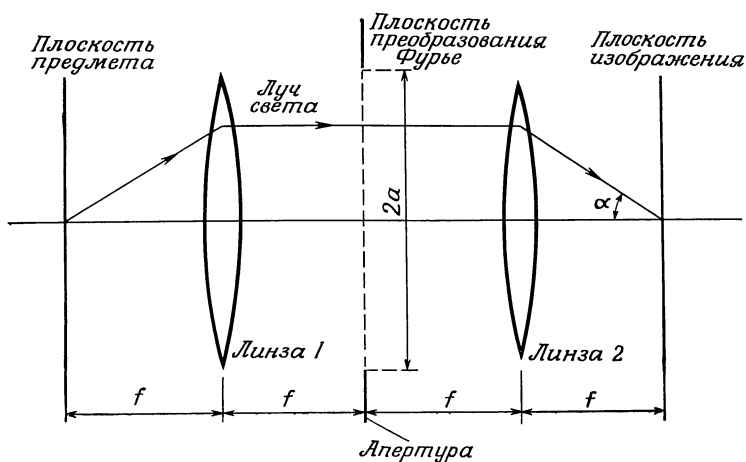
малыми искажениями. Однако их способность формирования изображения по сравнению со стеклянными линзами остается плохой. Их малые потери являются привлекательным свойством, если необходимо передать свет через последовательность многих линз. Мы еще вернемся к этому вопросу в следующей главе.

4.6. ПРЕДЕЛЬНАЯ РАЗРЕШАЮЩАЯ СПОСОБНОСТЬ ПРИ ФОРМИРОВАНИИ ИЗОБРАЖЕНИЯ

В разд. 4.2 было получено выражение (4.2.11) для предельной разрешающей способности при формировании изображения. Вывод был основан на принципе неопреде-

¹⁾ Автор имеет здесь в виду линзы с большим (слабые линзы) и малым (сильные линзы) фокусным расстоянием.— *Прим. ред.*

ленности квантовой теории лучевой оптики. В настоящем разделе рассмотрим критерий Рэля для разрешающей способности оптических приборов и сделаем некоторые замечания по дискуссии, проводимой в последнее время в литературе и лишь добавляющей значительную путаницу в рассматриваемый предмет. Утверждают, что рэлеевское



Ф и г. 4.6.1. Схематическое изображение оптической системы, формирующей изображение.

ограничение является только удобным критерием разрешающей способности оптических приборов и что возможно, хотя бы в принципе, получить неограниченное разрешение [7]. Это утверждение разобрано в статье Торальдо ди Франчия [37], который показал, что если даже можно было бы получить бесконечное разрешение на строго математической основе, практическая реализация этого математического идеала тем не менее невозможна.

Следуя методике Торальдо ди Франчия, рассмотрим систему формирования изображения, показанную на фиг. 4.6.1. Ради простоты будем считать эту систему двумерной. Существенные моменты относительно критерия разрешения можно получить из двумерной задачи; от трехмерной задачи отличие будет весьма незначительным и скажется на численных значениях коэффициентов,

Используя обозначения разд. 4.3, запишем преобразование Фурье поля предмета (объекта) $f(u)$ в задней фокальной плоскости первой линзы в соответствии с формулой (4.3.32) в виде

$$F(v) = \frac{1}{\sqrt{2\pi}} \int_{-u_0/2}^{u_0/2} f(u) e^{iuv} du. \quad (4.6.1)$$

Предполагаем, что поле объекта существует только в ограниченной области и поэтому интеграл берется в пределах от $-1/2 u_0$ до $1/2 u_0$. Вторая линза подвергает поле в плоскости преобразования Фурье еще одному фурье-преобразованию. Однако мы предполагаем, что в плоскости преобразования Фурье имеется апертура, ограничивающая область, через которую может проходить свет. Наличие такой апертуры является неизбежным. Даже при отсутствии реальной апертуры в задней фокальной плоскости первой линзы все равно имеется эффективная апертура, обусловленная конечными размерами двух линз. Оптическая система с неограниченными апертурами невозможна. Используя ту же формулу еще раз и предполагая, что апертура в плоскости преобразования Фурье простирается от $-v_0$ до v_0 , получаем распределение поля \bar{f} в плоскости изображения:

$$\bar{f}(w) = \frac{1}{\sqrt{2\pi}} \int_{-v_0}^{v_0} F(v) e^{ivw} dv. \quad (4.6.2)$$

Подстановка выражения (4.6.1) в (4.6.2) приводит к следующему соотношению:

$$\bar{f}(w) = \int_{-u_0/2}^{u_0/2} \frac{\sin v_0(u+w)}{\pi(u+w)} f(u) du. \quad (4.6.3)$$

Для того чтобы иметь возможность представить аргументацию Рэлея, допустим, что в плоскости объекта находится точечный источник (в двумерном случае ему, конечно, соответствует линейный источник):

$$f(u) = \delta(u). \quad (4.6.4)$$

Точечный источник описывается δ -функцией, которая в начале координат дает для поля бесконечно большое

значение. Изображение этого точечного источника следует из формул (4.6.3) и (4.6.4):

$$\bar{f}_\delta(w) = \frac{\sin v_0 w}{\pi w}. \quad (4.6.5)$$

Для исследования предельного случая, при котором два объекта возможно разрешить, нужна по крайней мере еще одна точка в плоскости объекта. Однако дальнейшее усложнение задачи излишне. Используя аргументацию Рэлея, можно сказать, что вторая точка не должна лежать настолько близко к первой точке, чтобы оба изображения полностью накладывались друг на друга. В качестве практического критерия можно допустить, что два объекта распознаются как различные, когда максимум поля изображения, созданного одним объектом, приходится на первый нуль дифракционной картины, представляющей изображение другой точки. Поскольку размеры изображения и объекта в рассматриваемом устройстве одинаковы (увеличение отсутствует), то можно сделать вывод, что изображения двух объектов являются разрешенными, когда два объекта разделены ($\Delta u = \Delta w$) по крайней мере в соответствии с условием

$$v_0 \Delta u \geq \pi. \quad (4.6.6)$$

Чтобы можно было воспользоваться этим результатом, необходимо перейти к обычным пространственным координатам, используя формулы (4.3.28) и (4.3.29):

$$\Delta u = \sqrt{\frac{2\pi}{\lambda f}} \Delta x \quad (4.6.7)$$

и

$$v_0 = \sqrt{\frac{2\pi}{\lambda f}} a. \quad (4.6.8)$$

Разнесение объектов выражается через реальную координату x с помощью Δx : величина a есть полуширина апертуры в плоскости преобразования Фурье. Используя соотношение (4.6.6), получаем критерий Рэлея для предельной разрешающей способности различения двух точек:

$$\Delta x \geq \frac{\lambda f}{2a}. \quad (4.6.9)$$

Критерий Рэлея можно выразить через угол α , под которым луч приходит в точку изображения:

$$\Delta x \geq \frac{\lambda}{2\alpha}. \quad (4.6.10)$$

Использование приближения $a/f = \alpha$ является логичным, так как вывод был основан на теории дифракции в предположении малых углов. За исключением множителя $(2\pi)^{-1}$, соотношение (4.6.10) совпадает с критерием разрешимости (4.2.11). Поскольку критерии для разрешения изображений в обоих случаях не идентичны, то совпадение между этими двумя формулами близко настолько, насколько это можно было ожидать.

Критерий Рэлея полезен для оценки величины разрешающей способности, которую можно достичь с помощью прибора, формирующего изображение. Множители $1/2$ или $1/4\pi$, появляющиеся в этих выражениях, не являются существенными. Важным является то, что предельная разрешающая способность по порядку величины равна $\lambda/\sin \alpha$ или $\lambda f/a$.

В последние годы в технических журналах стали появляться статьи, оспаривающие тот факт, что рэлеевский предел представляет непреодолимый барьер для возможности разрешения изображений. Эта точка зрения недавно была отражена Гудманом в его книге по фурье-оптике [7]. Аргумент приводится следующий. Фурье-преобразование поля объекта конечных размеров является аналитической функцией. Даже когда из-за наличия апертуры (реальной или эффективной) в плоскости преобразования используется лишь часть этого фурье-преобразования, можно восстановить всю функцию. Можно полностью построить аналитическую функцию, если известна функция в малой окрестности точки. Это вытекает из того факта, что любая аналитическая функция может быть выражена степенным рядом по крайней мере в конечной области. Знание функции и всех ее производных в одной точке можно распространить на область, определяемую радиусом сходимости степенного ряда. Зная функцию и ее производные вблизи границы области сходимости, можно определить новые коэффициенты и получить новый степенной ряд, который расширит область определения функции. Таким образом,

по значениям аналитической функции в малой первоначальной области можно построить всю эту функцию. Этот процесс известен как аналитическое продолжение [38].

Критики критерия Рэлея утверждают, что, поскольку преобразование Фурье (аналитическая функция) изображения конечных размеров известно в ограниченной области, можно построить полное преобразование Фурье с помощью аналитического продолжения и, таким образом, по рэлеевскому пределу полностью восстановить изображение объекта с любой степенью точности.

Математически этот аргумент является строгим. Имеются и другие вполне строгие положения, которые, однако, невозможно реализовать на практике. Рассмотрим, например, сигнал, переданный удаленным радиопередатчиком. Игнорируя квантовую природу излучения, можно утверждать, что сигнал этого передатчика идет в пространстве произвольно далеко. Конечно, количество энергии, достигающей некоторой точки пространства, убывает с ростом расстояния между передатчиком и приемником. Однако решение задачи с помощью теории Максвелла дает гарантию, что в принципе сигнал уходит в пространство на любое расстояние. Единственная проблема — это обнаружение сигнала. Практически предел обнаружения сигнала связан с неизбежным наличием шума. Если достаточно далеко разнести передатчик и приемник, то можно достичь такого пункта, где мощность сигнала исчезнет в тепловом шуме. Возможно, конечно, выделить сигнал из шума посредством специальных методов. Однако утверждать, что можно преодолеть произвольное расстояние с помощью передатчика данной конечной мощности и сигнала с данной полосой, поскольку сигнал по крайней мере в принципе всегда есть, бессмысленно, так как наличие шума накладывает определенные пределы на возможности, при которых можно выделить сигнал. Ограничения приема сигналов в присутствии шума рассмотрены Шенноном [39].

Утверждение, что рэлеевское ограничение не имеет отношения к теоретическому восстановлению изображения объекта, имеет ту же природу. В этот вопрос внес ясность Торальдо ди Франчия. Мы рассмотрим его аргумент, который основан на привлечении некоторых функций, называемых вытянутыми сфероидальными функциями

ми. Свойства этих функций исследованы в ряде статей Ландау, Поллака, Слепьяна и Зонненблика [40, 41, 62, 63]. Интересующее нас свойство этих функций $\psi_i(u)$ заключается в том, что они образуют полную систему функций одновременно как в неограниченной области $-\infty \leq u \leq \infty$, так и в конечной области $-u_0/2 \leq u \leq u_0/2$. Это свойство выражается условиями ортогональности ¹⁾. Первое из них распространяется на неограниченную область

$$\int_{-\infty}^{\infty} \psi_i(u) \psi_j(u) du = \delta_{ij} \quad (4.6.11)$$

(δ_{ij} — символ Кронекера), а второе — на конечную область

$$\int_{-u_0/2}^{u_0/2} \psi_i(u) \psi_j(u) du = \lambda_i \delta_{ij}. \quad (4.6.12)$$

Функции ψ являются решениями интегрального уравнения

$$\lambda_i \psi_i(v) = \int_{-u_0/2}^{u_0/2} \frac{\sin v_0(v-u)}{\pi(v-u)} \psi_i(u) du. \quad (4.6.13)$$

Параметр λ_i есть собственное значение интегрального уравнения (4.6.13). Собственное значение λ_i , так же как и функции $\psi_i(u)$, зависит от константы c , определяемой соотношением

$$c = \frac{1}{2} v_0 u_0. \quad (4.6.14)$$

Величина интервала $(-u_0/2, u_0/2)$, внутри которого функции $\psi_i(u)$ ортогональны, определяет зависимость

¹⁾ Здесь допущена неточность. Если система функций полна в интервале $(-\infty, \infty)$, то она автоматически полна в любом конечном интервале. Необычное свойство сфероидальных функций заключается в том, что они оказываются ортогональными как в интервале $(-\infty, \infty)$, так и в некотором конечном интервале. При этом свойство ортогональности является самостоятельным свойством, не связанным с полнотой. Заметим, что в физике под полнотой системы функций понимается возможность разложения по системе этих функций, т. е. разложимость. — *Прим. ред.*

этих функций от параметра c . Для дальнейшего полезно знать величину параметра c . С помощью формул (4.3.28) и (4.3.29) можно представить константу c в виде

$$c = \frac{\pi}{\lambda f} a D. \quad (4.6.15)$$

Размер объекта D соответствует u_0 , а величина a есть полуширина апертуры. В разд. 2.2 было отмечено, что приближение Френеля в теории дифракции справедливо при пренебрежении в разложении (2.2.23) членами, которые после умножения на $k = 2\pi/\lambda$ (не следует путать длину волны λ с собственным значением λ_i) много меньше единицы. Потребуем

$$\frac{1}{8} k \frac{a^4}{f^3} = \frac{\pi}{4} \frac{a^4}{\lambda f^3} \ll 1, \quad (4.6.16)$$

чтобы гарантировать справедливость приближения Френеля. При $f = 1$ м и $\lambda = 0,5$ мкм необходимо иметь $a = 1,5$ см, чтобы выражение в левой части (4.6.16) равнялось приблизительно 0,1. Беря для простоты $a = D$, из формулы (4.6.15) получим

$$c = 1410. \quad (4.6.17)$$

Число c можно сделать еще больше даже при сравнительно малых апертурах. Между прочим, величина c представляет собой степень справедливости приближения Френеля в задачах дифракции. Приближение Френеля точнее при меньших значениях c . Параметр c дает порядок величины второго члена в выражении (2.2.23) (после умножения на $k = 2\pi/\lambda$), который можно записать в виде

$$N = \frac{1}{2} k \frac{a^2}{L} = \pi \frac{a^2}{\lambda L}. \quad (4.6.18)$$

Величина N называется числом Френеля. Здесь вместо f стоит L для обозначения более общего расстояния вдоль оптической оси. Выбор фокусного расстояния f в формуле (4.6.15) был продиктован обычной задачей, схематически изображенной на фиг. 4.6.1. При $a = D$ и $L = f$ параметр c совпадает с числом Френеля N . В рассмотренном примере получаем $N = 1410$.

После этого отступления в область справедливости приближения Френеля вернемся к рассмотрению собствен-

ных значений λ_i . Слепян и Зонненблик [41] предложили приближенное выражение для λ_i , которое справедливо для больших c и i :

$$\lambda_v = \frac{1}{1 + e^{\pi b}}, \quad (4.6.19a)$$

где

$$b = \frac{\frac{\pi}{2} v - c + \frac{\pi}{4}}{0,2886 + 2 \ln 2 + \frac{1}{2} \ln c}. \quad (4.6.19b)$$

Отсюда видно, что $\lambda_v \approx 1$ при $v \ll (2/\pi) c$, но $\lambda_v \ll 1$ при $v \gg (2/\pi) c$. Таким образом, переход из области, где λ_v близки к единице, в область, где λ_v спадают до весьма малых значений, происходит при $v = v_c$, где

$$v_c = \frac{2}{\pi} c. \quad (4.6.20)$$

Чтобы показать, насколько быстро изменяется собственное значение, рассмотрим пример. Для $v = 890$ получаем

$$\lambda_{890} = 0,96, \quad (4.6.21)$$

тогда как для $v = 920$ имеем

$$\lambda_{920} = 10^{-5}. \quad (4.6.22)$$

При меньших значениях c переход еще более крутой. С помощью таблиц Слепяна и Зонненблика находим, что для $c = 30$

$$\lambda_{19} = 0,356, \quad (4.6.23)$$

$$\lambda_{25} = 4,74 \cdot 10^{-6}, \quad (4.6.24)$$

$$\lambda_{30} = 1,32 \cdot 10^{-11}. \quad (4.6.25)$$

Теперь мы подготовлены к обсуждению аргумента Торальдо ди Франчия. Каждую функцию можно разложить в ряд по полной системе ортогональных функций ψ . Поле объекта, в частности, может быть представлено выражением

$$f(u) = \sum_{i=0}^{\infty} a_i \psi_i(u), \quad (4.6.26)$$

коэффициенты которого определяются формулой

$$a_i = \int_{-\infty}^{\infty} f(u) \psi_i(u) du. \quad (4.6.27)$$

Аналогичным образом можно представить выходное поле (поле изображения):

$$\bar{f}(-w) = \sum_{i=0}^{\infty} A_i \psi_i(w), \quad (4.6.28)$$

где

$$A_i = \int_{-\infty}^{\infty} \bar{f}(-w) \psi_i(w) dw. \quad (4.6.29)$$

Исходя из выражения (4.6.13), можно получить связь между коэффициентами разложений a_i и A_i . Подставляя выражение (4.6.26) в (4.6.3) и используя формулу (4.6.13), получаем

$$\bar{f}(-w) = \sum_{i=0}^{\infty} a_i \lambda_i \psi_i(w); \quad (4.6.30)$$

сравнение формул (4.6.28) и (4.6.30) позволяет установить, что

$$A_i = \lambda_i a_i. \quad (4.6.31)$$

Система, формирующая изображение, позволяет наблюдать поле изображения $\bar{f}(w)$, а поле объекта $f(u)$ не известно. Критики рэлеевского ограничения разрешающей способности изображения утверждают, что поле изображения можно в принципе построить из поля объекта независимо от того, насколько мала апертура системы, формирующей изображение. То, что это утверждение математически верно, можно видеть из соотношения (4.6.31). Поскольку известно поле изображения, то в принципе известны и все коэффициенты A_i . Следовательно, коэффициенты разложения поля объекта (4.6.26) получаются как

$$a_i = \frac{A_i}{\lambda_i}. \quad (4.6.32)$$

Подстановка выражения (4.6.32) в формулу (4.6.26) дает точное поле изображения независимо от размера апертуры

и длины волны света, применяемого в системе, формирующей изображение. Это доказывает, что с точки зрения математики не существует ограничения, обусловленного конечными размерами апертуры рассматриваемой системы.

Обратимся теперь к практической стороне вопроса и посмотрим, есть ли надежда использовать полученный математический результат. Для того чтобы иметь возможность всегда восстановить изображение, нужно предположить, что все коэффициенты A_i действительно могут быть измерены как угодно точно. Однако рассмотрение собственных значений λ_i показывает, что при $i \gg (2/\pi)$ с собственные значения становятся малыми. Это говорит о том, что коэффициенты разложения поля изображения должны быть много меньше соответствующих коэффициентов разложения поля объекта. Таким образом, мы сталкиваемся с задачей измерения весьма малых величин с очень большой точностью. Так как, согласно формуле (4.6.32), коэффициенты a_i получаются из A_i путем умножения на очень большое число, очевидно, что ошибка при попытке измерить A_i также возрастает из-за слишком большой величины $1/\lambda_i$. Для точного определения поля объекта требуется исключительно точно измерить коэффициенты разложения поля изображения. Таким образом, очевидно существование аналогии между задачей определения радиосигнала на некотором расстоянии от передатчика и задачей восстановления поля объекта по полю изображения. Невозможно получить изображение без мешающего действия какого-либо шума. Это может быть тепловой шум в фотодетекторе при зондировании поля изображения, в телевизионной трубке или даже в фотографической эмульсии. Общеизвестен факт присутствия шумов при электрическом детектировании. Фотографические эмульсии вызывают «шум» из-за конечного размера зерен кристаллов галоида серебра. Но даже если удалось бы сделать идеальную эмульсию, все равно невозможно получить размеры зерен меньше размера отдельной молекулы, т. е. шум в фотоэмульсиях также неизбежен, как и шум в электрических цепях. Следовательно, нет никакой надежды точно восстановить поле объекта, зная поле изображения. Некоторая информация всегда теряется, даже если в строго математическом смысле имеется возможность построить

поле объекта по полю изображения, несмотря на дифракцию. Это математическое положение имеет ограниченное применение в реальном мире. Только значительно более тщательное изучение может показать, возможно ли улучшить предельную разрешающую способность, определяемую обычно с помощью критерия Рэля. Связь между проведенным обсуждением и критерием Рэля не простая и не явная, даже когда возможно определить практическую предельную разрешающую способность на основе соотношения (4.6.32). Наша цель состояла в том, чтобы показать, что аргументы против истинного ограничения, предсказываемого критерием Рэля, имеют весьма сомнительную ценность, если они основаны на утверждении, что аналитическую функцию можно в принципе построить по ее значениям внутри конечной области. Если воздействие шума, препятствующее процессу обнаружения, не учитывается, то нельзя получать надежные сведения.

ЛИНЗОВЫЕ ВОЛНОВОДЫ

5.1. ВВЕДЕНИЕ

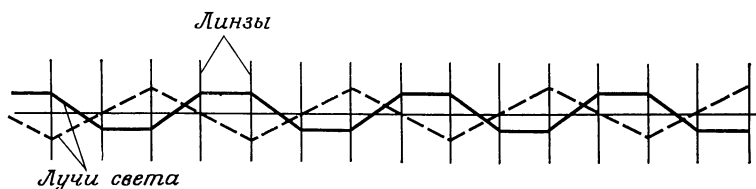
Параллельный пучок света нельзя передавать на произвольные расстояния без изменения его поперечных размеров (см. разд. 2.3, 2.4, 3.6). Строго говоря, параллельный пучок света возможен только в пределе для очень коротких длин волн или в случае бесконечно широкой апертуры (поперечного сечения) пучка. Оба эти условия не могут быть полностью реализованы. Однако для коротких длин волн света относительно легко получить большие (по сравнению с длиной волны) апертуры и, таким образом, воспроизвести достаточно параллельный пучок света. Если приемник света находится на не слишком большом расстоянии от источника, то практически весь переданный свет будет принят. Но если это расстояние окажется слишком большим, то весь передаваемый свет не будет принят. Например, когерентный пучок света, ограниченный круглой апертурой диаметром 30 см и передаваемый с Земли, будет на Луне засвечивать круг радиусом 1 км (без учета влияния атмосферы Земли).

Передача света между наземными станциями связи затруднена не только из-за дифракции света, но и из-за влияния земной атмосферы. Лазерный луч, который передается на несколько километров через земную атмосферу, приходит к приемнику с большими искажениями. Если такой луч принимается на экран, то его изображение создает впечатление движущегося пламени. Причиной такого явления являются случайные неоднородности (турбулентности) в атмосфере, которые приводят к отклонению и расхождению светового пучка, что ограничивает возможность параллельного распространения света. Если, кроме того, принять во внимание различные препятствия на пути пучка света, то становится ясным, что для передачи света

в атмосфере Земли необходима специальная направляющая система. Существует много различных способов передачи света. В этой книге будут рассмотрены лишь некоторые из них. Настоящая глава посвящена одному из наиболее перспективных способов передачи света с помощью линзового волновода, предложенного Губо [43] (Губо назвал его линзовым лучеводом). Как следует из названия, в таком волноводе для создания условий передачи света используются линзы. Мы будем исследовать линзовые волноводы с помощью методов как геометрической, так и волновой оптики. Кроме принципа работы регулярного прямолинейного волновода, будут рассмотрены также изогнутый линзовый волновод и волновод со статистическими нерегулярностями.

5.2. ЛУЧЕВАЯ ОПТИКА РЕГУЛЯРНОГО ЛИНЗОВОГО ВОЛНОВОДА

Схематически линзовый волновод изображен на фиг. 5.2.1. Расстояние между линзами равно двойному фокусному расстоянию, что соответствует особому случаю

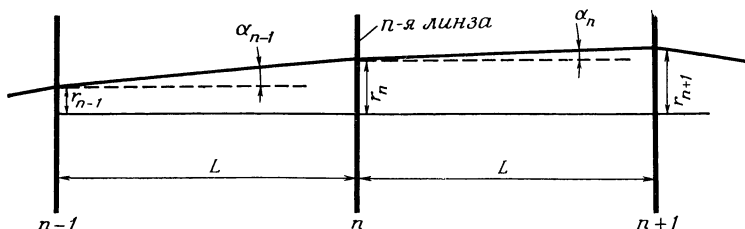


Ф и г. 5.2.1. Траектории лучей двух типов в конфокальном линзовом волноводе.

линзового волновода. Будут рассмотрены в дальнейшем также и другие варианты. Устройство, показанное на фиг. 5.2.1, называется конфокальным линзовым волноводом, так как фокусы соседних линз совпадают. Конфокальная геометрия приводит к чрезвычайно простой траектории луча. На фигуре показаны два луча, проходящие через волновод.

Анализ линзового волновода будет ограничен параксиальным приближением. Только в этом приближении уравнение луча можно решить аналитически. Работа

линзового волновода может быть понята с помощью квантовой теории световых лучей. В соответствии с принципом неопределенности (3.6.56) точное положение луча не известно, если лучи предполагаются параллельными. Для строго коллимированного пучка можно точно определить угол его наклона (в терминах квантовой теории световых лучей — импульс луча). Согласно принципу неопределенности $\Delta x = \infty$, если $\Delta p_x = 0$. Для того чтобы иметь



Ф и г. 5.2.2. Положение и углы наклона луча в линзовом волноводе.

пучок конечной протяженности по ширине (т. е. чтобы неопределенность его положения Δx имела конечную величину), необходимо допустить разброс импульсов лучей Δp_x . Линзовый волновод заставляет луч света отклоняться и сужаться при прохождении его от линзы к линзе. Элементарные лучи полного пучка никогда не имеют одинакового наклона, так что существует конечная неопределенность Δp_x лучевого «импульса». Это дает возможность определить положение луча внутри заданного интервала Δx без нарушения принципа неопределенности для световых лучей.

Для того чтобы иметь возможность определить траектории луча в линзовом волноводе, получим разностное уравнение для положения луча на каждой линзе [44]. Рассмотрим три линзы волновода (фиг. 5.2.2). Пусть расстояние от оптической оси луча в сечении n -й линзы равно r_n . Угол между лучом и направлением оси волновода обозначим через α_n . Все линзы имеют одинаковое фокусное расстояние f и расположены на расстоянии L друг от друга. Уравнение линзы в параксиальном приближении (4.2.4)

позволяет записать

$$\alpha_n - \alpha_{n-1} = -\frac{r_n}{f}. \quad (5.2.1)$$

Оба угла здесь положительные, что привело к различию в знаке в формулах (4.2.4) и (5.2.1). Положение луча в n -й линзе может быть определено по положению луча в $(n-1)$ -й линзе и углу наклона луча:

$$r_n = r_{n-1} + \alpha_{n-1}L. \quad (5.2.2)$$

Это равенство справедливо только для параксиальных лучей. Чтобы получить разностное уравнение для величин r_n , нужно из формул (5.2.1) и (5.2.2) исключить углы. Это можно сделать, используя формулу (5.2.2) с индексом $(n+1)$ вместо n :

$$r_{n+1} = r_n + \alpha_n L. \quad (5.2.3)$$

Вычитая (5.2.2) из (5.2.3) и используя равенство (5.2.1), получаем нужное уравнение

$$r_{n+1} + \left(\frac{L}{f} - 2\right)r_n + r_{n-1} = 0. \quad (5.2.4)$$

Решение разностного уравнения дает возможность определить положение луча на произвольной промежуточной линзе, если известно положение луча на первых двух линзах.

Решение уравнения (5.2.4) может быть найдено в виде

$$r_n = Ae^{in\Theta}. \quad (5.2.5)$$

Подстановка в (5.2.4) дает соотношение для определения величины Θ :

$$\cos \Theta = 1 - \frac{L}{2f}. \quad (5.2.6)$$

Выражения (5.2.5) и (5.2.6) являются решениями разностного уравнения (5.2.4). Другим решением является решение, полученное путем изменения знака перед Θ , так как выражение (5.2.6) справедливо как для положительных, так и для отрицательных Θ . Наиболее полное общее решение можно записать в виде

$$r_n = Ae^{in\Theta} + Be^{-in\Theta}. \quad (5.2.7)$$

Константы A и B могут быть выражены через положения луча на первых двух линзах:

$$r_1 = Ae^{i\Theta} + Be^{-i\Theta}, \quad (5.2.8)$$

$$r_2 = Ae^{2i\Theta} + Be^{-2i\Theta}. \quad (5.2.9)$$

Решением этих двух уравнений являются выражения

$$A = -\frac{e^{-i\Theta}}{2i} \frac{r_1 e^{-i\Theta} - r_2}{\sin \Theta}, \quad (5.2.10)$$

$$B = \frac{e^{i\Theta}}{2i} \frac{r_1 e^{i\Theta} - r_2}{\sin \Theta}. \quad (5.2.11)$$

Решение (5.2.7) можно теперь записать в конечном виде

$$r_n = -r_1 \frac{\sin(n-2)\Theta}{\sin \Theta} + r_2 \frac{\sin(n-1)\Theta}{\sin \Theta}. \quad (5.2.12)$$

Выражение (5.2.12) позволяет определить траекторию каждого луча в линзовом волноводе в параксиальном приближении. В частности, можно определить траектории двух лучей, показанных на фиг. 5.2.1. Для случая $L = 2f$ из формулы (5.2.6) получим, что $\cos \Theta = 0$, $\sin \Theta = 1$. Подстановка этих значений позволяет упростить выражение (5.2.12) для конфокального линзового волновода:

$$r_n = r_1 \sin n \frac{\pi}{2} - r_2 \cos n \frac{\pi}{2}. \quad (5.2.13)$$

Траектория луча, показанная сплошной линией на фиг. 5.2.1, имеет начальные условия

$$r_1 = r_0, \quad r_2 = -r_0, \quad (5.2.14)$$

поэтому можно записать

$$r_n = r_0 \left(\sin n \frac{\pi}{2} + \cos n \frac{\pi}{2} \right). \quad (5.2.15)$$

Легко видеть, что выражение (5.2.15) описывает траекторию, соответствующую сплошной линии на фиг. 5.2.1.

Траектория луча, показанная пунктирной линией, имеет начальные условия

$$r_1 = -r_0, \quad r_2 = 0, \quad (5.2.16)$$

так что эта траектория в конфокальном волноводе описывается выражением

$$r_n = -r_0 \sin n \frac{\pi}{2}. \quad (5.2.17)$$

Нетрудно проверить, что выражение (5.2.17) описывает траекторию луча, показанную на фиг. 5.2.1 пунктирной линией.

Наиболее важным результатом проведенного анализа является выражение (5.2.6), так как оно позволяет определить область изменения отношения L/f , которая допустима для линзовых волноводов. Из решения (5.2.12) видно, что лучи в волноводе могут существовать только при действительных значениях параметра Θ . Если Θ — комплексная величина, то тригонометрические функции становятся гиперболическими и лучи будут неограниченно отклоняться от оси. Таким образом, направляемые лучи не могут существовать в линзовом волноводе. Поскольку величина синусоидальной функции лежит между значениями -1 и 1 , то мы получаем допустимую область изменения отношения L/f в виде

$$0 \leq \frac{L}{f} \leq 4. \quad (5.2.18)$$

При значениях L/f , лежащих вне этой области, лучи в волноводе не распространяются. Если раздвинуть линзы на расстояние, хотя бы на немного превышающее $L = 4f$, то получим волновод, амплитуда отклонения луча в котором будет увеличиваться и в конце концов достигнет края линз, а затем области вне линз. В результате луч выйдет из волновода.

Рассмотрим этот случай более подробно. При $L > 4f$ выражение (5.2.6) становится отрицательным и больше единицы по абсолютной величине. Решение возможно только в форме

$$\Theta = i\theta' + \pi, \quad (5.2.19)$$

где θ' — действительная величина. При $L > 4f$ из формулы (5.2.6) следует

$$\operatorname{ch} \theta' = \frac{L}{2f} - 1 \quad \text{для} \quad \frac{L}{f} > 4. \quad (5.2.20)$$

Этот результат показывает, что имеются формальные решения вне области, где возможно существование направляемых волн. Уравнение соответствующих траекторий получим из формул (5.2.12) и (5.2.19):

$$r_n = (-1)^n \left[r_1 \frac{\operatorname{sh} (n-2) \theta'}{\operatorname{sh} \theta'} + r_2 \frac{\operatorname{sh} (n-1) \theta'}{\operatorname{sh} \theta'} \right]. \quad (5.2.21)$$

Так как гиперболический синус не осциллирует, а растет монотонно с увеличением аргумента, то выражение (5.2.21) описывает лучи, которые осциллируют около оптической оси лишь за счет коэффициента $(-1)^n$, при этом амплитуда отклонения безгранично растет при распространении лучей вдоль линзового волновода.

Чтобы сравнить режим работы волновода в области, где лучи поддерживаются волноводом, с режимом, когда лучи не поддерживаются волноводом, запишем выражение (5.2.12) для $\Theta = \pi$ в виде

$$r_n = (-1)^n [(n-2)r_1 + (n-1)r_2]. \quad (5.2.22)$$

Это выражение также получается из уравнения для лучей, не поддерживаемых волноводом (5.2.21). Переход между двумя этими режимами работы, таким образом, непрерывен. Критическая граница области с $\Theta = \pi$, или $\theta' = 0$, интересна тем, что для нее возможны оба режима работы волновода. Будут ли распространяться лучи в волноводе или вне его, зависит только от начальных условий. Для общего случая задания величин r_1 и r_2 мы видим, что амплитуда растет линейно, так что луч выходит за линзы. Однако в случае, когда $r_1 = -r_2 = r_0$, мы еще получаем направляемый луч

$$r_n = -(-1)^n r_0. \quad (5.2.23)$$

Анализ поведения лучей в линзовом волноводе дает достаточно много информации о возможных областях существования направляемых и ненаправляемых лучей или волн.

Другим интересным примером линзового волновода является последовательность линз, подобная показанной на фиг. 5.2.1, но отличающаяся тем, что фокусное расстояние линз изменяется от линзы к линзе. Пусть четные линзы имеют фокусное расстояние f_1 , а нечетные линзы — f_2 . Можно даже предположить случай, когда f_1 или f_2 (но не оба) становятся отрицательными, т. е. фокусирующие линзы чередуются с расфокусирующими. Мы увидим, что направляемые лучи могут существовать даже и в этом случае. Это так называемый случай с переменной фокусировкой [45, 46].

Уравнение луча можно получить из фиг. 5.2.3. Используя параксиальное уравнение линзы, подобное (5.2.1),

находим

$$\alpha_{2n} - \alpha_{2n-1} = -\frac{r_n}{f_1} \quad (5.2.24)$$

и

$$\alpha_{2n+1} - \alpha_{2n} = -\frac{\rho_{n+1}}{f_2}. \quad (5.2.25)$$

Положения лучей на смежных линзах опять связаны соотношением (5.2.2):

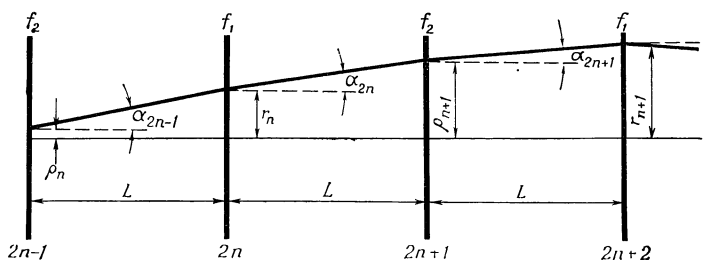
$$r_n = \rho_n + \alpha_{2n-1}L, \quad (5.2.26)$$

$$\rho_{n+1} = r_n + \alpha_{2n}L. \quad (5.2.27)$$

Вычитая (5.2.27) из (5.2.26), получим с учетом (5.2.24)

$$\rho_{n+1} + \rho_n = \left(2 - \frac{L}{f_1}\right) r_n. \quad (5.2.28)$$

Далее, заменив в формуле (5.2.26) n на $n + 1$ и вычитая



Ф и г. 5.2.3. Определение углов наклона и положений луча для линзового волновода с чередующимися линзами с различными фокусными расстояниями.

из полученного уравнения (5.2.27), найдем с помощью формулы (5.2.25)

$$r_{n+1} + r_n = \left(2 - \frac{L}{f_2}\right) \rho_{n+1}. \quad (5.2.29)$$

Заменим в этом соотношении n на $n - 1$ и прибавим полученное выражение к (5.2.29). Используя соотношение (5.2.28), чтобы исключить ρ , получим в результате уравнение, содержащее только r :

$$r_{n+1} + \left[2 - \left(2 - \frac{L}{f_1}\right) \left(2 - \frac{L}{f_2}\right)\right] r_n + r_{n-1} = 0. \quad (5.2.30)$$

Аналогично, заменив в (5.2.28) n на $n + 1$, прибавив полученное выражение к исходному и используя (5.2.29) для исключения r , получим уравнение, содержащее только ρ :

$$\rho_{n+1} + \left[2 - \left(2 - \frac{L}{f_1} \right) \left(2 - \frac{L}{f_2} \right) \right] \rho_n + \rho_{n-1} = 0. \quad (5.2.31)$$

Уравнения для параксиальных лучей (5.2.30) и (5.2.31) подобны по форме уравнению (5.2.4) простого линзового волновода. Решения получим непосредственно из формул (5.2.5) и (5.2.6)

$$r_n = A e^{in\Delta} + B e^{-in\Delta} \quad (5.2.32)$$

и

$$\rho_n = C e^{in\Delta} + D e^{-in\Delta} \quad (5.2.33)$$

с определением параметра Δ как

$$\cos \Delta = -\frac{1}{2} \left[2 - \left(2 - \frac{L}{f_1} \right) \left(2 - \frac{L}{f_2} \right) \right]. \quad (5.2.34)$$

До обсуждения этого решения проверим, справедливы ли полученные результаты в случае $f_1 = f_2$:

$$\cos \Delta = 2 \left(\frac{L}{2f} - 1 \right)^2 - 1 = 2 \cos^2 \Theta - 1,$$

или

$$\cos \Delta = \cos 2\Theta. \quad (5.2.35)$$

Это позволяет записать для $f_1 = f_2$

$$\begin{aligned} r_n &= A e^{i2n\Theta} + B e^{-i2n\Theta}, \\ \rho_n &= C e^{i\Theta} e^{i(2n-1)\Theta} + D e^{-i\Theta} e^{-i(2n-1)\Theta}. \end{aligned} \quad (5.2.36)$$

Следует помнить, что переменные r_n и ρ_n отличаются от тех, которые использовались при описании простого линзового волновода. Из фиг. 5.2.3 видно, что r_n соответствует r_{2n} , а ρ_n соответствует r_{2n-1} . Сравнение формул (5.2.7) и (5.2.36) показывает, что случай $f_1 = f_2$ сводится к предыдущему случаю простого линзового волновода с $C e^{i\Theta} = A$ и $D e^{-i\Theta} = B$.

Формула (5.2.32) формально идентична формулам (5.2.7) и (5.2.12), так что можно непосредственно записать

$$r_n = -r_1 \frac{\sin(n-2)\Delta}{\sin \Delta} + r_2 \frac{\sin(n-1)\Delta}{\sin \Delta}. \quad (5.2.37)$$

Аналогично для ρ_n имеем

$$\rho_n = -\rho_1 \frac{\sin(n-2)\Delta}{\sin\Delta} + \rho_2 \frac{\sin(n-1)\Delta}{\sin\Delta}. \quad (5.2.38)$$

Рассмотрим линзовый волновод, состоящий из ряда линз с различными (чередующимися) фокусными расстояниями. Уравнение (5.2.37) описывает положение луча в четных линзах при условии, что известно его положение во второй и четвертой линзах (r_1 и r_2). Аналогично уравнение (5.2.38) описывает положение луча в нечетных линзах при условии, что известно его положение в первой и третьей линзах. Траектории луча через четные и нечетные линзы не зависят друг от друга. Положения луча в четных и нечетных линзах связаны формулами (5.2.28) и (5.2.29). Таким образом, можно предсказать положение луча в любой линзе волновода.

Остается изучить условие, при котором возможна передача лучей [46]. Используя обозначения

$$x = 1 - \frac{L}{2f_1} \quad (5.2.39)$$

и

$$y = 1 - \frac{L}{2f_2}, \quad (5.2.40)$$

можно записать равенство (5.2.34) в виде

$$\cos\Delta = 2xy - 1. \quad (5.2.41)$$

Косинусоидальная функция не может быть больше 1 по абсолютной величине. Чтобы определить границы изменения допустимых значений x и y , положим $\cos\Delta = 1$; тогда получим

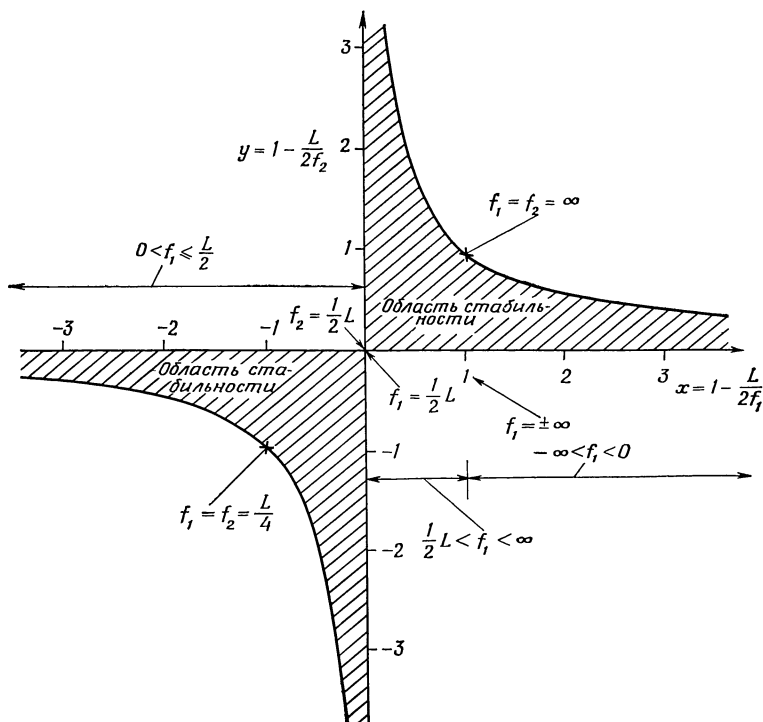
$$xy = 1. \quad (5.2.42)$$

При $\cos\Delta = -1$ из формулы (5.2.41) имеем

$$xy = 0. \quad (5.2.43)$$

Границы допустимых значений величин x и y показаны графически на фиг. 5.2.4. Эти гиперболы определяются формулой (5.2.42). Уравнение (5.2.43) устанавливает, что оси x и y являются граничными линиями. Допустимые значения x и y лежат в заштрихованной области на фигуре 5.2.4.

Некоторые точки на этой диаграмме имеют практический интерес. В точке $x = 1, y = 1$, например, линзовый волновод не работает, поскольку оба типа линз имеют



Фиг. 5.2.4. Области стабильности линзового волновода с чередующимися линзами с различными фокусными расстояниями или лазерного резонатора с неодинаковыми зеркалами. Вне заштрихованной области волновод или резонатор имеют большие потери [46].

бесконечное фокусное расстояние. Точка $x = -1, y = -1$ является пределом $f = L/4$, соответствующим простому линзовому волноводу с одинаковыми линзами. Фокусное расстояние этих линз принимает наименьшее значение, при котором еще существуют направляемые лучи. На фиг. 5.2.4 вдоль оси x показана область изменения фокусных расстояний четных линз (f_1). Точка $x = 1$ соответ-

ствуется фокусному расстоянию, равному ∞ . Справа от этой точки расположена область отрицательных фокусных расстояний. Это область попеременного фокусирования и расфокусирования: $f_1 < 0$ и $f_2 > 0$. Подобные области существуют, конечно, и вдоль оси y . Отрицательная ось x принадлежит области от $f_1 = 0$ ($x = -\infty$) до $f_1 = L/2$ ($x = 0$).

Точка $x = y = 0$ заслуживает особого внимания. Как показано на фиг. 5.2.4, она соответствует конфокальному линзовому волноводу с $f_1 = f_2 = L/2$. При изучении свойств простого линзового волновода конфокальная геометрия рассматривалась только из-за простоты траектории луча. Исследование линзовых волноводов, в которых используется два типа линз, представляет волновод в совершенно новом свете. Конфокальный линзовый волновод лежит на границе между областями стабильных и нестабильных (расходящихся) траекторий луча. Этот факт не очень важен для реальных линзовых волноводов. Весьма мала вероятность того, что фокусное расстояние, которое имеет случайные отклонения от линзы к линзе, будет изменяться так, чтобы настоящая теория стала применимой. Однако результаты, полученные для линзовых волноводов, имеют очень важные применения в теории лазерных резонаторов. Конфокальный лазерный резонатор также описывается точкой $x = 0$, $y = 0$, и нестабильность, определяемая сдвигом расстояний его зеркал от этой точки, имеет существенное значение для лазерного резонатора.

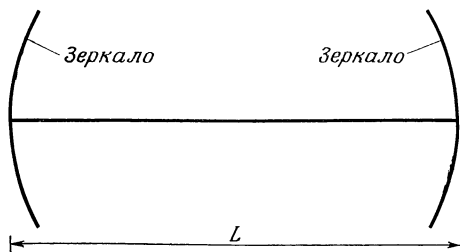
5.3. ЛАЗЕРНЫЕ РЕЗОНАТОРЫ [47—49]

На фиг. 5.3.1 схематически изображен типичный лазерный резонатор. Два зеркала с большой отражательной способностью расположены напротив друг друга. Обычно эти зеркала вогнутые, причем они могут иметь как одинаковую, так и различную кривизну. Опишем работу такого лазерного резонатора в терминах линзового волновода с учетом того факта, что лазерный резонатор может быть представлен так, как показано на фиг. 5.3.2.

Вогнутое зеркало оптически эквивалентно системе из одного плоского зеркала и линзы, расположенной непо-

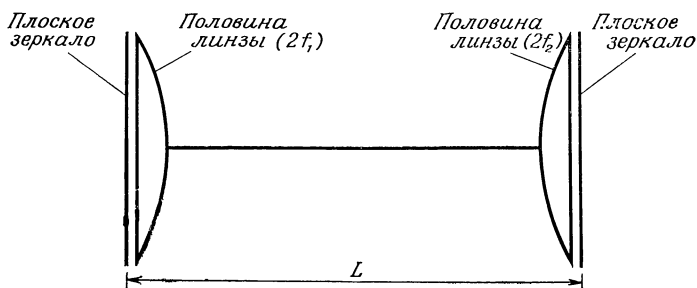
средственно перед ним. Резонаторы, показанные на фигурах 5.3.1 и 5.3.2, оптически эквивалентны.

При рассмотрении волновой оптики линзовых волноводов мы увидим, что в этих структурах существуют моды. Мода — это распределение поля, которое повторяет само



Ф и г. 5.3.1. Схематическое изображение лазерного резонатора.

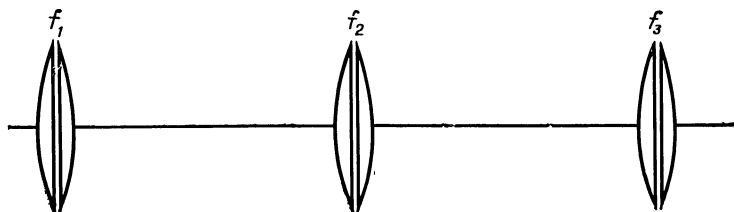
себя на каждой линзе, за исключением, может быть, фазы. Мода, распространяющаяся в линзовом волноводе, полностью эквивалентна моде лазерного резонатора при



Ф и г. 5.3.2. Эквивалентная система лазерного резонатора. Изогнутые зеркала заменены плоскими зеркалами с линзами.

условии, что можно пренебречь потерями в каждом устройстве. Рассмотрим резонансную моду резонатора, представленного на фиг. 5.3.1 или 5.3.2. Для простоты начнем рассмотрение с середины резонатора. Стоячая волна поля резонансной моды может быть разложена на две бегущие волны, распространяющиеся в противоположных направлениях. Следуя за одной из этих волн, бегущей вправо,

можно наблюдать, как она распространяется в направлении к зеркалу, отражается от него, идет в противоположном направлении к другому зеркалу. Рассматривая получающееся распределение поля, можно сказать, что резонатор точно соответствует линзовому волноводу. В самом деле, система на фиг. 5.3.2 преобразуется в линзовый волновод, если не учитывать того, что волна меняет свое направление в плоскости зеркала, и считать, что она распространяется без отражения при условии, что непосредственно за первой стоит другая идентичная линза.



Фиг. 5.3.3. Развернутый вдоль оси лазерный резонатор, иллюстрирующий эквивалентность между лазерными резонаторами и линзовыми волноводами.

Реальная отраженная волна пересекает каждую линзу, как видно из фиг. 5.3.2, дважды. Волна, приближающаяся к системе линза — зеркало, проходит через линзу, отражается от плоского зеркала и проходит еще раз через линзу. Эта волна распространяется так же, как в структуре, показанной на фиг. 5.3.3. Линзы, показанные на фиг. 5.3.2, имеют фокусные расстояния $2f_1$ и $2f_2$. Но поскольку эти линзы пересекаются волной дважды, они действуют подобно двум близко расположенным линзам, которые в сумме имеют фокусные расстояния f_1 или f_2 . Резонатор в развернутом виде показан на фиг. 5.3.3; он подобен линзовому волноводу.

До сих пор внимание концентрировалось на одной из бегущих волн стоячей волны резонатора. Другая волна претерпевает те же изменения, но движется в противоположном направлении. Обе бегущие волны накладываются друг на друга в каждой точке пространства, образуя в резонаторе стоячую волну. Конечно, необходимо, чтобы

волны точно повторялись при движении туда и обратно, создавая внутри резонатора интерференционную картину. Это требование приводит к тому, что резонатор может работать только на ряде дискретных частот, зависящих от его длины. В этом состоит отличие поля резонатора от моды, распространяющейся в соответствующем линзовом волноводе. Линзовый волновод используется в непрерывном спектре частот, тогда как резонатор может работать только на определенных дискретных частотах. Однако линзовый волновод способен поддерживать моды, как и резонатор. Картина поля идентична картине в соответствующем резонаторе. Сходство линзового волновода и резонатора лазера обнаруживается еще сильнее, если рассмотреть задачи методами волновой оптики, как это сделано в разд. 5.6 и 6.6. Резонатор лазера может быть описан интегральным уравнением. Итерационное решение этого уравнения эквивалентно процессу распространения волны по линзовому волноводу.

Установив эквивалентность лазерных резонаторов и линзовых волноводов, можно использовать результаты рассмотрения линзового волновода для случая резонатора.

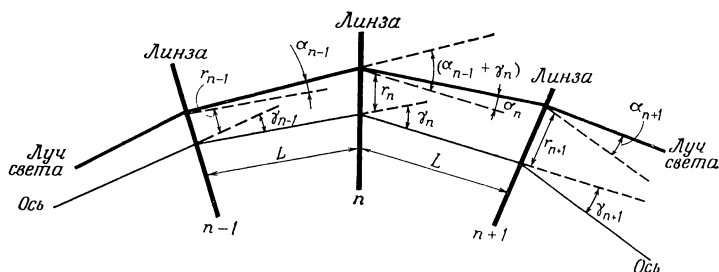
Мы видели на фиг. 5.2.4, что конфокальный линзовый волновод соответствует точке $x = y = 0$. Эта особая точка находится между стабильной и нестабильной областями структуры. Ранее отмечалось, что эта нестабильность не является очень существенной для реального линзового волновода, если даже существует опасность попадания фокусных расстояний некоторых линз в запретную зону (фиг. 5.2.4). Рассмотрим теперь лазерный резонатор, сконструированный для работы в режиме конфокального резонатора. В этом случае резонатор будет соответствовать конфокальному линзовому волноводу. Все множество линз волновода соответствует двум зеркалам резонатора, так что любое случайное отклонение от равенства фокусных расстояний зеркал может вызвать работу устройства либо в режиме внутри заштрихованной области (фиг. 5.2.4), либо вне ее. В последнем случае потери в резонаторе будут очень быстро возрастать и работа будет нарушена. По этим соображениям обычно лазерные резонаторы не конструируются с конфокальными системами, а рассчитываются для работы в точке, соответствующей заштрихованной

области (фиг. 5.2.4), чтобы устранить возможность работы в области больших потерь, обусловленных случайным отклонением зеркал от строгой идентичности. Для некоторых специальных применений в лазерах с высоким усилением предлагается работать в области неустойчивости резонатора для обеспечения требуемой селекции мод.

Мы еще вернемся к обсуждению лазерных резонаторов при рассмотрении линзовых волноводов. Однако представление об эквивалентности линзовых волноводов и лазерных резонаторов может быть получено непосредственно из геометрической оптики.

5.4. ЛИНЗОВЫЙ ВОЛНОВОД С ИСКРИВЛЕННОЙ ОСЬЮ

В предыдущих разделах были рассмотрены линзовые волноводы, оси которых представляют собой идеальные прямые линии. Разумеется, это является идеализацией,



Ф и г. 5.4.1. Линзовый волновод с изогнутой осью.

не реализуемой на практике. Одним из важных свойств линзовых волноводов является возможность передачи света по искривленным путям. Ниже изложена лучевая теория линзовых волноводов с произвольно искривленными осями [50, 51].

Геометрия участка линзового волновода с искривленной осью показана на фиг. 5.4.1. В этом разделе мы предполагаем, что линзы волновода одинаковы. Снова обозначим через r_n положение луча на линзе, измеренное относительно ее центра. Прямые линии, соединяющие центры линз, образуют ломаную ось структуры. Угол γ_n указывает изменение направления оси волновода у n -й линзы.

Сложение уравнений системы приводит к соотношению

$$a_n e^{i\theta} - a_{n-1} e^{-i\theta} = L \sum_{v=2}^{n-1} \gamma_v e^{-iv\theta}. \quad (5.4.5)$$

Пробное решение (5.4.4) позволяет привести неоднородное разностное уравнение второго порядка (5.4.3) к неоднородному разностному уравнению первого порядка. Такое понижение порядка уравнения совершенно аналогично методу вариации произвольных постоянных в дифференциальных уравнениях. Уравнение (5.4.5) может быть записано в виде системы

$$\begin{array}{lcl} a_3 e^{i\theta} & = L \sum_{v=2}^2 \gamma_v e^{-iv\theta} & \\ a_4 e^{i\theta} - a_3 e^{-i\theta} & = L \sum_{v=2}^3 \gamma_v e^{-iv\theta} & e^{2i\theta} \\ a_5 e^{i\theta} - a_4 e^{-i\theta} & = L \sum_{v=2}^4 \gamma_v e^{-iv\theta} & e^{4i\theta} \\ \dots & & \dots \\ a_n e^{i\theta} - a_{n-1} e^{-i\theta} & = L \sum_{v=2}^{n-1} \gamma_v e^{-iv\theta} & e^{i2(n-3)\theta} \end{array}$$

Очевидно, что все члены, кроме a_n в левых частях этих уравнений, взаимно уничтожатся, если умножить каждое уравнение на экспоненциальный множитель, стоящий правее вертикальной линии, и сложить все уравнения. Эта процедура приведет к уравнению

$$a_n e^{i(2n-5)\theta} = L \sum_{v=2}^{n-1} \gamma_v e^{-iv\theta} \sum_{\mu=v+1}^n e^{i2(\mu-3)\theta}. \quad (5.4.6)$$

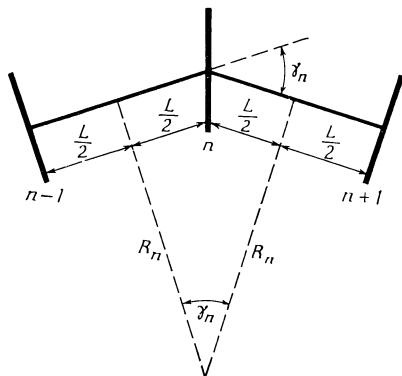
После выполнения суммирования и подстановки результата в уравнение (5.4.4) получаем частное решение неоднородного разностного уравнения второго порядка в следующем виде:

$$r_n = \frac{L}{\sin \theta} \sum_{v=2}^{n-1} \gamma_v \sin(n-v)\theta \quad \text{для } n \geq 3. \quad (5.4.7)$$

Полное решение неоднородного разностного уравнения (5.4.3) является суммой решения (5.2.12) однородного уравнения и частного решения (5.4.7) неоднородного уравнения:

$$r_n = \frac{1}{\sin \Theta} \left[-r_1 \sin (n-2) \Theta + r_2 \sin (n-1) \Theta + \right. \\ \left. + L \sum_{v=2}^{n-1} \gamma_v \sin (n-v) \Theta \right] \text{ для } n \geq 3. \quad (5.4.8)$$

Это решение для траектории луча в линзовом волноводе с искривленной осью получено в параксиальном приближении. Однако оно справедливо для волноводов, оси



Ф и г. 5.4.2. Определение радиуса кривизны оси изогнутого линзового волновода.

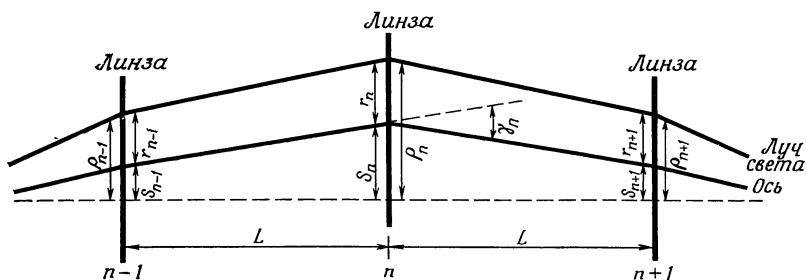
которых изменяются произвольным образом, но так, что угол γ_n остается малым.

Угол γ_n может быть использован для получения радиуса кривизны линзового волновода. На фиг. 5.4.2 показано, как угол γ_n связан с радиусом кривизны оси волновода. Эта связь имеет следующий вид:

$$\gamma_n = \frac{L}{R_n}. \quad (5.4.9)$$

Разумеется, это соотношение справедливо лишь в параксиальном приближении.

Другой способ описания линзового волновода с искривленной осью связан с рассмотрением положений как луча, так и центров линз. Это особенно полезно при рассмотрении линзовых волноводов, оси которых незначительно отклоняются от прямой линии. Подобная ситуация изображена на фиг. 5.4.3, где смещение линз сильно увеличено. Положение луча рассматривается относительно прямой линии отсчета, а не центра каждой линзы. Расстояние



Фиг. 5.4.3. Другое описание линзового волновода с искривленной осью. Положения центров линз и луча рассматриваются относительно пунктирной линии отсчета.

от линии отсчета до центра линзы обозначим через S_n . Тогда положение луча ρ_n определяется равенством

$$\rho_n = r_n + S_n. \quad (5.4.10)$$

Соотношение между расстоянием S_n и углом γ_n может быть получено следующим образом. Рассмотрим треугольник, образованный участками прямых, соединяющих центры линз с номерами $n-1$, n и $n+1$, и третьей прямой линией (на фиг. 5.4.3 не показана), соединяющей центр $(n-1)$ -й линзы с центром $(n+1)$ -й линзы. Внутренние углы этого треугольника приблизительно равны $(S_n - S_{n-1})/L$ у линзы $n-1$, $\pi - \gamma_n$ у линзы n и $(S_n - S_{n+1})/L$ у линзы $n+1$. Так как сумма этих углов должна равняться π , то получаем следующее выражение:

$$\gamma_n = \frac{1}{L} (2S_n - S_{n-1} - S_{n+1}). \quad (5.4.11)$$

Подстановка выражений (5.4.10) и (5.4.11) в (5.4.7) дает

$$\rho_n = S_n + \frac{1}{\sin \Theta} \sum_{v=2}^{n-1} (2S_v - S_{v-1} - S_{v+1}) \sin(n-v)\Theta. \quad (5.4.12)$$

Перегруппировка членов в сумме приводит к выражению

$$\begin{aligned} \rho_n = \frac{1}{\sin \Theta} \{ & -S_1 \sin(n-2)\Theta + S_2 \sin(n-1)\Theta + \\ & + \sum_{v=2}^{n-1} S_v [2 \sin(n-v)\Theta - \sin(n-v-1)\Theta - \sin(n-v+1)\Theta] \}. \end{aligned} \quad (5.4.13)$$

Нашей задачей является нахождение частного решения неоднородного разностного уравнения, которое соответствует этой модифицированной задаче. Первые два члена до знака суммы в формуле (5.4.13) являются решениями однородного разностного уравнения и могут быть использованы в качестве решения, которое должно быть добавлено для получения полного решения неоднородного разностного уравнения. Опустив первые два члена в формуле (5.4.13) и используя формулу для суммы синусов, можно записать частное решение неоднородной задачи в виде

$$\rho_n = \frac{2(1 - \cos \Theta)}{\sin \Theta} \sum_{v=2}^{n-1} S_v \sin(n-v)\Theta, \quad n \geq 3. \quad (5.4.14)$$

Прибавим решение однородного уравнения по аналогии с методом получения соотношения (5.4.8). В итоге получим полное решение задачи

$$\begin{aligned} \rho_n = \frac{1}{\sin \Theta} \Big[& -\rho_1 \sin(n-2)\Theta + \rho_2 \sin(n-1)\Theta + \\ & + \frac{L}{f} \sum_{v=2}^{n-1} S_v \sin(n-v)\Theta \Big]. \end{aligned} \quad (5.4.15)$$

Формула (5.2.6) использовалась для выражения $\cos \Theta$ через L/f . Наша вторая формулировка задачи линзового волновода с искривленной осью имеет тот недостаток, что она непригодна для описания волноводов, оси которых

существенно отклоняются от начального направления. В этих случаях должна использоваться формула (5.4.8). Однако выражение (5.4.15) является полезным при статистическом анализе траекторий лучей в линзовом волноводе с малыми случайными смещениями линз.

Теперь мы в состоянии предсказать траекторию любого луча в линзовом волноводе с произвольной деформацией оси и с достаточно большим радиусом кривизны, чтобы выполнялись приближения, которые были использованы при получении соотношения (5.4.8).

Существует несколько простых случаев, которые легко могут быть описаны нашей теорией. Рассмотрим линзовый волновод, ось которого свернута в кольцо. Можно опустить индекс n у радиуса кривизны и вместо выражения (5.4.9) записать

$$\gamma_n = \frac{L}{R} . \quad (5.4.16)$$

Суммирование в формуле (5.4.8) легко осуществляется; в результате имеем:

$$r_n = \frac{1}{\sin \Theta} \left\{ -r_1 \sin(n-2)\Theta + r_2 \sin(n-1)\Theta - \right. \\ \left. - \frac{L^2}{2R} \left[\frac{\cos \frac{\Theta}{2}}{\sin \frac{\Theta}{2}} \{ \cos(n-1)\Theta - 1 \} + \sin(n-1)\Theta \right] \right\} . \quad (5.4.17)$$

Единственным членом, не зависящим от n , является

$$r_c = \frac{L^2}{2R} \frac{\cos \frac{\Theta}{2}}{\sin \Theta \sin \frac{\Theta}{2}} . \quad (5.4.18)$$

Учет выражения (5.2.6) преобразует этот член к простому виду

$$r_c = \frac{Lf}{R} . \quad (5.4.19)$$

Легко можно показать, что $r_n = r_c$, если в формуле (5.4.17) положить $r_1 = r_2 = r_c$. Но вместо этого докажем более общее положение, что траектория луча в кольцевом линзовом волноводе подобна траектории луча в идеально

прямом волноводе с той лишь разницей, что луч осциллирует не около оси волновода, а около линии, которая параллельна оси и смещена от нее на величину r_c .

Для доказательства этого утверждения введем новую переменную

$$w_n = r_n - r_c, \quad (5.4.20)$$

которая представляет собой положение луча по отношению к линии, смещенной от оси волновода на расстояние r_c . Подстановка выражения (5.4.20) в (5.4.17) дает

$$w_n = \frac{1}{\sin \Theta} \left\{ -w_1 \sin (n-2) \Theta + w_2 \sin (n-1) \Theta + \right. \\ \left. + \frac{L^2}{2R} \left[\frac{\cos \frac{\Theta}{2}}{\sin \Theta \sin \frac{\Theta}{2}} (1 - \cos \Theta) - 1 \right] \sin (n-1) \Theta \right\}. \quad (5.4.21)$$

Здесь $\sin (n-2) \Theta$ заменен выражениями, содержащими $\sin (n-1) \Theta$ и $\cos (n-1) \Theta$. Член, содержащий $\cos (n-1) \Theta$, отсутствует. С помощью тригонометрических формул сложения легко показать, что выражение в квадратных скобках в (5.4.21) равно нулю. Следовательно, преобразование (5.4.20) приводит (5.4.17) к более простому выражению

$$w_n = \frac{1}{\sin \Theta} [-w_1 \sin (n-2) \Theta + w_2 \sin (n-1) \Theta]. \quad (5.4.22)$$

Это выражение, описывающее траекторию луча в линзовом волноводе, который изогнут по дуге окружности, с формальной точки зрения аналогично выражению (5.2.12). Луч внутри такого волновода проходит так же, как внутри прямого оптического волновода, если траектория луча определяется не по отношению к оси волновода, а по отношению к линии, смещенной относительно оси волновода на величину r_c .

Такой результат заслуживает внимания. Он подтверждает, что линзовый волновод способен передавать пучки света по изогнутым траекториям. Кроме того, он показывает, что свет может проходить в изогнутых линзовых волноводах без осцилляций ($w_1 = w_2 = 0$), если он следует

по траектории $w_n = 0$ или $r_n = r_c$, где r_c определяется формулой (5.4.19). Такая траектория луча смещается сильнее для меньших значений радиуса кривизны R , т. е. для более резко изогнутых волноводов. Очевидно, что расположение линз на небольших расстояниях друг от друга (L мало) позволяет использовать более резкие изгибы. Способность линзовых волноводов передавать световые лучи по искривленным траекториям ограничивается только размерами линз. Размах колебания траектории луча, разумеется, не может быть больше радиуса a линз. Действительно, величина a должна быть значительно больше, чем r_c , так как луч должен иметь запас пространства, чтобы не выйти за пределы линзы. Траектории лучей в линзовых волноводах, имеющих ось более сложной формы, могут быть получены из уравнения (5.4.8), но сумма, входящая в это уравнение, обычно не может быть вычислена в конечном виде. При вычислении траектории луча на вычислительной машине можно использовать уравнение (5.4.8) или разностное уравнение (5.4.3).

Другим отклонением от прямолинейности, которое легко может быть исследовано с помощью рассмотренных выше аналитических выражений, является обычный излом оси прямого оптического волновода. В этом случае

$$\gamma_v = \gamma \delta_{v\mu}, \quad (5.4.23)$$

что означает, что все углы γ_n , за исключением одного, равны нулю. Если луч начинается на оси ($r_1 = r_2 = 0$), уравнение траектории луча в точках за изломом получается из (5.4.8) в виде

$$r_n = \frac{\gamma L}{\sin \Theta} \sin (n - \mu) \Theta. \quad (5.4.24)$$

Излом на угол γ вызывает вблизи излома оси колебания луча относительно оси волновода с амплитудой $\gamma L / \sin \Theta$.

В качестве последнего примера рассмотрим случай прямого оптического волновода с одной смещенной линзой. В этом случае имеем

$$S_v = S \delta_{vu} \quad (5.4.25)$$

и для луча, находящегося вначале на оси ($\rho_1 = \rho_2 = 0$),

из формулы (5.4.15) получим

$$\rho_n = \frac{LS}{f \sin \Theta} \sin(n - \mu) \Theta. \quad (5.4.26)$$

Одна смещенная линза, так же как и излом оси, вызывает колебания пучка относительно оси с амплитудой $SL/(f \sin \Theta)$.

Обсуждение рассматриваемого вопроса в данном разделе показало, что линзовые волноводы способны передавать пучки света по изогнутым траекториям, радиус кривизны которых определяется фокусным расстоянием и радиусом линз. Однако любое отклонение от прямолинейности или правильной круговой формы вызывает колебания луча. Маловероятно, чтобы эти колебания уменьшались в последующих точках оси волновода. Конечно, не исключено, что другой случайный перекося какой-нибудь линзы уменьшит амплитуду колебаний луча. Однако в среднем колебания луча имеют тенденцию к возрастанию, что будет видно при рассмотрении этого вопроса в следующем разделе.

Свойство изломов оси или смещенных линз изменять траекторию луча света может быть использовано для возвращения луча на ось волновода, если луч начал колебаться при своем распространении. Однако подобное подавление колебаний возможно только для лучей, траектории которых известны. В общем же случае это не возможно, что будет показано в следующем разделе. Подавление колебаний луча требует наличия датчиков, определяющих положение светового луча по крайней мере на двух линзах. Данные, получаемые с этих датчиков, могут быть использованы для управления последующими линзами, возвращающими пучок на осевую траекторию.

5.5. ЛИНЗОВЫЕ ВОЛНОВОДЫ СО СЛУЧАЙНЫМ СМЕЩЕНИЕМ ЛИНЗ

Из предыдущего раздела видно, что хотя линзовые волноводы и обладают способностью передавать лучи вдоль искривленного пути, однако лучи в этом случае следуют по колеблющимся траекториям. Так как радиусы линз имеют конечные размеры, амплитуда колебаний луча

в линзах не может расти выше значений радиуса линз, в противном случае луч будет потерян. Мы изучим статистическое поведение лучевых траекторий пучков света в линзовом волноводе для того, чтобы понять, насколько опасно хаотическое смещение линз.

Начнем рассмотрение статистики линзового волновода, предполагая, что имеет место случайное расположение линз и что между смещениями соседних линз корреляции нет [50, 52]. Предположение о случайности смещений линз является близким к реальности, так как невозможно сконструировать линзовый волновод с идеально установленными линзами. Если не учитывать корреляцию между линзами, что является идеализацией, то это представляет первое приближение к любой реальной задаче.

Рассмотрим прямолинейный линзовый волновод. Лучи вводятся в такой волновод по оси, поэтому имеем

$$\rho_1 = \rho_2 = 0. \quad (5.5.1)$$

Введем дисперсию σ_n^2 положения луча ¹⁾

$$\sigma_n^2 = \langle \rho_n^2 \rangle. \quad (5.5.2)$$

Символ $\langle \rangle$ означает усреднение по ансамблю. Из формулы (5.4.15) получаем

$$\sigma_n^2 = \frac{L^2}{f^2 \sin^2 \Theta} \sum_{v=2}^{n-1} \sum_{\mu=2}^{n-1} \langle S_v S_\mu \rangle \sin(n-v) \Theta \sin(n-\mu) \Theta. \quad (5.5.3)$$

Утверждение о случайности смещений линз при отсутствии корреляции между их положениями можно математически записать в форме

$$\langle S_v S_\mu \rangle = s^2 \delta_{v\mu}, \quad (5.5.4)$$

где s — среднеквадратичная величина смещения линз. Условие (5.5.4) позволяет получить из (5.5.3)

$$\sigma_n^2 = \frac{s^2 L^2}{f^2 \sin^2 \Theta} \sum_{v=2}^{n-1} \sin^2(n-v) \Theta. \quad (5.5.5)$$

После вычисления суммы можно записать

$$\sigma_n^2 = \frac{s^2 L^2}{2f^2 \sin^2 \Theta} \left[n - 2 - \frac{\cos(n-1) \Theta \sin(n-2) \Theta}{\sin \Theta} \right]. \quad (5.5.6)$$

¹⁾ Дисперсия случайного отклонения луча от среднего положения, в данном случае от оси. — Прим. ред.

Практический интерес представляет случай большого числа линз n . Если n достаточно велико, то можно пренебречь членами внутри скобок, и тогда получим для модуля отклонения (корня квадратного из дисперсии) выражение

$$\sigma_n = \frac{sL}{\sqrt{2} f \sin \Theta} \sqrt{n} \quad \text{при } n \gg 1. \quad (5.5.7)$$

Используя формулу (5.2.6), запишем это выражение в виде

$$\sigma_n = \frac{s \sqrt{\frac{L}{f}}}{\sqrt{2 - \frac{L}{2f}}} \sqrt{n}. \quad (5.5.8)$$

В наиболее практически интересном случае конфокального линзового волновода ($L = 2f$) выражение для σ_n упрощается:

$$\sigma_n = s \sqrt{2n}. \quad (5.5.9)$$

Можно заметить, что линзовый волновод с $L = 4f$ представляет собой предельный случай, когда еще определенные лучи могут передаваться.

Этот результат следует и из выражения (5.5.8), которое становится неограниченным при $L = 4f$. Пучок в этом случае нестабилен, и потери имеют место всегда, если линзы хоть слегка не выставлены в линию.

Для того чтобы почувствовать влияние на распространение луча случайного смещения линз, рассмотрим пример. Предположим, что необходимо передать пучок света с помощью линзового волновода на расстояние свыше 10 км. Используя газовые линзы с низкими потерями, целесообразно расстояние между ними делать в 1 м. Тогда полное число линз будет $n = 10\,000$. Если принять, что отклонение пучка света от оптической оси равно приблизительно 3 мм (приемлемое допущение для газовых линз), то необходимо обеспечить

$$s < 2,12 \cdot 10^{-2} \text{ мм} = 21,2 \text{ мкм}. \quad (5.5.10)$$

Это означает, что для того, чтобы пучок света прошел путь длиной 10 км в линзовом волноводе, линзы в среднем не должны отклоняться более чем на 20 мкм от своего номинального положения.

Напомним, что одной из замечательных особенностей газовых линз является то, что они позволяют делать промежутки между ними весьма малыми без введения дополнительных потерь в передающую линию. Наше рассмотрение искривленных линзовых волноводов показало, что, располагая линзы близко друг к другу, можно осуществить довольно крутые изгибы. Этот результат приводит к выводу, что при построении линзового волновода нужно располагать линзы очень близко друг к другу. Полученный же выше результат, однако, предостерегает от использования слишком большого числа линз. Допустимое отклонение в 20 мкм, несомненно, практически трудновыполнимо. Предположим, что можно допустить значение $s = 0,1$ мм. Это означает, что можно пропустить пучок света через 450 линз, т. е. на расстояние 450 м при интервале между линзами 1 м.

Соотношение (5.4.15) со случайным S_v является идеальным приложением для центральной предельной теоремы теории вероятностей [54]. Центральная предельная теорема устанавливает, что случайная переменная, которая представляет собой сумму большого числа некоррелированных случайных переменных, распределена по закону Гаусса независимо от распределения случайных переменных в сумме. Поэтому можно предположить, что вероятность нахождения луча в некотором заданном месте в линзовом волноводе со случайным смещением линз подчиняется гауссову распределению. Этот результат справедлив, конечно, только после того, как световой луч пройдет большое число линз. Однако он оказывается справедливым и для произвольного значения n , если распределение вероятностей смещения линз также подчиняется закону Гаусса. Гауссово распределение вероятностей полностью определяется его дисперсией. Таким образом, вероятность нахождения луча в положении r в интервале dr определяется [53] выражением

$$p_n(r) dr = \frac{1}{\sqrt{2\pi} \sigma_n} e^{-(r^2/2\sigma_n^2)} dr. \quad (5.5.11)$$

Это распределение вероятностей нахождения луча при заданном радиусе на n -й линзе справедливо, конечно, только для бесконечно больших линз. Выражение (5.5.11)

приближенно справедливо и для линз с ограниченным радиусом a при условии, что $\sigma_n \ll a$.

Если требуется рассмотреть реальный случай волноводов с линзами, радиус которых не обязательно много больше среднеквадратичного отклонения смещения пучка, то необходимо знать распределение вероятностей, которое по крайней мере приближенно справедливо для такого случая. Чтобы понять ограничения применимости распределения вероятностей (5.5.11), важно учесть, что луч, который наблюдается при данном радиусе r , может прийти многими путями. Возможно, что этот луч пройдет через n -ю линзу при максимальном отклонении своей волнообразной траектории в тот момент, когда мы зафиксируем его при значении r у n -й линзы, но более вероятно, однако, что этот луч пересечет n -ю линзу во время какой-нибудь другой части цикла траектории. Другими словами, мы можем выделить в некоторой заданной точке те лучи, которые с наибольшей вероятностью достигают значений, больших чем r , в течение последующего полупериода своей колеблющейся траектории. Если амплитуда отклонения луча A (т. е. амплитуда колеблющейся траектории) больше a , то луч не достигнет n -й линзы в волноводе с линзами ограниченного радиуса a . Распределение вероятностей (5.5.11) особенно чувствительно к присутствию линз с ограниченной апертурой, так как оно содержит информацию о лучах, проходящих через линзу при $r < a$, которые, однако, имеют амплитуду A , большую чем радиус линзы a , и которые на предыдущих линзах должны были выйти из волновода.

Предыдущее рассмотрение показывает, что более разумно для расчетов использовать распределение вероятностей амплитуд траекторий, чем распределение вероятностей положения лучей. Распределение вероятностей амплитуды колеблющегося процесса, мгновенное значение которого является гауссовым, хорошо изучено в теории тепловых шумов в электрических цепях. Это так называемое распределение Рэлея [55]:

$$P_n(A) dA = \frac{A}{\sigma_n^2} e^{-A^2/2\sigma_n^2} dA. \quad (5.5.12)$$

Нетрудно доказать справедливость этого утверждения. Уравнение траектории луча может быть записано

в виде

$$r_n = A \sin(n\Theta + \phi). \quad (5.5.13)$$

В линзовых волноводах со случайным смещением линз A и ϕ — случайные переменные. Для всякого луча они представляют собой медленно меняющиеся функции от n . Распределение вероятностей фазы должно быть постоянной величиной. Рассматривая однозначную часть траектории луча, которая простирается от $\sin(n\Theta + \phi) = -1$ до $\sin(n\Theta + \phi) = +1$, можно с уверенностью полагать, что распределение вероятностей ϕ есть

$$\Phi(\phi) = \frac{1}{\pi}. \quad (5.5.14)$$

Считая A постоянной, а ϕ меняющейся, рассмотрим область значений r_n внутри интервала от $-A$ до A . Вероятность нахождения r_n в интервале dr есть $G_A(r_n) dr$. G_A — условная плотность вероятности нахождения r_n при заданном A . Соотношение между G_A и Φ задается условием, состоящим в том, что если ϕ находится в интервале $d\phi$, то радиус r может быть найден как соответствующая точка в интервале dr . Выразим это условие уравнением

$$G_A(r) dr = \Phi(\phi) d\phi, \quad (5.5.15)$$

или

$$G_A(r) = \frac{1}{\frac{dr}{d\phi}} \Phi(\phi). \quad (5.5.16)$$

Используя формулы (5.5.13) и (5.5.14) и исключая ϕ , получаем

$$G_A(r) = \begin{cases} \frac{1}{\pi \sqrt{A^2 - r^2}} & \text{для } |r| < A, \\ 0 & \text{для } |r| > A. \end{cases} \quad (5.5.17)$$

Плотность распределения вероятностей положения луча r определяется выражением [54, 55]

$$p_n(r) = \int_0^\infty G_A(r) P_n(A) dA, \quad (5.5.18)$$

где $P_n(A)$ — плотность распределения вероятностей амплитуды луча A . Используя формулу (5.5.17), можно выра-

зить это соотношение в виде

$$p_n(r) = \frac{1}{\pi} \int_r^{\infty} \frac{P_n(A)}{\sqrt{A^2 - r^2}} dA. \quad (5.5.19)$$

Читатель может проверить, что два распределения вероятностей (5.5.11) и (5.5.12) удовлетворяют уравнению (5.5.19).

Соотношение (5.5.19) пригодно не только в случае неограниченно больших линз, но также и для реального случая линз с ограниченным радиусом. Однако в последнем случае $p_n(r)$ и $P_n(A)$ уже не определяются выражениями (5.5.11) и (5.5.12).

Для практических целей более важно знать вероятность того, что положение луча (или амплитуда траектории луча) меньше определенного значения. Интегральная вероятность нахождения луча в области с радиусом, меньшим чем r_n , равна

$$W_n(r_n) = 2 \int_0^{r_n} p_n(r) dr. \quad (5.5.20)$$

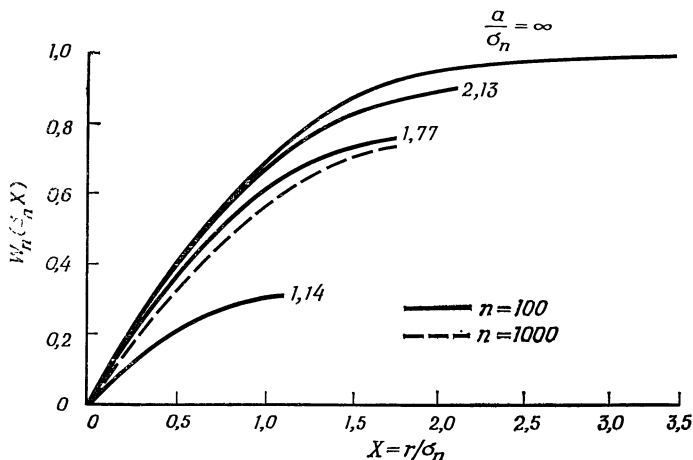
Поскольку p_n — четная функция, то можно ограничить область интегрирования положительными значениями r . Этим объясняется наличие множителя 2 в формуле (5.5.20). Интегральная вероятность того, что амплитуда траектории меньше A , есть

$$V_n(A) = \int_0^A P_n(x) dx. \quad (5.5.21)$$

Множитель 2 в этом случае отсутствует, так как вероятность P_n определена только для положительных значений A .

Поскольку оказалось очень сложным сформулировать аналитическое выражение для различных распределений вероятностей в случае линз с ограниченным радиусом, то автор провел ряд экспериментов по численному моделированию, чтобы получить распределение вероятностей положения луча в реальном линзовом волноводе с ограниченными линзовыми апертурами [53]. На ЭВМ моделировался конфокальный линзовый волновод со 100 линзами.

Линзы предполагались хаотически смещенными по закону случайных чисел. В каждом эксперименте начальное положение луча находилось на оси волновода и далее он проходил через линзы. Если отклонение луча принимало значение, большее чем радиус линзы, то далее этот



Ф и г. 5.5.1. Интегральная вероятность положения луча W_n для различных значений отношения a/σ_n (a — радиус линзы, σ_n — среднеквадратичное отклонение луча) [53].

луч не рассматривался. В каждом случае бралось новое распределение линз. Таким образом, оказалось возможным вычислить вероятность нахождения луча на 100-й линзе с радиусом, меньшим заданного радиуса r . Эта вероятность показана на фиг. 5.5.1. На этой фигуре изображено несколько кривых, соответствующих разным значениям a/σ_n . Величина σ_n , согласно формуле (5.5.9), указывает степень разброса положений линз. Горизонтальная ось графика нормирована относительно σ_n . Пунктирная кривая соответствует $a/\sigma_n = 1,77$, но она была получена при прохождении луча через 1000 линз вместо 100. Две кривые, соответствующие одинаковым значениям a/σ_n , достаточно близки. Это говорит о том, что нормированные кривые вероятности не очень сильно зависят от числа линз.

Из фиг. 5.5.1 видно, что вероятность W_n существенно зависит от отношения радиуса линзы к среднеквадратич-

ному отклонению. Именно этот результат и предполагалось получить. Теперь целесообразно выяснить, является ли распределение вероятностей V_n для A менее чувствительным к влиянию апертуры линз. С помощью формулы (5.5.19) можно получить соотношение между V_n и W_n . Изменим переменную r нижнего предела интеграла в (5.5.19) путем введения новой переменной интегрирования $u = A/r$:

$$p_n(r) = \frac{1}{\pi} \int_1^{\infty} \frac{P_n(ru)}{\sqrt{u^2 - 1}} du. \quad (5.5.22)$$

Подстановка (5.5.22) в (5.5.20) приводит к выражению

$$W_n(r) = \frac{2}{\pi} \int_1^{\infty} \frac{V_n(ru)}{u \sqrt{u^2 - 1}} du. \quad (5.5.23)$$

Соотношение (5.5.21) было использовано для представления одного из двух интегралов, которые появляются при этом вычислении, через V_n . Переход к переменной интегрирования A позволяет записать

$$W_n(r) = \frac{2r}{\pi} \int_r^{\infty} \frac{V_n(A)}{A \sqrt{A^2 - r^2}} dA. \quad (5.5.24)$$

Интегральная вероятность V_n обладает свойством

$$V_n(A) = V_n(a) \quad \text{для} \quad A \geq a. \quad (5.5.25)$$

Поскольку все лучи находятся в интервале $-a \leq r \leq a$, получаем дополнительное условие

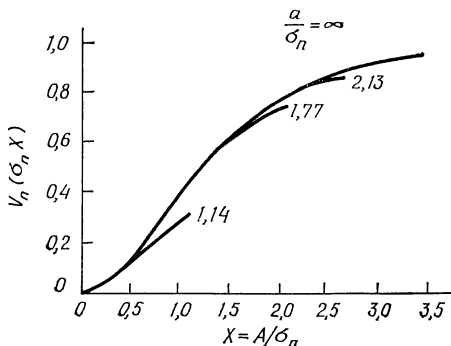
$$V_n(a) = W_n(a) \quad \text{для} \quad \begin{cases} r > a, \\ A > a. \end{cases} \quad (5.5.26)$$

Таким образом, можно осуществить интегрирование по A в два приема. Сначала проинтегрируем (5.5.24) от r до a , а к этому интегралу прибавим интеграл от a до ∞ . В последнем интеграле $V_n(A)$ предполагается постоянной величиной $W_n(a)$, так что интегрирование может быть доведено до конца. Получаем следующее интегральное

уравнение для $V_n(A)$:

$$W_n(r) - \frac{2}{\pi} W_n(a) \arcsin \frac{r}{a} = \frac{2r}{\pi} \int_r^a \frac{V_n(A)}{A \sqrt{A^2 - r^2}} dA. \quad (5.5.27)$$

Это интегральное уравнение может быть использовано для определения функции $V_n(A)$ по известной функции $W_n(r)$. Результат численного решения этого интегрального уравнения показан на фиг. 5.5.2. Представленные здесь кривые получены с использованием функции $W_n(r)$, графики которой изображены на фиг. 5.5.1 [53].



Ф и г. 5.5.2. Интегральная вероятность амплитуды колебания пучка для различных значений отношения a/σ_n [53].

Видно, что $V_n(A)$ намного меньше зависит от отношения a/σ_n , чем W_n . Действительно, для вероятностей $V_n(A)$, больших чем 0,8 (или 80%), функция $V_n(\sigma_n X)$ очень мало отличается от соответствующей функции для неограниченно большой линзы. Поэтому можно использовать функцию, которая является результатом подстановки выражения (5.5.12) в (5.5.21):

$$V_n(A) = 1 - e^{-A^2/2\sigma_n^2} \quad \text{для} \quad a/\sigma_n \rightarrow \infty. \quad (5.5.28)$$

Таким образом, строго доказано положение, что интегральная вероятность $V_n(A)$ амплитуды луча менее чувствительна к наличию линзовых апертур, чем вероятность $W_n(r)$ положения луча.

При проектировании линзовых волноводов для связи необходимо выполнить требование, чтобы световой пучок достигал конца волновода с большой вероятностью. С точки зрения волновой оптики не реально предполагать, что весь пучок теряется, если лучи пучка частично выходят за линзы. Распределение поля в линзовом волноводе имеет некоторую ширину, так что часть поля остается внутри волновода, даже если световой луч геометрической оптики попадает на край линзы. Однако можно ожидать больших потерь при передаче световой волны, если лучевая оптика предсказывает, что световой луч не достигнет конца линзового волновода.

Рассмотрим требования к линзовым волноводам, выполнение которых обеспечивает прохождение светового луча с 99 %-ной вероятностью. Возьмем в качестве примера опять конфокальный линзовый волновод ($L = 2f$). Требование достижения лучом последней линзы волновода с амплитудой A , меньшей чем a , при 99 %-ной вероятности ведет к условию $A = a = 3\sigma_n$. При $A/\sigma_n = 3$, согласно формуле (5.5.28), $V_n(A)$ равняется 0,989. Величина σ_n связана с среднеквадратичным смещением линз s посредством формулы (5.5.9). Отношение s/a , таким образом, однозначно определено требованием пропускания луча с 99 %-ной вероятностью. Для рассматриваемого конфокального линзового волновода получаем следующие значения допустимого отношения среднеквадратичной величины смещения линз к их радиусу:

$$\frac{s}{a} = \frac{1}{3\sqrt{2n}} = \begin{cases} 2,36 \cdot 10^{-2} & \text{для } n=100, \\ 6,47 \cdot 10^{-3} & \text{для } n=1000, \\ 2,36 \cdot 10^{-3} & \text{для } n=10\,000. \end{cases} \quad (5.5.29)$$

До сих пор предполагалось, что смещения линз некоррелированы друг с другом. Это допущение является слишком строгим для многих практических ситуаций. Рассмотрим, например, линзовый волновод, у которого линзы вмонтированы в трубу. Ось трубы не может быть идеально прямой, и всегда имеют место случайные смещения линз. Вследствие жесткости трубы случайные отклонения от прямой расположения ближайших линз не являются произвольными. Из-за жестко-

сти трубы смещения соседних линз коррелированы. Если одна линза отклоняется в определенном направлении от своего номинального положения, то соседние линзы будут с большой вероятностью смещаться в том же направлении из-за жесткости трубы. На эти коррелированные отклонения от прямой случайным образом накладываются некоррелированные отклонения от номинального положения, обусловленные установкой каждой отдельной линзы внутри трубы. Такие более или менее коррелированные отклонения от идеальной установки линз можно статистически описать с помощью коэффициента корреляции.

Рассмотрим среднее по ансамблю произведение смещений двух линз

$$R_{\nu\mu} = \langle S_\nu S_\mu \rangle. \quad (5.5.30)$$

При отсутствии корреляции $R_{\nu\mu}$ обращается в нуль при $\nu \neq \mu$. Если, однако, существует зависимость между отклонениями линзы ν и линзы μ от их номинального положения, т. е. если определенное значение S_ν будет вызывать отклонение S_μ в определенную сторону, то $R_{\nu\mu}$ будет не равен нулю при разных ν и μ . Нет причины, однако, ожидать, что коэффициент корреляции $R_{\nu\mu}$ будет иметь значительную величину, когда разность $\nu - \mu$ очень велика. Если коэффициент корреляции не стремится к нулю при $\nu - \mu \rightarrow \infty$, то можно заключить, что имеют место систематические смещения линз. Для истинно случайных процессов следует ожидать, что $R_{\nu\mu} \rightarrow 0$ при $\nu - \mu \rightarrow \infty$. Наконец, естественно допустить, что коэффициент корреляции зависит от $\rho = \nu - \mu$. Процессы, для которых такое предположение справедливо, называются стационарными случайными процессами [54, 55]. Для стационарного случайного процесса можно записать

$$R_{\nu\mu} = R_{\nu-\mu}, \quad (5.5.31)$$

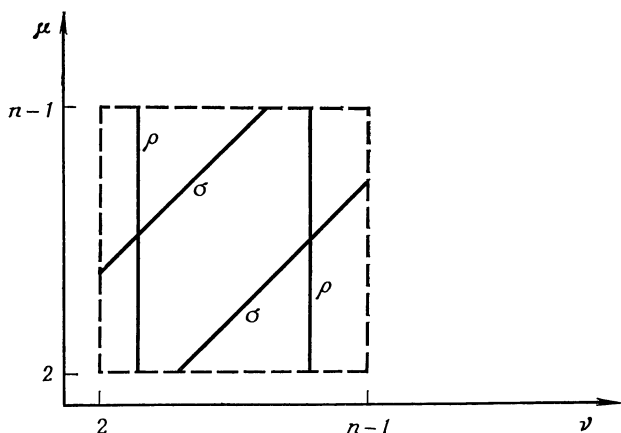
т. е. $R_{\nu\mu}$ зависит только от разности между ν и μ , а не от самих значений.

После такого отступления, касающегося коэффициента корреляции, можно рассчитать вариант положения луча для случая, когда между смещениями линз имеет место корреляция [52]. Вернемся к уравнению (5.5.3) и предста-

вим его с помощью коэффициента корреляции в следующем виде:

$$\sigma_n^2 = \frac{L^2}{2f^2 \sin^2 \Theta} \sum_{\nu=2}^{n-1} \sum_{\mu=2}^{n-1} R_{\nu-\mu} \{ \cos(\nu - \mu) \Theta - \cos[2n - (\nu + \mu)] \Theta \}. \quad (5.5.32)$$

Произведение синусоидальных функций преобразовано в сумму косинусов. Введем новые индексы суммирования $\sigma = \nu - \mu$ и $\rho = \nu$. Область суммирования по



Ф и г. 5.5.3. Диаграмма, показывающая переход от старых индексов суммирования ν , μ к новым σ , ρ .

индексам ν и μ представляет собой квадрат, который изображен на фиг. 5.5.3 пунктирной линией. Следует помнить, что все индексы суммирования могут принимать только положительные значения. Этот факт не отражен на рисунке. Сплошные линии, обозначенные через σ и ρ , ограничивают области в пространстве ν , μ , где σ или ρ принимают постоянные значения. Из большого числа параллельных линий, на которых σ и ρ могут принимать целые значения, изображена лишь одна пара. Чертеж позволяет определить область изменения новых индексов суммирования. Переход к этим новым индексам

суммирования приводит к следующему выражению:

$$\sigma_n^2 = \frac{L^2}{2f^2 \sin^2 \Theta} \left\{ \sum_{\sigma=-(n-3)}^0 R_\sigma \sum_{\rho=2}^{\sigma+n-1} [\cos \sigma \Theta - \right. \\ \left. - \cos (2n - 2\rho + \sigma) \Theta] + \right. \\ \left. + \sum_{\sigma=1}^{n-3} R_\sigma \sum_{\rho=2+\sigma}^{n-1} [\cos \sigma \Theta - \cos (2n - 2\rho + \sigma) \Theta] \right\}. \quad (5.5.33)$$

Пределы суммирования определяются следующим образом. Когда линия σ совпадает с диагональю квадрата на фиг. 5.5.3, σ принимает нулевое значение. Если линия σ движется вверх в квадрате, значение σ уменьшается и становится равным $\sigma = -(n - 3)$, когда линия σ достигает верхнего левого угла квадрата с координатами $\nu = 2$ и $\mu = n - 1$. Эти пределы указаны в первой сумме по σ в формуле (5.5.33). Чтобы определить пределы суммирования по ρ , нужно фиксировать величину σ . Значения ρ на линии σ находятся путем рассмотрения точки пересечения линии σ с левой стороной квадрата. Здесь $\nu = 2$, так что $\rho = 2$. Это нижний предел первой суммы по ρ . Верхний предел находится путем рассмотрения точки пересечения линий σ с верхней стороной квадрата. Здесь $\mu = n - 1$ и $\sigma = \nu - (n - 1)$, так что $\rho = \sigma + n - 1$. Это и есть верхний предел первой суммы по ρ . Пределы второй двойной суммы находятся подобным же образом в нижней правой части квадрата. Две двойные суммы могут быть сведены в одну путем введения нового индекса суммирования $\rho' = \rho - \sigma$ во второй сумме:

$$\sigma_n^2 = \frac{L^2}{2f^2 \sin^2 \Theta} \sum_{\sigma=-(n-3)}^{n-3} R_\sigma \sum_{\rho=2}^{n-1-|\sigma|} [\cos \sigma \Theta - \\ - \cos (2n - 2\rho - |\sigma|) \Theta]. \quad (5.5.34)$$

Суммирование по ρ приводит к результату

$$\sigma_n^2 = \frac{L^2}{2f^2 \sin^2 \Theta} \sum_{\sigma=-(n-3)}^{n-3} R_\sigma \left[(n - |\sigma| - 2) \cos \sigma \Theta - \right. \\ \left. - \frac{\cos (n - 1) \Theta \sin (n - 2 - |\sigma|) \Theta}{\sin \Theta} \right]. \quad (5.5.35)$$

В наиболее интересном случае, когда R_σ быстро уменьшается с ростом σ , в частности, когда n настолько велико, что значениями σ , для которых R_σ дает заметный вклад, можно пренебречь по сравнению с n , имеем приближенное выражение

$$\sigma_n^2 = \frac{nL^2}{2f^2 \sin^2 \Theta} \sum_{\sigma=-\infty}^{\infty} R_\sigma \cos \sigma \Theta. \quad (5.5.36)$$

Пределы суммирования произвольны, пока они велики, потому что R_σ близко к нулю при больших значениях σ .

При отсутствии корреляции имеем

$$R_\sigma = s^2 \delta_{0\sigma}. \quad (5.5.37)$$

В таком предельном случае выражение (5.5.35) сводится к (5.5.6), а (5.5.36) — к (5.5.7).

Истинная форма коэффициента корреляции зависит от конструктивных особенностей волновода. Чтобы получить конкретные выводы из нашей теории, необходимо использовать математические модели. В частности, наиболее простой коэффициент корреляции равен

$$R_\sigma = s^2 e^{-|\sigma|/B}. \quad (5.5.38)$$

Множитель s определяет среднеквадратичное значение смещения линз. Множитель B , стоящий в показателе экспоненты, имеет физический смысл интервала корреляции. Величина B определяет экспоненциальное уменьшение R_σ с ростом значений $|\sigma|$. Если коэффициент корреляции уменьшается быстро с ростом σ , то интервал корреляции между линзами мал. Если спад функции происходит медленно, мы говорим о большом интервале корреляции. Можно определить величину

$$D = BL \quad (5.5.39)$$

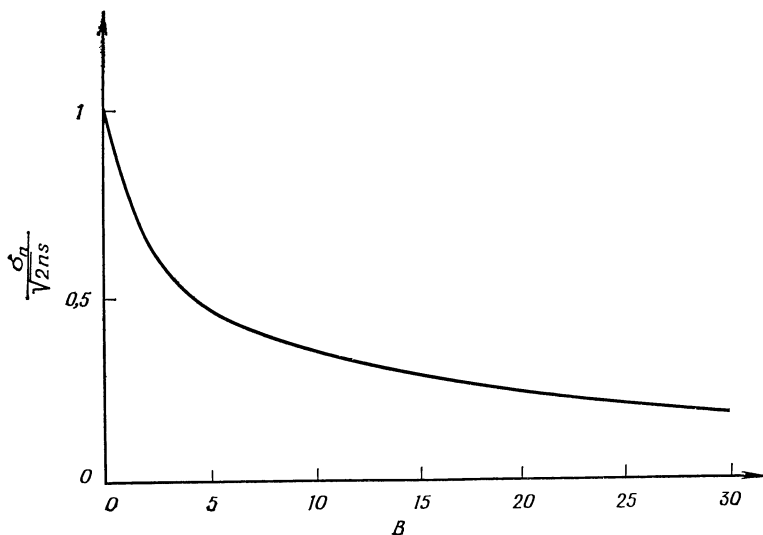
как корреляционную длину. Коэффициент корреляции, а также величины s и B описывают статистику разброса линз. Подстановка выражения (5.5.38) в (5.5.36) дает следующее значение среднеквадратичного смещения луча:

$$\sigma_n = \frac{L \sqrt{n} s}{\sqrt{2} f \sin \Theta} \left[\frac{\operatorname{sh} \frac{1}{B}}{\operatorname{ch} \frac{1}{B} - \cos \Theta} \right]^{1/2}. \quad (5.5.40)$$

Для конфокального расположения линз ($\cos \Theta = 0$, $\sin \Theta = 1$, $L = 2f$) это выражение упрощается к виду

$$\sigma_n = \sqrt{2n} s \left(\operatorname{th} \frac{1}{B} \right)^{1/2}. \quad (5.5.41)$$

В случае, когда интервал корреляции обращается в нуль ($B = 0$), (5.5.41) превращается в (5.5.9). Видно, что увеличение интервала корреляции улучшает характеристики



Ф и г. 5.5.4. Графическое представление функции (5.5.41).

линзового волновода за счет уменьшения среднеквадратичного отклонения луча. В пределе при $B \rightarrow \infty$ получаем идеальный линзовый волновод с $\sigma_n = 0$. Степень улучшения при увеличении B до полностью коррелированных смещений линз показана на фиг. 5.5.4. Для больших значений B кривая спадает как $1/\sqrt{B}$. В конфокальном линзовом волноводе, статистика которого описывается коэффициентом корреляции (5.5.38), отклонение пучка в 5 раз меньше, чем в конфокальном волноводе с полностью некоррелированными линзами (для одинаковых значений s), если корреляция между линзовыми смещениями охва-

тывает интервал примерно в 25 линз. Проведенное рассмотрение выявило некоторые общие свойства коррелированных смещений линз. Истинные детали процесса, связанные с отклонением луча, вызванным случайным смещением линз, зависят от различных обстоятельств в каждом отдельном случае. Получение соответствующего коэффициента корреляции для реального линзового волновода представляет собой трудную задачу.

Важной чертой волновода со случайным смещением линз является рост отклонения пучка пропорционально корню квадратному из числа линз. Если линзы оказываются смещенными систематически в соответствии с синусоидальной функцией, то смещение пучка может быть гораздо более сильным и расти прямо пропорционально числу линз [52]. Чтобы убедиться в этом, предположим, что линзы располагаются по следующему закону:

$$S_v = A \sin v\psi. \quad (5.5.42)$$

Подстановка этого выражения в (5.4.15) при $r_1 = r_2 = 0$ приводит к соотношению

$$\rho_n = \frac{AL}{2f \sin \Theta} \left[\frac{\cos \left(n \frac{\psi - \Theta}{2} + \frac{\psi + \Theta}{2} \right) \sin (n-2) \frac{\psi + \Theta}{2}}{\sin \frac{\psi + \Theta}{2}} - \frac{\cos \left(n \frac{\psi + \Theta}{2} + \frac{\psi - \Theta}{2} \right) \sin (n-2) \frac{\psi - \Theta}{2}}{\sin \frac{\psi - \Theta}{2}} \right]. \quad (5.5.43)$$

Пока ψ отличается от Θ , отклонение пучка ρ_n не слишком существенно. Однако, если $\psi = \Theta$, ситуация изменится радикально. В этом случае выражение (5.5.43) переходит в

$$\rho_n = -\frac{(n-2)AL}{2f \sin \Theta} \left[\cos n\Theta - \frac{\cos \Theta \sin (n-2)\Theta}{(n-2) \sin \Theta} \right]. \quad (5.5.44)$$

Амплитуда колебания луча растет линейно с числом линз n . Это показывает, что при синусоидальном смещении линз происходит более сильное отклонение пучка от оси, чем при случайном смещении линз. Так, например, для случая конфокального волновода имеем

$$\rho_n = -(n-2) A \cos n \frac{\pi}{2}. \quad (5.5.45)$$

Используя те же числа, которые фигурировали ранее в этой главе при обсуждении случайного смещения линз, находим, что если при $n = 10\,000$ ограничить смещение луча не более чем 3 мм, то можно допустить максимальное смещение линз только на

$$A \leq 3 \cdot 10^{-4} \text{ мм} = 0,3 \text{ мкм.} \quad (5.5.46)$$

Это действительно очень маленькая величина, равная примерно половине длины волны красного света.

Смещение линз по синусоидальному закону с таким же периодом, как период структуры ($\psi = \Theta$), не всегда реализуется. Необходимо помнить, однако, что произвольное смещение линз всегда может быть выражено в виде ряда Фурье и что фурье-компонента с периодом $\psi = \Theta$ обычно отлична от нуля. Ограничение (5.5.46) может быть, следовательно, интерпретировано как ограничение на фурье-амплитуду критической фурье-компоненты.

Последнее замечание может вызвать у читателя недоумение, поскольку оно вроде бы противоречит предыдущему. Мы нашли, что случайное смещение линз приводит к росту амплитуды раскачки луча пропорционально корню квадратному из n . Сейчас же мы как будто бы утверждаем, что луч растет пропорционально n за счет синусоидальной фурье-компоненты смещения линз. Кажущееся несоответствие исчезнет, если исследовать зависимость амплитуд фурье-компонент случайных функций от их длины. Можно показать, что амплитуды фурье-компонент случайных функций уменьшаются обратно пропорционально корню квадратному из длины функций (в данном случае числа линз), так что зависимость ρ_n (или среднеквадратичной величины) от корня квадратного из n также вытекает из фурье-анализа.

В этом разделе были выяснены общие свойства случайных процессов. Мы вновь встретимся с ними, когда будем обсуждать потери на рассеяние в диэлектрических волноводах.

Было показано, что световые лучи внутри линзового волновода распространяются по колеблющимся траекториям. Всякое желательное или нежелательное отклонение оси волновода от идеальной прямолинейности вынуждает световой луч осциллировать. Этот процесс напоминает

движение гармонического осциллятора под действием внешних сил. Колебания светового луча не опасны, когда их амплитуда остается значительно меньше радиуса линзы. Но если амплитуда колебаний возрастает так, что луч может выходить за линзу, то это приводит к потерям. Было бы весьма желательно изобрести такой линзовый волновод, который бы сам гасил колебания луча по мере распространения его по волноводу. Колебание луча в таком волноводе соответствовало бы затухающему гармоническому осциллятору.

Нетрудно построить математическую модель (к сожалению, неверную, но правдоподобную) линзового волновода, в котором лучи совершали бы затухающие колебания [24]. Используя тонкие линзы, обладающие формой, показанной на фиг. 4.5.11, которые описываются формулой (4.2.4), можно получить колеблющиеся траектории лучей с затухающими амплитудами. Однако рассмотрение теоремы Лиувилля в разд. 3.7 показало, что формулой такой линзы (4.2.4) можно пользоваться только тогда, когда световой луч проходит через тонкую линзу при нормальном или близком к нему падении. Поэтому вместо формулы (4.2.4) нужно использовать (4.2.3). Лучи, полученные с помощью точной формулы тонкой линзы, уже не имеют колеблющихся траекторий с затухающими амплитудами, а ведут себя подобно лучам, рассмотренным в разд. 5.3.

Читатель может подумать, что трудности конструирования линзового волновода с затухающими колебаниями луча связаны с отсутствием изобретательности. Теорема Лиувилля утверждает, что всякие попытки изобрести пассивный волновод с подавлением колебаний луча бесполезны. Предположим, что мы имеем устройство, подавляющее раскачку лучей. Точки в фазовом пространстве, соответствующие связке лучей, заполняют определенный объем. Напомним, что координатами фазового пространства являются положение и наклон лучей. Если «подавитель» колебаний луча работает, то будет иметь место тенденция к уменьшению амплитуды колебаний всех лучей. Это должно привести к сокращению объема в фазовом пространстве, что противоречит теореме Лиувилля, т. е. «подавитель» колебаний луча действует как ослабитель

лучевых осцилляций и, следовательно, невозможен. Таким образом, теорема Лиувилля позволяет строить лишь устройства, у которых наклон лучей соответствует лучевой амплитуде. Такое устройство может быть реализовано, например, путем постепенного увеличения оптической силы линз вдоль оси волновода и уменьшения расстояния между линзами таким образом, чтобы их фокусы все время оставались совпадающими. Однако уменьшение амплитуды лучей за счет изменения их наклона оказывается малым, поскольку невозможно неограниченно увеличивать оптическую силу линз. С некоторого момента необходимо вернуться к линзам с малой оптической силой. Но как только луч выйдет из волновода, линзы которого имеют большую оптическую силу, амплитуда его увеличится из-за уменьшения наклона и в результате не получится никакого выигрыша. Аналога затухающего гармонического осциллятора для световых лучей не существует. Можно уменьшить амплитуду лучей за счет потерь мощности, передаваемой пучком. Однако подавление лучевых колебаний за счет потерь мощности пучка, по-видимому, не очень выгодно ¹⁾.

Единственное практическое решение проблемы осциллирующих лучей возможно с помощью активного устройства, изменяющего направление луча [56]. Такое устройство должно быть чувствительно к положению отдельного луча и заставлять линзы или призмы изменять свое положение таким образом, чтобы световой пучок двигался обратно к оси. Оно не нарушает теорему Лиувилля, так как не действует на связку лучей с целью уменьшения ее объема в фазовом пространстве, а влияет на каждый отдельный луч, фазовый пространственный объем которого равняется нулю.

5.6. НОРМАЛЬНЫЕ МОДЫ ЛИНЗОВОГО ВОЛНОВОДА

В предыдущих разделах мы видели, что линзовые волноводы обладают способностью направлять световые лучи. Однако известно, что лучевая оптика дает лишь прибли-

¹⁾ Демпфирование раскочки пучка с помощью диафрагм рассмотрено в работе [37*].— *Прим. ред.*

женное описание распространения света, поэтому необходимо подтвердить справедливость полученных ранее результатов более строго. Мы увидим, что анализ с позиций волновой оптики эти результаты подтверждает. Дополнительно к информации о траекториях лучей мы получим также сведения, относящиеся к распределению поля и потерям световых пучков. Рассмотрение будет относиться к случаю волновода с правильно установленными линзами. Исследование случая хаотических смещений линз методами волновой оптики является гораздо более сложным. Результаты, полученные с помощью лучевой оптики, показали важность правильной расстановки линз.

Исследование линзового волновода с позиций волновой оптики базируется на скалярной волновой теории. Такое приближение является достаточным для выявления наиболее существенных явлений. Так, с его помощью можно показать, что в линзовом волноводе могут распространяться нормальные моды, а распределения полей более общего вида могут быть представлены как суперпозиции нормальных мод.

Для определения нормальных мод в линзовом волноводе воспользуемся тем обстоятельством, что распределение поля на каждой линзе повторяется с точностью до фазового множителя. Такое определение нормальной моды отличается от того, которое применяется в теории цилиндрических металлических волноводов [2, 58]. В этой теории типы волн имеют одно и то же распределение поля в любом сечении волновода (с точностью до фазового множителя). Такое различие можно объяснить следующим образом. Цилиндрический металлический волновод является структурой, инвариантной по отношению к произвольному смещению вдоль оси. При смещении металлического волновода на произвольную величину по оси z мы не обнаружим какого-либо изменения структуры, т. е. система инвариантна относительно произвольных смещений в направлении оси z . Линзовый же волновод имеет существенно отличные свойства. Произвольное смещение волновода вдоль его оси изменяет внешний вид волновода относительно неподвижного наблюдателя. Лишь при смещении волновода на величины, кратные расстоя-

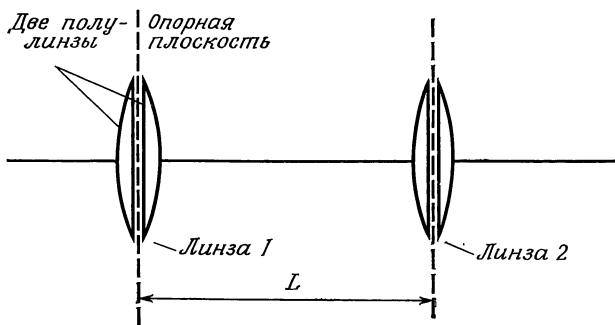
нию между линзами, его конфигурация сохраняется. Линзовый волновод является периодической структурой, и поэтому он инвариантен лишь относительно смещений, кратных периоду. Из теории групп следует, что любая функция $\exp(i\beta z)$ есть представление группы непрерывных смещений [59]. Но представления периодических структур должны быть периодическими функциями с периодом, равным периоду структуры.

Для читателя, знакомого с волновой механикой, может оказаться полезной следующая аналогия. Электрон в свободном пространстве можно описать с помощью плоских волн. Эти плоские волны могут рассматриваться в качестве нормальных мод, которые имеют одинаковый вид в любом поперечном сечении, перпендикулярном направлению распространения. Вдоль направления распространения изменяется только фаза. В периодическом кристалле волновая функция электронов совершенно иная. Она имеет вид функции Блоха [60] и представляется в виде произведения функции, период которой равен периоду кристаллической решетки, на фазовый множитель, зависящий от z . Нормальные моды в линзовом волноводе являются аналогами функции Блоха электронов в кристаллических решетках.

Задача нахождения модовых решений для системы апертур (отверстий в экране) привела Губо к изобретению линзового волновода. Путешествуя во время отпуска, он пытался решить задачу нахождения полей в системе периодических апертур. Он обратил внимание на то, что поля периодически повторяются у каждой апертуры системы, и сформулировал задачу, но не смог решить полученное интегральное уравнение. Губо догадался, что интегральное уравнение легко можно решить, если преобразовать его определенным образом. При рассмотрении физического смысла введенного преобразования он установил, что оно подразумевает включение линз в апертуры. Таким образом, он не только упростил решение математической задачи, но и изобрел важную и полезную систему передачи света с низкими потерями.

Найдем нормальные модовые решения линзового волновода с помощью дифракционной теории Кирхгофа — Гюйгенса [47, 48]. Рассмотрим две соседние линзы волно-

вода бесконечной длины, изображенные на фиг. 5.6.1. Линзовый волновод можно описать различными способами. Можно, например, выбрать опорную плоскость ¹⁾ посередине между линзами и рассматривать, как меняется поле в соседних плоскостях. Можно опорную плоскость выбрать непосредственно у линзы справа или слева от нее. Воспользуемся методом, наиболее удобным для описания лазерных



Фиг. 5.6.1. Линзовый волновод, для которого каждая линза представлена в виде двух полулинз. Распределение поля рассматривается в опорных плоскостях между полулинзами.

резонаторов. Каждую линзу разделим пополам и опорную плоскость разместим посередине между двумя половинами линзы. Передача поля через половину линзы описывается преобразованием (4.3.1) и (4.3.8):

$$\bar{\psi}_1 = e^{-i(h/4f)(a^2 - x^2 - y^2)} \psi_1(x, y). \quad (5.6.1)$$

Поле в опорной плоскости 1 обозначено через ψ_1 , а поле непосредственно справа от первой линзы — через $\bar{\psi}_1$. Радиус линзы обозначен через a [вместо ρ в (4.3.8)]. Вместо 2, входящей в формулу (4.3.8), стоит 4, так как фокусное расстояние половины линзы в 2 раза больше, чем фокусное расстояние f целой линзы. Поле $\bar{\psi}_2$ непосредственно у второй линзы слева может быть получено преобразованием поля $\bar{\psi}_1$ с помощью дифракционного интеграла Кирхгофа — Гюйгенса (2.2.31) в приближении Френеля.

¹⁾ То есть плоскость, в которой рассматривается поле. — Прим. ред.

Выбрав левую опорную плоскость за начало отсчета и полагая $\cos \alpha = \cos \beta = 1$, получаем

$$\bar{\psi}_2 = \frac{i}{\lambda L} e^{-ikL} \int_{A_1} \bar{\psi}_1 \exp \left\{ -i \frac{k}{2L} [(x' - x)^2 + (y' - y)^2] \right\} dx dy. \quad (5.6.2)$$

Поле ψ_2 во второй опорной плоскости окончательно получается с помощью преобразования, подобного (5.6.1):

$$\psi_2(x', y') = e^{-i(k/4f)(a^2 - x'^2 - y'^2)} \bar{\psi}_2. \quad (5.6.3)$$

Так как целью этих вычислений является нахождение модового решения, то необходимо потребовать, чтобы поле ψ_2 равнялось ψ_1 с точностью до множителя γ . Для линз с очень большой апертурой ($a \rightarrow \infty$) потерь мощности при распространении волны от первой опорной плоскости до второй быть не должно. В этом случае γ — просто фазовый множитель с единичным модулем. Если линзы имеют достаточно малые апертуры, часть поля не перехватывается ими, вследствие чего через линзу в опорной плоскости 2 проходит меньше мощности, чем через линзу в опорной плоскости 1. В этом случае γ является комплексной величиной, модуль которой меньше единицы. Условие для модового решения имеет следующий вид:

$$\psi_2 = \gamma \psi_1. \quad (5.6.4)$$

Уравнение (5.6.4) формулирует задачу на собственные значения.

Ему удовлетворяет не любое поле. Поле общего характера не удовлетворяет соотношению (5.6.4)¹⁾. Уравнение (5.6.4) имеет бесчисленное множество решений, и произвольное поле, возбужденное в линзовом волноводе, может быть описано как суперпозиция бесконечного числа мод [решений (5.6.4)] данной структуры. С помощью формул (5.6.1) — (5.6.3) можно записать соотношение (5.6.4) в виде

$$\gamma \psi(x', y') = \frac{i}{\lambda L} e^{-ik(L + a^2/2f)} \int_{A_1} \psi(x, y) \exp \left\{ i \frac{k}{2L} [(x^2 + y^2 + x'^2 + y'^2) \left(\frac{L}{2f} - 1 \right) + 2xx' + 2yy'] \right\} dx dy. \quad (5.6.5)$$

¹⁾ Автор здесь имеет в виду, что уравнение (5.6.4) справедливо только для параксиальных полей (волн). — *Прим. ред.*

Индекс 1 у функции ψ опущен. В общем виде уравнение (5.6.5) решать трудно. Но если ограничиться случаем конфокального линзового волновода

$$L = 2f, \quad (5.6.6)$$

то уравнение упрощается и аналитическое решение легко получается. Введя фазовый множитель перед интегралом в собственное значение γ и обозначив новое собственное значение через κ^2 , получаем следующее уравнение для мод конфокального линзового волновода в простой форме ($k = 2\pi/\lambda$):

$$\kappa^2 \psi(x', y') = \frac{1}{2\lambda f} \int_{A_1} \psi(x, y) e^{i(\pi/\lambda f)(xx' + yy')} dx dy. \quad (5.6.7)$$

Задача еще более упрощается, если будем искать решение в виде [47]

$$\psi(x, y) = f(x) g(y). \quad (5.6.8)$$

Теперь уравнение (5.6.7) разбивается на два:

$$\kappa f(x') = \frac{1}{\sqrt{2\lambda f}} \int_{-a}^a f(x) e^{i(\pi/\lambda f)xx'} dx \quad (5.6.9)$$

и

$$\kappa g(y') = \frac{1}{\sqrt{2\lambda f}} \int_{-a}^a g(y) e^{i(\pi/\lambda f)yy'} dy. \quad (5.6.10)$$

При этом полагается, что площадь интегрирования является квадратом. Апертуры линз имеют несколько необычную форму. Но в случае, когда пределы интегрирования гораздо больше площади, занимаемой полем, форма апертуры линзы не имеет значения, так как можно увеличить область интегрирования вплоть до $(-\infty, +\infty)$ без изменения величины интеграла. Для линз небольшого размера, когда интенсивность поля на границе апертуры не мала, решение зависит от формы апертуры. В этом случае наши результаты будут соответствовать линзам с квадратной диафрагмой. Результаты для линз с обычными круглыми апертурами получены численным решением интегрального уравнения (5.6.7) [48]. Результаты этих вычислений рассмотрим в этом разделе несколько позже.

Уравнения (5.6.9) и (5.6.10) имеют одинаковый вид (при $a \rightarrow \infty$ они являются преобразованиями Фурье). Поэтому достаточно рассмотреть решение лишь одного из этих уравнений. Разделение двумерного уравнения (5.6.7) сводит задачу к одномерной. Можно получить уравнение (5.6.9) при рассмотрении задачи о волноводе с цилиндрическими линзами.

Модовые решения дают распределения поля, которые концентрируются около оси линзового волновода. Если апертуры линз достаточно велики, то модовые решения не зависят от размера апертур. Поэтому без ущерба для точности можно расширить пределы интегрирования от $\pm a$ до $\pm \infty$. Следовательно, для линз с большими апертурами имеем

$$\kappa f(x') = \frac{1}{\sqrt{2\lambda f}} \int_{-\infty}^{\infty} f(x) e^{i(\pi/\lambda f)xx'} dx. \quad (5.6.11)$$

Приближение Френеля, использованное для получения уравнения (5.6.11), остается справедливым, несмотря на бесконечные пределы интегрирования, если рассматриваются волны с концентрацией поля вблизи оси. Если решение уравнения (5.6.11) распространить на большие расстояния от оптической оси, то оно становится несправедливым, поскольку при этом не будут выполняться условия, позволяющие использовать приближение Френеля.

Решение интегрального уравнения (5.6.11) может быть получено с помощью таблиц Градштейна и Рыжика [61], где имеется интегральное соотношение

$$(i^n) e^{-(y^2/2)} H_n(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ixy} e^{-(x^2/2)} H_n(x) dx. \quad (5.6.12)$$

Функции $H_n(x)$ называются полиномами Эрмита. Они определяются соотношением

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}. \quad (5.6.13)$$

Первые пять полиномов Эрмита имеют вид

$$\begin{aligned} H_0(x) &= 1, \\ H_1(x) &= 2x, \\ H_2(x) &= 4x^2 - 2, \\ H_3(x) &= 8x^3 - 12x, \\ H_4(x) &= 16x^4 - 48x^2 + 12. \end{aligned} \quad (5.6.14)$$

Сравнение выражений (5.6.14) и (5.6.12) показывает, что решением уравнения (5.6.7) является функция

$$f_n(x) = A_n H_n \left(\sqrt{\frac{\pi}{\lambda f}} x \right) e^{-(\pi x^2 / 2\lambda f)} \text{ для } n = 0, 1, 2, \dots, \infty \quad (5.6.15)$$

с собственными значениями

$$\kappa_n = i^n. \quad (5.6.16)$$

Функции (5.6.15) называются полиномами Эрмита — Гаусса. Они образуют полную систему функций ¹⁾. Их удобно нормировать таким образом, чтобы

$$\int_{-\infty}^{\infty} f_n^2(x) dx = 1. \quad (5.6.17)$$

С помощью таблиц Градштейна и Рыжика [61] находим, что

$$A_n = \frac{1}{(2^n n!)^{1/2} (\lambda f)^{1/4}}. \quad (5.6.18)$$

Особое значение имеет мода низшего порядка ($n = 0$). Для нее

$$f_0(x) = \frac{e^{-(\pi x^2 / 2\lambda f)}}{(\lambda f)^{1/4}}. \quad (5.6.19)$$

Нормальные моды ортогональны друг к другу и при нормировке (5.6.17) образуют ортонормированную систему функций, т. е.

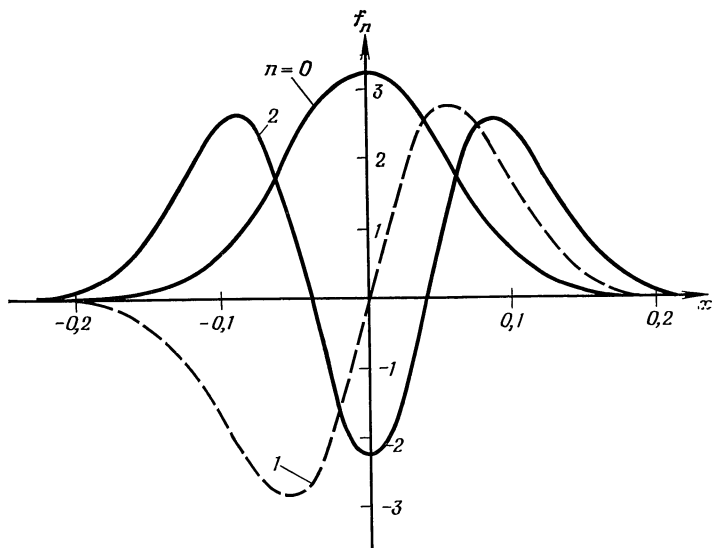
$$\int_{-\infty}^{\infty} f_n(x) f_m(x) dx = \delta_{nm}. \quad (5.6.20)$$

¹ В классе квадратично-интегрируемых функций. — *Прим. ред.*

Их значения быстро уменьшаются с ростом x по экспоненциальному закону. Полуширина интервала существенных значений экспоненциальной функции определяется параметром

$$w' = \sqrt{\frac{2\lambda f}{\pi}}. \quad (5.6.21)$$

При $x = w'$ поле, описываемое выражением (5.6.19), уменьшается в e раз по сравнению с его максимальным



Ф и г. 5.6.2. Графическое представление первых трех функций Эрмита — Гаусса.

значением при $x = 0$. Параметр w' не определяет наименьшую ширину пучка. Он задает ширину пучка на линзах. Пучок наиболее плотно концентрируется вокруг оси посередине между линзами. Мы вернемся к обсуждению формы распределения поля между линзами в главе, посвященной гауссовым пучкам.

На фиг. 5.6.2 показаны три первые функции f_n . Моды в линзовом волноводе с очень большими апертурами полу-

чаются из формул (5.6.8), (5.6.15) и (5.6.18):

$$\begin{aligned} \psi_{nm} = & \frac{1}{(2^{n+m} n! m! \lambda f)^{1/2}} H_n \left(\sqrt{\frac{\pi}{\lambda f}} x \right) \times \\ & \times H_m \left(\sqrt{\frac{\pi}{\lambda f}} y \right) e^{-(\pi/2\lambda f)(x^2+y^2)}. \end{aligned} \quad (5.6.22)$$

Собственные значения вытекают из (5.6.16)

$$\kappa_{nm}^2 = \kappa_n \kappa_m = i^{n+m}. \quad (5.6.23)$$

Важно иметь в виду, что полученное решение относится к случаю конфокального линзового волновода. Общее решение представлено в разд. 6.3 формулой (6.3.20).

После нахождения решения задачи о модах в конфокальных линзовых волноводах с очень большими апертурами линз рассмотрим случай, когда апертуры невелики и имеют прямоугольную форму. Интегральное уравнение (5.6.9) может быть решено с помощью вытянутых сфероидальных функций, которые подробно исследованы в работах Слепяна и Поллака [62], Ландау и Поллака [63], а также Слепяна и Зонненблика [41] (см. также разд. 4.6). Эти авторы исследовали функции, являющиеся решениями интегрального уравнения (5.6.9). Для того что бы увидать эквивалентность этого интегрального уравнения уравнению (29) на стр. 58 в статье [62], необходимо рассмотреть следующие соотношения между использованными там и здесь обозначениями:

$$\begin{aligned} \psi_n(c, t) &= f_n(x'), \quad \omega T/2\Omega = x, \quad t = x', \\ T &= 2a, \quad c = \pi a^2/\lambda f. \end{aligned}$$

Собственное значение κ_n связано с параметром λ_n , используемым в [62], соотношением

$$\kappa_n = i^n \sqrt{\lambda_n}. \quad (5.6.24)$$

Не следует путать параметр λ_n с длиной волны λ . Параметр λ_n подробно табулирован в [41]. Условие нормировки (5.6.17) сохраняется. Однако дополнительно к условию ортогональности (5.6.20) существует еще одно условие ортогональности модовых функций f_n на отрезке от $-a$

до $+a$ [ср. (4.6.12) в разд. 4.6]

$$\int_{-a}^a f_n(x) f_m(x) dx = \lambda_n \delta_{nm}. \quad (5.6.25)$$

Двойная ортогональность на бесконечности и на конечном отрезке является наиболее замечательным свойством мод конфокального линзового волновода с квадратными апертурами.

Мощность, проходящая через апертуру линзы, пропорциональна величине

$$P = \int_{-a}^a dx \int_{-a}^a dy |\psi_{nm}(x, y)|^2. \quad (5.6.26)$$

Поле первой опорной плоскости (фиг. 5.6.1) преобразуется следующим образом:

$$\psi_{nm} \rightarrow \kappa_{nm}^2 \psi_{nm}. \quad (5.6.27)$$

Это соотношение следует из интегрального уравнения (5.6.7). Мощность P_2 , проходящая через опорную плоскость 2, связана с мощностью P_1 , проходящей через опорную плоскость 1, соотношением

$$P_2 = |\kappa_{nm}^2|^2 P_1. \quad (5.6.28)$$

С помощью уравнений (5.6.23) и (5.6.24) можно найти связь между P_1 и P_2 через параметр λ_n :

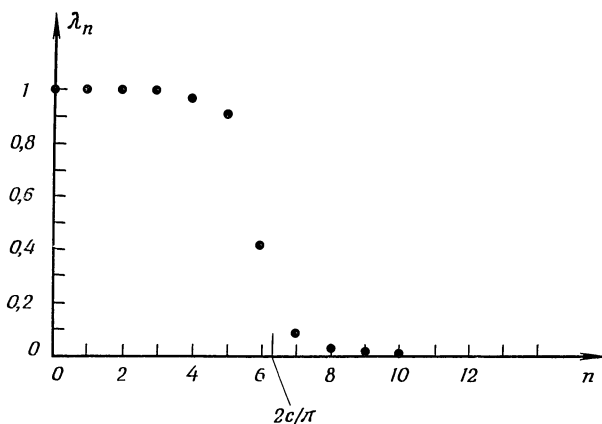
$$P_2 = \lambda_n \lambda_m P_1. \quad (5.6.29)$$

Параметр λ_n описывает потери в линзовом волноводе. Свойства λ_n уже подробно исследованы в разд. 4.6, где было показано, что параметр λ_n для заданных значений [ср. (4.6.15)]

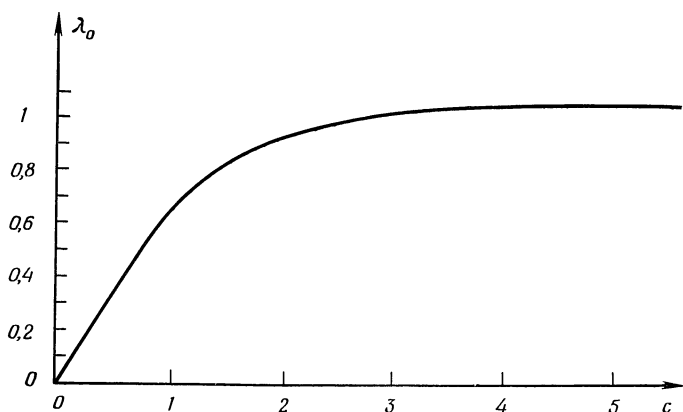
$$c = \pi \frac{a^2}{\lambda f} \quad (5.6.30)$$

близок к единице для $n < 2c/\pi$ и быстро уменьшается с ростом n при $n > 2c/\pi$. Поэтому потери нормальных мод малы при малых номерах, но весьма быстро растут для мод, номера которых больше $2c/\pi$. На фиг. 5.6.3 показана зависимость λ_n от n для $c = 10$, полученная

из таблиц [41]. Зависимость величины λ_0 для моды низшего порядка от параметра c показана на фиг. 5.6.4.



Ф и г. 5.6.3. Графическое представление зависимости собственных значений λ_n от порядка моды для $c = 10$. На оси n показана точка $2c/\pi$.



Ф и г. 5.6.4. Собственное значение моды низшего порядка как функции c .

Потери энергии мод обусловлены дифракцией. Поле распространяется, переходя от одной линзы к другой. Некоторая часть его перехватывается апертурой следующей

линзы, и таким образом поддерживается поток электромагнитной энергии вдоль линзового волновода. Если апертуры линз достаточно велики, то потери мощности из-за дифракции очень малы. Из фиг. 5.6.4 видно, что дифракционные потери для моды низшего порядка становятся малыми при $c > 4$.

Связь между интегральным уравнением (5.6.9) и уравнением (4.6.13) может быть установлена следующим образом. Из уравнения (5.6.9) находим $f(x)$ и подставляем в подынтегральное выражение того же самого уравнения. Таким образом, получаем

$$\kappa_n^2 f_n(x') = \frac{1}{2\lambda f} \int_{-a}^a f_n(x'') \int_{-a}^a e^{i(\pi/\lambda f)x(x'+x'')} dx dx''. \quad (5.6.31)$$

В результате интегрирования по x имеем

$$\kappa_n^2 f_n(x') = \frac{1}{\pi} \int_{-a}^a \frac{\sin \frac{\pi a}{\lambda f} (x' + x'')}{x' + x''} f_n(x'') dx''. \quad (5.6.32)$$

Функции $f_n(x)$ обладают следующим свойством:

$$f_n(-x) = (-1)^n f_n(x). \quad (5.6.33)$$

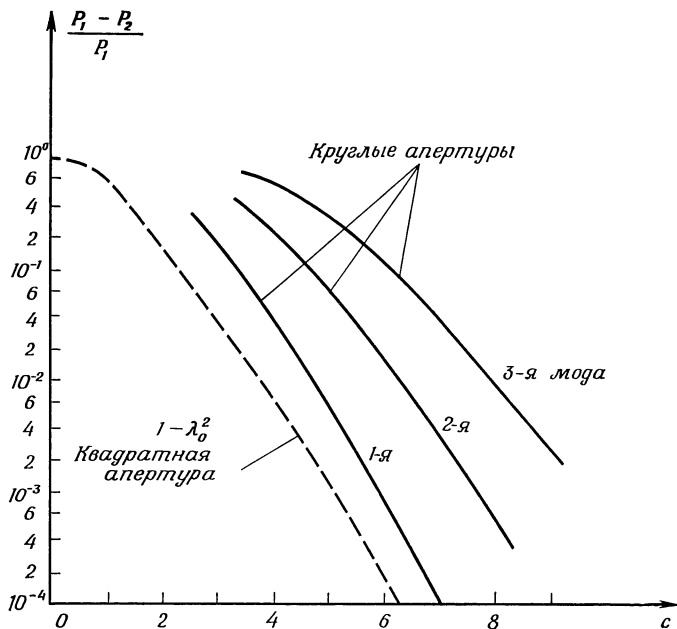
Это означает, что они являются четными или нечетными в зависимости от того, четное или нечетное n . В справедливости этого свойства можно убедиться, если вспомнить, что при $a \rightarrow \infty$ данная волновая функция при конечных значениях a должна непрерывно стремиться к значению, определяемому выражением (5.6.15). Свойство четности или нечетности сохраняется в процессе предельного перехода, т. е. функции (5.6.15) обладают свойством (5.6.33), что можно увидеть из (5.6.13) или (5.6.14). Используя уравнение (5.6.33), можно выражение (5.6.32) переписать в следующем виде:

$$(-1)^n \kappa_n^2 f_n(x') = \int_{-a}^a \frac{\sin \frac{\pi a}{\lambda f} (x' - x'')}{\pi (x' - x'')} f_n(x'') dx''. \quad (5.6.34)$$

Так как κ_n и λ_n связаны соотношением (5.6.24), то функции f_n действительно удовлетворяют уравнению (4.6.13),

что подтверждает тем самым их идентичность функциям ψ_n из разд. 4.6.

Таким образом, мы достаточно полно описали моды конфокального линзового волновода с квадратными апертурами. Моды конфокального линзового волновода с круглыми апертурами выражаются через гиперсфероидальные



Ф и г. 5.6.5. Сравнение потерь для первых трех мод линзового волновода с круглыми апертурами с потерями для моды низшего порядка волновода с квадратной апертурой [48].

функции [120]. Потери для такого волновода рассчитаны Фоксом и Ли [48]. Их расчеты выполнялись применительно к лазерным резонаторам. Однако в разд. 5.3 мы видели, что существует тесная связь между лазерными резонаторами и линзовыми волноводами. В обоих случаях дифракционные потери идентичны. В разд. 6.6 мы снова вернемся к аналогии между линзовыми волноводами и лазерными резонаторами. На фиг. 5.6.5 показаны поте-

ри для первых трех мод конфокального линзового волновода с круглыми апертурами. Пунктирной линией представлены потери мощности для моды низшего порядка линзового волновода с квадратной апертурой. При сравнении необходимо иметь в виду, что величина a , входящая в параметр s , является радиусом апертуры для случая линзового волновода с круглыми апертурами, а в случае квадратных апертур a есть полуширина стороны квадрата [см. формулу (5.6.30)]. Поэтому для тех же значений s площадь линзы с квадратной апертурой несколько больше. Дифракционные потери для линз с квадратными апертурами будут соответственно меньше.

5.7. ТРАЕКТОРИЯ ВОЛНЫ В КОНФОКАЛЬНОМ ЛИНЗОВОМ ВОЛНОВОДЕ

В предыдущем разделе мы нашли нормальные моды линзового волновода. Однако еще не было показано, что результаты лучевого рассмотрения могут быть получены из волновой теории. Установив, что траектории геометрических лучей соответствуют траекториям распространения световых волн в линзовом волноводе, можно с большим основанием использовать методы лучевой оптики. Несомненно, что при изучении линзовых волноводов методы лучевой оптики являются существенно более простыми, чем методы волновой оптики, поэтому желательнее там, где это возможно, использовать методы лучевой оптики.

Рассмотрим путь светового пучка в линзовом волноводе с бесконечными апертурами, полагая, что поле проходит через линзы на некотором расстоянии относительно оси. Начнем с предположения, что поле на первой опорной плоскости (фиг. 5.6.1) описывается выражением

$$\psi_1 = F_1(x) g_1(x), \quad (5.7.1)$$

где

$$F_1(x) = \frac{1}{(\lambda f)^{1/4}} e^{-(\pi/2\lambda f)(x+\xi)^2}, \quad (5.7.2)$$

а

$$g_1(y) = \frac{1}{(\lambda f)^{1/4}} e^{-(\pi/2\lambda f)y^2}. \quad (5.7.3)$$

Поле на следующей опорной плоскости задается соотношением (5.6.7) с тем исключением, что мы уже не требуем повторяемости поля, а используем интеграл в правой части этого соотношения для вычисления поля ψ_2 во второй опорной плоскости. Поскольку распределение поля записывается в виде произведения двух функций, одна из которых зависит от x , а другая — от y , интеграл представляет собой произведение интегралов по x и y . Функция g в выражении (5.7.3) является модой низшего порядка интеграла (5.6.10). Она удовлетворяет этому уравнению при $a \rightarrow \infty$ и распространяется через систему линз без искажения. Поэтому зависимость поля волны от y всегда определяется выражением (5.7.3) с точностью до фазового множителя. Следовательно, достаточно исследовать лишь зависимость функции (5.7.1) от x .

Функция F на следующей линзе определяется выражением

$$F_2 = DF_1, \quad (5.7.4)$$

где через D обозначен интегральный оператор правой части уравнения (5.6.9). Для нормальной моды получаем

$$\kappa_n f_n = Df_n. \quad (5.7.5)$$

Распределение поля на линзе с номером m дается выражением

$$F_m = D^{m-1}F_1. \quad (5.7.6)$$

Символ D^m обозначает m -кратное применение оператора D . Для нормальной моды имеем

$$\kappa_n^{m-1} f_n = D^{m-1} f_n. \quad (5.7.7)$$

Трудность рассматриваемой задачи состоит в том, что функция (5.7.2) не описывает нормальную моду линзового волновода. Она будет соответствовать нормальной моде лишь при $\xi = 0$. Для ненулевых значений ξ выражение (5.7.2) описывает распределение поля, как у нормальной моды, но смещенное на величину $-\xi$ от оси волновода.

Чтобы найти распределение поля (5.7.6) на линзе с номером m , представим начальное поле F_1 в виде суперпозиции нормальных мод. Пусть апертуры линз настолько велики, что все нормальные моды, дающие заметный вклад

в разложение, можно приближенно описать с помощью выражения (5.6.15).

В общем виде можно записать

$$F_1 = \sum_{v=0}^{\infty} a_v f_v. \quad (5.7.8)$$

С помощью соотношения (5.7.7) из формул (5.7.6) и (5.7.8) получим распределение поля на линзе с номером m :

$$F_m = \sum_{v=0}^{\infty} a_v \kappa_v^{m-1} f_v. \quad (5.7.9)$$

Таким образом, в принципе задача решена. Остается определить коэффициенты разложения a_v и сумму ряда (5.7.9).

Эти коэффициенты находятся из условия ортогональности (5.6.20). Умножая (5.7.8) на f_μ и интегрируя, получаем

$$a_\mu = \frac{1}{\sqrt{\lambda f}} \frac{1}{\sqrt{2^\mu \mu!}} \int_{-\infty}^{\infty} H_\mu \left(\sqrt{\frac{\pi}{\lambda f}} x \right) e^{-(\pi/2\lambda f)[x^2 + (x+\xi)^2]} dx. \quad (5.7.10)$$

Интеграл может быть найден с помощью таблиц Градштейна и Рыжика [61]:

$$a_\mu = \frac{1}{\sqrt{2^\mu \mu!}} \left(\frac{\pi}{\lambda f} \right)^{\mu/2} (-\xi)^\mu e^{-(\pi/4\lambda f)\xi^2}. \quad (5.7.11)$$

Используя выражения (5.6.15) и (5.6.16), получаем из (5.7.9)

$$F_m = \frac{1}{(\lambda f)^{1/4}} \sum_{\mu=0}^{\infty} \frac{i^{\mu(m-1)}}{2^\mu \mu!} \left(\frac{\pi}{\lambda f} \right)^{\mu/2} (-\xi)^\mu H_\mu \left(\sqrt{\frac{\pi}{\lambda f}} x \right) \times \\ \times e^{-(\pi/2\lambda f)(x^2 + 1/2\xi^2)}. \quad (5.7.12)$$

Сумму в выражении (5.7.12) можно оценить с помощью производящих функций полиномов Эрмита [61]:

$$e^{-t^2 + 2tz} = \sum_{\mu=0}^{\infty} \frac{t^\mu}{\mu!} H_\mu(z). \quad (5.7.13)$$

Таким образом,

$$F_m = \frac{1}{(\lambda f)^{1/4}} \exp \left\{ -\frac{\pi}{2\lambda f} \left[x^2 + \frac{1}{2} (i^{2(m-1)} + 1) \xi^2 + 2i^{m-1} x \xi \right] \right\}. \quad (5.7.14)$$

Выражение в квадратных скобках может быть записано в следующем виде:

$$\begin{aligned} x^2 + \frac{1}{2} (i^{2(m-1)} + 1) \xi^2 + 2i^{m-1} x \xi &= \\ &= x^2 + 2x\xi \frac{e^{im(\pi/2)} - e^{-im(\pi/2)}}{2i} + \xi^2 \frac{2 - e^{2im(\pi/2)} - e^{-2im(\pi/2)}}{4} + \\ &\quad + 2x\xi \frac{e^{im(\pi/2)} + e^{-im(\pi/2)}}{2i} = \\ &= \left(x + \xi \sin m \frac{\pi}{2} \right)^2 - 2ix\xi \cos m \frac{\pi}{2}. \quad (5.7.15) \end{aligned}$$

Поэтому распределение поля на линзе с номером m будет

$$F_m = \frac{1}{(\lambda f)^{1/4}} \exp \left\{ -\frac{\pi}{2\lambda f} \times \right. \\ \left. \times \left[\left(x + \xi \sin m \frac{\pi}{2} \right)^2 - 2ix\xi \cos m \frac{\pi}{2} \right] \right\}. \quad (5.7.16)$$

Мнимая часть показателя экспоненты указывает на зависимость фазового сдвига поля от x на каждой четной линзе. Физически это означает, что фазовые фронты у четных линз являются наклонными. Фазовые фронты у нечетных линз параллельны опорным плоскостям. Разумеется, эти результаты относятся только к конфокальному линзовому волноводу, возбуждаемому таким образом, что фазовый фронт поля у первой линзы параллелен опорной плоскости. Максимум распределения поля будет в той точке, где действительная часть показателя экспоненты равна нулю, т. е. при

$$x = -\xi \sin m \frac{\pi}{2}. \quad (5.7.17)$$

Сравнение с траекторией геометрооптических лучей (5.2.17) в конфокальных линзовых волноводах доказывает тот важный факт, что максимум распределения поля перемещается в волноводе так же, как луч в случае геометрооптического решения.

Таким образом, для частного случая конфокального линзового волновода и выбранной лучевой траектории доказано, что «центр тяжести» распределения поля распространяется в линзовом волноводе так же, как геометрооптический луч. Этот важный результат свидетельствует о том, что методы лучевой оптики дают хорошее приближение при решении волновых задач, так как он показывает, как перемещается вдоль линзового волновода «центр тяжести» распределения светового поля. Поскольку методы геометрической оптики существенно проще волновой оптики, то предпочтительнее исследовать поведение полей внутри линзового волновода именно этими методами. Мы доказали также, что волновое поле при распространении по волноводу сохраняет первоначальную форму. Это объясняется тем, что начальное поле выбрано таким образом, что оно соответствует смещенной моде нулевого порядка. В главе, посвященной гауссовым пучкам, будет показано, что в случае более общего распределения поля при движении центра тяжести вдоль лучевой траектории в процессе распространения волны вдоль волновода происходит периодическое изменение ширины траектории пучка. Частота этих изменений вдвое выше частоты осцилляций самого луча.

Важно понимать также, что поле сохраняет свою форму только в волноводе с идеальными линзами. В следующем разделе мы покажем, что в волноводе с неидеальными линзами в процессе распространения волнового поля происходит его деформация. Это наблюдается тогда, когда фокусное расстояние является функцией радиуса. Как мы видели в разд. 4.5, таким свойством обладают газовые линзы.

В разд. 3.6 было показано, что лучи в диэлектрике, переменный показатель преломления которого содержит только члены второго порядка по x и y , распространяются подобно центру тяжести волновых полей. Этот факт установлен с помощью теоремы Эренфеста из квантовой теории световых лучей. Эта теорема может быть теперь применена и к линзовому волноводу. Из нее следует, что только в линзовых волноводах с идеальными (квадратичными) линзами, т. е. линзами, описываемыми формулами (4.2.4) и (4.3.8), центр тяжести светового поля пере-

мещается как геометрический луч. В волноводе с произвольными (неквадратичными) линзами, например с газовыми, отсутствует простая связь между траекторией волнового поля и лучевыми решениями геометрической оптики. Для небольших участков волновода еще можно приближенно считать, что лучи геометрической оптики описывают траекторию светового поля. Когда же поле проходит через сотни и тысячи неидеальных линз, трудно ожидать, что геометрическое приближение справедливо. Результаты геометрических расчетов в таких случаях уже невозможно уверенно интерпретировать. В то же время эксперимент, выполненный со 100 стеклянными линзами из обычного промышленного стекла, которые имели фокусное расстояние 2,5 см, показал, что наблюдавшаяся траектория пучка лучей не разрушается после прохождения через линзовый волновод и что она соответствует предсказываемой из геометрических соображений. С другой стороны, моделирование с помощью ЭВМ процессов в линзовом волноводе, содержащем несколько сотен газовых линз, показало, что интерпретация волнового поля с помощью лучей не является очевидной. Таким образом, в случае линз с большими искажениями геометрическое приближение в общем случае неприемлемо.

5.8. РАЗРУШЕНИЕ ПУЧКА В НЕИДЕАЛЬНЫХ ЛИНЗОВЫХ ВОЛНОВОДАХ

В линзовом волноводе, содержащем хорошо выполненные тонкие линзы, световой пучок распространяется практически без искажений. Пучок может распространяться вдоль колеблющейся траектории, но если начальная форма поля соответствует одной из нормальных мод линзового волновода, то изменения поля не происходит даже в случае криволинейной траектории. Результаты моделирования на ЭВМ, а также экспериментальные исследования показали, что световой пучок в волноводе с неидеальными линзами не сохраняет своей формы и может значительно искажаться. Газовые линзы далеки от идеальности и деформация светового пучка происходит в процессе его распространения даже через сравнительно небольшое число линз [28].

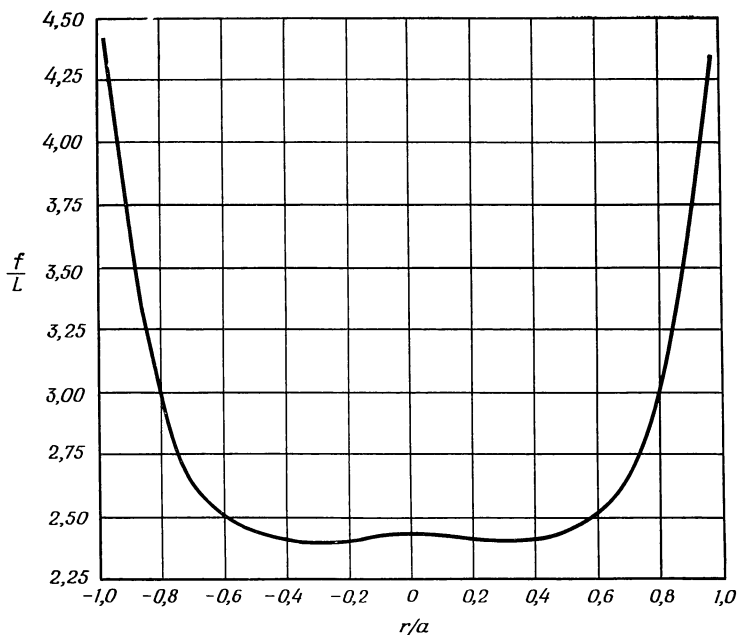
Искажения световых пучков в линзовом волноводе нежелательны по двум причинам. Во-первых, иногда требуется обнаружить слабый пучок света на выходном конце волновода. Эта задача упрощается, если входящий пучок имеет вид известной моды. Тогда можно воздействовать на него устройствами, трансформирующими пучок, чтобы согласовать его с местным гетеродинным пучком для более эффективного гетеродинного детектирования. Во-вторых, бывает желательно сфокусировать световую энергию в маленькое пятно на поверхности фотодиода. Фокусирование до малых пятен также по существу требует, чтобы пучки были одномодовыми. Может оказаться необходимым использовать устройства, изменяющие направление распространения пучка, чтобы избежать настолько большого увеличения амплитуды его расстройки, что пучок уже не сможет распространяться внутри линзового волновода [56, 57]. Но такие устройства работают на вполне определенных пучках. Искаженный пучок нельзя повернуть к оси.

Проблема искажений является наиболее серьезной в случае внеосевых пучков. Пучок, распространяющийся вдоль оси, заметно не искажается. В частности, всегда можно найти нормальную моду для любой периодической структуры, даже если линзы плохо формируют изображение и поэтому не могут считаться идеальными. Однако, как показано в разд. 5.5, осевые пучки не могут достаточно долго находиться на оси волновода из-за неизбежных смещений линз. Не удивительно, что внеосевые пучки в неидеальном волноводе искажаются. Линза называется неидеальной тогда, когда ее фокусное расстояние зависит от расстояния между лучом и осью волновода. Если же фокусное расстояние линзы одинаково для всех лучей, форма волны не искажается, даже если луч распространяется вдоль внеосевой колеблющейся траектории. Мода «подстраивается» к линзовому волноводу, если линзы имеют одинаковые фокусные расстояния независимо от их радиуса. Однако это не так, когда линзы неидеальны. Фокусное расстояние линз в этом случае зависит от радиального расстояния, на котором луч пересекает линзу. Пучок, являющийся модой при распространении вдоль оси, не соответствует волноводу, если он смещен от оси.

Следует отметить, что искажения возникают также в идеальном линзовом волноводе, если распределение поля не является модой данного волновода. Это обстоятельство в некоторой степени поясняет тот факт, почему распределение поля существенно искажается при распространении внеосевого пучка в неидеальном линзовом волноводе.

В разд. 4.5 было показано, что газовые линзы могут быть представлены в виде эквивалентных тонких линз, форма которых показана на фиг. 4.5.11. При моделировании на ЭВМ волновода с газовыми линзами использовались такие эквивалентные тонкие линзы [64]. Процесс прохождения волновых полей через искривленные линзы типа изображенной на фиг. 4.5.11 весьма сложен. Для описания этого процесса нужно использовать двумерный дифракционный интеграл Кирхгофа — Гюйгенса вида (2.2.41). Двумерная задача рассматривалась потому, что для решения полной трехмерной задачи необходимо двойное интегрирование, а это не только требует очень большого машинного времени, но и оказывается невозможным из-за недостаточной емкости памяти ЭВМ типа IBM 7094. Дифракционный интеграл становится сложнее из-за того, что обе функции и их производные должны быть известными на поверхности интегрирования. Упрощение (2.2.30) (или его двумерного эквивалента) возможно только тогда, когда производная функции, нормальная к фазовым фронтам, может быть приближенно заменена умножением на $-ik$. Производные, которые следует подставить в дифракционный интеграл, берутся в направлении нормали к поверхности интегрирования, т. е. в рассматриваемом случае к поверхности искривленной линзы. Это направление не перпендикулярно фазовым фронтам поля, так что эти производные не так просто получить. При численном определении производных требуется вычислять поле на двух близких замкнутых поверхностях. Эта процедура не только удваивает время вычисления, но и требует вдвое большей точности. Дополнительное машинное время нужно для того, чтобы вычислять не только саму функцию, но и ее производные [которые получаются дифференцированием уравнения (2.2.41)] на обеих поверхностях линзы. Для определения поля на следующей линзе необходимо выполнить численное интегрирование двух интегралов

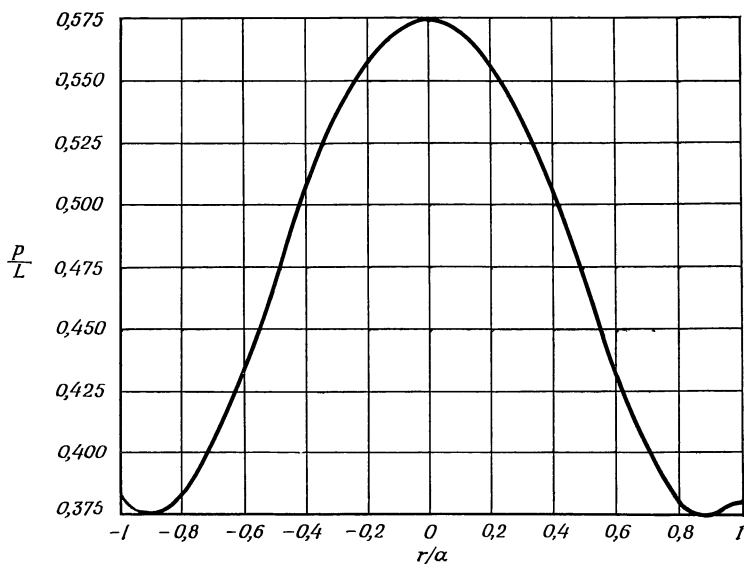
типа (2.2.41). Оценивание фазового преобразования, осуществляемого линзой, основано на уравнении (3.7.37) и подробно описано в [64]. При прохождении поля через линзу возникают специфические явления. Требуемый фазовый сдвиг может быть получен по законам лучевой



Ф и г. 5.8.1. Зависимость фокусного расстояния от положения луча в газовой линзе [64].

оптики. Однако если подсчитать прошедшую мощность на бесконечно малом расстоянии от линзы, то окажется, что закон сохранения энергии нарушается. Изменение мощности невелико. При моделировании на ЭВМ это учитывалось путем введения новой амплитуды в каждой точке поля после прохождения его через линзу так, чтобы плотность потока мощности в каждой точке линзы сохранялась. Этот пример показывает, что нужно быть очень осторожным при рассмотрении свойств тонкого оптического преобразователя. Выше уже отмечалось, какой опасно-

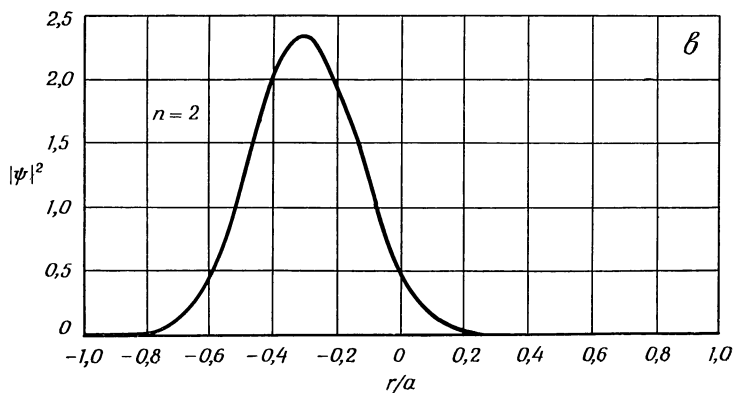
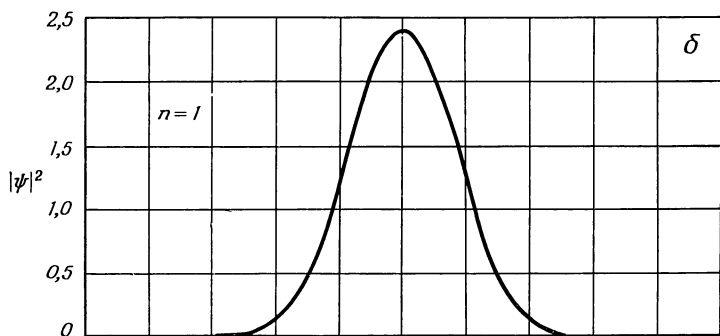
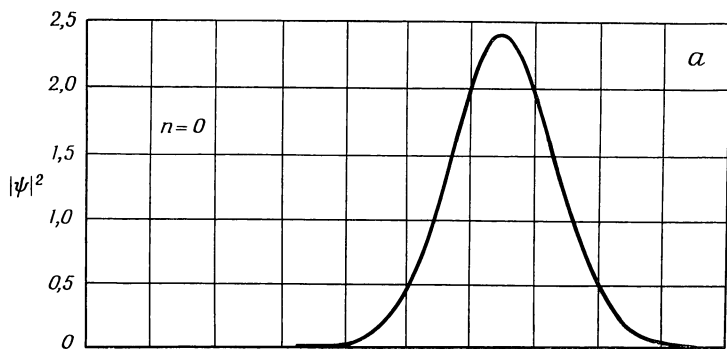
стью для изобретателей устройства, сдерживающего колебания луча, было игнорирование теоремы Лиувилля, которая является как бы проводником к правильному результату [24] (см. разд. 3.7 и конец разд. 5.5). Однако даже эта теорема не гарантирует сохранения энергии.



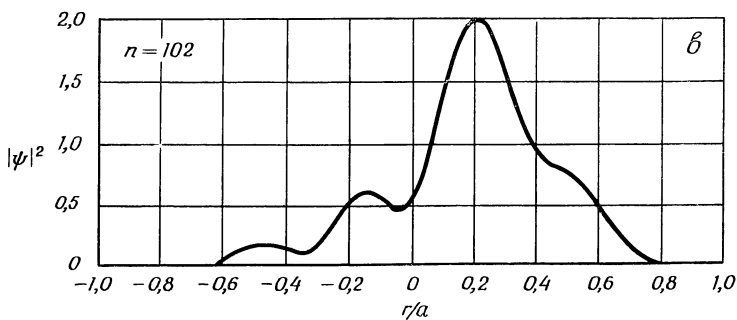
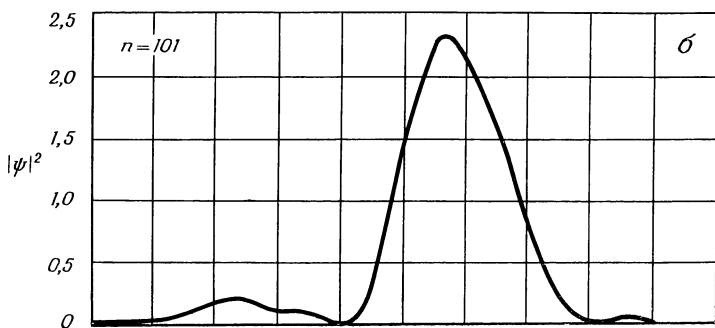
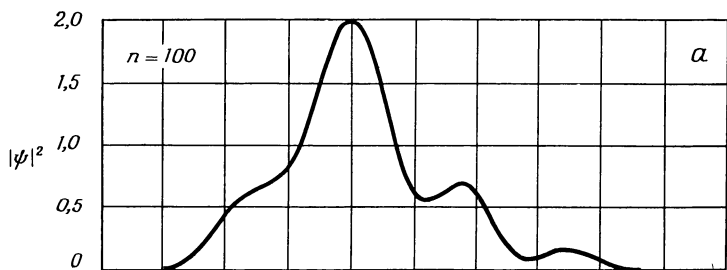
Фиг. 5.8.2. Форма главной поверхности газовой линзы [64].

Для численных расчетов было выбрано распределение поля низшей моды вдоль оси линзового волновода. Параметры линз были взяты такими же, как у реальной газовой линзы. Результаты моделирования на ЭВМ приведены на фиг. 5.8.1 — 5.8.5 [64].

На фиг. 5.8.1 представлена зависимость фокусного расстояния линзы, отсчитываемого от главной поверхности линзы, от радиального расстояния. На фиг. 5.8.2 показана форма главной поверхности линзы. Максимум распределения поля на первой линзе несколько смещен с оси. На фиг. 5.8.3 показано его положение у первых трех линз волновода. Пока не заметно какого-

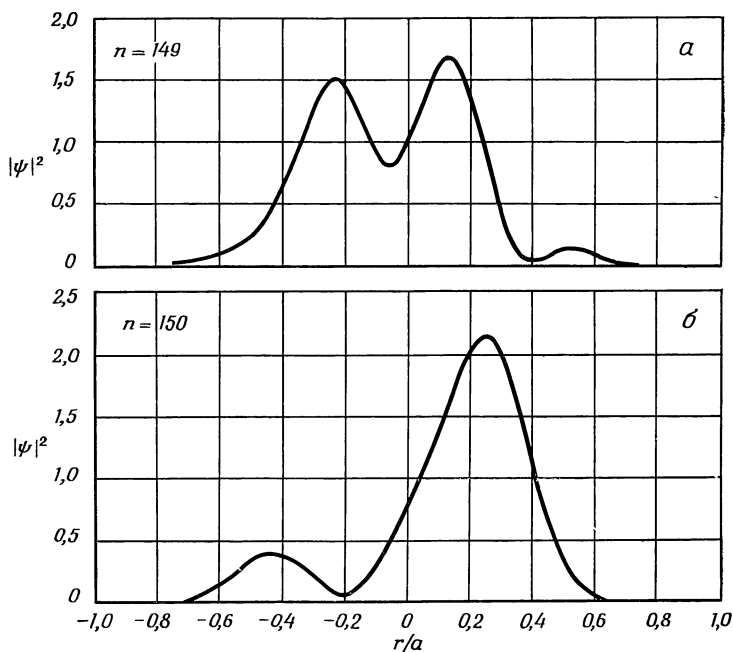


Ф и г. 5.8.3. Распределение квадрата поля на первых трех линзах волновода с газовыми линзами [64].



Ф и г. 5.8.4. Распределения квадрата поля на трех линзах за 100-й газовой линзой волновода. Видно появление искажений формы [64].

либо искажения. Это свидетельствует о том, что отдельные линзы значительных искажений не вносят. Искажение поля, обусловленное совместным действием многих линз,



Ф и г. 5.8.5. Разделение максимумов поля на 149-й газовой линзе волновода. На следующей 150-й линзе искажение поля существенно меньше [64].

показано на фиг. 5.8.4, где представлено распределение плотности мощности у линзы с номером 100 и двух последующих линз. Все еще отчетливо выделяется максимум, но видно, что часть поля движется к противоположному краю линзы. У линзы с номером 149 поле распадается на два пучка, близких по величине интенсивности. Неожиданной является форма распределения поля у следующей линзы (фиг. 5.8.5). Поле оказывается снова сфокусированным и выглядит почти неискаженным. Эти два распределения попеременно имеют место у следующих линз. Таким

образом, лишь после прохождения порядка 150 неидеальных газовых линз световой пучок разделяется на два пучка. Степень разрушения не связана с характером несовершенства линз. Всякая неидеальность, как бы мала она ни была, приводит в конце концов к подобной деформации пучка. Значительное различие в поведении линзовых волноводов, обусловленное отличием в степени искажения линз, связано с числом линз, через которые должен пройти пучок, прежде чем его разрушение станет заметным. Линзы с плоской главной поверхностью, имеющие неоднородность по фокусным расстояниям, также приводят к подобным деформациям пучка. Маркатили [65] аналитически для непрерывно фокусирующей среды с переменным показателем преломления получил такие же деформации пучка, которые были здесь рассмотрены. Экспериментальное исследование искажений световых пучков, распространяющихся в волноводе с газовыми линзами, проводилось Кайзером [28].

ГАУССОВЫ ПУЧКИ

6.1. ВВЕДЕНИЕ

Анализ распространения волн в линзовом волноводе показал, что моды линзового волновода с очень большими апертурами описываются произведениями полиномов Эрмита на функцию Гаусса. Функция Гаусса имеет вид $\exp(-ax^2)$. Если бы линзовый волновод был единственным устройством, где имеет место распределение поля по Эрмиту — Гауссу, едва ли нужно было бы его описанию посвящать целую главу. Однако, как уже отмечалось, лазерные резонаторы по своим свойствам близки к линзовым волноводам. Следствием такой связи является то, что моды лазера также описываются функциями Эрмита — Гаусса. Поскольку лазеры являются хорошим источником когерентного света, не удивительно, что гауссовы пучки приобрели важное значение с появлением лазера. На выходе хорошего одномодового лазера имеет место чистый гауссов пучок. Мода низшего порядка лазера и линзового волновода описывается просто гауссовой функцией, так как полином Эрмита низшего порядка равен единице [(см. 5.6.14)].

При использовании линзовых волноводов для целей связи важно, чтобы мода лазера эффективно их возбуждала. Это означает, что желательно ввести в волновод моду низшего порядка таким образом, чтобы она возбуждала также только моду низшего порядка, т. е. возникает проблема согласования гауссовых пучков. В этой главе также рассматриваются вопросы преобразования гауссовых пучков. Прежде чем можно будет рассмотреть согласование гауссовых пучков, необходимо более подробно изучить их свойства. В частности, важно узнать, как эти пучки распространяются в свободном пространстве,

Исследование распространения гауссова пучка будет проводиться двумя способами: с помощью дифракционного интеграла Кирхгофа — Гюйгенса и с использованием скалярного волнового уравнения в параксиальном приближении.

Ниже будет видно, что все моды типа Эрмита — Гаусса распространяются в основном таким же образом, как гауссов пучок. Поэтому достаточно остановиться на свойствах гауссовой моды низшего порядка.

Процесс преобразования гауссовых пучков обладает замечательным свойством: изменения радиуса кривизны фронта и ширины пучка формально подчиняются закону преобразования световых лучей.

6.2. РАСПРОСТРАНЕНИЕ ГАУССОВЫХ ПУЧКОВ В СВОБОДНОМ ПРОСТРАНСТВЕ

Предположим, что получен световой пучок с гауссовым распределением интенсивности, который имеет плоский фазовый фронт в плоскости $z = 0$. Такой пучок может быть описан выражением

$$\psi(x, y, 0) = A e^{-(x^2 + y^2)/w_0^2}. \quad (6.2.1)$$

Параметр w_0 есть половина ширины пучка в плоскости $z = 0$. Он равен расстоянию от максимума распределения поля до точки, где функция уменьшается в e раз по сравнению с максимальным значением.

Чтобы найти распределение поля вдоль оси z , рассмотрим дифракционный интеграл в виде (5.6.2):

$$\begin{aligned} \psi(x', y', z) = & \frac{iA}{\lambda z} e^{-ikz} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2 + y^2)/w_0^2} \times \\ & \times \exp \left\{ -i \frac{k}{2z} [(x' - x)^2 + (y' - y)^2] \right\} dx dy. \end{aligned} \quad (6.2.2)$$

Двойной интеграл может быть записан как произведение интеграла по x на интеграл по y . Оба интеграла имеют одинаковый вид. Таким образом, достаточно рассмотреть

один из них. Интеграл по x имеет вид

$$I_x = \int_{-\infty}^{\infty} e^{-(x^2/w_0^2)} e^{-i(k/2z)(x'-x)^2} dx = \sqrt{\pi} \frac{(w_0 \sqrt{2z})}{\sqrt{2z + ikw_0^2}} \times \\ \times \exp \left[-i \frac{2kzx'^2}{4z^2 + (kw_0^2)^2} \right] \exp \left[-\frac{(kw_0x')^2}{4z^2 + (kw_0^2)^2} \right]. \quad (6.2.3)$$

Этот интеграл легко берется с помощью таблиц, если представить экспоненту в виде $-(ax + b)^2 + c$ и ввести новую переменную интегрирования $u = ax + b$.

Полное поле на произвольном расстоянии z может быть записано в виде

$$\psi(x', y', z) = \frac{2i\pi w_0^2}{\lambda(2z + ikw_0^2)} A e^{-ikz} \exp \left[-i \frac{2kzr'^2}{4z^2 + (kw_0^2)^2} \right] \times \\ \times \exp \left[-\frac{(kw_0r')^2}{4z^2 + (kw_0^2)^2} \right], \quad (6.2.4)$$

$$r'^2 = x'^2 + y'^2. \quad (6.2.5)$$

Первый экспоненциальный множитель характеризует фазу плоской волны, второй — кривизну фазового фронта, а последний множитель определяет распределение интенсивности в поперечном сечении.

Введем параметр ширины пучка для произвольного z как расстояние (в поперечном направлении), на котором амплитуда поля уменьшается в e раз по сравнению с максимальным значением. Квадрат половины ширины пучка определяется следующим образом:

$$w^2(z) = w_0^2 \left[1 + \left(\frac{2z}{kw_0^2} \right)^2 \right]. \quad (6.2.6)$$

Подставляя $k = 2\pi/\lambda$, можно записать это выражение в следующем виде:

$$w^2(z) = w_0^2 \left[1 + \left(\frac{\lambda z}{\pi w_0^2} \right)^2 \right]. \quad (6.2.7)$$

Напомним, что пучок в точке $z = 0$ имеет плоский фазовый фронт и половину ширины, равную w_0 . От этой точки он распространяется в направлении роста z , увеличиваясь по ширине. Если взять дифракционный интеграл для определения поля влево от точки $z = 0$, то результат будет тот же самый. Равенство (6.2.7) справедливо как для

положительных, так и для отрицательных значений z и характеризует ширину пучка во всем пространстве. Видно, что точка $z = 0$, в которой гауссов пучок имеет плоский фронт, является также точкой, в которой он имеет наименьшее значение w_0 . Минимум полуширины пучка w_0 может быть любым. Однако расширение поля за пределами наиболее узкой его области зависит от минимальной ширины пучка. Можно определить половину угла расхождения пучка в точках, удаленных от плоскости $z = 0$, с помощью выражения

$$\theta = \lim_{z \rightarrow \infty} \frac{w(z)}{z}. \quad (6.2.8)$$

Половина угла расхождения определяется формулой

$$\theta = \frac{\lambda}{\pi w_0}. \quad (6.2.9)$$

Эта формула для полуширины гауссова светового пучка очень сходна с соответствующим выражением (2.4.11) для расширения пучка, получаемого от равномерно освещенной круглой апертуры. Однако гауссов пучок расширяется несколько медленнее, чем пучок, получаемый в случае равномерно освещенной апертуры.

Гауссов пучок в данной точке пространства не определяется полностью только шириной. Чтобы описать пучок полностью, необходимо знать второй параметр — радиус кривизны его волнового фронта. Мы получим его из выражения кривизны фазового фронта (6.2.4). Радиус кривизны поля может быть определен с помощью фиг. 6.2.1.

Фаза волны постоянна на сферической поверхности. Фазовый сдвиг между криволинейной поверхностью и плоскостью, касающейся этой поверхности в точке $r' = 0$, определяется уравнением

$$kd = \frac{2kzr'^2}{4z^2 + (kw_0^2)^2}. \quad (6.2.10)$$

Расстояние d показано на фиг. 6.2.1. Правая часть уравнения определяется из выражения (6.2.4). В соответствии с фиг. 6.2.1 для параксиальной области имеем соотношение

$$r'^2 + R^2 = (R + d)^2 \quad (6.2.11)$$

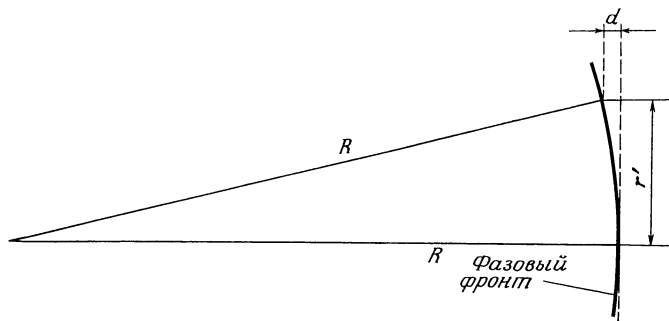
или с некоторым приближением

$$R = \frac{r'^2}{2d}. \quad (6.2.12)$$

Используя формулу (6.2.10), получаем выражение для радиуса кривизны гауссова пучка

$$R(z) = z \left[1 + \left(\frac{\pi w_0^2}{\lambda z} \right)^2 \right]. \quad (6.2.13)$$

Ширина пучка и радиус кривизны фазового фронта полностью определяют гауссов пучок в данной точке на оси. Выражение (6.2.13) справедливо как для положительных, так и для отрицательных значений z . Изменение знака $R(z)$, которое имеет место при перемене знака



Ф и г. 6.2.1. Сравнение фронтов сферической и плоской волн. Фазовый сдвиг равен kd .

z , означает изменение кривизны фазового фронта на обратное, так как при этом мы проходим область сужения пучка.

Несмотря на то что функции (6.2.7) и (6.2.13) непрерывны по z , нет оснований считать, что они точно описывают гауссов пучок в области $z = 0$, так как френелевское приближение дифракционного интеграла несправедливо для поля в ближней зоне, т. е. вблизи $z = 0$. Однако удивительно то, что обе формулы справедливы и в точке $z = 0$. Ширина пучка становится равной $w(0) = w_0$, а радиус кривизны принимает значение $R(0) = \infty$. Это означает, что эти формулы справедливы для всех значений z , несмотря на ограничения, накладываемые на дифракционный интеграл, из которого они были получены. В следующем

разделе будет показано, что (6.2.4) является приближенным решением волнового уравнения. На общее решение волнового уравнения не распространяется ограничение, связанное с ближним полем, которое накладывается на дифракционный интеграл, справедливый во френелевском приближении, поэтому можно быть уверенным, что выражения (6.2.4), (6.2.7) и (6.2.13) справедливы для всех значений z .

Решение (6.2.4) можно записать в более простой форме с помощью выражений для ширины пучка и радиуса кривизны фазовых фронтов. Преобразуем комплексный множитель перед экспоненциальной функцией:

$$\frac{2i\pi w_0^2}{\lambda(2z + ikw_0^2)} = \frac{1}{1 - i \frac{2z}{kw_0^2}} = \frac{w_0}{w(z)} e^{i \operatorname{arctg}(\lambda z / \pi w_0^2)}. \quad (6.2.14)$$

Используя выражения (6.2.6), (6.2.13) и (6.2.14), можно записать (6.2.4) в следующем виде:

$$\begin{aligned} \psi(x, y, z) = & \frac{w_0}{w(z)} A \exp \left\{ -i \left[kz + \frac{\pi r^2}{\lambda R(z)} - \right. \right. \\ & \left. \left. - \operatorname{arctg} \left(\frac{\lambda z}{\pi w_0^2} \right) \right] \right\} e^{-[r/u(z)]^2}. \end{aligned} \quad (6.2.15)$$

Это выражение не только проще по виду, но оно также непосредственно дает значение фазового сдвига поля в направлении оси z . В сужении (т. е. в наиболее узком месте пучка) в точке $z = 0$ фаза поля такая же, как и у плоской волны. По мере увеличения z фаза волны сдвигается относительно фазы плоской волны и закон ее изменения определяется функцией $1/R(z)$ и арктангенсом.

Однако в дальней области при $z \rightarrow \infty$ кривизна фронта поля $1/R(z)$ обращается в нуль, а значение функции arctg приближается к величине $\pi/2$ и фаза поля снова приближается к фазе плоской волны. Следуя работе Когельника [67], можно ввести комплексные параметры пучка

$$\frac{1}{q} = \frac{1}{R} + i \frac{\lambda}{\pi w^2(z)} \quad (6.2.16)$$

и

$$P = -\operatorname{arctg} \left(\frac{\lambda z}{\pi w_0^2} \right) + i \ln \left[\frac{w_0}{w(z)} \right]. \quad (6.2.17)$$

С учетом этих обозначений волновое поле можно записать в следующем виде:

$$\psi(x, y, z) = A \exp \left[-i \left(P + \frac{\pi}{\lambda q} r^2 \right) \right] e^{-ikz}. \quad (6.2.18)$$

Чтобы получить параметр q в несколько иной форме, которая будет более удобной при последующем рассмотрении, разделим выражение (6.2.7) на (6.2.13):

$$\frac{w^2(z)}{R(z)} = \left(\frac{\lambda}{\pi w_0} \right)^2 z. \quad (6.2.19)$$

Из (6.2.16) получим

$$q = \frac{R \left[1 + i \frac{\lambda R}{\pi w^2(z)} \right]}{1 + \left[\frac{\lambda R}{\pi w^2(z)} \right]^2}. \quad (6.2.20)$$

С учетом формул (6.2.13) и (6.2.19) будем иметь окончательно

$$q = z + i \frac{\pi w_0^2}{\lambda}. \quad (6.2.21)$$

Начало отсчета $z = 0$ находится в наиболее узкой части пучка. Таким образом, мы нашли весьма простое выражение для параметра q .

6.3. ДВОЙСТВЕННАЯ ПРИРОДА ГАУССОВЫХ ПУЧКОВ

Попытаемся вывести закон распространения гауссова пучка в свободном пространстве с помощью волнового уравнения. В качестве исходного возьмем уравнение вида (2.2.2):

$$\nabla^2 \psi + k^2 \psi = 0. \quad (6.3.1)$$

Решение будем искать в форме

$$\psi = u(x, y, z) e^{-ikz}, \quad (6.3.2)$$

т. е. нас интересуют решения, слабо отличающиеся от плоской волны. Быстрое изменение функции по z отражено в (6.3.2) экспоненциальным множителем. При этом можно считать, что функция u в зависимости от z меняется медленно по сравнению с быстрым изменением экспоненциального множителя. Применяя оператор ∇^2 к функции

(6.3.2), получаем

$$\nabla^2 \psi = \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} - 2ik \frac{\partial u}{\partial z} - k^2 u \right) e^{-ikz}. \quad (6.3.3)$$

Наше предположение относительно зависимости функции u от z дает возможность пренебречь второй производной функции u по z , сохранив первую производную и член $k^2 u$. Таким образом, получаем уравнение в частных производных для u из приведенного волнового уравнения [68]

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - 2ik \frac{\partial u}{\partial z} - k^2 u = 0. \quad (6.3.4)$$

Это уравнение для параксиальной области можно получить также путем подстановки выражения (6.3.2) в волновое уравнение (3.6.9). Однако из-за другого выбора знаков в (3.5.13) функция Гамильтона (3.5.17) описывает волну, распространяющуюся в отрицательном направлении оси z , поэтому в формулах (6.3.2) и (6.3.4) нужно взять k с противоположным знаком.

Легко показать, что гауссов пучок, который был получен в предыдущем разделе, является решением дифференциального уравнения в частных производных или, с учетом приближений по этапам от (6.3.1) до (6.3.4), решением скалярного волнового уравнения. Чтобы доказать это, представим выражение (6.2.18) в форме (6.3.2), откуда получим

$$u = A \exp \left[-i \left(P + \frac{\pi}{\lambda q} r^2 \right) \right], \quad (6.3.5)$$

где P и q являются функциями z . Нетрудно определить производные, которые входят в уравнение (6.3.4). Подставив выражение (6.3.5) в (6.3.4), получим

$$2k \left(P' + \frac{i}{q} \right) + \left(\frac{kr}{q} \right)^2 (1 - q') = 0. \quad (6.3.6)$$

Штрих означает производную по z . Каждый из двух членов должен независимо друг от друга обращаться в нуль, поскольку левая часть равна нулю при всех значениях z и r . В результате имеем два уравнения

$$q' = 1 \quad (6.3.7)$$

и

$$P' = -\frac{i}{q}. \quad (6.3.8)$$

Из формулы (6.2.21) следует, что равенство (6.3.7) выполняется. С помощью формул (6.2.7), (6.2.13), (6.2.16), (6.2.17) и (6.2.19) можно убедиться, что равенство (6.3.8) также удовлетворяется.

Таким образом, мы доказали, что поле (6.2.4), которое было получено с помощью приближения Френеля для дифракционного интеграла Кирхгофа — Гюйгенса, действительно является приближенным решением приведенного волнового уравнения. Приближение, необходимое для получения уравнения (6.3.4) из (6.3.1), оказывается вполне допустимым, поэтому следует ожидать, что гауссов пучок (6.2.15) является приемлемым приближением при всех значениях z .

Рассмотренное здесь гауссово распределение поля очень сходно с модой низшего порядка в линзовом волноводе, с которой мы встретились ранее в гл. 5. В гл. 5 также было выяснено, что моды более высокого порядка линзовых волноводов есть произведения гауссовой функции на полиномы Эрмита. Для изучения распространения мод высокого порядка линзового волновода в свободном пространстве произведем подстановку пробного решения

$$u(x, y, z) = f\left(\frac{x}{w}\right) g\left(\frac{y}{w}\right) \exp\left\{-i\left[P + \frac{\pi}{\lambda q} r^2 + \Phi(z)\right]\right\} \quad (6.3.9)$$

в дифференциальное уравнение (6.3.4). Параметры w , P и q являются функциями z и определяются в (6.2.7), (6.2.16) и (6.2.17). Уравнение (6.3.4) может быть преобразовано к следующему виду:

$$\begin{aligned} \frac{f''}{f} - 2ikx \frac{w}{q} \frac{f'}{f} + 2ikx \frac{\partial w}{\partial z} \frac{f'}{f} + \\ + \frac{\ddot{g}}{g} - 2iky \frac{w}{q} \frac{g'}{g} + 2iky \frac{\partial w}{\partial z} \frac{g'}{g} - 2kw^2 \frac{\partial \Phi}{\partial z} = 0. \end{aligned} \quad (6.3.10)$$

При упрощении этого уравнения было использовано равенство (6.3.6). Штрихи и точки обозначают производ-

ные по соответствующим аргументам функций. Аргументом f является x/w , а аргументом g является y/w .

С помощью формул (6.2.7), (6.2.16) и (6.2.21) можно исключить q и $\partial w/\partial z$ из уравнения (6.3.10), что дает

$$\frac{f''}{f} - 4\xi \frac{f'}{f} + \frac{\ddot{g}}{g} - 4\eta \frac{\dot{g}}{g} - 2kw^2 \frac{\partial \Phi}{\partial z} = 0, \quad (6.3.11)$$

где

$$\xi = \frac{x}{w} \quad \text{и} \quad \eta = \frac{y}{w}. \quad (6.3.12)$$

Напомним, что мы пытаемся получить решение приближенного волнового уравнения, которое описывает распространение пучка Эрмита — Гаусса в свободном пространстве. Функции f и g должны быть идентичны полиномам Эрмита. Дифференциальное уравнение для полинома Эрмита порядка n имеет вид [61]

$$\frac{d^2 H_n}{dt^2} - 2t \frac{dH_n}{dt} + 2nH_n = 0, \quad (6.3.13)$$

где n — целое положительное число.

Дифференциальное уравнение (6.3.11) содержит как раз такие члены, которые подобны членам дифференциального уравнения (6.3.13). Чтобы получить большее сходство, введем переменные

$$t = \sqrt{2}\xi \quad \text{и} \quad \tau = \sqrt{2}\eta. \quad (6.3.14)$$

Дифференциальное уравнение (6.3.11) примет вид

$$\frac{1}{f} \left(\frac{d^2 f}{dt^2} - 2t \frac{df}{dt} \right) + \frac{1}{g} \left(\frac{d^2 g}{d\tau^2} - 2\tau \frac{dg}{d\tau} \right) - kw^2 \frac{d\Phi}{dz} = 0. \quad (6.3.15)$$

Очевидно, что можно приравнять f и g следующим полиномам Эрмита:

$$f = H_n \left(\sqrt{2} \frac{x}{w} \right) \quad (6.3.16)$$

и

$$g = H_m \left(\sqrt{2} \frac{y}{w} \right) \quad (6.3.17)$$

при условии

$$kw^2 \frac{d\Phi}{dz} = -2(m+n). \quad (6.3.18)$$

Требование, чтобы f и g были полиномами Эрмита, приводит к определению фазы Φ в формуле (6.3.9), которая до сих пор оставалась произвольной. Потребовав, чтобы $\Phi(0) = 0$, получим с учетом выражения (6.2.7) решение дифференциального уравнения (6.3.18):

$$\Phi = -(m+n) \operatorname{arctg} \left(\frac{\lambda z}{\pi w_0^2} \right). \quad (6.3.19)$$

Суммируя результаты, получим следующее приближенное решение приведенного волнового уравнения (6.3.1):

$$\begin{aligned} \psi(x, y, z) = & A \frac{w_0}{w} H_n \left(\sqrt{2} \frac{x}{w} \right) H_m \left(\sqrt{2} \frac{y}{w} \right) \times \\ & \times \exp \left\{ -i \left[kz - (m+n+1) \operatorname{arctg} \left(\frac{\lambda z}{\pi w_0^2} \right) + \frac{\pi r^2}{\lambda R} \right] \right\} e^{-(r/w)^2}. \end{aligned} \quad (6.3.20)$$

Это решение приводит к важному результату, что параметр ширины пучка w (6.2.7) и кривизна фазового фронта R (6.2.13) одинаковы для всех мод пучка Эрмита — Гаусса. Фаза же волны зависит от модовых чисел n и m . Преобразование мод Эрмита — Гаусса в процессе их распространения может быть проведено независимо от значения модового числа, поскольку все моды имеют один и тот же параметр ширины пучка и одну и ту же кривизну фазового фронта. Исследование преобразования гауссовых пучков может быть таким образом ограничено рассмотрением моды низшего порядка. Общее решение (6.3.20) будет представлять интерес при изучении полей лазерных резонаторов.

Моды пучка Эрмита — Гаусса образуют полную систему ортогональных функций. Поэтому любое волновое поле можно представить как разложение в ряд по этим модам. Следует, однако, напомнить, что поле, записанное в виде (6.3.20), есть лишь приближенное решение волнового уравнения. Степень приближения ухудшается с увеличением числа $m+n$, и решение перестает быть верным, если вклад члена с множителем $m+n+1$ в экспоненте становится сравнимым со значением kz . Полное поле, представленное в виде разложения в ряд по модам (6.3.20), является, следовательно, не строгим, а приближенным решением волнового уравнения. Если члены высокого

порядка дают заметный вклад в разложение, то, чтобы ряд в целом был решением волнового уравнения, необходимо ввести дополнительные коэффициенты к членам ряда, причем эти члены в отдельности уже не будут удовлетворять волновому уравнению.

6.4. ПРЕОБРАЗОВАНИЕ ГАУССОВЫХ ПУЧКОВ

Часто бывает желательно ввести гауссов пучок света в оптическую систему таким образом, чтобы в ней возбудилась только одна мода. Примерами таких систем являются линзовые волноводы и оптические резонаторы. Такое согласование моды требует в общем случае преобразования гауссова пучка. Пусть, например, необходимо поле на выходе лазера преобразовать в гауссов пучок, который дает моду линзовой линии. В предыдущем разделе мы убедились, что параметры гауссова пучка (ширина пучка и кривизна фазового фронта) одинаковы для всех мод. Поэтому преобразование пучка, которое будет рассмотрено в данном разделе, справедливо одновременно для всех мод пучка Эрмита — Гаусса.

Начнем рассмотрение с преобразования гауссовых пучков от одной точки к другой вдоль оптической оси системы. Преобразование пучка по мере распространения его в свободном пространстве выражается формулами (6.2.7) и (6.2.13). Оба параметра пучка уже вошли в комплексный параметр q [см. формулу (6.2.16) или (6.2.21)]. Для любого данного комплексного значения q полуширина пучка w и кривизна фазового фронта R определены однозначно. Обозначая комплексный параметр на оптической оси пучка в точке z_1 через q_1 , а в точке z_2 — через q_2 , можно записать преобразование пучка в свободном пространстве простым соотношением

$$q_2 = q_1 + d_1, \quad (6.4.1)$$

где $d_1 = z_2 - z_1$.

Далее необходимо установить, как изменяется комплексный параметр пучка при его прохождении через идеально тонкую линзу. Пусть q_2 — параметр слева от линзы, а q_3 — параметр справа от линзы. Линза просто меняет фазу пучка, но не влияет на его ширину.

Параметр ширины пучка w , таким образом, остается неизменным, тогда как кривизна фазового фронта изменяется. Если выражение (6.2.15) представляет собой поле слева от линзы, то поле справа от линзы получим путем умножения его на величину фазового сдвига (4.3.8):

$$e^{i\gamma} = e^{i(\pi/\lambda f)r^2}. \quad (6.4.2)$$

Несущественный постоянный фазовый сдвиг здесь опущен. Произведение (6.2.15) на (6.4.2) снова должно дать гауссов пучок с другим радиусом кривизны R_3 (R_2 — радиус кривизны фронта слева от линзы). Таким образом, получаем

$$\frac{1}{R_3} = \frac{1}{R_2} - \frac{1}{f}. \quad (6.4.3)$$

Уравнение (6.4.3) при $w_3 = w_2$ дает закон преобразования гауссова пучка, проходящего через идеальную линзу. Предполагается, что ось гауссова пучка совпадает с осью линзы и пучок падает перпендикулярно к поверхности линзы.

Зная комплексный параметр пучка (6.2.16), можно записать закон преобразования пучка линзой следующим образом:

$$\frac{1}{q_3} = \frac{1}{q_2} - \frac{1}{f}, \quad (6.4.4)$$

или

$$q_3 = \frac{q_2}{-\frac{1}{f} q_2 + 1}. \quad (6.4.5)$$

При проведении преобразования пучка с расстояния d_1 слева от линзы на расстояние d_2 справа от линзы еще раз нужно применить закон преобразования. Параметр пучка на расстоянии d_2 за линзой обозначим через q_4 . Тогда имеем

$$q_4 = q_3 + d_2. \quad (6.4.6)$$

Из формул (6.4.1), (6.4.5) и (6.4.6) получаем

$$q_4 = \frac{\left(1 - \frac{d_1}{f}\right) q_1 + d_1 + d_2 - \frac{d_1 d_2}{f}}{-\frac{1}{f} q_1 + 1 - \frac{d_1}{f}}. \quad (6.4.7)$$

Закон преобразования гауссовых пучков обладает интересным матричным свойством, которое ясно видно, если записать формулу (6.4.7) в следующем виде [67]:

$$q_4 = \frac{A_4 q_1 + B_4}{C_4 q_1 + D_4}, \quad (6.4.8)$$

где

$$\begin{pmatrix} A_4 & B_4 \\ C_4 & D_4 \end{pmatrix} = \begin{pmatrix} 1 - \frac{d_2}{f} & d_1 + d_2 - \frac{d_1 d_2}{f} \\ -\frac{1}{f} & 1 - \frac{d_1}{f} \end{pmatrix}. \quad (6.4.9)$$

Произведем матричные преобразования, которые также приведут к выражению (6.4.7). Это может быть осуществлено путем сохранения формы записи (6.4.8) для всех преобразований и введения матрицы

$$\begin{pmatrix} A_1 & B_1 \\ C_1 & D_1 \end{pmatrix} = \begin{pmatrix} 1 & d_1 \\ 0 & 1 \end{pmatrix} \quad (6.4.10)$$

для преобразования (6.4.1), матрицы

$$\begin{pmatrix} A_2 & B_2 \\ C_2 & D_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \quad (6.4.11)$$

для преобразования (6.4.5), вносимого линзой, и, наконец, матрицы

$$\begin{pmatrix} A_3 & B_3 \\ C_3 & D_3 \end{pmatrix} = \begin{pmatrix} 1 & d_2 \\ 0 & 1 \end{pmatrix} \quad (6.4.12)$$

для преобразования (6.4.6). Нетрудно проверить, используя закон перемножения матриц, что матрица (6.4.9), описывающая преобразование (6.4.7) или (6.4.8), получена перемножением трех матриц:

$$\begin{pmatrix} A_4 & B_4 \\ C_4 & D_4 \end{pmatrix} = \begin{pmatrix} A_3 & B_3 \\ C_3 & D_3 \end{pmatrix} \begin{pmatrix} A_2 & B_2 \\ C_2 & D_2 \end{pmatrix} \begin{pmatrix} A_1 & B_1 \\ C_1 & D_1 \end{pmatrix}. \quad (6.4.13)$$

В общем случае, если имеют место два преобразования

$$q_k = \frac{A_j q_j + B_j}{C_j q_j + D_j} \quad (6.4.14)$$

и

$$q_j = \frac{A_i q_i + B_i}{C_i q_i + D_i}, \quad (6.4.15)$$

можно получить матрицу, соответствующую полному преобразованию

$$q_k = \frac{A_k q_i + B_k}{C_k q_i + D_k}, \quad (6.4.16a)$$

путем перемножения матриц первичных преобразований:

$$\begin{pmatrix} A_k & B_k \\ C_k & D_k \end{pmatrix} = \begin{pmatrix} A_j & B_j \\ C_j & D_j \end{pmatrix} \begin{pmatrix} A_i & B_i \\ C_i & D_i \end{pmatrix}. \quad (6.4.16b)$$

Это легко проверяется подстановкой (6.4.15) в (6.4.14). Закон преобразования параметров q гауссовых пучков есть закон $ABCD$ Когельника [49, 67], дающий очень удобный способ описания прохождения гауссовых пучков через сложные линзовые системы. Итак, было показано, что закон преобразования гауссова пучка (закон $ABCD$) справедлив при распространении гауссовых пучков в свободном пространстве и при прохождении через тонкие линзы. Кроме того, показано, что прохождение гауссовых пучков через ряд участков свободного пространства и линзы описывается с помощью произведений соответствующих матриц. Таким образом, можно утверждать, что закон $ABCD$ справедлив для любой оптической системы, которая может быть представлена рядом тонких линз и участками свободного пространства. Это открывает широкие возможности применения данного закона. Большая часть оптических устройств состоит из набора линз. Поскольку закон $ABCD$ справедлив для гауссова пучка, проходящего через любой элемент системы, а матрица пучка, проходящего через комбинацию этих элементов, получается путем перемножения отдельных матриц, то видно, что закон $ABCD$ справедлив для всей оптической системы. Линзовые волноводы представляют собой набор тонких линз и промежутков между ними, поэтому закон $ABCD$ справедлив и для них. Даже сплошная среда с квадратичным законом изменения показателя преломления (квадратичная среда) может рассматриваться как последовательность бесконечно близко расположенных линз. В этом случае закон $ABCD$ можно также применить, считая сначала, что число линз конечно, а затем устремляя число линз к бесконечности. Таким образом, мы видим, что закон $ABCD$ справедлив и для квадратичной среды.

При этом важно помнить, что закон $ABCD$ не справедлив для гауссова пучка, проходящего через среду с произвольным (неквадратичным) законом изменения показателя преломления, и для пучка, проходящего через линзы, обладающие аберрациями. Его применимость ограничена лишь системами, которые могут быть аппроксимированы комбинацией идеальных (квадратичных) тонких линз. Системы более общего характера меняют форму пучка (см. разд. 5.8), вследствие чего пучок, выходящий из такой системы, не является больше гауссовым и, следовательно, не может быть описан только двумя параметрами: полушириной и кривизной фазового фронта.

Существует интересная связь между лучевой оптикой и оптикой гауссовых пучков. Световые лучи также характеризуются двумя параметрами — положением пучка r и тангенсом угла его наклона r' . Прохождение светового луча от одной точки пространства к другой или его прохождение через линзу или любое другое оптическое устройство может быть описано следующим матричным уравнением:

$$\begin{pmatrix} r_2 \\ r'_2 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} r_1 \\ r'_1 \end{pmatrix}. \quad (6.4.17)$$

Здесь мы рассматриваем лучи, находящиеся в плоскости, проходящей через оптическую ось системы. Двух параметров — отклонения луча от оптической оси r и тангенса угла наклона луча r' — достаточно для полного описания поведения луча. В осесимметричных оптических системах, рассматриваемых здесь, всегда имеют место лучевые траектории, целиком лежащие в плоскости. Для пучка в свободном пространстве соотношение между его наклоном и положением в двух различных точках на расстоянии d вдоль оптической оси может быть выражено в параксиальном приближении как

$$r_2 = r_1 + r'_1 d_1 \quad (6.4.18)$$

и

$$r'_2 = r'_1. \quad (6.4.19)$$

Видно, что матрица, соответствующая этому преобразованию, та же самая, что и матрица (6.4.10) преобразования гауссова пучка между двумя точками пространства.

Прохождение луча через идеально тонкую линзу описывается преобразованием

$$r_2 = r_1 \quad (6.4.20)$$

и формулой (4.2.4), которая опять-таки в параксиальном приближении может быть записана в виде

$$r'_2 = r'_1 - \frac{r_1}{f}. \quad (6.4.21)$$

В соответствии с обозначениями, используемыми в разд. 4.2, $r' = -\alpha$. В матричных обозначениях это преобразование совпадает с преобразованием гауссова пучка линзой (6.4.14). Многократное преобразование световых лучей, очевидно, описывается произведением матриц отдельных преобразований. Таким образом, преобразование гауссовых пучков, проходящих через оптические системы, представляющие собой комбинацию тонких линз, может быть получено, если имеются решения соответствующих задач лучевой оптики.

Полученные результаты позволяют сразу же решить задачу об осевом распространении гауссовых пучков через линзовые волноводы. Матрица, описывающая распространение луча в идеальном линзовом волноводе, выводится из формулы (5.2.7). Вместо того чтобы выражать положение луча на n -й линзе через его положение на первых двух линзах, как это было сделано ранее, выразим его через положение и наклон на одной первой линзе:

$$r_1 = Ae^{i\theta} + Be^{-i\theta}. \quad (6.4.22)$$

Выражение для тангенса угла наклона луча r'_1 следует из формулы (6.4.21) при $r'_2 = (r_2 - r_1)/L$

$$\begin{aligned} r'_1 &= \frac{r_2 - r_1}{L} + \frac{r_1}{f} = \frac{1}{L} \left\{ A \left[\left(\frac{L}{f} - 1 \right) e^{i\theta} + e^{2i\theta} \right] + \right. \\ &\quad \left. + B \left[\left(\frac{L}{f} - 1 \right) e^{-i\theta} + e^{-2i\theta} \right] \right\} = \\ &= \frac{1}{L} [A(e^{i\theta} - 1) + B(e^{-i\theta} - 1)]. \end{aligned} \quad (6.4.23)$$

Эта формула получена с помощью (5.2.6). Формулы (6.4.22) и (6.4.23) могут быть использованы для полу-

чения выражений для A и B :

$$A = \frac{(1 - e^{-i\theta}) r_1 + L r'_1 e^{-i\theta}}{2i \sin \theta} \quad (6.4.24)$$

и

$$B = \frac{(e^{i\theta} - 1) r_1 - L r'_1 e^{i\theta}}{2i \sin \theta}. \quad (6.4.25)$$

Положение луча на n -й линзе, выраженное через положение и наклон луча на 1-й линзе, получается из (5.2.7):

$$r_n = \frac{[\sin n\theta - \sin(n-1)\theta] r_1 + L r'_1 \sin(n-1)\theta}{\sin \theta}, \quad (6.4.26)$$

$$\begin{aligned} r'_n &= \frac{r_n - r_{n-1}}{L} = \\ &= \frac{[\sin n\theta - 2 \sin(n-1)\theta + \sin(n-2)\theta] r_1 + \\ &\quad + L r'_1 [\sin(n-1)\theta - \sin(n-2)\theta]}{L \sin \theta}. \end{aligned} \quad (6.4.27)$$

Используя формулы суммы и разности синусов, получим матричное выражение для преобразования луча в следующем виде:

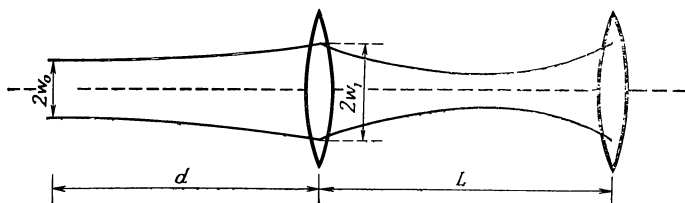
$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} \frac{\cos\left(n\theta - \frac{\theta}{2}\right)}{\cos \frac{\theta}{2}} & L \frac{\sin(n-1)\theta}{\sin \theta} \\ \frac{2 \sin[(n-1)\theta](-1 + \cos \theta)}{L \sin \theta} & \frac{\cos\left[(n-1)\theta - \frac{\theta}{2}\right]}{\cos \frac{\theta}{2}} \end{pmatrix}. \quad (6.4.28)$$

Такой же результат может быть получен путем построения матрицы одного элемента линзового волновода, которая описывает прохождение луча от положения слева от одной линзы в положение слева от следующей линзы. Матричная запись положения луча на n -й линзе, следовательно, получается возведением в $(n-1)$ -ую степень матрицы для одного элемента линзового волновода. Этот расчет предлагается читателю в качестве упражнения.

Знание матрицы $ABCD$ линзового волновода дает возможность рассчитать изменение параметра ширины гауссова пучка по мере распространения пучка по волно-

воду. Волноводная мода имеет одно и то же значение ширины пучка на каждой линзе. Следует ожидать, что гауссов пучок, который не является модой для данной направляющей структуры, будет меняться по ширине от линзы к линзе.

Параметр ширины пучка может быть получен из мнимой части выражения (6.2.16) для $1/q$. Из выражения (6.2.21) видно, что q является функцией расстояния z , отсчитываемого от наиболее узкой части пучка. Пусть гауссов пучок входит в линзовый волновод так, как показано на фиг. 6.4.1. Наиболее узкая часть входящего пучка



Фиг. 6.4.1. Ввод гауссова пучка в линзовый волновод.

находится на расстоянии d левее первой линзы. Прежде чем рассчитать изменение параметра ширины пучка на каждой линзе, необходимо знать его связь с комплексным параметром пучка q_1 на первой линзе. Эту связь можно найти путем подстановки $z = d$ в (6.2.21). С учетом (6.2.7) запишем

$$q_1 = d + ib, \quad (6.4.29)$$

где

$$b = \frac{\pi w_0^2}{\lambda}. \quad (6.4.30)$$

Из формулы (6.2.7) получаем

$$d^2 = \left(\frac{\pi w_0}{\lambda} \right)^2 (w_1^2 - w_0^2) = b^2 \left(\frac{w_1^2}{w_0^2} - 1 \right). \quad (6.4.31)$$

Теперь можно вычислить $1/q_n$ на n -й линзе с помощью матрицы (6.4.28). Из формулы (6.4.14) имеем

$$\frac{1}{q_n} = \frac{Cq_1 + D}{Aq_1 + B} = \frac{Cd + D + ibC}{Ad + B + ibA}. \quad (6.4.32)$$

Мнимая часть выражения $1/q_n$ имеет вид

$$\operatorname{Im} \frac{1}{q_n} = \frac{BC - AD}{(dA + B)^2 + b^2 A^2} b. \quad (6.4.33)$$

Выражение в числителе представляет собой отрицательное значение детерминанта матрицы $ABCD$. Непосредственный расчет показывает, что его значение равно -1 (см. выкладки в конце этого раздела). Используя формулу (6.2.16), получим требуемое выражение для параметра ширины пучка на n -й линзе:

$$\frac{\pi}{\lambda} w_n^2 = b \left[\left(\frac{d}{b} A + \frac{B}{b} \right)^2 + A^2 \right]. \quad (6.4.34)$$

Представим это выражение в более удобной форме. Выражения (6.4.30) и (6.4.31) позволяют записать (6.4.34) в виде

$$w_n^2 = w_1^2 A^2 + \frac{2\lambda}{\pi} \sqrt{\frac{w_1^2}{w_0^2} - 1} AB + \left(\frac{\lambda}{\pi w_0} \right)^2 B^2. \quad (6.4.35)$$

Матричные элементы A и B получаются из формулы (6.4.28) путем преобразования квадратов и произведений синусов и косинусов в функции двойного угла:

$$\begin{aligned} w_n^2 = \frac{1}{2 \cos^2 \frac{\theta}{2}} \left\{ w_1^2 [\cos \{(2n-1)\theta\} + 1] + \right. \\ \left. + \frac{\lambda L}{\pi} \sqrt{\frac{w_1^2}{w_0^2} - 1} \frac{\sin \left(2n - \frac{3}{2} \right) \theta - \sin \frac{\theta}{2}}{\sin \frac{\theta}{2}} + \right. \\ \left. + \left(\frac{\lambda L}{\pi w_0} \right)^2 \frac{1 - \cos (2n-2)\theta}{4 \sin^2 \frac{\theta}{2}} \right\}. \quad (6.4.36) \end{aligned}$$

Введем координату длины z_n вдоль линзового волновода, определив ее как

$$z_n = nL, \quad (6.4.37)$$

и длину периода колебания Λ луча

$$\frac{\theta}{L} = \frac{2\pi}{\Lambda}. \quad (6.4.38)$$

Выражение (6.4.26) показывает, что световой луч колеблется в линзовом волноводе с периодом Λ . Однако из

(6.4.36) видно, что ширина гауссова пучка внутри линзового волновода колеблется с периодом $\Lambda/2$. Таким образом, колебание ширины пучка происходит вдвое быстрее, чем колебание луча.

Соотношение (6.4.36) может быть записано в форме, которая более явно показывает зависимость от параметров колебаний:

$$w_n^2 = U^2 \cos(2n-1)\theta + V^2 \sin(2n-1)\theta + W^2, \quad (6.4.39)$$

где

$$U^2 = \frac{1}{2 \cos^2 \frac{\theta}{2}} \left[w_1^2 - \frac{\lambda L}{\pi} \sqrt{\frac{w_1^2}{w_0^2} - 1} \left(\frac{\lambda L}{\pi w_0} \right)^2 \frac{\cos \theta}{4 \sin^2 \frac{\theta}{2}} \right], \quad (6.4.40)$$

$$V^2 = \frac{\cos \frac{\theta}{2}}{2 \cos^2 \frac{\theta}{2} \sin \frac{\theta}{2}} \frac{\lambda L}{\pi} \left[\sqrt{\frac{w_1^2}{w_0^2} - 1} - \frac{1}{2} \frac{\lambda L}{\pi w_0^2} \right] \quad (6.4.41)$$

и

$$W^2 = \frac{1}{2 \cos^2 \frac{\theta}{2}} \left[w_1^2 - \frac{\lambda L}{\pi} \sqrt{\frac{w_1^2}{w_0^2} - 1} + \left(\frac{\lambda L}{\pi w_0} \right)^2 \frac{1}{4 \sin^2 \frac{\theta}{2}} \right]. \quad (6.4.42)$$

Гауссов пучок является модой линзового волновода, если его ширина остается неизменной на каждой линзе. Гауссов пучок, введенный в линзовый волновод, характеризуется полушириной w_0 в наиболее узкой части пучка и полушириной w_1 на первой линзе. Эти два параметра можно варьировать как угодно. Нормальная мода имеет место при условии $U^2 = V^2 = 0$, при этом $w_n = w_1 = W$. Требование $V^2 = 0$ приводит к условию

$$\sqrt{\frac{w_1^2}{w_0^2} - 1} = \frac{\lambda L}{2\pi w_0^2}. \quad (6.4.43)$$

Требование $W = w_1$ с учетом (6.4.43) дает

$$w_1 = \frac{1}{2 \sin \frac{\theta}{2}} \frac{\lambda L}{\pi w_0}. \quad (6.4.44)$$

Легко показать подстановкой (6.4.43) и (6.4.44) в (6.4.40), что $U^2 = 0$. Именно таким способом можно решить

совместно три уравнения с двумя переменными. Подстановка выражения (6.4.44) в (6.4.43) позволяет определить

$$w_1^2 = \frac{1}{\sin \theta} \frac{\lambda L}{\pi} \quad (6.4.45)$$

и

$$w_0^2 = \frac{\lambda L}{2\pi} \operatorname{ctg} \frac{\theta}{2}. \quad (6.4.46)$$

Таким образом можно решить задачу о нахождении мод линзового волновода общего вида с большой апертурой. В разд. 5.6 было получено решение интегрального уравнения только для частного случая конфокального линзового волновода. Теория преобразования гауссовых мод Когельника [67] дает возможность решить общую задачу простым и изящным путем. Моды линзового волновода описываются выражением (6.3.20), где w_0 и w определяются формулами (6.4.46) и (6.2.7), а R определяется выражением (6.2.13). Применяя эти формулы, нужно помнить, что начало координаты длины z находится в наиболее узкой части пучка. Вследствие симметрии системы оказывается, что наиболее узкая часть пучка расположена посередине между линзами.

Чтобы выразить параметр ширины пучка через размеры волновода и фокусные расстояния линз, исключим θ из (6.4.45) и (6.4.46) с учетом (5.2.6):

$$w_1 = \left(\frac{2\lambda f}{\pi \sqrt{\frac{4f}{L} - 1}} \right)^{1/2}. \quad (6.4.47)$$

Здесь w_1 — половина ширины пучка на линзе. Полуширина пучка в середине промежутка между двумя линзами равна

$$w_0 = \left(\frac{\lambda L}{2\pi} \sqrt{\frac{4f}{L} - 1} \right)^{1/2}. \quad (6.4.48)$$

В частном случае конфокального линзового волновода при $L = 2f$ имеем

$$w_0 = \sqrt{\frac{\lambda f}{\pi}} \quad (6.4.49)$$

и

$$w_1 = \sqrt{2} w_0. \quad (6.4.50)$$

Формула (6.4.50) согласуется с (5.6.21).

Поскольку существует взаимосвязь между линзовым волноводом и лазерным резонатором, полученное решение задачи о модах произвольного линзового волновода применимо также к модам лазерных резонаторов с зеркалами произвольной кривизны. И конечно, это позволяет решить задачу о модах в линзовом волноводе с чередующимися линзами, имеющими разные фокусные расстояния (см. разд. 5.2). Однако мы не будем рассматривать здесь этот вопрос. Связь между модами лазерных резонаторов и линзовых волноводов будет рассматриваться более подробно в разд. 6.6.

Задача о модах может быть также решена другим способом. Используя фиг. 6.4.1, можно определить матрицу преобразования лучей или параметров пучка для одного элемента периодической системы. За начало отсчета возьмем плоскость непосредственно слева от первой линзы и будем перемещаться по направлению к плоскости слева от следующей линзы. Матрица для элементов волновода получается перемножением (6.4.10) и (6.4.11):

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} 1 & L \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{pmatrix} \begin{pmatrix} 1 - \frac{L}{f} & L \\ -\frac{1}{f} & 1 \end{pmatrix}. \quad (6.4.51)$$

Условие образования моды находится из требования, чтобы преобразованный комплексный параметр на второй линзе соответствовал этому же параметру на первой линзе. Условие образования моды может быть, следовательно, выражено формулой

$$q = \frac{Aq + B}{Cq + D}. \quad (6.4.52)$$

Решение этой задачи приводит к тем же результатам, которые были получены выше. Сам расчет предлагается в качестве упражнения для читателя.

Закончим этот раздел доказательством того, что определитель матрицы $ABCD$ всегда равен 1. Доказательство основано на теореме Лиувилля.

Поскольку матрица $ABCD$ гауссова пучка определяет также характер поведения световых лучей, можно использовать наши знания лучевой оптики для изучения свойств матрицы.

Запишем (6.4.17) в виде

$$r_2 = Ar_1 + Br'_1, \quad (6.4.53)$$

$$r'_2 = Cr_1 + Dr'_1. \quad (6.4.54)$$

В интересующем нас случае при наличии двух измерений можно записать определитель (3.7.30) в форме, справедливой для параксиального приближения:

$$\begin{vmatrix} \frac{\partial r_2}{\partial r_1} & \frac{\partial r'_2}{\partial r_1} \\ \frac{\partial r_2}{\partial r'_1} & \frac{\partial r'_2}{\partial r'_1} \end{vmatrix} = 1. \quad (6.4.55)$$

Здесь r' — тангенс угла наклона луча. Используя соотношения (6.4.53) и (6.4.54) для определения производных в (6.4.55), получаем

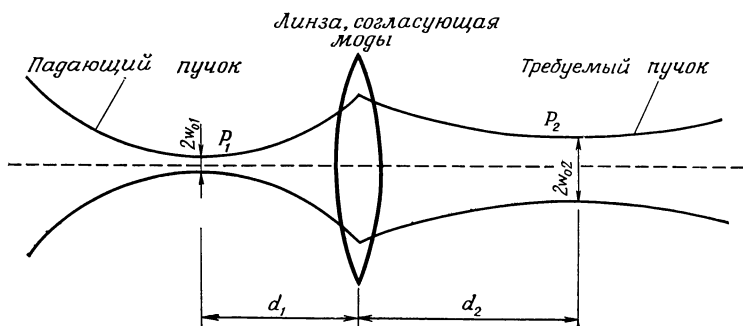
$$\begin{vmatrix} \frac{\partial r_2}{\partial r_1} & \frac{\partial r'_2}{\partial r_1} \\ \frac{\partial r_2}{\partial r'_1} & \frac{\partial r'_2}{\partial r'_1} \end{vmatrix} = \begin{vmatrix} A & C \\ B & D \end{vmatrix} = \begin{vmatrix} A & B \\ C & D \end{vmatrix} = 1.$$

Таким образом, мы установили факт, оказавшийся не случайным, что определители матриц, встречававшихся в этом разделе, равны единице.

6.5. СОГЛАСОВАНИЕ МОД

В предыдущем разделе мы видели, что гауссов пучок может быть введен в линзовый волновод. Если пучок не является модой волновода, то его поперечное сечение по мере распространения через волновод периодически меняется. Установлено, что период колебаний ширины пучка в 2 раза меньше периода колебания внеосевого луча. Если гауссов пучок выбран так, что его полуширина в наиболее узкой части определяется формулой (6.4.48), а полуширина на первой линзе дается формулой (6.4.47), то он распространяется через линзовый волновод без изменения параметров. Полуширина пучка на всех линзах в волноводе имеет то же значение, что и на первой линзе. Пучок такого вида согласован с линзовым волноводом и является его модой.

В большинстве реальных ситуаций экспериментатор имеет в своем распоряжении гауссов пучок, который образуется на выходе лазера. Мода линзового волновода или резонатора, которую мы желаем возбудить с помощью лазерного луча, к сожалению, чаще всего не совпадает с пучком, генерируемым лазером. Лазер и линзовый волновод (или другое устройство), таким образом, оказываются несогласованными. Однако всегда имеется возможность преобразовать с помощью линзы гауссов пучок



Ф и г. 6.5.1. Согласование гауссовых пучков с помощью линзы.

любого лазера. Мы видели, что гауссов пучок произвольной моды характеризуется двумя параметрами. Для данного рассмотрения выберем минимальную ширину пучка и расстояние от наиболее узкой части до выбранной точки отсчета. Эти два параметра однозначно характеризуют пучок. Соответствующая схема изображена на фиг. 6.5.1. Поскольку точка, от которой отсчитывается расстояние до наиболее узкого места пучка, выбрана, то можно выбрать и положение линзы, согласующей пучок. Минимальные значения параметров пучка w_{01} и w_{02} определяются характером задачи. Пусть мы имеем в распоряжении определенную линзу с фокусным расстоянием f . Проблема согласования мод заключается в определении расстояний d_1 и d_2 , где d_1 — расстояние от линзы до наиболее узкого места падающего пучка, а d_2 — расстояние от линзы до наиболее узкого места требуемого пучка.

Комплексный параметр пучка q принимает значение [см. (6.2.21)]

$$q_{01} = i \frac{\pi}{\lambda} w_{01}^2 = ib_1 \quad (6.5.1)$$

в точке P_1 и

$$q_{02} = i \frac{\pi}{\lambda} w_{02}^2 = ib_2 \quad (6.5.2)$$

в точке P_2 (см. фиг. 6.5.1). Цель задачи состоит в определении такого преобразования, для которого справедливо следующее выражение:

$$q_{02} = \frac{Aq_{01} + B}{Cq_{01} + D}. \quad (6.5.3)$$

Требуемая матрица $ABCD$ уже была составлена в разд. 6.4 [см. формулу (6.4.9)]. Элементы матрицы являются вещественными величинами, тогда как параметры q_{01} и q_{02} мнимые. Используя обозначения (6.5.1) и (6.5.2), можно разделить (6.5.3) на вещественную и мнимую части:

$$B + b_1 b_2 C = 0 \quad (6.5.4)$$

и

$$b_1 A - b_2 D = 0. \quad (6.5.5)$$

Подстановка элементов матрицы из (6.4.9) и выделение d_1 и d_2 из полученных уравнений приводят к следующим выражениям:

$$d_1 = f \pm \frac{w_{01}}{w_{02}} \sqrt{f^2 - f_0^2} \quad (6.5.6)$$

и

$$d_2 = f \pm \frac{w_{02}}{w_{01}} \sqrt{f^2 - f_0^2}, \quad (6.5.7)$$

где

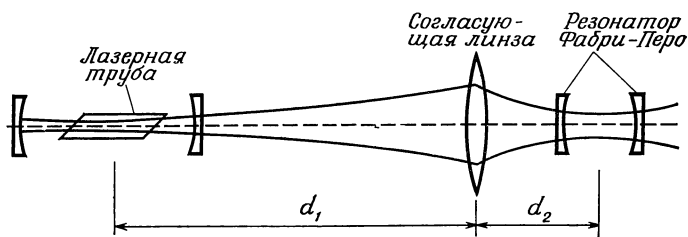
$$f_0 = \frac{\pi}{\lambda} w_{01} w_{02}. \quad (6.5.8)$$

Знаки правых частей должны быть или положительными, или оба отрицательными. При этом можно удовлетворить условию согласования мод двумя различными способами. Фокусное расстояние согласующей линзы произвольно. Необходимо только, чтобы

$$f \geq f_0, \quad (6.5.9)$$

т. е. фокусное расстояние не должно быть слишком малым.

Согласование мод особенно важно при работе со сканирующим интерферометром Фабри — Перо [70], который очень схож с лазерным резонатором. Он часто используется для контроля частоты многомодового лазера. Для этой цели резонатор интерферометра периодически меняет свою длину во времени. При точном согласовании свет может пройти через интерферометр только на своих резонансных частотах. По показаниям на выходе детектора,



Ф и г. 6.5.2. Согласование лазерного пучка с интерферометром Фабри — Перо посредством линзы.

который устанавливается за сканирующим интерферометром Фабри — Перо, с помощью осциллографа, который синхронизирован со сканирующим резонатором, можно на экране катодно-лучевой трубки получить спектр излучения на выходе лазера. Разрешение такого устройства чрезвычайно высоко.

Интересным с точки зрения нашего рассмотрения является факт многомодовости резонатора интерферометра. При отсутствии полного согласования входящий лазерный пучок возбуждает большое число мод в резонаторе. Каждая мода имеет свою резонансную частоту, поэтому каждая частотная компонента на выходе лазера представится в виде соответствующего сигнала на экране осциллографа. При работе в режиме многомодовости возможно использование сканирующего интерферометра Фабри — Перо в качестве анализатора спектра. Очевидно, что необходимо тщательное согласование мод лазера с модами резонатора интерферометра с тем, чтобы избежать многочастотности выходных сигналов. Согласование мод осуществляется установкой соответствующей линзы,

которая удовлетворяет условию (6.5.9). По отношению к лазерному пучку и интерферометру линза располагается в соответствии с формулами (6.5.6) и (6.5.7). Если нет других оптических элементов на пути лазерного пучка, наиболее узкая часть пучка находится в середине промежутка между двумя зеркалами при условии, что лазерный резонатор имеет два одинаковых зеркала. То же самое справедливо для пучка в резонаторе Фабри—Перо. Расстояния d_1 и d_2 должны быть, следовательно, измерены от центра лазерного резонатора и центра резонатора Фабри—Перо. Ширина пучка в сужениях определяется радиусами кривизны и расстоянием между зеркалами резонатора лазера и резонатора интерферометра Фабри—Перо. Схема согласования показана на фиг. 6.5.2. Заметим, что, вообще говоря, необходимо учитывать искажение пучка, имеющее место при прохождении пучка через плоские поверхности за пределами резонаторных зеркал. На фигуре это не отражено.

6.6. РЕЗОНАТОРЫ ЛАЗЕРА

Краткое обсуждение условий стабильности лазерных резонаторов в гл. 5.3 основывалось на результатах исследования передачи волноводов с неодинаковыми линзами.

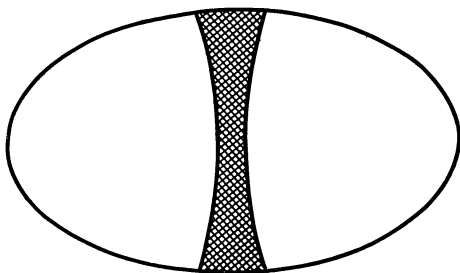
При рассмотрении типов колебаний (мод) лазерных резонаторов в этом разделе также будет использована аналогия между лазерными резонаторами и линзовыми волноводами. Однако, перед тем как воспользоваться результатами, полученными при исследовании линзовых волноводов, рассмотрим несколько аспектов теории лазерных резонаторов.

Е. А. Маркатили указал¹⁾ на то, что лазерный резонатор может быть рассмотрен как часть эллиптического объемного резонатора²⁾. Если рассмотреть колебания эллиптического резонатора, показанного на фиг. 6.6.1, то можно найти решения, для которых поле сосредоточено в заштрихованной области. В этой области содержится

¹⁾ Частное сообщение.

²⁾ Теория резонаторов с эллиптическими зеркалами изложена в монографии [1*].— *Прим. ред.*

основная часть энергии поля. Существуют, конечно, и другие моды, которые заполняют остальную часть объема эллиптического резонатора. Отсечка частей резонатора, где поле сосредоточенных в заштрихованной области мод отсутствует, не влияет на распределение поля этих мод. Данное рассмотрение поясняет связь между закрытыми резонаторами и открытыми резонаторами в лазерах.



Ф и г. 6.6.1. Эллиптический металлический резонатор, используемый для объяснения мод в открытом резонаторе лазера.

В разд. 5.6 было выведено интегральное уравнение для мод в линзовых волноводах. Чтобы показать взаимосвязь между двумя структурами, рассмотрим задачу о собственных значениях мод в лазерных резонаторах. Оказывается, что задача о колебаниях лазера математически идентична задаче о модах линзового волновода.

Рассмотрим лазерный резонатор, представленный на фиг. 6.6.2. Пусть существует поле на поверхности первого зеркала. Это поле распространяется ко второму зеркалу. Распространение поля от опорной плоскости зеркала 1 к опорной плоскости зеркала 2 описывается дифракционным интегралом (5.6.2). Распространение поля от изогнутой поверхности первого зеркала до его опорной плоскости может быть приближенно описано просто фазовым сдвигом опорной плоской волны при ее движении от поверхности зеркала до плоскости отсчета. Этот фазовый сдвиг равен

$$\phi = -ikd. \quad (6.6.1)$$

В параксиальном приближении d может быть выражено через радиус первого зеркала a , радиус кривизны его

поверхности R и расстояние r от рассматриваемой точки до оси структуры:

$$d = \frac{a^2 - r^2}{2R}. \quad (6.6.2)$$

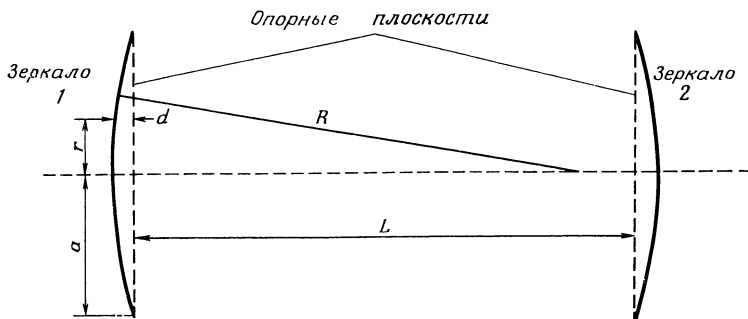
Радиус кривизны зеркала связан с фокусным расстоянием изогнутого отражателя соотношением

$$f = \frac{1}{2} R. \quad (6.6.3)$$

Поле ψ_1 на изогнутой поверхности зеркала 1 трансформируется в поле $\bar{\psi}_1$ в опорной плоскости в соответствии с выражением

$$\bar{\psi}_1 = e^{-i(k/4f)(a^2 - r^2)} \psi_1. \quad (6.6.4)$$

Аналогичное выражение связывает поле в опорной плоскости второго зеркала с полем на его изогнутой поверхности. Преобразование (6.6.4) оказывается таким же,



Ф и г. 6.6.2. Лазерный резонатор и опорные плоскости, используемые для расчета мод.

как преобразование (5.6.1), описывающее прохождение поля через половину линзы волновода.

Для того чтобы имели место колебания резонатора, потребуем, чтобы поле на первом зеркале совпадало с полем на втором зеркале с точностью до фазового сдвига. Это условие существования колебаний резонатора приводит снова к выражению (5.6.5), что доказывает математическую эквивалентность задачи о резонаторе задаче

о линзовом волноводе. Все моды резонатора автоматически совпадают с модами линзового волновода.

При формулировке задачи, связанной с модами в лазерных резонаторах, использовано параксиальное приближение, которое было также применено для получения мод в линзовом волноводе. В этом приближении сферические зеркала не отличаются от параболических зеркал. Такая аппроксимация справедлива только, когда кривизна зеркал мала и математически определяется условиями

$$\frac{a^2}{2R} \ll a \quad (6.6.5)$$

и

$$a \ll L. \quad (6.6.6)$$

Радиус a в этих соотношениях означает не радиус линзы, а размер поля, т. е. радиус, на котором поле спадает до малой величины по сравнению с величиной поля на оси.

Решение уравнения (5.6.5) является решением и задачи о модах в линзовом волноводе и задачи о модах в лазерном резонаторе. Приближенное решение уравнения собственных значений для больших зеркал или линз дано в (6.3.20). Для зеркал или линз конечных размеров было также рассмотрено несколько приближенных методов решения задачи (5.6.5) (см. [115] и [116]). Как показано в разд. 5.6, в случае конфокальных линз ($L = 2f$) существует точное решение. Аналитическое решение для конфокального случая с квадратными отверстиями представляется через сфероидальные волновые функции [47]. Решения для конфокального линзового волновода с круглыми отверстиями могут быть выражены с помощью гиперсфероидальных функций [120].

Численное решение общей задачи о собственных значениях (5.6.5) может быть получено методом последовательных приближений [48]. Этот метод заключается в следующем. Произвольная функция ψ подставляется в правую часть соотношения (5.6.5), и определяется (численно) величина интеграла, т. е. получается новая функция $\gamma\psi$. Конечно, величина γ не известна. Однако поскольку

постоянный множитель не влияет на форму функции, то тот факт, что величина γ не известна, не играет существенной роли. Первое приближение функции ψ используется, чтобы получить второе приближение новой оценкой интеграла. Этот процесс, если он повторяется несколько раз, обычно приводит к решению интегрального уравнения в первом порядке приближения. В математической постановке вопрос о том, является ли итерационный процесс сходящимся к определенной функции или нет, не должен нас особенно волновать, поскольку в тех практических случаях, когда пытались получить решение, это удавалось сделать. Кроме того, имеются физические причины, по которым процесс последовательного приближения должен быть сходящимся.

Процесс последовательного приближения может быть оправдан следующим образом. Пусть линзовый волновод возбуждается полем с некоторой структурой и передает его. С другой стороны, поле, которое возникает внутри лазерного резонатора, распространяется между его двумя зеркалами. В обоих случаях форма первичного распределения поля будет меняться либо при прохождении от линзы к линзе, либо при каждом прохождении от зеркала к зеркалу. Описание процесса распространения поля точно соответствует математическому процессу последовательного приближения, поскольку при вычислении мы применяем дифракционный интеграл (5.6.5), чтобы получить распределение поля на каждой последующей линзе или последующем зеркале резонатора. В системах с конечными апертурами часть мощности теряется после каждого прохождения волны через волновод или от одного зеркала до другого. Однако если проследить этот процесс достаточно далеко, то можно обнаружить, что в конце концов устанавливается распределение поля с наименьшим порядком структуры. То, что лазерные генераторы могут работать также на модах высокого порядка, вызвано механизмом усиления колебаний. Мода, которая обладает наибольшим усилением, имеет преимущество по сравнению с другими модами. Часто имеющее место достаточное усиление позволяет генерировать в резонаторе одновременно более чем одну моду. В отличие от этого в пассивных структурах с потерями мода низшего порядка, имеющая наиболее

низкие потери¹⁾, может быть получена методом последовательных приближений.

Такой процесс селекции мод может быть легко понят, если произвольное входное поле представить как линейную суперпозицию бесконечного множества мод. Когда поле проходит через линзовый волновод, форма каждой моды из общей суммы не изменяется. Но после прохождения от одной линзы до другой каждая мода умножается на свое собственное значение. После прохождения через n линз v -я мода оказывается умноженной на γ_v^n . Величины γ_v^n становятся бесконечно малыми для больших значений n . Только для моды низшего порядка величина γ_v остается еще конечной [ср. формулу (5.7.9)].

Последние соображения, основанные на представлении входного поля в виде суперпозиции мод структуры, являются в некоторой степени некорректными, поскольку используется понятие моды для объяснения того, почему поле моды может быть получено методом последовательного приближения²⁾. Однако такой подход помогает понять, почему мода низшего порядка, имеющая наименьшие потери, «выживает», пройдя через несколько линз, и почему любое произвольное распределение поля преобразуется в итоге в моду низшего порядка. Только в случае, когда амплитудный коэффициент при моде низшего порядка в первоначальном разложении равен нулю, получаем следующую моду с наиболее низкими потерями. Эксперимент показал, что при возбуждении волновода полем с произвольным распределением всегда устанавливается мода с низшими потерями. Экспериментально также установлено, что моды всегда существуют в периодических и почти периодических структурах (линзовые волноводы могут рассматриваться как периодические структуры). Именно практический опыт дает основание считать, что

¹⁾ Речь идет о дифракционных потерях из-за конечности апертуры линз или зеркал. — *Прим. ред.*

²⁾ Некорректность здесь состоит в том, что объяснение существования мод (определенных устойчивых структур поля) основывается на представлении произвольного поля в виде суперпозиции этих мод, т. е. априори допускается существование этих мод. — *Прим. ред.*

моды должны существовать даже тогда, когда это трудно доказать математически.

Вернемся к обсуждению итерационного процесса. Повторение итерации до тех пор, пока изменение формы поля перестанет наблюдаться, позволяет найти не только форму решения низшего порядка, но и соответствующее ему собственное значение. Для того чтобы получить величину γ , необходимо проследить уменьшение амплитуды поля и изменение его фазы после нескольких итераций. Итерационное решение задачи о собственных значениях (5.6.5) было получено Фоксом и Ли [48]. Некоторые их результаты для конфокальных линзовых волноводов или лазерных резонаторов с круглыми отверстиями приведены на фиг. 5.6.5. Моды высшего порядка были получены тем же методом с использованием первоначального распределения поля, не содержащего моды низшего порядка. Для больших значений s (5.6.30) итерационный процесс сходится медленно. Требуется большое количество итераций, и численный анализ становится очень трудоемким. Другие аппроксимирующие методы получения собственных значений и собственных функций состоят в преобразовании интегрального уравнения (5.6.5) к системе алгебраических уравнений, которая затем решается стандартными матричными методами [116].

Читателя может удивить, почему при обсуждении колебаний резонатора до сих пор не было никакого упоминания об условии резонанса. Решение уравнения собственных значений методом последовательных приближений не включает условия резонанса. Однако истинная ситуация в физическом резонаторе незначительно отличается от математических допущений, которые ведут к задаче (5.6.5). Поле имеется не только на зеркалах резонатора, но и в пространстве между ними. Действительно нужно рассмотреть поле между зеркалами, которое распространяется до тех пор, пока не отразится от второго зеркала и затем интерферирует само с собой. При такой интерференции накапливается энергия в резонаторе. Эта более полная картина процесса работы резонатора приводит к условию резонанса. Здесь необходимо потребовать, чтобы после полного цикла полю сообщалось следующее приращение в фазе по отношению к

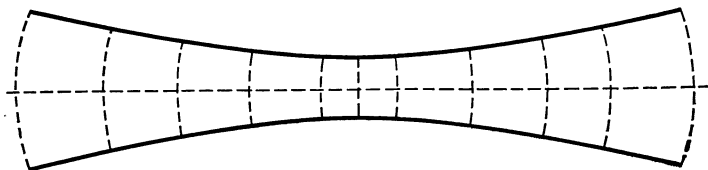
исходному полю:

$$2\Delta\phi = 2N\pi. \quad (6.6.7)$$

Фазовый сдвиг $\Delta\phi$ добавляется при прохождении волны от одного зеркала до другого. Удвоенный фазовый сдвиг, или фазовый сдвиг за один полный цикл, должен быть равен целому числу, умноженному на 2π .

Накопление энергии поля в реальном лазерном резонаторе может быть пояснено тем же методом, который был описан выше — методом последовательных приближений. Атомы усиливающей среды лазера испускают энергию спонтанно. Небольшое количество излученной энергии движется, например, направо, достигает одного из зеркал и отражается назад в область усиления. Далее это излучение пересекает область усиления между зеркалами и достигает противоположного зеркала несколько усиленным. Волна, которая проходит весь цикл с наибольшим усилением, растет быстрее, чем другие типы волн (моды) в резонаторе. Максимальная энергия усиливающейся моды ограничивается процессом насыщения. Усиливающая среда обеспечивает линейное усиление только для относительно слабых электромагнитных полей. По мере роста поля достигается режим насыщения усиливающей среды и усиление уменьшается. Поле внутри резонатора может существовать в течение произвольного отрезка времени только тогда, когда оно за период колебания увеличивает свою энергию на большую величину, чем теряет из-за потерь. Только определенные типы волн получают достаточное усиление, чтобы их потери после достижения насыщения усиливающей среды компенсировались. Большая часть мод имеет большие потери, чем усиление, даваемое насыщенной усиливающей средой. Эти моды затухают, практически не возбудившись. Таким образом, относительно небольшое число мод поддерживается любым лазером. Введением апертур на пути лазерного луча внутри резонатора можно увеличить потери для мод высшего порядка настолько, чтобы возбудилась только мода низшего порядка. Каждое решение уравнения собственных значений (5.6.5) для моды низшего порядка справедливо для нескольких частот, каждая из которых удовлетворяет условию резонанса (6.6.7). Поэтому обычно

существует несколько так называемых продольных мод резонатора, которые соответствуют решениям уравнения собственных значений (5.6.5) при одинаковой форме поля в поперечной плоскости (такая форма поля часто называется поперечной модой), но несколько отличающихся резонансных частотах. Число продольных мод, которые могут одновременно возбуждаться, зависит от ширины полосы частот кривой усиления. Частотное разделение продольных мод резонатора определяется расстоянием между зеркалами резонатора.



Ф и г. 6.6.3. Гауссов пучок.

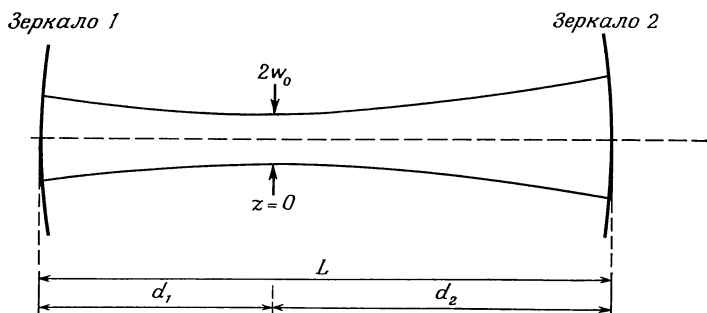
Сплошные линии показывают ширину пучка, пунктирные — поверхности постоянной фазы волны.

Разнесение зеркал приводит к сближению резонансных частот резонатора. Если желательно работать на одном типе колебаний, то необходимо сконструировать такой короткий лазер, чтобы только один тип колебаний соответствовал кривой усиления лазерной среды. Однако такой метод уменьшает усиление лазера. Если желательно иметь на выходе большую мощность и требуется работа на одной частоте, то должны быть использованы другие методы селекции мод [71].

Большинство лазеров работает с достаточно большими апертурами зеркал, так что структура их мод хорошо описывается аппроксимацией Эрмита — Гаусса, обсуждавшейся в предыдущем разделе.

Мы показали в разд. 6.5, что пучки Эрмита — Гаусса имеют область наибольшего сужения и расширяются от этой области в обоих направлениях. Фазовый фронт волны пучка плоский в области сужения и приблизительно сферический во всех других частях вдоль пучка. Типичный гауссов пучок показан на фиг. 6.6.3. Сплошные линии отмечают ширину луча, а пунктирные линии,

перпендикулярные оси, показывают фазовые фронты в различных точках вдоль пучка. Пучок Эрмита — Гаусса является приближенным решением волнового уравнения. Таким образом, можно получить приемлемое решение задачи о колебаниях резонатора с большими зеркалами, помещая зеркала в любом месте на пути пучка. Необходимо только, чтобы поверхность зеркал повторяла форму фазовых фронтов. В этом случае поле при отражении назад перейдет само в себя и сформируется соответствующая



Ф и г. 6.6.4. Лазерный резонатор с зеркалами различной кривизны.

мода резонатора. Данное рассмотрение показывает, что большое количество резонаторов пригодно для формирования пучка Эрмита — Гаусса. Размещая зеркала соответствующей кривизны в произвольном месте на пути луча, получим приемлемый для данной моды лазерный резонатор. Нет необходимости в том, чтобы оба зеркала имели одинаковую кривизну, если их кривизна соответствует кривизне фазового фронта пучка Эрмита — Гаусса в месте их установки. В действительности, конечно, именно зеркала формируют колебания пучка Эрмита — Гаусса. Но для того, чтобы проследить процесс формирования, удобно поменять местами причину и следствие и считать, что пучок существует между зеркалами, предназначенными отражать его самого в себя.

Типичный лазерный резонатор показан на фиг. 6.6.4. Зеркала, конечно, могут быть выбраны одинаковой

кривизны или одно может быть вогнутым, а другое — выпуклым. В последнем случае зеркала должны быть расположены с одной стороны от области сужения пучка.

Фазовый сдвиг пучка Эрмита — Гаусса на оси ($r = 0$) для одного пробега через резонатор, согласно (6.3.20), равен

$$\Delta\phi = \phi(d_2) - \phi(d_1) = kL - (m + n + 1) \left(\arctg \frac{\lambda d_2}{\pi w_0^2} + \arctg \frac{\lambda d_1}{\pi w_0^2} \right). \quad (6.6.8)$$

Зеркала расположены на расстояниях d_1 и d_2 от области сужения пучка: $L = d_1 + d_2$. Условие резонанса вытекает из равенства (6.6.7):

$$2\pi \frac{L}{\lambda} - (m + n + 1) \left(\arctg \frac{\lambda d_2}{\pi w_0^2} + \arctg \frac{\lambda d_1}{\pi w_0^2} \right) = N\pi, \quad (6.6.9)$$

где $k = 2\pi/\lambda$. Длина волны в этих соотношениях является не длиной волны моды гауссова пучка, а длиной волны, которую имела бы соответствующая плоская волна, если бы она распространялась в той же среде и с той же частотой, что и мода гауссова пучка. Длина волны гауссова пучка могла бы быть определена как расстояние вдоль оси, которое соответствует фазовому сдвигу 2π . Эта величина, однако, не постоянна и зависит, хотя и незначительно, от положения вдоль оси моды Эрмита — Гаусса.

Условие резонанса (6.6.9) можно выразить через радиусы кривизны лазерных зеркал. Чтобы получить соответствующую формулу, выразим расстояния d_1 , d_2 от плоскости сужения пучка до зеркал, полуширины пучка w_1 и w_2 на зеркалах 1 и 2, полуширину w_0 пучка в области сужения через радиусы кривизны R_1 , R_2 и расстояние L между двумя зеркалами. Зависимость между радиусами кривизны фазовых фронтов, которые равны радиусам кривизны зеркал, и расстояниями зеркал от плоскости сужения пучка задается формулой (6.2.13), откуда

$$R_1 = d_1 + \left(\frac{\pi w_0^2}{\lambda} \right)^2 \frac{1}{d_1} \quad (6.6.10)$$

и

$$R_2 = d_2 + \left(\frac{\pi w_0^2}{\lambda} \right)^2 \frac{1}{d_2}. \quad (6.6.11)$$

Исключение w_0 из этих уравнений дает

$$\frac{R_1 - d_1}{R_2 - d_2} = \frac{d_2}{d_1}. \quad (6.6.12)$$

Используя равенство

$$d_1 + d_2 = L, \quad (6.6.13)$$

сразу же получим

$$d_1 = \frac{L(R_2 - L)}{R_1 + R_2 - 2L} \quad (6.6.14)$$

и

$$d_2 = \frac{L(R_1 - L)}{R_1 + R_2 - 2L}. \quad (6.6.15)$$

Подстановка выражений (6.6.14) и (6.6.15) в (6.6.10) или (6.6.11) дает возможность получить полуширину пучка в наиболее узком его месте:

$$w_0^4 = \left(\frac{\lambda}{\pi} \right)^2 \frac{L(R_1 - L)(R_2 - L)(R_1 + R_2 - L)}{(R_1 + R_2 - 2L)^2}. \quad (6.6.16)$$

В частном случае при $R_1 = R_2 = 2f$ выражение (6.6.16) становится идентичным выражению (6.4.48). Подстановка выражений (6.6.14) и (6.6.16) в (6.2.19) (где $z = -d_1$) дает

$$w_1^4 = \left(\frac{\lambda R_1}{\pi} \right)^2 \frac{L(R_2 - L)}{(R_1 - L)(R_1 + R_2 - L)}. \quad (6.6.17)$$

Аналогично используя выражение (6.6.15) или просто меняя индексы 1 и 2 местами, получим

$$w_2^4 = \left(\frac{\lambda R_2}{\pi} \right)^2 \frac{L(R_1 - L)}{(R_2 - L)(R_1 + R_2 - L)}. \quad (6.6.18)$$

Формулы (6.6.14) — (6.6.18) дают возможность выразить параметры гауссова пучка через параметры резонатора.

Эти выражения могут быть использованы для получения условия резонанса (6.6.9) в форме, включающей в явном виде параметры резонатора. Используем теорему сложения арктангенсов

$$\arctg \alpha + \arctg \beta = \arccos \left(\frac{1 - \alpha\beta}{\sqrt{1 + \alpha^2 + \beta^2 + \alpha^2\beta^2}} \right), \quad (6.6.19)$$

где

$$\alpha = \frac{\lambda}{\pi} \frac{d_1}{w_0^2} = \sqrt{\frac{L(R_2 - L)}{(R_1 - L)(R_1 + R_2 - L)}} \quad (6.6.20)$$

и

$$\beta = \frac{\lambda}{\pi} \frac{d_2}{w_0^2} = \sqrt{\frac{L(R_1 - L)}{(R_2 - L)(R_1 + R_2 - L)}}. \quad (6.6.21)$$

После алгебраических преобразований запишем соотношение (6.6.19) в следующем виде:

$$\begin{aligned} \operatorname{arctg} \left(\frac{\lambda}{\pi} \frac{d_1}{w_0^2} \right) + \operatorname{arctg} \left(\frac{\lambda}{\pi} \frac{d_2}{w_0^2} \right) = \\ = \arccos \sqrt{\left(1 - \frac{L}{R_1} \right) \left(1 - \frac{L}{R_2} \right)}. \end{aligned} \quad (6.6.22)$$

Условие резонанса (6.6.19) примет вид

$$\frac{2L}{\lambda} - \frac{m+n+1}{\pi} \arccos \sqrt{\left(1 - \frac{L}{R_1} \right) \left(1 - \frac{L}{R_2} \right)} = N. \quad (6.6.23)$$

Резонансная частота, следовательно, равна [49, 72]

$$\nu = \frac{c}{2L} \left\{ N + \frac{m+n+1}{\pi} \arccos \sqrt{\left(1 - \frac{L}{R_1} \right) \left(1 - \frac{L}{R_2} \right)} \right\}. \quad (6.6.24)$$

Для конфокального резонатора с

$$R_1 = R_2 = L \quad (6.6.25)$$

выражение (6.6.24) преобразуется к виду

$$\nu = \frac{c}{2L} \left[N + \frac{1}{2} (m+n+1) \right]. \quad (6.6.26)$$

Целое число N определяет число полуволин, которое укладывается между зеркалами. Число N является, таким образом, порядком продольных мод резонатора. Целые числа n и m определяют порядок поперечной вариации распределения поля эрмита — гауссовой моды. При $n = m = 0$ получается чисто гауссова мода.

Важно помнить, что эрмита-гауссовы моды могут быть использованы при рассмотрении резонаторов с большими зеркалами, для которых дифракционные потери незначительны. Если зеркала оказываются малыми, так что дифракционные потери должны быть учтены, то аппроксимация поля моды эрмита-гауссовым колебанием в таком резонаторе будет несправедлива. Решение в этом случае

должно быть получено путем численного решения задачи о собственных значениях (5.6.5) (см. [48]).

Проблема стабильности колебаний лазерного резонатора обсуждалась в разд. 5.2. На фиг. 5.2.4 заштрихована область стабильности для колебаний лазерного резонатора. За пределами этой области потери моды становятся очень большими. Однако в особых случаях может быть использован и режим работы резонаторов в нестабильной области [417].

РАСПРОСТРАНЕНИЕ СВЕТА В СРЕДАХ С КВАДРАТИЧНЫМ ЗАКОНОМ ИЗМЕНЕНИЯ ПОКАЗАТЕЛЯ ПРЕЛОМЛЕНИЯ

7.1. ВВЕДЕНИЕ

В гл. 5 были рассмотрены свойства линзового волновода, способного передавать электромагнитные волны оптического диапазона. Такой волновод, однако, не является единственно возможным средством передачи световых лучей. Если расстояние между линзами уменьшить до нуля, то получим непрерывную направляющую среду. Такая среда может быть описана с помощью переменного показателя преломления, который имеет максимальное значение на оси и уменьшается монотонно при удалении от оптической оси волновода. Среда с показателем преломления вида

$$n = n_0 - \frac{1}{2} n_1 (x^2 + y^2) \quad (7.1.1)$$

называется квадратичной средой. Распространение электромагнитных волн оптического диапазона в квадратичной среде весьма сходно с распространением волн в линзовом волноводе [69, 76, 77, 78].

Оптические волноводы со средой, показатель преломления которой меняется по квадратичному закону, будут иметь важное практическое значение, когда появятся более совершенные методы их построения. Но уже и существующие оптические волокна имеют показатель преломления, изменяющийся по закону, близкому к квадратичному [23, 112]. Целью настоящей главы является ознакомление читателя с вопросами распространения света в непрерывных диэлектрических средах с квадратичным изменением параметров.

Квадратичные среды обладают тем свойством, что эффективный центр мощности параксиальных пучков движется в них по законам оптических лучей (см. разд. 3.6). Геометрооптический метод исследования таких структур физически оправдан. Он дает основные представления о свойствах распространения световых пучков.

Модель среды с квадратичным законом вида (7.1.1) имеет один серьезный недостаток. Для достаточно больших x и y показатель преломления становится меньше единицы и даже достигает отрицательных значений. Хотя существуют среды, показатель преломления которых может быть меньше единицы, например ионизированные газы, все же для большинства физических сред показатель преломления оказывается больше единицы. Поэтому необходимо учитывать, что распределение показателя преломления (7.1.1) может достигать нефизической области. До тех пор пока световые лучи не достигают области с $n < 1$ или пока поле внутри этой области мало, можно использовать квадратичный показатель преломления (7.1.1) как хорошее первое приближение. При этом следует ограничиваться рассмотрением параксиальных лучей. Параксиальные лучи распространяются по траекториям, которые всегда почти параллельны оптической оси. Эти лучи с меньшей вероятностью могут удаляться от оптической оси, чем лучи с большим углом наклона к оси.

Среда с квадратичным законом изменения показателя преломления (7.1.1) может рассматриваться как первое приближение для среды с симметричным законом распределения показателя преломления. Если разложить произвольное распределение показателя преломления в ряд Тейлора, то можно получить члены порядка выше второго. Но во многих практических случаях ряды Тейлора сходятся быстро, так что первых двух членов достаточно для первого приближения. Можно еще считать, что коэффициенты при x^2 и y^2 разные. Однако поскольку такое обобщение не дает больших преимуществ, то можно ограничиться распределением показателя преломления вида (7.1.1). Как было показано, среда с законом изменения показателя преломления $1/ch\ x$ является идеальной фокусирующей структурой даже для непараксиальных лучей

[75, 118, 119]. Однако здесь мы ограничимся рассмотрением только квадратичной среды и лучей в ней в параксиальном приближении.

7.2. ЛУЧЕВАЯ ОПТИКА КВАДРАТИЧНОЙ СРЕДЫ

Выведем траекторию луча в квадратичной среде, беря в качестве исходной траекторию луча в линзовом волноводе с бесконечно близко расположенными линзами. В уравнении (5.2.4) устремим L к нулю и запишем

$$\frac{r_{n+1} - r_n}{L} = \frac{dr}{dz} \quad (7.2.1)$$

и

$$\frac{(r_{n+1} - r_n)/L - (r_n - r_{n-1})/L}{L} = \frac{d^2r}{dz^2}. \quad (7.2.2)$$

Разностное уравнение (5.2.4) можно, таким образом, записать в виде

$$\frac{d^2r}{dz^2} = -\frac{1}{Lf} r. \quad (7.2.3)$$

Допустим теперь, что $f \rightarrow \infty$, так что

$$\frac{1}{Lf} = \frac{n_1}{n_0}. \quad (7.2.4)$$

Этот предельный переход переводит разностное уравнение для траектории луча линзового волновода в параксиальное дифференциальное уравнение

$$\frac{d^2r}{dz^2} = -\frac{n_1}{n_0} r \quad (7.2.5)$$

для квадратичной среды. Дифференциальное уравнение (7.2.5), по-видимому, идентично параксиальному дифференциальному уравнению (3.5.36), в котором выражение (7.1.1) включено в правую часть. Таким образом, можно сделать вывод, что линзовый волновод становится квадратичной средой в случае бесконечно близко расположенных линз. Решение для траектории луча линзового волновода может быть преобразовано подобным же образом к решению для случая квадратичной среды. Из формулы (5.2.6) получим $\cos \theta = 1$ в пределе при $L \rightarrow 0$ и $f \rightarrow \infty$, и,

кроме того,

$$\sin \theta = \theta = \sqrt{\frac{L}{f}} = \sqrt{\frac{n_1}{n_0}} L. \quad (7.2.6)$$

Используя выражение

$$z = nL, \quad (7.2.7)$$

получим из формулы (6.4.26)

$$r = r_0 \cos \sqrt{\frac{n_1}{n_0}} z + \sqrt{\frac{n_0}{n_1}} r'_0 \sin \sqrt{\frac{n_1}{n_0}} z. \quad (7.2.8)$$

Очевидно, что (7.2.8) является решением уравнения параксиальных лучей (7.2.5). Положение луча r_1 и наклон¹⁾ r'_1 выражены здесь соответственно через r_0 и r'_0 , чтобы показать, что они являются функциями положения и наклона при $z = 0$. Приведенный вывод траектории луча в квадратичной среде подчеркивает близкую связь между такой средой и линзовым волноводом. Разумеется, уравнение (7.2.5) и его решение (7.2.8) могут быть получены из уравнения параксиальных лучей (3.5.36) и формулы (7.1.1) для показателя преломления без ссылки на линзовый волновод. Итак, можно сделать вывод, что среда с квадратичным законом изменения показателя преломления ведет себя как световод.

Можно найти также траекторию луча в квадратичной среде с искривленной осью, используя результаты, полученные для линзового волновода. Из формул (5.4.8) и (5.4.9) с помощью (7.2.6) и (7.2.7) и с учетом того, что $dz = L$ при $L \rightarrow 0$, получим

$$r(z) = r_0 \cos \sqrt{\frac{n_1}{n_0}} z + \sqrt{\frac{n_0}{n_1}} r'_0 \sin \sqrt{\frac{n_1}{n_0}} z + \\ + \sqrt{\frac{n_0}{n_1}} \int_0^z \frac{1}{R(u)} \sin \sqrt{\frac{n_1}{n_0}} (z-u) du. \quad (7.2.9)$$

Решение однородного уравнения [первые два члена (7.2.9)] записано в форме (7.2.8). Функция $R(u)$ является радиусом кривизны изгиба оси волновода, образованного квадратичной средой, при $z = u$.

¹⁾ Величина r' является тангенсом угла наклона траектории. — Прим. ред.

Используя выражение (5.4.15) для линзового волновода с искривленной осью, получим для квадратичной среды с искривленной осью

$$\rho(z) = \rho_0 \cos \sqrt{\frac{n_1}{n_0}} z + \sqrt{\frac{n_0}{n_1}} \rho'_0 \sin \sqrt{\frac{n_1}{n_0}} z + \\ + \sqrt{\frac{n_1}{n_0}} \int_0^z S(u) \sin \sqrt{\frac{n_1}{n_0}} (z-u) du. \quad (7.2.10)$$

Функция $S(z)$ описывает отклонение оси квадратичной среды от прямой линии.

Эквивалентность двух решений (7.2.9) и (7.2.10) для траектории луча в квадратичной среде с изогнутой осью может быть легко продемонстрирована. Радиус кривизны изгиба оси квадратичной среды связан с линейным отклонением оси от идеальной прямой $S(z)$ следующим приближенным соотношением:

$$\frac{1}{R(z)} = -\frac{d^2 S}{dz^2}. \quad (7.2.11)$$

Это приближение справедливо, если первая производная dS/dz много меньше единицы. Отрицательный знак в (7.2.11) соответствует положительному радиусу кривизны на фиг. 5.4.2. Подстановка (7.2.11) в (7.2.9) приводит после интегрирования по частям к результату

$$r(z) = r_0 \cos \sqrt{\frac{n_1}{n_0}} z + \sqrt{\frac{n_0}{n_1}} r'_0 \sin \sqrt{\frac{n_1}{n_0}} z - S(z) + \\ + S(0) \cos \sqrt{\frac{n_1}{n_0}} z + \sqrt{\frac{n_0}{n_1}} \left(\frac{dS}{dz} \right)_{z=0} \sin \sqrt{\frac{n_1}{n_0}} z + \\ + \sqrt{\frac{n_1}{n_0}} \int_0^z S(u) \sin \sqrt{\frac{n_1}{n_0}} (z-u) du. \quad (7.2.12)$$

В соответствии с формулой (5.4.10) имеем соотношение

$$\rho(z) = r(z) + S(z). \quad (7.2.13)$$

Отсюда видно, что выражение (7.2.12) идентично (7.2.10), что показывает эквивалентность этих двух решений.

Среднеквадратичное отклонение луча σ , вызванное случайными отклонениями оси квадратичной среды, по-

лучим из формулы (5.5.36). Это выражение является аппроксимацией более общего выражения (5.5.35). Однако в большинстве практических случаев это приближение хорошо оправдывается и удобно для использования. В пределе для очень малых промежутков между линзами выражение (5.5.36) принимает вид

$$\sigma^2(z) = \frac{1}{2} \frac{n_1}{n_0} z \int_{-\infty}^{\infty} R_c(u) \cos \sqrt{\frac{n_1}{n_0}} u du. \quad (7.2.14)$$

Функция корреляции смещения оптической оси $S(z)$ определяется как среднее по ансамблю произведение $S(z)$ и $S(z - u)$

$$R_c(u) = \langle S(z) S(z - u) \rangle. \quad (7.2.15)$$

Выражение (7.2.14) показывает, что дисперсия отклонения пучка пропорциональна расстоянию вдоль волновода, на котором пучок еще наблюдается. Кроме того, он пропорционален преобразованию Фурье корреляционной функции. Преобразование Фурье корреляционной функции содержит только члены с косинусами, так как $R_c(u)$ — четная функция.

Поведение квадратичной среды во всех отношениях очень похоже на поведение линзового волновода. Поэтому нет необходимости в детальном рассмотрении траектории луча в квадратичной среде. Все результаты для квадратичной среды могут быть получены из соответствующих результатов для линзового волновода.

7.3. МОДЫ КВАДРАТИЧНОЙ СРЕДЫ

Рассмотрим волноводные свойства квадратичной среды с точки зрения волновой оптики. Строго говоря, нужно найти решение уравнений Максвелла в среде с показателем преломления (7.1.1). Однако, как видно из разд. 1.3, скалярному волновому уравнению (1.3.6) должна удовлетворять каждая составляющая электромагнитного поля. Для оптической среды с переменным показателем преломления волновое уравнение не является точным эквивалентом уравнений Максвелла. Приближение, которое имеет место при использовании волнового уравнения вместо

уравнений Максвелла, является приемлемым, если показатель преломления (или диэлектрическая проницаемость) меняется незначительно на расстоянии порядка длины оптической волны. Это требование означает, что величина R (1.3.28) должна быть много меньше единицы. Такое требование с физической точки зрения также естественно, как и параксиальное приближение, которое было использовано при решении задачи о распространении луча в квадратичной среде. Если нас удовлетворяют решения, которые вытекают из параксиального приближения, то для определения свойств распространения волны в квадратичной среде вместо уравнений Максвелла можно использовать волновое уравнение. Таким образом, нужно рассмотреть распространение скалярных волн в квадратичной среде [68, 69]. Скалярная волновая задача является хорошим приближением к задаче об электромагнитном поле, пока не рассматриваются поляризационные эффекты. Известно, что каждая составляющая электромагнитного поля должна быть приближенным решением скалярного волнового уравнения. Скалярная волновая задача в квадратичной среде применима также к проблеме распространения звука в океане. В океане имеются слои с такими изменениями плотности, которые приводят к появлению направленных звуковых волн. В первом приближении можно допустить, что на некоторой глубине океан подобен квадратичной среде.

Предположив, что волны имеют гармонический во времени характер, можно использовать приведенное волновое уравнение в форме

$$\nabla^2 \psi + n^2 k_0^2 \psi = 0 \quad (7.3.1)$$

с постоянной распространения в свободном пространстве

$$k_0 = \omega \sqrt{\epsilon_0 \mu_0} = \frac{2\pi}{\lambda} \quad (7.3.2)$$

и относительной диэлектрической проницаемостью (квадрат показателя преломления)

$$n^2 = n_0^2 - n_0 n_1 (x^2 + y^2). \quad (7.3.3)$$

Если считать исходным показатель преломления (7.1.1), то (7.3.3) является приближенным соотношением. С другой стороны, можно также считать выражение (7.3.3)

точным и рассматривать (7.1.4) как некоторое приближение. Такой подход имеет то преимущество, что n^2 , а не n является в данном случае основной величиной. Попробуем решить приведенное волновое уравнение с помощью пробной функции

$$\psi = f(x) g(y) e^{-i\beta z}. \quad (7.3.4)$$

Здесь не известны не только функции f и g , но также постоянная распространения β . Решение вида (7.3.4) есть мода квадратичной среды. Подставляя выражения (7.3.3) и (7.3.4) в (7.3.1), получаем

$$\left(\frac{1}{f} \frac{\partial^2 f}{\partial x^2} + n_0^2 k_0^2 - \beta^2 - k_0^2 n_0 n_1 x^2 \right) + \\ + \left(\frac{1}{g} \frac{\partial^2 g}{\partial y^2} - k_0^2 n_0 n_1 y^2 \right) = 0. \quad (7.3.5)$$

Это уравнение должно быть справедливым для всех значений x и y , что возможно лишь тогда, когда зависящая от x и зависящая от y части порознь равны константе. Сумма этих двух постоянных должна равняться нулю. Таким образом, получаем два уравнения

$$\frac{d^2 f}{dx^2} + (n_0^2 k_0^2 - \kappa^2 - \beta^2 - k_0^2 n_0 n_1 x^2) f = 0 \quad (7.3.6)$$

и

$$\frac{d^2 g}{dy^2} + (\kappa^2 - k_0^2 n_0 n_1 y^2) g = 0. \quad (7.3.7)$$

Чтобы преобразовать эти два уравнения к хорошо известной стандартной форме, введем следующие новые переменные и параметры:

$$\xi = \sqrt{k_0} (n_0 n_1)^{1/4} x, \quad (7.3.8)$$

$$\eta = \sqrt{k_0} (n_0 n_1)^{1/4} y, \quad (7.3.9)$$

$$\sigma = \frac{n_0^2 k_0^2 - \kappa^2 - \beta^2}{k_0 \sqrt{n_0 n_1}} \quad (7.3.10)$$

и

$$\rho = \frac{\kappa^2}{k_0 \sqrt{n_0 n_1}}. \quad (7.3.11)$$

Дифференциальные уравнения запишем теперь в виде

$$\frac{d^2 f}{d\xi^2} + (\sigma - \xi^2) f = 0 \quad (7.3.12)$$

и

$$\frac{d^2 g}{d\eta^2} + (\rho - \eta^2) g = 0. \quad (7.3.13)$$

Дифференциальные уравнения (7.3.12) и (7.3.13) хорошо исследованы в теории гармонического осциллятора в квантовой механике [22]. Они определяют энергетическую собственную функцию и собственные значения, которые допустимы для гармонического осциллятора. В нашей теории, так же как и в теории гармонического осциллятора в квантовой механике, необходимо потребовать выполнения граничных условий. Моды в квадратичной среде должны распространяться вблизи оси структуры. Это значит, что нужно потребовать, чтобы функции стремились к нулю при $x \rightarrow \infty$ и $y \rightarrow \infty$. Такие же требования накладываются на волновую функцию гармонического осциллятора. Они играют роль граничных условий. В работах по квантовой механике показано, что ограниченные решения такого типа могут существовать только при следующих условиях:

$$\sigma = 2p + 1, \quad \text{где } p = 0, 1, 2, 3, \dots, \quad (7.3.14)$$

$$\rho = 2q + 1, \quad \text{где } q = 0, 1, 2, 3, \dots \quad (7.3.15)$$

Эти условия вместе с (7.3.10) и (7.3.11) определяют возможные значения постоянной распространения β . Они, следовательно, являются собственными значениями уравнений мод квадратичной среды. Мы еще вернемся к этому важному понятию. Решения дифференциальных уравнений (7.3.12) и (7.3.13) представляют собой хорошо известные функции Эрмита — Гаусса, которые уже встречались нам неоднократно. Чтобы показать это, положим

$$f = H_p(\xi) e^{-(1/2)\xi^2} \quad (7.3.16)$$

и

$$g = H_q(\eta) e^{-(1/2)\eta^2}. \quad (7.3.17)$$

Подстановка в (7.3.12) и (7.3.13) дает

$$\frac{d^2 H_p}{d\xi^2} - 2\xi \frac{dH_p}{d\xi} + 2pH_p(\xi) = 0 \quad (7.3.18)$$

и

$$\frac{d^2 H_q}{d\eta^2} - 2\eta \frac{dH_q}{d\eta} + 2qH_q(\eta) = 0. \quad (7.3.19)$$

Эти уравнения идентичны дифференциальному уравнению (6.3.13) для полиномов Эрмита. Таким образом, показано, что модовыми решениями для волн в квадратичной среде опять-таки являются функции Эрмита — Гаусса. Этот результат едва ли является неожиданным, так как мы видели, что квадратичная среда аналогична линзовому волноводу, моды которого описываются функциями Эрмита — Гаусса. Общее решение для мод в квадратичной среде есть, следовательно,

$$\psi_{pq}(x, y, z) = \frac{\sqrt{2/\pi}}{\sqrt{2^{p+q} p! q!} w} H_p \left(\sqrt{2} \frac{x}{w} \right) H_q \left(\sqrt{2} \frac{y}{w} \right) \times \\ \times e^{-(x^2+y^2)/w^2} e^{-i\beta_{pq} z} \quad (7.3.20)$$

с полушириной модового пучка

$$w^2 = \frac{2}{k_0 \sqrt{n_0 n_1}}. \quad (7.3.21)$$

Модовая функция (7.3.20) здесь надлежащим образом нормирована.

Выражение для постоянной распространения β_{pq} получается из формул (7.3.10), (7.3.11), (7.3.14) и (7.3.15):

$$\beta_{pq} = \{n_0^2 k_0^2 - k_0 \sqrt{n_0 n_1} [(2p+1) + (2q+1)]\}^{1/2}. \quad (7.3.22)$$

Для сравнения с модами линзового волновода положим в формулах (6.4.47) и (6.4.48) $L \rightarrow 0$ и $f \rightarrow \infty$. Используя выражение (7.3.4), действительно получим полуширину мод квадратичной среды (7.3.21). При сравнении необходимо помнить, что показатель преломления квадратичной среды на оси есть n_0 . Это означает, что в формулах (6.4.47) и (6.4.48) нужно использовать $\lambda = 2\pi/n_0 k_0$. Сравнение формул (7.3.20) и (6.3.20) показывает, что моды квадратичной среды действительно тождественны модам линзового волновода с $L \rightarrow 0$. Для полноты сравнения покажем, что фазовый множитель в (6.3.20) переходит в (7.3.22) при отсутствии деления среды на линзы. Прежде всего по указанной выше причине необходимо в (6.3.20) k заменить на $n_0 k$. Положим также $m = p$ и $n = q$ и в пределе при $L \rightarrow 0$ ($z < L$) получим из (6.3.20) с учетом (7.3.21)

$$\beta_L = n_0 k_0 - \frac{1}{2} \sqrt{\frac{n_1}{n_0}} (2p+1 + 2q+1). \quad (7.3.23)$$

Согласно формуле (6.6.10), радиус фазового фронта при $d_1 = L = 0$ становится бесконечным. Две постоянные распространения (7.3.22) и (7.3.23), очевидно, совпадают только в первом приближении. Величину (7.3.23) можно получить из (7.3.22), если разложить в ряд квадратный корень и сохранить только первые два члена разложения. Из выражения (7.3.23) видно, что в первом приближении и при пренебрежении дисперсией среды моды квадратичной среды имеют замечательное свойство — их групповая скорость $(\partial\beta/\partial\omega)^{-1}$ не зависит от порядка моды.

Таким образом, мы получили интересный результат: моды квадратичной среды совпадают с модами линзового волновода с точностью до фазового множителя в пределе при уменьшении до нуля расстояний между линзами. Фазовые множители совпадают только в первом приближении. Этот результат обусловлен тем, что моды квадратичной среды (7.3.20) — (7.3.22) являются точными решениями волнового уравнения (7.3.1) с показателем преломления (7.3.3). Ранее же мы отмечали, что моды Эрмита — Гаусса (6.3.20) являются только приближенным решением волнового уравнения. Приближение вносится тогда, когда вторая производная от u по z в (6.3.3) не учитывается. Таким образом, решение не является строгим, если оно получается путем перехода к пределу при стремлении расстояний между линзами к нулю. Однако в большинстве практических случаев величина

$$Q = \frac{\sqrt{n_0 n_1}}{n_0^2 k_0} \quad (7.3.24)$$

обычно очень мала, так что величины (7.3.22) и (7.3.23) находятся в хорошем согласии, по крайней мере для малых модовых чисел. Но и настоящая теория оказывается несправедливой, когда величина Q (7.3.24) достигает единицы. Скалярное волновое уравнение можно использовать только в том случае, когда показатель преломления изменяется очень мало на расстоянии порядка длины волны. Это требование также приводит к $Q \ll 1$. Чтобы убедиться в этом, образуем в соответствии с формулой (1.3.28) и с учетом (7.3.3) выражение

$$R = \frac{n_0^2 - n^2}{2\pi n_0^2} = \frac{1}{2\pi} \frac{n_0 n_1 x^2}{n_0^2}. \quad (7.3.25)$$

Используем расстояние порядка одной длины волны ($x = \lambda/n_0$) для точки, в которой сравниваются n^2 и n_0^2 . Для упрощения положим $y = 0$. Таким образом, имеем

$$R = \frac{n_0 n_1}{2\pi_0 n_0^4} \lambda^2 = 2\pi \frac{n_0 n_1}{n_0^4 k_0^2}. \quad (7.3.26)$$

Сравнение с выражением (7.3.24) приводит к

$$R = 2\pi Q^2. \quad (7.3.27)$$

Этот оценочный расчет показывает, что Q^2 является величиной того же порядка, что и R . Так как R должно быть много меньше единицы (чтобы можно было применять волновое уравнение), требуется, чтобы $Q \ll 1$. Согласие между (7.3.22) и (7.3.23) является, следовательно, хорошим для мод низшего порядка, т. е. пока применим скалярный волновой подход. Однако любопытен тот факт, что модовое решение Эрмита — Гаусса (7.3.20) с фазовой постоянной (7.3.22) является точным решением скалярного волнового уравнения. Приближенный же характер рассматриваемой теории связан с вопросом применимости скалярного волнового уравнения.

7.4. ВНЕОСЕВЫЕ ПУЧКИ В КВАДРАТИЧНОЙ СРЕДЕ

В разд. 3.6 в общем виде было показано, что центр интенсивности светового поля движется в соответствии с законами лучевой (геометрической) оптики в квадратичной среде. Доказательство было проведено в параксиальном приближении. В настоящем разделе прямым вычислением покажем, что гауссов пучок может распространяться в квадратичной среде и вне оси. Максимум распределения поля движется подобно световому лучу. Ширина распределения поля меняется периодически с периодом, равным половине длины периода колебаний луча.

Предположим, что при $z = 0$ в квадратичной среде имеет место следующее распределение поля:

$$F(x, y, 0) = A e^{-[(x-\xi)/W]^2 - (y/w)^2} \quad (7.4.1)$$

Зависимый от y множитель соответствует моде низшего порядка. Множитель, зависящий от x , показывает, что распределение поля сдвинуто на величину ξ вдоль оси

x . [Здесь ξ имеет иное значение, чем в (7.3.8).] Дополнительно введена произвольная полуширина W пучка в направлении оси x . Первоначальное распределение поля при $z = 0$ может быть разложено в бесконечный ряд с помощью мод (7.3.20) квадратичной среды:

$$F(x, y, z) = \sum_{p=0}^{\infty} C_p \psi_{p0}(x, y, z). \quad (7.4.2)$$

Разложение (7.4.2) дает поле во всех точках вдоль оси z , если только определены коэффициенты разложения при $z = 0$. Используя ортогональность мод, получим

$$C_p = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy F(x, y, 0) \psi_{p0}^*(x, y, 0). \quad (7.4.3)$$

В случае когда n_0 и n_1 в (7.3.3) являются комплексными величинами, ψ^* означает не комплексно-сопряженное поле, а только изменение знака в экспоненте $\exp[i(\omega t - \beta z)]$. [Сравни с последующим рассмотрением формулы (10.2.13).] Определение C_p основано на ортогональности функций Эрмита — Гаусса, которые являются чисто математическими выражениями, справедливыми даже при комплексных параметрах, если только $\exp(-x^2/w^2) \rightarrow 0$ при $x \rightarrow \infty$. Интегралы могут быть взяты с помощью таблиц интегралов Градштейна и Рыжика [61]:

$$C_p = \frac{\sqrt{\pi} A}{2^{p/2} \sqrt{p!}} \frac{wW}{\sqrt{w^2 + W^2}} \left(\frac{w^2 - W^2}{w^2 + W^2} \right)^{p/2} \times \\ \times H_p \left(\frac{\sqrt{2} w \xi}{\sqrt{w^4 - W^4}} \right) e^{-\xi^2/(w^2 + W^2)}. \quad (7.4.4)$$

Распределение поля в некоторой произвольной точке квадратичной среды может быть теперь выражено уравнением

$$F(x, y, z) = \frac{\sqrt{2} W A}{\sqrt{w^2 + W^2}} e^{-\xi^2/(w^2 + W^2)} \times \\ \times e^{-(x^2 + y^2)/w^2} \cdot e^{-i(n_0 k_0 - \sqrt{n_1/n_0})z} \times \\ \times \sum_{p=0}^{\infty} \left\{ \frac{1}{p!} \left[\frac{1}{2} \sqrt{\frac{w^2 - W^2}{w^2 + W^2}} e^{i \sqrt{n_1/n_0} z} \right]^p H_p \left(\sqrt{2} \frac{x}{w} \right) \times \right. \\ \left. \times H_p \left(\frac{\sqrt{2} w \xi}{\sqrt{w^4 - W^4}} \right) \right\}. \quad (7.4.5)$$

Постоянная распространения β_{p0} здесь взята в виде (7.3.23). Бесконечный ряд можно оценить с помощью одной из производящих функций полиномов Эрмита [73]. В соответствии с этим получим

$$F(x, y, z) = \frac{WA}{\sqrt{W^2 \cos \gamma z - i w^2 \sin \gamma z}} e^{-\xi^2/(w^2 + W^2)} \times \\ \times e^{-(x^2 + y^2)/w^2} e^{-i(n_0 k_0 - \gamma/2)z} \times \\ \times \exp \left\{ \frac{2x\xi - \left[\left(1 - \frac{W^2}{w^2} \right) x^2 + \frac{w^2}{w^2 + W^2} \xi^2 \right] e^{i\gamma z}}{W^2 \cos \gamma z - i w^2 \sin \gamma z} \right\}, \quad (7.4.6)$$

где

$$\gamma = \sqrt{\frac{n_1}{i n_0}}. \quad (7.4.7)$$

Выражение (7.4.6) является приближенным решением приведенного волнового уравнения. Оно удовлетворяет уравнению (7.3.1) (в параксиальном приближении), даже если n_0 и n_1 комплексные величины. Упростим дальнейшее рассмотрение, исследуя это уравнение в двух частных случаях. Сначала рассмотрим случай, когда полуширина смещенного начального распределения поля равна полуширине моды структуры

$$W = w. \quad (7.4.8)$$

При таком условии выражение (7.4.6) можно представить как

$$F(x, y, z) = A \exp \left[-\frac{(x - \xi \cos \gamma z)^2 + y^2}{w^2} \right] e^{i\phi(z)} \times \\ \times e^{-i(n_0 k_0 - \gamma)z}, \quad (7.4.9)$$

где

$$\phi(z) = \frac{1}{w^2} \left(2x\xi \sin \gamma z - \frac{1}{2} \xi^2 \sin^2 \gamma z \right). \quad (7.4.10)$$

Выражение (7.4.9) показывает, что распределение поля вне оси имеет форму моды структуры и перемещается в квадратичной среде без существенных искажений. Центр распределения поля следует по траектории луча в соответствии с выражением (7.2.8):

$$x = \xi \cos \gamma z. \quad (7.4.11)$$

Наклоненные фазовые фронты внеосевого поля описываются фазовой функцией (7.4.10).

Далее рассмотрим осевое поле, ширина которого не соответствует модам структуры. Полагая

$$\xi = 0, \quad (7.4.12)$$

получим из выражения (7.4.6)

$$F(x, y, z) = A \sqrt{\frac{W}{\bar{W}(z)}} e^{-y^2/w^2} e^{-[x/\bar{W}(z)]^2} e^{i\psi(z)} \times \\ \times e^{-i(n_0 k_0 - \gamma)z}. \quad (7.4.13)$$

Квадрат полуширины пучка в направлении x определяется как

$$\bar{W}^2(z) = \frac{1}{2W^2} [(W^4 + w^4) + (W^4 - w^4) \cos 2\gamma z], \quad (7.4.14)$$

а фазовый фронт — как

$$\psi(z) = \frac{1}{2} \left\{ \arctg \left[\left(\frac{w}{W} \right)^2 \operatorname{tg} \gamma z \right] - \gamma z + \right. \\ \left. + \frac{W^4 - w^4}{w^2 W^2 \bar{W}^2(z)} x^2 \sin 2\gamma z \right\}. \quad (7.4.15)$$

Выражение (7.4.14) показывает, что ширина распределения поля меняется периодически с периодом, равным половине длине периода колебаний луча.

Результат (7.4.14) может быть также получен из рассмотрения гауссовых пучков в линзовом волноводе. В разд. 6.4 была исследована ширина гауссова пучка, перемещающегося в линзовом волноводе. Чтобы можно было сравнить выражения (6.4.36) и (7.4.14), необходимо ввести пучок в линзовый волновод так, чтобы радиус кривизны фазового фронта был бесконечным, так как предполагалось, что луч в квадратичной среде при $z = 0$ имеет плоский фазовый фронт. Записав $w_1 = w_0 = W$, положив $L \rightarrow 0$ и $f \rightarrow \infty$ и используя формулы (7.2.4), (7.2.6) и (7.3.21), нетрудно (6.4.36) преобразовать к форме (7.4.14).

Наше решение (7.4.6) является приближенным. Моды квадратичной среды являются точными решениями волнового уравнения, но само волновое уравнение лишь приближенно описывает электромагнитное поле. Чтобы просум-

мировать ряды в (7.4.5), необходимо использовать приближение (7.3.23) для постоянной распространения. Это приближение эквивалентно параксиальному приближению лучевой оптики, поскольку решение (7.4.9) описывает такое распределение поля, при котором центр интенсивности движется подобно параксиальному лучу.

Было показано, что, согласно параксиальному приближению, гауссов пучок движется через квадратичную среду без искажения поля. Пучок имеет траекторию, колеблющуюся относительно оптической оси, но сам остается хорошо коллимированным. Его ширина может изменяться периодически, если она не совпадает с шириной мод структуры, однако никакого разрушения поля пучка не происходит. Маркатили [65] показал, что подобная стабильность гауссовых пучков имеет место только в среде с показателем преломления, определяемым формулой (7.3.3). Любые члены высшего порядка в разложении показателя преломления по степеням x и y вызывают распадение пучка, подобное описанному в разд. 5.8. Распадение пучка на несколько отчетливых максимумов может произойти лишь в волноводе значительной длины. И это обстоятельство не зависит от степени аберрации среды. Такое распадение пучка должно происходить в диэлектрической среде, диэлектрическая проницаемость которой не соответствует строго параболическому закону. Сходство между непрерывной диэлектрической средой и линзовыми волноводами не ограничивается только квадратичной средой. Тот факт, что гауссовы пучки не сохраняют свою форму в среде, отличающейся от квадратичной, означает на практике, что их невозможно передавать без искажений на сколь угодно большие расстояния. Ни одна среда не может быть абсолютно квадратичной. Наше обсуждение распространения световых пучков в идеальных линзовых волноводах и в идеальной квадратичной среде является, следовательно, лишь идеализацией, которая, конечно, не встречается в реальных волноводах. Однако знание идеального случая весьма полезно. Хорошо передающая среда является в некотором смысле достаточным приближением к идеальной, так что можно ожидать, что поведение гауссовых пучков будет следовать нашим предсказаниям по крайней мере на некотором конечном

участке пути. На очень больших расстояниях поведение пучка существенно отличается от идеального случая. Что происходит в реальной ситуации, было показано в разд. 5.8.

7.5. КВАДРАТИЧНАЯ СРЕДА С ПОТЕРЯМИ ИЛИ УСИЛЕНИЕМ

В предыдущих разделах были рассмотрены колебания в квадратичной среде, которая не вносила ни затухания, ни усиления. Такое предположение, естественно, является идеализацией. Ни в одной волноведущей среде не существует точной квадратичной зависимости ее диэлектрической проницаемости от радиуса; не существует и диэлектрического материала, в котором полностью отсутствовало бы затухание волн. Большинство диэлектриков имеет относительно высокие потери в оптическом диапазоне. Потери в 100—1000 дБ/км не так уже необычны. В разд. 2.6 было показано [формулы (2.6.18) и (2.6.10)], что диэлектрическую среду с потерями можно описать комплексной диэлектрической проницаемостью.

Важно также рассмотреть диэлектрические среды не только с затуханием, но и с усилением. Активный материал в лазерах может быть описан как диэлектрическая среда, усиливающая, а не поглощающая. Усиление также может быть описано с помощью комплексной диэлектрической проницаемости. Проявляет ли среда усиливающие или поглощающие свойства, определяется знаком мнимой части показателя преломления.

В настоящем разделе будут исследованы моды в квадратичной среде с комплексной диэлектрической проницаемостью. При рассмотрении среды с квадратичной зависимостью действительной части диэлектрической проницаемости от радиуса было установлено, что направленность волн осуществляется благодаря тому, что действительная часть показателя преломления максимальна на оптической оси и уменьшается по мере удаления от нее. Такой механизм направленности настолько хорошо известен, что кажется неожиданной возможность получения направляемых мод в среде с постоянным значением действительной части показателя преломления. Однако, даже

если действительная часть показателя преломления имеет наименьшее значение на оси и увеличивается при удалении от оси, можно получить стабильную направляемую моду. Для этого необходимо, чтобы потери в среде имели наименьшее значение вблизи оптической оси и увеличивались по мере удаления к периферии. В случае среды с усилением стабильные моды получаются, когда максимум усиления имеет место на оси. Зависимость действительной части показателя преломления от радиуса при этом является в некоторой степени произвольной. Таким образом, нет особой необходимости обеспечивать уменьшение действительной части показателя преломления с увеличением расстояния от оптической оси. Оказывается, что для существования стабильных мод мнимая часть показателя преломления важнее, чем ее действительная часть [74].

Выше мы неоднократно упоминали о существовании устойчивых направляемых мод, но не пояснили, что понимается под этим термином. Назовем моду устойчивой, или стабильной, если она продолжает распространяться вдоль оси структуры даже в том случае, если ее слегка сместить с оптической оси. Нестабильная же мода может существовать лишь как математическое решение задачи, но она не распространяется даже при незначительном смещении от оптической оси. Нестабильные моды имеют место, когда потери в среде уменьшаются с увеличением расстояния от оптической оси или в случае, когда среда является активной, но ее усиление на оси наименьшее. Действительная часть показателя преломления, таким образом, не имеет большого значения в вопросе устойчивости моды. Важно отметить в этой связи, что вопрос стабильности моды до некоторой степени является академическим. Представим себе, например, диэлектрическую среду, действительная часть показателя преломления которой ведет себя достаточно хорошо, чтобы обеспечить направленность моды. Действительная часть показателя преломления такой среды и потери в ней уменьшаются с увеличением расстояния от оси. По нашему критерию о модовой стабильности такая мода не является стабильной. Однако очевидно, что эта нестабильность не имеет решающего значения, если величина потерь

поперек среды изменяется очень медленно. В принципе моды такой структуры должны быть нестабильными, но если они распространяются на достаточно большое расстояние вдоль волновода и при этом разрушение их структуры оказывается незначительным, то эту нестабильность нельзя наблюдать. Вопрос о стабильности мод является, таким образом, скорее принципиальным, чем имеющим практическое значение. Однако в среде с усилением нестабильность моды приводит к более серьезным последствиям. Поскольку поле растет, то нестабильность моды становится заметной, если мода распространяется на большое расстояние вдоль структуры.

Мода в среде с неправильной действительной частью показателя преломления (термин «неправильная» здесь означает, что показатель преломления увеличивается при удалении от оси) ведет себя совершенно отлично от моды, которая образуется при нормальном распределении показателя преломления. Направленность моды в нормальной среде типа описанной в предыдущем разделе осуществляется за счет полного внутреннего отражения. В терминах лучевой оптики это означает отклонение лучей по направлению к оси структуры. Никакого излучения энергии от оси здесь не происходит. В неправильной среде направленность моды возможна из-за мнимой части показателя преломления. Эта направленность совсем иного типа. Представим себе среду, которая обеспечивает некоторое усиление вблизи оси. Пусть усиление уменьшается с увеличением расстояния от оси, а потери, наоборот, возрастают при удалении от оси структуры. Структура такого типа обладает модами в математическом смысле. Задача о модах имеет устойчивые математические решения. Однако поле такой моды не остается вблизи оси направляющей среды и имеет место излучение энергии в радиальном направлении от оси. Возможность существования устойчивой моды обусловлена не полным внутренним отражением энергии поля обратно к оси волновода, а скорее возникновением энергии поля в активной среде, достаточной для поддержания модового распределения, которое имеет более высокую плотность энергии поля на оси. При удалении от оси плотность энергии поля уменьшается. Данное рассуждение справедливо, пока среда

в целом является более поглощающей, чем усиливающей. Если потери достаточно быстро увеличиваются при удалении от оси, то существуют устойчивые решения для мод. Форма поля в этом случае сохраняется потому, что энергия рассеивается при удалении от оси больше, чем непосредственно на оси, в результате чего максимум поля находится на оси. Некоторое отражение энергии по направлению к оси при этом имеет место даже в среде с постоянной действительной частью показателя преломления. Изменение мнимой части показателя также приводит к отражению энергии в диэлектрике.

После этих вводных замечаний вернемся к задаче о модах в диэлектрической среде с комплексным показателем преломления. К счастью, нет необходимости вновь решать волновое уравнение. Решение, полученное в разд. 7.3, справедливо для любых возможных значений n_0 и n_1 . Оно справедливо и тогда, когда эти постоянные принимают комплексные значения. Необходимо только выяснить пригодность комплексных параметров для наших решений. Комплексные значения n_0 и n_1 делают ξ и η , согласно формулам (7.3.8) и (7.3.9), также комплексными. Полиномы Эрмита (7.3.16) и (7.3.17) являются аналитическими функциями, которые остаются полиномами (5.6.14) даже при комплексных значениях своего аргумента.

Допустим, что n_0 имеет комплексное значение

$$n_0 = n_{0r} - in_{0i}. \quad (7.5.1)$$

Положительные значения n_{0i} указывают, что среда обладает потерями, тогда как отрицательные значения n_{0i} означают, что среда имеет усиление. Для всех реальных физических сред можно считать, что

$$|n_{0i}| \ll n_{0r}, \quad (7.5.2)$$

поэтому можно приближенно записать

$$n_0^2 = n_{r0}^2 - 2in_{0r}n_{0i}. \quad (7.5.3)$$

Аналогичным образом введем комплексную величину

$$\sqrt{n_0 n_1} = a + ib. \quad (7.5.4)$$

Величина (7.3.21) является теперь комплексной. Ее смысл как величины квадрата полуширины пучка сохраняется для реальной части. Имея в виду, что w является комплексной величиной, можно записать решение для моды (7.3.20) в квадратичной среде с комплексной диэлектрической проницаемостью в виде

$$\psi(x, y, z) = \frac{\sqrt{2}}{\sqrt{\pi 2^{p+q} p! q! w}} H_p \left(\sqrt{2} \frac{x}{w} \right) \times \\ \times H_q \left(\sqrt{2} \frac{y}{w} \right) e^{-(k_0/2)ar^2} e^{-i(k_0/2)br^2} e^{-i\beta_p q z}. \quad (7.5.5)$$

Потребуем, чтобы поле направляемой моды спадало с ростом величины

$$r = (x^2 + y^2)^{1/2}. \quad (7.5.6)$$

Это требование показывает, что квадратичная среда должна удовлетворять условию

$$a > 0. \quad (7.5.7)$$

Если это условие удовлетворяется, то получим направляемую моду независимо от значения диэлектрической проницаемости.

Из модового решения (7.5.5) видно, что фазовые фронты в среде с комплексным показателем преломления не являются плоскими. Их форма более сложная, поскольку они определяются не только экспоненциальными функциями, но и комплексными значениями полиномов Эрмита. Для моды низшего порядка ($p = 0$ и $q = 0$) фазовый фронт при положительных значениях b является вогнутым, если смотреть в положительном направлении оси z . Это означает, что энергия на оси при увеличении b уменьшается. Для того чтобы увидеть связь этой величины с показателем преломления, рассмотрим выражение (7.1.1). Используя формулы (7.5.1) и (7.5.4), запишем вещественную часть выражения (7.1.1)

$$\operatorname{Re} n = n_{0r} - \frac{1}{2n_{0r}} (a^2 - b^2) r^2 \quad (7.5.8)$$

и мнимую часть

$$\operatorname{Im} n = -n_{0i} - \frac{1}{n_{0r}} a b r^2. \quad (7.5.9)$$

Произведениями n_{0i} на a^2 , b^2 и ab мы пренебрегли, предположив, что они значительно меньше, чем члены в (7.5.8) и (7.5.9).

Поскольку a — величина положительная, то положительное значение b означает, что при увеличении r в среде увеличиваются потери. Предположим, что величина n_{0i} отрицательная. Тогда среда усиливает на оси при $r = 0$. Для положительного значения b усиление уменьшается при удалении от оси. Энергия, возникающая вблизи оси, уходит от нее и поглощается.

Поведение мнимой части показателя преломления (увеличивается или уменьшается она с ростом величины r) полностью зависит от знака b , так как величина a должна быть положительной, чтобы имела место направленность моды. Поведение действительной части показателя преломления зависит от того, больше a^2 чем b^2 или меньше. Среда обладает нормальными волноведущими свойствами ($\text{Re } n$ уменьшается с ростом r), если $a^2 > b^2$, и аномальными свойствами ($\text{Re } n$ возрастает с увеличением r), если $a^2 < b^2$. Направляемая мода существует во всех случаях как при нормальном, так и при аномальном поведении действительной части показателя преломления, а также в случае среды, для которой мнимая часть увеличивается или уменьшается с ростом r . Этот неожиданный результат, однако, не дает однозначного ответа о стабильности моды. Он говорит лишь о том, что математические модовые решения существуют во всех четырех случаях.

Для того чтобы решить вопрос о стабильности моды, исследуем поведение моды низшего порядка ($p = q = 0$) при ее смещении с оптической оси. Соответствующий расчет уже был проведен в разд. 7.4. Необходимо только рассмотреть вопрос о справедливости уравнения (7.4.9) при комплексных значениях n_0 и n_1 .

Из формул (7.3.21) и (7.5.4) имеем

$$\frac{1}{w^2} = \frac{k_0}{2} (a + ib), \quad (7.5.10)$$

а из (7.4.7) и (7.5.4) приближенно получаем

$$\gamma = \frac{a + ib}{n_{0r}}. \quad (7.5.11)$$

Вновь мы пренебрегли произведениями малой величины n_{0i} на малые величины a и b . Синус и косинус принимают следующие комплексные значения:

$$\cos \gamma z = \cos \frac{a}{n_{0r}} z \operatorname{ch} \frac{b}{n_{0r}} z - i \sin \frac{a}{n_{0r}} z \operatorname{sh} \frac{b}{n_{0r}} z \quad (7.5.12)$$

и

$$\sin \gamma z = \sin \frac{a}{n_{0r}} z \operatorname{ch} \frac{b}{n_{0r}} z + i \cos \frac{a}{n_{0r}} z \operatorname{sh} \frac{b}{n_{0r}} z. \quad (7.5.13)$$

Для абсолютной величины функции $F(x, y, z)$, определяемой формулой (7.4.9), получаем

$$\begin{aligned} |F(x, y, z)| = & A e^{-(ak_0/2)(x^2+y^2)} e^{-[n_{0i}k_0+(b/n_{0r})]z} \times \\ & \times \exp [x \xi k_0 e^{-(b/n_{0r})z} \left(a \cos \frac{a}{n_{0r}} z - b \sin \frac{a}{n_{0r}} z \right)] \times \\ & \times \exp \left\{ -\frac{k_0}{4} \xi^2 \left[a + e^{-2(b/n_{0r})z} \times \right. \right. \\ & \left. \left. \times \left(a \cos 2 \frac{a}{n_{0r}} z - b \sin 2 \frac{a}{n_{0r}} z \right) \right] \right\}. \quad (7.5.14) \end{aligned}$$

Устойчивость или неустойчивость смещенной моды определяется множителем $\exp(-bz/n_{0r})$. При положительных b экспоненциальный множитель уменьшается с ростом z , так что распределение поля принимает следующее устойчивое значение:

$$\begin{aligned} |F(x, y, z)| = & A e^{-(ak_0/4)\xi^2} e^{-(ak_0/2)(x^2+y^2)} \times \\ & \times e^{-[n_{0i}k_0+(b/n_{0r})]z}. \quad (7.5.15) \end{aligned}$$

Из этой формулы видно, что для положительных b внеосевое поле движется вдоль оси и уменьшается или растет в зависимости от того, является ли величина $[n_{0i}k_0 + b/n_{0r}]$ положительной или отрицательной. Положительные значения b и, следовательно, устойчивое распространение мод получаются в среде, усиление которой уменьшается с ростом r или потери в которой увеличиваются с ростом r .

В противном случае величина b отрицательная и множитель $\exp(-bz/n_{0r})$ увеличивается с ростом z . Распределение поля в этом нестабильном случае весьма сложное. Очевидно, однако, что поле простирается все дальше и дальше от оси. Это видно из третьей экспоненты в выра-

жении (7.5.14), которая стремится противодействовать первой экспоненте с ее зависимостью от $-r^2$, заставляющей моду оставаться сконцентрированной вблизи оси структуры. Хотя направляемая мода и возможна при $\xi \rightarrow 0$, но достаточно бесконечно малого смещения, чтобы появилась возможность ее полного расплывания. Как упоминалось во вводной части этого раздела, подобное расплывание слабо зависит от констант, в частности от b . Почти стабильная мода все же полезна для практических целей. В принципе, однако, моды с отрицательными b являются неустойчивыми.

Такое поведение стабильных и нестабильных мод не ограничивается квадратичной средой с комплексным показателем преломления. Можно показать, что в принципе такое же поведение имеет место в слоистой среде, в частности в рассматриваемой в следующей главе. Автором был проведен численный анализ уравнения с комплексными собственными значениями для слоистого диэлектрического волновода. Было показано, что формальное модовое решение для него также имеет место во всех четырех случаях распределения показателя преломления. Существование стабильных и нестабильных направляемых мод в диэлектрических волноводах с комплексным показателем преломления является характерным, таким образом, не только для квадратичной среды. Это было продемонстрировано для слоистых диэлектрических волноводов.

Возвращаясь к модам квадратичной среды, рассмотрим процесс затухания или усиления мод для случая комплексного показателя преломления. Постоянную распространения (7.3.23) можно разложить на действительную и мнимую части:

$$\operatorname{Re} \beta_{pq} = n_{0r} k_0 - \frac{a}{n_{0r}} (p + q + 1) \quad (7.5.16)$$

и

$$\operatorname{Im} \beta_{pq} = -n_{0i} k_0 - \frac{b}{n_{0r}} (p + q + 1). \quad (7.5.17)$$

Произведениями n_{0i} на a и b снова пренебрегаем. Затухание или усиление определяется мнимой частью β_{pq} . Моды

усиливаются, если

$$\operatorname{Im} \beta_{pq} > 0, \quad (7.5.18)$$

и испытывают поглощение, если

$$\operatorname{Im} \beta_{pq} < 0. \quad (7.5.19)$$

Вопрос о том, что именно имеет место для данной моды — поглощение или усиление, зависит от трех факторов: знака n_{0i} , знака b и величины модового числа $p + q + 1$.

Рассмотрим устойчивые моды ($b > 0$) при $n_{0i} > 0$. В этом случае среда обладает потерями на оси, причем с увеличением r потери растут. Величина $\operatorname{Im} \beta_{pq}$ отрицательная; с ростом модовых чисел она растет по абсолютной величине. Это означает, что устойчивые моды в квадратичной среде с потерями испытывают тем большее затухание, чем выше порядок моды. Такое поведение является типичным для нормальной направляющей среды. Оно обуславливает стабильность мод. Если произвольное смещенное относительно оси распределение поля разложить в ряд по нормальным модам, то наиболее быстро при увеличении z затухают высшие члены разложения и на достаточно большом расстоянии вдоль волновода остается лишь мода низшего порядка. Любое произвольное распределение поля переходит в моду низшего порядка, что и проявляется при рассмотрении разложенного в ряд поля.

Рассмотрим далее случай $b > 0$ при $n_{0i} < 0$. Среда такого рода приводит к усилению на оси, но оно уменьшается с ростом r и при достаточно большом r усиление переходит в поглощение. Усиление или затухание мод в этом случае зависит от соотношения между $n_{0i}k_0$ и b ($p + q + 1$). Усиление имеет место при

$$|n_{0i}k_0| > \frac{b}{n_{0r}}(p + q + 1), \quad (7.5.20)$$

а затухание — при

$$|n_{0i}k_0| < \frac{b}{n_{0r}}(p + q + 1). \quad (7.5.21)$$

Может оказаться, что усиление на оси будет недостаточным для роста каких-либо мод. Мода низшего порядка возрастает при достаточно большом осевом усилении.

Однако с ростом порядка моды усиление уменьшается и переходит в затухание при достаточно больших значениях $p + q + 1$. Снова оказывается, что любое произвольное распределение поля в конце концов приводит к моде низшего порядка, что проявляется при рассмотрении разложения поля в ряд.

Рассмотрим теперь нестабильный случай: $b < 0$. Сначала допустим, что среда имеет поглощение на оси ($n_{0i} > 0$). При достаточно малом значении b моды низшего порядка испытывают затухание, которое уменьшается с возрастанием порядка моды. При достаточно больших значениях порядка моды испытывают усиление. Если $|b|$ настолько велико, что все моды усиливаются, то вновь получаем, что степень усиления увеличивается с ростом порядка моды. Такая аномальная ситуация объясняет, почему распределение поля при $b < 0$ является неустойчивым. Действительно, рассмотрим разложение в ряд произвольного входного поля. По мере продвижения поля вдоль волновода моды высокого порядка в разложении растут быстрее и становятся доминирующими. Сходимость ряда является неустойчивой и поэтому не удивительно, что имеет место полное разрушение поля входного пучка, который первоначально был смещен лишь на малую величину от оси структуры.

Следует заметить, что квадратичная среда с неограниченно возрастающим усилением по мере роста r не имеет физического смысла. В любой физически реальной среде возможен рост усиления лишь на начальном участке возрастания r . Поэтому усиление мод высокого порядка не может неограниченно увеличиваться с ростом модового числа. В конечном счете у мод высокого порядка усиление должно постепенно уменьшаться и даже переходить в затухание. Разложения произвольных полей в ряд не расходятся в реальных случаях. Однако, если даже разложения поля сходятся в физически реальной среде, все же может происходить значительное искажение поля при увеличении усиления с ростом модовых чисел по крайней мере для мод низкого порядка. Таким образом, наше рассмотрение стабильных и нестабильных направляющих сред дает некоторое представление о том, что можно ожидать на практике.

7.6. ЛИНЗОВЫЕ СВОЙСТВА КВАДРАТИЧНОЙ СРЕДЫ

Закончим главу о квадратичной среде рассмотрением ее линзовых свойств [76]. Для упрощения допустим, что луч света входит в квадратичную среду и выходит из нее с небольшим дополнительным отклонением по углу. Это допущение оправдано следующей аргументацией. Показатель преломления в реальной квадратичной среде изменяется весьма незначительно. Если поместить участок квадратичной среды в однородную среду, показатель преломления которой согласован с показателем квадратичной среды на оси, то можно ожидать достаточно хорошего согласования показателей преломления во всех точках поверхности раздела между этими двумя средами. Таким образом, преломление при прохождении границы раздела однородной и квадратичной сред является очень слабым и им можно пренебречь. Если же не согласовать показатели, то будет иметь место преломление лучей, проходящих через границу раздела сред. Подобные эффекты аналогичны преломлению на границе раздела двух однородных сред. Известно, что входящие из одной однородной среды в другую лучи из точечного источника распространяются во второй среде так, как будто источник излучения находится ближе или дальше своего действительного местоположения. Для простоты изложения пренебрежем этими тривиальными эффектами преломления.

Легко убедиться, что квадратичная среда должна обладать свойством формирования изображения. Рассмотрим разложение (7.4.2) произвольного поля в точке $z = 0$ в ряд по модам (7.3.20). По мере проникновения поля в среду его форма меняется, поскольку меняются относительные фазы различных членов разложения. Фаза членов ряда дается выражением (7.3.23)

$$\beta_{pq}z = n_0 k_0 z - (p + q + 1) \gamma z, \quad (7.6.1)$$

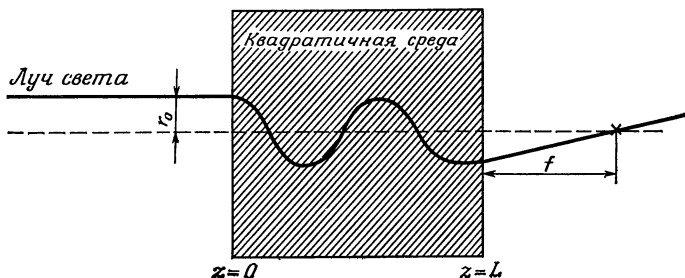
где γ определяется формулой (7.4.7). Для произвольных значений z фазы мод различны, так что суперпозиция отдельных мод разложения всегда приводит к изменению формы поля. Однако вдоль оси структуры имеются точки, где выполняется соотношение

$$\gamma z = 2N\pi, \quad N = 1, 2, 3, \dots \quad (7.6.2)$$

В этих точках все фазы кратны 2π , за исключением общего фазового множителя $n_0 k_0 z$. В точках, определяемых соотношением (7.6.2), члены ряда складываются точно таким же образом, как в начальной плоскости $z = 0$. Начальное распределение поля, таким образом, периодически восстанавливается в плоскостях

$$z = \frac{2N\pi}{\gamma}. \quad (7.6.3)$$

Во всех этих плоскостях имеет место точное воспроизведение первоначального распределения поля. Таким образом,



Ф и г. 7.6.1. Квадратичная среда действует как линза. Все горизонтальные входные лучи проходят через фокус.

квадратичная среда способна формировать изображение объектов, размещенных на ее входе. Во всех рассмотренных случаях двумерное изображение не искажается и не увеличивается. Из формулы (7.4.11) можно видеть, что плоскости изображения появляются через интервалы, кратные периоду колебаний луча.

Участок квадратичной среды конечной длины действует, таким образом, как линза. Чтобы убедиться в этом, рассмотрим фиг. 7.6.1. Луч света, входящий в квадратичную среду при $z = 0$ с тангенсом угла наклона $r'(0) = 0$ и на расстоянии r_0 от оси, выходит из среды при $z = L$ [см. формулу (7.2.8) или (7.4.11)] на расстоянии

$$r = r_0 \cos \gamma L. \quad (7.6.4)$$

Наклон луча в точке $z = L$ равен

$$r' = -\gamma r_0 \sin \gamma L. \quad (7.6.5)$$

Пренебрегая добавочным преломлением на граничной поверхности, мы видим, что луч пересекает оптическую ось на расстоянии f от границы квадратичной среды,

$$f = -\frac{r}{r'} = \frac{\operatorname{ctg} \gamma L}{\gamma}. \quad (7.6.6)$$

Примечательно, что расстояние f не зависит от расстояния r_0 , на котором луч входит в квадратичную среду. Все параллельные входные лучи пересекают оптическую ось в одной точке. Таким образом, квадратичная среда ведет себя как линза. Фокусное расстояние f может быть положительным или отрицательным. Положительное значение f указывает, что среда действует как положительная линза, способная формировать действительные изображения. Отрицательные значения f соответствуют тому, что среда ведет себя как отрицательная или рассеивающая линза, способная давать мнимые изображения. Действие среды как положительной или отрицательной линзы зависит от ее длины и фокусирующей способности, описываемой величиной γ согласно (7.4.7). Положительная линза, показанная на фиг. 7.6.1, может быть воспроизведена с помощью двух обычных тонких линз.

Существуют промышленные образцы тонких оптических волокон, радиальный градиент показателя преломления которых соответствует квадратичной среде. Эти волокна ведут себя как линзы [23, 112], т. е. можно увидеть изображения реальных объектов, размещенных перед волокном. Такие волокна весьма удобны для наблюдения недоступных объектов, например, таких, которые расположены за препятствием. Они могут также найти важные применения в медицине для наблюдения органов внутри организма. Они, конечно, могут быть использованы для передачи света на некоторое расстояние. Оптические волокна с квадратичной средой имеют слишком большое затухание, чтобы их можно было применять в световых волноводах для дальней оптической связи. Однако есть надежда, что можно уменьшить затухание до такого значения, когда станет возможной дальняя световая передача через такую квадратичную среду. Возможные колебания пучка в результате случайных отклонений оптической оси волокон от прямой требуют большого внимания при укладке таких оптических волноводов.

ОПТИЧЕСКИЕ ВОЛОКНА И ДИЭЛЕКТРИЧЕСКИЕ ВОЛНОВОДЫ

8.1. ВВЕДЕНИЕ

Термин «оптическое волокно» характеризует один из типов диэлектрических волноводов, направляющих световые волны [79]. Эти волноводы называются волокнами из-за их нитевидности. Обычно различают два типа оптических волокон. В предыдущей главе упоминалось оптическое волокно, показатель преломления материала которого изменяется по квадратичному закону в радиальном направлении. Оптическое волокно более простого типа действует также на основе принципа радиального изменения показателя преломления. Изменение показателя преломления по сечению не является у него плавным, а существуют резко выраженные области сечения с различными, но постоянными показателями преломления. Настоящая глава посвящена обсуждению свойств оптических волокон именно такого типа. В этой и последующих главах, где используется термин «оптическое волокно», имеется в виду волокно, у которого показатель преломления в поперечном сечении изменяется так, как показано на фиг. 8.1.1.

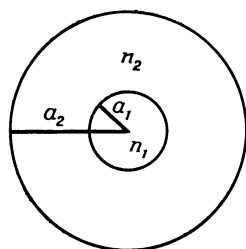
Оптический волновод, изображенный на фиг. 8.1.1, называется волокном в оболочке. Диэлектрический цилиндр из материала с показателем преломления n_1 окружен concentрическим диэлектрическим цилиндром из материала с показателем преломления n_2 . Эти показатели преломления удовлетворяют соотношению

$$n_1 > n_2. \quad (8.1.1)$$

Внешняя диэлектрическая область с показателем преломления n_2 не играет существенной роли при распространении

нии волны в диэлектрическом волноводе такого типа. Оптическое волокно поддерживает волны, даже когда $n_2 = 1$. Существуют две причины, по которым предпочтительнее использовать оптическое волокно в оболочке, чем без нее. Поле диэлектрического волновода не полностью находится внутри диэлектрической области с показателем преломления n_1 , частично оно выходит во внешнюю область, где спадает экспоненциально. Поскольку волокно должно поддерживаться в пространстве, то лучше

Ф и г. 8.1.1. Поперечное сечение круглого оптического волокна в оболочке.



окружить сердцевину внешней оболочкой, чтобы избежать рассеяния и искажения поля поддерживающими опорами, возмущающими поле волновода. Вследствие того, что поле спадает экспоненциально внутри области 2 с показателем преломления n_2 , вне оболочки поле практически отсутствует.

Вторая причина использования оболочки связана с механизмом распространения поля в волокне. На данной частоте оптическое волокно способно направлять конечное число типов волн (мод) [80, 81]. Если диаметр сердцевины a_1 много больше длины волны λ направляемого излучения, то возможно существование большого числа направляемых мод. Однако для целей передачи света часто желательно ограничить число возможных мод настолько, насколько это возможно. Путем правильного выбора размеров волновода можно добиться одномодового режима его работы. Размеры сердцевины при одномодовом режиме работы существенно зависят от отношения n_1/n_2 . Для распространения только одной моды необходимо, чтобы отношение n_1/n_2 было как можно ближе к единице, а размер a_1 был как можно меньше. Практически трудно изготовить оптическое волокно с малым диаметром сердце-

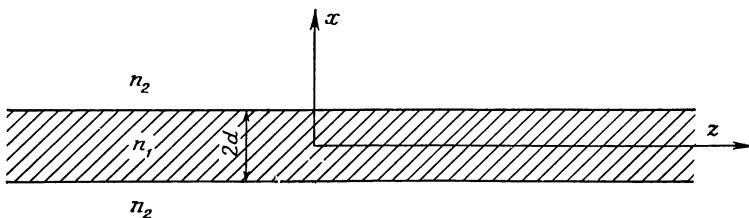
вины, например 0,5 мкм. Однако внутренняя сердцевина может быть довольно большой и тем не менее будет возможен одномодовый режим работы, если отношение n_1/n_2 окажется достаточно близким к единице. Можно допустить размер $2a_1$ в несколько микронов и при этом все же получить одномодовый волновод.

Оптические волокна имеют некоторые общие черты с полыми металлическими волноводами. И те и другие могут поддерживать ограниченное число направляемых мод на любой заданной частоте. В обеих структурах возможно преобразование мод, если волноводы отклоняются от идеальной прямолинейной геометрии [82]. Однако, если в металлических волноводах распространяются только направляемые моды и преобразование мод связано с обменом мощности между счетным числом этих мод, спектр волн диэлектрического волновода и оптических волокон наряду с конечным числом направляемых мод с дискретным спектром имеет континуум ненаправляемых мод излучения. Моды излучения также являются решениями уравнений Максвелла и удовлетворяют граничным условиям, налагаемым на поле при наличии диэлектрических поверхностей раздела. Однако если в металлических волноводах направляемые моды имеют только дискретный набор постоянных распространения, то в диэлектрическом волокне, кроме них, существует бесконечное множество волн излучения с постоянными распространения, образующими непрерывный спектр. При отклонении от идеальной геометрии происходит не только перераспределение мощности между направляемыми модами диэлектрического волновода, но и часть мощности рассеивается в непрерывный спектр мод излучения. Рассеяние мощности в непрерывный спектр представляет собой излучение поля во внешнюю среду.

Рассмотрим моды оптического диэлектрического волновода на основе уравнений Максвелла. Поляризационные эффекты оказываются важными для этих структур и в дальнейшем учитываются. Было бы неправильно изучать распространение мод в таком слоистом диэлектрическом волноводе, используя лишь скалярное волновое уравнение. Наряду с изучением свойств регулярных структур мы коснемся в гл. 9 вопросов преобразования мод и эф-

фактов излучения, вызванных нерегулярностью¹⁾ границ волновода.

Математическое рассмотрение потерь на излучение в оптических волокнах с цилиндрической геометрией оказывается довольно сложным. Существуют родственные структуры, геометрия и математическое рассмотрение которых намного проще, чем цилиндрических оптических волокон. Такой структурой является плоский



Ф и г. 8.1.2. Продольное сечение плоского волновода.

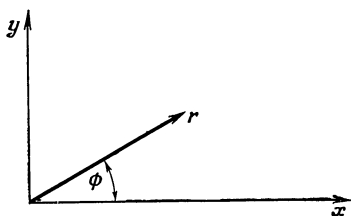
волновод, или слой. Схематически плоский волновод показан на фиг. 8.1.2. Предполагается, что волны распространяются в направлении оси z . Слой неограниченно простирается в направлении осей z и y (перпендикулярно к плоскости фигуры). Как в плоском, так и в круглом оптических волокнах могут распространяться направляемые моды, но, кроме них, существуют также моды излучения. Явления излучения поля и преобразования мод легче изучать на модели более простого плоского волновода. Результаты, полученные при рассмотрении плоского волновода, непосредственно применимы к круглым оптическим волокнам. Величины потерь на излучение, вызванные шероховатостью поверхности, почти одинаковы для плоского волновода и цилиндрического волокна. По этой причине рассмотрим распространение мод в плоском волноводе достаточно подробно, чтобы на его примере изучить свойства диэлектрического оптического волокна. Такой подход позволяет обойтись без сложных математических выкладок и продемонстрировать основные

¹⁾ Под нерегулярностью в теории волноводов понимается изменение параметров волновода вдоль его оси. — *Прим. ред.*

черты волоконной оптики, а также получить численные результаты, которые оказываются применимыми и к более сложным для анализа цилиндрическим волокнам.

8.2. НАПРАВЛЯЕМЫЕ ВОЛНЫ (МОДЫ) КРУГЛЫХ ОПТИЧЕСКИХ ВОЛОКОН

В этом разделе получим выражения для направляемых мод оптического волокна в оболочке. Оболочка выполняет свою роль полностью, если ее радиус настолько велик, что поле вблизи поверхности раздела между оболочкой и окружающим воздухом практически отсутствует. Таким образом, достаточно предположить, что радиус оболочки



Фиг. 8.2.1. Декартова и полярная системы координат.

неограничен: $a_2 \rightarrow \infty$. Разница между модами волокна с оболочкой и без нее будет незначительной для любого хорошо выполненного оптического волокна.

Уравнения поля для оптических систем цилиндрической геометрии выведены в разд. 1.4. Запишем эти уравнения в цилиндрических координатах r , ϕ , z . Для этого воспользуемся преобразованием координат. Координата z в направлении оптической оси системы является общей для декартовой и цилиндрической систем координат и нет необходимости преобразовывать ее из одной системы в другую. Соотношение между двумя системами координат в поперечной плоскости показано на фиг. 8.2.1. Из фигуры видно, что между координатами имеют место следующие соотношения:

$$x = r \cos \phi, \quad (8.2.1)$$

$$y = r \sin \phi. \quad (8.2.2)$$

Преобразование компонент F_x , F_y вектора \mathbf{F} из декартовой системы координат в компоненты F_r , F_ϕ цилиндриче-

ской системы координат показано на фиг. 8.2.2. Математически это преобразование имеет вид

$$F_r = F_x \cos \phi + F_y \sin \phi, \quad (8.2.3)$$

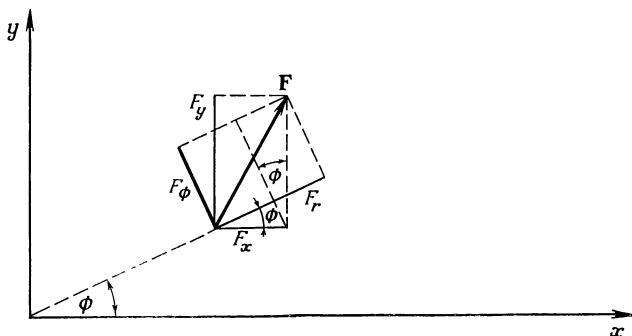
$$F_\phi = -F_x \sin \phi + F_y \cos \phi. \quad (8.2.4)$$

Производные компонент поля по r и ϕ получаются следующим образом:

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial r} = \frac{\partial f}{\partial x} \cos \phi + \frac{\partial f}{\partial y} \sin \phi, \quad (8.2.5)$$

$$\frac{\partial f}{\partial \phi} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \phi} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \phi} = -r \frac{\partial f}{\partial x} \sin \phi + r \frac{\partial f}{\partial y} \cos \phi. \quad (8.2.6)$$

Функция f обозначает либо E_z , либо H_z . С помощью преобразований (8.2.3) — (8.2.6) можно в выражениях



Ф и г. 8.2.2. Вектор \mathbf{F} и его компоненты, параллельные декартовым и полярным координатам.

(1.4.16) — (1.4.19) перейти к цилиндрическим координатам:

$$E_r = -\frac{i}{\kappa^2} \left(\beta \frac{\partial E_z}{\partial r} + \omega \mu \frac{1}{r} \frac{\partial H_z}{\partial \phi} \right), \quad (8.2.7)$$

$$E_\phi = -\frac{i}{\kappa^2} \left(\beta \frac{1}{r} \frac{\partial E_z}{\partial \phi} - \omega \mu \frac{\partial H_z}{\partial r} \right), \quad (8.2.8)$$

$$H_r = -\frac{i}{\kappa^2} \left(\beta \frac{\partial H_z}{\partial r} - \omega \varepsilon \frac{1}{r} \frac{\partial E_z}{\partial \phi} \right), \quad (8.2.9)$$

$$H_\phi = -\frac{i}{\kappa^2} \left(\beta \frac{1}{r} \frac{\partial H_z}{\partial \phi} + \omega \varepsilon \frac{\partial E_z}{\partial r} \right). \quad (8.2.10)$$

Здесь

$$\kappa^2 = k^2 - \beta^2 \quad (8.2.11)$$

и

$$k^2 = \omega^2 \epsilon \mu. \quad (8.2.12)$$

Параметр β является постоянной распространения (волновым числом) в направлении оси z . В итоге мы должны перейти в волновых уравнениях (1.4.22) и (1.4.23) к цилиндрическим координатам. Осуществим обратное преобразование в (8.2.1) и (8.2.2)

$$r = \sqrt{x^2 + y^2}, \quad (8.2.13)$$

$$\phi = \arctg \frac{y}{x} \quad (8.2.14)$$

и с помощью полученных выражений вычислим первые производные от E_z

$$\frac{\partial E_z}{\partial x} = \frac{\partial E_z}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial E_z}{\partial \phi} \frac{\partial \phi}{\partial x} = \frac{x}{r} \frac{\partial E_z}{\partial r} - \frac{y}{r^2} \frac{\partial E_z}{\partial \phi}, \quad (8.2.15)$$

$$\frac{\partial E_z}{\partial y} = \frac{\partial E_z}{\partial r} \frac{\partial r}{\partial y} + \frac{\partial E_z}{\partial \phi} \frac{\partial \phi}{\partial y} = \frac{y}{r} \frac{\partial E_z}{\partial r} + \frac{x}{r^2} \frac{\partial E_z}{\partial \phi}, \quad (8.2.16)$$

а также вторые производные

$$\begin{aligned} \frac{\partial^2 E_z}{\partial x^2} = & \left(\frac{1}{r} - \frac{x^2}{r^3} \right) \frac{\partial E_z}{\partial r} + \frac{x}{r} \left(\frac{x}{r} \frac{\partial^2 E_z}{\partial r^2} - \frac{y}{r^2} \frac{\partial^2 E_z}{\partial r \partial \phi} \right) + \\ & + \frac{2xy}{r^4} \frac{\partial E_z}{\partial \phi} - \frac{y}{r^2} \left(\frac{x}{r} \frac{\partial^2 E_z}{\partial r \partial \phi} - \frac{y}{r^2} \frac{\partial^2 E_z}{\partial \phi^2} \right), \end{aligned} \quad (8.2.17)$$

$$\begin{aligned} \frac{\partial^2 E_z}{\partial y^2} = & \left(\frac{1}{r} - \frac{y^2}{r^3} \right) \frac{\partial E_z}{\partial r} + \frac{y}{r} \left(\frac{y}{r} \frac{\partial^2 E_z}{\partial r^2} + \frac{x}{r^2} \frac{\partial^2 E_z}{\partial r \partial \phi} \right) - \\ & - \frac{2xy}{r^4} \frac{\partial E_z}{\partial \phi} + \frac{x}{r^2} \left(\frac{y}{r} \frac{\partial^2 E_z}{\partial \phi \partial r} + \frac{x}{r^2} \frac{\partial^2 E_z}{\partial \phi^2} \right). \end{aligned} \quad (8.2.18)$$

Подстановка этих выражений в (1.4.22) приводит к волновому уравнению в цилиндрической системе координат

$$\frac{\partial^2 E_z}{\partial r^2} + \frac{1}{r} \frac{\partial E_z}{\partial r} + \frac{1}{r^2} \frac{\partial^2 E_z}{\partial \phi^2} + \kappa^2 E_z = 0. \quad (8.2.19)$$

Аналогичное уравнение для H_z следует из (1.4.23)

$$\frac{\partial^2 H_z}{\partial r^2} + \frac{1}{r} \frac{\partial H_z}{\partial r} + \frac{1}{r^2} \frac{\partial^2 H_z}{\partial \phi^2} + \kappa^2 H_z = 0. \quad (8.2.20)$$

Волновые уравнения для E_z и H_z являются здесь строгими, так как они применяются в областях с однородным показателем преломления. Итак, получены все уравнения, необходимые для решения задачи о распространении мод в оптическом волокне с оболочкой. Для z -й компоненты электрического поля решение будем искать в виде

$$E_z = AF(r) e^{iv\phi}. \quad (8.2.21)$$

Множитель, зависящий от координаты z и времени,

$$e^{i(\omega t - \beta z)} \quad (8.2.22)$$

умножается на все компоненты поля и поэтому опущен здесь и далее во всех выражениях. Постоянная v в формуле (8.2.21) может быть как положительным, так и отрицательным числом, но обязательно целым, чтобы имела место периодичность по ϕ с периодом 2π .

Подстановка выражения (8.2.21) в (8.2.19) приводит к дифференциальному уравнению для $F(r)$

$$\frac{d^2 F}{dr^2} + \frac{1}{r} \frac{dF}{dr} + \left(\kappa^2 - \frac{v^2}{r^2} \right) F = 0, \quad (8.2.23)$$

которое является хорошо изученным уравнением Бесселя [41]. Поскольку это дифференциальное уравнение второго порядка, то должны иметь место два линейно независимых решения. Существует несколько способов выбора двух независимых решений (8.2.23). Можно использовать функцию Бесселя $J_v(\kappa r)$ и функцию Неймана $N_v(\kappa r)$. Эти функции для очень больших действительных значений их аргумента ведут себя подобно косинусам и синусам. Функция Бесселя J_v остается конечной в начале координат, тогда как функция Неймана N_v имеет особенность при $r = 0$. Другой набор линейно независимых решений состоит из функций Ханкеля первого и второго рода $H_v^{(1)}(\kappa r)$ и $H_v^{(2)}(\kappa r)$. Эти функции имеют особенность при $r = 0$. Для больших действительных значений аргумента $H_v^{(1)}$ представляет бегущую волну (экспоненту), которую при выборе временной зависимости (8.2.22) рассматривают как волну, распространяющуюся в направлении уменьшения r . Функция Ханкеля второго рода $H_v^{(2)}$ в этом случае представляет волну, распространяющуюся в направлении увеличения r , т. е. от оси структуры.

Представляют интерес также функции Ханкеля для мнимых значений κ

$$\kappa = i\gamma. \quad (8.2.24)$$

В этом случае $H_v^{(1)}(i\gamma r)$ становится пропорциональной $e^{-\gamma r}$ для больших значений аргумента, а $H_v^{(2)}(i\gamma r)$ пропорциональна $e^{\gamma r}$. Очевидно, что только функция $H_v^{(1)}(i\gamma r)$ с экспоненциальным спаданием является с физической точки зрения подходящей для описания направляемых мод вне сердцевины волокна. Функция $H_v^{(2)}(i\gamma r)$ должна быть отброшена, поскольку она экспоненциально растет с ростом r и поэтому не может правильно описывать распределение поля волны, которое концентрируется у сердцевины волокна.

Все четыре рассмотренные функции и любая их линейная комбинация известны как цилиндрические функции. Если специально не выделять какую-либо из них, то решение (8.2.23) можно записать как $Z(\kappa r)$. Обозначения, используемые здесь для функций Неймана, такие же, как у Янке и Эмде [41] и Градштейна и Рыжика [61]. Другие авторы иногда используют для функции Неймана символ Y_v . Функцию Ханкеля с мнимым аргументом часто обозначают новым символом. Она пропорциональна так называемой модифицированной функции Ханкеля K_v . Различные обозначения цилиндрических функций могут внести путаницу, поэтому более разумно использовать одинаковые символы для одинаковых функций, ясно указывая их аргументы, вместо того чтобы менять символы, используемые для самих функций. Каждая из четырех функций может быть выражена как линейная комбинация двух других функций. Свойства этих функций изложены в ряде книг, например [8, 11, 61].

Для описания поведения поля во внутренних и внешних областях волокна требуются различные функции. Существование физической направляемой волны накладывает на эти функции, являющиеся решениями (8.2.23), ограничения, которые состоят в том, что внутри сердцевины при $r = 0$ они должны оставаться конечными, а вне сердцевины при $r \rightarrow \infty$ описывать спадающее поле.

Имея теперь набор всех необходимых величин, можно записать формулы для электромагнитного поля оптического

волокна. Компоненты, соответствующие координатам r и ϕ , сразу получаются из z -й компоненты с помощью выражений (8.2.7) — (8.2.10).

Обозначим радиус сердцевинны через a . Для $r < a$ получим

$$E_z = AJ_v(\kappa r) e^{iv\phi}, \quad (8.2.25)$$

$$H_z = BJ_v(\kappa r) e^{iv\phi}, \quad (8.2.26)$$

$$E_r = -\frac{i}{\kappa^2} \left[\beta \kappa AJ'_v(\kappa r) + i\omega\mu_0 \frac{v}{r} BJ_v(\kappa r) \right] e^{iv\phi}, \quad (8.2.27)$$

$$E_\phi = -\frac{i}{\kappa^2} \left[i\beta \frac{v}{r} AJ_v(\kappa r) - \kappa\omega\mu_0 BJ'_v(\kappa r) \right] e^{iv\phi}, \quad (8.2.28)$$

$$H_r = -\frac{i}{\kappa^2} \left[-i\omega\varepsilon_1 \frac{v}{r} AJ_v(\kappa r) + \kappa\beta BJ'_v(\kappa r) \right] e^{iv\phi}, \quad (8.2.29)$$

$$H_\phi = -\frac{i}{\kappa^2} \left[\kappa\omega\varepsilon_1 AJ'_v(\kappa r) + i\beta \frac{v}{r} BJ_v(\kappa r) \right] e^{iv\phi}. \quad (8.2.30)$$

Штрих означает дифференцирование относительно аргумента κr функции Бесселя (а не r). Соотношение между κ , β и k_1 имеет вид

$$\kappa^2 = k_1^2 - \beta^2, \quad (8.2.31a)$$

где

$$k_1^2 = \omega^2 \varepsilon_1 \mu_0. \quad (8.2.31b)$$

Диэлектрическая постоянная сердцевинны связана с показателем преломления соотношением

$$n_1^2 = \frac{\varepsilon_1}{\varepsilon_0}. \quad (8.2.32)$$

Поле вне сердцевинны, $r > a^1$), определяется следующим образом:

$$E_z = CH_v^{(1)}(i\gamma r) e^{iv\phi}, \quad (8.2.33)$$

$$H_z = DH_v^{(1)}(i\gamma r) e^{iv\phi}, \quad (8.2.34)$$

$$E_r = -\frac{1}{\gamma^2} \left[\beta \gamma CH_v^{(1)'}(i\gamma r) + \omega\mu_0 \frac{v}{r} DH_v^{(1)}(i\gamma r) \right] e^{iv\phi}, \quad (8.2.35)$$

$$E_\phi = -\frac{1}{\gamma^2} \left[\beta \frac{v}{r} CH_v^{(1)}(i\gamma r) - \gamma\omega\mu_0 DH_v^{(1)'}(i\gamma r) \right] e^{iv\phi}, \quad (8.2.36)$$

$$H_r = -\frac{1}{\gamma^2} \left[-\omega\varepsilon_2 \frac{v}{r} CH_v^{(1)}(i\gamma r) + \gamma\beta DH_v^{(1)}(i\gamma r) \right] e^{iv\phi}, \quad (8.2.37)$$

$$H_\phi = -\frac{1}{\gamma^2} \left[\gamma\omega\varepsilon_2 CH_v^{(1)'}(i\gamma r) + \beta \frac{v}{r} DH_v^{(1)}(i\gamma r) \right] e^{iv\phi}. \quad (8.2.38)$$

¹) В уравнениях везде опущен множитель (8.2.22),

Штрих опять означает дифференцирование относительно аргумента, которым в этом случае является $i\gamma r$. Величины γ , β и k_2 связаны соотношением

$$\gamma^2 = \beta^2 - k_2^2, \quad (8.2.39a)$$

где

$$k_2^2 = \omega^2 \varepsilon_2 \mu_0. \quad (8.2.39b)$$

Постоянные A , B , C и D не определяются уравнениями Максвелла. Выражения для компонент поля, разумеется, являются решениями уравнений Максвелла, но, чтобы они правильно описывали моды оптического волокна [80], они должны еще удовлетворять граничным условиям (1.5.3) и (1.5.4). Так как существуют две тангенциальные компоненты для электрического поля и две для магнитного поля, применение граничных условий дает четыре уравнения. Поскольку в выражениях для компонент поля четыре неопределенные константы, то число уравнений совпадает с числом неизвестных. Однако наряду с амплитудными коэффициентами существует постоянная распространения β , которую также необходимо определить. Ее нахождение не представляет трудности. Граничные условия приводят к четырем однородным уравнениям. Однородные уравнения имеют решения только в том случае, если определитель системы уравнений равен нулю. Это дополнительное условие является достаточным для определения постоянной распространения β . Условие равенства определителя нулю называется уравнением собственных значений, так как из него определяются собственные значения β оператора волнового уравнения¹⁾. Из уравнения собственных значений определяются собственные значения постоянной распространения направляемых мод, или, что то же самое, волновых чисел этих мод.

Граничные условия позволяют приравнять выражения для компонент поля E_z , E_ϕ , H_z и H_ϕ на границе при

¹⁾ Правильнее сказать: «собственные значения оператора, включающего дифференциальный оператор волнового уравнения и граничные условия». Термин «собственное значение» заимствован из математики, точнее, из ее раздела, называемого функциональным анализом, где решается задача о нахождении собственных значений, а также собственных функций линейных операторов. Здесь рассматривается частный случай такой задачи.— *Прим. ред.*

$r = a$. Получается следующая система уравнений:

$$AJ_v(\kappa a) - CH_v^{(1)}(i\gamma a) = 0, \quad (8.2.40)$$

$$\frac{\beta}{\kappa^2} \frac{v}{a} AJ_v(\kappa a) + i \frac{\omega \mu_0}{\kappa} BJ'_v(\kappa a) + \frac{\beta v}{\gamma^2 a} CH_v^{(1)'}(i\gamma a) - \frac{\omega \mu_0}{\gamma} DH_v^{(1)'}(i\gamma a) = 0, \quad (8.2.41)$$

$$BJ_v(\kappa a) - DH_v^{(1)}(i\gamma a) = 0, \quad (8.2.42)$$

$$-i \frac{\omega \varepsilon_1}{\kappa} AJ'_v(\kappa a) + \frac{\beta}{\kappa^2} \frac{v}{a} BJ_v(\kappa a) + \frac{\omega \varepsilon_2}{\gamma} CH_v^{(1)'}(i\gamma a) + \frac{\beta}{\gamma^2} \frac{v}{a} DH_v^{(1)'}(i\gamma a) = 0. \quad (8.2.43)$$

Уравнения (8.2.40) и (8.2.42) связывают коэффициенты A , C и B , D :

$$C = \frac{J_v(\kappa a)}{H_v^{(1)}(i\gamma a)} A \quad (8.2.44)$$

и

$$D = \frac{J_v(\kappa a)}{H_v^{(1)}(i\gamma a)} B. \quad (8.2.45)$$

Соотношение между коэффициентами A и B получается подстановкой (8.2.44) и (8.2.45) в (8.2.43):

$$B = \frac{i}{v} \frac{a \kappa \gamma [\varepsilon_1 \gamma J'_v(\kappa a) H_v^{(1)}(i\gamma a) + i \varepsilon_2 \kappa J_v(\kappa a) H_v^{(1)'}(i\gamma a)]}{\omega (\varepsilon_1 - \varepsilon_2) \mu_0 \beta J_v(\kappa a) H_v^{(1)}(i\gamma a)} A. \quad (8.2.46)$$

Равенство

$$\kappa^2 + \gamma^2 = k_1^2 - k_2^2 = \omega^2 (\varepsilon_1 - \varepsilon_2) \mu_0 \quad (8.2.47)$$

было использовано для упрощения этого выражения. Уравнение (8.2.41) можно было бы использовать вместо (8.2.43). Полученное при этом выражение для B/A может быть преобразовано к виду (8.2.46) с помощью уравнения собственных значений [см. (8.2.53)]. Уравнение собственных значений получается из требования, чтобы определитель системы уравнений (8.2.40) — (8.2.43) равнялся нулю:

$$\begin{vmatrix} J_v(\kappa a) & 0 & -H_v^{(1)}(i\gamma a) & 0 \\ \frac{v}{a} \frac{\beta}{\kappa^2} J_v(\kappa a) & i \frac{\omega \mu_0}{\kappa} J'_v(\kappa a) & \frac{v}{a} \frac{\beta}{\gamma^2} H_v^{(1)}(i\gamma a) & -\frac{\omega \mu_0}{\gamma} H_v^{(1)'}(i\gamma a) \\ 0 & J_v(\kappa a) & 0 & -H_v^{(1)}(i\gamma a) \\ -i \frac{\omega \varepsilon_1}{\kappa} J'_v(\kappa a) & \frac{v}{a} \frac{\beta}{\kappa^2} J_v(\kappa a) & \frac{\omega \varepsilon_2}{\gamma} H_v^{(1)'}(i\gamma a) & \frac{v}{a} \frac{\beta}{\gamma^2} H_v^{(1)}(i\gamma a) \end{vmatrix} = 0. \quad (8.2.48)$$

Вычисление определителя приводит к уравнению собственных значений

$$\left[\frac{\varepsilon_1}{\varepsilon_2} \frac{a\gamma^2}{\kappa} \frac{J_{\nu}^{(1)'}(\kappa a)}{J_{\nu}(\kappa a)} + i\gamma a \frac{H_{\nu}^{(1)'}(i\gamma a)}{H_{\nu}^{(1)}(i\gamma a)} \right] \left[\frac{a\gamma^2}{\kappa} \frac{J_{\nu}'(\kappa a)}{J_{\nu}(\kappa a)} + i\gamma a \frac{H_{\nu}^{(1)'}(i\gamma a)}{H_{\nu}^{(1)}(i\gamma a)} \right] = \\ = \left[\nu \left(\frac{\varepsilon_1}{\varepsilon_2} - 1 \right) \frac{\beta k_2}{\kappa^2} \right]^2. \quad (8.2.49)$$

Снова формула (8.2.47) была использована для упрощения этого уравнения. Приведенные здесь соотношения представляют собой решение задачи о распространении направляемых мод в оптическом волокне с оболочкой. Приближенное решение уравнения собственных значений дано в разд. 8.6. В формуле (8.2.21) вместо $e^{i\nu\phi}$ можно использовать $e^{-i\nu\phi}$. Это привело бы к изменению знака ν во всех формулах (исключая индекс ν цилиндрических функций), которые, однако, остались бы теми же самыми, потому что ν входит в уравнение (8.2.23) только в виде ν^2 . Такое изменение не повлияло бы на собственное значение β . Прибавляя новые моды с противоположным знаком ν к старым модам, получим выражения для поля, которые будут содержать $\cos \nu\phi$ и $\sin \nu\phi$ вместо экспоненциальной функции. Вычитание же новых мод из старых даст другой набор мод с синусом и косинусом от $\nu\phi$. Выражения для мод часто записываются через синус и косинус от $\nu\phi$ вместо экспоненциальной функции, использованной здесь. Эта форма описания поля используется ниже (формулы (8.6.59) — (8.6.64)).

Оказывается, что моды диэлектрического волновода имеют шесть компонент поля и невозможно разделить их на поперечно-электрические и поперечно-магнитные моды. Моды диэлектрических волноводов являются, таким образом, более сложными, чем моды полых металлических волноводов. Моды оптического волокна являются гибридными. Исключение составляет случай $\nu = 0$, для которого правая часть уравнения (8.2.49) равна нулю. При этом справедливы два различных уравнения собственных значений:

$$\text{ТМ-моды: } \frac{\varepsilon_1}{\varepsilon_2} \frac{\gamma}{\kappa} \frac{J_1(\kappa a)}{J_0(\kappa a)} + i \frac{H_1^{(1)}(i\gamma a)}{H_0^{(1)}(i\gamma a)} = 0 \quad (8.2.50)$$

и

$$\text{ТЕ-моды: } \frac{\gamma}{\kappa} \frac{J_1(\kappa a)}{J_0(\kappa a)} + i \frac{H_1^{(1)}(i\gamma a)}{H_0^{(1)}(i\gamma a)} = 0. \quad (8.2.51)$$

Для получения этих результатов было использовано соотношение

$$Z'_0 = -Z_1. \quad (8.2.52)$$

Для мод, удовлетворяющих уравнению собственных значений (8.2.51), из выражения (8.2.46) находим, что $B = \infty$, если $\nu = 0$. Чтобы сохранить B конечным, необходимо положить $A = 0$. Это означает, что в этом случае продольная компонента E_z исчезающе мала. Моды становятся поперечно-электрическими, или ТЕ-модами. В случае когда удовлетворяется уравнение собственных значений (8.2.50), необходимо сначала воспользоваться равенством (8.2.49), чтобы исключить ν из знаменателя (8.2.46). В результате имеем

$$B = i\nu \frac{\omega(\epsilon_1 - \epsilon_2) \beta J_\nu(\kappa a) H_\nu^{(1)}(i\gamma a)}{\kappa \gamma a [\gamma J'_\nu(\kappa a) H_\nu^{(1)}(i\gamma a) + i\kappa J_\nu(\kappa a) H_{\nu'}^{(1)}(i\gamma a)]} A. \quad (8.2.53)$$

Аналогичное выражение можно получить, если использовать уравнение (8.2.41) вместо (8.2.43). Теперь можно устремить $\nu \rightarrow 0$, откуда $B = 0$ при условии, что знаменатель (8.2.53) не равен нулю. Действительно, если мода удовлетворяет уравнению собственных значений (8.2.50), знаменатель (8.2.53) не равен нулю. Равенство $B = 0$ означает, что продольная компонента магнитного поля H_z исчезающе мала. Моды, удовлетворяющие уравнению собственных значений (8.2.50), являются, таким образом, поперечно-магнитными, или ТМ-модами.

Важным параметром любой моды является частота отсечки, или критическая частота¹⁾. Мода перестает существовать как физическая волновая структура или отсекается, когда ее поле больше не уменьшается при удалении от сердцевинны. Степень уменьшения поля с увеличением r определяется значением величины γ . Ранее упоминалось, что функция $H_\nu^{(1)}(i\gamma r)$ экспоненциально убывает с увели-

¹⁾ В английской технической литературе применяется первый термин, в русской используются оба термина, но чаще последний. При переводе мы сохранили первый термин (второй дан здесь для пояснения), считая, что это не приведет к недоразумениям. — *Прим. ред.*

чением значений аргумента. Асимптотическое приближение для больших значений аргумента имеет вид [11, 61]

$$H_v^{(1)}(i\gamma r) = \sqrt{\frac{2}{\pi i \gamma r}} e^{-i(\pi v/2 + \pi/4)} e^{-\gamma r} \text{ для } \gamma r \gg 1. \quad (8.2.54)$$

При больших значениях γ поле плотно сконцентрировано внутри и вблизи сердцевины. С уменьшением γ поле перераспределяется в пространство вне сердцевины. При $\gamma = 0$ поле выходит из волновода. Частота, при которой это происходит, называется частотой отсечки. Условием отсечки является, таким образом, соотношение

$$\gamma = \sqrt{\beta^2 - k_2^2} = 0. \quad (8.2.55)$$

Решения уравнения собственных значений на частотах, больших частоты отсечки, можно получить так же, как это было сделано Шлезингером, Дамантом и Вигантом [83]. Как и эти авторы, преобразуем уравнение собственных значений к другому виду. Введем обозначения

$$J^+ = \frac{1}{\kappa a} \frac{J_{v+1}(\kappa a)}{J_v(\kappa a)}, \quad (8.2.56)$$

$$J^- = \frac{1}{\kappa a} \frac{J_{v-1}(\kappa a)}{J_v(\kappa a)}, \quad (8.2.57)$$

$$H^+ = \frac{1}{i\gamma a} \frac{H_{v+1}^{(1)}(i\gamma a)}{H_v^{(1)}(i\gamma a)}, \quad (8.2.58)$$

$$H^- = \frac{1}{i\gamma a} \frac{H_{v-1}^{(1)}(i\gamma a)}{H_v^{(1)}(i\gamma a)}. \quad (8.2.59)$$

Используя следующие функциональные соотношения для цилиндрических функций [11, 61]

$$Z'_v = \frac{1}{2} (Z_{v-1} - Z_{v+1}) \quad (8.2.60)$$

и обозначение

$$\varepsilon = \frac{\varepsilon_1}{\varepsilon_2}, \quad (8.2.61)$$

запишем уравнение собственных значений (8.2.49) после деления на $a^4 \gamma^4$ в следующей форме:

$$\begin{aligned} [\varepsilon (J^- - J^+) - (H^- - H^+)] [(J^- - J^+) - (H^- - H^+)] = \\ = \left[\frac{2v(\varepsilon - 1)\beta k_2}{a^2 \gamma^2 \kappa^2} \right]^2. \end{aligned} \quad (8.2.62)$$

Перегруппировка членов приводит к

$$\begin{aligned} & -(\epsilon J^- - H^-)(J^+ - H^+) - (\epsilon J^+ - H^+)(J^- - H^-) + \\ & + (\epsilon J^+ - H^+)(J^+ - H^+) + (\epsilon J^- - H^-)(J^- - H^-) = \\ & = \left[\frac{2\nu(\epsilon - 1)\beta k_2}{a^2\gamma^2\kappa^2} \right]^2. \end{aligned} \quad (8.2.63)$$

Используя функциональное соотношение цилиндрических функций [41, 61]

$$Z_{\nu+1}(z) + Z_{\nu-1}(z) = \frac{2\nu}{z} Z_{\nu}(z), \quad (8.2.64)$$

получим

$$J^+ + J^- = \frac{2\nu}{(\gamma a)^2} \quad (8.2.65)$$

и

$$H^+ + H^- = -\frac{2\nu}{(\gamma a)^2}. \quad (8.2.66)$$

С помощью этих соотношений уравнение (8.2.63) преобразуется к виду

$$\begin{aligned} & -2(\epsilon J^- - H^-)(J^+ - H^+) - 2(\epsilon J^+ - H^+)(J^- - H^-) + \\ & + \frac{2\nu}{a^2} \left(\frac{\epsilon}{\kappa^2} + \frac{1}{\gamma^2} \right) [(J^+ - H^+) + (J^- - H^-)] = \\ & = \left[\frac{2\nu(\epsilon - 1)\beta k_2}{a^2\gamma^2\kappa^2} \right]^2. \end{aligned} \quad (8.2.67)$$

Далее с помощью формул (8.2.65) и (8.2.66) найдем

$$(J^+ - H^+) + (J^- - H^-) = \frac{2\nu}{a^2} \left[\frac{1}{\kappa^2} + \frac{1}{\gamma^2} \right]. \quad (8.2.68)$$

Используя соотношения между постоянными κ , γ , β , k_1 и k_2 , получим уравнение

$$\left(\frac{\epsilon}{\kappa^2} + \frac{1}{\gamma^2} \right) \left(\frac{1}{\kappa^2} + \frac{1}{\gamma^2} \right) = \left[\frac{(\epsilon - 1)\beta k_2}{\kappa^2\gamma^2} \right]^2. \quad (8.2.69)$$

Уравнение собственных значений (8.2.67) принимает простую форму

$$(\epsilon J^- - H^-)(J^+ - H^+) + (\epsilon J^+ - H^+)(J^- - H^-) = 0. \quad (8.2.70)$$

Уравнения собственных значений (8.2.49) и (8.2.70) являются, конечно, полностью эквивалентными. Для изучения решений в области отсечки лучше пользоваться уравнением (8.2.70).

Известно, что условием отсечки является $\gamma = 0$. Поскольку аргументы функций Ханкеля в этом случае исчезающе малы, необходимо использовать приближение для малых значений аргументов [11, 61]

$$H_0^{(1)}(i\gamma a) = 1 + \frac{2i}{\pi} \ln \frac{i\Gamma\gamma a}{2} = \frac{2i}{\pi} \ln \frac{\Gamma\gamma a}{2}, \quad \text{где} \quad \Gamma = 1,781672, \quad (8.2.71)$$

и

$$H_v^{(1)}(i\gamma a) = -\frac{i(v-1)!}{\pi} \left(\frac{2}{i\gamma a} \right)^v \quad \text{для } v = 1, 2, 3, \dots \quad (8.2.72)$$

Из этих уравнений следуют соотношения

$$H^+ = -\frac{2v}{(a\gamma)^2} \quad \text{для } v = 1, 2, 3, \dots, \quad (8.2.73)$$

$$H^- = -\ln \frac{\Gamma\gamma a}{2} \quad \text{для } v = 1, \quad (8.2.74)$$

$$H^- = \frac{1}{2(v-1)} \quad \text{для } v = 2, 3, 4, \dots \quad (8.2.75)$$

Из формулы (8.2.70) для малых значений γ и $v \neq 0$ с помощью (8.2.73) получим

$$(\epsilon J_{v-1} - \kappa a H^- J_v) (a^2 \gamma^2 J_{v+1} + 2v\kappa a J_v) + \\ + (\epsilon a^2 \gamma^2 J_{v+1} + 2v\kappa a J_v) (J_{v-1} - \kappa a H^- J_v) = 0. \quad (8.2.76)$$

При $\gamma \rightarrow 0$ необходимо отдельно рассмотреть случаи $v = 1$ и $v > 1$. Начнем с $v = 1$. Из равенства (8.2.76) при $\gamma \rightarrow 0$ получим

$$[2(\kappa a) J_1(\kappa a)]^2 \ln \frac{\Gamma\gamma a}{2} = 0. \quad (8.2.77)$$

Решение этого уравнения имеет вид

$$J_1(\kappa a) = 0. \quad (8.2.78)$$

Другое возможное решение $\kappa a = 0$ также учтено в (8.2.78)

Для $v > 1$ из равенства (8.2.76) в пределе при $\gamma \rightarrow 0$ найдем

$$J_v(\kappa a) \left[(\epsilon + 1) J_{v-1}(\kappa a) - \frac{\kappa a}{v-1} J_v(\kappa a) \right] = 0. \quad (8.2.79)$$

Это уравнение допускает два решения:

$$J_\nu(\kappa a) = 0 \text{ для } \kappa a \neq 0 \text{ и } \nu = 2, 3, 4, \dots \quad (8.2.80)$$

и

$$(\varepsilon + 1) J_{\nu-1}(\kappa a) = \frac{\kappa a}{\nu-1} J_\nu(\kappa a) \text{ для } \nu = 2, 3, 4, \dots \quad (8.2.81)$$

Как отмечено в (8.2.80), решение $\kappa a = 0$ должно быть здесь исключено. Этот очень важный результат следует из (8.2.70). Чтобы показать это, необходимо использовать приближение для функций Бесселя от малого аргумента [44, 61]

$$J_0(\kappa a) = 1 \quad (8.2.82)$$

и

$$J_\nu(\kappa a) = \frac{1}{\nu!} \left(\frac{\kappa a}{2} \right)^\nu \text{ для } \nu = 1, 2, 3, \dots \quad (8.2.83)$$

С учетом этих приближений найдем

$$J^+ = \frac{1}{2(\nu+1)} \text{ для } \nu = 1, 2, 3, \dots \quad (8.2.84)$$

и

$$J^- = \frac{2\nu}{(a\kappa)^2} \text{ для } \nu = 1, 2, 3, \dots \quad (8.2.85)$$

Если γ и κ становятся одновременно исчезающе малыми, то равенство (8.2.70) для $\nu > 1$ принимает вид

$$\left[\frac{2\varepsilon\nu}{(a\kappa)^2} - \frac{1}{2(\nu-1)} \right] \left[\frac{1}{2(\nu+1)} + \frac{2\nu}{(a\gamma)^2} \right] + \\ + \left[\frac{\varepsilon}{2(\nu+1)} + \frac{2\nu}{(a\gamma)^2} \right] \left[\frac{2\nu}{(a\kappa)^2} - \frac{1}{2(\gamma-1)} \right] = 0. \quad (8.2.86)$$

Когда κ и γ стремятся к нулю, уравнение (8.2.86) переходит в

$$\frac{4\nu^2(\varepsilon+1)}{(a^2\kappa\gamma)^2} = 0 \text{ для } \nu = 2, 3, 4, \dots \quad (8.2.87)$$

То, что значение $\kappa a = 0$ не удовлетворяет этому уравнению, доказывает тот факт, что оно не может быть и решением уравнения отсечки (8.2.80). С другой стороны, при $\nu = 1$ из формулы (8.2.70) имеем

$$\left[\frac{2\varepsilon\nu}{(a\kappa)^2} - \ln \frac{2}{a\gamma\Gamma} \right] \left[\frac{1}{2(\nu+1)} + \frac{2\nu}{(a\gamma)^2} \right] + \\ + \left[\frac{\varepsilon}{2(\nu+1)} + \frac{2\nu}{(a\gamma)^2} \right] \left[\frac{2\nu}{(a\kappa)^2} - \ln \frac{2}{a\gamma\Gamma} \right] = 0. \quad (8.2.88)$$

Когда κ и γ стремятся к нулю, получаем

$$\frac{4\nu}{(a\gamma)^2} \left[\frac{\nu(\varepsilon+1)}{(a\kappa)^2} - \ln \frac{2}{a\gamma\Gamma} \right] = 0. \quad (8.2.89)$$

Это уравнение удовлетворяется, поскольку логарифм стремится к бесконечности, когда $\gamma \rightarrow 0$. Решение имеет вид

$$\kappa a = 0 \quad \text{для} \quad \nu = 1. \quad (8.2.90)$$

Теперь осталось исследовать случай $\nu = 0$. Для малых значений γ равенство (8.2.50) принимает вид

$$\frac{1}{\varepsilon} \frac{\kappa}{\gamma} \frac{J_0(\kappa a)}{J_1(\kappa a)} = -a\gamma \ln \frac{2}{a\gamma\Gamma}. \quad (8.2.91)$$

Поскольку γa стремится к нулю быстрее, чем логарифм к бесконечности, произведение в правой части стремится к нулю при $\gamma \rightarrow 0$. Решение этого уравнения имеет, таким образом, вид

$$\left. \begin{array}{l} \text{TE-} \\ \text{TM-} \end{array} \right\} \text{моды: } J_0(\kappa_c a) = 0. \quad (8.2.92)$$

Стремление κ к нулю не приводит к решению (8.2.91). Поскольку равенство (8.2.51) отличается от (8.2.50) только множителем ε , условие отсечки волн (8.2.92) справедливо для обеих поляризаций.

Таким образом, находим полное решение задачи об отсечке волн для всех мод. Суммируем результаты решения уравнения собственных значений направляемых мод (с $\nu \neq 0$) оптического волокна для условия отсечки мод. Имеем следующие условия:

$$\text{HE}_{11}: \kappa_c a = 0 \quad \text{для} \quad \nu = 1, \quad (8.2.93)$$

$$\left. \begin{array}{l} \text{EH}_{\nu\mu}: \\ \text{(для } \nu=1) \text{ HE}_{1\mu}: \end{array} \right\} \kappa_c a = w_{\nu\mu} \quad \text{для} \quad \nu = 1, 2, 3, \dots \quad (8.2.94)$$

и подразумеваемое условие ($\varepsilon = \varepsilon_1/\varepsilon_2$)

$$\begin{aligned} \text{HE}_{\nu\mu}: (\varepsilon + 1) J_{\nu-1}(\kappa_c a) &= \frac{a\kappa_c}{\nu-1} J_{\nu}(\kappa_c a) \\ &\text{для } \nu = 2, 3, 4, \dots \end{aligned} \quad (8.2.95)$$

Параметр $w_{\nu\mu}$ является μ -м корнем уравнения

$$J_{\nu}(w_{\nu\mu}) = 0, \quad w_{\nu\mu} \neq 0. \quad (8.2.96)$$

Наиболее важным для данного рассмотрения режима отсечки направляемых мод является факт существования одной моды, частота отсечки которой равняется нулю. Частота отсечки f_c определяется при $\gamma = 0$ с помощью формул (8.2.31), (8.2.39) и (8.2.55)

$$f_c = \frac{\kappa_c}{2\pi \sqrt{(\varepsilon_1 - \varepsilon_2) \mu_0}}. \quad (8.2.97)$$

Только несимметричная мода низшего порядка при $\nu = 1$ может иметь нулевое значение κ_c и, следовательно, частоту отсечки $f_c = 0$. Эта мода, таким образом, может существовать при любой частоте и при любом диаметре стержня. Все другие моды не распространяются на частотах ниже их частот отсечки. Таким образом, имеется возможность работать с оптическим волокном, по которому распространяется лишь одна мода. Для этого требуется, чтобы волокно было достаточно тонким и все другие направляемые моды с более высокими частотами отсечки не могли распространяться.

Из формул (8.2.94) и (8.2.95) видно, что имеются два типа мод для каждого целого значения $\nu > 1$. Моды, частоты отсечки которых определяются условием (8.2.94), обозначаются как $EH_{\nu\mu}$, а моды, частоты отсечки которых определяются условием (8.2.95), обозначаются как $HE_{\nu\mu}$ [80]. Исключением являются моды $HE_{1\mu}$, частоты отсечки которых также определяются условием (8.2.94), и мода HE_{11} , частота отсечки которой определяется условием (8.2.93). Оба типа мод имеют шесть компонент поля. Существование двух типов мод аналогично наличию двух типов (ТЕ и ТМ) мод металлических волноводов.

При $\nu = 1$ имеет место только одно условие отсечки. В этом случае, однако, существует также два типа мод. Оба типа ($EH_{1\mu}$ и $HE_{1\mu}$) мод имеют одинаковую частоту отсечки. При частотах, отличных от частоты отсечки, оба типа мод имеют различные постоянные распространения. Поэтому они не являются вырожденными. При $\nu = 0$ мы имеем также невырожденные ТЕ- и ТМ-моды с одинаковыми значениями частот отсечки, которые задаются условием (8.2.92).

В табл. 8.2.1 приведены значения $\kappa_c a$ для различных сочетаний ν и μ , рассчитанные для HE-мод при $\varepsilon_1/\varepsilon_2 = 1,1$.

Таблица 8.2.1

Критические значения параметра k_2a для некоторых мод. Для НЕ-мод (кроме мод $HE_{1\mu}$, которые не зависят от ε) вычисления проводились для $\varepsilon_1/\varepsilon_2 = 1,1$

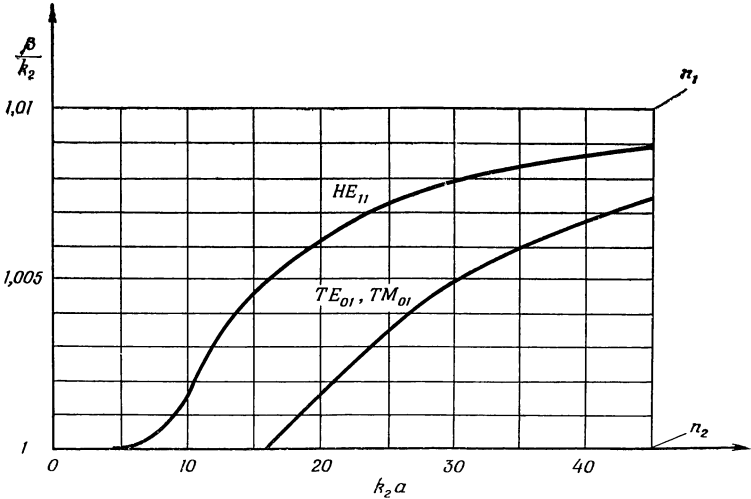
$\nu \backslash \mu$	1	2	3	
0	2,405	5,52	8,654	ТЕ, ТМ
1	0,000	3,832	7,016	НЕ
1	3,832	7,016	10,173	ЕН
2	2,445	5,538	8,665	НЕ
2	5,136	8,417	11,620	ЕН

Видно, что с увеличением частоты появляются новые моды. Первой модой является мода HE_{11} , распространяющаяся без отсечки. По мере увеличения частоты возникает возможность существования TE_{01} - и TM_{01} -мод. Интервал частот, при которых в волноводе распространяется одна мода, определяется неравенством

$$0 < f < \frac{2,405}{2\pi a \sqrt{(\varepsilon_1 - \varepsilon_2) \mu_0}}. \quad (8.2.98)$$

На фиг. 8.2.3 показана зависимость отношения β/k_2 от k_2a для мод HE_{11} , TE_{01} , TM_{01} . (Приближенные решения уравнения собственных значений будут получены в разделе 8.6.) Графики на фиг. 8.2.3 построены для $n_2 = 1$ и $n_1 = 1,01$. Постоянные распространения мод TE_{01} и TM_{01} почти одинаковы в силу малой разности показателей преломления сред волновода, поэтому на фигуре они представлены одной линией. Существование мод высшего порядка возможно при значениях k_2a , больших приведенных на данной фигуре. Их критические частоты больше критической частоты моды TE_{01} . Как уже отмечалось, мода HE_{11} отсечки не имеет. Это видно из фигуры, хотя на ней график для моды HE_{11} практически обрывается при $k_2a = 5$. Это вызвано тем, что при $k_2a < 5$ отношение β/k_2 становится очень близким к единице. В действитель-

ности же кривая проходит до $k_2 a = 0$. Все кривые начинаются на линии $\beta/k_2 = 1$ и стремятся асимптотически к значению $\beta/k_2 = n_1/n_2$. Чтобы показать асимптотическое



Ф и г. 8.2.3. Зависимость отношения постоянной распространения направляемых мод β к постоянной распространения плоской волны k_2 (в среде оболочки) от относительного радиуса волокна $k_2 a$.

поведение кривой для моды HE_{11} более ясно, на фиг. 8.2.4 приведена качественная диаграмма для зависимости круговой частоты ω от β . Диаграмма такого типа называется $(\omega - \beta)$ -диаграммой. Она удобна тем, что тангенс угла наклона

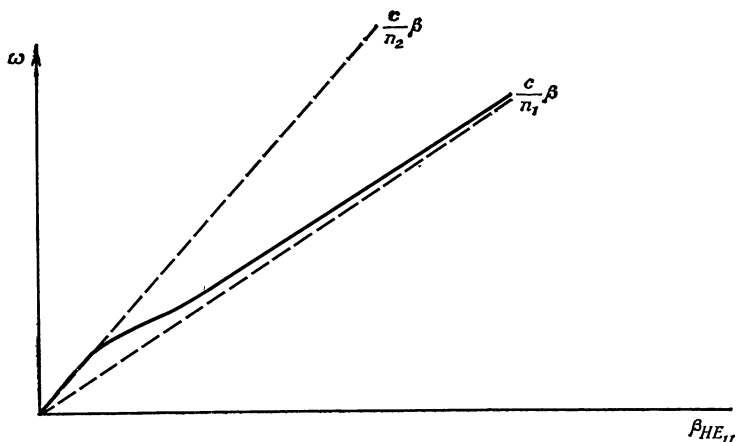
$$v_g = \frac{\partial \omega}{\partial \beta} \quad (8.2.99)$$

определяет групповую скорость моды в оптическом волокне. Очевидно, что групповая скорость является функцией частоты. Кривая имеет точку перегиба, где вторая производная

$$\frac{\partial^2 \omega}{\partial \beta^2} = \frac{\partial v_g}{\partial \beta} = \frac{\partial v_g}{\partial \omega} v_g \quad (8.2.100)$$

равна нулю. В окрестности этой точки групповая скорость приближенно не зависит от частоты. Это означает, что

профиль импульсов, распространяющихся на этой моде, почти не искажается. Точка перегиба появляется на частоте, при которой в волокне может распространяться более чем одна мода. При этом возможное число мод, однако, еще невелико. Поэтому можно работать с волокном на этой частоте и передавать импульсы с минимальным



Ф и г. 8.2.4. Частотная диаграмма для моды HE_{11} .

искажением. Дисперсия диэлектрической среды сердцевины и оболочки при этом не учитывается. Их учет несколько сместит точку перегиба [94].

Суммируем полученные результаты. Оптическое волокно в оболочке (или без нее) способно поддерживать направляемые моды. Число возможных мод зависит от значения $V = a\sqrt{k_1^2 - k_2^2}$ [см. формулу (8.2.97)]. При больших значениях V может распространяться много мод. Но можно изготовить такое оптическое волокно, в котором при заданной частоте будет распространяться только одна мода HE_{11} . Электромагнитная энергия направляемых мод переносится внутри сердцевины и вне ее. Чем выше частота моды по сравнению с частотой отсечки, тем большая энергия концентрируется внутри сердцевины. Моды волокна обладают дисперсией. Групповая скорость, так

же как и фазовая скорость, зависит от рабочей частоты. Однако для моды HE_{11} можно выбрать рабочую частоту таким образом, что групповая скорость будет приблизительно постоянной внутри узкого (относительно величины круговой частоты рабочего диапазона) диапазона частот. Частотный диапазон, внутри которого мода HE_{11} имеет постоянную групповую скорость, может быть достаточно широким, если его сравнивать с полосой частот в диапазоне СВЧ.

Оптические волокна очень перспективны как световые волноводы (световоды) для целей связи. Во время подготовки данной книги к печати было опубликовано сообщение о волокнах с потерями в 20 дБ/км [95]¹⁾. Теоретическое рассмотрение показывает, что снижение потерь в стекле возможно приблизительно до $5 \div 10$ дБ/км в диапазоне частот видимого света. При таких потерях оптические волокна можно было бы применять для целей передачи света на большие расстояния. Важным применением оптических волокон является также передача света и изображений на короткие расстояния. Можно изготавливать связки (пучки), содержащие большое число оптических волокон, которые способны передавать достаточно контрастное изображение [79]. Каждое волокно при этом переносит свет, соответствующий одной точке изображения. Передача изображения в этом случае существенно отличается от передачи изображения через оптические материалы с переменным показателем преломления, которые обсуждались в гл. 7. Передача некогерентного света через связки оптических волокон без формирования изображения находит применение при контроле источников света. Например, наличие света от источника, расположенного сзади автомобиля, может контролироваться посредством оптического волокна (световода), передающего часть этого света к приборной панели.

¹⁾ Согласно последним данным, опубликованным в *BSTJ*, 1973, т. 52, № 2, стр. 265, на длине волны 1,1 мкм получены волокна из кварца с затуханием 2,5 дБ/км в многоволновом режиме работы. — *Прим. ред.*

8.3. НАПРАВЛЯЕМЫЕ МОДЫ ПЛОСКОГО ВОЛНОВОДА

В предыдущем разделе были рассмотрены моды круглых оптических волокон. Математическое описание распространения этих мод довольно сложное. Еще труднее анализировать потери на излучение круглого оптического волокна. Однако для получения данных о передающих свойствах диэлектрических волноводов нет необходимости изучать сложные направляющие структуры. Существуют более простые структуры диэлектрических волноводов, физические свойства которых близки к свойствам круглых диэлектрических волноводов, а анализ проще. Поэтому рассмотрим моды идеализированной модели плоского диэлектрического волновода. Эта структура наиболее простая и на ее примере легче изучать преобразование мод и излучение из диэлектрического волновода. Результаты, полученные с помощью модели плоского волновода, обычно непосредственно применимы к круглому оптическому волокну¹⁾.

Плоский волновод схематически изображен на фиг. 8.1.2. Анализ плоского волновода как некоторой приближенной модели оптического волокна аналогичен рассмотрению проблемы дифракции, которую иногда удобно свести к двумерному случаю. Действительно, продольное сечение оптического волокна имеет сходство с плоским волноводом, который может рассматриваться как его двумерный аналог.

Для простоты в дальнейшем будем считать, что плоский волновод является бесконечно протяженным в направлении оси y и вариации поля в этом направлении отсутствуют. Математически это условие выражается соотношением

$$\frac{\partial}{\partial y} = 0. \quad (8.3.1)$$

В общем случае моды оптического волокна гибридные, но при $\nu = 0$ они разделяются на ТЕ- и ТМ-моды. Ограничение (8.3.1) также позволяет представить поле плоского волновода в виде ТЕ- и ТМ-мод.

¹⁾ Это не совсем так. Потери на излучение, например в изогнутых волноводах [111*], различны для плоских и круглых волноводов. — *Прим. ред.*

Начнем с изучения свойств ТЕ-мод. Для них $E_z = 0$. Подставив (8.3.1) в формулы (1.4.16) — (1.4.19), получим, что только компоненты H_z , H_x и E_y не равны нулю. Используя уравнения Максвелла (1.4.13) и (1.4.15), выразим составляющие H_z и H_x через E_y :

$$H_x = -\frac{i}{\omega\mu} \frac{\partial E_y}{\partial z} \quad (8.3.2)$$

и

$$H_z = \frac{i}{\omega\mu} \frac{\partial E_y}{\partial x}. \quad (8.3.3)$$

Компонента E_y получается как решение приведенного волнового уравнения¹⁾

$$\frac{\partial^2 E_y}{\partial x^2} + \frac{\partial^2 E_y}{\partial z^2} + n^2 k_0^2 E_y = 0, \quad (8.3.4)$$

где

$$n^2 = \frac{\varepsilon}{\varepsilon_0} \quad (8.3.5)$$

и

$$k_0 = \omega \sqrt{\varepsilon_0 \mu_0} = \frac{2\pi}{\lambda_0}. \quad (8.3.6)$$

Для случая зависимости от координаты z и времени в виде

$$e^{i(\omega t - \beta z)} \quad (8.3.7)$$

из (8.3.4) получим

$$\frac{\partial^2 E_y}{\partial x^2} + (n^2 k_0^2 - \beta^2) E_y = 0. \quad (8.3.8)$$

Решение этого уравнения внутри слоя отличается от решения в окружающей среде (оболочке). Решение задачи можно упростить путем разделения мод на четные и нечетные. Естественно, четные и нечетные моды можно получить, используя основные выражения поля и строго решая задачу определения собственных значений. Однако априорное введение четных и нечетных мод существенно упрощает решение задачи (сравните получение ТЕ- и ТМ-мод в разд. 8.2).

¹⁾ Заметим, что уравнение (8.3.4) удовлетворяется даже на границе раздела при $x = \pm d$. В силу ограничения (8.3.1) член $\mathbf{E} \cdot \nabla \varepsilon$ в (1.3.4) равен нулю и остается только компонента E_y вектора \mathbf{E} .

ЧЕТНЫЕ НАПРАВЛЯЕМЫЕ ТЕ-МОДЫ

Внутри волновода, $|x| < d$, составляющие поля четных мод имеют вид ¹⁾

$$E_y = A_e \cos \kappa x \quad (8.3.9)$$

и

$$H_z = -\frac{i\kappa}{\omega\mu_0} A_e \sin \kappa x, \quad (8.3.10)$$

где

$$\kappa^2 = n_1^2 k_0^2 - \beta^2. \quad (8.3.11)$$

Составляющая H_x находится из (8.3.2). Поле вне волновода, $|x| > d$, имеет вид

$$E_y = A_e \cos \kappa d e^{-\gamma(|x|-d)} \quad (8.3.12)$$

и

$$H_z = \frac{x}{|x|} \frac{i\gamma}{\omega\mu_0} A_e \cos \kappa d e^{-\gamma(|x|-d)}, \quad (8.3.13)$$

где

$$\gamma^2 = \beta^2 - n_2^2 k_0^2. \quad (8.3.14)$$

Значения κ^2 и γ^2 положительны, так как $n_1 > n_2$. Для положительных значений γ поле вне волновода уменьшается при увеличении $|x|$. Условие существования моды, таким образом, можно записать в виде

$$\gamma > 0. \quad (8.3.15)$$

Амплитудная постоянная в (8.3.12) выбирается из условия непрерывности составляющей E_y при $x = \pm d$. Необходимо также потребовать непрерывности составляющей H_z на границе раздела сред. Таким образом, из формул (8.3.10) и (8.3.13) получаем уравнение собственных значений

$$\operatorname{tg} \kappa d = \frac{\gamma}{\kappa}. \quad (8.3.16)$$

Амплитудный коэффициент можно выразить через мощность P , переносимую модой. Из (1.2.12) с помощью

¹⁾ Множитель (8.3.7) опущен во всех составляющих поля.

формул (8.3.2) и (8.3.7) находим

$$\begin{aligned} P &= \frac{1}{2} \int_{-\infty}^{\infty} (\mathbf{E} \times \mathbf{H}^*)_z dx = -\frac{1}{2} \int_{-\infty}^{\infty} E_y H_x^* dx = \\ &= \frac{\beta}{\omega \mu_0} \int_0^{\infty} |E_y|^2 dx. \end{aligned} \quad (8.3.17)$$

Величина P выражает мощность, переносимую волной по волноводу в направлении оси z через единицу его ширины (по оси y). Подстановка выражений (8.3.9) и (8.3.12) в (8.3.17) приводит к соотношению

$$A_e = \left(\frac{2\omega \mu_0}{\beta d + \frac{\beta}{\gamma}} P \right)^{1/2}. \quad (8.3.18)$$

Чтобы получить A_e в таком простом виде, было использовано уравнение (8.3.16).

НЕЧЕТНЫЕ НАПРАВЛЯЕМЫЕ ТЕ-МОДЫ

Составляющие поля и уравнение собственных значений нечетных мод находятся так же, как в случае четных мод. Выражения для составляющих поля внутри волновода, $|x| < d$, имеют вид

$$E_y = A_0 \sin \kappa x \quad (8.3.19)$$

и

$$H_z = \frac{i\kappa}{\omega \mu_0} A_0 \cos \kappa x. \quad (8.3.20)$$

Вне волновода, $|x| > d$, имеем

$$E_y = \frac{x}{|x|} A_0 \sin \kappa d e^{-\gamma(|x|-d)} \quad (8.3.21)$$

и

$$H_z = \frac{-i\gamma}{\omega \mu_0} A_0 \sin \kappa d e^{-\gamma(|x|-d)}. \quad (8.3.22)$$

Постоянные κ и γ определяются формулами (8.3.11) и (8.3.14). Составляющая E_y опять является непрерывной на границе раздела в силу выбора амплитудных коэффициентов. Требование непрерывности для составляющей

H_z приводит к уравнению собственных значений

$$\operatorname{tg} \kappa d = -\frac{\kappa}{\gamma}. \quad (8.3.23)$$

Амплитудный коэффициент может быть выражен через переносимую мощность с помощью формул (8.3.17), (8.3.19), (8.3.21) и (8.3.23). В результате получается

$$A_0 = \left(\frac{2\omega\mu_0}{\beta d + \frac{\beta}{\gamma}} P \right)^{1/2}. \quad (8.3.24)$$

Соотношения (8.3.18) и (8.3.24) внешне одинаковы. Однако постоянные β и γ , входящие в них, различны: в одном случае они являются решениями уравнения собственных значений для четных мод, в другом — для нечетных мод.

СОБСТВЕННЫЕ ЗНАЧЕНИЯ ТЕ-МОД

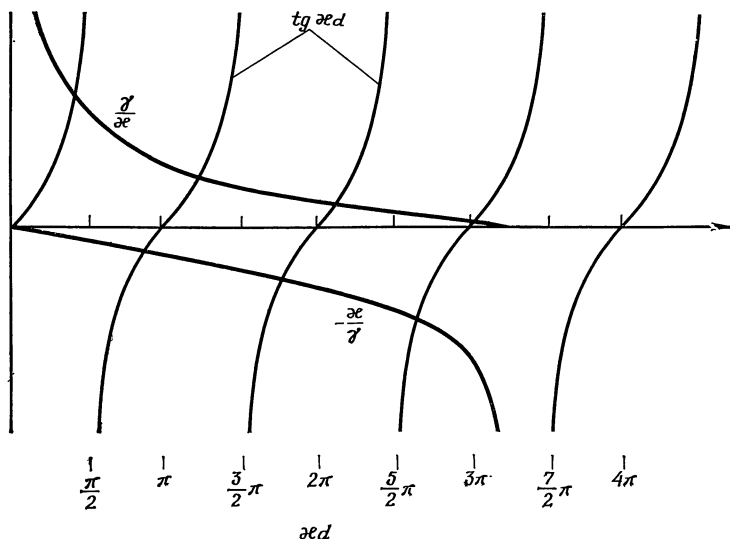
Уравнения собственных значений для четных и нечетных ТЕ-мод плоского диэлектрического волновода намного проще соответствующих уравнений для круглого оптического волокна. Решения уравнений (8.3.16) и (8.3.23) можно представить наглядно, для этого необходимо построить графики правых и левых частей этих уравнений. На фиг. 8.3.1 представлены эти графики. Решения уравнений находятся на пересечениях кривых. Из графиков видно, что при данной частоте существует ограниченное число решений. Величина γ/κ получается из формул (8.3.11) и (8.3.14)

$$\frac{\gamma}{\kappa} = \frac{\sqrt{(n_1^2 - n_2^2) k_0^2 - \kappa^2}}{\kappa}. \quad (8.3.25)$$

Фиксируя d и изменяя частоту, получим точку, в которой кривая γ/κ касается оси κd , а кривая $-\kappa/\gamma$ уходит в минус бесконечность. Кривые γ/κ и $-\kappa/\gamma$ пересекают во многих точках графики $\operatorname{tg} \kappa d$. Это показывает, что в плоском волноводе может распространяться много направляемых мод. Из фигуры также видно, что четная ТЕ-мода низшего порядка может распространяться при сколь угодно малой частоте. Это единственная мода волновода, которая не имеет отсечки. Все ТЕ- и ТМ-моды оптического волокна имеют частоту отсечки, не равную нулю. Только мода

HE_{11} не имеет отсечки. В целях сравнения необходимо предположить, что четная ТЕ-мода низшего порядка плоского диэлектрического волновода соответствует моде HE_{11} круглого оптического волокна.

Условие отсечки для четных ТЕ-мод получается из требования, чтобы величина γ/κ обращалась в нуль на оси κd в том месте, где тангенс равняется нулю. Таким



Ф и г. 8.3.1. Графическое решение уравнений собственных значений (8.3.16) и (8.3.23).

образом, необходимо, чтобы $\kappa_c d = v\pi$ (v — целое) в точке, где $\gamma_c d = 0$. Тогда из формулы (8.3.25) получим условие отсечки

$$\sqrt{n_1^2 - n_2^2} k_0 d = v\pi = \kappa_c d. \quad (8.3.26)$$

При отсечке $\beta = n_2 k_0$, а $\gamma = 0$ и

$$\frac{\beta_c}{\kappa_c} = \frac{n_2}{\sqrt{n_1^2 - n_2^2}}. \quad (8.3.27)$$

Заметим, что соотношение (8.3.27) справедливо как для четных ТЕ-мод, так и для четных ТМ-мод [см. ниже формулу (8.3.38)].

Механизм распространения моды в плоском волноводе с физической точки зрения можно объяснить следующим образом. Рассмотрим четную ТЕ-моду. Восстанавливая множитель (8.3.7), перепишем выражение (8.3.9) в виде

$$E_y = \frac{1}{2} A_e [e^{i(\omega t + \kappa x - \beta z)} + e^{i(\omega t - \kappa x - \beta z)}]. \quad (8.3.28)$$

Из этого выражения видно, что компонента E_y внутри диэлектрического волновода может быть представлена как суперпозиция двух плоских волн. Направление распространения этих плоских волн определяется соотношением

$$\operatorname{tg} \alpha = \pm \frac{\beta}{\kappa}. \quad (8.3.29)$$

Угол α образован направлением распространения плоской волны и нормалью к поверхности раздела волновода с внешней средой. Вследствие явления полного внутреннего отражения плоские волны не могут покидать среду, если их угол падения на поверхность раздела больше угла полного внутреннего отражения. В пределе, определяемом формулой (1.6.23), получим

$$\operatorname{tg} \alpha_i = \frac{\sin \alpha_i}{\sqrt{1 - \sin^2 \alpha_i}} = \frac{n_2}{\sqrt{n_1^2 - n_2^2}}. \quad (8.3.30)$$

Сравнивая формулы (8.3.27) и (8.3.30), видим, что правые части у них одинаковы. Выражение (8.3.30) характеризует предельный угол α_i полного внутреннего отражения, при котором плоские волны еще распространяются внутри диэлектрического волновода.

Существование направляемых мод физически можно объяснить следующим образом. Внутри диэлектрического волновода плоская волна распространяется под определенным углом к поверхности раздела сред волновода. От этой поверхности плоская волна полностью отражается и при дальнейшем своем распространении, последовательно отражаясь от двух (верхней и нижней) поверхностей раздела, колеблется между ними. Экспоненциально спадающее поле, существующее вне волновода, встречалось в разд. 1.6 как присущее волне, которая полностью отражается от поверхности раздела диэлектриков. Такое объяснение механизма

распространения моды в плоском диэлектрическом волноводе справедливо для других диэлектрических волноводов, и в частности для круглого оптического волокна. Из-за сложной геометрии круглого диэлектрического волновода механизм полного внутреннего отражения не является таким очевидным, как в плоском волноводе. Однако такие рассуждения не объясняют наличие отсечки у одних мод и отсутствие у других. Угол, под которым плоские волны распространяются внутри плоского волновода, определяется только путем полного решения задачи о распространении волн.

ЧЕТНЫЕ ТМ-МОДЫ

Кроме ТЕ-мод, в плоском диэлектрическом волноводе существуют также ТМ-моды, для которых $H_z = 0$. Составляющие поля ТМ-мод можно найти из формул (1.4.16) — (1.4.19) и (8.3.1). Получим E_z , E_x и H_y . Обе электрические компоненты выразим через компоненту H_y с помощью формул (1.4.10) и (1.4.12)

$$E_x = \frac{i}{n^2 \omega \epsilon_0} \frac{\partial H_y}{\partial z} \quad (8.3.31)$$

и

$$E_z = -\frac{i}{n^2 \omega \epsilon_0} \frac{\partial H_y}{\partial x}. \quad (8.3.32)$$

Составляющая H_y получается как решение приведенного волнового уравнения¹⁾

$$\frac{\partial^2 H_y}{\partial x^2} + (n^2 k_0^2 - \beta^2) H_y = 0. \quad (8.3.33)$$

Составляющие поля четных ТМ-мод внутри волновода, $|x| < d$, имеют вид

$$H_y = B_e \cos \kappa x \quad (8.3.34)$$

и

$$E_z = \frac{i\kappa}{n^2 \omega \epsilon_0} B_e \sin \kappa x, \quad (8.3.35)$$

¹⁾ Заметим, что уравнение (8.3.33) не удовлетворяется на границе раздела диэлектриков при $x = \pm d$ [сравните с формулой (9.5.2)]. Поверхности раздела диэлектриков $x = \pm d$ учитываются граничными условиями.

где κ определяется из (8.3.11). Составляющие поля вне волновода, $|x| > d$, имеют вид

$$H_y = B_e \cos \kappa x e^{-\gamma(|x|-d)} \quad (8.3.36)$$

и

$$E_z = \frac{x}{|x|} \frac{i\gamma}{n_2^2 \omega \epsilon_0} B_e \cos \kappa x e^{-\gamma(|x|-d)}. \quad (8.3.37)$$

Уравнение собственных значений находится из условия непрерывности составляющей E_z при $x = \pm d$

$$\operatorname{tg} \kappa d = \frac{n_1^2}{n_2^2} \frac{\gamma}{\kappa}. \quad (8.3.38)$$

Мощность, проходящая по волноводу через единицу его ширины (по оси y), получается из формулы (1.2.12)

$$\begin{aligned} P &= \frac{1}{2} \int_{-\infty}^{\infty} (\mathbf{E} \times \mathbf{H}^*)_z dx = \frac{1}{2} \int_{-\infty}^{\infty} E_x H_y^* dx = \\ &= \frac{\beta}{\omega \epsilon_0} \int_0^{\infty} \frac{1}{n^2} |H_y|^2 dx. \end{aligned} \quad (8.3.39)$$

Тогда амплитудный коэффициент имеет вид

$$B_e = \left\{ \frac{2\omega \epsilon_0 n_1^2 P}{\beta \left[d + \frac{(n_1 n_2)^2}{\gamma} \frac{\kappa^2 + \gamma^2}{n_2^2 \kappa^2 + n_1^4 \gamma^2} \right]} \right\}^{1/2}. \quad (8.3.40)$$

НЕЧЕТНЫЕ ТМ-МОДЫ

Составляющие поля нечетных ТМ-мод внутри волновода имеют вид

$$H_y = B_0 \sin \kappa x \quad (8.3.41)$$

и

$$E_z = -\frac{i\kappa}{n_1^2 \omega \epsilon_0} B_0 \cos \kappa x. \quad (8.3.42)$$

Составляющие поля для $|x| > d$ имеют вид

$$H_y = \frac{x}{|x|} B_0 \sin \kappa x e^{-\gamma(|x|-d)} \quad (8.3.43)$$

и

$$E_z = \frac{i\gamma}{n_2^2 \omega \epsilon_0} B_0 \sin \kappa x e^{-\gamma(|x|-d)}. \quad (8.3.44)$$

Уравнение собственных значений получается обычным способом с учетом граничных условий

$$\operatorname{tg} \kappa d = -\frac{n_2^2}{n_1^2} \frac{\kappa}{\gamma}. \quad (8.3.45)$$

Коэффициент B_0 при этом может быть представлен в виде

$$B_0 = \left\{ \frac{2\omega\epsilon_0 n_1^2 P}{\beta \left[d + \frac{(n_1 n_2)^2}{\gamma} \frac{\kappa^2 + \gamma^2}{n_2^2 \kappa^2 + n_1^4 \gamma^2} \right]} \right\}^{1/2}. \quad (8.3.46)$$

Опять оказалось, что выражения для амплитудных коэффициентов четных и нечетных ТМ-мод аналогичны.

При условии (8.3.1) ТЕ- и ТМ-моды обеспечивают полный набор направляемых типов волн плоского диэлектрического волновода. Однако условие (8.3.1) является дополнительным ограничением в постановке задачи. Если его не вводить, считая, что поля изменяются в направлении оси y , можно было бы получить намного больше мод. Эти дополнительные моды являются гибридными, они имеют все шесть компонент поля. Но для упрощения описания потерь на излучение и явления преобразования мод мы не будем изучать более сложные гибридные моды прямоугольного диэлектрического волновода¹⁾. Наложение условия (8.3.1) ограничивает задачу до такой степени, что ТЕ- и ТМ-моды полностью описывают волны, направляемые плоским волноводом.

8.4. МОДЫ ИЗЛУЧЕНИЯ ПЛОСКОГО ВОЛНОВОДА

Система направляемых мод, найденная выше, достаточна для описания распределения любого направляемого поля плоского волновода [при условии (8.3.1)]. Но эта система оказывается недостаточно полной для описания явления излучения из волновода. С помощью только этих волн нельзя учесть поле излучения. Поэтому систему направляемых волн диэлектрического волновода, имеющую дискретный спектр, необходимо дополнить системой волн с непрерывным спектром (модами излучения). Ана-

¹⁾ О модах прямоугольного диэлектрического волновода см. в работах [82*, 90*].— *Прим. ред.*

логичная ситуация имеет место в квантовой механике для атома водорода. Атом водорода имеет бесконечное число связанных состояний. Кроме них, энергетический спектр атома водорода содержит континуум несвязанных состояний, описывающих явление атомного рассеяния. Моды излучения открытых волноводов подробно рассмотрены в книге Шевченко [113].

Существование мод излучения диэлектрического волновода можно проиллюстрировать следующим образом. Пусть плоская волна падает извне на плоский диэлектрический волновод. Плоская волна испытывает при этом отражение и преломление таким образом, что возникает стоячая волна внутри и вне волновода. Поле излучения при удалении от диэлектрического слоя в этом случае не уменьшается. Оно не связано с волноводом и поэтому не убывает во всем пространстве. Однако поле излучения также является решением задачи, так как оно удовлетворяет уравнениям Максвелла и граничным условиям на поверхности волновода. Можно возбуждать симметричные и несимметричные моды излучения, используя два источника с разных сторон от слоя. То, что волны излучения не исчезают при удалении от волновода в бесконечность и, следовательно, переносят неограниченную мощность, не приводит к физически противоречивым результатам. Дело в том, что никакой реальный источник поля не возбудит только одну моду излучения. Всегда возбуждается совокупность этих мод. Поскольку спектр непрерывный, то это означает, что всегда поле выражается в виде интеграла от этих волн, а интеграл, выражающий собой вдали от источника поле излучения, на бесконечности спадает, а точнее, удовлетворяет условию излучения. Используя эти вводные замечания, рассмотрим моды излучения математически.

ЧЕТНЫЕ ТЕ-МОДЫ ИЗЛУЧЕНИЯ

Число составляющих поля ТЕ-мод излучения (мод непрерывного спектра) такое же, как у направляемых мод (мод дискретного спектра) этого типа. Составляющая E_y должна удовлетворять уравнению (8.3.8), а составляющие магнитного поля определяются из формул (8.3.2) и (8.3.3). Единственное различие между направляемыми

модами и модами излучения состоит в том, что для последних не требуется экспоненциального спада поля при удалении от волновода. В то же время при их изучении недостаточно лишь положить в формуле (8.3.14) величину γ мнимой. При мнимых значениях γ волны распространяются вне волновода. Но для удовлетворения граничным условиям нужно включить в рассмотрение стоячие волны, необходимость которых математически объясняется следующим образом. Если бы в формулировке задачи о модах излучения, рассмотренной в предыдущем разделе, положить величину γ мнимой, то пришлось бы решать однородное уравнение, которое является следствием условия непрерывности составляющей H_z при $x = \pm d$. В итоге опять получилось бы уравнение собственных значений (8.3.16). Однако оказывается, что не существует решений уравнения (8.3.16) при действительных β и мнимых γ . Чтобы получить решения для составляющих поля, отличные от найденных в предыдущем разделе, необходимо удовлетворить граничным условиям без введения ограничения на возможные значения β . Последнее означает, что нужно попытаться получить неоднородное уравнение для амплитудных коэффициентов. Его можно получить только путем добавления еще одной волны к полю вне волновода. Полагая величину γ в формуле (8.3.12) мнимой, получаем волну, распространяющуюся при $|x| > d$. Добавление к ней другой волны, распространяющейся в противоположном направлении, дает дополнительный амплитудный коэффициент и приводит, таким образом, краевую задачу к неоднородному уравнению¹⁾.

Снова запишем составляющие поля в виде²⁾

$$E_y = C_e \cos \sigma x, \quad (8.4.1)$$

$$H_z = -\frac{i\sigma}{\omega\mu_0} C_e \sin \sigma x, \quad (8.4.2)$$

1) В [113] дан более простой и более корректный метод построения волн непрерывного спектра (мод излучения) с помощью условия ограниченности поля волн в поперечной бесконечности. Это условие однозначно определяет как функции поля самих волн, так и их спектр: поперечные волновые числа ρ могут быть лишь вещественными величинами, а не произвольными, как ниже неверно утверждает автор. При этом и числа β не являются произвольными. — *Прим. ред.*

2) Множитель (8.3.7) опущен.

где

$$\sigma^2 = n_1^2 k_0^2 - \beta^2 \quad (8.4.3)$$

внутри слоя, а вне слоя при $|x| > d$

$$E_y = D_e e^{-i\rho|x|} + F_e e^{i\rho|x|}, \quad (8.4.4)$$

$$H_z = \frac{x}{|x|} \frac{\rho}{\omega \mu_0} (D_e e^{-i\rho|x|} - F_e e^{i\rho|x|}), \quad (8.4.5)$$

где

$$\rho^2 = n_2^2 k_0^2 - \beta^2. \quad (8.4.6)$$

Сравнение с формулами (8.3.9) — (8.3.14) даст дополнительный неопределенный амплитудный коэффициент в выражениях поля. Граничные условия, требующие непрерывности составляющих E_y и H_z при $x = \pm d$, приводят к следующей системе уравнений (можно рассматривать только $x = d$, так как граничные условия при $x = -d$ идентичны):

$$D_e e^{-i\rho d} + F_e e^{i\rho d} = C_e \cos \sigma d, \quad (8.4.7)$$

$$D_e e^{-i\rho d} - F_e e^{i\rho d} = -i \frac{\sigma}{\rho} C_e \cos \sigma d. \quad (8.4.8)$$

Поскольку имеются только два уравнения для отыскания трех неопределенных констант, необходимо рассматривать одну из них, например C_e , как заданную и решать полученную неоднородную систему уравнений. Неоднородная система уравнений имеет решения, только когда ее детерминант не равен нулю. В этом случае мы не получаем уравнения собственных значений и постоянная распространения β останется произвольной. Коэффициенты D_e и F_e можно получить из формул (8.4.7) и (8.4.8)

$$D_e = \frac{C_e}{2} e^{i\rho d} \left(\cos \sigma d - i \frac{\sigma}{\rho} \sin \sigma d \right) = \bar{D}_e C_e, \quad (8.4.9)$$

$$F_e = D_e^*. \quad (8.4.10)$$

Выразим амплитудный коэффициент C_e через мощность, переносимую модой излучения. Ранее отмечалось, что мощность каждой моды излучения не ограничена (беско-

нечна). Тогда вместо (8.3.17) можно потребовать, чтобы

$$P\delta(\rho - \rho') = \frac{\beta}{\omega\mu_0} \int_0^\infty E_y(\rho) E_y^*(\rho') dx. \quad (8.4.11)$$

Это соотношение является условием ортогональности двух различных мод с параметрами ρ и ρ' . Для разных значений этих параметров интеграл справа в (8.4.11) равен нулю, а для $\rho = \rho'$ принимает бесконечное значение. Вычислять интеграл необходимо с особым вниманием, так как с его помощью определяется δ -функция. Проведем это подробно. Из формул (8.4.1), (8.4.4), (8.4.9) и (8.4.10) получаем

$$I = \int_0^\infty E_y(\rho) E_y^*(\rho') dx = C_e C_e^{*'} \left\{ \int_0^d \cos \sigma x \cos \sigma' x dx + \right. \\ \left. + \int_d^\infty (\bar{D}_e e^{-i\rho x} + \bar{D}_e^* e^{i\rho x}) (\bar{D}_e^{*'} e^{i\rho' x} + \bar{D}_e' e^{-i\rho' x}) dx \right\}. \quad (8.4.12)$$

Интегрирование не представляет особых трудностей. Символ ∞ в данном случае характеризует очень большое число, стремящееся к бесконечности:

$$I = C_e C_e^{*'} \left\{ \frac{\sigma \sin \sigma d \cos \sigma' d - \sigma' \cos \sigma d \sin \sigma' d}{\sigma^2 - \sigma'^2} - \right. \\ - \frac{\bar{D}_e \bar{D}_e^{*'}}{i(\rho - \rho')} (e^{-i(\rho - \rho')\infty} - e^{-i(\rho - \rho')d}) - \\ - \frac{\bar{D}_e \bar{D}_e'}{i(\rho - \rho')} (e^{-i(\rho + \rho')\infty} - e^{-i(\rho + \rho')d}) + \\ + \frac{\bar{D}_e^* \bar{D}_e^{*'}}{i(\rho - \rho')} (e^{i(\rho + \rho')\infty} - e^{i(\rho + \rho')d}) + \\ \left. + \frac{\bar{D}_e^* \bar{D}_e'}{i(\rho - \rho')} (e^{i(\rho - \rho')\infty} - e^{i(\rho - \rho')d}) \right\}. \quad (8.4.13)$$

Используя формулу (8.4.9) и соотношение

$$\sigma^2 - \sigma'^2 = \rho^2 - \rho'^2 = \beta'^2 - \beta^2, \quad (8.4.14)$$

получаем

$$\begin{aligned}
 I = & \frac{1}{2} C_e C_e^* \left[\left(\cos \sigma d \cos \sigma' d + \right. \right. \\
 & \left. \left. + \frac{\sigma \sigma'}{\rho \rho'} \sin \sigma d \sin \sigma' d \right) \frac{\sin(\rho - \rho')(\infty - d)}{\rho - \rho'} + \right. \\
 & + \left(\frac{\sigma}{\rho} \sin \sigma d \cos \sigma' d - \frac{\sigma'}{\rho'} \cos \sigma d \sin \sigma' d \right) \frac{\cos(\rho - \rho')(\infty - d)}{\rho - \rho'} + \\
 & + \left(\cos \sigma d \cos \sigma' d - \frac{\sigma}{\rho} \frac{\sigma'}{\rho'} \sin \sigma d \sin \sigma' d \right) \frac{\sin(\rho + \rho')(\infty - d)}{\rho + \rho'} + \\
 & \left. + \left(\frac{\sigma}{\rho} \sin \sigma d \cos \sigma' d + \frac{\sigma'}{\rho'} \cos \sigma d \sin \sigma' d \right) \frac{\cos(\rho + \rho')(\infty - d)}{\rho + \rho'} \right]. \quad (8.4.15)
 \end{aligned}$$

Следует заметить, что это выражение имеет смысл только тогда, когда используется как подынтегральная функция некоторого интеграла. Дельта-функция является сингулярной и, строго говоря, определена только в том случае, когда она входит как множитель в подынтегральное выражение. Выражение (8.4.15) может быть использовано подобным образом только в качестве множителя выражения, которое в дальнейшем будет интегрироваться. Очевидно, что множители при \cos в (8.4.15) остаются ограниченными для всех значений ρ и ρ' . Так как аргумент \cos стремится к бесконечности, то этот член с \cos никогда не сможет дать вклад, если окажется под интегралом. Таким образом, члены, содержащие \cos , не дают вклада и могут быть опущены. Члены, содержащие \sin , выражаются через δ -функцию следующим образом:

$$\delta(x) = \lim_{K \rightarrow \infty} \frac{\sin Kx}{\pi x}. \quad (8.4.16)$$

Значения ρ и ρ' положительные. Член с $\sin(\rho + \rho') \infty$ также не дает вклада. Для действительного C_e получаем

$$I = \frac{\pi}{2} C_e^2 \left(\cos^2 \sigma d + \frac{\sigma^2}{\rho^2} \sin^2 \sigma d \right) \delta(\rho - \rho'). \quad (8.4.17)$$

Объединение выражений (8.4.11) и (8.4.17) позволяет выразить амплитудный коэффициент C_e через мощность P :

$$C_e = \left[\frac{2\rho^2 \omega \mu_0 P}{\pi \beta (\rho^2 \cos^2 \sigma d + \sigma^2 \sin^2 \sigma d)} \right]^{1/2}. \quad (8.4.18)$$

Таким образом, получены четные ТЕ-моды излучения и доказано условие ортогональности для них. Не повторяя аналогичных расчетов для других мод излучения, запишем конечные результаты.

НЕЧЕТНЫЕ ТЕ-МОДЫ ИЗЛУЧЕНИЯ

Составляющие поля для $|x| < d$ имеют вид

$$E_y = C_0 \sin \sigma x, \quad (8.4.19)$$

$$H_z = \frac{i\sigma}{\omega\mu_0} C_0 \cos \sigma x. \quad (8.4.20)$$

Вне волновода, $|x| > d$, имеем

$$E_y = \frac{x}{|x|} (D_0 e^{-i\rho|x|} - F_0 e^{i\rho|x|}), \quad (8.4.21)$$

$$H_z = \frac{\rho}{\omega\mu_0} (D_0 e^{-i\rho|x|} - F_0 e^{i\rho|x|}). \quad (8.4.22)$$

Соотношения между постоянными C_0 , D_0 и F_0 имеют вид

$$D_0 = \frac{1}{2} C_0 e^{i\rho d} \left(\sin \sigma d + i \frac{\sigma}{\rho} \cos \sigma d \right), \quad (8.4.23)$$

$$F_0 = D_0^*. \quad (8.4.24)$$

Амплитудный коэффициент связан с мощностью, переносимой модой. (Выражение «мощность, переносимая модой» использовано здесь не совсем точно, так как P — это фактически мощность, которая получается в результате интегрирования мощности поля внутри бесконечно малого интервала по ρ).

$$C_0 = \left[\frac{2\rho^2 \omega \mu_0 P}{\pi \beta (\rho^2 \sin^2 \sigma d + \sigma^2 \cos^2 \sigma d)} \right]^{1/2}. \quad (8.4.25)$$

ЧЕТНЫЕ ТМ-МОДЫ ИЗЛУЧЕНИЯ

Составляющие поля для $|x| < d$ имеют вид

$$H_y = C_e \cos \sigma x, \quad (8.4.26)$$

$$E_z = \frac{i\sigma}{n_1^2 \omega \epsilon_0} C_e \sin \sigma x. \quad (8.4.27)$$

Составляющие поля вне волновода, $|x| > d$, имеют вид

$$H_y = D_e e^{-i\rho|x|} + F_e e^{i\rho|x|}, \quad (8.4.28)$$

$$E_z = -\frac{x}{|x|} \frac{\rho}{n_2^2 \omega \epsilon_0} (D_e e^{-i\rho|x|} - F_e e^{i\rho|x|}). \quad (8.4.29)$$

Постоянные связаны соотношениями

$$F_e = D_e^* \quad (8.4.30)$$

и

$$D_e = \frac{1}{2} C_e e^{i\rho d} \left[\cos \sigma d - i \left(\frac{n_2}{n_1} \right)^2 \frac{\sigma}{\rho} \sin \sigma d \right], \quad (8.4.31)$$

а C_e выражается через мощность как

$$C_e = \left[\frac{2n_1^2 \rho^2 \omega \varepsilon_0 P}{\beta \pi \left(\frac{n_1^2}{n_2^2} \rho^2 \cos^2 \sigma d + \frac{n_2^2}{n_1^2} \sigma^2 \sin^2 \sigma d \right)} \right]^{1/2}. \quad (8.4.32)$$

НЕЧЕТНЫЕ ТМ-МОДЫ ИЗЛУЧЕНИЯ

Составляющие поля внутри волновода, $|x| < d$, имеют вид

$$H_y = C_0 \sin \sigma x, \quad (8.4.33)$$

$$E_z = -\frac{i\sigma}{n_1^2 \omega \varepsilon_0} C_0 \cos \sigma x. \quad (8.4.34)$$

Для $|x| > d$ имеем

$$H_y = \frac{x}{|x|} (D_0 e^{-i\rho|x|} + F_0 e^{i\rho|x|}), \quad (8.4.35)$$

$$E_z = -\frac{\rho}{n_2^2 \omega \varepsilon_0} [D_0 e^{-i\rho|x|} - F_0 e^{i\rho|x|}], \quad (8.4.36)$$

где

$$F_0 = D_0^*, \quad (8.4.37)$$

$$D_0 = \frac{1}{2} C_0 e^{i\rho d} \left[\sin \sigma d + i \frac{n_2^2}{n_1^2} \frac{\sigma}{\rho} \cos \sigma d \right] \quad (8.4.38)$$

и

$$C_0 = \left[\frac{2n_1^2 \rho^2 \omega \varepsilon_0 P}{\pi \beta \left(\frac{n_1^2}{n_2^2} \sigma^2 \sin^2 \sigma d + \frac{n_2^2}{n_1^2} \sigma^2 \cos^2 \sigma d \right)} \right]^{1/2}. \quad (8.4.39)$$

МОДЫ ИЗЛУЧЕНИЯ КРУГЛОГО ОПТИЧЕСКОГО ВОЛОКНА

Моды излучения плоского волновода достаточно просто описываются с помощью тригонометрических функций. В отличие от них моды оптического волокна выражаются намного сложнее через функции Бесселя. Для случая $\nu = 0$ (что соответствует $\partial/\partial \phi = 0$) эти моды рассмотрены в работе [99], а для $\nu = 1$ — в [96].

Моды излучения круглого оптического волокна могут быть выражены через функции Бесселя и Неймана. Имеется интересная математическая особенность, которую стоит обсудить. Мы уже видели, что граничные условия при $r = a$ дают четыре уравнения для определения амплитудных коэффициентов в выражениях составляющих направляемых мод. Выражения для мод излучения содержат шесть неопределенных коэффициентов: два в выражениях поля внутри стрежня и четыре в выражениях поля вне стрежня. Две дополнительные постоянные являются следствием того, что моды излучения содержат приходящие и уходящие вдоль радиуса волны. Один из шести коэффициентов можно связать с мощностью, переносимой модой. Однако второй коэффициент остается неопределенным¹⁾. Физическая причина этой математической загадки объясняется следующим образом. Подобной проблемы не было при изучении мод излучения плоского волновода в силу разделения их на четные и нечетные. В случае круглого волновода такое разделение невозможно. Однако имеются две независимые смешанные моды, которые необходимо рассмотреть отдельно. Каждая мода содержит неопределенный амплитудный коэффициент. Если распространяются две моды, то имеем два неопределенных коэффициента. Именно такая картина имеет место в случае оптического волокна. Но эти моды можно разделить посредством следующей процедуры. Если ввести произвольное линейное соотношение между этими двумя неопределенными коэффициентами, то получим одну моду излучения. Оставшийся неопределенный коэффициент обозначим через A_1 . Другая мода получается при использовании другого линейного соотношения между неопределенными коэффициентами. Оставшийся при этом неопределенный коэффициент обозначим через A_2 . Таким образом, имеем две произвольные моды 1 и 2. Чтобы полная система мод была ортогональной, моды 1 и 2 должны быть ортогональными. Выбор ортогональных векторов до некоторой степени произволен. Произвольность существует всегда, когда из двух линейно независимых и неортогональных векторов

¹⁾ В [113, заключение] дано условие, определяющее обсуждаемый неопределенный коэффициент (см. также [31*]). — *Прим. ред.*

образуют два ортогональных вектора. Два вектора можно объединить разными способами в два ортогональных вектора. Это обстоятельство можно использовать для упрощения соотношений между двумя дополнительными неопределенными коэффициентами. Результат процедуры ортогонализации мод излучения оптического волокна можно найти в [96]. Получающиеся уравнения достаточно сложные и неприменимы для простого объяснения физических процессов, которые приводят к излучению из диэлектрического волновода.

ОБЛАСТИ СУЩЕСТВОВАНИЯ НАПРАВЛЯЕМЫХ МОД И МОД ИЗЛУЧЕНИЯ

Итак, особенностью диэлектрических волноводов является наличие дискретного спектра направляемых мод и непрерывного спектра мод излучения. Необходимо исследовать диапазон возможных собственных значений (постоянных распространения) мод каждой части спектра. Для этого рассмотрим постоянные γ и ρ . Известно, что направляемые моды существуют для действительных значений γ . Используя соотношения (8.3.14) и (8.3.11) для действительных значений γ и κ , получим интервал возможных значений β для направляемых мод дискретного спектра

$$n_2 k_0 < |\beta_g| < n_1 k_0. \quad (8.4.40)$$

Моды излучения получаются из (8.4.6) для действительных значений ρ . Диапазон возможных значений β мод излучения состоит из двух частей. Для действительных значений β из формулы (8.4.6) получаем

$$0 \leq |\beta_r| < n_2 k_0. \quad (8.4.41)$$

Этот диапазон не перекрывается с диапазоном постоянных распространения направляемых мод. Оба диапазона стыкуются в точке $n_2 k_0$. Существует также еще один диапазон возможных значений β_r для действительных значений ρ . Этот диапазон задается мнимыми значениями β :

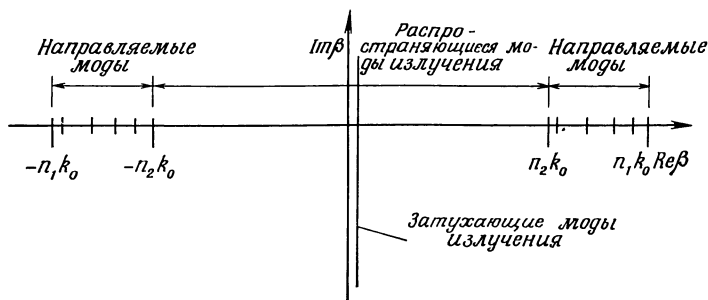
$$\beta_r = -i |\beta_r|, \quad (8.4.42a)$$

где

$$0 < |\beta_r| < \infty. \quad (8.4.42b)$$

Этот диапазон значений β соответствует затухающим модам. Эти моды упоминались в разд. 1.3 и 1.6. Оказалось, что затухающие моды в диэлектрическом волноводе также возможны, как и моды излучения.

Диапазоны возможных значений постоянных распространения направляемых мод и мод излучения, показанные на фиг. 8.4.1, одинаковы для круглого оптического волокна и плоского диэлектрического волновода.



Фиг. 8.4.1. Спектры постоянных распространения β направляемых мод и мод излучения диэлектрических волноводов.

Затухающие моды непрерывного спектра необходимы для описания поля вблизи волновода ¹⁾. Однако они не уносят мощность и поэтому не являются важными при изучении потерь на излучение направляемых мод диэлектрических волноводов.

8.5. СООТНОШЕНИЯ ОРТОГОНАЛЬНОСТИ

Моды плоского волновода, как и моды круглого оптического волокна, ортогональны друг к другу [97, 113]. Вместо доказательства ортогональности мод посредством прямых вычислений дадим общее доказательство, справедливое во всех системах цилиндрической конфигурации. «Цилиндрическая конфигурация» означает, что моды не изменяются (исключая фазу) вдоль оси системы.

¹⁾ Это не совсем правильно. Такие моды необходимы для описания локального поля вблизи источника. Они имеют место и при возбуждении волн в свободном пространстве при отсутствии волновода. — Прим. ред.

Используя зависимость от времени (8.3.7), запишем уравнения Максвелла (1.2.1) и (1.2.2) в виде

$$\nabla \times \mathbf{H} = i\omega \varepsilon \mathbf{E}, \quad (8.5.1)$$

$$\nabla \times \mathbf{E} = -i\omega \mu_0 \mathbf{H}. \quad (8.5.2)$$

Предположим, что электрические и магнитные поля относятся к индивидуальным модам структуры, которые обозначаются индексами ν и μ . Применим формулу (8.5.1) к ν -й моде и умножим ее комплексно-сопряженную величину на \mathbf{E}_μ . Формулу (8.5.2) применим к μ -й моде и умножим на \mathbf{H}_ν^* . Вычитание двух полученных уравнений и интегрирование приводят к соотношению

$$\begin{aligned} i\omega \int_{-\infty}^{\infty} (\varepsilon \mathbf{E}_\mu \cdot \mathbf{E}_\nu^* - \mu_0 \mathbf{H}_\mu \cdot \mathbf{H}_\nu^*) dx dy = \\ = - \int_{-\infty}^{\infty} [\mathbf{E}_\mu \cdot (\nabla \times \mathbf{H}_\nu^*) - \mathbf{H}_\nu^* (\nabla \times \mathbf{E}_\mu)] dx dy. \end{aligned} \quad (8.5.3)$$

Это же выражение можно записать в виде

$$\begin{aligned} i\omega \int_{-\infty}^{\infty} (\varepsilon \mathbf{E}_\mu \cdot \mathbf{E}_\nu^* - \mu_0 \mathbf{H}_\mu \cdot \mathbf{H}_\nu^*) dx dy = \\ = - \int_{-\infty}^{\infty} \nabla \cdot (\mathbf{H}_\nu^* \times \mathbf{E}_\mu) dx dy. \end{aligned} \quad (8.5.4)$$

Подставляя в (8.5.4) зависимость от координаты z (8.3.7), получим

$$\begin{aligned} i\omega \int_{-\infty}^{\infty} (\varepsilon \mathbf{E}_\mu \cdot \mathbf{E}_\nu^* - \mu_0 \mathbf{H}_\mu \cdot \mathbf{H}_\nu^*) dx dy = \\ = i(\beta_\mu - \beta_\nu) \int_{-\infty}^{\infty} (\mathbf{H}_\nu^* \times \mathbf{E}_\mu)_z dx dy - \\ - \int_{-\infty}^{\infty} \nabla_t \cdot (\mathbf{H}_\nu^* \cdot \mathbf{E}_\mu) dx dy. \end{aligned} \quad (8.5.5)$$

Оператор ∇_t является поперечной частью оператора ∇ . Интегрирование осуществляется по всему бесконечному

поперечному сечению в плоскости x, y . Ко второму интегралу в правой части (8.5.5) можно применить теорему

$$\int_{-\infty}^{\infty} \nabla_t \cdot (\mathbf{H}_v^* \times \mathbf{E}_\mu) dx dy = \int_C (\mathbf{H}_v^* \times \mathbf{E}_\mu) \cdot \mathbf{n} ds. \quad (8.5.6)$$

Интегрирование справа осуществляется по контуру бесконечно большого круга C , у которого \mathbf{n} — внешняя нормаль. Если одна или обе моды являются направляемыми, то их поле на бесконечности исчезает и интеграл в правой части (8.5.6) обращается в нуль. Однако интеграл обращается в нуль и в том случае, когда обе моды относятся к спектру излучения структуры. Последнее можно объяснить следующим образом. Моды излучения являются функциями координат x и y . Соотношения ортогональности для мод излучения полезны только в том случае, если они оказываются под интегралом, который берется по v или μ . Индексы мод излучения изменяются непрерывно. Для $v \neq \mu$ интегрирование по бесконечному интервалу v или μ в результате дает нуль для всех выражений, которые содержат быстро меняющиеся сомножители с бесконечно большими аргументами. Аналогичные соображения приводят к дельта-функции (8.4.16), равной нулю для $x \neq 0$. Таким образом, интеграл справа в (8.5.6) обращается в нуль для всех мод, для которых $v \neq \mu$. Получаем важное соотношение

$$\begin{aligned} (\beta_v - \beta_\mu) \int_{-\infty}^{\infty} (\mathbf{E}_\mu \times \mathbf{H}_v^*)_z dx dy = \\ = \omega \int_{-\infty}^{\infty} (\epsilon \mathbf{E}_\mu \cdot \mathbf{E}_v^* - \mu_0 \mathbf{H}_\mu \cdot \mathbf{H}_v^*) dx dy. \end{aligned} \quad (8.5.7)$$

Индекс z отмечает z -ю компоненту векторного произведения. Возьмем комплексно-сопряженную величину от (8.5.7), поменяем индексы v и μ и вычтем полученное уравнение из первоначального уравнения (8.5.7). В результате получим

$$(\beta_v - \beta_\mu) \int_{-\infty}^{\infty} [(\mathbf{E}_\mu \times \mathbf{H}_v^*)_z + (\mathbf{E}_v^* \times \mathbf{H}_\mu)_z] dx dy = 0. \quad (8.5.8)$$

При $\beta_\nu \neq \beta_\mu$

$$\int_{-\infty}^{\infty} [(\mathbf{E}_\mu \times \mathbf{H}_\nu^*)_z + (\mathbf{E}_\nu^* \times \mathbf{H}_\mu)_z] dx dy = 0. \quad (8.5.9)$$

Как видно из рассмотрения (1.4.16) — (1.4.23), уравнения Максвелла инвариантны относительно преобразований вида

$$\begin{aligned} \beta &\rightarrow -\beta, \\ E_z &\rightarrow -E_z, \\ H_z &\rightarrow +H_z, \\ E_x &\rightarrow +E_x, \\ E_y &\rightarrow +E_y, \\ H_x &\rightarrow -H_x, \\ H_y &\rightarrow -H_y. \end{aligned} \quad (8.5.10)$$

Преобразование (8.5.10) переводит любое данное решение уравнений Максвелла в другое решение этих уравнений. Соотношение (8.5.9) удовлетворяется для любых двух модовых решений уравнений Максвелла. Преобразуем μ -ю моду (но не ν -ю) согласно (8.5.10). Так как уравнение (8.5.9) удовлетворяется для любых двух модовых решений, то оно распространяется и на новую комбинацию моды ν и преобразованной моды μ . Таким образом, имеем

$$\int_{-\infty}^{\infty} [(\mathbf{E}_\mu \times \mathbf{H}_\nu^*)_z - (\mathbf{E}_\nu^* \times \mathbf{H}_\mu)_z] dx dy = 0. \quad (8.5.11)$$

Прибавив (8.5.11) к (8.5.9), получим

$$\int_{-\infty}^{\infty} (\mathbf{E}_\mu \times \mathbf{H}_\nu^*)_z dx dy = 0 \quad \text{для } \nu \neq \mu. \quad (8.5.12)$$

Заметим, что переход от (8.5.9) к (8.5.12) невозможен при $\beta_\mu = -\beta_\nu$. Выражение (8.5.12) представляет требуемое соотношение ортогональности между двумя модами в любом диэлектрическом волноводе без потерь. Если волновод с потерями, собственные значения β_ν и β_μ становятся комплексными и вывод соотношений ортогональности не может

быть проведен так, как здесь. Соотношение (8.5.12) справедливо для любых двух мод ν , μ , где ν -я и μ -я — направляемые моды, или одна из них направляемая, а другая — мода излучения, или обе являются модами излучения. Равенство (8.5.12) выполняется, когда правая часть равенства (8.5.7) обращается в нуль как при $\nu \neq \mu$, так и при $\nu = \mu$. Однако мода не ортогональна такой же моде, распространяющейся в противоположном направлении, т. е. при $\beta_\mu = -\beta_\nu$.

Выражение (8.5.12) можно непосредственно использовать для установления нескольких полезных соотношений для конкретного случая мод плоского волновода. Формулу (8.3.12) можно расширить таким образом, что для двух любых направляемых ТЕ-мод независимо от того, четные они, нечетные или смешанные, имеет место соотношение

$$\frac{\beta_\nu}{2\omega\mu_0} \int_{-\infty}^{\infty} E_{\nu y} E_{\mu y}^* dx = P\delta_{\nu\mu}, \quad (8.5.13)$$

где $\delta_{\nu\mu}$ — символ Кронекера, который равен нулю для $\nu \neq \mu$ и единице для $\nu = \mu$. Соотношение ортогональности для четных или нечетных ТЕ-мод излучения уже установлено в (8.4.11).

Для направляемых ТМ-мод в качестве расширения (8.3.39) получим

$$\frac{\beta_\nu}{2\omega\epsilon_0} \int_{-\infty}^{\infty} \frac{1}{n^2} H_{\nu y} H_{\mu y}^* dx = P\delta_{\nu\mu}. \quad (8.5.14)$$

Соотношение ортогональности для ТМ-мод излучения имеет вид

$$\frac{\beta}{2\omega\epsilon_0} \int_{-\infty}^{\infty} \frac{1}{n^2} H_y(\rho) H_y^*(\rho') dx = P\delta(\rho - \rho'). \quad (8.5.15)$$

Соотношения ортогональности (8.5.13) и (8.5.14) справедливы также, если одна из мод является модой излучения, а другая — направляемой модой. В этом случае правые части обоих выражений обращаются в нуль.

Любое произвольное распределение поля плоского волновода можно выразить в виде ортогональных мод

волновода. Можно, например, записать

$$E_y = \sum_{\nu} c_{\nu} E_{\nu y} + \sum \int_0^{\infty} q(\rho) E_y(\rho) d\rho. \quad (8.5.16)$$

Первая сумма распространяется на все четные и нечетные ТЕ-моды, а комбинация из суммы и интеграла распространяется на все моды излучения. Символ суммирования показывает, что в разложении поля должны использоваться четные и нечетные моды. Другие составляющие поля можно представить подобным образом.

Коэффициенты разложения легко получаются с помощью соотношений ортогональности. Для случая (8.5.16), используя формулу (8.5.13), получим

$$c_{\nu} = \frac{\beta_{\nu}}{2\omega\mu_0 P} \int_{-\infty}^{\infty} E_y E_{\nu y}^* dx. \quad (8.5.17)$$

Аналогично с помощью формулы (8.4.11) находим

$$q(\rho) = \frac{\beta}{2\omega\mu_0 P} \int_{-\infty}^{\infty} E_y E_y^*(\rho) dx. \quad (8.5.18)$$

В общем случае коэффициенты разложения c_{ν} и $q(\rho)$ являются функциями от z .

Для составляющей H_y получим разложение в виде ТМ-мод

$$H_y = \sum_{\nu} d_{\nu} H_{\nu y} + \sum \int_0^{\infty} p(\rho) H_y(\rho) d\rho, \quad (8.5.19)$$

где

$$d_{\nu} = \frac{\beta_{\nu}}{2\omega\epsilon_0 P} \int_{-\infty}^{\infty} \frac{1}{n^2} H_y H_{\nu y}^* dx \quad (8.5.20)$$

и

$$p(\rho) = \frac{\beta}{2\omega\epsilon_0 P} \int_{-\infty}^{\infty} \frac{1}{n^2} H_y H_y^*(\rho) dx. \quad (8.5.21)$$

Мощность P полагалась одинаковой для всех мод. Удобно использовать для P единицу мощности, например 1 Вт.

Фактический вклад мощности каждой моды в разложениях поля (8.5.16) и (8.5.19) определяется значениями коэффициентов разложения.

Для поля, составляющими которого являются H_z , H_x , E_y , полная переносимая мощность дается выражением

$$P_t = P \left\{ \sum_v |c_v|^2 + \sum_0 \int_0^\infty |q(\rho)|^2 d\rho \right\}. \quad (8.5.22)$$

Это соотношение получается из первой или второй части (8.3.17) с помощью (8.3.2) и соотношений ортогональности (8.5.13) и (8.4.11).

Разложение произвольных полей по ТЕ- или ТМ-модам возможно только для таких полей, которые удовлетворяют условию (8.3.1). Поля более общего вида можно разложить только по всем модам плоского волновода. Выражения (8.5.16) и (8.5.19) формально, конечно, правильные. Необходимо только в общем случае рассмотреть суммирование разложений по всем возможным модам волновода ¹⁾.

Возможность разложения произвольных полей по модам направляющей структуры чрезвычайно полезна для многих задач. Вопросы преобразования мод и потерь на излучение будем излагать, используя метод разложения по модам. Задача решения уравнений Максвелла для произвольных полей сводится при этом к системе обыкновенных дифференциальных уравнений для коэффициентов разложения c_v и $q(\rho)$ [113]. Точное решение системы этих уравнений достаточно сложно, однако они оказываются более удобными при решении методом возмущений ²⁾.

8.6. ПОЛЕЗНЫЕ ПРИБЛИЖЕНИЯ

Свойства направляемых мод определяются значениями их постоянных распространения. Постоянная распространения находится из трансцендентного уравнения

¹⁾ Если сказать точнее, то необходимо еще суммирование или интегрирование по модам, соответствующим вариации поля по второй поперечной координате. — *Прим. ред.*

²⁾ В общем случае методом последовательных приближений. — *Прим. ред.*

собственных значений, поэтому в явном виде она обычно не получается. Однако для случаев, имеющих практический интерес, можно найти полезные приближенные решения. В настоящем разделе получены некоторые приближенные решения уравнения собственных значений (8.2.70) круглого оптического волокна и уравнений собственных значений (8.3.16) и (8.3.23) ТЕ-мод плоского волновода. Соответствующие результаты для ТМ-мод аналогичны и приводятся здесь без вывода.

Начнем с рассмотрения плоского волновода. Постоянные распространения четных ТЕ-мод получаются из уравнения

$$\operatorname{tg} \kappa d = \frac{\gamma}{\kappa}, \quad (8.6.1)$$

а нечетных ТЕ-мод — из уравнения

$$\operatorname{tg} \kappa d = -\frac{\kappa}{\gamma}. \quad (8.6.2)$$

Вводя обозначение

$$V = \sqrt{n_1^2 - n_2^2} k_0 d, \quad (8.6.3)$$

в котором k_0 определяется формулой (8.3.6), получим из (8.3.11) и (8.3.14) следующее соотношение между постоянными κ и γ :

$$(\kappa d)^2 + (\gamma d)^2 = V^2. \quad (8.6.4)$$

При отсечке имеем

$$\gamma d = 0. \quad (8.6.5)$$

Из формулы (8.6.4) видно, что κd и V при отсечке одинаковы:

$$(\kappa d)_c = V_c = v \frac{\pi}{2}, \quad v = 0, 1, 2, 3, \dots \quad (8.6.6)$$

Последняя часть соотношения (8.6.6) получается с помощью фиг. 8.3.1, где представлены графически решения уравнений собственных значений (8.6.1) и (8.6.2). Четные значения v принадлежат четным ТЕ-модам, а нечетные — нечетным. Выражая κd в виде

$$\kappa d = v \frac{\pi}{2} + \eta, \quad (8.6.7)$$

где $\eta \ll 1$, для четных ТЕ-мод (ν — четное) из формулы (8.6.4) получим

$$\gamma d = \nu \frac{\pi}{2} \eta. \quad (8.6.8)$$

Из формул (8.6.4) и (8.6.6) находим, что для первого порядка по η

$$\eta = V - \nu \frac{\pi}{2}. \quad (8.6.9)$$

Таким образом, получаем приближенное решение (8.6.4) для четных ТЕ-мод плоского волновода при условиях, близких к отсечке,

$$\gamma d = \nu \frac{\pi}{2} \left(V - \nu \frac{\pi}{2} \right). \quad (8.6.10)$$

Аналогично, используя нечетные значения ν , опять получим соотношение (8.6.8) из уравнения собственных значений (8.6.2) для нечетных ТЕ-мод. Таким образом, вблизи отсечки решение (8.6.10) справедливо для четных и нечетных ТЕ-мод плоского волновода. Четные моды получаются для четных значений ν , а нечетные моды — для нечетных значений ν .

Единственным исключением из этого правила является четная ТЕ-мода низшего порядка, частота отсечки которой равна нулю. Для $V_c = 0$ наше рассмотрение неверно. Около отсечки для четной моды низшего порядка $\kappa d \ll 1$ и из формулы (8.6.1) получаем

$$\gamma d = (\kappa d)^2. \quad (8.6.11)$$

Если заменить κd с помощью формулы (8.6.4), то получим квадратное уравнение для γd с решением

$$\gamma d = \frac{1}{2} \left(\sqrt{4V^2 + 1} - 1 \right). \quad (8.6.12)$$

Постоянная распространения получается из (8.3.14). С помощью формулы (8.6.10) около отсечки для четных и нечетных ТЕ-мод, исключая моду низшего порядка, получаем приближенное выражение

$$\beta d = \sqrt{(n_2 k_0 d)^2 + \nu^2 \frac{\pi^2}{4} \left(V - \nu \frac{\pi}{2} \right)^2}. \quad (8.6.13)$$

Для четной ТЕ-моды низшего порядка из формул (8.3.14) и (8.6.12) находим

$$\beta d = \sqrt{(n_2 k_0 d)^2 + \frac{1}{2}(1 + 2V^2 - \sqrt{1 + 4V^2})}. \quad (8.6.14)$$

Параметр V определяется формулой (8.6.3). Можно также получить приближенные решения уравнений собственных значений вдали от отсечки. Для этого используем метод, который был успешно применен Снайдером [90, 92].

Рассмотрим κd как функцию от V и продифференцируем (8.6.1) по V

$$\frac{1}{\cos^2 \kappa d} \frac{\partial (\kappa d)}{\partial V} = -\frac{\gamma d}{(\kappa d)^2} \frac{\partial (\kappa d)}{\partial V} + \frac{V - (\kappa d) \frac{\partial (\kappa d)}{\partial V}}{(\kappa d)(\gamma d)}. \quad (8.6.15)$$

Для выражения γd через κd и V было использовано соотношение (8.6.4). Из формул (8.6.1) и (8.6.4) получим

$$\cos^2 \kappa d = \frac{(\kappa d)^2}{V^2} = \frac{\kappa^2}{(n_1^2 - n_2^2) k_0^2}. \quad (8.6.16)$$

Это позволяет представить равенство (8.6.15) в виде

$$\frac{\partial (\kappa d)}{\partial V} = \frac{\kappa d}{V(1 + \gamma d)}. \quad (8.6.17)$$

Вдали от отсечки $\gamma d \gg 1$ и $V \gg 1$, так что из (8.6.4) приближенно получим $\gamma d = V$. Уравнение (8.6.17) тогда примет вид

$$\frac{\partial (\kappa d)}{\partial V} = \frac{\kappa d}{V(1 + V)}. \quad (8.6.18)$$

При $V \rightarrow \infty$ из фиг. 8.3.1 видно, что в случае четных ТЕ-мод

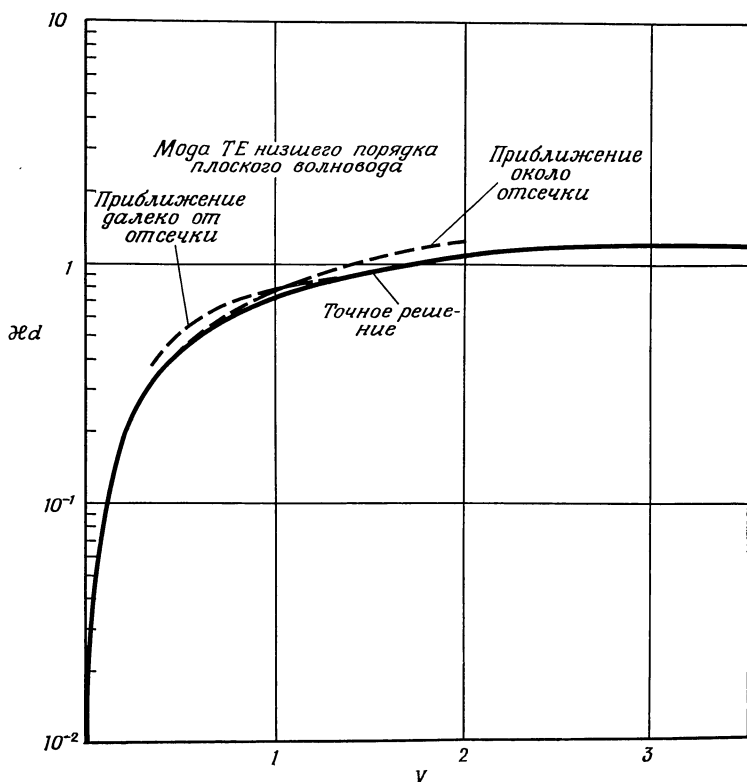
$$\kappa d = (v + 1) \frac{\pi}{2}, \quad (8.6.19)$$

где $v = 0, 2, 4, \dots$. Таким образом, находим, что решение дифференциального уравнения (8.6.18) имеет вид

$$\kappa d = (v + 1) \frac{\pi}{2} \frac{V}{1 + V}. \quad (8.6.20)$$

Выражение (8.6.20) определяет не только собственные значения уравнения (8.6.1) для четных ТЕ-мод плоского

волновода далеко от отсечки, но, как можно показать прямой подстановкой его в (8.6.2), оно также приближенно удовлетворяет уравнению (8.6.2) для нечетных ТЕ-мод при $V \gg 1$. Снова получаем значения для четных мод



Фиг. 8.6.1. Зависимость κd от нормированного частотного параметра V для ТЕ-моды низшего порядка плоского волновода (сплошная линия). Приближенному решению соответствуют пунктирные линии.

в случае четных v и для нечетных мод при нечетных v . Постоянная распространения β находится из формулы (8.3.11)

$$\beta = \sqrt{n_1^2 k_0^2 - \kappa^2} = \sqrt{n_2^2 k_0^2 + \gamma^2}. \quad (8.6.21)$$

На фиг. 8.6.1 сравниваются приближенные значения κd для четной ТЕ-моды низшего порядка, полученные из

формул (8.6.4) и (8.6.12) для малых V и из (8.6.20) с $v = 0$ для больших значений V . На этой фигуре приведено также точное решение уравнения (8.6.1). Как и следовало ожидать, приближенные значения κd совпадают с точными решениями в областях их применимости.

Представляем без вывода результаты соответствующих вычислений для четных и нечетных ТМ-мод плоского волновода, полученные из формул (8.3.38) и (8.3.45). Около отсечки имеем

$$\gamma d = \frac{n_2^2}{n_1^2} v \frac{\pi}{2} \left(V - v \frac{\pi}{2} \right) \quad (8.6.22)$$

с четными целыми значениями v ($v \neq 0$) для четных мод и с нечетными целыми значениями v для нечетных мод. Для четной моды низшего порядка

$$\gamma d = \frac{1}{2} \frac{n_1^2}{n_2^2} \left(\sqrt{4 \frac{n_2^4}{n_1^4} V^2 + 1} - 1 \right). \quad (8.6.23)$$

Далеко от отсечки

$$\kappa d = (v+1) \frac{\pi}{2} \frac{V}{\frac{n_1^2}{n_2^2} + V} \quad (8.6.24)$$

с четными v для четных мод и с нечетными v для нечетных мод плоского волновода.

Аналогичные приближения для γd около отсечки и для κd далеко от отсечки можно получить в случае круглого оптического волокна. Для этого удобно использовать уравнение собственных значений в виде (8.2.70). Около отсечки, где $\gamma d \ll 1$, уравнение (8.2.70) можно аппроксимировать с помощью (8.2.76), в котором H^- дается формулой (8.2.75) для $v = 2, 3, 4, \dots$. Используем приближение

$$\kappa a = V - \frac{(a\gamma)^2}{2V_c}, \quad (8.6.25)$$

где V_c — значение V при отсечке. Получаем $V_c = \kappa_c a$ из формулы (8.2.94) для ЕН-мод и из (8.2.95) для НЕ-мод. Используя разложения функций Бесселя с помощью (8.6.25) по членам вида $(a\gamma)^2/2V_c$, из формул (8.2.75) и (8.2.76) найдем около отсечки следующее приближение

для ЕН- и НЕ-мод круглого оптического волокна:

$$(\gamma a)^2 = \frac{V_c J_v(V) \left[\frac{V}{v-1} J_v(V) - (\varepsilon+1) J_{v-1}(V) \right]}{\left\{ \frac{v-2}{2(v-1)} (\varepsilon+1) J_v^2(V_c) + \left[\frac{2\varepsilon}{V_c} + \frac{V_c}{v-1} \left(1 + \frac{\varepsilon+1}{4v} \right) \right] \times \right.} \\ \left. \times J_v(V_c) J_{v-1}(V_c) - \frac{2\varepsilon+v(\varepsilon+1)}{2v} J_{v-1}^2(V_c) \right\}}, \quad (8.6.26)$$

в котором использовано обозначение

$$\varepsilon = \frac{n_1^2}{n_2^2}. \quad (8.6.27)$$

Соотношение (8.6.26) справедливо для ЕН- и НЕ-мод при $v \geq 2$. Заметим, что функция Бесселя в числителе зависит от V , а в знаменателе — от V_c . При значении V вблизи отсечки согласно (8.2.94) для ЕН-мод или (8.2.95) для НЕ-мод один из двух членов числителя является малым. Уравнение (8.6.26) несправедливо для $v = 0$. Важное приближение получается для $v = 1$ с использованием (8.2.74) в (8.2.75). Вблизи отсечки

$$\frac{\varepsilon+1}{\kappa a} \frac{J_0(\kappa a)}{J_1(\kappa a)} + 2 \ln \frac{\Gamma \gamma a}{2} = 0, \quad (8.6.28)$$

где $\Gamma = 1,781672$. Вблизи отсечки можно использовать приближение

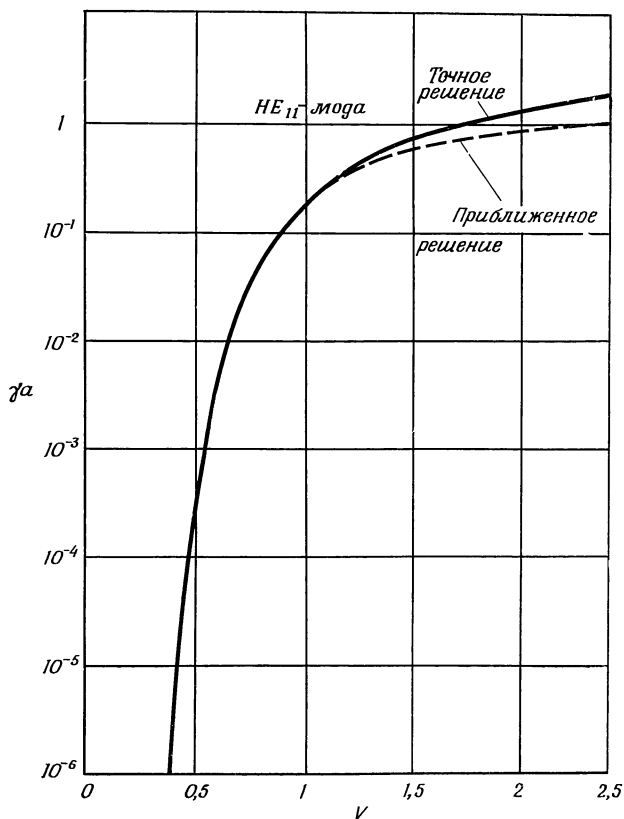
$$\kappa a \approx V. \quad (8.6.29)$$

Таким образом, для мод $ЕН_{1\mu}$ и $НЕ_{1\mu}$ получаем приближение

$$\gamma a = 1,122 \exp \left[-\frac{\varepsilon+1}{2V} \frac{J_0(V)}{J_1(V)} \right]. \quad (8.6.30)$$

Насколько хорошо это приближение, видно из фиг. 8.6.2 для моды $НЕ_{11}$ ($V_c = 0$) в случае, когда $\varepsilon - 1 \ll 1$. Даже для $V = 1$ приближение не отличается от точного решения. Из фиг. 8.6.2 видно, как очень быстро γa уменьшается с уменьшением V . Очевидно, что представление о том, что мода $НЕ_{11}$ направляется волноводом даже при нулевой частоте, в действительности чересчур академично. При значениях V ниже 0,5 мода едва ли может считаться направляемой, так как ее поле простирается вне волокна

- на расстояние, в 1000 раз превышающее радиус сердцевины волокна. Соотношение (8.6.30) можно также аппроксимировать, используя приближения функций Бесселя для



Ф и г. 8.6.2. Зависимость γa от V для моды HE_{11} круглого оптического волокна.

малых аргументов (8.2.82) и (8.2.83). В этом случае согласие с точным решением недостаточно хорошее. К сожалению, этот метод нельзя использовать в случае $v = 0$ для ТЕ- и ТМ-мод.

Используя аналогичные приближения, получим из уравнения (8.2.70) около отсечки

$$\text{для ТМ-мод} \quad (\gamma a)^2 = \frac{-V_c(V - V_c)}{\frac{1}{2} - \varepsilon \ln \frac{\Gamma \gamma a}{2}}. \quad (8.6.31)$$

Уравнение (8.6.31) применимо к ТЕ-модам, если положить $\varepsilon = 1$. Несмотря на то что это уравнение проще исходного уравнения собственных значений (8.2.70), оно требует применения приближенных методов для получения зависимости γa от V . Однако уравнение (8.6.31) облегчает вычисление $V - V_c$ как функции γa и графическое решение становится простым. Для ТЕ- и ТМ-мод при решении уравнения (8.6.31) V_c берется из (8.2.92).

Для ТЕ-мод плоского и круглого волноводов можно выразить κa и γa (или κd и γd) в виде функций от V и нет необходимости задавать значения $\varepsilon = n_1^2/n_2^2$. Для всех других мод необходимо знать ε и V , чтобы получить решения для κa и γa . Однако во многих интересных для практики случаях отношение n_1/n_2 близко к единице и хорошее приближение получается, если в уравнении (8.2.70) положить $\varepsilon = 1$. Это приближение особенно полезно для оптических волноводов в оболочке, показатель преломления сердцевины которых близок к показателю преломления оболочки, т. е. $(n_1/n_2) - 1 < 0,01$. Для важных случаев этого типа можно аппроксимировать (8.2.70), полагая $\varepsilon = 1$ и сразу же получая два уравнения собственных значений с помощью (8.2.56)–(8.2.59):

$$\text{для НЕ-мод} \quad \frac{J_{\nu-1}(\kappa a)}{\kappa a J_{\nu}(\kappa a)} = \frac{H_{\nu-1}^{(1)}(i\gamma a)}{i\gamma a H_{\nu}^{(1)}(i\gamma a)} \quad (8.6.32)$$

и

$$\text{для ЕН-мод} \quad \frac{J_{\nu+1}(\kappa a)}{\kappa a J_{\nu}(\kappa a)} = \frac{H_{\nu+1}^{(1)}(i\gamma a)}{i\gamma a H_{\nu}^{(1)}(i\gamma a)}. \quad (8.6.33)$$

С помощью формулы (8.2.64) условия отсечки (8.2.95) для случая $n_1 \approx n_2$ упрощаются:

$$\text{для НЕ-мод} \quad J_{\nu-2}(V_c) = 0, \quad (8.6.34)$$

а условие отсечки (8.2.94) остается неизменным:

$$\text{для ЕН-мод} \quad J_v(V_c) = 0. \quad (8.6.35)$$

Вдали от отсечки ($\gamma a \rightarrow \infty$) из формул (8.6.32) и (8.6.33) находим условия

$$\text{для НЕ-мод} \quad J_{v-1}(\kappa a) = 0 \quad \text{при} \quad V \rightarrow \infty \quad (8.6.36)$$

и

$$\text{для ЕН-мод} \quad J_{v+1}(\kappa a) = 0 \quad \text{при} \quad V \rightarrow \infty. \quad (8.6.37)$$

Приближенные уравнения собственных значений (8.6.32) и (8.6.33) намного удобнее для получения решения, чем исходное уравнение (8.2.70). Выше были получены приближения около отсечки (8.6.26), (8.6.30) и (8.6.31). Но можно еще получить очень полезные приближения для области вдали от отсечки. Снова используем метод, впервые примененный Снайдером [90, 92]. Начнем с НЕ-мод, предположив, что $V \gg 1$ и $\gamma a \gg 1$. Тогда с помощью формулы (8.2.54) можно записать

$$\frac{H_{v-1}^{(1)}(i\gamma a)}{H_v^{(1)}(i\gamma a)} = i \quad \text{для} \quad V \rightarrow \infty. \quad (8.6.38)$$

Теперь получим (8.6.32) в приближенном виде

$$\gamma a J_{v-1}(\kappa a) = \kappa a J_v(\kappa a). \quad (8.6.39)$$

Далее запишем производную (8.6.39) по V . После дифференцирования получаем приближенное равенство $\gamma a = V$ и исключаем функции Бесселя из уравнения с помощью равенства (8.6.39). Тогда находим

$$\frac{d(\kappa a)}{dV} = \frac{\kappa a}{V[V - 2(v-1)]}. \quad (8.6.40)$$

Решение этого дифференциального уравнения дает искомое приближение

$$\text{для НЕ-мод} \quad \kappa a = (\kappa a)_\infty \left[1 - \frac{2(v-1)}{V} \right]^{1/[2(v-1)]} \quad \text{при} \quad V \gg 1. \quad (8.6.41)$$

Значение $(\kappa a)_\infty$ является корнем уравнения (8.6.36). Корни функций Бесселя табулированы и представлены в [8, 11].

Решение для $v = 1$ осуществляется или путем предельного перехода при $v \rightarrow 1$ в (8.6.41), или путем интегриро-

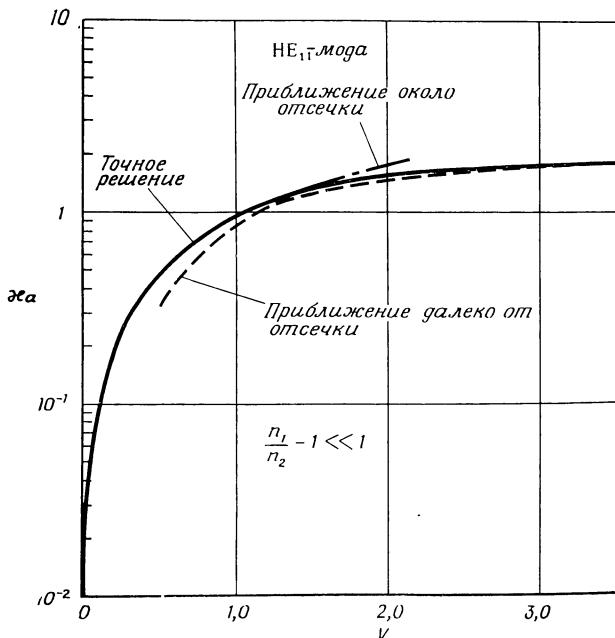
вания (8.6.40), что дает

$$\text{для HE}_{11}\text{-мод } \kappa a = (\kappa a)_{\infty} e^{-1/V} \text{ при } V \gg 1. \quad (8.6.42)$$

Аналогичная процедура позволяет получить приближенное решение уравнения (8.6.33):

$$\text{для EH-мод } \kappa a = (\kappa a)_{\infty} \left[1 - \frac{2(v+1)}{V} \right]^{1/[2(v+1)]} \text{ при } V \gg 1. \quad (8.6.43)$$

Это уравнение справедливо для всех положительных значений v . Постоянная $(\kappa a)_{\infty}$ получается из формулы (8.6.37) как корень функции Бесселя.



Фиг. 8.6.3. Зависимость κa от V для моды HE_{11} круглого оптического волокна. Пунктирной и штрих-пунктирной линиями представлены приближенные решения.

Фиг. 8.6.3 иллюстрирует точность приближения (8.6.42) для основной моды HE_{11} . Сплошной кривой представлено точное решение (8.6.32), а пунктирной —

приближенное решение для $V \gg 1$. Штрих-пунктирной кривой представлены значения κa , полученные из формулы (8.6.30) с помощью (8.6.4). Даже при $V = 1,75$, т. е. в той точке, где, согласно обоим приближениям, происходит наибольшее отклонение от точной кривой, отклонение не превышает 7%. Для многих целей оба приближения обеспечивают достаточную точность решений уравнения собственных значений.

Для области вдали от отсечки в случае малой разности показателей преломления $n_1 - n_2$ можно привести простые приближенные выражения для компонент электромагнитного поля мод круглого волновода. Эти выражения получены Снайдером [90, 92]. Предположим, что волновод работает в режиме, далеком от отсечки, т. е. можно считать $\beta = n_1 k$ и $\gamma a \gg 1$. Уравнения собственных значений (8.6.32) и (8.6.33) тогда примут вид

$$\text{для НЕ-мод} \quad \gamma a J_{\nu-1}(\kappa a) = \kappa a J_{\nu}(\kappa a), \quad (8.6.44)$$

$$\text{для ЕН-мод} \quad \gamma a J_{\nu+1}(\kappa a) = -\kappa a J_{\nu}(\kappa a). \quad (8.6.45)$$

При получении этих уравнений использовалось приближение (8.2.54). Используя функциональное соотношение [11, 61]

$$\begin{aligned} J'_{\nu}(\kappa a) &= -\frac{\nu}{\kappa a} J_{\nu}(\kappa a) + J_{\nu-1}(\kappa a) = \\ &= \frac{\nu}{\kappa a} J_{\nu}(\kappa a) - J_{\nu+1}(\kappa a) \end{aligned} \quad (8.6.46)$$

и приближения, упомянутые ранее, получим следующее приближение для (8.2.46):

$$B = \mp i \frac{|\nu|}{\nu} n \sqrt{\frac{\epsilon_0}{\mu_0}} A. \quad (8.6.47)$$

Верхний знак справедлив для НЕ-мод, а нижний — для ЕН-мод. Показатель преломления взят при $n_1 \approx n_2$. Снова, используя формулу (8.6.46), из (8.2.25) — (8.2.30) с помощью (8.6.47) получим для $r \leq a$ ¹⁾

$$E_z = A J_{\nu}(\kappa r) e^{i\nu\phi} = \frac{1}{2} F_z^{(1)} e^{i\nu\phi}, \quad (8.6.48)$$

$$H_z = \mp i \frac{|\nu|}{\nu} n \sqrt{\frac{\epsilon_0}{\mu_0}} A J_{\nu}(\kappa r) e^{i\nu\phi} =$$

¹⁾ Множитель $\exp[i(\omega t - \beta z)]$ опущен.

$$= \mp i \frac{|\nu|}{2\nu} G_{z|\nu|}^{(1)} e^{i\nu\phi}, \quad (8.6.49)$$

$$E_r = \mp i \frac{nk_0 A}{\kappa} J_{\nu \mp 1}(\kappa r) e^{i\nu\phi} = \mp \frac{i}{2} F_{|\nu| \mp 1}^{(1)} e^{i\nu\phi}, \quad (8.6.50)$$

$$E_\phi = \frac{|\nu|}{\nu} \frac{nk_0 A}{\kappa} J_{\nu \mp 1}(\kappa r) e^{i\nu\phi} = \frac{|\nu|}{2\nu} F_{|\nu| \mp 1}^{(1)} e^{i\nu\phi}, \quad (8.6.51)$$

$$\begin{aligned} H_r &= -\frac{|\nu|}{\nu} \sqrt{\frac{\varepsilon_0}{\mu_0} \frac{n^2 k_0 A}{\kappa}} J_{\nu \mp 1}(\kappa r) e^{i\nu\phi} = \\ &= -\frac{|\nu|}{2\nu} G_{|\nu| \mp 1}^{(1)} e^{i\nu\phi}, \end{aligned} \quad (8.6.52)$$

$$\begin{aligned} H_\phi &= \mp i \sqrt{\frac{\varepsilon_0}{\mu_0} \frac{n^2 k_0 A}{\kappa}} J_{\nu \mp 1}(\kappa r) e^{i\nu\phi} = \\ &= \mp \frac{i}{2} G_{|\nu| \mp 1}^{(1)} e^{i\nu\phi}. \end{aligned} \quad (8.6.53)$$

Из формул (8.2.33) — (8.2.38) получим составляющие поля для $r > a$

$$E_z = A \sqrt{\frac{a}{r}} J_\nu(\kappa a) e^{-\gamma(r-a)} e^{i\nu\phi} = \frac{1}{2} F_{z|\nu|}^{(2)} e^{i\nu\phi}, \quad (8.6.54)$$

$$\begin{aligned} H_z &= \mp i \frac{|\nu|}{\nu} n \sqrt{\frac{\varepsilon_0}{\mu_0}} A \sqrt{\frac{a}{r}} J_\nu(\kappa a) e^{-\gamma(r-a)} e^{i\nu\phi} = \\ &= \mp i \frac{|\nu|}{2\nu} G_z^{(2)} e^{i\nu\phi}, \end{aligned} \quad (8.6.55)$$

$$\begin{aligned} E_r &= \mp i \frac{nk_0 A}{\kappa} \sqrt{\frac{a}{r}} J_{\nu \mp 1}(\kappa a) e^{-\gamma(r-a)} e^{i\nu\phi} = \\ &= \mp \frac{i}{2} F_{|\nu| \mp 1}^{(2)} e^{i\nu\phi}, \end{aligned} \quad (8.6.56)$$

$$\begin{aligned} E_\phi &= \frac{|\nu|}{\nu} \frac{nk_0 A}{\kappa} \sqrt{\frac{a}{r}} J_{\nu \mp 1}(\kappa a) e^{-\gamma(r-a)} e^{i\nu\phi} = \\ &= \frac{|\nu|}{2\nu} F_{|\nu| \mp 1}^{(2)} e^{i\nu\phi}, \end{aligned} \quad (8.6.57)$$

$$\begin{aligned} H_r &= -\frac{|\nu|}{\nu} \sqrt{\frac{\varepsilon_0}{\mu_0} \frac{n^2 k_0 A}{\kappa}} \sqrt{\frac{a}{r}} J_{\nu \mp 1}(\kappa a) e^{-\gamma(r-a)} e^{i\nu\phi} = \\ &= -\frac{|\nu|}{2\nu} G_{|\nu| \mp 1}^{(2)} e^{i\nu\phi}, \end{aligned} \quad (8.6.58)$$

$$\begin{aligned} H_\phi &= \mp i \sqrt{\frac{\varepsilon_0}{\mu_0} \frac{n^2 k_0 A}{\kappa}} \sqrt{\frac{a}{r}} J_{\nu \mp 1}(\kappa a) e^{-\gamma(r-a)} e^{i\nu\phi} = \\ &= \mp \frac{i}{2} G_{|\nu| \mp 1}^{(2)} e^{i\nu\phi}. \end{aligned} \quad (8.6.59)$$

При выводе этих выражений были использованы соотношения (8.2.44) и (8.2.45), чтобы выразить C и D через A и B . Все функции Ханкеля аппроксимировались с помощью (8.2.54), а приближенные уравнения собственных значений (8.6.44) и (8.6.45) использовались для выражения $J_\nu(\kappa a)$ через $J_{\nu \mp 1}$. Верхний знак снова относится к НЕ-модам, а нижний — к ЕН-модам. Абсолютные значения ν соответствуют функциональным соотношениям (8.6.46) и обозначают порядок функции Бесселя. Как отмечалось в разд. 8.2 ниже уравнения (8.2.49), ν , входящее в экспоненту $\exp(i\nu\phi)$, можно заменить на его отрицательное значение, не изменяя порядка функции Бесселя. Когда мы хотим заменить $\exp(i\nu\phi)$ на $\exp(-i\nu\phi)$, то заменяем ν его отрицательным значением везде, кроме порядка функций Бесселя и тех мест, где записано абсолютное значение ν . Заменяя ν на $-\nu$ в формулах (8.6.48) — (8.6.59), складывая и вычитая уравнения, получаем компоненты электромагнитного поля мод волновода в следующем виде:

$$E_z = F_{zv} \begin{cases} \cos \nu\phi \\ i \sin \nu\phi \end{cases}, \quad (8.6.60)$$

$$H_z = \mp G_{zv} \begin{cases} -\sin \nu\phi \\ i \cos \nu\phi \end{cases}, \quad (8.6.61)$$

$$E_r = \mp F_{v \mp 1} \begin{cases} i \cos \nu\phi \\ -\sin \nu\phi \end{cases}, \quad (8.6.62)$$

$$E_\phi = F_{v \mp 1} \begin{cases} i \sin \nu\phi \\ \cos \nu\phi \end{cases}, \quad (8.6.63)$$

$$H_r = -G_{v \mp 1} \begin{cases} i \sin \nu\phi \\ \cos \nu\phi \end{cases}, \quad (8.6.64)$$

$$H_\phi = \mp G_{v \mp 1} \begin{cases} i \cos \nu\phi \\ -\sin \nu\phi \end{cases}. \quad (8.6.65)$$

(Выражения в скобках умножаются на F и G .) Два возможных знака и два типа тригонометрических функций в этих выражениях используются независимо друг от друга. Верхний знак относится к НЕ-модам, а нижний знак — к ЕН-модам. Два набора тригонометрических

функций выражают различные поляризации векторного поля. Каждая поляризация предполагает любой набор мод (НЕ или ЕН). Выражения для F и G вида (8.6.48) — (8.6.59) справедливы только вдали от отсечки и при малой разности $n_1 - n_2$. Однако выражения (8.6.60) — (8.6.65) являются точными и не зависят от приближения. В разд. 10.4 такая запись используется для описания моды HE_{11} .

Приближенные выражения составляющих поля полезны для получения простых выражений для полей в декартовых координатах. Используя обратные преобразования (8.2.3) и (8.2.4)

$$F_x = F_r \cos \phi - F_\phi \sin \phi \quad (8.6.66)$$

и

$$F_y = F_r \sin \phi + F_\phi \cos \phi, \quad (8.6.67)$$

можно записать поперечные компоненты поля (8.6.62) — (8.6.65) в следующем виде:

$$E_x = \mp F_{v \mp 1} \begin{cases} i \cos(v \mp 1) \phi \\ -\sin(v \mp 1) \phi \end{cases}, \quad (8.6.68)$$

$$E_y = F_{v \mp 1} \begin{cases} i \sin(v \mp 1) \phi \\ \cos(v \mp 1) \phi \end{cases}, \quad (8.6.69)$$

$$H_x = -G_{v \mp 1} \begin{cases} i \sin(v \mp 1) \phi \\ \cos(v \mp 1) \phi \end{cases}, \quad (8.6.70)$$

$$H_y = \mp G_{v \mp 1} \begin{cases} i \cos(v \mp 1) \phi \\ -\sin(v \mp 1) \phi \end{cases}. \quad (8.6.71)$$

Верхний и нижний знаки относятся соответственно к НЕ- и ЕН-модам. Из этих уравнений видно, что моды $HE_{1\mu}$ имеют только по одной поперечной составляющей E и H . Для НЕ-мод при $v = 1$ с верхними тригонометрическими функциями получаем составляющие E_x и H_y мод $HE_{1\mu}$. Нижний набор тригонометрических функций описывает поляризацию, плоскость которой повернута на 90° , так как в этом случае имеем только E_y и H_x . Таким образом, поле моды HE_{11} особенно простое вдали от отсечки и для волноводов с малыми значениями разности $n_1 - n_2$.

Можно сформировать комбинации из НЕ- и ЕН-мод, которые проще, чем поля (8.6.68) — (8.6.71). Для этого важно выявить вырождение мод НЕ порядка $\nu = \nu' + 1$ с модами ЕН порядка $\nu = \nu' - 1$. Это вырождение не точное. Оно приближенно выполняется в случае малых разностей $n_1 - n_2$, но не зависит от значения γa . Заменяя в формуле (8.6.32) ν на $\nu' + 1$ и используя (8.2.65) для исключения $J_{\nu'+1}$, получаем

$$\kappa a \frac{J_{\nu'-1}(\kappa a)}{J_{\nu'}(\kappa a)} = i\gamma a \frac{H_{\nu'-1}^{(1)}(i\gamma a)}{H_{\nu'}^{(1)}(i\gamma a)}. \quad (8.6.72)$$

Такое же уравнение получается, если в (8.6.33) заменить ν на $\nu' - 1$. Для приближений уравнений собственных значений в форме (8.6.32) и (8.6.33) моды НЕ порядка $\nu = \nu' + 1$ и моды ЕН порядка $\nu = \nu' - 1$ имеют одинаковые постоянные распространения. Такая же связь, очевидно, существует в (8.6.41) и (8.6.43).

Используем нижний набор тригонометрических функций в (8.6.68) — (8.6.71) для НЕ- и ЕН-мод и образуем следующие комбинации:

$$(E_x)_{\text{НЕ}\nu'+1} + (E_x)_{\text{ЕН}\nu'-1} = 0, \quad (8.6.73)$$

$$(E_y)_{\text{НЕ}\nu'+1} + (E_y)_{\text{ЕН}\nu'-1} = 2F_{\nu'} \cos \nu' \phi, \quad (8.6.74)$$

$$(H_x)_{\text{НЕ}\nu'+1} + (H_x)_{\text{ЕН}\nu'-1} = -2G_{\nu'} \cos \nu' \phi, \quad (8.6.75)$$

$$(H_y)_{\text{НЕ}\nu'+1} + (H_y)_{\text{ЕН}\nu'-1} = 0. \quad (8.6.76)$$

Полученные новые моды имеют только четыре не равные нулю составляющие поля E_z , E_y , H_z и H_x и их структура более простая, чем у первоначальных НЕ- и ЕН-мод. Вычитание НЕ- и ЕН-мод приводит к полю, повернутому на 90° по отношению к полю (8.6.73) — (8.6.76). Однако, поскольку вырождение между модами неточное, две моды распространяются с мало отличающимися фазовыми скоростями и линейно-поляризованное поле становится вследствие этого эллиптически-поляризованным. Но после прохождения модами расстояния $\Lambda = 2\pi N/(\beta_2 - \beta_1)$, где N — произвольное целое число, β_2 и β_1 — мало отличающиеся постоянные распространения первоначальных мод ЕН и НЕ, поле опять становится линейно-поляризованным и имеет вид (8.6.73) — (8.6.76).

НЕРЕГУЛЯРНЫЕ ДИЭЛЕКТРИЧЕСКИЕ ВОЛНОВОДЫ

9.1. ВВЕДЕНИЕ

Диэлектрический волновод является по существу многомодовой структурой. Даже в случае, когда по нему распространяется только одна направляемая мода, кроме нее всегда существует непрерывный спектр мод излучения. Если волновод идеально прямолинеен и можно пренебречь его диэлектрическими потерями, то направляемые моды распространяются без искажения и без затухания. Изучение мод такого идеального волновода является первым шагом в исследовании его свойств. Чтобы оценить возможности реального волновода, необходимо изучить его поведение в случае, когда имеют место отклонения от идеальной прямолинейной геометрии. Дело в том, что невозможно создать диэлектрические волноводы, в которых не наблюдались бы случайные отклонения.

К нерегулярностям ¹⁾ диэлектрических волноводов, которые необходимо учитывать, относятся тепловые потери в материале диэлектрика, отклонения от идеальной прямолинейности, неоднородность материала диэлектрика, отклонения поверхности раздела сердцевина — оболочка от идеальной плоскости в плоских волноводах или от идеальной цилиндрической поверхности в круглых оптических волокнах. Было бы слишком утомительно описывать влияние всех этих нерегулярностей в данной книге.

¹⁾ В оригинале используется термин «imperfections» — несовершенности, хотя речь идет в основном о неоднородностях параметров волновода вдоль его оси, для которых в теории волноводов принят более узкий термин «нерегулярности». Что касается тепловых потерь, то если они постоянны (регулярны) вдоль оси, к ним термин «нерегулярность» не относится. — *Прим. ред.*

Тепловые потери в материале волновода аналогичны тем, которые имеют место в сплошной диэлектрике. Особенно это справедливо, если материалы оболочки и сердцевины имеют одинаковые потери. В настоящей главе из всех упомянутых нерегулярностей будет рассмотрено влияние неровностей границы раздела сред волновода и кривизны оси волновода.

9.2. ПЛОСКИЙ ВОЛНОВОД С НЕРОВНОЙ ГРАНИЦЕЙ

Моды плоского волновода были получены в предыдущей главе путем решения уравнений поля внутри и вне волновода и дальнейшего определения амплитудных коэффициентов и постоянных распространения с учетом граничных условий. Такой способ скрывает тот факт, что функции поля плоского волновода на самом деле являются решениями уравнений (8.3.2) — (8.3.4) с изменяющимся n . Если бы изменение показателя преломления рассматривалось как непрерывное и мы были бы в состоянии решить приведенное волновое уравнение (8.3.4) для такого распределения показателя преломления, то не было бы необходимости обращаться к граничным условиям. Можно было бы аппроксимировать распределение показателя преломления гладкой функцией, решить волновое уравнение (8.3.4), а затем перейти к распределению показателя преломления с резкой границей. Достаточно напомнить, что задача о распространении мод в среде с квадратичным изменением показателя преломления ранее рассматривалась для приведенного волнового уравнения. Единственными граничными условиями, с которыми мы тогда столкнулись, были условия на бесконечности.

Предлагаемая теория плоского волновода с неровной стенкой основана на нахождении решений приведенного волнового уравнения во всем пространстве без учета граничных условий, кроме условий на бесконечности. Ранее было показано, что любое произвольное распределение поля можно представить в виде суперпозиции мод идеального прямолинейного волновода. Таким образом, и поле нерегулярного волновода можно выразить через моды регулярного волновода. Полученное разложение подстав-

ляется в приведенное волновое уравнение, а коэффициенты разложения определяются так, чтобы распределение поля приближенно было решением уравнения (8.3.4).

Вместо разложения поля нерегулярного волновода по модам регулярного волновода можно использовать так называемые локальные нормальные моды [113], которые сходны с модовыми решениями регулярного волновода. Фактически по форме они идентичны точным решениям для мод регулярного волновода. Однако размеры волновода, входящие как параметры в выражения для мод и уравнение собственных значений, являются функциями длины z в соответствии с точной формой нерегулярного волновода. Таким образом, локальные нормальные моды удовлетворяют граничным условиям на деформированной поверхности раздела сердцевина — оболочка нерегулярного волновода, являются взаимно ортогональными друг другу при интегрировании по сечению, но не удовлетворяют волновому уравнению или уравнениям Максвелла. В данной книге локальные нормальные моды не используются, вместо них применяется разложение произвольного поля по модам регулярного волновода. Для рассматриваемых задач такая процедура оказывается более простой¹⁾.

Потребуем, чтобы отклонения стенок волновода от прямолинейной геометрии подчинялись ограничению (8.3.1). Это условие означает, что толщина слоя волновода не меняется в направлении оси y . Отклонения стенок волновода в направлении оси x предполагаются достаточно малыми, так что их можно описать с помощью теории возмущений [84, 89, 98]. Плоский волновод с нерегулярными стенками изображен на фиг. 9.2.1, где отклонения стенок сильно увеличены. Распределение показате-

¹⁾ Здесь автор дает слишком упрощенное изложение метода, использующего локальные нормальные моды. Этот метод получил название «метода поперечных сечений» [59*, 113]. С его помощью получают строгие уравнения для поля в нерегулярных волноводах. Излагаемый автором метод является приближенным, который применим в случае малых по величине нерегулярностей. Его можно рассматривать как приближенный частный случай метода поперечных сечений. — *Прим. ред.*

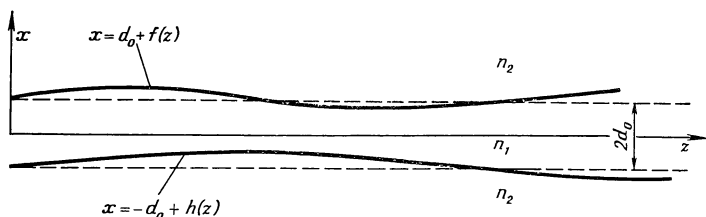
ля преломления в волноводе с нерегулярными стенками можно записать в виде

$$n^2(x, z) = n_0^2(x, z) + \eta(x, z). \quad (9.2.1)$$

Распределение показателя преломления n_0 описывает регулярный плоский волновод:

$$n_0(x, z) = \begin{cases} n_2 & \text{для } |x| > d_0, \\ n_1 & \text{для } |x| < d_0. \end{cases} \quad (9.2.2)$$

В случае регулярного волновода $\eta = 0$. Для нерегуляр-



Ф и г. 9.2.1. Плоский волновод с нерегулярной границей раздела сердцевина — оболочка.

ного волновода получаем следующие выражения:

$$\eta = \left\{ \begin{array}{ll} 0 & \text{для } x > d_0 + f(z) \\ n_1^2 - n_2^2 & \text{для } d_0 < x < d_0 + f(z) \\ 0 & \text{для } -d_0 + h(z) < x < d_0 \\ -(n_1^2 - n_2^2) & \text{для } -d_0 < x \leq -d_0 + h(z) \\ 0 & \text{для } -\infty < x < -d_0 \end{array} \right\} \text{ для } f > 0, h > 0. \quad (9.2.3)$$

Подобные соотношения справедливы также для тех значений z , при которых $f(z) < 0$, $h(z) > 0$, и для всех других возможных комбинаций. Для применения теории возмущений разность $n_1^2 - n_2^2$ не обязательно должна быть величиной малого порядка. Достаточно, чтобы область, в пределах которой η отличается от нуля, была очень узкой.

Подставив формулы (8.5.16) и (9.2.1) в (8.3.4), получим

$$\sum_{\nu} \left[\frac{\partial^2 c_{\nu}}{\partial z^2} - 2i\beta_{\nu} \frac{\partial c_{\nu}}{\partial z} + c_{\nu} k_0^2 \eta \right] E_{\nu y} +$$

$$+ \sum \int_0^{\infty} \left[\frac{\partial^2 q(\rho)}{\partial z^2} - 2i\beta \frac{\partial q(\rho)}{\partial z} + q(\rho) k_0^2 \eta \right] E_y(\rho) d\rho = 0. \quad (9.2.4)$$

При выводе этого уравнения использовалось то обстоятельство, что моды $E_{\nu y}$ и $E_y(\rho)$ удовлетворяют волновому уравнению для регулярного плоского волновода. Предполагалось также, что поле нерегулярного волновода имеет только три составляющие: H_x , H_z и E_y . Составляющая E_y уже определялась в (8.5.16), а составляющие магнитного поля можно получить с помощью (8.3.2) и (8.3.3). Задача ставится следующим образом. Мода ТЕ низшего порядка из регулярной части волновода падает на его нерегулярную часть. Нужно определить, как мощность, переносимая падающей модой ТЕ регулярного волновода, преобразуется в другие волноводные моды и излучается.

Уравнение (9.2.4) содержит производные только относительно координаты z . Однако это уравнение зависит еще от координаты x через зависимость от x нормальных мод и η . Для получения системы связанных дифференциальных уравнений умножим обе части уравнения (9.2.4) на

$$\frac{\beta_{\mu}}{2\omega\mu_0} E_{\mu y}^*, \quad (9.2.5)$$

где μ — целое число, μ_0 — магнитная проницаемость, и проинтегрируем по всему поперечному сечению волновода. Используя соотношение ортогональности (8.5.13), получим

$$\frac{\partial^2 c_{\mu}}{\partial z^2} - 2i\beta_{\mu} \frac{\partial c_{\mu}}{\partial z} + \sum_{\nu} c_{\nu} F_{\nu\mu}(z) +$$

$$+ \sum \int_0^{\infty} q(\rho) G_{\mu}(\rho, z) d\rho = 0, \quad (9.2.6)$$

где

$$F_{\nu\mu}(z) = \frac{\beta_{\mu} k_0^2}{2\omega\mu_0 P} \int_{-\infty}^{\infty} E_{\mu y}^* \eta(x, z) E_{\nu y} dx \quad (9.2.7)$$

и

$$G_{\mu}(\rho) = \frac{\beta_{\mu} k_0^2}{2\omega\mu_0 P} \int_{-\infty}^{\infty} E_{\mu y}^* \eta(x, z) E_y(\rho) dx. \quad (9.2.8)$$

Аналогично получим

$$\begin{aligned} \frac{\partial^2 q(\rho')}{\partial z^2} - 2i\beta' \frac{\partial q(\rho')}{\partial z} + \sum_{\nu} c_{\nu} \bar{F}_{\nu}(\rho', z) + \\ + \sum \int_0^{\infty} q(\rho) \bar{G}(\rho, \rho') d\rho, \end{aligned} \quad (9.2.9)$$

где

$$\bar{F}_{\nu}(\rho', z) = \frac{\beta' k_0^2}{2\omega\mu_0 P} \int_{-\infty}^{\infty} E_y^*(\rho') \eta(x, z) E_{\nu y} dx \quad (9.2.10)$$

и

$$\bar{G}(\rho, \rho') = \frac{\beta' k_0^2}{2\omega\mu_0 P} \int_{-\infty}^{\infty} E_y^*(\rho') \eta(x, z) E_y(\rho) dx. \quad (9.2.11)$$

Постоянная распространения β' мод излучения зависит от ρ' . Очевидно, что наличие возмущения η связывает коэффициенты разложения c_{ν} и $q(\rho)$. Система интегродифференциальных уравнений (9.2.6) и (9.2.9), конечно, достаточно сложна.

Перед тем как пытаться решить систему уравнений, рассмотрим ее решение при $\eta = 0$. В этом случае получаются несвязанные уравнения

$$\frac{\partial^2 c_{\mu}}{\partial z^2} - 2i\beta_{\mu} \frac{\partial c_{\mu}}{\partial z} = 0 \quad (9.2.12)$$

и

$$\frac{\partial^2 q(\rho')}{\partial z^2} - 2i\beta' \frac{\partial q(\rho')}{\partial z} = 0. \quad (9.2.13)$$

Так как оба уравнения одинаковы, то достаточно рассмотреть одно из них. Решение уравнения (9.2.12) имеет вид

$$c_{\mu} = A + B e^{2i\beta_{\mu} z}. \quad (9.2.14)$$

Моды, которые входили в ряды разложения (8.5.16), имели зависимость от координаты z и времени в виде (8.3.7).

Так как величина β_μ всегда берется положительной, то все моды распространяются в положительном направлении z . В разложения, однако, эти моды входят умноженными на c_ν . Пространственно-временная зависимость комбинации $c_\nu E_{\nu y}$ дается соотношением

$$c_\nu e^{i(\omega t - \beta_\nu z)} = A e^{i(\omega t - \beta_\nu z)} + B e^{i(\omega t + \beta_\nu z)}. \quad (9.2.15)$$

Таким образом, видно, что, хотя мы начали с мод, распространяющихся в положительном направлении оси z , в итоге получили моды, распространяющиеся как в положительном, так и в отрицательном направлениях оси z .

Теперь рассмотрим следующие неоднородные дифференциальные уравнения:

$$\frac{\partial^2 c_\mu}{\partial z^2} - 2i\beta_\mu \frac{\partial c_\mu}{\partial z} = \phi_\mu(z) \quad (9.2.16)$$

и

$$\frac{\partial^2 q(\rho')}{\partial z^2} - 2i\beta' \frac{\partial q(\rho')}{\partial z} = \psi_{\rho'}(z). \quad (9.2.17)$$

Уравнение (9.2.16) имеет решение

$$c_\mu(z) = A_\mu + B_\mu e^{2i\beta_\mu z} + \frac{1}{2i\beta_\mu} \left[- \int_0^z \phi_\mu(\xi) d\xi + e^{2i\beta_\mu z} \int_0^z e^{-2i\beta_\mu \xi} \phi_\mu(\xi) d\xi \right]. \quad (9.2.18)$$

Удобно разделить коэффициент c_μ на две части. Первая часть

$$c_\mu^{(+)} = A_\mu - \frac{1}{2i\beta_\mu} \int_0^z \phi_\mu(\xi) d\xi \quad (9.2.19)$$

соответствует волнам, распространяющимся в положительном направлении оси z . Вторая часть

$$c_\mu^{(-)} = \left[B_\mu + \frac{1}{2i\beta_\mu} \int_0^z e^{-2i\beta_\mu \xi} \phi_\mu(\xi) d\xi \right] e^{2i\beta_\mu z} \quad (9.2.20)$$

соответствует волнам, распространяющимся в отрицательном направлении оси z . Может показаться, что такое разде-

ление коэффициента на «положительную» и «отрицательную» части зависит не только от множителей перед интегралами, но и от способа интегрирования. Вообще говоря, это действительно так. Однако оказывается, что в данном методе разделение на две части, соответствующие двум волнам, распространяющимся в противоположных направлениях, имеет физический смысл ¹⁾.

Коэффициенты A_μ и B_μ определяются из начальных условий. Предположим, что из регулярной части плоского волновода на его нерегулярную часть протяженностью L падает четная мода ТЕ низшего порядка. За нерегулярным участком волновода снова следует регулярный участок бесконечной протяженности. Такая модель позволит вычислить вклад в потери от нерегулярного участка волновода длиной L . Сделанное предположение дает возможность сформулировать определенные начальные условия. Допустим, что при $z = 0$ только четная мода ТЕ низшего порядка распространяется направо. Это позволяет утверждать без доказательства, что амплитуда c_0 этой моды должна равняться единице:

$$c_0^{(+)}(0) = 1, \quad (9.2.21)$$

тогда как для всех других мод амплитуды должны обращаться в нуль:

$$c_\mu^{(+)}(0) = 0 \quad \text{для} \quad \mu \neq 0. \quad (9.2.22)$$

Подобным образом постулируем, что при $z = L$ не появляется мод, распространяющихся в отрицательном направлении от z :

$$c_\mu^{(-)}(L) = 0 \quad \text{для всех} \quad \mu. \quad (9.2.23)$$

Начальные условия (9.2.21) и (9.2.22) приводят к

$$A_\mu = \delta_{0\mu}, \quad (9.2.24)$$

а условие (9.2.23) дает

$$B_\mu = -\frac{1}{2i\beta_\mu} \int_0^L e^{-2i\beta_\mu \xi} \phi_\mu(\xi) d\xi. \quad (9.2.25)$$

Аналогичные выражения получаются и при решении уравнения (9.2.17).

¹⁾ Пояснение этого вопроса см. в [113, гл. 1].— *Прим. ред.*

Выражение (8.5.22) определяет полную мощность, переносимую всеми модами. Если исключить мощность падающей моды ТЕ и расширить (8.5.22) таким образом, чтобы оно включало как мощность, проходящую через бесконечную плоскость при $z = 0$ налево, так и мощность, проходящую через бесконечную плоскость при $z = L$ направо, то получим величину мощности, которую падающая мода передает прямым и обратным направляемым модам и прямым и обратным модам излучения. Относительные потери мощности падающей моды ТЕ окончательно можно определить из следующего соотношения:

$$\frac{\Delta P}{P} = \sum_{v=0}^{\infty} [(1 - \delta_{0v}) |c_v^{(+)}(L)|^2 + |c_v^{(-)}(0)|^2] + \\ + \sum_1^2 \int_0^{\infty} [|q^{(+)}(\rho, L)|^2 + |q^{(-)}(\rho, 0)|^2] d\rho. \quad (9.2.26)$$

С помощью формул (9.2.19), (9.2.20), (9.2.24) и (9.2.25) это соотношение можно переписать в виде

$$\frac{\Delta P}{P} = \sum_{v=0}^{\infty} \left[(1 - \delta_{0v}) \left| \frac{1}{2i\beta_v} \int_0^L \phi_v(\xi) d\xi \right|^2 + \right. \\ \left. + \left| \frac{1}{2i\beta_v} \int_0^L e^{-2i\beta_v \xi} \phi_v(\xi) d\xi \right|^2 \right] + \\ + \sum_1^2 \int_0^{\infty} \left[\left| \frac{1}{2i\beta(\rho)} \int_0^L \psi_\rho(\xi) d\xi \right|^2 + \right. \\ \left. + \left| \frac{1}{2i\beta(\rho)} \int_0^L e^{-2i\beta(\rho)\xi} \psi_\rho(\xi) d\xi \right|^2 \right] d\rho. \quad (9.2.27)$$

Сумма от 1 до 2 перед знаком интеграла соответствует вкладу в потери четных и нечетных мод излучения. Предполагается, что сумма по направляемым модам содержит четные и нечетные моды.

Наша теория полезна в том случае, когда можно получить выражения для функций $\phi_v(z)$ и $\psi_\rho(z)$. Выполним это с помощью теории возмущений. Теория возмущений пред-

полагает, что величина мощности, переносимой падающей четной модой ТЕ низшего порядка, не уменьшается очень сильно при прохождении модой нерегулярного участка волновода. Другими словами, предполагается, что $c_0^{(+)}(z)$ остается почти равной единице. Это ограничивающее допущение будет ослаблено для некоторых возможных частных случаев. Поскольку амплитуда падающей моды не отклоняется существенно от своего начального единичного значения, то можно предположить, что амплитуды нежелательных мод не принимают больших значений. Такое предположение является достаточно обоснованным, если только величины, определяемые выражениями (9.2.7) (9.2.8), (9.2.10) и (9.2.11), малы. Произведения коэффициентов разложения на эти величины являются малыми величинами второго порядка и ими можно пренебречь. Единственный член в суммах (9.2.6) и (9.2.9), который нужно учитывать, есть произведение $c_0 = 1$ на малые величины $F_{\nu\mu}(z)$. С учетом этих предположений и путем сравнения формул (9.2.6) и (9.2.16) получим

$$\phi_\mu = -c_0 F_{0\mu} = -\frac{\beta_\mu k_0^2}{2\omega\mu_0 P} \int_{-\infty}^{\infty} E_{\mu y}^* \eta E_{0y} dx \quad (9.2.28)$$

и аналогичным образом

$$\begin{aligned} \psi_{\rho'}(z) &= -c_0 \bar{F}_0(\rho', z) = \\ &= -\frac{\beta(\rho') k_0^2}{2\omega\mu_0 P} \int_{-\infty}^{\infty} E_y^*(\rho') \eta E_{0y} dx. \end{aligned} \quad (9.2.29)$$

Область, в которой $\eta(x)$ дает вклад в интегралы, предполагается очень узкой. Это допущение находится в соответствии с принятыми постулатами и физически означает, что границы волновода очень незначительно отклоняются от их идеальной прямолинейной формы. Поскольку η отличается от нуля только в пределах узкой области, составляющие поля под интегралами можно считать постоянными внутри области значений x , где η дает вклад в интегралы. Заменим поля их значениями при $x = \pm d$. Выражение (9.2.28) с учетом формул (8.3.9), (8.3.18) и (9.2.3)

принимает вид

$$\begin{aligned} \phi_{\mu e}(z) = \\ = -(n_1^2 - n_2^2) \frac{\beta_\mu k_0^2 \cos \kappa_0 d_0 \cos \kappa_\mu d_0}{\sqrt{\left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0}\right) \left(\beta_\mu d_0 + \frac{\beta_\mu}{\gamma_\mu}\right)}} [f(z) - h(z)] e^{i(\beta_\mu - \beta_0)z} \end{aligned} \quad (9.2.30)$$

для случая, когда μ дает четную моду ТЕ, и

$$\begin{aligned} \phi_{\mu 0}(z) = \\ = -(n_1^2 - n_2^2) \frac{\beta_\mu k_0^2 \cos \kappa_0 d_0 \sin \kappa_\mu d_0}{\sqrt{\left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0}\right) \left(\beta_\mu d_0 + \frac{\beta_\mu}{\gamma_\mu}\right)}} [f(z) + h(z)] e^{i(\beta_\mu - \beta_0)z} \end{aligned} \quad (9.2.31)$$

для случая, когда μ дает нечетную моду ТЕ. Зависимость от координаты z типа (8.3.7) в этих выражениях, таким образом, восстановлена. Функцию (9.2.29) можно аналогичным образом преобразовать к виду

$$\begin{aligned} \psi_{\rho e} = -(n_1^2 - n_2^2) \frac{\sqrt{\beta(\rho)} k_0^2 \rho \cos \kappa_0 d_0 \cos \sigma d_0}{\sqrt{\pi \left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0}\right) (\rho^2 \cos^2 \sigma d_0 + \sigma^2 \sin^2 \sigma d_0)}} \times \\ \times [f(z) - h(z)] e^{i[\beta(\rho) - \beta_0]z} \end{aligned} \quad (9.2.32)$$

для четных мод излучения и

$$\begin{aligned} \psi_{\rho 0} = -(n_1^2 - n_2^2) \frac{\sqrt{\beta(\rho)} k_0^2 \rho \cos \kappa_0 d_0 \sin \sigma d_0}{\sqrt{\pi \left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0}\right) (\rho^2 \sin^2 \sigma d_0 + \sigma^2 \cos^2 \sigma d_0)}} \times \\ \times [f(z) + h(z)] e^{i[\beta(\rho) - \beta_0]z} \end{aligned} \quad (9.2.33)$$

для нечетных мод излучения. Собственные значения β_μ и параметры κ_μ и γ_μ являются решениями уравнения собственных значений (8.3.16), если они относятся к четным модам, и уравнения (8.3.23), если они соответствуют нечетным модам. Для простоты два типа собственных значений обозначены одинаково. Напомним, что функции $f(z)$ и $h(z)$ вошли в выражения из (9.2.3). Эти функции описывают отклонения верхней и нижней границ плоского волновода.

Интересно рассмотреть коэффициенты, которые входят в выражение (9.2.26) для относительных потерь мощности. Для амплитуд четных направляемых мод ТЕ из формул (9.2.19), (9.2.24) и (9.2.30) получим

$$c_{\mu e}^{(+)}(L) = (n_1^2 - n_2^2) \frac{L k_0^2 \cos \kappa_0 d \cos \kappa_\mu d_0}{2i \sqrt{\left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0}\right) \left(\beta_\mu d_0 + \frac{\beta_\mu}{\gamma_\mu}\right)}} \times \\ \times [F(\beta_0 - \beta_\mu) - H(\beta_0 - \beta_\mu)], \quad (9.2.34)$$

где

$$F(\beta_0 - \beta_\mu) = \frac{1}{L} \int_0^L f(z) e^{-i(\beta_0 - \beta_\mu)z} dz \quad (9.2.35)$$

и

$$H(\beta_0 - \beta_\mu) = \frac{1}{L} \int_0^L h(z) e^{-i(\beta_0 - \beta_\mu)z} dz. \quad (9.2.36)$$

Коэффициент для нечетных направляемых мод ТЕ имеет вид

$$c_{\mu o}^{(+)}(L) = (n_1^2 - n_2^2) \frac{L k_0^2 \cos \kappa_0 d_0 \sin \kappa_\mu d_0}{2i \sqrt{\left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0}\right) \left(\beta_\mu d_0 + \frac{\beta_\mu}{\gamma_\mu}\right)}} \times \\ \times [F(\beta_0 - \beta_\mu) + H(\beta_0 - \beta_\mu)]. \quad (9.2.37)$$

Коэффициент для четных мод излучения определяется выражением

$$q_e^{(+)}(\rho, L) = \frac{n_1^2 - n_2^2}{2i} \frac{L \rho k_0^2 \cos \kappa_0 d_0 \cos \sigma d_0}{\sqrt{\pi \beta \left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0}\right) (\rho^2 \cos^2 \sigma d_0 + \sigma^2 \sin^2 \gamma d_0)}} \times \\ \times [F(\beta_0 - \beta) - H(\beta_0 - \beta)], \quad (9.2.38)$$

а для нечетных мод

$$q_o^{(+)}(\rho, L) = \frac{n_1^2 - n_2^2}{2i} \frac{L \rho k_0^2 \cos \kappa_0 d_0 \sin \sigma d_0}{\sqrt{\pi \beta \left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0}\right) (\rho^2 \sin^2 \sigma d_0 + \sigma^2 \cos^2 \sigma d_0)}} \times \\ \times [F(\beta_0 - \beta) + H(\beta_0 - \beta)], \quad (9.2.39)$$

где $\cos \kappa_0 d_0$ можно выразить через κ и k_0 с помощью формулы (8.6.16). Соответствующие коэффициенты $c_{\mu}^{(-)}(0)$ и т. д. получаются из установленных ранее коэффициентов простой заменой β_μ на $-\beta_\mu$ или β на $-\beta$ в функциях F и H .

Следует обратить внимание в этих выражениях на их зависимость от F и H , которые являются просто фурье-преобразованиями функций искривления стенки f и h . Важно отметить, что эти фурье-преобразования берутся по той же области, где имеют место потери. Приведенные выражения справедливы для потерь плоского волновода длиной L . Эта же длина используется как период интегрирования для коэффициентов Фурье. Такая зависимость от длины L преобразований Фурье окажется важной при рассмотрении потерь в волноводах как с систематическими, так и со случайными отклонениями стенки.

Таким образом, получены все выражения, необходимые для вычисления потерь, которые испытывает четная ТЕ-мода низшего порядка при распространении через участок плоского волновода с нерегулярной стенкой.

9.3. ПЛОСКИЙ ВОЛНОВОД С СИНУСОИДАЛЬНЫМИ ИСКАЖЕНИЯМИ СТЕНКИ

Применим теорию возмущений, развитую в предыдущем разделе, к специальному случаю. Предположим, что функции искажения стенки имеют следующий вид [98]:

$$f(z) = a \sin \theta z \quad (9.3.1)$$

и

$$h(z) = -a \sin(\theta z + \phi). \quad (9.3.2)$$

При

$$\phi = 0 \quad (9.3.3)$$

формулы (9.3.1) и (9.3.2) описывают плоский волновод, толщина которого меняется периодически. Если

$$\phi = \pi, \quad (9.3.4)$$

то плоский волновод имеет постоянную толщину, но периодически изменяющееся направление. Коэффициенты Фурье этих функций искажения следуют из (9.2.35):

$$F(\beta_0 - \beta_\mu) = \frac{a}{iL} e^{i[\theta - (\beta_0 - \beta_\mu)]L/2} \frac{\sin[\theta - (\beta_0 - \beta_\mu)] \frac{L}{2}}{\theta - (\beta_0 - \beta_\mu)} \quad (9.3.5)$$

и из (8.6.38):

$$H(\beta_0 - \beta_\mu) = -\frac{a}{iL} e^{i\{\theta - (\beta_0 - \beta_\mu)L/2 + \phi\}} \frac{\sin[\theta - (\beta_0 - \beta_\mu)] \frac{L}{2}}{\theta - (\beta_0 - \beta_\mu)}. \quad (9.3.6)$$

Членами с $\theta + (\beta_0 - \beta_\mu)$ в этих двух выражениях мы пренебрегли. Так как β_0 соответствует моде низшего порядка, то оно больше β_μ любой другой моды и $\beta_0 - \beta_\mu > 0$. Предположим, что L много больше длины волны в волноводе. Очевидно, что коэффициенты Фурье имеют большие значения, если только знаменатели в формулах (9.3.5) и (9.3.6) близки к нулю. Во всех других случаях коэффициенты Фурье пренебрежимо малы. По этой причине членами с $\theta + (\beta_0 - \beta_\mu)$ можно пренебречь, хотя они и не равны строго нулю. Исчезающе малые коэффициенты Фурье дают в результате исчезающе малые амплитуды мод. Данное рассмотрение позволяет сформулировать следующий важный результат. Две моды в волноводе являются связанными из-за синусоидального отклонения стенки только в том случае, если частота θ механического искажения стенки связана с постоянными распространения этих мод следующим соотношением:

$$\theta = \beta_0 - \beta_\mu. \quad (9.3.7)$$

Это утверждение справедливо как для направляемых паразитных мод, так и для мод излучения. Таким образом, синусоидальное искажение стенки связывает падающую моду либо с другой направляемой модой, либо с узкой областью мод непрерывного спектра.

Сначала предположим, что равенство (9.3.7) удовлетворяется для направляемой паразитной моды. Из формул (9.3.5) и (9.3.6) с учетом (9.3.7) и $\phi = 0$ получим

$$F(\beta_0 - \beta_\mu) = \frac{a}{2i} \quad (9.3.8)$$

и

$$H(\beta_0 - \beta_\mu) = -\frac{a}{2i}. \quad (9.3.9)$$

Из формулы (9.2.37) видно, что в этом случае нет связи с нечетными модами. Симметричное изменение толщины плоского волновода связывает симметричные падающие

моды только с другими симметричными модами. Если положить $\phi = \pi$, то знак в (9.3.9) будет положительным и связь будет возможна только с нечетными модами.

Используем полученные соотношения для вычисления с помощью формулы (9.2.26) относительных потерь мощности падающей моды при связи с одной паразитной модой ($\phi = 0$):

$$\left(\frac{\Delta P}{P}\right)_g = a^2 L^2 \frac{[(n_1^2 - n_2^2) k_0^2 \cos \kappa_0 d \cos \kappa_\mu d]^2}{4 \left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0}\right) \left(\beta_\mu d_0 + \frac{\beta_\mu}{\gamma_\mu}\right)}. \quad (9.3.10)$$

Связь падающей моды с модами излучения подчиняется такому же закону. Как уже отмечалось, изменения толщины связывают падающую моду только с четными модами излучения, а периодические изменения направления волновода связывают ее только с нечетными модами. Потери мощности в модах излучения получаются не так просто, так как даже строго синусоидальное искажение стенки связывает падающую моду с бесконечным числом мод излучения. Если соотношение (9.3.7) удовлетворяется для мод излучения, то θ попадает внутрь диапазона волновых чисел мод излучения. Поскольку мощность, рассеянную назад, необходимо рассмотреть наряду с мощностью, рассеянной вперед, диапазон возможных мод излучения, который представляет интерес, определяется соотношением

$$-n_2 k_0 < \beta < n_2 k_0. \quad (9.3.11)$$

Диапазон частот механического изменения стенки, в котором может иметь место связь падающей моды с модами излучения, задается неравенством

$$\beta_0 - n_2 k_0 < \theta < \beta_0 + n_2 k_0. \quad (9.3.12)$$

Синусоидальные искажения стенки, частоты которых выпадают из диапазона (9.3.12), не вызывают потерь мощности на излучение. Этот результат справедлив, строго говоря, только в пределе для бесконечно длинных участков волновода и выполняется с хорошим приближением для длинных по сравнению с длиной волны конечных участков волновода.

Формулу (9.2.26) для потерь мощности, вызванных излучением, можно записать в более удобном виде. Интеграл

в (9.2.26) включает распространяющиеся моды излучения и затухающие моды излучения. Так как последние мощности не переносят, то можно ограничить диапазон интегрирования распространяющимися модами излучения. Сделаем замену переменной. Из формулы (8.4.6) получим

$$d\rho = -\frac{\beta}{\rho} d\beta. \quad (9.3.13)$$

Интегрирование производится от $\beta = 0$ до $\beta = n_2 k_0$. Из предыдущего видно, что коэффициент $q^{(-)}(\rho)$ получается из $q^{(+)}(\rho)$ изменением постоянной распространения β , которая входит в коэффициенты Фурье, на $-\beta$. Таким образом, можно объединить два интеграла в (9.2.26) в один, расширяя пределы интегрирования от $-n_2 k_0$ до $n_2 k_0$. Относительные потери мощности, вызванные только излучением, запишутся в виде

$$\frac{\Delta P}{P} = \int_{-n_2 k_0}^{n_2 k_0} [|q_e|^2 + |q_0|^2] \frac{|\beta|}{\rho} d\beta. \quad (9.3.14)$$

В индексах «+» и «-» больше нет необходимости, и поэтому они опущены. Суммирование в (9.2.26) распространялось на четные и нечетные моды под знаком интеграла.

В частном случае при $\phi = 0$ имеют место только коэффициенты разложения q_e для четных мод. Используя формулы (9.3.5) и (9.3.6), получим

$$\begin{aligned} \frac{\Delta P}{P} = & \frac{a^2 (n_1^2 - n_2^2)^2 k_0^4 \cos^2 \kappa_0 d_0}{\pi \left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0} \right)} \times \\ & \times \int_{-n_2 k_0}^{n_2 k_0} \frac{\rho \sin^2 \Gamma \frac{L}{2} \cos^2 \sigma d_0}{\Gamma^2 (\rho^2 \cos^2 \sigma d_0 + \sigma^2 \sin^2 \sigma d_0)} d\beta, \end{aligned} \quad (9.3.15)$$

где

$$\Gamma = \theta - (\beta_0 - \beta). \quad (9.3.16)$$

Для малых значений L интегрировать затруднительно. Однако, если L много больше длины волны излучения, можно получить хорошее приближение. Согласно (8.4.16),

для больших L

$$\delta(\Gamma) \approx \frac{1}{\pi} \frac{\sin \Gamma \frac{L}{2}}{\Gamma}, \quad (9.3.17)$$

т. е. фактически под интегралом стоит квадрат дельта-функции. Однако квадрат дельта-функции ведет себя подобно дельта-функции. Это позволяет вынести множители при дельта-функции из-под интеграла, так как интегрирование дельта-функции производится по бесконечно малому интервалу. Результирующий интеграл имеет вид

$$\int_{-n_2 k_0}^{n_2 k_0} \frac{\sin^2 \Gamma \frac{L}{2}}{\Gamma^2} d\beta \approx \int_{-\infty}^{\infty} \frac{\sin^2 \Gamma \frac{L}{2}}{\Gamma^2} d\Gamma = \frac{\pi L}{2}. \quad (9.3.18)$$

Переход здесь к бесконечным пределам интегрирования не является существенным, так как основной вклад дает интегрирование только в непосредственной окрестности $\Gamma = 0$.

Относительные потери мощности, вызванные периодическим изменением толщины плоского волновода, в конечном счете определяются выражением

$$\frac{\Delta P}{P} = 2\alpha L = L \frac{a^2 (n_1^2 - n_2^2)^2 k_0^4 \rho_0 \cos^2 \kappa_0 d_0 \cos^2 \sigma_0 d_0}{2 \left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0} \right) (\rho_0^2 \cos^2 \sigma_0 d_0 + \sigma_0^2 \sin^2 \sigma_0 d_0)}, \quad (9.3.19)$$

где

$$\rho_0 = \sqrt{n_2 k_0^2 - (\beta_0 - \theta)^2}, \quad (9.3.20)$$

$$\sigma_0 = \sqrt{n_1^2 k_0^2 - (\beta_0 - \theta)^2}, \quad (9.3.21)$$

α — амплитудный коэффициент потерь.

Существует очень интересное различие между потерями мощности для одной направляемой моды (9.3.10) и потерями мощности для континуума мод излучения. Потери, вызванные направляемой модой, пропорциональны L^2 , а потери для непрерывного спектра мод излучения пропорциональны L . Данный расчет обоих видов потерь справедлив до тех пор, пока $\Delta P/P \ll 1$.

Позднее мы исследуем связь падающей моды с направляемыми модами более подробно. Потери на излучение можно объяснить следующим образом. Излучение, вызванное

искажением стенки, покидает волновод и исчезает в пространстве. Ниже в этом разделе будет рассчитана структура поля излучения в дальней зоне. При потере малой части мощности падающая мода практически не изменяется и ее амплитуда уменьшается незначительно. Преобразованная в излучение мощность больше не направляется волноводом, а излучается в пространство. Это не имеет места для случая связи падающей моды с направляемой, поскольку переданная направляемой моде мощность снова взаимодействует с мощностью падающей моды. Однако в случае излучения такое взаимодействие очень слабое. Этот результат позволяет расширить диапазон применения теории возмущений. Используя формулу (9.3.19) неоднократно к уменьшенному значению мощности падающей моды, получим на конце волновода длиной D мощность

$$P = P_0 e^{-2\alpha D}. \quad (9.3.22)$$

Коэффициент затухания α определяется из формулы (9.3.19). Выражение (9.3.22) справедливо для малых значений $2\alpha\lambda$, где λ — длина волны, и точность его увеличивается с ростом угла, под которым излучение покидает волновод.

Из анализа структуры поля в дальней зоне получаем угол излучения [100]. Для коэффициента разложения $q_e^{(+)}$ в предположении (9.3.3) из формул (9.2.38), (9.3.5) и (9.3.6) находим следующее выражение:

$$q_e^{(+)}(\rho, L) = -(n_1^2 - n_2^2) \times \\ \times \frac{apk_0^2 \cos \kappa_0 d_0 \cos \sigma d_0 e^{i\Gamma L/2}}{\sqrt{\pi \beta \left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0} \right) (\rho^2 \cos^2 \sigma d_0 + \sigma^2 \sin^2 \sigma d_0)}} \frac{\sin \Gamma \frac{L}{2}}{\Gamma}, \quad (9.3.23)$$

где Γ определяется формулой (9.3.16). Составляющая электрического поля излучения получается из (8.5.16) в виде

$$E_y = \int_0^\infty q_e^{(+)}(\rho, L) E_y(\rho, z) d\rho. \quad (9.3.24)$$

Выражения (9.3.23) и (9.3.24) описывают поле излучения плоского волновода с синусоидальным изменением его толщины в точке пространства с координатами x, z . излу-

чение вызывается нерегулярным участком длиной L плоского волновода. Нас интересует только поле в дальней зоне, поэтому можно считать, что

$$L \ll z. \quad (9.3.25)$$

Поле моды излучения $E_y(\rho)$ вне слоя волновода определяется формулами (8.4.4), (8.4.9), (8.4.10) и (8.4.18). Объединение этих выражений приводит к следующему результату:

$$E_y(\rho) = \frac{\rho \sqrt{2\omega\mu_0 P} \left[\cos \sigma d_0 \cos \rho (|x| - d_0) - \frac{\sigma}{\rho} \sin \sigma d_0 \sin \rho (|x| - d_0) \right]}{\sqrt{\pi\beta (\rho^2 \cos^2 \sigma d_0 + \sigma^2 \sin^2 \sigma d_0)}} e^{-i\beta z}. \quad (9.3.26)$$

Окончательно выражение для составляющей электрического поля излучения получается из формул (9.3.23), (9.3.24) и (9.3.26) в виде

$$E_y = - \frac{ak_0^2 \sqrt{2\omega\mu_0 P} (n_1^2 - n_2^2) \cos \kappa_0 d_0}{\pi \sqrt{\beta_0 d_0 + \frac{\beta_0}{\gamma_0}}} \times \\ \int_0^\infty \frac{e^{i\Gamma L/2} \rho^2 \cos \sigma d_0 \left[\cos \sigma d_0 \cos \rho (|x| - d_0) - \frac{\sigma}{\rho} \sin \sigma d_0 \sin \rho (|x| - d_0) \right]}{\beta (\rho^2 \cos^2 \sigma d_0 + \sigma^2 \sin^2 \sigma d_0)} \times \\ \times \frac{\sin \Gamma \frac{L}{2}}{\Gamma} e^{-i\beta z} d\rho. \quad (9.3.27)$$

Интеграл в этом выражении достаточно сложен и не приходится рассчитывать на получение общего решения. Однако можно получить его приближенное значение методом стационарной фазы [9] и определить, таким образом, поле в дальней зоне. Этот полезный приближенный метод вычисления интегралов, имеющих быстро осциллирующие функции в подынтегральных выражениях, был рассмотрен в разд. 2.3. Предполагается, что x и z становятся большими по сравнению с L . Это заставляет функции, содержащие x и z в аргументах, осциллировать намного быстрее функций с L и d_0 в аргументах. Главный вклад в интеграл

имеем в области, где аргументы быстро осциллирующих функций принимают стационарные значения. Таким образом, нужно проанализировать выражения вида

$$\rho x + \beta z = f(\rho). \quad (9.3.28)$$

Используя формулу (8.4.6), получим

$$\frac{df}{d\rho} = x - \frac{\rho}{\beta} z \quad (9.3.29)$$

и

$$\frac{d^2 f}{d\rho^2} = -\frac{n_2^2 k_0^2}{\beta^3} z. \quad (9.3.30)$$

Требование стационарности фазы $df/d\rho = 0$ приводит к условию

$$\frac{\rho_s}{\beta_s} = \frac{x}{z} = \operatorname{tg} \alpha_s. \quad (9.3.31)$$

Разлагая $f(\rho)$ в ряд Тейлора в окрестности стационарной точки, получаем

$$\begin{aligned} \int_0^\infty e^{-i(\rho x + \beta z)} d\rho &\approx e^{-i(\rho_s x + \beta_s z)} \times \\ &\times \int_{-\infty}^\infty e^{i(n_0^2 k_0^2 / 2\beta_s^3) z (\rho - \rho_s)^2} d(\rho - \rho_s) = \\ &= (1+i) \sqrt{\pi} \frac{\beta_s^{3/2}}{n_2 k_0 \sqrt{z}} e^{-i(\rho_s x + \beta_s z)}. \end{aligned} \quad (9.3.32)$$

Для оценки интеграла была использована формула (2.3.23). В (9.3.27) необходимо было вычислить лишь интеграл (9.3.32). Разложим \sin и \cos по экспоненциальным функциям. В разложении необходимо рассмотреть только ту его часть, которая содержит функции вида (9.3.28). Экспоненциальные функции с аргументами $(\beta z - \rho x)$ дают незначительные вклады при x и z больше нуля. Множители при экспоненциальных функциях можно рассматривать как постоянные в точке $\rho = \rho_s$ и их можно вынести из-под интеграла. Окончательное интегрирование

выполнено в (9.3.32). Таким образом, имеем

$$F_y = - \frac{[a(1+i)k_0\rho_s \sqrt{\beta_s} \sqrt{2\omega\mu_0 P} (n_1^2 - n_2^2) \times \\ \times \cos \kappa_0 d_0 \cos \sigma_s d_0 e^{i\rho_s d_0} e^{i\Gamma_s L/2}]}{\left[2n_2 \sqrt{\pi} \sqrt{\left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0} \right) \times} \right.} \times \\ \left. \times (\rho_s \cos \sigma_s d_0 + i\sigma_s \sin \sigma_s d_0) \right]} \times \frac{\sin \Gamma_s \frac{L}{2}}{\Gamma_s} \frac{e^{-in_2 k_0 (x \sin \alpha_s + z \cos \alpha_s)}}{\sqrt{z}}, \quad (9.3.33)$$

где

$$\rho_s = n_2 k_0 \sin \alpha_s \quad (9.3.34)$$

и

$$\beta_s = n_2 k_0 \cos \alpha_s, \quad (9.3.35)$$

а Γ_s получается из (9.3.16) заменой β на β_s . Выражение для составляющей поля в дальней зоне указывает на несколько интересных особенностей. Видно, что поле распространяется примерно так же, как плоская волна под углом α_s , под которым поле приходит в точку x, z из источника в области нерегулярного участка ¹⁾. Для поля в дальней зоне источник ограниченной протяженности подобен точечному источнику, расположенному в начале координат. Последний множитель выражения (9.3.33) представляет собой цилиндрическую волну. Это становится более очевидным, если ввести в соответствии с (9.3.31)

$$x = r \sin \alpha_s \quad (9.3.36)$$

и

$$z = r \cos \alpha_s. \quad (9.3.37)$$

Тогда получается соотношение

$$\sqrt{\beta_s} \frac{e^{-in_2 k_0 (x \sin \alpha_s + z \cos \alpha_s)}}{\sqrt{z}} = \sqrt{n_2 k_0} \frac{e^{-in_2 k_0 r}}{\sqrt{r}}. \quad (9.3.38)$$

Появление здесь вместо сферической волны цилиндрической связано с двумерной геометрией волновода. Наиболее

¹⁾ Из этого примера видно, что суперпозиция стоячих волн излучения в интегральном разложении дает распространяющуюся волну.

важным множителем в формуле (9.3.33) является

$$G = \frac{\sin \Gamma_s \frac{L}{2}}{\Gamma_s}. \quad (9.3.39)$$

Известно, что он является приближением дельта-функции для возрастающих значений L . Чтобы величина G/π была хорошей аппроксимацией дельта-функции, необходимо, чтобы L было много больше длины волны λ . Это условие не противоречит условию (9.3.25), при котором было получено поле в дальней зоне. Множитель (9.3.39) определяет диаграмму направленности излучения участка плоского волновода с синусоидальным возмущением стенки. Главный лепесток этой диаграммы становится уже с увеличением L . Направление главного лепестка в пространстве задается формулой

$$\Gamma_s = \theta - (\beta_0 - \beta_s) = 0. \quad (9.3.40)$$

Поэтому основное излучение происходит только в направлении, определяемом из формулы (9.3.35):

$$\cos \alpha_s = \frac{\beta_0 - \theta}{n_2 k_0}. \quad (9.3.41)$$

Излучение направлено перпендикулярно поверхности волновода при $\theta = \beta_0$ и по касательной к его поверхности при $\theta = \beta_0 - n_2 k_0$. Другими словами, излучение перпендикулярно поверхности при $\beta_s = 0$ и приближается к касательному с увеличением β_s . При $\beta_s = n_2 k_0$ излучение главного лепестка параллельно поверхности волновода. Большие значения постоянных распространения мод излучения невозможны. При $\beta_s > n_2 k_0$ получаем направляемые моды. Данное рассмотрение достаточно ясно показывает, каким образом поле излучения становится все более направленным вперед и как его направление достигает своего предела — направления, параллельного оси волновода, которое сливается с направляемой модой. При рассмотрении направления излучения лепестка диаграммы направленности и изменения этого направления предполагалось, что можно непрерывно перестраивать механическую частоту θ . Точность приближения в формуле (9.3.22) возрастает при увеличении угла α_s , определенного в (9.3.41).

Читатель может удивиться, что обсуждению плоского волновода с синусоидальными искажениями стенки отве-

дено так много места. К тому же совершенно невероятно, чтобы такой случай мог иметь место на практике, если не позаботиться специально о создании волновода с подобной геометрией. Но все недоразумения исчезают, если функцию, описывающую любое произвольное искажение стенки, разложить в ряд или интеграл Фурье по синусоидальным функциям. Мы видели, что данное синусоидальное искажение стенки связывает падающую направляемую моду только с одной направляемой модой или только с очень узкой областью непрерывного спектра мод излучения. Таким образом, наше рассмотрение применимо к каждой фурье-компоненте разложения произвольной функции искажения стенки. При определении потерь на излучение для произвольной деформации стенки введем фурье-компоненту $F^2 d\theta$ вместо амплитуды a^2 и будем интегрировать (9.3.19) по интервалу (9.3.12). При этом выражения (9.2.38) и (9.2.39) подставляются в формулу (9.3.14). При $\Gamma = 0$ из (9.3.16) получим $d\beta = -d\theta$. Величина $|F|^2$ (и $|H|^2$) называется спектром мощности функции искажения стенки. Этот термин заимствован из анализа Фурье электрических сигналов. Квадрат абсолютного значения коэффициента разложения Фурье изменяющегося во времени сигнала пропорционален плотности мощности при данной частоте. Ограничиваясь для простоты искажением стенки только с одной стороны волновода, $h(z) = 0$, из формул (9.2.38), (9.2.39) и (9.3.14), получим

$$\frac{1}{L} \frac{\Delta P}{P} = \int_{-n_2 k_0}^{n_2 k_0} I(\beta) |F(\beta_0 - \beta)|^2 L d\beta, \quad (9.3.42)$$

где

$$I(\beta) = (n_1^2 - n_2^2)^2 \frac{\rho k_0^4 \cos^2 \kappa_0 d_0}{4\pi \left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0} \right)} \times \\ \times \left[\frac{\cos^2 \sigma d_0}{\rho^2 \cos^2 \sigma d_0 + \sigma^2 \sin^2 \sigma d_0} + \frac{\sin^2 \sigma d_0}{\rho^2 \cos^2 \sigma d_0 + \sigma^2 \sin^2 \sigma d_0} \right]. \quad (9.3.43)$$

Результаты, полученные при рассмотрении влияния строго синусоидального искажения стенки, помогают интуитивно предсказывать влияние более сложного искажения стенки,

которое всегда можно представить в виде интеграла Фурье. Если спектр функции искажения стенки известен, можно вычислить потери на излучение с помощью формулы (9.3.42). Обычно спектр мощности функции искажения стенки не известен, поэтому необходимо изучить некоторые модели, чтобы понять и оценить ожидаемые потери. Вместо изучения моделей с различными «спектрами мощности» мы в следующем разделе воспользуемся статистическим подходом.

С помощью теории возмущений было показано, что потери на излучение, испытываемые направляемой модой, можно вычислить даже для достаточно длинных участков волноводов. Теперь покажем, что аналогичный подход можно использовать для изучения взаимодействия между направляемыми модами. Нельзя, конечно, непосредственно распространить формулу для потерь (9.3.10) на (9.3.22). Потери по формуле (9.3.10) не зависят линейно от L , поэтому из нее нельзя определить потери на единицу длины. Физические соображения, использованные для получения соотношения (9.3.22) из (9.3.19), неприменимы в данном случае к направляемым модам, поскольку мощность, переходящая от падающей моды к паразитной моде, остается в волноводе и продолжает распространяться, взаимодействуя с падающей модой. Существуют определенные закономерности связи направляемых мод, вызванной строго синусоидальным искажением стенки, которые можно использовать для получения решения, справедливого за пределами применимости теории возмущений. Выше мы видели [см. формулу (9.3.7)], что при синусоидальном искажении стенки эффективно связаны только две направляемые моды. Поэтому пренебрежем влиянием всех других направляемых мод и поля излучения и будем считать, что взаимодействуют только две направляемые моды, связанные синусоидальным искажением стенки. Обозначая амплитудные коэффициенты этих двух мод через c_0 и c_1 , из уравнения (9.2.6) получим

$$\frac{\partial^2 c_0}{\partial z^2} - 2i\beta_0 \frac{\partial c_0}{\partial z} = -c_0 F_{00} - c_1 F_{10} \quad (9.3.44)$$

и

$$\frac{\partial^2 c_1}{\partial z^2} - 2i\beta_1 \frac{\partial c_1}{\partial z} = -c_0 F_{01} - c_1 F_{11}. \quad (9.3.45)$$

Эта система уравнений применима только в особом случае синусоидально изменяющейся функции искажения стенки. Здесь диагональные члены F_{00} и F_{11} изменяются синусоидально в зависимости от z . Они представляют собой синусоидально изменяющиеся поправки к постоянной распространения для регулярного волновода, стенка которого имеет среднее положение. Так как рассматривается незначительное отклонение стенки, то этими диагональными членами можно пренебречь. Предположим, что обе моды распространяются в положительном направлении оси z . В отсутствие связи коэффициенты c_0 и c_1 являются постоянными. Для слабой связи, которая рассматривается в данном случае, c_0 и c_1 изменяются на длине волны очень незначительно. Поэтому можно пренебречь вторыми производными коэффициентов c_0 и c_1 по сравнению с много большими произведениями первых производных на 2β . Пренебрегая указанными малыми членами, получим следующую упрощенную систему уравнений [сравните с уравнениями связанных волн (2.7.15) и (2.7.16)]:

$$\frac{\partial c_0}{\partial z} = -\kappa c_1 \quad (9.3.46)$$

и

$$\frac{\partial c_1}{\partial z} = \kappa c_0. \quad (9.3.47)$$

Коэффициенты связи в (9.3.46) и (9.3.47) можно определить из формулы (9.2.7). Однако проще воспользоваться формулами (9.2.34), (9.3.8) и (9.3.9) при $L = z$. Учитывая, что коэффициент c_0 равен единице в формуле (9.2.34), путем дифференцирования по z и сравнения с (9.3.47), получим

$$\kappa = -(n_1^2 - n_2^2) \frac{ak_0^2 \cos \kappa_0 d_0 \cos \kappa_1 d_0}{2 \sqrt{\left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0}\right) \left(\beta_1 d_0 + \frac{\beta_1}{\gamma_1}\right)}}. \quad (9.3.48)$$

Члены типа $\cos \kappa d_0$ можно исключить с помощью формулы (8.6.16). Коэффициент связи является вещественной величиной. Коэффициент связи для соотношения (9.3.46) получается заменой c_0 на c_1 в (9.2.34). Как видно из формулы (9.3.48), коэффициент κ инвариантен относительно перестановки индексов 0 и 1. Изменение в знаке в (9.3.46) можно получить, рассматривая коэффициенты Фурье

в формулах (9.2.35) и (9.3.8). При выводе равенства (9.3.8) мы пренебрегли членом вида

$$\frac{\sin x \frac{L}{2}}{x}, \quad (9.3.49)$$

где $x = \theta + (\beta_0 - \beta_1)$, поскольку x конечная величина. Однако после перестановки индексов 0 и 1 этот отброшенный член становится существенным, а членом, содержащим $x = \theta - (\beta_0 - \beta_1)$, можно пренебречь. Два оставленных члена имеют противоположные знаки, поскольку они получены из разложения функции $\sin \theta z$ на экспоненциальные компоненты. Данное рассмотрение объясняет различие знаков в формулах (9.3.46) и (9.3.47).

Исключая c_1 из формул (9.3.46) и (9.3.47), получим

$$\frac{d^2 c_0}{dz^2} + \kappa^2 c_0 = 0 \quad (9.3.50)$$

и аналогично

$$\frac{d^2 c_1}{dz^2} + \kappa^2 c_1 = 0. \quad (9.3.51)$$

Так как предполагалось, что c_0 — амплитудный коэффициент падающей моды, причем $c_0 = 1$ при $z = 0$, то имеем решение

$$c_0 = \cos \kappa z \quad (9.3.52)$$

и

$$c_1 = \sin \kappa z. \quad (9.3.53)$$

Для малых значений κz из формулы (9.3.53) при $z = L$ получаем

$$\frac{\Delta P}{P} = |c_1|^2 = \kappa^2 L^2, \quad (9.3.54)$$

что хорошо согласуется с (9.3.10).

Решения (9.3.52) и (9.3.53) справедливы за пределами применимости теории возмущений. Они хорошо описывают взаимодействие мощностей двух направляемых мод, которые связаны посредством синусоидального искажения стенки. Теперь стало ясно, почему приближенные относительные потери мощности (9.3.10) пропорциональны L^2 , а не L .

Две направляемые моды могут полностью обмениваться своей мощностью. Из формул (9.3.52) и (9.3.53) видно, что

мода с индексом 1 растет за счет моды с индексом 0. После прохождения расстояния

$$D = \frac{\pi}{2\kappa} \quad (9.3.55)$$

падающая мода полностью передает свою мощность моде с индексом 1.

Теория взаимодействия двух мод, связанных посредством синусоидального искажения стенки, не может быть распространена на случай произвольной функции искажения стенки. Хотя только одна фурье-компонента функции искажения стенки ответственна за связь двух мод, данное рассмотрение оказывается несправедливым, если синусоидальная функция (9.3.1) является одной из многих фурье-компонент полной функции искажения стенки. Существуют по крайней мере две причины, в силу которых простая теория связи неприменима в этом более сложном случае. Прежде всего отметим, что гораздо большее число мод связано с падающей направляемой модой. Моды теряют мощность и на излучение за счет фурье-компонент, соответствующих непрерывному спектру, и из-за связи с другими направляемыми модами. Таким образом, в общем случае нельзя ограничиваться рассмотрением связи только двух направляемых мод, а необходимо принимать во внимание и другие конкурирующие процессы. Однако, если даже потери на излучение были бы малы и кроме падающей существовала только одна направляемая мода, мы не смогли бы использовать простую теорию связи мод, поскольку направление передачи мощности зависит критически от фазового соотношения между двумя связанными модами и фазы синусоидальной функции изменения стенки. Фурье-компонента является функцией z , так как она получается интегрированием в точке z , в которой рассматриваются обе моды. Не только амплитуда a фурье-компоненты изменяет свое значение с увеличением z , но, что более важно, и ее фаза не имеет постоянного значения, необходимого для непрерывного обмена энергией. Для случайной функции $f(z)$ фаза фурье-компоненты изменяется случайно как функция z . Это заставляет направление передачи мощности между двумя модами меняться случайным образом то в одну, то в другую сторону, так что полная передача

мощности между модами невозможна. Если две моды даже и обмениваются своими мощностями, следует рассматривать это как чисто случайное событие.

Теорию связи между двумя направляемыми модами, вызванной строго синусоидальным искажением стенки, нельзя непосредственно применять при изучении потерь моды, вызванных случайным искажением стенки. Однако эта теория имеет важное значение при конструировании модовых ответвителей, осуществляющих передачу мощности между двумя направляемыми модами.

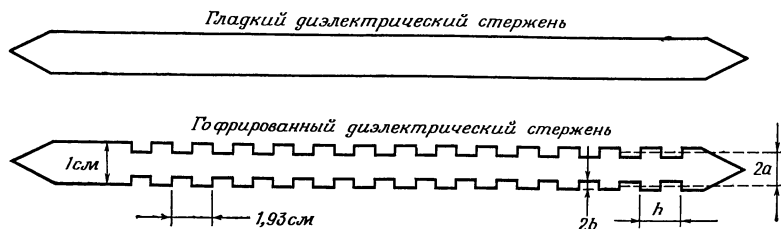
Теория связи мод, потери на излучение и направляющие свойства волновода с синусоидально искаженной стенкой проверялись экспериментально в миллиметровом диапазоне волн [99]. Все аспекты этой теории оказались в хорошем согласии с теорией возмущений. В эксперименте вместо плоского волновода использовался круглый волновод из тефлона, который возбуждался симметричной электрической TE_{01} -модой. Выбор волновода был продиктован экспериментальными удобствами. Плоский волновод является очень простым устройством с точки зрения его математического описания. Однако TE - и TM -моды плоского волновода трудно возбудить в реальном прямоугольном волноводе, так как для этого требуется, чтобы отсутствовали вариации поля вдоль оси y . С другой стороны, симметричные моды круглого волновода возбуждаются довольно легко. К тому же по структуре поля они очень похожи на TE - и TM -моды двумерного слоя с одним лишь исключением, что TE -мода низшего порядка круглого волновода не является основной, т. е. она не существует на очень низких частотах и в очень тонких стержнях. Результаты эксперимента сравнивались с результатами теории TE -мод круглого волновода (а не плоского волновода), которая в общих чертах и даже количественно почти совпадает с теорией плоского волновода.

Теория TE -мод круглого диэлектрического стержня изложена в [99]. Здесь обсуждаются только результаты эксперимента.

По причине простого технологического выполнения изготавливался тефлоновый стержень с периодическими прямоугольными пазами. Поскольку глубина пазов составляла 76 и 230 мкм, то трудно было бы изготовить на станке

стержень со строго синусоидальным изменением диаметра. Более того, нет необходимости воспроизводить точно форму синусоидального изменения диаметра. Мы уже видели, что фурье-компонента действует так же, как реальное синусоидальное искажение стенки при условии, что ее амплитуда и фаза не зависят от z .

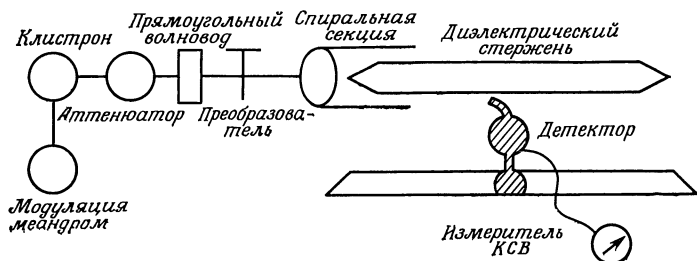
Измерения затухания основывались на сравнении потерь гладкого и гофрированного стержней. Два стержня,



Фиг. 9.3.1. Гладкий и гофрированный тефлоновые стержни, использованные в микроволновом эксперименте [99].

используемые в эксперименте, показаны на фиг. 9.3.1. Стержни подвешивались в воздухе на тонких нейлоновых нитях, которые не были заметной помехой при распространении волны. Мода TE_{01} возбуждалась на частоте ~ 50 ГГц клистроном и передавалась от прямоугольного волновода к круглому волноводу через преобразователь. Преобразователь был сконструирован таким образом, чтобы возбуждалась преимущественно TE_{01} -мода круглого полого металлического волновода. Для подавления нежелательных мод между диэлектрическим волноводом и преобразователем вставлялся отрезок круглого спирального волновода. Схема эксперимента приведена на фиг. 9.3.2. Пазы были разнесены так, что условие (9.3.7) удовлетворялось для TE_{01} - и TE_{02} -мод. Диаметр стержня был выбран так, чтобы на частоте выше 50 ГГц могла распространяться TE_{02} -мода. Частота отсечки была ниже этой частоты. В результате на частотах выше 50 ГГц наблюдалась связь двух направляемых мод, а на частотах ниже 50 ГГц наблюдалась связь с модами излучения. Было замечено, что оба типа связи исключают друг друга при строго синусоидаль-

ном возмущении стенки. Периодические пазы аппроксимируют реальные синусоидальные изменения стенки достаточно хорошо. Круглый стержень поддерживает, кроме указанных, и другие моды, например основную HE_{11} -моду, которая всегда присутствует в диэлектрическом стержне и не имеет отсечки. Однако эта мода не возбуждалась. Из семейства же TE_{0n} -мод были возможны только TE_{01} - и TE_{02} -моды.

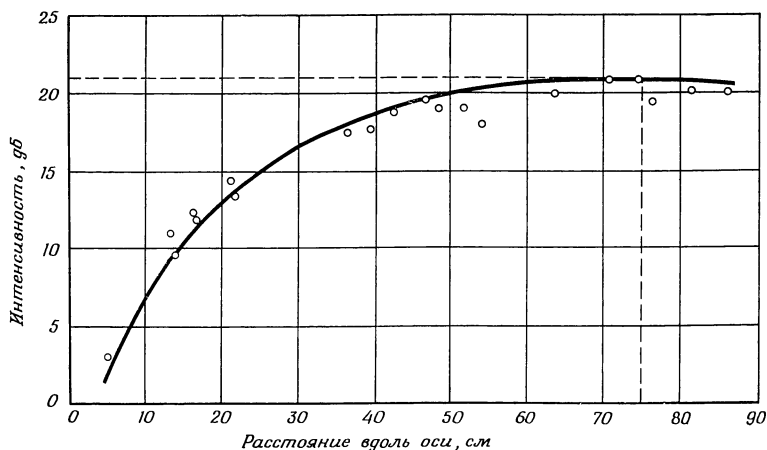


Ф и г. 9.3.2. Блок-схема микроволнового эксперимента [99].

Мощность, переносимую TE_{02} -модой, достаточно легко отделить от мощности TE_{01} -моды. Мода TE_{02} распространялась в режиме, очень близком к отсечке, и поэтому ее поле простиралось в поперечном направлении намного дальше от стержня, чем поле моды TE_{01} . Детектор состоял из L-образного отрезка прямоугольного волновода, который мог передвигаться на различные расстояния от поверхности исследуемого волновода, как показано на фиг. 9.3.2. Уже в нескольких миллиметрах от стержня мощность моды TE_{01} не обнаруживалась, а мощность моды TE_{02} принималась достаточно хорошо. Мощность миллиметровой волны, отобранная L-образным куском волновода, подавалась на диодный детектор. Мощность кликстрона модулировалась меандром. Модулированный сигнал усиливался низкочастотным усилителем, связанным с диодным детектором. При возбуждении TE_{01} -моды диэлектрического стержня диэлектрический стержень вставлялся в круглый металлический волновод. Такое сочленение достаточно эффективно возбуждало TE_{01} -моду диэлектрического волновода и значительно слабее TE_{02} -моду. Присутствие TE_{02} -моды

и других нежелательных мод малой мощности, возбуждаемых источником, несколько влияло на результаты измерений, однако это не было серьезной помехой при измерениях.

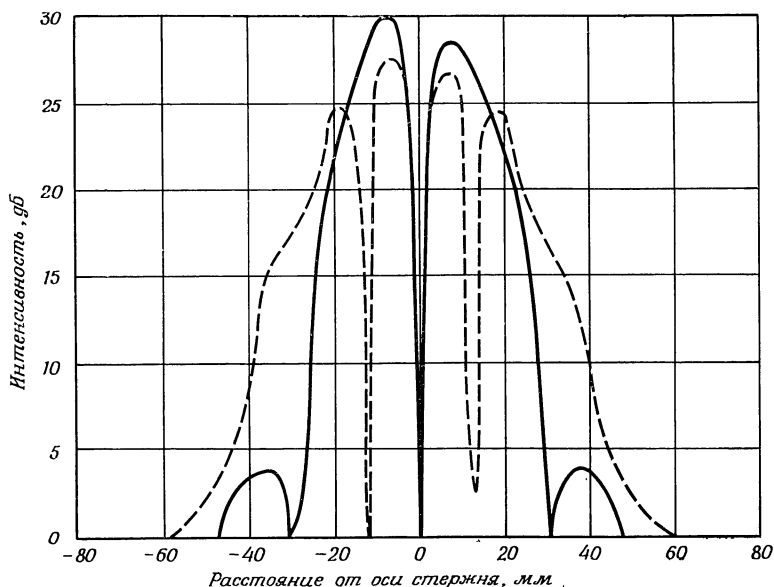
Возрастание мощности TE_{02} -моды как результат периодического возмущения стенки показано на фиг. 9.3.3. В данном случае пазы имели глубину 76 мкм, что делало



Ф и г. 9.3.3. Кривая нарастания мощности моды TE_{02} как результат связи между модами TE_{01} и TE_{02} . Глубина паза $7,6 \cdot 10^{-3}$ см. [99].

их едва видимыми. Тем не менее вся мощность переходила от падающей моды TE_{01} к моде TE_{02} после прохождения расстояния в 75 см. Это расстояние, рассчитанное по формуле (9.3.55) с величиной κ , полученной не из (9.3.48), а из соответствующей теории круглого стержня [99], составляет 80 см. Незначительное расхождение теории с экспериментом можно объяснить наличием механического допуска ± 10 мкм. Разброс экспериментальных данных вызван интерференцией мощности моды TE_{02} с другими нежелательными модами. Для подтверждения того, что измеренная мощность действительно была мощностью TE_{02} -моды, были сняты картины поля на концах гладкого и гофрированного стержней. Результаты показаны на

фиг. 9.3.4. Поперечное распределение поля моды, полученное на конце гладкого стержня (сплошная кривая), показывает характерную форму TE_{01} -моды с незначительной добавкой мощности TE_{02} -моды, которая возбуждалась источником. При замене гладкого стержня гофрированным распределение поля соответствовало распределению

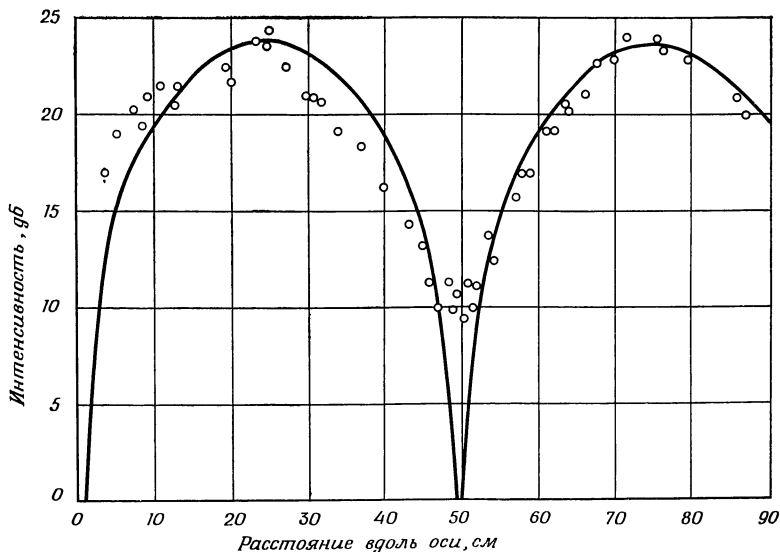


Фиг. 9.3.4. Картины распределения поперечного поля на конце гладкого стержня (сплошная кривая) и на конце гофрированного стержня (пунктирная кривая).

Сплошная кривая соответствует интенсивности моды TE_{01} с малой добавкой интенсивности моды TE_{02} . Пунктирная кривая соответствует моде TE_{02} [99].

поля TE_{02} -моды (пунктирная кривая). Гофрированный стержень вставлялся в металлический волновод настолько, что длина оставшегося свободного конца соответствовала длине D полного преобразования моды. Из фигуры видно, что преобразование моды осуществлялось практически полностью. Незначительная асимметрия и искажение поперечного профиля поля TE_{02} -моды, по-видимому, связаны с наличием паразитной моды, возбуждаемой источником.

Потери на излучение для пазов глубиной 76 мкм слишком малы, чтобы их можно было наблюдать. Поэтому пазы необходимо было углубить. Для изучения обмена мощностью между двумя направляемыми модами при наличии излучения был использован стержень с пазами глубиной 230 мкм. Результаты приведены на фиг. 9.3.5. Полученное полное преобразование на длине 25 см хорошо согласуется

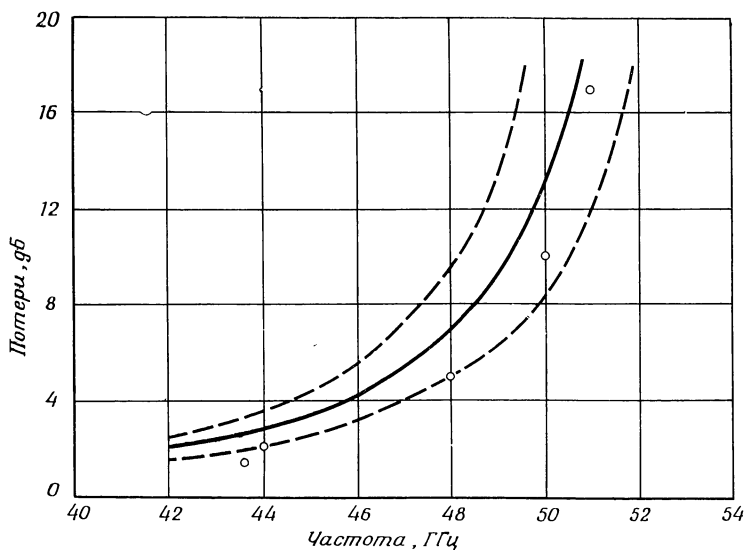


Фиг. 9.3.5. Интенсивность моды TE_{02} в гофрированном стержне. Обмен энергией проходит через два полных цикла. Глубина паза $2,3 \cdot 10^{-2}$ см [99].

с теорией. Таким образом, коэффициент связи, полученный из теории возмущений, предсказывает преобразование моды с такой высокой точностью. Полное преобразование моды осуществляется, как и предсказано теорией, на расстоянии в 45 длин волн. Фиг. 9.3.5 также убедительно демонстрирует справедливость введенного в теории допущения, что в результате синусоидального возмущения стенки в рассматриваемом случае взаимодействуют только две моды. Связь мод осуществляется при

прохождении двух полных циклов без потери мощности на излучение.

Понижение частоты ниже частоты отсечки для моды TE_{02} делает основной модой TE_{01} . Возможны, конечно, другие моды, особенно HE_{11} , но они не связаны с модой TE_{01} вследствие симметрии структуры волновода. Ниже частоты

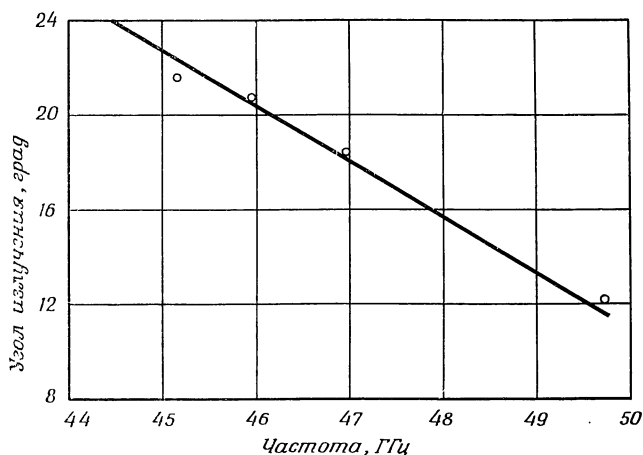


Ф и г. 9.3.6. Потери на излучение в зависимости от частоты.

Пунктирные кривые представляют результаты расчета потерь в предположении, что радиус волновода равен максимальному или минимальному значению на вершине или дне пазов. Сплошной кривой показаны результаты расчета с использованием среднего радиуса гофрированного стержня.

отсечки для моды TE_{02} падающая мода TE_{01} может быть связана только с модами излучения. Зависимость потерь на излучение моды TE_{01} от частоты для стержня с глубиной паза 230 мкм показана на фиг. 9.3.6. Точками отмечены результаты измерений. Сплошная кривая получена теоретически в предположении, что эффективным радиусом стержня является среднее значение радиусов до верхней и нижней границ пазов. Две пунктирные кривые соответствуют теоретическим результатам в случаях, когда в качестве радиуса волновода были использованы радиусы

сы до верхних и нижних границ пазов. Почти все экспериментальные точки лежат между этими двумя пунктирными кривыми, демонстрируя тем самым согласие теории и эксперимента. Точность экспериментальных данных не более чем 1 дБ. Высокие потери, имеющие место при частоте, близкой к частоте отсечки моды TE_{02} , далеко выходят



Ф и г. 9.3.7. Зависимость угла главного лепестка поля в дальней зоне от частоты [99].

—— теоретический график; ○ измеренные точки.

за рамки применимости простой теории возмущений. Эксперимент подтверждает возможность использования формулы (9.3.22) для произвольных значений потерь.

Зависимость угла главного лепестка диаграммы излучения от частоты дана на фиг. 9.3.7. Экспериментальные точки опять близки к расчетным.

Закончим этот раздел несколькими примерами численных расчетов характеристик излучения и преобразования мод нерегулярного волновода оптического диапазона. Выше было дано доказательство справедливости применения теории возмущений в миллиметровом диапазоне волн. Однако для целей данной книги важнее знать применимость теории для передачи света через оптические фибровые линии. Пусть длина волны света в свободном про-

странстве составляет

$$\lambda = 1 \text{ мкм.} \quad (9.3.56)$$

Эта длина волны, соответствующая инфракрасному спектру, выбрана в связи с простотой использования округленного числа.

Начнем с вопроса, какая амплитуда a синусоидальных искажений стенки, или фурье-компонента произвольного искажения стенки, вызовет потери мощности, составляющие 10%, в ближайшую направляемую моду после прохождения расстояния 1 см ($\Delta P/P = 0,1$, $L = 1$ см). Положим $\kappa_0 d = 3,0$. При $\lambda = 1$ мкм это приводит к полуширине слоя $d = 0,477$ мкм. При этих условиях возможно существование трех ТЕ-мод. Их постоянные распространения получаются из формулы (8.3.16) при $n_1 = 1,5$ и $n_2 = 1$:

$$\begin{aligned} \beta_0 d &= 4,336 \text{ (четная мода),} \\ \beta_1 d &= 3,831 \text{ (нечетная мода),} \\ \beta_2 d &= 3,051 \text{ (четная мода).} \end{aligned} \quad (9.3.57)$$

По формуле (9.3.10) вычислим требуемое значение амплитуды синусоидального возмущения, связывающего моду с индексом 0 с модой с индексом 2 ($\phi = 0$, толщина изменяется):

$$a = 2,5 \cdot 10^{-5} \text{ мкм} = 0,25 \text{ \AA}. \quad (9.3.58)$$

Это необычайно малое число показывает чрезвычайно высокие требования, которые связь мод накладывает на допуски к размерам волновода. Для $n_1 = 1,01$, $n_2 = 1$ и $k_0 d = 23$ получим полуширину волновода $d = 3,66$ мкм и постоянные распространения трех возможных ТЕ-мод

$$\begin{aligned} \beta_0 &= 23,199 \text{ (четная мода),} \\ \beta_1 &= 23,112 \text{ (нечетная мода),} \\ \beta_2 &= 23,002 \text{ (четная мода).} \end{aligned} \quad (9.3.59)$$

Требуемая амплитуда синусоидального возмущения толщины стенки ($\phi = 0$), которая вызывает потери мощности, составляющие 10%, за счет связи моды с индексом 0 с четной модой с индексом 2 при $L = 1$ см равна

$$a = 1,17 \cdot 10^{-2} \text{ мкм} = 117 \text{ \AA}. \quad (9.3.60)$$

Оба волновода допускают распространение трех направляемых волн. Их размеры для $\lambda = 1$ мкм существенно отличаются. При этих условиях волновод с малой разницей показателей преломления его сред менее чувствителен к изменению толщины, чем волновод с большим различием показателей преломления сред. Однако даже этот второй пример, представляющий собой типичное волокно в оболочке, показывает, что требования к допускам размеров являются очень высокими. Из данных примеров видно, что маловероятно, что в многомодовом оптическом волокне будет распространяться только одна направляемая мода. Неизбежные искажения стенки и другие виды искажений заставляют мощность падающей моды очень быстро рассеиваться в другие распространяющиеся моды. Волокно, может, конечно, работать также при таких условиях, когда обеспечено распространение одной направляемой моды. Если радиус или толщина волновода меньше критического значения для любой высшей моды, то может распространяться только эта основная мода. Случайные искажения стенки в этом случае не вызовут связь этой моды с другой волноводной модой, но приведут к потерям на излучение.

Оценим потери на излучение. Используя формулу (9.3.56), $n_1 = 1,01$, $n_2 = 1$, $k_0 d = 15$, получим $d = 2,39$ мкм. Необходимо найти такую амплитуду синусоидальных искажений толщины стенки ($\phi = 0$), которая вызывает потери мощности в 10% на длине $L = 1$ см. Из формулы (9.3.19) вычислим

$$a = 5,46 \cdot 10^{-2} \text{ мкм} = 546 \text{ \AA}. \quad (9.3.61)$$

Потери мощности, вызванные излучением, таким образом, являются при тех же допусках не такими большими, как потери на связь с другими направляемыми модами. Однако необходимо помнить, что амплитуда (9.3.61) относится только к одной фурье-компоненте. Если одна фурье-компонента может вызвать такие потери любой заданной направляемой моды, то большой диапазон механических частот внесет более существенный вклад в потери на излучение. Для определения полных потерь на излучение необходимо учесть потери, вызванные всеми фурье-компонентами. Вместо того чтобы следовать такому способу рас-

чета, обратимся к статистическому рассмотрению потерь на излучение, вызванных случайными искажениями стенки.

Между прочим из теории плоского волновода и из теории ТЕ-мод в круглом диэлектрическом волокне [99] получаются почти одинаковые числовые результаты. Такое совпадение для различных моделей указывает на то, что результаты, полученные для простой модели плоского волновода, применимы для других типов волноводов и, в частности, для круглого оптического волокна.

9.4. СЛУЧАЙНЫЕ ИСКАЖЕНИЯ СТЕНКИ

Основные формулы для потерь мощности, вызванных преобразованием направляемой моды в другие направляемые моды и моды излучения, могут быть использованы для вычисления средних потерь в плоском волноводе со случайными искажениями стенки.

Очевидно, что обе стенки слоя дают вклад в общие потери. Однако для простоты предположим в этом разделе, что одна из стенок волновода гладкая, $h(z) = 0$. Результаты, полученные при таком предположении, дают правильный порядок величины потерь мощности. Кроме того, они позволяют оценить влияние другой стенки. Если обе стенки слоя искажены случайным образом и функция искажения содержит широкий спектр частот механического возмущения с равномерным распределением «мощности», то потери удваиваются. Для узкого спектра функций искажения стенок с сильной корреляцией между этими функциями потери могут быть в 4 раза больше вычисленных в предположении, что у волновода одна из стенок гладкая.

Начнем с рассмотрения потерь мощности моды на преобразование в другие направляемые моды. Вклад в потери мод, согласно формуле (9.2.26), определяется как квадрат абсолютного значения их коэффициентов разложения. Средний вклад в потери, вносимый четной направляемой модой, согласно формуле (9.2.34), равен

$$\langle |c_{\mu e}|^2 \rangle = (n_1^2 - n_2^2)^2 \frac{L^2 k_0^4 \cos^2 \kappa_0 d_0 \cos^2 \kappa_\mu d_0}{4 \left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0} \right) \left(\beta_\mu d_0 + \frac{\beta_\mu}{\gamma_\mu} \right)} \langle |F|^2 \rangle. \quad (9.4.1)$$

Символ $\langle \rangle$ означает усреднение по ансамблю. Здесь подразумевается, что рассматривается много статистически

идентичных систем и их потери на преобразование мод. Среднее по значениям потерь всех систем дается выражением (9.4.1). Многие системы являются эргодическими, т. е. среднее, полученное по измерениям потерь от разных участков одного и того же волновода, такое же, как среднее, полученное по измерениям от различных волноводов. Для эргодических систем среднее по ансамблю можно заменить средним по измерениям потерь в разных точках вдоль волновода.

Представляющая интерес величина в (9.4.1) есть усредненный по ансамблю коэффициент Фурье

$$\langle |F|^2 \rangle = \frac{1}{L^2} \int_0^L dz \int_0^L dz' \langle f(z) f(z') \rangle e^{-i(\beta_0 - \beta_\mu)(z - z')}. \quad (9.4.2)$$

Среднее значение произведения функции $f(z)$ на такую же функцию называется автокорреляционной функцией от $f(z)$. Автокорреляционная функция не зависит от положения z для стационарного случайного процесса. Она зависит только от разности $z - z'$:

$$R(z - z') = \langle f(z) f(z') \rangle. \quad (9.4.3)$$

Интеграл можно упростить введением новых переменных

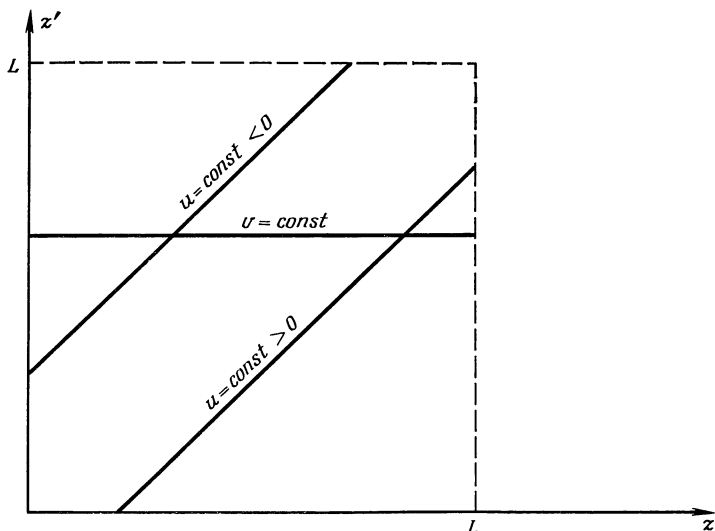
$$\begin{aligned} u &= z - z', \\ v &= z'. \end{aligned} \quad (9.4.4)$$

В результате интеграл (9.4.2) принимает вид

$$\begin{aligned} \langle |F|^2 \rangle &= \frac{1}{L^2} \int_0^L du R(u) e^{-i(\beta_0 - \beta_\mu)u} \int_0^{L-u} dv + \\ &+ \frac{1}{L^2} \int_{-L}^0 du R(u) e^{-i(\beta_0 - \beta_\mu)u} \int_{-u}^L dv. \end{aligned} \quad (9.4.5)$$

Новые пределы интегрирования представлены на фиг. 9.4.1. Линия $u = 0$ является диагональю квадрата, идущей от начала координат в верхний правый угол. Ниже этой диагонали u пробегает значения от $u = 0$ до $u = L$ (первый интеграл). Соответствующий интервал изменения v начинается от нижней стороны квадрата при $v = 0$ и заканчи-

вается в точке, где линия $u = \text{const}$ пересекает вертикаль $z = L$. В последней точке $v = z' = L - u$. Второй двойной интеграл охватывает область квадрата над диагональю. В верхнем левом углу $u = -L$, так что диапазон значений u охватывает интервал от $u = -L$ до $u = 0$ (первый интеграл второго слагаемого). Диапазон изменения v для



Ф и г. 9.4.1. Преобразование координат z, z' в координаты u и v .

этого слагаемого начинается с пересечения линии $u = \text{const}$ и вертикали $z = 0$. В этой точке $v = -u$. Значения v увеличиваются до тех пор, пока линия $u = \text{const}$ не достигнет вершины квадрата при $v = L$.

Проведя интегрирование по v , получим

$$\langle |F|^2 \rangle = \frac{1}{L^2} \int_{-L}^L (L - |u|) R(u) e^{-i(\beta_0 - \beta_u)u} du. \quad (9.4.6)$$

Это соотношение пока еще строгое. Можно, однако, ввести полезное приближение. Для чисто случайных процессов $R(u)$ уменьшается до нуля с увеличением расстояния u между двумя точками z и z' . Если бы это не выполнялось,

то деформация стенок волновода была бы систематической, что противоречит предположению о случайной деформации стенок. Допустим, что L намного больше расстояния, в пределах которого $R(u)$ еще существенно отличается от нуля. Поэтому можно пренебречь $|u|$ по сравнению с L . Далее, так как $R(u)$ стремится к нулю для больших значений u , можно увеличить диапазон интегрирования до ∞ . Получим приближение

$$\langle |F|^2 \rangle = \frac{1}{L} \int_{-\infty}^{\infty} R(u) e^{-i(\beta_0 - \beta_\mu)u} du. \quad (9.4.7)$$

Таким образом, среднее значение «спектра мощности» $|F|^2$ функции искажения стенки пропорционально преобразованию Фурье от корреляционной функции.

Величина (9.4.7) обратно пропорциональна L . Относительные потери мощности (9.4.1) на преобразование в μ -ю моду оказываются пропорциональными L . Это дает возможность определить потери на единицу длины. Данный результат интересен. Определяемые формулой (9.3.10) относительные потери мощности, вызванные синусоидальным искажением стенки, пропорциональны L^2 (см. разд. 9.3). Средние же потери, вызванные случайным искажением стенки, пропорциональны L . Природа процесса потерь полностью изменилась. Этот результат находится в согласии с нашим предположением относительно вкладов в потери различных фурье-компонент функции искажения стенки. В частности, в разд. 9.3 указывалось, что полный обмен мощностью между двумя направляемыми модами невозможен, если функция искажения стенки является случайной, потому что фаза преобразования Фурье становится случайной переменной. Настоящий результат доказывает, что полный обмен мощностью невозможен даже в случае, когда только одна фурье-компонента функции искажения стенки ответственна за связь между двумя модами. Потери мощности на преобразование в данную направляемую моду становятся по существу потерями на единицу длины. Фурье-компонента случайной функции обратно пропорциональна корню квадратному из длины, тогда как фурье-компонента строго периодической функции не зависит от длины,

Если необходимо получить более конкретную, чем такой общий результат, информацию, требуется знать корреляционную функцию искажения стенки. Эту функцию получить не проще, чем «спектр мощности» функции искажения стенки. Действительно, одна из них является фурье-преобразованием от другой. Однако известны некоторые общие черты корреляционной функции, так что можно использовать модельные представления и с их помощью изучать потери мощности. Как уже отмечалось, функция корреляции при больших значениях u должна стремиться к нулю. Она также должна быть симметричной функцией u , так как не имеет значения, расположена ли точка z' справа или слева от точки z . Корреляционная функция простейшего типа имеет вид

$$R(u) = A^2 e^{-(|u|/B)}. \quad (9.4.8)$$

Постоянная A является среднеквадратичным отклонением стенки волновода от прямолинейного направления, поскольку $R(0) = \langle f^2(z) \rangle = A^2$. Постоянная B определяет скорость уменьшения $R(u)$ при увеличении u и называется интервалом корреляции.

Из формул (9.4.7) и (9.4.8) получаем

$$\langle |F|^2 \rangle = \frac{2A^2}{BL} \frac{1}{(\beta_0 - \beta_\mu)^2 + \frac{1}{B^2}}. \quad (9.4.9)$$

Потери мощности на преобразование в μ -ю четную направляемую моду находятся из формул (9.4.1) и (9.4.9):

$$\langle |c_{\mu e}|^2 \rangle = \frac{LA^2 (n_1^2 - n_2^2)^2 k_0^4 \cos^2 \kappa_0 d_0 \cos^2 \kappa_\mu d_0}{2B \left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0} \right) \left(\beta_\mu d_0 + \frac{\beta_\mu}{\gamma_\mu} \right) \left[(\beta_0 - \beta_\mu)^2 + \frac{1}{B^2} \right]}. \quad (9.4.10)$$

Потери на преобразование в нечетные направляемые моды получаются заменой $\cos \kappa_\mu d_0$ на $\sin \kappa_\mu d_0$. Полные потери, вызванные всеми направляемыми модами, являются суммой вкладов потерь всех четных и нечетных направляемых мод.

Средние относительные потери на излучение получают путем усреднения по ансамблю (9.3.42) с использо-

ванием формул (9.3.43) и (9.4.9)

$$\begin{aligned} \left\langle \frac{\Delta P}{P} \right\rangle = 2\alpha L = L \frac{A^2 (n_1^2 - n_2^2)^2 k_0^4 \cos^2 \kappa_0 d_0}{2\pi B \left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0} \right)} \times \\ \times \int_{-n_2 k_0}^{n_2 k_0} \frac{\rho}{(\beta_0 - \beta)^2 + \frac{1}{B^2}} \left[\frac{\cos^2 \sigma d_0}{\rho^2 \cos^2 \sigma d_0 + \sigma^2 \sin^2 \sigma d_0} + \right. \\ \left. + \frac{\sin^2 \sigma d_0}{\rho^2 \sin^2 \sigma d_0 + \sigma^2 \cos^2 \sigma d_0} \right] d\beta. \quad (9.4.11) \end{aligned}$$

Выражение (9.4.11) дает возможность определить величину потерь мощности на единицу длины 2α , которая, согласно (9.3.22), дает полные потери на излучение (α — коэффициент, определяющий спад амплитуды).

Величина (9.4.11) должна вычисляться численными методами. Рассмотрим два различных частных случая. Начнем с рассмотрения случая, когда

$$n_1 = 1,5, \quad n_2 = 1. \quad (9.4.12)$$

Удобно ввести безразмерную величину

$$\frac{d_0^3}{A^2 L} \frac{\Delta P}{P}. \quad (9.4.13)$$

Это выражение для нормированных относительных потерь мощности имеет то преимущество, что в нем исключена зависимость от среднеквадратичного отклонения A . Зависимость нормированных потерь от нормированного интервала корреляции B/d_0 приведена на фиг. 9.4.2—9.4.4 для нескольких значений параметра $k_0 d_0$.

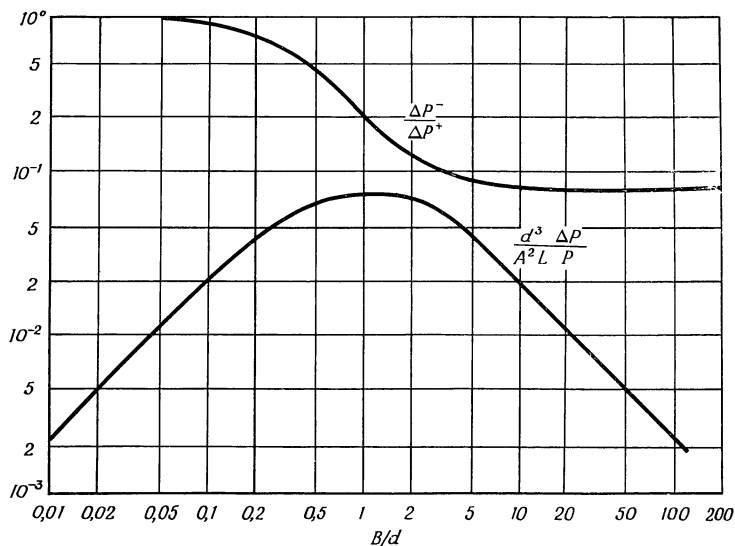
На этих фигурах приведены также графики отношения рассеянных назад и вперед мощностей $\Delta P^-/\Delta P^+$. Число возможных направляемых мод увеличивается с ростом $k_0 d_0$. Фиг. 9.4.2 соответствует случаю, когда распространяется только падающая ТЕ-мода. Такой волновод поддерживает только основную моду. Следующие две фигуры даны для двух- и трехмодового волноводов. Относительные потери мощности на излучение показаны сплошной линией, а относительные потери мощности на преобразование в другие моды даны пунктирной линией,

Случай малой разности показателей преломления сред волновода часто встречается в оптических волокнах. Результаты для

$$n_1 = 1,01, \quad n_2 = 1 \quad (9.4.1')$$

приведены на фиг. 9.4.5—9.4.7.

Все представленные зависимости имеют одни и те же характерные особенности. Нормированные относительные

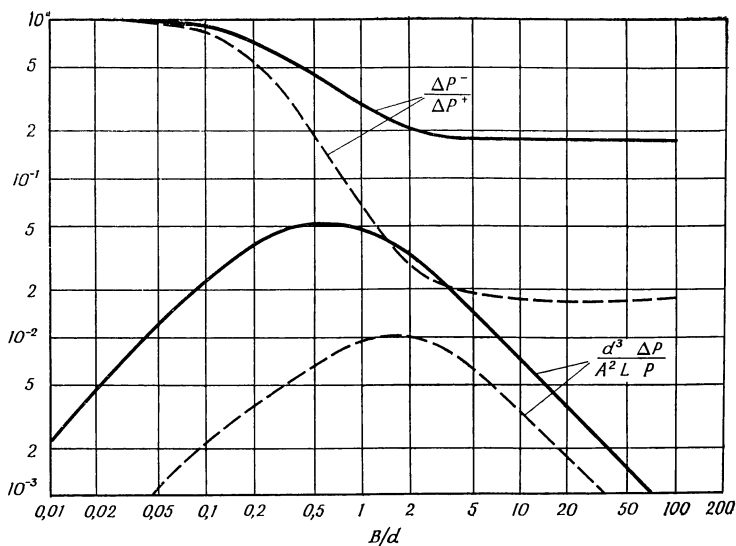


Ф и г. 9.4.2. Зависимость нормированных потерь на излучение $(d^3/A^2L)(\Delta P/P)$ и отношения мощностей $\Delta P^-/\Delta P^+$, рассеянных назад и вперед, от нормированного интервала корреляции B/d для $n_1 = 1,5$, $n_2 = 1$, $kd = 1,3$ [98].

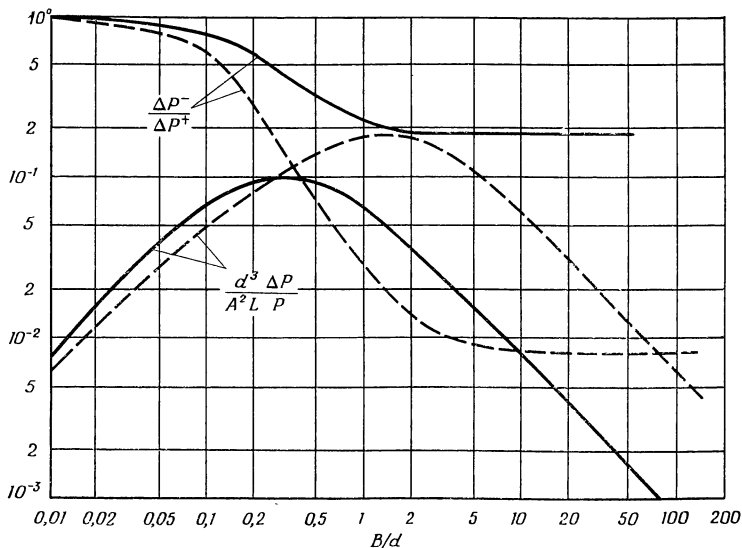
Волновод одномодовый.

d — полуширина сердцевин; A — среднеквадратичное отклонение одной стороны поверхности раздела сердцевина — оболочка; L — длина части волновода, дающей вклад в потери; n_1 — показатель преломления сердцевин; n_2 — показатель преломления оболочек; k — постоянная распространения в свободном пространстве.

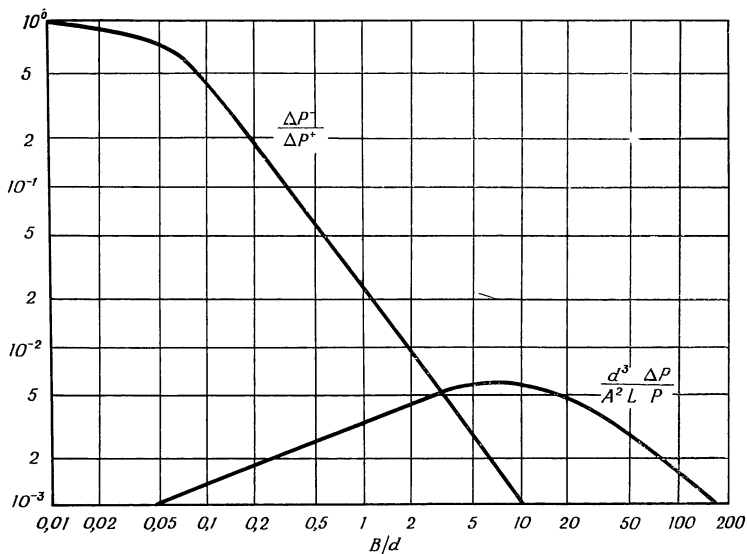
потери мощности имеют максимум вблизи $B/d_0 = 1$. Потери мощности оказываются наибольшими, если интервал корреляции приблизительно равен полутолщине волновода или длине волны, поскольку полутолщина волновода имеет порядок длины волны. Отношение рассеянных назад



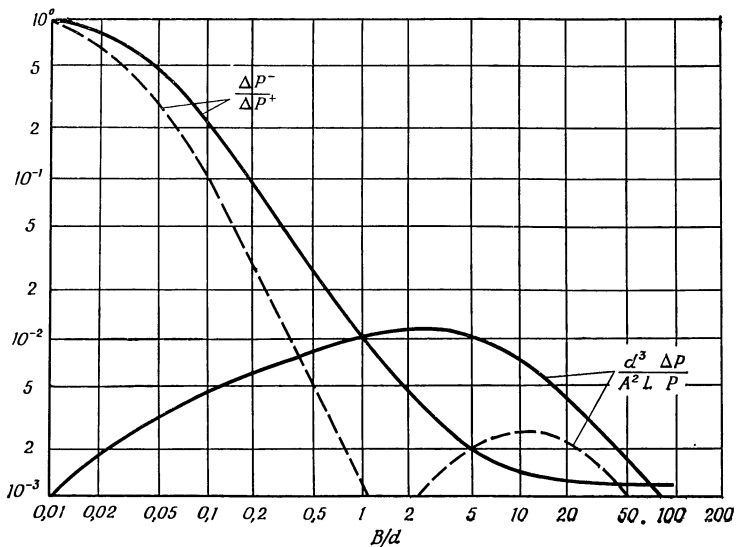
Ф и г. 9.4.3. То же, что и на фиг. 9.4.2, но для случая, когда могут распространяться две моды. Здесь $n_1 = 1,5$, $n_2 = 1$, $kd = 1,8$ [98].
 — — — мощность потерь на преобразование в паразитную моду; — потери на излучение.



Ф и г. 9.4.4. То же, что и на фиг. 9.4.2, но для случая, когда могут распространяться три моды. Здесь $n_1 = 1,5$, $n_2 = 1$, $kd = 3$ [98].
 — — — мощность потерь на преобразование в две паразитные волноводные моды; — потери на излучение.



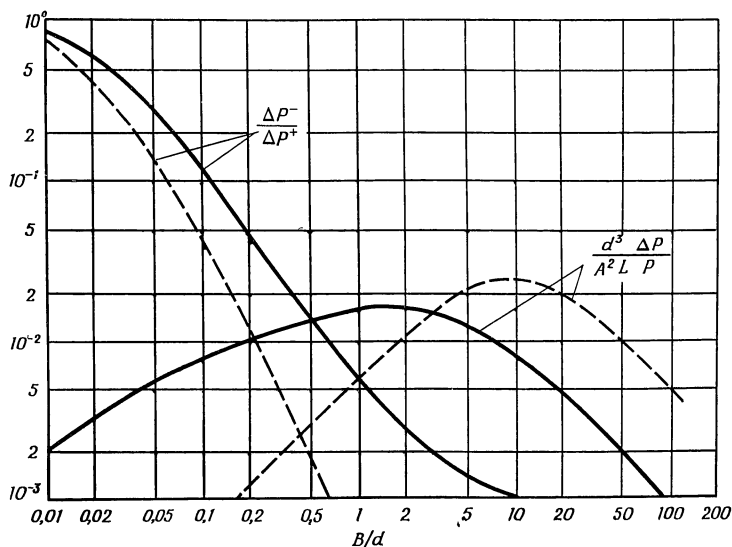
Ф и г. 9.4.5. То же, что и на фиг. 9.4.2. Параметры $n_1 = 1,01$, $n_2 = 1$, $kd = 8$ обеспечивают одномодовый режим работы [98]



Ф и г. 9.4.6. То же, что и на фиг. 9.4.2. Распространяются две моды. Здесь $n_1 = 1,01$, $n_2 = 1$, $kd = 15$ [98].

— — — мощность потерь на преобразование в паразитную моду;
 ————— потери на излучение.

и вперед мощностей $\Delta P^-/\Delta P^+$ становится равным единице для малых значений B и уменьшается к некоторому конечному значению (показано не на всех фигурах) для больших значений B . Аналогичная картина наблюдается в случае рассеяния света на малых частицах. Таким образом, можно считать, что интервал корреляции определяет размер



Ф и г. 9.4.7. То же, что и на фиг. 9.4.2, но могут распространяться три моды. Здесь $n_1 = 1,01$, $n_2 = 1$, $kd = 23$ [98].

частиц рассеивателей на поверхности волновода. Малый интервал корреляции соответствует частицам малого размера, а большой интервал корреляции — частицам большого размера. Эти частицы не должны пониматься буквально. Термин «частица» просто используется здесь для указания области искажения поверхности, в пределах которой имеет место корреляция. Рассеяние частиц, много меньших длины волны, как известно, является изотропным. Большие же частицы рассеивают преимущественно в прямом направлении.

Зависимость нормированных относительных потерь мощности от интервала корреляции определяется видом

функции корреляции. Приведенные зависимости справедливы только для экспоненциальной функции корреляции (9.4.8). Другие функции и комбинации различных функций также можно было бы (и следовало бы) использовать. Было рассмотрено несколько других функций, монотонно убывающих и симметричных относительно u . Получено, что форма кривой относительных потерь мощности слева от максимума, так же как положение и высота максимума, не зависят от вида функции корреляции. Форма кривой потерь справа от пика очень сильно зависит от вида функции корреляции.

Аналогичные кривые для относительных потерь мощности были получены для ТЕ- и ТМ-мод [99] и для основной моды HE_{11} круглого волновода [96]. Форма этих кривых хорошо согласуется с приведенными здесь кривыми для плоского волновода. Однако пики кривых, характеризующих потери круглого волновода, приблизительно в 4 раза превышают соответствующие пики для плоского волновода. Объяснение этому достаточно простое. Приведенные кривые справедливы, когда искажена только одна из двух сторон слоя. Если бы обе стенки были деформированы совершенно одинаково, то потери должны были бы быть в 4 раза выше. Именно этот случай имеет место для круглого стержня¹⁾. Согласие между теорией плоского волновода и более сложной теорией круглого стержня, таким образом, хорошее.

В заключение рассмотрим в этом разделе некоторые численные примеры с тем, чтобы получить представление о том, какие потери мощности можно ожидать от случайных искажений стенки волновода. Пусть показатели преломления сред волновода такие, как в (9.4.12), $n_0 d_0 = 1,3$ и может распространяться только основная направляемая мода. Кроме того, будем считать, что функция корреляции принимает наихудшее в смысле потерь мощности значение. Другие возможные допущения зависят от формы функции корреляции и поэтому не имеет смысла их рассматривать. Для $\lambda = 1$ мкм получим $d_0 = 0,207$ мкм. Нормированные

¹⁾ Деформация стенки круглого волновода, по-видимому, была чисто радиальная и не имела зависимости по углу. — *Прим. ред.*

потери в пике кривых на фиг. 9.4.2 составляют величину

$$\frac{d^3}{A^2 L} \frac{\Delta P}{P} = 7 \cdot 10^{-2}. \quad (9.4.15)$$

Опять оценим величину среднеквадратичного отклонения A , при котором потери на излучение составляют 10% после прохождения волной нерегулярного участка волновода длиной $L = 1$ см. В результате получим

$$A = 1,12 \cdot 10^{-3} \text{ мкм} = 11,2 \text{ \AA}. \quad (9.4.16)$$

Требуемый относительный допуск составит

$$\frac{A}{d} = 5,4 \cdot 10^{-3} = 0,54\%. \quad (9.4.17)$$

Из фиг. 9.4.3 видно, что потери на излучение преобладают над потерями на преобразование в одну паразитную направляемую моду. Для данного примера при $k_0 d_0 = 1,8$ и $\lambda = 1$ мкм полуширина волновода $d = 0,286$ мкм. Нормированные относительные потери на излучение, равные 0,05 при $B/d_0 = 1$, вызывают потери на излучение в 10% для $L = 1$ см, если

$$A = 2,16 \cdot 10^{-3} \text{ мкм} = 21,6 \text{ \AA}. \quad (9.4.18)$$

Требуемый относительный допуск составит

$$\frac{A}{d_0} = 7,6 \cdot 10^{-3} = 0,76\%. \quad (9.4.19)$$

Добавление потерь направляемых мод не изменит существенно этих результатов.

Потери на излучение, представленные на фиг. 9.4.4, несколько меньше, чем потери на преобразование в две паразитные моды. При $k_0 d_0 = 3$ и $\lambda = 1$ мкм полуширина волновода $d_0 = 0,477$ мкм. Объединение нормированных потерь на излучение и на преобразование при $B/d_0 = 1$ дает величину, равную 0,25. Среднеквадратичное отклонение, приводящее к потерям мощности, составляющим 10% при $L = 1$ см, равно

$$A = 2,08 \cdot 10^{-3} = 20,8 \text{ \AA}. \quad (9.4.20)$$

Требуемый относительный допуск составляет

$$\frac{A}{d_0} = 4,35 \cdot 10^{-3} = 0,44\%. \quad (9.4.21)$$

Если провести сравнение, то увидим, что требуемый относительный допуск почти одинаков во всех трех рассмотренных случаях.

Далее рассмотрим плоский волновод, у которого разница показателей преломления (9.4.14) его сред составляет 1%. Для основной моды (фиг. 9.4.5) при $k_0 d_0 = 8$ с $d_0 = 1,28$ мкм и для потерь мощности в пике при $B/d_0 = 10$, составляющих $6 \cdot 10^{-3}$, получим

$$A = 5,9 \cdot 10^{-2} \text{ мкм} = 590 \text{ \AA}. \quad (9.4.22)$$

Требуемый относительный допуск составляет

$$\frac{A}{d_0} = 4,6 \cdot 10^{-2} = 4,6\%. \quad (9.4.23)$$

Случай, когда кроме падающей моды, могут распространяться две паразитные направляемые моды, представлен на фиг. 9.4.7. Для $k_0 d_0 = 23$ и $d_0 = 3,66$ мкм полные потери на излучение и преобразование моды составляют $3,4 \cdot 10^{-2}$, и требуемый допуск, который получается из обычных условий, равен

$$A = 0,12 \text{ мкм} \quad (9.4.24)$$

или требуемый относительный допуск

$$\frac{A}{d_0} = 3,3 \cdot 10^{-2} = 3,3\%. \quad (9.4.25)$$

Требуемые относительные допуски для волновода с малой разностью показателей преломления сред также почти одинаковы для случаев, когда волновод направляет одну моду или три моды. Требуемый допуск для такого волновода намного мягче, чем для волновода с большой разностью показателей преломления сред. Из этого следует, что волновод с большой разностью показателей преломления сред более чувствителен к потерям, вызываемым случайным искажением стенки, чем волновод с малой разностью показателей преломления сред.

В качестве примера рассмотрим случай волоконного волновода для передачи света на большие расстояния. Пусть показатели преломления сред волновода имеют значения, указанные в (9.4.14), и по волноводу распространяется основная мода при условиях, приведенных на фиг. 9.4.5. Единственное отличие данного примера от рас-

смотренного ранее и приводящего к (9.4.22) состоит в том, что потери на излучение теперь составляют 10 дБ/км. При этих условиях необходимо потребовать, чтобы среднеквадратичное отклонение стенки было порядка

$$A = 9 \cdot 10^{-4} \text{ мкм} = 9 \text{ \AA}. \quad (9.4.26)$$

Видно, что требуемые допуски при передаче света на большие расстояния оказываются чрезвычайно строгими даже для менее чувствительного волновода с малой разностью показателей преломления сред. Оценки оказались такими из-за того, что интервал корреляции функции искажения стенки взят порядка полуширины волновода. На практике же отклонения стенок тянутых стеклянных волокон имеют больший интервал корреляции, что приводит к меньшим потерям на излучение и преобразование. Одним из способов вычисления интервала корреляции является определение отношения мощностей прямого и обратного рассеяния света. Если это отношение равняется примерно единице, то интервал корреляции короткий. Преобладающее рассеяние вперед соответствует более длинному интервалу корреляции. Некоторые предварительные измерения дали потери мощности за счет рассеяния порядка 40 дБ/км [105]. Этот результат показывает, что допуски на отклонения стенок волновода оказались чрезвычайно малыми, если интервал корреляции был порядка ширины волновода. Однако более вероятно, что интервал корреляции был намного длиннее. Приведенная здесь модель не слишком поучительна, так как характер функции корреляции существенно влияет на величину потерь. Чтобы можно было теоретически предсказывать потери, необходимо больше знать о «спектре мощности» функции искажения стенки или же о корреляционной функции.

9.5. СКАЧКООБРАЗНЫЕ И ПЛАВНЫЕ ИЗМЕНЕНИЯ ТОЛЩИНЫ ПЛОСКОГО ВОЛНОВОДА

В разд. 9.2 коэффициенты разложения поля (8.5.16) были найдены из требования, чтобы компонента электрического поля была решением волнового уравнения. Этот метод был применен при исследовании полей поперечно-электрического типа. Подобный метод, конечно, может

быть использован и при исследовании полей других типов. Однако в общем случае процедура определения амплитудных коэффициентов является намного более сложной. Для иллюстрации этого рассмотрим волновое уравнение для составляющей H_y поперечно-магнитных полей. Волновое уравнение в форме (8.3.33) справедливо для ТМ-мод лишь когда диэлектрические среды волновода однородные. Это уравнение можно было использовать в разд. 8.3, так как его решения отыскивались отдельно в каждой однородной среде волновода и затем сшивались на границах раздела диэлектриков при $x = \pm d$.

Метод, развитый в разд. 9.2, не требует использования каких-либо граничных условий, но при этом волновое уравнение должно удовлетворяться везде, в том числе и на границах раздела сред диэлектриков. Соответствующее приведенное волновое уравнение для \mathbf{H} имеет вид

$$\nabla^2 \mathbf{H} + \frac{1}{\varepsilon} (\nabla \varepsilon) \times (\nabla \times \mathbf{H}) + \omega^2 \varepsilon \mu_0 \mathbf{H} = 0. \quad (9.5.1)$$

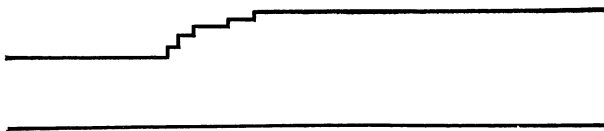
Вывод этого уравнения предоставляется читателю в качестве упражнения. Для поля поперечно-магнитного типа, единственной составляющей которого является H_y , в случае, когда диэлектрическая постоянная изменяется только в направлении координаты x , уравнение (9.5.1) принимает вид

$$\nabla^2 H_y - \frac{1}{\varepsilon} \frac{\partial \varepsilon}{\partial x} \frac{\partial H_y}{\partial x} + \omega^2 \varepsilon \mu_0 H_y = 0. \quad (9.5.2)$$

На поверхности раздела диэлектриков производная от диэлектрической проницаемости по координате x представляется дельта-функцией, а производная $\partial H_y / \partial x$ имеет скачок. Волновое уравнение для составляющей E_y поля поперечно-электрического типа не является таким сложным, так как формула (1.3.4) содержит лишь производную $\partial \varepsilon / \partial y$, но она в силу условия (8.3.1) обращается в нуль. Уравнение (9.5.2) можно решать методом, изложенным в разд. 9.2. Однако эта процедура достаточно сложна. Необходимо аппроксимировать ε быстро меняющейся, но непрерывной функцией. Другим возможным методом решения является использование локальных нормальных мод [113], о которых кратко упоминалось в начале разд. 9.2. Существуют и более простые методы, которые могут быть

использованы. Рассмотрение одного такого метода составляет предмет этого раздела.

Любую произвольную деформацию границы раздела сред можно аппроксимировать последовательностью малых скачков [101], как показано на фиг. 9.5.1. В пределе при бесконечно большом числе скачков и нулевой высоте каждого скачка последовательностью скачков можно аппроксимировать любую произвольную функцию. Преобразование моды многими скачками является суперпозицией полей, порождаемых каждым индивидуальным



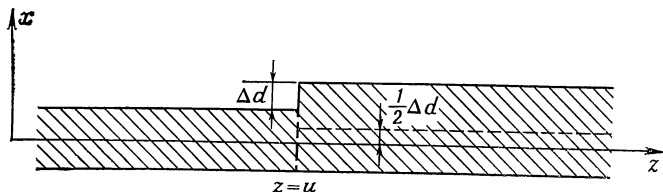
Ф и г. 9.5.1. Произвольная деформация границы раздела сердцевина — оболочка аппроксимируется последовательностью скачкообразных изменений.

скачком. Такой метод достаточно прост и удобен ¹⁾. Его можно применить к любому типу поля, в том числе и к более сложным полям круглого оптического волокна [96]. Справедливость метода подтверждена экспериментально. Кроме того, этот метод приводит к тем же результатам, что и метод разд. 9.2. Применим метод последовательных скачков к ТЕ- и ТМ-модам плоского волновода, а также рассмотрим возможность его применения к другим типам волноводов. Метод малых скачков также является приближенным методом. В дальнейшем будем фиксировать приближения по мере их появления.

Начнем с рассмотрения одного скачка, изображенного на фиг. 9.5.2. Пусть в волноводе распространяется четная ТЕ-мода низшего порядка. При прохождении скачка мода теряет мощность на преобразование в другие направляемые моды и на излучение. Для простоты ограничимся рас-

¹⁾ Этот метод получил название метода малых неоднородностей [59*]. Метод поперечных сечений [59*, 113], или, как называет его автор, метод локальных нормальных волн является развитием и математической формализацией метода малых неоднородностей. — *Прим. ред.*

смотрением случая, когда до и после скачка существует только одна направляемая мода. Поле в волноводе можно вычислить с помощью разложения по нормальным модам. Коэффициенты разложения находятся из условия непрерывности касательных составляющих электрического



Ф и г. 9.5.2. Схематическое изображение одного скачка поверхности раздела сердцевина — оболочка плоского волновода.

и магнитного полей E_y и H_x в области скачка при $z = u$. Это условие приводит к уравнению для компонент E_y

$$E_1^{(i)} + a_r E_1^{(r)} + \int_0^\infty q_e^{(-)}(\rho) E_{e1}^{(r)}(\rho) d\rho + \int_0^\infty q_0^{(-)}(\rho) E_{01}^{(r)}(\rho) d\rho = \\ = c_t E_2^{(t)} + \int_0^\infty q_e^{(+)}(\rho) E_{e2}^{(t)}(\rho) d\rho + \int_0^\infty q_0^{(+)}(\rho) E_{02}^{(t)}(\rho) d\rho \quad (9.5.3)$$

и для компонент H_x

$$H_1^{(i)} + a_r H_1^{(r)} + \int_0^\infty q_e^{(-)}(\rho) H_{e1}^{(r)}(\rho) d\rho + \int_0^\infty q_0^{(-)}(\rho) H_{01}^{(r)}(\rho) d\rho = \\ = c_t H_2^{(t)} + \int_0^\infty q_e^{(+)}(\rho) H_{e2}^{(t)}(\rho) d\rho + \int_0^\infty q_0^{(+)}(\rho) H_{02}^{(t)}(\rho) d\rho. \quad (9.5.4)$$

Разложения в левых частях этих уравнений представляют поля слева от скачка. Поле слева от скачка является суперпозицией падающей волны (индекс i сверху), отраженной направляемой моды (индекс r сверху) и отраженных четных и нечетных мод излучения. Постоянная a_r является коэффициентом отражения направляемой моды. Коэффици-

циенты $q_e^{(-)}$ и $q_0^{(-)}$ являются амплитудами отраженных четных и нечетных мод излучения. Индексом 1 отмечено поле в волноводе слева от скачка. Правые части уравнений представляют прошедшую направляемую моду и рассеянные вперед моды излучения. Индексом t отмечены моды, распространяющиеся в положительном направлении z . Постоянная c_t — коэффициент прохождения направляемой моды; $q_e^{(+)}$ и $q_0^{(+)}$ — амплитуды распространяющихся вперед четных и нечетных мод излучения. В разд. 8.6 коэффициенты разложения определялись из условия удовлетворения поля волновому уравнению. В данном разделе коэффициенты разложения определим из условия непрерывности поперечных составляющих поля в плоскости скачка.

Моды волновода слева от скачка взаимно ортогональны. Аналогичное условие справедливо для мод справа от скачка. Однако моды волновода слева и справа от скачка не ортогональны из-за разных размеров волновода по обе стороны от скачка. Но для малой высоты скачка имеет место приближенная ортогональность этих мод по обе стороны от скачка. С помощью условия ортогональности мод справа от скачка выделим коэффициент разложения $q_e^{(+)}$. Коэффициенты $q_e^{(-)}$ и $q_0^{(-)}$ малы. Можно пренебречь разницей между модами излучения по обе стороны скачка, но при этом необходимо помнить, что $E^{(t)}(\rho)$ зависит от $\exp(-i\beta z)$, а $E^{(r)}(\rho)$ — от $\exp(i\beta z)$. Определим $q_e^{(+)}$, пренебрегая амплитудным коэффициентом a_r , который пренебрежимо мал для малой величины скачка. Умножив (9.5.3) на

$$\frac{\beta}{2\omega\mu_0} E_{e2}^{(t)*}(\rho') \quad (9.5.5)$$

и затем проинтегрировав по x , с помощью формулы (8.4.11) при $z = u$ получим

$$q_e^{(+)}(\rho) = \frac{\beta}{2\omega\mu_0 P} \int_{-\infty}^{\infty} E_1^{(i)} E_{e2}^{(t)*}(\rho) dx + q_e^{(-)}(\rho) e^{2i\beta u}. \quad (9.5.6)$$

Коэффициент $q_e^{(+)}$ может быть получен также из (9.5.4). Выразим с помощью формулы (8.3.2) компоненты H_x ,

входящие в это уравнение, через E_y :

$$\begin{aligned} \beta_0 E_1^{(i)} - a_r \beta_0 E_1^{(r)} - \int_0^\infty \beta q_e^{(-)}(\rho) E_{e1}^{(r)}(\rho) d\rho - \\ - \int_0^\infty \beta q_0^{(-)}(\rho) E_{01}^{(r)}(\rho) d\rho = \beta_0 c_t E_2^{(t)} + \int_0^\infty \beta q_e^{(+)}(\rho) E_{e2}^{(t)}(\rho) d\rho + \\ + \int_0^\infty \beta q_0^{(+)}(\rho) E_{02}^{(t)}(\rho) d\rho. \quad (9.5.7) \end{aligned}$$

Умножая на

$$\frac{1}{2\omega\mu_0} E_{e2}^{(t)*}(\rho) \quad (9.5.8)$$

и интегрируя по x , с помощью формулы (8.4.11) при $z = u$ получим

$$q_e^{(+)}(\rho) = \frac{\beta_0}{2\omega\mu_0 P} \int_{-\infty}^\infty E_1^{(i)} E_{e2}^{(t)*}(\rho) dx - q_e^{(-)}(\rho) e^{2i\beta u}. \quad (9.5.9)$$

При этом коэффициентом отражения a_r снова пренебрегаем. При сложении (9.5.6) и (9.5.9) $q_e^{(-)}(\rho)$ исчезает и мы получаем

$$q_e^{(+)}(\rho) = \frac{\beta_0 + \beta}{4\omega\mu_0 P} \int_{-\infty}^\infty E_1^{(i)} E_{e2}^{(t)*}(\rho) dx \quad \text{при } z = u. \quad (9.5.10)$$

Вычитание (9.5.6) и (9.5.9) дает выражение для $q_e^{(-)}(\rho)$, которое очень похоже на (9.5.10), за исключением того что β заменяется на $-\beta$ и $E^{(t)*}(\rho)$ на $E^{(r)*}(\rho)$ (объясните, почему). Выражение для коэффициента разложения нечетных мод излучения получается из (9.5.10) заменой E_{e2} на E_{02} .

Предположим, что высота скачка Δd намного меньше, чем полуширина слоя d . Выражение для y -составляющей электрического поля мод излучения волновода справа от

скачка имеет вид [см. формулы (8.4.1) — (8.4.10)]

$$E_{e2}^{(t)*}(\rho, x') = e^{i\beta z} \begin{cases} C_e \cos \sigma x' & \text{для } |x'| < d_2, \\ C_e \left[\cos \sigma d_2 \cos \rho (|x'| - d_2) - \right. \\ \left. - \frac{\sigma}{\rho} \sin \sigma d_2 \sin \rho (|x'| - d_2) \right] & \text{для } |x'| > d_2. \end{cases} \quad (9.5.11)$$

Координата x' отсчитывается от пунктирной линии на фиг. 9.5.2. Разложим это выражение в ряд по степеням Δd и сохраним только первые два члена:

$$E_{e2}^{(t)*}(\rho) = E_{e1}^{(t)*}(\rho, x') + \frac{1}{2} \left(\frac{\partial E_{e2}^{(t)*}(\rho, x')}{\partial d_2} \right)_{d_2=d_1} \Delta d, \quad (9.5.12)$$

где

$$d_2 = d_1 + \frac{1}{2} \Delta d. \quad (9.5.13)$$

Координата x' связана с x соотношением

$$x' = x - \frac{1}{2} \Delta d. \quad (9.5.14)$$

Второй член в правой части (9.5.12) всегда является малой величиной первого порядка. Разность x и x' того же порядка. Замена x' на x в этом члене вызовет изменение Δd только на величину второго порядка малости, которой можно пренебречь в приближении первого порядка. Таким образом, из формулы (9.5.12) получаем

$$E_{e2}^{(t)*}(\rho) = E_{e1}^{(t)*}(\rho, x) - \frac{1}{2} \left(\frac{\partial E_{e1}^{(t)*}(\rho, x')}{\partial x'} \right)_{x'=x} \Delta d + \\ + \frac{1}{2} \left(\frac{\partial E_{e2}^{(t)*}(\rho, x)}{\partial d_2} \right)_{d_2=d_1} \Delta d. \quad (9.5.15)$$

Первый член в (9.5.15) не дает вклада в (9.5.10) в силу ортогональности мод. В результате имеем

$$q_e^{(+)}(\rho) = \Delta d \frac{\beta_0 + \beta}{8\omega\mu_0 P} \int_{-\infty}^{\infty} E_1^{(i)} \left[\left(\frac{\partial E_{e2}^{(t)*}}{\partial d_2} \right)_{d_2=d_1} - \left(\frac{\partial E_{e1}^{(t)*}}{\partial x'} \right)_{x'=x} \right] dx. \quad (9.5.16)$$

Пусть падающая мода является четной функцией от x . Мода излучения $E_e(\rho)$ также является четной функцией. Производная $\partial E_{e1}(\rho, x')/\partial x'$ от четной функции должна быть нечетной функцией. Произведение направляемой моды с производной моды излучения по x' есть нечетная функция, которая не вносит вклада в симметричный интеграл. Таким образом, соотношение (9.5.16) принимает вид

$$q_e^{(+)}(\rho) = \Delta d \frac{\beta_0 + \beta}{8\omega\mu_0 P} \int_{-\infty}^{\infty} E_1^{(i)} \left(\frac{\partial E_{e2}^{(t)*}}{\partial d_2} \right)_{d_2=d_1} dx. \quad (9.5.17)$$

Амплитуда C_e , согласно выражению (8.4.18), зависит от d . Однако компонента электрического поля ортогональна падающей направляемой моде, если она взята при $d_2 = d_1$ и $x' = x$, независимо от того, входит ли в формулу (9.5.11) коэффициент C_e или его производная. Член, содержащий производную от C_e , не дает вклада в (9.5.17) и можно считать C_e не зависящим от d_2 . Постоянная распространения β , так же как σ и ρ , не зависит от d_2 . Окончательно производная моды излучения принимает простой вид

$$\begin{aligned} & \left(\frac{\partial E_{e2}^{(t)*}}{\partial d_2} \right)_{d_2=d_1} = \\ & = e^{i\beta z} \begin{cases} 0 & \text{для } |x| < d, \\ -C_e \frac{n_1^2 - n_2^2}{\rho} k_0^2 \cos \sigma d_1 \sin \rho (|x| - d_1) & \text{для } |x| > d. \end{cases} \end{aligned} \quad (9.5.18)$$

Это простое выражение, полученное при отбрасывании части членов, является типичным для ТЕ-мод плоского волновода. Соответствующие выражения для моды HE_{11} круглого волокна намного сложнее [96]. Выражение

$$\sigma^2 - \rho^2 = (n_1^2 - n_2^2) k_0^2 \quad (9.5.19)$$

было использовано при упрощении (9.5.18).

Выражения (8.3.12), (8.3.18) и (9.5.18) использовались для получения из (9.5.17) выражения

$$q_e(\rho, u) = - \frac{(n_1^2 - n_2^2) k_0^2 \rho \cos \chi d \cos \sigma d e^{-i(\beta_0 - \beta)u} \Delta d}{2(\beta_0 - \beta) \sqrt{\pi |\beta| \left(\beta_0 d_0 + \frac{\beta_0}{\gamma_0} \right) (\rho^2 \cos^2 \sigma d + \sigma^2 \sin^2 \sigma d)}}. \quad (9.5.20)$$

Индекс сверху «+» здесь опущен, так как оба выражения $q_e^{(+)}$ и $q_e^{(-)}$ отличаются только знаком постоянной распространения ρ моды излучения; β_0 — постоянная распространения падающей направляемой моды.

Мы нашли амплитуду моды излучения, которая возбуждается в месте малого скачка плоского волновода. Используем полученный результат для вычисления потерь на излучение при искажении стенки более общего вида. Для этого рассмотрим Δd как бесконечно малый прирост в серии скачков, приближенно описывающих функцию $f(z)$. Можно записать

$$\Delta d = \frac{df}{dz} dz. \quad (9.5.21)$$

Каждый скачок последовательности порождает моды излучения. Скачки расположены при разных положениях $z = u$. В точке, расположенной вдали от возмущенной области, все моды излучения, которые возбуждены, складываются так, что для данной моды излучения, соответствующей ρ ($q = \hat{q} du$), можно записать

$$\begin{aligned} \bar{E}_e(\rho, x, z) &= \int_0^L \hat{q}_e(\rho, u) E_e(\rho, x, z) du = \\ &= E_e(\rho, x, z) \int_0^L \hat{q}_e(\rho, u) du. \end{aligned} \quad (9.5.22)$$

Черточка над составляющей поля означает суммарное поле от всех вкладов по всем скачкам. Выражение для моды можно вынести из-под интеграла, так как оно в любом случае не содержит зависимости от положения u на каждой ступеньке. Рассмотрение показывает, что эффективная амплитуда моды излучения, соответствующая ρ , представляется интегралом от (9.5.20). Записывая z вместо u , полу-

ЧИМ

$$q_e(\rho) = -\frac{n_1^2 - n_2^2}{2\sqrt{\pi|\beta|}} \times$$

$$\times \rho k_0^2 \int_0^L \frac{\cos \kappa d \cos \sigma de}{(\beta_0 - \beta) \sqrt{\left(\beta_0 d + \frac{\beta_0}{\gamma}\right) (\rho^2 \cos^2 \sigma d + \sigma^2 \sin^2 \sigma d)}} \left(\frac{df}{dz} - \frac{dh}{dz} \right) dz, \quad (9.5.23)$$

где

$$d = d_0 + f(z). \quad (9.5.24)$$

Интеграл $-i \int_0^z (\beta_0 - \beta) du$ здесь необходим, потому что β_0^0 является функцией от z , а приращение фазы составляет $-i (\beta_0 - \beta) du$. Аналогичным образом получим соответствующую амплитуду нечетных мод излучения

$$q_0(\rho) = \frac{n_1^2 - n_2^2}{2\sqrt{\pi|\beta|}} \times$$

$$\times \rho k_0^2 \int_0^L \frac{\cos \kappa d \sin \sigma de}{(\beta_0 - \beta) \sqrt{\left(\beta_0 d_0 + \frac{\beta_0}{\gamma}\right) (\rho^2 \sin^2 \sigma d + \sigma^2 \cos^2 \sigma d)}} \left(\frac{df}{dz} + \frac{dh}{dz} \right) dz. \quad (9.5.25)$$

Можно также провести аналогичные вычисления для ТМ-мод без дополнительных сложностей. Записывая p вместо q в выражениях для амплитуд ТМ-мод излучения, возбужденных направляемой четной ТМ-модой низшего порядка, для четных мод излучения получим

$$p_e(\rho) = -\frac{n_1^2 - n_2^2}{2\sqrt{\pi|\beta|}} \frac{n_2^2}{n_2^2} \times$$

$$\times \rho \int_0^L \frac{\sqrt{\gamma} (\beta_0 \beta \cos \sigma d + \gamma \sigma \sin \sigma d) \cos \kappa de}{(\beta_0 - \beta) \sqrt{\beta_0 \left[\gamma d + (n_1 n_2)^2 \frac{\kappa^2 + \gamma^2}{n_2^4 \kappa^2 + n_1^4 \gamma^2} \right] \times \left(\frac{n_1^2}{n_2^2} \rho^2 \cos^2 \sigma d + \frac{n_2^2}{n_1^2} \sigma^2 \sin^2 \sigma d \right)}} \left(\frac{df}{dz} - \frac{dh}{dz} \right) dz \quad (9.5.26)$$

и для нечетных мод излучения

$$\rho_0(\rho) = -\frac{n_1^2 - n_2^2}{2\sqrt{\pi}|\beta|n_2^2} \times$$

$$\times \rho \int_0^L \frac{V\sqrt{\gamma}(\beta_0\beta \sin \sigma d - \gamma\sigma \cos \sigma d) \cos \kappa de}{(\beta_0 - \beta) \sqrt{\beta_0 \left[\gamma d + (n_1 n_2)^2 \frac{\kappa^2 + \gamma^2}{n_2^4 \kappa^2 + n_1^4 \gamma^2} \right] \times}} \left(\frac{df}{dz} + \frac{dh}{dz} \right) dz. \quad (9.5.27)$$

$$\times \left(\frac{n_1^2}{n_2^2} \rho^2 \sin^2 \sigma d + \frac{n_2^2}{n_1^2} \sigma^2 \cos^2 \sigma d \right)$$

В эти выражения мы ввели функцию искажения нижней стенки волновода $h(z)$. Таким образом, наша теория справедлива для произвольных деформаций обеих стенок при условии, что волновод поддерживает только ТЕ- или ТМ-моду низшего порядка. Если размеры волновода увеличиваются таким образом, что могут распространяться дополнительные направляемые моды, то эти моды также необходимо учесть.

Соотношения (9.5.23) — (9.5.27) справедливы для очень малых отклонений стенки, которые рассмотрены выше, и приближенно выполняются для больших отклонений стенок. Эти соотношения использовались для предсказания потерь на излучение достаточно больших скачков в виде резких сужений волновода. Даже в этих случаях было получено хорошее согласие с экспериментальными данными [96]. Постоянная распространения β_0 падающей направляемой моды и параметры κ, γ являются функциями от z , так как они зависят от ширины волновода. Данная теория применима в тех случаях, когда излучение покидает сужающийся волновод достаточно быстро и не может вернуться обратно к направляемой моде. Это условие удовлетворяется с хорошим приближением как для слабых длинных сужений, так и для крутых сужений ¹⁾. Применимость для последних подтверждена экспериментально в [96].

Сравнение выражений (9.2.38) и (9.5.23) не показывает,

¹⁾ В обоих случаях это справедливо лишь, когда излученная мощность составляет малую часть от мощности падающей волны. Если это не так, то полученные формулы несправедливы (см. [113]). — *Прим. ред.*

что они тождественны. Однако легко продемонстрировать их идентичность в случае малого искажения волновода, для которого было получено выражение (9.2.38). Выражения разд. 9.2 не удовлетворяются для больших отклонений стенок волновода. Теория же, представленная в настоящем разделе, имеет более широкую область применимости.

Предположим, что отклонение стенок от их идеальной формы мало. При этом d должно приближенно равняться d_0 , а β_0 , κ и γ являются постоянными. Тогда после вынесения постоянных из-под интеграла (9.5.23) останется интеграл

$$I = \int_0^L e^{-i(\beta_0 - \beta)z} \frac{d}{dz} [f(z) - h(z)] dz. \quad (9.5.28)$$

Осуществляя интегрирование по частям, получим

$$I = [f(L) - h(L)] e^{-i(\beta_0 - \beta)L} - [f(0) - h(0)] + \\ + i(\beta_0 - \beta) \int_0^L [f(z) - h(z)] e^{-i(\beta_0 - \beta)z} dz. \quad (9.5.29)$$

В разд. 9.2 предполагалось, что волновод регулярный при $z = 0$ и при $z = L$. Таким образом, имеем $f(L) = 0$, $f(0) = 0$, $h(L) = 0$ и $h(0) = 0$. Члены в первых двух квадратных скобках не дают вклада. Используя формулы (9.2.35) и (9.2.36), получим

$$I = i(\beta_0 - \beta) L [F(\beta_0 - \beta) - H(\beta_0 - \beta)]. \quad (9.5.30)$$

Подстановка выражения (9.5.30) в (9.5.23) приводит к (9.2.38).

Полученный результат является важным по ряду причин. Прежде всего, он доказывает, что теория потерь на излучение при искажениях стенки, основанная на синтезе произвольных функций как последовательности малых скачков, приводит к таким же результатам, что и метод разд. 9.2. Кроме того, теория данного раздела была распространена также и на ТМ-моды, хотя получение соответствующих выражений методом разд. 9.2 было значительно более трудным. Подтверждение того, что теория потерь на излучение, основанная на потерях от индивидуального скачка, справедлива, позволяет рекомендовать ее для решения более сложных задач таким же путем.

Таким образом, можно вычислить потери на излучение моды HE_{11} для случайных искажений стенки также, как для скачков и плавных сужений [96]. Наконец, мы получили результаты, применимость которых шире аналогичных результатов, полученных в предыдущих разделах. Соотношения (9.5.23) — (9.5.27) приближенно выполняются для всех типов сужений при условии, что волновод поддерживает только основную моду ¹⁾. При попытке распространить изложенную теорию на случай многомодовых волноводов возникает трудность, связанная с тем, что, возбуждвшись, высшие моды распространяются по волноводу, сохраняя способность к взаимодействию с исходной падающей модой. При этом теория возмущений оказывается неприменимой [91]. При наличии потерь на излучение преобразованная мощность уходит из волновода, так что падающая мода просто испытывает затухание.

Закончим этот раздел применением теории определения потерь на излучение из-за случайных искажений стенки для ТМ-мод. Обсудим также результаты подобных вычислений применительно к основной моде HE_{11} круглого оптического волокна.

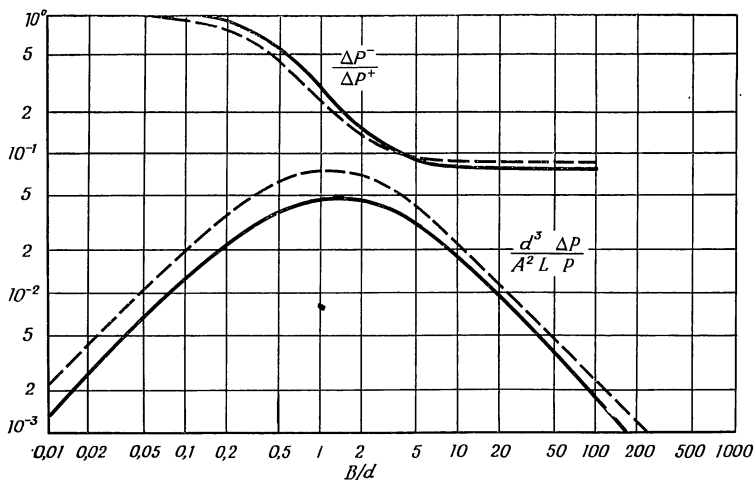
Используя процедуру, аналогичную той, которая выше привела от выражения (9.2.38) к (9.5.23), можно получить амплитуды ТМ-мод излучения, соответствующие (9.2.38) и (9.2.39). Эти выражения можно затем использовать для получения относительных потерь мощности на излучение за счет случайных дефектов стенки. Эквивалент формулы (9.4.11) для ТМ-мод имеет вид

$$\begin{aligned} \left\langle \frac{\Delta P}{P} \right\rangle &= \frac{A^2 \gamma L (n_1^2 - n_2^2)^2}{2\pi B \beta_0 n_2^4} \times \\ &\times \int_{-n_2 k_0}^{n_2 k_0} \left\{ \frac{\rho \cos^2 \kappa d}{\left[(\beta_0 - \beta)^2 + \frac{1}{B^2} \right] \left[\gamma d + n_1^2 n_2^2 \frac{\kappa^2 + \gamma^2}{n_2^2 \kappa^2 + n_1^4 \gamma^2} \right]} \times \right. \\ &\times \left[\frac{(\beta_0 \beta \cos \sigma d + \gamma \sigma \sin \sigma d)^2}{\frac{n_1^2}{n_2^2} \rho^2 \cos^2 \sigma d + \frac{n_2^2}{n_1^2} \sigma^2 \sin^2 \sigma d} + \right. \\ &\left. \left. + \frac{(\beta_0 \beta \sin \sigma d - \gamma d \cos \sigma d)^2}{\frac{n_1^2}{n_2^2} \rho^2 \sin^2 \sigma d + \frac{n_2^2}{n_1^2} \sigma^2 \cos^2 \sigma d} \right] \right\} d\beta. \quad (9.5.31) \end{aligned}$$

¹⁾ См. примечание на стр. 493. — *Прим. ред.*

Результаты численных расчетов с использованием этого выражения приведены на фиг. 9.5.3 и 9.5.4.

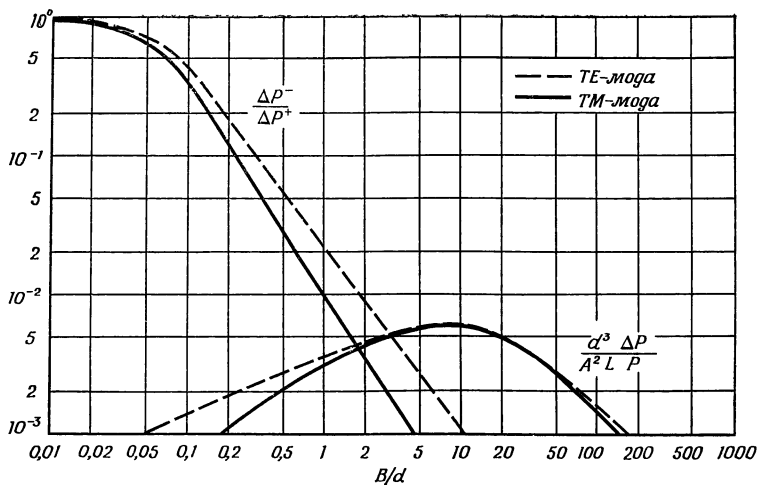
На этих фигурах потери ТМ-моды представлены сплошной кривой и для сравнения пунктирной кривой показаны потери ТЕ-моды. На фиг. 9.5.3 приведены результаты для $n_1 = 1,5$, $n_2 = 1$ и $k_0 d = 1,3$, а на фиг. 9.5.4 для $n_1 = 1,5$, $n_2 = 1$, $k_0 d = 1,3$. Примечательно, что потери ТЕ- и ТМ-мод почти одинаковы. Это связано с тем, что потери на



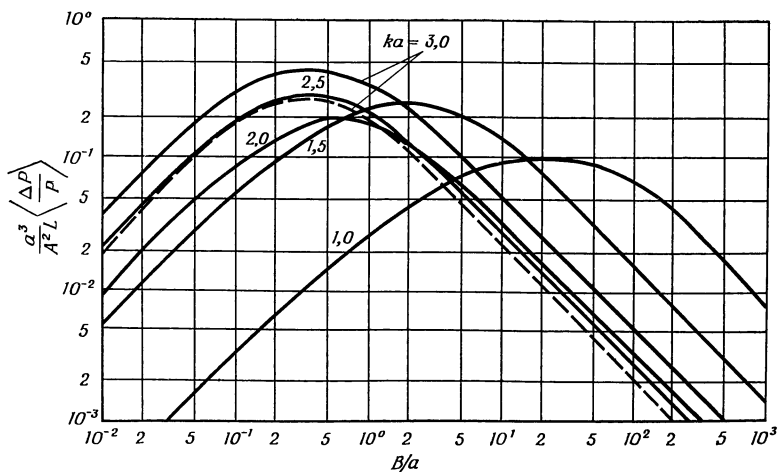
Ф и г. 9.5.3. Потери на излучение ТМ-моды (сплошные кривые) и ТЕ-моды (пунктирные кривые) плоского волновода, вызванные случайным искажением одной стороны поверхности раздела сердцевина — оболочка. Здесь $n_1 = 1,5$, $n_2 = 1$, $kd = 1,3$ [101].

излучение, вызванные случайными дефектами стенки, практически не зависят от поляризации падающей моды.

Соответствующая теория потерь на излучение за счет дефектов стенки круглого диэлектрического волновода (оптического волокна) была разработана для основной моды HE_{11} [96]. Эта теория также основана на вычислении потерь для одного скачка. Рассмотрение искажения стенки было ограничено симметричным изменением диаметра стержня. Полученная теория крайне сложна, поэтому приведем здесь лишь некоторые ее результаты. Потери



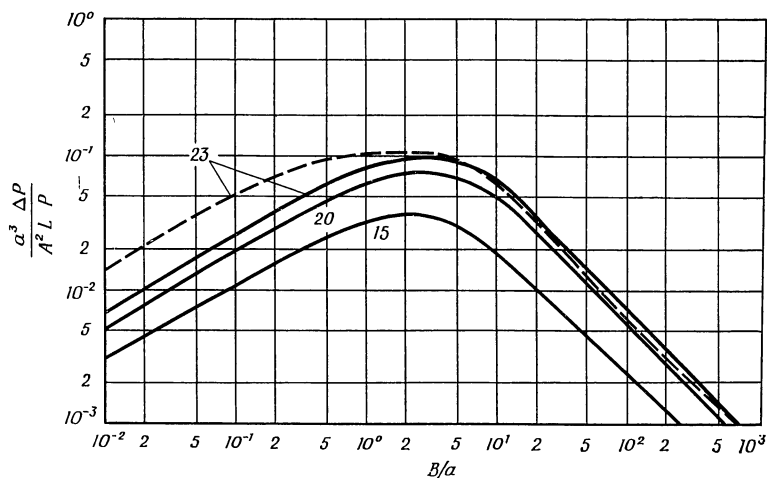
Ф и г. 9.5.4. То же, что на фиг. 9.5.3, но для $n_1 = 1,01$, $n_2 = 1$, $kd = 8$ [101].



Ф и г. 9.5.5. Потери на излучение, вызванные случайными изменениями радиуса сердцевины круглого оптического волокна. Здесь $n_1 = 1,432$, $n_2 = 1$, а значения ka являются параметрами; B — интервал корреляции.

— мода HE_{11} ; — — мода TE_{01} .

моды HE_{11} , вызванные случайными искажениями стенки с экспоненциальной функцией корреляции (9.4.8), представлены на фиг. 9.5.5 и 9.5.6. Сплошной кривой на каждой фигуре показаны потери моды HE_{11} , пунктирной — потери симметричной электрической моды TE_{01} . Последняя мода в некоторых отношениях соответствует TE -моду плоского волновода, но имеет частоту отсечки. Кривые на фиг. 9.5.5 соответствуют $n_1 = 1,432$, $n_2 = 1$. Значения



Фиг. 9.5.6. То же, что и на фиг. 9.5.5, но для $n_1 = 1,01$, $n_2 = 1$. Числовые значения на кривых соответствуют различным значениям $k_0 a$ [401].

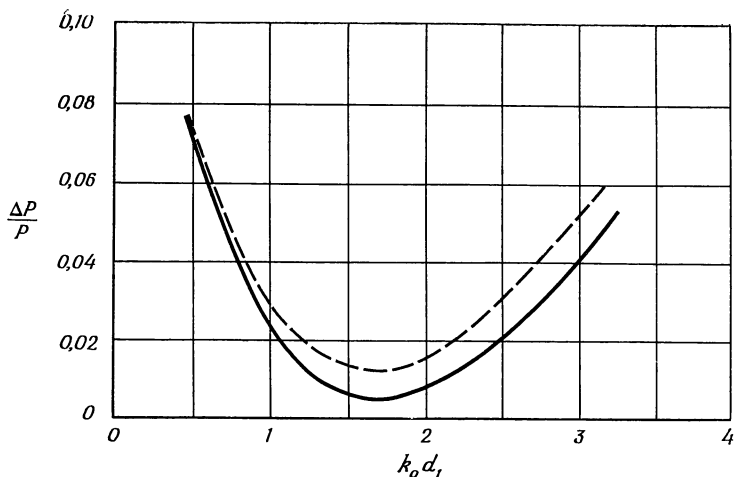
$k_0 a$ отмечены как параметр на каждой кривой. Кривые, приведенные на фиг. 9.5.6, рассчитаны для $n_1 = 1,01$, $n_2 = 1$; кривые для мод HE_{11} и TE_{01} очень близки. Сравнение полученных данных с результатами для плоского волновода показывает, что потери мод круглого стержня в 4 раза выше соответствующих потерь для плоского волновода. Причина такой разницы была объяснена ранее. Она вызвана тем, что плоский волновод имел одну гладкую стенку. Круглый стержень с изменяющимся диаметром соответствует плоскому слою с симметричным изменением толщины. Необходимо, чтобы обе стенки слоя деформировались совершенно одинаково. Потери для этого

случая находятся из фиг. 9.4.2—9.4.7, 9.5.3 и 9.5.4 умножением их на четыре. После такой поправки получается близкое согласие результатов для обоих случаев. Это показывает, что потери моды, вызванные случайными искажениями стенок диэлектрических волноводов, почти не зависят от типа волновода и от типа распространяющейся по нему моды.

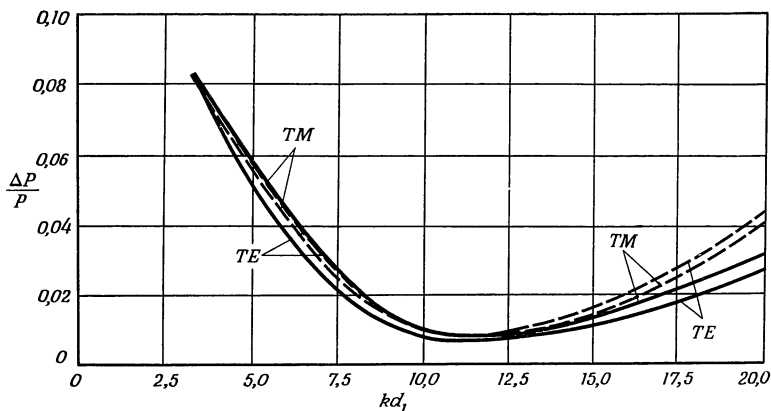
Теория настоящего раздела применима к скачкам и плавным сужениям диэлектрических волноводов. Потери на больших скачках можно рассчитать двумя различными методами. Для вычисления относительных потерь мощности на крутом сужении можно использовать выражения (9.5.23) — (9.5.27) с учетом (9.2.26). Другой путь соответствует более прямому методу. Для вычисления приближенных значений коэффициента прохождения направляемой моды c_t и коэффициента отражения a_r можно использовать формулы (9.5.3) и (9.5.4). Потери, вызванные излучением, затем можно получить из выражения

$$\frac{\Delta P}{P} = 1 - |c_t|^2 - |a_r|^2. \quad (9.5.32)$$

Эта формула основана на том факте, что потери на излучение составляют величину, которая получается после вычитания из полной мощности P прошедшей мощности $|c_t|^2 P$ и отраженной мощности $|a_r|^2 P$. Поскольку оба метода расчета являются приближенными, трудно ожидать точного совпадения результатов. Однако совпадение результатов конкретных расчетов обоими методами оказывается удивительно хорошим даже для больших скачков при отношении толщин слоя 2 : 1. Фиг. 9.5.7 демонстрирует это сравнение потерь на излучение ТЕ-моды от скачка в плоском волноводе [101]. Произведение $k_0 d_1$ используется на этой фигуре как независимая переменная, где полуширина d_1 соответствует более широкой части волновода. Удивительно, что потери на излучение основной моды ТЕ плоского волновода могут быть такими малыми. Потери ТМ-моды оказываются несколько выше [101]. Сплошной кривой представлены результаты расчета по методу потерь на излучение, а пунктирной — результаты расчета по формуле (9.5.32). Согласие между результатами вычислений этими двумя метода-



Ф и г. 9.5.7. Относительные потери на излучение, вызванные одним скачком, для моды ТЕ низшего порядка плоского волновода [401]. Сплошная и пунктирная кривые соответствуют разным приближенным методам вычисления потерь. Отношение двух полуширин сердцевины до и после скачка $d_2/d_1 = 0,5$, $n_1 = 1,432$, $n_2 = 1$.



Ф и г. 9.5.8. Потери на излучение ТЕ- и ТМ-мод низшего порядка от одного скачка плоского волновода, $d_2/d_1 = 0,5$, $n_1 = 1,01$, $n_2 = 1$. Сплошные и пунктирные кривые соответствуют разным приближенным методам расчета [401].

ми очень хорошее для рассматриваемого размера скачка сечения.

Еще лучшее согласие получается для случая $n_1 = 1,01$, $n_2 = 1$ и $d_2/d_1 = 0,5$. Результаты вычислений для этого случая приведены на фиг. 9.5.8. Сплошные кривые опять получены из теории потерь на излучение, а пунктирные представляют результаты расчета по формуле (9.5.32). Потери ТЕ- и ТМ-мод почти одинаковы для волновода с малой разностью показателей преломления сред.

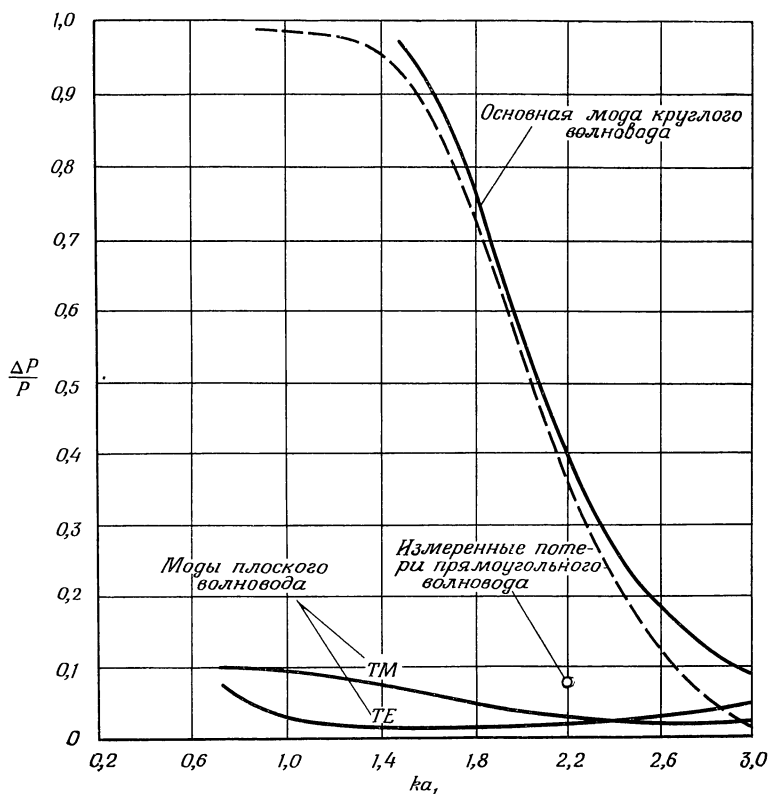
Потери от скачка сечения плоского волновода намного ниже соответствующих потерь для круглого диэлектрического волновода. Для вычисления потерь от резких изменений сечения круглого диэлектрического волновода использовалась теория моды HE_{11} [96]. Результаты этих вычислений приведены на фиг. 9.5.9 и 9.5.10.

На фиг. 9.5.9 показаны потери на излучение моды HE_{11} круглого волновода и потери ТЕ- и ТМ-мод диэлектрического слоя при $n_1 = 1,432$, $n_2 = 1$ и $a_2/a_1 = 0,5$. Радиус круглого волновода обозначен через a ; для плоского волновода $a = d$. Результаты расчета обоими методами представлены сплошной и пунктирной кривыми. Согласие двух методов опять очень хорошее. Из графиков четко видно, насколько потери моды HE_{11} круглого стержня выше потерь ТЕ- и ТМ-мод плоского волновода. Кривые на фиг. 9.5.10 даны для $n_1 = 1,01$, $n_2 = 1$. Из них видно, что при больших значениях $k_0 a_1$ потери мод плоского волновода приближаются к потерям моды круглого волновода.

Теория потерь на скачке проверялась экспериментально в СВЧ-диапазоне [96]. В эксперименте использовался резонатор с гладким тефлоновым стержнем для калибровки, у которого $n_1 = 1,432$, а $n_2 = 1$, и два тефлоновых стержня со скачками диаметра. Результаты измерения потерь на скачках на фиг. 9.5.11 отмечены крестиками. Кривая построена по результатам расчета обоими методами. Совпадение результатов в этом случае настолько сильное, что обе кривые практически сливаются. Согласие данных эксперимента и теории также очень хорошее. При сравнении результатов на фиг. 9.5.9 и 9.5.11 следует учесть, что $ka_2 = (a_2/a_1) ka_1$.

Для проверки потерь, предсказанных теорией плоского волновода, использовался резонатор для измерения

потерь на излучение на скачке с отношением толщин 2 : 1 прямоугольного диэлектрического волновода [96], поскольку очень трудно реализовать настоящий плоский волно-

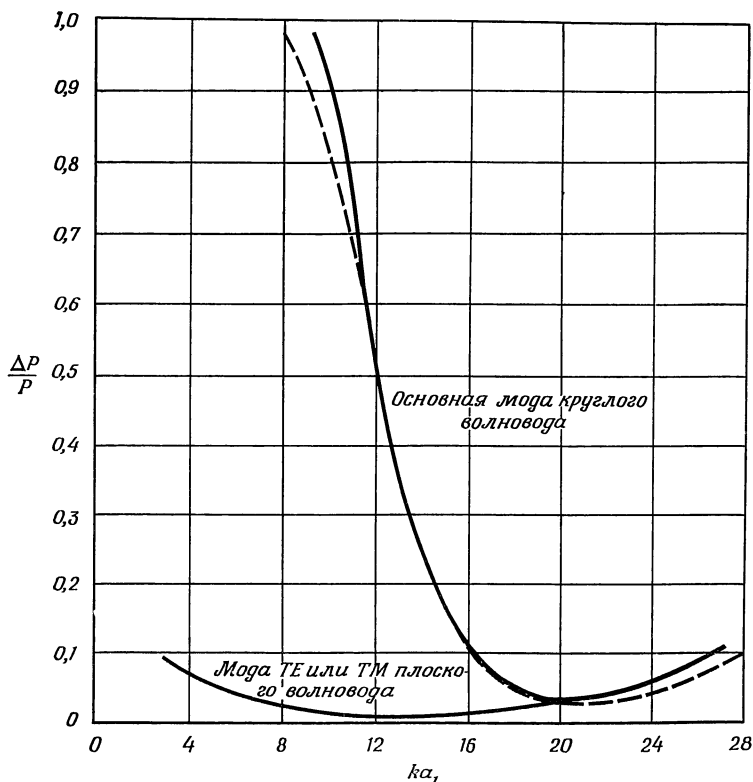


Фиг. 9.5.9. Потери на излучение, вызванные одним скачком ширины сердцевины плоского волновода и диаметра круглого оптического волокна [96].

Высота скачка $d_2/d_1 = 0,5$, $n_1 = 1,432$, $n_2 = 1$. Сплошные и пунктирные кривые соответствуют двум разным приближенным методам. Круглом отмечены измеренные потери от скачка для прямоугольного диэлектрического волновода.

вод. В прямоугольном диэлектрическом волноводе отношение широкой стороны к узкой составляло 2 : 1. Скачок осуществлялся по узкой стороне волновода без измене-

ния его широкой стороны. Такой прямоугольный диэлектрический волновод приближенно соответствует плоскому волноводу ¹⁾. Экспериментальный результат, соответст-

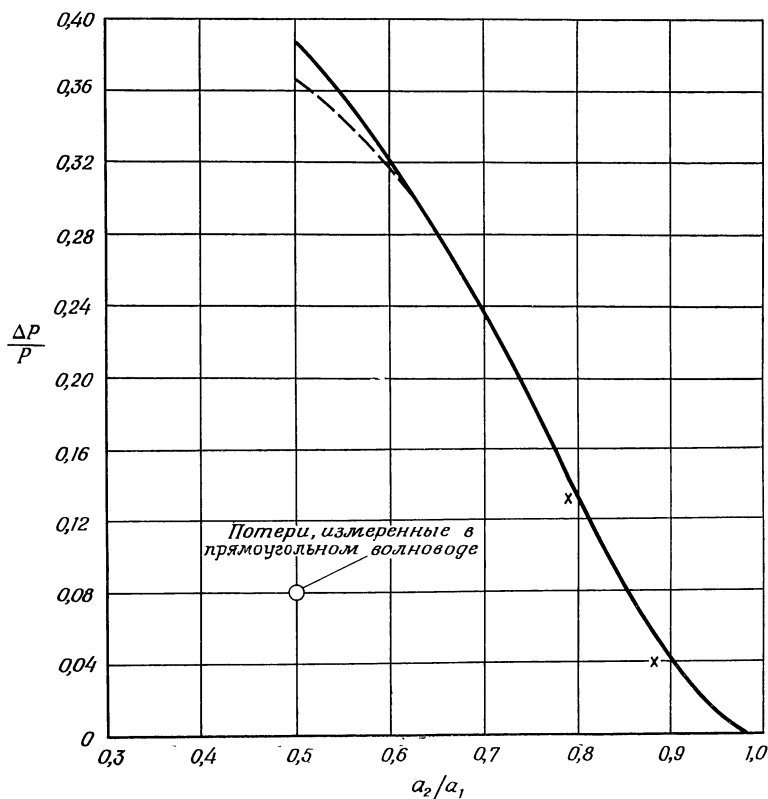


Ф и г. 9.5.10. То же, что и на фиг. 9.5.9 для $d_2/d_1 = 0,5$, $n_1 = 1,01$, $n_2 = 1$ [96].

вующий прямоугольному волноводу, отмечен на фиг. 9.5.9 и 9.5.11 кружком. Потери для волновода такого типа на-

¹⁾ Как показано в [82*, 101*, 104*], замедление, энергетические характеристики и затухание волн прямоугольного диэлектрического волновода близки к соответствующим величинам плоского волновода при отношении сторон 10:1 и более.— *Прим. ред.*

много ниже, чем для круглого волновода, но существенно выше, чем для плоского волновода. Этот результат пока-



Ф и г. 9.5.11. Зависимость потерь на излучение моды HE_{11} круглого волокна от отношения a_2/a_1 [96].

Крестиками представлены измеренные значения, соответствующие расчетным параметрам. Кругиком отмечен результат измерения для прямоугольного диэлектрического волновода. Здесь $n_1 = 1,432$, $n_2 = 1$, $ka_2 = 1,1$.

зывает, что прямоугольный диэлектрический волновод не так «терпим» к скачкам, как плоский волновод. Однако, он близок к плоскому волноводу по порядку величины

потерь, которые намного меньше, чем у круглого волновода ¹⁾.

Потери на излучение можно уменьшить до любого желаемого уровня, если скачок, или резкое сужение, волновода растянуть в плавный протяженный переход. Эта идея является основанием для рассмотрения механизма уменьшения потерь на полученных таким образом длинных волноводных переходах [113]. Потери на переходах определяются амплитудами мод излучения [формулы (9.5.23)—(9.5.27)].

Если переход настолько короткий, что $\int_0^z (\beta_0 - \beta) du \ll 1$,

и если пренебречь зависимостью от z других множителей в подынтегральном выражении, то интеграл приближенно будет пропорционален алгебраической сумме $f(L) - f(0) \pm (h(L) - h(0))$. Это означает, что для коротких переходов высота скачка является единственным фактором, определяющим потери на излучение. Потери от резких скачков и коротких переходов одинаковы. По этой причине приближенная теория потерь на излучение для переходов применима для резких скачков. Как только переход становится таким длинным, что $\int_0^z (\beta_0 - \beta) du$ оказывается порядка единицы и больше,

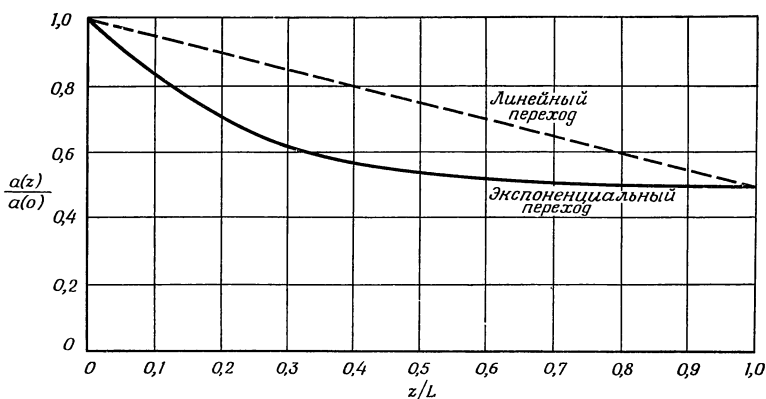
необходимо учитывать еще один эффект. Экспоненциальная функция с экспонентой вида $-i \int_0^z (\beta_0 - \beta) du$ является

осциллирующей функцией (точнее говоря, ее действительная и мнимая части — осциллирующие функции). Быстро осциллирующие функции подынтегрального выражения уменьшают величину интеграла. Это наиболее наглядно демонстрируется на примере дельта-функции (8.4.16),

¹⁾ Потери на излучение от скачка сечения прямоугольного волновода зависят от соотношения сторон. Если сечение квадратное и скачок имеет место одновременно в обоих поперечных направлениях, то, по-видимому, потери будут такими же, как у круглого волновода. — *Прим. ред.*

которая дает вклад в интеграл только в области своего полюса, а быстро осциллирующая синусоидальная функция уменьшает интеграл от любой другой функции, стоящей вместе с ней под интегралом в качестве сомножителя.

Экспоненциальная функция в интегралах (9.5.23)—(9.5.27) не осциллирует очень быстро. Однако если первые и вторые производные функций $f(z)$ и $h(z)$ малы, то влияние экспоненциальной функции становится преобладающим. Действительно, для бесконечно длинных переходов интегралы обращаются в нуль ¹⁾.



Ф и г. 9.5.12. Профили линейного и экспоненциального переходов.

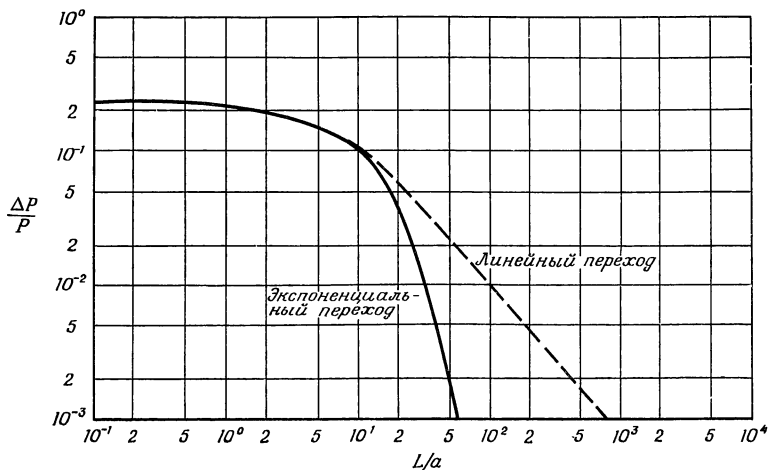
Величина $\int_0^z (\beta_0 - \beta) du$ в экспоненциальной функции

зависит от z . Постоянная распространения β_0 , имеющая большее значение в широкой части перехода, уменьшается в узкой его части. Большое значение этой величины приводит к уменьшению интеграла. Поэтому вклад в интеграл

¹⁾ Бесконечный переход при условии конечного изменения сечения делает нерегулярный участок волновода регулярным, поэтому потери на излучение и преобразование на нем отсутствуют.— *Прим. ред.*

в узкой части линейного перехода больше, чем в широкой.

Очевидно, что преимущество протяженного перехода можно усилить, формируя переход таким образом, чтобы потери на излучение от него были равномерно распределены по всей его длине. Анализ показывает, что переход с экспоненциальной формой изменения толщины дает более равномерный поток излучения по всей его длине, чем линейный переход. Профили обоих типов переходов приведены на фиг. 9.5.12. Разрыв первой производной экспоненциального перехода при $z = 0$ не дает вклада в интеграл. Поэтому безразлично, будем ли его сглаживать, чтобы



Ф и г. 9.5.13. Зависимость относительных потерь на излучение линейного перехода (пунктирная линия) и экспоненциального перехода (сплошная кривая) от отношения длины перехода к радиусу сердцевины L/a .

Отношение радиуса сердцевины на узком конце перехода к радиусу на его широком конце $a_2/a_1 = 0,5$; $n_1 = 1,432$; $n_2 = 1$, $ka_1 = 2,5$.

придать непрерывность первой производной, или оставим его форму такой, как показано на фигуре. Зависимость потерь на излучение моды HE_{11} круглого волновода для обоих переходов от длины перехода приведена на фиг. 9.5.13.

Представленные кривые демонстрируют некоторые интересные особенности. Потери на излучение постоянны от $L/a = 0$ до $L/a = 1$. Плавные переходы такой короткой длины приводят к таким же потерям, как резкие скачки толщины. При увеличении длины перехода потери на излучение быстро уменьшаются. Потери от линейного перехода при $L/a = 400$ составляют только 1% от потерь скачка. Кривые также демонстрируют справедливость нашего предположения о низких потерях экспоненциального перехода. Выигрыш при использовании экспоненциального перехода оказывается значительным. Однако, увеличивая длину линейного перехода, можно получить те же преимущества, что и в случае экспоненциального перехода. Длина этих переходов не обязательно должна быть очень большой. Например, для $\lambda = 1$ мкм радиус волновода приблизительно равен 1 мкм. Переход с $L/a = 1000$ имеет длину всего 1 мм.

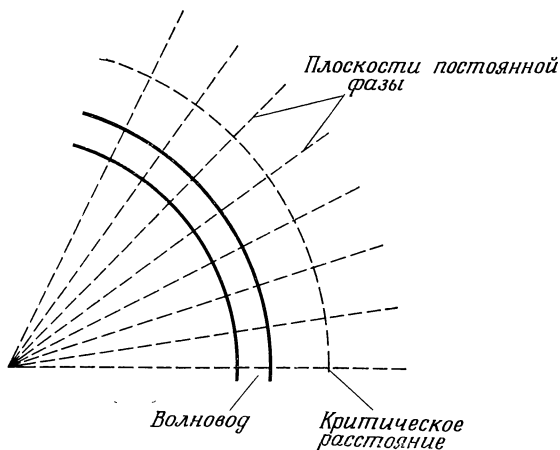
9.6. ПОТЕРИ НА ИЗЛУЧЕНИЕ В ИЗОГНУТЫХ ВОЛНОВОДАХ

Диэлектрические волноводы не могут направлять электромагнитную энергию по изгибам без потерь мощности на излучение [114]. Практически потери на изгибе могут быть малы, если радиус кривизны изгиба достаточно велик, и велики, если волновод сильно изогнут [106].

Рассмотрение потерь на изгибе затруднительно. Поэтому, следуя нашей обычной практике, рассмотрим основные физические принципы, ответственные за формирование потерь на изгибе плоского волновода.

Появление потерь на изгибе можно объяснить следующим образом. Часть мощности, направляемой диэлектрическим волноводом, распространяется вне области сердцевины. Вдали от отсечки эта часть мощности может быть очень малой, но тем не менее она существует. Поле вне сердцевины волновода спадает экспоненциально в поперечном направлении от оси волновода. Теоретически некоторое конечное количество мощности присутствует даже на весьма значительных расстояниях от волновода. Для изогнутого волновода плоскости постоянной фазы принимают форму, которая схематически показана на

фиг. 9.6.1. Вблизи оси волновода плоскости постоянной фазы (или фазовые фронты) движутся со скоростью, совпадающей с фазовой скоростью моды прямого волновода. Эта скорость меньше скорости плоских волн в среде, окружающей сердцевину. Можно изобразить движение фазовых фронтов так, как если бы их плоскости были нанизаны на стержень, расположенный в центре кривизны, и вращались вокруг этого центра. Очевидно, что скорость фазового фронта на границе волновода, которая ближе к центру

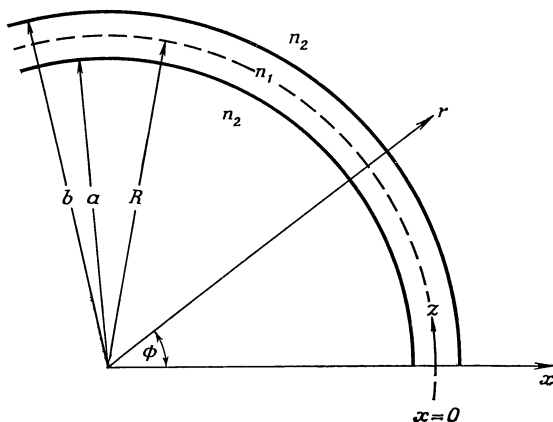


Ф и г. 9.6.1. Изогнутый плоский волновод. Критическое расстояние соответствует границе, где скорость фазовых фронтов равна скорости света.

кривизны, оказывается меньше фазовой скорости моды в прямом волноводе, а скорость фазового фронта на другой границе волновода выше, чем скорость моды прямого волновода. На некотором критическом расстоянии скорость фазового фронта равняется скорости плоской волны в среде, окружающей волновод. Электромагнитные волны не могут распространяться со скоростью, превышающей скорость плоской волны в среде. Поэтому часть поля отделяется от поля направляемой моды и излучается в пространство. Такое грубое рассмотрение качественно объясня-

ет, почему на изгибе диэлектрических волноводов мощность должна теряться на излучение ¹⁾.

Математическое рассмотрение проблемы потерь на изгибе основано на использовании того, что для очень слабых изгибов распределение поля и фазовая скорость около сердцевины волновода почти такие же, как для прямого волновода. Известен также точный вид, который поле должно принимать в цилиндрической системе координат. Используем систему координат, приведенную на фиг. 9.6.2.



Ф и г. 9.6.2. Система координат, используемая для описания потерь на излучение изогнутого плоского волновода.

Ось y (ее обычно используют как ось z) цилиндрической системы координат выходит из плоскости чертежа. Решение уравнений Максвелла в цилиндрических координатах дается в виде

$$E_y = BH_v^{(2)}(n_2 k_0 r) e^{i(\omega t - \nu \phi)}, \quad (9.6.1)$$

$$H_r = B \frac{\omega \epsilon_0}{k_0^2} \frac{\nu}{r} H_v^{(2)}(n_2 k_0 r) e^{i(\omega t - \nu \phi)}, \quad (9.6.2)$$

$$H_\phi = -iB \frac{n_2 \omega \epsilon_0}{k_0} H_v^{(2)'}(n_2 k_0 r) e^{i(\omega t - \nu \phi)}, \quad (9.6.3)$$

¹⁾ Изложенная здесь идея физической причины излучения с изгиба, впервые высказанная в работе [114], была использована для количественного расчета потерь на излучение круглого диэлектрического волновода в [111*]. — Прим. ред.

где $k_0 = \omega \sqrt{\epsilon_0 \mu_0}$ и предполагается, что отсутствует зависимость поля от y . Это решение справедливо только для $r \gg b$, т. е. вне сердцевины волновода. Внутри сердцевины поле описывается суперпозицией функций Ханкеля первого и второго рода, а вне сердцевины для $r \leq a$ выражается через функцию Бесселя $J_\nu(n_2 k_0 r)$. Штрих у функции Ханкеля в (9.6.3) обозначает ее производную по аргументу $n_2 k_0 r$. Выбор функций определяется требованием, согласно которому поле должно спадать экспоненциально в направлении увеличения r при $r > b$ и превратиться в распространяющуюся волну при $r \gg b$. Функция Ханкеля второго рода удовлетворяет этим требованиям. Внутри волновода нужно использовать функции Ханкеля первого и второго рода или функции Бесселя и Неймана. Для $r < a$ необходимо взять цилиндрическую функцию, которая не имеет особенностей при $r = 0$. Единственной функцией, удовлетворяющей этому условию, является $J_\nu(n_2 k_0 r)$. Порядок ν цилиндрической функции в данной задаче не обязательно должен быть целым числом, так как нет необходимости, чтобы функции изменялись периодически по ϕ с периодом 2π . Для данного рассмотрения не требуется определять поле внутри сердцевины волокна и для $r < a$.

Представление решения в виде (9.6.1) — (9.6.3) определяется требованиями геометрии и уравнениями Максвелла. Если положить, что поле не зависит от y и изогнутый слой неограниченно простирается в направлении оси y , поле не может иметь другой формы. В то же время известно, что в пределе при $R \rightarrow \infty$ поле около сердцевины должно принимать вид (8.3.12), соответствующий симметричной ТЕ-моду прямолинейного плоского волновода. Это условие позволяет определить постоянную B . Зная этот параметр, с помощью уравнений поля можно найти мощность, текущую в радиальном направлении на бесконечном расстоянии от волновода. Полученный результат можно использовать для определения потерь направляемой моды.

Строгий метод решения задачи, конечно, требует определения неизвестных амплитудных коэффициентов из граничных условий. В результате можно получить уравнение собственных значений для определения ν . Это уравнение не имеет действительных решений. Комплексные значения ν определяют потери моды. Такая процедура при-

водит к точному решению задачи [106, 114]. Однако она намного сложнее следующего приближенного метода. Обойдем необходимость вычисления v из уравнения собственных значений, полагая, что постоянная распространения волны в изогнутом волноводе почти такая же, как постоянная распространения соответствующей моды в прямолинейном волноводе. Сравнение с модами прямолинейного волновода дает соотношение

$$v\phi = \beta z. \quad (9.6.4)$$

Координата отсчитывается здесь вдоль дуги изгиба оси волновода, $z = R\phi$. Порядок функции Бесселя определяется, таким образом, выражением

$$v = \beta R. \quad (9.6.5)$$

Предполагается, что постоянная распространения β такая же, как у четной ТЕ-моды прямолинейного плоского волновода. Она получается как решение уравнения собственных значений (8.3.16).

Для определения константы B необходимо аппроксимировать функцию Ханкеля. Из формулы (9.6.5) следует, что v — величина того же порядка, что и аргумент функции Бесселя $n_2 k_0 r$. Введя ось x , как показано на фиг. 9.6.2, получим соотношение

$$r = R + x. \quad (9.6.6)$$

Отношение порядка к аргументу есть

$$\text{ch } \alpha = \frac{v}{n_2 k_0 r} = \frac{\beta}{n_2 k_0} \frac{1}{1 + \frac{x}{R}} \approx \frac{\beta}{n_2 k_0} \left(1 - \frac{x}{R}\right). \quad (9.6.7)$$

Постоянная распространения моды больше постоянной распространения плоских волн в среде вне сердцевины:

$$\beta > n_2 k_0. \quad (9.6.8)$$

При $x = 0$ отношение (9.6.7) больше единицы. Однако, так как x возрастает, величина (9.6.7) должна в конечном счете стать меньше единицы. Однако для больших R и малых x можно считать, что отношение (9.6.7) больше

единицы ¹⁾. Необходимое приближение для функции Ханкеля можно найти в [61]

$$H_v^{(2)}(n_2 k_0 r) = -i \frac{e^{v(\alpha - \text{th } \alpha)}}{\sqrt{\frac{\pi}{2} v \text{th } \alpha}}. \quad (9.6.9)$$

Аргумент α гиперболического тангенса определяется из формулы (9.6.7). В результате имеем

$$u = \text{th } \alpha = \frac{\sqrt{\text{ch}^2 \alpha - 1}}{\text{ch } \alpha} = \frac{\sqrt{v^2 - (n_2 k_0 R)^2}}{v}. \quad (9.6.10)$$

При этом величину $v \text{th } \alpha$ в знаменателе (9.6.9) достаточно аппроксимировать следующим образом:

$$v \text{th } \alpha = \sqrt{v^2 - (n_2 k_0 R)^2} = R \sqrt{\beta^2 - (n_2 k_0)^2}.$$

Тогда из (8.3.14) получим

$$v \text{th } \alpha = \gamma R. \quad (9.6.11)$$

Аргумент экспоненциальной функции в (9.6.9) необходимо аппроксимировать более точно. Используем соотношение

$$\alpha = \text{Ar th } u = \frac{1}{2} \ln \left(\frac{1+u}{1-u} \right) = u + \frac{1}{3} u^3 + \frac{1}{5} u^5 + \frac{1}{7} u^7 + \dots, \quad (9.6.12)$$

где u определяется формулой (9.6.10). Согласно (9.6.7), $\text{ch } \alpha$ тоже не отличается от единицы, так что $\text{th } \alpha = u$ есть малая величина, удовлетворяющая правой части разложения (9.6.12). Можно записать

$$\alpha - \text{th } \alpha = \frac{1}{3} u^3 + \frac{1}{5} u^5 + \frac{1}{7} u^7 + \dots. \quad (9.6.13)$$

С помощью формул (9.6.5), (9.6.6) и (9.6.10) получим

$$\begin{aligned} u^p &= \left[\frac{\sqrt{\beta^2 - (n_2 k_0 R/R)^2}}{\beta} \right]^p \approx \left[\frac{\sqrt{\gamma^2 - 2(n_2 k_0)^2 x/R}}{\beta} \right]^p \approx \\ &\approx \left(\frac{\gamma}{\beta} \right)^p \left(1 - p \frac{(n_2 k_0)^2 x}{\gamma^2 R} \right). \end{aligned} \quad (9.6.14)$$

¹⁾ Для нашей приближенной теории требуется выполнение двух условий $\beta/[n_2 k_0 (1 + \frac{d}{R})] > 1$ и $\beta/[n_1 k_0 (1 - \frac{d}{R})] < 1$. Если одно или оба условия не выполняются, наша теория несправедлива.

Подстановка в (9.6.13) дает

$$\alpha - \text{th } \alpha = \frac{1}{3} \left(\frac{\gamma}{\beta} \right)^3 + \frac{1}{5} \left(\frac{\gamma}{\beta} \right)^5 + \frac{1}{7} \left(\frac{\gamma}{\beta} \right)^7 + \dots$$

$$\dots - \left(\frac{n_2 k_0}{\beta} \right)^2 \frac{\gamma}{\beta} \frac{x}{R} \left[1 + \left(\frac{\gamma}{\beta} \right)^2 + \left(\frac{\gamma}{\beta} \right)^4 + \left(\frac{\gamma}{\beta} \right)^6 + \dots \right]. \quad (9.6.15)$$

Бесконечные ряды в правой части (9.6.15) можно снова свернуть с помощью (9.6.12); выражение в квадратных скобках является просто геометрической прогрессией. Таким образом, можно записать

$$\alpha - \text{th } \alpha = \left(\text{Ar th } \frac{\gamma}{\beta} \right) - \frac{\gamma}{\beta} - \frac{(n_2 k_0)^2}{\beta^3 R} \frac{\gamma x}{1 - (\gamma/\beta)^2}. \quad (9.6.16)$$

Используя равенство

$$1 - \left(\frac{\gamma}{\beta} \right)^2 = \frac{(n_2 k_0)^2}{\beta^2}, \quad (9.6.17)$$

из формул (9.6.5), (9.6.9), (9.6.11), (9.6.16) и (9.6.17) окончательно получим

$$H_v^{(2)}(n_2 k_0 r) =$$

$$= - \frac{i}{\sqrt{\frac{\pi}{2} \gamma R}} \exp \left(\beta R \text{Ar th } \frac{\gamma}{\beta} - \gamma R \right) e^{-\gamma x}. \quad (9.6.18)$$

Видно, что поле в виде (9.6.1), если аппроксимировать его с помощью (9.6.18), зависит от x так же, как поле в формуле (8.3.12). Сравнение двух выражений дает амплитудный коэффициент B в виде

$$B = i \sqrt{\frac{\pi}{2} \gamma R} \sqrt{\frac{2\omega\mu_0 P}{\beta d + \frac{\beta}{\gamma}}} \cos \kappa d e^{\gamma d} e^{-[\beta \text{Ar th } (\gamma/\beta) - \gamma]R}. \quad (9.6.19)$$

Вычисление потерь на изгибе теперь не представляет труда. Коэффициент потерь мощности определяется отношением вклада мощности от единичной длины волновода в поле излучения на бесконечности к мощности P ,

проходящей в волноводе:

$$2\alpha = \frac{(LS_r)_{r \rightarrow \infty}}{P}. \quad (9.6.20)$$

Так как α — амплитудный коэффициент затухания, то коэффициент потерь мощности есть 2α . Параметр L является длиной дуги между двумя радиусами r , ведущими к конечным точкам дуги единичной длины вдоль оси волновода. Составляющая S_r — компонента вектора плотности потока мощности (вектора Пойнтинга) вдоль координаты r . Таким образом, имеем

$$L = \frac{r}{R} \quad (9.6.21)$$

и

$$\begin{aligned} S_r &= -\frac{1}{2} E_y H_\phi^* = \\ &= -i |B|^2 \frac{n_2 \omega \epsilon_0}{2k_0} H_v^{(2)}(n_2 k_0 r) [H_v^{(2)'}(n_2 k_0 r)]^*. \end{aligned} \quad (9.6.22)$$

Звездочка обозначает комплексно-сопряженную величину.

Поскольку аргумент в данном случае намного больше порядка, необходимое приближение функции Ханкеля для большого аргумента имеет вид [11, 61]

$$H_v^{(2)}(n_2 k_0 r) = \frac{\sqrt{2}}{\sqrt{\pi n_2 k_0 r}} e^{-in_2 k_0 r} e^{i(\pi/4)} e^{iv(\pi/2)}. \quad (9.6.23)$$

Подстановка (9.6.19), (9.6.21), (9.6.22) и (9.6.23) в (9.6.20) дает окончательный результат

$$2\alpha = \frac{\gamma^2}{\beta(1+\gamma d)} \frac{\kappa^2}{(n_1^2 - n_2^2) k_0^2} e^{2\gamma d} e^{-2[\beta \operatorname{Arth}(\gamma/\beta) - \gamma]R}. \quad (9.6.24)$$

При получении выражения (9.6.24) было использовано соотношение

$$\kappa^2 + \gamma^2 = (n_1^2 - n_2^2) k_0^2 \quad (9.6.25)$$

и выражение (8.6.16)

$$\cos^2 \kappa d = \frac{\kappa^2}{\kappa^2 + \gamma^2}. \quad (9.6.26)$$

Для малых значений γ/β можно получить упрощенное приближенное выражение

$$2\alpha = \frac{\gamma^2}{\beta(1+\gamma d)} \frac{\kappa^2}{(n_1^2 - n_2^2) k_0^2} e^{2\gamma d} e^{-(2/3)(\gamma^3/\beta^2)R}. \quad (9.6.27)$$

У читателя может возникнуть вопрос, почему потери мощности при некотором z находятся путем вычисления излученной мощности на бесконечном расстоянии от волновода. Мощность, протекающая при данном z_1 и $x \rightarrow \infty$, должна включать мощность от другого значения z_2 , расположенного на большом расстоянии от точки z_1 ¹⁾. Справедливость используемой нами процедуры следует из предположения, что амплитуда направляемой волны не изменяется, хотя она используется для вычисления потерь мощности. Допущение постоянства амплитуды типично для любой теории возмущений первого порядка. Так как радиально текущая мощность не зависит от z (или от угла ϕ), не имеет значения, какой элемент волновода фактически дает вклад. Относительные потери мощности на единицу длины постоянны и не зависят от положения вдоль оси волновода.

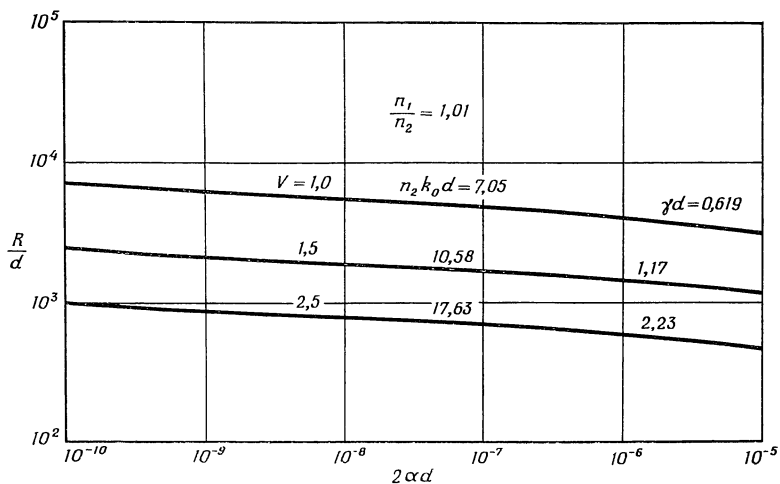
Данный приближенный метод расчета потерь на изгибе очень хорошо согласуется с результатом, полученным Маркатили [106], который использовал приближенное решение комплексного уравнения собственных значений для расчета потерь волны на изгибе. Его выражения существенно сложнее полученных здесь. Однако результаты расчета потерь обоими методами хорошо согласуются. Изложенная теория не ограничивается малой разностью $n_1 - n_2$.

Значения γ (или β) должны получаться из решения уравнения собственных значений (8.3.16) или из приближенных уравнений (8.6.12) и (8.6.20).

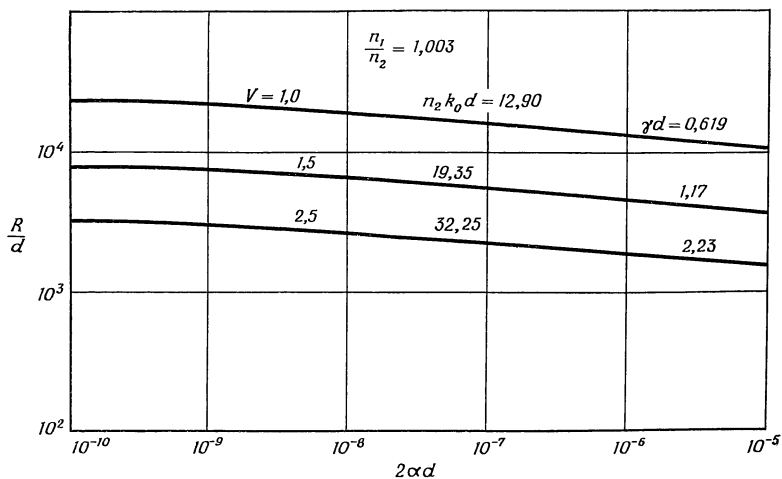
Был рассмотрен только простейший случай четной ТЕ-моды плоского волновода. Теория потерь на изгибе моды HE_{11} круглого волокна намного сложнее²⁾. Заманчиво попытаться с помощью теории плоского волновода по крайней мере оценить ожидаемые потери на изгибе для моды HE_{11} круглого волокна. В этом случае требуется использовать величины κ , γ и β для круглого волок-

1) Здесь автор касается вопроса локальности процесса излучения на изгибе. Только при условии локальности потери могут быть выражены с помощью экспоненциальной функции. — *Прим. ред.*

2) См. ссылку [111*]. — *Прим. ред.*



Ф и г. 9.6.3. Зависимость отношения радиуса кривизны изгиба волновода R к полуширине слоя d от величины потерь на излучение 2α в случае $n_1/n_2 = 1,01$.



Ф и г. 9.6.4. То же, что и на фиг. 9.6.3 для $n_1/n_2 = 1,003$,

на вместо этих величин при $a = d$ для плоского волновода ¹⁾.

Графики зависимости нормированного радиуса кривизны R/d от потерь на изгибе $2\alpha d$ приведены на фиг. 9.6.3 и 9.6.4. Отношение показателей преломления сердцевин и оболочки $n_1/n_2 = 1,01$ для кривых на фиг. 9.6.3, а на фиг. 9.6.4 $n_1/n_2 = 1,003$. Здесь приведена зависимость, обратная (9.6.27), с учетом приближенного решения (8.6.12), чтобы показать, какой радиус кривизны допустим, если заданы приемлемые потери на изгибе. Интересно, что R/d мало изменяется в зависимости от $2\alpha d$. Следовательно, незначительное изменение радиуса кривизны вызывает существенное изменение коэффициента потерь. Типичные рабочие условия могут иметь место при отношении показателей преломления $n_1/n_2 = 1,003$, полуширине слоя $d = 1$ мкм и $n_2 k d = 12,9$. Потери на изгибе при этом изменяются от $2\alpha d = 10^{-10}$, или $2\alpha = 0,434$ дБ/км, до $2\alpha d = 10^{-5}$, или $2\alpha = 43,4$ дБ/м, при изменении радиуса кривизны от $R = 2,2$ м до $R = 1$ см.

¹⁾ Результаты, рассчитанные по формуле плоского волновода (9.6.24), хорошо согласуются с экспериментальными данными для моды HE_{11} [25].

СВЯЗЬ МЕЖДУ ДИЭЛЕКТРИЧЕСКИМИ ВОЛНОВОДАМИ

10.1. ВВЕДЕНИЕ

В разд. 2.7 и 9.3 было показано, что две волны могут быть связаны настолько эффективно, что возможна полная передача мощности от одной из них к другой. Этот тип связи вызывается синусоидальными отклонениями границы раздела сред сердцевины и оболочки в волноводе или синусоидальным изменением вдоль оси показателя преломления волновода. В этой главе будет рассмотрена связь между направляемыми модами двух различных волноводов. Связь между волноводами также может быть вызвана изменениями геометрической формы или неоднородностями диэлектрической среды волновода. Мощность в одном из волноводов рассеивается нерегулярностями такого типа и излучается в окружающее волновод пространство. Если регулярный волновод поместить в поле такого излучения, то оно пройдет сквозь волновод и частично отразится им, но эта мощность не перейдет ни в какую моду второго волновода. Однако если второй волновод тоже имеет нерегулярности, то часть этой мощности может частично преобразоваться в направляемые моды этого волновода [104]. Такой тип связи называется связью через рассеяние. При этом механизме связи не происходит полного обмена мощностью между волноводами.

Два диэлектрических волновода, даже если оба они регулярные, могут эффективно обмениваться мощностью совсем по другой причине [103]. В гл. 9 показано, что электромагнитное поле направляемых мод в диэлектрических волноводах спадает экспоненциально в поперечном направлении от оси волновода. Если два диэлектрических волновода расположены рядом, то часть поля одного

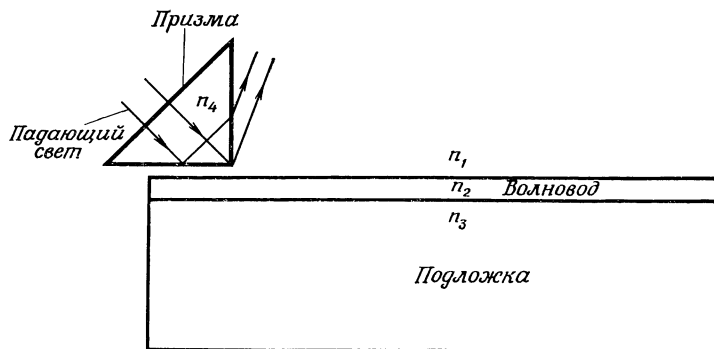
волновода достигает другого волновода. Наличие второго волновода несколько искажает поле направляемой моды первого волновода. Это искаженное поле можно представить как суперпозицию полей двух мод, по одной от каждого волновода, плюс малое дополнительное поле. Так как поле моды второго волновода необходимо для описания искажения поля первого волновода, то ясно, что обе моды взаимодействуют. В этой главе будет показано, что полная передача мощности возможна с помощью этого типа связи. Для этого требуется, чтобы фазовые скорости мод в волноводах при отсутствии связи были одинаковыми. Моды с разными фазовыми скоростями не могут взаимодействовать эффективно, они могут обмениваться лишь очень малой мощностью.

Связь мод двух различных диэлектрических волноводов имеет два важных приложения. Нежелательным явлением является перекрестная связь (переходная помеха) между двумя передающими линиями [84] — [87]. Для передачи сигнала часто желательно иметь пучок из многих оптических волокон в одном кабеле. Однако, поскольку волокна находятся в непосредственной близости друг от друга, некоторые из них могут обмениваться мощностью, так что часть сигнала, передаваемая в одном волноводе, может перейти в соседний волновод и быть помехой распространяющемуся в нем сигналу. Оба отмеченных здесь типа связи оказываются нежелательными из-за того, что они вызывают переходную помеху. Однако связь между волноводами дает возможность передавать энергию от одного волновода к другому контролируемым образом, что можно использовать, например, для возбуждения волновода. Важным примером реализации связи такого типа является призмный ответвитель, который применяется для возбуждения направляемых мод в тонких диэлектрических пленках [107, 108]¹⁾. Хотя призмный ответвитель использует для возбуждения не моду волновода, а излучение от лазера, по принципу своей работы он имеет близкое отношение к механизму связи между регулярными волноводами. В обоих случаях имеет

¹⁾ Впервые такой возбудитель был предложен и исследован Йогансеном [98*]. — *Прим. ред.*

место «нарушение» полного внутреннего отражения от границы волновода, которое вызывает связь. Рассмотрим кратко призмный ответвитель. На фиг. 10.1.1 изображена его принципиальная схема.

Луч лазера падает сквозь призму на нижнюю ее грань, как показано на фигуре. При отсутствии волновода луч полностью отражается от поверхности раздела стекло — воздух. Однако из рассмотрения, проведенного в разд. 1.6,



Ф и г. 10.1.1. Схематическое изображение призмного ответвителя, используемого для возбуждения направляемых мод в тонких диэлектрических пленках.

известно, что экспоненциально спадающее поле проходит в воздушное пространство под призмой. Тонкопленочный волновод расположен в непосредственной близости от призмы, так что спадающее поле может достигать его с некоторой интенсивностью. Показатель преломления материала волновода меньше, чем у призмы ($n_2 < n_4$), так что если волноведущая среда достаточно протяженная, то направляемая ею волна будет возбуждаться через призму. Несмотря на то что среда с n_2 является тонкой пленкой, при возбуждении через призму в направляемую моду может быть передана от падающего поля значительная мощность, если фазовая скорость падающей волны вдоль грани призмы равна фазовой скорости направляемой моды. Таким образом можно передать в пленку $\sim 80\%$ падающей мощности. Если воздушный зазор между призм-

мой и тонкой пленкой сужается, то, как показано в [109, 121], может быть передано больше 90% мощности.

Связь с помощью спадающего поля такого рода обладает свойством направленности. Связь волноводов также можно использовать для передачи мощности от одного волновода к другому так, что в другом волноводе она будет распространяться только в том же направлении. Устройство, использующее такую связь, называется направленным ответвителем.

10.2. УРАВНЕНИЯ СВЯЗАННЫХ ВОЛН

Уравнения связанных волн [88] частного вида уже встречались в разд. 9.3 [уравнения (9.3.46), (9.3.47)] и в разд. 2.7 [уравнения (2.7.15), (2.7.16)]. Хотя вид уравнений связанных волн довольно очевиден и их можно получить интуитивно, приведем вывод этих уравнений для случая двух произвольных диэлектрических волноводов. Пусть показатель преломления является неоднородным только в узкой области и имеет постоянное значение вне волновода. Распределения квадратов показателей преломления двух волноводов показаны на фиг. 10.2.1. Координатная ось x соответствует поперечному направлению. Если оба волновода расположены рядом, квадрат показателя преломления общей для обоих волноводов среды выразится в виде

$$n^2 = (n_1^2 - n_3^2) + (n_2^2 - n_3^2) + n_3^2. \quad (10.2.1)$$

Величина n_3^2 имеет постоянное значение вне области волноводов. Поскольку $n_2^2 - n_3^2 = 0$ вне второго волновода, то формула (10.2.1) правильно описывает распределение квадрата показателя преломления около и внутри первого волновода. Аналогично поскольку $n_1^2 - n_3^2 = 0$ вне первого волновода, то формула (10.2.1) справедлива вблизи и внутри второго волновода.

Обозначим электромагнитное поле каждого волновода при отсутствии другого волновода индексами 1 и 2. Тогда составляющая электрического поля запишется как

$$E_v = \hat{E}_v e^{i(\omega t - \beta_v z)}, \quad v = 1, 2, \quad (10.2.2)$$

а составляющая магнитного поля — в виде

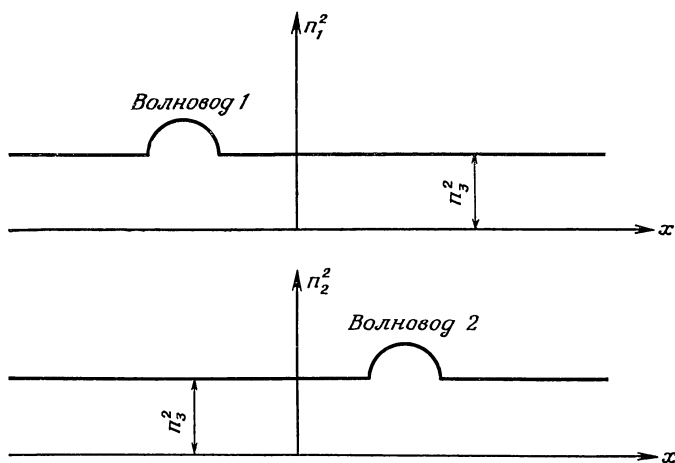
$$\mathbf{H}_v = \hat{\mathbf{H}}_v e^{i(\omega t - \beta_v z)}, \quad v = 1, 2. \quad (10.2.3)$$

Эти составляющие удовлетворяют уравнениям Максвелла в форме

$$\nabla_t \times \mathbf{H}_v - i\beta_v (\mathbf{z} \times \mathbf{H}_v) - i\omega \varepsilon_0 n_v^2 \mathbf{E}_v = 0, \quad (10.2.4)$$

$$\nabla_t \times \mathbf{E}_v - i\beta_v (\mathbf{z} \times \mathbf{E}_v) + i\omega \mu_0 \mathbf{H}_v = 0, \quad (10.2.5)$$

где $v = 1$ или 2 , а \mathbf{z} — единичный вектор в направлении координаты z , ∇_t — поперечная составляющая векторного оператора ∇ .



Ф и г. 10.2.1. Распределения показателей преломления в поперечном сечении для двух волноводов [103].

Если волноводы расположены вблизи друг от друга, то полное поле можно приближенно представить в виде суперпозиции невозмущенных полей каждого волновода. Однако, чтобы учесть влияние волноводов друг на друга, необходимо предположить возможность изменения амплитуд поля с расстоянием. Тогда полное поле выражается в виде

$$\mathbf{E} = A_1(z) \mathbf{E}_1 + A_2(z) \mathbf{E}_2 \quad (10.2.6)$$

и

$$\mathbf{H} = A_1(z) \mathbf{H}_1 + A_2(z) \mathbf{H}_2 \quad (10.2.7)$$

Эти соотношения не являются строгими. Для точного выражения полей необходимы дополнительные малые члены. Полные электрическое \mathbf{E} и магнитное \mathbf{H} поля удовлетворяют уравнениям Максвелла

$$\nabla \times \mathbf{H} = i\omega\epsilon_0 n^2 \mathbf{E} \quad (10.2.8)$$

и

$$\nabla \times \mathbf{E} = -i\omega\mu_0 \mathbf{H}, \quad (10.2.9)$$

в которых n^2 выражается формулой (10.2.1). Подстановка выражений (10.2.6) и (10.2.7) в (10.2.8) и (10.2.9) приводит к уравнениям

$$\begin{aligned} A_1 [\nabla_t \times \mathbf{H}_1 - i\beta_1 (\mathbf{z} \times \mathbf{H}_1)] + \frac{\partial A_1}{\partial z} (\mathbf{z} \times \mathbf{H}_1) - i\omega_0 n^2 A_1 \mathbf{E}_1 + \\ + A_2 [\nabla_t \times \mathbf{H}_2 - i\beta_2 (\mathbf{z} \times \mathbf{H}_2)] + \\ + \frac{\partial A_2}{\partial z} (\mathbf{z} \times \mathbf{H}_2) - i\omega\epsilon_0 n^2 A_2 \mathbf{E}_2 = 0 \end{aligned} \quad (10.2.10)$$

и

$$\begin{aligned} A_1 [\nabla_t \times \mathbf{E}_1 - i\beta_1 (\mathbf{z} \times \mathbf{E}_1)] + \frac{\partial A_1}{\partial z} (\mathbf{z} \times \mathbf{E}_1) + \\ + i\omega\mu_0 A_1 \mathbf{H}_1 + A_2 [\nabla_t \times \mathbf{E}_2 - i\beta_2 (\mathbf{z} \times \mathbf{E}_2)] + \\ + \frac{\partial A_2}{\partial z} (\mathbf{z} \times \mathbf{E}_2) + i\omega\mu_0 A_2 \mathbf{H}_2 = 0. \end{aligned} \quad (10.2.11)$$

Используя (10.2.1), (10.2.4) и (10.2.5), можно упростить эти уравнения, так что они принимают вид

$$\begin{aligned} \frac{\partial A_1}{\partial z} (\mathbf{z} \times \mathbf{H}_1) - i\omega\epsilon_0 (n_2^2 - n_3^2) A_1 \mathbf{E}_1 + \\ + \frac{\partial A_2}{\partial z} (\mathbf{z} \times \mathbf{H}_2) - i\omega\epsilon_0 (n_1^2 - n_3^2) A_2 \mathbf{E}_2 = 0 \end{aligned} \quad (10.2.12)$$

и

$$\frac{\partial A_1}{\partial z} (\mathbf{z} \times \mathbf{E}_1) + \frac{\partial A_2}{\partial z} (\mathbf{z} \times \mathbf{E}_2) = 0. \quad (10.2.13)$$

Следующим шагом при получении уравнений связанных волн является скалярное умножение уравнения (10.2.12) на \mathbf{E}_1 и аналогично уравнения (10.2.13) на \mathbf{H}_1 и вычитание одного из другого. Верхний индекс «—»

означает, что угловая частота ω и постоянная распространения β заменены в этих величинах на отрицательные значения. Это необходимо для устранения временной зависимости из уравнений. Если ограничиться рассмотрением действительных значений показателей преломления, то вместо \mathbf{E}_1^- и \mathbf{H}_1^- можно использовать комплексно-сопряженные значения полей. Однако мы предполагаем, что показатели преломления могут быть комплексными, что означает наличие потерь в средах, составляющих волновод. Проинтегрируем полученное уравнение по бесконечному поперечному сечению

$$\begin{aligned} \int_{-\infty}^{\infty} \int \left\{ \frac{\partial A_1}{\partial z} [\mathbf{E}_1^- \cdot (\mathbf{z} \times \mathbf{H}_1) - \mathbf{H}_1^- \cdot (\mathbf{z} \times \mathbf{E}_1)] + \right. \\ \left. + \frac{\partial A_2}{\partial z} [\mathbf{E}_1^- \cdot (\mathbf{z} \times \mathbf{H}_2) - \mathbf{H}_1^- \cdot (\mathbf{z} \times \mathbf{E}_2)] - \right. \\ \left. - i\omega\epsilon_0 (n_2^2 - n_3^2) A_1 \mathbf{E}_1^- \cdot \mathbf{E}_1 - \right. \\ \left. - i\omega\epsilon_0 (n_1^2 - n_3^2) A_2 \mathbf{E}_1^- \cdot \mathbf{E}_2 \right\} dx dy = 0. \quad (10.2.14) \end{aligned}$$

Можно упростить это уравнение, пренебрегая малыми членами. Выражение $n_2^2 - n_3^2$ равно нулю вне области второго волновода. В области же, где это выражение отлично от нуля, поле первого волновода всегда очень слабое. Содержащий это выражение член умножается на квадрат \mathbf{E}_1 . В результате получается величина второго порядка малости, которой можно пренебречь. Произведение $\mathbf{E}_1^- \cdot (\mathbf{z} \times \mathbf{H}_1)$ является членом нулевого порядка. Произведение $(n_1^2 - n_3^2) \mathbf{E}_1^- \cdot \mathbf{E}_2$ есть малая величина первого порядка, так как разность квадратов показателей преломления не равна нулю в области первого волновода, где \mathbf{E}_1 величина нулевого, а \mathbf{E}_2 — величина первого порядка малости. Такое сравнение порядков величин наводит на мысль, что производные $\partial A / \partial z$ являются величинами первого порядка малости. Поскольку произведение $\mathbf{E}_1^- \cdot (\mathbf{z} \times \mathbf{H}_2)$ есть малая величина первого порядка, из-за того, что поля волноводов перекрываются лишь незначительно, член с $\partial A_2 / \partial z$ есть малая величина второго порядка и им можно пренебречь. Сохраняя только члены

первого порядка, получим уравнение

$$\frac{\partial A_1}{\partial z} = ic_1 A_2 e^{i(\beta_1 - \beta_2)z}. \quad (10.2.15)$$

Аналогично, умножая (10.2.12) и (10.2.13) на E_2^- и H_2^- , получим уравнение

$$\frac{\partial A_2}{\partial z} = ic_2 A_1 e^{-(\beta_1 - \beta_2)z}, \quad (10.2.16)$$

где

$$c_1 = -\omega \varepsilon_0 \frac{\int_{-\infty}^{\infty} \int (n_1^2 - n_2^2) \hat{E}_1^- \cdot \hat{E}_2 \, dx \, dy}{\int_{-\infty}^{\infty} \int \mathbf{z} \cdot (\hat{E}_1^- \times \hat{H}_1 + \hat{E}_1 \times \hat{H}_1^-) \, dx \, dy} \quad (10.2.17)$$

и

$$c_2 = -\omega \varepsilon_0 \frac{\int_{-\infty}^{\infty} \int (n_2^2 - n_1^2) \hat{E}_2^- \cdot \hat{E}_1 \, dx \, dy}{\int_{-\infty}^{\infty} \int \mathbf{z} \cdot (\hat{E}_2^- \times \hat{H}_2 + \hat{E}_2 \times \hat{H}_2^-) \, dx \, dy}. \quad (10.2.18)$$

Коэффициенты связи (10.2.17) и (10.2.18) не зависят от z . Чтобы подчеркнуть, что член $\exp(\pm i\beta_{\nu} z)$ был отброшен, мы использовали запись вида \hat{E}_1 , \hat{E}_1^- и т. д., введенную в (10.2.2) и (10.2.3). Напомним, что величины с отрицательными верхними индексами получаются при замене знака ω и β . При переходе от (10.2.14) к (10.2.15)–(10.2.18) использовалось то обстоятельство, что амплитуды A_1 и A_2 не зависят от поперечных координат x и y и поэтому их можно вынести из-под знака интеграла. Члены со смешанными скалярно-векторными произведениями преобразовывались с помощью хорошо известного векторного тождества.

Как показало исследование, уравнения связанных волн (10.2.15) и (10.2.16) не являются точными. В добавление к тому, что мы пренебрегли членами второго порядка малости, мы ограничили еще рассмотрение только двумя модами. Даже если два волновода расположены близко друг к другу и поддерживают только одну направляемую моду, все равно сохраняется возможность существования связи с модами излучения. Если волноводы много-

модовые, то все моды в некоторой степени связаны между собой [85—87]. Таким образом, при учете в уравнениях связанных волн только двух мод мы получаем приближение. Однако вскоре мы увидим, что только моды с одинаковыми фазовыми постоянными распространения могут обмениваться значительной величиной энергии. Ограничиваясь в уравнениях связанных волн только двумя модами, получаем фактически очень хорошее приближение, позволяющее изучать обмен энергией между двумя модами с высокой точностью. Мы уже сталкивались с подобной ситуацией в разд. 9.3.

Уравнения связанных волн часто записывают в несколько другом виде. Введя амплитуды волны

$$a_v = A_v e^{-i\beta_v z}, \quad v=1, 2, \quad (10.2.19)$$

формулы (10.2.15) и (10.2.16) можно записать в достаточно известной форме [88]:

$$\frac{\partial a_1}{\partial z} = -i\beta_1 a_1 + ic_1 a_2 \quad (10.2.20)$$

и

$$\frac{\partial a_2}{\partial z} = -i\beta_2 a_2 + ic_2 a_1. \quad (10.2.21)$$

Эти уравнения имеют такой ясный физический смысл, что их можно было бы записать без вывода. Однако наше рассмотрение имеет преимущество в том, что значения коэффициентов связи определяются достаточно точно. Поскольку коэффициенты $n_1(x, y)$ и $n_2(x, y)$, как и постоянный показатель преломления окружающей среды n_3 , могут быть комплексными, то коэффициенты связи c_1 и c_2 в общем случае также являются комплексными. В случае сред без потерь показатели преломления вещественные и выражения (10.2.17) и (10.2.18) можно упростить. Для вещественных значений n_1 , n_2 и n_3 вместо E_1^- можно использовать комплексно-сопряженную величину E_1^{*+} ¹⁾. Выражение в знаменателе можно интерпретировать как $4P$,

¹⁾ Из приведенного выше вывода не так очевидно, что составляющая E_1^- использована здесь вместо E_1^{*+} . Этот результат более наглядно представлен в [103]. Он следует из того, что E_1^- должна быть решением уравнений Максвелла, в которых ω и β заменены их отрицательными значениями, а n_1 и n_2 остаются неизменными.

где P — мощность моды в волноводе 1 (при $A_1 = 1$). Если к тому же два волновода одинаковы (распределение n_1 в волноводе 1 такое же, как n_2 в волноводе 2), то вместо (10.2.17) и (10.2.18) получим

$$c_2 = -\frac{\omega \varepsilon_0}{4P} \int_{-\infty}^{\infty} \int (n_2^2 - n_3^2) \hat{\mathbf{E}}_2^* \hat{\mathbf{E}}_1 dx dy \quad (10.2.22)$$

и

$$c_1 = c_2. \quad (10.2.23)$$

Легко показать, что для вещественных значений c_1 и c_2 условие (10.2.23) представляет собой требование сохранения мощности, переходящей из моды в моду. С помощью формул (10.2.20) и (10.2.21) получаем соотношение

$$\frac{d}{dz} (|a_1|^2 + |a_2|^2) = 2 \operatorname{Re} [i (c_2 - c_1^*) a_1 a_2^*],$$

где Re — обозначает вещественную часть при условии, что β_1 и β_2 вещественные. Слева в этом соотношении стоит производная по z от полной мощности, переносимой в обоих волноводах. Она обращается в нуль, если мощность сохраняется. Поскольку $a_1(0)$ и $a_2(0)$ можно выбрать произвольно, получим следующее условие сохранения мощности:

$$c_1 = c_2^*. \quad (10.2.24)$$

Равенства (10.2.24) и (10.2.23) совпадают для вещественных значений c_1 и c_2 .

Легко видеть, что значительная часть мощности передается из волновода в волновод, если только $\beta_1 = \beta_2$. Предположим, что $A_2 = 0$ при $z = 0$. Тогда из формулы (10.2.16) получим

$$A_2(L) = ic_2 \int_0^L A_1(z) e^{-i(\beta_1 - \beta_2)z} dz. \quad (10.2.25)$$

Если $\beta_1 - \beta_2 \neq 0$, функция $A_1(z)$ умножается на $\cos(\beta_1 - \beta_2)z$ и $\sin(\beta_1 - \beta_2)z$. Обе функции осциллируют и уменьшают интеграл в (10.2.25). Однако, если $\beta_1 - \beta_2 = 0$, интеграл становится пропорциональным L (по крайней мере в начале участка, когда $A_1(z)$ еще не изменится существенно). Такое рассмотрение показывает,

что величина $A_2(z)$ может быть значительной, если постоянные распространения обеих мод одинаковы. Это является причиной того, что теория связанных волн, в которой рассматриваются только две моды, применима даже в случае многомодовых волноводов, поскольку значительный обмен мощностью возможен только между модами с равными фазовыми скоростями. Однако, если коэффициенты связи не постоянны и зависят от координаты z , двухмодовая теория не применима к многомодовым волноводам.

Уравнение связанных волн с постоянными коэффициентами имеет следующее простое решение при $\beta_1 = \beta_2 = \beta$:

$$a_1(z) = \frac{1}{2} \left\{ a_1(0) [e^{i\Delta\beta z} + e^{-i\Delta\beta z}] + \right. \\ \left. + \sqrt{\frac{c_1}{c_2}} a_2(0) [e^{i\Delta\beta z} - e^{-i\Delta\beta z}] \right\} e^{-i\beta z} \quad (10.2.26)$$

и

$$a_2(z) = \frac{1}{2} \left\{ a_2(0) [e^{i\Delta\beta z} + e^{-i\Delta\beta z}] + \right. \\ \left. + \sqrt{\frac{c_2}{c_1}} a_1(0) [e^{i\Delta\beta z} - e^{-i\Delta\beta z}] \right\} e^{-i\beta z}, \quad (10.2.27)$$

где

$$\Delta\beta = \sqrt{c_1 c_2}. \quad (10.2.28)$$

Члены в этом решении сгруппированы так, чтобы лучше были видны первоначальные значения a_1 и a_2 при $z = 0$. Перегруппировав члены, получим, что решения представляют собой суперпозицию двух новых мод с фазовыми постоянными

$$\beta_+ = \beta + \Delta\beta \quad (10.2.29)$$

и

$$\beta_- = \beta - \Delta\beta. \quad (10.2.30)$$

Два связанных волновода, таким образом, обладают нормальными модами с несколько измененными постоянными распространения (10.2.29) и (10.2.30). Если коэффициенты связи вещественны, то $\Delta\beta$ — также вещественная величина. Полагая для простоты $a_2(0) = 0$, из формул (10.2.26) и (10.2.27) с использованием (10.2.24) получим

$$a_1(z) = a_1(0) \cos(\Delta\beta z) e^{-i\beta z} \quad (10.2.31)$$

и

$$a_2(z) = ia_1(0) \sin(\Delta\beta z) e^{-i\beta z}. \quad (10.2.32)$$

Это решение ясно показывает, что при вещественных значениях $\beta_1 = \beta_2$ происходит непрерывный обмен энергией между двумя волноводами. Если, однако, величина $\Delta\beta$ не является вещественной, то $\exp(-i\Delta\beta z)$ уменьшается, а $\exp(+i\Delta\beta z)$ возрастает (или наоборот). Для достаточно больших значений z один из двух множителей становится пренебрежимо малым. Тогда имеем

$$a_1(z) = \frac{1}{2} a_1(0) e^{-i(\beta + \Delta\beta)z} \quad (10.2.33)$$

и

$$a_2(z) = \frac{1}{2} \sqrt{\frac{c_2}{c_1}} a_1(0) e^{-i(\beta + \Delta\beta)z} \quad (10.2.34)$$

при условии, что $a_2(0) = 0$. Хотя обе амплитуды мод уменьшаются, при комплексном $\Delta\beta$ оба волновода перенесут равные количества мощности (при условии, что c_1/c_2 не отличается существенно от единицы) после прохождения модами достаточно большого расстояния, несмотря на то, что первоначально возбуждался только один из волноводов. Периодический обмен мощностью при этом прекращается и мощности в обоих волноводах выравниваются.

С целью изучения перекрестной связи достаточно рассмотреть передачу небольшой мощности, так как эта связь обычно является нежелательной и всегда ее стараются сделать по возможности малой. Сначала предположим, что первоначально вся мощность сосредоточена в волноводе 1, $a_2(0) = 0$ и $\Delta\beta z \ll 1$ на конце участка связи волноводов при $z = L$. Из формулы (10.2.27) находим

$$\frac{a_2(L)}{a_1(0) e^{-i\beta L}} = i \sqrt{\frac{c_2}{c_1}} \Delta\beta L. \quad (10.2.35)$$

Из выражения (10.2.26) получаем в том же приближении

$$a_1(L) = a_1(0) e^{-i\beta L}. \quad (10.2.36)$$

Мощность в каждом волноводе пропорциональна квадрату амплитуды поля. Отношение мощностей на концах двух волноводов равно

$$C = \frac{|a_2(L)|^2}{|a_1(L)|^2} = \left| \frac{c_2}{c_1} \right| (|\Delta\beta| L)^2. \quad (10.2.37)$$

В большинстве практических случаев можно считать, что соотношение (10.2.24) выполняется по крайней мере приближенно, так что окончательно получаем достаточно хорошее приближение

$$C = |\Delta\beta|^2 L^2. \quad (10.2.38)$$

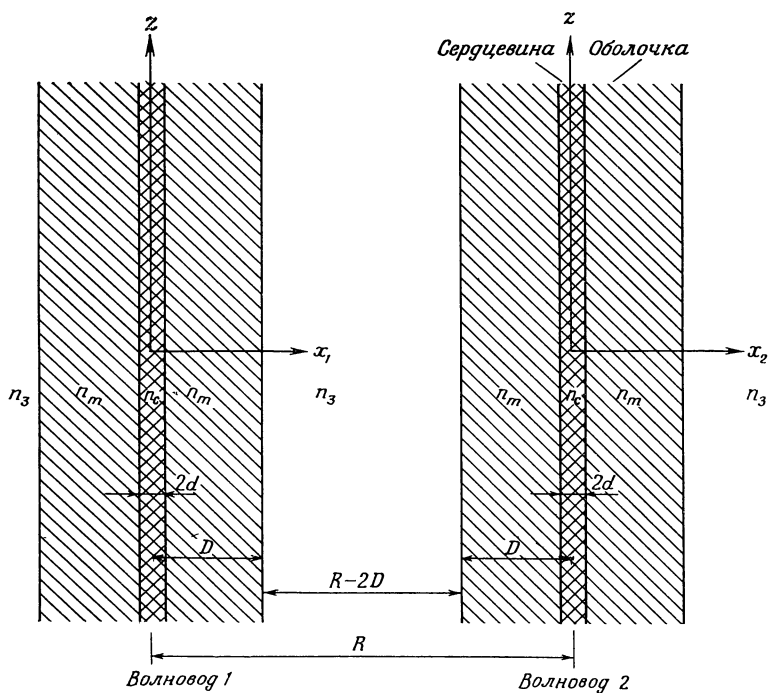
Важно помнить, что это соотношение справедливо только при $C \ll 1$.

Итак, в этом разделе показано, что два диэлектрических волновода влияют друг на друга, если они располагаются достаточно близко друг от друга. Связь между волноводами можно выразить с помощью двух уравнений связанных волн с постоянными коэффициентами связи. Связь вызывается тем, что часть поля направляемой моды одного волновода достигает другого волновода. Рассматриваемая теория связи справедлива для любого типа диэлектрических волноводов независимо от распределения показателя преломления по сечению волновода. Не обязательны также идентичность волноводов и мод в них, а также допустимо наличие потерь в диэлектрической среде волновода и в окружающей среде. Единственное требование для осуществления связи состоит в том, чтобы постоянные распространения двух мод были почти одинаковыми. Если в каждом волноводе может существовать только одна мода, это требование можно значительно ослабить. Однако моды с разными постоянными распространения не связаны эффективно при постоянных коэффициентах связи. Таким образом, действительно интересным является случай, когда обе моды имеют одинаковые фазовые скорости.

10.3. СВЯЗАННЫЕ ПЛОСКИЕ ВОЛНОВОДЫ

Применим теперь общую теорию связи к двум конкретным случаям. В этом разделе рассмотрим связь двух одинаковых параллельных плоских волноводов, изображенных на фиг. 10.3.1. Коэффициенты связи для задачи двух связанных одинаковых мод можно найти, используя формулу (10.2.22) из предыдущего раздела. Применение формулы (10.2.22) вместо (10.2.17) допустимо, если считать, что показатели преломления сердцевины и оболочки n_c и n_m — величины вещественные. Окружающая среда

может иметь комплексный показатель преломления n_3 , но для того чтобы можно было применять формулу (10.2.22), необходимо предположить, что оболочка является достаточно толстой и потери моды, вызванные окружающей средой, незначительны. Это вполне естественное предположение, так как волноводы целесообразно использовать



Ф и г. 10.3.1. Схематическое изображение двух связанных плоских волноводов в оболочке [103].

тогда, когда они направляют волны с низкими потерями.

Проблема связи волноводов может быть решена более прямым путем. Сначала можно решить задачу о направляемых модах объединенной структуры (фиг. 10.3.1). Для волноводов в оболочке такой подход является трудно осуществимым. Однако в предельном случае при $n_m = n_3$ этим способом легко решить задачу о модах комбинированной

структуры. Читателю предлагается проделать такой расчет в качестве упражнения. С приемлемыми приближениями такой прямой метод приводит к выражению (10.3.12).

Изложим кратко решение задачи связи двух плоских волноводов в оболочке. Даже с помощью формулы (10.2.22) эту задачу трудно решить строго [при рассмотрении величины (10.2.22) как точной]. Однако если допустить, что поля мод спадают до малых значений на границе оболочки при $x = D$, то можно существенно упростить теорию. Начнем с рассмотрения ТЕ-мод плоского волновода. Составляющая электрического поля выражается в виде

$$\hat{E}_y = \begin{cases} A \cos \kappa x, & 0 \leq x \leq d, \\ Be^{\gamma x} + Ce^{-\gamma x}, & d \leq x \leq D, \\ Fe^{-\rho x}, & D \leq x < \infty. \end{cases} \quad (10.3.1)$$

Функция $\hat{E}_y(x)$ предполагается непрерывной и симметричной относительно нуля. Параметры, входящие в (10.3.1), равны

$$k_0 = \omega \sqrt{\varepsilon_0 \mu_0}, \quad (10.3.2)$$

$$\kappa = \sqrt{n_c^2 k_0^2 - \beta^2}, \quad (10.3.3)$$

$$\gamma = \sqrt{\beta^2 - n_m^2 k_0^2}, \quad (10.3.4)$$

$$\rho = \sqrt{\beta^2 - n_s^2 k_0^2}. \quad (10.3.5)$$

Соотношение между \hat{E}_y и E_y задается формулой (10.2.2). Составляющая магнитного поля получается из (10.3.1) с помощью формул (8.3.2) и (8.3.3). Теперь положим

$$e^{-\gamma(D-d)} \ll 1. \quad (10.3.6)$$

Пренебрегая квадратами этих величин, получим ¹⁾

$$B = 0, \quad (10.3.7)$$

$$C = Ae^{\gamma d} \cos \kappa d, \quad (10.3.8)$$

$$F = \frac{2\kappa\gamma A}{(\gamma + \rho) \sqrt{\kappa^2 + \gamma^2}} e^{\rho D} e^{-\gamma(D-d)}. \quad (10.3.9)$$

¹⁾ Приближение $B = 0$ было сделано после вычисления F .

Коэффициент A можно выразить через мощность P , переносимую модой:

$$A = \sqrt{\frac{2\omega\mu_0 P}{\beta d + \frac{\beta}{\gamma}}}. \quad (10.3.10)$$

Уравнение собственных значений моды плоского волновода в оболочке отличается от (8.3.16) лишь членом, пропорциональным квадрату (10.3.6). В духе нашего приближения можно считать величину (8.3.16) неизменной. Поле моды (10.3.1) выражено в системе координат с центром в сердцевине каждого волновода.

Поскольку согласно формулам (10.2.23) и (10.2.28) $\Delta\beta = c_1$, то непосредственное вычисление интеграла из (10.2.22) дает

$$\Delta\beta = \frac{4\kappa^2\gamma^3\rho}{\beta(1+\gamma d)(\gamma+\rho)^2(\kappa^2+\gamma^2)} e^{-2\gamma(D-d)} e^{-\rho(R-2D)}. \quad (10.3.11)$$

Показатели преломления сердцевины и оболочки предполагаются вещественными, но n_3 и, следовательно, ρ могут быть комплексными. Это дает возможность изучить развязку двух волноводов, которая возникает из-за потерь в окружающей среде.

В частном случае, когда оболочка и окружающая среда имеют одинаковые показатели преломления, $n_m = n_3$, получим, что $\gamma = \rho$ и соотношение (10.3.11) принимает вид

$$\Delta\beta = \frac{\kappa^2\gamma^2}{\beta(1+\gamma d)(\kappa^2+\gamma^2)} e^{-\gamma(R-2d)}. \quad (10.3.12)$$

Как упоминалось ранее, это выражение можно получить непосредственно из решения задачи о модах комбинированной структуры, составленной из плоских волноводов. Моды комбинированной структуры имеют согласно формулам (10.2.29) и (10.2.30) малоотличающиеся постоянные распространения. Разность этих постоянных представлена в (10.3.12). Такое же выражение можно найти в книге Капани [79] ¹⁾.

¹⁾ В [100*] дано строгое численное решение этой задачи с применением ЭВМ. *Прим. ред.*

Так как связь между двумя волноводами можно уменьшать, увеличивая потери в окружающей среде, то важно знать влияние этих внешних потерь на потери направляемых мод. Потери направляемой моды можно вычислить с помощью простого метода теории возмущений, аналогичного тому, который применялся в разд. 9.6 для вычисления потерь на изгибе. Метод очень прост, ограничимся его кратким изложением и приведем конечный результат.

Выражения для поля (10.3.1) позволяют вычислить x — составляющую вектора Пойнтинга S_x на границе оболочки при $x = D$. Потери мощности, таким образом, определяются выражением

$$2\alpha = \frac{2S_x}{P}, \quad (10.3.13)$$

где α — амплитудный коэффициент потерь. Множитель 2 в левой части необходим, так как 2α определяет потери мощности, множитель 2 в правой части отражает тот факт, что мощность выходит по обе стороны волновода. Наличие второго волновода не учитывается. Его влияние незначительно, так как потери мощности вызваны излучением, которое проходит через второй волновод почти без отражения. Однако даже малая отраженная мощность практически не возвращается к моде первого волновода. Несложное вычисление дает результат в виде

$$2\alpha = \frac{8\kappa^2 \gamma^3 \operatorname{Im}(\rho) e^{-2\gamma(D-d)}}{\beta(1+\gamma d)(\kappa^2 + \gamma^2) |\gamma + \rho|^2}. \quad (10.3.14)$$

Выражение (10.3.14) отражает несколько интересных свойств затухания моды. Видно, что затухание экспоненциально уменьшается с увеличением толщины оболочки. Таким образом, потери можно сделать произвольно малыми, выполняя оболочку достаточно толстой. Кроме того, затухание моды зависит от мнимой части ρ . Если ρ действительное, потери моды отсутствуют. Напомним, что это имеет место в случае отсутствия потерь в материалах сердцевин и оболочки. Наличие потерь в этих материалах вызывает дополнительные потери моды. Параметр ρ может быть действительным, когда величина n_3 действи-

тельная и, кроме того, $\beta > n_3 k_0$. Но если

$$\beta < n_3 k_0, \quad (10.3.15)$$

то параметр ρ мнимый и моды испытывают потери даже при действительном n_3 . Этот случай является другим примером «нарушения» полного внутреннего отражения. Экспоненциально уменьшающееся поле в оболочке вне сердцевины, вызванное полным внутренним отражением на границе раздела сердцевина — оболочка, переходит в поле излучения, если показатель преломления окружающей среды достаточно велик. Для малых значений n_3 потери моды определяются потерями в окружающей среде.

Аналогичные выражения можно получать для ТМ-мод плоского волновода. Моды ТЕ и ТМ не связаны между собой, так как скалярное произведение их электрических полей обращается в нуль. Изменение постоянной распространения, вызванное связью двух симметричных ТМ-мод плоских волноводов, есть

$$\Delta\beta = \frac{n_c^2 n_m^2 \kappa^2 \gamma^2 [2e^{(\gamma-\rho)(D-d)} - 1] e^{-2\gamma(D-d)} e^{-\rho(R-2D)}}{\beta [(n_m^4 \kappa^2 + n_c^4 \gamma^2) \gamma d + n_c^2 n_m^2 (\kappa^2 + \gamma^2)]}. \quad (10.3.16)$$

Видно, что связь для ТМ- и ТЕ-мод не одинакова. Только в пределе при $n_c \approx n_m \approx n_3$ выражения (10.3.11) и (10.3.16) совпадают.

Выражение для потерь ТМ-мод в окружающей среде имеет вид

$$2\alpha = \frac{8n_c^2 n_m^4 |n_3|^4 \kappa^2 \gamma^3 \operatorname{Im} \left(\frac{\rho}{n_3^2} \right) e^{-2\gamma(D-d)}}{\beta [\gamma d (n_m^4 \kappa^2 + n_c^4 \gamma^2) + n_c^2 n_m^2 (\kappa^2 + \gamma^2)] |n_m^2 \rho + n_3^2 \gamma|^2}. \quad (10.3.17)$$

10.4. СВЯЗЬ МОД HE_{11} ДВУХ КРУГЛЫХ ВОЛОКОН

Аналогично тому, как был исследован в предыдущем разделе случай плоских волноводов, кратко рассмотрим здесь связь между модами HE_{11} двух круглых оптических волокон.

Схематически два волокна представлены на фиг. 10.4.1.

Выражения для поля моды волокна можно получить из выражений для продольных составляющих электри-

ческого и магнитного полей внутри сердцевины

$$\begin{aligned}\hat{E}_z &= A J_1(\kappa r) \cos \phi, \\ \hat{H}_z &= B J_1(\kappa r) \sin \phi,\end{aligned}\quad (10.4.1)$$

где $0 \leq r \leq a$.

В оболочке имеем

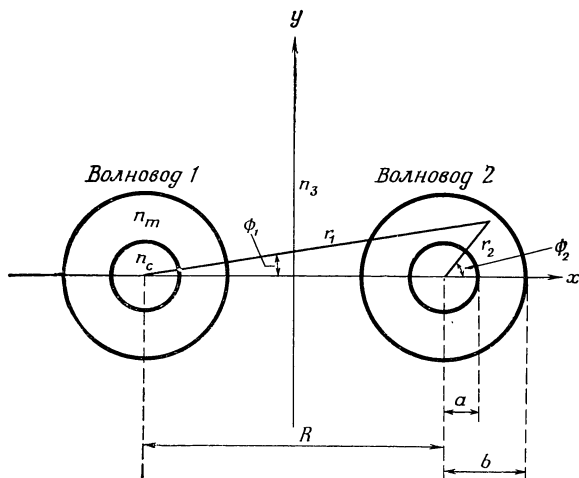
$$\begin{aligned}\hat{E}_z &= [C H_1^{(1)}(i\gamma r) + D H_1^{(2)}(i\gamma r)] \cos \phi, \\ \hat{H}_z &= [F H_1^{(1)}(i\gamma r) + G H_1^{(2)}(i\gamma r)] \sin \phi,\end{aligned}\quad (10.4.2)$$

где $a \leq r \leq b$,

и, наконец, составляющие поля вне оболочки (в окружающей среде) определяются выражениями

$$\begin{aligned}\hat{E}_z &= M H_1^{(1)}(i\gamma r) \cos \phi, \\ \hat{H}_z &= N H_1^{(1)}(i\gamma r) \sin \phi,\end{aligned}\quad (10.4.3)$$

где $b \leq r < \infty$.



Ф и г. 10.4.1. Поперечное сечение двух связанных круглых оптических волокон в оболочке [103].

Остальные составляющие поля получаются из формул (8.2.7) — (8.2.10), в которые вместо κ^2 подставляются κ^2 , $-\gamma^2$ или $-\rho^2$. В таком же приближении, в каком

рассматривался плоский волновод, коэффициенты B , C и F оказываются связанными соотношениями

$$B = \sqrt{\frac{\varepsilon_0}{\mu_0} \frac{a^2 k_0 \kappa^2 \gamma^2}{\beta (\kappa^2 + \gamma^2)}} \left\{ \frac{n_c^2}{\kappa a} \left[\frac{1}{\kappa a} - \frac{J_0(\kappa a)}{J_1(\kappa a)} \right] + \frac{n_m^2}{\gamma a} \left[\frac{1}{\gamma a} - \frac{i H_0^{(1)}(i \gamma a)}{H_1^{(1)}(i \gamma a)} \right] \right\} A \quad (10.4.4)$$

и

$$\frac{C}{A} = \frac{F}{B} = \frac{J_1(\kappa a)}{H_1^{(1)}(i \gamma a)}. \quad (10.4.5)$$

Считая, что поле на границе раздела между оболочкой и окружающей средой мало, имеем ¹⁾

$$D = G = 0. \quad (10.4.6)$$

Приближенные выражения для коэффициентов M и N получаются заменой функций Ханкеля их приближениями для большого аргумента

$$M = \frac{\left(\frac{\rho}{\gamma} \right)^{3/2} e^{(\rho - \gamma)b} C \left\{ \left[\left(\frac{\beta}{b} \right)^2 \left(1 - \frac{\gamma^2}{\rho^2} \right)^2 \left(\frac{\gamma}{\rho} - 1 \right) - k_0^2 \gamma^2 \left(1 + \frac{\gamma}{\rho} \right) \times \right. \right. \\ \left. \left. \times \left(n_m^2 + n_3^2 \frac{\gamma^2}{\rho^2} + \frac{\gamma}{\rho} (n_m^2 - n_3^2) \right) \right] + 2 \omega \mu_0 F \gamma \frac{\beta}{b} \left(1 - \frac{\gamma^2}{\rho^2} \right) \right\}}{\left[\frac{\beta}{b} \left(1 - \frac{\gamma^2}{\rho^2} \right) \right]^2 - k_0^2 \gamma^2 \left(1 + \frac{\gamma}{\rho} \right) \left(n_m^2 + n_3^2 \frac{\gamma}{\rho} \right)} \quad (10.4.7)$$

и

$$N = - \sqrt{\frac{\rho}{\gamma}} e^{(\rho - \gamma)b} \frac{2 F k_0^2 \gamma^2 \left(n_m^2 + n_3^2 \frac{\gamma}{\rho} \right) + 2 \omega n_m^2 \varepsilon_0 \gamma \frac{\beta}{b} \left(\frac{\gamma^2}{\rho^2} - 1 \right) C}{\left[\frac{\beta}{b} \left(1 - \frac{\gamma^2}{\rho^2} \right) \right]^2 - k_0^2 \gamma^2 \left(1 + \frac{\gamma}{\rho} \right) \left(n_m^2 + n_3^2 \frac{\gamma}{\rho} \right)}. \quad (10.4.8)$$

Коэффициент A связан с мощностью, переносимой модой, соотношением

$$P = \frac{\pi}{4} \sqrt{\frac{\varepsilon_0}{\mu_0}} A^2 \left\{ \frac{k_0 \beta}{\kappa^4} [(a \kappa)^2 (J_0^2(\kappa a) + J_1^2(\kappa a)) - 2 J_1^2(\kappa a)] \times \right. \\ \left. \times \left(n_c^2 + \frac{\mu_0}{\varepsilon_0} \frac{B^2}{A^2} \right) + \frac{k_0 \beta}{\gamma^4} \times \right.$$

¹⁾ Так же как и в случае плоского волновода, при вычислении M и N можно положить $D = G = 0$.

$$\times \left[(a\gamma)^2 \left(1 - \left(\frac{iH_0^{(1)}(i\gamma a)}{H_1^{(1)}(i\gamma a)} \right)^2 \right) + 2 \right] J_1^2(\kappa a) \left(n_m^2 + \frac{\mu_0}{\varepsilon_0} \frac{B^2}{A^2} \right) + \\ + 2 \sqrt{\frac{\mu_0}{\varepsilon_0}} \frac{B}{A} \left[\frac{\beta_0^2 + n_c^2 k_0^2}{\kappa^4} - \frac{\beta_0^2 + n_m^2 k_0^2}{\gamma^4} \right] J_1^2(\kappa a) \}. \quad (10.4.9)$$

Параметры κ , γ и ρ снова определяются выражениями (10.3.3), (10.3.4) и (10.3.5). Для приближенных расчетов можно использовать здесь уравнение собственных значений (8.2.49), так как уравнение собственных значений волновода с оболочкой для моды HE_{11} отличается от уравнения (8.2.49) для случая бесконечно протяженной оболочки только членом, который представляет собой квадрат от (10.3.6).

Вычисление интеграла в формуле (10.2.22) значительно сложнее, чем для случая плоского волновода. Этот интеграл не может быть вычислен точно. Используя систему координат с началом на оси второго волновода, можно записать соотношение (10.2.22) в виде

$$\Delta\beta = -\frac{\omega\varepsilon_0}{4P} \int_0^b r_2 dr_2 \int_0^{2\pi} (n_2^2 - n_3^2) \hat{\mathbf{E}}_2^* \cdot \hat{\mathbf{E}}_1 d\phi_2. \quad (10.4.10)$$

Так как поле каждого волновода записывается в системе координат с началом на его оси, необходимо преобразовать поля к общей для обоих волноводов системе координат, прежде чем пытаться вычислять скалярное произведение векторов электрического поля. Запишем составляющие электрического поля в системе координат каждого волновода в виде [ср. с формулами (8.6.60), (8.6.62) и (8.6.63), множитель i включен в функцию F]

$$\begin{aligned} \hat{E}_{zv} &= F_{zv}(r_v) \cos \phi_v, \\ \hat{E}_{rv} &= F_{rv}(r_v) \cos \phi_v, \\ \hat{E}_{\phi v} &= F_{\phi v}(r_v) \sin \phi_v. \end{aligned} \quad (10.4.11)$$

Индекс v принимает значения 1 или 2 в соответствии с волноводом, которому принадлежит составляющая поля или начало координат. Цель записи (10.4.11) состоит в том, чтобы выделить зависимость каждой составляющей поля от r и ϕ . Можно показать, что скалярное произведение двух

векторов поля при таком расположении двух волноводов, как показано на фиг. 10.4.1, выражается следующим образом:

$$\begin{aligned}\hat{\mathbf{E}}_2^* \cdot \hat{\mathbf{E}}_1 = & F_{z1}(r_1) F_{z2}^*(r_2) (\cos \phi_1 \cos \phi_2 + \\ & + [F_{r1}(r_1) F_{r2}^*(r_2) \cos \phi_1 \cos \phi_2 + \\ & + F_{\phi 1}(r_1) F_{\phi 2}^*(r_2) \sin \phi_1 \sin \phi_2] \cos(\phi_2 - \phi_1) + \\ & + [F_{\phi 1}(r_1) F_{r2}^*(r_2) \sin \phi_1 \cos \phi_2 - \\ & - F_{r1}(r_1) F_{\phi 2}^*(r_2) \cos \phi_1 \sin \phi_2] \sin(\phi_2 - \phi_1). \quad (10.4.12)\end{aligned}$$

Для того чтобы вычислить интегралы в (10.4.10), предположим, что два волновода расположены далеко друг от друга. Оказывается, что получающееся в этом случае решение нашей задачи дает хорошие результаты, даже когда волноводы находятся достаточно близко. Приближения, необходимые при вычислении интеграла, основаны на предположении, что $R \gg a$. Первое следствие этого допущения состоит в том, что угол ϕ_1 мало отличается от нуля. Отсюда ясно, что только те области интегрирования дают значительный вклад, где ϕ_1 мало. Это связано с тем, что поле моды сильно сконцентрировано вблизи сердцевин волновода. Даже если оболочка достаточно толстая, область, близкая к сердцевине волновода, дает более значительный вклад в интеграл, поскольку в ней сосредоточено основное поле. Поэтому можно положить $\phi_1 = 0$. Это допущение сводит скалярное произведение (10.4.12) к более простому виду

$$\begin{aligned}\hat{\mathbf{E}}_2^* \cdot \hat{\mathbf{E}}_1 = & F_{z1}(r_1) F_{z2}^*(r_2) \cos \phi_2 + \\ & + F_{r1}(r_1) [F_{r2}^*(r_2) \cos^2 \phi_2 - F_{\phi 2}^*(r_2) \sin^2 \phi_2]. \quad (10.4.13)\end{aligned}$$

Следующий шаг приближенного вычисления интеграла состоит в использовании приближенного выражения для радиуса r_1

$$r_1 = \sqrt{R^2 + r_2^2 + 2r_2 R \cos \phi_2} \approx R + r_2 \cos \phi_2. \quad (10.4.14)$$

Функции $F(r_1)$ в (10.4.13), согласно формулам (10.4.3) и (8.2.7), являются функцией Ханкеля первого порядка и ее производной. Так как or_1 велико, можно использовать приближение функций Ханкеля для больших значений

аргументов с учетом формулы (10.4.14):

$$H_1^{(1)}(i\rho r_1) \approx -\sqrt{\frac{2}{\pi\rho r_1}} e^{-\rho r_1} \approx -\sqrt{\frac{2}{\pi\rho R}} e^{-\rho R} e^{-\rho r_2 \cos \phi_2}, \quad (10.4.15)$$

$$H_1^{(1)'}(i\rho r_1) \approx iH_1^{(1)}(i\rho r_1). \quad (10.4.16)$$

Используя приближенное скалярное произведение векторов электрического поля и функций Ханкеля, можно в формуле (10.4.10) произвести интегрирование по ϕ_2 . При этом необходимо взять интегралы

$$\int_0^{2\pi} e^{-\rho r_2 \cos \phi_2} \cos \phi_2 d\phi_2 = 2\pi i J_1(i\rho r_2), \quad (10.4.17)$$

$$\begin{aligned} \int_0^{2\pi} e^{-\rho r_2 \cos \phi_2} \cos^2 \phi_2 d\phi_2 = \\ = 2\pi \left[J_0(i\rho r_2) + \frac{i}{\rho r_2} J_1(i\rho r_2) \right], \end{aligned} \quad (10.4.18)$$

и

$$\int_0^{2\pi} e^{-\rho r_2 \cos \phi_2} \sin^2 \phi_2 d\phi_2 = -\frac{2\pi i}{\rho r_2} J_1(i\rho r_2). \quad (10.4.19)$$

Оставшийся интеграл содержит только произведения цилиндрических функций, интегрирование которых не представляет особых трудностей. Таким образом, в результате получаем приближенную оценку выражения (10.4.10) в виде

$$\begin{aligned} \Delta\beta_h = \sqrt{\frac{\varepsilon_0}{\mu_0}} \frac{A^2}{2P} \frac{M}{A} \frac{1}{k_0} \sqrt{\frac{2\pi}{\rho R}} e^{-\rho R} \left\{ \frac{e^{(\rho-\gamma)b} J_1(\kappa a)}{\pi \sqrt{\rho\gamma} H_1^{(1)}(i\gamma a)} \times \right. \\ \times \left[\left(1 + \frac{\beta^2}{\gamma\rho} \right) (\gamma + \rho) + \frac{\beta^2}{b\gamma^2\rho^2} (\rho^2 - \gamma^2) \left(1 - \frac{k_0}{\beta} \sqrt{\frac{\mu_0}{\varepsilon_0}} \frac{B}{A} \right) \right] + \\ + iJ_1(i\rho a) \left[a \left(\frac{\beta^2}{\kappa} - \kappa \right) J_0(\kappa a) + \right. \\ \left. + a \left(\frac{\beta^2}{\gamma} + \gamma \right) \frac{iH_0^{(1)}(i\gamma a)}{H_1^{(1)}(i\gamma a)} J_1(\kappa a) - \right. \\ \left. - \frac{\beta^2}{\kappa^2\gamma^2} (\kappa^2 + \gamma^2) \left(1 - \frac{k_0}{\beta} \sqrt{\frac{\mu_0}{\varepsilon_0}} \frac{A}{B} \right) J_1(\kappa a) \right] \right\}. \end{aligned} \quad (10.4.20)$$

Индекс h при $\Delta\beta$ указывает на то, что векторы электрических полей мод обоих волноводов, показанных на фиг. 10.4.1, поляризованы горизонтально. Это означает, что вблизи оси каждого волновода вектор электрического поля параллелен линии, соединяющей центры волноводов. В случае вертикальной поляризации векторы электрического поля ориентированы перпендикулярно (или вертикально) к линии, соединяющей центры волноводов. Связь вертикально поляризованных мод отличается от связи горизонтально поляризованных мод несущественно. Подобное вычисление приводит к результату

$$\begin{aligned} \Delta\beta_v = & \frac{A^2}{2P} \frac{N}{A} \frac{1}{\rho} \sqrt{\frac{2\pi}{\rho R}} e^{-\rho R} \left\{ \frac{e^{(\rho-\gamma)b}}{\pi\gamma b \sqrt{\rho\gamma}} \frac{J_1(\kappa a)}{H_1^{(1)}(i\gamma a)} \times \right. \\ & \times \left[k_0 \sqrt{\frac{\mu_0}{\epsilon_0}} \frac{B}{A} \left((\gamma + \rho)b + \frac{\rho^2 - \gamma^2}{\gamma\rho} \right) - \frac{\beta}{\rho\gamma} (\rho^2 - \gamma^2) \right] + \\ & + i\rho J_1(i\rho a) \left[k_0 a \sqrt{\frac{\mu_0}{\epsilon_0}} \frac{B}{A} \left(\frac{1}{k_0} J_0(\kappa a) + \right. \right. \\ & \left. \left. + \frac{1}{\gamma} \frac{iH_0^{(1)}(i\gamma a)}{H_1^{(1)}(i\gamma a)} J_1(\kappa a) \right) + \right. \\ & \left. \left. + \frac{\beta}{\kappa^2\gamma^2} (\kappa^2 + \gamma^2) \left(1 - \frac{k_0}{\beta} \sqrt{\frac{\mu_0}{\epsilon_0}} \frac{B}{A} \right) J_1(\kappa a) \right] \right\}. \quad (10.4.21) \end{aligned}$$

В этом выражении показатели преломления n_c и n_m должны быть действительными, а n_z может быть комплексным. Коэффициенты A , B , M и N были определены ранее в этом разделе. Параметры k_0 , κ , γ и ρ задаются с помощью формул (10.3.2) — (10.3.5).

Вертикально поляризованная мода не связана с горизонтально поляризованной модой.

Несмотря на очевидную грубость приближений, выражения (10.4.20) и (10.4.21) очень хорошо описывают связь между двумя модами HE_{11} круглых оптических волокон в оболочке. Результаты по теории связи мод HE_{11} были опубликованы ранее Джонсом [84], который провел сравнение своих теоретических результатов с измерениями Брэйси и др. [110]. Теоретические результаты Джонса хорошо согласуются с экспериментальными данными. Однако его теория не включает случая волокон в оболочке с потерями в окружающей среде. Автором были проведены

численные расчеты для случая $n_m = n_3$, $\rho = \gamma$. Сравнение полученных результатов с результатами Джонса показало, что они хорошо согласуются, даже когда две сердцевинки касаются друг друга. Для особого случая волокон без оболочки теория Джонса полностью согласуется с представленной здесь теорией и обе теории хорошо согласуются с экспериментальными измерениями Брэиси [110].

Как и для плоских волноводов, желательно знать, какие потери моды HE_{11} вызываются потерями окружающей среды, имеющей большой показатель преломления, при котором выполняется условие (10.3.15), и какие потери связаны с поглощением в окружающей среде. Используя подход, описанный в разд. 10.3, без особых трудностей получим

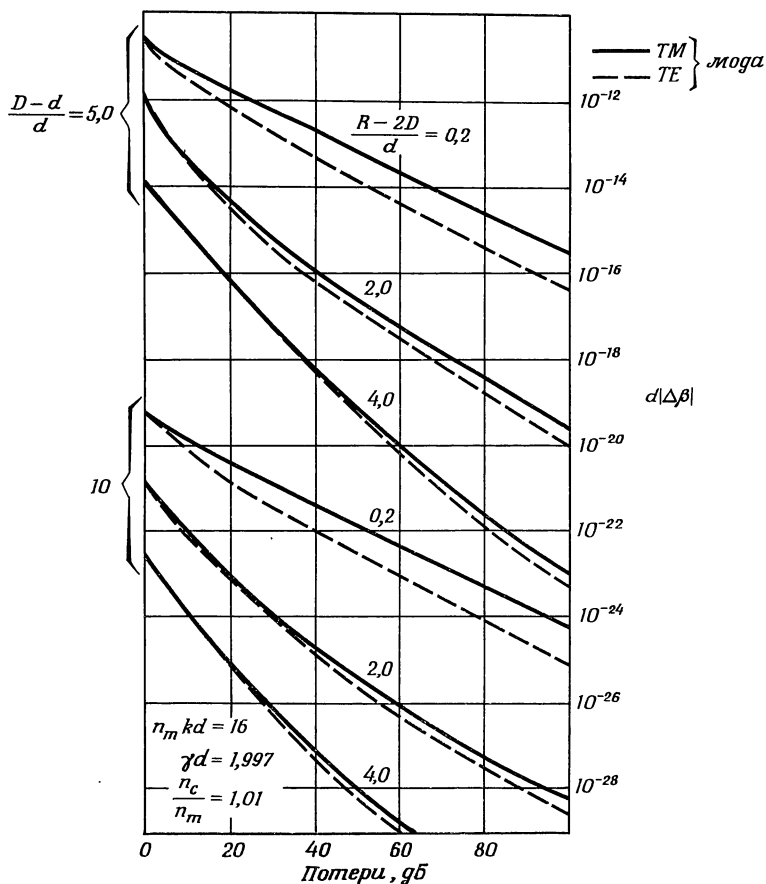
$$2\alpha = \sqrt{\frac{\varepsilon_0}{\mu_0}} \frac{e^{-2\rho_r b}}{|\rho| P} \left\{ \frac{\beta}{b} \frac{1}{|\rho|^4} \times \right. \\ \times \operatorname{Im} \left[\sqrt{\frac{\mu_0}{\varepsilon_0}} MN^* (\rho^{*2} - \rho^2) \right] - k_0 |M|^2 \operatorname{Im} \left(\frac{n_3^2}{\rho} \right) - \\ \left. - \frac{\mu_0}{\varepsilon_0} k_0 |N|^2 \operatorname{Im} \left(\frac{1}{\rho} \right) \right\}. \quad (10.4.22)$$

Здесь символ $\operatorname{Im}(\)$ обозначает мнимую часть; ρ_r — действительная часть ρ . Коэффициенты M^2 , N^2 и MN^* выражаются через коэффициент, который содержит в качестве множителя P , поэтому мощность P из выражения исчезает.

10.5. ПЕРЕКРЕСТНАЯ СВЯЗЬ

Результаты двух предыдущих разделов дают возможность рассмотреть перекрестную связь плоских и круглых волноводов в оболочке [103]. На фиг. 10.5.1 показана зависимость произведения $|\Delta\beta|d$ от потерь в окружающей среде для плоского волновода. Под потерями в окружающей среде понимаются потери, которые испытывает плоская волна при распространении перпендикулярно оси волновода от границы оболочки первого волновода к границе оболочки второго волновода. Сплошные кривые на фиг. 10.5.1 относятся к ТМ-модам, а пунктирные — к ТЕ-модам. Поляризации мод ТМ и ТЕ соответствуют

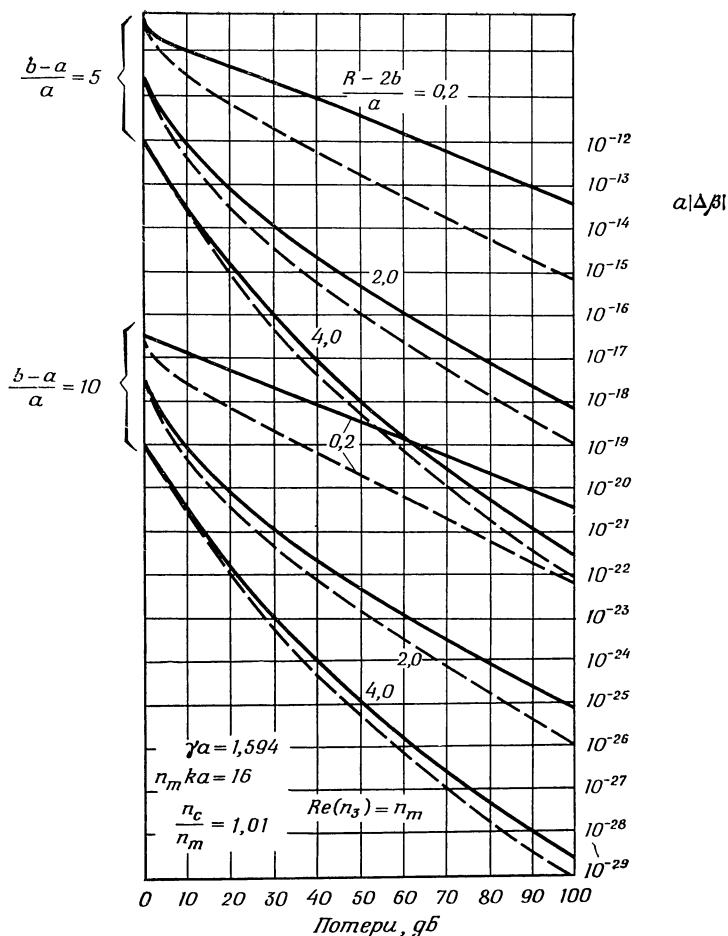
горизонтальной и вертикальной поляризациям в круглом волокне. В качестве параметров взяты относительная



Ф и г. 10.5.1. Зависимость параметра перекрестной связи $|\Delta\beta|d$ двух связанных плоских волноводов в оболочке от величины потерь плоской волны в материале вне оболочек волноводов.

Параметр $(D-d)/d$ — относительная толщина оболочки, параметр $(R-2D)/d$ — относительное расстояние между оболочками двух волноводов.

толщина оболочки и относительное расстояние между оболочками волноводов. Относительная толщина оболоч-



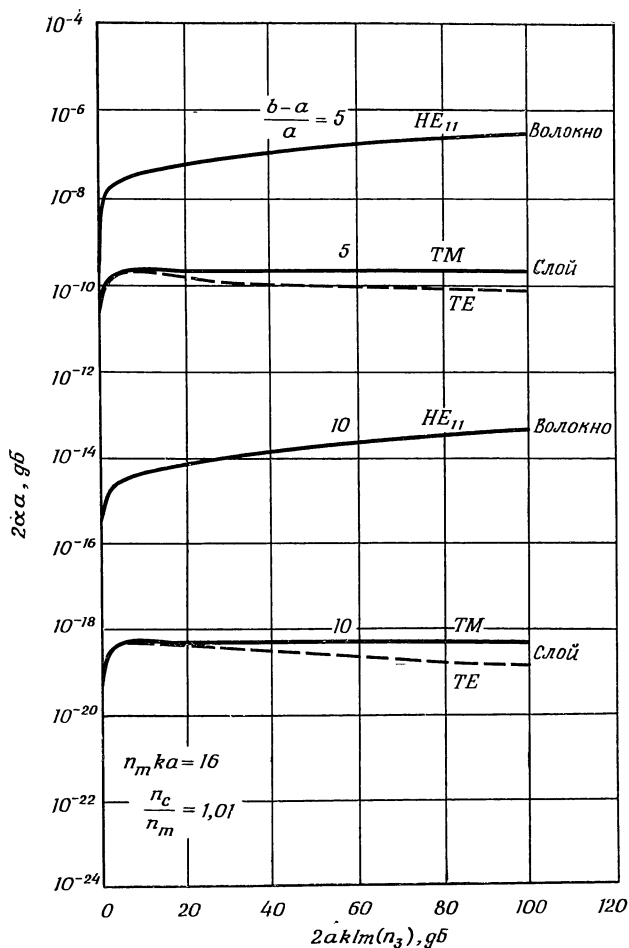
Ф и г. 10.5.2. Зависимость параметра связи $|\Delta\beta| a$ двух оптических волокон в оболочке от потерь плоской волны в материале вне оболочек двух волноводов [103].

Параметр $(b-a)/a$ — относительная толщина оболочки. Параметр $(R-2b)/a$ — относительное расстояние между оболочками двух волноводов. — — — вертикальная поляризация; ————— горизонтальная поляризация.

ки $(D - d)/d = 5$ и 10. Видно, насколько эффективно увеличение толщины оболочки уменьшает связь между двумя волноводами. Вторым параметром — относительное расстояние между двумя оболочками волноводов $(R - 2D)/d$ — имеют значения 0,2, 2,0 и 4. Уменьшение связи с увеличением значений этого параметра опять связано с увеличением расстояния между сердцевинами волноводов. Потери, представленные на оси абсцисс, даны для различных расстояний. Эти потери всегда имеют место при приближении одного волновода к другому. Их можно регулировать, осуществляя необходимую развязку волноводов. На фиг. 10.5.1 связь уменьшается с увеличением потерь в окружающей среде. Действительная часть показателя преломления окружающей среды n_3 выбиралась равной показателю преломления оболочки n_m .

На фиг. 10.5.2 приведены аналогичные данные для круглого волокна в оболочке. Из кривых видно, что горизонтально поляризованные моды связаны сильнее, чем вертикально поляризованные, а плоские волноводы связаны слабее круглых волокон. Различие в связи вызвано тем, что $\gamma_{\text{слоя}} > \gamma_{\text{волокна}}$. С увеличением величины γ поле концентрируется плотнее около сердцевин и уменьшается та его часть, которая достигает другого волновода.

На фиг. 10.5.3 показана зависимость потерь мод от потерь $2ak_0 \operatorname{Im}(n_3)$ в окружающей среде. В случае плоского волновода радиус сердцевин a необходимо заменить на полутолщину слоя d . Аналогично толщина оболочки волокна $(b - a)/a$ в этом случае заменяется толщиной оболочки $(D - d)/d$. Потери для плоского волновода меньше, так как $\gamma_{\text{слоя}} > \gamma_{\text{волокна}}$. Наиболее характерной особенностью кривых фиг. 10.5.3 является слабая зависимость потерь мод от потерь в окружающей среде после первоначального быстрого нарастания от нулевого значения. Поскольку окружающая среда имеет потери в несколько децибелл на радиус сердцевин, потери мод определяются преимущественно толщиной оболочки. Такой характер зависимостей дает возможность рассмотреть одновременно проблемы потерь мод и перекрестной связи. Рассмотрим пример. Пусть волновод имеет $a = 1$ мкм ($d = 1$ мкм) и потери мод, вызванные потерями в окружающей среде, составляют 1 дБ/км. Это озна-



Ф и г. 10.5.3. Потери мод, вызванные потерями в среде вне оболочки.

По оси абсцисс отложены потери плоской волны во внешней среде на расстоянии, равном радиусу сердцевины (или полуширине сердцевины слоя). Параметр $(b - a)/a$ представляет собой относительную ширину оболочки. Сплошные кривые соответствуют ТМ-модам, пунктирные — ТЕ-модам.

чает, что необходимо иметь $2a\alpha < 10^{-9}$ дБ. Из кривых на фиг. 10.5.3 только самая верхняя кривая для моды HE_{11} показывает превышение допустимых потерь моды. Далее предположим, что перекрестная связь для волновода длиной 1 км составляет $C = 10^{-6}$. Согласно выражению (10.2.38), для этого необходимо, чтобы $|\Delta\beta| a < 10^{-12}$. Все кривые на фиг. 10.5.1, кроме самой верхней, удовлетворяют этому требованию. Ни одна из кривых верхнего набора на фиг. 10.5.2 для $(b - a)/a = 5$ не может быть использована до тех пор, пока не будут обеспечены высокие потери в окружающей среде. Такое сравнение потерь при перекрестной связи позволяет сделать следующий вывод. Если развязывать два параллельных волновода путем увеличения потерь в окружающей среде, то одновременно будут увеличиваться потери мод до недопустимо высокого уровня. Или, другими словами, если толщина оболочки недостаточна для получения нужной изоляции мод, то она также недостаточна, чтобы предохранить моды волновода от потерь в окружающей среде. Это заключение может быть доказано более основательно, чем только на примере представленных здесь двух простых случаев [103]. Таким образом, увеличивая потери в окружающей среде, нельзя уменьшить перекрестную связь двух диэлектрических волноводов. Если толщина оболочки достаточна для развязки двух волноводов, то потери в окружающей среде не нужны. Потери же мод, вызванные потерями в окружающей среде, становятся слишком большими.

Здесь мы не рассматриваем перекрестную связь за счет рассеяния. Интересующие читателя результаты можно найти в [104]. Заключение относительно желаемых потерь в окружающей среде справедливо и для перекрестной связи за счет рассеяния. Если рассеяние на дефектах волновода является источником перекрестной связи, то существует тесная связь между величиной возникающей перекрестной связи и соответствующими потерями мод. Можно опять сделать вывод, что если уровень перекрестной связи достаточно высок, то потери за счет рассеяния направляемых мод также высоки и недопустимы. В принципе, конечно, можно подавить поперечную связь за счет рассеяния увеличением потерь в окружающей среде.

Эти потери не сказываются на характере распространения мод, если оболочка достаточно толстая. Однако если потери в окружающей среде необходимы для надежной изоляции волновода, то волновод оказывается настолько плохим, что добавочные потери за счет рассеяния делают его совершенно непригодным. Однако есть одно исключение. Если потери за счет рассеяния вызываются не случайными дефектами волновода, а систематической синусоидальной деформацией на протяжении всей длины обоих волноводов, то потери мод могут быть не слишком большими, а перекрестная связь при этом может быть значительной. В этом случае потери в окружающей среде будут желательны для развязки волноводов.

Результаты, изложенные в этой главе, можно использовать, разумеется, и для расчета характеристик направленных ответвителей. Выражения для коэффициентов связи (10.2.17) и (10.2.18) являются достаточно общими и могут применяться к любому типу диэлектрического волновода.

ЛИТЕРАТУРА

1. **Born M., Wolf E.** Principles of Optics, 3rd ed., New York, Pergamon Press, 1964; русский перевод: **Борн М., Вольф Э.** Основы оптики, изд-во «Наука», 1970.
2. **Stratton J. A.** Electromagnetic Theory, New York, McGraw-Hill, 1941; русский перевод: **Страттон Дж. А.** Теория электромагнетизма, М.—Л., ОГИЗ — Гостехиздат, 1948.
3. **Durney C. H., Johnson C. C.** Introduction to Modern Electromagnetics, New York, McGraw-Hill, 1969.
4. **Courant R., Hilbert D.** Methods of Mathematical Physics, Vol. II, Partial Differential Equations, New York, Interscience Publ., 1962; русский перевод: **Курант Р., Гильберт Д.** Методы математической физики, т. II, Уравнения с частными производными, изд-во «Мир», 1964.
5. **Dirac P. A. M.** The Principles of Quantum Mechanics, 4th ed., New York, Oxford, 1958; русский перевод: **Дирак П. А. М.** Принципы квантовой механики, Физматгиз, 1960.
6. **Courant R., Hilbert D.** Methods of Mathematical Physics, Vol. I, New York, Interscience Publ., 1953; русский перевод: **Курант Р., Гильберт Д.** Методы математической физики, т. I, М.—Л., ГИТТЛ, 1951.
7. **Goodman J. W.** Introduction to Fourier Optics, New York, McGraw-Hill, 1968; русский перевод: **Гудмен Дж.** Введение в Фурье-оптику, изд-во «Мир», 1970.
8. **Abromovitz M., Stegun I. A.** Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables, National Bureau of Standards Applied Mathematics Series, v. 55, Washington, D.C., National Bureau of Standards, 1965.
9. **Mathews J., Walker R. L.** Mathematical Methods of Physics, New York, Benjamin, 1965.
10. **Marcuse D.** Engineering Quantum Electrodynamics, New York, Harcourt, Brace and World, 1970.
11. **Jahnke E., Emde P.** Tables of Functions with Formulas and Curves, 4th ed., New York, Dover Publ., 1945; русский перевод: **Янке Е., Эмде Ф., Лёш Ф.** Специальные функции. Формулы, графики, таблицы, изд-во «Наука», 1968.
12. **Heitler W.** The Quantum Theory of Radiation, 3rd ed., New York, Oxford, 1957.

13. **Kogelnik H.** Coupled Wave Theory for Thick Hologram Gratings — «BSTJ», 1969, v. 48, № 9, p. 2909—2947.
14. **Keller J. B., Streifer W.** Complex Rays with an Application to Gaussian Beams — «J. Opt. Soc. Am.», 1971, v. 61, № 1, p. 40—43.
15. **Herzberger M.** Modern Geometrical Optics, New York, Interscience Publ., 1958; русский перевод: **Герцбергер М.** Современная геометрическая оптика, ИЛ, 1962.
16. **Luneberg R. K.** Mathematical Theory of Optics, Berkeley and Los Angeles, Univ. of California Press, 1964.
17. **Kaplan W.** Advanced Calculus, Reading, Mass., Addison-Wesley Publ. Co., 1953.
18. **Kline M., Kay I. W.** Electromagnetic Theory and Geometrical Optics, New York, Interscience Publ., 1965.
19. **Goldstein H.** Classical Mechanics, Reading, Mass., Addison-Wesley Publ. Co, 1950.
20. **Гельфанд И. М., Фомин С. В.** Вариационное исчисление, Физматгиз, 1961.
21. **Фок В. А.** Теория пространства, времени и тяготения, Физматгиз, 1961.
22. **Schiff L. I.** Quantum Mechanics, 2nd ed., New York, McGraw-Hill, 1955; русский перевод: **Шифф Л.** Квантовая механика, ИЛ, 1959.
23. **Kitano I., Koizumi K., Matsumura H., Uchida T., Furukawa M.** A Light-Focusing Fibre Guide Prepared by Ion Exchange Techniques — «J. Japan Soc. Appl. Phys. (Suppl.)», 1970, v. 39, p. 63—70.
24. **Marcuse D.** Physical Limitations on Ray Oscillation Suppressors — «BSTJ», 1966, v. 45, № 5, p. 743—751.
25. **Tolman R. C.** The Principles of Statistical Mechanics, New York, Oxford Univ. Press, 1938.
26. **Courant R.** Differential and Integral Calculus, v. 2, New York, Wiley, 1936.
27. **Berreman D. W.** A Lens or Light Guide Using Convectively Distorted Thermal Gradients in Gases — «BSTJ», 1964, v. 43, № 4, p. 1469—1475.
28. **Kaiser P.** Measured Beam Deformation in a Guide Made of Tubular Gas Lenses — «BSTJ», 1968, v. 47, № 2, p. 179—194.
29. **Kaiser P.** An Optical Beam Waveguide Made of Counterflow Gas Lenses (будет опубликовано).
30. **Beck A. C.** An Experimental Gas Lens Optical Transmission Line — «IEEE Trans. Microwave Theory and Techniques», 1967, v. MTT-15, № 7, p. 433—434.
31. **Marcuse D., Miller S. E.** Analysis of a Tubular Gas Lens — «BSTJ», 1964, v. 43, № 4, p. 1759—1782.
32. **Jacob M.** Heat Transfer, Vol. 1, New York, Wiley, 1949.

33. **Gloge D.** Deformation of Gas Lenses by Gravity — «BSTJ», 1967, v. 46, № 2, p. 357—365.
34. **Marcuse D.** Theory of a Thermal Gradient Gas Lens — «IEEE Trans. Microwave Theory and Techniques», 1965, v. MTT-13, № 6, p. 734—739.
35. **Steier W. H.** Measurements on a Thermal Gradient Gas Lens — «IEEE Trans. Microwave Theory and Techniques», 1965, v. MTT-13, № 6, p. 740—748.
36. **Marcuse D.** Comparison between a Gas Lens and Its Equivalent Thin Lens — «BSTJ», 1966, v. 45, № 8, p. 1339—1344.
37. **Toraldi di Francia G.** Degrees of Freedom of an Image — «J. Opt. Soc. Am.», 1969, v. 59, № 7, p. 799—804.
38. **Spiegel M. R.** Theory and Problems of Complex Variables with an Introduction to Conformal Mapping and Its Applications, New York, Schaum Publ. Co., 1964.
39. **Shannon C. E.** A Mathematical Theory of Communications — «BSTJ», 1948, v. 27, p. 379—423, 623—656.
40. **Slepian D.** Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty-IV: Extension to Many Dimensions, Generalized Prolate Spheroidal Functions — «BSTJ», 1964, v. 43, № 6, p. 3009—3057.
41. **Slepian D., Sonnenblick E.** Eigenvalues Associated with Prolate Spheroidal Wave Functions of Zero Order — «BSTJ», 1965, v. 44, № 8, p. 1745—1759.
42. **Gloge D., Marcuse D.** Formal Quantum Theory of Light Rays — «J. Opt. Soc. Am.», 1969, v. 59, № 12, p. 1629—1631.
43. **Goubau G., Schwering F.** On the Guided Propagation of Electromagnetic Beam Waves — «IRE Trans. on Antennas and Propagation», 1961, v. AP-9, p. 248—256; русский перевод: **Губо Г., Шверинг Ф.** Направленное распространение электромагнитных волновых пучков — «Зарубежная радиоэлектроника», 1961, № 11, стр. 3—15.
44. **Pierce J. R.** Theory and Design of Electron Beams, New York, Van Nostrand Reinhold, 1954; русский перевод: **Пирс Дж. Р.** Теория и расчет электронных пучков, изд-во «Советское радио», 1960.
45. **Miller S. E.** Alternating Gradient Focusing and Related Properties of Conventional Convergent Lens Focusing — «BSTJ», 1964, v. 43, № 4, pt. II, p. 1741—1758.
46. **Boyd G. D., Kogelnik H.** Generalized Confocal Resonator Theory — «BSTJ», 1962, v. 41, № 4, p. 1347—1369.
47. **Boyd G. D., Gordon J. P.** Confocal Multimode Resonator for Millimeter through Optical Wavelength Masers — «BSTJ», 1961, v. 40, № 2, p. 489—508; русский перевод: **Бойд Дж., Гордон Дж.** Конфокальный резонатор со многими типами колебаний для квантовых генераторов миллиметрового и оптического диапазонов, сб. «Лазеры», ИЛ, 1963.

48. Fox A. G., Li T. Resonant Modes in a Maser Interferometer — «BSTJ», 1961, v. 40, № 2, p. 453—488; русский перевод: Фокс А., Ли Т. Резонансные типы колебаний в интерферометре квантового генератора, сб. «Лазеры», ИЛ, 1963.
49. Kogelnik H., Li T. Laser Beams and Resonators — «Applied Optics», 1966, v. 5, № 10, p. 1550—1567.
50. Hirano J., Fukatsu Y. Stability of Light Beams in a Beam Waveguide — «Proc. IEEE», 1964, v. 52, № 11, p. 1284—1292.
51. Marcuse D. Propagation of Light Rays through a Lens Waveguide with Curved Axis — «BSTJ», 1964, v. 43, № 2, p. 741—753.
52. Marcuse D. Statistical Treatment of Light-Ray Propagation in Beam-Waveguides — «BSTJ», 1965, v. 44, № 9, p. 2065—2081.
53. Marcuse D. Probability of Ray Position in Beam Waveguides — «IEEE Trans. on Microwave Theory and Techniques», 1967, v. MTT-15, № 3, p. 167—171.
54. Papoulis A. Probability, Random Variables and Stochastic Processes, New York, McGraw-Hill, 1965.
55. Beckmann P. Probability in Communication Engineering, New York, Harcourt, Brace and World, 1967.
56. Marcatili E. A. J. Ray Propagation in Beam-Waveguides with Redirectors — «BSTJ», 1966, v. 45, № 1, p. 105—115.
57. Marcatili E. A. J. Effect of Redirectors, Refocusers, and Mode Filters on Light Transmission through Aberrated and Misaligned Lenses — «BSTJ», 1967, v. 46, № 8, p. 1733—1752.
58. Johnson C. C. Field and Wave Electrodynamics, New York, McGraw-Hill, 1965.
59. Heine V. Group Theory in Quantum Mechanics, New York, Pergamon Press, 1960.
60. Ziman J. M. Electrons and Phonons, New York, Oxford Univ. Press, 1960.
61. Градштейн И. С., Рыжик И. М. Таблицы интегралов, сумм, рядов и произведений, изд-во «Наука», 1971.
62. Slepian D., Pollak H. O. Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty (I) — «BSTJ», 1961, v. 40, № 1, p. 43—63.
63. Landau H. J., Pollak H. O. Prolate Spheroidal Wave Functions, Fourier Analysis and Uncertainty (II) — «BSTJ», 1961, v. 40, № 1, p. 65—84.
64. Marcuse D. Deformation of Fields Propagating through Gas Lenses — «BSTJ», 1966, v. 45, № 8, p. 1345—1368.
65. Marcatili E. A. J. Off-Axis Wave-Optics Transmission in a Lens-Like Medium with Aberrations — «BSTJ», 1967, v. 46, № 1, p. 149—166.
66. Eichmann G. Quasi-Geometric Optics of Media with Inhomogeneous Index of Refraction — «J Opt. Soc. Am.», 1971, v. 61, № 2, p. 161—168.

67. **Kogelnik H.** Imaging of Optical Mode Resonators with Internal Lenses — «BSTJ», 1965, v. 44, № 3, p. 455—494.
68. **Kogelnik H.** On the Propagation of Gaussian Beams of Light through Lenslike Media Including Those with Loss or Gain Variation — «Appl. Optics», 1965, v. 4, № 12, p. 1562—1569.
69. **Tien P. K., Gordon J. P., Whinnery J. R.** Focusing of a Light Beam of Gaussian Field Distribution in Continuous and Periodic Lens-Like Media — «Proc. IEEE», 1965, v. 53, № 2, p. 129—136.
70. **Fork R. L., Herriot D. R., Kogelnik H.** A Scanning Spherical Mirror Interferometer for Spectral Analysis of Laser Radiation — «Appl. Optics», 1964, v. 3, № 12, p. 1471—1484.
71. **Smith P.** Stabilized Single Frequency Output from a Long Laser Cavity — «IEEE J. of Quantum Electronics», 1965, v. QE-1, № 8, p. 343—348.
72. **Boyd G. D., Kogelnik H.** Generalized Confocal Resonator Theory — «BSTJ», 1962, v. 41, № 4, p. 1347—1369.
73. **Erdelyi A., Magnus W., Oberhettinger F., Tricomi F. G.** Higher Transcendental Functions, Vol. II, New York, McGraw-Hill, 1953.
74. **Marcuse D.** Modes and Pseudomodes in Dielectric Waveguides — «IEEE Trans. on Microwave Theory and Techniques», 1970, v. MTT-18, № 1, p. 62, 63.
75. **Hirsch P., Carter A. H.** Mathematical Models for the Prediction of SOFAR Propagation Effects — «J. Acoustical Soc. Am.», 1965, v. 37, № 1, p. 90—94.
76. **Miller S. E.** Light Propagation in Generalized Lenslike Media — «BSTJ», 1965, v. 44, № 9, p. 2017—2064.
77. **Gordon J. P.** Optics of General Guiding Media — «BSTJ», 1966, v. 45, № 2, p. 321—332.
78. **Marcatili E. A. J.** Modes in a Sequence of Thick Astigmatic Lens-Like Focusers — «BSTJ», 1964, v. 43, № 6, p. 2887—2904.
79. **Kapany N. S.** Fiber Optics, New York, Academic Press, 1967; русский перевод: **Капани Н. С.** Волоконная оптика, изд-во «Мир», 1969.
80. **Snitzer E.** Cylindrical Dielectric Waveguide Modes — «J. Opt. Soc. Am.», 1961, v. 51, № 5, p. 491—498.
81. **Collin R. E.** Field Theory of Guided Waves, New York, McGraw-Hill, 1960.
82. **Rowe H. E., Warters W. D.** Transmission in Multimode Waveguide with Random Imperfections — «BSTJ», 1962, v. 41, № 3, p. 1031—1170.
83. **Schlesinger S. P., Diamant P., Vigants A.** On Higher Order Hybrid Modes of Dielectric — «IEEE Trans. on Microwave Theory and Techniques», 1960, v. MTT-8, № 2, p. 252, 253.
84. **Jones A. L.** Coupling of Optical Fibers and Scattering in Fibers — «J. Opt. Soc. Am.», 1965, v. 55, № 3, p. 261—274.

85. **Vanclooster R., Phariseau P.** The Coupling of Two Parallel Dielectric Fibers I-Basic Equations — «Physica», 1970, v. 47, № 4, p. 485—500.
86. **Vanclooster R., Phariseau P.** The Coupling of Two Parallel Dielectric Fibers II-Characteristic of Coupling in Two Fibers — «Physica», 1970, v. 47, № 4, p. 501—514.
87. **Vanclooster R., Phariseau P.** Light Propagation in Fiber Bundles — «Physica», 1970, v. 49, № 4, p. 493—501.
88. **Miller S. E.** Coupled Wave Theory and Waveguide Applications — «BSTJ», 1954, v. 33, № 3, p. 661—720.
89. **Snyder A. W.** Radiation Loss Due to Variations of Radius on Dielectric or Optical Fibers — «IEEE Trans. on Microwave Theory and Techniques», 1970, v. MTT-18, № 9, p. 608—615.
90. **Snyder A. W., De La Rue R.** Asymptotic Solutions of Eigenvalue Equations for Surface Waveguide Structures — «IEEE Trans. on Microwave Theory and Techniques», 1970, v. MTT-18, № 9, p. 650, 651.
91. **Snyder A. W.** Coupling of Modes on a Tapered Dielectric Cylinder — «IEEE Trans. on Microwave Theory and Techniques», 1970, v. MTT-18, № 7, p. 383 —392.
92. **Snyder A. W.** Asymptotic Expressions for Eigenfunctions and Eigenvalues of a Dielectric or Optical Waveguide — «IEEE Trans. on Microwave Theory and Techniques», 1969, v. MTT-17, № 12, p. 1130—1138.
93. **Snyder A. W.** Excitation and Scattering of Modes on a Dielectric or Optical Fiber — «IEEE Trans. on Microwave Theory and Techniques», 1969, v. MTT-17, p. 1138—1144.
94. **Dyott R. B., Stern J. R.** Group Delay in Glass Fibre Waveguide, Conference on Trunk Telecommunications by Guided Waves, Conference Publication № 71, London, Institute of Electrical Engineers, p. 176—181, 1970.
95. **Kapron F. P., Keck D. B., Maurer R. D.** Radiation Losses in Glass Optical Waveguides — «Appl. Phys. Letters», 1970, v. 17, № 10, p. 423—425.
96. **Marcuse D.** Radiation Losses of the Dominant Mode in Round Dielectric Waveguides — «BSTJ», 1970, v. 49, № 8, p. 1665—1693.
97. **Adler R. B.** Waves on Inhomogeneous Cylindrical Structures — «Proc. IRE», 1952, v. 40, № 3, p. 339—348.
98. **Marcuse D.** Mode Conversion Caused by Surface Imperfections of a Dielectric Slab Waveguide — «BSTJ», 1969, v. 48, № 10, p. 3187—3215.
99. **Marcuse D., Derosier R. M.** Mode Conversion Caused by Diameter Changes of a Round Dielectric Waveguide — «BSTJ», 1969, v. 48, № 10, p. 3217—3232.
100. **Marcuse D.** Radiation Losses of Dielectric Waveguides in Terms of the Power Spectrum of the Wall Distortion Function — «BSTJ», 1969, v. 48, № 10, p. 3233—3242,

101. **Marcuse D.** Radiation Losses of Tapered Dielectric Slab Waveguides — «BSTJ», 1970, v. 49, № 2, p. 273—290; Erratum — «BSTJ», 1970, v. 49, № 5, p. 949.
102. **Marcuse D.** Compression of a Bundle of Light Rays — «Appl. Optics», 1971, v. 10, № 3, p. 494—497.
103. **Marcuse D.** The Coupling of Degenerate Modes in Two Parallel Dielectric Waveguides — «BSTJ», 1971, v. 50, № 6, p. 1791—1816.
104. **Marcuse D.** Crosstalk Caused by Scattering in Slab Waveguides — «BSTJ», 1971, v. 50, № 6, p. 1817—1831.
105. **Tynes A. R., Pearson A. D., Bisbee D. L.** Loss Mechanisms and Measurements in Clad Glass Fibers and Bulk Glass — «J. Opt. Soc. Am.», 1971, v. 61, № 2, p. 143—153.
106. **Marcatili E. A. J.** Bends in Optical Dielectric Guides — «BSTJ», 1969, v. 48, № 7, p. 2103—2132.
107. **Harris J. H., Shubert R., Polky J. N.** Beam Coupling to Films — «J. Opt. Soc. Am.», 1970, v. 60, № 8, p. 1007—1016.
108. **Tien P. K., Ulrich R., Martin R. J.** Modes of Propagating Light Waves in Thin Deposited Semiconductor Films — «Appl. Phys. Letters», 1969, v. 14, № 9, p. 291—294.
109. **Ulrich R.** Optimum Excitation of Optical Surface Waves — «J. Opt. Soc. Am.», 1971, v. 61, № 11, p. 1467—1477.
110. **Bracey M. F., Cullen A. L., Gillespie E. F. T., Staniforth J. A.** Surface-Wave Research in Sheffield — «IRE Trans. on Antennas and Propagation», 1959, v. AP-7, Suppl., p. S249.
111. **Schweber S. S.** An Introduction to Relativistic Quantum Field Theory, New York, Harper and Row, 1961.
112. **Pearson A. D., French W. G., Rawson E. G.** Preparation of a Light Focusing Glass Rod by Ion Exchange Techniques — «Appl. Phys. Letters», 1969, v. 15, № 2, p. 76, 77.
113. **Шевченко В. В.** Плавные переходы в открытых волноводах, изд-во «Наука», 1969.
114. **Миллер М. А., Таланов В. И.** Поверхностные электромагнитные волны, направляемые границей с малой кривизной — «ЖТФ», 1956, т. 26, № 12, стр. 2755.
115. **Bergstein L., Marom E.** Angular Spectra of Optics Cavities — «J. Opt. Soc. Am.», 1966, v. 56, № 1, p. 16—32.
116. **Sanderson R. L., Streifer W.** Comparison of Laser Mode Calculations — «Appl. Optics», 1969, v. 8, № 1, p. 131—136.
117. **Siegman A. E.** Unstable Optical Resonator for Laser Applications — «Proc. IEEE», 1965, v. 53, № 3, p. 277—287.
118. **Kawakami S., Nishizawa J.** An Optical Waveguide with the Optimum Distribution of the Refractive Index with Reference to Waveform Distortion — «IEEE Trans. Microwave Theory and Techniques», 1968, v. 16, № 10, p. 814—818.
119. **Rawson E. G., Herriott D. R., McKenna J.** Analysis of Refractive Index Distributions in Cylindrical, Graded Index Glass Rods

- (Grin Rods) Used as Image Relays — «Appl. Optics», 1970, v. 9, № 3, p. 753—759.
120. Heurtley J. C. Hyperspheroidal Functions — Optical Resonators with Circular Mirrors — «Proceedings of the Symposium on Quasi-Optics», Polytechnic Press of the Polytechnic Institute of Brooklyn, p. 367—375, 1964.
121. Harris J. H., Shubert R. Variable Tunneling Excitation of Optical Surface Waves — «IEEE Trans. Microwave Theory and Techniques», 1971, v. 19, № 3, p. 269—276.

ДОПОЛНИТЕЛЬНАЯ ЛИТЕРАТУРА

Общие вопросы теории оптических и квазиоптических волноводов

- 1*. Вайнштейн Л. А. Открытые резонаторы и открытые волноводы, изд-во «Советское радио», 1966.
- 2*. Каценеленбаум Б. З. Высокочастотная электродинамика, изд-во «Наука», 1966.
- 3*. Каценеленбаум Б. З. Квазиоптические методы формирования и передачи миллиметровых волн. — «УФН», 1964, т. 83, № 1, стр. 81.
- 4*. Таланов В. И. Работы Л. И. Мандельштама по теории оптического изображения и современная квазиоптика. — «УФН», 1965, т. 87, № 9, стр. 23.
- 5*. Ардашев А. Ю., Кашин В. А., Скроцкий Г. В. Электромагнитное поле когерентного светового пучка. — «Радиотехника и электроника», 1968, т. 13, № 7, стр. 1320.
- 6*. Бондаренко Н. Г., Таланов В. И. Некоторые вопросы теории квазиоптических систем. — «Изв. вузов, Радиофизика», 1964, т. 7, № 2, стр. 313.
- 7*. Быков В. П. Построение волнового поля по лучевой картине. — «ЖТФ», 1970, т. 40, № 10, стр. 2035.
- 8*. Ваганов Р. Б., Матвеев Р. Ф. Теоретическое исследование протяженных слабо нерегулярных лучеводов. — «Радиотехника и электроника», 1968, т. 13, № 2, стр. 232.
- 9*. Ваганов Р. Б. Траектория энергетического центра волнового пучка. — «Радиотехника и электроника», 1970, т. 15, № 8, стр. 2016.
- 10*. Ваганов Р. Б., Мериакри В. В. Несимметричный волновой пучок. — «Радиотехника и электроника», 1970, т. 15, № 10, стр. 2174.
- 11*. Вардья В. П., Догадкин А. Б., Дяченко А. А., Коршунов И. П., Матвеев Р. Ф., Шушпанов О. Е. Полигонная квазиоптическая линия для исследования распространения лазерных пучков на большие расстояния. — «Радиотехника и электроника», 1973, т. 18, № 2, стр. 391.

- 12*. Vershinina L. N., Kazantsev Y. N., Meriakri V. V., Shevchenko V. V. Submillimetre Quasi-Optical Transmission Lines, European Microwave Conference, 1969, London, 251.
- 13*. Власов С. Н., Петрищев В. А., Таланов В. И. Усредненное описание волновых пучков в линейных и нелинейных средах (метод моментов).— «Изв. вузов, Радиофизика», 1971, т. 14, № 9, стр. 1353.
- 14*. Власов С. Н., Петрищев В. А., Таланов В. И. Теория периодической самофокусировки световых пучков.— «Изв. вузов, Радиофизика», 1970, т. 13, № 6, стр. 908.
- 15*. Власов С. Н., Петрищев В. А., Таланов В. И. Нелинейные квазиоптические системы.— «Изв. вузов, Радиофизика», 1972, т. 15, № 12, стр. 1162.
- 16*. Власов С. Н., Таланов В. И. О связи лучевого и волнового описания электромагнитных пучков в квазиоптических системах.— «Изв. вузов, Радиофизика», 1965, т. 8, № 1, стр. 195.
- 17*. Зятицкий В. А. О средних потерях на преобразование в лучевом (бесконечные неидеальные корректоры).— «Радиотехника и электроника», 1967, т. 12, № 7, стр. 1192.
- 18*. Измestьев А. А. Однопараметрические волновые пучки в свободном пространстве.— «Изв. вузов, Радиофизика», 1970, т. 13, № 9, стр. 1380.
- 19*. Измestьев А. А. Производящие функции параксимальных гауссовых пучков.— «Радиотехника и электроника», 1971, т. 16, № 1, стр. 175.
- 20*. Каценеленбаум Б. З., Семенов В. В. Синтез фазовых корректоров, формирующих заданное поле.— «Радиотехника и электроника», 1967, т. 12, № 2, стр. 244.
- 21*. Коршунов И. П. Лазерный пучок, частично закрытый экраном.— «Радиотехника и электроника», 1970, т. 15, № 10, стр. 2176.
- 22*. Маненков А. Б. Возбуждение открытых однородных волноводов.— «Изв. вузов, Радиофизика», 1970, т. 13, № 5, стр. 739.
- 23*. Матвеев Р. Ф. К вопросу о многолучевой передаче по световодам.— «Радиотехника и электроника», 1971, т. 16, № 10, стр. 1950.
- 24*. Матвеев Р. Ф. О влиянии турбулентности воздуха на раскачку луча в световодных линиях.— «Радиотехника и электроника», 1972, т. 17, № 5, стр. 1073.
- 25*. Скроцкая Е. Г., Скроцкий Г. В. Электромагнитное поле узкого светового пучка, отраженного и преломленного на плоской границе раздела сред.— «Радиотехника и электроника», 1973, т. 18, № 3, стр. 483.
- 26*. Таланов В. И. Операторный метод описания волновых пучков в комбинированных квазиоптических системах.— «Изв. вузов, Радиофизика», 1965, т. 8, № 2, стр. 260.

- 27*. Хапалюк А. П., Нестеренко Г. М. О границах применимости параболического уравнения для описания световых пучков.— «Вестн. Белорусского ун-та», 1970, сер. 1, № 1, стр. 52.
- 28*. Хлопов Г. И. К вопросу о представлении поля квазиоптического пучка в виде углового спектра плоских волн.— «Радиотехника», 1972, вып. 20, стр. 168.
- 29*. Шатров А. Д. Дискретные представления поля в задаче возбуждения диэлектрической пластины.— «Радиотехника и электроника», 1970, т. 15, № 10, стр. 1806.
- 30*. Шевченко В. В. Наглядная классификация волн, направляемых регулярными открытыми волноводами.— «Радиотехника и электроника», 1969, т. 14, № 10, стр. 1768.
- 31*. Schevchenko V. V. Gradual Transitions in the Open Waveguides, European Microwave Conference, 1969, London, 272.
- 32*. Шевченко В. В. О разложении полей открытых волноводов по собственным и несобственным волнам.— «Изв. вузов, Радиофизика», 1971, т. 14, № 8, стр. 1242.

Линзовые и диафрагменные волноводы

- 33*. Баскаков С. И. Возбуждение лучевого волновода.— «Радиотехника и электроника», 1964, т. 9, № 4, стр. 607.
- 34*. Ваганов Р. Б. Потери энергии на преобразование в волны высших типов при деформациях и смещениях линз в конфокальном лучевом волноводе.— «Радиотехника и электроника», 1964, т. 9, № 11, стр. 1958.
- 35*. Ваганов Р. Б. Потери на неоднородном участке лучевода и возможность их компенсации.— «Радиотехника и электроника», 1965, т. 10, № 12, стр. 2146.
- 36*. Ваганов Р. Б. Диэлектрическая призма в лучевом волноводе.— «Радиотехника и электроника», 1969, т. 14, № 3, стр. 393.
- 37*. Ваганов Р. Б. Демпфирование раскочки пучка в диафрагмированном лучевом волноводе.— «Изв. вузов, Радиофизика», 1969, т. 12, № 10, стр. 1546.
- 38*. Ваганов Р. Б., Войтович Н. Н. Нерегулярности лучевода диафрагменного типа.— «Радиотехника и электроника», 1966, т. 11, № 2, стр. 339.
- 39*. Вардя В. П. Система автоматической подстройки световодной линии связи.— «Вопросы радиоэлектроники, сер. Техника проводной связи», 1969, № 8, стр. 83.
- 40*. Вершинина Л. Н., Лагунов А. А., Шевченко В. В. Исследование квазиоптических линий в субмиллиметровом диапазоне волн.— «Радиотехника и электроника», 1968, т. 13, № 2, стр. 346.
- 41*. Войтович Н. Н. Открытые резонаторы и линии с корректорами в виде конусов и усеченных конусов.— «Радиотехника и электроника», 1966, т. 11, № 3, стр. 488.

- 42*. **Войтович Н. Н.** К вопросу об исследовании лучевых волноводов со случайными сдвигами фаз.—«Радиотехника и электроника», 1966, т. 11, № 8, стр. 1496.
- 43*. **Войтович Н. Н.** Исследование устойчивости лучеводов к поперечным сдвигам линз с помощью моделирования на ЭВМ. — «Изв. вузов, Радиофизика», 1967, т. 10, № 9, стр. 1087.
- 44*. **Войтович Н. Н., Догадкин А. Б., Каценеленбаум Б. З.** Линзово-волноводная линия. — «Радиотехника и электроника», 1964, т. 9, № 9, стр. 1707.
- 45*. **Войтович Н. Н., Семенов В. В.** Квазиоптические линии-фидеры.—«Радиотехника и электроника», 1970, т. 15, № 4, стр. 697.
- 46*. **Dogadkin A. B.** Measurement of Quasi-Optical Transmission Line Parameters, Symp. of Microwave Measurements, 1966, Budapest.
- 47*. **Дяченко А. А., Шушпанов О. Е.** Определение голограмм для анализа волновых пучков в квазиоптических линиях передачи.— «Радиотехника и электроника», 1968, т. 13, № 11, стр. 2067.
- 48*. **Зяцкий В. А.** Регулярные лучеводы второго рода.— «Радиотехника и электроника», 1970, т. 15, № 8, стр. 1604.
- 49*. **Зяцкий В. А.** К теории обобщенных регулярных волноводов (регулярные лучеводы третьего рода).— «Радиотехника и электроника», 1970, т. 15, № 9, стр. 1816.
- 50*. **Коршунов И. П.** Экспериментальное исследование нерегулярных квазиоптических линзовых линий.— «Радиотехника и электроника», 1968, т. 13, № 2, стр. 348.
- 51*. **Коршунов И. П.** Полигонная квазиоптическая линзовая линия.—«Радиотехника и электроника», 1972, т. 17, № 12, стр. 2597.
- 52*. **Коршунова Е. Н., Матвеев Р. Ф.** Сравнение нерегулярных световодных линий из газовых и стеклянных линз.— «Вопросы радиоэлектроники, сер. Техника проводной связи», 1969, № 8, стр. 67.
- 53*. **Коршунова Е. Н., Сивов А. Н.** Границы применимости асимптотического описания тонкой линзы.— «Радиотехника и электроника», 1972, т. 17, № 4, стр. 863.
- 54*. **Макеев Ю. В., Смоловик В. С.** Определение амплитуд типов волн в линзовом волноводе.— «Электроника СВЧ», 1970, № 1, стр. 117.
- 55*. **Таланов В. И., Шеронова Н. М.** Влияние случайных фазовых аберраций корректоров на характеристики лучевого волновода при высокой степени селективности мод.— «Изв. вузов, Радиофизика», 1969, т. 12, № 2, стр. 273.
- 56*. **Шевченко В. В.** Неотражающие линзы для квазиоптических линий.— «Радиотехника и электроника», 1969, т. 14, № 10, стр. 1764.

- 57*. Шушпанов О. Е. Комбинированные корректоры в квазиоптических линиях.— «Радиотехника и электроника», 1968, т. 13, № 7, стр. 1232.
- 58*. Шушпанов О. Е. Квазиоптическая линия передачи из неидеальных корректоров.— «Радиотехника и электроника», 1968, т. 13, № 11, стр. 1950.

Полые и зеркальные волноводы

- 59*. Каценеленбаум Б. З. Теория нерегулярных волноводов с медленно меняющимися параметрами, Изд-во АН СССР, 1961.
- 60*. Ваганов Р. Б., Догадкин А. Б., Каценеленбаум Б. З. Перископическая зеркальная линия.— «Радиотехника и электроника», 1965, т. 10, № 9, стр. 1672.
- 61*. Горшкова Н. К., Дяченко А. А., Зятицкий В. А., Каценеленбаум Б. З., Колесникова Н. А. Основы статистического анализа распространения луча света в слабдеформированной зеркальной трубе круглого сечения.— «Радиотехника и электроника», 1966, т. 11, № 1, стр. 42.
- 62*. Дяченко А. А., Шушпанов О. Е. Исследование зеркальных перископических линий передачи.— «Изв. вузов, Радиофизика», 1968, т. 11, № 5, стр. 707.
- 63*. Казначеев Ю. И., Горшкова Н. К., Колесникова Н. А. Энергетические потери в плоском изгибе оптического волновода.— «Радиотехника и электроника», 1964, т. 9, № 6, стр. 1076.
- 64*. Каценеленбаум Б. З. Передача миллиметровых волн при помощи отражений от ряда фокусирующих зеркал.— «Радиотехника и электроника», 1963, т. 8, № 9, стр. 1516.
- 65*. Ривлин Л. А., Шильдяев В. С. Полигармонические волноводы для когерентного света.— «Изв. вузов, Радиофизика», 1968, т. 11, № 4, стр. 572.

Газово-линзовые волноводы и волноводы с фокусирующей средой

- 66*. Аэротермооптика и лучеводы, сб. докладов под ред. Лыкова А. В., Минск, изд-во «Наука и техника», 1970.
- 67*. Берковский Б. М., Мартыненко О. Г., Жилкин А. М., Порохов О. П. Теплогидродинамические световоды, Минск, изд-во «Наука и техника», 1969.
- 68*. Мартыненко О. Г., Колесников П. М., Колпачиков В. Л. Введение в теорию конвективных газовых линз, Минск, изд-во «Наука и техника», 1972.
- 69*. Герценштейн Ф. Э. Газовые линзы для оптических лучевых волноводов, сб. «Проблемы передачи информации лазерным излучением», Киев, 1969.

- 70*. Герценштейн Ф. Э. Распределение температуры в циркуляризованной газовой линзе с учетом гравитационных сил.— «Инж.-физич. журнал», 1969, т. 17, № 1, стр. 57.
- 71*. Герценштейн Ф. Э., Фрадин А. З. О влиянии гравитационных сил на фокусирующие свойства газовых линз.— «Труды учебных ин-тов связи», 1968, вып. 41.
- 72*. Мартыненко О. Г., Берковский Б. М., Ватутин И. А. К теории термодинамической газовой линзы.— «Инж.-физич. журнал», 1969, т. 17, № 1, стр. 145.
- 73*. Шевченко В. В. Квазиволноводные (вытекающие) волны в слоисто-неоднородных волноводах.— «Изв. вузов, Радиофизика», 1969, т. 12, № 9, стр. 1389.

Диэлектрические, включая пленочные, волноводы

- 74*. Взятых В. Ф. Диэлектрические волноводы, изд-во «Советское радио», 1970.
- 75*. Доклады научно-технической конференции, сб. под ред. Взятых В. Ф., Изд-во МЭИ, 1969.
- 76*. Совещание по диэлектрическим волноводам, сб. докладов под ред. Взятых В. Ф., Изд-во МЭИ, 1968.
- 77*. Аникин В. И., Дерюгин Л. Н., Сотин В. Е. Резонансное возбуждение плоского оптического волновода через критический слой ограниченным пучком.— «Изв. вузов, Радиоэлектроника», 1971, т. 14, № 4, стр. 371.
- 78*. Аникин В. И. Исследование пленочных функциональных микроволноводов среднего ИК диапазона.— «Изв. вузов, Радиоэлектроника», 1973, т. 16, № 8, стр. 5.
- 79*. Беланов А. С. О распределении продольного потока энергии в диэлектрических волноводах оптического диапазона, сб. «Радиоэлектроника оптического диапазона», 1967, стр. 17.
- 80*. Беланов А. С., Взятых В. Ф. Об условиях минимального взаимодействия между круглыми диэлектрическими волноводами, сб. «Вопросы технической физики», 1969, стр. 40.
- 81*. Беланов А. С., Ежов Г. И., Черный В. В. Уравнения и параметры собственных волн плоского несимметричного диэлектрического волновода, сб. «Радиоэлектроника оптического диапазона», 1970, стр. 3.
- 82*. Беланов А. С., Черный В. В. Составляющие поля и характеристическое уравнение собственных волн прямоугольного диэлектрического волновода, сб. «Взаимодействие излучения с веществом», 1972, стр. 239.
- 83*. Блох К. И., Беланов А. С. Волноводные свойства оптических волокон и волоконных элементов, сб. «Оптические волокна и волоконные элементы», изд-во «Химия», 1972, стр. 155.
- 84*. Будаган И. Ф., Дубровин В. Ф., Мировицкий Д. И. О методе парциальных волн в теории функциональных узлов интег-

- ральной оптики.— «Оптика и спектроскопия», 1972, т. 33, № 4, стр. 736.
- 85*. Будагян И. Ф., Дубровин В. Ф., Камлюк С. Н., Мировицкий Д. И., Усатюк В. В. Голографические рефрактометр и рефлектометр.— «Приборы и техника эксперимента», 1972, № 6, стр. 174.
- 86*. Взятых В. Ф., Беланов А. С. О максимальной концентрации энергии круглым диэлектрическим волноводом.— «Радиотехника и электроника», 1966, т. 11, № 4, стр. 748.
- 87*. Взятых В. Ф., Беланов А. С. Расчет коэффициента связи круглых диэлектрических волноводов.— «Радиотехника и электроника», 1968, т. 13, № 4, стр. 725.
- 88*. Веселов Г. И., Крехтунов В. М. О собственных колебаниях бесконечной системы диэлектрических волноводов.— «Радиотехника и электроника», 1969, т. 14, № 8, стр. 1399; «Радиотехника», 1970, т. 25, № 8, стр. 52.
- 89*. Веселов Г. И., Воронина Г. Г. Новый метод расчета открытых эллиптических волноводов, Теория электромагнитных волн (тезисы докладов международного симпозиума), изд-во «Наука», 1971, стр. 46.
- 90*. Веселов Г. И., Воронина Г. Г. К расчету открытого диэлектрического волновода прямоугольного сечения.— «Изв. вузов, Радиофизика», 1971, т. 14, № 12, стр. 1891.
- 91*. Гончаренко А. М. Электромагнитные свойства плоского анизотропного волновода.— «ЖТФ», 1967, т. 37, № 5, стр. 822.
- 92*. Гусак М. А., Гончаренко А. М. Влияние анизотропии на типы волн диэлектрического волновода.— «Вестн. АН БССР, сер. Физ.-мат. наук», 1965, № 1, стр. 90.
- 93*. Дерюгин Л. Н., Марчук А. Н., Сотин В. Е. Свойства плоских несимметричных диэлектрических волноводов на подложке из диэлектрика.— «Изв. вузов, Радиоэлектроника», 1967, т. 10, № 2, стр. 134.
- 94*. Дерюгин Л. Н., Марчук А. Н., Сотин В. Е. Резонансное возбуждение плоского диэлектрического волновода через закрытый слой плоской волной.— «Изв. вузов, Радиоэлектроника», 1970, т. 13, № 9, стр. 973.
- 95*. Дерюгин Л. Н., Марчук А. Н., Сотин В. Е. Излучение плоского диэлектрического волновода.— «Изв. вузов, Радиоэлектроника», 1970, т. 13, № 9, стр. 309.
- 96*. Дерюгин Л. Н., Чехлова Т. К. Исследование оптических микроволноводов на желатиновых пленках.— «Оптика и спектроскопия», 1973, т. 25, № 2, стр. 362.
- 97*. Добромислов В. С., Взятых В. Ф. Диэлектрические резонаторы с волнами шепчущей галереи.— «Труды МЭИ», 1973, вып. 161, стр. 78.
- 98*. Иогансен Л. В. Теория резонансных электромагнитных систем с полным внутренним отражением.— «ЖТФ», 1962, т. 32,

- № 4, стр. 406; 1963, т. 33, № 11, стр. 1323; 1966, т. 36, № 12, стр. 2056; 1968, т. 38, № 5, стр. 781.
- 99*. Любимов Л. А., Веселов Г. И., Бей Н. А. Диэлектрический волновод эллиптического сечения. — «Радиотехника и электроника», 1961, т. 6, № 11, стр. 1871.
- 100*. Мальцев В. П., Нефедов Е. И., Шевченко В. В. Биеция поверхностных волн в связанных пленочных световодах и их применение. — «Изв. вузов, Радиоэлектроника», 1970, т. 13, № 11, стр. 1381.
- 101*. Мальцев В. П., Миронов В. Л., Шевченко В. В. Поверхностные электромагнитные волны в пленках конечной ширины. — «Радиотехника и электроника», 1972, т. 17, № 8, стр. 1734.
- 102*. Мировицкий Д. И., Дубровин В. Ф. Кольцевые гибридные соединения на диэлектрических волноводах. — «Радиотехника и электроника», 1970, т. 15, № 12, стр. 2613.
- 103*. Мировицкий Д. И., Дубровин В. Ф., Баскаков В. В. Синфазные направленные разветвители поверхностной волны на диэлектрических волноводах. — «Антенны», 1971, № 12, стр. 65.
- 104*. Рожков Г. Д., Рябов Б. А. Характеристики диэлектрического волновода сильно вытянутой формы сечения (экспериментальные результаты). — «Труды МЭИ», 1973, вып. 161, стр. 90.
- 105*. Розанов М. Я., Шер Ю. А. Общий метод анализа связанных оптопленочных линий. — «Радиотехника и электроника», 1968, т. 13, № 3, стр. 565.
- 106*. Семенов Н. А. Закон парциальных мощностей (диэлектрического волновода). — «Радиотехника и электроника», 1963, т. 8, № 8, стр. 1476.
- 107*. Семенов Н. А., Черенков Г. А., Купалова М. Л. Исследование неоднородностей диэлектрического волновода, Использование оптических квантовых генераторов в современной науке и технике, Изд-во ЛДНТП, 1973.
- 108*. Семенов Н. А., Черенков Г. А., Ермохин М. И. Экспериментальное исследование оптических диэлектрических волноводов, Использование оптических квантовых генераторов в современной науке и технике, Изд-во ЛДНТП, 1973.
- 109*. Черенков Г. А. Теоретическое исследование потерь на рассеяние от внутренних нерегулярностей диэлектрического волновода. — «Сб. трудов ВЗЭИС», 1969, вып. 4, ч. 2, стр. 155.
- 110*. Черенков Г. А., Семенов Н. А. Анализ характеристик возбудителя диэлектрического волновода, сб. «Методы и устройства управления оптическим излучением», Тула, Изд-во ТПИ, 1973.
- 111*. Шевченко В. В. Потери на излучение в изогнутых волноводах поверхностных волн. — «Изв. вузов, Радиофизика», 1971, т. 14, № 5, стр. 768.
- 112*. Шевченко В. В. О поведении волновых чисел волн диэлектрических волноводов за критическим значением (среды с потерями). — «Изв. вузов, Радиофизика», 1972, т. 15, № 2, стр. 257.

**Газово-диэлектрические, металло-диэлектрические
и многослойные волноводы**

- 113*. Беланов А. С., Ежов Г. И. Уравнения и параметры собственных волн пятислойного диэлектрического волновода, сб. «Радиоэлектроника оптического диапазона», 1970, стр. 18.
- 114*. Беланов А. С., Ежов Г. И. Составляющие поля и характеристическое уравнение собственных волн круглого слоистого диэлектрического волновода, сб. «Взаимодействие излучения с веществом», 1972, стр. 205.
- 115*. Belanov A. S., Ešov G. I., Chernij W. W. Zur Ausbreitung der Wellen im geschichteten runden dielektrischen Wellenleiter, Archiv für Elektronik und Übertragungstechnik, 1973, 11 S. 494.
- 116*. Беланов А. С., Черный В. В. О прямоугольном диэлектрическом волноводе на подложке из металла, сб. «Взаимодействие излучения с веществом», 1972, стр. 263.
- 117*. Взятыхшев В. Ф., Рожков Г. Д., Рябов Б. А. Некоторые экспериментальные исследования композиционного диэлектрического волновода.— «Труды МЭИ», 1972, № 100, стр. 114.
- 118*. Зятицкий В. А., Казанцев Ю. Н. Задачи возбуждения волноводов класса «диэлектрический канал». — «Изв. вузов, Радиофизика», 1971, т. 14, № 10, стр. 1570.
- 119*. Казанцев Ю. Н. Полый диэлектрический световод с газовым наполнителем. — «Радиотехника и электроника», 1967, т. 12, № 6, стр. 1107.
- 120*. Казанцев Ю. Н. Собственные типы волн в газово-диэлектрических световодах (волноводах). — «Радиотехника и электроника», 1958, т. 13, № 7, стр. 1227.
- 121*. Казанцев Ю. Н. Затухание электромагнитных волн в газово-диэлектрическом волноводе. — «Радиотехника и электроника», 1969, т. 14, № 7, стр. 1165.
- 122*. Казанцев Ю. Н. Затухание собственных волн в широком волноводе с диэлектрическим покрытием конечной толщины. — «Радиотехника и электроника», 1970, т. 15, № 1, стр. 207.
- 123*. Казанцев Ю. Н. Затухание электромагнитных волн в диэлектрических каналах прямоугольного сечения. — «Радиотехника и электроника», 1970, т. 15, № 6, стр. 1140.
- 124*. Казанцев Ю. Н., Харлашкин О. А. Широкие волноводы прямоугольного сечения с малыми потерями. — «Радиотехника и электроника», 1971, т. 16, № 6, стр. 1063.
- 125*. Казанцев Ю. Н. Плавный изгиб газово-диэлектрического волновода. — «Радиотехника и электроника», 1973, т. 18, № 7.
- 126*. Маненков А. Б. Открытые резонаторы с диэлектрическими стенками. — «Изв. вузов, Радиофизика», 1971, т. 14, № 4, стр. 606.
- 127*. Марчук А. Н., Сотин В. Е. Потери в слоистых диэлектрических волноводах, Труды УДН им. П. Лумумбы, 1967, вып. 23, стр. 130,

- 128*. Мальцев В. П., Нефедов Е. И., Шевченко В. В. Вытекающие волны в волноводе из двух разделенных слоев.— «Изв. вузов, Радиофизика», 1969, т. 12, № 12 стр. 1855.
- 129*. Мелехин В. Н., Маненков А. Б. Диэлектрические трубы — открытые волноводы с малыми потерями и редким спектром.— «Электроника больших мощностей», 1969, № 6, стр. 161.
- 130*. Пружановский В. А. Металло-оптический волновод.— «ЖТФ», 1968, т. 38, № 6, стр. 1035.
- 131*. Пружановский В. А. Диаграмма излучения коаксиального металло-оптического волновода.— «ЖТФ», 1968, т. 38, № 10, стр. 1741.
- 132*. Пружановский В. А. Несимметричные волны в коаксиальном металло-оптическом волноводе.— «ЖТФ», 1970, т. 40, № 6, стр. 1161.
- 133*. Пружановский В. А. Волны низших типов в круглом металло-оптическом волноводе.— «ЖТФ», 1971, т. 41, № 10, стр. 2088.
- 134*. Рожков Г. Д., Беланов А. С., Взятых В. Ф. Расчет многослойного диэлектрического волновода.— «Труды МЭИ», 1972, вып. 100, стр. 102.
- 135*. Рожков Г. Д., Взятых В. Ф., Лопатникова Н. Д. Расчет характеристик основных волн в плоском экранированном диэлектрическом волноводе.— «Труды МЭИ», 1973, вып. 161, стр. 85.
- 136*. Черенков Г. А. Анализ круглого двухслойного диэлектрического волновода, окруженного безграничной диэлектрической средой.— «Сб. трудов ВЗЭИС», 1971, вып. 6, стр. 15.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Аберрация (aberration) 173, 182, 190, 191
- Аббе теорема синусов (Abbe's sine theorem) 161
- условие синусов (Abbe's sine condition) 163
- Автокорреляционная функция (autocorrelation function) 471
- Амплитуда поля (field amplitude) 190
- Амплитудно-передаточная функция (amplitude transmission function) 22, 183, 187
- Апертура в непрозрачном экране (aperture in opaque screen) 59
- круглая (circular aperture) 227, 283
- линзы (lens aperture) 179, 186
- Аргумент Торальдо де Франчия (Toraldo di Francia's argument) 223
- Берремана линза (Berreman's lens) 191
- Бесселя дифференциальное уравнение (Bessel's differential equation) 63, 74, 76
- Блоха функция (Bloch function) 272
- Бриллюэна рассеяние (Brillouin scattering) 100
- Брэгга дифракция (Bragg's diffraction) 49, 50, 88, 98
- рассеяние (Bragg's scattering) 96
- теория возмущений (Bragg's perturbation theory) 88
- — связанных волн (Bragg's coupled wave theory) 49
- условие (Bragg's condition) 96, 97, 98, 104
- Брэгга эффективность преобразования (Bragg's conversion efficiency) 81
- Брюстера угол (Brewster's angle) 44, 45, 46, 47
- Вариационный принцип (variation principle) 123, 124
- Возмущение синусоидальное (sinusoidal perturbation) 454, 462, 468
- случайное (random perturbation) 470, 473
- стенки (wall perturbation) 462, 463
- Волна направляемая (guided wave) 372
- нераспространяющаяся (evanescent wave) 23
- отраженная (reflected wave) 34, 35, 39, 45
- прошедшая (transmitted wave) 35
- плоская (plane wave) 22, 27, 33, 48, 117, 143
- — дифрагированная (diffracted plane wave) 102
- — произвольной формы (random plane wave) 20
- поперечная (transverse wave) 45
- световая (light wave) 50
- связанная (coupled wave) 49
- синусоидальная (sinusoidal wave) 20, 22
- цилиндрическая (cylindrical wave) 453
- Волновод диэлектрический (dielectric waveguide) 369
- — изогнутый (bend dielectric waveguide) 249, 508

- Волновод конфокальный линзовый (confocal lens waveguide) 232, 241, 257, 267
 — одномодовый (single mode waveguide) 370
 — плоский (slab waveguide) 392
 — — с нерегулярностью границ (slab waveguide with imperfect boundary) 371
 — цилиндрический (cylindrical waveguide) 371
 Волновое уравнение в полярных координатах (wave equation in polar coordinates) 374
 — — неоднородное (inhomogeneous wave equation) 61
 — — приведенное (reduced wave equation) 51, 115, 399
 — — скалярное (scalar wave equation) 18
 Волокно без оболочки (uncoated fibre) 372
 — в оболочке (cladded fibre) 368, 369, 375, 542
 — круглое (circular fibre) 371, 372
 — оптическое (optical fibre) 368, 369, 372
 — цилиндрическое (cylindrical fibre) 371, 372
 — — оптическое (cylindrical optical fibre) 371
 Вычислительное устройство (computer) 185, 189
- Газовая линза (gas lens) 191, 192, 194, 202
 Гамильтона механика (Hamilton's mechanics) 124, 128
 — оператор (Hamilton's operator) 136
 — принцип (Hamilton's principle) 114, 124, 125
 — уравнение (Hamilton's equation) 128, 129
 — — движения (Hamilton's equation of motion) 128
 Гамильтона — Якоби уравнение (Hamilton — Jacobi equation) 131, 132
- Гамильтониан в параксиальном приближении (Hamiltonian paraxial approximation) 129
 — нерелятивистский (Hamiltonian nonrelativistic) 129, 130
 — релятивистский (relativistic hamiltonian) 129
 Гармонический осциллятор (harmonic oscillator) 269, 347
 Гаусса кривизна фазового фронта (Gaussian phase front curvature) 300
 — преобразование (Gaussian transformation) 302
 Гелий-неоновый лазер (helium-neon laser) 192
 Геометрическая оптика (geometrical optics) 113, 114, 115, 119, 122, 131
 Гиперсфероидальная функция (hiperspheroidal function) 328
 Голография (holography) 50, 98, 183, 187
 Грина функция (Green's function) 51, 52, 54, 55, 62, 63
- Диаграмма поля в дальней зоне (far field pattern) 60, 71
 Диаметр связки волокон (bundle diameter) 520
 Диапазон механических частот (range of mechanical frequencies) 115, 391, 469
 — направляемых мод и мод излучения (range of guided modes and radiation modes) 410
 Дирака дельта-функция (Dirac's delta function) 51, 67
 — уравнение (Dirac's equation) 136
 Диспергирующая среда (dispersive medium) 20, 22
 Дисперсия (dispersion) 20, 23, 252, 390
 Дисторсия (искажение) луча (distortion of ray) 343
 — поля (distortion of field) 295, 296
 — фокусного расстояния (distortion of focal length) 210, 211

- Дифракционная решетка (diffraction grating) 50, 76, 77, 78, 94
- Дифракционный интеграл (diffraction integral) 174
- — двумерный (diffraction two-dimensional integral) 175
- — приближенный (diffraction approximate integral) 177
- Дифракция на круглом отверстии (diffraction by circular aperture) 49, 73
- — щели (diffraction by a slit) 49, 65
- Диэлектрический волновод (dielectric waveguide) 433
- излучатель (dielectric radiator) 65
- Допплера эффект (Doppler effect) 100
- Закон идеального газа (ideal gas law) 193
- преобразования гауссова пучка (Gaussian beam transformation law) 132
- формирования изображения (law of image formation) 177
- Затухание волны (wave attenuation) 355
- Зеркало (mirror) 101, 155
- Зоммерфельда условие излучения (Sommerfeld's radiation condition) 63
- Излучение (radiation) 72, 393, 402, 403, 454
- Изображение (image) 169, 190
- мнимое (virtual image) 169, 170
- «Импульс» световых лучей («momentum» of light rays) 170
- Импульсная амплитудная модуляция (pulse amplitude modulation) 88
- Интегральная теорема Коши (Cauchy's integral theorem) 95
- Интегральное представление функции Бесселя (integral representation of Bessel function) 86
- Интерференция (interference) 331, 463
- Квантовая теория лучей (quantum theory of rays) 114, 135
- Кирхгофа — Гюйгенса дифракционный интеграл (Kirchhoff — Huygens diffraction integral) 50, 55
- теория дифракции (Kirchhoff — Huygens diffraction theory) 272
- Клейна — Гордона уравнение (Klein — Gordon equation) 136, 137
- Когельника закон ABCD (Kogelnik's ABCD law) 312
- комплексный параметр пучка (Kogelnik's complex beam parameter) 303
- Когерентный свет (coherent light) 189, 190
- Колебание траектории луча (oscillation of ray trajectory) 251
- Коэффициент отражения (reflection coefficient) 39, 44
- связи (coupling coefficient) 457
- Круговая частота (angular frequency) 16
- Лазерный резонатор (laser resonator) 192, 238, 239, 241, 325
- Лагранжа уравнение (Lagrange equation) 128, 129, 209
- функция (Lagrange function) 125
- Лагранжиан (Lagrangian) 125, 128, 132
- Ламинарное течение (laminar flow) 192
- Ламинарный поток (laminar flow) 196
- Лепесток (lobe) 71
- диаграммы излучения (radiation lobe) 72, 454
- — направление (direction of radiation lobe) 454

- Линза газовая (gas lens) 192, 289
— сложная (complex lens) 166
— тонкая (thin lens) 166, 167, 168
— цилиндрическая (cylindrical lens) 174
Линзовое свойство среды с квадратичным законом изменения параметров (lens property of square law media) 339, 365
Линзовый волновод из круглых линз (circular lens waveguide) 175
— — конфокальный (confocal lens wave guide) 232, 241
— — с искривленной осью (lens waveguide with curved axis) 242, 243
— — с наклонной линзой (lens waveguide with tilt)
— — со случайно смещенной линзой (lens waveguide with random lens displacement) 248
— — со смещенными линзами (lens waveguide with displaced lenses) 248
Лиувилля теорема (Liouville's theorem) 161, 162, 163, 165, 269
Лорентца преобразование (Lorentz transformation) 114
Луч комплексный (complex ray) 116
— направляемый в линзовом волноводе (guided ray in lens waveguide)
— ненаправляемый (unguided ray) 216
Лучевая оптика (ray optics) 113
— — квадратичной среды (ray optics of square law medium) 148, 341
— — тонкой линзы (ray optics of thin lens) 167
Магнитная индукция (magnetic flux density) 16
Максвелла уравнение (Maxwell's equation) 113, 114, 123, 136
Максвелла уравнение в интегральной форме (Maxwell's equation in integral form) 14, 15
Метод стационарной фазы (stationary phase method) 67, 70, 71
Мода (mode) 27
— гибридная (hibrid mode) 380, 400
— затухающая (evanescent mode) 411
— излучения (radiation mode) 393, 402, 403, 409
— — волокна (radiation mode of fibre) 401
— квадратичной среды (mode of square law medium) 348
— лазерного резонатора (mode of laser resonator) 325
— линзового волновода (mode of lens waveguide) 401
— локальная, нормальная (local normal mode) 435
— направляемая (guided mode) 30, 372, 415
— — круглого волокна (guided mode of round fibre) 372
— ненаправляемая (unguided mode) 370
— неустойчивая (unstable mode) 361
— нечетная (odd mode) 468
— отраженная (reflected mode) 487
— паразитная (unwanted mode) 456, 461, 482
— поперечная (transverse mode) 380
— продольная (longitudinal mode) 333, 337
— произвольная (random mode) 409
— пучка Эрмита — Гаусса (Hermite—Gaussian beam mode) 308
— смешанная (mixed mode) 409, 415
— согласованная (matching mode) 321
— устойчивая (stable mode) 361
— четная (even mode) 468

- Монохроматический свет (monochromatic light) 190
Мощность моды (modal power) 404, 417
— — излучения (power of radiation mode) 404
— — рассеянная (scattered power) 476
Нарушение закона сохранения энергии (violation of energy conservation) 292
Неймана функция (Neumann function) 63, 375
Нестабильность траектории луча (unstable ray trajectories) 238
Неустойчивость резонатора (instability of resonator) 325
Оболочка (cladding) 518, 536
Оптика волновая (wave optics) 166, 167
— волоконная (fiber optics) 368
— пучка (beam optics) 298
— Фурье (Fourier optics) 178
Оптическая длина пути (optical path length) 121, 124, 187
Оптический волновод (optical waveguide) 368
— преобразователь (optical transformer) 182, 183, 184, 186, 187, 419
Оптическое аналоговое вычислительное устройство (optical analog computer) 166, 185, 199
Ортогональность векторов (orthogonality of vectors) 409, 410
— сфероидальных функций (orthogonality of spheroidal functions) 200
Отверстие в непрозрачном экране (hole in opaque screen) 73, 116
Ответвитель направленный (directional coupler) 549
— призмный (prism coupler) 520, 521
Отражение внутреннее (internal reflection) 398, 399
Отсечка (cutoff) 326, 386, 396
— условие (cutoff condition) 386
Отсечка, частота (cutoff frequency) 381, 382, 387, 397
Ошибка (error) 150
Параксиальное волновое уравнение (paraxial wave equation) 119, 140
— приближение (paraxial approximation) 114, 119, 129, 174, 179
— — уравнения луча (paraxial approximation of ray equation) 119
— уравнение линзы (paraxial lens equation) 204, 233
Переключатель (коммутатор) (commutator) 145
Перекрестная связь (cross talk) 520, 543, 548
Плавный переход (gradual taper) 483, 505
Плотность потока мощности (power flow density) 16
— тока (current density) 151, 152
Показатель преломления (refraction index) 368, 369
— — комплексный (complex index of refraction) 434
— — непрерывное изменение (index discontinuity) 434
— — параболическое распределение (parabolic index distribution) 119
Поле излучения (radiation field) 402
— оптического волокна (field optical fibre)
Положение луча (ray position) 149, 255
Попеременная фокусировка (alternating focusing) 233
Постоянная идеального газа (ideal gas constant) 193
— Планка (Planck's constant) 136, 137
— распространения (propagation constant) 378, 417
Потери мощности (losses of power) 282, 284, 444, 469
— — в линзовом волноводе (losses of power in lens waveguide) 284

- Потери на изгибе (bending losses) 508, 516
— на излучение ТЕ-моды (radiation loss of TE mode) 401
— — — ТМ-моды (radiation loss of TM mode) 402
Предельная разрешающая способность (resolution limit) 171
Представление дельта-функции (representation of delta function) 204
Преломление (refraction) 33, 125
Приближение дифракционного интеграла (approximation of diffraction integral) 56
— для моды HE_{11} (approximation for HE_{11} mode) 427
— — поля в дальней зоне (far field approximation)
— — функции Бесселя (approximation for Bessel function) 385, 424, 513
— — — Ханкеля (approximation for Hankel function) 513
Приближенное решение для ТЕ-мод (approximate solution for TE modes) 418, 419
— уравнение собственных значений (approximate eigenvalue equation) 380, 418
Принцип Гюйгенса (Huygens principle) 59, 60
— неопределенности (uncertainty principle) 148, 216
Пространственная когерентность (spatial coherence) 190
— фильтрация (spatial filtering) 190
Пучок (beam) 151, 190
— колебание ширины (beam width oscillation) 321
— наиболее узкое место (минимальная ширина пучка) (beam waist minimum) 304, 303
— параметр ширины (beam width parameter) 300, 319
— ширина (beam width) 302
— конфокального линзового волновода (beam width of confocal lens waveguide) 302
Радиус кривизны гауссова пучка (radius of curvature of Gaussi — an beam) 518
— — линзового волновода (radius of curvature of lens wave guide) 245
Разложение поля по модам (field expansion using modes) 370
Разрешающая способность (resolving power) 215
Распределение поля (field distribution) 189, 415
Распространение гауссова пучка в пространстве (propagation of Gaussian beam in space) 298, 299
Рассеяние мощности (power scattering) 370
— света (light scattering) 97, 98, 191, 479
— прямое (в прямом направлении) (forward scattering) 479
Резонансная частота резонатора лазера (resonant frequency of laser cavity) 337
Релятивистская механика (relativistic mechanics) 114, 132
— энергия частиц (relativistic energy of particles) 135, 136
Решение параксиального уравнения луча (solution of paraxial ray equation) 119, 204
Решетка трехмерная (three dimensional grating) 97, 100
Рэля критерий (Rayleigh criterion) 216, 218, 219, 226
— предельная разрешающая способность (Rayleigh resolution limit) 216
— распределение (Reyleigh distribution) 255
Свет некогерентный (incoherent light) 189
Световод (light guide) 391
Световой волновод (light waveguide) 148, 391
— луч (light ray) 116, 117, 120, 121
— пучок (light beam) 116, 117

- Свертка (convolution) 186, 188, 189
- Связанный волновод (coupled waveguide) 531
- Связка лучей (bundle of rays) 269
- Связь мод HE_{11} (coupling of HE_{11} modes) 536
- Сердцевина волокна (core of fibre) 369, 376
- Система формирования изображения (image forming system) 121
- Скалярная волновая теория (scalar wave theory) 271
- Скорость групповая (group velocity) 349, 389
- света (velocity of light) 124
- фазовая (phase velocity) 20, 509
- Смещение линз по синусоидальному закону (sinusoidal lens displacement) 248, 261, 543
- Собственное значение (eigenvalue) 141, 220, 223, 225
- — моды TE (eigenvalue of TE mode) 396
- — — TM (eigenvalue of TM mode) 400
- Соотношение между входным и выходным полем линзы (relation between input and output of lens) 176
- Сохранение мощности (conservation of power) 528
- Спектр механических частот (spectrum of mechanical frequencies) 470
- мощности (power spectrum) 473
- непрерывный (continuous spectrum) 370, 403
- функции искажения стенки (spectrum of wall distortion) 470, 473
- Спектральный анализ (spectral analysis) 80
- Среда неоднородная (inhomogeneous medium) 122
- однородная (homogeneous medium) 27, 105, 114
- Статистическая механика (statistical mechanics) 114
- Стационарная точка (stationary point) 68, 69
- фаза (stationary phase) 67
- Стержень гофрированный (corrugated rod) 461, 463
- круглый (circular rod) 498
- лазера (laser rod) 460, 461
- тефлоновый (teflon rod) 460
- Сужение (переход) (taper) 149, 499
- экспоненциальной формы (taper with exponential shape) 508
- Суперпозиция нормальных мод (superposition of normal modes) 271, 285
- Теорема дивергенции (divergence theorem) 152, 195
- Теплопроводность (thermal conductivity) 194
- Траектория волны в линзовом волноводе (wave trajectory in lens waveguide) 284
- — в фазовом пространстве (wave trajectory in phase space) 151
- Уравнение переноса тепла (heat transfer equation) 195
- разностное (difference equation) 244
- собственных значений (eigenvalue equation) 141, 378, 380
- — — графическое решение (graphical solution of eigenvalue equation) 396
- Условие граничное (boundary condition) 119
- начальное (initial condition) 440
- резонанса (resonance condition) 335
- согласования мод (mode matching condition) 321, 324
- Усреднение по ансамблю (ensemble average) 252, 474

- Фабри — Перо интерферометр (Fabry — Perot interferometer) 324, 325
— резонатор (Fabry — Perot cavity) 324, 325
Фазовая модуляция (phase modulation) 88
— решетка (phase grating) 84, 87, 96
— скорость (phase velocity) 115
Фазовый корректор (phase transformer) 173
Ферма принцип (Fermat's principle) 123, 124, 125, 126, 131
Фильтр пропускания верхних частот (high pass filter) 186
— — нижних частот (low pass filter) 184
Фокальная точка (focal point) 168
Фокусное расстояние газовой линзы (focal length of gas lens) 169, 177, 204
Формула тонкой линзы (thin lens formula) 167, 269
Фотографическая камера (photographic camera) 169
Фраунгофера дифракция (Fraunhofer diffraction) 60, 61, 72, 182
— область (Fraunhofer region) 182
Френеля дифракция (Fresnel diffraction) 60, 61
— приближение (Fresnel approximation) 222
Фурье амплитуда (Fourier amplitude) 268
— интеграл (Fourier integral) 20
— компонента (Fourier component) 268
— преобразование оптическое (optical Fourier transform) 178, 179, 181, 187, 188, 189
— ряд (Fourier series) 268
Ханкеля приближение (Hankel approximation) 64
Ханкеля функция (Hankel function) 63, 375
Центральная предельная теорема (central limit theorem) 254
Цилиндрическая симметрия (cylindrical symmetry) 26
— функция (cylindrical function) 63, 380, 383
Частотная модуляция (frequency modulation) 88
Численное решение (numerical solution) 204, 338
Шум (noise) 220, 225
Щель в непрозрачном (светонепроницаемом) экране (slit in opaque screen) 64, 184
— узкая (narrow slit) 72, 149
Эйконала уравнение (eikonal equation) 116, 117, 118, 119, 129, 131
Эйлера уравнение (Euler equation) 126, 128
Экран непрозрачный (opaque screen) 184
— с отверстием (screen with a hole) 259
Энергия частиц релятивистская (energy of particle, relativistic) 129
Эренфеста теорема (Ehrenfest's theorem) 138, 144, 145
Эрмита полином (Hermite polynomial) 276, 277
— пучок (Hermite beam) 307, 308
Эрмита — Гаусса пучка распространение (Hermite — Gaussian beam propagation) 307
— функция (Hermite — Gaussian function) 278
Эффект гравитационный (gravitational effect) 196
Якобиан (Jacobian) 155

ОГЛАВЛЕНИЕ

Предисловие редактора русского издания	5
Предисловие автора к русскому изданию	9
Предисловие	11
Глава 1. Волновая оптика	14
1.1. Введение	14
1.2. Уравнения Максвелла	15
1.3. Волновое уравнение	17
1.4. Оптические системы с цилиндрической симметрией . .	26
1.5. Граничные условия	30
1.6. Отражение и преломление на границе диэлектриков	33
Глава 2. Теория дифракции	49
2.1. Введение	49
2.2. Дифракционный интеграл Кирхгофа — Гюйгенса . .	50
2.3. Дифракция на щели в непрозрачном экране	65
2.4. Дифракция на круглом отверстии	73
2.5. Дифракционная решетка	76
2.6. Дифракция Брэгга. Теория возмущений	88
2.7. Дифракция Брэгга. Теория связанных волн	102
Глава 3. Геометрическая оптика	113
3.1. Введение	113
3.2. Вывод уравнений лучевой оптики из волнового уравнения	114
3.3. Граничные условия для световых лучей	120
3.4. Принцип Ферма	123
3.5. Гамильтонова формулировка лучевой оптики	127
3.6. Квантовая теория световых лучей	135
3.7. Теорема Лиувилля	150
Глава 4. Линзы	166
4.1. Введение	166
4.2. Лучевая оптика тонких линз	167
4.3. Волновая оптика тонких линз	171
4.4. Оптическое преобразование Фурье и пространственная фильтрация	182
4.5. Газовые линзы	190
4.6. Предельная разрешающая способность при формировании изображения	215
Глава 5. Линзовые волноводы	227
5.1. Введение	227
5.2. Лучевая оптика регулярного линзового волновода	228
5.3. Лазерные резонаторы	238
5.4. Линзовый волновод с искривленной осью	242
5.5. Линзовые волноводы со случайным смещением линз	251

5.6. Нормальные моды линзового волновода	270
5.7. Траектория волны в конфокальном линзовом волно- воде	284
5.8. Разрушение пучка в неидеальных линзовых волно- водах	289
Глава 6. Гауссовы пучки	298
6.1. Введение	298
6.2. Распространение гауссовых пучков в свободном про- странстве	299
6.3. Двойственная природа гауссовых пучков	304
6.4. Преобразование гауссовых пучков	309
6.5. Согласование мод	321
6.6. Резонаторы лазера	325
Глава 7. Распространение света в средах с квадратичным законом изменения показателя преломления	339
7.1. Введение	339
7.2. Лучевая оптика квадратичной среды	341
7.3. Моды квадратичной среды	344
7.4. Внеосевые пучки в квадратичной среде	350
7.5. Квадратичная среда с потерями или усилением	355
7.6. Линзовые свойства квадратичной среды	365
Глава 8. Оптические волокна и диэлектрические волноводы	368
8.1. Введение	368
8.2. Направляемые волны (моды) круглых оптических волокон	372
8.3. Направляемые моды плоского волновода	392
8.4. Моды излучения плоского волновода	401
8.5. Соотношения ортогональности	411
8.6. Полезные приближения	417
Глава 9. Нерегулярные диэлектрические волноводы	
9.1. Введение	433
9.2. Плоский волновод с неровной границей	434
9.3. Плоский волновод с синусоидальными искажения- ми стенки	445
9.4. Случайные искажения стенки	470
9.5. Скачкообразные и плавные изменения толщины пло- ского волновода	483
9.6. Потери на излучение в изогнутых волноводах	508
Глава 10. Связь между диэлектрическими волноводами	519
10.1. Введение	519
10.2. Уравнения связанных волн	522
10.3. Связанные плоские волноводы	531
10.4. Связь мод HE_{11} двух круглых волокон	536
10.5. Перекрестная связь	543
Литература	550
Предметный указатель	567

