МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ ИМЕНИ М.В.ЛОМОНОСОВА
ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ

КОРОЛЕНКО П.В.

ОПТИКА КОГЕРЕНТНОГО ИЗЛУЧЕНИЯ

МОСКВА
1997

УДК 535.4

Оптика когерентного излучения. К о р о л е н к о П. В., учебное пособие.

Систематически изложены вопросы формирования и распространения световых когерентных пучков в разнообразных оптических системах и передающих средах. Рассмотрены современные способы записи и обработки оптической информации. Дается представление о возможностях компьютерных методов. Изложены основные принципы и тенденции обновления элементной базы устройств когерентной оптики на основе использования голограмм, киноформов, дифракционных и адаптивных оптических элементов.

Пособие подготовлено в обычном печатном и электронном вариантах, что существенно расширяет возможности как изложения, так и освоения учебного материала.

Для студентов и аспирантов, специализирующихся в области физической и вычислительной оптики, систем оптической связи, а также лазерной физики.

Вариант учебного пособия для системы INTERNET подготовила электронная группа в составе: Маганова М.С., Петрова Г.В., СердюченкоА.Ю., Федотов Н.Н. (руководитель группы), Эмбаухов С.В. Его можно получить по адресу: `ftp://optics.npi.msu.su/pub/co/`

Оглавление

ОГЛАВЛЕНИЕ 3
ВВЕДЕНИЕ 7
ГЛАВА 1. ОБЩЕТЕОРЕТИЧЕСКИЕ ПОЛОЖЕНИЯ 10
1.1. Когерентность 10
1.1.1. Степень когерентности светового пучка 11
1.1.2. Методы измерения пространственной и временной когерентности 12
1.2. Основные положения волновой оптики 17
1.2.1. Волновое уравнение 17
1.2.2. Теория дифракции по Кирхгофу 20
1.2.3. Дифракция Френеля 25
1.2.4. Дифракция Фраунгофера. Элементы Фурье-оптики 27
1.2.5. Принцип Бабине. Эффект Талбота 32
1.2.6. О возможности обобщения метода Кирхгофа на случай векторных полей 33
1.3. Методы геометрической оптики 35
1.3.1. Геометрооптическое приближение 35
1.3.2. Лучи и фронты 36
1.3.3. Принцип Ферма 38
1.3.4. Лучевые трубки 39
1.3.5. Точка стационарной фазы. Область влияния 41
1.3.6. Условие применимости геометрической оптики 44
1.3.7. Каустики 45
1.3.8. Элементы гамильтоновой оптики 48
Литература к главе 1 50
ГЛАВА 2. ФОРМИРОВАНИЕ И РАСПРОСТРАНЕНИЕ ВОЛНОВЫХ ПУЧКОВ 51
2.1 Моды свободного пространства 51
2.1.1 Параболическое приближение 51
2.1.2 Свойства основной моды 52
2.1.3. Моды высших порядков 55
2.1.4. Преобразование волновых пучков с помощью линз 59

2.1.5. Геометрооптическое описание распространения и преобразования волновых пучков 61
2.1.6. Расчет поля дифрагированной волны методом разложения по модам свободного пространства 65
2.2. Моды оптических резонаторов 67
2.2.1. Геометрооптическое описание внутрирезонаторных полей 67
2.2.2. Моды устойчивых резонаторов в приближении бесконечной апертуры 69
2.2.3. Согласование резонаторов 74
2.2.4. Моды резонаторов при ограниченной апертуре зеркал 75
2.2.5. Неустойчивые резонаторы 78
2.2.6. Резонаторы, применяемые для селекции мод 80
2.2.7. Кольцевые резонаторы 84
2.2.8. Модульные системы 86
2.3. Волноводное распространение излучения 88
2.3.1. Квадратичные среды 88
2.3.2. Оптические волокна 92
2.3.3. Полые волноводы 94
2.4. Распространение когерентного излучения в среде со случайными неоднородностями 97
2.4.1. Борновское приближение и приближение Рытова 97
2.4.2. Флуктуации уровня и фазы 100
2.4.3. Характеристики излучения в случайной среде с гауссовой функцией корреляции 103
2.4.4. Колмогоровская модель турбулентной атмосферы 107
2.5. Формирование спекл-полей при взаимодействии света с диффузными объектами 110
2.5.1. Физическая природа спеклов и их размеры 110
2.5.2. Спекл-фотография и спекл-интерферометрия 111
2.6. Стохастизация световых пучков в каналах с регулярным распределением неоднородностей. Оптический хаос и фрактальные структуры лучей 117
2.6.1. Уравнения траектории луча 117
2.6.2. Нелинейный лучевой резонанс 119
2.6.3. Фрактальные лучевые структуры 120

2.7. Пучки с винтовыми дислокациями волнового фронта. Элементы сингулярной оптики 123
2.7.1. Общая характеристика дислокаций волнового фронта 123
2.7.2. Методы регистрации ВД 125
2.7.3. Оптические вихри в случайно-неоднородных средах 127
2.7.4. Генерация винтовых полей в лазерах. 132
Литература к главе 2 138
ГЛАВА 3. ЗАПИСЬ И ОБРАБОТКА ОПТИЧЕСКОЙ ИНФОРМАЦИИ 140
3.1. Общая характеристика оптических систем 140
3.2. Однолинзовая система 144
3.2.1. Линзы как элементы, выполняющие преобразование Фурье 144
3.2.2. Формирование изображения 147
3.3. Получение изображений в сложных системах 153
3.3.1. Дифракционно-ограниченные системы. Теории Аббе и Релея 153
3.3.2. Роль аподизации 156
3.4. Учет аберраций 157
3.5. Голографическая запись информации 159
3.5.1. Принцип голографической записи. 159
3.5.2. Голограммы Фурье 161
3.6. Оптическая фильтрация и распознавание образов 168
3.6.1 Применение системы 4-F 168
3.6.2. Голографический метод синтезирования пространственных фильтров и проблема апостериорной обработки информации 171
3.7. Сопоставление методов когерентной и некогерентной оптики 173
3.8. Характеристики качества изображения 175
Литература к главе 3 177
ГЛАВА 4. МЕТОДЫ КОМПЬЮТЕРНОЙ ОПТИКИ 178
4.1. Задачи компьютерной оптики 179
4.2. Цифровая голография 180
4.2.1. Общая процедура изготовления синтезированной голограммы 181

4.2.2. Получение цифровой голограммы Фурье и ее бинаризация 183
4.2.3. Киноформ 187
4.3. Фазовая проблема в оптике. Создание на основе решения обратных задач нового класса оптических элементов 189
4.3.1. Извлечение фазовой информации из данных об интенсивности 189
4.3.2. Особенности расчета характеристик фокусаторов и корректоров излучения 191
4.3.3. Дифракционные оптические элементы 193
4.3.4. Создание фокусаторов на основе управляемых зеркал. 197
4.4. Фокусировка излучения при наличии случайных помех. Использование методов адаптивной оптики 199
4.5. Оптические элементы для анализа и формирования поперечного состава излучения 202
4.6. Цифровая обработка полей в оптических системах 204
4.6.1. Виды обработки оптических полей 205
4.6.2. Автоматизированная измерительная система для диагностики структуры лазерных пучков 206
Литература к главе 4 211
ПРИЛОЖЕНИЕ 1. АЛГОРИТМ БЫСТРОГО ПРЕОБРАЗОВАНИЯ ФУРЬЕ 213
ПРИЛОЖЕНИЕ 2. ОЦЕНКА ФАЗОВЫХ ВОЗМУЩЕНИЙ ПО ЗНАЧЕНИЯМ ЧИСЛА ШТРЕЛЯ 217
1. О фазовой проблеме. 217
2. Описание метода. 218
3. Метод оценки точности. 218
4. Результаты и выводы. 219
Список литературы 222

ВВЕДЕНИЕ

Оптика когерентного излучения является частью современной физической оптики. Ее предмет составляют физические процессы, связанные с формированием и распространением когерентного излучения в разнообразных оптических системах и передающих средах. Бурное развитие оптики когерентного излучения в последние десятилетия непосредственным образом обусловлено достижениями лазерной физики. Ведущиеся широким фронтом исследования уникальных характеристик лазерных пучков обогатили знания о свойствах когерентного света. При этом процесс изучения новых оптических явлений и закономерностей с использованием лазеров происходил так быстро, что стал наблюдаться определенный разрыв между вновь развиваемыми теоретическими представлениями и традиционными положениями классической оптики. Этому способствовал и тот факт, что в физике лазеров новые данные очень часто возникали на стыке различных научных направлений, и их интерпретацией занимались исследователи, представляющие разные школы и специальности. Следует учитывать также произошедшее в "лазерную эпоху" необычайно широкое внедрение оптических методов исследования в самые разные научные области, часто значительно отличающиеся как природой изучаемых объектов, так и используемым теоретическим аппаратом. Такое экстенсивное расширение оптических понятий и представлений, все возрастающая неопределенность в характеристике предмета современной когерентной оптики, отсутствие единой теоретической основы стали негативно сказываться на процессе сопоставления и обобщения данных, полученных различными авторами, и определили, в конечном счете, устойчивую тенденцию к объединению различных частных теорий на основе известных положений классической оптики. Естественно, такое объединение с самого начала предполагало и определенную модернизацию этих положений, и расширение понятийного и математического аппаратов. Материал данного учебного пособия по замыслу автора должен отражать указанную тенденцию.

Все наиболее важные оптические явления, происходящие в когерентном свете, рассмотрены в учебном пособии с единых позиций: как правило, на основе анализа различных форм решения приведенного волнового уравнения. При этом во всех случаях, оправданных с точки зрения более полного и точного описания физической сути рассматриваемых процессов, делается попытка связать новые сведения с ранее известными классическими представлениями, как с точки зрения терминологии, так и методов описания. Даже при изложении новых, во

многом находящихся в стадии становления разделов когерентной оптики, к которым может быть отнесена, в частности, значительная часть компьютерной оптики, освещение соответствующего теоретического и математического аппаратов производится с позиции дальнейшей эволюции ранее известных подходов и методов, в том числе и методов фурье-оптики.

Если попытаться сформулировать наиболее общую задачу, решаемую в рамках когерентной оптики, то это можно сделать следующим образом. Пусть в плоскости $П_1$ (**рис. В1**) задано амплитудно-фазовое распределение когерентного поля световой волны, движущейся в направлении, показанном стрелкой. Требуется найти распределение амплитуды и фазы в плоскости $П_2$ в предположении, что между плоскостями $П_1$ и $П_2$ располагается какая-либо оптическая система или передающая среда. Сформулированная задача является основной прямой задачей оптики когерентного излучения. Несмотря на простоту ее формулировки, ее решение в общем случае является нетривиальным. В отдельных случаях решение этой задачи требует использования весьма сложных методов и подходов. Исходя из приведенной схемы, можно

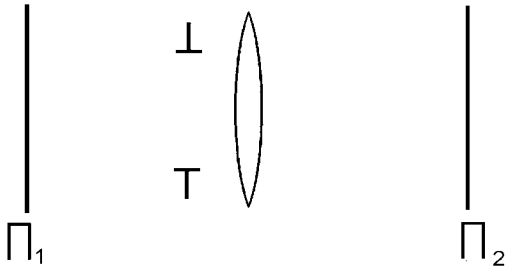


Рис. В1. *Общая схема задач, решаемых в оптике когерентного излучения*

сформулировать два варианта обратных задач. Согласно первому варианту по известному распределению поля в плоскости $П_2$ и известным параметрам системы, располагающейся между плоскостями $П_1$ и $П_2$, следует определить распределение поля в плоскости $П_1$. Второй вариант сводится к определению характеристик системы между заданными плоскостями $П_1$ и $П_2$ по известным на них распределениям поля. С указанными основными физическими задачами когерентной оптики граничит широкий круг задач, связанных с разнообразными прикладными вопросами. Он включает проблему управления параметрами излучения, разработку методов обработки световых структур и реконструкции оптических изображений, а также записи и обработки оптической формации. Приведенная схема и

классификация решаемых задач составляет методическую основу анализа рассматриваемых явлений.

В материале учебного пособия естественным образом нашли свое отражение научные интересы автора, а также его коллег, работающих в области когерентной оптики в ряде ведущих российских вузов и институтов РАН. В нем учтены многолетние традиции преподавания курса оптики когерентного излучения на кафедре оптики и спектроскопии физического факультета МГУ, а также опыт использования приобретенных знаний выпускниками кафедры на практике, в ходе научно-исследовательских работ в различных НИИ и ОКБ. Именно исходя из запросов практики, в пособие включены некоторые разделы, которые обычно включаются в руководства по статистической оптике. К ним, в частности, относятся элементы теории когерентности и оптики случайно-неоднородных сред. Это связано с тем, что при распространении излучения через некоторые оптические системы и передающие среды происходит заметное изменение степени его когерентности. Благодаря влиянию ответственных за это физических факторов в когерентных световых колебаниях появляется случайная составляющая, без учета которой невозможно корректное описание изучаемых оптических явлений. Однако, несмотря на стремление автора максимально обобщить современное понимание предмета когерентной оптики и ее содержательной части, круг вопросов, включенных в пособие, и характер их освещения не может претендовать на исчерпывающую полноту, хотя бы из естественных ограничений объема пособия. В частности, по последней причине, исключены из рассмотрения разнообразные нелинейные эффекты, происходящие в поле когерентного излучения. Предполагается, что читатель сможет самостоятельно удовлетворить свой интерес к слабо освещенным вопросам, используя приводимые в пособии развернутые библиографические сведения. Для удобства обращения к используемым источникам информации заголовок каждого параграфа содержит соответствующие литературные ссылки. Дополнительную информацию о новых направлениях физической оптики и наиболее интересных научных результатах, полученных в последнеее время, можно получить из приложения "Семинарий". Семинарий содержит постоянно обновляемое изложение докладов, сделанных на семинаре по когерентной оптике кафедры оптики и спектроскопии физического факультета МГУ.

Автор полагает, что данное учебное пособие помимо использования непосредственно в учебном процессе при подготовке специалистов-оптиков может оказаться полезным справочным руководством в области физической оптики для широкого круга исследователей и инженеров- практиков.

ГЛАВА 1. ОБЩЕТЕОРЕТИЧЕСКИЕ ПОЛОЖЕНИЯ

Теоретической основой анализа оптических явлений в когерентном свете являются положения классической оптики, достаточно полно изложенные в известных трудах [1,2], а также в последующих монографиях и учебниках (см., например, [3 - 6]). При этом особенно большую роль играют те разделы оптики, в которых рассматриваются процессы распространения, интерференции и дифракции излучения. В данной главе мы рассмотрим эти процессы и явления, используя подход, основанный на анализе решений приведенного волнового уравнения. Однако, прежде чем приступить к изложению основ теории дифракции и интерференции, уточним фундаментальное понятие когерентности, к которому нам придется постоянно апеллировать в процессе изложения материала учебного пособия.

1.1. Когерентность [7]

В оптике понятие когерентности вводится для характеристики скоррелированности световых колебаний в различных точках пространства и в различные моменты времени. Поэтому наиболее логично степень когерентности определять посредством корреляционной функции светового поля. Рассмотрим для простоты поляризованное поле, вектор напряженности электрического поля **E** в котором колеблется в определенном направлении. Если вектор напряженности содержит компоненту, случайным образом изменяющуюся по пространственным координатам **r** и по времени t , то можно построить следующую корреляционную функцию

$$B(\mathbf{r}_1, t_1, \mathbf{r}_2, t_2) = \left\langle \mathbf{E}(\mathbf{r}_1, t_1)\mathbf{E}^*(\mathbf{r}_2, t_2) \right\rangle, \qquad \textbf{(1.1.1)}$$

где угловые скобки означают усреднение по всему пространству и по всему интервалу времени наблюдения, а "звездочка" при втором множителе обозначает комплексно сопряженную величину. Для полей, статистические характеристики которых во времени не меняются (такие поля называются стационарными),

$$B(\mathbf{r}_1, t_1, \mathbf{r}_1, t_2) = B(\mathbf{r}_1, \mathbf{r}_1, t_2 - t_1). \qquad (1.1.2)$$

Принято выделять также статистически однородные поля, для которых корреляционная функция зависит лишь от разности $\mathbf{r}_2$ - $\mathbf{r}_1$

$$B(\mathbf{r}_1, t_1, \mathbf{r}_2, t_2) = B(\mathbf{r}_2 - \mathbf{r}_1, t_1, t_2). \qquad (1.1.3)$$

Однородное случайное поле называется изотропным, если корреляционная функция зависит лишь от абсолютного значения расстояния между двумя точками $\rho = |\mathbf{r}_2 - \mathbf{r}_1|$ Таким образом, для стационарных, однородных и изотропных полей с изменяющимся по случайному закону вектора **E**

$$B(\mathbf{r}_1, t_1, \mathbf{r}_2, t_2) = B(\mathbf{r}_2 - \mathbf{r}_1, t_2 - t_1), \qquad (1.1.4)$$

где $\tau = t_2 - t_1$. Корреляционная функция $B(\rho, \tau)$ принимает максимальное значение при $\rho = \tau = 0$.

1.1.1. Степень когерентности светового пучка

Введем теперь применительно к световому пучку нормированную корреляционную функцию

$$\gamma(\mathbf{r}_2, \mathbf{r}_1, t_2, t_1) = \frac{B(\mathbf{r}_2, \mathbf{r}_1, t_2, t_1)}{\sqrt{I(\mathbf{r}_1, t_1) I(\mathbf{r}_2, t_2)}}, \qquad (1.1.5)$$

где $I(\mathbf{r}_1, t_1) I(\mathbf{r}_2, t_2)$ - значение интенсивности пучка в указанных пространственных точках и в указанные моменты времени. В случае стационарности поле светового пучка

$$\gamma(\mathbf{r}_2, \mathbf{r}_1, \tau = t_2 - t_1) = \frac{B(\mathbf{r}_2, \mathbf{r}_1, \tau)}{\sqrt{I(\mathbf{r}_1, t_1) I(\mathbf{r}_2, t_2)}}. \qquad (1.1.6)$$

Построенную таким образом величину γ называют комплексной степенью когерентности, так как корреляционные функции в общем случае комплексны. Абсолютную величину γ называют модулем степени когерентности или просто степенью когерентности. Степень когерентности всегда удовлетворяет неравенству

$$|\gamma(\mathbf{r}_2, \mathbf{r}_1, \tau)| \le 1. \qquad (1.1.7)$$

$|\gamma|$ при $\tau = 0$ дает значение степени пространственной когерентности, а при $\mathbf{r}_2 = \mathbf{r}_1$ - значение степени временной когерентности. Значения $\rho = \rho_k$ и $\tau = \tau_k$, при которых степени пространственной и временной

когерентности уменьшаются в два раза называются соответственно размером зоны когерентности и временем когерентности.

1.1.2 Методы измерения пространственной и временной когерентности

Рассмотрим кратко оптические методы экспериментального определения пространственных и временных корреляционных функций, или, в терминах оптики, методы измерения пространственной и временной когерентности световых полей. Исторически понятие когерентности возникло в оптике в связи с интерпретацией результатов интерференционных опытов. Классические интерференционные опыты Юнга и Майкельсона оказываются прямыми методами измерения пространственных и временных корреляционных функций; распределение средней интенсивности в интерференционной картине непосредственно дает корреляционную функцию поля. Одновременно эти опыты можно рассматривать как схемы, поясняющие физический смысл пространственных и временных корреляционных функций. Обратимся к их рассмотрению. Начнем с определения пространственной когерентности с помощью интерферометра Юнга.

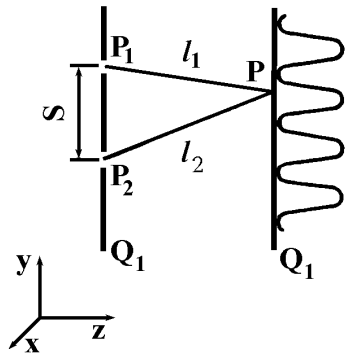


Рис. 1.1.1 *Схема интерферометра Юнга.*

Интерферометр Юнга представляет собой непрозрачный экран, в котором на некотором расстоянии s друг от друга вырезаны два малых отверстия P_1 и P_2 (**рис. 1.1.1**). Пусть на такой экран перпендикулярно падает случайная линейно поляризованная волна, поле которой $E(\mathbf{r},t)$ будем считать стационарным и однородным. Волновые пучки, исходящие из отверстий P_1 и P_2, интерферируют на экране Q_2, расположенном на некотором расстоянии от экрана Q_1.

Обозначим комплексное поле в точке P_j (i=1,2) через $E(P_j,t)$, а расстояние между точкой P_j и произвольной точкой P экрана Q_2

через $l_j=P_jP$. Суммарное электрическое поле в точке P от двух отверстий равно

$$E(P,t) = K_1E(P_1,t-t_1) + K_2E(P_2,t-t_2), \qquad (1.1.8)$$

где $t_j=l_j/c$ - время запаздывания (дисперсией среды пренебрегаем).

Коэффициенты передачи K_1 и K_2 являются комплексными величинами, их абсолютные значения зависят от формы и размеров отверстий.

Для средней интенсивности в точке P получаем

$$I(P) = \left\langle \left|E(P,t)\right|^2 \right\rangle = K_1K_1^*I_1 + K_2K_2^*I_2 +$$
$$K_1K_2^*\left\langle E(P_1,t-t_1)E^*(P_2,t-t_2)\right\rangle +$$
$$K^*_1K_2\left\langle E^*(P_1,t-t_1)E(P_2,t-t_2)\right\rangle$$

или

$$I(P) = |K_1|^2 I_1(P_1)+|K_2|^2 I_2(P_2) + 2|K_1K_2|\mathrm{Re}B(s,t). \qquad (1.1.9)$$

Здесь $I(P_1), I_2(P_2)$ - интенсивности светового поля в точках P_1 и P_2, $B(s,\tau)$ - пространственно-временная корреляционная функция:

$$B(s,\tau) = B(s,t) = \left\langle E(P_1,t-t_1)E^*(P_2,t-t_2)\right\rangle, \qquad (1.1.10)$$

t_2-t_2=τ, s- расстояние между точками P_1 и P_2 на экране Q_1; при этом учтена статистическая стационарность и однородность поля.

Если открыто лишь одно из отверстий в экране Q_1, то в точке P интенсивность, очевидно, равна

$$I_1 = I_1(P) = |K_1|^2 I_1(P_1) \quad \text{Î Ë· Ó} \quad I_2 = I_2(P) = |K_2|^2 I_2(P_2).$$

Пользуясь этими обозначениями, выражение (1.1.9) можно переписать в виде

$$I(P) = I_1 + I_2 + 2\sqrt{I_1I_2}\,\mathrm{Re}\,\gamma(s,t),$$
$$\gamma(s,t) = \frac{B(s,t)}{\sqrt{I_1I_2}}, \qquad (1.1.11)$$

где $\gamma(s,t)$ - комплексная степень когерентности.

Для электромагнитного поля вида

$$\mathbf{E} = \mathbf{e}A(x,y,z,t)\exp i(\omega_0 t - k_0 z)$$

(**e** - единичный вектор поляризации волны, A - медленно меняющаяся амплитуда волны) корреляционная функция (1.1.4) равна

$$B(\mathrm{s},\tau)=\left\langle A(r,t)\,A^{*}(r+s,t+\tau)\right\rangle \exp i(-\omega_0\tau)=B_{\perp}(\mathrm{s},\tau)\exp i(-\omega_0\tau).$$

Индекс $\perp$ означает, что корреляция оценивается в направлении, перпендикулярном оси z.

Следовательно,

$$\gamma(\mathrm{s},\tau)=\left|\gamma(\mathrm{s},\tau)\right|\exp i\left[\alpha(\mathrm{s},\tau)-\omega_0\tau\right], \qquad (1.1.12)$$

где $\alpha(s,\tau)=\arg B_{\perp}(s,\tau)$. Таким образом, выражение (1.1.11) принимает вид

$$\mathrm{I(P)}=\mathrm{I}_1+\mathrm{I}_2+2\sqrt{\mathrm{I}_1\mathrm{I}_2}\left|\gamma(\mathrm{s},\tau)\right|\cos\left[\alpha(\mathrm{s},\tau)-\delta\right]. \qquad (1.1.13)$$

Параметры δ и τ равны соответственно:

$$\tau=\frac{I_2-I_1}{c},\quad \delta=\omega_0\tau=\frac{2\pi}{\lambda_0}(I_2-I_1),$$

где λ_0- средняя длина волны. При $\tau<<\tau_к$ зависимость от τ в (1.1.13) входит только через δ, так что максимальные и минимальные значения интенсивности на экране Q_2 (**рис. 1.1.1**) определяются выражением

$$I_{\substack{\max\\ \min}}=I_1+I_2\pm 2\sqrt{I_1I_2}\left|\gamma(s)\right|. \qquad (1.1.14)$$

Контраст интерференционной картины, следуя Майкельсону, обычно характеризуют величиной

$$v=\frac{I_{\max}-I_{\min}}{I_{\max}+I_{\min}}, \qquad (1.1.15)$$

которую называют видностью. В соответствии с (1.1.14) для видности в окрестности точки P имеем

$$v(P)=\frac{2\sqrt{I_1I_2}}{I_1+I_2}\left|\gamma(s)\right|=2\left\{\sqrt{\frac{I_1}{I_2}}+\sqrt{\frac{I_2}{I_1}}\right\}^{-1}\left|\gamma(s)\right|. \qquad (1.1.16)$$

Если интенсивности интерферирующих пучков одинаковы (I_1=I_2), то значение (1.1.16) максимально и

$$v=\left|\gamma(S)\right|=\frac{\left|B_{\perp}(s)\right|}{\sqrt{I_1I_2}}, \qquad (1.1.17)$$

т.е. видность интерференционной картины просто равна степени пространственной когерентности. На **рис. 1.1.2** приведено

распределение интенсивности в интерференционной картине для различных видностей.

В общем случае видность (1.1.15) дает информацию о степени пространственно-временной когерентности. Если время задержки $\tau \sim \tau_к$, то видность будет зависеть от τ:

$$v = v(\tau) = |\gamma(s, \tau)|. \qquad (1.1.18)$$

Таким образом, если временная задержка меньше времени корреляции $\tau \ll \tau_к$, то интерферометр Юнга позволяет определить поперечную пространственную когерентность. Если мы хотим измерить не искаженную пространственной статистикой временную корреляционную функцию поля, следует обратиться к другой интерференционной схеме - интерферометру Майкельсона.

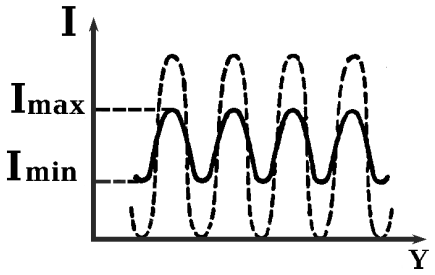


Рис. 1.1.2. *Распределение интенсивности в интерференционной картине для различных видностей*

Понятие временной когерентности прямо связано с интерференционным экспериментом, схема которого изображена на **рис. 1.1.3**. Волна падает на наклонную полупрозрачную пластинку *П* интерферометра Майкельсона, формирующую два пучка. Эти пучки отражаются от зеркал $З_1$ и $З_2$. Затем один из них, пройдя через пластинку *П*, а другой, отразившись от нее, поступают на экран *Q*, где интерферируют. В плоскости экрана расположен детектор, измеряющий интенсивность (например, фотодетектор, величина тока которого пропорциональна средней интенсивности).

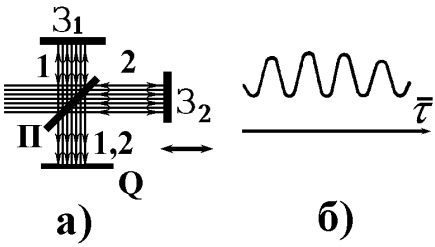


Рис. 1.1.3. *Схема интерферометра Майкельсона (а) и изменение интерференционного сигнала на экране Q (б).*

Если напряженность электрических полей пучков равна соответственно E_1 и E_1, то поле на экране *Q* равно

$$E(r,t) = E_1(r,t-t_1) + E_2(r,t-t_2), \qquad (1.1.19)$$

где $t_j=2l_j/c$, l_j- расстояние от зеркала 3_j до пластинки $П$.

Выражение (1.1.19) аналогично (1.1.8).Поэтому расчеты, подобные выполненным выше, приводят к выражению для средней интенсивности

$$I = I_1 + I_2 \pm 2\sqrt{I_1 I_2}\,|\gamma(\tau)|\cos[\alpha(\tau) - \omega_0\tau], \qquad (1.1.20)$$

которое сходно с (1.1.13) ($\tau=t_2-t_1$).

Таким образом, изменяя временную задержку в схеме интерферометра Майкельсона от $\tau=0$ до $\tau\to\infty$, из графика распределения средней интенсивности в интерференционной картине (интерферограмме) можно непосредственно определить временную корреляционную функцию светового поля.

Как и для интерферометра Юнга, для интерферометра Майкельсона можно ввести понятие видности интерференционной картины. В данном случае им удобно пользоваться, если волна квазимонохроматическая, т.е. $\frac{\Delta\omega}{\omega_0} << 1$; для такой волны, используя (1.1.15), для видности интерференционной картины в интерферометре Майкельсона вблизи заданного значения τ при $I_1=I_2$ имеем

$$v = |\gamma(\tau)|. \qquad (1.1.21)$$

Кроме интерферометров Юнга и Майкельсона существует большое число и других схем, используемых для измерения временной и пространственной когерентности оптических полей. Все многообразие интерферометров базируется на двух методах: методе деления амплитуды и методе деления волнового фронта. В методе деления амплитуды исходный пучок делится на частично отражающих или частично пропускающих оптических элементах. В методе деления волнового фронта пучок, проходя через отверстия, делится на несколько пучков.

Согласно такой классификации интерферометр Юнга - это интерферометр с делением фронта, интерферометр Майкельсона - интерферометр с делением амплитуды. Очевидно, интерферометр Майкельсона обладает большей светосилой, чем схема Юнга.

Проведение измерений пространственной и временной когерентности имеет большое значение для постановки

экспериментов в области когерентной оптики. Как правило, такие измерения должны проводиться по отношению к любым источникам излучения, используемых в оптических исследованиях.

1.2. Основные положения волновой оптики [1-6, 8-12]

Волновой оптикой называют раздел физической оптики, изучающей явления, в которых проявляется волновая природа света. В этом разделе в кратком изложении сформулированы основные теоретические положения волновой оптики и приведены наиболее важные соотношения и уравнения, положенные в основу всего дальнейшего рассмотрения.

1.2.1. Волновое уравнение

По своей физической природе световые волны являются волнами электромагнитными. Поэтому волновая оптика непосредственно основывается на уравнениях Максвелла.

Уравнения Максвелла связывают вектор напряженности электрического поля **E** и вектор электрической индукции **D** с вектором напряженности магнитного поля **H** и вектором индукции **B**. В отсутствие токов и свободных электрических зарядов они имеют вид:

$$\mathrm{rot}\mathbf{H} = \frac{\partial \mathbf{D}}{\partial t}, \qquad \textbf{(1.2.1)}$$

$$\mathrm{rot}\,\mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \qquad \textbf{(1.2.2)}$$

$$\mathrm{div}\mathbf{B} = 0, \qquad \textbf{(1.2.3)}$$

$$\mathrm{div}\mathbf{D} = 0, \qquad \textbf{(1.2.4)}$$

$$\mathbf{D} = \varepsilon\varepsilon_0\mathbf{E}, \qquad \textbf{(1.2.5)}$$

$$\mathbf{B} = \mu\mu_0\mathbf{H}, \qquad \textbf{(1.2.6)}$$

где ε_0 и μ_0 - соответственно электрическая и магнитная постоянные, ε и μ - относительные соответственно диэлектрическая и магнитная проницаемости среды.

Подставим величину (1.2.6) в уравнение (1.2.2) и, предполагая, что величина μ не зависит от пространственных координат, возьмем ротор от левой и правой частей этого уравнения

$$\mathrm{rot}(\mathrm{rot}\,\mathbf{E}) = -\mu\mu_0 \frac{\partial}{\partial t}\mathrm{rot}\,\mathbf{H}\,. \qquad \textbf{(1.2.7)}$$

Используя уравнения (1.2.1) и (1.2.5), а также принимая во внимание, что

$$\mathrm{rot}(\mathrm{rot})\mathbf{E} = \nabla \mathrm{div}\mathbf{E} - \Delta\mathbf{E}, \; \mathrm{div}\mathbf{E} = \nabla \cdot \mathbf{E}, \qquad \textbf{(1.2.8)}$$

можно переписать уравнение (1.2.7) в виде

$$\Delta\mathbf{E} + \nabla\left(\mathbf{E} \cdot \frac{\nabla\varepsilon}{\varepsilon}\right) = \varepsilon\mu\varepsilon_0\mu_0 \frac{\partial^2\mathbf{E}}{\partial t^2}. \qquad \textbf{(1.2.9)}$$

Присутствующий в соотношениях (1.2.8), (1.2.9) оператор ∇ является вектором с компонентами $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$. Если величина ε постоянна в пространстве, градиент ε обращается в нуль, и уравнение (1.2.9) принимает вид волнового уравнения

$$\Delta\mathbf{E} = \varepsilon\mu\varepsilon_0\mu_0 \frac{\partial^2\mathbf{E}}{\partial t^2}. \qquad \textbf{(1.2.10)}$$

Применяя операцию rot к обеим частям уравнения (1.2.1), можно получить аналогичное уравнение для **B**

$$\Delta\mathbf{B} = \varepsilon\mu\varepsilon_0\mu_0 \frac{\partial^2\mathbf{B}}{\partial t^2}. \qquad \textbf{(1.2.11)}$$

Из (1.2.11) следует, что каждая декартова компонента $\Psi_\omega(x,y,z,t)$ вектора $\mathbf{E}$ или **B** удовлетворяют скалярному волновому уравнению

$$\Delta\Psi_\omega = \frac{1}{v^2}\frac{\partial^2\Psi_\omega}{\partial t^2}, \qquad \textbf{(1.2.12)}$$

где величина v имеет физический смысл скорости света в среде с параметрами ε и μ

$$v = (\varepsilon\mu\varepsilon_0\mu_0)^{-1/2}. \qquad \textbf{(1.2.13)}$$

Частным решением уравнения (1.2.12) может служить плоская волна произвольной формы, распространяющаяся в направлении **n**. Если в каждой точке пространства величина Ψ_ω меняется во времени по гармоническому закону, то плоская волна может быть описана выражением

$$\Psi_\omega = a \cdot \cos\left(\omega t - \frac{\omega}{v}\overline{\mathbf{nr}}\right), \qquad \textbf{(1.2.14)}$$

где ω - циклическая частота, *a*- амплитуда.

Вводя волновой вектор $\mathbf{k} = \omega\frac{\mathbf{n}}{\mathrm{v}}$, модуль которого равен $k=2\pi/\lambda$, выражению (1.2.14) можно придать вид

$$\Psi_\omega = a\cdot\cos(\omega t - \mathbf{kr}). \qquad \textbf{(1.2.15)}$$

В расчетах удобно пользоваться комплексным представлением плоской гармонической волны

$$\Psi_\omega = a\cdot\exp[-i(\omega t - \mathbf{kr})]. \qquad \textbf{(1.2.16)}$$

В комплексной форме могут быть представлены также расходящиеся и сходящиеся сферические волны, которые имеют соответственно вид

$$\Psi_\omega = \left(a/r\right)\exp\left[-i\left(\omega t - kr\right)\right], \qquad \textbf{(1.2.17)}$$

$$\Psi_\omega = \left(a/r\right)\exp\left[-i\left(\omega t + kr\right)\right]. \qquad \textbf{(1.2.18)}$$

Поскольку нас интересует преимущественно монохроматическое излучение, то есть излучение определенной частоты, мы будем в дальнейшем опускать экспоненциальный множитель

$$\exp\left[-i\left(\omega t\right)\right]. \qquad \textbf{(1.2.19)}$$

Если зависимость от времени представляется в форме (1.2.19), то дифференцирование по времени заменяется умножением на -iω и волновое уравнение (1.2.12) принимает вид

$$\Delta\Psi + k^2\Psi = 0, \qquad \textbf{(1.2.20)}$$

где под Ψ следует понимать комплексную амплитуду волны. Уравнение (1.2.20) называется приведенным волновым уравнением Гельмгольца. В нашем курсе это уравнение играет фундаментальную роль. В последующих разделах нам придется неоднократно к нему обращаться в процессе анализа особенностей распространения волновых пучков в различных средах и оптических системах.

До сих пор обсуждение касалось волнового уравнения (1.2.12), полученного как частный случай уравнения (1.2.9). Поскольку первое из них является существенно более простым и удобным, возникает вопрос о возможности его применения также для случая неоднородных сред. Для ответа на этот вопрос следует определить, в каких случаях можно пренебречь вторым членом в уравнении (1.2.9). Доминирующими членами в уравнении (1.2.9) являются первый член в левой части и член в правой части, порядок величин которых одинаковый. Возьмем отношение второго члена левой части уравнения (1.2.9) к члену, стоящему в правой части

$$R = \left[\nabla\left(\mathbf{E}\cdot\frac{\nabla\varepsilon}{\varepsilon}\right)\right] \Big/ \left(\varepsilon\mu\varepsilon_0\mu_0\frac{\partial^2 E}{\partial t^2}\right). \qquad \textbf{(1.2.21)}$$

Несложный анализ, выполненный в [8], показывает, что порядок R определяется соотношением

$$R \approx \frac{\lambda}{2\pi}\frac{\varepsilon_2-\varepsilon_1}{\varepsilon\cdot\Delta S}, \qquad \textbf{(1.2.22)}$$

где ε_1, ε_2 - значения относительной диэлектрической проницаемости в двух близких точках, разделенных расстоянием ΔS, ε - ее среднее значение. Выбирая $\Delta S=\lambda$, получаем

$$R = \frac{1}{2\pi}\frac{\varepsilon_2-\varepsilon_1}{\varepsilon}. \qquad \textbf{(1.2.23)}$$

Для того чтобы пренебречь вторым членом в левой части уравнения (1.2.9), необходимо потребовать выполнения условия $R<<1$. Это означает, что относительное изменение ε на расстоянии длины волны должно быть много меньше 1. Для большинства неоднородных оптических сред такое условие хорошо выполняется, что позволяет ограничиваться решением уравнения (1.2.10) вместо (1.2.9). Только на границе раздела двух областей с различными диэлектрическими проницаемостями, например, на границе между стеклянными линзами и воздухом, величина R становится большой. Однако даже в этих случаях следует решать уравнение (1.2.10) или (1.2.12), так как оно справедливо всюду, кроме границы раздела сред. На практике обычно решают волновое уравнение в различных однородных областях и сшивают эти решения посредством граничных условий.

1.2.2. Теория дифракции по Кирхгофу

В основе волновых представлений о распространении когерентого излучения лежит теория дифракции. Под дифракцией света обычно понимают отклонения от простых законов распространения света, описываемых геометрической оптикой. Дифракцию можно наблюдать, когда на пути распространения света находятся непрозрачные препятствия или когда свет проходит через отверстия в экранах. С дифракцией непосредственно связан физический механизм, обуславливающий перераспределение интенсивности в поперечном сечении пространственно-неоднородных лазерных пучков при их распространении в свободном пространстве. В большинстве случаев при описании дифракции можно не учитывать поляризации световой волны. Поэтому в основу теории дифракции мы положим скалярное уравнение Гельмгольца (1.2.20).

Пусть в пространстве распространяется монохроматическая волна

$$E(x,y,z,t) = \Psi(x,y,z)e^{-i\omega t}, \qquad \textbf{(1.2.24)}$$

амплитуда которой $\Psi(x,y,z)$ удовлетворяет уравнению Гельмгольца (1.2.20). Окружим точку наблюдения P произвольной поверхностью S (**рис. 1.2.1**.).

Определим возмущение в точке P в зависимости от возмущения на границе выделенной области. Воспользуемся для этого известной теоремой Грина

$$\int_V (\Psi\Delta\chi - \chi\Delta\Psi)dV = -\int_S \left(\Psi\frac{\partial\chi}{\partial n} - \chi\frac{\partial\Psi}{\partial n}\right)dS, \qquad \textbf{(1.2.25)}$$

где χ - некоторая вспомогательная величина, которая также как и Ψ имеет непрерывные частные производные первого и второго порядков внутри объема V, ограниченного поверхностью S, и на самой поверхности S. Потребуем также, чтобы функция χ удовлетворяла уравнению (1.2.20). Операция $\partial/\partial n$ в формуле (1.2.25) означает дифференцирование по внутренней нормали **n** к поверхности S. В качестве вспомогательной функции χ рассмотрим функцию $\chi = \frac{1}{r}\exp(i\mathbf{kr})$, где r - расстояние между произвольной точкой объема V и точкой P (радиус-вектор **r** будем считать направленным от точки P, как от начала координат). Функция χ представляет собой функцию изменения амплитуды поля точечного источника (т.е. так называемую функцию Грина свободного пространства). Для того чтобы эта функция удовлетворяла условиям теоремы Грина, нужно исключить из области V точку P, где функция χ обращается в бесконечность. С этой целью окружим точку P сферой бесконечно малого радиуса R и исключим ее из области V. Тогда формула (1.2.25) примет вид

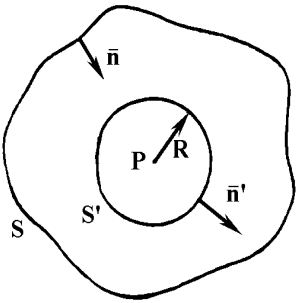


Рис. 1.2.1. *К выводу дифракционного интеграла Кирхгофа-Гюйгенса.*

$$\int_{V'} (\Psi\Delta\chi - \chi\Delta\Psi)dV = -\int_S \left(\Psi\frac{\partial\chi}{\partial n} - \chi\frac{\partial\Psi}{\partial n}\right)dS - \int_{S'} \left(\Psi\frac{\partial\chi}{\partial n} - \chi\frac{\partial\Psi}{\partial n}\right)dS'.$$

(1.2.26)

Здесь V' означает объем V без объема сферы, S' - площадь сферы. Поскольку функции Ψ и χ удовлетворяют уравнению Гельмгольца, то

объемный интеграл в выражении (1.2.26) равен нулю. Тогда из (1.2.26) следует, что

$$\int_S \left(\Psi\frac{\partial\chi}{\partial n}-\chi\frac{\partial\Psi}{\partial n}\right)dS=-\lim_{R\to 0}\int_{S'}\left(\Psi\frac{\partial}{\partial n}\left(\frac{e^{ikr}}{R}\right)-\frac{e^{ikr}}{R}\frac{\partial\Psi}{\partial n}\right)dS'=$$

$$=-\lim_{R\to 0}\int_{S'}\left(\Psi ik\frac{e^{ikr}}{R}-\Psi\frac{e^{ikr}}{R^2}-\frac{e^{ikr}}{R}\frac{\partial\Psi}{\partial n}\right)R^2 d\Omega=4\pi\Psi(P). \qquad \textbf{(1.2.27)}$$

При выводе соотношения (1.2.27) мы перешли от интегрирования по поверхности S′ к интегрированию по телесному углу Ω. Таким образом, возмущение в точке P будет равно

$$\Psi(P)=\frac{1}{4\pi}\int_S\left(\Psi\frac{\partial}{\partial n}\left(\frac{e^{ikr}}{r}\right)-\frac{e^{ikr}}{r}\frac{\partial\Psi}{\partial n}\right)dS. \qquad \textbf{(1.2.28)}$$

Выражение (1.2.28) известно как дифракционный интеграл Кирхгофа-Гюйгенса.

Дифракционный интеграл (1.2.28) широко используется при решении многих дифракционных задач. Следует, однако, иметь в виду, что возможны и другие математические подходы к анализу дифракции. Они, прежде всего, связаны с выбором другой вспомогательной функции χ. В частности, функция χ может быть выбрана так, чтобы она обращалась в ноль на поверхности *S*. Такой подход в какой-то степени упрощает задачу, так как в выражении (1.2.28) обращается в ноль член, содержащий $\partial\Psi/\partial n$. Однако цена этого упрощения состоит в усложнении функции χ.

Для большинства оптических задач выполняется условие

$$k >> \frac{1}{r}, \qquad \textbf{(1.2.29)}$$

тогда, пренебрегая производной от $1/\mathrm{r}$ по сравнению с производной от $\exp(ikr)$, формулу (1.2.28) можно записать в виде

$$\Psi(P)=\frac{1}{4\pi}\int_S\left(ik\frac{\partial r}{\partial n}\Psi-\frac{\partial\Psi}{\partial n}\right)\frac{e^{ikr}}{r}dS. \qquad \textbf{(1.2.30)}$$

Рассмотрим классическую для теории дифракции задачу о прохождении плоской волны через отверстие площадью *A* в бесконечном непрозрачном экране (**рис. 1.2.2**). Будем считать, что поверхность, по которой происходит интегрирование, включает экран и бесконечную полусферу радиуса $\widetilde{R}$, ограничивающую пространство справа от экрана. Обозначим координаты точки *P* как x′, y′, z′, угол наклона волнового

вектора k плоской волны и оси z через γ, а угол, задающий направление на точку P, через α. Из **рис. 1.2.2** следует, что

$$\frac{\partial}{\partial \mathrm{n}} = \frac{\partial}{\partial \mathrm{z}}, \qquad \textbf{(1.2.31)}$$

поэтому

$$\frac{\partial \Psi}{\partial n} = \frac{\partial \Psi}{\partial z} = ik_z \Psi = ik\Psi \cos\gamma \qquad \textbf{(1.2.32)}$$

и

$$\frac{\partial \mathbf{r}}{\partial \mathbf{n}} = \mathbf{grad} \cdot \mathbf{n} = -\cos\alpha. \qquad \textbf{(1.2.33)}$$

С помощью формул (1.2.32) и (1.2.33) интеграл (1.2.30) можно переписать в виде

$$\Psi(x', y', z') = \frac{ik}{4\pi} \int_A (\cos\gamma + \cos\alpha) \Psi(x, y, z) \frac{e^{ikr}}{r} dS. \qquad \textbf{(1.2.34)}$$

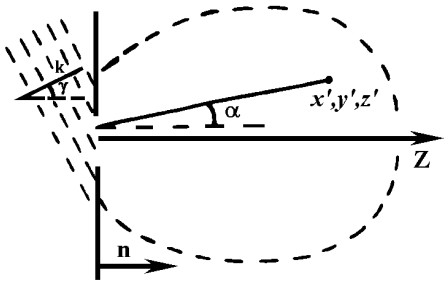


Рис. 1.2.2. *Дифракции плоской волны на отверстии в бесконечном экране.*

В таком виде этот интеграл известен как формула (интеграл) Френеля-Кирхгофа. В выражении (1.2.34) мы ограничились интегрированием лишь по площади отверстия, считая, что интеграл по бесконечной полусфере обращается в ноль. Последнее утверждение выглядит вполне обоснованным, если предположить, что падающая волна представляет собой очень длинный, но все же конченый цуг волн. Конечный же цуг волн не может достичь бесконечной полусферы за конечное время, тем самым интеграл по поверхности этой полусферы равен нулю.

Более точный математический анализ показывает, что интеграл по бесконечной полусфере стремится к нулю, если выполняется так называемое условие Зоммерфельда, согласно которому

$$\lim_{R\to\infty} = R\left(\frac{\partial \Psi}{\partial n} - jk\Psi\right) = 0.$$

Гораздо боле уязвимым предположением является использованное при получении формулы (1.2.34) второе предположение о равенстве нулю

функции Ψ и ее производной $\partial\Psi/\partial n$ на непрозрачном экране. Дело в том, что равенство нулю решения волнового уравнения и его производной на любом конечном интервале приводит к обращению его в ноль во всем объеме. Однако, несмотря на явный математический изъян, формула (1.2.34) приводит к результатам, хорошо согласующимся с данными экспериментов.

Во многих практических случаях, когда отверстие в экране мало и точка *P* располагается вблизи оси, можно считать, что

$$\cos\alpha \approx 1. \qquad \textbf{(1.2.35)}$$

Одновременно, если экран освещается волной, падающей на него перпендикулярно, можно положить, что

$$\cos\gamma = 1. \qquad \textbf{(1.2.36)}$$

Тогда формула (1.2.34) преобразуется в выражение

$$\Psi(x', y', z') = -\frac{i}{\lambda}\int_A \Psi(x, y, z)\frac{e^{ikr}}{r}dS, \qquad \textbf{(1.2.37)}$$

которое является математическим обобщением принципа Гюйгенса-Френеля. Из него видно, что недостаточно просто предполагать, как это делал в свое время Гюйгенс, что падающая волна выполняет роль источника сферических волн с амплитудами, пропорциональными амплитуде падающей волны в каждой точке. Необходимо потребовать, чтобы фаза вторичного источника отставала от фазы падающей волны на $\pi/2$ (из-за наличия в правой части (1.2.37) множителя -*i*).

Поскольку в большинстве практических случаев выполняются соотношения

$$|x' - x| << |z' - z| \ \textbf{и}\ |y' - y| << |z' - z|, \qquad \textbf{(1.2.38)}$$

можно построить выражение для величины *r*, ограничиваясь первыми двумя членами ее разложения в ряд Тейлора

$$r = z' - z + \frac{1}{2}\frac{(x' - x)^2 + (y' - y)^2}{z' - z} + \ldots \qquad \textbf{(1.2.39)}$$

Используя это выражение, получим следующее приближение для формулы Френеля-Кирхгофа (1.2.34):

$$\Psi(x', y', z') = -\frac{i}{4\lambda}e^{ikz'}\int_A (\cos\gamma + \cos\alpha)\frac{\Psi}{z' - z}\times$$
$$\times\exp\left[-ik\left(z - \frac{(x' - x)^2 + (y' - y)^2}{2(z' - z)}\right)\right]dS \qquad \textbf{(1.2.40)}$$

Принято говорить о двух случаях применения интеграла (1.2.40): дифракции Френеля и дифракции Фраунгофера. Дифракция Френеля имеет место, когда поле рассчитывается на небольшом расстоянии от отверстия и член $x^2 + y^2$, появляющийся в показателе степени экспоненты, следует принимать во внимание. Дифракция же Фраунгофера наблюдается вдали от отверстия, когда этот член пренебрежительно мал.

Вторичные сферические волны, излучаемые каждой точкой в плоскости отверстия, являются в определенном смысле абстракцией и вводятся в приведенном выше подходе к решению дифракционных задач, главным образом, для удобства описания. Более физический подход развит в работах Зоммерфельда. Зоммерфельд рассматривал высказанную еще в 1802 г. Томасом Юнгом идею, заключающуюся в следующем: наблюдаемое поле является суперпозицией падающей волны, прошедшей через отверстие без искажения, и дифрагированной волны, источником которой служит край отверстия. Однако на этом подходе мы подробно останавливаться не будем.

1.2.3. Дифракция Френеля

Анализ дифракции Френеля в общем случае или применительно к прохождению света через сколь-нибудь сложные неоднородные структуры представляет собой непростую задачу. Поэтому мы ограничимся рассмотрением дифракции Френеля на квадратном отверстии со стороной *l*. Представим выражение (1.2.40) в виде произведения двух интегралов, считая углы γ и α пренебрежимо малыми:

$$\Psi(x', y') = \frac{\exp(ikz)}{i\lambda z}\Phi(x')\Phi(y'), \qquad \textbf{(1.2.41)}$$

где

$$\Phi(x') = \int_{-a/2}^{a/2} \exp\left[i\frac{k}{2z}(x - x')^2 \right] dx,$$

$$\Phi(y') = \int_{-a/2}^{a/2} \exp\left[i\frac{k}{2z}(y - y')^2 \right] dy.$$

Интегралы существенно упрощаются после замены переменных

$$\xi = \sqrt{\frac{k}{\pi(z'-z)}}(x - x'), \quad \eta = \sqrt{\frac{k}{\pi(z'-z)}}(y - y'),$$

и переходят в следующие интегралы

$$\Phi(x') = \sqrt{\frac{\pi(z'-z)}{k}} \int_{\xi_1}^{\xi_2} \exp\left(i\frac{\pi}{2}\xi^2\right) d\xi,$$

$$\Phi(y') = \sqrt{\frac{\pi(z'-z)}{k}} \int_{\eta_1}^{\eta_2} \exp\left(i\frac{\pi}{2}\eta^2\right) d\eta,$$

где пределы интегрирования определяются соотношениями

$$\xi_1 = -\sqrt{\frac{k}{\pi(z'-z)}}\left(\frac{a}{2}+x'\right), \quad \xi_2 = \sqrt{\frac{k}{\pi(z'-z)}}\left(\frac{a}{2}-x'\right),$$

$$\eta_1 = -\sqrt{\frac{k}{\pi(z'-z)}}\left(\frac{a}{2}+y'\right), \quad \eta_2 = \sqrt{\frac{k}{\pi(z'-z)}}\left(\frac{a}{2}-y'\right).$$

Интегралы *Φ(x')* и *Φ(y')* можно выразить через табулированные функции, известные под названием интегралов Френеля. Последние определяются выражениями

$$C(\alpha) = \int_0^{\alpha} \cos\frac{\pi t^2}{2} dt, \quad S(\alpha) = \int_0^{\alpha} \sin\frac{\pi t^2}{2} dt. \qquad \textbf{(1.2.43)}$$

Учитывая, что

$$\int_{\xi_1}^{\xi_2} \exp\left(i\frac{\pi}{2}\xi^2\right) d\xi = \int_0^{\xi_2} \exp\left(i\frac{\pi}{2}\xi^2\right) d\xi - \int_0^{\xi_1} \exp\left(i\frac{\pi}{2}\xi^2\right) d\xi, \qquad \textbf{(1.2.44)}$$

можно записать

$$\Phi(x') = \sqrt{\frac{\pi(z'-z)}{k}} \left\{\left[C(\xi_2) - C(\xi_1)\right] + i\left[S(\xi_2) - S(\xi_1)\right]\right\},$$

$$\Phi(y') = \sqrt{\frac{\pi(z'-z)}{k}} \left\{\left[C(\eta_2) - C(\eta_1)\right] + i\left[S(\eta_2) - S(\eta_1)\right]\right\}. \qquad \textbf{(1.2.45)}$$

Наконец, подставляя (1.2.45) в (1.2.41), получаем распределение комплексного поля

$$\psi(x', y') = \frac{\exp(ikz)}{i\lambda z} \left\{\left[C(\xi_2) - C(\xi_1)\right] + i\left[S(\xi_2) - S(\xi_1)\right]\right\} \times$$
$$\times \left\{\left[C(\eta_2) - C(\eta_1)\right] + i\left[S(\eta_2) - S(\eta_1)\right]\right\}. \qquad \textbf{(1.2.46)}$$

и соответствующее распределение интенсивности

$$I(x',y') = \frac{1}{4}\left\{\left[C(\xi_2)-C(\xi_1)\right]^2 + i\left[S(\xi_2)-S(\xi_1)\right]^2\right\}\times$$
$$\times\left\{\left[C(\eta_2)-C(\eta_1)\right]^2 + i\left[S(\eta_2)-S(\eta_1)\right]^2\right\}. \qquad \textbf{(1.2.47)}$$

Для интерпретации этих выражений удобно воспользоваться графическим построением, которое называется спиралью Корню; спираль Корню (**рис. 1.2.3**) представляет собой совместный график зависимости $C(\alpha)$ и $S(\alpha)$ от параметра α.

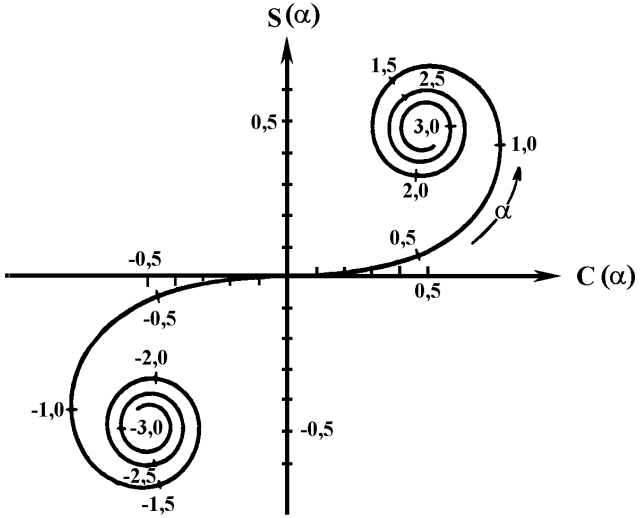


Рис. 1.2.3. *Спираль Корню.*

Заметим, что величину $C(\alpha)+iS(\alpha)$ можно считать комплексным фазором, соединяющим начало координат с точкой α на спирали. Следовательно, величина $\left\{\left[C(\xi_2)-C(\xi_1)\right]+i\left[S(\xi_2)-S(\xi_1)\right]\right\}$ представляет собой фазор, определяемый участком спирали между точками ξ_1 и ξ_2. Используя подобный графический метод, можно вычислить значение выражений (1.2.46) и (1.2.47) в каждой точке дифракционной картины.

1.2.4. Дифракция Фраунгофера. Элементы фурье-оптики

Скалярную теорию дифракции можно облечь в иную форму, если применить двумерный анализ Фурье. В дальнейшем мы воспользуемся тем, что при фурье-анализе распределения светового поля в любой плоскости

пространственные составляющие его фурье-образа можно отождествить с плоскими волнами, идущими в разных направлениях. Складывая амплитуды этих волн и учитывая их фазовые набеги, можно вычислить амплитуду поля в любой интересующей нас точке пространства. Раздел физической оптики, в котором для описания преобразований структуры светового поля используется двумерный анализ Фурье, называется фурье-оптикой. Особенно удобно использовать аппарат фурье-оптики для характеристики дифракции Фраунгофера.

Пусть в некоторой плоскости $\{x,y,0\}$ задана комплексная амплитуда $\Psi(x,y,0)$ световой волны, распространяющейся в положительном направлении оси z. Рассмотрим трансформацию поля волны при ее распространении. Выясним сначала, какой физический смысл имеет фурье-образ функции $\Psi(x,y,0)$.

В плоскости $\{x,y,0\}$ двумерный фурье-образ функции Ψ имеет вид

$$A(f_x,f_y)=\int_{-\infty}^{\infty}\int\Psi(x,y,0)\exp\left[-i2\pi(f_x x+f_y y)\right]dxdy. \quad \textbf{(1.2.48)}$$

С помощью фурье-образа $A(f_x,f_y)$ функцию ψ можно представить в виде совокупности простых экспоненциальных функций

$$\Psi(x,y,0)=\int_{-\infty}^{\infty}\int A(f_x,f_y)\exp\left[i2\pi(f_x x+f_y y)\right]df_x df_y. \quad \textbf{(1.2.49)}$$

Из формулы (1.2.16) следует, что для комплексной амплитуды плоской волны справедливо выражение

$$\Psi(x,y,z)=\alpha\exp\left[i\frac{2\pi}{\lambda}(\alpha x+\beta y+\gamma z)\right], \quad \textbf{(1.2.50)}$$

где α, β, γ - направляющие косинусы нормали к фронту плоской волны, причем $\gamma=\sqrt{1-\alpha^2-\beta^2}$. Тем самым комплексную экспоненциальную функцию $\exp\left[i2\pi(f_x x+f_y y)\right]$, входящую в выражение (1.2.42), можно рассматривать как плоскую волну с направляющими косинусами

$$\alpha=\lambda f_x,\beta=\lambda f_y,\gamma=\sqrt{1-(\lambda f_x)^2-(\lambda f_y)^2} \quad \textbf{(1.2.44)}$$

и комплексной амплитудой, равной $A(f_x,f_y)df_x df_y$, где $f_x=\frac{\alpha}{\lambda}$, $f_y=\frac{\beta}{\lambda}$. Таким образом, можно считать, что выражение (1.2.42) задает угловой спектр возмущения $\Psi(x,y,0)$ в плоскость, параллельную первоначальной, но смещенную от нее на расстояние z, ее угловые составляющие сохранят

свои амплитуды. Все изменения сведутся лишь к изменению фаз угловых составляющих спектра, поскольку плоские волны, распространяясь под разными углами, проходят разные расстояния. Иная ситуация будет иметь место, если на пути распространения волны будут находиться какие-либо препятствия.

Предположим, что на пути распространения волны с угловым спектром $A_i\left(f_x, f_y\right)$ находится экран с амплитудным коэффициентом пропускания *t(x,y)*. В плоскости непосредственно за экраном распределение поля можно записать в виде

$$\Psi_t\left(x, y, 0\right) = \Psi_i\left(x, y, 0\right) \cdot t\left(x, y\right), \quad (1.2.45)$$

где Ψ_i - поле падающей волны. По теореме свертки - важнейшей теореме анализа Фурье, - для углового спектра пропущенной экраном волны $A_t\left(f_x, f_y\right)$ будет справедливо выражение

$$A_t\left(f_x, f_y\right) = \int_{-\infty}^{\infty}\int A_i\left(f_x', f_y'\right) T\left(f_x - f_x', f_y - f_y'\right) df_x' df_y', \quad (1.2.46)$$

где *T* - фурье-образ функции t(x,y).

Если на экран перпендикулярно падает плоская волна единичной амплитуды, то ее угловой спектр будет определяться δ-функцией

$$A_i\left(f_x', f_y'\right) = \delta\left(f_x', f_y'\right). \quad (1.2.47)$$

В этом случае (1.2.46) упрощается

$$A_t = T\left(f_x, f_y\right). \quad (1.2.48)$$

Таким образом, для рассмотренного частного случая угловой спектр дифрагированной волны представляет собой фурье-образ функции пропускания *t(x,y)*. Как правило, помещение на пути распространения волны какой-либо ограничивающей апертуры приводит к существенному уширению углового спектра, причем это уширение тем больше, чем меньше размер апертуры.

Применим теперь анализ Фурье для описания дифракции Фраунгофера на отверстии. Разложим квадратичные члены в экспоненте, стоящей под интегралом в выражении (1.2.40), ограничиваясь случаем, когда $\cos\alpha \approx \cos\gamma \approx 1$:

$$\Psi\left(x', y', z'\right) = \frac{i\exp(ikl)}{\lambda l}\exp\left[i\frac{k}{2l}\left(x'^2 + y'^2\right)\right]\times$$

$$\times \int_{-\infty}^{\infty}\int \left\{ \Psi(x,y) \exp\left[i\frac{k}{2l}\left(x^2+y^2\right)\right]\right\} \exp\left[-i\frac{2\pi}{\lambda l}\left(x'x+y'y\right)\right] dxdy, \quad \textbf{(1.2.49)}$$

$$l = z' - z.$$

Учтем также, что при дифракции Фраунгофера

$$l >> \frac{k\left(x^2+y^2\right)_{\max}}{2}. \quad \textbf{(1.2.50)}$$

Тогда можно считать, что квадратичная фазовая функция $\exp\left[i\frac{k}{2l}\left(x^2+y^2\right)\right]$ приблизительно равна единице по всему отверстию, и в точке с координатами x', y' поле равно

$$\Psi(x',y') = \frac{i}{\lambda l}\exp(ikl)\exp\left[i\frac{k}{2l}\left(x'^2+y'^2\right)\right] \times \int_{-\infty}^{\infty}\int \Psi(x,y)\exp\left[-i\frac{2\pi}{\lambda l}\left(x'x+y'y\right)\right] dxdy. \quad \textbf{(1.2.51)}$$

Поскольку последний интеграл представляет собой фактически фурье-образ функции $\Psi(x,y)$ с пространственными частотами $f_x = \frac{x'}{\lambda l}, f_y = \frac{y'}{\lambda l}$, то, обозначая фурье-образ как $F\{\Psi(x,y)\}$, имеем

$$\Psi(x',y') = \frac{i}{\lambda l}\exp(ikl)\exp\left[i\frac{k}{2l}\left(x'^2+y'^2\right)\right] F\{\Psi(x,y)\}. \quad \textbf{(1.2.52)}$$

Отсюда видно, что расчет распределения поля дифрагированной волны фактически сводится к нахождению фурье-образа поля сразу за экраном. Если экран освещается плоской когерентной волной с единичной амплитудой, то $F\{\Psi(x,y)\} = T$. Ниже приводятся фурье-образы функций пропускания для наиболее важных в практическом отношении случаев.

1. Прямоугольное отверстие

$$t(x,y) = \mathrm{rect}\left(x/l_x\right)\mathrm{rect}\left(y/l_y\right), \quad \textbf{(1.2.53)}$$

l_x, l_y - размеры отверстий соответственно в направлении x и y,

$$T = l_x l_y \,\mathrm{sinc}\left(l_x f_x\right)\mathrm{sinc}\left(l_y f_y\right), \quad \textbf{(1.2.54)}$$

где $\sin c\theta = \sin c(\pi\theta)/\pi\theta$.

2. Круглое отверстие

$$t(x,y) = circ\left(\frac{\sqrt{x^2+y^2}}{D/2}\right), \qquad \textbf{(1.2.55)}$$

D -диаметр отверстия,

$$T = \left(\frac{D}{2}\right)^2 \frac{J_1\left(\pi D\sqrt{x'^2+y'^2}/\lambda l\right)}{D\sqrt{x'^2+y'^2}/\lambda l}, \qquad \textbf{(1.2.56)}$$

где J_1 - функция Бесселя первого порядка.

3. Синусоидальная амплитудная решетка

$$t(x,y) = \left[\frac{1}{2} + \frac{m}{2}\cos(2\pi f_0 x)\right]\mathrm{rect}\left(\frac{x}{L}\right)\mathrm{rect}\left(\frac{y}{L}\right); \qquad \textbf{(1.2.57)}$$

L - размер квадратной решетки, f_0 - ее частота, m - разность между максимальным и минимальным пропусканием.

$$T = \frac{L^2}{2}\mathrm{sinc}(Lf_y)\left\{\mathrm{sinc}(Lf_x) + \frac{m}{2}\mathrm{sinc}[L(f_x+f_0)] + \frac{m}{2}\mathrm{sinc}[L(f_x-f_0)]\right\}. \qquad \textbf{(1.2.58)}$$

4. Дифракционная решетка - непрозрачный экран размером L, имеющий N щелей шириной a (щели располагаются строго периодически в направлении y, на расстоянии b одна за другой, так что период решетки составляет $d=a+b$)

$$T = \mathrm{sinc}(af_x)\frac{\mathrm{sinc}(Lf_x)}{\mathrm{sinc}(df_x)}. \qquad \textbf{(1.2.59)}$$

5. "Мягкая" диафрагма с гауссовым профилем пропускания

$$t(x,y) = \exp\left[-\pi\left(x^2+y^2\right)\right], \qquad \textbf{(1.2.60)}$$

$$T = \exp\left[-\pi\left(f_x^2+f_y^2\right)\right]. \qquad \textbf{(1.2.61)}$$

На последнем примере следует остановиться особо. Из формул (1.2.60) и (1.2.61) следует, что, если световая волна имеет в поперечном сечении гауссово распределение амплитуды (такую волну можно получить, например, пропуская плоскую однородную волну через диафрагму с профилем (1.2.60), то ее фурье-образ также будет характеризоваться функцией Гаусса. Благодаря этому обстоятельству, "гауссовый" световой пучок, распространяясь в свободном пространстве, будет сохранять неизменной форму распределения амплитуды поля. Более подробно свойства гауссова пучка будут рассмотрены в следующей главе.

1.2.5. Принцип Бабине. Эффект Талбота

Рассмотрим в силу практической значимости более подробно два случая дифракции на плоских экранах. Пусть задан некоторый экран. Заменой отверстий на непроницаемые участки и наоборот можно получить так называемый дополнительный экран. Если Ψ_1 и Ψ_2 - дифрагированные поля на этих двух экранах, то имеет место следующее соотношение (принцип Бабине)

$$\Psi_1(\mathbf{r}') + \Psi_2(\mathbf{r}') = \Psi(\mathbf{r}'), \qquad \textbf{(1.2.62)}$$

где $\Psi(\mathbf{r}')$ - поле в отсутствие экрана. Соотношение (1.2.62) непосредственно следует из (1.2.40), если интегрирование в этом соотношении выполнить по всей плоскости. Остается лишь предположить, что поля $\Psi_1(\mathbf{r})$ и $\Psi_2(\mathbf{r})$ на отверстиях первого и второго экрана совпадают с полем $\Psi(\mathbf{r})$, которое имеет место в отсутствие экрана. Вообще говоря, принцип Бабине выполняется лишь приближенно, так как $\Psi_1(\mathbf{r})$ и $\Psi_2(\mathbf{r})$ не равны $\Psi(\mathbf{r})$, но нарушение (1.2.62) существенно лишь вблизи границ диафрагм.

Второй случай относится к дифракционным полям, имеющим вид периодических функций. Он имеет интересные приложения для теории решеток и теории модульных лазерных систем.

Пусть цилиндрическое поле $\Psi(x,z)$, фаза которого не зависит от координаты y, в плоскости z=0 записывается в виде

$$\Psi(x,0) = f(x) = \sum_{n=-N}^{+N} f_n \exp\left(in2\pi\frac{x}{d}\right), \qquad \textbf{(1.2.63)}$$

где $f(x) = f(x+d)$ - периодическая функция координаты *x*, содержащая столько гармоник *N*, что $\frac{N\lambda}{d} << 1$. Тогда для ближней зоны дифракции можно использовать дифракционную формулу Френеля (1.2.40), и мы имеем

$$\Psi(x',z) = \frac{\exp(i\pi/4 - ikz)}{\sqrt{\lambda z}} \times$$

$$\times \sum_{n=-N}^{+N} f_n \int_{-\infty}^{+\infty} \exp\left[-ik\frac{(x-x')^2}{2z} + in2\pi\frac{x}{d}\right] dx = \qquad \textbf{(1.2.64)}$$

$$= \exp(-ikz) \sum_{n=-N}^{+N} f_n \exp\left[in2\pi\frac{x'}{d} + i\left(\frac{n2\pi}{d}\right)^2 \frac{z}{2k}\right].$$

Следует заметить, что в частном случае $z = z_q = \frac{qd^2}{\lambda}$ (q=1, 2, ...) имеет место соотношение $\exp\left(i\left(\frac{n2\pi}{d}\right)^2 \frac{z_q}{2k}\right) = \exp\left(in^2\pi q\right) = \exp\left(in\pi q\right)$, так что

$$\Psi\left(x', z_q\right) = \exp\left(-ikz_q\right) f\left(x + \frac{1}{2}qd\right). \qquad \textbf{(1.2.65)}$$

Отсюда следует, что во всех плоскостях z_q распределение интенсивности поля одинаково. Это свойство называют эффектом Талбота, или эффектом самовоспроизведения. Оно было впервые замечено Талботом в 1836 г. и имеет весьма важные применения в фурье-спектроскопии, интерферометрии и оптике лазеров.

1.2.6. О возможности обобщения метода Кирхгофа на случай векторных полей

Математическая не строгость теории дифракции приводит к тому, что она не может быть использована для расчета характеристик поля в непосредственной близости от отверстий в непрозрачных экранах. Кроме того, приемлемая точность расчетов может быть обеспечена лишь в тех случаях, когда размеры отверстий немного превосходят длину волны. Указанные ограничения сужают диапазон возможных приложений скалярной теории. Например, она не может быть использована для расчета характеристик дифракционных решеток высокого разрешения.

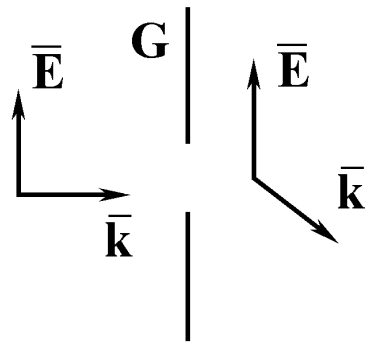


Рис. 1.2.4. *Появление продольных компонентов поля при дифракции поляризованной волны на отверстии в непрозрачном экране G.*

Применение скалярных методов для описания дифракции линейно поляризованной волны приводит к еще одному серьезному противоречию. оно состоит в том, что скалярные методы приводят к выводу о наличии продольных компонент поля за экраном (**рис. 1.2.4**.). Тем самым существует необходимость построения теории дифракции на основе более последовательного векторного

подхода.

Формальное обобщение метода Кирхгофа на случай векторных полей можно осуществить, записав для каждого компонента вектора $\mathbf{E}$ интеграл Кирхгофа (1.2.25), а затем, сложив их векторно. В результате этой процедуры получается следующее выражение для $\mathbf{E}$:

$$\mathbf{E} = \frac{1}{4\pi}\iint_S \left(\mathbf{E}\frac{\partial \chi}{\partial n} - \chi \frac{\partial \mathbf{E}}{\partial n} \right) ds = \frac{1}{4\pi}\iint_S \left[\mathbf{E}(\mathbf{n}\nabla\chi) - \chi(\mathbf{n}\nabla)\mathbf{E} \right] ds, \quad \textbf{(1.2.66)}$$

Учитывая, что $rot\mathbf{E} = i\frac{\omega}{c}\mu\mathbf{H}$ и $div\mathbf{E} = 0$, путем математических преобразований, получаем

$$\mathbf{E}(x,y,z) = \frac{1}{4\pi}\iint_S \left\{ i\frac{\omega}{c}\mu[\mathbf{nH}]\chi + [[\mathbf{nE}]\,grad\chi] + (\mathbf{nE})\,grad\chi \right\} ds, \quad \textbf{(1.2.67)}$$

где μ - магнитная проницаемость среды.

Аналогичное выражение можно получить и для вектора $\mathbf{H}$.

Определяемый формулой (1.2.67) вектор $\mathbf{E}$ не будет удовлетворять уравнениям Максвелла. Причина состоит в том, что тангенциальные компоненты векторов $\mathbf{E}$ и $\mathbf{H}$ терпят разрыв при переходе через границу контура отверстия. Для того, чтобы удовлетворить условиям непрерывности, необходимо ввести некоторое распределение зарядов и токов на контуре отверстия в экране.

Электродинамический анализ показывает, что самосогласованное поле дифрагированной волны с учетом дополнительных источников поля, обусловленных указанными зарядами, получается добавлением к поверхностному интегралу (1.2.67) интеграла по контуру отверстия

$$\mathbf{E}(x,y,z) = \frac{1}{4\pi}\iint_S \left\{ i\frac{\omega}{c}\mu[\mathbf{nH}]\chi + [[\mathbf{nE}]\,\mathrm{grad}\,\chi] + (\mathbf{nE})\,grad\chi \right\} ds +$$

$$+ \frac{c}{4\pi i\omega\varepsilon}\oint_C (\mathbf{Ht})\,\mathrm{grad}\,\chi\, dl, \quad \textbf{(1.2.68)}$$

где τ - единичный вектор, касательный к элементу контура отверстия dl. Обход контура диафрагмы осуществляется против часовой стрелки, если смотреть из точки, в которой определяется поле. С увеличением расстояния от точки наблюдения до отверстия вклад интеграла по контуру в величину поля снижается и на расстояниях многих длин волн им вообще можно пренебречь.

1.3. Методы геометрической оптики [1,9,10,13,14]

Возможны две точки зрения на место геометрической оптики в системе современных оптических представлений. Согласно первой из них геометрическая оптика рассматривается как самостоятельный раздел оптики, основанный на определенной системе постулатов. К наиболее важным из них относятся законы прямолинейного распространения света, законы его отражения и преломления. В такой постановке геометрическая оптика является основой вычислительной оптики [11], на базе которой осуществляются расчеты разнообразных оптических элементов и систем. Согласно второй точки зрения основные выражения и соотношения аппарата геометрической оптики являются по своей сути приближенными решениями волновых уравнений, во многих случаях облегчающих их анализ. Исходя из целевой установки данной книги мы будем придерживаться второй точки зрения. При этом сосредоточимся на вопросах распространения света в неоднородной среде, показатель преломления которой плавно меняется в пространстве. Световое поле представляется в форме локально плоской волны. В приближении геометрической оптики амплитуда этой волны не зависит от частоты, а частота, которая считается большой величиной, входит только в фазовый множитель.

1.3.1. Геометрооптическое приближение

В однородной (n=const) среде простейшее решение волнового уравнения - плоская волна

$$\Psi = Ae^{-iknz}. \qquad \textbf{(1.3.1)}$$

Фаза ее постоянна на плоскостях z=const, т.е. фазовые фронты плоские. Нормали к фронтам параллельны.

Пусть показатель преломления среды есть функция координат

$$\boldsymbol{n=n(}\mathrm{r}\boldsymbol{)},$$

тогда амплитуда A становится функцией координат, а волновые фронты перестают быть плоскими. Нормали к ним не параллельны, но близки к параллельным. Будем считать, что все характерные масштабы изменения и амплитуды поля, и показателя преломления среды велики по сравнению с длиной волны $\lambda(\lambda=2\pi/k)$. Если l - наименьший из этих масштабов, то предполагается, что выполнено неравенство

$$\mu = \frac{1}{knl} << 1. \qquad \textbf{(1.3.2)}$$

Это неравенство - не самое сильное условие применимости лучевой или геометрической оптики, законы которой мы ниже сформулируем. Представим поле световой волны в неоднородной среде в виде почти плоской волны

$$\Psi = A(k,\mathbf{r})e^{-ikS(\mathrm{r})}. \qquad \textbf{(1.3.3)}$$

Здесь $A(k,\mathbf{r})$ - амплитуда волны; $kS(\mathbf{r})$ - фаза; величина $S(\mathbf{r})$ называется эйконалом. Термин "почти плоская волна" оправдан тем, что в области порядка $2\pi/kn$ поле имеет вид.(1.3.1).

Амплитуду A и эйконал S будем искать из требования, чтобы решение, записанное в форме (1.3.3), удовлетворяло волновому уравнению, и при этом используем условие, что k велико. Это можно сделать несколькими способами. Простейший из них состоит в том, чтобы искать A в виде лучевого разложения по обратным степеням k:

$$A(k,\mathbf{r}) = \sum_m \frac{A_m(\mathbf{r})}{(ik)^m}. \qquad \textbf{(1.3.4).}$$

Фактически это разложение проводится по возрастающим степеням безразмерного малого параметра μ (1.3.2).

Подставляя (1.3.3), (1.3.4) в волновое уравнение

$$\Delta\Psi + k^2 n^2 \Psi = 0 \qquad \textbf{(1.3.5)}$$

и приравнивая нулю коэффициенты при одинаковых степенях k, получим в нулевом приближении уравнение эйконала

$$(\nabla S)^2 = n^2, \qquad \textbf{(1.3.6)}$$

а в следующих приближениях- систему рекуррентных уравнений для амплитуд

$$2\nabla A_0 \nabla S + A_0 \Delta S = 0, \qquad \textbf{(1.3.7)}$$

$$2\nabla A_1 \nabla S + A_1 \Delta S = -\Delta A_0, \qquad \textbf{(1.3.8)}$$

которые называются уравнениями переноса.

Приближение, при котором в (1.3.4) сохраняется только нулевой член, называется геометрооптческим приближением. Поля в этом приближении, т.е. поля геометрической оптики

$$u = A_0(\mathbf{r})e^{-ikS(\mathbf{r})}, \qquad \textbf{(1.3.9)}$$

содержат частоту только множителем в фазе. В отличие от (1.3.3), в (1.3.9) A_0 уже не есть функция частоты.

1.3.2. Лучи и фронты

Решение задачи по определению поля в неоднородной среде следует начинать с нахождения эйконала S. Зная эту функцию, можно построить

волновые фронты: они задаются уравнением S=const, а затем и лучи - линии, перпендикулярные волновому фронту.

Следует заметить, что вообще эйконал в большей степени определяет световое поле, чем амплитуда A_0 (**r**). Это объясняется тем, что перед $S(k,\mathbf{r})$ стоит большой множитель k. Поэтому все изменения u при малом изменении координат определяются главным образом изменением S, а не A_0.

Уравнение (1.3.6) решается в наиболее общем виде с помощью метода характеристик. Этот метод сводит уравнение в частных производных к системе обыкновенных дифференциальных уравнений. Обозначим $\nabla S \equiv \mathbf{p}$, введем параметр τ вдоль направления **p**, связанный с длиной дуги σ на луче условием $d\tau = d\sigma/n$. Вектор **r** определяет точку на луче, а вектор $d\mathbf{r}/d\tau$ - касательный лучу. Уравнение (1.3.6) выглядит в новых обозначениях как $|\mathbf{p}|=n$. Можно показать, что оно эквивалентно следующей системе уравнений:

$$а)\quad \frac{d\mathbf{r}}{d\tau} = \mathbf{p}, \quad б)\quad \frac{d\mathbf{p}}{d\tau} = \frac{1}{2}\nabla n^2, \qquad \textbf{(1.3.10)}$$

$$\frac{dS}{d\tau} = n^2. \qquad \textbf{(1.3.11)}$$

Уравнения (1.3.10) определяют геометрию лучей

$$\mathbf{r} = \mathbf{r}(\tau), \quad \mathbf{p} = \mathbf{p}(\tau), \qquad \textbf{(1.3.12)}$$

т.е. координаты **r** и направления **p** луча в точке с параметром τ (этот параметр пропорционален времени прохождения волны вдоль луча). Очевидно, что касательный к лучу вектор (1.3.10,а) параллелен ∇S, т.е. перпендикулярен к волновой поверхности. Таким образом, луч есть нормаль к поверхности равной фазы.

Вектор **p**, определяющий направление луча, изменяется вдоль луча, согласно (1.3.11,б) в направлении градиента показателя преломления. Иными словами, преломление или рефракция криволинейного луча в неоднородной среде происходит в область возрастания n.

В частном случае однородной среды $\nabla n \equiv 0$, так что **p**=const, и лучи являются прямыми линиями.

Для эйконала из (1.3.11) получаем

$$S = S(\tau_0) = +\int_{\tau_0}^{\tau} n^2 d\tau. \qquad \textbf{(1.3.13)}$$

Здесь $S(\tau_0)$ - значение эйконала при $\tau = \tau_0$; интегрирование ведется вдоль геометрооптического луча. Заметим, что в геометрической оптике

физическое значение имеет лишь разность эйконалов $S - S(\tau)$, а не величина S.

Для однородной среды $S = S(\tau_0) + n^2(\tau - \tau_0)$.

Выбрав направление оси z вдоль луча, умножив затем эйконал - его часто называют "оптический путь" - на частоту k, получим привычное выражение для фазы плоской волны $kn(z-z_0)$, так как, очевидно, в данном случае $n\tau = z$.

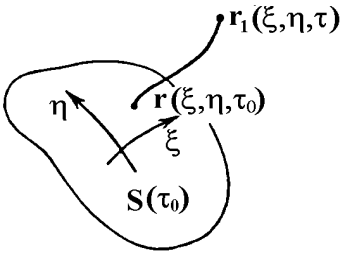


Рис. 1.3.1. *Лучевые координаты.*

Для построения лучей и фронтов удобно пользоваться системой лучевых координат (**рис. 1.3.1**).

Такими координатами являются две координаты ξ, η на поверхности любого, принятого за начальный, волнового фронта $S(\tau_0)$, характеризующие данный луч и постоянные вдоль луча, и длина дуги σ, отсчитываемая вдоль луча, либо вместо σ - введенный выше параметр τ (время).

1.3.3. Принцип Ферма

Из уравнения эйконала вытекает важное физическое положение, известное, как принцип Ферма, согласно которому оптический луч всегда выбирает траекторию с минимальной длиной оптического пути. Только в очень редких случаях условие минимума заменяется условием максимума. Таким образом, лучевые траектории являются экстремалями функционала Ферма

$$\int n d\sigma \qquad \textbf{(1.3.14)}$$

Принцип Ферма определяет лучевую структуру поля не только в плавно неоднородной среде. Луч, соединяющий две точки r_0 и r_1, выделяется из всех кривых, проходящих через эти точки, тем, что эйконал $S(r_1) - S(r_0)$ (1.3.13) экстремален. Из принципа экстремальности может быть выведен закон зеркального отражения

$$\varphi_1 = \varphi_3 \qquad \textbf{(1.3.15)}$$

и закон преломления

$$\sin\varphi_1 / \sin\varphi_2 = n_2 / n_1 \qquad \textbf{(1.3.16)}$$

на резкой границе между двумя плавно-неоднородными средами.

Здесь n_1 - показатель преломления среды, из которой на границу падает луч. Он преобразуется в два луча - преломленный и отраженный в среду n_2. Углы φ_1, φ_2, φ_3 - соответственно углы с нормалью падающего, преломленного и отраженного лучей. Все три луча и нормаль к поверхности расположены в плоскости падения. Хотя эти законы получены для случая падения плоской волны на плоскую границу раздела однородных сред, они выполняются и для неплоской границы между плавно неоднородными средами, если поле сохраняет лучевую структуру (1.3.9).

1.3.4. Лучевые трубки

Перейдем от лучевой структуры поля, т.е. системы волновых фронтов S=const и лучей ∇S, к определению амплитуд. В лучевых координатах ξ, η, τ уравнения переноса (1.3.7), (1.3.8) сводятся к системе обыкновенных дифференциальных уравнений, и можно выписать в общем виде их решения.

Решение уравнений переноса для двух первых членов ряда (1.3.4) имеет вид:

$$A_0(\tau) = A_0(\tau_0)\left|\frac{D(\tau_0)}{D(\tau)}\right|^{1/2}, \qquad \textbf{(1.3.17)}$$

$$A_1(\tau) = A_1(\tau_0)\left[\frac{D(\tau_0)}{D(\tau)}\right]^{1/2} - \frac{1}{2}\int_{\tau_0}^{\tau}\left[\frac{D(\bar{\tau})}{D(\tau)}\right]^{1/2}\Delta A_0 d\bar{\tau}. \qquad \textbf{(1.3.18)}$$

Здесь

$$D(\tau) = \frac{\partial(x,y,z)}{\partial(\xi,\eta,\tau)} = \begin{vmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} & \frac{\partial x}{\partial \tau} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} & \frac{\partial y}{\partial \tau} \\ \frac{\partial z}{\partial \xi} & \frac{\partial z}{\partial \eta} & \frac{\partial z}{\partial \tau} \end{vmatrix} \qquad \textbf{(1.3.19)}$$

- якобиан перехода от декартовых координат к лучевым; параметр τ_0, вообще говоря, произволен. Якобиан (1.3.19) легко вычисляется, если известны уравнения семейства лучей.

Например, цилиндрическая волна есть геометрооптическое поле, если исключить область порядка λ вблизи начала координат. Лучи совпадают с радиусами: под лучевыми координатами ξ, η, τ следует

понимать φ, z, $\mathbf{r}/n$. $D(\tau)=n\mathbf{r}$; и $A_0(\tau) = A_0(\tau_0)\sqrt{\frac{n(\tau_0)\mathbf{r}(\tau_0)}{n(\tau)\mathbf{r}(\tau)}}$. Разумеется, цилиндрическая волна может распространятся только в аксиально-симметричной среде (в частности, в среде с постоянным n). В сферически симметричной среде распространяется сферическая волна, $\tau = \rho n,\ D(\tau) = n\rho^2$, здесь координата ρ отсчитывается от центра волны, так что радиус кривизны волнового фронта $R=\rho$. В общем случае поверхности с двумя главными радиусами кривизны $D(\tau) = nR_1R_2$.

Рис. 1.3.2. *Лучевая трубка.*

Определим теперь понятие лучевой трубки. На начальном волновом фронте $\tau=\tau_0$ возьмем исходную поверхность $s(\tau_0)$ бесконечно малого размера (**рис. 1.3.2**).

Лучи, выходящие с контура поверхности, образуют стенки лучевой трубки. Величина $D(\tau)/D(\tau_0)$ пропорциональна отношению площадей элементарной лучевой трубки:

$$\frac{D(\tau)}{D(\tau_0)} = \frac{n(\tau)s(\tau)}{n(\tau_0)s(\tau_0)}, \qquad \textbf{(1.3.20)}$$

и может быть названа расходимостью лучей. Отсюда для нулевого, т.е. геометрооптического, приближения из (1.3.17), (1.3.20) получаем

$$A_0(\tau) = A_0(\tau_0)\left[\frac{n(\tau_0)s(\tau_0)}{n(\tau)s(\tau)}\right]^{1/2}. \qquad \textbf{(1.3.21)}$$

Таким образом, в геометрической оптике амплитуда на луче определяется только амплитудой на том же луче в любой точке, откуда этот луч пришел, геометрической расходимостью лучевой трубки и изменением n вдоль луча. Формула (1.3.17), таким образом, означает, что в лучевой трубке сохраняется поток энергии

$$A_0^2 ns = \text{const} \qquad \textbf{(1.3.22)}$$

Все лучевое поле можно себе представить состоящим из тонких трубок, причем в каждой трубке распространяется та энергия, которая была в ее начале. Распространение энергии в трубке происходит независимо от

соседних. Геометрическая оптика, которая ограничивает ряд (1.3.4) первым членом, допускает, что на протяжении произвольно длинной границы между лучевыми трубками с разной интенсивностью и, в частности, на границе между освещенной областью и теневой - не будет никакого обмена энергией.

Предположение о том, что взаимодействие между лучевыми трубками пренебрежимо мало, может оказаться неверным при продвижении вдоль трубки на достаточно большое расстояние. Действительно, уже для первого коэффициента лучевого разложения A_1 (1.3.18) кроме слагаемого, учитывающего геометрическую расходимость лучей, есть еще интегральное слагаемое, которое содержит производные амплитуды предыдущего приближения A_0. Если бы A_1 и A_0 были величинами одного порядка, то влияние A_1 на суммарное поле, как это следует из лучевого разложения (1.3.4), было бы в k раз меньше, чем влияние A_0. Но эффект взаимодействия между лучевыми трубками из-за интегрального, накапливающегося характера A_1 на достаточно длинном пути может существенно превзойти изменение A_0, связанное с изменением сечения трубки или показателя преломления n вдоль луча.

Итак, геометрическая оптика не дает правильного решения не только в случае, если член лучевого разложения A_1/k становится сравнимым с геометрооптическим членом A_0. Геометрическая оптика не может так же ничего сказать о поле в области тени, куда не проникают лучи. Наконец упомянем третий случай отказа геометрической оптики.

Он относится к ситуации, когда выделенная на заданном волновом фронте лучевая трубка при подходе к некоторой точке схлопывается, т.е. площадь трубки $s(\tau)$ становится равной нулю. При этом нулевой член лучевого разложения (1.3.21) становится бесконечно большим. Это означает, что структура поля локально не близка к плоской волне и основные геометрооптические представления теряют свой смысл.

1.3.5. Точка стационарной фазы. Область влияния

Пусть вдоль луча $\mathbf{r}_0$, $\mathbf{r}_0$ (см. **рис. 1.3.3**) распространяется световая волна, имеющая геометрооптическую структуру (1.3.9).

Основной вклад в поле $\Psi(r_1)$ дает окрестность точки так называемой стационарной фазы $\mathbf{r}=r_\nu$. Фаза вблизи стационарной точки r_ν квадратично зависит от расстояния до этой точки:

$$S(\mathbf{r}_1,\mathbf{r}) \approx S(\mathbf{r}_1,\mathbf{r}_\nu) + \frac{1}{2}\frac{d^2S}{dx^2}\bigg|_{x_\nu}(x-x_\nu)^2. \qquad \textbf{(1.3.27)}$$

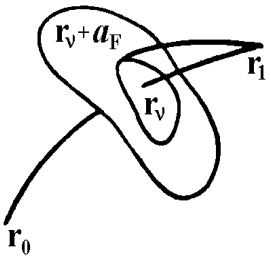


Рис. 1.3.3. *Зона влияния на поле в точке r_1 около точки стационарной фазы r_ν*

Окрестность точки стационарной фазы представляет собой ту область влияния, которая формирует поле в точке наблюдения. Область влияния, светящееся пятнышко, которое можно наблюдать, если глаз поместить в точку $\mathbf{r}_1$, можно назвать первой зоной Френеля. Уточним это понятие.

Предположим, что вблизи точки стационарной фазы падающее поле имеет форму сферической волны с радиусом кривизны R. Найдем разность эйконалов вдоль луча из r_ν в r_1 и вдоль луча, испускаемого в ту же точку r_1 точечным источником, мысленно помещенным на волновую поверхность на краю зоны Френеля в точке $r_\nu + a_\mathrm{F}$. По определению, будем считать a_F размером первой зоны Френеля, если эта разность эйконалов, умноженная на k, равна по модулю π:

$$k\left|S\left(r_\nu + a_F, r_1\right) - S\left(r_\nu, r_1\right)\right| = \pi. \qquad \textbf{(1.3.29)}$$

Особенно просто определить область влияния в однородной среде. В параксиальном приближении

$$a_\mathrm{F} << \max\left(|R|, z\right), \qquad \textbf{(1.3.30)}$$

оставляя в (1.3.29) только квадратичные по a_F члены, получим явную формулу для радиуса первой зоны Френеля

$$a_F = \left(\frac{\lambda}{n}\left|\frac{1}{R} + \frac{1}{z}\right|^{-1}\right)^{1/2}, \qquad \textbf{(1.3.31)}$$

где z- расстояние вдоль луча от точки стационарной фазы до точки наблюдения. Радиус кривизны волнового фронта считаем положительным ($R>0$) для расходящейся волны, отрицательным ($R<0$) для сходящейся.

Область влияния изменяется при изменении расстояния от выбранной точки наблюдения r_1 (**рис. 1.3.4**).

Если $z \to 0$, то светящееся пятнышко стягивается в точку, т.е. практически передача световой энергии идет по законам геометрической оптики, которые не учитывают никаких нелокальных воздействий.

Величина a_F при данном $R>0$ максимальна при отнесении z на бесконечность и равна при этом

$$a_F = \sqrt{\lambda R / n}, \quad z \to \infty. \qquad \textbf{(1.3.32)}$$

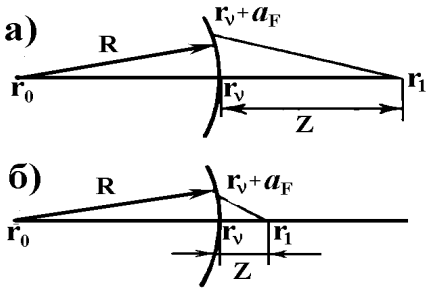


Рис. 1.3.4. *Уменьшение первой зоны Френеля с приближением точки наблюдения к выбранному волновому фронту.*

Если $R=\infty$, т.е. фронт в районе точки стационарной фазы плоский, то размер области влияния

$$a_F = \sqrt{\lambda z / n}. \qquad \textbf{(1.3.33)}$$

Если $R<0$, волна сходящаяся, то

$$a_F = \left(\frac{\lambda}{n} \left| \frac{1}{z} + \frac{1}{|R|} \right|^{-1} \right)^{1/2}. \qquad \textbf{(1.3.34)}$$

Очевидно, что на расстоянии $z=|R|$ размер светящегося пятна растет и заполняет всю поверхность. Таким образом, на поле в фокусе влияет вся светящаяся поверхность. При $z \to \infty$ размер пятна стремится к тому же пределу (1.3.32): $a_F \to \sqrt{\lambda |R| / n}$, как и для расходящейся волны. Заметим, что лишь для плоской волны такого предела не существует- пятно растет неограниченно с увеличением z (1.3.33).

В более общем случае неоднородной среды и двух радиусов кривизны волнового фронта в выбранной на луче точке r_ν также существует область вокруг луча, влияющая на формирование поля в точке r_1 . Форма этой области, которая находится двумерным методом стационарной фазы, может быть довольно сложной.

1.3.6. Условие применимости геометрической оптики

Неравенство (1.3.2), при выполнении которого волну можно считать почти плоской, а среду почти однородной, является необходимым, но недостаточным условием применимости геометрической оптики. Достаточные же условия применимости должны тем или иным способом учитывать накапливающиеся погрешности, обусловленные тем, что поле нулевого приближения (1.3.3) не является точным решением волнового уравнения. Корректный учет такого рода погрешностей в общем виде представляет собой весьма трудную задачу, еще ждущую своего решения. Однако обобщая результаты многих работ, выполненных в указанном направлении, можно сформулировать некий эвристический критерий выполнимости геометрической оптики. Этот критерий требует, чтобы вблизи луча на расстоянии, много большем линейного размера первой зоны Френеля, не было резких изменений ни свойств среды, ни свойств поля. Если обозначить поперечный масштаб изменения этих свойств (точнее, наименьший из масштабов) через $l_{\perp}$, то условие применимости геометрической оптики запишется в виде неравенства

$$a_F << l_{\perp}. \qquad \textbf{(1.3.35)}$$

Это условие намного более жесткое, чем условие (1.3.2) превышения масштабов среды и поля над длиной волны. Так, если волна почти плоская, то $a_{\mathrm{F}} \approx \sqrt{\lambda z / n}$, а этот размер в $\sqrt{z / \lambda n}$ больше длины волны, и при $z \to \infty$ становится как угодно большим. Растет до бесконечности область влияния и при приближении к фокусу: $z \to |R|, \quad R < 0$. Поэтому в реальной ситуации, когда нельзя обеспечить неизменные свойства среды на бесконечной поверхности, излучение плоского поля при достаточно больших z теряет лучевую структуру. В окрестности фокуса (а не только в фокальной плоскости) поле также принципиально не может быть геометрооптическим.

Область влияния для сходящихся или расходящихся лучей при $z \to \infty$ приближается к конечному пределу (1.3.32), поэтому, в отличие от плоской волны, они могут быть описаны в представлениях геометрической оптики, если только выполнено условие (1.3.35).

При фиксированной точке наблюдения, расположенной на луче, можно провести огибающую первых зон Френеля, выделив так называемый френелевский объем. Если, например, точка наблюдения расположена на расстоянии L от выделенной поверхности волнового фронта с радиусом кривизны R, то радиус вытянутой поверхности, ограничивающей френелевский объем,

$$a_F(z) = \left(\frac{\lambda z |L + R - z|}{|L + R|} \right)^{1/2}, \qquad \textbf{(1.3.36)}$$

где z отсчитывается от точки наблюдения вдоль луча (**рис. 1.3.5**). Эта формула получена из (1.3.31) очевидной заменой в ней R, т. е. радиуса кривизны на расстоянии z от точки наблюдения, на $L + R - z$.

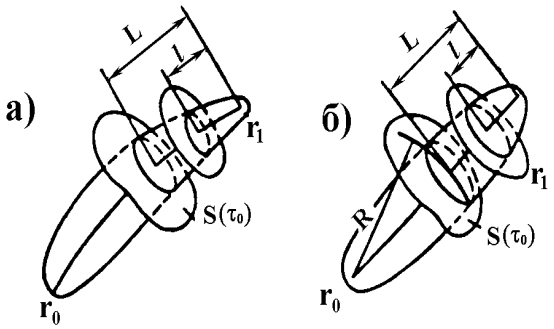


Рис. 1.3.5. *Френелевский объем (огибающая первых зон Френеля) для луча, соединяющего точки r_0 и r_1. а) луч криволинейный б) луч - прямая линия.*

1.3.7. Каустики

При анализе лучевой картины светового поля принято выделять важный структурный элемент, называемый каустикой. Каустика - это поверхность (или линия), огибающая систему лучей (**рис. 1.3.6**). Для плоской волны каустики нет. Каустика цилиндрической волны вырождается в фокальную линию (ось системы координат). Каустика сферической волны вырождается в точку ñ фокус. Каустика может сформироваться как в неоднородной среде, так и в однородной. Пример каустики в однородной среде приведен на **рис. 1.3.7**, где лучи ñ нормали к волновому фронту, который несколько отличен от сферического.

На каустике пересекаются бесконечно близкие лучи, поперечный размер лучевой трубки уменьшается до нуля, и лучевые разложения (1.3.4) поэтому не могут быть использованы, так как все коэффициенты A_i на каустике обращаются в бесконечность.

Каустика разделяет часть пространства, заполненную лучами, от каустической тени. В освещенной части через каждую точку проходит два луча ñ один из них уже коснулся каустики, другой еще нет. При подходе к каустике со стороны освещенной стороны наблюдается рост амплитуды поля, локальный максимум; при переходе через каустику и удалении от нее в область тени поле спадает. В направлении нормали к каустике поле в

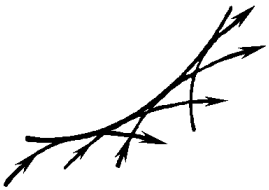


Рис. 1.3.6. *Каустика в неоднородной среде.*

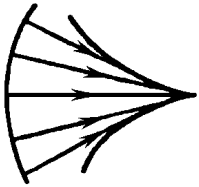


Рис. 1.3.7. *Каустика лучей - нормалей к волновой поверхности близ максимума ее кривизны.*

освещенной части имеет, из-за интерференции двух лучевых полей, характер стоячей волны. Вдоль каустики поле имеет характер бегущей волны.

Представление о структуре поля вблизи каустики можно получить, рассчитав распределение интенсивности в так называемом линейном слое. Пусть слева, со стороны пространства n=1, падает на плоскую границу раздела z=0 плоская волна (**рис. 1.3.8**). В области z>0 среда имеет линейно убывающую с ростом z диэлектрическую постоянную $n^2(z)$:

$$n^2(z) = \frac{z_0 - z}{z_0}, \qquad \textbf{(1.3.37)}$$

где z_0 - характеристика среды, та плоскость, на которой n=0. Предположим сначала, что плоская волна падает перпендикулярно к линейному слою. Тогда волновое уравнение для амплитуды поля

$$\frac{d^2\Psi}{dz^2} + k^2 n^2 \Psi = 0 \qquad \textbf{(1.3.38)}$$

после замены переменных

$$t = k^{2/3} z_0^{-1/3}\left(z - z_0\right) \qquad \textbf{(1.3.39)}$$

принимает вид

$$\frac{d^2\Psi}{dt^2} - t\Psi = 0. \qquad \textbf{(1.3.40)}$$

Решение уравнения (1.3.40) пропорционально функции Эйри $v(t)$, которая определяется интегралом

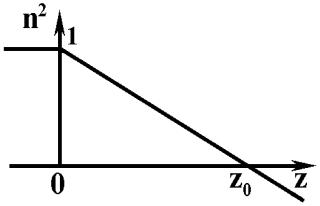


Рис. 1.3.8. *Показатель преломления слоя.*

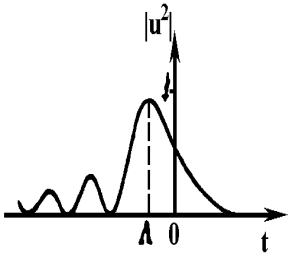


Рис. 1.3.9. *Структура светового поля вблизи каустики.*

$$v(t)=\frac{1}{\sqrt{\pi}}\int_0^{\infty}\cos\left(\frac{x^3}{3}+tx\right)dx. \qquad \textbf{(1.3.41)}$$

Искомое поле

$$\Psi(t)=Av(t), \qquad \textbf{(1.3.42,a)}$$

причем коэффициент пропорциональности определяется из требования непрерывности поля и его производной при $t=t_0$, где t_0ñ значение t (1.3.39) при z=0:

$$A=2\left[v(t_0)-i(kz_0)^{-1/3}v'(t_0)\right]^{-1}. \qquad \textbf{(1.3.42,b)}$$

Предполагается, что амплитуда падающего поля равна единице.

Вид функции Эйри приведен на **рис. 1.3.9**. Значение функции вблизи нуля порядка единицы, $v(0)$=0.63, максимум несколько сдвинут в сторону освещенной области и достигается при $t\approx$-1: v(ñ1.02)=0.95.

Слева от точки $z=z_0$ (т.е. при t<0) эта функция при больших значениях аргумента ($|t|$>>1) описывает распределение поля, схожее со стоячей волной.

Вблизи каустики и на каустике поле не может быть описано с помощью геометрооптических лучей. На каустике ñ потому, что сечение лучевой трубки обращается в нуль, коэффициент расходимости (1.3.20) равен нулю, и все амплитудные коэффициенты в лучевом разложении неограниченно растут. В окрестности каустики лучевое разложение неприменимо потому, что лучи становятся неразличимыми, так как разность эйконалов двух пересекающихся лучей, один из которых коснулся каустики, а другой еще нет, меньше $\lambda/2$.

Несмотря на то, что геометрооптические представления не применимы для прикаустической области, к их достоинствам следует все же

отнести возможность нахождения расположения каустической области, а в некоторых случаях - и ее ширины.

1.3.8. Элементы гамильтоновой оптики

Как уже указывалось, поведение лучей в световом потоке может быть определено путем использования вариационного принципа, предполагающего нахождение экстремалей функционала (1.3.14). Решение вариационной задачи оказывается более легким, если перейти к новой переменной интегрирования. Используем определение элемента длины

$$ds = \sqrt{dx^2 + dy^2 + dz^2} = \sqrt{1 + x'^2 + y'^2}\, dz, \qquad \textbf{(1.3.43)}$$

где

$$x' = \frac{dx}{dz}, \quad y' = \frac{dy}{dz}, \qquad \textbf{(1.3.44)}$$

чтобы выразить (1.3.14) в виде

$$\int_{P_1}^{P} L(x, y, x', y', z)\, dz = \min. \qquad \textbf{(1.3.45)}$$

Здесь P_1 и P_2начальная и конечная точки луча.

Функция L задается соотношением

$$L(x, y, x', y', z) = n(x, y, z)\sqrt{1 + x'^2 + y'^2}. \qquad \textbf{(1.3.46)}$$

Уравнение (1.3.45) имеет точно ту же форму, что и принцип наименьшего действия Гамильтона, достаточно подробно рассматриваемого в курсах теоретической механики. Единственное отличие принципа Ферма от принципа Гамильтона заключается в том, что в принципе Ферма вместо переменной 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 (1.3.45) хорошо известно в вариационном исчислении и нет необходимости приводить его здесь. Оно дается уравнениями Эйлера для вариационной задачи

$$\frac{d}{dz}\frac{\partial L}{\partial x'} - \frac{\partial L}{\partial x} = 0, \qquad \textbf{(1.3.47)}$$

$$\frac{d}{dz}\frac{\partial L}{\partial y'} - \frac{\partial L}{\partial y} = 0. \qquad \textbf{(1.3.48)}$$

Подстановка выражения (1.3.46) приводит к соотношениям

$$\frac{d}{dz}\frac{nx'}{\sqrt{1+x'^2+y'^2}} = \sqrt{1+x'^2+y'^2}\,\frac{\partial n}{\partial x}, \qquad \textbf{(1.3.49)}$$

$$\frac{d}{dz}\frac{ny'}{\sqrt{1+x'^2+y'^2}} = \sqrt{1+x'^2+y'^2}\,\frac{\partial n}{\partial y}. \qquad \textbf{(1.3.50)}$$

Используя известную из классической механики связь между уравнениями Эйлера (1.3.47), (1.3.48) и уравнениями Гамильтона, а также соотношения (1.3.49), (1.3.50) можно формально уравнениям лучей придать вид гамильтоновых уравнений:

$$\frac{dx}{dz} = \frac{\partial H}{\partial p_x}., \qquad \textbf{(1.3.51)}$$

$$\frac{dy}{dz} = \frac{\partial H}{\partial p_y}., \qquad \textbf{(1.3.52)}$$

$$\frac{dp_x}{dz} = -\frac{\partial H}{\partial x}., \qquad \textbf{(1.3.53)}$$

$$\frac{dp_y}{dz} = -\frac{\partial H}{\partial y}. \qquad \textbf{(1.3.54)}$$

Здесь *x, y, z* - соответственно поперечные и продольная координаты точки на луче, а "импульсы" p_x и p_y задаются формулами

$$p_x = \frac{nx'}{\sqrt{1+x'^2+y'^2}}, \qquad \textbf{(1.3.55)}$$

$$p_y = \frac{ny'}{\sqrt{1+x'^2+y'^2}}. \qquad \textbf{(1.3.56)}$$

Из последних соотношений видно, что в параксиальном приближении значения импульсов определяют значения углов наклона лучей. Сам же гамильтониан выражается через импульсы с помощью соотношения

$$H = -\sqrt{n^2 - p_x{}^2 - p_y{}^2}. \qquad \textbf{(1.3.57)}$$

Представление геометрической оптики в гамильтоновой форме не просто расширяет ее аппарат, Такое представление позволяет перенести на многие геометрооптические задачи элементы анализа, применяемого к поведению динамических систем. В частности, как мы увидим в следующей главе, динамическая теория перехода детерминированных систем к хаосу позволяет вскрыть один весьма необычный механизм стохастизации излучения в регулярно- неоднородных средах.

ЛИТЕРАТУРА К ГЛАВЕ 1

1. Борн М., Вольф Э. Основы оптики. - М.: Наука, 1970, 856 с.

2. Зоммерфельд А. Оптика. - ИЛ, 1953, 487 с.

3. Гудмен Дж. Введение в фурье-оптику. - М.: Мир, 1970, 364 с.

4. Королев Ф.А. Теоретическая оптика. - М.: Высшая школа, 1968, 556 с.

5. Матвеев А.И. Оптика. - М.: Высшая школа, 1986, 352 с.

6. Калитеевский Н.Л., Волновая оптика. - М.: Высшая школа, 1978, 384 с.

7. Ахманов С.А., Дьяков Ю.Е., Чиркин А.С. Введение в статистическую радиофизику и оптику. М.: Наука, 1981, 640 с.

8. Маркузе Д. Оптические волноводы. - М.: Мир, 1974, 576 с.

9. Солимено С., Крозиньяни Б., Ди Порто П. Дифракция и волноводное распространение излучения. - М.: Мир, 1989, 664 с.

10. Ваганов Р.Б., Каценеленбаум Б.З. Основы теории дифракции. - М.: Наука, 1982, 272 с.

11. Виноградова М.Б., Руденко О.В., Сухоруков А.П. Теория волн - М.: Наука, 1990, 432 с.

12. Семенов А.А. Теория электромагнитных волн. - М.: Изд-во МГУ, 1968, 316 с.

13. Кравцов Ю.А., Орлов Ю.И. Геометрическая оптика неоднородных сред. - М.: Наука, 1980, 304 с.

14. Русинов М.М., Грамматин А.П., Иванов П.Д. и др. Вычислительная оптика. Справочник. - Ленинград: Машиностроение, 1984, 424 с.

ГЛАВА 2. ФОРМИРОВАНИЕ И РАСПРОСТРАНЕНИЕ ВОЛНОВЫХ ПУЧКОВ

В настоящее время наиболее распространенным источником когерентного излучения является лазер. Как правило, лазер генерирует излучение в виде слабо расходящегося волнового пучка, размеры которого в поперечной плоскости велики по сравнению с длиной волны. Настоящая глава посвящена анализу процессов формирования и распространения узконаправленных пучков лазерного излучения. Центральным понятием излагаемой теории является понятие моды - такого волнового пучка, который сохраняет в процессе распространения форму распределения амплитуды и фазы световых колебаний в поперечном сечении.

2.1 Моды свободного пространства [1-5]

2.1.1 Параболическое приближение

Волновой лазерный пучок в силу своей высокой направленности имеет много общего с плоской волной. Отличие же его от плоской волны состоит в том, что распределение интенсивности в нем неоднородно (мощность пучка, в основном, сконцентрирована вблизи оси), а фазовый фронт несколько отличается от плоского. Поэтому решение приведенного волнового уравнения, описывающее распространение такого пучка, будем искать в виде

$$\Psi = u(x, y, z)e^{ikz}, \qquad (2.1.1)$$

где *u(x,y,z)* - медленно меняющаяся комплексная функция, которая и определяет свойства лазерного пучка, отличающие его от плоской волны. Быстрые осцилляции по *z* обусловлены присутствием в (2.1.1) экспоненциального множителя. Применяя оператор Δ к функции Ψ, имеем

$$\Delta\Psi = \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} + 2ik\frac{\partial u}{\partial z} - k^2 u\right)e^{ikz}. \qquad (2.1.2)$$

Если в выражении (2.1.2) пренебречь второй производной *u* по *z* по сравнению с первой, то на основании (1.2.20) получим уравнение

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + 2ik\frac{\partial u}{\partial z} = 0. \qquad (2.1.3)$$

Полученное уравнение относится к уравнениям параболического типа, а само приближение, в рамках которого оно было получено, называется параболическим приближением. Нетрудно показать, что уравнению (2.1.3) будет удовлетворять так называемый гауссов пучок, амплитуда которого меняется по поперечной координате по гауссовому закону.

2.1.2 Свойства основной моды

Для гауссова пучка можно записать выражение

$$u = a \exp\left[i\left(p + \frac{k}{2q} r^2 \right)\right], \qquad \textbf{(2.1.4)}$$

где $r^2 = x^2 + y^2$. Параметр P - комплексный фазовый сдвиг при распространении света вдоль оси z, а q - комплексный параметр пучка, определяющий гауссово распределение поля по координате r, где r - расстояние от оси. Кроме того, q определяет кривизну волнового фронта, который вблизи оси является сферическим. Подставив выражение (2.1.4) в (2.1.3), получим

$$2k\left(p' + \frac{i}{q} \right) + \left(\frac{kr^2}{q} \right)\left(1 - q'\right) = 0 . \qquad \textbf{(2.1.5)}$$

Штрих означает производную по z. Уравнение (2.1.5) эквивалентно двум уравнениям

$$q = 1, \qquad \textbf{(2.1.6)}$$

$$p' = -\frac{i}{q} . \qquad \textbf{(2.1.7)}$$

Интегрируя (2.1.6), получаем

$$q_2 = q_1 + z . \qquad \textbf{(2.1.8)}$$

Это уравнение устанавливает весьма простое соотношение между параметром пучка в разных сечениях, отстоящих друг от друга на расстоянии z.

Когерентный световой пучок с гауссовым распределением поля имеет фундаментальное значение в теории волновых пучков. Этот пучок называют основной модой в отличие от других мод более высокого порядка, которые будут рассматриваться ниже. Вследствие особой важности рассмотрим свойства гауссова пучка с длиной волны λ более подробно. Для этого выразим комплексный параметр q через два действительных параметра пучка R и w

$$\frac{1}{q} = \frac{1}{R} + i\frac{\lambda}{\pi w^2}. \qquad \textbf{(2.1.9)}$$

Физический смысл этих параметров становится ясным при подстановке соотношения (2.1.9) в (2.1.4). Видно, что R есть радиус кривизны волнового фронта, а w характеризует изменение поля E в поперечной плоскости. Распределение поля в этой плоскости, как видно из *рис.* 2.1.1, подчиняется закону Гаусса и w равно расстоянию, на котором амплитуда поля убывает в e раз по сравнению с полем на оси.

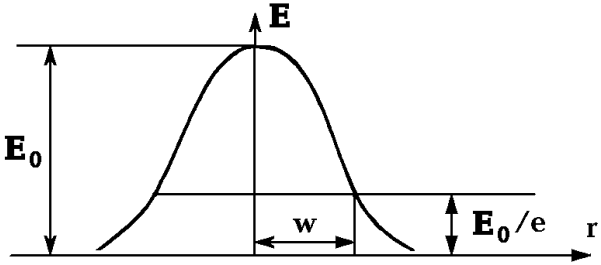


Рис. 2.1.1. *Поперечное распределение амплитуды поля для пучка основной моды.*

Важно отметить, что гауссов характер распределение поля будет иметь в любой плоскости, будет меняться лишь ширина этого распределения. Параметр w принято называть радиусом пучка, а $2w$ - диаметром пучка. В некоторой плоскости, называемой горловиной пучка, гауссов пучок стягивается к минимальному диаметру $2w_0$. В этой плоскости, от которой целесообразно отсчитывать расстояние z, фазовый фронт является плоским и комплексный параметр пучка становится чисто мнимым

$$q_0 = \frac{\pi w_0^2}{i\lambda}. \qquad \textbf{(2.1.10)}$$

На расстоянии z от горловины

$$q = q_0 + z = \frac{\pi w_0^2}{i\lambda} + z. \qquad \textbf{(2.1.11)}$$

Из сопоставления соотношений (2.1.9) и (2.1.11) легко получить следующие важные в практическом отношении выражения:

$$w^2(z) = w_0^2\left[1+\left(\frac{\lambda z}{\pi w_0^2}\right)^2\right], \qquad \textbf{(2.1.12)}$$

$$R(z) = z\left[1+\left(\frac{\pi w_0^2}{\lambda z}\right)^2\right]. \qquad \textbf{(2.1.13)}$$

Изменение радиуса, задаваемое выражением (2.1.12), графически иллюстрируется *рис. 2.1.2*.

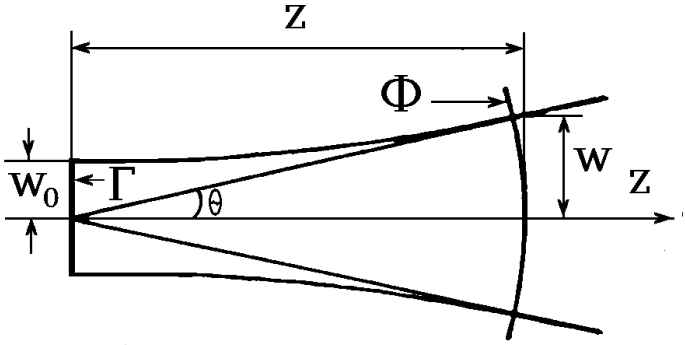


Рис. 2.1.2. *Продольная структура гауссова пучка: Ф - фазовый фронт, Г - горловина пучка.*

Образующая пучка *w(z)* представляет собой гиперболу, асимптота которой наклонена к оси под углом

$$\theta = \lambda/\pi w_0 . \qquad \textbf{(2.1.14)}$$

Этот угол равен углу дифракции основной моды в дальней зоне.

Для расчета комплексного фазового сдвига на расстоянии z от горловины пучка воспользуемся соотношениями (2.1.7) и (2.1.11); в результате получим

$$p' = -\frac{i}{q} = -\frac{i}{z+i\left(\pi w_0^2/\lambda\right)}. \qquad \textbf{(2.1.15)}$$

Интегрирование уравнения (2.1.15) дает

$$iP(z) = \ln\left[1-i\left(\lambda z/\pi w_0^2\right)\right] = \ln\sqrt{1+\left(\lambda z/\pi w_0^2\right)^2} - i\,\mathrm{arctg}\left(\lambda z/\pi w_0^2\right). \qquad \textbf{(2.1.16)}$$

Действительная часть P представляет собой разность фаз Φ между гауссовым пучком и идеальной плоской волной, а мнимая - амплитудный

фактор w_0/w, который характеризует уменьшение интенсивности на оси из-за расширения пучка. С учетом полученных соотношений выражение (2.1.1) принимает вид

$$\Psi(r,z) = \frac{w_0}{w}\exp\left\{i(kz-\Phi) - r^2\left(\frac{1}{w^2} - \frac{ik}{2R}\right)\right\}, \qquad \textbf{(2.1.17)}$$

где

$$\Phi = \operatorname{arctg}\left(\lambda z/\pi w_0^2\right). \qquad \textbf{(2.1.18)}$$

Из формулы (2.1.18) видно, что Φ растет с увеличением z и уменьшением минимального радиуса пучка w_0. Максимальное значение Φ равно $\pi/2$. Наличие в показателе экспоненты выражения (2.1.17) члена $ikr^2/2R$ обусловлено отставанием по фазе световых колебаний на периферии гауссова пучка из-за кривизны волнового фронта.

2.1.3. Моды высших порядков

В предыдущем параграфе рассматривалось лишь одно решение уравнения (2.1.3), а именно гауссов пучок, являющийся основной модой свободного пространства. Существуют, однако, и другие решения уравнения (2.1.3), которым соответствуют пучки с сохраняющейся формой распределения амплитуды поля по поперечному сечению - высшие моды свободного пространства. Все решения (2.1.3) образуют полную ортогональную систему функций, поэтому любое произвольное распределение монохроматического поля может быть разложено по модам свободного пространства.

В прямоугольной системе координат *x, y, z* решение уравнения (2.1.3) может быть записано в виде

$$u = g\left(\frac{x}{w}\right)h\left(\frac{y}{w}\right)\exp\left\{i\left[p + \frac{k}{2q}\left(x^2+y^2\right)\right]\right\}, \qquad \textbf{(2.1.19)}$$

где g - функция x и z, а h - функция y и z. Для действительных g и h это решение описывает моды, поперечное распределение поля которых связано с радиусом гауссова пучка $w(z)$. Подставляя (2.1.19) в (2.1.3), можно убедиться, что функция g и h удовлетворяют тому же самому дифференциальному уравнению, что и полиномы Эрмита $H_n(t)$

$$\frac{d^2H_n}{dt^2} - 2t\frac{dH_n}{dt} + 2nH_n = 0, \qquad \textbf{(2.1.20)}$$

где n - целое число, а $t = \sqrt{2}\,x/w$ для функции g и $t = \sqrt{2}\,y/w$ для функции h. Таким образом

$$g \cdot h = H_m\left(\sqrt{2}\,\frac{x}{w}\right) \cdot H_n\left(\sqrt{2}\,\frac{y}{w}\right), \qquad \textbf{(2.1.21)}$$

где m, как и n - целое число.

Физический смысл индексов m и n, называемых поперечными индексами моды, заключается в том, что они показывают, сколько раз поле меняет знак соответственно в направлении x и y. Важно отметить, что моды всех порядков характеризуются одним и тем же масштабным параметром $w(z)$. Полиномы Эрмита низших порядков равны

$$H_0(t) = 1,$$
$$H_1(t) = 2t,$$
$$H_2(t) = 4t^2 - 2,$$
$$H_3(t) = 8t^3 - 12t.$$

Для математического описания мод более высоких порядков можно использовать выражение (2.1.17), если в правую его часть вставить произведение $g \cdot h$. Распределение поля в модах свободного пространства будет определяться, таким образом, произведением функций Эрмита и Гаусса

$$\Psi(x,y,z) = \frac{w_0}{w} H_m\left(\sqrt{2}\,\frac{x}{w}\right) H_n\left(\sqrt{2}\,\frac{y}{w}\right) \times \exp\left\{ i(kz - \Phi) - \left(x^2 + y^2\right)\left(\frac{1}{w^2} - \frac{ik}{2R}\right)\right\} \qquad \textbf{(2.1.22)}$$

В частном случае m=0, n=0 мы имеем гауссов пучок - основную моду свободного пространства. Параметр $R(z)$ в (2.1.22) для всех мод одинаков. Это означает, что кривизна волнового фронта одинакова для всех мод и закон его изменения один и тот же. Однако фазовый сдвиг Φ зависит от поперечного индекса. Можно найти, что

$$\Phi(m,n,z) = (m+n+1)\,arctg\left(\frac{\lambda z}{\pi w_0^2}\right). \qquad \textbf{(2.1.23)}$$

Из выражения (2.1.23) видно, что фазовая скорость с ростом индекса увеличивается.

Если решать уравнение (2.1.3) в цилиндрической системе координат r, φ, z, подобное решение записывается в виде

$$\Psi = g\left(\frac{r}{w}\right)\exp\left\{i\left(p+\frac{k}{2q}r^2+l\varphi\right)\right\}. \qquad \textbf{(2.1.24)}$$

После некоторых преобразований можно получить

$$g = \left(\sqrt{2}\,\frac{r}{w}\right)^l \cdot L_p^l\left(2\,\frac{r^2}{w^2}\right), \qquad \textbf{(2.1.25)}$$

где L_p^l - обобщенный полином Лагерра, а p и l - соответственно угловой и радиальный индексы, показывающие, сколько раз поле меняет знак в радиальном и азимутальном направлениях.

Полиномы Лагерра низших порядков равны:

$$L_0^l(t) = 1,$$

$$L_1^l(t) = l + 1 - t$$

$$L_2^l(t) = \frac{1}{2}(l+1)(l+2) - (l+2)t + \frac{1}{2}t^2$$

В цилиндрической системе координат, тем самым, моды будут описываться лагерро - гауссовыми функциями

$$\Psi(r,\varphi) = \frac{w_0}{w}\left(\sqrt{2}\,\frac{r}{w}\right)^l \cdot L_p^l\left(2\,\frac{r^2}{w^2}\right)\exp\left\{i(kz-\Phi) - r^2\left(\frac{1}{w^2}-\frac{ik}{2R}\right)\right\}\begin{Bmatrix}\cos l\varphi\\ \sin l\varphi\end{Bmatrix}. \qquad \textbf{(2.1.26)}$$

Как и для случая прямоугольных координат, параметры w и R одинаковы для всех цилиндрических волн, а разность фаз, как и ранее, зависит от индексов моды и определяется уравнением

$$\Phi(p,l,z) = (2p+l+1)\,arctg\left(\frac{\lambda z}{\pi w_0^2}\right) \qquad \textbf{(2.1.27)}$$

Графические распределения амплитуды поля для некоторых низших мод приведены на *рис. 2.1.3.* Для наглядности под каждым графиком приведена картина, наблюдаемая на экране при падении на него светового пучка, соответствующего определенной моде.

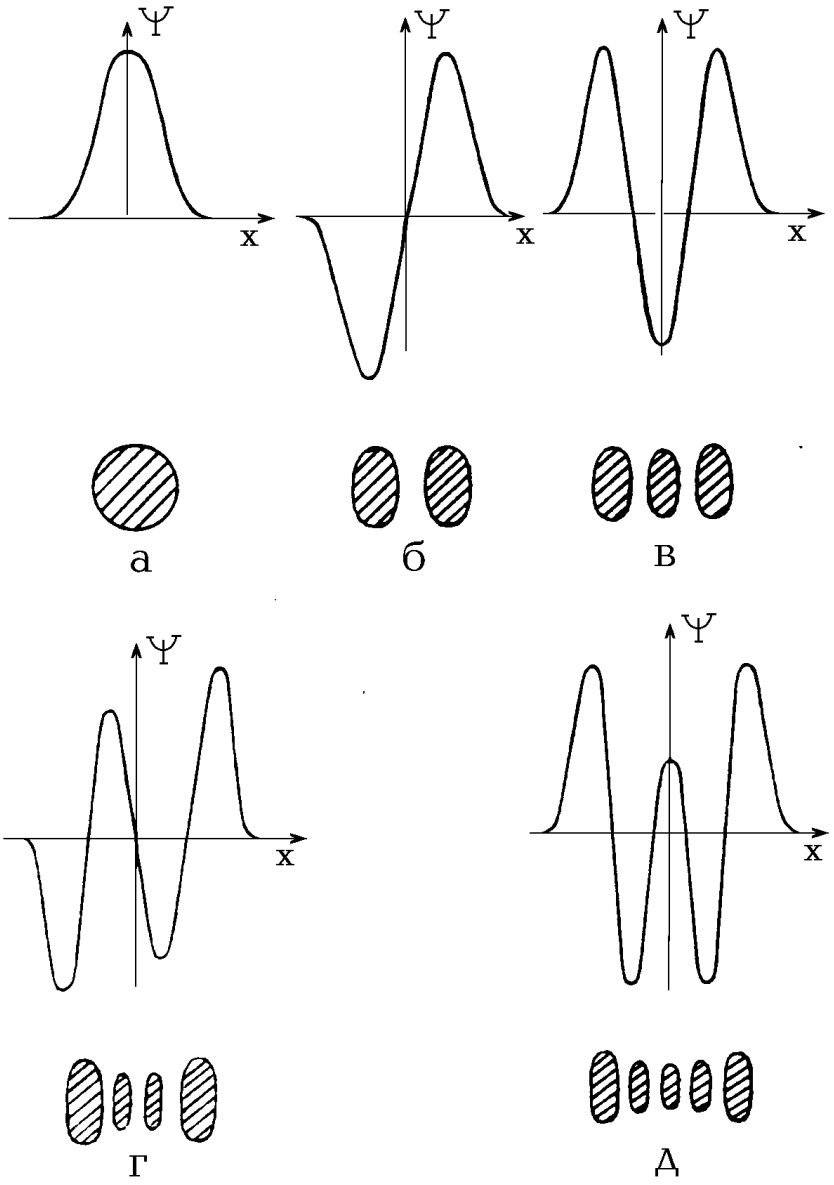


Рис. 2.1.3. *Распределение амплитуды поля низших TEM_{mn}-мод: а - TEM_{00}; б - TEM_{10}; в - TEM_{20}; г - TEM_{30}; д - TEM_{40}.*

Как эрмито-гауссовые, так и лагерро-гауссовые моды, реализуемые на практике, характеризуются, как правило, большим значением радиуса кривизны волнового фронта. Поэтому их с хорошей степенью приближения можно отнести к поперечным электромагнитным волновым пучкам вида TEM. С учетом поперечных индексов эти пучки часто обозначаются как пучки TEM_{mn} или TEM_{pl}.

Следует, однако, помнить, что световые поля, записанные в виде (2.1.22) и (2.1.26), являются лишь приближенными решениями волнового уравнения. Степень приближения ухудшается с увеличением чисел *(m+n)* или *(2p+l)* и решение перестает быть верным, если величина, содержащая в качестве сомножителя сумму $m+n+1$ (либо $2p+l+1$) становится сравнимой со значением kz. Представление же произвольного поля в виде разложения по модам (2.1.22) и (2.1.26) 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 вводить дополнительные коэффициенты к членам ряда.

2.1.4. Преобразование волновых пучков с помощью линз

Как известно, линзы широко применяются либо для фокусировки лазерного пучка в пятна небольших размеров, либо для соответствующего преобразования диаметра и кривизны волнового фронта пучка с целью ввода в данную оптическую систему. Идеальная линза или система линз не изменяет поперечного распределения поля моды свободного пространства. Иначе говоря, входная основная гауссова мода после прохождения линзовой системы сохраняется, а моды высших порядков преобразуются на выходе в моды тех же порядков. Однако при этом параметры мод *R(z)* и *w(z)* претерпят изменения. Рассмотрим соотношение между входными параметрами, обозначаемыми индексом 1, и соответствующим выходными параметрами с индексом 2.

Идеальная тонкая линза с фокусным расстоянием f преобразует сферическую волну радиуса R_1, слева от линзы в сферическую волну с радиусом R_2 справа так, что

$$\frac{1}{R_2} = \frac{1}{R_1} - \frac{1}{f}. \qquad \textbf{(2.1.28)}$$

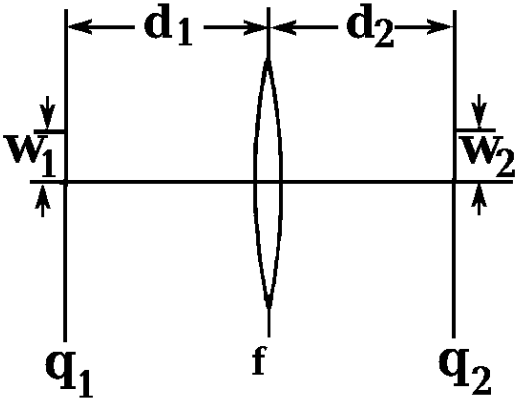


Рис. 2.1.5. *К выводу формулы (2.1.30)*

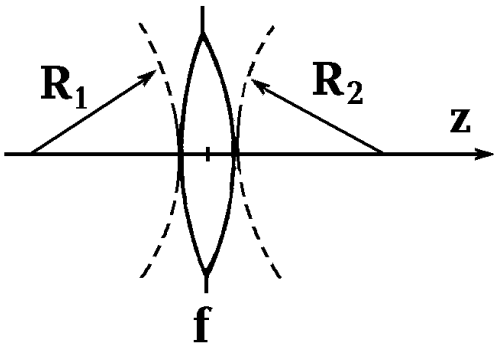


Рис. 2.1.4. *Преобразование волнового фронта тонкой линзой. Пунктир -волновые фронты.*

Это преобразование поясняется *рис. 2.1.4.* Радиус кривизны считается положительным, когда волновой фронт выпуклый, если смотреть из точки $z \to \infty$. Линзовая система преобразует волновой фронт моды так же, как и сферическую волну. Если линза тонкая и диаметр пучка слева и справа от нее одинаков, то соотношение между комплексными параметрами входного и выходного пучка запишется по аналогии с (2.1.28) в виде

$$\frac{1}{q_2} = \frac{1}{q_1} - \frac{1}{f}, \qquad \textbf{(2.1.29)}$$

где $q_{1,2}$ измерены непосредственно у линзы.

Если же q_1 и q_2 измерять на расстоянии d_1 и d_2 от линзы, как показано на *рис. 2.1.5*, то соотношение между ними принимает вид

$$q_2 = \frac{\left(1 - d_2/f\right)q_1 + \left(d_1 + d_2 - d_1 d_2/f\right)}{-\left(q_1/f\right) + \left(1 - d_1/f\right)} \qquad \textbf{(2.1.30)}$$

Эта формула получается непосредственно из выражений (2.1.8) и (2.1.9).

Более сложные оптические системы, такие, как комбинация линз, газовые линзы или толстые линзы, следует рассматривать как последовательность тонких линз, расположенных на различных расстояниях. Для расчета прохождения пучка лазера через сложные системы, таким образом, достаточно последовательного приложения к системе соотношений (2.1.28) и (2.1.29).

2.1.5. Геометрооптическое описание распространения и преобразования волновых пучков

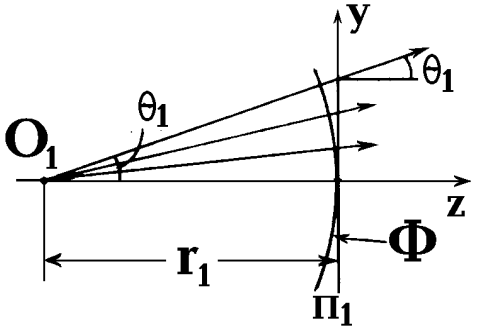


Рис. 2.1.6. *Пучок лучей, соответствующий волновому фронту Ф в плоскости поперечного сечения $П_1$.*

Решение многих важных в практическом отношении задач существенным образом упрощается, если воспользоваться связью, которая существует между оптикой гауссовых пучков и геометрической оптикой. Сферическому волновому фронту гауссова пучка в любом поперечном сечении можно поставить в соответствие пучок лучей, исходящих из одной и той же точки O_1 (*рис. 2.1.6*).

Для каждого параксиального луча отношение координаты y_1 точки его пересечения с плоскостью $П_1$ к лучевому углу θ_1 равно r_1, т.е.

$$r_1 = y_1/\theta_1 . \qquad \textbf{(2.1.31)}$$

Кривизну волнового фронта считают положительной, когда волна расходится. При этом для радиуса кривизны фронта принято использовать ì приведенное значениеî $R = (r/n) = (y/v)$, где n - показатель преломления среды, v = θn. Использование приведенных значений удобно тем, что при пересечении пучком границы раздела двух сред значение R не изменяется.

Рассмотрим, как преобразуется радиус кривизны R при прохождении пучком некоторой оптической системы. Пусть входная плоскость этой системы совпадает с плоскостью $П_1$; выходную же обозначим через $П_2$. Параметры любого луча из рассматриваемого лучевого пучка будут преобразовываться системой в соответствии с соотношениями

$$y_2 = Ay_1 + Bv_1, \qquad \textbf{(2.1.32)}$$

$$v_2 = Cy_1 + Dv_1. \qquad \textbf{(2.1.33)}$$

Параметры y_2 и v_2 соответствуют плоскости $П_2$. Величины A, B, C, D характеризуют конкретную оптическую систему. Уравнения (2.1.32) и (2.1.33) можно записать в матричной форме

$$\begin{vmatrix} y_2 \\ v_2 \end{vmatrix} = \begin{vmatrix} AB \\ CD \end{vmatrix} \begin{vmatrix} y_1 \\ v_1 \end{vmatrix}, \qquad \textbf{(2.1.34)}$$

где матрица $ABCD$ является матрицей передачи луча. Разделим уравнение (2.1.32) на уравнение (2.1.33)

$$\frac{y_2}{v_2} = \frac{Ay_1 + Bv_1}{Cy_1 + Dv_1} = \frac{A(y_1/v_1) + B}{C(y_1/v_1) + D}. \qquad \textbf{(2.1.35)}$$

Заменяя значения y/v на R, получаем следующее важное соотношение, связывающее значения приведенного радиуса на выходе и входе системы

$$R_2 = \frac{AR_1 + B}{CR_1 + D}. \qquad \textbf{(2.1.36)}$$

Преобразование (2.1.36) называют правилом *ABCD*. Поскольку преобразование оптическими элементами параметра гауссова пучка q аналогично преобразованию радиуса кривизны волнового фронта R, правилу *ABCD* можно придать вид

$$q_2 = \frac{Aq_1 + B}{Cq_1 + D}. \qquad \textbf{(2.1.37)}$$

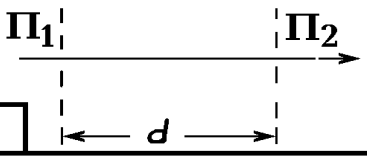
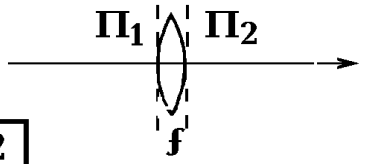
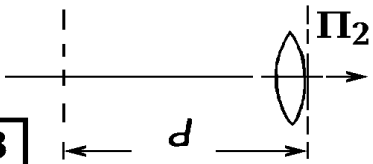
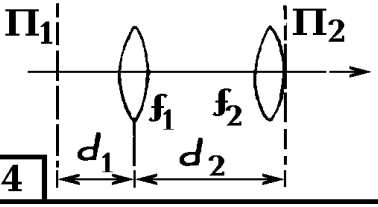
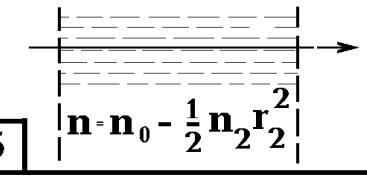
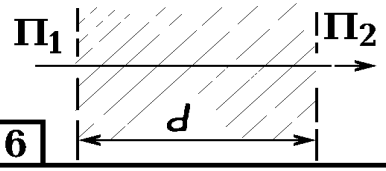
Тем самым, зная матрицу передачи луча, можно в соответствии с (2.1.37) определить трансформацию гауссова пучка той или иной оптической системой. Так как матрицы передачи луча рассчитаны в настоящее время для многих оптических систем, использование правила *ABCD* существенным образом упрощает описание происхождения гауссовых пучков через

различные среды и устройства. Матрицы передачи светового луча для шести простейших оптических систем приведены в *табл. 2.1.1*. Плоскости $П_1$ и $П_2$ помечены пунктиром. Матрица № 1 описывает прохождение луча в свободном пространстве на расстоянии *d*, матрица № 2 - прохождение луча через тонкую линзу с фокусным расстоянием *f*. Матрица № 3 представляет собой комбинацию первых двух. Последовательное соединение двух систем типа № 3 описывается матрицей № 4, которая получается простым перемножением матриц. Под номером 5 представлена матрица передачи луча для линзоподобной квадратичной по показателю преломления *n* среды длиной *d*. Матрица № 6 соответствует диэлектрику с показателем преломления *n* и длиной *d*. Сравнивая матрицы № 1 и № 6, можно заметить, что оптически более плотная среда уменьшает эффективную длину пути параксиальных лучей, хотя, как известно, оптическая длина пути увеличивается.

Геометрооптические методы расчета характеристик эрмито-гауссовых или лагерро-гауссовых пучков, распространяющихся в оптических системах, широко используются в современной оптике. Обширная практика их применения, многократные сопоставления результатов расчетов с экспериментальными данными, не оставляют сомнений в их правильности и эффективности.

Однако приведенное выше физическое обоснование указанных методов, апеллирующее к сходству изменений комплексного параметра гауссова пучка *q* и радиуса кривизны волнового фронта *R*, нельзя считать достаточно строгим. Более последовательное и строгое доказательство применимости геометрического подхода можно осуществить, используя формализм расчета полей от источников, расположенных в комплексной плоскости. В рамках этой теории, с которой можно более подробно ознакомиться по монографии [5], гауссов пучок оказывается эквивалентным волне от точечного источника, расположенного в точке с координатами (0,0,-*ib*), где *i* - мнимая единица, $b=\pi w_0^2 i\lambda$. "Сферический характер" такой волны делает более наглядным сходство преобразования в оптических системах сферических волн и гауссовых пучков.

Таблица 2.1. Матрицы передачи для простейших оптических систем.

Оптическая система	Матрица передачи луча
1 	$\begin{vmatrix} 1 & d \\ 0 & 1 \end{vmatrix}$
2 	$\begin{vmatrix} 1 & 0 \\ -\frac{1}{f} & 1 \end{vmatrix}$
3 	$\begin{vmatrix} f & d \\ -\frac{1}{f} & 1-\frac{1}{f} \end{vmatrix}$
4 	$\begin{vmatrix} 1-\frac{d_2}{f_1} & d_1+d_2-\frac{d_1 d_2}{f_1} \\ -\frac{1}{f_1}-\frac{1}{f_2}+\frac{d_2}{f_1 f_2} & 1-\frac{d_1}{f_1}-\frac{d_2}{f_2}-\frac{d_2}{f_1}+\frac{d_1 d_2}{f_1 f_2} \end{vmatrix}$
5 	$\begin{vmatrix} \cos d\sqrt{\frac{n_2}{n_1}} & \frac{1}{\sqrt{n_0 n_2}}\sin d\sqrt{\frac{n_2}{n_0}} \\ -\sqrt{n_0 n_2}\sin d\sqrt{\frac{n_2}{n_1}} & \cos d\sqrt{\frac{n_2}{n_0}} \end{vmatrix}$
6 	$\begin{vmatrix} 1 & \frac{d}{n} \\ 0 & 1 \end{vmatrix}$

2.1.6. Расчет поля дифрагированной волны методом разложения по модам свободного пространства

Скалярная теория дифракции, рассмотренная в первой главе, дает универсальный подход расчета поля дифракции волновых пучков на различного рода препятствиях. Однако составленные на основе дифракционных интегралов программы машинного расчета поля дифракции для ближней и дальней зон весьма сложны и имеют существенные отличия.

Свойство ортогональности мод свободного пространства дает возможность использовать для расчета поля дифракции иной подход, основанный на разложении поля пучка, претерпевающего дифракцию, по модам свободного пространства. Оно во многих случаях оказывается более простым и позволяет на основе одного алгоритма рассчитывать поля в ближней и дальней зонах дифракции. Этот метод особенно эффективен в случае слабого дифрагирования лазерных пучков. Плавный профиль изменения амплитуды лазерного пучка в плоскости апертурного разложения позволяет исключить в разложении моды со слишком высокими индексами, что сокращает время расчета и повышает его точность.

Пусть лазерная волна падает на отверстие произвольной формы в непрозрачном экране. Будем считать, что поперечное распределение амплитуды поля пучка в полярных координатах определяется функцией $u(r, \varphi)$. Представим поле дифрагировавшей волны в виде суперпозиции лагерро-гауссовых TEM_{pl}-мод свободного пространства (p,l - радиальный и угловой индексы соответственно), поперечное распределение поля которых характеризуется нормированными ортогональными функциями ψ_{pl}. Тогда амплитуду светового поля $u_D(r,\varphi,z)$ на расстоянии z от отверстия можно представить в виде ряда

$$\mathrm{u_D}(\mathrm{r},\varphi,z) = \sum_{p,l}^{\infty} C_{pl}\,\psi_{pl}(\mathrm{r},\varphi,z), \qquad \textbf{(2.1.38)}$$

ось z совпадает с направлением распространения TEM_{pl}-волн (в плоскости отверстия z=0). Коэффициенты разложения можно найти из соотношения

$$C_{pl} = \int_S u_D(r,\varphi)\,\psi_{pl}(r,\varphi)\,dS, \qquad \textbf{(2.1.39)}$$

где интегрирование ведется по площади отверстия S. Минимальный радиус пучка мод TEM_{pl} в практических расчетах выбирается из соображений быстрой сходимости ряда (2.1.38). Таким образом, алгоритм расчета поля дифракции как в ближней, так и в дальней зонах сводится к двум

операциям: нахождению коэффициентов разложения и суммированию амплитуд TEM_{pl}-волн в плоскости наблюдения.

Рассмотрим частный случай нормального падения основной волны лазера TEM_{00} на на круглую центрированную диафрагму радиуса r_0. Световое поле на отверстии описываетс функцией

$$u_A(r) = \frac{1}{\overline{w}}\sqrt{\frac{2}{\pi}}\,\mathbf{exp}\left(\frac{-r^2}{\overline{w}^2}\right), \qquad \textbf{(2.1.40)}$$

где $\overline{w}$ - радиус пучка в плоскости отверстия. В силу аксиальной симметрии задачи отличными от нуля будут только те коэффициенты разложения $C_{p0}=C_p$, азимутальный индекс которых равен нулю. Переходя к безразмерным параметрам $\kappa = \frac{z}{\overline{w}}$ и $\chi = \frac{\lambda}{\overline{w}}$ и используя выражения для ψ_{pl}, поперечное распределение интенсивности волны на расстоянии z от диафрагмы можно представить в виде (радиус пучка в плоскости диафрагмы считается минимальным, $\overline{w}=\overline{w}_0$)

$$I(x) = u_D u_D^* = $$

$$= \sum_{p,q} C_p C_q L_p\left(2x^2\right) L_q\left(2x^2\right)\mathbf{exp}\left(-2x^2\right)\mathbf{cos}\left(\theta_{pq}\right)\frac{2}{\pi w^2(z)}, \qquad \textbf{(2.1.41)}$$

где

$$x = \frac{z}{w}, \quad \theta_{pq} = 2\left(p-q\right)arctg\left(\frac{\kappa\chi}{\pi}\right),$$

$$w(z) = w_0\sqrt{1+\left(\frac{\kappa\chi}{\pi}\right)^2}.$$

Для улучшения сходимости ряда (2.1.41) целесообразно проводить разложение по модам TEM_{p0} с минимальным радиусом пучка, равным $w_0 = r_0$, если $\alpha<1$, и $\overline{w}=w_0$, если $\alpha \geq 1$,где $\alpha = \frac{r_0}{\overline{w}}$ - праметр ограничения пучка.

Расчеты по формуле (2.1.41)показывают, что приемлемую точность в расчетах можно достичь, не привлекая в разложении моды с индексами p и l,значения которых превышают 10.

2.2. Моды оптических резонаторов [1-4, 6-10]

В лазерных устройствах широко используются оптические резонаторы, состоящие чаще всего из двух плоскопараллельных или сферических зеркал. Оптические резонаторы обеспечивают необходимую для лазерной генерации положительную обратную связь и позволяют повысить плотность мощности светового поля до уровня, при котором происходит эффективный съем энергии с активной лазерной среды. Геометрия резонатора во многом определяет структуру лазерных пучков. В данном параграфе будут рассмотрены пространственные спектральные характеристики собственных типов колебаний (мод) оптических резонаторов, наиболее часто используемых в лазерной технике.

2.2.1. Геометрооптическое описание внутрирезонаторных полей

Как известно, традиционная схема лазера включает в себя двухзеркальный резонатор и располагающуюся в его внутренней полости активную среду. Рассмотрим симметричный лазерный резонатор со сферическими зеркалами (*рис. 2.2.1*). Луч, идущий вблизи оси резонатора, усиливается в активной среде и испытывает периодические отражения от зеркал. При каждом отражении луч частично проходит через зеркало и

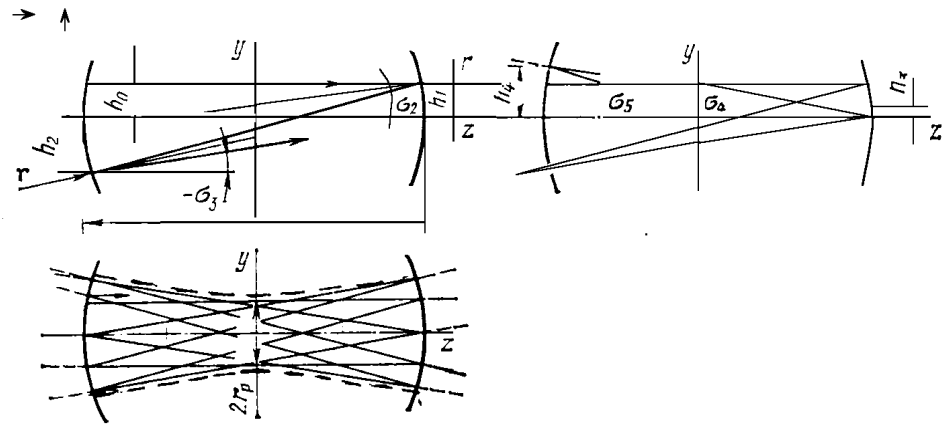


Рис. 2.2.1. *Ход луча в симметричном резонаторе; пунктир - каустика.*

покидает резонатор. Отраженный луч усиливается и при следующем отражении снова частично выходит из резонатора и так далее. Таким образом, пучок света, выходящий из лазерного резонатора, можно

представить в виде совокупности лучей (лучевого пакета), являющихся продолжениями первоначального луча после каждого его отражения от зеркал. Траекторию меридионального луча несложно определить, используя матричный метод, описанный в разделе 2.1.5. При использовании приведенных в указанном разделе соотношений надо воспользоваться тем, что с точки зрения воздействия на луч зеркало эквивалентно линзе с фокусным расстоянием $f=R/2$. Оценивая изменение координат луча путем последовательного применения матриц зеркала и свободного промежутка между зеркалами, можно показать, что поперечная координата h_k луча при k-м отражении, начиная с $k=0$, определяется соотношением:

$$h_k = h_0 \frac{\cos[(k+0.5)\omega]}{\cos(0.5\omega)} - \sigma_0 d_c \frac{\sin(k\omega)}{\sin\omega}, \qquad \textbf{(2.2.1)}$$

где h_0 и σ_0 - поперечные линейная и угловая координаты луча при первом отражении, $\cos\omega=g$ ($0<g<1$), где g - так называемый g-фактор резонатора, выражаемый через его длину d и радиус кривизны зеркала R посредством соотношения. Неравенство определяет условие устойчивости симметричного резонатора. Если оно нарушено, то луч после определенного числа отражений обязательно выйдет за границы апертуры зеркал. Такие резонаторы называются неустойчивыми.

В общем случае, когда ω не кратно целому числу π, точки, в которых происходит отражение луча на зеркалах резонатора, не повторяются. Луч при этом, несмотря на дискретный характер отражения, стремится полностью зачертить некоторую область внутри резонатора. Граница этой области является огибающей лучевого пакета (каустикой), т.е. описывает форму лазерного пучка в геометрооптической трактовке.

Приведенный лучевой анализ структуры поля в устойчивом симметричном резонаторе может быть распространен и на случай несимметричного резонатора с разными радиусами кривизны зеркал. Опуская получающиеся при этом соотношения, отметим лишь, что условие устойчивости резонатора с радиусами кривизны зеркал R_1 и R_2 имеет вид:

$$\mathbf{0<g_1g_2<1}, \qquad \textbf{(2.2.2)}$$

где $g_1=1-d/R_1$ и $g_2=1-d/R_2$ - g-факторы несимметричного резонатора. Лучевые модели описания внутрирезонаторных полей получили широкое распространение прежде всего в силу их простоты и наглядности. Однако не всегда при этом дается отчет в том, что указанные модели не удовлетворяют условиям выполнимости геометрооптического приближения. Некорректность лучевых моделей особенно отчетливо проявляется применительно к устойчивым резонаторам. В устойчивом

резонаторе лазерный пучок формируется в прикаустической зоне и значительные изменения амплитуды поля происходят на масштабе, меньшем первой зоны Френеля. Отступление от классической трактовки геометрической оптики приводит к тому, что распространяющиеся в резонаторе лучи оказываются неперпендикулярными волновому фронту. В этой связи в когерентной оптике сложилось направление, рассматривающее внутрирезонаторные лучевые модели в виде неких модификаций геометрической оптики. Разработанная в рамках данного направления теория [8,9] позволяет по определенным правилам восстанавливать амплитудно-фазовый профиль пучка внутри резонатора на основе структуры лучевых пакетов.

2.2.2. Моды устойчивых резонаторов в приближении бесконечной апертуры

В волновой трактовке свойства устойчивых резонаторов во многих случаях целесообразно рассматривать в приближении бесконечной апертуры зеркал. Такое приближение является оправданным, когда поле лазерного излучения концентрируется вблизи оси резонатора и его величина у краев зеркал пренебрежимо мала. Последнее обстоятельство освобождает от рассмотрения эффектов дифракции на внешней апертуре зеркал и позволяет свести моды устойчивых резонаторов к изученным уже в разделе 2.1 модам свободного пространства. В этом разделе было показано, что эрмито-гауссовые или лагерро-гауссовые моды свободного пространства имеют область наибольшего сужения (горловину) и расширяются от этой области в обоих направлениях. Типичный гауссов пучок показан на *рис. 2.2.2,а*. Сплошные линии характеризуют ширину пучка, а пунктирные линии, перпендикулярные оси, показывают фазовые фронты в различных точках вдоль пучка. Такой же пучок может существовать и внутри устойчивого резонатора, если зеркала поместить в тех местах, где радиусы кривизны фазового фронта пучка совпадают с радиусами кривизны (*рис. 2.2.2, б*).

При каждом отражении от зеркал пучок будет переходить сам в себя, что и обеспечит формирование моды резонатора. Поскольку все принадлежащие к одному семейству моды Эрмита-Гаусса и Лагерра-Гаусса характеризуются одними и теми же значениями радиуса кривизны волнового фронта, можно утверждать, что устойчивому сферическому резонатору можно поставить в соответствие целый набор собственных мод TEM_{mn}, различающихся поперечными индексами m и n. Их структура определяется выражениями (2.1.12, 2.1.13, 2.1.22, 2.1.26). Входящие в эти

выражения минимальный радиус гауссова пучка w_0 и длина волны λ могут быть найдены из условия резонанса, обеспечивающего интерференцию на

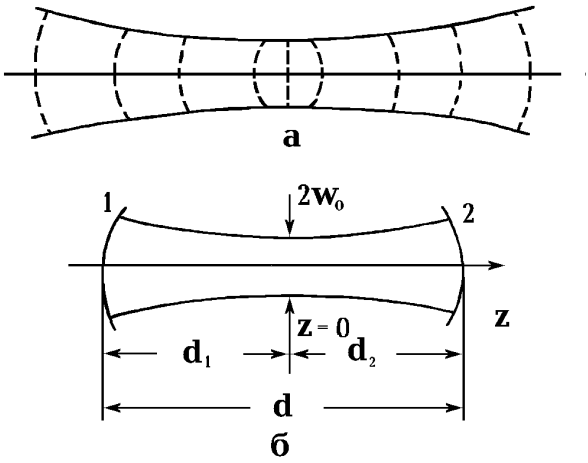


Рис. 2.2.2. *Гауссов пучок: а - в свободном пространстве; б - внутри резонатора; 1,2 - зеркала*

усиление лучей, отраженных от зеркал резонатора.

Условие резонанса будет выполняться, если фазовый сдвиг внутрирезонаторного пучка за двойной проход по резонатору будет равен целому числу, умноженному на 2π. Тем самым, при резонансе должно выполняться равенство

$$2\Delta\Phi = 2N\pi, \qquad \textbf{(2.2.3)}$$

где $\Delta\Phi$ - сдвиг фазы при прохождении волны от одного зеркала к другому, N - целое число. На оси резонатора (r=0) согласно (2.1.23)

$$\Delta\Phi = \Phi(d_2) - \Phi(d_1) = kd - (m+n+1)\left(arctg\frac{\lambda d_2}{\pi w_0^2} + arctg\frac{\lambda d_1}{\pi w_0^2} \right). \qquad \textbf{(2.2.4)}$$

Считается, что зеркала расположены на расстояниях d_1 и d_2 от области сужения пучка и $d=d_1+ d_2$ (см. *рис. 2.2.2,б*). С учетом (2.2.4) условие резонанса (2.2.3) принимает вид

$$2\pi\frac{d}{\lambda} - (m+n+1)\left(arctg\frac{\lambda d_2}{\pi w_0^2} + arctg\frac{\lambda d_1}{\pi w_0^2} \right) = N\pi. \qquad \textbf{(2.2.5)}$$

Длина волны λ в приведенных 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 быть определена как расстояние вдоль оси, соответствующее фазовому сдвигу 2π. Ее величина, однако, не является постоянной и зависит, хотя и незначительно, от положения вдоль оси моды Эрмито-Гаусса.

Условие резонанса (2.2.5) можно выразить непосредственно через параметры резонатора. Выразим сначала расстояния d_1 и d_2, радиусы пучков w_1 и w_2 на зеркалах 1 и 2, минимальный радиус пучка w_0 через радиусы кривизны зеркал R_1 и R_2 и расстояние d между ними. Используя зависимость (2.1.13), можно записать

$$R_1 = d_1 + \frac{\pi w_0^4}{\lambda^2 d_1}, \qquad \textbf{(2.2.6)}$$

$$R_2 = d_2 + \frac{\pi w_0^4}{\lambda^2 d_2}. \qquad \textbf{(2.2.7)}$$

Освобождаясь от величины w_0 в этих уравнениях, получаем

$$\left(R_1 - d_1\right)\Big/\left(R_2 - d_2\right) = d_2/d_1 . \qquad \textbf{(2.2.8)}$$

Присовокупив к (2.2.8) равенство

$$d_1 + d_2 = d , \qquad \textbf{(2.2.9)}$$

можно определить неизвестные параметры d_1 и d_2

$$d_1 = d\left(R_2 - d\right) / \left(R_1 + R_2 - 2d\right), \qquad \textbf{(2.2.10)}$$

$$d_2 = d\left(R_2 - d\right) / \left(R_1 + R_2 - 2d\right). \qquad \textbf{(2.2.11)}$$

Подстановкой (2.2.10) и (2.2.11) в (2.2.6) или (2.2.7) находим последний неизвестный параметр - минимальный радиус пучка w_0

$$w_0^4 = \left(\frac{\lambda}{\pi}\right)^2 \frac{d\left(R_1 - d\right)\left(R_2 - d\right)\left(R_1 + R_2 - d\right)}{\left(R_1 + R_2 - 2d\right)^2}. \qquad \textbf{(2.2.12)}$$

Используя выражения (2.2.10), (2.2.11) и (2.1.12), легко найти величину радиусов пучков на зеркалах резонатора

$$w_1^4 = \left(\frac{\lambda R_1}{\pi}\right)^2 \frac{d\left(R_2 - d\right)}{\left(R_1 - d\right)\left(R_1 + R_2 - d\right)}, \qquad \textbf{(2.2.13)}$$

$$w_2^4 = \left(\frac{\lambda R_2}{\pi}\right)^2 \frac{d\left(R_1 - d\right)}{\left(R_2 - d\right)\left(R_1 + R_2 - d\right)}. \qquad \textbf{(2.2.14)}$$

Вышеприведенные выражения позволяют условие резонанса (2.2.5) записать в форме, включающей в явном виде параметры резонатора. Воспользуемся известной тригонометрической формулой

$$arctg\alpha + arctg\beta = \arccos\left(\frac{1-\alpha\beta}{\sqrt{1+\alpha^2+\beta^2+\alpha^2\beta^2}}\right), \qquad \textbf{(2.2.15)}$$

где

$$\alpha = \frac{\lambda}{\pi}\frac{d_1}{w_0^2} = \sqrt{\frac{d\left(R_2 - d\right)}{\left(R_1 - d\right)\left(R_1 + R_2 - d\right)}}, \qquad \textbf{(2.2.16)}$$

$$\beta = \frac{\lambda}{\pi}\frac{d_2}{w_0^2} = \sqrt{\frac{d\left(R_1 - d\right)}{\left(R_2 - d\right)\left(R_1 + R_2 - d\right)}}. \qquad \textbf{(2.2.17)}$$

Путем алгебраических упрощений выражение (2.2.15) можно привести к виду

$$arctg\left(\frac{\lambda}{\pi}\frac{d_1}{w_0^2}\right) + arctg\left(\frac{\lambda}{\pi}\frac{d_2}{w_0^2}\right) = \arccos\sqrt{\left(1-\frac{d}{R_1}\right)\left(1-\frac{d}{R_2}\right)}. \qquad \textbf{(2.2.18)}$$

С учетом (2.2.18) условие резонанса (2.2.5) примет вид

$$\frac{2d}{\lambda} - \frac{m+n+1}{\pi}\arccos\sqrt{\left(1-\frac{d}{R_1}\right)\left(1-\frac{d}{R_2}\right)} = N. \qquad \textbf{(2.2.19)}$$

Выражение (2.2.19) позволяет определить резонансную частоту (собственную частоту резонатора)

$$\nu = \frac{c}{2d}\left\{N + \frac{m+n+1}{\pi}\arccos\sqrt{\left(1-\frac{d}{R_1}\right)\left(1-\frac{d}{R_2}\right)}\right\}. \qquad \textbf{(2.2.20)}$$

Заметим, что последнее выражение сочетается с условием устойчивости (2.2.1), поскольку член под знаком корня может быть только действительной величиной, а модуль корня должен быть меньше единицы. Число N определяет число полуволн, укладывающихся вдоль оси резонатора. Поэтому число N часто называют порядком продольных мод резонатора или продольным индексом. Индексы же m и n называют индексами резонаторных мод, поскольку они определяют число поперечных

вариаций поля эрмито-гауссовых мод. При $m=n=0$ имеет место чисто гауссова мода. Таким образом, чтобы охарактеризовать пространственную структуру собственной моды резонатора помимо параметров d_1, d_2, w_0, $w_{1,2}$ нужно обязательно задать один продольный N и два поперечных индекса m и n.

Если зафиксировать поперечные индексы m и n, то из (2.2.20) легко установить, что частотный интервал между соседними "продольными" модами (продольный индекс которых отличается на единицу) равен $c/2d$.

Если структуру поля в резонаторе описывать с помощью лагерро-гауссовых мод, то их параметры по-прежнему будут определяться выражениями (2.2.10) - (2.2.14). Резонансные же частоты будут определяться выражением

$$\nu = \frac{c}{2d}\left\{ N + \frac{2p+l+1}{\pi}\,\mathbf{arccos}\sqrt{\left(1-\frac{d}{R_1}\right)\left(1-\frac{d}{R_2}\right)}\right\}. \quad \textbf{(2.2.21)}$$

Для конфокального резонатора ($d=R_1=R_2=b$) параметры собственной моды принимают значения

$$d_1 = d_2 = \frac{b}{2},\ w_0^2 = \frac{\lambda b}{2\pi},\quad w_{1,2}^2 = \frac{\lambda b}{\pi}, \qquad \textbf{(2.2.22)}$$

собственная же частота моды будет определяться выражением

$$\nu = \frac{c}{2d}\left(N + \frac{1}{2}(m+n+1)\right). \qquad \textbf{(2.2.23)}$$

Параметр b называют конфокальным параметром. Самая важная особенность конфокального резонатора состоит в том, что в нем достигается высокая степень вырождения собственных мод: моды, имеющие различный набор индексов m, n, N могут иметь совпадающие частоты. Действительно, из (2.2.23) видно, что значение собственной частоты резонатора ν не изменится, если сумму поперечных индексов $m+n$ увеличить на целое число $2K$ (K=1, 2, 3...), а индекс N уменьшить на K. Как следует из (2.2.23), минимальный частотный интервал между четными и нечетными модами резонатора, сумма поперечных индексов которых $m+n$ является соответственно четной и нечетной, равен $c/4d$.

2.2.3. Согласование резонаторов

Установив зависимость между параметрами генерируемого в лазере светового пучка от параметров лазерного резонатора, рассмотрим важную в практическом отношении задачу согласования моды одного резонатора с модой другого посредством тонкой линзы. Такая задача часто возникает, когда выходной лазерный пучок подается в пассивный сферический

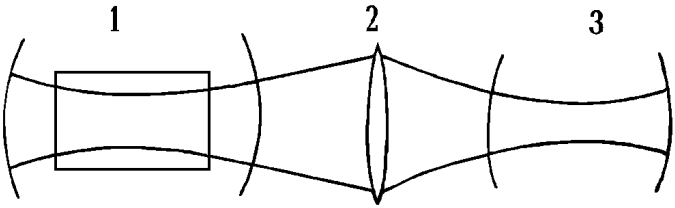


Рис. 2.2.3. *Согласование резонатора лазера с интерферометром*
1 - лазер; 2 - линза; 3 - интерферометр.

резонатор, выполняющий роль интерферометра, или в регенеративный оптический квантовый усилитель (невозбужденный лазер). Если не обеспечить согласование параметров лазерного пучка с параметрами собственной моды внешней системы, то возникает преобразование моды.

Рассмотрим согласование резонаторов лазера и интерферометра, когда известны положения горловин их собственных мод и минимальные радиусы пучков (*рис. 2.2.3*). Предположим, что минимальный радиус пучка лазера равен w_{10}, а расстояние от его горловины до линзы есть z_1. Соответствующие величины для собственного пучка интерферометра пусть будут равны w_{20} и z_2. В горловинах комплексные параметры пучков будут чисто мнимые и равны

$$q_1 = -i\pi w_{01}^2/\lambda,\ q_2 = -i\pi w_{02}^2/\lambda. \qquad \textbf{(2.2.24)}$$

На основании уравнений (2.1.11) и (2.1.28) получаем следующее соотношение

$$q_2 = \left[1/\left(q_1 + z\right) - 1/f\right]^{-1} + z_2. \qquad \textbf{(2.2.25)}$$

где f - фокусное расстояние линзы. Так как параметры пучка q_1 и q_2 являются чисто мнимыми, уравнение (2.2.25) можно разделить на вещественную и мнимую части

$$\left(z_1 - f\right)\left(z_2 - f\right) = f^2 + q_1 q_2, \qquad \textbf{(2.2.26)}$$

$$q_2(f-z_1)-q_1(f-z_2)=0. \qquad \textbf{(2.2.27)}$$

Найдем теперь, например, величину z_2 путем исключения z_1. При этом получим следующее квадратное уравнение:

$$z_2^2-2fz_2-q_2^2-f^2\left(\frac{q_2}{q_1}-1\right)=0. \qquad \textbf{(2.2.28)}$$

Решение уравнения (2.2.28) будет действительным, если

$$f>f_0, \text{ где } f_0=\left(|q_1|\cdot|q_2|\right)^{1/2}. \qquad \textbf{(2.2.29)}$$

Уравнения (2.2.26) и (2.2.27) могут быть представлены в виде

$$(f-z_1)(f-z_2)=(w_{10}/w_{20})^2, \qquad \textbf{(2.2.29a)}$$

$$(z_1-f)(z_2-f)=f^2-f_0^2. \qquad \textbf{(2.2.30)}$$

Обращает на себя внимание сходство уравнения (2.2.30) с известной в геометрической оптике формулой линзы Ньютона. Отличие состоит лишь в наличии члена $f_0{}^2$.

2.2.4. Моды резонаторов при ограниченной апертуре зеркал

С уменьшением размеров зеркал резонатора следует считаться с проявлениями эффектов дифракции. Последние приведут к потерям энергии на внешней апертуре зеркал. В этом случае мода резонатора представляет собой определенную конфигурацию медленно затухающего поля с относительным распределением амплитуды, не изменяющейся во времени. В случае лазерной генерации потери энергии компенсируются за счет активной среды, что обеспечивает существование стационарного поля.

Для расчета поля резонаторов с конечной апертурой зеркал может быть привлечен принцип Гюйгенса в формулировке Френеля-Кирхгофа (1.2.34). Если распределение поля на зеркале 1 (см. *рис. 2.2.2*) задать функцией u_1, то поле на зеркале 2 будет определяться дифракционным интегралом

$$u_2=\frac{-ik}{4\pi}\int_A u_1\frac{e^{ikr}}{r}(1+\cos\alpha)\,dS, \ \cos\gamma=1, \qquad \textbf{(2.2.31)}$$

где A - площадь каждого из зеркал. Отразившись от зеркала 2, световая волна начнет распространяться в обратном направлении. Таким образом, она будет распространяться в резонаторе вперед и назад, попеременно отражаясь от его зеркал. После q проходов связь поля у одного зеркала с

полем у другого зеркала будет опять же определяться выражением (2.2.31), где поле u_1 следует заменить на u_q, а u_2 на u_{q+1}.

После большого числа проходов распределение поля у зеркал будет подвергаться незначительным изменениям от отражения к отражению и со временем станет стационарным. На этой стадии поля около зеркал становятся одинаковыми с точностью до комплексной постоянной, что позволяет записать соотношение

$$u_q = \left(\frac{1}{\gamma}\right)^q v, \qquad \textbf{(2.2.32)}$$

где v - функция распределения, не изменяющаяся от отражения к отражению, и γ - комплексная постоянная, которая не зависит от координат. Подставляя (2.2.32) в (2.2.31), получаем интегральное уравнение

$$v = \gamma \int_A K v dS, \qquad \textbf{(2.2.33)}$$

где ядро

$$K = \frac{-ik}{4\pi R}(1+\cos\alpha)e^{ikR}.$$

Функцию распределения поля v можно рассматривать как моду резонатора. Дифракционные потери моды определяются выражением

$$\theta = 1 - |\gamma|^2, \qquad \textbf{(2.2.34)}$$

где θ - энергия, теряемая при одном прохождении из-за дифракции на зеркалах. Фазовые набеги моды связаны с величиной γ соотношением

$$\textbf{β=угол γ,} \qquad \textbf{(2.2.35)}$$

где β - фазовый сдвиг, претерпеваемый волной при распространении от одного зеркала к другому, в дополнение к фазовому сдвигу плоской волны, определяемому как $2\pi d/\lambda$. Если суммарные потери в резонаторе малы, добротность резонатора может быть представлена в виде

$$\theta = \frac{2\pi d}{\lambda \theta_t}, \qquad \textbf{(2.2.36)}$$

где θ_t - суммарные потери, включая дифракционные потери, потери на пропускание, поглощение и рассеяние. Резонансная частота будет определяться соотношением

$$\nu = \frac{c}{2d}\left(N + \frac{\pi}{\beta}\right). \qquad \textbf{(2.2.37)}$$

Интегральное уравнение (2.2.33) можно решить численным методом последовательных приближений, который во многом аналогичен

описанному выше физическому процессу возбуждения начального распределения поля световой волны в резонаторе и распространению его взад и вперед между зеркалами. Сначала на зеркале 1 задается произвольное распределение поля, а затем с помощью преобразования (2.2.31) последовательно находятся поля на зеркалах после каждого нового прохода. Расчет показывает, что после 300 отражений световой волны, флуктуации, наблюдающиеся от прохода к проходу, составляют менее 0.03 % от средней величины, тем самым распределение амплитуды и фазы после 300 проходов можно поставить в соответствие определенной моде резонатора.

На *рис. 2.2.4* приведено полученное в результате расчета относительное распределение поля моды TEM_{00} симметричного резонатора с круглыми зеркалами радиуса a. На *рис. 2.2.5* для такого же резонатора приведена зависимость дифракционных потерь моды TEM_{00} от числа Френеля $N_{фр}=a^2/\lambda d$ (число $N_{фр}$ равно числу зон Френеля, перекрываемых зеркалом резонатора при помещении точки наблюдения в центр противоположного зеркала). Цифрами у кривых обозначена величина g-фактора резонатора. Из *рис. 2.2.4 и 2.2.5* видно, что конечный

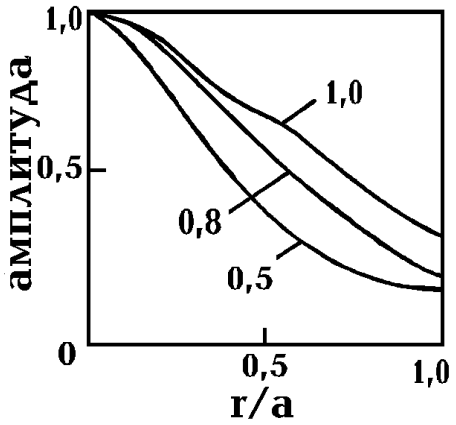


Рис. 2.2.4. *Относительное распределение поля моды* TEM_{00} ($N_{Фр}=1$).

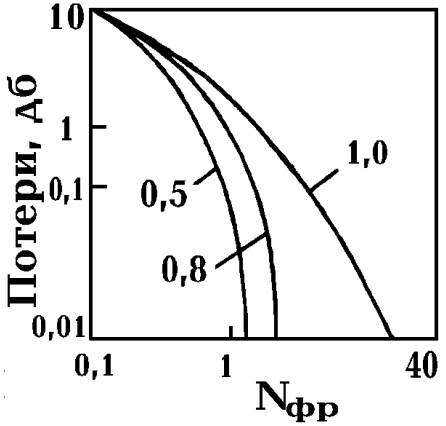


Рис. 2.2.5. *Относительные потери моды* TEM_{00} *на проход.*

размер зеркал приводит к заметной деформации формы моды по отношению к рассчитанной в рамках модели бесконечной апертуры. При этом имеет место быстрый рост дифракционных потерь моды с уменьшением числа Френеля. В области $N_{фр}\geq 10$ потери имеют

незначительную величину. При этом распределения полей собственных мод резонатора хорошо описываются эрмито-гауссовыми или лагерро-гауссовыми функциями.

2.2.5. Неустойчивые резонаторы

Несмотря на широкую область применений устойчивых резонаторов, они обладают одним очень серьезным недостатком. Он состоит в весьма малых поперечных размерах основной моды, что связано с фокусирующим действием лазерных зеркал. Так, при длине резонатора порядка 1м и для длины волны, лежащей в видимом диапазоне спектра, радиус пучка основной моды имеет порядок 1мм. В неустойчивых резонаторах, *g*-факторы которых меняются в областях $g_1g_2>1$ и $g_1g_2>0$, поле не фокусируется вблизи оси и с хорошим приближением распределение его амплитуды можно считать однородным, а волновой фронт сферическим. Однако, в случае неустойчивых резонаторов, возникает другая проблема, которая связана с тем, что лучи стремятся покинуть резонатор, увеличивая потери энергии. Тем не менее этот факт можно использовать даже в качестве преимущества, если эти лучи, уходящие из резонатора, включить в полезное выходное излучение лазера.

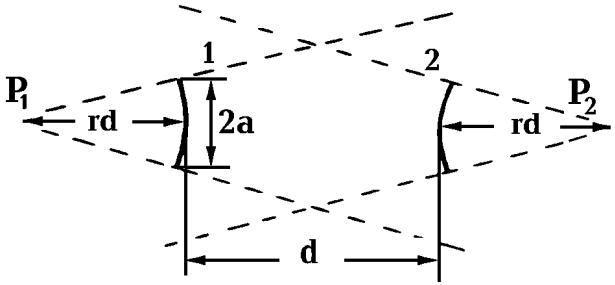


Рис. 2.2.6. *Симметричный двухторцовый неустойчивый резонатор*

Для описания полей в неустойчивых резонаторах в силу более медленного, чем в устойчивых зонаторах, поперечного изменения амплитуды и фазы вполне подходит геометрооптическое приближение. Рассмотрим симметричный двуторцовый неустойчивый резонатор (*рис. 2.2.6*). Как и прежде, будем предполагать, что мода образована суперпозицией двух сферических волн постоянной интенсивности. Центры P_1 и P_2, из которых исходят эти волны, не совпадают с центрами кривизны зеркал 1 и 2, но их координаты нетрудно вычислить, используя следующий

принцип самосогласования: сферическая волна, исходящая из точки P_1, после отражения от зеркала 2 должна давать сферическую волну, выходящую из точки P_2, и наоборот. Чисто геометрическое рассмотрение приводит к следующему выражению для показанной на *рис. 2.2.6* величины r:

$$r = \left\{ \left[g^2 \left(g^2 - 1 \right) \right]^{1/2} + g^2 - g \right\}^{-1}. \qquad \textbf{(2.2.38)}$$

Нетрудно видеть, что после того, как пучок пройдет от одного зеркала до другого, размер пятна от каждой сферической волны увеличивается в M раз, причем величина M определяется выражением

$$M = g + \left(g^2 - 1 \right)^{1/2}. \qquad \textbf{(2.2.39)}$$

Величину M называют однопроходным коэффициентом увеличения. Считая поперечное распределение освещенности однородным, потери за один проход можно записать в виде

$$\gamma = \frac{S_2 - S_1}{S_2} = \frac{M^2 - 1}{M^2}, \qquad \textbf{(2.2.40)}$$

где S_1 и S_2 - площади поперечного сечения пучка, исходящего из точки P_1, соответственно на зеркалах 1 и 2.

Резонатор, показанный на *рис. 2.2.6* , редко используется на практике. Гораздо шире применяются асимметричные конфокальные резонаторы. Одна из возможных конфигураций такого резонатора приведена на *рис. 2.2.7.*

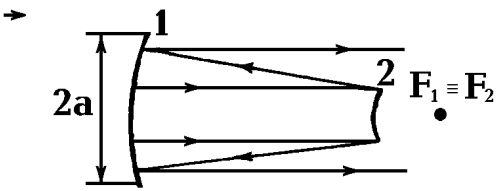


Рис. 2.2.7. *Конфокальный неустойчивый резонатор.*

Мода неустойчивого конфокального резонатора представляет собой суперпозицию сферической волны, исходящей из общего фокуса и плоской волны. Лучи последней, покидая резонатор, формируют на выходе узконаправленное излучение. Таким образом, помимо хорошего заполнения активного вещества излучением неустойчивые резонаторы обеспечивают малую угловую расходимость выходного излучения, приближающуюся к дифракционному пределу.

Вышеприведенное рассмотрение свойств оптических резонаторов основывалось на предположении, что активная среда, находящаяся внутри

резонатора, не вносит существенных искажений в структуру собственных мод. Накопленный в настоящее время обширный экспериментальный материал подтверждает справедливость этого предположения для широкого класса лазеров. Однако в ряде случаев, когда распределение показателя преломления и коэффициента усиления в сильной степени неоднородно, следует учитывать влияние активной среды. Поскольку неоднородности активной среды чаще всего носят квадратичный характер, для их учета следует владеть теорией распространения света в квадратичных средах. Основы этой теории приведены в разделе 2.3.1.

2.2.6. Резонаторы, применяемые для селекции мод

Как следует из разделов 2.2.2 и 2.2.3, в общем случае в лазерном резонаторе может возбуждаться большое число мод, отличающихся как продольными, так и поперечными индексами. Многомодовый характер генерации существенным образом усложняет пространственные характеристики и спектральные характеристики излучения. На практике обычно не составляет труда выделить одну поперечную моду. Так, основная мода TEM_{00} легко выделяется помещением в резонатор диафрагмы, размеры которой обеспечивают подавление всех высших мод. Однако даже в случае генерации лазера на основной моде в контур рабочей линии усиления попадает, как правило, большое число частот. Для того, чтобы улучшить монохроматичность излучения, необходимо проводить селекцию продольных мод, что представляет задачу гораздо более сложную. В настоящее время разработан целый ряд способов селекции продольных мод. Целесообразность применения каждого из них определяется конкретными свойствами лазерной среды и требованиями, предъявляемыми к спектральному составу генерации. Все существующие методы основаны на создании таких условий, когда минимальными потерями будут обладать моды, частоты которых располагаются в узком спектральном интервале. Эти моды будут присутствовать в спектре генерации, остальные же будут подавляться из-за недостатка усиления. Мы ограничимся качественным описанием некоторых резонаторных устройств, позволяющих селектировать продольные моды.

Связанные резонаторы. Уже на начальном этапе развития лазерной техники было обнаружено, что селектирующими свойствами обладает система из двух оптических связанных открытых резонаторов. Простейшей системой такого типа является трехзеркальная система, схема которой приведена на *рис. 2.2.8, а*. Зеркала 1 и 2 формируют основной резонатор с активным веществом 4, зеркала 2 и 3 - дополнительный. Выходное зеркало

1 и среднее зеркало 2 выполнены полупрозрачными. Возможность селекции частот в такой системе легко объяснима, если интерферометр, образованный зеркалами 2 и 3, рассматривать как единый отражатель, эффективный коэффициент отражения которого R^* зависит от частоты ν. Эта зависимость, фактически характеризующая отражательную способность интерферометра Фабри-Перо, хорошо известна и качественно приведена на *рис. 2.2.8, в*. Расстояние между соседними минимумами связано с длиной дополнительного резонатора d_g соотношением $\Delta\nu=c/2d_g$. Их же ширина увеличивается с ростом коэффициента пропускания зеркала 2. Генерация будет иметь место на частотах, соответствующих максимальным значениям R^*. Если в отсутствии зеркала 3 спектр генерации имел вид, показанный на *рис. 2.2.8, б*, то в трехзеркальной схеме происходит его сужение (*рис. 2.2.8, г*). Поскольку спектральный интервал, где R^* принимает высокие значения, довольно большой, селектирующая способность трехзеркальной системы невелика.

Существенно более хорошими селектирующими свойствами обладает так называемая система Смита, показанная на *рис. 2.2.9, а*.

Основной резонатор этой системы, образованный зеркалами 1, 2, содержит активное вещество 5. Дополнительный резонатор сформирован зеркалами 2, 3, а также полупрозрачной пластиной 4, располагающейся внутри основного резонатора. Как и в предыдущем случае, зеркала 2, 3 и пластину 4 можно рассматривать как единый отражатель с изменяющимися по частоте эффективным коэффициентом отражения R^*. Принципиальное отличие системы Смита от трехзеркальной системы состоит в том, что поведение R^* совпадает с максимумом внутрирезонаторной мощности. Зависимость $R^*(\nu)$ приведена на *рис. 2.2.9, в*. Так как область максимальных значений R^* сужается, то, как видно из *рис. 2.2.9, б* и *г*, селектирующая способность улучшается, и появляется возможность реализовать одночастотный режим генерации.

Резонаторы с поглощающей пленкой. Селективность резонаторов, содержащих помещаемую между зеркалами поглощающую пленку, обусловлена эффектом подавления тех продольных мод, пучности стоячих волн которых внутри резонатора совпадают с положением пленки. Этот эффект иллюстрирует *рис. 2.2.10*, где показана структура поля стоячих волн внутри резонатора. Пунктир соответствует подавляемой моде, непрерывные кривые характеризуют структуру стоячей селектируемой моды. Узел ее поля совпадает с пленкой, и в отличие от других продольных мод она будет испытывать минимальные потери.

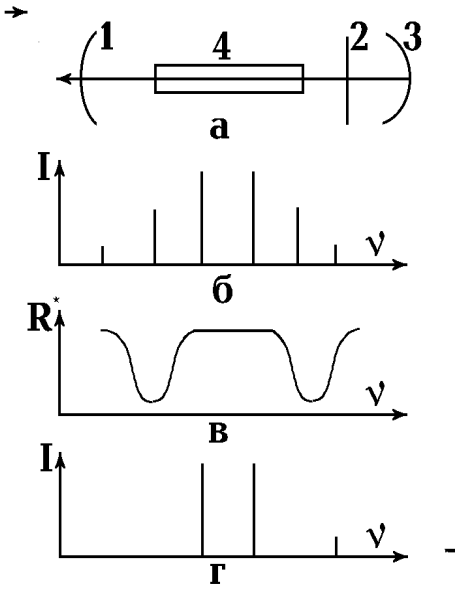


Рис. 2.2.8. *Трехзеркальная система связанных резонаторов: а - оптическая схема; б - спектр генерации без третьего зеркала; в - частотная зависимость эффективного коэффициента отражения; г - спектр генерации трехзеркальной системы; I - интенсивность излучения*

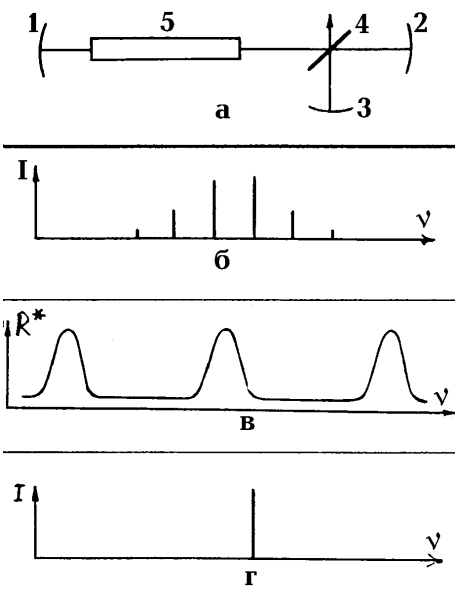


Рис. 2.2.9. *Система связанных резонаторов Смита: а-оптическая схема; б-спектр генерации в основном резонаторев-спектральная зависимость эффективного коэффициента отражения; г-спектр генерации в системе Смита.*

Резонаторы с дисперсионными элементами. Весьма эффективный способ сужения спектра лазерной генерации основан на использовании в лазерном резонаторе элементов, коэффициенты пропускания которых характеризуются резко выраженными зависимостями от частоты световых колебаний. Чаще всего в роли таких элементов используются дифракционные решетки и наклонные интерферометры Фабри-Перо. На *рис. 2.2.11*. Изображен лазер, резонатор которого содержит такие элементы. Дифракционная решетка в нем одновременно выполняет функции резонаторного зеркала. Наличие в резонаторе наклонного интерферометра позволяет дополнительно сузить спектральный интервал, выделяемый дифракционной решеткой. Достоинством приведенной схемы является возможность простым вращением решетки перестраивать частоту генерации лазера.

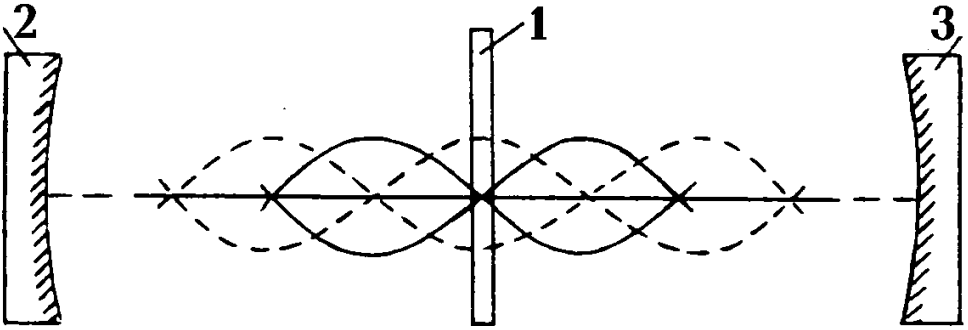


Рис. 2.2.10. *Подавление продольных мод поглощающей пленкой: 1-поглощающей пленка; пунктир - стоячая волна подавляемой моды; непрерывные кривые - стоячая волна селектируемой моды; 2,3 - зеркала резонатора*

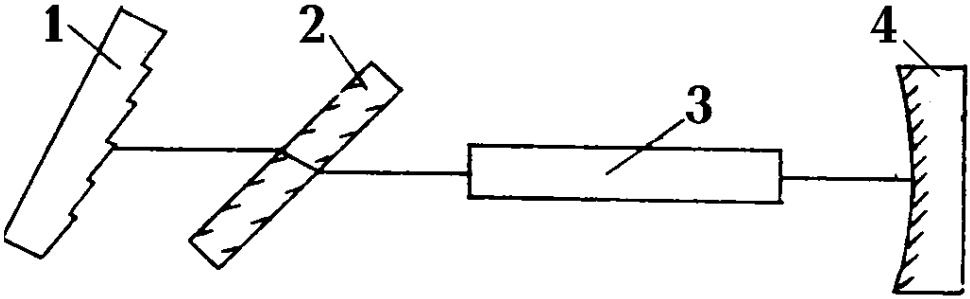


Рис. 2.2.11. *Резонатор лазера с селективными элементами: 1-дифракционная решетка; 2-наклонный интерферометр Фабри-Перо; 3-активный эемент; 4-отражающее зеркало.*

2.2.7. Кольцевые резонаторы

Все вышеперечисленные резонаторы объединены тем свойством, что в них формируются стоячие волны. Существуют, однако, такие резонаторы, энергия в которых накапливается в виде энергии бегущих волн. К ним, в первую очередь, следует отнести так называемые кольцевые резонаторы, подобные изображенным на *рис. 2.2.12*. Осевым контуром

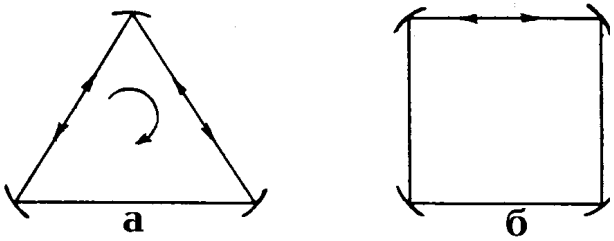


Рис. 2.2.12. *Кольцевые резонаторы:*
а - трехзеркальный;
б - четырехзеркальный.

кольцевого резонатора принято называть луч, который, испытав отражения от резонаторных зеркал, замыкается сам на себя.

Траектория этого луча представляет собой ломаную линию, лежащую в некоторой плоскости, называемой плоскостью контура резонатора. Вдоль контура резонатора, навстречу друг другу, распространяются две бегущие волны. Рассматривая движение гауссова пучка по контуру и приравнивая его фазовый набег за один оборот целому числу π, можно аналогично тому, как было сделано в разделе 2.2.2 для линейного резонатора, найти параметры пучка с самосогласованными параметрами, являющегося собственной модой резонатора. Единственной особенностью является астигматизм, проявляющийся при отражении пучка от сферического зеркала. Наличие астигматизма приводит к тому, что радиусы пучков в плоскости контура резонатора и в перпендикулярной ей плоскости оказываются различными. Расчет, выполненный для симметричного трехзеркального кольцевого резонатора (см. *рис. 2.2.12,а*), показывает, что перетяжки собственной моды резонатора расположены на серединах сторон треугольника. При этом минимальный радиус пучка в плоскости контура резонатора

$$w_0' = \sqrt[4]{\left(\sqrt{3}R - a\right)a / 4k^2}\,, \qquad \textbf{(2.2.41)}$$

а в перпендикулярной плоскости

$$w_0'' = \sqrt[4]{\left(4\sqrt{3}R - a\right)a / 4k^2}\,, \qquad \textbf{(2.2.42)}$$

где a - сторона треугольника, R - радиус кривизны зеркал. Спектр собственных частот такого резонатора определяется формулой

$$\nu = \frac{c}{a}\left[(2N-n)\frac{1}{6} + \frac{n+\frac{1}{2}}{2\pi}\,\mathbf{arccos}\left(1-\frac{2a}{R\sqrt{3}}\right) + \frac{m+\frac{1}{2}}{2\pi}\,\mathbf{arccos}\left(1-\frac{a\sqrt{3}}{2R}\right)\right].$$

(2.2.43)

Здесь n - определяет число нулей поля в плоскости контура, а m - в перпендикулярной плоскости.

Кольцевые резонаторы нашли широкое применение в лазерных гироскопах. Принцип работы последних основан на различии путей волн, распространяющихся по и против часовой стрелки, в случае вращения резонатора вокруг оси, перпендикулярной плоскости контура (*рис. 2.2.12, а*). Это различие является непосредственным следствием известного из специальной теории относительности эффекта сокращения линейных размеров тел при движении. Разница в оптических путях приводит к относительным сдвигам частот встречных волн и возникновению между ними биений. По частоте же биений можно судить об угловой скорости вращения гироскопа.

2.2.8. Модульные системы

Создание мощных лазеров требует использования больших объемов активного вещества. В то же время увеличение размеров активной зоны лазера, как правило, приводит к снижению качества излучения из-за расширения модового состава генерации и усугубляющегося негативного влияния неоднородностей лазерно-активной среды. Один из вариантов решения проблемы сочетания требований качества и мощности излучения состоит в использовании так называемой модульной лазерной системы.

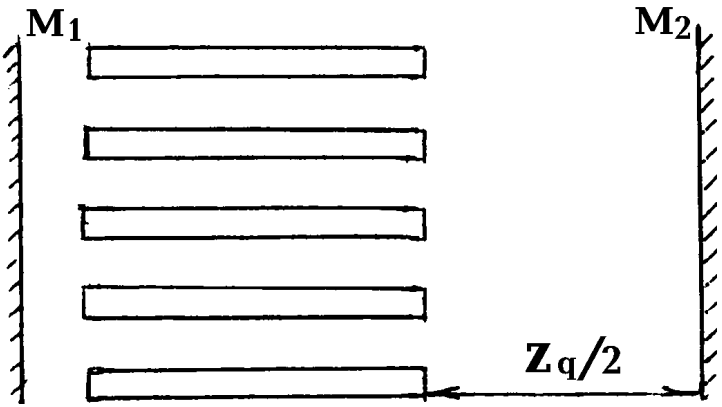


Рис. 2.2.13. *Лазерная модульная система.*

Активный модуль такой системы состоит из сборки малоапертурных активных элементов, количество которых может достигать нескольких десятков, а в некоторых случаях и нескольких сотен. На *рис. 2.2.13* приведена упрощенная двумерная схема сборки активных элементов (обычно применяется трехмерная, объемная сборка с существенно большими количеством элементов). Сборка помещена в плоскопараллельный резонатор, образованный зеркалами M_1 и M_2. В силу небольшой апертуры активных элементов через них без существенных потерь будут распространяться световые пучки, сформированные лишь низшими модами с невысокой угловой расходимостью. Если сборка включает N активных элементов, то направляемая на мишень мощность, выходящая из резонатора модульной системы, оказывается в N раз больше плотности мощности, получаемой от отдельного элемента. Этот результат соответствует случаю, когда из-за несфазированности световых колебаний в отдельных элементах, выходящее из модульной системы излучение является некогерентным.

Иная ситуация имеет место, если генерация в отдельных модулях сфазирована и излучение когерентно. В этом случае максимальная плотность мощности на мишени оказывается в N^2 раз больше, чем для отдельнного элемента. Для того, чтобы обеспечить последний, несомненно, более выгодный режим работы лазерной модульной системы, необходимо ввести оптическую связь между элементами сборки. Такую связь, в принципе, можно обеспечить, направляя с помощью системы дополнительных полупрозрачных зеркал часть излучения от каждого элемента в соседний. Однако, при большом числе элементов в сборке такой способ оказывается слишком громоздким и сложным в осуществлении.

Весьма эффективным оказался иной подход к осуществлению оптической связи, использующий эффект Талбота (см. п. 1.2.4). Этот эффект может быть использован, если расположение элементов в сборке будет носить строго регулярный, периодический характер. Такая сборка может рассматриваться в виде периодической решетки. Согласно эффекту Талбота, при когерентном излучении периодической решетки существует расстояние z_q, на котором структура решетки будет воспроизводиться. Отсюда ясно, что, располагая одно из плоских зеркал резонатора на расстоянии $z_q/2$ от сборки, можно обеспечить возвращение излучения, дифрагировавшего на торцах элементов, в каждый элемент. Необходимая оптическая связь между отдельными каналами модуля достигается при этом за счет дифракционного перераспределения световой энергии в возвращенном пучке. Поскольку режим генерации с самовоспроизводящимся по эффекту Талбота полем характеризуется большей добротностью резонатора, лазерная модульная система с резонатором, построенным по схеме *рис. 2.2.13*, будет генерировать когерентное излучение с высокой плотностью мощности в дальней зоне.

2.3. Волноводное распространение излучения [2,12-14]

Использованное в предыдущих разделах понятие моды является не совсем строгим, так как требование сохранения в процессе распространения формы амплитудно-фазового профиля не сочеталось с требованием неизменности поперечных размеров светового поля. Тем не менее, в некоторых оптических системах и средах возможно распространение волновых пучков, удовлетворяющих одновременно двум сформулированным выше требованиям. Такие волновые пучки представляют собой истинные моды с точки зрения их корректного определения. Если световое поле в оптической системе представимо в виде такого рода мод, то говорят о возможности волноводного распространения излучения.

2.3.1. Квадратичные среды

Анализ волноводного распространения излучения начнем с пучков в квадратичных средах. Квадратичной средой называется среда, показатель преломления которой меняется в поперечном направлении по квадратичному закону

$$n = n_0 - \frac{1}{2} n_1 \left(x^2 + y^2 \right). \qquad \textbf{(2.3.1)}$$

Квадратичная среда обладает волноводными свойствами, распространение в ней световых волн во многом сходно с распространением света в линзовом волноводе, состоящем из последовательности собирающих линз. Модель квадратичной среды широко используется как при анализе распространения излучения через лазерные активные элементы, так и при изучении распространения света в некоторых типах оптических волокон. Однако эта модель имеет один серьезный недостаток. Как видно из (2.3.1) при больших значениях поперечных координат x и y показатель преломления становится меньше единицы и даже достигает отрицательных значений. Модель квадратичной среды будет, тем самым, иметь смысл для пучков, основная часть энергии которых концентрируется вблизи оси и не выходит за пределы области, где $n>1$.

Для большинства реальных оптических сред, свойства которых близки к квадратичной модели (2.3.1), относительные изменения показателя преломления по длине световой волны малы и определяемый выражением (1.2.21) параметр $R<<1$. Поэтому для определения мод квадратичной среды можно воспользоваться приведенным уравнением в форме

$$\Delta\Psi + n^2 k^2 \Psi = 0, \qquad \textbf{(2.3.2)}$$

где k_0 выражается через длину волны λ в свободном пространстве $k_0 = 2\pi/\lambda$, а

$$n^2 = n_0^2 - n_0 n_1 \left(x^2 + y^2\right). \qquad \textbf{(2.3.3)}$$

Если считать исходным показатель преломления (2.3.1), то (2.3.3) является приближенным соотношением. Как и для свободного пространства, уравнение (2.3.2) решается с помощью пробной функции

$$\Psi = g\left(\frac{x}{w}\right) h\left(\frac{y}{w}\right) e^{i\beta z}. \qquad \textbf{(2.3.4)}$$

Опуская промежуточные выкладки, запишем сразу общий вид получающегося решения

$$\Psi_{pq}(x,y,z) = \frac{\sqrt{2/\pi}}{w\sqrt{2^{p+q}\,p!\,q!}} H_p\left(\sqrt{2}\,\frac{x}{w}\right)\times$$
$$\times H_q\left(\sqrt{2}\,\frac{y}{w}\right)\exp\left(-\frac{x^2+y^2}{w^2}\right)\exp\left(i\beta_{pq} z\right). \qquad \textbf{(2.3.5)}$$

Мы видим, что модовыми решениями для волн в квадратичной среде опять-таки являются функции Эрмита-Гаусса. Радиус пучка основной моды оказывается равным

$$w^2 = \frac{2}{k_0\sqrt{n_0 n_1}} \qquad \textbf{(2.3.6)}$$

и не зависит от продольной координаты z. Постоянная распространения β_{pq} будет определяться выражением

$$\beta_{pq} = \left\{n_0^2 k_0^2 - k_0\sqrt{n_0 n_1}\left[\left(2p+1\right)+\left(2q+1\right)\right]\right\}^{1/2}. \qquad \textbf{(2.3.7)}$$

Заметим, что в отличие от модового решения для свободного пространства, решение (2.3.5) является точным решением скалярного волнового уравнения. При этом моды среды, квадратичной по показателю преломления, всегда имеют плоский волновой фронт.

Рассмотренные моды являются модами квадратичной среды, которая не вносит ни затухания ни усиления. Такое предположение, естественно, является идеализацией. Чтобы учесть затухание или усиление, в дополнение к вышеприведенному анализу рассмотрим квадратичные среды с комплексной диэлектрической проницаемостью. Если по квадратичному закону изменяется лишь действительная часть показателя преломления, мы имеем ситуацию, когда направленность волн определяется

фокусирующими свойствами среды, обладающей на оси максимальным значением действительного показателя преломления. Применительно к средам с комплексной диэлектрической проницаемостью можно показать, что в тех случаях, когда усиление уменьшается с увеличением расстояния от оси, а потери, наоборот, возрастают при удалении от оси, в среде могут распространяться направленные моды, даже если действительная часть показателя преломления однородна или увеличивается при удалении от оси.

Рассматривая диэлектрическую среду с комплексным показателем преломления, нет необходимости вновь решать волновое уравнение. Выше приведенное решение справедливо также и для комплексных значений n_0 и n_1. Полиномы Эрмита являются аналитическими функциями даже при комплексных значениях своего аргумента.

Допустим, что n_0 имеет комплексное значение

$$n_0 = n_{0\mathrm{r}} + in_{0\mathrm{i}}. \qquad \textbf{(2.3.8)}$$

Положительные и отрицательные значения $n_{0\mathrm{i}}$ относятся к средам соответственно с потерями и усилением. В реальных физических средах всегда

$$\left|n_{0\mathrm{i}}\right| << n_{0\mathrm{r}}, \qquad \textbf{(2.3.9)}$$

так что можно записать

$$n_0^2 = n^2{}_{0r} + 2in_{0r}n_{0i}. \qquad \textbf{(2.3.10)}$$

Введем еще одну комплексную величину

$$\sqrt{n_0 n_1} = a + ib. \qquad \textbf{(2.3.11)}$$

Величина (2.3.6) является теперь комплексной. Ее смысл как квадрата радиуса пучка сохраняется для реальной части. Помня, что w является комплексной величиной, воспользуемся выражением (2.3.5) для описания моды в квадратичной среде с комплексной диэлектрической проницаемостью

$$\psi(x,y,z) = \frac{\sqrt{2}}{w\sqrt{\pi 2^{p+q}\,p!q!}} H_p\left(\sqrt{2}\,\frac{x}{w}\right) H_q\left(\sqrt{2}\,\frac{y}{w}\right) \times$$
$$\times \exp\left(-\frac{k_0}{2}ar^2\right)\exp\left(i\frac{k_0}{2}br^2\right)\exp\left(i\beta_{pq}z\right). \qquad \textbf{(2.3.12)}$$

Чтобы поле направленной моды спадало с ростом величины $r = \sqrt{\left(x^2 + y^2\right)}$, необходимо потребовать, чтобы выполнялось условие

$$a > 0. \qquad \textbf{(2.3.13)}$$

Из выражения (2.3.12), описывающего структуру моды, видно, что волновые фронты в среде с комплексным показателем преломления не

являются плоскими. Их форма является более сложной, поскольку они определяются не только экспоненциальными функциями, но и комплексными значениями полиномов Эрмита. Для основной моды $p=0,\quad q=0$ волновой фронт при положительных значениях b является вогнутым, если смотреть в положительном направлении оси z. Для того, чтобы установить связь величины b с показателем преломления, воспользуемся формулами (2.3.8) и (2.3.11) и запишем вещественную и мнимую части выражения (2.3.1)

$$\mathrm{Re}\, n = n_{0r} - \frac{1}{2n_{0r}}\left(a^2 - b^2\right)r^2, \qquad \textbf{(2.3.14)}$$

$$\mathrm{Im}\, n = n_{0i} + \frac{1}{n_{0r}}abr^2. \qquad \textbf{(2.3.15)}$$

Выражения (2.3.14) и (2.3.15) записаны в пренебрежении малыми членами с произведением $n_{0\mathrm{i}}$ на a^2, b^2 и ab.

Так как a - величина положительная, то положительное значение b означает, что с увеличением r в среде растут потери. Предположим, что величина $n_{0\mathrm{i}}$ отрицательная. Тогда среда усиливает на оси при r=0. Для положительного b усиление падает при удалении от оси. Поскольку при этом у лучей, перпендикулярных волновому фронту, появляется радиальная составляющая, направленная к периферии пучка, энергия, возникающая вблизи оси, будет уходить от нее и поглощаться.

Если поведение мнимой части показателя преломления полностью зависит от знака b, так как всегда $a>0$, то поведение действительной части показателя преломления зависит от того, больше a^2, чем b^2 или меньше. Re n уменьшается с ростом r, если $a^2 > b^2$, и увеличивается, если $a^2 < b^2$. Направленная волноводная мода может существовать как в первом, так и во втором случаях.

Заметим, что теоретически нельзя исключать возможность существования моды в среде с $b<0$, однако, анализ решения волнового уравнения (2.3.12) на устойчивость показывает, что при $b<0$ мода среды является неустойчивой, даже малые возмущения в системе могут привести к разрушению модовой структуры.

В заключение рассмотрим процесс затухания или усиления мод для случая комплексного показателя преломления. Постоянную распространения (2.3.7) можно разложить на действительную и мнимую части:

$$\mathrm{Re}\,\beta_{pq} = n_{0r}k_0 - \frac{a}{n_{0r}}\left(p+q+1\right), \qquad \textbf{(2.3.16)}$$

$$\mathrm{Im}\,\beta_{pq} = n_{0i}k_0 - \frac{b}{n_{0r}}\left(p+q+1\right). \qquad \textbf{(2.3.17)}$$

Затухание или усиление определяется мнимой частью β_{pq}. Моды усиливаются, если $I_m\beta_{pq}<0$, и испытывают поглощение, если $I_m\beta_{pq}>0$. Вопрос, усиливается или поглощается мода, зависит от трех факторов: знака n_{0i}, знака b и величины $p+q+1$.

2.3.2. Оптические волокна

Для передачи когерентного излучения в волноводном режиме между элементами различного рода оптических устройств или систем широко используются оптические волокна. Оптическое волокно *рис. 2.3.1* представляет собой внутреннюю диэлектрическую среду (стекло, кварц и т.п.), в которой содержится основная часть световой энергии, передаваемой по волокну в волноводном режиме. Эта внутренняя среда называется сердцевиной. Сердцевина может быть окружена слоем с более низким показателем преломления, называемым оболочкой. Для защиты от внешних воздействий сердцевину с оболочкой часто покрывают защитным слоем пластмассы. Обычно оптические волокна имеют круглую форму. Существует два основных типа круглых волокон. К первому типу относится волокно со скачком показателя преломления (*рис. 2.3.1, а*). В нем показатель преломления сердцевины характеризуется постоянным значением, и волноводное распространение излучения обеспечивается эффектом полного внутреннего отражения между сердцевиной и оболочкой. Второй тип волокон имеет сердцевину, показатель преломления которой изменяется в зависимости от расстояния r от оптической оси по параболическому закону

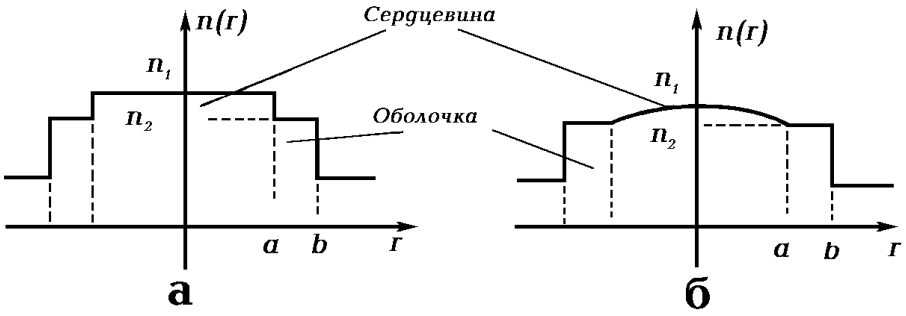


Рис. 2.3.1. *Скачкообразное (а) и градиентное (б) изменение показателя преломления в оптических волокнах.*

$$n(r) = n(0) - [n(0) - n(a)]\left(\frac{r}{a}\right)^2. \qquad \textbf{(2.3.18)}$$

(*рис. 2.3.1, б*) Волокна этого типа обычно называют волокнами с градиентом показателя преломления или градиентными волокнами.

Каждому типу волокон может быть поставлена в соответствие система волноводных мод. Отдельные моды или их суперпозиции определяют структуру излучения, распространяющегося по волокну. На градиентные волокна полностью переносится теория волноводных мод, изложенная в разделе 2.3.1 для сред, квадратичных по показателю преломления.

Что касается характеристик мод волокон со скачком показателя преломления, то их определяют на основе поиска решений приведенного волнового уравнения с постоянным коэффициентом для сердцевины и оболочки с последующей процедурой их "сшивания". Среди мод такого волокна выделим в качестве наиболее часто встречающихся на практике поперечные поляризованные моды типа EH, структура которых определяется выражениями:

при $r<a$

$$\begin{cases} E_x = 0; \quad E_y = AJ_\nu\left(u\dfrac{r}{a}\right)\Big|^{\cos\nu\varphi}_{\sin\nu\varphi}, \\ H_y = 0; \quad H_x = -An_1\sqrt{\dfrac{\varepsilon}{\mu}}J_\nu\left(u\dfrac{r}{a}\right)\Big|^{\cos\nu\varphi}_{\sin\nu\varphi}; \end{cases} \qquad \textbf{(2.3.19)}$$

при $r>a$

$$\begin{cases} E_x = 0; \quad E_y = A\dfrac{J_\nu(u)}{K_\nu(\omega)}K_\nu\left(\omega\dfrac{r}{a}\right)\Big|^{\cos\nu\varphi}_{\sin\nu\varphi}, \\ H_y = 0; \quad H_x = -An_1\sqrt{\dfrac{\varepsilon}{\mu}}\dfrac{J_\nu(u)}{K_\nu(\omega)}K_\nu\left(\omega\dfrac{r}{a}\right)\Big|^{\cos\nu\varphi}_{\sin\nu\varphi}. \end{cases} \qquad \textbf{(2.3.20)}$$

Здесь A - постоянная величина; E_x, E_y и H_x, H_y - поперечные составляющие электрического и магнитного полей; J_ν - функции Бесселя; K_ν - модифицированные функции Бесселя;

$$u = ka\sqrt{n_1^2 - 1}; \quad \omega = ka\sqrt{n_2^2 - 1}; \quad \nu = 0,1,2,\ldots$$

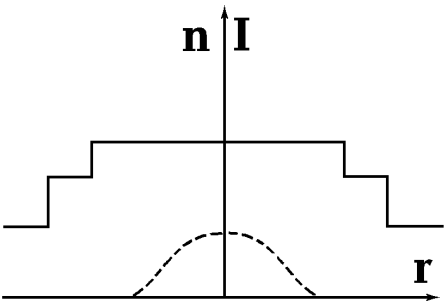


Рис. 2.3.2. *Амплитудный профиль низшей моды оптического волокна: непрерывные линии-профиль показателя преломления n; пунктир-распределение интенсивнсти I основной моды.*

Низшая из приведенных мод хорошо описывается гауссовой кривой. Это, в частности, видно из *рис. 2.3.2*, где приведен профиль интенсивности для низшей моды. Подобное распределение интенсивности обеспечивает хорошее ее согласование посредством линзы с основной модой лазера, излучение которого направляется в волокно.

2.3.3. Полые волноводы

В лазерной физике и технике широко используются газовые и жидкостные лазеры, в которых активная среда находится внутри полого диэлектрического волновода. Такие лазеры способны генерировать одночастотное излучение, перестраиваемое в широком спектральном интервале. Это делает их весьма ценным инструментом при решении многих задач когерентной оптики. Полым диэлектрическим волноводам, в роли которых часто выступают обычные стеклянные, кварцевые или керамические трубки, можно поставить в соответствие целую систему волноводных мод. Не останавливаясь подробно на выводе выражений для мод полого волновода, исходя из уравнения Гельмгольца, ограничимся краткой характеристикой их свойств.

Можно выделить три основных типа мод:

1) поперечные электрические моды TE_{0m}, в которых электрическое поле направлено по касательной к поверхности волновода;

2) поперечные магнитные моды TE_{0m}, в которых электрическое поле перпендикулярно к волноводной поверхности;

3) смешанные моды EH_{nm}, в которых электрическое поле имеет радиальную и тангенциальную составляющие.

В практическом отношении наиболее интересны смешанные гибридные EH_{nm} моды ($n\neq0$; $m\geq1$), так как именно эти моды возбуждаются в большинстве лазерных волноводов. Если показатель преломления стенок волновода $n_0<2.02$, то наименьшими потерями обладает волноводная мода EH_{11}, у которой распределение поля излучения описывается бесселевой функцией нулевого порядка $J_0(2.405r/a)$ (r - радиальная компонента, a -

радиус круглого волновода). Эта мода линейно поляризована и наилучшим образом связана с основной модой свободного пространства TEM_{00}.

Тангенциальная и радиальная составляющие электрического поля гибридных мод более высокого порядка даются выражениями

$$E_\theta = J_{n-1}(u_{nm}\, r/a)\cos(n\theta);$$
$$E_r = J_{n-1}(u_{nm}\, r/a)\sin(n\theta), \qquad (2.3.21)$$

где u_{nm} - m-ый корень уравнения $J_{n-1}(u_{nm}) = 0$, θ - азимутальная координата, $J_n(x)$ - функция Бесселя n-го порядка. Считается, что все рассматриваемые моды удовлетворяют соотношению

$$a >> \lambda |n_c| u_{nm}, \qquad \textbf{(2.3.22)}$$

где λ - длина волны излучения. Помимо моды EH_{11} линейно поляризованными будут моды EH_{1m}, а также суперпозиции мод $TE_{0m} + EH_{2m}$ и $EH_{n-1,m} + EH_{n+1,m}$.

Константа распространения гибридных мод круглого волновода имеет вид

$$\gamma = \frac{2\pi}{\lambda}\left[1 - \frac{1}{2}\left(\frac{u_{nm}\lambda}{2\pi a}\right)^2\left(1 - \frac{in_c'\lambda}{\pi a}\right)\right]. \qquad \textbf{(2.3.23)}$$

Набег фазы β_{nm} и коэффициент затухания α_{nm} даются действительной и мнимой частями последнего выражения:

$$\beta_{nm} = \mathrm{Re}\{\gamma\} = \frac{2\pi}{\lambda}\left\{1 - \frac{1}{2}\left(\frac{u_{nm}\lambda}{2\pi a}\right)^2\left[1 + \mathrm{Im}\left(\frac{n_c'}{\pi a}\right)\right]\right\}; \qquad \textbf{(2.3.24)}$$

$$\alpha_{nm} = \mathrm{Im}\{\gamma\} = \left(\frac{u_{nm}}{2\pi}\right)^2 \frac{\lambda^2}{a^3}\mathrm{Re}(n_c'), \qquad \textbf{(2.3.25)}$$

где

$$n_c' = \frac{1/2\left(n_c^2 + 1\right)}{n_c^2 - 1}. \qquad \textbf{(2.3.26)}$$

Теоретические оценки показывают, что моды полых диэлектрических волноводов имеют весьма малые потери. Так, мода EH_{11} с длиной λ=1 мкм распространяется в стеклянном волноводе с показателем преломления n_c=1.5 и внутренним радиусом α=1 мм испытывая потери, равные всего лишь 1.85 дБ/км. Однако потери в волноводе критичны по отношению к изгибам. Так, потери для моды EH_{11} удваиваются, если радиус кривизны волновода составит 10 км.

Если отношение λ/a пренебрежимо мало, то волновые фронты волноводных мод можно с большой степенью точности считать плоскими. В силу этого обстоятельства эффективная обратная связь в волноводном резонаторе может быть осуществлена с помощью плоских зеркал с высоким коэффициентом отражения, помещенных непосредственно на торцах волновода (одно из зеркал - выходное - обычно выполняется полупрозрачным). В такой системе при отражении от зеркала резонатора не происходит изменение модового состава излучения. Из условия резонанса, согласно которому вдоль волновода должно укладываться целое число полуволн, можно получить следующее выражение для собственных частот резонатора:

$$\nu_{mnq} = \frac{qc}{2L}\frac{\lambda_{mn}}{\lambda_0}, \qquad \textbf{(2.3.27)}$$

где q=0, 1, 2...., L - длина резонатора, c и λ_0 - соответственно скорость и длина волны света в свободном пространстве, $\lambda_{mn} = \dfrac{2\pi}{\beta_{nm}}$.

Распределение поля в сечении генерируемого таким лазером светового пучка будет определяться дифракцией волноводной моды на торце волновода. Профиль распределения интенсивности центрального дифракционного пятна очень близок к гауссовому, в результате чего 98% энергии волноводной моды переходит в моду свободного пространства TEM_{00}. При этом минимальный радиус пучка гауссовой моды w_0 связан с радиусом круглого волновода a соотношением

$$w_0 = 0.7032a \qquad \textbf{(2.3.28)}$$

(горловина TEM_{00} моды лежит в плоскости торца волновода).

2.4. Распространение когерентного излучения в среде со случайными неоднородностями [15-19]

Рассмотрим теперь распространение излучения в среде, показатель преломления которой изменяется случайным образом. При падении световой волны на такую случайную среду амплитуда и фаза волны претерпевают флуктуации, обусловленные флуктуациями показателя преломления среды.

2.4.1. Борновское приближение и приближение Рытова

В случайной среде относительная диэлектрическая проницаемость ε_r и показатель преломления n меняются от точки к точке и эти изменения нельзя предсказать; более того, даже если бы они и были известны, практически невозможно описать их значения во всех точках пространства. Поэтому среду необходимо описывать статистически и искать статистические закономерности поведения волны в такой среде. В соответствии с этим диэлектрическую проницаемость следует задавать как случайную функцию радиус-вектора **r**:

$$\varepsilon_r = \varepsilon_r(\mathbf{r}) = n^2(\mathbf{r}) \qquad \textbf{(2.4.1)}$$

Для неоднородного распределения показателя преломления приведенное волновое уравнение имеет вид

$$\nabla^2\psi(\mathbf{r}) + k_0^2 n^2 \psi(\mathbf{r}) - 2\nabla\left(\frac{\nabla n}{n}\cdot\psi(\mathbf{r})\right) = 0. \qquad \textbf{(2.4.2)}$$

Это уравнение непосредственно получается из уравнения (1.2.9) в предположении монохроматичности излучения. Показатель преломления n можно представить в виде суммы среднего значения $\langle n\rangle$ и флуктуаций n_1. Как уже указывалось в комментарии к уравнению (1.2.9) (см. соотношения (1.2.21)-(1.2.23)), последним членом в (2.4.2) можно пренебречь, если длина волны λ много меньше расстояния, на котором заметным образом меняется показатель преломления. Для волн оптического диапазона это допущение чаще всего выполняется, что позволяет придать волновому уравнению более простую форму:

$$\left(\nabla^2 + k_0^2 n^2\right)\psi(\mathbf{r}) = 0. \qquad \textbf{(2.4.3)}$$

Используя среднее волновое число $k^2 = k_0^2\langle n\rangle^2$, где k_0 - волновое число для вакуума, можно записать k_1

$$\left[\nabla^2 + k^2\left(1+n_1\right)^2\right]\psi(\mathbf{r}) = 0. \qquad \textbf{(2.4.4)}$$

Поставим задачу найти приближенное решение уравнения (2.4.4) для малых значений n_1. Это можно сделать двумя способами. Один из них основан на разложении в ряд самого поля:

$$\psi = \psi_0 + \psi_1 + \psi_2 + \ldots \qquad \textbf{(2.4.5)}$$

Нулевой член ряда описывает падающую или невозмущенную волну, слагаемые первого порядка - однократно рассеянное поле, второго порядка - двукратно рассеянное и т.д. Модель однократного и многократного рассеяния иллюстрирует *рис. 2.4.1*. Другой способ использует разложение в ряд показателя экспоненты:

$$\psi = \exp\left(\tilde{\varphi}_0 + \tilde{\varphi}_1 + \tilde{\varphi}_2 + \ldots\right) \qquad \textbf{(2.4.6)}$$

Разложение (2.4.5) представляет собой так называемое борновское приближение, а разложение (2.4.6) называется разложением Рытова. Рассмотрим первые приближения этих разложений.

Начнем с борновского приближения. Запишем уравнение (2.4.4) в виде

$$(\nabla^2 + k^2)\psi = -k^2 \delta n \psi, \qquad \textbf{(2.4.7)}$$

где

$$\delta n = \left(1 + n_1\right)^2 - 1 = 2n_1 + n_1^2.$$

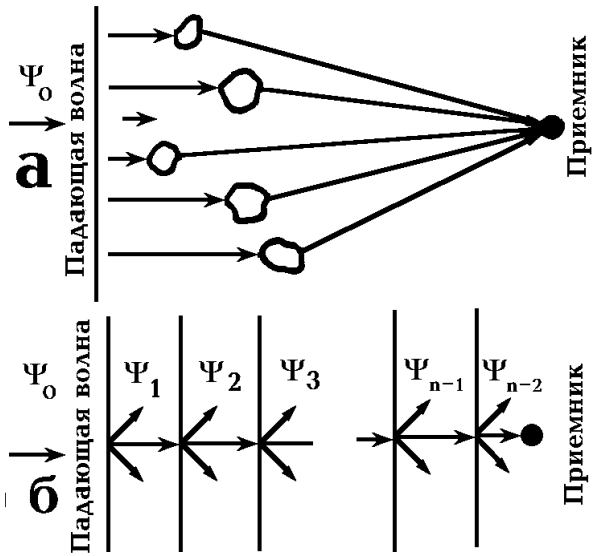


Рис. 2.4.1 *Модели однократного (а) и многократного (б) рассеяния*

Это уравнение можно свести к следующему интегральному уравнению для ψ:

(2.4.8)

где $\psi_0(\mathbf{r})$ - поле в отсутствие флуктуаций $(\delta n = 0)$, а

$$G(\mathbf{r}-\mathbf{r}') = \frac{\exp(ik|\mathbf{r}-\mathbf{r}'|)}{4\pi|\mathbf{r}-\mathbf{r}'|} - \qquad \textbf{(2.4.9)}$$

функция точечного источника.

Из уравнения (2.4.8) видно, что поле рассеянной волны, наблюдаемое в точке *r*, обусловлено сферической волной $\dfrac{\exp(ik|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|}$, исходящей из точки *r*'. Амплитуда сферической волны пропорциональна произведению амплитуды поля распространяющегося в среде излучения на флуктуирующую часть показателя преломления $n_1(r')$, а фаза определяется общим числом длин волн, укладывающихся на пути от источника до рассеивателя и далее до приемника.

Если подставить в интеграл ψ_0, то получится первая итерация ψ_1 борновского приближения. Продолжая последовательность итераций, можно получить выражение для ψ в виде ряда.

Рассмотрим теперь приближение Рытова. Поле $\psi(\mathbf{r})$ можно записать в виде

$$\psi(\mathbf{r}) = e^{\tilde{\varphi}(\mathbf{r})} \qquad \textbf{(2.4.10)}$$

и искать решение для $\tilde{\varphi}(\mathbf{r})$ в виде ряда. Этот подход, известный под названием метода Рытова, широко используется в задачах распространения волн в пределах прямой видимости. Имеется ряд теоретических и экспериментальных подтверждений того, что в задачах распространения в пределах прямой видимости первое приближение Рытова является более точным, чем борновское приближение.

Используя уравнение (2.4.7) и представление (2.4.10), а также полагая, что

$$\tilde{\varphi} = \tilde{\varphi}_0 + \tilde{\varphi}_1, \qquad \textbf{(2.4.11)}$$

можно получить следующее интегральное уравнение для $\tilde{\varphi}_1$:

$$\tilde{\varphi}_1(\mathbf{r}) = \frac{1}{\psi_0(\mathbf{r})}\int_{V'} G(|\mathbf{r}-\mathbf{r}'|)\left(\nabla\tilde{\varphi}_1\cdot\nabla\tilde{\varphi}_1 + k^2\delta n\right)\psi_0(\mathbf{r}')dV'. \qquad \textbf{(2.4.12)}$$

Уравнения типа (2.4.12) обычно решаются методом итераций с представлением решения в виде ряда. Упростим уравнение (2.4.12), полагая, что

$$\overline{\left|\nabla\tilde{\varphi}_1\right|} << k^2\delta n, \qquad \textbf{(2.4.12a)}$$

или

$$\overline{\left|\lambda\nabla\tilde{\varphi}_1\right|^2} << \delta n. \qquad \textbf{(2.4.12b)}$$

Это предположение означает, что изменения $\tilde{\varphi}_1$ на расстояниях порядка длины волны должны быть малыми по сравнению с δn и происходить достаточно плавно. Тем самым и метод, основанный на использовании допущения (2.4.12a,b), получил название метода плавных возмущений (МПВ). Таким образом, пренебрегая под знаком интеграла (2.4.12) произведением $\nabla\tilde{\varphi}_1 \cdot \nabla\tilde{\varphi}_1$, находим первую итерацию

$$\tilde{\varphi}_{10}(\mathbf{r}) = \frac{1}{\psi_0(\mathbf{r})}\int_{V'} G(\mathbf{r}-\mathbf{r}')\delta n(\mathbf{r}')\psi_0(\mathbf{r}')dV', \qquad \textbf{(2.4.13)}$$

которая представляет собой первое приближение Рытова и широко используется в теории слабых флуктуаций.

Итак, приближение Рытова представимо в виде:

$$\psi(\mathbf{r}) = \exp(\tilde{\varphi}_0 + \tilde{\varphi}_{10}) = \psi_0(\mathbf{r})\exp\left[\tilde{\varphi}_{10}(\mathbf{r})\right]. \qquad \textbf{(2.4.14)}$$

Представление $\tilde{\varphi}_1$ в виде итерационного ряда может быть найдено из следующего уравнения, получающегося из (2.4.12) и (2.4.13):

$$\tilde{\varphi}_1(\mathbf{r}) = \tilde{\varphi}_{10}(\mathbf{r}) + \frac{1}{\psi_0(\mathbf{r})}\int_{V'} G(\mathbf{r}-\mathbf{r}')\nabla\tilde{\varphi}_1(\mathbf{r}')\cdot\nabla\tilde{\varphi}_1(\mathbf{r}')\psi_0(\mathbf{r}')dV'. \qquad \textbf{(2.4.15)}$$

Подставляя $\tilde{\varphi}_{10}$ под знаком интеграла, приходим к следующей итерации. Продолжая эту процедуру, можно найти выражение для $\tilde{\varphi}_1$ в виде ряда.

2.4.2. Флуктуации уровня и фазы

Рассмотрим первое приближение Рытова для слабо неоднородной среды. В этом случае удобно использовать приближенное равенство

$$\delta n = 2n_1 + n_1^2 \approx 2n_1 \qquad \textbf{(2.4.16)}$$

и записать

$$\psi(\mathbf{r}) = \psi_0(\mathbf{r})\exp\left[\tilde{\varphi}_1(\mathbf{r})\right], \qquad \textbf{(2.4.17)}$$

$$\tilde{\varphi}_1(\mathbf{r}) = \int_{V'} h(\mathbf{r},\mathbf{r}')n_1(\mathbf{r}')dV', \qquad \textbf{(2.4.18)}$$

где

$$h(\mathbf{r},\mathbf{r}') = 2k^2 G(\mathbf{r}-\mathbf{r}')\frac{\psi_0(\mathbf{r}')}{\psi_0(\mathbf{r})}. \qquad \textbf{(2.4.19)}$$

Соотношение (2.4.18) устанавливает связь поля $\tilde{\varphi}_1(\mathbf{r})$ с флуктуациями показателя преломления $n_1(\mathbf{r})$.

Найдем теперь выражения для амплитуды A и фазы S поля $\psi(\mathbf{r})$. Полагая

$$\psi(\mathbf{r}) = A(\mathbf{r})\,e^{iS(r)},\ \psi_0(\mathbf{r}) = A_0(\mathbf{r})\exp\left[iS_0(\mathbf{r})\right] \qquad \textbf{(2.4.20)}$$

получаем

$$\tilde{\varphi}_1(\mathbf{r}) = \chi + iS_1 = \ln\left(\frac{A}{A_0}\right) + i\left(S - S_0\right). \qquad \textbf{(2.4.21)}$$

Вещественная часть $\tilde{\varphi}_1$, обозначенная через χ, представляет собой так называемые флуктуации уровня. Мнимая часть, обозначенная через S_1, описывает флуктуации фазы. Следует отметить, что при $|\chi| \prec\prec 1$ величина χ приближенно равна $\dfrac{A - A_0}{A_0}$.

Рассмотрим теперь распространение плоской волны, падающей на неоднородную среду с границей при z=0 (*рис. 2.4.2*). Точку наблюдения возьмем на расстоянии L от плоскости z=0. В случае плоской волны в отсутствие случайной среды падающая волна имеет вид

$$\psi_0(\mathbf{r}) = e^{ikz}. \qquad \textbf{(2.4.22)}$$

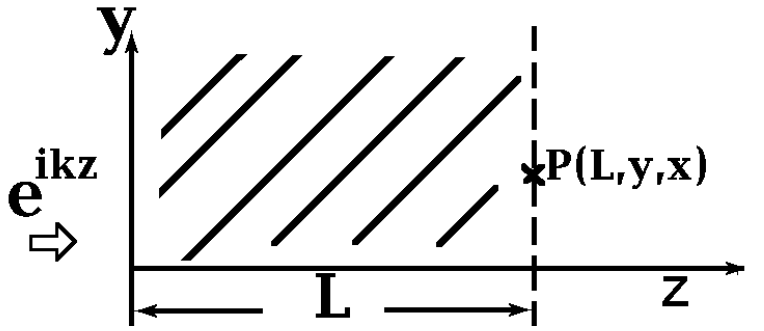


Рис.2.4.2. *Падение плоской волны на случайную среду [P(L,x,y)-точка наблюдения].*

Подставляя (2.4.22) в (2.4.18) и (2.4.19), получаем выражение для $\tilde{\varphi}_1$. Дальнейшие математические выкладки можно значительно упростить, если сделать несколько разумных предположений. Во-первых, предположим, что

обратное рассеяние пренебрежимо мало, так что интегрирование в (2.4.18) ограничено пределами от z=0 до z=L. Заметим также, что основной вклад в $\tilde{\varphi}_1(\mathbf{r})$ дает область $y' \approx y$ и $x' \approx x$ и, следовательно $|y'-y|$ и $|x'-x|$ предполагаются малыми по сравнению с расстоянием $|z-z'|$. Это заведомо верно при $\lambda < l_0$, где l_0 - характерный размер неоднородности, поскольку в этом случае рассеяние на неоднородности размером l_0 сосредоточено в направлении вперед внутри угла порядка λ / l_0.

В этих предположениях имеем

$$\tilde{\varphi}_1(L,y,x) = \int_0^L dz' \int_{-\infty}^{\infty} dy' \int_{-\infty}^{\infty} dx' h(L-z', y-y', z-z') n_1(x', y', x'), \tag{2.4.23}$$

где

$$h(L-z', y-y', x-x') = \frac{k^2}{2\pi} \frac{1}{(L-z')} \exp\left[i \frac{k}{2} \frac{(y'-y)^2 + (x'-x)^2}{L-z'} \right] \tag{2.4.24}$$

Это выражение получается путем приближенной замены $G(\mathbf{r}-\mathbf{r}')$ в (2.4.19) выражением

$$\frac{\exp(ik|\mathbf{r}-\mathbf{r}'|)}{4\pi|\mathbf{r}-\mathbf{r}'|} \approx$$

$$\approx \frac{1}{4\pi(z-z')} \exp\left[ik \left((z-z') + \frac{(y-y')^2 + (x-x')^2}{2(z-z')} \right) \right]. \tag{2.4.24}$$

Выражения (2.4.17) и (2.4.18) являются исходными для дальнейшего анализа флуктуаций амплитуды и фазы плоской волны.

Выпишем действительную и мнимую части (2.4.23)

$$\chi(\mathbf{r}) = \int h_r(\mathbf{r}-\mathbf{r}') n_1(\mathbf{r}') dV',$$
$$S_1(\mathbf{r}) = \int h_i(\mathbf{r}-\mathbf{r}') n_1(\mathbf{r}') dV' \tag{2.4.25}$$

и образуем корреляционные функции

$$B_\chi(\mathbf{r}_1, \mathbf{r}_2) = \langle \chi(\mathbf{r}_1) \chi(\mathbf{r}_2) \rangle,$$
$$B_s(\mathbf{r}_1, \mathbf{r}_2) = \langle S_1(\mathbf{r}_1) S_1(\mathbf{r}_2) \rangle. \tag{2.4.26}$$

В выражениях (2.4.26) угловыми скобками обозначена процедура усреднения по пространству. Корреляционные функции (2.4.26) могут быть представлены в виде:

$$B_\chi(\mathbf{r}_1,\mathbf{r}_2)=\int dV_1'\int dV_2'\,h_r(\mathbf{r}_1-\mathbf{r}_1')h_r(\mathbf{r}_2-\mathbf{r}_2')B_n(\mathbf{r}_1',\mathbf{r}_2'), \quad \textbf{(2.4.27)}$$

$$B_s(\mathbf{r}_1,\mathbf{r}_2)=\int dV_1'\int dV_2'\,h_i(\mathbf{r}_1-\mathbf{r}_1')h_r(\mathbf{r}_2-\mathbf{r}_2')B_n(\mathbf{r}_1',\mathbf{r}_2'), \quad \textbf{(2.4.28)}$$

$$\text{где}\quad B_n(\mathbf{r}_1',\mathbf{r}_2')=\left\langle n_1(\mathbf{r}_1')n_1(\mathbf{r}_2')\right\rangle -$$

корреляционная функция флуктуаций показателя преломления.

Вычисляя интеграл в (2.4.27), можно найти корреляционную функцию флуктуаций амплитуды $B\gamma$ для заданной B_n.

2.4.3. Характеристики излучения в случайной среде с гауссовой функцией корреляции

При конкретных расчетах следует считаться с тем, что интегралы (2.4.27) и (2.4.28) обычно очень трудно вычислить, и интегрирование оказывается возможным только для корреляционной функции B_n специального вида. В частности, расчеты можно осуществить в относительно простой форме, если использовать гауссову корреляционную функцию

$$\left\langle n_1(\mathbf{r}_1)n_1(\mathbf{r}_2)\right\rangle=\left\langle n_1^2\right\rangle\exp\left(-\frac{\left|\mathbf{r}_1-\mathbf{r}_2\right|^2}{l^2}\right). \quad \textbf{(2.4.29)}$$

Подобная аппроксимация корреляционной функции удачно описывает многие встречающиеся на практике случаи, кроме того, ее часто применяют, если точный вид корреляционной функции неизвестен. В формуле (2.4.29) параметр l следует рассматривать как характерный радиус корреляции флуктуаций показателя преломления. Подставляя (2.4.29) в (2.4.28) и (2.4.27), после ряда преобразований получаем при $L<<\dfrac{l^2}{\lambda}$

$$B_\chi(\rho)=\left\langle n_1^2\right\rangle\frac{8\sqrt{\pi}}{3}\left(\frac{L}{l}\right)^3\exp\left(-\frac{\rho^2}{l^2}\right)\left[1-2\frac{\rho^2}{l^2}+\frac{1}{2}\left(\frac{\rho^2}{l^2}\right)^2\right], \quad \textbf{(2.4.30)}$$

$$B_s(\rho)=\left\langle n_1^2\right\rangle k^2 Ll\sqrt{\pi}\exp\left(-\frac{\rho^2}{l^2}\right). \quad \textbf{(2.4.31)}$$

а при $L>>\dfrac{l^2}{\lambda}$

$$B_\chi(\rho) = B_s(\rho) = \langle n_1^2 \rangle \frac{1}{2} k^2 Ll\sqrt{\pi} \exp\left(-\frac{\rho^2}{l^2}\right). \qquad \textbf{(2.4.32)}$$

где ρ - величина смещения в перпендикулярном направлении. Формулы (2.4.30) и (2.4.31) относятся к случаю, когда радиус корреляции l намного превосходит радиус первой зоны Френеля $\sqrt{\lambda L}$, т.е. к области выполнимости геометрооптического приближения. В этой области дисперсии флуктуаций уровня и фазы соответственно равны

$$\sigma_\chi^2 = \langle n_1^2 \rangle \frac{8\sqrt{\pi}}{3}\left(\frac{L}{l}\right)^3, \qquad \textbf{(2.4.33)}$$

$$\sigma_s^2 = \langle n_1 \rangle^2 k^2 Ll\sqrt{\pi}, \qquad \textbf{(2.4.34)}$$

где $\sigma_\chi^2 = B_\chi(L,0)$ - дисперсия флуктуаций уровня, а $\sigma_s^2 = B_s(L,0)$ - дисперсия флуктуаций фазы.

Область $L >> \frac{l^2}{\lambda}$, для которой справедлива формула (2.4.32), является областью, где существенное проявление находят эффекты дифракции. В ней радиус корреляции l много меньше размера зоны Френеля $\sqrt{\lambda L}$. Для указанной области

$$\sigma_\chi^2 = \sigma_s^2 = \langle n_1 \rangle^2 \frac{1}{2} k^2 Ll\sqrt{\pi}. \qquad \textbf{(2.4.35)}$$

Соотношения (2.4.33)-(2.4.35) иллюстрируются на *рис. 2.4.3*.

Из рисунка видно, что на малых расстояниях изменения амплитуды несущественны, в то время как фаза заметно меняется с расстоянием, проходимым волной. По мере увеличения расстояния амплитудные флуктуации нарастают, и в конечном счете амплитудные и фазовые флуктуации становятся одинаковыми.

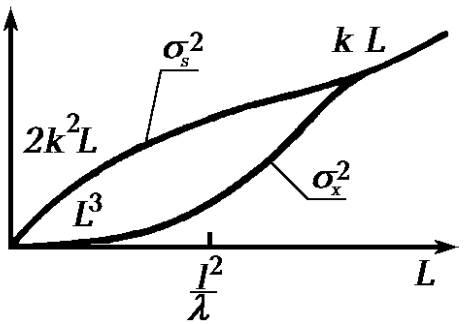


Рис.2.4.3. *Зависимость дисперсии уровня и дисперсии фазы от расстояния.*

Поведение коэффициентов корреляции изменения фазы и уровня

$$R_s = \frac{B_s(\rho, l)}{\sigma_s^2} \quad и \quad R_\chi = \frac{B_\chi(\rho, l)}{\sigma_\chi^2}$$

показано на *рис. 2.4.4*.

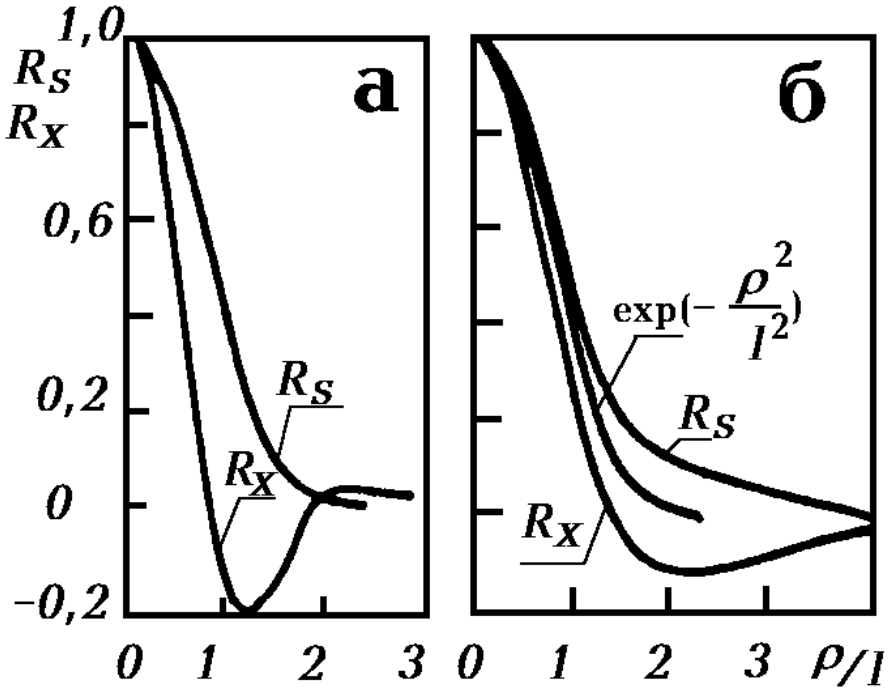


Рис.2.4.4 *Поведение коэффициентов корреляции изменения уровня и фазы в случае малых D (D<<1) (а) и в случае больших D (D=10) (б), $D=4L/kl^2$.*

Коэффициент R_S имеет гауссов вид и совпадает с коэффициентом корреляции для флуктуаций показателя преломления. Коэффициент R_χ для уровней отличается от гауссова. Однако и в этом случае корреляция между флуктуациями уровня простирается на расстояние того же порядка, что и корреляция между флуктуациями показателя преломления. Приведенные выше соотношения и комментарии к ним относятся к случаю слабых флуктуаций, когда при прохождении излучения через случайно неоднородную среду проявляются эффекты однократного рассеяния. Приближение однократного рассеяния остается справедливым до тех пор,

пока отношение величины флуктуаций к среднему уровню амплитуды и фазы не превысит значения 0.4÷0.5. При увеличении флуктуаций, происходящем с увеличением $\langle n_1^2 \rangle$ или L, формула (2.4.35) перестает быть справедливой. Так, она предсказывает увеличение дисперсии флуктуаций с тостом L до бесконечности, что лишено физического смысла.

Снять ограничения на уровень флуктуаций и уточнить выражения для статистических характеристик излучения позволяет учет многократного рассеяния на основе высших приближений в ряде (2.4.5). Поскольку теория, учитывающая высшие приближения, очень сложна, ограничимся качественной графической иллюстрацией некоторых ее результатов.

На *рис. 2.4.5* приведена зависимость нормированной дисперсии флуктуаций интенсивности

$$\sigma_I^2 = \frac{\langle (I - \langle I \rangle)^2 \rangle}{\langle I \rangle^2}$$

в условиях многократного рассеяния от аналогичной величины, вычисленной по методу Рытова.

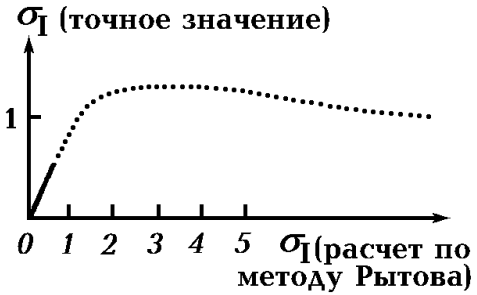


Рис.2.4.5. *Зависимость значений дисперсии флуктуаций интенсивности, полученных с учетом многократных рассеяний, от значений дисперсии, вычисленных по методу Рытова.*

Из рисунка видно, что в режиме сильных флуктуаций имеет место насыщение флуктуаций интенсивности, которые оказываются ограниченными определенным уровнем. Что касается поведения функции корреляции интенсивности, то в режиме сильных флуктуаций оно претерпевает значительные изменения.

На *рис. 2.4.6* приведены рассчитанные для плоской волны графики коэффициентов корреляции интенсивности, определяемых выражением

$$b_I(z_1, \rho_1, \rho_2) = \frac{\langle I_1 I_2 \rangle - \langle I_1 \rangle \langle I_2 \rangle}{\langle I_1 \rangle \langle I_2 \rangle}. \qquad \textbf{(2.4.36)}$$

Один из графиков относится к режиму слабых флуктуаций, другой-сильных.

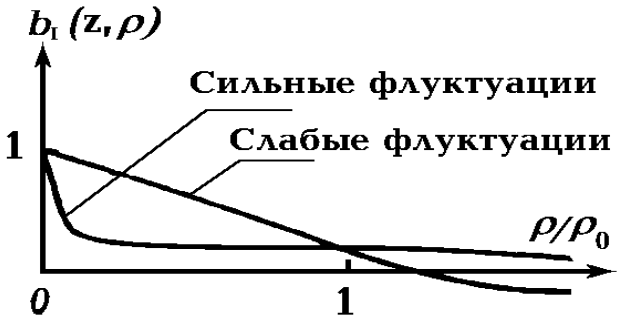


Рис.2.4.6. *Корреляционные функции интенсивности (ρ_0 - радиус корреляции для слабых флуктуаций).*

Как видно из рисунка, при сильных флуктуациях корреляция характеризуется малым радиусом корреляции и длинным хвостом. Резкое уменьшение радиуса корреляции при сильных флуктуациях приводит к тому, что распределение интенсивности в волне характеризуется многочисленными случайно расположенными в поперечном сечении световыми пятнами.

Таким образом, прохождение излучения через случайно неоднородную среду может привести к заметным пространственным флуктуациям в распределении амплитуды и фазы и значительному ухудшению пространственной когерентности.

2.4.4. Колмогоровская модель турбулентной атмосферы

Одним из наиболее важных с практической точки зрения случаев распространения когерентного света через случайно-неоднородную среду является его распространение через турбулентную атмосферу. Для корректного описания этого процесса гауссова модель часто оказывается слишком грубой. В основе представлений об атмосферной турбулентности лежит теория Колмогорова. Согласно этой теории турбулентные вихри, обусловливающие возникновение неоднородностей, можно характеризовать двумя масштабами: внешним масштабом турбулентности L_0 и внутренним масштабом l_0. Внешний масштаб турбулентности определяет размеры наиболее крупных вихрей, которые приобретают энергию за счет ветрового сноса и градиента температуры. Внутренний масштаб соответствует самым мелким вихрям с размерами порядка миллиметра, диссипация энергии

которых из-за вязкости превосходит кинетическую энергию. В так называемом инерционном интервале (L_0> размер вихря <l_0) происходит перманентный процесс зарождения крупных вихрей, их дробления, а затем и полного исчезновения в результате проявления эффектов вязкости.

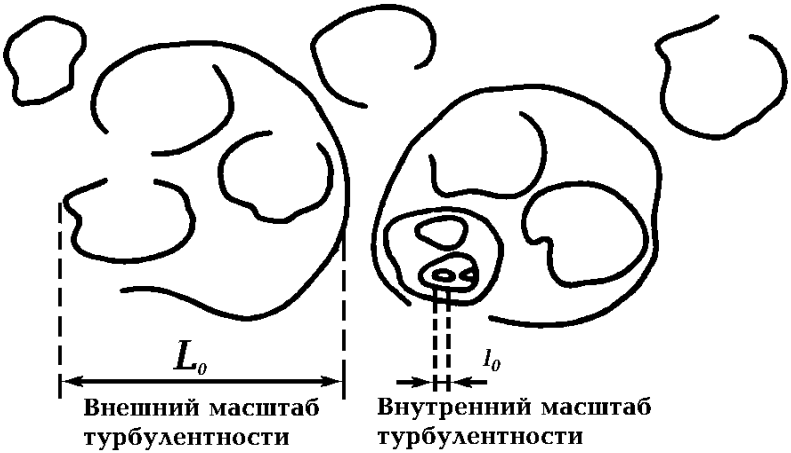


Рис.2.4.7. *Турбулентные вихри; показаны внешний и внутренний масштабы турбулентности.*

Структуру турбулентных образований в атмосфере иллюстрирует *рис. 2.4.7*.

Колмогоров установил, что структурная функция $D_{\text{n}}(\mathbf{r})$ флуктуаций показателя преломления в инерционном интервале определяется выражением

$$D_n(r) = \left\langle \left| n_1\left(\mathbf{r}+\mathbf{r}_1\right) - n_1\left(\mathbf{r}_1\right) \right|^2 \right\rangle = C_n^2 r^{2/3}, \qquad \textbf{(2.4.37)}$$

где C_n^2 - коэффициент, называемый структурной характеристикой флуктуаций показателя преломления. Коэффициент C_n^2 связан с дисперсией флуктуаций показателя преломления соотношением

$$C_n^2 = 2\left\langle n_1^2 \right\rangle L_0^{-2/3}. \qquad \textbf{(2.4.38)}$$

Выражение (2.4.37) известно, как "закон двух третей" Колмогорова. Используя стандартную связь между структурной и корреляционной функциями, можно перейти от выражения (2.4.37) к выражению для корреляционной функции флуктуаций показателя преломления

$$B_n(r) = \left\langle n_1 \right\rangle^2 - 0.5 D_n(r). \qquad \textbf{(2.4.39)}$$

На основе этого выражения, используя процедуру, аналогичную той, которая была использована выше для гауссовой корреляционной функции, можно рассчитать для рассматриваемой модели турбулентности корреляционные функции для уровня и фазы. Однако из-за сложности получающихся выражений мы их здесь приводить не будем. Приведем лишь вытекающие из них формулы для флуктуаций уровня χ и фазы $\varphi = S - S_0$ для дифракционной зоны.

$$\sigma_\chi^2 = 0.3k^{7/6}L^{11/6}C_n^2, \qquad \textbf{(2.4.40)}$$

$$\sigma_\varphi^2 = 0.4k^2 L L_0^{5/3} C_n^2. \qquad \textbf{(2.4.41)}$$

Из них видно, что флуктуации амплитуды с увеличением L нарастают быстрее.

До сих пор предполагалось, что флуктуации показателя преломления не зависят от времени. Такое описание атмосферной турбулентности является достаточным, если постоянная времени аппаратуры, регистрирующей параметры излучения, меньше характерного времени изменения показателя преломления. В тех случаях, когда обеспечить необходимое быстродействие приемной системы сложно, нужно учитывапть временные характеристики флуктуаций волны. Делается это обычно на основе гипотезы о "замороженной" турбулентности. Согласно этой гипотезе все временные изменения $n_1(\rho, t)$ обусловлены лишь перемещением (например, под действием ветра) пространственных возмущений с некоторой постоянной скоростью v_0, т.е.

$$n_1(\bar{\rho}, t) = n_1(\bar{\rho} - \bar{v}_0 t, 0). \qquad \textbf{(2.4.42)}$$

Наличие ветра приводит к зависимости от времени величин χ и φ. Если скорость ветра перпендикулярна к направлению распространения волны, то расчет, основанный на применении модели (2.4.42), дает следующие зависимости для спектральной плотности флуктуаций χ и φ:

$$w_\chi(\omega) \sim C_n^2 k^{2/3} L^{7/3} / v_0, \quad \omega \to 0, \qquad \textbf{(2.4.43)}$$

$$w_\chi(\omega) \sim C_n^2 k^{2/3} L^{7/6} (\omega / \omega_t)^{-8/3} / v_0, \quad \omega \to \infty, \qquad \textbf{(2.4.44)}$$

$$w_\varphi(\omega) \sim C_n^2 k^{2/3} L^{7/3} (\omega / \omega_t)^{-8/3} / v_0, \quad \omega \to 0. \qquad \textbf{(2.4.45)}$$

При частотах $\omega < \omega_t = \sqrt{\dfrac{k v_0^2}{L}}$ спектр флуктуаций постоянен, а при $\omega > \omega_t$ убывает по закону $\omega^{-8/3}$. Асимптотическое выражение для спектральной плотности флуктуаций φ при $\omega \to \infty$ совпадает с (2.4.44).

2.5. Формирование спекл-полей при взаимодействии света с диффузными объектами [20-22]

Когда наблюдатель рассматривает или фотографирует в когерентном свете диффузно отражающий или пропускающий объект, то структура регистрируемого излучения кажется ему зернистой. Создается впечатление, что она покрыта множеством мелких, хаотически расположенных светлых и темных пятнышек - так называемых спеклов. Поля с подобной структурой называют спекловыми или спекл-полями.

2.5.1. Физическая природа спеклов и их размеры

Физическая природа спеклов очень проста. Они являются результатом интерференции многих световых волн от различных точек объекта. Если предположить, что спекл-поле формируется в результате равномерного о.свещения диффузора (например, матового стекла) шириной

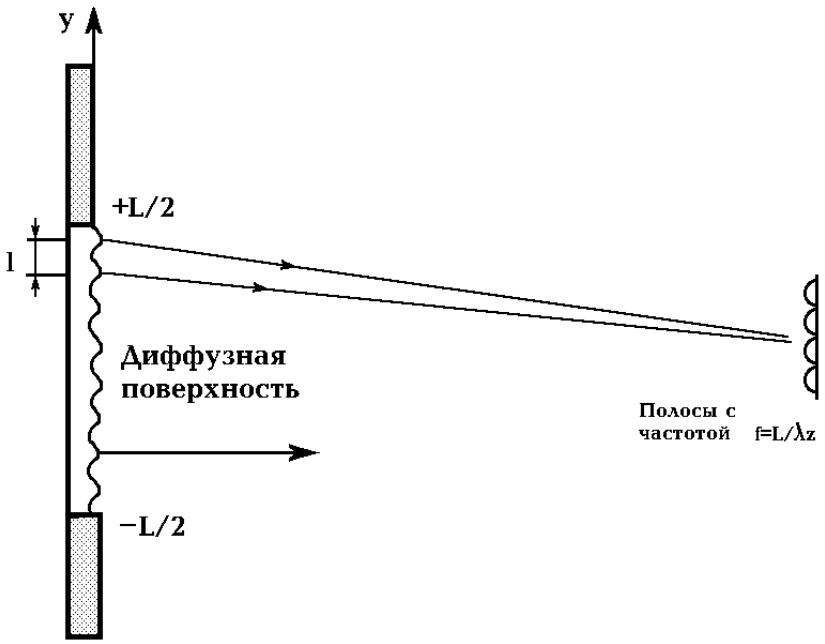


Рис.2.5.1.*Образование спеклов. Свет, исходящий из двух точек диффузора, разделенных расстоянием l, дает на экране интерференционные полосы с частотой* $f = \frac{l}{\lambda z}$.

L, то размер спеклов можно оценить из следующих соображений.

Для простоты рассмотрим зависимость интенсивности только от координаты y. Спекл-структура, наблюдаемая в плоскости на расстоянии z от диффузора, представляет собой суперпозицию интерференционных картин, возникающих при рассеянии света каждой парой точек на диффузоре. Две любые точки, разделенные расстоянием l, дают интерференционные полосы с частотой $f = \frac{l}{\lambda z}$. Наиболее тонкие полосы, т.е. полосы с наибольшей пространственной частотой $f_{\max} = \frac{L}{\lambda z}$ будут образованы крайними точками диффузора. Для меньших расстояний l существует большое количество пар точек, дающих полосы с частотой, определяемой расстоянием между ними. Число пар таких точек, разделенных расстоянием l, пропорционально L-l. Различные интерференционные полосы будут иметь случайные по отношению друг к другу фазы, поэтому при формировании усредненной по ансамблю освещенности вклад от интерференционных картин с разной частотой полос пропорционален соответствующему числу пар рассеивающих точек. Поскольку число последних пропорционально разности L-l, которая в свою очередь пропорциональна $f_{\max}$ -f, распределение освещенности по частоте полос будет линейным. Средняя частота полос равна

$$\langle f \rangle = \frac{1}{3} f_{\max} = \frac{1}{3}\frac{L}{\lambda z}, \qquad \textbf{(2.5.1)}$$

и, следовательно, распределение освещенности в "типичном спекле" запишется следующим образом:

$$I(y) = 1 + \cos\left(2\pi \frac{Ly}{3\lambda z}\right). \qquad \textbf{(2.5.2)}$$

За ширину спекла принимают расстояние между точками, где I падает до половины своего максимального значения, т.е. $1.5\left(\frac{\lambda z}{L}\right)$. Итак, можно считать, что размер типичного спекла (или, что одно и то же, средний размер спекла) равен

$$b_s \approx 1{,}5\frac{\lambda z}{L}. \qquad \textbf{(2.5.3)}$$

2.5.2. Спекл-фотография и спекл-интерферометрия

В оптике когерентного излучения очень часто спеклы рассматриваются как оптический шум, который приводит к ухудшению

качества изображения и снижению четкости интерференционной картины. Однако с помощью специальных методов это явление может быть успешно использовано для создания основ измерительной техники нового типа.

Спекл-фотография - это метод измерения плоских перемещений, деформаций, поворотов и вибраций, обладающий умеренной чувствительностью. Чтобы дать основные представления о спекл-фотографии рассмотрим схему измерения плоского перемещения, которая показана на *рис. 2.5.2,а*.

Линза с фокусным расстоянием f и диаметром D образует изображение поверхности объекта в плоскости фотослоя. Расстояние до объекта l_0 и до изображения l_i связаны уравнением

$$\frac{1}{l_0}+\frac{1}{l_i}=\frac{1}{f}. \qquad \textbf{(2.5.4)}$$

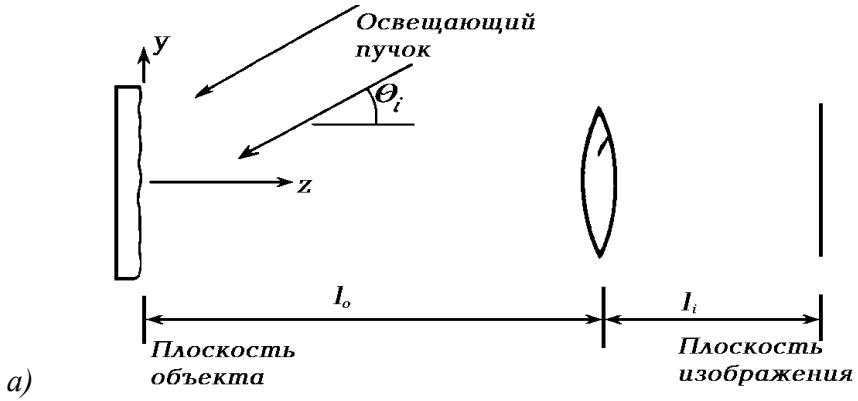
Образованное в плоскости фотослоя изображение промодулировано случайной картиной спеклов, имеющих характерный размер b_S, определяемый апертурой линзы:

$$b_s \approx 1.22\lambda\frac{l_i}{D}. \qquad \textbf{(2.5.5)}$$

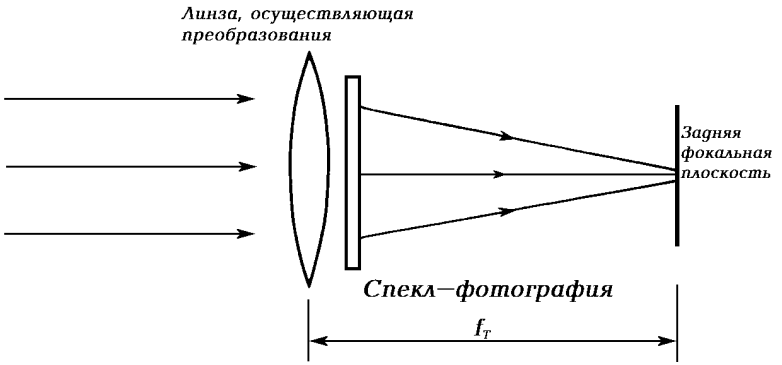
Если объект сместится в вертикальном направлении на величину L_Y, относительная фаза для каждого из множества лучей, участвующих в образовании каждого спекла, останется неизменной.

Следовательно, картина спеклов просто сместится в плоскости фотопластинки как целое на величину ML_Y, где M - поперечное увеличение оптической системы. Аналогично спеклы сместятся на величину ML_X, если объект переместится в горизонтальном направлении на L_X. Перемещение картины спеклов для таких перемещений в плоскости не зависит от угла освещения θ_i.

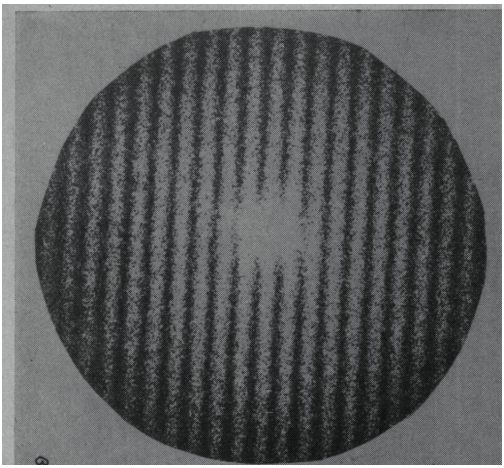
Чтобы измерить плоское перемещение объекта, пластинку экспонируют дважды - один раз до перемещения и один раз после него. Если предположить, что величина смещения L превышает размер спеклов b_S, то на проявленной фотопластинке получается фотография пары идентичных спекл-картин, смещенных на расстояние ML. Расстояние ML для каждой пары спеклов можно непосредственно измерить путем микроскопического исследования пластинки. Альтернативным способом является когерентно-оптическая обработка фотоснимка, в результате которой смещение может быть представлено в виде картины интерференционных полос.



а)



б)

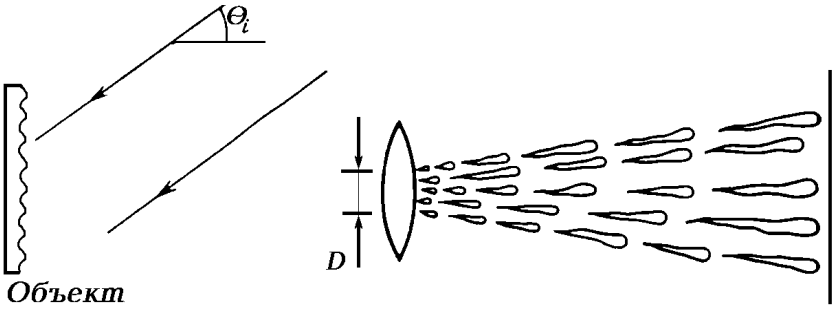


в)

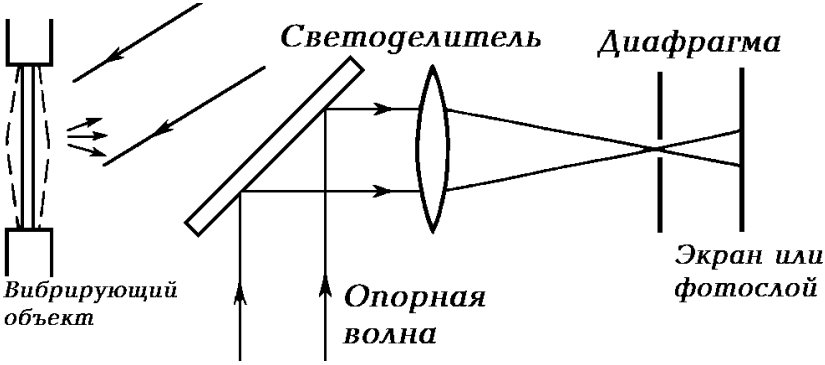
Рис.2.5.2. *Измерение смещения в плоскости с помощью двухэкспозиционной спекл-фотографии: а - схема регистрации; б - схема оптической обработки; в - картина полос, образующихся в задней фокальной плоскости оптической системы, изображенной на рис.б.*

Пластинку помещают в сходящийся лазерный пучок, образованный линзой с фокусным расстоянием f_T, как показано на *рис. 2.5.2,б*.

Распределение освещенности в задней фокальной плоскости линзы состоит из яркого центрального пятна, окруженного картиной спеклов, промодулированной полосами с косинусоидальным распределением интенсивности. Яркое центральное пятно образовано недифрагированным светом, прошедшим через фотопластинку; модулированная картина спеклов образуется светом, дифрагирующим на спекл-структуре, зарегистрированной на фотопластинке. Полосы с косинусоидальным распределением образуются в результате того, что каждая пара соответственных спеклов действует как пара идентичных источников когерентного света, которые образуют полосы Юнга (*рис.* 2.5.3,в).



а)



б)

Рис.2.5.3. *Измерение вибраций с помощью спекл-интерферометрии: а - схематическое представление продольной структуры спеклов; б - схема спекл-интерферометра для наблюдения вибраций поверхности.*

Ориентация полос нормальна к вектору **L** перемещения в плоскости. Величину **L** можно определить с помощью уравнения

$$I = I_0 \cos^2\left[\frac{\pi b}{\lambda}\left(\frac{y}{l_0}+\frac{y_s}{l_s}\right)\right]. \qquad \textbf{(2.5.6)}$$

Согласно которому если спеклы в каждой паре на фотоснимке разделены промежутком d_s, расстояние между полосами составляет $d_f=\lambda f_T/d_S$. Поэтому перемещение объекта в плоскости равно

$$L=\frac{\lambda f_T}{M d_f}, \qquad \textbf{(2.5.7)}$$

где λ - длина волны лазерного излучения, используемого для образования полос; f_T - фокусное расстояние линзы, осуществляющей преобразование; M - увеличение оптической системы, использованной при получении спекл-фотографии; d_f - расстояние между полосами.

В отличие от спекл-фотографии спекл-интерферометрия представляет собой класс измерительных методов с существенно более высокой чувствительностью, в которых происходит когерентное сложение (интерференция) поля, имеющего спекл-структуру, с плоской опорной волной или с другим полем, имеющим спекл-структуру.

Здесь мы ограничимся рассмотрением метода с использованием опорной волны. Предположим, что лазерный свет рассеивается шероховатой поверхностью в направлении экрана или фотопленки. Характерный размер спеклов при этом равен $b_s \approx 1.22\lambda\frac{\lambda z}{D}$, где z - расстояние от объекта до плоскости наблюдения и D - диаметр объекта. Если для получения изображения используется линза, то $b_s \approx 1.22\lambda\frac{\lambda f}{D}$, где D/f - относительное отверстие линзы. Если на диффузный свет наложить плоскую когерентную волну, интенсивность которой равна средней интенсивности спекл-картины, то это приведет к очень существенным изменениям в поведении картины спеклов при перемещении объекта в направлении к наблюдателю или от него. На основе статистических исследований продольной структуры спеклов, спеклы можно представить как образования, имеющие вытянутую структуру вправо от линзы, строящей изображение (*рис. 2.5.4*). Следовательно, если объект движется вдоль оси, то изменения структуры спекл-картины в плоскости изображения будут незначительны. Другими словами, небольшое смещение в осевом

направлении не приводит к изменению относительной фазы световых лучей, рассеянных отдельными точками поверхности, По указанной причине методы спекл-фотографии практически не чувствительны к нормальным смещениям. Теперь предположим, что на световое поле, имеющее спекл-структуру, наложена опорная волна, распространяющаяся в направлении z. Тогда наблюдаемая спекл-картина является результатом интерференции поля спеклов с опорной волной. Если объект сместится на расстояние z, то относительная фаза этих двух полей изменится на

$$\Delta\varphi = \frac{2\pi}{\lambda}\left(1+\cos\theta_i\right)\Delta z, \qquad \textbf{(2.5.8)}$$

где θ - угол падения излучения на объект.

Поэтому освещенность спекл-картины будет периодически изменяться при движении объекта в осевом направлении. Если

$$\Delta z = \frac{N\lambda}{1+\cos\theta_i}, \quad N = 0,\pm1,\pm2,\ldots, \qquad \textbf{(2.5.9)}$$

то картина спеклов будет идентична картине, соответствующей исходному положению объекта, для которого Δz=0. Если

$$\Delta z = \frac{\left(N+\frac{1}{2}\right)\lambda}{1+\cos\theta_i}, \quad N = 0,\pm1,\pm2,\ldots, \qquad \textbf{(2.5.10)}$$

то контраст будет обращенным, то есть области, которые первоначально были темными, теперь станут светлыми и наоборот. Поэтому, если объект медленно движется в осевом направлении, то кажется, что спеклы мерцают. Основанный на этом эффекте спекл-интерферометр можно использовать для наблюдения модовой структуры колебаний поверхностей. Исследуемая поверхность освещается лазерным светом, и ее изображение строится с использованием апертуры переменного диаметра, для того чтобы можно было изменять размеры спеклов. Свет, рассеянный узловыми областями поверхности, образует в плоскости наблюдения отчетливую неподвижную спекл-картину. Однако освещенность спеклов в других точках изображения периодически флуктуирует при вибрации поверхности, и если изображение наблюдают визуально или фотографируют со временем экспозиции, превышающим период колебаний, то освещенность спеклов усредняется, создавая относительно однородную засветку. Такое изображение называют спекл- интерферограммой. Наблюдатель может обнаружить узловые области, поскольку спеклы в этих местах имеют высокий контраст. На изображении вибрирующих участков контраст спеклов низкий. Таким

образом по структуре спекл интерферограммы можно судить о динамике смещения отдельных участков поверхности.

2.6. Стохастизация световых пучков в каналах с регулярным распределением неоднородностей. Оптический хаос и фрактальные структуры лучей [23-25]

Рассмотренные в разделах 2.4-2.5 процессы стохастизации излучения непосредственным образом обусловлены случайным распределением неоднородностей среды или неровностей отражающих поверхностей. Существует, однако, принципиально иной механизм стохастизации изначально регулярных световых пучков, который может проявляться даже в средах с регулярным изменением показателя преломления. Этот механизм представляет собой частный (оптический) случай физического сценария перехода к динамическому хаосу детерминированных нелинейных систем.

2.6.1. Уравнения траектории луча

Для описания траектории луча в волноводе воспользуемся гамильтоновым формализмом, изложенным в разделе 1.3.8. Пусть ось z совпадает с осью волноводного канала, а координата луча есть $\mathbf{r}=(x,y,z)$. Придавая уравнениям (1.3.51-1.3.57) векторный вид, получим, что координаты луча (x,y,z) связаны гамильтоновыми уравнениями

$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{r}}, \quad \dot{\mathbf{r}} = -\frac{\partial H}{\partial \mathbf{p}} \qquad \textbf{(2.6.1)}$$

с гамильтонианом

$$H = -\sqrt{n^2(\mathbf{r},z) - \mathbf{p}^2}\,. \qquad \textbf{(2.6.2)}$$

Точка обозначает дифференцирование по z. Эта переменная, таким образом, играет роль времени. Импульс $\boldsymbol{p}$ равен

$$\mathbf{p} = \frac{n\dot{\mathbf{r}}}{\sqrt{1+\dot{\mathbf{r}}^2}}\,. \qquad \textbf{(2.6.3)}$$

Параметр $n = n(\mathbf{r},z)$ является показателем преломления. Записанные уравнения упрощаются, если показатель преломления слабо отличается от постоянного значения. Тогда можно записать

$$n^2(\mathbf{r},z) = n^2(\mathbf{r}) + \varepsilon\upsilon(\mathbf{r},z), \qquad \textbf{(2.6.4)}$$

где $n(\mathbf{r})$ соответствует регулярному (однородному по z) случаю, а возмущене $\varepsilon\upsilon$ учитывает влияние неоднородности. Величина $\varepsilon<<1$ - безразмерный параметр возмущения. Благодаря его малости можно записать H в виде

$$H = H_0(\mathbf{r},\mathbf{p}) + \varepsilon V(\mathbf{r},\mathbf{p},z), \qquad \textbf{(2.6.5)}$$

где

$$H_0(\mathbf{r},\mathbf{p}) = -\sqrt{n^2(\mathbf{r}) - \mathbf{p}^2},$$
$$V(\mathbf{r},\mathbf{p},z) = \frac{\upsilon(\mathbf{r},z)}{2H_0}. \qquad \textbf{(2.6.6)}$$

Гамильтониан H_0 определяет невозмущенные траектории луча. Система уравнений (2.6.1), (2.6.5) и (2.6.6) показывает, что мы пришли к обычной динамической задаче о влиянии нестационарного возмущения на частицу, совершающую финитные колебания, которые описываются гамильтонианом H_0. Неоднородность вдоль переменной z эквивалентна нестационарности динамической задачи.

Наиболее простым является плоский случай, когда n не зависит от y. Уравнения (2.6.1), (2.6.5) и (2.6.6) переходят в следующие:

$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial x}, \quad \dot{x} = -\frac{\partial H}{\partial \mathbf{p}}, \quad p \equiv p_x,$$
$$H = H_0 + \varepsilon V(x,p,z), \quad H_0 = -\sqrt{n^2(x) - p^2}. \qquad \textbf{(2.6.7)}$$

Опишем сначала невозмущенное движение луча. Пусть значение n_∞ определяет соответствующие асимптотики n(x) при $x \to \pm\infty$ и n(x) имеет простой вид горба с одним максимумом. Тогда гамильтониану H_0

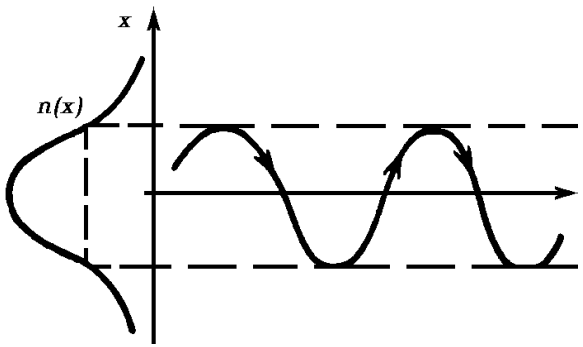


Рис.2.6.1. *Траектория луча (справа от оси x) в волноводном канале распространения и профиль показателя преломления (слева от оси x).*

соответствует движение в простой потенциальной яме. Финитным периодическим траекториям эквивалентной динамической системы соответствуют лучи, распространяющиеся в естественном волноводном канале по траекториям, вид которых показан на *рис. 2.6.1*.

2.6.2. Нелинейный лучевой резонанс

Введем в области финитного движения переменные действие - угол (I,θ) по стандартным правилам теории динамических систем. В этих переменных система (2.6.7) принимает канонический вид

$$H = H_0(I) + \varepsilon V(I,\theta,z),$$
$$\dot{I} = -\varepsilon \frac{\partial V}{\partial \theta}, \quad \dot{\theta} = \omega(I) + \varepsilon \frac{\partial V}{\partial I}. \qquad \textbf{(2.6.8)}$$

Частота

$$\omega(I) = \frac{dH_0}{dI} \qquad \textbf{(2.6.9)}$$

характеризует число колебаний луча между стенками волноводного канала, приходящееся на единицу длины пути. Иными словами, $2\pi/\omega$ есть пространственный период луча в волноводе. Уравнения (2.6.8-2.6.9) являются нелинейными, поэтому основанная на их использовании область геометрической оптики называют нелинейной лучевой динамикой.

Представим теперь, что возмущение обладает пространственной периодичностью по оси z. Будем считать, что пространственный период возмущения составляет величину $2\pi/\kappa$. В настоящее время теория решений уравнений 2.6.8-2.6.9 хорошо разработана. Она предсказывает, что при выполнении условия

$$m\omega(I) + s\kappa = 0 \qquad \textbf{(2.6.10)}$$

в системе возникает нелинейный резонанс (который развивается не во времени, а в пространстве). Ширина этого резонанса равна $\delta\omega=\kappa/m$. При наличии резонанса траектории лучей приобретают неустойчивый характер и обнаруживают многочисленные неупорядоченные пересечения. Это влечет стохастизацию поля и формирование спеклоподобной структуры поля в поперечном сечении волновода. Важно отметить, что возникающие искажения волновых пучков нельзя скомпенсировать никакими известными методами, включая методы адаптивной оптики и методы обращения волнового фронта. Такая ситуация обусловлена прежде всего тем, что в силу принципиальной неустойчивости лучевых траекторий форма амплитудно-фазовых распределений поля в волноводе оказывается чрезвычайно критичной к малым изменениям начальных условий.

2.6.3. Фрактальные лучевые структуры

Использование методов традиционной статистической физики для описания стохастизации световых пучков под влиянием случайных неоднородностей или в результате проявления нелинейного лучевого резонанса не всегда приводит к исчерпывающим результатам. Это во многом связано с тем, что статистические методы не учитывают свойства масштабной инвариантности (скейтлинга), которыми при определенных условиях могут обладать амплитудно-фазовые распределения или лучевые структуры световых пучков. Указанный пробел восполняет применение фрактальных моделей. В математике фрактал представляет собой множество точек в метрическом пространстве, для которого невозможно определить какую-либо из традиционных мер с целой размерностью - длину, площадь или объем (их размерности - соответственно первая степень, квадрат и куб длины). Измерение, например, длины фрактальной кривой может дать бесконечный результат, а заметаемой ею площади - нулевой. Задача измерения таких множеств решается введением мер Хаусдорфа с любой (в том числе нецелой) размерностью. Наибольшая размерность меры Хаусдорфа называется размерностью Хаусдорфа-Безиковича (РХБ) этого множества. Используя эти представления фрактал можно определить, как масштабно-инвариантный, т.е. самоподобный объект, РХБ которого превышает топологическую размерность (1 - для линии, 2 - для поверхности и т.д.).

Успех в применении фрактальных моделей в физике обусловлен прежде всего тем, что фрактальные формы присущи огромному числу процессов и структур. Весьма эффективными оказались фрактальные представления и при анализе процессов формирования и распространения световых пучков. Не выходя далеко за рамки обсуждаемой темы, отметим, что фрактальные структуры присутствуют и в картине лучей, распространяющихся в продольно неоднородном волноводе. Их появление является прямым следствием возникновения нелинейных резонансов.

Как уже указывалось, в продольно однородных волноводах лучи периодически колеблются вблизи оси волновода, не покидая его. Захват лучей связан либо с наличием отражающих стенок, либо с неоднородным поперечным распределением показателя преломления. Длина цикла луча определяется начальным углом наклона луча к оси волновода. При наличии продольных неоднородностей (неровности стенок, колебания оси, изменения показателя преломления) становится возможным захват лучей в нелинейные резонансы. Рассмотрим волновод с однородным заполнением и

абсолютно отражающими стенками; одна его стенка плоская, а другая имеет периодические неровности вида

$$f(z) = \frac{4b}{L}\xi(1-\xi), \qquad \textbf{(2.6.11)}$$

где *b, L* ñ соответственно амплитуда и период неровностей, $\xi = \frac{z}{L}$ ñ дробная часть нормированной на период продольной координаты *z*. При *b=0* ширина волновода ñ *h*. Лучи в таком волноводе распространяются, попеременно отражаясь от его стенок. Распространение луча можно описать нелинейным отображением, определяющим угол и продольную координату отражения луча от плоской стенки через угол и продольную координату предыдущего отражения от плоской стенки. Если амплитуда неровностей равна нулю, то длина цикла луча *D* ñ расстояние между двумя последовательными отражениями от стенки ñ постоянна и равна $D = 2hctg\theta_0$, где θ_0 ñ исходный угол выхода луча. Неровности оказывают наиболее сильное влияние на лучи, находящиеся в нелинейном резонансе с периодом неровностей, что для некоторых целых чисел m и n означает выполнение равенства $\frac{2n\pi}{D} = \frac{2\pi m}{L}$, или

$$tg\,\theta_0^{(m,n)} = \frac{m}{n}\frac{2h}{L}, \qquad \textbf{(2.6.12)}$$

обеспечивающего резонанс между какими-либо гармониками неровностей и траектории луча. Лучи с углом выхода вблизи одного из резонансных углов $\theta_0^{(m,n)}$ захватываются в резонанс и имеют одинаковые средние периоды отражений и времена распространения. На *рис. 2.6.2, а* показана зависимость пространственной частоты к колебаний луча от угла выхода θ_0. Эта кривая состоит из ступенек с постоянной величиной *k*, расположенных вблизи резонансных углов выхода. Распределение ступенек по углу выхода фрактально, в том смысле, что при увеличении разрешения *r* по углу число *N(r)* промежутков между ступеньками степенным образом зависит от разрешения. Фрактальность иллюстрируется на *рис. 2.6.2, а* двумя врезками, показывающими увеличено малый участок кривой и график зависимости *N(r)*. На *рис. 2.6.2, б* представлена аналогичная ступенчатая зависимость длины луча (времени распространения сигнала по лучу).

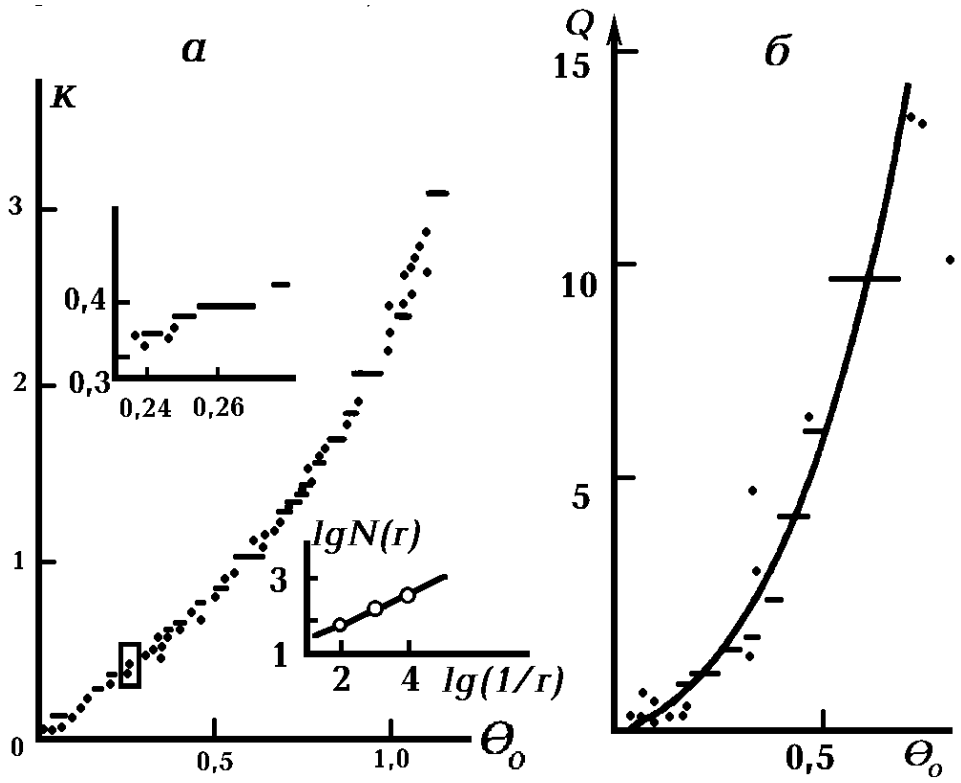


Рис.2.6.2. *(a) Зависимость пространственной частоты колебаний луча от угла выхода луча из источника. Шаг по углу 0.01; на врезке - увеличенный участок графика, обведенный прямоугольником, шаг по углу 0.001. Параметры волновода: h/L=1/3, b/L=0.001. (b) Зависимость длины луча (за вычетом расстояния вдоль волновода z) от угла выхода луча. Параметры волновода: h/L=1/3, b/L=0.005.*

2.7. Пучки с винтовыми дислокациями волнового фронта. Элементы сингулярной оптики [21, 26, 27]

Среди волновых пучков с разнообразной структурой амплитудно-фазового профиля особое место занимают пучки с винтовыми возмущениями волнового фронта. Такого рода возмущения обусловливают вихревой характер распространения световой энергии, что позволяет говорить о существовании своеобразных оптических вихрей.

В силу ряда аспектов фундаментального характера, а также возможности оригинальных технических приложений изучение оптических вихрей ведется у нас в стране и за рубежом весьма бурными темпами. В настоящее время в оптике фактически сформировалась новая область, называемая "оптикой винтовых полей" или "сингулярной оптикой". В рамках этой области рассматриваются свойства оптических вихрей, а также физический механизм их образования. Данный раздел вводит читателя в новую область когерентной оптики, знакомя с условиями возникновения и основными особенностями винтовых световых полей.

2.7.1. Общая характеристика дислокаций волнового фронта

Как известно, волновой фронт световых пучков, близких по своим свойствам к плоской волне выглядит как семейство непересекающихся поверхностей (*рис*. 2.7.1, а). Расстояние между соседними поверхностями равно длине волны λ.

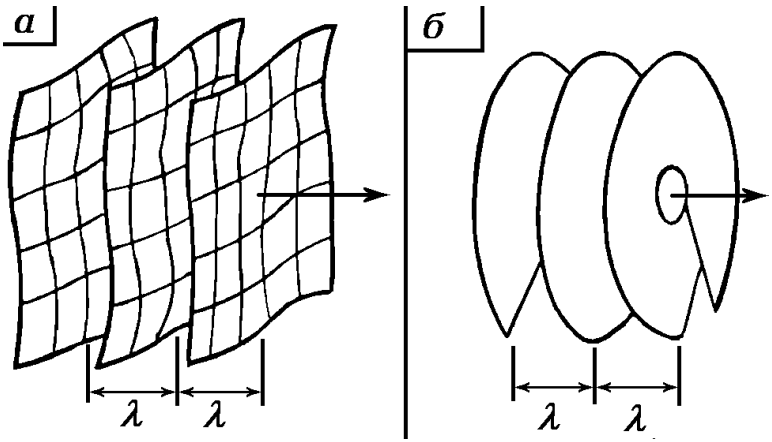


Рис. 2.7.1. *Структура волновых фронтов в отсутствие (а) и при наличии винтовой дислокации (б).*

Имеющие место в реальных пучках отклонения волновых фронтов от плоской формы называются оптическими аберрациями. Аберрации заметно ухудшают свойства световых пучков. Их учет и минимизация составляют важную задачу классической теории аберраций, широко привлекаемую для расчета разнообразных лазерных систем. Однако все аберрации, рассматриваемые в классической теории, деформируют волновой фронт без изменения его топологии.

Иная картина наблюдается при наличии в лазерном пучке оптических вихрей. Если такие вихри появились, то на поверхности волнового фронта присутствуют особые точки, которые во многих отношениях аналогичны известным в физике твердого тела дефектам кристаллической решетки - *винтовым дислокациям* и имеют такое же название. В сам**о**й особой точке амплитуда световых колебаний обращается в нуль, а значение фазы не определено. В окрестности ее происходят резкие коллапсирующие фазовые изменения. Из-за наличия такой особенности функция фазового распределения относится к классу сингулярных функций, что и стало причиной появления упомянутого выше термина "сингулярная оптика". Основное свойство винтовой дислокации (ВД) состоит в том, что при обходе вокруг нее фаза изменяется ровно на 2π. На поверхности волнового фронта может возникать как единичная ВД, так и целая система дислокаций. В зависимости от направления закрутки винта, ВД подразделяются на левые (отрицательные) и правые (положительные). Появление ВД кардинальным образом меняет топологию волнового фронта. Эквифазная поверхность перестает быть многолистной (см. *рис. 2.7.1, а*), и осуществляется переход к единой поверхности со специфической винтовой структурой. Это иллюстрирует *рис. 2.7.1, б*, на котором изображен волновой фронт лазерного пучка с ВД, расположенной на оси. Направление распространения световой энергии задается вектором Умова-Пойнтинга, перпендикулярным, как известно, поверхности волнового фронта в каждой точке. Следовательно, в окрестности ВД будет происходить "завихрение" энергетического потока.

В окрестности ВД амплитуду световых колебаний u можно представить в виде

$$\mathrm{u} = \mathrm{C_x} x + i C_y y, \qquad \textbf{(2.7.1)}$$

где x и y - декартовы координаты в плоскости, перпендикулярной направлению распространения лазерного пучка; C_x, C_y - произвольные константы. Использованное в формуле (2.7.1) комплексное представление амплитуды указывает, что световые колебания в окрестности ВД можно представить в виде суммы гармонических колебаний, сдвинутых по фазе на

$\pi/2$ (умножение на мнимую единицу i второго члена эквивалентно указанному сдвигу фаз). Как видно из формулы (2.7.1), амплитуда поля u меняет знак при одновременном изменении знаков x и y, что свидетельствует о противофазности световых колебаний по разные стороны от ВД. При этом в центре ВД - точке (0,0) - одновременно обращаются в нуль действительная и мнимая части комплексной амплитуды. Последнее свойство при теоретическом анализе вихревых полей широко используется для идентификации ВД. Если $C_x=\pm C_y=C$, то используя формулу Эйлера, можно записать:

$$u = C \cdot r \cdot \exp(\pm i\alpha), \qquad \textbf{(2.7.2)}$$

где $r = \sqrt{x^2 + y^2}$ - полярный радиус, α - азимутальный угол. Выражение (2.7.2) описывает регулярную ВД с равномерным азимутальным вращением фазы.

Приведенные выше характеристики относятся к так называемым дислокациям первого порядка, азимутальное изменение фазы в окрестности которых составляет 2π. Однако физически представляется возможным формирование ВД более высокого порядка. Скорость азимутального вращения фазы у них выше и изменение фазы за полный обход по азимуту равно $2\pi\ell$, где целое число ℓ определяет порядок дислокации. Для описания регулярных ВД высших порядков используется выражение

$$u = C \cdot r^{\ell} \cdot \exp(\pm i\ell\alpha). \qquad \textbf{(2.7.3)}$$

2.7.2. Методы регистрации ВД

Образование ВД на волновом фронте лазерных пучков является чисто фазовым эффектом. Поэтому на основе анализа только лишь распределения интенсивности в лазерном пучке зарегистрировать особенности распределения фазы световых колебаний в области ВД не представляется возможным. Единственным способом, обеспечивающим надежную идентификацию ВД, является способ, основанный на использовании интерферометрической информации. Интерферограммы поперечного сечения пучка могут быть получены разными способами. Самым удобным с точки зрения обработки является способ, основанный на регистрации структуры интерференции исследуемого поля с плоской или сферической однородными волнами (естественно, исследуемая и опорная волны должны быть взаимно когерентными).

Рассмотрим интерференционные "портреты" ВД первого и второго порядков. Если в качестве опорной волны взять соосную сферическую

волну с амплитудой a и радиусом кривизны волнового фронта R, то суммарная комплексная амплитуда поля u_Σ в области интерференции будет иметь вид:

$$u_\Sigma = C \cdot r^{\ell} \cdot \exp(i\ell\alpha) + a \cdot \exp\left(i\frac{\pi r^2}{\lambda R} + i\beta \right), \qquad \textbf{(2.7.4)}$$

где β - расстройка по фазе между интерферирующими волнами. Присутствующий в показателе второй экспоненты фактор $\pi r^2/\lambda R$ определяет фазовый сдвиг между сферической и плоской волнами.

Переходя к безразмерной величине радиус-вектора $r' = \sqrt{\pi r^2/\lambda R}$, и приравнивая к нулю действительную и мнимую части формулы (2.7.4), легко установить, что темные линии на регистрируемой интерференционной картине будут удовлетворять условию:

$$r' = \sqrt{\ell\alpha - \beta}. \qquad \textbf{(2.7.5)}$$

На *рис. 2.7.2, а, б* показана структура темных линий на интерферограмме для ВД первого и второго порядков. Здесь же для сравнения пунктиром изображено первое темное кольцо интерференционной картины, получающейся в том случае, если бы со сферической волной вместо волны с дислокацией интерферировала плоская волна. Из рисунков видно, что в области ВД формируются спиральные интерференционные полосы. ВД первого порядка соответствует ординарная спираль, а дислокации второго порядка - двойная (двухзаходная). Параметр β не влияет на общую структуру интерферограмм. Его изменение приводит лишь к повороту спиралей вокруг оси. Если ВД смещена относительно оси опорной сферической волны, то структура интерферограмм меняется (см. *рис. 2.7.2, в, г*). ВД первого порядка порождает на интерферограмме одну дополнительную полосу, а ВД второго порядка - две полосы.

Рассмотрим теперь вопрос о причинах и физическом механизме формирования вихревых лазерных полей. Оптико-физические процессы, вызывающие появление оптических вихрей весьма разнообразны. Излучение с вихревой структурой может при определенных условиях формироваться в результате интерференции лазерных пучков с исходно регулярным волновым фронтом, при их прохождении через случайно-неоднородные и нелинейные среды, а также через волоконные многомодовые световоды или специальным образом изготовленные голограммы. Кроме того, возможно возбуждение вихревых полей непосредственно в лазерах. Мы ограничимся более подробным

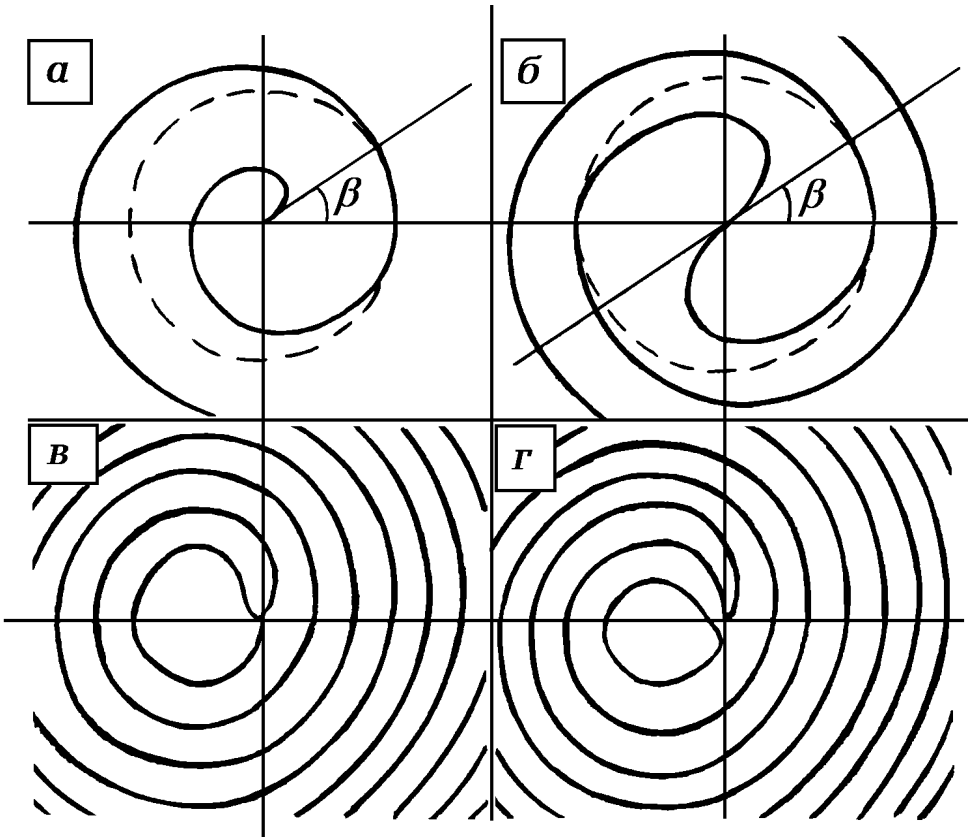


Рис. 2.7.2. *Интерференционные "портреты" винтовых дислокаций первого (а,в) и второго (б,г) порядков, опорная волна соосна (а,б) и смещена(в,г)*

рассмотрением механизма образования оптических вихрей в случайно-неоднородных средах и в лазерных резонаторах.

2.7.3. Оптические вихри в случайно-неоднородных средах

Наиболее часто вихревые возмущения фазы лазерных пучков, в частности, ВД наблюдаются при распространении излучения через передающие среды со случайными неоднородностями показателя преломления. Так, ВД могут быть зарегистрированы на волновом фронте в результате прохождения достаточно больших расстояний в атмосфере. В последнем случае оптическая неоднородность воздуха обязана развитию турбулентных образований в приземном слое атмосферы из-за наличия там температурных градиентов.

Рис. 2.7.3 иллюстрирует трансформацию волнового фронта (изображен пунктиром) лазерного излучения в результате прохождения им случайно-неоднородной среды. Имея в плоскости $П_1$ плоский волновой фронт, излучение на своем пути от $П_1$ к $П_2$ встречает неоднородности показателя преломления. По мере приближения к плоскости $П_2$ в световой волне возникают изменения фазы, связанные с ускорением или замедлением скорости распространения различных участков волнового фронта. Возмущения эквифазной поверхности в плоскости $П_2$ вызывают отклонение локальных нормалей к волновому фронту от первоначального направления волны. Если изгибы эквифазной поверхности значительны, то нормали пересекаются, вызывая интерференцию различных участков волнового фронта в плоскости наблюдения $П_3$.

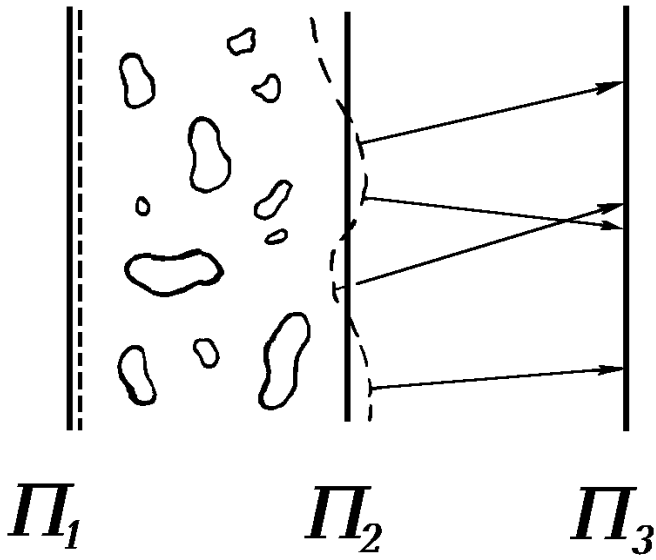


Рис. 2.7.3. *Трансформация волнового фронта в случайно-неоднородной среде.*

Интерференция парциальных волн в плоскости $П_3$ вызывает глубокую хаотическую модуляцию амплитуды $a(x,y)$ и фазы $\varphi(x,y)$, в результате чего распределение интенсивности имеет спеклоподобный вид. Одновременно с этим появляются пересечения линий смены знака действительной и мнимой частей комплексной амплитуды

$$u = a(x,y)\exp\left[i\varphi(x,y)\right]. \qquad \textbf{(2.7.6)}$$

А такие точки пересечения, как мы видели выше, и есть дислокации фазы. Эту ситуацию иллюстрирует *рис. 2.7.4*, где показаны пересечения нулевых линий (точки абсолютного нуля поля), определяющие положение ВД. Таким образом, наличие в передающей среде неоднородностей, создает условия для стохастизации светового поля. Одним из признаков стохастичности поля является формирование вихревой структуры поля.

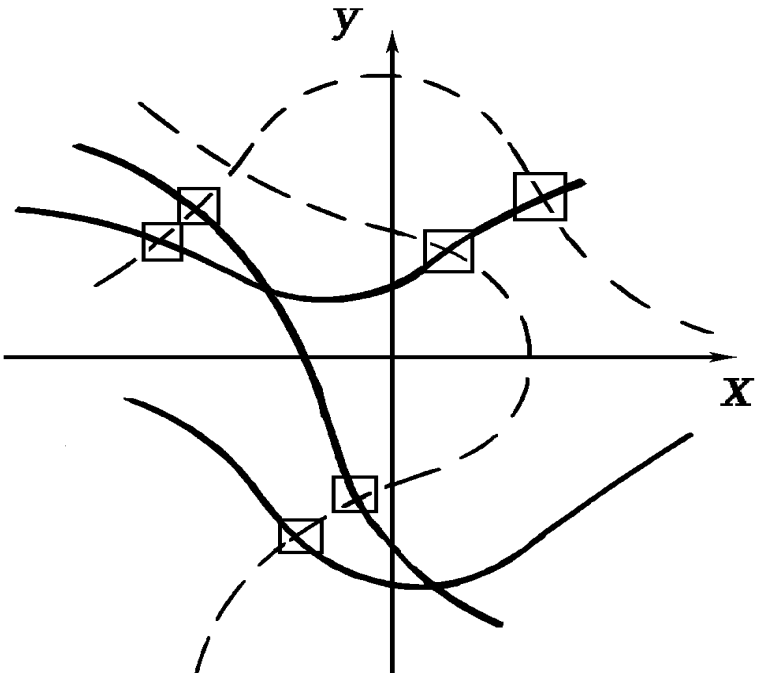


Рис. 2.7.4. *Образование винтовых дислокаций в результате пересечения нулевых линий действительной (пунктир) и мнимой (непрерывные кривые) частей комплексной амплитуды, расположение дислокаций помечено прямоугольниками.*

Эксперименты полностью подтверждают правильность изложенных выше теоретических представлений. На *рис. 2.7.5, а, б* приведены фотографии лазерного пучка, до и после прохождения атмосферной трассы длиной 600 м в условиях развитой мелкомасштабной турбулентности. Исходная структура пучка (*рис. 2.7.5, а*) характеризуется высокой

однородностью распределения интенсивности. После прохождения трассы (*рис. 2.7.5, б*) структура пучка претерпевает качественные изменения. Распределение интенсивности приобретает спеклоподобный вид. При этом на границах спеклов формируются ВД. На *рис. 2.7.5, в* приведена интерферограмма фрагмента поперечной структуры, по которой можно судить о структуре фазы. Наличие многочисленных разветвлений интерференционных полос свидетельствует о присутствии большого числа ВД (для удобства наблюдения часть их помечена на рисунке прямоугольниками). Анализ структуры интерферограммы путем ее сопоставления с расчетными интерферограммами (*рис. 2.7.2, в, г*) показывает, что все ВД имеют первый порядок (ℓ=1), причем количество положительных и отрицательных ВД одинаково. Тем самым можно говорить о существовании своеобразного закона сохранения общего топологического заряда поверхности волнового фронта.

Рис. 2.7.5. *Фотографии (а,б) и интерферограмма (в) лазерного пучка; а - исходный пучок, б - пучок после прохождения атмосферной трассы.*

Отсутствие ВД более высоких порядков объясняется тем, что они крайне неустойчивы даже к малым случайным возмущениям поля. Незначительное "шевеление" волнового фронта вызывает распад ВД высокого порядка на соответствующее число ВД с ℓ=1.

Количество ВД впрямую связано с интенсивностью турбулентных процессов в атмосфере, которую можно регистрировать по уровню локальных флуктуаций температуры и показателя преломления. Эта связь проявляется настолько отчетливо, что были предложены и нашли практическую реализацию методы диагностики турбулентных состояний атмосферы на основе регистрации и подсчета числа ВД.

Указанные методы могут служить положительным примером использования уникальных свойств вихревых полей. Однако часто приходится сталкиваться и с негативными эффектами, вызванными появлением оптических вихрей. Так, присутствие ВД на волновом фронте излучения серьезным образом усложняет работу адаптивных устройств, используемых в оптических линиях связи для компенсации фазовых искажений. В таких устройствах в качестве основного элемента часто используется гибкое отражающее зеркало. С помощью специальной системы управления кривизна отдельных участков зеркала адаптируется под изгибы волнового фронта, падающего на зеркало излучения, что позволяет компенсировать фазовые возмущения. Но обладая высокой эффективностью при компенсации обычных аберраций, такое адаптивное устройство оказывается не в состоянии ликвидировать возмущения винтового типа, так как отражающая поверхность зеркала не может менять своей топологии. Тем не менее, ситуация не является безнадежной. В настоящее время разработаны методы, основанные на использовании эффектов обращения волнового фронта в нелинейных средах, которые способны с успехом бороться и с топологическими деформациями волнового фронта.

Для рассмотренных выше процессов формирования вихрей в световых пучках с исходно плоским фронтом принципиально важным является наличие в каналах распространения неоднородностей, изменяющих скорость световой волны. В определенном смысле ситуация аналогична развитию турбулентности в потоках жидкости или газа, когда присутствие препятствий или ограничивающих поток поверхностей приводит к локальным изменениям скорости и переходу от ламинарного к турбулентному движению при достаточно высокой скорости потока. В оптике безвихревые пучки с непересекающимися лучами могут быть соотнесены с ламинарными потоками жидкости или газа, а пучки с оптическими вихрями - с турбулентными потоками.

Оптические вихри могут образовываться и непосредственно в источниках когерентного излучения - лазерах. Здесь механизм их формирования имеет ряд принципиальных отличий, хотя дефекты и неоднородности в оптических элементах лазера играют не последнюю роль. Рассмотрим этот вопрос более подробно.

2.7.4. Генерация винтовых полей в лазерах.

Винтовые М-моды

Как указывалось в разделе 2.1, для устойчивых резонаторов можно найти решение волнового уравнения в виде самовоспроизводящихся при последовательном отражении от зеркал световых структур. В цилиндрической системе координат эти структуры соответствуют модам Лагерра-Гаусса. Распределение их полей описывается выражением (2.1.26). Несложно показать, что это выражение допускает обобщение, согласно которому внурирезонаторное поле можно представить в виде:

$$u_{pl} \sim f_{p\ell}\left(\frac{r}{w}\right)\cdot\left(a_{p\ell}\cdot e^{i\ell\alpha}+b_{p\ell}\cdot e^{-i\ell\alpha}\right), \qquad \textbf{(2.7.7)}$$

где w - параметр ширины распределения; $f_{p\ell}(t)\sim t^{\ell}\cdot L_p^{\ell}\left(2t^2\right)\cdot e^{-t^2}$, L_p^{ℓ} - полиномы Лагерра; $a_{p\ell}$ и $b_{p\ell}$ - произвольные постоянные; индекс ℓ, как и прежде, характеризует азимутальные изменения фазы. Вид показателей у экспоненциальных множителей указывает на возможность возбуждения в резонаторе лазера винтовых полей. Однако реализовать эту возможность на практике бывает не просто. Первое препятствие состоит в том, что обычно в резонаторе отсутствует предпочтительное направление закрутки оптического вихря, в результате чего осуществляется равенство коэффициентов $a_{p\ell}$ и $b_{p\ell}$, и формула (2.7.7) принимает вид:

$$u_{p\ell} \sim f_{p\ell}\left(\frac{r}{w}\right)\cdot a_{p\ell}\cdot\begin{cases}\cos\ell\alpha\\ \sin\ell\alpha\end{cases}, \qquad \textbf{(2.7.8)}$$

полностью соответствующий формуле (2.1.26). Поле (2.7.8) характеризуется знакопеременной амплитудой с многочисленными кольцевыми радиальными узловыми линиями. Исчезновение винтовой структуры поля происходит вследствие наложения и интерференции вихрей разного знака.

Вторая трудность формирования пучков с оптическими вихрями связана с многомодовым характером лазерной генерации. Как правило, в резонаторе лазера возбуждается целая система мод с разными индексами p и ℓ. В общем случае частоты этих мод не совпадают, что снижает когерентные свойства излучения и не позволяет говорить о единой для генерируемого излучения эквифазной поверхности.

Указанные препятствия для генерации пучков с оптическими вихрями, тем не менее, можно обойти. Для этого нужно прежде всего

обеспечить совпадение частот генерируемых мод. Проще всего это делается путем перехода к такой геометрической конфигурации резонатора, которая обеспечивает равенство частот всех генерируемых мод, то есть их частотное вырождение по индексам p и ℓ. В частности, эффект вырождения достигается, как уже указывалось в разделе 2.2.2, в конфокальном резонаторе. На *рис. 2.7.6, а* представлена фотография излучения лазера с конфокальным резонатором. Видно, что из-за интерференции большого числа мод с одинаковыми частотами и неким разбросом фаз (в общем, случайным) формируется спеклоподобное распределение интенсивности. Однако в отличие от хаотической спекл-структуры поля в случайно-неоднородных средах, приведенное распределение обнаруживает четко выраженную центральную симметрию. Интерферометрический анализ таких пучков показывает, что на границах между спеклами формируется целая система ВД. Об этом свидетельствует фрагмент интерферограммы пучка, приведенный на *рис. 2.7.6, б*. Как и на предыдущем рисунке, для удобства идентификации ВД их положение отмечено прямоугольниками.

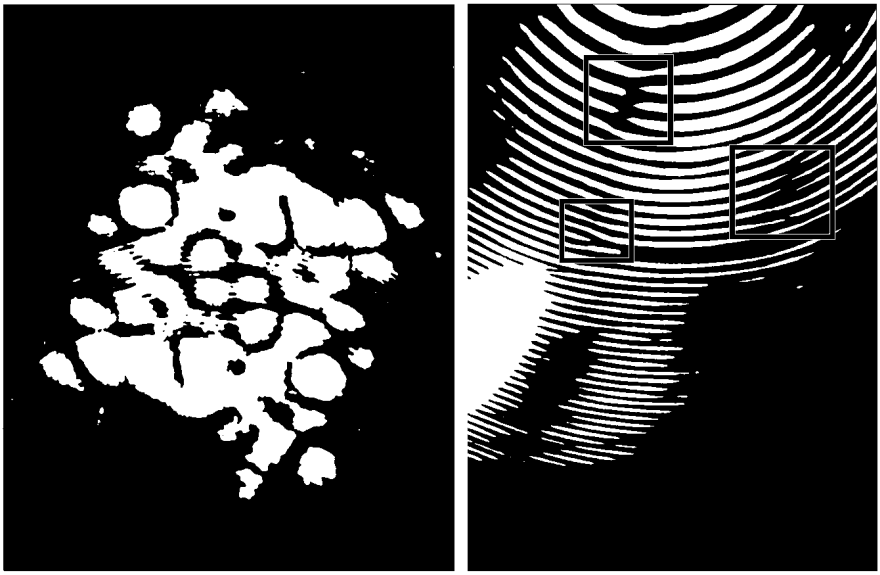


Рис. 2.7.6. *Структура (а) и интерферограмма (б) излучения лазера с конфокальным резонатором.*

Аналогично расположению спеклов, распределение ВД в выходных пучках характеризуется высокой степенью упорядоченности. Регулярный характер распределения ВД в лазерных пучках стимулировал появление в литературе термина "фазовый сингулярный кристалл". Фазовый

сингулярный кристалл служит своеобразным остовом для вихревого лазерного пучка, определяя все его свойства, которые проявляются при распространении в свободном пространстве или оптических системах. Важное значение дислокационной структуры излучения определило развитие в рамках сингулярной оптики своеобразной "кристаллографии", предмет которой составляет расчет и анализ характеристик фазовых сингулярных кристаллов.

Остановимся, наконец, на возможности генерации в лазерах регулярных винтовых полей. Эксперименты показали, что такие поля не сложно получать во многих типах лазеров при не слишком высоком превышении порога самовозбуждения. Вначале на отражающее покрытие одного из зеркал на оси резонатора наносят маленькое пятнышко из поглощающего материала. Это подавляет возбуждение мод с максимальным значением интенсивности на оси, обладающих, как правило, наибольшим усилением. Затем уменьшают размеры внутрирезонаторной диафрагмы до тех пор, пока излучение лазера на выходе не примет кольцевую форму (см. *рис. 2.7.7, а*). Это и есть пучок с винтовой структурой структурой волнового фронта. Его интерферограмма приведена на *рис. 2.7.7, б*. Ее сравнение с расчетной интерференционной структурой на *рис. 2.7.1, а* позволяет утверждать, что в центре сфотографированного пучка находится ВД с топологическим зарядом, равным единице. Варьируя размеры поглощающей зоны на поверхности зеркала и внутрирезонаторной диафрагмы, в принципе можно получать регулярные винтовые моды с более высоким топологическим зарядом. То, что лазер в таких условиях генерирует лишь одну из двух возможных винтовых мод (правую или левую) объясняется неравенством их потерь. Вблизи порога самовозбуждения из-за всегда присутствующих слабых паразитных отражений от элементов лазера добротность одной из винтовых мод может случайным образом оказаться выше, и в результате межмодовой конкуренции в резонаторе будет формироваться мода, соответствующая ВД определенного знака. При увеличении накачки лазера и значительном превышении порога самовозбуждения указанные факторы нивелируются, и появляется мода с противоположной закруткой. Интерферируя между собой, моды будут формировать поле, описываемое формулой (2.1.26) с нулевым значением индекса p. Такое поле характеризуется системой располагающихся по диаметру пучка узловых линий, количество которых соответствует топологическому заряду ℓ. В качестве иллюстрации на *рис. 2.7.7, в* приведено поле данного типа с ℓ=5.

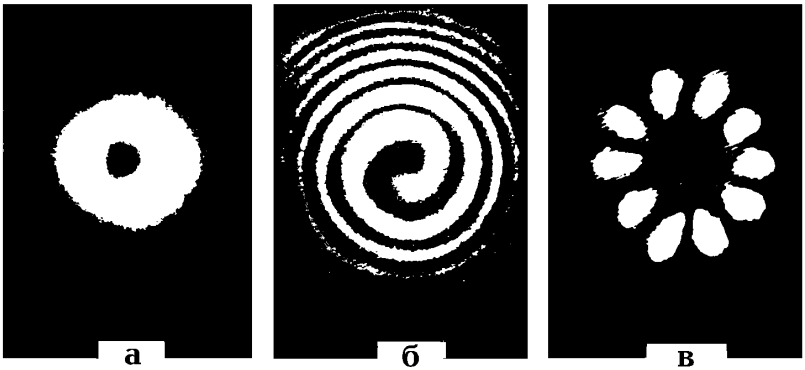


Рис. 2.7.7. *Структура поля регулярной винтовой моды (а,б) и суперпозиции винтовых мод (в); а,в - распределение интенсивности, б - интерферограмма.*

Регулярные винтовые поля, возбуждаемые в лазерных резонаторах, допускают довольно простую геометрооптическую интерпретацию, если им поставить в соответствие объемные лучевые пакеты. В отличие от плоского пакета, рассмотренного в разделе 2.2.1, точки отражения луча на зеркалах в объемном пакете располагаются по окружности *(рис. 2.7.8)*. Расчеты с использованием аппарата матричной оптики показывают, что его структура определяется через *g*-факторы резонатора посредством выражения

$$\frac{K}{N} = \frac{1}{\pi}\arccos\left(g_1 g_2\right), \qquad \textbf{(2.7.9)}$$

где N - число отражений луча на каждом из зеркал, K - число оборотов луча по азимуту, необходимых для полного замыкания траектории.

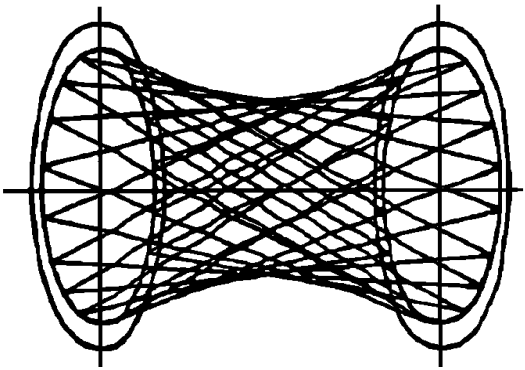


Рис.2.7.8. *Структура объемного лучевого пакета.*

Если в направлении лучей, формирующих объемный лучевой пакет, распространяется гауссов пучок, то принято говорить о возбуждении в резонаторе винтовой многоходовой моды (М-моды). При достаточно большом расстоянии между точками поворота лучей на зеркалах распределение поля каждой из М-мод в плоскости зеркал будет представлять круговую систему пятен с гауссовой формой распределения интенсивности. Возбуждение в резонаторе винтовых М-мод весьма широко используется в лазерной технике для съема энергии с лазерно-активных сред кольцевой формы. Примером может служить организация оптического тракта коаксиального CO_2-лазера.

Упрощенная схема коаксиального CO_2-лазера приведена на *рис. 2.7.9*.

Активная среда возбуждается ВЧ разрядом в промежутке между двумя цилиндрическими электродами 2 и 3. Устойчивый резонатор, образованный зеркалами 1 и 4, обеспечивает возбуждение М-мод с круговым расположением световых пятен на зеркалах. Их образующие пучки распространяются в межэлектродном зазоре, формируя внутрирезонаторное поле с кольцевой геометрией. Для вывода излучения одно из зеркал изготовляется полупрозрачным; в ряде случаев вывод энергии осуществляется через радиальную полупрозрачную зону или щель на одном из зеркал.

Мощные вихревые пучки с кольцевой формой, генерируемые в лазерах, широко используются в лазерной технологии обработки материалов, так как создают более равномерное распределение температуры. Кроме того, при импульсной генерации вихревого пучка появляется возможность пропускать его через воздушную среду в волноводном режиме, исключающем его дифракционное расплывание. Это связано с тем, что из-за обращения интенсивности в нуль в центре пучка температура воздуха на оси пучка оказывается ниже (а показатель преломления выше), чем в остальных областях его сечения. Тем самым создаются условия для постоянной подфокусировки пучка в процессе его распространения.

Отметим в заключение, что изучение винтовой дислокационной формы волновых фронтов приобрело в последнее время актуальность не только в области оптических исследований, но и в других разделах физики, в которых большую роль играют волновые процессы. Это прежде всего относится к радиофизике и акустике. Есть сообщения о регистрации ВД в ионосферных радиосигналах, а также в акустических сигналах, распространяющихся в океанических волноводах. Исследование ВД

стимулировало развитие ряда новых научных направлений. Среди них в качестве наиболее яркого примера можно выделить развитие дислокационной томографии океана. Она позволяет определять гидродинамические возмущения различных типов путем регистрации вихрей в акустической волне, распространяющейся в океане. Не исключено, что методами дислокационной томографии удастся обнаруживать зарождение даже мезо-масштабных неоднородностей типа синоптических турбулентных образований. Таким образом, концепция фазовых сингулярностей и связанных с ними вихревых структурных элементов оказалась весьма плодотворной при анализе сложных волновых процессов самой разной физической природы.

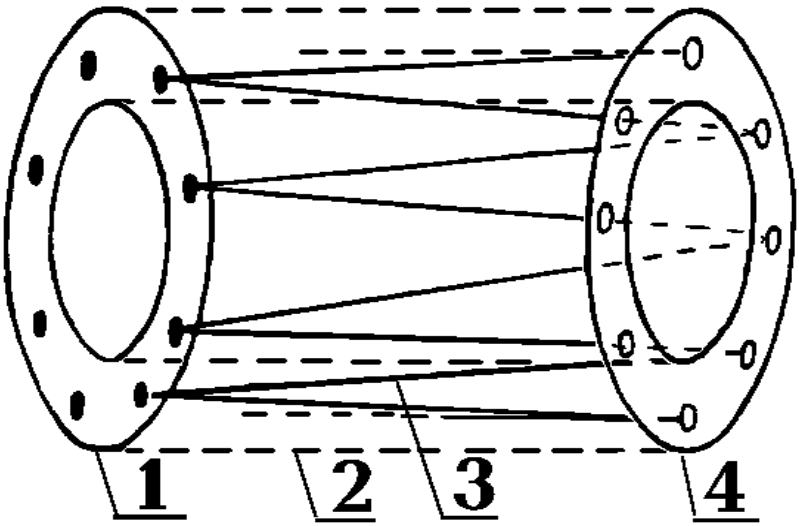


Рис. 2.7.9. *М-мода в резонаторе коаксиального CO_2-лазера:*

1,4-зеркала резонатора;
2,3-цилиндрические электроды.

ЛИТЕРАТУРА К ГЛАВЕ 2

1. Когельник Г., Ли Т. Световые пучки, резонаторы и типы колебаний. //Справочник по лазерам под ред. А.М.Прохорова. М.: Сов. радио, 1978, с.11-24.

2. Маркузе Д. Оптические волноводы. - М.: Мир, 1974, 576 с.

3. Карлов Н.В. Лекции по квантовой электроники. - М.: Наука, 1983, 302 с.

4. Солимено С., Крозиньяни Б, Ди Порто П. Дифракция и волноводное распространение оптического излучения. - М.: Мир, 1989, 662с.

5. Короленко П.В., Маркова С.Н., Хапаев А.М.//Известия высших учебных заведений. Радиофизика. 1978, т.21, №11, с. 1644-1647

6. Пахомов И.И., Цибуля А.Б., Расчет оптических систем лазерных приборов. - М.: Радио и связь, 1986, 152 с.

7. Короленко П.В. Оптика когерентного излучения. - М.: Изд-во Московского университета, 1989, 96 с.

8. Быков В.П., Вайнштейн. Геометрическая оптика открытых резонаторов. //ЖЭТФ, 1964, т.47, вып. 8, с. 508-517.

9. Бабич В.М., Булдырев В.С. Асимтотические методы в задачах дифракции коротких волн. - М.: Наука, 1972.

1О. Быков В.П. Специальные оптические резонаторы. //Справочник по лазерам под ред. А.М.Прохорова. М.: Сов. радио, 1978, с.24-46.

11. Лиханский В.В., Напартович А.П. Излучение оптически связанных лазеров. УФН, 1990, т. 160, вып.3, с. 101-143.

12. Козанне А., Флере Ж., Мэтр Г., Руссо М. Оптика и связь. - М.: Мир, 1984, 504 с.

13. Хаус Х. Волны в оптоэлектронике. - М.: Мир, 1988.

14. Marcatili E.A., Shmeltzer R. Hollow metallic and dielectric waiveguides for long distance optical transmission and their losses // Bell Syst. Tech. J., 1963, v.43, π7, 1783-1809.

15. Исимару А. Распространение и рассеяние волн в случайно-неоднородных средах, т.2.-М.: Мир, 1981, 318 с.

16. Сб. "Распространение лазерного пучка в атмосфере" под ред. Д.Стробена.-М.: Мир, 1981, 416 с.

17. Рытов С.М., Кравцов Ю.А., Татарский В.И. Введение в статистическую радиофизику, ч.2. - М.: Наука, 1978, 464 с.

18. Чернов Л.А. Волны в случайно-неоднородных средах. - М.: Наука, 1975, 172 с.

19. Семенов А.А., Арсеньян Т.И. Флуктуации электромагнитных волн на приземных трассах.- М.: Наука, 1978, 272 с.

20. Вест Ч. Голографическая интерферометрия. - М.: Мир, 1982, 504 с.

21. Зельдович Б.Я., Пилипецкий Н.Ф., Шкунов В.В. Обращение волнового фронта. - М.: Наука, 1985, 228 с.

22. Франсон М. Оптика спеклов. - М.: Мир, 1980, 174 с.

23. Заславский Г.М., Сагдеев Р.З. Введение в нелинейную физику. - М.: Наука, 1988, 368 с.

24. Адуллаев С.С., Заславский Г.М. Стохастическая неустойчивость лучей и спекл-структура поля в неоднородных средах. //ЖЭТФ, 1984, т.87, в.3 (9), с. 763-775.

25. Зосимов В.В., Лямшев Л.М. Фракталы в волновых процессах.//УФН, 1995, т.165, №4, с. 361-401.

26. Короленко П.В. Оптические вихри. //Соросовский образовательный журнал, 1998, т. - , № - , с. - .

27. Короленко П.В., Федотов Н.Н., Шарков В.Ф. Основные свойства и перспективы использования лазеров на М-модах. //Квантовая электроника, 1995, т.22, №6, с. 562-566.

ГЛАВА 3. ЗАПИСЬ И ОБРАБОТКА ОПТИЧЕСКОЙ ИНФОРМАЦИИ

В настоящее время существует большое количество способов записи и обработки получаемой в когерентном свете оптической информации о структуре того или иного физического объекта. Самый распространенный из них состоит в получении с помощью оптической системы изображения интересующего объекта, его регистрации с использованием возможностей фото- и видеотехники и в последующей апостериорной обработке изображения. Другой способ, также получивший широкое распространение, основан на получении голограммы объекта. Этот способ, в отличие от первого, позволяет регистрировать информацию не только о распределении интенсивности света, отраженного или излучаемого объектом, но и о распределении фазы световых колебаний. Последнее обстоятельство создает дополнительные возможности по корректировке характеристик изображения. Однако свойства когерентных оптических систем , даже состоящих из традиционных оптических элементов (линзы, зеркала, киноформы, диафрагмы, маски и т.д.), не сводятся только к способности формировать оптические изображения. В ряде случаев их можно рассматривать как некие оптические процессоры, осуществляющие определенные математические преобразования, например, фурье-преобразования, по отношению к двумерной функции, определяющей распределение комплексной амплитуды на входе системы. Несмотря на кажущееся различие указанных вариантов использования оптических систем, их теоретическое описание имеет много общего.

В связи с этим изложение материала данной главы начнем с построения общей теории оптических систем.

3.1. Общая характеристика оптических систем [1-3]

В самом общем виде функциональная схема записи и обработки оптической информации приведена на *рис. 3.1.1*. Плоская монохроматическая волна 1 освещает объект 2, который размещают во входной (предметной) плоскости системы 3. Излучение, прошедшее объект или отраженное от него, попадает во входное отверстие (входной зрачок) 7 оптической системы 4. Пройдя элементы оптической системы, излучение выходит из выходного отверстия (выходного зрачка) 8 и формирует в выходной плоскости (плоскости изображений) 5 изображение объекта.

Вблизи плоскости изображений располагается светочувствительный элемент 6 регистрирующей системы (фотопластина, матрица фотоприемников и т.д.)

В тех случаях, когда оптическая система играет роль оптического процессора, у входной плоскости системы вместо объекта 2 располагается преобразователь входных сигналов. Он, пространственно модулируя падающую на него световую волну, преобразует информацию, поступающую от некоторого источника во входной оптический сигнал. В одном случае преобразователь может представлять собой слайд, на котором информация записана в виде изменяющего коэффициента пропускания. В другом случае в качестве преобразователя может использоваться слой жидкости, рельеф поверхности которой изменяется под действием

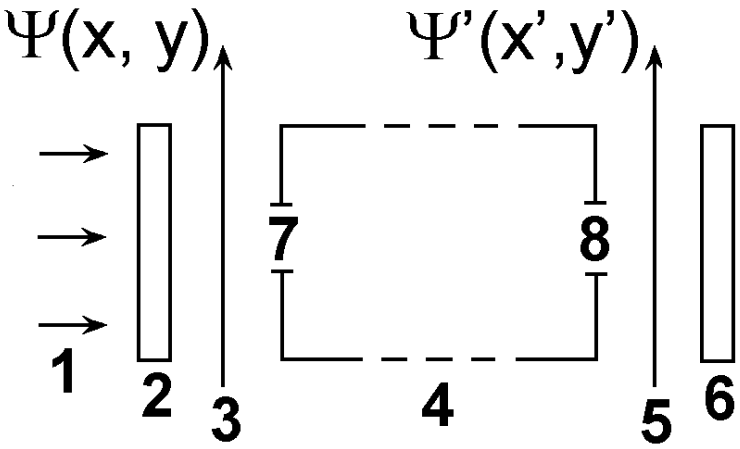


Рис. 3.1.1. *Схема оптической обработки информации*
1-плоская монохроматическая волна; 2-объект; 3-входная (предметная) плоскость; 4-оптическая система; 5-выходная плоскость (плоскость изображения); 6-светочувствительный элемент; 7-входное отверстие (входной зрачок); 8-выходное отверстие (выходной зрачок).

ультразвуковых волн или электронного пучка.

Свяжем с входной плоскостью оптической системы прямоугольную систему координат *(x,y)*, а с выходной - систему координат *(x',y')*. Если в системе отсутствуют нелинейные оптические элементы, между входным сигналом $\psi(x,y)$ и выходным сигналом $\psi'(x',y')$ (см. *рис. 3.1.1*) может быть записана в виде

$$\psi'(x',y') = \hat{L}\{\psi(x,y)\}, \qquad \textbf{(3.1.1)}$$

где $\hat{L}$ - некий линейный оператор. Линейность преобразования означает, что выходной сигнал от суммы входных сигналов равен сумме выходных сигналов от каждого входного сигнала в отдельности.

Используя свойство δ-функции Дирака, можно представить функцию $\psi(x,y)$ в виде

$$\psi(x,y) = \iint_{\infty} \psi(\xi,\eta)\delta(\xi - x,\eta - y)\,d\xi d\eta. \qquad \textbf{(3.1.2)}$$

Подставляя (3.1.2) в (3.1.1) и учитывая линейность оператора $\hat{L}$, можно записать

$$\begin{aligned} \psi'(x',y') &= \hat{L}\left\{\iint_{\infty} \psi(\xi,\eta)\delta(\xi - x,\eta - y)\,d\xi d\eta\right\} = \\ &= \iint_{\infty} \psi(\xi,\eta)\,\hat{L}\{\delta(\xi - x,\eta - y)\}\,d\xi d\eta. \end{aligned} \qquad \textbf{(3.1.3)}$$

Таким образом, выходной сигнал $\psi'(x',y')$ может быть представлен в виде суммы (интеграла) элементарных откликов $L\{\delta(\xi - x,\eta - y)\}$ с весовыми коэффициентами $\psi(x,y)$. Поскольку δ- функция $\delta(\xi - x,\eta - y)$ моделирует узкий и высокий "импульс" в точке $\xi = x, \eta = y$ (точечный источник света), функцию

$$h(x',y';\xi,\eta) = L\{\delta(\xi - x,\eta - y)\} \qquad \textbf{(3.1.4)}$$

называют импульсным откликом системы. Учитывая последнее обозначение, выражению (3.1.3) придаем вид

$$\psi'(x',y') = \iint_{\infty} \psi(x,y)h(x',y';x,y)\,dxdy. \qquad \textbf{(3.1.5)}$$

Импульсный отклик $h(x',y',x,y)$ называют также переходной функцией данной оптической системы, а выражение (3.1.5) получило название "интеграл суперпозиции".

Качество оптических систем во многом определяется тем, в какой степени им присуще свойство изопланарности. Под изопланарностью понимают инвариантность к пространственным смещениям, обеспечивающую выполнение соотношения

$$\hat{L}\{\psi(x+\Delta x, y+\Delta y)\} = \psi'(x'+\Delta x', y'+\Delta y'). \qquad \textbf{(3.1.6)}$$

Форма выходного сигнала в изопланарной системе, тем самым, не зависит от пространственных смещений входного сигнала. Хотя реальные оптические системы редко бывают изопланарными по всей плоскости

входного сигнала (изопланарность имеет место лишь на отдельных участках), все рассматриваемые системы мы в дальнейшем будем считать изопланарными. Это допущение позволяет нам представить переходную функцию системы $h(x',y',x,y)$ в более простом виде, считая ее зависящей от разности значений соответствующих координат

$$h(x',y';x,y) \to h(x'-x,y'-y). \qquad \textbf{(3.1.7)}$$

Интеграл суперпозиции (3.1.5) при этом принимает вид интеграла свертки

$$\psi'(x',y') = \iint_{\infty} \psi(x,y)h(x'-x,y'-y)dxdy. \qquad \textbf{(3.1.8)}$$

Применим к левой и правой части соотношения (3.1.8) преобразование Фурье

$$F\{\psi'(x',y')\} = F\{\iint_{\infty} \psi(x,y)h(x'-x,y'-y)dxdy\}. \qquad \textbf{(3.1.9)}$$

Из теоремы свертки следует, что

$$F\{\iint_{\infty} \psi(x,y)h(x'-x,y'-y)dxdy\} = F\{\psi(x,y)\}F\{h(x,y)\}, \textbf{(3.1.10)}$$

поэтому выражение (3.1.9) может быть представлено в виде

$$F\{\psi'\} = F\{\psi\}F\{h\} \qquad \textbf{(3.1.11)}$$

Введем следующие обозначения:

$$F\{\psi'\} = S'(f_x,f_y), \quad F\{\psi\} = S(f_x,f_y), \quad F\{h\} = H(f_x,f_y). \qquad \textbf{(3.1.12)}$$

Функцию $H(f_x,f_y)$ принято называть передаточной функцией системы.

Перепишем выражение (3.1.11) с учетом обозначений (3.1.12)

$$S(f_x,f_y)H(f_x,f_y) = S'(f_x,f_y). \qquad \textbf{(3.1.13)}$$

Из последнего соотношения видно, что спектр пространственных частот выходного сигнала равен произведению спектра частот входного сигнала и передаточной функции. Выражение (3.1.13), как мы убедимся ниже, играет фундаментальную роль в теории систем записи и обработки оптической информации. Теоретический анализ мы начнем с простейшей оптической системы, состоящей всего лишь из одной линзы.

3.2. Однолинзовая система [1-4]

3.2.1. Линзы как элементы, выполняющие преобразование Фурье

Пусть оптическая система состоит из одного элемента - тонкой идеальной (без аберраций) линзы. Покажем, что такая простейшая система может выполнять функцию оптического процессора, выполняющего преобразование Фурье. Будем считать, что линза с фокальным расстоянием F и апертурой D располагается в плоскости (ξ, η) между входной *(x,y)* и выходной *(x',y')* плоскостями соответственно на расстоянии d_0 и d_1 (см. *рис. 3.2.1*).

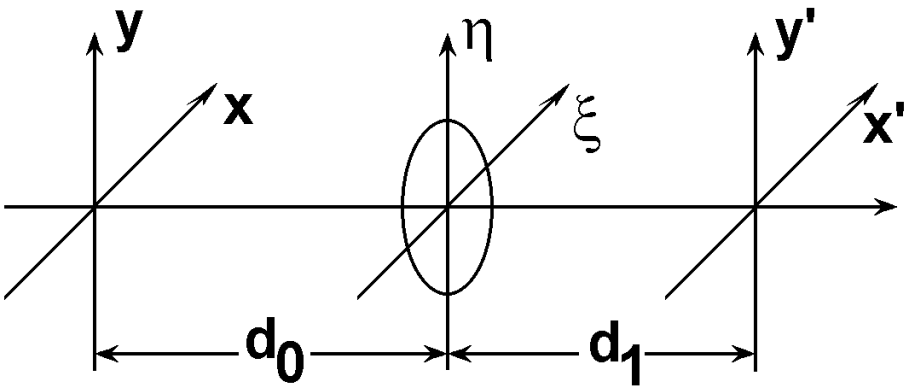


Рис.3.2.1. *Однолинзовая система.*

Как уже ранее отмечалось, линза является элементом, осуществляющим квадратичную фазовую модуляцию. Это означает, что распределение поля падающей на линзу волны $\Psi_\xi(\xi,\eta)$ будет связано с распределением поля световых колебаний за линзой $\Psi'_\xi(\xi,\eta)$ соотношением

$$\Psi'_\xi(\xi,\eta) = \Psi_\xi(\xi,\eta)T(\xi,\eta) \qquad \textbf{(3.2.1),}$$

где так называемая модуляционная характеристика линзы $T(\xi, \eta)$ равна

$$T(\xi,\eta) = \exp\left[-ik\frac{\xi^2+\eta^2}{2F}\right]P(\xi,\eta), \qquad \textbf{(3.2.2)}$$

где

$$P(\xi,\ \eta)=\begin{cases}1 \quad \text{при} \quad \sqrt{\xi^2+\eta^2}\le\dfrac{D}{2},\\ 0 \quad \text{при} \quad \sqrt{\xi^2+\eta^2}>\dfrac{D}{2}.\end{cases} \qquad \textbf{(3.2.3)}$$

Рассчитаем теперь переходную функцию однолинзовой системы. Для простоты будем считать, что апертура линзы существенно превосходит апертуру падающего на него светового пучка. Возьмем за основу формулу (1.2.40) гл. 1, исключив из нее легко учитываемый постоянный фазовый множитель *-i*exp*(ikz)*. В соответствии с этой формулой входной сигнал $\psi(x,y)$ после прохождения расстояния d_0 преобразуется в сигнал

$$\psi_\xi(\xi,\eta)=\frac{1}{\lambda d_0}\iint_\infty \psi(x,\ y)\exp\left\{ik\frac{(\xi-x)^2+(\eta-y)^2}{2d_0}\right\}dxdy. \quad \textbf{(3.2.4).}$$

Согласно (3.2.1 - 3.2.3), сразу за линзой сигнал принимает вид

$$\psi'_\xi(\xi,\eta)=\psi_\xi(\xi,\eta)\exp\left\{ik\frac{\xi^2+\eta^2}{2F}\right\}. \qquad \textbf{(3.2.5).}$$

Еще раз воспользовавшись формулой (1.2.40), получаем выражение для сигнала в выходной плоскости (после прохождения на расстоянии d_1)

$$\psi'(x',y')=\frac{1}{\lambda d_1}\iint_\infty \psi_\xi'(\xi,\eta)\exp\left\{ik\frac{(\xi-x')^2+(\eta-y')^2}{2d_1}\right\}d\xi d\eta. \textbf{(3.2.6).}$$

Используя выражения (3.2.4-3.2.6), запишем теперь в явном виде связь между входным $\psi(x,y)$ и выходным $\psi(x',y')$ сигналами

$$\psi'(x',y')=\frac{1}{\lambda^2 d_0 d_1}\iint\limits_\infty\iint \psi(x,y)\exp\left\{\mathrm{ik}\frac{(\xi-x)^2+(\eta-y)^2}{2d_0}\right\}\times$$
$$\times\exp\left\{-\mathrm{ik}\frac{\xi^2+\eta^2}{2F}\right\}\exp\left\{\mathrm{ik}\frac{(\xi-x')^2+(\eta-y')^2}{2d_1}\right\}d\xi d\eta dxdy. \qquad \textbf{(3.2.7).}$$

Отсюда видно, что переходная функция однолинзовой системы равна

$$\mathrm{h}=\frac{1}{\lambda^2 d_0 d_1}\iint_\infty \exp\left\{\mathrm{ik}\frac{(\xi-x)^2+(\eta-y)^2}{2d_0}\right\}\times$$
$$\times\exp\left\{-\mathrm{ik}\frac{\xi^2+\eta^2}{2F}\right\}\exp\left\{\mathrm{ik}\frac{(\xi-x')^2+(\eta-y')^2}{2d_1}\right\}d\xi d\eta. \qquad \textbf{(3.2.8)}$$

Положим $d_0=d_1=F$. В этом случае (3.2.8) принимает вид

$$h = \frac{1}{\lambda^2 F^2}\exp\left\{ik\frac{x^2 + x'^2 + y^2 + y'^2}{2F}\right\}\times$$
$$\times\iint_{\infty}\exp\left\{ik\frac{x^2 + h^2}{2F}\right\}\exp\left\{-ik\frac{x(x+x') + h(y+y')}{F}\right\}dxdh. \quad \textbf{(3.2.9)}$$

Используя далее соотношение

$$\iint_{\infty}\exp[\mathrm{ia}(\xi^2+\eta^2)]\exp[\mathrm{ib}(\mathrm{u}\xi+\mathrm{v}\eta)]\mathrm{d}\xi\mathrm{d}\eta = \quad \textbf{(3.2.10)}$$
$$= \frac{\pi}{\mathrm{a}}\exp\left[-\mathrm{ib}^2\,\frac{u^2+v^2}{4a}\right]$$

и, полагая $pa = \frac{k}{2F}$, $b = -\frac{k}{F}$, $u = x + x'$, $v = y + y'$, получаем для переходной функции простое выражение

$$\mathrm{h} = \frac{1}{\lambda\mathrm{F}}\exp\left[\frac{-\mathrm{i}2\pi(\mathrm{xx'}+\mathrm{yy'})}{\lambda\mathrm{F}}\right]. \quad \textbf{(3.2.11)}$$

Если теперь ввести обозначения

$$\frac{\mathrm{x'}}{\lambda\mathrm{F}} = \mathrm{f_x}, \quad \frac{\mathrm{y}}{\lambda\mathrm{F}} = \mathrm{f_y}, \quad \textbf{(3.2.12)}$$

то интеграл суперпозиции (3.2.7) можно привести к виду

$$\mathrm{S}(\lambda\mathrm{F}f_{\mathrm{x}}, \lambda\mathrm{F}f_{\mathrm{y}}) = \frac{1}{\lambda\mathrm{F}}\iint_{\infty}\psi(\mathrm{x},\mathrm{y})\,\mathbf{exp}\left[-i2\pi\left(f_x x + f_y y\right)\right]dxdy. \quad \textbf{(3.2.13)}$$

Отсюда видно, что при выполнении условия $d_0=d_1=F$ линза выполняет Фурье-преобразование сигнала: ее задняя фокальная плоскость является спектральной плоскостью входного сигнала. Таким образом, линза может предельно просто выполнять математическую операцию, представляющую трудность даже для сложных электронных устройств.

Заметим, что спектр с точностью до легко учитываемого фазового множителя будет формироваться в фокальной плоскости даже в том случае, когда $d_0>F$. При этом условии выражение (3.2.7) приводится к виду (промежуточные выкладки мы опускаем)

$$\psi'(\mathrm{x}',\mathrm{y}') = \frac{1}{\lambda\mathrm{F}}\exp\left[\mathrm{ik}\left(1-\frac{\mathrm{d}_0}{\mathrm{F}}\right)\frac{x'^2+y'^2}{2F}\right]\times$$
$$\times\iint_{\infty}\psi(\mathrm{x},\mathrm{y})\exp\left[-\mathrm{ik}\,\frac{\mathrm{xx'}+\mathrm{yy'}}{\mathrm{F}}\right]\mathrm{dxdy}. \quad \textbf{(3.2.14)}$$

Фазовый множитель

$$\exp\left[ik\left(1-\frac{d_0}{F}\right)\frac{x'^2+y'^2}{2F}\right], \qquad \textbf{(3.2.15)}$$

не зависящий от вида входного сигнала, легко учитывается при последующей обработке.

3.2.2. Формирование изображения [1]

Конечно, наиболее известным свойством линз является их способность формировать изображение. Если предмет пемещен перед линзой и освещен, то при определенных условиях в другой плоскости возникает распределение интенсивности, которое очень напоминает предмет. Это распределение называется изображением предмета. Изображение может быть действительным в том смысле, что в плоскости за линзой возникает действительное распределение, и мнимым в том смысле, что свет за линзой кажется исходящим из новой плоскости, расположенной перед линзой.

Предположим, что плоский предмет, находящийся на расстоянии d_0 перед положительной линзой, освещен монохроматическим светом. Обозначим комплексное поле непосредственно за предметом через $\psi(x,y)$ Распределение поля, которое возникает на расстоянии d_1 за линзой, обозначим через $\psi'(x',y')$. Наша задача - определить условия, при которых распределение поля ψ' можно с уверенностью назвать "изображением" распределения поля в плоскости предмета ψ.

Ввиду линейности явления распространения волн поле ψ' можно представить в виде интеграла суперпозиции (3.1.5). Тем самым свойства системы, создающей изображение, будут полностью описаны с помощью импульсного отклика h (см. (3.2.8)).

Чтобы оптическая система давала высококачественное изображение, поле ψ' должно как можно меньше отличаться от ψ. Это означает, что импульсный отклик должен приближенно походить на δ-функцию, т.е.

$$h(x',y';x,y) \approx K\delta(x'\pm Mx, y'\pm My), \qquad \textbf{(3.2.16)}$$

где K - комплексная постоянная, M - увеличение системы, а знак плюс или минус учитывает возможность как прямого, так и обратного изображения. Поэтому "плоскостью изображения" мы будем называть ту плоскость, где (3.2.16) выполняется лучше всего.

Преобразуем (3.2.8) к виду

$$h(x',y';x,y)=\frac{1}{\lambda^2 d_0 d_1}\exp\left[i\frac{k}{2d_1}\left(x'^2+y'^2\right)\right]\times$$
$$\times\exp\left[i\frac{k}{2d_1}\left(x^2+y^2\right)\right]\int_{-\infty}^{\infty}\int P(\xi,\eta)\times$$
$$\times\exp\left[i\frac{k}{2}\left(\frac{1}{d_0}+\frac{1}{d_1}-\frac{1}{f}\right)\left(\xi^2+\eta^2\right)\right]\times \qquad \textbf{(3.2.17)}$$
$$\times\exp\left\{-ik\left[\left(\frac{x}{d_0}+\frac{x'}{d_1}\right)x+\left(\frac{y}{d_0}+\frac{y'}{d_1}\right)y\right]\right\}d\xi d\eta.$$

Соотношения (3.2.17) играет важную роль при определении зависимости между ψ и ψ'. Однако без дальнейших упрощений трудно определить условия, при которых распределение ψ' можно с уверенностью назвать изображением распределения ψ.

Самые неприятные члены в приведенном выше выражении для импульсного отклика - это члены, содержащие квадратичные фазовые множители. Заметим, что два из них:

$$\exp\left[i\frac{k}{2d_1}\left(x'^2+y'^2\right)\right];\quad \exp\left[i\frac{k}{2d_1}\left(x^2+y^2\right)\right] \qquad \textbf{(3.2.18)}$$

не зависят от координат линзы (ξ, η). Эти члены определяют фазовое искривление в плоскостях *(x',y')* и *(x,y)*. Если бы мы решили рассматривать формирование изображения между двумя сферическими поверхностями, а не между двумя плоскостями, эти члены можно было бы исключить. Однако можно показать, что и в случае формирования изображения между двумя плоскостями оба эти члена несущественны.

Опуская множитель $\exp\left[i\frac{k}{2d_1}\left(x'^2+y'^2\right)\right]$, заметим, что в подавляющем большинстве представляющих интерес случаев конечной целью задачи формирования изображения является получение некоторого распределения света, которое будет воспринято детектором, реагирующим только на интенсивность (например, фотопленкой). Так как рассматриваемый член изменяет только распределение фазы, он никак не будет влиять на результаты измерения интенсивности и, следовательно, может быть опущен.

К сожалению, от фазового множителя $\exp\left[i\frac{k}{2d_1}\left(x^2+y^2\right)\right]$ не удается освободиться столь же просто, поскольку он зависит от переменных интегрирования *(x,y)* интеграла суперпозиции. Однако в большинстве случаев, представляющих интерес, от него тоже можно избавиться. Если система, создающая изображение, ведет себя приблизительно так же, как идеальная система, для которой справедливо соотношение (3.2.16), то амплитуда волны в точке с координатами *(x',y')* будет определяться вкладом только очень малой области в пространстве предмета с центром в точке, соответствующей идеальному геометрическому изображению (*рис. 3.2.2*).

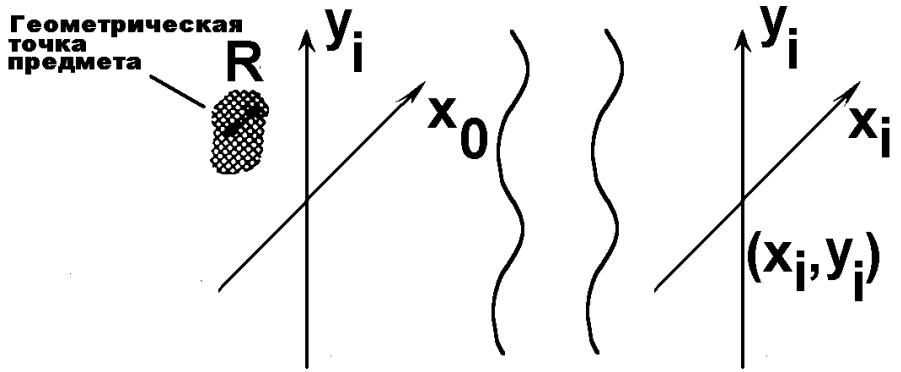


Рис.3.2.2. *Область R, в которой функция h для точки с координатами (x',y') имеет значительную величину.*

Если внутри этой малой области аргумент $\exp\left[i\frac{k}{2d_1}\left(x^2+y^2\right)\right]$ изменяется не более чем на долю радиана, то можно использовать приближение

$$\exp\left[i\frac{k}{2d_0}\left(x^2+y^2\right)\right]\approx\exp\left[i\frac{k}{2d_0}\left(\frac{x'^2+y'^2}{M^2}\right)\right]. \qquad \textbf{(3.2.19)}$$

Теперь экспоненциальный член можно опустить, так как он не зависит от *(x, y)* и следовательно, не влияет на результат измерения интенсивности в плоскости *(x',y')*.

Воспользовавшись приведенными соображениями, перепишем выражение для импульсного отклика в виде

$$h(x',y';x,y)\approx$$

$$\approx\frac{1}{\lambda^2 d_0 d_1}\int\limits_{-\infty}^{\infty}\int P(\xi,\eta)\exp\left[i\frac{k}{2}\left(\frac{1}{d_0}+\frac{1}{d_1}-\frac{1}{f}\right)\times\left(\xi^2+\xi^2\right)\right]\times \quad \textbf{(3.2.20)}$$

$$\times\exp\left\{-ik\left[\left(\frac{x}{d_0}+\frac{x'}{d_1}\right)\xi+\left(\frac{y}{d_0}+\frac{y'}{d_1}\right)\eta\right]\right\}d\xi d\eta.$$

Чтобы получить совсем простой результат, рассмотрим случай, когда плоскость наблюдения расположена на таком расстоянии d_1 от линзы, что удовлетворяется соотношением

$$\left(\frac{1}{d_0}+\frac{1}{d_1}-\frac{1}{f}\right)=0. \quad \textbf{(3.2.21)}$$

Это соотношение известно из геометрической оптики, где оно называется формулой линзы. Соотношение (3.2.21) определяет расположенную за линзой точку, в которой пересекаются лучи, исходящие из одной точки предмета (точка изображения). В приближении геометрической оптики выполнение формулы линзы означает, что импульсный отклик системы достаточно близок к идеальному. Предположение о выполнении формулы линзы позволяет свести импульсный отклик к виду

$$h(x',y';x,y)\approx\frac{1}{\lambda^2 d_0 d_1}\int\limits_{-\infty}^{\infty}\int P(\xi,\eta)\times$$

$$\times\exp\left\{-ik\left[\left(\frac{x}{d_0}+\frac{x'}{d_1}\right)\xi+\left(\frac{y}{d_0}+\frac{y'}{d_1}\right)\eta\right]\right\}d\xi d\eta. \quad \textbf{(3.2.22)}$$

Определяя увеличение системы как

$$M=\frac{d_1}{d_0}, \quad \textbf{(3.2.23)}$$

находим последний упрощенный вид импульсного отклика

$$h(x',y';x,y)\approx\frac{1}{\lambda^2 d_0 d_1}\int\limits_{-\infty}^{\infty}\int P(\xi,\eta)\times$$

$$\times\exp\left\{-i\frac{2\pi}{\lambda d_1}\left[\left(x'+Mx\right)\xi+\left(y'+My\right)\eta\right]\right\}d\xi d\eta. \quad \textbf{(3.2.24)}$$

Таким образом, если формула линзы справедлива, то импульсный отклик соответствует картине дифракции Фраунгофера на апертуре линзы, причем центр картины находится в точке изображения $(x_i=-Mx,\ y_i=-My)$.

Перейдем теперь к анализу соотношения между предметом и изображением. Рассмотрим сначала свойства изображения, предсказываемые геометрической оптикой. Чтобы найти это идеальное изображение, допустим, что длина волны λ стремится к нулю. В этом случае дифракционные эффекты становятся несущественными. Производя замену переменных

$$\tilde{\xi}=\frac{\xi}{\lambda d_1},\ \ \tilde{\eta}=\frac{\eta}{\lambda d_1}, \qquad \textbf{(3.2.25)}$$

выражение для импульсного отклика (3.2.24) можно переписать в виде

$$h(\xi',\eta';\xi,\eta)\approx M\iint_{\infty} P(\lambda d_1\tilde{\xi},\lambda d_1\tilde{\eta})\times$$
$$\times\exp\left\{-i2\pi\left[(\xi'+M\xi)\tilde{\xi}+(\eta'+M\eta)\tilde{\eta}\right]\right\}d\tilde{\xi}d\tilde{\eta}. \qquad \textbf{(3.2.26)}$$

Так как λ стремится к нулю, область значений $(\tilde{\xi},\tilde{\eta})$, где функция зрачка P равна единице, будет безгранично увеличиваться, что дает возможность заменить P единицей, оставив те же пределы интегрирования. Таким образом,

$$h(x',y';x,y)\to$$
$$\to M\iint_{\infty}\exp\left\{-i2\pi\left[(x'+Mx)\tilde{\xi}+(y'+My)\tilde{\eta}\right]\right\}d\tilde{\xi}d\tilde{\eta}= \qquad \textbf{(3.2.27)}$$
$$M\delta(x'+Mx,y'+My)=\frac{1}{M}\delta\left(\frac{x'}{M}+x,\frac{y'}{M}+y\right).$$

Представляя этот результат в интеграл суперпозиции (3.2.16), получаем соотношение, связывающее распределения амплитуды в точках предмета и в точках изображения

$$\psi'(x',y')=\frac{1}{M}\psi\left(-\frac{x'}{M},-\frac{y'}{M}\right). \qquad \textbf{(3.2.28)}$$

Отсюда следует, что изображение, получаемое в приближении геометрической оптики, представляет собой точную копию изображения.

Выводы геометрической оптики, конечно, приближенны. Более точное представление о соотношении между предметом и изображением можно получить только при учете дифракционных эффектов. Чтобы найти

такое соотношение, вернемся к выражению (3.2.26) для импульсного отклика и произведем следующую дополнительную замену переменных:

$$\tilde{x} = -Mx, \tilde{y} = -My. \qquad \textbf{(3.2.29)}$$

Импульсный отклик в этом случае будет равен

$$h(x', y'; \tilde{x}, \tilde{y}) = M \iint_{\infty} P(\lambda d_1 \tilde{\xi}, \lambda d_1 \tilde{\eta}) \times \times \exp\{-i2\pi[(x' - \tilde{x})\tilde{\xi} + (y' - \tilde{y})\tilde{\eta}]\} d\tilde{x} d\tilde{y}. \qquad \textbf{(3.2.30)}$$

Заметим, что h теперь пространственно-инвариантная величина, зависящая только от разности координат $(x' - \tilde{x}, y' - \tilde{y})$. С введением еще одного определения

$$\tilde{h} = \frac{1}{M} h \qquad \textbf{(3.2.31)}$$

распределение поля в плоскости изображения принимает вид

$$\Psi'(x', y') = \iint_{\infty} \tilde{h}(x' - \tilde{x}, y' - \tilde{y}) \times \times \left[\frac{1}{M} \Psi\left(-\frac{\tilde{x}}{M}, -\frac{\tilde{y}}{M}\right)\right] d\tilde{x} d\tilde{y}. \qquad \textbf{(3.2.32)}$$

В этом выражении мы узнаем свертку импульсного отклика $\tilde{h}$ и идеального изображения. Для удобства определим новую функцию

$$\Psi_{\text{i}}(x', y') = \frac{1}{M} \Psi\left(-\frac{x'}{M}, -\frac{y'}{M}\right). \qquad \textbf{(3.2.33)}$$

Свертку (3.2.32) тогда можно переписать в упрощенном виде:

$$\psi'(x', y') = \tilde{h}(x', y') * \psi_i(x', y'), \qquad \textbf{(3.2.34)}$$

где

$$\tilde{h}(x', y') = \iint_{\infty} P(\lambda d_i \tilde{\xi}, \lambda d_i \tilde{\eta}) \times \times \exp\{-i2\pi[x'\tilde{\xi} + y'\tilde{\eta}]\} d\tilde{\xi} d\tilde{\eta}. \qquad \textbf{(3.2.35)}$$

Соотношения (3.2.34) и (3.2.35) представляют собой конечный результат настоящего анализа. Они показывают, что при учете дифракционных эффектов изображение нельзя больше считать точной копией предмета. Полученное изображение дает несколько сглаженный облик предмета, что является следствием неравенства нулю ширины импульсного отклика $\tilde{h}$. Это сглаживание может привести к значительному ослаблению мелких деталей предмета и соответственно к потере точности воспроизведения изображения. Точно такое же явление можно наблюдать в случае, когда

электрический сигнал проходит через линейную электрическую схему. Если длительность импульсного отклика схемы велика по сравнению с "временем пульсаций" входного сигнала, то схема будет сглаживать входной сигнал. Таким образом, быстрые изменения входного сигнала не будут воспроизводиться на выходе.

3.3. Получение изображений в сложных системах [1, 8]

3.3.1. Дифракционно-ограниченные системы. Теории Аббе и Релея

Предположим, что рассматриваемая оптическая система состоит не из одной линзы, а из нескольких линз, среди которых могут быть как положительные, так и отрицательные. Линзы могут и не быть "тонкими". Будем предполагать, однако, что система в конечном счете дает действительное изображение, но фактически это не ограничение, так как если система дает мнимое изображение, то оно может быть преобразовано в итоге в действительное, например глазом. Значит, в подобном случае нам следует включить глаз в качестве конечного элемента в нашу систему.

При рассмотрении свойств системы линз будем считать, что все элементы, создающие изображение, помещены в один "черный" ящик" и что основные свойства системы можно полностью описать, определяя только конечные свойства этого устройства.

Согласно *рис. 3.1.1*, входом этого черного ящика служит входной зрачок, представляющий собой отверстие конечных размеров (эффективнное или действительное), через которое должен проходить свет прежде, чем он достигнет элементов, создающих изображение, а выходом - выходной зрачок (также эффективный или действительный), представляющий собой отверстие конечных размеров, через которое свет проходит после создающих изображение элементов на пути к плоскости изображения. Обычно считают, что путь света между входной и выходной плоскостями может быть достаточно полно описан в приближениях геометрической оптики. Таким образам, конечный размер обоих зрачков можно найти, строя по законам геометрической оптики проекцию наименьшей апертуры системы соответственно в пространстве предметов и пространстве изображений. Поскольку размеры получающихся зрачков определяются размерами изображения эффективного отверстия, существующего где-то внутри системы, они могут быть меньше действительных физических размеров отверстий в входной и выходной

плоскостях. Заметим, что при таком определении входной зрачок всегда является изображением выходного зрачка и наоборот.

Оптическая система называется дифракционно ограниченной, если она преобразует расходящуюся сферическую волну, исходящую из любого точечного источника, в новую идеальную сферическую волну, которая сходится в точке, лежащей в плоскости изображения. Таким образом, конечное свойство дифракционно ограниченной системы линз заключается в том, что она преобразует расходящуюся сферическую волну, падающую на входной зрачок, в сходящуюся сферическую волну, выходящую через выходной зрачок. Для любой реальной оптической системы это свойство выполняется в лучшем случае только для конечной области в плоскости предмета. Если рассматриваемый предмет не выходит за пределы этой области, систему можно отнести к дифракционно ограниченной. Если в действительности фронт волны от точечного источника после выходного зрачка значительно отличается от идеальной сферической формы, то говорят, что оптическая система имеет аберрации.

Геометрическая оптика с достаточной точностью описывает прохождение света от входного зрачка к выходному, поэтому дифракционные эффекты играют заметную роль только на пути света от предмета к входному зрачку и от выходного зрачка к изображению. Действительно, все ограничения, налагаемые дифракцией, можно связать с любым из этих двух участков пути распространения света. Утверждения о том, что разрешение изображения ограничивается входным зрачком конечных размеров или выходным зрачком конечных размеров, полностью эквивалентны. Основная причина эквивалентности заключается в том, что один зрачок представляет собой просто изображение другого.

Представление о том, что обсуждаемые дифракционные эффекты обусловлены входным зрачком конечных размеров, было впервые введено Эрнстом Аббе в 1873г. Согласно теории Аббе, только определенная часть дифракционных максимумов, созданных сложным предметом, пропускается входным зрачком конечных размеров. Не пропускаются зрачком те максимумы, которые соответствуют высокочастотным составляющим предмета. Это положение иллюстрирует *рис. 3.3.1*, где предметом служит простая решетка, а оптическая система состоит из одной положительной линзы.

В 1896 г. Релей высказал другую, фактически эквивалентную точку зрения, в соответствии с которой дифракционные эффекты обусловлены выходным зрачком конечных размеров. Тем самым, на сложные оптические системы могут быть перенесены соотношения (3.2.34) и (3.2.35), полученные для однолинзовой системы, путем замены апертурной функции

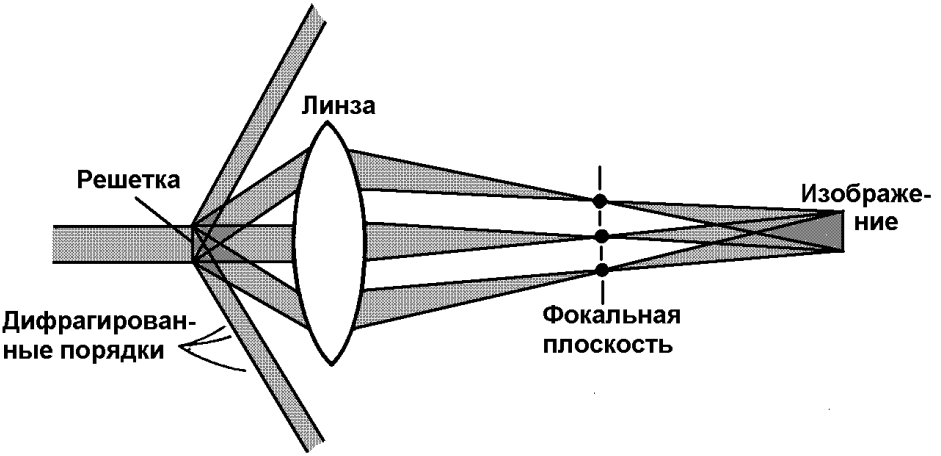


Рис. 3.3.1. *Формирование изображения по Аббе*

линзы P на апертурную функцию выходного зрачка системы. К сложным системам, формирующим изображение непосредственно применимо понятие передаточной функции.

Передаточная функция H определяется как фурье-образ переходной функции $\widetilde{h}$, которая в свою очередь определяется преобразованием Фурье от фукции выходного зрачка (переходная функция будет выражаться формулой (3.2.35) в предположении,что функция P относится к выходному зрачку)

Таким образом, мы приходим к выводу, что для дифракционно ограниченной системы

$$H\left(f_x, f_y\right) = F\left\{F\left\{P\left(\lambda d_1 \widetilde{\xi}, \lambda d_1 \widetilde{\eta}\right)\right\}\right\} = P\left(-\lambda d_1 f_x, -\lambda d_1 f_y\right). \quad \textbf{(3.3.1)}$$

Это крайне важное соотношение, так как оно дает информацию относительно поведения дифракционно ограниченных когерентных систем в частотной области. Так как функция зрачка P всегда равна или единице или нулю, то же самое справедливо и для передаточной функции. Это, естественно, означает, что в частотной области дифракционно ограниченная система имеет конечную полосу пропускания, внутри которой все частотные составляющие пропускаются без искажения амплитуды и фазы. На границе этой полосы пропускания частотный отклик сразу падает до нуля, в силу чего частотные составляющие вне полосы пропускания полностью подавляются.

3.3.2. Роль аподизации

Поскольку функция зрачка системы играет принципиальную роль в формировании структуры изображения, возникает вопрос о возможности подбора такого амплитудного пропускания зрачка системы, при котором ослабляются боковые лепестки дифракционной картины резко очерченной диафрагмы. Появление боковых лепестков в дифракционной картине аналогично эффекту оптических выбросов или эффекту Гиббса. Как известно, эффект Гиббса полностью исчезает, если от зрачка, амплитудное пропускание которого описывается прямоугольным импульсом, перейти к зрачку, описываемому треугольным импульсом. Наиболее подходящей формой зрачка является такая, амплитудное пропускание которой описывается функцией Гаусса. Действительно, в этом случае картина дифракции далекого поля описывается фурье- образом зрачка, а фурье-образ функции Гаусса равен функции Гаусса. Боковые лепестки при этом полностью исчезают. Процесс аподизации сопровождается неизбежным уширением основного пика дифракционной картины.

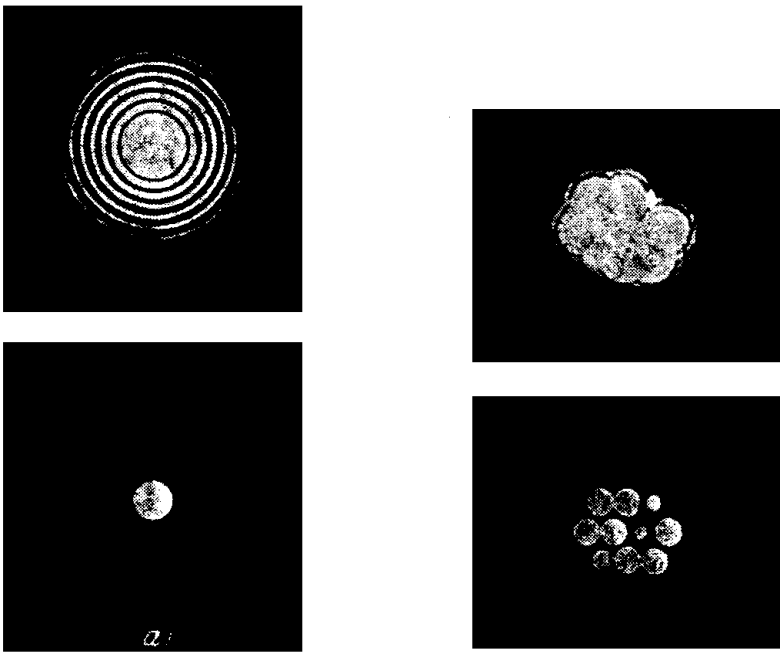


Рис. 3.3.2. *Влияние аподизации на вид картины дифракции далекого поля. Вверху - дифракция без аподизации; внизу - после аподизации.*

Метод аподизации, основанный на сглаживании функции пропускания зрачка системы, является весьма эффективным способом улучшения пространственной структуры оптического сигнала. На *рис. 3.3.2* приведены результаты аподизации картины дифракции далекого поля. С помощью аподизации удается разделить изображение двух близко расположенных предметов, когда они очень сильно отличаются между собой по интенсивности.

3.4. Учет аберраций [5]

Если оптическая система является дифракционно ограниченной, то импульсный отклик (при когерентном освещении), как мы видели, представляет собой картину дифракции Фраунгофера на выходном отверстии с центром в точке идеального изображения. Это обстоятельство подсказывает удобный прием, который позволит непосредственно учесть аберрации в наших предыдущих результатах. В частности, в случае искажения волнового фронта можно представить, что выходной зрачок освещается идеальной сферической волной, но в пределах отверстия находится фазовая пластинка, деформирующая выходящий из зрачка фронт волны. Если фазовая ошибка в точке (ξ,η) выходного зрачка изображается как $kW(x,h)$, где $k = 2\pi / \lambda$, а *W*- эффективная погрешность длины пути, то комплексный коэффициент пропускания *P* воображаемой фазовой пластинки можно представить в виде

$$\widetilde{P}(\xi,\eta) = P(\xi,\eta)\exp\left[jkW(\xi,\eta)\right]. \qquad \textbf{(3.4.1)}$$

Комплексную функцию *P* можно считать обобщенной функцией зрачка. Импульсный отклик когерентной системы с аберрациями представляет собой просто картину дифракции Фраунгофера на отверстии с коэффициентом пропускания *P*.

Передаточная функция при наличии аберрации запишется следующим образом:

$$H\left(f_\xi, f_\eta\right) = \widetilde{P}\left(\lambda d_1 f_\xi, \lambda d_1 f_\eta\right) = \\ = P\left(\lambda d_1 f_\xi, \lambda d_1 f_\eta\right)\exp\left[jkW\left(\lambda d_1 f_\xi, \lambda d_1 f_\eta\right)\right], \qquad \textbf{(3.4.2)}$$

где f_ξ и f_η - пространственные частоты в направлениях ξ и η.

Очевидно, что в случае когерентного освещения ограничение полосы пропускания передаточной функции, которое обусловлено конечным размером выходного зрачка, не зависит от наличия аберраций. Аберрации вводят только фазовые искажения в пределах полосы

пропускания. Разумеется, фазовые искажения могут оказывать определенное влияние на точность воспроизведения.

В теории оптических систем эффективную погрешность длины пути W принято оценивать по расстояниям отделяющих волновой фронт точечного источника вблизи выходного зрачка реальной системы (с аберрациями) от сферического волнового фронта, формируемого идеальной системой (без аберраций). Тем самым, при оценке аберраций сферический волновой фронт идеальной системы выступает в качестве сферы сравнения. Если ввести в плоскости выходного зрачка полярную систему координат с радиальной координатой $r = \dfrac{\sqrt{\xi^2 + \eta^2}}{R}$ (R - внешняя апертура выходного зрачка) и азимутальной координатой φ, то W можно представить в виде:

$$W = dr^2 + s_1 r^4 + s_2 r^6 + \left(Kr + C_1 r^3 + C_2 r^5\right)\cos\varphi + \\ + ar^2 \cos 2\varphi. \qquad \textbf{(3.4.3)}$$

Коэффициенты в этом выражении имеют следующий смысл:

d представляет дефокусировку или кривизну поля, т.е. максимальное отклонение на краю зрачка *(r=1)* сферы сравнения от волновой поверхности (которая является сферической, если эта аберрация единственная);

s_1 - коэффициент сферической аберрации 3-го порядка. Это максимальное отклонение на краю зрачка деформированной волновой поверхности от сферы сравнения, имеющей центр в параксиальном фокусе (острие геометрической каустики);

s_2 - коэффициент сферической аберрации 5-го порядка при тех же условиях;

K - параметр, соответствующий возможному покачиванию сферы сравнения;

C_1 - коэффициент комы 3-го порядка, т.е. максимальное отклонение на краю отверстия от сферы сравнения, центр которой совпадает с параксиальным фокусом волновой поверхности, соответствующей коме (если эта аберрация единственная);

C_2 - коэффициент комы 5-го порядка при тех же условиях;

a - коэффициент астигматизма, т.е. максимальное отклонение астигматической волновой поверхности от сферы сравнения.

3.5. Голографическая запись информации [2, 6-9]

3.5.1. Принцип голографической записи.

В отличие от фотографирования, регистрирующего лишь интенсивность волны, идущей от объекта, метод голографии позволяет записывать как амплитуду, так и фазу световых колебаний в плоскости наблюдения. В основе голографической записи лежит идея, согласно которой для выявления фазовой информации надо создать интерференцию исследуемой (объектной) волны с некоторой вспомогательной (опорной) волной. Будем описывать световые колебания объектной O и опорной R волн в точке M, лежащей в плоскости наблюдения, соответственно в виде выражений

$$O = o(M)\exp\{i\varphi_o(M)\}, \qquad \textbf{(3.5.1)}$$

$$R = r(M)\exp\{i\varphi_r(M)\}, \qquad \textbf{(3.5.2)}$$

o, r и φ_O, φ_r - соответственно амплитуды и фазы колебаний. Если в плоскости наблюдения расположить фотопластинку с идеальной линейной по интенсивности фотоэмульсией, то она зарегистрирует следующее распределение интенсивности:

$$I = |O+R|^2 = |O|^2 + |R|^2 + OR^* + O^*R. \qquad \textbf{(3.5.3)}$$

После соответствующей обработки пропускание пластинки станет пропорционально I.

Если теперь эту пластинку осветить опорной волной R, то поле на выходе пластинки будет иметь вид

$$RI = \left(|O|^2 + |R|^2\right)R + O|R|^2 + O^*R^2. \qquad \textbf{(3.5.4)}$$

Последнее выражение составляет теоретическую основу голографии. В нем первое слагаемое $\left(|O|^2 + |R|^2\right)R$ представляет собой поле опорной волны R, амплитуда которой промодулирована коэффициентом $\left(|O|^2 + |R|^2\right)$; второе слагаемое $|R|^2 O = r^2 o\exp\{i\varphi_o\}$ - поле объектной волны, амплитуда которой промодулирована коэффициентом r^2. Если в качестве опорной волны используется плоская волна, то $r(x,y) = r_o$ и модуляция исчезает; объектная волна при этом равномерно ослабевает в соответствии с коэффициентом $r_o{}^2$. Наконец, третье слагаемое в выражении (3.5.4) $O^*R^2 = r^2 o\exp\{2\varphi_r - \varphi_0\}$ - описывает волну, комплексно-сопряженную с объектной волной. Оно несет информацию, хотя и очень близкую к

информации об объекте, но отличающуюся от объектной волны наличием обратной фазы.

Остановимся теперь на конкретных схемах голографической записи и восстановления изображений. Наибольшее распространение получила схема, предложенная в 1962 г. американскими учеными Лейтом и Упатниексом. Ее упрощенный вариант приведен на *рис. 3.5.1*.

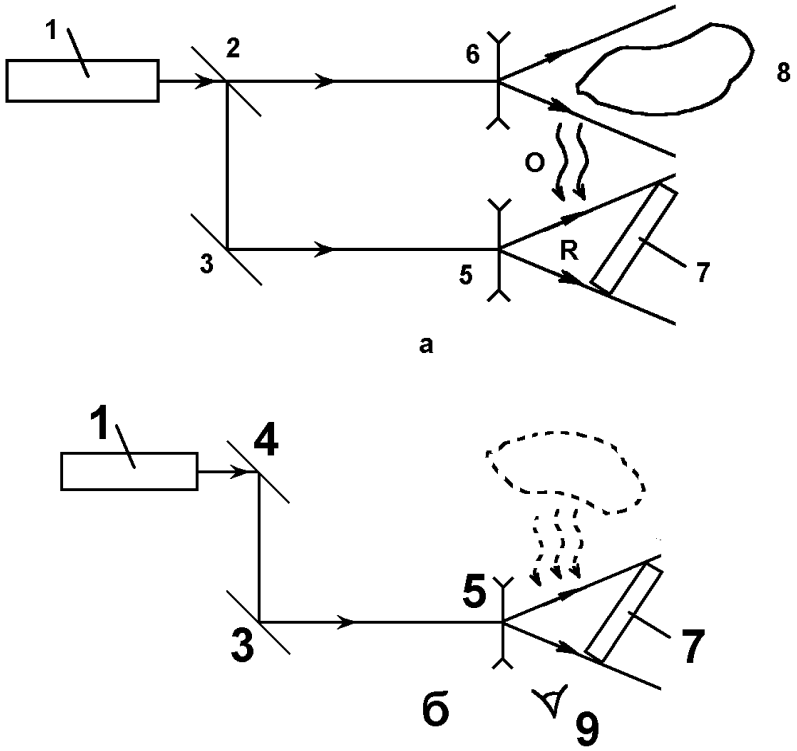


Рис. 3.5.1. *Схема записи (а) и восстановления (б) изображения: 1 - лазер; 2 - светоделительная пластина; 3,4 - поворотные зеркала; 5,6 - линзы; 7 - голографическая пластина; 8 - объект; 9 - наблюдатель; O - объектная волна; R - опорная волна.*

Как видно из рисунка, при восстановлении изображения используется пучок, который при записи выполняет роль опорного. Положение восстановленного изображения полностью соответствует положению объекта.

Другая схема, ориентированная на получение так называемых голограмм Фурье, будет далее рассмотрена более подробно.

3.5.2. Голограммы Фурье

Рассмотрим приведенную на *рис. 3.5.2* схему получения голограммы Фурье с точечным опорным источником, расположенным на оси системы. Такая схема рассчитана на получение голограмм плоских предметов, как правило - изображений различных объектов на фотопленке. Голограммы Фурье широко используются для оптической обработки информации, а также в системах голографической памяти.

На схеме плоскость, в которой размещается пленка с изображениями объектов, обозначена буквой *П*, а плоскость, в которой формируется голограмма, - буквой *Г*. Плоскости *П* и *Г* совпадают с фокальными плоскостями линзы $Л_3$. На плоскость *П* падает плоская волна когерентного света, создаваемая источником *S*. От этого же источника в плоскости *П* с помощью линз $Л_1$ и $Л_2$ формируется точечный источник S_0, создающий опорную волну. Опорный и предметный пучки собираются

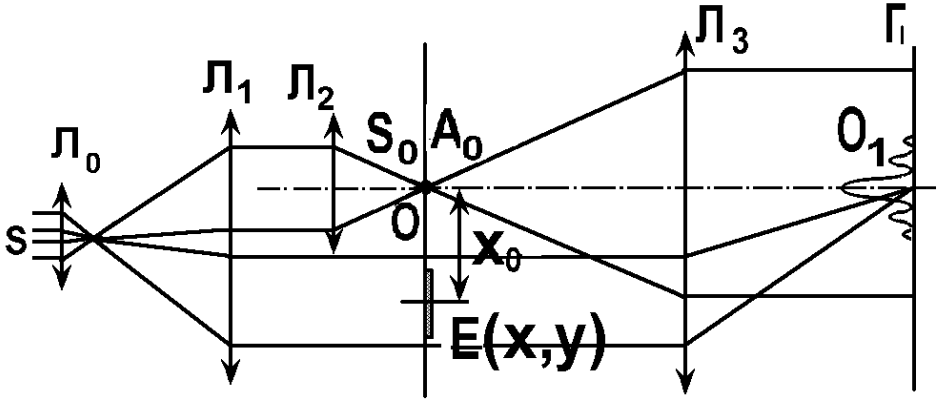


Рис. 3.5.2. *Схема получения голограммы Фурье.*

линзой $Л_3$ и в плоскости *Г* создают интерференционную картину, которая регистрируется помещенной в этой плоскости фотопленкой.

Выберем в плоскости *П* систему координат *Oxy*, а в плоскости *Г* - систему $O_1x_1y_1$. Расположим начала координат этих систем на оптической оси линзы $Л_3$. Точечный опорный источник S_0 поместим в начало

координат системы *Oxy*. Комплексное световое поле, образованное источником *S* в плоскости фотопленки, обозначим двумерной функцией *h(x,y)*. Комплексное световое поле в плоскости *Г* обозначим через $g(x_1,y_1)$. Линза $Л_3$ выполняет над функцией *h(x,y)* двумерное преобразование Фурье, так что функции *h(x,y)* и $g(x_1,y_1)$ связаны следующей зависимостью:

$$g(x_1,y_1)=\frac{1}{\lambda f_1}\int_{-\infty}^{\infty}\int h(x,y)\exp\left[-i\frac{2\pi}{\lambda f_1}(xx_1+yy_1)\right]dxdy, \quad \textbf{(3.5.5)}$$

где λ - длина волны когерентного света источника *S*, f_1 - фокусное расстояние линзы $Л_3$. Запишем (3.5.5) в несколько иной форме:

$$g(x_1,y_1)=\int_{-\infty}^{\infty}\int h(x,y)\,\mathbf{exp}\left[-i2\pi\left(\frac{x_1}{\lambda f_1}x+\frac{y_1}{\lambda f_1}y\right)\right]dxdy, \quad \textbf{(3.5.6)}$$

из которой видно, что в подынтегральном выражении роль пространственных частот играют величины

$$p=\frac{x_1}{\lambda f_1},\quad q=\frac{y_1}{\lambda f_1}. \quad \textbf{(3.5.7)}$$

Иначе:

$$g(x_1,y_1)=\frac{1}{\lambda f_1}\int_{-\infty}^{\infty}\int h(x,y)\,\mathbf{exp}\left[-i2\pi(px+qy)\right]dxdy. \quad \textbf{(3.5.8)}$$

Таким образом,

$$g(x_1,y_1)=\frac{1}{\lambda f_1}H(p,q), \quad \textbf{(3.5.9)}$$

где *H(p,q)* - двумерное преобразование Фурье функции *h(x,y)*, а *p* и *q* определяются равенствами (3.5.7). Будем называть *H(p,q)* комплексным спектром функции *h(x,y)*, а его составляющие *A(p,q)* и $\varphi(p,q)$ в соответствии с выражением

$$H(p,q)=A(p,q)\cdot\mathbf{exp}\left[i\varphi(p,q)\right]$$

- амплитудным и фазовым спектрами.

Вернемся к схеме *рис. 3.5.2* и составим выражение для функции *h(x,y)*. Как видно из рисунка, эта функция образуется двумя компонентами: полем опорного источника и полем, прошедшим пленку с изображением объекта. Так как мы рассматриваем опорный источник в виде идеальной точки, обозначим его поле с помощью двумерной дельта- функции Дирака с амплитудой волны A_0, т.е.

$$E_0(x,y) = A_0\delta(x,y). \qquad \textbf{(3.5.10)}$$

Поле центрированного объекта обозначим через $E^0(x,y)$. Ввиду того, что в принятой схеме объект смещен относительно начала координат на величину x_0 и y_0, то поле будет характеризоваться функцией $E_0(x-x_0, y-y_0)=E(x,y)$. Таким образом, получим:

$$h(x,y) = A_0\delta(x,y) + E^0(x-x_0, y-y_0). \qquad \textbf{(3.5.11)}$$

Поле в плоскости Γ определим в соответствии с выражением (3.5.8). Как известно, преобразование Фурье суммы функций равно сумме преобразований Фурье слагаемых, поэтому получим комплексные спектры слагаемых в (3.5.11) раздельно. Для первого слагаемого, вспоминая, что по одному из определений дельта- функции

$$\int_{-\infty}^{\infty} f(x)\delta(x)dx = f(0),$$

получим:

$$H_0(p,q) = \int_{-\infty}^{\infty}\int A_0\delta(x,y)\,\mathbf{exp}\left[-j2\pi(px+qy)\right]dxdy = A_0. \qquad \textbf{(3.5.12)}$$

Для второго слагаемого

$$H_E(p,q) = \int_{-\infty}^{\infty}\int E^0(x-x_0, y-y_0)\,\mathbf{exp}\left[-i2\pi(px+qy)\right]dxdy.$$

Делая подстановку

$$x - x_0 = u, \quad y - y_0 = v,$$

получим

$$H_E(p,q) = \int_{-\infty}^{\infty}\int E^0(u,v)\,\mathbf{exp}\left[-i2\pi(pu+qv)\right]dudv\,\mathbf{exp}\left[-i2\pi(px_0+qy_o)\right]$$

или

$$H_E(p,q) = \mathbf{exp}\left[-i2\pi(px_0+qy_0)\right]H_E^0(p,q), \qquad \textbf{(3.5.13)}$$

где $H_E^0(p,q)$ - комплексный спектр центрированного объекта.

Таким образом, учитывая, что

$$g(x_1,y_1) = \frac{1}{\lambda_1 f_1}\left[H_0(p,q) + H_E(p,q)\right],$$

из (3.5.12) и (3.5.13) получим:

$$g(x_1,y_1) = \frac{1}{\lambda f_1}\left\{A_0 + \mathbf{exp}\left[-i2\pi(px_0+qy_0)\right]H_E^0(p,q)\right\}. \qquad \textbf{(3.5.14)}$$

Комплексный спектр $H_E^0(p,q)$ центрированного объекта через амплитудный и фазовый спектры выразится следующим образом:

$$H_E^0(p,q) = A_E(p,q)\,\mathbf{exp}\left[i\varphi_E(p,q)\right].$$

Поэтому для поля (3.5.14) будем иметь:

$$g(x_1,y_1) = \frac{1}{\lambda f_1}\left\{A_0 + A_E(p,q)\,\mathbf{exp}\left[-i\left(2\pi(px_0+qy_0) - \varphi_E(p,q)\right)\right]\right\}. \quad \textbf{(3.5.15)}$$

Голограмма получается регистрацией поля в плоскости Γ на фотопленку. Ввиду того, что фотографическая эмульсия реагирует на интенсивность света $I(p,q)$, прозрачность пленки $\tau(p,q)$ (амплитудное пропускание) является функцией интенсивности:

$$\tau(p,q) = k\left[I(p,q)\right]^{-\frac{\gamma}{2}}, \quad \textbf{(3.5.16)}$$

где γ - коэффициент контрастности пленки. Коэффициент k определяется чувствительностью фотослоя и временем экспозиции.

Интенсивность светового поля в плоскости Γ равна квадрату модуля функции $g(x_1,y_1)$, т.е.

$$I(p,q) = g(x_1,y_1)\cdot g^*(x_1,y_1), \quad \textbf{(3.5.17)}$$

где

$$x_1 = p\lambda f_1, \quad y_1 = q\lambda f_1,$$

а звездочкой обозначена комплексно сопряженная величина. Подставив в (3.5.17) вместо $g(x_1,y_1)$ ее значение из (3.5.15), получим:

$$J(p,q) = \frac{1}{\lambda^2 f_1^2}\left[A_0 + A_E(p,q)e^{-i\tilde{\varphi}(p,q)}\right]\left[A_0 + A_E(p,q)e^{i\tilde{\varphi}(p,q)}\right] =$$

$$= \frac{1}{\lambda^2 f_1^2}\left[A_0^2 + A_E^2(p,q) + A_0A_Ee^{i\tilde{\varphi}(p,q)} + A_0A_Ee^{-i\tilde{\varphi}(p,q)}\right]. \quad \textbf{(3.5.18)}$$

Здесь

$$\tilde{\varphi}(p,q) = 2p(px_0 + qy_0) - \varphi_E(p,q) \quad \textbf{(3.5.19)}$$

Учитывая, что $\dfrac{e^{i\tilde{\varphi}} + e^{-i\varphi}}{2} = \cos\tilde{\varphi}$, получим:

$$I(p,q) = \left(\frac{A_0}{\lambda f_1}\right)^2\left[1 + \frac{A_E^2(p,q)}{A_0^2} + 2\frac{A_E(p,q)}{A_0}\,\mathbf{cos}\,\tilde{\varphi}(p,q)\right]. \quad \textbf{(3.5.20)}$$

Функцию прозрачности голограммы получим, подставив (3.5.20) в (3.5.16). При A_0, достаточно больших по сравнению с $A_E(p,q)$, разлагая (3.5.16) в степенной ряд и ограничиваясь линейным приближением, получим:

$$\tau(p,q) = k\left(\frac{A_0}{\lambda f_1}\right)^{-\gamma}\left[1-\frac{\gamma}{2}\cdot\frac{A_E^2(p,q)}{A_0^2}-\gamma\frac{A_E(p,q)}{A_0}\cos\tilde{\varphi}(p,q)\right]. \quad \textbf{(3.5.21)}$$

Запишем полученное выражение в более удобной для анализа форме:

$$\tau(p,q) = k_\Gamma\left[\alpha-\beta A_E^2(p,q)-A_E(p,q)\cos\tilde{\varphi}(p,q)\right]. \quad \textbf{(3.5.22)}$$

Здесь

$$k_\Gamma = \frac{k\gamma A_0^{-\gamma-1}}{\lambda^{-\gamma}f_1^{-\gamma}}, \quad \alpha = \frac{A_0}{\gamma}, \quad \beta = \frac{1}{2A_0}.$$

Соотношение (3.5.22) является уравнением голограммы Фурье, полученной по схеме *рис. 3.5.2*. но представляет собой связь амплитудного пропускания $\tau(p,q)$ голограммы с пространственно-частотными характеристиками голографируемого плоского предмета.

С точностью до разрешающей способности фотопленки и в пределах линейности ее характеристики на голограмме зафиксирована вся информация о предмете, содержащаяся в его амплитудном $A_E(p,q)$ и фазовом $\varphi_E(p,q)$ спектрах. Для сведения к минимуму нелинейных искажений при формировании голограммы необходимо правильно выбирать входящие в уравнение параметры $k\Gamma$, α и β, а следовательно, и определяющие их амплитуду A_0 опорного пучка и коэффициент контрастности γ. Обычно принимают $A_0>10A_E$ и $\gamma\approx 2$.

Проанализируем уравнение голограммы (3.5.22).

В правой части уравнения содержится три слагаемых. Постоянная составляющая $k_\Gamma\alpha=\tau_0$ определяет среднюю прозрачность голограммы, которая получилась бы в случае перекрывания пучка от предмета, т.е. когда $A_E(p,q)=0$.

Вторая составляющая $k_\Gamma\beta A_E^2(p,q)$ характеризует дополнительную неравномерную засветку голограммы пучком от предмета. Так как эта составляющая вычитается из τ_0, то она уменьшает прозрачность голограммы, особенно в тех местах, где велико значение амплитудного спектра предмета. Ввиду того, что для большинства предметов наибольшую

энергию несут низкочастотные составляющие спектра, потемнение голограммы Фурье за счет второй составляющей сосредоточено вблизи начала координат частотной плоскости.

Вторая составляющая содержит лишь часть информации о предмете, так как в ней отсутствует фазовый спектр. Полную информацию содержит третья составляющая, возникающая благодаря интерференции предметного пучка с опорным. Из-за наличия в ней функции $\cos\tilde{\varphi}$ она знакопеременная. При положительном значении косинуса она уменьшает прозрачность голограммы, при отрицательном- увеличивает. Эта составляющая представляет собой косинусную волну, промодулированную по амплитуде и фазе. Вектор несущей частоты косинусоиды имеет составляющие

$$\omega_p = 2\pi x_0, \omega_q = 2\pi y_0, \qquad \textbf{(3.5.23)}$$

зависящие от смещения предмета относительно опорного источника. Направление фронта волны косинусоиды получим из соотношения (3.5.19) для фазы. Полагая $\varphi_E(p,q)=0$, из (3.5.19) получим:

$$\tilde{\varphi}_0(p,q) = 2\pi x_0\left(p + \frac{y_0}{x_0}q\right). \qquad \textbf{(3.5.24)}$$

Так как фронт волны соединяет точки с одинаковой фазой, то, положив

$$p + \frac{y_0}{x_0}q = z = const, \qquad \textbf{(3.5.25)}$$

получим уравнение фронта косинусоиды:

$$p = -\frac{y_0}{x_0}q + z.$$

Это прямая, имеющая наклон $k=-y_0/x_0$. Косинусная волна на голограмме проявляется в виде периодических интерференционных полос и особенно хорошо видна при равномерных $A_E(p,q)$ и $\varphi_E(p,q)$, т.е. когда она не сильно искажена модуляцией.

Таким образом, объект на голограмме представляется в виде суперпозиции элементарных косинусоид. Этим и объясняется, почему этот способ записи получил название голографии Фурье.

Следует отметить, что фазовый спектр $\varphi_E(p,q)$, входящий в выражение для $\tilde{\varphi}$, есть фазовый спектр центрированного объекта, ибо смещение объекта на величины x_0 и y_0 от центра учитывается первым слагаемым в (3.5.19). Поэтому для всех объектов, имеющих центральную симметрию (двумерный аналог четности функций), $\varphi_E(p,q)=0$.

Отметим также, что голограмма Фурье любого вещественного объекта имеет центральную симметрию. Это следует из того, что уравнение голограммы таких объектов инвариантно по отношению к перемене знака пространственных частот, ибо входящие в него члены $A_E(p,q)$ и $\cos\tilde{\varphi}(p,q)$ не изменяют знака при изменении знаков p и q, первый вследствие центральной симметрии, а второй - вследствие четности.

При восстановлении голограммы она освещается параллельным пучком света (*рис. 3.5.3*) Каждая косинусоидальная решетка при этом

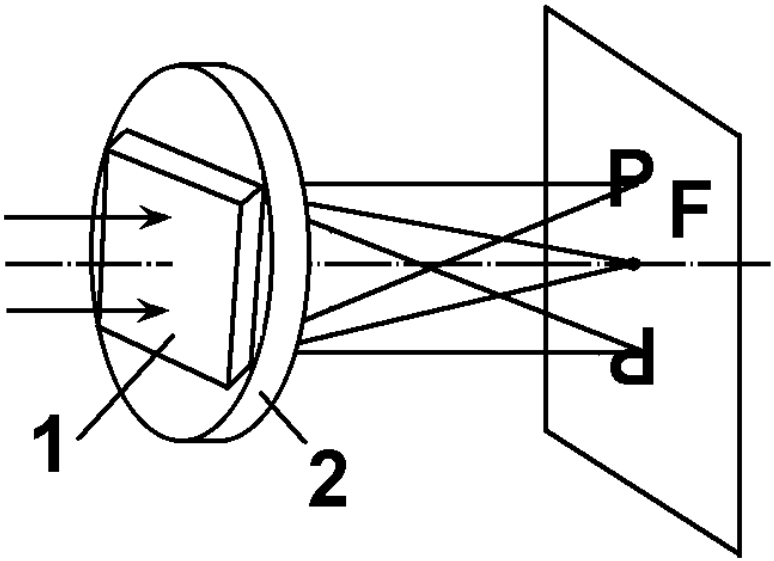


Рис.3.5.3. *Восстановление голограммы Фурье: 1-голографическая фотопластинка; 2-линза*

сформирует изображение соответствующей точки в бесконечности. Если же сразу за голограммой установить линзу, то изображения всех точек пернесутся из бесконечности в фокальную плоскость линзы. Изображение объекта и сопряженное ему изображение будут располагаться симметрично относительно оси. На оси будет располагаться светлое пятно, обусловленное наличием первых двух слагаемых в уравнении голограммы.

В заключение заметим, что воспроизведение большого диапазона значений прозрачности голограммы Фурье, полученной по схеме *рис. 3.5.2*, является сложной задачей, особенно при синтезе цифровых голограмм.

Одним из путей устранения этой трудности является переход к схеме с рассеивателем. Эта схема отличается от предыдущей лишь тем, что перед предметом, установленным в плоскости *П*, помещается рассеиватель - тонкая прозрачная пластинка с неровной поверхностью (матовое стекло), не изменяющая амплитуду, но изменяющая случайным образом фазу в каждой точке падающей на предмет волны. Введенная таким образом случайная фаза не влияет на изображение предмета, получаемое при восстановлении, так как фотослой, регистрирующий изображение, чувствителен лишь к интенсивности восстановленной волны. В то же время она существенным образом перераспределяет энергию в частотной плоскости, т.е. в плоскости голограммы, значительно уменьшая требуемую фотографическую ширину фотослоя.

Изложенные выше свойства голограмм играют важную роль при разработке методов оптической фильтрации, краткая характеристика которых приведена в следующем параграфе.

3.6. Оптическая фильтрация и распознавание образов [2, 3]

3.6.1 Применение системы 4-F

Рассмотрим устройство, состоящее из двух линз (*рис. 3.6.1*). Расположим их так, чтобы предметная плоскость *(x,y)* совпадала с передней фокальной плоскостью линзы 1, а задняя фокальная плоскость линзы 1 совпадала с передней фокальной плоскостью линзы 2. Плоскость изображений *(x',y')* совместим с задней фокальной плоскостью линзы 2.

Рис. 3.6.1 иллюстрирует случай, когда линзы имеют одинаковые фокусные расстояния *F*; выделенная на нем плоскость (ξ_1, η_1) является плоскостью линзы 1, плоскость (ξ_2, η_2) является спектральной плоскостью, а плоскость (ξ_3, η_3) - плоскостью линзы 2. Рассматриваемая оптическая система известна в литературе как система 4-F. Она осуществляет два последовательных преобразования Фурье.

Сначала входной сигнал *(x,y)* подвергается Фурье-преобразованию линзой 1. Световое поле в плоскости (ξ_2, η_2) распределено в соответствии со спектром сигнала $S(f_x, f_y)$

$$\psi_{\xi}(\xi_2,\eta_2)=\frac{1}{\lambda F}\iint_{\infty}\psi(x,y)\,\mathbf{exp}\left[-i2\pi\left(f_x x+f_y y\right)\right]dxdy= \tag{3.6.1}$$
$$=\frac{1}{\lambda F}\mathrm{S}(\mathrm{f_x},\mathrm{f_y}).$$

Здесь мы воспользуемся тем, что $\xi_2=\lambda F f_x$, $\eta_2=\lambda F f_y$.

Сигнал $\psi\xi(\xi_2,\eta_2)$, пропорциональный спектру $S(f_x,f_y)$, подвергается аналогичному преобразованию Фурье при прохождении через линзу 2. Тем самым в плоскости изображений распределение поля будет иметь вид

$$\psi'(x',y')=\frac{1}{\lambda^2F^2}\iint_{\infty}S\left(f_x,f_y\right)\mathbf{exp}\left[-i2\pi\left(f_x x'+f_y y'\right)\right]d\xi_2 d\eta_2= \tag{3.6.2}$$
$$\iint_{\infty}S\left(f_x,f_y\right)\mathbf{exp}\left[-i2\pi\left(f_x x'+f_y y'\right)\right]df_x df_y=\psi(-x',-y').$$

Таким образом, система 4-F восстанавливает в плоскости изображений входной сигнал в перевернутом виде.

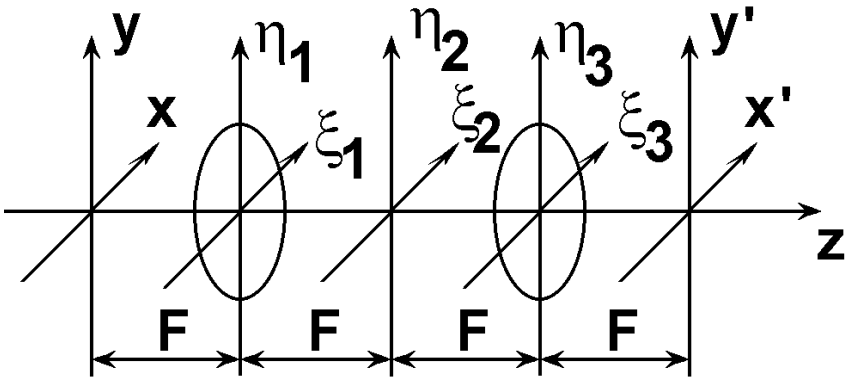


Рис. 3.6.1. *Система 4-F*

Систему 4-F можно с успехом использовать для пространственной фильтрации. Для этого в плоскости (ξ_2,η_2) помещают пространственный фильтр. Если фильтр имеет модуляционную характеристику $T(\xi_2,\eta_2)$, то после прохождения фильтра спектр $S(f_x,f_y)$ входного сигнала принимает вид

$$S_1(f_x,f_y)=S(f_x,f_y)T(\lambda T(_x,\lambda\lambda F_y). \tag{3.6.3}$$

При этом в плоскости изображений будет формироваться сигнал $\psi_1'(x', y')$, соответствующий спектру $S_1(f_x, f_y)$. Тем самым, меняя модуляционную характеристику фильтра, можно обеспечить требуемое преобразование оптического пространственного сигнала.

Пространственная фильтрация широко используется для решения задач, связанных с проблемой распознавания образов. Предположим, что на входе системы 4-F помимо полезного сигнала присутствуют посторонние сигналы (помехи) $\psi_i(x, y)$ и

$$\psi(x, y) = \psi_a(x, y) + \sum_{i=1}^{n} \psi_i(x, y). \qquad \textbf{(3.6.4)}$$

Тогда в спектральной плоскости будет формироваться сумма спектров:

$$S(f_x, f_y) = S_a(f_x, f_y) + \sum_{i=1}^{n} S_i(f_x, f_y). \qquad \textbf{(3.6.5)}$$

Если в спектральной плоскости расположить фильтр с модуляционной характеристикой

$$T(f_x, f_y) = S_a(f_x, f_y), \qquad \textbf{(3.6.6)}$$

то сразу за ним спектр сигнала $\psi(x, y)$ принимает вид

$$S_1(f_x, f_y) = S(f_x, f_y)T(f_x, f_y) = \left|S_a(f_x, f_y)\right|^2 + \sum_{i=1}^{n} S_i(f_x, f_y) S_a^*(f_x, f_y). \qquad \textbf{(3.6.7)}$$

Волновой фронт поля $S_a(f_x, f_y)$ в плоскости (ξ_2, η_2) является плоским, поэтому, проходя через линзу 2, это световое поле должно фокусироваться в точку на плоскости изображения. Второе слагаемое в (3.6.7) представляет поле не с плоским фронтом, его преобразование линзой 2 дает на выходной плоскости некоторое распределенное изображение.

Пространственный фильтр, превращающий парциальную световую волну, несущую информацию о полезном сигнале, в плоскую волну, называют согласованным с полезным (распознаваемым) сигналом.

Если во входной волне содержится полезный сигнал, то присутствие в системе согласованного фильтра приводит к появлению в выходной плоскости яркой светящейся точки. Фиксируя появление этой точки, можно регистрировать наличие на входе системы распознаваемого сигнала.

3.6.2. Голографический метод синтезирования пространственных фильтров и проблема апостериорной обработки информации

Эффективность пространственной фильтрации во многом определяется качеством используемых фильтров с заданной переходной функцией. Для синтезирования пространственных фильтров успешно используются голографические методы. На *рис. 3.6.2* изображена голографическая схема получения пространственных фильтров, известная как схема Ван дер Люгта. Она включает следующие элементы: пластину 2, на которой записан оптический сигнал $h(\xi,\eta)$, описывающий переходную функцию системы; собирающую линзу 3 с фокусным расстоянием F, производящую Фурье-преобразование сигнала $h(\xi,\eta)$; голограмму 4 с записью пространственного фильтра; призму 5, формирующую из части

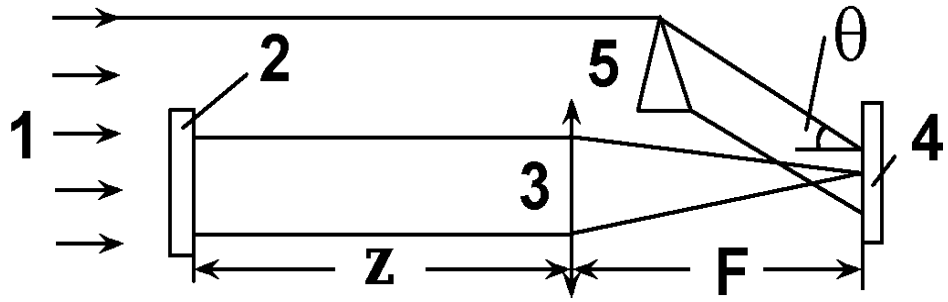


Рис.3.6.2. *Схема Ван дер Люгта для синтеза пространственных фильтров*

падающего на систему светового потока 1 плоскую волну, падающую на голограмму под углом Θ. При записи голограммы объектной волной является волна

$$O = \frac{1}{\lambda F} H\left(\frac{x}{\lambda F}, \frac{y}{\lambda F}\right),$$

где

$$H(f_x, f_y) = \iint_{\infty} \mathrm{h}(\xi,\eta)\exp\left[-\mathrm{i}2\pi\left(f_x\xi + f_y\eta\right)\right]d\xi\mathrm{d}\eta,$$

а опорной волной - волна

$$R = \mathbf{exp}\left(-i2\pi\alpha y\right),$$

где

$$\alpha = \frac{\sin\Theta}{\lambda}.$$

Рассчитаем модуляционную характеристику голограммы

$$T(x,y) \sim \left| \frac{1}{\lambda F} H\left(\frac{x}{\lambda F}, \frac{y}{\lambda F}\right) + \exp(-i2\pi\alpha y) \right|^2 =$$

$$= 1 + \frac{1}{(\lambda F)^2}|H|^2 + \frac{1}{\lambda F} H + \exp(i2\pi\alpha y) + \qquad \textbf{(3.6.8)}$$

$$+ \frac{1}{\lambda F} H^* \exp(-i2\pi\alpha y).$$

В случае использования этой голограммы в качестве пространственного фильтра для некоторого входного сигнала существенным является наличие в характеристике (3.6.8) третьего и четвертого членов. Они действуют как пространственные фильтры H и H^*. Их действие приводит к тому, что после второго Фурье-преобразования в выходной плоскости системы (*рис. 3.6.1*) будут присутствовать сигналы

$$\iint_{\infty} HS\left[-i2\pi\left(\frac{x}{\lambda F}x' + \frac{y}{\lambda F}(y' - \alpha\lambda F)\right)\right]\frac{dxdy}{\lambda^2 F^2} = \qquad \textbf{(3.6.9)}$$

$$= \psi_1'(-x', -y' + \alpha\lambda F),$$

$$\iint_{\infty} H^* S\left[-i2\pi\left(\frac{x}{\lambda F}x' + \frac{y}{\lambda F}(y' - \alpha\lambda F)\right)\right]\frac{dxdy}{\lambda^2 F^2} = \qquad \textbf{(3.6.10)}$$

$$= \psi_2'(-x', -y' - \alpha\lambda F).$$

Из выражений (3.6.9) и (3.6.10) видно, что выходной сигнал, полученный при использовании фильтра H, будет сдвинут вверх по вертикали на расстояние $\alpha\lambda F$, сигнал же, полученный при использовании фильтра H^*, будет на то же расстояние сдвинут вниз. Таким образом появляется возможность рассматривать эти сигналы порознь.

Рассмотренный голографический метод получения пространственных фильтров снимает проблему синтеза оптических масок в пространстве частот. Трудности же синтеза оптических масок в пространстве координат менее серьезны, поскольку требуемые переходные функции (импульсные отклики), как правило, имеют простую форму и необходимые маски с пропусканием $h(\xi, \eta)$ несложно изготовить с помощью простых фотографических средств.

Синтезирование пространственных фильтров голографическими методами позволяет успешно решать важную в практическом отношении

задачу улучшения качества оптических изображений, выполняемого уже после того, как получено изображение. Такая задача относится к задачам, связанным с апостериорной ("послеопытной") обработкой информации.

Пусть некоторая оптическая система формирует изображение интересующего нас объекта. Идущая от объекта световая волна может рассматриваться как входной сигнал $\psi(x,y)$, а изображение объекта - как выходной сигнал $\psi'(x',y')$. Если бы система не вносила искажений, то имело бы место равенство $\psi=\psi'$. Однако в реальной ситуации следует считаться с тем, что любая система вносит определенные искажения, которые описываются некоторой переходной функцией (или передаточной функцией H). С помощью голографических методов можно синтезировать пространственный фильтр с модуляционной характеристикой H^*. Совмещая его с фильтром, синтезированным обычным фотоспособом, получим составной фильтр с модуляционной характеристикой

$$T = \frac{H^*}{|H|^2}. \qquad \textbf{(3.6.11)}$$

Для апостериорной обработки искаженного сигнала ψ' воспользуемся схемой 4-F (*рис. 3.6.1*) В ее входной плоскости разместим искажение ψ', а в спектральной плоскости разместим фильтр с модуляционной характеристикой (3.6.11). Тогда в выходной плоскости системы 4-F сформируется сигнал $\overline{\psi}' = F^{-1}\{S' T\}$. Так как $S'=SH$, то согласно (3.6.11)

$$S' T = \frac{SHH^*}{|H|^2} = S. \qquad \textbf{(3.6.12)}$$

Это означает, что $\overline{\psi}' = \psi$ Таким образом, благодаря модуляционной характеристике (3.6.11) фильтр "учел" и автоматически "вычел" те искажения, которые оптическая система внесла при построении изображения.

3.7. Сопоставление методов когерентной и некогерентной оптики [2]

Когерентная оптическая фильтрация зарекомендовала себя удобным средством обработки изображений, что объясняется ее высоким быстродействием, двумерностью, а также относительной универсальностью, поскольку могут быть реализованы почти любые линейные пространственно-инвариантные операции фильтрации над комплексными сигналами. Тем не менее, когерентные методы имеют и

слабые стороны. К ним, в первую очередь, относится требование когерентности источника света, что не позволяет обрабатывать самосветящиеся объекты, такие, например, как изображения на телевизионном экране. Другим недостатком когерентных систем является их большая чувствительность к различного рода шумам (например, связанным с зернистостью фотоэмульсий, дефектами оптических поверхностей и т.д.). Эти шумы негативно сказываются на качестве выходных изображений.

Указанные недостатки методов обработки информации в когерентном свете стимулируют разработку альтернативных способов, основанных на принципах некогерентной оптики. Однако, поскольку подробное изучение этих способов выходит за рамки книги, ограничимся самыми общими замечаниями.

В отличие от когерентного оптического сигнала ψ(x,y) некогерентный сигнал $\tilde{\psi}(x,y)$ описывается не комплексной, а действительной функцией. Информация содержится в интенсивности световой волны

$$\tilde{\psi}(x,y) = \left|\psi(x,y)\right|^2. \qquad \textbf{(3.7.1)}$$

Фазовая информация в световой волне утрачивается.

Интеграл свертки примет вид

$$\tilde{\psi}'(x',y') = \iint_{\infty} \left|\psi(x,y)h(x'-x,y'-y)\right|^2 dxdy = \\ = \iint_{\infty} \left|\tilde{\psi}(x,y)\tilde{h}(x'-x,y'-y)\right|^2 dxdy. \qquad \textbf{(3.7.2)}$$

Здесь функция

$$\tilde{h}(x'-x,y'-y) = \left|h(x'-x,y'-y)\right|^2 \qquad \textbf{(3.7.3)}$$

является некогерентной функцией системы. Аналогично в случае некогерентных изопланарных систем вводят некогерентную передаточную функцию системы

$$\tilde{H}(f_x,f_y) = \iint_{\infty} \tilde{h}(x,y)\,\mathbf{exp}\left[-i2\pi(f_x x + f_y y)\right]dxdy. \qquad \textbf{(3.7.4)}$$

Используя (3.7.3) и известную в Фурье-анализе теорему автокорреляции, можно построить выражение

$$\tilde{H}(f_x,f_y) = \iint_{\infty} H(f_x',f_y')H^*(f_x'+f_x,f_y'+f_y)df_x'df_y', \qquad \textbf{(3.7.5)}$$

которое определяет связь между передаточными функциями для когерентного и некогерентного сигналов.

3.8. Характеристики качества изображения[10]

В заключение данной главы рассмотрим дополнительно вопрос о количественных характеристиках качества изображений формируемых оптическими системами.

Если бы волновой интервал светового излучения был бесконечно мал, а оптическая система (объектив) совершенна, то функция $\psi_0(x,y)$ преобразовалась бы идеальным образом в распределение освещенности $\psi'_0(x',y')$; единственным различием их было бы линейное увеличение. Реальное же распределение освещенности в изображении $\psi'_i(x',y')$ отличается от идеального ψ_0 из-за поперечного дифракционного рассеяния света, а так же несовершенств объектива. Качество изображения - мера степени отличия ψ_i от ψ_0.

Принято использовать три параметра для сравнения распределений в объекте (предмете) и изображении. Этими параметрами являются:

относительное содержание (емкость) структуры T_L

$$T_L = \frac{\iint \psi'^2_i(x',y')dx'dy'}{\iint \psi'^2_0(x',y')dx'dy'}; \qquad \textbf{(3.8.1)}$$

правдоподобие изображения Φ_L

$$\Phi_L = 1.0 - \frac{\iint \left[\psi'_0(x',y') - \psi'_i(x',y')\right]^2 dx'dy'}{\iint \psi'^2_0(x',y')dx'dy'} \qquad \textbf{(3.8.2)}$$

$$Q_L = \frac{\iint \psi'_0(x',y')\psi'_i(x',y')dx'dy'}{\iint \psi'^2_0(x',y')dx'dy'}. \qquad \textbf{(3.8.3)}$$

Пределы интегрирования здесь формально $-\infty, +\infty$, но обычно достаточно интегрировать в пределах формата данного изображения. Заметим, что

$$Q_L = \frac{T_L + \Phi_L}{2}. \qquad \textbf{(3.8.4)}$$

Относительное содержание структуры наиболее близко подходит к тому, что обычно подразумевается под качеством изображения; реально же величина T_L представляет искажения вариаций функции ψ_0 в процессе ее регистрации. Ограничения применимости величины T_L связаны с искажениями, вносимыми оптической системой или приемником изображения. Величина T_L также "не чувствительна" к искажениям

координатной сетки из-за дисторсии. Эти искажения хорошо описываются величиной Φ_L - поэтому она широко используется в приложениях, где подобие координатной сетки принципиально важно. Корреляционное качество отражает сочетание требований как к подобию координатной сетки, так и к структурному содержанию изображения.

Параметры качества представляют собой удобные оценочные характеристики для случаев, когда параметры поверхности предмета априорно хорошо известны; указанные параметры - одночисловые, они легко сравнимы (для разных систем) и непосредственно взаимосвязаны (для данной системы).

ЛИТЕРАТУРА К ГЛАВЕ 3

1. Гудмен Дж. Введение в фурье-оптику. - М.: Мир, 1970, 364 с.

2. Дьяков В.А., Тарасов Л.В. Оптическое когерентное излучение.- М.: Советское Радио, 1974; 169 с.

3. Парыгин В.Н., Балакший В.И. Оптическая обработка информации.- М.: Изд. МГУ, 1987, 141 с.

4. Матвеев А.Н. Оптика.- М.: Высшая школа, 1985, 351 с.

5. Марешаль А., Франсон М. Структура оптического изображения.- М.: Мир, 1964. 295 с.

6. Козанне А., Флере Ж., Мэтр Г., Руссо М. Оптика и связь.- М.: Мир, 1984, 504 с.

7. Франсон М. Голография.- М.: Мир, 1972, 246 с.

8. Сороко Л.М. Основы голографии и когерентной оптики.- М.: Наука, 1971, 616 с.

9. Федоров Б.Ф. Эльман Р.И. Цифровая голография. - М.: Наука, 1976, 152 с.

10. Уэзерелл У. Оценка качества изображения. //Сб. Проектирование оптических систем. Под ред. Р. Шеннона и Дж. Вайнта.-М.: Мир, 1983, 430 с.

ГЛАВА 4. МЕТОДЫ КОМПЬЮТЕРНОЙ ОПТИКИ

Термин "компьютерная оптика" является относительно новым и не приобрел еще строгого определения. Разные авторы очень часто вкладывают в него различное содержание. Можно сказать, что в самом широком смысле слова "компьютерная оптика" - это компьютеры в оптике и оптика в компьютерах. Сюда относятся численные решения задач дифракции и фокусировки излучения, автоматизированное проектирование и гибкое автоматизированное производство оптических систем, обработка изображений, оптический вычислительный эксперимент, оптические процессоры и запоминающие устройства, цифровая голография.

Очень часто формулировка предмета компьютерной оптики как научного направления сужается и в нее вкладывается более конкретный смысл. При этом считается, что компьютерная оптика - это получение на основе применения ЭВМ оптических элементов, осуществляющих требуемое преобразование волновых полей. Именно на такую направленность компьютерной оптики и сориентирован, в основном, материал настоящей главы.

Компьютерная оптика возникла на стыке физической оптики и информатики и стала формироваться как новое научное направление, объединяющее теорию, методы и технические средства обработки оптических сигналов с использованием ЭВМ, и отражающее тот факт, что современная оптика и оптические приборы становятся все в большей степени цифровыми.

Какие качественно новые свойства придают компьютеры оптическим системам? Главных свойств два. Во-первых, это способность к адаптации и гибкость в перенастройке. Благодаря тому, что компьютер способен перестраивать структуру обработки сигнала без перестройки своей физической структуры, он является идеальным средством адаптивной обработки оптических сигналов и быстрой перестройки ее на решение разных задач. Здесь речь идет прежде всего об информационной адаптации. Заметим попутно, что эта способность ЭВМ к адаптации и перестройке нашла так же применение в активной и адаптивной оптике для управления световыми пучками как переносчиками энергии.

Во-вторых, это простота и естественность получения и переработки количественной информации, содержащейся в оптических сигналах, соединение оптических систем с другими информационными системами. Цифровой сигнал, представляющий в компьютере оптический сигнал,- это

переносимая оптическим сигналом информация, так сказать, в чистом виде, лишенная своей физической оболочки.

Благодаря универсальному характеру цифровой сигнал представляет собой идеальное средство объединения различных информационных систем. Теоретической базой компьютерной оптики являются теории информации, цифровой обработки сигналов, статистических решений, теория систем и преобразований в оптике.

4.1. Задачи компьютерной оптики [1,2]

Одно из основных предназначений компьютерной оптики состоит в расширении гаммы конструктивных элементов оптических систем. Помимо традиционных линз, призм и зеркал с помощью современных вычислительных средств и систем управления могут быть созданы оптические элементы с более широкими функциональными возможностями. Типичным представителем семейства элементов компьютерной оптики является плоский оптический элемент- киноформ. Сочетание киноформных корректоров с обычными линзами позволяет проектировать оптические системы со слабыми сферическими аберрациями и новыми оптико-физическими свойствами.

Существо подхода к созданию элементов компьютерной оптики состоит в следующем. Оптический элемент, работающий на пропускание или на отражение излучения, характеризуется амплитудно-фазовой функцией пропускания или отражения. Эта характеристика должна быть определена, исходя из решаемой задачи преобразования волнового поля. Для простейших случаев может быть известно ее аналитическое выражение, например, фазовая функция сферической или цилиндрической линзы. В общем же случае требуется применение ЭВМ для определения характеристики оптического элемента. При этом ЭВМ может использоваться как для численных расчетов в рамках прямой задачи, так и для решения обратных задач. Таким образом, на этапе проектирования, компьютер используется для определения характеристики создаваемого оптического элемента.

После того, как указанная характеристика сформирована в памяти ЭВМ, возникает задача переноса ее на физическую среду с помощью программно-управляемого технологического автомата. На этом этапе роль компьютера также очень велика.

Созданный оптический элемент необходимо далее экспериментально исследовать и аттестовать. Экспериментальные данные при этом регистрируются, как правило, в виде различного рода распределений интенсивности света: теневых картин, интрферограмм, голограмм. При этом

компьютер необходим для обработки, отображения и интерпретации экспериментальных данных, поскольку визуальные наблюдения и ручная обработка не позволяют получать количественные результаты.

Необходимо отметить, что в компьютерной оптике перспективным методом исследования является вычислительный эксперимент, в котором ключевую роль играет компьютер. Процесс создания элементов компьютерной оптики носит сложный итерационный характер и на компьютер возлагается также функция обеспечения диалога с проектировщиком, технологом и исследователем.

В компьютерной оптике можно выделить следующие основные направления:

- цифровая голография;
- решение обратных задач теории дифракции и создание на их основе фокусаторов и корректоров излучения;
- создание оптических элементов для анализа и формирования поперечно-модового состава излучения;
- создание корректоров волновых фронтов и пространственных фильтров для оптических систем обработки информации и оптико-цифровых процессоров;
- цифровая обработка полей в оптических системах.

Характеристику вышеуказанных направлений мы начнем с цифровой голографии. Именно работы по цифровой голографии во многом стимулировали появление компьютерной оптики как самостоятельного научного направления на стыке квантовой электроники, вычислительной математики и информатики.

4.2. Цифровая голография [3-5]

Цифровой голографией называется метод получения и восстановления голограмм, при котором основная роль отводится компьютеру. Роль компьютера заключается в расчете распределения коэффициента прозрачности или преломления по полю голограммы, которое затем записывается в оптической запоминающей среде. С помощью компьютера рассчитывается и восстанавливается изображение, которое записано на такой синтезированной голограмме и которое можно было бы получить оптическим путем.

Имеется ряд веских оснований для такого синтеза голограмм и, в частности, то обстоятельство, что геометрические размеры голографического объекта в этом случае не ограничиваются такими факторами, как когерентность освещения, вибрация или турбулентность

воздуха, и появляется возможность исследовать путем моделирования некоторые голографические эффекты.

Еще более существенным моментом, стимулирующим синтезирование голограмм с помощью компьютеров, является возможность создать оптический волновой фронт для такого объекта, который физически не существует. Потребность в формировании волнового фронта, соответствующего объекту, определяемому расчетным путем, возникает в любом случае, когда требуется визуально отобразить в трех измерениях результаты того или иного трехмерного исследования, например, при моделировании разрабатываемых конструкций. Иногда волновой фронт от синтезированной голограммы может служить интерференционным эталоном для контроля сложной оптической поверхности в процессе ее обработки. Другая область применения таких голограмм связана с экспериментами по пространственной фильтрации. В некоторых случаях изготовить фильтр с заданной функцией оптическими методами бывает затруднительно, в то же время компьютер решает подобные задачи сравнительно легко.

4.2.1. Общая процедура изготовления синтезированной голограммы

Для того, чтобы получить синтезированную голограмму, поступают следующим образом:

1) Задавшись объектом, голограмму которого нужно получить, рассчитывают с помощью компьютера комплексную амплитуду испускаемого им света в плоскости, находящейся на определенном расстоянии от него. Эта плоскость будет плоскостью голограммы.
2) Рассчитанная таким образом комплексная амплитуда кодируется так, чтобы она была действительной и положительной функцией. Например, производят сложение амплитуды света, испускаемого объектом, с какой-нибудь комплексной амплитудой, которая играет роль когерентного фона. Результирующая интенсивность будет в этом случае действительной и положительной функцией.
3) Соответствующее устройство, управляемое компьютером, изображает графически распределение значений этой функции в некоторой плоскости. Это может быть, например, электронно-лучевая трубка, печатающее устройство и т.п.
4) Полученный чертеж фотографируется; негатив и представляет собой синтетическую голограмму. Для того, чтобы голограмма хорошо дифрагировала свет, нужно, чтобы структура чертежа была достаточно

тонкой. Поэтому обычно фотографируют чертеж со значительным уменьшением.

Для формирования голограммы применяются компьютерные дисплеи, штриховые печатающие устройства, плоттеры. Этап фотографического уменьшения, разумеется, может быть исключен, если применить специальные выходные устройства, позволяющие осуществить непосредственную запись голограммы требуемого размера. Быстродействие современных компьютеров достаточно для расчета синтетической голограммы, идентичной голограмме, полученной при записи интерференционной картины, созданной реальным объектом. Тем не менее, в большинстве случаев рассчитываются голограммы, где отсутствуют полутона и вся голограмма состоит из светлых участков (апертур) на черном фоне. Такая голограмма называется бинарной. Бинарную голограмму с помощью компьютера можно рассчитать и построить в увеличенном масштабе за несколько минут.

Фотографическое уменьшение и репродуцирование бинарных голограмм легче и более точно, чем серых голограмм. На качество бинарной голограммы совершенно не влияют нелинейные фотографические эффекты, поэтому в процессе фотоуменьшения бинарных голограмм требуется значительно менее строгий контроль величины экспозиции и режима проявления.

Другое преимущество бинарной голограммы в сравнении с серой голограммой состоит в том, что она направляет на восстанавливаемое изображение большую часть из падающего на нее света. Если в обычной голограмме светоотдача, или эффективность, равна 6,2%, то светоотдача бинарной голограммы достигает 10%. Помимо более высокой светоотдачи преимущество бинарной голограммы состоит в том, что при восстановлении возникает меньше шумов от света, рассеянного зернистой структурой фотоэмульсии. Бинарная голограмма может быть вычерчена плоттером. Восстановленное с бинарной голограммы в когерентном свете изображение имеет все свойства изображения, получаемого с обычной голограммы.

Бинарные голограммы являются эффективным промежуточным звеном, позволяющим осуществлять связь между цифровой и оптической формами представления информации. Один из методов цифровой голографии позволяет получать голограммы, которые при восстановлении падающий на голограмму свет направляют на создание одного изображения, т.е. имеют эффективность около 100%.

4.2.2. Получение цифровой голограммы Фурье и ее бинаризация

Рассмотрим более подробно процедуру получения цифровой голограммы. Сделаем это на примере голограммы Фурье, принцип регистрации которой был рассмотрен в параграфе 3.5.2. Как и всякие другие цифровые модели, цифровые модели голограмм воспроизводят процесс лишь приближенно, однако наиболее существенные свойства, подлежащие исследованию, представляются четко выделенными, в явном виде, что часто нельзя сделать в реальном процессе. Одно из основных приближений связано с переходом от непрерывных величин к дискретным, с которыми работает ЭВМ. Этот переход, уменьшая точность результатов, в то же время не вносит принципиальных изменений в процесс, так как с уменьшением шага дискретизации модель все более приближается к непрерывной. Степень такого приближения ограничена лишь возможностями ЭВМ. Кроме того, есть разумный предел плотности дискретизации, определяемый разрешающей способностью оптических элементов и фотоматериалов, участвующих в голографическом процессе. Этот предел для функций с ограниченным спектром определяется известной специалистам теоремой Котельникова, из которой следует, что если функция имеет спектр, ограниченный частотой f_0, то она может быть представлена с большой точностью в точках x_m, отстоящих одна от другой на расстоянии $\Delta x = 1/2f_0$. Теорема Котельникова легко распространяется на двумерные функции. В этом случае отсчеты берут в узлах прямоугольной сетки с размерами ячеек

$$\Delta x = 1/2p_0, \quad \Delta y = 1/2q_0.$$

Итак, переходя к цифровому моделированию голографического процесса, заменим части плоскостей П и Г (см**. рис. 3.5.2**), ограниченные прямоугольными апертурами, сетками. В узлах этих сеток зададим отсчеты поля. Эти сетки в плоскости предметов обозначим $\sigma_П$, а в плоскости голограммы - $\sigma_Г$. Для удобства последующих преобразований расположение сеток в плоскостях П и Г выберем таким, как показано на **рис. 4.2.1.** Правомерность такого выбора будет видна из дальнейшего. Чтобы параметры сеток отвечали теореме Котельникова, необходимо выполнение следующих соотношений:

$$\Delta x = 1/2P;\ \Delta y = 1/2Q;\ M = X/\Delta x = 2PX;\ N = Y/\Delta y = 2QY. \qquad \textbf{(4.2.1)}$$

При этом суммарное число узлов сетки σ_Π равно MN. Перейдем в плоскости П к новым координатам. Приняв размеры сетки X=У=1, получаем:

$$P = M/2; \quad Q = N/2; \quad \Delta x = 1/M; \quad \Delta y = 1/N. \qquad \textbf{(4.2.2)}$$

Следовательно, координаты узлов сетки σ_Π выразятся так:

$$x_m = m\Delta x = m/M, \quad m = 0,\ldots,M; \qquad y_n = n\Delta y = n/N, \quad n = 0,\ldots,N. \qquad \textbf{(4.2.3)}$$

Число узлов сетки σ_Γ выбирают так, чтобы было обеспечено взаимно однозначное соответствие между изображениями, заданными на σ_Π и его дискретным преобразованием Фурье, заданным на σ_Γ. Это число узлов также оказывается равным MN. Последнее определено тем, что в системе, состоящей из MN точек, полной является система тригонометрических функций с частотами

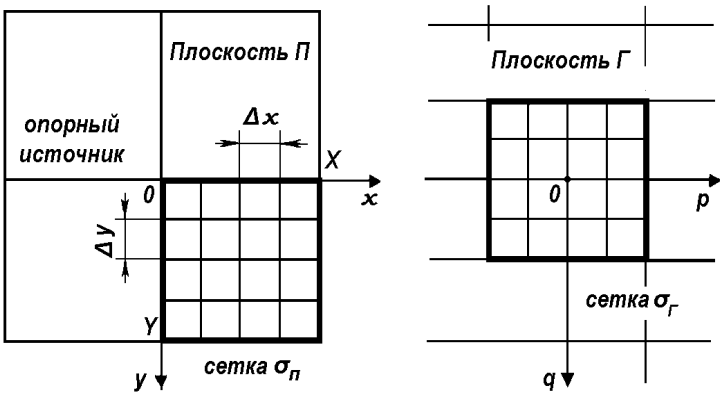


Рис.4.2.1 *Расположение сеток.*

$$P = 0,\pm 1,\ldots,\pm\frac{M}{2}-1,\frac{M}{2}; \qquad Q = Q,\pm 1,\ldots,\pm\frac{N}{2}-1,\frac{N}{2}. \qquad \textbf{(4.2.4)}$$

Соотношения между размерами сеток σ_Π и σ_Γ получим из (4.2.1) с учетом того, что $p = x_1/\lambda f_1$ и $q = y_1/\lambda f_1$ согласно (3.5.7)

$$2X = (\lambda f_1 M)/X; \quad 2Y_1 = (\lambda f_1 N)/Y. \qquad \textbf{(4.2.5)}$$

Выбор сеток в плоскостях П и Г означает, что все непрерывные функции в этих плоскостях могут быть представлены своими дискретными

значениями в узлах сетки. Эти значения теперь являются функциями номеров узлов, т.е. m и n в плоскости П, p и q в плоскости Г. Для отличия от непрерывных величин аргументы дискретных величин будем обозначать индексами, например E_{mn}, вместо $E(x_m,y_n)$, A_{pq} вместо A(p,q). Установим соответствие между основными физическими величинами, рассмотренными ранее, и их цифровыми моделями. Поле в плоскости П представим так:

$$h_{mn} = A_0\delta_{mn} + E_{mn}, \qquad \textbf{(4.2.6)}$$

дискретное преобразование Фурье от h_{mn} определит соотношение:

$$H_{pq} = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} h_{mn}\exp\left[-i2\pi\left(p\frac{m}{M}+q\frac{n}{N}\right)\right] \qquad \textbf{(4.2.7)}$$

Примем с учетом (4.2.6)

$$H_{pq} = A_0 + A_{pq}\exp(ipq), \qquad \textbf{(4.2.8)}$$

Цифровая модель голограммы Фурье, являющаяся аналогом ранее рассмотренной модели (3.5.22), будет иметь вид

$$\tau_{pq} = k_{\Gamma}\left(\alpha - \beta A_{pq}^2 - A_{pq}\cos\vartheta_{pq}\right), \qquad \textbf{(4.2.9)}$$

где

$$\vartheta_{pq} = 2\pi\left(p\frac{m_0}{M}+q\frac{n_0}{N}\right) - \varphi_{pq}. \qquad \textbf{(4.2.10)}$$

Величину H_{pq} можно интерпретировать как коэффициент двойного ряда Фурье от дискретной функции, заданной на двумерном интервале MN . При этом в уравнении голограммы последнее слагаемое является не чем иным, как косинусным коэффициентом Фурье a_{pq} изображения предмета. С учетом изложенного уравнение цифровой голограммы Фурье, удобное для расчетов на ЭВМ, принимает вид:

$$\tau_{pq} = k_{\text{„}}\left(\alpha - \beta\text{¿}_{pq}^{2} - a_{pq}\right). \qquad \textbf{(4.2.11)}$$

Здесь в общем случае имеем

$$a_{pq} = 2\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} E_{mn}\cos\left[2\pi\left(p\frac{m}{M}+q\frac{n}{N}\right)-\Phi_{mn}\right], \qquad \textbf{(4.2.12)}$$

$$b_{pq} = 2\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} E_{mn}\sin\left[2\pi\left(p\frac{m}{M}+q\frac{n}{N}\right)-\Phi_{mn}\right], \qquad \textbf{(4.2.13)}$$

$$A_{pq}^2 = \frac{1}{4}\left(a_{pq}^2 + b_{pq}^2\right). \qquad \textbf{(4.2.14)}$$

В двух первых формулах последние члены в прямоугольных скобках используются при наличии рассеивателя со случайной фазой. Если рассеиватель не используют, то они равны нулю и формула упрощается.

При компьютерном расчете структуры голограммы исходной информацией является изображение, которое разбивают на отдельные участки в соответствии с выбранной сеткой (т.е. из изображения делают выборку значений E_{mn} в узлах сетки), а также задаваемые параметры *M, N,* k_Γ, α, β. В результате расчета должны быть получены величины τ_{pq} прозрачности голограммы в узлах сетки σ_Γ.

Основой вычисления является выполнение дискретного преобразования Фурье (ДПФ), причем двумерное преобразование выполняется в два этапа: сначала по строкам, а затем по столбцам. Последовательность вычислений показана на **рис. 4.2.2.** Для выполнения одномерных преобразований используется алгоритм быстрого преобразования Фурье (БПФ).

В приложении 1 содержится краткое описание процедур ДПФ и БПФ, которые широко вошли в практику компьютерных расчетов. Для удобства вычислений матрицу C_{pq}, полученную после преобразования строк, транспонируют и повторное преобразование также выполняют по строкам. В результате двойного БПФ получают коэффициенты a_{pq} и b_{pq} по которым и определяют значения A^2_{pq}. Результаты вычислений вместе с заданными параметрами используют для расчета прозрачности голограммы по ее формуле. Эти значения и выдает машина.

Отпечатанную цифровую голограмму затем фотографируют с соответствующим уменьшением и используют для восстановления изображения оптическим путем. Очень часто голограмму Фурье пеставляют в двоичном (бинарном) виде. В этом случае ее прозрачность имеет только два значения: 0 или 1. Двоичную голограмму рассчитывают следующим образом. Прозрачность голограммы как функцию пространственных частот обозначим через Θ_{pq}. Выберем некоторый порог A'. Если τ_{pq} больше или равно A', то величине Θ_{pq} сопоставим единицу, в противном случаеĥ нуль. Это возможно записать как

$$\Theta_{pq} = \begin{cases} 1, \text{ при } \tau_{pq} \geq A'; \\ 0, \text{ при } \tau_{pq} < A'. \end{cases} \qquad \textbf{(4.2.17)}$$

В данном случае 1 соответствует уровню белого, а 0 - черного. Окончательно получим

$$\Theta_{pq} = \begin{cases} 1, \text{ при } \beta A_{pq}^2 + a_{pq} \le A_{\text{пор}}, \\ 0, \text{ при } \beta A_{pq}^2 + a_{pq} > A_{\text{пор}}, \end{cases} \qquad \textbf{(4.2.18)}$$

$$\text{где} \quad A_{\text{пор}} = \frac{A'}{k_{\Gamma}} - \alpha.$$

В выборе параметров β и $A_{\text{пор}}$ имеется определенный произвол. В общем случае их увеличение приводит к снижению доли высоких пространственных частот в голограмме. Сама же двоичная голограмма в большой степени подчеркивает высокие пространственные частоты.

4.2.3. Киноформ

Часто встречаются случаи, когда комплексная амплитуда объектной световой волны $\psi(\xi,\eta) = A(\xi,\eta)\exp[i\varphi(\xi,\eta)]$ в плоскости регистрации голограммы практически постоянна по модулю. В таких случаях изображение интересующего объекта может быть восстановлено с использованием только фазовой информации $\varphi(\xi,\eta)$. Как правило, это имеет место, когда голографируемый объект является диффузным или освещен диффузно рассеянным светом. Однако освещение объекта световой волной со специально выбранным детерминированным распределением фаз также приводит к объектной световой волне практически постоянной амплитуды в плоскости регистрации голограммы. Таким образом, в указанных здесь случаях, записав только фазовую информацию

Рис. 4.2.2. *Последовательность вычислений голограммы Фурье*

об объекте, можно восстановить трехмерное изображение интересующего объекта. Получаемая при этом запись называется киноформом. Киноформ не является голограммой в полном смысле этого слова, так как он содержит не полную информацию об объекте, а только фазовую. Киноформ обладает тем замечательным свойством, что в отличие от других типов голограмм при идеальном изготовлении восстанавливает только одно изображение - мнимое или действительное. Это означает, что весь световой поток, дифрагированный киноформом, концентрируется на одном изображении.

Процесс изготовления киноформа выглядит следующим образом.

На компьютере рассчитываются дискретные значения фазы $\varphi(\xi,\eta)$ объектной световой волны. Полученные значения фазы обрабатываются таким образом, чтобы их отклонения от начальной фазы лежали в интервале от 0 до 2π радиан по всей области выборки, т.е. из каждого значения фазы вычитаются величины, кратные 2π радианам. В результате получается двумерный массив, состоящий из дискретных значений фазы

$$\hat{\varphi}(\xi,\eta)=\sum_{n=-N}^{N}\sum_{m=-M}^{M}\varphi_{nm}(\bmod 2\pi)\delta(\xi-n\Delta\xi,\eta-m\Delta\eta),$$

$$\varphi_{mn}(\bmod 2\pi)=\varphi-2\pi j, \qquad \textbf{(4.2.19)}$$

$$j2\pi\le\varphi\le(j+1)2\pi;\quad j=0,1,2,\ldots$$

Данный массив кодируется массивом значений яркости в многоградационной шкале, который уже отображается в виде картины на выходное устройство компьютера, например на дисплей. Полученная картина фотографируется с необходимым уменьшением и конечный фотоснимок отбеливается в дубящем отбеливателе. При отбеливании градации фотографического почернения превращаются в соответствующее распределение значений оптической толщины. Полученный таким образом киноформ имеет функцию пропускания

$$t(\xi,\eta)\sim rect\left[\frac{\xi}{\Delta\xi}\right]rect\left[\frac{\eta}{\Delta\eta}\right]*\exp[\pm ic\varphi(\bmod 2\pi)]\times$$

$$\times\sum_{n=-N}^{N}\sum_{m=-M}^{M}\delta(\xi-n\Delta\xi,\eta-m\Delta\eta), \qquad \textbf{(4.2.20)}$$

Знак показателя экспоненциального сомножителя определяется тем, что используется в качестве киноформа- негатив или позитив фотоснимка картины киноформа. Соответственно и изображение, восстанавливаемое киноформом, будет мнимым или действительным.

Из рассмотрения функции пропускания киноформа (4.2.20) следует, что для восстановления исходного волнового фронта без искажений

необходимо, чтобы константа *c* равнялась единице. Это означает, что свет, падающий на участок с фазой $\hat{\varphi} = 0$, будет задерживаться ровно на одну длину волны по сравнению со светом, падающим на участок с фазой $\hat{\varphi} = 2\pi$. Если такое согласование фаз было достигнуто, то весь свет, падающий на киноформ, будет участвовать в формировании единственного (действительного или мнимого) изображения записанного обекта. В противном случае киноформ подобен осевой голограмме, в которой действительное и мнимое изображения частично накладываются; часть света дифрагирует в нулевой порядок, создавая яркое пятно в центре изображения. Качество изображения резко ухудшается. На практике согласование фаз достигается путем тщательного контроля процессов экспонирования и проявления уменьшенных фотоснимков киноформа, а также отбеливания.

4.3. Фазовая проблема в оптике. Создание на основе решения обратных задач нового класса оптических элементов [1, 2, 6-9]

4.3.1. Извлечение фазовой информации из данных об интенсивности

Как известно, проблема получения фазовых характеристик световых полей возникает в разнообразных оптических исследованиях. Трудности непосредственного измерения фазы в оптическом диапазоне заставляют оптиков искать обходные пути: пытаться извлекать фазовую информацию из данных об интенсивности. Разумеется, попытки найти простые рецепты решения задачи были обречены на неудачу. Однако за последние 25 лет наметилось серьезное продвижение в проблеме восстановления фазовых характеристик световых полей. Работы, по восстановлению фазовых характеристик по характеристикам интенсивности с помощью ЭВМ, набирают размах. Важно подчеркнуть, что при этом привлекаются дополнительные данные о поле, например, во многих случаях используются два (а не одно) распределения интенсивности, относящиеся к двум сечениям поля.

Допустим, что монохроматическое поле, несущее информацию об исследуемом объекте, было зарегистрировано в двух плоскостях: в плоскости изображения и в фурье-плоскости. Будем считать, что поле зависит только от одной пространственной координаты. Обозначим поле в плоскости изображения через $E(x) = a(x)\exp(i\varphi(x))$, в фурье-плоскости -

через $F(x) = A(x)\exp\left(i\Phi(x)\right)$. Для операции преобразования Фурье введем символ

$$LE(x) = \frac{1}{\sqrt{2\pi}}\int E(x')\exp(-ixx')dx', \qquad \textbf{(4.3.1)}$$

так что

$$F(x) = LE(x). \qquad \textbf{(4.3.2)}$$

Пусть при регистрации поля получена информация о функциях *a(x)* и *A(x)*, а информация о фазовых множителях отсутствует. Задача состоит в построении комплексной функции по заданному ее модулю и модулю ее фурье-образа. Для решения этой задачи была предложена итерационная процедура, которая заключалась в следующем. В качестве пробной функции бралась функция, модуль которой совпадал с заданным в плоскости изображения модулем *a(x)*, а фазовый множитель $\exp\left(i\psi(x)\right)$ брался произвольным; в работе он строился с помощью генератора случайных чисел. Пробную функцию удобно обозначить через $y_0(x)$, в дальнейшем индекс будет совпадать с числом выполненных итераций, перед началом итераций имеем

$$y_0(\mathrm{x}) = a(x)\exp\left(i\psi(x)\right). \qquad \textbf{(4.3.3)}$$

Для этой функции строилось ее преобразование Фурье

$$Ly_0 = A_0(x)\exp\left(i\Phi_0(x)\right). \qquad \textbf{(4.3.4)}$$

Затем у полученной функции модуль заменялся на правильный, т.е. на *A(x)*, а фаза сохранялась. Так получали функцию в фурье-плоскости

$$Y_0(x) = A(x)\exp\left(i\Phi_0(x)\right), \qquad \textbf{(4.3.5)}$$

что можно представить как

$$Y_0 = \frac{ALy_0}{\left|Ly_0\right|}. \qquad \textbf{(4.3.6)}$$

Затем выполнялось обратное преобразование Фурье и снова исправлялся модуль, при этом получалась функция в плоскости изображения в первом приближении

$$y_1 = \frac{aL^{-1}Y_0}{\left|L^{-1}Y_0\right|}. \qquad \textbf{(4.3.7)}$$

На следующем шаге выполнялись такие же преобразования. Всю итерационную процедуру можно представить формулами

$$Y_n = \frac{ALy_n}{\left|Ly_n\right|}, \quad y_{n+1} = \frac{aL^{-1}Y_n}{\left|L^{-1}Y_n\right|}. \qquad \textbf{(4.3.8)}$$

Если процесс окажется сходящимся, то функцию, получающуюся в качестве предела итераций, будем обозначать через y

$$y = \lim_{n\to\infty} y_n . \qquad \textbf{(4.3.9)}$$

Описанный алгоритм расчета распределения фазы известен в литературе как алгоритм Гершберга-Сэкстона Практика его использования показала, что во многих случаях процесс действительно сходился, и функция y совпадала с исходным комплексным полем $a(x)\exp(i\varphi(x))$. Несмотря на то, что в настоящее время разработан ряд более совершенных и корректных методов решения фазовой проблемы алгоритм Гершберга-Сэкстона остается одним из самых популярных.

4.3.2. Особенности расчета характеристик фокусаторов и корректоров излучения

К фазовой проблеме очень часто относят широкий класс обратных задач теории дифракции, направленных на расчет фазовой функции оптических элементов, обеспечивающих формирование определенных световых структур.

Если задана фазовая функция оптического элемента, то в принципе всегда можно решить прямую задачу дифракции волн на оптическом элементе и получить распределение поля в интересующей нас области. Сложнее обстоит дело с решением обратной задачи.

Для уяснения физической сущности проблемы рассмотрим **рис. 4.3.1**. Плоский оптический элемент *Ф*, расположенный в области *G* плоскости *u=(u,v)*, освещается пучком *E* монохроматического излучения длины волны. Требуется сформировать в области *D* плоскости *x=(x,y)* волновое поле *I(x,z)*. Фазовая функция оптического элемента полностью определяет поведение пучка за плоскостью u и, в частности, в интересующей нас области *x*. Задача состоит в отыскании фазовой функции оптического элемента *(u,y)*, обеспечивающего формирование волнового поля. С математической точки зрения обратная задача является некорректной: во-первых, решение может вообще не существовать; во-вторых, оно может быть неоднозначным; в-третьих, оно может быть неустойчивым.

Отсутствие решения, как правило, обусловлено ограничениями, налагаемыми на область фокусировки фундаментальными физическими законами. Например, хорошо известно, что нельзя получить в области фокусировки пятно, диаметр которого *D* меньше, чем дифракционный предел, определяемый поперечным размером *d* линзы (фокусатора), длиной волны и фокусным расстоянием *f*, то есть *Df/d*. Аналогичные ограничения накладываются на минимально достижимую толщину линии при

фокусировке в отрезок на плоскости, перпендикулярной оси распространения. Решение задачи фокусировки может отсутствовать не только по причине противоречия законам дифракции. В случае, когда область фокусировки не расположена в плоскости перпендикулярной оптической оси, а является, например, трехмерной пространственной кривой, вполне определенные ограничения накладываются из-за того, что в силу закона прямолинейного распространения света перенос энергии осуществляется вдоль оптической оси. Нельзя, в частности, требовать, чтобы при фокусировке в тонкий цилиндр, расположенный вдоль оптической оси, энергия целиком была сосредоточена в нем, а интенсивность вне тонкого цилиндра равнялась нулю. Ясно, что энергия к более удаленным от фокусатора точкам цилиндра переносится через периферийные области пространственных сечений, более близких к фокусатору.

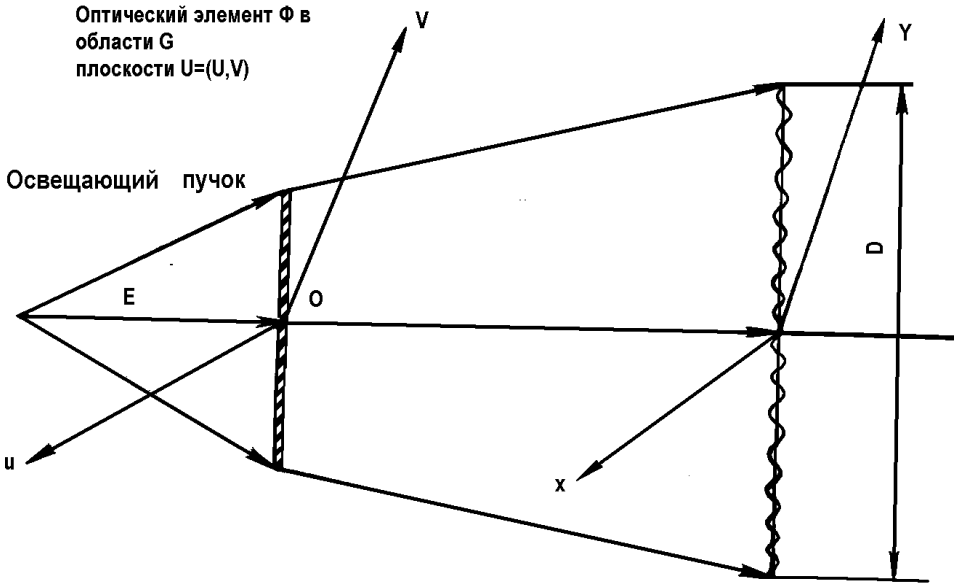


Рис. 4.3.1 *Постановка задачи формирования оптического поля*

Успехи в решении обратных задач привели к качественному обновлению элементной базы современных оптических систем. Так, имеются практические достижения в создании киноформных линз большой светосилы. Сообщается о создании "безаберрационных" объективов, содержащих несколько киноформных линз и обладающих прекрасными массогабаритными характеристиками. Наконец, есть несомненные достижения в области создания различных типов фокусаторов. В самом

широком смысле термин "фокусатор" используется среди специалистов для обозначения элементов компьютерной оптики, обеспечивающих концентрацию световой энергии в пределах пространственной области с заранее заданной пространственной конфигурацией. Чаще всего речь идет о фокусировке излучения в некоторую фокальную кривую с заданным распространением интенсивности на ней. Зачем нужны фокусаторы? Прежде всего для лазерных технологических установок в промышленности и медицине. Лазер без фокусатора - это только генератор излучения, лазер с фокусатором - это уже компонент гибкой производственной системы с программируемым режимом технологических операций. Но производство - не единственная область применения фокусаторов излучения. Закон пространственного распределения энергии в фокальной области оптической системы определяет режим нагрева мишеней при лазерном управляемом термодинамическом синтезе, течение химических реакций, стимулированных лазерным излучением. В оптическом приборостроении часто требуется сложная форма фокальной кривой.

На сегодняшний день созданы фокусаторы излучения в видимом и инфракрасном диапазонах с регулировкой интенсивности вдоль фокальной линии. Эти результаты являются наглядной иллюстрацией достижений компьютерной оптики.

Огромные возможности открывает компьютерная оптика для получения оптических элементов, позволяющих корректировать амплитудно-фазовое распределение поля в световых пучках. Такого рода корректоры позволяют сформировать волновой фронт заданной формы. К числу корректоров принадлежат, в частности, компенсаторы- элементы, преобразующие плоский или сферический волновой фронт в асферический произвольного порядка. Основное назначение компенсаторов - контроль оптических поверхностей. При этом компенсатор формирует эталонный волновой фронт для интерферометрического исследования изготавливаемой оптической поверхности или же играет роль "нулевой линзы", сводя асферическую задачу к сферической.

4.3.3. Дифракционные оптические элементы

Фокусаторы и корректоры излучения чаще всего выполняются в виде дифракционных зонных пластинок. Их расчет основан на процедуре приведения их фазовой функции к интервалу $[0, 2\pi)$, которая использовалась ранее при вычислении пропускания киноформов. Последнее обстоятельство привело к тому, что указанный класс оптических дифракционных элементов часто относят к так называемой киноформной оптике. Для уяснения основных принципов построения дифракционных

элементов кратко рассмотрим характеристики простейших фазовых элементов, являющихся аналогами классических линз и оптических призм.

Плоско-сферическая линза

Как следует из выражения (3.2.2), фазовая функция сферической линзы в параксиальном приближении имеет вид

$$\varphi(\xi,\eta) = \varphi(r) = -k\frac{\xi^2+\eta^2}{2F} \text{ при} \le \frac{D}{2}, \qquad \textbf{(4.3.10)}$$

где F - фокусное расстояние; D - диаметр линзы; $r = \sqrt{\xi^2+\eta^2}$.

Приведение фазовой функции (4.3.10) к интервалу [0, 2π) показано на **рис. 4.3.2.**

Если материал линзы имеет коэффициент преломления n, то максимальная высота рельефа составляет

$$h_{\max} = \frac{\lambda}{n-1}$$

и имеет порядок длины волны. Высота микрорельефа определяется по формуле

$$h(r) = \frac{\lambda}{n-1}\frac{1}{2\pi}\varphi(r) \bmod 2\pi \qquad \textbf{(4.3.11)}$$

Радиусы зон Френеля можно найти из соотношения

$$\varphi(r_j) = -2\pi j,$$

откуда следует

$$r_j = \sqrt{2\lambda F j} \qquad \textbf{(4.3.12)}$$

Число полных зон j_0 на линзе определяется из условия $r_{j0} \le \frac{D}{2}$ и удовлетворяет соотношению

$$j_0 = \left\rceil \frac{D^2}{8\lambda F} \right\lceil, \qquad \textbf{(4.3.13)}$$

где $\rceil\ \lceil$ - означает целую часть числа с округлением в меньшую сторону.

Плоская цилиндрическая линза

Рассмотрим цилиндрическую линзу (**рис. 4.3.3**), описываемую фазовой функцией

$$\varphi(\xi) = -k\frac{\xi^2}{2F} \text{при} |\xi| \le D/2. \qquad \textbf{(4.3.14)}$$

Приведение фазовой функции к интервалу [0, 2π] аналогично проведенному выше для сферической линзы. Границы зон в данном случае - прямые линии, а расстояния между ними определяются формулой (4.3.12) при замене r_j на ξ_j. На **рис. 4.3.4** приведен фотошаблон плоской цилиндрической линзы.

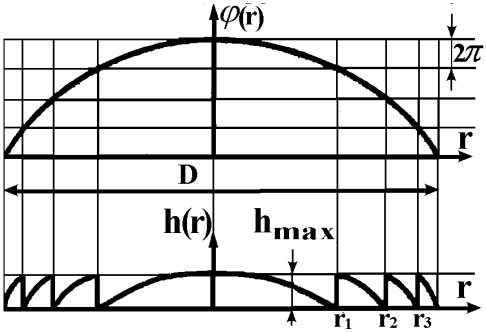


Рис. 4.3.2. *Приведение фазовой функции линзы к интервалу [0,2π]*

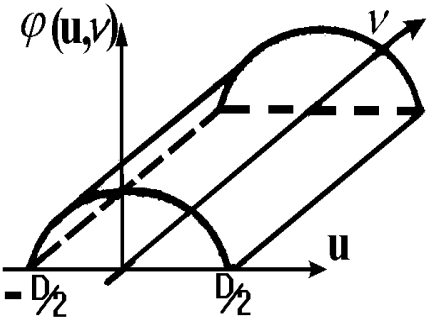


Рис. 4.3.3. *Цилиндрическая линза*

Плоская призма

Рассмотрим призму с углом α (**рис. 4.3.5**).

Призма обеспечивает фазовый сдвиг, линейно зависящий от координаты, и характеризуется фазовой функцией

$$\varphi(\xi) = \frac{2\pi}{\lambda}\xi\theta, \qquad \textbf{(4.3.15)}$$

где $\theta = (n-1)tg\alpha$.

Приводя фазовую функцию $\varphi(\xi)$ к интервалу [0, 2π) (см. **рис. 4.3.6**) получим 1-*D* дифракционную решетку "с блеском" (**рис. 4.3.7**).

Максимальная высота рельефа:

$$h_{\max} = \frac{\lambda}{n-1}$$

и уравнение высоты микрорельефа имеет вид

$$h(\xi) = \frac{\lambda}{n-1}\frac{1}{2\pi}\varphi(\xi) \bmod 2\pi. \qquad \textbf{(4.3.16).}$$

Различным углам отклонения β соответствуют различные периоды решетки

$$d = \lambda / \sin\beta, \qquad \textbf{(4.3.17),}$$

где $\sin\beta = (n-1)tg\alpha$.

Бинарная 1-*D* амплитудная дифракционная решетка получается при замене линейно-меняющейся фазовой функции в пределах одного периода на двоичную функцию, принимающую значения $\pm\pi/2$ (*рис. 4.3.8*).

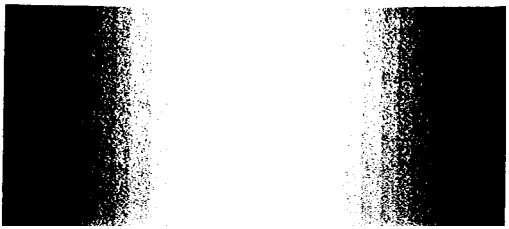


Рис. 4.3.4. *Фотошаблон плоской цилиндрической линзы.*

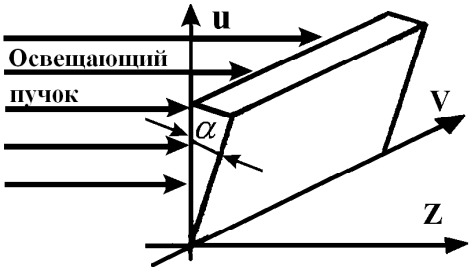


Рис. 4.3.5. *Призма.*

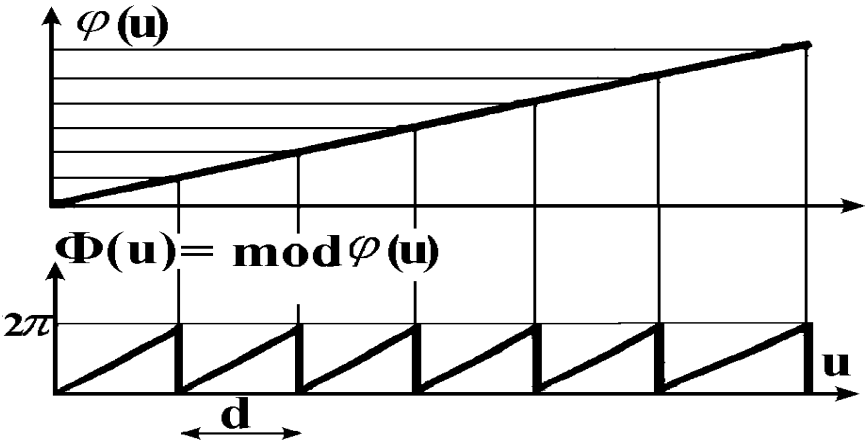


Рис. 4.3.6. *Фазовая функция дифракционной решетки "с блеском".*

Рис. 4.3.7. *Дифракционная решетка "с блеском".*

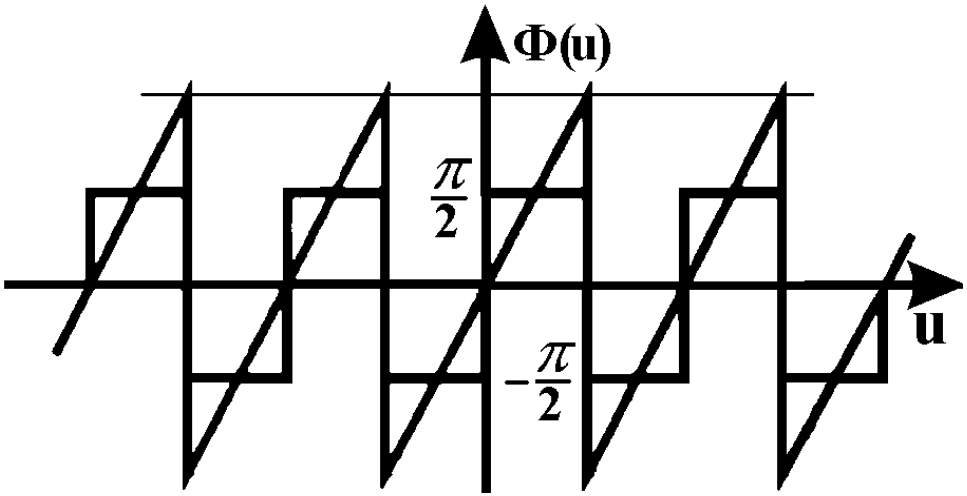


Рис. 4.3.8. *Сравнение бинарной дифракционной решетки и решетки "с блеском".*

4.3.4. Создание фокусаторов на основе управляемых зеркал.

Альтернативным техническим решением по отношению к плоским дифракционным элементам киноформного типа является использование гибких управляемых зеркал, поверхность которых может принимать ту или иную форму в зависимости от управляющих напряжений, приложенных к зеркалу. Гибкое зеркало, управляемое ЭВМ, позволяет по заданной программе изменять интенсивность в зоне фокусировки.

Задача расчета формы поверхности гибкого зеркала имеет особенности. Функцию $u(\bar{\mathbf{r}})$, описывающую профиль зеркала, к которому приложены управляющие воздействия a_j, можно представить в виде:

$$u(\bar{\mathbf{r}}) = \sum_{j=1}^{N} a_j S_j(\bar{\mathbf{r}}), \qquad \textbf{(4.3.18)}$$

где $S_j(\bar{\mathbf{r}})$ - функции отклика зеркала, зависящие от конструкции и технологии его изготовления. Функция отклика $S_j(\bar{\mathbf{r}})$ описывает форму поверхности зеркала при единичном управляющем воздействии, приложенном к j-му приводу. При этом, естественно, предполагается линейность отклика. Расчет требуемого профиля поверхности зеркала сводится к определению конечного числа управляющих воздействий $\{a_j\}$.

Общая схема гибкого зеркала мембранного типа приведена на **рис. 4.3.9** (патент фирмы "Perkin-Elmer").

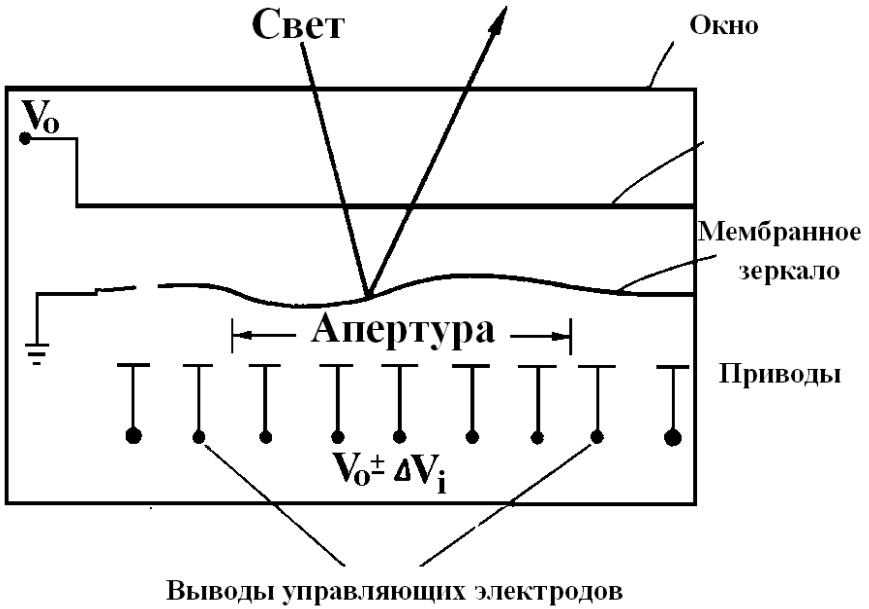


Рис. 4.3.9. *Схема управляемого зеркала мембранного типа.*

Мембрана толщиной обычно 0.5-1.5 мкм помещается между прозрачным электродом, к которому приложено напряжение смещения (V_0), и группой приводов, представляющей собою набор проводящих прокладок, к которым приложено напряжение $V_0 \pm V_s$, где V_s - напряжение, создаваемое сигналом. Мембрана заземлена, а расстояние между ней и электродами составляет 50-100 мкм. При отсутствии сигнала V_s суммарное усилие, приложенное к мембране, равно нулю, и в этом случае мембрана не испытывает никаких отклонений. плоская форма мембраны сохраняется с точностью до среднеквадратичного отклонения λ/20. Если к какому-либо из электродов приложить напряжение V_s, то мембрана отклонится, причем центр деформации локализуется над данным электродом. При напряжении менее 100 В величина прогиба мембраны составляет обычно ±0.5λ. Используя для отклонения мембраны большое число приводов, нетрудно получать прогиб величиной во много длин волн. Общий вид описанного мембранного зеркала приведен на **рис. 4.3.10**.

Перспективным представляется создание комбинированных формирующих оптических элементов, состоящих из киноформа и гибкого

зеркала. Эти элементы удачно дополняют друг друга. Гибкое зеркало позволяет создать сравнительно плавную оптическую поверхность с нужной глубиной рельефа Задача киноформного элемента- обеспечить требуемое пространственное разрешение при сравнительно небольшой глубине рельефа. Число зон при этом становится небольшим, и эффективность киноформа повышается (меньше сказывается пространственная дискретизация фазовой функции). Кроме того, комбинированный оптический элемент позволяет в некоторой степени скорректировать небольшие отклонения параметров формирующей системы от расчетных

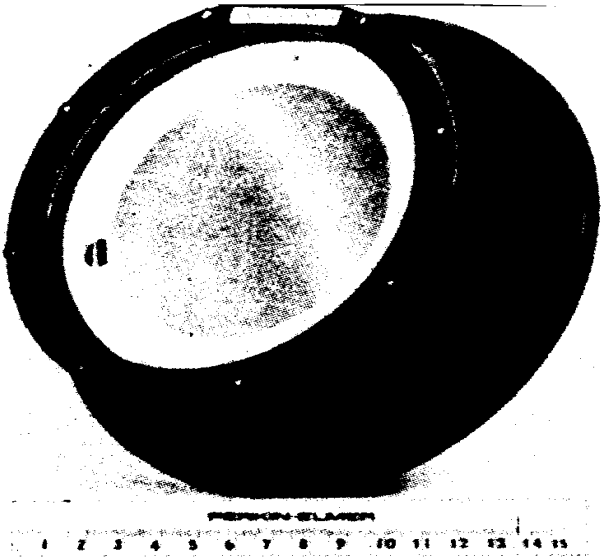


Рис. 4.3.10. *Общий вид мембранного зеркала.*

(например, ввести коррекцию угла падения излучения на элемент, расстояния до плоскости фокусировки, ширины светового пучка на элементе и т.д.).

4.4. Фокусировка излучения при наличии случайных помех. Использование методов адаптивной оптики [7-9]

Очень часто излучение, идущее от фокусатора, достигает заданного объекта, пройдя через область пространства с нестационарной случайно-неоднородной средой. В этом случае эффективно реализовать

фокусирующие свойства гибкого зеркала можно лишь перманентно меняя форму его отражающей поверхности с целью компенсации влияния неоднородностей. Как известно, оптические элементы, характеристики которых могут под воздействием управляющих сигналов изменяться во времени, составляют элементную базу адаптивной оптики. Таким образом, проблема создания фокусаторов, способных работать в условиях случайных помех, оказывается тесно связанной с необходимостью использования методов адаптивной оптики. Среди этих методов важную роль играет метод фазового сопряжения.

Метод фазового сопряжения следует из принципа оптической обратимости. В приближении геометрической оптики его можно сформулировать так: если волну $\psi = A\exp(ik\varphi)$ $\left(k = \frac{2\pi}{\lambda}\right)$, прошедшую участок с неоднородным показателем преломления, послать обратно по тому же самому пути, заменив фазу этой волны φ на $u=-\varphi$, то на выходе из среды волна восстановит свой первоначальный неискаженный фазовый профиль. Замена фазы φ на $-\varphi$ эквивалентна операции сопряжения комплексной амплитуды волны: $\psi^* = A\exp(-ik\varphi)$. Этим объясняется название метода фазового сопряжения. **Рис. 4.4.1** иллюстрирует сказанное.

Пусть плоская волна проходит участок оптически неоднородной среды, например кусок стекла. В результате фаза волны искажается и на

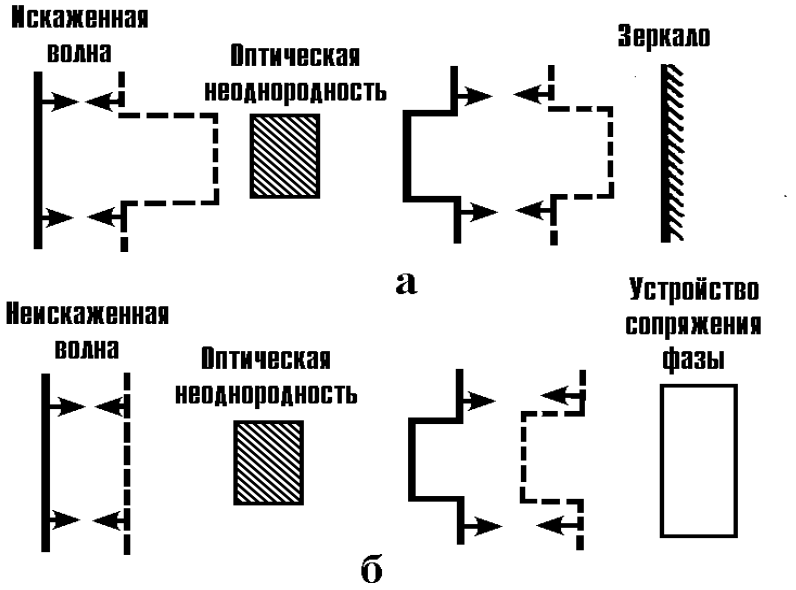


Рис. 4.4.1. *Компенсация возмущающего фактора оптической неоднородности.*

выходе из среды приобретает характерную впадину (сказалось относительное увеличение оптической длины пути и времени распространения волны через стекло). При отражении от обычного зеркала запаздывание на особом участке сохраняется и при повторном прохождении фазовая неоднородность удваивается (**рис. 4.4.1, а**). Для того чтобы скомпенсировать первоначальное отставание по фазе при обратном распространении, необходимо обратить фазу, т.е. сформировать волну с профилем фазы -φ. Фаза этой волны на месте впадины имеет выступ, равный ей по значению **(рис. 4.4.1, б**). Этот особый участок теперь уже опережает по фазе остальные. После повторного прохождения волной неоднородности опережающий участок отстанет по фазе ровно настолько, чтобы скомпенсировать введенное возмущение. Волна в результате двукратного распространения останется плоской (**рис. 4.4.1, б**).

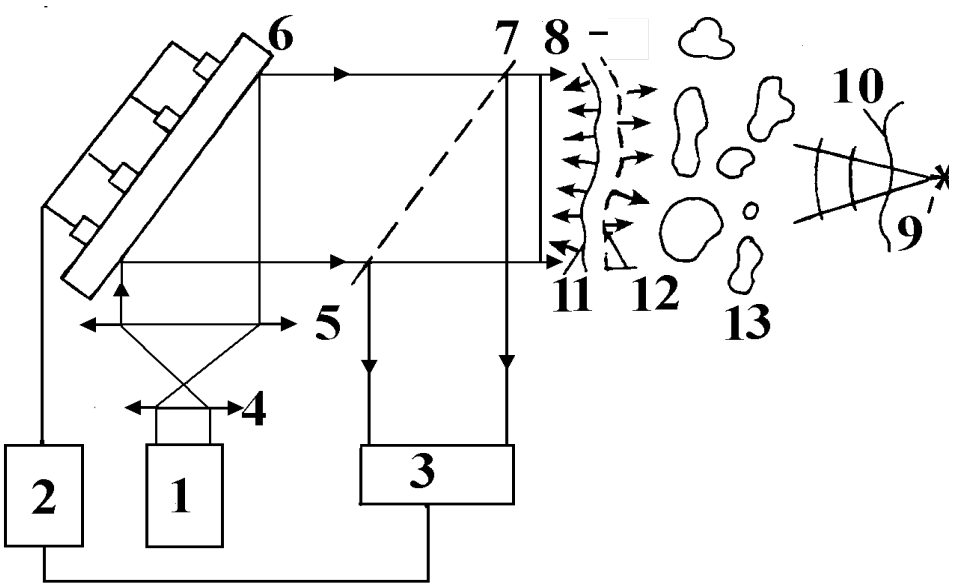


Рис. 4.4.2. *Оптическая фокусирующая система, реализующая принцип фазового сопряжения. 1-лазер; 2-ЭВМ; 3-датчик волнового фронта; 4, 5-линзы телескопа; 6-мембранное зеркало; 7-ответвитель; 8-начальный фронт волны; 9-мишень; 10-фронт в области мишени; 11-фронт отраженной волны; 12-скорректированный фронт; 13-случайные неоднородности.*

Рассмотрим реализацию идеи фазового сопряжения на примере фокусатора, использующего для коррекции фазы гибкое зеркало (**рис. 4.4.2**).

Световой пучок от лазера 1, пройдя формирующий телескоп, образованный линзами 4 и 5, и отразившись от гибкого зеркала 6, выходит из системы, имея изначально плоский волновой фронт 8. Пройдя область со случайными неоднородностями (в частности, это может быть турбулентная атмосфера) вблизи мишени 9 волна будет иметь значительно возмущенный фронт 10. Часть энергии, отразившись от мишени в виде сферической волны, которая в данной оптической системе будет играть роль опорного пучка. К фокусатору этот пучок подойдет с сильно искривленным волновым фронтом 11. Степень отклонения этого фронта от фронта плоской волны определяется от датчика фазы 3, излучение на который подается при помощи светоделительной пластины 7. Датчик фазы, представляющий собой интерферометрическое или голографическое устройство, регистрирует фазовый профиль, пришедшей от мишени волны. Поступающая от датчика фазы информация перерабатывается ЭВМ 2 и в виде управляющих сигналов подается на гибкое зеркало. Это приводит к формированию на выходе фокусатора волны с обращенным (фазово-сопряженным) по отношению к опорному пучку фронтом 12. При распространении этой волны фазовые неоднородности будут компенсироваться. В результате излучение полностью сфокусируется на мишени.

4.5. Оптические элементы для анализа и формирования поперечного состава излучения [1]

Оптические элементы этого класса позволили решить целый ряд практически интересных задач, связанных с задачами анализа и формирования поперечно-модового состава лазерного излучения. Как известно, каждой моде соответствует известная математическая функция двух переменных. Многомодовое излучение падает на транспарант, пропускание которого определяется функцией $\psi_k(u,v)$. Компьютер формирует в своей памяти двумерный массив чисел, соответствующих моде с номером k, а фотопостроитель переводит числа в значения оптической плотности на фотоматериале. В результате получаем набор плоских оптических элементов, соответствующих различным модовым функциям. Используя их как оптические элементы, можно построить прибор для анализа и формирования поперечно-модового состава излучения. Рассмотрим принцип действия анализатора. Многомодовое излучение

падает на транспарант, пропускание которого определяется функцией $\psi_k(u,v)$. Интенсивность света в фокусе выходной линзы при этом равна интенсивности *k*-й моды. Меняя транспаранты, можно измерить интенсивность различных мод и таким образом решить задачу определения поперечного состава. На практике на одном оптическом элементе с использованием голографического метода можно записать сразу несколько модовых функций. При освещении такого транспаранта многомодовым излучением одномоментно в различных точках фокальной плоскости линзы измеряется интенсивность различных мод. Описанный оптический элемент подобен дифракционной решетке, разделяющей по углам излучение различных длин волн. Однако в данном случае решается гораздо более серьезная задача: по углам разделяются поперечно-модовые составляющие излучения. Имея набор оптических элементов, согласованных с модовыми функциями, можем решить также и задачу формирования требуемого поперечно-модового состава лазерного излучения. Как практически использовать выше означенные возможности? Поскольку при распространении излучения в волоконном световоде (ВОЛС) поперечно-модовая структура его обладает устойчивостью, имеется возможность уплотнить канал передачи информации. Для этого нужно использовать в качестве переносчиков сообщения колебания, соответствующие различным поперечным модам. Число одновременно возбуждаемых на входе ВОЛС колебаний может составлять от нескольких единиц до нескольких десятков. Каждое колебание-переносчик несет свое сообщение и распространяется по волоконному световоду независимо от других переносчиков. На выходе ВОЛС проводится анализ поперечно-модового состава излучения и индивидуальная демодуляция сообщения в каждой из выделяемых мод. Пропускная способность ВОЛС при модовом уплотнении резко возрастает. При этом усложнение аппаратуры связи, особенно на приемном конце, не очень существенно.

Теперь уместным будет обсудить физический смысл поперечно-модовых разложений излучения и ответить на вопрос: "Являются ли поперечные моды выдумкой математиков или же действительно излучение в неоднородной среде состоит из множества мод?"

Аналогичный вопрос ставился в дискуссии о природе белого света: "В самом деле в солнечном свете есть монохроматические волны различного цвета или мы только математически выражаем этот процесс суммой синусоид?" Современная точка зрения на последний вопрос состоит в том, что спектральное разложение приобретает конкретный физический смысл при взаимодействии излучения со спектральным прибором. В этом случае оно оказывается физически адекватным представлением, соответствующим

сущности физической задачи. Точно так же и разложение излучения по модам становится физически адекватным и целесообразным при взаимодействии излучения с анализатором мод. Примечательно, что этот шаг удалось сделать только на основе применения компьютера.

В природе нам не известны генераторы эталонных поперечных мод, подобные генераторам монохроматического излучения. Отсутствуют также оптические элементы, подобные призмам и дифракционным решеткам, но предназначенные для проведения поперечно-модового анализа. Таким образом, компьютерная оптика восполняет существенный пробел путем создания искусственных эталонов физических величин по их математическим моделям. Вполне возможно, что в дальнейшем будут открыты новые физические явления и созданы соответствующие приборы без применения компьютеров. Однако, это уже ни в коей мере не повлияет на оценку роли компьютерной оптики в задаче анализа и формирования поперечно-модового состава излучения.

4.6. Цифровая обработка полей в оптических системах [10-13]

Важнейшей задачей оптики всегда было повышение качества и информационной производительности оптических приборов. В настоящее время современная оптико-электронная техника, по существу, решила проблему формирования изображений высокого качества и большой информационной емкости для большинства практических задач. Важнейшей стала проблема эффективности использования заключенного в них огромного объема информации, т.е. проблема обработки изображений, голограмм и интерферограмм.

Фундаментальным методом решения этой проблемы является развитие на основе теории информации и теории сигналов информационных аспектов теории оптического изображения, оптических систем и привлечение современных методов обработки сигналов, из которых важнейшими в настоящее время являются средства компьютерной техники. Оптические приборы с вычислительными устройствами теперь усиливают не только оптические свойства зрения, но и его аналитические возможности. Это привело к тому, что в рамках общего научного направления обработки оптических полей сформировалась еще одна дисциплина, называемая иконикой. В настоящее время иконика объединяет теоретические и экспериментальные исследования, направленные на всестороннее изучение свойств изображений в "тесной увязке" их со зрительным восприятием.

Техническое обеспечение процедуры компьютерной обработки оптических полей помимо персональных и специализированых компьютеров включает сканеры, ПЗС-матрицы, телекамеры и т.п. На начальных этапах компьютерная оптика заимствовала эти устройства из других областей, таких как машинная графика и автоматизированное проектирование. Но в последнее время стали создаваться и получили широкое распространение специализированные технические средства компьютерной оптики, среди которых особое место занимают параллельные процессоры для обработки многомерных сигналов.

4.6.1. Виды обработки оптических полей

Как научное направление, обработка оптических полей соприкасается с другими ветвями информационных и компьютерных наук - с распознаванием образов, искусственным интеллектом, компьютерным зрением, телевидением, интроскопией, акустоскопией, радиоголографией, томографией. Задачи, решаемые в рамках данного направления, можно классифицировать следующим образом.

Синтез изображений по сигналам, получаемым с датчиков физических полей. Это задача цифровой обработки сигналов датчиков, направленная на их преобразование в форму, пригодную для визуализации. Сюда, например, относится томографический синтез, цифровое восстановление акустических и радиоголограмм, формирование изображений в оптических и других системах с кодированной апертурой и т.д.

Коррекция искажений. Это задача компьютерной обработки искаженного изображающей системой или датчиком изображения или сигнала, направленная на получение изображения (сигнала), соответствующего выходу идеальной изображающей системы. Такова, например, обработка изображения для повышения его резкости и подавления случайных помех.

Препарирование изображений. Идеальная изображающая система не обязательно дает изображение, наилучшим образом отвечающее требованиям конкретной задачи визуального анализа и выделения информации, поскольку требования к идеальным характеристикам практически являются компромиссом между требованиями широкого класса задач. Для отдельных конкретных задач могут потребоваться дополнительные преобразования сигнала, облегчающие визуальный анализ путем подчеркивания одних особенностей и деталей изображения и устранения других, изменения пространственных соотношений, измерения и визуализации количественных характеристик и т.п. Такие преобразования, являющиеся инструментом визуального анализа изображений

профессиональным пользователем-экспертом в конкретной области применения данной визуальной информации, и называются препарированием изображения. Существенная особенность препарирования изображения - диалоговый режим обработки под управлением и наблюдением пользователя, неформализованный опыт и знания которого как бы включаются при этом в систему обработки.

Характерной особенностью препарирования изображения является также то, что в работу пользователю-эксперту предлагается для сравнительного анализа не само исходное (скорректированное) изображение, а варианты в виде цветных, стереоскопических, подвижных (анимационные) изображений объекта исследования по результатам выявления существенных признаков.

Кодирование изображений. Это вид обработки, при котором осуществляется преобразование оптических сигналов и изображений в форму, удобную для передачи по каналам связи и архивного хранения. Важной составной частью кодирования является, в частности, сжатие данных, т.е. кодирование, снижающее требования к пропускной способности каналов связи и емкости запоминающих устройств.

Измерение изображений полей. Под этой задачей подразумевается обработка оптических сигналов, в результате которой получаются количественные данные, предназначенные для принятия решений и управления автоматическими исполнительными механизмами.

Последний тип обработки оптических полей является необходимым атрибутом различных автоматизированных измерительных систем. Ниже в качестве примера рассматривается одна из таких систем, программное обеспечение которой было разработано на кафедре оптики и спектроскопии физического факультета МГУ с целью обработки изображений лазерных пучков со сложной пространственной структурой.

4.6.2. Автоматизированная измерительная система для диагностики структуры лазерных пучков

В автоматизированной измерительной системе (АИС) можно выделить две основные части: устройство ввода изображения (УВИ) и персональный компьютер (ПК), снабженный соответствующим программным обеспечением. УВИ регистрирует в телевизионном стандарте распределение интенсивности светового пучка и после оцифровки видеосигналов вводит информацию о структуре поля в память ПК. Качество и особенности структуры изображения можно оценивать либо визуально, наблюдая его на мониторе компьютера, либо путем расчета по

определенным программам ряда характеристик, о которых будет сказано ниже.

Важнейшим элементом УВИ является камера с матричным фотоприемником. Фотоприемник представляет собой матрицу полупроводниковых элементов, работающих на принципе зарядовой связи (ПЗС-матрица). Снимаемое с каждого элемента напряжение линейно зависит от его освещенности, то есть от интенсивности соответствующего участка падающего излучения.

При "считывании" напряжения с каждого элемента матрицы на выходе приемника формируется последовательность электрических сигналов, в которой закодировано распределение интенсивности. Для оцифровки сигналов от камеры в данной системе применяется компьютерная плата "Видеопорт-2", позволяющая фиксировать процессы

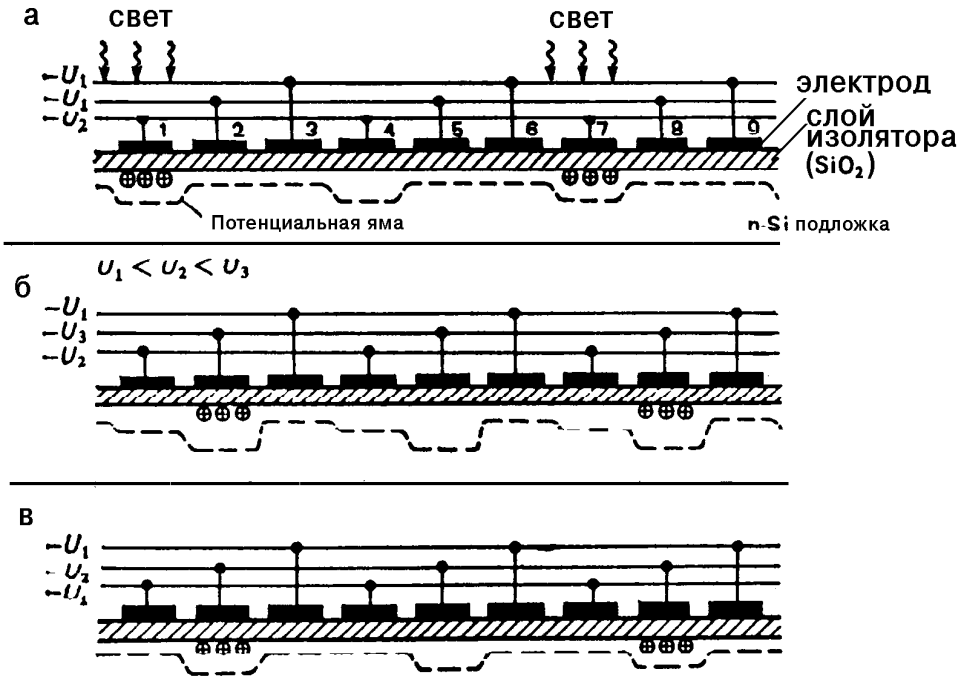


Рис. 4.6.1 *Схема регистрации излучения ПЗС-матрицей*

длительностью приближенно 20 мс. Время ввода оцифрованного сигнала в компьютер зависит только от быстродействия последнего. Программа ввода

и первичной обработки изображений с применением платы "Видеопорт-2" работает под управлением операционной системы MS-DOS и позволяет вводить и сохранять на диске компьютера черно-белые изображения размером до 512*512 точек с цветовым разрешением до 64 градаций яркости.

Матрица имеет форму прямоугольника с длиной диагонали 8.8 мм и несет 704S560 элементов. Столь плотное расположение элементов матрицы и их малые размеры обеспечивают очень высокую разрешающую способность - 50 линий на миллиметр. Это означает, что матрица может различать две линии в структуре изображения, расположенные на расстоянии 20 мкм друг от друга. Укажем для сравнения, что средняя толщина человеческого волоса - 50 мкм.

ПЗС-матрица выполнена на полупроводниковой подложке n-типа нанесением металлических электродов с изоляцией от полупроводника тонким слоем диэлектрика.

Принцип работы ПЗС-матрицы демонстрирует **рис. 4.6.1.** Если на электрод подается отрицательное напряжение, то на поверхности полупроводника формируется обедненный слой, в котором нет свободных зарядов, способных перемещаться. Фактически обедненный слой представляет собой потенциальную яму, в которой могут накапливаться неосновные носители заряда из числа носителей, возникающих под действием света.

Передача заряда осуществляется под действием приложенного к управляющим контактам напряжения трехфазного сигнала. Сначала, как показано на **рис. 4.6.1,а**, накапливаются дырки в потенциальной яме, созданной высоким отрицательным напряжением, приложенным к контакту 1.

Если теперь приложить высокое отрицательное напряжение к контакту 2, то потенциальная яма углубится в сторону этого контакта, и туда же переместятся дырки (**рис. 4.6.1,б**). Сразу же конфигурацию потенциальных ям приводят к виду, показанному на **рис. 4.6.1,в.** Затем следует новый цикл. При повторении циклов заряд последовательно перемещается и воспринимается как изменяющийся во времени сигнал, несущий информацию об интенсивности и пространственном распределении излучения.

Помимо ПЗС-матрицы в УВИ входит еще несколько электронных узлов системы "Видеопорт-2", обеспечивающих в определенном режиме считывание и оцифровку видеосигналов и их ввод в память компьютера. Эти узлы и их функциональные связи в упрощенном виде изображены на **рис. 4.6.2.** Сигналы с камеры попадают в 6-разрядный аналогово-цифровой

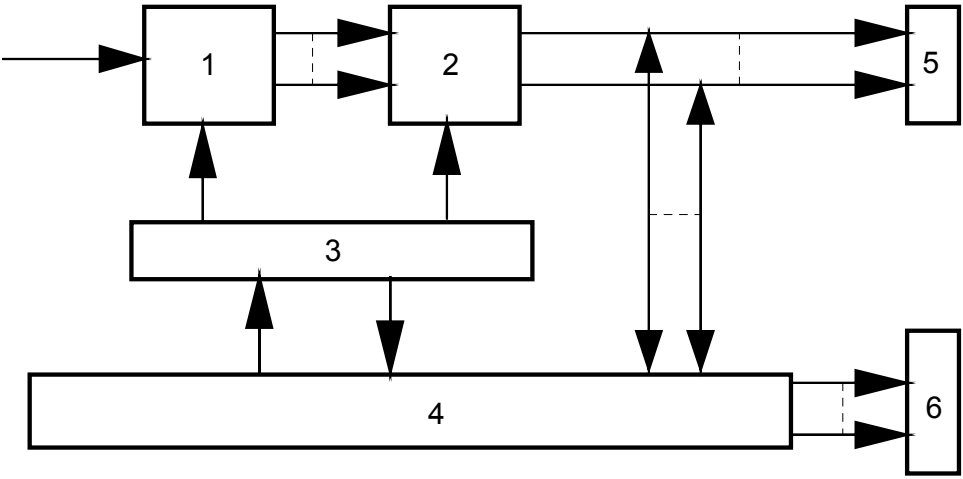


Рис. 4.6.2 *Блок-схема устройства для обработки видеосигнала.*
1- АЦП; 2- буфер; 3- генератор; 4- схема управления и адресации; 5,6- шина данных и адресная шина компьютера.

преобразователь (АЦП) 1, что позволяет получать до 64 градаций яркости. Оцифрованные данные накапливаются в буфере 2. В нем каждой точке изображения соответствует одна ячейка, а каждой строке изображения - строка матрицы. Такое построение упрощает схему управления, так как переключение осуществляется по строчным синхроимпульсам (ССИ). По кадровым синхроимпульсам (КСИ) происходит опознание начала кадра.

Генератор формирует сигналы оцифровки для АЦП, сигналы записи-чтения для блока памяти, импульсы переключения для устройства адресации.

Схема адресации и управления опознает начало кадра, начало строки и, соответственно, управляет работой генератора и буфера.

Устройство работает в двух режимах: режим записи сигнала в буфер и режим передачи содержимого в компьютер.

При помощи компьютера можно управлять такими параметрами изображения, как контрастность и яркость, изменяя опорное напряжение на АЦП и постоянную составляющую сигнала соответственно.

Полученная информация о видеосигнале хранится и обрабатывается в виде массива с размерностью, зависящей от выбранного формата ввода (количества точек по вертикали и по горизонтали), элементами которого являются целые числа от 0 до 255. Меньшие значения соответствуют темным, а большие- светлым областям изображения.

Автоматизированная измерительная система имеет развитое программное обеспечение (ПО). Пользователю предлагается обширное меню, включающее проведение целого ряда операций над структурой оптического изображения и вычислений его характеристик. Все управление работой УВИ и ходом операций осуществляется с клавиатуры ПК или манипулятором "мышь". Все предлагаемые процедуры обработки изображений можно разбить на две группы. К первой относятся операции изменения яркости и контрастности изображений, а также проведение бинаризации и просмотр отдельных структурных фрагментов. Чаще всего указанные операции осуществляются на предварительной стадии обработки оптических полей. Ко второй группе относятся операции, обеспечивающие определение характеристик оптических полей. С их помощью определяются следующие характеристики.

1. Статистика точек введенного изображения по их интенсивности.

2. Распределение интенсивности по сечению пучка, профили пучка, поперечные размеры пучка.

3. Среднее значение интенсивности <I> по заранее заданной площади на приемной апертуре оптической системы.

4. Относительная дисперсия флуктуаций интенсивности

$$\sigma^2 = \left\langle \frac{(I - \langle I \rangle)^2}{\langle I \rangle^2} \right\rangle, \qquad \textbf{(4.6.1)}$$

где $I = I(x,y)$ - распределение интенсивности в сечении светового пучка в декартовых координатах.

5. Координаты центра тяжести пучка:

$$x_c = \frac{\int_{y_1}^{y_2}\int_{x_1}^{x_2} I(x,y)\,x\,dx\,dy}{\int_{y_1}^{y_2}\int_{x_1}^{x_2} I(x,y)\,dx\,dy}, \quad y_c = \frac{\int_{y_1}^{y_2}\int_{x_1}^{x_2} I(x,y)\,y\,dx\,dy}{\int_{y_1}^{y_2}\int_{x_1}^{x_2} I(x,y)\,dx\,dy}, \qquad \textbf{(4.6.2)}$$

где x_1, y_1, x_2, y_2 - координаты границы рабочего поля обработки.

6. Двумерный коэффициент автокорреляции интенсивности:

$$R(\xi,\eta) = \frac{\left\langle \tilde{I}(x+\xi, y+\eta)\cdot \tilde{I}(x,y) \right\rangle_{x,y}}{\sigma_I^2}, \qquad \textbf{(4.5.3)}$$

где $\tilde{I} = I - \langle I \rangle$, x,y; ξ, η - координаты в поперечном сечении;
и тот же коэффициент в одномерном выражении:

$$R(r) = \left\langle R(\xi,\eta)\right\rangle\Big|_{\xi^2+\eta^2=r^2}, \qquad \textbf{(4.6.4)}$$

Физический смысл функции *R(r)* - степень зависимости значения в двух точках, разделенных расстоянием *r*. Значение функции нормировано, то есть, ее максимум равен единице. Максимум этот достигается в точке r=0 (по определению), а также на таких значениях *r*, на которых поле не изменяется. Когда значение *R* убывает до величины, близкой к нулю, это означает, что на таких расстояниях значения поля не коррелируют, то есть, успевают измениться достаточно сильно.

7. Радиус зоны корреляции интенсивности $r_{кор}$ - значение аргумента $r_{кор}$, при котором $r_{кор}$ 1/2.

8. Пространственные Фурье-спектры распределения интенсивности и характерные пространственные частоты по горизонтали $f_Г$ и по вертикали $f_В$. Процедура расчета пространственных Фурье-спектров распределения интенсивности основывается на типовом дискретном преобразовании Фурье. Спектры могут характеризовать распределение интенсивности в любом заранее заданном скане сечения лазерного пучка, кроме того, может быть рассчитан усредненный по системе сканов фурье-спектр. По виду рассчитанных спектров можно определить характерные пространственные частоты по горизонтали $f_Г$ и по вертикали $f_В$.

9. Количество и площадь выбросов интенсивности - таких участков обрабатываемого светового поля, интенсивность которых превосходит заранее задаваемый уровень.

9. Фрактальную размерность структуры светового пучка.

Приведенный перечень измеряемых характеристик свидетельствует о том, что автоматизированная система измерений позволяет осуществлять многопараметрический анализ структуры световых полей со сложной пространственной структурой. Она была с успехом использована для исследования возмущений амплитудно-фазового профиля лазерных пучков при их распространении в турбулентной атмосфере.

ЛИТЕРАТУРА К ГЛАВЕ 4

1. Сисакян И.Н., Сойфер В.А. Компьютерная оптика. Достижения и проблемы //сб. "Компьютерная оптика" под ред. акад. Велихова Е.П. и акад. Прохорова А.М., 1987, в.1, с.5-19.

2. Сойфер В.А. Компьютерная оптика //Соросовский образовательный журнал, 1998

3. Франсон М. Голография.- М.: Мир, 1972, 248 с.

4. Горохов Ю.Г., Неплюев Л.Н. Голография в приборах и устройствах.- М.: Энергия,1974, 80 с.

5. Федоров Б.Ф., Цибулькин Л.М. Голография.- М.: Радио и связь, 1989, 140 с.

6. Кузнецова Т.И. О фазовой проблеме в оптике //УФН, 1988, т.154, в. 4, с. 677-690.

7. Воронцов М.А., Шмальгаузен В.И. Принципы адаптивной оптики.- М.: Наука, 1985, 336 с.

8. Воронцов М.А., Корябин А.В., Шмальгаузен В.И. Управляемые оптические системы. - М.: Наука, 1988, 270 с.

9. Гроссо Р., Еллин М. Мембранное зеркало как элемент адаптивной оптической системы //Сб. статей "Адаптивная оптика" под ред. Э.А. Витриченко - М.: Мир, 1980, с. 428-447.

10. Ярославский Л.П. Цифровая обработка полей в оптических системах. Цифровая оптика. //сб. "Новые физические принципы оптической обработки информации" под ред. С.А. Ахманова и М.А. Воронцова, - М.: Наука. Гл. ред. физ.-мат. лит., 1990, 400 с.

11. Мирошников М.М., Нестерук В.Ф. Развитие методологии иконики и ее структурной схемы //Труды Государственного оптического института им. С.И. Вавилова, 1982, т. 57, в. 185, с. 7- 13.

12. Сойфер В.А. Компьютерная обработка изображений. Часть 1. Математические модели //Соросовский образовательный журнал, 1996, №2, с. 118-124.

13. Сойфер В.А. Компьютерная обработка изображений. Часть 2. Методы и алгоритмы //Соросовский образовательный журнал, 1996, №3, с. 110-121.

Приложение 1

Алгоритм быстрого преобразования Фурье

Непосредственное вычисление коэффициентов Фурье по формуле

$$c_{\mathrm{pn}} = \sum_{k=0}^{M-1} \left(E_{\mathrm{mn}}\right)_{\Phi} \exp\left(-i2\pi p \frac{m}{M}\right) \qquad \textbf{(П.1.1)}$$

требует M операций умножения и столько же операций сложения комплексных чисел $\left(E_{\mathrm{mn}}\right)_{\Phi}$ и $\exp\left(-i2\pi p \frac{m}{M}\right)$ для вычисления одного коэффициента c_{pn}. Для вычисления M коэффициентов Фурье строки требуется M^2 операций. При квадратном изображении оно содержит M строк, следовательно для вычисления всей матрицы $\|c_{\mathrm{pn}}\|$ требуется M^3 операций. Столько же операций необходимо для преобразования по столбцам, т.е. для получения комплексного спектра $\|c_{\mathrm{pn}}\|$ требуется $2M^3$ операций умножения комплексных чисел и $2M^3$- сложения. Учитывая, что каждое комплексное умножение содержит четыре умножения и два сложения действительных чисел, а каждое комплексное сложение - два сложения действительных чисел, а также принимая, что умножение эквивалентно по времени выполнения полутора операциям сложения, получим, что пара комплексных операций умножения и сложения эквивалентна десяти приведенным операциям сложения действительных чисел.

Таким образом, для вычисления комплексного спектра изображения требуется $20M^3$ приведенных операций сложения, для реализации которых на современных компьютерах потребуется несколько часов. Положение изменилось после того, как был предложен алгоритм быстрого преобразования Фурье (БПФ). Этот алгоритм позволил значительно сократить (в сотни раз) количество вычислительных операций. Ниже излагается сущность БПФ.

Пусть функция E_{m} задана в M равноотстоящих точках числовой прямой, причем $M=2^{\mu}$, $\mu=1,2,\ldots$ Дискретное преобразование Фурье этой функции имеет вид:

$$c_{\mathrm{p}} = \sum_{k=0}^{M-1} E_{\mathrm{m}} \exp\left(-i2\pi p \frac{m}{M}\right), \quad p = 0,1,\ldots,\frac{M}{2}-1. \qquad \textbf{(П.1.2)}$$

Разобьем последовательность, состоящую из M значений функции E_{m}, на две последовательности F_{m} и G_{m}, каждая из которых состоит из $M/2$ значений. Разбиение произведем так, что в первую последовательность войдут значения E_{m}, заданные в точках с четными номерами, а во вторую- с нечетными:

$$F_{\mathrm{k}} = E_{2\mathrm{k}}, \quad G_{\mathrm{k}} = E_{2\mathrm{k}+1}, \quad p = 0,1,\ldots,\frac{M}{2}-1. \qquad \textbf{(П.1.3)}$$

Рассматривая F_{k} и G_{k} как новые функции, запишем для них дискретные преобразования Фурье:

$$\left(c_{\mathrm{p}}\right)_{\mathrm{ev}} = \sum_{k=0}^{\frac{M}{2}-1} F_{\mathrm{k}}\left(-i2\pi p \frac{k}{M/2}\right), \left(c_{\mathrm{p}}\right)_{\mathrm{od}} = \sum_{k=0}^{\frac{M}{2}-1} G_{\mathrm{k}}\left(-i2\pi p \frac{k}{M/2}\right)$$

$$p = 0,1,\ldots,\frac{M}{2}-1. \qquad \textbf{(П.1.4)}$$

Здесь через $\left(c_{\mathrm{p}}\right)_{\mathrm{ev}}$ и $\left(c_{\mathrm{p}}\right)_{\mathrm{od}}$ обозначены коэффициенты Фурье последовательностей, составленных из четно и нечетно пронумерованных значений.

Сумму в (П.1.2) разделим на две части, в одной из которых суммируются E_{m}, заданные на четных точках, во второй- на нечетных:

$$c_{\mathrm{p}} = \sum_{k=0}^{\frac{M}{2}-1} E_{2\mathrm{k}} \exp\left(-i2\pi p \frac{2k}{M}\right) + \sum_{k=0}^{\frac{M}{2}-1} F_{2\mathrm{k}+1} \exp\left(-i2\pi p \frac{2k+1}{M}\right) =$$

$$= \sum_{k=0}^{\frac{M}{2}-1} F_{\mathrm{k}} \exp\left(-i2\pi p \frac{k}{M/2}\right) + \exp\left(-i2\pi p \frac{p}{M}\right) \sum_{k=0}^{\frac{M}{2}-1} G_{\mathrm{k}} \exp\left(-i2\pi p \frac{k}{M/2}\right),$$

$$p = 0,1,\ldots,\frac{M}{2}-1. \qquad \textbf{(П.1.5)}$$

Как видно из (П.1.4), суммы в правой части (П.1.5) для k от нуля до (M/2)-1 равны $\left(c_{\mathrm{p}}\right)_{\mathrm{ev}}$ и $\left(c_{\mathrm{p}}\right)_{\mathrm{od}}$. Следовательно, для половины всех значений справедливо равенство

$$c_{\mathrm{p}} = \left(c_{\mathrm{p}}\right)_{\mathrm{ev}} + \exp\left(-i2\pi\frac{p}{M}\right)\cdot\left(c_{\mathrm{p}}\right)_{\mathrm{od}}, p = 0,1,\ldots,\frac{M}{2}-1. \quad \textbf{(П.1.6)}$$

Учитывая, что $\left(c_{\mathrm{p}}\right)_{\mathrm{ev}}$ и $\left(c_{\mathrm{p}}\right)_{\mathrm{od}}$ являются величинами периодическими с периодом P=M/2, можно записать:

$$\left(c_{\mathrm{p}+\frac{\mathrm{M}}{2}}\right)_{\mathrm{ev}} = \left(c_{\mathrm{p}}\right)_{\mathrm{ev}}, \quad \left(c_{\mathrm{p}+\frac{\mathrm{M}}{2}}\right)_{\mathrm{od}} = \left(c_{\mathrm{p}}\right)_{\mathrm{od}}.$$

Поэтому для второй половины коэффициентов Фурье получим:

$$c_{\mathrm{p}+\frac{\mathrm{M}}{2}} = \left(c_{\mathrm{p}}\right)_{\mathrm{ev}} + \exp\left(-i2\pi\frac{p+(m/2)}{M}\right)\cdot\left(c_{\mathrm{p}}\right)_{\mathrm{od}}.$$

Но

$$\exp\left(-i2\pi\frac{p+(m/2)}{M}\right) = \exp\left(-i2\pi\frac{p}{M}\right)\cdot\exp(-i\pi) = -\exp\left(-i2\pi\frac{p}{M}\right),$$

и поэтому будем иметь:

$$c_{\mathrm{p}+\frac{\mathrm{M}}{2}} = \left(c_{\mathrm{p}}\right)_{\mathrm{ev}} - \exp\left(-i2\pi\frac{p}{M}\right)\cdot\left(c_{\mathrm{p}}\right)_{\mathrm{od}}, p = 0,1,\ldots,\frac{M}{2}-1. \textbf{(П.1.7)}$$

Итак окончательно для вычисления всех M коэффициентов Фурье последовательности, заданной на M точках, имеем два соотношения:

$$c_{\mathrm{p}} = \left(c_{\mathrm{p}}\right)_{\mathrm{ev}} + \exp\left(-i2\pi\frac{p}{M}\right)\cdot\left(c_{\mathrm{p}}\right)_{\mathrm{od}}, \quad \textbf{(П.1.8)}$$

$$c_{\mathrm{p}+\frac{\mathrm{M}}{2}} = \left(c_{\mathrm{p}}\right)_{\mathrm{ev}} - \exp\left(-i2\pi\frac{p}{M}\right)\cdot\left(c_{\mathrm{p}}\right)_{\mathrm{od}}, p = 0,1,\ldots,\frac{M}{2}-1. \textbf{(П.1.9)}$$

Эти соотношения позволяют получить М значений c_{p}, имея по M/2 коэффициентов Фурье $\left(c_{\mathrm{p}}\right)_{\mathrm{ev}}$ и $\left(c_{\mathrm{p}}\right)_{\mathrm{od}}$ подпоследовательностей, заданных на четных и нечетных точках. Коэффициенты $\left(c_{\mathrm{p}}\right)_{\mathrm{ev}}$ и $\left(c_{\mathrm{p}}\right)_{\mathrm{od}}$, в свою очередь, могут быть получены по формулам (П.1.8) и (П.1.9) в результате аналогичной процедуры разбиения F_{k} и G_{k} каждой на две части по M/4 значений. Процедура разбиения подпоследовательностей может

продолжаться до тех пор, пока в полученных группах останется по одному числу. Этими числами являются элементы исходной последовательности E_m.

Таким образом, числа исходной последовательности являются первичным материалом для подстановки в систему (П.1.8), (П.1.9) рекуррентных соотношений для вычисления c_p. Весь процесс вычисления c_p проходит за μ итераций. Так как на каждой итерации из двух подпоследовательностей получается одна с двое большим количеством членов, то k-ая итерация переводит $2^{\mu-(k-1)}$ подпоследовательностей с 2^{k-1} значениями каждая в $2^{\mu-k}$ подпоследовательностей с 2^k значениями каждая (k=1,2,…,μ). Обозначив коэффициенты Фурье, получаемые при k-ый итерации для i-ой подпоследовательности, через , из (П.1.8) и (П.1.9) получим:

$$\left(c_p\right)_i^{(k)} = \left(c_p\right)_{2i}^{(k-1)} + \exp\left(-i2\pi\frac{p}{2^k}\right)\cdot\left(c_p\right)_{2i+1}^{(k-1)}, \qquad \textbf{(П.1.10)}$$

$$\left(c_{p+2^{\cdots}}\right)_i^{(k)} = \left(c_p\right)_{2i}^{(k-1)} - \exp\left(-i2\pi\frac{p}{2^k}\right)\cdot\left(c_p\right)_{2i+1}^{(k-1)},$$

$$k = 1,2,\ldots,\mu, \quad i = 0,1,\ldots,2^{\mu-k}-1, p = 0,1,\ldots,2^{k-1}-1., p = 0,1,\ldots,2^{k-1}-1. \qquad \textbf{(П.1.11)}$$

Соотношения (П.1.10) и (П.1.11) являются окончательными рекуррентными уравнениями для вычисления всех 2μ значений c_p с помощью μ итераций, причем при k=1 величины $\left(c_p\right)^{(k-1)} = \left(c_p\right)^0$ есть значения исходной последовательности E_m.

Приложение 2. Элементы фазовой проблемы: оценка фазовых возмущений по значениям числа Штреля (обзор литературы)

Доклад М.С. Магановой на семинаре лаборатории когерентной оптики 31 марта 1998г.

1. О фазовой проблеме.

Многие десятилетия особое место в оптике занимает, так называемая, фазовая проблема, суть которой состоит в возможности точного определения фазовых характеристик светового пучка, по известным распределениям интенсивности.

Существуют различные подходы к данной проблеме. Один из них изложен в известной монографии М.Борна и Э.Вольфа [1]. Согласно этому подходу для слабых возмущений волнового фронта существует простая однозначная связь [1] между значением числа Штреля *i* и дисперсией фазовых флуктуаций ϕ_{RMS}. Число Штреля определяется как отношение максимального значения интенсивности деформированного фазовыми возмущениями профиля пучка к максимальному значению интенсивности пучка, не искаженного фазовыми возмущениями. На основе расчета дифракционного интеграла М.Борн и Э.Вольф показали, что [1] :

$$i(P) = \frac{I(P)}{I_0} \approx 1 - \left(\frac{2\pi}{\lambda}\right)^2 (\Delta\phi_{\mathrm{P}})^2,$$

где $\phi_{RMS} = (\Delta\phi_{\mathrm{P}})^2 \equiv \langle\phi^2\rangle - \langle\phi\rangle^2$-дисперсия фазовых флуктуаций в плоскости P, число Штреля *i* и определяется в дальней зоне по отношению к распределению в плоскости P.

Целью данного сообщения является анализ точности метода , предложенного М.Борном и Э.Вольфом, а так же рассмотрение возможности применимости этого метода для случая более глубоких возмущений фазы, когда дисперсия фазовых флуктуаций, не является малой величиной. Заметим, что рассматриваемый метод не исчерпывает глубины фазовой проблемы, хотя и позволяет определенной статистической

характеристике распределения интенсивности, к которой надо относить число Штреля, ставить в соответствие другую статистическую характеристику - дисперсию фазы. Естественно метод не позволяет определять распределение фазы в плоскости P и какие-либо особенности структуры волнового фронта.

2. Описание метода.

После работы М.Борна и Э.Вольфа были предприняты попытки обобщения метода для более глубоких значений возмущения волнового фронта путем использования экспоненциального вида зависимости числа Штреля от дисперсии фазы [3],[4]:

$$I/I_0 = e^{-\phi_{RMS}^2},$$

Эта зависимость использовалась при оценке аберраций в лазерно-активных средах, а так же при определении степени возмущения волнового фронта в атмосфере. Однако использованию этой зависимости не предшествовал анализ ее точности [2]. Такой анализ содержится в работе [2]. Остановимся подробнее на ее результатах.

3. Метод оценки точности.

Роберт Д. Квинелль в своей работе [2] осуществил проверку экспоненциальной зависимости на основе численного моделирования. Численное моделирование проводилось путем расчета дифракционного интеграла по круглой диафрагме, располагающейся в плоскости P. При расчете диафрагма была разбита на систему, состоящую из 41-ой круговой зоны (см. *рис. 1*). При переходе от одной зоны к другой задавался случайный разброс фазы, величина которого варьировалась в процессе расчета. Из распределения интенсивности в дальней зоне дифракции определялось число Штреля,

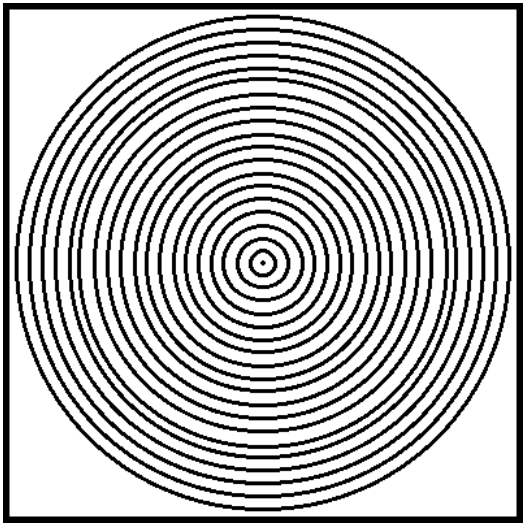


Рис. 1 *Структура зон в плоскости диафрагмы P*

которому ставилось в соответствие значение числа Штреля, получаемое из экспоненциальной зависимости (для одинакового уровня дисперсии фазы).

4. Результаты и выводы.

Метод оценки на основе экспоненциальной аппроксимации оказался менее точным, чем ожидалось. Как видно из *рис.2,* начиная с уровня 1 рад., отклонения от "истинного" значения числа Штреля становятся весьма значительными. Это подтверждается *рис.3*, где приведена зависимость величины ошибки от дисперсии фазы.

Такое заключение привело Роберта Д. Квинелли к поискам более точной аппроксимации. В результате методом "научного подбора" была предложена следующая зависимость:

$$I \sim \frac{1+\cos(T)}{2} - \frac{\sin^2(T)}{12},$$

где $T \equiv 1.75\phi_{RMS}$

Сопоставление этой формулы с точными расчетами привел к более удовлетворительным результатам (см. *рис.4*).

В качестве модификации рассматриваемого метода автор предложил для оценки фазовых возмущений использовать связь между мощностью, содержащейся в телесном угле с определенным раскрывом и дисперсией фазы. Графически эта связь показана на *рис.5.*

Т.о. существует возможность путем модификации расчетных соотношений распространить метод оценки дисперсии фазы по значениям числа Штреля на режим сильных фазовых флуктуаций.

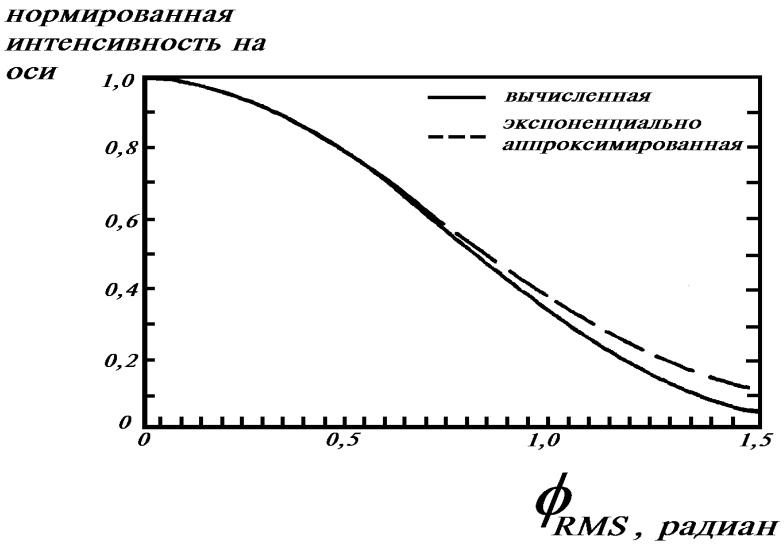


Рис.2 *Сопоставление интенсивностей Штреля, вычисленных точно и аппроксимированной экспоненциально*

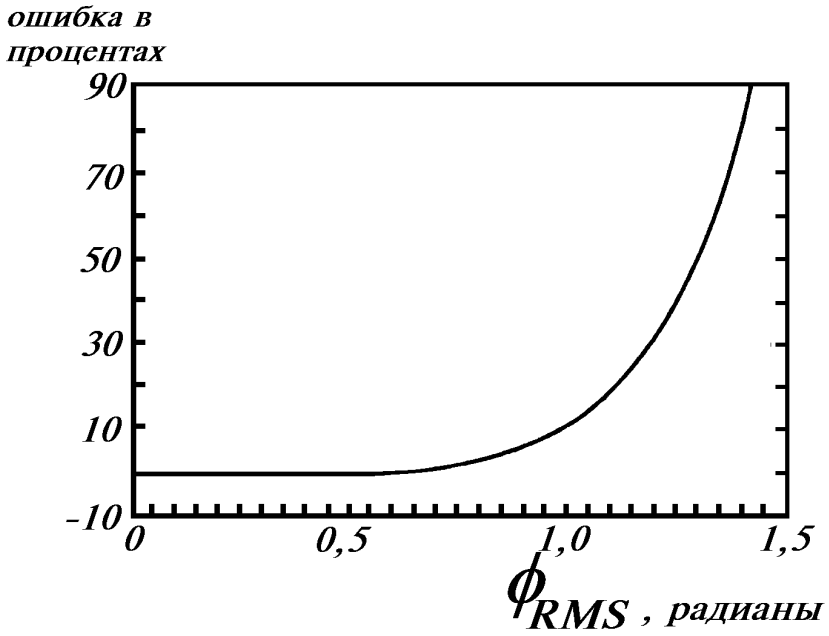


Рис.3 *Относительная ошибка экспоненциальной аппроксимации*

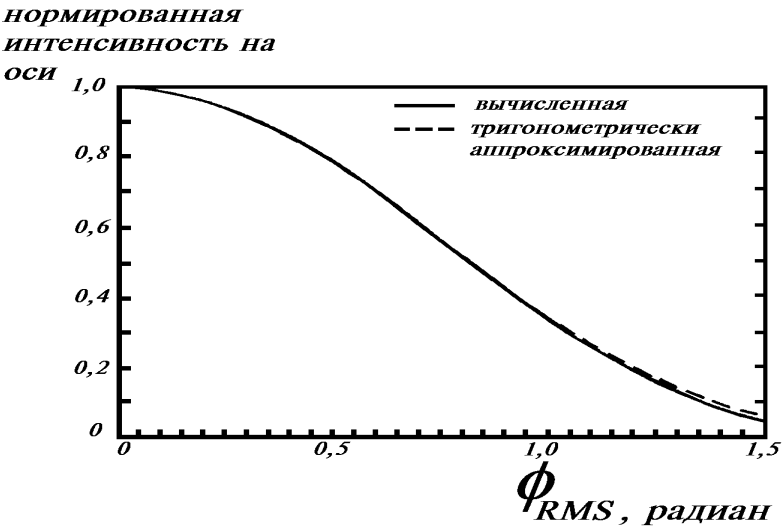


Рис.4 *Сопоставление интенсивностей, рассчитанной точно и аппроксимированной тригонометрические*

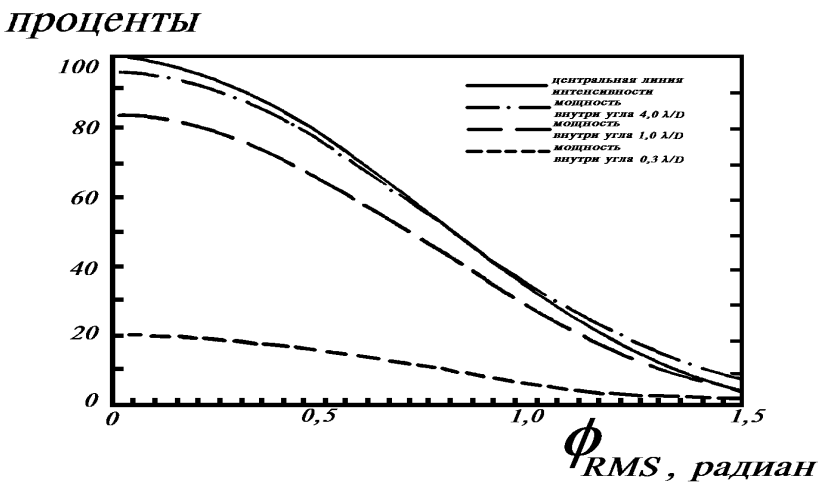


Рис.5 *Распределение интенсивностей по телесным углам*

Список литературы

1. М.Борн и Э.Вольф "*Основы оптики*", с.500-508

2. Robert D. Quinnell "*Limitation on the use of root-mean-square (rms) phase to describe beam quality characteristics*", SPIE, vol.293, Wavefront Distortions in Power Optics (1981), p.12-19

3. А.С. Башкин, В.В. Лобачев, И.А. Федоров "*Анализ пространственных масштабов оптических неоднородностей в активных средах мощных проточных лазерных усилителей*", "Квантовая электроника", **24**, 2 (1997), с.173-175

4. В.П. Лукин, Е.В. Носов, Б.В. Фортес "*Эффективный внешний масштаб атмосферной турбулентности", "Оптика атмосферы и океана*", **10**, 2, (1997), с.162-171