

РАДИООПТИКА И ГОЛОГРАФИЯ

Учебно-методический комплекс

2011

УДК 681.7(075)+535(075)
ББК 22.343я73
Р15

Авторы-составители: заведующий кафедрой инженерной физики УО «ВГУ им. П.М. Машерова», доктор технических наук, профессор **А.С. Ключников**; преподаватель кафедры инженерной физики УО «ВГУ им. П.М. Машерова» **Е.А. Краснобаев**; студент V курса физического факультета УО «ВГУ им. П.М. Машерова» **А.В. Боярин**

Рецензенты:

декан физического факультета УО «ВГУ им. П.М. Машерова», кандидат физико-математических наук, доцент *Ю.И. Бохан*; директор НУП «Информационно-обучающие технологии», кандидат технических наук *В.А. Ключников*

Учебно-методический комплекс по дисциплине специализации «Радиооптика и голография» предназначен для студентов специальности 1-31 04 01-02 «Физика (производственная деятельность)» и составлен в соответствии с учебной программой указанного курса для специализации 1-31 04-01-02-16 «Компьютерное моделирование физических процессов».

УДК 681.7(075)+535(075)
ББК 22.343я73

© УО «ВГУ им. П.М. Машерова», 2011

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	5
СТРУКТУРНОЕ СОДЕРЖАНИЕ КУРСА	8
ЛЕКЦИОННАЯ ЧАСТЬ	9
ЛЕКЦИЯ 1. Понятие о волнах. Примеры волновых движений	9
1.1. Волновое уравнение.....	12
1.2. Однородные и неоднородные волны.	14
1.3. Векторные волны. Поляризация.	15
1.4. Цилиндрические и сферические волны.	19
ЛЕКЦИЯ 2. Плоские электромагнитные волны в изотропных средах	21
ЛЕКЦИЯ 3. Особенности распространения волновых пучков.	
Дифракция.....	25
3.1. Угловой спектр плоских волн.	27
ЛЕКЦИЯ 4. Элементы теории излучения волн. Акустические излучатели.....	30
4.1. Элементарный электрический излучатель. Полуволновой диполь.....	33
ЛЕКЦИЯ 5. Виды антенн, применяемых в радиодиапазоне.	38
5.1. Распространение радиоволн вдоль земной поверхности. Функция ослабления.	42
5.2. Зоны видимости радиолокационных станций.....	44
ЛЕКЦИЯ 6. Двумерные преобразования Фурье и их свойства.....	46
6.1. Физическая интерпретация двумерных преобразований Фурье в теории волновых процессов.....	47
6.2. Пространственно-частотный анализ волновых процессов.....	49
6.3. Свободное пространство.	51
ЛЕКЦИЯ 7. Понятие теории выборки	54
7.1. Расчет систем, формирующие и преобразующие волновые фронты.....	55
7.2. Пространственно-частотный подход к дифракционным явлениям.....	55
ЛЕКЦИЯ 8. Оптические системы обработки информации	61
8.1. Теория фокусирующей линзы.....	61
ЛЕКЦИЯ 9. Расчет прохождения электромагнитных волн в простейших фокусирующих системах.....	65
9.1. Явления, связанные с конечностью размеров волнового фронта.....	69
ЛЕКЦИЯ 10. Физическая реализация интегральных преобразований в оптических системах.....	70
10.1. Геометрические и физические характеристики радиооптических систем.	72
10.2. Оптическая фильтрация. Фильтр Вандер Люгта.	77

ЛЕКЦИЯ 11. Распознавание образов в радиооптических системах.....	79
11.1. Амплитудный, фазовый и цифровой фильтры.	88
11.2. Оптические и оптико-цифровые процессы.	90
ЛЕКЦИЯ 12. Пространственная фильтрация и обработка волновых фронтов.....	92
12.1. Влияние явлений дискретизации на структуру волновых полей.....	92
12.2. Анализ и синтез волновых полей на ЭВМ	94
12.3. Поиск и распознавание образов.....	96
ЛЕКЦИЯ 13. Расчеты геометрических параметров голографических изображений.	100
13.1. Требования к когерентности и типы голограмм.....	104
ПРАКТИЧЕСКАЯ ЧАСТЬ	106
ЛАБОРАТОРНАЯ РАБОТА № 1. Основные характеристики объемных голограмм	106
ЛАБОРАТОРНАЯ РАБОТА № 2. Оптическая пространственная фильтрация.....	115
ЛАБОРАТОРНАЯ РАБОТА № 3. Запись оптических голограмм.....	124
ЛАБОРАТОРНАЯ РАБОТА № 4. Формирование акустических голограмм с оптическим восстановлением	128
КОНТРОЛЬНЫЕ ТЕСТЫ	131
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	136
СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ.....	136

ВВЕДЕНИЕ

Изучая акустику, оптику, радиофизику, динамическую теорию упругости, гидродинамику и другие разделы физики, мы замечаем некоторую закономерность, обусловленную тем, что со всех этих разделах рассматриваются волновые движения. Для описания вопросов, связанных с распространением всевозможных колебаний, вводится понятие «волна», охватывающее весьма обширный круг явлений. Так, можно сказать, что волны – это пространственно-временной процесс, описывающий распространение колебаний, т.е. периодически изменяющееся во времени чередование максимумов и минимумов любой физической величины (температуры, плотности вещества, плотности тока, напряженности электрического или магнитного поля и т.д.). *Волны* – это изменения состояния среды (возмущения), распространяющиеся в этой среде и несущие с собой энергию. Например, удар по концу стального стержня вызывает в этом месте сжатие, распространяющееся затем вдоль стержня со скоростью 5 км/с. Такая волна называется *упругой*. Упругие волны возникают в твердых, жидких и газообразных средах (сейсмические волны в земной коре, звуковые и ультразвуковые волны в жидкостях и газах и др.). При распространении упругих волн происходит перенос энергии упругой деформации. Теория волн в упругих средах имеет широкое применение в сейсмологии, акустике твердых тел и подводной акустике.

Акустика океана составляет важнейший раздел океанологии. В современной акустике наряду с традиционными областями (архитектурной, музыкальной акустикой, теорией дифракции звука и т.д.) развиваются новые направления: гидроакустика, биоакустика, ультразвуковая технология и многие другие. В связи с этим разработка единого подхода к изучению этих явлений имеет большое значение.

Примером волн являются морские волны, образующиеся на поверхности моря или океана под действием разного рода сил, выводящих частицы воды из состояния равновесия. Такими силами, под действием которых частицы воды совершают колебательные движения, являются приливообразующие силы Луны и Солнца, ветер, колебание атмосферного давления, подводные землетрясения и деформации дна. Изучение морских волн представляет большой практический интерес. Несколько лет назад возникло одно из самых грандиозных и смелых направлений в энергетике океана – волновая энергетика, изучающая возможность использования энергии волн для выработки электрической энергии. Существует ряд мест с высоким уровнем прилива, где планируется установка мощных приливных электростанций (ПЭС). В настоящее время существует несколько действующих опытных ПЭС. Теория морских волн является составной частью более общей теории волн в жидкостях. Волны в жидкостях изучаются в различных

областях науки и техники, таких, как акустика, метеорология, астрофизика, теория корабля, биология и т.д. Теория волн в жидкости охватывает самые разнообразные явления. К таким явлениям можно отнести цунами – гигантские волны, возникающие при подводных землетрясениях и извержениях подводных вулканов, ветровые и взрывные волны, волны в ионосфере, приливно-отливные движения и т.д.

В XIX в. ученые, сопоставляя известные из опыта свойства света со свойствами электромагнитных волн, следующими из теории Максвелла, пришли к выводу, что свет – это электромагнитные волны, т.е. в световой волне распространяются электрические и магнитные поля. К электромагнитным волнам относятся также радиоволны всех диапазонов, рентгеновские лучи, гамма-излучение и др. В настоящее время использование сложнейших электромагнитных процессов получило самое широкое распространение. Трудно перечислить те области науки и техники, в которых произошла революция благодаря открытию электромагнитных волн.

В атмосфере, кроме упругих звуковых волн, существуют волны нескольких типов, различные по происхождению и характеру. К ним относятся волны, образующиеся на границе двух воздушных слоев, движущихся с различными скоростями и имеющих различные плотности и температуры. При этом в «гребнях» волн происходит охлаждение воздуха и в результате конденсации водяных паров образуются облака. В «долинах» волн воздух нагревается и удаляется от состояния насыщения. При этом небо между «гребнями» остается чистым, в результате чего образуются гряды волнистых облаков. Аналогичный процесс происходит при движении воздуха около гор. В этом случае на вершинах также образуются облака. Колебательные движения продолжаются довольно долго после того, как масса воздуха миновала горное препятствие. Волны такого типа широко распространены и оказывают влияние на летательные аппараты.

И, наконец, существуют *волны де Бройля*, или *волны материи*, связанные с любой движущейся микрочастицей и отражающие ее квантовую природу. В 1924 г. французский физик Луи де Бройль выдвинул гипотезу о том, что всем без исключения материальным частицам (электронам, протонам, атомам и т.д.), помимо корпускулярных свойств, присущи волновые свойства.

Большой интерес представляют и паводковые волны в реках, ледниках, волны в потоке транспорта, звуковые удары и т.п. Явления возникновения звукового удара и пробки в потоке транспорта можно рассматривать в чисто описательном плане. С другой стороны, они весьма интересны для изучения и могут быть рассмотрены как относящиеся к волновым.

Несмотря на разнообразие природы волн их волновая картина является универсальной в том плане, что все они подчиняются практически одним и тем же законам.

Звуковая волна – это механическое явление, однако «поведение» ее существенно отличается от движения материальных тел. Обычные в законах механики понятия траектории, скорости, ускорения лишь частично применимы к волне. Для частиц же среды, в которой распространяется волна, эти понятия вполне корректны. Волне, выходящей из некоторого источника, нельзя приписать какую-либо траекторию, так как она распространяется равномерно во все стороны. Изучение волны в целом как самостоятельного объекта оказывается намного проще изучения поведения каждой частицы в среде, так как удастся найти простые законы распространения волны в различных средах, отражения от границ раздела сред, рассеяния от различных препятствий и т.д.

При изучении электромагнитных волн не учитывается механика частиц среды, следовательно, электрические и магнитные явления нельзя отождествлять с механическим поведением тел. Однако законы электромагнитных волн оказываются весьма близкими к законам упругих волн.

Таким образом, изучение общих закономерностей волн различной природы и составляет предмет курса теории волновых процессов. Основные свойства волн, независимо от их природы, проявляются в том, что при волновом движении осуществляется перенос энергии без переноса вещества. Так, например, после прохождения по поверхности жидкости волны, образовавшейся в результате бросания камня, частицы жидкости колеблются в вертикальной плоскости, оставаясь приблизительно в том же месте, что и до прохождения волны.

Волны различаются ориентацией возмущения по отношению к направлению распространения волны. Если возмущение направлено вдоль направления распространения волны, волна называется *продольной*. Так, например, звуковая волна в газе распространяется в том же направлении, в каком происходит смещение частиц газа. У *поперечной волны* возмущение перпендикулярно направлению распространения. Примером такой волны может служить распространяющаяся вдоль струны волна, смещение точек которой при этом происходит в направлении, перпендикулярном струне.

В жидкостях и газах упругие силы возникают только при сжатии и не возникают при сдвиге, поэтому упругие деформации в жидкостях и газах распространяются только в виде продольных волн. В твердых телах в связи с возникновением упругих сил при сжатии и сдвиге существуют как продольные, так и поперечные волны. Электромагнитная волна, распространяющаяся в безграничном изотропном пространстве, является поперечной волной; векторы напряженности электрического и магнитного полей такой волны колеблются в плоскостях, перпендикулярных к направлению распространения волны.

СТРУКТУРНОЕ СОДЕРЖАНИЕ КУРСА

№ п/п	Вид занятия	Тема занятия и изучаемые вопросы	Количество часов
1.	Лекция 1	Понятие о волнах. Примеры волновых движений	2
2.	Лекция 2	Плоские электромагнитные волны в изотропных средах	2
3.	Лекция 3	Особенности распространения волновых пучков. Дифракция	2
4.	Лекция 4	Элементы теории излучения волн. Акустические излучатели	2
5.	Лекция 5	Виды антенн, применяемых в радиодиапазоне	2
6.	Лекция 6	Двумерные преобразования Фурье и их свойства	2
7.	Лекция 7	Понятие теории выборки	2
8.	Лекция 8	Оптические системы обработки информации	2
9.	Лекция 9	Расчет прохождения электромагнитных волн в простейших фокусирующих системах	2
10.	Лекция 10	Физическая реализация интегральных преобразований в оптических системах	2
11.	Лекция 11	Распознавание образов в радиооптических системах	2
12.	Лекция 12	Пространственная фильтрация и обработка волновых фронтов	2
13.	Лекция 13	Расчеты геометрических параметров голографических изображений	2
14.	Лабораторная работа 1	Основные характеристики объемных голограмм	4
15.	Лабораторная работа 2	Оптическая пространственная фильтрация	4
16.	Лабораторная работа 3	Запись оптических голограмм	4
17.	Лабораторная работа 4	Формирование акустических голограмм с оптическим восстановлением	4
	Всего		52

ЛЕКЦИОННАЯ ЧАСТЬ

ЛЕКЦИЯ 1. Понятие о волнах. Примеры волновых движений

В общем случае волны могут иметь различные вид и зависимость от временной и пространственных координат. Описание таких волн общего вида является довольно сложной, а подчас и неразрешимой задачей. В связи с этим для качественного и количественного описания явлений, возникающих при распространении волны, рассмотрим частный случай одномерной волны. *Одномерные волны* – это волны, все характеристики которых зависят от времени и одной координаты. Примером таких волн могут служить волны в одномерной среде (в струне, стержне, жидкости, заполняющей узкую трубу, и т.п.), в двухмерных (на пластине) и трехмерных (плоские электромагнитные волны в изотропных неограниченных средах и т.п.) средах.

Пусть некоторая скалярная физическая величина s , характеризующая волну, зависит от времени t и одной координаты z декартовой системы координат:

$$s = f(z, t). \quad (1.1)$$

Эта формула свидетельствует о том, что величина s в любой данный момент времени принимает постоянные значения на системе параллельных плоскостей ($z = \text{const}$), причем значения s на различных плоскостях, вообще говоря, различны, следовательно, с течением времени значения s на каждой из этих плоскостей меняются. Если зависимость s от z представляет собой функцию с периодически изменяющимися максимумами и минимумами, расположение которых различно в различные моменты времени, то процесс, описываемый выражением (1.1), является *волновым*, а функция s называется *плоской волной*.

Рассмотрим частный случай плоской волны, когда переменные функции $f(z, t)$ входят через линейную комбинацию:

$$s = F(kz - \omega t), \quad (1.2)$$

где k, ω – некоторые произвольные постоянные.

Выясним некоторые свойства плоской волны (1.2). Рассмотрим графики функций $s(z, t)$ и $s(z, t)$, называемые соответственно *пространственным* и *временным профилем* волны s . В первом случае получим:

$$s(z, t_1) = F(kz - \omega t_1) = F[k(z - vt_1)], \quad (1.3)$$

где

$$v = \omega / k \quad (1.4)$$

В частности, при $t = 0$

$$s(z,0) = F(kz). \quad (1.5)$$

Из сравнения выражений (1.3) и (1.5) видно, что пространственный профиль волны в момент времени $t = t_1$ отличается от профиля волны в момент $t = 0$ только смещением вправо на расстояние v, t_1 . Во втором случае имеем

$$s(z_1, t) = F(kz_1 - \omega t) = F[-\omega(t - z_1/v)]. \quad (1.6)$$

В частности, при $z = 0$

$$s(0, t) = F(-\omega t). \quad (1.7)$$

Очевидно, что график функции (1.6) точно воспроизводит график функции (1.7) с опозданием на время z_1/v .

Таким образом, волна, описываемая выражением (1.2), распространяется вдоль оси z в сторону положительных значений со скоростью v . Аналогично можно убедиться, что выражение $s = F(-kz - \omega t)$ описывает плоскую волну, распространяющуюся вдоль оси z в направлении, противоположном направлению этой оси, со скоростью v , определяемой выражением (1.4). Такие волны называются *бегущими*.

В теории волн особо важно представление о *гармонической волне*, т.е. бесконечной синусоидальной волне, в которой все изменения состояния среды происходят по закону синуса или косинуса:

$$s = A \cos(kz - \omega t). \quad (1.8)$$

Величина A называется *амплитудой волны*. Это значение, которое s принимает при наибольших отклонениях от положения равновесия. Основными параметрами гармонической волны являются *длина волны* λ , – расстояние между двумя максимумами или минимумами возмущения – и *период* T – время, за которое частицы среды совершают одно полное колебание. Бесконечная волна, таким образом, обладает строгой периодичностью в пространстве и времени. Периодичность в пространстве можно обнаружить, например, на моментальной фотографии пространственного профиля упругой волны. Следя за движением определенной частицы среды, можно обнаружить периодичность волны во времени.

Связь между длиной волны и периодом определяется соотношением

$$T = \lambda / v, \quad (1.9)$$

где v – скорость распространения волны.

Вместо периода T часто пользуются *частотой колебаний* f , равной числу колебаний в единицу времени:

$$f = 1/T. \quad (1.10)$$

Из соотношений (9) и (10) получаем

$$\lambda f = v. \quad (1.11)$$

Для определения постоянной k в формуле (1.8) мгновенное значение волны. Например, при $t = 0$

$$S = A \cos \left(\frac{2\pi}{\lambda} z \right). \quad (1.12)$$

Пространственный период волны, т.е. расстояние λ , таков, что при любом z выполняется соотношение

$$s = (z + \lambda) = s(z). \quad (1.13)$$

Подставив выражение (1.2) в (1.3), получим $k\lambda = 2\pi$, т.е. $k = 2\pi/\lambda$. Величина k называется *волновым числом*.

Рассматривая волну для фиксированного значения $z = 0$ и используя вытекающее из временной периодичности волны соотношение $s(t + T) = s(t)$, находим $\omega T = 2\pi$, откуда определяем постоянную ω :

$$\omega = 2\pi/T = 2\pi f, \quad (1.14)$$

называемую *круговой частотой* и равную числу колебаний за 2π единиц времени.

Гармонические волны называются также *монохроматическими*, поскольку колебания происходят с вполне определенной частотой. Интерес к гармонической зависимости волны от времени обусловлен, во-первых, тем, что большинство источников создают волны, зависимость которых от времени близка к гармонической; во-вторых, почти любой немонхроматический процесс с произвольной зависимостью от времени можно разложить в интеграл Фурье, т.е. представить как наложение гармонических колебаний; в-третьих, существует удобный математический аппарат, позволяющий из уравнений, описывающих волновые процессы, исключить время и тем самым существенно их упростить.

Аргумент косинуса плоской гармонической волны называется *фазой колебаний*. Понятие фазы введено для определения положения точки на волне. Фаза колебаний зависит от выбора начала отсчета времени. Если отсчет ведется от момента, когда s проходит через максимум, мы можем считать, что фаза равняется нулю. Если отсчет вести с момента, когда s проходит, возрастая, через нуль, то фаза равна $\pi/2$. Одно и то же значение фазы может соответствовать различным началам отсчета времени, отстоящим друг от друга на целое число периодов. При фиксированном отсчете времени фаза определена с точностью до целого кратного числа 2π .

Из опыта нам известно, что бегущие по воде волны распространяются с постоянной скоростью. В случае гармонических волн эта скорость называется *фазовой*. Это скорость распространения заданной фазы в пространстве, которую легко определить. Зафиксировав момент времени $t = t_1$ и точку $z = z_1$, отметим, на какое расстояние dz переместится точка с фазой $kz_1 - \omega t_1$ через время dt :

$$kz_1 - \omega t_1 = k(z_1 + dz) - \omega(t_1 + dt)$$

или

$$dz = \frac{\omega}{k} dt.$$

Отсюда получим выражение для фазовой скорости:

$$v_{\phi} = dz / dt = \omega / k = v.$$

Полученный результат показывает, что входившая в исходные уравнения для плоской волны скорость распространения v является фазовой.

Следует отметить, что перемещение фазы не связано с перемещением каких-либо материальных точек волны. Фаза – понятие математическое, и поэтому фазовая скорость может принимать любые значения без нарушения основных законов физики. Действительно, картина волны, движущейся вдоль оси z , не изменится, если мы будем считать, что возмущение перемещается вдоль оси z с фазовой скоростью $v_{\phi 1}$ или вдоль некоторой другой оси z' со скоростью $v_{\phi 2}$.

1.1. Волновое уравнение

В теории волн большое значение имеет уравнение в частных производных второго порядка

$$\frac{\partial^2 F}{\partial z^2} - \left(\frac{k}{\omega}\right)^2 \frac{\partial^2 F}{\partial t^2} = 0, \quad (1.15)$$

называемое *волновым уравнением*.

Легко убедиться, что функция $F(kz - \omega t)$ и $F(-kz - \omega t)$, описывающие бегущие плоские волны, удовлетворяют уравнению (1.15). Действительно, продифференцировав функцию $F(kz - \omega t)$ дважды по z и t , получим:

$$\partial^2 F / \partial z^2 = k^2 F''(kz - \omega t), \quad (1.16)$$

$$\partial^2 F / \partial t^2 = \omega^2 F''(kz - \omega t). \quad (1.17)$$

Здесь штрихи обозначают дифференцирование по аргументу $kz - \omega t$.

Подставив в уравнение (1.16) значение $F''(kz - \omega t)$, определенное из выражения (1.17), получим волновое уравнение.

Можно доказать, что уравнение (1.15) не имеет решений, отличных от тех, которые могут быть представлены функциями $F(\pm kz - \omega t)$ или суперпозицией таких функций.

Наиболее общим аналитическим выражением для плоских волн, распространяющихся вдоль какой-то произвольной оси z' , является функция

$$F[\pm k(\alpha x + \beta y + \gamma z) - \omega t], \quad (1.18)$$

где α, β, γ – направляющие косинусы единичного вектора вдоль оси z' , удовлетворяющие условию

$$\alpha^2 + \beta^2 + \gamma^2 = 1. \quad (1.19)$$

Легко убедиться, что функция (1.18) является решением волнового уравнения

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

Таким образом, если какая-нибудь физическая величина s удовлетворяет уравнению вида

$$\nabla^2 s - \left(\frac{k}{\omega}\right)^2 \frac{\partial^2 s}{\partial t^2} = 0, \quad (1.20)$$

то можно с уверенностью утверждать, что процесс изменения величины s носит характер плоской волны, распространяющейся вдоль определенного направления в ту или другую сторону со скоростью

$$v = \omega/k.$$

Для гармонических волн волновое уравнение можно существенно упростить, применив метод комплексных амплитуд. Уравнение гармонической волны можно представить следующим образом

$$s = A \cos(kz - \omega t) = \operatorname{Re}[(Ae^{jkz})e^{-j\omega t}] = \operatorname{Re}(Ae^{-j\omega t}). \quad (1.21)$$

Величина $A = Ae^{jkz}$ называется *комплексной амплитудой* гармонической волны (1.8). Таким образом, если известна комплексная амплитуда некоторой величины s , то мгновенное значение ее представляет собой действительную часть произведения комплексной амплитуды на $e^{-j\omega t}$. В связи с простотой перехода от комплексных амплитуд к мгновенным значениям волновое уравнение (1.20) можно записать не для величины s , а для комплексной амплитуды. Смысл такой записи состоит в упрощении волнового уравнения, так как в него не входят производные по времени. Для этого подставим выражение (1.21) в уравнение (1.20) и получим выражение, взяв производную по t , сократив на $e^{-j\omega t}$ и отбросив знак Re , будем иметь

$$\nabla^2 A + k^2 A = 0. \quad (1.22)$$

Мы получили уравнение, не зависящее от времени. Оно называется *однородным, волновым уравнением Гельмгольца* относительно комплексной амплитуды гармонической волны.

Уравнение (1.22) описывает процесс распространения плоской гармонической волны в однородной изотропной среде. Процесс распространения волн в присутствии источников колебаний или каких-либо внешних сил описывается неоднородным волновым уравнением Гельмгольца:

$$\nabla^2 A + k^2 A = f(z), \quad (1.23)$$

где $f(z)$ – функция, характеризующая внешние воздействия.

При описании волн в реальных средах с учетом поглощений или в диспергирующих средах волновое уравнение дополняется неко-

торым линейным членом, зависящим от конкретных условий распространения волны в данной среде. В неоднородных средах, свойства которых зависят от координат, волновые процессы описываются волновым уравнением вида (1.23), в котором величина k является функцией координат. Всевозможные нелинейные эффекты, возникающие при возбуждении в среде сильных полей, описываются неоднородным волновым уравнением, правая часть которого представляет собой сумму некоторых линейных операторов. Итак, любая задача по определению величин, характеризующих волновой процесс в различных средах, сводится к получению того или иного волнового уравнения, его решения и анализа полученных результатов.

1.2. Однородные и неоднородные волны

Одной из важнейших характеристик волн является вид поверхностей равных фаз, т.е. поверхностей, у которых в данный момент времени фазы в любой точке одинаковы. Эти поверхности называют *фронтом волны*.

Форма фазовой поверхности зависит от условий возникновения и распространения волн. У плоских волн поверхности равных фаз представляют собой плоскости. Гармоническая волна, описываемая уравнением (1.8), где A , k , ω – постоянные, является бегущей волной: во всех точках плоскости $z = \text{const}$ значения амплитуд и фаз одинаковы. Плоские волны, у которых плоскости равных амплитуд и фаз совпадают, называются *однородными плоскими волнами*.

Можно легко показать, что при некоторых определенных условиях в плоской гармонической волне плоскости равных амплитуд и фаз могут не совпадать. Такие волны называются *неоднородными плоскими волнами*. Они играют большую роль при полном отражении. Возникают неоднородные плоские волны при распространении в сложных оптических системах и неоднородных средах. В настоящее время вопросы, связанные с поведением неоднородных плоских волн, весьма актуальны в связи с развитием интегральной и планарной оптики. Не задаваясь пока вопросами возникновения и поведения таких волн, выясним принципиальную возможность их существования. Для этого рассмотрим одно из частных решений волнового уравнения (1.22), имеющее вид

$$A = C \exp[jk(\alpha x + \beta y + \gamma z)], \quad (1.24)$$

где C – постоянная; α , β , γ – коэффициенты, удовлетворяющие соотношению (1.19). Если α , β и γ – вещественные числа, удовлетворяющие соотношению (1.19), то функция (1.24) является однородной плоской волной. Действительно, в этом случае при введении новой координаты z' с помощью соотношения $z' = \alpha x + \beta y + \gamma z$ выражение (1.24) принимает вид $A = Ce^{jkz'}$, и, следовательно, волна

$$s = \operatorname{Re}(Ae^{-j\omega t}) = C \cos(kz' - \omega t) \quad (1.25)$$

является однородной плоской волной, распространяющейся вдоль оси z' .

Если среди чисел α , β и γ , удовлетворяющих соотношению (1.19), есть хотя бы одно комплексное число, то волну (1.24) уже нельзя привести к однородной плоской волне (1.25). Положим

$$\alpha = \alpha' + j\alpha'', \beta = \beta' + j\beta'', \gamma = \gamma' + j\gamma'',$$

где α' , β' , γ' – действительные, а α'' , β'' , γ'' – мнимые части. Тогда выражение (1.24) примет вид

$$A = C \exp[jk(\alpha'x + \beta'y + \gamma'z)] \exp[-k(\alpha''x + \beta''y + \gamma''z)].$$

Очевидно, что у такой волны поверхности равных фаз определяются выражением

$$\alpha'x + \beta'y + \gamma'z = \text{const}, \quad (1.26)$$

а поверхности равных амплитуд – выражением

$$\alpha''x + \beta''y + \gamma''z = \text{const}. \quad (1.27)$$

Плоскости (1.26) и (1.27) взаимно перпендикулярны, так как из условия (19) следует, что $\alpha'\alpha'' + \beta'\beta'' + \gamma'\gamma'' = 0$.

1.3. Векторные волны. Поляризация

Если предположить, что некоторая векторная величина s (напряженность электрического или магнитного поля, скорость и т.д.) изменяется в пространстве и во времени по закону

$$s = F(kz - \omega t), \quad (1.28)$$

то очевидно, что функция (1.28) описывает плоскую векторную волну. В зависимости от ориентации вектора s относительно направления распространения волны z векторные волны могут быть продольными и поперечными. Векторные волны, так же, как и скалярные, описываются волновым уравнением (1.22). Описание продольных векторных волн ничем не отличается от описания скалярных волн, поскольку вместо вектора s можно рассматривать его, проекцию на ось z . Таким образом, все направления, перпендикулярные к направлению распространения волны, физически равноправны. Поперечные векторные волны коренным образом отличаются от скалярных и продольных векторных волн.

Рассмотрим вопрос об ориентации векторов поля при заданном направлении распространения, или, как принято говорить, о поляризации волны. Плоскость, проведенная через вектор s и вектор, совпадающий с направлением распространения, называется *плоскостью поляризации*. Если, например, вектор s коллинеарен оси x , то плоскость поляризации есть плоскость xOz . При этом говорят, что волна поляризована в плоскости xOz . Если вектор s , изменяясь по абсолютной ве-

личине, не изменяет своего направления в пространстве, то поляризация называется *линейной*.

Запишем векторное волновое уравнение Гельмгольца для комплексной амплитуды $\dot{s}$ вектора s :

$$\nabla^2 \dot{s} + k^2 \dot{s} = 0. \quad (1.29)$$

Для одномерной волны, распространяющейся вдоль оси z , уравнение (1.29) принимает вид

$$d^2 \dot{s} / dz^2 + k^2 \dot{s} = 0. \quad (1.30)$$

Для случая линейной поляризации, когда вектор s не изменяет своего направления в пространстве (т.е. если в выражении комплексной амплитуды $\dot{s} = a \dot{s}$ единичный вектор a , указывающий направление, постоянен), уравнение (1.30) является скалярным волновым уравнением Гельмгольца, решение которого запишется в виде

$$s = a[s_1 \cos(kz - \omega t + \varphi_1) + s_2 \cos(-kz - \omega t + \varphi_2)]. \quad (1.31)$$

Формула (1.31) описывает не что иное, как наложение двух плоских однородных гармонических волн с амплитудами s_1 и s_2 и начальными фазами φ_1 и φ_2 , распространяющихся в противоположных направлениях z и $-z$. Следовательно, любую плоскую векторную волну, распространяющуюся вдоль оси z , можно записать в виде

$$s = a s_1 e^{jkz}.$$

Рассмотрим две волны, распространяющиеся в одном направлении, но поляризованные в двух взаимно перпендикулярных плоскостях xOz и yOz . Пусть эти волны различаются начальными амплитудами и фазами. Их комплексные амплитуды запишутся соответственно в виде:

$$\dot{s}_1 = x_0 s_1 e^{jkz}; \quad \dot{s}_2 = y_0 s_2 e^{jkz}.$$

При сложении этих волн образуется поле с комплексной амплитудой

$$s = x_0 s_1 \cos(kz - \omega t + \varphi) + y_0 s_2 \cos(kz - \omega t + \varphi). \quad (1.32)$$

Для случая, когда начальные фазы волн совпадают ($\varphi_x = \varphi_y = \varphi$), выражение (1.32) принимает вид

$$s = (x_0 s_1 + y_0 s_2) \cos(kz - \omega t + \varphi) = x'_0 s_m \cos(kz - \omega t + \varphi), \quad (1.33)$$

где $x'_0 = x_0 \cos \vartheta + y_0 \sin \vartheta$; $\tan \vartheta = s_2 / s_1$; $s_m = \sqrt{s_1^2 + s_2^2}$ — *полная амплитуда волны*.

Из выражения (1.33) видно, что суммарный вектор s не изменяет своего направления в пространстве.

Следовательно, суперпозиции двух волн, линейно поляризованных во взаимно перпендикулярных плоскостях и совпадающих по фазе во времени, дает линейно поляризованную волну. Плоскость поляризации суммарной волны составляет угол ϑ с плоскостью xOz .

Складывая волны с различными амплитудами s_1 и s_2 , можно получать волны, поляризованные во всевозможных плоскостях, проходящих через ось z .

Рассмотрим суперпозицию поляризованных в двух взаимно перпендикулярных плоскостях волн с равными амплитудами и фазами, различающимися на $-\pi/2$, т.е. $s_1 = s_1 e^{i\varphi_1}$ и $s_2 = s_1 e^{i(\varphi_1 - \pi/2)}$. Тогда выражение для суммарной волны (1.32) можно записать в виде

$$\mathbf{s} = s_1 [\mathbf{x}_0 \cos(kz - \omega t + \varphi_1) + \mathbf{y}_0 \sin(kz - \omega t + \varphi_1)]. \quad (1.34)$$

Соотношение компонент s_y и s_x в этом случае не остается постоянным в пространстве и времени. Очевидно, что при фиксированном значении координаты z вектор $\mathbf{s}$ с течением времени будет вращаться с угловой скоростью ω вокруг направления распространения волны. Относительно наблюдателя, смотрящего в направлении распространения волны, вектор $\mathbf{s}$ будет вращаться против часовой стрелки, так как с течением времени угол линейно уменьшается.

При фиксированном моменте времени $t = \text{const}$ угол ϑ линейно возрастает по закону kz с увеличением координаты z , изменяясь на 2π на расстоянии λ . Таким образом, вектор $\mathbf{s}$ в момент $t = \text{const}$ равномерно поворачивается с увеличением z в направлении от x_0 к y_0 (по часовой стрелке относительно направления z_0) и делает на расстоянии λ один оборот. При этом концы вектора $\mathbf{s}$, относящиеся к различным точкам оси z , расположены на правовинтовой круговой спирали. Такая волна представляет собой *волну круговой поляризации с левым направлением вращения*.

Если рассматривать суперпозицию волн с равными амплитудами и фазами, отличающимися на $\pi/2$, то вместо выражения (1.34) будем иметь

$$\mathbf{s} = s_1 [\mathbf{x}_0 \cos(kz - \omega t + \varphi_1) - \mathbf{y}_0 \sin(kz - \omega t + \varphi_1)]. \quad (1.35)$$

У этой волны круговой поляризации в плоскости $z = \text{const}$ вектор $\mathbf{s}$ равномерно вращается в направлении от x_0 к y_0 , т.е. по часовой стрелке относительно наблюдателя, так как с течением времени угол ϑ возрастает, а в момент времени $t = -\text{const}$ концы вектора расположены на левовинтовой круговой спирали. Такая волна называется *волной круговой поляризации с правым направлением вращения*. Иногда целесообразно рассматривать линейно поляризованную волну как результат суперпозиции двух круговых волн.

Рассмотрим случай распространения линейно поляризованных волн с различными начальными амплитудами и фазами. Пусть вдоль оси z распространяются две плоские волны:

$$\mathbf{s}_1 = \mathbf{x}_0 s_1 \cos(kz - \omega t + \varphi_1); \quad \mathbf{s}_2 = \mathbf{y}_0 s_2 \cos(kz - \omega t + \varphi_2). \quad (1.36)$$

Из выражений (1.36) видно, что вектор поля $\mathbf{s} = \mathbf{s}_1 + \mathbf{s}_2$ суммарной волны имеет две проекции s_x и s_y , причем $z=0$

$$s_x = s_1 \cos(\varphi_1 - \omega t); s_y = s_2 \cos(\varphi_2 - \omega t). \quad (1.37)$$

Будем рассматривать величины s_x и s_y как координаты точки на плоскости xOy . Плотность потока мощности такой волны пропорциональна сумме $s_1^2 + s_2^2$.

Исключив из выражений (1.37) время с помощью соотношений $\varphi_1 - \omega t = \alpha$, $\varphi_2 - \varphi_1 = \varphi$, после элементарных преобразований получим

$$\frac{s_x^2}{s_1^2} + \frac{s_y^2}{s_2^2} - 2 \frac{s_x s_y}{s_1 s_2} \cos \varphi = \sin^2 \varphi.$$

Это равенство представляет собой уравнение эллипса, повернутого относительно координат x и y на некоторый угол. Таким образом, в общем случае годографом вектора поля плоской монохроматической волны является эллипс, называемый *поляризационным эллипсом*, внутри которого вектор s вращается с периодически изменяющейся скоростью, причем полный оборот происходит за период несущей частоты. Волна такого типа называется *волной эллиптической поляризации*.

Частным случаем волн эллиптической поляризации являются волны круговой и линейной поляризации. Поляризационный эллипс этих волн вырождается в окружность или прямую линию. Для количественной характеристики поляризации волны пользуются геометрическими параметрами эллипса (ориентацией и направлением вращения вектора s). Форму эллипса характеризуют величиной, называемой *коэффициентом эллиптичности*, величина которого определяется по формуле $|r| = b/a$, где b , a – соответственно малая и большая полуоси эллипса. Величина r считается положительной, если волна правополяризованная, т.е. конец вектора s обходит эллипс по часовой стрелке при наблюдении вдоль направления распространения, и отрицательной при левополяризованной волне. Поскольку $|r| \leq 1$, удобно ввести угол $\alpha = \text{arctgr } r$ ($-\pi/4 \leq \alpha \leq \pi/4$), называемый *углом эллиптичности*, абсолютная величина которого однозначно определяет форму эллипса, а знак указывает направление вращения вектора поля.

Ориентация поляризационного эллипса определяется углом β между большой осью эллипса и осью абсцисс выбранной прямоугольной системы координат. Очевидно, что значения угла β ограничиваются пределами $0 \leq \beta \leq \pi$.

Параметры поляризационного эллипса являются важными характеристиками эллиптически поляризованной волны, поскольку мощность принимаемого сигнала сильно зависит от параметров и ориентации относительно друг друга поляризационных эллипсов принимаемой волны и принимающего устройства.

1.4. Цилиндрические и сферические волны

Волновому уравнению могут удовлетворять волны различных типов. Широкое практическое применение находят цилиндрические и сферические волны. Рассмотрение большого количества радиофизических задач сводится к изучению поля, описываемого одной составляющей вектора s . Это соответствует случаю линейных токов или постоянного в пространстве направления вектора поляризации, случаю анизотропии, создаваемой магнитным полем, направленным по одной из координатных осей, и т.д. Такие явления хорошо описываются цилиндрическими волнами.

В теории распространения волн, как правило, надо учитывать конечную удаленность источника волн как от приемника, так и от всевозможных препятствий. Классической и простейшей задачей такого рода является задача о поле точечного излучателя, расположенного на конечном расстоянии от препятствия, иными словами – задача о распространении сферической волны.

Для описания цилиндрических волн удобно пользоваться цилиндрической системой координат, связанной с декартовыми координатами x, y, z формулами преобразования: $x = \rho \cos \varphi, y = \rho \sin \varphi, z = z$, где $0 \leq \rho \leq \infty, 0 \leq \varphi \leq 2\pi, -\infty \leq z \leq \infty$. Координатными поверхностями в этой системе являются плоскости $z = \text{const}$, перпендикулярные к оси z , плоскости $\varphi = \text{const}$, проходящие через ось z , образующую угол φ с плоскостью xOz , а также коаксиальные цилиндры, определяемые уравнением $x^2 + y^2 = \rho_0^2$ ($\rho_0 = \text{const}$), ось вращения которых совпадает с осью z . Волновое уравнение (1.15) для функции $s(x, y, t)$, не зависящей от z , в цилиндрической системе координат принимает вид

$$\frac{\partial^2 s}{\partial t^2} - \left(\frac{\omega}{k}\right)^2 \left[\frac{1}{\rho} \left(\frac{\partial}{\partial \rho} \rho \frac{\partial s}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 s}{\partial \varphi^2} \right] = 0. \quad (1.38)$$

Для частного случая, наиболее часто встречающегося в практике, когда функция s не зависит от азимутального угла φ уравнение (1.38) запишем в виде

$$\frac{\partial^2 s}{\partial t^2} - \left(\frac{\omega}{k}\right)^2 \left(\frac{\partial^2 s}{\partial \rho^2} + \frac{1}{\rho} \frac{\partial s}{\partial \rho} \right) = 0. \quad (1.39)$$

Произведя в уравнении (1.39) замену переменных $u = s\sqrt{\rho}$, получим волновое уравнение

$$\frac{\partial^2 u}{\partial t^2} - \left(\frac{\omega}{k}\right)^2 \frac{\partial^2 u}{\partial \rho^2} = 0,$$

одним из частных решений которого является

$$u(\rho, t) = F(t - k\rho/\omega).$$

Отсюда следует выражение для функции

$$s(\rho, t) = \frac{1}{\sqrt{\rho}} F(t - \rho/v), \quad (1.40)$$

где v – скорость распространения волны вдоль направления, перпендикулярного к оси z . В данном случае поверхность равной фазы удовлетворяет уравнению $t - \rho/v = \text{const} = t_0$ или $\rho = v(t - t_0)$ и представляет собой поверхность с круговым сечением, ось которой совпадает с осью z , т.е. с координатной поверхностью коаксиальных цилиндров. В отличие от плоской волны в цилиндрической волне амплитуда обратно пропорциональна корню квадратному из расстояния ρ .

Для описания сферических волн пользуются сферической системой координат, связанной с декартовой следующими соотношениями: $x = r \sin \theta \cos \varphi$, $y = r \sin \theta \sin \varphi$, $z = r \cos \theta$, где $0 \leq r \leq \infty$, $0 \leq \theta \leq \pi$, $0 \leq \varphi \leq 2\pi$. Координатные поверхности в этой системе представляют собой: семейство плоскостей ($\varphi = \text{const}$); семейство конусов с вершиной в центре сферы и углом при вершине ($2\theta = \text{const}$); семейство концентрических сфер с центром в начале координат ($r = \text{const}$). Волновое уравнение

$$\nabla^2 s - \left(\frac{k}{\omega}\right)^2 \frac{\partial^2 s}{\partial t^2} = 0$$

для функции, зависящей от всех координат $s(x, y, z, t)$, в сферической системе координат имеет вид, в частном случае, когда функция s не зависит от угловых координат:

$$\frac{\partial^2 s}{\partial t^2} - \left(\frac{\omega}{k}\right)^2 \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial s}{\partial r} \right) \right] = 0. \quad (1.41)$$

Произведя замену переменных $u = rs$, из (1.41) получим уравнение

$$\frac{\partial^2 u}{\partial t^2} - \left(\frac{\omega}{k}\right)^2 \frac{\partial^2 u}{\partial r^2} = 0,$$

частное решение которого имеет вид:

$$u = F(t - r/v); \quad s = r^{-1} F(t - r/v), \quad (1.42)$$

где v – скорость распространения волны вдоль направления, определяемого единичным вектором, коллинеарным радиус-вектору.

Уравнение волнового фронта $r = v(t - t_0)$ есть уравнение сферической поверхности, радиус которой увеличивается пропорционально времени. Таким образом, функция $s(r, t)$ представляет собой сферическую волну, распространяющуюся со скоростью v . Значение функции s в сферической волне уменьшается обратно пропорционально расстоянию r .

ЛЕКЦИЯ 2. Плоские электромагнитные волны в изотропных средах

Согласно положениям макроскопической электродинамики, электромагнитное поле в любой среде определяется в каждый момент времени векторами $\mathbf{E}$ и $\mathbf{D}$, характеризующими электрическое поле, и векторами $\mathbf{H}$ и $\mathbf{B}$, определяющими магнитное поле. В уравнениях электродинамики, помимо этих четырех векторов, фигурируют плотности заряда ρ и тока $\mathbf{j}$, характеризующие источники поля, под воздействием которых в окружающей среде возникает электромагнитное поле. Поскольку анализ электромагнитных явлений возможен только на основе полной системы уравнений электродинамики, обратимся к системе уравнений Максвелла:

$$\text{rot } \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{j}; \text{div } \mathbf{D} = \rho; \mathbf{D} = \epsilon \mathbf{E}; \text{div } \mathbf{D} = \rho; \mathbf{D} = \epsilon \mathbf{E}; \quad (2.1)$$

$$\text{rot } \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} + \mathbf{j}_m; \text{div } \mathbf{B} = \rho_m; \mathbf{B} = \mu \mathbf{H}, \quad (2.2)$$

где $\mathbf{E}$, $\mathbf{H}$ – мгновенные значения векторов напряженности соответственно электрического и магнитного полей; $\mathbf{D}$, $\mathbf{B}$ – мгновенные значения векторов электрической и магнитной индукции; ϵ , μ – диэлектрическая и магнитная проницаемость среды; $\mathbf{j}$, $\mathbf{j}_m$ – мгновенные значения векторов объемной плотности электрического и магнитного токов; ρ , ρ_m – объемные плотности электрических и магнитных зарядов.

Величина плотности электрического тока $\mathbf{j}$ в уравнениях (2.1) представляет собой сумму токов первичных (сторонних) источников $\mathbf{j}^{\text{ст}}$ и вторичных источников, возбуждаемых самим полем, $\mathbf{j}^{\text{в}}$. Плотность тока вторичных источников определяется, согласно закону Ома, выражением $\mathbf{j}^{\text{в}} = \sigma \mathbf{E}$, где σ – проводимость среды.

При анализе конкретных электродинамических систем во избежание чрезмерного усложнения задачи выделяют область первичных источников. Поля, созданные сторонними токами, в процессе решения задачи считаются заданными и не зависящими от порождаемых ими полей.

Величины ϵ , μ и σ , входящие в уравнения электродинамики, в большинстве сред являются скалярными. Такие среды называются изотропными, и здесь зависимость между электрической индукцией и напряженностью электрического поля и между магнитной индукцией и напряженностью магнитного поля в материальных уравнениях системы (2.2) является линейной. Однако существуют среды, в которых эта зависимость имеет другой вид. Например, в сегнетоэлектриках нарушается линейность соотношения $\mathbf{D} = \epsilon \mathbf{E}$, а в ферромагнитных веще-

ствах – линейность соотношения $\mathbf{B} = \mu\mathbf{H}$, и, таким образом, параметры сред зависят от электрической и магнитной напряженности поля. Существуют среды, параметры которых изменяются в различных направлениях. Такие среды называются анизотропными. Диэлектрическая и магнитная проницаемость таких сред определяются тензорами.

Материальные уравнения Максвелла для векторов $\mathbf{D}$ и $\mathbf{B}$ в этом случае запишутся в виде:

$$\mathbf{D} = \tilde{\epsilon}\mathbf{E}; \mathbf{B} = \tilde{\mu}\mathbf{H}, \quad (2.3)$$

где $\tilde{\epsilon}$, $\tilde{\mu}$ – тензоры соответственно диэлектрической и магнитной проницаемости. В тензорном виде уравнения (2.3) имеют вид:

$$D_i = \epsilon_{ij} E_j; B_j = \mu_{ij} H_j. \quad (2.4)$$

Для случая гармонической зависимости от времени, когда $\mathbf{E} = \mathbf{E}_m \cos(\varphi - \omega t)$; $\mathbf{H} = \mathbf{H}_m \cos(\varphi - \omega t)$, где $\mathbf{E}_m$, $\mathbf{H}_m$ – максимальные амплитуды, система уравнений (2.1) значительно упрощается. Действительно, применив метод комплексных амплитуд, например, к первому уравнению системы (2.1), получим

$$\text{rot} \dot{\mathbf{H}} = -j\omega \dot{\mathbf{D}} + \dot{\mathbf{j}},$$

где $\dot{\mathbf{H}}, \dot{\mathbf{D}}, \dot{\mathbf{j}}$ – комплексные амплитуды соответствующих векторов.

Таким образом, мы получили уравнения относительно комплексных амплитуд, не зависящие от времени. Очевидно, что все уравнения Максвелла легко приводятся к комплексной форме заменой функций поля их комплексными амплитудами.

$$\dot{\epsilon} = \epsilon + \sigma / \omega,$$

где $\dot{\epsilon}$ – комплексная диэлектрическая проницаемость.

Диэлектрическая и магнитная проницаемость в общем случае – комплексные величины: $\dot{\epsilon} = \epsilon' + j\epsilon''$, $\dot{\mu} = \mu' + j\mu''$. Тогда полная система уравнений Максвелла в комплексной форме запишется следующим образом:

$$\left. \begin{aligned} \text{rot} \dot{\mathbf{H}} &= -j\omega \dot{\epsilon} \dot{\mathbf{E}} + \dot{\mathbf{j}}^{\text{ct}}; \quad \text{div} \dot{\mathbf{D}} = \dot{\rho} \quad \dot{\mathbf{D}} = \dot{\epsilon} \dot{\mathbf{E}}; \\ \text{rot} \dot{\mathbf{E}} &= j\omega \dot{\mu} \dot{\mathbf{H}} + \dot{\mathbf{j}}_m^{\text{ct}}; \quad \text{div} \dot{\mathbf{B}} = \dot{\rho}_m \quad \dot{\mathbf{B}} = \dot{\mu} \dot{\mathbf{H}} \end{aligned} \right\} \quad (2.5)$$

Рассмотрим случай, когда электромагнитное поле создается только электрическими токами ($\rho_m = 0$). Для однородной среды, в которой $\epsilon = \text{const}$, $\mu = \text{const}$, применив к первому уравнению системы (2.5) операцию rot , исключив $\text{rot} \dot{\mathbf{E}}$ при помощи второго уравнения системы, исключив из системы (2.5) вектор $\mathbf{H}$, получим уравнение для вектора $\dot{\mathbf{E}}$:

$$\nabla^2 \dot{\mathbf{E}} + k^2 \dot{\mathbf{E}} = \text{grad} \frac{\dot{\rho}}{\dot{\epsilon}} - j\omega \dot{\mathbf{j}}^{\text{ct}}.$$

Аналогичные уравнения получаются для случая, когда поле создается только магнитными токами.

Очевидно, что неоднородность уравнений Гельмгольца обусловливается наличием в рассматриваемой области сторонних токов и зарядов. На практике неоднородные уравнения для полей обычно не употребляются, так как для определения поля, созданного токами и зарядами, удобнее исходить непосредственно из уравнений (2.5) или вводить некоторые вспомогательные величины.

Таким образом, для определения векторов $\dot{\mathbf{E}}$ и $\dot{\mathbf{H}}$ свободного электромагнитного поля в однородной среде получаем следующие векторные волновые уравнения:

$$\nabla^2 \dot{\mathbf{E}} + k^2 \dot{\mathbf{E}} = 0; \nabla^2 \dot{\mathbf{H}} + k^2 \dot{\mathbf{H}} = 0. \quad (2.6)$$

При использовании волновых уравнений надо учитывать, что решения уравнения для $\dot{\mathbf{E}}$ и $\dot{\mathbf{H}}$ не независимы, а связаны между собой уравнениями (2.5), т.е. уравнения (2.6) не эквивалентны уравнениям (2.5). Это видно хотя бы из того, что система (2.6) может иметь решение (например $\dot{\mathbf{E}} = 0$ и $\dot{\mathbf{H}} = 0$), а система (2.5) – нет. На практике для определения одного из векторов поля пользуются уравнениями (2.6), а второй определяется уравнениями (2.5). Если отвлечься от процессов возбуждения и излучения электромагнитных волн и исходить из факта их существования, можно с помощью уравнений Максвелла (2.6) и волновых уравнений (2.6) исследовать общие свойства поля в среде, свободной от источников.

Рассмотрим электромагнитное поле в однородной среде, изменяющееся вдоль одной координаты. Пусть это будет координата z выбранной декартовой системы координат. Исследуем такое поле с помощью уравнений (2.5) и (2.6), отбросив производные по координатам x и y . Воспользуемся волновым уравнением для электрического вектора:

$$d^2 \dot{\mathbf{E}} / dz^2 + k^2 \dot{\mathbf{E}} = 0. \quad (2.7)$$

Если вектор $\dot{\mathbf{E}} = a\dot{\mathbf{E}}$ не изменяет своего направления в пространстве, то решение уравнения (2.7) так же, как и решение уравнения (1.30), запишется в виде

$$\dot{\mathbf{E}} = a(\dot{E}_1 e^{jkz} + \dot{E}_2 e^{-jkz}). \quad (2.8)$$

Для полного описания плоской волны обратимся теперь к уравнениям Максвелла (2.5).

$$\left. \begin{aligned} d\dot{H}_x / dz &= -j\omega \epsilon \dot{E}_y; & d\dot{E}_x / dz &= j\omega \mu \dot{H}_y \\ d\dot{H}_y / dz &= -j\omega \epsilon \dot{E}_x; & d\dot{E}_y / dz &= -j\omega \mu \dot{H}_x \end{aligned} \right\}. \quad (2.9)$$

Из уравнений (2.9) очевидно, что в плоской волне отсутствуют составляющие поля $\dot{E}_z$ и $\dot{H}_z$. Так как направление вектора $\mathbf{E}$ в пространстве неизменно и определяется единичным вектором $\mathbf{a}$, мы вправе выбрать $\mathbf{a} = \mathbf{x}_0$, где $\mathbf{x}_0$ – единичный вектор нашей системы коор-

динат. Таким образом, векторы $\mathbf{E}$ и $\mathbf{H}$ перпендикулярны друг к другу и к направлению распространения волны, т.е. они лежат в плоскости фронта волны. Уравнения (2.9) при этом принимают вид:

$$\left. \begin{aligned} d\dot{H}_y / dz &= j\omega\epsilon\dot{E}_x; \quad d\dot{E}_x / dz = j\omega\mu\dot{H}_y \\ \dot{E}_y &= \dot{E}_z = \dot{H}_x = \dot{H}_y = 0 \end{aligned} \right\}. \quad (2.10)$$

Исследуем волну, распространяющуюся вдоль оси z . Для этого в выражении (2.8) следует положить $\dot{E}_2 = 0$. Учитывая выбранную ориентацию $\mathbf{a} = \mathbf{x}_0$, получим $\dot{\mathbf{E}} = x_0 \dot{E}_1 e^{jkz}$. Подставив во второе уравнение системы (2.9) величину $\dot{E}_x = \dot{E}_1 e^{jkz}$, получим

$$\frac{\dot{E}_x}{\dot{H}_y} = \frac{\dot{E}}{\dot{H}} = \frac{\omega\mu}{k} \sqrt{\frac{\omega}{\mu}} = W. \quad (2.11)$$

Эта величина называется *волновым сопротивлением*. Поскольку волновое число вещественно, электрическое и магнитное поля колеблются в фазе. Для вакуума

$$W = W_0 = \sqrt{\mu_0 / \epsilon_0} = 120\pi. \quad (2.12)$$

Запишем окончательные выражения для векторов $\mathbf{E}$ и $\mathbf{H}$ плоской однородной волны в непоглощающей среде, полученные из решения уравнений (2.5)

$$\mathbf{E} = x_0 E_1 \cos(kz - \omega t + \varphi_1); \quad \mathbf{H} = y_0 \frac{E_1}{W} \cos(kz - \omega t + \varphi_1). \quad (2.13)$$

Поскольку речь идет о свободных колебаниях, не возбуждаемых каким-то конкретным источником, амплитуда E_1 и начальная фаза φ_1 остаются неопределенными. Вектор Пойнтинга направлен вдоль оси z и определяется выражением

$$\mathbf{\Pi} = [\mathbf{E}, \mathbf{H}]. \quad (2.14)$$

Напомним, что полученное из теории Максвелла численное значение v_0 , близкое к скорости света, стало аргументом в пользу электромагнитной природы света. Плоская волна, таким образом, представляет собой электромагнитное поле, векторы напряженности которого не имеют составляющих вдоль направления распространения. Такое поле в общем виде может быть представлено как сумма двух волн, распространяющихся навстречу друг другу. Волну такого типа называют поперечно-электромагнитной волной.

ЛЕКЦИЯ 3. Особенности распространения волновых пучков. Дифракция

В природе практически не существует плоских волн с неограниченной волновой поверхностью. Во всех опытах, проводимых в оптике, акустике, радиофизике, волновой процесс всегда ограничен в пространстве размерами приборов и апертур излучающих систем. При прохождении волн через различные отверстия также происходит ограничение волнового процесса. Такой ограниченный в пространстве волновой процесс называется волновым пучком. В некоторых случаях можно считать, что волновой пучок распространяется прямолинейно вдоль некоторого луча, подобно тому как в геометрической оптике распространяется пучок света. Однако при прохождении пучка вблизи края какого-либо непрозрачного экрана его путь не остается строго прямолинейным. Оказывается, что волна частично проникает в область тени, т.е. возникает дифракция. В широком смысле дифракцией называется поведение волн в некоторой области, имеющей границу с теми или иными свойствами. Как только на пути распространения волн возникает препятствие, образуется рассеянное поле. Задача теории дифракции состоит в определении поля вокруг препятствия, образованного в результате взаимодействия падающей и рассеянной волн. К дифракционным относятся задачи о прохождении волн сквозь отверстия, огибании волнами различных препятствий, отражении и поглощении волн телами и т.д.

С математической точки зрения теория дифракции, состоящая в решении линейных дифференциальных уравнений в частных производных с линейными краевыми условиями, настолько сложна, что до настоящего времени лишь небольшое число задач дифракции имеет строгое решение. В связи с этим первостепенное значение приобретают приближенные методы решения дифракционных задач. При решении практических задач в оптике, акустике, электродинамике применяют метод Кирхгофа, основанный на принципе Гюйгенса–Френеля, согласно которому поле от источника Q в некоторой точке наблюдения P равно сумме волн вторичных источников, расположенных на некоторой волновой поверхности между точкой P и источником Q . Кирхгоф дал строгую формулировку этого принципа, основываясь на волновом уравнении.

Рассмотрим некоторую монохроматическую волну $u(x, y, z, t) = u_s(x, y, z)e^{-j\omega t}$. Очевидно, что в вакууме часть этой волны удовлетворяет уравнению Гельмгольца

$$(\nabla^2 + k^2)u_s = 0. \quad (3.1)$$

Если функция u_s и любая другая функция W вместе со своими

первыми и вторыми производными однозначны и непрерывны внутри некоторой области V и на поверхности S , ограничивающей объем V , то, согласно теореме Грина,

$$\int_V (u_s \nabla^2 W - W \nabla^2 u_s) dV = \int_S \left(W \frac{\partial u_s}{\partial n} - u_s \frac{\partial W}{\partial n} \right) dS \quad (3.2)$$

где n – внутренняя к поверхности S нормаль.

В качестве второй функции выберем функцию Грина $W = e^{jkr} / r$, которая непрерывна во всех точках объема V , за исключением точки P , которую мы окружим сферой радиуса r_0 . Будем устремлять r_0 к нулю. Тогда уравнение (3.2) примет вид

$$\int_V (u_s \nabla^2 W - W \nabla^2 u_s) dV = \int_S \left(W \frac{\partial u_s}{\partial n} - u_s \frac{\partial W}{\partial n} \right) dS + \left(W \frac{\partial u_s}{\partial n} - u_s \frac{\partial W}{\partial n} \right) dS. \quad (3.3)$$

Еще выражение (3.3) можно записать в виде:

$$u_P = -\frac{1}{4\pi} \int_S (u_s \nabla^2 W - W \nabla^2 u_s) dV + \frac{1}{4\pi} \int_S \left(W \frac{\partial u_s}{\partial n} - u_s \frac{\partial W}{\partial n} \right) dS. \quad (3.4)$$

Если функция W удовлетворяет волновому уравнению (3.1), то очевидно, что $u_s \nabla^2 W - W \nabla^2 u_s = 0$, и уравнение (3.4) преобразуется к виду

$$u_P = \int_S \left(W \frac{\partial u_s}{\partial n} - u_s \frac{\partial W}{\partial n} \right) dS. \quad (3.5)$$

Пользуясь формулой (3.5), можно найти функцию u_P в любой точке объема по заданному распределению функции u_s и ее производной на поверхности, ограничивающей объем. Формула (3.5) является строгим определением принципа Гюйгенса–Френеля. Так как поле в точке P однозначно определяется значениями u_s и $\partial u_s / \partial n$ на поверхности S , то очевидно, что между ними должна существовать общая функциональная зависимость. Эту зависимость можно получить, устремив точку наблюдения к поверхности. При этом оказывается, что значения u_s и $\partial u_s / \partial n$ связаны между собой интегральным преобразованием. Следовательно, в строгой формулировке принцип Гюйгенса–Френеля может быть применен для решения задач дифракции только совместно с решением некоторого интегрального уравнения.

Если в формуле (3.5) вместо u_s и $\partial u_s / \partial n$ подставить произвольные значения u_0 и $\partial u_0 / \partial n$, то придем к решению, называемому интегралом Кирхгофа. Здесь u_0 и $\partial u_s / \partial n$ – уже не краевые значения функции и ее нормальной производной, а некоторые разумно выбранные приближенные их значения. Кирхгоф принял следующие предпо-

ложения, называемые граничными условиями Кирхгофа:

1. На отверстии распределение поля и его производная имеют точно такие же значения, какие они имели бы в отсутствие экрана, т.е. $u_0 = u_{\text{отв}}$, $\partial u_0 / \partial n = \partial u_{\text{отв}} / \partial n$.

2. На той части поверхности, которая лежит в области геометрической тени, распределение поля и его производная тождественно равны нулю.

Для нахождения поля в точке P строим поверхность, ограничивающую объем с точкой P , таким образом, чтобы источник и дифрагирующее тело исключить из объема. Поле в точке P можно вычислить с помощью интеграла Кирхгофа. Для этого поверхность S разобьем на три части: S_0 – поверхность отверстия, S_1 – поверхность геометрической тени и S_2 – поверхность, уходящая в бесконечность. Согласно условию Зоммерфельда, заключающемуся в том, что поле на бесконечности равно нулю, интеграл по S_2 будет равен нулю. Интеграл по S_1 равен нулю в силу граничных условий Кирхгофа. Следовательно, выражение (3.5) принимает вид

$$u_P = \frac{1}{4\pi} \int_{S_0} \left(W \frac{\partial u_{\text{отв}}}{\partial n} - u_{\text{отв}} \frac{\partial W}{\partial n} \right) dS. \quad (3.6)$$

Таким образом, сущность метода Кирхгофа состоит в следующем. Вспомогательная поверхность, замыкающаяся на бесконечности, плотно облегает теневую сторону дифрагированного тела; интегрирование в формуле (3.6) распространяется на освещенную часть (отверстие) этой поверхности в предположении, что значения u_0 и $\partial u_0 / \partial n$ определяются в ней только падающей волной независимо от дифрагирующего тела.

3.1. Угловой спектр плоских волн

При решении задач о дифракции часто используется метод разложения поля в плоскости отверстия по плоским волнам. Предположим, что на отверстие в непрозрачном экране, расположенном в плоскости $z = 0$, падает в направлении оси z плоская монохроматическая волна. Необходимо определить поле в произвольной точке P за экраном. Запишем поле в точке P в виде двумерного Фурье-преобразования:

$$u(x, y, z) = \iint_{-\infty}^{\infty} A(k_x, k_y, z) \exp[j(k_x x + k_y y)] dk_x dk_y. \quad (3.7)$$

Функция $A(k_x, k_y, z)$ называется угловым спектром функции $u(x, y, z)$ в плоскости, параллельной плоскости xOy и расположенной на расстоянии z от нее. Подынтегральное выражение в преобразова-

нии (3.7) представляет собой комплексную амплитуду плоской гармонической волны с амплитудой $A(k_x, k_y, z)$, зависящей от направления распространения, так как

$$k_x = k \cos \alpha, k_y = k \cos \beta, k_z = \sqrt{k^2 - k_x^2 - k_y^2} = k \cos \gamma,$$

где $\cos \alpha, \cos \beta, \cos \gamma$ – направляющие косинусы нормали к фронту плоской волны. Таким образом, выражение (3.7) представляет собой разложение комплексной амплитуды поля в плоскости, параллельной плоскости xOy , в угловой спектр плоских волн.

Рассмотрим, каким образом отверстие в экране влияет на угловой спектр. Для этого будем характеризовать экран с отверстием амплитудными коэффициентами пропускания $g(x, y)$, равными единице в пределах отверстия и нулю вне отверстия. Предположим, что при падении плоской волны на отверстие выполняются граничные условия Кирхгофа. Тогда поле в плоскости, расположенной за экраном, можно записать в виде

$$u_0(x, y, 0) = u_i(x, y, 0)g(x, y),$$

где $u_i(x, y, 0)$ – поле волны, падающей на экран.

Если угловой спектр падающей волны обозначить через $A_i(k_x, k_y)$, а спектр коэффициента пропускания – через $G(k_x, k_y)$, то, согласно теореме свертки анализа Фурье, угловой спектр поля за экраном определится соотношением

$$A_0(k_x, k_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A_i(\xi, \eta) G(k_x - \xi, k_y - \eta) d\xi d\eta.$$

Пусть, например, плоская волна нормально падает на бесконечно длинную узкую щель шириной a , прорезанную в непрозрачном экране, расположенном в плоскости xOy . Если щель параллельна оси y , а плоскость падения совпадает с плоскостью xOz , то Фурье-образ коэффициента пропускания имеет вид

$$G(k_x) = A_0(k_x) = \frac{1}{2\pi} \int_{-a/2}^{a/2} e^{-jk_x x} dx = \frac{\sin(k_x a / 2)}{\pi k_x}.$$

Ширину углового спектра можно определить из условия обращения в нуль функции $A_0(kx)$:

$$\frac{ak}{2}(\sin \theta - \sin \theta_0) = \pi. \quad (3.8)$$

Ширина спектра зависит от угла падения плоской волны на щель и отношения λ/a . Для определения поля в точке P установим функциональную зависимость между угловыми спектрами плоских волн $A(k_x, k_y, z)$ и $A_0(k_x, k_y)$. Для этого, подставив выражение (3.6) в волновое уравнение Гельмгольца $\nabla^2 u + k^2 u = 0$, которому функция

$u(x, y, z)$ удовлетворяет во всех точках, где нет источников, получим дифференциальное уравнение для функции $A(k_x, k_y, z)$

$$A(k_x, k_y, z) = A_0(k_x, k_y) \exp \left[\pm z \sqrt{k^2 - k_x^2 - k_y^2} \right], \quad (3.9)$$

откуда видно, что при $k_x^2 + k_y^2 < k^2$ распространение волны на расстояние r проявляется только в изменении относительных фаз между плоскими волнами, распространяющимися под различными углами к оси z . Если $k_x^2 + k_y^2 > k^2$, то квадратный корень в уравнении (3.9) будет мнимым. Тогда очевидно, что выражение (3.9) описывает неоднородные волны, амплитуда которых затухает по мере удаления от плоскости $z = 0$. Предельный случай $k_x^2 + k_y^2 = k^2$ соответствует плоским волнам, распространяющимся перпендикулярно к оси z .

ЛЕКЦИЯ 4. Элементы теории излучения волн. Акустические излучатели

Основная задача теории излучения волн заключается в том, чтобы по известным источникам найти создаваемое ими поле. Источником энергии волн является передатчик, энергия от которого поступает в излучающее устройство (антенну) по соединительной линии. Эффективность излучения в большой степени зависит от конструктивных особенностей антенны и ее основных характеристик. Важное значение при этом имеет понятие направленности излучения. Излучение реальных антенн в окружающее пространство неодинаково в различных направлениях. Степень различия интенсивности описывается диаграммой направленности.

Диаграммой направленности передающей антенны называется зависимость амплитуды поля, излучаемого антенной, от направления на точку наблюдения при условии, что эта точка перемещается на фиксированном расстоянии от центра антенны. Это значит, что в выражении для любой физической величины s

$$s(\theta, \varphi) = s_0 f(\theta, \varphi) e^{j\psi(\theta, \varphi)},$$

где θ, φ – координаты точки наблюдения в сферической системе координат, диаграммой направленности является функция $f(\theta, \varphi)$. Для удобства сравнения диаграммы направленности различных антенн нормируются. Нормированной диаграммой направленности называется отношение

$$E(\theta, \varphi) = f(\theta, \varphi) / f_{\max}(\theta, \varphi).$$

Иногда пользуются диаграммой направленности по интенсивности излучения, равной отношению интенсивности $J(\theta, \varphi)$ к максимальной интенсивности $J_{\max}$. Естественно, что $F^2(\theta, \varphi) = J(\theta, \varphi) / J_{\max}$. Излучатели с резко выраженной направленностью позволяют увеличить дальность действия волн без дополнительного увеличения мощности.

Для оценки выигрыша в мощности за счет направленности излучения антенны вводится понятие коэффициента направленного действия (КНД), представляющего собой отношение интенсивности (плотности потока энергии) в некоторой точке поля (θ', φ') к интенсивности ненаправленного излучателя J_0 в точке, находящейся на том же расстоянии от антенны, когда мощности этих излучателей одинаковы:

$$D(\theta', \varphi') = \frac{J(\theta', \varphi')}{J_0} = \frac{J(\theta', \varphi')}{P / 4\pi r^2}, \quad (4.1)$$

где P – мощность излучателя, определяемая через интенсивность выражением

$$P = \int_0^{2\pi} \int_0^\pi J(\theta, \varphi) r^2 \sin \theta d\theta d\varphi. \quad (4.2)$$

Подставив выражение (4.2) в формулу (4.1) и разделив полученное равенство на $J_{\max}$, будем иметь

$$D_{\max} = \frac{4\pi F^2(\theta', \varphi')}{\int_0^{2\pi} \int_0^\pi F^2(\theta, \varphi) \sin \theta d\theta d\varphi}.$$

В направлении максимального излучения антенны КНД принимает максимальное значение:

$$D_{\max} = \frac{D(\theta', \varphi')}{F^2(\theta', \varphi')} = \frac{4\pi}{\int_0^{2\pi} \int_0^\pi F^2(\theta, \varphi) \sin \theta d\theta d\varphi}. \quad (4.3)$$

Изучая плоские акустические волны, мы убедились, что реальную плоскую волну можно создать только в ограниченной области, например, в среде, заполняющей трубу с абсолютно жесткими стенками. Для создания плоской волны в безграничном пространстве понадобилась бы колеблющаяся плоскость бесконечных размеров. Все акустические излучатели, представляющие собой любое колеблющееся тело конечных размеров, создают вдали от тела сферическую волну. Естественно, что вблизи от источника волна может быть и не сферической. Например, фронт волны вблизи кварцевой пластинки, колеблющейся с ультразвуковой частотой, представляет собой участки плоскости. Однако при удалении от источника на большое расстояние волна становится сферической. Таким образом, реальная волна, созданная акустическим излучателем, по мере распространения все более приближается к сферической, а при некоторых видах колебаний сферическая волна излучается прямо с поверхности тела.

Сферическая волна, амплитуда которой вдоль фронта одинакова во всех точках, называется сферически-симметричной волной. Согласно решению (1.41), выражения для сходящихся и расходящихся волн имеют вид:

$$p = \frac{1}{r} f(t - r/c); \quad p = \frac{1}{r} g(t + r/c), \quad (4.4)$$

где c — скорость звука.

В отличие от плоских волн профиль сферической волны не остается неизменным. Постоянным в сферической волне остается профиль, определяемый величинами p, r .

Скорость частиц в сферически-симметричной волне направлена по радиус-вектору точки, в которой находится частица, и зависит, как и давление, от времени и расстояния от центра волны. Для ее определения проинтегрируем по времени линеаризованное уравнение движения Эйлера $\rho \partial v / \partial t + \nabla p = 0$.

Если до какого-то момента времени возмущение в данной точке отсутствовало, то скорость бегущих волн, описываемых выражениями (4.4), определяется соотношением

$$v = \pm \frac{p}{\rho c} + \frac{1}{\rho r} \int_{-\infty}^t p dt. \quad (4.5)$$

Из соотношения (4.5) видно, что в отличие от плоской волны скорость частицы сферической волны не пропорциональна давлению в тот же момент времени, а стремится к значению $\pm p/\rho c$ при $r \rightarrow \infty$. При конечном значении r скорость частиц определяется двумя слагаемыми, первое из которых воспроизводит профиль давления и изменяется обратно пропорционально расстоянию r , а второе – обратно пропорционально r^2 . Для центра волны, таким образом, второе слагаемое в соотношении (4.5) является главным, а при удалении от центра им можно пренебречь.

Для несжимаемой жидкости (при замене $t \pm r/c$ на t в аргументе p), когда вся среда движется синфазно и волны как таковой нет, первое слагаемое в соотношении (4.5) обращается в нуль, а второе сохраняется. В связи с этим слагаемые называют соответственно волновым и неволновым, а расстояния, на которых эти члены играют существенную роль, – волновой (дальней) и неволновой (ближней) зонами. Область $kr \leq 1$ является неволновой, область $kr \geq 1$ – волновой, а при $kr = 1$, когда оба члена в соотношении (4.5) равны по модулю, область следует считать промежуточной.

Для создания сферически-симметричной расходящейся волны можно поместить в среду пульсирующую сферу, т.е. сферу с проницаемыми стенками, внутри которой попеременно возникает то избыток, то недостаток некоторого количества вещества данной среды. Такой излучатель является идеальным для создания сферически-симметричной волны. Если радиус сферы мал по сравнению с радиусом неволновой зоны, то излучатель называется монополем. Поскольку скорость частиц на поверхности сферы в неволновой зоне примерно обратно пропорциональна r^2 , а поверхность сферы пропорциональна r^2 , то общее количество втекающего или вытекающего через поверхность сферы вещества, необходимого для создания поля, не зависит от радиуса. В связи с этим при малом радиусе сферы говорят о точечном монополе.

Задача об излучении сферических волн сводится к решению волнового уравнения относительно потенциала колебательной скорости

$$\nabla^2 \varphi - \frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2} = 0, \quad (4.6)$$

которое должно удовлетворять граничному условию на поверхности колеблющегося тела

$$\partial\phi/\partial n = v(t) \quad (4.7)$$

и условию излучения $\phi = 0$ при $r \rightarrow \infty$.

Для пульсирующей сферы радиуса a , скорость колебания точек которой по нормали изменяется по закону $v_n(t) = v_0 \cos \omega t$ и не зависит от угловых координат, решение волнового уравнения (4.6) в сферической системе координат принимает вид

$$\dot{\phi} = \frac{1}{r} (Ae^{jkr} + Be^{-jkr}). \quad (4.8)$$

Скорость частиц и звуковое давление выражаются формулами:

$$\begin{aligned} \dot{v} &= \frac{d\dot{\phi}}{dr} = \frac{1}{r^2} \{ [\dot{A}(jkr - 1)]e^{jkr} - [\dot{B}(1 + jkr)]e^{-jkr} \}; \\ \dot{p} &= j\omega\rho\dot{\phi} = j\frac{\omega\rho}{r} (\dot{A}e^{jkr} + \dot{B}e^{-jkr}). \end{aligned} \quad (4.9)$$

Из условия излучения следует, что $\dot{B} = 0$, а из условия (4.7) $\partial\phi/\partial n = \dot{v}_0$ при $r = a$ получим выражение для постоянной $\dot{A}$:

$$\dot{A} = \frac{\dot{v}_0 a^2}{jka - 1} e^{-jka}.$$

Выполняя умножение и деление выражений для $\dot{v}$ и $\dot{p}$ на j и переходя от комплексных амплитуд к действительным их значениям в предположении, что фаза скорости принимается за начальную, т.е. $\arctg(1/kr) - \arctg(1/ka) = \text{const} = 0$, получаем:

$$\begin{aligned} v &= \frac{Q}{4\pi r^2} \sqrt{\frac{1 + k^2 r^2}{1 + k^2 a^2}} \cos(kr_1 - \omega t); \\ p &= \frac{Q}{4\pi r^2} \frac{\rho c k r}{\sqrt{1 + k^2 a^2}} \cos(kr_1 - \omega t - \psi), \end{aligned}$$

где $Q = 4a^2 v_0$ – амплитуда объемной скорости;

Интенсивность излучения выражается формулой

$$J = \frac{Q^2 k^2 \rho c}{32\pi^2 r^2 (1 + k^2 a^2)}.$$

Мощность излучения пульсирующей сферы

$$P = 4\pi r^2 J.$$

4.1. Элементарный электрический излучатель. Полуволновой диполь

Простейшим излучателем электрического типа является элементарный электрический диполь, представляющий собой короткий по сравнению с длиной волны провод, по всей длине которого ток имеет постоянные амплитуду и фазу. Несмотря на некоторую идеализацию,

анализ этого простейшего излучателя дает возможность понять основные процессы, происходящие в реальных антенных системах, наметить общий ход их расчета. Каким бы ни было распределение амплитуды и фазы тока вдоль какого-нибудь провода, в пределах отрезка $l \leq \lambda$ их можно принять неизменными. Таким образом, сложные антенные системы можно считать составленными из элементарных излучателей. Практическим аналогом элементарного излучателя является устройство, состоящее из двух коротких проводов, на концах которых расположены шары, представляющие собой сосредоточенные емкости. Сторонний ток $I^{\text{ст}}$, проходящий по проводам, переходит в ток смещения, замыкающийся через распределенные емкости C_1 и C_2 , существующие между проводами вибратора. Подавляющая часть тока смещения проходит через сосредоточенную емкость между шарами, так как она существенно превышает распределенные емкости. Впервые такая схема была предложена немецким физиком Генрихом Герцем. Излучатель называется диполем Герца.

Перейдем к анализу поля элементарного электрического вибратора, расположенного в безграничной однородной изотропной непроводящей среде, характеризуемой параметрами ε и μ . Для этого удобно напряженность электрического поля выразить через векторный потенциал, который легко получить из уравнений Максвелла следующим образом. Поскольку дивергенция ротора любого вектора равна нулю, на основании уравнения Максвелла $\text{div} \mathbf{B} = 0$ удобно представить $\mathbf{B}$ как ротор некоторого вектора $\mathbf{A}$, называемого векторным потенциалом:

$$\mathbf{B} = \text{rot } \mathbf{A} \text{ или } \mathbf{H} = \frac{1}{\mu} \text{rot } \mathbf{A}. \quad (4.10)$$

Подставив выражение (4.10) в уравнение Максвелла для вектора $\mathbf{E}$, получим

$$\mathbf{E} = - \left(\frac{\partial \mathbf{A}}{\partial t} + \text{grad } \varphi \right). \quad (4.11)$$

Подставляя соотношения (4.10) и (4.11) в уравнение Максвелла и учитывая, что $\mathbf{D} = \varepsilon \mathbf{E}$, получаем

$$- \text{grad} \left(\text{div} \mathbf{A} + \varepsilon \mu \frac{\partial \varphi}{\partial t} \right) + \nabla^2 \mathbf{A} - \varepsilon \mu \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu \mathbf{j}. \quad (4.12)$$

Уравнение (4.12) при наложении условия

$$\text{div} \mathbf{A} + \varepsilon \mu \frac{\partial \varphi}{\partial t} = 0 \quad (4.13)$$

принимает вид

$$\nabla^2 \mathbf{A} - \varepsilon \mu \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu \mathbf{j}. \quad (4.14)$$

Подставив в уравнение Максвелла $\text{div} \mathbf{E} = \rho$ выражение (4.11) и используя соотношение (4.13), получим

$$\nabla^2 \varphi - \varepsilon \mu \frac{\partial^2 \varphi}{\partial t^2} = -\frac{1}{\varepsilon} \rho \quad (4.15)$$

Уравнения (4.14) и (4.15) описывают волновой процесс и являются соответственно волновыми. Они имеют следующие решения:

$$A = \frac{\mu}{4\pi} \int_V \frac{1}{r} \mathbf{j}(t - r/v) dV; \quad \varphi = \frac{1}{4\pi\varepsilon} \int_V \frac{1}{r} \rho(t - r/v) dV, \quad (4.16)$$

где dV – элемент объема, в котором заданы плотность тока $\mathbf{j}$ и плотность заряда ρ ; r – расстояние от элемента объема до точки, в которой определяются $\dot{A}$ и $\dot{\varphi}$; v – скорость распространения колебаний.

При расчете полей вокруг линейных проводов, когда токи проводимости и заряды сосредотачиваются только вдоль осей проводов, в выражениях (4.16) объемные интегралы могут быть заменены линейными. В случае гармонических колебаний выражения для комплексных амплитуд $\dot{A}$ и $\dot{\varphi}$ принимают вид

$$\dot{A} = \frac{\mu}{4\pi} \int_l \frac{\dot{\mathbf{I}}}{r} e^{jkr} dl; \quad \dot{\varphi} = \frac{1}{4\pi\varepsilon} \int_l \frac{\dot{\sigma}}{r} e^{jkr} dl, \quad (4.17)$$

где $\dot{\mathbf{I}}$ – ток проводимости, проходящий по проводу; $\dot{\sigma}$ – линейная плотность заряда.

Из соотношения (4.13) при гармонических колебаниях следует

$$\dot{\varphi} = -j \frac{1}{\varepsilon \mu \omega} \text{div} \dot{A}. \quad (4.18)$$

Подставляя равенство (4.18) в (4.11)

$$\dot{E} = j\omega \dot{A} + j \frac{1}{\varepsilon \mu \omega} \text{grad div} \dot{A}. \quad (4.19)$$

Это уравнение совместно с выражением (4.10) позволяет вычислить все составляющие электромагнитного поля, если известен векторный потенциал $\mathbf{A}$.

Поскольку для элементарного излучателя ток остается постоянным по всей его длине, значение $\dot{\mathbf{I}}$ в выражении (4.17) можно вынести за знак интеграла. Предполагая, кроме того, что $l = r$ можно вынести за знак интеграла и члены, зависящие от r . Таким образом, получаем

$$\dot{A} = \frac{\mu}{4\pi} \frac{\dot{\mathbf{I}}}{r} e^{jkr} l. \quad (4.20)$$

Систему координат выберем таким образом, чтобы ось диполя совпадала с осью z , а начало координат – с серединой диполя. При таком выборе системы координат имеем $A_x = A_y = 0$, $A_z = 0$. Тогда на основании формул (4.10) и (4.19) получим следующие выражения для

составляющих **E** и **H**

$$\begin{aligned}
\dot{E}_x &= \frac{xz\dot{I}l}{4\pi\epsilon\omega r^2} \left(j\frac{3}{r^3} + \frac{3k}{r^2} + j\frac{k^2}{r} \right) e^{jkr}; \\
\dot{E}_y &= \frac{yz\dot{I}l}{4\pi\epsilon\omega r^2} \left(j\frac{3}{r^3} + \frac{3k}{r^2} + j\frac{k^2}{r} \right) e^{jkr}; \\
\dot{E}_z &= \frac{\dot{I}l}{4\pi\epsilon\omega r^2} \left[j\frac{3z^2 - r^2}{r^3} + \frac{(3z^2 - r^2)k}{r^2} + j\frac{(z^2 - r^2)k^2}{r} \right] e^{jkr}; \\
\dot{H}_x &= -\frac{y\dot{I}l}{4\pi r} \left(\frac{1}{r^2} - j\frac{k}{r} \right) e^{jkr}; \quad \dot{H}_y = -\frac{x\dot{I}l}{4\pi r} \left(\frac{1}{r^2} - j\frac{k}{r} \right) e^{jkr}; \quad \dot{H}_z = 0.
\end{aligned} \tag{4.21}$$

Вследствие осевой симметрии элементарного излучателя удобно пользоваться формулами, определяющими поле в сферической системе координат. В этом случае напряженности полей E и H характеризуются составляющими $E_r, E_\theta, E_\varphi, H_r, H_\theta, H_\varphi$. Если ввести в рассмотрение составляющую

$$E_R = E_x \cos \varphi + E_y \sin \varphi, \tag{4.22}$$

то получим формулы для поля элементарного электрического излучателя в сферической системе координат с учетом того, что $k = 2\pi / \lambda$:

$$\begin{aligned}
\dot{E}_r &= \frac{1}{4\pi\epsilon} \frac{2\dot{I}l}{\omega} \left(\frac{2\pi}{\lambda} \right)^3 \cos \theta \left[j \left(\frac{\lambda}{2\pi r} \right)^3 + \left(\frac{\lambda}{2\pi r} \right)^2 \right] e^{jkr}; \\
\dot{E}_\theta &= \frac{\dot{I}l}{4\pi\epsilon\omega} \left(\frac{2\pi}{\lambda} \right)^2 \sin \theta \left[j \left(\frac{\lambda}{2\pi r} \right)^3 + \left(\frac{\lambda}{2\pi r} \right)^2 - j\frac{\lambda}{2\pi r} \right] e^{jkr}; \\
\dot{H}_\varphi &= \frac{\dot{I}l}{4\pi} \left(\frac{2\pi}{\lambda} \right)^2 \sin \theta \left[\left(\frac{\lambda}{2\pi r} \right)^2 - j\frac{\lambda}{2\pi r} \right] e^{jkr}.
\end{aligned} \tag{4.23}$$

В случае дальней (волновой) зоны можно пренебречь заключенными в квадратные скобки слагаемыми $(\lambda/2\pi r)$ в степенях выше первой; кроме того, составляющая E_r мала по сравнению с E_θ . В результате, подставляя $2\pi / \lambda = \omega \sqrt{\epsilon\mu}$, получаем:

$$\begin{aligned}
\dot{E}_\theta &= -j \frac{\dot{I}l\omega}{4\pi r} \sin \theta e^{jkr}; \quad \dot{H}_\varphi = -j \frac{\dot{I}l\omega}{4\pi r} \sqrt{\epsilon\mu} \sin \theta e^{jkr}; \\
E_r &= E_\varphi = H_r = H_\theta = 0.
\end{aligned} \tag{4.24}$$

Таким образом, в дальней зоне существуют фактически две составляющие поля E_θ и H_φ . Обе они не зависят от координаты φ , и диаграмму направленности в плоскости, перпендикулярной к оси диполя, можно изобразить в виде окружности, радиус которой в определенном масштабе равен $|\dot{H}_\varphi|$ или $|\dot{E}_\theta|$, вычисленному при фиксированных значениях радиуса-вектора r и угла θ .

Распространение волны сопровождается переносом энергии. Поток энергии за единицу времени через шаровую поверхность, в центре которой расположен излучатель, представляет собой излученную мощность

$$P_{\Sigma} = \int_S \Pi_{n\text{cp}} ds, \quad (4.25)$$

где $\Pi_{n\text{cp}}$ – среднее за период значение составляющей вектора Пойнтинга Π_n , нормальной к элементарной площадке $ds = r^2 \sin \theta d\theta d\varphi$. Энергия, проходящая через поверхность сферы в направлении r , определяется составляющими векторов напряженности электрического и магнитного полей, перпендикулярных к r , т.е. $\dot{E}_\theta$ и $\dot{H}_\varphi$. Имеем $\Pi_n = E_\theta H_\varphi$. Поскольку среднее за период значение произведения двух величин A и B , являющихся гармоническими функциями времени, равно действительной части произведения $\dot{A}\dot{B}^*/2$, где A и B – комплексные амплитуды, то

$$\Pi_{n\text{cp}} = \dot{E}_\theta \dot{H}_\varphi^*/2. \quad (4.25)$$

Подставляя выражения (4.23) и (4.25) в формулу (4.24), получаем

$$P_{\Sigma} = \frac{\pi}{3} \left(\frac{l}{\lambda} \right)^2 \sqrt{\frac{\mu}{\epsilon}} I^2. \quad (4.26)$$

Присоединение генератора к вибратору может быть произведено, например, посредством двухпроводного фидера. Под воздействием ЭДС генератора в вибраторе возникают электрические токи, которые распределяются вдоль вибратора таким образом, что создаваемое в окружающем пространстве электромагнитное поле удовлетворяет уравнениям Максвелла и граничным условиям. Задача об излучении волн сводится к установлению связи между током в вибраторе и созданным им электромагнитным полем. Наибольшее применение получил вибратор, длина которого $2l = \lambda/2$. Такой вибратор, называемый полуволновым диполем, является основным излучающим элементом сложных антенн. Если полуволновой диполь расположить вдоль оси, то нормированная диаграмма направленности в меридиональной плоскости описывается выражением

$$F(\theta) = \cos \frac{\pi}{2} (\cos \theta) / \sin \theta.$$

В плоскости, перпендикулярной к оси диполя, диаграмма направленности представляет собой, так же, как и у элементарного диполя, окружность. Коэффициент направленного действия полуволнового диполя равен 1,64.

ЛЕКЦИЯ 5. Виды антенн, применяемых в радиодиапазоне

Подобно тому, как все радиоволны классифицируются по диапазонам, среди антенн, применяемых в радиодиапазоне, следует выделить антенны ультракоротких волн с длиной волны от 1 мм до 10 м ($f = 30 \text{ МГц} - 300 \text{ ГГц}$), антенны коротких волн с длиной волны от 10 до 100 м ($f = 3-30 \text{ МГц}$), антенны средних ($\lambda = 100-1000 \text{ м}$) и длинных ($\lambda = 1-10 \text{ км}$) волн.

Ультракороткие волны широко применяются в различных областях радиотехники. Метровые волны ($\lambda = 1-10 \text{ м}$) используются в телевидении и радиовещании, для радиосвязи на расстояния 1000–2000 км (за счет рассеяния волн на неоднородностях ионосферы), для радионавигации. Дециметровые ($\lambda = 10 \text{ см} - 1 \text{ м}$) и сантиметровые (1–10 см) волны используются для радиорелейной связи, связи с космическими объектами, связи при помощи искусственных спутников Земли, применяемых в качестве ретрансляторов, для радиотелеметрии, а также дальней связи с помощью волн, рассеянных на неоднородностях в тропосфере.

Ультракороткие волны плохо огибают сферическую поверхность Земли и различные неровности земной поверхности, а также сильно поглощаются почвой и атмосферой. Вследствие этого устойчивая радиосвязь между пунктами может существовать только в пределах прямой видимости, на расстоянии r , определяемом по формуле $r = 3570 \left(\sqrt{h_{\text{пер}}} + \sqrt{h_{\text{пр}}} \right)$, где $h_{\text{пер}}$ и $h_{\text{пр}}$ – высоты подвеса соответственно передающей и приемной антенн.

В диапазоне ультракоротких волн применяется много типов антенн. Простейшими из них являются антенны, представляющие собой различные симметричные и несимметричные вибраторы. Вертикальный несимметричный вибратор может располагаться непосредственно над металлической поверхностью, например, над корпусом машины, или устанавливаться на мачте. Плечо вибратора часто выполняют в виде трубки или штыря длиной $0,25\lambda$.

Для расширения полосы применяется наиболее удобный в конструктивном отношении петлевой вибратор, предложенный советским радиотехником А.А. Пистолькорсом. В диапазоне метровых волн в качестве направленных антенн большое распространение получили антенны типа «волновой канал». Такая антенна состоит из активного вибратора, рефлектора и нескольких директоров. Вибратор излучает электромагнитное поле как в направлении рефлектора, так и в направлении директоров. Под действием этого поля в рефлекторе наводится ток, который создает вторичное поле – поле излучения рефлектора.

Если длину рефлектора выбрать равной $(0,51-0,53)\lambda$, а расстояние между рефлектором и активным вибратором – $(0,15-0,25)\lambda$, то результирующее поле за рефлектором будет равно разности полей вибратора и рефлектора, а в направлении директоров – их сумме. Директоры возбуждаются результирующим полем вибратора и рефлектора. Для того чтобы вторичные поля директоров складывались с результирующим полем в направлении излучения, длины директоров выбираются равными $(0,41-0,45)\lambda$, а расстояния между ними – $(0,1-0,34)\lambda$.

В дециметровом и сантиметровом диапазонах широкое распространение получили антенны, в которых вместо проволочных элементов используются излучающие поверхности. К ним относятся рупорные, линзовые и зеркальные антенны.

Рупорная антенна представляет собой волновод с плавно меняющимся сечением. При расширении узкой стенки волновода рупор называется Е-секторальным, широкой стенки – Н-секторальным, обеих стенок – пирамидальным.

Круглый волновод при плавном увеличении сечения переходит в конический рупор. Коэффициент направленного действия рупорной антенны может быть вычислен по формуле

$$D_{\max} = 4\pi \frac{s}{\lambda^2} v, \quad (5.1)$$

где s – площадь раскрыва рупора; v – коэффициент использования поверхности: $v = 0,6-0,8$.

Линзовая антенна состоит из облучателя, расположенного в фокусе линзы. В качестве облучателя часто используется укороченный рупор, создающий сферический фронт волны и имеющий широкую диаграмму направленности. При установке линзы фронт волны преобразуется в плоский, а диаграмма направленности – в узкую.

Зеркальные антенны состоят из облучателя и рефлектора. В зависимости от вида рефлектора различают следующие типы зеркальных антенн: параболические зеркала (параболоид вращения, усеченный парабоид, параболический цилиндр), сферические зеркала, плоские и уголкового зеркала. Зеркальная параболическая антенна состоит из металлической поверхности, выполненной в виде парабоида вращения, и антенны-облучателя, установленного в фокусе парабоида. Аналогично линзовой антенне в раскрыве зеркала сферический фронт облучателя преобразуется в плоский фронт волны, а диаграмма направленности – в узкую диаграмму. Коэффициент направленного действия можно вычислить по формуле. Величина коэффициента использования поверхности зеркальных антенн составляет обычно $0,5-0,6$. Однако путем применения облучателей со специальной диаграммой направленности, как это имеет место в системе «Орбита», удастся получить $v = 0,70-0,75$. Практически можно изготовить парабоиды

вращения очень больших относительных размеров, имеющие весьма узкие диаграммы направленности и коэффициент направленного действия порядка миллионов единиц. Такие параболоиды применяются для космической радиосвязи. Например, прецизионный радиотелескоп Крымской астрофизической обсерватории имеет диаметр раскрытия 22 м, наименьшая длина волны рабочего диапазона 8 мм. Антенна Боннского радиотелескопа имеет диаметр раскрытия 100 м и работает в диапазоне $1\text{ см} < \lambda < 25\text{ см}$.

Короткие волны, используемые для радиосвязи (телефон, телеграф, фототелеграф) и радиовещания, при распространении вдоль земной поверхности сильно затухают. Поэтому связь при помощи земных волн возможна лишь на небольшие расстояния. Связь на большие расстояния осуществляется посредством ионосферных волн, которые достигают точки приема путем однократного отражения от верхних слоев ионосферы или многократных отражений от ионосферы и поверхности Земли. Следовательно, направление максимального излучения антенны должно составлять некоторый угол с линией горизонта.

На коротких волнах применяют в основном антенны с горизонтальной поляризацией. Горизонтально поляризованные волны при отражении от полупроводящей почвы при углах наклона $20\text{--}30^\circ$ испытывают меньшее поглощение, чем вертикально поляризованные. Прием на горизонтально поляризованные антенны сопровождается меньшими помехами, поскольку местные помехи атмосферного происхождения имеют преимущественно вертикальную поляризацию. Вертикальные вибраторы создают более интенсивное излучение вдоль земной поверхности, чем горизонтальные антенны, и поэтому применяются при связи на небольшие расстояния.

Простейшей антенной в диапазоне коротких волн является горизонтальный симметричный вибратор, выполненный из медного или биметаллического провода диаметром 3, 4 или 6 мм. Такой вибратор является узкополосной антенной и применяется для работы на фиксированной частоте.

Для расширения диапазона рабочих частот советский радиофизик С.И. Надененко предложил использовать вибратор, получивший название диполя Надененко. Диполь выполняется из параллельных медных или биметаллических проводов радиуса 2–3 мм.

В качестве направленных антенн в диапазоне коротких волн используется синфазная решетка. Полотно такой решетки включает n этажей (рядов) вибраторов. Каждый этаж состоит из n_z полуволновых вибраторов. Расстояние между этажами составляет половину длины волны. Питание антенны производится вертикальным фидером, который перекрещивается при переходе от этажа к этажу. При распределении энергии от одного этажа к другому волна, пройдя в

вертикальном фидере расстояние $0,5\lambda$, отстает по фазе на $kl = \pi$. Если перекрещивания не делать, то вибраторы соседних этажей будут возбуждены в противофазах. За полотном решетки на расстоянии $(0,2-0,3)\lambda$ расположено полотно рефлектора, представляющего собой точную копию полотна антенны. Коэффициент направленного действия горизонтальной антенны с учетом влияния земли можно определить по формуле $D_{\max} \approx 8n_b n_z$.

Средние и длинные волны могут распространяться вблизи поверхности земли (земная волна) и путем отражения от нижней границы ионосферы. Земные волны, на которых осуществляется связь в дневное время, сравнительно мало поглощаются в почве и хорошо огибают сферическую поверхность Земли. При этом связь осуществляется на расстояния до 2000–3000 км в длинноволновом диапазоне и 500–700 км в средневолновом. Связь на средних волнах устойчива только при отсутствии интерференции между земной и ионосферной волнами, что имеет место на расстояниях примерно 50–200 км от передатчика (зона уверенного приема). Далее расположена зона ближнего замирания (фединг), где поля земной и ионосферной волн соизмеримы. Прием в зоне пространственной волны возможен только в ночное время. При этом возникает явление дальнего замирания, так как в точку приема могут прийти волны, отраженные один или два раза от ионосферы. Условия распространения длинных и средних волн обуславливают ряд требований к антенным устройствам. Поскольку в дневное время прием волн возможен только за счет поверхностной волны, антенны должны обеспечивать интенсивное горизонтальное излучение. Такое излучение дает вертикальный несимметричный вибратор. Для борьбы с ближним замиранием на средних волнах эффективным средством является использование антифединговых антенн. Такая антенна в отличие от элементарного вибратора имеет диаграмму направленности в вертикальной плоскости, сильно прижатую к поверхности земли. Такая диаграмма показана сплошной линией, а штриховой – диаграмма элементарного электрического излучателя. Благодаря такому распределению на расстояниях до нескольких сотен километров от передатчика поле поверхностной волны намного превышает поле ионосферной. Замирания при этом практически устраняются, зона уверенного приема увеличивается. Такие диаграммы направленности имеют несимметричные вертикальные антенны высотой $(0,53-0,58)\lambda$.

Для средних и особенно длинных волн вертикальные антенны в виде металлических мачт-излучателей, установленных на изоляторах, в ряде случаев построить весьма трудно, поскольку стоимость мачты прямо пропорциональна третьей степени ее высоты. Максимальные размеры вертикальной антенны ограничиваются в основном высотами 250–350 м.

Длинные волны в настоящее время используются главным образом для радиотелеграфной связи (радиосвязь с подводными лодками, радионавигация и т.д.). Для радиовещания используется часть длинноволнового диапазона, соответствующая длинам волн $\lambda = 1050\text{--}2000$ м. Средние волны применяются для связи с передвижными радиостанциями, в авиационной радиослужбе и для радиовещания (200–550 м).

5.1. Распространение радиоволн вдоль земной поверхности.

Функция ослабления

Влияние земли на распространение волн сводится к следующему. При прохождении волн между двумя пунктами может иметь место отражение волны от земной поверхности. В этом случае в некоторой точке А, расположенной над поверхностью земли, поле будет представлять собой сумму прямой и отраженной волн, имеющих разный сдвиг фаз в различных точках пространства. В зависимости от разности фаз этих волн дальность действия, например, радиолокационной станции в одних направлениях увеличивается, а в других значительно уменьшается. Это явление существенно сказывается на работе радиолокационных станций метрового диапазона, имеющих, как правило, антенны с широкой диаграммой направленности.

Процесс распространения радиоволн над поверхностью земли неразрывно связан с явлением поглощения энергии волны вследствие проникновения радиоволн в толщу земли и дифракционных потерь энергии волны. Воздействие земли необходимо учитывать особенно в случае размещения антенных устройств в непосредственной близости к ее поверхности, что обычно имеет место в передвижных радиоустройствах. Потери энергии при излучении здесь настолько значительны, что в ряде случаев определяют выбор важнейших параметров радиоустройства в целом. В связи с этим возникает необходимость определения поля элементарного диполя в точке, расположенной на поверхности земли. При решении такой задачи необходимо знать электрические параметры (диэлектрическую проницаемость и удельную проводимость) земной поверхности.

В случае идеально проводящей земли, решение получается по методу зеркальных изображений. В этом случае токи в земле существуют только в виде поверхностных токов проводимости. Граничные условия у поверхности земли, требующие равенства нулю тангенциальной составляющей вектора **E** и нормальной составляющей вектора **H**, удовлетворяются, если системы токов земли заменить зеркальным изображением токов вибратора. В случае горизонтального вибратора ток в зеркальном изображении имеет амплитуду, равную амплитуде тока в вибраторе, но сдвинутую относительно него по фазе на π . В случае вертикального вибратора ток в зеркальном изображении равен

току вибратора по амплитуде и по фазе. Поскольку поля вибратора и его зеркального изображения удовлетворяют граничному условию на поверхности раздела воздух – земля, можно утверждать, что поле, создаваемое проходящими по поверхности токами, такое же, как и поле, создаваемое зеркальным изображением. Это дает основание при расчетах заменить земную поверхность зеркальным изображением вибратора.

Очевидно, что поле горизонтального вибратора, расположенного на идеально проводящей земле, равно нулю всюду, а поле вертикального вибратора в какой-либо точке над поверхностью земли удваивается по сравнению с полем в той же точке в свободном пространстве.

В общем случае земная поверхность не является идеальным проводником. Большая часть (71%) поверхности земного шара покрыта водой. С точки зрения электрических параметров соленую воду океанов, морей и некоторых озер следует отличать от пресной воды озер и рек. В то время как относительные диэлектрические проницаемости их равны ($\varepsilon_r = 80$), удельные проводимости соленой и пресной воды равны соответственно 4 и 10^{-3} (Ом•м).

Естественно, что поле в земле отлично от нуля, и, следовательно, часть электромагнитной энергии из атмосферы просачивается в землю. Поле над поверхностью земли, таким образом, уменьшается по сравнению с полем над идеально проводящей землей. При этом вектор напряженности $\mathbf{E}$ вертикального диполя должен иметь горизонтальную составляющую, поскольку только при ее существовании возможен поток электромагнитной энергии из атмосферы в землю. Следовательно, будут иметь место наклон вектора $\mathbf{E}$ вперед и уменьшение вертикальной составляющей напряженности поля E_z на реальной земле по сравнению с напряженностью $E_0 = E_{z0}$ на идеально проводящей земле. Очевидно, что, сравнивая эти составляющие, можно судить о степени влияния земли на распространение радиоволн. Для этого вертикальную составляющую представляют в виде

$$E_z = E_{z0} F, \quad (5.2)$$

где F – так называемая функция ослабления, модуль которой не может быть больше единицы по физическому смыслу. Единице он равен в случае идеально проводящей почвы. Следовательно, $|F| \leq 1$.

Функцию ослабления можно найти, решая строго задачу об излучении вертикального диполя, расположенного на границе раздела двух сред. При этом находят решения уравнений Максвелла для каждой среды, которые затем связывают между собой с помощью граничных условий. Впервые такую задачу решил в 1909 г. немецкий физик и математик А. Зоммерфельд. Однако применение строгих методов

исследования полей вблизи земной поверхности приводит к громоздким и малонаглядным результатам. В связи с этим возникла потребность в развитии приближенных методов анализа, позволяющих получать решения более простым и наглядным способом. Такой метод был создан и получил признание, поскольку оказался эффективным при решении ряда сложных задач. Сущность метода состоит в исключении полей в нижнем полупространстве ($z < 0$) путем применения к полям приближенных граничных условий при $z = 0$.

5.2. Зоны видимости радиолокационных станций

Пусть на расстоянии r от некоторой точки, в которой расположена радиолокационная станция (РЛС), находится цель. Плотность потока энергии, создаваемого антенной РЛС в месте расположения цели, определяется выражением

$$P_{\text{ц}} = P_{\Sigma} G / (4\pi r^2), \quad (5.3)$$

где G – коэффициент усиления (КУ) антенны, определяемый произведением КНД на коэффициент полезного действия, характеризующий потери мощности в проводниках и изоляторах антенны. КУ антенны в некотором направлении, характеризуемом углами θ и φ , так же, как и КНД, определяется соотношением $G(\theta, \varphi) = G_{\text{max}} F^2(\theta, \varphi)$, где G_{max} – максимальное значение КУ, а $F(\theta, \varphi)$ – нормированная диаграмма направленности антенны. Тогда выражение (5.3) принимает вид

$$P_{\text{ц}} = \frac{P_{\Sigma} G_{\text{max}} F^2}{4\pi r^2}, \quad (5.4)$$

$$F^2 = F^2(\theta, \varphi). \quad (5.5)$$

Цель обнаруживается благодаря тому, что часть мощности, рассеянная ею, попадает в приемную антенну РЛС. Для характеристики радиолокационных целей вводится понятие эффективной площади рассеяния. Тогда мощность, излучаемая целью в результате вторичного рассеяния, определяется произведением $P_{\text{нс}}$. Следовательно, в месте расположения приемной антенны РЛС цель создает поток энергии, плотность которого

$$P_{\text{пр}} = \frac{\sigma P_{\Sigma} G}{(4\pi r^2)^2}. \quad (5.6)$$

Мощность, принимаемая антенной, пропорциональна плотности потока энергии. Коэффициент пропорциональности характеризует эффективность антенны и, поскольку имеет размерность площади, называется эффективной площадью приемной антенны. Она обозначается через $A_{\text{эф}}$ и имеет следующую связь с КУ антенны:

$$A_{\text{эф}} = A_{\text{эф max}} F^2(\theta, \varphi). \quad (5.7)$$

Используя соотношения (5.6) и (5.7), получаем выражение для мощности, принимаемой антенной РЛС:

$$P_{\text{пр}} = \frac{\sigma P_{\Sigma} G_{\text{max}} A_{\text{эф max}}}{(4\pi r^2)^2} F^4(\theta, \varphi). \quad (5.8)$$

Эта формула называется уравнением радиолокации.

Важной характеристикой в радиолокации является максимальная дальность действия РЛС, определяемая из выражения (5.8), если положить $F(\theta, \varphi) = 1$ и $P_{\text{пр}} = P_{\text{пр min}}$ (где $P_{\text{пр min}}$ – минимальная мощность, принимаемая антенной РЛС), дальность действия РЛС, как это следует из формул (5.8), определяется выражением

$$r_{\text{max}}(\theta, \varphi) = r_{0\text{max}} F(\theta, \varphi), \quad (5.9)$$

представляющим собой уравнение поверхности в координатах r, θ, φ с началом координат в месте расположения антенны РЛС. Очевидно, что эта поверхность делит все пространство вокруг РЛС на область, в которой цель не видна, и область, в которой она видна (ее называют зоной видимости РЛС). Эта поверхность относится к семейству поверхностей одинаковых значений напряженности поля. Действительно, поскольку плотность потока энергии на расстоянии r , согласно соотношениям (5.2) и (5.5), то расстояние до точки на этой поверхности $E_0 = \text{const}$

$$r(\theta, \varphi) = \frac{\sqrt{60 P_{\Sigma} G_{\text{max}}}}{E_0} F(\theta, \varphi). \quad (5.10)$$

Тогда из выражений (5.9) и (5.10) можно определить E_0 , представляющее собой граничное значение напряженности поля на поверхности зоны видимости:

$$E_0 = E_{\text{гр}} = \sqrt{60 P_{\Sigma} G_{\text{max}}} / r_{0\text{max}}.$$

При учете влияния земли в уравнении радиолокации (5.8) необходимо вместо диаграммы направленности антенны в свободном пространстве $F(\theta, \varphi)$ брать произведение $F(\theta, \varphi)\Phi(\theta, \varphi)$, где $\Phi(\theta, \varphi)$ – интерференционный множитель. Например, если диаграмма направленности антенны симметрична относительно максимума излучения, направленного параллельно поверхности земли, то дальность действия РЛС определяется соотношением $r_{\text{max}}(\theta) = r_{0\text{max}} F(\theta)\Phi(\theta)$, где $\Phi(\theta)$ определяется выражением (5.3). Пользуясь этим соотношением, нетрудно построить диаграмму дальности РЛС.

ЛЕКЦИЯ 6. Двумерные преобразования Фурье и их свойства

В импульсной технике при разложении непериодической временной функции $f(t)$ в интеграл Фурье применяется формула

$$\Phi = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt,$$

где $\Phi(t)$ – ее комплексный частотный спектр.

Для случая функции двух переменных $g(x, y)$ ее образ, или пространственно-частотный спектр, представляется так:

$$\Phi\{g(x, y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) e^{-j(\omega_x x + \omega_y y)} dx dy = G(\omega_x, \omega_y), \quad (6.1)$$

где $\omega_x = 2\pi f_x$ и $\omega_y = 2\pi f_y$ – пространственные частоты. Обычно в (6.1) вместо ω_x и ω_y используют линейные пространственные частоты и тогда (6.1) принимает вид

$$G(f_x, f_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) e^{-j2\pi(f_x x + f_y y)} dx dy. \quad (6.2)$$

Для существования прямого преобразования Фурье комплексная функция $g(x, y)$ должна удовлетворять следующим условиям Дирихле:

- 1) быть абсолютно интегрируемой на бесконечной плоскости (x, y) ;
- 2) иметь конечное число разрывов и конечное число минимумов и максимумов на любом прямоугольнике конечных размеров этой плоскости;
- 3) не должна иметь разрывов второго рода.

Функции, не удовлетворяющие этим условиям, определяются как предел последовательности функций, удовлетворяющих им. Тогда каждая функция последовательности имеет свой Фурье-образ. Получаем последовательность Фурье-образов, предел которой назовем обобщенным Фурье-образом исходной функции, не удовлетворяющей условиям Дирихле. С обобщенными Фурье-образами можно оперировать, как с обычными. Существуют специальные таблицы Фурье-образов двумерных функций. Например, δ – функция Дирака.

Выражение (6.2) представляет собой разложение двумерной функции $g(x, y)$ как линейной комбинации элементарных функции вида

$$e^{-j2\pi(f_x x + f_y y)}.$$

Из определения прямого двумерного преобразования Фурье (6.2) вытекает несколько математических следствий, полезных при нахождении Фурье-образов функции двух переменных (аналогичные

следствия имеются для случая одномерного преобразования при разложении в спектр временной функции).

Частным случаем преобразования Фурье в полярных координатах для функции, обладающих круговой симметрией, $g(\rho, \varphi) = g_\rho(\rho)$, является преобразование Фурье–Бесселя

$$\Phi\{g_\rho(\rho)\} = 2\pi \int_0^\infty \rho g_\rho(\rho) I_0(2\pi\rho r) dr,$$

где $I_0(...)$ – функция Бесселя. Обратное преобразование Фурье–Бесселя имеет вид

$$g_\rho(\rho) = 2\pi \int_0^\infty r G_0(r) I_0(2\pi\rho r) dr.$$

Для преобразования Фурье–Бесселя применимы те же теоремы, что и для Фурье-преобразования.

Подобно таблицам оригиналов и изображений в операционном исчислении существуют таблицы преобразований Фурье для функции, удовлетворяющих условиям Дирихле.

Пространственно-частотное задание комплексных двумерных функции иногда значительно упрощает математический расчет систем формирования волновых фронтов, так как преобразования ведутся над элементарными функциями вида

$$e^{-j2\pi(f_x x + f_y y)},$$

а сама функция, описывающая волновой процесс, является как бы весовым множителем разложения (6.2).

6.1. Физическая интерпретация двумерных преобразований Фурье в теории волновых процессов

Исследуем обратное

$$f(t) = \int_{-\infty}^{\infty} \Phi(f) e^{j2\pi f t} dt$$

и прямое

$$\Phi(t) = \int_{-\infty}^{\infty} f(t) e^{j2\pi f t} dt$$

одномерные преобразования Фурье. Первое из преобразований можно рассматривать как представление функции $f(t)$ линейной комбинацией бесконечного набора функции вида $e^{j2\pi f t}$ с весовым комплексным множителем $\Phi(\omega)$.

Аналогично двумерное обратное преобразование Фурье можно рассматривать как линейную комбинацию элементарных функции вида $e^{j2\pi(f_x x + f_y y)}$ с комплексными весовыми множителями $G(f_x, f_y)$.

Фазовый множитель периодически повторяется через целое число 2π :

$$2\pi(f_x x + f_y y) = n \cdot 2\pi, \quad n = (1, 2, 3, \dots),$$

откуда

$$y = \frac{n}{f_y} - \frac{f_x}{f_y} x.$$

Получаем уравнение линии, вдоль которой фазовый множитель экспоненты постоянен.

Таким образом, любой паре пространственных частот в плоскости (f_x, f_y) в пространственной плоскости (x, y) соответствует прямая линия, направленная под углом $\varphi = \arctan f_x / f_y$ к оси x и отсекающая по оси f_y расстояние n .

В курсе «Теория антенн» изучается процесс излучения электромагнитных волн и их формирования на различных расстояниях от излучателя. Известно, что в дальней зоне векторы $\mathbf{E}$ и $\mathbf{H}$ распространяющейся волны взаимно перпендикулярны и перпендикулярны к направлению распространения. Такие волны называются плоскими. Если рассматривать процесс существования поля в любом однородном пространстве, анализ уравнений Максвелла с соответствующими граничными условиями приводит к уравнению плоской волны, распространяющейся вдоль оси z :

$$E = E_0 \cos(\omega t - kz), \quad H = H_0 \cos(\omega t - kz), \quad (6.3)$$

где $E_0, H_0 = |E_0|/|z_0|$ – амплитуды волны;

$$k = \frac{2\pi}{\lambda} \text{ – волновое число;}$$

$$\omega = 2\pi f \text{ – круговая частота;}$$

$$|z_0| \text{ – волновое сопротивление среды;}$$

z – направление распространения.

При практическом применении уравнения плоской электромагнитной волны направление распространения последней не всегда совпадает с одной из координатных осей выбранной системы координат. Поэтому удобнее пользоваться уравнением плоской волны, распространяющейся в произвольном направлении.

Пусть плоская волна движется в направлении, составляющем с осями x, y, z углы α, β, γ соответственно. Уравнение плоской волны принимает вид

$$E = E_0 e^{-j \frac{2\pi}{\lambda} (x \cos \alpha + y \cos \beta + z \cos \gamma)}. \quad (6.4)$$

Если функция $E(x, y, z)$ удовлетворяет условиям Дирихле, что почти всегда имеет место в реальных физических задачах, то в плоскости $z = 0$ поле этой волны $E(x, y, z=0)$ можно представить в виде обрат-

ного Фурье-образа частотного спектра $E(f_x, f_y)$ следующим образом:

$$E(x, y, z = 0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(f_x, f_y) e^{j2\pi(f_x x + f_y y)} df_x df_y. \quad (6.5)$$

Сопоставляя экспоненты в выражениях (6.4) и (6.5), следует заметить, что функцию $e^{j2\pi(f_x x + f_y y)}$ можно рассматривать в плоскости $z = 0$ как плоскую волну, распространяющуюся в пространстве с направляющими косинусами:

$$\cos \alpha = \lambda f_x, \cos \beta = \lambda f_y, \cos \gamma = \sqrt{1 - \lambda^2(f_x^2 + f_y^2)}.$$

С учетом этого Фурье-образ плоской электромагнитной волны в плоскости $z = 0$

$$E\left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda}\right) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(x, y, z = 0) e^{-j2\pi\left(\frac{\cos \alpha}{\lambda}x + \frac{\cos \beta}{\lambda}y\right)} dx dy \quad (6.6)$$

есть ее угловой спектр, физический смысл которого очевиден из равенства (6.5).

Плоская электромагнитная волна, распространяющаяся в пространстве под углами α, β, γ к осям x, y, z , в плоскости ($z = 0$) эквивалентна суперпозиции бесконечного числа плоских электромагнитных волн с комплексными амплитудами, распространяющимися под углами $\gamma = \arccos \sqrt{1 - \lambda^2(f_x^2 + f_y^2)}$ к оси z .

Таким образом, совокупность пары пространственных частот f_x, f_y указывает направление распространения в пространстве элементарной плоской волны, т.е. третью пространственную частоту. Пространственно-частотный спектр электромагнитной волны зависит от длины волны излучения: чем больше длина волны, тем уже ее частотный спектр. Последнее соотношение позволяет установить ширину спектра волнового фронта $E(x, y)$ в плоскости ($z = 0$).

6.2. Пространственно-частотный анализ волновых процессов

В радиоэлектронных схемах входной и выходной сигналы представляют собой действительные временные функции (напряжения или токи). В системах, формирующих волновой фронт, входные и выходные функции можно рассматривать как действительные (интенсивность), а также как комплексные (поле) функции двух переменных. В системе (радиоэлектронной или оптической) определенный входной сигнал вызывает на выходе определенный выходной сигнал. Однако могут быть случаи, когда входные сигналы не дают сигналов на выходе. В отдельных системах независимо от вида входного сигнала на выходе появляется один и тот же сигнал.

Пусть на входе системы имеется сигнал $g_1(x_1, y_1)$, а на выходе $g_2(x_2, y_2)$. Тогда можно ввести некоторый математический оператор, посредством которого из входного сигнала получается выходной:

$$M\{g_1(x_1, y_1)\} = g_2(x_2, y_2). \quad (6.7)$$

При нахождении частотных спектров временных функций применяется понятие δ -функции Дирака:

$$\delta(t) = \begin{cases} \infty, & t = 0 \\ 0, & t \neq 0 \end{cases}.$$

Определим понятие системы математически, так как физический смысл этого понятий известен из теории линейных цепей (интеграл Дюамеля).

Будем называть систему линейной, если для всех входных функций $\varphi(x, y), \psi(x, y), f(x, y) \dots$ и для всех комплексных постоянных a, b, c выполняется свойство суперпозиции:

$$\begin{aligned} M\{a\varphi(x, y) + b\psi(x, y) + cf(x, y)\} = \\ = aM\{\varphi(x, y)\} + bM\{\psi(x, y)\} + cM\{f(x, y)\}. \end{aligned} \quad (6.8)$$

На оснований этого свойства и определения (6.7) входной сигнал линейной системы можно представить как суперпозицию некоторых элементарных сигналов, а ее отклик на весь входной сигнал – как суперпозицию откликов на каждый из элементарных входных сигналов. Поэтому при описаний линейных систем важно найти удобную форму элементарных входных сигналов.

Оператор M можно ввести под знак интеграла и использовать свойство (6.8):

$$g_2(x_2, y_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(\xi, \zeta) M\{\delta(x_1 - \xi, y_1 - \zeta)\} d\xi d\zeta. \quad (6.9)$$

Введем обозначение:

$$h(x_2, y_2; \xi, \zeta) = M\{\delta(x_1 - \xi, y_1 - \zeta)\}. \quad (6.10)$$

Эту функцию интерпретируют как импульсный отклик систем в точке с координатами (x_1, y_2) на элементарный сигнал в виде двумерной δ -функции с координатами ξ, ζ . При описании временных функций аналогично вводилось понятие переходной характеристики цепи. Согласно (6.10), равенство (6.9) переписывается в виде

$$g_2(x_2, y_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(\xi, \zeta) h(x_2, y_2; \xi, \zeta) d\xi d\zeta. \quad (6.11)$$

Последнее соотношение называется двумерным суперпозиционным интегралом и означает, что можно найти выходной сигнал системы, если известен ее импульсный отклик на некоторые элементарные воздействия, по которым можно разложить входной сигнал. Для систем, формирующих волновой фронт, физический смысл этого интеграла

можно интерпретировать так: действие формирующих волновой фронт устройств может быть описано полностью, если определены в этой системе изображения точечных источников исходного волнового фронта.

При описании линейных электрических цепей используется понятие инвариантной линейной цепи. Это такая цепь, отклик которой в момент t на воздействие, поступившее на вход цепи в момент τ , зависит только от промежутка времени $t-\tau$. В линейных системах, формирующих волновой фронт, вводят понятие пространственной инвариантности (изопланарности), т.е. для таких систем $h(x_2, y_2; \xi, \zeta) = h(x_2 - \xi, y_2 - \zeta)$, или, интерпретируя это соотношение, можно сказать: изображение точки формируемого волнового фронта меняет только свое положение при прохождении через систему. Двумерный суперпозиционный интеграл для изопланарных систем принимает вид

$$g_2(x_2, y_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(\xi, \zeta) h(x_2 - \xi, y_2 - \zeta) d\xi d\zeta. \quad (6.12)$$

Последний интеграл называют двумерной сверткой функций $g_1(\xi, \zeta)$ и импульсного отклика системы h , его обозначают $g_2 = g_1 \otimes h$. Применяя к левой и правой частям равенства (6.12) преобразование Фурье и используя теорему свертки, видим, что

$$G_2(f_x, f_y) = G_1(f_x, f_y) H(f_x, f_y), \quad (6.13)$$

где $H(f_x, f_y)$ – Фурье-образ импульсного отклика системы.

Равенство (6.13) означает, что для нахождения сформированного системой волнового фронта нужно определить Фурье-образ входного фронта и перемножить его с Фурье-образом импульсного отклика системы, после чего найти обратное преобразование Фурье полученного произведения.

Рассмотрим конкретные примеры пространственно-частотного анализа прохождения электромагнитной волны сквозь некоторые системы. В первом примере в качестве системы возьмем определенный участок свободного пространства и проследим за изменениями, происходящими с распространяющимся в нем волновым фронтом. В качестве второго примера разберем пространственно-частотный анализ прохождения электромагнитного волнового фронта сквозь слой диэлектрика, имеющего плоскую или цилиндрическую форму.

6.3. Свободное пространство

Пусть есть угловой спектр волны $E(x, y, z)$ в произвольной плоскости $z = \text{const}$. Обратное Фурье-преобразование должно иметь вид

$$E(x, y, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E_z\left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda}, z\right) e^{j2\pi\left(\frac{\cos \alpha}{\lambda}x + \frac{\cos \beta}{\lambda}y\right)} d\left(\frac{\cos \alpha}{\lambda}\right) d\left(\frac{\cos \beta}{\lambda}\right), \quad (6.14)$$

а уравнение плоской волны, распространяющейся к оси z под углом

$\arccos \sqrt{1 - (\cos^2 \alpha + \cos^2 \beta)}$, имеет вид:

$$\mathbf{E}_z \left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda}, z \right) = \mathbf{E}_z \left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda}, z \right) e^{jk \sqrt{1 - (\cos^2 \alpha + \cos^2 \beta)} z}. \quad (6.15)$$

Заметим, что при $\cos^2 \alpha + \cos^2 \beta > 1$ показатель экспоненты в выражении (6.15) есть действительное отрицательное число. С увеличением расстояния z плоская волна, описываемая уравнением (6.15), быстро затухает. Очевидно, что при $\cos^2 \alpha + \cos^2 \beta = 1$.

$\mathbf{E}$ не зависит от z . Это значит, что выражение (6.15) описывает плоские волны, распространяющиеся перпендикулярно к оси z .

Распространение плоской волны от плоскости $z = 0$ до плоскости z рассматривается как процесс воздействия двумерной функции $E(x, y, z = 0)$ на систему, формирующую волновой фронт $E(x, y, z)$. Распространение электромагнитной волны в однородной среде является линейным процессом, что следует из линейности уравнения Максвелла (волнового уравнения) согласно принципу Гюйгенса.

Таким образом, процесс распространения электромагнитной волны в свободном пространстве можно рассматривать как действие линейного фильтра пространственных частот с конечной полосой пропускания по ним, которая в области частот равна кругу радиуса λ . Модуль коэффициента передачи этого фильтра равен единице, а фазовые сдвиги зависят от пространственной частоты и стремятся к нулю при ее уменьшении.

Резюмируя, можно дать простое физическое истолкование процессу распространения углового спектра плоской электромагнитной волны от плоскости $z = 0$ до произвольного $z = \text{const}$. Распространение волны на расстояние z приводит к изменению фазы ее пространственных составляющих за счет различия в длине пути.

Пространственно-частотный анализ распространения электромагнитных волн в свободном пространстве позволяет в приближениях геометрической оптики более просто находить математические выражения, описывающие влияние различных ограничивающих отверстий в непрозрачных экранах на структуру распространяющейся волны.

Пусть в металлическом плоском экране прорезано квадратное отверстие размером $1a$ и $2b$ по осям x и y соответственно. На экран падает плоская электромагнитная волна $E(x, y, z = 0)$. Поле волны за экраном можно представить в виде

$$E_{\text{пр}}(x, y, z = 0) = E(x, y, z = 0)T(x, y), \quad (6.16)$$

где $T(x, y)$ – коэффициент прохождения отверстия. В данном случае

$$T(x, y) = \begin{cases} 1, & |x| \leq a, |y| \leq b \\ 0, & |x| > a, |y| > b \end{cases}. \quad (6.17)$$

Применяя к левой и правой частям равенства (6.16) преобразование Фурье, учитывая теорему свертки, имеем угловой спектр падающей волны, характеризующий само отверстие и определяемый в данном случае геометрическими размерами:

$$\tau = \left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda} \right) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} T(x, y) e^{-j2\pi \left(\frac{\cos \alpha}{\lambda} x + \frac{\cos \beta}{\lambda} y \right)} dx dy. \quad (6.18)$$

Ввиду (6.17) выражение можно преобразовать:

$$\tau = \left(\frac{\cos \alpha}{\lambda}, \frac{\cos \beta}{\lambda} \right) = \int_{-a}^a e^{-j2\pi \frac{\cos \alpha}{\lambda} x} dx \int_{-b}^b e^{-j2\pi \frac{\cos \beta}{\lambda} y} dy. \quad (6.19)$$

Таким образом, угловой спектр прошедшей отверстие волны определяется сверткой угловых спектров падающей волны и отверстия (6.18). Это означает, что пространственное ограничение распространяющегося волнового фронта приводит к расширению его пространственно-частотного спектра, которое обратно пропорционально размерам отверстия, что следует из интеграла (6.19), и соответствует физическим представлениям о диаграмме направленности излучения поверхностных антенн СВЧ-диапазона. Радиотехническим аналогом этого примера является расширение частотного спектра временного импульсного сигнала с уменьшением длительности импульса.

ЛЕКЦИЯ 7. Понятие теории выборки

При обработке временных сигналов и регистрации волновых фронтов часто нет возможности производить непрерывную регистрацию как во времени, так и в пространстве. В связи с этим возникает задача о степени достоверности, с которой по зарегистрированным дискретным значениям исследуемой функции можно определить ее аналитическую форму. В статистической радиофизике такие задачи решаются посредством теории выборки, одномерный вариант которой наиболее полно разработан советским ученым, академиком В.А. Котельниковом. Задачи теории выборки – определение законов расположения дискретных точек отсчета исследуемого волнового процесса, при которых последний будет восстановлен с наибольшей достоверностью.

Рассмотрим волновой процесс, описываемый комплексной функцией $f(x, y)$, дискретные значения которой заданы на плоскости xOy , причем точки отсчетов расположены по некоторому закону.

Если функция $f(x, y)$ задана аналитически или известен ее спектр $F(f_x, f_y)$, все необходимые математические операции проводятся посредством Фурье-преобразований и их свойств. В данном случае аналитических выражений для прямого и обратного преобразований Фурье не существует, поэтому потребовалась разработка другого исследовательского аппарата.

Очевидно, что чем ближе точки отсчета расположены друг к другу, тем с большим приближением дискретные значения функции $f_{\text{дискр}}(x, y)$ описывают аналитическую зависимость $f(x, y)$. Под $f_{\text{дискр}}(x, y)$ будем понимать численное задание функции $f(x, y)$ в точках отсчета.

Если в каждой точке отсчета взять преобразование Фурье от соответствующего значения функции, в плоскости пространственно-частотных координат f_x, f_y получим частотный спектр дискретного набора $f_{\text{дискр}}(x, y)$, который обозначим: $F_{\text{дискр}}(x, y)$. В теории выборки показывается, что при некоторых специальных законах расположения точек отсчетов $f_{\text{дискр}}(x, y)$ в плоскости (x, y) для определенного вида функций $f(x, y)$ их спектр $F_{\text{дискр}}(x, y)$ представляет собой дискретный набор спектров $F(f_x, f_y)$ аналитической функции $f(x, y)$. Если эти спектры не перекрываются (случай функций с ограниченным спектром), путем фильтрации всех составляющих пространственно-частотного спектра $F_{\text{дискр}}(f_x, f_y)$ кроме одной области находим

Фурье-образ аппроксимируемой функции $f(x, y)$. Вычисляя ее обратный Фурье-образ, восстанавливаем функцию $f(x, y)$ по ее дискретным значениям. Таким образом возможно точное восстановление функций лишь с ограниченным спектром.

Поясним вышеизложенное на эксперименте с прохождением волнового фронта электромагнитной волны через фокусирующую систему. Волновой фронт $f(x, y)$, помещенный в переднюю фокальную плоскость фокусирующей системы, преобразуется и во второй фокальной плоскости, последняя математически описывается преобразованием Фурье от исходного вида $f(x, y)$. Вторая фокусирующая система осуществляет обратное преобразование Фурье, т.е. восстанавливает исходный волновой фронт $f(x, y)$. При дискретном задании преобразуемого волнового фронта, если функция $f(x, y)$ имеет ограниченный спектр, путем установки в плоскости пространственных частот специального фильтра (экрана с отверстием нужных размеров) производится вышеуказанная пространственная фильтрация волнового фронта $F_{\text{дискр}}(f_x, f_y)$, которая позволяет выделить из него пространственный спектр $F(f_x, f_y)$. Обратное Фурье-преобразование, выполняемое фокусирующей системой L_2 , дает возможность восстановить точное представление волнового фронта $f(x, y)$.

Задачей теории выборки является изучение оптимальных законов расположения точек отсчета при дискретизации волновых фронтов, позволяющих восстановить аналитическую форму исследуемого волнового процесса максимально возможной точностью. Оптические системы обработки предварительно преобразованных волновых фронтов любого диапазона являются наиболее перспективными.

7.1. Расчет систем, формирующие и преобразующие волновые фронты

Теория прохождения волновых фронтов сквозь различные системы до недавнего времени основывалась на математическом аппарате теории дифракции. Трактовка последнего на математическом языке двумерных Фурье-преобразований позволила широко использовать их для расчета оптических систем построения изображений и обработки информации.

7.2. Пространственно-частотный подход к дифракционным явлениям

При обычной трактовке связь между распределением амплитуды и фазы электромагнитной волны в исходной и произвольной плоскостях пространства в скалярном приближении описывается интегра-

лом Френеля–Кирхгофа. Начало скалярной теории дифракции положено итальянским ученым Гримальди (1665), который экспериментально установил отсутствие четкой границы между тенью и освещенным участком плоскости наблюдения, расположенной непосредственно за отверстием светопроницаемого экрана. На основе господствовавшей тогда корпускулярной теории света эти явления не могли быть объяснены, поэтому предпринимались попытки трактовки отклонения световых лучей от прямолинейного закона распространения. Развитие волновой теории света на основе принципа Гюйгенса для построения сложного волнового фронта из элементарных волновых фронтов точечных источников с учетом интерференции Юнга привело к созданию первого математического аппарата теории дифракции – интегралу Гельмгольца–Кирхгофа:

$$g_2(x, y) = \frac{1}{4\pi} \iint_S \left[\frac{\partial g_1(\zeta, \eta)}{\partial n} \left(\frac{e^{jkr}}{r} \right) - g_1(\zeta, \eta) \frac{\partial}{\partial n} \left(\frac{e^{jkr}}{r} \right) \right] d\zeta d\eta, \quad (7.1)$$

где $g_2(x, y)$ – поле в произвольной точке наблюдения; S – поверхность интегрирования, на которой задано исходное распределение поля $g_1(x, y)$; $\frac{\partial}{\partial n}$ – производная по направлению внешней нормали к поверхности S ; k – волновое число; $\mathbf{r}$ – вектор направления между произвольными точками интегрируемой поверхности и точкой наблюдения. При выводе интеграла (7.1) применяется теорема Грина, которая математически строго обоснована для замкнутых объемов, ограниченных поверхностью S . Следовательно, использование интеграла возможно лишь для точек наблюдения x, y , в окрестностях которых исходный волновой фронт $g_1(x, y)$ задан в полном телесном угле 4π . На практике же, как правило, известен волновой фронт лишь в ограниченном телесном угле, характеризуемом элементом поверхности Σ . В таком случае интеграл (7.1) разбивается на два, в одном из которых интегрирование проводится по поверхности Σ , в другом – по поверхности S_1 .

$$\frac{\partial}{\partial n} \left(\frac{e^{jkr}}{r} \right) = \left(j - \frac{1}{r} \right) \frac{e^{jkr}}{r} \cos(\hat{\mathbf{n}}, \hat{\mathbf{r}}) = jk \frac{e^{jkr}}{r} \cos(\hat{\mathbf{n}}, \mathbf{r}). \quad (7.2)$$

Для обоснования правомерности применения к интегралу (7.1) упрощения (7.2) необходимо устремить r к бесконечности, что растянет размеры поверхности Σ за пределы задания исходного волнового фронта $g_1(\zeta, \eta)$. Кирхгоф ввел упрощение, предполагающее на этой дополнительной поверхности интегрирования обращение в нуль $g_1(\zeta, \eta)$ и ее первой производной, известное как граничное условие Кирхгофа.

Вклад же в интеграл (7.1) интегрирования по сферической поверхности S_1 можно устранить путем дополнительного предположения. Наличие исходного волнового фронта на поверхности Σ приведет в силу его распространения (волновой процесс распространяется со скоростью света в свободном пространстве) к существованию функции $g_1(\zeta, \eta)$ и на поверхности S_1 . Интеграл (7.1) по поверхности S_1 будет обращаться в нуль только в предположении, что радиус последней выбран настолько большим, что волновой фронт $g_1(\zeta, \eta)$ не достиг ее поверхности.

Согласно принципу Гюйгенса–Френеля, исходный волновой фронт может достичь поверхности S_1 за счет распространения элементарных возмущений в виде сферических волн от каждой точки исходного волнового фронта $g_1(\zeta, \eta)$. Можно показать, что конечная скорость распространения сферической волны приводит к тому, что возмущение $g_1(\zeta, \eta)$ не достигнет поверхности S_1 ранее, чем оно обратится на этой поверхности в нуль в силу выполнения условия Зоммерфельда. Следовательно, последнее условие в сочетании с граничным условием Кирхгофа позволяет вместо выражения (7.1) использовать следующий вид интеграла Гельмгольца–Кирхгофа:

$$g_2(x, y) = \frac{1}{4\pi} \iint_{\Sigma} \left[\frac{\partial g_1(\zeta, \eta)}{\partial n} \left(\frac{e^{jkr}}{r} \right) - g_1(\zeta, \eta) \frac{\partial}{\partial n} \left(\frac{e^{jkr}}{r} \right) \right] d\zeta d\eta \quad (7.3)$$

При практическом применении выражения (7.3) в аналитическом расчете волновых процессов появляется ошибка, так как невозможно каким-то образом вырезать из волнового фронта ограничивающим устройством его часть, не исказив краями устройства исходной структуры волны. Однако при размерах апертуры ограничивающего устройства в несколько длин волн расчетные результаты, как правило, соответствуют экспериментальным.

Наличие исходного волнового фронта $g_1(\xi, \eta)$ на поверхности Σ , в свою очередь, обусловлено приходом в произвольную точку интегрирования возмущения от какого-либо первичного источника. Пусть этот источник является точечным и расположен в точке с координатами x_0, y_0, z_0 , тогда зададим вектор, соединяющий первичный источник сферической волны единичной амплитуды и произвольную точку интегрируемой поверхности Σ . Применяя к уравнению (7.3) соотношение (7.2) при тех же предположениях относительно $r_0 = \sqrt{(\zeta - x_0)^2 + (\eta - y_0)^2 + z_0^2}$, получаем

$$g_2(x, y) = \frac{j}{2\lambda} \times \iint_{\Sigma} \frac{e^{jk} \sqrt{(\zeta - x_0)^2 + (\eta - y_0)^2 + z_0^2} + \sqrt{(\zeta - x_0)^2 + (\eta - y_0)^2 + z_0^2}}{\sqrt{(\zeta - x_0)^2 + (\eta - y_0)^2 + z_0^2} \sqrt{(\zeta - x_0)^2 + (\eta - y_0)^2 + z_0^2}} \times [\cos(n, r_0) - \cos(n, r)] d\zeta d\eta. \quad (7.4)$$

Соотношение (7.4) называется дифракционной формулой Френеля–Кирхгофа. Это соотношение можно записать следующим образом:

$$g_2(x, y) = \iint_{\Sigma} A(\zeta, \eta) \frac{e^{jk} \sqrt{(\zeta - x_0)^2 + (\eta - y_0)^2 + z_0^2}}{\sqrt{(\zeta - x_0)^2 + (\eta - y_0)^2 + z_0^2}} d\zeta d\eta. \quad (7.5)$$

Последнее соотношение является математической формулировкой принципа Гюйгенса и означает, что волновой фронт в произвольной, но фиксированной плоскости $z = \text{const}$ есть суперпозиция элементарных, сферических волн с комплексными амплитудами $A(\zeta, \eta)$, которые, в свою очередь, образованы элементарным точечным источником сферической волны в нашем предположении, единичной амплитуды.

Теория Френеля–Кирхгофа была развита в предположении пространства исходной сферической волны от элементарного источника только в одном направлении при выборе функции Грина в виде $\frac{e^{jkr}}{r}$ и одновременном обращении в нуль $g_1(\zeta, \eta)$ и ее первой производной. Зоммерфельд с физической точки зрения более строго подошел к этой задаче и ограничился требованием, чтобы $g_1(\zeta, \eta)|_{\Sigma} = 0$ либо $\frac{\partial g_1(\zeta, \eta)}{\partial n}|_{\Sigma} = 0$.

В предположении, которое было принято при выводе (7.5) последнее соотношение имеет вид

$$g_2(x, y) = \frac{1}{j\lambda} \iint_{\Sigma} \frac{e^{jk} \sqrt{(\zeta - x_0)^2 + (\eta - y_0)^2 + z_0^2} + \sqrt{(\zeta - x_0)^2 + (\eta - y_0)^2 + z_0^2}}{\sqrt{(\zeta - x_0)^2 + (\eta - y_0)^2 + z_0^2} \sqrt{(\zeta - x_0)^2 + (\eta - y_0)^2 + z_0^2}} \times \cos(n, r) d\zeta d\eta \quad (7.6)$$

и называется дифракционной формулой Рэлея–Зоммерфельда. Отличается оно от дифракционной формулы Френеля–Кирхгофа тем, что не учитывает наклонный характер прихода сферической волны в точку (ζ, η) от источника, расположенного в точке x_0, y_0, z_0 . В предположении, что интегрируемая поверхность возбуждается плоской вол-

ной, фазовый множитель e^{jkr_0} и множитель направления $\cos(n, r_0)$ выражениях (7.6) и (7.2) обращаются в единицу, и формулы Френеля–Кирхгофа и Рэлея–Зоммерфельда совпадают.

Таким образом, на основании соотношения (7.6) можно рассчитать волновой фронт в произвольной плоскости пространства, если он задан в некоторой плоскости ($z = 0$ в данном случае, что не меняет общности результатов). Частные применения общего соотношения (7.6) дают дифракционные формулы Френеля–Кирхгофа и Рэлея–Зоммерфельда. Выражение (7.6), являющееся упрощением интеграла Гельмгольца–Кирхгофа, более общее и поэтому будет использоваться нами в дальнейшем. Формула (7.6) и суперпозиционный интеграл аналогичны, если в (7.6) расширить пределы интегрирования до ∞ , что правомерно ввиду выполнения условия Кирхгофа. Функцию

$$\frac{1}{j\lambda} \frac{e^{jk} \sqrt{(\zeta - x_0)^2 + (\eta - y_0)^2 + z_0^2}}{\sqrt{(\zeta - x_0)^2 + (\eta - y_0)^2 + z_0^2}} \cos(n, r) = h(x, y; \zeta, \eta) \quad (7.7)$$

можно интерпретировать как импульсный отклик в точке $P(x, y, z)$ на воздействие из точки $A(\zeta, \eta, \xi = 0)$.

Формула (7.7) преобразуется так же, как и интеграл (7.2), т.е. при дополнительных ограничениях на расстояние r до точки наблюдения и на направление на точку наблюдения.

При дальнейшем изложении будем рассматривать волновые процессы, происходящие вблизи оптической оси линейных инвариантных систем, т.е. параксиальное приближение. В этом случае $\cos(n, r) = 1$, а в знаменателе вместо $r = \sqrt{(\zeta - x_0)^2 + (\eta - y_0)^2 + z_0^2}$ будем брать $r = z$, так как для амплитудного множителя их различие не существенно.

В приближении Френеля импульсный отклик имеет вид

$$h(x, y; \zeta, \eta) = \frac{e^{jkz}}{j\lambda z} e^{j\frac{k}{2z}[(x-\zeta)^2 + (y-\eta)^2]}. \quad (7.8)$$

Импульсный отклик свободного пространства

$$h(x, y; \zeta, \eta) = \frac{e^{jkz}}{j\lambda z} e^{j\frac{k}{2z}(x+y)^2} e^{-\frac{k}{z}(x\zeta + \eta y)}. \quad (7.9)$$

При этом поле в точке наблюдения определится как Фурье-образ распределения поля в дифрагирующей апертуре по тем же пространственным частотам. При подстановке импульсного отклика в суперпозиционный интеграл можно убедиться, что последний является интегралом свертки, поэтому

$$\Phi\{g_2(x, y)\} = \Phi\{g_1(\zeta, \eta)\} \Phi\{h(x - \zeta, y - \eta)\},$$

где

$$\Phi\{h(x - \varsigma, y - \eta)\} = H\left(\frac{x}{\lambda z}, \frac{y}{\lambda z}\right) e^{jkz} e^{-j\pi\left(\frac{x^2}{\lambda^2 z} + \frac{y^2}{\lambda^2 z}\right)} \quad (7.10)$$

есть передаточная функция свободного пространства в приближении Френеля. Выражение (7.10) является приближением Френеля применительно к полученной выше передаточной функции свободного пространства. Учитывая, что дифракция Фраунгофера является предельным случаем дифракции Френеля, выражение (7.10) справедливо и для импульсного отклика (7.9).

При вычислениях суперпозиционных интегралов в одинаковой степени используются импульсные отклики в приближениях Френеля и Фраунгофера (7.9).

Приведенные сведения из теории дифракции не претендуют на полноту, а лишь указывают на фундаментальность математического аппарата Фурье-преобразований, который присутствовал в разработанной ранее скалярной геометрической теории дифракции. Скалярная теория дифракции была существенно дополнена так называемой физической теорией дифракции, одной из разновидностей которой является метод краевых волн. Геометрическая теория дифракции, основанная на лучевой трактовке, также была существенно дополнена физической теорией дифракции, учитывающей различие в структуре волны, дифрагируемой (переизлучаемой) характерными деталями объектов (краями, изломами и другими структурными неоднородностями).

ЛЕКЦИЯ 8. Оптические системы обработки информации

8.1. Теория фокусирующей линзы

Рассмотрим следующую радиооптическую задачу. На фокусирующую линзу толщиной Δ падает волновой фронт $a(x, y)$. Требуется найти волновой фронт $b(x, y)$ на произвольном расстоянии B от линзы.

Волновой фронт непосредственно на входе и выходе линзы обозначим $a'(x, y)$ и $a''(x, y)$. Таким образом, наша общая задача разобьется на три независимые подзадачи:

- 1) найти $a'(x, y)$, зная $a(x, y)$;
- 2) найти $a''(x, y)$, зная $a'(x, y)$;
- 3) найти $b(x, y)$, зная $a''(x, y)$ как функцию $a(x, y)$.

В первой подзадаче из дифракционного интеграла Гельмгольца–Кирхгофа в приближении Френеля для параксиального случая, который, как правило, реализуется на практике в связи с незначительными размерами линз, имеем

$$a'(x, y) = \frac{e^{jkA}}{j\lambda A} \exp\left[j\frac{k}{2A}(x^2 + y^2)\right] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} a(\xi, \eta) \times \\ \times \exp\left[j\frac{k}{2A}(\xi^2 + \eta^2)\right] + \exp\left[-j \cdot 2\pi\left(\frac{x}{\lambda A}\xi + \frac{y}{\lambda A}\eta\right)\right] d\xi d\eta. \quad (8.1)$$

Для нахождения $a''(x, y)$ воспользуемся понятием комплексного коэффициента пропускания линзы $t(x, y) = |t(x, y)|e^{j\varphi(x, y)}$, где $|t(x, y)| \approx 1$ – его амплитуда

$$\varphi(x, y) = \frac{2\pi}{\lambda}nc(x, y) + \frac{2\pi}{\lambda}[C - c(x, y)] \equiv \frac{2\pi}{\lambda}[(n-1)c(x, y) + C] \quad (8.2)$$

есть фазовый множитель; n – показатель преломления; $c(x, y)$ – толщина линзы в произвольной точке (x, y) ; $C = \text{const}$.

Из геометрической оптики известно, что

$$c(x, y) = C - \frac{x^2 - y^2}{2} \left(\frac{1}{R_1} - \frac{1}{R_2} \right), \quad (8.3)$$

где R_1, R_2 – радиусы кривизны левой и правой частей линзы. Следовательно,

$$a''(x, y) = a'(x, y)t(x, y) \approx \\ \approx a'(x, y) \exp\left(j\frac{2\pi}{\lambda}C\right) \times \exp\left[j\frac{2\pi}{\lambda}(n-1)c(x, y)\right] = \\ = a'(x, y) \exp\left(j\frac{2\pi}{\lambda}C\right) \times \\ \times \exp\left\{j\frac{2\pi}{\lambda}(n-1)\left[C - \frac{x^2 - y^2}{2} \left(\frac{1}{R_1} - \frac{1}{R_2} \right)\right]\right\}. \quad (8.4)$$

Учитывая, что

$$\frac{1}{F} = (n-1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right),$$

из (8.4) получаем

$$a'(x, y) = a(x, y) \exp \left(j \frac{2\pi}{\lambda} nC \right) \exp \left[-j \frac{\pi}{\lambda F} (x^2 + y^2) \right]. \quad (8.5)$$

После подстановки (8.1) в (8.5) имеем

$$\begin{aligned} a''(x, y) = & \exp \left(j \frac{2\pi}{\lambda} nC \right) \exp \left[-j \frac{\pi}{\lambda F} (x^2 + y^2) \right] \frac{e^{jkA}}{j\lambda A} \times \\ & \times \exp \left[j \frac{k}{2A} (x^2 + y^2) \right] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} a(\xi, \eta) \exp \left[j \frac{k}{2A} (\xi^2 + \eta^2) \right] \times \\ & \times \exp \left[-j \cdot 2\pi \left(\frac{x}{\lambda A} \xi + \frac{y}{\lambda A} \eta \right) \right] d\xi d\eta. \end{aligned} \quad (8.6)$$

Далее опять воспользуемся интегралом (8.4), который в нашем случае будет иметь вид

$$\begin{aligned} b(x, y) = & \frac{e^{jkB}}{j\lambda B} \exp \left[j \frac{k}{2B} (x^2 + y^2) \right] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} a''(\xi, \eta) \exp \left[j \frac{k}{2B} (\xi^2 + \eta^2) \right] \times \\ & \times \exp \left[-j \cdot 2\pi \left(\frac{x}{\lambda B} \xi + \frac{y}{\lambda B} \eta \right) \right] d\xi d\eta. \end{aligned} \quad (8.7)$$

Используя соотношения (8.1), (8.4) и (8.7), в двух частных случаях, когда $A \rightarrow \infty$ (на линзу падает плоская волна) или $A = F$ (исходный волновой фронт находится в передней фокальной плоскости), для $B = F$ (волновой фронт в ее задней фокальной плоскости) получаем выражение с точностью до множителей, стоящих перед интегралами (8.1), (8.4), (8.7):

$$\begin{aligned} b_F(x_F, y_F) = & \frac{\exp[jk/2F(x_F^2 + y_F^2)]}{j\lambda F} \times \\ & \times \exp \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} a(x, y) \exp \left[-j \frac{2\pi}{\lambda F} (xX_F + yY_F) \right] dx dy, \end{aligned} \quad (8.8)$$

где индекс F у величин x, y означает принадлежность этих величин к задней фокальной плоскости линзы.

Таким образом, фронт $b(x, y)$ в этих случаях является преобразованием Фурье от исходного волнового фронта по пространственным частотам

$$f_x = \frac{x_F}{\lambda F}; \quad f_y = \frac{y_F}{\lambda F}, \quad (8.9)$$

который можно обозначить

$$A\left(\frac{x_F}{\lambda F}, \frac{y_F}{\lambda F}\right) = b_F(x_F, y_F).$$

Это замечательное свойство фокусирующей линзы дает возможность выполнять над комплексной двумерной функцией такую сложную математическую операцию, как двумерное преобразование Фурье.

Пользуясь указанным дифракционным интегралом и определениями прямого и обратного преобразований Фурье, нетрудно из соотношений (8.1), (8.6), (8.7) получить вывод, что для произвольного расстояния A существует плоскость за линзой (некоторое расстояние B), в которой волновой фронт есть Фурье-образ исходного $a(x, y)$.

Аналогично доказывается, что при расположении исходного волнового фронта между линзой и задней фокальной плоскостью линзы в последней также формируется Фурье-образ первого, но с масштабным изменением, определяемым отношением F/A , которое можно регулировать, изменяя A .

При изложении теории линзы мы не накладывали ограничений на форму и размеры входной апертуры, предполагая ее бесконечной. На практике в соотношение (8.8) под знак интеграла следует вводить еще один множитель, задающий форму и размеры апертуры $t(x, y)$ с конечной плоскостью интегрирования, на которой он обращается в нуль. Это приводит к некоторым отклонениям физического от точного математического преобразования Фурье, выполняемого линзой. Однако в большинстве практических приложений эти отклонения несущественны или же их можно учесть.

Возвратимся к исходной задаче в радиооптической трактовке. Найденная нами зависимость (8.7), описывающая распространение волнового фронта через линзу, эквивалентна соотношению (8.10) для линейных инвариантных систем

$$b(x, y) \int_{-\infty}^{\infty} \int h(x_B, y_B; x_A, y_A) a(x_A, y_A) dx_A dy_A, \quad (8.10)$$

где индексы A и B означают принадлежность координат x, y к произвольной точке соответствующей области.

Подставляя (8.1) в (8.4), а (8.4) в (8.7) и сопоставляя результат с (8.10), получаем в приближении Френеля для исходного волнового фронта $a(x, y)$ в виде δ -функции с координатами x_A, y_A отклик участка пространства с линзой:

$$\begin{aligned}
h(x_B, y_B; x_A, y_A) = & \frac{1}{\lambda^2 AB} \exp \left[j \frac{k}{2B} (x_B^2 + y_B^2) \right] \times \\
& \exp \left[j \frac{k}{2B} (x_A^2 + y_A^2) \right] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp \left[j \frac{k}{2} \left(\frac{1}{A} + \frac{1}{B} - \frac{1}{F} \right) (x^2 + y^2) \right] \times \\
& \times \exp \left\{ -jk \left[\left(\frac{x_A}{A} + \frac{x_B}{B} \right) x + \left(\frac{y_A}{A} + \frac{y_B}{B} \right) y \right] \right\} dx dy.
\end{aligned} \quad (8.11)$$

В точке, удовлетворяющей формуле линзы $(1/A) + (1/B) = 1/F$, пренебрегая для случаев регистрации интенсивности волнового фронта экспонентами, стоящими перед интегралом (8.11), получаем

$$\begin{aligned}
h(x_B, y_B; x_A, y_A) = & \frac{1}{\lambda^2 AB} \times \\
& \times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp \left\{ -j \cdot 2\pi \left[\left(\frac{x_A}{\lambda A} + \frac{x_B}{B} \right) x + \left(\frac{y_A}{\lambda A} + \frac{y_B}{B} \right) y \right] \right\} dx dy.
\end{aligned} \quad (8.12)$$

Введя в рассмотрение увеличение линзы $M = B/A$, (8.12) упростим:

$$\begin{aligned}
h(x_B, y_B; x_A, y_A) = & \frac{1}{\lambda^2 AB} \times \\
& \times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp \left\{ -j \frac{2\pi}{\lambda B} [(x_B + Mx_A)x + (y_B + My_A)y] \right\} dx dy.
\end{aligned} \quad (8.13)$$

Подстановкой (8.13) в (8.10) заканчивается решение задачи о прохождении волнового фронта в простейшей РОС – участке свободного пространства, содержащего линзу. Численная реализация этих расчетов производится путем дискретизации указанных интегральных выражений и использования программ БПФ для пространственных частот

$$f_x = \frac{(x_B + Mx_A)}{\lambda B}; \quad f_y = \frac{(y_B + My_A)}{B\lambda}.$$

Поскольку фокусирующая линза является линейной инвариантной системой, то в соответствии с (8.10) и теоремой свертки (8.12) на практике удобно сначала найти отклик согласно (8.13). Затем, находя Фурье-образы отклика (передаточную характеристику данного участка пространства с линзой) и исходного волнового фронта $a(x, y)$, можно получить их произведение $A(f_x, f_y) \cdot H(f_x, f_y)$. Обратное Фурье-преобразование последнего является искомым волновым фронтом $b(x, y)$ в плоскости ξ, η в соответствующем масштабе, определяемом увеличением M . Легко убедиться, что для $A = B = F$ это увеличение равно единице.

ЛЕКЦИЯ 9. Расчет прохождения электромагнитных волн в простейших фокусирующих системах

Рассмотрим фокусирующую систему в виде сферической линзы, образованной двумя тонкими долями сферических поверхностей с одинаковыми площадями оснований и радиусами кривизны r_1 и r_2 . Площадь оснований этих поверхностей, т.е. размер апертуры, будем считать значительно больше квадрата длины волны излучения, проходящего сквозь линзу перпендикулярно к ее апертуре. Из-за отличия диэлектрической постоянной материала линзы от диэлектрической постоянной свободного пространства, в котором пусть находится линза, происходит замедление (ускорение) фазовой скорости распространяющейся волны. Вследствие этого спектральные составляющие распространяющегося волнового фронта претерпевают преломления на обеих границах раздела в зависимости от угла их падения на эту границу.

Пусть на линзу, расположенную в плоскости xOy прямоугольной декартовой системы координат, падает плоская электромагнитная волна $g_1(x, y)$, распространяющаяся вдоль оси z . В соответствии с принятым выше допущением о тонкости линзы будем считать, что луч, падающий в произвольную точку входной плоскости 1-1 нашей системы, попадает в произвольную точку ее выходной плоскости 2-2 с теми же координатами (плоскости 1-1 и 2-2 перпендикулярны к оптической оси системы и касательны к сферическим поверхностям, образующим линзу). В параксиальном приближении для амплитудного множителя коэффициента прохождения электромагнитной волны отличие в углах падения для каждой составляющей спектра несущественно, поэтому будем учитывать лишь фазовый сдвиг в распространяющуюся волну. Фазовый сдвиг между плоскостями 1-1 и 2-2, как известно, определяется соотношением

$$g_2(x, y) = g_1(x, y)e^{jkn l}, \quad (9.1)$$

где $g_2(x, y)$ – выходной сигнал; $g_1(x, y)$ – входной сигнал; l – расстояние от произвольной точки входной плоскости системы до произвольной точки выходной плоскости системы (в нашем приближении относительно тонкости линзы расстояние от плоскости 1-1 до плоскости 2-2); n – показатель преломления среды, в которой распространяется электромагнитная волна.

Для произвольной точки апертуры

$$l = l_1(x, y) + l_2(x, y), \quad (9.2)$$

где $l_1(x, y)$ – толщина материала линзы в произвольной точке апертуры. Ввиду того, что на участке $l_1(x, y)$ волна распространяется в среде с показателем преломления n , а на участке $l_2(x, y)$ – в свободном про-

странстве, соотношение (9.1) с учетом (9.2) принимает вид

$$g_2(x, y) = g_1(x, y) e^{jkl} e^{jk(n-l)l_1(x, y)} \quad (9.3)$$

Теперь учтем, что

$$l = |z_1| + |z_2| = \text{const}, \quad l_1(x, y) = l'(x, y) + l''(x, y). \quad (9.4)$$

Так как левая поверхность линзы имеет уравнение $x^2 + y^2 + z_1^2 = r_1^2$, а правая – уравнение $x^2 + y^2 + z_2^2 = r_2^2$ и согласно формуле линзы, $\frac{1}{F} = (n-1) \left(\frac{1}{r_1} - \frac{1}{r_2} \right)$, где F – фокусное расстояние лин-

зы, поэтому последнее соотношение принимает вид

$$g_2(x, y) = g_1(x, y) e^{-\pi \frac{x^2 + y^2}{\lambda F}}. \quad (9.5)$$

Заметим, что для правильного пользования выражением (9.5) применительно к произвольным линзам следует радиусы кривизны поверхностей, направленных выпуклой частью против оси z , считать положительными, вдоль оси z – отрицательными.

Дадим физическую интерпретацию преобразования (9.5), выполняемого линзой над исходным волновым фронтом.

Рассмотрим случай, когда линза имеет идеальную сферическую форму и неограниченные размеры. При этом соотношение (9.5) выполняется достаточно строго. Нарушение этих условий приводит к отклонению фронта волны $g_2(x, y)$ от сферического, т.е. к появлению aberrаций. Эти явления достаточно подробно изучаются в оптике, поэтому мы на них останавливаться не будем.

Представляет определенный интерес выражение для поля электромагнитной волны, прошедшей линзовую систему, в ее фокальной плоскости:

$$g(x_F, y_F) = \frac{e^{jkF}}{j\lambda F} e^{j\frac{k}{2F}(x_F^2 + y_F^2)} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(x, y) e^{-j2\pi \left(\frac{x_F}{\lambda F} x + \frac{y_F}{\lambda F} y \right)} dx dy. \quad (9.6)$$

Есть преобразование Фурье от волнового фронта, падающего на линзу по пространственным частотам:

$$f_x = \frac{\cos \alpha}{\lambda}, \quad f_y = \frac{\cos \beta}{\lambda},$$

так как

$$\frac{x_F}{F} = \cos \alpha, \quad \frac{y_F}{F} = \cos \beta, \quad (9.7)$$

что следует из элементарных геометрических соображений для параксиального приближения.

Расширение пределов интегрирования до ∞ вместо конечной площади апертуры оправдано при обращении в нуль подынтегральной

функции за ее пределами. На практике волновые фронты конечных размеров выделяют из неограниченных волновых фронтов путем применения ограничивающих апертур с некоторым комплексным коэффициентом пропускания $t(x, y)$. В таком случае выражение (9.6) имеет вид

$$g(x_F, y_F) = \frac{e^{jkF}}{j\lambda F} e^{j\frac{\pi}{\lambda F}(x_F^2 + y_F^2)} \iint_{-\infty}^{\infty} t(x, y) e^{-j2\pi\left(\frac{x_F}{\lambda F}x + \frac{y_F}{\lambda F}y\right)} dx dy. \quad (9.8)$$

В дальнейшем пространственное ограничение волнового фронта будем считать задаваемым функцией $t(x, y)$ выражением (9.8).

Формулами (9.7) установлена связь между протяженностью в пространстве Фурье-образа объекта и его пространственными частотами, которая зависит от фокусного расстояния линзы F .

Таким образом, если посредством специального транспаранта с заданными амплитудными и фазовыми свойствами, помещенного в непосредственном контакте с линзой, задать некоторый исходный волновой фронт, то последняя с точностью до множителя выполняет двумерное преобразование Фурье над исходным волновым фронтом. Если в фокальной плоскости линзы регистрируется только интенсивность волнового фронта, этот множитель не играет никакой роли. И от него можно избавиться, если применить устройство, вносящее фазо-

вый сдвиг в проходящую волну $e^{-j\frac{\pi}{\lambda F}(x_F^2 + y_F^2)}$, согласно выражению (9.6). Таким устройством является линза с фокусным расстоянием F .

Следовательно, система, состоящая из двух линз с фокусными расстояниями F , расположенными на расстоянии F друг от друга, преобразует входной сигнал $t(x, y)$ в его Фурье-образ.

Если транспарант расположить в передней фокальной плоскости некоторой линзы, волновой фронт в ее задней фокальной плоскости представляет собой также Фурье-образ волнового фронта, пропущенного транспарантом. Физически выражения (9.6) и (9.8) означают, что амплитуды и фазы волнового фронта в точках пространства с координатами x_F, y_F определяются амплитудами и фазами спектральных

пространственных частот $\frac{x_F}{\lambda F}, \frac{y_F}{\lambda F}$, распространяющихся за линзой.

Если транспарант расположен в непосредственном контакте с линзой, его размеры должны быть меньше или равны размерам линзы. При размещении его в передней фокальной плоскости линза преобразовывает только часть пространственного спектра транспаранта. Итак, размеры Фурье-образа в задней фокальной плоскости пространственно ограничены. Это ограничение называют виньетированием.

При расположении транспаранта между линзой и ее задней фокальной плоскостью на расстоянии l от последней распределение поля в задней фокальной плоскости совпадает с выражением (9.8), но вме-

сто F фигурирует величина l , которая часто применяется при изменении масштаба Фурье-образов исследуемых волновых фронтов путем изменения расстояния l .

До сих пор мы рассматривали специальный случай падения на транспарант плоской электромагнитной волны. Пусть на линзу падает произвольный волновой фронт $g_1(x, y)$ и транспарант с коэффициентом пропускания $g_1(x, y)$ расположен на расстоянии d от линзы с фокусным расстоянием F . Будем искать волновой фронт $g_2(x, y)$ в произвольной плоскости, находящейся на расстоянии l от линзы.

Система представляет собой участок пространства от плоскости x_1Oy_1 до плоскости x_2Oy_2 , содержит линзу и состоит из трех последовательных подсистем «пл. x_1Oy_1 – пл. а-а», «пл. а-а – пл. б-б», «пл. б-б – пл. x_2Oy_2 ». Для всей системы, если она линейна и инвариантна, применим интеграл суперпозиции

$$g_2(x_2, y_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(x_1, y_1) h(x_2 - x_1, y_2 - y_1) dx_1 dy_1. \quad (9.9)$$

Для нахождения импульсного отклика системы рассчитаем прохождение сферической волны из произвольной точки плоскости x_1Oy_1 в произвольную точку плоскости x_2Oy_2 через все три подсистемы.

Поле в произвольной точке плоскости а-а в приближении Френеля запишется

$$g'_1(x, y) = \frac{e^{jkd}}{j\lambda d} e^{j\frac{k}{2d}[(x-x_1)^2 + (y-y_1)^2]}. \quad (9.10)$$

Согласно (9.5), волновой фронт в плоскости б-б

$$g'_2(x, y) = g'_1(x, y) e^{-j\frac{\pi}{\lambda F}(x^2 + y^2)}. \quad (9.11)$$

Проведя ряд преобразований можно получить формальное решение нашей задачи:

$$g_2(x_2, y_2) = \frac{1}{\lambda^2 dl} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g_1(x_1, y_1) \times \\ \times \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-j\frac{2\pi}{\lambda} \left(\frac{x_2 + Mx_1}{\lambda l} x + \frac{y_2 + My_1}{\lambda l} y \right)} dx dy \right] dx dy. \quad (9.12)$$

Из последнего соотношения для случаев, рассмотренных выше, могут быть получены выражения, показывающие, что волновой фронт $g_2(x_2, y_2)$ является Фурье-образом волнового фронта $g_1(x_1, y_1)$ по пространственным частотам $\frac{x_F}{\lambda F}, \frac{y_F}{\lambda F}$.

9.1. Явления, связанные с конечностью размеров волнового фронта

Согласно выражению (9.8), с точностью до множителя, стоящего перед интегралом, волновой фронт в фокальной плоскости линзы представляет собой двумерный Фурье-образ исходного волнового фронта транспаранта $t(x, y)$. Если на нем записана плоская волна $t(x, y) = 1$, интеграл имеет вид двумерной δ -функции:

$$\delta(f_x, f_y) = \delta\left(\frac{x_F}{\lambda F}, \frac{y_F}{\lambda F}\right) = \begin{cases} 1, & x_F = y_F = 0, \\ 0, & x_F = y_F \neq 0. \end{cases}$$

Это означает, что падающая на линзу волна преобразуется в точку, расположенную на оптической оси ее фокальной плоскости. Если не учитывать конечность пределов интегрирования по апертуре транспаранта, то значение интеграла является не δ -функцией, а некоторым математическим выражением, имеющим конечную область существования в плоскости $x_F O y_F$.

Таким образом, конечность размеров транспаранта ограничивает спектр пространственных частот исходного волнового фронта и приводит к абберациям в оптических системах формирования и обработки волновых фронтов. Ограниченность размеров линзы также влияет на структуру выполняемого ею Фурье-преобразования исходного волнового фронта.

При рассмотрении дифракции электромагнитной волны на синусоидальной решетке видно, что углы дифракции первой и минус первой гармоник определяются соотношением $\gamma = \pm \arccos \lambda f$. Учитывая конечность размеров синусоидальной решетки и преобразующей линзы, находим, что максимальный угол дифракции, определяемый размером, линзы b :

$$\gamma_{\text{макс}} = \arctg \frac{b}{d} \approx \arcsin \frac{b}{d}.$$

Поэтому максимальная пространственная частота лучей, дающих вклад в изображение, находится из неравенства $f_{\text{макс}} = b/d\lambda$. Лучи, излучаемые краями транспаранта, участвуют в формировании изображения, если идут под углами

$$\gamma' \leq \arcsin \frac{b-a}{a} \left(f'_{\text{макс}} = \frac{b-a}{\lambda d} \right).$$

Следовательно, для повышения информативности системы нужно увеличивать размер линзы относительно размеров транспаранта и уменьшать расстояние между ними.

ЛЕКЦИЯ 10. Физическая реализация интегральных преобразований в оптических системах

Согласно изложенному выше, на выходе первой из приведенных оптических систем получим исходный волновой фронт $a(x, y)$. Если же в заднем фокусе первой из линз, совпадающем с передним фокусом второй линзы, одновременно со спектром $A(f_x, f_y)$ воспроизвести каким-то образом спектр некоторой функции $H(f_x, f_y) = F\{h(x, y)\}$, то на выходе оптической системы получим свертку функций $b(x, y) = a(x, y) \otimes h(x, y)$.

Физическая реализация предлагаемого процесса математической обработки волнового фронта осуществляется путем записи на некоторый носитель – транспарант – оптической информации (фото-пленка, фотопластинка, фотополупроводник и др.) и пространственной информации о его амплитуде и фазе. Если распределение амплитуды адекватно распределению интенсивности волнового фронта, то регистрация фазового распределения представляет значительные технические трудности. Для его воспроизведения нужно, чтобы фазовый набег волны в каждой точке плоскости x, y соответствовал реальному распределению фазы. Это распределение можно воспроизвести либо модуляцией коэффициента преломления носителя информации, либо изменением толщины носителя при неизменном коэффициенте преломления согласно воспроизводимому фазовому распределению. В связи с малостью длины оптических длин волн электромагнитного излучения, резкими градиентами фазы в большинстве приложений прямые методы регистрации и воспроизведения фазовых распределений оптических волновых полей практически не реализуемы. Поэтому естественным стремлением в развитии технических средств радиооптики было найти способ записи на амплитудных транспарантах фазовой и поляризационной информации о волновых фронтах. К этому вопросу вернемся несколько позднее, а сейчас остановимся на физической реализации интегральных преобразований в радиооптике с помощью простейших фокусирующих систем.

Оптические методы обработки информации в радиоэлектронных системах занимают значительное место. Преимущества их заключаются в двумерности обрабатываемой информации, возможности одновременной ее обработки по всей площади, мгновенности обработки, простоте операций Фурье-преобразования, корреляционных и интегральных преобразований над сигналами, пространственно-частотной фильтрации и т.д.

Пусть имеются диапозитивы с коэффициентами пропускания по интенсивности $T_1(x, y)$ и $T_2(x, y)$. В некоторой точке пространства за

этими диапозитивами интенсивность волны определяется выражением

$$I(x, y) \sim \iint_S T_1(x, y) T_2(x, y) dx dy.$$

Помимо математической операции интегрального умножения оптические устройства позволяют осуществлять и более сложные операции.

Если представить, что транспарант $T_1(x, y)$ перемещается в плоскости $x_1 O y_1$ со скоростями v_x, v_y , то интенсивность выходного сигнала

$$I(x_3, y_2, \tau) \sim \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} T_1(x - v_x \tau, y - v_y \tau) T_1(x, y) dx dy.$$

При этом последнее интегральное преобразование является сверткой функции пропускания в каждой точке выходной плоскости $x_2 O y_2$. Однако операцию свертки можно выполнить и без движения транспарантов. Требуемое распределение интенсивности волнового поля в плоскости регистрации при осуществлении операции свертки должно иметь вид

$$I(x, y) k \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\xi, \eta) h(x - \xi, y - \eta) d\xi d\eta \right|^2,$$

где k – коэффициент пропорциональности; ξ, η – координаты входной плоскости, в которой расположен транспарант.

В плоскости $x_1 O y_1$ моделируется некоторый сигнал $g(\xi, \eta)$ в виде транспаранта с коэффициентом пропускания $|g(\xi, \eta)|$. Для источника света с единичной амплитуды линза L_1 служит коллиматором источника s , линза L_2 выполняет операцию $\Phi\{g(\xi, \eta)\} = G(x_2 / \lambda F, y_2 / \lambda F)$ с некоторой комплексной постоянной k_1 , т.е. в плоскости $x_2 O y_2$ имеем вместо $g(\xi, \eta)$ сигнал $k_1 G(x_2 / \lambda F, y_2 / \lambda F)$.

Если в плоскости $x_2 O y_2$ поместить транспарант, представляющий собой с точностью до постоянной k_2 Фурье-образ некоторой функции $h(x, y)$, т.е. с коэффициентом пропускания $k_2 H(x_2 / \lambda F, y_2 / \lambda F)$, то распределение света за плоскостью будет пропорционально произведению $G(x_2 / \lambda F, y_2 / \lambda F) \cdot H(x_2 / \lambda F, y_2 / \lambda F)$.

При этом распределение интенсивности волнового поля в плоскости $x_3 O y_3$

$$I(x_3, y_3) \sim \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\xi, \eta) h(x_3 - \xi, y_3 - \eta) d\xi d\eta \right|^2,$$

т.е. является сверткой сигналов $g(\xi, \eta), h(x, y)$.

Заметим, что изложенные рассуждения были проведены нами для когерентных волновых фронтов. Если условие когерентности обяза-

тельно, суперпозиционный интеграл, свертка, импульсный отклик и передаточная функция фокусирующих систем рассчитываются аналогичным образом, но вместо комплексных величин напряженности поля электромагнитной волны берутся действительные величины – интенсивности входного и выходного сигналов. При этом передаточную функцию системы называют оптической передаточной функцией (ОПФ), а ее модуль – модуляционной передаточной функцией (МПФ) [5]. На практике широкое распространение получили как когерентные, так и некогерентные оптические системы. Целесообразность выбора типа оптической системы определяется, исходя из конкретных решаемых задач: для одних входных волновых фронтов более предпочтительна когерентная система, для других – некогерентная. В некоторых оптических системах нет плоскостей, в которых можно регистрировать пространственные частоты волнового фронта, поэтому они не могут применяться для пространственно-частотной фильтрации волновых фронтов. Однако корреляционные и интегральные преобразования могут выполняться ими в ряде практических применений.

10.1 Геометрические и физические характеристики радиооптических систем

Из предыдущего материала видно, что математические преобразования над волновыми фронтами в РОС основаны на преобразованиях Фурье над двумерными комплексными функциями, операциях перемножения волновой информации, записанной на транспарантах, с последующим Фурье-преобразованием, дающим возможность физически реализовать операцию свертки. Нетрудно убедиться, что, выбирая определенным образом вид транспаранта $H(f_x, f_y)$ относительно входного сигнала $a(x, y)$, можно реализовать и другие сложные математические операции над последним, например, корреляцию, кросс-корреляцию и т.д. Процесс оптической обработки волновых фронтов $a(x, y)$, $g_1(\xi, \eta)$ посредством транспарантов $H(f_x, f_y), T(x_2, y_2)$ называют пространственной фильтрацией в области частотных спектров. Если $H(f_x, f_y) = F\{a^*(x, y)\}$, $T(x_2, y_2) = F\{g_2^*(\xi, \eta)\}$, фильтрация называется согласованной, имеющей важное практическое применение в распознавании образов, на чем мы остановимся в дальнейшем более подробно.

Сейчас же рассмотрим некоторые важные для практики геометрические и физические характеристики РОС, реализующих Фурье-преобразования.

Из свойства подобия преобразования Фурье следует, что если в РОС с фокусирующей линзой исходный волновой фронт (транспа-

рант) уменьшается в размерах, то $b(x, y)$ изображение увеличивается в той или иной пропорции. Обратный эффект очевиден. Из этого же свойства следует, что симметрия или геометрическое подобие исходного волнового фронта в различных сечениях обуславливает симметрию или геометрическое подобие изображения, которые могут быть нарушены лишь специальным выбором фильтра-транспаранта $H(f_x, f_y)$.

Смещение исходного волнового фронта на величины α, β во входной плоскости не приводит к изменению распределения интенсивности волны в плоскости пространственного спектра. Лишь после второго преобразования Фурье произойдет его сдвиг в пространстве на величины α, β за счет фазовой экспоненты $\exp[-j \cdot 2\pi(f_x\alpha + f_y\beta)]$ волны, распространяющейся от плоскости фильтрации до выходной апертуры x_2, y_2 . Это означает, что фильтруемую информацию можно сдвигать в исходной плоскости, не меняя геометрического положения, РОС, а отфильтрованную информацию фиксировать в сдвинутой точке выходной плоскости (выходного зрачка). Таким образом, имея распределенный во входной плоскости сколь угодно большой набор информации (например, текст или цифры), результат его оптической обработки – фильтрации – в выходной плоскости получается мгновенно (со скоростью света) в соответствующих ее точках.

Из теоремы свертки следует также, что преобразованием Фурье произведения двух функций является их свертка, и наоборот, преобразованием Фурье от свертки двух функций является произведение их спектров.

При изложении теории линзы и доказательстве свойства ее выполнять двумерное Фурье-преобразование над исходным волновым фронтом мы пренебрегали третьей компонентой электромагнитной волны – ее продольной составляющей. Между тем нахождение плоскости за линзой, в которой выполняется математически точное преобразование Фурье, чрезвычайно важно. На положение этой плоскости относительно продольного направления оптической оси РОС влияют многие физические факторы – ограничение размеров входной апертуры, когерентность и ширина спектра излучения, неоднородности материалов линз и транспарантов, вибрации отдельных частей РОС и др. Все эти факторы, в конечном счете, приводят к отклонениям действительного распределения волнового фронта в выходной апертуре от расчетного – к абберациям. Для их устранения применяются различные технические приемы, например, защита от вибраций, расположение линз и транспарантов в специальных кюветах, фазокомпенсирующие фильтры, зонирование линз и др.

В заключение данного параграфа рассмотрим явления, связанные с конечностью размеров волнового фронта. Волновой фронт в

фокальной плоскости линзы представляет собой двумерный Фурье-образ исходного волнового фронта транспаранта $t(x, y)$. Если на нем записана плоская волна $t(x, y) = 1$, интеграл имеет вид двумерной δ -функции:

$$\delta(f_x, f_y) = \delta\left(\frac{x_F}{\lambda F}, \frac{y_F}{\lambda F}\right) = \begin{cases} 1, & x_F = y_F = 0; \\ 0, & x_F = y_F \neq 0. \end{cases}$$

Это значит, что падающая на линзу волна преобразуется в точку, расположенную на оптической оси ее фокальной плоскости. Если же учитывать конечность пределов интегрирования по апертуре транспаранта, то значение интеграла является не δ -функцией, а некоторым математическим выражением, имеющим конечную область существования в плоскости $x_F O y_F$.

Таким образом, конечность размеров транспаранта ограничивает спектр пространственных частот исходного волнового фронта и приводит к абберациям в оптических системах формирования и обработки волновых фронтов. Ограниченность размеров линзы также влияет на структуру выполняемого ею Фурье-преобразования исходного волнового фронта.

Следовательно, для повышения информативности системы нужно увеличивать размер линзы относительно размеров транспаранта и уменьшать расстояние между ними.

Пусть имеется комплексный исходный волновой фронт

$$g(x, y) = |g(x, y)| e^{j\varphi(x, y)}$$

в плоскости xOy , при помощи модуля $|g(x, y)|$ и фазы $\varphi(x, y)$ которого закодирована какая-то информация. При распространении этого волнового фронта на некоторое расстояние требуется зарегистрировать обусловленную его распространением амплитудно-фазовую структуру поля в плоскости регистрации $\xi O \eta$. Волновой фронт $g(\xi, \eta)$ определится посредством двумерного суперпозиционного интеграла через волновой фронт $g(x, y)$ и импульсный отклик участка свободного пространства между плоскостями xOy и $\xi O \eta$. По виду волнового фронта $g(\xi, \eta)$ требуется восстановить вид волнового фронта $g(x, y)$, т.е. выделить закодированную информацию. Очевидно, что в случаях, поддающихся аналитическому решению, удобно, если $g(x, y)$ и $g(\xi, \eta)$ связаны каким-либо легко реализуемым обратным математическим преобразованием. Тогда вид волнового фронта $g(x, y)$ однозначно находится по виду волнового фронта $g(\xi, \eta)$ и наоборот. Представляет практический интерес случай, когда $g(\xi, \eta)$ и $g(x, y)$ связаны прямым и обратным двумерными Фурье-преобразованиями.

К сожалению, не существует индикаторов электромагнитного поля, позволяющих одновременно регистрировать его распределение по всей плоскости $\xi O \eta$ с необходимой чувствительностью и разрешающей способностью в реальном времени. Поэтому производится выборочная регистрация поля в ряде дискретных точек этой плоскости. Дискретизация волновых фронтов вносит в их структуру искажения, которые необходимо учитывать и устранять.

Определим вспомогательную двумерную функцию, называемую гребенчатой, следующим образом:

$$\text{combxcomby} = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(x-m)\delta(y-n), \quad (10.1)$$

где m и n – целые числа единиц расстояния вдоль осей x и y соответственно. Очевидно, что прямоугольную сетку выборок функций $g(x, y)$ с помощью гребенчатой функции можно представить в виде

$$g_b(x, y) = g(x, y) \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(x-ma)\delta(y-nb), \quad (10.2)$$

где a и b – периоды выборки вдоль осей x и y соответственно. Это выражение можно преобразовать к виду

$$\begin{aligned} g_b(x, y) &= g(x, y) \frac{1}{ab} \text{comb} \frac{x}{a} \text{comb} \frac{y}{b} = \\ &= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} g(ma, nb) \delta(x-ma) \delta(y-nb). \end{aligned} \quad (10.3)$$

Если по функции $g(\xi, \eta)$ в плоскости $\xi O \eta$, заданной своими дискретными значениями (10.3), восстановить вид функции $g(x, y)$ в плоскости $x O y$, последняя также будет дискретной.

Согласно интегралу свертки, свойству гребенчатой функции со сдвигом по аргументу, симметрии и фильтрующему свойству δ -функции получим:

$$\begin{aligned} \Phi\{g_b(x, y)\} &= \Phi\{g(x, y)\} \otimes \Phi\left\{\frac{1}{ab} \text{comb} \frac{x}{a} \text{comb} \frac{y}{b}\right\} = \\ &= \frac{1}{ab} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} G\left(\xi - \frac{m}{a}, \eta - \frac{n}{b}\right), \end{aligned} \quad (10.4)$$

где $G(\dots)$ – Фурье-образ функции $g(x, y)$ по пространственным частотам ξ, η .

Таким образом, при дискретизации исходного волнового фронта его Фурье-образ с точностью до множителя $1/ab$ становится набором дискретных фурье-образов дискретизированного волнового фронта, сдвинутых в плоскости частот на целое число величин $1/a, 1/b$. Слагаемые суммы (10.4) не перекрываются, если выполняются соотношения

$$a \leq \frac{1}{\xi_{\max}}; \quad b \leq \frac{1}{\eta_{\max}}, \quad (10.5)$$

где $\xi_{\max}$, $\eta_{\max}$ – максимальные пространственные частоты исходного волнового фронта.

Если периоды выборок взяты согласно (10.5) (что возможно лишь для волновых фронтов с существующими максимальными пространственными частотами, т.е. для всегда имеющих место в практике пространственно ограниченных волновых фронтов), из выражения (10.3), используя понятие прямоугольной функции

$$\text{rect}_{\xi} \text{rect}_{\eta} = \begin{cases} 1 & |\xi| \leq 1/2, |\eta| \leq 1/2 \\ 0 & \text{в остальных случаях} \end{cases}$$

и фильтрующие свойства δ -функции, находим:

$$g(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} g\left(\frac{m}{\xi_{\max}}, \frac{n}{\eta_{\max}}\right) \frac{\sin \pi(\xi_{\max} x - m)}{\pi(\xi_{\max} x - m)} \times \frac{\sin(\eta_{\max} y - n)}{\pi(\eta_{\max} y - n)}. \quad (10.6)$$

Это означает, что по дискретным значениям волнового фронта $g(\xi, \eta)$ (первый множитель под знаком двойной суммы), полученным в плоскости $\xi O \eta$, можно восстановить вид исходного волнового фронта $g(x, y)$ посредством функций вида $\sin \varphi / \varphi$, построенных в каждой точке выборки. На практике число отсчетов всегда ограничено, поэтому точного значения волнового фронта $g(x, y)$ по выборкам воспроизвести нельзя. Однако это можно сделать со сколь угодно малым приближением путем увеличения количества выборок $g(m/\xi_{\max}, n/\eta_{\max})$. Учитывая (10.6), связь между $g(\xi, \eta)$ и $g(x, y)$ при помощи прямого преобразования Фурье и таблицы Фурье-образов специальных функций, получаем

$$g(\xi, \eta) = \begin{cases} \frac{1}{\xi_{\max} \eta_{\max}} \times \\ \times \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} g\left(\frac{m}{\xi_{\max}}, \frac{n}{\eta_{\max}}\right) \exp\left[-j2\pi\left(\frac{m\xi}{\xi_{\max}} + \frac{n\eta}{\eta_{\max}}\right)\right], & \left|\frac{\xi}{\xi_{\max}}\right| \leq \frac{1}{2}, \\ \text{или} \\ g(\xi, \eta) = \left|\frac{\eta}{\eta_{\max}}\right| \leq 1/2, & g(\xi, \eta) = 0 \text{ в остальных случаях,} \end{cases} \quad (10.7)$$

т.е. другое представление анализируемого волнового фронта $g(\xi, \eta)$ через его выборочные значения.

В дальнейшем под синтезом волновых полей будем понимать измерение (расчет) физических и геометрических параметров источника по излучаемому или рассеиваемому волновому фронту, под ана-

лизом – нахождение излучаемого или рассеиваемого источником волнового фронта по его физическим и геометрическим свойствам, заданным математически либо путем физического описания.

10.2. Оптическая фильтрация. Фильтр Вандер Люгта

Принципы изменения структуры оптических изображений путем воздействия на Фурье-спектры были сформулированы и осуществлены в конце XIX – начале XX в. Э. Аббе и А. Портером. Суть проведенных ими экспериментов заключалась в следующем: если перед линзой поместить проволочную сетку из двух перекрывающихся под прямым углом систем параллельных проводов, то в задней фокальной плоскости формируется ее Фурье-образ, все компоненты которого складываются и формируют изображение сетки. Если же в фокальной плоскости затемнять различные участки ее, можно проводить фильтрацию получаемого изображения. Например, помещая в фокальной плоскости горизонтальную щель в непрозрачном для света экране, вместо изображения сетки имеем набор вертикальных проволок. Вертикальная же щель позволяет их отфильтровать и получить горизонтальные проволоки исходной сетки.

Для оптической фильтрации различных спектральных компонент в плоскости пространственных частот А. Вандер Люгтом (1963) был предложен метод получения пространственно-частотных фильтров, которые синтезируются с помощью оптической системы. Одна часть пучка света проходит через транспарант с коэффициентом пропускания $h(x, y)$ (линза Л осуществляет фурье-преобразование $H(x_2/\lambda F, y_2/\lambda F)$ с точностью до множителя $1/\lambda F$, другая падает под некоторым углом θ от призмы П в плоскость x_2Oy_2 :

$$u_{\Pi}(x_2, y_2) = |u_{\Pi}| \exp(-jk \sin \theta y_2) = |u_{\Pi}| \exp\left(-j \cdot 2\pi \frac{\sin \theta}{\lambda} y_2\right),$$

где $(\sin \theta)/\lambda$ – пространственная частота вдоль оси y .

Суммарная интенсивность света в этой плоскости

$$\begin{aligned} I(x_2, y_2) &= \left| |u_{\Pi}| \exp\left(-j \cdot 2\pi \frac{\sin \theta}{\lambda} y_2\right) + \frac{1}{\lambda F} H\left(\frac{x_2}{\lambda F}, \frac{y_2}{\lambda F}\right) \right|^2 = \\ &= |u_{\Pi}|^2 + \frac{1}{\lambda^2 F^2} |H|^2 + \frac{2u_{\Pi}}{\lambda F} |H| \cos\left[2\pi \frac{\sin \theta}{\lambda} y_2 - \psi\left(\frac{x_2}{\lambda F}, \frac{y_2}{\lambda F}\right)\right]. \end{aligned}$$

Из последнего соотношения следует, что можно получить транспарант с записью комплексной функции $H(f_x, f_y)$ на обычную пленку, регистрирующую только интенсивность света. Амплитудная и фазовая информация о комплексной функции записана в виде амплитудной и фазовой модуляций несущей пространственной частоты

$(\sin\theta/\lambda)$. Существуют и другие схемы получения такой интенсивности. После обработки диапозитива его коэффициент пропускания пропорционален интенсивности $I(x, y)$. Третье слагаемое соотношения с точностью до экспоненциального множителя представляет собой интересующую нас функцию H . Установим этот транспарант в плоскость x_2Oy_2 системы обработки информации. Если в плоскость x_1Oy_1 подается фильтруемая функция $g(x_1y_1)$, в плоскости x_2Oy_2 формируется распределение света $(1/\lambda F)G(x_2/\lambda F, y_2/\lambda F)$.

Следующая линза системы обработки выполняет Фурье-преобразование поля u_2 с множителем $1/\lambda F$. Считая, что система координат в плоскости x_3Oy_3 инверсно преобразована, вместо

$$\Phi\{\Phi\{g(x_1, y_1)\}\} = g(-x_2, -y_3)$$

будем писать $g(x_3, y_3)$. Кроме того, учтем, что

$$\begin{aligned}\Phi\left\{\exp\left(j \cdot 2\pi \frac{\sin\theta}{\lambda} y_3\right)\right\} &= \delta(x_3, y_3 + F \sin\theta); \\ \Phi\left\{\exp\left(j \cdot 2\pi \frac{\sin\theta}{\lambda} y_3\right)\right\} &= \delta(x_3, y_3 - F \sin\theta).\end{aligned}$$

Используя фильтрующие свойства δ -функций, это выражение превратится в интегралы, один из которых представляет собой свертку функций g и h , смещенную по оси y в плоскости x_3Oy_3 на величину $F \sin\theta$, а второй есть кросс-корреляция g и h , смещенная по оси y на величину $F \sin\theta$. Первые же два слагаемых дают фон, не смещенный из начала координат.

$$\begin{aligned}\frac{|u_n|}{\lambda F} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\xi, \eta) h^*(\xi - x_3, \eta - y_3 + F \sin\theta) d\xi d\eta; \\ \frac{|u_n|}{\lambda F} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\xi, \eta) h(x_3 - \xi, y_3 - \eta + F \sin\theta) d\xi d\eta.\end{aligned}$$

Таким образом, при формировании волнового фронта в предметной плоскости в системе имеется возможность путем манипуляций в частотной плоскости отфильтровать различные пространственно-частотные составляющие передаваемого изображения или отдельные его детали. Для проведения операции фильтрации требуется перемножить частотный спектр фильтруемой функции с частотным спектром некоторой заданной функции – заданного импульсного отклика. Получение спектра импульсного отклика методом Вандер Люгта позволяет отказаться от использования при фильтрации фазосдвигающих элементов, так как фазовая информация кодируется в фильтре в виде модуляции несущей пространственной частоты, т.е. голографическим путем.

ЛЕКЦИЯ 11. Распознавание образов в радиооптических системах

Новым направлением в кибернетике является распознавание образов, т.е. объектов, явлений, ситуаций. Это направление возникло при разработке читающих автоматов, медицинских диагностических машин, устройств, прогнозирующих метеорологические процессы, и т.д.

Первые попытки распознавания объектов были основаны на сравнении свойств объектов со свойствами эталонов, а многие последующие – на использовании принципов построения биологических систем.

Теория распознавания оперирует образами, подразделяющимися на объекты, явления и ситуации. Образами могут быть звуки речи, подводные лодки, космические корабли, прогноз погоды, знаки текста, данные диагностики заболеваний и т.п. Признаки образов вводятся с помощью датчиков-рецепторов, которые являются неотъемлемой частью распознающего устройства и могут быть оптическими, акустическими, радиоэлектронными, гидравлическими, электромагнитными, химическими и т.д. В распознающем устройстве эти признаки обрабатываются, и автоматически принимается решение, к какому классу следует отнести образ.

Все объекты по основным признакам делятся на классы. Не все признаки дают одинаковый результат при распознавании. Использование отдельного признака не позволяет добиться того результата, который можно получить, применяя одновременно всю совокупность признаков.

Достоверность распознавания в основном определяется точностью измерения признаков: с увеличением точности достоверность возрастает; при увеличении числа признаков, одновременно используемых для распознавания, достоверность также увеличивается. Иногда распознавание можно рассматривать как статистическую задачу, в которой вероятность распознавания определяется дифференциальными законами распределения признаков объектов и помех.

В теории распознавания любой объект может иметь бесконечно большое число признаков. Признаки делятся на объективные и описательные. Объективные признаки можно выразить количественно, к описательным относятся признаки, которые количественно охарактеризовать нельзя. Для удобства описательные признаки делятся на градации (группы), каждой из которых присваивается то или иное числовое значение. Признаки схожих объектов незначительно отличаются друг от друга в количественном отношении.

Распознавание объекта – это сравнение его признаков с признаками одного или нескольких заранее определенных классов для отне-

сения объектов к тому или иному классу. Существуют правила количественной классификации распознавания. Использование тех или иных правил зависит от конкретных задач распознавания. Правильность установления величины количественного критерия распознавания определяется практикой. По мере накопления сведений о правильно распознанных объектах можно установить оптимальную величину распознающего порога, причем самое важное при выборе порога – получить максимально верное распознавание классов.

Очень часто в теории распознавания приходится разделять объекты на два класса, т.е. использовать дихотомию. *Дихотомия* – это деление всего алфавита классов объектов на два противоречащих друг другу видовых понятия, из которых одно представляет собой отрицание признаков другого. В качестве воспринимающих устройств (рецепторов) системы обычно применяют радиотехнические, оптоэлектронные, акустические и оптические устройства, позволяющие заблаговременно обнаружить объекты на больших расстояниях.

В устройстве сравнения значение поступающих реализаций сравнивается в каждый момент времени с описанием всех классов – эталонов, которые хранятся в блоке памяти. Сопоставление найденных вероятностных характеристик по заданному критерию и отнесение объекта к одному из эталонных классов выполняются в блоке принятия решения.

При распознавании образов в РОС широко используются согласованные фильтры. Пространственный фильтр называется *согласованным с сигналом* $g(x, y)$, если его импульсный отклик имеет вид $h(x, y) = g^*(-x, -y)$. При подаче на вход фильтра сигнала $g(x, y)$ на его выходе появится сигнал

$$f(x, y) = \int \int_{-\infty}^{\infty} g(\xi, \eta) g^*(\xi - x, \eta - y) d\xi d\eta,$$

т.е. кросскорреляционная функция функций g и g^* . Если этот отклик заменить его Фурье-образом и расположить в системе обработки в частотной плоскости второй линзы, а в плоскости $x_1 O y_1$ поместить транспарант с коэффициентом пропускания $g(x, y)$, за фильтром частотный спектр будет пропорционален $G(x_2 / \lambda F, y_2 / \lambda F) \times G^*(x_2 / \lambda F, y_2 / \lambda F)$, т.е. действительной величине $|G|^2$. Это значит, что на последнюю линзу падает плоская волна и линза фокусирует ее в точку – фокус линзы. Если же входной сигнал отличен от $g(x, y)$, свет не будет собираться в яркую точку фокуса последней линзы. Если входной сигнал смещен относительно оптической оси, яркая точка в фокальной плоскости последней линзы тоже смещается на соответствующее расстояние. Указанное свойство согласованного фильтра может использоваться для обнаружения среди

всех входных сигналов сигнала, с которым фильтр согласован, т.е. для распознавания образов.

На вход системы распознавания подаются одновременно или последовательно распознаваемые волновые фронты. На выходе каждого из фильтров, входящих в систему, находится устройство отклика к полной энергии, приходящей от образа, с которым фильтр согласован (различным образом может соответствовать различный уровень энергии). Нормировка осуществляется как электронными методами путем определенных оптико-электронных преобразований, так и оптическим методом путем создания в выходной плоскости соответствующего фона. При наличии на выходе сигнала, согласованного с одним из фильтров G_k^* , на выходе последнего будет наибольший сигнал.

Очевидно, что таким согласованным фильтром может быть фильтр Вандер Люгта, так как на выходе оптической системы, включающей его в соответствующей точке, формируется кросскорреляционный сигнал. Для согласованной фильтрации следует синтезировать фильтры по методу Вандер Люгта. Опознаваемые сигналы следует подавать на вход последовательно. На одном транспорте можно изготовить несколько фильтров. При этом одна пространственная несущая частота опорной волны модулируется волновыми фронтами от различных объектов.

Для практической реализации фильтрации важна точность установки фильтров и подача входных сигналов, подвергающих Фурье-преобразованиям.

При фильтрации сигнала в присутствие шума $g(x, y) = s(x, y) + n(x, y)$ на изготовленный фильтр контактно накладывается диапозитив со спектральной плоскостью шума. В выходной плоскости появляется яркая точка, свидетельствующая о присутствии сигнала и указывающая его положение на выходной плоскости корреляра.

В реальном случае полученное отношение сигнал–шум заметно ниже теоретического. Это происходит в основном из-за шумов фотоэмульсии транспаранта, на который записан фильтр, фазовых шумов фотоматериала, а также ошибок фотообработки, юстировки и аберраций оптической системы.

Наилучшие результаты получаются с чисто амплитудным фильтром. Эксперимент показал, что распознавание методом согласованной оптической фильтрации менее критично к деформации размеров объекта, а также к его повороту. Применение оптимальной фильтрации перспективно при распознавании букв негативного текста, отпечатков пальцев, для поиска карточек в картотеках и т.д.

Рассмотрим практические аспекты пространственной фильтрации. В некоторой системе обозначим входную плоскость P_1 , плоскость преобразования – P_2 с координатами x_1, y_1 и x_2, y_2 соответственно. Поместим входной транспарант в передней фокальной плоскости линзы L_1 , осуществляющей преобразование Фурье, и осветим его однородным по амплитуде и фазе пучком когерентного света. Полагая, что входной пучок имеет единичную амплитуду, записываем комплексное пропускание в плоскости P_2 как $f(x_1, y_1) = A(x_1, y_1) \times e^{-i\varphi(x_1, y_1)}$, где $A(x_1, y_1)$ – положительная величина, представляющая собой амплитуду, а $\varphi(x_1, y_1)$ описывает фазу входной функции (для физической определенности можно отождествить A с электрическим или магнитным вектором электромагнитной волны). Это означает, что если ко всем фазам в наших вычислениях добавить произвольно постоянную фазу, физически ничего бы не изменилось.

В плоскости P_2 будем иметь Фурье-образ функции $f(x_1, y_1)$, который по определению имеет вид

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-2\pi i(x_1 u + y_1 v)} dx_1 dy_1,$$

где u, v – координаты пространственной частоты в плоскости P_2 ; $x_2 = u\lambda f_1, y_2 = v\lambda f_1$ – координаты расстояния в плоскости P_2 ; λ – длина волны используемого света; f_1 – фокусное расстояние преобразующей линзы L_1 .

Если за плоскостью преобразования P_2 поместить вторую преобразующую линзу L_2 , то распределение амплитуд света в выходной плоскости $P_3(x_3, y_3)$ будет двойным преобразованием Фурье входного транспаранта (x_1, y_1) , где координаты в плоскостях P_1 и P_3 связаны соотношениями $x_3 = (f_2 / f_1)x_1$ и $y_3 = (f_2 / f_1)y_1$.

Выходное изображение является полностью перевернутым изображением входного транспаранта, причем коэффициент увеличения равен отношению фокусных расстояний линз f_2/f_1 .

Вспомним, что входной свет когерентен и распределение, сформированное в P_2 , есть преобразование Фурье $F(u, v)$ от $f(x_1, y_1)$. Если в P_2 поместим на оптической осп непрозрачный точечный экран, то задержим постоянную составляющую входного изображения и пропустим только высокочастотные компоненты света.

В плоскости P_3 будет получено распределение $f(x_1, y_1)$, прошедшее фильтрацию. Если в P_2 находится линейный фильтр, пропускание которого увеличивается с увеличением расстояния от центра, то в P_3 наблюдается продифференцированный входной сигнал. Помещая непрозрачные точечные экраны в определенных местах в плоскости P_2 , можно удалять различные частотные составляющие, присутствующие

во входном изображении. Распознавание образов (изображений или их фрагментов) и знаков (букв, цифр и символов) является одним из наиболее привлекательных областей применения оптической обработки информации. Задача оптической системы, предназначенной для распознавания образов или знаков, состоит в обнаружении интересующего нас образа во входном изображении. При распознавании знаков обычно используется постоянный набор эталонных функций, задача же системы состоит в том, чтобы определить, какая из этого класса эталонных функций присутствует на входе (а также в каком месте входной плоскости она находится).

Ключевой операцией является корреляция входной и эталонной функций или двух входных функций. В наиболее распространенной схеме оптического коррелятора корреляция осуществляется перемножением Фурье-образа входной и эталонной функции с последующим преобразованием Фурье полученного произведения. При этом эталонная функция записывается в виде своего комплексно-сопряженного Фурье-образа. Поскольку эталонная функция помещается в частотной плоскости коррелятора, она, в сущности, является пространственным фильтром. Амплитудное пропускание записанной эталонной функции в общем случае имеет комплексный характер и, следовательно, подобно амплитудному пропусканию голограммы. Но цель пространственного фильтра-голограммы состоит в определении соответствия (согласования) между входным образом (или его частью) и эталонной функцией (а не в формировании эстетически приятного изображения на выходе, как в голографии). Таким образом, комплексный эталонный Фурье-образ, расположенный в частотной плоскости, можно назвать *согласованным пространственным фильтром* (СПФ).

Все вариации фазы, присутствующие в волновом фронте, падающем на плоскость P_2 , устраняются фильтром, после чего за плоскостью P_2 распространяется плоская волна. Линза L_2 фокусирует эту волну в точку в плоскости P_3 . Итак, оптически согласованный фильтр можно рассматривать как фазокомпенсирующую пластину.

Если входной сигнал отличается от $h(x_1, y_1)$, то кривизна волнового фронта не будет точно скомпенсирована фильтром и распределение в плоскости P_3 будет иметь вид пятна с меньшей интенсивностью или пятно вообще будет отсутствовать.

Если входной сигнал $h(x_1, y_1)$ помещен в центре плоскости P_1 , то яркая точка света в выходной плоскости будет располагаться в центре плоскости P_3 .

Если входной сигнал $h(x_1, y_1)$ смещен вверх или вниз (или влево-вправо) относительно центра плоскости P_1 , то плоский волновой фронт за плоскостью P_2 будет идти под углом к оптической оси. Фурье-образ по-прежнему будет находиться на оси в плоскости P_2 , но

будет иметь постоянный фазовый сдвиг. Плоский волновой фронт за плоскостью P_2 будет образовывать угол с оптической осью. Линза L_3 просто фокусирует этот плоский волновой фронт в точку на плоскости P_3 . По мере того, как $h(x_1, y_1)$ сдвигается вверх или вниз (или влево-вправо) относительно центра плоскости P_3 положение яркого выходного светового пятна в плоскости P_3 сдвигается вниз или вверх (или вправо-влево) соответственно.

Итак, координаты яркого выходного светового пятна в плоскости P_3 пропорциональны координатам сигнала $h(x_1, y_1)$ во входной плоскости (с коэффициентом увеличения f_2/f_1), а интенсивность его в плоскости P_3 пропорциональна степени согласования входной и фильтрующей функций. Так как система линейна, справедлив принцип суперпозиции и приведенные выше соображения справедливы для множества сигналов $h_n(x_1, y_1)$, находящихся в различных местах плоскости P_1 . Множество выходных световых точек появляется в плоскости P_3 в соответствующих местах, и интенсивность их отражает степень согласования каждого входного сигнала с фильтрующей функцией. Из этого краткого описания очевидны свойства оптического процессора – высокая скорость и параллельность обработки.

Огромное практическое значение имеют вопросы синтеза пространственного согласованного фильтра и оптимизации его параметров.

Из всего многообразия систем пространственной фильтрации (оптических корреляторов) для практического рассмотрения выберем коррелятор с частотной плоскостью (фильтр Вандер Люгта), так как он является первым и самым распространенным оптическим коррелятором.

Допустим, нам надо выяснить, присутствует ли функция $h(x_1, y_1)$ (или ее фрагмент) во входном сигнале $g(x_1, y_1)$. Для этого необходимо получить взаимную корреляцию функций $g(x_1, y_1)$, $h(x_1, y_1)$ на выходе системы. Прежде всего, надо синтезировать в плоскости P_2 пространственный фильтр $H^*(u, v)$, согласованный $h(x_1, y_1)$. Для синтеза этого фильтра во входную плоскость P_1 помещают транспарант с записью функции $h^*(x_1, y_1)$, а в плоскости P_2 записывают интерференционную картину, создаваемую Фурье-образом $H^*(u, v)$ функции $h^*(x_1, y_1)$ и плоской опорной волной по методу Вандер Люгта.

Для настройки элементов схемы и обеспечения визуального контроля картин распределения света в различных плоскостях можно использовать переменный аттенуатор (ПА). Переменный светоделитель (ПСД) позволяет управлять отношением интенсивностей опорного и объектного пучков.

Для выравнивания длин оптических путей, проходимых опорным и объектным пучками, применяются зеркала. Угол между опор-

ным и объектным пучками составляет обычно 15–20°. При этом значении угла пространственные частоты интерференционной картины достаточно высоки. Например, при использовании излучения с $\lambda = 632,8$ мкм пространственная несущая частота равна 412 мм^{-1} . Если же разрешающая способность используемого материала недостаточна для записи столь высоких пространственных частот, между опорным и объектным пучками следует установить меньший угол. При этом для сведения пучков под малым углом требуется светоделитель и дополнительные оптические элементы. Пространственные микроструктурные неоднородности пучка лазера устраняются точечными диафрагмами. Коллимирующая оптика обеспечивает формирование однородных плоских волн.

Если во входной плоскости P_1 поместить транспарант с амплитудным пропусканием $h(x_1, y_1)$, то распределение комплексных амплитуд в плоскости P_2 будет равно $u_2(x_2, y_2) = H(u, v)$, где $H(u, v)$ – комплексный Фурье-образ функции $h(x_1, y_1)$. Таким образом,

$$H(u, v) = \frac{1}{j\lambda f_1} \iint_{-\infty}^{\infty} h(x, y) \exp[-2\pi j(ux + vy)] dx dy.$$

Предположим для простоты, что любые преобразования Фурье, реализуемые с помощью линз, выполняются точно, и, следовательно, пренебрежем постоянными коэффициентами амплитудного пропускания, а также множителем $1/j\lambda f_1$ перед интегралом. Координаты пространственных частот (u, v) в плоскости P_2 связаны с пространственными координатами (x_2, y_2) этой плоскости соотношениями $x_2 = f_1 \lambda u$ и $y_2 = f_1 \lambda v$, где λ – длина волны света; f_1 – фокусное расстояние Фурье-преобразующей линзы.

Распределение комплексных амплитуд, создаваемое в плоскости P_2 опорным пучком с постоянной амплитудой r_0 , который падает на эту плоскость под углом θ к объектному пучку, имеет вид

$$u_r = r_0 \exp(-i2\pi\lambda\alpha x_2),$$

где $\alpha = \sin \theta / \lambda$ – пространственная частота, связанная с наклонным падением опорной волны.

Регистрируемое фотоэмульсией распределение интенсивности в плоскости P_2 будет равно

$$I(x_2, y_2) = (u_2 + u_r)^2 = r_0^2 + |H|^2 + r_0 H \exp(i2\pi\alpha x_2) + r_0 H^* \exp(-i2\pi\alpha x_2).$$

Предположим, что мы осуществили линейную запись распределения интенсивности $I(x_2, y_2)$ и получили голограмму-фильтр, амплитудная прозрачность I которой с точностью до постоянного множителя определяется выражением

$$t(x_2, y_2) = r_0^2 + |H|^2 + r_0 H \exp(i2\pi\alpha x_2) + r_0 H^* \exp(-i2\pi\alpha x_2).$$

Последний член в этом выражении и является искомым членом, пропорциональным H^* .

После того как согласованный пространственный фильтр изготовлен и вновь установлен в плоскости P_2 , мы поворачиваем его на 180° вокруг оптической оси. Благодаря этому корреляция будет наблюдаться под углом $+15...-20^\circ$ по отношению к оси объектного пучка, т.е. в пределах стола, а не под углом $+15...-20^\circ$ в направлении продолжения опорного пучка. При выполнении операции корреляционного анализа опорный пучок блокируется, а во входную плоскость помещают транспарант с амплитудным пропусканием $g_1(x_1, y_1)$. Волна, несущая информацию о корреляции входного и эталонного изображений, будет исходить из плоскости P_2 под углом $+15...-20^\circ$ к оптической оси коррелятора. Вторая Фурье-преобразующая линза L_2 устанавливается вдоль этого направления на фокусном расстоянии от плоскости P_2 , а не вдоль сигнального участка, как это обычно делают. Такая схема коррелятора существенно уменьшает требования к размерам апертур и углам поля зрения используемых линз.

Рассмотрим сначала, что получится, если согласованный пространственный фильтр поворачивать на 180° , а линзу L_2 установить так, чтобы ее оптическая ось совпадала с осью L_1 .

В плоскости P_2 формируется Фурье-образ $G(u, v)$ входной функции $g(x_1, y_1)$. После прохождения этого распределения через согласованный фильтр мы имеем $G(u, v)t(x_2, y_2)$. Поскольку плоскость P_2 совмещена с передней фокальной плоскостью Фурье-преобразующей линзы L_2 с фокусным расстоянием f_2 , распределение комплексных амплитуд в задней фокальной плоскости этой линзы представляет собой Фурье-образ произведения Gt . Таким образом,

$$u_3(x_3, y_3) = r_0^2 g \delta(x, y) + [h \otimes h * g] \delta(x, y) + \\ + r_0 [h \otimes g \otimes \delta(x_3 + \alpha \lambda f_2, y_3)] + r_0 [g * h \otimes \delta(x_3 - \alpha \lambda f_2, y_3)],$$

где $\otimes$ обозначает свертку, а $*$ – корреляцию.

В этом выражении два первых слагаемых соответствуют распределениям, локализованным на оптической оси в выходной плоскости. Третье слагаемое представляет собой свертку $h \otimes g$, которая формируется пучком, исходящим из плоскости P_2 под углом $+\theta$, и это есть распределение в выходной плоскости в окрестности точки с координатами $-\alpha \lambda f_2, 0$.

Распределение, описываемое последним слагаемым с интересующей нас корреляцией $g * h$, создается пучком, исходящим из плоскости P_2 под углом $-\theta$, и локализуется в выходной плоскости вокруг точки с координатами $\alpha \lambda f_2, 0$.

Итак, при установке фотопластинки с записанным на ней согласованным фильтром в плоскости P_2 ее поворачивают на 180° вокруг оптической оси, а линзу L_2 помещают в пучок, выходящий из плоскости P_2 под углом $+\theta$, чтобы сформировать на выходе распределение, описываемое последним слагаемым. Тогда распределение, формируемое в задней фокальной плоскости линзы L_2 на ее оптической оси, представляет собой интересующую нас функцию взаимной корреляции g^*h . Следует заметить, что положение яркого пятна в выходной плоскости P_3 однозначно связано с положением эталонной функции h во входном изображении g ; интенсивность этого пятна указывает на степень соответствия между h и g .

Имеются и другие задачи радиооптики, связанные с обнаружением, корреляционной обработкой информации. Многие понятия и определения радиотехники и оптики одинаковы и различаются лишь количеством переменных.

Функции

$$B(x_1, y_1, z_1; x_2, y_2, z_2) = \langle p(x_1, y_1, z_1) p^*(x_2, y_2, z_2) \rangle$$

называют по определению *корреляционными функциями* комплексных амплитуд волнового поля. В случае $z_1 = z_2$ они являются *поперечными корреляционными функциями* – функциями волновых полей, проходящих через один фильтр. *Продольной корреляционной функцией* – корреляционной функцией одинаковых сигналов – называют функцию, проходящую через разные фильтры.

Для однородного процесса поперечная корреляционная функция

$$B(x_1, y_1, z_1; x_2, y_2, z_2 = z_1) = B(x_2 - x_1, y_2 - y_1, 0).$$

Если процесс однороден только по двум координатам, то поперечная корреляционная функция

$$B = B_z(x_2 - x_1, y_2 - y_1) = B_z(\xi, \eta).$$

Для однородных поперечных функций корреляции

$$B_z(\xi, \eta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} B(f_x, f_y) \exp[j(f_x \xi + f_y \eta) 2\pi] df_x df_y, \quad (11.1)$$

где

$$B(f_x, f_y) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} B_z(\xi, \eta) \exp[-j(f_x \xi + f_y \eta) 2\pi] d\xi d\eta$$

является угловым спектром мощности волнового поля.

Из теории радиоэлектронных систем известно, что спектр сигнала на выходе системы $s_2(\omega) = |k(j\omega)|^2 s_1(\omega)$, где $|k(j\omega)|$ – передаточная характеристика. По аналогии для РОС

$$B(f_x, f_y) = |h(f_x, f_y)|^2 B_{z=0}(f_x, f_y). \quad (11.2)$$

Из (11.1) и (11.2) следует

$$B_z(\xi, \eta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(\xi - \xi', \eta - \eta') B_0(\xi', \eta') d\xi' d\eta',$$

где

$$H(\xi - \xi', \eta - \eta') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |h(f_x, f_y)|^2 e^{j2\pi(f_x\xi + f_y\eta)} df_x df_y.$$

11.1. Амплитудный, фазовый и цифровой фильтры

Вернемся к выражению, связывающему спектры входного $A(f_x, f_y)$ и выходного $B(f_x, f_y)$ сигналов РОС с ее передаточной характеристикой $H(f_x, f_y)$. Для реальных РОС отклик (обратный Фурье-образ передаточной характеристики), а, следовательно, и сама система физически реализуются с учетом aberrаций и шумов, вызванных процессом распространения волны в среде с помехами.

В общем случае

$$H(f_x, f_y) = H_a(f_x, f_y) \cdot H_l(f_x, f_y) H_n(f_x, f_y),$$

где индексы «а», «л», «н» относятся к передаточным характеристикам атмосферы, линзы и носителя, на котором записывается информация.

Согласно схеме формирования изображения в РОС, сигнал с учетом шума

$$s(x, y) = a(x, y) \otimes h(x, y) + n(x, y).$$

Важнейшим элементом схемы обработки является фильтр-транспарант. Характеристикой последнего в области пространственных частот служит комплексная функция

$$\begin{aligned} \Phi(f_x, f_y) &= |\Phi(f_x, f_y)| \times \\ &\times \exp[-j \cdot 2\pi(f_x x + f_y y)] = B \exp[j\varphi(f_x, f_y)]. \end{aligned} \quad (11.3)$$

Идеальным случаем является процедура фильтрации, когда сигнал ошибки $O(x, y) = 0$, т.е. имеют место полное подавление шума, устранение влияния aberrаций и фильтрация искажений, вносимых в сигнал носителем информации. Естественно, что практически добиться этого полностью невозможно, однако в устранении первых двух источников помех достигнуты значительные успехи. Практические приложения пространственной фильтрации волновых полей, помимо распознавания образов, включают улучшение качества изображений (фотографий), повышение их резкости, восстановление изображений, искаженных средой, с известными и в некоторых случаях неизвестными физическими характеристиками и т.д. Качество фильтрации в первую очередь определяется свойствами РОС и фильтра, который управляет амплитудами и фазами пространственного спектра, проходящего через транспарант волнового фронта.

Критерием качества фильтрации является среднеквадратичная ошибка

$$E = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \omega(f_x, f_y) O(f_x, f_y) df_x df_y, \quad (11.4)$$

где $\omega(f_x, f_y)$ – весовая функция ошибок, выбираемая в каждом случае фильтрации с учетом ее цели по критерию приоритета высоких или низких пространственных частот; $O(f_x, f_y)$ – спектр ошибки [12]:

$$O(f_x, f_y) = N(f_x, f_y) |H(f_x, f_y)|^2 + A(f_x, f_y) |1 - H(f_x, f_y)\Phi(f_x, f_y)|^2 \quad (11.5)$$

в предположении, что РОС линейна, а сигнал и шум стационарны и независимы.

Для так называемого винеровского фильтра

$$\Phi(f_x, f_y) = \frac{H^*(f_x, f_y)}{|H(f_x, f_y)|^2 + N(f_x, f_y)/A(f_x, f_y)}. \quad (11.6)$$

Подставив (11.5) в (11.4), получим

$$E = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \omega(f_x, f_y) \frac{A(f_x, f_y)N(f_x, f_y)}{|H(f_x, f_y)|^2 + N(f_x, f_y)/A(f_x, f_y)} df_x df_y. \quad (11.7)$$

В случае отсутствия или низкого уровня шумов ($N(f_x, f_y) \approx 0$) имеем так называемый обратный фильтр

$$\Phi(f_x, f_y) = 1/H(f_x, f_y) \quad (11.8)$$

со среднеквадратичной ошибкой

$$E = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \omega(f_x, f_y) A(f_x, f_y) |1 - H(f_x, f_y)\Phi(f_x, f_y)|^2 df_x df_y. \quad (11.9)$$

Минимальная ошибка достигается при минимальном значении разности $|1 - H(f_x, f_y)\Phi(f_x, f_y)|^2$. Очевидно, что это условие выполняется при $\Phi(f_x, f_y) = H^*(f_x, f_y)$, т.е. когда фазовые множители фильтра по каждой пространственной частоте противоположны по знаку фазовым множителям передаточной функции РОС. Такой фильтр называется *фазовым*, или *чисто фазовым*.

Для нахождения оптимального фазового фильтра приравняем нулю производную от выражения (11.5)

$$(d/d\varphi)[N(f_x, f_y)B^2 + A(f_x, f_y)(1 - H(f_x, f_y)Be^{j\varphi}) \times (1 - H^*(f_x, f_y)Be^{-j\varphi})] = 0. \quad (11.10)$$

Учитывая, что из (11.10) следует $\varphi(f_x, f_y) = \psi(f_x, f_y)$, т.е. указанный выше случай минимизации ошибки фильтрации. Для такого

фильтра

$$E = \int_{-\infty}^{\infty} W(f_x, f_y) \{N(f_x, f_y) + A(f_x, f_y)[1 - |H(f_x, f_y)|^2]\} df_x df_y,$$

т.е. фильтрация шума не реализуется.

Оптимальный амплитудный, или чисто амплитудный, фильтр находится из равенства нулю производной от (11.5) по A при $\varphi = 0$, в результате чего получится фильтр

$$\Phi(f_x, f_y) = B(f_x, f_y) = \frac{\operatorname{Re}\{H(f_x, f_y)\}}{|H(f_x, f_y)|^2 + N(f_x, f_y)/A(f_x, f_y)}, \quad (11.11)$$

как винеровский. Среднеквадратичная ошибка фильтрации

$$E = W(f_x, f_y) \times \int_{-\infty}^{\infty} \frac{A(f_x, f_y) \{N(f_x, f_y) + A(f_x, f_y) |H(f_x, f_y)|^2 \sin^2[\varphi(x, y)]\}}{N(f_x, f_y) + A(f_x, f_y) |H(f_x, f_y)|^2} df_x df_y. \quad (11.12)$$

Выражение (11.12) сводится к виду (11.7) при $\varphi(x, y) \rightarrow 0$, а в случае отсутствия шума – к виду

$$E = \int_{-\infty}^{\infty} W(f_x, f_y) A(f_x, f_y) \sin^2[\varphi(f_x, f_y)] df_x df_y.$$

Амплитудные, фазовые и амплитудно-фазовые (комплексные) фильтры могут быть изготовлены в виде соответствующих транспарантов или синтезированы на ЭВМ в цифровом виде и затем выведены на экран дисплея для фотографирования и изготовления соответствующего транспаранта. При цифровой обработке радиооптической информации они синтезируются в процессе фильтрации. Технологические сложности изготовления фильтров с развитием вычислительной техники, опто- и микроэлектроники, как отмечалось выше, обусловили все большее распространение цифровых и оптико-цифровых методов обработки и фильтрации.

11.2. Оптические и оптико-цифровые процессы

Устройства для выполнения интегральных и других математических преобразований над волновыми фронтами в оптическом диапазоне волн (комплексными функциями) называют *оптическими процессорами*. Управляя амплитудами и фазами пространственных частот распространяющегося в оптической системе волнового фронта, можно создать процессор для выполнения широкого диапазона математических преобразований над закодированной в волновом поле информацией. При этом реализуется огромная скорость (скорость света) обработки, практически неограниченная многоканальность. Рассмотрен-

ные выше линзовые системы и являются оптическими процессорами для выполнения преобразований Фурье, свертки, корреляции, кросс-корреляции в плоскостях. Придавая линзам другую форму (цилиндрическую, коническую и т.д.), можно конструировать оптические процессоры со специальными видами математических преобразований в разных сечениях этих плоскостей. Различают одно- и многолинзовые оптические процессоры, используемые в различных практических приложениях – неразрушающем контроле материалов и изделий, радиолокации, космических исследованиях, биологии, медицине, криминалистике, обработке радио-, акустических, звуковых, речевых сигналов и т.д. Все оптические приборы (бинокли, лупы, микроскопы и др.) являются оптическими процессорами, использующими некогерентное излучение, – некогерентными процессорами. Оптические процессоры, работающие в когерентном излучении, называются *когерентными*. Естественно, что с развитием лазерной техники последним отдают предпочтение, так как они позволяют реализовать фазовую фильтрацию, увеличить разрешение и улучшить другие характеристики оптических систем обработки информации.

Применение ЭВМ для синтеза фильтров и волновых фронтов вообще с применением модуляции и развертки луча лазеров сканирующими системами позволило реализовать гибридные оптико-цифровые когерентные процессоры для обработки информации. Оператор с пульта управления может задавать программу обработки, регулировать смену входных сигналов, синтезировать с помощью ЭВМ входной сигнал на транспаранте-детекторе T_1 (отключая лазер непосредственно или через ЭВМ), фильтр на транспаранте-детекторе T_2 , регистрировать и обрабатывать выходной сигнал с транспаранта-детектора T_3 .

Различные научно-технические приложения оптической и оптико-цифровой обработки информации описаны в обширной специальной литературе.

ЛЕКЦИЯ 12. Пространственная фильтрация и обработка волновых фронтов

При кодировании и обработке информации, передаваемой на различные расстояния посредством распространяющихся в пространстве волновых фронтов, все более широкое применение находят электронные вычислительные машины (ЭВМ) и оптические системы обработки информации (оптические вычислительные устройства).

12.1. Влияние явлений дискретизации на структуру волновых полей

При регистрации волновых фронтов любого диапазона электромагнитных волн представляет интерес амплитудное и фазовое распределения напряженности поля. Идеальным регистрирующим устройством является такое, посредством которого можно зарегистрировать непрерывное распределение амплитуды и фазы поля по плоской, сферической, цилиндрической или другим поверхностям. В дальнейшем для простоты будем говорить о плоской поверхности. Практическое значение имеет следующая общая задача теории и техники волновых процессов.

Имеется комплексный исходный волновой фронт

$$g(x, y) = |g(x, y)|e^{j\varphi(x, y)}$$

в плоскости (xOy) , при помощи модуля $|g(x, y)|$ и фазы $\varphi(x, y)$, которого закодирована некоторая информация. При распространении этого волнового фронта на некоторое расстояние требуется зарегистрировать обусловленную его распространением амплитудно-фазовую структуру поля в плоскости регистрации $\xi O\eta$. Волновой фронт $g(\xi, \eta)$ определится посредством двумерного суперпозиционного интеграла через волновой фронт $g(x, y)$ и импульсный отклик участка свободного пространства между плоскостями xOy и $\xi O\eta$. По виду волнового фронта требуется восстановить вид волнового фронта, т.е. выделить закодированную j информацию. Очевидно, что в случаях, поддающихся аналитическому решению, удобно, если $g(x, y)$ и $g(\xi, \eta)$ связаны каким-либо легко реализуемым обратным математическим преобразованием. Тогда вид волнового фронта $g(x, y)$ однозначно находится по виду волнового фронта $g(\xi, \eta)$ и наоборот. Представляет практический интерес случай, когда $g(\xi, \eta)$ и $g(x, y)$ связаны прямым и обратным двумерными Фурье-преобразованиями.

К сожалению, не существует индикаторов электромагнитного поля, позволяющих одновременно зарегистрировать его распределение по всей плоскости $\xi O\eta$ с необходимой чувствительностью и раз-

решающей способностью в реальном масштабе времени. Поэтому производится выборочная регистрация поля в ряде дискретных точек этой плоскости. Дискретизация волновых фронтов вносит в их структуру некоторые искажения, которые необходимо учитывать и устранять.

Определим вспомогательную двумерную функцию, называемую гребенчатой, следующим образом:

$$\text{comb}x\text{comb}y = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(x-m)\delta(y-n), \quad (12.1)$$

где m и n – целые числа единиц расстояния вдоль осей x и y соответственно. Очевидно, что прямоугольную сетку выборок функций $g(x, y)$ с помощью гребенчатой функции (12.1) можно представить в виде

$$g_b(x, y) = g(x, y) \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(x-ma)\delta(y-nb), \quad (12.2)$$

где a и b – периоды выборки вдоль осей x и y соответственно.

Выражение (12.2) может иметь вид

$$g_b(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} g(ma, nb)\delta(x-ma)\delta(y-nb). \quad (12.3)$$

Согласно интегралу свертки, свойству гребенчатой функции со сдвигом по аргументу, симметрии и фильтрующему свойству δ -функции, получим:

$$\Phi\{g_b(x, y)\} = \frac{1}{ab} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} G\left(\xi - \frac{m}{a}, \eta - \frac{n}{b}\right), \quad (12.4)$$

где $G(\dots)$ – Фурье-образ функции $g(x, y)$ по пространственным частотам ξ, η .

Таким образом, при дискретизации исходного волнового фронта его Фурье-образ с точностью до множителя $1/ab$ становится набором дискретных Фурье-образов дискретизированного волнового фронта, сдвинутых в плоскости частот на целое число величин $1/a, 1/b$. Слагаемые не перекрываются, если выполняются соотношения

$$a \leq \frac{1}{\xi_{\text{макс}}}, \frac{b}{\eta_{\text{макс}}}, \quad (12.5)$$

где $\xi_{\text{макс}}, \eta_{\text{макс}}$ – максимальные пространственные частоты исходного волнового фронта.

Если периоды выборок взяты согласно (12.5) (что возможно, лишь для волновых фронтов с существующими максимальными пространственными частотами, т.е. для всегда имеющих место в практике пространственно-ограниченных волновых фронтов), из выражения (12.4), используя понятие прямоугольной функции и фильтрующие свойства δ -функции, находим:

$$g(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} g\left(\frac{m}{\xi_{\max}}, \frac{n}{\eta_{\max}}\right) \times \frac{\sin \pi(\xi_{\max} x - m)}{\pi(\xi_{\max} x - m)} \cdot \frac{\sin \pi(\eta_{\max} y - n)}{\pi(\eta_{\max} y - n)}. \quad (12.6)$$

Это означает, что по дискретным значениям волнового фронта $g(\xi, \eta)$ (первый множитель под знаком двойной суммы), полученным в плоскости (ξ, η) , можно восстановить вид исходного волнового фронта $g(x, y)$ посредством функции вида $\frac{\sin \varphi}{\varphi}$, построенной в каждой точке выборки. На практике число отсчетов m и n всегда ограничено, поэтому точного значения волнового фронта $g(x, y)$ по выборкам воспроизвести нельзя. Однако это можно сделать со сколь угодно близким приближением путем увеличения числа выборок $g\left(\frac{m}{\xi_{\max}}, \frac{n}{\eta_{\max}}\right)$.

12.2. Анализ и синтез волновых полей на ЭВМ

Учитывая (12.6), связь между $g(\xi, \eta)$ и $g(x, y)$ при помощи прямого преобразования Фурье и таблицы Фурье-образов специальных функций, получаем

$$g(x, y) = \begin{cases} \frac{1}{\xi_{\max} \eta_{\max}} \sum_{m=-\infty}^m \sum_{n=-\infty}^n g\left(\frac{m}{\xi_{\max}}, \frac{n}{\eta_{\max}}\right) e^{-j2\pi\left(\frac{m\xi}{\xi_{\max}} + \frac{n\eta}{\eta_{\max}}\right)} \\ \text{при } \left|\frac{\xi}{\xi_{\max}}\right| \leq \frac{1}{2}, \left|\frac{\eta}{\eta_{\max}}\right| \leq \frac{1}{2} \\ 0 \text{ в остальных случаях} \end{cases}, \quad (12.7)$$

т.е. другое представление анализируемого волнового фронта через его выборочные значения.

До сих пор мы ограничивали волновой фронт $g(\xi, \eta)$, т.е. тот, по которому требуется восстановить исходный волновой фронт $g(x, y)$. Если же наложить ограничения на размеры волнового фронта $g(x, y)$,

т.е. при $-\frac{A}{2} \leq x \leq \frac{A}{2}, -\frac{B}{2} \leq y \leq \frac{B}{2}$, можно получить

$$g(x, y) = \frac{1}{AB} \text{rect} \frac{x}{A} \text{rect} \frac{y}{B} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} G\left(\frac{m}{A}, \frac{n}{B}\right) e^{j2\pi\left(\frac{mx}{A} + \frac{ny}{B}\right)}, \quad (12.8)$$

где $G(\dots)$ – Фурье-образ волнового фронта $g(x, y)$.

Выражения (12.7) и (12.8) лежат в основе анализа и синтеза волновых полей. Реализация операции (12.8) является задачей анализа,

операции (12.8) – задачей синтеза. Наиболее широкое развитие задачи анализа и синтеза получили в теории и технике антенн применительно к задачам построения антенн по заданной диаграмме направленности и к расчету поля излучения в дальней зоне по заданному амплитудно-фазовому распределению поля в раскрыве. До недавнего времени развитие теории анализа и синтеза волновых полей на ЭВМ шло в основном по пути использования теории аппроксимации аналитических функций, итерационных и вариационных методов решения интегральных уравнений и т.д.

На практике, как правило, встречаются случаи, когда вид функций $g(x, y)$ и $g(\xi, \eta)$ не позволяет установить аналитическую взаимосвязь между ними даже с использованием приближенных методов, развитых в теории синтеза антенн. Поэтому необходима численная реализация соотношений (12.7) и (12.8) на ЭВМ, в которых ξ, η, x, y меняются дискретным образом, а число точек отсчетов ограничено. Соотношение (12.8) означает, что для пространственно-ограниченного волнового фронта его вид может быть восстановлен из выборок пространственно-частотного спектра, взятых через интервалы $1/A, 1/B$. Из соотношения (12.7) следует, что для волнового фронта с ограниченным пространственным спектром последний можно найти по выборкам спектра, взятым через интервалы $1/\xi_{\max}, 1/\eta_{\max}$.

При реализации соотношений (12.7) и (12.8) на ЭВМ время расчета может быть большим, поэтому разрабатываются различные пути его уменьшения. Одним из таких путей является вычисление $g(\xi, \eta)$ только в точках отсчета $\xi = \frac{p}{A}, \eta = \frac{q}{B}$, где p и q – целые числа. При этом выражение (12.7) имеет вид

$$g\left(\frac{p}{A}, \frac{q}{B}\right) = \frac{1}{\xi_{\max} \eta_{\max}} \sum_{m=-M/2}^{M/2} \sum_{n=-N/2}^{N/2} g\left(\frac{m}{\xi_{\max}}, \frac{n}{\eta_{\max}}\right) e^{-j2\pi\left(\frac{mp}{\xi_{\max} A} + \frac{nq}{\eta_{\max} B}\right)}, \quad (12.9)$$

где $M \times N$ – число отсчетов ограниченной в пространстве функции $g(x, y)$. Выражение (12.9) является дискретным Фурье-преобразованием. Общее число выборок волнового фронта $g(x, y)$:

$$L = M \times N = \frac{A}{a} \cdot \frac{B}{b} = A \xi_{\max} B \eta_{\max},$$

общее число выборок $g(\xi, \eta)$ есть

$$L' = \frac{\xi_{\max}}{1/A} \cdot \frac{\eta_{\max}}{1/B} = A \xi_{\max} B \eta_{\max}.$$

При расчетах на ЭВМ $L \times L$ дискретных значений выборок $g(\xi, \eta)$ требуется L^4 машинных операций. Прямые методы вычислений становятся очень продолжительными. В настоящее время существ-

вуют программы для проведения быстрых преобразований Фурье (БПФ), позволяющие резко сократить число необходимых операций с L^4 до $4L^2 \log_2 L$, что особенно эффективно при больших L . С помощью БПФ Фурье-образ выборки с 1000×1000 элементами вычисляется за несколько минут, в то время как прямое вычисление проводилось бы около года.

В дальнейшем под синтезом волновых полей будем понимать измерение или расчет физических или геометрических параметров источника по излучаемому или рассеиваемому волновому фронту, под анализом – нахождение излучаемого или рассеиваемого источником волнового фронта по его физическим и геометрическим свойствам, заданным математически или путем физического описания. При решении этих задач с помощью ЭВМ возможно производить промежуточную коррекцию получаемых результатов, позволяющую учесть различные искажающие факторы на любой стадии анализа и синтеза, осуществлять ввод исходных данных ввода и вывода результатов в любом удобном виде (таблицы, графики, текст и т.д.), синтезировать и анализировать как реальные, так и воображаемые источники волновых полей, осуществлять более широкий круг преобразований, чем с помощью реальных технических устройств. В работе изложен современный уровень развития научных исследований в области разработки и применения ЭВМ для анализа и синтеза волновых полей.

12.3. Поиск и распознавание образов

Новым направлением в кибернетике является в настоящее время распознавание образов, т.е. объектов, явлений, ситуаций. Это направление возникло сравнительно недавно при разработке читающих автоматов, медицинских диагностических машин, устройств, прогнозирующих метеорологические процессы, и т.д.

Первые попытки распознавания объектов были основаны на распознавании и сравнении свойств объектов с эталонами, а многие последующие – на использовании принципов построения биологических систем.

Теория распознавания оперирует с «образами», подразделяющимися на «объекты», «явления» и «ситуации». Образами могут быть звуки речи, подводные лодки, космические корабли, прогноз погоды, знаки текста, диагностика заболеваний и т.п.

Признаки образов вводятся с помощью датчиков-рецепторов, которые являются неотъемлемой частью распознающего устройства и могут быть оптическими, акустическими, радиолокационными, гидравлическими, электромагнитными, химическими и т.д. В распознающем устройстве эти признаки обрабатываются, и автоматически принимается решение, к какому классу следует отнести образ.

Все объекты по основным признакам делятся на классы. Не все признаки дают одинаковый результат при распознавании. Использование отдельного признака не позволяет добиться того результата, который можно получить, применяя одновременно всю совокупность признаков.

Достоверность распознавания в основном определяется точностью измерения признаков: с увеличением точности достоверность возрастает; при увеличении числа признаков, одновременно используемых для распознавания, достоверность также увеличивается. Иногда распознавание можно рассматривать как статистическую задачу, в которой вероятность распознавания определяется дифференциальными законами распределения признаков объектов и помех.

В теории распознавания любой объект может иметь бесконечно большое число признаков. Признаки делятся на объективные и описательные. Объективные признаки можно выразить количественно, к описательным относятся признаки, которые количественно охарактеризовать нельзя. Описательные признаки делятся для удобства на градации (группы), каждой из которых присваивается то или иное числовое значение. Признаки схожих объектов количественно незначительно отличаются друг от друга.

Распознавание объекта можно определить как сравнение его признаков с признаками одного или нескольких заранее определенных классов для отнесения объектов к тому или иному классу. Существуют определенные правила количественной классификации распознавания. Использование тех или иных правил определяется конкретными задачами распознавания. Правильность установления величины количественного критерия распознавания определяется практикой. По мере накопления сведений о правильно распознанных объектах можно установить оптимальную величину, распознающего порога, причем самое важное при выборе порога – получить максимально верное распознавание классов.

Очень часто в теории распознавания приходится разделять объекты на 2 класса, т.е. использовать дихотомию. Дихотомия – это деление всего алфавита классов объектов на два противоречащих друг другу видовых понятия, из которых одно представляет отрицание признаков другого. В качестве воспринимающих устройств (рецепторов) системы обычно применяют радиотехнические, оптико-электронные, акустические и оптические устройства, позволяющие заблаговременно обнаружить объекты на больших дальностях.

В устройстве сравнения значение поступающих реализаций сравнивается в каждый момент времени с описанием всех классов – эталонов, которое хранится в блоке памяти. Сопоставление найденных вероятностных характеристик по заданному критерию и отнесе-

ние объекта к одному из эталонных классов выполняются в блоке принятия решения.

При распознавании образов голографическими методами широко используются согласованные фильтры. Пространственный фильтр называется согласованным с сигналом $g(x, y)$, если его импульсный отклик имеет вид $h(x, y) = g^*(-x, -y)$. При подаче на вход фильтра сигнала $g(x, y)$ на его выходе появится сигнал

$$f(x, y) = \iint_{-\infty}^{\infty} g(\xi, \zeta) g^*(\xi - x, \zeta - y) d\xi d\zeta,$$

т.е. кросс-корреляционная функция функций g и g^* . Если этот отклик заменить его Фурье-образом и расположить в системе обработки в частотной плоскости второй линзы, а в плоскости x_1Oy_1 поместить транспарант с коэффициентом пропускания $g(x, y)$, за фильтром частотный спектр будет пропорционален

$$G\left(\frac{x_2}{\lambda F}, \frac{y_2}{\lambda F}\right) \cdot G^*\left(\frac{x_2}{\lambda F}, \frac{y_2}{\lambda F}\right),$$

т.е. действительной величине $|G|^2$. Это значит, что на последнюю линзу падает плоская волна и линза фокусирует ее в точку – фокус линзы. Если же входной сигнал отличен от $g(x, y)$, свет не будет собираться в яркую точку фокуса последней линзы. Если входной сигнал смещен относительно оптической оси, яркая точка в фокальной плоскости последней линзы тоже смещается на соответствующее расстояние. Указанное свойство согласованного фильтра может использоваться для обнаружения среди всех входных сигналов сигнала, с которым фильтр согласован, т.е. для распознавания образов.

На вход системы распознавания подаются одновременно или последовательно распознаваемые волновые фронты. На выходе каждого из фильтров, входящих в систему, находится устройство для нормировки отклика к полной энергии, приходящей от образа, с которым фильтр согласован (различным распознаваемым образам может соответствовать различный уровень энергии). Нормировка осуществляется как электронными методами путем определенных оптико-электронных преобразований, так и оптическим путем создания в выходной плоскости соответствующего фона. При наличии на входе сигнала, согласованного с одним из фильтров G^* , на выходе последнего будет наибольший сигнал.

Очевидно, что таким согласованным фильтром может быть фильтр Вандер Люгта, так как на выходе оптической системы, включающей его, формируется в соответствующей точке кросс-корреляционный сигнал. Для согласованной фильтрации следует синтезировать фильтры по методу Вандер Люгта и поместить их в кана-

лах последней схемы. Оpoznаваемые сигналы необходимо подавать на вход последовательно. На одном транспаранте можно изготовить несколько фильтров. При этом одна пространственная несущая частота опорной волны модулируется волновыми фронтами от различных объектов. Можно получить фильтр также путем использования различных пространственных опорных частот. При изготовлении фильтра таким образом, волновые фронты разных объектов составляют различные углы с опорной волной, поэтому при фильтрации кросс-корреляционная функция будет появляться под различными углами к оптической оси фильтрующей системы в плоскости наблюдения. На одном транспаранте Вандер Люгту удавалось записать до девяти фильтров.

Для практической реализации фильтрации важна точность установки фильтров и подачи входных сигналов, подвергающихся Фурье-преобразованиям. При фильтрации (опознавании) сходных объектов возможны некоторые ошибки.

При фильтрации сигнала в присутствии известного однородного шума $g(x, y) = S(x, y) + n(x, y)$ на изготовленный фильтр контактно накладывается диапозитив со спектральной плотностью шума. В результате в выходной плоскости появляется яркая точка, свидетельствующая о присутствии сигнала и указывающая его положение на входной плоскости коррелятора.

В реальном случае полученное отношение сигнал–шум заметно ниже теоретического. Это происходит в основном из-за шумов фотоэмульсии транспаранта, на которой записан фильтр, фазовых шумов фотоматериала, а также ошибок фотообработки, юстировки и aberrаций оптической системы.

Наилучшие результаты получаются с чисто амплитудным фильтром. Эксперимент показал, что распознавание методом согласованной оптической фильтрации менее критично к деформации формы и размерам объекта, а также к его повороту. Применение оптимальной фильтрации перспективно при распознавании букв негативного текста, отпечатков пальцев, поиска карточек в картотеках и т.д.

ЛЕКЦИЯ 13. Расчеты геометрических параметров голографических изображений

Рассмотрим процесс записи голографического изображения точечного источника сферической волны, расположенного в точке с координатами (x_0, y_0, z_0) относительно произвольной плоскости декартовой системы координат (x, y, z) , где регистрируется голограмма. Координаты источника опорной сферической волны выберем также фиксированными, но произвольными относительно этой плоскости (x_i, y_i, z_i) . Если обозначить длину волны излучения λ , то в квадратичном приближении Френеля суммарное распределение волнового поля определится выражением

$$g(x, y) = A \exp \left\{ j \frac{\pi}{\lambda z_i} [(x - x_i)^2 + (y - y_i)^2] \right\} + a \exp \left\{ j \frac{\pi}{\lambda z_0} [(x - x_0)^2 + (y - y_0)^2] \right\}, \quad (13.1)$$

а уравнение голограммы примет вид

$$J(x, y) = |A|^2 + |a|^2 + Aa^* \exp \left\{ j \frac{\pi}{\lambda z_i} [(x - x_i)^2 + (y - y_i)^2] \right\} + \exp \left\{ -j \frac{\pi}{\lambda z_0} [(x - x_0)^2 + (y - y_0)^2] \right\} + A^* a \exp \left\{ -j \frac{\pi}{\lambda z_i} [(x - x_i)^2 + (y - y_i)^2] \right\} + \exp \left\{ j \frac{\pi}{\lambda z_0} [(x - x_0)^2 + (y - y_0)^2] \right\}. \quad (13.2)$$

где A и a – комплексные амплитуды источников опорной и восстанавливающей волн.

В пропорциональном коэффициенте пропускания голограммы $j(x, y)$ нас интересуют слагаемые, за счет перемножения которых с восстанавливающей волновой длиной λ , так же в квадратичном приближении Френеля

$$B(x, y) = B \exp \left\{ j \frac{\pi}{\lambda z_b} [(x - x_b)^2 + (y - y_b)^2] \right\}. \quad (13.3)$$

Геометрическая схема восстановления аналогична схеме записи, с той лишь разницей, что в плоскости xuz расположена голограмма с коэффициентом пропускания $T(x, y) = \alpha \tau |J(x, y)|^2$, в произвольной точке (x_g, y_g, z_g) расположен источник восстанавливающей волны длиной λ_i , а объект и опорный источник отсутствуют.

В произвольной точке (x, y) плоскости волна от реконструированного из голограммы объекта в приближении Френеля для точечно-

го источника должна определяться выражением

$$u(x, y)|u(x, y)|\exp\left[j\frac{\pi}{\lambda_1 z_{\text{н}}}(x^2 + y^2)\right]\exp\left[-j\frac{2\pi}{\lambda_i z_{\text{н}}}(x_{\text{н}}x + x_{\text{н}}y)\right], \quad (13.4)$$

поэтому должны выполняться тождественные равенства

$$u'(x, y) = C'(x, y) \text{ и } u''(x, y) = C''(x, y). \quad (13.5)$$

Подставим в (13.5) выражения (13.5) и учтем (13.3). Затем преобразуем экспоненты двух тождественных равенств, получаем координаты восстанавливаемых ДИ и МИ точечного объекта:

$$\begin{aligned} z'_{\text{н}} &= \frac{1}{(1/z_0) - (\lambda_{\text{в}}/\lambda z_0) + (\lambda_{\text{в}}/\lambda z_i)}; \\ z''_{\text{н}} &= \frac{1}{(1/z_{\text{в}}) - (\lambda_{\text{в}}/\lambda z_0) + (\lambda_{\text{в}}/\lambda z_i)}; \\ x'_{\text{н}} &= \frac{\lambda_{\text{в}} z'_{\text{н}}}{\lambda z_0} x_0 + \frac{\lambda_{\text{в}} z'_{\text{н}}}{\lambda z_i} x_j - \frac{z'_{\text{н}}}{z_{\text{в}}} x_{\text{в}}; \\ x''_{\text{н}} &= -\frac{\lambda_{\text{в}} z''_{\text{н}}}{\lambda z_0} x_0 + \frac{\lambda_{\text{в}} z''_{\text{н}}}{\lambda z_i} x_j - \frac{z''_{\text{н}}}{z_{\text{в}}} x_{\text{в}}; \\ y'_{\text{н}} &= \frac{\lambda_{\text{в}} z'_{\text{н}}}{\lambda z_0} y_0 + \frac{\lambda_{\text{в}} z'_{\text{н}}}{\lambda z_i} y_j - \frac{z'_{\text{н}}}{z_{\text{в}}} y_{\text{в}}; \\ y''_{\text{н}} &= -\frac{\lambda_{\text{в}} z''_{\text{н}}}{\lambda z_0} y_0 + \frac{\lambda_{\text{в}} z''_{\text{н}}}{\lambda z_i} y_j - \frac{z''_{\text{н}}}{z_{\text{в}}} y_{\text{в}}; \end{aligned} \quad (13.6)$$

где $x'_{\text{н}}, y'_{\text{н}}, z'_{\text{н}}; x''_{\text{н}}, y''_{\text{н}}, z''_{\text{н}}$ – координаты восстанавливаемых из голограммы.

Нетрудно видеть, что изменения координат объекта вызывают изменения координат изображения:

$$M_x = M_y = \frac{|\Delta x_{\text{н}}|}{|\Delta x_0|} = \frac{|\Delta y_{\text{н}}|}{|\Delta y_0|} = \frac{\lambda_{\text{в}} z_{\text{н}}}{\lambda z_0}, \quad (13.7)$$

где $z_{\text{н}}$ определяется формулами (13.6). Для одного из изображений

$$\begin{aligned} M' &= \left| \frac{1}{1 - (z_0/z_i) - (\lambda z_0/\lambda_{\text{в}} z_{\text{в}})} \right|, \\ M'' &= \left| \frac{1}{1 - (z_0/z_i) - (\lambda z_0/\lambda_{\text{в}} z_{\text{в}})} \right|. \end{aligned} \quad (13.8)$$

При использовании плоской опорной и восстанавливающей волн $z_i = \infty, z_{\text{в}} = \infty$ увеличения $M' = M'' = 1$, хотя $\lambda \neq \lambda_{\text{в}}$. Однако, $z'_{\text{н}}$ и $z''_{\text{н}}$ выражаются формулами

$$z'_{\text{н}} = \frac{\lambda_{\text{в}}}{\lambda} z_0; \quad z''_{\text{н}} = \frac{\lambda_{\text{в}}}{\lambda} z_0. \quad (13.9)$$

Изменение координаты объекта дает изменение координаты изображения, откуда увеличение в продольном направлении.

Пусть объект представляет собой четыре светящиеся точки. Его голограмму получим с помощью плоской опорной волны. Проведем математическое вычисление для одной точки.

Голограмма точечного объекта представляет математический интерес, так как волновой фронт, отражаемый или излучаемый объектом, можно рассматривать как сумму волновых фронтов, излучаемых каждой отдельной точкой, на которые можно разбить объект.

Обозначим угол падения сферической волны на плоскость голограммы в произвольной точке через θ , расстояние от объекта до этой плоскости – через H , а путь, проходимый сферической волной до произвольной точки плоскости голограммы, – через R . Тогда уравнение голограммы (переменная интенсивность) запишем в виде

$$I = A^2 + \frac{a^2}{R^2} + 2\frac{Aa}{R} \cos(kR - kH),$$

где A и a – амплитуды опорной и объектной волн соответственно. Коэффициент прозрачности голограммы – T .

Распределение поля на неосвещенной стороне голограммы, если для восстановления изображения использовать ту же волну A , равно

$$\begin{aligned} E &= A\sqrt{T} = A\sqrt{1 - \alpha\tau} \approx A\left(1 - \frac{1}{2}\alpha\tau\right) = \\ &= A\left[1 - \frac{1}{2}\alpha\tau\left(A^2 + \frac{a^2}{R^2}\right)\right] - A^2\frac{\alpha\tau}{R}a \cos(kR - kH). \end{aligned}$$

Первое из слагаемых есть плоская волна, прошедшая через голограмму без искажения волнового фронта, а второе слагаемое

$$\frac{A^2\alpha\tau}{R} \cos(kR - kH) = \frac{A^2\alpha\tau}{R} \left[\frac{e^{jk(R-H)}}{R} + \frac{e^{-jk(R-H)}}{R} \right]$$

представляет собой расходящуюся и сходящуюся световые волны. Источник расходящейся сферической волны находится в точке $k(R-H) = 0$, $R = H$, т.е. на расстоянии H от голограммы, источник сходящейся волны – в точке $R = -H$, т.е. в точке с обратной стороны голограммы на расстоянии H . С точностью до постоянного множителя $1/2\alpha\tau A^2$, где α – постоянная, характерная для пленки, на которой регистрируется голограмма; τ – время экспозиции, расходящаяся и сходящаяся сферические волны дают зарегистрированный голограммой волновой фронт голографируемого точечного источника сферической волны. Аналогично вычисляются остальные три точки. При освещении этой голограммы плоской восстанавливающей волной получим сопряженные изображения каждой из этих точек.

Это и дает ощущение реальности предмета. Кроме того, перспектива изображения меняется при перемещении наблюдателя из исходного положения к последующим.

Таким образом, сквозь голограмму, как сквозь ограничивающую прозрачную структуру, наблюдается МИ предмета, ДИ как бы вывернуто по отношению к исходному предмету. Такое изображение называется псевдоскопическим.

Следует отметить еще одну особенность голографического изображения. Как видно, при использовании плоских опорной и восстанавливающей волн, что приводит к искажению формы восстановленного изображения в продольном направлении. Предмет либо сплюснут вдоль оси z ($\lambda < \lambda_{\text{в}}$), либо вытянут $\lambda > \lambda_{\text{в}}$.

Укажем на возможность получения цветных голограмм и восстановления цветных изображений. Как известно, ощущение цветовой реальности изображений можно передать, используя сочетания трех цветов: красного, желтого и синего, причем $\lambda_{\text{кр}} \neq \lambda_{\text{жел.}} \neq \lambda_{\text{син.}}$. Поэтому освещение голограммы, полученной в монохроматическом свете, смесью этих цветов даст три изображения, расположенных на различных расстояниях $z_{\text{и}}$ от голограммы. Если же записать голограмму, применив смесь этих цветов, для каждого цвета получим свою интерференционную картину, т.е. сама голограмма не будет цветной. Осветим ее смесью этих цветов. Каждый цвет даст три изображения. Поскольку для одной из длин волн смеси в голограмме есть только одна голограмма с такой же длиной волны, то из всех девяти изображений только три будут совмещены в точке $z_{\text{и}} = z_0$, $x_{\text{и}} = x_0$, $y_{\text{и}} = y_0$, соответствующей положению объекта при получении голограммы. Они и дадут нам цветные изображения объекта. Остальные шесть монохроматических изображений будут разнесены под различными углами к оси z .

Таким образом, освещая определенную часть голограммы восстанавливающей волной, получаем изображения объекта с различной перспективой.

Если в схеме голографирования точечного источника со сферической опорной волной предположить, что $z_i = z_0 = z$, то суммарную интенсивность поля в плоскости голограммы можно привести к виду

$$I(x, y) = A^2 + |a(x_0, y_0)|^2 + 2A|a(x, y)| \times \\ \times \cos \left[2\pi \frac{x_0 - x_i}{\lambda z} x + 2\pi \frac{y_0 - y_i}{\lambda z} y + \varphi(x_i, y_i, x_0, y_i) \right],$$

где $\varphi(x_i, y_i, x_0, y_i)$ не зависит от x, y . Это выражение можно рассматривать как Фурье-преобразование исходного волнового фронта с пространственными частотами

$$f_x = \frac{x_0 - x_i}{\lambda z}; \quad f_y = \frac{y_0 - y_i}{\lambda z}.$$

При освещении такой голограммы плоской восстанавливающей волной оба изображения находятся в бесконечности. Чтобы их сфокусировать, за голограммой следует расположить собирающую линзу.

Известно, что при оценке разрешающей способности оптических приборов пользуются критерием Рэля: две спектральные линии считаются разрешенными, если максимум интенсивности для одной длины волны λ_1 совпадает с минимумом для другой λ_2 . В этом случае при равной интенсивности глубина провала между максимумами составляет 20%. Естественно, что критерий Рэля условен, так как в отдельных случаях можно обнаружить и меньший провал.

Применительно к голографии теоретические расчеты показывают следующее:

- наилучшая разрешающая способность имеет место при расположении объекта в центре голограммы;
- при увеличении размеров голограммы линейная разрешающая способность стремится к предельному значению $\lambda/2$;
- при голографировании объемных объектов более близкие к голограмме участки разрешаются лучше, чем более удаленные;
- при малых размерах голограммы и расположения объекта на оси, выходящей из ее центра, продольная разрешающая способность ухудшается пропорционально квадрату расстояния между объектом и голограммой;
- разрешающая способность в продольном направлении имеет одинаковое значение с поперечным, если расстояние между голограммой и объектом в два раза меньше размеров голограммы.

13.1. Требования к когерентности и типы голограмм

Для получения устойчивой интерференционной картины голограммы требуется применение высококогерентных источников излучения. Нарушение когерентности лазерного излучения может происходить из-за различия во времени испускания света атомами активного вещества и из-за разной длины оптического пути, идущего от различных его участков.

Временная (частотная) когерентность. Излучение света возбужденным атомом активного вещества происходит не мгновенно, а за некоторое конечное время τ . Следовательно, ему соответствует определенный спектр частот с шириной $\Delta\nu \sim 1/\tau$. Связь между скоростью распространения электромагнитной волны, ее периодом (частотой) и длиной имеет вид $\lambda = cT = c/\nu$. Для источников высокомонохроматических

$$\frac{1}{2} \lambda \Delta\nu = \frac{1}{2} \nu \Delta\lambda, \quad (13.10)$$

где $\Delta\lambda = \lambda_2 - \lambda_1$; $\Delta\nu = \nu_1 - \nu_2$. Отсюда $\Delta\lambda = \lambda^2 \Delta\nu / c$.

На расстоянии $L = \lambda^2 / 2\Delta\lambda = c / 2\Delta\nu$ согласно (13.10) разность хода лучей составит половину длины волны. Таким образом, максимум интерференционной картины на одном краю спектра совпадает с минимумом ее на другом, и интерференционной картины не получается. Обычно расстояние, на котором исчезает интерференционная картина, определяют путем сравнения разности хода волн, соответствующих середине спектра и его краю, т.е. вместо $\Delta\lambda$ берут $\Delta\lambda/2$. Поэтому интерференционная картина исчезает на вдвое большем расстоянии $c / \Delta\nu = \lambda^2 / \Delta\lambda = L_{\text{ког}}$, называемом длиной когерентности. Следовательно, голографируемый объект должен располагаться на расстоянии от голограммы, меньшем длины когерентности. Кроме того, эти же ограничения вводятся и на глубину голографируемой сцены. Для повышения временной когерентности источников (увеличения длины когерентности) необходимо сузить спектр их излучения.

Пространственная когерентность источника. Она оценивается качеством интерференционной картины между лучами, исходящими из различных его точек. Чем меньше разность хода между этими лучами, тем больше пространственная когерентность.

Для получения устойчивости интерференционной картины разность хода между крайними лучами объекта, падающими на голограмму, не должна превышать длины когерентности лазера, т.е. $L/A = \sin \theta$ или $A = L/\sin \theta$, где A – размер голограммы, θ – угловые размеры источника.

Отсюда следует, что при голографировании с использованием источника со слабыми требованиями на его когерентность размеры голограммы очень малы, что ведет к снижению ее разрешающей способности и ухудшению качества изображения. Требования на временную когерентность источника должны быть высокими.

Голографирование можно осуществлять в пространственно-некогерентном свете. Для этого свет, исходящий из конкретной точки объекта, нужно разделить на две части. Вследствие этого каждая из частей будет иметь устойчивую разность фаз. Пара лучей, исходящих из другой точки объекта, также будет иметь свою постоянную разность фаз и, следовательно, устойчивую интерференционную картину, так как свет от различных точек некогерентно-освещенного объекта не будет давать устойчивой интерференционной картины.

Некогерентная голография имеет тот недостаток, что каждая голограмма этой общей голограммы, соответствующая точке голографируемого объекта, получается путем интерференции очень слабых световых пучков.

ПРАКТИЧЕСКАЯ ЧАСТЬ

ЛАБОРАТОРНАЯ РАБОТА № 1. Основные характеристики объемных голограмм

Цель работы: знакомство с основными характеристиками трехмерных голограмм и их экспериментальным определением.

Объект исследования: пропускающие голограммы-решетки, имеющие различную толщину и зарегистрированные на одной пространственной частоте.

Задачи, решаемые в работе:

1. Ознакомиться с методикой измерения и оценки основных параметров голограммы.
2. Измерить и оценить параметры голограмм-решеток, имеющих одинаковый период, но различную толщину.
3. Оценить толщину и селективность виртуальных голограмм с различной пространственной частотой и степенью объемности.
4. Проанализировать полученные результаты с точки зрения влияния толщины и пространственной частоты голограммы-решетки на ее селективность и степень объемности; результаты анализа оформить в виде вывода.

СВЕДЕНИЯ ИЗ ТЕОРИИ

Перед началом работы необходимо ознакомиться с определениями и значениями следующих терминов, приведенных в терминологическом словаре данного пособия: голограмма, пропускающая голограмма, двумерная голограмма, трехмерная голограмма, фазовая голограмма, селективность, условие Брэгга, дифракционная эффективность, пространственная частота голограммы.

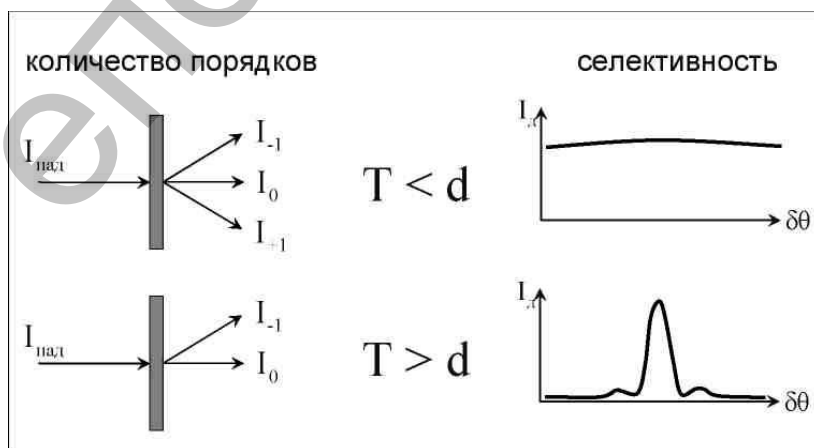


Рис. 1. Влияние толщины регистрирующей среды на свойства голограмм (селективность продемонстрирована на примере угловой селективности).

Трехмерные голограммы, толщина (Т) которых намного превышает их период ($T \gg d$), обладают характерными свойствами, обуславливающими их широкое применение и использование; в первую очередь, это – формирование только одного порядка дифракции и селективность. На рис. 1 схематически показано основное различие двумерных и трехмерных голограмм, определяемое соотношением между толщиной голограммы и ее периодом.

При рассмотрении свойств трехмерных голограмм именно толщина является одним из определяющих параметров – в зависимости от величины Т различают тонкослойные трехмерные голограммы (с толщиной менее 20 мкм) и объемные голограммы (с толщиной более 1 мм). Толщина и пространственная частота голограммы определяют степень объемности голограммы – важный для практики параметр. Как правило, вопрос о степени объемности голограммы решается экспериментально по наличию +1-го и высших порядков дифракции и их эффективности. В качестве теоретического критерия степени объемности для голограмм-решеток используют параметр Клейна

$$Q = 2\pi\lambda T / (nd)^2, \quad (1)$$

где λ – длина волны излучения, Т – толщина голограммы, d – пространственный период голограммы, иначе, период зарегистрированной интерференционной картины ($d=1/v$, где v – пространственная частота голограммы), n – средний показатель преломления голограммы (для голограмм, используемых в работе, $n=1,5$). Объемные голограммы обладают высокой селективностью – угловой ($\Delta\theta$) и спектральной ($\Delta\lambda$). Чем выше величина Т·v для данной голограммы, тем выше ее селективность и меньше значения $\Delta\theta$ и $\Delta\lambda$. Табл. 1 иллюстрирует основные различия между двумерными, трехмерными тонкослойными и объемными голограммами.

Исследование селективности голограмм в данной работе производится на примере исследования угловой селективности – проводится измерение контура угловой селективности голограмм (-1)-го порядка дифракции и дается оценка угловой селективности данной голограммы ($\Delta\theta$).

Таблица 1

Сравнение свойств двумерных и трехмерных голограмм

Свойство голограммы	Двумерная голограмма (2D)	Трехмерная голограмма (3D)	
		Тонкослойная	Объемная
Количество порядков дифракции	≥ 2	1	1

Селективность	Не обладают селективностью	Селективность имеется	Селективность высокая
Степень объемности, Q (по Клейну)	<10	10–1000	>1000

Для используемых в работе голограмм-решеток связь между основными параметрами определяется с помощью формул теории связанных волн:

$$\Delta\theta = 2\pi n_0 / (\pi v T), \quad (2)$$

где n_0 – показатель преломления среды, T – толщина голограммы, v – пространственная частота голограммы.

Толщина голограммы, определенная расчетным путем по формуле (2), называется эффективной толщиной голограммы – $T_{\text{эф}}$, в отличие от геометрической толщины ($T_{\text{геом}}$), измеренной в физическом эксперименте. Эффективная толщина голограммы может не совпадать с геометрической толщиной из-за неравномерности распределения оптических параметров голограммы по ее толщине. Как правило, $T_{\text{эф}} \leq T_{\text{геом}}$.

ЭКСПЕРИМЕНТАЛЬНАЯ УСТАНОВКА

Для измерения параметров голограммы используется установка, оптическая схема которой приведена на рис. 2. Освещение голограммы производится коллимированным монохроматическим излучением полупроводникового лазера (ИИ) с длиной волны $\lambda = 650$ нм.

Голограмма (Г) располагается на столике, позволяющем осуществлять поворот образца вокруг вертикальной оси с необходимой точностью. Вращение столика осуществляется микрометрическим винтом (МВ). Излучение лазера должно проходить через ось вращения столика. Приемник излучения (ПИ) подключен к измерительному прибору – мультиметру. Если нет дополнительных рекомендаций преподавателя, следует работать по шкале измерения токов (А). Показания амперметра прямо пропорциональны потоку излучения, падающему на приемник излучения. Для определения дифракционной эффективности измерения потока излучения проводятся в двух положениях приемника излучения: (1) и (3). При измерении контура угловой селективности приемник излучения находится в положении (3).

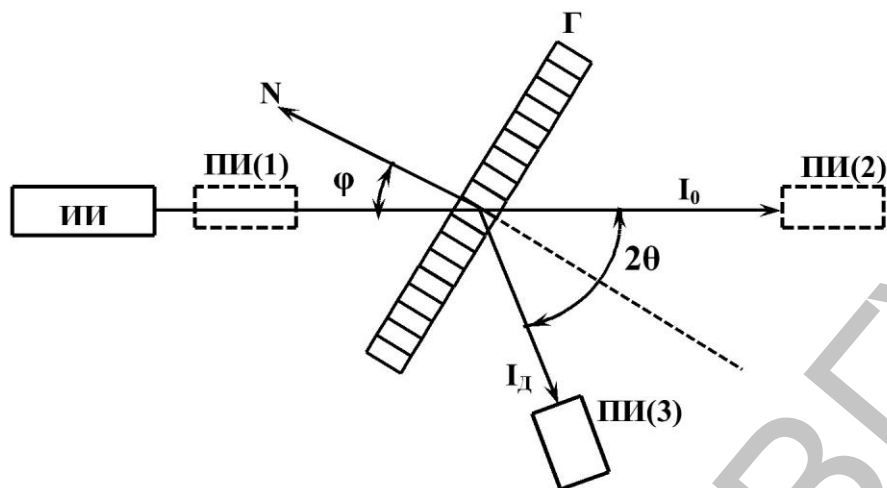


Рис. 2. Оптическая схема установки для измерения дифракционной эффективности и контура угловой селективности пропускающей голограммы-решетки. ИИ – источник излучения; ПИ – приемник излучения, который может быть установлен в положения 1, 2, 3; Г – голограмма; N – нормаль к поверхности голограммы; I_0 и I_d – нулевой и дифрагированный пучки излучения; φ – угол поворота образца относительно падающего пучка; 2θ – угол между нулевым и дифрагированным пучками излучения.

В работе необходимо провести измерения дифракционной эффективности и контуров угловой селективности исследуемых голограмм; рассчитать значения угловой селективности, эффективной толщины, параметра объемности исследуемых голограмм и виртуальных голограмм и проанализировать полученные результаты.

Сравнение контуров угловой селективности голограмм с различной толщиной, близких по параметрам к используемым в данной работе, приведено на рис. 3.

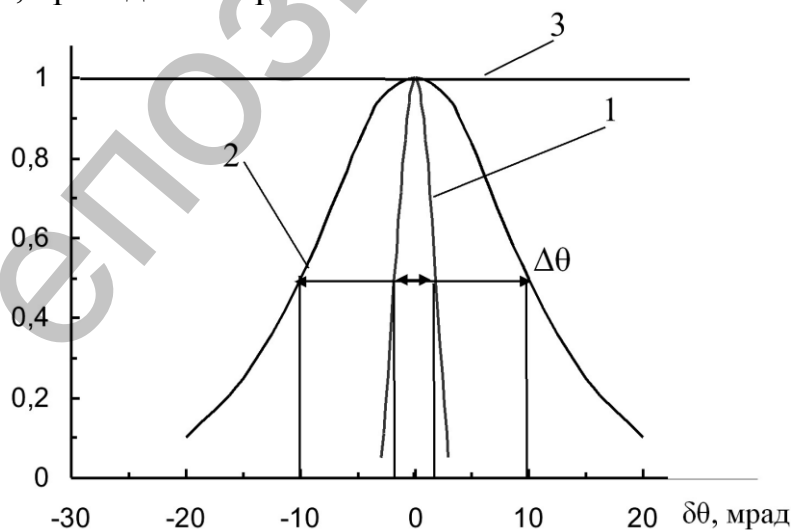


Рис. 3. Сравнение контуров угловой селективности голограмм с различной толщиной, нормированных на 1 и сориентированных по углу Брэгга ($\theta_{бр}$); $\delta\theta$ – отклонение от угла Брэгга; $\Delta\theta$ – угловая селективность голограмм.

Как видно, сравнение удобно проводить, если каждый контур нормирован на «1» и сориентирован по углу Брэгга ($\theta_{\text{бр}}$). Из данных, приведенных на рис. 3 видно, что наиболее высокой селективностью обладает голограмма № 1 ($\Delta\theta$ – минимально), а голограмма № 3 – неселективна (в рассматриваемом диапазоне углов поворота голограммы значение $I_{\text{д}}$ не меняется).

Исходные данные, необходимые для проведения расчетов и оформления отчета по работе, следует определить в эксперименте и уточнить у преподавателя:

- номер исследуемого образца и голограммы,
- длина волны лазерного излучения,
- угол между нулевым и дифрагированным пучками – 2θ (определить самостоятельно в эксперименте),
- пространственная частота исследуемой голограммы – ν , мм^{-1} (рассчитать, используя угол 2θ и условие Брэгга),
- цена деления барабана микрометрического винта в угловых единицах – q , рад/мм (определить, измерив расстояние от вертикальной оси вращения столика с голограммой до горизонтальной оси вращения сердцевинки микровинта).

ПОРЯДОК ВЫПОЛНЕНИЯ РАБОТЫ

1. Ознакомиться с экспериментальным стендом и принципиальной оптической схемой установки для проведения измерений параметров пропускающих голограмм-решеток.

2. Включить источник излучения и установить экран за поворотным столиком для наблюдения пучков излучения.

3. Установить образец № 1 (№ 2, № 3) на поворотный столик таким образом, чтобы излучение лазера попадало на исследуемую голограмму.

4. Поворачивая голограмму относительно падающего излучения, найти положение, при котором дифрагированное излучение имеет максимальную яркость при визуальном контроле.

5. Проанализировать картину дифракции, наблюдаемую на экране и проидентифицировать наблюдаемые порядки дифракции (I_0, I_1, I_2 и т.д.).

Измерение дифракционной эффективности

1. Подготовить табл. 2 для записи результатов измерений.

2. Установить приемник излучения в положение «3», соответствующее измерению интенсивности дифрагированного пучка (-1)-го порядка (рис. 2). Проверить перпендикулярность положения фотоприемника относительно падающего пучка.

3. Поворачивая голограмму относительно падающего пучка, найти максимальное значение интенсивности дифрагированного пучка I_d и записать его. (Измерения проводятся по шкале мультиметра «А»).

4. Установить фотоприемник в положение «1» и измерить интенсивность падающего излучения $I_{пад}$.

5. Если имеются несколько порядков дифракции, определить аналогичным образом значение дифракционной эффективности наиболее ярких из них, например, +1 и ± 2 порядков, при положении голограммы, в котором производились измерения ДЭ (-1)-го порядка. Уточнить задание у преподавателя.

6. После измерения ДЭ голограммы необходимо, не изменяя юстировку голограммы относительно падающего пучка, провести измерение контура угловой селективности данной голограммы.

Таблица 2

Измерение дифракционной эффективности голограмм

№	№ дифр. порядка	Поток излучения при $\theta = \theta_{max}$ Бр, I, мкА		ДЭ, %
			$I_{пад}$	
1	2	3	4	5
Голограмма № 1. Образец № ...				
Голограмма № 2. Образец № ...				
Голограмма № 3. Образец № ...				

Измерение контура угловой селективности

Измерения контуров угловой селективности проводятся для всех голограмм только в (-1) порядке дифракции.

1. Подготовить табл. 3 для записи результатов измерений контура угловой селективности. Данные измерений по каждой голограмме заносить в таблицу.

2. Установить фотоприемник в положение (3) (рис. 2) для измерения интенсивности дифрагированного пучка (-1)-го порядка дифракции.

3. Определить цену деления барабана микрометрического винта (МВ) в угловых единицах (q , рад/мм), измерив расстояние от вертикальной оси вращения столика с голограммой до горизонтальной оси вращения сердцевинки микрометрического винта. Занести результат в исходные данные к табл. 3.

4. Определить диапазон углов поворота, необходимый для построения контура угловой селективности данной голограммы и рассчитать, с каким шагом необходимо проводить измерения, чтобы при построении контура имелось 20–30 экспериментальных точек. Провести измерение контура угловой селективности голограммы, заполнив столбцы 1 и 2 табл. 3.*

* При измерении контура угловой селективности голограммы № 1 заполняются столбцы 1 и 2 табл. 3.

При измерении контура угловой селективности голограммы № 2 заполняются столбцы 1 и 2 табл. 3.

При измерении контура угловой селективности голограммы № 3 заполняются столбцы 1 и 2 табл. 3. Учитывая низкую селективность данной голограммы, оценку селективности проводить по измерениям интенсивности дифрагированного пучка (I_d) в диапазоне углов поворота голограммы не более 25–30 миллирадиан.

Таблица 3

Данные для построения контура угловой селективности голограммы № 1. Исходные данные: цена деления барабана

МВ - q = ____ радиан/мм (или q^{-1} = ____ рад/деление МВ)

Отсчет по барабану МВ, N_i , мм	Интенсивность дифрагированного пучка		$\delta\theta$, рад
	I_i , мкА	$(I_d)_{отн} = I_i / I_{max}$	
1	2	3	4

ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ

1. Рассчитать значения ДЭ для исследуемых голограмм в соответствующих дифракционных порядках, используя формулу

$$\text{ДЭ} = I_{\text{д}} / I_{\text{пад}} \quad (3)$$

Данные занести в табл. 3.

2. Рассчитать значения $(I_{\text{д}})_{\text{отн}}$, заполнив столбец 3 в табл. 3.

3. Рассчитать $\delta\theta$ для каждой голограммы, заполнив столбец 4 в табл. 3. Расчет производить с учетом следующих факторов: при величине $(I_{\text{д}})_{\text{отн}} = 1$ выполняются условия Брэгга для дифракции излучения на голограмме, при этом $\theta = \theta_{\text{бр}}$, и, следовательно, $\delta\theta = 0$, и $N_i = N_0$. Тогда значение $\delta\theta$ для всех остальных значений N_i можно вычислить, используя формулу:

$$\delta\theta = (N_i - N_0) \cdot q [\text{радиан}], \quad (4)$$

где рекомендуется использовать $(N_i - N_0) [\text{мм}]$ и $q [\text{радиан/мм}]$.

4. Построить графики зависимости $I_{\text{д}}(\delta\theta)$ для каждой голограммы и определить $\Delta\theta$ для каждой из них. Данные занести в табл. 4. В отчете привести график, аналогичный приведенному на рис. 3.

5. Рассчитать $T_{\text{эфф}}$ для каждой исследуемой голограммы, используя формулу (2). Данные занести в табл. 4. Для голограммы 3 значение T^* уточнить у преподавателя.

6. Рассчитать параметр объемности Q для каждой исследуемой голограммы, используя формулу (1). Данные занести в табл. 4.

7. Рассчитать аналогичным образом неизвестные параметры модельных двумерных и трехмерных голограмм, заполнив табл. 4.

8. Проанализировать полученные результаты и сделать выводы о влиянии толщины голограммы и ее пространственной частоты (периода решетки) на такие параметры, как селективность голограммы и параметр объемности.

Таблица 4

Сравнение параметров пропускающих голограмм

Голограмма	$\nu, \text{мм}^{-1}$	Кол-во дифр. поряд- ков	$\Delta\theta,$ рад	$T_{\text{эфф}}$	$Q,$ параметр объемн.
1	360	1	?	?	?

Окончание табл. 4

2	360	1	?	?	?
3	360	>2	?	T*	?
Двумерная (2D)	360	>2	$>90^0$	?	≤ 10
Трехмерная (3D)	360	1	?	?	≥ 10
Двумерная (2D)	1500	>2	?	?	≤ 10
Трехмерная (3D)	1500	1	?	?	≥ 10
Объемная (3D)	360	1	?	?	≥ 1000
Объемная (3D)	1500	1	?	?	≥ 1000

* Значение T* уточнить у преподавателя.

КОНТРОЛЬНЫЕ ВОПРОСЫ

1. Какие типы голограмм Вы знаете, и чем они отличаются друг от друга?
2. Как изменяется характер дифракции монохроматического излучения на объемной голограмме-решетке при изменении длины волны?
3. Что такое «селективность голограммы»?
4. Может ли быть разная угловая селективность у объемных голограмм-решеток с одинаковой толщиной?
5. Как влияет на селективность голограммы-решетки изменение периода решетки?

ЛАБОРАТОРНАЯ РАБОТА № 2. Оптическая пространственная фильтрация

Цель работы: изучение роли дифракционных явлений в формировании оптических изображений.

ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ

При изучении явлений дифракции часто сталкиваются со следующей проблемой. Известно комплексное распределение амплитуд $F(\eta, \xi)$ в каком-нибудь отверстии T плоскости A (рис. 4). Нужно найти распределение амплитуд в плоскости B , находящейся на расстоянии l от A .

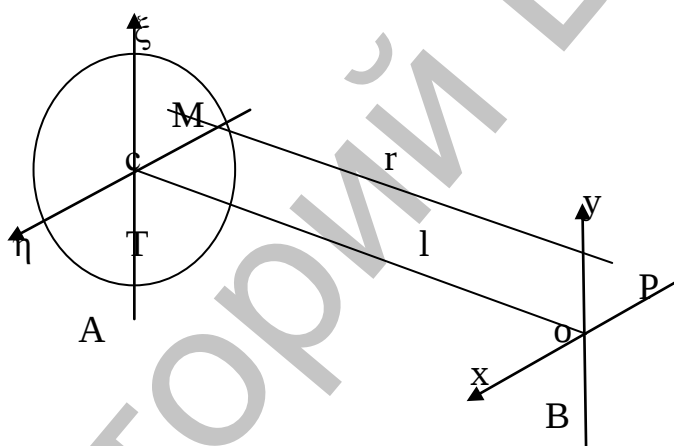


Рис. 4.

Формула Френеля–Кирхгофа дает возможность решить такую задачу. Пусть λ – длина волны используемого света, а r – расстояние от некоторой точки $M(\eta, \xi)$ плоскости A до точки $P(x, y)$ плоскости B . Амплитуда $f(x, y)$ в точке P равна

$$f(x, y) = \frac{1}{j\lambda} \iint_T F(\eta, \xi) \frac{e^{jKr}}{r} d\eta d\xi,$$

где T – отверстие в плоскости A . Предполагается, что линейные размеры отверстия T малы по сравнению с расстоянием l . В том приближении

$$r = \sqrt{l^2 + (\eta - x)^2 + (\xi - y)^2} \approx l \left[1 + \frac{1}{2} \left(\frac{x - \eta}{l} \right)^2 + \frac{1}{2} \left(\frac{y - \xi}{l} \right)^2 \right].$$

Тогда выражение для амплитуды в точке Р запишется в виде

$$f(x, y) = \frac{e^{jKl}}{j\lambda l} \iint_T F(\eta, \xi) \exp \left\{ j \frac{K}{2l} (x - \eta)^2 + (y - \xi)^2 \right\} d\eta d\xi. \quad (1)$$

Полагая, что амплитуда $F(\eta, \xi)$ тождественно равна нулю всюду, кроме отверстия Т, получим

$$f(x, y) = \frac{e^{jKl}}{j\lambda l} \exp \left[j \frac{K}{2l} (x^2 + y^2) \right] \times \\ \times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\eta, \xi) \exp \left[j \frac{K}{2l} (\eta^2 + \xi^2) \right] \times \exp \left[-j \frac{K}{l} (x\eta + y\xi) \right] d\eta d\xi. \quad (2)$$

Теперь видно, что функция $f(x, y)$ с точностью до множителя, стоящего перед интегралом, представляет собой преобразование Фурье функции $F(\eta, \xi) \exp[j(K/2l)(\eta^2 + \xi^2)]$. Если $l \gg 1/2K(\eta^2 + \xi^2)$, то функция $f(x, y)$ будет преобразованием Фурье функции $F(\eta, \xi)$, и мы переходим к дифракции Фраунгофера:

$$f(x, y) = A \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\eta, \xi) \exp \left[-j \frac{K}{l} (x\eta + y\xi) \right] d\eta d\xi. \quad (3)$$

Рассмотрим линзу, толщина которой в центре равна e_0 (рис. 5а), а на некотором расстоянии h от центра она равна e . Если n – показатель преломления материала линзы, то оптический путь луча, проходящего сквозь нее по оптической оси, будет ne_0 . Так как линза тонкая, мы пренебрегаем отклонением луча между плоскостями А и В. Координаты точек пересечения произвольного луча с плоскостями А и В практически совпадают (рис. 5б).

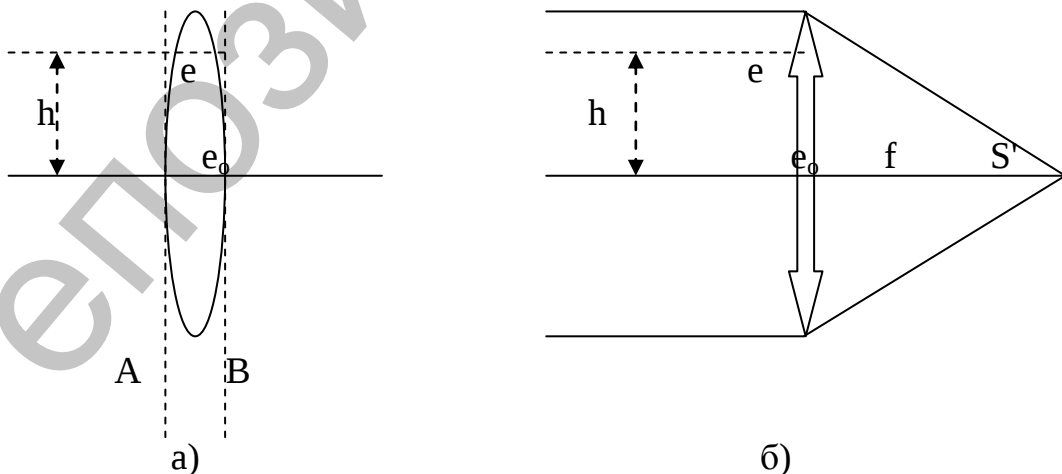


Рис. 5.

Луч, параллельный оптической оси и находящийся на расстоянии h от нее, проходит оптический путь $ne + (e_0 - e)$. Следовательно, тонкая линза вносит изменение фазы $\phi = K[ne + (e_0 - e)]$.

Пусть $F(\eta, \xi)$ – амплитуда падающей волны перед самой линзой (в плоскости A на рис. 5а), тогда амплитуда волны $F'(\eta, \xi)$ сразу после линзы (в плоскости B) будет

$$F'(\eta, \xi) = F(\eta, \xi) e^{jKe_o} e^{jK(\eta^2 + \xi^2)}.$$

Классический расчет позволяет выразить e как функцию e_o радиусов R_1 и R_2 поверхностей линзы, в которой толщина ее равна e . Имеем

$$e = e_o - \frac{\eta^2 + \xi^2}{2} \left(\frac{1}{R_1} - \frac{1}{R_2} \right),$$

поэтому амплитуда запишется в виде

$$F'(\eta, \xi) = F(\eta, \xi) e^{jKe_o} \times \exp \left\{ jK(\eta^2 + \xi^2) \left[e_o - \frac{\eta^2 + \xi^2}{2} \left(\frac{1}{R_1} - \frac{1}{R_2} \right) \right] \right\}, \quad (4)$$

или, вводя сюда фокусное расстояние линзы f , получим

$$F'(\eta, \xi) = F(\eta, \xi) e^{jKe_o} \exp \left[-j \frac{K}{2f} (\eta^2 + \xi^2) \right]. \quad (5)$$

Если на линзу падает плоская волна, как показано на рис. 5б, то $F(\eta, \xi) = \text{const}$, и воздействие линзы на волну выразится формулой

$$F'(\eta, \xi) = e^{jKe_o} \exp \left[-j \frac{K}{2f} (\eta^2 + \xi^2) \right].$$

В рассматриваемом здесь приближении функция

$$\exp \left[-j \frac{K}{2f} (\eta^2 + \xi^2) \right]$$

описывает сферическую волну, сходящуюся при $f > 0$ и расходящуюся при $f < 0$.

Рассмотрим амплитуду волны в фокальной плоскости линзы, когда объект расположен вплотную к линзе (рис. 6). $F(\eta, \xi)$ – амплитуда волны, прошедшей сквозь объект. Из (5) следует, что амплитуда волны после линзы равна (с точностью до множителя e^{jKne})

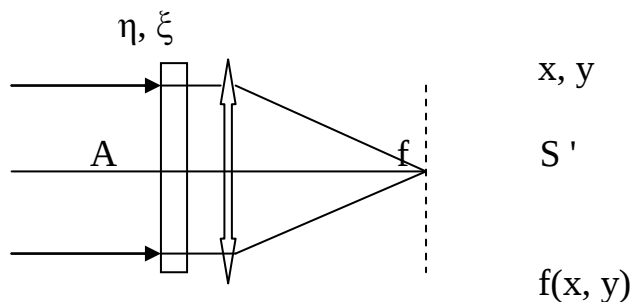


Рис. 6.

$$F(\eta, \xi) \rightarrow F(\eta, \xi) \exp \left[-j \frac{K}{2f} (\eta^2 + \xi^2) \right]. \quad (6)$$

Чтобы найти амплитуду $f(x, y)$ в фокальной плоскости, расположенной на расстоянии f от плоскости линзы, в которой амплитуда равна $F(\eta, \xi)$, нужно применить формулу Френеля–Кирхгофа (2). Предположим, что объект меньше линзы и что конечные размеры линзы не играют роли. Используя (6), получаем

$$f(x, y) = \frac{\exp \left[j \frac{K}{2f} (x^2 + y^2) \right]}{j\lambda f} \times \int_{-\infty}^{\infty} \int F(\eta, \xi) \exp \left[-j \frac{K}{f} (\eta x + y \xi) \right] d\eta d\xi. \quad (7)$$

С точностью до множителя $\exp[j(K/2f)(x^2 + y^2)]$ это выражение представляет собой преобразование Фурье функции $F(\eta, \xi)$. Таким образом, из выражения (7) следует, что распределение комплексных амплитуд на поверхности сферы радиусом f , касающейся фокальной плоскости в точке S' , представляет собой преобразование Фурье амплитуды $F(\eta, \xi)$. Следует отметить, что при решении дифракционных задач обычно ищут распределение интенсивности ff^* , так что множитель перед интегралом, учитывающий фазу исчезает.

Пусть теперь объект помещен на расстоянии d от линзы и пусть $F_o(\eta_o, \xi_o)$ – амплитуда света, прошедшего сквозь объект (рис. 7).

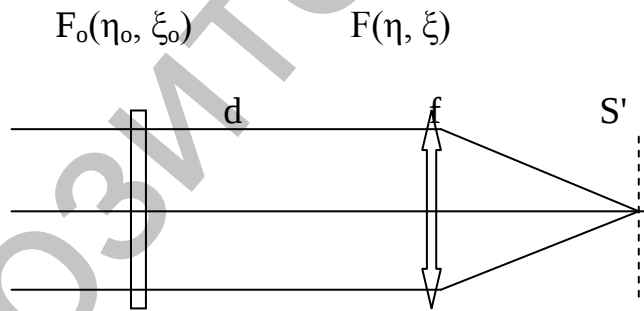


Рис. 7.

Амплитуда $F(\eta, \xi)$ сразу после прохождения тонкой линзы выражается по формуле Френеля–Кирхгофа (1), которую здесь мы запишем в виде

$$F(\eta, \xi) = \frac{e^{jKd}}{j\lambda d} \int_{-\infty}^{\infty} \int F_o(\eta_o, \xi_o) \exp \left\{ j \frac{K}{2d} (\eta - \eta_o)^2 + (\xi - \xi_o)^2 \right\} d\eta_o d\xi_o.$$

Как и раньше, мы пренебрегаем конечными размерами линзы. Опуская промежуточные расчеты, амплитуда $f(x, y)$ в фокальной плоскости линзы запишется следующим образом:

$$f(\eta, \xi) = \frac{\exp \left[j \frac{K}{2f} \left(1 - \frac{d}{f} \right) (\eta^2 + \xi^2) \right]}{j\lambda f} \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_o(\eta_o, \xi_o) \exp \left[-j \frac{K}{f} (\eta \eta_o + \xi \xi_o) \right] d\eta_o d\xi_o$$

Из этого выражения видно, что если объект с амплитудным пропусканием $F_o(\eta_o, \xi_o)$ расположен в передней фокальной плоскости линзы ($d=f$), то распределение амплитуд $f(x,y)$ в задней фокальной плоскости представляет собой точное преобразование Фурье (Фурье-образ) функции прозрачности объекта.

Рассмотрим процесс оптической фильтрации при когерентном освещении. На рис. 8 показана классическая схема оптической фильтрации. Объектом $F_o(\eta_o, \xi_o)$ служит диапозитив или негатив, например, фотопластинка с изображением пейзажа. Объект освещен параллельным пучком монохроматического света. Если он находится от линзы O_2 на расстоянии f , равном ее фокусному расстоянию, то преобразование Фурье объекта лежит в задней фокальной плоскости линзы O_2 . Это преобразование Фурье есть не что иное, как картина дифракции на объекте. Поместим в фокальной плоскости линзы O_2 фильтр, который изменит распределение комплексных амплитуд в этой плоскости. Таким фильтром может быть маленький непрозрачный экран, который задерживает часть дифрагированного света в непосредственной близости от оптической оси линзы, соответствующую крупным деталям объекта (т.е. низким пространственным частотам). Экран не мешает распространению лучей, дифрагированных мелкими деталями объекта (высокие пространственные частоты) и распространяется дальше от оси. Если третья линза O_3 образует в плоскости A' изображение объекта A , то крупные детали не будут воспроизведены на его изображении. Зато мелкие детали будут относительно усилены и видны лучше, что создаст впечатление увеличения четкости фотографии.

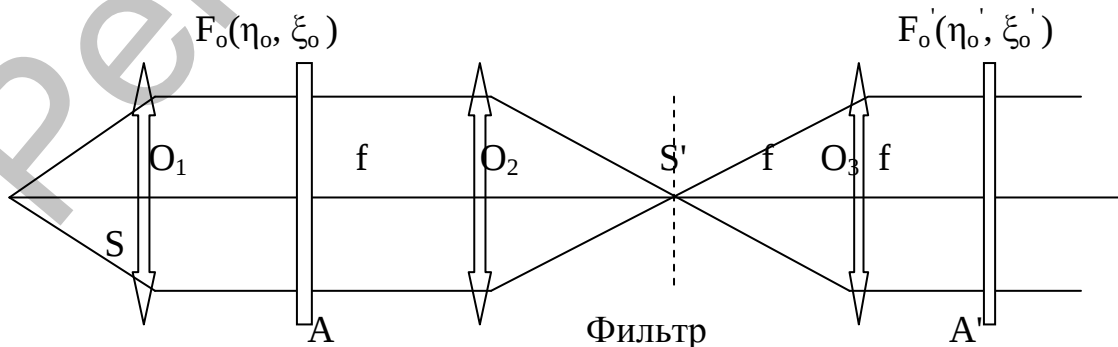


Рис. 8.

Здесь приведен только один пример, но оптическая фильтрация более эффективна, если использовать разные фильтры. В некоторых областях, например, при распознавании образов используются фильтры, которые регистрируют одновременно и фазу, и амплитуду. Эти фильтры представляют собой голограммы.

Рассмотрим дифракцию светового пучка на объекте, обладающем простейшей структурой. Таким объектом может служить, например, одномерная дифракционная решетка.

Если на решетку падает плоская волна, то после решетки мы будем иметь множество плоских волн, распространяющихся в направлениях, удовлетворяющих условию $d \sin \varphi = m \lambda$, где d – период решетки, $m = 0, 1, 2, \dots$ порядок дифракционных максимумов (рис. 9).

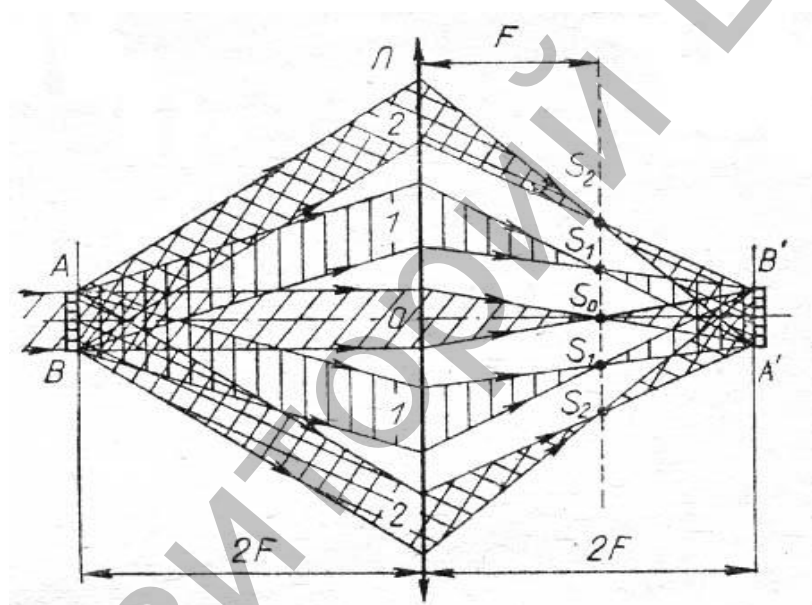


Рис. 9.

Если на пути преобразованного волнового фронта поместить собирающую линзу L , то каждая из плоских волн $0, 1, 2, \dots$ дает в фокальной плоскости этой линзы яркую точку S_0, S_1, S_2 , и т.д. Плоской волне, распространяющейся по какому-либо направлению, принято ставить в соответствие определенное значение пространственной частоты, которая определяется проекцией волнового вектора $\vec{k}$ на некоторое направление, например, на ось, перпендикулярную плоскости решетки. Можно сказать, таким образом, что в фокальной плоскости линзы мы наблюдаем ряд дифракционных максимумов, соответствующих множеству пространственных частот, – так называемый пространственный спектр объекта.

Собственно изображение объекта можно наблюдать за фокальной плоскостью линзы на экране или при помощи другой линзы. В

предлагаемой установке объект АВ (рис. 9) расположен на расстоянии порядка $2F$ от линзы; следовательно, его действительное изображение А'В' будет наблюдаться также на расстоянии порядка $2F$ за линзой.

Теория и эксперимент показывают, что сходство объекта и его оптического изображения будет тем лучше, чем полнее используется пространственный спектр объекта. С другой стороны, если отдельные пространственные частоты исключить из процесса формирования изображения, то могут возникнуть существенные искажения изображения. Процесс изменения изображения объекта, подчеркивание его отдельных структурных особенностей путем выделения тех или иных пространственных частот носит название оптической пространственной фильтрации.

Получение возможно большей информации о свойствах объекта имеет прямое отношение к проблеме разрешающей способности оптических приборов. В частности, объектив микроскопа можно рассматривать как фильтр пространственных частот. Если оправа объектива выделяет из пространственного спектра объекта только один максимум нулевого порядка (что равносильно отсутствию дифракции на объекте), то, несмотря на то, что сквозь объектив проходит значительная световая энергия, никакой информации о структуре объекта мы не получим. Для получения минимальной информации, или как говорят, для разрешения объекта, необходимо, чтобы фильтр пропускал, по крайней мере, максимумы первых порядков, т.е. чтобы выполнялось условие

$$d \geq \frac{\lambda}{n \sin u},$$

где d – размер объекта, $n \sin u$ – числовая апертура.

ОПИСАНИЕ ЭКСПЕРИМЕНТАЛЬНОЙ УСТАНОВКИ

Схема экспериментальной установки представлена на рис. 10. Источником света может служить либо лазер 1, либо точечный источник света, помещенный в фокусе линзы и дающей параллельный пучок лучей света.

Для просвечивания протяженных объектов узкий пучок лазерного излучения расширяется телескопической системой 2. Расширенный пучок света проходит через объект 3 (двумерную дифракционную решетку) и через объектив 4 с фокусным расстоянием $F_2 = \dots$ мм. В фокальной плоскости объектива помещаются различные фильтры пространственных частот 5 – ирисовая диафрагма, или экраны с отверстиями, прорезями, сделанными в соответствии с расположением дифракционных максимумов в пространственном спектре объекта.

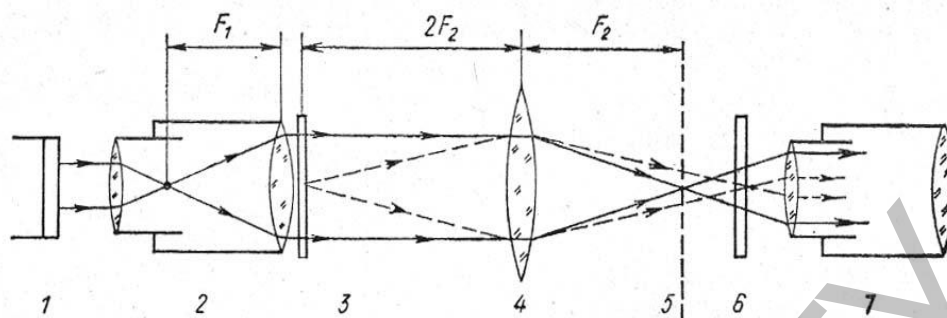


Рис. 10.

Наблюдение изображения производится с помощью микроскопа 7. Для ослабления чрезмерной яркости картины перед микроскопом поставлен поляризатор 6. Объектами, оптическое изображение которых исследуется в работе, служат двумерные дифракционные решетки и сетки.

ИЗМЕРЕНИЯ И ОБРАБОТКА РЕЗУЛЬТАТОВ

Задание 1. Работа с решеткой

Поместите в держатель образца дифракционную решетку и, перемещая микроскоп (7), получите резкое изображение решетки в его поле зрения. Положение всех остальных элементов оптической схемы строго фиксированы и изменять их в процессе работы нельзя.

На экране (6) получите пространственный спектр объекта. Измерьте расстояние между максимумами дифракционной картины. Зная фокусное расстояние объектива (4) и длину волны излучения, определите период d дифракционной решетки. Проверьте результат, измерив период непосредственно с помощью микроскопа. Цена деления окулярной шкалы 36 мкм. Зарисуйте пространственный спектр и соответствующее изображение объекта.

Замените экран ирисовой диафрагмой и, уменьшая отверстие, перекройте все максимумы дифракционной картины, кроме нулевого. Что наблюдается при такой фильтрации? Увеличивая диаметр ирисовой диафрагмы, наблюдайте за изменениями, происходящими в изображении. Зарисуйте выделенные пространственные спектры и соответствующие изображения объекта при различном количестве 2, 3, 4, ... максимумов проходящих через отверстие диафрагмы.

Замените ирисовую диафрагму пространственным фильтром с прорезями, которые пропускают только четные максимумы. Зарисуйте выделенный пространственный спектр и изображение объекта. Измерьте период решетки при такой фильтрации.

Задание 2. Работа с двумерной решеткой (сеткой)

В держателе образца установите двумерную решетку (мелкоячеистую сетку) и получите на экране пространственный спектр структуры. Убрав экран и открыв ирисовую диафрагму, наблюдайте изображение в микроскоп и получите четкое изображение объекта. Зарисуйте пространственный спектр и соответствующее изображение объекта. Измерьте период структуры в горизонтальном и вертикальном направлениях.

Уменьшая диаметр ирисовой диафрагмы, наблюдайте за изменением изображения объекта. Зарисуйте выделенные спектры и соответствующие изображения объекта при различных диаметрах диафрагмы, вырезающей области с различными 0, 1, 2, ... порядками дифракционных максимумов.

Установите пространственный фильтр с прорезями, которые выделяют только вертикальные или горизонтальные максимумы дифракционного спектра. Зарисуйте выделенные пространственные спектры и соответствующие получившиеся изображения объекта.

Установите пространственный фильтр с прорезями, которые пропускают весь горизонтальный спектр и четные порядки вертикального спектра. Зарисуйте выделенный пространственный спектр и изображение объекта.

ЛАБОРАТОРНАЯ РАБОТА № 3. Запись оптических голограмм

Цель работы:

1. Снять голограмму Френеля полупрозрачного объекта, расположенного на разных расстояниях от регистрирующей пластинки, при различных углах падения опорной волны, и голограмму трехмерного объекта.

2. Исследовать влияние соотношения интенсивностей опорной и предметной волн на качество голограмм и приобрести навыки экспериментального изготовления голограмм.

СВЕДЕНИЯ ИЗ ТЕОРИИ

Запишем уравнение голограммы в координатах. Пусть поле $u(x, y)$ в плоскости x, y (объектная волна) образуется в результате дифракции плоской волны на транспаранте, расположенном в плоскости x_1, y_1 . Запишем объектную волну в плоскости фотопластинки в виде

$$u(x, y) = a(x, y) \exp(-j\phi(x, y)),$$

где $a(x, y)$ – амплитудный множитель волнового фронта; $\phi(x, y)$ – фазовый множитель.

Если под некоторым углом относительно нормали к фотопластинке на нее падает плоская волна от того же источника (опорная волна)

$$u_0(x, y) = a_0 \exp(-j\omega_0 x),$$

где $\omega_0 = 2\pi \sin \Theta / \lambda$, то результатом интерференции волн $u(x, y)$ и $u_0(x, y)$ будет суммарное распределение интенсивности волнового поля:

$$I(x, y) = [u_0(x, y) + u(x, y)]^2 = a_0^2 + a^2(x, y) + a_0 a(x, y) \exp(j\omega_0 x) + \exp(-j\omega_0 x)$$

Распределение интенсивности содержит полную информацию как об амплитуде, так и о фазе объектной волны, попавшей на фотопластинку.

Коэффициент пропускания по интенсивности экспонированной, проявленной и зафиксированной фотопластинки связан с интенсивностью следующим соотношением:

$$T = I^{-\gamma},$$

где γ – коэффициент контрастности фотоматериала, зависящий от условий эксперимента и режима обработки.

Наиболее целесообразно выбирать значение коэффициента контрастности $\gamma = -2$. Такая голограмма соответствует так называемой синусоидальной дифракционной решетке. Как известно амплитудная

дифракционная решетка с синусоидальным пропусканием не образует дифракционных порядков выше первого.

В случае, когда фотопластинка проявлена до коэффициента контрастности $\gamma = -2$, амплитудное пропускание голограммы равно

$$t_A(x, y) = \sqrt{T(x, y)} = \sqrt{I(x, y)}.$$

Будем освещать голограмму волной, идентичной опорной с единичной амплитудой. Тогда за голограммой появляется волна

$$u_{\text{обл}}(x, y) = \exp(-j\omega_0 x) T(x, y) \approx I(x, y) \exp(-j\omega_0 x) = \exp(-j\omega_0 x)$$

$$\left[\frac{1}{a_0^2} + a_0^2 \right] \exp(-j\omega_0 x) + a_0 \exp(-j\omega_0 x) \exp(-j\varphi(x, y)) \exp(-j\omega_0 x) + \exp(-j\omega_0 x) \exp(-j\varphi(x, y)) \exp(-j\omega_0 x).$$

Первая волна $\exp(-j\omega_0 x) \left[\frac{1}{a_0^2} + a_0^2 \right]$ сохраняет направление восстанавливающей (плоской) и является неизбежным световым фоном (нулевым дифракционным порядком), вторая $a_0 \exp(-j\varphi(x, y))$, очевидно, повторяет с точностью амплитудного множителя a_0 волну объекта. Она создает его мнимое изображение. В направлении, ему противоположном, распространяется комплексно-сопряженная ей волна $\exp(-j\varphi(x, y)) \exp(-j\omega_0 x) a_0$, которая создает действительное изображение объекта. Это две волны первого дифракционного порядка.

В случае коэффициента контрастности, отличного $\gamma = -2$, кроме дифракции нулевого и первого порядков возникли дифракции более высоких порядков. Это легко видеть, разлагая выражение в ряд и используя формулы:

$$T(x, y) = a_0^{-\gamma} \left[1 + \left(\frac{1}{a_0^2} + \frac{1}{a_0} \right) \exp(-j\omega_0 x) + \frac{1}{a_0} \exp(-j\omega_0 x) \right].$$

Такое положение дает для отношения интенсивностей первого и нулевого порядков величину $I_1/I_0 = \gamma^2/4a_0^2$, где I_1 – интенсивность первой волны; I_0 – интенсивность второй или третьей волны этого равенства.

В пределах линейной контрастно-частотной характеристики с увеличением коэффициента контрастности возрастает дифракционная эффективность голограммы (надо отметить, что она является одним из основных параметров голограммы), т.е. в первый дифракционный порядок уходит все больше световой энергии восстанавливающей волны. При $\gamma = -2$ исчезают нелинейные искажения изображения объекта. Разложение в ряд позволяет оценить отношение сигнал/шум в изображении первого порядка и определяется величиной

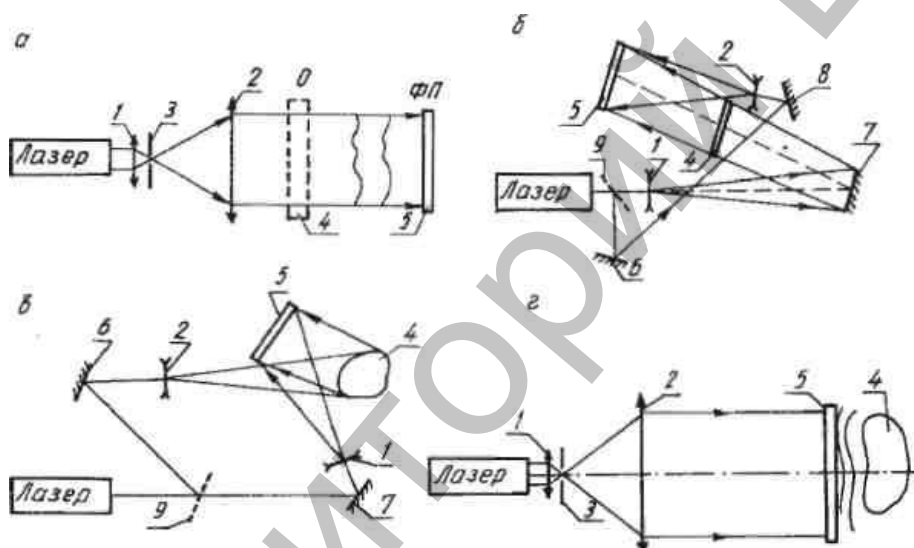
$$S/N = \frac{za_0^2}{S} + 2\gamma^2,$$

где S – амплитуда сигнала; N – амплитуда шума.

Отсюда можно сделать вывод, что для уменьшения нелинейных искажений (шума) следует на стадии записи использовать предметную волну малой интенсивности. Обычно берут a_0^2/u^2 равным от 3 до 10. При уменьшении отношения u^2/a_0^2 уменьшается амплитуда второго и высших дифракционных порядков.

Разновидности топографических схем

Схема записи голограмм Габора включает в себя систему линз 1 и 2 с диафрагмой 3 для формирования однородной плоской световой волны от лазера. Эта световая волна, проходя через полупрозрачный объект 4, образует интерференционную картину в результате суперпозиции, ослабленной объектом волны u_0 и волны u , характеризующей рассеивающие свойства объекта.



а – без пространственного разделения опорной и объектной волн; *б* – с пространственным разделением опорной и объектной волн для голографирования полупрозрачных объемов; *в* – с пространственным разделением опорной и объектной волн для трехмерных объектов; *г* – схема Денисюка.

1, 2 – линзы; 3 – диафрагма; 4 – объект; 5 – фотопластинка; 6, 7, 8 – зеркала; 9 – амплитудный делитель.

Рис. 11. Голографические схемы.

Более широкое применение находят схемы с пространственным разделением опорной и объектной волн. При голографировании прозрачных объектов может быть использована схема, изображенная на рис. 11б. В ней применяются направленные источники когерентного света. Предполагается, что голографируемый объект 4 не обладает свойствами диффузно рассеивающего тела. Необходимый угол между опорной и объектной волнами устанавливается выбором нужной ориентации зеркал 6 и 7.

ОПИСАНИЕ ЛАБОРАТОРНОЙ УСТАНОВКИ

В состав установки для съема голограмм входят: лазер ЛГ-38, электромеханический затвор, диафрагмы, полупрозрачные зеркала с разным коэффициентом пропускания по интенсивности, глухие зеркала, длиннофокусные и короткофокусные линзы, люксметр, держатели оптических элементов, кассета для пластинок 9Х12, полупрозрачный транспарант типа слайда и трехмерный рассеивающий объект. Все элементы устанавливаются на голографическом столе, оптической скамье или другом защищенном от вибраций устройстве.

РАБОТА С УСТАНОВКОЙ

1. Получите у преподавателя объекты голографирования (полупрозрачный транспарант и трехмерный рассеивающий объект и выберите схему для записи: а) полупрозрачного транспаранта; б) трехмерного рассеивающего объекта.

2. Подберите оптические элементы на основе выбранной схемы.

3. Включите лазер, предварительно ознакомившись с порядком и правилами его эксплуатации.

4. Установите детали согласно выбранной схеме для голографирования полупрозрачного объекта и убедитесь в жесткости крепления всех элементов схемы.

5. Оцените с помощью выбранной схемы разрешающую способность регистрирующего материала. Получите у преподавателя фотопластинку с нужным разрешением.

6. Изготовьте проявляющий и фиксирующий растворы согласно рецептурным данным для фотоматериала.

7. Установите фотопластинку в кассету, предварительно закрыв луч лазера затвором.

8. Выберите экспозицию и произведите съемку.

9. Повторите вышесказанное для трехмерных объектов.

ЗАДАНИЕ ПО РАБОТЕ И ПОРЯДОК ЕГО ВЫПОЛНЕНИЯ

1. Снимите голограмму полупрозрачного объекта, расположенного на расстоянии 10 см при угле 60° между опорным и объектным лучами, меняя при каждой экспозиции соотношение интенсивностей лучей в следующем порядке:

а) интенсивности равны; б) интенсивность предметного пучка в 2 раза больше интенсивности опорного пучка; в) интенсивность опорного пучка в 5 раз превышает интенсивность предметного пучка.

2. Снимите голограмму при соотношении интенсивностей 5:1 при разных углах падения между опорным и объектным лучами при различных экспозициях.

3. Полученные голограммы проявите, закрепите, высушите согласно технологии обработки голографических материалов.

ЛАБОРАТОРНАЯ РАБОТА № 4. Формирование акустических голограмм с оптическим восстановлением

Цель работы:

1. Экспериментально получить, преобразовать в оптический диапазон и восстановить акустические голограммы звуковых волн, излучаемых динамиком и рассеиваемых простейшим объектом.
2. Исследовать влияние различных условий эксперимента на решающую способность формируемой голограммы и качество восстановленного изображения.

СВЕДЕНИЯ ИЗ ТЕОРИИ

В акустической голографии, как и в голографии диапазона СВЧ, используются в основном аппаратные методы записи голограмм. Следует отметить высокую чувствительность аппаратного метода по сравнению с другими. Методы, подобные фотографическому (термический, метод деформации поверхности), имеют чувствительность порядка $1\text{--}10^{-3}$ Вт/см². Чувствительность же аппаратного метода может достигать $10^{-11}\text{--}10^{-12}$ Вт/см².

Простейшая схема установки для записи акустических голограмм реализуется просто. Объект освещается акустической волной, излучаемой динамиком, который возбуждается генератором. Рассеянная волна вызывает электрические колебания в звукоприемнике. Сигнал от звукоприемника поступает на смеситель. На второй вход смесителя поступает сигнал постоянной амплитуды и меняющийся посредством фазовращателя по определенному закону фазы. Он после смесителя представляет собой интерференционный сигнал предметной и опорной волн. Усилитель мощности нагружен лампочкой, движущейся совместно с микрофоном-звукоприемником. Интерференционная картина (звуковая голограмма) фиксируется на фотопленке при помощи фотоаппарата. Одновременно происходит необходимое уменьшение размеров голограммы.

Для имитации плоской опорной волны, падающей под углом к оси звукоприемной апертуры, фазу электрического сигнала необходимо менять по закону $\varphi = \omega_s t + kx \sin \Theta + \varphi_0$, где ω_s – частота генератора; t – время; k – волновое число; x – положение звукоприемника; φ_0 – начальная фаза.

Данное выражение указывает и другой способ поворота электрической опорной волны на заданный угол, если записать его в виде $\varphi = \omega_s + kv \sin \Theta \cdot \vec{t} + \varphi_0 = \omega_s + \Delta\omega \cdot \vec{t} + \varphi_0$, где $\Delta\omega$ – линейная скорость звукоприемного элемента по плоскости апертуры.

Отсюда видно, что электрическую опорную волну можно повернуть, используя два генератора со взаимной расстройкой $\Delta\omega$.

Этим способом можно имитировать и сферическую опорную волну радиуса R , если $\Delta\omega$ менять по закону $\Delta\omega = \frac{2\pi}{\lambda} \sqrt{R^2 - a^2}$, где a – смещение источника сферической волны с оси апертуры.

Дальность действия голографических установок ограничивается нестабильностью аппаратуры и внешними условиями.

Допустимая ошибка фазы между опорной и предметной волнами за время съема голограммы не должна превышать значения $\Delta\psi = \pi/5$. Изменение фазы может быть вызвано уходом частоты генератора $\Delta\psi = 2\pi N \Delta f / f$, где N – число длин волн, укладываемых на расстоянии от объекта до приемной апертуры.

Типичная система для получения акустических голограмм посредством радиоэлектронных устройств и их восстановления состоит из следующих основных узлов: высокостабильного генератора гармонических колебаний перестраиваемого в диапазоне рабочих частот; усилителя мощности; источника звука, излучающего стационарное звуковое поле в направлении наблюдения; устройства, моделирующего опорный сигнал; входного устройства, формирующего приемную апертуру; устройства обработки, в котором коммутируется, сливается и преобразуется принятый каждым микрофоном сигнал; устройства отображения и фиксации голограммы синхронизирующего устройства; оптического устройства для восстановления изображения.

ЗАДАНИЕ ПО РАБОТЕ И ПОРЯДОК ЕГО ВЫПОЛНЕНИЯ

1. Снимите голограмму точечного источника при соотношении $u_{on} / u_{пред}$ для значений $\lambda = 3$ см, 1,5 см, $\alpha = 0$ (схема Габора); $R = 3$ м, $\lambda = 3$ см, 2 см, $\alpha = 0,2$, $R = 2$ м, 3 м; $\lambda = 1,5$ см, $\alpha = 0,3$, $R = 3$ м.

2. Снимите голограмму трех точечных источников $\lambda = 1,5$ см, $\alpha = 0,3$, $R = 3$ м, $u_{on} / u_{пред} = 5$.

3. Снимите голограмму объекта при $\lambda = 1,5$ см, $R_1 = 1,5$, $R_2 = 3$ м. $\alpha_1 = 0,3$, $\alpha_2 = 0,5$ для трех значений отношения $u_{on} / u_{пред} = 1, 5, 10$.

4. Рассчитайте коэффициенты продольного и поперечного увеличений, координаты изображения для плоской восстанавливающей волны $\lambda = 0,6$ мкм. Рассчитайте местоположение растровых порядков и разрешающую способность голограмм.

5. Восстановите голограммы плоской восстанавливающей волной. Сравните параметры изображения с результатами расчета. Как влияет дискретизация приемной апертуры на разрешающую способность восстановленного изображения?

ОПИСАНИЕ ЛАБОРАТОРНОЙ УСТАНОВКИ

В данной установке фазовращатель отсутствует, следовательно, в устройстве имитируется плоская опорная волна, нормально падающая к апертуре (рис. 12). Для уменьшения помех применен предварительный усилитель 1, размещенный непосредственно у приемника звука 2, а также резонансный усилитель 4, повышающий добротность устройства. Усилитель постоянного тока собран по схеме с модулятором 5 и последующим усилителем мощности 6. Голограмма записывается при открытом затворе фотоаппарата 11.

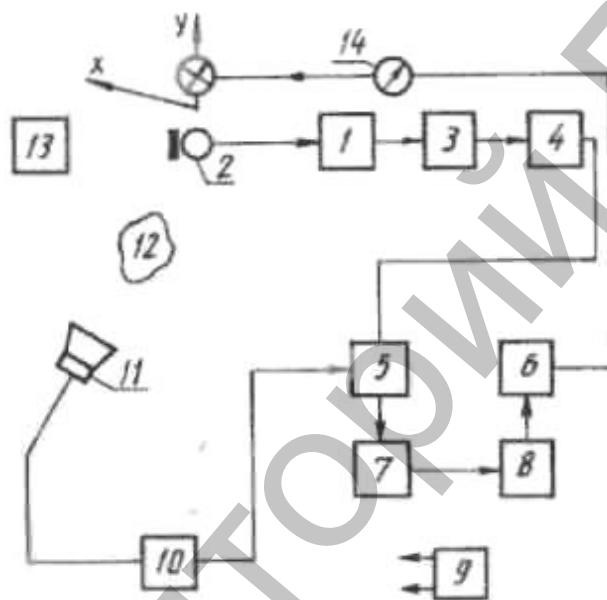


Рис. 12. Схема установки.

КОНТРОЛЬНЫЕ ТЕСТЫ

1. Одномерные волны – это:

- волны, все характеристики которых зависят от времени и одной координаты;
- волны, все характеристики которых зависят только от времени;
- волны, все характеристики которых зависят от времени и двух координат.

2. Гармоническая волна – это:

- бесконечная синусоидальная волна, в которой все изменения состояния среды происходят по закону синуса или косинуса;
- бесконечная синусоидальная волна, в которой все изменения состояния среды происходят по закону Гаусса;
- бесконечная стоячая волна.

3. Длина волны – это:

- скорость распространения волны;
- время, за которое частицы среды совершают одно полное колебание;
- расстояние между двумя максимумами или минимумами возмущения.

4. Фронт волны – это:

- передний край волны;
- поверхности, у которых в данный момент времени фазы в любой точке одинаковы;
- форма фазовой поверхности волны.

5. Поляризации волны – это:

- ориентации векторов поля при заданном направлении распространения;
- суперпозиции двух волн;
- угловая скорость в направлении распространения волны.

6. Явление, возникающее при прохождении пучка света вблизи края какого-либо непрозрачного экрана, при котором его путь не остается строго прямолинейным:

- дифракция;
- дисперсия;
- дисторсия.

7. В голографии на светочувствительном материале регистрируется:

- распределение яркости света, адекватное его распределению в плоскости объекта;
- интерференционная картина объектной и вспомогательной (опорной) волн.

8. Для получения устойчивой интерференционной картины голограммы требуется:

- большая разность оптических путей двух волн;
- высококогерентные источники излучения;
- одинаковое значение разрешающей способности в продольном и в поперечном направлении.

9. Существующие способы записи голограмм:

- схема Денисюка;
- схема Лейта–Упатниекса;
- схема Айзека Раби.

10. Назовите фотоматериалы для записи голограмм:

- бромид серебра;
- йодистый азот;
- сульфат аммония.

11. Бинарные голограммы – это:

- голограммы, с полным пропусканием;
- голограммы, в которых участки с полным пропусканием чередуются с участками, абсолютно непрозрачными для света;
- голограммы, абсолютно непрозрачные для света.

12. Можно ли голографирование осуществлять в пространственно-некогерентном свете:

- нельзя;
- можно, но свет исходящий из конкретной точки объекта, нужно разделить на две части;
- можно, только когда фазовая информация о волновом фронте объекта кодируется в зоне чередования участков голограммы с различным амплитудным пропусканием.

13. Зачем используются лазеры в голографии:

- для термического выжигания изображений на фотопластинках;
- для исследования химических реакций;
- для получения когерентного излучения.

14. Распознавание объекта – это:

- сравнение его признаков с признаками одного или нескольких заранее определенных классов для отнесения объектов к тому или иному классу;
- наиболее распространенная область применения оптической фильтрации;
- получение спектра импульсного отклика методом Вандер Люгта.

15. Оптическая фильтрация – это:

- синтез волновых полей;
- принцип изменения структуры оптических изображений путем воздействия на Фурье-спектры;

- дискретизация волновых фронтов.

16. Оптические процессоры – это:

- устройства для выполнения интегральных и других математических преобразований над волновыми фронтами в оптическом диапазоне волн;
- устройства для управления модуляцией лазеров;
- устройства для получения решений дифференциальных уравнений третьего порядка.

17. Промышленный робот – это:

- робот, подражающий движениям живых организмов;
- автоматическая машина, состоящая из манипулятора и устройства программного управления его движением, предназначенная для замены человека при выполнении основных и вспомогательных операций в производственных процессах;
- человекообразный робот, способный самостоятельно прибраться в квартире и вернуться на место для подзарядки без участия человека.

18. Манипулятор промышленного робота – это:

- рука человека, предназначенная специально для выполнения какого-то определенного набора операций;
- совокупность пространственного рычажного механизма и системы приводов, осуществляющая под управлением программируемого автоматического устройства или человека-оператора действия (манипуляции), аналогичные действиям руки человека.

19. Искусственный интеллект – это:

- наука и технология создания интеллектуальных машин и систем, особенно интеллектуальных компьютерных программ, направленных на то, чтобы понять человеческий интеллект;
- попытки смоделировать с помощью искусственных систем психофизиологическую деятельность человеческого мозга с целью создания искусственного разума;
- создание программ, позволяющих с использованием ЭВМ воспроизводить не саму мыслительную деятельность, а являющиеся ее результатами процессы.

20. Уравнение голограммы выражает:

- коэффициент пропускания фотопластинки;
- интенсивность интерференционной картины на поверхности;
- отношение дифракционных порядков.

21. Оптическая голография применяется для:

- трехмерного восстановления объектов;
- производства микросхем;
- исследования вибраций, деформаций;
- получения химических углеродистых соединений.

22. Цифровая голография позволяет осуществлять:

- синтез голограмм по их математической модели;
- сверхплотную запись информации на фотополимеры;
- улучшение качества голограмм.

23. Чем отличается пассивная радиолокация от активной:

- частотным диапазоном;
- происхождением принимаемого сигнала (собственное излучение объекта или облучение внешним способом);
- разрешающей способностью РЛС.

24. Укажите верное утверждение:

- ширина характеристики направленности антенны, определяющая разрешающую способность РЛС, обратно пропорциональна отношению раскрыва антенны к длине волны;
- ширина характеристики направленности антенны, определяющая разрешающую способность РЛС, пропорциональна отношению раскрыва антенны к длине волны;
- ширина характеристики направленности антенны, определяющая разрешающую способность РЛС, пропорциональна отношению скорости антенны к ее длине.

25. Разрешающая способность РЛС – это:

- способность отдельного наблюдения целей;
- отношение ширины РЛС к ее высоте;
- фокусирование антенны по дальности.

26. Радио- и звуковидение – это:

- обнаружение и распознавание объектов, источниками волновых фронтов которых является радио- и звуковые изображения;
- преобразование волновой информации в звук и радио;
- исследование радио- и звуковых волн по их интерференционным картинам.

27. Использование СВЧ голографии дает возможность:

- голографирования в темноте и за непрозрачными для света препятствиями;
- получения голограмм и интерферограмм больших сцен и крупномасштабных процессов;
- обходиться без источников когерентного излучения.

28. Лазерные линии связи:

- технология беспроводной оптики, позволяющая организовать высокоскоростные каналы связи путем передачи направленного инфракрасного излучения через атмосферу;
- технология беспроводной связи основанная на принципе передачи информации в виде голограмм.

29. Фазовая голограмма:

- полученная в регистрирующей среде, в которой под воздействием интерференционного поля произошли изменения показателя преломления или толщины материала;
- полученная в регистрирующей среде, в которой под воздействием интерференционного поля произошли изменения коэффициента поглощения;
- полученная с оптическими параметрами, неизменными во времени.

30. Амплитудная голограмма:

- полученная в регистрирующей среде, в которой под воздействием интерференционного поля произошли изменения показателя преломления или толщины материала;
- полученная в регистрирующей среде, в которой под воздействием интерференционного поля произошли изменения коэффициента поглощения;
- полученная с оптическими параметрами, неизменными во времени.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Ключников А.С. Теория волновых процессов. – Мн.: БГУ им. Ленина, 1977.
2. Ключников А.С. Радиооптика и голография. – Мн.: Издательство «Университетское», 1989.

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

1. Кольер Р., Беркхард К., Лин Л. Оптическая голография. – М.: Мир, 1973. – 686 с.
2. Оптическая голография / под ред. Г. Колфилда. – М.: Мир, 1982. – 735 с.
3. Миллер М. Голография. – Л.: Машиностроение, 1979. – 140 с.
4. Ландсберг Г.С. Оптика. – М.: ФИЗМАТЛИТ, 2003. – 848 с.
5. Фризер Х. Фотографическая регистрация информации. – М.: Мир, 1978. – 670 с.
6. Ахманов С.А., Никитин С.Ю. Физическая оптика. – М.: Издательство МГУ им. М.В. Ломоносова; Наука, 2004. – 654 с.
7. Saxby G. Practical holography. – Taylor&Francis, 2003. – 482 p.
8. Янг М. Оптика и лазеры, включая волоконную оптику и оптические волноводы. – М.: Мир, 2005. – 542 с.
9. Гудмен Дж. Введение в Фурье-оптику. – М.: Мир, 1970. – 364 с.
10. Palmer K. Diffraction grating handbook. – Richardson grating laboratory, 2000. – 144 p.
11. Horimai H., Tan X., Li J. Collinear holography. Applied optics. – 2005. – Vol. 44, № 13. – P. 2575–2579.
12. Ludman J., Caulfield H.J., Riccobono J. Holography for the new millennium. – Springer, 2002. – 344 p.