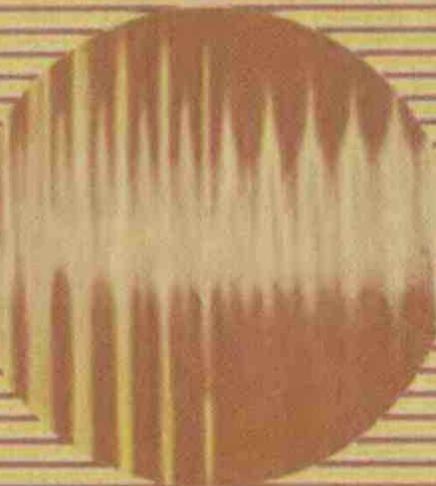


А.Н.Матвеев

Оптика



Высшая
школа

А.Н.Матвеев

Оптика

Допущено
Министерством высшего и среднего
специального образования СССР
в качестве учебного пособия
для студентов физических специальностей
вузов



Москва
„Высшая школа“
1985

ББК 22.34
М 33
УДК 535

Р е ц е н з ен ты:
акад. АН УССР А. И. Ахиезер; кафедра общей физики Ленинградского государственного университета им. А. А. Жданова (зав. кафедрой — проф. Н. И. Калининский)

Матвеев А. Н.

М 33 Оптика: Учеб. пособие для физ. спец. вузов. — М.: Высш. шк., 1985.—351 с., ил.
В пер.: 1р. 40 к.

Изложение материала ведется в рамках электромагнитной теории с использованием преобразования Фурье как для монохроматического, так и для немонхроматического и хаотичного излучений. Наряду с традиционными в книге освещены вопросы, получившие особенное значительное развитие в последние 15—20 лет: Фурье-оптика, Фурье-анализ случайных сигналов, матричные методы в геометрической оптике, голограммия, лазеры, нелинейные явления и др.

Книга представляет собой четвертый том курса общей физики для университетов и вузов. Предыдущие тома — «Механика и теория относительности», «Молекулярная физика», «Электричество и магнетизм» — вышли соответственно в 1976, 1981 и 1983 гг.

M 1704050000-176
001(01)-85 75-84

ББК 22.34
535

Оглавление

1 Электромагнитные волны

Предисловие	9
§ 1. Оптический диапазон электромагнитных волн	12
Длины волн видимого диапазона. Частоты волн видимого диапазона. Оптический и другие диапазоны электромагнитных волн. Почему мы видим именно в видимом диапазоне? Почему микроволновая область непригодна для зрения? Ночное видение	
§ 2. Свойства электромагнитных волн	17
Электромагнитная природа света. Волновое уравнение. Плоские волны. Сферические волны. Плоские гармонические волны. Волновой вектор. Представление плоской волны в комплексной форме. Представление сферической волны в комплексной форме. Плоская электромагнитная волна. Инвариантность плоской волны. Инвариантность фазы. Четырехмерный волновой вектор. Формулы преобразования частоты и направления распространения плоской волны. Эффект Доплера	
§ 3. Плотность потока энергии и импульса электромагнитных волн. Давление света	26
Плотность потока энергии. Распределение плотности потока по сечению пучка Гауссов пучок. Плотность импульса электромагнитной волны. Давление света. Действие светового давления на малые частицы. Лазерный термояд. Преобразование амплитуды и нормали плоской электромагнитной волны. Энергия цуга плоских волн. Импульс цуга плоских волн	
§ 4. Суперпозиция электромагнитных волн	33
Суперпозиция векторов поля волны. Суперпозиция бегущих плоских монохроматических электромагнитных волн. Биенция. Стоячие волны. Преобразование энергии в стоячей электромагнитной волне. Экспериментальное доказательство электромагнитной природы света	
§ 5. Поляризация электромагнитных волн	37
Поляризация. Линейная поляризация. Суперпозиция линейно поляризованных волн. Эллиптическая и круговая поляризации. Изменение вектора напряженности в пространстве при эллиптической и круговой поляризациях. Вырожденный случай эллиптической поляризации. Число независимых поляризаций. Линейно поляризованная волна как суперпозиция волн с круговой поляризацией	
§ 6. Усреднения	40
Операция усреднения. Усреднение гармонических функций. Усреднение квадратов гармонических функций. Линейность операции усреднения. Вычисления с комплексными скалярными величинами. Вычисления с комплексными векторными величинами	
§ 7. Фотометрические понятия и величины	44
Энергетические и фотометрические величины. Энергетические величины. Энергетическая сила излучения. Энергетическая яркость. Энергетическая светимость. Энергетическая освещенность. Фотометрические величины. Световой поток. Яркость. Светимость. Освещенность. Световая экспозиция. Соотношения между энергетическими и световыми характеристиками излучения	
Задачи	
§ 8. Спектральный состав функций	54
Ряд Фурье в действительной форме. Ряд Фурье в комплексной форме. Интеграл Фурье в действительной форме. Интеграл Фурье в комплексной форме. Спектр амплитуд и спектр faz. Нахождение спектра амплитуд и faz из ряда Фурье в комплексной форме. Непрерывный спектр. Спектр прямоугольных импульсов. Спектр пилообразных импульсов. Спектр изолированного прямоугольного импульса. Спектр экспоненциально убывающей функции. Соотношение между продолжительностью импульса и шириной спектра. Смещение начала отсчета времени. Смещение спектра по частотам. Отрицательные частоты. Теорема Парсеваля. Теорема Планшереля	56

2 Немонохроматическое и хаотическое излучение

**3
Распространение
света
в изотропных
средах**

§ 9. Естественная ширина линии излучения	63
Классическая модель излучателя. Спектральный состав излучения. Лоренцева форма и ширина линии излучения. Время излучения. Форма линии поглощения. Квантовая интерпретация формы линии излучения	
Квазимохроматическая волна	
§ 10. Уширение спектральных линий	69
Причины уширения. Однородное и неоднородное уширение. Естественная ширина линии излучения как однородное уширение. Ударное уширение. Доплеровское уширение. Форма составной линии излучения	
§ 11. Модулированные волны	73
Модуляция. Модуляция амплитуды. Модуляция частоты и фазы. Спектр колебания с гармонической модуляцией частоты	
§ 12. Волновые пакеты	75
Волновой пакет, образованный двумя волнами. Групповая скорость. Суперпозиция колебаний с эквидистантными частотами. Квазиплоская волна	
§ 13. Хаотический свет	78
Суперпозиция волн со случайными фазами. Время разрешения. Усреднение по периоду колебаний. Влияние увеличения промежутка времени на результат усреднения. Время когерентности. Длина когерентности. Гауссов свет. Флуктуации плотности потока энергии хаотического света. Поляризация.	
§ 14. Фурье-анализ случайных процессов	82
Спектр мощности. Автокорреляционная функция. Теорема Винера—Хинчина. Интервал корреляции. Связь интервала корреляции с нормированным спектром мощности	
Задачи	86
§ 15. Распространение света в диэлектриках	88
Монохроматические волны. Дисперсия. Нормальная дисперсия. Аномальная дисперсия. Рассеяние света. Распространение волнового пакета. Замещение световой волны в среде. Дисперсия света в межзвездном пространстве. Окраска тел	
§ 16. Отражение и преломление света на границе между диэлектриками. Формулы Френеля	95
Границные условия. Постоянство частоты волны при отражении и преломлении. Плоскость падающего, отраженного и преломленного лучей. Соотношения между углами падения, отражения и преломления. Разложение плоской волны на две с взаимно перпендикулярными линейными поляризациями. Вектор E перпендикулярен плоскости падения. Формулы Френеля для перпендикулярных составляющих векторов поля. Вектор E лежит в плоскости падения. Формулы Френеля для параллельных составляющих векторов поля. Явление Брюстера. Соотношения между фазами волн при отражении или преломлении. Степень поляризации	
§ 17. Полное отражение света	104
Формулы для углов $\theta_{\text{пад}} > \theta_{\text{пред}}$. Волна во второй среде. Глубина проникновения. Фазовая скорость. Отраженная волна	
§ 18. Энергетические соотношения при преломлении и отражении света	107
Плотности потоков энергии. Коэффициент отражения. Коэффициент пропускания. Закон сохранения энергии. Поляризация света при отражении и преломлении	
§ 19. Распространение света в проводящих средах	109
Комплексная диэлектрическая проницаемость. Глубина проникновений. Физическая причина поглощения. Фазовая скорость и длина волны. Соотношения между фазами колебаний векторов поля. Соотношения между амплитудами векторов поля. Среды с малой электропроводимостью. Среды с большой электропроводимостью	
§ 20. Отражение света от поверхности проводника	113
Границные условия. Соотношения между амплитудами волн. Коэффициент отражения. Связь между отражательной и поглощающей способностями	
Задачи	115

§ 21. Приближение геометрической оптики	118
Уравнение эйконала. Луч света. Область применимости лучевого приближения. Принцип Ферма. Вывод закона преломления из принципа Ферма. Распространение луча в среде с переменным показателем преломления.	
§ 22. Линзы, зеркала и оптические системы	123
Параксиальное приближение. Преломление на сферической поверхности. Матричные обозначения. Распространение луча в линзе. Преломление луча на второй сферической поверхности. Преломление луча линзой. Распространение луча через оптическую систему. Отражение от сферических поверхностей	
§ 23. Оптическое изображение	127
Матрица оптической системы. Преобразование луча от плоскости предмета к плоскости изображения. Кардинальные элементы оптической системы. Физический смысл постоянных Гаусса. Построение изображений. Уравнение линзы. Тонкие линзы. Система тонких линз. Использование ЭВМ	
§ 24. Аберрации оптических систем	134
Источники aberrаций. Точные матрицы преобразований. Сферическая aberrация. Кома. Аберрации, обусловленные внеосевыми наклонными лучами. Хроматическая aberrация. Иммерсионный объектив. Условие Аббе	
§ 25. Оптические приборы	140
Диафрагмирование. Основные понятия, связанные с диафрагмированием. Глаз как оптическая система. Фотоаппарат. Лупа. Микроскоп. Зрительная труба. Проекционные устройства	
Задачи	145
§ 26. Двухлучевая интерференция, осуществляемая делением амплитуды	148
Определение интерференции. Интенсивность при суперпозиции двух монохроматических волн. Способы получения когерентных волн в оптике. Интерференция монохроматических волн, распространяющихся строго вдоль оси интерферометра Майкельсона. Интерференция монохроматических волн, распространяющихся под углом к оси интерферометра. Причина размытия полос интерференции. Интерференция немонохроматического света. Принцип Фурье-спектроскопии. Видимость при гауссовой форме линии. Видимость при лоренцевой форме линии. Интерферометр Майкельсона с линейными полосами. Интерференционная картина от белого света. Интерферометр Маха—Цендера. Интерферометр Тваймана—Грина. Интерферометр Жамена	
§ 27. Двухлучевая интерференция, осуществляемая делением волнового фронта	162
Принцип Гюйгенса. Схема Юнга. Интерференция при белом свете. Источник конечного размера. Источник с однородным распределением интенсивности излучения. Временная и пространственная когерентности. Угол и ширина когерентности. Звездный интерферометр. Измерение диаметров звезд. Измерение расстояния между компонентами двойной звезды. Бипризма Френеля. Билинза Бийе. Зеркало Ллойда. Бизеркало Френеля. Закон сохранения энергии в явлениях интерференции	
§ 28. Многолучевая интерференция, осуществляемая делением амплитуды	171
Интерферометр Фабри—Перо. Распределение интенсивности в интерференционной картине. Интерференционные кольца. Разрешающая способность. Факторы, ограничивающие разрешающую способность. Дисперсионная область. Сканирующий интерферометр Фабри—Перо. Интерференционные фильтры. Пластика Люммера—Герке. Эшелон Майкельсона	
§ 29. Интерференция в тонких пленках	180
Оптическая длина пути. Отражение от параллельных поверхностей. Линии равного наклона. Роль размера источника. Роль толщины пленки и монохроматичности излучения. Линии равной толщины. Кольца Ньютона. Учет многократных отражений. Слой с нулевой отражательной способностью. Слой с высокой отражательной способностью.	

6 Дифракция

Матричный метод расчета многослойных пленок. Многослойные диэлектрические зеркала. Полупрозрачные материалы

§ 30. Частичная когерентность и частичная поляризация

Частичная когерентность. Функция взаимной когерентности. Комплексная степень когерентности. Степень когерентности. Опыт Брауна и Твисса. Частичная поляризация. Матрица когерентности квазимонохроматической плоской волны. Комплексная степень когерентности взаимно перпендикулярных проекций напряженности электрического поля волны. Естественный (неполяризованный) свет. Полностью поляризованный свет. Степень поляризации световой волны. Выражение степени поляризации через экстремальные значения интенсивности. Представления естественного света. Соотношение между степенью поляризации и степенью когерентности. Теорема Ван-Циттерта—Цернике.

Задачи

204

208

213

219

232

234

236

239

247

249

7 Основные понятия Фурье-оптики

Линза как элемент, осуществляющий преобразование Фурье

Фазовое преобразование, осуществляющее тонкой линзой. Расчет функции толщины. Виды линз. Линза как элемент, осуществляющий преобразование Фурье

§ 35. Дифракционное образование изображений линзой

Фурье-преобразование амплитуд между фокальными плоскостями линзы. Формирование изображения линзой. Предел разрешающей способности оптических приборов. Метод темного поля. Метод фазового контраста

§ 37. Пространственная фильтрация изображений

Существо пространственной фильтрации изображений. Пространственная фильтрация изображений дифракционной решетки. Эксперимент Аббе—Портера

§ 38. Голография

Синхронное детектирование. Голограмма плоской волны. Восстановление изображения. Голограмма точечного объекта. Голограмма произвольного объекта. Требования к фотопластинкам и времени экспозиции. Объемное воспроизведение предмета. Толстослойные гологramмы (метод Денисюка). Условие Вульфа—Брэгга. Получение голограммы и восстановление плоской волны. Получение голограммы и восстановление сферической волны. Получение голограммы и восстановление изображения произвольного объекта. Цветное объемное изображение. Особенности голограмм как носителей информации. Применения голографии

8	Распространение света в анизотропных средах	262
	§ 39. Описание анизотропных сред	262
	Источники анизотропии. Описание анизотропной диэлектрической среды. Тензор диэлектрической проницаемости	
	§ 40. Распространение плоской электромагнитной волны в анизотропной среде	264
	Плоская электромагнитная волна в анизотропной среде. Зависимость фазовой скорости от направлений распространения волны и колебаний вектора D . Уравнение Френеля. Типы возможных волн	
	§ 41. Ход лучей в анизотропной среде	267
	Зависимость лучевой скорости от направления. Эллипсоид лучевых скоростей. Анализ хода лучей с помощью эллипса лучевых скоростей. Оптическая ось. Двусосные и одноосные кристаллы. Эллипсоид волновых нормалей. Лучевая поверхность	
	§ 42. Двойное лучепреломление	272
	Обыкновенный и необыкновенный лучи. Сущность двойного лучепреломления. Построение Гюйгенса. Оптическая ось перпендикулярна поверхности кристалла. Оптическая ось параллельна поверхности кристалла. Оптическая ось под углом к поверхности кристалла. Закон Малюса. Поляризация при двойном лучепреломлении. Поляризатор. Поляризационные и двоякопереломляющие призмы. Призма Николя. Двоякопереломляющие призмы. Полихроизм	
	§ 43. Интерференция поляризованных волн	276
	Интерференция волн при взаимно перпендикулярных направлениях линейной поляризации. Пластина в четверть волны. Пластина в полволны. Пластина в целую волну. Анализ линейно поляризованного света. Анализ эллиптически поляризованного света. Анализ циркулярно поляризованного света. Компенсаторы. Цвета кристаллических пластинок. Явления в сходящихся лучах	
	§ 44. Вращение плоскости поляризации	281
	Вращение плоскости поляризации в кристаллических телах. Вращение плоскости поляризации в аморфных веществах. Феноменологическая теория вращения плоскости поляризации. Оптическая изомерия. Вращение плоскости поляризации в магнитном поле	
	§ 45. Искусственная анизотропия	285
	Анизотропия при деформации. Анизотропия, создаваемая в веществе электрическим полем. Анизотропия, создаваемая в веществе магнитным полем. Эффект Поккельса	
	Задачи	288
9	Рассеяние света	290
	§ 46. Природа процессов рассеяния	290
	Природа рассеяния. Типы рассеяния. Многократное рассеяние	
	§ 47. Рэлеевское рассеяние и рассеяние Ми	291
	Модель элементарного рассеивателя. Рэлеевское рассеяние. Закон Рэлея. Угловое распределение и поляризация света при рэлеевском рассеянии. Ослабление интенсивности света. Рассеяние Ми. Распределение интенсивности по углам и поляризация излучения в рассеянии Ми. Проявления рассеяния Ми	
	§ 48. Рассеяние Мандельштама—Бриллюэна	297
	Компоненты Мандельштама—Бриллюэна. Несмешенная компонента. Явление Мандельштама—Бриллюэна в твердых телах	
	§ 49. Комбинационное рассеяние	298
	Классическая интерпретация. Экспериментальные факты. Квантовая интерпретация. Применения комбинационного рассеяния	
10	Генерация света	302
	§ 50. Излучение абсолютно черного тела	302
	Плотность излучения. Равновесная плотность излучения. Первый закон Кирхгофа. Поглощающая способность и энергетическая светимость. Второй закон Кирхгофа. Абсолютно черное тело. Концентрация мод колебаний. Формула Рэлея—Джинса. Формула Вина. Формула Планка. Закон Стефана—Больцмана. Закон смещения Вина. Элементарная квантовая теория. Спонтанные и вынужденные переходы. Коэффициенты Эйнштейна	

**11
Нелинейные
явления в
оптике**

§ 51. Оптические усилители	309
Прохождение света через среду. Закон Бургера. Условия усиления. Воздействие светового потока на заселенность уровней. Условия насыщения. Создание инверсной заселенности	
§ 52. Лазеры	312
Принципиальная схема лазера. Порог генерации. Условия стационарной генерации. Добротность. Непрерывные и импульсные лазеры. Повышение мощности излучения. Метод модулированной добротности	
§ 53. Лазерное излучение	315
Моды излучения. Резонатор с прямоугольными плоскими зеркалами. Аксиальные (продольные) моды. Ширина линий излучения. Боковые моды. Цилиндрический резонатор со сферическими зеркалами. Синхронизация мод. Продолжительность импульса. Осуществление синхронизации мод. Лазерные спеклы	
§ 54. Характеристики некоторых лазеров	321
Разнообразие лазеров. Рубиновый лазер. Гелий-неоновый лазер. CO ₂ -лазер с замкнутым объемом. Проточный CO ₂ -лазер. Т-лазер. Газодинамические лазеры. Лазеры на красителях	
Задачи	326
§ 55. Нелинейная поляризованность	328
Линейная поляризованность. Нелинейная поляризованность. Квадратичная нелинейность. Нелинейная восприимчивость. Комбинационные частоты	
§ 56. Генерация гармоник	330
Волна линейной поляризованности. Волны нелинейной поляризованности. Условие пространственного синхронизма. Длина когерентности. Осуществление пространственного синхронизма. Векторное условие пространственного синхронизма. Генерация суммарных и разностных частот. Спонтанный распад фотона. Параметрическое усиление света. Параметрические генераторы света	
§ 57. Самовоздействие света в нелинейной среде	338
Нелинейная поправка к показателю преломления. Самофокусировка и дефокусировка пучка. Длина самофокусировки. Пороговая мощность. Основные причины возникновения нелинейности показателя преломления. Инерционность	
Приложение. Единицы СИ, используемые в книге	342

Предисловие

Стремительное внедрение достижений науки в производство — необходимая черта сегодняшнего дня. Одним из ярких примеров здесь является именно оптика. Бурное развитие лазерной техники стало не только научным достижением, но и во многих отраслях революционизировало промышленную технологию. Важнейшее значение приобрела сегодня проблема подготовки квалифицированных кадров для отраслей промышленности, внедряющих новейшие научные разработки.

«Значительно улучшить подготовку в вузах и техникумах специалистов для ведущих отраслей народного хозяйства...», — призвал июньский (1983 г.) Пленум ЦК КПСС.

Применение лазеров и их использование совместно с ЭВМ создали весьма благоприятные условия для развития оптики. Высокая когерентность лазерного излучения позволяет изучать и воспроизвести в оптическом диапазоне широкий класс явлений, недоступных для исследований при малых степенях когерентности излучения. Высокая плотность энергии лазерного излучения дает возможность исследовать нелинейные оптические процессы в условиях, недоступных при прежних методах исследования. Возможность генерации коротких и сверхкоротких лазерных импульсов открыла путь к исследованию быстропротекающих процессов, включая внутримолекулярные. Использование ЭВМ в громадной степени ускорило оптические исследования, поскольку во многих случаях оно свело их либо к прямому расчету, либо к постановке численных экспериментов.

Все это за последние 25 лет привело значительному развитию оптики, существенно расширились ее приложения. Начало этому процессу было положено важными работами, приведшими к созданию квантовых генераторов излучения. Наряду с фундаментальными работами по мазерам и лазерам советскими физиками внесен большой вклад в развитие многих важных разделов оптики. Например, таких, как рассеяние света, голограмия, оптические системы, нелинейная оптика и т. д. В этом развитии оптики фундаментальные основы ее, естественно, не претерпели существенных изменений. В ряде случаев они были прояснены, а в других случаях — обогащены проникновением понятий, методов, математических приемов и т. д. из других областей науки (например, теории случайных процессов, физики линейных и нелинейных колебаний, матричных методов расчета и т. д.).

Содержание книги достаточно полно отражено в оглавлении. Несколько больше внимания, чем обычно, уделено статистическим свойствам света и спектральному представлению. Дифракция изложена в рамках интеграла Кирхгофа. На материале геометрической оптики и интерференции в тонких пленках показана эффективность матричных методов. Дифракционная теория формирования изображений, пространственная фильтрация изображений, голограмия и другие аналогичные вопросы представлены единообразно в рамках Фурье-оптики. Анализ частичной когерентности и частичной поляризации проводится в рамках первой корреляционной функции.

Математическая сторона излагаемого материала представлена в возможно простой форме, совместимой с достаточной строгостью изложения. В необходимых случаях даются математические пояснения и приводятся более детальные расчеты. Громоздкость некоторых из этих расчетов не должна создавать у студента впечатления о слож-

ности используемого математического аппарата. Чтобы рассеять это впечатление, ему необходимо лишь запастись терпением и провести самостоятельно эти расчеты.

Наиболее существенное отличие курса оптики от курсов механики, молекулярной физики и электричества состоит в том, что его фундаментальные основы лежат вне курса. Это обстоятельство приводит к значительному усилению роли дедуктивного метода изложения. Поэтому изложение в основном носит дедуктивный характер, а анализ экспериментальных данных в большинстве случаев (хотя и не всегда) призван либо продемонстрировать согласие выводов теории с результатами опытов, либо объяснить наблюдаемые явления.

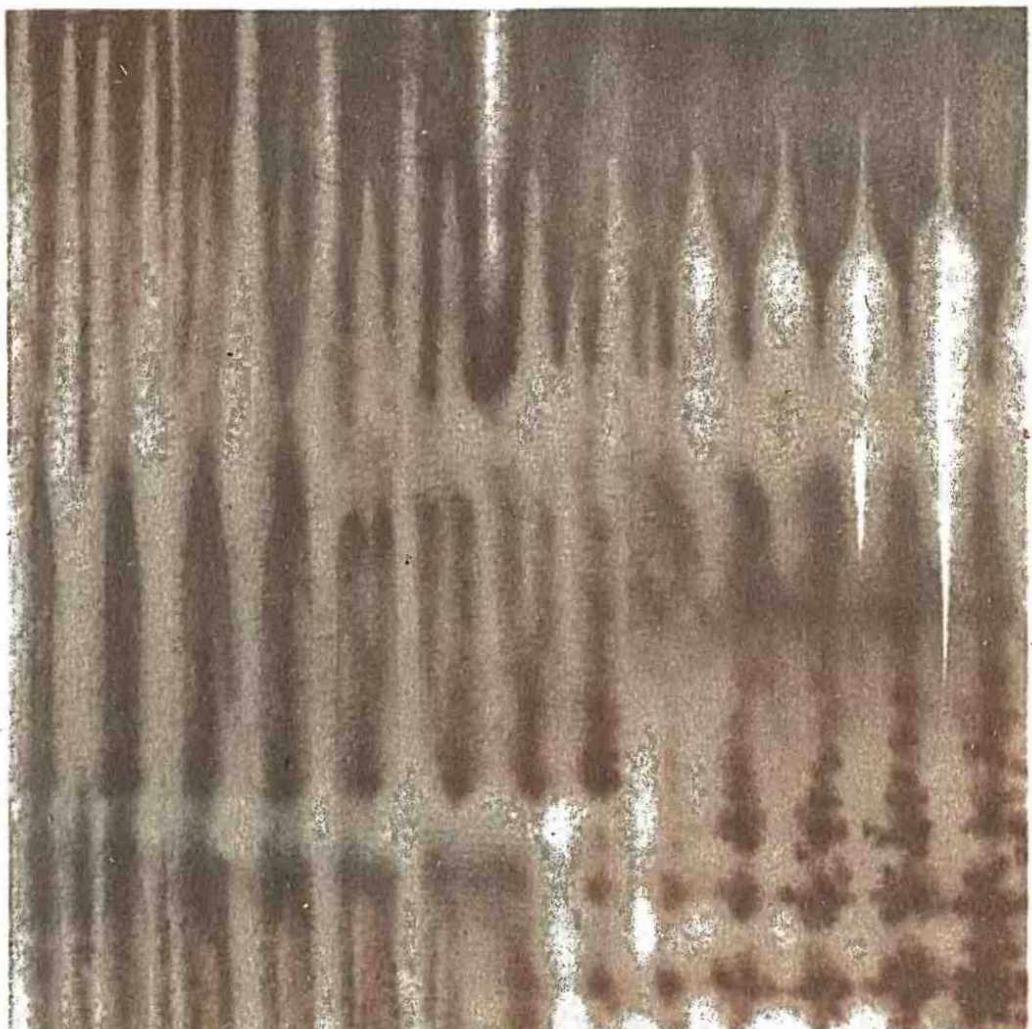
Книга написана на основе многолетнего опыта преподавания автора на физическом факультете Московского государственного университета им. М. В. Ломоносова. Автор благодарен своим коллегам, дискуссии с которыми оказали влияние на общий план книги.

Автор благодарен академику АН УССР А. И. Ахиезеру, а также проф. Н. И. Калитеевскому и сотрудникам возглавляемой им кафедры за внимательное рецензирование рукописи и ценные замечания.

Автор

Основная идея:
Основываясь
на электромагнитной
природе света,
изучить свойства
монохроматических
световых волн
с помощью
уравнений Максвелла.

Электромагнитные волны



§ 1 Оптический диапазон электромагнитных волн

Анализируются факторы, делающие волны видимого

диапазона наиболее подходящими для зрения.

Описываются характеристики оптического диапазона.

Длины волн видимого диапазона. Видимый диапазон включает электромагнитные волны, воспринимаемые человеческим глазом. Граница диапазона этих волн зависит от индивидуальных особенностей глаза и варьируется приблизительно в пределах

$$\lambda = 0,38 \div 0,76 \text{ мкм}. \quad (1.1)$$

Частоты волн видимого диапазона. В оптике используется как круговая частота

$$\omega = 2\pi/T, \quad (1.2)$$

где T — период колебаний волны, так и частота

$$v = 1/T, \quad (1.3)$$

связанные очевидным соотношением

$$\omega = 2\pi v. \quad (1.4)$$

Частота выражается в герцах, а круговая частота — в секундах в минус первой степени. Принимая во внимание, что

$$\lambda = cT, \quad (1.5)$$

где $c = 3 \cdot 10^8$ м/с — скорость света в вакууме, для границ видимого диапазона получаем:

$$v = (4 \div 8) \cdot 10^{14} \text{ Гц}; \quad (1.6) \quad \omega = (2,5 \div 5,0) \cdot 10^{15} \text{ с}^{-1}. \quad (1.7)$$

Оптический и другие диапазоны электромагнитных волн. Теоретически мыслимым является существование всех частот от $v = 0$ до $v = \infty$. Однако корпускулярные свойства излучения накладывают на эти возможности ограничения. Как показывается в квантовой теории, электромагнитное излучение существует в виде «порций» энергии (квантов). Энергия кванта излучения связана с его частотой формулой

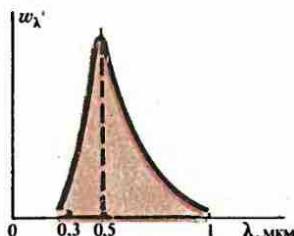
$$E = \hbar\omega = h\nu, \quad (1.8)$$

где $\hbar = 6,62 \cdot 10^{-34}$ Дж с — постоянная Планка; $\hbar = h/(2\pi) = 1,05 \cdot 10^{-34}$ Дж·с называется также постоянной Планка (новой).

Обе эти величины встречаются одинаково часто в зависимости от обстоятельств. Как видно (в 1.8), постоянную \hbar удобно выбирать в случае оперирования с круговыми частотами, а h — при использовании частот v .

Из (1.8) следует, что бесконечные частоты $v = \infty$ невозможны, поскольку соответствующие кванты излучения обладали бы бесконечной энергией. Соотношение (1.8) также дает ограничение на малые частоты, если существует минимально возможное значение энергии кванта E_0 . А это означает, что и частота не может быть меньше $v_0 = E_0/h$. В настоящее время в физике нет никаких свидетельств ограничения снизу энергии фотонов электромагнитного излучения. Следовательно, частоты электромагнитных волн не ограничены снизу. Минимальная частота (около 8 Гц) наблюдается в стоячих электромагнитных волнах между ионосферой и земной поверхностью. Отсюда можно заключить [см. (1.8)], что минимальная энергия квантов электромагнитного излучения по крайней мере меньше 10^{-33} Дж.

Всевозможные частоты электромагнитных волн подразделяются на следующие диапазоны:



1
Спектр излучения Солнца

Название диапазона волн	Границы диапазона	
	по длине волн λ	по энергии E квантов
Гамма-излучение	<0,0012 нм	> 1 МэВ
Излучение:		
рентгеновское	0,0012—12 нм	100 эВ — 1 МэВ
ультрафиолетовое	12—380 нм	3,2—100 эВ
видимое	380—760 нм	1,6—3,2 эВ
инфракрасное	760— 10^6 нм = 1 мм	$1,2 \cdot 10^{-3}$ — 1,6 эВ
Радиоволны	> 1 мм	$< 1,2 \cdot 10^3$ эВ

Каждый из диапазонов имеет свои характерные особенности. С увеличением частоты волны усиливается проявление корпускулярных свойств излучения. Волны разных диапазонов различаются также методами генерации излучения. Каждый из диапазонов служит предметом изучения соответствующего раздела физики.

Видимый диапазон и примыкающие к нему диапазоны ультрафиолетового и инфракрасного излучений в совокупности составляют диапазон электромагнитных волн, изучаемый в оптике. Кванты излучения видимого диапазона называются фотонами. Они имеют энергию в интервале

$$E = (2,6 \div 5,2) \cdot 10^{-19} \text{ Дж} = 1,6 \div 3,2 \text{ эВ}. \quad (1.9)$$

Почему мы видим именно в видимом диапазоне? При температуре выше 0 К все материальные тела излучают электромагнитные волны, которые поглощаются и отражаются (рассеиваются) материальными телами. Интенсивность излучения отражения и поглощения зависит от частоты излучения, температуры, свойств вещества и других факторов. Наиболее интенсивным источником электромагнитного излучения, определяющим «радиационную обстановку» вблизи земной поверхности, является Солнце. Температура поверхности Солнца составляет около 6000 К. Спектр его излучения является спектром излучения абсолютно черного тела (см. § 50). Максимум интенсивности излучения по длинам волн приходится примерно на длину волны 0,5 мкм (рис. 1).

При прохождении света через атмосферу Земли в результате рассеяния и поглощения состав солнечного спектра существенно меняется в зависимости от толщины проходящего светом слоя воздуха, запыленности и других факторов. В результате спектр у поверхности Земли обрывается примерно на волне $\lambda \approx 0,3$ мкм. Волны с меньшей длиной волны от Солнца поверхности Земли не достигают. Это обусловли-

- Длины волн видимого диапазона заключены примерно в пределах 0,38—0,76 мкм, а оптический диапазон включает также и инфракрасную и ультрафиолетовую области спектра. Видимый диапазон наиболее подходит для зрения потому, что на меньшие длины волн днем вблизи поверхности Земли приходится слишком мало доли энергии, а на больших длинах волн зрению мешают шумы.

- Какие изменения претерпевает солнечный спектр при прохождении через атмосферу? Укажите границы видимого диапазона по частотам и круговым частотам. Укажите границы оптического диапазона длин волн.

вается поглощением их озоном O_3 в верхних слоях атмосферы. Ослабление волн зависит главным образом от высоты Солнца над уровнем горизонта, что наглядно характеризуется следующими данными:

Цвет, область спектра	Доля излучения, достигающего поверхности Земли при положении Солнца		Диапазон длин волн, мкм	Энергия излучения, %	
	в зените	при заходе или восходе		в спектре Солнца	у поверх- ности Земли
Красный, 0,65 мкм	0,96	0,21	0,3—0,4	5	1
Зеленый, 0,52 мкм	0,9	0,024	0,4—0,75	52	40
Фиолетовый, 0,41 мкм	0,76	0,000065	0,75—2,3	43	59

Таким образом, при прохождении атмосферы наиболее сильно ослабляется коротковолновая часть спектра. Это является, в частности, причиной покраснения Солнца при восходе и заходе. За счет поглощения в атмосфере доля энергии ультрафиолетовой части спектра уменьшается, а инфракрасной и микроволновой — увеличивается. Например, если взять энергию, приходящуюся на интервал длин волн от 0,3 до 2,3 мкм за 100%, то распределение энергии в солнечном спектре до прохождения атмосферы и у поверхности Земли характеризуется данными, приведенными в таблице выше.

Видение предметов осуществляется посредством отраженного света. Поэтому наиболее подходящим для зрения является интервал вблизи длины волны, на которую приходится максимальная интенсивность излучения, т. е. $\lambda \approx 0,5$ мкм, причем интервал должен быть таким, чтобы на него приходилась значительная часть полной энергии излучения. Этим условиям полностью удовлетворяет видимый диапазон электромагнитных волн. Является вполне естественным, что в результате эволюции именно в этом диапазоне развилась способность человека к зрению.

Однако достаточно много энергии (больше 50%) приходится на микроволновую часть спектра. Поэтому энергетические соображения в принципе не исключают возможности развития способности зрения у человека в этом диапазоне. Тем не менее эта область спектра непригодна для зрения.

Почему микроволновая область непригодна для зрения? Волновые свойства излучения приводят к ухудшению качества зрения с увеличением длины волны, поскольку при этом ухудшается разрешающая способность как оптических приборов (см. § 36), так и глаза. Кроме того, с увеличением длины волны необходимо увеличивать геометрические размеры приемных устройств, в том числе и биологических элементов, связанных со зрением. Это, безусловно, отрицательный фактор в отборе, осуществляемом в процессе эволюции.

Однако главная причина непригодности микроволнового диапазона для зрения связана с корпускулярными свойствами электромагнитного излучения и существованием больших «шумов» в этом диапазоне, которые делают невозможным зрение глазами, имеющими температуру порядка температуры тела человека или животных.

Прежде всего покажем невозможность зрения в микроволновом диапазоне с помощью отраженного солнечного излучения, а затем рассмотрим условия, при которых в микроволновом диапазоне можно видеть предметы («ночное видение»).

Как известно из молекулярной физики, средняя концентрация фотонов, приходящихся на одну моду колебаний (см. § 50) с частотой ω , в равновесном излучении при температуре T задается формулой

$$\langle n \rangle = \frac{1}{\exp[\hbar\omega/(kT)] - 1},$$

где $k = 1,38 \cdot 10^{-23}$ Дж/К — постоянная Больцмана.

Речь идет о различных типах колебаний, имеющих одну и ту же частоту, но отличающихся друг от друга поляризацией, направлением распространения соответствующих волн и другими особенностями, и о колебаниях различных частот. Средняя концентрация фотонов, приходящаяся на частоту ω , равна сумме средних концентраций фотонов, приходящихся на различные моды или типы колебаний, имеющих эту частоту, т. е. выражается формулой (1.10). Сравнение концентраций фотонов с различной частотой при разных температурах в равновесных условиях сводится к сравнению средних концентраций (1.10), поскольку коэффициенты пропорциональности, учитывающие число мод, одинаковы.

Вблизи поверхности Земли имеются фотоны солнечного излучения и фотоны излучения поверхности Земли и находящихся на ней предметов. Считая для определенности температуру у поверхности Земли равной $T_3 = 300$ К, для концентрации этих фотонов, приходящихся на одну моду колебаний с частотой ω , можем написать

$$\langle n_3 \rangle = \frac{1}{\exp[\hbar\omega/(kT_3)] - 1}. \quad (1.11)$$

Температура поверхности Солнца равна примерно $T = 6000$ К. Вблизи поверхности Солнца концентрация фотонов равновесного теплового излучения задается формулой (1.10) при $T = T_C$. Энергия излучения, дошедшего до границ атмосферы Земли, имеет то же распределение по частотам, что и вблизи поверхности Солнца, но соответственно меньшую концентрацию. При прохождении атмосферы в результате рассеяния и поглощения спектральный состав света изменяется. Затем происходит отражение и поглощение света поверхностью Земли и находящимися на ней предметами. В результате этого спектральный состав излучения, обусловленного солнечным освещением, весьма сложно зависит от условий его образования. Однако для оценки спектрального состава излучения с точностью до порядков величин можно пренебречь всеми этими изменениями и считать, что он у поверхности Земли примерно таков же, как у исходного солнечного излучения. Поэтому для средней концентрации $\langle n_C \rangle$ фотонов солнечного излучения вблизи поверхности Земли в соответствии с (1.10) можно написать

$$\langle n_C \rangle = \left(\frac{r_C}{R}\right)^2 \frac{1}{\exp[\hbar\omega/(kT_C)] - 1}, \quad (1.12)$$

где r_C — радиус Солнца, R — радиус земной орбиты. Множитель $(r_C/R)^2$ учитывает ослабление плотности солнечного излучения при удалении от Солнца, обратно пропорциональное квадрату расстояния. Учитывая, что $r_C = 695,5 \cdot 10^3$ км, $R = 149,5 \cdot 10^6$ км, находим $(r_C/R)^2 \approx 2,16 \cdot 10^{-5}$.

Решающее значение для эффективного зрения имеет соотношение между потоком фотонов, несущих информацию о предмете, и потоком тепловых фотонов, которые никакой информации не несут и создают просто фоновый шум. Чем больше превышение потока фотонов, несущих информацию, над потоком фотонов, создающих шум, тем лучше условия зрения. Поток фотонов, несущих информацию о предмете, возникает за счет отражения предметом излучения солнечного происхождения. Можно считать, что он пропорционален средней концентрации фотонов $\langle n_C \rangle$. Поток тепловых фотонов, создающих шум (будем называть их шумовыми), пропорционален $\langle n_3 \rangle$.

Оценим эффективность зрения в видимом диапазоне. В качестве примера рассмотрим фотоны с длиной волны $\lambda = 0,5$ мкм. Их энергия равна $\hbar\omega = 2\pi\hbar c/\lambda \approx 4 \cdot 10^{-19}$ Дж $\approx 2,5$ эВ. Учитывая, что $kT_C = 8,3 \cdot 10^{-20}$ Дж $\approx 0,52$ эВ, $kT_3 = 0,4 \cdot 10^{-20}$ Дж $\approx 0,026$ эВ, для таких фотонов находим $\hbar\omega/(kT_C) = 4,8$ и $\hbar\omega/(kT_3) = 96$.

Для концентрации фотонов в соответствии с (1.11) и (1.12) находим:

$$\langle n_3 \rangle \approx e^{-96} \text{ м}^{-3} \approx 10^{-4.2} \text{ м}^{-3}, \quad (1.13) \quad \langle n_C \rangle \approx 2 \cdot 10^{-5} e^{-4} \text{ м}^{-3} \approx 10^{-7} \text{ м}^{-3}. \quad (1.14)$$

Отношение числа шумовых фотонов к числу фотонов, несущих информацию, составляет $\langle n_3 \rangle / \langle n_C \rangle \approx 10^{-35}$, т. е. совершенно ничтожно. Практически на этой длине волны никакого шума нет.

Теперь рассмотрим ситуацию в микроволновом диапазоне (например, для $\lambda = 2$ мкм). В этом случае $\hbar\omega = 0,625$ эВ, $\hbar\omega/(kT_C) \approx 1,2$, $\hbar\omega/(kT_3) \approx 24$ и, следовательно,

$$\langle n_3 \rangle \approx e^{-24} \text{ м}^{-3} \approx 10^{-10} \text{ м}^{-3}, \quad (1.15) \quad \langle n_C \rangle \approx 2 \cdot 10^{-5} \cdot 0,43 \text{ м}^{-3} \approx 10^{-5} \text{ м}^{-3}. \quad (1.16)$$

Поэтому $\langle n_3 \rangle / \langle n_C \rangle \approx 10^{-5}$ (ср. предыдущий результат), т. е. условия для зрения неизмеримо хуже. Однако на первый взгляд кажется, что в абсолютном смысле ситуация здесь не так уж плоха, поскольку в данной моде один шумовой фотон приходится на 10^5 фотонов излучения, имеющего солнечное происхождение. Но это не так. Надо принять во внимание, что видение осуществляется с помощью отраженных от предмета лучей. От точки предмета лучи достигают глаза, распространяясь в очень малом телесном угле $\Delta\Omega \approx S/r^2$ (S — площадь зрачка, r — расстояние от зрачка до предмета). Кроме того, отраженный луч в этом телесном угле содержит не всевозможные моды данной частоты, а лишь некоторые, обусловленные характером отражения. Тепловые же фотоны в полости глаза присутствуют во всех модах и падают на все точки сетчатки глаза со всех направлений, т. е. из телесного угла 2π . Это обстоятельство увеличивает относительное число тепловых фотонов приблизительно в $a = 2\pi/\Delta\Omega = 2\pi r^2/S$ раз. Считая, что радиус зрачкового отверстия равен 2 мм, находим $S = \pi(0,2)^2 \text{ см}^2$. Поэтому для $r = 100$ см получаем $a = 0,5 \cdot 10^6$. Это приводит к тому, что вместо числа 10^5 получается $\langle n_3 \rangle / \langle n_C \rangle = 10$, что делает зрение на таких длинах волн невозможным. Заметим, что все сказанное об относительном увеличении числа шумовых квантов применимо и к видимому диапазону, однако ничтожное малое значение отношения $\langle n_3 \rangle / \langle n_C \rangle \approx 10^{-35}$ не увеличивается настолько, чтобы зрение стало невозможным.

Резюмируя, можно сказать, что видимый диапазон наиболее подходит для зрения, потому что на более короткие волны приходится слишком малая доля энергии, а на более длинных волнах зрению мешают тепловые шумы.

Ночное видение. Ночью фотоны солнечного излучения вблизи поверхности Земли почти полностью отсутствуют (возможное лунное освещение и свечение неба здесь не принимаются во внимание). Однако тепловое излучение материальных тел при температуре 300 К наиболее интенсивно вблизи волны $\lambda \approx 10$ мкм. Термодинамическое равновесие между излучением и материальными телами у поверхности Земли с наступлением темноты не устанавливается, поскольку условия непрерывно изменяются, и со стороны неба система открыта. Поэтому все предметы и земная поверхность представляются «светящимися» на длине волны $\lambda = 10$ мкм. Распределение энергии излучения по спектру существенно зависит от поглощающей способности воздушной среды и может быть учтено.

Представим себе, что хрусталик человеческого глаза способен фокусировать излучение с длиной волны $\lambda = 10$ мкм на сетчатке, которая в состоянии воспринимать это излучение и вырабатывать соответствующие нервные импульсы для создания зрительного ощущения. Спрашивается, будет ли человек видеть окружающие его предметы? Из изложенного выше следует, что поток шумовых фотонов на любой участок сетчатки глаза существенно превосходит поток фотонов, образующих на сетчатке глаза изображение предметов, и поэтому человек не в состоянии видеть окружающие его предметы. Для того чтобы зрение стало возможным, необходимо уменьшить плотность шумовых фотонов внутри глаза, т. е. существенно уменьшить температуру глаза. Расчет показывает, что необходимые для этого температуры очень малы и составляют несколько десятков кельвин. Поэтому осуществить ночное зрение можно лишь с помощью приборов, поддерживаемых при достаточно низкой температуре. Получаемое в таких приборах изображение в микроволновом диапазоне преобразуется в изображение в длинах волн видимого диапазона и наблюдается глазом при обычных температурах. В процессе

преобразования изображений соответствующие сигналы могут быть усилены или подвергнуты соответствующей обработке, что позволяет получить высококачественное видимое изображение.

§ 2 Свойства электромагнитных волн

С помощью уравнений Максвелла выводятся основные свойства электромагнитных волн.

Электромагнитная природа света. Существование электромагнитных волн было теоретически предсказано Максвеллом (1862—1864) как прямое следствие из уравнений электромагнитного поля. Скорость электромагнитных волн в вакууме оказалась равной величине $1/\sqrt{\epsilon_0\mu_0}$ (в современных обозначениях), называемой в то время **электродинамической постоянной**. Ее числовое значение ($3,1 \cdot 10^8$ м/с) было получено несколько раньше (1856) из электромагнитных измерений В. Е. Вебера (1804—1891) и Р. Г. Кольрауша (1809—1858). Оно почти совпадало со скоростью света в вакууме, равной, по измерениям И. Л. Физо (1819—1896) в 1849 г., $c = 3,15 \cdot 10^8$ м/с. Другое важное совпадение в свойствах электромагнитных волн и света обусловлено поперечностью волн: Поперечность электромагнитных волн следует из уравнений Максвелла, а поперечность световых волн — из экспериментов по поляризации света (Юнг, 1817). Эти два факта привели Максвелла к заключению, что свет представляет собой электромагнитные волны.

Существование электромагнитных волн экспериментально было доказано в 1888 г. Г. Р. Герцем (1857—1894). Длина волн, генерированных и детектированных, составляла примерно 66 см. С помощью металлического зеркала Герц наблюдал отражение и преломление волн, изучил их поляризацию, получил стоячие волны, доказав тем самым их способность к интерференции.

Волновое уравнение. Уравнения Максвелла для вакуума при отсутствии токов ($j=0$) и зарядов ($\rho=0$) имеют следующий вид:

$$\text{rot } \mathbf{H} = \partial \mathbf{D} / \partial t, \quad (2.1)$$

$$\text{rot } \mathbf{E} = -\partial \mathbf{B} / \partial t, \quad (2.2)$$

$$\text{div } \mathbf{B} = 0, \quad (2.3)$$

$$\text{div } \mathbf{D} = 0, \quad (2.4)$$

$$\mathbf{D} = \epsilon_0 \mathbf{E},$$

$$\mathbf{B} = \mu_0 \mathbf{H}, \quad (2.5)$$

где ϵ_0 и μ_0 — соответственно **электрическая** и **магнитная** постоянные. Применяя к обеим частям уравнения (2.1) операцию **rot**, получаем

$$\frac{1}{\mu_0} \text{rot rot } \mathbf{B} = -\epsilon_0 \frac{\partial}{\partial t} (\text{rot } \mathbf{E}), \quad (2.6)$$

где учтены соотношения (2.5) и принято во внимание, что порядок дифференцирований по независимым переменным (пространственным координатам и времени) можно изменить. Принимая во внимание векторное равенство

$$\text{rot rot } \mathbf{B} = \text{grad div } \mathbf{B} - \nabla^2 \mathbf{B} \quad (2.7)$$

и заменяя в правой части (2.6) **rot E** его выражением (2.2), получаем уравнение для **B**:

$$\nabla^2 \mathbf{B} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{B}}{\partial t^2} = 0. \quad (2.8)$$

Аналогично, применяя операцию **rot** к обеим частям равенства (2.2), получаем уравнение для **E**:

$$\nabla^2 \mathbf{E} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0.$$

Оператор

$$\square = \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}, \quad (2.10)$$

где $c = 1/\sqrt{\epsilon_0 \mu_0}$ — скорость света в вакууме, называется оператором Д'Аламбера. С его помощью волновые уравнения (2.8) и (2.9) могут быть записаны в форме

$$\square \mathbf{E} = 0, \quad \square \mathbf{B} = 0. \quad (2.11)$$

Плоские волны. Большую роль в физике играет волновое уравнение. Для скалярной функции Φ оно имеет вид

$$\nabla^2 \Phi - \frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2} = 0. \quad (2.12)$$

Найдем общее решение этого уравнения для случая, когда Φ зависит только от одной из декартовых координат, например z , т. е. $\Phi = \Phi(z, t)$. Это означает, что Φ имеет постоянное значение в точках плоскости, перпендикулярной оси Z . В этом случае уравнение (2.12) принимает вид

$$\frac{\partial^2 \Phi}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2} = 0. \quad (2.13)$$

Используя новые независимые переменные

$$\xi = z - ct, \quad \eta = z + ct, \quad (2.14)$$

получаем

$$\frac{\partial \Phi}{\partial z} = \frac{\partial \Phi}{\partial \xi} \frac{\partial \xi}{\partial z} + \frac{\partial \Phi}{\partial \eta} \frac{\partial \eta}{\partial z} = \frac{\partial \Phi}{\partial \xi} + \frac{\partial \Phi}{\partial \eta}, \quad (2.15a)$$

$$\frac{\partial \Phi}{\partial t} = \frac{\partial \Phi}{\partial \xi} \frac{\partial \xi}{\partial t} + \frac{\partial \Phi}{\partial \eta} \frac{\partial \eta}{\partial t} = -c \frac{\partial \Phi}{\partial \xi} + c \frac{\partial \Phi}{\partial \eta}. \quad (2.15b)$$

Разделив обе части уравнения (2.15b) на c и вычитая их почленно из левых и правых частей уравнения (2.15a), находим

$$\frac{\partial}{\partial z} - \frac{1}{c} \frac{\partial}{\partial t} = 2 \frac{\partial}{\partial \xi}. \quad (2.16)$$

Аналогично, почленное сложение правых и левых частей тех же уравнений дает

$$\frac{\partial}{\partial z} + \frac{1}{c} \frac{\partial}{\partial t} = 2 \frac{\partial}{\partial \eta}. \quad (2.17)$$

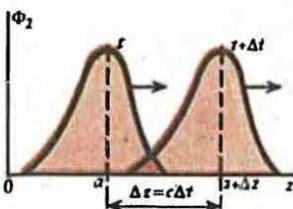
Тогда

$$\left(\frac{\partial}{\partial z} - \frac{1}{c} \frac{\partial}{\partial t} \right) \left(\frac{\partial}{\partial z} + \frac{1}{c} \frac{\partial}{\partial t} \right) = \frac{\partial^2}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}.$$

С учетом (2.16) и (2.17) преобразуем уравнение (2.13) к виду

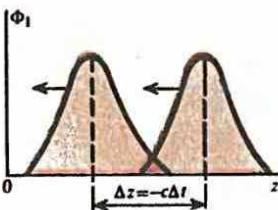
$$\frac{\partial^2 \Phi}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2} = 4 \frac{\partial}{\partial \xi} \frac{\partial}{\partial \eta} \Phi = 0. \quad (2.18)$$

Интегрируя (2.18) по ξ , получаем независимую от ξ функцию, которая в данном случае может зависеть только от η , т. е. является произвольной функцией $\Psi(\eta)$. После этого уравнение (2.18) принимает вид



2

Волна движется в направлении положительных значений z : $\Phi = \Phi_2(z - ct)$



3

Волна движется в направлении отрицательных значений z : $\Phi = \Phi_1(z + ct)$

○ Почему не может существовать универсального соотношения между частотой волны и волновым числом?

Какое универсальное соотношение существует между частотой и волновым числом в изотропной среде с постоянной скоростью распространения волны?

Откуда следует инвариантность плоской волны?

Как преобразуются частота волны и волновой вектор при переходе между инерциальными системами отсчета?

$$\frac{\partial \Phi}{\partial \eta} = \Psi(\eta). \quad (2.19)$$

Интегрируя (2.19) по η , получаем

$$\Phi = \int \Psi(\eta) d\eta = \Phi_1(\eta) + \Phi_2(\xi), \quad (2.20)$$

где $\Phi_1(\eta)$ — первообразная функция в интеграле от $\Psi(\eta)$ по $d\eta$, $\Phi_2(\xi)$ — постоянная интегрирования. Как видно по ходу решения, функции Φ_1 и Φ_2 произвольны. С учетом (2.14) общее решение (2.20) уравнения (2.13) может быть записано в виде

$$\Phi(z, t) = \Phi_1(z + ct) + \Phi_2(z - ct). \quad (2.21)$$

Выясним физический смысл этого решения. Сначала проанализируем решение

$$\Phi = \Phi_2(z - ct). \quad (2.22)$$

График Φ_2 как функции от z в моменты времени t и $t + \Delta t$ изображен на рис. 2. Видно, что значение аргумента функции в точке z в момент t совпадает со значением аргумента функции в точке $z + \Delta z$ в момент $t + \Delta t$, если $\Delta z = c\Delta t$, поскольку

$$z - ct = z + \Delta z - c(t + \Delta t) \quad (\Delta z = c\Delta t). \quad (2.23)$$

Поэтому график функции для $t + \Delta t$ получается из графика для t смещением всех точек кривой в направлении положительных значений оси Z на $\Delta z = c\Delta t$. Следовательно, скорость волны равна $v = \Delta z / \Delta t = c$. Функция $\Phi_2(z - ct)$ описывает волну произвольной формы, движущуюся со скоростью c в направлении положительных значений оси Z . В процессе движения значение Φ_2 в каждой точке волны и форма волны не изменяются.

Физический смысл Φ_1 , т. е. решения

$$\Phi = \Phi_1(z + ct), \quad (2.24)$$

выясняется аналогично. Учитывая, что

$$z + ct = z + \Delta z + c(t + \Delta t) \quad (\Delta z = -c\Delta t), \quad (2.25)$$

заключаем, что функция $\Phi_1(z + ct)$ описывает волну произвольной формы, движущуюся со скоростью c в направлении отрицательных значений оси Z (рис. 3). Значение Φ_1 в каждой точке волны и форма волны в процессе движения не изменяются.

Волна, описываемая формулой (2.21), является суперпозицией двух волн, движущихся в противоположных направлениях. В этом случае уже нельзя говорить о скорости или направлении волны. В простейшем случае получается стоячая волна, а в общем случае — сложное электромагнитное поле, которое требует специального изучения.

Значение функции Φ для фиксированных z и t является постоянным на плоскости, перпендикулярной оси Z . Поэтому такие волны называются плоскими.

Сферические волны. Если волна от точечного источника изотропна, то решение уравнения (2.12) необходимо искать в виде $\Phi = \Phi(r, t)$, где r — расстояние от точечного источника, принятого за начало координат. Учитывая, что в сферической системе координат (r, θ, ϕ)

$$\nabla^2 \Phi = \frac{1}{r} \frac{\partial^2}{\partial r^2} (r \Phi) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \Phi}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \Phi}{\partial \phi^2}, \quad (2.26)$$

а искомое решение не зависит от угловых переменных, уравнение (2.12) примет вид

$$\frac{\partial^2}{\partial r^2} (r \Phi) - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} (r \Phi) = 0, \quad (2.27)$$

т. е. имеет тот же вид, что и (2.13), но с $z \rightarrow r$, $\Phi \rightarrow r \Phi$. Поэтому вместо (2.21) имеем

$$r \Phi(r, t) = \Phi_1(r + ct) + \Phi_2(r - ct),$$

где Φ_1 и Φ_2 — по-прежнему произвольные функции своего аргумента. Следовательно, общее сферически симметричное решение уравнения (2.12) имеет вид

$$\boxed{\Phi(r, t) = \frac{\Phi_1(r + ct)}{r} + \frac{\Phi_2(r - ct)}{r}} \quad (2.28)$$

Физический смысл отдельных слагаемых, входящих в (2.26), выясняется так же, как и для (2.21). Второе слагаемое представляет собой волну, движущуюся в направлении увеличения значений r , т. е. от центра. Такая волна называется расходящейся. Первое слагаемое описывает волну, движущуюся в направлении уменьшения значений r , т. е. к центру. Такая волна называется сходящейся. Общее решение (2.28) является суперпозицией сходящейся и расходящейся волн.

Значение Φ в фиксированный момент времени на сфере постоянного радиуса является постоянным. Такие волны называются сферическими.

Плоские гармонические волны. Если Φ_1 и Φ_2 в (2.21) являются гармоническими функциями своего аргумента, то волна называется гармонической. Запишем для примера функцию Φ_2 в виде

$$\Phi_2(z - ct) = \Phi_2[-c(t - z/c)] = A \cos \omega(t - z/c),$$

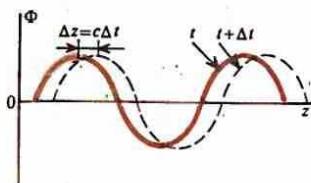
где A — постоянная, ω — частота гармонической функции. Волна, описываемая функцией

$$\boxed{\Phi(z, t) = A \cos \omega(t - z/c)}, \quad (2.29)$$

называется плоской гармонической волной. Она распространяется в направлении положительных значений оси Z . Постоянная A называется амплитудой волны, ω — ее частотой. Поскольку волна движется, ее называют также бегущей. Аналогичное (2.29) выражение для волны можно написать с использованием синуса. Общее выражение для бегущей волны, распространяющейся в положительном направлении оси Z , имеет вид

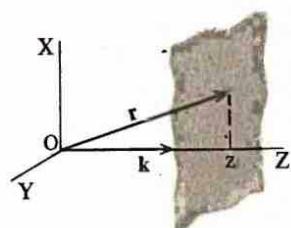
$$\boxed{\Phi(z, t) = A \cos \omega(t - z/c) + B \sin \omega(t - z/c)}. \quad (2.30)$$

Выбором подходящего начала отсчета времени бегущую плоскую волну всегда можно представить в виде (2.29) или аналогичным выражением с использованием синуса. Волна, распро-



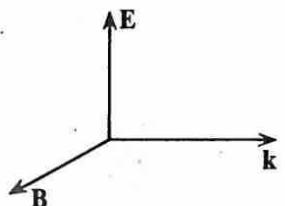
4

Гармоническая плоская волна в два последовательных промежутка времени



5

К записи плоской волны в векторных обозначениях



6

Взаимная ориентировка векторов плоской электромагнитной волны

строящаяся в направлении отрицательных значений оси Z , описывается очевидно, функцией

$$\Phi(z, t) = A \cos(\omega(t + z/c)) + B \sin(\omega(t + z/c)). \quad (2.31)$$

На рис. 4 изображена плоская гармоническая волна в два последовательных промежутка времени t и $t + \Delta t$. Для наглядности можно представить, что это волна на поверхности воды, а Φ характеризует отклонение частиц поверхности воды от горизонтальной плоскости. Конечно, при такой интерпретации c является не скоростью света, а скоростью распространения волны относительно воды. Положительные значения Φ соответствуют «горбам» на поверхности воды, а отрицательные — «впадинам». На рисунке изображена небольшая часть волны, включающая в себя два «горба» и одну «впадину». Если следить за какой-то фиксированной точкой среды, то будем наблюдать ее колебание по гармоническому закону с течением времени. Например, в точке $z=0$ этот закон описывается функцией

$$\Phi(0, t) = A \cos \omega t. \quad (2.32)$$

Если наблюдатель «сел» на какую-то точку волны, например на вершину «горба», и движется вместе с волной, то никаких изменений в картине он не видит — перед его взором будет находиться неизменная по времени совокупность «горбов» и «впадин», составляющих плоскую волну.

Аргумент гармонической функции в (2.29) называется фазой волны. Волна, у которой поверхностями постоянных фаз являются плоскости, называется плоской. Учитывая, что

$$\lambda = cT = 2\pi c/\omega, \quad (2.33)$$

запишем (2.29) в виде

$$\Phi(z, t) = A \cos(\omega t - \omega z/c) = A \cos(\omega t - kz), \quad (2.34)$$

где $k = \omega/c = 2\pi/\lambda$ — волновое число.

Волновой вектор. Чтобы освободиться от использования системы координат, запишем (2.34) с помощью векторных обозначений. Пусть вектор \mathbf{k} равен по модулю волновому числу и направлен параллельно оси Z в сторону положительных значений (рис. 5). Такой вектор называется волновым. Принимая во внимание, что $\mathbf{k} \cdot \mathbf{r} = kz$, запишем для произвольной точки, характеризуемой радиусом-вектором \mathbf{r} , вместо (2.34) выражение

$$\Phi(\mathbf{r}, t) = A \cos(\omega t - \mathbf{k} \cdot \mathbf{r}). \quad (2.35)$$

Эта формула не зависит от системы координат и характеризует плоскую волну, распространяющуюся в направлении вектора \mathbf{k} .

Аналогичное (2.35) выражение для волны можно также написать с использованием синуса:

$$\Phi'(\mathbf{r}, t) = A' \sin(\omega t - \mathbf{k} \cdot \mathbf{r}), \quad (2.36)$$

что опять-таки при подходящем начале отсчета времени может быть сведено к (2.35), поскольку $\sin(\alpha + \pi/2) = -\cos \alpha$.

Представление плоской волны в комплексной форме. Принимая во внимание формулу Эйлера $e^{ia} = \cos a + i \sin a,$

представим выражения (2.35) и (2.36) формулами

$$\Phi(\mathbf{r}, t) = A \operatorname{Re}[e^{-i(\omega t - \mathbf{k} \cdot \mathbf{r})}], \quad (2.38a)$$

$$\Phi'(\mathbf{r}, t) = -A' \operatorname{Im}[e^{-i(\omega t - \mathbf{k} \cdot \mathbf{r})}], \quad (2.38b)$$

где Re и Im — вещественная и мнимая части комплексного числа. В расчетах удобно пользоваться комплексным представлением плоской волны в виде

$$\Phi(\mathbf{r}, t) = A e^{-i(\omega t - \mathbf{k} \cdot \mathbf{r})}, \quad (2.39)$$

обозначая комплексную величину тем же символом, что и действительную. Это упрощает написание формул и не приводит к путанице. В тех случаях, когда путаница все же возможна, будем в явном виде указывать, о каком представлении идет речь.

Величина A в (2.39) может быть как действительной, так и комплексной или мнимой. Учитывая, что в общем случае

$$A = |A| e^{i\varphi}, \quad \operatorname{tg} \varphi = \operatorname{Im} A / \operatorname{Re} A, \quad (2.40)$$

запишем выражение (2.39) в виде

$$\Phi(\mathbf{r}, t) = |A| e^{-i(\omega t - \mathbf{k} \cdot \mathbf{r} - \varphi)}, \quad (2.41)$$

где $|A|$ — амплитуда плоской волны. Поэтому и в (2.39) $|A|$ — амплитуда плоской волны, $\omega t - \mathbf{k} \cdot \mathbf{r} - \varphi$ — фаза, где $\operatorname{tg} \varphi = \operatorname{Im} A / \operatorname{Re} A$.

Представление сферической волны в комплексной форме. Из способа записи плоской волны в комплексной форме очевидно, что расходящаяся и сходящаяся сферические волны согласно (2.28) могут быть представлены соответственно в виде

$$\Phi(r, t) = (A/r) e^{-i(\omega t - kr)}, \quad (2.42)$$

$$\Phi(r, t) = (A/r) e^{-i(\omega t + kr)} \quad (2.43)$$

Следует обратить внимание, что в показателе экспоненты этих формул стоит не скалярное произведение $\mathbf{k} \cdot \mathbf{r}$, а произведение положительных величин k и r .

Плоская электромагнитная волна. Для анализа структуры плоской электромагнитной волны удобно записать уравнения Максвелла в символьической форме с помощью оператора

$$\nabla = \mathbf{i}_x \frac{\partial}{\partial x} + \mathbf{i}_y \frac{\partial}{\partial y} + \mathbf{i}_z \frac{\partial}{\partial z}. \quad (2.44)$$

Принимая во внимание, что

$$\text{rot } \mathbf{A} = \nabla \times \mathbf{A},$$

$$\text{div } \mathbf{A} = \nabla \cdot \mathbf{A},$$

(2.45)

уравнения Максвелла (2.1) — (2.5) можно записать так:

$$\nabla \times \mathbf{B} = \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}, \quad (2.46) \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad (2.47)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (2.48) \quad \nabla \cdot \mathbf{E} = 0. \quad (2.49)$$

Будем искать решение этих уравнений в виде

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 e^{-i(\omega t - \mathbf{k} \cdot \mathbf{r})} \quad (2.50) \quad \mathbf{B}(\mathbf{r}, t) = \mathbf{B}_0 e^{-i(\omega t - \mathbf{k} \cdot \mathbf{r})}, \quad (2.51)$$

где \mathbf{E}_0 и \mathbf{B}_0 — постоянные векторы, не зависящие от координат и времени. Компоненты этих векторов могут быть комплексными.

Подставляя выражения (2.50) и (2.51) в уравнения (2.46) — (2.49) и учитывая, что

$$\nabla e^{i\mathbf{k} \cdot \mathbf{r}} = i\mathbf{k} e^{i\mathbf{k} \cdot \mathbf{r}}, \quad \frac{\partial}{\partial t} e^{-i\omega t} = -i\omega e^{-i\omega t}, \quad (2.52)$$

получаем следующие соотношения:

$$-\mathbf{k} \times \mathbf{B} = \omega \mu_0 \epsilon_0 \mathbf{E}, \quad (2.53) \quad \mathbf{k} \times \mathbf{E} = \omega \mathbf{B}, \quad (2.54)$$

$$\mathbf{k} \cdot \mathbf{B} = 0, \quad (2.55) \quad \mathbf{k} \cdot \mathbf{E} = 0. \quad (2.56)$$

Из соотношений (2.55) и (2.56) следует, что векторы \mathbf{E} и \mathbf{B} плоской волны перпендикулярны вектору \mathbf{k} , т. е. направлению распространения. Это означает, что электромагнитная волна является поперечной. Соотношения (2.53) и (2.54) показывают, что векторы \mathbf{E} и \mathbf{B} взаимно перпендикулярны. Таким образом, \mathbf{E} , \mathbf{B} и \mathbf{k} составляют тройку взаимно перпендикулярных векторов (рис. 6).

Поперечность световых колебаний была открыта в 1817 г. Т. Юнгом (1773—1829). С помощью этого представления он объяснил отсутствие интерференции лучей света, поляризованных во взаимно перпендикулярных плоскостях, обнаруженное в 1816 г. экспериментально в совместной работе Д. Ф. Араго (1786—1853) и О. Ж. Френеля (1788—1827).

Взяв от обеих частей (2.54) модули $|\mathbf{k} \times \mathbf{E}| = \omega |\mathbf{B}|$ и учитывая, что $|\mathbf{k} \times \mathbf{E}| = |\mathbf{k}| |\mathbf{E}|, |\mathbf{k}| = k = \omega/c$, находим следующее соотношение между напряженностью электрического поля и магнитной индукцией плоской волны в вакууме:

$$E = cB. \quad (2.57)$$

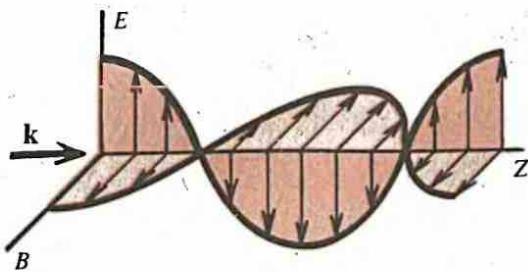
Поскольку \mathbf{k} , ω , μ_0 , ϵ_0 в (2.53) и (2.54) — вещественные величины, из (2.50) и (2.51) заключаем, что \mathbf{E} и \mathbf{B} в плоской волне изменяются в одинаковой фазе, т. е. одновременно достигают максимальных и нулевых значений (рис. 7).

Инвариантность плоской волны. Основными свойствами плоской волны являются взаимная перпендикулярность векторов \mathbf{E} и \mathbf{B} волны и соблюдение соотношения (2.57) между ними. Прежде всего возникает вопрос об инвариантности плоской волны, т. е. вопрос о том, что плоская в одной системе координат волна является плоской волной во всех других системах координат, движущихся относительно первой равномерно и прямолинейно. Ответ на этот вопрос основывается на инвариантах электромагнитного поля. Как показано в курсе электричества и магнетизма, при преобразованиях Лоренца инвариантными являются следующие величины, характеризующие электромагнитное поле:

$$I_1 = c^2 B^2 - E^2,$$

$$I_2 = \mathbf{E} \cdot \mathbf{B}.$$

(2.58)



Если в некоторой системе координат электромагнитное поле составляет плоскую волну, то $I_1 = 0$, $I_2 = 0$. Ввиду инвариантности этих величин относительно преобразований Лоренца они равны нулю также и во всех других системах координат. А это в соответствии с (2.58) означает, что электромагнитное поле в других системах координат также является плоской волной. Тем самым доказана инвариантность плоской волны.

Инвариантность фазы. Ясно, что утверждение о равенстве нулю векторов поля волны в некоторой пространственно-временной точке имеет объективный смысл независимо от того, в какой системе координат эта пространственно-временная точка рассматривается. Другими словами, векторы поля в этой пространственно-временной точке во всех системах координат равны нулю. А это означает, что фаза волны во всех системах координат одна и та же, что доказывает ее инвариантность. Инвариантность фазы следует из формул преобразования векторов поля. Написав формулы преобразования векторов электромагнитного поля плоской волны и подставив в них выражения вида (2.50) и (2.51), сразу заключаем, что для справедливости формул преобразования в любых пространственно-временных точках необходима инвариантность фаз волн.

Четырехмерный волновой вектор. Фазу можно представить в виде

$$\mathbf{k} \cdot \mathbf{r} - \omega t = k_1 x_1 + k_2 x_2 + k_3 x_3 + k_4 x_4, \quad (2.59)$$

где

$$x_1 = x, \quad x_2 = y, \quad x_3 = z, \quad x_4 = ic, \quad (2.60a)$$

$$k_1 = k_x, \quad k_2 = k_y, \quad k_3 = k_z, \quad k_4 = i\omega/c. \quad (2.60b)$$

Правая часть (2.59) имеет вид скалярного произведения. Поскольку совокупность величин (2.60a) составляет четырехмерный вектор, из инвариантности скалярного произведения в правой части (2.59) (инвариантность фазы) следует, что совокупность величин (2.60b) также составляет четырехмерный вектор. Это позволяет по известным формулам теории относительности преобразовать эти величины от одной системы координат к другой.

Формулы преобразования частоты и направления распространения плоской волны. Обозначив единичный вектор в направлении распространения плоской волны, представим волновой вектор в виде

$$\mathbf{k} = n\omega/c.$$

Поскольку совокупность (k_1, k_2, k_3, k_4) составляет четырехмерный вектор, получаем следующие формулы преобразования:

$$\omega n_x = \omega' (\beta + n'_x) / \sqrt{1 - \beta^2}, \quad \omega n_y = \omega' n'_y, \quad (2.61)$$

$$\omega n_z = \omega' n'_z, \quad \omega = \omega' (1 + \beta n'_x) / \sqrt{1 - \beta^2},$$

где величины со штрихами относятся к системе координат K' , которая движется относительно неподвижной системы координат K в направлении положительных значений оси X со скоростью v ($\beta = v/c$). Оси X и X' совмещены, оси Z' , Y' соответственно параллельны осям Z и Y .

Эффект Доплера. Согласно формулам (2.61), частота зависит от относительной скорости источника и наблюдателя. Из последнего из уравнений (2.61), воспользовавшись принципом относительности, получаем формулу

$$\omega = \omega' \sqrt{1 - \beta^2} / (1 - \beta n_x), \quad (2.62)$$

которая описывает эффект, открытый Доплером (1803—1853) в 1842 г. При малых скоростях ($\beta = v/c \ll 1$) формула (2.62) упрощается:

$$\omega = \omega' (1 + \beta n_x). \quad (2.63)$$

Если волна распространяется вдоль оси X , то $n_x = \pm 1$ и (2.63) принимает вид

$$\Delta\omega/\omega' = (\omega - \omega')/\omega' = \pm v/c. \quad (2.64)$$

Знак «+» относится к случаю, когда источник приближается к наблюдателю, а знак «—», — когда удаляется от наблюдателя. Этот случай изменения частоты называется **продольным** эффектом Доплера.

Если скорость источника перпендикулярна направлению наблюдения ($n_x = 0$), то проявляется **поперечный** эффект Доплера:

$$\omega = \omega' \sqrt{1 - \beta^2}. \quad (2.65)$$

Он обусловлен множителем $\sqrt{1 - \beta^2}$ в (2.62) и является чисто релятивистским эффектом, связанным с замедлением течения времени движущегося наблюдателя. Наличие этого эффекта было подтверждено экспериментально.

Пример 2.1. В направлении положительных значений оси X со скоростью v движется зеркало, отражающая плоская поверхность которого перпендикулярна оси X . На зеркало под углом $\theta_0 < \pi/2$ к положительному направлению оси X падает луч света. Найти направление отраженного от зеркала луча. Луч лежит в плоскости XY .

Проекции единичного вектора \mathbf{n} , направленного вдоль падающего на зеркало луча, равны $(\cos \theta_0, \sin \theta_0, 0)$. Если бы зеркало было неподвижным, то проекции единичного вектора, характеризующего отраженный луч, вследствие равенства угла отражения и угла падения были бы равны $(-\cos \theta_0, \sin \theta_0, 0)$. При отражении от движущегося зеркала явление усложняется.

Решим задачу, воспользовавшись принципом относительности: рассмотрим отражение в системе K' , где зеркало покоится и закон отражения известен. Из (2.61) и (2.62) получаем формулы преобразования от системы K к системе K' :

- Существование электромагнитных волн следует из уравнений Максвелла. Электромагнитная природа света установлена в результате совпадения свойств электромагнитных волн, описываемых уравнениями Максвелла, и свойств света.
- Волны классифицируются по форме поверхностей постоянной фазы (плоские, сферические, цилиндрические).
- Простейшими и наиболее важными электромагнитными волнами являются гармонические, но есть и другие.
- Представление волн в комплексной форме очень удобно для проведения расчетов.

$$26 \quad \omega' = \omega_0(1 - \beta n_x) / \sqrt{1 - \beta^2} = \omega_0(1 - \beta \cos \theta_0) / \sqrt{1 - \beta^2},$$

$$\text{---} \quad n'_x = (n_x - \beta) / (1 - \beta n_x) = (\cos \theta_0 - \beta) / (1 - \beta \cos \theta_0), \quad (2.66)$$

$$\text{---} \quad n'_y = n_y \sqrt{1 - \beta^2} / (1 - \beta n_x) = \sin \theta_0 \sqrt{1 - \beta^2} / (1 - \beta \cos \theta_0),$$

где ω_0 — частота света в системе K до отражения от движущегося зеркала. При отражении от неподвижного зеркала в системе K' частота не изменяется, а угол отражения равен углу падения. Следовательно, для отраженного луча на основании (2.66) можно написать:

$$\omega'_{\text{отр}} = \omega_0(1 - \beta \cos \theta_0) / \sqrt{1 - \beta^2}, \quad n'_{x \text{ отр}} = -(\cos \theta_0 - \beta) / (1 - \beta \cos \theta_0),$$

$$n'_{y \text{ отр}} = \sin \theta_0 \sqrt{1 - \beta^2} / (1 - \beta \cos \theta_0). \quad (2.67)$$

Теперь по формулам (2.61) можно вернуться в систему K :

$$\omega_{\text{отр}} = \omega_{\text{отр}} \frac{1 + \beta n'_x}{\sqrt{1 - \beta^2}} = \omega_0 \frac{1 - 2\beta \cos \theta_0 + \beta^2}{1 - \beta^2},$$

$$\cos \theta_{\text{отр}} = n_{x \text{ отр}} = \frac{\beta + n'_{x \text{ отр}}}{1 + \beta n'_{x \text{ отр}}} = -\frac{\cos \theta_0 - 2\beta + \beta^2 \cos \theta_0}{1 - 2\beta \cos \theta_0 + \beta^2}, \quad (2.68)$$

$$\sin \theta_{\text{отр}} = n_{y \text{ отр}} = \frac{n'_{y \text{ отр}} \sqrt{1 - \beta^2}}{1 + \beta n'_{x \text{ отр}}} = \frac{\sin \theta_0 (1 - \beta^2)}{1 - 2\beta \cos \theta_0 + \beta^2}.$$

В нерелятивистском случае, когда $\beta \ll 1$, эти равенства с точностью до величины порядка β можно представить так:

$$\omega_{\text{отр}} \approx \omega_0(1 - 2\beta \cos \theta_0), \quad \cos \theta_{\text{отр}} \approx -\cos \theta_0 + 2\beta \sin^2 \theta_0,$$

$$\sin \theta_{\text{отр}} \approx \sin \theta_0 (1 + 2\beta \cos \theta_0). \quad (2.69)$$

Видно, что луч от движущегося зеркала отражается под большим углом, чем от неподвижного.

§ 3 | Плотность потока энергии и импульса электромагнитных волн. Давление света

Общие формулы для плотности потока энергии и импульса электромагнитного поля детализируются применительно к оптическому диапазону.

Плотность потока энергии. Плотность потока энергии электромагнитного поля определяется вектором Пойнтинга

$$\boxed{\mathbf{S} = \mathbf{E} \times \mathbf{H}}, \quad (3.1)$$

модуль которого в случае плоской волны может быть представлен в виде

$$S = |\mathbf{S}| = |\mathbf{E} \times \mathbf{H}| = |\mathbf{E}| |\mathbf{H}| = EB / \mu_0 = [1 / (\mu_0 c)] E^2, \quad (3.2)$$

где учтено соотношение (2.57). Принимая во внимание, что $1/\mu_0 = \epsilon_0 c^2$, запишем соотношение (3.2) в форме

$$\boxed{S = c \epsilon_0 E^2}. \quad (3.3)$$

В (3.1) — (3.3) входят мгновенные значения величин. Однако векторы электромагнитной волны светового диапазона колеблются с частотами порядка 10^{15} Гц, поэтому нельзя следить за изменением величин по времени. Можно наблюдать и измерять лишь средние значения ве-

личин по очень большому числу периодов колебаний. Поэтому от мгновенных величин необходимо перейти к средним.

Учитывая, что $E = E_0 \cos \omega t$, где E_0 — амплитуда напряженности электрического поля волны, находим среднюю по времени плотность потока энергии:

$$\langle S \rangle_t = c \epsilon_0 E_0^2 \langle \cos^2 \omega t \rangle_t = c \cdot \frac{1}{2} \epsilon_0 E_0^2 \quad (3.4)$$

(индекс t у угловых скобок означает, что усреднение производится по времени; в большинстве случаев этот индекс не будет выписываться, поскольку и без обозначения бывает ясно, об усреднении по какой переменной идет речь).

Распределение плотности потока по сечению пучка. Обычно в эксперименте используют пучки света конечного поперечного сечения, по которому плотность потока распределена неравномерно. Мощность потока энергии в пучке по определению равна

$$P = \int |S| d\sigma, \quad (3.5)$$

где σ — площадь поперечного сечения пучка, перпендикулярного направлению распространения света. Поскольку поперечные размеры пучка бывают обычно порядка 1 мм — 1 см, т. е. $r_0 \approx 10^3 \div 10^4$ мкм, а длина волны λ имеет порядок 1 мкм, заключаем, что $\lambda/r_0 \approx 10^{-3} \div 10^{-4}$. Следовательно, дифракционные эффекты (см. § 33) в пучках малы и для расчета плотности потока энергии можно использовать формулы для плоских электромагнитных волн.

Гауссов пучок. Чаще всего пучок имеет круговое сечение, распределение плотности энергии по которому аксиально симметричное и гауссово. Такой пучок называется гауссовым. Амплитуды плоских электромагнитных волн, составляющих пучок, распределены по закону,

$$E_0(r) = A_0 e^{-r^2/(2r_0^2)}, \quad (3.6)$$

где A_0 — постоянная, равная амплитуде волн в центре пучка ($r=0$); r — расстояние от центра. Средняя плотность потока энергии равна

$$S = \frac{1}{2} \epsilon_0 A_0^2 e^{-r^2/r_0^2} = S_0 e^{-r^2/r_0^2}, \quad (3.7)$$

где $S_0 = \epsilon_0 A_0^2 / 2$ — средняя плотность потока энергии в центре пучка. Для упрощения написания угловые скобки, обозначающие средние величины, здесь не выписываются.

Мощность потока энергии в пучке в соответствии с (3.5) равна

$$P = S_0 \int_0^\infty e^{-r^2/r_0^2} \cdot 2\pi r dr = S_0 \pi r_0^2. \quad (3.8)$$

Фактически плотность энергии в гауссовом пучке распределена по экспоненциальному закону (3.7). На расстоянии r_0 она убывает в $e=2,7$ раза. По обычной договоренности об обращении с экспоненциально убывающими величинами можно сказать, что радиус пучка равен r_0 . В соответствии с (3.8) для оценочных расчетов полагают, что вся энергия пучка сосредоточена в пределах кругового сечения радиусом r_0 , а плотность по сечению постоянна и равна плотности в центре пучка. Гауссово распределение плотности потока и эффективное распределение показаны на рис. 8 соответственно сплошной и штрихованной линией. Такой подход обычно обобщается и на другие, не гауссовые пучки. Мощность потока энергии пучка представляется в виде

$$P = S_{\text{зф}} \pi r_{\text{зф}}^2, \quad (3.9)$$

где $S_{\text{зф}}$ — эффективная плотность потока энергии в пучке, $r_{\text{зф}}$ — его эффективный радиус. Если импульс света имеет энергию W , то его эффективная длительность $\tau_{\text{зф}}$ связана с мощностью потока энергии в импульсе соотношением

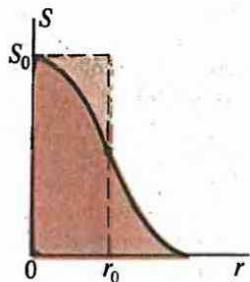
$$W = P \tau_{\text{зф}}. \quad (3.10)$$

В настоящее время с помощью лазеров получены громадные плотности потока энергии, порядка $S \sim 10^{20} \text{ Вт}/\text{м}^2$. Это означает [см. (3.3)], что напряженности электрического поля в волне имеют порядок $E \approx 10^9 \text{ В/м}$, т. е. достигают значений, характерных для внутриатомных полей. Длительность импульсов $\tau = 10^{-14} \text{ с}$ и менее.

Плотность импульса электромагнитной волны. Электромагнитная волна обладает не только энергией, но и импульсом. В теории электричества и магнетизма было показано, что плотность импульса G электромагнитной волны связана с плотностью потока энергии S в ней соотношением

$$G = S/c^2.$$

(3.11)



Гауссово (сплошная линия) и эффективное (пунктирная линия) распределения плотности потока энергии в пучке

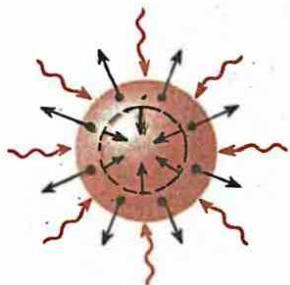


Схема осуществления лазерного термояда

Давление света. При поглощении или отражении света телом последнему, по закону сохранения, сообщается некоторый импульс, равный разности импульсов пучка света до и после поглощения или отражения. В результате на тело действует соответствующая сила и возникает световое давление. Идея о существовании давления света была высказана еще Кеплером для объяснения отклонения хвостов комет от Солнца. Однако эта идея вызывала много споров, поскольку доказать существование светового давления долго не удавалось. Сторонники волновой теории долго считали, что световое давление не существует, а отсутствие экспериментального доказательства наличия светового давления выдвигали в качестве аргумента против корпускулярной теории, согласно которой световое давление, безусловно, должно существовать. Однако электромагнитная теория света предсказывает существование светового давления. Если электромагнитная волна падает нормально на плоскую поверхность твердого тела и полностью поглощается им, то в 1 с на 1 м² поверхности тела в соответствии с (3.11) передается импульс G и, следовательно, световое давление равно

$$p = c G = S/c.$$

(3.12)

При полном отражении поверхности тела передается импульс, в два раза больший, и, следовательно, в два раза большее давление. Можно подсчитать давление при частичном поглощении потока энергии. Если плотность потока поглощаемой энергии равна $S_{\text{пог}} = aS$, то по закону сохранения энергии плотность потока отражаемой энергии $S_{\text{отр}} = (1 - a)S$ и, следовательно, давление может быть представлено в виде $p = aS/c + 2(1 - a)S/c = (2 - a)S/c$. При падении потока энергии под

○ Если мгновенные значения величин нельзя измерить, что означает утверждение об их существовании?

Перечислите основные параметры гауссова пучка.

Каковы причины возникновения радиометрических сил и как они были устранены в опытах Лебедева?

Перечислите основные факторы, приводящие к сжатию мишени в лазерном термояде. Можно ли сказать, что все эти факторы сводятся к световому давлению?

По каким формулам преобразуются энергия и импульс цуга плоских волн при переходе от одной инерциальной системы отсчета к другой?

углом к нормали необходимо принять во внимание лишь нормальную составляющую плотности потока энергии. Световое давление в обычных условиях очень мало. Например, на земной орбите плотность потока солнечного излучения составляет примерно $S = 1400 \text{ Вт}/\text{м}^2$, поэтому $p = 1400/(3 \cdot 10^8) \text{ Па} = 0,5 \cdot 10^{-5} \text{ Па}$. Учитывая, что атмосферное давление составляет около 10^5 Па , находим, что световое давление Земли примерно в 10^{10} раз меньше атмосферного. Поэтому обнаружить его было очень трудно. Первый шаг в правильном направлении сделал У. Крукс (1832—1919). Он использовал крутильные весы, лепестки которых облучались светом. По закручиванию нити можно было судить о действующей на лепестки силе. Однако из-за радиометрического эффекта ему не удалось измерить световое давление. Это впервые удалось сделать П. Н. Лебедеву (1866—1912) в 1900 г. Радиометрические силы были им уменьшены в результате создания глубокого вакуума в сосуде, в котором находились крутильные весы. Благодаря этому сила светового давления стала играть доминирующую роль в закручивании нити крутильных весов и была измерена.

Экспериментальная проверка достоверности исключения радиометрических сил состоит в следующем. Если поверхность, на которую направляется луч, отражающая, то давление света на нее в два раза больше, чем если бы она была полностью поглощающей, а радиометрическое действие меньше. Если же поверхность полностью поглощает излучение, то световое давление на нее в два раза меньше, чем давление на полностью отражающую поверхность, а радиометрические силы — больше. П. Н. Лебедев действительно наблюдал этот эффект увеличения в два раза светового давления при освещении отражающего крыльышка крутильных весов по сравнению с освещением поглощающего (черного) крыльышка, что и доказывает исключение радиометрического действия.

Действие светового давления на малые частицы. Световое давление пропорционально площади, т. е. квадрату линейных размеров частицы, а масса частицы пропорциональна объему, т. е. кубу линейных размеров. Это означает, что при фиксированной плотности сила, возникающая за счет светового давления и приходящаяся на единицу массы, изменяется обратно пропорционально линейным размерам частицы, т. е. растет с уменьшением ее размеров. Пусть, например, у шарообразной частицы $\rho = 10^3 \text{ кг}/\text{м}^3$, $r = 10^{-5} \text{ м}$. Масса частицы $m = 4/3\pi r^3 \rho = 4 \cdot 10^{-12} \text{ кг}$. При потоке $S = 1400 \text{ Вт}/\text{м}^2$ солнечного излучения на орбите Земли световое давление $p = S/c = 0,5 \text{ мкПа}$ и поэтому сила, действующая на полностью поглощающую излучение частицу, $F = p\pi r^2 = 1,5 \cdot 10^{-15} \text{ Н}$. Она сообщает частице ускорение $a = F/m = 0,4 \cdot 10^{-3} \text{ м}/\text{с}^2$. Это большое ускорение. Ускорение, сообщаемое Солнцем частице на орбите Земли, равно $a_c = v^2/R = [(30 \times 10^3)^2 / (15 \cdot 10^{10})] \text{ м}/\text{с}^2 = 0,6 \cdot 10^{-2} \text{ м}/\text{с}^2$, т. е. примерно в $(a_c/a) \approx 15$ раз большее ускорения вследствие светового давления. Для частицы с меньшими в 15 раз линейными размерами сила, обусловленная световым давлением примерно уравновесит силу притяжения Солнца. Поскольку как плотность потока излучения, так и сила тяжести убывают с расстоянием по одному и тому же закону (обратно пропорционально квадрату расстояния от Солнца), сила притяжения Солнца и сила за счет светового давления равны друг другу по абсолютному значению для всех расстояний, и поэтому такая частица в поле тяготения Солнца движется так, как будто это поле отсутствует. Наиболее наглядным проявлением светового давления является ориентировка хвостов комет при их прохождении вблизи Солнца («вблизи» — в астрономическом масштабе расстояний).

Лазерный термояд. Этим термином обозначается идея осуществления управляемого термоядерного синтеза с помощью лазерного излучения. Для этого необходимо добиться сближения легких ядер (дейтерия, трития) на столь малое расстояние, чтобы между ними произошла ядерная реакция слияния. Но для сближения необходимо преодолеть кулоновскую силу отталкивания положительно заряженных ядер. Для этого им необходимо сообщить достаточно большую кинетическую энергию (нагреть вещество).

Идея лазерного термояда состоит в следующем. Предположим, что на небольшой шарик, содержащий ядерное горючее (ядра дейтерия, трития), одновременно со всех сторон направля-

ются мощные короткие импульсы лазерного излучения (рис. 9, волнистые стрелки). Энергия и импульс этого излучения передаются веществу в поверхностном слое шарика, он нагревается и приобретает направленное движение к центру. Энергия и импульс этого вещества передаются внутренней части шарика, которая сильно сжимается и нагревается. Частицы верхних слоев шарика приобретают скорость от центра шарика (как бы «испаряются» с поверхности). Таким образом, вещество внутренней области шарика очень сильно сжимается, что сопровождается огромным повышением температуры, а вещество внешних слоев шарика разлетается с очень большими скоростями (рис. 9; прямые стрелки). Если плотность и температура сжатого вещества шарика достигнут необходимых для осуществления ядерной реакции значений, то произойдет небольшой термоядерный взрыв, вроде взрыва маленькой водородной бомбы. Выделенная при этом энергия превращается в основном в кинетическую энергию продуктов ядерной реакции слияния, которая, в принципе, может быть преобразована в другие формы энергии и целесообразно использована.

Для осуществления термоядерной реакции вещество шарика необходимо сжать в несколько сотен раз, а температуру поднять на много десятков миллионов градусов. Это можно сделать лишь с помощью очень мощных лазерных импульсов излучения.

Однако схема осуществления ядерного термояда проста лишь в принципе. Ее техническое осуществление чрезвычайно сложно и требует глубоких научных исследований. Оказалось, что процесс сжатия и нагревания вещества шариков очень сложен и зависит от многих факторов. Далее задача усложняется тем, что осуществление миниатюрного термоядерного взрыва шариков еще не означает овладения управляемой термоядерной реакцией. Необходимо, чтобы общий энергетический баланс работы установки был положителен, т. е. чтобы получаемая в результате работы установки энергия в форме, пригодной для использования, была больше энергии, необходимой для функционирования установки.

Преобразование амплитуды и нормали плоской электромагнитной волны. Пусть волновой вектор плоской волны лежит в плоскости $X Y$. Для напряженностей электрического поля и индукции магнитного поля можем написать:

$$E_x = -E_0 n_y e^{i\varphi}, \quad E_y = E_0 n_x e^{i\varphi}, \quad B_z = (E_0/c) e^{i\varphi}, \quad (3.13)$$

где φ — фаза, E_0 — амплитуда плоской волны. Фаза волны инвариантна относительно преобразований Лоренца, поэтому по формулам преобразования векторов поля находим:

$$E_0 = E'_0 (1 + \beta n_x) / \sqrt{1 - \beta^2}, \quad (a) \quad n_x E_0 = E'_0 (n'_x + \beta) / \sqrt{1 - \beta^2}, \quad (b)$$

$$n_y E_0 = n'_y E'_0. \quad (b) \quad (13.14)$$

Как обычно, величины со штрихами относятся к системе координат K' , движущейся со скоростью v в положительном направлении оси X — системы K . Оси X' и X систем координат совмещены.

Почленным делением второго и третьего уравнения (3.14) на первое находим формулы для преобразования проекций единичного вектора нормали плоской волны:

$$n_x = (n'_x + \beta) / (1 + \beta n'_x), \quad (a) \quad n_y = n'_y \sqrt{1 - \beta^2} / (1 + \beta n'_x). \quad (b) \quad (3.15)$$

Разделив почленно уравнение (3.14б) на первое из уравнений (2.61), получим

$$E_0 / \omega = E'_0 / \omega'. \quad (3.16)$$

Это соотношение между амплитудой и частотой плоской волны инвариантно относительно преобразований Лоренца.

Энергия цуга плоских волн. Заключенная в объеме V энергия в соответствии с (3.4) равна

$$W = \frac{1}{2} \epsilon_0 E_0^2 V. \quad (3.17)$$

Объем, занимаемый цугом волн, движется со скоростью c . Поэтому с этим объемом нельзя связать систему координат и говорить о величине этого объема в состоянии покоя. Однако найти формулу преобразования объема цуга волн можно. Введем некоторый вспомогательный объем V_0 , который в системе координат K' движется со скоростью u'_x , а в системе K — со скоростью u_x . Учитывая релятивистское сокращение длин, запишем

$$V' = V_0 \sqrt{1 - u'^2/c^2}, \quad V = V_0 \sqrt{1 - u_x^2/c^2} \quad (3.18)$$

и, следовательно,

$$\frac{V'}{V} = \frac{\sqrt{1 - u'^2/c^2}}{\sqrt{1 - u_x^2/c^2}}. \quad (3.19)$$

Отсюда с учетом формулы сложения скоростей теории относительности находим

$$V' = \frac{V \sqrt{1 - \beta^2}}{1 - vu_x/c^2}. \quad (3.20)$$

Заменим u_x в (3.20) по формуле

$$u_x = un_x, \quad (3.21)$$

где n_x — проекция единичного вектора, характеризующего направление скорости \mathbf{u} , на ось X . Тогда

$$V' = \frac{V \sqrt{1 - \beta^2}}{1 - vun_x/c^2}. \quad (3.22)$$

При $v \rightarrow c$ получим закон преобразования объемов цуга волн, движущихся со скоростью света:

$$V' = \frac{V \sqrt{1 - \beta^2}}{1 - \beta n_x}. \quad (3.23)$$

Формулу (3.14а) на основании принципа относительности можно переписать в виде

$$E'_0 = E_0 (1 - \beta n_x) / \sqrt{1 - \beta^2}. \quad (3.24)$$

Из сравнения (3.23), (3.24), (3.17) и (2.62) получаем

$$\boxed{W'/\omega' = W/\omega,} \quad \boxed{V'\omega' = V\omega.} \quad (3.25)$$

Это означает, что энергия цуга плоских волн прямо пропорциональна частоте. Этот результат используется в квантовой теории света.

Импульс цуга плоских волн. Он равен

$$\mathbf{G} = (1/c^2) \int \mathbf{E} \times \mathbf{H} dV = \mathbf{n} (\epsilon_0/c) \int E^2 dV = \mathbf{n} W/c, \quad (3.26)$$

где \mathbf{n} — единичный вектор в направлении распространения волны. Все величины в (3.26) имеют мгновенное значение (т. е. усреднение по периоду колебаний не предусмотрено). При выводе (3.26) приняты во внимание соотношения

$$EH/c^2 = \epsilon_0 E^2/c = w/c. \quad (3.27)$$

С учетом (3.25) равенство (3.26) можно записать в виде

$$\mathbf{G} = \mathbf{n}\omega, \quad W = \omega c, \quad (3.28)$$

где $a = W/(\omega c)$ — на основании (3.25) — постоянная величина. Из (2.60б) заключаем, что совокупность величин

$$(\omega n_x, \omega n_y, \omega n_z, -i\omega) \quad (3.29)$$

образует четырехмерный вектор. Поэтому из (3.28) следует, что совокупность величин

$$(G_x, G_y, G_z, -iW/c), \quad (3.30)$$

характеризующая полный импульс и полную энергию пуга плоских волн, также образует четырехмерный вектор.

Пример 3.1. Рубиновый лазер излучает импульсы линейно поляризованного света с гауссовым распределением амплитуд по круговому сечению пучка. Энергия в импульсе, продолжительность импульса и эффективный радиус пучка равны соответственно 1 Дж, 100 мкс, 5 мм. Считая мгновенную мощность в течение импульса излучения постоянной, определить ее значение, плотность потока энергии на оси пучка, амплитуду напряженности электрического поля и амплитуды индукции и напряженности магнитного поля на расстоянии 2 мм от оси пучка.

Ввиду постоянства мгновенной мощности средняя мощность в импульсе равна ее мгновенному значению $P = W/\tau = 10^4$ Вт [см. (3.10)]. Плотность потока энергии на оси пучка равна [см. (3.8)] $S_0 = P/(\pi r_0^2) = 1,27 \cdot 10^8$ Вт/м². Тогда амплитуда напряженности электрического поля на оси [см. (3.4)] равна $E_0(0) = [2S_0/(c\varepsilon_0)]^{1/2} = 3,1 \cdot 10^5$ В/м. Напряженность электрического поля на расстоянии 2 мм от оси [см. (3.6)]

$$E_0(2 \cdot 10^{-3} \text{ м}) = E_0(0) \exp[-r^2/(2r_0^2)] = 2,86 \cdot 10^5 \text{ В/м}.$$

Отсюда для амплитуд индукции и напряженности магнитного поля получаем значения $0,95 \cdot 10^{-3}$ Тл и $7,56 \cdot 10^2$ А/м.

● В экспериментах мгновенные значения величин в пространственных точках не могут быть измерены. Всегда измеряются средние значения величин по некоторому интервалу времени и области пространства.

Экспериментально в настоящее время средние значения величин на интервалах времени, меньших периода светового колебания, не измеряются. Всегда измеряются средние по многим периодам. Измеренное в эксперименте среднее значение напряженности электрического поля световой волны всегда равно нулю. Простейшие величины, среднее значение которых отлично от нуля, пропорциональны квадрату напряженности поля. Важнейшей из них является объемная плотность энергии, связанная простыми соотношениями с плотностью потока энергии и объемной плотностью импульса волны.

Световое давление является проявлением объемной плотности импульса у волны и закона сохранения импульса при взаимодействии волны с веществом.

Прослеживается переход от принципа суперпозиции для напряженностей электромагнитного поля волны к принципу суперпозиции для электромагнитной волны.

Суперпозиция векторов поля волны. Напряженность электрического поля и магнитная индукция равны соответственно сумме напряженностей и магнитных индукций всех полей в данной точке независимо от их происхождения. В частности, эти поля могут принадлежать плоским электромагнитным волнам всевозможных частот и направлений распространения. Однако полученная в результате сложения полей совокупность электрического и магнитного полей, вообще говоря, не составляет бегущую электромагнитную волну, даже если слагаемые поля принадлежат к бегущим электромагнитным волнам.

Суперпозиция бегущих плоских монохроматических электромагнитных волн. Пусть имеются две волны с одинаковым волновым вектором \mathbf{k} и одинаковой частотой ω , поля которых описываются векторами $\mathbf{E}_1, \mathbf{B}_1$ и $\mathbf{E}_2, \mathbf{B}_2$. Соотношения (2.53)–(2.56) для них имеют следующий вид:

$$\begin{aligned} -\mathbf{k} \times \mathbf{B}_1 &= \omega \mu_0 \epsilon_0 \mathbf{E}_1, & -\mathbf{k} \times \mathbf{B}_2 &= \omega \mu_0 \epsilon_0 \mathbf{E}_2, \\ \mathbf{k} \times \mathbf{E}_1 &= \omega \mathbf{B}_1, & \mathbf{k} \times \mathbf{E}_2 &= \omega \mathbf{B}_2, \\ \mathbf{k} \cdot \mathbf{B}_1 &= 0, & \mathbf{k} \cdot \mathbf{B}_2 &= 0, \\ \mathbf{k} \cdot \mathbf{E}_1 &= 0, & \mathbf{k} \cdot \mathbf{E}_2 &= 0, \end{aligned} \quad (4.1)$$

а формула (2.57) дает равенства

$$\mathbf{E}_1 = c \mathbf{B}_1, \quad \mathbf{E}_2 = c \mathbf{B}_2. \quad (4.2)$$

Складывая почленно левые и правые части равенств (4.1) и (4.2) и обозначая

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2, \quad \mathbf{B} = \mathbf{B}_1 + \mathbf{B}_2, \quad (4.3)$$

приходим к соотношениям

$$\begin{aligned} -\mathbf{k} \times \mathbf{B} &= \omega \mu_0 \epsilon_0 \mathbf{E}, & \mathbf{k} \times \mathbf{E} &= \omega \mathbf{B}, & \mathbf{k} \cdot \mathbf{B} &= 0, & \mathbf{k} \cdot \mathbf{E} &= 0, \\ \mathbf{E} &= c \mathbf{B}, \end{aligned} \quad (4.4)$$

которые доказывают, что поле, описываемое векторами \mathbf{E}, \mathbf{B} , представляет плоскую монохроматическую бегущую волну с волновым вектором \mathbf{k} и частотой ω . Тем самым доказано, что две плоские монохроматические бегущие волны с одинаковой частотой, распространяющиеся в одном и том же направлении, в результате сложения дают плоскую монохроматическую волну той же частоты, распространяющуюся в том же направлении. Если слагаемые волны имеют разные частоты или различные направления распространения, то в результате их сложения не будет получена плоская монохроматическая бегущая волна.

Биения. Рассмотрим случай сложения двух монохроматических волн, имеющих частоты ω_1 и ω_2 и распространяющихся в одном направлении. Векторы \mathbf{E} в этих волнах коллинеарны. Для определенности ось Z совместим с направлением распространения волн, а X совместим с направлением вектора \mathbf{E} волны, т. е. предположим, что $\mathbf{E} = (E_x, 0, 0)$, $\mathbf{B} = (0, B_y, 0)$. Чтобы не загромождать изложения, будем следить за вектором \mathbf{E} , поскольку поведение вектора \mathbf{B} определяется по вектору \mathbf{E} с помощью соотношений между векторами плоской волны. Для про-

34 стоты допустим, что амплитуды напряженностей электрического поля слагаемых волн одинаковы:

$$E_{1x} = E_0 \cos(\omega_1 t - k_1 z), \quad E_{2x} = E_0 \cos(\omega_2 t - k_2 z). \quad (4.5)$$

В соответствии с принципом суперпозиции имеем

$$E = E_x = E_{1x} + E_{2x} = E_0 \cos(\omega_1 t - k_1 z) + E_0 \cos(\omega_2 t - k_2 z) =$$

$$= 2E_0 \cos\left(\frac{\omega_1 - \omega_2}{2} t - \frac{k_1 - k_2}{2} z\right) \\ \cos\left(\frac{\omega_1 + \omega_2}{2} t - \frac{k_1 + k_2}{2} z\right), \quad (4.6)$$

где использована формула сложения косинусов. Учитывая, что $k_1 = \omega_1/c$, $k_2 = \omega_2/c$, представим (4.6) в виде

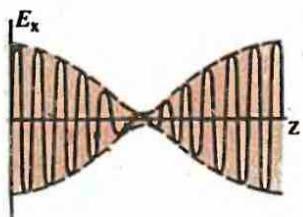
$$E = 2E_0 \cos[(\omega_1 - \omega_2)(t - z/c)/2] \\ \cos[(\omega_1 + \omega_2)(t - z/c)/2], \quad (4.7)$$

который показывает, что результирующее электромагнитное поле распространяется без затухания в направлении положительных значений оси Z со скоростью c (об этом свидетельствует наличие комбинации $t - z/c$ в аргументе функции). В этом смысле речь идет о бегущей волне, однако не монохроматической. Учитывая, что в пределах оптического диапазона всегда соблюдается соотношение $|\omega_1 - \omega_2| \ll \omega_1 + \omega_2$, можно дать следующую наглядную интерпретацию такой волны: сомножитель $2E_0 \cos[(\omega_1 - \omega_2)(t - z/c)/2]$ представляет изменяющуюся амплитуду гармонической волны с частотой $(\omega_1 + \omega_2)/2$ [последний сомножитель в (4.7)]. Таким образом, волна в некоторый фиксированный момент времени имеет вид, изображенный на рис. 10. Сплошной линией показаны колебания частоты $(\omega_1 + \omega_2)/2$, а пунктирной — огибающая амплитуды колебаний, изменяющихся от максимального значения $2E_0$ до нуля. Если амплитуды E_{01} и E_{02} полей слагаемых волн не равны друг другу, то амплитуда суммарной волны изменяется от $E_{01} + E_{02}$ до $|E_{01} - E_{02}|$.

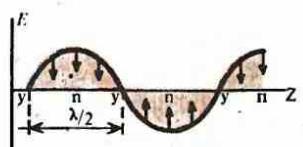
Гармонические колебания с медленно изменяющейся амплитудой называются биениями. Понятие «медленно изменяющаяся амплитуда» определяется относительно основного гармонического колебания: амплитуда мало меняется в течение многих периодов основного гармонического колебания (рис. 10).

Поскольку по определению амплитудой колебаний называется положительная величина максимального отклонения от положения равновесия, из (4.7) заключаем, что частота биений равна

$$\Omega = |\omega_1 - \omega_2|. \quad (4.8)$$



10
Биения



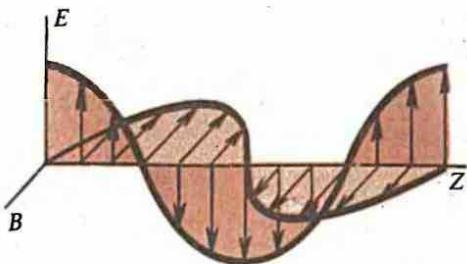
11
Стоячая волна напряженности E

○ Какое свойство электромагнитных волн обеспечивает соблюдение для них принципа суперпозиции как прямого следствия справедливости принципа суперпозиции для напряженности электрического поля и индукции магнитного поля?

В каких областях пространства движется энергия в стоячей электромагнитной волне? Рассмотрите вопрос о движении энергии при наличии биений.

12

Стоячая электромагнитная волна



Стоящие волны. Рассмотрим суперпозицию двух монохроматических волн одинаковой частоты, распространяющихся навстречу друг другу. Будем считать, что векторы напряженности электрического поля в этих волнах коллинеарны и колеблются с одинаковой амплитудой. По-прежнему ось Z располагаем по направлению распространения волны, а ось X — коллинеарно направлению векторов E волн. Имеем

$$E_1 = E_{1x}(z, t) = E_0 \cos(\omega t - kz), \quad (4.9)$$

$$E_2 = E_{2x}(z, t) = E_0 \cos(\omega t + kz + \delta), \quad (4.10)$$

где положительный знак при kz в (4.10) учитывает, что волна с E_2 распространяется в направлении отрицательных значений оси Z ; δ — сдвиг фаз.

В результате суперпозиции этих двух бегущих волн образуется волна, напряженность поля которой равна

$$E = E_1 + E_2 = 2E_0 \cos(kz + \delta/2) \cos(\omega t + \delta/2). \quad (4.11)$$

Видно, что эта волна не является бегущей, поскольку отсутствует характерный для нее множитель $t \pm z/c$. Сомножитель $2E_0 \cos(kz + \delta/2)$ с точностью до знака можно рассматривать как амплитуду колебаний напряженности поля в точке z . Она изменяется от точки к точке по гармоническому закону. Напряженность во всех точках изменяется с одинаковой частотой в одной и той же фазе [сомножитель $\cos(\omega t + \delta/2)$]. Такая волна называется стоячей (рис. 11). В точках оси Z , удовлетворяющих условию $\cos(kz + \delta/2) = 0$, напряженность E тождественно равна нулю. Эти точки называются узлами (буква «у» на рисунке). Точки, для которых $\cos(kz + \delta/2) = \pm 1$, имеют максимальную амплитуду колебаний напряженности. Они называются пучностями (буква «п» на рисунке). Расстояние Δz между узлами (или между пучностями) находится, очевидно, из условия $k\Delta z = \pi$ и равно $\Delta z = \lambda/2$, т. е. половине длины бегущей волны. Стоящая волна при не равных нулю напряженностях электрического поля показана на рис. 11. Видно, что мгновенный снимок стоячей волны совпадает со снимком бегущей волны. Однако между бегущей и стоячей волнами имеется глубокое различие. Колебания напряженности во всех точках стоячей волны в некоторый момент времени находятся в одной и той же фазе, а колебания напряженности электрического поля в различных точках бегущей волны не совпадают по фазе. В частности, у стоячей волны имеется такой момент времени, когда напряженность E во всех точках оси Z равна нулю [при $\cos(\omega t + \delta/2) = 0$].

Магнитная индукция B полей волн также складывается в соответствии с принципом суперпозиций полей. Так как векторы E , B , k плоской электромагнитной волны образуют правовинтовую тройку векторов (рис. 6), векторы B_1 и B_2 волн, распространяющихся в положительном и отрицательном направлениях оси Z , на основании (4.9) и (4.10) имеют соответственно вид

$$B_1 = B_{1y}(z, t) = E_{1x}/c = (E_0/c) \cos(\omega t - kz), \quad (4.12)$$

$$B_2 = B_{2y}(z, t) = -E_{2x}/c = -(E_0/c) \cos(\omega t + kz + \delta) \quad (4.13)$$

(знак минус в правой части учитывает, что правовинтовая тройка векторов составляется вектором \mathbf{E} в положительном направлении оси X , вектором \mathbf{B} в отрицательном направлении оси Y , вектором \mathbf{k} в отрицательном направлении оси Z). Магнитная индукция поля результирующей волны равна

$$B = B_y = B_1 + B_2 = (2E_0/c) \sin(kz + \delta/2) \sin(\omega t + \delta/2), \quad (4.14)$$

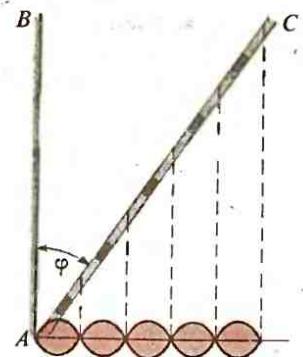
т. е. выражается формулой, аналогичной (4.11), но с заменой косинуса на синус. Это означает, что вектор \mathbf{B} также образует стоячую волну, узлы которой совпадают с пучностями стоячей волны \mathbf{E} . Векторы \mathbf{E} и \mathbf{B} расположены во взаимно перпендикулярных плоскостях. Из сравнения (4.11) и (4.14) также видно, что по времени колебания электрического и магнитного полей стоячей электромагнитной волны отличаются по фазе на четверть периода колебаний. Это означает, что если, например, напряженность электрического поля стоячей волны достигает максимума, то магнитная индукция в это время равна нулю; если же напряженность электрического поля стоячей волны достигает половины максимальной величины, то магнитная индукция также достигает половины максимальной величины, а отличаются они тем, что, например, напряженность электрического поля находится в фазе роста ее абсолютного значения, а индукция — в фазе уменьшения. Стоячая электромагнитная волна показана на рис. 12 для такого момента времени, когда ни напряженность электрического поля, ни индукция не достигают своих максимальных абсолютных значений.

Преобразование энергии в стоячей электромагнитной волне. Плотность потока энергии волн описывается вектором Пойнгтига (3.1). Следовательно, поток энергии отсутствует в точках, где либо \mathbf{E} , либо \mathbf{B} равны нулю. Это означает, что поток энергии в стоячей электромагнитной волне отсутствует через узлы и пучности в волне, поскольку пучность напряженности электрического поля совпадает с узлом индукции магнитного поля и наоборот. Поэтому с течением времени энергия движется между соседними узлами и пучностями, превращаясь из энергии магнитного поля в энергию электрического поля и обратно. С помощью (4.11) и (4.14), пользуясь формулой для объемной плотности энергии электромагнитного поля

$$w = \frac{1}{2} (\mathbf{E} \cdot \mathbf{D} + \mathbf{H} \cdot \mathbf{B}), \quad (4.15)$$

можно проверить, что энергия стоячей волны, заключенная между соседними узлами и пучностями, с течением времени сохраняет постоянное значение, что и доказывает утверждение о взаимопревращении энергии электрического и магнитного полей стоячей волны.

Экспериментальное доказательство электромагнитной природы света: Стоячие волны света получаются в результате сложения волн падающего и отраженного от некоторой поверх-



Опыт Винера по доказательству электромагнитной природы света

Из справедливости принципа суперпозиции для электрического и магнитного полей не следует автоматически его справедливость для волн. Принцип суперпозиции для электромагнитных волн требует специального доказательства.

Электромагнитная волна и объемная плотность ее энергии не могут существовать без движения. Движение является способом их существования. Поэтому даже в стоячей электромагнитной волне как напряженность поля волны, так и объемная плотность энергии волны постоянно изменяются. В опытах Винера важно не только экспериментальное прямое доказательство электромагнитной природы света, но и установление того факта, что на фотозенулюсию действует напряженность электрического поля волны.

ности пучка света. Законы отражения будут рассмотрены несколько позднее (см. § 16, 20). Здесь лишь заметим, что для изменения направления движения волны при отражении необходимо изменить взаимную ориентацию векторов E и B волны так, чтобы векторное произведение $E \cdot B$ было направлено в сторону движения отраженного луча, в то время как до отражения оно направлено по падающему лучу. Поэтому при отражении либо вектор E падающей волны, либо вектор B должны изменить свое направление на обратное.

Если стоячие волны света создать в фотоэмульсии, то максимальные почернения должны наблюдаться в пучностях стоячих волн. Если фотографическое действие электрического и магнитного векторов одинаково, то и почернение в их пучностях одинаково, а если различно, то и почернения различны. Таким образом, исследование стоячих волн в фотоэмульсиях позволяет проверить электромагнитную теорию света и одновременно установить, какой из векторов волны и в какой степени обладает фотографическим действием. Такие опыты впервые были поставлены в 1890 г. О. Винером, получившим систему стоячих волн в воздухе отражением от металлического зеркала. Поскольку расстояние между пучностями очень мало (порядка 0,3 мкм), Винер исследовал почернение в тонком светочувствительном слое AC (порядка $\lambda/20$), расположенному под малым углом ϕ к поверхности AB зеркала (рис. 13). Если расстояния между пучностями по нормали к поверхности зеркала равны $\lambda/2$, то в наклонном тонком светочувствительном слое эти расстояния равны $d = \lambda/(2 \sin \phi)$, т. е. при малых углах могут быть сделаны достаточно большими и их можно измерить.

Опыты Винера доказали, что фотографическое действие обусловлено электрической напряженностью E поля, а не индукцией B . Предсказания электромагнитной теории света в опытах Винера полностью подтвердились.

§ 5 Поляризация электромагнитных волн

Показывается, что электромагнитные волны имеют две независимые поляризации, и анализируются различные представления волни.

Поляризация. Для продольных волн все направления, перпендикулярные линии распространения волн, эквивалентны. Для поперечных волн они не эквивалентны. Электромагнитные волны являются поперечными, и их свойства зависят от ориентировки векторов E и B , характеризуемой понятием поляризации.

Линейная поляризация. Если в процессе распространения волн вектор E лежит в одной и той же плоскости, параллельной направлению распространения волн, то волны называются линейно поляризованными (см. рис. 7).

Плоскостью поляризации называется плоскость, в которой лежат векторы B и k (вектор B , а не E !). Однако это понятие сейчас употребляется редко. Чаще говорят о плоскости колебаний вектора E .

Суперпозиция линейно поляризованных волн. Рассмотрим суперпозицию двух линейно поляризованных волн одной и той же частоты, распространяющихся в одном направлении. Для определенности будем считать, что колебания E первой волны лежат в плоскости XZ , а второй — в плоскости YZ (рис. 14). Тогда можно записать:

$$E_{1x}(z, t) = E_{10} \sin(\omega t - kz), \quad E_{1y} = E_{1z} = 0, \quad (5.1)$$

$$E_{2y}(z, t) = E_{20} \sin(\omega t - kz + \delta), \quad E_{2x} = E_{2z} = 0, \quad (5.2)$$

где δ — сдвиг фаз между колебаниями.

Исследуем напряженность $E = E_1 + E_2$ электрического поля суммарной волны в плоскости, перпендикулярной направлению распространения волн при фиксированном значении z . С течением времени конец вектора E описывает в плоскости XY некоторую замкнутую кривую.

Найдем уравнение этой кривой. Перепишем (5.2) в виде

$$E_y = E_{20} \sin(\omega t - kz) \cos \delta + E_{20} \cos(\omega t - kz) \sin \delta$$

и с помощью (5.1) исключим из этого равенства

$$\sin(\omega t - kz) \text{ и } \cos(\omega t - kz)$$

$$E_y = E_{20}(E_x/E_{10}) \cos \delta + E_{20} \sin \delta \sqrt{1 - E_x^2/E_{10}^2}. \quad (5.3)$$

Напомним, что амплитуды E_{10} и E_{20} предполагаются положительными числами. Перенося первое слагаемое правой части (5.3) на левую сторону, возводя обе части в квадрат, деля их на E_{20}^2 , после перегруппировки членов приводим уравнение к виду

$$\frac{E_x^2}{E_{10}^2} + \frac{E_y^2}{E_{20}^2} - 2 \frac{E_x}{E_{10}} \frac{E_y}{E_{20}} \cos \delta = \sin^2 \delta. \quad (5.4)$$

Рассмотрим различные случаи, описываемые этим уравнением.

Эллиптическая и круговая поляризации. Если $\cos \delta = 0$, $\sin \delta = \pm 1$, то уравнение (5.4) принимает вид

$$E_x^2/E_{10}^2 + E_y^2/E_{20}^2 = 1. \quad (5.5)$$

При $E_{10} \neq E_{20}$ это выражение является уравнением эллипса с центром в начале координат и осями, направленными вдоль осей системы координат. Полуоси эллипса (рис. 15) равны E_{10} (по оси X) и E_{20} (по оси Y). Условие $\cos \delta = 0$ соблюдается при

$$\delta = \pi/2 + n\pi \quad (n = 0, \pm 1, \pm 2, \dots). \quad (5.6)$$

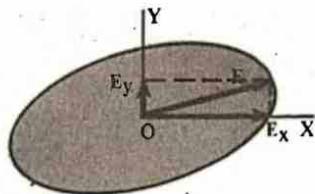
Тогда уравнения (5.1) и (5.2) принимают следующий вид ($z=0$):

$$E_x = E_{10} \sin \omega t, \quad (5.7)$$

$$E_y = E_{20} \sin(\omega t + \pi/2 + n\pi) = (-1)^{n+1} E_{20} \cos \omega t. \quad (5.8)$$

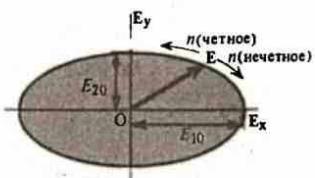
Видно, что конец вектора \mathbf{E} вращается по часовой стрелке при нечетном n и против часовой стрелки при четном n . В первом случае говорят о правой эллиптически поляризованной волне, а во втором — о левой эллиптически поляризованной волне. Отметим, что наблюдение за вращением вектора \mathbf{E} ведется со стороны, в которую движется волна (ось Z на рис. 15 направлена к нам).

При $E_{10} = E_{20}$ эллипс становится окружностью. Соответствующая волна называется **поляризованной по кругу** или **волной с циркулярной поляризацией**. Как и в случае волн с эллиптической поляризацией, могут быть волны с левой и правой циркулярной поляризацией.



14

Суперпозиция линейно поляризованных плоских волн, приводящая к эллиптически поляризованной волне



15

Эллиптическая поляризация в частном случае, когда главные оси эллипса коллинеарны направлениям колебаний E

○ Зависит ли ориентация главных осей эллипса при эллиптической поляризации от разности фаз линейно поляризованных во взаимно перпендикулярных направлениях волн, в результате суперпозиции которых образовалась эллиптически поляризованная волна?

Можно ли получить циркулярно поляризованную волну при разных амплитудах слагаемых линейно поляризованных волн?

Можно ли получить эллиптически поляризованную волну при сложении линейно поляризованных волн с одинаковыми амплитудами? Чем определяется направление циркулярной поляризации?

При $\cos \delta \neq 0$ уравнение (5.4) также описывает эллипс, однако его главные оси не совпадают с осями координат. Как видно из уравнений (5.1) и (5.2), максимальные и минимальные значения составляющих E_x и E_y равны $\pm E_{20}$ и $\pm E_{10}$, поэтому эллипс вписан в прямоугольник со сторонами $2E_{10}$ и $2E_{20}$ с центром в начале координат (рис. 16). Ориентация эллипса и его параметры зависят от δ . В частности, следует обратить внимание, что для $\cos \delta \neq 0$ получается эллиптически поляризованная волна даже при $E_{10} = E_{20}$. Направление вращения суммарного вектора \mathbf{E} определяется значением δ .

Изменение вектора напряженности в пространстве при эллиптической и круговой поляризациях. Для фиксированного момента времени множество точек, образуемых концом вектора \mathbf{E} , лежит на винтообразной линии. На рис. 17 показано пространственное изменение вектора \mathbf{E} при круговой поляризации волны. Представить себе эллиптически поляризованную волну при фиксированном t можно так: на поверхности прямого эллиптического цилиндра проведена винтовая линия, начало вектора \mathbf{E} находится в точках оси цилиндра, конец — на винтовой линии, причем сам вектор везде перпендикулярен оси.

Вырожденный случай эллиптической поляризации. При $\cos \delta = \pm 1$, $\sin \delta = 0$ соотношение (5.4) превращается в равенство

$$\frac{E_x^2}{E_{10}^2} + \frac{E_y^2}{E_{20}^2} \mp 2 \frac{E_x}{E_{10}} \frac{E_y}{E_{20}} = \left(\frac{E_x}{E_{10}} \mp \frac{E_y}{E_{20}} \right)^2 = 0,$$

которое при $\cos \delta = +1$ и $\cos \delta = -1$ описывает прямые

$$E_x/E_{10} - E_y/E_{20} = 0, \quad (5.9a) \quad E_x/E_{10} + E_y/E_{20} = 0. \quad (5.9b)$$

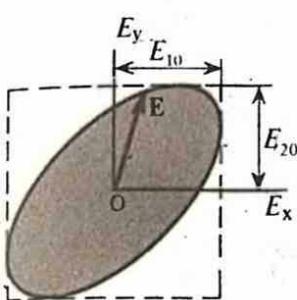
Конец суммарного вектора \mathbf{E} движется по соответствующей прямой (рис. 18, а, б). Получается линейно поляризованная волна, которая является предельным случаем эллиптической поляризованной волны (см. рис. 16) при равенстве нулю одной из полуосей эллипса. При $\cos \delta = +1$ линейное колебание суммарного вектора \mathbf{E} происходит в первом и третьем, а при $\cos \delta = -1$ во втором и четвертом квадрантах.

Число независимых поляризаций. Изложенное показывает, что электромагнитная волна с любой поляризацией может быть представлена в виде суперпозиции двух линейно поляризованных волн, плоскости колебаний электрического вектора которых взаимно перпендикулярны. Поэтому можно сказать, что электромагнитные волны обладают двумя независимыми состояниями поляризации.

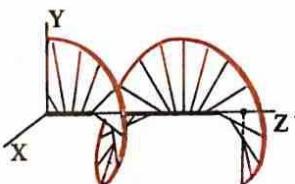
Линейно поляризованная волна как суперпозиция волн с круговой поляризацией. Рассмотрим суперпозицию волн с левой и правой круговыми поляризациями. Пусть при некотором фиксированном значении z векторы напряженности \mathbf{E}_1 и \mathbf{E}_2 волны заданы соотношениями (рис. 19)

$$E_{1x} = E_0 \cos \omega t, \quad E_{1y} = E_0 \sin \omega t, \quad (5.10)$$

$$E_{2x} = E_0 \cos \omega t, \quad E_{2y} = -E_0 \sin \omega t. \quad (5.11)$$



16
Эллиптическая поляризация в общем случае.



17
Изменение вектора \mathbf{E} в пространстве в фиксированный момент времени при круговой поляризации волны

Первая волна обладает левой, а вторая — правой круговой поляризацией. Амплитуды волн одинаковы. В результате суперпозиции волн получается волна с проекциями напряженности

$$E_x = E_{1x} + E_{2x} = 2E_0 \cos \omega t, \quad (5.12)$$

$$E_y = E_{1y} + E_{2y} = 0, \quad (5.13)$$

т. е. линейно поляризованная волна. В данном случае линия колебаний вектора \mathbf{E} совпадает с осью X . Если между колебаниями [см. (5.10) и (5.11)] имеется сдвиг фаз, то линия колебаний суммарной напряженности образует угол с осью X , определяемый сдвигом фаз слагаемых волн.

Пример 5.1. Определить характеристики волны, получаемой в результате суперпозиции двух волн с одинаковой амплитудой, поляризованных по правому и левому кругу, если в начальный момент разность фаз волн равна δ .

Напряженности левополяризованной и правополяризованной волн записываются в виде

$$E_{1x} = E_0 \cos \omega t, \quad E_{1y} = E_0 \sin \omega t, \quad (5.14)$$

$$E_{2x} = E_0 \cos(-\omega t + \delta), \quad E_{2y} = E_0 \sin(-\omega t + \delta).$$

Удобно произвести сложение описываемых формулами (5.14) колебаний, представив их в виде комплексных векторов

$$E_1 = E_{1x} + iE_{1y} = E_0 e^{i\omega t}, \quad E_2 = E_{2x} + iE_{2y} = E_0 e^{-i(\omega t - \delta)}, \quad (5.15)$$

результат сложения которых дается формулой

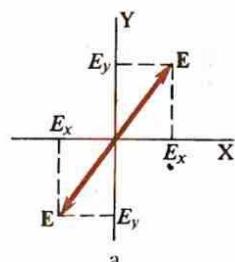
$$E = E_1 + E_2 = E_0 e^{i\delta/2} [e^{i(\omega t - \delta/2)} + e^{i(\omega t - \delta/2)}] = \\ = 2E_0 e^{i\delta/2} \cos(\omega t - \delta/2), \quad (5.16)$$

поэтому

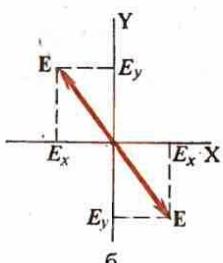
$$E_x = 2E_0 \cos(\delta/2) \cos(\omega t - \delta/2), \quad (5.17)$$

$$E_y = 2E_0 \sin(\delta/2) \cos(\omega t - \delta/2).$$

Из (5.17) следует, что $E_y/E_x = \tan(\delta/2)$, т. е. результирующая волна является линейно поляризованной, причем линия колебаний вектора \mathbf{E} образует с осью X угол $\delta/2$.



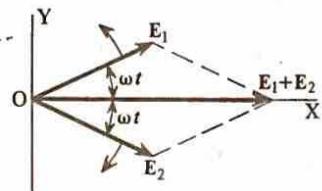
a



b

b18

18
Вырожденный случай сложения двух линейно поляризованных волн



19
Образование линейно поляризованной волны в результате суперпозиции циркулярно поляризованных волн

- Из принципа суперпозиции следует, что любую линейно поляризованную волну можно представить в виде суммы двух волн с взаимно перпендикулярными плоскостями линейной поляризации. Если возникает разность фаз колебаний между этими волнами, то результирующая волна является либо эллиптически поляризованной, либо циркулярно поляризованной, в зависимости от разности фаз.

§ 6 Усреднения

Показывается, как результат усреднения зависит от интервала усреднения и даются правила вычислений при комплексной форме представления волн.

Операция усреднения. В физических теориях пользуются величинами, относящимися к точке пространства и моменту времени. Однако в физическом эксперименте измеряют средние значения величин по некоторому объему и промежутку времени. Например, мгновенным значением величины называется

се предельное значение при стремлении промежутка времени устремления к нулю. Однако в эксперименте это предельное значение достигнуто быть не может. Результатом измерения всегда является среднее по некоторому малому промежутку времени. Все сказанное об усреднении по времени относится также и к усреднениям по объемам.

Среднее значение физической величины $f(x, y, z, t)$ определяется формулой

$$\langle f(x, y, z, t) \rangle_{\Delta V, \Delta t} = \frac{1}{\Delta V \Delta t} \int_V f(x, y, z, t) dx dy dz dt, \quad (6.1)$$

где ΔV и Δt — соответственно объем и интервал времени усреднения.

Результат усреднения зависит от размеров области усреднения. Масштабы изменения f , меньшие области усреднения, не фиксируются в усредненных величинах. Например, нельзя изучить ход изменения напряженности электрического поля во времени в гармоническом колебании, если за промежуток времени усреднения взять период колебаний.

Если в области усреднения ΔV в любой момент промежутка времени Δt усреднения величина f изменяется незначительно и этим изменением можно пренебречь, то в (6.1) при интегрировании по $dx dy dz$ величину f можно считать постоянной и выполнить усреднение по (x, y, z) :

$$\langle f(x, y, z, t) \rangle_{\Delta V, \Delta t} = \frac{1}{\Delta t} \int_{\Delta t} dt \frac{1}{\Delta V} \int_{\Delta V} f(x, y, z, t) dx dy dz = \frac{1}{\Delta t} \int_{\Delta t} f(x, y, z, t) dt, \quad (6.2)$$

где $\frac{1}{\Delta V} \int_{\Delta V} dx dy dz = 1$. Таким образом, все операции усреднения в этом случае сводятся к усреднению по времени:

$$\langle f(t) \rangle_{\Delta t} = \frac{1}{\Delta t} \int_{\Delta t} f dt, \quad (6.3)$$

где пространственные переменные для упрощения написания формулы не выписаны, поскольку операция усреднения от них не зависит.

Усреднение гармонических функций. По определению,

$$\langle e^{i\omega t} \rangle_{\Delta t} = \frac{1}{\Delta t} \int_{-\Delta t/2}^{\Delta t/2} e^{i\omega t} dt = \frac{1}{i\omega \Delta t} \left[e^{i\omega t} \right]_{-\Delta t/2}^{\Delta t/2} = \frac{\sin(\omega \Delta t/2)}{(\omega \Delta t/2)} e^{i\omega t}, \quad (6.4)$$

где $e^{ia} - e^{-ia} = 2i \sin a$. Разделяя в (6.4) действительные и мнимые части, находим:

$$\langle \cos \omega t \rangle_{\Delta t} = \frac{\sin(\omega \Delta t/2)}{(\omega \Delta t/2)} \cos \omega t, \quad (6.5)$$

$$\langle \sin \omega t \rangle_{\Delta t} = \frac{\sin(\omega \Delta t/2)}{(\omega \Delta t/2)} \sin \omega t. \quad (6.6)$$

Результатом усреднения гармонической функции является гармоническая функция с той же частотой, но с амплитудой, умноженной на

$$A(\xi) = \sin \xi / \xi \quad (\xi = \omega \Delta t / 2) \quad (6.7)$$

(рис. 20). Видно, что амплитуда усредненной гармонической функции быстро убывает с увеличением Δt . При $\xi = \omega \Delta t / 2 = \pi$; $\Delta t = 2\pi/\omega = T$ (T — период гармонических колебаний) амплитуда обращается в нуль. В частности, в оптическом диапазоне электромагнитных колебаний ($T \approx 10^{-15}$ с) приборов, которые регистрируют напряженность электрического или индукцию

42

магнитного поля за промежуток времени, меньший 10^{-15} с, в настоящее время не существует. Поэтому невозможно экспериментально зафиксировать ход изменения электрической напряженности и магнитной индукции световой волны. В оптике изучаются лишь усредненные по многим периодам колебаний значения физических величин.

Усреднение квадратов гармонических функций. В этом случае имеем

$$\begin{aligned} <\cos^2 \omega t>_{\Delta t} &= <\frac{1}{2}(1 + \cos 2\omega t)>_{\Delta t} = \\ &= \frac{1}{2}\left[1 + \frac{\sin(\omega \Delta t)}{(\omega \Delta t)} \cos 2\omega t\right], \end{aligned} \quad (6.8)$$

$$\begin{aligned} <\sin^2 \omega t>_{\Delta t} &= <\frac{1}{2}(1 - \cos 2\omega t)>_{\Delta t} = \\ &= \frac{1}{2}\left[1 - \frac{\sin(\omega \Delta t)}{(\omega \Delta t)} \cos 2\omega t\right], \end{aligned} \quad (6.9)$$

где использованы соотношения для двойного угла и формулы (6.5) и (6.6). График средней величины квадрата гармонической функции приведен на рис. 21. Видно, что среднее значение колеблется около $1/2$ с амплитудой, определяемой характерным множителем $A(\xi)$. При увеличении интервала времени усреднения амплитуда колебаний уменьшается и среднее значение квадрата гармонической функции стремится к постоянному значению $1/2$. Эта ситуация в оптическом диапазоне всегда осуществляется.

Линейность операции усреднения. Из определения (6.1) следует, что

$$<\mu_1 f_1 + \mu_2 f_2> = <\mu_1 f_1> + <\mu_2 f_2> = \mu_1 <f_1> + \mu_2 <f_2>, \quad (6.10)$$

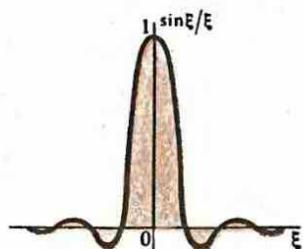
где μ_1, μ_2 — постоянные. Равенство (6.10) показывает, что операция усреднения является линейной.

Вычисления с комплексными скалярными величинами. Для упрощения вычислений колебания и волны обычно представляются в комплексной форме. Вычисление средних по времени приходится производить от произведений действительных или мнимых частей комплексных величин. Однако произведение действительных частей двух комплексных величин не равно действительной части их произведения, т. е.

$$\operatorname{Re} A \operatorname{Re} B \neq \operatorname{Re}(AB),$$

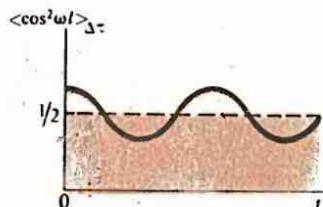
и поэтому нахождение среднего от произведения действительных частей двух комплексных величин не сводится к вычислению среднего от действительной части их произведения. Но вычисление все же значительно упрощается, поскольку зависимость от времени рассматриваемых величин имеет вид

$$\begin{aligned} A(x, y, z, t) &= A_0(x, y, z) e^{-i\omega t}, \\ B(x, y, z, t) &= B_0(x, y, z) e^{-i\omega t}. \end{aligned} \quad (6.12)$$



20

График амплитуды усредненной гармонической функции



21

График колебаний усредненного квадрата гармонической функции

● **Результат усреднения зависит от интервала усреднения. С помощью средних значений нельзя изучать изменения, которые происходят на интервалах, не имеющих интервала усреднения.**

● **Докажите, что среднее значение производной равно производной от среднего значения.**

Почему среднее значение произведения вещественных частей двух комплексных чисел не равно вещественной части среднего значения их произведения?

Учитывая равенства:

$$\operatorname{Re} A = \frac{1}{2}(A + A^*), \quad \operatorname{Re} B = \frac{1}{2}(B + B^*), \quad (6.13)$$

получаем

$$\operatorname{Re} A \operatorname{Re} B = \operatorname{Re}(A_0 e^{-i\omega t}) \operatorname{Re}(B_0 e^{-i\omega t}) = \frac{1}{2} \operatorname{Re}(A_0 B_0^* + A_0 B_0 e^{-2i\omega t}), \quad (6.14)$$

где принято во внимание соотношение

$$\operatorname{Re}(AB^*) = \operatorname{Re}(A^*B). \quad (6.15)$$

В результате усреднения (6.14) по времени находим

$$\langle \operatorname{Re} A \operatorname{Re} B \rangle = \frac{1}{2} \operatorname{Re}(A_0 B_0^*) = \frac{1}{2} \operatorname{Re}(AB^*) = \frac{1}{2} \operatorname{Re}(A^*B), \quad (6.16)$$

где A и B имеют вид (6.12).

Вычисления с комплексными векторными величинами. Если векторы поля волны представлены в комплексном виде, т. е.

$$\mathbf{E} = E_0 e^{-i(\omega t - \mathbf{k} \cdot \mathbf{r})}, \quad (6.17)$$

$$\mathbf{H} = H_0 e^{-i(\omega t - \mathbf{k} \cdot \mathbf{r})}, \quad (6.18)$$

то плотность потока энергии определяется формулой

$$\mathbf{S} = \operatorname{Re} \mathbf{E} \times \operatorname{Re} \mathbf{H}. \quad (6.19)$$

Представим не зависящие от времени части векторов в виде

$$E_0 e^{i\mathbf{k} \cdot \mathbf{r}} = E_R + iE_I,$$

$$H_0 e^{i\mathbf{k} \cdot \mathbf{r}} = H_R + iH_I, \quad (6.20)$$

где индексами R и I отмечены действительные и мнимые части величин, стоящих в левой части равенства. Имеем

$$\mathbf{E} = (E_R + iE_I) e^{-i\omega t}, \quad \mathbf{H} = (H_R + iH_I) e^{-i\omega t} \quad (6.21)$$

и, следовательно,

$$\operatorname{Re} \mathbf{E} = \operatorname{Re}[(E_R + iE_I)(\cos \omega t - i \sin \omega t)] = E_R \cos \omega t + E_I \sin \omega t; \quad (a) \quad (6.22)$$

$$\operatorname{Re} \mathbf{H} = H_R \cos \omega t + H_I \sin \omega t. \quad (b)$$

Подставляя эти выражения в (6.19), находим

$$\mathbf{S} = (E_R \times H_R) \cos^2 \omega t + (E_I \times H_I) \sin^2 \omega t + [(E_R \times H_I) + (E_I \times H_R)] \sin \omega t \cos \omega t. \quad (6.23)$$

При усреднении по времени третье слагаемое в правой части равенства исчезает, а средние от квадратов синуса и косинуса равны $\frac{1}{2}$. Поэтому

$$\langle \mathbf{S} \rangle = \frac{1}{2} (E_R \times H_R) + \frac{1}{2} (E_I \times H_I). \quad (6.24)$$

Соотношению (6.24) можно придать более простую форму, если использовать комплексно-сопряженные величины. Из (6.21) следует, что

$$\mathbf{H}^* = (H_R - iH_I) e^{i\omega t}. \quad (6.25)$$

44 Вычислим векторное произведение:

$$\mathbf{E} \times \mathbf{H}^* = (\mathbf{E}_R + i\mathbf{E}_I) e^{-i\omega t} \times (\mathbf{H}_R - i\mathbf{H}_I) e^{i\omega t} = (\mathbf{E}_R \times \mathbf{H}_R) + \mathbf{E}_I \times \mathbf{H}_I + \\ + i[(\mathbf{E}_I \times \mathbf{H}_R) - (\mathbf{E}_R \times \mathbf{H}_I)]. \quad (6.26)$$

Сравнивая (6.26) с (6.24), получаем

$$\langle \mathbf{S} \rangle = \frac{1}{2} \operatorname{Re}(\mathbf{E} \times \mathbf{H}^*). \quad (6.27)$$

Аналогично вычисляются средние объемные плотности энергии электрического и магнитного полей волны:

$$\langle w_{\text{эл}} \rangle = (\epsilon/4) \mathbf{E} \cdot \mathbf{E}^*, \quad (6.28)$$

$$\langle w_m \rangle = (\mu/4) \mathbf{H} \cdot \mathbf{H}^* = [1/(4\mu)] \mathbf{B} \cdot \mathbf{B}^*. \quad (6.29)$$

§ 7

Фотометрические понятия и величины

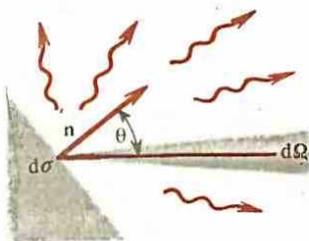
Дается способ пересчета энергетических величин в фотометрические и наоборот.

Энергетические и фотометрические величины. Физические приборы и человеческий глаз в оптическом диапазоне регистрируют средние значения измеряемых величин по большому числу периодов колебаний. Средние значения напряженности электрического поля и индукции магнитного поля равны нулю и не могут быть зафиксированы. Простейшими регистрируемыми величинами являются те, которые зависят от квадратов напряженности, т. е. энергетические величины (объемная плотность энергии излучения, плотность потока энергии излучения, мощность излучения и др., получаемые на их основе). Их измеряют с помощью физических приборов.

Следует заметить, что по своему действию электрическое поле волны незквивалентно магнитному полю. Например, известно, что почернение фотопластинки под влиянием света происходит в результате действия электрического поля волны. Однако учитывая, что объемные плотности энергии электрических и магнитных полей волны равны, всегда можно действие волны характеризовать энергетическими величинами.

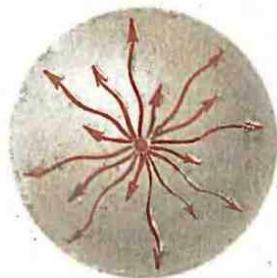
Человеческий глаз также регистрирует усредненные значения, т. е. он реагирует на энергетические характеристики излучения в видимом диапазоне или, другими словами, реагирует на энергетические характеристики света. Однако ощущение, вызываемое светом, зависит не только от энергетических характеристик света, но и от других обстоятельств, в первую очередь от длины волны света. Например, максимальной чувствительностью глаз обладает к зеленому свету с длиной волны 555 нм. К границам видимого диапазона чувствительность глаза уменьшается до нуля. Например, чтобы излучение с длиной волны 760 нм создало у человека такое же ощущение яркости, как излучение с длиной волны 555 нм, необходимо увеличить мощность излучения примерно в 20 000 раз.

Во многих случаях интерес представляют не сами энергетические характеристики света, а те субъективные ощущения, которые с ними связаны. Например, необходимо определить освещенность письменного стола, которая наиболее благоприятна для работы. С помощью энергетических характеристик света этого сделать нельзя, потому что одна и та же мощность излучения, направляемого на стол, вызывает совершенно различные ощущения освещенности стола при различных спектральных составах света. Для решения таких вопросов приходится пользоваться иными, отличными от энергетических величинами, называемыми фотометрическими. Энергетические и фотометрические величины взаимосвязаны.



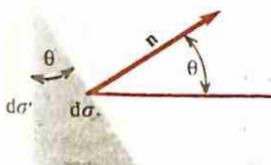
22

К определению величин, характеризующих излучение от элемента поверхности



23

Точечный источник излучения, равномерно испускаемого по всем направлениям



24

К определению энергетической яркости

Энергетические величины. Определения энергетических величин основываются на мощности излучения. Если в течение времени dt испускается энергия dW в форме излучения, то мощность излучения равна

$$P = dW/dt. \quad (7.1a)$$

Она распределяется по всевозможным длинам волн.

Спектральной плотностью мощности излучения является величина

$$P_\lambda = dP/d\lambda, \quad (7.1b)$$

где dP — мощность, приходящаяся на интервал длин волн $(\lambda, \lambda + d\lambda)$. Ясно, что

$$dP = P_\lambda d\lambda. \quad (7.1c)$$

Излучают поверхности материальных тел. Элементарным излучателем является элемент поверхности тела с площадью $d\sigma$ (рис. 22). В этом параграфе: $d\sigma$ — площадь элемента поверхности, поскольку буквой S обозначается плотность потока энергии (3.4); dP — мощность излучения, испускаемого элементарным источником.

Энергетическая сила излучения. Энергетической силой излучения dI элементарного источника называется отношение мощности dP излучения в элемент телесного угла $d\Omega$ к $d\Omega$:

$$dI = dP/d\Omega. \quad (7.2a)$$

Для спектральной плотности излучения эта формула принимает вид

$$dI_\lambda = dP_\lambda/d\Omega, \quad (7.2b)$$

где $dI_\lambda = d(dI)/d\lambda$ — спектральная плотность энергетической силы излучения, приходящейся на интервал длин волн $(\lambda, \lambda + d\lambda)$.

Ясно, что dI зависит, вообще говоря, от направления излучения, т. е. от ориентировки элемента телесного угла $d\Omega$ относительно элементарного излучателя. Если элементарным излучателем является элемент поверхности тела, то dI зависит, в частности, от угла θ между нормалью n к поверхности и направлением, в котором ориентирован элемент телесного угла (рис. 22), и также от аксиального угла, характеризующего вращение вокруг нормали как оси. Нормаль, от поверхности направленная в сторону испускания излучения. Кроме того, очевидно, что dI также пропорциональна площади $d\sigma$ элемента поверхности.

Энергетическая сила точечного источника излучения, равномерно испускаемого по всем направлениям (рис. 23), равна

$$\text{d}I = I_0 = \text{const.}$$

(7.3)

Из (7.2а) следует соотношение

$$P = \int_{\Omega=4\pi} dI = \int_{\Omega=4\pi} I_0 d\Omega = 4\pi I_0,$$

связывающее энергетическую силу точечного источника с полной мощностью его излучения.

Энергетическая яркость. Излучение с элемента поверхности $d\sigma$ испускается по всевозможным направлениям, характеризуемым углом θ между нормалью и к элементу поверхности и направлением распространения излучения (рис. 24). Проекция $d\sigma$ на поверхность, перпендикулярную направлению распространения излучения, равна

$$d\sigma' = d\sigma \cos \theta.$$

Энергетической яркостью поверхности в точке элемента поверхности $d\sigma$ называется отношение энергетической силы излучения dI с этого элемента поверхности к площади $d\sigma'$:

$$L = \frac{dI}{d\sigma'} = \frac{dI}{d\sigma \cos \theta} = \frac{dP}{d\Omega d\sigma \cos \theta}.$$

Для спектральной плотности эта формула имеет вид

$$L_\lambda = \frac{dI_\lambda}{d\sigma'} = \frac{dI_\lambda}{d\sigma \cos \theta} = \frac{dP_\lambda}{d\Omega d\sigma \cos \theta}.$$

Ясно, что энергетическая яркость зависит от направления испускания излучения и, вообще говоря, различна для разных точек поверхности.

Энергетическая светимость. Мощность излучения с элемента поверхности по всем направлениям, отнесенная к площади элемента, называется энергетической светимостью:

$$M = \frac{dP}{d\sigma} = \int_{\Omega=2\pi} \frac{dI d\Omega}{d\sigma} = \int_{\Omega=2\pi} L \cos \theta d\Omega,$$

где интегрирование распространяется по телесному углу 2π , включающему в себя все направления от элемента $d\sigma$ в сторону испускания излучения.

Спектральная плотность энергетической светимости определяется по формуле

$$M_\lambda = \frac{dP_\lambda}{d\sigma} = \int_{\Omega=2\pi} \frac{dI_\lambda d\Omega}{d\sigma} = \int_{\Omega=2\pi} L_\lambda \cos \theta d\Omega.$$

Если спектральная плотность энергетической яркости L_λ не зависит от направления (т. е. $L_\lambda = \text{const}$), то интеграл в (7.7б) можно вычислить. Направив ось Z сферической системы координат по нормали к элементу поверхности и обозначив φ аксиальный угол, запишем (7.7б) в виде

$$M_\lambda = L_\lambda \int_0^{2\pi} d\varphi \int_0^{\pi/2} \sin \theta \cos \theta d\theta.$$

(7.8)

Выполнив интегрирование, получим

$$M_\lambda = \pi L_\lambda.$$

(7.9 а)

Из (7.9 а) следует равенство $\int_0^\infty M_\lambda d\lambda = \pi \int_0^\infty L_\lambda d\lambda$, которое с учетом (7.6 а), (7.6 б), (7.7 а), (7.7 б) записывается в виде

$$M = \pi L,$$

(7.9 б)

где

$$M = \int_0^\infty M_\lambda d\lambda,$$

$$L = \int_0^\infty L_\lambda d\lambda$$

(7.9 в)

— энергетическая светимость и энергетическая яркость поверхности.

Энергетическая освещенность. Все предыдущие величины характеризовали процесс излучения. Теперь рассмотрим падение излучения на элемент поверхности. Это явление характеризуется величиной, называемой энергетической освещенностью. Она равна отношению мощности излучения dP , падающего на элемент поверхности, к площади элемента $d\sigma$:

$$E = dP/d\sigma.$$

(7.10 а)

При расчетах нормаль n к поверхности считается направленной в ту сторону поверхности $d\sigma$, откуда падает излучение (рис. 25).

Спектральная плотность энергетической освещенности дается формулой

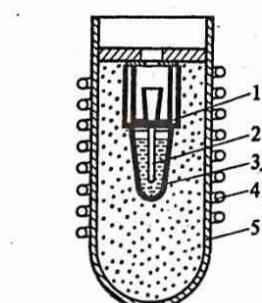
$$E_\lambda = dP_\lambda / d\sigma.$$

(7.10 б)

Фотометрические величины. Они определяются аналогично энергетическим, но исходя из силы света как основной величины. Единица силы света — кандела определяется с помощью черного излучателя, принятого в качестве основного эталона, работающего при температуре затвердевания платины. Этот эталон был утвержден в 1967 г. решением XIII Генеральной конференции по мерам и весам. Он состоит из закрытой снизу керамической трубы 2 диаметром до 2 мм и длиной 40 мм (рис. 26). Эта трубка помещена в тигель 3 для расплава, заполненный чистой платиной. Для термоизоляции тигель помещен в сосуд 5 с порошком тория. Платина расплавляется индукционными токами, возбуждаемыми переменным током, который протекает по обмотке 4. При охлаждении платина затвердевает и ее температура устанавливается и сохраняется на значении 2045 К. Трубка и тигель для расплава сверху закры-

25

К расчету энергетической освещенности



26

Эталон силы света

ты крышкой I с отверстием, через которое выходит излучение, определяющее единицу силы света. Это излучение направляется на поверхность, играющую роль фотометра. Сила света от другого источника определяется из сравнения освещенностей, создаваемых им и эталоном.

Кандела (кд) — это сила света, излучаемого перпендикулярно поверхности черного излучателя с площадью $\frac{1}{6} \cdot 10^{-5} \text{ м}^2$ при температуре затвердевания платины, находящейся под давлением 101 325 Па.

Кандела является основной световой единицей. На основе канделы определяют все другие фотометрические величины. Будем обозначать их теми же буквами, что и энергетические величины, с добавлением индекса V . Названия фотометрических величин в большинстве случаев получаются из названий энергетических заменой слова излучение на свет или соответствующих производных от них, а также отбрасыванием прилагательного энергетический. Каждой фотометрической величине соответствует энергетическая. Их свойства аналогичны. Сила света обозначается dI_V . Она соответствует энергетической силе излучения dI [см. (7.2)].

Световой поток. Световым потоком называется произведение силы света dI_V источника на телесный угол $d\Omega$, в котором испущен свет:

$$d\Phi_V = dI_V d\Omega. \quad (7.11)$$

Световой поток аналогичен мощности излучения в энергетическом определении [см. (7.2)], однако его обычно обозначают не P_V , а Φ_V . Из (7.11) видно, что если точечный источник силой света I_{0V} излучает во всех направлениях, то полный поток его излучения равен $\Phi_V = 4\pi I_{0V}$ [см. (7.4)].

Спектральная плотность светового потока определяется формулой

$$d\Phi_{V\lambda} = dI_{V\lambda} d\Omega. \quad (7.12)$$

Яркость. Она вводится аналогично определению (7.6) энергетической яркости:

$$L_V = \frac{dI_V}{d\sigma'} = \frac{dI_V}{d\sigma \cos \theta} = \frac{d\Phi_V}{d\Omega d\sigma \cos \theta}, \quad \text{a)} \quad L_{V\lambda} = \frac{d\Phi_{V\lambda}}{d\Omega d\sigma \cos \theta}. \quad \text{б)} \quad (7.13)$$

Перепишем (7.13 а) в виде

$$dI_V = L_V d\sigma \cos \theta. \quad (7.14)$$

Зависимость L_V от углов θ обуславливается свойствами поверхности. Если L_V не зависит от углов, то $L_V d\sigma = (dI)_{V0} = \text{const}$ и (7.14) принимает вид

$$dI_V(\theta) = (dI)_{V0} \cos \theta. \quad (7.15)$$

где dI_{V0} — сила света в направлении угла θ , $(dI)_{V0}$ — сила света по нормали к поверхности. Зависимость (7.15) называется законом Ламберта, а поверхности, излучение которых характеризуется условием $L_V = \text{const}$, — ламбертовскими. Излучение от таких поверхностей имеет диффузный характер, и поэтому их называют также диффузно излучающими.

В качестве примера приведем значения яркостей некоторых источников: ночного безлунного неба $\sim 10^{-4}$ кд/м², полной луны — 10^3 кд/м², безоблачного ясного дневного неба — 10^7 кд/м², спиралей лампочки накаливания — 10^6 кд/м², Солнца — 10^9 кд/м².

Светимость. Аналогично (7.7) имеем

$$M_V = \frac{d\Phi_V}{d\sigma} = \int_{\Omega=2\pi} L_V \cos \theta d\Omega,$$

$$M_{V\lambda} = \int_{\Omega=2\pi} L_{V\lambda} \cos \theta d\Omega.$$

(7.16)

Освещенность. В соответствии с (7.10) освещенность и спектральная плотность освещенности определяются соотношениями

$$E_V = \frac{d\Phi_V}{d\sigma}, \quad (7.17)$$

$$E_{V\lambda} = \frac{d\Phi_{V\lambda}}{d\sigma}.$$

Световая экспозиция. Эта величина определяется формулой

$$dH_V = \Phi_V dt, \quad (7.19)$$

где dt — промежуток времени, в течение которого излучается или воспринимается световой поток Φ_V .

Соотношения между энергетическими и световыми характеристиками излучения. Одна и та же мощность излучения в различных интервалах длин волн вызывает совершенно различные зрительные ощущения яркости источника. Если длины волн лежат вне видимой части спектра, то источник остается невидимым независимо от его мощности. Наиболее ярким при фиксированной мощности он представляется при излучении на длине волны 555 нм. Если ставится вопрос об освещении помещения, то нельзя ограничиться только обсуждением энергетической стороны дела. Надо принять во внимание спектральный состав излучения источников света и связь истинной мощности излучения в различных интервалах длин волн с субъективным зрительным ощущением мощности.

Прежде всего необходимо оценить соотношения между спектральной плотностью светового потока $\Phi_{V\lambda}$ и спектральной мощностью P_λ при одинаковом зрительном ощущении на различных длинах волн. Количественно это соотношение характеризуют величиной

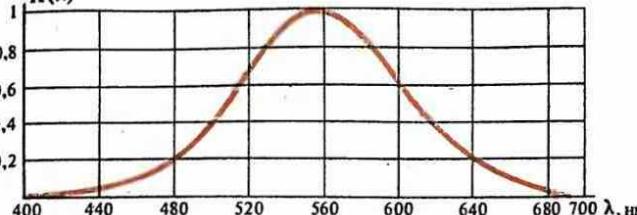
$$V(\lambda) = \Phi_{V\lambda} / P_\lambda, \quad (7.20)$$

называемой **спектральной световой эффективностью**. Световой поток Φ_V оценивается по зрительному ощущению мощности излучения в люменах, а P выражается в ваттах. Спектральная световая эффективность зависит от длины волны. Она равна нулю для длин волн вне видимой части спектра. Ее значение в видимом диапазоне испытывает небольшие изменения для различных индивидуумов. В результате измерений, проведенных с привлечением большого числа людей с нормальным зрением, получена некоторая среднестатистическая кривая $V(\lambda)$, которая принята в качестве стандартной.

Для длины волны 555 нм $V(\lambda)$ достигает максимального значения. Вместо $V(\lambda)$ удобно пользоваться безразмерной величиной, называемой относительной **спектральной световой эффективностью** и равной

$$K(\lambda) = V(\lambda) / V(555 \text{ нм}). \quad (7.21)$$

Кривая $K(\lambda)$ приведена на рис. 27. Наиболее «экономно» для освещения пользоваться излучением с длиной волны 555 нм. Однако это неприемлемо, поскольку все окружающие предметы были бы окрашены в зеленый цвет различной интенсивности. Необходимо освещение осуществлять излучением со спектральным составом «среднего солнечного света», который отличает

$K(\lambda)$ 

27

Кривая относительной световой эффективности

ется от спектрального состава солнечного излучения (см. § 1). Принимая экономичность источника освещения на длине волны 555 нм за единицу, для экономичности источника, генерирующего средний солнечный свет, получаем значение 0,35. Экономичность солнечного излучения составляет примерно 0,14, ламп накаливания — 0,02, а люминесцентных ламп — 0,06.

Если спектральная плотность энергетической мощности излучения P_λ известна, то полная мощность излучения равна

$$P = \int_0^{\infty} P_\lambda d\lambda.$$

(7.22a)

Спектральная плотность светового потока в соответствии с (7.20) и (7.21) определяется формулой

$$\Phi_{\nu\lambda} = V(555 \text{ нм}) K(\lambda) P_\lambda,$$

(7.22б)

а световой поток выражается в виде

$$\Phi_\nu = \int_{\lambda_1}^{\lambda_2} \Phi_{\nu\lambda} d\lambda = V(555 \text{ нм}) \int_{\lambda_1}^{\lambda_2} K(\lambda) P_\lambda d\lambda.$$
(7.23)

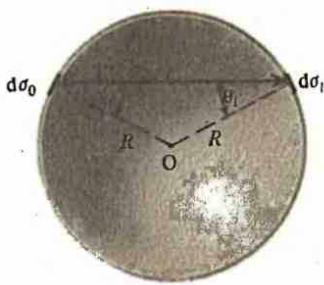
Эти формулы связывают значения энергетических и фотометрических величин. Множитель $K(\lambda)$ равен нулю вне диапазона видимых волн. Применим формулу (7.23) к излучению эталона, с помощью которого определяется кандела. Поскольку эталон излучает как абсолютно черное тело с температурой 2045 К, можно правую часть (7.23) вычислить с помощью формулы Планка для этой температуры, а левую часть — с помощью определения эталона. Переходим в (7.23) к силе излучения по формулам (7.2) и (7.11):

$$dI_\nu = V(555 \text{ нм}) \int_{\lambda_1}^{\lambda_2} K(\lambda) dI_\lambda d\lambda.$$
(7.24)

Эталон определен через силу света в направлении, перпендикулярном излучающей поверхности. Поверхность абсолютно черного тела излучает как ламбертовская. Поэтому для силы света, излучаемой с площади $d\sigma$ в угол 2π , отнесенной к площади поверхности, находим

$$\frac{dI_\nu}{d\sigma} = \frac{2\pi}{(1/6)10^{-5}} \frac{\text{кд}}{\text{м}^2} = 2\pi \cdot 6 \cdot 10^5 \frac{\text{кд} \cdot \text{ср}}{\text{м}^2 \cdot \text{ср}} = 2\pi \cdot 6 \cdot 10^5 \text{ лм}/(\text{м}^2 \cdot \text{ср}).$$
(7.25)

Энергетическая сила света, излучаемая с поверхности $d\sigma$, отнесенная к площади поверхности, равна



28

К расчету освещенности шаровой полости

$$\frac{dI}{d\sigma} = \frac{M}{2\pi}, \quad (7.26)$$

поскольку по определению энергетическая светимость является полной мощностью, излученной в полупространство и отнесенной к площади поверхности, а энергетическая сила света является полной мощностью, отнесеной к телесному углу. С учетом (7.25) и (7.26) равенство (7.24) можем записать в виде

$$2\pi \cdot 6 \cdot 10^5 \text{ лм}/(\text{м}^2 \cdot \text{ср}) = V(555 \text{ нм}) \int_{\lambda_1}^{\lambda_2} K(\lambda) \frac{c}{8\pi} w_\lambda d\lambda, \quad (7.27)$$

где для установления связи M_λ с спектральной плотностью w_λ использовано соотношение (50.24). Запишем выражение (50.28) в виде

$$w_\lambda = \frac{16\pi^2 c \hbar}{\lambda^5} \frac{1}{\exp[2\pi\hbar c/(kT\lambda)] - 1} \quad (7.28)$$

и подставим его в формулу (7.27) с $T = 2045$ К. После этого в (7.27) можно произвести интегрирование, поскольку $K(\lambda)$ известно. Вычисления приводят к следующему результату:

$$V(555 \text{ нм}) = 1/M = 683 \text{ лм/Вт}, \quad (7.29)$$

$$M(555 \text{ нм}) = 1,46 \cdot 10^{-3} \text{ Вт/лм}. \quad (7.30)$$

Постоянные $V(555 \text{ нм})$ и $M(555 \text{ нм})$ называются соответственно фотометрическим эквивалентом излучения и энергетическим эквивалентом света.

Пример 7.1. Элемент $d\sigma_0$ шаровой поверхности радиусом R испускает свет в соответствии с законом Ламберта. Определить освещенность различных точек поверхности полости (рис. 28).

Вычислим освещенность элемента поверхности $d\sigma_1$. Нормали к элементам поверхности $d\sigma_0$ и $d\sigma_1$ проведем в направлении центра полости, а углы между этими нормалью и линией, соединяющей элементы поверхности, обозначим соответственно θ_0 и θ_1 . Яркость элемента поверхности $d\sigma_0$ обозначим L_V . Сила света с элемента поверхности дается формулой

$$dI_V = L_V d\sigma_0 \cos \theta_0, \quad (7.31)$$

причем L_V не зависит от угла θ_0 .

На основании (7.17) и (7.11) для освещенности элемента $d\sigma_1$ получаем

$$E_V = dI_V \cos \theta_1 / r^2, \quad (7.32)$$

где r — расстояние между элементами $d\sigma_0$ и $d\sigma_1$. Линия, соединяющая элементы $d\sigma_0$ и $d\sigma_1$, является хордой к окружности в сечении полости плоскостью, проходящей через центр шара и элементы $d\sigma_0$ и $d\sigma_1$. Видно, что $\theta_0 = \theta_1$ и $r = 2R \cos \theta_0$. Следовательно, (7.32) принимает вид

Можно ли указать, какова будет освещенность письменного стола, если известно, что весьма велика энергетическая светимость источника освещения?

Как можно увеличить светимость, не изменяя мощности источника излучения?

Если яркость участка поверхности не зависит от направления (ламбертовская поверхность), то значит ли это, что световой поток от этого участка во всех направлениях одинаков?

$$E_V = \frac{I_V \cos \theta_0}{4R^2 \cos^2 \theta_0} = \frac{I_V d\sigma_0 \cos^2 \theta_0}{4R^2 \cos^2 \theta_0} = \frac{L_V d\sigma_0}{4R^2}, \quad (7.33)$$

где использовано соотношение (7.14). Таким образом, энергетическая освещенность всех точек поверхности полости одинакова.

Пример 7.2. Абсолютно черное тело при температуре затвердевания платины при давлении 101 325 Па излучает с плоской круговой площадки радиусом 3 мм. Найти освещенность другой площадки, расположенной в направлении, образующем угол 45° с нормалью к излучающей поверхности, на расстоянии 0,75 м от нее. Угол между нормалью к площадке, воспринимающей излучение, и линией, соединяющей испускающую и воспринимающую излучение площадки, равен 30° .

Излучение абсолютно черного тела при температуре затвердевания платины с большой точностью эквивалентно излучению эталона, определяющего кандидатуру. Площадь излучающей площадки равна $2,83 \cdot 10^{-5} \text{ м}^2$. Следовательно, по нормали к площадке испускается свет силой $I_V(0^\circ) = 2,83 \cdot 10^{-5} \cdot 6 \cdot 10^5 \text{ кд} = 17 \text{ кд}$. По закону Ламберта сила света под углом 45° к нормали равна $I_V(45^\circ) = I_V(0^\circ) \cos 45^\circ = 12 \text{ кд}$. Освещенность площадки, находящейся на расстоянии r от излучателя и наклоненной под углом 30° к световому потоку, равна

$$E_V = \frac{I_V(45^\circ) \cos 30^\circ}{r^2} = \frac{12 \cdot \cos 30^\circ}{(0,75)^2} \frac{\text{кд}}{\text{м}^2} = 18,5 \text{ лк.} \quad (7.34)$$

Пример 7.3. Найти яркость поверхности Солнца, зная, что в ясный день прямое падение солнечного излучения создает на поверхности, перпендикулярной направлению на Солнце, освещенность $7 \cdot 10^4 \text{ лк}$. Поглощением и рассеянием солнечных лучей в атмосфере пренебречь.

Поскольку Солнце излучает как абсолютно черное тело, у него яркость $L_V = \text{const}$ и, следовательно, между яркостью и светимостью имеется аналогичное (7.9) соотношение. Отсюда для яркости поверхности Солнца находим

$$L_V = \frac{M_V}{\pi} = \frac{1}{\pi} \left(\frac{d\Phi_V}{d\sigma} \right)_0, \quad (7.35)$$

где принята во внимание формула (7.16). Учитывая, что световой поток от Солнца убывает обратно пропорционально квадрату расстояния, а излучение Солнца изотропно, для освещенности поверхности, перпендикулярной световому потоку, на основании (7.17) получаем

$$E_V = \frac{d\Phi_V}{d\sigma} = \left(\frac{d\Phi_V}{d\sigma} \right)_0 \left(\frac{r}{R} \right)^2, \quad (7.36)$$

где r — радиус Солнца, R — радиус орбиты Земли. Известно, что диаметр Солнца с Земли виден под углом $32'$. Следовательно, $r/R = 4,7 \cdot 10^{-3}$. Из (7.35) с учетом (7.36) находим

$$L_V = \frac{1}{\pi} E_V \left(\frac{R}{r} \right)^2 = \frac{7 \cdot 10^4}{3,14 (4,7 \cdot 10^{-3})^2} \frac{\text{кд}}{\text{м}^2} = 10^9 \text{ кд/м}^2. \quad (7.37)$$

Пример 7.4. На плоскую поверхность под углом $\theta = 45^\circ$ падает излучение, содержащее длины волн между $\lambda_1 = 555 \text{ нм}$ и $\lambda_2 = 600 \text{ нм}$. Спектральная плотность потока энергии излучения в направлении поверхности по шкале длин волн на длине волны 555 нм равна $s_{\lambda_0} = 10^4 \text{ Вт/м}^3$ и спадает до нуля на волне 600 нм по линейному закону. Считать, что относительная спектральная световая эффективность в этом интервале длин волн убывает также по линейному закону от $K_1 = 1$ до $K_2 = 0,5$. Найти освещенность поверхности.

Зависимость спектральной плотности потока энергии излучения от длины волны в заданном диапазоне дается формулой

$$s_{\lambda} = s_{\lambda 0} \frac{\lambda_2 - \lambda}{\lambda_2 - \lambda_1}, \quad (7.38)$$

а относительная спектральная световая эффективность

$$K(\lambda) = \frac{K_1(\lambda_2 - \lambda)}{\lambda_2 - \lambda_1} - \frac{K_2(\lambda_1 - \lambda)}{\lambda_2 - \lambda_1}.$$

Мощность падающего на поверхность $d\sigma'$ в интервале длин волн $d\lambda$ излучения равна

$$dP = s_{\lambda} d\lambda d\sigma' = s_{\lambda 0} \frac{\lambda_2 - \lambda}{\lambda_2 - \lambda_1} d\lambda d\sigma', \quad (7.39)$$

где $d\sigma' = d\sigma \cos \theta$ — площадь проекции элемента $d\sigma$ поверхности на плоскость, перпендикулярную падающему на нее потоку энергии излучения. Для спектральных плотностей падающего светового потока имеем

$$d\Phi_{V\lambda} = dP_{\lambda} V(555 \text{ нм}) K(\lambda), \quad (7.40)$$

где $V(555 \text{ нм})$ определяется в (7.29). Из (7.40) с учетом (7.39) следует, что

$$E_{V\lambda} = \frac{d\Phi_{V\lambda}}{d\sigma} = s_{\lambda} V(555 \text{ нм}) K(\lambda) \cos \theta,$$

и поэтому освещенность поверхности равна

$$\begin{aligned} E_V &= \int_{\lambda_1}^{\lambda_2} E_{V\lambda} d\lambda = V(555 \text{ нм}) \int_{\lambda_1}^{\lambda_2} s_{\lambda} K(\lambda) \cos \theta d\lambda = \\ &= \frac{V(555 \text{ нм}) s_{\lambda 0} \cos \theta}{(\lambda_2 - \lambda_1)^2} \int_{\lambda_1}^{\lambda_2} (\lambda_2 - \lambda) [K_1(\lambda_2 - \lambda) - K_2(\lambda_1 - \lambda)] d\lambda = \\ &= \frac{V(555 \text{ нм}) s_{\lambda 0} \cos \theta}{(\lambda_2 - \lambda_1)^2} \left[\frac{1}{2} (K_1 + K_2) (\lambda_2 - \lambda_1) (\lambda_1^2 + \lambda_2^2 - \lambda_1 \lambda_2) + \right. \\ &\quad \left. + \frac{1}{3} (K_1 - K_2) (\lambda_2^3 - \lambda_1^3) \right]. \end{aligned} \quad (7.41)$$

Подставив числовые значения величин, входящих в формулу и выполнив вычисления, найдем

$$\begin{aligned} E_V &= \frac{683 \cdot 10^4 \cdot 0,701}{(45 \cdot 10^{-9})^2} [0,5 \cdot 1,5 \cdot 45 \cdot 10^{-9} (3,08 \cdot 10^5 + 3,6 \cdot 10^5 - 3,3 \cdot 10^5) \cdot 10^{-18} + \\ &+ \frac{1}{3} 0,5 (2,16 \cdot 10^8 - 1,71 \cdot 10^8) \cdot 10^{-27}] = 54 \text{ лк.} \end{aligned}$$

При определении энергетических величин исходят из мощности излучения, а при определении фотометрических величин — из силы света. Мощности излучения соответствует не сила света, а световой поток.

Перерасчет энергетических величин в фотометрические осуществляется посредством спектральных характеристик излучения.

Задачи

- 1.1.** На вход выпрямителя подается синусоидальное напряжение $U = U_0 \sin t$. Найти ряды Фурье для напряжения на выходе, если выпрямитель исключает полупериоды с отрицательным значением напряжения.
- 1.2.** Лампа без рефлекторов находится над горизонтальным столом на высоте 0,5 м от его поверхности. Ее полный световой поток 1000 лм равномерно распределяется по всем направлениям. Какова освещенность стола по вертикали под лампой?
- 1.3.** Определить энергетическую светимость поверхности Солнца, если на земной орбите плотность потока энергии солнечного излучения составляет $1400 \text{ Вт}/\text{м}^2$, а диаметр Солнца виден с Земли под углом $32'$.
- 1.4.** Освещенность земной поверхности вертикально падающими лучами Солнца составляет $7 \cdot 10^4 \text{ лк}$ (влиянием атмосферы пренебрегаем). Найти силу света, испускаемого Солнцем.
- 1.5.** Пользуясь данными об излучении Солнца, приведенными в задаче 1.3, найти энергетическую яркость поверхности Солнца.
- 1.6.** Электрический пробой в воздухе наступает примерно при напряженности электрического поля 3 МВ/м. При какой плотности потока энергии плоских электромагнитных волн можно наблюдать искру в воздухе?
- 1.7.** Точечный источник мощностью 1 Вт излучает равномерно по всем направлениям фотонов с длиной волны 0,5 мкм. Какова плотность потока фотонов на расстоянии 1 км от источника?
- 1.8.** Абсолютно черное тело с температурой 1500 К испускает световой поток в полупространство. Найти освещенность в 1 м на перпендикуляре к поверхности при нормальном падении лучей на площадку.
- 1.9.** Найти амплитуды напряженностей электрического и магнитного полей волны, плотность потока энергии которой $1 \text{ Вт}/\text{м}^2$.
- 1.10.** Найти амплитуду напряженности магнитного поля волны, распространяющейся в воде ($n=1,33$), и плотность потока энергии волны, если амплитуда электрического поля 20 кВ/м.
- 1.11.** В воде ($n=1,33$) распространяется эллиптически поляризованная волна, плотность потока энергии волны которой $20 \text{ Вт}/\text{м}^2$. Эксцентриситет эллипса, описываемого концом вектора напряженности, равен 0,6. Найти амплитуды колебаний компонент напряженности электрического вектора по главным направлениям эллипса.
- 1.12.** Чему равно световое давление при нормальном падении света на полностью отражающую поверхность при плотности потока энергии волн $1 \text{ Вт}/\text{м}^2$?
- 1.13.** Чему равны напряженности электрического и магнитного полей волны с длиной 555 нм, если при нормальном падении она создает освещенность 500 лк?

Ответы

- 1.1.** $U_{\text{вых}} = \frac{U_0}{\pi} \left(1 + \frac{\pi}{1 \cdot 2} \sin t - \frac{2}{1 \cdot 3} \cos 2t - \frac{2}{3 \cdot 5} \cos 4t - \frac{2}{5 \cdot 7} \cos 6t - \dots \right)$. **1.2.** 321 лк.
- 1.3.** $63,4 \text{ МВт}/\text{м}^2$. **1.4.** $1,6 \cdot 10^{27} \text{ кд}$. **1.5.** $2 \cdot 10^7 \text{ Вт}/(\text{м}^2 \cdot \text{ср})$. **1.6.** $1,2 \cdot 10^{10} \text{ Вт}/\text{м}^2$. **1.7.** $2 \cdot 10^{11} \text{ фотонов}/(\text{м}^2 \cdot \text{с})$.
- 1.8.** 17,3 лк. **1.9.** $27,45 \text{ В/м}$; $0,073 \text{ А/м}$. **1.10.** $70,6 \text{ А/м}$; $706 \text{ кВт}/\text{м}^2$. **1.11.** $78,9 \text{ В/м}$; $73,4 \text{ В/м}$. **1.12.** $0,66 \cdot 10^{-8} \text{ Па}$.
- 1.13.** $23,5 \text{ В/м}$; $7,6 \cdot 10^{-2} \text{ А/м}$.

Основная идея:
реальнонос излучение
имеет конечную
продолжительность и
происходит со случайно
изменяющимися
амплитудой и фазой.
Его спектральный состав
анализируется с помощью
Фурье-преобразований.

Немонохроматическое и хаотическое излучение



§ 8 Спектральный состав функций

Теория рядов и интегралов Фурье применяется для анализа спектрального состава функций. Анализируется соотношение между продолжительностью импульса и шириной спектра.

Ряд Фурье в действительной форме. В математике доказывается, что при весьма общих условиях, которые в физических задачах обычно удовлетворяются, периодическая функция $f(t)$ с периодом T (рис. 29) [$f(t+T)=f(t)$] может быть представлена рядом Фурье:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos n\omega t + b_n \sin n\omega t), \quad (8.1)$$

где

$$\omega = 2\pi/T,$$

(a)

$$a_n = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \cos n\omega t dt, \quad (8.2)$$

$$b_n = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \sin n\omega t dt. \quad (b)$$

Периодическая функция координат $f(z)$ с периодом L (рис. 30) [$f(z+L)=f(z)$] имеет вид, аналогичный (8.1), но с заменой $\omega \rightarrow k = 2\pi/L$, $T \rightarrow L$, $t \rightarrow z$.

Ряд Фурье в комплексной форме. С помощью формулы Эйлера $e^{ia} = \cos a + i \sin a$ ряд (8.1) представим

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{in\omega t}, \quad (8.3)$$

где

$$c_n = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-in\omega t} dt. \quad (8.4)$$

Коэффициенты a_n , b_n и c_n связаны равенствами:

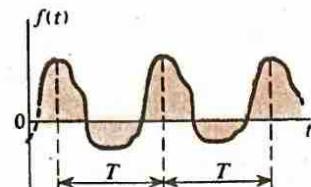
$$c_0 = \frac{1}{2} a_0, \quad c_n = \frac{1}{2} (a_n - ib_n) \quad (n = 1, 2, 3, \dots), \quad (8.5)$$

$$c_{-n} = \frac{1}{2} (a_{-n} + ib_{-n}) \quad (n = -1, -2, -3, \dots).$$

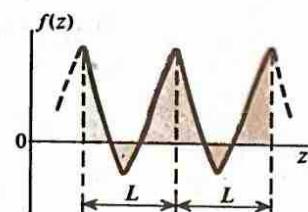
Интеграл Фурье в действительной форме. Непериодическую функцию $f(t)$ (рис. 31) нельзя представить рядом Фурье. Если функция $f(t)$ кусочно-непрерывна, имеет конечное число экстремумов и абсолютно интегрируема на интервале $(-\infty, \infty)$, то она может быть представлена интегралом Фурье

$$f(t) = \frac{1}{\pi} \int_0^{\infty} d\omega \int_{-\infty}^{\infty} f(\tau) \cos [\omega(t-\tau)] d\tau, \quad (8.6)$$

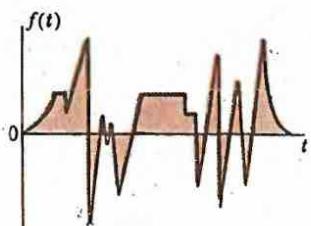
получаемым из (8.1) при $T \rightarrow \infty$.



29
Периодическая функция времени



30
Периодическая функция координат



31
Непериодическая функция времени

Интеграл Фурье в комплексной форме. С помощью формулы Эйлера (8.6) преобразуется к виду

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega, \quad (8.7)$$

где

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \quad (8.8)$$

— Фурье-образ функции $f(t)$.

Спектр амплитуд и спектр фаз. Преобразуем слагаемые ряда Фурье (8.1) следующим образом:

$$\begin{aligned} a_n \cos n\omega t + b_n \sin n\omega t &= \sqrt{a_n^2 + b_n^2} [(a_n / \sqrt{a_n^2 + b_n^2}) \cos n\omega t + (b_n / \sqrt{a_n^2 + b_n^2}) \sin n\omega t] = \\ &= A_n \cos(n\omega t - \varphi_n), \end{aligned} \quad (8.9)$$

$$\cos \varphi_n = a_n / \sqrt{a_n^2 + b_n^2}, \quad \sin \varphi_n = b_n / \sqrt{a_n^2 + b_n^2}, \quad A_n = \sqrt{a_n^2 + b_n^2}, \quad \operatorname{tg} \varphi_n = b_n / a_n. \quad (8.10)$$

В результате формула (8.1) принимает вид

$$f(t) = \frac{A_0}{2} + \sum_{n=1}^{\infty} A_n \cos(n\omega t - \varphi_n), \quad (8.11)$$

где $A_0 = a_0$. Совокупность величин A_n называется спектром амплитуд функции $f(t)$, а совокупность φ_n — спектром фаз. Частоты по определению имеют положительные значения.

Нахождение спектра амплитуд и фаз из рядов Фурье в комплексной форме. Для вещественной функции $f(t)$ в (8.3) выполняется равенство $c_n^* = c_{-n}$. Преобразуя (8.3) к виду (8.1) и сравнивая коэффициенты при $\cos n\omega t$ и $\sin n\omega t$, находим:

$$a_0 = 2c_0, \quad a_n = 2\operatorname{Re} c_n, \quad b_n = -2\operatorname{Im} c_n, \quad (8.12)$$

где Re и Im — действительная и мнимая части комплексного числа. Из (8.10) с учетом (8.12) получаем

$$A_0 = 2c_0, \quad A_n = 2|c_n|, \quad \operatorname{tg} \varphi_n = -\operatorname{Im} c_n / \operatorname{Re} c_n. \quad (8.13)$$

Непрерывный спектр. Фурье-образ (8.8) называется комплексным спектром или просто спектром функции $f(t)$. Он полностью определяет функцию $f(t)$ и эквивалентен амплитудному и фазовому спектрам. Имеем

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega = \frac{1}{2\pi} \int_0^{\infty} [F(\omega) e^{i\omega t} + F(-\omega) e^{-i\omega t}] d\omega. \quad (8.14)$$

Для вещественной функции $f(t)$ соблюдается соотношение $F^*(\omega) = F(-\omega)$ и можно написать $F(\omega) e^{i\omega t} + F(-\omega) e^{-i\omega t} = F(\omega) e^{i\omega t} + [F(\omega) e^{i\omega t}]^* = 2\operatorname{Re} [F(\omega) e^{i\omega t}]$.

Следовательно, (8.14) принимает вид

$$2 \quad f(t) = \frac{1}{\pi} \int_0^\infty [\operatorname{Re} F(\omega) \cos \omega t - \operatorname{Im} F(\omega) \sin \omega t] d\omega. \quad (8.15)$$

Преобразуя подынтегральное выражение в (8.15) в соответствии с формулами (8.9) и (8.10), находим

$$f(t) = \frac{1}{2\pi} \int_0^\infty A(\omega) \cos [\omega t - \varphi(\omega)] d\omega, \quad (8.16)$$

где

$$A(\omega) = 2|F(\omega)|, \quad \operatorname{tg} \varphi(\omega) = -\operatorname{Im} F(\omega) / \operatorname{Re} F(\omega). \quad (8.17)$$

Величины $A(\omega)$ и $\varphi(\omega)$ представляют амплитудный и фазовый спектры функции $f(t)$. Наличие множителя $1/(2\pi)$ в (8.16) показывает, что $A(\omega)$ является плотностью амплитуд, отнесенной к частотам $v = \omega/(2\pi)$, а не к круговым частотам, поскольку формула (8.16) может быть представлена в виде

$$f(t) = \int_0^\infty A(2\pi v) \cos [2\pi vt - \varphi(2\pi v)] dv. \quad (8.18)$$

Таким образом, периодические функции характеризуются дискретными спектрами, а непериодические — непрерывными.

Спектр прямоугольных импульсов. Пусть имеется бесконечная последовательность прямоугольных импульсов величиной U_0 и продолжительностью τ , повторяющихся через промежутки времени T (рис. 32). Поскольку функция четна, $b_n = 0$, а коэффициенты при косинусе равны

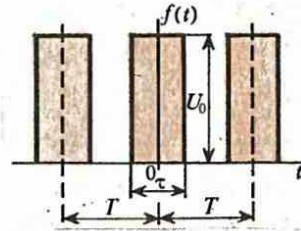
$$a_n = \frac{2}{T} U_0 \int_{-\pi/2}^{\pi/2} \cos n\omega t dt = 2 \frac{\tau}{T} U_0 \frac{\sin(n\omega\tau/2)}{n\omega\tau/2}, \quad (8.19)$$

где $\omega = 2\pi/T$. Отсюда следует, что

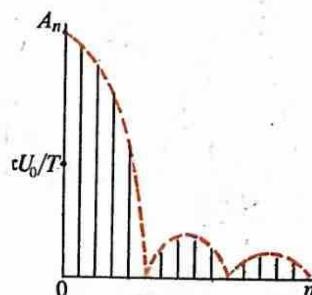
$$A_n = 2 \frac{\tau}{T} U_0 \left| \frac{\sin(n\omega\tau/2)}{n\omega\tau/2} \right|, \quad \varphi_n = 0. \quad (8.20)$$

Спектр амплитуд изображен на рис. 33. Учитывая, что $n\omega\tau/2 = n\pi\tau/T$, заключаем, что число гармоник между нулями амплитуды зависит от отношения τ/T : чем меньше продолжительность импульсов по сравнению с периодом их повторения, тем больше число гармоник ($\Delta n = T/\tau$).

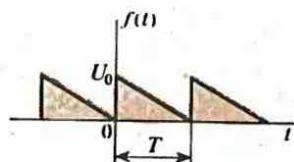
Полезно заметить, что ряд Фурье зависит не только от вида кривой, которую этот ряд представляет, но и от положения начала координат относительно этой кривой. Например, если



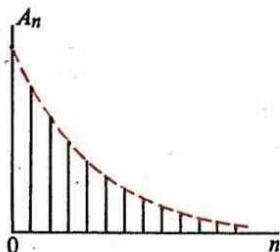
32
Бесконечная последовательность прямоугольных импульсов



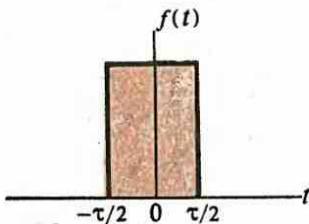
33
Спектр амплитуд бесконечной последовательности прямоугольных импульсов



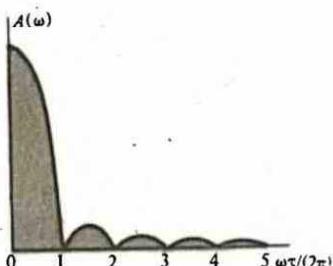
34
Бесконечная последовательность пилообразных импульсов



35 Спектр амплитуд бесконечной последовательности пилообразных импульсов



36 Изолированный прямоугольный импульс



37 Амплитудный спектр изолированного прямоугольного импульса

начало координат на рис. 32 расположить не в середине импульса, а ближе к одной из его границ, то периодическая функция относительно этого начала не будет обладать определенной четностью. Следовательно, ее ряд Фурье будет отличаться от (8.32), поскольку в нем наряду с членами, содержащими $\cos n\omega t$, присутствуют члены с $\sin n\omega t$.

Спектр пилообразных импульсов. Рассмотрим бесконечную последовательность пилообразных импульсов (рис. 34).

Имеем:

$$a_0 = \frac{2}{T} U_0 \int_{-\tau/2}^{\tau/2} f(t) dt = \frac{2}{T} U_0 \int_0^{\tau} \left(1 - \frac{t}{\tau}\right) dt = U_0,$$

$$a_n = \frac{2}{T} U_0 \int_0^{\tau} \left(1 - \frac{t}{\tau}\right) \cos n\omega t dt = 0, \quad (8.21)$$

$$b_n = \frac{2}{T} U_0 \int_0^{\tau} \left(1 - \frac{t}{\tau}\right) \sin n\omega t dt = \frac{U_0}{\pi} \frac{1}{n}.$$

Спектры амплитуд и фаз даются формулами

$$A_n = U_0 / (n\pi), \quad \varphi_n = \pi/2. \quad (8.22)$$

Спектр амплитуд изображен на рис. 35: амплитуды убывают обратно пропорционально номеру гармоники.

Спектр изолированного прямоугольного импульса. Этот импульс изображен на рис. 36. Имеем

$$F(\omega) = U_0 \int_{-\tau/2}^{\tau/2} e^{-i\omega t} dt = U_0 \tau \frac{\sin(\omega\tau/2)}{\omega\tau/2}. \quad (8.23)$$

Следовательно,

$$A(\omega) = 2U_0 \tau \left| \frac{\sin(\omega\tau/2)}{\omega\tau/2} \right|, \quad \varphi(\omega) = 0. \quad (8.24)$$

Амплитудный спектр $A(\omega)$ показан на рис. 37.

Спектр экспоненциально убывающей функции. Найдем спектр функции

$$f(t) = \begin{cases} 0 & \text{при } -\infty < t < 0, \\ U_0 e^{-i\omega t} & \text{при } 0 < t < \infty, \end{cases}$$

график которой изображен на рис. 38. Имеем

$$F(\omega) = U_0 \int_0^{\infty} e^{-i\omega t} dt = U_0 / (1 + i\omega\tau).$$

Следовательно,

$$2 A(\omega) = 2 U_0 \tau / \sqrt{1 + \omega^2 \tau^2}, \quad \varphi(\omega) = \omega \tau. \quad (8.25)$$

Амплитудный $A(\omega)$ и фазовый $\varphi(\omega)$ спектры показаны на рис. 39 и 40.

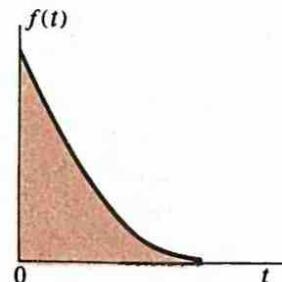
Соотношение между продолжительностью импульса и шириной спектра. Продолжительностью импульса называется промежуток времени Δt , в течение которого импульс существенно отличается от нуля. Шириной спектра называется интервал частот $\Delta\omega$, на котором амплитуда спектра существенно отлична от нуля. В этих определениях имеется неопределенность, а именно не уточнено, что понимать под словами «существенно отлична от нуля». В зависимости от определения этого понятия несколько изменяется соотношение между продолжительностью конкретного импульса и шириной его спектра. При выбранном определении данного понятия это соотношение изменяется для различных импульсов в зависимости от формы. Поэтому универсального соотношения между продолжительностью импульса и шириной спектра не существует. Однако есть универсальная закономерность в соотношениях между продолжительностью импульса и шириной спектра, которая соблюдается при различных определениях понятия «существенно отлична от нуля» и для импульсов различной формы. Эта закономерность гласит: ширина спектра обратно пропорциональна продолжительности импульса. Выведем эту закономерность на примере прямоугольного изолированного импульса (см. рис. 36), когда определение его продолжительности не вызывает сомнений, — продолжительность импульса $\Delta t = \tau$. С другой стороны, ширина спектра этого импульса (см. рис. 37) также определяется довольно естественно: это интервал частот от нуля до частоты, при которой амплитуда обращается первый раз в нуль, поскольку последующие максимумы амплитуд незначительны по сравнению с основным максимумом при $\omega = 0$. Поэтому для ширины спектра $\Delta\omega$ можем написать равенство

$$\Delta\omega\tau/2 = \pi. \quad (8.26)$$

Следовательно, между шириной спектра по частотам $\Delta\omega = \Delta\omega/(2\pi)$ и продолжительностью импульса $\Delta t = \tau$ существует соотношение

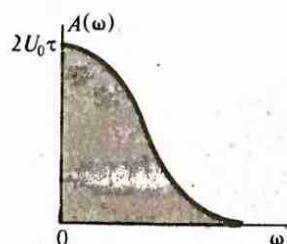
$$\Delta\omega\Delta t \approx 1.$$

Вместо знака строгого равенства использован знак приближенного равенства, чтобы подчеркнуть неопределенность



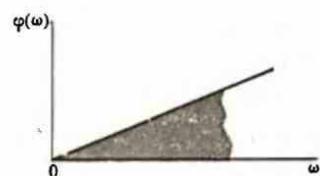
38

График экспоненциально убывающей функции



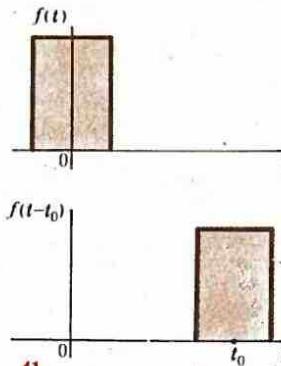
39

Амплитудный спектр экспоненциально убывающей функции



40

Фазовый спектр экспоненциально убывающей функции



41 Смещение начала отсчета времени

определенний Δv и Δt , в результате которой соотношение справедливо лишь с точностью до множителя порядка единицы. Соотношение (8.27) принимается в качестве универсального соотношения между продолжительностью импульса и шириной спектра.

Основной вывод из (8.27) заключается в том, что чем короче продолжительность импульса, тем более широким спектром частот он обладает. Другими словами, нельзя надеяться представить очень короткий импульс набором гармонических функций с небольшим интервалом частот. Если $\Delta t \rightarrow 0$, то в спектре присутствуют всевозможные частоты от малых до очень больших.

Смещение начала отсчета времени. Пусть вместо функции $f(t)$ имеется функция $f(t - t_0)$, где t_0 — постоянная. Это изменение аргумента не изменяет формы импульса, а изменяет лишь начало отсчета времени (рис. 41). При этом изменяется образ Фурье-функции

$$\begin{aligned} F_{t_0}(\omega) &= \int_{-\infty}^{\infty} f(t - t_0) e^{-i\omega t} dt = e^{-i\omega t_0} \int_{-\infty}^{\infty} f(\xi) e^{-i\omega \xi} d\xi = \\ &= F(\omega) e^{-i\omega t_0}, \end{aligned} \quad (8.28)$$

где произведен переход к новой переменной интегрирования $\xi = t - t_0$. Таким образом, смещение начала отсчета времени в точку t_0 изменяет фазу образа Фурье на $-\omega t_0$, т. е. оставляет без изменения амплитудный спектр и изменяет фазовый.

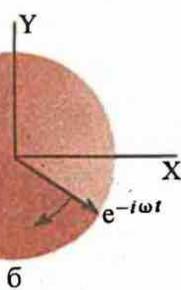
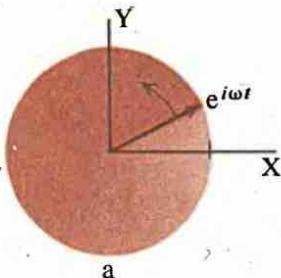
Смещение спектра по частотам. Аналогично может быть выяснено влияние сдвига частот, т. е. замены $F(\omega) \rightarrow F(\omega - \omega_0)$:

$$\begin{aligned} f_{\omega_0}(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega - \omega_0) e^{i\omega t} d\omega = \frac{1}{2\pi} e^{i\omega_0 t} \int_{-\infty}^{\infty} F(\xi) e^{i\xi t} d\xi = \\ &= f(t) e^{i\omega_0 t}, \end{aligned} \quad (8.29)$$

где $\xi = \omega - \omega_0$. Таким образом, смещение спектра на ω_0 эквивалентно модуляции временной функции гармоническим множителем с частотой ω_0 .

Отрицательные частоты. Комплексный спектр (8.8) полностью определяет как спектр амплитуд $A(\omega)$, так и спектр фаз $\phi(\omega)$ посредством соотношений (8.16) и (8.17). Однако в большинстве случаев удобнее обсуждать спектр функции, пользуясь непосредственно выражением $F(\omega)$ без перехода к величинам $A(\omega)$ и $\phi(\omega)$. Поскольку аргумент $F(\omega)$ принимает как положительные, так и отрицательные значения, возникает вопрос о смысле отрицательных частот.

Примем во внимание, что $e^{i\omega t}$ описывает комплексный единичный вектор (рис. 42, a), проведенный из начала координат и врачающийся около этого начала от оси X к оси Y при увеличении t . Это направление принимается за положительное, поскольку оно связано правилом правого винта с направлением



оси Z . Комплексный единичный вектор $e^{-i\omega t}$ при увеличении t вращается в отрицательном направлении (рис. 42, б). Поскольку функция $f(t)$ через образ Фурье выражается формулой (8.7), заключаем, что $F(\omega)$ при $\omega > 0$ описывает плотность спектральной компоненты частотой ω с положительным направлением вращения, а $F(-\omega)$ — плотность спектральной компоненты той же частоты ω , но с отрицательным направлением вращения. Таким образом, обращение к отрицательным частотам связано с изменением базисных функций, с помощью которых осуществляется Фурье-преобразование, а именно с переходом к вращающимся комплексным векторам как базисным функциям Фурье-преобразования. Все сказанное об отрицательных частотах в связи с (8.7), разумеется, полностью сохраняет свое значение и для рядов Фурье в комплексной форме.

Теорема Парсеваля. Исходя из представления периодической функции в виде ряда Фурье (8.3), найдем интеграл по периоду от $|f|^2$:

$$\int_{-T/2}^{T/2} f(t)f^*(t) dt = \sum_{n, n' = -\infty}^{\infty} c_n c_{n'}^* \int_{-T/2}^{T/2} \exp[i(n-n')\omega t] dt.$$

Учитывая, что $\int_{-T/2}^{T/2} \exp[i(n-n')\omega t] dt = T\delta_{nn}$, получаем равенство

$$\frac{1}{T} \int_{-T/2}^{T/2} |f(t)|^2 dt = \sum_{n = -\infty}^{\infty} |c_n|^2, \quad (8.30)$$

которое называется **теоремой Парсеваля**. Для ряда Фурье (8.1) в вещественной форме получаем

$$\frac{1}{T} \int_{-T/2}^{T/2} |f(t)|^2 dt = \frac{a_0^2}{4} + \frac{1}{2} \sum_{n=1}^{\infty} (a_n^2 + b_n^2) = \frac{A_0^2}{4} + \frac{1}{2} \sum_{n=1}^{\infty} A_n^2. \quad (8.31)$$

Теорема Планшереля. Если $f(t)$ представлено в виде (8.7), то для интеграла от квадрата модуля $|f(t)|^2$ аналогично предыдущему случаю получаем формулу

$$\int_{-\infty}^{\infty} |f(t)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |F(\omega)|^2 d\omega = \int_{-\infty}^{\infty} |F(2\pi v)|^2 dv, \quad (8.32)$$

выражающую **теорему Планшереля**.

- Коэффициенты ряда Фурье или Фурье-образ функции зависят от положения начала отсчета времени.

Спектральный состав функции однозначно определяется коэффициентами ряда Фурье.

Амплитудный и фазовый спектры описываются вещественными числами.

Смещение начала отсчета времени приводит к умножению Фурье-образа функции на фазовый множитель, зависящий от смещения и частоты.

Смещение спектра по частотам приводит к умножению функции на фазовый множитель, зависящий от смещения и времени.

§ 9 | Естественная ширина линии излучения

Вычисляется спектр излучения изолированного, неподвижного атома.

Классическая модель излучателя. Простейшей классической моделью излучателя является электрон, колеблющийся около положения равновесия по гармоническому закону. Уравнение свободных колебаний электрона имеет вид

$$m\ddot{x} + m\omega_0^2 x = 0, \quad (9.1)$$

где $m = 9,1 \cdot 10^{-31}$ кг — масса электрона, ω_0 — круговая частота колебаний, имеющая для оптической области спектра излучения порядок 10^{15} с^{-1} . Решение уравнения (9.1) имеет вид

$$x(t) = A \cos(\omega_0 t - \varphi), \quad (9.2)$$

где A — амплитуда колебаний.

Полная энергия колеблющегося электрона равна

$$W = m\dot{x}^2/2 + m\omega_0^2 x^2/2 = (m\omega_0^2/2) A^2. \quad (9.3)$$

При ускоренном движении заряда, как известно, мощность излучения равна

$$P = \frac{1}{6\pi\varepsilon_0} \frac{e^2}{c^3} \dot{x}^2. \quad (9.4)$$

Из-за потери энергии электроном на излучение его энергия уменьшается, а колебания являются затухающими. Закон сохранения энергии для колеблющегося электрона с учетом потерь энергии на излучение имеет вид

$$\frac{dW}{dt} = -\langle P \rangle = -\frac{1}{12\pi\varepsilon_0} \frac{e^2}{c^3} \omega_0^4 A^2, \quad (9.5)$$

где $\langle P \rangle$ — среднее по периоду колебаний, вычисленное с помощью (9.4) и (9.2) с учетом того, что уменьшение амплитуды за период колебаний мало по сравнению с амплитудой или, что то же самое, испущенная за период энергия мала по сравнению со всей энергией излучения. Это условие всегда хорошо соблюдается.

Подставляя в (9.5) выражение для A из (9.3), получаем

$$\frac{dW}{dt} = -\gamma W, \quad (9.6)$$

где

$$\boxed{\gamma = \frac{1}{6\pi\varepsilon_0} \frac{e^2 \omega_0^2}{mc^3}.} \quad (9.7)$$

Решение уравнения (9.6) имеет вид

$$W(t) = W_0 e^{-\gamma t}, \quad (9.8)$$

где $W_0 = W(0)$, и, следовательно, на основании (9.3) амплитуда колебаний электрона изменяется по закону

$$A = A_0 e^{-\gamma t/2}, \quad (9.9)$$

а отклонение x электрона от положения равновесия описывается вместо (9.2) формулой

$$x(t) = A_0 e^{-\gamma t/2} \cos(\omega_0 t - \varphi). \quad (9.10)$$

Необходимо отметить, что все вычисления были проведены в предположении малости затухания амплитуды за период колебаний $T = 2\pi/\omega_0$:

$$\gamma T/2 = \pi\gamma/\omega_0 \ll 1. \quad (9.11)$$

Таким образом, условие малости затухания имеет вид

$$\gamma \ll \omega_0. \quad (9.12)$$

Подставив значения постоянных ϵ_0, e, m, c в выражение (9.7), видим, что для оптических частот ($\omega_0 \sim 10^{15} \text{ с}^{-1}$) условие (9.12) всегда хорошо выполняется. В типичных условиях амплитуда уменьшается в два раза в течение нескольких миллионов колебаний. Другими словами, $\gamma \sim 10^8 \text{ с}^{-1}$. Изобразить графически такое медленное изменение амплитуды колебаний затруднительно, поэтому на рис. 43 колебания электрона показаны при сильно увеличенном значении γ (в сотни тысяч раз).

С учетом затухания колебаний вместо (9.1) можно написать уравнение

$$m\ddot{x} + m\gamma\dot{x} + m\omega_0^2x = 0. \quad (9.13)$$

Прямой проверкой убеждаемся, что решение (9.13) при условии $\gamma \ll \omega_0$ совпадает с (9.10).

Спектральный состав излучения. Если электрон начинает колебаться и излучать в момент $t=0$, то смещение $x(t)$ может быть представлено в виде

$$x(t) = \begin{cases} 0 & \text{при } t < 0, \\ e^{-\gamma t/2}(A_0 e^{i\omega_0 t} + A_0^* e^{-i\omega_0 t}) & \text{при } t > 0, \end{cases} \quad (9.14)$$

поскольку $x(t)$ — вещественная величина.

Излученная в интервале $0 < t < \infty$ энергия дается выражением

$$W = - \int_0^\infty \frac{dW}{dt} dt = \frac{1}{6\pi\epsilon_0} \frac{e^2}{c^3} \int_{-\infty}^\infty \dot{x}^2 dt, \quad (9.15)$$

где использована формула (9.4) и учтено, что $\dot{x}=0$ при $t < 0$.

Представим энергию излучения (9.15) распределенной по частотам. Поскольку затухание мало ($\gamma \ll \omega_0$), можно считать, что

$$\ddot{x} \approx -\omega_0^2 x, \quad (9.16)$$

а $x(t)$ представим в виде интеграла Фурье:

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} d\omega \int_{-\infty}^{\infty} x(\xi) e^{-i\omega\xi} d\xi = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega, \quad (9.17)$$

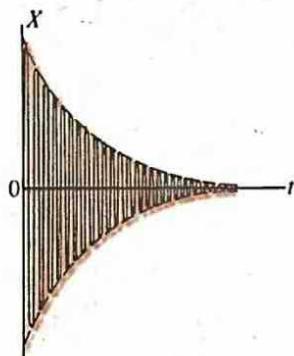
где введено обозначение

$$F(\omega) = \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt = \int_0^{\infty} x(t) e^{-i\omega t} dt \quad (9.18)$$

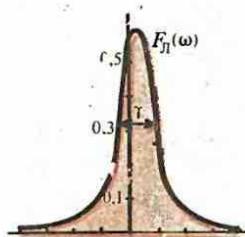
с учетом того, что $\dot{x}=0$ при $t < 0$.

Подставляя \ddot{x} из (9.16) в (9.15) и выражая $x(t)$ по формуле (9.17), получаем

$$W = \frac{1}{6\pi\epsilon_0} \frac{e^2}{c^3} \omega_0^4 \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega \int_{-\infty}^{\infty} F(\xi) e^{i\omega\xi} d\xi =$$



43
Затухающие колебания электрона



44
Лоренцева форма линии

$$= \frac{1}{24\pi^2\epsilon_0} \frac{e^2}{c^3} \omega_0^4 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\omega) F(\xi) d\omega d\xi \int_{-\infty}^{\infty} e^{i(\omega+\xi)t} dt. \quad (9.19)$$

Принимая во внимание соотношение

$$\int_{-\infty}^{\infty} e^{i(\omega+\xi)t} dt = 2\pi\delta(\omega+\xi), \quad (9.20)$$

приводим (9.19) к виду

$$\begin{aligned} W &= \frac{1}{12\pi^2\epsilon_0} \frac{e^2}{c^3} \omega_0^4 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\omega) F(\xi) \delta(\omega+\xi) d\omega d\xi = \\ &= \frac{1}{12\pi^2\epsilon_0} \frac{e^2}{c^3} \omega_0^4 \int_{-\infty}^{\infty} F(\omega) F(-\omega) d\omega = \\ &= \frac{1}{6\pi^2\epsilon_0} \frac{e^2}{c^3} \omega_0^4 \int_0^{\infty} F(\omega) F(-\omega) d\omega = \int_0^{\infty} w(\omega) d\omega, \end{aligned} \quad (9.21)$$

где учтено, что $F(\omega)F(-\omega)$ является четной функцией ω , и введено обозначение

$$w(\omega) = \frac{1}{6\pi^2\epsilon_0} \frac{e^2}{c^3} \omega_0^4 F(\omega) F(-\omega). \quad (9.22)$$

Функция $w(\omega)$ характеризует распределение энергии по частотам, являясь плотностью энергии волн, приходящейся на частоту ω . Подставляя выражение (9.14) в (9.18), получаем

$$\begin{aligned} F(\omega) &= \int_0^{\infty} A_0 \exp \{[-\gamma/2 + i(\omega_0 - \omega)]t\} dt + \\ &+ \int_0^{\infty} A_0^* \exp \{[-\gamma/2 - i(\omega_0 + \omega)]t\} dt = \\ &= \frac{A_0}{\gamma/2 - i(\omega_0 - \omega)} + \frac{A_0^*}{\gamma/2 + i(\omega_0 + \omega)}. \end{aligned} \quad (9.23)$$

В общем виде выражение для $F(\omega)F(-\omega)$, входящее в формулу (9.22), получается громоздким. Однако при условии $\gamma \ll \omega_0$ слагаемые в правой части (9.23) сильно различаются по своему значению на частотах $\omega \approx \omega_0$. Слагаемое, в которое входит разность $\omega_0 - \omega$, много больше слагаемого, в которое входит сумма $\omega_0 + \omega$. Поэтому для $F(\omega)$ достаточно сохранить лишь первое слагаемое с A_0 , а для $F(-\omega)$ — лишь второе слагаемое с A_0^* и можно написать

$$\begin{aligned} F(\omega) F(-\omega) &= \frac{A_0}{\gamma/2 - i(\omega_0 - \omega)} \frac{A_0^*}{\gamma/2 + i(\omega_0 + \omega)} = \\ &= \frac{A_0 A_0^*}{(\omega_0 - \omega)^2 + (\gamma/2)^2}. \end{aligned} \quad (9.24)$$

- Зависит ли от сдвига начала отсчета времени спектр:
- а) амплитудный? б) фазовый?
- Что такое комплексный спектр функций? Как он связан с амплитудным и фазовым спектрами?

Тогда формула (9.22) принимает вид

$$w(\omega) = \frac{1}{6\pi^2\epsilon_0} \frac{e^2}{c^3} \omega_0^4 \frac{A_0 A_0^*}{(\omega_0 - \omega)^2 + (\gamma/2)^2} \quad (9.25)$$

Максимум интенсивности излучения лежит вблизи частоты $\omega = \omega_0$, а при удалении от нее интенсивность излучения быстро уменьшается. На частотах

$$\omega_2 = \omega_0 + \gamma/2, \quad \omega_1 = \omega_0 - \gamma/2, \quad (9.26)$$

объемная плотность излучения в два раза меньше объемной плотности излучения на частоте ω_0 . Это означает, что основная часть энергии излучения приходится на интервал частот

$$\delta\omega = \omega_2 - \omega_1 = \gamma, \quad (9.27)$$

называемый шириной линии излучения.

Лоренцева форма и ширина линии излучения. Формулу (9.22) удобно представить в виде

$$w = 2mA_n A_n^* \omega_0^2 F_L(\omega), \quad (9.28)$$

где

$$F_L(\omega) = \frac{\gamma/(2\pi)}{(\omega_0 - \omega)^2 + (\gamma/2)^2} \quad (9.29)$$

— нормированная функция лоренцевой формы линии излучения (рис. 44). Функция (9.29) удовлетворяет условию нормировки

$$\int_{-\infty}^{\infty} F_L(\omega) d\omega = 1. \quad (9.30 \text{ а})$$

Главный вклад в интеграл дает область интегрирования вблизи $\omega \approx \omega_0$. Вклад в интеграл от области значений $\omega < 0$ имеет порядок величины членов, отброшенных при вычислении $F(\omega) F(-\omega)$ по формуле (9.24), исходя из (9.23). Поэтому с той же точностью, с какой $F(\omega) F(-\omega)$ выражается формулой (9.24), условие нормировки (9.30а) имеет вид

$$\int_0^{\infty} F_L(\omega) d\omega = 1. \quad (9.30 \text{ б})$$

Лоренцева форма линии излучения образуется при «естественных» условиях излучения, когда единственным фактором, влияющим на излучение осциллятора, является радиационное затухание. Поэтому эта форма линии часто называется естественной формой линии излучения, а ширина линии излучения — γ -естественной шириной.

Естественная ширина линии излучения очень мала. Для оптических частот $\gamma/\omega_0 \sim 10^{-7}$. Если при анализе некоторого вопроса изменение частоты примерно на 1/10 000 000 от частоты ω_0 не имеет значения, то можно считать, что все излучение приходится на частоту ω_0 , и рассматривать излучение как монохроматическое. Такое приближение справедливо в большинстве случаев, рассматриваемых в оптике. Лишь в некоторых ситуациях приходится принимать во внимание конечность естественной ширины линии излучения.

Время излучения. Как видно из (9.14), амплитуда колебания электрона убывает в $e = 2,7$ раза в течение времени

$$\tau = 2/\gamma = 2/(\delta\omega),$$

(9.31)

которое называется временем излучения. Видно, что чем меньше время излучения, тем больше ширина линии излучения. Стого монокроматическое излучение возможно лишь при бесконечно большом времени излучения.

Форма линии поглощения. Уравнение движения упругосвязанного электрона, находящегося под действием электрического поля E , имеет вид

$$m\ddot{x} + m\gamma\dot{x} + m\omega_0^2x = eE, \quad (9.32)$$

где предполагается, что напряженность E направлена по оси X . Если

$$E(t) = E_0 e^{i\omega t}, \quad (9.33)$$

то в установившемся режиме $x(t)$ является гармонической функцией с той же частотой ω :

$$x(t) = x_0 e^{i\omega t}. \quad (9.34)$$

Поскольку $\dot{x} = i\omega x$, $\ddot{x} = -\omega^2 x$, уравнение (9.32) принимает вид

$$m(-\omega^2 + i\omega\gamma + \omega_0^2)x = eE, \quad (9.35)$$

откуда

$$x = \frac{e}{m} \frac{E}{\omega_0^2 - \omega^2 + i\omega\gamma}. \quad (9.36)$$

Колеблющийся электрон является источником излучения, причем энергия ему передается падающей волной. Следовательно, происходит поглощение энергии упругосвязанным электроном. Интенсивность поглощения пропорциональна квадрату амплитуды (9.36) колебаний электрона, т. е. xx^* . Поэтому если на упругосвязанные электроны падает излучение с непрерывным спектром частот, то возникает линия поглощения, форма которой задается выражением

$$\Phi(\omega) = \text{const } xx^* = \frac{\text{const}}{(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2}. \quad (9.37)$$

При $\omega_0 \ll \gamma$ эта функция отлична от нуля лишь вблизи частоты ω_0 , когда можно знаменатель в (9.37) преобразовать следующим образом:

$$(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2 = (\omega_0 - \omega)^2(\omega_0 + \omega)^2 + \omega^2\gamma^2 \approx 4\omega_0^2(\omega_0 - \omega)^2 + \omega_0^2\gamma^2 = 4\omega_0^2[(\omega_0 - \omega)^2 + (\gamma/2)^2]. \quad (9.38)$$

Тогда (9.37) принимает вид

$$\Phi(\omega) = \frac{\text{const}}{(\omega_0 - \omega)^2 + (\gamma/2)^2}. \quad (9.39)$$

Следовательно, линия поглощения имеет лоренцеву форму с той же шириной, что и линия излучения.

Квантовая интерпретация формы линии излучения. В квантовой картине излучение осуществляется посредством испускания фотонов. Частота ω фотона связана с его энергией E соотношением

$$E = \hbar\omega. \quad (9.40)$$

Испускание фотона происходит в результате перехода излучающего атома из состояния атома с энергией E_n в состояние с энергией E_m , причем $E = E_n - E_m$. Следовательно, излучаемая частота

$$\omega_{nm} = (E_n - E_m) / \hbar \quad (9.41)$$

имеет вполне определенное значение. Однако при квантовом переходе излучаются фотоны различных частот, потому что энергия возбужденного состояния E_n не имеет вполне определенного значения. Другими словами, возбужденный уровень энергии имеет конечную ширину. Время излучения в квантовой картине интерпретируется как время нахождения атома в возбужденном состоянии, а соотношение (9.31) сводится к утверждению, что ширина энергетического уровня обратно пропорциональна времени жизни электрона на этом уровне. В основном состоянии с наименьшей энергией электрон может находиться бесконечно долго, если атом изолирован, и, следовательно, ширина основного энергетического уровня равна нулю. Если линия излучения образуется за счет переходов электрона из некоторого возбужденного состояния в основное, то лоренцева форма линии является отражением распределения электронов по энергиям в пределах возбужденного энергетического уровня. Если атом находится в поле излучения или взаимодействует с другими атомами, то он может перейти из основного состояния в возбужденное. Следовательно, его время жизни в основном состоянии конечно и поэтому основной уровень не бесконечно узок, а имеет конечную ширину. Это приводит к соответствующему уширению линии излучения.

Квазимохроматическая волна. Если ширина линии излучения достаточно мала, то волну можно представить в виде (2.39), понимая под ω частоту в центре линии, т. е. считать волну монохроматической с частотой, равной средней частоте немонохроматической волны. Это выполняется при условии

$$\Delta\omega \ll \omega, \quad (9.42)$$

называемом условием квазимохроматичности.

С квазимохроматической волной можно обращаться как с монохроматической в течение промежутков времени, меньших времени когерентности (см. § 13, 26, 30). Если же при рассмотрении некоторого явления необходимо принять во внимание изменения в волне в течение промежутков времени, больших времени когерентности, то ее представление в виде монохроматической волны вида (2.39) становится невозможным.

- Естественная ширина линии может быть реализована только в излучении неподвижного и изолированного от окружения атома. Естественная ширина линии излучения в классической интерпретации обусловлена конечной продолжительностью времени излучения.
- Естественная ширина линии излучения в квантовой интерпретации обусловлена конечной шириной энергетических уровней.

- Как между собой связаны формы линии излучения и поглощения? Как форма линии поглощения связана с резонансной амплитудой характеристикой линейного осциллятора с затуханием? Почему нельзя присвоить отдельному фотону спектр частот, соответствующий естественной форме линии излучения?

Дается классификация уширения спектральных линий и описываются доплеровское и ударное уширение.

Причины уширения. Естественная форма линии излучения возникает в идеальных условиях, когда излучающий атом покойится и не подвергается в процессе излучения действию каких-либо внешних сил. Идеальные условия полностью никогда не реализуются, потому что всегда имеется тепловое движение атома и взаимодействие с другими атомами. Эти обстоятельства, вообще говоря, влияют на форму линии излучения. Поэтому в реальных условиях естественная форма линии обычно не наблюдается.

Однородное и неоднородное уширение. Факторы, приводящие к уширению линий излучения, можно разделить на две группы. Одна группа вызывает в излучении каждого атома одинаковое изменение линии излучения. Такое уширение линий излучения называется однородным.

Другая группа факторов вызывает в излучении отдельных атомов различные изменения линий излучения. В этом случае уширение линии, наблюдаемое в излучении совокупности атомов, происходит за счет различных изменений линий излучения отдельных атомов. Такое уширение линий излучения называется неоднородным.

Естественная ширина линии излучения как однородное уширение. Примером такого уширения является естественная ширина линий излучения, поскольку это уширение одинаково в излучении всех атомов данного сорта и определяется лишь временем излучения. Наблюданная форма линии от излучения совокупности атомов совпадает с формой линии от излучения отдельного атома.

Ударное уширение. В газе при комнатной температуре и нормальном атмосферном давлении время, в течение которого излучение атома не нарушается взаимодействиями, имеет порядок $\tau_0 \sim 10^{-11}$ с. Если учесть, что время τ естественного излучения в (9.31) имеет порядок 10^{-8} с, то в процессе излучения атом испытывает свыше сотни нарушений процесса излучения. При каждом таком нарушении происходит изменение режима излучения, как бы его прерывание. После этого излучение продолжается с прежней частотой ω . Это означает, что фаза испускаемой волны в момент нарушения режима излучения изменяется случайным образом. Весь процесс излучения разбивается как бы на отдельные акты излучения, продолжительность которых равна промежуткам времени между последовательными нарушениями режима излучения. Это приводит к уширению линии излучения. Поскольку время между нарушениями режима излучения примерно в 100 раз меньше всей продолжительности излучения, ширина линий оказывается в рассматриваемых условиях примерно в 100 раз больше естественной ширины линии излучения. Это уширение называется ударным, поскольку является следствием соударений атомов.

Для вычисления ударного уширения примем естественную ширину линии излучения равной нулю. Ввиду случайного характера соударений время между последовательными столкновениями подчиняется распределению Пуассона, т. е. вероятность того, что время между последовательными столкновениями заключено между ξ и $\xi + d\xi$, равна

$$p(\xi) d\xi = (1/\tau_0) \exp(-\xi/\tau_0) d\xi, \quad (10.1)$$

где τ_0 — среднее время между столкновениями. Между столкновениями затухание амплитуды колебаний электрона и, следовательно, амплитуд напряженности испускаемой волны невелико. Поэтому в процессе излучения между столкновениями в моменты t_0 и $t_0 + \xi$ амплитуда напряженности электрического поля испускаемой волны примерно постоянна:

70 $E(t) = E_0 \exp(i\omega_0 t + i\varphi)$ ($t_0 < t < t_0 + \xi$), (10.2)

где φ — случайная фаза.

2 Спектральный состав напряженности электрического поля волны в рассматриваемый промежуток времени дается формулой

$$E_\omega = \int_{-\infty}^{\infty} E(t) e^{-i\omega t} dt = \int_{t_0}^{t_0 + \xi} E_0 \exp(i\omega_0 t - i\omega t + i\varphi) dt = \\ = E_0 \exp[i(\omega_0 - \omega)t_0 + i\varphi] \frac{\exp[i(\omega_0 - \omega)\xi] - 1}{i(\omega_0 - \omega)}. \quad (10.3)$$

Отсюда следует, что спектральный состав мощности излучения $w(\omega)$, усредненной по периоду колебаний (рис. 45), дается соотношением

$$w_1(\omega) \sim E_\omega E_\omega^* = 4E_0^2 \frac{\sin^2[(\omega_0 - \omega)\xi/2]}{(\omega_0 - \omega)^2}. \quad (10.4)$$

Формула (10.4) описывает интенсивность излучения в течение промежутка ξ отдельного атома. Одновременно другие атомы также излучают, причем их времена свободного пробега распределены в соответствии с (10.1). Поэтому для нахождения спектрального состава полной интенсивности излучения от всех атомов надо (10.4) усреднить по ξ с учетом (10.1):

$$w(\omega) \sim \frac{1}{\tau_0} \int_0^\infty \frac{\sin^2[(\omega_0 - \omega)\xi/2]}{(\omega_0 - \omega)^2} e^{-\xi/\tau_0} d\xi = \frac{1}{2} \frac{1}{(\omega_0 - \omega)^2 + (1/\tau_0)^2}. \quad (10.5)$$

Это означает, что ударное уширение приводит к лоренцевой форме линии с шириной $\gamma_{yw} = 2/\tau_0$.

Ударное уширение является однородным, потому что оно характеризуется средним промежутком времени τ между столкновениями, одинаковыми для всех атомов, и нет экспериментального способа отнести излучение определенной частоты к какой-то определенной группе атомов.

Хотя ударная ширина при нормальных условиях на два порядка больше естественной, ее абсолютное значение во многих случаях пренебрежимо мало и можно считать, что речь идет об излучении одной частоты ω_0 .

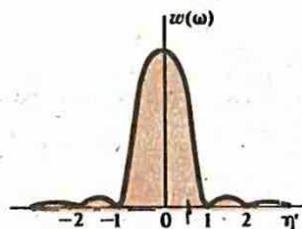
Доплеровское уширение. Если в процессе излучения атом движется, то излучаемая им частота не равна частоте излучения покоящегося атома. Законы сохранения энергии и импульса атома при излучении имеют вид

$$E_2 + \frac{1}{2} M v_2^2 = E_1 + \frac{1}{2} M v_1^2 + \hbar\omega, \quad (10.6)$$

$$M v_2 = M v_1 + \hbar k, \quad (10.7)$$

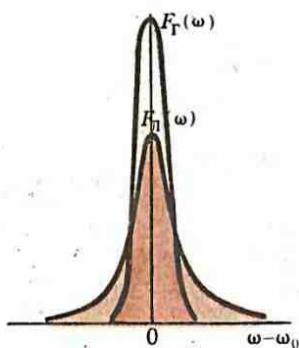
где E_2 — энергия возбужденного состояния атома; E_1 — энергия нижнего уровня, на который электрон совершает переход при излучении; M — масса атома; v_2 и v_1 — скорости электрона до и после акта излучения; $\hbar\omega$ и $\hbar k$ — энергия и импульс излученного фотона. Обозначая ω_0 частоту излучения атомом, который до акта и после акта излучения поконится, имеем

$$\hbar\omega_0 = E_2 - E_1. \quad (10.8)$$



45

Спектральный состав излучения между последовательными столкновениями [$\eta' = (\omega - \omega_0)\tau/(2r)$]



46

Лоренцева и гауссова линии излучения, нормированные на единицу и имеющие одинаковую ширину

Возводя (10.7) в квадрат и исключая из (10.6) — (10.8) величины E_1 , E_2 , v_1 , приходим к уравнению

$$\hbar(\omega_0 - \omega) = \hbar^2 k^2 / (2M) - \hbar v_{2z} \cdot \mathbf{k}. \quad (10.9)$$

Направляя ось Z декартовой системы координат вдоль волнового вектора k излученного фотона и принимая во внимание, что $k = \omega/c$, перепишем (10.9) в виде

$$\omega_0 - \omega = \hbar \omega^2 / (2Mc^2) - \omega v_{2z} / c. \quad (10.10)$$

Скорости атомов нерелятивистские и при комнатных температурах $v_{2z}/c \sim 10^{-5}$, а для $\omega \approx 10^{15} \text{ с}^{-1}$ и типичных масс атомов значение $\hbar \omega^2 / (2Mc^2) \approx 10^{-9}$. Поэтому уменьшаемым в правой части (10.10) можно пренебречь по сравнению с вычитаемым и написать:

$$\omega = \frac{\omega_0}{1 - v_{2z}/c} \approx \omega_0 \left(1 + \frac{v_{2z}}{c}\right). \quad (10.11)$$

Формула (10.11) выражает эффект Доплера: в направлении скорости атома частота испущенного фотона увеличивается на $\Delta\omega = \omega v/c$, а при излучении в противоположном направлении — на столько же уменьшается.

Доля атомов, скорости которых заключены между v_{2z} и $v_{2z} + dv_{2z}$, в соответствии с распределением Максвелла пропорциональна

$$dv_{2z} \exp[-Mv_{2z}^2 / (2kT)], \quad (10.12)$$

где k — постоянная Больцмана. Из (10.11) имеем

$$v_{2z} = (\omega - \omega_0) c / \omega_0, \quad dv_{2z} = (c/\omega_0) d\omega. \quad (10.13)$$

Тогда из (10.12) получается следующее распределение энергии излучения по частотам:

$$w(\omega)d\omega \sim \exp[-Mc^2(\omega - \omega_0)^2 / (2\omega_0^2 kT)]d\omega. \quad (10.14)$$

Описываемая этой функцией форма линии излучения называется гауссовой. Максимум интенсивности в гауссовой линии излучения приходится на частоту $\omega = \omega_0$. Шириной Δ гауссовой линии называется расстояние между частотами, соответствующими половине максимальной интенсивности. Эти частоты ω_1 и ω_2 находятся как корни уравнения

$$1/2 = \exp[-Mc^2(\omega - \omega_0)^2 / (2\omega_0^2 kT)]. \quad (10.15)$$

Отсюда следует, что

$$\Delta = 2\omega_0 [2kT \ln 2 / (Mc^2)]^{1/2}. \quad (10.16)$$

Гауссово распределение удобно представить в виде

$$F_{\Gamma}(\omega) = [1/(\sigma \sqrt{2\pi})] \exp [-(\omega - \omega_0)^2/(2\sigma^2)], \quad (10.17)$$

где

$$\sigma = [kT\omega_0^2/(Mc^2)]^{1/2} = \Delta/(2\sqrt{2\ln 2}). \quad (10.18)$$

При таком представлении оно нормировано на единицу:

$$\int_{-\infty}^{\infty} F_{\Gamma}(\omega) d\omega = 1. \quad (10.19)$$

На рис. 46 представлены лоренцева (9.29) и гауссова (10.17) линии, имеющие одинаковую ширину $\gamma = \Delta$. Площадь под этими кривыми равна единице.

При комнатной температуре и нормальном атмосферном давлении доплеровская ширина линии больше естественной ширины примерно на два порядка. Она равна при этих условиях по порядку величины ширине линии за счет ударного уширения.

Форма составной линии излучения. При одновременном действии нескольких факторов уширения линии излучения результирующая линия связана с составляющими простой формулой в предположении, что центральная частота ω_0 у всех линий одна и та же. Если $F_1(\omega)$ и $F_2(\omega)$ характеризуют две линии излучения с одинаковой центральной частотой ω_0 , то форма составной линии излучения определяется так:

$$F(\omega) = \int_{-\infty}^{\infty} F_1(\xi) F_2(\omega + \omega_0 - \xi) d\xi, \quad (10.20)$$

причем центральная частота составной линии та же ω_0 , что и центральная частота слагающих.

Лоренцева и гауссова линии обладают важной особенностью. При сложении лоренцевых линий с шириной γ_1 и γ_2 получается лоренцева линия с шириной $\gamma = \gamma_1 + \gamma_2$. При сложении гауссовых линий с шириной Δ_1 и Δ_2 получается гауссова линия с шириной $\Delta = \sqrt{\Delta_1^2 + \Delta_2^2}$.

● Форма неоднородно уширенной линии излучения совокупности атомов не совпадает с формой неоднородно уширенных линий излучения отдельных атомов.

При ударном уширении форма линии лоренцева, при доплеровском — гауссова.

○ Почему ударное уширение однородное?

Почему доплеровское уширение неоднородное?

Каковы порядки величин доплеровского и ударного уширений линий при нормальных условиях в газе?

Равна ли частота столкновений атомов, приводящая к ударному уширению линий, газодинамической частоте столкновений? Почему? Каков смысл частоты столкновений в этих случаях?

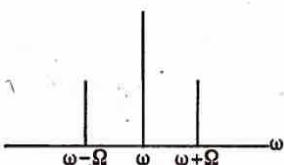
§ 11 | Модулированные волны

Описывается амплитудная, частотная и фазовая модуляции волн.



47

Колебание с гармонически модулированной амплитудой



48

Спектр колебаний с гармонически модулированной амплитудой

Модуляция. Гармоническое колебание, описывающее волну, характеризуется амплитудой, частотой и фазой. Изменение этих параметров в процессе колебания называется модуляцией, а волны, получающиеся в процессе модуляции, называются модулированными.

Модуляция амплитуды. Колебание с модулированной амплитудой может быть представлено в виде

$$f(t) = [A_0 + a(t)] \cos \omega t, \quad (11.1)$$

где $a(t)$ описывает модуляцию; ω — частота гармонического колебания; $|a(t)| < A_0$.

Прежде всего рассмотрим случай, когда $a(t)$ является гармонической функцией

$$a(t) = a_0 \cos \Omega t \quad (11.2)$$

с частотой $\Omega \ll \omega$. Тогда равенство (11.1) принимает вид

$$f(t) = (A_0 + a_0 \cos \Omega t) \cos \omega t. \quad (11.3)$$

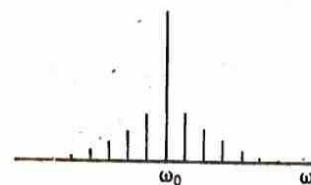
График этого колебания изображен на рис. 47. С помощью формул для косинуса суммы и разности углов выражение (11.3) преобразуется к виду

$$f(t) = A_0 \cos \omega t + \frac{1}{2}a_0 \cos(\omega - \Omega)t + \frac{1}{2}a_0 \cos(\omega + \Omega)t, \quad (11.4)$$

из которого можно заключить, что спектральный состав колебания сводится к трем частотам: ω , $\omega - \Omega$, $\omega + \Omega$ (рис. 48). Частота ω называется несущей, а частоты $\omega \pm \Omega$ называются боковыми.

Если $a(t)$ является не гармонической, но периодической функцией с периодом $T = 2\pi/\Omega$, то ее можно представить в виде ряда Фурье по частотам, кратным Ω . Подставив ряд Фурье в формулу (11.1) и преобразовав каждый из членов ряда после умножения на $\cos \omega t$, аналогично тому, как это было сделано при переходе от (11.3) к (11.4), получим ряд, в который входят частоты ω и $\omega \pm n\Omega$ ($n = 1, 2, \dots$), т. е. спектр состоит из набора частот, отстоящих друг от друга на Ω в обе стороны от несущей частоты ω_0 (рис. 49). Ширина спектра определяется шириной спектра функции $a(t)$. Если $a(t)$ является непериодической функцией, которая представляется интегралом Фурье, то ее спектр непрерывный. Подставляя в этом случае в (11.1) выражение для $a(t)$ в виде интеграла Фурье и преобразовывая гармонические составляющие аналогично предыдущим случаям, получим для модулированного колебания непрерывный спектр, простирающийся в обе стороны от несущей частоты ω_0 (рис. 50). Таким образом, и в этом случае ширина спектра определяется шириной спектра $a(t)$.

Все изложенное справедливо для колебаний, например, электрической напряженности в фиксированной пространствен-



49

Спектр колебания с периодически модулированной амплитудой

- Частотная и фазовая модуляции эквивалентны только тогда, когда они гармонические. При негармонической модуляции структура сигналов этих модуляций совершенно различна.

ной точке. Если скорость распространения волн не зависит от частоты (среда без дисперсии), то колебание без изменения формы переносится в другие точки со скоростью распространения волны и поэтому форма колебания по времени в данной точке легко пересчитывается на пространственную форму колебаний (см. § 5).

Модуляция частоты и фазы. Эти два вида модуляции целесообразно рассматривать совместно, поскольку описывающие их формулы тесно связаны друг с другом, хотя структура сигналов, модулированных по частоте и фазе, существенно различна.

Соотношение между частотной и фазовой модуляциями получается как следствие записи фазы колебания через зависящую от времени (модулированную) частоту по формуле

$$\sin \Phi(t) = \sin \left[\int_0^t \omega(t) dt \right], \quad (11.5)$$

где $\Phi(t)$ — фаза колебания, $\omega(t)$ — модулированная частота. Начальная фаза колебаний считается равной нулю. Из равенства

$$\Phi(t) = \int_0^t \omega(t) dt \quad (11.6)$$

следует, что мгновенное значение частоты $\omega(t)$ связано с фазой $\Phi(t)$ соотношением

$$\omega(t) = d\Phi(t)/dt. \quad (11.7)$$

Формулы (11.6) и (11.7) позволяют от формул, описывающих частотную модуляцию, перейти к формулам, описывающим модуляцию по фазе.

Рассмотрим гармоническую модуляцию частоты. В этом случае имеем

$$\omega(t) = \omega_0 + \Delta\omega \cos \Omega t, \quad (11.8)$$

где ω_0 — постоянная частота, около которой происходят колебания частоты с амплитудой $\Delta\omega$ и частотой колебаний Ω . В соответствии с (11.6) имеем

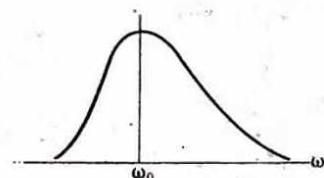
$$\Phi(t) = \int_0^t \omega(t) dt = \omega_0 t + (\Delta\omega/\Omega) \sin \Omega t. \quad (11.9)$$

Это означает, что фаза $\Phi(t)$ также модулирована по гармоническому закону с той же частотой Ω и амплитудой модуляции $\Delta\omega/\Omega$. Если фаза модулирована по гармоническому закону $\Phi(t) = \omega_0 t + \Delta\Phi \sin \Omega t,$

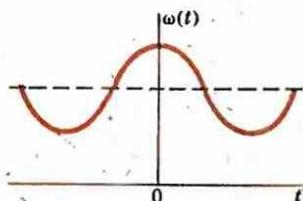
то в результате дифференцирования (11.10) по времени с учетом (11.7) приходим к формуле

$$\omega(t) = d\Phi/dt = \omega_0 + \Delta\Phi \Omega \cos \Omega t, \quad (11.11)$$

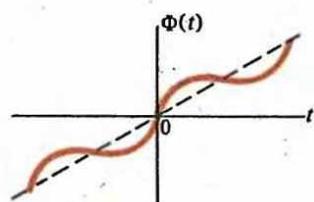
показывающей, что частота при этом оказывается модулированной также по гармоническому закону с той же частотой Ω и амплитудой $\Delta\Phi\Omega$. Гармоническая модуляция частоты $\omega(t)$ и фазы $\Phi(t)$ показана на рис. 51, 52.



Сплошной спектр колебания с не-периодически модулированной ам-плитудой



Гармоническая модуляция частоты



Гармоническая модуляция фазы

- Чем отличаются способы осущ-ществления частотной и фазовой модуляций?
- Как зависит пропускная способность линии связи от не-сущей частоты?

Таким образом, частотная и фазовая модуляции полностью эквивалентны только тогда, когда они гармонические. При негармонической модуляции такая эквивалентность невозможна, структура сигналов, модулированных по частоте и по фазе, оказывается совершенно различной. При частотной модуляции медленным изменениям сигнала (т. е. низким частотам в сигнале) соответствуют большие колебания по фазе [$\Delta\Phi_{\max} = \Delta\omega/\Omega$ в (11.9)], а быстрым изменениям — малые. При фазовой модуляции, наоборот, медленным изменениям сигнала соответствуют малые амплитуды колебаний частоты ($\Delta\omega = \Delta\Phi\Omega$), а быстрым изменениям — большие. Частотная и фазовая модуляции отличаются также по способу осуществления. При частотной модуляции образуется прямое воздействие на частоту колебаний генератора, при фазовой частоте колебаний генератора постоянна, а фаза модулируется при движении сигнала после генератора.

Спектр колебания с гармонической модуляцией частоты. Рассмотрим спектральный состав частотно-модулированного сигнала с гармоническим законом модуляции

$$E(t) = E_0 \sin [\omega_0 t + (\Delta\omega/\Omega) \sin \Omega t]. \quad (11.12)$$

Считая, что амплитуда модуляции мала ($\Delta\omega/\Omega \ll 1$), разложим (11.12) в ряд Тейлора по $(\Delta\omega/\Omega) \sin \Omega t$ и ограничимся членами первого порядка:

$$E(t) \approx E_0 [\sin \omega_0 t + (\Delta\omega/\Omega) \sin \Omega t \cos \omega_0 t] = E_0 [\sin \omega_0 t + (\Delta\omega/(2\Omega)) \sin (\omega_0 + \Omega)t - - - (\Delta\omega/(2\Omega))] \sin (\omega_0 - \Omega)t. \quad (11.13)$$

Таким образом, в спектре в первом приближении присутствуют лишь частоты ω_0 и $\omega_0 + \Omega$, $\omega_0 - \Omega$, т. е. он аналогичен спектру сигнала, модулированного по амплитуде с той же частотой. Однако такое соответствие справедливо лишь при малых глубинах модуляции. При увеличении $\Delta\omega/\Omega$ существенную роль начинают играть и другие составляющие спектра в частотно-модулированном сигнале. Поэтому, вообще говоря, сигнал, частота которого модулирована по гармоническому закону, содержит в своем спектре бесконечное число частот и этим принципиально отличается от амплитудно-модулированного по гармоническому закону сигнала. Частотная модуляция отличается от амплитудной также и тем, что при амплитудной модуляции связь между спектром сигнала и спектром модулированного колебания линейна, а при частотной модуляции — нелинейна. При амплитудной модуляции добавление новой частоты в спектр сигнала добавляет соответствующую частоту в спектр модулированного колебания, не изменяя амплитуд остальных частот. При частотной модуляции добавление новой частоты приводит не только к добавлению в спектр модулированного колебания многих новых частот, но и к изменению амплитуды существующих.

Спектр колебания (11.12) при произвольных значениях $\Delta\omega/\Omega$ выражается посредством функции Бесселя $J_n(x)$ с целым индексом n и здесь не рассматривается.

§ 12 | Волновые пакеты

Дается характеристика волновых пакетов и квазиплоской волны.

Волновой пакет, образованный двумя волнами. Электромагнитные волны распространяются со скоростью света независимо от частоты только в вакууме. В среде скорость электромагнитной волны меньше скорости света и зависит от частоты. Зависимость скорости волны от частоты называется дисперсией.

Рассмотрим суперпозицию двух волн, частоты которых ω_1 и ω_2 , а волновые числа k_1 и k_2 : $E_1(z, t) = E_0 \cos(\omega_1 t - k_1 z)$, $E_2(z, t) = E_0 \cos(\omega_2 t - k_2 z)$, (12.1)

считая, что они имеют одинаковую поляризацию и распространяются в одном направлении. Фазовая скорость волны определяется из условия

$$\omega t - kz = \text{const.} \quad (12.2)$$

Дифференцируя (12.2) по t , получаем

$$v_\phi = v = dz/dt = \omega/k$$

(12.3)

(индекс « ϕ » у фазовой скорости для упрощения в дальнейшем указывать не будем). Фазовую скорость в вакууме обозначим c . Фазовые скорости волн в (12.1), вообще говоря, могут быть и различными. Напряженность образовавшейся в результате суперпозиции волны описывается формулой

$$E(z, t) = E_1 + E_2 = 2E_0 \cos [(\omega_1 - \omega_2)t/2 - (k_1 - k_2)z/2] \times \\ \times \cos [(\omega_1 + \omega_2)t/2 - (k_1 + k_2)z/2]. \quad (12.4)$$

Форма такой волны показана на рис. 10. Если дисперсия отсутствует, то напряженность имеет вид (4.7). Волна без изменения формы распространяется со скоростью света в направлении положительных значений оси Z , причем огибающая амплитуд, обозначенная на рис. 10 пунктирной линией, движется со скоростью света.

Групповая скорость. Суперпозиция двух или большего числа волн с различными частотами составляет группу волн, или волновой пакет. Скоростью группы волн или групповой скоростью называется скорость движения максимума огибающей амплитуды группы волн. Из условия постоянства фазы огибающей амплитуды волны (12.4), записанного в виде

$$\frac{1}{2}(\omega_1 - \omega_2)t - \frac{1}{2}(k_1 - k_2)z = \text{const}, \quad (12.5)$$

после дифференцирования по t находим групповую скорость:

$$v_r = dz/dt = (\omega_1 - \omega_2)/(k_1 - k_2). \quad (12.6)$$

Если дисперсия отсутствует, то $\omega_1 = ck_1$, $\omega_2 = ck_2$ и из (12.6) получаем $v_r = c$, т. е. групповая скорость совпадает с фазовой. При наличии дисперсии групповая скорость отличается от фазовой. В результате огибающая амплитуд и слагаемые волны движутся с различными скоростями, что приводит к изменению формы огибающей в процессе распространения волны, т. е. при наличии дисперсии волновой пакет распространяется с изменением формы.

Если частоты слагаемых волн близки друг к другу ($\omega_2 \rightarrow \omega_1$), то для групповой скорости из (12.6) получается формула

$$v_r = d\omega/dk. \quad (12.7)$$

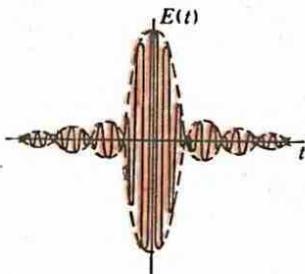
Она справедлива не только для двух волн с бесконечно близкими частотами, но и для произвольного волнового пакета, образованного суперпозицией бесконечного числа волн с близкими частотами, поскольку является дифференциальной.

Суперпозиция колебаний с эквидистантными частотами. Пусть происходит N колебаний одннаковой амплитуды E_0 , частота которых различается на $\delta\omega$. Результат суперпозиции этих колебаний представляется формулой

$$E(t) = E_0 \cos \omega t + E_0 \cos (\omega + \delta\omega)t + E_0 \cos (\omega + 2\delta\omega)t + \dots + E_0 \cos [\omega + (N-1)\delta\omega]t. \quad (12.8)$$

Суммирование этого ряда проще произвести в экспоненциальном представлении гармонических функций:

$$E(t) = \operatorname{Re} \left[E_0 \sum_{n=0}^{N-1} e^{i(\omega t + n\delta\omega t)} \right] = E_0 \operatorname{Re} \left[e^{i\omega t} \sum_{n=0}^{N-1} e^{in\delta\omega t} \right] = E_0 \operatorname{Re} \left[e^{i\omega t} \frac{1 - e^{iN\delta\omega t}}{1 - e^{i\delta\omega t}} \right] = \\ = E_0 \operatorname{Re} \left[e^{i\omega t} \frac{e^{iN\delta\omega t/2} (e^{-iN\delta\omega t/2} - e^{iN\delta\omega t/2})}{e^{i\delta\omega t/2} (e^{-i\delta\omega t/2} - e^{i\delta\omega t/2})} \right] = E_0 \operatorname{Re} \left[e^{i[\omega t + (N-1)\delta\omega t/2]} \frac{\sin(N\delta\omega t/2)}{\sin(\delta\omega t/2)} \right] =$$



53

Форма волнового пакета с эквидистантными частотами

$$= E_0 \frac{\sin(N\delta\omega t/2)}{\sin(\delta\omega t/2)} \cos <\omega>t, \quad (12.9)$$

где $<\omega> = \omega + (N - 1)\delta\omega/2$ — средняя частота волнового пакета. Принимая во внимание, что $N\delta\omega = \Delta\omega$ — полная ширина частот волнового пакета, выражение (12.9) можно представить в виде

$$E(t) = E_0 \frac{\sin(\Delta\omega t/2)}{\sin[\Delta\omega t/(2N)]} \cos(<\omega>t). \quad (12.10)$$

В большинстве случаев, представляющих практический интерес, $N \gg 1$ и поэтому в течение многих периодов изменения аргумента $\Delta\omega t/2$ у синуса в числите формулы аргумент у синуса в знаменателе формулы остается малым ($\Delta\omega t/2N \ll 1$), так что можно считать $\sin[\Delta\omega t/(2N)] \approx \Delta\omega t/(2N)$. Поэтому (12.10) можно записать в виде

$$E(t) = E_0 N \frac{\sin(\Delta\omega t/2)}{(\Delta\omega t/2)} \cos(<\omega>t). \quad (12.11)$$

График этой функции приведен на рис. 53. Огибающая пунктирная кривая представляет изменяющуюся амплитуду колебаний в волновом пакете, основная частота которых $<\omega>$. Энергия такого волнового пакета сосредоточена в сравнительно небольшом интервале частот. Поэтому волновые пакеты называются также импульсами. Мы будем использовать оба эти названия в зависимости от обстоятельств. Максимальная амплитуда образуется в точке $t=0$, когда все колебания складываются в одинаковой фазе. Через промежуток времени Δt , определяемый условием

$$\Delta\omega\Delta t/2 = \pi, \quad (12.12)$$

амплитуда колебаний обращается в нуль. Это время принимается за меру длительности центрального импульса. Поэтому между частотным интервалом слагаемых колебаний $\Delta\nu = 2\pi\Delta\omega$ и временной продолжительностью импульса существует соотношение

$$\Delta\nu\Delta t \approx 1, \quad (12.13)$$

где использован знак приблизительного равенства, что учитывает произвольность в определении продолжительности импульса. Такое соотношение уже было получено [см.(8.56)] при анализе спектрального состава прямоугольного импульса. Ввиду универсальности соотношения (12.13) его часто называют теоремой о ширине частотной полосы.

Квазиплоская волна. Плоской волной, представленной формулой вида (2.50), может быть лишь пространственно не ограниченная во всех направлениях волна. Ограничение волны в направлении распространения приводит к ее немонохроматичности, характеризуемой шириной спектра частот $\Delta\omega$ (см. § 9).

- Групповой скоростью называется скорость максимума амплитуды группы волн. С этой скоростью движется энергия волнового пакета. При наличии дисперсии групповая скорость отличается от фазовой.

Соотношение между шириной линии излучения и продолжительностью импульса может быть представлено в виде соотношения между шириной волнового пакета и его пространственной протяженностью заменой $\omega \rightarrow k, t \rightarrow z$.

- В чем заключается теорема о ширине частотной полосы? Какова ширина частотной полосы импульса, представляющего δ -функцией?

Ограничение волны в перпендикулярных направлениях приводит к возникновению конечной ширины спектра волновых чисел Δk_x и Δk_y (волна распространяется вдоль оси Z). Другими словами, волна с конечным поперечным сечением пучка не может распространяться строго в одном направлении, характеризуемом вектором k ; имеется некоторый разброс направлений волновых векторов от среднего направления. Это явление называется дифракцией (см. гл. 6). Следовательно, плоских волн с конечным поперечным сечением не существует. Однако если разброс направлений волновых векторов невелик, волна с большой точностью может считаться плоской и быть представленной в форме (2.50), где под k понимается средний волновой вектор волны. Такая волна называется квазиплоской.

Для получения математического критерия квазиплоской волны поступаем так же, как и для установления критерия (9.42) квазимохроматической волны. Волну представляем в виде интеграла Фурье:

$$E(x, y, z, t) = \frac{1}{(2\pi)^4} \iiint F(k_x, k_y, k_z, \omega) e^{-i(\omega t - \mathbf{k} \cdot \mathbf{r})} dk_x dk_y dk_z d\omega. \quad (12.14)$$

Если линейные размеры поперечного сечения волны, распространяющейся в направлении оси Z , велики по сравнению с длиной волны, то амплитуда F в (12.14) отлична от нуля лишь в узком интервале волновых чисел $\Delta k_x, \Delta k_y$ вблизи значений $k_x = 0, k_y = 0$. Если

$$|2\Delta k_x| \ll k_z, \quad |2\Delta k_y| \ll k_z, \quad (12.15)$$

то волна называется квазиплоской. Ее с достаточной точностью можно представить в виде (2.50), понимая под k средний волновой вектор. Квазиплоскую волну можно считать плоской лишь на участке фронта волны, линейные размеры которого меньше ширины когерентности (см. § 27, 30). Если же при рассмотрении некоторого явления необходимо учесть изменения по фронту волны на участках, больших ширины когерентности, то ее нельзя считать плоской и представлять формулой (2.50).

§ 13 | Хаотический свет

Анализируются волны со случайными амплитудами и фазами.

Суперпозиция волн со случайными фазами. В классической картине изолированным неподвижным атомом излучается цуг волн с экспоненциально убывающей амплитудой продолжительностью порядка 10^{-8} с. Ширина линии излучения имеет порядок 10^8 Гц (см. § 9). Форма линии — лоренцева. В результате взаимодействий с другими атомами в процессе излучения происходит ударное уширение линии (см. § 10). Напряженность поля волн в некоторой точке в момент времени t дается выражением

$$E_i(t) = E_0 e^{-i[\omega(t-t'_i) - \phi'_i]}, \quad (13.1)$$

где t'_i — момент начала испускания соответствующего цуга волн после столкновения, ϕ'_i — случайная начальная фаза излучения. Ее можно объединить с членом $-\omega t'_i$, который зависит от расстояния между атомом и точкой наблюдения, являющимся также случайной величиной. Поэтому напряженность поля излучения атома в некоторой точке может быть представлена в виде

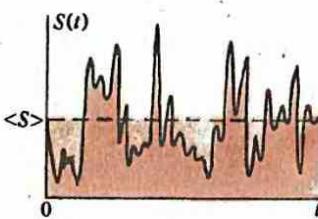
$$E_i(t) = E_0 e^{-i(\omega t - \phi_i)}, \quad (13.2)$$

где $\phi_i = \omega t'_i + \phi'_i$ — случайная фаза. Эффект Доплера, приводящий к изменению частоты, пока не принимаем во внимание. Его нетрудно учесть как дополнительный фактор уширения линий излучения. Для упрощения предположим, что все волны имеют одинаковые амплитуду и поляризацию. Суммарная напряженность от всех излучателей равна



54

Сложение мгновенных значений фазовых множителей



55

Флуктуации плотности потока энергии при усреднении по промежуткам времени, много меньшим t_0

○ Какая величина играет роль образа Фурье функции, описывающей стационарный случайный процесс?

Что такое спектр мощности? Каким образом по спектру мощности можно определить автокорреляционную функцию и интервал корреляции?

$$E(t) = E_0 \sum_i e^{-i(\omega t - \varphi_i)} = E_0 e^{-i\omega t} \sum_i e^{i\varphi_i}, \quad (13.3)$$

где φ_i — случайная фаза цуга волн, пришедших от i -го атома. Стоящий под знаком суммы ряд может быть представлен как векторное сложение комплексных чисел. Каждый из членов ряда $\exp(-i\varphi_i)$ по модулю равен единице. На рис. 54 показан результат сложения этих членов для некоторого момента времени t . Аналитически их сумму можно записать в виде

$$\sum_i e^{i\varphi_i} = a(t) e^{i\varphi(t)}, \quad (13.4)$$

где $a(t)$ — амплитуда суммы экспоненциальных слагаемых, $\varphi(t)$ — ее фаза. Таким образом,

$$E(t) = E_0 a(t) e^{-i[\omega t - \varphi(t)]} \quad (13.5)$$

где ω — несущая частота волны, у которой модулированы амплитуды и фаза. Спектральный состав этой волны такой же, как и у линии с ударным уширением. Это является следствием того, что значительные изменения фазы и амплитуды в (13.4) происходят в течение времени t_0 , равного времени между возмущениями. Поэтому временные параметры, определяющие ударное уширение и ширину спектра модулированной волны (13.5), одинакового порядка, а потому в основном одинаковы и спектральный состав излучения.

Время разрешения. Каждый прибор измеряет лишь некоторое среднее значение величины по малому промежутку времени, называемому временем разрешения. Время разрешения лучших приборов для измерения напряженности электрического поля по порядку величины равно 10^{-9} с. Поскольку время t_0 , в течение которого амплитуда $a(t)$ и фаза $\varphi(t)$ в (13.5) существенно изменяются, имеет порядок 10^{-11} , заключаем, что в течение многих десятков и даже сотен периодов колебаний волны эти величины могут считаться практически постоянными. Это означает, что при усреднении (13.5) по периоду колебаний или многим периодам $\langle E \rangle = 0$. Поэтому экспериментально можно изучать не средние величины напряженности поля волны (13.5), а средние величины от квадрата напряженности, т. е. потоки энергии волн. Результат измерения потока энергии волн в эксперименте зависит от времени разрешения прибора.

Усреднение по периоду колебаний. Хотя приборов, которые имели бы разрешение в один период световых колебаний, нет, целесообразно изучить этот вопрос теоретически, чтобы затем обсудить результаты усреднений интенсивности по более продолжительным периодам. На основании (3.4) с учетом сказанного о скорости изменения $a(t)$ и $\varphi(t)$ для средней плотности потока энергии волн, напряженность электрического поля которых дается выражением (13.5), получаем формулу

$$S = c(1/2) \epsilon_0 E_0^2 a^2(t), \quad (13.6)$$

где усреднение произведено по одному периоду или по промежутку времени, много меньшему t_0 . Знак усреднения у S не указывается для упрощения написания и не оговаривается.

что это средняя плотность. Изменение плотности потока энергии волн по времени определяется величиной $a^2(t)$.

Изучение $a^2(t)$ для столь коротких времен усреднения может быть проведено лишь теоретически методами математического моделирования. Основные особенности поведения $a^2(t)$ со временем хорошо видны на рис. 55. Пунктиром обозначены средние значения $a^2(t)$ при усреднении по промежуткам времени, большим τ_0 .

На рис. 55 видны большие флуктуации интенсивности $S(t)$, имеющие порядок средней величины интенсивности. Временной масштаб флуктуаций определяется значением τ_0 . Это является другим выражением утверждения о том, что в течение времени τ_0 происходит существенное изменение $a(t)$ и $a^2(t)$.

Влияние увеличения промежутка времени на результат усреднения. При увеличении промежутка времени усреднения кривая, изображенная на рис. 55, сглаживается: высота пиков уменьшается, и резкие изменения ослабляются. При приближении времени усреднения к τ_0 сплошная кривая приближается к пунктирной. При временах усреднения порядка τ_0 изменения плотности потока энергии волн полностью исчезают. Значит, все эти изменения происходят в промежутки времени, меньшие τ_0 , а τ_0 является масштабом флуктуаций. Дальнейшее увеличение промежутка времени усреднения не изменяет среднего значения. Таким образом, время τ_0 является характерным временем для рассматриваемого процесса.

Время когерентности. Время τ_0 является характерным масштабом случайных флуктуаций фазы и амплитуды волн светового пучка с ударным уширением линий. При других механизмах уширения линий и их комбинаций также имеются характерные времена случайных флуктуаций фазы и амплитуды светового пучка. Эти характерные времена называются временами когерентности τ_k . Они играют большую роль в явлениях интерференции (см. гл. 5).

Длина когерентности. Так называется путь l_k , проходимый световым пучком за время когерентности:

$$l_k = c\tau_k. \quad (13.7)$$

Для ударного уширения длина когерентности в тысячи раз больше длины световой волны. При других механизмах уширения линий длина когерентности также много больше длины световой волны. Длина когерентности является пространственным масштабом флуктуаций интенсивности светового пучка. От картины изменения флуктуации во времени (рис. 55) можно перейти к картине их изменения в пространстве, приняв ось t за ось Z с положительным направлением этой оси в сторону отрицательных значений t .

Гауссов свет. С математической точки зрения случайный процесс сложения фазовых множителей (см. рис. 54) является задачей на случайное блуждание, рассматриваемой, например, в теории броуновского движения. В данном случае средняя длина шага равна единице. Обозначим p плотность вероятности того, что конечная точка случайного блуждания характеризуется полярными координатами (a, φ) после n шагов. Тогда вероятность того, что после n шагов конечная точка находится на площадке $dad\varphi$, дается формулой

$$dP = padad\varphi. \quad (13.8)$$

Теория случайного блуждания для p дает выражение

$$p = p(a) = (1/\pi n) \exp(-a^2/n) \quad (13.9)$$

с нормировкой

$$\int dP = \int_0^\infty da \int_0^{2\pi} p(a)d\varphi = 2\pi \int_0^\infty ap(a)da = 1. \quad (13.10)$$

Поскольку $dad\varphi$ — элемент площади, формула (13.8) показывает, что амплитуды $a(t)$ распределены по гауссову закону (13.9). Это распределение справедливо для любого источника хаотического излучения. Поэтому свет от хаотического источника называют часто гауссовым.

Флуктуации плотности потока энергии хаотического света. При усреднении по промежутку времени, большему времени когерентности τ_k (в случае ударного уширения $\tau_k = \tau_0$), получаем

$$\langle a^2(t) \rangle = \left\langle \left| \sum_i e^{i\varphi_i(t)} \right|^2 \right\rangle = n, \quad (13.11)$$

так как среднее от перекрестных членов равно нулю; $\langle |e^{i\varphi_i(t)}|^2 \rangle = 1$; n — число атомов, излучение которых наблюдается. При усреднении по очень большому в сравнении с периодом колебаний промежутку времени плотность потока S [см.: (13.6)] играет роль мгновенного значения плотности потока энергии волн, а усредненная в соответствии с (13.11) — средней плотности. Средняя плотность потока энергии волн на основе (13.6) с учетом (13.11) равна

$$\langle S \rangle = 1/2c\varepsilon_0 E_0^2 n. \quad (13.12)$$

Из сравнения (13.12) и (13.6) заключаем, что

$$a^2(t) = Sn / \langle S \rangle. \quad (13.13)$$

Тогда формула (13.8) принимает вид

$$dP = \frac{1}{\pi n} \exp\left(-\frac{a^2}{n}\right) adad\varphi = \frac{1}{2\pi} \frac{1}{\langle S \rangle} \exp\left(-\frac{S}{\langle S \rangle}\right) dS d\varphi. \quad (13.14)$$

Отсюда следует, что распределение плотности вероятностей S дается формулой

$$p'(S) = (1/\langle S \rangle) \exp(-S/\langle S \rangle) \quad (13.15)$$

с нормировкой

$$\int_0^\infty p'(S) dS = 1. \quad (13.16)$$

Величина флуктуаций плотности потока энергии может быть вычислена с помощью (13.14) и характеризуется среднеквадратичным отклонением от равновесного значения:

$$\sqrt{\langle (\Delta S)^2 \rangle} = \sqrt{\langle (S - \langle S \rangle)^2 \rangle} = \sqrt{\langle S^2 \rangle - (\langle S \rangle)^2}, \quad (13.17)$$

где $\langle (S - \langle S \rangle)^2 \rangle = \langle S^2 \rangle - (\langle S \rangle)^2$. Учитывая, что

● **Случайные флуктуации фазы и амплитуды волн светового пучка** характеризуются некоторым минимальным временным интервалом τ_k , при усреднении по которому изменения плотности потока энергии полностью сглаживаются. Это характерное время называется временем когерентности.

Амплитуды хаотического излучения распределены по закону Гаусса.

В хаотическом свете значение флуктуации интенсивности равно среднему значению интенсивности.

○ По каким промежуткам времени надо производить усреднения, чтобы найти флуктуации интенсивности и среднюю интенсивность светового потока?

Какие обстоятельства приводят к гауссову распределению амплитуд волн хаотического света?

Какой смысл имеет «мгновенное» значение интенсивности светового потока?

Учитывая, что при фиксированной частоте объемная плотность излучения пропорциональна концентрации фотонов, найдите на основании формулы (13.19) флуктуации числа фотонов, попадающих на детектор в промежуток времени фиксированной продолжительности. Какой должна быть продолжительность этого промежутка времени, чтобы утверждение о флуктуациях могло быть экспериментально проверено?

$$\frac{1}{\langle S \rangle} \int_0^{\infty} S^2 \exp \left(-\frac{S}{\langle S \rangle} \right) dS = 2(\langle S \rangle)^2, \quad (13.18)$$

находим

$$\sqrt{(\Delta S)^2} = \langle S \rangle, \quad (13.19)$$

т. е. флуктуация плотности потока энергии равна ее среднему значению. Таким образом, отклонение плотности потока энергии волны от среднего значения в световом пучке очень большое, что можно качественно видеть на рис. 56. Отметим еще раз, что «мгновенное значение» плотности потока энергии в световом пучке необходимо измерять за промежутки времени, много меньшие времени когерентности, а вся совокупность измерений должна занимать время, большее времени когерентности τ_k . При выводе соотношения (13.19) предполагалось, что мгновенное значение плотности потока энергии S является средним по периоду светового колебания. В действительности таких детекторов не существует. При учете конечного времени отклика детектора формула (13.19) немножко изменяется, поскольку величина флуктуаций несколько сглаживается.

Поляризация. Монохроматические электромагнитные волны (см. § 5) всегда обладают определенной поляризацией, т. е. являются поляризованными.

У хаотического света, представляющего собой суперпозицию волн со случайными фазами и амплитудами, конец вектора напряженности поля волн описывается в плоскости, перпендикулярной направлению распространения волн, нерегулярную, чрезвычайно хаотическую линию. Если эти колебания вектора напряженности не имеют никаких преимущественных направлений, то волна называется неполяризованной. Если же в колебаниях вектора напряженности имеется некоторая регулярность, хотя и не такая, как в поляризованной волне, то волна называется частично поляризованной. Количественная теория частичной поляризации основывается на теории когерентности взаимно перпендикулярных компонент напряженности поля волн. Поэтому целесообразно ее изложить вместе с теорией когерентности (см. § 30).

§ 14 | Фурье-анализ случайных процессов

Показывается, как можно применить Фурье-преобразование к функциям, не удовлетворяющим математическим критериям применимости Фурье-преобразования. Излагаются понятия и методы Фурье-анализа случайных процессов.

Спектр мощности. Большинство случайных процессов стационарны по времени, т. е. их общий характер с течением времени не изменяется. Это означает, что функции, описывающие эти процессы, не имеют образа Фурье, поскольку они не абсолютно интегрирумы (функция не стремится к нулю при $t \rightarrow \pm \infty$). Следовательно, применить обычные методы и понятия спектрального анализа к этим функциям нельзя. Да это и нецелесообразно, поскольку в случайных процессах интересны лишь средние характеристики, а фазовые соотношения между гармоническими составляющими в спектральном разложении не имеют значения. Кроме того, полностью не известна функциональная зависимость случайных функций от времени. Поэтому в Фурье-анализе случайных процессов используются более подходящие для этих целей величины и понятия.

Чтобы исключить проблему расходимости интеграла, определим $F_T(\omega)$ — образ Фурье случайного стационарного процесса $f(t)$ на конечном, достаточно большом промежутке времени T :

$$F_T(\omega) = \int_{-T/2}^{T/2} f(t) e^{-i\omega t} dt. \quad (14.1)$$

Как показывает (9.22), распределение энергии по частотам в спектре излучения характеризуется квадратом модуля спектральной плотности $|F(\omega)|^2 = F(\omega)F^*(\omega)$. В однородном процессе распределение энергии по частотам с течением времени не изменяется, а полная энергия излучения пропорциональна времени. Отсюда можно заключить, что

$$\int_{-\infty}^{\infty} |F_T(\omega)|^2 d\omega$$

растет пропорционально T и, следовательно, $|F_T(\omega)| \sim \sqrt{T}$. Поэтому величина

$$F_c(\omega) = \lim_{T \rightarrow \infty} \frac{F_T(\omega)}{\sqrt{T}} = \lim_{T \rightarrow \infty} \frac{1}{\sqrt{T}} \int_{-T/2}^{T/2} f(t) e^{-i\omega t} dt \quad (14.2)$$

конечна, не зависит от времени и может быть использована для характеристики спектральных свойств случайных процессов. Квадрат модуля этой величины

$$w_c(\omega) = F_c(\omega) F_c^*(\omega) \quad (14.3)$$

называется спектром мощности. С помощью (14.2) получаем

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} w_c(\omega) d\omega &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} dt f(t) \int_{-T/2}^{T/2} dt' f(t') - \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp[-i\omega(t-t')] d\omega = \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} dt f(t) \int_{-T/2}^{T/2} dt' f(t') \delta(t-t') = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} |f(t)|^2 dt = \langle |f(t)|^2 \rangle. \end{aligned}$$

Таким образом, спектр мощности связан со средним квадратом модуля случайной функции соотношением

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} w_c(\omega) d\omega = \langle |f|^2 \rangle_t. \quad (14.4)$$

В спектре мощности отсутствуют фазовые характеристики процесса, и поэтому он содержит меньше информации, чем амплитудно-фазовый спектр. Спектр мощности содержит характеристики процесса, усредненные по большому интервалу времени. Поэтому из него выпадают мелкомасштабные характеристики процесса.

Автокорреляционная функция. Важнейшей характеристикой случайного процесса является взаимосвязь значений описывающей этот процесс функции в различных пространственных и временных точках. Взаимосвязь между значениями функции в одной точке в различные моменты времени описывается автокорреляционной функцией. Рассмотрим для примера стационарную случайную величину. Нас интересует связь значений этой функции в момент времени t , т. е. $f_1 = f(t)$, с ее значением в момент времени $t + \tau$, т. е. $f_2 = f(t + \tau)$, где τ — некоторая фиксированная величина. Другими словами, надо изучить связь значений этой функции в моменты времени, разделенные интервалом τ . Сами по себе моменты времени произвольны из-за стационарности процесса, важно лишь, чтобы между ними был интервал τ . Из-за случайного характера $f(t)$ при каждом измерении («испытании») значения $f(t)$ и $f(t + \tau)$ будут находиться в случайном соотношении друг с другом. Поэтому взаимосвязь значений f в моменты времени, разделенные интервалом τ , может характеризоваться лишь статистически.

Простейшей статистической характеристикой случайной величины является ее среднее значение

$$\langle f(t) \rangle_a = \langle f_1 \rangle, \quad \langle f(t + \tau) \rangle_a = \langle f_2 \rangle, \quad (14.5)$$

где индекс «а» у угловых скобок означает среднее по ансамблю или, иначе, математическое ожидание по множеству реализаций от $f(t)$ и $f(t+\tau)$. Для стационарного процесса, очевидно, это среднее не зависит от t , т. е. $\langle f_1 \rangle = \langle f_2 \rangle = \langle f \rangle$.

Взаимосвязь значений f в моменты времени, разделенные интервалом τ , описывается автокорреляционной функцией, определяемой формулой

$$\Gamma_{11}(\tau) = \langle [f(t) - \langle f \rangle] [f(t+\tau) - \langle f \rangle] \rangle_a. \quad (14.6)$$

Если $p(f_1, f_2)$ является плотностью вероятности того, что $f(t)$ принимает значение f_1 , а $f(t+\tau)$ — значение f_2 , то (14.6) представляется в виде

$$\Gamma_{11}(\tau) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (f_1 - \langle f \rangle) (f_2 - \langle f \rangle) p(f_1, f_2) df_1 df_2. \quad (14.7a)$$

Если процесс является эргодическим, то усреднение по ансамблю может быть заменено усреднением по времени и вместо (14.7) можно написать

$$\begin{aligned} \Gamma_{11}(\tau) &= \langle [f(t) - \langle f \rangle] [f(t+\tau) - \langle f \rangle] \rangle_t = \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} [f(t) - \langle f \rangle] [f(t+\tau) - \langle f \rangle] dt, \end{aligned} \quad (14.7b)$$

где $\langle f \rangle$ — среднее значение f по времени.

Для упрощения формул рассмотрим в качестве случайной величины функцию $f(t) - \langle f \rangle$, среднее значение которой равно нулю. Чтобы не вводить дополнительных символов, обозначим ее также $f(t)$. Формула (14.7b) для такой функции принимает вид

$$\Gamma_{11}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} f(t) f(t+\tau) dt, \quad (14.8)$$

т. е. совпадает с (14.7b) при $\langle f \rangle = 0$. Все результаты, следующие из формулы (14.8) для $\langle f \rangle = 0$, без труда переносятся на общий случай (14.7b). В дальнейшем, если не будет оговорено противное, используется определение (14.8), т. е. предполагается, что $\langle f \rangle = 0$.

Автокорреляционная функция $\Gamma_{11}(\tau)$ при $\tau = 0$ равна среднему значению квадрата $f(t)$, т. е. некоторой положительной величине. При небольших значениях τ она продолжает сохранять положительное, отличное от нуля значение, хотя и уменьшается с увеличением τ . Про область значений τ , при которых $\Gamma_{11}(\tau) \neq 0$, говорят, что в ней корреляция имеет конечную величину. При увеличении τ корреляция ослабевает, т. е. всегда $\Gamma_{11}(\tau) < \Gamma_{11}(0)$. Из (14.8) имеем

$$\Gamma_{11}(-\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} f(t) f(t-\tau) dt = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\tau/2}^{T/2-\tau} f(t') f(t') dt' = \Gamma_{11}(\tau), \quad (14.9)$$

где переход от первого интеграла ко второму осуществлен заменой переменных $t = t' + \tau$. Таким образом, **автокорреляционная функция симметрична относительно нуля**:

$$\Gamma_{11}(-\tau) = \Gamma_{11}(\tau). \quad (14.10)$$

Если значения функции $f(t)$ и $f(t+\tau)$ не связаны друг с другом, т. е. независимы друг от друга, то автокорреляционная функция равна нулю. В частности, при $\tau \rightarrow \infty$ она стремится к нулю.

Теорема Винера — Хинчина. Важное значение автокорреляционной функции обуславливается, в частности, ее связью со спектром мощности, которая устанавливается теоремой Винера — Хинчина: **спектр мощности является образом Фурье автокорреляционной функции, а автокорреляционная функция является образом Фурье спектра мощности.**

Для доказательства преобразуем выражение (14.3) с учетом (14.2):

$$w_c(\omega) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} f(t') e^{-i\omega t'} dt' \int_{-T/2}^{T/2} f(t) e^{i\omega t} dt, \quad (14.11)$$

где учтено, что функция f действительна ($f^* = f$). Для дальнейших преобразований заметим, что значения этих интегралов при стационарной функции не зависят от пределов интегрирования, лишь бы интервал интегрирования был T . Поэтому в первом интеграле можно заменить переменную интегрирования $t' = t + \tau$ и интегрировать по τ ($dt' = d\tau$), оставив без изменения пределы интегрирования:

$$w_c(\omega) = \lim_{T \rightarrow \infty} \int_{-T/2}^{T/2} dt e^{-i\omega\tau} \frac{1}{T} \int_{-T/2}^{T/2} f(t) f(t + \tau) dt. \quad (14.12)$$

Внутренний интеграл в (14.12) на основании (14.8) равен $\Gamma_{11}(\tau)$. Поэтому (14.12) превращается в соотношение

$$w_c(\omega) = \int_{-\infty}^{\infty} \Gamma_{11}(\tau) e^{-i\omega\tau} d\tau, \quad (14.13)$$

которое доказывает первую часть теоремы Винера — Хинчина. Вторая часть теоремы есть просто обратное преобразование (14.13):

$$\Gamma_{11}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} w_c(\omega) e^{i\omega\tau} d\omega. \quad (14.14)$$

Взяв в (14.2) комплексно-сопряженные величины от обеих частей равенства и принимая во внимание, что для действительной функции f выполняется равенство $f = f^*$, заключаем, что $F_c^*(\omega) = F_c(-\omega)$ и поэтому (14.3) является четной функцией ω , т. е. $w_c(\omega) = w_c(-\omega)$. Корреляционная функция $\Gamma_{11}(\tau)$ согласно (14.10) также является четной функцией τ . Поэтому, представив в подынтегральных выражениях (14.13) и (14.14) $\exp(\pm i\omega\tau) = \cos \omega\tau \pm i \sin \omega\tau$, приходим к выводу, что интегралы, содержащие синус, исчезают и соотношения принимают вещественную форму:

$$w_c(\omega) = 2 \int_0^{\infty} \Gamma_{11}(\tau) \cos \omega\tau d\tau, \quad (14.15) \qquad \Gamma_{11}(\tau) = \frac{1}{\pi} \int_0^{\infty} w_c(\omega) \cos \omega\tau d\omega. \quad (14.16)$$

Теорема Винера — Хинчина позволяет находить спектр мощности, если известна автокорреляционная функция, которая может быть экспериментально измерена. Обычно она представляет собой быстро затухающую функцию, благодаря чему вычисление интеграла (14.15) не составляет труда. Тем самым спектр мощности оказывается измеренным экспериментально.

Интервал корреляции. Из (14.8) видно, что

Функции, описывающие стационарные случайные процессы, не имеют образов Фурье. Для них определяется специальная величина, играющая роль образа Фурье. Квадрат модуля этой величины называется спектром мощности. В спектре мощности отсутствуют фазовые характеристики случайного процесса. Спектр мощности содержит характеристики случайного процесса, усредненные по большому интервалу времени. Поэтому малконастабильные характеристики процесса в спектре мощности отсутствуют. Спектр мощности является образом Фурье автокорреляционной функции, и наоборот. Экспериментально определив спектр мощности, можно найти автокорреляционную функцию. Интервал корреляции равен значению нормированного спектра мощности при нулевой частоте.

$$\Gamma_{11}(0) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} f^2(t) dt = \langle f^2 \rangle_t. \quad (14.17)$$

Поэтому вместо корреляционной функции $\Gamma_{11}(\tau)$ удобно пользоваться безразмерной функцией

$$\gamma_{11}(\tau) = \Gamma_{11}(\tau) / \Gamma_{11}(0) = \Gamma_{11}(\tau) / \langle f^2 \rangle, \quad (14.18)$$

получившей название **нормированной корреляционной функции**. Из (14.18) видно, что $\gamma_{11}(0) = 1$. Мерой корреляции является величина

$$\tau_x = \int_{-\infty}^{\infty} \gamma_{11}(\tau) d\tau = 2 \int_0^{\infty} \gamma_{11}(\tau) d\tau \quad (14.19)$$

— интервал корреляции.

Связь интервала корреляции с нормированным спектром мощности. По аналогии с нормированной корреляционной функцией вводится нормированный спектр мощности

$$w_n(\omega) = w_c(\omega) / \langle f^2 \rangle = w_c(\omega) / \Gamma_{11}(0). \quad (14.20)$$

Между интервалом корреляции и нормированным спектром мощности существует важное соотношение. Для его вывода выразим в (14.19) $\gamma_{11}(\tau)$ через нормированный спектр мощности по формуле (14.14):

$$\tau_x = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w_n(\omega) e^{i\omega\tau} d\omega d\tau = \int_{-\infty}^{\infty} w_n(\omega) \delta(\omega) d\omega = w_n(0). \quad (14.21)$$

Таким образом, интервал корреляции равен значению нормированного спектра мощности при нулевой частоте.

Задачи

- 2.1. Найти длину когерентности излучения рубинового лазера ($\lambda = 693,6$ нм), если ширина линии излучения в длинах волн равна $\Delta\lambda = 1,6 \cdot 10^{-17}$ м.
- 2.2. Записать ряд Фурье для периодической функции с периодом T , которая в интервале $(-T/4, +T/4)$ равна двум, а вне интервала до его границ — нулю.
- 2.3. Записать ряд Фурье для периодической функции с периодом T , которая в интервале $(0, T/2)$ равна двум, а в интервале $(T/2, T)$ — нулю.
- 2.4. Записать ряд Фурье для периодической функции с периодом T , заданной на участке $(0, T)$ формулой $f(t) = t/T$.
- 2.5. Записать ряд Фурье для периодической функции с периодом T , заданной формулой $f(t) = \begin{cases} -2t/T & \text{при } -T/2 < t < 0, \\ 2t/T & \text{при } 0 < t < T/2. \end{cases}$
- 2.6. Найти Фурье-образ функции $f(t) = \exp(-at^2) \cos \omega_0 t$ ($a > 0$).
- 2.7. Найти Фурье-образ функции $f(z) = 4 \exp(-z^2/a^2)$.
- 2.8. Проведите расчет, доказывающий, что составная линия из двух лоренцевых линий является лоренцевой с шириной $\gamma = \gamma_1 + \gamma_2$, а составная линия из двух гауссовых линий является гауссовой с шириной $\Delta = \sqrt{\Delta_1^2 + \Delta_2^2}$.

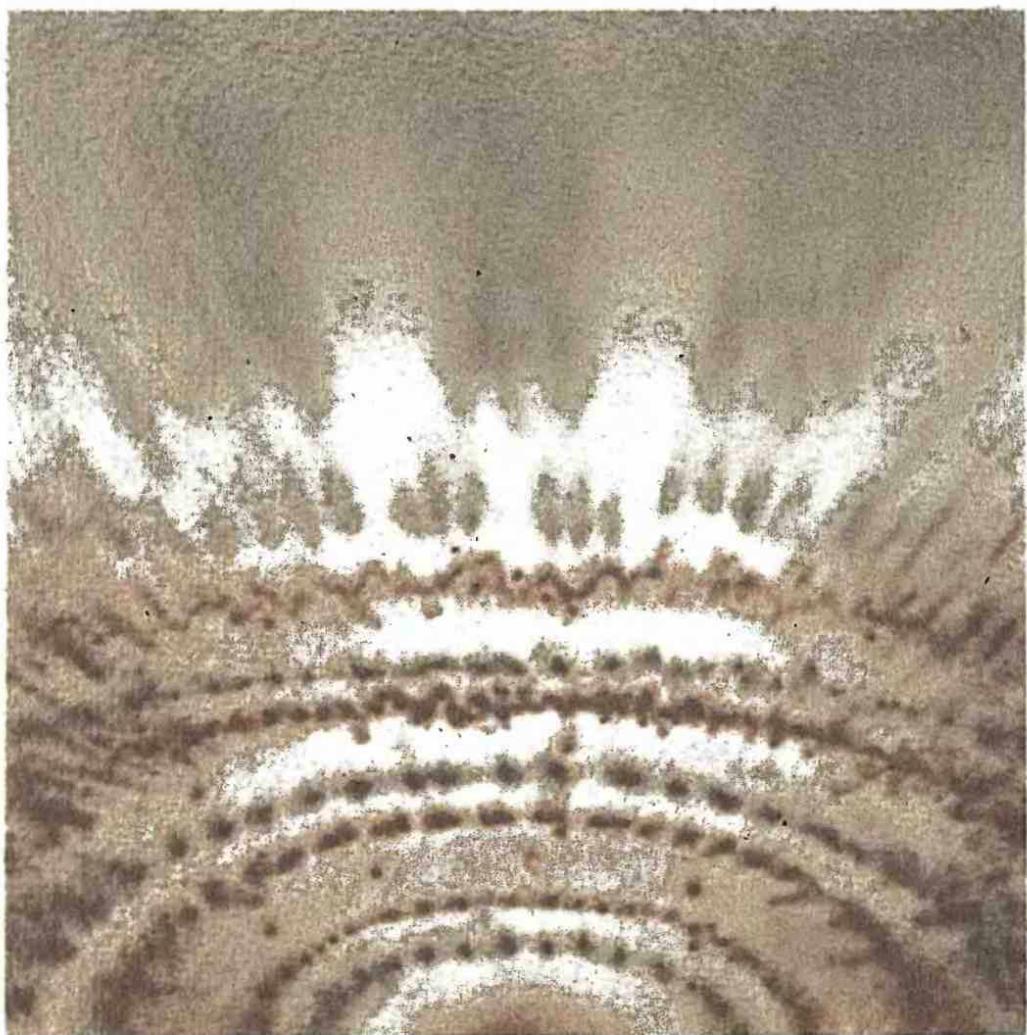
Ответы

- 2.1. $3 \cdot 10^4$ м.
- 2.2. $1 + (4/\pi) [\cos \omega t - 1/3 \cos 3\omega t + 1/5 \cos 5\omega t - \dots]$.
- 2.3. $1 + (4/\pi) (\sin \omega t + 1/3 \sin 3\omega t + 1/5 \sin 5\omega t + \dots)$.
- 2.4. $f(t) = 1/2 - (1/\pi) (\sin \omega t + 1/2 \sin 2\omega t + \dots)$.
- 2.5. $f(t) = 1/2 - (4/\pi^2) [\cos \omega t + (1/3^2) \cos 3\omega t + (1/5^2) \cos 5\omega t + \dots]$.
- 2.6. $\sqrt{\pi/a} \exp[-(\omega - \omega_0)^2/(4a)]$.
- 2.7. $aA \sqrt{\pi} \exp[-(k^2 a^2/2)]$.

Основная идея:
свойства среды
описываются скалярными
величинами —
дизэлектрической и
магнитной проницаемостями,
электропроводимостью.

Поведение света
на границе между
различными средами
определяется
граничными условиями.

Распространение света в изотропных средах



Описывается микроскопический механизм возникновения дисперсии и ее проявления. Изучается распространение импульсов в диспергирующей среде.

Монохроматические волны. В однородных изотропных диэлектриках диэлектрическая проницаемость ϵ не зависит от координат. Кроме того, считается, что она не зависит также и от времени. В этом случае уравнения Максвелла аналогичны уравнениям (2.1) — (2.5), но с заменой $\epsilon_0 \rightarrow \epsilon$, т. е. изменяется лишь первое уравнение (2.5):

$$\mathbf{D} = \epsilon \mathbf{E}. \quad (15.1)$$

Поэтому все дальнейшие результаты § 2 для электромагнитных волн в вакууме справедливы для диэлектрика, но с заменой $\epsilon_0 \rightarrow \epsilon$. Это приводит лишь к изменению скорости волн. Из уравнений (2.8) и (2.9) с заменой $\epsilon_0 \rightarrow \epsilon$ для скорости электромагнитных волн в диэлектрике получаем выражение

$$v = 1/\sqrt{\epsilon\mu_0} = 1/\sqrt{\epsilon_r \epsilon_0 \mu_0} = c/\sqrt{\epsilon_r} = c/n, \quad (15.2)$$

где $\epsilon_r = \epsilon/\epsilon_0$ — относительная диэлектрическая проницаемость, $1/\sqrt{\epsilon_0 \mu_0} = c$ — скорость света, $n = \sqrt{\epsilon_r}$ — коэффициент преломления диэлектрика относительно вакуума или просто коэффициент (показатель) преломления.

Электромагнитные волны в диэлектриках аналогичны волнам в вакууме (см. § 2). Нет необходимости повторять сказанное. Отметим лишь, что длина волны λ связана с частотой ω соотношением (2.39), но с заменой $c \rightarrow v$:

$$\lambda = vT = 2\pi v/\omega, \quad (15.3)$$

а волновое число дается выражением

$$k = 2\pi/\lambda = \omega/v. \quad (15.4)$$

Соответственно изменяется и выражение для волнового вектора.

Структура электромагнитной плоской волны, описываемая в диэлектрике уравнениями (2.53) — (2.56) с заменой $\epsilon_0 \rightarrow \epsilon$, аналогична вакууму, а соотношение (2.57) принимает вид

$$\mathbf{E} = v\mathbf{B}. \quad (15.5)$$

Плотность потока энергии волн в диэлектриках дается формулой (3.1). Для $|\mathbf{S}|$ вместо (3.2) находим

$$S = |\mathbf{S}| = |\mathbf{E}||\mathbf{H}| = EB/\mu_0 = E^2/(\mu_0 v), \quad (15.6)$$

где учтено (15.5). Поскольку в диэлектрике $1/\mu_0 = \epsilon v^2$ (см. 15.2) принимает вид

$$S = v\epsilon E^2. \quad (15.7)$$

В диэлектрике объемная плотность энергии электромагнитного поля выражается формулой $w = 1/2 (\mathbf{E} \cdot \mathbf{D} + \mathbf{B} \cdot \mathbf{H}) = \epsilon E^2$,

где $\mathbf{B} \cdot \mathbf{H} = E^2/(\mu_0 v^2) = \epsilon E^2$. Поэтому интерпретация соотношения (15.7) весьма наглядна, если его записать в виде

$$S = vw. \quad (15.9)$$

Для среднего по времени значения плотности потока энергии волн вместо (3.4) получим $\langle S \rangle = v\epsilon E_0^2/2$,

где E_0 — амплитуда напряженности электрического поля волны.

Все сказанное в §3 о распределении плотности потока энергии по сечению пучка, о мощности потока энергии и других аналогичных вопросах сохраняет свое значение для однородного диэлектрика и не нуждается в повторении.

Дисперсия. В диэлектрике скорость электромагнитных волн зависит от частоты. Это явление называется дисперсией. Влияние дисперсии проявляется лишь в распространении немонокроматических волн, поскольку различные частоты, составляющие волну, распространяются с различной скоростью.

Дисперсия является следствием зависимости поляризованности атомов от частоты. В электрическом поле волны с напряженностью

$$E = E_0 e^{-i\omega t} \quad (15.11)$$

движение электрона описывается уравнением (9.32). Решение этого уравнения дается формулой (9.36). Дипольный момент атома, электрон которого сместился из положения равновесия ($x = 0$) в точку x , равен

$$p = ex = e^2 E / [m(\omega_0^2 - \omega^2 - i\gamma\omega)]. \quad (15.12)$$

Поэтому зависящая от частоты и изменяющаяся по времени поляризованность равна

$$P = Np = e^2 NE / [m(\omega_0^2 - \omega^2 - i\gamma\omega)], \quad (15.13a)$$

где N — концентрация электронов с собственной частотой колебаний ω_0 . Запишем (15.13a) в виде

$$P = \epsilon_0 \chi^{(1)} E, \quad (15.13b)$$

$$\text{где } \chi^{(1)} = e^2 N / [m\epsilon_0(\omega_0^2 - \omega^2 - i\gamma\omega)] \quad (15.13c)$$

— комплексная диэлектрическая восприимчивость (линейная). Если имеются электроны с другой собственной частотой колебаний, то необходимо добавить соответствующий член в правую часть (15.13). Принимая во внимание соотношение

$$D = \epsilon_0 E + P \quad (15.14)$$

и учитывая, что $D = \epsilon_0 E$, а P дается равенством (15.13a), из (15.14) после сокращения на общий множитель E находим

$$\epsilon_{\omega} = \epsilon_0 + e^2 N / [m(\omega_0^2 - \omega^2 - i\gamma\omega)], \quad (15.15)$$

где ϵ_{ω} — диэлектрическая проницаемость, зависящая от частоты. Из (15.15) следует, что

$$\epsilon_{r\omega} = \epsilon_{\omega} / \epsilon_0 = 1 + e^2 N / [m\epsilon_0(\omega_0^2 - \omega^2 - i\gamma\omega)], \quad (15.16)$$

где $\epsilon_{r\omega}$ — относительная диэлектрическая проницаемость, зависящая от частоты. Как видно из (15.2), это означает, что и коэффициент преломления $n'_{\omega} = \sqrt{\epsilon_{r\omega}}$ а следовательно, и скорость электромагнитных волн зависят от частоты. Этим объяснен механизм возникновения дисперсии.

На основании (15.16) для коэффициента преломления можно написать соотношение

$$n'^2 = \epsilon_{r\omega} = 1 + e^2 N / [m\epsilon_0(\omega_0^2 - \omega^2 - i\gamma\omega)], \quad (15.17)$$

из которого следует, что n'_{ω} является комплексной величиной. Представим ее в виде

$$n'_{\omega} = n_{\omega} + i\xi_{\omega}. \quad (15.18)$$

где n_{ω} и ξ_{ω} — соответственно действительная и мнимая части n'_{ω} . Подставляя (15.18) в (15.17) и приравнивая между собой действительные и мнимые части полученного равенства, находим два уравнения:

$$n_{\omega}^2 - \xi_{\omega}^2 = 1 + \frac{e^2 N}{\epsilon_0 m} \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2}, \quad (15.19)$$

$$2n_{\omega}\xi_{\omega} = \frac{e^2 N}{\epsilon_0 m} \frac{\omega\gamma}{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2}. \quad (15.20)$$

Поскольку γ очень мала [см. (9.7)], в оптической области, за исключением частот $\omega \approx \omega_0$, $\gamma^2 \omega^2 \ll (\omega_0^2 - \omega^2)^2$ (15.21)

и можно считать, что $\xi_{\omega} = 0$. Тогда формула (15.19) приобретает вид

$$n_{\omega}^2 = 1 + e^2 N / [\epsilon_0 m (\omega_0^2 - \omega^2)]. \quad (15.22)$$

Если кроме электронов с собственной частотой колебаний ω_0 имеются электроны с другими частотами собственных колебаний, то в правую часть (15.22) необходимо добавить соответствующие этим частотам члены. Обозначая N_i концентрацию электронов в диэлектрике, собственная частота колебаний которых ω_{0i} , получаем

$$n_{\omega}^2 = 1 + \frac{e^2}{\epsilon_0 m} \sum_i \frac{N_i}{\omega_{0i}^2 - \omega^2}. \quad (15.23)$$

Однако дисперсия создается не только в результате колебаний электронов. Надо принять во внимание также и колебания ионов. Ввиду большой массы ионов частоты их собственных колебаний значительно меньше частот собственных колебаний электронов. Колебания ионов учитываются дополнением правой части (15.23) соответствующими слагаемыми. Например, наличие однозарядных ионов с собственной частотой ω_{0j} и концентрацией N_j дает дополнительный член $e^2 N_j / [\epsilon_0 M_j (\omega_{0j}^2 - \omega^2)]$, в котором M_j — масса иона. Чтобы не усложнять формулы, эти члены в (15.23) не выписываются. Собственные частоты ионов лежат в далекой инфракрасной области и не оказывают существенного влияния на ход дисперсионной кривой в видимой области спектра. Однако они играют главную роль в объяснении отлиния значения статической диэлектрической проницаемости от значения диэлектрической проницаемости в видимом диапазоне частот. Формула (15.23) нуждается также еще в одном уточнении: Под действием электромагнитного поля не все электроны колеблются одинаково, часть электронов не колеблется вообще и т. д. Следовательно, не все N_i электронов дают вклад в дисперсию, а лишь их доля. Поэтому в уточненной формуле в (15.23) вместо N_i должна стоять величина $N_i f_i$, в которой $f_i \ll 1$ называется силой осциллятора. Только с учетом силы осциллятора формула (15.23) дает удовлетворительное согласие с экспериментом. Теоретически сила осциллятора может быть рассчитана только в рамках квантовой теории. Как в (15.23), так и в последующих формулах сила осциллятора в явном виде не учитывается.

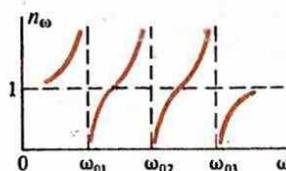
Иллюстрацией вклада ионов в дисперсию может служить вода. У нее для оптических частот $\sqrt{\epsilon_r} = n_{\omega} = 1,33$, а статическое значение $\sqrt{\epsilon_r} \approx 9$. Это объясняется колебаниями ионов. При учете колебаний ионов в правой части (15.23) при $\omega \rightarrow 0$ появляются члены вида $e^2 N_j f_j / (\epsilon_0 M_j \omega_{0j}^2)$, имеющие большие числовые значения, поскольку собственные частоты ω_{0j} колебаний ионов малы. Вблизи минимальной частоты ω_{0j} дисперсионная кривая поднимается, а при $\omega \rightarrow 0$ несколько снижается, но все же пересекает ось $\omega = 0$ при $\epsilon_r \approx 81$, чем и объясняется большое статическое значение относительной диэлектрической проницаемости воды.

Нормальная дисперсия. Если показатель преломления близок к единице, например для разреженных газов, то $(n_{\omega}^2 - 1) = (n_{\omega} - 1)(n_{\omega} + 1) \approx 2(n_{\omega} - 1)$ и формула (15.23) упрощается:

$$n_{\omega} = 1 + \frac{e^2}{2\epsilon_0 m} \sum_i \frac{N_i}{\omega_{0i}^2 - \omega^2}. \quad (15.24)$$

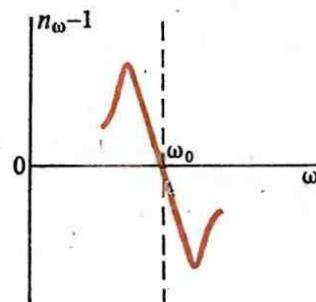
Графическая зависимость n_{ω} от частоты ω называется дисперсионной кривой (рис. 56). Если коэффициент преломления растет с частотой, то дисперсия называется нормальной. На рисунке изображены участки дисперсионной кривой, соответствующие нормальной дисперсии. Можно сказать, что нормальная дисперсия наблюдается во всей прозрачной области.

Для малых частот $\omega \ll \omega_{0i}$ формула (15.24) дает статическое значение показателя преломления



56

Нормальная дисперсия



57

Аномальная дисперсия

- Нормальная дисперсия наблюдается во всей прозрачной области. Аномальная дисперсия наблюдается в областях поглощения. Форма импульса при распространении в диспергирующей среде изменяется. Передний фронт импульса в диспергирующей среде движется со скоростью света в вакууме. Лишь последующие части движутся со скоростью, характерной для среды. Замещение в среде первичной волны вторичной, которая распространяется со скоростью, характерной для среды, происходит на некоторой длине пути. Для воздуха этот путь равен примерно 0,5 м, а для нежезвездной среды — около двух световых лет.

$$n = 1 + \frac{e^2}{2\epsilon_0 m} \sum_l \frac{N_l}{\omega_{0l}^2}, \quad (15.25)$$

которое может существенно отличаться от значения показателя преломления для оптических частот.

Для больших частот ($\omega \gg \omega_{0l}$) показатель преломления стремится к единице, оставаясь меньше ее, поскольку в этом случае (15.24) принимает вид

$$n_\omega = 1 - \frac{e^2}{2\epsilon_0 m} \frac{1}{\omega^2} \sum_l N_l. \quad (15.26)$$

Следовательно, для коротковолнового излучения диэлектрик является оптически менее плотной средой, чем вакуум. В частности, в этом случае от поверхности диэлектрика может наблюдаться полное отражение (см. § 17), как, например, при рентгеновском излучении. Из (15.26) видно, что при очень больших частотах характер связи электронов в атомах не играет роли, а показатель преломления зависит лишь от общего числа колеблющихся электронов в единице объема.

Аномальная дисперсия. Дисперсионная кривая на рис. 56 построена в соответствии с формулой (15.23), которая получена из (15.17) без учета затухания колебаний ($\gamma = 0$). Это привело к обращению в бесконечность n_ω в точках $\omega = \omega_{0l}$. Если принять во внимание затухание ($\gamma \neq 0$), то дисперсионная кривая в окрестностях точек ω_{0l} становится непрерывной и не уходит на бесконечность при $\omega = \omega_{0l}$. Представим показатель преломления в виде (15.18). Если $|n'_\omega|$ мало отличается от единицы, то из (15.17) получаем

$$n'_\omega = n_\omega + i\xi_\omega = \sqrt{\epsilon_r \omega} \approx 1 + e^2 N / [2\epsilon_0 m (\omega_0^2 - \omega^2 - i\gamma\omega)], \quad (15.27)$$

откуда в результате разделения действительных и мнимых частей следует

$$n_\omega = 1 + \frac{e^2}{2\epsilon_0 m} N \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2}, \quad (15.28)$$

$$\xi_\omega = \frac{e^2}{2\epsilon_0 m} N \frac{\gamma\omega}{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2}. \quad (15.29)$$

Дисперсионная кривая вблизи резонансной частоты $\omega = \omega_0$ представлена на рис. 57. Вблизи резонансной частоты показатель преломления с увеличением частоты уменьшается. Это явление называется аномальной дисперсией.

Рассеяние света. Электрический вектор плоской волны, распространяющейся в направлении оси Z , представляется в виде

$$E(z, t) = E_0 e^{-i(\omega t - kz)}, \quad (15.30)$$

где

$$k = \omega/v = \omega \sqrt{\epsilon_r}/c. \quad (15.31)$$

Замена $\sqrt{\epsilon_{r\omega}}$ в (15.31) ее выражением из (15.27) приводит к соотношению

$$k = \omega n_0/c + i\omega \xi_0/c, \quad (15.32)$$

подстановка которого в выражение (15.30) дает формулу

$$E(z, t) = E_0 e^{-\omega \xi_0 z/c} e^{i(\omega t - \omega n_0 z/c)}, \quad (15.33)$$

показывающую, что мнимая часть коэффициента преломления ξ_0 описывает затухание плоской волны в диэлектрике. Оно вызвано тем, что энергия волны расходуется на возбуждение вынужденных колебаний электронов. Колеблющиеся электроны, в свою очередь, излучают электромагнитные волны той же частоты во всех направлениях. Таким образом, чистый итог всей этой цепочки процессов состоит в том, что при прохождении через диэлектрик происходит рассеяние электромагнитных волн. Ввиду малости γ это рассеяние невелико.

Распространение волнового пакета. В общем случае волновой пакет является суперпозицией группы волн, частоты и волновые числа которых заключены в достаточно узких интервалах, и может быть представлен в виде

$$E(z, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(k) e^{-i(\omega t - kz)} dk, \quad (15.34)$$

где амплитуда $F(k)$ отлична от нуля лишь в узком интервале значений вблизи k_0 . В качестве независимой переменной выбрано волновое число k , а частота является функцией от k , т. е. $\omega = \omega(k)$. В (15.34) предполагается, что поляризация всех волн группы одинакова. Амплитуда $F(k)$ характеризует распределение волн по волновым числам и, следовательно, частотам. Она может быть найдена из (15.34) при $t=0$ с помощью преобразования Фурье:

$$F(k) = \int_{-\infty}^{\infty} E(z, 0) e^{-ikz} dz, \quad (15.35)$$

т. е. $F(k)$ является образом Фурье $E(z, 0)$.

Если распределение амплитуд волнового пакета по частотам имеет резкий максимум вблизи волнового числа k_0 , то функцию $\omega(k)$ можно вблизи k_0 разложить в ряд Тейлора по отклонению k от k_0 , т. е. по $k - k_0$, и ограничиться первым членом:

$$\omega(k) \approx \omega_0 + (k - k_0) d\omega_0/dk_0, \quad (15.36)$$

где $\omega_0 = \omega(k_0)$, $d\omega_0/dk_0 = (d\omega/dk)|_{k=k_0}$. Формула (15.34) принимает вид

$$E(z, t) \approx \frac{1}{2\pi} \exp \left[-i \left(\omega_0 - k_0 \frac{d\omega_0}{dk_0} \right) t \right] \int_{-\infty}^{\infty} F(k) \exp \left[ik \left(z - \frac{d\omega_0}{dk_0} t \right) \right] dk, \quad (15.37)$$

где знак приблизительного равенства подчеркивает, что в расчете учитываются лишь члены первого порядка в разложении Тейлора для $\omega(k)$. Формула (15.34) показывает, что интеграл (15.37) равен $E[z - (d\omega_0/dk_0)t, 0]$ и поэтому [см. (15.37)]

$$E(z, t) \approx E[z - (d\omega_0/dk_0)t, 0] \exp \left\{ -i[\omega_0 - k_0(d\omega_0/dk_0)]t \right\}. \quad (15.38)$$

Отсюда видно, что квадрат амплитуды импульса

$$|E|^2 = |E[z - (d\omega_0/dk_0)t, 0]|^2. \quad (15.39)$$

Значит, в первом приближении импульс движется без изменения формы с групповой скоростью

$$v_r = d\omega_0/dk_0 \quad (15.40)$$

[см. (12.7)]. Участ членов более высокого порядка в разложении Тейлора для $\omega(k)$ показывает, что форма импульса в процессе распространения изменяется.

Все формулы, характеризующие движение импульса, были в приближении (15.36). В этом приближении может случиться, что групповая скорость в области аномальной дисперсии окажется больше скорости света. Это невозможно, поскольку с групповой скоростью движется энергия, а движение энергии и массы со скоростями, превышающими скорость света, недопу-

стимо. Разрешение противоречия состоит в том, что в области аномальной дисперсии приближение (15.36) является недостаточным и необходимо использовать следующие члены разложения.

Дисперсия возникает в результате интерференции первичной и вторичной волн. Поэтому передний фронт светового импульса распространяется в среде со скоростью света в вакууме, поскольку вторичные волны не могут его догнать. Эта часть импульса прибывает первой. Ее амплитуда мала. Затем прибывает вторая часть импульса, имеющая более значительную амплитуду и продолжительность. Затем прибывает основной сигнал. Ясно, что скорость сигнала не является точно определенным понятием, поскольку за сигнал можно было бы принять часть импульса, прибывающей в точку приема первой. Обычно, говоря о скорости сигнала, имеют в виду групповую скорость на частоте, соответствующей максимальной амплитуде в сигнале. Однако при достаточной чувствительности детектора за скорость сигнала можно было бы принять скорость предшественников основного сигнала. В этом случае скорость сигнала может быть сколь угодно близкой к скорости света в вакууме, хотя сигнал и распространяется в среде.

Замещение световой волны в среде. Входящая в диэлектрическую среду световая волна возбуждает колебание диполей, которые излучают вторичную волну. Вторичная волна распространяется в среде со скоростью, характерной для среды, а первичная волна после входа в среду продолжает распространяться с начальной скоростью. Первичная и вторичная волны складываются между собой. В момент излучения вторичная волна имеет ту же фазу колебаний, что и первичная, но вследствие неодинаковых фазовых скоростей распространения в среде между волнами образуется разность фаз, в результате чего вторичная волна гасит первичную. В итоге первичная волна замещается вторичной, которая распространяется со скоростью, характерной для среды.

Если показатель преломления среды n , то на длине x между волнами образуется разность фаз $\Delta\phi = \omega|n - 1|x/c$. В качестве длины замещения принимается расстояние, на котором $\Delta\phi \approx 1$. Поэтому длина замещения равна

$$L = c/(\omega|n - 1|) = \lambda/(2\pi|n - 1|). \quad (15.41)$$

Для прозрачных тел типа стекол ($n \approx 1,5 - 2$) в световом диапазоне длина замещения $L \approx (2 - 1) \cdot 10^{-7}$ м, а для воздуха и газов $L \approx 0,5$ мм. В межзвездной среде значение $(n - 1)$ чрезвычайно мало и поэтому длина замещения значительна. Расчеты показывают, что длина замещения для света в межзвездной среде составляет примерно два световых года.

Дисперсия света в межзвездном пространстве. Дисперсия света в межзвездной среде чрезвычайно мала. Верхний предел дисперсии можно оценить из наблюдений пульсаров. У пульсаров периодически изменяется яркость свечения, причем периоды пульсаций составляют в ряде случаев тысячные доли секунды. При наличии дисперсии волны различных длин имеют различное время распространения от момента излучения до наблюдателя. Поэтому одновременно наблюдаются волны различной частоты, излученные пульсаром в различные моменты времени. Это означает, что при наличии дисперсии кривая изменения суммарной яркости должна «смазываться». Отсутствие такого «смазывания» позволяет оценить верхний предел дисперсии.

Пусть $v(\omega)$ — скорость волн частотой ω , Δt — продолжительность импульса излучения, L — расстояние от пульсара до наблюдателя. Условие отсутствия «смазывания» состоит в том, что разность времен распространения волн различной длины должны быть меньше

- В какой области лежат частоты собственных колебаний ионов? Какой физический смысл имеет утверждение, что для очень коротковолнового излучения диэлектрик является оптически менее плотной средой, чем вакуум?

Можно ли в принципе в среде передать сигнал со скоростью, большей групповой скорости света?

продолжительности импульса:

$$\left| \frac{L}{v(\omega_1)} - \frac{L}{v(\omega_2)} \right| \leq \Delta t. \quad (15.42)$$

Это неравенство целесообразно преобразовать к виду

$$\left| \frac{v(\omega_2) - v(\omega_1)}{c} \right| \leq \frac{\Delta t}{L}, \quad (15.43)$$

где $c^2 \approx v(\omega_1)v(\omega_2)$. Для пульсара в крабовидной туманности известны следующие данные: $L \approx 6 \cdot 10^3$ св. лет, $\Delta t = 3 \cdot 10^{-3}$ с. Поэтому результаты наблюдений этого пульсара позволяют утверждать, что $\Delta v/c < 10^{-14}$. Учитывая, что интервал наблюдаемых частот простирается от $4 \cdot 10^8$ до $3 \cdot 10^{20}$ Гц, с большой точностью заключаем, что дисперсия практически отсутствует.

Окраска тел. Наличие резонансных частот обуславливает возникновение окраски тел в результате селективного поглощения света вблизи резонансных частот ω_0 . Это может происходить как в толще вещества, так и в поверхностном слое, если при рассеянии свет проникает в толщу вещества и испытывает при этом селективное поглощение.

Если частота ω_0 (см. рис. 56) находится в ультрафиолетовой области, то в видимой части спектра нет селективного поглощения света и вещество является почти бесцветным и прозрачным. Например, у стекла ω_0 лежит в ультрафиолетовой области, сравнительно близко к границе видимого спектра. Благодаря этому оно прозрачно, но имеет достаточно большой показатель преломления (около 1,5) и сильно поглощает ультрафиолетовые лучи. Когда собственные частоты ω_0 , попадают в видимую часть спектра, соответствующие участки видимой части спектра поглощаются и вещество оказывается окрашенным в дополнительный цвет. Например, окраска фиолетовых чернил в сосуде возникает в результате того, что при прохождении белого света через их толщу поглощается желтоватая часть спектра. На листе бумаги фиолетовые чернила также имеют фиолетовый цвет. Это обусловлено тем, что волокна поверхностного слоя бумаги пропитаны чернилами. Отражение (рассеяние) света происходит в некоторой толщине поверхностного слоя бумаги, в котором пропитанные чернилами волокна поглощают желтоватую часть спектра. Однако если капля чернил попадает на поверхность, которая не пропитывается чернилами (например, на поверхность стекла), то после высыхания чернильное пятно имеет желтоватый цвет. Возникновение этого цвета обусловлено не селективным поглощением, а селективным отражением света. Очень эффективно различие окраски материалов при селективном поглощении и отражении света проявляется у золота. Толстый кусок золота имеет желто-красноватую окраску. Однако если приготовить очень тонкий золотой лепесток, то при наблюдении на просвет его цвет глубоко голубой. В первом случае окраска золота обусловливается селективным отражением, а во втором — селективным поглощением.

При отражении света (см. § 16, 20) от поверхности наиболее интенсивно отражаются те участки спектра, которые при прохождении толщины вещества наиболее сильно поглощаются. Поэтому цвет вещества, возникающий за счет селективного отражения, является дополнительным к цвету того же вещества, возникающего в результате селективного поглощения.

Пример 15.1. Выразить групповую скорость (15.40) в виде функции от показателя преломления и длины волны и найти групповую скорость в воде для $\lambda = 656,3$ нм, если известно, что при 20°C показатель преломления для этой длины волны равен 1,3311, а для $\lambda = 643,8$ нм он равен 1,3314.

Поскольку $\omega = 2\pi c/(n\lambda)$, $k = 2\pi/\lambda$, из (15.40) получаем

$$v_g = \frac{d[2\pi c/(n\lambda)]}{d(2\pi/\lambda)} = c \frac{(nd\lambda + \lambda dn)/(n^2\lambda^2)}{d\lambda/n^2} = \frac{c}{n} \left(1 + \frac{\lambda}{n} \frac{dn}{d\lambda} \right). \quad (15.44)$$

Подставляя числовые значения в (15.44), находим

$$v_g = \frac{3 \cdot 10^8}{1,3311} \left(1 + \frac{656,3}{1,3311} \frac{0,0003}{12,5} \right) \text{ м/с} = 2,28 \cdot 10^8 \text{ м/с}.$$

Показывается, как граничные условия для векторов электрического и магнитного полей волн полностью определяют законы отражения и преломления.

Граничные условия. Диэлектрическая проницаемость диэлектриков различна. Поведение волны на границе полностью определяется граничными условиями для векторов поля волны, которые при отсутствии свободных зарядов и токов проводимости имеют вид

$$D_{2n} = D_{1n}, \quad B_{2n} = B_{1n}; \quad (a) \quad E_{2t} = E_{1t}, \quad H_{2t} = H_{1t}, \quad (b)$$
 (16.1)

где индексы n и t обозначают соответственно нормальную и тангенциальную составляющие вектора. Величины, относящиеся к падающей, отраженной и преломленной волнам, будем обозначать соответственно индексами «пд», «от», «пр». Величины, относящиеся к среде, в которой распространяются падающая и отраженная волны, обозначаются индексом 1, а к среде, в которой распространяется преломленная волна, — индексом 2. В частности, диэлектрические проницаемости этих сред равны ϵ_1 и ϵ_2 .

Напряженности электрического поля падающей, отраженной и преломленной плоских волн выражаются формулами

$$\begin{aligned} E_{\text{пд}} &= E_{\text{пд}}^{(0)} e^{-i(\omega_{\text{пд}} t - \mathbf{k}_{\text{пд}} \cdot \mathbf{r})}, \\ E_{\text{от}} &= E_{\text{от}}^{(0)} e^{-i(\omega_{\text{от}} t - \mathbf{k}_{\text{от}} \cdot \mathbf{r})}, \\ E_{\text{пр}} &= E_{\text{пр}}^{(0)} e^{-i(\omega_{\text{пр}} t - \mathbf{k}_{\text{пр}} \cdot \mathbf{r})}, \end{aligned}$$
 (16.2)

причем волновые числа связаны со скоростями распространения волн в средах соотношениями $k_{\text{пд}} = \omega_{\text{пд}} / v_1$, $k_{\text{от}} = \omega_{\text{от}} / v_1$, $k_{\text{пр}} = \omega_{\text{пр}} / v_2$,

где $v_1 = (\epsilon_1 \mu_1)^{-1}$, $v_2 = (\epsilon_2 \mu_2)^{-1}$ — скорости распространения волн в первой и второй средах. Выражения для вектора индукции магнитного поля волны получаются по общему правилу: вектор \mathbf{B} перпендикулярен \mathbf{E} , колеблется в одинаковой с \mathbf{E} фазе (в диэлектриках) и по модулю равен $|\mathbf{B}| = |\mathbf{E}|/v$, где v — скорость распространения волны.

Законы отражения и преломления получаются как следствие граничных условий (16.1) для векторов (16.2) и соответствующих векторов магнитного поля.

Постоянство частоты волны при отражении и преломлении. Граничное условие для тангенциальных составляющих напряженности электрического поля имеет вид

$$[E_{\text{пд}}^{(0)} e^{-i(\omega_{\text{пд}} t - \mathbf{k}_{\text{пд}} \cdot \mathbf{r})} + E_{\text{от}}^{(0)} e^{-i(\omega_{\text{от}} t - \mathbf{k}_{\text{от}} \cdot \mathbf{r})}]_t = [E_{\text{пр}}^{(0)} e^{-i(\omega_{\text{пр}} t - \mathbf{k}_{\text{пр}} \cdot \mathbf{r})}]_t,$$
 (16.4)

причем начало отсчета радиуса-вектора произвольно. Радиус-вектор можно представить в виде $\mathbf{r} = \mathbf{r}_n + \mathbf{r}_t$, где \mathbf{r}_n и \mathbf{r}_t — соответственно нормальная к поверхности и тангенциальная составляющие радиуса-вектора. Другими словами, \mathbf{r}_n соединяет начало отсчета радиус-вектора по нормали с соответствующей точкой плоскости раздела, а \mathbf{r}_t лежит в плоскости раздела, соединяя конец вектора \mathbf{r}_n с концом вектора \mathbf{r} . Из равенства $\mathbf{k} \cdot \mathbf{r} = \mathbf{k}_n \cdot \mathbf{r}_n + \mathbf{k}_t \cdot \mathbf{r}_t$, где \mathbf{k}_n и \mathbf{k}_t — нормальная и тангенциальная составляющие волнового вектора волны, заключаем, что при переходе от одной точки поверхности к другой скалярное произведение $\mathbf{k}_n \cdot \mathbf{r}_n$ постоянно, а все изменение значения $\mathbf{k} \cdot \mathbf{r}$ сводится к изменению $\mathbf{k}_t \cdot \mathbf{r}_t$. Поэтому по причине, которая сейчас станет очевидной, целесообразно начало отсчета радиусов-векторов взять в точке плоскости раздела сред. Эта точка в остальном произвольна. Чтобы не усложнять формулы, будем по-прежнему обозначать радиус-вектор \mathbf{r} без индекса t , хотя лежит \mathbf{r} в плоскости раздела сред. Равенство (16.4) может тождественно соблюдаться при произвольных и независимых изменениях t и \mathbf{r} лишь в том случае, если

$$\omega_{\text{нд}} t = \omega_{\text{от}} t = \omega_{\text{нр}} t, \quad (16.5) \quad \mathbf{k}_{\text{нд}} \cdot \mathbf{r} = \mathbf{k}_{\text{от}} \cdot \mathbf{r} = \mathbf{k}_{\text{нр}} \cdot \mathbf{r}. \quad (16.6)$$

Из (16.5) следует, что

$$\omega_{\text{нд}} = \omega_{\text{от}} = \omega_{\text{нр}}, \quad (16.7)$$

т. е. частота электромагнитной волны при отражении и преломлении не изменяется.

Плоскость падающего, отраженного и преломленного лучей. Направление, характеризуемое волновым вектором \mathbf{k} , будем называть лучом. Выберем начало отсчета вектора \mathbf{r} в плоскости раздела сред так, чтобы он был перпендикулярен вектору $\mathbf{k}_{\text{нд}}$, т. е. выполнялось условие

$$\mathbf{k}_{\text{нд}} \cdot \mathbf{r} = 0. \quad (16.8)$$

Из (16.6) следует, что

$$\mathbf{k}_{\text{от}} \cdot \mathbf{r} = \mathbf{k}_{\text{нр}} \cdot \mathbf{r} = 0, \quad (16.9)$$

т. е. векторы $\mathbf{k}_{\text{от}}$ и $\mathbf{k}_{\text{нр}}$ также перпендикулярны \mathbf{r} . Значит, волновые векторы падающей, отраженной и преломленной волн лежат в одной плоскости.

Соотношения между углами падения, отражения и преломления. Поместим начало прямоугольной декартовой системы координат в точку падения луча, ось Z направим в сторону среды, в которой распространяется преломленный луч, а плоскость XZ будем считать совпадающей с плоскостью, в которой лежат падающий, отраженный и преломленный лучи (рис. 58). Векторы $\mathbf{k}_{\text{нд}}$, $\mathbf{k}_{\text{от}}$, $\mathbf{k}_{\text{нр}}$ приложены к точке O (для упрощения чертежа они разнесены от этой точки вдоль прямых, характеризующих распространение "соответствующих волн"). Единичный вектор \mathbf{n} направлен во вторую среду по нормали к поверхности раздела. Единичный вектор \mathbf{t} лежит в плоскости раздела вдоль оси X . Он является тангенциальным вектором. Напомним, что единичные векторы безразмерны. Углы $\theta_{\text{нд}}$, $\theta_{\text{от}}$ и $\theta_{\text{нр}}$, отсчитываемые от перпендикуляров к поверхности раздела, называются соответственно углами падения, отражения и преломления. Плоскость, в которой лежат вектор падающей волны и нормаль к поверхности раздела в точке падения луча, называется **плоскостью падения** луча.

Начало отсчета \mathbf{r} поместим в точку O' , расположенную на оси X системы координат при отрицательном значении x . Следовательно, вектор $\mathbf{r} = r\mathbf{t}$ ориентирован в сторону положительных значений оси X . Из рис. 60 видно, что

$$\mathbf{k}_{\text{нд}} \cdot \mathbf{r} = \mathbf{k}_{\text{нд}} \cdot r\mathbf{t} = k_{\text{нд}} r \sin \theta_{\text{нд}}, \quad \mathbf{k}_{\text{от}} \cdot \mathbf{r} = \mathbf{k}_{\text{от}} \cdot r\mathbf{t} = k_{\text{от}} r \sin \theta_{\text{от}}, \quad (16.10)$$

$$\mathbf{k}_{\text{нр}} \cdot \mathbf{r} = \mathbf{k}_{\text{нр}} \cdot r\mathbf{t} = k_{\text{нр}} r \sin \theta_{\text{нр}}.$$

Тогда [см. (16.6)]

$$k_{\text{нд}} \sin \theta_{\text{нд}} = k_{\text{от}} \sin \theta_{\text{от}} = k_{\text{нр}} \sin \theta_{\text{нр}}. \quad (16.11)$$

Эти соотношения с учетом (16.3) упрощаются:

$$\sin \theta_{\text{нд}} / v_1 = \sin \theta_{\text{от}} / v_1 = \sin \theta_{\text{нр}} / v_2. \quad (16.12)$$

Из (16.12) следует, что $\sin \theta_{\text{нд}} = \sin \theta_{\text{от}}$, т. е.

$$\theta_{\text{нд}} = \theta_{\text{от}}, \quad (16.13)$$

$$\sin \theta_{\text{нд}} / \sin \theta_{\text{нр}} = v_1 / v_2 = n_{12}, \quad (16.14)$$

где $n_{12} = v_1 / v_2$ — показатель преломления второй среды относительно первой. Показатели преломления первой и второй сред относительно вакуума по определению равны

$$n_1 = c/v_1, \quad n_2 = c/v_2 \quad (16.15)$$

и поэтому

$$n_{12} = n_2/n_1. \quad (16.16)$$

Равенство (16.13) показывает, что угол отражения равен углу падения, а соотношение (16.14) выражает закон преломления Снеллиуса: отношение синуса угла падения к синусу угла преломления равно показателю преломления среды с преломленным лучом относительно среды с падающим лучом

Показатель преломления зависит от длины волны и от температуры, а для газов — и от давления. Например, показатель преломления воды при 20°C для длин волн 678,0; 589,3; 480,0 и 404,7 нм равен соответственно 1,3308; 1,3330; 1,3374 и 1,3428. Если нет необходимости учитывать показатель преломления с большой точностью, то можно для света принять $n = 1,33$. Показатель преломления газов при нормальных условиях отличается от единицы на 10^{-3} или 10^{-4} . Например, при 0°C и атмосферном давлении у азота $n = 1,000297$, у кислорода $n = 1,000272$, у воздуха $n = 1,000292$. Обычно для воздуха принимается $n = 1,0003$. У стекол марки «флинт» показатель преломления заключен между 1,6 и 1,9; у стекол марки «крон» — между 1,5 и 1,6, у алмаза показатель преломления равен 2,4.

На рис. 58 и 59 показан ход лучей при $n_2 > n_1$ и $n_2 < n_1$. Во втором случае видно, что при некотором угле падения $\theta_{\text{пад}} = \theta_{\text{пред}}$, называемом предельным углом, угол преломления становится равным $\pi/2$, т. е. преломленный луч движется вдоль поверхности раздела (рис. 60) и нет никакого преломленного луча во второй среде. Из (16.14) с учетом (16.16) получаем

$$\sin \theta_{\text{пред}} = n_2/n_1 \quad (n_2 < n_1). \quad (16.17)$$

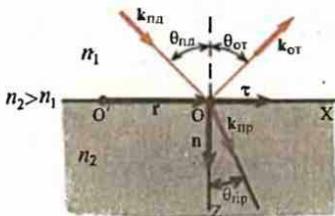
При углах падения, больших $\theta_{\text{пред}}$, уравнение (16.14) не имеет решения в области действительных значений угла преломления $\theta_{\text{пр}}$. Поэтому возникающую при этом ситуацию нельзя изобразить аналогично тому, как это сделано на рис. 58 и 59. Этот случай требует особого рассмотрения (см. §17).

Разложение плоской волны на две с взаимно перпендикулярными линейными поляризациями. Для того чтобы сделать легко обозримым вопрос об энергетических соотношениях при отражении и преломлении, целесообразно общий случай падающей волны, когда вектор $E_{\text{пад}}$ направлен под произвольным углом к плоскости падения, разбить на два: когда вектор $E_{\text{пад}}$ лежит в плоскости падения и когда он перпендикулярен ей. Для этого надо доказать, что плоскую волну можно представить в виде суммы плоских волн с взаимно перпендикулярными поляризациями, причем сумма плотностей потока энергии этих волн должна быть равной плотности потока энергии исходной волны. Просто из принципа суперпозиции это утверждение не следует.

Будем обозначать с индексом \parallel компоненты векторов, лежащих в плоскости падения, а с индексом \perp — в перпендикулярной плоскости. Поскольку отраженный и преломленный

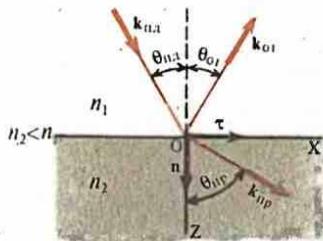
58

Преломление света при $n_2 > n_1$



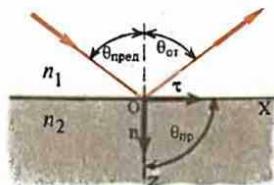
59

Преломление света при $n_2 < n_1$



60

При предельном угле падения угол преломления становится равным $\pi/2$

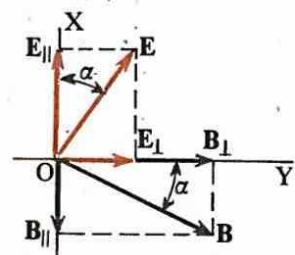


лучи лежат также в плоскости падения, на основании принципа суперпозиции векторы электрической напряженности падающей, отраженной и преломленной волн имеют вид

$$\begin{aligned} \mathbf{E}_{\text{пад}} &= \mathbf{E}_{\text{пад}\parallel} + \mathbf{E}_{\text{пад}\perp}, & \mathbf{E}_{\text{от}} &= \mathbf{E}_{\text{от}\parallel} + \mathbf{E}_{\text{от}\perp}, \\ \mathbf{E}_{\text{пр}} &= \mathbf{E}_{\text{пр}\parallel} + \mathbf{E}_{\text{пр}\perp}. \end{aligned} \quad (16.18)$$

Аналогично можно представить и векторы индукции \mathbf{B} . На рис. 61 приведено это разложение векторов поля на составляющие. Ось Z направлена перпендикулярно плоскости рисунка от нас. Вдоль оси Z распространяется волна. Плоскость XZ является плоскостью падения. Из рис. 61 видно, что между составляющими векторов поля соблюдаются следующие соотношения:

$$\begin{aligned} E_{\parallel}/B_{\perp} &= E \cos \alpha / (B \cos \alpha) = E/B, \\ E_{\perp}/B_{\parallel} &= E \sin \alpha / (B \sin \alpha) = E/B. \end{aligned} \quad (16.19)$$



61

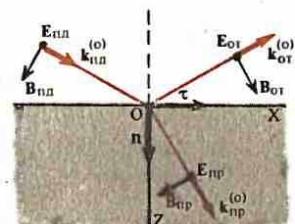
Разложение плоской волны на две волны, поляризованные во взаимно перпендикулярных плоскостях

Кроме того, видно, что векторы $(E_{\parallel}, B_{\perp}, \hat{k})$, $(E_{\perp}, B_{\parallel}, \hat{k})$ составляют правовинтовые тройки векторов. Отсюда заключаем, что эти векторы сами образуют плоские волны. Равенство $E^2 = E_{\parallel}^2 + E_{\perp}^2$ показывает, что плотность потока энергии исходной волны равна сумме плотностей потока энергии волн, на которые она разлагается. Тем самым доказано, что плоскую волну, вектор \mathbf{E} которой произвольно ориентирован относительно плоскости падения, можно разложить на сумму волн, у одной из которых напряженность электрического поля лежит в плоскости падения, а у другой — перпендикулярно ей. Изучив поведение этих волн при отражении и преломлении прямым применением принципа суперпозиции с учетом аддитивности плотности потока энергии, получим все характеристики волны с произвольной ориентировкой напряженности электрического поля.

Вектор \mathbf{E} перпендикулярен плоскости падения (рис. 62). Векторы $\mathbf{k}_{\text{пад}}^{\circ}$, $\mathbf{k}_{\text{от}}^{\circ}$, $\mathbf{k}_{\text{пр}}^{\circ}$ являются единичными, т. е. волновые векторы волн связаны с ними соотношениями $\mathbf{k}_{\text{от}} = \mathbf{k}_{\text{от}}^{\circ} \mathbf{k}_{\text{пад}}^{\circ}$ и т. д. Значения остальных величин такие же, как на рис. 60. Будем считать, что $\mathbf{E}_{\text{пад}}$ направлен к нам. Для упрощения формул не будем указывать индекс \perp , поскольку рассматриваются только напряженности, перпендикулярные плоскости падения. Если вектор $\mathbf{E}_{\text{пад}}$ направлен к нам, то вектор $\mathbf{B}_{\text{пад}}$ направлен так, как на рис. 62. Как направлены напряженности $\mathbf{E}_{\text{от}}$ и $\mathbf{E}_{\text{пр}}$, заранее не известно. Поэтому им можно присвоить произвольное направление и решать задачу. Если в результате решения у них получится отрицательный знак, то их действительное направление противоположно выбранному. Будем считать, что $\mathbf{E}_{\text{от}}$ и $\mathbf{E}_{\text{пр}}$ также направлены к нам. Векторы $\mathbf{B}_{\text{от}}$ и $\mathbf{B}_{\text{пр}}$ при этом ориентированы так, как на рис. 62.

Границные условия (16.16) для непрерывности тангенциальных составляющих напряженности электрического и магнитного полей имеют вид

$$\mathbf{E}_{\text{пад}} + \mathbf{E}_{\text{от}} = \mathbf{E}_{\text{пр}}, \quad (a) \quad (\mathbf{H}_{\text{пад}} + \mathbf{H}_{\text{от}}) \cdot \tau = \mathbf{H}_{\text{пр}} \cdot \tau. \quad (b) \quad (16.20)$$



62

Напряженность электрического поля падающей волны перпендикулярна плоскости падения

Для дальнейших преобразований векторы \mathbf{H} удобно выразить через векторы \mathbf{E} :

$$\mathbf{H} = \mathbf{k}^0 \times \mathbf{E}/Z,$$

(16.21)

$$\text{дл } Z = \sqrt{\mu/\epsilon},$$

(16.22)

— волновое сопротивление среды. Заметим, что для вакуума $\sqrt{\mu_0/\epsilon_0} = 377$ Ом. Учитывая, что вектор $\mathbf{E}_{\text{нд}}$ перпендикулярен плоскости чертежа и направлен к нам (рис. 62), заключаем

$$\tau = \mathbf{E}_{\text{нд}} \times \mathbf{n}/E_{\text{нд}}.$$

Поэтому тангенциальная составляющая \mathbf{H} падающей, отраженной и преломленной волн может быть представлена в виде

$$\mathbf{H} \cdot \tau = (\mathbf{k}^0 \times \mathbf{E}) \cdot (\mathbf{E}_{\text{нд}} \times \mathbf{n}) / (ZE_{\text{нд}}) = -\mathbf{E} \cdot (\mathbf{k}^0 \times \mathbf{n}) / Z,$$

где под \mathbf{H} , \mathbf{E} , Z , k можно понимать соответствующие величины в падающей, отраженной и преломленной волнах.

При преобразованиях в (16.24) учтена известная из векторной алгебры формула

$$(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) = (\mathbf{A} \cdot \mathbf{C})(\mathbf{B} \cdot \mathbf{D}) - (\mathbf{A} \cdot \mathbf{D})(\mathbf{B} \cdot \mathbf{C})$$

и принято во внимание, что $\mathbf{k}^0 \cdot \mathbf{E}_{\text{нд}} = 0$, $\mathbf{E} \cdot \mathbf{E}_{\text{нд}} = E^2$ для всех волн. Волновое сопротивление среды для падающей и отраженной волн равно Z_1 , а для преломленной — Z_2 . С учетом (16.24) граничные условия (16.21) принимают вид

$$(1/Z_1)[E_{\text{нд}}(\mathbf{k}_{\text{нд}}^0 \cdot \mathbf{n}) + E_{\text{от}}(\mathbf{k}_{\text{от}}^0 \cdot \mathbf{n})] = (1/Z_2)E_{\text{пр}}(\mathbf{k}_{\text{пр}}^0 \cdot \mathbf{n}).$$

Формулы Френеля для перпендикулярных составляющих векторов поля. Уравнения (16.20а) и (16.26) можно переписать в виде системы двух уравнений

$$1 + E_{\text{от}}/E_{\text{нд}} = E_{\text{пр}}/E_{\text{нд}},$$

$$(\mathbf{k}_{\text{нд}}^0 \cdot \mathbf{n}) + (E_{\text{от}}/E_{\text{нд}})(\mathbf{k}_{\text{от}}^0 \cdot \mathbf{n}) = (Z_1/Z_2)(\mathbf{k}_{\text{пр}}^0 \cdot \mathbf{n})(E_{\text{пр}}/E_{\text{нд}})$$

относительно неизвестных отношений $E_{\text{от}}/E_{\text{нд}}$ и $E_{\text{пр}}/E_{\text{нд}}$. Решение системы:

$$(E_{\text{от}}/E_{\text{нд}})_\perp = [Z_2(\mathbf{k}_{\text{нд}}^0 \cdot \mathbf{n}) - Z_1(\mathbf{k}_{\text{пр}}^0 \cdot \mathbf{n})]/[Z_2(\mathbf{k}_{\text{нд}}^0 \cdot \mathbf{n}) + Z_1(\mathbf{k}_{\text{пр}}^0 \cdot \mathbf{n})],$$

$$(E_{\text{пр}}/E_{\text{нд}})_\perp = 2Z_2(\mathbf{k}_{\text{нд}}^0 \cdot \mathbf{n})/[Z_2(\mathbf{k}_{\text{нд}}^0 \cdot \mathbf{n}) + Z_1(\mathbf{k}_{\text{пр}}^0 \cdot \mathbf{n})],$$

где $\mathbf{k}_{\text{от}}^0 \cdot \mathbf{n} = -\mathbf{k}_{\text{нд}}^0 \cdot \mathbf{n}$, поскольку угол падения равен углу отражения.

В соответствии с (16.14) угол преломления определяется соотношением

$$\sin \theta_{\text{пр}} = (n_1/n_2) \sin \theta_{\text{нд}}.$$

При $n_2 > n_1$ это уравнение относительно $\theta_{\text{пр}}$ имеет действительное решение для всех углов падения $\theta_{\text{нд}}$. Если же $n_2 < n_1$, то оно имеет действительное решение лишь для углов падения, меньших критического [см. (16.17)]. Поэтому формулы (16.28) удобнее выразить через углы падения и преломления. Учитывая, что $\mathbf{k}_{\text{нд}}^0 \cdot \mathbf{n} = \cos \theta_{\text{нд}}$, $\mathbf{k}_{\text{пр}}^0 \cdot \mathbf{n} = \cos \theta_{\text{пр}}$, перепишем (16.28) в виде

$$\left(\frac{E_{\text{от}}}{E_{\text{нд}}} \right)_\perp = \frac{Z_2 \cos \theta_{\text{нд}} - Z_1 \cos \theta_{\text{пр}}}{Z_2 \cos \theta_{\text{нд}} + Z_1 \cos \theta_{\text{пр}}} = \frac{\cos \theta_{\text{нд}} - (\mu_1/\mu_2) \sqrt{n_{12}^2 - \sin^2 \theta_{\text{нд}}}}{\cos \theta_{\text{нд}} + (\mu_1/\mu_2) \sqrt{n_{12}^2 - \sin^2 \theta_{\text{нд}}}},$$

$$\left(\frac{E_{\text{пр}}}{E_{\text{нд}}} \right)_\perp = \frac{2Z_2 \cos \theta_{\text{нд}}}{Z_2 \cos \theta_{\text{нд}} + Z_1 \cos \theta_{\text{пр}}} = \frac{2 \cos \theta_{\text{нд}}}{\cos \theta_{\text{нд}} + (\mu_1/\mu_2) \sqrt{n_{12}^2 - \sin^2 \theta_{\text{нд}}}}$$

С помощью соотношения

$$Z_1/Z_2 = \mu_1 \sin \theta_{\text{нп}} / (\mu_2 \sin \theta_{\text{нп}}) \quad (16.31)$$

формулам (16.30) можно придать иной вид:

$$(E_{\text{от}}/E_{\text{нп}})_\perp = (\mu_2 \operatorname{tg} \theta_{\text{нп}} - \mu_1 \operatorname{tg} \theta_{\text{нп}}) / (\mu_2 \operatorname{tg} \theta_{\text{нп}} + \mu_1 \operatorname{tg} \theta_{\text{нп}}), \quad (16.32a)$$

$$(E_{\text{нп}}/E_{\text{нп}})_\perp = 2\mu_2 \operatorname{tg} \theta_{\text{нп}} / (\mu_2 \operatorname{tg} \theta_{\text{нп}} + \mu_1 \operatorname{tg} \theta_{\text{нп}}). \quad (16.32b)$$

Если магнитные свойства сред по разные стороны границы раздела одинаковы ($\mu_1 = \mu_2$), то из (16.32a) и (16.32b) получаем соотношения

$$\left(\frac{E_{\text{от}}}{E_{\text{нп}}} \right)_\perp = - \frac{\sin(\theta_{\text{нп}} - \theta_{\text{нп}})}{\sin(\theta_{\text{нп}} + \theta_{\text{нп}})}, \quad (16.32v)$$

$$\left(\frac{E_{\text{нп}}}{E_{\text{нп}}} \right)_\perp = \frac{2 \cos \theta_{\text{нп}} \sin \theta_{\text{нп}}}{\sin(\theta_{\text{нп}} + \theta_{\text{нп}})} = \frac{(n_1/n_2) \sin 2\theta_{\text{нп}}}{\sin(\theta_{\text{нп}} + \theta_{\text{нп}})}, \quad (16.32g)$$

называемые формулами Френеля для колебаний в волне, перпендикулярных плоскости падения О. Ж. Френель (1788–1827) вывел их задолго (1818) до появления электромагнитной теории света Максвелла, рассматривая свет как колебания упругой среды — эфира.

При нормальном падении из (16.32) получаем

$$(E_{\text{от}}/E_{\text{нп}})_\perp = (\mu_2 n_1 - \mu_1 n_2) / (\mu_2 n_1 + \mu_1 n_2), \quad (16.33a)$$

$$(E_{\text{нп}}/E_{\text{нп}})_\perp = 2\mu_2 n_1 / (\mu_2 n_1 + \mu_1 n_2). \quad (16.33b)$$

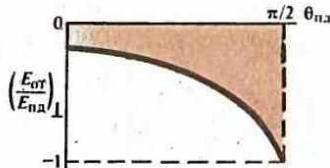
Для сред с одинаковыми магнитными свойствами ($\mu_1 = \mu_2$) по разные стороны границы раздела находим

$$(E_{\text{от}}/E_{\text{нп}})_\perp = (n_1 - n_2) / (n_1 + n_2), \quad (16.33v)$$

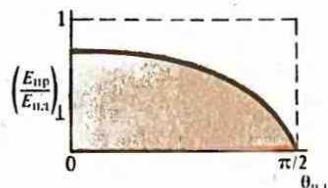
$$(E_{\text{нп}}/E_{\text{нп}})_\perp = 2n_1 / (n_1 + n_2). \quad (16.33g)$$

Формулы (16.33v) и (16.33g) можно, конечно, получить из (16.32v) и (16.32g), раскрыв в них неопределенности обычным, известным из математики способом.

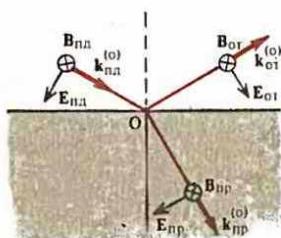
Изменения $(E_{\text{от}}/E_{\text{нп}})_\perp$ и $(E_{\text{нп}}/E_{\text{нп}})_\perp$ в зависимости от $\theta_{\text{нп}}$ при $n_1 < n_2$ в соответствии с формулами (16.32v) и (16.32g) показаны на рис. 63 и 64. На рис. 63 видно, что $(E_{\text{от}}/E_{\text{нп}})_\perp$ является отрицательной величиной для всех углов падения. Это означает, что направление вектора $F_{\text{от}}$ противоположно тому, которое указано на рис. 62, т. е. при отражении света от границы с средой с большим показателем преломления фаза отраженной волны изменяется на π (напряженность электрического поля изменяет направление на обратное). На рис. 64 видно, что $(E_{\text{нп}}/E_{\text{нп}})_\perp$ всегда положительно. Следовательно, при преломлении в этом случае не происходит изменения фазы.



Зависимость $(E_{\text{от}}/E_{\text{нп}})_\perp$ от угла падения при $n_1 < n_2$



Зависимость $(E_{\text{нп}}/E_{\text{нп}})_\perp$ от угла при $n_1 < n_2$



Напряженность электрического поля падающей волны лежит в плоскости падения

В случае $n_1 > n_2$ при углах падения, меньших предельного, видно, что $\theta_{\text{пад}} < \theta_{\text{пр}}$. Следовательно, в формуле (16.32в) $\sin(\theta_{\text{пад}} - \theta_{\text{пр}}) < 0$ и $(E_{\text{от}}/E_{\text{пад}})_\perp > 0$. Это означает, что при отражении света от границы со средой с меньшим показателем преломления не происходит изменения фазы у вектора E . Преломленная волна также не претерпевает изменения фазы.

Вектор E лежит в плоскости падения (рис. 65). Векторы B в всех волнах считаются направленными от нас за плоскость рисунка. Вместо (16.21) и (16.23) в этом случае имеем

$$E = ZH \times k^{(0)}, \quad (16.34) \quad \tau = n \times H_{\text{пад}} / H_{\text{пад}}. \quad (16.35)$$

Отсюда следует, что

$$E \cdot \tau = -ZH(k^{(0)} \cdot n) = -E(k^{(0)} \cdot n). \quad (16.36)$$

Поэтому граничные условия для тангенциальных составляющих напряженности электрического и магнитного полей

$$(E_{\text{пад}} + E_{\text{от}}) \cdot \tau = E_{\text{пр}} \cdot \tau, \quad (16.37)$$

$$H_{\text{пад}} + H_{\text{от}} = H_{\text{пр}} \quad (16.38)$$

принимают вид

$$E_{\text{пад}}(k_{\text{пад}}^{(0)} \cdot n) + E_{\text{от}}(k_{\text{от}}^{(0)} \cdot n) = (E_{\text{пад}} - E_{\text{от}})(k_{\text{пад}}^{(0)} \cdot n) = E_{\text{пр}}(k_{\text{пр}}^{(0)} \cdot n), \quad (16.39)$$

$$(E_{\text{пад}} + E_{\text{от}})/Z_1 = E_{\text{пр}}/Z_2. \quad (16.40)$$

Формулы Френеля для параллельных составляющих векторов поля. Разрешая эти уравнения относительно $E_{\text{от}}/E_{\text{пад}}$ и $E_{\text{пр}}/E_{\text{пад}}$, в полной аналогии с (16.28) получаем

$$(E_{\text{от}}/E_{\text{пад}})_{\parallel} = [Z_1(k_{\text{пад}}^{(0)} \cdot n) - Z_2(k_{\text{пр}}^{(0)} \cdot n)]/[Z_1(k_{\text{пад}}^{(0)} \cdot n) + Z_2(k_{\text{пр}}^{(0)} \cdot n)], \quad (16.41\alpha)$$

$$(E_{\text{пр}}/E_{\text{пад}})_{\parallel} = 2Z_2(k_{\text{пад}}^{(0)} \cdot n)/[Z_1(k_{\text{пад}}^{(0)} \cdot n) + Z_2(k_{\text{пр}}^{(0)} \cdot n)]. \quad (16.41\beta)$$

При действительных углах преломления (т. е. для $n_2 > n_1$, а для $n_2 < n_1$ при $\theta_{\text{пад}} < \theta_{\text{пред}}$) получаем аналогично (16.30) формулы

$$\left(\frac{E_{\text{от}}}{E_{\text{пад}}} \right)_{\parallel} = \frac{Z_1 \cos \theta_{\text{пад}} - Z_2 \cos \theta_{\text{пр}}}{Z_1 \cos \theta_{\text{пад}} + Z_2 \cos \theta_{\text{пр}}} = \frac{(\mu_1/\mu_2)n_{12}^2 \cos \theta_{\text{пад}} - \sqrt{n_{12}^2 - \sin^2 \theta_{\text{пад}}}}{(\mu_1/\mu_2)n_{12}^2 \cos \theta_{\text{пад}} + \sqrt{n_{12}^2 - \sin^2 \theta_{\text{пад}}}}, \quad (16.42\alpha)$$

$$\left(\frac{E_{\text{пр}}}{E_{\text{пад}}} \right)_{\parallel} = \frac{2Z_2 \cos \theta_{\text{пад}}}{Z_1 \cos \theta_{\text{пад}} + Z_2 \cos \theta_{\text{пр}}} = \frac{2n_{12} \cos \theta_{\text{пад}}}{(\mu_1/\mu_2)n_{12}^2 \cos \theta_{\text{пад}} + \sqrt{n_{12}^2 - \sin^2 \theta_{\text{пад}}}}. \quad (16.42\beta)$$

С помощью (16.31) при $\mu_1 = \mu_2$ формулы (16.42) можно представить в виде

$$\left(\frac{E_{\text{от}}}{E_{\text{пад}}} \right)_{\parallel} = \frac{\sin 2\theta_{\text{пад}} - \sin 2\theta_{\text{пр}}}{\sin 2\theta_{\text{пад}} + \sin 2\theta_{\text{пр}}} = \frac{\operatorname{tg}(\theta_{\text{пад}} - \theta_{\text{пр}})}{\operatorname{tg}(\theta_{\text{пад}} + \theta_{\text{пр}})}, \quad (16.43\alpha)$$

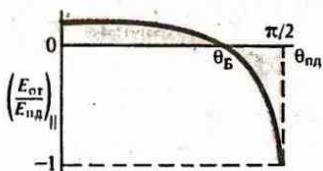
$$\left(\frac{E_{\text{пр}}}{E_{\text{пад}}} \right)_{\parallel} = \frac{4 \cos \theta_{\text{пад}} \sin \theta_{\text{пр}}}{\sin 2\theta_{\text{пад}} + \sin 2\theta_{\text{пр}}} = \frac{2 \cos \theta_{\text{пад}} \sin \theta_{\text{пр}}}{\sin(\theta_{\text{пад}} + \theta_{\text{пр}}) \cos(\theta_{\text{пад}} - \theta_{\text{пр}})}. \quad (16.43\beta)$$

Эти формулы называются формулами Френеля для лежащих в плоскости падения компонент напряженности электрического поля волны. Согласно с (16.32) они дают полное решение задачи о поведении полей электромагнитной волны при отражении и преломлении на границе между диэлектриками при произвольной ориентировке векторов волн относительно плоскости падения в падающей волне.

Для нормального падения ($\theta_{\text{пад}} = 0$) формулы упрощаются:

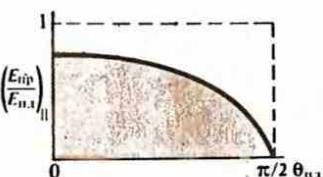
3)
$$\left(\frac{E_{\text{от}}}{E_{\text{пад}}}\right)_{\parallel} = \frac{\mu_1 n_2 - \mu_2 n_1}{\mu_1 n_2 + \mu_2 n_1}, \quad \left(\frac{E_{\text{от}}}{E_{\text{пад}}}\right)_{\perp} = \frac{n_2 - n_1}{n_2 + n_1} \quad (\mu_1 = \mu_2), \quad (16.44a)$$

$$\left(\frac{E_{\text{пр}}}{E_{\text{пад}}}\right)_{\parallel} = \frac{2\mu_2 n_1}{\mu_1 n_2 + \mu_2 n_1}, \quad \left(\frac{E_{\text{пр}}}{E_{\text{пад}}}\right)_{\perp} = \frac{2n_1}{n_2 + n_1}, \quad (\mu_1 = \mu_2). \quad (16.44b)$$



66

Зависимость $(E_{\text{от}}/E_{\text{пад}})_{\parallel}$ от угла падения при $n_1 < n_2$.



67

Зависимость $(E_{\text{пр}}/E_{\text{пад}})_{\perp}$ от угла падения при $n_1 < n_2$.

● При отражении изменяется на π фаза колебаний либо вектора напряженности электрического поля волны, либо вектора напряженности магнитного поля, что определяется формулами Френеля. Фазы колебаний векторов волн при преломлении не изменяются.

При падении под углом Брюстера отражается только волна, в которой вектор \mathbf{E} колеблется перпендикулярно плоскости падения.

● Сформулируйте граничные условия для векторов поля волны, которые полностью определяют законы отражения и преломления. Какие физические факторы определяют значения угла Брюстера? Почему экспериментальная проверка формулы Френеля может быть выполнена наиболее эффективно при углах Брюстера?

На первый взгляд кажется, что формула (16.44a) противоречит (16.33a), поскольку при нормальном падении на поверхность раздела плоскость падения может считаться ориентированной произвольно и, следовательно, не должно быть различия в формулах для параллельных и перпендикулярных составляющих, а (16.44a) и (16.33a) имеют противоположные знаки. Однако в действительности никакого противоречия между ними нет. Дело в том, что формула (16.33a) выведена для случая, когда векторы $E_{\text{пад}}$ и $E_{\text{от}}$ при нормальном падении считаются совпадающими по направлению (см. рис. 64) и число $E_{\text{пад}}/E_{\text{от}}$ принимается положительным. Вывод же формул (16.44a) основывался для случая, когда векторы $E_{\text{пад}}$ и $E_{\text{от}}$ при нормальном падении считались направленными противоположно (рис. 65) и, следовательно, принималось, что число $E_{\text{от}}/E_{\text{пад}}$ отрицательно. Различие знаков в формулах (16.33a) и (16.44a) отражает эту разницу в исходных предположениях при выводе формул.

Изменения $(E_{\text{от}}/E_{\text{пад}})_{\parallel}$ и $(E_{\text{пр}}/E_{\text{пад}})_{\parallel}$ в зависимости от угла падения $\theta_{\text{пад}}$ при $n_1 < n_2$ в соответствии с формулами (16.43) показаны на рис. 66 и 67. Рис. 67 аналогичен рис. 64. В совокупности они показывают, что при любых ориентировках напряженности электрического поля в падающей волне при преломлении не происходит изменения фазы. Кроме того, видно, что при падении волны касательно поверхности раздела ($\theta_{\text{пад}} = \pi/2$) преломленная волна отсутствует.

Явление Брюстера. Из анализа хода лучей, показанных на рис. 66, можно заключить, что волны, электрический вектор которой лежит в плоскости падения при угле падения θ_B , отраженная волна полностью отсутствует. Это явление было открыто экспериментально в 1815 г. Д. Брюстером (1781—1868) и называется **явлением Брюстера**. Угол θ_B находится по формуле (16.43a) из условия $(E_{\text{от}}/E_{\text{пад}})_{\parallel} = 0$, т. е. когда знаменатель правой части равенства обращается в бесконечность. Таким образом, этот угол находится из условия $\theta_B + \theta_{\text{пр}B} = \pi/2$, в котором $\theta_{\text{пр}B}$ является углом преломления, соответствующим $\theta_{\text{пад}} = \theta_B$. По закону Снеллиуса, $n_1 \sin \theta_B = n_2 \sin \theta_{\text{пр}B} = n_2 \sin(\pi/2 - \theta_B) = n_2 \cos \theta_B$ и тогда получаем уравнение $\tan \theta_B = n_2/n_1$. (16.45)

Эта формула называется **законом Брюстера**, а угол θ_B , при котором отсутствует отраженная волна, — **углом Брюстера**. На рис. 66 видно, что при $\theta_{\text{пад}} < \theta_B$ отношение $(E_{\text{от}}/E_{\text{пад}})_{\parallel}$ положительно, а при $\theta_{\text{пад}} > \theta_B$ отрицательно. Это означает, что при переходе через угол Брюстера фаза отраженной волны скачком изменяется на π .

Поскольку в падающей под углом Брюстера волне лежащая в плоскости падения компонента вектора $E_{\text{пад}}$ не отражается, в отраженной волне имеется лишь компонента напряженности электрического поля, перпендикулярная плоскости падения, т. е. отраженный свет полностью поляризован. Отражение под углом Брюстера является одним из способов получения линейно поляризованного света.

Явление Брюстера обусловлено поперечностью электромагнитных волн. Под влиянием падающей волны электроны среды начинают колебаться и излучают вторичные волны, которые складываются с первоначальной. На длине замещения (см. § 15) произойдет полная замена падающей волны волной, излучаемой колеблющимися электронами среды. Линия колебаний электронов коллинеарна вектору E волны. При угле Брюстера, когда угол между преломленной и отраженной волнами должен был бы составить 90° , электроны среды, порождающие преломленную волну, колеблются вдоль линии, параллельной направлению, в котором должна распространяться отраженная волна. Вдоль линии своих колебаний электрон не может излучить электромагнитную волну. Поэтому отраженная волна отсутствует.

Изучение отражения света под углами, близкими к углу Брюстера, позволяет подвергнуть формулы Френеля экспериментальной проверке с чрезвычайно большой точностью. Это связано с тем, что интенсивность отраженной под углом Брюстера волны, в которой вектор колебляется в плоскости падения, равна нулю. Такое утверждение удобно для экспериментальной проверки, потому что не требует точного соблюдения угла падения. Достаточно непрерывно изменять угол падения вблизи угла Брюстера. При прохождении угла Брюстера интенсивность волны с соответствующей поляризацией должна обратиться в нуль. Кроме того, нулевые измерения обычно более точны. Второе утверждение, удобное для экспериментальной проверки, состоит в том, что отраженная под углом Брюстера волна является точно линейно поляризованной с вектором напряженности электрического поля, колеблющимся перпендикулярно плоскости падения. Здесь также возможны измерения типа нулевых.

Исследования отражения света под углом Брюстера показали, что имеются небольшие отклонения от предсказаний формул Френеля. Оказалось, что не существует такого угла падения, при котором интенсивность отраженной волны с электрическим вектором, колеблющимся в плоскости падения, была бы равна нулю, а электрический вектор отраженной волны колебался бы по линии, перпендикулярной плоскости падения. Если в падающей линейно поляризованной волне вектор E перпендикулярен плоскости падения, то отраженная под углом Брюстера волна является эллиптически поляризованной, что находится в противоречии с формулами Френеля, которые предсказывают линейную поляризацию. Ясно также, что существование у эллиптически поляризованной волны компоненты вектора E в плоскости падения объясняет отсутствие угла, при котором интенсивность отраженной волны соответствующей поляризации была бы равна нулю.

Отступление от формул Френеля объясняется тем, что отражение и преломление происходят не на математической граничной поверхности, а в тонком переходном слое между различными средами. Свойства переходного слоя отличаются от свойств сред, которые он разграничивает. Благодаря этому отражение и преломление не сводятся к скачкообразному и мгновенному изменению параметров, описывающих волны, как предполагается при выводе формул Френеля. Толщина переходного слоя, обуславливающего процесс преломления и отражения, имеет порядок межатомных расстояний. В таком слое процесс отражения и преломления не может быть описан с помощью уравнений Максвелла, сформулированных в приближении сплошной среды. Надо использовать уравнения Максвелла для вакуума, а переходный слой моделировать с учетом его молекулярной структуры. При таком подходе удалось объяснить отступление от формул Френеля. Отклонения от формул Френеля для углов, отличных от угла Брюстера, обнаружить затруднительно.

Соотношения между фазами волн при отражении или преломлении. Направление движения волн связано с взаимным расположением векторов E и H в волне правилом правого винта. Отсюда следует, что при отражении разность фаз колебаний E и H изменяется на π , а при пре-

ломлении — остается неизменной. Поведение фаз волн при отражении и преломлении анализируется с помощью формул Френеля и может быть описано следующим образом.

3 Преломление происходит при всех условиях без изменения фазы волн. При отражении происходит изменение фаз, зависящее от условий.

При падении волн на границу раздела со стороны оптически менее плотной среды при любых углах падения фаза перпендикулярной составляющей напряженности электрического поля в отраженной волне изменяется на π , т. е. $E_{\text{от}}^{\perp}$ и $E_{\text{пп}}^{\perp}$ имеют на границе раздела противоположные фазы. Фаза параллельной составляющей E'' изменяется лишь при углах падения, меньших угла Брюстера, а при углах падения, больших θ_B , фаза параллельной составляющей $E''_{\text{от}}$ остается той же (т. е. в этом случае на π изменяется фаза напряженности магнитного поля).

При падении волн на границу раздела со стороны оптически более плотной среды при любых углах падения фаза перпендикулярной составляющей напряженности электрического поля волны та же (т. е. в этом случае изменяется на π фаза напряженности магнитного поля). Фаза параллельной составляющей постоянна лишь при углах падения, меньших угла Брюстера. При углах падения, больших угла Брюстера, фаза параллельной составляющей электрического поля в отраженной волне изменяется на π (и, следовательно, фаза напряженности магнитного поля в отраженной волне не изменяется).

Степень поляризации. Для описания изменения поляризации света при отражении и преломлении пользуются степенью поляризации

$$P = (I^{\perp} - I^{\parallel}) / (I^{\perp} + I^{\parallel}). \quad (16.46)$$

Для неполяризованного света $P=0$. Для полностью поляризованного света, когда плоскость колебаний вектора E перпендикулярна плоскости падения, $P=1$. Если вектор E лежит в плоскости падения, то $P=-1$. Пользуясь определением (16.46), удобно обсуждать вопросы поляризации света при анализе отражения и преломления. Однако необходимо помнить, что в общей теории частично поляризованного света (см. §30) под степенью поляризации понимают величину (30.63), а не (16.46), хотя они одинаково называются и одинаково обозначены.

§ 17 | Полное отражение света

Рассматривается полное отражение света и изучаются свойства волны, проникающей во вторую среду.

Формулы для углов $\theta_{\text{пп}} > 0_{\text{пред}}$. Полное отражение происходит при $n_1 > n_2$, $\theta_{\text{пп}} > 0_{\text{пред}}$ [см. рис. 59 и (16.17)]. Его важнейшей физической характеристикой является отсутствие преломленной волны. Уравнение (16.14) в этом случае не имеет решения в области действительных значений угла преломления $\theta_{\text{пп}}$. Для анализа полного отражения необходимо законы преломления выразить формулами, справедливыми и для $n_1 > n_2$, $\theta_{\text{пп}} > 0_{\text{пред}}$. Прежде всего учтем, что равенства (16.11) могут быть представлены в виде

$$k_{\text{пп}} = k_{\text{от}} = k_{\text{пр}}, \quad (17.1)$$

где $k_{\text{пп}} = k_{\text{пп}} \sin \theta_{\text{пп}}$, $k_{\text{от}} = k_{\text{от}} \sin \theta_{\text{от}}$, $k_{\text{пр}} = k_{\text{пр}} \sin \theta_{\text{пр}}$ — соответственно тангенциальные компоненты волновых векторов падающей, отраженной и преломленной волн. Соотношения между нормальными компонентами волновых векторов этих волн (положительная нормаль считается направленной в сторону второй среды, см. рис. 59) получаются из очевидных равенств

$$k_{\text{пп}}^2 = k_{\text{пп}}^2 + k_{\text{напп}}^2, \quad k_{\text{от}}^2 = k_{\text{от}}^2 + k_{\text{напт}}^2, \quad k_{\text{пр}}^2 = k_{\text{пр}}^2 + k_{\text{напр}}^2 \quad (17.2)$$

в комбинации с соотношениями (17.1)

$$k_{\text{напп}} = -k_{\text{напн}}, \quad k_{\text{напр}}^2 = k_{\text{2}}^2 - k_{\text{напн}}^2, \quad (17.3)$$

где k_2 — модуль волнового вектора во второй среде (волновое число), а знак минус в первом из равенств (17.3) учитывает, что отраженная волна движется противоположно направлению нормали, выбранной за положительную, а преломленная — в направлении положительной нормали. Примем во внимание, что $k_{\text{нр},t} = k_1 \sin \theta_{\text{нд}}$, где k_1 — волновое число в первой среде. Равенства (17.3) при условии (17.1) принимают вид

$$k_{\text{нр},t} = k_1 \sin \theta_{\text{нд}}, \quad (17.4) \quad k_{\text{нр},n}^2 = k_2^2 - k_1^2 \sin^2 \theta_{\text{нд}} = k_2^2 [1 - (n_1/n_2)^2 \sin^2 \theta_{\text{нд}}]. \quad (17.5)$$

В этом виде формулы справедливы для произвольных значений угла падения $\theta_{\text{нд}}$ и отношения n_1/n_2 . Угол $\theta_{\text{нр}}$ определяется из соотношений

$$k_{\text{нр},t} = k_2 \sin \theta_{\text{нр}}, \quad k_{\text{нр},n} = k_2 \cos \theta_{\text{нр}}. \quad (17.6)$$

В том случае, когда этот угол имеет геометрическую интерпретацию, показанную на рис. 58 и 59, его смысл очевиден — это угол преломления. При полном отражении он не имеет геометрической интерпретации, но формулы (17.4) и (17.5) остаются справедливыми, а $\theta_{\text{нр}}$ по-прежнему определяется равенством (17.6).

Подставляя $k_{\text{нр},t}$ и $k_{\text{нр},n}$ из (17.6) в формулы (17.4) и (17.5), получаем законы преломления:

$$n_1 \sin \theta_{\text{нд}} = n_2 \sin \theta_{\text{нр}}, \quad (17.7) \quad \cos^2 \theta_{\text{нр}} = 1 - (n_1/n_2)^2 \sin^2 \theta_{\text{нд}}. \quad (17.8)$$

Эти два соотношения находятся в полном согласии между собой, поскольку $\cos^2 \theta_{\text{нр}} = 1 - \sin^2 \theta_{\text{нр}}$. Соотношение (17.7) совпадает с законом Снеллиуса (16.14). Соотношение (17.8) позволяет найти $\cos \theta_{\text{нр}}$ не только в случае, когда $\theta_{\text{нр}}$ имеет геометрическую интерпретацию, но и в случае, когда такая интерпретация отсутствует при полном отражении.

Волна во второй среде. Расположим оси координат так, как указано на рис. 60. Напряженности электрического поля в падающей и преломленной волнах выражаются формулами

$$E_{\text{нд}} = E_{\text{нд},x}^{(o)} \exp [-i(\omega t - k_{\text{нд}} \cdot r)] = E_{\text{нд}}^{(o)} \exp \{-i[\omega t - k_1(x \sin \theta_{\text{нд}} + z \cos \theta_{\text{нд}})]\}, \quad (17.9)$$

$$E_{\text{нр}} = E_{\text{нр}}^{(o)} \exp [-i(\omega t - k_{\text{нр}} \cdot r)] = E_{\text{нр}}^{(o)} \exp \{-i[\omega t - k_{\text{нр}} x - k_{\text{нр},z} z]\}, \quad (17.10)$$

где в соответствии с (17.4) можно принять $k_{\text{нр},x} = k_{\text{нр},t} = k_1 \sin \theta_{\text{нд}}$. Значение $k_{\text{нр},z} = k_{\text{нр},n}$ вычисляется из (17.5). Учитывая, что заключенная в (17.5) в квадратные скобки величина отрицательна, имеем

$$k_{\text{нр},n}^2 = k_{\text{нр},z}^2 = -k_2^2 [(n_1/n_2)^2 \sin^2 \theta_{\text{нд}} - 1] = -k_2^2 [(\sin \theta_{\text{нд}} / \sin \theta_{\text{нр}})^2 - 1] = -s^2, \quad (17.11)$$

где s — действительная величина. Отсюда $k_{\text{нр},z} = \pm is$, что после подстановки в (17.10) приводит к появлению действительного множителя $\exp (\pm sz)$. Следовательно, общее решение для преломленной волны $E_{\text{нр}}$ состоит из двух слагаемых, один из которых содержит множитель $\exp (sz)$, а второй — $\exp (-sz)$. Множитель $\exp (sz)$ возрастает бесконечно при $z \rightarrow \infty$, что соответствует безграничному росту амплитуды во второй среде при удалении от границы с первой средой. Ясно, что это неприемлемо из физических соображений (закон сохранения энергии) и должно быть отвергнуто. Остается слагаемое с множителем $\exp (-sz)$, которое имеет вид

$$E_{\text{нр}} = E_{\text{нр}}^{(o)} e^{-sz} \exp [-i(\omega t - k_1 x \sin \theta_{\text{нд}})], \quad (17.12)$$

где

$$s = k_2 [(n_1/n_2)^2 \sin^2 \theta_{\text{нд}} - 1]^{1/2} = (n_2 \omega / c) [(\sin \theta_{\text{нд}} / \sin \theta_{\text{нр}})^2 - 1]^{1/2}. \quad (17.13)$$

Глубина проникновения. Формула (17.12) описывает волну, распространяющуюся вдоль поверхности раздела двух сред в направлении положительных значений оси X , причем амплитуда этой волны убывает экспоненциально при удалении от границы раздела во вторую среду. Амплитуда убывает в e раз на расстоянии

$$\Delta = \frac{1}{s} = \frac{\lambda_2 / (2\pi)}{[(n_1/n_2)^2 \sin^2 \theta_{\text{нд}} - 1]^{1/2}},$$

(17.14)

называемым глубиной проникновения (см. § 19).

Фазовая скорость. Для определения фазовой скорости преломленной волны в направлении оси X надо, как обычно, фазу волны (17.12) считать постоянной и продифференцировать ее по времени:

$$v_{2x} = \frac{\omega}{k_1 \sin \theta_{\text{пп}}} = \frac{c}{n_1 \sin \theta_{\text{пп}}} = \frac{(c/n_2)}{(n_1/n_2) \sin \theta_{\text{пп}}} = \left(\frac{\sin \theta_{\text{пред}}}{\sin \theta_{\text{пп}}} \right) v_2, \quad (17.15)$$

где $v_2 = c/n_2$ — фазовая скорость плоской электромагнитной волны во второй среде. Поскольку $\theta_{\text{пп}} > \theta_{\text{пред}}$, скорость v_{2x} меньше v_2 . Волна, описываемая формулой (17.12), не является плоской, так как ее амплитуда не постоянна в плоскости, перпендикулярной направлению распространения. Фазовая скорость этой волны также непостоянна и зависит как от свойств среды, так и от угла падения.

Отраженная волна. Амплитуда отраженной волны может быть найдена по формулам (16.28а) и (16.41а), принимающим с учетом (17.7), (17.8) и (17.13) следующий вид:

$$\left(\frac{E_{\text{от}}}{E_{\text{пп}}} \right)_{\perp} = \frac{Z_2 \cos \theta_{\text{пп}} - Z_1 [1 - (n_1/n_2)^2 \sin^2 \theta_{\text{пп}}]^{1/2}}{Z_2 \cos \theta_{\text{пп}} + Z_1 [1 - (n_1/n_2)^2 \sin^2 \theta_{\text{пп}}]^{1/2}} = \frac{Z_2 \cos \theta_{\text{пп}} - iZ_1 (s/k_2)}{Z_2 \cos \theta_{\text{пп}} + iZ_1 (s/k_2)}, \quad (17.16)$$

$$\left(\frac{E_{\text{от}}}{E_{\text{пп}}} \right)_{\parallel} = \frac{Z_1 \cos \theta_{\text{пп}} - Z_2 [1 - (n_1/n_2)^2 \sin^2 \theta_{\text{пп}}]^{1/2}}{Z_1 \cos \theta_{\text{пп}} + Z_2 [1 - (n_1/n_2)^2 \sin^2 \theta_{\text{пп}}]^{1/2}} = \frac{Z_1 \cos \theta_{\text{пп}} + iZ_2 (s/k_2)}{Z_1 \cos \theta_{\text{пп}} - iZ_2 (s/k_2)}. \quad (17.17)$$

Тогда

$$(E_{\text{от}}/E_{\text{пп}})_{\perp} = e^{2i\varphi_{\perp}}, \quad (E_{\text{от}}/E_{\text{пп}})_{\parallel} = e^{2i\varphi_{\parallel}}, \quad (17.18)$$

где

$$\operatorname{tg} \varphi_{\perp} = - \frac{Z_1 (s/k_2)}{Z_2 \cos \theta_{\text{пп}}}, \quad \operatorname{tg} \varphi_{\parallel} = - \frac{Z_2 (s/k_2)}{Z_1 \cos \theta_{\text{пп}}}. \quad (17.19)$$

Следовательно, амплитуды параллельных и перпендикулярных плоскости падения составляющих напряженности электрического поля плоской волны при отражении не изменяются по абсолютному значению, но испытывают различные изменения по фазе. Равенство абсолютных значений амплитуды падающей и отраженной волн означает равенство плотностей потока

● При полном отражении луч во вторую среду проникает на глубину порядка длины волны и движется в ней параллельно границе раздела с фазовой скоростью, меньшей фазовой скорости волны во второй среде. Он возвращается в первую среду в точке, смещенной относительно точки входа.

○ Каким образом можно наблюдать проникновение волны во вторую среду при полном отражении?

Возможно ли полное отражение при падении волны на границу раздела со стороны оптически менее плотной среды?

Зависит ли фазовая скорость волны, проникшей во вторую среду и движущейся параллельно границе раздела, от угла падения?

энергии в них. Следовательно, поток энергии падающей волны полностью отражается и никакая часть энергии не уходит во вторую среду. Поэтому такое отражение называется полным. Однако формула (17.12) показывает, что энергия все же проникает во вторую среду. Истолкование этого проникновения с учетом того, что энергия полностью отражена, состоит в следующем: волна и соответствующая доля энергии проникают через границу раздела во вторую среду на небольшую глубину Δ , движутся вдоль поверхности раздела и затем возвращаются в первую среду. Места входа энергии во вторую среду и ее возвращения в первую несколько смещены друг относительно друга.

Проникновение волны во вторую среду можно наблюдать экспериментально. Интенсивность электромагнитной волны во второй среде имеет заметное значение лишь на расстояниях, меньших длины волны. Вблизи поверхности раздела, на которую падает свет, испытывающий полное отражение, на стороне второй среды наблюдается свечение тонкого флуоресцирующего слоя, если во второй среде имеется некоторое количество флуоресцирующего вещества. Это явление демонстрируется на лекциях. Смещение точки выхода отраженного луча относительно входа также экспериментально наблюдается в лабораторных условиях.

Заход электромагнитной волны во вторую среду наблюдался экспериментально также в микроволновой области и в области более длинных волн. В области сантиметровых волн антенный детектор может быть помещен на различных расстояниях от поверхности раздела, что позволяет не только доказать факт захода волны во вторую среду, но и изучить количественно зависимость амплитуды волны от расстояния до поверхности раздела.

Полное отражение обуславливает возникновение миражей, когда земная поверхность сильно нагрета. Например, если шоссе сильно нагрето, то температура воздуха максимальна у поверхности шоссе и убывает при удалении от поверхности. Следовательно, показатель преломления воздуха минимальен вблизи поверхности шоссе и возрастает при удалении от поверхности. Вследствие этого лучи, идущие под достаточно малым углом к поверхности шоссе, испытывают полное отражение. Может случиться, что шофер автомашины, сконцентрировав свое внимание на подходящем участке поверхности шоссе, увидит достаточно далеко идущую впереди автомашину в перевернутом состоянии. Аналогично, полное отражение света, идущего, например, от облаков приводит к возникновению впечатления о наличии луж на поверхности разогретого шоссе и т. д. Вертикальные градиенты температуры в комбинации с конвективными потоками воздуха могут создать весьма причудливые миражи.

§ 18 | Энергетические соотношения при преломлении и отражении света

Излагается способ перехода от характеристик для векторов волны при отражении и преломлении к характеристикам плотностей потока энергии.

Плотности потоков энергии. Плотность потока энергии плоской электромагнитной волны, усредненная по периоду, дается формулой (15.10), которую в векторной форме удобно представить в виде

$$\mathbf{S} = \frac{1}{2} v \epsilon E^{(0)} \mathbf{k}^{(0)}, \quad (18.1)$$

где $\mathbf{k}^{(0)}$ — единичный вектор в направлении распространения волны; знак усреднения при \mathbf{S} не показан для упрощения написания формул. Плотность потока (18.1) можно разложить на нормальную и тангенциальную к поверхности раздела составляющие:

$$\mathbf{S} = \mathbf{S}_n + \mathbf{S}_t, \quad (18.2)$$

где

$$\mathbf{S}_n = (\mathbf{S} \cdot \mathbf{n}) \mathbf{n}, \quad \mathbf{S}_t = (\mathbf{S} \cdot \boldsymbol{\tau}) \boldsymbol{\tau}. \quad (18.3)$$

Коэффициент отражения. Коэффициентом отражения называется абсолютное значение отношения нормальных компонент плотностей потоков энергии в отраженной и падающей волнах:

$$\rho = |\mathbf{S}_{\text{отр}}| / |\mathbf{S}_{\text{пад}}| = |\mathbf{S}_{\text{отр}} \cdot \mathbf{n}| / |\mathbf{S}_{\text{пад}} \cdot \mathbf{n}|. \quad (18.4)$$

Подставляя (18.1) в (18.4) и принимая во внимание, что $|\mathbf{k}_{\text{пад}}^{(0)} \cdot \mathbf{n}| = |\mathbf{k}_{\text{отр}}^{(0)} \cdot \mathbf{n}| = \cos \theta_{\text{пад}} = \cos \theta_{\text{отр}}$, получаем

$$\rho = |E_{\text{отр}} / E_{\text{пад}}|^2. \quad (18.5)$$

Учитывая полученные ранее соотношения для напряженности электрического поля в отраженной и падающей волнах при различных ориентировках $\mathbf{E}_{\text{пад}}$ относительно плоскости падения, можно найти коэффициенты отражения ρ_{\perp} и ρ_{\parallel} для перпендикулярной и параллельной ориентировки вектора $\mathbf{E}_{\text{пад}}$ относительно плоскости падения. С учетом (16.30а) и (16.41а) на основании (18.5) запишем:

$$\rho_{\perp} = [(Z_2 \cos \theta_{\text{пад}} - Z_1 \cos \theta_{\text{пр}}) / (Z_2 \cos \theta_{\text{пад}} + Z_1 \cos \theta_{\text{пр}})]^2, \quad (18.6\text{a})$$

$$\rho_{\parallel} = [(Z_1 \cos \theta_{\text{пад}} - Z_2 \cos \theta_{\text{пр}}) / (Z_1 \cos \theta_{\text{пад}} + Z_2 \cos \theta_{\text{пр}})]^2. \quad (18.6\text{b})$$

Правые части этих равенств могут быть выражены также с помощью формул Френеля (16.32а) и (16.43а). При нормальном падении волны на поверхность раздела с помощью (16.33а) и (16.44а) находим

$$\rho_{\text{норм}} = [(n_1 - n_2) / (n_1 + n_2)]^2. \quad (18.7)$$

Оценим значение коэффициента отражения при нормальном падении из воздуха на поверхность стекла, для которого $n_2/n_1 \approx 1,5$. Формула (18.7) показывает, что в этом случае $\rho_{\text{норм}} = 0,04$, т. е. отражается примерно 4% падающей энергии и, следовательно, 96% проходит через стекло.

Коэффициент пропускания. Плотность потока энергии в преломленной волне характеризуется **коэффициентом пропускания**, который определяется аналогично (18.4):

$$\tau = |\mathbf{S}_{\text{пр}} \cdot \mathbf{n}| / |\mathbf{S}_{\text{пад}} \cdot \mathbf{n}|. \quad (18.8)$$

Учитывая, что

$$k_{\text{пад}}n = k_1 \cos \theta_{\text{пад}}, \quad k_{\text{пад}}\tau = -k_1 \sin \theta_{\text{пад}}, \quad k_{\text{пр}}n = k_2 \cos \theta_{\text{пр}}, \quad k_{\text{пр}}\tau = k_2 \sin \theta_{\text{пр}},$$

с помощью этих формул и формулы (18.1) приведем выражение (18.8) к виду

$$\tau = \frac{Z_1 \cos \theta_{\text{пр}}}{Z_2 \cos \theta_{\text{пад}}} \left| \frac{E_{\text{пр}}}{E_{\text{пад}}} \right|^2. \quad (18.9)$$

Из (18.9) с помощью (16.30б) и (16.41б) аналогично (18.6) получаем:

$$\tau_{\perp} = 4Z_1 Z_2 \cos \theta_{\text{пад}} \cos \theta_{\text{пр}} / (Z_2 \cos \theta_{\text{пад}} + Z_1 \cos \theta_{\text{пр}})^2, \quad (18.10\text{a})$$

$$\tau_{\parallel} = 4Z_1 Z_2 \cos \theta_{\text{пад}} \cos \theta_{\text{пр}} / (Z_1 \cos \theta_{\text{пад}} + Z_2 \cos \theta_{\text{пр}})^2. \quad (18.10\text{b})$$

Правые части (18.10) при $\mu_1 = \mu_2$ с помощью формул Френеля могут быть представлены при необходимости как функции только углов падения и преломления. Для нормального падения коэффициент пропускания равен

$$\tau_{\text{норм}} = 4n_1 n_2 / (n_2 + n_1)^2 \quad (\mu_1 = \mu_2). \quad (18.11)$$

Закон сохранения энергии. Прямой проверкой с помощью формул (18.5), (18.6), (18.8) и (18.10) убеждаемся, что коэффициенты отражения и пропускания удовлетворяют соотношению

$$\rho + \tau = 1, \quad (18.12)$$

которое выражает закон сохранения энергии при отражении и преломлении.

Поляризация света при отражении и преломлении. Естественный свет является неполяризованным. Ввиду различия ρ_{\perp} , τ_{\perp} и ρ_{\parallel} , τ_{\parallel} отраженный и преломленный лучи частично поляризованы. Поляризация при отражении была экспериментально обнаружена в 1808 г. Э. Л. Малюсом (1775—1812). Он наблюдал через кристалл исландского шпата двойное лучепреломление (см. § 42) луча солнца, отраженного от поверхности стеклянной пластинки. При вращении пластинки вокруг луча как оси он заметил, что относительные интенсивности двух, возникающих в результате двойного лучепреломления лучей изменяются. Это свидетельствует о частичной поляризации луча солнца при отражении от поверхности стекла. Теоретического объяснения поляризации при отражении Малюс не предложил. Поляризация света при преломлении экспериментально была обнаружена в 1811 г. Э. Л. Малюсом и Ж. Б. Био (1774—1862).

§ 19 | Распространение света в проводящих средах

Излагается метод учета проводимости среды и влияние проводимости на свойства волн.

Комплексная диэлектрическая проницаемость. В однородной проводящей среде проводимость $\gamma \neq 0$ и поэтому уравнения Максвелла (2.47)—(2.49) остаются без изменения, а вместо уравнения (2.46) можно написать

$$\nabla \times \mathbf{B} = \mu \gamma \mathbf{E} + \mu \epsilon \partial \mathbf{E} / \partial t, \quad (19.1)$$

где принят во внимание закон Ома в дифференциальной форме ($\mathbf{j} = \gamma \mathbf{E}$).

Подставляя в уравнения (19.1) и (2.47)—(2.49) выражения (2.50) и (2.51), получаем вместо (2.53) соотношение

$$-\mathbf{k} \times \mathbf{B} = \omega \mu [\epsilon + i\gamma/\omega] \mathbf{E}, \quad (19.2)$$

а уравнения (2.54)—(2.56) остаются без изменения. При $\gamma = 0$ эти уравнения превращаются в уравнения для распространения света в диэлектрике (см. § 15). Как видно из (19.2), все формулы для распространения света в диэлектриках превращаются в формулы для распространения света в проводящих средах, если в них диэлектрическую проницаемость ϵ заменить на комплексную диэлектрическую проницаемость:

$$\epsilon_{\omega} = \epsilon + i\gamma/\omega. \quad (19.3)$$

В диэлектрической среде между волновым числом k , частотой ω и диэлектрической проницаемостью ϵ существует соотношение (15.4), которое удобно записать в виде

$$k^2 = \omega^2 \epsilon \mu, \quad (19.4)$$

110 где $\epsilon\mu = 1/\nu^2$. Для проводящей среды ϵ заменяется на ϵ_ω , k на k_ω , определяемые в соответствии с (19.4) из равенства

3 $k_\omega^2 = \omega^2 \epsilon_\omega \mu = \omega^2 \epsilon \mu + i\omega \gamma \mu$. (19.5)

Представив k_ω в виде комплексного числа

$$k_\omega = k + is, \quad (19.6)$$

равенство (19.5) можно переписать:

$$k^2 + 2iks - s^2 = \omega^2 \epsilon \mu + i\omega \gamma \mu. \quad (19.7)$$

Приравнивая между собой действительные и мнимые части этого равенства, находим:

$$k^2 - s^2 = \omega^2 \epsilon \mu, \quad (19.8)$$

$$2ks = \omega \gamma \mu. \quad (19.9)$$

Решение данной системы уравнений имеет вид

$$k^2 = (\omega^2 \epsilon \mu / 2) (\sqrt{1 + (\gamma/\epsilon\omega)^2} + 1), \quad (19.10)$$

$$s^2 = (\omega^2 \epsilon \mu / 2) (\sqrt{1 + (\gamma/\epsilon\omega)^2} - 1). \quad (19.11)$$

Глубина проникновения. При распространении в проводящей среде векторы поля плоской волны даются формулами (2.50) и (2.51), но с заменой k на k_ω , т. е. для распространения в направлении положительных значений оси Z представляются в виде

$$\mathbf{E} = E_0 e^{-i(\omega t - k_\omega z)} = E_0 e^{-sz} e^{-i(\omega t - kz)}, \quad (19.12)$$

$$\mathbf{B} = B_0 e^{-i(\omega t - k_\omega z)} = B_0 e^{-sz} e^{-i(\omega t - kz)}, \quad (19.13)$$

где для k_ω использовано выражение (19.6). Таким образом, амплитуда плоской волны в процессе распространения уменьшается, т. е. волна распространяется с поглощением. По общему правилу глубина проникновения

$$\Delta = 1/s. \quad (19.14)$$

На больших глубинах волна практически отсутствует.

Для видимого диапазона и хороших проводников формула (9.11) значительно упрощается. Проводимость металлов $\gamma \approx 10^7$ См/м, а ϵ может быть принята равной ϵ_0 . Следовательно, для этого случая

$$\gamma/(\epsilon\omega) \approx 2 \cdot 10^2 \gg 1. \quad (19.15)$$

С учетом (9.15) в формуле (19.11) можно пренебречь единицей по сравнению с $\gamma/(\epsilon\omega)$ и записать выражение для s в виде

$$s = \sqrt{\omega \mu / 2}. \quad (19.16)$$

Следовательно, глубина проникновения равна

$$\boxed{\Delta = \sqrt{2/(\omega \mu)}}. \quad (19.17)$$

Поскольку $\omega = 2\pi/(\lambda \sqrt{\epsilon\mu})$, формулу (19.17) можно также представить в виде

$$\Delta = \sqrt{\frac{\lambda}{\pi \gamma}} \left(\frac{\epsilon_0}{\mu_0} \right)^{1/4} \left(\frac{\epsilon_r}{\mu_r} \right)^{1/4}, \quad (19.18)$$

где $(\mu_0/\epsilon_0)^{1/2} = 377$ Ом.

Рассмотрим в качестве примера медь, для которой $\gamma = 5 \cdot 10^7 \text{ См/м}$. Для длины волн порядка световых ($\lambda \approx 1 \text{ мкм}$) глубина проникновения $\Delta \approx 4 \text{ нм}$, т. е. составляет тысячные доли длины волны. В этом случае говорить о распространении света в проводящей нет смысла. Однако если проводимость среды не очень велика, то свет может распространяться на довольно большие расстояния. Амплитуда волн при этом постепенно затухает.

Физическая причина поглощения. Физической причиной затухания электромагнитных волн в проводящей среде является преобразование электромагнитной энергии волн в джоулеву теплоту: электрическая напряженность волны возбуждает в проводящей среде точки проводимости, которые по закону Джоуля—Ленца нагревают вещество среды.

Фазовая скорость и длина волны. Из (19.12) и (19.13) с учетом (19.10) для фазовой скорости находим выражение

$$v = \frac{\omega}{k} = \frac{1}{\sqrt{\mu\epsilon}} \left\{ \frac{2}{[1 + (\gamma/\omega\epsilon)^2]^{1/2} + 1} \right\}^{1/2}. \quad (19.19)$$

Эта скорость меньше скорости света в непроводящей среде с тем же значением ϵ (при одинаковых магнитных проницаемостях), т. е. наличие в среде проводимости уменьшает фазовую скорость. Длина волны в проводящей среде равна

$$\lambda = \frac{2\pi}{k} = \frac{2\pi}{\omega\sqrt{\mu\epsilon}} \left\{ \frac{2}{[1 + (\gamma/\omega\epsilon)^2]^{1/2} + 1} \right\}^{1/2}, \quad (19.20)$$

т. е. уменьшается по сравнению с длиной волны в непроводящей среде с теми же значениями ϵ и магнитной проницаемости.

Из (19.19) видно, что скорость волны в проводящей среде зависит от частоты, т. е. наблюдается явление дисперсии. Поэтому распространение немонокроматических волн в ней происходит, как в диспергирующей среде.

Соотношения между фазами колебаний векторов поля. Комплексную величину k_ω целесообразно представить в экспоненциальной форме:

$$k_\omega = |k_\omega| e^{i\varphi}, \quad |k_\omega| = \sqrt{k^2 + s^2}, \quad (19.21)$$

где $\operatorname{tg} \varphi = s/k$ [см. (19.6)]. Соотношение (2.54) может быть записано в виде

$$\mathbf{B} = \frac{|k_\omega|}{\omega} e^{i\varphi} \mathbf{k}^{(0)} \times \mathbf{E}, \quad (19.22)$$

где $\mathbf{k}^{(0)}$ — единичный вектор в направлении распространения волны, в данном случае в направлении оси Z . Векторы \mathbf{E} и \mathbf{B} перпендикулярны этой оси.

Напряженность электрического поля волны задается формулой (19.12), в которой без ограничения общности E_0 можно считать действительной величиной, поскольку выбор начала отсчета времени всегда произведен. Подставляя (19.12) в (19.22), приходим к равенству

$$\mathbf{B} = \frac{|k_\omega|}{\omega} \mathbf{k}^{(0)} \times \mathbf{E}_0 e^{-sz} e^{-i(\omega t - kz - \varphi)}. \quad (19.23)$$

Взяв действительные части выражений (19.12) для \mathbf{E} и (9.23) для \mathbf{B} , найдем (обозначая действительные \mathbf{E} и \mathbf{B} теми же буквами) соотношения

$$\mathbf{E} = \mathbf{E}_0 e^{-sz} \cos(\omega t - kz),$$

$$\mathbf{B} = \frac{|k_\omega|}{\omega} \mathbf{k}^{(0)} \times \mathbf{E}_0 e^{-sz} \cos(\omega t - kz - \varphi), \quad (19.24)$$

которые показывают, что электрический и магнитный векторы волны колеблются с разностью фаз ϕ . Из (19.21) следует, что

$$\operatorname{tg} \phi = s/k = \sqrt{1 + (\omega/\gamma)^2} - \sqrt{\omega/\gamma}, \quad (19.25)$$

т. е. угол ϕ положителен. Это означает, что фаза колебаний \mathbf{B} отстает от фазы колебаний вектора \mathbf{E} .

Соотношения между амплитудами векторов поля. Из (19.12) и (19.23) следует, что

$$|\mathbf{B}|/|\mathbf{E}| = |k_\omega|/\omega = \sqrt{\mu\epsilon} [1 + (\gamma/(\omega\epsilon))^2]^{1/4}. \quad (19.26)$$

Сравнив это соотношение с (15.5) ($v = 1/\sqrt{\mu\epsilon}$), видим, что в проводящей среде значение B относительно E больше, чем в непроводящей с теми же диэлектрической и магнитной проницаемостями.

Среды с малой электропроводимостью. В «плохом» проводнике $\gamma \rightarrow 0$ и, следовательно, $\omega\epsilon/\gamma \rightarrow \infty$. Все формулы, в которые входит $\gamma/(\omega\epsilon) \ll 1$, можно аппроксимировать разложением в ряд по $\gamma/(\omega\epsilon)$, ограничившись первым неисчезающим членом разложения:

$$k = \omega \sqrt{\mu\epsilon} [1 + \frac{1}{8} (\frac{\gamma}{\omega\epsilon})^2], \quad s = \frac{\gamma}{2} \sqrt{\frac{\mu}{\epsilon}}, \quad (19.27)$$

$$\operatorname{tg} \alpha = -\frac{1}{2} \frac{\gamma}{\omega\epsilon}, \quad |k_\omega| = \omega \sqrt{\mu\epsilon} [1 + \frac{1}{4} (\frac{\gamma}{\omega\epsilon})^2], \quad (19.28)$$

$$v = \frac{1}{\sqrt{\mu\epsilon}} [1 - \frac{1}{8} (\frac{\gamma}{\omega\epsilon})^2], \quad \lambda = \frac{2\pi}{\omega \sqrt{\mu\epsilon}} [1 - \frac{1}{8} (\frac{\gamma}{\omega\epsilon})^2], \quad (19.29)$$

$$\frac{|\mathbf{B}_0|}{|\mathbf{E}_0|} = \sqrt{\mu\epsilon} [1 + \frac{1}{4} (\frac{\gamma}{\omega\epsilon})^2]. \quad (19.30)$$

При $\gamma=0$ эти формулы переходят в соответствующие равенства для диэлектрика и в очень ясной форме представляют влияние проводимости среды на характеристики волны. Видно, что наличие проводимости среды приводит к увеличению волнового числа, появлению сдвига фаз между векторами поля и затуханию волны, уменьшению фазовой скорости и длины волны, увеличению $|B|$ относительно $|E|$.

Среды с большой электропроводимостью. В «хорошем» проводнике $\gamma \rightarrow \infty$ и, следовательно, $\omega\epsilon/\gamma \rightarrow 0$. В результате разложения в ряд соответствующих выражений получаем:

$$k = \sqrt{\frac{\omega\gamma\mu}{2}} \left(1 + \frac{1}{2} \frac{\omega\epsilon}{\gamma}\right), \quad s = \sqrt{\frac{\omega\gamma\mu}{2}} \left(1 - \frac{1}{2} \frac{\omega\epsilon}{\gamma}\right), \quad (19.31)$$

● **Электропроводимость среды приводит к поглощению в среде энергии волны.** Электропроводимость среды учитывается минимой частью комплексной диэлектрической проницаемости.

Длина волны в проводящей среде уменьшается по сравнению с длиной волны в непроводящей среде с теми же диэлектрическими и магнитными проницаемостями.

Колебания вектора напряженности электрического поля и вектора индукции магнитного поля волны в проводящей среде происходят с разностью фаз.

○ **В чем состоит физическая причина возникновения разности фаз колебаний вектора напряженности электрического поля и вектора индукции магнитного поля волны в проводящей среде?**

Какими физическими факторами обуславливается уменьшение длины волны в проводящей среде по сравнению с длиной волны в непроводящей?

$$\operatorname{tg} \varphi = \frac{\omega \epsilon}{\gamma} - 1, \quad |k_\omega| = \sqrt{\omega \gamma \mu}, \quad (19.32)$$

$$v = \sqrt{\frac{2\omega}{\gamma\mu}} \left(1 - \frac{1}{2} \frac{\omega\epsilon}{\gamma}\right), \quad \lambda = 2\pi \sqrt{\frac{2}{\omega\gamma\mu}} \left(1 - \frac{\omega\epsilon}{\gamma}\right). \quad (19.33)$$

$$\frac{|B_0|}{|E_0|} = \sqrt{\frac{\gamma\mu}{\omega}} = \frac{\sqrt{\mu\epsilon}}{\sqrt{\omega\epsilon/\gamma}}. \quad (19.34)$$

Видно, что разность фаз между векторами поля в хорошем проводнике близка к $-\pi/4$, а фазовая скорость волны в «хорошем» проводнике значительно меньше фазовой скорости волны в диэлектрике. Затухание волн в хорошем проводнике очень большое.

Пример 19.1. Волна с круговой частотой $\omega = 2,5 \cdot 10^{15}$ с⁻¹ распространяется в проводящей среде. Установлено, что интенсивность волны в среде убывает в e раз на пути длиной $l = 20$ см, а длина волны в среде равна $\lambda = 0,5$ мкм. Найти показатель преломления среды и ее удельную электропроводимость.

Показатель преломления дается формулой

$$n = c/v = \lambda_0/\lambda = 2\pi c/(\omega\lambda) = 1,5.$$

Принимая во внимание закон (19.12), получаем

$$2sl = 1, \quad (19.35)$$

где s определено равенством (19.11). Разделив обе части (19.11) на $(\omega \sqrt{\epsilon_0 \mu_0})^2 = (\omega/c)^2$, получим

$$[c/(2\omega)]^2 = (\epsilon_r/2) [\sqrt{1 + [\gamma/(\epsilon\omega)]^2} - 1], \quad (19.36)$$

где $\epsilon_r = \epsilon/\epsilon_0$ — относительная диэлектрическая проницаемость среды, связанная с показателем преломления соотношением $n = \sqrt{\epsilon_r}$ [среда предполагается немагнитной ($\mu = \mu_0$)]. Учитывая, что $c/(2\omega) = 3 \cdot 10^8 / (2 \cdot 0,2 \cdot 10^{15} \cdot 2,5) = 3 \cdot 10^{-7} \ll 1$, из (19.36) получаем неравенство $[\gamma/(\epsilon\omega)]^2 \ll 1$, на основании которого правую часть (19.36) можно разложить в ряд по $[\gamma/(\epsilon\omega)]^2$, ограничившись первым неисчезающим членом. В результате имеем

$$\frac{c}{2\omega} = \frac{\sqrt{\epsilon_r}}{2} \frac{\gamma}{\epsilon\omega} = \frac{1}{2} \frac{\gamma}{\epsilon_0\omega} \frac{1}{\sqrt{\epsilon_r}} = \frac{1}{2} \frac{\gamma}{\epsilon_0\omega} \frac{1}{n}.$$

Следовательно,

$$\gamma = \epsilon_0 c n / l = 8,85 \cdot 10^{-12} \cdot 3 \cdot 10^8 \cdot 1,5 / 0,2 \text{ См/м} = 0,02 \text{ См/м}. \quad (19.37)$$

§ 20 Отражение света от поверхности проводника

Излагается способ учета проводимости среды при отражении.

Границные условия. При падении на поверхность проводника волна частично проникает в проводник и там поглощается, а частично отражается. Для упрощения формул ограничимся рассмотрением случая нормального падения.

Диэлектрическую и магнитную проницаемости среды, из которой на поверхность проводника падает световая волна, обозначим ϵ_1, μ_1 , а проводника — ϵ_2, μ_2 (диэлектрическая среда и проводник считаются немагнитными). Электропроводимость проводника γ . Ориентировки системы координат и векторов, характеризующих волны, указаны на рис. 68. Напряженности электрического поля волн можно записать в виде

$$\mathbf{E}_{\text{пд}} = \mathbf{i}_x E_{\text{пд}}^{(0)} e^{-i(\omega t - k_1 z)}, \quad \mathbf{E}_{\text{от}} = \mathbf{i}_x E_{\text{от}}^{(0)} e^{-i(\omega t + k_1 z)}, \quad (20.1)$$

$$\mathbf{E}_{\text{пр}} = \mathbf{i}_x E_{\text{пр}}^{(0)} e^{-i(\omega t - k_2 \omega z)},$$

где \mathbf{i}_x — единичный вектор в направлении положительных значений оси X , а значения

$$k_1 = n_1 \omega / c, \quad k_{2\omega} = k_2 + i s_2 \quad (20.2)$$

даются соотношениями (15.4) ($v = c/n_1$), (19.6), (19.10) и (19.11). Индукции соответствующих магнитных полей находятся из формулы (16.22), записанной в виде

$$\mathbf{B} = \mathbf{k} \times \mathbf{E} / \omega. \quad (20.3)$$

С учетом ориентировки векторов (рис. 68) на основании (20.3) имеем

$$\begin{aligned} \mathbf{B}_{\text{пд}} &= \mathbf{i}_y k_1 E_{\text{пд}} / \omega, & \mathbf{B}_{\text{от}} &= -\mathbf{i}_y k_1 E_{\text{от}} / \omega, \\ \mathbf{B}_{\text{пр}} &= \mathbf{i}_y k_{2\omega} E_{\text{пр}} / \omega, \end{aligned} \quad (20.4)$$

где \mathbf{i}_y — единичный вектор в направлении положительных значений оси Y .

Непрерывность тангенциальных составляющих напряженности электрического и магнитного полей выражается двумя уравнениями:

$$E_{\text{пд}}^{(0)} + E_{\text{от}}^{(0)} = E_{\text{пр}}^{(0)}, \quad k_1 (E_{\text{пд}}^{(0)} - E_{\text{от}}^{(0)}) / \mu_1 = k_{2\omega} E_{\text{пр}}^{(0)} / \mu_2. \quad (20.5)$$

Соотношения между амплитудами волн. Решая уравнения (20.5) относительно неизвестных отношений $E_{\text{от}}^{(0)}/E_{\text{пд}}^{(0)}$ и $E_{\text{пр}}^{(0)}/E_{\text{пд}}^{(0)}$, находим:

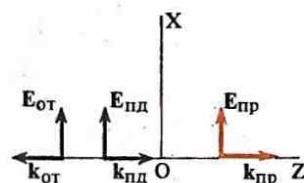
$$\frac{E_{\text{от}}^{(0)}}{E_{\text{пд}}^{(0)}} = \frac{1 - (k_{2\omega}/k_1)(\mu_1/\mu_2)}{1 + (k_{2\omega}/k_1)(\mu_1/\mu_2)}, \quad (20.6)$$

$$\frac{E_{\text{пр}}^{(0)}}{E_{\text{пд}}^{(0)}} = \frac{2}{1 + (k_{2\omega}/k_1)(\mu_1/\mu_2)}. \quad (20.7)$$

Поскольку $k_{2\omega}$ — комплексная величина, из (20.6) и (20.7) следует, что фазы всех волн сдвинуты друг относительно друга. Сдвиг фаз можно найти из (20.6) и (20.7), представив их правые части в виде произведения модуля величины на фазовый множитель аналогично (19.21).

Коэффициент отражения. Из (20.6) в соответствии с определением (18.5) находим выражение для коэффициента отражения:

$$\rho = \left| \frac{E_{\text{от}}^{(0)}}{E_{\text{пд}}^{(0)}} \right| = \frac{[1 - \mu_1 k_2 / (\mu_2 k_1)]^2 + [\mu_1 s_2 / (\mu_2 k_1)]^2}{[1 + \mu_1 k_2 / (\mu_2 k_1)]^2 + [\mu_1 s_2 / (\mu_2 k_1)]^2}, \quad (20.8)$$



Ориентировка векторов поля при анализе отражения света от поверхности проводника

При отражении от проводящей поверхности между параллельными и перпендикулярными компонентами векторов поля волны возникает разность фаз.

Какими факторами определяется поляризация волн, отраженных от проводящей поверхности? К какому виду относится эта поляризация? Как коэффициент отражения зависит от электропроводимости и каково его значение у хороших проводников?

где k_2 и s_2 определяются формулами (19.10) и (19.11). Разделив числитель и знаменатель формулы (20.8) на $\mu_1 s_2 / (\mu_2 k_1)$, преобразуем ее к виду

$$\rho = \frac{\{[\mu_2 k_1 / (\mu_1 s_2)] - (k_2 / s_2)\}^2 + 1}{\{[\mu_2 k_1 / (\mu_1 s_2)] + (k_2 / s_2)\}^2 + 1}. \quad (20.9)$$

Из (19.10) и (19.11) видно, что для идеальных проводников $\gamma \rightarrow \infty$, $s_2 \rightarrow \infty$, $k_2 / s_2 \rightarrow 1$, $k_1 / s_2 \rightarrow 0$ и, следовательно, из (20.9) получаем

$$\rho \rightarrow 1; \quad (20.10)$$

следовательно, у очень хороших проводников коэффициент отражения близок к единице.

Связь между отражательной и поглощающей способностями. Из формулы (19.17) следует, что чем больше проводимость, т. е. чем лучше проводник, тем сильнее он поглощает свет. А из формулы (20.10) видно, что чем больше проводимость, т. е. чем лучше проводник, тем он лучше отражает свет. Поэтому получается, что чем лучше свет отражается проводником, тем лучше он им и поглощается. Это приводит к различию в цвете проводника, если его наблюдать в отраженном и в проходящем свете (см. § 15).

Задачи

- 3.1. Найти выражение для объемной плотности энергии плоской волны в проводнике, распространяющейся в направлении оси Z .
- 3.2. Найти выражения для коэффициента преломления очень хорошего проводника.
- 3.3. Найти отношение $\langle w_m \rangle / \langle w_s \rangle$ объемной плотности энергии магнитного поля волны к объемной плотности энергии электрического поля в проводящей среде.
- 3.4. Идеальный проводник, плоская поверхность которого перпендикулярна оси Z , движется в положительном направлении оси Z со скоростью v . В том же направлении распространяется плоская электромагнитная волна, вектор напряженности электрического поля которой коллинеарен оси Y ($E_x = 0$, $E_y = E_0 \cos \omega(t - z/c)$, $E_z = 0$). Найти напряженность электрического поля волны, образованной в результате суперпозиции падающей и отраженной волн ($v \ll c$).
- 3.5. Естественный свет падает на поверхность стекла ($n = 1,5$) под углом Брюстера. Определить степень поляризации преломленного луча [см. (16.45)].
- 3.6. Найти степень поляризации узкого пучка лучей естественного света после прохождения плоскопараллельной пластины из стекла ($n = 1,5$), если пучок падает на пластину под углом Брюстера.
- 3.7. Угол между плоскостью колебаний вектора напряженности электрического поля падающей на поверхность раздела электромагнитной волны и плоскостью падения называют азимутом колебаний. Найти азимуты колебаний $a_{\text{от}}$ и $a_{\text{пр}}$ в отраженной и преломленной волнах, если азимут колебаний в падающей волне $a_{\text{пад}}$, а угол падения $\theta_{\text{пад}}$.
- 3.8. Найти относительное отклонение групповой скорости от фазовой в среде с $n = 1,5$, у которой для $\lambda = 0,5 \text{ мкм}$ $dn/d\lambda = 3 \cdot 10^4 \text{ м}^{-1}$.
- 3.9. На поверхность стекла ($n = 1,5$) под углом 40° из воздуха падает линейно поляризованная волна, вектор напряженности электрического поля которой перпендикулярен плоскости падения. Найти коэффициенты отражения и пропускания.
- 3.10. На поверхность стекла ($n = 1,6$) под углом 25° из воздуха падает линейно поляризованная волна, электрический вектор которой колеблется в плоскости падения. Определить коэффициенты отражения и пропускания.
- 3.11. На поверхность стекла ($n = 1,65$) под углом 35° из воздуха падает линейно поляризованная волна, электрический вектор которой колеблется в плоскости, образующей угол 30° с плоскостью падения. Найти коэффициенты отражения и пропускания.
- 3.12. Естественный свет падает из стекла с $n = 1,65$ под углом 40° на границу с некоторым раствором, показатель преломления n_2 которого зависит от концентрации растворенного вещества и может изменяться в широких пределах. При каком показателе преломления n_2 отраженный свет линейно поляризован и каков при этом коэффициент отражения?

- 3.13.** Из стекла ($n = 1,55$) на границу раздела с воздухом под углом 60° падает световая волна. Найти критический угол и сдвиги фаз колебаний напряженности электрического поля φ_1 [см. (17.19)].
- 3.14.** Найти область углов падения линейно поляризованной волны из воздуха на поверхность воды ($n = 1,33$), при которой коэффициент отражения больше 0,5. Плоскость колебаний электрического вектора волны перпендикулярна плоскости падения.
- 3.15.** Естественный свет падает под углом Брюстера на стеклянную пластину ($n = 1,65$). Найти коэффициент отражения.
- 3.16.** Вывести формулу для сдвига фаз колебаний перпендикулярной и лежащей в плоскости падения компонент электрического поля волны при полном отражении.
- 3.17.** Используя условие примера 19.1, найти сдвиг фаз между колебаниями напряженности электрического поля и индукции магнитного поля волны.
- 3.18.** При какой удельной электропроводимости затухание в среде ($n = 1,5$) равно 1 дБ/км?
- 3.19.** При 20°C показатель преломления бензола для длины волны 589 нм составляет 1,5, а плотность равна $0,879 \text{ г/см}^3$. Пользуясь формулой Клаузиуса — Моссotti, рассчитать показатель преломления паров бензола при температуре кипения (80°C).

Ответы

- 3.1.** $[k^2/(2\mu\omega^2)]|E_0(0)|^2 \exp(-2sz)$. **3.2.** $\mu, \gamma/(2\omega\epsilon_0)$. **3.3.** $[1 + (\gamma/(\omega\epsilon)^2)]^{1/2}$. **3.4.** $E_x = E_0 \{ \cos \omega(t - z/c) - \cos \omega [(1 - 2v/c)(t - z/c)] \}$, $E_z = 0$. **3.5.** $P = [4n^2 - (1 + n^2)^2]/[4n^2 + (1 + n^2)^2] = -0,08$. **3.6.** $P = [16n^4 - (1 + n^2)^4]/[16n^4 + (1 + n^2)^4] = 0,16$. **3.7.** $\operatorname{tg} \alpha_{np} = \cos(\theta_{nd} - \theta_{np}) \operatorname{tg} \alpha_{nd}$, $\operatorname{tg} \alpha_{otp} = -\cos(\theta_{nd} - \theta_{np}) \operatorname{tg} \alpha_{nd}/\cos(\theta_{nd} + \theta_{np})$. **3.8.** $(v_r - c/n)/(c/n) = (\lambda/n) (dn/d\lambda) = 10^{-2}$. **3.9.** 0,08; 0,92. **3.10.** 0,04; 0,96. **3.11.** 0,05; 0,95. **3.12.** 1,385; 0,03. **3.13.** 40,18°; 50,5°; 70,2°. **3.14.** $\theta_{nd} > 81,2^\circ$. **3.15.** $1/2 [(1 - n^2)/(1 + n^2)]^2 = 0,107$. **3.16.** $\varphi_\perp - \varphi_1 = \arctg [\cos \theta_{nd} \sqrt{\sin^2 \theta_{nd} - n^2} / \sin^2 \theta_{nd}]$. **3.17.** $\varphi = \arctg s/k = \arctg [\gamma/(2\epsilon_0\omega n^2)] = \frac{1}{2} 10^{-7}$. **3.18.** $0,92 \cdot 10^{-6} \text{ См/м}$. **3.19.** $n = 1 + 1,36 \cdot 10^{-3}$.

Основная идея:
в оптическом диапазоне
с достаточно
большой точностью
распространение волн
можно представить
как движение энергии
волн по лучам,
описываемым с помощью
геометрических
соотношений.

Геометрическая оптика



§ 21

Приближение геометрической оптики

Излагается способ перехода от уравнений, описывающих волны, к уравнениям, описывающим лучи, и анализируется понятие луча.

Уравнение эйконала. Любая из компонент амплитуды полей световых волн в вакууме удовлетворяет волновому уравнению (2.12). В среде, скорость распространения электромагнитных волн в которой v , волновое уравнение принимает для любой из компонент вид

$$\nabla^2\Phi - \frac{1}{v^2} \frac{\partial^2\Phi}{\partial t^2} = \nabla^2\Phi - \frac{n^2}{c^2} \frac{\partial^2\Phi}{\partial t^2} = 0, \quad (21.1)$$

где введен коэффициент преломления $n=c/v$ среды относительно вакуума. Для монохроматических волн частотой ω полагаем

$$\Phi(\mathbf{r}, t) = \Psi(\mathbf{r}) e^{-i\omega t} \quad (21.2)$$

и после подстановки (21.2) в (21.1) находим уравнение для зависящей только от координат амплитуды $\Psi(\mathbf{r})$:

$$\nabla^2\Psi + n^2 k_0^2 \Psi = 0, \quad (21.3)$$

где $k_0 = \omega/c$ — волновое число, соответствующее вакууму. Оно, очевидно, связано с волновым числом k в среде соотношением

$$k = nk_0. \quad (21.4)$$

Принимая во внимание соотношение

$$\frac{1}{\Psi} \frac{\partial^2\Psi}{\partial x^2} = \frac{\partial^2}{\partial x^2} (\ln \Psi) + \left(\frac{\partial}{\partial x} \ln \Psi \right)^2 \quad (21.5)$$

и аналогичные соотношения для производных по y и z , запишем уравнение (21.3) после деления на Ψ в виде

$$\frac{1}{\Psi} \nabla^2\Psi + n^2 k_0^2 = \nabla^2(\ln \Psi) + [\text{grad}(\ln \Psi)]^2 + n^2 k_0^2 = 0. \quad (21.6)$$

Для решения этого уравнения полагаем

$$\Psi(\mathbf{r}) = A(\mathbf{r}) e^{iS(\mathbf{r})}. \quad (21.7)$$

Подставляя (21.7) в (21.6), находим

$$\nabla^2(\ln A) + [\text{grad}(\ln A)]^2 - (\text{grad } S)^2 + n^2 k_0^2 + i[\nabla^2 S + 2 \text{grad}(\ln A) \cdot \text{grad } S] = 0. \quad (21.8)$$

Приравнивая к нулю действительную и мнимую части, получаем два уравнения для определения $A(\mathbf{r})$ и $S(\mathbf{r})$:

$$\nabla^2(\ln A) + [\text{grad}(\ln A)]^2 - (\text{grad } S)^2 + n^2 k_0^2 = 0, \quad (\text{a})$$

$$\nabla^2 S + 2 \text{grad}(\ln A) \cdot \text{grad } S = 0. \quad (\text{б}) \quad (21.9)$$

Эти уравнения значительно упрощаются для оптического диапазона длин волн. Амплитуда волны существенно изменяется лишь на расстояниях, много больших длины волны, т. е. на расстояниях l , имеющих порядок размера линз, оптических приборов и т. д., удовлетворяющих условию

$$\lambda \ll l. \quad (21.10)$$

Сравнивая (21.7) с представлением плоской волны в комплексной форме (2.39), видим, что $|\operatorname{grad} S|^2 \sim k^2 \sim 1/\lambda^2$. (21.11)

С другой стороны, сумма первых двух членов в (21.9а) имеет порядок

$$\nabla^2(\ln A) + [\operatorname{grad}(\ln A)]^2 = (1/A)\nabla^2A \sim 1/l^2 \ll 1/\lambda^2. \quad (21.12)$$

Поэтому этими членами можно пренебречь по сравнению с двумя последними и записать уравнение (21.9а) в виде соотношения

$$(\operatorname{grad} S)^2 = n^2 k_0^2, \quad (21.13)$$

называемого **уравнением эйконала**.

Луч света. Градиент от функции S направлен по нормали к поверхности $S = \text{const}$. Поэтому эйконал S описывает поверхности постоянной фазы волны, а $\operatorname{grad} S$ приводит к понятию луча, т. е. к представлению о движении световой энергии в данной точке в определенном направлении. Лучом называется линия, касательная к которой совпадает в каждой точке с вектором $\operatorname{grad} S$.

Распространение света рассматривается как движение световой энергии по лучам. Плоскость, перпендикулярная лучам света (т. е. плоскость $S = \text{const}$), называется волновым фронтом.

Область применимости лучевого приближения. Анализ распространения света в лучевом приближении составляет предмет геометрической оптики. Как видно из (21.12), его справедливость оправдана всегда, когда $\nabla^2 A/A$ является малой относительно $1/\lambda^2$ величиной. Физически этот член описывает искривление световых лучей материальными объектами, т. е. дифракцию световых лучей. Поэтому можно сказать, что в геометрической оптике не учитываются эффекты дифракции (см. гл. 6).

Принцип Ферма. В однородной среде $S = k \cdot r$ ($k = \text{const}$). Лучи являются прямыми параллельными линиями, а фронт волны — плоскостью, перпендикулярной этим линиям. Для неоднородной среды вопрос значительно усложняется.

Пусть точки P_1 и P_2 соединяются лучом L (рис. 69). Вычислим изменение фазы вдоль луча. Для каждой точки луча имеем

$$dS = \operatorname{grad} S \cdot d\mathbf{r} = |\operatorname{grad} S| |d\mathbf{r}| = k_0 n(\mathbf{r}) dl, \quad (21.14)$$

где учтено, что $d\mathbf{r}$ направлен по лучу и, следовательно, совпадает с $\operatorname{grad} S$, dl — элемент длины пути. Таким образом, для изменения фазы находим

$$S = \int_{P_1}^{P_2} dS = k_0 \int_{P_1}^{P_2} n(\mathbf{r}) dl, \quad (21.15)$$

причем путь P_1P_2 совпадает с лучом.

Из (21.15) с учетом определения луча и волнового фронта следует, что оптические длины путей вдоль различных лучей между точками волнового фронта в два момента времени одинаковы. Если точки P_1 и P_2 соединить другой, отличной от луча кривой, все точки которой будут находиться в непосредственной близости от луча, и вычислить интеграл (21.15), то получится, конечно, другое значение интеграла. Принцип Ферма утверждает, что интеграл (21.15) вдоль луча имеет стационарное значение, т. е. первая вариация δS относительно соседних путей интегрирования равна нулю. Эту формулировку можно изменить. Учтем, что $dl/v = dt$ есть время прохождения пути dl со скоростью v , и примем во внимание соотношение $n(\mathbf{r}) = c/v(\mathbf{r})$. Тогда равенство (21.15) принимает вид

$$S = k_0 c \int_{P_1}^{P_2} \frac{dl}{v} = \omega \int_{P_1}^{P_2} \frac{dl}{v}, \quad (21.16)$$

где интеграл

$$\int_{P_1}^{P_2} \frac{dl}{v} \quad (21.17)$$

дает время, затрачиваемое светом на прохождение пути от P_1 до P_2 . Поэтому принцип Ферма может быть высказан в форме утверждения, что лучом, соединяющим две точки, является тот путь, который делает стационарным время, затрачиваемое светом на его прохождение.

Он был открыт П. Фермой (1601—1665) в 1657 г. как «принцип наименьшего времени» в такой формулировке: «Природа всегда следует наикратчайшему пути». Однако это не следует понимать как утверждение, что лучом, соединяющим две точки, является тот путь, на прохождение которого затрачивается меньше времени, чем на прохождение по соседним путям между теми же точками. Утверждение Ферма, что «природа всегда следует наикратчайшему пути», верно, но таких наикратчайших путей, соединяющих две точки, может существовать много. Формулировка о стационарности времени прохождения пути между двумя точками, с одной стороны, утверждает экстремальный характер этого времени (максимальность или минимальность), а с другой стороны, не исключает наличия нескольких путей с одинаковым временем прохождения.

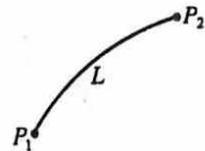
Такая ситуация является типичной для геометрической оптики при построении изображений.

В геометрической оптике все лучи от точки предмета идут по различным путям и встречаются в точке изображения. Но все они затрачивают одно и то же время на прохождение своего пути. Другими словами, оптические длины всех путей, соединяющих точку предмета с точкой изображения, одинаковы. Это утверждение называется принципом таутокронизма.

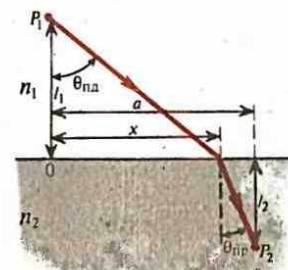
Вывод закона преломления из принципа Ферма. Для иллюстрации применения принципа Ферма выведем с его помощью закон преломления. Пусть требуется соединить лучом две точки P_1 и P_2 , находящиеся в однородных средах с показателями преломления n_1 и n_2 , разделенных плоской границей (рис. 70). В каждой однородной среде луч является прямой линией. Пусть x является координатой входа луча из первой среды во вторую. Полное время распространения света от P_1 к P_2 , очевидно, равно

$$t = (n_1/c) \sqrt{l_1^2 + x^2} + (n_2/c) \sqrt{l_2^2 + (a-x)^2}, \quad (21.18)$$

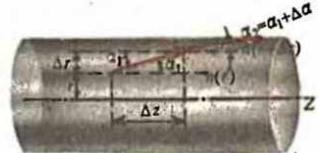
где $v = c/n$, и зависит от переменной x . Условие стационарности $\partial t / \partial x = 0$ принимает вид



Траектория луча света в неоднородной среде



К выводу закона преломления с помощью принципа Ферма



К выводу уравнения луча в световоде

$$n_1 x / (c \sqrt{l_1^2 + x^2}) - n_2 (a - x) / (c \sqrt{l_2^2 + (a - x)^2}) = 0. \quad (21.19)$$

Учитывая, что $x / \sqrt{l_1^2 + x^2} = \sin \theta_{\text{пп}}$, $(a - x) / \sqrt{l_2^2 + (a - x)^2} = \sin \theta_{\text{пп}}$, из (21.19) получаем закон преломления

$$\sin \theta_{\text{пп}} / \sin \theta_{\text{пп}} = n_2 / n_1, \quad (21.20)$$

который, конечно, совпадает с законом Снеллиуса (16.14).

Таким образом, с помощью принципа Ферма, зная закон изменения показателя преломления $n(r)$ в среде, можно построить лучи и тем самым решить задачу о распространении света в среде в тех условиях, когда справедливо приближение геометрической оптики.

Прямолинейное распространение света, отражение и преломление были известны еще древним грекам. Первые систематические описания этих явлений, дошедшие до нас, принадлежат Емпедоклу (490—430 гг. до н. э.) и Евклиду (300 г. до н. э.). Им был известен закон отражения света. Закон преломления света был установлен экспериментально в 1621 г. В. Снеллиусом (1591—1626).

Распространение луча в среде с переменным показателем преломления. Рассмотрим распространение луча в среде, изменение показателя преломления которой аксиально-симметрично относительно оси, которая принимается за ось Z (рис. 71). Луч предполагается распространяющимся в положительном направлении оси Z вблизи оси (параксиальный луч). Расстояние от оси Z обозначается r . Запишем закон Снеллиуса для преломления на бесконечно тонком слое Δr , в котором показатель преломления изменяется от $n(r)$ до $n(r + \Delta r)$:

$$n(r) \cos \alpha_1 = n(r + \Delta r) \cos (\alpha_1 + \Delta \alpha). \quad (21.21)$$

Разлагая $n(r + \Delta r)$ в правой части (21.21) в ряд Тейлора по Δr , ограничиваясь линейным по Δr членом и пользуясь тригонометрической формулой для косинуса суммы двух углов, получаем

$$n(r) \cos \alpha_1 = [n(r) + \Delta r \frac{\partial n}{\partial r}] (\cos \alpha_1 \cos \Delta \alpha - \sin \alpha_1 \sin \Delta \alpha). \quad (21.22)$$

В параксиальном приближении можно принять, что $\sin \Delta \alpha \approx \Delta \alpha$, $\cos \Delta \alpha \approx 1$. Тогда с точностью до величины первого порядка по $\Delta \alpha$ из (21.22) находим

$$\frac{\partial n}{\partial r} = n(r) \frac{\Delta \alpha}{\Delta r} \operatorname{tg} \alpha_1, \quad (21.23)$$

где члены с $\Delta \alpha \Delta r$, являющиеся членами второго порядка малости, отброшены. Поскольку $\operatorname{tg} \alpha_1 = \Delta r / \Delta z$, в параксиальном приближении можем написать

$$\operatorname{tg} \alpha_1 \frac{\Delta \alpha}{\Delta r} = \frac{\Delta r}{\Delta z} \frac{\Delta \alpha}{\Delta r} = \frac{\Delta \alpha}{\Delta z} = \frac{\Delta}{\Delta z} \left(\frac{\Delta r}{\Delta z} \right) \rightarrow \frac{d^2 r}{dz^2}. \quad (21.24)$$

С учетом (21.24) из (21.23) находим уравнение распространения луча:

$$\frac{d^2 r}{dz^2} = \frac{1}{n(r)} \frac{dn(r)}{dr}. \quad (21.25)$$

Пример 21.1. Рассмотреть распространение луча света в тонком диэлектрическом волокне (световое волокно), показатель преломления которого изменяется по закону $n(r) = n_0(1 - ar^2/2)$. Предполагается, что $ar^2/2 \ll 1$ ($a > 0$) на всем сечении волокна.

Уравнение (21.25) с точностью до линейных по r членов принимает вид

$$\frac{d^2r}{dz^2} = -ar. \quad (21.26)$$

Общее решение этого уравнения известно:

$$r(z) = A_1 \cos(\sqrt{a}z) + A_2 \sin(\sqrt{a}z). \quad (21.27)$$

Постоянные A_1 и A_2 определяются начальными условиями. Период этого колебания в направлении оси Z равен $l = 2\pi/\sqrt{a}$ (рис. 72). Световые волокна широко применяются для управления движением световых пучков. Они действуют как световоды. При изгибе волокон, если только радиус кривизны не чрезвычайно мал (порядка длины волны света), световой пучок следует за изгибами волокна. Большим достоинством световых волокон является малая величина потерь энергии при распространении в них световых пучков. Эти потери значительно меньше, чем потери в проводах при передаче соответствующей энергии с помощью переменных токов. Поэтому их выгодно применять для передачи информации. Однако главное преимущество использования света для передачи информации связано с большой частотой света, благодаря чему по световому пучку в световоде можно передать очень большой объем информации. Световод толщиной в человеческий волос в состоянии обеспечить передачу информации, эквивалентную многим сотням телефонных линий. Немаловажными преимуществами световодов являются также их малый диаметр, их изготовление из диэлектрических материалов, не поддающихся коррозии и стойким к другим вредным воздействиям, технологичность изготовления.

Пример 21.2. Уравнение (21.25) справедливо лишь в параксиальных приближениях, когда угол α и отклонение луча от оси Z малы. Рассмотреть распространение луча без этих ограничений, пользуясь непосредственно законом преломления (21.20), если $n = n_0 \sqrt{1 - r^2/a^2}$. Угол между осью r , перпендикулярной оси Z , и лучом обозначить $\beta = \beta(r, z)$. При $z = 0$ луч пересекает ось Z под углом $\beta_0 = \beta(0, 0)$.

Закон преломления (21.20), записанный в виде $n_1 \sin \beta_1 = n_2 \sin \beta_2$, показывает, что при распространении луча $n(r) \sin [\beta(r, z)] = \text{const}$. Отсюда с учетом начальных условий получаем уравнение для определения луча $n \sin \beta = n_0 \sin \beta_0 = \text{const}$, которое с учетом $n = n_0 \sqrt{1 - r^2/a^2}$ имеет вид:

$$\sqrt{1 - r^2/a^2} \sin [\beta(r, z)] = n_0 \sin \beta_0. \quad (21.28)$$

Видно, что $|r| < |a|$, т. е. луч не может удалиться от оси Z больше, чем на $|a|$. Можно показать, что уравнение луча дается формулой $r = a \cos \beta_0 \sin [z/(a \sin \beta_0)]$. Однако это не просто.



Траектория луча в световоде

При достаточно малой длине волны можно пользоваться понятием луча. Волновым фронтом при этом является поверхность, ортогональная семейству лучей.

Принцип Ферма утверждает не минимальность времени движения по лучу, а стационарность этого времени.

Опишите какую-либо ситуацию, когда очевидно, что время движения по лучу от одной точки к другой не будет минимальным, а будет стационарным.

Параксиальное приближение. Рассмотрим прохождение лучей через линзу, не накладывая ограничений на ее толщину. Предположим, что ось Z , в положительном направлении которой распространяется свет, совпадает с осью линзы (рис. 73) и луч света лежит в плоскости XZ . Линза ограничена сферическими поверхностями, радиусы кривизны которых r_1 и r_2 . Величины, относящиеся к сферической поверхности, которую лучи при своем движении встречают первой, обозначим с индексами 1, а второй — с индексами 2. Показатель преломления материала линзы n_2 , а показатель преломления среды, в которой находится линза, n_1 . В расчет входят малые углы, по порядку величины равные углам между лучами и осью Z . Весь расчет проводится в параксиальном приближении, т. е. в предположении достаточной близости лучей к оси Z . Принимая во внимание разложения

$$\begin{aligned}\sin \alpha &= \alpha - \alpha^3/(3!) + \alpha^5/(5!) - \dots, \\ \cos \alpha &= 1 - \alpha^2/(2!) + \alpha^4/(4!) - \dots, \\ \operatorname{tg} \alpha &= \alpha + \alpha^3/3 + 2\alpha^5/15 + \dots,\end{aligned}\tag{22.1}$$

заключаем, что при малых углах α можно считать, что $\sin \alpha \approx \alpha$, $\cos \alpha \approx 1$, $\operatorname{tg} \alpha \approx \alpha$. Например, если угол равен 6° , то $\alpha = 2\pi/6/360 \approx 0,1$ и, следовательно, $\sin \alpha \approx \alpha$ с точностью примерно до 10^{-4} , $\operatorname{tg} \alpha \approx \alpha$ — до 10^{-3} и $\cos \alpha \approx 1$ — до 10^{-2} . Поэтому параксиальное приближение с достаточной точностью дает хорошие результаты в довольно широком интервале условий.

Преломление на сферической поверхности. Закон Снеллиуса для преломления в точке P_1 в параксиальном приближении имеет вид

$$n_1 \theta_1 = n_2 \theta'_1 .\tag{22.2a}$$

Углы наклона α_1 и α'_1 падающего и преломленного в точке P_1 лучей удовлетворяют соотношениям

$$\theta_1 = \alpha_1 + \varphi, \quad (22.2b) \qquad \theta'_1 = \alpha'_1 + \varphi, \quad (22.2b)$$

которые позволяют представить (22.2a) в виде

$$n_1(\alpha_1 + \varphi) = n'_1(\alpha'_1 + \varphi),\tag{22.3}$$

где для сохранения единства обозначений положено $n'_1 = n_2$. Принимая во внимание, что $\varphi = x_1/r_1$ (параксиальное приближение), вместо (22.3) получаем

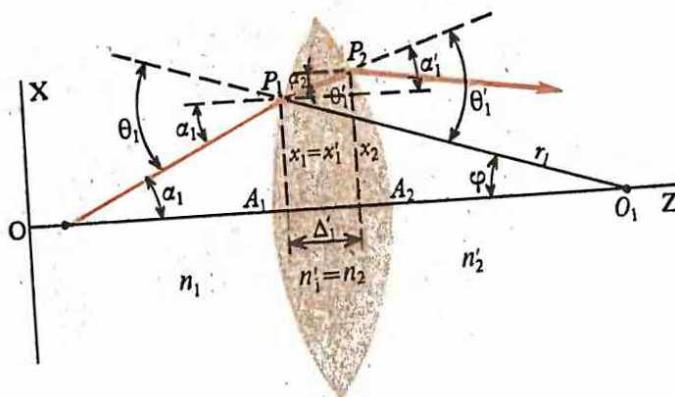
$$n_1(\alpha_1 + x_1/r_1) = n'_1(\alpha'_1 + x_1/r_1),\tag{22.4}$$

или

$$n'_1 \alpha'_1 = (n_1 - n'_1)(x_1/r_1) + n_1 \alpha_1.\tag{22.5}$$

Соотношение (22.5) позволяет найти угол наклона преломленного луча α'_1 , если известны угол наклона падающего луча α_1 и расстояние x_1 от оси до точки падения. Расстояние x'_1 для преломленного луча в точке P_1 равно, очевидно, x_1 , т. е. $x'_1 = x_1$.

Матричные обозначения. Для упрощения записи удобно формулу (22.5) совместно с равенством $x'_1 = x_1$ представить в матричном виде:



73.
К расчету прохождения лучей че-
рез линзу

$$\begin{pmatrix} n'_1 \alpha'_1 \\ x'_1 \end{pmatrix} = \begin{pmatrix} 1 & -k_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} n_1 \alpha_1 \\ x_1 \end{pmatrix}, \quad (22.6)$$

где

$$k_1 = (n_1 - n'_1)/r_1,$$

k_1 — преломляющая сила поверхности 1. Стоящая в (22.6) двухрядная матрица

$$R_1 = \begin{pmatrix} 1 & -k_1 \\ 0 & 1 \end{pmatrix}, \quad (22.8)$$

называется преломляющей матрицей поверхности 1.

Распространение луча в линзе. Преломленный луч пересекает вторую поверхность на расстоянии

$$x_2 = x'_1 + \Delta \operatorname{tg} \alpha'_1 = x'_1 + \Delta \alpha'_1 \quad (22.9)$$

от оси. В параксиальном приближении можно считать Δ равным толщине $A_1 A_2$ линзы, а углу наклона α'_1 , если угол наклона α_2 луча во второй среде ($n'_1 = n_2$) является просто изменением обозначений). Следовательно, (22.9) в матричной форме с учетом того, что $n'_1 = n_2$, $\alpha'_1 = \alpha_2$, имеет вид

$$\begin{pmatrix} n_2 \alpha_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{\Delta_1}{n'_1} & 1 \end{pmatrix} \begin{pmatrix} n'_1 \alpha'_1 \\ x'_1 \end{pmatrix} \quad (22.10)$$

а двухрядная матрица

$$T_{21} = \begin{pmatrix} 1 & 0 \\ \Delta_1/n'_1 & 1 \end{pmatrix} \quad (22.11)$$

называется передаточной матрицией T_{21} . Она описывает распространение луча от первой преломляющей поверхности до второй внутри линзы. Отметим, что величины $n_2 \alpha_2, x_2$ описывают луч в точке второй поверхности до преломления.

Преломление луча на второй сферической поверхности. Оно описывается аналогично (22.6) с помощью преломляющей

Матричный метод позволяет следить за движением лучей при прохождении через оптическую систему и хорошо приспособлен для расчетов на электронных вычислительных машинах.

- Выпишите преломляющую матрицу сферической поверхности.
- Как строится передаточная матрица?
- Как образуется матрица, описывающая отражение от сферической поверхности?

матрицы второй поверхности:

$$R_2 = \begin{pmatrix} 1 & -k_2 \\ 0 & 1 \end{pmatrix}, \quad (22.12)$$

где

$$k_2 = (n'_2 - n_2)/r_2, \quad (22.13)$$

причем r_2 берется со знаком минус, поскольку центр кривизны этой поверхности находится от нее в направлении отрицательных значений оси Z (эта поверхность для падающего на нее луча не выпуклая, а вогнутая). Следовательно, для преломления на второй поверхности можем написать

$$\begin{pmatrix} n'_2 a'_2 \\ x'_2 \end{pmatrix} = R_2 \begin{pmatrix} n_2 a_2 \\ x_2 \end{pmatrix}. \quad (22.14)$$

Преломление луча линзой. Комбинируя (22.6), (22.10) и (22.14), получаем связь между характеристиками луча на входе в линзу и на выходе из нее в виде

$$\begin{pmatrix} n'_2 a'_2 \\ x'_2 \end{pmatrix} = R_2 T_{21} R_1 \begin{pmatrix} n_1 a_1 \\ x_1 \end{pmatrix}. \quad (22.15)$$

Распространение луча через оптическую систему. Распространение луча справа от линзы описывается передаточной матрицей T_{23} , которая строится аналогично (22.11), но вместо Δ' в нее входит Δ'_2 — расстояние от точки A_2 (рис. 73) до точки оси Z , в плоскости которой мы хотим определить параметры луча. Если луч на своем пути встречает другую линзу, то преломление на первой поверхности этой линзы описывается с помощью преломляющей матрицы этой поверхности и т. д. Таким образом, расчет распространения луча через оптическую систему сводится к перемножению матриц, выражающих преломляющие силы поверхностей линз и передаточных матриц. При этом необходимо помнить о знаках: если встречаемая лучом преломляющая поверхность выпуклая, то ее радиус кривизны надо брать с положительным знаком, а если вогнутая — с отрицательным; углы a , отсываемые от оси Z против часовой стрелки, положительны, а по часовой стрелке — отрицательны; расстояния, отсываемые слева направо, положительны, а справа налево — отрицательны; расстояния от оси Z , отсываемые вверх, положительны, вниз — отрицательны. Напомним также, что в качестве Δ линзы берется ее толщина $A_1 A_2$ на оси.

Отражение от сферических поверхностей. Отражение от сферических поверхностей рассматривается как преломление в среду с отрицательным показателем преломления $-n$, если n — показатель преломления среды, из которой луч падает на отражающую поверхность. В остальном матрица, описывающая отражение, полностью аналогична матрице, описывающей преломление. Например, отражение от вогнутой поверхности сферического зеркала с радиусом кривизны r_2 в среду с показателем преломления n_2 описывается матрицей вида (22.12) с $n'_2 = -n_2$, т. е.

$$R_2 = \begin{pmatrix} 1 & -(-n_2 - n_2)/r_2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2n_2/r_2 \\ 0 & 1 \end{pmatrix}. \quad (22.16)$$

Поскольку центр кривизны поверхности расположен в направлении отрицательных значений z , то $r_2 = -|r_2|$. После отражения луч движется в направлении отрицательных значений оси Z . Это учитывается в следующей передаточной матрице тем, что длины пути и показатель преломления берутся с отрицательным знаком. Если на пути луча встречается преломляющая поверхность, то показатель преломления среды, из которой на эту поверхность падает луч, по-прежнему считается отрицательным, показатель преломления среды, в которую входит преломленный

луч, берется с отрицательным знаком (поскольку после преломления луч продолжает двигаться в направлении отрицательных значений по оси Z), а знак радиуса кривизны поверхности определяется обычным правилом. Например, если после отражения от вогнутой поверхности, которое описывается матрицей (22.16), луч света, двигаясь в направлении отрицательных значений z , встречает вогнутую поверхность с радиусом кривизны r_1 (поверхность вогнута, если смотреть в направлении отрицательных значений оси Z), то r_1 берется с положительным знаком по обычному правилу, поскольку ее центр кривизны расположен в направлении положительных значений оси Z от поверхности. Если показатель преломления второй среды n_1 , то матрица преломления имеет вид

$$R_3 = \begin{pmatrix} 1 & -\frac{n_1 - (-n_2)}{r_1} \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{n_1 - n_2}{r_1} \\ 0 & 1 \end{pmatrix}. \quad (22.17)$$

Таким образом, отражение от зеркал анализируется матричным методом аналогично преломлению. Надо лишь внимательно следить за знаками величин, которые входят в матрицы отражения. Оптические системы, в которые входят зеркала, рассчитываются при этом по общим правилам матричного метода.

Пример 22.1. Имеется двояковыпуклая линза, одна из сферических поверхностей которой посеребрена и является отражающей. Для определенности считаем, что у линзы, изображенной на рис. 75, посеребрена правая сферическая поверхность радиусом r_2 . Линза находится в воздухе ($n = 1$), показатель преломления вещества линзы $n_2 > 1$. Радиусы кривизны поверхностей r_1 и r_2 (r_2 , по общему правилу, отрицательная величина, т. е. $r_2 = -|r_2|$). Луч света падает слева. Найти передаточную матрицу от входа луча в линзу до выхода из линзы через ту же поверхность.

В соответствии с изложенными правилами матрица преломления на первой поверхности имеет вид

$$R_1 = \begin{pmatrix} 1 & -(n_2 - 1)/r_1 \\ 0 & 1 \end{pmatrix}. \quad (22.18)$$

Движение луча внутри линзы после преломления на первой поверхности до встречи с отражающей поверхностью описывается матрицей

$$T_{21} = \begin{pmatrix} 1 & 0 \\ \Delta'_1/n_2 & 1 \end{pmatrix}. \quad (22.19)$$

Отражение от второй поверхности обратно внутрь линзы описывается матрицей (22.16). Движение луча внутри линзы после отражения до встречи с первой поверхностью (при втором проходе луча она обозначена индексом 3) описывается матрицей

$$T_{32} = \begin{pmatrix} 1 & 0 \\ -\Delta'_1/(-n_2) & 1 \end{pmatrix}. \quad (22.20)$$

Матрица преломления на первой поверхности при движении луча в отрицательном направлении дается выражением (22.17) с $n_1 = 1$; т. е.

$$R_3 = \begin{pmatrix} 1 & -(n_2 - 1)/r_1 \\ 0 & 1 \end{pmatrix}. \quad (22.21)$$

Следовательно, искомая передаточная матрица равна

$$S_{31} = R_3 T_{32} R_2 T_{21} R_1. \quad (22.22)$$

Перемножив матрицы, находим по общим правилам все характеристики оптической системы.

Оптическое изображение

Дается матричное описание построения изображений в оптических системах.

Матрица оптической системы. Соотношение (22.15) может быть представлено в следующем виде:

$$\begin{pmatrix} n'_2 a'_2 \\ x'_2 \end{pmatrix} = S_{21} \begin{pmatrix} n_1 a_1 \\ x_1 \end{pmatrix}; \quad (23.1)$$

$$S_{21} = R_2 T_{21} R_1 = \begin{pmatrix} 1 & -k_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \Delta'_1/n'_1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -k_1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} b & -a \\ -d & c \end{pmatrix}, \quad (23.2)$$

$$a = k_1 + k_2 - k_1 k_2 \Delta'_1/n'_1, \quad (a) \quad b = 1 - k_2 \Delta'_1/n'_1, \quad (b)$$

$$c = 1 - k_1 \Delta'_1/n'_1, \quad (в) \quad d = -\Delta'_1/n'_1, \quad (г) \quad (23.3)$$

где a, b, c, d — **постоянные Гаусса**. Не все они независимы между собой. Непосредственной проверкой убеждаемся, что детерминанты R_1, R_2 и T_{21} равны единице и поэтому также $\det S_{21} = bc - ad = \det(R_2 T_{21} R_1) = 1$.

Независимыми являются только три из четырех постоянных Гаусса, в качестве которых обычно выбираются a, b, c .

Преобразование луча от плоскости предмета к плоскости изображения. В плоскости, отстоящей от A_1 влево на расстояние l (плоскость предмета), заданы параметры луча $(n_1 a_1, x)$. В плоскости, отстоящей от A_2 вправо на расстоянии l' (плоскость изображения), луч характеризуется параметрами $(n'_2 a'_2, x')$. По общему правилу, изложенному в § 22, между этими параметрами существует соотношение

$$\begin{pmatrix} n'_2 a'_2 \\ x' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ l'/n'_2 & 1 \end{pmatrix} \begin{pmatrix} b & -a \\ -d & c \end{pmatrix} \begin{pmatrix} 0 & 0 \\ -l/n_1 & 1 \end{pmatrix} \begin{pmatrix} n_1 a_1 \\ x \end{pmatrix}, \quad (23.5)$$

где принято во внимание, что расстояние до плоскости предмета отсчитывается влево от линзы и, следовательно, входит в формулы с отрицательным знаком в последней квадратной матрице в (23.5). Матрицы в (23.5) можно перемножить между собой и записать формулу в виде

$$\begin{pmatrix} n'_2 a'_2 \\ x' \end{pmatrix} = Q_{21} \begin{pmatrix} n_1 a_1 \\ x \end{pmatrix}, \quad (23.6)$$

где

$$Q_{21} = \begin{pmatrix} b + al/n_1 & a \\ bl'/n'_2 - d + al'l/(n'_2 n_1) - cl/n_1 & c - al'/n'_2 \end{pmatrix} \quad (23.7)$$

является **матрицей преобразования от предмета к изображению**.

Из (23.6) видно, что **увеличение $M = x'/x$** зависит от a_1 . Другими словами, лучи, выходящие из одной точки предмета под разными углами, не сходятся в одной точке плоскости изображения. Следовательно, точка предмета в плоскости изображения будет представлена размытым

пятном. Под изображением понимается такое отображение плоскости предмета на плоскость изображения, когда все лучи, исходящие от точки предмета, сходятся в одной точке плоскости изображения и все точки отображаются с одинаковым увеличением. Для этого необходимо отображение сделать независимым от угла a_1 , что достигается обращением в нуль члена в матрице (23.7), ответственного за такую зависимость:

$$bl'/n'_2 - d + al'/l(n'_2 n_1) - cl/n_1 = 0. \quad (23.8)$$

Для увеличения можно написать

$$M = x'/x = c - al'/n'_2. \quad (23.9)$$

Так как детерминанты всех сомножителей, образующих матрицу Q_{21} , равны единице, то $\det Q_{21} = 1$. Тогда при выполнении условия (23.8) с учетом (23.9) получаем, что

$$b + al/n_1 = 1/M \quad (23.10)$$

и, следовательно,

$$Q_{21} = \begin{pmatrix} 1/M & -a \\ 0 & M \end{pmatrix} \quad (23.11)$$

Кардинальные элементы оптической системы. Найдем плоскости, увеличение для точек которых $M=1$. Плоскость H предмета расположена от точки A_1 в соответствии с формулой (23.10) на расстоянии (рис. 74)

$$l_H = n_1(1 - b)/a, \quad (23.12)$$

а плоскость H' изображения расположена от A_2 в соответствии с формулой (23.9) на расстоянии (рис. 74)

$$l'_H = n'_2(c - 1)/a. \quad (23.13)$$

Плоскости H и H' называются **главными плоскостями**, а их пересечения с осью системы, которая совпадает с осью Z , — **главными точками системы**.

Найдем точку на оси системы, в которой сходятся лучи, бывшие до прохождения оптической системы параллельными осями. Из (23.9) при $M=0$ для определения этой точки получаем уравнение $c - al_F/n'_2 = 0$,

из которого следует, что

$$l_F = n'_2 c/a. \quad (23.15)$$

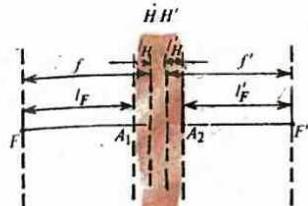
Эта точка называется **фокусом оптической системы**. Ее расстояние до главной плоскости H' называется **фокусным расстоянием** (рис. 74):

$$f' = l_F - l'_H = n'_2/a. \quad (23.16)$$

Аналогично находим точку, выйдя из которой лучи после прохождения оптической системы становятся параллельными оптической оси. Из (23.10) при $M=\infty$ для этой точки получаем $l_F = -n_1 b/a$.

Эта точка также называется **фокусом оптической системы**. Ее расстояние до главной плоскости H называется **фокусным расстоянием** (рис. 74):

$$f = l_F - l_H = -n_1/a. \quad (23.18)$$



74

Кардинальные элементы оптической системы

75

Пример расположения главных плоскостей внутри линзы

Плоскости, проходящие через фокусы перпендикулярно оптической оси, называются **фокальными**. Главные и фокальные плоскости являются **кардинальными элементами оптической системы**. Если кардинальные элементы известны, то можно ответить на вопрос о действии оптической системы на входящие в нее лучи, не зная деталей прохождения лучей через систему.

Физический смысл постоянных Гаусса. Для воздуха можно считать $n_1 = n_2 = 1$. Тогда соотношения (23.16) и (23.18) приводят к равенствам

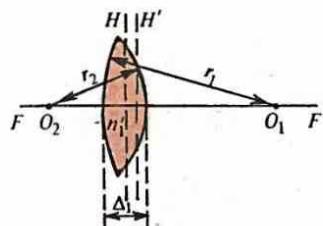
$$-f = f' = 1/a, \quad (23.19)$$

т. е. a является величиной, обратной фокусному расстоянию. Из (23.17) с учетом (23.19) следует, что

$$b = l_F/f \quad (23.20)$$

и, аналогично,

$$c = l'_F/f'. \quad (23.21)$$



Таким образом, b и c характеризуют взаимное расположение главных и фокальных плоскостей. Например, если главная плоскость H' лежит внутри линзы, то величина c определяет ту долю фокусного расстояния, которая приходится на область вне линзы (рис. 75). Следует отметить, что главные плоскости достаточно толстой собирающей линзы лежат внутри линзы асимметрично относительно центра, если радиусы кривизны $|r_1|$ и $|r_2|$ преломляющих поверхностей различны.

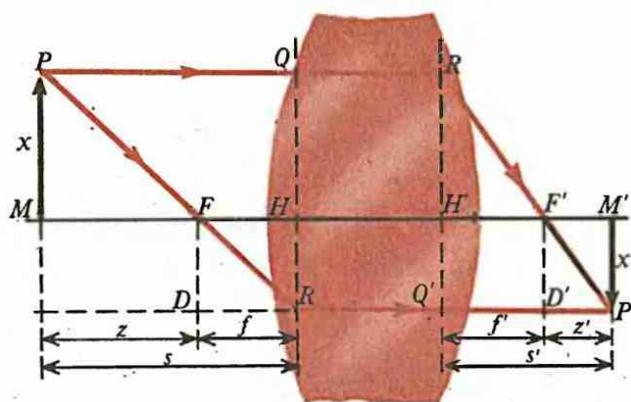
Построение изображений. Изображение точки строится по двум лучам, один из которых направлен параллельно оси, а другой идет по линии, проходящей через фокус. Первый луч проводится параллельно оптической оси до пересечения с главной плоскостью H' (рис. 76), а затем проходит через фокус F' . Второй луч проводится до пересечения с главной плоскостью H , затем параллельно оптической оси до пересечения с первым лучом. Точка пересечения является изображением исходной точки. Если расстояние от предмета до главной плоскости меньше фокусного расстояния, то не сам второй луч, а его продолжение в обратном направлении должно проходить через фокус. Из системы выходят расходящиеся лучи (рис. 77). Пересечение этих лучей дает мнимое изображение точки. Все остальные случаи построений сводятся к этим двум.

Уравнение линзы. Из подобия треугольников RDF , RQP и FMP (рис. 76) следует

$$(x + x')/s = x'/f = x/z, \quad (23.22)$$

а подобие треугольников $P'D'F'$, $F'H'R'$ и $P'Q'R'$ позволяет написать

$$(x + x')/s' = x/f' = x'/z'. \quad (23.23)$$



Построение изображений по кардинальным элементам оптической системы

Из этих соотношений находим выражение

$$x = z x' / f = x' f' / z', \quad (23.24)$$

а из последнего равенства (23.24) следует, что

$$z z' = f f'. \quad (23.25)$$

Равенство (23.25) называется **уравнением линзы в форме Ньютона**. Из (23.22) и (23.23) получаем

$$(x + x')/x' = s/f = s'/z' = s'/(s' - f'), \quad (23.26)$$

где $z' = s' - f'$. Равенство $s/f = s'/(s' - f')$, записанное в виде

$$f'/s' + f/s = 1, \quad (23.27)$$

называется **уравнением линзы в форме Гаусса**. Оно, конечно, эквивалентно (23.25). Из (23.22) и (23.23) следует

$$x + x' = s x' / f = s' x / f. \quad (23.28)$$

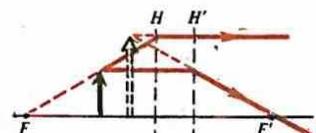
Из последнего равенства (23.28) находим **увеличение линзы**:

$$M = x'/x = s' f / (s f'). \quad (23.29)$$

Тонкие линзы. Равенство (23.3a) с учетом (23.16) и (23.18) приводит к соотношениям

$$a = n_1/f = n'_2/f' = k_1 + k_2 - k_1 k_2 \Delta_1/n'_1. \quad (23.30)$$

Принимая во внимание значения k_1 и k_2 по (22.7) и (22.13) и считая, что линза находится в воздухе ($n_1 = n'_2 = 1$), преобразуем (23.30) к виду



Случай минного изображения

Матрица оптической системы строится по правилу перемножения матриц, описывающих прохождение луча через составные части оптической системы.

$$\frac{1}{f'} = (n'_1 - 1) \left[\frac{1}{r_1} - \frac{1}{r_2} + \frac{(n'_1 - 1)\Delta'_1}{n'_1 r_1 r_2} \right]. \quad (23.31)$$

Формула (23.31) выражает фокусное расстояние линзы через показатель преломления материала линзы и ее геометрические параметры.

Тонкими линзами по определению называются такие, для которых можно пренебречь третьим членом в квадратных скобках в (23.31). Для тонких линз это уравнение принимает вид

$$\frac{1}{f'} = (n'_1 - 1) \left(\frac{1}{r_1} - \frac{1}{r_2} \right). \quad (23.32)$$

Для тонких линз третий член в квадратных скобках (23.31) должен быть много меньше каждого из двух первых членов. Поскольку n'_1 имеет порядок единицы, заключаем, что толщина Δ'_1 линзы должна быть много меньше каждого из радиусов кривизны $|r_1|$ и $|r_2|$ поверхностей линзы, т. е.

$$\Delta'_1 \ll |r_1|, \quad \Delta'_1 \ll |r_2|. \quad (23.33)$$

При этом условия (23.36) и (23.38) принимают вид

$$b = 1, \quad c = 1, \quad (23.34)$$

а равенства (23.12) и (23.13) сводятся к соотношению

$$l_H = l'_H = 0, \quad (23.35)$$

т. е. тонкая линза представляется не имеющей толщины и с ней совпадают обе главные плоскости. Фокусное расстояние при этом становится равным отрезку от линзы до фокуса, а матрица (23.2) принимает вид

$$S_{21} = \begin{pmatrix} 1 & -1/f' \\ 0 & 1 \end{pmatrix}, \quad (23.36a)$$

где $a = 1/f'$. Уравнение (23.27) для тонкой линзы упрощается:

$$\frac{1}{s'} + \frac{1}{s} = \frac{1}{f'}, \quad (23.36b)$$

Обратная фокусному расстоянию линзы величина

$$\Phi = 1/f' \quad (23.36b)$$

называется оптической силой и выражается в диоптриях. Диоптрия равна оптической силе линзы с фокусным расстоянием 1 м.

Система тонких линз. Матрица (23.36a) преобразует параметры луча, входящего в тонкую линзу, в параметры выходящего луча. Распространение луча до тонкой линзы и после нее описывается передаточными матрицами T вида (22.11). Если имеются, например, две линзы, то матрица преобразования луча от входа в первую линзу до выхода из второй равна произведению матриц, описывающих преобразование луча в линзах, и матрицы T , описывающей распространение луча между линзами, взятыми справа налево в том порядке, в каком луч распространяется в системе линз. Матрицы S и T удобно снабжать индексами, показывающими, между какими плоскостями данной матрицей осуществляется преобразование луча. При этом разные стороны

тонкой линзы целесообразно обозначать различными индексами. Например, обозначая плоскость тонкой линзы, через которую входит луч, A_1 , а выходит — A_2 , матрицу (23.36) следует записать так:

$$S_{21} = \begin{pmatrix} 1 & -1/f' \\ A_2 & 1 \end{pmatrix}_{A_1} \quad (23.37)$$

Аналогичные обозначения используются также и для передаточной матрицы T .

Пусть система состоит из двух линз с фокусными расстояниями $f'_1 > 0$ и $f'_2 < 0$ (собирающая и рассеивающая) и расстоянием между линзами d (рис. 78). Линзы на этом рисунке вычерчены пунктирной линией, чтобы подчеркнуть, что их толщиной мы пренебрегаем. Передаточная матрица S_{41} , которая преобразует параметры луча от входа в первую линзу на плоскости A_1 к выходу из второй линзы на плоскости A_4 , на основании (23.37) и (22.11) имеет вид

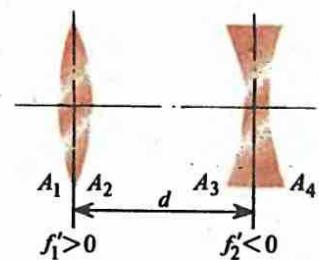
$$S_{41} = \begin{pmatrix} 1 & -1/f'_2 \\ A_4 & 1 \end{pmatrix}_{A_3} \cdot \begin{pmatrix} 1 & 0 \\ d & 1 \end{pmatrix}_{A_3} \cdot \begin{pmatrix} 1 & -1/f'_1 \\ 0 & 1 \end{pmatrix}_{A_1}, \quad (23.38)$$

где предполагается, что показатель преломления среды между линзами равен 1. Чтобы построить изображение, создаваемое этой системой, надо перемножить матрицы (23.38) и найти значения постоянных Гаусса a, b, c, d , определенных в (23.2):

$$S_{41} = \begin{pmatrix} b & -a \\ -d & c \end{pmatrix}. \quad (23.39)$$

Зная постоянные Гаусса, по формулам (23.12), (23.13), (23.15), (23.17) вычисляем l'_H, l'_H, l_F, l'_F , строим кардинальные плоскости системы и изображение в соответствии с общими правилами, рассмотренными в связи с рис. 76. На рис. 79 построены кардинальные плоскости системы, показанной на рис. 78. Напомним, что при отрицательных значениях l_H, l'_H, l_F, l'_F они откладываются влево от точек своего отсчета, а при положительных — вправо. На рис. 79 принято, что l_F, l_H и l'_H отрицательны, а l'_F положительна.

Использование ЭВМ. Представление в матричной форме преобразований, которые претерпевает луч в оптической системе, делает очень удобным использование ЭВМ для анализа оптических систем, поскольку программы вычислений с матрицами являются стандартными. В связи с этим в настоящее время проектирование и расчет оптических систем производят



78

К нахождению передаточной матрицы для системы двух линз



79

Построение кардиональных плоскостей системы из собирающей и рассеивающей линз

○

Каков физический смысл постоянных Гаусса и как с их помощью образуется матрица оптической системы?

Почему из четырех постоянных Гаусса независимыми являются только три?

Дайте определение кардиональных элементов оптической системы.

с использованием ЭВМ, хотя для анализа общих закономерностей полностью сохраняют свою эффективность аналитические и графические методы.

Пример 23.1. Рассчитать элементы двояковыпуклой линзы, для которой $r_2 = -2$; $r_1 = +3$; $n'_1 = 1,5$; $\Delta'_1 = 1$ (см. рис. 75). Линза находится в вакууме (воздухе). Длины можно задать в любых единицах. Поэтому в условии задачи дано лишь число единиц длины без указания наименования единицы.

В соответствии с изложенным в этом параграфе правилом обозначения встречающихся в расчете единиц можно написать:

$$\Delta_2 = \Delta'_1 = 1; \quad n_1 = 1; \quad n_2 = n'_1 = 1,5; \quad n'_2 = 1. \quad (23.40)$$

Из (22.7) и (23.13) получаем:

$$k_1 = (n'_1 - n_1) / r_1 = 0,17; \quad k_2 = (n'_2 - n_2) / r_2 = 0,25.$$

Отсюда по формулам (23.3 а—г) находим постоянные Гаусса:

$$a = 0,17 + 0,25 - 0,17 \cdot 0,25 \cdot 1/1,5 = 0,39;$$

$$b = 1 - 0,25 \cdot 1/1,5 = 0,84; \quad (23.41a)$$

$$c = 1 - 0,17 \cdot 1/1,5 = 0,87;$$

$$d = -1/1,5 = -0,67.$$

Хорошой проверкой правильности вычислений постоянных Гаусса может быть равенство единице детерминанта их матрицы: $bc - ad = 0,99 \approx 1$. Из равенств (23.12) и (23.13) получаем

$$l_H = 1(1 - 0,84)/0,39 = 0,41; \quad l'_H = 1(0,87 - 1)/0,39 = -0,33 \quad (23.41b)$$

и отсюда с помощью (23.17) и (23.15) находим

$$l_F = -1 \cdot 0,84/0,39 = -2,15; \quad l'_F = 1 \cdot 0,87/0,39 = 2,23.$$

Из (23.41б) заключаем, что кардинальные плоскости расположены внутри линзы, а фокусные расстояния в соответствии с (23.18) и (23.16) равны: $f = -2,15 - 0,41 = -2,56$; $f' = 2,23 - (-0,34) = 2,56$, что также свидетельствует о правильности вычислений.

Пример 23.2. Рассчитать систему, состоящую из двояковыпуклой и двояковогнутой линз (см. рис. 78). Линзы считаются тонкими, $d = 10$, $f_1 = 15$, $f_2 = -12$. Показатель преломления в пространстве вне линз равен единице.

Прежде всего по формуле (23.38) находим передаточную матрицу:

$$S_{41} = \begin{pmatrix} 1 & 1/12 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 10 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1/15 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 1,83 & -0,04 \\ 10 & 0,34 \end{pmatrix}. \quad (23.42)$$

Следовательно, постоянные Гаусса равны

$$a = 0,04; \quad b = 1,83; \quad c = 0,34; \quad d = -10. \quad (23.43)$$

С их помощью, так же как в примере 23.1, находим все характеристики оптической системы. Далее можно анализировать ход лучей и построение изображений графическими методами. Но это можно сделать и с помощью матрицы Q_{21} по формуле (23.6).

Описывается аберрации оптических систем и методы их уменьшения или устранения.

Источники аберраций. В определении понятия изображения содержится требование того, чтобы все лучи, выходящие из какой-то точки предмета, сходились в одной и той же точке в плоскости изображения и чтобы все точки предмета отображались с одинаковым увеличением в одной и той же плоскости.

Для параксиальных лучей условия отображения без искажений соблюдаются с большой точностью, однако не абсолютно. Другими словами, параксиальное приближение описывает параксиальные лучи приближенно, хотя и с большой точностью. Поэтому полученная в параксиальном приближении идеальная картина изображений в действительности не осуществляется на практике. Отклонения фактически получаемого изображения от идеального называются аберрациями. Для параксиальных лучей аберрации малы и ими пренебрегают. Если же лучи не параксиальны, то аберрации становятся значительными и сильно искажают изображение. Поэтому первый источник аберраций состоит в том, что линзы, ограниченные сферическими поверхностями, преломляют лучи не совсем так, как это принимается в параксиальном приближении. Например, фокусы для лучей, падающих на линзу на разных расстояниях от оси линзы, различны и т. д. Такие аберрации называют геометрическими. Их можно классифицировать по определенным признакам. Например, параксиальное приближение основывается на том, что точные формулы разложения синуса в ряд (22.1) обрываются на первом члене, пропорциональном a . Не учтенный в параксиальном приближении член $\sim a^3$ приводит к аберрациям третьего порядка.

Следующий член в разложении описывает аберрации пятого порядка, которые обычно меньше аберраций третьего порядка, и т. д.

Второй источник аберраций связан с дисперсией света. Поскольку показатель преломления зависит от частоты, то и фокусное расстояние и другие характеристики системы зависят от частоты. Поэтому лучи, соответствующие излучению различной частоты, исходящие из одной точки предмета, не сходятся в одной точке плоскости изображения даже тогда, когда лучи, соответствующие каждой частоте, осуществляют идеальное отображение предмета. Такие аберрации называются хроматическими.

Изучение геометрических аберраций сводится к учету тех факторов, которыми пренебрегает параксиальное приближение. В принципиальном смысле это просто, но чрезвычайно трудоемко и громоздко. Поэтому ограничимся изложением сути, не вдаваясь в детали математической стороны дела. Это касается также и хроматических аберраций. Точные расчеты проводят на ЭВМ.

Точные матрицы преобразований. Для нахождения точной матрицы преобразования параметров луча при преломлении на сферической поверхности необходимо вместо приближенного расчета, начинаящегося с формулы (22.2), провести точный расчет (см. рис. 73). Вместо формулы (22.2) необходимо записать закон Снеллиуса:

$$n_1 \sin \theta_1 = n'_1 \sin \theta'_1. \quad (24.1)$$

Точный учет (а не в параксиальном приближении) геометрических соотношений между углами и длинами на рис. 73 приводит вместо (22.5) к соотношению

$$n'_1 \sin a'_1 = (n_1 \cos \theta_1 - n'_1 \cos \theta'_1) (x'_1 / r_1) + n_1 \sin a_1, \quad (24.2)$$

которое в параксиальном приближении совпадает с (22.5). С учетом $x'_1 = x_1$ соотношение (24.2) можно представить в матричном виде, аналогичном (22.6):

$$\begin{pmatrix} n'_1 \sin a'_1 \\ x'_1 \end{pmatrix} = \begin{pmatrix} 1 & -\tilde{k}_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} n_1 \sin a_1 \\ x_1 \end{pmatrix}. \quad (24.3)$$

где

$$\tilde{k}_1 = (n'_1 \cos \theta'_1 - n_1 \cos \theta_1) / r_1.$$

Следовательно, точная матрица преломления

$$\tilde{R}_1 = \begin{pmatrix} 1 & -\tilde{k}_1 \\ 0 & 1 \end{pmatrix} \quad (24.4)$$

аналогична (22.8), но с заменой значения k_1 . Для луча после преломления до встречи со второй преломляющей поверхностью вместо (22.9) можем записать

$$x_2 = x'_1 + L'_1 \sin \alpha'_1, \quad (24.5)$$

где L'_1 — длина отрезка $|P_1 P_2|$. Далее, так же как и при выводе (22.10), учитываем, что $n'_1 = n_2$, $\alpha'_1 = \alpha_2$ и, следовательно,

$$n'_1 \sin \alpha'_1 = n_2 \sin \alpha_2. \quad (24.6)$$

В матричной форме равенства (24.5) и (24.6) записывают в виде

$$\begin{pmatrix} n_2 \sin \alpha_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ L'_1/n'_1 & 1 \end{pmatrix} \begin{pmatrix} n'_1 \sin \alpha'_1 \\ x'_1 \end{pmatrix}. \quad (24.7)$$

Поэтому матрица

$$\tilde{T}_{21} = \begin{pmatrix} 1 & 0 \\ L'_1/n'_1 & 1 \end{pmatrix}, \quad (24.8)$$

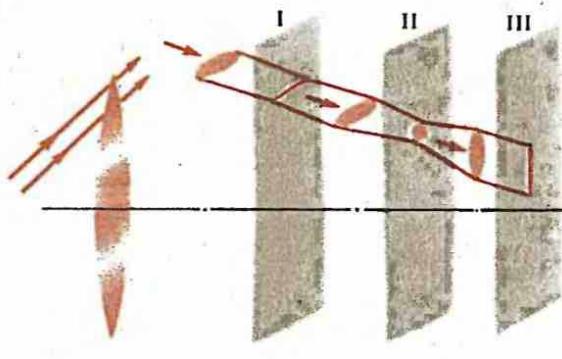
аналогичная матрице (22.11), является передаточной матрицей при точном учете распространения света под произвольным углом в среде. В формуле (24.8) значение L'_1 не является постоянным и его надо каждый раз вычислять по уже известным на предшествующем шаге вычислений значениям α'_1 , x'_1 и параметрам, характеризующим вторую преломляющую поверхность линзы.

Для анализа преломления на второй поверхности линзы используется матрица вида (24.4) с соответствующим значением k_1 , а распространение луча до линзы и после линзы описывается матрицами вида (24.8). Используются эти матрицы аналогично тому, как это было разобрано для параксиального приближения в связи с (23.5).

Сферическая aberrация

Наиболее существенные aberrации сводятся к тому, что круглое сечение пучка под углом к оси после линзы не сохраняется (рис. 80), лучи, параллельные оптической оси, не пересекаются после линзы в одной точке (рис. 81) и точка на оси системы дает изображение в виде причудливой фигуры (комы). Начнем описание со второй из этих aberrаций, называемой сферической. Ее можно трактовать либо как поперечную, либо как продольную (рис. 81). Если в рассматриваемом вопросе существенно то, что луч пересекает ось не в параксиальном фокусе, то говорят о продольной сферической aberrации, а если существенно отклонение луча от оси в параксиальной фокальной плоскости, то говорят о поперечной сферической aberrации.

Пучок параллельных оси лучей после преломления образует совокупность конусов, вершины которых расположены на оси (рис. 82). Огибающая эту совокупность конусов поверхность называется каустической, а сечение этой поверхности любой плоскостью, проходящей через луч — каустической кривой. На рис. 82 изображена каустическая поверхность при сферической aberrации. Ее сечения плоскостями, перпендикулярными оси, являются окружностями различного радиуса. Параллельный пучок лучей создается светящейся точкой, расположенной на оси на очень большом расстоянии от линзы. Поэтому светящиеся кружки играют роль изображений точки в различных плоскостях. Фокус F' определен в параксиальном приближении

*P?*

80

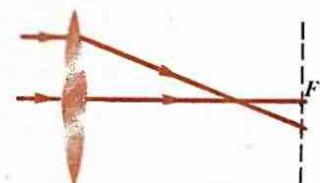
Астигматизм наклонных пучков

и играет роль фокуса лишь для параксиальных лучей (т. е. тех лучей, которые прошли линзу вблизи ее оси). Наиболее яркое и маленькое изображение точки линзой достигается в плоскости M' , которая не проходит через параксиальный фокус F' . Следовательно, для уменьшения поперечной сферической aberrации данной линзы необходимо выбрать соответствующую фокусировку этой линзы, т. е. получать изображение не из расчета, что ее фокус в F' , а из расчета, что ее фокус в M' . Собирающие линзы имеют отрицательную продольную сферическую aberrацию, т. е. непараксиальные лучи пересекают ось ближе к линзе, чем параксиальный фокус. Рассеивающие линзы обладают сферической aberrацией противоположного знака.

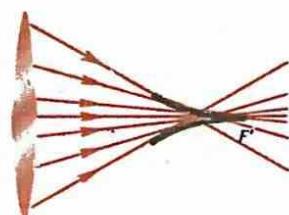
Соответствующим подбором поверхностей и систем линз сферическая aberrация может быть практически ликвидирована. То же самое касается и сферической aberrации зеркал.

Кома. Если светящаяся точка, посылающая широкий пучок лучей, расположена не на оси оптической системы, то ее изображение не является светящимся кружком, как в предыдущем случае, а представляется в виде довольно сложной асимметричной фигуры. Иногда эта фигура напоминает комету с хвостом, отчего и произошло название этого вида aberrации. Соответствующим подбором характеристик системы кома может быть значительно ослаблена.

Аберрации, обусловленные внеосевыми наклонными лучами. Плоскость, проходящая через ось системы, называется меридиональной. Если в ней под достаточно большим углом к оси падает цилиндрический пучок лучей, то после преломления он не останется цилиндрическим. Лучи, лежащие в меридиональной плоскости, преломляются не так, как параллельные им лучи, но лежащие в стороне от меридиональной плоскости. В результате этого после преломления лучи пучка не параллельны друг другу. Поэтому сечение пучка лучей изменяется с расстоянием от линзы после преломления. На некотором расстоянии от линзы сечение является отрезком линии, направленным перпендикулярно меридиональной плоскости (рис. 80), затем эта линия



81

Продольная $F'A$ и поперечная $F'B$ aberrации

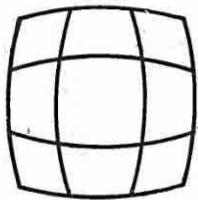
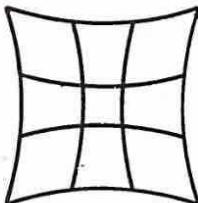
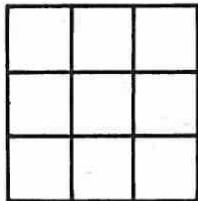
82

Каустическая поверхность при сферической aberrации



83

Поверхности меридианальных и сагиттальных фокусов.



переходит в эллипс, параметры которого меняются по мере удаления от линзы. На некотором расстоянии сечение становится круговым, а затем снова эллиптическим и, наконец, превращается в отрезок линии, лежащей в меридианальной плоскости. Такой вид aberrации называется **астигматизмом наклонных пучков**.

Интерпретируем описанную картину преломления пучков другими терминами. В результате прохождения через линзу пучок фокусируется в **меридианальной плоскости** и в плоскости, перпендикулярной меридианальной и параллельной оси называемой **сагиттальной**. Фокусы меридианальной и сагиттальной фокусировок различны. Меридианальный фокус расположен на рис. 80 в плоскости I , а сагиттальный — в плоскости III . В плоскости II лучи верхней половины первоначального цилиндрического пучка находятся в нижней половине кружка, а нижней — в верхней. Лучи правой половины первоначального цилиндрического пучка находятся в правой половине кружочка, а левой — в левой. Положение плоскостей, в которых осуществляются меридианальная и сагиттальная фокусировки, зависит от угла наклона падающего пучка к оптической оси. Поэтому поверхности, на которых лежат фокусы, создаваемые меридианальной и сагиттальной фокусировками, не совпадают между собой и не являются плоскостями. Очевидно, что эти поверхности касаются лишь в точке F' на оптической оси, будучи ей в этой точке перпендикулярными (рис. 83). Этот вид aberrации называется **искривлением поверхности изображения**. Он устранился при выполнении условия Петцвала, на выводе и обсуждении которого остановливаться не будем.

Увеличение системы зависит, вообще говоря, от угла наклона луча к оптической оси. При больших углах эта зависимость становится заметной и приводит к тому, что изображение теряет свое подобие предмету. В результате, например, сетка из прямых линий превращается в сетку из кривых линий (рис. 84). Такая aberrация называется **дисторсией**.

Ослабление геометрических aberrаций достигается подбором линз, их характеристик и т. д. В настоящее время удается устранить все aberrации до уровня практических потребностей.

Хроматическая aberrация. Она устраниется подбором такой комбинации линз, которая сводила бы к минимуму несовпадения изображений в различных длинах волн. Точно совместить изображения для всех длин волн спектра невозможно. Обычно ограничиваются точным совмещением изображения для каких-то двух длин волн, а для остальных совмещение осуществляется с той или иной степенью точности. Этот процесс называется **ахроматизацией оптической системы**. Изображения для двух длин волн совпадут, если у системы для этих длин волн одинаковые кардинальные элементы. Это достигается при одинаковости трех постоянных Гаусса, т. е. для ахроматизации надо

84

Дисторсия

иметь как минимум три свободных параметра. Значения параметров могут быть найдены как решение системы трех уравнений, выражающих условие совпадения кардинальных элементов для обоих длин волн. Всегда можно подобрать такую оптическую систему, в которой имелись бы три необходимых свободных параметра. Задача облегчается тем, что для практических целей достаточно осуществить лишь частичную ахроматизацию, добившись либо совпадения фокусов, не заботясь о некотором расхождении главных плоскостей, либо совпадения главных плоскостей, не заботясь о некотором расхождении фокусов. Выбор параметров ахроматизации зависит от назначения прибора.

Очевидно, что ахроматизацию можно в принципе осуществить и по трем и большему числу длин волн. Для этого надо создать систему, имеющую достаточное число свободных параметров, и выбрать соответствующим образом эти параметры. Ахроматизация по большему, чем два, числу длин волн также используется в оптике. Выбор длин волн, для которых осуществляется точное совмещение изображений, зависит от назначения прибора. Для приборов визуального наблюдения выбирают длины волн вблизи максимальной чувствительности глаза, т. е. по разные стороны желто-зеленой области спектра. Чаще всего берутся лучи с $\lambda = 656,3$ нм и $\lambda = 486,1$ нм. В фотоаппаратах длины волн выбираются ближе к синей области спектра, поскольку фотопластинка чувствительнее к таким длинам волн.

Иммерсионный объектив. Чтобы использовать широкие пучки и при этом избегать сферической aberrации, применяют иммерсионные объективы. Их принцип действия наиболее четко проявляется при построении изображения точки, расположенной внутри сферической линзы (рис. 85). Точка P выбирается на расстоянии $n' r/n$ от центра O сферической линзы, где n и n' — показатели преломления линзы и среды относительно вакуума. Изображение образуется в результате преломления сферической поверхностью, на которую луч падает с вогнутой стороны. Поэтому радиус кривизны этой поверхности входит в формулы с отрицательным знаком. Величины, отсчитываемые от точки A_1 влево, также отрицательны.

По теореме о синусах для треугольников $P'QO$ и PQO можем написать

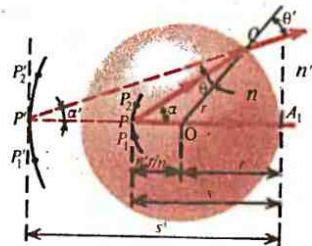
$$(s' - r) \sin a' = r \sin \theta', \quad (24.9) \quad (s - r) \sin a = r \sin \theta. \quad (24.10)$$

Закон Снеллиуса для преломления в точке Q имеет вид
 $n' \sin \theta' = n \sin \theta. \quad (24.11)$

Учитывая, что, по условию, $s - r = n' r/n$, из (24.11) и (24.10) заключаем

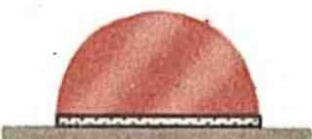
$$\theta' = a, \quad (24.12)$$

а из треугольников $P'QO$ и PQO с общим углом следует равенство



85

Построение изображения точки, расположенной внутри сферической линзы



86

Принцип устройства иммерсионного объектива микроскопа

● **Аберрациями называются отклонения фактически полученного изображения от идеального. Изучение геометрических aberrаций сводится к учету тех факторов, которыми преобладает параксиальное приближение.**

Источником хроматической aberrации является дисперсия света.

$$a + \theta = a' + \theta',$$

что с учетом (24.12) дает

$$\theta = a'. \quad (24.14)$$

В результате (24.9) принимает вид

$$s' = r \sin \theta' / \sin a' + r = r \sin \theta' / \sin \theta + r = r(1 + n/n'), \quad (24.15)$$

т. е. s' не зависит от угла a . Это означает, что продолжения всех лучей, вышедших из точки P , пересекаются в точке P' или, другими словами, точка P' является изображением точки P , при чем сферическая aberrация полностью отсутствует, поскольку в расчете не предполагалась малость углов.

Объект не может быть помещен в стекло линзы. Однако можно создать ситуацию, эквивалентную нахождению объекта в точке P (рис. 85). Для этого надо объект поместить в масло с показателем преломления, равным показателю преломления стекла линзы, и переднюю поверхность линзы погрузить в масло (рис. 86). В результате получится иммерсионный объектив микроскопа.

Из (24.14) видно, что иммерсионный объектив, не создавая сферической aberrации, уменьшает углы лучей с осью после выхода из иммерсионного объектива, поскольку всегда $a' < a$. Это позволяет без уменьшения интенсивности светового потока создать для последующих элементов оптической системы более благоприятные условия формирования изображения без сферической aberrации. Можно повторным применением приема уменьшить углы между лучами и осью еще более значительно. Единственным ограничением для сколь угодно значительного уменьшения углов является возникающая при этом хроматическая aberrация.

Условие Аббе. Рассмотрим изображение отрезка $P_1 P_2$ окружности радиусом $n'r/n$ с центром в точке O (см. рис. 85). Он изображается отрезком $P'_1 P'_2$ окружности, причем каждая точка отрезка отображается широким пучком без сферической aberrации. Необходимо далее, чтобы все точки отрезков изображались с одинаковым увеличением. В результате изображение широкими пучками осуществляется без сферической aberrации и комы. Обозначим длины дуг $P_1 P_2$ и $P'_1 P'_2$ соответственно x и x' . Из рис. 85 видно, что

$$\frac{x'}{x} = \frac{s' - r}{s - r} = \frac{s' - r}{rn'/n} = \frac{n \sin a}{n' \sin a'}, \quad (24.16)$$

где использовано равенство (24.15). Соотношение (24.16), записанное в виде

$$x' n' \sin a' = x n \sin a, \quad (24.17)$$

называется **уравнением Аббе**. Если оптическая система сконструирована так, что увеличение равно $x'/x = \text{const}$ для любых углов a , то изображение осуществляется без сферической aberrации и комы. В оптической системе (см. рис. 85) это условие соблюдается, поскольку на основании (24.11), (24.12) и (24.14) вместо (24.16) можем написать

$$x'/x = n \sin a / (n' \sin a') = n \sin \theta' / (n' \sin \theta) = (n/n')^2 = \text{const}. \quad (24.18)$$

Задача конструктора оптической системы состоит в том, чтобы обеспечить постоянство увеличения для всех лучей, проходящих через систему, в том числе и для лучей, идущих от предмета под очень большим углом к оси системы. Для соблюдения уравнения Аббе при очень больших увеличениях микроскопов используются иммерсионные объективы.

Анализируется роль диафрагмирования и описываются распространенные оптические приборы.

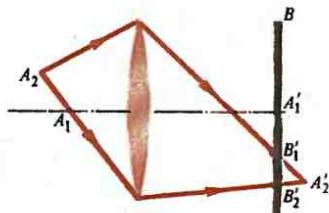
Диафрагмирование. Для получения удовлетворительного изображения в оптической системе необходимо использовать пучки ограниченной ширины. Это связано, во-первых, с тем, что праксиальное приближение ограничивает ширину допустимых пучков. Во-первых, если даже допустить, что создана идеальная система, дающая без aberrации отображение точек предмета в точки изображения при неограниченной ширине пучка, то эта система без ограничения ширины пучка не даст удовлетворительного изображения предмета (рис. 87). Дело в том, что изображение должно быть получено в одной плоскости, а предмет — пространственный, поэтому идеальные отображения его точек займут некоторую область пространства. Например, если изображение точки A_1 образовалось в точке A'_1 плоскости B , принятой за плоскость изображений, то изображение точки A_2 образуется в точке A'_2 вне плоскости изображений. В плоскости B изображений точка A_2 даст светлое пятно, линейный размер которого в плоскости рисунка равен длине отрезка $B'_1B'_2$. Чем шире пучок света из точки A_2 , участвующий в формировании изображения A'_2 , тем больше размеры светлого пятна, изображающего точку A_2 в плоскости B . При неограниченной ширине пучка даже идеальная система не дает изображения предмета в плоскости изображений. Ограничение сечений световых пучков называется **диафрагмированием**. Оно осуществляется с помощью диафрагм.

Основные понятия, связанные с диафрагмированием. Ширина пучков, проходящих через диафрагмированную систему, различна для различных точек предмета. Для точек предмета, лежащих на оси оптической системы, диафрагмирование характеризуется апертурной диафрагмой, входным и выходным зрачками.

Апертурной называется диафрагма, которая осуществляет максимальное ограничение пучка, исходящего от точки предмета, лежащей на оптической оси системы (D_1D_2 на рис. 88). Если бы оправа линзы L_1 закрывала кольцевые области $K_1D'_1$ и $K_2D'_2$, то апертурной диафрагмой по-прежнему была бы диафрагма D_1D_2 . Если бы оправа линзы L_1 закрывала ее кольцевые области $K_1D''_1$ и $K_2D''_2$, то апертурной диафрагмой была бы диафрагма $D''_1D''_2$, а не D_1D_2 .

Входным зрачком называется изображение апертурной диафрагмы, осуществляемое той частью оптической системы, которая находится перед ней (B_1B_2 на рис. 88). Если апертурная диафрагма расположена перед первой линзой или образована оправой первой линзы, то входной зрачок совпадает с апертурной диафрагмой.

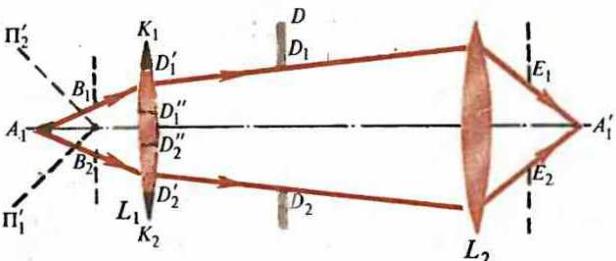
Выходным зрачком называется изображение апертурной диафрагмы, осуществляемое той частью оптической системы,



87

Иллюстрация невозможности получения плоского изображения предмета в идеальной системе посредством широких пучков

$$M = Dd/(f_1 f_2)$$



88

К определению апертурной диафрагмы входного и выходного зрачков

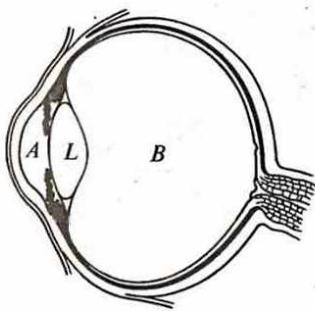
которая находится после диафрагмы ($E_1 E_2$ на рис. 88). Можно также сказать, что выходной зрачок есть изображение входного зрачка, осуществляемое всей системой. Если апертурная диафрагма расположена за системой или образована оправой последней линзы, то выходной зрачок совпадает с апертурной диафрагмой.

Лучи от точек предмета, лежащие вне оси, прошедшие через входной зрачок, могут частично или даже полностью задерживаться на их путях различными частями оптической системы. Вследствие этого изображение соответствующих точек ослаблено или полностью отсутствует. Полем зрения называется наибольшая область $P'_1 P'_2$ в плоскости предмета, точки которой отображаются оптической системой без заметного ослабления. Поле зрения ограничивается диафрагмой поля зрения.

Для нахождения диафрагмы поля зрения строится изображение каждой из диафрагм системы, осуществляемое частью оптической системы перед диафрагмой. То из изображений диафрагм, которое из центра входного зрачка видно под наибольшим углом (углом зрения), называется диафрагмой поля зрения. Аналогично относительно выходного зрачка определяются диафрагма поля зрения со стороны изображения и угол изображения.

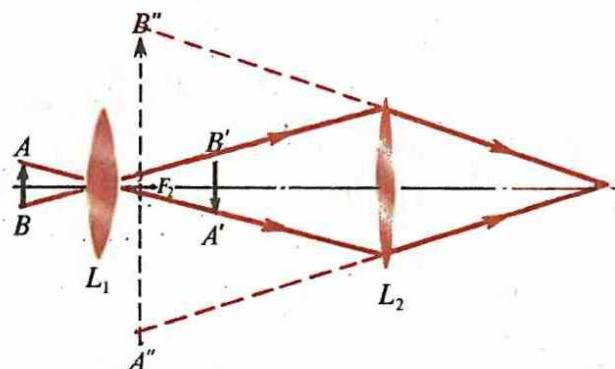
Глаз как оптическая система. В оптическом отношении глаз является оптической системой с переменным фокусным расстоянием (рис. 89). Оптическая система состоит из совокупности преломляющих сред, включающих в себя водянистую влагу A , хрусталик L и стекловидное тело B . Изменение фокусного расстояния системы осуществляется за счет мышечного усилия, изменяющего кривизну поверхности хрусталика. Изображение предметов фокусируется на заднюю стенку глаза (сетчатку), где световая энергия воспринимается чувствительными элементами нервной системы человека. Переданная в мозг информация о распределении светового потока возникает в сознании человека в виде изображения предмета. Более подробно останавливаться на физиологических аспектах зрения здесь не представляется целесообразным.

Фокусировка глаза на предмет называется аккомодацией. Наиболее удаленная и самая близкая точки, на которые может



89

Глаз как оптическая система



быть аккомодирован глаз, называются дальней и ближней точками аккомодации глаза. При нормальном зрении дальняя точка лежит на бесконечности, а ближняя — на расстоянии 10—20 см. С возрастом диапазон аккомодации сужается.

Апертурная диафрагма осуществляется радужной оболочкой. Входным зрачком является изображение зрачка в передней оптической части глаза, т. е. в камере с водянистой влагой. Входной зрачок практически совпадает с реальным зрачком.

Глаз как оптическая система представляется в виде приведенного глаза, построенного из однородного преломляющего вещества и имеющего следующие данные:

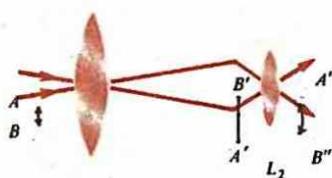
Оптическая сила, дптр	58,48
Длина глаза, мм	22
Радиус кривизны, мм:	
сетчатки	9,7
преломляющей поверхности	5,6
Показатель преломления среды	1,33

Фотоаппарат. Оптической системой фотоаппарата является объектив, обеспечивающий фокусировку изображения объекта на фотопластинку.

Лупа. Простейшая оптическая система с небольшим фокусным расстоянием (1 см или несколько больше), состоит из одной или нескольких линз. Лупа располагается между глазом и предметом. Мнимое изображение образуется либо на расстоянии наилучшего зрения глаза (около 25 см), либо на бесконечности. В обоих случаях увеличение практически одно и то же и равно D/f , где D — расстояние наилучшего зрения, f — фокусное расстояние линзы.

Микроскоп (рис. 90) используется для получения больших увеличений изображений близко расположенных предметов. Действительное изображение $A'B'$ предмета AB , полученное с помощью объектива L_1 , дает в окуляре L_2 мнимое изображение $A''B''$. Если фокусные расстояния объектива и окуляра f_1 и f_2 , а d — расстояние между фокусами, то **увеличение микроскопа**

$$M = Dd/(f_1 f_2). \quad (25.1)$$



Для получения достаточной яркости изображения и увеличения разрешающей способности микроскопа (см. § 36) необходимо пользоваться широкими пучками. При этом для ликвидации сферической aberrации употребляются иммерсионные объективы (см. § 24).

Зрительная труба (рис. 91) используется для получения больших увеличений изображений отдаленных предметов. Действительное изображение предмета в длиннофокусном объективе L_1 рассматривается через окуляр как лупу. Увеличение системы $M = f_1/f_2$, (25.2)

где f_1 и f_2 — фокусные расстояния объектива и окуляра.

Зрительные трубы, используемые для наблюдения астрономических объектов, называются **телескопами**. Для них задний фокус объектива может считаться совмещенным с передним фокусом окуляра (рис. 92), в результате чего формулу (25.2) можно представить в виде

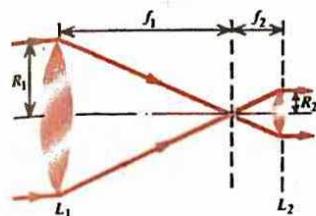
$$M = f_1/f_2 = R_1/R_2, \quad (25.3)$$

где R_1 и R_2 — радиусы объектива и окуляра.

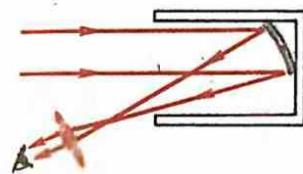
Наблюдение с помощью зрительной трубы ведется глазом. Поэтому для наиболее полного использования светового потока, поступающего в телескоп, необходимо, чтобы выходной зрачок телескопа был равен или меньше зрачка глаза, размеры которого при ночных наблюдениях 6—8 мм, а при дневных — 2—3 мм. Если выходной зрачок больше зрачка глаза, то часть светового потока теряется на радужной оболочке и не участвует в формировании изображения в глазе. Поэтому для эффективного использования поверхности объектива необходимо окуляр подбирать так, чтобы выходной зрачок имел нужные размеры, при этом увеличение трубы оказывается фиксированным. Брать большой объектив при слишком малом увеличении нецелесообразно. При том же самом диаметре объектива ночью целесообразно пользоваться меньшим увеличением, чем днем. Следовательно, если задан диаметр объектива, то целесообразное увеличение оказывается заключенным в довольно узких пределах.

Для многих целей необходимо обеспечить у зрительной трубы большой угол зрения (например, у биноклей). Чтобы обеспечить высокое качество изображения, необходимо устранить астигматизм наклонных пучков, кривизну поля и хроматизм. Поэтому окуляры обычно выполняют в виде сложных систем, состоящих из нескольких линз.

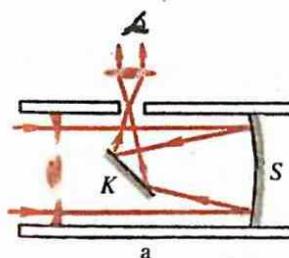
Получение максимального увеличения при фиксированном размере выходного зрачка достигается увеличением диаметра объектива. Кроме того, увеличение диаметра объектива позволяет повысить способность трубы различать слабосветящиеся объективы. Поэтому стремятся всемерно увеличить диаметр телескопов. Изготовление телескопов с линзовыми объективами большого диаметра очень трудно технически. Легче изго-



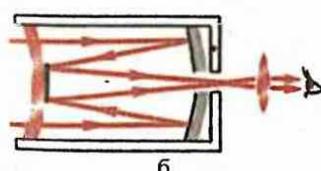
92
Ход лучей в телескопе



93
Схема простейшего телескопа-рефлектора



a



б

94
Система Шмидта (а) и менисковая система Д. Д. Максутова (б)

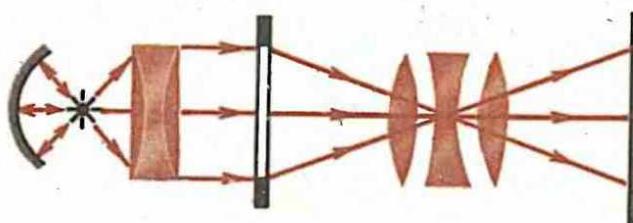


Схема проекционного устройства

тывать вогнутое зеркало большого диаметра с заданной кривизной поверхности. Поэтому наиболее большие телескопы являются не рефракторами, а рефлекторами (рис. 93).

Рефлекторы свободны от хроматической aberrации, но при сферической форме зеркала обладают значительной сферической aberrацией. Для устранения сферической aberrации пользуются различными методами. Можно перейти к несферическим зеркалам, в частности к зеркалам с поверхностью параболоида вращения. Комбинация двух несферических зеркал позволяет добиться очень хороших результатов.

Можно использовать смешанные системы, в которых линзовая оптика комбинируется с зеркальной. Наиболее давно известной системой такого рода является камера Шмидта. На пути лучей к сферическому зеркалу S (рис. 94, а) помещается корректирующая пластиинка K , которая так корректирует фазу проходящего луча, что в значительной степени устраняются сферическая aberrация и кома. Очень совершенной системой такого рода является менисковая система Максутова (рис. 94, б).

Менисковая система Д. Д. Максутова показанная на рис. 94, б, включает в себя лишь сферическое зеркало и мениск, ограниченные сферическими поверхностями. Это очень сильно упрощает расчет и изготовление системы, что весьма существенно при ее практической реализации. В то же время эта система весьма эффективна по своим оптическим параметрам.

Проекционные устройства предназначаются для получения действительных увеличенных изображений на экране (диакинопроекторы и т. д.). Проецируемый светящийся (или освещаемый) предмет помещают около главной фокальной плоскости проекционного объектива, который может перемещаться для резкой наводки (рис. 95). Освещение объекта обычно осуществляется с помощью короткофокусного конденсора большого размера с короткофокусным расстоянием, чтобы пропустить через проецируемый объект значительный световой поток. Световой поток от конденсора сходится на входном зрачке проекционного объектива. Объектив и конденсор должны быть согласованы так, чтобы обеспечивалось резкое изображение предмета и возможно полнее использовался световой поток, проходящий через конденсор.

$$M = f_1 / f_2$$

- 4.1.** Чему равен диаметр изображения Солнца, образуемый вогнутым зеркалом с радиусом кривизны 10 м? Угол, под которым виден диаметр Солнца с Земли, равен 32°.
- 4.2.** Монета лежит в воде на глубине h . На какой глубине она кажется для наблюдателя, смотрящего в вертикальном направлении сверху через поверхность воды? Показатели преломления воды и воздуха соответственно $n = 1,33$ и $n' = 1$.
- 4.3.** Микроскоп с фокусным расстоянием объектива $f_{об} = 4$ мм и окуляра $f_{ок} = 25$ мм и расстоянием между их фокусами $l = 200$ мм используется в качестве проекционного. На каком расстоянии от окуляра необходимо расположить экран, чтобы получить изображение объекта с увеличением $M = 2000$?
- 4.4.** Радиус кривизны вогнутой поверхности плоскогубчатой линзы равен 5 см, толщина линзы 0,5 см. Луч падает на линзу со стороны вогнутой поверхности на расстоянии 1 см от оси, имея наклон 0,1 рад к оси. Найти расстояние от оси до точки выхода луча из линзы и его угол наклона.
- 4.5.** Двояковыпуклая линза (см. рис. 75), параметры которой $r_1 = 6$, $r_2 = -3$, $\Delta i = 1,5$, $n_i = 1,5$, находится в воздухе ($n_i = 1$). Найти фокусное расстояние линзы и положение кардинальных плоскостей.
- 4.6.** Оптическая система (см. рис. 78) характеризуется следующими данными: $f_1 = 30$, $f_2 = -20$, $d = 24$. В приближении тонких линз найти параметры оптической системы. Определить увеличение.
- 4.7.** Луч, выходящий из плоскопараллельной пластины, параллелен падающему на неё лучу. Толщина и показатель преломления материала пластины равны d и n . Найти боковое смещение луча, вышедшего из пластины.
- 4.8.** Световод круглого сечения диаметром d находится в воздухе. Чему равен минимально допустимый радиус кривизны внешней стенки световода? Показатель преломления материала световода n .
- 4.9.** На прозрачный шар радиусом R с показателем преломления n падает вертикальный луч света на расстоянии $d < R$ от вертикального диаметра. При каких значениях n и d точка пересечения луча с вертикальным диаметром шара находится внутри шара?
- 4.10.** На стеклянную полусферу радиусом R , лежащую плоской поверхностью на столе, вертикально падает пучок света круглого сечения радиусом $r < R$. Ось пучка совпадает с вертикалью, проходящей через центр основания полусфера. Чему равен радиус светлого пятна на столе?
- 4.11.** Между двояковыпуклой линзой и зеркалом, на котором получается действительное изображение, помещается прямоугольный сосуд с прозрачными стенками. Внутри сосуда налита вода ($n = 1,33$). В воду опускается другой сосуд дном вниз так, что через его верхние края вода внутрь сосуда не заливается. Второй сосуд тоже прямоугольный, и его две противоположные плоскопараллельные стенки перпендикулярны оптической оси системы. Две плоскопараллельные стенки большого сосуда также перпендикулярны оптической оси системы. После этого на экране получают действительное изображение точки, расположенной на оптической оси. Стенки сосуда, опущенного в воду, сделаны из тонкого стекла ($n = 1,5$). Поворотом опущенного в воду сосуда вокруг вертикальной оси можно изменять угол $\theta_{под}$ между нормалью к его плоской вертикальной поверхности и оптической осью. При некотором угле $\theta_{кр}$ изображение точки исчезает. Найдите этот угол.
- 4.12.** Призма находится в воздухе. Показатель преломления материала призмы $n = 1,5$, углы при основании 45° . Луч света падает на боковую поверхность по нормали к ней. Найти коэффициент пропускания призмы.
- 4.13.** На поверхности воды ($n = 1,33$) плавает круглый диск радиусом R . К центру диска на очень тонкой нити подвешен тяжелый грузик малого размера, который вытягивает нить по вертикали. Расстояние от поверхности воды до грузика h . Толщину диска считать исчезающей малой. При каком условии грузик будет виден наблюдателю, находящемуся в воздухе?
- 4.14.** Чему равен минимальный угол отклонения луча, прошедшего через равностороннюю призму ($n = 1,5$)? Каким станет минимальный угол отклонения, если призму погрузить в воду ($n = 1,33$)?
- 4.15.** Экран расположен на расстоянии $l = 2,5$ м от вогнутого зеркала с радиусом кривизны $r_2 = -5$ м. На каком расстоянии от зеркала необходимо поместить небольшой предмет, чтобы получить его четкое изображение на экране, и каково при этом увеличение?
- 4.16.** Узкий пучок света падает по нормали к поверхности прозрачного шара, находящегося

в воздухе. При каком показателе преломления вещества шара пучок света фокусируется на противоположной поверхности шара?

4.17. Выведите формулу для определения фокусного расстояния системы из двух тонких линз с фокусными расстояниями f_1 и f_2 , находящихся на расстоянии d друг от друга.

4.18. Плоскопараллельная пластиинка толщиной d помещена между линзой и изображением предмета. Показатель преломления материала пластиинки $n > 1$. На какое расстояние в результате этого переместится изображение?

4.19. Объектив фотоаппарата изготовлен из двух тонких линз — плосковогнутой, сделанной из стекла с показателем преломления n_1 , и двояковыпуклой, сделанной из стекла с показателем преломления n_2 . Радиус кривизны вогнутой поверхности плосковогнутой линзы R . Радиусы кривизны поверхностей двояковыпуклой линзы R и R' . Поверхности линз с радиусом кривизны R соприкасаются. Опре-

делить фокусное расстояние объектива фотоаппарата.

4.20. Телеобъектив состоит из двух линз — собирающей с фокусным расстоянием 20 см и расположенной сзади нее на расстоянии 15,5 см рассеивающей линзы с фокусным расстоянием 5 см. Телеобъектив направлен на башню высотой 30 м, расположенную на расстоянии 3 км. Какой величины будет изображение башни и каково расстояние от первой линзы до изображения?

4.21. Линза представляет собой стеклянный шар радиуса R . Показатель преломления стекла n . Доказать, что главные плоскости этой линзы совпадают и проходят через центр шара. Найти фокусное расстояние линзы.

4.22. Расстояние между вогнутым сферическим зеркалом радиуса R и экраном равно d . На каком расстоянии от сферического зеркала на оси оптической системы необходимо поместить маленький объект, чтобы получить его четкое изображение на экране? Какое при этом будет увеличение?

Ответы

- 4.1. 47 мм. 4.2. $h' = -n'h/n = -0,75h$. 4.3. $Mf_{\text{ок}}/l + f_{\text{ок}}(1 - f_{\text{ок}}/l) = 1,028 \text{ м}$. 4.4. 1 см; 0,0 рад. 4.5. $f = 4,26$; $l_H = 0,72$; $l'_H = -0,36$. 4.6. $l_H = -51,42$; $l'_H = -34,28$; $f_F = -94,28$; $l'_F = 8,58$; $M = -1,67$. 4.7. $d \sin(\theta_{\text{пд}} - \theta_{\text{пр}})/\cos \theta_{\text{пр}}$. 4.8. $d/(n-1)$. 4.9. $(d/R)^2 > 1 - (n^2 - 1)^2/4$; $\sqrt{2} < n < 2$. 4.10. $(r/n)/\{(1 - (r/R)^2)^{1/2}[n^2 - (r/R)^2]^{1/2} + (r/R)^2\}$. 4.11. $48^\circ 36'$. 4.12. 0,92. 4.13. $h \geq R/\sqrt{n^2 - 1} \approx 0,882R$. 4.14. $37^\circ; 8^\circ 27'$. 4.15. 0,275 м. 4.16. 2,777. $1/f = 1/f_1 + 1/f_2 - d/(f_1 f_2)$. 4.18. $-d(1 - 1/n)$. 4.19. $[(n_2 - n_1)/R + (n_2 - 1)/R']^{-1}$. 4.20. 2 см; 60,5 см. 4.21. $nR/[2(n-1)]$. 4.22. $R/(2-R); -(2-R)/R$.

Основная идея:
интенсивность волны,
являющейся результатом
суперпозиции двух
или нескольких волн,
определяется
соотношением их фаз
и поляризованностей.

Интерференция



§ 26

Двухлучевая интерференция, осуществляемая делением амплитуды

Выводится общая формула для двухлучевой интерференции и изучаются ее применения в схемах опытов с делением амплитуды. Изучается видимость интерференционной картины для различных условий экспериментов. Обсуждается временная когерентность.

Определение интерференции. Интерференцией называется изменение средней плотности потока энергии, обусловленное суперпозицией электромагнитных волн. Поскольку [см. (3.4)] плотность потока энергии и объемная плотность энергии пропорциональны квадрату амплитуды электромагнитной волны, а коэффициенты пропорциональности постоянны для заданной среды, будем характеризовать плотность потока энергии и объемную плотность энергии монохроматической волны величиной

$$I = \langle \operatorname{Re} E \cdot \operatorname{Re} E \rangle = \frac{1}{2} \operatorname{Re} (E^* E) = \frac{1}{2} E_0^2, \quad (26.1)$$

называемой интенсивностью света. В (26.1) использована формула (6.19), причем E_0 является действительной величиной — амплитудой световой волны. Умножив (26.1) на ϵ , получим среднюю объемную плотность энергии, а умножив на $v\epsilon$, получим плотность потока энергии, где ϵ — диэлектрическая проницаемость среды, а v — скорость света в среде.

Интенсивность при суперпозиции двух монохроматических волн. Будем считать, что обе волны поляризованы линейно в одном и том же направлении и имеют одинаковые амплитуды. Тогда их можно представить в виде

$$E_1 = E_0 e^{-i(\omega t - \varphi_1)}, \quad E_2 = E_0 e^{-i(\omega t - \varphi_2)}, \quad (26.2)$$

причем E_0 предполагается действительной величиной.

В результате суперпозиции волн напряженность электрического поля равна

$$E = E_1 + E_2, \quad (26.3)$$

а интенсивность на основе (26.1) дается формулой

$$\begin{aligned} I &= \langle \operatorname{Re} (E_1 + E_2) \operatorname{Re} (E_1 + E_2) \rangle = \frac{1}{2} \operatorname{Re} [(E_1^* + E_2^*) (E_1 + E_2)] = \\ &= \frac{1}{2} \operatorname{Re} (E_1^* E_1 + E_2^* E_2 + E_1^* E_2 + E_2^* E_1), \end{aligned} \quad (26.4)$$

где использовано соотношение (6.19). Из (26.2) следует

$$E_1^* E_1 = E_0^2, \quad E_2^* E_2 = E_0^2, \quad E_1^* E_2 = E_0^2 e^{i(\varphi_2 - \varphi_1)}, \quad E_2^* E_1 = E_0^2 e^{-i(\varphi_2 - \varphi_1)}$$

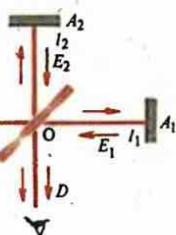
и поэтому

$$I = E_0^2 (1 + \cos \delta) = 2I_0 (1 + \cos \delta); \quad (26.5)$$

где $\delta = \varphi_2 - \varphi_1$; $I_0 = E_0^2/2$ — интенсивность каждой из световых волн, интерференцию которых выражает формула (26.5).

Соотношение между суммарной интенсивностью и слагаемыми интенсивностями зависит от разности фаз между волнами. Если складываются волны различных амплитуд

$$E_1 = E_{01} e^{-i(\omega t - \varphi_1)}, \quad E_2 = E_{02} e^{-i(\omega t - \varphi_2)}, \quad (26.6)$$



96

Схема деления амплитуды волны в интерферометре Майкельсона

то $E_1^*E_1 = E_{01}^2$, $E_2^*E_2 = E_{02}^2$, $E_1^*E_2 = E_{01}E_{02}e^{i(\varphi_2 - \varphi_1)}$, $E_2^*E_1 = E_{02}E_{01}e^{-i(\varphi_2 - \varphi_1)}$, поэтому формула (26.4) принимает вид

$$I = \frac{1}{2}(E_{01}^2 + E_{02}^2 + 2E_{01}E_{02}\cos\delta) = I_1 + I_2 + 2\sqrt{I_1}\sqrt{I_2}\cos\delta, \quad (26.7)$$

где $I_1 = E_{01}^2/2$, $I_2 = E_{02}^2/2$ — интенсивности слагаемых волн. Из (26.7) видно, что суммарная интенсивность изменяется от минимального значения

$$I_{\min} = (\sqrt{I_1} - \sqrt{I_2})^2 \quad (\cos\delta = -1) \quad (26.8)$$

до максимального

$$I_{\max} = (\sqrt{I_1} + \sqrt{I_2})^2 \quad (\cos\delta = 1). \quad (26.9)$$

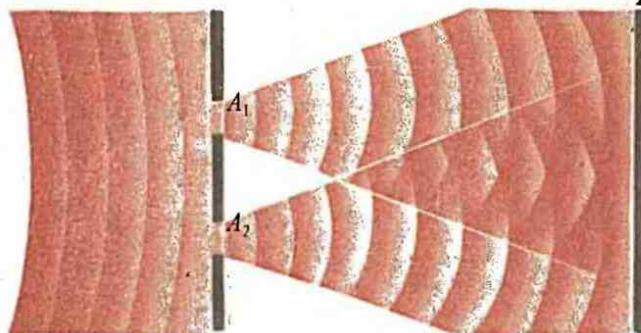
Способы получения когерентных волн в оптике. Чтобы осуществить двухлучевую интерференцию, необходимо иметь две монохроматические волны одинаковой частоты. Такие волны, по определению, имеют бесконечную продолжительность во времени. Ясно, что в природе они не существуют. Поэтому приходится ограничиться квазимонохроматическими волнами.

Можно получить волны, пригодные к интерференции, если они возникают в результате разделения одной и той же волны на две части. Обе части волны в отношении изменения их фазы по времени являются точными копиями исходной. Однако полной аналогии с интерференцией монохроматических волн здесь не получится, поскольку каждая из волн имеет конечное время когерентности (см. § 13), в течение которого эти волны действительно могут интерферировать. Поэтому картина интерференции монохроматических волн является лишь первым приближением в изучении интерференции волн от реальных источников.

Получение волн для реализации интерференции в оптике осуществляется двумя способами: 1) делением амплитуды волны; 2) делением фронта волны.

Сущность этих способов ясна из рис. 96 и 97. На рис. 96 изображена схема интерферометра Майкельсона, с помощью которого осуществляется интерференция делением амплитуды волны. Волна, исходящая из источника S_0 , падает на полупрозрачную пластинку O , расположенную под углом 45° к направлению распространения луча. На пластинке волна разделяется на две части: отраженная волна идет в направлении к A_2 , а прошедшая через пластинку — в направлении A_1 . После отражения от зеркал A_1 и A_2 они снова частично отражаются, а частично проходят через пластинку O . Волны, распространяющиеся в направлении D , могут между собой интерферировать. Ясно, что на пластинке O происходит деление амплитуды, поскольку фронты волн на ней сохраняются, меняя лишь направление своего движения.

$$I = \frac{1}{2} E_0^2$$



97

Схема деления фронта волны экраном с двумя щелями

На рис. 97 изображена схема деления фронта волны. Волна от источника падает на непроницаемую перегородку, в которой имеются две щели: A_1 и A_2 . Те части фронта волны, которые попадают на непроницаемые части перегородки, отражаются от нее или поглощаются. К экрану D проходят лишь те части фронта исходной волны, которые попадают на щели A_1 и A_2 . После щелей распространяются две волны, которые могут между собой интерферировать. Ясно, что экран делит фронт волны на части, откуда и происходит название этого способа получения интерференции волн.

Интерференция монохроматических волн, распространяющихся строго вдоль оси интерферометра Майкельсона. Разность хода лучей приобретается за счет разницы в длинах плечей l_1 и l_2 интерферометра Майкельсона:

$$\Delta = 2(l_1 - l_2).$$

(26.10)

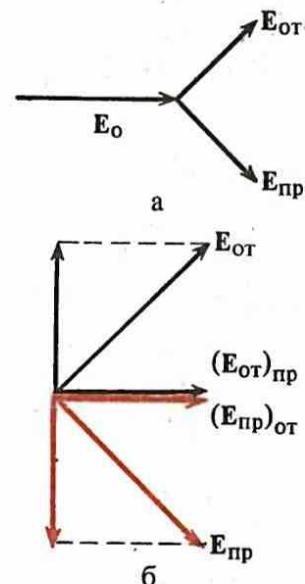
Возникшая при этом разность фаз равна

$$\delta = 2\pi\Delta/\lambda = k\Delta.$$

(26.11)

В этом расчете не приняты во внимание изменения фаз волн при отражениях от зеркал A и B , отражениях и преломлениях в пластинке O и за счет распространения света внутри пластины. В принципе эти факторы нетрудно учесть. Однако здесь это не делается, поскольку их учет никаких интересных для интерференции моментов не содержит.

Считаем, что при каждом падении волны на пластину O плотность потока энергии делится на две равные части. Если амплитуда исходной волны E_0 , то в направлении D в результате двух падений волны на пластину O распространяются волны с амплитудами $E_0/2$:



98

Векторные диаграммы сложения комплексных амплитуд волн, отразившихся от пластины и прошедших через нее

$$E_1 = \frac{1}{2} E_0 e^{-i(\omega t - \varphi_1)}, \quad E_2 = \frac{1}{2} E_0 e^{-i(\omega t - \varphi_2)}, \quad (26.12)$$

причем $\varphi_2 - \varphi_1 = \delta$. Отсюда для интенсивности на экране D в соответствии с (26.5) получаем

$$I = \frac{1}{4} E_0^2 (1 + \cos \delta) = \frac{1}{2} I_0 (1 + \cos \delta), \quad (26.13)$$

где $I_0 = E_0^2 / 2$ — интенсивность входящей в интерферометр волны от источника S_0 . Таким образом, при $\delta = (2m+1)\pi$ ($m = 0, \pm 1, \pm 2, \dots$) интенсивность равна нулю. Это означает, что никакого потока энергии в направлении экрана D от пластины O нет, т. е. весь поток энергии возвращается в направлении источника S_0 . Если же $\delta = 2m\pi$ ($m = 0, \pm 1, \pm 2, \dots$), то интенсивность на экране D равна I_0 , т. е. вся энергия, идущая от S_0 , попадает на экран и нет потока энергии, возвращающегося в направлении источника S_0 . Таким образом, поток энергии к D зависит от δ .

Это можно понять, только рассматривая все волновое поле как целое. Нельзя представить рассматриваемое явление как результат последовательности событий, происходящих на лучах. Ситуация здесь аналогична той, которая имеет место в квантовой механике и дискутируется обычно в связи с парадоксом Эйнштейна — Подольского — Розена. Поэтому необходимо ее кратко описать.

С учетом закона сохранения энергии для лучей (но не граничных условий для волн) векторная диаграмма сложения комплексных амплитуд лучей, отразившихся от пластины и прошедших через нее, изображена на рис. 98, *a*, где E_0 , $E_{\text{от}}$, $E_{\text{пр}}$ — комплексные амплитуды падающего, отраженного и прошедшего луча, $|E_{\text{от}}| = |E_{\text{пр}}| = |E_0| / \sqrt{2}$, $E_0 = E_{\text{от}} + E_{\text{пр}}$. Если разность хода лучей при возвращении к пластине составляет целое число длин волн, то фазовое отношение между ними, принятное на рис. 98, *a*, не изменится и каждый из них разделится на два [рис. 98, *b*]; $(E_{\text{от}})_{\text{от}}$ и $(E_{\text{пр}})_{\text{от}}$ — комплексные амплитуды отраженных лучей, которые при первом прохождении пластины были соответственно отраженным и преломленным; $(E_{\text{от}})_{\text{пр}}$ и $(E_{\text{пр}})_{\text{пр}}$ — комплексные амплитуды преломленных лучей, которые при первом прохождении пластины были соответственно отраженным и преломленным: $E_{\text{от}} = (E_{\text{от}})_{\text{от}} + (E_{\text{от}})_{\text{пр}}$, $E_{\text{пр}} = (E_{\text{пр}})_{\text{от}} + (E_{\text{пр}})_{\text{пр}}$, $|(E_{\text{от}})_{\text{от}}| = |(E_{\text{от}})_{\text{пр}}| = |E_{\text{от}}| / \sqrt{2}$, $|(E_{\text{пр}})_{\text{от}}| = |(E_{\text{пр}})_{\text{пр}}| = |E_{\text{пр}}| / \sqrt{2}$.

Из рис. 98, *b* видно, что $(E_{\text{от}})_{\text{от}} + (E_{\text{пр}})_{\text{пр}} = 0$, $(E_{\text{от}})_{\text{пр}} + (E_{\text{пр}})_{\text{от}} = 2(E_{\text{от}})_{\text{пр}} = 2(E_{\text{пр}})_{\text{от}}$.

Следовательно, в направлении источника S_0 (см. рис. 96) луча нет, а в направлении наблюдателя D имеется луч с амплитудой $|(E_{\text{от}})_{\text{пр}} + (E_{\text{пр}})_{\text{от}}| = 2|(E_{\text{от}})_{\text{пр}}| = \sqrt{2}|E_{\text{от}}| = E_0$, т. е. с амплитудой первоначального луча. В уравнении (26.13) это соответствует $\cos \delta = 1$.

В этом рассуждении, чтобы избежать понятия волны, необходимо было понятия фазы и амплитуды отнести к точке луча. При этом уже разделение луча на два содержит в себе противоречие с поведением фаз при делении волн. Понятия фазы и амплитуды нельзя отнести к точке луча. Противоречие сохраняется и в том случае, когда луч делится с учетом граничных условий для волн. В этом случае на диаграмме рис. 98, *a* (для напряженности электрического поля, перпендикулярной плоскости падения) направление комплексной амплитуды $E_{\text{от}}$ меняется на обратное, $E_0 + E_{\text{от}} = E_{\text{пр}}$ и соответствующим образом изменяется диаграмма на рис. 98, *b*. Это означает, что некоторые проблемы интерпретации квантовой механики содержатся уже в классической физике.

Интерференция монохроматических волн, распространяющихся под углом к оси интерферометра. Чтобы рассчитать разность хода между лучами, приходящими в D и порожденными лучом, который от S_0 распространяется под углом к оси S_0O (см. рис. 96), проще всего построить эквивалентную схему изображения точечного источника в зеркалах A_1 и A_2 , которые наблю-

● Образование интерференционной картины нельзя представить как результат последовательности событий, происходящих на лучах.

даются в D , причем все элементы системы удобно расположить вдоль одной оси (рис. 99). A'_1 представляет зеркало A_1 (см. рис. 96) интерферометра Майкельсона, а A'_2 — зеркало A_2 . Точечный источник S_0 (см. рис. 96) представлен на рис. 99 точкой S'_0 . Точки S'_1 и S'_2 являются изображениями точки S_0 в зеркалах A_2 и A_1 соответственно. Расстояние между зеркалами A_1 и A_2 равно $d = l_1 - l_2$, а расстояние между изображениями S'_1 и S'_2 в два раза больше. Отсюда для разности хода получаем формулу

$$\Delta = 2d \cos \theta = 2(l_1 - l_2) \cos \theta. \quad (26.14)$$

Разность фаз между лучами равна

$$\delta = k\Delta = (4\pi/\lambda)(l_1 - l_2) \cos \theta. \quad (26.15)$$

Условие максимума интенсивности при интерференции ($d > 0$)

$$\delta = 2\pi m \quad (m = 0, 1, 2, \dots) \quad (26.16)$$

или, поскольку $\delta = 2\pi\Delta/\lambda$,

$$\Delta = m\lambda. \quad (m = 0, 1, 2, 3 \dots). \quad (26.17)$$

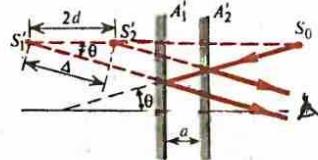
От точечного источника, расположенного на оси OS_0 , лучи аксиально-симметрично расходятся под всевозможными углами θ . Разность хода зависит только от этого угла (все остальные параметры фиксированы). Следовательно, картина аксиально-симметрична. Лучи, идущие под углами, соответствующими условиям (26.17), образуют при интерференции окружность максимальной интенсивности. Таким образом, интерференционная картина состоит из чередующихся между собой окружностей изменяющейся интенсивности (интерференционных колец, рис. 100). Целое число m в (26.17) называется порядком интерференции. Подставив в (26.17) выражение для Δ из (26.14), запишем условие (26.17) в виде

$$2d \cos \theta = m\lambda. \quad (26.18)$$

Отсюда видно, что уменьшающимся значениям m соответствуют увеличивающиеся значения угла θ , т.е. кольца больших радиусов соответствуют меньшим порядкам интерференции. При $m=1$ равенство (26.18) удовлетворяется при углах θ , близких к $\pi/2$. Это означает, что первый порядок интерференции теоретически реализуется кольцом очень большого радиуса, практически равного бесконечности, и наблюдение начинается с кольца большого радиуса, соответствующего высокому порядку интерференции. С увеличением порядка радиус интерференционного кольца уменьшается.

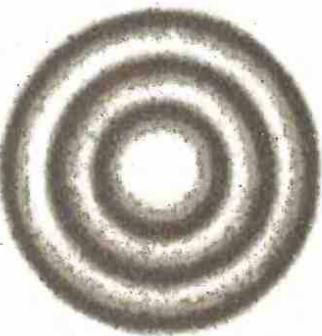
При увеличении разности хода d полосы интерференции размываются и интерференционная картина пропадает.

Причина размывания полос интерференции. При интерференции от строго монохроматического источника света никакого размывания полос при увеличении разности хода не должно быть. Причина размывания связана с конечным временем и дли-



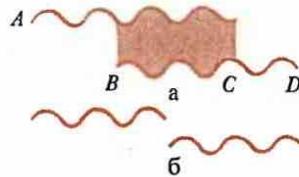
99

К расчету разности хода лучей в интерферометре Майкельсона



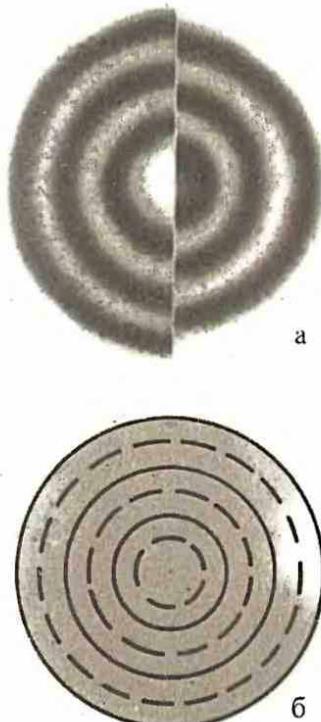
100

Изменение интенсивности излучения в интерферометре Майкельсона



101

Интерпретация смазывания картины интерференции конечной длиной цуга волн



102

Интерпретация смазывания картины интерференции немонокроматично излучения

ной когерентности света, излучаемого источником. Объяснить это можно двумя полностью эквивалентными способами.

Если интерпретировать волну с конечным временем и длиной когерентности как гармоническую волну определенной частоты и конечной продолжительности, то две интерферирующие волны с разностью хода представляются, аналогично, отрезками синусоид, сдвинутых (в пространстве и времени) друг относительно друга (рис. 101). Интерференционный вклад, пропорциональный $\cos \delta$, возникает лишь за счет участков волн BC (рис. 101, а), а в остальное время, соответствующее участкам AB и CD , интерференционный член отсутствует и наблюдается лишь сложение интенсивностей соответствующих волн. Таким образом, благодаря сдвигу синусоид интенсивность интерференционной картины ослабляется и суммарная интенсивность приближается к сумме интенсивностей двух интерферирующих волн. Когда разность хода становится столь большой, что отрезки синусоид на экран приходят в разное время и не перекрываются (рис. 101, б), интерференционная картина полностью пропадает.

Другая интерпретация волны с конечным временем и длиной когерентности состоит в том, что эта волна немонокроматическая, а ширина ее спектра связана с протяженностью волны по времени соотношением (12.13), в котором Δt при рассматриваемой интерпретации играет роль времени когерентности. От различных частот немонокроматической волны в интерферометре Майкельсона образуются кольца различных радиусов (рис. 102). На рис. 102, а изображены интерференционные кольца одинакового порядка от двух волн. Сплошные кольца соответствуют интерференции волн с меньшей длиной. Видно, что максимумы и минимумы интерференционной картины волн различной длины сдвинуты друг относительно друга и, следовательно, интерференционная картина размывается. При достаточно большой разнице в длинах волн максимумы в интерференционной картине одной длины волны попадают на минимумы другой, и интерференционная картина полностью пропадает (рис. 102, б).

Подтвердим эквивалентность этих подходов количественно. В направлении, определяемом углом θ , интерференционная картина полностью размывается при разности хода падающих лучей, равной длине когерентности, т. е.

$$2d \cos \theta \sim l_{\text{kog}} = c \tau_{\text{kog}}.$$

С другой стороны, картина размывается при совпадении максимума от волны длиной $\lambda + \Delta\lambda$ с минимумом от волны длиной λ . Это осуществляется при

$$\begin{aligned} 2d \cos \theta &= m(\lambda + \Delta\lambda) = (m + \frac{1}{2})\lambda, \quad \text{т. е. } m\Delta\lambda = \lambda/2, \quad 2d \cos \theta = \\ &= \lambda(\lambda + \Delta\lambda)/(2\Delta\lambda) \approx \lambda^2/(2\Delta\lambda) = \lambda v/(2|\Delta v|) \sim c \tau_{\text{kog}}, \quad \text{где учтено,} \\ &\text{что } \Delta\lambda/\lambda = -\Delta v/v, \quad \text{а соотношение } \Delta v \approx 1 \quad \text{определенено лишь} \\ &\text{с точностью до множителя порядка 1.} \end{aligned}$$

Интерференция немонохроматического света. Если излучение немонохроматично и имеет непрерывный спектр, то в интервале волновых чисел от k до $k + dk$ интенсивность пропорциональна интервалу dk :

$$dI_0 = (1/\pi) F(k) dk, \quad (26.18')$$

где множитель $1/\pi$ введен для того, чтобы согласовать $F(k)$ с определениями, используемыми в теории интегралов Фурье. В соответствии с (26.13) в результате интерференции при разности хода лучей Δ интенсивность описывается формулой

$$dI = 1/2 dI_0 (1 + \cos \delta) = [1/(2\pi)] F(k) (1 + \cos k\Delta) dk. \quad (26.19)$$

Полная интенсивность равна интегралу от (26.19) по всем волновым числам:

$$I(\Delta) = \int_{k=0}^{\infty} dI = [1/(2\pi)] \int_0^{\infty} F(k) (1 + \cos k\Delta) dk. \quad (26.20)$$

Спектр волн, образующих линию излучения, сосредоточен в очень малом интервале частот вблизи частоты с максимальной плотностью излучения даже тогда, когда уширение линии относительно велико (см. § 10). Поэтому $F(k)$ в (26.20) отлична от нуля лишь в узком интервале волновых чисел вблизи k_0 , соответствующего центру линии излучения. Переходя в (26.20) к новой переменной интегрирования

$$k' = k - k_0, \quad (26.21)$$

можно пределы интегрирования по k' считать равными $\pm \infty$ и записать (26.20) в виде

$$\begin{aligned} I(\Delta) &= [1/(2\pi)] \int_{-\infty}^{\infty} F(k_0 + k') \{1 + \cos [(k_0 + k')\Delta]\} dk' = \\ &= [1/(2\pi)] \int_{-\infty}^{\infty} \Phi(k') \{1 + \cos [(k_0 + k')\Delta]\} dk', \end{aligned} \quad (26.22)$$

где $\Phi(k') = F(k_0 + k')$. С помощью тригонометрического соотношения $\cos [(k_0 + k')\Delta] = \cos (k_0\Delta) \cos (k'\Delta) - \sin (k_0\Delta) \sin (k'\Delta)$ преобразуем (26.22) к равенству

$$I(\Delta) = Q + C \cos \beta - S \sin \beta, \quad (26.23)$$

где

$$\beta = k_0\Delta, \quad (a) \quad Q = [1/(2\pi)] \int_{-\infty}^{\infty} \Phi(k') dk' \quad (6)$$

$$C = [1/(2\pi)] \int_{-\infty}^{\infty} \Phi(k') \cos (k'\Delta) dk', \quad (b) \quad (26.24)$$

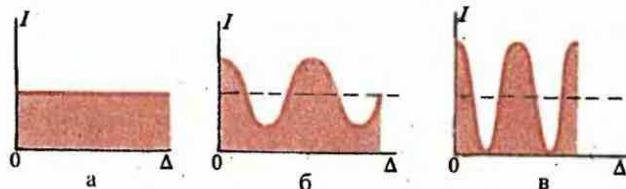
$$S = [1/(2\pi)] \int_{-\infty}^{\infty} \Phi(k') \sin (k'\Delta) dk'. \quad (c)$$

● **Интерференция определяется разностью фаз интерферирующих лучей.**

В оптике когерентные лучи получают делением амплитуды и делением фронта волны. Интерференцию также можно наблюдать от двух независимых лазерных источников.

Разрывание полос интерференции в интерферонете Майкельсона обусловлено конечностью времени и длины когерентности света, излучаемого источником, т. е. обусловлено временной когерентностью.

Зависимость интенсивности излучения от Δ при различных значениях видимости: $V = 0$ (а); $V = -0,5$ (б); $V = 1$ (в)



Из (26.23) видно, что интенсивность изменяется в зависимости от Δ по гармоническому закону. Выражение (26.23) удобнее представить в виде

$$I(\Delta) = Q + \sqrt{C^2 + S^2} \cos(\beta + \alpha), \quad (26.25a)$$

где α определяется из равенства

$$\cos \alpha = C / \sqrt{C^2 + S^2}, \quad \sin \alpha = S / \sqrt{C^2 + S^2}. \quad (26.25b)$$

Ясно, что максимумы и минимумы интенсивности достигаются соответственно при $\cos(\beta + \alpha) = 1$ и $\cos(\beta + \alpha) = -1$:

$$I_{\max} = Q + \sqrt{C^2 + S^2}, \quad I_{\min} = Q - \sqrt{C^2 + S^2}. \quad (26.26)$$

Чем больше разница между максимальной и минимальной интенсивностями, тем отчетливее интерференционная картина. Отчетливость интерференционной картины количественно характеризуется ее видимостью

$$V = (I_{\max} - I_{\min}) / (I_{\max} + I_{\min}). \quad (26.27)$$

Максимальная видимость ($V=1$) достигается при $I_{\min} = 0$, а минимальная ($V=0$) — при $I_{\max} = I_{\min}$ т. е. когда интерференционная картина отсутствует. В общем случае для видимости из (26.27) на основании (26.26) можно написать выражение

$$V = \sqrt{C^2 + S^2} / Q. \quad (26.28)$$

Изменение интенсивности I в зависимости от Δ при различных значениях видимости показано на рис. 103, а, б, в.

Принцип Фурье-спектроскопии. Равенство (26.20) при $\Delta=0$ принимает вид

$$I(0) = 2[1/(2\pi)] \int_0^\infty F(k) dk, \quad (26.29)$$

поэтому (26.20) можно представить в виде

$$I(x) = 1/2 I(0) + [1/(2\pi)] \int_0^\infty F(k) \cos kx dk, \quad (26.30)$$

где $x = \Delta$. Из (26.30) следует, что функция

$$f(x) = I(x) - I(0)/2, \quad (26.31)$$

называемая **интерферограммой**, связана с $F(k)$ равенством

$$f(x) = \frac{1}{2\pi} \int_0^\infty F(k) \cos kx dk. \quad (26.32)$$

которое является интегралом Фурье для действительных $F(k)$.
Обратное к (26.32) преобразование Фурье имеет вид

$$F(k) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(x) \cos kx dx. \quad (26.33)$$

Таким образом, представив интенсивность интерферограммы в виде функции разности хода лучей $\Delta = x$ и вычислив из (26.31) $f(x)$, с помощью (26.33) получаем спектральный состав излучения. Такой метод изучения спектров называется Фурье-спектроскопией. В Фурье-спектрометре создается изменяющаяся разность хода лучей и регистрируется соответствующая интенсивность интерферирующих лучей. После этого, обычно с помощью ЭВМ, получается вся необходимая информация о спектре излучения.

Из (26.31) видно, что кривые $I(x)$ и $f(x)$ имеют одинаковую форму и отличаются лишь нулевой точкой отсчета. На рис. 104 в качестве примера показано соотношение между этими кривыми. Поскольку результатом измерения в эксперименте является $I(x)$, а не $f(x)$, то $f(x)$ рассматривается лишь как вспомогательная величина для вычисления спектра и спектр соотносится не с $f(x)$, а с $I(x)$. На рис. 105 показаны характерные графики, иллюстрирующие соотношение между спектром и интенсивностью двухлучевого интерференционного сигнала.

Видимость при гауссовой форме линии. Формула (10.17), описывающая гауссову форму линии как функцию частоты, может быть выражена через волновые числа ($k = \omega/c$) и представлена в виде

$$F(k) = \frac{1}{\sigma' \sqrt{2\pi}} \exp \left[-\frac{(k - k_0)^2}{2\sigma'^2} \right], \quad (26.34)$$

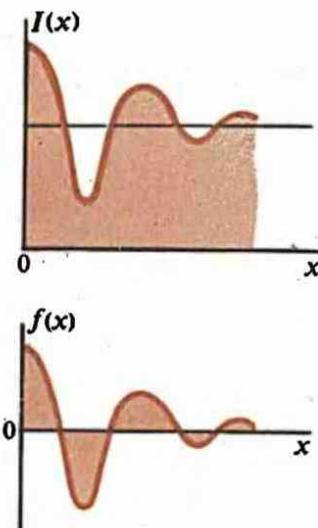
где $\sigma' = \sigma c$ [см. (10.17)]. Поэтому $\Phi(k')$, входящая в формулы (26.22) — (26.24), равна

$$\Phi(k') = F(k_0 + k') = [1/(\sigma' \sqrt{2\pi})] \exp [-k'^2/(2\sigma'^2)] \quad (26.35)$$

Учитывая, что $\Phi(k') = \Phi(-k')$, из (26.24г) получаем $S=0$ и поэтому

$$V = \frac{C}{S} = \frac{\int_{-\infty}^{\infty} \exp [-k'^2/(2\sigma'^2)] \cos (k'\Delta) dk'}{\int_{-\infty}^{\infty} \exp [k'^2/(2\sigma'^2)] dk'} = \exp (-2\pi^2 \Delta^2 \sigma'^2), \quad (26.36)$$

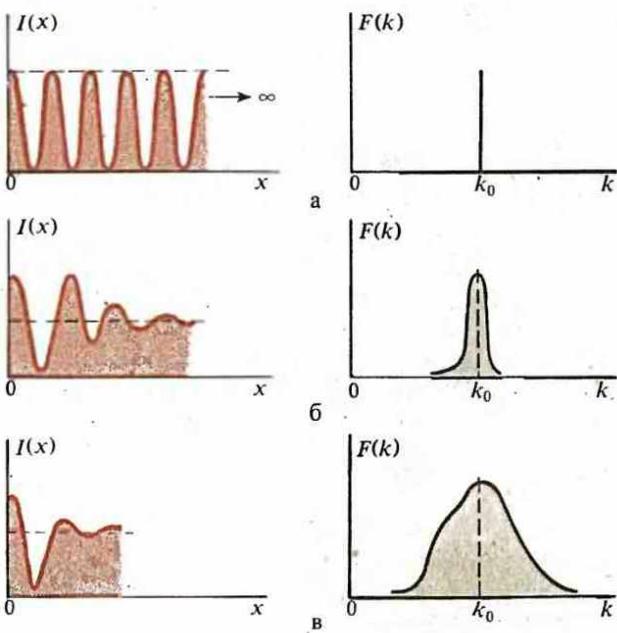
т. е. видимость описывается гауссовой функцией:



104

Соотношение между интенсивностью $I(x)$ и функцией $f(x)$

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}$$



105

Соотношение между интерференционным сигналом и спектром для монохроматического излучения (а), узкой (б) и широкой (в) линии излучения

Видимость при лоренцевой форме линии. В соответствии с (9.39) имеем

$$\Phi(k') = A/(k'^2 + a^2), \quad (26.37)$$

где $a = \gamma/(2c)$, A — постоянная. Отсюда $Q = 0$,

$$C = \frac{A}{2\pi} \int_{-\infty}^{\infty} \frac{\cos(\Delta k') dk'}{k'^2 + a^2} = \frac{A}{2a} e^{-|\Delta a|}. \quad (26.38)$$

Учитывая, что $Q = C(\Delta = 0)$, для видимости получаем

$$V = C/Q = \exp(-|\Delta a|). \quad (26.39)$$

Из сравнения (26.36) и (26.39) заключаем, что видимость $V(\Delta)$ интерференционной картины, соответствующей малым Δ , лучше для гауссовой формы линий, а соответствующей большими Δ — лучше для лоренцевой формы линий (рис. 106).

Интерферометр Майкельсона с линейными полосами. Из соотношения для разности фаз

$$\delta = 2\pi\Delta/\lambda = (4\pi/\lambda)(l_1 - l_2) \cos\theta$$

следует, что при больших разностях хода лучей расстояние между полосами мало и полосы сливаются друг с другом.

- Что такое видимость интерференционной картины?
- В чем состоит принцип метода Фурье-спектроскопии?
- Каким образом контролируется качество изготовления зеркал, линз и призм с точностью до долей длины волны?

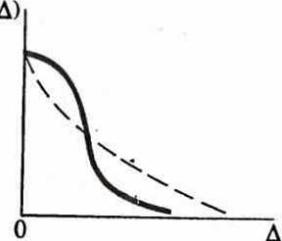
Для наблюдения интерференции при этих условиях пользуются несколько модифицированным интерферометром Майкельсона (рис. 107). Зеркало на пути одного из лучей интерферометра Майкельсона не строго перпендикулярно лучу. Следовательно, строго коллимированный пучок (т. е. плоская волна) падает на зеркало под небольшим углом и разность хода зависит от угла между лучами, который изменяется в плоскости падения луча. Поэтому полосы интерференции являются прямыми линиями и не сливаются друг с другом при большой разности хода.

Интерференционная картина от белого света. Каждая длина волны производит в интерферометре свою систему полос, причем полосы наблюдаются лишь при небольшой разности хода лучей в несколько длин волн (малая длина когерентности!). При нулевой разности хода лучей интерференция каждой длины волны происходит с одинаковым усилием и поэтому в центре наблюдается белое пятно. При разности хода лучей $\Delta = \lambda/2$ в результате интерференции гасятся. Поэтому белое пятно оказывается окруженным темной кольцевой областью. Дополнительная разность хода лучей, которая образуется в результате прохождения пластины, разделяющей лучи, или дополнительными прозрачными пластинами на пути лучей, может сделать центральное пятно темным, а темную кольцевую область — светлой.

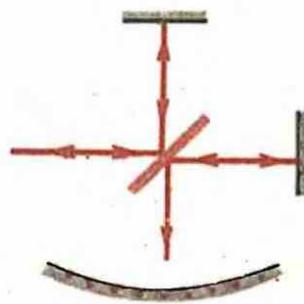
За темной кольцевой областью интерференционные круговые полосы различных длин волн начинают перекрываться. В результате образуются окрашенные кольца. Далее максимумы и минимумы интенсивности интерференционных колец различных порядков перекрываются и вся картина смазывается — никаких колец не наблюдается.

По интерференционной картине от белого света находят порядок наблюданной интерференции. Определить порядок интерференции по картине интерференции от монохроматического света невозможно, хотя положение интерференционных колец и можно фиксировать с большой точностью.

В связи с этим сделаем несколько замечаний о юстировке интерферометра Майкельсона. В разделительной пластине O (см. рис. 96) конечной толщины отражающей является обычно дальняя от источника поверхность. Сначала с помощью линейки уравниваются расстояния от центра ближней к источнику поверхности пластины до центров зеркал. Луч, идущий к D после отражения от A_2 , проходит через разделительную пластину три раза, а отраженный от A_1 — только раз. Поэтому для выравнивания оптических длин в плечо l_1 под углом 45° ставится пластина, идентичная разделительной, но без отражающего покрытия. Через эту пластину луч проходит дважды. Далее на оси прибора между источником S_0 и O помещают маленький («точечный») предмет, два изображения которого в зеркалах A_1 и A_2 наблюдаются из D . После этого от источника на разделительную пластину направляется монохрома-



Зависимость видимости от Δ для излучения с гауссовыми (сплошные) и лоренцевыми (пунктир) линиями



Модифицированный интерферометр Майкельсона, в котором полосы интерференции линейны

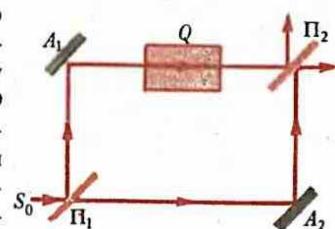
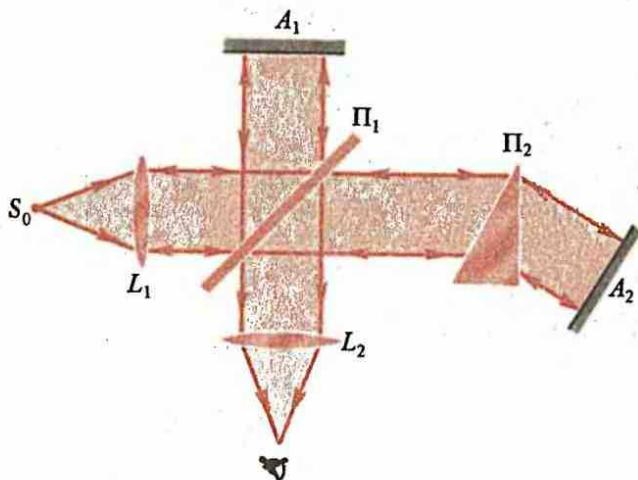


Схема интерферометра Маха — Цендера



109

Схема интерферометра Тваймана — Грина для проверки качества призмы P_2

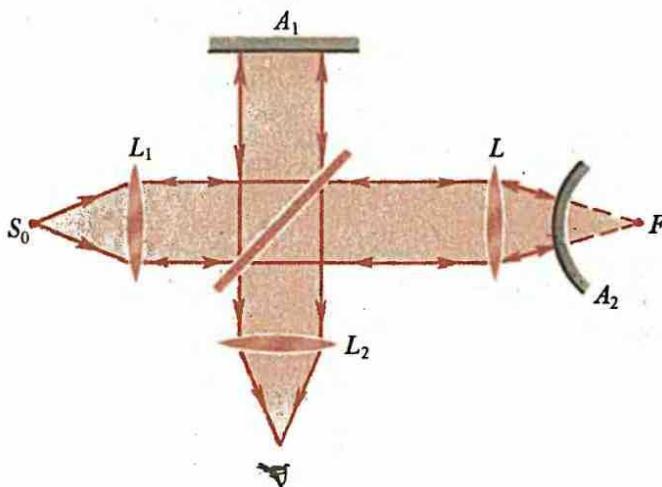
тический пучок света, поворотом зеркал добиваются совмещения между собой изображений предмета и совмещения предмета с центром возникающей системы интерференционных колец, а затем перемещением зеркал добиваются максимизации радиуса интерференционных колец, что соответствует равенству нулю разности оптических длин путей. Дополнительный контроль проводится освещением белым светом.

Явление интерференции в виде возникновения цветной окраски тонких пленок было впервые обнаружено Р. Гуком (1635—1703) и Р. Бойлем (1627—1691).

Интерферометр Маха — Цендера. Схема этого интерферометра показана на рис. 108. Полупрозрачной пластиной P_1 луч света S_0 разделяется на два. После отражения от зеркал A_1 и A_2 лучи света снова соединяются полупрозрачной пластиной P_2 в результате частичного отражения и прохождения через нее. Интерференция этих лучей приводит к возникновению картины, аналогичной наблюдаемой в интерферометре Майкельсона. Если на пути одного из лучей помещена ячейка Q с газом или веществом, показатель преломления которого отличен от единицы, то интерференционная картина изменится. По изменению интерференционной картины и длине пути светового луча в ячейке можно с большой точностью определить относительный показатель преломления, что позволяет изучать физические процессы, которые приводят к изменению показателя преломления.

Интерферометр Тваймана — Грина используется для контроля качества различных компонент оптических приборов. Схема интерферометра для исследования качества призмы показана на рис. 109.

От точечного монохроматического источника S_0 , помещенного в фокусе линзы L_1 , образуется плоская монохроматическая волна, которая падает на пластину P_1 и разделяется на две. После отражения от зеркал A_1 и A_2 они сводятся пластины P_1 и направляются к линзе L_2 . Если пластина P_1 и зеркала A_1 и A_2 оптически достаточно совершенны и сохраняют постоянство разности хода различных лучей пучка с точностью до долей длины волны, то на апертуре линзы L_2 будет наблюдаться равномерная освещенность. Если призма P_2 является оптически высокосовершенной, то разность хода интерферирующих лучей сохраняется постоянной по всему сечению пучка и, следовательно, апертура линзы L_2 будет по всему сечению освещена равномерно. Если же равномерность освещения нарушается и наблюдается интерференционная картина,



то призма недостаточно оптически совершенна. По наблюдаемой интерференционной картине определяют характер несовершенств призмы и устраниют их.

Проверка качества линзы осуществляется с помощью интерферометра Тваймана — Грина по схеме, показанной на рис. 110. Поверхность зеркала A_2 является сферической, отполированной с точностью до долей длины волны. Центр сферического зеркала находится в фокусе F исследуемой линзы L .

Интерферометр Тваймана — Грина позволяет обнаруживать, очень незначительные несовершенства элементов оптических приборов. С его помощью удается создать почти совершенные зеркала, линзы и призмы.

Интерферометр Жамена (рис. 111). В результате отражения лучей от передней и задней граней пластин Π_1 и Π_2 образуются четыре луча $1—4$, из которых два (луча 2, 3) пространственно совмещены и могут интерферировать. Разность хода, возникающая между лучами, отраженными от двух стенок толстой пластины, можно найти с помощью рис. 112. Толщина пластины — d , показатель преломления вещества пластины относительно вакуума — n . Из рис. 112 видно, что оптическая разность хода лучей равна

$$\Delta = (|AB| + |BC|)n - |AD|. \quad (26.40)$$

Учитывая, что

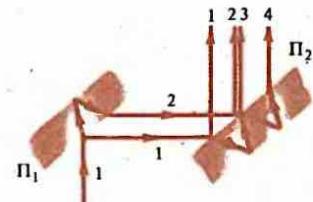
$$(|AB| + |BC|) = 2d / \cos \theta_{\text{пр}},$$

$|AD| = 2d \operatorname{tg} \theta_{\text{пр}} \sin \theta_{\text{пр}}, \sin \theta_{\text{пр}} / \sin \theta_{\text{нл}} = n$,
находим

$$\Delta = 2d(n / \cos \theta_{\text{пр}} - \operatorname{tg} \theta_{\text{пр}} n \sin \theta_{\text{пр}}) = 2nd \cos \theta_{\text{пр}}. \quad (26.41)$$

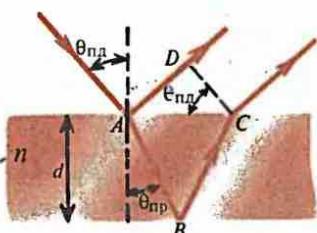
110

Схема интерферометра Тваймана — Грина для проверки качества линзы



111

Схема интерферометра Жамена



112

К расчету разности хода лучей, отраженных от двух стенок пластины

Разность хода лучей в интерферометре Жамена образуется в результате прохождения и отражения лучей в двух пластинах: Π_1 и Π_2 . На основании (26.41) она равна

$$\Delta = 2nd(\cos \theta_{\text{пр}1} - \cos \theta_{\text{пр}2}), \quad (26.42)$$

где $\theta_{\text{пр}1}$ и $\theta_{\text{пр}2}$ — углы преломления лучей в первой и второй пластинах.

Пример 26.1. Интерференционная картина в интерферометре Майкельсона наблюдается в фокальной плоскости линзы с фокусным расстоянием f . Определить радиус r_p светлого кольца, если известно, что центральное светлое пятно соответствует порядку интерференции m_0 . Длина волны λ .

Очевидно, что разность длин плеч интерферометра равна $l_1 - l_2 = d = m_0\lambda/2$. Условие m -го порядка интерференции имеет вид

$$2d \cos \theta = m\lambda. \quad (26.43)$$

Отсюда следует, что

$$1 - \cos \theta_p = 1 - m/m_0 = (m_0 - m)/m_0 = p/m_0, \quad (26.44)$$

где порядок интерференции уменьшается при увеличении угла θ ; $p = m_0 - m$ — номер p -го кольца в интерференционной картине. Учитывая, что $\theta_p \ll 1$, и принимая во внимание разложение $\cos \theta_p = 1 - \theta_p^2/2$, находим из (26.44) для угла θ_p отклонения луча, образующего p -е светлое кольцо, выражение

$$\theta_p = \sqrt{2p/m_0}. \quad (26.45)$$

Радиус кольца равен

$$r_p = 1 - f\theta_p = f\sqrt{2p/m_0}. \quad (26.46)$$

Если порядок m_0 интерференции не известен, а известна разность длин плеч в интерферометре, то вместо (26.46) находим

$$r_p = 1 - f\sqrt{p\lambda/d}. \quad (26.47)$$

Пример 26.2. Полученная в эксперименте методами Фурье-спектроскопии зависимость интенсивности $I(x)$ [см. (26.30) и далее] аппроксимируется с большой точностью формулой

$$I(x) = A + Be^{-ax^2}, \quad (26.48)$$

где A , B , a — постоянные. Найти спектральный состав излучения.

Поскольку нас интересует только спектральный состав излучения, существенные лишь относительные изменения интенсивностей. Поэтому в качестве $f(x)$ [см. (26.31)] можно взять

$$f(x) = (I(x) - B)/A = e^{-ax^2}. \quad (26.49)$$

Отсюда по формуле (26.33) находим

$$F(k) = 4 \int_0^\infty e^{-ax^2} \cos kx dx = 2\sqrt{\pi/a} e^{-k^2/(4a^2)}, \quad (26.50)$$

т. е. амплитуды экспоненциально убывают с ростом частоты. Форма спектра зависит лишь от относительного изменения амплитуд, и поэтому абсолютное значение $F(k)$ в (26.50) не имеет никакого значения, что уже было использовано в (26.49).

При практической реализации метода Фурье-спектроскопии необходимо принять во внимание, что (26.50), строго говоря, не дает точной формы спектра, потому что имеются многочисленные искажения зарегистрированной интенсивности (26.48). Они обусловлены конечно-сторонними амплитудами, несовершенством зеркал, шумовыми эффектами и рядом других факторов. Учет всех этих факторов составляет главную задачу при использовании метода.

Изучаются применения общей формулы для двухлучевой интерференции в схемах опытов с делением волнового фронта. Исследуется роль размеров источника в интерференции, обсуждается пространственная когерентность.

Принцип Гюйгенса. При обосновании волновой теории света Гюйгенс высказал принцип, позволивший ему просто и наглядно решить некоторые задачи, связанные с распространением и преломлением света. Он состоит в следующем. Если в некоторый момент времени известен фронт световой волны, то для определения положения фронта волны через промежуток времени Δt надо каждую точку фронта рассматривать как источник сферической волны и построить около источника сферу радиусом $c\Delta t$ (c — скорость света). Поверхность, огибающая вторичные сферические волны, представляет фронт волны через промежуток времени Δt .

Главная слабость принципа Гюйгенса в том, что он не учитывает важнейшего свойства волнового движения — явления интерференции. При учете интерференции принцип получил название принципа Гюйгенса—Френеля (см. § 31).

Схема Юнга. Простейший способ деления волнового фронта изображен на рис. 97. Щели A_1 и A_2 в соответствии с принципом Гюйгенса могут рассматриваться как источники волн. Эти источники, волны порождаются одной и той же первичной волной и поэтому взаимно когерентны. Между порожденными ими волнами наблюдается интерференция.

Рассмотрим интерференцию, возникающую в результате выделения с помощью щелей S_1 и S_2 двух участков волнового фронта излучения от точечного источника S (рис. 113). Такая схема интерференции была впервые осуществлена Т. Юнгом в 1801 г.

На рис. 113 изображено сечение волнового фронта и экрана со щелями плоскостью, проходящей через точечный источник и перпендикулярной экрану и щелям. Для того чтобы картина интерференции была одинаковой во всех плоскостях, параллельных плоскости рисунка, необходимо вместо точечного источника взять линейный источник в виде бесконечной нити, перпендикулярной плоскости рисунка, от которой распространяется цилиндрическая волна. Однако если нас интересует картина интерференции лишь в плоскости рисунка, источник можно считать точечным.

Однако Юнг в объяснении интерференции ограничился лишь качественными соображениями. Поэтому его идеи не получили общего признания. Волновая теория света восторжествовала лишь после работ Френеля.

Для расчета интерференции необходимо прежде всего найти разность хода лучей. Из рис. 113 видно, что

$$\Delta = \sqrt{l^2 + (y + d/2)^2} - \sqrt{l^2 + (y - d/2)^2}. \quad (27.1)$$

В случае $d \ll l$ можем записать

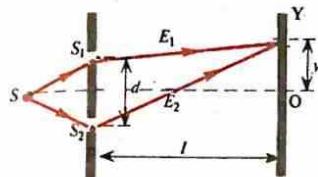
$$\sqrt{l^2 + (y \pm d/2)^2} = l[1 + (y \pm d/2)^2/(2l^2)] \quad (27.2)$$

и, следовательно, с точностью до величины первого порядка по d/l находим

$$\Delta = dy/l. \quad (27.3)$$

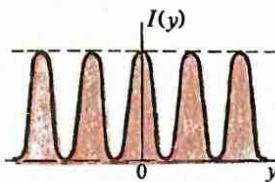
Отсюда следует, что разность фаз между лучами равна

$$\delta = 2\pi\Delta/\lambda = [2\pi d/(\lambda l)]y, \quad (27.4)$$



113

Схема интерференционного опыта Юнга



114

Изменение интенсивности излучения при интерференции от двух щелей для монохроматического света без учета изменения интенсивности при дифракции

а изменение интенсивности в зависимости от y в соответствии с (26.5) дается формулой

$$I(y) = 2I_0 \{1 + \cos[(2\pi d/(\lambda l))y]\}. \quad (27.5)$$

График изменения интенсивности показан на рис. 114. Заметим, что этот график не учитывает явлений дифракции в каждой из щелей. Если учесть дифракцию, то интенсивности в максимумах не будут постоянными (см. § 33).

Интерференция при белом свете. Каждая длина волн создает свою систему интерференционных полос, причем центральный максимум для всех длин волн совпадает. Поэтому центральный максимум имеет вид белой полосы. Первые минимумы всех длин волн весьма близки и не перекрываются с интерференционными полосами высших порядков. Поэтому к белой центральной полосе прилегают черные полосы. Следующие интерференционные полосы по тем же причинам, которые были подробно рассмотрены ранее (см. § 26), окрашены. Еще более дальние полосы смазываются, и интерференционная картина пропадает.

Источник конечного размера. Его можно представить как сумму некогерентных между собой точечных источников.

Пусть u — текущая координата точки протяженного источника от оси OO , относительно которой ведется отсчет (рис. 115). Для простоты рассматриваем линейный источник. В соответствии с (27.3) разность хода лучей от точки с координатой u и до точки, характеризуемой координатой y , равна

$$\Delta = (d/l)y + (d/L)u,$$

а возникшая за счет этого разность фаз

$$\delta = (kd/l)y + (kd/L)u \quad (k = 2\pi/\lambda). \quad (27.6)$$

Если $i(u)du$ — интенсивность света, испущенного с участка источника длиной du в точке u источника, то вклад от этого участка в полную интенсивность интерференционной картины в точке y на основании (27.5) выражается формулой

$$dI(y) = 2i(u)(1 + \cos \delta)du, \quad (27.7)$$

где δ дается равенством (27.6). Отсюда для полной интенсивности находим выражение

$$I(y) = 2 \int_{-\infty}^{\infty} i(u)(1 + \cos \delta)du. \quad (27.8)$$

Принимая во внимание, что

$$\cos \delta = \cos(kdy/l + kdu/L) = \cos(kdy/l) \cos(kdu/L) -$$

$$- \sin(kdy/l) \sin(kdu/L),$$

представим (27.8) в форме

$$I(y) = 2 \int_{-\infty}^{\infty} i(u)du + 2 \cos(kdy/l) \int_{-\infty}^{\infty} i(u) \cos(kdu/L)du - \\ - 2 \sin(kdy/l) \int_{-\infty}^{\infty} i(u) \sin(kdu/L)du. \quad (27.9)$$

115

К расчету интерференционной картины от протяженного источника

5

$$Q = 2 \int_{-\infty}^{\infty} i(u) du,$$

$$C = 2 \int_{-\infty}^{\infty} i(u) \cos(kdu/L) du,$$

$$S = 2 \int_{-\infty}^{\infty} i(u) \sin(kdu/L) du$$

это выражение можно записать аналогично (26.23) или (26.25а):

$$I(y) = Q + C \cos(kdy/l) - S \sin(kdy/l), \quad (27.11)$$

$$I(y) = Q + \sqrt{C^2 + S^2} \cos[(kdy/l) + \alpha], \quad (27.12)$$

где α дается равенством (26.25б).

Следовательно, интенсивность изменяется по гармоническому закону, а максимумы и минимумы интенсивности равны:

$$I_{\max} = Q + \sqrt{C^2 + S^2}, \quad (a)$$

$$I_{\min} = Q - \sqrt{C^2 + S^2}. \quad (b)$$

Видимость интерференционной картины определяется соотношением (26.27), которое с учетом (27.13а) и (27.13б) приводит к выражению

$$V = \sqrt{C^2 + S^2 / Q}, \quad (27.14)$$

совпадающему с (26.28).

Анализ интерференционной картины сводится к вычислению Q , C , S . Рассмотрим некоторые важные случаи.

Источник с однородным распределением интенсивности излучения. Распределение яркости описывается формулой

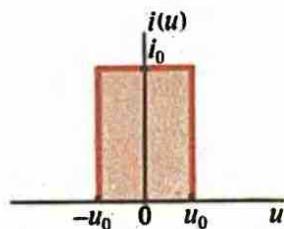
$$i(u) = \begin{cases} 0 & \text{при } -\infty < u < -u_0, \\ i_0 & \text{» } -u_0 < u < u_0, \\ 0 & \text{» } u_0 < u < \infty. \end{cases} \quad (27.15)$$

Ее график изображен на рис. 116. Длина светящегося источника равна $2u_0$. По формулам (27.10) находим:

$$Q = 2 \int_{-u_0}^{u_0} i_0 du = 4i_0 u_0,$$

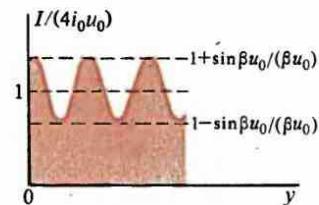
$$C = 2 \int_{-u_0}^{u_0} i_0 \cos(kdu/L) du = 4i_0 u_0 \sin(\beta u_0)/(\beta u_0), \quad (27.16)$$

$$S = 2 \int_{-u_0}^{u_0} i_0 \sin(kdu/L) du = 0,$$



116

График распределения яркости в источнике конечного размера



117

Распределение интенсивности в интерференционной картине от источника конечного размера

где $\beta = kd/L$. Следовательно, формула (27.11) принимает вид

$$I(y) = 4i_0 u_0 [1 + \{\sin(\beta u_0) / (\beta u_0)\} \cos(kdy/l)]. \quad (27.17)$$

График распределения интенсивности показан на рис. 117.

В соответствии с (27.14) и (27.16) видимость выражается формулой

$$V = |\sin(\beta u_0) / (\beta u_0)|. \quad (27.18)$$

Для точечного источника $u_0 \rightarrow 0$ и, следовательно, $V \rightarrow 1$. Это на основании (27.17) означает, что источник, описываемый ступенчатой функцией (27.15), дает такие же интерференционные полосы, как и точечный источник, но их видимость уменьшается, т. е. видимость полос зависит от размера источника.

Это обстоятельство позволяет сделать важное заключение: измеряя видимость полос, можно определить угловой размер источника. Практически такая возможность осуществляется в звездном интерферометре Майкельсона.

Временная и пространственная когерентности. При анализе явлений в интерферометре Майкельсона необходимо было принять во внимание временную когерентность. При анализе явлений интерференции делением волнового фронта необходимо учесть корреляцию фаз по фронту волны в одни и те же промежутки времени. Эта корреляция описывается понятием пространственной когерентности.

Как видно из (27.18), точечный монохроматический источник ($u_0 \rightarrow 0$) дает интерференционную картину с видимостью $V = 1$. Источник конечных размеров ($2u_0 \neq 0$), состоящий из точечных монохроматических не когерентных между собой источников, дает интерференционную картину с меньшей, чем единица, видимостью. Излучение источника конечных размеров не является когерентным, хотя оно и монохроматическое. Степень когерентности этого излучения можно характеризовать видимостью порождаемой им интерференционной картины. Если видимость равна нулю, то излучение полностью некогерентно, если $V = 1$, то излучение когерентно.

При промежуточных значениях видимости излучение частично когерентно.

Когерентного излучения, для которого $V = 1$, не существует, хотя к этому пределу в специальных случаях (например, излучение лазеров в одномодовом режиме) и можно очень близко подойти. Однако без специальных излучателей этот случай представляется лишь идеализацией. Поэтому необходимо условиться, при какой видимости излучение называть когерентным. Общепринятого соглашения по этому вопросу нет. Некоторые считают, что излучение может называться когерентным, если $V = 0,5$. Другие для V принимают значения $0,9; \pi/4; e^{-1}$ и т. д. В зависимости от этого получаются несколько различные величины, которыми характеризуется пространственная когерентность, — угол и ширина когерентности. Но это не имеет принципиального значения, потому что значения этих величин при различных определениях отличаются множителями порядка единицы.

Угол и ширина когерентности. Значение видимости, при котором излучение принимается когерентным, обозначим V_{kog} . Поэтому с учетом (27.18) условие когерентности принимает вид

$$|\sin(\beta u_0) / (\beta u_0)| = V_{\text{kog}}, \quad (27.19)$$

где $\beta = kd/L$. Решение уравнения (27.19) относительно βu_0 можно записать в форме

$$\beta u_0 = \theta_0, \quad (27.20)$$

где θ_0 зависит от V_{kog} . Например, при $V_{\text{kog}} = 0,5$ получаем $\theta_0 = 1,9$.

Угол когерентности⁴ $\varphi_{\text{кор}}$ определяется как максимальный угловой размер источника, излучение которого в соответствии с принятым критерием когерентности можно считать когерентным. Поскольку во всех представляющих интерес случаях $u_0 \ll L$, угол когерентности можно представить в виде

$$\varphi_{\text{кор}} = 2u_0/L, \quad (27.21)$$

где u_0/L должно удовлетворять равенству (27.20). Подставляя u_0/L из (27.20) в (27.21), получаем

$$\varphi_{\text{кор}} = 2\theta_0/(kd) = (\theta_0/\pi)(\lambda/d). \quad (27.22)$$

При $\theta_0 = 1,9$ имеем $\theta_0/\pi = 0,6$. При других значениях $V_{\text{кор}}$ получаются другие значения θ_0 и θ_0/π , но все они имеют одинаковый порядок величины. Характер зависимости угла когерентности от длины волны и расстояния между центрами щелей при этом одинаков.

Вместо угла когерентности можно пользоваться понятием ширины когерентности, определяя ее как то расстояние по фронту волны, на котором излучение в точках фронта может рассматриваться когерентным в соответствии с принятым критерием. Равенство (27.20) определяет расстояние $d_{\text{кор}}$ между центрами щелей, на которые падает фронт волны, удовлетворяющее этому условию. Из (27.20) получаем

$$d_{\text{кор}} = (\theta_0 L)/(ku_0) = (\theta_0/2\pi)(L\lambda/u_0). \quad (27.23)$$

При $\theta_0 = 1,9$ получаем $\theta_0/(2\pi) \approx 0,3$.

Эти соображения о пространственной когерентности в комбинации с тем, что было раньше (см. § 26) сказано о временной когерентности, свидетельствуют о сложности явления частичной когерентности. Адекватное описание этого явления возможно лишь в рамках общей теории случайных процессов (см. § 30). Связь явления частичной когерентности с теорией случайных процессов обусловлена физической природой излучения. В каждой точке напряженность электрического поля волны является суперпозицией напряженностей электрических полей от многих независимых излучателей, частоты, амплитуды и фазы волн от которых между собой не связаны. Поэтому суммарная напряженность электрического поля не представляет собой монохроматического излучения, а изменение амплитуд и фаз этого излучения имеет случайный характер.

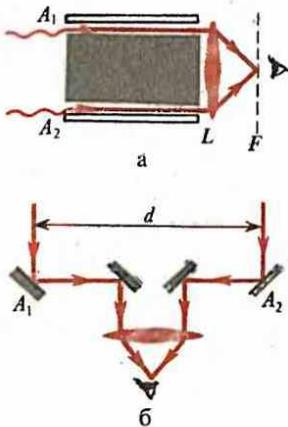
Аналогичными свойствами обладает, конечно, и индукция магнитного поля волны.

Причиной уменьшения видимости при увеличении размеров источника является смещение интерференционных полос от различных точек источника друг относительно друга. В результате на минимумы интенсивности от одних точек попадают области с ненулевой интенсивностью от других и вся картина смазывается. Ясно, что интерференционная картина будет сильно смазана, если в фиксированной точке у экрана разности хода лучей от крайних точек источника различаются на полволны. Поэтому в качестве условия когерентности можно принять соглашение, что разности хода лучей от крайних точек источника должны отличаться меньше чем на полволны. Это условие записывается в виде

$$\delta\Delta = (d/L)\delta u = (d/L)2u_0 = \lambda/2,$$

где учтено, что расстояние между крайними точками источника $\delta u = 2u_0$. Отсюда угол когерентности

$$\varphi_{\text{кор}} = 2u_0/L = \lambda/(2d). \quad (27.24)$$



118

Схема интерферометра Рэлея (а); схема звездного интерферометра Майкельсона (б)

Из (27.24) находим ширину когерентности

$$d_{\text{ког}} = \lambda L / (4u_0). \quad (27.25)$$

Формулы (27.4) и (27.5) дают хорошее согласие соответственно с (27.22) и (27.23).

Основной физический вывод из сказанного о пространственной когерентности может быть сформулирован так: излучение от некогерентного протяженного монохроматического источника может рассматриваться как когерентное на площадке, линейные размеры которой имеют порядок ширины когерентности, причем источник излучения при этом из центра, площадки виден под углом когерентности. Например, светящаяся булавочная головка на вытянутой руке создает на входном зрачке человеческого глаза когерентное излучение.

Звездный интерферометр. Измерение диаметров звезд. Из (27.18) следует, что видимость зависит не только от размеров $2u_0$ источника, но и от расстояния между щелями. Точнее говоря, она зависит от расстояния между щелями и угла, под которым виден источник из средней точки между щелями, поскольку

$$\beta u_0 = kdu_0/L = (k/2)d\phi, \quad (27.26)$$

где $\phi = 2u_0/L$ — угол, в котором наблюдается источник. Поэтому, измерив видимость интерференционной картины и зная расстояние d между щелями, можно определить угловые размеры источника излучения, например звезды. Это позволяет создать звездный интерферометр, с помощью которого можно измерить угловые размеры астрономических объектов. Идея звездного интерферометра была выдвинута впервые в 1868 г. Физо.

Прибором, принципиально пригодным для этой цели, является интерферометр Рэлея (рис. 118, а). Из плоской волны, идущей от удаленной звезды, в A_1 и A_2 выделяются два параллельных пучка света, которые, выйдя из прибора, дают в фокальной плоскости F линзы L дифракционную картину, позволяющую измерить угловой размер источника. Однако применить интерферометр Рэлея для измерения угловых размеров астрономических объектов оказалось невозможным. Интерференционные полосы получаются очень узкими и проводить измерение трудно. Большие осложнения также связаны с обеспечением точности взаимного положения трубок на больших расстояниях между ними.

Развитие идеи Физо для измерения угловых размеров астрономических объектов принадлежит Майкельсону, который практически реализовал ее в 1920 г.

Схема звездного интерферометра показана на рис. 118, б. Роль щелей выполняют подвижные зеркала A_1 и A_2 , с помощью которых можно изменять базу d (расстояние между щелями). Отраженные от зеркал лучи после вторичного отражения направляются через линзу в прибор, регистрирующий интенсив-

• Видимость интерференции от двух щелей обусловлена конечным значением ширины когерентности световой волны, т. е. обусловлена пространственной когерентностью.
Определение угловых размеров звезд в звездном интерферометре основывается на измерении видимости интерференционной картины

нность интерферирующих лучей. Наблюдаемая в центре интенсивность определяется формулой (27.17) при $y=0$:

$$I(d) = 4i_0 u_0 \left[1 + \frac{\sin(kd\varphi/2)}{kd\varphi/2} \right], \quad (27.27)$$

где принято во внимание соотношение (27.26). Видимость интерференционной картины на основании (23.18) дается функцией

$$V(d) = \left| \frac{\sin(kd\varphi/2)}{kd\varphi/2} \right|, \quad (27.28)$$

график которой представлен на рис. 119.

При увеличении расстояния d от нуля яркость центрального пятна постепенно уменьшается и при $d_0 = 2\pi/(k\varphi)$ пятно полностью размывается и видимость интерференционной картины становится равной нулю. По значению d_0 находим угловой размер наблюдаемой звезды:

$$\varphi = 2\pi/(kd_0) = \lambda/d_0. \quad (27.29)$$

Первой звездой, угловой размер которой был измерен таким способом, была Бетельгейзе ($\varphi = 0,047''$). Расстояние до этой звезды было известно по земному параллаксу. По известным угловым размерам и расстоянию можно рассчитать линейные размеры звезды. Диаметр Бетельгейзе оказался примерно в 300 раз больше диаметра Солнца. Так же были измерены диаметры некоторых других звезд и объектов Солнечной системы.

Измерение расстояния между компонентами двойной звезды. Если размеры звезд двойной системы много меньше расстояния между ними, то звезды можно считать точечными источниками и рассматривать интерференционную картину от двух точечных источников, находящихся на расстоянии $2u_0$ друг от друга, т. е. на расстоянии между звездами двойной системы (рис. 120). Распределение яркости излучателя дается формулой

$$i(u) = i_0 \delta(u + u_0) + \varepsilon i_0 \delta(u - u_0), \quad (27.30)$$

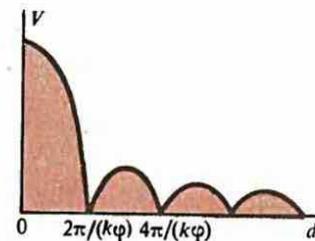
где ε характеризует соотношение яркостей компонент двойной системы; δ — дельта-функция. По формуле (27.9) с учетом (27.30) находим

$$I(y) = 2i_0 [(1 + \varepsilon) + (1 + \varepsilon) \cos \beta u_0 \cos(kdy/l) - (1 - \varepsilon) \sin \beta u_0 \sin(kdy/l)], \quad (27.31)$$

где β имеет то же значение, что и в (27.16). Отсюда по формуле (27.14) получаем

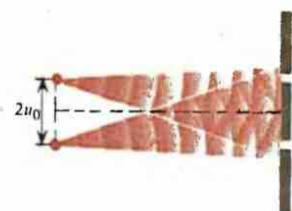
$$V = [(1 + \varepsilon^2) + 2\varepsilon \cos 2\beta u_0]^{1/2}/(1 + \varepsilon), \quad (27.32)$$

причем, аналогично (27.26), $2\beta u_0 = kd\varphi$, где φ — угловое расстояние между компонентами двойной звезды. График видимости $V(d)$ показан на рис. 121. При увеличении d от 0 до $\lambda/(2\varphi)$ видимость уменьшается от 1 до $(1 - \varepsilon)/(1 + \varepsilon)$. Задача заклю-



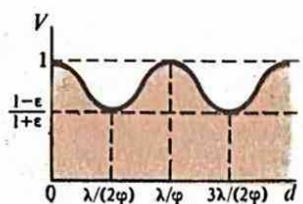
119

Видимость интерференционной картины в звездном интерферометре



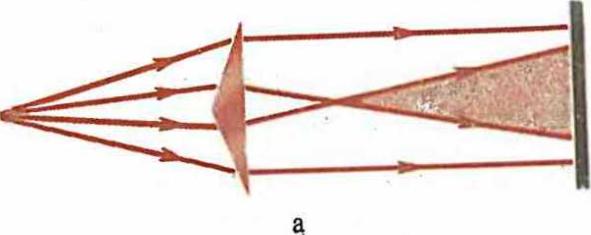
120

Схема интерференции от излучения двойной звезды

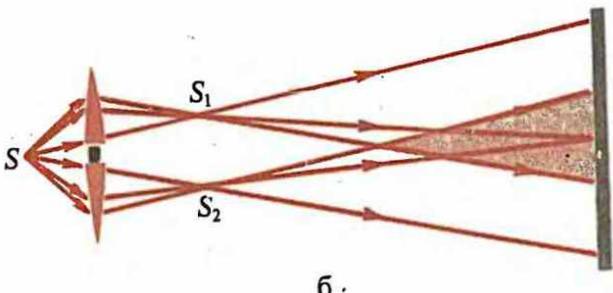


121

Видимость интерференционной картины от излучения двойной звезды



а



б.

122

Бипризма Френеля (а); билинза
Бийе (б)

чается в том, чтобы зафиксировать значение d_0 , при котором видимость достигает первого минимума. Угловое расстояние между компонентами двойной системы вычисляется по формуле

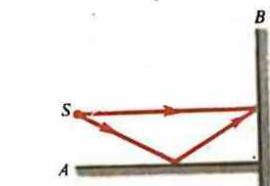
$$\varphi = \lambda / (2d_0). \quad (27.33)$$

Бипризма Френеля. Схема деления волнового фронта бипризмой Френеля показана на рис. 122, а. Падающий на призму волновой фронт преломляется в различных направлениях между компонентами двойной системы. Интерференционная картина возникает в области пересечения преломленных фронтов.

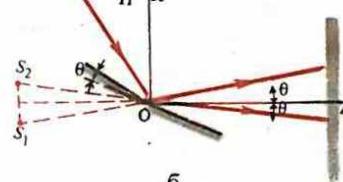
Билинза Бийе. Линза разрезается по диаметру, и половинки линзы разводятся на некоторое расстояние. Промежуток между разведенными линзами закрывается непрозрачным экраном. Такая система называется билинзой Бийе (рис. 122, б). От точечного источника S образуются два действительных изображения: S_1 и S_2 . Расходящиеся от S_1 и S_2 волны в области их перекрытия создают интерференционную картину.

Зеркало Ллойда. Схема деления волнового фронта и осуществления интерференции с помощью зеркала Ллойда показана на рис. 123, а. Одна часть волнового фронта от источника S падает непосредственно на экран B , а другая — после отражения от зеркала A . В области пересечения фронтов происходит интерференция.

Бизеркало Френеля (рис. 123, б). Свет, испущенный точечным источником S , отражается от двух зеркал, расположенных под углом θ , близким к 0. В области перекрытия отраженных



а



б

123

Зеркало Ллойда (а); бизеркало
Френеля (б)

пучков происходит интерференция. Область интерференции от прямых лучей источника закрывается экраном P . В точках S_1 и S_2 расположены мнимые изображения источника.

Закон сохранения энергии в явлениях интерференции. Обычно явление интерференции сопровождается перераспределением энергии между различными участками волнового поля при безусловном выполнении закона сохранения энергии. В применении к интерферометру Майкельсона это было подробно рассмотрено при обсуждении формулы (26.13).

Однако отсюда не следует, что явление интерференции не влияет на сам процесс излучения электромагнитных волн. Из решения задачи 5.1. видно, что при расстоянии $\lambda/2$ между двумя источниками излучения мощность их совместного излучения не равна сумме мощностей излучения этих источников в отдельности. Таким образом, источники излучения влияют друг на друга — их излучение является **коллективным**. Когда расстояние между источниками много больше длины испускаемой волны, их взаимное влияние пренебрежимо мало.

Если рассчитывать полную интенсивность интерференционной картины волнового поля от двух источников, то она, вообще говоря, не равна сумме полных интенсивностей интерференционных картин от двух изолированных источников. Но это не означает нарушения закона сохранения энергии. **Различие в энергии полей полностью объясняется изменением мощности излучателей.** Например, если полная интенсивность увеличивается в результате интерференции, то мощность взаимодействующих между собой через поле излучателей должна увеличиться. Если по своим внутренним свойствам излучатели не в состоянии увеличить мощность излучения, то полная интенсивность интерференционной картины соответствующим образом уменьшается.

Пример 27.1. Проанализировать интерференцию, создаваемую с помощью бизеркала Френеля (рис. 123, б).

После отражения от зеркал напряженности электрического поля волн, распространяющихся в указанных на рис. 123, б стрелками направлениях, даются выражениями

$$E_1 = E_0 \exp [-i(\omega t - kz \cos \theta + kx \sin \theta)], \quad (27.34)$$

$$E_2 = E_0 \exp [-i(\omega t - kz \cos \theta - kx \sin \theta)]. \quad (27.35)$$

Отсюда

$$E = E_1 + E_2 = E_0 \exp [-i(\omega t - kz \cos \theta)] [\exp (ikx \sin \theta) + \exp (-ikx \sin \theta)]. \quad (27.36)$$

И, следовательно, интенсивность интерференции дается формулой

$$I = \frac{1}{2} EE^* = 2E_0^2 \cos^2 (kx \sin \theta) = E_0^2 [1 + \cos (2kx \sin \theta)]. \quad (27.37)$$

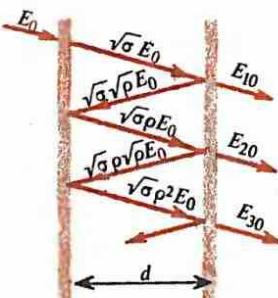
Геометрическое значение x и $\sin \theta$ очевидно из рис. 123, б. Видно, что картина интерференции точно такая же, как если бы когерентные источники (щели) находились в точках S_1 и S_2 . Прямые лучи от источника S в область, где наблюдается интерференция, задерживаются экраном P .

○ При каких условиях излучение от некогерентного протяженного источника может рассматриваться как когерентное?

В чем состоит причина уменьшения видимости интерференционной картины при увеличении размеров источника?

§ 28 Многолучевая интерференция, осуществляемая делением амплитуды

Излагается общий метод исследования многолучевой интерференции и его реализации для конкретных интерферометров. Определяются разрешающая способность и дисперсионная область спектральных аппаратов.



124

Схема лучей в интерферометре Фабри — Перо

Интерферометр Фабри — Перо. Рассмотрим последовательные частичные отражения и прохождения света через две стеклянные пластины, внутренние поверхности которых строго параллельны друг другу (рис. 124), отполированы с большой точностью (от $1/20$ до $1/200$ длины волны) и покрыты сильно отражающими пленками. Пленки могут быть металлическими (серебро, золото, алюминий) или состоять из нескольких диэлектрических слоев, подобранных так, чтобы получился очень большой коэффициент отражения (см. § 29). Внешние поверхности стеклянных пластин наклонены под небольшим углом (порядка $0,1^\circ$) к внутренним поверхностям, чтобы отражения от них уводились в сторону и не смешивались с лучами, отраженными от внутренних рабочих поверхностей. Однако энергия, связанная с этими отражениями, незначительна и в последующем расчете не учитывается. Кроме того, нет необходимости также учитывать поглощение света при прохождении света через стеклянную пластину. Ослабление амплитуды при отражении характеризуется коэффициентом отражения ρ [см. (18.5)]. Отношение амплитуды отраженной волны к амплитуде падающей равно $\sqrt{\rho}$ (рис. 124). Для характеристики прохождения волны через пластину пользоваться коэффициентом пропускания τ [см. (18.9)] неудобно, поскольку он связывает амплитуду волны внутри стекла с амплитудой волны вне стекла, а в данном случае удобнее связать между собой амплитуды волн по разные стороны стеклянной пластины. Обозначим отношение модуля амплитуды прошедшей через пластину волны к модулю амплитуды падающей $\sqrt{\sigma}$:

$$|E_{0\text{пр}} / E_{0\text{пад}}| = \sqrt{\sigma}. \quad (28.1)$$

Иначе говоря, σ является коэффициентом пропускания пластины. Поскольку поглощением в пластине и отражением на внешней поверхности пластины пренебрегаем, закон сохранения энергии записывается аналогично (8.12):

$$\rho + \sigma = 1. \quad (28.2)$$

На основании рис. 124 заключаем:

$$E_{10} = \sigma E_0,$$

$$E_{20} = \sigma \rho E_0,$$

$$E_{30} = \sigma \rho^2 E_0,$$

$$\dots \dots \dots$$

$$E_{0N} = \sigma \rho^{N-1} E_0,$$

где E_0 — амплитуда падающей волны, E_{10}, \dots, E_{N0} — амплитуды соответствующих волн, интерферирующих между собой. Разность хода для соседних лучей легко находится с помощью построения, изображенного на рис. 125. Видно, что $\Delta = 2l - p$. Учитывая равенства $l = d/\cos \theta$, $p = 2s \sin \theta = 2(dtg \theta) \sin \theta$, получаем

$$\Delta = 2d/\cos \theta - 2d \sin^2 \theta / \cos \theta = 2d \cos \theta. \quad (28.4)$$

Если прошедшие волны перекрываются, либо с помощью линзы сводятся в одну точку, наблюдается интерференция. Максимальная интенсивность возникает при $\Delta = m\lambda$. Разность фаз между соседними лучами равна

$$\delta = k\Delta = (4\pi/\lambda) d \cos \theta. \quad (28.5)$$

При интерференции N прошедших лучей комплексная амплитуда в соответствии с (28.3) дается выражением

$$E = \sigma E_0 + \sigma E_0 e^{i\delta} + \sigma E_0 e^{2i\delta} + \dots + \sigma E_0 e^{(N-1)i\delta} = \\ = \sigma E_0 (1 + \rho e^{i\delta} + \rho^2 e^{2i\delta} + \dots + \rho^{N-1} e^{(N-1)i\delta}). \quad (28.6)$$

Плоская волна пространственно не ограничена и поэтому на выходе пространственно не ограниченного интерферометра между собой интерферирует бесконечное число лучей. Однако конечность поперечных размеров светового пучка и интерферометра обуславливает конечность числа интерферирующих лучей при угле падения, не равном нулю. Это число интерферирующих лучей всегда велико, и поэтому велико число членов геометрической прогрессии в (28.6), а ее последний член много меньше первого. Поэтому можно суммирование распространить до бесконечности и написать

$$E = \sigma E_0 / (1 - \rho e^{i\delta}). \quad (28.7)$$

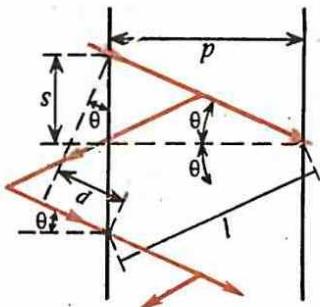
Следовательно, интенсивность интерференционной картины равна

$$I = E^* E / 2 = (\sigma^2 E_0^2 / 2) / (1 + \rho^2 - 2\rho \cos \delta). \quad (28.8)$$

Обозначая $E_0^2/2 = I_0$, учитывая, что $1 - \cos \delta = 2 \sin^2(\delta/2)$, и принимая во внимание закон сохранения энергии, преобразуем (28.8) к виду

$$I = \frac{I_0}{1 + 4\rho \sin^2(\delta/2) / (1 - \rho)^2}. \quad (28.9)$$

Распределение интенсивности в интерференционной картине. При $\sin(\delta/2) = 0$, т. е. когда $\delta = 2\pi n$ ($n = 1, 2, \dots$), наблюдается



125

К расчету разности хода между соседними лучами в интерферометре Фабри — Перо

$$E = \frac{\sigma E_0}{1 - \rho e^{i\delta}}$$

полное прохождение ($I_{\max} = I_0$), хотя каждая из поверхностей, через которые проходит свет, обладает большой отражательной способностью. Физическая причина возникновения этого явления заключается в том, что соседние волны, отраженные от каждой из поверхностей, имеют у первой поверхности разность фаз π и взаимно ослабляются, в результате чего отраженная волна отсутствует. У второй поверхности разность у «соседних» волн 2π . Они усиливают друг друга и проходят через эту поверхность без ослабления. «Соседними» для краткости названы волны, разность фаз между которыми минимальна.

При $\sin^2(\delta/2) = 1$, т. е. при $\delta = \pi(2m + 1)$ ($m = 1, 2, \dots$), интенсивность проходящего света достигает минимального значения

$$I_{\min} = I_0(1 - \rho^2)/(1 + \rho)^2. \quad (28.10)$$

При большой отражательной способности поверхностей $\rho \approx 1$ и поэтому интенсивность прошедшего света близка к нулю ($I_{\min} \approx 0$). Это означает, что практически весь свет отражается. Физическая причина этого явления состоит в том, что разность фаз между соседними волнами, прошедшими через вторую поверхность, равна π и они взаимно ослабляются, в результате чего прошедшая волна практически отсутствует.

Это замечательное явление возникновения практического полного пропускания или полного отражения как интерференционного эффекта имеет важное применение для создания высокoeffективных зеркал и фильтров (см. § 29).

Интерференционные кольца. Если на интерферометр падает пучок не абсолютно параллельных лучей, то в нем присутствуют лучи со всевозможными углами падения θ . Обычно распределение лучей по углам аксиально-симметрично. Расстояние d между пластинами интерферометра фиксировано. Формула (28.5) показывает, что разность фаз δ определяется в этом случае только углом θ (для данной длины волны λ) и, следовательно, интерференционная картина аксиально-симметрична. Она состоит из интерференционных колец. Поскольку лучи, соответствующие заданному углу θ , параллельны между собой, интерференционные кольца на экране наблюдаются с помощью линзы в ее фокальной плоскости. При наблюдении колец без экрана глаз адекватируется на бесконечность.

Для описания интерференционной картины удобнее пользоваться не углом θ , а разностью фаз δ , которая с углом θ связана соотношением (28.5). Из (28.9) следует, что видимость интерференционной картины зависит от множителя

$$A = 4\rho/(1 - \rho)^2. \quad (28.11)$$

Этот множитель сильно изменяется с изменением коэффициента отражения. Например, при $\rho = 0,5$ $A = 8$ и при $\rho = 0,9$ $A = 360$. Записав формулу (28.9) в виде

$$\frac{I}{I_0} = \frac{1}{1 + A \sin^2(\delta/2)}, \quad (28.12)$$

заключаем, что в первом случае это отношение изменяется от 1 до $1/9$, т. е. интенсивность в максимуме интерференционных колец больше интенсивности в минимуме в 9 раз. Во втором случае интенсивности отличаются в 361 раз. Этот пример иллюстрирует сильную зависимость видимости интерференционной картины от отражательной способности пластин. Ширина интерференционных максимумов также уменьшается с увеличением видимости, что обуславливается просто законом сохранения энергии при интерференции. На рис. 126 показаны графики распределения интенсивностей в интерференционной картине при различных отражательных способностях, которые можно построить по формуле (28.12).

Разрешающая способность. Если интерферометр облучается двумя волнами с длинами λ и $\lambda + \Delta\lambda$, то каждая из них создает свою картину интерференционных колец. Если длины волн

очень близки друг к другу, то интерференционные картины почти совпадают между собой и наличие двух накладывающихся друг на друга интерференционных картин обнаружить нельзя. В этом случае говорят, что волны длиной $\lambda_1 = \lambda$ и $\lambda_2 = \lambda - \Delta\lambda$ разрешить не удается. Разрешающая способность интерферометра характеризуется той минимальной разницей $\Delta\lambda = \lambda_1 - \lambda_2$ в длинах волн, при которой возможно их разрешение.

На рис. 127 изображено распределение интенсивностей от волн с близкими длинами для близких значений разности хода:

$$\begin{aligned}\delta_1 &= (2\pi/\lambda_1)2d\cos\theta_1 = (2\pi/\lambda)2d\cos\theta_1, \\ \delta_2 &= (2\pi/\lambda_2)2d\cos\theta_2 = [2\pi/(\lambda - \Delta\lambda)]2d\cos\theta_2.\end{aligned}\quad (28.13)$$

Эти графики строятся с помощью формулы (28.12). Пунктирной кривой изображена суммарная интенсивность от интерференционных картин двух длин волн. Значения δ_1 и δ_2 в (28.13) определяются из условия, что они соответствуют максимумам интенсивности соответствующих интерференционных колец:

$$\sin(\delta_1/2) = 0, \quad \sin(\delta_2/2) = 0. \quad (28.14)$$

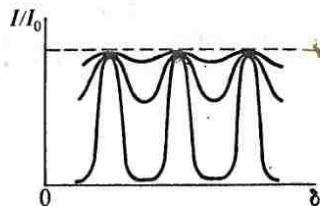
Мы пользуемся представлением интерференционной картины в виде функции от δ . Это очень удобно, так как позволяет одновременно проанализировать зависимость интерференционной картины от всех параметров, которыми определяется δ , т. е. от угла θ , расстояния между пластинами d и длины волны λ , а также от их скомбинированного изменения. .

Установим, при каких условиях распределение интенсивностей от двух волн, изображенное на рис. 127 пунктирной линией, можно считать разрешенным на сумму двух распределений интенсивности от каждой из волн. В теории интерферометра Фабри — Перо принимается, что условием разрешения является пересечение кривых, описывающих распределение интенсивностей от каждой из волн на половине их максимального значения, т. е. в точке δ_0 , где $I/I_0 = 1/2$ для обеих кривых. Отсюда в соответствии с (28.12) условие разрешения можно записать в виде

$$\begin{aligned}A \sin^2[(\delta_1 + \varepsilon_1)/2] &= 1, \\ A \sin^2[(\delta_2 - \varepsilon_2)/2] &= 1,\end{aligned}\quad (28.15)$$

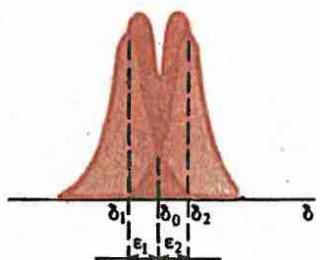
где $\varepsilon_1 = \delta_0 - \delta_1$, $\varepsilon_2 = \delta_2 - \delta_0$. Примем во внимание, что ε_1 и ε_2 имеют очень малые значения, поскольку длины волн близки друг к другу и расстояние между максимумами, т. е. $\delta_2 - \delta_1$, очень мало. Разложим синусы в (28.15) в ряд Тейлора по $\varepsilon_1/2$ и $\varepsilon_2/2$, ограничившись первым членом:

$$\begin{aligned}\sin[(\delta_1 + \varepsilon_1)/2] &= \sin(\delta_1/2) + (\varepsilon_1/2)\cos(\delta_1/2) + \dots, \\ \sin[(\delta_2 - \varepsilon_2)/2] &= \sin(\delta_2/2) - (\varepsilon_2/2)\cos(\delta_2/2) + \dots.\end{aligned}\quad (28.16)$$



126

Изменение интенсивности в интерферометре Фабри — Перо



127

К определению разрешающей способности оптических приборов

Подставляя эти выражения в (28.15) и принимая во внимание равенства (28.14) и следующие из них равенства

$$\cos^2(\delta_1/2) = 1, \quad \cos^2(\delta_2/2) = 1, \quad (28.17)$$

вместо (28.15) получаем

$$A(\varepsilon_1/2)^2 = 1, \quad A(\varepsilon_2/2)^2 = 1, \quad (28.18)$$

откуда

$$\varepsilon_1 = \varepsilon_2 = 2/\sqrt{A}, \quad \varepsilon_1 + \varepsilon_2 = \delta_2 - \delta_1 = 4/\sqrt{A}. \quad (28.19)$$

Углы θ_1 и θ_2 в (28.13), которые соответствуют наблюдаемым в интерферометре Фабри—Перо кольцам интерференции, достаточно малы. Так как $\cos \theta = 1 - \theta^2/2 + \dots$, то с точностью до величин второго порядка малости относительно θ_1 и θ_2 можно в (28.13) положить $\cos \theta_1 = 1$, $\cos \theta_2 = 1$, поэтому

$$\delta_2 - \delta_1 = 4\pi d \left(\frac{1}{\lambda - \Delta\lambda} - \frac{1}{\lambda} \right) = \frac{4\pi d}{\lambda} \frac{\Delta\lambda}{\lambda}, \quad (28.20)$$

где использовано приближение $(\lambda - \Delta\lambda)\lambda = \lambda^2$, т. е. с учетом множителя $\Delta\lambda$ в числителе принятые во внимание лишь члены нулевого порядка по $\Delta\lambda$ в знаменателе. С помощью (28.19) равенство (28.20) можно представить в виде

$$\frac{\Delta\lambda}{\lambda} = \frac{\lambda}{\pi d} \frac{1}{\sqrt{A}} = \frac{\lambda}{2\pi d} \frac{1 - \rho}{\sqrt{\rho}}, \quad (28.21)$$

где для A использовано выражение (28.11).

Разрешающая способность спектрального аппарата по определению равна

$$\Upsilon = \lambda/\Delta\lambda, \quad (28.22)$$

где $\Delta\lambda$ — минимальная разница в длинах волн, которые могут считаться разрешенными в соответствии с принятым критерием. Разрешающая способность интерферометра Фабри—Перо на основании (28.21) дается выражением

$$\Upsilon = \frac{2\pi d}{\lambda} \frac{\sqrt{\rho}}{1 - \rho}. \quad (28.23)$$

Факторы, ограничивающие разрешающую способность. Формально по формуле (28.23) можно получить сколь угодно большую разрешающую способность, если только выбрать d достаточно большим, а коэффициент отражения ρ — достаточно близким к единице. Однако добиться сколь угодно большой разрешающей способности нельзя из-за ряда обстоятельств.

Прежде всего пластины интерферометра не могут быть сделаны абсолютно плоскими. Как уже отмечалось, они могут быть отполированы с точностью до $1/200$ длины волны, но не с абсолютной точностью. Отклонение поверхности пластин от идеальной плоскости эквивалентно изменению d , что приводит к соответствующему изменению δ . Если изменение d составляет α ($\alpha \ll 1$), то δ изменяется на $4\pi\alpha$. Чтобы не нарушить условий разрешения, это изменение должно быть меньше $\delta_2 - \delta_1$, т. е. должно удовлетворять условию

$$4\pi\alpha < (\delta_2 - \delta_1) = 2(1 - \rho)/\sqrt{\rho}, \quad (28.24)$$

где учтены равенства (28.19) и (28.21). Отсюда следует, что на допустимые отклонения поверхности пластины от идеальной плоскости накладываются очень жесткие требования:

$$\alpha < \frac{1 - \rho}{2\pi \sqrt{\rho}}.$$

(28.25)

Напомним, что α характеризует отклонение поверхности пластины от идеальной плоскости в долях длины волны, т. е. требования действительно очень жесткие.

Второе ограничение обусловлено необходимостью добиться параллельности пластин с той же точностью, как и отклонения их поверхности от идеальной плоскости, что очевидно из анализа физического содержания предшествующего ограничения.

Расстояние между пластинами также не может быть сделано очень большим. Расстояние между пластинами, умноженное на число эффективных отражений, должно быть меньше, чем временная длина когерентности исследуемого света.

Использование больших расстояний d также связано с некоторыми практическими трудностями. Из (28.5) следует, что условие максимума интенсивности может быть представлено в виде

$$2d \cos \theta = m\lambda \quad (m = 0, 1, 2, 3, \dots). \quad (28.26)$$

Это значит, что порядок интерференции растет с уменьшением диаметра кольца. Центральное пятно ($\theta = 0$) соответствует порядку интерференции

$$m_{\max} = 2d/\lambda. \quad (28.27)$$

Первое кольцо имеет порядок интерференции $m_{\max} = 1$. Оно реализуется за счет лучей, угол падения θ_1 которых удовлетворяет условию

$$2d \cos \theta_1 = (m_{\max} - 1)\lambda! \quad (28.28)$$

Этот угол мал и поэтому, полагая в (28.28)

$$\cos \theta_1 = 1 - \theta_1^2/2 \quad (28.29)$$

и учитывая (28.27), находим

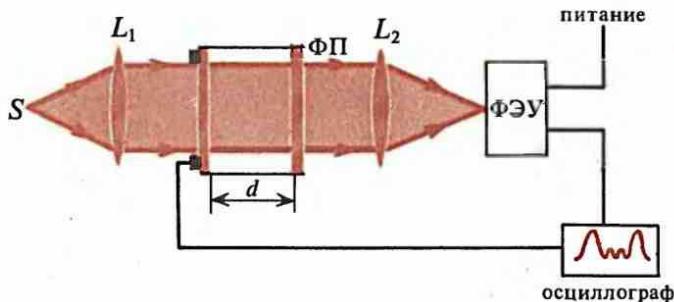
$$\theta_1 = \sqrt{\lambda/d}. \quad (28.30)$$

Это означает, что с увеличением d радиус первого кольца и расстояние между кольцами уменьшаются, что ухудшает условия анализа интерференционной картины, особенно при использовании в технике фотометрирования интерференционной картины. Анализ значительно облегчается при использовании сканирующего интерферометра Фабри—Перо (рис. 128) выбором подходящего интервала изменений d .

Дисперсионная область. При увеличении разности длин волн $\Delta\lambda$ интерференционные полосы от волн с различной длиной волны разделяются (рис. 129, а). При дальнейшем увеличении $\Delta\lambda$ разделение увеличивается (рис. 129, б) и при достаточно больших $\Delta\lambda$ наступает перекрытие интерференционных полос соседних порядков (рис. 129, в). При этом интерпретация интерференционной картины становится затруднительной. Разность длин волн $\Delta\lambda$, при которой наступает перекрытие полос соседних порядков интерференции, называется дисперсионной обла-

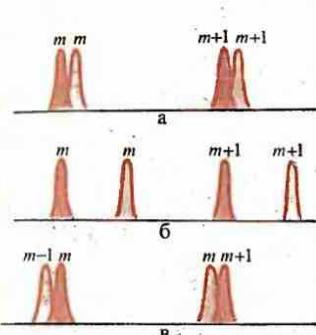
• **Отражение и преломление зависят не только от свойств поверхности раздела и характеристик падающей на поверхность раздела волны, но и от характеристик других волн, которые падают на поверхность раздела.**

Разрешающая способность интерферометра Фабри—Перо возрастает при увеличении расстояния между пластинами, а дисперсионная область при этом уменьшается.



128

Схема сканирующего интерферометра Фабри — Перо



129

К определению дисперсионной области

стью. Из (28.26) видно, что условие совпадения максимумов соседних порядков записывается в виде

$$2d \cos \theta = (m+1)\lambda, \quad 2d \cos \theta = m(\lambda + \Delta\lambda). \quad (28.31)$$

Отсюда следует, что

$$(m+1)\lambda = m(\lambda + \Delta\lambda) \quad (28.32)$$

или

$$\lambda = m\Delta\lambda \quad (28.33)$$

и дисперсионная область равна

$$G = \Delta\lambda = \lambda/m. \quad (28.34)$$

Для интерферометра Фабри — Перо $m \approx 2d/\lambda$ при $\theta \approx 0$ и поэтому дисперсионная область

$$G = \lambda^2/(2d) \quad (28.35)$$

достаточно мала, поскольку используемые порядки интерференции велики. Требования повышения разрешающей способности и расширения дисперсионной области противоречат друг другу: для повышения разрешающей способности приходится переходить к более высоким порядкам интерференции (увеличивая d) и тем самым уменьшать дисперсионную область.

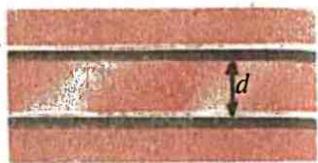
Сканирующий интерферометр Фабри — Перо. Измерение распределения интенсивности в интерференционной картине чрезвычайно трудоемко. Надо фотографировать картину и затем профотометрировать пленку, чтобы по распределению почертения определить распределение интенсивности. Во время фотографической выдержки температура несколько колеблется, что влечет за собой движение колец. Интерференционная картина размывается и разрешение ухудшается. Интерференционная картина чувствительна к изменению температуры порядка 0,1 К. Поэтому необходимо в течение всего эксперимента обеспечить достаточную стабильность температуры.

Чтобы избежать этих трудностей, используется сканирующий интерферометр Фабри — Пере (рис. 128). Свет от точечного источника S , расположенного в фокусе линзы L_1 , после прохождения линзы образует параллельный пучок лучей ($\theta \approx 0$). После прохождения интерферометра Фабри — Пере ФП выходит пучок лучей, который линзой L_2 фокусируется на приемник фотоумножителя ФЭУ, где и осуществляется интерференция, причем на приемник ФЭУ попадает лишь центральная область картины. Интенсивность света, воспринимаемого приемником, определяется формулой (28.9), в которой d дается выражением (25.5) с $\cos \theta = 1$. Фотоумножитель вырабатывает электрический сигнал, пропорциональный воспринимаемой интенсивности, который направляется в записывающее устройство. Этим устройством может быть, например, осциллограф.

Сигнал зависит от λ и d . При изменении d изменяется длина волны, которая дает на приемнике фотоумножителя максимум интенсивности. Поэтому величина наблюдаемого сигнала при каждом значении d позволяет непосредственно сделать заключение об интенсивности волны соответствующей длины в падающем на интерферометр излучении. Одна из пластин интерферометра монтируется на кольцо из пьезоэлектрического материала. Напряжение, подаваемое на пьезоэлектрическое кольцо, подбирается так, чтобы соответствующее изменение d обеспечивало прохождение всей дисперсионной области около длины волны λ , при которой возникает максимум интенсивности в центре интерференционной картины, регистрируемой приемником фотоумножителя. Сигнал с фотоумножителя подается на осциллограф, а развертка осциллографа синхронизируется с частотой колебаний пьезоэлектрического кольца. В результате на экране осциллографа можно визуально наблюдать картину распределения интенсивности по длинам волн в некотором масштабе. Наблюдаемые величины затем пересчитываются на длины волн и определяется искомый спектр излучения.

Интерференционные фильтры. Как было отмечено, интерферометр Фабри — Пере при определенных условиях пропускает без ослабления световую волну определенной длины. Это происходит тогда, когда в формуле (28.9) $\sin(\delta/2) = 0$, причем $\delta/2 = 2\pi d/\lambda$ ($\cos \theta = 1$). Поскольку значение $4\rho/(1 - \rho)^2$ в (28.9) весьма велико, знаменатель становится очень большим уже при небольших, отличных от нуля, значениях $\sin(\delta/2)$. Следовательно, волны, длина волн которых лишь немногого отличается от λ , проходят с большим ослаблением (интерферометр Фабри — Пере действует как узкополосный фильтр). Практически наиболее простой узкополосный фильтр осуществляется следующим образом (рис. 130). Берется диэлектрическая пластина с небольшим показателем преломления такой толщины d , чтобы оптическая длина пути в ней была равна половине длины волны, т. е.

$$nd = \lambda/2.$$



130

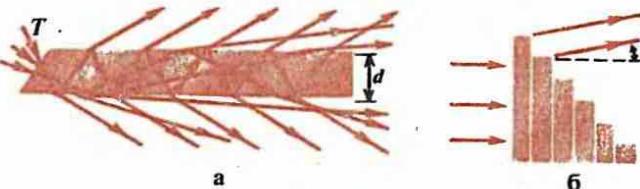
Схема интерференционного фильтра

○ В чем состоит физическая причина возникновения почти полного отражения или почти полного пропускания волны в интерферометре Фабри — Пере?

Какие факторы ограничивают разрешающую способность интерферометра Фабри — Пере?

Почему дисперсионная область интерферометра Фабри — Пере невелика? Что происходит с разрешающей способностью интерферометра Фабри — Пере при увеличении дисперсионной области?

Опишите принцип действия интерференционных фильтров.



131

Пластинка Люммера — Герке (а);
шелоп Майкельсона (б)

а

б

Поверхности пластины покрываются тонкими металлическими слоями с достаточно большой отражательной способностью, но частично пропускающие свет. В результате получается миниатюрный интерферометр Фабри — Перо, причем для выбранной длины волны $\delta/2 = \pi$, $\sin(\delta/2) = 0$. Следовательно, при достаточно высокой отражательной способности металлических слоев эта система действует как узкополосный фильтр. Типичная ширина полосы пропускания, определяемая на половине максимальной интенсивности, составляет около 20 нм, что для такой простой системы является хорошим показателем. Падение света предполагается нормальным к поверхности. Ширина полосы пропускания может уменьшаться примерно вдвое, если имеется интерференционный максимум второго порядка. Однако в этом случае наряду с волной длины λ будет пропущена и волна длины 2λ (см. также § 29).

Для того чтобы защитить от повреждения металлические слои, интерференционный фильтр с обеих сторон закрывается стенками. Существенных изменений в описанную картину функционирования фильтра это не вносит. Предохранительные стекла на рис. 130 обозначены точками.

Пластинка Люммера — Герке. Сравнительно несложным прибором для осуществления многолучевой интерференции делением амплитуды является пластинка Люммера — Герке. Она представляет собой плоскопараллельную пластину из стекла толщиной d (рис. 131, а). Впуск света внутрь пластинки производится либо черезрезанный торец T , либо через дополнительную призму, причем условия впуска подбирают такими, чтобы луч внутри пластины многократно отражался от ее поверхностей под углами, близкими к углам полного отражения. Обозначая θ угол отражения для разности хода между лучами, возникающей между последовательными отражениями от одной и той же поверхности пластины, найдем выражение

$$\Delta = 2nd \cos \theta, \quad (28.36)$$

которое получается аналогично (28.4) и выведено в § 29. Условие максимума имеет вид

$$2nd \cos \theta = m\lambda. \quad (28.37)$$

В фокальной плоскости линзы образуются интерференционные полосы, параллельные поверхности пластины. Обычно толщина d пластины бывает от 3 до 10 мм, а угол θ близок к $\pi/4$. В результате порядок m интерференции оказывается очень высоким — десятки тысяч. Теория интерференции с пластинкой Люммера — Герке совершенно аналогична теории интерферометра Фабри — Перо.

Шелоп Майкельсона (рис. 131, б). Этот прибор собирается из стеклянных пластин одинаковой толщины, сложенных в виде лестницы с одинаковыми выступами. Разность хода между лучами, выходящими из соответствующих точек соседних ступенек, образуется за счет различных оптических длин путей.

Пример 28.1. Линия $\lambda = 546,1$ нм в излучении ртути является дублетом, длины волн в котором λ_1 и λ_2 очень близки друг к другу. Считая, что $|1/\lambda_1 - 1/\lambda_2| = 0,1 \text{ см}^{-1}$, определить, какие требования надо предъявить к коэффициенту отражения пластин интерферометра Фабри — Перо и допуску на отклонение пластин от абсолютной плоскости. Расстояние между пластинами равно 1 мм.

Разрешающая способность интерферометра дается формулой (28.23). Представив ее в виде

$\frac{\Delta\lambda}{\lambda^2} = \frac{1}{2\pi d} \frac{1-\rho}{\sqrt{\rho}}$ и учитывая, что $|1/\lambda_1 - 1/\lambda_2| = \Delta\lambda/\lambda^2 = 0,1 \text{ см}^{-1}$, находим $\rho = 0,93$. Требование допустимого отклонения пластин от абсолютной плоскости дается формулой (28.25), которая приводит к оценке $\alpha < 1/90$.

§ 29 | Интерференция в тонких пленках

Рассматривается применение общих методов многолучевой интерференции к исследованию интерференции в тонких пленках. Описываются матричный метод исследования и его применение к многослойным пленкам — интерференционным фильтрам и диэлектрическим зеркалам.

Оптическая длина пути. При распространении света в среде его скорость уменьшается. Это влечет за собой уменьшение длины волны, поскольку частота не изменяется. Следовательно, волновое число $k_{\text{cp}} = 2\pi/\lambda_{\text{cp}}$ в среде связано с волновым числом k в вакууме соотношением

$$k_{\text{cp}} = 2\pi/\lambda_{\text{cp}} = 2\pi/(vT) = 2\pi n/(cT) = n2\pi/\lambda = nk, \quad (29.1)$$

где $v = c/n$ — скорость волны в среде, n — показатель преломления среды относительно вакуума. Разность фаз, возникающая за счет прохождения волной геометрического расстояния Δ_r в среде, равна

$$\delta = k_{\text{cp}} \Delta_r = nk \Delta_r = k \Delta, \quad (29.2)$$

где $\Delta = n \Delta_r$ — оптическая длина пути. Таким образом, при расчете изменения фаз в среде удобно считать длину волны и волновой вектор равными их значениям в вакууме, а в качестве длины пути брать оптическую длину пути, равную геометрической длине, умноженной на показатель преломления.

Отражение от параллельных поверхностей. Обозначим n_1 и n_2 показатели преломления среды и пленки толщиной d (рис. 132). Рассчитаем разность оптических путей лучей, отразившихся от нижней и верхней поверхностей пленки. Из рис. 132 заключаем, что

$$\Delta = (|AB| + |BC|)n_2 - |AD|n_1. \quad (29.3)$$

Учитывая, что $|AB| + |BC| = 2d/\cos\theta_{\text{пп}}$, $|AD| = 2d \tan\theta_{\text{пп}} \sin\theta_{\text{пп}}$, и принимая во внимание закон преломления $\sin\theta_{\text{пп}}/\sin\theta_{\text{нд}} = n_2/n_1$, получаем

$$\Delta = 2d[n_2/\cos\theta_{\text{пп}} - \tan\theta_{\text{пп}}(n_2/n_1) \sin\theta_{\text{пп}} n_1] = 2dn_2 \cos\theta_{\text{пп}}. \quad (29.4)$$

Эта формула совпадает с (28.4), но с заменой геометрической толщины d на «оптическую толщину» $n_2 d$.

Формула (29.4) показывает, что все лучи, падающие на пленку под одним и тем же углом, разделяются на два луча и после отражений от поверхностей пленки имеют одно и то же направление распространения, будучи параллельными друг другу. Между лучами, отразившимися от разных поверхностей пленки, возникает разность хода (29.4), и они могут интерферировать между собой. Максимум интерференции наступает при условии

$$\Delta = m\lambda. \quad (29.5)$$

Подчеркнем, что λ в этой формуле является длиной волны в вакууме, поскольку особенности распространения света в среде учтены тем, что под Δ понимается оптическая длина.

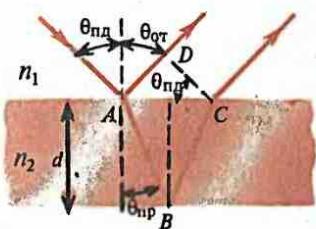
Параллельные лучи не пересекаются на конечном расстоянии. Следовательно, интерференционные полосы не могут возникнуть на конечном расстоянии. Говорят, что они локализованы на бесконечности. Наблюдать интерференцию можно двумя способами: либо фокусировать лучи с помощью линзы, в результате чего в фокальной плоскости линзы возникнут полосы интерференции, либо увидеть их непосредственно глазами, аккомодированными на бесконечность. В этом случае полосы образуются в фокусе зрачка на сетчатке глаза и создают соответствующее зрительное впечатление.

Утверждение о том, что интерференционные полосы локализованы на бесконечности, имеет не только математический, но и физический смысл. Интенсивность светового поля в некоторой области пространства может быть измерена с помощью фотоэлемента. Перемещая фотоэлемент в пространстве, можно исследовать, как изменяется интенсивность светового поля. Если фотоэлемент перемещать в фокальной плоскости линзы, где образованы полосы интерференции, то он показывает изменение интенсивности светового поля в соответствии с наблюдавшейся интерференционной картиной. Если же перемещать фотоэлемент в пространстве, когда полосы интерференции локализованы на бесконечности, то никакого интерференционного изменения интенсивности светового поля не будет отмечено.

Линии равного наклона. Если на пластинку падает пучок непараллельных лучей, то в отраженном пучке будут присутствовать лучи различных направлений распространения с соответствующими различными углами преломления. Те из них, для которых удовлетворяется условие (29.5), дают при интерференции максимум интенсивности. Следовательно, если с помощью линзы в ее фокальной плоскости образовать интерференционную картину, то интерференционная линия определенной интенсивности соответствует определенному углу $\theta_{\text{пр}}$ в (29.4) или, что то же самое, определенному углу падения или отражения. Другими словами, эта линия соответствует определенному углу наклона образующих ее лучей к поверхности пластины. Поэтому такие интерференционные линии называются линиями равного наклона. Они локализованы на бесконечности.

Если пучок света, падающий на пленку, аксиально-симметричен, то линии равного наклона являются окружностями. На рис. 133 показана схема реализации линий интерференции равного наклона на экране с помощью линзы.

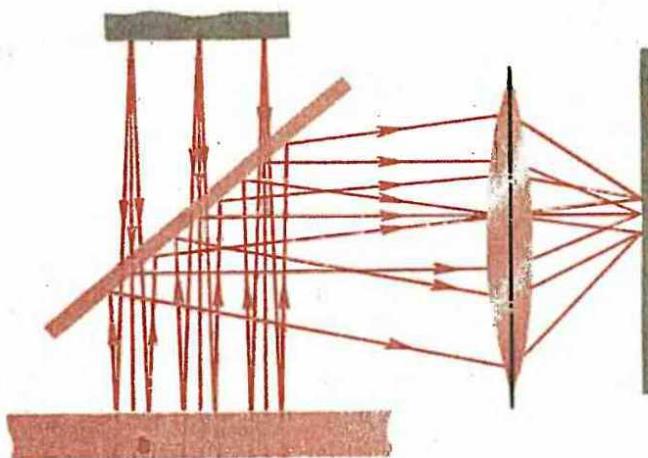
Линии равного наклона, которые в случае аксиальной симметрии падающего на пластинку излучения являются окружностями, можно наблюдать глазом (рис. 134). Если при смещении



1.32

К расчету оптической длины пути в тонких пленках

$$\Delta = (|AB| + |BC|)n_2 - |AD|n_1$$



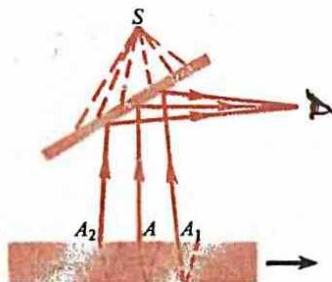
133

Схема образования интерференционных линий равного наклона на экране

пленки параллельно самой себе ее толщина изменяется на видимом участке, то радиус интерференционных колец также изменяется, причем изменение радиуса колец легко заметить при весьма незначительных изменениях толщины пленки. Это дает весьма эффективный способ контроля толщины пленки (пластин), широко используемого в производстве. Точность такого контроля очень большая, поскольку локальное изменение толщины пластины на долю длины волны приводит к заметному невооруженным глазом изменению форме линии.

Роль размера источника. Различные точки источника излучают некогерентно. Однако интерференционные картины, образуемые любой точкой источника при отражении под одинаковым углом, идентичны друг другу и не зависят от точки поверхности пленки, в которой произошло отражение. Интерференционные полосы от излучения различных точек источника накладываются друг на друга без смазывания картины интерференции. Следовательно, конечность размеров источника не смазывает картину интерференции линий равного наклона и не является ограничивающим интерференцию фактором.

Роль толщины пленки и монохроматичности излучения. При увеличении толщины пленки видимость интерференционной картины ухудшается ввиду конечного значения временной когерентности. Причина этого аналогична той, которая была рассмотрена при анализе ухудшения видимости в интерферометре Майкельсона, (см. § 26). Увеличение ширины линий излучения (немонохроматичность) ухудшает видимость интерференционной картины.



134

Наблюдение глазом линий одинакового наклона. Способ контроля толщины пластины

Для наблюдения интерференции в тонких пленках необходимо, чтобы их толщина не превосходила определенного значения. Для наблюдения интерференции в белом свете перекрытие полос от различных длин волн не должно смазывать цветовую гамму белого света. Известно, что глаз различает цвета, соответствующие разнице длии волн примерно на $\Delta\lambda \approx 10^{-2}$ мкм. Поэтому для наблюдения интерференции надо обеспечить дисперсионную область такой же величины. Из (28.34) следует, что порядок интерференции должен быть не больше

$$m_{\max} = \lambda/\Delta\lambda = 0,5 \cdot 10^{-6}/10^{-8} \approx 50, \quad (29.6)$$

где для примера $\lambda = 0,5$ мкм. Из условия (29.5) находим

$$d \approx m\lambda/(2n_2) = [50 \cdot 0,5 \cdot 10^{-6}/(2 \cdot 1,5)] \text{ м} = 8 \cdot 10^{-6} \text{ м} = 8 \text{ мкм}, \quad (29.7)$$

где $\cos\theta_{\text{пр}} = 1$ и $n = 1,5$. Этот расчет можно, конечно, провести, основываясь на времени когерентности. Из соотношения $\lambda = c/v$ следует, что $\Delta\lambda/\lambda = -\Delta v/v$. Ширина спектра Δv связана с временем когерентности τ соотношением $\Delta v = \tau c$. Скорость распространения света в пластине равна c/n_2 , а проходимое за время τ расстояние должно быть не меньше чем $2d$. Поэтому для определения максимальной толщины пластины получаем уравнение $2d = \tau c/n_2 = \tau c/(\Delta\lambda v)$, из которого следует $d = \lambda^2/(2\Delta\lambda n_2) = m\lambda/(2n_2)$, что совпадает с (29.7). Таким образом, для наблюдения интерференции в белом свете толщина пленок должна быть достаточно малой. Поэтому речь идет об интерференции в тонких пленках, а не просто в плоскопараллельных пластинах.

При наблюдении интерференции от монохроматического источника допустимая толщина пленки увеличивается. Она зависит от ширины линии излучения или, что одно и то же, от времени и длины когерентности. Например, ширина зеленой линии излучения ртутной лампы в длинах волн $\Delta\lambda \approx 0,01$ нм ($\lambda = 502,564$ нм). Отсюда для максимально допустимой толщины пластины по формулам (29.6) и (29.7) находим $d = 8$ мм.

Длина когерентности лазерного излучения может достигать очень больших значений (многие километры). Поэтому с его помощью можно наблюдать картину интерференции в очень толстых пластинах. Ограничения возможностей наблюдения обусловливаются не степенью когерентности, а неоднородностями материала, качеством поверхностей пластины и другими аналогичными факторами.

Линии равной толщины. Если толщина пластины переменна, то от различных участков ее поверхности пары лучей с одинаковой разностью фаз распространяются в разных направлениях и, следовательно, картина интерференции лучей равного наклона не возникает. Однако появляется другая интерференционная картина, локализованная на поверхности пластины. Образующие ее интерференционные полосы называются линиями равной толщины ввиду того, что интенсивность полос одинакова в тех областях, в которых одинакова толщина пластины.

Фокусируем линзу так, чтобы на экране получить изображение небольшого участка поверхности (рис. 135). Освещенность участка зависит от толщины пластины на этом участке, поскольку у ее поверхности интерферируют лучи с почти одинаковой разностью хода при приближительно вертикальном падении. Различие в разности хода лучей, приходящих в данную точку поверхности, обусловливается лишь отличием их углов преломления ($\theta_{\text{пр}}$). Однако оно невелико, поскольку линейные размеры наблюдаемой области малы. Следовательно, линзой на экран проецируется интерференционная картина, которая возникает у поверхности пластины. Если по поверхности пластины перемещать фотозлемент, то он отметит изменение интенсивности светового поля, соответствующее изменению интенсивности в наблюдаемой интерференционной картине. Светлые и темные полосы на поверхности пластины можно видеть также и непосредственно глазом, аккомодированным на поверхность пластины. Линия одинаковой интенсивности интерференции совпадает с линией постоянной толщины пластины. Область одинаковой интенсивности соответствует области, в которой толщина пластины постоянна.

Контрастность интерференционной картины зависит от порядка интерференции, т. е. от толщины пластины. При наблюдении в белом свете видны цветные полосы. Ограничения на толщину пластины такие же, как в только что рассмотренном случае линий равного наклона.

Кольца Ньютона. Полосы равной толщины образуются за счет воздушной прослойки между плоской поверхностью стекла, на которую выпуклой стороной положена плосковыпуклая линза (рис. 136), и поверхностью линзы. Линии постоянной толщины d воздушной прослойки являются окружностями. Поэтому и интерференционные линии равной толщины — окружности. Они называются **кольцами Ньютона**.

Из рис. 136 видно, что $R^2 - (R - d)^2 = r^2$, $2Rd - d^2 = r^2$. Переписав это соотношение в виде $2Rd/r^2 - d^2/r^2 = 1$ и учитывая, что $d^2/r^2 \ll 1$, находим

$$d = r^2/(2R). \quad (29.8)$$

При отражении на границе с более плотной средой фаза напряженности электрического поля волны изменяется на π [см. (16.33а), рис. 65]. Следовательно, условие образования темного кольца имеет вид

$$2r_m^2/(2R) + \lambda/2 = m\lambda + \lambda/2.$$

Отсюда радиус m -го темного кольца

$$r_m = \sqrt{R\lambda m}. \quad (29.9)$$

В области соприкосновения линзы с поверхностью стекла наблюдается темное пятно. Оно окружено последовательностью светлых и темных колец. В белом свете возникают цветные кольца. Условия наблюдения в принципиальном отношении такие же, как и в случае тонких пленок, т. е. с увеличением d картина смазывается.

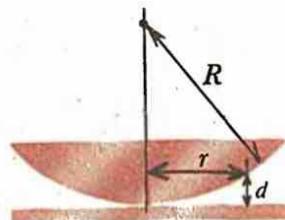
В проходящем свете возникает дополнительная интерференционная картина. В частности, при монохроматическом освещении на месте светлых колец образуются темные, и наоборот.

Учет многократных отражений. До сих пор при анализе интерференции в тонких пленках нами рассматривалась лишь двухлучевая интерференция, возникающая в результате одного отражения от поверхностей пленки (см. рис. 132). Такое приближение дает хорошие результаты и является вполне оправданным, если коэффициент отражения на поверхностях пленки мал. При не очень малых коэффициентах необходимо учитывать многократные отражения и рассматривать интерференцию в тонких пленках и пластинах не как двухлучевую, а как многолучевую.



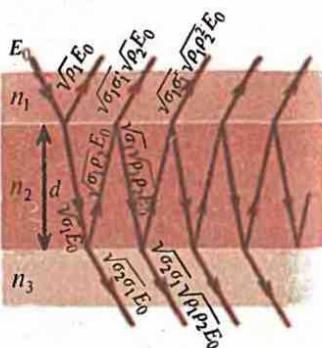
135

Наблюдение линий одинаковой толщины



136

К расчету радиусов колец Ньютона



137

К расчету коэффициентов отражения и прохождения для слоев

На рис. 137 изображена картина многократных отражений в плоскопараллельной пластине толщиной d с показателем преломления n_2 , граничащей со средами, показатели преломления которых n_1 и n_3 . Коэффициенты отражения от верхней и нижней границ пластины обозначены соответственно ρ_1 и ρ_2 . По формуле (18.5) с их помощью можно связать амплитуды падающей и отраженной волн. Амплитуды падающей и преломленной волн связаны посредством коэффициента пропускания τ формулой (18.9). Этую формулу удобно представить так:

$$|E_{\text{пп}}^0| = \sqrt{\tau v_2 \cos \theta_{\text{пп}} / (v_1 \cos \theta_{\text{пп}})} |E_{\text{пп}}^0| = \sqrt{\sigma} |E_{\text{пп}}^0|, \quad (29.10)$$

где

$$\sigma = \tau v_2 \cos \theta_{\text{пп}} / (v_1 \cos \theta_{\text{пп}}). \quad (29.11)$$

Коэффициент отражения ρ зависит от угла падения. При одном и том же угле падения коэффициент отражения, вообще говоря, различен для падения волны на границу из различных сред. Как видно из (18.6), при определенном угле падения $\theta_{\text{пп}}$ коэффициент отражения при падении волны на границу раздела из первой среды равен коэффициенту отражения при падении волны на границу раздела из второй среды лишь при условии, что угол падения во втором случае равен углу преломления в первом. Однако $\sqrt{\rho}$, как видно из (16.30а) и (16.42а), отличаются знаком для различных направлений движения волны к границе раздела. Другими словами, $\sqrt{\rho}$ необходимо брать равным правым частям равенств (16.30а) и (16.42а) при расчете преломления для соответствующих составляющих напряженности E поля волны. Что касается σ [см. (29.11)], то он при указанных условиях различен для преломления на одной и той же границе при движении волны с разных сторон. Поэтому коэффициент σ для преломления на верхней поверхности пластины при движении луча из среды в пластину обозначим σ_1 , при движении из пластины в среду — σ'_1 , для преломления на нижней поверхности пластины при движении луча из пластины в среду — σ_2 .

Амплитуды прошедших и отраженных от поверхностей пластины лучей находятся с помощью $\sqrt{\rho_1}$, $\sqrt{\rho_2}$, $\sqrt{\sigma_1}$, $\sqrt{\sigma'_1}$, $\sqrt{\sigma_2}$ по тем правилам, которые только что были объяснены. Разность фаз между соседними лучами, выходящими из пластины на основании (29.4), равна

$$\delta = (4\pi/\lambda) n_2 d \cos \theta_{\text{пп}}. \quad (29.12)$$

Проще сначала рассчитать суммарную комплексную амплитуду напряженности прошедшей через пластину волны. Она равна (рис. 137)

$$\begin{aligned} E &= \sqrt{\sigma_1 \sigma_2} E_0 + \sqrt{\sigma_1 \sigma_2} \sqrt{\rho_1 \rho_2} E_0 e^{i\delta} + \sqrt{\sigma_1 \sigma_2} (\sqrt{\rho_1 \rho_2})^2 E_0 e^{2i\delta} + \dots = \\ &= \sqrt{\sigma_1 \sigma_2} E_0 \sum_{m=0}^{\infty} (\sqrt{\rho_1 \rho_2})^m e^{mi\delta} = \sqrt{\sigma_1 \sigma_2} E_0 / (1 - \sqrt{\rho_1 \rho_2} e^{i\delta}), \end{aligned} \quad (29.13)$$

где сумма геометрической прогрессии распространена до бесконечности, поскольку для плоской волны, не ограниченной в пространстве, число отражений, вносящих вклад в интерференцию, бесконечно. В пространственно ограниченных световых пучках число отражений, которые дают вклад в интерференцию, наблюдаемую с помощью апертур конечного размера, ограничено. Однако и в этом случае обычно можно распространить суммирование по бесконечному числу отражений, поскольку главный вклад в интерференцию дает первое небольшое число членов, а вклад остальных членов незначителен. Взяв квадраты модулей от обеих частей (29.13), находим

$$|E|^2 = \frac{\sigma_1 \sigma_2 |E_0|^2}{1 + \rho_1 \rho_2 - 2 \sqrt{\rho_1 \rho_2} \cos \delta}. \quad (29.14)$$

Для того чтобы правильно применить (29.13) и полученные из нее последующие формулы, необходимо подчеркнуть смысл величин $\sqrt{\rho_1}$ и $\sqrt{\rho_2}$. Из вывода формулы ясно, что $\sqrt{\rho_1}$ характеризует отражение волн, движущихся в пластине и отраженных от границы пластины с первой средой обратно в пластину, а $\sqrt{\rho_2}$ относится к отражению волны, движущейся в пластине, от границы с третьей средой. Другими словами, $\sqrt{\rho_1}$ и $\sqrt{\rho_2}$ описывают отражение волн в пластине от ее поверхностей.

Плотность потока электромагнитной энергии в среде равна

$$S = EH = E^2 / (\mu_0 v), \quad (29.15)$$

где $\mu = \mu_0$ (среда немагнитна), v — скорость распространения волны в среде. Нормальные компоненты потока энергии волн из первой среды в пластину и из пластины в третью среду

$$S_{1n} \sim |E_0|^2 \cos \theta_{\text{пад}} / v_1, \quad (29.16)$$

$$S_{3n} \sim |E|^2 \cos \theta_{\text{пр3}} / v_3, \quad (29.17)$$

где $\theta_{\text{пад}}$ и $\theta_{\text{пр3}}$ — соответственно углы падения из первой среды на поверхность пластины и преломления из пластины в третью среду.

Обозначим T коэффициент пропускания пластины в целом:

$$T = S_{3n} / S_{1n} = |E|^2 v_1 \cos \theta_{\text{пр3}} / (|E_0|^2 v_3 \cos \theta_{\text{пад}}). \quad (29.18)$$

Подставляя в (29.18) $|E|^2 / |E_0|^2$ из (29.14), находим

$$T = \frac{\sigma_1 \sigma_2 v_1 \cos \theta_{\text{пр3}} / (v_3 \cos \theta_{\text{пад}})}{1 + \rho_1 \rho_2 - 2 \sqrt{\rho_1 \rho_2} \cos \delta} \quad (29.19)$$

Запишем σ_1 и σ_2 в соответствии с их определением [см. (29.11)]:

$$\sigma_1 = \tau_1 v_2 \cos \theta_{\text{пад}} / (v_1 \cos \theta_{\text{пр2}}); \quad \sigma_2 = \tau_2 v_3 \cos \theta_{\text{пад2}} / (v_2 \cos \theta_{\text{пр3}}). \quad (29.20)$$

Принимая во внимание, что $\cos \theta_{\text{пр2}} = \cos \theta_{\text{пад2}}$, получаем

$$\sigma_1 \sigma_2 v_1 \cos \theta_{\text{пр3}} / (v_3 \cos \theta_{\text{пад}}) = \tau_1 \tau_2. \quad (29.21)$$

Тогда (29.19) принимает вид

$$T = \tau_1 \tau_2 / (1 + \sqrt{\rho_1 \rho_2} - 2 \sqrt{\rho_1 \rho_2} \cos \delta). \quad (29.22)$$

Для каждой из поверхностей пластин соблюдается закон сохранения энергии:

$$\tau_1 = 1 - \rho_1, \quad \tau_2 = 1 - \rho_2. \quad (29.23)$$

При отсутствии поглощения коэффициент отражения всей пластины

$$R = 1 - T = (\rho_1 + \rho_2 - 2 \sqrt{\rho_1 \rho_2} \cos \delta) / (1 + \rho_1 \rho_2 - 2 \sqrt{\rho_1 \rho_2} \cos \delta), \quad (29.24)$$

где учтены равенства (29.23).

- Линии равного наклона локализуются на бесконечности, а линии равной толщины — на поверхности пластины.
Конечность разнеров источника не смазывает картины интерференции в случае линий равного наклона и не является ограничивающим интерференцию фактором.
Допустимая толщина пленок для наблюдения интерференции в белом свете лимитируется способностью глаза различать близкие цвета.

Слой с нулевой отражательной способностью. При $R=0$ отражение от слоя отсутствует. Для этого необходимо выполнение условия

$$\rho_1 + \rho_2 - 2\sqrt{\rho_1 \rho_2} \cos \delta = 0. \quad (29.25)$$

Рассмотрим случай нормального падения и выберем слой с оптической толщиной в четверть длины волны, т. е. полагаем $n_2 d = \lambda/4$, $\delta = \pi$, $\cos \delta = -1$. Формула (29.24) принимает вид $R = (\sqrt{\rho_1} + \sqrt{\rho_2})^2 / (1 + \sqrt{\rho_1} \sqrt{\rho_2})^2$.

Коэффициенты $\sqrt{\rho_1}$ и $\sqrt{\rho_2}$ описывают отражения волны, движущейся в пластине (среда 2), от границы со средами 1 и 3. Следовательно, на основании (16.33а) они равны

$$\sqrt{\rho_1} = (n_2 - n_1) / (n_2 + n_1), \quad \sqrt{\rho_2} = (n_2 - n_3) / (n_2 + n_3). \quad (29.27)$$

Подставляя (29.27) в (29.26), находим

$$R = [(n_2^2 - n_1 n_3) / (n_2^2 + n_1 n_3)]^2. \quad (29.28)$$

Коэффициент отражения равен нулю при выполнении условия

$$n_2 = \sqrt{n_1 n_3}. \quad (29.29)$$

Например, баритовый флинт БФ-1 имеет для длины волны 589,3 нм показатель преломления $n_3 = 1,53$. Для того чтобы ликвидировать отражение от поверхности этого стекла в воздухе ($n_1 = 1$), необходимо на его поверхность наложить слой вещества с показателем преломления $n_2 = \sqrt{1,53} \approx 1,23$ с оптической толщиной в четверть длины волны. Показатели преломления большинства твердых веществ заключены между 1,5 и 2,2, и подбор вещества со столь низким показателем преломления является не простой задачей. Достаточно хорошо этому требование удовлетворяют виллиолент NaF , имеющий $n=1,33$, и криолит Na_3AlF_6 примерно с тем же показателем преломления. Коэффициент отражения при этом становится равным 0,008, что значительно меньше показателя отражения от поверхности стекла, равного приблизительно 0,04. Коэффициент отражения для других волн несколько увеличивается, но остается достаточно малым практически для всего светового диапазона.

Описанный способ уменьшения показателя отражения от стеклянных поверхностей называется просветлением оптики.

Слой с высокой отражательной способностью. Физическая причина малой отражательной способности слоя в четверть длины волны с показателем преломления меньше показателя преломления материала, на который этот слой наложен, состоит в следующем. При отражении от обеих поверхностей слоя волна изменяет фазу на π и, следовательно, разность фаз между отраженными волнами образуется исключительно за счет двойного прохода пластиинки в четверть длины волны, т. е. составляет π . Благодаря этому отраженные волны ослабляют друг друга. Чтобы увеличить коэффициент отражения, необходимо обеспечить усиление волн, отраженных от различных поверхностей слоя. Это можно сделать слоем с оптической толщиной в четверть длины волны, если на одной из поверхностей происходит изменение фазы E на π , а на другой отражение происходит без изменения фазы. Для этого показатель преломления слоя должен быть больше или меньше показателей преломления соседних сред (рис. 138). В волне, представляющей лучом 1, при отражении фаза E изменяется на π , а в волне, представляющей лучом 2, отражение происходит без изменения фазы. Следовательно, полная разность фаз между отраженными лучами равна 2π и они усиливают друг друга. Коэффициент отражения увеличивается. Например, если на поверхность стекла ($n_3 = 1,5$) наложить слой SiO ($n_2 = 2$) оптической толщиной в четверть длины волны, то по формуле (29.28) получаем $R=0,2$. Это значительно больше, чем коэффициент отражения от поверхности стекла, равный 0,04.

Формула (29.24) упрощается для случаев, когда с обеих сторон пластины находится одна и та же среда, т. е. когда $n_1 = n_3 = n$. Тогда, очевидно, $\rho_1 = \rho_2 = \rho$ и формула принимает вид

$$R = 2\rho(1 - \cos \delta)/(1 + \rho^2 - 2\rho \cos \delta). \quad (29.30)$$

Если $\cos \delta = 1$, то отражение отсутствует. Это наблюдается при условии $\delta = 2\pi m (m = 1, 2, \dots)$, которое для нормального падения лучей принимает вид $n_2 d = m\lambda/2$. Максимальный коэффициент отражения получается, если $\cos \delta = -1$. Это наблюдается при условии $\delta = (2m+1)\pi (m = 0, 1, 2, \dots)$, которое для нормального падения лучей принимает вид $n_2 d = (2m+1)\lambda/4$. Максимальный коэффициент отражения равен

$$R_{\max} = 4\rho/(1 + \rho)^2. \quad (29.31)$$

Например, коэффициент отражения ρ от поверхности стекла ($n = 1,5$) в воздухе ($n_1 = n_2 = 1$) равен 0,04, а максимальный коэффициент отражения от стеклянной пластины в воздухе по формуле (29.31) оказывается равным $R_{\max} = 0,15$, т. е. примерно в четыре раза больше, чем при отражении от одной поверхности.

Однако все сказанное справедливо лишь для достаточно малых толщин пластины, когда длина пути света в пластине меньше длины когерентности. Например, при работе с изолированными спектральными линиями от обычных источников излучения, когда имеется как ударное, так и доплеровское уширение линии, толщина пластинок обычно должна быть меньше миллиметра. Если используется излучение лазера, то толщина пластинок может составлять сантиметры в зависимости от степени когерентности лазерного излучения.

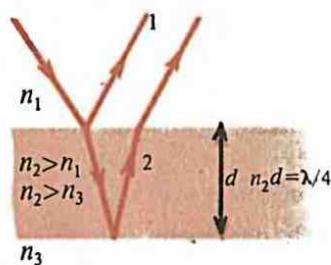
Матричный метод расчета многослойных пленок. Нахождение коэффициента отражения для пленки, состоящей из нескольких слоев, является очень трудоемкой задачей. Наиболее просто она решается матричным методом, который здесь будет описан без вывода.

Обозначим величины, относящиеся к среде, из которой свет попадает на пленку, с индексом 0; относящиеся к отдельным слоям многослойной пластины — с индексами 1, 2, 3, ..., m ; относящиеся к среде, в которую свет выходит из пленки, — с индексом «кон» (рис. 139). Каждому слою сопоставляется двухрядная квадратная матрица

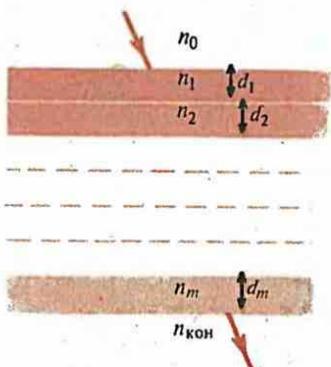
$$M_j = \begin{pmatrix} \cos \delta_j & i \sin \delta_j / n_j \\ i n_j \sin \delta_j & \cos \delta_j \end{pmatrix}, \quad (29.32)$$

в которой

$$\delta_j = (2\pi/\lambda) n_j d_j \cos \theta_{\text{пр}}, \quad (29.33)$$



Условие возникновения большой отражательной способности слоя



Обозначения величин, используемых в матричном методе расчета многослойных пленок

○ Каким образом создается слой с нулевой отражательной способностью?

Почему слой с оптической толщиной в четверть длины волны в одних случаях является слоем с нулевой отражательной способностью, а в других — слоем с очень высокой отражательной способностью?

Как устроены диэлектрические зеркала с очень высоким коэффициентом отражения?

где n_j , d_j , $\cos \theta_{np}$ — соответственно показатель преломления, толщина и угол преломления, относящиеся к j -му слою пленки. Необходимо отметить, что δ_j в (29.33) в два раза меньше δ в (29.12).

Среда, из которой свет падает на пленку, описывается двухрядной матрицей

$$\begin{pmatrix} n_0 & -1 \\ n_0 & 1 \end{pmatrix}, \quad (29.34)$$

а среда, в которую свет выходит из пленки, описывается матрицей с одним столбцом

$$\begin{pmatrix} 1 \\ n_{\text{кон}} \end{pmatrix}. \quad (29.35)$$

Коэффициент отражения вычисляется по формуле

$$R = |a/b|^2, \quad (29.36)$$

в которой

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} n_0 & -1 \\ n_0 & 1 \end{pmatrix} \begin{pmatrix} \cos \delta_1 & \frac{i \sin \delta_1}{n_1} \\ i n_1 \sin \delta_1 & \cos \delta_1 \end{pmatrix} \dots \begin{pmatrix} \cos \delta_m & \frac{i \sin \delta_m}{n_m} \\ i n_m \sin \delta_m & \cos \delta_m \end{pmatrix} \begin{pmatrix} 1 \\ n_{\text{кон}} \end{pmatrix}. \quad (29.37)$$

Многослойные диэлектрические зеркала. Для многих оптических применений необходимы отражающие поверхности с максимально большим коэффициентом отражения, например 0,99 и даже больше. Кроме того, часто необходимо, чтобы не отраженный этой поверхностью свет возможно меньше поглощался. Последнему условию не удовлетворяют металлические отражающие поверхности, поскольку они сильно поглощают не отраженную ими часть света. Этим требованиям удовлетворяют многослойные диэлектрические зеркала.

Элементом диэлектрического зеркала является пара слоев, каждый из которых имеет оптическую толщину в четверть длины волны. Один из слоев делается из материала с большим показателем преломления n_b , а другой — с малым n_m . Поскольку, по условию, $\delta_b = \delta_m = \pi/2$, матрица, описывающая прохождение света через пару слоев, имеет вид

$$\begin{pmatrix} 0 & -i/n_b \\ -i n_b & 0 \end{pmatrix} \begin{pmatrix} 0 & -i/n_m \\ -i n_m & 0 \end{pmatrix} = \begin{pmatrix} -n_m/n_b & 0 \\ 0 & -n_b/n_m \end{pmatrix}. \quad (29.38)$$

Матрица, описывающая последовательность N таких пар, дается выражением

$$\begin{pmatrix} -n_m/n_b & 0 \\ 0 & -n_b/n_m \end{pmatrix} \dots \begin{pmatrix} -n_m/n_b & 0 \\ 0 & -n_b/n_m \end{pmatrix} = \begin{pmatrix} (-n_m/n_b)^N & 0 \\ 0 & (-n_b/n_m)^N \end{pmatrix}, \quad (29.39)$$

Поэтому в соответствии с (29.37) имеем

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} (-n_m/n_b)^N & 0 \\ 0 & (-n_b/n_m)^N \end{pmatrix} \begin{pmatrix} 1 \\ n_{\text{кон}} \end{pmatrix}. \quad (29.40)$$

Отсюда по формуле (29.36) получаем

$$R = \left[\frac{n_{\text{кон}} - (n_m/n_b)^{2N}}{n_{\text{кон}} + (n_m/n_b)^{2N}} \right]^2 \quad (29.41)$$

Так как $n_m/n_b < 1$, то при увеличении N коэффициент отражения R приближается к единице. В принципе при достаточно большом значении N он может быть сделан как угодно близким к единице. Факторами, ограничивающими число пар, являются поглощение и рассеяние света при прохождении через слои. В настоящее время получены значения R , большие 0,995.

Полупрозрачные материалы. Подбором вещества и толщины слоев пластины можно добиться, чтобы она обладала достаточно большим коэффициентом отражения и пропускания (например, $R = 0,5$; $T = 0,5$). Если из такого полупрозрачного материала изготовить, например, стекла легковой автомашины, то со стороны улицы днем нельзя будет увидеть, что происходит в кабине, поскольку отраженный от стекол световой поток значительно сильнее светового потока, идущего из кабины машины. Из кабины же хорошо видно все, что происходит на улице.

§ 30 | Частичная когерентность и частичная поляризация

Излагаются общая теория частичной когерентности и частичной поляризации и их связь с общей теорией случайных процессов.

Частичная когерентность. При анализе двухлучевой интерференции, осуществляющейся делением амплитуды (см. § 26), было выяснено, что видимость интерференционной картины для строго монохроматических волн равна единице. Для квазимонохроматического излучения видимость при увеличении разности хода лучей ухудшается и при достаточно большой разности хода, превосходящей «временную длину когерентности», обращается в нуль. **При видимости, заключенной между 0 и 1, говорят, что волны частично когерентны.**

Аналогично обстоит дело в двухлучевой интерференции, осуществляющейся делением волнового фронта (см. § 27). Если источник, освещающий щели, монохроматический и точечный, то видимость интерференционной картины равна единице и можно говорить, что вторичные волны, исходящие из щелей, когерентны. Однако если источник точечный, но немонохроматический, или монохроматический, но протяженный, то видимость интерференционной картины ухудшается. В обоих случаях интерферирующие волны лишь частично когерентны либо некогерентны, если интерференционная картина пропадает совсем.

В первом случае говорится о временной когерентности излучения, а во втором — о пространственной (см. § 26, 27). Временная когерентность характеризуется временем и длиной когерентности, а пространственная — шириной, радиусом и углом когерентности. Можно также говорить об объеме когерентности. Все эти понятия выражают идею частичной когерентности, для которой необходимо дать общее количественное выражение.

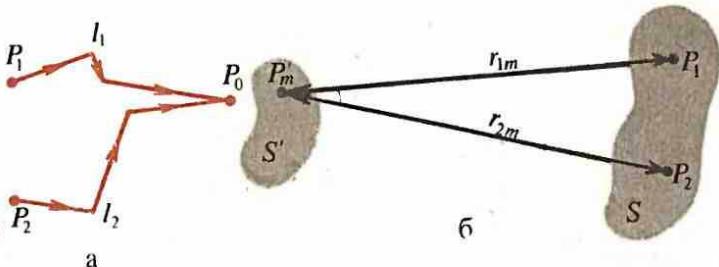
Функция взаимной когерентности. Схема опыта по осуществлению интерференции показана на рис. 140, *a*. В точке P_0 осуществляется интерференция лучей, исходящих из точек P_1 и P_2 . Пути лучей от точек P_1 и P_2 к P_0 изображены ломанными линиями, чтобы подчеркнуть возможность управления их движением с помощью зеркал, линз и других приспособлений. Чтобы не усложнять изложение несущественными уточнениями, будем считать, что l_1 и l_2 являются длинами путей, а скорость света при распространении по ним равна c . Следовательно, время, затрачиваемое лучом для прохождения путей l_1 и l_2 , равно соответственно $t_1 = l_1/c$ и $t_2 = l_2/c$.

Напряженность электрического поля в точках P_1 и P_2 обозначим $E_1(P_1, t')$, $E_2(P_2, t')$.

Амплитуды волн, пришедших из P_1 и P_2 в точку P_0 , изменяются. Эти изменения зависят от расстояний, пройденных волнами, размеров отверстий или других особенностей устройств, с помощью которых осуществляется дифракционный опыт. Однако они не влияют на характер изменения напряженности полей в зависимости от времени. Поэтому напряженности полей в точке P_0 получаются из напряженностей полей в точках P_1 и P_2 умножением на некоторые постоянные, учитывающие только что перечисленные факторы. В эти постоянные включается

140

К расчету степени когерентности (а); к доказательству теоремы
Ван-Цигтерта — Цернике (б)



в качестве множителя $-i$, если $E_1(P_1, t')$, $E_2(P_2, t')$ являются полями в точках P_1 и P_2 , а распространяющиеся к P_0 волны рассматриваются как волны от вторичных источников (см. § 32). Кроме того, необходимо учесть время прохождения путей l_1 и l_2 . Следовательно, полная напряженность поля в точке P_0

$$E(P_0, t) = a_1 E_1(P_1, t - l_1/c) + a_2 E_2(P_2, t - l_2/c), \quad (30.1)$$

где a_1 и a_2 — постоянные, определяющие изменения характеристик волн при переходе от точек P_1 и P_2 к точке P_0 , а аргументы $t - l_1/c$ и $t - l_2/c$ описывают запаздывание волн.

Распределение интенсивности в интерференционной картине описывается формулой

$$I(P_0) = \langle \operatorname{Re} E \operatorname{Re} E^* \rangle = \frac{1}{2} |a_1|^2 \operatorname{Re} \langle |E_1|^2 \rangle + \frac{1}{2} |a_2|^2 \operatorname{Re} \langle |E_2|^2 \rangle + |a_1 a_2| \operatorname{Re} \langle E_1 E_2^* \rangle, \quad (30.2)$$

причем аргументы у функций E_1 и E_2 не выписаны для упрощения написания формул. Они же, что и в (30.1). В (30.2) учтено, что a_1 и a_2 либо действительны, либо мнимы и, следовательно, $a_1 a_2^* = a_1 a_2$.

Мы рассматриваем стационарные процессы, когда средние величины не зависят от момента времени, относительно которого произведено усреднение. Следовательно,

$$\frac{1}{2} \langle E_1(P_1, t - l_1/c) E_1^*(P_1, t - l_1/c) \rangle = \frac{1}{2} \langle E_1(P_1, t) E_1^*(P_1, t) \rangle = I_1(P_1), \quad (30.3)$$

$$\frac{1}{2} \langle E_2(P_2, t - l_2/c) E_2^*(P_2, t - l_2/c) \rangle = I_2(P_2), \quad (30.4)$$

$$\boxed{\langle E_1(P_1, t - l_1/c) E_2^*(P_2, t - l_2/c) \rangle = \langle E_1(P_1, t) E_2^*(P_2, t + \tau) \rangle = \Gamma_{12}(\tau)}, \quad (30.5)$$

где $\tau = (l_1 - l_2)/v$, $I_1(P_1)$ и $I_2(P_2)$ — интенсивности волн в точках P_1 и P_2 . Функция $\Gamma_{12}(\tau)$ является основной величиной, характеризующей частичную когерентность. Она описывает взаимную когерентность световых колебаний в точках P_1 и P_2 в моменты времени, разделенные промежутком τ , и называется взаимной функцией когерентности волнового поля. Точки поля, к которым относится эта функция, обозначены буквами с индексами. Если точки P_1 и P_2 совпадают ($P_1 = P_2$), то соответствующая функция

$$\boxed{\Gamma_{11}(\tau) = \langle E_1(P_1, t) E_1^*(P_1, t + \tau) \rangle} \quad (30.6)$$

называется функцией автокогерентности световых колебаний в точке P_1 . В общей теории стационарных случайных процессов $\Gamma_{12}(\tau)$ называется взаимной корреляционной функцией величин $E_1(t)$ и $E_2(t)$, а $\Gamma_{11}(\tau)$ — автокорреляционной функцией величины $E_1(t)$.

192 С учетом (30.3) — (30.5) выражение (30.2) может быть представлено в виде

$$I(P_0) = |\alpha_1|^2 I_1(P_1) + |\alpha_2|^2 I_2(P_2) + |\alpha_1 \alpha_2| \operatorname{Re} \Gamma_{12}(\tau). \quad (30.7)$$

5 Комплексная степень когерентности. Учтем, что

$$|\alpha_1|^2 I_1(P_1) = I_1(P_0), \quad |\alpha_2|^2 I_2(P_2) = I_2(P_0), \\ |\alpha_1 \alpha_2| = 2 \sqrt{I_1(P_0)} \sqrt{I_2(P_0)} / \sqrt{\langle |E_1(P_1, t)|^2 \rangle \langle |E_2(P_2, t)|^2 \rangle}, \quad (30.8)$$

и представим (30.7) в виде

$$I(P_0) = I_1(P_0) + I_2(P_0) + 2 \sqrt{I_1(P_0)} \sqrt{I_2(P_0)} \operatorname{Re} \gamma_{12}(\tau), \quad (30.9)$$

где

$$\gamma_{12}(\tau) = \Gamma_{12}(\tau) / \sqrt{\langle |E_1(P_1, t)|^2 \rangle \langle |E_2(P_2, t)|^2 \rangle} \quad (30.10)$$

— комплексная степень когерентности. Формула (30.9) выражает общий закон интерференции для стационарных волновых полей. Для частных случаев эта формула была получена ранее [см., например, (26.7)]. При интерференции, осуществляющейся делением амплитуды, например, в интерферометре Майкельсона, в формулу (30.9) вместо $\gamma_{12}(\tau)$ входит $\gamma_{11}(\tau)$, относящаяся к точке P_1 , в которой происходит деление амплитуды.

Степень когерентности. Понятие когерентности связано с видимостью интерференционной картины. Вычислим видимость интерференционной картины излучения квазимонохроматического источника со средней частотой ω с помощью (30.9). Напряженности волн в точках P_1 и P_2 имеют вид

$$E_1(P_1, t') = E_{01}(t') \exp \{-i[\omega t' - \varphi_1(t')]\}, \quad (30.11)$$

$$E_2(P_2, t') = E_{02}(t') \exp \{-i[\omega t' - \varphi_2(t')]\}, \quad (30.12)$$

где амплитуды E_1 , E_2 и фазы φ_1 и φ_2 являются случайными функциями, причем E_{01} и E_{02} можно считать вещественными. Комплексная степень когерентности дается формулой

$$\gamma_{12}(\tau) = \langle E_{01}(t) E_{02}(t+\tau) \exp \{i[\omega t - \psi(\tau)]\} \rangle / \sqrt{\langle |E_{01}|^2 \rangle \langle |E_{02}|^2 \rangle}, \quad (30.13)$$

где

$$\psi = \varphi_2(t+\tau) - \varphi_1(t). \quad (30.14)$$

Множитель $\exp(i\omega t)$ можно вынести из-под знака усреднения и преобразовать (30.13) следующим образом:

$$\gamma_{12}(\tau) = (\langle E_{01} E_{02} \cos \psi \rangle - i \langle E_{01} E_{02} \sin \psi \rangle) e^{i\omega\tau} / \sqrt{\langle |E_{01}|^2 \rangle \langle |E_{02}|^2 \rangle} = \\ = |\gamma_{12}(\tau)| e^{-i\varphi} e^{i\omega\tau} = |\gamma_{12}(\tau)| e^{i(\omega\tau - \varphi)}, \quad (30.15)$$

где

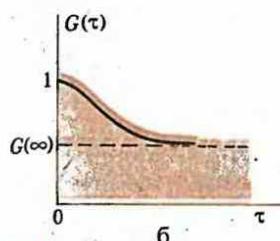
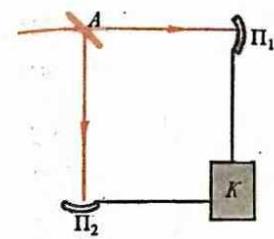
$$\operatorname{tg} \varphi = \langle E_{01} E_{02} \sin \psi \rangle / \langle E_{01} E_{02} \cos \psi \rangle, \quad (30.16)$$

$$|\gamma_{12}(\tau)| = \sqrt{(\langle E_{01} E_{02} \cos \psi \rangle)^2 + (\langle E_{01} E_{02} \sin \psi \rangle)^2} / \sqrt{\langle |E_{01}|^2 \rangle \langle |E_{02}|^2 \rangle}. \quad (30.17)$$

● Взаимная функция когерентности волнового поля и функция автокогерентности световых колебаний в общей теории стационарных случайных процессов называются соответственно взаимной корреляционной функцией и автокорреляционной функцией. Комплексная степень когерентности содержит информацию о флуктуациях амплитуды и фазы волны.

При одинаковой интенсивности интерферирующих пучков света видимость равна степени когерентности.

Теория частичной поляризации основана на изучении взаимной функции когерентности взаимно перпендикулярных компонент напряженности электрического поля волны.



141

Схема опыта Брауна и Твисса (а); зависимость корреляции интенсивностей от t (б)

Какая информация теряется при переходе от комплексной степени когерентности к степени когерентности?

Каким образом видимость интерференционной картины связана со степенью когерентности при произвольном соотношении интенсивностей интерферирующих пучков света?

Что такое матрица когерентности квазимохроматической волны?

Приведите определение степени поляризации и запишите ее выражение через экстремальные значения интенсивности.

Выпишите представления естественного света в виде суммы перпозиции линейно поляризованных волн и волн с циркулярной поляризацией.

С помощью неравенства Шварца нетрудно показать, что $|\gamma_{12}(\tau)| \ll 1$. Формула (30.9) приобретает вид

$$I = I_1 + I_2 + 2\sqrt{I_1}\sqrt{I_2}|\gamma_{12}(\tau)|\cos(\omega t - \phi). \quad (30.18)$$

Для квазимохроматического света величины $|\gamma_{12}(\tau)|$ и ϕ являются медленно изменяющимися функциями по сравнению с $\cos(\omega t)$, как это следует из анализа интерференционных картин (см. § 26, 27). Поэтому максимумы и минимумы интенсивности в (30.18) достигаются при значениях $\cos(\omega t - \phi)$, равных ± 1 , и видимость интерференционной картины равна

$$V = 2\sqrt{I_1}\sqrt{I_2}|\gamma_{12}(\tau)|/(I_1 + I_2). \quad (30.19)$$

Эта формула непосредственно связывает видимость интерференционной картины со степенью когерентности. При одинаковой интенсивности интерферирующих пучков ($I_1 = I_2$) видимость равна степени когерентности:

$$V = |\gamma_{12}(\tau)|. \quad (30.20)$$

Опыт Брауна и Твисса. В этом опыте была исследована корреляция интенсивности в световом пучке. Световой поток S (рис. 141, а) разделяется полупрозрачной пластиной A на две части, которые направляются к фотоприемникам P_1 и P_2 , проходя разные длины путей. Ток от приемников, пропорциональный световому потоку, направляется в коррелятор K , где в соответствующих электрических цепях вырабатывается ток, равный произведению сил токов. Измеряемой величиной является

$$G(\tau) = \frac{1}{\langle I^2 \rangle} \frac{1}{T} \int_0^T I(t)I(t+\tau)dt. \quad (30.21)$$

Поскольку $I \sim E^2$, здесь дело идет о корреляционной функции четвертого порядка относительно напряженности поля. На рис. 141, б. изображена зависимость $G(\tau)$, найденная в опытах Брауна и Твисса. При очень малых τ значение $G(\tau)$ близко к единице, при увеличении τ оно уменьшается. При больших τ функция $G(\tau)$ практически постоянна.

Для объяснения такого поведения $G(\tau)$ необходимо принять во внимание флуктуации интенсивности светового пучка. Если бы флуктуаций не было, то при всех значениях τ было бы $G(\tau) = 1$. Однако при наличии флуктуаций ситуация меняется. Для флуктуаций можно определить характерный масштаб времени. Если τ меньше характерного времени флуктуаций, то в корреляторе все время регистрируются примерно одинаковые силы токов и $G(\tau)$ близка к единице. При увеличении τ корреляция между силами токов в корреляторе нарушается, максимумы тока в одном канале попадают на минимумы в другом и т. д.,

в результате чего $G(t)$ уменьшается. Когда t превосходит характерное для флуктуаций время, его увеличение не вносит изменений в соотношение токов в каналах и значение $G(t)$ остается постоянным. Функция $G(t)$ дает информацию о статистических свойствах излучения.

Частичная поляризация. В квазимохроматической волне конец вектора напряженности электрического поля описывает чрезвычайно нерегулярную кривую в плоскости, перпендикулярной направлению распространения волны (см. § 13). В волне, распространяющейся в направлении положительных значений z , проекции напряженности электрического поля в фиксированной точке могут быть представлены аналогично (13.5) в виде

$$E_x(t) = E_{0x}(t) \exp \{-i[\omega t - \phi_x(t)]\}, \quad (30.22)$$

$$E_y(t) = E_{0y}(t) \exp \{-i[\omega t - \phi_y(t)]\}, \quad (30.23)$$

где E_{0x} , E_{0y} , ϕ_x , ϕ_y — случайные амплитуды и фазы. Для строго монохроматических волн эти величины постоянны, а волна имеет вполне определенную поляризацию. Квазимохроматическая волна не обладает определенной поляризацией. Однако если в плоскости X Y наблюдается некоторая регулярность в колебаниях вектора напряженности, то можно говорить о частичной поляризации. Количественная теория частичной поляризации основывается на анализе корреляции между взаимно перпендикулярными проекциями напряженности электрического поля волны и характеризуется матрицей когерентности.

Матрица когерентности квазимохроматической плоской волны. Вычислим интенсивности световых колебаний в направлении, составляющем угол θ с положительным направлением оси X . Проекция напряженности электрического поля на указанное направление представляется формулой

$$E(t, \theta, \varepsilon) = E_x \cos \theta + E_y \sin \theta e^{i\varepsilon}, \quad (30.24)$$

где ε — сдвиг фаз между проекциями, являющийся независимым параметром. Его значение задается по усмотрению экспериментатора. При $\varepsilon > 0$ у-проекция напряженности электрического поля волны запаздывает относительно x -проекции, а при $\varepsilon < 0$ опережает ее. Для интенсивности волны получаем формулу

$$I(\theta, \varepsilon) = \langle \operatorname{Re} E \operatorname{Re} E \rangle = \frac{1}{2} \operatorname{Re} \langle EE^* \rangle = \frac{1}{2} \cos^2 \theta \langle E_x E_x^* \rangle + \frac{1}{2} \sin^2 \theta \langle E_y E_y^* \rangle + \sin \theta \cos \theta \operatorname{Re} \langle E_x E_y^* e^{-i\varepsilon} \rangle = I_{xx} \cos^2 \theta + I_{yy} \sin^2 \theta + 2 \sin \theta \cos \theta \operatorname{Re}(I_{xy} e^{-i\varepsilon}), \quad (30.25)$$

где

$I_{xx} = \frac{1}{2} \langle E_x E_x^* \rangle$,	$I_{xy} = \frac{1}{2} \langle E_x E_y^* \rangle$,
$I_{yx} = \frac{1}{2} \langle E_y E_x^* \rangle$,	$I_{yy} = \frac{1}{2} \langle E_y E_y^* \rangle$.

(30.26)

Матрица

$$J = \begin{vmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{vmatrix} \quad (30.27)$$

называется матрицей когерентности квазимохроматической плоской волны. Она эрмитова, поскольку $I_{xy} = I_{yx}^*$.

Комплексная степень когерентности взаимно перпендикулярных проекций напряженности электрического поля волны. В полной аналогии с (30.13) комплексная степень когерентности μ_{xy} взаимно перпендикулярных проекций напряженности электрического поля волны определяется формулой

$$\mu_{xy} = \langle E_x E_y^* \rangle / \sqrt{\langle E_x E_x^* \rangle \langle E_y E_y^* \rangle} = I_{xy} / \sqrt{I_{xx} I_{yy}}, \quad (30.28)$$

С учетом (30.28) выражение (30.25) можно представить в виде

$$I(\theta, \epsilon) = I_{xx} \cos^2 \theta + I_{yy} \sin^2 \theta + 2 \sin \theta \cos \theta \sqrt{I_{xx}} \sqrt{I_{yy}} \operatorname{Re}(\mu_{xy} e^{-i\psi}), \quad (30.29)$$

аналогичном (30.9).

Для квазимонохроматической волны, описываемой уравнениями (30.22) и (30.23), имеем $I_{xy} = \langle E_{0x} E_{0y} e^{i\psi} \rangle = \langle E_{0x} E_{0y} \cos \psi \rangle + i \langle E_{0x} E_{0y} \sin \psi \rangle$, где $\psi = \phi_x - \phi_y$. Следовательно, $I_{xy} = |\mu_{xy}| e^{i\psi}$,

$$|I_{xy}| = \sqrt{(\langle E_{0x} E_{0y} \cos \psi \rangle)^2 + (\langle E_{0x} E_{0y} \sin \psi \rangle)^2}, \quad (30.30)$$

где $\operatorname{tg} \phi = \langle E_{0x} E_{0y} \sin \psi \rangle / \langle E_{0x} E_{0y} \cos \psi \rangle$. Отсюда комплексная степень когерентности

$$\mu_{xy} = |\mu_{xy}| e^{i\psi}, \quad (30.31a)$$

где

$$|\mu_{xy}| = |I_{xy}| / \sqrt{I_{xx} I_{yy}} \quad (30.31b)$$

степень когерентности взаимно перпендикулярных проекций напряженности электрического поля волны. Формула (30.29) приобретает вид

$$\begin{aligned} I(\theta, \epsilon) &= I_{xx} \cos^2 \theta + I_{yy} \sin^2 \theta + 2 \sin \theta \cos \theta \sqrt{I_{xx}} \sqrt{I_{yy}} |\mu_{xy}| \cos(\epsilon - \phi) = \\ &= I_{xx} \cos^2 \theta + I_{yy} \sin^2 \theta + 2 \sin \theta \cos \theta \sqrt{I_{xy} I_{yx}} \cos(\epsilon - \phi). \end{aligned} \quad (30.32)$$

Комплексная степень когерентности μ_{xy} содержит полную информацию о взаимной когерентности x - и y -проекций напряженности электрического поля волны. Необходимо обратить внимание на то, что степень когерентности $|\mu_{xy}|$ зависит, вообще говоря, от направления осей координат. Лишь для полностью неполяризованного света степень когерентности равна нулю для всех направлений осей в плоскости, перпендикулярной направлению распространения волн.

Естественный (неполяризованный) свет. У него интенсивность одинакова по всем направлениям и не зависит от разности фаз, которую можно ввести между взаимно перпендикулярными проекциями напряженности электрического поля волны, т. е. не зависит от θ и ϵ . Другими словами, $I(\theta, \epsilon) = \text{const}$. Из (30.32) видно, что это возможно только при

$$\mu_{xy} = 0, \quad I_{xy} = I_{yx} = 0, \quad (30.33) \quad I_{xx} = I_{yy} = \frac{1}{2} I, \quad (30.34)$$

где $I = I_{xx} + I_{yy}$ — полная интенсивность света.

Таким образом, у неполяризованного света степень когерентности взаимно перпендикулярных проекций напряженности электрического поля равна нулю, а **матрица когерентности** имеет вид

$$J = \frac{1}{2} I \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \quad (30.35)$$

Полностью поляризованный свет. Прежде всего рассмотрим монохроматические волны. Они описываются формулами (30.22) и (30.23) с независимыми от времени E_{0x} , E_{0y} , ϕ_x , ϕ_y . В этом случае средние равны мгновенным значениям величин и **матрица когерентности** имеет вид

$$J = \frac{1}{2} \begin{vmatrix} E_{0x}^2 & E_{0x} E_{0y} e^{i\psi} \\ E_{0x} E_{0y} e^{-i\psi} & E_{0y}^2 \end{vmatrix} \quad (30.36)$$

Величина $\psi = \phi_x - \phi_y$ является постоянной разностью фаз между взаимно перпендикулярными проекциями напряженности электрического поля волны. **Комплексная степень когерентности** по формуле (30.28) выражается формулой

$$\mu_{xy} = e^{i\psi}$$

$$(30.37)$$

и, следовательно,

$$|\mu_{xy}| = 1.$$

(30.38)

Это означает, что степень когерентности при переходе от полностью неполяризованного света к полностью поляризованному изменяется от 0 до 1. Значения коэффициента между 0 и 1 описывают частичную поляризацию. Однако количественно поляризация характеризуется степенью поляризации P [см. (30.63)], которая не равна $|\mu_{xy}|$.

Определитель матрицы когерентности (30.36) полностью поляризованного света

$$\det J = I_{xx}I_{yy} - I_{xy}I_{yx} = 0. \quad (30.39)$$

Это равенство является важнейшим условием полной поляризации.

Полной поляризацией могут обладать не только монохроматические, но и квазимохроматические волны, для которых условие (30.39) полной поляризации соблюдается в случае, если

$$E_{0y}(t)/E_{0x}(t) = \text{const} = a, \quad \varphi_x - \varphi_y = \psi = \text{const}. \quad (30.40)$$

Имеем $I_{xx} = \langle E_{0x}^2 \rangle$, $I_{xy} = a \langle E_{0x}^2 \rangle e^{i\psi}$, $I_{yx} = a \langle E_{0x}^2 \rangle$, $I_{yy} = a^2 \langle E_{0x}^2 \rangle$ и, следовательно, уравнение (30.39) удовлетворяется.

Это можно понять из физических соображений. Равенство (30.39) показывает, что I_{xx} , I_{xy} , I_{yx} , I_{yy} у полностью поляризованного света не являются независимыми. Условие (30.40) выражает ту зависимость проекций E_{0x} от E_0 , и их фаз, которая обеспечивает выполнение соотношения (30.9), характеризующего полностью поляризованный свет.

Линейно поляризованный свет определяется условием $\psi = m\pi$ ($m = 0, \pm 1, \pm 2, \dots$). Матрица когерентности (30.36) принимает вид

$$J = \frac{1}{2} \begin{vmatrix} E_{0x}^2 & (-1)^m E_{0x} E_{0y} \\ (-1)^m E_{0x} E_{0y} & E_{0y}^2 \end{vmatrix} \quad (30.41)$$

Направление колебаний вектора \mathbf{E} определяется соотношением $E_x/E_y = (-1)^m E_{0x}/E_{0y}$. Если направление колебаний электрического вектора совпадает с осью X ($E_{0y} = 0$), то

$$J = I \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix}, \quad (30.42)$$

где $I = E_{0x}^2/2$ — интенсивность волны. При направлении колебаний по оси Y

$$J = I \begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix}. \quad (30.43)$$

Если же вектор напряженности электрического поля волны колеблется в направлении биссектрис углов между осями координат, т. е. под углами $\pi/4$ и $3\pi/4$ к оси X , то $E_{0x} = E_{0y}$, $m=0$ и $E_{0x} = E_{0y}$, $m=1$, а матрицы когерентности имеют вид

$$J = \frac{1}{2} I \begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix}, \quad J = \frac{1}{2} I \begin{vmatrix} 1 & -1 \\ -1 & 1 \end{vmatrix}. \quad (30.44)$$

У поляризованного по кругу света $E_{0x} = E_{0y}$, $\psi = m\pi/2$ ($m = \pm 1, \pm 3, \dots$) и, следовательно, матрица когерентности записывается так:

$$J = \frac{1}{2} I \begin{vmatrix} 1 & \pm i \\ \mp i & 1 \end{vmatrix}, \quad (30.45)$$

где верхние и нижние знаки при i относятся к круговым поляризациям различного направления

Степень поляризации световой волны. Независимые световые волны между собой не интерферируют и, следовательно, средние значения произведений проекций различных волн равны нулю. Это приводит к заключению, что матрица когерентности результирующей волны равна сумме матриц когерентности слагаемых волн. Для доказательства рассмотрим суперпозицию независимых волн. Проекции напряженностей электрического поля результирующей волны даются равенствами

$$E_x = \sum_{n=1}^N E_x^{(n)}, \quad E_y = \sum_{m=1}^N E_y^{(m)}. \quad (30.46)$$

Отсюда для средних значений произведений проекций получаем

$$\langle E_i E_j^* \rangle = \sum_{n=1}^N \sum_{m=1}^N \langle E_i^{(n)} E_j^{(m)*} \rangle = \sum_{n=1}^N \langle E_i^{(n)} E_j^{(n)*} \rangle + \sum_{n \neq m}^N \langle E_i^{(n)} E_j^{(m)*} \rangle, \quad (30.47)$$

где $i = x, y; j = x, y$. Члены последней суммы в (30.47) равны нулю, поскольку являются средними от произведений проекций напряженностей электрического поля различных независимых волн. Поэтому приходим к равенству

$$\langle E_i E_j^* \rangle = \sum_{n=1}^N \langle E_i^{(n)} E_j^{(n)*} \rangle, \quad (30.48)$$

которое доказывает, что матрица когерентности результирующей волны равна сумме матриц когерентности для слагаемых волн. Это подсказывает возможность представления матрицы когерентности произвольной волны в виде суммы матриц полностью поляризованной и полностью неполяризованной волн. Для осуществления такой возможности необходимо доказать единственность представления матрицы

$$J = \begin{vmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{vmatrix} \quad (30.49)$$

в виде

$$J = J^{(1)} + J^{(2)}, \quad (30.50)$$

где

$$J^{(1)} = \begin{vmatrix} A & 0 \\ 0 & A \end{vmatrix} \quad (30.51)$$

в соответствии с (30.35) является матрицей когерентности полностью неполяризованного света, а

$$J^{(2)} = \begin{vmatrix} B & D \\ D^* & C \end{vmatrix}, \quad (30.52)$$

удовлетворяющая условию (30.39) в виде

$$BC - DD^* = 0, \quad (30.53)$$

является матрицей когерентности полностью поляризованного света. Из (30.50) имеем равенства

$$A + B = I_{xx}, \quad D = I_{xy}, \quad D^* = I_{yx}, \quad A + C = I_{yy}, \quad (30.54)$$

с помощью которых уравнение (30.53) приводится к виду

$$A^2 - (I_{xx} + I_{yy})A + I_{xx}I_{yy} - I_{xy}I_{yx} = 0. \quad (30.55)$$

198 Отсюда находим корни уравнения:

5 $A = \frac{1}{2} (I_{xx} + I_{yy}) \pm \frac{1}{2} \sqrt{(I_{xx} + I_{yy})^2 - 4 \det J},$ (30.56)

где

$$\det J = I_{xx}I_{yy} - I_{xy}I_{yx} \quad (30.57)$$

— определитель матрицы (30.49)

Чтобы проанализировать решения этого уравнения, необходимо учесть некоторые неравенства. Из (30.28) получаем

$$1 - \mu_{xy}\mu_{xy}^* = (I_{xx}I_{yy} - I_{xy}I_{yx})/(I_{xx}I_{yy}) = \det J/(I_{xx}I_{yy}), \quad (30.58)$$

$$\mu_{xy} = I_{xy}/\sqrt{I_{xx}I_{yy}}. \quad (30.59)$$

Из формулы (30.59) в соответствии с неравенством Шварца следует, что $|\mu_{xy}| \leq 1$. Учитывая также, что I_{xx} и I_{yy} неотрицательны, из (30.58) находим $0 < \det J \leq I_{xx}I_{yy}$. Кроме того, из неравенства $(I_{xx} - I_{yy})^2 \geq 0$ получаем $(I_{xx} + I_{yy})^2 \geq 4I_{xx}I_{yy}$ и поэтому $0 < \det J \leq I_{xx}I_{yy} \leq 1/4(I_{xx} + I_{yy})^2$. Это означает, что оба корня (30.56) вещественные и положительны. Непосредственной проверкой убеждаемся, что при положительном знаке корня в (30.56) для B и C получаются отрицательные значения, которые неприемлемы, поскольку B и C , по определению, положительны. Следовательно, необходимо в (30.56) взять отрицательный знак у корня. Решение единствено и имеет следующий вид:

$$A = \frac{1}{2} (I_{xx} + I_{yy}) - \frac{1}{2} \sqrt{(I_{xx} + I_{yy})^2 - 4 \det J}, \quad (a)$$

$$B = \frac{1}{2} (I_{xx} - I_{yy}) + \frac{1}{2} \sqrt{(I_{xx} + I_{yy})^2 - 4 \det J}, \quad (b)$$

$$C = \frac{1}{2} (I_{yy} - I_{xx}) + \frac{1}{2} \sqrt{(I_{xx} + I_{yy})^2 - 4 \det J}, \quad (c)$$

$$D = I_{xy}, \quad D^* = I_{yx}. \quad (g)$$

Полная интенсивность волны и интенсивность поляризованной части

$$I = I_{xx} + I_{yy}, \quad (30.61)$$

$$I_{\text{пол}} = B + C = \sqrt{(I_{xx} + I_{yy})^2 - 4 \det J}. \quad (30.62)$$

Степенью поляризации P волны называется отношение интенсивности поляризованной части к полной интенсивности:

$$P = \frac{I_{\text{пол}}}{I} = \sqrt{1 - 4 \det J / (I_{xx} + I_{yy})^2}. \quad (30.63)$$

Из рассмотренных в связи с решением (30.60) неравенств следует, что $0 < P \leq 1$. Из теории матриц известно, что $\det J$ и $\text{Sp } J = I_{xx} + I_{yy}$ инвариантны при вращении системы координат. Отсюда с учетом (30.63) заключаем, что P не зависит от направления осей X и Y .

Выражение степени поляризации через экстремальные значения интенсивности. Исследуем изменение интенсивности в зависимости от θ и ε . Воспользовавшись тригонометрической фор-

мулой для синусов и косинусов двойных углов, запишем (30.32) в виде

$$I(\theta, \varepsilon) = \frac{1}{2}(I_{xx} + I_{yy}) + \frac{1}{2}(I_{yy} - I_{xx}) \cos 2\theta + \sqrt{I_{xy}I_{yx}} \cos(\varepsilon - \varphi) \sin 2\theta. \quad (30.64)$$

Эту формулу можно преобразовать аналогично тому, как (26.23) была преобразована в (26.25а):

$$I(\theta, \varepsilon) = Q + \sqrt{C^2 + S^2} \cos(2\theta - a), \quad (30.65)$$

где $Q = (I_{xx} + I_{yy})/2$, $C = (I_{yy} - I_{xx})/2$, $S = \sqrt{I_{xy}I_{yx}} \cos(\varepsilon - \varphi)$, $\operatorname{tg} a = S/C$.

Амплитуда колебаний интенсивности зависит от $\cos(\varepsilon - \varphi)$, достигая максимального значения при $|\cos(\varepsilon - \varphi)| = 1$. Максимумы и минимумы интенсивности достигаются при углах θ , для которых $\cos(2\theta - a) = 1$ и $\cos(2\theta - a) = -1$ соответственно. Следовательно,

$$I(\theta, \varepsilon)|_{\max} = Q + \sqrt{C^2 + S^2}, \quad I(\theta, \varepsilon)|_{\min} = Q - \sqrt{C^2 + S^2}.$$

При нахождении экстремумов посредством только что использованных приемов необходимо соблюдать осторожность. Дело в том, что зависимость I в (30.65) от θ и ε содержит не только в аргументах $\cos(\varepsilon - \varphi)$ и $\cos(2\theta - a)$, но и в величинах I_{xx} , I_{xy} , I_{yx} , I_{yy} . Оправданием использованного приема является то обстоятельство, что зависимость I от ε и θ посредством $\cos(\varepsilon - \varphi)$ и $\cos(2\theta - a)$ много сильнее зависимости посредством I_{xx} , I_{xy} , I_{yx} , I_{yy} и при нахождении экстремумов последней зависимостью можно пренебречь.

Примем во внимание, что

$$\begin{aligned} \sqrt{C^2 + S^2} &= \frac{1}{2} \sqrt{(I_{yy} - I_{xx})^2 + 4I_{xy}I_{yx}} = \frac{1}{2} \sqrt{(I_{yy} + I_{xx})^2 - 4(I_{xx}I_{yy} - I_{xy}I_{yx})} = \\ &= \frac{1}{2} \sqrt{(I_{yy} + I_{xx})^2 - 4\det J}, \end{aligned}$$

где $\det J = I_{xx}I_{yy} - I_{xy}I_{yx}$ — определитель матрицы когерентности. Поэтому

$$\frac{I(\theta, \varepsilon)|_{\max} - I(\theta, \varepsilon)|_{\min}}{I(\theta, \varepsilon)|_{\max} + I(\theta, \varepsilon)|_{\min}} = \frac{\sqrt{C^2 + S^2}}{Q} = \sqrt{1 - 4\det J/(I_{xx} + I_{yy})^2}. \quad (30.66)$$

Сравнивая правые части (30.66) и (30.63), заключаем, что

$$P = (I(\theta, \varepsilon)|_{\max} - I(\theta, \varepsilon)|_{\min}) / (I(\theta, \varepsilon)|_{\max} + I(\theta, \varepsilon)|_{\min}), \quad (30.67)$$

причем под экстремальными значениями здесь понимаются абсолютные максимумы и минимумы относительно обеих независимых переменных θ и ε . Полностью поляризованный свет характеризуется значением $P = 1$. Из (30.63) следует, что при этом $\det J = I_{xx}I_{yy} - I_{xy}I_{yx} = 0$. Это означает, что $|\mu_{xy}| = 1$, т. е. E_x и E_y взаимно полностью когерентны. Полностью неполяризованный свет характеризуется значением $P = 0$. Из (30.63) следует, что

$$(I_{xx} + I_{yy})^2 - 4(I_{xx}I_{yy} - I_{xy}I_{yx}) = (I_{xx} - I_{yy})^2 + 4I_{xy}I_{yx} = 0. \quad (30.68)$$

Поскольку $I_{xy}I_{yx} = I_{xy}I_{xy}^*$ является положительной величиной, каждое из слагаемых в правой части (30.68) должно быть равно нулю в отдельности:

$$I_{xx} = I_{yy}, \quad I_{xy} = I_{yx} = 0 \quad (30.69)$$

Следовательно, $|\mu_{xy}| = 0$, E_x и E_y взаимно полностью некогерентны. При $0 < P < 1$ свет частично поляризован, $0 < |\mu_{xy}| < 1$, E_x и E_y частично взаимно когерентны.

Представления естественного света. Пользуясь возможностью представления матрицы когерентности для результирующей волны в виде суммы матриц когерентности слагаемых волн, можно матрицу когерентности (30.35) естественного света представить в виде

200
5

$$\frac{1}{2} I \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = \frac{1}{2} I \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} + \frac{1}{2} I \begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix}, \quad (30.70)$$

$$\frac{1}{2} I \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = \frac{1}{4} I \begin{vmatrix} 1+i & -i \\ -i & 1 \end{vmatrix} + \frac{1}{4} I \begin{vmatrix} 1-i & i \\ i & 1 \end{vmatrix} \quad (30.71)$$

В первом случае волна естественного света интенсивности I представлена в виде суммы двух линейно поляризованных волн с интенсивностями $I/2$. Электрические векторы колеблются во взаимно перпендикулярных направлениях в плоскости, перпендикулярной направлению распространения. Во втором случае волна естественного света эквивалентна двум независимым волнам, поляризованным по правому и левому кругу. Интенсивность каждой из слагаемых волн равна половине интенсивности I естественного света.

Соотношение между степенью поляризации и степенью когерентности. Из (30.59) находим, что

$$|\mu_x|^2 = 1 - \det J / (I_{xx} I_{yy}). \quad (30.72)$$

Квадрат степени поляризации (30.63) выражается формулой

$$P^2 = 1 - \det J / [(I_{xx} + I_{yy})^2 / 4]. \quad (30.73)$$

Из неравенства $(I_{xx} - I_{yy})^2 \geq 0$ следует $(I_{xx} + I_{yy})^2 \geq 4I_{xx}I_{yy}$, и поэтому сравнение (30.72) с (30.73) показывает, что

$$P \geq |\mu_{xy}|, \quad (30.74)$$

причем знак равенства достигается только при $I_{xx} = I_{yy}$. Чтобы проанализировать, когда это происходит, рассмотрим преобразование проекций напряженностей поля при повороте осей координат. Если ось X' новой системы координат составляет с осью X угол α (начала систем координат совпадают, а поворот осуществляется в плоскости $X Y$), то

$$E_{x'} = E_x \cos \alpha + E_y \sin \alpha, \quad E_{y'} = -E_x \sin \alpha + E_y \cos \alpha. \quad (30.75)$$

Тогда

$$I_{x'x'} = I_{xx} \cos^2 \alpha + I_{yy} \sin^2 \alpha + (I_{xy} + I_{yx}) \cos \alpha \sin \alpha,$$

$$I_{y'y'} = I_{xx} \sin^2 \alpha + I_{yy} \cos^2 \alpha - (I_{xy} + I_{yx}) \sin \alpha \cos \alpha. \quad (30.76)$$

Из условия $I_{x'x'} = I_{y'y'}$ находим

$$\operatorname{tg} 2\alpha = (I_{yy} - I_{xx}) / (I_{xy} + I_{yx}). \quad (30.77)$$

Ввиду того что $I_{xy} = I_{yx}$, решение уравнения (30.77) дает два взаимно ортогональных направления, для которых интенсивности $I_{x'x'}$ и $I_{y'y'}$ одинаковы, а степень когерентности $\mu_{x'y'}$ достигает максимального значения, равного значению степени поляризации P волны.

Теорема Ван-Циттерта — Цернике. В § 26, 27 были рассмотрены конкретные случаи проявления временной и пространственной когерентности. Поскольку степень когерентности определяет видимость интерференционной картины, важно уметь находить степень когерентности излучения, не зная видимости интерференционной картины. Для квазимонохроматического

излучения от источника небольших размеров, находящегося на достаточно большом расстоянии, такую возможность представляет теорема Ван-Циттерта — Цернике. Она справедлива не для любого достаточно малого источника на достаточно большом расстоянии, а лишь при условии, что разность времен распространения света от любой точки источника до двух точек наблюдения, для которых определяется взаимная степень когерентности излучения, меньше времени когерентности излучения из соответствующих точек источника.

Для упрощения расчета предполагаем, что **светимость поверхности S'** (см. рис. 140, б) создается находящимися на ней точечными квазимохроматическими источниками с одинаковой средней частотой излучения, между собой не когерентными. От каждого из них в направлении поверхности S испускается волна. Если напряженность электрического поля волны, испускаемая источником в точке P'_m , пропорциональна $E'_m(t')$, то амплитуда волны в точке P_1 на плоскости наблюдения S равна

$$E_{1m}(P_1, t) = E'_m(t - r_{1m}/v)/r_{1m}, \quad (30.78)$$

где v — скорость распространения волны в среде между точками P'_m и R_m . Наличие $t - r_{1m}/v$ в аргументе у E'_m учитывает запаздывание колебаний в точке P_1 по сравнению с точкой P'_m , а r_{1m} в знаменателе принимает во внимание уменьшение амплитуды напряженности в процессе распространения волны. Нормировочная постоянная в (30.78) не выписана. Обозначение точки P_{1m} в (30.78) дается индексом 1 у E_{1m} .

Для квазимохроматических излучателей

$$E'_m(t') = E'_{0m}(t') e^{-i\omega t'}. \quad (30.79)$$

Тогда выражение (30.78) принимает вид

$$E_{1m}(t) = E'_{0m}(t - r_{1m}/v) \exp(-i\omega(t - r_{1m}/v)) r_{1m}^{-1}. \quad (30.80)$$

Напряженность электрического поля в точке P_2 дается аналогичной (30.80) формулой

$$E_{2m}(t) = E'_{0m}(t - r_{2m}/v) \exp(-i\omega(t - r_{2m}/v)) r_{2m}^{-1}. \quad (30.81)$$

В точки P_1 и P_2 в момент времени t приходит излучение от всех источников светящейся поверхности S' . Поэтому суммарная напряженность электрических полей в этих точках

$$E_1(t) = \sum_m E_{1m}(t) = \sum_m E'_{0m}(t - r_{1m}/v) \exp(-i\omega(t - r_{1m}/v)) r_{1m}^{-1}, \quad (30.82)$$

$$E_2(t) = \sum_{m'} E_{2m'}(t) = \sum_{m'} E'_{0m'}(t - r_{2m'}/v) \exp(-i\omega(t - r_{2m'}/v)) r_{2m'}^{-1}, \quad (30.83)$$

где суммирование распространяется по всем излучателям на площади S' поверхности.

В полной аналогии с (30.5), изменяя соответствующим образом начало отсчета времени, для взаимной корреляционной функции $\Gamma_{12}(t)$ величин E_1 и E_2 получаем выражение

$$\Gamma_{12}(t) = \langle E_1(t') E_2^*(t') \rangle = \sum_{m, m'} \langle E'_{0m}(t - r_{1m}/v) E'_{0m'}^*(t - r_{2m'}/v) r_{1m}^{-1} r_{2m'}^{-1} e^{i\omega t_{mm'}} \rangle \quad (30.84)$$

где $t_{mm'} = (r_{1m} - r_{2m'})/v$. Поскольку точечные источники излучения на S' являются между собой статистически независимыми, имеем

$$\langle E'_{0m}(t - r_{1m}/v) E'_{0m'}^*(t - r_{2m'}/v) \rangle = \langle E'_{0m}(t - r_{1m}/v) E'_{0m'}^*(t - r_{2m'}/v) \rangle \delta_{mm'} \quad (30.85)$$

202 и формула (30.84) принимает вид

$$5 \quad \Gamma_{12} = \sum_m \langle E'_{0m}(t - r_{1m}/v) E'^*_m(t - r_{2m}/v) \rangle r_{1m}^{-1} r_{2m}^{-1} e^{i\omega t_m} \quad (30.86)$$

где $\omega t_m = \omega(r_{1m} - r_{2m})/v = k(r_{1m} - r_{2m})$, $k = \omega/v = 2\pi/\lambda$ — волновое число.

Рассматриваемые процессы излучения и наблюдения предполагаются происходящими в стационарных условиях. Поэтому в правой части (30.86) под знаком усреднения можно заменить начало отсчета времени и написать [см. (30.5)]

$$\langle E'_{0m}(t - r_{1m}/v) E'^*_m(t - r_{2m}/v) \rangle = \langle E'_{0m}(t) E'^*_m(t + \tau_m) \rangle.$$

Предполагается, что время когерентности излучения всех точечных источников больше, чем разность времен хода от источника до точек наблюдения, т. е. $\tau_{\text{кор}} > |\tau_m|$. Вычисляя (30.86), при этом условии с достаточно большой точностью можно положить $\tau_m = 0$ и принять, что

$$\langle E'_{0m}(t) E'^*_m(t + \tau_m) \rangle = \langle E'_{0m}(t) E'^*_m(t) \rangle = 2I'_m, \quad (30.87)$$

где I'_m — интенсивность излучения, соответствующая амплитуде напряженности E'_{0m} . С учетом (30.87), переходя в (30.86) от суммирования к интегрированию, получаем

$$\boxed{\Gamma_{12} = 2Q \int_S \frac{I'(P') e^{ik(r_1 - r_2)}}{r_1 r_2} dS'}, \quad (30.88)$$

где dS' — площадь элемента поверхности S' в точке P' ; r_1 и r_2 — расстояния от P' до точек P_1 и P_2 соответственно. Постоянная Q является нормировочной. Она возникает в результате перехода от суммирования к интегрированию. Ее точное значение не нужно для дальнейших вычислений, поскольку она сокращается с соответствующими постоянными, которые в этих вычислениях появляются в знаменателях.

Из (30.88) для функции автокогерентности (30.6) световых колебаний в точках P_1 и P_2 получаем:

$$\boxed{\Gamma_{11} = 2Q \int_S \frac{I'(P')}{r_1^2} dS' = 2Q I(P_1)}, \quad (30.89)$$

$$\boxed{\Gamma_{22} = 2Q \int_S \frac{I'(P')}{r_2^2} dS' = 2Q I(P_2)}, \quad (30.90)$$

— интенсивности излучения в точках P_1 и P_2 при отсутствии интерференции излучения от различных источников на S' .

Формула (30.10) для комплексной степени когерентности с учетом (30.88), (30.89), (30.90), а также с помощью соотношений

$$I(P_1) = I_1 = \int_S \frac{I'(P')}{r_1^2} dS', \quad (30.91)$$

$$I(P_2) = I_2 = \int_S \frac{I'(P')}{r_2^2} dS' \quad (30.92)$$

принимает вид

$$\boxed{\gamma_{12} = \frac{\Gamma_{12}}{\sqrt{\Gamma_{11} \Gamma_{22}}} = \frac{1}{\sqrt{I_1 I_2}} \int_S \frac{I'(P') e^{ik(r_1 - r_2)}}{r_1 r_2} dS'}. \quad (30.93)$$

Она позволяет вычислить комплексную степень когерентности излучения в точках волнового поля, если известно распределение интенсивности излучения на поверхности протяженного квазимонокроматического источника, в предположении, что излучение источника удовлетворяет перечисленным ранее требованиям. Формула (30.93) выражает теорему Ван-Циттерта — Цернке.

Пример 30.1. Найти взаимную функцию когерентности $\Gamma_{12}(\tau)$ волнового поля для хаотического светового излучения с лоренцевой формой линии. Световой поток образуется совокупностью плоских волн, распространяющихся в направлении положительных значений оси Z .

Для упрощения расчета удобно считать излучение заключенным в полости очень больших размеров (много большей длины когерентности в направлении оси Z). Границы полости можно считать прозрачными. Электрическое поле в точке z внутри полости в момент времени t представляется в виде суммы нормальных мод излучения:

$$E(z, t) = \sum_k E_k \exp(-i(\omega_k t - kz)), \quad (30.94)$$

где $\omega_k = ck$. Амплитуды E_k являются комплексными величинами, абсолютные значения которых и фазы определяются только статистически.

По определению взаимной функции когерентности волнового поля,

$$\Gamma_{12}(\tau) = \langle E(z_1, t_1) E^*(z_2, t_2) \rangle = \sum_{k, k'} \langle E_k E_{k'}^* \rangle \exp[-i(\omega_k t_1 - \omega_{k'} t_2 - kz_1 + kz_2)]. \quad (30.95)$$

Поскольку амплитуды E_k и $E_{k'}$ при $k \neq k'$ — случайные независимые величины с нулевыми средними значениями, имеем

$$\langle E_k E_{k'}^* \rangle = \langle |E_k|^2 \rangle \delta_{kk'}. \quad (30.96)$$

Поэтому выражение (30.95) принимает вид

$$\Gamma_{12}(\tau) = \sum_k \langle |E_k|^2 \rangle \exp(-i\omega_k \tau), \quad (30.97)$$

где $\tau = t_1 - t_2 - (z_1 - z_2)/c$.

В соответствии с (9.28) для лоренцевой формы линии можем написать

$$\langle |E_k|^2 \rangle = A \frac{1}{(\omega_0 - \omega_k)^2 + (\gamma/2)^2}, \quad (30.98)$$

где A — постоянная, определяемая средним потоком энергии

$$\begin{aligned} \langle S \rangle &= 1/2 \sqrt{\epsilon_0/\mu_0} \langle |E(z, t)|^2 \rangle = \\ &= 1/2 \sqrt{\epsilon_0/\mu_0} \sum_{k, k'} \langle E_k E_{k'}^* \rangle \exp[(\omega_k - \omega_{k'})t - (k - k')z] = 1/2 \sqrt{\epsilon_0/\mu_0} \sum_k \langle |E_k|^2 \rangle, \end{aligned} \quad (30.99)$$

где принято во внимание соотношение (39.96). При большой длине полости ($L \rightarrow \infty$) можно по обычному правилу в (30.99) перейти от суммы к интегралу:

$$\sum_k \rightarrow \frac{L}{\pi} \int_0^\infty dk \rightarrow \frac{L}{\pi c} \int_0^\infty d\omega_k. \quad (30.100)$$

При таком переходе с учетом (30.98) при $\gamma \ll \omega_0$ находим

$$\langle S \rangle = \frac{1}{2} \sqrt{\frac{\epsilon_0}{\mu_0}} \frac{AL}{\pi c} \int_0^\infty \frac{d\omega_k}{(\omega_0 - \omega_k)^2 + (\gamma/2)^2} = \frac{1}{2} \sqrt{\frac{\epsilon_0}{\mu_0}} \frac{2AL}{\gamma c}. \quad (30.101)$$

Подставляя A из (30.101) в (30.98), получаем

$$5 \quad \langle |E_k|^2 \rangle = \frac{\gamma c}{L} \sqrt{\frac{\mu_0}{\epsilon_0}} \langle S \rangle \frac{1}{(\omega_0 - \omega_k)^2 + (\gamma/2)^2}.$$

Тогда (30.97) принимает вид

$$\Gamma_{12}(\tau) = \frac{\gamma c}{L} \sqrt{\frac{\mu_0}{\epsilon_0}} \langle S \rangle \sum_k \frac{\exp(-i\omega_k \tau)}{(\omega_0 - \omega_k)^2 + (\gamma/2)^2}. \quad (30.102)$$

Переходя в (30.102) от суммы к интегралу, по формуле (30.100) находим

$$\Gamma_{12}(\tau) = \frac{\gamma}{\pi} \sqrt{\frac{\mu_0}{\epsilon_0}} \langle S \rangle \int_0^\infty \frac{\exp(-i\omega_k \tau) d\omega_k}{(\omega_0 - \omega_k)^2 + (\gamma/2)^2}.$$

С помощью теории вычетов находим значение интеграла:

$$\int_{-\infty}^{\infty} \frac{\exp(-i\omega_k \tau) d\omega_k}{(\omega_0 - \omega_k)^2 + (\gamma/2)^2} = \frac{2\pi}{\gamma} \exp\left(-i\omega_0 \tau - \frac{\gamma|\tau|}{2}\right),$$

поэтому

$$\Gamma_{12}(\tau) = 2 \sqrt{\mu_0/\epsilon_0} \langle S \rangle \exp\left(-i\omega_0 \tau - \frac{\gamma|\tau|}{2}\right). \quad (30.103)$$

Таким образом, $\Gamma_{12}(\tau)$ экспоненциально уменьшается как при увеличении ширины линии, так и при увеличении $|\tau|$.

Задачи

- 5.1. Расстояние между точечными источниками $\lambda/2$, а исходящие из них волны имеют разность фаз π и в отдельности дают интенсивность I_0 . Найти распределение интенсивностей в интерференционной картине.
- 5.2. Расстояние между точечными источниками $\lambda/4$, а исходящие из них волны имеют разность фаз $\pi/2$ и в отдельности дают интенсивность I_0 . Найти распределение интенсивностей в интерференционной картине.
- 5.3. Радиоастрономический интерферометр по принципу действия эквивалентен дифракционной решетке, в которой роль щелей выполняют приемники радиоизлучения, от которых, как от источников вторичных волн, образуется интерференционная картина. Линейная цепочка радиоастрономического интерферометра состоит из $N = 40$ приемников-источников, находящихся на расстоянии $d = 10$ м друг от друга и работающих на длине волн $\lambda = 21$ см. Найти угловую ширину центрального максимума и угловое расстояние между центральным максимумом и главными максимумами первого порядка.
- 5.4. Кольца Ньютона наблюдаются при освещении D -линией натрия с $\lambda = 589$ нм. При медленном удалении линзы от поверхности, на

которой она первоначально лежала, кольца Ньютона как бы стягиваются к центру и в центре происходит последовательное потемнение и просветление. Число стягиваемых к центру колец можно подсчитать. Когда число стянутых к центру колец стало приближаться к 500, видимость дифракционной картины ухудшилась и дифракционная картина исчезла при 500 стянутых кольцах. При дальнейшем увеличении расстояния между линзой и поверхностью видимость улучшается и достигает своего максимального значения при 1000 колцах. Дальнейшие минимумы и максимумы получаются при 1500 и 2000 колцах соответственно. Такое поведение интерференционной картины обусловлено тем, что D -линия натрия является дублетом. Найти расстояние между линиями дублета.

- 5.5. В интерферометре Маха—Цендера (см. рис. 108) длина пути света в ячейке Q равна $l = 20,4$ см. Из ячейки воздух практически полностью откачен. В центре интерференционной картины наблюдается темное пятно. Длина волны монохроматического источника света равна $\lambda = 589$ нм. В ячейку через кран медленно впускается воздух. В результате в объеме внутри ячейки давление

- поднимается, радиусы колец в интерференционной картине изменяются, а в ее центре наблюдается последовательное увеличение и уменьшение освещенности. Экспериментатор считает число просветлений. К тому моменту, когда давление воздуха в ячейке стало равным атмосферному, экспериментатор насчитал 101 просветление, а в установившейся при атмосферном давлении интерференционной картине в центре наблюдается потемнение. Чему равен показатель преломления воздуха при атмосферном давлении?
- 5.6.** На сколько в сторону сместится центральный максимум в интерференционной картине от двух щелей, если одну из них закрыть тонкой плоскопараллельной пластинкой толщиной q с показателем преломления n ? Расстояние от щелей до экрана l (см. рис. 113).
- 5.7.** На горизонтальном стеклянном столе лежит плоскопараллельная пластина. Одна сторона пластины касается стола, а другая несколько приподнята, в результате чего между плоскостями пластины и стола образуется очень маленький угол α . Сверху на пластину и стол падает по нормали плоская монохроматическая волна с длиной волны 589 нм. Расстояние между линиями одинаковой толщины равно 5 мм. Определить угол α .
- 5.8.** В интерферометре Тваймана — Грина (см. рис. 110) собирающая линза L отсутствует, а отражение происходит непосредственно от выпуклого зеркала с радиусом кривизны $R = 5$ м. Источник света монохроматический с длиной волны 550 нм. Интерференционная картина наблюдается невооруженным глазом на экране без линзы. Известно, что расстояния от центра делительной пластиинки до A_1 и до выпуклого зеркала A_2 равны и поэтому в центре интерференционной картины наблюдается светлое пятно. Определить радиус девятого светлого кольца.
- 5.9.** Найти радиус четвертого кольца в интерференционной картине от интерферометра Фабри — Перо в фокальной плоскости линзы с фокусным расстоянием $f = 0,5$ м. Расстояние между пластиинами интерферометра $d = 1$ см, длина волны монохроматического излучения $\lambda = 693$ нм.
- 5.10.** Через центр тонкой линзы с фокусным расстоянием 10 см проведен диаметр. Из линзы параллельно диаметру вырезана симметрично относительно диаметра полоска толщиной 1 мм, после чего оставшиеся части линзы плотно прижимнуты друг к другу. Определить длину волны точечного источника света,

- расположенного на расстоянии 8 см от линзы, если на экране, находящемся на расстоянии 160 см от линзы, наблюдается интерференционная картина, в которой расстояние между соседними максимумами равно 0,3 мм. При каком расстоянии между зеркалами в интерферометре Фабри — Перо можно разделить дублет линии натрия ($\lambda = 589$ нм, $\Delta\lambda = 0,6$ нм)?
- 5.12.** В интерферометре Майкельсона наблюдаются кольца равного наклона от ртутной лампы, испускающей свет с длиной волны $\lambda = 435,8$ нм. При смещении подвижного зеркала интерферометра на 47,5 мм интерференционная картина размывается. Найти ширину линии.
- 5.13.** Какое число штрихов на 1 мм должна иметь дифракционная решетка, чтобы при фотографировании дублета натрия ($\Delta\lambda = 0,6$ нм; $\lambda = 589$ нм) во втором порядке камерой с фокусным расстоянием 100 см получились на пластинке линии дублета на расстоянии 1 мм?
- 5.14.** Мыльная пленка толщиной, существенно меньшей длины волны, находится в воздухе. Темной или светлой она представляется в отраженном свете?
- 5.15.** Лучи, выходящие из пластиинки Люммера — Герке, можно считать практически скользящими вдоль поверхности. Определить дисперсионную область пластиинки с учетом дисперсии света. Показатель преломления вещества пластиинки n , толщина пластиинки h , длина волны света λ .
- 5.16.** Рассеянный монохроматический свет с $\lambda = 0,5$ мкм падает на тонкую пленку с показателем преломления $n = 1,5$. Определить толщину пленки, если известно, что угол между лучами, образующими соседние максимумы вблизи угла отражения 45° , равен 2° .
- 5.17.** В интерферометре Жамена полосы наблюдаются в фокальной плоскости объектива с фокусным расстоянием 15 см. Показатель преломления вещества пластиин $n = 1,5$, толщина пластиин 2 см. Пластиин образуют между собой малый угол $1'$. Свет падает на пластиин под углом 45° . Расстояние между полосами интерференции 5 мм. Найти длину волны света, создающего интерференционную картину.
- 5.18.** Билинза изготовлена из линзы с фокусным расстоянием $f = 15$ см, а ее половинки раздвинуты на расстояние $a = 1$ мм друг от друга. Определить расстояние между интерференционными полосами, если источником света является щель, расположенная на рас-

стоянии $L = 30$ см от билинзы, а экран от билинзы находится на расстоянии $l = 5$ м. Длина волны света $\lambda = 0,5$ мкм.

- 5.19. На каком минимальном расстоянии должны находиться щели в опыте Юнга для того, чтобы наблюдать интерференционную картину от излучения Солнца? Угловой размер Солнца $32'$. Длину волны считать равной $0,55$ мкм [см. (27.25)].

- 5.20. Бипризма Френеля с малым преломляющим углом α расположена на расстоянии L от источника и на расстоянии l от экрана. Показатель преломления вещества бипризмы n , длина волны света λ . Найти число полос интерференций.

- 5.21. Кольца Ньютона наблюдаются между двумя плосковыпуклыми линзами, касающимися друг друга своими выпуклыми поверхностями. Радиусы кривизны поверхностей R_1 и R_2 . Найти радиус m -го темного кольца в отраженном свете. Длина волны λ .

- 5.22. При измерении угловых размеров астрономического объекта с помощью звездного интерферометра обнаружено, что видимость интерференционной картины принимает последовательные минимальные размеры при увеличении расстояния между зеркалами на 15 см. Наблюдение ведется на волне длиной волны 550 мкм. Каков угловой размер объекта?

- 5.23. В пластинке Люммера — Герке наблюдается интерференция N лучей, показатель преломления стекла 1,71, длина волны $\lambda = 540$ нм. Определить наблюдаемый порядок интерференции и разрешающую способность пластиинки при этих условиях.

- 5.24. При наблюдении по схеме Юнга интерференции от Солнца на волне $\lambda = 0,6$ мкм с

изменяющимся между щелями расстоянием отмечено, что видимость достигает минимума при расстоянии между щелями 0,064 мм. Чему равен угловой размер Солнца?

- 5.25. Точечный источник (щель), освещаящий зеркало Ллойда, помещен на высоте 1 см над его плоскостью на расстоянии 5,5 м от экрана, направленного перпендикулярно плоскости зеркала и плотно прилегающего к нему с противоположной от источника света стороны. Длина зеркала 50 см. Найти расстояние между максимумами в интерференционной картине и число светлых и темных полос. Длина волны света, 0,5 мкм.

- 5.26. Острый угол при вершине бипризмы Френеля составляет $20'$, толщиной бипризмы в расчете можно пренебречь. Точечный источник расположен на оси симметрии бипризмы, перпендикулярной наибольшей стороне, на расстоянии 10 см от бипризмы, а экран перпендикулярен этой оси и находится на расстоянии 2,9 м от бипризмы. Длина волны света 0,5 мкм. Сколько белых и темных полос образуется на экране?

- 5.27. Билинза Биц с фокусным расстоянием 20 см имеет щелевую между половинками линзы 0,5 мм. Расстояние от билинзы до точечного источника (горизонтальная щель) 40 см, до экрана — 60 см. Найти расстояние между максимумами в интерференционной картине и число светлых полос на экране (диаметр билинзы достаточно велик, например больше 1 см). Длина волны $\lambda = 0,5$ мкм.

- 5.28. Найти степень когерентности для хаотического света с гауссовым распределением по частотам, описываемым формулой (10.17).

Ответы

- 5.1. $4I_0 \sin^2\left(\frac{\pi}{2} \sin \theta\right)$. 5.2. $2I_0 [1 + \sin\left(\frac{\pi}{2} \sin \theta\right)]$. 5.3. $3'36''$; $1^{\circ}12'$. 5.4. $\Delta\lambda = 0,6$ нм. 5.5. 1,000293. 5.6. $lg(n-1)/a$. 5.7. $0,59 \cdot 10^{-4}$ рад. 5.8. 5 мм. 5.9. 8,3 мм. 5.10. 0,6 мкм. 5.11. 0,29 мм. 5.12. 0,002 нм. 5.13. 600. 5.14. Темной. 5.15. $[k^2 \sqrt{n^2 - 1}/(2h)] [Nn(dn/dk) - (n^2 - 1)]^{-1}$. 5.16. 18,7 мкм. 5.17. 434 нм. 5.18. $\lambda/[l(L-f) - Lf]/(Ld) = 1,8$ мм. 5.19. 0,03 мм. 5.20. $4L k a^2 (n-1)^2 / [l(L+f)]$. 5.21. $\sqrt{ml/(1/R_1 + 1/R_2)}$. 5.22. $3,7 \cdot 10^{-6}$ рад. 5.23. $1,03 \cdot 10^4$, $6,68 \cdot 10^6$. 5.24. $32'$. 5.25. 0,14 мкм; 7 темных, 7 светлых и 1 темная центральная. 5.26. 7 светлых, 6 темных. 5.27. 0,1 мм; 13. 5.28. $\exp(-\sigma^2 t^2/2)$.

Основная идея:
интерференция излучения
от вторичных
источников участка
волнового фронта
приводит
к возникновению
дифракции.

Дифракция



На примере зон Френеля для различных ситуаций излагается основная идея анализа излучения вторичных источников.

Принцип Гюйгенса — Френеля (1818). Представление о том, что каждая точка волнового фронта является источником вторичных волн [принцип Гюйгенса, см. (8.27)], было дополнено Френелем в виде утверждения, что эти источники когерентны между собой, а испускаемые ими вторичные волны интерферируют (рис. 142). Таким образом, при анализе распространения волн необходимо принять во внимание их фазу и амплитуду, что позволяет рассматривать вопрос об интенсивности света. Для Френеля было ясно, что амплитуда вторичной волны зависит от угла между нормалью к фронту первичной волны и направлением на точку фронта вторичной волны, причем в направлении нормали амплитуда максимальна, а в перпендикулярном направлении, т. е. по касательной к исходному волновому фронту, она равна нулю. Более точно характер этой зависимости в то время не был известен.

Зоны Френеля. Типичный пример распространения пучков света конечных размеров изображен на рис. 143. Сферическая (или плоская) волна падает на непрозрачный экран с отверстием. Требуется найти распределение интенсивности света за экраном. Для решения этой задачи с помощью принципа Гюйгенса — Френеля делаются два предположения: 1) непроницаемые части экрана не являются источниками вторичных волн; 2) в отверстии точки волнового фронта являются такими же источниками вторичных волн, какими они были бы при отсутствии непроницаемых частей экрана.

Пусть A — источник сферической волны (рис. 144), S — волновой фронт в некоторый момент времени. Найдем интенсивность волны в точке B с помощью принципа Гюйгенса — Френеля. Для решения разобьем поверхность M на кольцеобразные зоны такого размера, чтобы расстояния от краев зоны до B отличались на $\lambda/2$. Обозначая M_0, M_1, M_2, \dots границы зон, запишем это условие в виде

$$M_1B - M_0B = \lambda/2,$$

$$M_2B - M_1B = \lambda/2,$$

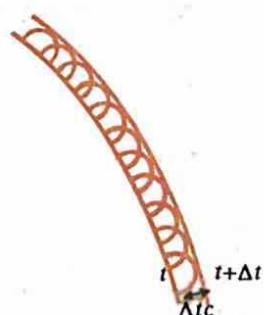
$$\dots$$

$$M_nB - M_{n-1}B = \lambda/2.$$

(31.1)

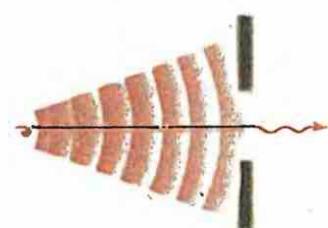
Радиусы зон находятся с помощью построения, показанного на рис. 145. Радиусы зон r_1, r_2, \dots, r_m , R — радиус кривизны фронта сферической волны, D — точка пересечения фронта с прямой AB , d_1, d_2, \dots, d_m — расстояния от D до проекции границы соответствующей зоны на прямую AB . Центральная зона называется *нулевой*. Для радиуса r_m находим уравнение

$$r_m^2 = R^2 - (R - d_m)^2 = [l + (m+1)\lambda/2]^2 - (l + d_m)^2, \quad (31.2)$$



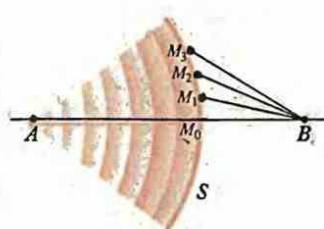
142

Иллюстрация принципа Гюйгенса



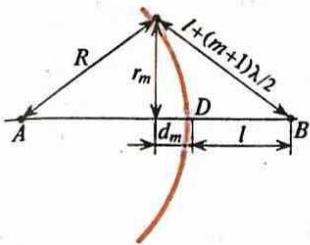
143

Распространение света через отверстие конечных размеров

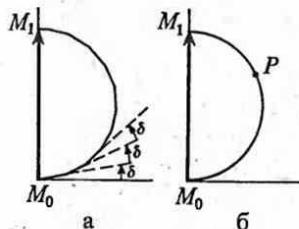


144

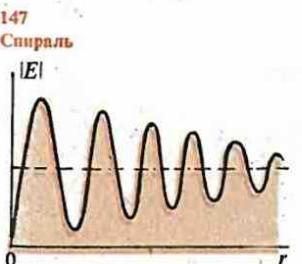
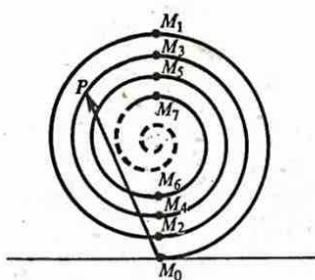
К определению зон Френеля



145
К расчету радиусов и площадей зон Френеля



146
Сложение комплексных амплитуд в нулевой зоне Френеля: при разбивке зоны условно на конечные участки (а); при непрерывном изменении фазы (б)



148
Зависимость амплитуды от радиуса отверстия

из которого с точностью до величин λ^2 следует, что

$$d_m = \frac{l(m+1)}{R+l} \frac{\lambda}{2}, \quad r_m^2 = \frac{Rl(m+1)\lambda}{R+l}. \quad (31.3)$$

Следовательно, площадь нулевой зоны

$$S_0 = \pi r_0^2 = \pi R\lambda/(R+l). \quad (31.4)$$

Для суммарной площади нулевой и первой зон получаем

$$S_{0+1} = \pi R/2\lambda/(R+l), \quad (31.5)$$

поэтому площадь первой зоны

$$S_1 = S_{(0+1)} - S_0 = \pi l\lambda/(R+l), \quad (31.6)$$

т. е. равна площади нулевой зоны. Такое же выражение [см. (31.3)] имеют площади всех остальных зон. В этом расчете пренебрегается кривизной поверхности волнового фронта и считается, что площадь кольцевой зоны на сферической поверхности волнового фронта равна площади ее проекции на плоскость, перпендикулярную прямой AB . Это не вносит в расчет существенной ошибки, если радиусы зон Френеля много меньше радиуса кривизны волнового фронта. Принимая во внимание малость длин волн, из (31.3) заключаем, что это условие хорошо выполняется для очень большого числа зон Френеля.

Графическое вычисление амплитуды. Для получения количественных результатов необходимо более строго рассчитать интерференцию вторичных волн в точке B .

Разделим каждую из зон на большое число n участков (подзон). Между началом и концом зоны фаза волн изменяется на π , и поэтому при переходе от одного участка к другому фаза изменяется на $\delta = \pi/n$. Обозначим амплитуду волны, приходящей в B от каждого участка, E_0 . Примем фазу волны, приходящей в B от нулевого участка, начинающегося в точке M_0 , за нуль. Тогда комплексная амплитуда волны в точке B от нулевой зоны с учетом интерференции равна

$$E = E_0 + E_0 e^{i\delta} + E_0 e^{i2\delta} + \dots \quad (31.7)$$

Аналитическое сложение амплитуд в (31.7) может быть выполнено графически (рис. 146, а). При увеличении числа разбиений до бесконечности ($n \rightarrow \infty$) ломаная кривая превращается в плавную (рис. 146, б). Длина $|M_0M_1|$ пропорциональна амплитуде волны в точке B (см. рис. 145), когда открыта часть нулевой зоны от центра до границы, соответствующей точке P . Длина $|M_0M_1|$ пропорциональна амплитуде при полном открытии нулевой зоны. Графическое построение амплитуды при учете вклада от последующих зон проводится аналогично. Необходимо лишь учесть, что значение E_0 при удалении от точки (т. е. увеличении фазы в экспоненциальном множителе) несколько уменьшается. Из-за этого непрерывная кривая не замыкается, а имеет вид спирали (рис. 147). Она позволяет определить амплитуду при открытии любого числа зон и их частей. Например, отрезок $|M_0P|$ на рис. 147 пропорционален амплитуде при открытии нулевой, первой, второй зон и части третьей зоны. Длины $|M_0M_1|$, $|M_0M_2|$, ... пропорциональны

210 амплитудам при открытии соответственно нулевой зоны, нулевой и первой зон и т. д. Когда открыты все зоны, амплитуда пропорциональна $|M_0 M_\infty| \approx |M_0 M_1|/2$. Это показывает, что при открытии только нулевой зоны амплитуды волн в B примерно в два раза, а интенсивность, в четыре раза больше, чем при открытии всех зон. При открытии нулевой и первой зон амплитуда пропорциональна $|M_0 M_2|$ и поэтому интенсивность очень мала. Интенсивность в точке B при непрерывном увеличении радиуса отверстия, на которое падает волна, непрерывно изменяется. С помощью спирали (рис. 147) можно построить график изменения амплитуды волны в зависимости от r (рис. 148).

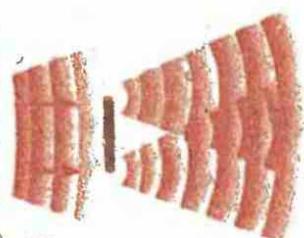
Пятно Пуассона. Если на пути световой волны стоит непрозрачный круглый экран (рис. 149), то за экраном в его тени на оси возникает светлое пятно, называемое **пятном Пуассона**. Необходимость возникновения светлого пятна очевидна из рассуждений по методу зон Френеля. Экран закрывает некоторое число зон Френеля начиная с нулевой. Однако следующие зоны после последней из закрытых создают в точке B освещенность, значение которой можно рассчитать с помощью спирали (рис. 147). Пусть, например, закрыто какое-то число зон. При не очень малых размерах экрана это число зон велико. Следовательно, точка, соответствующая радиусу экрана, лежит на спирали (рис. 147) где-то на небольшом расстоянии от M_∞ (точка P). Амплитуда волны в точке B , очевидно, пропорциональна $|PM_\infty|$. Таким образом, получается, что волна как бы огибает непрозрачный экран. Явление египания волнами препятствий называется **дифракцией**.

Интенсивность пятна Пуассона весьма слаба при больших размерах непрозрачного экрана, поскольку точка P (см. рис. 147) оказывается весьма близко к M_∞ . Кроме того, необходимо, чтобы свет обладал достаточно большой степенью когерентности, потому что в противном случае не будет происходить интерференция лучей от различных участков зон. Для наблюдения дифракции необходимо брать достаточно малые экраны. Однако при использовании лазерного излучения удается наблюдать дифракцию на сравнительно больших препятствиях и очень ярко демонстрировать это на больших экранах.

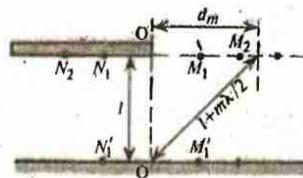
Дифракция на прямолинейном крае полубесконечного экрана. Рассмотрим плоскую волну, падающую на полубесконечный экран (рис. 150). Отсчет зон начинается от края экрана, принимая ближайшую зону за первую. Расстояние от начала до дальней границы m -й зоны обозначим d_m . Из рис. 150 видно, что с точностью до величин λ^2

$$d_m^2 = (l + m\lambda/2)^2 - l^2 \approx lm\lambda, \quad d_m = \sqrt{lm\lambda}. \quad (31.8)$$

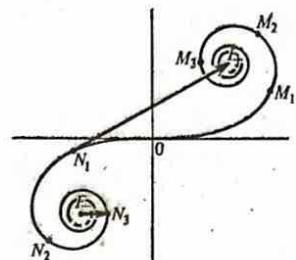
* Зоны Френеля представляют собой полосы с границами, параллельными краю экрана. Площади зон относятся как их ширины, т. е. находятся в отношении



149
Возникновение пятна Пуассона

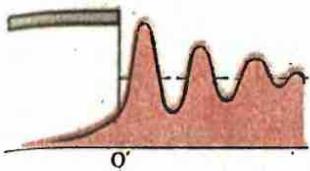


150
Расчет зон Френеля при дифракции на прямолинейном крае полу бесконечного непроницаемого экрана



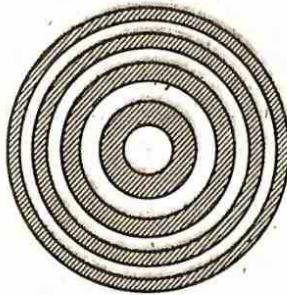
151
Сpirаль Корнио

$$d_1 : (d_2 - d_1) : (d_3 - d_2) : \dots = 1 : (\sqrt{2} - 1) : (\sqrt{3} - \sqrt{2}) : \dots \quad (31.9)$$



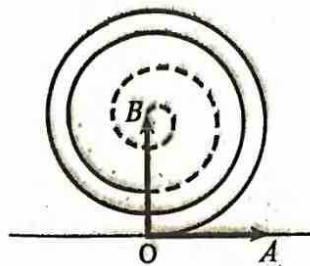
152

Распределение освещенности за полубесконечным непроницаемым экраном



153

Зонная пластинка



154

Отличие вычисленной по методу Френеля фазы волн от действительной

Поскольку $(\sqrt{2} - 1) \approx 0,41$, $(\sqrt{3} - \sqrt{2}) \approx 0,32$, видно, что площади зон Френеля при удалении от края экрана быстро убывают. Это означает, что при векторном сложении амплитуд по формуле (31.7) значения E_0 при увеличении фазы убывают значительно быстрее, чем в случае круглого отверстия. Поэтому при графическом суммировании амплитуд получается отличная от рассмотренной ранее спираль, называемая спиралью Корни.

Длины прямолинейных отрезков $|OM_1|$, $|OM_2|$, ... на рис. 150 пропорциональны освещенности в точке O , когда на рис. 150 открыты зоны, находящиеся на $|OM_1|$, $|OM_2|$, Если открыть все полупространство вправо от O , то освещенность в точке O пропорциональна $|OF_+|$. Длины прямолинейных отрезков $|ON_1|$, $|ON_2|$, ... (рис. 151) пропорциональны освещенности в точке O , когда на рис. 150 открыты зоны, находящиеся на $|ON_1|$, $|ON_2|$, Если открыть все полупространство влево от O , т. е. когда непроницаемый экран находится вправо от точки O , то освещенность в точке O пропорциональна $|OF_-|$. Если экрана нет вообще, т. е. открыто все пространство, то освещенность в точке O пропорциональна $|F-F_+|$.

Определение освещенности в других точках производится с помощью спирали Корни, считая, что точка O этой спирали соответствует рассматриваемой точке. Например, для нахождения освещенности в точке M' (см. рис. 150) считаем, что точка O спирали Корни (рис. 151) соответствует точке M' . Следовательно, край экрана на рис. 150 соответствует на спирали Корни некоторая точка слева от O , например N_1 , и поэтому освещенность дается длиной прямолинейного отрезка $|N_1 F_+|$. Для определения освещенности в точке N'_1 (см. рис. 150) считаем точку O спирали Корни соответствующей точке N'_1 . Следовательно, край экрана соответствует некоторой точке справа от точки O на спирали Корни, например точке M_1 . Следовательно, освещенность в N'_1 дается длиной прямолинейного отрезка $M_1 F_+$. Этим способом можно найти освещенность в любой точке не только для бесконечного экрана, закрывающего все полупространство, но и для любой системы экранов в виде бесконечно длинных непроницаемых полос, ребра которых параллельны друг другу. Для этого необходимо произвести векторное сложение амплитуд освещенностей от открытых участков, пропускающих волну. Амплитуда волны от каждого участка находится с помощью спирали Корни. Однако геометрическое построение при этом оказывается довольно громоздким, поэтому обычно предпочитают более эффективные аналитические методы (см. § 34). Для полубесконечного же экрана геометрическое построение с помощью спирали Корни выполнить очень просто. На рис. 152 приведен график изменения амплитуды за полубесконечным экраном.

$$r_m = [Rl/(m+1)\lambda]/(R+l) \quad (m = 0, 1, 2, \dots). \quad (31.10)$$

Закроем все нечетные зоны, оставив открытыми четные. В результате получается пластинка, называемая зонной (рис. 153). Амплитуда прошедшей через пластинку световой волны на оси в точке B может быть рассчитана с помощью спирали, изображенной на рис. 147. Амплитуда от нулевой открытой круглой зоны дается вектором $\overrightarrow{M_0 M_1}$, от второй кольцеобразной открытой зоны — вектором $\overrightarrow{M_2 M_3}$, от четвертой — $\overrightarrow{M_4 M_5}$ и т. д. Все эти векторы имеют одинаковое направление, т. е. фазы комплексных амплитуд отличаются на $2\pi m$ (m — целое число). Поэтому осуществляется интерференция волн с усилением. Следовательно, в точке B на оси происходит значительное усиление интенсивности света, т. е. в этой точке свет фокусируется. Зонная пластинка ведет себя как линза.

Будем считать, что падающие на пластинку лучи параллельны, т. е. $R = \infty$. Тогда точка на оси, в которой собираются лучи, ведет себя как фокус линзы, совпадающей по положению с зонной пластинкой, если ее фокусное расстояние $f = l$. При $R \rightarrow \infty$ формула (31.10) принимает вид

$$r_m^2 = (m+1)\lambda \quad (31.11)$$

и, следовательно, фокусное расстояние равно

$$f = l = r_m^2 / [(m+1)\lambda]. \quad (31.12)$$

Соотношение (31.10) может быть представлено в виде формулы

$$\frac{1}{R} + \frac{1}{l} = \frac{1}{r_m^2 / [(m+1)\lambda]} = \frac{1}{f}, \quad (31.13)$$

показывающей, что для источника и точки на оси за зонной пластинкой, в которой лучи собираются, соблюдается соотношение (23.27), характерное для предмета и его изображения с помощью линзы. Это означает, что зонная пластинка действует как собирающая линза. С ее помощью можно, например, формировать изображение и выполнять другие операции, характерные для линзы. Это является экспериментальным подтверждением правильности идеи зон Френеля.

Трудности метода зон Френеля. Метод зон Френеля приводит к результатам, которые хорошо согласуются с экспериментом для практически важных случаев, когда размеры препятствий многое больше длины волны. Однако метод имеет существенные недостатки:

● По своему физическому содержанию принцип Гюйгенса выражает взгляд на свет как на непрерывный процесс в пространстве. Принцип Гюйгенса эквивалентен переходу к описанию распространения света с помощью дифференциальных уравнений, которые, однако, в то время еще не были известны.

Представление Гюйгенса о том, что каждая точка волнового фронта является источником вторичных волн, было дополнено утверждением, сделанным Френелем, что эти источники когерентны между собой, а испускаемые ими вторичные волны интерферируют.

● Как качественно зависит интенсивность пятна Пуассона от расстояния до непрозрачного экрана?

Почему интенсивность в фокусах зонной пластинки максимальна для самого дальнего от пластинки фокуса? Перечислите основные трудности метода зон Френеля.

1. Он не решает вопроса о законе ослабления амплитуды вторичных волн в зависимости от направления распространения. Этую зависимость приходится постулировать, чтобы обеспечить отсутствие обратной волны.

2. Метод Френеля дает неправильную фазу волны. Фаза на фронте волны принимается по определению равной нулю. Поэтому амплитуда волны задается вектором \vec{OA} (рис. 154). Вычисленная по методу Френеля амплитуда задается вектором \vec{OB} , т. е. вычисленная по методу Френеля фаза отличается от фактической фазы волны на $\pi/2$. Хотя для многих практически важных явлений, зависящих от модуля амплитуды, эта разница в фазах несущественна, она все же с теоретической точки зрения имеет принципиальный характер и должна быть объяснена. Это удалось сделать лишь в более строгой теории дифракции, основанной на интеграле Кирхгофа.

§ 32 | Приближение Кирхгофа

Выводятся формулы теории дифракции в приближении Кирхгофа.

Формула Грина. Из теоремы Гаусса — Остроградского

$$\int_V \operatorname{div} \mathbf{A} dV = \oint_S \mathbf{A} \cdot d\mathbf{S} \quad (32.1)$$

положив

$$\mathbf{A} = \Phi \operatorname{grad} G - G \operatorname{grad} \Phi, \quad (32.2)$$

находим

$$\int_V (\Phi \nabla^2 G - G \nabla^2 \Phi) dV = \oint_S \left(\Phi \frac{\partial G}{\partial n} - G \frac{\partial \Phi}{\partial n} \right) dS. \quad (32.3)$$

Здесь $(\partial G / \partial n) dS = \operatorname{grad} G \cdot d\mathbf{S}$, $(\partial \Phi / \partial n) dS = \operatorname{grad} \Phi \cdot d\mathbf{S}$, т. е. $\partial G / \partial n$ и $\partial \Phi / \partial n$ являются производными по длине параллельно внешней нормали к замкнутой поверхности S ; V — объем, ограниченный поверхностью S .

Формула (32.3) называется второй формулой Грина. Она применима, когда функции Φ и G , их первые и вторые частные производные непрерывны внутри объема V и на поверхности S .

Теорема Гельмгольца — Кирхгофа. Распространение световых волн в свободном пространстве описывается уравнением (2.12). Полагая в случае монохроматических волн

$$\Phi(\mathbf{r}, t) = \Phi(\mathbf{r}) e^{-i\omega t}, \quad (32.4)$$

находим для зависящей только от пространственных координат амплитуды уравнения

$$\nabla^2 \Phi + k^2 \Phi = 0, \quad (32.5)$$

где $k^2 = \omega^2/c^2 = (2\pi/\lambda)^2$, λ — длина волны.

Применим формулу (32.3) к объему V , ограниченному поверхностью S (рис. 155). Обозначим P_0 фиксированную точку внутри объема, а P_1 — переменную точку (любую точку пространства, отличную от P_0). Принимая P_0 за начало отсчета, координаты точки P_1 можно характеризовать радиусом-вектором \mathbf{r}_{01} . Непосредственной проверкой убеждаемся, что функция

$$G(P_1) = \exp(ikr_{01})/r_{01} \quad (32.6)$$

удовлетворяет уравнению (32.5) для всех точек, отличных от P_0 . В обозначении $G(P_1)$ в (32.6) учтено, что точка P_0 принята за начало координат и нет необходимости ее указывать, а при подстановке $G(P_1)$ в (32.5) производные вычисляются по координатам точки P_1 . Функция Φ , описывающая электромагнитную волну, непрерывна во всех точках.

Поскольку в точке $P_1 = P_0$ функция $G(P_1)$ обращается в бесконечность, а ее производные терпят разрыв, к ней нельзя применить формулу Грина для всего объема V . Поэтому окружим точку P_0 небольшой сферой, радиусом ϵ с центром в точке P_0 . Обозначим площадь поверхности сферы S_ϵ , а ограничивающий ее объем — V_ϵ . Вне объема V_ϵ функция $G(P_1)$ удовлетворяет всем условиям применимости (32.3). Учитывая, что

$$\int_{V - V_\epsilon} (\Phi \nabla^2 G - G \nabla^2 \Phi) dV = \int_{V - V_\epsilon} (-k^2 \Phi G + k^2 \Phi G) dV = 0, \quad (32.7)$$

получаем

$$\int_{S + S_\epsilon} (\Phi \frac{\partial G}{\partial n} - G \frac{\partial \Phi}{\partial n}) dS = 0. \quad (32.8)$$

Заметим, что внешняя нормаль n на поверхности S_ϵ , ограничивающей объем $V' = V - V_\epsilon$, направлена внутрь объема V_ϵ (рис. 155). Из (32.7) и (32.3) следует, что

$$-\int_{S_\epsilon} (\Phi \frac{\partial G}{\partial n} - G \frac{\partial \Phi}{\partial n}) dS = \int_S (\Phi \frac{\partial G}{\partial n} - G \frac{\partial \Phi}{\partial n}) dS. \quad (32.9)$$

Для точек P_1 на поверхности S имеем

$$\frac{\partial r_{01}}{\partial n} = \text{grad}_1 r_{01} \cdot n = \cos(\hat{n}, \mathbf{r}_{01}) \quad (32.10)$$

и, следовательно,

$$\partial G / \partial n = \cos(\hat{n}, \mathbf{r}_{01}) (ik - 1/r_{01}) G. \quad (32.11)$$

В (32.10) индекс 1 в обозначении операции градиента показывает, что градиент вычисляется по координатам точки P_1 , т. е. $\text{grad}_1 r_{01} = \mathbf{r}_{01}/r_{01}$. Очевидно, что $\text{grad}_0 r_{01} = \mathbf{r}_{01}/r_{01}$ и, следовательно, $\text{grad}_1 r_{01} = -\text{grad}_0 r_{01}$. Для точек P_1 на поверхности S_ϵ

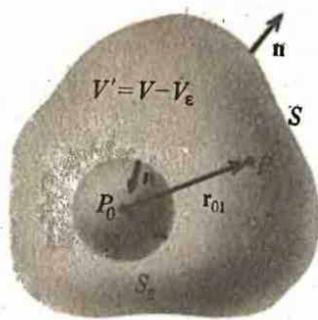
$$G(P_1) = e^{ik\epsilon}/\epsilon, \quad \partial G / \partial n = -\partial G / \partial \epsilon = (1/\epsilon - ik) G. \quad (32.12)$$

При $\epsilon \rightarrow 0$, поскольку Φ и ее производные непрерывны, получаем

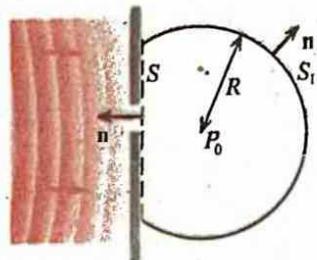
$$\begin{aligned} \int_{S_\epsilon \rightarrow 0} (G \frac{\partial \Phi}{\partial n} - \Phi \frac{\partial G}{\partial n}) dS &= \left\{ \left[\frac{e^{ik\epsilon}}{\epsilon} \frac{\partial \Phi}{\partial n} - \right. \right. \\ &\quad \left. \left. - \Phi \frac{e^{ik\epsilon}}{\epsilon} \left(\frac{1}{\epsilon} - ik \right) \right] 4\pi\epsilon^2 \right\}_{\epsilon \rightarrow 0} = -4\pi\Phi(P_0). \end{aligned} \quad (32.13)$$

Поэтому [см. (32.9)]

$$\Phi(P_0) = \frac{1}{4\pi} \int_S \left[\frac{e^{ikr_{01}}}{r_{01}} \frac{\partial \Phi}{\partial n} - \Phi \frac{\partial}{\partial n} \left(\frac{e^{ikr_{01}}}{r_{01}} \right) \right] dS. \quad (32.14)$$



К выводу интегральной теоремы Кирхгофа — Гельмгольца



К выводу условия излучения

Эта формула позволяет вычислить значение функции Φ в любой точке внутри объема, если известны значения функции и ее производной по нормали на поверхности, ограничивающей этот объем. Она называется интегральной теоремой Гельмгольца — Кирхгофа и является основой скалярной теории дифракции.

Функция Φ в (32.14) должна удовлетворять волновому уравнению. Поэтому значения Φ и $\partial\Phi/\partial n$ на поверхности нельзя задать произвольно. Следовательно, (32.14) не является формулой, по которой можно вычислить $\Phi(P_0)$, а представляет собой интегральное уравнение относительно Φ и кажущаяся простота вычисления $\Phi(P_0)$ обманчива.

Физический смысл функции $G = \exp(ikr_{01})/r_{01}$, входящей в (32.14), ясен из сопоставления этой функции с правой частью (2.43). Видно, что G представляет зависящую от координат часть расходящейся из точки P_1 сферической волны.

Условие излучения. Рассмотрим типичную ситуацию, в которой изучается дифракция (рис. 156). Имеется большой (бесконечный) экран S с одним или несколькими отверстиями, на который падает свет, и наблюдается распределение интенсивности света, прошедшего через отверстия, т. е. дифракция в пространстве за экраном. Значение $\Phi(P_0)$ может быть выражено по формуле (32.14), причем ограничивающая объем поверхность состоит из двух частей: плоской поверхности S и части поверхности S_1 сферы радиусом R с центром в точке P_0 . Интеграл (32.14) представляется в виде суммы двух интегралов — по поверхностям S и S_1 . Оценим интеграл по S_1 при $R \rightarrow \infty$. Так как для точек поверхности S_1

$$\frac{\partial}{\partial n} \left(\frac{e^{ikr_{01}}}{r_{01}} \right) = \frac{\partial}{\partial R} \left(\frac{e^{ikR}}{R} \right) = \left(ik - \frac{1}{R} \right) \frac{e^{ikR}}{R}, \quad (32.15)$$

то

$$Q = \int_{S_1} \left[\frac{e^{ikr_{01}}}{r_{01}} \frac{\partial\Phi}{\partial n} - \Phi \frac{\partial}{\partial n} \left(\frac{e^{ikr_{01}}}{r_{01}} \right) \right] dS = \frac{e^{ikR}}{R} \int_{S_1} \left[\frac{\partial\Phi}{\partial n} - \Phi \left(ik - \frac{1}{R} \right) \right] dS. \quad (32.16)$$

Обозначим Ω телесный угол, под которым видна поверхность S_1 из точки P_0 . Учитывая, что в (32.16) $dS = R^2 d\Omega$, находим

$$Q = e^{ikR} \int_{S_1} R \left[\frac{\partial\Phi}{\partial n} - \left(ik - \frac{1}{R} \right) \Phi \right] d\Omega. \quad (32.17)$$

Этот интеграл равен нулю при $R \rightarrow \infty$, если выполняется условие излучения

$$\lim_{R \rightarrow \infty} R \left[\frac{\partial\Phi}{\partial n} - \left(ik - \frac{1}{R} \right) \Phi \right] = 0. \quad (32.18)$$

При этом формула (32.14) для $\Phi(P_0)$ в ситуации, изображенной на рис. 156, не изменяет своего вида, но под S понимается не вся замкнутая поверхность, охватывающая точку P_0 , а только плоская бесконечная поверхность, обозначенная на рис. 156 также S .

Можно указать на некоторые соображения, показывающие, что условие (32.18) в реальных физических ситуациях действительно выполняется. При конечном размере отверстия и конечном расстоянии между отверстием и точкой P_0 функция Φ должна при $R \rightarrow \infty$ описывать сферическую волну, т. е. должна иметь вид

$$\Phi(R) \sim e^{ikR}/R. \quad (32.19)$$

Так как

$$\frac{\partial\Phi}{\partial n} = \frac{\partial\Phi}{\partial R} = \left(ik - \frac{1}{R} \right) \Phi, \quad (32.20)$$

то подынтегральное выражение в (32.17), а следовательно, и интеграл действительно равны нулю, т. е. условие излучения (32.18) выполняется.

Приближение Кирхгофа. Для того чтобы формулу (32.14) использовать не как интегральное уравнение для нахождения Φ , а как формулу для вычисления $\Phi(P_0)$ по известным значениям Φ и $\partial\Phi/\partial n$ в точках плоского экрана, Кирхгоф предложил следующие правила для определения их значений в плоскости экрана:

1. На отверстиях Φ и $\partial\Phi/\partial n$ имеют такие же значения, какие они имели бы при отсутствии непрозрачных частей экрана.

2. На непрозрачных частях экрана $\Phi = 0$, $\partial\Phi/\partial n = 0$.

Выбор граничных условий в соответствии с этими правилами приводит к решению задач дифракции в приближении Кирхгофа.

Границные условия Кирхгофа никогда точно не выполняются: а) на краях отверстий должны соблюдаться определенные граничные условия, которые в принципе можно найти в соответствии с электромагнитной теорией света; б) за экраном не может быть резкой тени, т. е. скачкообразного обращения Φ в нуль.

Однако при линейных размерах отверстий, много больших длины волны, краевыми эффектами можно пренебречь и граничные условия Кирхгофа являются хорошим приближением к действительности.

Оптическое приближение. Поскольку в видимом диапазоне $\lambda \approx 1$ нм, практически во всех представляющих интерес случаях соблюдается условие

$$1/r_0 \ll k, \quad (32.21)$$

называемое условием оптического приближения. При его выполнении можно положить

$$\frac{\partial}{\partial n} \left(\frac{e^{ikr_{01}}}{r_{01}} \right) = \cos(\hat{n}, \hat{r}_{01}) \left(ik - \frac{1}{r_{01}} \right) \frac{e^{ikr_{01}}}{r_{01}} \approx ik \cos(\hat{n}, \hat{r}_{01}) e^{ikr_{01}} / r_{01} \quad (32.22)$$

и представить формулу (32.14) в виде

$$\Phi(P_0) = \frac{1}{4\pi} \int_S \frac{e^{ikr_{01}}}{r_{01}} \left[\frac{\partial\Phi}{\partial n} - ik\Phi \cos(\hat{n}, \hat{r}_{01}) \right] dS. \quad (32.23)$$

Формула дифракции Френеля — Кирхгофа. Пусть на отверстие падает сферическая волна, исходящая из точки P_2 (рис. 157):

$$\Phi(P_1) = A e^{ikr_{12}} / r_{12}. \quad (32.24)$$

Учитывая, что

$$\partial r_{12} / \partial n = \cos(\hat{n}, \hat{r}_{12}) = -\cos(\hat{n}, \hat{r}_{12}), \quad (32.25)$$

находим в оптическом приближении

$$\partial\Phi/\partial n = -ik \cos(\hat{n}, \hat{r}_{12}) e^{ikr_{12}} / r_{12}. \quad (32.26)$$

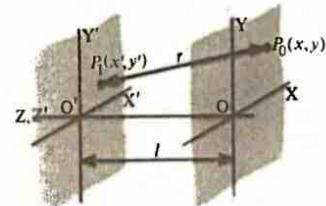
Следовательно, формула (32.23) принимает вид

$$\Phi(P_0) = -A \frac{ik}{4\pi} \int_{S_0} \frac{e^{ik(r_{01} + r_{12})}}{r_{01} r_{12}} [\cos(\hat{n}, \hat{r}_{12}) + \cos(\hat{n}, \hat{r}_{01})] dS \quad (32.27)$$

(S_0 — площадь отверстия), поскольку подынтегральное выражение на непрозрачных частях экрана равно нулю. Равенство (32.27) называется **формулой дифракции Френеля — Кирхгофа**.

Теорема взаимности Гельмгольца. Сумма косинусов в подынтегральном выражении (32.27) не изменяется, если поменять местами точку наблюдения и точечный источник. Это означает, что точечный источник, помещенный в P_2 , дает в точке P_0 такой же эффект, как и эффект, создаваемый в точке P_2 точечным источником равной интенсивности, помещенным в P_0 . Это утверждение составляет содержание теоремы взаимности Гельмгольца.

157
К выводу формулы дифракции
Френеля — Кирхгофа



158
Расположение систем координат
в плоскостях источника и дифракционной картины

Вторичные источники. Перепишем формулу (32.27) в виде

$$\Phi(P_0) = \int_{S_0} \Phi'(P_1) \frac{e^{ikr_{01}}}{r_{01}} dS, \quad (32.28)$$

где

$$\Phi'(P_1) = -A \frac{ik}{4\pi} \frac{e^{ikr_{12}}}{r_{12}} [\cos(\hat{n}, \hat{r}_{12}) + \cos(\hat{n}, \hat{r}_{01})]. \quad (32.29)$$

В соответствии с принципом Гюйгенса — Френеля поле в точке P_0 создается вторичными источниками на отверстии S_0 , интенсивность которых характеризуется формулой (32.29). Видно, что отличие этих источников от волны $A \exp(ikr_{12})/r_{12}$ заключается в следующем: 1) амплитуда вторичной волны отличается от амплитуды падающей волны множителем $k/(4\pi)$; 2) зависимость амплитуды вторичной волны от направления распространения дается множителем $\cos(\hat{n}, \hat{r}_{12}) + \cos(\hat{n}, \hat{r}_{01})$, который отличается от множителя, предлагавшегося Френелем; 3) фаза вторичного источника отличается от фазы падающей волны на $\pi/2$ ввиду наличия множителя $-i$. Тем самым разрешается трудность с фазой волны, указанная в конце § 31 в связи с методом зон Френеля.

Таким образом, в приближении Кирхгофа в рамках электромагнитной теории света удается преодолеть трудности метода зон Френеля и ответить на вопросы, которые теория Френеля была не в состоянии разрешить (амплитуда вторичных источников, характер зависимости амплитуды излучения вторичных источников от направления, фаза).

Приближение Френеля. В типичных условиях дифракционная картина наблюдается в плоскости, параллельной экрану с отверстиями. Плоскость, в которой наблюдается дифракционная картина, будем называть плоскостью дифракционной картины, а другую плоскость — плоскостью источников. В каждой плоскости введем прямоугольные декартовы системы координат, оси X , Y которых параллельны, а оси Z совпадают, причем положительное направление оси Z согласуется по обычному правилу с направлением осей X , Y (рис. 158). Единичный вектор нормали [см. (32.27)] направлен в сторону, из которой приходит излучение. Функция $e^{ikr_{01}}/r_{01}$ описывает волну, движущуюся к точке P_0 . Обозначим (x, y) координаты точки P_0 в плоскости дифракционной картины, (x', y') — координаты точки P_1 интегрирования в плоскости источников, $dS' = dx'dy'$ — элемент

- Применение формулы Грина не дает решения задачи о волновом поле, а дает интегральное уравнение для определения волнового поля.

Применимость приближения Кирхгофа обусловливается малостью длины световых волн по сравнению с линейными размерами экрана и отверстий в экранах.

- При каких условиях можно применять формулу Грина? Каков физический смысл условия излучения? Каковы условия применимости приближения Френеля?

площади на поверхности источников, а

$$\Psi(x', y') = A e^{ikr_{12}} / r_{12} \quad (32.30)$$

— амплитуду источников.

Тогда [см. (32.27)]

$$\Phi(x, y) = \frac{-ik}{4\pi} \int_S \Psi(x', y') \frac{e^{ikr}}{r} [\cos(\hat{n}, \hat{r}_{12}) + \cos(\hat{n}, \hat{r}_{01})] dS, \quad (32.31)$$

где

$$r = \sqrt{l^2 + (x - x')^2 + (y - y')^2}, \quad (32.32)$$

l — расстояние между плоскостями. Член с косинусами в подынтегральном выражении является медленно меняющейся функцией по сравнению с быстро осциллирующим множителем $\exp(ikr)$. Он не оказывает практического влияния на картину интерференции, не изменяя ее видимости, и лишь слабо влияет на среднюю яркость. Кроме того, в большинстве практически важных случаев углы (\hat{n}, \hat{r}_{12}) и (\hat{n}, \hat{r}_{01}) изменяются в небольших пределах вблизи нулевого значения. В этом предположении, называемом приближением малых углов, без существенного искажения результатов можно косинусы принять равными единице:

$$\cos(\hat{n}, \hat{r}_{12}) = 1, \quad \cos(\hat{n}, \hat{r}_{01}) = 1 \quad (32.33)$$

и представить формулу (32.31) в виде, удобном для применений:

$$\Phi(x, y) = \frac{k}{2\pi i} \int_S \Psi(x', y') (e^{ikr}/r) dx' dy'. \quad (32.34)$$

Вынос $\cos(\hat{n}, \hat{r}_{12}) + \cos(\hat{n}, \hat{r}_{01})$ из-под знака интеграла с использованием равенств (32.33) возможен благодаря тому, что эта функция меняется очень медленно по сравнению с быстро осциллирующим экспоненциальным множителем $\exp(ikr)$. Строго говоря, при выносе этой функции за знак интеграла необходимо заменить ее некоторым значением по теореме о среднем. Однако точное значение этой величины несущественно, поскольку, будучи множителем, она не оказывает влияния на видимость интерференционной картины и играет роль масштабного множителя. Расчет точного значения интенсивностей в дифференциальной картине в подавляющем большинстве случаев не представляет интереса. Поэтому целесообразно не загромождать формулы несущественными множителями и воспользоваться равенствами (32.33).

Дальнейшее упрощение формулы производится также в предположении малых углов, которое математически формулируется в виде неравенств

$$|x - x'| \ll 1, \quad |y - y'| \ll 1. \quad (32.35)$$

При выполнении условий (32.35) выражение (32.32) можно разложить в ряд по (32.35) и ограничиться членами второго порядка:

$$r = l[1 + (x - x')^2/l^2 + (y - y')^2/l^2]^{1/2} = l + \frac{(x - x')^2 + (y - y')^2}{2l}. \quad (32.36)$$

Подставляя разложение (32.36) в (32.34), находим

$$\Phi(x, y) = \frac{k}{2\pi i} \frac{e^{ikl}}{l} \int_S \Psi(x', y') \exp\left\{\frac{ik[(x - x')^2 + (y - y')^2]}{2l}\right\} dx' dy', \quad (32.37)$$

где медленно изменяющаяся величина $r \approx l$ в знаменателе вынесена из-под знака интеграла, поскольку она не оказывает влияния на видимость интерференционной картины, а лишь очень слабо меняет ее общую яркость.

В большинстве практически важных случаев для описания дифракции достаточно пользоваться приближением (32.37), называемым **приближением Френеля**. Дифракция, рассматриваемая в этом приближении, называется **дифракцией Френеля**.

При определенных условиях возможны дальнейшее упрощение формулы (32.37) и переход к приближению **Фраунгофера**.

§ 33 | Дифракция Фраунгофера

Рассматривается дифракция в дальней зоне и обсуждается критерий дальности. Изучается дифракционная решетка и обсуждается дифракция на периодических структурах и неоднородностях среды.

Область дифракции Фраунгофера. Представляя показатель экспоненты в (32.37) в виде

$$\frac{ik[(x-x')^2 + (y-y')^2]}{2l} = \frac{ik(x^2+y^2)}{2l} + \frac{ik(x'^2+y'^2)}{2l} - \frac{ik(xx'+yy')}{l}, \quad (33.1)$$

получаем

$$\Phi(x, y) = \frac{k}{2\pi i} \frac{e^{ik/l}}{l} \exp\left[\frac{-ik(x^2+y^2)}{2l}\right] \times \\ \times \int_S \Psi(x', y') \exp\left[\frac{-ik(x'^2+y'^2)}{2l}\right] \exp\left[-\frac{ik(xx'+yy')}{l}\right] dx' dy'. \quad (33.2)$$

Распределение интенсивности в дифракционной картине определяется квадратом модуля $\Phi(x, y)$. Следовательно, экспоненциальные множители перед интегралом в (33.2) никакого влияния на распределение интенсивности в дифракционной картине не оказывают, поскольку по модулю равны единице. Интегрирование в (33.2) подразумевается по всей площади S' плоскости, учитывая, что $\Psi(x', y')=0$ в точках непрозрачных частей экрана. Другими словами, интегрирование по x' и y' производится в пределах от $-\infty$ до ∞ . Поэтому $\Phi(x, y)$ с точностью до множителей, один из которых фазовый, зависящий от (x, y) , является образом Фурье-функции:

$$\Psi(x', y') \exp[ik(x'^2+y'^2)/(2l)] \dots \quad (33.3)$$

Изучение дифракции в математическом смысле сводится к применению теории преобразований Фурье.

Наличие экспоненциального множителя в (33.3) несколько усложняет выполнение преобразования Фурье и тем самым затрудняет анализ дифракционной картины. Простейший случай дифракции осуществляется при устранении этого множителя. Если отверстие достаточно мало, а расстояние до плоскости дифракционной картины достаточно велико ($l \rightarrow \infty$), то

$$\exp[ik(x'^2+y'^2)/(2l)] \rightarrow 1. \quad (33.4)$$

Дифракция при выполнении этого условия называется **фраунгоферовой** или просто **дифракцией Фраунгофера**.

Однако не только при бесконечных l можно пренебречь экспоненциальным членом в (33.3). Достаточно, чтобы этот член не осцилировал, т. е. показатель экспоненты не превосходил

220 $\pi/2$ или даже был $\pi/2$. Например, $\cos(\pi/4) = \sin(\pi/4) = 0,7$ и в качестве критерия возможности пренебрежения влиянием экспоненциального множителя в (33.3) на дифракцию можно принять, например, условие

$$\frac{k(x'^2 + y'^2)}{2l} = \frac{k\rho'^2}{2l} < \frac{\pi}{4}, \quad (33.5)$$

где

$$\rho'^2 = x'^2 + y'^2 \quad (33.6)$$

— максимальное расстояние от центра до края отверстия, на котором происходит дифракция. Из (33.6) заключаем, что при расстояниях

$$l > l_{\min} = 2k\rho'^2/\pi = 4\rho'^2/\lambda \quad (33.7)$$

наблюдается дифракция Фраунгофера. Другими словами, область дифракции Фраунгофера простирается от бесконечности до некоторого минимального расстояния [см. (33.7)]. Например, при $\lambda \approx 0,5$ мкм, $\rho' = 1$ см = 10^{-2} м находим $l_{\min} = 800$ м, т. е. дифракцию Фраунгофера можно наблюдать начиная с очень больших расстояний. Однако уже при $\rho' = 1$ мм = 10^{-3} м $l_{\min} = 8$ м, что ближе к лабораторным условиям. При $\rho' = 0,1$ мм получается $l_{\min} = 8$ см и дифракция Фраунгофера может изучаться совсем в небольшой области пространства.

Множитель перед интегралом в (33.2) не оказывает влияния на распределение интенсивности в дифракционной картине. Для упрощения написания формул дифракции Фраунгофера целесообразно приравнять его к единице и не выписывать. Тогда формула (33.2) принимает вид

$$\Phi(x, y) = \int_S \Psi(x', y') \exp \left[-\frac{ik(xx' + yy')}{l} \right] dx' dy'. \quad (33.8)$$

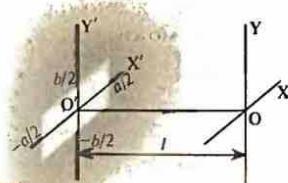
При таком написании формулы следует помнить, что она может служить для вычисления относительных, а не абсолютных величин интенсивностей в дифракционной картине. Кроме того, в ней Φ и Ψ имеют различные размерности. Например, если Φ — амплитуда, то Ψ имеет размерность амплитуды, деленной на площадь, поскольку левая и правая части (33.8) имеют одинаковую размерность.

Из (33.8) видно, что дифракционная картина зависит от x/l и y/l . Это означает, что дифракционная картина увеличивается подобно самой себе пропорционально расстоянию. Принимая во внимание, что в чистом виде дифракция Фраунгофера осуществляется на бесконечности ($l \rightarrow \infty$), заключаем, что она является дифракцией в параллельных лучах. Это обстоятельство обосновывает для дифракции Фраунгофера название дифракции в параллельных лучах.

Дифракцию Фраунгофера в области, удовлетворяющей условию (33.7), можно наблюдать на экране без всяких дополнительных устройств. Однако проще ее наблюдать в фокальной плоскости собирающей линзы, узловатченной на пути дифрагированных лучей.

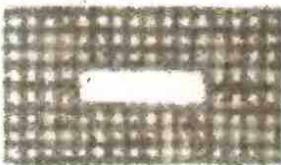
Дифракция на прямоугольном отверстии. Считаем, что начало координат расположено в центре прямоугольного отверстия со сторонами a и b (рис. 159). На отверстие слева падает плоская волна вдоль оси Z . На отверстии фаза и амплитуда плоской волны постоянны. Обозначим комплексную амплитуду волны на отверстии A_0 . Принимая во внимание, что Ψ' в (33.8) имеет размерность амплитуды, деленной на площадь, полагаем

$$\Psi' = A_0/(ab). \quad (33.9)$$



159

К расчету дифракции на прямоугольном отверстии



160

Дифракция от прямоугольного отверстия

Тогда [см. формулу (33.8)]

$$A(x, y) = (A_0/(ab)) \int_{-a/2}^{a/2} \exp(-ikxx'/l) dx' \int_{-b/2}^{b/2} \exp(-ikyy'/l) dy', \quad (33.10)$$

где $A(x, y) = \Phi(x, y)$ — амплитуда. Элементарное интегрирование в (33.10) приводит к формуле

$$A = A_0 (\sin \alpha/\alpha) (\sin \beta/\beta), \quad (33.11)$$

где

$$\alpha = kax/(2l), \quad \beta = kby/(2l). \quad (33.12)$$

Наблюдаемая интенсивность пропорциональна $|A|^2$ и с точностью до постоянного множителя равна

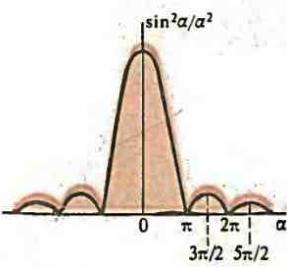
$$I(x, y) = |A_0|^2 \frac{\sin^2 \alpha}{\alpha^2} \frac{\sin^2 \beta}{\beta^2}. \quad (33.13)$$

Полосы нулевой интенсивности определяются условиями $\sin \alpha = 0, \sin \beta = 0$ и представляют собой прямые линии, параллельные осям X и Y и образующие систему прямоугольников, внутри которых интенсивность отлична от нуля (рис. 160). Зависимость множителя $\sin^2 \alpha/\alpha^2$ от α показана на рис. 161. Максимумы, кроме центрального, достигаются при значениях $\alpha = 3\pi/2, 5\pi/2$ и т. д. Поэтому интенсивности в последовательных максимумах относятся как $1 : [2/(3\pi)]^2 : [2/(5\pi)]^2 : \dots = 1 : 0,04 : 0,016 : \dots$, т. е. интенсивности убывают очень быстро. Поэтому основная энергия прошедшего через отверстие света сосредоточена в центральном светлом прямоугольном пятне.

Из (33.12) следует, что α и β зависят только от отношений x/l и y/l . Это означает, что размеры дифракционной картины растут пропорционально расстоянию от отверстия, т. е. сама дифракционная картина из центра отверстия видна под постоянным углом (рис. 162). По смыслу приближения Фраунгофера этот угол считается малым. Поэтому α и β можно преобразовать к виду

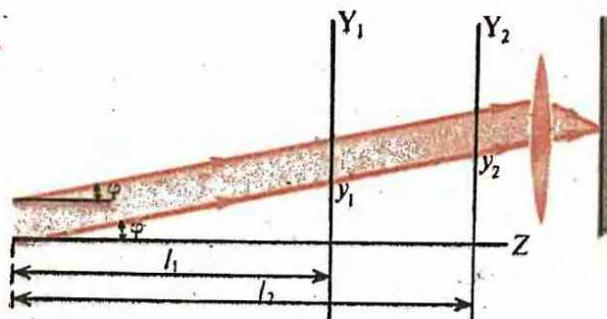
$$\begin{aligned} \alpha &= kax/(2l) = (ka/2) \operatorname{tg} \phi_x = (ka/2) \sin \phi_x, \\ \beta &= (ka/2) \sin \phi_y, \end{aligned} \quad (33.14)$$

где $\operatorname{tg} \phi_x = \sin \phi_x \approx x/l$, $\sin \phi_y \approx y/l$ из-за малости углов ϕ_x и ϕ_y . Запись (33.14) показывает, что дифракционная картина действительно зависит только от угла между лучом и осью Z , т. е. $A(x, y) = A(\phi_x, \phi_y)$. Например, на рис. 162 изображены лучи в плоскости $X = 0$. В этом случае $\phi_x = 0$, $\phi_y = \phi$ и $A = A(\phi)$. Следовательно, дифрагированные лучи, исходящие из различных точек отверстия и образующие в результате интерференции на бесконечности точку дифракционной картины в направлении угла ϕ , параллельны друг другу и направлены под углом ϕ к оси. Чтобы образовать интерференционную картину на конечном расстоянии, надо на конечном расстоянии свести параллельные лучи в точку. Для этого на их пути нужно поместить



161

Ход функции $\sin^2 \alpha / \alpha^2$



тить собирающую линзу, в фокальной плоскости которой и образуется дифракционная картина.

Дифракция на щели. При увеличении одного из размеров отверстия период чередования дифракционных полос, перпендикулярных длинной стороне, уменьшается. При достаточно большом размере длинной стороны отверстия этот период становится настолько малым, что глаз не в состоянии его различать. В дифракционной картине остаются лишь полосы, параллельные длинной стороне отверстия, которые образуют дифракционную картину от щели (рис. 163, 164). Формула (33.11) может в этом случае быть записана в виде

$$A_\varphi = A_0 \sin \beta / \beta, \quad (33.15)$$

где $\beta = kb y / (2) = kb \sin \varphi / 2$. Полезно формулу (33.15) получить методом зон Френеля. Примем за начало отсчета фаз дифрагированных волн фазу волны в средней точке щели. Волна от точки y' (рис. 165) распространяется в направлении, определяемом углом φ , и приобретает по сравнению с волной от точки O разность фаз $\delta = k\Delta = ky' \sin \varphi$. Амплитуда волны, приходящаяся на элемент щели шириной dy' , равна, очевидно, $A_0 dy' / b$. Поэтому при суперпозиции этих параллельных волн на бесконечности комплексная амплитуда равна

$$A_\varphi = (A_0/b) \int_{-b/2}^{b/2} e^{-iky' \sin \varphi} dy' = A_0 \frac{\sin \beta}{\beta}, \quad \beta = kb \sin \varphi / 2 \quad (33.16)$$

[см. (33.15)]. Чтобы осуществить дифракцию на конечном расстоянии, надо воспользоваться собирающей линзой.

Угол $\Delta\varphi$, под которым из центра щели видна первая темная полоса, определяется из условия

$$\beta = kb \sin \Delta\varphi / 2 = \pi. \quad (33.17)$$

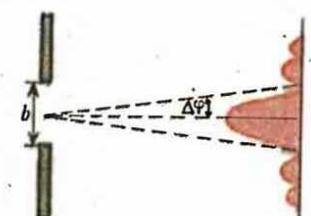
Считая $\Delta\varphi$ малым и полагая $\sin \Delta\varphi \approx \Delta\varphi$, находим, что угол, под которым из центра щели видна центральная светлая полоса, равен

$$\theta_0 = 2\Delta\varphi = 2\lambda/b. \quad (33.18)$$

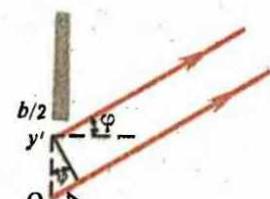
Дифракция Фраунгофера — дифракция в параллельных лучах



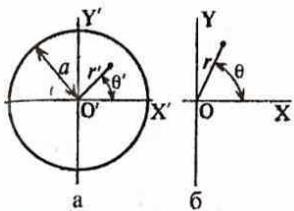
Дифракция на щели (полосы параллельны краям щели)



Распределение интенсивности при дифракции на щели

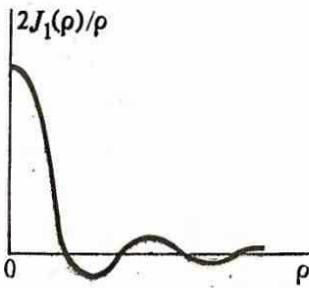


К расчету дифракции на щели методом зон Френеля



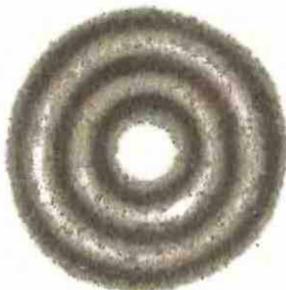
166

Полярные системы координат в плоскостях источников и дифракционной картины



167

График функции $J_1(\rho)/\rho$, определяющей картину дифракции от круглого отверстия



168

Дифракция от круглого отверстия

При уменьшении ширины щели угловые размеры светлой полосы увеличиваются и при $b \rightarrow \lambda$ интенсивность плавно уменьшается от центра к периферии без каких-либо колебаний.

Дифракция на круглом отверстии. Обозначим a радиус круга с центром в начале координат плоскости X' , Y (рис. 166). Расчет удобно вести как в плоскости источников, так и в плоскости дифракционной картины в полярных координатах:

$$x' = r' \cos \theta', \quad x = r \cos \theta, \quad y' = r' \sin \theta', \quad y = r \sin \theta. \quad (33.19)$$

Учитывая, что

$$dx'dy' = r' dr' d\theta', \quad \sin \theta \sin \theta' + \cos \theta \cos \theta' = \\ = \cos (\theta' - \theta), \quad (33.20)$$

запишем формулу (33.8) в виде

$$A(r, \theta) = (A_0/(\pi a^2)) \int_0^{2\pi} r' dr' \int_0^{2\pi} \exp [ikr'(r/l) \cos(\theta - \theta')] d\theta', \quad (33.21)$$

где считается, что на круглое отверстие падает плоская волна, фронт которой параллелен плоскости отверстия, и, следовательно, $\Psi = A_0/(\pi a^2)$. Внутренний интеграл равен

$$\int_0^{2\pi} \exp [ikr'(r/l) \cos(\theta - \theta')] d\theta' = \int_0^{2\pi} \exp [ikr' \cos \theta'] d\theta' = 2\pi J_0(\eta),$$

где $J_0(\eta)$ — функция Бесселя нулевого порядка, $\eta = krr'/l$.

Тогда [см. (33.21)]

$$A = (2A_0/a^2) \int_0^a r' dr' J_0(krr'/l). \quad (33.22)$$

Из теории функций Бесселя известно соотношение

$$\int x J_0(x) dx = x J_1(x), \quad (33.23)$$

где $J_1(x)$ — функция Бесселя первого порядка. Учитывая (33.23), получаем

$$\int_0^a r' J_0(krr'/l) dr' = [ia/(kr)] J_1(kra/l), \quad (33.24)$$

тогда [см. (33.22)]

$$A(r) = A_0 \frac{2J_1(kar/l)}{kar/l}. \quad (33.25)$$

Функции Бесселя хорошо изучены. График функции $2J_1(\rho)/\rho$ приведен на рис. 167. В центре дифракционной картины имеется светлое круглое пятно, окруженное темными и светлыми дифракционными кольцами. При этом максимумы интенсивности в светлых кольцах быстро убывают (рис. 168). Радиус r первого темного кольца находится из условия

$$J_1(kar_1/l) = 0. \quad (33.26)$$

Принимая во внимание, что первый корень уравнения $J_1(x) = 0$ равен $x_1 = 3,832$, находим угол $\Delta\varphi_1$, под которым это кольцо видно из центра круглого отверстия:

224 $\Delta\phi_1 = (r_1/l) = 3,832/(ka) = 0,61\lambda/a$.

(33.27)

Поэтому угловой размер светлого пятна, наблюдаемого из центра круглого отверстия, равен

$$\theta_0 = 2\Delta\phi_1 = 1,22\lambda/a.$$
(33.28)

Эта формула играет большую роль в вопросе о разрешающей силе оптических приборов (см. § 36). Второе кольцо видно под углом

$$\Delta\phi_2 = 1,116\lambda/a,$$
(33.29)

однако интенсивность вне центрального светлого пятна очень сильно убывает. В пределах центрального пятна сосредоточено около 84% всей энергии, проходящей через отверстие. Энергией, приходящейся на область вне центрального пятна, можно в большинстве случаев пренебречь.

Дифракционная решетка. Она представляет собой совокупность периодически повторяющихся щелей (рис. 169). Обозначим a , b и $d = a + b$ соответственно ширину непроницаемой части решетки, ширину отверстия и длину периода решетки. Дифракционную картину можно найти с помощью формулы (33.8) аналогично тому, как это было сделано для одной щели. Однако более наглядно и проще провести анализ методом сложения комплексных амплитуд.

Амплитуду волны, дифрагированной каждой из щелей в направлении, характеризуемом углом φ [см. (33.15)], представим в виде

$$A_\varphi^{(0)} = A_0 \sin \beta / \beta, \quad \beta = kb \sin \varphi / 2,$$
(33.30)

где индекс (0) у A_φ отмечает, что это амплитуда волны от одной щели. Дифрагированные от щелей волны интерферируют между собой и образуют дифракционную картину. Таким образом, дифракционная картина от решетки является результатом дифракции волн на каждой щели и интерференции волн от различных щелей.

Рассмотрим интерференцию волн от щелей. Разность хода волн от двух соседних щелей (рис. 169) и разность фаз между ними равны

$$\Delta = d \sin \varphi,$$
(33.31)

$$\delta = k\Delta = kd \sin \varphi.$$
(33.32)

Взяв за начало отсчета фазу волны, дифрагированной в направлении, определяемом углом φ от первой щели, т. е. принимая ее амплитуду равной (33.30), находим для амплитуд волн, дифрагированных от второй, третьей щелей и т. д., выражения $A_\varphi^0 e^{-i\delta}$, $A_\varphi^0 e^{-i2\delta}$, В результате суперпозиции волн от всех N щелей образуется волна с амплитудой

$$A_\varphi = A_\varphi^0 + A_\varphi^0 e^{-i\delta} + A_\varphi^0 e^{-i2\delta} + \dots + A_\varphi^0 e^{-(N-1)\delta} = A_\varphi^0 (1 - e^{-iN\delta}) / (1 - e^{-i\delta}),$$
(33.33)

где использована формула для геометрической прогрессии

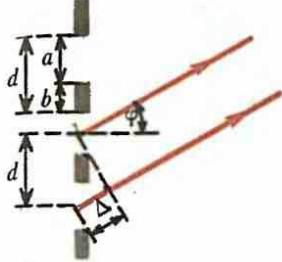
$$1 + x + x^2 + \dots + x^n = (1 - x^{n+1}) / (1 - x).$$

Учитывая, что

$$\frac{1 - e^{-iN\delta}}{1 - e^{-i\delta}} = \frac{e^{-iN\delta/2}}{e^{-i\delta/2}} \frac{e^{iN\delta/2} - e^{-iN\delta/2}}{e^{i\delta/2} - e^{-i\delta/2}} = \frac{e^{-iN\delta/2}}{e^{-i\delta/2}} \frac{\sin(N\delta/2)}{\sin(\delta/2)},$$
(33.34)

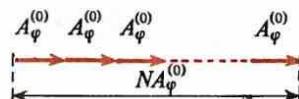
и используя (33.30), окончательно находим

$$|A_\varphi| = |A_0| |\sin \beta / \beta| |\sin N\alpha / \sin \alpha|, \quad \beta = kb \sin \varphi / 2, \quad \alpha = \delta / 2 = kd \sin \varphi / 2.$$
(33.35)



169

Дифракционная решетка



170

Сложение комплексных амплитуд, приводящее к образованию главных максимумов

Формула (33.35) полностью описывает дифракционную картину от щели, поскольку интенсивность картины определяется $|A_\phi|^2$. Из (33.31) видно, что при

$$d \sin \varphi = m\lambda \quad (m = 0, \pm 1, \pm 2, \dots) \quad (33.36)$$

волны от соседних щелей усиливают друг друга, т. е. волны от всех щелей усиливают друг друга. Это означает, что соотношение (33.36) определяет направления, по которым образуются главные максимумы. При этом условии $a = kd \sin \varphi/2 = \pi m$ и, следовательно,

$$|\sin Na / \sin a| = N, \quad (33.37)$$

Графически сложение комплексных амплитуд от отдельных щелей, приводящее к образованию главных максимумов, показано на рис. 170. Видно, что $N|A_\phi^{(0)}|$ является действительно максимальной амплитудой, которая может быть образована из амплитуд волн, дифрагированных на N щелях в направлении, определяемом углом φ . Отсюда ясно, почему эти максимумы называются главными. Однако $|A_\phi|$ амплитуд главных максимумов не одинаков. Из (33.35) следует, что он модулируется множителем $|\sin \beta/\beta|$, т. е. амплитуда главных максимумов модулируется дифракцией от отдельных щелей. Максимальное значение $|\sin \beta/\beta|$ равно единице, оно достигается при условии $\beta = 0$, которое соответствует центральному максимуму ($\varphi = 0$). Амплитуда всех остальных главных максимумов меньше. Если главный максимум приходится на направление, для которого $\sin \beta = 0$, то он отсутствует. Это может случиться, когда b и d соизмеримы.

Минимумы излучения образуются тогда, когда в результате сложения комплексных амплитуд получается результирующая нулевая амплитуда (рис. 171). Для различных δ (т. е. при различных φ) ломаная кривая может быть замкнута один раз, два раза и т. д., т. е. разность фаз волн от крайних щелей решетки равна $2\pi, 4\pi, \dots$. Поэтому условие минимумов амплитуд в дифракционной картине записывается в виде

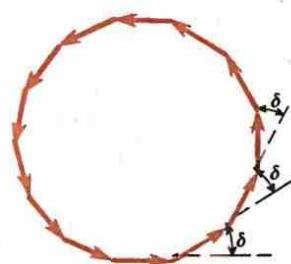
$$N\delta = 2\pi n \quad (n = 0, 1, 2, \dots) \quad (33.38)$$

или, поскольку $\delta = kd \sin \varphi = 2\pi d \sin \varphi / \lambda$,

$$d \sin \varphi = (n/N)\lambda \quad (n \neq 0, N, 2N, \dots), \quad (33.39)$$

так как при $n = 0, N, 2N, \dots$ отношение n/N становится целым числом ($m = n/N$) и (33.39) превращается в условие главного максимума. Условие (33.38) для минимума может быть, конечно, получено и из (33.35):

$$\sin(N\delta/2) = 0. \quad (33.40)$$



171

Сложение комплексных амплитуд, приводящее к образованию минимумов



Общий характер дифракционной картины от шести щелей (без соблюдения масштабов)

Отсюда следует, что $N\delta/2 = n\pi$ ($n = 0, 1, 2, \dots$) совпадает с (33.38). Если $n/N = m$, то $\sin Na/\sin a$ остается конечным и обуславливает возникновение главного максимума.

Из (33.39) заключаем, что между двумя главными максимумами имеется $N - 1$ минимумов. Ясно, что между минимумами должны быть максимумы, которые называются вторичными. Следовательно, между двумя главными максимумами имеется $N - 2$ вторичных максимумов. Это позволяет, зная число N щелей, выяснить общий характер дифракционной картины, а по формуле (33.35) найти относительные значения модулей амплитуд и интенсивностей различных частей дифракционной картины.

На рис. 172 показан общий характер распределения модуля амплитуды при шести щелях ($N = 6$) без соблюдения масштабов.

Дифракция белого света на решетке. Каждая из волн различной длины дает свою дифракционную картину. Из условия $d \sin \phi = m\lambda$ видно, что угол ϕ для фиксированного m увеличивается с увеличением λ . В нулевом порядке интерференций ($m=0$) центральный максимум $\phi=0$ совпадает для всех волн. Поэтому в центре образуется белая полоса. Затем идет первый порядок интерференции ($m=1$). Линии интерференции первого порядка не перекрываются линиями интерференции второго порядка. Перекрытие различных порядков наступает при $m\lambda = m'\lambda'$, где λ, λ' — длины волн; m, m' — порядки их интерференции. Для видимого спектра

$$1 \cdot 0,76 \cdot 10^{-6} = 2 \cdot 0,38 \cdot 10^{-6}, \quad (33.41)$$

т. е. длинноволновая граница первого порядка спектра попадает на самое начало коротковолновой границы второго порядка спектра и, следовательно, первый порядок видимой части спектра не перекрывается вторым порядком. Второй и третий порядки спектра перекрываются. Как видно из (33.41), это обстоятельство обусловлено тем, что видимая часть спектра занимает одну октаву.

Дисперсионная область. Для дифракционной решетки она определяется так же, как в случае интерферометра Фабри — Пере (см. § 28), и находится по формуле (28.34). У дифракционной решетки обычно наблюдаются спектры низких порядков ($m = 1, 2, 3, \dots$), поэтому дисперсионная область оказывается очень большой ($\Delta\lambda = \lambda, \Delta\lambda = \lambda/2, \Delta\lambda = \lambda/3$). В частности, в первом порядке дисперсионная область решетки совпадает со всем видимым спектром.

Разрешающая способность. Для дифракционной решетки она определяется так же, как и в случае интерферометра Фабри — Пере (см. § 28), однако в качестве условия разрешения линий принимается условие Рэлея: линии считаются разрешенными, если максимум интенсивности одной попадает на минимум интенсивности другой.

Из (33.39) видно, что главный максимум возникает при $n = mN$, а соседний минимум — при $n = mN + 1$.

Следовательно, по (33.39) для соответствующих углов можно написать

$$d \sin \phi_{\max} = mN\lambda/N = m\lambda, \quad (33.42)$$

$$d \sin \phi_{\min} = (mN + 1)\lambda/N = (m + 1/N)\lambda.$$

Для длины волны $\lambda_1 = \lambda + \Delta\lambda$ условие главного максимума порядка m имеет вид

$$d \sin \phi_{\max}^{(1)} = m\lambda_1. \quad (33.43)$$

$$d \sin \phi_{\max}^{(1)} = m\lambda_1.$$

По условию Рэлея линии считаются разрешенными, если $\varphi_{\min} = \varphi_{\max}^{(1)}$. С учетом (33.42) и (33.43) это условие записывается в виде

$$(m + 1/N)\lambda = m(\lambda + \Delta\lambda). \quad (33.44)$$

Отсюда для разрешающей способности решетки [см. (рис. 28.22)] находим

$$Y = \lambda/\Delta\lambda = Nm. \quad (33.45)$$

Так как число N щелей в решётке достигает многих десятков тысяч, то её разрешающая способность большая. Для повышения разрешающей способности необходимо переходить к большим значениям порядка m интерференции. Однако m ограничены углами отклонения $\varphi \leq \pi/2$ и, следовательно, $\sin \varphi$ в (33.36) не может быть больше единицы. Поэтому $m_{\max} = d/\lambda$ и

$$Y_{\max} = Nd/\lambda, \quad (33.46)$$

равно числу волн, укладывающихся на полной ширине решетки. Следовательно, максимальная разрешающая способность не зависит от числа щелей и от того, как на каждом периоде d решётки распределяются между собой прозрачная и непрозрачная части. Решетки с большими N облашают существенными преимуществами. У них значительно лучше угловая дисперсия и отношение изменения $d\varphi$ в угле при изменении длины волны λ к $\delta\lambda$, равно

$$D = \delta\varphi/\delta\lambda. \quad (33.47)$$

Из (33.36) находим $d \cos \varphi \delta\varphi = m \delta\lambda$ и, следовательно,

$$D = m/(d \cos \varphi). \quad (33.48)$$

Поэтому у дифракционной решетки с малым d угловая дисперсия выше. Это означает, что заданная угловая дисперсия на ней достигается при меньших порядках m интерференции, что очень важно, поскольку на меньших порядках интерференции интенсивности максимумов больше. Кроме того, при малых m больше дисперсионная область. Поэтому практическое значение имеют решетки лишь с достаточно большим числом N периодов и большой общей шириной. Лучшие решетки имеют до 15 см ширины и содержат около 100 000 периодов.

Отражательные дифракционные решетки. Если вместо щелей на решётке находятся хорошо отражающие участки, а вместо непроницаемых участков — участки, поглощающие свет, то в отражённом свете наблюдается дифракционная картина, аналогичная картине в проходящем свете. Действующие в отраженном свете решётки называются **отражательными**. Если непроницаемая часть решетки достаточно хорошо отражает свет, а щель хорошо пропускает, то решётка может одновременно действовать в качестве дифракционной решётки как в проходящем, так и в отраженном свете. Период d обеих решёток одинаков, а смысл величин a и b противоположен: в формулах для интенсивности они меняются местами.

Дифракция на щели с непрерывным изменением фазы волны. Допустим, что в щель помещена призма с показателем преломления n (рис. 173). При прохождении плоской волны через призму на выходе из щели у волны образуется сдвиг фаз, изменяющийся вдоль щели. Как и прежде, начало координат предполагается находящимся в середине щели. Считая угол a у вершины призмы малым, можно толщину призмы в точке с координатой y' считать равной $ab/2 + ay'$. Оптическая толщина равна $(ab/2 + ay')n$, а сдвиг фаз

- **Дифракция Френеля** осуществляется в ближней зоне, а дифракция Фраунгофера — в дальней.
Дифракция Фраунгофера описывается с помощью преобразования Фурье.

228 $\delta = kn(ab/2 + ay') + k[ab - (ab/2 + ay')]$, (33.49)

где второе слагаемое учитывает сдвиг фаз при распространении луча на той части длины ab , которую он проходит в воздухе. Поэтому Ψ в формуле (33.8), записанной для одной переменной (щель предполагается бесконечно длинной), имеет вид

$$\Psi(y') = (A_0/b) e^{i\delta} \quad (33.50)$$

и, следовательно, амплитуда дифрагированной волны [см. (33.16)]

$$A_\varphi = (A_0/b) \exp[ik(n+1)ab/2] \times \\ \times \int_{-b/2}^{b/2} \exp\{-i[k \sin \varphi - k(n-1)a]y'\} dy' = A_0 e^{ik(n+1)ab/2} \sin \beta'/\beta' \\ \beta' = kb[\sin \varphi - (n-1)a]/2, \quad (33.51)$$

откуда

$$|A_\varphi| = |A_0| \left| \frac{\sin \beta'}{\beta'} \right| = |A_0| \frac{\sin \{kb[\sin \varphi - (n-1)a]/2\}}{kb[\sin \varphi - (n-1)a]/2}. \quad (33.52)$$

Видно, что центральный максимум дифракционной картины располагается на дифракционном угле $\varphi = (n-1)a$, а в остальном характер дифракции существенно не меняется (рис. 173).

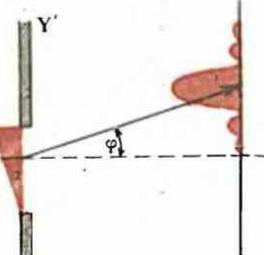
Фазовые решетки. Если во все щели дифракционной решетки поставить призмы, как это только что было рассмотрено, то максимум дифракции сместится на дифракционный угол $(n-1)a$. Направление на главные максимумы не изменяется, поскольку они обусловлены интерференцией между волнами от различных щелей, а условия этой интерференции не меняются (разность хода между лучами от соседних щелей по-прежнему $d \sin \varphi$). Поэтому общий характер интерференционной картины постоянен, изменяется лишь распределение интенсивностей: максимальной интенсивностью обладает не главный максимум на угле $\varphi = 0$, а тот главный максимум, который попадает на угол $\varphi = (n-1)a$ или находится вблизи этого угла. Это позволяет работать с более высокими порядками m интерференции, что улучшает разрешающую способность, и в то же время избегать потерь в интенсивности.

Решетки, при прохождении через которые меняется фаза волны, называются **фазовыми**. Вариация фазы вдоль щели может быть самой разнообразной. Она подбирается исходя из требований задачи, для решения которой предназначена фазовая решетка.

Амплитудно-фазовые решетки. В них при прохождении через щель света существенно изменяется не только фаза, но и амплитуда волн, т. е. происходит поглощение света.

Наклонное падение лучей на решетку. При наклонном падении лучей на решетку (рис. 174) условием главных максимумов по-прежнему является равенство разности хода лучей в соседних щелях целому числу длин волн [см. (33.36)]. На рис. 174 видно, что разность хода между лучами равна

$$\Delta = \Delta'' - \Delta' = d \sin \theta - d \sin \varphi,$$



173

Дифракция на щели с непрерывным изменением фазы волны

- При каких условиях дифракция Фраунгофера наблюдается на малых расстояниях? Чем объясняется большая дисперсионная область дифракционной решетки?
- Можете ли Вы описать возможновение дифракции на решетке с помощью представлений о дифракции на непрерывно изменяющихся структурах?

поэтому условие главного максимума имеет вид

$$d(\sin \theta - \sin \varphi) = m\lambda. \quad (33.53)$$

Будем отсчитывать углы α от направления падающего луча. Как видно из рис. 174

$$\alpha = \theta - \varphi. \quad (33.54)$$

Обычно углы дифракции малы, т. е. $\alpha \ll 1$, поэтому

$$\begin{aligned} \sin \theta - \sin \varphi &= 2 \cos [(\theta + \varphi)/2] \sin [(\theta - \varphi)/2] = \\ &= 2 \cos \theta \sin (\alpha/2) = \cos \theta \sin \alpha. \end{aligned} \quad (33.55)$$

Обозначая $d' = d \cos \theta$ (рис. 174), запишем условие (33.53) в форме

$$d' \sin \alpha = m\lambda, \quad (33.56)$$

совпадающей с (33.36). Это означает, что дифракция при наклонном падении лучей на решетку происходит так же, как при прямом, однако в качестве периода выступает проекция периода решетки на перпендикулярное падающему лучу направление (рис. 174), т. е. дифракция происходит как бы на решетке с меньшим периодом.

Это позволяет наблюдать дифракцию от волн с очень короткой длиной, пользуясь решеткой с большим периодом, когда при нормальном падении волн на решетку никакой дифракции практически не наблюдается. Напомним, что для наблюдения дифракции при нормальном падении лучей период решетки должен быть немного больше длины волны. Например, в обычной патефонной пластинке расстояние между бороздками, в которых осуществляется запись звука, равно примерно 40—50 мкм. При косом падении лучей белого света на пластинку эффективное расстояние между бороздками может стать равным нескольким микрометрам, и пластинка будет отражать свет как отражательная дифракционная решетка с образованием спектра в первом порядке и цветных полос в высших порядках интерференции.

Дифракция на непрерывных периодических и непериодических структурах. Дифракционная решетка является периодической структурой, у которой коэффициент пропускания τ амплитуды равен 1 на щелях и 0 на непрозрачных частях (рис. 175). Если амплитуда падающей на решетку волны A_0 , то амплитуда выходящей из решетки волны равна

$$A = A_0 \tau = \begin{cases} A_0 & \text{(на щелях),} \\ 0 & \text{(на непрозрачных частях).} \end{cases}$$

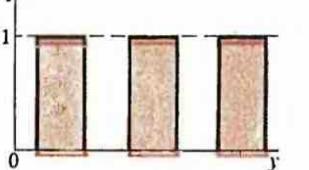
Предположим, что вместо решетки имеется слой вещества, коэффициент пропускания которого изменяется по закону (рис. 176)

$$\tau(y') = 1/2[1 + \cos(2\pi y'/d)], \quad (33.57)$$

174

Наклонное падение лучей на решетку

175

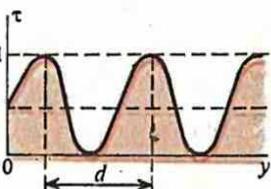


175

Изменение коэффициента пропускания амплитуды для дифракционной решетки

176

Непрерывная структура с гармонически изменяющимся коэффициентом пропускания амплитуды



причем считается, что слой бесконечен, т. е. $-\infty < y' < \infty$. Если на слой вещества падает волна с амплитудой A_0 , то на выходе имеется волна с амплитудой

$$A(y') = (A_0/2) [1 + \cos(2\pi y'/d)]. \quad (33.58)$$

По формуле (33.8) для одного измерения найдем распределение амплитуды A по углам дифракции:

$$A_\varphi = \frac{B}{2\pi} \int_{-\infty}^{\infty} [1 + \cos(2\pi y'/d)] e^{-iky' \sin \varphi} dy', \quad (33.59)$$

где постоянная B включает в себя как A_0 , так и некоторый нормировочный множитель, который нет необходимости более точно определять, потому что все расчеты ведутся с точностью до относительного изменения амплитуды; множитель $1/(2\pi)$ введен для удобства. Принимая во внимание, что

$$\cos(2\pi y'/d) = \frac{1}{2} (e^{i2\pi y'/d} + e^{-i2\pi y'/d}), \quad (33.60)$$

представим выражение (33.59) в виде

$$\begin{aligned} A_\varphi = & \frac{B}{2\pi} \int_{-\infty}^{\infty} \exp(-iky' \sin \varphi) dy' + \frac{B}{4\pi} \int_{-\infty}^{\infty} \exp[i(\frac{2\pi}{d} - k \sin \varphi)y'] dy' + \\ & + \frac{B}{4\pi} \int_{-\infty}^{\infty} \exp[-i(\frac{2\pi}{d} + k \sin \varphi)y'] dy', \end{aligned} \quad (33.61)$$

откуда

$$A_\varphi = B\delta(k \sin \varphi) + (B/2)\delta(2\pi/d - k \sin \varphi) + (B/2)\delta(2\pi/d + k \sin \varphi). \quad (33.62)$$

Следовательно, дифракционная картина состоит из трех линий: линии на угле $\varphi = 0$, которая возникает за счет постоянной составляющей амплитуды в (33.58) и соответствует просто прямолинейному распространению волны через однородную среду, и двух линий, дифрагированных под углами $\varphi_{(+)}$ и $\varphi_{(-)}$, определяемых из условий

$$2\pi/d - k \sin \varphi_{(+)} = 0, \quad 2\pi/d + k \sin \varphi_{(-)} = 0. \quad (33.63)$$

Учитывая, что $k = 2\pi/\lambda$, можно эти условия записать в виде

$$d \sin \varphi_{(+)} = \lambda, \quad d \sin \varphi_{(-)} = -\lambda. \quad (33.64a)$$

Таким образом, на гармонической структуре наблюдается лишь дифракция первого порядка, причем угол дифракции связан с периодом гармонической структуры соотношением, аналогичным (33.36) для главного максимума первого порядка у дифракционной решетки с тем же периодом.

Теперь можно рассмотреть дифракцию на решетке как дифракцию на гармонических структурах. Периодическую функцию, выражающую коэффициент пропускания решетки с периодом d , можно разложить в ряд Фурье по гармоническим функциям, периоды которых равны $d, d/2, d/3, \dots$. Дифракция волны длиной волны λ , падающей на решетку, сводится к дифракции на гармонических структурах, составляющих решетку, которые были получены разложением коэффициента пропускания решетки в ряд Фурье. При дифракции на гармонической составляющей структуры с длиной d/m возникают два дифракционных максимума, условия которых в соответствии с (33.64) имеют вид

$$(d/m) \sin \varphi_m = \lambda, \quad (d/m) \sin \varphi_{-m} = -\lambda$$

(33.64б) 231

или

$$d \sin \varphi = m\lambda \quad (m = \pm 1, \pm 2, \dots),$$

(33.65)

что совпадает с (33.36). Тем самым положение главных максимумов решетки полностью описано. В принципе можно описать таким методом и изменение интенсивности, если принять во внимание значения коэффициентов в ряде Фурье.

Теперь ясно, как можно рассмотреть дифракцию на произвольной периодической структуре. Надо представить ее характеристики рядом Фурье, рассмотреть дифракцию первого порядка, описываемую отдельными членами ряда Фурье. Совокупность этих дифракций первого порядка составляет всю дифракционную картину на периодической структуре. Ясно, в принципе, что дифракцию на непериодической структуре можно рассмотреть аналогично. Надо вместо ряда Фурье использовать интеграл Фурье.

Дифракция на ультразвуковых волнах. Ультразвуковыми называются колебания с частотой порядка 10^8 Гц. В жидкости скорость звука $v \approx 10^3$ м/с, и поэтому длина ультразвуковой волны $d = v/\nu = 10^{-5}$ м = 10 мкм. Уплотнения и разрежения в ультразвуковой волне, распространяющейся в жидкости, создают фазовую гармоническую решетку. При гармонической модуляции фазы возникает дифракция, аналогичная той, которая была рассмотрена для гармонической модуляции амплитуды. Поэтому должна наблюдаться дифракция первого порядка, которую очень удобно воспроизвести с помощью ультразвуковой установки, схема которой изображена на рис. 177. Пьезодатчик P создает ультразвуковые волны, на которых происходит дифракция волн, испускаемых источником S . Имеются два дифракционных максимума первого порядка в полном соответствии с (33.64а) и центральный максимум.

Сравнение характеристик спектральных аппаратов. Одно из главных применений интерферометров и дифракционных решеток сводится к анализу спектрального состава излучения, т. е. эти приборы используются в качестве спектральных аппаратов. Какой из аппаратов наиболее целесообразно применять, зависит от конкретных условий.

Высокая разрешающая способность достигается как в интерферометрах Фабри—Перо и Майкельсона (порядка 10^6), так и в дифракционных решетках (порядка 10^5) в других интерферометрах. Однако такая высокая разрешающая способность в них достигается за счет различных факторов. В интерферометре Фабри—Перо и Майкельсона она достигается за счет высоких порядков интерференции (порядка 10^5) при сравнительно небольшом числе интерферирующих лучей (несколько десятков в интерферометре Фабри—Перо и два луча в интерферометре Майкельсона), а в дифракционной решетке — за счет большого числа интерферирующих лучей (порядка 10^5) при малом порядке интерференции (несколько единиц). Благодаря этому дисперсионная область очень мала у интерферометра Фабри—Перо (порядка 10^{-2} нм) и интерферометра Майкельсона (порядка 10^{-3} нм) и очень велика у дифракционной решетки (порядка 10^3 нм). Поэтому если исследуемое излучение имеет большую дисперсионную область, а его необходимо исследовать с помощью приборов высокого разрешения с малой дисперсионной областью, то приходится комбинировать между собой различные спектральные аппараты. При этом получаются одновременно и широкая дисперсионная область и большое разрешение.

Спектральные аппараты должны обеспечивать возможность работы со слабыми интенсивностями исследуемого излучения. В этом отношении интерферометр Фабри—Перо существенно превосходит дифракционную решетку, особенно если пользоваться фотозелектрической регистрацией в схеме сканирующего интерферометра Фабри—Перо. Разрешающая способность в Фурье-спектроскопии определяется максимальной разностью хода, которая может быть обеспечена механизмом подвижного зеркала, и достигает больших значений.

$\tau(y') = \cos(\pi y' / b)$. Начало отсчета находится в центре щели; b — ширина щели, φ — угол дифракции.

Имеем

$$\begin{aligned} A_\varphi &= B \int_{-b/2}^{b/2} \cos(\pi y' / b) e^{-ik \sin \varphi \cdot y'} dy' = (B/2) \int_{-b/2}^{b/2} e^{i(\pi/b - k \sin \varphi) y'} dy' + (B/2) \int_{-b/2}^{b/2} e^{-i(\pi/b + k \sin \varphi) y'} dy' = \\ &= -\frac{2\pi B b \cos(kb \sin \varphi / 2)}{(kb \sin \varphi)^2 - \pi^2}. \end{aligned} \quad (33.66)$$

При $\varphi \rightarrow 0$ амплитуда

$$A_0 = 2Bb/\pi. \quad (33.67)$$

Отсюда для распределения интенсивности находим

$$I_\varphi = \frac{I_0}{[1 - 4b^2 \sin^2 \varphi / \lambda^2]} \cos^2(\pi b \sin \varphi / \lambda), \quad (33.68)$$

где I_0 — интенсивность в направлении $\varphi = 0$.

§ 34 | Дифракция Френеля

Рассматривается дифракция в ближней зоне и даётся количественное описание спирали Корнио.

Область дифракции Френеля. Область дифракции Френеля расположена вблизи объекта, на которой происходит дифракция, и простирается до расстояний, с которых дифракцию можно рассматривать как фраунгоферову [см. (33.7)]. Однако полезно заметить, что дифракция Фраунгофера в строгом смысле слова имеет место лишь на бесконечности.

Для рассмотрения дифракции Френеля необходимо пользоваться формулой (32.37), в которой для упрощения написания можно отбросить множитель, стоящий перед интегралом, поскольку он не оказывает влияния на относительное распределение интенсивностей в дифракционной картине.

Дифракция на прямоугольном отверстии. Расположение систем координат и размеры прямоугольного отверстия даны на рис. 159. Считая, что на отверстие падает плоская волна с амплитудой A_0 , полагаем в (32.37) $\Psi = A_0/(ab)$ и записываем эту формулу в виде

$$A(x, y) = [A_0/(ab)] \int_{-a/2}^{a/2} \exp[ik(x - x')^2/(2l)] dx' \int_{-b/2}^{b/2} \exp[ik(y - y')^2/(2l)] dy'. \quad (34.1)$$

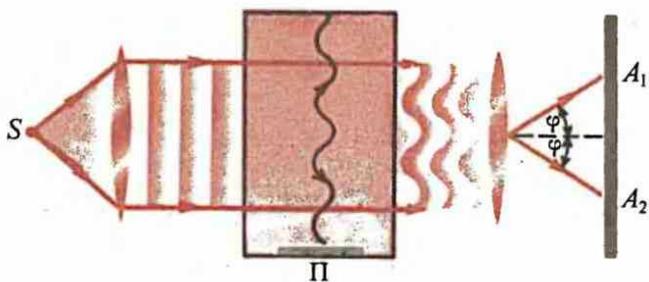
Вычислим первый из перемножаемых интегралов:

$$Q = \int_{-a/2}^{a/2} \exp\left[-\frac{ik(x - x')^2}{2l}\right] dx'. \quad (34.2)$$

Переходя к новой переменной интегрирования

$$\xi = (x - x') [k/(\pi l)]^{1/2}, \quad d\xi = -dx' [k/(\pi l)]^{1/2} \quad (34.3)$$

находим



177

Схема установки для наблюдения дифракции на ультразвуковых волнах

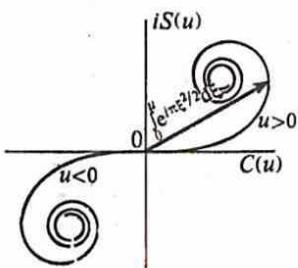
$$\begin{aligned} Q &= -(\pi l/k)^{1/2} \int_{(x-a/2)/\sqrt{k/l(\pi l)}}^{(x+a/2)/\sqrt{k/l(\pi l)}} e^{i\pi\xi^2/2} d\xi = \\ &= (\pi l/k)^{1/2} \left[\int_0^{(x+a/2)/\sqrt{k/l(\pi l)}} e^{i\pi\xi^2/2} d\xi - \int_0^{(x-a/2)/\sqrt{k/l(\pi l)}} e^{i\pi\xi^2/2} d\xi \right]. \end{aligned} \quad (34.4)$$

Интегралы Френеля. Вычислим

$$\int_0^u e^{i\pi\xi^2/2} d\xi = \int_0^u \cos(\pi\xi^2/2) d\xi + i \int_0^u \sin(\pi\xi^2/2) d\xi. \quad (34.5)$$

Интегралы

$$C(u) = \int_0^u \cos(\pi\xi^2/2) d\xi, \quad S(u) = \int_0^u \sin(\pi\xi^2/2) d\xi \quad (34.6)$$



178

Построение спирали Корнио с помощью интегралов Френеля

называются **интегралами Френеля** и находятся численными методами. Имеются таблицы этих интегралов, которые графически представляются в виде спиралей Корнио (см. рис. 151). Ее более точное построение с помощью (34.6) показано на рис. 178. Полезно заметить, что

$$C(\infty) = S(\infty) = 0,5, \quad C(-\infty) = S(-\infty) = -0,5.$$

Спираль Корнио. Из определения (34.5) и правила построения спирали Корнио следует, что **комплексное число**

$$\int_0^u e^{i\pi\xi^2/2} d\xi \quad (34.7)$$

изображается комплексным вектором, начинающимся в начале координат и оканчивающимся в точке спирали Корнио, соответствующей значению u (рис. 178). Тем самым обосновывается метод анализа дифракции с помощью спирали Корнио (см. § 31). Он позволяет найти интегралы (34.1) и проанализировать количественно распределение интенсивности в дифракционной картине.

Задачи

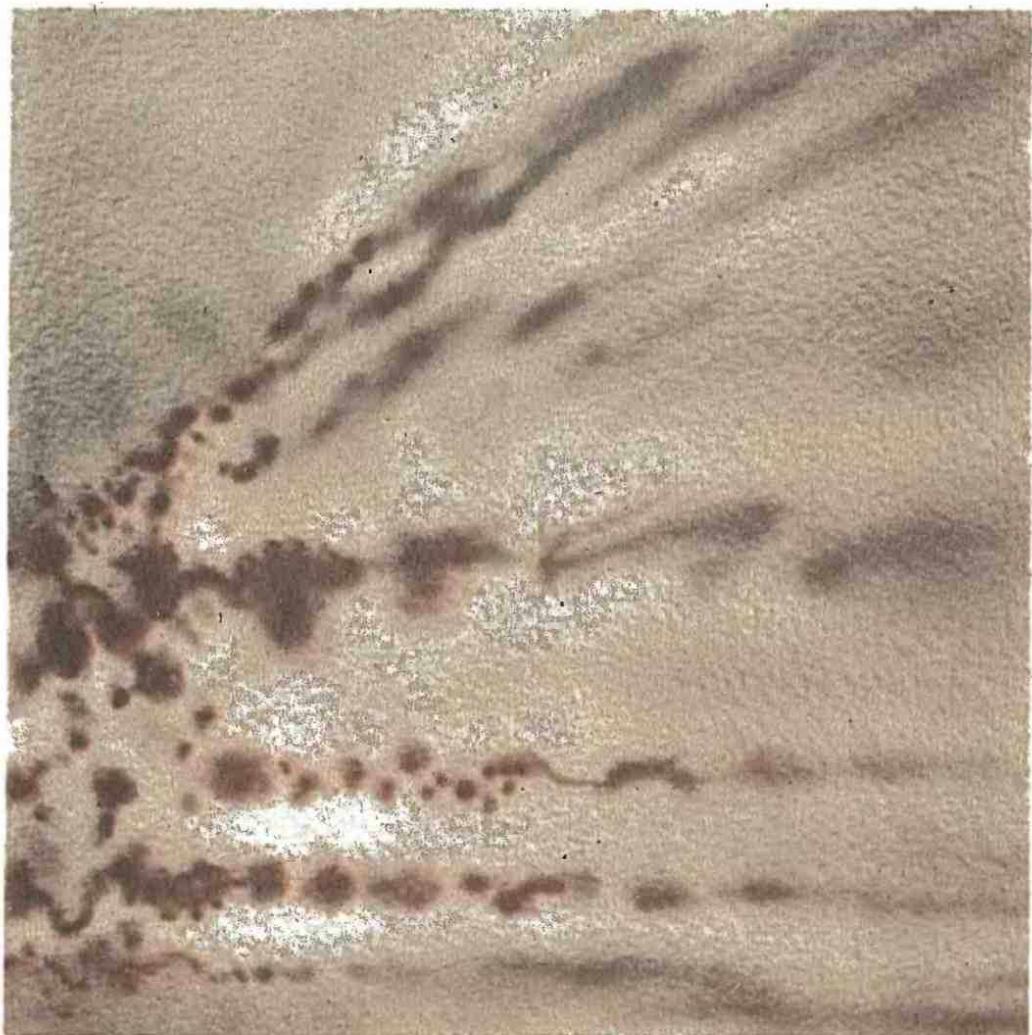
- 6.1. Длины волн дублета в излучении натрия равны $\lambda_1 = 589$ нм и $\lambda_2 = 589,6$ нм. Чему равно минимальное число щелей в дифракционной решетке, чтобы разрешить этот дублет во втором порядке? в первом порядке?
- 6.2. Через стеклянную пластину, покрытую порошком ликоподия, пропускается монохроматическая волна, и на экране за пластиной с помощью линзы с фокусным расстоянием $f = 1$ м наблюдается интерференционная картина. Центральное темное пятно имеет радиус 3 см. Длина волны излучения 600 нм. Найти радиус частиц ликоподия.
- 6.3. Дифракционная решетка имеет 200 штрихов на миллиметр. Под каким углом должно падать на решетку излучение, чтобы в перпендикулярном решетке направлении наблюдался дифракционный максимум второго порядка для длины волны $\lambda = 0,5$ мкм?
- 6.4. Свет падает по нормали к решетке. Найти отношение интенсивностей главного максимума I_m m -го порядка к интенсивности I_0 центрального максимума.
- 6.5. Параллельный пучок монохроматического света падает на проволоку диаметром 1 мм, натянутую перпендикулярно направлению распространения света. На экране, расположенному перпендикулярно направлению распространения света, на расстоянии 1 м от проволоки наблюдаются дифракционные полосы, расстояние между которыми 0,5 мм. Найти длину волны света.
- 6.6. При каком условии m -й главный максимум для дифракционной решетки с периодом d и шириной щели b обращается в нуль?
- 6.7. Диск диаметром 0,5 см, имеющий неровности порядка 10 мкм, расположен на расстоянии 1 м от точечного монохроматического источника света, излучающего на длине волны 0,5 мкм. Точечный источник находится на перпендикуляре к плоскости диска, проходящей через его центр. Считая, что пятно Пуассона видно лишь до тех пор,
- пока неровности краев 'диска перекрывают соответствующую зону Френеля не более чем на $1/4$ ее ширины, найти минимальное расстояние, на котором можно видеть пятно Пуассона.
- 6.8. Точечный монохроматический источник света ($\lambda = 0,5$ мкм) расположен на расстоянии 0,5 м от плоского экрана с круглым отверстием диаметром 0,5 мм на перпендикуляре к плоскости экрана, проходящей через центр отверстия. На каком расстоянии от отверстия на перпендикуляре расположена точка, имеющая максимум освещенности?
- 6.9. Найти условие, при котором все четные максимумы в дифракционной картине от решетки пропадают.
- 6.10. Поверхность Луны облучается пучком света, диаметр которого на выходе из лазера равен 0,5 см. Расходимость пучка чисто дифракционная, влиянием атмосферы пренебречь. Чему равен диаметр лазерного пучка при достижении поверхности Луны, если расстояние до Луны $368 \cdot 10^3$ км, а длина волны лазерного излучения 632,8 нм?
- 6.11. Входной зрачок глаза имеет радиус 1,5 мм. На каком расстоянии от глаза светящийся диск диаметром 1 см, испускающий излучение с длиной волны $\lambda = 0,55$ мкм, создает на входном зрачке глаза когерентное поле?
- 6.12. В эшелоне Майкельсона 20 плоскопараллельных пластинок толщиной 1 мм (показатель преломления $n = 1,5$). Высота ступенек 0,5 см. Найти угловую ширину дифракционных максимумов и порядок интерференции. Длина волны $\lambda = 0,55$ мкм.
- 6.13. Чему равна угловая дисперсия дифракционной решетки, имеющей $4 \cdot 10^3$ штрихов на 1 см, в спектре второго порядка?
- 6.14. Какова угловая разрешающая способность человеческого глаза (диаметр зрачка 4 мм) и зрительной трубы (диаметр входного зрачка 10 см) при длине волны 0,6 мкм?

Ответы

- 6.1. 491; 982. 6.2. 12 мкм. 6.3. $23^\circ 30'$. 6.4. $I_m/I_0 = [\sin(m\pi/b)/(m\pi/b)]^2$. 6.5. 0,5 мкм. 6.6. $m = nd/b$ (n — целое число). 6.7. 0,67 м. 6.8. 0,17 м. 6.9. $d = 2b$. 6.10. 118,58 км. 6.11. 110 м. 6.12. $0,27 \cdot 10^{-4}$ рад; $4,5 \cdot 10^3$. 6.13. $0,8 \cdot 10^5$ рад/м = 0,08 рад/мкм. 6.14. $1,83 \cdot 10^4$ рад = $38''$; $7,32 \cdot 10^{-6}$ рад = $1,5''$.

Основная идея:
объект рассматривается
как источник
вторичных волн,
дифракционная картина
от которого
описывается методами
теории дифракции.

**Основные
понятия
Фурье-оптики**



Линза как элемент, осуществляющий преобразование Фурье

Показывается, что линза осуществляет преобразование Фурье распределения амплитуд между передней и задней фокальными плоскостями линзы с точностью до несущественного для распределения интенсивностей фазового множителя.

Фазовое преобразование, осуществляющее тонкой линзой. Тонкой называется линза, у которой можно считать одинаковыми координаты (x', y') входящего и выходящего лучей. Показатель преломления материала линзы n (рис. 179).

Плоскости, перпендикулярные осям Z и касающиеся левой и правой поверхностей линзы, будем называть соответственно левой и правой плоскостью. Координаты точек в этих плоскостях обозначаем (x', y') , потому что по условию тонкой линзы луч света всегда соединяет эти точки. Длину луча в линзе, характеризуемую координатами (x', y') , обозначим $\Delta(x', y')$, а максимальную толщину линзы — Δ_0 (рис. 179).

Полное изменение фазы волны при прохождении пути от левой плоскости до правой равно

$$\delta(x', y') = kn\Delta(x', y') + k(\Delta_0 - \Delta(x', y')) = \\ = k\Delta_0 + k(n-1)\Delta(x', y'), \quad (35.1)$$

где $\Delta(x', y')$ — функция толщины линзы. Величина

$$\tau(x', y') = e^{i\delta(x', y')} = e^{ik\Delta_0} e^{ik(n-1)\Delta(x', y')} \quad (35.2)$$

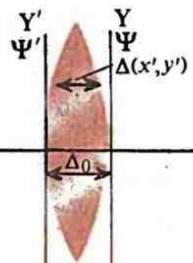
является коэффициентом пропускания линзы. Для линзы он чисто фазовый, поскольку в ней поглощение света пренебрежимо мало.

В геометрической оптике в качестве основного использовалось понятие луча. Задача заключалась в выяснении влияния различных факторов на луч света. Это действие сводилось к изменению направления луча под влиянием этих факторов. Например, преломление на поверхностях двояковыпуклой линзы приводит к тому, что луч света после прохождения линзы отклоняется от первоначального направления к ее оси. Понятие луча было связано с волновой природой света посредством его определения: луч является линией, касательная к которой в каждой точке совпадает по направлению с нормалью к фронту волны. При дальнейшем рассмотрении каких-либо ссылок на волновую природу света делать нет необходимости.

В Фурье-оптике в качестве основного используется понятие волны. Задача заключается в выяснении влияния различных факторов на волновой фронт, фазу и амплитуду волны. Такой подход является более общим, чем подход в геометрической оптике, поскольку зная поведение волновых фронтов всегда можно выяснить не только характеристики соответствующих лучей, но и другие свойства излучения, в частности, распределение амплитуд.

Прежде всего рассмотрим прохождение электромагнитной волны сквозь тонкую линзу.

Амплитуда $\Psi'(x', y')$ волны на входе в линзу на левой плоскости и амплитуда $\Psi(x', y')$ волны на выходе из линзы

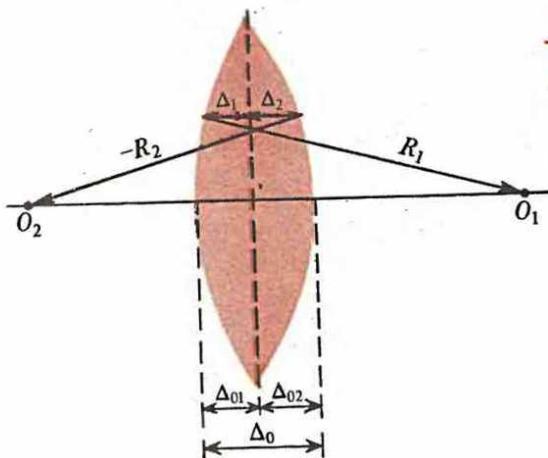


179

К расчету фазового преобразования, осуществляемого линзой

Линза осуществляет такое преобразование фазы плоской волны, в результате которого волна становится сферической сходящейся или расходящейся.

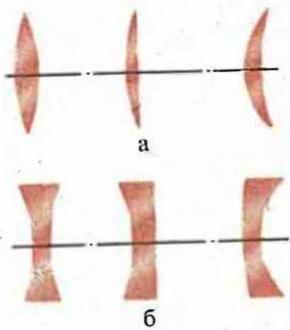
С точностью до несущественных масштабных и фазовых множителей распределение амплитуд в фокальной плоскости линзы является образом Фурье распределения амплитуд на входе в линзу.



180
К расчету функции толщины линзы

на правой плоскости связаны соотношением

$$\Psi(x', y') = \tau(x', y') \Psi'(x', y'). \quad (35.3)$$



181
Линзы с положительным ($f > 0$)
(а) и отрицательным ($f < 0$) (б)
фокусными расстояниями

Расчет функции толщины. При расчете функции толщины радиус кривизны выпуклой поверхности по ходу луча считают положительным, радиус кривизны вогнутой поверхности — отрицательным (рис. 180). Предполагается, что луч распространяется слева направо. Тогда

$$\begin{aligned} \Delta &= \Delta_1 + \Delta_2, \quad \Delta_0 = \Delta_{01} + \Delta_{02}, \quad \rho^2 = x'^2 + y'^2, \\ \Delta_1 &= \Delta_{01} - (R_1 - \sqrt{R_1^2 - \rho^2}), \\ \Delta_2 &= \Delta_{02} - (-R_2 - \sqrt{R_2^2 - \rho^2}), \end{aligned} \quad (35.4)$$

откуда

$$\begin{aligned} \Delta &= \Delta_1 + \Delta_2 = \Delta_0 - R_1(1 - \sqrt{1 - \rho^2/R_1^2}) + \\ &+ R_2(1 - \sqrt{1 - \rho^2/R_2^2}), \end{aligned} \quad (35.5)$$

где в последнем слагаемом правой части равенства при извлечении квадратного корня из R_2^2 учтено условие отрицательности R_2 , т. е. $\sqrt{R_2^2} = -R_2$. Разлагая корни в ряд по $(\rho/R)^2$ и ограничиваясь членами порядка $(\rho/R)^2$, получаем

$$\Delta = \Delta_0 - \frac{x'^2 + y'^2}{2} \left(\frac{1}{R_1} - \frac{1}{R_2} \right). \quad (35.6)$$

На основании (23.32) можно записать

$$1/f = (n-1)(1/R_1 - 1/R_2) \quad (35.7)$$

и представить (35.6) в виде

$$\Delta = \Delta_0 - (x'^2 + y'^2)/[2f(n-1)], \quad (35.8)$$

где f — фокусное расстояние линзы. Подставляя (35.8) в (35.1),

находим

$$\delta(x', y') = kn\Delta_0 - k(x'^2 + y'^2)/(2f), \quad (35.9)$$



182
Плоская волна, выходящая из линзы в виде сферической сходящейся
волны (при $f > 0$)

и, следовательно, (35.2) принимает вид

$$\tau(x', y') = \exp(ikn\Delta_0) \exp\{-ik(x'^2 + y'^2)/(2f)\}. \quad (35.10)$$

Виды линз. Линзы с положительным фокусным расстоянием ($f > 0$) изображены на рис. 181, а, с отрицательным ($f < 0$) — на рис. 181, б. Первый сомножитель в (35.10) не приводит к искажению формы плоской волны при прохождении линзы; второй переводит плоскую волну в сферическую, причем она может быть расходящейся или сходящейся в зависимости от знака экспоненты. Если экспонента отрицательна (т. е. $f > 0$), то периферийные части волны испытывают при прохождении линзы меньшую задержку фазы, чем центр волны. В результате на выходе из линзы образуется сходящаяся сферическая волна (рис. 182). При $f < 0$ она выходит сферически расходящейся (рис. 183). Такая волна является сферической только в параксиальном приближении. Аберрации здесь не учитываются.

Линза как элемент, осуществляющий преобразование Фурье. Поместим предмет с амплитудным коэффициентом пропускания $\tau_0(x', y')$ непосредственно перед линзой (рис. 184) и направим на него плоскую монохроматическую волну. На передней плоскости перед линзой образуется световое поле $A_0 \tau_0(x', y') = \Psi'(x', y')$, где A_0 — амплитуда плоской волны. Конечный размер апертуры линзы учитывается функцией зрачка, которая определяется условием

$$P(x', y') = \begin{cases} 1 & \text{внутри апертуры линзы,} \\ 0 & \text{вне апертуры линзы.} \end{cases} \quad (35.11)$$

Образующееся на выходе из линзы световое поле на основании (35.3) с учетом (35.10) описывается функцией

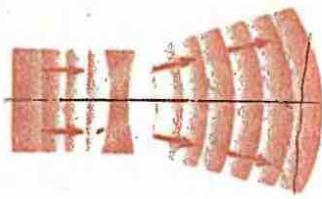
$$\Psi(x', y') = \exp(ikn\Delta_0) \Psi'(x', y') P(x', y') \times \exp[-ik(x'^2 + y'^2)/(2f)]. \quad (35.12)$$

Дифракционная картина в фокальной плоскости линзы при $z_0 = f$ с помощью (32.37) описывается формулой

$$\Phi(x, y) = \frac{k}{2\pi i} \frac{\exp[i(kf + kn\Delta_0)]}{f} \iint_{-\infty}^{\infty} \Psi'(x', y') P(x', y') \times \exp\left[-\frac{ik(x'^2 + y'^2)}{2f}\right] \times \exp\left\{\frac{ik[(x - x')^2 + (y - y')^2]}{2f}\right\} dx' dy', \quad (35.13)$$

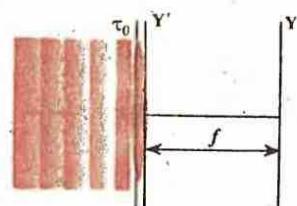
где пределы интегрирования распространены от $-\infty$ до ∞ , потому что функция зрачка P под интегралом автоматически обеспечивает интегрирование по площади апертуры.

Формула (35.13) с учетом (33.1) принимает вид



183

Плоская волна, выходящая из линзы в виде сферической расходящейся волны (при $f < 0$)



184

Образование изображения от предмета, помещенного непосредственно перед линзой

- Каким образом учитывается наличие диафрагмы на входе в линзу? Благодаря какому фактору именно в фокальной плоскости линзы, а не в какой-то другой, возникает образ Фурье распределения амплитуд на входе в линзу?

$$\Phi(x, y) = \frac{k}{2\pi i} \frac{\exp[i(kf + kn\Delta_0)]}{f} \exp\left[\frac{ik(x^2 + y^2)}{2f}\right] \times \\ \times \iint_{-\infty}^{\infty} \Psi'(x', y') P(x', y') \exp\left[-\frac{ik(xx' + yy')}{f}\right] dx' dy'. \quad (35.14)$$

Видно, что распределение амплитуд в фокальной плоскости с точностью до несущественных для дифракционной картины фазовых и масштабных множителей является образом Фурье распределения амплитуд на входе в линзу, т. е. линза является элементом, осуществляющим преобразование Фурье.

§ 36

. Дифракционное образование изображений линзой

Прослеживается процесс превращения дифракционной картины в изображение предмета. Обсуждается предел разрешающей способности оптических приборов.

Фурье преобразование амплитуд между фокальными плоскостями линзы. Изложенные в предыдущем параграфе соображения показывают, что в процессе распространения волны распределение амплитуд в плоскости, перпендикулярной направлению распространения, претерпевает изменение от плоскости к плоскости. Последовательно применяя формулы, описывающие эти изменения, можно найти формулы преобразования распределения амплитуд между двумя любыми плоскостями. Можно также найти распределение интенсивностей в этих плоскостях. Связь между распределениями амплитуд в общем случае получается довольно сложной, а распределение интенсивностей ничем не похожи друг на друга. Однако в определенных условиях связь между распределениями амплитуд оказывается достаточно простой и сводится в своей существенной части к преобразованию Фурье. Ясно, что наиболее простые случаи следует рассмотреть в первую очередь. Затем будут рассмотрены условия, при которых распределения интенсивностей в двух плоскостях достаточно хорошо похожи друг на друга. В этом случае говорят о дифракционном образовании изображения, поскольку все рассмотрение основывается на волновых понятиях без какого-либо обращения к лучам. Поместим плоский предмет с амплитудным коэффициентом пропускания $\tau_0(x_0, y_0)$ перед линзой на расстоянии L (рис. 185) и направим на него плоскую монохроматическую волну. На задней плоскости предмета образуется световое поле

$$\Psi''(x_0, y_0) = A_0 \tau_0(x_0, y_0), \quad (36.1)$$

где A_0 — амплитуда падающей на предмет плоской волны. Это поле в передней плоскости линзы создает дифракционную картину $\Psi'(x', y')$, которая с помощью формулы (32.37) представляется как

$$\Psi'(x', y') = \frac{k}{2\pi i} \frac{e^{ikL}}{L} \iint \Psi''(x_0, y_0) \exp\left\{-\frac{ik[(x' - x_0)^2 + (y' - y_0)^2]}{2L}\right\} dx_0 dy_0, \quad (36.2)$$

где пределами интегрирования по x_0, y_0 являются $-\infty, \infty$, а конечный размер предмета учитывается тем, что коэффициент пропускания $\tau_0(x_0, y_0)$ вне предмета полагается равным нулю. Если необходимо учесть вклад в дифракцию от волнового фронта вне предмета, то $\tau_0(x_0, y_0)$ для соответствующих частей волнового фронта следует принять равным единице.

Световое поле на выходе из линзы дается соотношением (13.12), а в фокальной плоскости — формулой (35.13), которая принимает вид

$$\Phi(x, y) = B \iint_{-\infty}^{\infty} \exp \left\{ -\frac{ik[(x-x')^2 + (y-y')^2]}{2f} \right\} dx' dy' P(x', y') \times \\ \times \exp \left[-\frac{ik(x'^2 + y'^2)}{2f} \right] \iint_{-\infty}^{\infty} \Psi''(x_0, y_0) \times \\ \times \exp \left\{ -\frac{ik[(x'-x_0)^2 + (y'-y_0)^2]}{2L} \right\} dx_0 dy_0, \quad (36.3)$$

где B — все постоянные множители перед интегралами использованных в (36.3) выражений.

Раскрывая в экспонентах подынтегрального выражения квадраты разностей координат, запишем (36.3) в виде

$$\Phi(x, y) = Be^{ik(x^2+y^2)/(2f)} \iint_{-\infty}^{\infty} P(x', y') \Psi''(x_0, y_0) e^{-ik(xx'+yy')/f} \times \\ \times e^{ik(x'^2+y'^2)/(2L)} e^{ik(x\delta+y\delta)/(2L)} e^{-ik(x'x_0+y'y_0)/L} dx' dy' dx_0 dy_0. \quad (36.4)$$

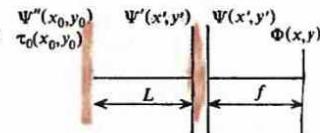
Функция P зрачка оказывает влияние на дифракционную картину. Если зрачок не очень мал, то $P=1$ для всех x', y' , потому что вне зрачка экспоненциальные функции, зависящие от этих переменных, сильно осциллируют и вклад в интеграл от области вне зрачка становится пренебрежимо малым. При этом условии интегрирование по dx' и dy' в (36.4) может быть выполнено с помощью известной из таблиц интегралов формулы

$$\int_{-\infty}^{\infty} \exp [i(a\xi^2 + 2b\xi)] d\xi = (1+i) \sqrt{\pi/(2a)} \exp(-ib^2/a).$$

В результате интегрирования (36.4) принимает вид

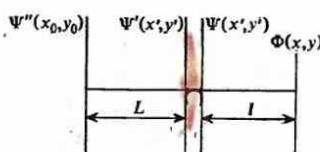
$$\Phi(x, y) = B_1 \exp \left[\frac{ik}{2f} \left(1 - \frac{L}{f} \right) (x^2 + y^2) \right] \times \\ \times \iint_{-\infty}^{\infty} \Psi''(x_0, y_0) \times \exp \left[-\frac{ik(xx_0+yy_0)}{f} \right] dx_0 dy_0, \quad (36.5)$$

где B_1 объединяет все постоянные множители. Это означает, что распределение амплитуд в фокальной плоскости линзы является Фурье-образом распределения амплитуд светового поля на поверхности объекта с точностью до несущественных фазового и масштабного множителей. Если объект расположен в передней фокальной плоскости линзы ($L=f$), то фазовый множитель пропадает. Это означает, что распределения амплитуд в фокальных плоскостях линзы связаны между собой преобразованием Фурье без каких-либо фазовых искажений, т. е. линза осуществляет преобразование Фурье между распределениями амплитуд светового поля в ее фокальных плоскостях.



185

Образование изображения от предмета, помещенного перед линзой на расстоянии L



186

К анализу формирования изображения линзой

Изображением называется дифракционная картина, распределение интенсивности света в которой с достаточной точностью воспроизводит распределение интенсивностей на изображаемом объекте.

Формирование изображения линзой. Для получения распределения амплитуд не в задней фокальной плоскости линзы, а в плоскости, расположенной на расстоянии l от линзы (рис. 186), воспользуемся формулами (36.1) — (36.3). В формуле (36.3) в первом экспоненциальном множителе подынтегрального выражения вместо расстояния f от линзы до задней фокальной плоскости должно теперь стоять расстояние l до плоскости, в которой определяется распределение амплитуд. Тогда [см. (36.4)]

$$\Phi(x, y) = B_2 \int_{-\infty}^{\infty} G(x, y, x_0, y_0) \Psi''(x_0, y_0) dx_0 dy_0, \quad (36.6)$$

где

$$G(x, y, x_0, y_0) = \exp\left[\frac{ik(x^2 + y^2)}{2l}\right] \exp\left[\frac{ik(x_0^2 + y_0^2)}{2L}\right] \int_{-\infty}^{\infty} P(x', y') \times \\ \times \exp\left[\frac{ik}{2}\left(\frac{1}{L} + \frac{1}{l} - \frac{1}{f}\right)(x'^2 + y'^2)\right] \exp\left\{-ik\left[\left(\frac{x_0}{L} + \frac{x}{l}\right)x' + \left(\frac{y_0}{L} + \frac{y}{l}\right)y'\right]\right\} dx' dy'. \quad (36.7)$$

Если распределение интенсивностей $|\Phi(x, y)|^2$ напоминает распределение интенсивностей $|\Psi''(x_0, y_0)|^2$ в плоскости предмета, то о дифракционной картине $|\Phi(x, y)|^2$ говорят как об изображении предмета. Оно может быть увеличенным или уменьшенным, прямым или перевернутым. Понятие изображения определено в геометрической оптике путем сопоставления каждой точке предмета точки изображения по правилам, рассмотренным в § 25. Поэтому вычислим (36.7) и (36.6) в приближении геометрической оптики:

$$\lambda \rightarrow 0, \quad k \rightarrow \infty. \quad (36.8)$$

Проанализируем интеграл, входящий в (36.7). Выберем расстояния L и l такими, чтобы член с экспонентой, в которую входит сумма $x'^2 + y'^2$, обратился в единицу. Для этого надо принять, что

$$1/L + 1/l - 1/f = 0. \quad (36.9)$$

Тогда

$$Q = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(x', y') \exp\left\{-ik\left[\left(\frac{x_0}{L} + \frac{x}{l}\right)x' + \left(\frac{y_0}{L} + \frac{y}{l}\right)y'\right]\right\} dx' dy'. \quad (36.10)$$

Обозначив

$$M = l/L, \quad (36.11)$$

представим (36.10) в форме

$$Q = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(x', y') \exp\left\{-\frac{ik}{l} [(x + Mx_0)x' + (y + My_0)y']\right\} dx' dy'. \quad (36.12a)$$

Вычислим этот интеграл в приближении геометрической оптики, когда $\lambda \rightarrow 0, k \rightarrow \infty$. Для этого перейдем в (36.12a) к новым переменным интегрирования: $\xi = (k/l)x'$, $\eta = (k/l)y'$. Тогда

● Точное изображение плоского объекта теоретически может быть получено лишь с помощью неограниченных пучков света. При наличии диафрагмы каждая точка объекта изображается дифракционной картиной диафрагмы. Это ограничивает предел разрешающей способности оптических приборов.

$$Q = \left(\frac{l}{k}\right)^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P[(l/k)\xi, (l/k)\eta] \exp[-i(x + Mx_0)\xi - i(y + My_0)\eta] d\xi d\eta. \quad (36.12 \text{б})$$

Аргументы у P при $k \rightarrow \infty$ для любых конечных значений ξ и η стремятся к нулю и, следовательно, $P \rightarrow 1$. Интеграл (36.12 б) при $P = 1$ выражается через δ -функцию:

$$Q = (l/k)^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp[-i(x + Mx_0)\xi - i(y + My_0)\eta] d\xi d\eta = 4\pi^2 (l/k)^2 \delta(x + Mx_0) \delta(y + My_0). \quad (36.12 \text{в})$$

Поэтому в пределе геометрической оптики (36.7) принимает вид

$$G(x, y, x_0, y_0) = 4\pi^2 (l/k)^2 \exp\left[\frac{-ik(x^2 + y^2)}{2l}\right] \times \\ \times \exp\left[\frac{-ik(x_0^2 + y_0^2)}{2L}\right] \delta(x + Mx_0) \delta(y + My_0). \quad (36.13)$$

Подставляя выражение (36.13) в (36.6) и интегрируя по $dx_0 dy_0$, получаем

$$\Phi(x, y) = B_4 \exp\left[\frac{-ik(x^2 + y^2)}{2l}\right] \exp\left[\frac{-ik(x^2 + y^2)}{2LM^2}\right] \times \\ \times \Psi''\left(-\frac{x}{M}, -\frac{y}{M}\right), \quad (36.14)$$

где в B_4 объединены все константы, несущественные для распределения амплитуд $\Phi(x, y)$.

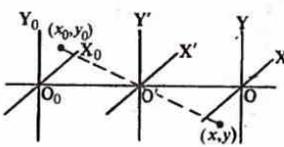
Из (36.14) получаем

$$|\Phi(x, y)|^2 = |B_4|^2 |\Psi''(-x/M, -y/M)|^2. \quad (36.15)$$

Формула (36.15) показывает, что в приближении геометрической оптики наблюдаемое в дифракционной картине распределение интенсивностей $|\Phi(x, y)|^2$ является перевернутым изображением распределения интенсивностей $|\Psi''|^2$ с увеличением M (рис. 187). Таким образом, изображение возникает при соблюдении условия (36.9), которое принято в геометрической оптике записывать в виде

$$1/L + 1/l = 1/f, \quad (36.16)$$

где L — расстояние от линзы до предмета, l — расстояние от линзы до изображения. Это означает, что формула (36.6) в приближении геометрической оптики описывает образование геометрических изображений в соответствии с правилами геометрической оптики.



187

К выводу условия образования изображения линзой



188

К расчету предельной разрешающей способности объектива телескопа

○ Чем отличаются изображения предмета, полученные методом фазового контраста, если дополнительный сдвиг фазы проходящего нимо объекта света составляет $\pi/2$ и $3\pi/2$?

Почему «метод темного поля» нельзя заменить «методом светлого поля», т. е. отмечать присутствие объекта не по рассеянному, а по задержанному им излучению?

Однако в случае конечных длин волн и учета конечных размеров апертуры линзы структура изображения усложняется. Каждая точка предмета уже не изображается точкой. Из сравнения (36.6), в котором в качестве G взято выражение (36.13), с (33.12а) заключаем, что изображение предмета искажается дифракцией Фраунгофера на апертуре линзы, учитываемой наличием функций зрачка $P(x', y')$ в формуле (36.12а). Каждая из точек предмета порождает в плоскости изображения картину дифракции Фраунгофера на апертуре линзы, т. е. точка изображается светлым пятном. Если отверстие круглое, то радиус центрального светлого пятна может быть вычислен с помощью (33.27).

Предел разрешающей способности оптических приборов. Разрешающая способность оптических приборов ограничивается дифракцией Фраунгофера на их входной апертуре, поскольку при этом каждая точка объекта изображается дифракционной картиной с центром в точке, соответствующей идеальному геометрическому изображению. Дифракционная картина состоит из центрального светлого пятна, окруженного дифракционными кольцами. **Изображением точки можно с достаточной точностью считать центральное светлое пятно. Две точки предмета будут разрешенными, если изображающие их светлые пятна четко разделены.** При этом в качестве критерия разрешения применяется критерий Рэлея.

Определим предельную разрешающую способность объектива телескопа. Предмет удален на бесконечность, а изображение предмета образуется в фокальной плоскости объектива с фокусным расстоянием f (рис. 188). На основании (33.27) заключаем, что радиус центрального светлого пятна равен

$$r_0 = 0,61\lambda/R, \quad (36.17)$$

где R — радиус объектива телескопа. Две близкие звезды, угловое расстояние между которыми θ , дают два дифракционных светлых пятна радиусом r_0 , расстояние между центрами которых $\Delta = \theta_0 f$. По критерию Рэлея они считаются разрешенными, если центр одного попадает на край другого, т. е. когда $\Delta = r_0$ и, следовательно,

$$\theta_0 = 0,61\lambda/R. \quad (36.18)$$

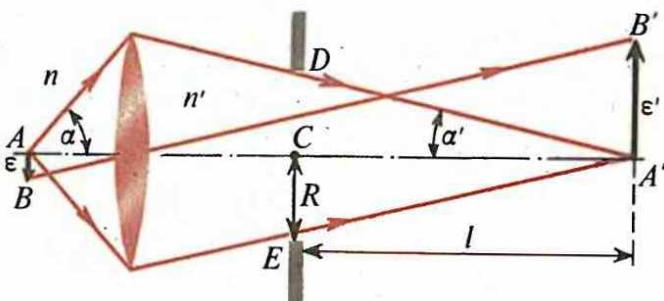
Разрешающей способностью Y объектива называется величина, обратная θ_0 :

$$Y = 1/\theta_0 = R/(0,61\lambda). \quad (36.19)$$

Для увеличения разрешающей способности необходимо увеличивать размер апертуры объектива.

Если угловое расстояние между точками меньше θ_0 , то они воспринимаются как одна точка и, следовательно, детали предмета с меньшими угловыми размерами не могут быть различими.

Формула (36.19) для разрешающей способности объектива получена на основании соотношения (33.27), которое было выведено в предположении, что амплитуда и фаза падающей волны постоянны во всех точках отверстия объектива (однородная апертура). При этом условии единственным способом увеличения разрешающей способности при фиксированной длине волны является увеличение радиуса объектива. Однако радиус центрального дифракционного максимума в дифракционной картине на круглом отверстии фиксированного радиуса может быть уменьшен специальным подбором распределения амплитуд и фаз излучения в плоскости объектива, вследствие чего увеличивается разрешающая способность объектива. Однако при этом интенсивность центрального максимума уменьшается. Следовательно, если допустимо уменьшение яркости изображения, то разрешающую способность объектива можно увеличить без увеличения его радиуса за счет соответствующей фазово-амплитудной модуляции падающего на объектив света.



Аналогично можно найти и предел разрешающей способности микроскопа, ход лучей в котором показан на рис. 189. Для исключения сферической aberrации (см. § 24) необходимо соблюдение условия Аббе (24.17), которое в обозначениях рис. 189 записывается в виде

$$\epsilon n \sin \alpha = \epsilon' n' \sin \alpha', \quad (36.20)$$

где n и n' — показатели преломления сред соответственно в пространстве предмета и пространстве изображения. Обозначая l — расстояние от плоскости отверстия до плоскости изображения и R — радиус входного зрачка, получаем, что радиус центрального светлого пятна дифракционного изображения точки равен

$$r_0 = l\phi = 0,61\lambda l/R, \quad (36.21)$$

где $R = DE/2$.

С помощью объектива микроскопа осуществляется первая стадия образования изображения. Изображение предмета, полученное с помощью объектива, затем увеличивается посредством окуляра. Однако ясно, что разрешающая способность определяется только первой стадией образования изображения. Поэтому на рис. 189 показан ход лучей только в объективе микроскопа.

Точки A и B будут различимы, если по критерию Рэлея расстояние между точками A' и B' их геометрического изображения больше или равно r_0 . Следовательно, максимальная разрешающая способность системы (рис. 189) определяется из условия

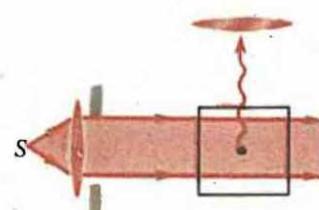
$$\epsilon' = r_0 = 0,61\lambda l/R. \quad (36.22)$$

Из (36.20) получаем

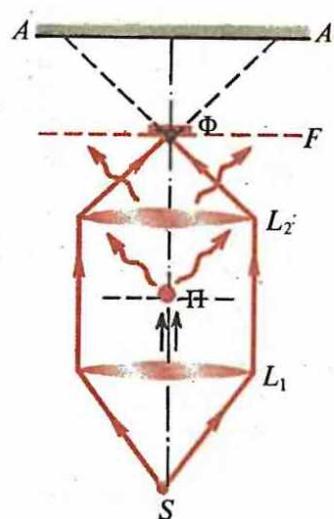
$$\epsilon = \epsilon' n' \sin \alpha' / (n \sin \alpha) = n' \lambda 0,61 / (n \sin \alpha), \quad (36.23)$$

где $R/d \approx \alpha \approx \sin \alpha'$. Формула (36.23) определяет максимальную разрешающую способность микроскопа. Так как обычно в пространстве изображения $n' = 1$, коэффициент преломления жидкостей, используемых для иммерсионных объективов (см. § 24), $n \approx 1,5$, а угол α близок к $\pi/2$, то (36.23) можно записать так: $\epsilon \approx 0,4\lambda$,

К расчету предельной разрешающей способности микроскопа



Метод темного поля



Метод фазового контраста

$$\varepsilon \approx 0,15 \div 0,3 \text{ мкм}.$$
(36.25)

Более мелкие детали в видимом диапазоне наблюдать нельзя.

Метод темного поля. Детали предметов, меньших значений ε , определяемых формулой (36.25), в оптическом диапазоне различать нельзя, но детектировать наличие частичек такого или меньшего размера можно. Это делается методом темного поля (рис. 190).

Наблюдаемые частицы освещаются белым светом от источника S так, чтобы прямые лучи не попадали в глаз наблюдателя, фокусирующего микроскоп на точку, в которой ожидается присутствие частицы. Дифрагированные частицей лучи попадают в микроскоп и фиксируются глазом в виде светящихся точек, форма которых не имеет ничего общего с формой частицы.

Имеются и другие схемы наблюдения, в которых прямые лучи устраниются из поля зрения специальными конденсорами, но физическая суть метода темного поля остается, конечно, неизменной.

Метод фазового контраста. Тонкие прозрачные объекты очень слабо поглощают свет и не создают сколько-нибудь заметной амплитудной модуляции проходящего света, в результате чего они оказываются практически невидимыми. Их удается наблюдать, воспользовавшись модуляцией фазы проходящего через них света. Наблюдение с использованием модуляции фазы называется **методом фазового контраста**.

Свет от точечного источника S (рис. 191) с помощью линзы L_1 превращается в пучок параллельных лучей. Большая часть пучка проходит мимо предмета P , а меньшая — через предмет. Амплитуду света, прошедшего мимо предмета, примем за 1, а амплитуду света, дифрагирующего на предмете, обозначим $\varepsilon \ll 1$, поскольку на предмет попадает лишь очень малая часть всего потока. Проходящая через предмет волна испытывает изменение фазы на δ , и, следовательно, ее комплексная амплитуда равна $\varepsilon e^{i\delta}$. В фокальной плоскости линзы L_2 образуется дифракционная картина, которая в плоскости AA' трансформируется в изображение предмета P .

Свет, проходящий мимо предмета, не испытывает изменения фазы, и поэтому полная интенсивность в плоскости изображения

$$2I = |1 + \varepsilon e^{i\delta}|^2 = 1 + \varepsilon^2 + 2\varepsilon \cos \delta \approx (1 + \varepsilon)^2 - \varepsilon \delta^2,$$
(36.26)

где в разложении $\cos \delta$ в ряд по δ первый член, зависящий от δ , имеет порядок δ^2 , т. е. чрезвычайно мал. Поэтому модуляция фазы не приводит к модуляции интенсивности, которую можно наблюдать.

Однако можно добиться того, чтобы модуляция интенсивности резко возросла и была пропорциональна величине модуляции фазы. Для этого свет, идущий мимо предмета, пропускают через пластину, которая изменяет его фазу на $\pi/2$, т. е. его комплексная амплитуда становится равной $e^{i\pi/2} = i$. Амплитуда проходящего через предмет света по-прежнему $\varepsilon e^{i\delta}$. Интенсивность интерференционной картины в плоскости изображения

$$2I = |i + \varepsilon e^{i\delta}|^2 = 1 + \varepsilon^2 + 2\varepsilon \sin \delta \approx 1 + \varepsilon^2 + 2\varepsilon \delta,$$
(36.27)

т. е. модуляция интенсивности линейно зависит от изменения фазы и позволяет наблюдать фазовый объект.

Чтобы увеличить контраст, надо уменьшить амплитуду света, минующего объект, сделав ее соизмеримой с амплитудой света, прошедшего через объект. Обозначим $\varepsilon' \ll 1$ уменьшенное значение амплитуды. Интенсивность наблюдаемой в плоскости изображения картины становится равной

$$2I = |\varepsilon' + \varepsilon e^{i\delta}|^2 = \varepsilon'^2 + \varepsilon'^2 + 2\varepsilon' \varepsilon \sin \delta,$$
(36.28)

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = \frac{\varepsilon\varepsilon'(\sin \delta_{\max} - \sin \delta_{\min})}{\varepsilon^2 + \varepsilon'^2 + \varepsilon\varepsilon'(\sin \delta_{\max} + \sin \delta_{\min})}, \quad (36.29)$$

где δ_{\max} и δ_{\min} — максимальное и минимальное изменения фазы в объекте. При $\varepsilon' = \varepsilon$ видимость и контраст изображения существенно увеличиваются.

Если фазу света, минущего предмет, изменить на $3\pi/2$, т. е. если комплексная амплитуда ослабленного света станет равной $\varepsilon' \exp(3\pi/2) = -i\varepsilon'$, то вместо (36.28) находим

$$2I = \varepsilon^2 + \varepsilon'^2 - 2\varepsilon\varepsilon' \sin \delta, \quad (36.30)$$

что дает контрастность такой же величины, но другого знака: что казалось темным, будет светлым, и наоборот.

При практической реализации метода фазового контраста свет, проходящий мимо предмета, ослабляется и испытывает изменение фазы на $\pi/2$ или $3\pi/2$ не в плоскости предмета, а в фокальной плоскости. Для этого в фокальной плоскости центральная часть закрывается небольшой пластиной Φ (рис. 191), которая одновременно уменьшает амплитуду света примерно до ε' и изменяет фазу на $\pi/2$ или $3\pi/2$. Толщина пластины подбирается экспериментально, по максимальному контрасту.

Метод фазового контраста был разработан Ф. Цернике в 1935 г. и явился важным шагом на пути к голограммии.

Пример 36.1. Определить разрешающую способность глаза. Радиус входного зрачка глаза $R = 1,5$ мм, показатель преломления n внутриглазной жидкости равен примерно показателю преломления воды.

Если бы все пространство было заполнено жидкостью с показателем преломления, равным показателю преломления глазной жидкости, то разрешающая способность глаза выражалась бы формулой (36.18), но с заменой длины волны в вакууме на длину волны λ' в жидкости:

$$\theta_0 = 0,61\lambda'/R, \quad (36.31)$$

где $\lambda' = \lambda/n$. Отметим, что θ_0 — угол между лучами на центр светлого пятна и его границу внутри глаза, равный минимальному углу между лучами, входящими в глаз из пространства, заполненного жидкостью. Из геометрической оптики известно, что если по разные стороны линзы находятся среды с показателями преломления n и n' , то малый угол θ между соседними лучами после прохождения линзы изменяется и становится равным

$$\theta' = \theta n/n'. \quad (36.32)$$

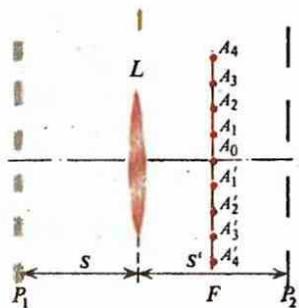
Если в пространстве вне глаза $n' = 1$ (воздух), то угол (36.31) увеличивается в n раз и, следовательно, минимально разрешимый угол θ_0 между приходящими в глаз лучами также увеличивается в n раз:

$$\theta_0 = n\theta' = 0,61\lambda/R. \quad (36.33)$$

Таким образом, разрешающая способность глаза не зависит от показателя преломления внутриглазной жидкости. Для $\lambda = 550$ нм находим $\theta_0 = 2,2 \cdot 10^{-4} = 0,8'$. Полезно заметить, что если исходить из структуры воспринимающих свет элементов глаза (палочки и колбочки), то минимально разрешимый угол также будет примерно равным $1'$. Следовательно, различные элементы глаза оптимально согласованы между собой.

Подчеркнем, что «экономия» в этом примере означает именно оптимальное согласование различных элементов глаза и отсутствие излишних неиспользованных возможностей, а отнюдь не «экономию» на создание возможностей и надежности функционирования системы в целом. Надежность жизненно важных систем всегда обеспечивается полностью, часто за счет дублирования функций, поскольку отсутствие надежности лишает организм возможности участвовать в естественном отборе и развиваться.

Излагается суть пространственной фильтрации изображений и ее применения.



192

К объяснению сущности пространственной фильтрации изображения

● **Пространственная фильтрация изображений основывается на возможности изменить изображение, воздействуя определенным образом на дифракционную картину предмета в фокальной плоскости линзы.**

● **Какие максимумы интенсивности дифракционной картины в эксперименте Аббе — Портера надо закрыть, чтобы в изображении сохранились лишь горизонтальные линии? В каких максимумах интенсивности интерференционной картины содержится информация о более мелких деталях объекта?**

Сущность пространственной фильтрации изображений. Дифракционное образование изображения (см. § 36) сводится к двум стадиям:

1) формированию в фокальной плоскости линзы дифракционной картины предмета;

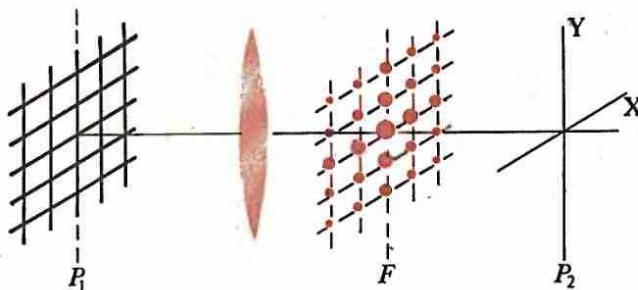
2) преобразованию дифракционной картины предмета в фокальной плоскости линзы в изображение предмета в плоскости изображений.

Вся информация, имеющаяся в изображении предмета, содержится также и в дифракционной картине предмета в фокальной плоскости. Если произвести в фокальной плоскости изменения дифракционной картины, например закрыть или, наоборот, усилить некоторые максимумы, то произойдет соответствующее изменение в изображении предмета. Внесение изменений в изображение предмета посредством модификации дифракционной картины предмета, из которой в последующем формируется изображение, называется **пространственной фильтрацией изображения**.

Пространственная фильтрация изображений дифракционной решетки. Наиболее наглядно существо пространственной фильтрации можно проиллюстрировать на примере изображения дифракционной решетки.

Дифракционная решетка P_1 (рис. 192), расстояние s которой до линзы больше фокусного, освещается параллельным пучком лучей. Линза L создает в плоскости P_2 , расстояние s' которой от линзы определяется соотношением (23.36б), действительное перевернутое изображение решетки, состоящее из темных и светлых полос с достаточно резкими границами. Резкость границ зависит от ширины щелей и непрозрачных промежутков между ними: чем больше ширина, тем резче граница. В фокальной плоскости F линзы создается картина дифракции Фраунгофера на решетке. Если ее наблюдать на реальной плоской поверхности, то она представляется в виде дифракционных полос (см. § 33). Если эта поверхность воображаемая, и лучи беспрепятственно проходят дальше, то в плоскости P_2 образуется изображение решетки.

На рис. 192 главные максимумы интенсивности дифракционной картины обозначены $A_0, A_1, A_2, \dots, A'_1, A'_2, \dots$. Наиболее полное изображение решетки получается тогда, когда вся дифракционная картина в плоскости F участвует в образовании изображения в плоскости P_2 . Если часть максимумов задержать, то изображение ухудшится, пропадут некоторые детали или даже полностью исказится изображение решетки. Например, устраним все максимумы, за исключением A_0 , т. е. сведем дифракционную картину к одному центральному мак-



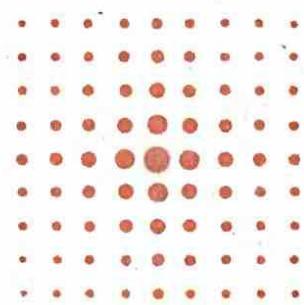
193
Схема эксперимента Аббе — Портера

симуму. Это соответствует лучу, прошедшему при отсутствии препятствий на его пути. В плоскости P_2 наблюдается равномерное освещение, изображение дифракционной решетки отсутствует, ее как бы не существует. Устраним все нечетные максимумы $A_1, A_3, \dots, A'_1, A'_3, \dots$. Оставшиеся максимумы $A_0, A_2, A_4, \dots, A'_2, A'_4, \dots$ соответствуют картине дифракции на решетке, период которой вдвое меньше. Поэтому в плоскости изображений возникает изображение дифракционной решетки с вдвое меньшим периодом, т. е. более частая решетка. Максимумы первых порядков определяют более крупные детали объекта, а информация о более мелких деталях передается через максимумы более высоких порядков. Если закрыть все максимумы, за исключением нулевого и первого порядков, то в плоскости изображений получается распределение амплитуд светового поля по гармоническому закону [см. (33.58)].

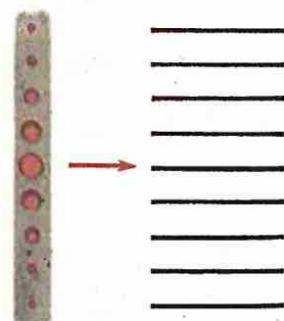
Изменение изображения посредством воздействия не на самоизображение, а на распределение амплитуд, из которых в последующем синтезируется изображение, составляет суть пространственной фильтрации. Основная задача при этом состоит в создании фильтра, который нужным образом изменяет проходящую через него волну.

Эксперимент Аббе — Портера. Первым, продемонстрировавшим возможность пространственной фильтрации изображений, был эксперимент Аббе — Портера, в котором в качестве предмета была взята плоская решетка, образованная штрихами, пересекающимися под прямым углом (рис. 193). Картина дифракции, возникающая в фокальной плоскости линзы, показана на рис. 194. В плоскости изображений из этой дифракционной картины образуется изображение решетки.

Если непрозрачной пластиною с щелью закрыть в фокальной плоскости все максимумы интенсивности, за исключением центрального и находящихся с ним на одной линии, параллельной одной из систем штрихов решетки (рис. 195), то в изображении сохраняется лишь перпендикулярная система штрихов решетки. Если закрыть центральный максимум, то получается изображение решетки с обращенным контрастом — на месте светлых линий возникают темные, и наоборот.



194
Картина дифракции на пространственной решетке



195
Фильтрация изображения пространственной решетки

Излагаются принципы получения тонкослойных и толстослойных голограмм и методы восстановления изображений. Обсуждается применение голографии.

Синхронное детектирование. Фотопластинка регистрирует среднее значение квадрата напряженности поля электромагнитной световой волны, т. е. интенсивность. Информация о фазе волны при этом теряется. Таким образом, содержащаяся в фотографии информация об объекте весьма ограничена. В частности, отсутствует информация о расстояниях различных частей объекта от фотопластинки и других важных характеристиках. Другими словами, обычная фотография не позволяет восстановить полностью тот волновой фронт, который на ней был зарегистрирован. В обычной фотографии содержится более или менее точная информация об амплитудах зафиксированных волн, но полностью отсутствует информация о фазах.

Голография позволяет устранить этот недостаток обычной фотографии и записать на фотопластинке информацию не только об амплитудах падающих на нее волн, но и о фазах, т. е. полную информацию. Волна, восстановленная с помощью такой записи, полностью идентична первоначальной, содержит в себе всю информацию, которую содержала первоначальная. Поэтому метод назван голографией, т. е. методом полной записи волн.

Идея метода аналогична синхронному детектированию, известному в радиотехнике. Информация передается в виде модулированного сигнала на высокой несущей частоте:

$$V_m = V_{0m} [1 + m f(t)] \cos \omega t, \quad (38.1)$$

где ω — несущая частота, $f(t)$ — передаваемый сигнал, m — глубина модуляции (рис. 196).

В модулированном сигнале (38.1) содержится информация не только об амплитудном, но и о фазовом спектре функции $f(t)$. Эту информацию можно выделить с помощью синхронного детектирования.

В радиотехнике известны устройства, которые позволяют перемножать напряжения подаваемых на них сигналов и получать на выходе их произведения. Если в такое устройство подать модулированный сигнал (38.1) и постоянный сигнал несущей частоты

$$V_n = V_{0n} \cos \omega t, \quad (38.2)$$

то на выходе получим сигнал

$$V_{\text{вых}} = a V_m V_n = a V_{0m} V_{0n} [1 + m f(t)] (1 + \cos 2\omega t)/2, \quad (38.3)$$

где a — постоянная. Высокочастотная составляющая с помощью соответствующего фильтра устраняется, а оставшийся низкочастотный сигнал

$$V_{\text{вых}}^{(n)} = A [1 + m f(t)] \quad (38.4)$$

содержит полную информацию об амплитудном и фазовом спектрах модулированного сигнала.

Логично использовать подобный синхронному детектированию прием и в световом диапазоне частот. Для этого необходимо информацию, содержащуюся в световой волне, записать в виде модулированной волны на некоторой несущей частоте. При записи модулированной волны на фотопластинку информация о фазе записываемой волны сохраняется. По записи на фотопластинке модулированной волны синхронным детектированием, т. е. ее облучением несущей частотой, можно восстановить исходную волну. Задача о полном восстановлении волнового фронта тем самым принципиально будет решена.

Для того чтобы осуществить этот метод в световом диапазоне, необходимо иметь излучение с достаточно высокой степенью когерентности. Такое излучение получается с помощью лазеров. Поэтому только после создания лазеров, дающих излучение с высокой степенью когерентности, удалось практически осуществить голографию.

Фотопластинка, на которой записана информация о модулированной световой волне, называется **гологриммой**. Задача голографии состоит в разработке методов записи голограмм и восстановлении по ним волнового фронта.

Идея голографии была выдвинута в 1920 г. польским физиком М. Вольфке (1883—1947), но была забыта. В 1947 г. независимо от Вольфке идею голографии предложил и обосновал английский физик Д. Габор, удостоенный за это в 1971 г. Нобелевской премии.

Голограмма плоской волны. Волна с несущей частотой называется **опорной**, а волна, содержащая информацию об объекте, — **сигнальной**.

Плоская сигнальная волна частотой ω (рис. 197) распространяется в положительном направлении оси Z , перпендикулярно которой в плоскости XY расположена фотопластинка. Опорная волна образуется делением волнового фронта и с помощью призмы P направляется на фотопластинку, перекрываясь на ней с сигнальной волной, также возникающей при делении волнового фронта. Угол наклона опорной волны с осью Z обозначен θ .

Сигнальная и опорная волны записываются в виде

$$E_c = E_1 e^{-i(\omega t - kz)}, \quad (38.5)$$

$$E_{\text{оп}} = E_0 e^{-i(\omega t - k_z z - k_x x)} \quad (38.6)$$

Следует заметить, что выражения (38.5) и (38.6) записаны с точностью до постоянной фазы. Можно было бы считать, что E_1 и E_0 в этих формулах являются комплексными и содержат в себе не выписанные в явном виде фазы. Однако это лишь усложняет написание последующих формул и не содержит в себе какой-либо существенной информации. Поэтому будем считать E_1 и E_0 вещественными.

Учитывая, что $k_x = k \sin \theta$, $k_z = k \cos \theta \approx k$ с точностью до величин второго порядка по углу θ , который предполагается малым, можем (38.6) представить в виде

$$E_{\text{оп}} = E_0 e^{-i(\omega t - kz - kx \sin \theta)}. \quad (38.7)$$

Полная амплитуда напряженности электрического поля в плоскости фотопластинки равна

$$E = E_c + E_{\text{оп}} = e^{-i(\omega t - kz)} (E_1 + E_0 e^{ikx \sin \theta}), \quad (38.8)$$

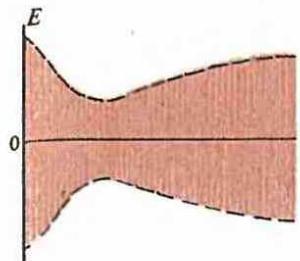
где E_1 и E_0 можно считать действительными амплитудами. Отсюда для распределения интенсивностей находим выражение

$$I(x) = \frac{1}{2} |E|^2 = \frac{1}{2} [E_1^2 + E_0^2 + 2E_0 E_1 \cos(kx \sin \theta)], \quad (38.9)$$

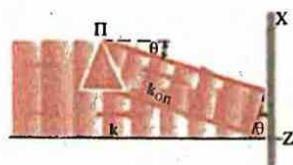
которое свидетельствует о том, что на фотопластинке в этом случае также записана и разность фаз между опорной и сигнальной волнами, т. е. фаза сигнальной волны, если считать фазу опорной волны заданной.

Восстановление изображения. Фотоэмulsionия состоит из частиц галоидного серебра, рассеянных в желатине. Все это находится на подложке из стекла или ацетата. При попадании света на частицу галоидного серебра в ней возникают центры восстановленного серебра. Это центры проявления. При проявлении частицы, в которых имеются центры проявления, восстанавливаются до металлического серебра. Там, где нет центров проявления, частицы остаются галоидными. После проявления при «фиксации» частицы галоидного серебра удаляются и в пластинке остается лишь металлическое серебро в мелких частицах, которые образуют почернение пластиинки. В теории фотографического процесса показывается, что плотность почернения пластиинки равна

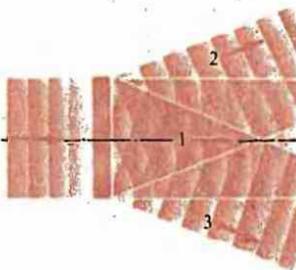
$$Q = \gamma \log |E|^2, \quad (38.10)$$



196 Амплитудно-модулированный сигнал



197 Схема записи голограммы плоской волны



198 Восстановление волны, записанной на голограмме

где коэффициент контрастности γ характеризует материал фотопластинки, а после проявления коэффициент пропускания пластины имеет вид

$$\tau = (|E|^2)^{-\gamma/2} \quad (38.11)$$

С учетом (38.9) из (38.11) находим

$$\tau = [E_0^2 + E_1^2 + 2E_0E_1 \cos(kx \sin \theta)]^{-\gamma/2}. \quad (38.12)$$

Принимая во внимание, что в обычных условиях $E_1 \ll E_0$, вместо (38.12) можно записать

$$\begin{aligned} \tau &= E_0^{-\gamma} [1 - \gamma E_1^2 / 2E_0^2 - (\gamma E_1 / E_0) \cos(kx \sin \theta)] = \\ &= (E_0^{-\gamma-2}/2) [2E_0^2 - \gamma E_1^2 - 2\gamma E_0 E_1 \cos(kx \sin \theta)]. \end{aligned} \quad (38.13)$$

Отбрасывая несущественный для дальнейшего рассмотрения масштабный множитель $(E_0^{-\gamma-2}/2)$, запишем выражение (38.13) в более удобном виде:

$$\tau = 2E_0^2 - \gamma E_1^2 - \gamma E_0 E_1 e^{ikx \sin \theta} - \gamma E_0 E_1 e^{-ikx \sin \theta}. \quad (38.14)$$

Направим на голограмму по пути сигнальной волны плоскую волну (рис. 198)

$$E_{\text{вос}} = E_2 e^{-i(\omega t - kz)}. \quad (38.15)$$

На выходе из голограммы возникает световое поле

$$\begin{aligned} E_{\text{вых}} &= E_{\text{вос}} \tau = E_2 (2E_0^2 - \gamma E_1^2) e^{-i(\omega t - kz)} - \\ &- \gamma E_0 E_1 E_2 e^{-i(\omega t - kz + kx \sin \theta)} - \gamma E_0 E_1 E_2 e^{-i(\omega t - kz - kx \sin \theta)} \end{aligned} \quad (38.16)$$

состоящее из трех плоских волн (рис. 198):

1) первый член в (38.16) описывает плоскую волну 1, распространяющуюся в направлении положительных значений оси Z как продолжение волны, падающей на голограмму;

2) второй член описывает волну 2, распространяющуюся под углом θ к оси Z с наклоном в сторону положительных значений оси X . Это видно из сравнения знаков $u kx \sin \theta$ в экспонентах волн 2 и волны, описываемой уравнением (38.7), распространяющейся с наклоном в сторону отрицательных значений оси X ;

3) третий член описывает волну 3, распространяющуюся под углом θ к оси Z с наклоном в сторону отрицательных значений оси X .

Это световое поле представляет собой дифракцию плоской волны, падающей на голограмму. Видно, что возникает лишь дифракция первого порядка, как это и должно быть, когда коэффициент пропускания (38.14) изменяется по гармоническому закону [ср. с (33.64)].

Голограмма точечного объекта. Схема получения голограммы показана на рис. 199. Опорная волна образуется после преломления в призме. Она описывается аналогично (38.7) формулой

$$E_{\text{оп}} = E_0 e^{-i(\omega t - kz - kx \sin \theta)}. \quad (38.17)$$

От точечного объекта исходит расходящаяся сферическая волна, которую в плоскости фотопластинки можно представить в виде

$$E_c = E_1 e^{-i(\omega t - kz - kx^2/(2z_0))}, \quad (38.18)$$

причем небольшим изменением амплитуды напряженности при удалении от точечного объекта при движении вдоль фотопластинки пренебрегаем.

Заметим, что в формулах (38.17) и (38.18) не выписаны в явном виде постоянные фазы, поскольку они не имеют какого-либо существенного физического значения. Наличие постоянной разности фаз между волнами приводит лишь к некоторому небольшому сдвигу дифракционной картины в пространстве, не изменяя существенно эту картину, и ее нет необходимости учитывать. Поэтому в выражениях для плоских волн начало отсчета системы координат не имеет значения, но для сферической волны (38.18), представляемой в экспоненте слагаемыми $-ikx^2/(2z_0)$, необходимо помнить, что начало отсчета совпадает с отверстием в экране. Поэтому в знаменателе этого выражения использовано обозначение z_0 , поскольку нас интересует дифракционная картина в плоскости пластиинки, находящейся на расстоянии z_0 от отверстия.

Полная напряженность поля на пластиинке

$$E = E_{\text{оп}} + E_c \quad (38.19)$$

принимает вид

$$E = e^{-i(\omega t - kz)} [E_0 e^{ikx \sin \theta} + E_1 e^{ikx^2/(2z_0)}], \quad (38.20)$$

откуда

$$\begin{aligned} |E|^2 &= E_0^2 + E_1^2 + E_0 E_1 (e^{i[kx \sin \theta - kx^2/(2z_0)]} + \\ &+ e^{-i[kx \sin \theta - kx^2/(2z_0)]}) = \\ &= E_0^2 + E_1^2 + 2E_0 E_1 \cos [kx \sin \theta - kx^2/(2z_0)]. \end{aligned} \quad (38.21)$$

Видно, что фазовые соотношения между волнами зафиксированы на пластиинке. Для коэффициента пропускания так же как в (38.19) получается выражение

$$\tau = 2E_0^2 - \gamma E_1^2 - \gamma E_0 E_1 e^{i(kx \sin \theta - kx^2/(2z_0))} - \gamma E_0 E_1 e^{-i(kx \sin \theta - kx^2/(2z_0))} \quad (38.22)$$

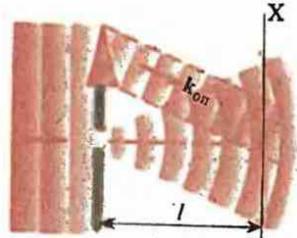
и поэтому при облучении голограммы плоской волной

$$E_{\text{вос}} = E_2 e^{-i(\omega t - kz)} \quad (38.23)$$

аналогично (38.16) на выходе из голограммы образуется световое поле

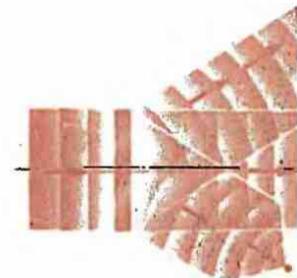
$$\begin{aligned} E_{\text{вых}} &= E_{\text{вос}} \tau = E_2 (2E_0^2 - \gamma E_1^2) e^{-i(\omega t - kz)} - \\ &- \gamma E_0 E_1 E_2 e^{-i[\omega t - kz - kx \sin \theta + kx^2/(2z_0)]} - \\ &- \gamma E_0 E_1 E_2 e^{-i[\omega t - kz + kx \sin \theta - kx^2/(2z_0)]}, \end{aligned} \quad (38.24)$$

состоящее из трех волн (рис. 200):



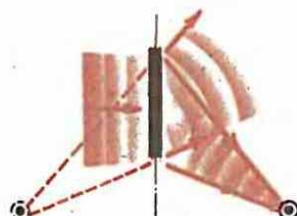
199

Схема получения голограммы точечного объекта



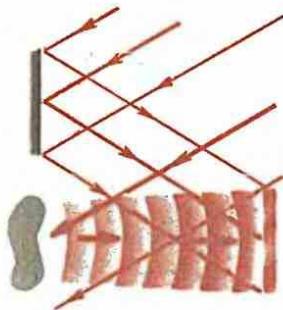
200

Восстановление волнового фронта от точечного объекта



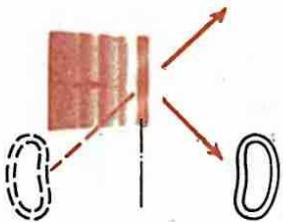
201

Действительное и мнимое изображения точки при восстановлении голограммы



202

Одна из схем записи голограммы



203

Восстановление действительного и
мнимого изображений предмета

1) первый член (в 38.24) представляет плоскую волну, распространяющуюся в том же направлении, что и волна, падающая на голограмму;

2) второй член в (38.24) представляет волну, распространяющуюся с наклоном к оси Z под углом θ в направлении отрицательного значения оси X (в экспоненте стоит $ikx \sin \theta$). Эта волна сферическая, причем вогнутость направлена в сторону распространения [в экспоненте стоит $-ikx^2/(2z_0)$], т. е. является сходящейся сферической волной. Она сходится в фокусе, расположенному в центре кривизны поверхности. Фокус служит действительным изображением точечного объекта. Видно, что он является зеркальным изображением точечного объекта, зафиксированного на голограмме, если плоскость зеркала совпадает с плоскостью голограммы, а точечный объект расположен относительно голограммы так, как он был расположен во время записи голограммы (рис. 201);

3) третий член в (38.24) представляет волну, распространяющуюся с наклоном к оси X под углом θ в направлении положительных значений оси X (в экспоненте стоит $-ikx \sin \theta$). Эта волна сферическая, с вогнутостью против направления распространения волны [в экспоненте стоит $ikx^2/(2z_0)$], т. е. является расходящейся сферической волной. Она абсолютно идентична сферической волне, распространявшейся от точечного объекта во время записи голограммы, и дает мнимое изображение точечного объекта, находящееся в той же точке, в которой был точечный объект во время записи голограммы. Если волна попадает в глаз человека, то возникающее ощущение идентично тому, которое возникает при попадании в глаз расходящейся сферической волны непосредственно от точечного объекта. Значит, голографическая запись позволяет полностью восстановить волновой фронт.

Голограмма произвольного объекта. Волновой фронт произвольного объекта складывается из волновых фронтов, рожденных его точками (см. § 31, 34, 36). Поэтому запись голограммы произвольного объекта осуществляется аналогично предыдущему случаю, необходимо лишь, чтобы сигнальная несущая информацию об объекте волна была когерентна с опорной. Для этого надо объект освещать волной когерентной с опорной. Это можно сделать различными способами.

На рис. 202 показан один из возможных способов записи голограммы. Видно, что опорная волна образуется отражением от зеркала той же волны, которая освещает голографируемый объект. Приходящая от объекта на фотопластинку волна может быть в плоскости пластины представлена в виде

$$E_c(x, y) = E_1(x, y) e^{-i[\omega t - kz - \varphi(x, y)]} \quad (38.25)$$

где $E_1(x, y)$ и $\varphi(x, y)$ описывают распределение амплитуд и фаз светового поля от источника. Опорная волна представляется в виде (38.19). Для квадрата модуля напряженности электричес-

ского поля, записанного на голограмме, вместо (38.21) получается выражение

$$|E|^2 = E_0^2 + E_1^2 + E_0 E_1 (e^{i(kx \sin \theta - \varphi)} + e^{-i(kx \sin \theta - \varphi)}). \quad (38.26)$$

При облучении голограммы монохроматической волной получаем аналогично (38.24) выражение для восстановленной волны

$$E_{\text{внж}} = E_{\text{вос}} \tau = E_2 (2E_0^2 - \gamma E_1^2) e^{-i(\omega t - kz)} - \gamma E_0 E_2 E_1 e^{-i(\omega t - kz - kx \sin \theta + \varphi)} - \gamma E_0 E_2 E_1 e^{-i(\omega t - kz + kx \sin \theta - \varphi)}, \quad (38.27)$$

которое аналогично (38.24) описывает проходящую волну и волны, дающие действительное и мнимое изображения предмета (рис. 203).

Волна, обуславливающая мнимое изображение предмета, является точной реконструкцией волны, исходящей непосредственно от предмета. Это мнимое изображение является объемным, и поэтому, изменяя угол зрения, можно посмотреть на предмет несколько сбоку. При перемещении головы видны боковые части предмета. Предмет можно также сфотографировать под различными ракурсами при условии, конечно, что объектив фотоаппарата находится в пределах реконструированной волны.

Сигнальная и опорная волны при записи голограммы должны быть когерентными между собой. Ширина когерентности должна быть, во всяком случае, не меньше размеров предмета, а длина когерентности — не менее разности хода сигнальной и опорной волн. В реальных условиях это означает, что при записи голограммы необходимо использовать излучение с высокой степенью временной и пространственной когерентности. Этим требованиям отвечает лазерное излучение. Восстанавливаются голограммы также с помощью лазеров. Однако при восстановлении голограммы частота лазерного излучения может отличаться от частоты, использованной при записи голограммы. Это следует из того факта, что восстановление голограммы сводится к дифракции падающей на голограмму волны. При увеличении длины волны дифракционные углы увеличиваются. Поэтому при восстановлении голограммы излучением с большей, чем при записи, длиной волны изображение увеличивается по сравнению с оригиналом.

Реконструкцию голограммы можно также осуществить и без лазера. Достаточно малый некогерентный источник, видимый под углом когерентности (27.24), создает на ширине когерентности (27.25) излучение с достаточно высокой степенью когерентности. Например, светящаяся булавочная головка с расстояния вытянутой руки создает на зрачке глаза световое поле с высокой степенью когерентности. Поэтому если голограмму поместить между светящейся булавочной головкой на вытянутой руке и глазом, то можно видеть восстановление голограммического изображения предмета, записанное на голограмме. Отличие от изображения, восстановленного с помощью лазера, состоит в меньшей четкости, т. е. в потере дифракционных максимумов высших порядков. Объемность изображения сохраняется.

Требования к фотопластинкам и времени экспозиции. Фотопластинка при записи голограммы так же, как при записи фотографии, регистрирует интенсивность светового потока, т. е. в обоих случаях она выполняет одну и ту же функцию. Различие состоит лишь в том, что на голограмме необходимо фиксировать значительно более мелкие подробности распределения интенсивности и значительно больший диапазон изменения интенсивности, чем на фотографии.

Фотопластинка должна обеспечить запись дифракционной картины, которая составляет голограмму. В голограмме плоской волны (38.9) условие максимумов имеет вид

$$\cos(kx \sin \theta) = 1, \quad (38.28)$$

а расстояние Δx между ними определяется соотношением

$$k\Delta x \sin \theta = 2\pi. \quad (38.29)$$

Отсюда следует, что $\Delta x = \lambda / \sin \theta$. Например, при $\theta = 15^\circ \sin \theta = 0,26$ и поэтому $\Delta x \approx 4\lambda \approx 2 \text{ мкм}$, т. е. пластина должна быть способна разрешить линии, расположенные на расстоянии 2 мкм. Обычно разрешающая способность фотопластинок выражается в числе линий на 1 мм длины, которые пластина может разрешить. В рассмотренном случае требуемая разрешающая способность составляет 500 линий/мм. Желательно иметь пластины с еще большей разрешающей способностью. Для этого приходится использовать очень мелкие зерна галоидного серебра, что уменьшает чувствительность пластины. Поэтому пластины с высокой разрешающей способностью обладают низкой чувствительностью и требуют больших времен экспозиции, достигающих нескольких секунд при небольших мощностях лазеров. В течение времени экспозиции необходимо обеспечить стационарность процесса экспозиции и относительную неподвижность приборов и предмета съемки с точностью до доли длины волны (обычно $\lambda/4$). При использовании импульсных лазеров большой мощности времена экспозиции могут быть уменьшены до продолжительности импульса (миллисекунды и меньше). В этих условиях можно снимать голограммы быстродвижущихся объектов.

Объемное воспроизведение предмета. Получаемое с помощью голограммы мнимое изображение предмета наблюдается как его объемная фотография. Действительное изображение предмета представляет собой в определенном смысле объемное воспроизведение предмета с точностью до зеркального отражения. Предмет представляется висящим в воздухе, его можно фотографировать и т. д. Таким образом, в принципе, можно осуществить пространственную объемную реконструкцию обстановки, например создать иллюзию нахождения в комнате предмета, которого в действительности там нет.

Толстослойные голограммы (метод Денисюка). В отличие от рассмотренных голограмм, которые записываются на обычных фотопластинках с тонким слоем эмульсии, советский ученый Ю. Н. Денисюк (р. 1927 г.) предложил в 1962 г. метод толстослойных голограмм, в которых интерференционная картина дифрагированных лучей является не двухмерной, а трехмерной и захватывает всю толщину эмульсии.

Проанализируем толстослойную голограмму плоской волны, которую можно записать по схеме, изображенной на рис. 197, но используя вместо обычной фотопластинку с толстым слоем эмульсии (рис. 204). Обозначим \mathbf{k} и \mathbf{k}_0 волновые векторы сигнальной и опорной волн. Аналогично (38.9) для квадрата модуля амплитуды напряженности поля можем записать

$$|E|^2 = |E_0 e^{i\mathbf{k}_0 \cdot \mathbf{r}} + E_1 e^{i\mathbf{k} \cdot \mathbf{r}}|^2 = E_0^2 + E_1^2 + 2E_0 E_1 \cos [(\mathbf{k}_0 - \mathbf{k}) \cdot \mathbf{r}]. \quad (38.30)$$

Условие максимумов $|E|^2$ записывается в виде

$$(\mathbf{k}_0 - \mathbf{k}) \cdot \mathbf{r} = 2\pi m \quad (m = 0, \pm 1, \pm 2, \dots). \quad (38.31)$$

Уравнение (38.31) представляет собой систему плоскостей, перпендикулярных вектору $\mathbf{k}_0 - \mathbf{k}$ (рис. 205). Расстояние d между плоскостями удовлетворяет на основании (38.31) условию

$$|\mathbf{k}_0 - \mathbf{k}|d = 2\pi. \quad (38.32)$$

Учитывая, что $|\mathbf{k}| = |\mathbf{k}_0|$, получаем

$$|\mathbf{k} - \mathbf{k}_0| = 2k \sin (\theta/2). \quad (38.33)$$

Тогда [см. (38.32)]

$$d = \lambda / [2 \sin (\theta/2)]. \quad (38.34)$$

В частности, при интерференции двух встречных волн ($\theta = \pi$) плоскости максимального почернения параллельны волновым фронтам интерферирующих волн (рис. 206). Из (38.34) заключаем, что расстояние между плоскостями $d = \lambda/2$.

Условие Вульфа — Брэгга. От каждой плоскости максимального почернения, в которой сосредоточен максимум плотности восстановленного серебра, волны частично отражаются и частично проходят через нее. Однако от системы параллельных плоскостей отражение возможно лишь в том случае, если отраженные от соседних плоскостей волны усиливают друг друга (см. § 29). Угол падения α , при котором происходит отражение от системы параллельных плоскостей (рис. 207), аналогично (29.4) и (29.5) определяется условием

$$2d \cos \alpha = m\lambda, \quad (38.35)$$

где угол преломления $\theta_{\text{пр}}$ равен углу падения, а показатель преломления вещества между отражающими поверхностями равен единице. **Равенство (38.35) называется условием Брэгга — Вульфа.** Оно было независимо сформулировано У. Л. Брэггом (1890—1971) и русским кристаллографом Ю. В. Вульфом (1863—1925).

Получение голограммы и восстановление плоской волны. На рис. 208 изображена голограмма плоской волны, волновой вектор k которой образует с волновым вектором k_0 опорной волны угол $\pi - \beta$. Плоскости максимального почернения расположены перпендикулярно направлению вектора $k_0 - k$. Нормаль к поверхностям составляет с вектором k угол 2β , расстояние между поверхностями в соответствии с (28.34) равно

$$d = \lambda / \{2 \sin [(\pi - 2\beta)/2]\} = \lambda / (2 \cos \beta).$$

Тогда [см. (38.35)]

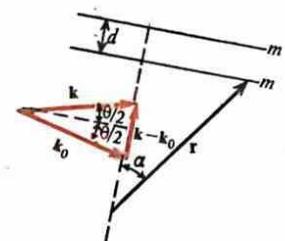
$$[2\lambda / (2 \cos \beta)] \cos \alpha = m\lambda \quad (38.36) \quad \text{или} \quad \cos \alpha = m \cos \beta. \quad (38.37)$$

Отсюда следует, что имеется только первый порядок отражения ($m = 1$) и $\alpha = \beta$. Следовательно, отражается только волна с волновым вектором k_0 , а отраженная волна имеет волновой вектор k . Другими словами, если голограмму плоской волны записать с помощью опорной волны той же частоты, то, облучая голограмму опорной волной, восстановим плоскую волну, информация о которой содержится в голограмме.

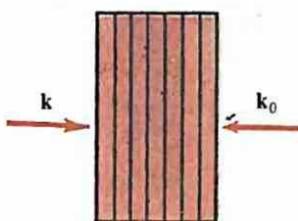
Получение голограммы и восстановление сферической волны. Сферическая волна на небольшом участке вдалеке от источника может рассматриваться как плоская, поэтому, облучая толстостойкую фотопластинку и точечный объект A одной и той же опорной волной с волновым вектором k_0 (рис. 209), получим в толще эмульсии совокупность поверхностей максимального почернения, расстояние между которыми удовлетворяет (38.34) с $\theta = \pi - 2\beta$. Из сказанного относительно равенства (38.37)



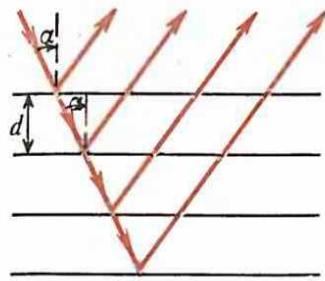
Запись голограммы плоской волны в толстостойкой эмульсии



К определению поверхностей максимального почернения в толстостойкой эмульсии

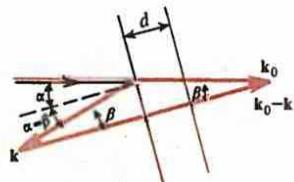


Пространственная картина интерференции в случае двух встречных волн



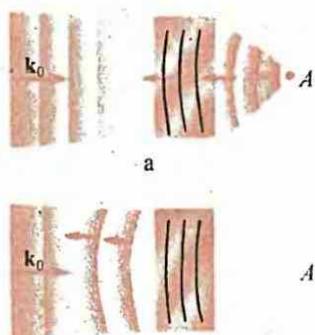
207

К выводу условия Вульфа — Брэгга



208

Схема отражения плоской волны от толстослойной голограммы



209

Запись толстослойной голограммы точечного объекта (а) и восстановление волны (б)

следует, что при облучении голограммы плоской волной с волновым вектором k_0 полностью восстанавливается записанная на голограмме сферическая волна как результат отражения плоской волны от дифракционной структуры, созданной в толще эмульсии при записи голограммы.

Получение голограммы и восстановление изображения произвольного объекта. При облучении фотопластинки и объекта одной и той же волной (рис. 210) каждая точка объекта создает в толще эмульсии дифракционную структуру, которая только что была рассмотрена (см. рис. 209). Совокупность дифракционных структур всех точек объекта составляет голограмму объекта.

Восстановление изображения производится облучением голограммы волной, совпадающей с опорной при записи голограммы (рис. 211). Изображение объекта — мнимое, расположенное в том месте, где находился реальный объект при записи голограммы. Такое восстановление изображения имеет существенный недостаток: восстанавливающая волна пространственно совпадает с восстановленной. Для устранения этого недостатка можно облучать голограмму волной с волновым вектором k_0 , направленным под подходящим углом к поверхности голограммы (рис. 212). Отраженная от дифракционной структуры волна образуется под углом отражения, равным углу падения. В результате восстановленная волна и восстанавливающая оказываются разделенными пространственно. Изображение объекта — мнимое, а его положение зависит от угла, под которым производится облучение голограммы.

Поскольку расстояние d между поверхностями максимального почернения примерно равно $\lambda/2$, заключаем, что в (38.35) $m = 1$. При восстановлении нет необходимости облучать голограмму монохроматическим светом. При облучении голограмм белым светом отразится лишь та волна, длина которой удовлетворяет условию (38.35). Это упрощает восстановление толстослойных голограмм и, кроме того, дает возможность получить цветное изображение.

Толстослойная голограмма позволяет создать действительное изображение предмета. Для этого необходимо облучать голограмму с той стороны (рис. 213), с которой находился предмет при записи по схеме рис. 210. Как видно на рис. 213, восстановленная волна является сходящейся и поэтому создает действительное изображение предмета.

Цветное объемное изображение. Хорошо известно, что любой цвет может быть представлен как смесь трех независимых цветов, в качестве которых обычно можно взять, например, красный, зеленый, синий. Цветное изображение получается в результате совмещения изображений в трех цветах.

Для получения цветного изображения при записи голограммы предмет облучается волнами, соответствующими трем цветам, принятым за независимые. На голограмме записываются три дифракционные структуры, соответствующие трем

волнам. Восстановление изображения осуществляется белым светом. В соответствии с (38.35) отражаются только волны с частотами, использованными при записи голограммы, причем их интенсивности соответствуют интенсивностям волн при записи. Это означает, что восстановленная с помощью белого света волна несет в себе цветное объемное изображение предмета.

Качество изображения зависит от разрешающей способности фотэмulsionии, от характера искажений дифракционной структуры, которые возникают из-за механических деформаций эмульсии при проявлении и сушке, и других причин. В настоящее время получают голограммы высокого качества.

Особенности голограмм как носителей информации. На любой участок голограммы попадает излучение от всех точек предмета. Это означает, что часть голограммы содержит информацию о предмете всю информацию, которую содержит вся голограмма. Если голограмму разделить на несколько частей, то каждая часть позволяет восстановить полное изображение предмета. Другими словами, часть голограммы содержит ту же информацию, что и целая голограмма. Поэтому ухудшение качества записи на отдельных участках голограммы (шарниры, затертости и т. д.) не ухудшают качества изображения. Запись информации на голограмме осуществляется с большим запасом надежности.

Объем информации, записанной на голограмме, несравненно больше объема, записанного на обычной фотографии. Емкость в сочетании с надежностью делает голограмму весьма перспективным носителем информации.

Применения голографии. Первоначальная задача голографии заключалась в получении объемного изображения. С развитием голографии на толстослойных пластинках возникла возможность создания объемных цветных фотографий. На этой базе исследуются пути реализации голографического кино, телевидения и т. д.

Из технологических применений наиболее значительно разработана **голографическая интерферометрия**. Восстановленная по голограмме волна дает копию объекта в тот момент времени, когда записывалась голограмма. Если эту волну сравнить с волной от объекта, восстановленной по голограмме и записанной в другой момент времени, то можно сделать заключение об изменениях в объекте за время между моментами записи голограмм. Поскольку голограмма фиксирует предмет с очень большой точностью, такой метод позволяет изучать с большой точностью явления, которые влияют на голограмму, например деформации, колебания и т. д. Метод называется голографической интерферометрией. На голограмму влияют не только пространственные перемещения частей предмета или его перемещение в целом, но и условия отражения и преломления света в предмете и другие факторы, приводящие к амплитудно-фазовой модуляции света. Поэтому посредством голографической интерферометрии изучают также распределение



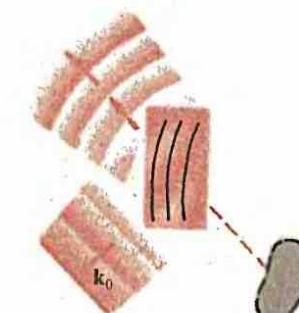
210

Запись толстослойной голограммы произвольного объекта



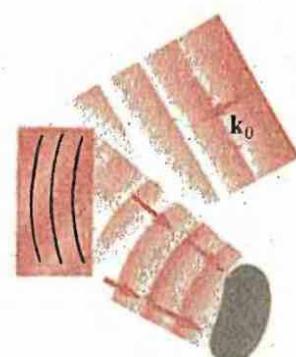
211

Восстановление изображения произвольного объекта



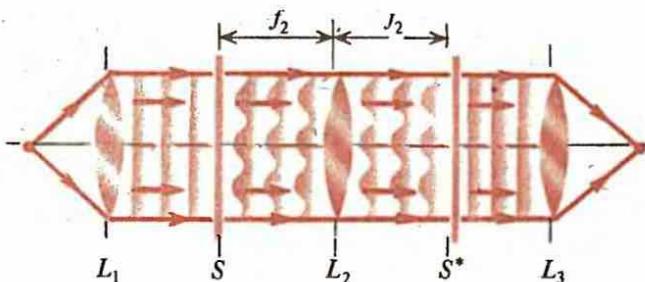
212

Способ разведения восстанавливающей и восстановленной волн



213

Получение действительного изображения с помощью толстослойной голограммы



214

Схема пространственной фильтрации изображения для решения задачи распознавания образов

напряжений в теле, крутильные моменты, распределение температур и т. д. Голография может применяться для обеспечения точности обработки деталей.

Это приводит часто к необходимости получить объемное изображение предмета, которого еще не существует, и следовательно, нельзя получить голограмму такого предмета оптическими методами. В этом случае голограмма рассчитывается на ЭВМ (**цифровая голограмма**) и результаты расчета соответствующим образом переносятся на фотопластинку. С полученной таким способом **машинной голограммы** объемное изображение предмета восстанавливается обычным оптическим способом. Поверхность предмета, полученного по машинной голограмме, используется как этalon, с которым методами голограммической интерференции производится сравнение поверхности реального предмета, изготовленного соответствующими инструментами. Голограммическая интерферометрия позволяет произвести сравнение поверхности изготовленного предмета и эталона с чрезвычайно большой точностью до долей длины волны. Это дает возможность изготавливать с такой же большой точностью очень сложные поверхности, которые было бы невозможно изготовить без применения цифровой голограммы и методов голограммической интерферометрии. Само собой разумеется, что для сравнения эталонной поверхности с изготавляемой не обязательно восстанавливать оптическим способом машинную голограмму. Можно снять голограмму предмета, перевести ее на цифровой язык ЭВМ и сравнить с цифровой голограммой. Оба эти пути в принципе эквивалентны.

Особенности голограмм как носителей информации делают весьма перспективными разработки по созданию голограммической памяти, которая характеризуется большим объемом, надежностью, быстрой считывания и т. д.

Голограммы также могут эффективно использоваться для решения различных задач, связанных с распознаванием образов. По своему физическому содержанию решение этой задачи сводится к осуществлению пространственной фильтрации изображений. На выходе линзы L_1 (рис. 214) формируется плоская волна, падающая на дифракционную структуру S . На выходе из S возникает сигнал, который требуется распознать. Пластинку с записью сигнала помещают в переднем фокусе линзы L_2 ,

- По физическому содержанию голография является реализацией принципа синхронного детектирования, давно известного в радиофизике, в видимом диапазоне электромагнитных волн. Для этого необходимо иметь свет с высокой степенью когерентности.

- Можно ли тонкослойные голограммы восстановить с помощью обычных, некогерентных источников света? Как это можно сделать? Какие изменения вносятся в изображение, если голограмма восстанавливается излучением с длиной волны, отличной от той, с которой она записывалась?

- Каким физическим фактором обусловливается возможность восстановления изображения, записанного на тонкослойной голограмме, с помощью излучения со сплошным спектром?

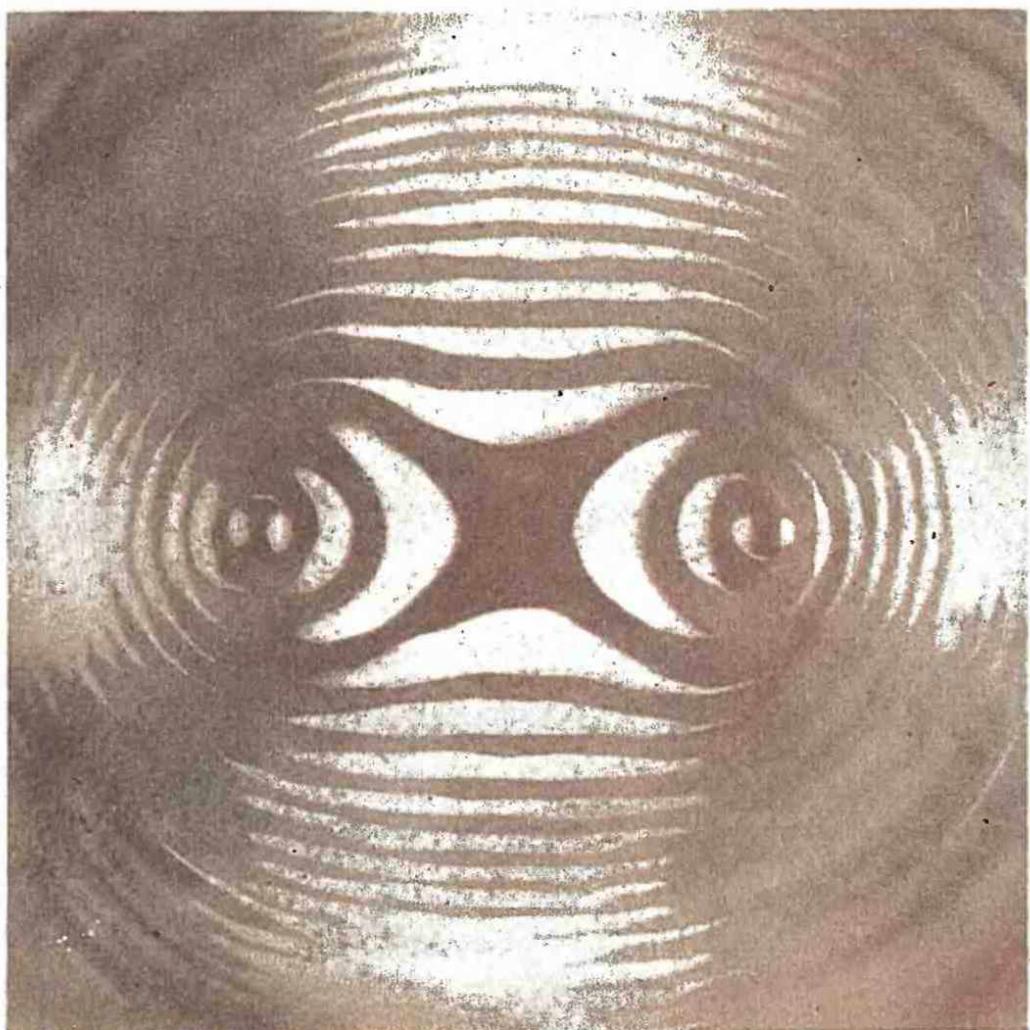
в заднем фокусе которой находится маска S^* для сигнала S . Маской называется дифракционная структура, при прохождении через которую волна снова становится плоской. Соотношение между записью на S и маской S^* определяется тем обстоятельством, что линза переводит дифракционную картину из одной фокальной плоскости в другую посредством преобразования Фурье (см. § 35). Поэтому маска S^* должна ликвидировать преобразование Фурье от S , отсюда её обозначение S^* . В фокусе линзы L_3 имеется приемник световой энергии. Если S является сигналом для маски S^* , то в фокусе концентрируется вся прошедшая через линзу энергиия. Если же S не представляет сигнала для S^* , то на выходе из S^* волна не плоская и, следовательно, световая энергия, прошедшая линзу L_3 , не сосредоточивается в фокусе, а распределяется по фокальной плоскости. Следовательно, концентрация поступившей энергии в фокусе не будет зафиксирована. Отсюда можно сделать заключение, что S не содержит сигнала, который ищется с помощью маски S^* . Таким образом, процедура распознавания заключается в «предъявлении» в качестве S различных образов и выборе из них того, который соответствует образу, фиксированному в маске S^* .

Эта принципиальная схема может быть усовершенствована. Например, если сигнал подать сразу на все маски, которым в принципе сигнал может соответствовать, то «ответит» та маска, которая соответствует сигналу. Благодаря этому образ, заключенный в S , оказывается «опознанным».

В принципиальной схеме (рис. 214) пространственной фильтрации для распознавания образов в качестве S и S^* могут быть и голограммы. Это расширяет возможности метода, потому что информационная емкость голограмм значительно больше информационной емкости фотографий.

Основная идея:
электрические свойства
среды характеризуются
тензором
диэлектрической
проницаемости.
Переход к главным
осям тензора делает
анализ распространения
света в анизотропной
среде наглядным.

Распространение света в анизотропных средах



Рассматривается метод описания анизотропной среды с помощью тензора диэлектрической проницаемости и осуществляется переход к главным осям тензора.

Источники анизотропии. Оптической анизотропией называется зависимость оптических свойств среды от направления. Она обусловлена зависимостью диэлектрических или магнитных свойств среды от направления. Полная анизотропия складывается из анизотропии свойств отдельных атомов и из анизотропии их упорядочения в пространстве.

Описание анизотропной диэлектрической среды. Анизотропия диэлектрических свойств среды означает, что зависимость поляризованности среды от напряженности электрического поля не может характеризоваться только одной скалярной величиной — диэлектрической восприимчивостью. В анизотропной среде проекции поляризованности связаны с проекциями напряженности электрического поля более сложными по сравнению с изотропной средой соотношениями:

$$\begin{aligned} P_x &= \epsilon_0 \alpha_{xx} E_x + \epsilon_0 \alpha_{xy} E_y + \epsilon_0 \alpha_{xz} E_z, \\ P_y &= \epsilon_0 \alpha_{yx} E_x + \epsilon_0 \alpha_{yy} E_y + \epsilon_0 \alpha_{yz} E_z, \\ P_z &= \epsilon_0 \alpha_{zx} E_x + \epsilon_0 \alpha_{zy} E_y + \epsilon_0 \alpha_{zz} E_z, \end{aligned} \quad (39.1)$$

где P_x, P_y, P_z — проекции поляризованности, ϵ_0 — диэлектрическая постоянная.

Совокупность величин $\alpha_{xx}, \alpha_{xy}, \dots$ называется тензором диэлектрической восприимчивости. Для упрощения написания формул будем нумеровать оси X, Y, Z соответственно индексами 1, 2, 3. Тогда формула (39.1) может быть записана в виде

$$P_i = \epsilon_0 \sum_j \alpha_{ij} E_j. \quad (39.2)$$

Соотношение между вектором смещения \mathbf{D} и поляризованностью \mathbf{P} :

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P}, \quad (39.3)$$

справедливое как для изотропной, так и анизотропной сред, принимает для анизотропной среды с учетом (39.2) вид

$$D_i = \epsilon_0 E_i + \epsilon_0 \sum_j \alpha_{ij} E_j = \sum_j \epsilon_0 (\delta_{ij} + \alpha_{ij}) E_j, \quad (39.4)$$

где δ_{ij} — символ Кронекера. Это соотношение удобнее представить в форме

$$D_i = \sum_j \epsilon_{ij} E_j, \quad (39.5)$$

где

$$\epsilon_{ij} = \epsilon_0 (\delta_{ij} + \alpha_{ij}) \quad (39.6)$$

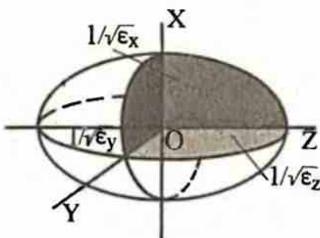
— тензор диэлектрической проницаемости.

Плотность электрической энергии в среде выражается формулой

$$w = \frac{1}{2} \mathbf{E} \cdot \mathbf{D} = \frac{1}{2} \sum_i E_i D_i = \frac{1}{2} \sum_{i,j} E_i \epsilon_{ij} E_j, \quad (39.7)$$

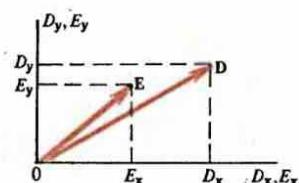
где для D_i использовано представление (39.5). Изменяя обозначение индексов суммирования ($i \leftrightarrow j$), запишем (39.7) в виде

$$w = \frac{1}{2} \sum_{i,j} E_j \epsilon_{ji} E_i. \quad (39.8)$$



215

Главные оси, относительно которых тензор диэлектрических проницаемостей диагонален



216

Неколлинеарность векторов \mathbf{D} и \mathbf{E} в анизотропной среде

Вычитая почленно (39.8) из (39.7), находим

$$0 = \frac{1}{2} \sum_{i,j} (\epsilon_{ij} - \epsilon_{ji}) E_i E_j. \quad (39.9)$$

Поскольку проекции E_i независимы, заключаем, что

$$\epsilon_{ij} - \epsilon_{ji} = 0, \quad (39.10)$$

т. е. тензор диэлектрической проницаемости является симметричным:

$$\epsilon_{ij} = \epsilon_{ji}. \quad (39.11)$$

Тензор диэлектрической проницаемости. Поскольку плотность электрической энергии w положительна, стоящая в правой части (39.8) квадратичная форма является положительно определенной. Для дальнейшего анализа удобно перейти к новым переменным:

$$x_i = E_i / \sqrt{2w}, \quad (39.12)$$

представив (39.8) равенством

$$\sum_{i,j} \epsilon_{ij} x_i x_j = 1. \quad (39.13)$$

Как известно из курса математики, квадратичная форма (39.13) может быть приведена к виду

$$\epsilon_{xx} x^2 + \epsilon_{yy} y^2 + \epsilon_{zz} z^2 = 1, \quad (39.14)$$

если соответствующим образом ориентировать оси координат. В системе координат, в которой квадратичная форма (39.13) принимает вид (39.14), тензор диэлектрической проницаемости является диагональным:

$$\epsilon_{ij} = \begin{pmatrix} \epsilon_x & 0 \\ 0 & \epsilon_y \\ 0 & 0 & \epsilon_z \end{pmatrix} \quad (39.15)$$

Уравнение (39.14) описывает эллипсоид, главные оси которого $1/\sqrt{\epsilon_x}$, $1/\sqrt{\epsilon_y}$, $1/\sqrt{\epsilon_z}$ (рис. 215). Оси X , Y , Z называют **главными осями тензора диэлектрических проницаемостей**. Уравнения (39.5), отнесенные к главным осям, принимают вид:

$$D_x = \epsilon_x E_x, \quad D_y = \epsilon_y E_y, \quad D_z = \epsilon_z E_z. \quad (39.16)$$

Поскольку, вообще говоря, $\epsilon_x \neq \epsilon_y \neq \epsilon_z$, векторы \mathbf{D} и \mathbf{E} не коллинеарны (рис. 216).

Анизотропия магнитных свойств среды описывается аналогично с помощью тензора магнитной восприимчивости. В дальнейшем будем считать среду немагнитной и ограничимся описанием анизотропии электрических свойств, поскольку именно этот случай реализуется в большинстве задач кристаллооптики. Магнитная анизотропия играет существенную роль при распространении света в прозрачных ферритах и ряде других сред.

- Какие существуют источники анизотропии электрических свойств среды?

Каким образом приводится тензор к своим главным осям?

Какая связь существует между вектором электрического смещения и напряженностью электрического поля в декартовой системе координат, оси которой совпадают с главными осями тензора диэлектрической проницаемости?

Каким физическим фактором определяется симметричность тензора диэлектрической проницаемости?

Описывается типы волн и их свойства
в анизотропной среде.

Плоская электромагнитная волна в анизотропной среде. Аналогично (2.50) и (2.51) векторы поля плоской электромагнитной волны представляются в виде

$$\mathbf{E} = \mathbf{E}_0 e^{-i(\omega t - \mathbf{k} \cdot \mathbf{r})}, \quad \mathbf{D} = \mathbf{D}_0 e^{-i(\omega t - \mathbf{k} \cdot \mathbf{r})}, \quad \mathbf{B} = \mathbf{B}_0 e^{-i(\omega t - \mathbf{k} \cdot \mathbf{r})}, \quad \mathbf{H} = \mathbf{H}_0 e^{-i(\omega t - \mathbf{k} \cdot \mathbf{r})}. \quad (40.1)$$

Подставляя (40.1) в уравнения Максвелла (2.1) — (2.5) и используя (2.45) и правила действия оператора ∇ (2.52), получаем уравнения

$$-\mathbf{k} \times \mathbf{H} = \omega \mathbf{D}, \quad (a) \quad \mathbf{k} \times \mathbf{E} = \omega \mu_0 \mathbf{H}, \quad (b) \quad \mathbf{k} \cdot \mathbf{D} = 0, \quad (v) \quad \mathbf{k} \cdot \mathbf{H} = 0. \quad (g) \quad (40.2)$$

Волновой вектор \mathbf{k} перпендикулярен поверхности одинаковой фазы, т. е. показывает направление распространения волнового фронта. Фазовая скорость v волны имеет направление по этому вектору, которое принимается за направление распространения волны и характеризуется единичным вектором $\mathbf{n} = \mathbf{k}/k$. Из, (40.2 в, г) видно, что волна распространяется перпендикулярно \mathbf{D} и \mathbf{H} . Выражение (3.1) для вектора Пойнтинга показывает, что поток энергии направлен перпендикулярно \mathbf{E} и \mathbf{H} . Направление потока энергии в волне называется лучом. Оно, вообще говоря, не совпадает с направлением движения волны. Обозначим единичный вектор в направлении луча $\tau = \mathbf{S}/S$. Как было отмечено в § 15, энергия электромагнитной волны движется с групповой скоростью. Поэтому можно сказать, что групповая скорость v_g волны совпадает по направлению с τ .

Поскольку в анизотропной среде векторы \mathbf{E} и \mathbf{D} не коллинеарны, направления распространения волны и луча не совпадают и, следовательно, групповая и фазовая скорости не совпадают по направлению. В этом состоит первая важная особенность распространения электромагнитной волны в анизотропной среде. Вторая важная особенность состоит в том, что скорость электромагнитных волн зависит от направления их распространения и поляризации.

Из (40.2) и (3.1) видно, что \mathbf{n} и τ перпендикуляры \mathbf{H} ; \mathbf{E} и \mathbf{D} также перпендикуляры \mathbf{H} и, кроме того, \mathbf{n} и τ перпендикуляры соответственно \mathbf{D} и \mathbf{E} . Это означает, что \mathbf{D} , \mathbf{E} , \mathbf{n} , τ лежат в одной и той же плоскости, перпендикулярной \mathbf{H} (рис. 217). Угол между \mathbf{D} и \mathbf{E} равен углу между \mathbf{n} и τ .

Зависимость фазовой скорости от направлений распространения волны и колебаний вектора \mathbf{D} . Рассмотрим волну, распространяющуюся в положительном направлении оси Z , которая является одной из главных осей тензора ϵ_{ij} диэлектрической проницаемости (рис. 215). Считаем, что вектор электрического смещения \mathbf{D} коллинеарен оси X ($D_x \neq 0, D_y = 0, D_z = 0$) и, следовательно, вектор \mathbf{H} коллинеарен оси Y . На основании (39.16) имеем $E_x = D_x/\epsilon_{xx}$, $E_y = E_z = 0$. Уравнения (40.2 а, б) принимают вид.

$$kH_y = \omega D_x, \quad (a) \quad kE_x = \omega \mu_0 H_y. \quad (b) \quad (40.3)$$

Перемножая между собой соответственно левые и правые части равенств (40.3 а, б), получаем

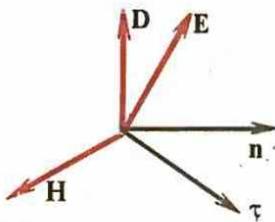
$$k^2 E_x H_y = \omega^2 \mu_0 \epsilon_x E_x H_y, \quad (40.4)$$

где $D_x = \epsilon_x E_x$. Отсюда

$$k^2 = \omega^2 \epsilon_x \mu_0 \quad (40.5)$$

и, следовательно, фазовая скорость

● В анизотропной среде в заданном направлении могут распространяться лишь две волны с взаимно перпендикулярными линейными поляризациями и различными скоростями.

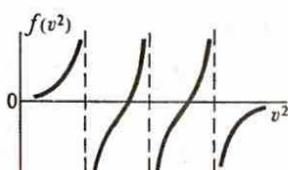


217

Взаимное расположение векторов плоской волны в анизотропной среде

$$v_x = \omega/k = 1/\sqrt{\epsilon_x \mu_0},$$

(40.6)



218

К графическому определению корней уравнения Френеля

где индекс x у фазовой скорости означает, что она является скоростью волны, векторы D и E которой коллинеарны оси X . Если векторы D и E волны коллинеарны оси Y , то аналогично (40.6) получаем

$$v_y = 1/\sqrt{\epsilon_y \mu_0}.$$

(40.7)

Поскольку, вообще говоря, $\epsilon_x \neq \epsilon_y$, заключаем, что фазовая скорость волны различна для этих двух направлений колебаний вектора E . Отсюда следует, что в направлении оси Z могут распространяться лишь волны, векторы D и E которых колеблются параллельно либо оси X , либо оси Y .

Уравнение Френеля. Аналогичные заключения могут быть сделаны и относительно волн, распространяющихся в направлении осей X и Y . Для того чтобы изучить поведение волн, распространяющихся в произвольном направлении, необходимо вместо уравнений (40.3 а, б), которые являются частным случаем уравнений (40.2 а, б), проанализировать общий случай. Подставляя выражение для H из (40.2б) в уравнение (40.2а) и пользуясь формулой разложения двойного векторного произведения, находим

$$\mathbf{v}(\mathbf{n} \cdot \mathbf{E}) - \mathbf{E} + \mu_0 v^2 \mathbf{D} = 0,$$

где $v = \omega/k$. Все дальнейшие вычисления удобно вести в прямоугольной декартовой системе координат, оси которой совпадают с главными осями тензора диэлектрической проницаемости, когда $D_i = \epsilon_i E_i$. В этой системе координат уравнение (40.8) записывается в виде трех скалярных уравнений:

$$n_i(\mathbf{n} \cdot \mathbf{E}) - E_i(1 - v^2 \mu_0 \epsilon_i) = 0 \quad (i = 1, 2, 3).$$

Пусть вектор E направлен по одной из главных осей тензора диэлектрической проницаемости, например X . Тогда $D_1 = \epsilon_1 E_1$, $D_2 = D_3 = 0$, $E_2 = E_3 = 0$. Ясно, что в этом случае векторы \mathbf{v} и \mathbf{n} совпадают и лежат в плоскости $X_2 X_3$ (напомним, что $X_1 = X$, $X_2 = Y$, $X_3 = Z$). Уравнения (40.9) сводятся к одному:

$$E_1(1 - v^2 \mu_0 \epsilon_1) = 0,$$

где фазовая скорость обозначена индексом 1, соответствующим волне, у которой векторы E и \mathbf{E} коллинеарны оси X . Поскольку $E_1 \neq 0$, из (40.10) заключаем, что

$$v_1 = 1/\sqrt{\epsilon_1 \mu_0}.$$

Этот частный случай совпадает с тем, который был в качестве простейшего примера рассмотрен, исходя из уравнений (40.3 а, б), а формула (40.11) совпадает с (40.6). Аналогичны формулы и для осей X_2 и X_3 :

$$v_i = 1 / \sqrt{\epsilon_i \mu_0} \quad (i = 2, 3). \quad (40.12)$$

Можно сказать, что v_i есть фазовая скорость волны, соответствующая оси X_i . Однако отметим, что она не является проекцией фазовой скорости волны на ось X_i , а характеризует фазовую скорость волны, векторы \mathbf{E} и \mathbf{D} которой коллинеарны оси X_i . Скорости v_i называют **главными скоростями распространения волны**. С учетом (40.12) уравнения (40.9) могут быть представлены в виде.

$$n_i(\mathbf{n} \cdot \mathbf{E}) - E_i(1 - v^2/v_i^2) = 0. \quad (40.13)$$

Умножая обе части (40.13) на $n_i/(1 - v^2/v_i^2)$ и суммируя по i , получаем равенство

$$\sum_{i=1}^3 \frac{n_i^2 v_i^2}{v_i^2 - v^2} - 1 = 0, \quad (40.14)$$

в котором произведено сокращение на множитель $\sum n_i E_i = \mathbf{n} \cdot \mathbf{E}$. Поскольку $\sum n_i^2 = 1$, уравнение (40.14) после приведения к общему знаменателю и деления на $v^2 \neq 0$ принимает вид

$$\sum_{i=1}^3 \frac{n_i^2}{v_i^2 - v^2} = 0. \quad (40.15)$$

Формула (40.15) называется **уравнением Френеля**, позволяющим найти фазовую скорость в направлении, характеризуемом направляющими косинусами n_1, n_2, n_3 . Отметим, что величины v_i в этом уравнении не являются проекциями вектора \mathbf{v} на ось координат, т. е. $n_i v \neq v_i$. Решение уравнения (40.15) дает фазовую скорость v как функцию n_i и v_i .

Фазовую скорость v удобно выразить в виде функции от направления вектора \mathbf{D} . Пусть $\mathbf{d} = \mathbf{D}/D$ — единичный вектор в направлении \mathbf{D} . Умножая скалярно обе части (40.8) на \mathbf{D} , получаем

$$-\sum \epsilon_i E_i^2 + \mu_0 v^2 D^2 = 0, \quad (40.16)$$

или

$$v^2 = (1/D^2) \sum_i v_i^2 D_i^2, \quad (40.17)$$

где $\epsilon_i E_i^2 / \mu_0 = v_i^2 D_i^2$. Поскольку $d_i = D_i/D$, равенство (40.17) принимает вид

$$v^2 = \sum_i d_i^2 v_i^2, \quad (40.18)$$

т. е. фазовая скорость полностью определяется направлением вектора \mathbf{D} .

Типы возможных волн. Уравнение (40.15) будет удовлетворено, если в его левой части имеются члены с различными знаками. Поэтому v^2 не может быть ни больше, ни меньше всех v_i^2 . Для нахождения корней уравнения (40.15) построим график функции

$$f(v^2) = \sum_i \frac{n_i^2}{v_i^2 - v^2} \quad (40.19)$$

Этот график представлен на рис. 218. Вертикальные пунктирные линии проведены через точки v_i^2 . На рисунке видно, что имеются два действительных значения v^2 , удовлетворяющих уравнению (40.15). Это означает, что в заданном направлении могут распространяться волны с двумя различными фазовыми скоростями v' и v'' , заключенными между наименьшей и средней

- В каких случаях векторы электрического смещения и напряженности электрического поля волны в анизотропной среде совпадают? Почему в общем случае в анизотропной среде нормаль к поверхности волнового фронта не совпадает с направлением потока энергии волны? Когда они совпадают?

и наибольшей из скоростей v_i . Докажем, что векторы \mathbf{D}' и \mathbf{D}'' этих волн взаимно перпендикулярны. Для этого умножим соотношение (40.8) для \mathbf{D}' и \mathbf{E}' скалярно на \mathbf{D}'' и вычтем из него почленно это же соотношение для \mathbf{D}'' , \mathbf{E}'' , умноженное скалярно на \mathbf{D}' . В результате получаем $(\mathbf{D}' \cdot \mathbf{E}'' - \mathbf{D}'' \cdot \mathbf{E}') - \mu_0(v'^2 - v''^2)\mathbf{D}' \cdot \mathbf{D}'' = 0$. (40.20)

Учтем, что

$$\mathbf{D}' \cdot \mathbf{E}'' = \sum_{i,j} \epsilon_{ij} E'_j E''_i = \sum_{i,j} \epsilon_{ji} E'_i E''_j = \mathbf{D}'' \cdot \mathbf{E}' . \quad (40.21)$$

Заключенная в первые скобки разность в (40.20) равна нулю, а разность $v'^2 - v''^2$ не равна нулю и поэтому

$$\boxed{\mathbf{D}' \cdot \mathbf{D}'' = 0,} \quad (40.22)$$

т. е. векторы \mathbf{D} двух волн, распространяющихся в данном направлении, взаимно перпендикулярны. Отсюда следует, что для каждого направления в перпендикулярной ему плоскости существуют два взаимно перпендикулярных направления, коллинеарно которым могут колебаться векторы \mathbf{D} . Фазовые скорости соответствующих волн различны. Других волн, которые распространялись бы в том же направлении, но имели другое направление колебаний вектора \mathbf{D} , не существует.

§ 41 | Ход лучей в анизотропной среде

С помощью лучевого эллипсоида анализируется ход лучей в анизотропной среде и дается определение одноосных и двуосных кристаллов.

Зависимость лучевой скорости от направления. Все результаты о направлении движения фронта волны и фазовой скорости были получены при анализе уравнений (40.2), в которые входят волновой вектор \mathbf{k} и частота ω , характеризующие фазовую скорость, и нормаль \mathbf{n} к поверхности фронта волны. Чтобы проанализировать вопрос о лучах света и групповой скорости v_r , необходимо эти уравнения преобразовать так, чтобы в формулы вошли τ и v_r . Для нахождения групповой скорости v_r заметим, что фронт волны распространяется в направлении \mathbf{n} , а энергия — в направлении τ . Поэтому «фронт потока энергии» расположен перпендикулярно τ . Отсюда заключаем (см. рис. 217), что групповая и фазовая скорости света в анизотропной среде связаны между собой соотношением

$$v_r = v \cos(\hat{\mathbf{n}}, \hat{\tau}) = v(\mathbf{n} \cdot \tau), \quad (41.1)$$

где $(\hat{\mathbf{n}}, \hat{\tau})$ — угол между \mathbf{n} и τ .

Умножая уравнения (40.2 а, б) слева векторно на τ , получаем

$$(\tau \cdot \mathbf{k}) \mathbf{H} = \omega \tau \times \mathbf{D}, \quad (\text{а}) \quad (\tau \cdot \mathbf{k}) \mathbf{E} = -\omega \mu_0 \tau \times \mathbf{H}. \quad (\text{б}) \quad (41.2)$$

Вычисляя выражение для \mathbf{H} из (41.2а) и подставляя его в (41.2б), находим

$$(\tau \cdot \mathbf{k}) \mathbf{E} = -\omega^2 \mu_0 [\tau(\tau \cdot \mathbf{D}) - \mathbf{D}] / (\tau \cdot \mathbf{k}). \quad (41.3)$$

Принимая во внимание, что $\mathbf{k} = \mathbf{n}k$, $v = \omega/k$, $v_r = v(\mathbf{n} \cdot \tau)$, уравнение (41.3) перепишем в виде

$$\tau(\tau \cdot \mathbf{D}) - \mathbf{D} + [1/(\mu_0 v_r^2)] \mathbf{E} = 0. \quad (41.4)$$

Оно подобно уравнению (40.8), если в последней произвести замены $\mathbf{n} \rightarrow \tau$, $\mathbf{E} \rightarrow \mathbf{D}$, $\mathbf{D} \rightarrow \mathbf{E}$, $\mu_0 v^2 \rightarrow 1/(\mu_0 v_r^2)$. Поэтому все результаты, полученные на основании (40.8) для волн, могут быть переформулированы для лучей с учетом указанной замены величин.

Пусть вектор \mathbf{E} направлен вдоль любой из главных осей тензора диэлектрической проницаемости (например, оси X). Поскольку в этом случае \mathbf{D} параллельно \mathbf{E} , получим для v_r , выраже-

268 ние совпадающие с (40.10), т. е. главные групповые скорости совпадают с главными фазовыми скоростями [см. (40.12)]. Следовательно, их можно обозначить не v_{ri} , а v_i . Вместо (40.13) получаем

$$t_i(\tau \cdot D) \rightarrow D_i(1 - v_i^2/v_r^2) = 0. \quad (41.5)$$

Уравнение Френеля (40.15) принимает для лучевых скоростей вид

$$\sum_{i=1}^3 t_i^2 v_i^2 / (v_i^2 - v_r^2) = 0. \quad (41.6)$$

Если единичный вектор в направлении E обозначить $I = E/E$, то уравнение, соответствующее (40.18), для лучевой скорости будет иметь вид

$$1/v_r^2 = \sum_i I_i^2/v_i^2. \quad (41.7)$$

Докажем, что скорость в направлении луча является действительно групповой. В уравнение (40.15) подставим $v = \omega/k$, тогда

$$\Sigma n_i^2 / (\omega^2 - v_i^2 k^2) = 0. \quad (41.8)$$

Рассматривая (41.8) как уравнение, которое в неявном виде определяет $\omega = \omega(k)$, и вычисляя из него $v_{ri} = d\omega/dk_i$, можно убедиться, что эти скорости удовлетворяют уравнению (41.6). Таким образом, входящие в (41.6) скорости действительно групповые.

Две волны, распространяющиеся в данном направлении с двумя различными групповыми скоростями, имеют взаимно перпендикулярные направления поляризации

$$E \cdot E'' = 0, \quad (41.9)$$

что доказывается аналогично выводу формулы (40.22).

Эллипсоид лучевых скоростей. Произведем в уравнении (39.14) замену переменных:

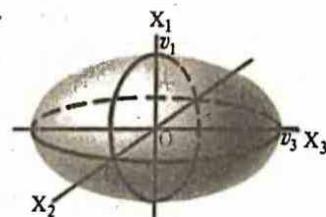
$$x_i \rightarrow \sqrt{\mu_0} x_i, \quad (41.10)$$

в результате имеем

$$x_1^2/v_1^2 + x_2^2/v_2^2 + x_3^2/v_3^2 = 1, \quad (41.11)$$

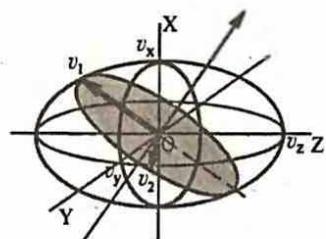
где $v_1 = 1/\sqrt{\epsilon_1 \mu_0}$, $v_2 = 1/\sqrt{\epsilon_2 \mu_0}$, $v_3 = 1/\sqrt{\epsilon_3 \mu_0}$. Уравнением (41.11) описывается эллипсоид (рис. 219), идентичный эллипсоиду (см. рис. 215), описываемому уравнением (39.14), но в других переменных. Эллипсоид, точки поверхности которого удовлетворяют уравнению (41.11), называется **эллипсоидом лучевых скоростей** [x_i в (41.11) имеют размерность скорости].

Анализ хода лучей с помощью эллипсоида лучевых скоростей. Решение уравнений (41.5) и (41.6) может быть выполнено с помощью геометрического построения, базирующегося на эллип-



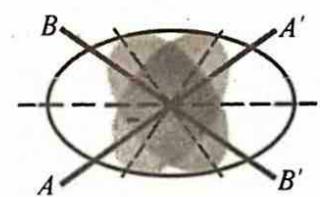
219

Эллипсоид лучевых скоростей



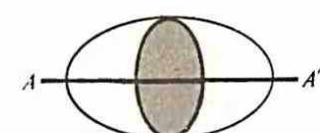
220

К анализу хода лучей света с помощью эллипсоида лучевых скоростей



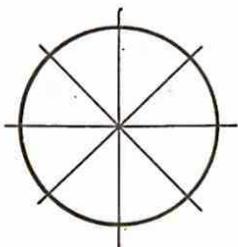
221

Эллипсоид лучевых скоростей двуосного кристалла (AA' и BB'—оптические оси)



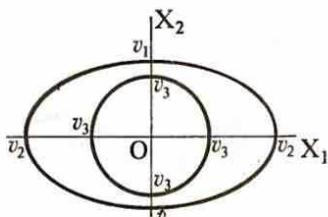
222

Эллипсоид лучевых скоростей однородного кристалла (AA'—оптическая ось)



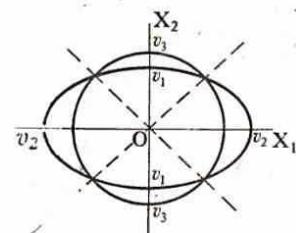
223

Эллипсоид лучевых скоростей оптически изотропной среды



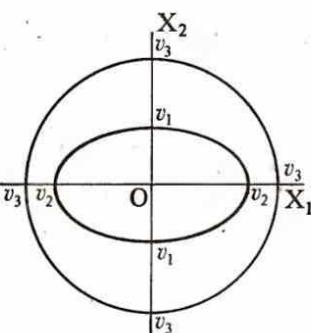
224

Сечения лучевой поверхности плоскостью $X_3 = 0$ при $v_3 < v_1 < v_2$



225

Сечения лучевой поверхности плоскостью $X_3 = 0$ при $v_1 < v_3 < v_2$



226

Сечения лучевой поверхности плоскостью $X_3 = 0$ при $v_1 < v_2 < v_3$

соиде лучевых скоростей (рис. 220), уравнение которого

$$x^2/v_x^2 + y^2/v_y^2 + z^2/v_z^2 = 1,$$

(41.12)

где v_x, v_y, v_z — главные лучевые скорости.

Направление луча задается единичным вектором τ . Через центр лучевого эллипсоида проведем плоскость, перпендикулярную τ . В сечении эллипса этой плоскостью образуется эллипс (рис. 220), главные полуоси которого v_1, v_2 . Вектор E световой волны, распространяющейся по лучу, может колебаться только параллельно главным осям этого эллипса, соответствующие групповые (лучевые) скорости равны длинам его главных полуосей v_1 и v_2 . Таким образом, в произвольном направлении τ возможно распространение лишь двух линейно поляризованных волн с различными лучевыми (групповыми) скоростями. Если сечение эллипса вырождается в окружность, то $v_1 = v_2$ и поэтому любой радиус окружности является главной полуосью. В этом случае вектор E может колебаться в любом направлении, перпендикулярном τ .

Оптическая ось. В направлении, перпендикулярном плоскости кругового сечения эллипса лучевых скоростей, всем лучам соответствует одна и та же лучевая скорость, а векторы E волн могут колебаться в любом направлении плоскости кругового сечения. Это означает, что для этих лучей анизотропия среды не проявляется и среда ведет себя как изотропная. Направление, перпендикулярное круговому сечению, называется оптической осью анизотропной среды.

Двусные и одноосные кристаллы. В аналитической геометрии доказывается, что эллипсоид с тремя различными по значению главными осями имеет два круговых сечения (рис. 221). Это означает, что если у эллипса лучевых скоростей все главные скорости v_x, v_y, v_z различны, то соответствующая среда имеет две оптические оси: AA' и BB' . Обычно анизотропия наблюдается в кристаллах, поэтому говорят об оптических осях кристалла. Кристаллы с двумя оптическими осями называются **двусными**.

Если у эллипса лучевых скоростей две главные скорости равны между собой, то он является эллипсоидом вращения вокруг третьей оси. В этом случае имеется только одна оптическая ось, совпадающая с осью вращения (т. е. с третьей главной осью эллипса). Такие кристаллы называются **одноосными** (рис. 222).

Если у эллипса лучевых скоростей все главные скорости равны, то он превращается в сферу. В этом случае все направления эквивалентны и соответствующий кристалл **оптически изотропен** (рис. 223).

Эллипсоид волновых нормалей. Вместо лучей можно рассматривать распространение волн, движение которых характеризуется единичным вектором n , нормальным к поверхности волнового фронта. При анализе хода лучей используются век-

торы E и H , перпендикулярные направлению луча τ . Анализ распространения волн проводится с помощью векторов D и H , перпендикулярных вектору n . Вместо эллипсаода лучевых скоростей необходимо пользоваться эллипсоидом волновых нормалей, уравнение для которого получается из квадратичной формы (39.8) аналогично тому, как был получен эллипсоид лучевых скоростей. В формуле (39.14) величины (x, y, z) в соответствии с (39.12) пропорциональны E_x, E_y, E_z . Чтобы в качестве независимых переменных взять D_x, D_y, D_z , необходимо заменить $x \rightarrow x/\varepsilon_x, y \rightarrow y/\varepsilon_y, z \rightarrow z/\varepsilon_z$. В результате вместо (39.14) имеем

$$x^2/\varepsilon_x + y^2/\varepsilon_y + z^2/\varepsilon_z = 1. \quad (41.13)$$

Аналогично тому, как из (39.14) было получено уравнение эллипсаода лучевых скоростей (41.12), из (41.13) находим соотношение

$$v_x^2 x^2 + v_y^2 y^2 + v_z^2 z^2 = 1, \quad (41.14)$$

которое называется **уравнением эллипсаода волновых нормалей**. Значения скоростей v_x, v_y, v_z здесь такие же, как и в (39.14). Для того чтобы с помощью этого эллипсаода анализировать распространение волн, удобно представить его в виде

$$\frac{x^2}{(1/v_x)^2} + \frac{y^2}{(1/v_y)^2} + \frac{z^2}{(1/v_z)^2} = 1. \quad (41.15)$$

Анализ распространений волн проводится аналогично анализу хода лучей, надо лишь вместо эллипсаода лучевых скоростей пользоваться эллипсоидом волновых нормалей. Направление распространения волны задается вектором n . Находится сечение эллипсаода (41.15) плоскостью, перпендикулярной n и проходящей через центр эллипсаода. Колебания вектора D возможны лишь в направлениях, параллельных главным осям эллипса в сечении эллипсаода. Фазовые скорости волн обратно пропорциональны длинам соответствующих главных осей эллипса. Однако для анализа распространения света в анизотропных средах удобнее пользоваться понятием лучевой поверхности, а не поверхности волнового фронта.

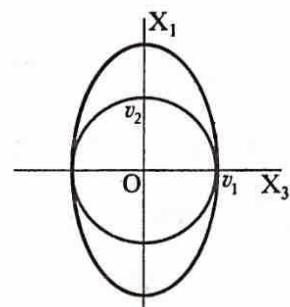
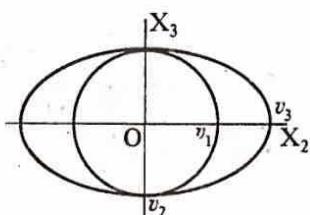
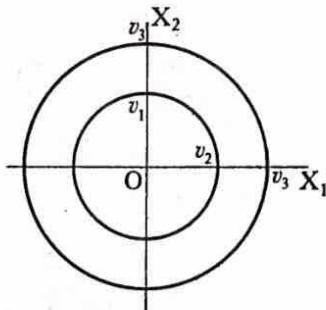
Лучевая поверхность. Лучи в анизотропной среде можно также рассматривать и без эллипсаода лучевых скоростей, непосредственно с помощью уравнения Френеля (41.6). Для этого перейдем к новым переменным

$$r = \tau v_r, \quad r_i = x_i = \tau_i v_r, \quad (41.16)$$

в которых уравнение (41.6) имеет вид

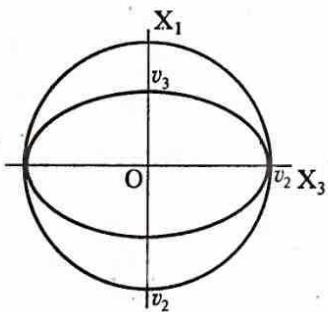
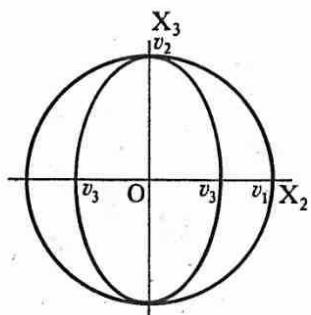
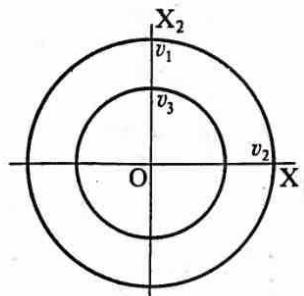
$$\sum x_i^2 v_i^2 / (r^2 - v_r^2) = 0. \quad (41.17)$$

Поверхность четвертого порядка, определяемая уравнением (41.17), называется **лучевой**. Расстояние r от начала координат до соответствующей точки поверхности пропорционально лучевой скорости в направлении τ . В каждом направлении луче-



Сечения лучевой поверхности одностороннего кристалла при $v_3 > v_1 = v_2$

• **Лучевой эллипсоид** дает полное решение задачи о распространении лучей света в анизотропной среде. Для этой цели можно также пользоваться эллипсоидом волновых нормалей и лучевой поверхностью.



228

Сечения лучевой поверхности однодиосного кристалла при $v_3 < v_1 = v_2$

- Что такое оптическая ось? Сколько оптических осей может существовать в кристалле? Что такое одноосные и двуосные кристаллы? Опишите метод анализа распространения лучей в анизотропной среде с помощью лучевого эллипсоида.

вая поверхность встречается два раза, что соответствует наличию двух скоростей распространения света в каждом направлении.

Рассмотрим сечение лучевой поверхности координатными плоскостями, например плоскостью $x_3 = 0$. При этом условии уравнение (41.17) распадается на два:

$$x_1 v_1^2 / (r^2 - v_1^2) + x_2 v_2^2 / (r^2 - v_2^2) = 0 \quad (r^2 = x_1^2 + x_2^2), \quad (a)$$

$$r^2 - v_3^2 = x_1^2 + x_2^2 - v_3^2 = 0. \quad (b)$$

Первое уравнение после элементарных преобразований принимает вид

$$x_1^2/v_1^2 + x_2^2/v_2^2 + x_3^2/v_3^2 = 1, \quad (41.19)$$

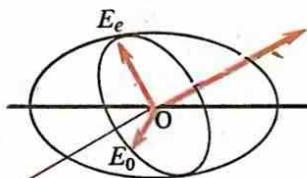
Таким образом, лучевая поверхность пересекает координатные плоскости по эллипсу (41.19) и окружности (41.18б). Их взаимное расположение определяется соотношением скоростей v_1, v_2, v_3 . При $v_3 < v_1 < v_2$ сечения имеют вид, показанный на рис. 224. Если $v_3 < v_2 < v_1$, то эллипс вытянут вдоль оси X_2 . Сечения при $v_1 < v_3 < v_2$ показаны на рис. 225. В случае $v_2 < v_3 < v_1$ эллипс вытягивается вдоль оси X_1 . При $v_1 < v_2 < v_3$ картина сечений изображена на рис. 226. При $v_2 < v_1 < v_3$ эллипс вытягивается вдоль оси X_2 .

Нетрудно видеть, что окружность описывает скорости того луча, электрический вектор которого колеблется параллельно главной оси, перпендикулярной рассматриваемой плоскости, в данном случае главной оси X_3 , т. е. электрический вектор колеблется перпендикулярно плоскости рисунка. Электрический вектор луча, описываемого эллипсом, колеблется в плоскости рисунка, в данном случае в плоскости X_1X_2 . На рис. 224—226 изображены возможные сечения лучевой поверхности координатными поверхностями при неравных v_1, v_2, v_3 . Поскольку оптическая ось определяется равенством скоростей для обоих лучей в направлении оси, она может быть найдена построением, указанным на рис. 225, где оптические оси изображены пунктирными линиями. При неравных v_1, v_2, v_3 кристалл имеет две оптические оси.

У одноосного кристалла две оси эллипсоида лучевых скоростей равны между собой. Положим $v_1 = v_2$. Тогда при $v_3 > v_1 = v_2$ эллипсоид лучевых скоростей сплюснут вдоль оси X_3 , при $v_3 < v_1 = v_2$ — вытянут. Сечения лучевой поверхности координатными плоскостями в этих случаях изображены на рис. 227, 228. Оптическая ось совпадает с главной осью лучевого эллипсоида. Кристаллы, для которых $v_3 < v_1 = v_2$, называют положительными, а для которых $v_3 > v_1 = v_2$ — отрицательными.

Лучевые поверхности можно построить также и не прибегая к решению уравнения (41.17), а исходя непосредственно из эллипсоида лучевых скоростей. Для этого из центра эллипсоида в каждом направлении откладывают два отрезка, равные главным осям эллипсов в сечениях эллипсоида, перпендикулярных соответствующим направлениям. Эти отрезки равны лучевым

272 скоростям. Концы отрезков лежат на лучевой поверхности. Такой метод, основанный на эллипсоиде лучевых скоростей, нагляден и более предпочтителен при качественном анализе лучевых поверхностей. Для количественных расчетов более целесообразно использовать уравнение (41.17).



228

Обыкновенный и необыкновенный лучи в отрицательном кристалле

§ 42 Двойное лучепреломление

С помощью лучевых поверхностей рассматриваются двойное лучепреломление и обусловленные им явления.

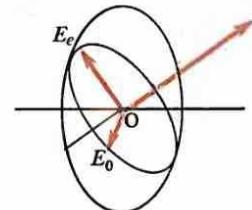
Обыкновенный и необыкновенный лучи. Через луч, направленный под углом к оптической оси (рис. 229), и оптическую ось можно провести плоскость, называемую главной (на рис. 229 она совпадает с плоскостью чертежа). Ясно, что у луча, вектор E_0 которого направлен перпендикулярно главной плоскости, скорость не зависит от направления и равна лучевой скорости, направленной коллинеарно оптической оси. Этот луч называется обыкновенным; величины, относящиеся к нему, обозначаются с индексом o , его скорость v_o , показатель преломления $n_o = c/v_o$. У луча, вектор E_e которого (рис. 229) лежит в главной плоскости, скорость зависит от направления, поскольку соответствующая главная ось эллипса в сечении эллипсоида изменяется с изменением направления луча. Этот луч называется необыкновенным; величины, относящиеся к нему, обозначаются с индексом e . Его скорость v_e , показатель преломления n_e .

У отрицательных кристаллов $v_0 < v_e$, у положительных — $v_0 > v_e$ (см. § 41). Соотношение главных осей эллипсоида лучевых скоростей у отрицательного кристалла показано на рис. 229, а у положительного — на рис. 230. Сечения лучевых поверхностей отрицательных кристаллов показаны на рис. 227, а положительных — на рис. 228.

Сущность двойного лучепреломления. Поскольку внутри кристалла возможно распространение лишь двух лучей с различными лучевыми скоростями, преломление луча на поверхности кристалла приводит к возникновению двух лучей внутри кристалла. Разделение луча, входящего в кристалл, на два называется **двойным лучепреломлением**.

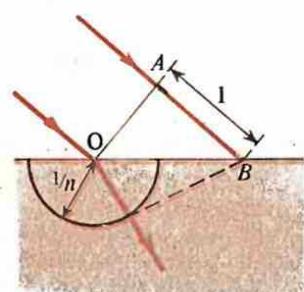
Двойное лучепреломление было открыто в 1669 г. Э. Бартолинусом (1625—1698). Оно впервые было объяснено Х. Гюйгенсом с помощью представления об эллиптической вторичной волне.

Построение Гюйгенса. Двойное лучепреломление анализируется с помощью построения, предложенного Гюйгенсом. Оно является обобщением построения Гюйгенса для изотропных сред, с помощью которого выводится закон Снеллиуса (рис. 231) и производится с помощью лучевых (не волновых!)



229

Обыкновенный и необыкновенный лучи в положительном кристалле

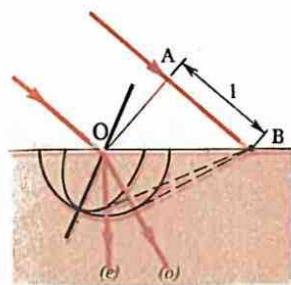


230

Обыкновенный и необыкновенный лучи в положительном кристалле

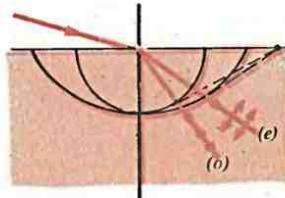
231

Построение Гюйгенса для изотропных сред



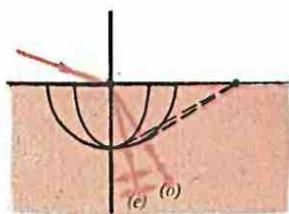
232

Построение Гюйгенса для положительного кристалла при произвольном угле между оптической осью и поверхностью кристалла



233

Двойное лучепреломление на поверхности отрицательного кристалла, когда оптическая ось перпендикулярна поверхности



234

Двойное лучепреломление на поверхности положительного кристалла, когда оптическая ось перпендикулярна поверхности

поверхностей. Построение Гюйгенса для положительного кристалла, оптическая ось которого направлена под произвольным углом к поверхности кристалла, изображено на рис. 232. Отрезок AB принимается за единицу. Точка O принимается за центр сечения лучевой поверхности. Радиус окружности сечения для обычного лука равен $1/n_o$, а эллиптическое сечение для необыкновенного лука чертится так, чтобы расстояние от центра до точки эллипса было равно $1/n_e = v_e/c$. После этого из B проводятся касательные к окружности и эллипсу. Прямые, проведенные через точку O и точки касания, являются искомыми лучами: обычный проходит через точку касания на окружности, а необыкновенный — на эллипсе.

На рис. 232 изображена наиболее простая ситуация, когда входящий в кристалл луч лежит в главной плоскости кристалла. Благодаря этому все построение Гюйгенса удается выполнить в плоскости чертежа. Если входящий в кристалл луч не лежит в главной плоскости, то построение Гюйгенса становится пространственным. В этом случае необходимо строить эллипсоиды, сферы и плоскости, но принцип нахождения преломленных лучей при этом не изменяется: преломленные лучи из точки O проходят через точки касания эллипсоида и сферы с соответствующими плоскостями.

Рассмотрим различные характерные случаи двойного лучепреломления.

Оптическая ось перпендикулярна поверхности кристалла. При нормальном падении луч направлен параллельно оптической оси и, следовательно, распространяется, как в изотропной среде, — двойного лучепреломления нет. При падении луча под углом к поверхности кристалла наблюдается двойное лучепреломление, характер которого зависит от типа кристалла. В отрицательном кристалле (рис. 233) обычный луч преломляется сильнее необыкновенного. Точки и стрелки на этом рисунке показывают направление колебаний электрического вектора волны. В положительном кристалле (рис. 234) сильнее преломляется необыкновенный луч.

Оптическая ось параллельна поверхности кристалла. При падении луча по нормали (рис. 235) в кристалле без пространственного разделения образуются два луча: в обычном луче вектор E перпендикулярен оптической оси (на рис. 235 это направление обозначено точками), а в необыкновенном — параллелен (обозначено стрелками). При выходе из пластины кристалла лучи приобретают разность фаз и в результате суперпозиции образуют эллиптически поляризованную волну (см. § 43). Если падающий свет естественный, то на выходе образуются эллиптически поляризованные волны со всевозможными ориентациями эллипсов и соотношениями их осей. Этот свет по своим свойствам является также естественным.

При падении под углом к поверхности кристалла особенности двойного лучепреломления зависят от угла между плоскостью падения и главной плоскостью. Если плоскость падения

и главная плоскость совпадают (рис. 236), то обыкновенный и необыкновенный лучи лежат в той же плоскости. На рис. 236 изображено двойное лучепреломление для отрицательного кристалла. Построение Гюйгенса для положительного кристалла аналогично.

Если плоскость падения перпендикулярна оптической оси (рис. 237), то обыкновенный и необыкновенный лучи находятся в плоскости падения, но показатель преломления для обоих лучей не зависит от направления. На рис. 237 оптическая ось перпендикулярна плоскости рисунка.

Если плоскость падения пересекает оптическую ось под углом, отличным от $\pi/2$, то картина двойного лучепреломления усложняется, поскольку в этом случае необходимо выполнить пространственное построение Гюйгенса. На рис. 238 показано сечение лучевых поверхностей на поверхности кристалла, которое совпадает с плоскостью рисунка. Стрелкой, оканчивающейся в точке O , показана проекция падающего луча на поверхность кристалла, а пунктирными стрелками, оканчивающимися на окружности и эллипсе, — проекции обыкновенного и необыкновенного лучей на поверхность кристалла. Основное заключение состоит в том, что обыкновенный луч лежит в плоскости падения, а необыкновенный — выходит из нее. Для получения более детальной наглядной информации необходимо построить пространственную модель.

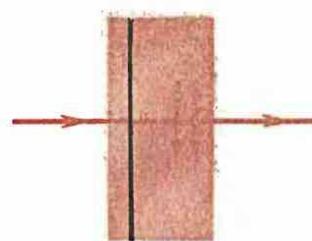
Оптическая ось под углом к поверхности кристалла. Наиболее простой случай, когда плоскость падения совпадает с главной плоскостью, показан на рис. 232 в связи с объяснением сущности построения Гюйгенса. При нормальном падении (рис. 239) обыкновенный луч сохраняет направление падающего, а необыкновенный меняет направление, в результате чего лучи расходятся.

Если плоскость падения не совпадает с главной плоскостью, то картина двойного лучепреломления усложняется, поскольку для ее анализа необходимо использовать пространственное построение Гюйгенса.

Закон Малиса. При анализе нормального падения луча на пластинку, вырезанную из кристалла параллельно оптической оси (рис. 235), необходимо прежде всего определить амплитуды колебаний в обыкновенном и необыкновенном лучах. Ответ на этот вопрос дается **законом Малиса**. Если β — угол между линией колебаний вектора E и оптической осью (рис. 240), I — интенсивность падающего луча, то интенсивности обыкновенного и необыкновенного лучей

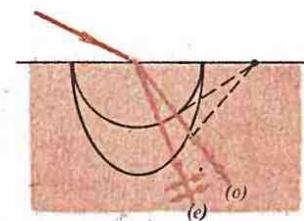
$$I_o = I \sin^2 \beta, \quad I_e = I \cos^2 \beta. \quad (42.1)$$

Это означает, что электрический вектор E падающей волны можно представить как сумму составляющих, перпендикулярных и параллельных оптической оси, которые являются электрическими векторами обыкновенной и необыкновенной волн.



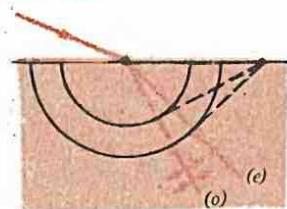
235

Двойное лучепреломление при нормальном падении на поверхность кристалла, когда оптическая ось параллельна поверхности



236

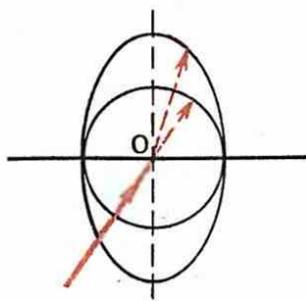
Двойное лучепреломление на поверхности отрицательного кристалла, когда оптическая ось параллельна поверхности, а плоскость падения совпадает с главной плоскостью



237

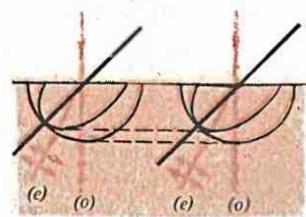
Двойное лучепреломление, когда плоскость падения перпендикулярна оптической оси

- В обыкновенном луче вектор напряженности электрического поля волны перпендикулярен главной плоскости, необыкновенном — лежит в главной плоскости.



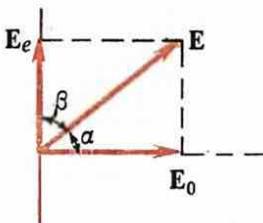
238

Проекция лучей на поверхность кристалла, когда оптическая ось не перпендикулярна плоскости падения



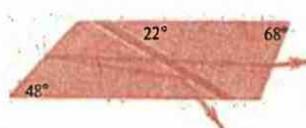
239

Двойное лучепреломление при нормальном падении, когда оптическая ось кристалла не параллельна его поверхности



240

Истекование закона Малюса



241

Призма Николя

Поляризация при двойном лучепреломлении. Поскольку обычный и необычный лучи обладают линейной поляризацией во взаимно перпендикулярных плоскостях, двойное лучепреломление может быть использовано для получения поляризованных лучей. Для этого необходимо развести друг от друга пространственно обычный и необычный лучи или ликвидировать один из них путем сильного поглощения.

Поляроид. Если на выходе из кристаллической пластинки (см. рис. 235) один из лучей сильно ослабляется в результате поглощения, то из пластины выходит линейно поляризованный свет. Такая пластина называется **поляроидом**. Хорошим поляроидом являются кристаллы турмалина. Уже при толщине кристалла турмалина около 1 мм в нем практически полностью поглощается обычный луч. В прошедшем луче (необычном) электрический вектор колеблется параллельно оптической оси (см. рис. 235). Хорошим поляроидом также является герапатит, в котором уже при толщине 0,1 мм практически полностью поглощается один из лучей.

Если поляроид используется для получения поляризованного света, то он называется **поляризатором**. Если же он используется для анализа поляризации света (см. § 43), то его называют **анализатором**.

Поляризационные и двоякотрепломлюющие призмы. Комбинация кристаллов, дающая поляризованный свет, называется **поляризационной** или **двойкотрепломлюющей** призмой. **Поляризационной** призма называется тогда, когда на выходе имеется один поляризованный луч, а **двойкотрепломлющей** — когда на выходе оба луча.

Призма Николя. Она является **поляризационной** призмой и изготавливается из исландского шпата. Кристаллы вырезают относительно оптической оси так, как указано на рис. 241; и склеивают канадским бальзамом по поверхности, отмеченной на рисунке более темным слоем. Коэффициент преломления канадского бальзама $n = 1,550$; он имеет числовое значение, заключенное между коэффициентами преломления обычного и необычного лучей. При соответствующем выборе направления падающего луча необычный луч проходит через призму, а обычный на поверхности склейки испытывает полное отражение и выводится из призмы или поглощается на ее заслоненной поверхности. Призма Николя (ее часто называют просто николем) является наиболее широко распространенной поляризационной призмой.

Двойкотрепломлюющие призмы. Двойкотрепломлюющая призма из стекла и исландского шпата показана на рис. 242. Точки на призме обозначено направление оптической оси кристалла. На рис. 243, *a, b, c* изображены двоякотрепломлюющие призмы, составленные из комбинации призм, изготовленных из исландского шпата, с различной взаимной ориентированной оптических осей.

Полихроизм. Поглощение обыкновенной и необыкновенной волн в кристалле зависит не только от длины волны, но и от ее направления. В результате этого при прохождении белого света возникает окраска кристалла, зависящая от направления распространения света. Это явление называется полихроизмом.

Пример 42.1. На рис. 243 изображена призма Волластона, изготовленная из исландского шпата. Показатель преломления обыкновенного луча $n_o = 1,658$, а необыкновенного, когда вектор \mathbf{E} коллинеарен оптической оси, $n_e = 1,486$. Угол $\alpha = 15^\circ$. Найти угол между выходящими из призмы лучами.

При переходе через границу раздела между средами с взаимно перпендикулярными оптическими осями обыкновенный в первой среде луч становится необыкновенным во второй и, наоборот, необыкновенный луч в первой среде становится обыкновенным во второй. Обозначая $\theta_{\text{пр } o}$ и $\theta_{\text{пр } e}$ углы преломления обыкновенного и необыкновенного лучей, можем записать

$$\frac{\sin \theta_{\text{пр } o}}{\sin \theta_{\text{пр } e}} = \frac{n_e}{n_o} < 1, \quad \frac{\sin \theta_{\text{пр } e}}{\sin \theta_{\text{пр } o}} = \frac{n_o}{n_e} > 1, \quad (42.2)$$

причем $\sin \theta_{\text{пр } o} = \sin \theta_{\text{пр } e} = \sin \alpha$.

Обозначая $\theta'_{\text{пр } o}$ и $\theta'_{\text{пр } e}$ углы преломления при выходе из призмы в воздух, запишем законы преломления в виде

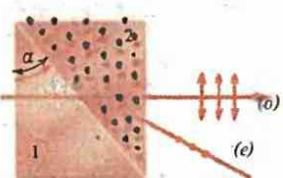
$$\sin(\alpha - \theta_{\text{пр } e}) / \sin \theta'_{\text{пр } e} = 1/n_e, \quad \sin(\alpha - \theta_{\text{пр } o}) / \sin \theta'_{\text{пр } o} = 1/n_o, \quad (42.3)$$

поскольку углы падения на границу при выходе из призмы обыкновенного и необыкновенного лучей равны соответственно $\alpha - \theta_{\text{пр } o}$ и $\alpha - \theta_{\text{пр } e}$. Из (42.3) с учетом (42.2) получаем для углов преломления значения $\theta'_{\text{пр } e} = 2^\circ 14'$ и $\theta'_{\text{пр } o} = 3^\circ 2'$, причем эти углы отсчитываются в разные стороны от нормали. Поэтому угол между вышедшими из призмы лучами $\theta'_{\text{пр } o} + \theta'_{\text{пр } e} = 5^\circ 16'$.

§ 43 Интерференция поляризованных волн

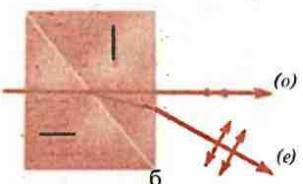
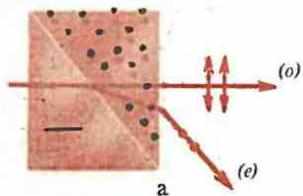
Рассматривается интерференция волн с взаимно перпендикулярными плоскостями поляризации.

Интерференция лучей при взаимно перпендикулярных направлениях линейной поляризации. Лучи с взаимно перпендикулярными направлениями поляризации могут быть получены по схеме рис. 235, если на пути луча до входа в кристаллическую пластинку поставить призму Николя N (рис. 244), которая создает линейно поляризованный свет, причем направление колебаний электрического вектора может меняться поворотом николя. Линейно поляризованный свет падает на кристаллическую пластинку, где распадается на обыкновенный и необыкновенный лучи. Для амплитуд этих лучей на основании закона Малюса имеем (рис. 245)



242

Двоякопереломляющая призма из стекла и исландского шпата



243

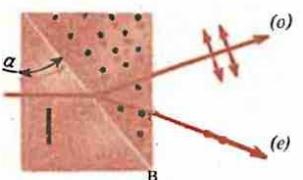
Двоякопереломляющие призмы из исландского шпата

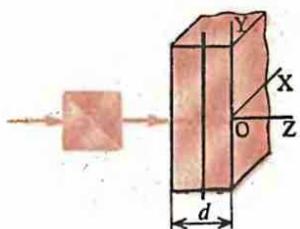
○ В каких случаях преломленный луч лежит в плоскости падения?

В каких случаях преломленный луч выходит из плоскости падения?

Перечислите известные вам поляризационные и двоякопереломляющие призмы.

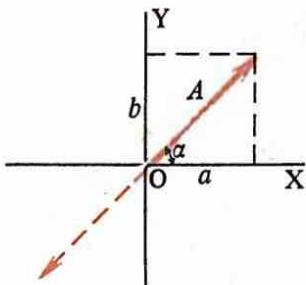
Для каких целей используются поляроиды?





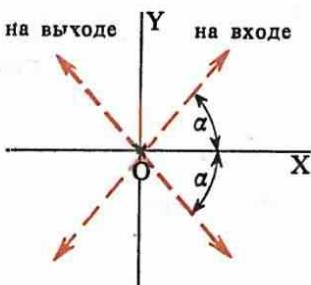
244

Схема осуществления интерференции лучей с взаимно перпендикулярными направлениями линейной поляризации



245

К определению амплитуд обыкновенной и необыкновенной волн на входе в пластинку



246

Поворот плоскости колебаний при выходе из пластины в полуволны

$$a = A \cos \alpha,$$

$$b = A \sin \alpha,$$

где A — амплитуда падающего луча. Представим колебания электрических векторов обычной (вдоль оси X) и необычайной волн на входе в пластинку в виде

$$x = a \cos \omega t',$$

$$y = b \cos \omega t',$$

где t' — текущее время, описывающее колебание на входе в пластинку. Обычный луч на прохождение пластины толщиной d затрачивает время $n_o d/c$, а необычайный — $n_e d/c$. Обозначая t текущее время на выходе из пластины, для обычного и необычайного лучей получаем следующие соотношения между временем на входе и временем на выходе:

$$t = t' + n_o d/c,$$

$$t = t' + n_e d/c.$$

(43.3)

Поэтому на выходе из пластины колебания электрических векторов обычной и необычайной волн на основании (43.2) и (43.3) описываются формулами

$$x = a \cos (\omega t - kn_o d),$$

(43.4)

$$y = b \cos (\omega t - kn_e d),$$

где $k = \omega/c$. Таким образом, между волнами образуется разность фаз $\phi = kd(n_e - n_o)$. Выражения (43.4) можно представить в виде

$$x = a \cos \xi,$$

$$y = b \cos (\xi - \phi),$$

(43.5)

где $\xi = \omega t - kn_o d$. Равенства (43.5) могут рассматриваться как параметрическое представление эллипса, уравнение которого имеет вид (см. § 5)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{2xy}{ab} \cos \phi = \sin^2 \phi.$$

(43.6)

Форма и ориентация эллипса зависят от a и ϕ (a зависит от соотношения между a и b).

Пластинка в четверть волны. При $d(n_e - n_o) = \pm \lambda/4$, $\phi = \pm \pi/2$ уравнение (43.6) принимает вид

$$x^2/a^2 + y^2/b^2 = 1.$$

(43.7)

Это означает, что главные оси эллипса совпадают с главными осями кристалла. При $a=b$, т. е. при $\phi=\pi/4$, уравнение (43.7) преобразуется в уравнение окружности

$$x^2 + y^2 = a^2,$$

(43.8)

т. е. на выходе из пластины возникает циркулярно поляризованная волна.

Знак разности $n_e - n_o$ зависит от типа кристалла. Для положительного кристалла $(n_e - n_o) > 0$, для отрицательного

$(n_e - n_o) < 0$. Направление вращения вектора \mathbf{E} определяется не только типом кристалла, но и толщиной пластинки d . Это следует из того, что уравнение (43.6) принимает вид (43.7) не только при $d(n_e - n_o) = \pm \lambda/4$, но и при $d(n_e - n_o) = \lambda/4 + m\lambda/2$ ($m = \pm 1, \pm 2, \dots$).

Пластинка в полволны. При $d(n_e - n_o) = \pm \lambda/2$, $\varphi = \pm \pi$ уравнение (43.6) принимает вид

$$(x/a + y/b)^2 = 0, \quad x/a + y/b = 0. \quad (43.9)$$

Следовательно, на выходе из пластины возникает линейно поляризованная волна, для которой лишь изменились квадранты плоскости XY , в которых происходит колебание (рис. 246), т. е. плоскость поляризации повернулась на угол 2α . Конечно, эта ситуация возникает не только при $d(n_e - n_o) = \pm \lambda/2$, но и при $d(n_e - n_o) = \lambda/2 + m\lambda$ ($m = \pm 1, \pm 2, \dots$).

Пластинка в целую волну. При $(n_e - n_o)d = m\lambda$, $\varphi = 2\pi m$ уравнение (43.6) принимает вид

$$(x/a - y/b)^2 = 0, \quad x/a - y/b = 0. \quad (43.10)$$

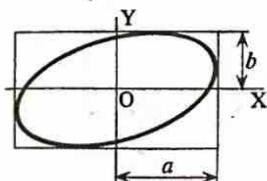
Значит, на выходе имеется линейно поляризованная волна с той же плоскостью колебаний вектора \mathbf{E} , что и на входе.

Анализ линейно поляризованного света. На пути луча ставится николь таким образом, чтобы его главное направление расположалось перпендикулярно лучу. При изменении этого направления (вращением николя) возможны два исхода: интенсивность пропускаемого николем света либо не изменяется, либо изменяется. Если интенсивность пропускаемого света не изменяется, то он либо вообще не поляризован, либо циркулярно поляризован. (Как отличить неполяризованный свет от циркулярно поляризованного, будет сказано позже). Если интенсивность пропускаемого света при вращении николя изменяется, то возможны два случая: либо интенсивность уменьшается до нуля, либо она нуля никогда не достигает. Если интенсивность уменьшается до нуля, то свет линейно поляризован, причем электрический вектор колеблется коллинеарно главному направлению николя в положении, при котором пропускаемая им интенсивность света максимальна. Если интенсивность не уменьшается до нуля, то свет либо эллиптически поляризован, либо частично поляризован. Причина изменений интенсивности пропускаемого через николь света при его повороте для эллиптической поляризации показана на рис. 247. Амплитуда пропускаемых николем колебаний равна расстоянию между касательными к эллипсу, перпендикулярными направлению пропускаемых колебаний. Видно, что максимальная пропускаемая амплитуда равна большой полуоси эллипса, минимальная — малой.

Анализ эллиптически поляризованного света. Эллиптически поляризованную волну можно рассматривать как результат суперпозиции двух линейно поляризованных волн с разностью

● **Поляризация волны, образовавшейся в результате интерференции двух волн с взаимно перпендикулярными направлениями линейной поляризации, зависит от разности фаз между ними и от соотношения их амплитуд**

○ **Опишите метод определения поляризации света. Чем отличается частично поляризованный свет от эллиптически поляризованного? В чем причина образования креста при наблюдении интерференции в сходящихся лучах?**



247

К анализу эллиптически поляризованного света

фаз $\pi/2$ при условии, что их направления поляризации совпадают с главными осями эллипса. Поэтому если на пути эллиптически поляризованного луча поставить пластинку в четверть волны так, чтобы главная ось пластины была параллельна главной оси эллипса, то из пластины выйдет линейно поляризованная волна. Главные же оси эллипса определяются по ориентировке николя, при которой он пропускает максимальную и минимальную интенсивности света. Поэтому, расположив главную ось пластины в четверть волны в одном из найденных направлений и проанализировав выходящий из нее свет на линейную поляризацию, делаем заключение о том, является ли свет эллиптически поляризованным или частично поляризованным. Если выходящий из пластины в четверть волны свет линейно поляризован, то входящий в нее свет эллиптически поляризован, а если частично поляризован, то и входящий в пластины свет частично поляризован.

Анализ циркулярио поляризованного света. Если при вращении николя интенсивность выходящего из него света не изменяется, то входящий в николь свет может быть либо циркулярио поляризованным, либо неполяризованным. Поставим на пути света пластины в четверти волны, причем в отличие от эллиптической поляризации теперь нет необходимости заботиться о выборе направления главной оптической оси пластины. Если выходящий из пластины в четверть волны свет линейно поляризован, то входящий циркулярио поляризован. В противном случае входящий свет просто неполяризован.

Компенсаторы. Для упрощения анализа эллиптически поляризованного света используется компенсатор, представляющий собой пластинку, вдоль которой непрерывно изменяется разность фаз между лучами на выходе из пластины. На рис. 248 изображен один из таких компенсаторов, для которого получим

$$\Delta = (n_e d_1 + n_o d_2) - (n_o d_1 + n_e d_2) = (n_e - n_o)(d_1 - d_2). \quad (43.11)$$

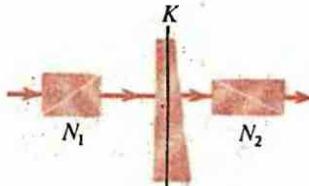
Поэтому между выходящими лучами образуются всевозможные разности фаз. В тех местах, где за счет распространения волны в компенсаторе разность фаз между входящими лучами становится равной $0, 2\pi, 4\pi, \dots$, возникает линейная поляризация. В других местах выходящий свет имеет эллиптическую поляризацию (в том числе циркулярионную) с изменяющимися параметрами эллипса. Посмотрев на поверхность компенсатора через николь, обнаружим периодическое изменение освещенности, которое и свидетельствует о поляризации входящего в компенсатор света. Создаваемая пластиною разность фаз между лучами известна [см. (43.11)]. Поэтому, зная распределение интенсивностей при различных положениях поляризатора (анализатора), можно определить характер поляризации входящего в компенсатор света.

Цвета кристаллических пластинок. Возьмем кристаллическую пластинку переменной толщины, оптическая ось в которой направлена перпендикулярно проходящему через нее лучу



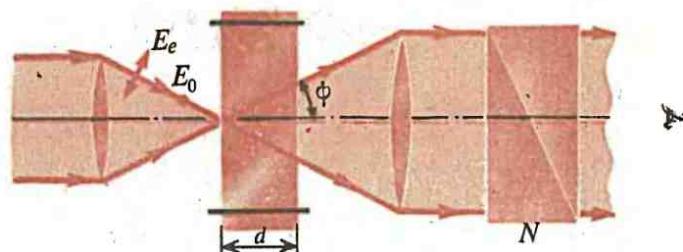
248

Компенсатор



249

К объяснению возникновения цветов тонких пластинок



К анализу явлений в сходящихся пучках

(рис. 249). Николь N_1 пропускает лишь колебания E , параллельные его главному направлению. На выходе из пластинки K из-за изменяющейся толщины образуется эллиптически поляризованный свет с различной ориентировкой эллипсов. Анализатор N_2 оставляет от каждого колебания лишь проекцию вектора E на направление, в котором он пропускает колебания. В результате этого при наблюдении через анализатор N_2 пластинки K , освещаемой монохроматическим светом, последняя будет казаться неравномерно освещенной. При освещении белым светом возникает цветовая окраска поверхности кристаллической пластины.

Явления в сходящихся лучах. На рис. 250 изображен ход лучей в плоскости чертежа. Кристаллическая пластина вырезана перпендикулярно оптической оси. Из рис. 250 ясно, что картина прохождения лучей через кристаллическую пластину аксиально-симметрична, если падающий пучок света неполяризован.

Направление колебаний электрического вектора необыкновенной волны лежит в плоскости рисунка (изображено стрелками), а обыкновенной — перпендикулярно (изображено точками). Длина пути луча в пластине $h = d/\cos \psi$, где ψ — угол между лучом и осью. Поэтому разность фаз между лучами, имеющими данное направление, равна

$$\Delta = (d/\cos \psi) (n_e - n_o). \quad (43.12)$$

Разность фаз между лучами на выходе из пластины составляет

$$\delta = \frac{2\pi}{\lambda} \frac{d}{\cos \psi} (n_e - n_o), \quad (43.13)$$

т. е. различна для различных углов. Кроме того, формула (43.13) показывает, что картина аксиально-симметрична. Следовательно, под различными углами к оси на выходе из пластины распространяются эллиптически поляризованные лучи с различными ориентациями эллипсов. Если излучение монохроматическое и на пути вышедших из пластины лучей стоит николь N_2 , то наблюдаются чередующиеся темные и светлые кольца, которые соответствуют проекциям электрического вектора на главное направление николя N_2 . При повороте николя на $\pi/2$ светлые кольца становятся темными, и наоборот. При освещении белым светом наблюдается совокупность окрашенных концентрических окружностей.

Если на пути падающего на пластину пучка лучей поставить николь N_1 , то наблюдаемая картина изменяется — концентрические окружности пересекаются крестом, причем крест светлый, если главные направления николей N_1 и N_2 параллельны, и темный — если они перпендикулярны. Это объясняется тем, что в падающем пучке в плоскости падения, совпадающей с главной плоскостью николя N_1 , имеются лишь необыкновенные лучи и, следовательно, отсутствует интерференция с обычными лучами. Поэтому на выходе из пластины в этой плоскости свет линейно поляризован. Если главное направление николя N_2 лежит в этой плоскости, то лучи в этой плоскости проходят без ослабления и поэтому наблюдается светлая полоса. Если николь N_2 скрещен с николем N_1 , то вместо светлой полосы наблюдается темная.

Аналогичным способом рассматриваются лучи, лежащие в плоскости, перпендикулярной плоскости рисунка. Лучи в этой плоскости при той же исходной поляризации являются только обычновенными. Из пластины выходит линейно поляризованный луч, который затем проходит через николь N_2 , поскольку направление колебаний его вектора E такое же, как и у необыкновенного луча, лежащего в перпендикулярной плоскости. В результате возникает светлая полоса, перпендикулярная светлой полосе от необыкновенного луча. Совокупностью этих двух светлых полос образуется крест, который пересекает концентрические окрашенные окружности.

§ 44

Вращение плоскости поляризации

Дается феноменологическое описание естественной и искусственной оптической активности веществ.

Вращение плоскости поляризации в кристаллических телах. При прохождении линейно поляризованного луча вдоль оптической оси кварцевой пластиинки (рис. 251) наблюдается поворот плоскости поляризации (Араго, 1811). Разделения луча на два при нормальном падении на пластинку, вырезанную перпендикулярно оптической оси, не происходит. Угол поворота плоскости поляризации определяется по взаимной ориентировке осей николей N_1 (поляризатор) и N_2 (анализатор). Установлено экспериментально, что угол поворота зависит от длины d пути в кристаллической пластине и от длины волны, т. е. имеется вращательная дисперсия:

$$\phi = ad, \quad (44.1)$$

где a — вращательная способность. Эта величина выражается в радианах на метр (в СИ) или в градусах на миллиметр ($^{\circ}/\text{мм}$). У кварца для красной области спектра $a \approx 15^{\circ}/\text{мм}$, для зеленой $a \approx 27^{\circ}/\text{мм}$, фиолетовой — $a \approx 51^{\circ}/\text{мм}$. Эти данные показывают, что вращательная дисперсия у кварца весьма значительна.

Направление вращения плоскости поляризации изменяется при изменении направления распространения света на обратное и может определяться правилом винта.

Экспериментально установлено, что существуют две модификации кристаллов кварца — правовращающая и левовращающая. Они характеризуются различными направлениями вращения плоскости поляризации, которые определяются в соответствии с правилом правого или левого винта при распространении света вдоль оптической оси. Исторически сложилось так, что направление вращения плоскости поляризации устанавливается для наблюдателя, к которому направлен луч света. Поэтому правовращающая модификация кварца (рис. 252, а) обуславливает левовинтовое вращение плоскости поляризации, а левовращающая (рис. 252, б) — правовинтовое.

Установлено, что если кристалл вращает плоскость поляризации, то у него всегда существуют обе модификации, причем

$$|\alpha_{(+)}| = |\alpha_{(-)}|, \quad (44.2)$$

где $\alpha_{(+)}$ и $\alpha_{(-)}$ — вращательные способности правовращающей и левовращающей модификаций кристалла. (Будем в дальнейшем эти модификации называть соответственно правый и левый, кристалл.)

Вращение плоскости поляризации в аморфных веществах. Во многих аморфных веществах, в том числе и в жидкостях (например, в скапидаре), также наблюдается вращение плоскости поляризации, причем в ряде случаев довольно значительное (например, в никотине $a \approx 1,64^{\circ}/\text{мм}$ для желтой области спектра).

Аморфные вещества также существуют в двух модификациях — право- и левовращающей. Если вращающее плоскость поляризации вещество находится в растворе, то угол поворота плоскости вращения при прохождении света через раствор определяется по закону Био (1832):

$$\varphi = aq/l, \quad (44.3)$$

где q — концентрация раствора, l — длина пути в растворе. Вращательная способность a зависит от длины волны и температуры раствора. Экспериментальное изучение этого явления показало, что зависимость a от температуры слабая, а от длины волны дается в грубом приближении соотношением $a \sim 1/\lambda^2$.

Вращение плоскости поляризации в аморфных веществах является молекулярным свойством. Молекуле можно соответствующим образом приписать определенную вращательную способность, которая не зависит от агрегатного состояния вещества. Кристаллическая структура вещества также может обуславливать вращательную способность. Вещества, способные вращать плоскость поляризации, называются оптически активными.

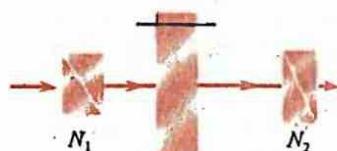
Феноменологическая теория вращения плоскости поляризации. Линейно поляризованная волна может рассматриваться как суперпозиция двух циркулярно поляризованных волн с противоположными направлениями вращения электрического вектора [см. (5.10) и (5.11) и рис. 19]. В основе теории лежит предположение о различной скорости распространения циркулярно поляризованных волн с различными направлениями вращения электрического вектора (Френель, 1817). У входящего в кристалл луча, описываемого формулами (5.10) и (5.11), на выходе из кристалла между колебаниями векторов E_1 и E_2 возникает постоянная разность фаз, которая обуславливает поворот плоскости поляризации.

Считая, что волны (см. рис. 19) распространяются в положительном направлении оси Z (к нам), и обозначая коэффициенты преломления циркулярно поляризованных волн с векторами E_1 и E_2 соответственно n_1 и n_2 , можно на основании (5.10) и (5.11) для выходящих из кристалла волн записать

$$E_{1x} = E_0 \cos(\omega t - kn_1 d), \quad E_{1y} = E_0 \sin(\omega t - kn_1 d), \quad (44.4)$$

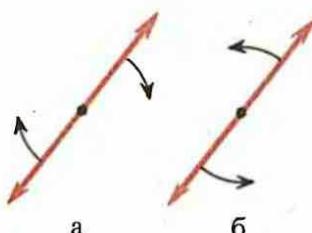
$$E_{2x} = E_0 \cos(\omega t - kn_2 d), \quad E_{2y} = -E_0 \sin(\omega t - kn_2 d), \quad (44.5)$$

где d — расстояние, пройденное волной во вращающей плоскости поляризации среде. Ясно, что теперь векторы E_1 и E_2 расположены симметрично относительно линии $y=0$, как на рис. 19, а относительно другой линии, которая и определяет поворот плоскости колебаний результирующего вектора $E_1 + E_2$. Чтобы найти положение этой плоскости, определим положения векторов E_1 и E_2 в некоторый момент времени, например $t = kn_1 d/\omega$. Тогда [см. (44.4) и (44.5)]



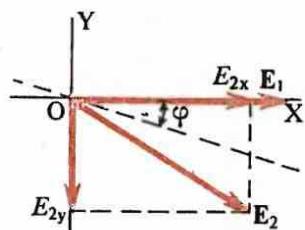
251

Вращение плоскости поляризации при прохождении света через пластинку кварца вдоль оптической оси



252

Вращение плоскости поляризации в правовращающем (а) и левовращающем (б) кварце



253

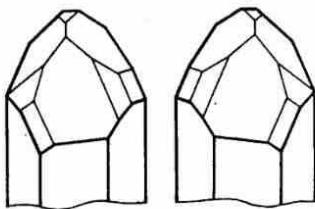
К расчету угла поворота плоскости поляризации

Вращение плоскости поляризации объясняется различной скоростью циркулярно поляризованных волн с правым и левым вращением.



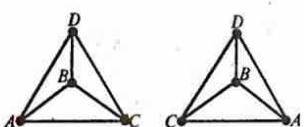
254

Эксперимент, доказывающий существование циркулярно поляризованных лучей в оптически активных веществах



255

Кристаллы правого и левого кварца



256

Зеркально-симметричные молекулы

- Что такое зеркальные изомеры? Приведите примеры пространственной структуры молекул, являющихся зеркальными изомерами. Существуют ли в живой природе правовращающие и левовращающие зеркальные изомеры? Какие существуют?

$$E_{1x} = E_0, \quad E_{1y} = 0,$$

$$E_{2x} = E_0 \cos [kd(n_1 - n_2)],$$

$$E_{2y} = -E_0 \sin [kd(n_1 - n_2)].$$

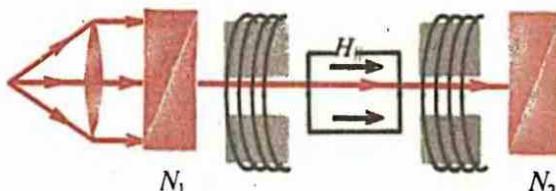
На рис. 253 изображено положение этих векторов для случая $(n_1 - n_2) > 0$. Если $(n_1 - n_2) < 0$, то вектор E_2 повернут относительно E_1 в другом направлении. Линия, относительно которой векторы E_1 и E_2 расположены симметрично, является гипотенузой угла между этими векторами, обозначенного 2ϕ на рис. 253, и, следовательно, плоскость поляризации поворачивается на угол

$$\phi = k(n_1 - n_2) d/2.$$

Поскольку направление вращения вектора E в циркулярно поляризованной волне принято характеризовать со стороны наблюдателя, к которому луч направлен, то n_1 является показателем преломления волны с левой круговой поляризацией, а n_2 — с правой. Отсюда заключаем, что для правовращающих веществ $n_1 > n_2$, а для левовращающих $n_1 < n_2$.

Справедливость теории подтверждается экспериментами по преломлению циркулярно поляризованных волн (рис. 254). В результате различного соотношения между скоростями волн с правой и левой круговой поляризацией в право- и левовращающих средах на границе между ними происходит двойное лучепреломление, причем от падающей линейно поляризованной волны возникают две циркулярно поляризованные волны, которые можно пространственно разделить (рис. 254). Эксперименты показали, что эти лучи действительно циркулярно поляризованы, а преломление на границах между средами происходит в соответствии с построением Гюйгенса, проведенным для скоростей на основе анализа вращения плоскости поляризации. Таким путем феноменологическая теория вращения плоскости поляризации была подтверждена экспериментально.

Оптическая изомерия. Правые и левые кристаллы имеют одинаковый химический состав, но различную форму. Они являются зеркальным отражением друг друга. На рис. 255 изображена часть граней правого и левого кристаллов кварца. Аналогичное различие наблюдается у молекул аморфных оптически активных веществ. На рис. 256 показаны две молекулы, в которых атомы A и C различного сорта. Ясно, что эти молекулы являются как бы зеркальным отображением друг друга. Они называются оптическими изомерами. Никакими пространственными движениями предмет не может быть совмещен со своим зеркальным отражением. Поэтому правые и левые кристаллы, правые и левые молекулы существуют в природе раздельно друг от друга. В частности, молекулы аминокислот, входящие в состав живых систем на Земле, оптически активны и все лево-



257

Схема наблюдения вращения плоскости поляризации при помещении оптически неактивного вещества в магнитное поле

вращающие. Поскольку строение молекул, особенно органических веществ, очень сложно, возникают возможности образования многих различных пар оптических изомеров.

Вращение плоскости поляризации в магнитном поле. Оптически неактивные вещества в магнитном поле становятся оптически активными и вращают плоскость поляризации света, распространяющегося в веществе вдоль силовых линий напряженности магнитного поля. Этот эффект был открыт в 1846 г. М. Фарадеем (1791—1867) и называется явлением Фарадея. Оно выражает влияние магнитного поля на вещество, в результате чего вещество становится оптически активным.

Схема наблюдения вращения плоскости поляризации оптически неактивным веществом в магнитном поле H_{\parallel} изображена на рис. 257. Угол поворота ϕ плоскости поляризации определяется соотношением

$$\phi = VH_{\parallel} l, \quad (44.9)$$

где V — постоянная Верде, характеризующая свойства вещества и зависящая от частоты света и температуры; H_{\parallel} — проекция вектора напряженности магнитного поля на направление распространения света.

Формула (44.9) описывает угол поворота плоскости поляризации в пара- и диамагнитных материалах.

Углы поворота ϕ невелики. Для большинства твердых тел при напряженности поля порядка 1 МА/м и $l \approx 0,1$ м угол поворота составляет $1-2^\circ$, для газов — еще меньше.

В ферромагнитных материалах вращение плоскости поляризации принято выражать не через напряженность магнитного поля, а через намагниченность J . При этом величина H в формуле (44.9) заменяется на J , а постоянная Верде V — на постоянную Кундта, значение которой приводится в таблицах.

Направление вращения плоскости поляризации принято определять для наблюдателя, смотрящего в направлении вектора напряженности магнитного поля (а в ферромагнетиках — вектора намагниченности).

Большинство веществ является правовращающими. Такие вещества называют положительными. К ним относятся все диамагнетики. Имеются и отрицательные вещества, которые обязательно содержат парамагнитные атомы.

Направление вращения плоскости поляризации в магнитном поле зависит только от направления распространения света. Это отличает вращение плоскости поляризации в магнитном поле от естественной оптической активности и показывает, что оно обусловлено действием магнитного поля на вещество.

Постоянная Верде выражается законом Био

$$V = A/\lambda^2 + B/\lambda^4, \quad (44.10)$$

где A и B слабо зависят от температуры.

Описывается различные методы создания искусственной анизотропии.

Анизотропия при деформации. Оптически изотропное тело при деформации сжатия или растяжения (рис. 258) приобретает свойства одноосного кристалла, оптическая ось которого коллинеарна с направлением деформирующих сил. Экспериментально установлена следующая связь между показателями преломления необыкновенной и обычной волн в направлении, перпендикулярном оптической оси:

$$n_e - n_o = b\sigma, \quad (45.1)$$

где $\sigma = F/S$ — напряжение, вызвавшее деформацию; b — постоянная, характеризующая свойства вещества. Разность $(n_e - n_o)$ может принимать как положительные, так и отрицательные значения и, кроме того, зависит от длины волны света.

При наблюдении прозрачного тела на просвет в скрещенных николях деформированное тело представляется окрашенным, причем окраска зависит от деформации. По распределению окраски можно судить о распределении деформаций в теле.

Анизотропия, создаваемая в веществе электрическим полем. Оптически изотропное вещество в электрическом поле (рис. 259) приобретает свойства одноосного кристалла с оптической осью, коллинеарной вектору напряженности электрического поля (явление Керра, 1875). При распространении света перпендикулярно оптической оси экспериментально установлено следующее соотношение между показателями преломления обычной и необыкновенной волн:

$$n_e - n_o = K\lambda E^2, \quad (45.2)$$

где K — постоянная Керра. Для жидкостей она обычно составляет несколько пикометров на вольт в квадрате. Например, для длины волны 546 нм для жидкого кислорода при 90 К и жидкого азота при 71,4 К она составляет соответственно 10,1 и 4,02 пм/ V^2 . Для газов постоянная Керра значительно меньше. Например, для кислорода и азота она соответственно равна $0,45 \times 10^{-15}$ и $0,3 \cdot 10^{-15} \text{ м}/V^2$ при нормальных условиях. Это обусловлено тем, что эффект Керра определяется свойствами молекул и, следовательно, усиливается с повышением концентрации молекул. Из (45.2) видно, что при изменении направления электрического поля на обратное оптические свойства вещества не меняются, т. е. оно действительно ведет себя как одноосный кристалл.

При прохождении пути l разность оптических путей обычного и необыкновенного лучей равна $\Delta = K\lambda/E^2$, а разность фаз между волнами

$$\varphi = k\Delta = 2\pi K\lambda/E^2. \quad (45.3)$$

Для большинства жидкостей $K > 0$. Постоянная Керра зависит также от длины волны (т. е. имеется дисперсия) и уменьшается при увеличении температуры.

Для объяснения явления Керра надо принять во внимание два физических фактора. Неполярные молекулы в электрическом поле приобретают дипольный момент в направлении поля, а сама молекула при этом переориентируется так, чтобы дипольный момент совпадал с направлением наибольшей поляризуемости молекулы. Следовательно, наибольший показатель преломления оказывается у волны, электрический вектор которой колеблется коллинеарно внешнему электрическому полю, т. е. у необыкновенной волны:

$$n_e > n_o, \quad K > 0. \quad (45.4)$$

Полярные молекулы во внешнем электрическом поле ориентируются своими постоянными дипольными моментами преимущественно в направлении напряженности поля. При этом

направление наибольшей поляризуемости молекулы составляет с электрическим полем, вообще говоря, некоторый угол, в зависимости от которого возможны различные соотношения между n_e и n_o , т. е. различные значения для K , включая $K=0$. Эффект Керра обладает очень малой инерционностью (оптическая анизотропия следует за изменением напряженности электрического поля с запаздыванием порядка 10^{-10} с), поэтому он позволяет создать быстродействующие модуляторы света, называемые ячейками Керра.

Размер ячейки Керра выбирается такой, чтобы ее оптическая длина составила полволны при определенной напряженности электрического поля. Если ячейку поместить между скрещенными николями, главные плоскости которых направлены под углом 45° к оптической оси, возникающей в ячейке Керра при наличии электрического поля, то при отсутствии поля ячейка не пропускает свет. При включении поля свет начинает проникать через нее и при определенном значении напряженности поля, когда ячейка действует как пластинка полволны, интенсивность прошедшего через систему света достигает максимального значения.

Анизотропия, создаваемая в веществе магнитным полем (явление Коттон—Мутона, 1910). Оптически изотропное вещество в магнитном поле приобретает свойства одноосного кристалла, ось которого коллинеарна направлению индукции магнитного поля (рис. 260). Зависимость разности ($n_e - n_o$) от B при распространении света перпендикулярно индукции магнитного поля выражается соотношением

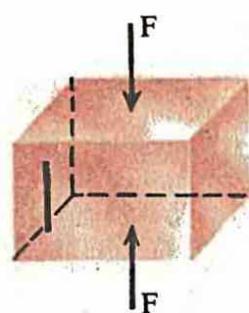
$$n_e - n_o = C\lambda B^2, \quad (45.5)$$

где C — постоянная, характеризующая свойства вещества. Этот эффект очень мал. Его объяснение аналогично объяснению явления Керра. Необходимо отметить, что этот эффект принципиально отличается от явления Фарадея (44.9), поскольку зависит от индукции магнитного поля квадратично, а не линейно.

Постоянная C определяется экспериментально. Например, для нитробензола $C = 2,25 \cdot 10^{-2} \text{ м}^{-1} \cdot \text{Тл}^{-2}$. В магнитном поле с индукцией 1 Тл при прохождении светом расстояния 1 м разность фаз равна 0,14 рад.

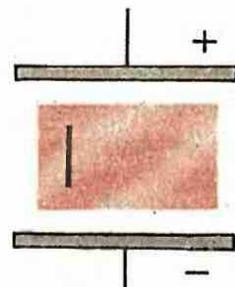
Эффект Покельса. В некоторых кристаллах при наложении внешнего электрического поля возникает двойное лучепреломление, которого нет в отсутствие поля, причем разность показателей преломления необыкновенного и обыкновенного лучей пропорциональна первой степени напряженности электрического поля. Это явление называется эффектом Покельса.

Оптическая ось кристалла ориентируется параллельно лучу света, и напряженность внешнего электрического поля также коллинеарна этому направлению. Это можно осуществить либо взяв прозрачные электроды, либо проделав в центрах электродов маленькие отверстия. При наличии внешнего электрического поля возникает вторая оптическая ось, лежащая в плоскости,



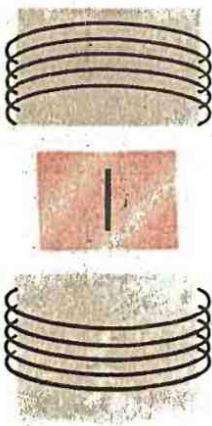
258

Направление оптической оси при деформации сжатия и растяжения



259

Направление оптической оси в веществе в электрическом поле



260

Направление оптической оси в веществе в магнитном поле

перпендикулярной первой оптической оси. Следовательно, индуцированная оптическая ось перпендикулярна направлению распространения света и относительно этой оси свет испытывает двойное лучепреломление. Для этой оси

$$n_e - n_o = aE, \quad (45.6)$$

где a — постоянная. Ориентировка индуцированной оси в плоскости, перпендикулярной направлению луча света, зависит от ориентировки кристалла. Разность потенциалов, которую необходимо приложить между электродами в эффекте Покельса, примерно на порядок меньше разности потенциалов, необходимой для получения в эффекте Керра одинакового двойного лучепреломления (при равных расстояниях между электродами). Это является важным преимуществом эффекта Покельса для практических применений. Например, значение a для кристалла дигидрофосфата калия (KDP, формула KH_2PO_4) составляет около $3,6 \cdot 10^{-11}$ м/В, для ниобата лития — $3,7 \times 10^{-10}$ м/В. Эффект Покельса столь же безынерционен, что и эффект Керра. Он используется для создания быстродействующих модуляторов света. Соответствующее устройство называется ячейкой Покельса.

Эффект Покельса возникает не только при приложении поля вдоль оптической оси кристалла, но и перпендикулярно ей. В первом случае говорят о продольном эффекте Покельса, во втором — о поперечном. Однако в обоих случаях луч света должен распространяться в направлении оптической оси, которую кристалл имеет при отсутствии внешнего электрического поля.

Для создания ячеек Покельса поперечный эффект имеет определенное преимущество перед продольным. Во-первых, электроды расположены параллельно пучку света и расстояние между ними может быть сделано достаточно малым, а длина вдоль луча — достаточно большой. Поэтому создать полу волновую ячейку можно при сравнительно небольшой разности потенциалов между электродами. При использовании продольного эффекта сдвиг фаз между обычновенной и необыкновенной волнами для фиксированной разности потенциалов не зависит от длины ячейки, потому что с увеличением длины уменьшается напряженность электрического поля, а сдвиг фаз остается неизменным. Следовательно, увеличить сдвиг фаз можно только в результате роста разности потенциалов между электродами. Во-вторых, технически проще осуществить ячейку с поперечным эффектом, чем с продольным. Поэтому обычно низковольтные ячейки Покельса основаны на поперечном эффекте. Однако для создания высокоскоростных ячеек предпочтительнее использовать продольный эффект, поскольку в этом случае электроды имеют меньший размер и меньшую электроемкость, что облегчает достижение высоких скоростей изменения потенциалов.

Ячейки Покельса применяются для тех же полей, что и ячейки Керра, и во многих случаях заменили их.

Задачи

- 8.1.** Между двумя николями помещена кристаллическая пластина. Главные плоскости николей образуют с оптической осью углы α и β . Оптическая ось параллельна плоскости кристалла, на которую падает свет. Найти интенсивность вышедшего из пластины света. Толщина пластиинки d .
- 8.2.** Пластиинка кварца толщиной 0,014 мм вырезана параллельно оптической оси. Известно, что $n_e = 1,5533$, $n_0 = 1,5442$ для $\lambda = 0,5 \text{ мкм}$. На пластиину падает линейно

поляризованный свет, плоскость колебаний электрического вектора образует с оптической осью угол 25° . Определить характер поляризации выходящего из пластины луча.

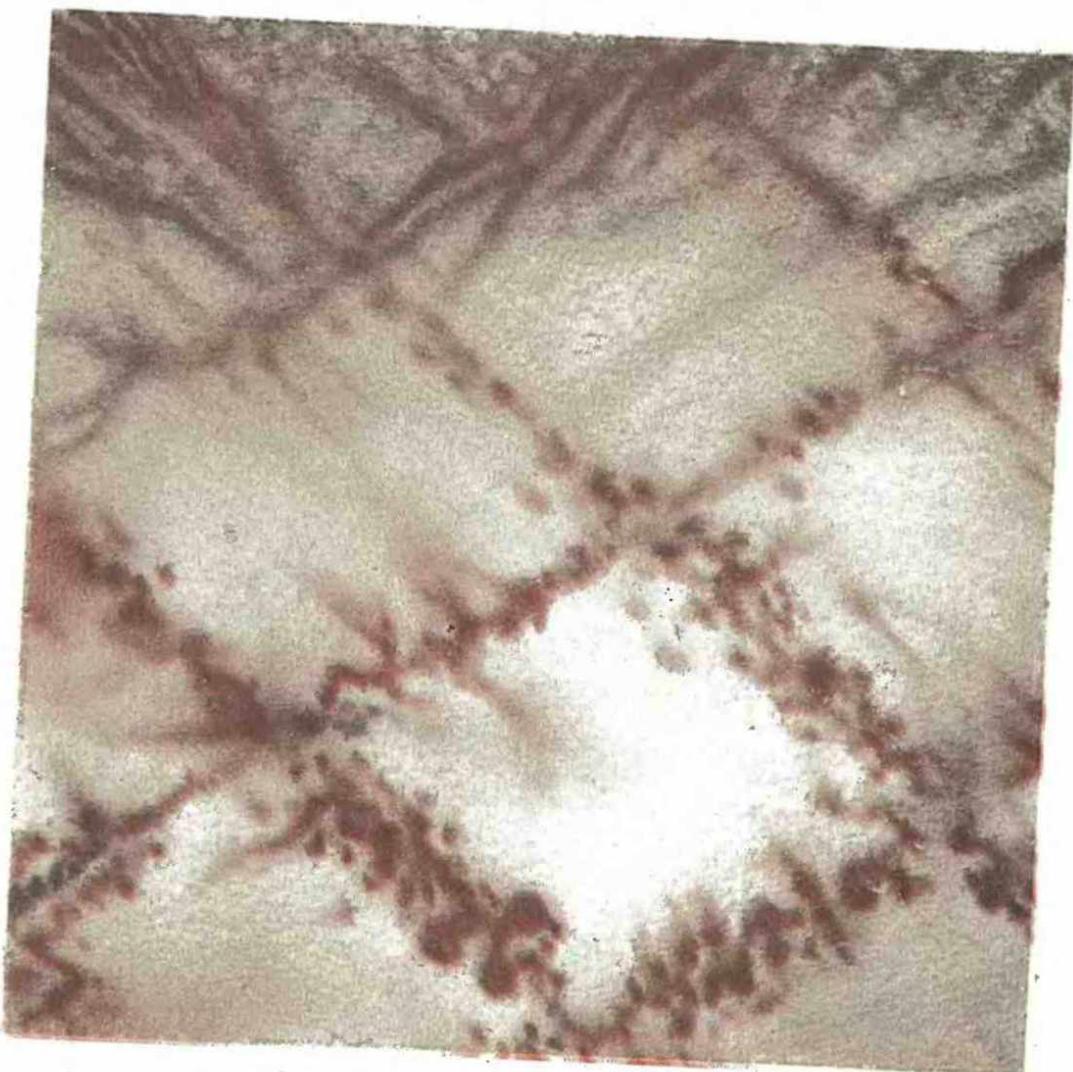
- 8.3.** Двоекропеломляющая призма (рис. 242) состоит из стеклянной призмы с $n = 1,66$ и призмы из исландского шпата с $n_o = 1,66$ и $n_e = 1,49$. Угол $\alpha = 30^\circ$. Найти угол между лучами на выходе из призмы, если в призму луч входит так, как изображено на рис. 242.

Ответы

- 8.1.** $I = I_0 [\cos^2(\alpha - \beta) - \sin 2\alpha \sin 2\beta \sin^2 \delta/2]$; $\delta = k(n_e - n_0)d$. **8.2.** Круговая поляризация. **8.3.** $5^\circ 45'$.

Основная идея:
атомы и молекулы
вещества
и их совокупности
под воздействием
падающего на них
излучения становятся
источниками
вторичного излучения.

Рассеяние света



§ 46 | Природа процессов рассеяния

Объясняется сущность процессов рассеяния света и лается их классификация.

Природа рассеяния. В состав среды входят молекулы или атомы основного вещества, составляющего среду, и посторонние частицы (пылинки, водяные капли и т. д.). Молекулы имеют размеры порядка 0,1 нм, а посторонние частицы, состоящие из агрегатов молекул, — в тысячи и десятки тысяч раз больше.

Процесс рассеяния света состоит в заимствовании молекулой или частицей энергии у распространяющейся в среде электромагнитной волны и излучении этой энергии в телесный угол, вершиной которого является молекула или частица. В этом смысле рассеяние света молекулой и частицей из громадного числа молекул осуществляется одинаково и различие состоит лишь в механизмах переизлучения.

Если среда рассматривается как непрерывная, то источником рассеяния выступают оптические неоднородности среды. В этом случае среда феноменологически характеризуется изменяющимся показателем преломления, а «размеры» областей, на которых происходит рассеяние, определяются расстояниями, на которых происходит значительное изменение показателя преломления. По своему физическому содержанию рассеяние является дифракцией волны на неоднородностях среды.

Типы рассеяния. Характер рассеяния в первую очередь зависит от соотношения между длиной волны и размером частиц. Если линейные размеры частицы меньше, чем примерно $1/\lambda$, длины волны, то рассеяние называется рэлеевским по имени Д. У. Рэлея (1842—1919), изучившего этот вид рассеяния. При больших размерах частиц принято говорить о рассеянии Ми. Хотя первоначально развитая Г. А. Ми (1908) теория относилась только к сферическим частицам, термин «рассеяние Ми» используется и для частиц неправильной формы. Для малых частиц теория Ми приводит к результатам теории Рэлея. Важным частным случаем оптической неоднородности является неоднородность оптических свойств среды, в которой распространяется звуковая волна. В результате этого возникают гармоническое распределение оптической неоднородности среды в пространстве и гармоническое, изменение оптических свойств во времени. В результате пространственной гармонической неоднородности оптических свойств наблюдается дифракция света на волне (см. § 33). В результате гармонического изменения оптических свойств во времени в каждой точке среды наблюдается изменение частоты дифрагированного света. Это изменение частоты дифрагированного на звуковой волне света получило название явления Мандельштама — Бриллюэна. Оно было независимо открыто Л. И. Мандельштамом (1879—1944) и Л. Бриллюэном (1889—1969).

Квантовые свойства молекул проявляются в комбинационном рассеянии света, характеризующемся изменением частоты рассеянного света по сравнению с частотой падающего. Ввиду специфически квантовой природы этого рассеяния оно также выделяется в отдельный тип.

Многократное рассеяние. Рассеянное частицей излучение может быть в свою очередь рассеяно другой частицей и т. д. В этом случае говорят о многократном рассеянии. Оно в каждом из последовательных актов осуществляется по законам однократного рассеяния. Окончательный результат получается суммированием результатов однократных рассеяний с учетом статистических характеристик их следования друг за другом.

- Процесс рассеяния сводится к генерации вторичных волн молекулами или частицами под действием падающего на них излучения. Для сплошной среды рассеяние сводится к дифракции волн на неоднородностях среды.

Рэлеевское рассеяние и рассеяние Ми

Описывается основные законы рассеяния Рэлея и Ми.

Модель элементарного рассеивателя. Электроны, попадающие в электрическое поле электромагнитной волны, совершают колебательное движение с частотой волны. Если волна распространяется в положительном направлении оси X (рис. 261), а электрический вектор колеблется в плоскости $Y=0$, то уравнение движения электрона в соответствии с (9.13) имеет вид

$$m\ddot{z} + m\omega_0^2 z = eE_0 \cos \omega t, \quad (47.1)$$

где $E_0 \cos \omega t$ — колебания напряженности электрического поля волны, коллинеарной оси Z в плоскости $Y=0$; m и e — масса и заряд электрона; ω_0 — собственная частота колебаний электрона, определяемая упругой силой, удерживающей электрон в положении равновесия. Затухание колебаний электрона за счет излучения не учитывается, поскольку оно мало.

Из (47.1) для отклонения электрона от положения равновесия находим

$$z = \frac{e}{m} \frac{E_0}{\omega_0^2 - \omega^2} \cos \omega t. \quad (47.2)$$

Колеблющийся электрон сам является излучателем. Его излучение рассеянное. Таким образом, моделью элементарного классического рассеивателя света является элементарный классический излучатель — электрический диполь, находящийся в поле электромагнитной волны.

Электрон входит в состав атома, являющегося электрически нейтральной системой. Поэтому можно считать, что колебания электрона в соответствии с (47.2) происходят около точки равновесия, в которой находится положительный заряд $|e|$. Этот заряд можно считать неподвижным, поскольку масса несущего его протона (или ядра) много больше массы электрона. Следовательно, (47.2) может быть записано в виде формулы для изменяющегося по времени дипольного момента:

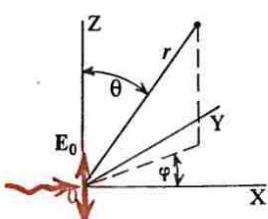
$$p_z(t) = ez = \frac{e^2}{m} \frac{1}{\omega_0^2 - \omega^2} E_0 \cos \omega t. \quad (47.3)$$

Поле излученной диполем электромагнитной волны в сферической системе координат с полярной осью, совпадающей с направлением диполя, описывается формулами (рис. 262)

$$E_\theta = cB_\phi = -\frac{1}{4\pi\epsilon_0} \frac{\omega^2}{c^2} \frac{\sin \theta}{r} p(t - r/c),$$

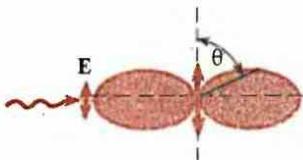
$$E_r = E_\phi = 0, \quad B_r = B_\phi = 0, \quad (47.4)$$

где θ, ϕ — соответственно полярный и аксиальный углы; r — расстояние от диполя до точки, в которой определяется поле (см. рис. 261). Подчеркнем, что по линии колебаний диполя излучение отсутствует.



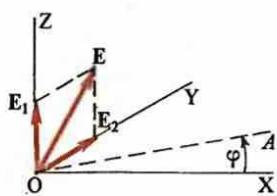
261

Элементарный рассеиватель света



262

Угловое распределение интенсивности рассеянного излучения от поляризованной волны



263

К расчету рассеяния неполяризованного излучения

Плотность потока энергии в направлении, характеризуемом углами θ и φ , в соответствии с (3.1) равна

$$S(\theta, \varphi, r, t) = E_\theta H_\varphi = \frac{E_\theta B_\varphi}{\mu_0} = \frac{\omega^4 \sin^2 \theta p^2}{16\pi^2 \epsilon_0 c^2 r^2}. \quad (47.5)$$

Усредняя S по периоду и учитывая (47.3), находим

$$\langle S \rangle_t = \frac{\pi^2 c \sin^2 \theta}{2\epsilon_0 r^2 \lambda^4} \left[\frac{e^2}{m(\omega_0^2 - \omega^2)} \right]^2 E_0^2, \quad (47.6)$$

где $\lambda = 2\pi c/\omega$.

Поток энергии $dP(\theta, \varphi)$ в телесный угол $d\Omega = d\sigma/r^2$, опирающийся на элемент площади сферы $d\sigma$ равен

$$dP(\theta, \varphi) = \langle S \rangle_t d\sigma = \langle S \rangle_t r^2 d\Omega. \quad (47.7)$$

Отсюда для интенсивности рассеяния $I_1(\theta, \varphi)$, определяемой как поток энергии, отнесенной к телесному углу $d\Omega$, находим

$$I_1(\theta, \varphi) = \frac{dP}{d\Omega} = \frac{\pi^2 c \sin^2 \theta}{2\epsilon_0 \lambda^4} \left[\frac{e^2}{m(\omega_0^2 - \omega^2)} \right]^2 E_0^2. \quad (47.8)$$

Заметим, что (47.8) выражает плотность энергии рассеянного потока от одного элементарного излучателя.

С помощью соотношения (3.4) эту формулу целесообразнее представить в виде

$$I_1(\theta, \varphi) = \frac{\pi^2 \sin^2 \theta}{\lambda^4} \left[\frac{e^2}{m(\omega_0^2 - \omega^2)} \right]^2 \langle S_0 \rangle, \quad (47.9)$$

где S_0 — среднее значение плотности потока энергии в падающей волне.

Рэлеевское рассеяние. Если размеры рассеивателя много меньше длины волны, то все элементарные диполи излучают когерентно. Под рэлеевским рассеянием обычно понимается рассеяние молекулами среды, потому что размеры обычных молекул (не макромолекул) всегда много меньше длины волны видимого света. Элементарные рассеиватели, принадлежащие различным молекулам, излучают некогерентно, потому что, во-первых, расстояние между молекулами может быть достаточно большим и, во-вторых, вследствие движения молекул происходят флуктуации плотности среды. С учетом этих обстоятельств заключаем, что интенсивность рассеянной волны от одной молекулы увеличивается пропорционально квадрату числа N_0 элементарных рассеивателей в ней. Концентрацию молекул обозначим N . Следовательно, в единице объема находится $N_0 N$ элементарных диполей. Из курса электричества известно соотношение

$$\frac{(n^2 - 1) 3\epsilon_0}{N_0 N (n^2 + 2)} = \frac{e^2}{m(\omega_0^2 - \omega^2)}, \quad (47.10)$$

где n — показатель преломления среды. Заменяя выражение, стоящее в квадратных скобках (47.9), его значением (47.10), получаем для интенсивности рассеяния от одной молекулы следующую формулу:

$$I(\theta, \varphi) = I_1(\theta, \varphi) N_0^2 = \frac{9\pi^2 \sin^2 \theta}{N^2 \lambda^4} \left(\frac{n^2 - 1}{n^2 + 2} \right)^2 \langle S_0 \rangle. \quad (47.11)$$

Заметим, что полученные формулы справедливы для случая, когда собственная частота ω_0 колебаний электронов много больше частот видимого спектра и ближнего ультрафиолета. Это условие в большинстве случаев соблюдается. Далее, при выводе формулы (47.12) предполага-

галось, что собственные частоты колебаний всех электронов в молекуле одинаковы. Это ограничение нетрудно снять, поскольку оно не имеет принципиального характера и его снятие лишь делает формулы более громоздкими. Читателю рекомендуется провести соответствующие вычисления в качестве упражнения.

Полная интенсивность рассеяния одной частицей по всем направлениям получается из (47.11) интегрированием по всем углам рассеяния:

$$I_0 = \int_{4\pi} I(\theta, \phi) d\Omega = \frac{18\pi^3}{N^2\lambda^4} \left(\frac{n^2 - 1}{n^2 + 1} \right)^2 \langle S_0 \rangle \int_0^\pi \sin^3 \theta d\theta \quad (47.12)$$

$$I_0 = \frac{24\pi^3}{N^2\lambda^4} \left(\frac{n^2 - 1}{n^2 + 2} \right)^2 \langle S_0 \rangle. \quad (47.13)$$

Поскольку различные молекулы рассеивают некогерентно, полная интенсивность рассеяния в единице объема вещества вычисляется умножением выражения (47.13) на концентрацию N молекул. Для не очень плотных газов показатель преломления $n \approx 1$ и, следовательно, в формуле (47.13) можно принять $n^2 + 2 \approx 3$, $(n^2 - 1) \approx 2(n - 1)$. Для интенсивности рассеяния в единице объема в этом случае получаем

$$I_V = I_0 N = \frac{32\pi^3 (n - 1)^2}{3N\lambda^4} \langle S_0 \rangle. \quad (47.14)$$

Закон Рэлея. Из (47.14) видно, что интенсивность рассеяния обратно пропорциональна четвертой степени длины волны. Такая зависимость рассеяния от длины волны называется законом Рэлея.

Законом Рэлея объясняется, например, голубой цвет неба и красноватый цвет Солнца на восходе и заходе. На восходе и заходе наблюдается свет, в котором в результате рассеяния по закону Рэлея коротковолновая часть спектра (фиолетовая) ослаблена значительно сильнее длинноволновой (красной) части. В результате интенсивность длинноволновой (красной) части спектра относительно возрастает и воспринимается глазом как красноватый цвет Солнца. Относительное изменение интенсивности различных частей спектра будет заметным лишь при достаточно большом рассеянии. Поэтому Солнце в зените, когда проходящая лучами толща атмосферы не очень велика и рассеяние света незначительно, не имеет красного цвета. Однако и в этом случае рассеяние и поглощение существенно изменяют спектральный состав излучения, достигающего поверхности Земли (см. § 1).

При наблюдении небосвода днем в глаз попадает рассеянное излучение, в котором более сильно присутствует коротковолновая часть спектра, соответствующая голубому цвету. Вне земной атмосферы небо представляется черным, а в глаз попадают лишь прямые лучи от звезд.

Угловое распределение и поляризация света при рэлеевском рассеянии. Угловое распределение рассеяния поляризованного излучения от отдельной молекулы описывается формулой (47.11). Оно аксиально-симметрично относительно линии, проходящей через элементарный рассеиватель в направлении колебаний электрического вектора падающей волны (рис. 262). Перпендикулярно направлению распространения падающей волны вдоль линии колебаний \mathbf{E} рассеяние отсутствует. Максимальное рассеяние наблюдается в плоскости, перпендикулярной направлению колебаний электрического вектора падающей волны. Рассеянное излучение поляризовано — электрический вектор колеблется в плоскости, проходящей через линию колебаний электрона элементарного рассеивателя. Если рассеяние от различных молекул можно считать некогерентным друг с другом, то полная интенсивность рассеяния в единице объема вычисляется умножением выражения (47.11) на концентрацию N молекул. Следовательно, свойства излучения, рассеянного от отдельной молекулы, полностью сохраняются для излучения, рассеянного в объеме.

Для расчета углового распределения рассеяния неполяризованной волны представим падающую волну в виде суперпозиции двух волн с взаимно перпендикулярными направлениями поляризации (см. § 4):

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2, \quad (47.15)$$

где $\mathbf{E}_1 \perp \mathbf{E}_2$. Отсюда находим соотношение между плотностями потока энергии этих волн:

$$\langle S_0 \rangle = \langle S_{01} \rangle + \langle S_{02} \rangle. \quad (47.16)$$

Ясно, что рассеяние неполяризованного света должно быть аксиально-симметрично относительно направления луча. Поэтому достаточно найти интенсивность рассеянного света в плоскости XY (рис. 261) в направлении, составляющем угол ϕ с осью X (рис. 263). Направим вектор \mathbf{E}_1 по оси Z, а вектор \mathbf{E}_2 — по оси Y. Тогда [см. (47.11)]

$$I = \frac{9\pi^2}{N^2\lambda^4} \left(\frac{n^2 - 1}{n^2 + 2} \right)^2 [\langle S_{01} \rangle \sin^2(\pi/2) + \langle S_{02} \rangle \cos^2(\phi)], \quad (47.17)$$

поскольку вклады в полную интенсивность от волн с взаимно перпендикулярными направлениями поляризации независимы. У неполяризованного света вектор E в плоскости YZ равновероятно распределен по всем углам. Поэтому, усредняя E^2 по углам, из (47.16) находим

$$\langle S_0 \rangle = \langle S_{01} \rangle + \langle S_{02} \rangle, \quad (47.18)$$

где внешними угловыми скобками обозначено усреднение по направлениям E. Очевидно,

$$\langle S_{02} \rangle = \langle S_{01} \rangle = \langle S_0 \rangle / 2 \quad (47.19)$$

и, следовательно, формула (47.17) принимает вид

$$I(\phi) = \frac{9\pi^2}{2N^2\lambda^4} \left(\frac{n^2 - 1}{n^2 + 2} \right)^2 (1 + \cos^2 \phi) \langle S_0 \rangle. \quad (47.20)$$

Зависимость $I(\phi)$ показана на рис. 264. Картина рассеяния аксиально-симметрична относительно направления распространения падающей волны. Рассеяния вперед и назад одинаково интенсивны и распределены симметрично относительно центра рассеяния.

В формуле (47.17) слагаемое с $\langle S_{01} \rangle$ описывает рассеянную линейно поляризованную волну, электрический вектор которой колеблется коллинеарно оси Z, а слагаемое с $\langle S_{02} \rangle$ — линейно поляризованную волну, электрический вектор которой колеблется коллинеарно оси Y. Чтобы освободиться в описании рассеяния от координатной системы, назовем плоскостью наблюдения плоскость, проходящую через падающий луч и точку наблюдения. Можно сказать, что слагаемое с $\langle S_{01} \rangle$ в (47.17) описывает рассеянную волну, электрический вектор которой колеблется перпендикулярно плоскости наблюдения, а с $\langle S_{02} \rangle$ — волну с электрическим вектором, колеблющимся в плоскости наблюдения. Рассеяние волны с направлением электрического вектора, перпендикулярного плоскости наблюдения, описывается в (47.20) слагаемым с единицей в последних круглых скобках, а параллельно плоскости наблюдения — слагаемым с $\cos^2 \phi$. Таким образом, при рассеянии неполяризованного света наблюдается частично поляризованное рассеянное излучение, степень поляризации которого зависит от угла ϕ . Степень поляризации определяется соотношением

$$P = \frac{I_L(\phi) - I_\parallel(\phi)}{I_L(\phi) + I_\parallel(\phi)} = \frac{\sin^2 \phi}{1 + \cos^2 \phi}, \quad (47.21)$$

т. е. лишь при углах $\phi = 0$ и $\phi = \pi$ обе компоненты поляризации присутствуют с одинаковой интенсивностью. При других углах ϕ более интенсивно присутствует рассеяние, в котором электрический вектор колеблется перпендикулярно плоскости наблюдения.

Из (47.21) следует, что в направлении, перпендикулярном падающему лучу, свет полностью линейно поляризован ($P = 1$). В частности, если в земной атмосфере наблюдается только рэлеевское рассеяние, то рассеянный атмосферой солнечный свет распространяющийся в плоскости, перпендикулярной направлению Земля—Солнце, должен быть полностью поляризованным. Практически полная степень поляризации никогда не наблюдается, поскольку одновременно с рэлеевским рассеянием солнечного света наблюдаются рэлеевское рассеяние света, диффузно отраженного от поверхности Земли, и рассеяние Ми на аэрозольных частицах в воздухе. Некоторую роль в деполяризации наблюдаемого света играет также слабая оптическая анизотропия молекул, составляющих воздух.

Ослабление интенсивности света. В результате рассеяния плотность потока энергии распространяющегося в среде света ослабляется. Если пучок света имеет поперечное сечение σ , то на длине пути dx он встречает $N\sigma dx$ рассеивателей, каждый из которых выводит из пучка мощность (47.13). Следовательно, мощность потока энергии в пучке уменьшается:

$$d < S_0 > \sigma = -I_0 N \sigma dx. \quad (47.22)$$

С учетом (47.13) равенство (47.22) принимает вид

$$d < S_0 > = -\gamma < S_0 > dx, \quad (47.23)$$

где коэффициент рассеяния

$$\gamma = \frac{24\pi^3}{N\lambda^4} \left(\frac{n^2 - 1}{n^2 + 1} \right)^2. \quad (47.24)$$

Поэтому закон ослабления плотности потока энергии в световом пучке из-за рассеяния выражается формулой

$$< S_0(x) > = < S_0(0) > e^{-\gamma x}. \quad (47.25)$$

Рассеяние Ми. Теория Рэлея хорошо описывает не только рассеяние на молекулах, но и на достаточно малых сферических частицах, радиус которых меньше примерно $0,03\lambda$. При увеличении размеров частиц становятся заметными отклонения от предсказаний теории Рэлея и необходимо пользоваться теорией Ми. Теория рассеяния Ми учитывает размеры частиц и выражает рассеяние в виде рядов, малым параметром в которых служит

$$a = ka = 2\pi a/\lambda, \quad (47.26)$$

где a — радиус сферической частицы. Теория рассеяния Ми относится собственно только к сферическим частицам. Однако термин «рассеяние Ми» употребляется также и для рассеяния на частицах других форм. Как видно из (47.26), имеет значение не абсолютный размер частиц, а соотношение размера частицы и длины волны. Закономерности рассеяния ультрафиолетовой части спектра на частицах определенного радиуса таковы же, как и рассеяния красной части

- Закономерности рассеяния зависят от размеров частиц, на которых происходит рассеяние света. При линейных размерах частиц, меньших примерно $1/\lambda$ длины волны, наблюдается рэлеевское рассеяние, при больших размерах — рассеяние Ми. Интенсивность рэлеевского рассеяния обратно пропорциональна четвертой степени длины волны. Интенсивность рассеяния Ми слабо зависит от длины волны, а для частиц линейные размеры которых много больше длины волны, практически не зависят от длины волны.

спектра на частотах, примерно в два раза больших. При $a < 0,2$ рассеяние становится рэлеевским, т. е. рэлеевское рассеяние является предельным случаем рассеяния Ми.

Распределение интенсивности по углам и поляризация излучения в рассеянии Ми. В больших частицах имеется много молекул. Элементарные диполи каждой молекулы под влиянием электромагнитной волны приходят в колебания и становятся источниками вторичных волн, составляющих рассеянное излучение. В этом отношении механизм рассеяния Ми аналогичен механизму рассеяния Рэлея. Различия обусловливаются лишь двумя обстоятельствами.

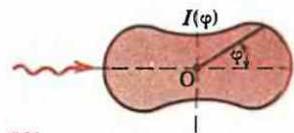
1. При рассеянии Рэлея все элементарные рассеиватели находятся в поле одной и той же волны и излучают когерентно. При рассеянии Ми необходимо учесть влияние переизлучения первичной волны элементарными рассеивателями, в результате чего элементарные рассеиватели находятся, вообще говоря, не в одинаковых электромагнитных полях. Другими словами, необходимо принять во внимание, что коэффициент прецессии в объеме частицы не равен единице.

2. В рассеянии Рэлея излучение от элементарных рассеивателей одной и той же частицы (молекулы) интерферирует между собой при одинаковой разности фаз независимо от направления. В рассеянии Ми необходимо учитывать различие в фазах излучения элементарных рассеивателей и разность фаз, вносимую в наблюдаемое излучение конечным расстоянием между элементарными рассеивателями. Последнее обстоятельство приводит к существенной зависимости распределения интенсивности излучения от направления, выражающей зависимость условий интерференции излучения элементарных рассеивателей от их взаимного расположения относительно точки наблюдения.

В математическом смысле теория Ми сводится к решению уравнений Максвелла с граничными условиями на поверхности сферической частицы произвольного радиуса, характеризующейся диэлектрической и магнитной проницаемостями и электропроводимостью. Решение получается в виде рядов, которые дают полную информацию о рассеянии. В целом получается довольно громоздкая и сложная теория, излагать которую в данной книге нет необходимости. Укажем лишь на некоторые важные результаты.

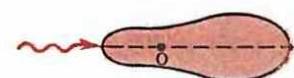
С увеличением размера частиц (точнее a/λ) появляется асимметрия рассеяния вперед и назад — превалирует рассеяние вперед (рис. 265), однако (при $a \approx \lambda/4$) без резких максимумов и минимумов. При дальнейшем увеличении размеров частиц ($a > \lambda$) наблюдается преимущественное рассеяние вперед со многими вторичными максимумами, распределение которых зависит от размеров частиц (рис. 266).

Рассеянный свет частично поляризован даже при неполяризованном, падающем на частицы излучении, как и при рэлеевском рассеянии. Характер поляризации зависит от оптических свойств частиц и направления, в котором наблюдается рассеян-



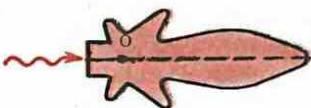
264

Угловое распределение интенсивности рассеяния неполяризованного света



265

Рассеяние Ми при $a \approx \lambda/4$



266

Рассеяние Ми при $a > \lambda$

- Опишите угловое распределение и поляризацию света при рэлеевском рассеянии. В каком направлении в рэлеевском рассеянии наблюдается полностью линейно поляризованный свет? В чем заключаются физические причины этого явления?

Опишите распределение интенсивности по углам и поляризацию света в рассеянии Ми.

Почему при достаточно большом размере частиц, на которых происходит рассеяние, рассеяние Ми очень слабо зависит от длины волны света?

ный свет. Если падающий свет поляризован, то поляризация рассеянного света зависит также и от его поляризации.

Важной особенностью рассеяния Ми является его слабая зависимость от длины волны для частиц, линейные размеры которых много больше длины волны, что существенно отличается от рассеяния Рэлея. Благодаря этому, например, облака являются белыми, а небо — голубым.

Проявления рассеяния Ми. Многообразие проявлений рассеяния Ми обусловливается многообразием частиц, на которых оно осуществляется. Небо, голубое в зените, постепенно сереет к горизонту. При задымлении атмосферы небо приобретает белесый оттенок. При полете в самолете на большой высоте четкая линия горизонта обычно не видна. Она застилается атмосферной дымкой. Все эти явления обусловлены рассеянием Ми на аэрозолях воздуха. Малая или почти полная непрозрачность тумана является следствием сильного рассеяния Ми малыми каплями воды. Сильное ослабление света от Солнца при заходе и восходе в значительной степени обусловлено также рассеянием Ми.

§ 48 | Рассеяние Мандельштама — Бриллюэна

Описывается основные закономерности рассеяния
Мандельштама — Бриллюэна.

Компоненты Мандельштама — Бриллюэна. При дифракции на звуковой волне возникают лишь два максимума первого порядка, описываемые формулой (33.64). Амплитуда дифрагированной волны изменяется вместе с коэффициентом пропускания и коэффициентом преломления среды, обусловленным изменением плотности среды в волне. Следовательно, амплитуда изменяется гармонически с частотой Ω звуковой волны. Поэтому наблюдаемая в направлении дифракционных максимумов напряженность электромагнитной волны описывается формулой

$$E(t) = A_0 \cos \Omega t \cos \omega t = (A_0/2)[\cos(\omega + \Omega)t + \cos(\omega - \Omega)t], \quad (48.1)$$

где ω — частота падающего света. Таким образом, в рассеянном свете должны наблюдаться две сателлитные частоты, расположенные симметрично относительно основной частоты (рис. 267). Сателлит с частотой $\omega - \Omega$ называется стоксовым, а с $\omega + \Omega$ — антистоксовым. Они являются компонентами рассеяния Мандельштама — Бриллюэна.

Поскольку оптическая длина пути в среде в n раз больше геометрической, где n — показатель преломления среды, формула (33.64) с учетом изменения скорости распространения света в среде представляется равенствами

$$nd \sin \phi_{(+)} = \lambda, \quad nd \sin \phi_{(-)} = -\lambda. \quad (48.2)$$

Частота звуковой волны может быть выражена в виде

$$\Omega = 2\pi v/d = 2\pi n \sin \phi / \lambda = 2\omega n(v/c) \sin(\phi/2), \quad (48.3)$$

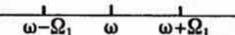
где v — скорость распространения звуковых волн в среде, $\sin \phi \approx 2 \sin(\phi/2)$ ввиду малости угла ϕ . Формула (48.3) называется формулой Мандельштама — Бриллюэна.



- Как образуются компоненты Мандельштама — Бриллюэна?
- Какие физические причины обуславливают присутствие несмещенной частоты в рассеянии?
- Каковы основные особенности явления Мандельштама — Бриллюэна в твердых телах?

Несмешенная компонента. В жидкостях в большинстве случаев наряду с частотами $\omega \pm \Omega$ наблюдается также и частота ω . Формально ее появление можно объяснить, если в (48.1) вместо $\cos \Omega t$ стоит $a + \cos \Omega t$, где a — приблизительно постоянная. Другими словами, наличие несмешенной частоты ω в дифрагировании свете обусловливается оптической характеристикой среды, которая не изменяется во времени по гармоническому закону, а является примерно постоянной. Такая постоянная составляющая оптической неоднородности возникает за счет флуктуаций в среде, которые выравниваются за короткие по сравнению с периодом звуковой волны промежутки времени, в частности флуктуаций энтропии, которые выравниваются посредством теплопроводности.

Явление Мандельштама — Бриллюэна в твердых телах. В аморфных твердых телах возможны как поперечные, так и продольные волны, распространяющиеся с различными скоростями. Каждая из волн приводит в рассеянном свете к возникновению двух сателлитов. Поэтому всего в рассеянном излучении наблюдается пять компонент, включая несмешенную. В кристаллических твердых телах число компонент увеличивается в соответствии с числом волн, распространяющихся с различными скоростями и различными направлениями колебаний, и числом электромагнитных волн, которые могут распространяться в кристалле в данном направлении. Расчет показывает, что в общем случае в кристалле возникают 24 смешанные компоненты.



§ 49 | Комбинационное рассеяние

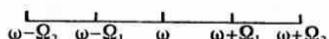
Описываются основные закономерности комбинационного рассеяния.

Классическая интерпретация. Допустим, что оптические свойства молекулы изменяются по гармоническому закону, в результате чего амплитуда рассеиваемого молекулой света также изменяется по гармоническому закону. Наблюдаемая напряженность электрического поля рассеянного света аналогично (48.1) равна

$$E(t) = A_0(1 + a \cos \Omega t) \cos \omega t, \quad (49.1)$$

где Ω — частота, характеризующая изменение оптических свойств молекулы, ω — частота падающего на молекулу света. Коэффициент a учитывает эффективность модуляции амплитуды падающего света молекулой.

Из (49.1) видно, что в рассеянном излучении присутствуют волны с частотами ω , $\omega + \Omega$, $\omega - \Omega$. Наличие смешенных частот в рассеянном молекулой излучении называется комбинационным рассеянием. Оно было открыто в 1928 г. Ч. В. Раманом, Г. С. Ландсбергом и Л. И. Мандельштамом. Каждая из спектральных линий первичного излучения в рассеянном излу-



чении сопровождается целой системой спутников, частоты которых отстоят от центральной частоты на величины, характерные для молекулы (рис. 268). Можно сказать, что молекулы обладают набором собственных частот колебаний ее оптических свойств $\Omega_1, \Omega_2, \dots$, которые в спектре рассеяния проявляются в соответствии с формулой (49.1). Частоты рассеянного света комбинируются из частоты падающего света и собственных частот колебаний молекулы.

Экспериментальные факты. Опыт показывает, что спутники сопровождают каждую линию падающего излучения, а частоты $\Omega_1, \Omega_2, \dots$ одинаковы для всех линий и характеризуют свойства молекулы. Система спутников симметрична относительно частоты падающего излучения.

Спутники со стороны больших частот называются фиолетовыми или антистоксовыми, а со стороны меньших — красными или стоксовыми. Опыт показывает, что ближайшие к центральной частоте стоксы спутники значительно интенсивнее, чем антистоксы, однако с повышением температуры это различие уменьшается, поскольку интенсивность антистоксовых спутников значительно растет.

Спектры излучения молекул называют полосатыми, потому что они имеют вид полос, состоящих из близко расположенных линий. Такой вид спектра обусловливается размыванием линейчатого электронного спектра излучения молекулы за счет энергетических переходов молекулы между колебательными и вращательными уровнями энергий. Энергетическое расстояние между колебательными уровнями значительно больше, чем между вращательными. Поэтому полоса в спектре образуется как бы в два этапа — на определенных расстояниях от частоты излучения в результате электронного перехода образуются линии колебательного спектра, а около каждой линии колебательного спектра образуются очень близко расположенные линии за счет вращательных переходов. Изучение спектров излучения молекул и их комбинационных спектров рассеяния показало, что комбинационные частоты $\Omega_1, \Omega_2, \dots$ всегда совпадают с соответствующими разностями частот колебательного спектра молекул или, другими словами, комбинационные частоты совпадают с собственными частотами колебаний молекул. Однако не всем собственным частотам колебаний молекул удается сопоставить комбинационную частоту в спектре комбинационного рассеяния и, кроме того, нет простой связи между интенсивностью линии поглощения в спектре молекулы и соответствующей линии комбинационного рассеяния.

Классическая интерпретация (49.1) комбинационного рассеяния позволяет понять смысл комбинационных частот, но не в состоянии объяснить многие количественные закономерности. В частности, непонятно, почему интенсивности стоксовых и антистоксовых компонент различны. Комбинационное рассеяние является квантовым по своей природе и может быть полностью описано лишь квантовой теорией.

Квантовая интерпретация. Энергетические уровни молекулы дискретны. Комбинационное рассеяние объясняется переходами молекул между колебательными уровнями. Молекула поглощает квант падающего излучения с энергией $\epsilon = \hbar\omega$. Часть энергии $\epsilon_k = \hbar\Omega$ она поглощает и переходит на более высокий колебательный уровень. Оставшаяся энергия испускается в виде кванта излучения с энергией $\epsilon = \epsilon - \epsilon_k = \hbar(\omega - \Omega)$, т. е. частоты $\omega_1 = \omega - \Omega$. Так образуется стоксова компонента. Антистоксова компонента образуется, если колебательные уровни мо-

- Какими физическими факторами обусловливается существование полосатых спектров молекул?
- Почему ближайшие к центральной частоте стоксы спутники значительно интенсивнее антистоксовых и почему с повышением температуры это различие уменьшается?
- Почему не всем собственным частотам колебаний молекул удается сопоставить комбинационную частоту в спектре комбинационного рассеяния?

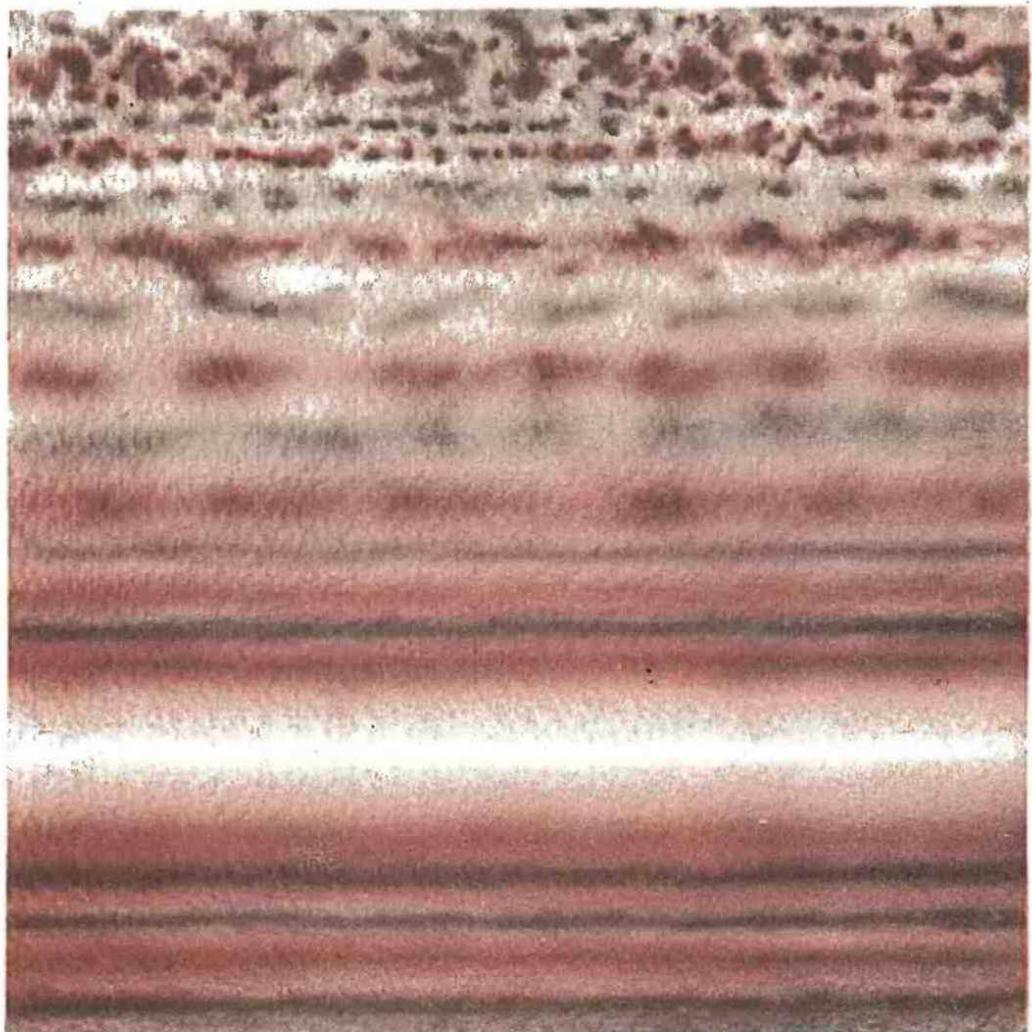
лекул достаточно сильно возбуждены при соответствующей температуре. Квант излучения $\varepsilon = \hbar\omega$ поглощается молекулой, находящейся на возбужденном колебательном уровне. После поглощения кванта излучения молекула переходит на более низкий колебательный уровень и освобождающуюся при этом энергию $\varepsilon_k = \hbar\Omega$ присоединяет к энергии поглощенного кванта, испуская квант с энергией $\varepsilon_2 = \varepsilon + \varepsilon_k = \hbar(\omega + \Omega)$, т. е. частоты $\omega_2 = \omega + \Omega$. Так образуется антистоксова компонента. Ясно, что при не очень высоких температурах число молекул, которые могут принять участие в испускании стоксовых компонент, значительно больше числа молекул, способных принять участие в испускании антистоксовых компонент. Этим объясняется, что интенсивность стоксовых компонент больше антистоксовых. **Различие интенсивностей стоксовых и антистоксовых компонент уменьшается с увеличением температуры**, потому что относительно увеличивается число возбужденных молекул, способных участвовать в испускании антистоксовых компонент.

Интенсивность линии комбинационного рассеяния обусловливается легкостью поляризуемости молекулы на соответствующей комбинационной частоте, а интенсивность линии поглощения определяется легкостью возбуждения колебаний молекулы излучением соответствующей частоты. Эти два физических фактора различны по своей природе, и поэтому между интенсивностями линий поглощения и соответствующих линий комбинационного рассеяния нет прямой связи.

Применение комбинационного рассеяния. Комбинационное рассеяние дает прямой метод исследования строения молекул, позволяя измерять частоты их собственных колебаний, изучать симметрию молекул, внутримолекулярные силы, молекулярную динамику и т. д. Спектры комбинационного рассеяния настолько характерны для молекулы, что с их помощью можно проводить анализ строения сложных молекулярных смесей, когда химические методы анализа не дают желаемых результатов.

Основная идея:
инверсная заселенность
уровней энергии
создает возможность
усиления и генерации
световых потоков
с большими временами
когерентности
и большой мощности.

Генерация света



Излагается классическая и элементарная квантовая теория излучения абсолютно черного тела. Анализируются свойства индуцированного излучения.

Плотность излучения. Повседневными генераторами света являются разогретые до достаточно высоких температур материальные среды, которые при этом могут быть в твердом, жидким или газообразном состоянии. Будем их все называть материальными телами или просто телами, в отличие от излучения, которое является также материальным объектом, но не телом. Все тела излучают, поглощают и отражают электромагнитные волны, однако не одинаково, и интенсивность этих процессов зависит от свойств тел, их температуры и частоты электромагнитных волн. Электромагнитное излучение в вакууме характеризуется **полной объемной плотностью излучения**

$$w = \frac{1}{2} (\epsilon_0 E^2 + B^2 / \mu_0), \quad (50.1)$$

которая определенным образом распределена по спектру частот. Распределение излучения по частотам описывается **спектральной плотностью излучения**

$$w_\omega = dw/d\omega. \quad (50.2)$$

Из (50.2) следует, что

$$w = \int_0^\infty w_\omega d\omega. \quad (50.3)$$

Равновесная плотность излучения. Предположим, что некоторый объем пространства с находящимися в нем материальными телами окружен замкнутой адиабатной оболочкой. По истечении достаточно большого промежутка времени между материальными телами в полости, замкнутой оболочкой, и излучением в полости установится термодинамическое равновесие. Все тела будут иметь одинаковую температуру T , а излучение в полости — определенную спектральную плотность излучения, называемую **равновесной**.

Первый закон Кирхгофа. Равновесная спектральная плотность w_ω зависит только от температуры T и не зависит от свойств и природы тел, находящихся в полости, и от свойств и природы стенок полости.

Справедливость этого утверждения доказывается с помощью второго начала термодинамики. Допустим, что w_ω зависит от свойств и природы полости и материальных тел, находящихся в ней. Возьмем две различные полости с одинаковой температурой. По предположению, w_ω в них различны. Соединим полости в одну. Ввиду различия w_ω между ними должен начаться обмен энергией излучения. Одна из полостей начнет нагреваться, другая — охлаждаться. Возникает разность температур, которую можно использовать для получения работы. При совершении работы разность температур между полостями будет ликвидирована и система придет в состояние термодинамического равновесия при более низкой, чем в исходном состоянии, температуре, поскольку часть энергии системы была затрачена на работу. Таким образом, производится работа за счет охлаждения адиабатно изолированной системы, что противоречит второму началу термодинамики. Тем самым первый закон Кирхгофа доказан и можно записать $w_\omega = w_\omega(T)$.

Поглощающая способность и энергетическая светимость. Принимается, что излучение с поверхности подчиняется закону Ламберта и поэтому спектральные плотности энергетической светимости и энергетической яркости связаны соотношением (7.9а). **Поглощающая способность** A_ω определяется как отношение энергии, поглощаемой участком поверхности тела в секунду в интервале частот $(\omega, \omega + d\omega)$, ко всей энергии излучения, падающей в секунду на этот

участок в том же интервале частот, причем предполагается, что излучение падает на поверхность изотропно.

Второй закон Кирхгофа. В состоянии равновесия поглощаемая в секунду участком поверхности энергия излучения должна быть равна энергии, излучаемой в тот же промежуток времени тем же участком поверхности. Это условие можно записать в виде формулы

$$\frac{M_\omega}{A_\omega} = \frac{c}{4} w_\omega(T), \quad (50.5)$$

в которой множитель $c/4$ учитывает связь плотности энергии с плотностью потоков энергии при их изотропном распределении.

Значение этого множителя получается из следующих соображений. Если бы движение энергии осуществлялось в одном измерении и плотность потока энергии распределялась одинаково между двумя направлениями, то плотность потока энергии в одном направлении была бы равна $sw_\omega/2$. Однако в трехмерном пространстве при изотропном распределении плотностей потоков энергии поток в заданном направлении образуется в результате сложения проекций плотностей потоков во всех направлениях на данное, причем необходимо учитывать только потоки с положительной проекцией. Поэтому плотность потока энергии в направлении, например оси Z , равна $\langle v_z \rangle w_\omega/2$, где $\langle v_z \rangle$ — среднее значение положительной проекции скорости потока на ось Z . Обозначая через θ , ϕ полярный и аксиальный углы в сферической системе координат, находим

$$\langle v_z \rangle = \frac{c}{2\pi} \int_0^{2\pi} d\phi \int_0^{\pi/2} \cos \theta \sin \theta d\theta = c/2,$$

где учтено, что опирающийся на полусферу телесный угол равен 2π . Отсюда получаем для плотности потока энергии в заданном направлении выражение $sw_\omega/4$, совпадающее с первой частью (50.5).

Направим ось Z сферической системы координат по нормали к излучающей поверхности, обозначив θ и ϕ полярный и аксиальный углы. Плотность потока энергии в телесный угол $d\Omega$ ввиду изотропного распределения излучения равна $[sw_\omega/(4\pi)]d\Omega$, а спектральная плотность мощности излучения dP_ω через элемент поверхности $d\sigma'$, перпендикулярный направлению движения потока энергии, дается формулой $dP_\omega = [sw_\omega/(4\pi)]d\Omega d\sigma'$. Эта мощность излучения испускается с площади $d\sigma$ светящейся поверхности, причем $d\sigma' = d\sigma \cdot \cos \theta$ (см. рис. 24). Отсюда для спектральной плотности энергетической светимости получаем

$$M_\omega = \frac{dP_\omega}{d\sigma} = \int_{\Omega=2\pi} \frac{sw_\omega}{4\pi} \cos \theta d\Omega = \frac{sw_\omega}{4\pi} \int_0^{2\pi} d\phi \int_0^{\pi/2} \cos \theta \sin \theta d\theta = \frac{c}{4} w_\omega,$$

как это и предусмотрено в (50.5). **Формула (50.5) выражает второй закон Кирхгофа.** Зная универсальную функцию $w_\omega(T)$, можно по поглощательной способности определить энергетическую светимость.

Абсолютно черное тело. Абсолютно черным называется тело, которое полностью поглощает все падающее на него излучение, т. е. у которого $A_\omega = 1$. Для абсолютно черного тела (50.5) принимает вид

$$M_\omega(T) = (c/4) w_\omega(T), \quad (50.6)$$

т. е. задача нахождения $w_\omega(T)$ сводится к определению закона излучения абсолютно черного тела.

Реализовать абсолютно черное тело можно в виде полости с небольшим отверстием

(рис. 269). Лучи, попадающие через отверстие внутрь полости, в результате многократных отражений с поглощением на внутренних стенах оболочки полости практически полностью поглощаются и не выходят через отверстие наружу. Это обстоятельство проявляется наглядно, например, при взгляде на открытые окна в доме, которые в светлый день кажутся темными. Излучение, исходящее из отверстия в оболочке полости, с достаточно большой точностью может рассматриваться как излучение абсолютно черного тела. Его излучение позволяет найти M_ω и с помощью (50.6) вычислить $w_\omega(T)$.

Классическая физика оказалась не в состоянии объяснить теоретически вид функции $w_\omega(T)$, измеренной экспериментально. Пределные случаи $w_\omega(T)$ при достаточно малых и достаточно больших частотах были теоретически обоснованы формулами Рэлея — Джинса и Вина. Общая формула как интерполяционная формула для предельных случаев была найдена Планком. Она положила начало развитию квантовой теории.

Концентрация мод колебаний. Будем считать, что полость имеет форму куба с ребром L (рис. 270). Стоячая волна может образоваться лишь в том случае, если бегущая волна после отражения от двух противоположных граней куба и прохождения пути $2L$ возвращается в исходную точку с фазой, отличающейся от первоначальных на $2\pi n$, где n — целое. Не ограничивая общности, можно считать, что двукратное отражение от граней либо не вносит в фазу волны каких-либо изменений, либо изменяет фазу на 2π . Поэтому условие образования стоячих волн в каждом из измерений куба имеет вид

$$k \cdot 2L = 2\pi n \quad (50.7)$$

или

$$k_x L = \pi n_x, \quad k_y L = \pi n_y, \quad k_z L = \pi n_z, \quad (50.8)$$

где n_x, n_y, n_z — целые числа.

Число волн dN , волновые числа которых заключены между $(k_x, k_x + dk_x), (k_y, k_y + dk_y), (k_z, k_z + dk_z)$, равно числу целых чисел, заключенных в интервале $(n_x, n_x + dn_x), (n_y, n_y + dn_y), (n_z, n_z + dn_z)$, поэтому

$$dN = dn_x dn_y dn_z = (L/\pi)^3 dk_x dk_y dk_z. \quad (50.9)$$

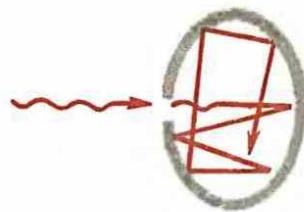
Расчет удобно вести в сферических координатах, считая, что по осям декартовой системы координат отложены k_x, k_y, k_z (рис. 271). Поскольку числа k_x, k_y, k_z положительны, в сферических координатах (50.9) принимает вид

$$dN = (L/\pi)^3 (1/8) 4\pi k^2 dk. \quad (50.10)$$

Учитывая, что $k = \omega/c$, находим концентрацию стоячих волн:

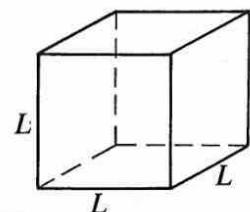
$$\frac{dN}{L^3} = \frac{1}{2} \frac{\omega^2}{\pi^2 c^3} d\omega. \quad (50.11)$$

Поскольку электромагнитная волна обладает двумя возможными поляризациями, то полная концентрация стоячих волн в два раза больше (50.11) и равна



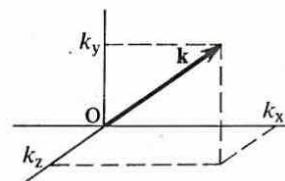
269

Реализация абсолютно черного тела



270

К расчету числа стоячих волн в кубической полости



271

Расчет числа стоячих волн в сферической системе координат

$$\frac{dN_{\text{пол}}}{L^3} = \frac{\omega^2}{\pi^2 c^3} d\omega.$$

Каждая из стоячих волн называется модой колебаний, а число мод (50.12) равно числу степеней свободы системы. Если $\langle \varepsilon \rangle$ является средней энергией, приходящейся на одну степень свободы, то плотность энергии стоячих волн равна

$$w_\omega(T) = \frac{dN_{\text{пол}}}{L^3} \langle \varepsilon \rangle = \frac{\omega^2}{\pi^2 c^3} \langle \varepsilon \rangle. \quad (50.13)$$

Таким образом, нахождение $w_\omega(T)$ свелось к определению средней энергии моды колебаний.

Формула Рэлея — Джинса. По теореме о равнораспределении энергии на одну степень свободы в классической статистической системе приходится энергия $kT/2$. У гармонического осциллятора средняя кинетическая энергия равна средней потенциальной, и поэтому его средняя энергия равна kT . Это энергия, приходящаяся на одну моду колебаний. В (50.13) положим

$$\langle \varepsilon \rangle = kT, \quad (50.14)$$

в результате получим

$$w_\omega(T) = \omega^2 kT / (\pi^2 c^3). \quad (50.15)$$

Равенство (50.15) называется формулой Рэлея — Джинса. Эта формула была предложена Д. У. Рэлеем (1842—1911) в 1900 г. и несколько более подробно обоснована Д. Д. Джинсом (1877—1946). Она дает достаточно хорошее согласие с экспериментом при малых ω . При больших ω спектральная плотность (50.15) значительно превосходит наблюдаемую, а при $\omega \rightarrow \infty$ получается недопустимое соотношение $w_\omega \rightarrow \infty$. Кроме того, полная объемная плотность излучения

$$w = \int_0^\infty w_\omega d\omega \rightarrow \infty, \quad (50.16)$$

что также недопустимо. Расходимость плотности энергии излучения (50.16) называется ультрафиолетовой катастрофой.

Формула Вина. В. Вин (1864—1928) предположил, что каждая мода колебаний является носителем энергии $\varepsilon(\omega)$, но не все моды данной частоты возбуждены. Относительное число $\Delta N/N$ возбужденных мод дается распределением Больцмана:

$$\Delta N/N = e^{-\varepsilon(\omega)/kT}. \quad (50.17)$$

Отсюда для средней энергии, приходящейся на моды с частотой ω , находим

$$\langle \varepsilon \rangle = \varepsilon(\omega) \Delta N/N = \varepsilon(\omega) e^{-\varepsilon(\omega)/kT}. \quad (50.18)$$

Из общих термодинамических соображений Вин заключил, что энергия моды частотой ω пропорциональна частоте, т. е. $\varepsilon(\omega) = \hbar\omega$. Коэффициент пропорциональности здесь дан в современных обозначениях в виде **постоянной Планка**, которая в то время не была известна. Фор-

● Направление распространения, поляризация и фаза волны вынужденного излучения совпадают с соответствующими характеристиками волны вынуждающего излучения.

Положение максимума спектральной плотности излучения абсолютно черного тела зависит от шкалы, для которой определяется спектральная плотность излучения. Максимум спектральной плотности излучения по шкале частот приходится на более длинные волны, чем по шкале длин волн.

мula (50.13) с учетом (50.18) принимает вид

$$w_{\omega}(T) = \frac{\hbar\omega^3 e^{-\hbar\omega/(kT)}}{(\pi^2 c^3)} / (\pi^2 c^3). \quad (50.19)$$

Она называется формулой Вина (1896) и дает хорошее совпадение с экспериментом в области достаточно больших частот. Если взять спектр солнечного излучения (см. рис. 1), то с помощью формулы Рэлея — Джинса удается описать лишь частоты, много меньшие тех, на которые приходится максимум плотности излучения, а с помощью формулы Вина — только большие частоты, далеко за максимумом. Промежуточную область долго не удавалось описать.

Формула Планка. М. Планк (1858—1947) предложил в 1900 г. интерполяционную формулу, которая при малых частотах переходит в формулу Рэлея — Джинса, а при больших — в формулу Вина. Она имеет вид

$$w_{\omega}(T) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{e^{\hbar\omega/(kT)} - 1}, \quad (50.20)$$

где $\hbar = 1,05 \cdot 10^{-34}$ Дж·с — постоянная Планка. При $\hbar\omega \ll kT$ (50.20) переходит в (50.15), а при $\hbar\omega \gg kT$ — в (50.19). Формула (50.20) полностью описывает излучение абсолютно черного тела и дает согласие с экспериментом.

Закон Стефана — Больцмана. Полная объемная плотность излучения равна

$$W = \int_0^{\infty} w_{\omega} d\omega = \frac{k^4 T^4}{\pi^2 c^3 \hbar^3} \int_0^{\infty} \frac{\xi^3 d\xi}{e^{\xi} - 1}. \quad (50.21)$$

Учитывая, что

$$\int_0^{\infty} \frac{\xi^3 d\xi}{e^{\xi} - 1} = \frac{\pi^4}{15}, \quad (50.22)$$

запишем (50.21) в виде

$$W = a T^4, \quad a = \frac{\pi^2 k^4}{15 c^3 \hbar^3} = 7,56 \cdot 10^{-16} \text{ Дж} \cdot \text{м}^{-3} \cdot \text{К}^{-4}. \quad (50.23)$$

Обычно вместо объемной плотности излучения используется понятие энергетической светимости, которая на основе (50.6) связана с величиной W соотношением

$$M = cw/4. \quad (50.24)$$

Это равенство с учетом (50.23) принимает вид

$$M = \sigma T^4, \quad (50.25)$$

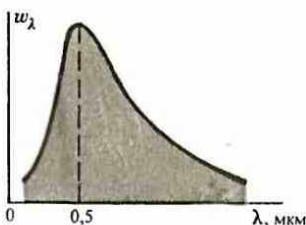
где

$$\sigma = ac/4 = 5,67032 \cdot 10^{-8} \text{ Вт} \cdot \text{м}^{-2} \cdot \text{К}^{-4}. \quad (50.26)$$

Соотношение (50.25) называется формулой Стефана — Больцмана, σ — постоянной Стефана — Больцмана.

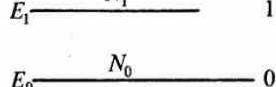
Закон смещения Вина. Максимум спектральной плотности излучения может быть найден из (50.20). Однако положение максимума зависит от шкалы, для которой определяется спектральная плотность. Для вычисления в шкале длин волн необходимо перейти к длинам волн $\lambda = 2\pi c/\omega$ и спектральной плотности излучения w_{λ} по шкале длин волн. Тогда

$$w_{\omega} d\omega = -w_{\lambda} 2\pi c d\lambda / \lambda^2 \quad (50.27)$$



272

Спектр излучения абсолютно черного тела на шкале длии волн



273

Двухуровневая система атомов

и поэтому распределение спектральной плотности излучения по длинам волн имеет вид (рис. 272)

$$W_\lambda = \frac{16\pi^2\hbar c}{\lambda^5 (\exp [2\pi c\hbar/(kT\lambda)] - 1)}. \quad (50.28)$$

Максимум спектральной плотности излучения находится из условия

$$\partial w_\lambda / \partial \lambda = 0, \quad (50.29)$$

дающего для определения λ_{\max} уравнение

$$5 = xe^x / (e^x - 1), \quad x = 2\pi c\hbar / (kT\lambda_{\max}). \quad (50.30)$$

Решением этого уравнения является $x = 4,965$. Поэтому λ_{\max} определяется соотношением

$$\lambda_{\max} T = 2\pi c\hbar / (kx) = 0,0029 \text{ м}\cdot\text{К}, \quad (50.31)$$

называемым законом смещения Вина Спектр излучения абсолютно черного тела по шкале длин волн показан на рис. 272. Из формулы (50.31) для λ_{\max} в спектре излучения Солнца ($T_{\text{пов}} \approx 6000$ К) получается значение около 0,550 мкм. Для расчета максимума по спектру частот вместо (50.29) надо решить уравнение $\partial w_\omega / \partial \omega = 0$. В результате для спектра излучения Солнца при частоте ω_{\max} получается значение около $2,1 \times 10^{15} \text{ с}^{-1}$, что соответствует длине волны около 0,88 мкм.

Элементарная квантовая теория. Получить формулу Планка в рамках классических представлений невозможно. Однако ее удается обосновать с помощью элементарных рассуждений, основывающихся на квантовых представлениях об излучении и поглощении света. Хотя эти рассуждения не дают полного количественного решения задачи об излучении абсолютно черного тела, они дают достаточно ясное представление о механизме динамического равновесия между излучением и материальными телами.

В термодинамическом равновесии находится оболочка полости тела и излучение в полости. Излучение представляется совокупностью квантов с энергией $\varepsilon = \hbar\omega$. Кванты могут поглощаться атомами, которые при этом переходят на более высокий энергетический уровень с энергией $E_1 = E_0 + \hbar\omega$, где E_0 — исходный энергетический уровень атома. При переходе атома с уровня E_1 на E_0 излучается квант с энергией $E_1 - E_0 = \hbar\omega$. Обозначим эти уровни индексами 0 и 1 (рис. 273) и назовем соответственно нижним и верхним уровнем.

Динамическое равновесие осуществляется посредством постоянного обмена квантами между полем излучения и материальными телами, в результате которого происходят переходы атомов между уровнями энергии и изменяется число квантов в поле излучения. Согласно принципу детального равновесия, обмен квантами должен уравновешиваться для каждой частоты в отдельности (точнее, для каждой моды излучения). Поэтому

- Связан ли расчет концентрации мод колебаний с конкретизацией граничных условий? Сформулируйте условия, при которых получаются формулы Рэлея — Джинса и Вина. Для какой шкалы спектральной плотности энергии сформулирован закон смещения Вина?

рассмотрим лишь одну частоту. Для других частот все рассуждения аналогичны, меняются лишь уровни атомов, переходы между которыми осуществляют динамическое равновесие.

Спонтанные и вынужденные переходы. С нижнего уровня на верхний по закону сохранения энергии переходы возможны только с поглощением кванта энергии, т. е. под влиянием излучения, падающего на атом. Такие переходы называются вынужденными. Переходы с верхнего уровня на нижний могут быть как вынужденными, под влиянием падающего на атом излучения, так и спонтанными, или самопроизвольными, происходящими независимо от падающего на атом излучения.

Обозначим $v_{10}^{(c)}$ частоту переходов атомов с верхнего уровня на нижний спонтанно, $v_{10}^{(v)}$ — вынужденно и $v_{01}^{(v)}$ — с нижнего уровня на верхний вынужденно. Условие динамического равновесия записывается в виде

$$v_{10}^{(c)} + v_{10}^{(v)} = v_{01}^{(v)}. \quad (50.32)$$

Коэффициенты Эйнштейна. В 1917 г. А. Эйнштейн (1879—1955) предложил элементарную квантовую картину динамического равновесия между излучением и материальной средой, приводящую к правильной формуле излучения абсолютно черного тела.

Обозначим A_{10} вероятность спонтанного перехода $1 \rightarrow 0$ в секунду, N_1 — концентрацию атомов на верхнем уровне. По определению этих величин можно записать

$$v_{10}^{(c)} = N_1 A_{10}. \quad (50.33)$$

Частота вынужденных переходов, очевидно, пропорциональна плотности излучения. Обозначим B_{10} и B_{01} вероятности вынужденных переходов $1 \rightarrow 0$ и $0 \rightarrow 1$ в секунду, отнесенные к спектральной плотности излучения w_ω ; N_0 — концентрацию атомов на нижнем уровне. По определению этих величин можно записать

$$v_{10}^{(v)} = N_1 w_\omega B_{10}, \quad v_{01}^{(v)} = N_0 w_\omega B_{01}. \quad (50.34)$$

С учетом (50.33) и (50.34) соотношение (50.32) принимает вид

$$N_1 A_{10} + N_1 w_\omega B_{10} = N_0 w_\omega B_{01}. \quad (50.35)$$

В равновесном состоянии справедливо распределение Больцмана, которое для концентраций атомов принимает вид

$$N_1 = A e^{-E_1/(kT)}, \quad N_0 = A e^{-E_0/(kT)}, \quad (50.36)$$

где A — нормировочная постоянная. Подставляя (50.36) в (50.35) и сокращая полученное выражение на общий множитель A , находим

$$A_{10} e^{-E_1/(kT)} + w_\omega B_{10} e^{-E_1/(kT)} = w_\omega B_{01} e^{-E_0/(kT)} \quad (50.37)$$

Величины A_{10} , B_{10} , B_{01} называют коэффициентами Эйнштейна.

Из физических соображений ясно, что при $T \rightarrow \infty$ должно быть $w_\omega \rightarrow \infty$. Тогда, разделив обе части (50.37) на w_ω , получим

$$B_{10} = B_{01}. \quad (50.38)$$

Поэтому соотношение (50.37) может быть записано в виде

$$w_\omega = \frac{A_{10}}{B_{10}} \frac{1}{e^{\hbar\omega/(kT)} - 1}, \quad (50.39)$$

где $\hbar\omega = E_L - E_0$. Теоретически значение отношения A_{10}/B_{10} в рамках элементарной квантовой теории найдено быть не может. Оно вычисляется в рамках строгой квантовой теории излучения. Однако значение этого отношения можно найти, если учесть, что (50.39) при малых частотах должно совпадать с формулой Рэлея — Джинса (50.15), а при больших частотах — с формулой Вина (50.19).

При $\hbar\omega \ll kT$ можем считать, что $\exp[\hbar\omega/(kT)] \approx 1 + \hbar\omega/(kT)$, и записать (50.39) в виде

$$w_\omega = \frac{A_{10}}{B_{10}} \frac{kT}{\hbar\omega}. \quad (50.40)$$

Сравнивая (50.40) с (50.15), находим

$$A_{10}/B_{10} = \hbar\omega^3/(\pi^2 c^3). \quad (50.41)$$

Формула (50.39) принимает вид

$$w_\omega = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{e^{\hbar\omega/(kT)} - 1}, \quad (50.42)$$

совпадающий с формулой Планка (50.20).

Хотя элементарная квантовая теория излучения абсолютно черного тела не позволяет теоретически вычислить значения коэффициентов Эйнштейна, она демонстрирует необходимость существования спонтанных и вынужденных переходов, причем для вероятностей вынужденных переходов соблюдается важное соотношение (50.38).

Испущенные в результате спонтанных переходов кванты имеют случайное направление распространения, случайную поляризацию и случайную фазу. Кванты, испущенные в результате вынужденных переходов, коррелируют по своим свойствам с излучением, которое вызывает переход. Вынужденное излучение обладает той же поляризацией, тем же направлением распространения и той же фазой, что и вынуждающее переход излучение. Это свойство вынужденного (или индуцированного) излучения чрезвычайно важно для его применений и проявлений.

§ 51 | Оптические усилители

Обсуждаются условия усиления светового потока при прохождении через среду.

Прохождение света через среду. При прохождении света через среду осуществляется обмен квантами между пучком света и атомами среды посредством вынужденных переходов и спонтанное испускание квантов. Обозначим частоту излучения, концентрации атомов на верхнем и нижнем уровнях соответственно ω , N_1 , N_0 (см. рис. 273). Спектральную объемную плотность излучения частоты ω обозначим w_ω . Она изменяется в результате вынужденного поглощения квантов атомами среды, благодаря чему плотность потока уменьшается, и вследствие вынужденного излучения атомов, приводящего к увеличению плотности w_ω . Закон сохранения энергии при вынужденных переходах на основании (50.34) записывается в виде

$$dw_\omega/dt = \hbar\omega w_\omega B_{10}N_1 - \hbar\omega w_\omega B_{01}N_0 = \hbar\omega w_\omega B(N_1 - N_0), \quad (51.1)$$

где $B = B_{10} = B_{01}$. С помощью обозначения

$$\alpha = \hbar\omega B(N_1 - N_0)/v \quad (51.2)$$

уравнение (51.1) может быть записано в виде

$$dw_\omega/dt = \alpha v w_\omega, \quad (51.3)$$

где v — скорость света с частотой ω в среде. Это уравнение может быть переписано в виде

$$dS/dt = \alpha v S, \quad (51.4)$$

где

$$S = v w_\omega, \quad (51.5)$$

— плотность потока энергии. Считая, что свет распространяется в направлении оси Z , можем написать

$$\frac{dS}{dt} = \frac{dS}{dz} \frac{dz}{dt} = \frac{dS}{dz} v, \quad (51.6)$$

где $dz/dt = v$ — скорость распространения света в среде. Ясно, что здесь речь идет о групповой скорости, с которой распространяется энергия. С учетом (51.6) уравнение (51.4) принимает вид

$$dS/dz = \alpha S \quad (51.7)$$

и имеет решение

$$S(z) = S_0 e^{\alpha z}, \quad (51.8)$$

где $S_0 = S(0)$.

Закон Бургера. Экспоненциальная зависимость (51.8) плотности потока от расстояния называется законом Бургера. В состоянии термодинамического равновесия концентрации атомов описываются распределением Больцмана. Из неравенства $E_1 > E_0$ следует $N_1 < N_0$ и поэтому $\alpha < 0$. Это означает, что плотность потока по мере прохождения света в среде уменьшается. Механизм уменьшения плотности состоит в следующем. В результате вынужденных переходов атомов с нижнего энергетического уровня на верхний плотность энергии потока уменьшается. При переходе атомов с верхнего уровня на нижний лишь часть квантов возвращается в поток, а именно кванты, испущенные в результате вынужденных переходов. Кванты, испущенные спонтанно, в поток не возвращаются, что и является причиной уменьшения его плотности.

Условия усиления. Если привести систему атомов в неравновесное состояние и тем самым достаточно сильно нарушить распределение Больцмана, то можно добиться изменения концентрации атомов на различных уровнях так, чтобы было $N_1 \approx N_0$ или даже $N_1 > N_0$. В первом случае $\alpha \approx 0$ и пучок через среду распространялся бы без поглощения, а во втором случае $\alpha > 0$ и пучок при прохождении усиливался бы, т. е. среда действовала бы как усилитель светового потока.

Это позволяет создавать генераторы и усилители волн, основанные на индуцированном излучении, которые для светового диапазона называются лазерами, а для микроволнового — мазерами. За фундаментальные исследования в области квантовой электроники, которые привели к созданию генераторов и усилителей нового типа — мазеров и лазеров, советским ученым А. М. Прохорову (р. 1916 г.) и Н. Г. Басову (р. 1922 г.) и американскому ученному Ч. Х. Таунсу (р. 1915 г.) в 1964 г. была присуждена Нобелевская премия. Идея о возможности использования индуцированного излучения для усиления светового потока была высказана в 1940 г. советским физиком В. А. Фабрикантом (р. 1907 г.).

Воздействие светового потока на заселенность уровней. Световой пучок, вызывая вынужденные переходы атомов между уровнями, изменяет их заселенность. Обозначим N полную концентрацию атомов. По определению,

$$N = N_1 + N_0. \quad (51.9)$$

Изменение заселенности верхнего уровня происходит за счет вынужденных и спонтанных переходов и может быть описано уравнением

$$dN_1/dt = -w_\omega B(N_1 - N_0) - N_1/\tau, \quad (51.10)$$

где N_1/τ учитывает частоту спонтанных переходов. Подставляя в (51.10) выражение для N_0 из (51.9), получим

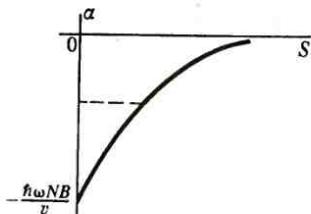
$$dN_1/dt = -SB(2N_1 - N)/v - N_1/\tau, \quad (51.11)$$

где $w_\omega = S/v$. При достижении стационарного состояния должно соблюдаться условие $dN_1/dt = 0$,

$$-SB(2N_1 - N)/v - N_1/\tau = 0, \quad (51.12)$$

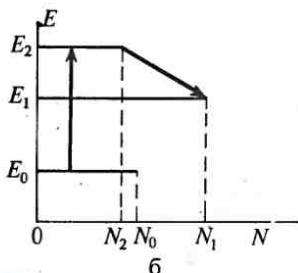
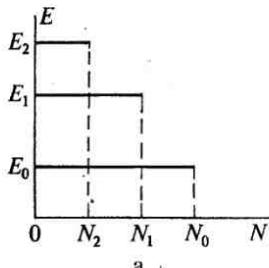
которое с учетом (51.11) принимает вид

$$-SB(2N_1 - N)/v - N_1/\tau = 0, \quad (51.13)$$



274

Зависимость показателя ослабления от плотности потока энергии в световом пучке



275

Создание инверсной заселенности в трехуровневой системе: а — нормальная заселенность; б — инверсная

- При каких условиях свет при прохождении через среду ослабляется? проходит без ослабления? усиливается? Как эти условия реализуются? Опишите трехуровневый механизм создания инверсной заселенности уровней. Какие при этом требования предъявляются к энергетическим уровням?

откуда следует

$$N_1 = \frac{N}{2} \frac{1}{1 + v/(2\tau BS)}. \quad (51.14)$$

Из (51.14) видно, что с увеличением плотности потока S заселенность верхнего уровня увеличивается. Это приводит к соответствующему изменению коэффициента.

$$\alpha = \hbar\omega B(N_1 - N_0)/v = \hbar\omega B(2N_1 - N)/v =$$

$$= -\frac{\hbar\omega NB}{v} \left[\frac{1}{1 + 2\tau BS/v} \right]. \quad (51.15)$$

Условия насыщения. График зависимости α от S показан на рис. 274. Видно, что при любых S происходит поглощение. При $S \rightarrow \infty$ оно полностью прекращается и наступает насыщение. Поскольку α стоит в экспоненте, принимается, что насыщение практически наступает при изменении α в два раза по сравнению с его минимальным значением, т. е. при увеличении знаменателя (51.15) в два раза. Поэтому **условие насыщения записывается в виде**

$$v/(2\tau BS) = 1, \quad (51.16)$$

и, следовательно, **насыщение наступает при плотности потока энергии**

$$S = v/(2\tau B). \quad (51.17)$$

Таким образом, световой поток выравнивает заселенность двух уровней, между которыми он обуславливает вынужденные переходы, а при достаточно большой плотности потока может даже почти сравнять их заселенность, но он не может создать инверсную заселенность между этими уровнями.

Создание инверсной заселенности. Инверсной заселенности энергетических уровней можно добиться с помощью некоторого воздействия на атомы, независимого от усиливаемого света.

Наиболее простой путь создания инверсной заселенности осуществляется в трехуровневых системах (рис. 275). На рис. 275, а изображено распределение заселенностей в равновесном состоянии системы. При воздействии на систему излучением большой мощности с частотой $\omega_n = (E_2 - E_0)/\hbar$ заселенности уровней E_0 и E_2 при выполнении условия (51.17) практически сравниваются. Допустим, что время жизни атомов на уровне E_2 очень мало и они спонтанно переходят на уровень E_1 , время жизни на котором у них достаточно велико. Ясно, что атомы на уровне E_1 будут накапливаться, в результате чего создается инверсная заселенность между уровнями E_1 и E_0 (рис. 275, б) Этот переход может быть использован для усиления света с частотой $\omega = (E_1 - E_0)/\hbar$.

Рассматриваются принципы и характеристики работы лазеров.

Принципиальная схема лазера. Среда с инверсной заселенностью, способная усиливать проходящий через нее световой поток, называется активной. Заполним пространство между пластины интерферометра Фабри — Перо активной средой (рис. 276). Между последовательными отражениями от зеркал при прохождении через активную среду световой поток усиливается. Эта система образует активный оптический резонатор. Усиление потока при прохождении через активную среду происходит в соответствии с формулой (51.8). При отражении от зеркал излучение частично ослабляется. Одно из зеркал делается с максимально возможным коэффициентом отражения, а через другое зеркало свет в определенной пропорции выходит из системы, образуя ее излучение, которое называется лазерным. Кроме потерь света при отражении от зеркал имеются потери за счет рассеяния в среде и других дифракционных эффектов. Для работы системы в качестве генератора света необходимо обеспечить определенный баланс между усилением светового потока при прохождении через активную среду и ослаблением за счет всех факторов, включая само лазерное излучение.

Порог генерации. Элементарный цикл работы лазера включает два последовательных прохождения через активную среду и соответствующие отражения от зеркал. Потери энергии могут быть учтены эффективными коэффициентами отражения ρ_1 и ρ_2 на зеркалах, причем они учитывают не только отражения от зеркал (вообще говоря, различные, поскольку через одно из них из лазера выходит излучение), но и другие потери, о которых говорилось выше. Таким образом, ρ_1 и ρ_2 меньше коэффициентов отражения только от зеркал резонатора. За один цикл проходят два отражения света и, следовательно, ослабление потока пропорционально $\rho_1 \rho_2$. За один цикл свет в активной среде проходит путь $2L$. Поэтому на основании (51.8) усиление потока за цикл пропорционально $\exp(a2L)$, где a — значение коэффициента усиления (51.2) за цикл. Полное усиление плотности потока энергии за один цикл описывается формулой

$$S = S_0 \rho_1 \rho_2 e^{2aL}, \quad (52.1)$$

где S_0 — плотность потока энергии в начале цикла (за начало цикла можно взять любой момент времени). Перепишем (52.1) в виде

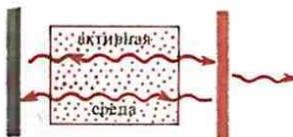
$$S = S_0 e^{2aL - 2f}, \quad (52.2)$$

где

$$2f = -\ln(\rho_1 \rho_2). \quad (52.3)$$

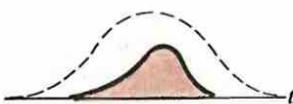
Из (52.2) видно, что генерация лазерного излучения начинается тогда, когда приобретаемая световым потоком в активной среде энергия за цикл превосходит потери энергии, включая энергию покинувшего систему лазерного излучения. На пороге генерации плотность потока энергии в системе не столь велика, чтобы изменить заселенность уровней, и поэтому для порога

- Генерация лазерного излучения начинается лишь тогда, когда передаваемая в резонаторе за цикл энергия от активной среды в световой поток начинает превосходить суммарные потери светового потока в резонаторе, включая энергию, уносимую лазерным излучением. Количественно начало генерации лазерного излучения характеризуется условием порога генерации.



276

Принципиальная схема лазера



277

Импульсная накачка лазера (пунктирная линия — импульс накачки; сплошная — импульс излучения)



278

Модуляция добротности вращающейся призмой

генерации можно считать $a = a_0$, где a_0 — коэффициент усиления в отсутствие светового потока. Следовательно, на основании (52.2) условие порога генерации имеет вид

$$a_0 L = f. \quad (52.4)$$

Условия стационарной генерации. При стационарной генерации потери энергии компенсируются за счет энергии, полученной световым пучком от активной среды. Стационарная генерация может осуществляться при значительных плотностях потока энергии, поэтому в качестве a следует взять его значение при наличии потока. Условие стационарной генерации имеет вид

$$aL = f. \quad (52.5)$$

Добротность. Поскольку лазер представляет собой оптический резонатор, в теории лазеров широко используется терминология теории колебаний. Потерю энергии за период колебаний принято характеризовать добротностью — отношением запасенной в системе энергии W к потерям энергии за одно колебание ΔW :

$$Q = W / \Delta W. \quad (52.6)$$

Запасенная в системе энергия равна

$$W = w\sigma L, \quad (52.7)$$

где σ — площадь поперечного сечения лазерного пучка. Потеря энергии за один цикл составляет

$$(w/2)\sigma L (1 - e^{-2f}), \quad (52.8)$$

где учтено, что объемная плотность энергии w складывается из равных объемных плотностей двух потоков энергии, движущихся в противоположных направлениях. Учитывая, что продолжительность одного цикла равна $2L/v$, и обозначая период лазерного излучения через $T = 2\pi/\omega$, находим, что потеря энергии за одно колебание выражается формулой

$$\Delta W = \frac{(w/2)\sigma L (1 - e^{-2f})}{2L/v} T = \frac{1}{2} w v \sigma f T, \quad (52.9)$$

где $\exp(-2f) \approx 1 - 2f$, поскольку $f \ll 1$. Отсюда по формуле (52.6) находим

$$Q = \frac{\sigma w L}{w v \sigma f T / 2} = \frac{2L}{\beta} = \frac{m}{f}, \quad (52.10)$$

где $m = 2L/\lambda = L/(\lambda/2)$ — число стоячих полуволн в резонаторе, λ — длина волны излучения. С помощью (52.10) условие (52.4) выражается особенно наглядно:

$$a_0 \lambda / 2 = 1/Q, \quad (52.11)$$

- Каков физический смысл порога генерации и условия стационарной генерации.

В чем состоит метод модуляции добротности для увеличения мощности излучения?

т. е. для осуществления генерации необходимо, чтобы усиление на пути в половину длины волны было равно (или больше) величине, обратной добротности.

Порог генерации тем выше, чем меньше добротность. Добротность же тем меньше, чем больше потери. Поэтому для осевых лучей порог генерации достигается раньше и энергия излучения лазера уносится преимущественно ими. Это означает, что излучение лазера сосредоточено в узком параллельном пучке лучей, угол расходимости которого обусловливается главным образом дифракцией.

Непрерывные и импульсные лазеры. Создание инверсной заселенности уровней называется накачкой. Накачка лазеров может быть самой разнообразной и будет рассмотрена в связи с конкретными типами лазеров. По характеру зависимости накачки от времени она может быть непрерывной и импульсной. Если накачка осуществляется импульсами, то и излучение лазера импульсное. После начала импульса накачки начинает изменяться заселенность уровней. Когда достигаются условия порога генерации (52.4), начинается испускание лазерного излучения (рис. 277).

При непрерывной накачке, которая постоянно обеспечивает соблюдение условия (52.5), излучение лазера непрерывно. Следует, однако, отметить, что при непрерывной накачке возможжен также импульсный режим излучения, так же как возможно излучение нескольких импульсов излучения при одном импульсе накачки.

Повышение мощности излучения. Для повышения мощности излучения необходимо увеличить число атомов, участвующих в усилении светового потока в резонаторе лазера за счет индуцированного излучения, и уменьшить длительность импульса.

Метод модулированной добротности. Чтобы увеличить число атомов, участвующих почти одновременно в усилении светового потока, необходимо задержать начало генерации, чтобы накопить как можно больше возбужденных атомов, создающих инверсную заселенность, для чего надо поднять порог генерации лазера и уменьшить добротность. Это можно сделать посредством увеличения потерь светового потока. Например, можно нарушить параллельность зеркал, что резко уменьшит добротность системы. Если при такой ситуации начать накачку, то даже при значительной инверсии заселенности уровней генерация не начинается, поскольку порог генерации высок. Поворот зеркала до параллельного другому зеркалу положения повышает добротность системы и тем самым понижает порог генерации. Когда добротность системы обеспечит начало генерации, инверсная заселенность уровней будет весьма значительной. Поэтому мощность излучения лазера сильно увеличивается. Такой способ управления генерацией лазера называется методом модулированной добротности.

Продолжительность импульса излучения зависит от того, в течение какого времени вследствие излучения инверсная заселенность изменится настолько, что система выйдет из условия генерации. Продолжительность зависит от многих факторов, но обычно составляет 10^{-7} — 10^{-8} с.

Очень распространено модулирование добротности с помощью вращающейся призмы (рис. 278). При определенном положении она обеспечивает полное отражение падающего вдоль оси резонатора луча в обратном направлении. Частота вращения призмы составляет десятки или сотни герц. Импульсы лазерного излучения имеют такую же частоту, хотя длительность каждого импульса составляет 10^{-7} — 10^{-8} с.

Более частое повторение импульсов может быть достигнуто модуляцией добротности с помощью ячейки Керра (см. § 45). Ячейку Керра и поляризатор помещают в резонатор. Поляризатор обеспечивает генерацию лишь излучения определенной поляризации, а ячейка Керра ориентирована так, чтобы при наложении на нее напряжения не проходил свет с этой поляризацией. При накачке лазера напряжение с ячейки Керра снимается в такой момент времени, чтобы начавшаяся при этом генерация была наиболее сильной.

Имеются также и другие способы введения потерь, приводящие к соответствующим методам модуляции добротности.

Моды излучения. В стационарном режиме излучения лазера в его оптическом резонаторе, по определению стационарности, должны образоваться стоячие волны. В самом общем виде стоячая волна может быть представлена как суперпозиция элементарных стоячих волн, называемых модами колебаний. Излучение лазера, соответствующее этим модам колебаний, называют **модами излучения лазера**.

Мода колебаний зависит от геометрических характеристик резонатора, от коэффициента преломления активной среды и, вообще говоря, от условий на граничных поверхностях резонатора. Рассмотрим для конкретности резонатор с прямоугольными плоскими зеркалами и цилиндрический резонатор со сферическими зеркалами.

Резонатор с прямоугольными плоскими зеркалами (рис. 279). Частное решение волнового уравнения (2.9) имеет вид

$$E = E_0 \begin{bmatrix} \sin \omega t \\ \cos \omega t \end{bmatrix} \begin{bmatrix} \sin k'_x x \\ \cos k'_x x \end{bmatrix} \begin{bmatrix} \sin k'_y y \\ \cos k'_y y \end{bmatrix} \begin{bmatrix} \sin k'_z z \\ \cos k'_z z \end{bmatrix}, \quad (53.1)$$

причем можно выбрать произвольную комбинацию сомножителей, по одному из каждой скобки (выбор комбинации определяется граничными условиями, которым надо удовлетворить); E_0 — постоянная амплитуда, определяющая поляризацию колебаний. Подстановка выбранной в (53.1) комбинации сомножителей в уравнение (2.9) приводит к равенству

$$k'^2 + k'^2 + k'^2 = k'^2 = \mu \epsilon \omega^2, \quad (53.2)$$

где волновое число k' относится к среде, т. е. $k' = 2\pi/\lambda_c$, λ_c — длина волны в среде. В уравнении (2.9) сделана замена $\epsilon_0 \rightarrow \epsilon$, $\mu_0 \rightarrow \mu$, чтобы записать его для среды. Учитывая, что $\mu_\epsilon = 1/v^2$, из (53.2) получаем

$$k' = \omega/v = \omega n/c = kn, \quad (53.3)$$

где $n = c/v$ — показатель преломления среды, $k = \omega/c$ — волновое число в вакууме.

Возникновение стоячих волн в каждом измерении прямоугольного резонатора происходит при выполнении **условия цикличности**

$$m\lambda_c = 2l, \quad (53.4)$$

где λ_c — длина волны в среде, l — размер резонатора в соответствующем измерении. Стоящая волна может быть представлена как сумма двух бегущих волн, распространяющихся в противоположных направлениях. Условие цикличности означает, что после двух последовательных отражений от зеркал и прохождения двух длин резонаторов все характеристики волны должны вернуться к своим исходным значениям. Отсюда следует, в частности, что эти волны должны иметь определенное состояние поляризации. Быть неполяризованными они не могут. Характер поляризации этих волн зависит от резонатора. В правую часть (53.4), строго говоря, необходимо включить изменение фазы волн при отражении от зеркал. Однако это обстоятельство не вносит ничего существенного в анализ процесса, а лишь делает громоздкими формулы и поэтому не учитывается.

Считая ось Z направленной перпендикулярно поверхностям зеркал резонатора, запишем условие (53.4) в виде

316 $k_x' a = m_x \pi,$

— $k_y' b = m_y \pi,$

10 $k_z' L = m_z \pi,$

(53.5)

где a и b — размеры резонатора в направлении осей X и Y .

Заметим, что условия (53.5) получены не как запись граничных условий, которые всегда необходимы для получения решения задачи, сформулированной в виде дифференциальных уравнений. Они получены как следствие более общего физического, а не математического требования (53.4) цикличности волнового поля, что весьма существенно, поскольку это требование в явном виде содержит предположение об образовании стоячих волн как результата суперпозиции бегущих волн.

Аксидальные (продольные) моды. Моды характеризуются набором чисел (m_x, m_y, m_z) . Главной называется мода $(0, 0, m_z)$. Она не имеет узлов в плоскости, перпендикулярной оси Z и описывает стоячую волну, являющуюся суперпозицией встречных бегущих волн, распространяющихся параллельно оси Z . Вне резонатора ей соответствует волна, распространяющаяся параллельно оси лазера. В теории волноводов эта мода называется поперечной электромагнитной модой и обозначается TEM_{00m_z} . Из (53.3) с учетом (53.5) для частот излучения этой моды получаем выражение

$$\omega_{m_z} = k_z' c/n = c \pi m_z / (L n). \quad (53.6)$$

Возможные круговые частоты аксиальных мод разделены интервалом

$$\Delta \omega = c \pi \cdot \Delta m_z / (L n) = c \pi / (L n), \quad (53.7)$$

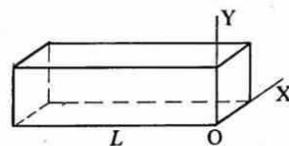
поскольку число m_z для соседних частот различается на единицу ($\Delta m_z = 1$). Расстояние по частотам равно $\Delta v = \Delta \omega / (2\pi) = c / (2Ln)$. Учитывая, что $v = c/\lambda$, запишем это соотношение в виде $\Delta v/v = \lambda/(2Ln)$.

Лазером генерируются не все частоты, а лишь те, которые удовлетворяют условиям порога генерации. На рис. 280 изображена кривая $a_0(\omega)$. Генерация возможна лишь для частот, заключенных между ω_1 и ω_2 . Генерируемые частоты расположены на равных расстояниях (53.7) друг от друга. Сколько генерируемых частот попадает в этот интервал, зависит от конкретных условий. У гелий-неонового лазера ($\lambda = 632,8$ нм) это число равно 5—10, в рубиновом лазере — несколько сотен, а в лазерах на красителях — несколько тысяч. Однако можно добиться одномодового, двухмодового режима и т. д.

Ширина линий излучения. Каждая из излучаемых линий, попадающих в полосу усиления, не является строго монохроматичной; она имеет конечную ширину. Ее ширина Δv может быть определена из общего условия

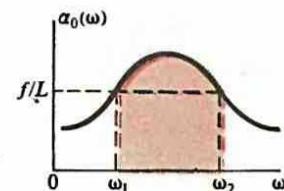
$$\Delta v t \approx 1, \quad (53.8)$$

где t — продолжительность излучения. В лазерах с модуляцией добротности $t \approx 10^{-8}$ с и, следовательно, $\Delta v = 10^8$ Гц.



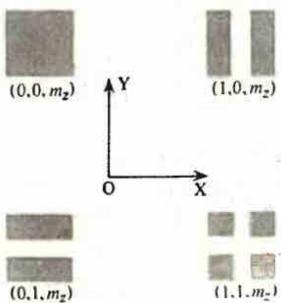
279

Прямоугольный оптический резонатор



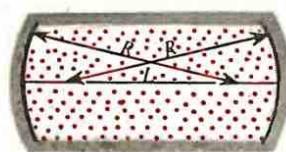
280

Метод определения частот генерируемых мод



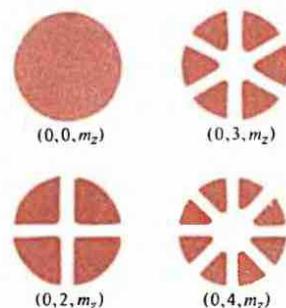
281

Моды лазерного излучения для прямоугольного резонатора



282

Цилиндрический резонатор со сферическими зеркалами



283

Моды лазерного излучения в цилиндрическом резонаторе со сферическими зеркалами

В непрерывном режиме теоретически можно для продолжительности излучения τ принять любые значения, а следовательно, можно теоретически получить сколь угодно тонкие линии. Однако практически происходят неконтролируемые флуктуации показателя преломления, изменения расстояния L в результате вибраций, теплового расширения стекны из-за колебаний температуры и т. д. Все это приводит к уширению линий. Например, из (53.6) следует, что $\Delta v/v = -\Delta L/L$. Поэтому небольшое изменение $\Delta L = \lambda/50$ при $L = 1$ м, $\lambda = 0,5$ мкм дает $\Delta v/v = 10^{-8}$. Однако все эти факторы не играют принципиальной роли, их действие может быть уменьшено совершенствованиями технической стороны дела. Принципиальное значение имеют броуновское движение зеркал и спонтанное излучение атомами среды. Однако уширение за счет этих факторов составляет очень малую величину порядка $\Delta v \approx 10^2 - 10^3$ Гц, что при $v \approx 10^{15}$ Гц дает $\Delta v/v = 10^{-13} - 10^{-16}$. В настоящее время этот предел почти достигнут.

Боковые моды. Значения m_x и m_y характеризуют число узлов стоячей волны в плоскости $X Y$. При $m_x = 0$, $m_y = 0$ узлов нет и поэтому на выходе из лазера аксиальная мода $(0, 0, m_z)$ дает распределение интенсивностей без узлов (рис. 281). Моды с $m_x \neq 0$ и $m_y \neq 0$ называются боковыми. Распределение интенсивности излучения этих мод на выходе из лазера характеризуется наличием линий нулевой интенсивности, соответствующих узлам стоячих волн при $m_x \neq 0$ и $m_y \neq 0$. Распределение интенсивности излучения на выходе из лазера в некоторых первых боковых модах показано на рис. 281. Из (53.2) с учетом (53.5) находим

$$\omega_{m_x, m_y, m_z} = \frac{\pi c}{n} \sqrt{\left(\frac{m_x}{a}\right)^2 + \left(\frac{m_y}{b}\right)^2 + \left(\frac{m_z}{L}\right)^2}. \quad (53.9)$$

Обозначим $\Delta \omega_x$ и $\Delta \omega_z$ изменение частоты при переходе от одной моды к другой при $\Delta m_x = 1$ и $\Delta m_z = 1$. Из (53.9) имеем

$$\Delta \omega_x = \frac{\pi c}{n} \left[\left(\frac{m_x}{a} \right)^2 + \left(\frac{m_y}{b} \right)^2 + \left(\frac{m_z}{L} \right)^2 \right]^{-1/2} \frac{m_x}{a^2}, \quad (53.10)$$

$$\Delta \omega_z = \frac{\pi c}{n} \left[\left(\frac{m_x}{a} \right)^2 + \left(\frac{m_y}{b} \right)^2 + \left(\frac{m_z}{L} \right)^2 \right]^{-1/2} \frac{m_z}{L^2}. \quad (53.11)$$

Из (53.11) при $m_x = m_y = 0$ получаем формулу (53.7). Дели почлененно левые и правые части (53.10) на соответствующие части (53.11), находим

$$\frac{\Delta \omega_x}{\Delta \omega_z} = \frac{m_x}{m_z} \left(\frac{L}{a} \right)^2. \quad (53.12)$$

Учитывая, что $m_z \sim L/\lambda$, $m_x \sim 1$, из (53.12) получаем

$$\frac{\Delta \omega_x}{\Delta \omega_z} \sim \frac{\lambda}{a} \frac{L}{a} \ll 1, \quad (53.13)$$

т. е. расстояние между частотами боковых мод значительно меньше, чем у аксиальных мод.

Цилиндрический резонатор со сферическими зеркалами.

Для стоячих волн в этом резонаторе поверхности зеркал являются поверхностями одинаковой фазы. Другими словами, волновой фронт изменяется вдоль оси Z и на зеркалах совпадает с поверхностью зеркал (рис. 282). При равных радиусах кривизны зеркал в середине резонатора волновой фронт плоский. Стоячую волну, как обычно, можно себе представить как суперпозицию двух волн, распространяющихся в противоположных направлениях. За один цикл, в течение которого волна дважды отражается от зеркал и дважды проходит через резонатор, все характеристики каждой из волн — отраженной и прошедшей — должны возвратиться к своим исходным значениям. Расчет показывает, что условие цикличности для отраженных волн имеет вид

$$[m_z + (1/\pi)(2p + q + 1) \arccos(1 - L/p)]\lambda_c = 2L, \quad (53.14)$$

где m_z, p, q — целые числа, характеризующие моду. Решение для электрического поля может быть записано в цилиндрических координатах с осью в направлении Z в виде

$$E(r, \phi, z) = E_0 \left[\frac{r}{\sigma} \sqrt{2} \right]^q L_q^p \left(2 \frac{r^2}{\sigma^2} \right) \exp \left[-\frac{r^2 \cos q\phi}{\sigma^2} \right] \sin kz. \quad (53.15)$$

Здесь L_q^p — полиномы Лагерра и σ определяется соотношением

$$\sigma^4 = \left(\frac{\lambda}{\pi} \right)^2 \frac{R^2 L}{(2R - L)}, \quad (53.16)$$

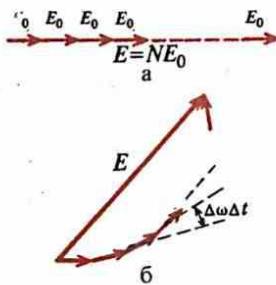
где R — радиус кривизны зеркал.

Наиболее важное значение имеют поперечные электромагнитные волны, которые обозначаются TEM_{pqm_z} . При $p = 0, q = 0$ мода является аксиальной. Условие (53.14) для аксиальной моды совпадает с условием (53.4) для аксиальной моды в прямоугольном резонаторе. Моды с $p \neq 0$ и $q \neq 0$ являются боковыми. Значение p определяет число узлов по радиусу, значение q равно половине числа аксиальных узлов. Распределение интенсивности излучения различных мод без узлов по радиусу ($p = 0$) на выходе из лазера показано на рис. 283. При $p = 1, q = 0$ в точке $r = 0$ возникает узел, который на рис. 283 в моде $(0, 0, m_z)$ может быть изображен темной точкой. При других значениях p возникают концентрические окружности нулевой интенсивности.

Синхронизация мод. Продолжительность импульсов излучения лазера, получаемых методом модуляции добротности, обычно колеблется в пределах 10^{-7} – 10^{-8} с. С помощью ячеек Керра удается получать импульсы продолжительностью $\tau = 10^{-9}$ с. Более короткие импульсы ($\sim 10^{-12}$ с) получаются в результате синхронизации мод. Суть синхронизации мод состоит в следующем.

В многомодовом режиме лазер излучает спектр частот, интервал между которыми для аксиальных мод равен $\Delta\omega$, определяемой по формуле (53.7). Если все моды складываются в одинаковой фазе, то амплитуда суммарной волны равна сумме амплитуд мод (рис. 284, а). Через промежуток времени Δt фазы волн изменяются на $\Delta\omega\Delta t, 2\Delta\omega\Delta t, \dots, (N-1)\Delta\omega\Delta t$, где N — число мод, участвующих в образовании суммарной амплитуды. Картина сложения комплексных амплитуд через Δt показана на рис. 284, б. При даль-

● Концентрация лазерного излучения в одном направлении в резонаторе с плоскими зеркалами обуславливается тем, что для этого направления добротность максимальна, а порог генерации минимальен. Расходимость пучка лазерного излучения проистекает главным образом из-за дифракции явлений.



284

Сложение амплитуд при синхронизации мод

нейшем увеличении Δt амплитуда уменьшается, а затем начинает возрастать. Исходная ситуация, когда амплитуды складываются в одинаковой фазе, повторится через промежуток времени T , удовлетворяющий очевидному условию.

$$\Delta\omega T = 2\pi. \quad (53.17)$$

Из (53.17) с учетом (53.7) получаем

$$T = 2\pi/\Delta\omega = 2nL/c, \quad (53.18)$$

т. е. T — продолжительность цикла. Таким образом, ситуация, при которой фазы различных мод оказываются в некоторый момент времени равными друг другу, повторяется через цикл. Согласование фаз различных мод называется синхронизацией мод. В тот момент, когда амплитуды мод складываются в одинаковой фазе, интенсивность лазерного излучения резко возрастает — наблюдается импульс излучения (рис. 285). Импульсы следуют друг за другом с частотой $1/T = c/(2nL)$.

Продолжительность импульса. Она определяется из условия обращения в нуль суммарной амплитуды. Обозначая ΔT — продолжительность импульса, N — число мод, участвующих в процессе синхронизации, условие обращения амплитуды в нуль аналогично (53.17) запишем в виде

$$N\Delta\omega\Delta T = 2\pi, \quad (53.19)$$

откуда

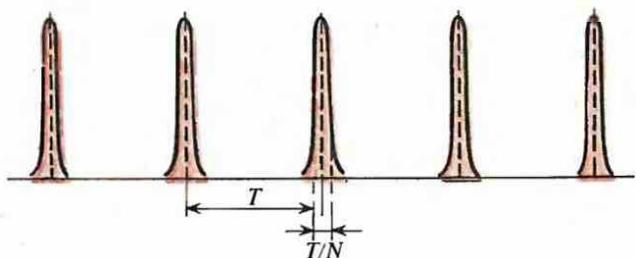
$$\Delta T = \frac{2\pi}{N\Delta\omega} = \left(\frac{2Ln}{c}\right) \frac{1}{N} = \frac{T}{N}. \quad (53.20)$$

Продолжительность импульса тем меньше, чем больше число синхронизированных мод. Например, для $n \approx 1$, $L = 10^{-1}$ м, $N = 10^3$ получаем $\Delta T \approx 0,6 \cdot 10^{-12}$ с. Методом синхронизации мод удается получить рекордно короткие импульсы, применение которых для изучения быстропротекающих процессов очень эффективно.

Мощность излучения в импульсе может быть оценена из закона сохранения энергии. Благодаря синхронизации мод практически вся энергия излучения, приходящаяся на промежуток времени $T = 2Ln/c$, испускается в импульсе продолжительностью $\Delta T = T/N$. Это означает, что мощность излучения в импульсе увеличивается в $T/\Delta T$ раз по сравнению со средней мощностью.

Осуществление синхронизации мод. Синхронизация мод может возникать самопроизвольно. В этом случае синхронизуется обычно небольшое число мод и поэтому продолжительность импульсов мало отличается от периода их повторений (продолжительности цикла). Для уменьшения продолжительности импульсов применяются различные методы, в частности модуля-

- Какие характеристики лазерного излучения обусловливаются аксиальными и боковыми модами?
- Чем определяется ширина линии излучения в лазере и каков теоретический предел относительной ширины линии излучения?
- Какие относительные ширины линий достигнуты в настоящее время?
- Как в лазерах осуществляется синхронизация мод?

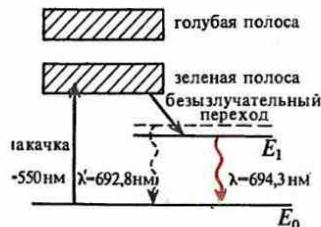


ции добротности с периодом цикла $2L n/c$. Для этого используются нелинейные эффекты, играющие в лазерах большую роль. С учетом нелинейности между различными модами имеется связь и их нельзя рассматривать как независимые. Синхронизация мод перестает быть полностью случайным процессом, и имеется возможность на нее воздействовать. В качестве примера рассмотрим генерацию лазерного излучения, когда в резонатор введен фильтр, коэффициент пропускания которого увеличивается с увеличением интенсивности проходящего через него света. Фильтр действует практически безынерционно. Если в результате случайной синхронизации небольшого числа мод произошло усиление интенсивности, то при прохождении через фильтр эти синхронизованные моды будут меньше всего ослаблены и получат преимущество по сравнению с другими модами, которые проходя фильтр в другие моменты времени при меньшей интенсивности поля и, следовательно, сильно ослабляются. При последующих прохождениях фильтра синхронизованные моды создают условия относительного усиления других мод, которые оказались случайно синхронизованными с уже существующими. Таким образом, усиливаются лишь моды, находящиеся в режиме синхронизации или впервые случайно попадающие в режим синхронизации. Благодаря этому увеличивается число мод, находящихся в режиме синхронизации, продолжительность импульса уменьшается, а его мощность соответствующим образом увеличивается. В настоящее время получены импульсы, продолжительность которых существенно меньше 10^{-12} с.

Лазерные спеклы. Поверхность большинства предметов, освещенных лучом лазера, представляется пятнистой. Пятна распределяются по поверхности случайно, их можно сфотографировать. Если фотоаппарат сфокусирован на точки до поверхности или за поверхностью, то на фотографии все равно получается изображение пятен. Если при наблюдении пятен глазом попытаться аккомодировать глаз на точки до поверхности или за поверхностью, зрительные впечатления о наличии пятен не изменятся. Если, глядя на поверхность, наблюдатель движется, то одни пятна исчезают, другие появляются — создается впечатление, что пятна мигают и движутся относительно поверхности. Эти пятна принято называть спеклами.

Их возникновение обусловлено большой степенью когерентности лазерного излучения. Большинство поверхностей, не отполированных специально с высокой оптической точностью, имеют случайные неровности, высота которых больше длины волны. Лазерное излучение с большой степенью когерентности отражается диффузно от поверхности. От различных точек поверхности распространяются волны с постоянными разностями фаз. При попадании на сетчатку глаза или на фотопластинку образуется интерференционная картина в виде чередующихся темных и светлых спеклов.

Если поверхность отполирована с оптической точностью до долей длины волны, то происходит зеркальное отражение лазерного излучения, не вносящее разности хода между различными лучами в пучке, и никаких спеклов не наблюдается. Не наблюдаются спеклы также и при диффузном отражении от поверхности жидкости, когда ее неровности с течением времени меняются и, следовательно, меняются разности хода лучей, отраженных от различных точек поверхности,



286

Схема энергетических уровней рубина

в результате чего происходит усреднение интенсивности интерференционной картины по времени.

С физической точки зрения возникновение спеклов является шумовым эффектом. Этот эффект может иметь полезные применения. С его помощью можно измерять малые смещения твердой поверхности, например в направлении, параллельном поверхности. Для этого делаются две последовательные фотографии спеклов. Если между моментами фотографирования поверхности не смешалась, то спеклы на обоих снимках накладываются друг на друга. При наличии смещения на фотографии видны две совершенно одинаковые картины спеклов, сдвинутые друг относительно друга. Зная расположение фотоаппарата относительно поверхности в моменты фотографирования, нетрудно по относительному сдвигу изображений спеклов на фотографии рассчитать перемещение поверхности. Аналогичным способом можно определить поворот поверхности вокруг некоторой оси. Известны также другие применения спеклов.

Спеклы возникают не только при диффузном отражении от поверхности, но и при прохождении света через рассеивающий объект, поскольку разности фаз волн, приходящих в данную точку от различных рассеивателей, неизменны по времени. Образующиеся при этом спеклы также имеют многие применения.

§ 54 | Характеристики некоторых лазеров

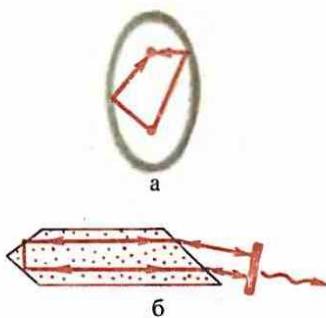
Даются краткие характеристики наиболее типичных лазеров.

Разнообразие лазеров. В настоящее время имеется громадное разнообразие лазеров, отличающихся между собой активными средами, мощностями, режимами работы и другими характеристиками. Нет необходимости все их описывать. Поэтому здесь дается краткое описание лазеров, которые достаточно полно представляют характеристики основных типов лазеров (режим работы, способы накачки и т. д.).

Рубиновый лазер. Первым квантовым генератором света был рубиновый лазер, созданный в 1960 г.

Рабочим веществом является рубин, представляющий собой кристалл оксида алюминия Al_2O_3 (корунд), в который при выращивании введен в виде примеси оксид хрома Cr_2O_3 . Красный цвет рубина обусловлен положительным ионом Cr^{+3} . В решете кристалла Al_2O_3 ион Cr^{+3} замещает ион Al^{+3} . Вследствие этого в кристалле возникают две полосы поглощения: одна — в зеленой, другая — в голубой части спектра (рис. 286). Густота красного цвета рубина зависит от концентрации ионов Cr^{3+} : чем больше концентрация, тем гуще красный цвет. В темно-красном рубине концентрация ионов Cr^{3+} достигает 1%.

Наряду с голубой и зеленой полосами поглощения имеется два узких энергетических уровня E_1 и E'_1 , при переходе с которых на основной уровень излучается свет с длинами волн 694,3 и 692,8 нм. Ширина линий составляет при комнатных темпера-



287

Схема накачки (а) и хода лучей (б) в рубиновом лазере

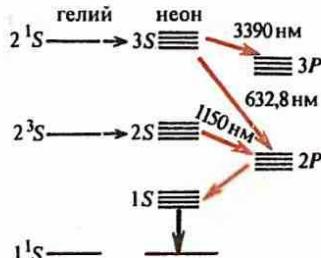
турах примерно 0,4 нм. Вероятность вынужденных переходов для линии 694,3 нм больше, чем для 692,8 нм, ввиду того, что $B \sim \omega^{-3}$. Поэтому проще работать с линией 694,3 нм. Однако можно осуществить генерацию и линии 692,8 нм, если использовать специальные зеркала, имеющие большой коэффициент отражения для излучения с $\lambda = 692,8$ нм и малый — для $\lambda = 694,3$ нм.

Схема энергетических уровней рубина показана на рис. 286. При облучении рубина белым светом голубая и зеленая части спектра поглощаются, а красная отражается. В рубиновом лазере используется оптическая накачка ксеноновой лампой, которая дает вспышки света большой интенсивности при прохождении через нее импульса тока, нагревающего газ до нескольких тысяч кельвин. Непрерывная накачка невозможна, потому что лампа при столь высокой температуре не выдерживает непрерывного режима работы. Возникающее излучение близко по своим характеристикам к излучению абсолютно черного тела. Излучение поглощается ионами Cr^+ , переходящими в результате этого на энергетические уровни в области полос поглощения. Однако с этих уровней ионы Cr^+ очень быстро в результате безызлучательного перехода переходят на уровни E_1 , E'_1 (рис. 286). При этом излишек энергии передается решетке, т. е. превращается в энергию колебаний решетки или, другими словами, в энергию фононов. Уровни E_1 и E'_1 метастабильны. Время жизни на уровне E_1 равно 4,3 мс. В процессе импульса накачки на уровнях E_1 и E'_1 накапливаются возбужденные атомы, создающие значительную инверсионную заселенность относительно уровня E_0 .

Кристалл рубина выращивается в виде круглого цилиндра. Для лазера обычно используют кристаллы размером: длина $L \approx 5$ см, диаметр $d \approx 1$ см. Ксеноновая лампа и кристалл рубина помещаются в эллиптическую полость с хорошо отражающей внутренней поверхностью (рис. 287, а). Чтобы обеспечить попадание на рубин всего излучения ксеноновой лампы, кристалл рубина и лампа, имеющая также форму круглого цилиндра, помещаются в фокусы эллиптического сечения полости параллельно ее образующим. Благодаря этому на рубин направляется излучение с плотностью, практически равной плотности излучения на источнике накачки.

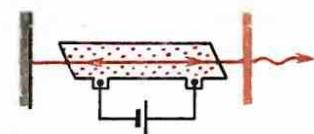
Один из концов рубинового кристалла срезан так (рис. 287, б), что от граней среза обеспечивается полное отражение и возвращение луча обратно. Такой срез заменяет одно из зеркал лазера. Второй конец рубинового кристалла срезан под углом Брюстера. Он обеспечивает выход из кристалла рубина без отражения луча с соответствующей линейной поляризацией. Второе зеркало резонатора ставится на пути этого луча.

Таким образом, излучение рубинового лазера линейно поляризовано.



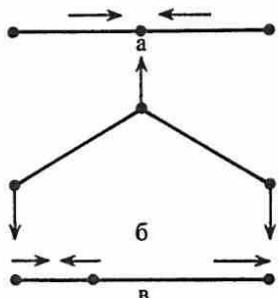
288

Схема энергетических уровней гелия и неона



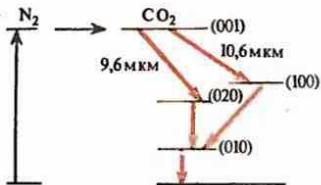
289

Схема гелий-неонового лазера



290

Моды колебаний молекулы CO_2 : симметричное растяжение (а); изгиб (б); несимметричное растяжение (в)



291

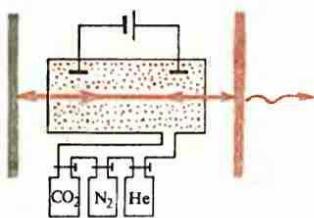
Схема энергетических уровней в CO_2 -лазере

Гелий-неоновый лазер. Активной средой является газообразная смесь гелия и неона. Генерация осуществляется за счет переходов между энергетическими уровнями неона, а гелий играет роль посредника, через который энергия передается атомам неона для создания инверсной заселенности.

Неон, в принципе, может генерировать лазерное излучение в результате более 130 различных переходов. Однако наиболее интенсивными являются линии с 632,8 нм, 1,15 и 3,39 мкм (рис. 288). Волна 632,8 нм находится в видимой части спектра, а 1,15 и 3,39 мкм — в инфракрасной.

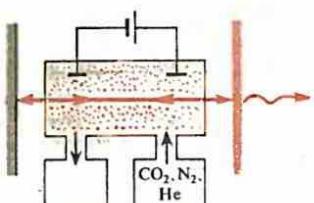
При пропускании тока через гелий-неоновую смесь газов электронным ударом атомы гелия возбуждаются до состояний 2^3S и 2^2S , которые являются метастабильными, поскольку переход в основное состояние из них запрещен квантово-механическими правилами отбора. При прохождении тока атомы накапливаются на этих уровнях. Когда возбужденный атом гелия сталкивается с невозбужденным атомом неона, энергия возбуждения переходит к последнему. Этот переход осуществляется очень эффективно вследствие хорошего совпадения энергии соответствующих уровней. Вследствие этого на уровнях $3S$ и $2S$ неона образуется инверсная заселенность относительно уровней $3P$ и $2P$, приводящая к возможности генерации лазерного излучения. Лазер может оперировать в непрерывном режиме. Типичная схема гелий-неонового лазера показана на рис. 289. Концы лазерной трубки закрыты соответствующим прозрачным материалом так, чтобы аксиальные моды падали на него под углом Брюстера. Благодаря этому обеспечивается полное пропускание одной из поляризаций света и устранение из лучка другой. Излучение гелий-неонового лазера линейно поляризовано. Обычно давление гелия в камере составляет 332 Па, а неона — 66 Па. Постоянное напряжение на трубке около 4 кВ. Одно из зеркал имеет коэффициент отражения порядка 0,999, а второе, через которое выходит лазерное излучение, — около 0,990. В качестве зеркал используют многослойные диэлектрики (см. § 29), поскольку более низкие коэффициенты отражения не обеспечивают достижения порога генерации.

CO_2 -лазер с замкнутым объемом. Молекулы углекислого газа, как и другие молекулы, имеют полосатый спектр, обусловленный наличием колебательных и вращательных уровней энергии. Молекула CO_2 является линейной с центром симметрии. Она имеет три фундаментальные моды колебаний (рис. 290). Энергия квантов фундаментальных мод колебаний равна: а) $1/\lambda_1 = 1337 \text{ cm}^{-1}$; б) $1/\lambda_2 = 667 \text{ cm}^{-1}$; в) $1/\lambda_3 = 2349 \text{ cm}^{-1}$. В каждой моде может быть один или несколько квантов. Колебательные состояния молекулы обозначаются количеством квантов в соответствующей фундаментальной моде колебаний. Например, (010) означает, что симметричные и антисимметрич-



292

Схема закрытого CO_2 -лазера



293

Схема проточного CO_2 -лазера

ные колебания не возбуждены, а в изгибной моде имеется один квант. Используемый в CO₂-лазере переход дает излучение с длиной волны 10,6 мкм, т. е. лежит в инфракрасной области спектра. Пользуясь колебательными уровнями, можно несколько варьировать частоту излучения в пределах примерно от 9,2 до 10,8 мкм. Энергия молекулам CO₂ передается от молекул азота N₂, которые сами возбуждаются электронным ударом при прохождении тока через смесь. Схема уровней показана на рис. 291.

Возбужденное состояние молекулы азота N₂ является метастабильным и отстоит от основного уровня на расстоянии 2318 см⁻¹, что весьма близко к энергетическому уровню (001) молекулы CO₂. Ввиду метастабильности возбужденного состояния N₂ при прохождении тока число возбужденных атомов накапливается. При столкновении N₂ с CO₂ происходит резонансная передача энергии возбуждения от N₂ к CO₂. Вследствие этого возникает инверсия заселенностей между уровнями (001), (100), (020) молекул CO₂. Обычно для уменьшения заселенности уровня (100), который имеет большое время жизни, что ухудшает генерацию при переходе на этот уровень, добавляют гелий. В типичных условиях смесь газов в лазере состоит из гелия (1330 Па), азота (133 Па) и углекислого газа (133 Па).

При работе CO₂-лазера происходит распад молекул CO₂ на CO и O, благодаря чему активная среда ослабляется. Далее CO распадается на C и O, а углерод осаждается на электродах и стенах трубки. Все это ухудшает работу CO₂-лазера. Чтобы преодолеть вредное действие этих факторов, в закрытую систему добавляют пары воды, которые стимулируют реакцию CO + O → CO₂. Используются платиновые электроды, материал которых является катализатором для этой реакции. Для увеличения запаса активной среды резонатор соединяется с дополнительными емкостями, содержащими CO₂, N₂, He, которые в необходимом количестве добавляются в объем резонатора для поддержания оптимальных условий работы лазера (рис. 292). Такой закрытый CO₂-лазер, в состоянии работать в течение многих тысяч часов.

Проточный CO₂-лазер. Важной модификацией является прочный CO₂-лазер, в котором смесь газов CO₂, N₂, He непрерывно прокачивается через резонатор (рис. 293) в аксиальном направлении. Такой лазер может генерировать непрерывное когерентное излучение мощностью свыше 50 Вт на метр длины своей активной среды.

T-лазер. Во многих практических приложениях важную роль играет CO₂-лазер, в котором рабочая смесь находится под атмосферным давлением и возбуждается поперечным электрическим полем (T-лазер; рис. 294). Поскольку электроды расположены параллельно оси резонатора, для получения больших значений напряженности электрического поля в резонаторе требуются сравнительно небольшие разности потенциалов между электродами, что дает возможность работать

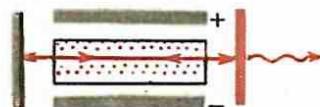
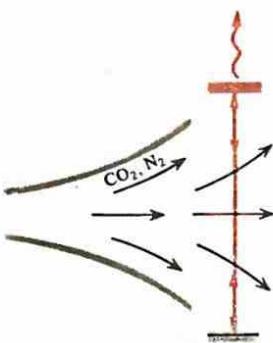


Схема Т-лазера, оперирующего при атмосферном давлении

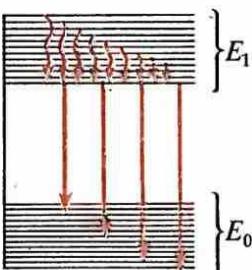
Селекция мод в лазерах осуществляется посредством изменения добротности резонатора для колебаний, соответствующих различным модам. Это дает возможность регулировать число и характеристики мод в излучении лазера и изменять частоту излучения, если селекция мод возможна в достаточно широком интервале частот.

○ Почему T-лазер может успешно работать при атмосферном давлении рабочей смеси? Какой механизм создания инверсной заселенности используется в газодинамических лазерах? Каким методом осуществляется селекция мод в лазерах на красителях?



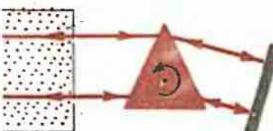
295

Схема газодинамического лазера



296

Схема энергетических уровней молекулы красителя



297

Метод отбора частот для генерации в лазере на красителе

в импульсном режиме при атмосферном давлении, когда концентрация CO_2 в резонаторе велика. Следовательно, удается получить большую мощность, достигающую обычно 10 МВт и больше в одном импульсе излучения продолжительностью менее 1 мкс. Частота повторения импульсов в таких лазерах составляет обычно несколько импульсов в минуту.

Газодинамические лазеры. Нагретая до высокой температуры (1000—2000 К) смесь CO_2 и N_2 при истечении с большой скоростью через расширяющееся сопло сильно охлаждается. Верхний и нижний энергетический уровни при этом термализуются с различной скоростью, в результате чего образуется инверсная заселенность. Следовательно, образовав на выходе из сопла оптический резонатор, можно за счет этой инверсной заселенности генерировать лазерное излучение (рис. 295). Действующие на этом принципе лазеры называются газодинамическими. Они позволяют получать очень большие мощности излучения в непрерывном режиме.

Лазеры на красителях. Красители являются очень сложными молекулами, у которых сильно выражены колебательные уровни энергии. Энергетические уровни в полосе спектра располагаются почти непрерывно (рис. 296). Вследствие внутримолекулярного взаимодействия молекула очень быстро (за времена порядка $10^{-11} — 10^{-12}$ с) переходит безызлучательно на нижний энергетический уровень каждой полосы. Поэтому после возбуждения молекул через очень короткий промежуток времени на нижнем уровне полосы E_1 сосредоточится все возбужденные молекулы. Они далее имеют возможность совершить излучательный переход на любой из энергетических уровней нижней полосы (рис. 296). Таким образом, возможно излучение практически любой частоты в интервале, соответствующем ширине нулевой полосы. А это означает, что если молекулы красителя взять в качестве активного вещества для генерации лазерного излучения, то в зависимости от настройки резонатора можно получить практически непрерывную перестройку частоты генерируемого лазерного излучения. Поэтому на красителях создаются лазеры с перестраиваемой частотой генерации. Накачка лазеров на красителях производится газоразрядными лампами или излучением других лазеров.

Выделение частот генерации достигается тем, что порог генерации создается только для узкой области частот. Например, положения призмы и зеркала (рис. 297) подбираются так, что в среду после отражения от зеркала благодаря дисперсии и разным углам преломления возвращаются лишь лучи с определенной длиной волн. Только для таких длин волн обеспечивается лазерная генерация. Вращая призму, можно обеспечить непрерывную перестройку частоты излучения лазера на красителях. Генерация осуществлена со многими красителями, что позволило получить лазерное излучение не только во всем оптическом диапазоне, но и на значительной части инфракрасной и ультрафиолетовой областей спектра.

10

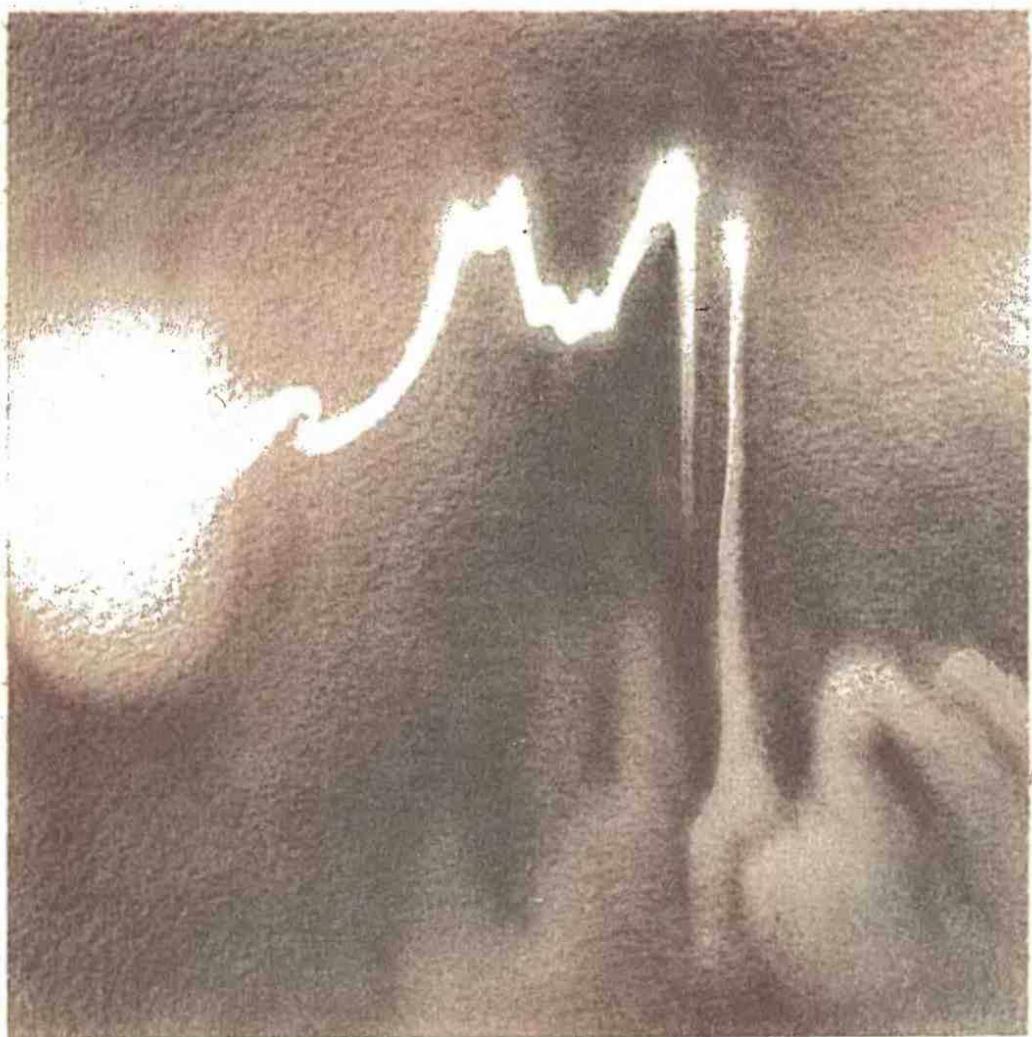
- 10.1** Имеются две шаровые полости большого радиуса с малыми круглыми отверстиями одинаковых диаметров $d = 1$ см и абсолютно отражающими наружными поверхностями. Плоскости круглых отверстий параллельны друг другу, и их центры лежат на прямой, перпендикулярной этим плоскостям. Расстояние между отверстиями $l = 10$ см. В одной из полостей поддерживается постоянная температура $T_0 = 2000$ К. Вычислить установившуюся температуру во второй полости.
- 10.2.** Зная, что Солнце излучает как абсолютно черное тело с температурой 6000 К, а его диаметр виден с Земли под углом $\Delta\theta_0 = 32'$, определить плотность потока энергии солнечного излучения на земной орбите.
- 10.3.** Найти температуру поверхности Сириуса, если известно, что максимум плотности потока его энергии излучения по шкале длин волн приходится на $\lambda_{\max} = 0,29$ мкм.
- 10.4.** Мощность излучения абсолютно черного тела, имеющего поверхность 10m^2 , равна 1 МВт. Найти температуру тела.
- 10.5.** Мощность излучения абсолютно черного тела 100 МВт, а максимальная плотность излучения приходится на волну 0,6 мкм. Найти площадь излучающей поверхности тела.
- 10.6.** Раскаленная до 2900 К ламбертовская поверхность излучает 200 Вт с 1 см². Найти отношение энергетической светимости этой поверхности к энергетической светимости поверхности абсолютно черного тела.
- 10.7.** Поток энергии в импульсе излучения лазера равен 10^{11} Вт/см². Найти амплитуду напряженности электрического поля в волне.
- 10.8.** Какой температуре абсолютно черного тела соответствует излучение рубинового лазера на волне $\lambda = 694,3$ нм при спектральной объемной плотности излучения на этой частоте $w_\omega = 10$ Дж·с·м⁻³?

Ответы

- 10.1.** $T_0 \pi^{1/4} [d/(2l)]^{1/2} = 595$ К. **10.2.** 1,4 кВт/м². **10.3.** 10^4 К. **10.4.** 1150 К. **10.5.** 3,2 м². **10.6.** 0,5. **10.7.** $6,13 \times 10^8$ В/м. **10.8.** $9,3 \cdot 10^{16}$ К.

Основная идея:
при достаточно больших
напряженностях
электрического поля
волны проявляется
нелинейный характер
зависимости
поляризованности
от напряженности.

Нелинейные явления в оптике



Описывается связь между поляризованностью и напряженностью электрического поля волны с учетом нелинейных членов.

Линейная поляризованность. Теория линейной поляризованности, устанавливающая зависимость показателя преломления от частоты, изложена в § 15. Поляризованность связана с напряженностью электрического поля соотношением

$$P = \epsilon_0 \alpha E, \quad (55.1)$$

где α — диэлектрическая восприимчивость. Сравнение (55.1) с (15.13) приводит к следующему выражению комплексной диэлектрической восприимчивости:

$$\alpha = \frac{e^2 N}{m \epsilon_0} \frac{1}{\omega_0^2 - \omega^2 + i\gamma\omega}. \quad (55.2)$$

Нелинейная поляризованность. Удерживающая электрон около положения равновесия сила $f(x) = -m\omega_0^2 x$ (см. (9.32)) подчиняется закону Гука лишь при не слишком больших x . При больших x наблюдаются отступления от закона Гука и колебания становятся нелинейными.

Функция $f(x)$ в общем случае может быть представлена в виде ряда Тейлора:

$$f(x) = f(0) + xf'(0) + \frac{x^2}{2!}f''(0) + \frac{x^3}{3!}f'''(0) + \dots. \quad (55.3)$$

Вместо (9.32) для описания движения электрона в поле световой волны с учетом нелинейности можно написать уравнение

$$m\ddot{x} + m\gamma\dot{x} = eE + f(0) + xf'(0) + \frac{x^2}{2!}f''(0) + \frac{x^3}{3!}f'''(0) + \dots. \quad (55.4)$$

Так как точка $x=0$ является равновесной, то $f(0)=0$. Сила f всегда направлена к точке равновесия и, следовательно, $f'(0) < 0$. Полагая $f'(0) = -m\omega_0^2$, перепишем (55.4) в виде

$$m\ddot{x} + m\gamma\dot{x} + m\omega_0^2 x = eE + \frac{x^2}{2!}f''(0) + \frac{x^3}{3!}f'''(0) + \dots. \quad (55.5)$$

Если в (55.5) можно пренебречь членами, квадратичными, кубическими и т. д. по x , то приходим к уравнению движения (9.32) линейного осциллятора. При учете этих членов осциллятор называется ангармоническим, а его колебания — ангармоническими колебаниями. Ясно, что для ангармонических колебаний зависимость $x(E)$, в линейном случае выраженная формулой (9.36), усложняется и не будет линейной. Поэтому поляризованность [см. (15.13)], которую удобно представить в виде

$$P = N|e|x, \quad (55.6)$$

перестает быть линейной функцией от E .

Квадратичная нелинейность. Если $f''(0) \neq 0$, то нелинейность колебаний проявляется в первую очередь за счет квадратичного по x члена. Уравнение (55.5) записывают в виде

$$\ddot{x} + \gamma\dot{x} + \omega_0^2 x = (e/m)E + \xi x^2, \quad (55.7)$$

где $\xi = f''(0)/(2m)$. Поскольку величина ξx^2 предполагается малой, решать уравнение (55.7) следует методом возмущений. Представим искомое решение в виде ряда

$$x = x_0 + x_1 + x_2 + \dots, \quad (55.8)$$

где x_1, x_2, \dots — малые величины порядка ξ, ξ^2, \dots относительно x_0 . Подставив (55.8) в (55.7) и приравняв между собой члены одинакового порядка по ξ , получим уравнения

$$\ddot{x}_0 + \gamma \dot{x}_0 + \omega_0^2 x_0 = (e/m) E,$$
(55.9)

$$\ddot{x}_1 + \gamma \dot{x}_1 + \omega_1^2 x_1 = \xi x_0^2,$$
(55.10)

где точками обозначены уравнения для x_2, x_3, \dots которые пока нами не рассматриваются.

В линейном случае (55.1) рассматривалась одна волна частотой ω , поскольку добавление волны другой частоты ничего не добавляло к картине образования поляризованности: поляризованность от двух волн равна сумме поляризованностей от каждой из волн. При учете нелинейности ситуация меняется. Из (55.9) видно, что x_0 выражается линейно через E , а из (55.10) следует, что x_1 зависит от x_0^2 и, следовательно, от E^2 . Если E выражается в виде суммы напряженностей полей с различными частотами, то x_1 зависит от попарных произведений этих напряженностей, т. е. является квадратичной функцией напряженностей. Поляризованность в соответствии с (55.6) описывается формулой

$$P = N|e|x_0 + N|e|x_1 = P_{\text{л}} + P_{\text{н}}^{(2)}, \quad (55.11)$$

$$P_{\text{л}} = N|e|x_0 = \epsilon_0 \alpha^{(1)} E \quad (55.12) \quad \text{и} \quad P_{\text{н}}^{(2)} = N|e|x_1 = \epsilon_0 \alpha^{(2)} E^2 \quad (55.13)$$

— линейная и квадратичная нелинейная поляризованности.

Если решать уравнения, обозначенные после (55.10) точками, то для x_2 получим решение, зависящее от E^3 , приводящее к поляризованности $P_{\text{н}}^{(3)}$, пропорциональной E^3 , и т. д. Символически этот результат запишется в виде

$$P = \epsilon_0 \alpha^{(1)} E + \epsilon_0 \alpha^{(2)} EE + \epsilon_0 \alpha^{(3)} EEE + \dots; \quad (55.14)$$

поляризованность представляется суммой членов, зависящих линейно, квадратично, кубично и т. д. от напряженности электрического поля.

Нелинейная восприимчивость. Чтобы найти нелинейную квадратичную поляризованность, необходимо решить уравнения (55.9) и (55.10). Будем считать, что E является суперпозицией нескольких гармонических вещественных полей с частотами $\omega_1, \omega_2, \omega_3, \dots$:

$$E = E(\omega_1)e^{i\omega_1 t} + E^*(\omega_1)e^{-i\omega_1 t} + E(\omega_2)e^{i\omega_2 t} + E^*(\omega_2)e^{-i\omega_2 t} + \dots + E(\omega_n)e^{i\omega_n t} + E^*(\omega_n)e^{-i\omega_n t}, \quad (55.15)$$

где пары комплексно-сопряженных членов одинаковых частот описывают вещественные поля соответствующей частоты. Для того чтобы при вычислениях сделать запись достаточно компактной, целесообразно воспользоваться специальными обозначениями. Комплексное число

$$E(\omega_m)e^{i\omega_m t} + E^*(\omega_m)e^{-i\omega_m t} \quad (55.16)$$

вещественно, если $E^*(\omega_m) = E(-\omega_m)$ [см. (8.35)]. Поэтому, вводя обозначение

$$\omega_{-m} = -\omega_m, \quad (55.17)$$

представим сумму (55.15) в виде

● Влияние среды на излучение сводится к учету переменной поляризованности среды под влиянием излучения, в результате чего среда сама становится источником излучения, складывающегося с первоначальным в соответствии с принципом суперпозиции. Нелинейные явления в сильных электромагнитных волнах возникают не в результате принципа суперпозиции для электромагнитного поля, а в результате нелинейности влияния электромагнитного поля на поляризованность среды.

○ Какие физические факторы приводят к возникновению нелинейной поляризованности?

По какому правилу образуются комбинационные частоты?

$$E = \sum_{m=-n}^n E(\omega_m) e^{i\omega_m t}. \quad (55.18)$$

Подставляя (55.18) в (55.9) и принимая во внимание, что для каждого члена суммы в (55.18) решение имеет вид, аналогичный (15.12), находим

$$x_0 = \frac{e}{m} \sum_{m=-n}^n \frac{E(\omega_m) e^{i\omega_m t}}{\omega_0^2 - \omega_m^2 + i\gamma\omega_m}. \quad (55.19)$$

Отсюда линейная поляризованность

$$P_L = N |e| x_0 = \epsilon_0 \sum_{m=-n}^n \alpha^{(1)}(\omega_m) E(\omega_m) e^{i\omega_m t}, \quad (55.20)$$

где

$$\alpha^{(1)}(\omega_m) = \frac{e^2 N}{m \epsilon_0} \frac{1}{\omega_0^2 - \omega_m^2 + i\gamma\omega_m} \quad (55.21)$$

— линейная диэлектрическая восприимчивость, для каждой частоты совпадающая с (15.13в). Подставляя (55.19) в (55.10), приходим к уравнению для x_1 :

$$\ddot{x}_1 + \gamma \dot{x}_1 + \omega_0^2 x_1 = \frac{\xi e^2}{m^2} \sum_{k=-n}^{k=n} \frac{E(\omega_m) E(\omega_k) e^{i(\omega_m + \omega_k)t}}{(\omega_0^2 - \omega_m^2 + i\gamma\omega_m)(\omega_0^2 - \omega_k^2 + i\gamma\omega_k)}, \quad (55.22)$$

которое решается аналогично уравнению (55.9) для x_0 :

$$x_1 = \frac{\xi e^2}{m^2} \sum_{m,k} \frac{E(\omega_m) E(\omega_k) e^{i(\omega_m + \omega_k)t}}{(\omega_0^2 - \omega_m^2 + i\gamma\omega_m)(\omega_0^2 - \omega_k^2 + i\gamma\omega_k)} \times \frac{1}{[\omega_0^2 - (\omega_m + \omega_k)^2 + i\gamma(\omega_m + \omega_k)]}, \quad (55.23)$$

причем пределы суммирования по m и k прежние и для упрощения написания формул не выписываются. Отсюда нелинейная восприимчивость

$$P_L^{(2)} = N |e| x_1 = \epsilon_0 \Sigma \alpha^{(2)}(\omega_m, \omega_k) E(\omega_m) E(\omega_k) e^{i(\omega_m + \omega_k)t}, \quad (55.24)$$

где

$$\alpha^{(2)}(\omega_m, \omega_k) = \frac{\xi_m \epsilon_0^2}{N^2 e^3} \alpha^{(1)}(\omega_m) \alpha^{(1)}(\omega_k) \alpha^{(1)}(\omega_m + \omega_k) \quad (55.25)$$

— восприимчивость второго порядка, описывающая нелинейную квадратичную поляризованность.

Комбинационные частоты. Из (55.24) видно, что квадратичная поляризованность содержит члены с любыми комбинационными частотами $\omega_m + \omega_k$. Например, для двух полей с частотами ω_1 и ω_2 она зависит от частот $\omega_1 + \omega_2$, $|\omega_1 - \omega_2|$, $2\omega_1$, $2\omega_2$.

При расчете поляризованности более высоких порядков получаем аналогичные выражения. Поляризованность зависит от комбинации соответствующего числа частот. Например, кубическая поляризованность зависит от комбинационных частот $\omega_m + \omega_k + \omega_p$ при всевозможных значениях индексов m, k, p .

§ 56 Генерация гармоник

Рассматриваются генерация гармоник и условия векторного и пространственного синхронизма.

Волна линейной поляризованности. Напряженность электрического поля волны, распространяющейся в среде, создает в точках среды поляризованность, распространяющуюся в пространстве в виде волны поляризованности. В каждой точке среды изменяющаяся поляризованность по-

рождает вторичную электромагнитную волну, которая складывается с волной, породившей поляризованность. Суммарная волна сама является источником поляризованности, которая, в свою очередь, порождает электромагнитную волну, вызывающую поляризованность, и так до бесконечности. Другими словами, волна поляризованности и электромагнитная волна взаимно обусловливают друг друга.

Выведем уравнение электромагнитной волны, учитывающее эту связь. В отсутствие источников ($j = 0, \rho = 0$) уравнения Максвелла имеют вид:

$$\operatorname{rot} \mathbf{H} = \partial \mathbf{D} / \partial t, \quad (56.1)$$

$$\operatorname{rot} \mathbf{E} = -\partial \mathbf{B} / \partial t = -\mu_0 \partial \mathbf{H} / \partial t; \quad (56.2)$$

среда предполагается немагнитной ($\mu = \mu_0$). Продифференцировав обе части уравнения (56.1) по t , получим

$$\operatorname{rot} (\partial \mathbf{H} / \partial t) = \partial^2 \mathbf{D} / \partial t^2. \quad (56.3)$$

Заменив $\partial \mathbf{H} / \partial t$ в (56.3) его выражением из (56.2), найдем уравнение

$$\operatorname{rot} \operatorname{rot} \mathbf{E} = -\mu_0 \partial^2 \mathbf{D} / \partial t^2, \quad (56.4)$$

которое с учетом равенства

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \quad (56.5)$$

принимает вид

$$\operatorname{rot} \operatorname{rot} \mathbf{E} + (1/c^2) \partial^2 \mathbf{E} / \partial t^2 + \mu_0 \partial^2 \mathbf{P} / \partial t^2 = 0, \quad (56.6)$$

где $c^2 = 1/(\epsilon_0 \mu_0)$. Уравнение (56.6) описывает распространение электромагнитной волны с учетом ее связи с поляризованностью.

При учете только линейной поляризованности уравнение (56.6) является линейным. В этом случае взаимообусловленность волны поляризованности и электромагнитной волны приводит к изменению скорости распространения электромагнитной волны, в результате чего волна поляризованности и электромагнитная волна распространяются с одной и той же фазовой скоростью в одинаковой фазе. Если, например, электромагнитная волна распространяется в направлении положительных значений z , то и волна поляризованности распространяется в том же направлении и с той же фазовой скоростью $c/n(\omega)$, где $n(\omega)$ — зависящий от частоты показатель преломления. Эти волны могут быть представлены в виде

$$\mathbf{E} = E_0 \cos \omega (t - zn/c), \quad (56.7)$$

$$\mathbf{P} = \epsilon_0 \alpha^1 E_0 \cos \omega (t - zn/c). \quad (56.8)$$

В каждой точке среды в результате изменения поляризованности порождается электромагнитная волна, амплитуда которой пропорциональна (56.8). Вторичные волны, возбуждаемые поляризованностью (56.8) в точках z' и z'' и затем распространяющиеся в направлении положительных значений z , записываются в виде

$$\mathbf{E}'(z, t) = A \cos \omega [t - z'n/c - (z - z')n/c], \quad (56.9a)$$

$$\mathbf{E}''(z, t) = A \cos \omega [t - z''n/c - (z - z'')n/c]. \quad (56.9b)$$

Поскольку

$$t - z'n/c - (z - z')n/c = t - zn/c, \quad t - z''n/c - (z - z'')n/c = t - zn/c,$$

заключаем, что обе волны (56.9a) и (56.9b), возникшие в различных точках, приходят в точку z в одной фазе и взаимно усиливаются. Именно это обстоятельство и обуславливает возможность распространения электромагнитной волны в среде и ее скорость. Можно сказать, что вторичные волны в линейной среде, излученные в различных точках, синхронны между собой.

Волны нелинейной поляризованности. Нелинейная квадратичная поляризованность содержит все возможные комбинационные частоты первичных электромагнитных волн. Следовательно, порождаемые ею вторичные волны имеют те же самые все возможные комбинационные частоты и распространяются с различными скоростями в соответствии с законом дисперсии.

Суперпозиция волн различных частот не представляет интереса, поскольку она не приводит к интерференции. Интерференция может происходить лишь между волнами одной и той же частоты, излученными в различных точках среды. Если в результате интерференции волны усиливаются, то можно говорить о существовании волны соответствующей частоты в среде, т. е. о генерации новой частоты, как о нелинейном эффекте распространения волн в среде. Если же такого усиления нет, то никакой генерации новой частоты не наблюдается, хотя в каждой точке среды эти частоты генерируются. Рассмотрим условия, при которых происходит генерация волн с частотами, отличающимися от частоты первичной электромагнитной волны. Они называются условиями пространственного синхронизма.

Условие пространственного синхронизма. Запишем в явном виде волны поляризованности, порожденные квадратичной нелинейностью поляризованности. Если имеются две первичные электромагнитные волны с частотами ω_1 и ω_2

$$E_1 = E_{01} \cos(\omega_1 t - k_1 z), \quad (56.10)$$

$$E_2 = E_{02} \cos(\omega_2 t - k_2 z), \quad (56.11)$$

то порождаемая ими волна квадратичной поляризованности представляется так:

$$P = 2aE^2 = 2a(E_1 + E_2)^2 = 2a(E_{01}^2 \cos^2(\omega_1 t - k_1 z) + E_{02}^2 \cos^2(\omega_2 t - k_2 z) + 2E_{01}E_{02} \cos(\omega_1 t - k_1 z) \cos(\omega_2 t - k_2 z)), \quad (56.12)$$

где волны взяты в действительном виде, а коэффициент $2a$ описывает поляризованность второго порядка и при необходимости может быть вычислен из (56.25). Учитывая тригонометрические соотношения

$$2 \cos^2 \alpha = 1 + \cos 2\alpha, \quad 2 \cos \alpha \cos \beta = \cos(\alpha + \beta) + \cos(\alpha - \beta),$$

представим (56.12) в виде

$$P = P_0 + P_{2\omega_1} + P_{2\omega_2} + P_{\omega_1 + \omega_2} + P_{\omega_1 - \omega_2}, \quad (56.13)$$

где

$$P_0 = a(E_{01}^2 + E_{02}^2), \quad (56.14)$$

$$P_{2\omega_1} = aE_{01}^2 \cos[2(\omega_1 t - k_1 z)], \quad (56.15)$$

$$P_{2\omega_2} = aE_{02}^2 \cos[2(\omega_2 t - k_2 z)], \quad (56.16)$$

$$P_{\omega_1 + \omega_2} = 2aE_{01}E_{02} \cos[(\omega_1 + \omega_2)t - (k_1 + k_2)z], \quad (56.17a)$$

$$P_{\omega_1 - \omega_2} = 2aE_{01}E_{02} \cos[(\omega_1 - \omega_2)t - (k_1 - k_2)z]. \quad (56.17b)$$

Таким образом, две электромагнитные гармонические волны порождают при наличии квадратичной нелинейности четыре волны поляризованности с частотами $2\omega_1$, $2\omega_2$, $\omega_1 + \omega_2$ и $|\omega_1 - \omega_2|$ и статическую поляризованность P_0 . Возникновение статической поляризованности называется оптическим детектированием по аналогии с детектированием радиосигналов выпрямлением тока, поскольку в продетектированном сигнале содержится постоянная составляющая.

Выход условий пространственного синхронизма проиллюстрируем на примере удвоения частоты ω_1 . Для осуществления удвоения частоты достаточно, чтобы в среде распространялась лишь одна волна с частотой ω_1 , а волна с частотой ω_2 может и отсутствовать. Волна (56.15) может быть представлена в виде

$$P_{2\omega_1}(z, t) = aE_{01}^2 \cos 2\omega_1 [t - zn(\omega_1)/c], \quad (56.18)$$

где $n(\omega_1)/c = k_1/\omega_1$, $n(\omega_1)$ — показатель преломления волны с частотой ω_1 . Порождаемые

волной поляризованности (56.18) в точках z' и z'' электромагнитные волны аналогично (56.9a) и (56.9b) описываются формулами

$$E'_{2\omega_1}(z, t) = A \cos \{2\omega_1 [t - z'n(\omega_1)/c - (z - z')n(2\omega_1)/c]\}, \quad (56.19a)$$

$$E''_{2\omega_1}(z, t) = A \cos \{2\omega_1 [t - z''n(\omega_1)/c - (z - z'')n(2\omega_1)/c]\}, \quad (56.19b)$$

где учтено, что после возникновения этих волн в точках z' и z'' они распространяются со скоростью $c/[n(2\omega_1)]$, отличной от скорости волны поляризованности $c/[n(\omega_1)]$. Учтем, что

$$z'n(\omega_1)/c + (z - z')n(2\omega_1)/c = z'\Delta n/c + zn(2\omega_1)/c,$$

$$z''n(\omega_1)/c + (z - z'')n(2\omega_1)/c = z''\Delta n/c + zn(2\omega_1)/c,$$

где

$$\Delta n = n(\omega_1) - n(2\omega_1), \quad (56.20)$$

и представим (56.19 а, б) в виде

$$E'_{2\omega_1}(z, t) = A \cos \{2\omega_1 [t - zn(2\omega_1)/c - z'\Delta n/c]\}, \quad (56.21a)$$

$$E''_{2\omega_1}(z, t) = A \cos \{2\omega_1 [t - zn(2\omega_1)/c - z''\Delta n/c]\}. \quad (56.21b)$$

Сравнение (56.21a) с (56.21b) показывает, что вторичные волны приходят в точку z в одинаковой фазе и усиливают друг друга лишь в том случае, если

$$\Delta n = 0. \quad (56.22)$$

Это условие с учетом (56.20) записывается в виде

$$n(\omega_1) = n(2\omega_1) \quad (56.23)$$

и называется условием пространственного синхронизма для удвоения частоты.

Впервые генерация второй гармоники была осуществлена в 1961 г.: была удвоена частота излучения рубинового лазера в нелинейном кристалле.

Длина когерентности. Разность фаз в точке z между волнами, описываемыми формулами (56.21a) и (56.21b), равна

$$\Delta\phi = 2\omega_1 \Delta n (z'' - z')/c = 2\omega_1 \Delta n l/c, \quad (56.24)$$

где

$$l = z'' - z' \quad (56.25)$$

— расстояние между точками z'' и z' , в которых эти волны генерированы. Ясно, что при $\Delta\phi = 0$ в результате интерференции модуль амплитуды суммарной волны имеет максимальное значение, равное сумме модулей амплитуд интерферирующих волн. При увеличении $\Delta\phi$ модуль амплитуды суммарной волны уменьшается и обращается в нуль при $\Delta\phi = \pi$. Следовательно, вторичные волны, генерированные на пути l_k , удовлетворяющем в соответствии с (56.24) условию $2\omega_1 \Delta n l_k / c = \pi$, (56.26)

дают в результате суперпозиции волну с отличной от нуля амплитудой. Из (56.26) следует, что

$$l_k = \lambda/(4\Delta n), \quad (56.27)$$

где $\omega_1/c = 2\pi/\lambda$, l_k — длина когерентности (характеризует расстояние, на котором разность фаз вторичных электромагнитных волн изменяется меньше чем на π).

Из (56.24) ясно, что амплитуда суммарной вторичной электромагнитной волны по мере распространения в среде изменяется периодически. Найдем закон этого изменения. Если первичная волна входит в среду в точке $z=0$, то напряженность суммарной вторичной волны в точке z внутри среды равна сумме напряженностей вторичных волн, генерированных на пути от 0 до z . Напряженность от излучателей между z' и $z' + dz'$ на основании (56.21a) дается выражением

$$dE_{2\omega_1}(z, t) = A \cos \{2\omega_1[t - zn(2\omega_1)/c - z'\Delta n/c]\} dz'. \quad (56.28)$$

Отсюда следует, что напряженность вторичной электромагнитной волны в точке z равна

$$E_{2\omega_1}(z, t) = A \int_0^z \cos \{2\omega_1[t - zn(2\omega_1)/c - z'\Delta n/c]\} dz'. \quad (56.29)$$

В результате интегрирования получаем

$$E_{2\omega_1}(z, t) = \frac{Ac}{2\omega_1\Delta n} (\sin 2\omega_1[t - zn(2\omega_1)/c] - \sin 2\omega_1[t - zn(2\omega_1)/c]). \quad (56.30)$$

где в соответствии с (56.20) $n(\omega_1) = n(2\omega_1) + \Delta n$. Воспользовавшись тригонометрической формулой для разности синусов, окончательно находим

$$E_{2\omega_1} = A \frac{\sin(\omega_1\Delta nz/c)}{\omega_1\Delta n/c} \cos 2\omega_1 \left[t - \frac{n(2\omega_1) + n(\omega_1)}{2c} z \right]. \quad (56.31)$$

Эта формула показывает, что вторичная электромагнитная волна имеет удвоенную частоту $2\omega_1$ и изменяющуюся амплитуду. Изменение амплитуды волны (рис. 298) по мере распространения в среде описывается функцией

$$A_{2\omega_1} = A \frac{\sin(\omega_1\Delta nz/c)}{\omega_1\Delta n/c}. \quad (56.32)$$

На пути Δz после входа волны в среду при условии

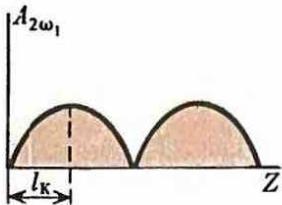
$$\omega_1\Delta n\Delta z/c \ll \pi/2, \quad (56.33)$$

амплитуда увеличивается, поскольку при этом значение синуса в (56.32) увеличивается от нуля до единицы. Путь, при прохождении которого амплитуда достигает максимального значения, на основании (56.33) дается выражением

$$\Delta z = \pi c / (2\omega_1\Delta n) = \lambda / (4\Delta n) = l_k, \quad (56.34)$$

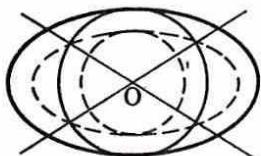
т. е. равен длине когерентности (56.27). Следовательно, если для генерации удвоенной частоты добиться соблюдения условия пространственного синхронизма (55.27) не удается, можно попытаться ее генерировать при прохождении излучения большой мощности через пластинку толщиной, равной длине когерентности. Как показывает опыт, при достаточно большой мощности излучения на выходе из пластиинки наблюдается удвоенная частота. Интенсивность генерации удвоенной частоты очень слабая, поскольку нелинейность среды, как правило, мала. Для увеличения интенсивности излучения необходимо увеличивать длину когерентности, уменьшая Δn . В пределе ($\Delta n \rightarrow 0$) длина когерентности $l_k \rightarrow \infty$, что означает реализацию условия пространственного синхронизма (56.23). Значение Δn фиксировано свойствами среды и не может уменьшаться произвольно. Однако можно подобрать условия в анизотропных средах, при которых

 Физическим содержанием процесса удвоения частоты является перекачка энергии из одной волны в другую с удвоенной частотой. Соблюдение условия пространственного синхронизма необходимо для того, чтобы такая перекачка энергии происходила в одном направлении достаточно долго. При отсутствии пространственного синхронизма перекачка энергии в волну с удвоенной частотой продолжается лишь небольшой промежуток времени, после чего начинается обратная перекачка энергии из волны с удвоенной частотой в волну с основной частотой, и эффект удвоения частоты уменьшается.



298

Изменение амплитуды второй гармоники с расстоянием



299

Определение направлений пространственной синхронизации в однородном кристалле

$\Delta n = 0$. Для этого использовать разность показателей преломления обыкновенного и необыкновенного лучей.

Формулой (56.32) описывается изменение амплитуды волны удвоенной частоты, показанное на рис. 298. Плотность потока энергии волны пропорциональна квадрату этой амплитуды. Следовательно, рассматриваемый процесс представляет собой преобразование энергии волны частоты ω в энергию волны частоты 2ω при росте амплитуды волны удвоенной частоты и обратное преобразование при уменьшении этой амплитуды. Другими словами, это есть процесс обмена энергией между двумя волнами с частотами ω и 2ω , происходящий посредством поляризации среды.

Осуществление пространственного синхронизма. Обыкновенный и необыкновенный лучи в одноосных кристаллах распространяются с различной скоростью (см. § 41). Скорость обыкновенного луча не зависит от направления, а скорость необыкновенного зависит от угла между лучом и оптической осью. В положительных кристаллах скорость обыкновенного луча больше скорости необыкновенного, а в отрицательных — меньше. На рис. 299 изображены сечения лучевых поверхностей отрицательного одноосного кристалла плоскостью, проходящей через оптическую ось. Поскольку частоты светового диапазона, ближней ультрафиолетовой и инфракрасной частей спектра лежат в области нормальной дисперсии, сплошной линией изображены сечения лучевых поверхностей для частоты ω , а пунктирной линией — для частоты 2ω . Видно, что в направлениях, обозначенных двусторонними сплошными стрелками, скорость обыкновенного луча (частота ω) равна скорости необыкновенного луча (частота 2ω), т. е. соответствующие показатели преломления равны

$$n_o(\omega) = n_e(2\omega). \quad (56.35)$$

Это равенство совпадает с условием пространственного синхронизма (56.23). Но отсюда сразу не следует, что можно повторить без изменения все рассуждения, приведшие к формуле (56.23), и утверждать возможность генерации второй гармоники необыкновенной волны. Дело в том, что поляризованности обыкновенной и необыкновенной волн взаимно перпендикулярны и поэтому волна поляризованности, связанная с обыкновенной волной, не может взаимодействовать с необыкновенной волной, и наоборот. Если это так, то нельзя повторять рассуждения, приведшие к условию (56.23), и утверждать, что (56.35) является условием пространственного синхронизма для генерации удвоенной частоты необыкновенной волны. Однако такое утверждение справедливо лишь для линейной поляризованности. Нелинейность делает зависимыми между собой поляризованности, связанные с обыкновенной и необыкновенной волнами. Благодаря этому все рассуждения, приведшие к формуле (56.23), справедливы для обыкновенной и необыкновенной волн и приводят к соотношению (56.35) как условию пространственного

синхронизма для генерации удвоенной частоты необыкновенной волны. Очевидно, что удвоенную частоту обыкновенной волны можно генерировать в положительном кристалле.

Эксперимент подтвердил высокую эффективность генерации второй гармоники в одноосных кристаллах при совпадении лучей с направлением осуществления пространственного синхронизма. Коэффициент преобразования энергии излучения исходной частоты в энергию излучения с удвоенной частотой достигает нескольких десятков процентов. Явление очень эффективно демонстрируется на лекциях. Для этого используется невидимое лазерное излучение с длиной волны 1,06 мкм, которое пропускается через изменяющий свою ориентацию кристалл. В момент совпадения направления пространственного синхронизма кристалла с лучом наблюдается возникновение в кристалле зеленого излучения (длина волны 0,53 мкм) в виде вспышки, если кристалл продолжает изменять свою ориентацию. Если же кристалл зафиксировать в этом положении, то возникающее зеленое излучение является постоянным или импульсным в зависимости от характера лазерного излучения, направляемого на кристалл.

Генерация высших гармоник. В нелинейной поляризованности третьего порядка присутствуют члены с утроенной частотой 3ω падающего излучения. Совершенно аналогично рассмотренному случаю это приводит к генерации третьей гармоники. Можно также генерировать четвертую и дальнейшие гармоники. Это требует совершенствования экспериментальных возможностей, но в принципиальном отношении не содержит в себе новых физических явлений.

Векторное условие пространственного синхронизма. На первый взгляд кажется, что благодаря наличию волн поляризованности (56.17а) и (56.17б) можно аналогично генерации волн с удвоенной частотой добиться генерации волн с частотами $\omega_1 + \omega_2$ и $|\omega_1 - \omega_2|$. Однако это не так. Дело в том, что монохроматические волны в квантовой интерпретации представляются фотонами определенной частоты и импульса. Поэтому образование поляризованностью $P_{\omega_1} + P_{\omega_2}$ электромагнитной волны с частотой $\omega_1 + \omega_2$ можно рассматривать как слияние фотонов с частотами ω_1 и ω_2 . При этом одновременно должны выполняться законы сохранения энергии и импульса. Энергии и импульсы исходных фотонов равны:

$$\epsilon_1 = \hbar\omega_1, \quad p_1 = \hbar k_1 = \hbar\omega_1 n(\omega_1)/c, \quad (56.36)$$

$$\epsilon_2 = \hbar\omega_2, \quad p_2 = \hbar k_2 = \hbar\omega_2 n(\omega_2)/c. \quad (56.37)$$

Если предположить, что электромагнитная волна, порождаемая поляризованностью (56.17а), также представляется фотоном, то его энергия и импульс равны

$$\epsilon = \hbar(\omega_1 + \omega_2), \quad p = \hbar(\omega_1 + \omega_2)n(\omega_1 + \omega_2)/c. \quad (56.38)$$

Видно, что закон сохранения энергии при этом выполняется:

$$\epsilon = \epsilon_1 + \epsilon_2, \quad (56.39)$$

а закон сохранения импульса не выполняется:

$$p \neq p_1 + p_2, \quad (56.40)$$

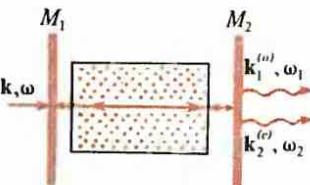
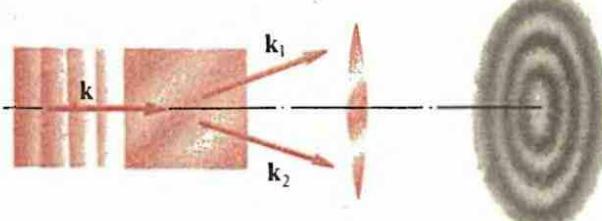
поскольку

$$\omega_1 n(\omega_1) + \omega_2 n(\omega_2) \neq (\omega_1 + \omega_2) n(\omega_1 + \omega_2). \quad (56.41)$$

В случае удвоения гармоник ($\omega_1 = \omega_2 = \omega$) при выполнении условия пространственного синхронизма $n(\omega) = n(2\omega)$ неравенство (56.41) превращается в равенство $2\omega n(\omega) = 2\omega n(2\omega)$ и закон сохранения импульса удовлетворяется.

300

Наблюдение спонтанного распада
фотона



301

Схема параметрического генера-
тора света

Для того чтобы применить механизм генерации удвоенной частоты к генерации суммарных и разностных частот, необходимо рассмотреть электромагнитные волны, распространяющиеся в произвольном направлении в среде. Вместо волн (56.10) и (56.11) в направлении оси Z возьмем волны, характеризуемые волновыми векторами \mathbf{k}_1 и \mathbf{k}_2 :

$$E_1 = E_{01} \cos(\omega_1 t - \mathbf{k}_1 \cdot \mathbf{r}), \quad (56.42)$$

$$E_2 = E_{02} \cos(\omega_2 t - \mathbf{k}_2 \cdot \mathbf{r}). \quad (56.43)$$

Для поляризованностей (56.15) — (56.17) получаем аналогичные соотношения, но с заменой $k_1 z \rightarrow \mathbf{k}_1 \cdot \mathbf{r}$, $k_2 z \rightarrow \mathbf{k}_2 \cdot \mathbf{r}$. В частности, имеем

$$P_{\omega_1 + \omega_2} = 2aE_{01}E_{02} \cos[(\omega_1 + \omega_2)t - (\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{r}]. \quad (56.44)$$

Порождаемая этой поляризованностью волна может рассматриваться как плоская монохроматическая волна с волновым вектором и частотой ω лишь в том случае, когда выполняются законы сохранения энергии и импульса:

$$\omega = \omega_1 + \omega_2, \quad (56.45) \quad \mathbf{k} = \mathbf{k}_1 + \mathbf{k}_2. \quad (56.46)$$

Равенство (56.46) называется условием векторного пространственного синхронизма. Его применение для удвоения частоты волн, распространяющихся в одном направлении, как это было объяснено в связи с (56.41), приводит к условию пространственного синхронизма (56.23).

При наличии дисперсии это условие вследствие (56.41) не может быть выполнено в изотропных средах даже для коллинеарных векторов \mathbf{k}_1 и \mathbf{k}_2 , а тем более оно не может быть выполнено для неколлинеарных. Поэтому следует искать условия удовлетворения векторного пространственного синхронизма в анизотропных средах. Пользуясь обыкновенным и необыкновенным лучами при определенных углах между волновыми векторами, удается удовлетворить этому условию. Например, в кристалле KDP (дигидрофосфат калия KH_2PO_4) оно может быть удовлетворено двумя способами:

$$\mathbf{k}^e = \mathbf{k}_1^{(o)} \pm \mathbf{k}_2^{(o)}, \quad (56.47)$$

$$\dot{\mathbf{k}} = \dot{\mathbf{k}}_1^{(o)} + \dot{\mathbf{k}}_2^{(o)}. \quad (56.48)$$

- В чем состоит физический смысл пространственного синхронизма? Каким образом осуществляется пространственная синхронизация? В чем состоит физический смысл векторного условия пространственной синхронизации? Опишите принцип параметрического усиления света и его применение в параметрических генераторах света.

Генерация суммарных и разностных частот. Если волны, характеризуемые векторами \mathbf{k}_1 и \mathbf{k}_2 , удовлетворяют условию векторного пространственного синхронизма (56.46), то волна поляризованности (56.44) с максимальной интенсивностью генерирует электромагнитную волну частотой $\omega_1 + \omega_2$. Но это не означает, что при другом соотношении векторов \mathbf{k}_1 и \mathbf{k}_2 не будет генерации частотой $\omega_1 + \omega_2$. Так же как и в случае удвоения частоты при невыполнении условия пространственного синхронизма, эта генерация происходит на небольших длинах когерентности и является слабой. Аналогично анализируется генерация разностной частоты $|\omega_1 - \omega_2|$ и других комбинационных частот.

Спонтанный распад фотона. При прохождении света через подходящим образом ориентированный кристалл, когда может быть выполнено условие (56.46), некоторые из фотонов падающего излучения распадаются на два фотона с меньшими частотами ω_1 и ω_2 , движущимися под соответствующими углами к исходному направлению (рис. 300). Это явление называется спонтанным распадом фотона. Оно наблюдается экспериментально. Картина аксиально-симметрична. В центре картины пятно, образованное фотонами, прошедшими через кристалл без распада, окруженное окрашенным кругом с изменяющейся от центра к периферии окраской, причем по мере удаления от оси длина волны уменьшается в полном согласии с требованием векторного пространственного синхронизма.

Параметрическое усиление света. Если в среде распространяются три волны, удовлетворяющие условию (56.46), то между ними происходит обмен энергией. Если одна из волн (например, \mathbf{k}) значительно мощнее, чем две другие, то энергия переходит от мощной волны в более слабые волны, в результате чего последние усиливаются. Это явление называется параметрическим усилением света, поскольку его можно рассматривать как модуляцию оптических параметров среды волной \mathbf{k} , приводящую к усилению волн \mathbf{k}_1 и \mathbf{k}_2 .

Параметрические генераторы света (ПГС). В основе их действия лежит параметрическое усиление света. ПГС — это лазер, в резонатор которого введен нелинейный кристалл, вырезанный таким образом, чтобы для осевых лучей выполнялось одно из векторных условий пространственного синхронизма: $\mathbf{k}_1^{(o)} + \mathbf{k}_2^{(o)} = \mathbf{k}^{(e)}$ или $\mathbf{k}_1^{(e)} + \mathbf{k}_2^{(e)} = \mathbf{k}^{(o)}$ (рис. 301). Рабочими частотами генератора являются ω_1 и ω_2 , в соответствии с которыми подбираются коэффициенты отражения зеркал M_1 и M_2 . Накачка генератора производится пучком света с частотой ω , для которого зеркало M_1 должно быть достаточно прозрачным.

При изменении ориентации кристалла или его параметров (например, за счет изменения температуры, изменения внешнего электрического поля и т. д.) изменяются условия пространственного синхронизма для осевых лучей, вследствие чего изменяется частота генерируемого излучения. Благодаря этому возможна плавная перестройка генерируемой частоты.

§ 57 | Самовоздействие света в нелинейной среде

Рассматриваются основные закономерности самофокусировки и дефокусировки света в нелинейной среде.

Нелинейная поправка к показателю преломления. Для изотропных сред или кристаллов с центром симметрии должно по определению соблюдаться соотношение

$$\mathbf{P}(-\mathbf{E}) = -\mathbf{P}(\mathbf{E}), \quad (57.1)$$

Отсюда следует, что в выражении (55.14) исчезают все члены, содержащие четное число множителей \mathbf{E} . Поэтому первая нелинейная поправка дается кубическим членом и поляризованность принимает вид

$$\mathbf{P} = \epsilon_0 \alpha \mathbf{E} + a_3 E^2 \mathbf{E}, \quad (57.2)$$

откуда

$$\mathbf{D} = \epsilon \mathbf{E} = \epsilon_0 \mathbf{E} + \mathbf{P} = [\epsilon_0 (1 + \alpha) + a_3 E^2] \mathbf{E}. \quad (57.3)$$

Тогда диэлектрическая проницаемость

$$\epsilon = \epsilon_0(1+\alpha) + a_3 E^2.$$

(57.4)

Показатель преломления определяется соотношением

$$n = \sqrt{\epsilon/\epsilon_0} = \sqrt{1+\alpha+a_3 E^2/\epsilon_0} \approx \sqrt{1+\alpha} \left(1 + \frac{a_3 E^2}{2\epsilon_0(1+\alpha)}\right),$$

так как нелинейная поправка к показателю преломления много меньше единицы:

$$a_3 E^2 / [\epsilon_0(1+\alpha)] \ll 1.$$

Таким образом,

$$n = n_0 + n_2 E_0^2.$$

Здесь произведено усреднение квадрата напряженности поля по периоду колебаний, и поэтому E_0 является амплитудой напряженности электрического поля волны; n_0 представляет линейный показатель преломления света, а $n_2 E_0^2$ описывает нелинейную поправку к показателю преломления.

Самофокусировка и дефокусировка пучка. Плотность потока энергии по сечению пучка не постоянна, следовательно, и показатель преломления (57.7) изменяется по сечению пучка. Будем для определенности считать сечение пучка круглым, а распределение интенсивности — гауссовым (см. § 3). В зависимости от характера нелинейности знак поправки в (57.7) может быть как положительным, так и отрицательным.

Если нелинейная поправка положительна, то скорость движения периферических участков пучка больше, чем центральных. В результате плоский волновой фронт становится вогнутым в сторону распространения пучка (рис. 302, а) и происходит его фокусировка к оси.

Если нелинейная поправка отрицательна, то скорость движения центральных участков пучка больше, чем периферических. В результате плоский волновой фронт становится выпуклым в сторону распространения пучка (рис. 302, б) и происходит его дефокусировка от оси.

Длина самофокусировки. Обозначим a радиус пучка, E_0 — амплитуду напряженности электрического поля на оси, а амплитуду на расстоянии a от оси будем считать равной нулю (рис. 303). По определению, длиной самофокусировки называется путь l_{cf} , при прохождении которого в нелинейной среде пучок сходится к оси или, как говорят, схлопывается.

На длине l_{cf} между осевым лучом и крайними лучами пучка образуется разность фаз

$$\Delta\phi = \omega \Delta n l_{cf} / c,$$

где

$$\Delta n = n_2 E_0^2,$$

Физическим содержанием эффектов самофокусировки и самодефокусировки является деформация фронта волн в результате различий в их скорости распространения, обусловленных изменениями показателя преломления среды под воздействием волн, т. е. среда играет как бы роль посредника, через которого волны действуют на себя.

- Чем определяется, будет ли пучок света самофокусироваться или дефокусироваться?
Какие физические факторы определяют пороговую мощность излучения?
В чем состоят основные причины возникновения нелинейности показателя преломления?

Из рис. 303 следует, что

$$b = l_{\text{сф}} - \sqrt{l_{\text{сф}}^2 - a^2} \approx a^2/(2l_{\text{сф}}), \quad (57.10)$$

где $a/l_{\text{сф}} \ll 1$. При изображенном на рис. 303 изгибе волнового фронта разность фаз равна

$$\Delta\phi_1 = \omega n_0 b/c \approx \omega n_0 a^2/(2cl_{\text{сф}}), \quad (57.11)$$

где для b использовано выражение (57.10). Разности фаз (57.8) и (57.11) должны быть равны

$$\omega \Delta n l_{\text{сф}}/c = \omega n_0 a^2/(2cl_{\text{сф}}), \quad (57.12)$$

откуда

$$l_{\text{сф}} = a \sqrt{n_0/(2\Delta n)} = (a/E_0) \sqrt{n_0/(2n_2)}. \quad (57.13)$$

Эксперимент хорошо подтверждает зависимость $l_{\text{сф}} \sim a/E_0$.

Эффект заметен уже при сравнительно малых $\Delta n/n$. Из (57.13) имеем

$$\Delta n/n = a^2/(2l_{\text{сф}}^2) \quad (57.14)$$

и, например, при $\Delta n/n \approx 0,5 \cdot 10^{-4}$ получаем $a/l_{\text{сф}} \approx 10^{-2}$, что нетрудно наблюдать.

Выразим E_0 через плотность потока энергии излучения, предполагая распределение амплитуд волн в пучке в виде

$$E(r) = E_0 \sqrt{1 - r^2/a^2}. \quad (57.15)$$

Для средней плотности потока энергии имеем выражение

$$\begin{aligned} < S > &= v \epsilon E^2 / 2 = [\epsilon c E_0^2 / (2n_0)] (1 - r^2/a^2) = \\ &= (\epsilon_0 n_0 c E_0^2 / 2) (1 - r^2/a^2). \end{aligned} \quad (57.16)$$

Мощность потока энергии в пучке равна

$$P = \int_{\sigma} < S > d\sigma = \pi \epsilon_0 n_0 c E_0^2 \int_0^a (1 - r^2/a^2) r dr, \quad (57.17)$$

где σ — площадь поперечного сечения пучка. Выполнив интегрирование, находим

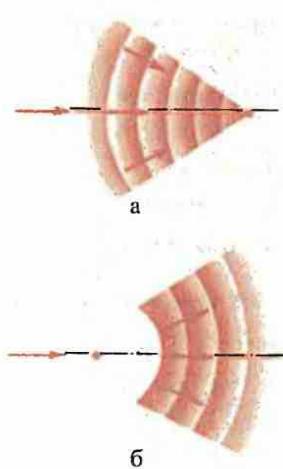
$$P = \pi \epsilon_0 n_0 c E_0^2 a^2 / 4 = \pi \epsilon_0 n_0 c a^4 / (8n_2 l_{\text{сф}}^2). \quad (57.18)$$

Например, у сероуглерода $n_0 = 1,62$; $n_2 = 0,22 \cdot 10^{-19} \text{ м}^2/\text{В}^2$, и для получения $l_{\text{сф}} = 10 \text{ см}$ при $a = 1 \text{ см}$ мощность составляет 10 МВт.

Пороговая мощность. Требуемая для самофокусировки мощность потока энергии в пучке уменьшается с уменьшением радиуса a пучка. Однако при уменьшении радиуса пучка увеличивается дифракционная расходимость, для преодоления которой необходимо увеличивать мощность потока энергии в пучке. Пороговой называется минимальная мощность, вызывающая склонение пучка. Она может быть найдена из следующих соображений.

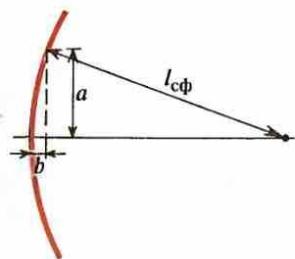
За счет нелинейности пучок склоняется на длине $l_{\text{сф}}$. Это означает, что периферийный луч пучка (рис. 304) отклоняется от направления оси на угол

$$\Delta\phi_1 \approx a/l_{\text{сф}}. \quad (57.19)$$



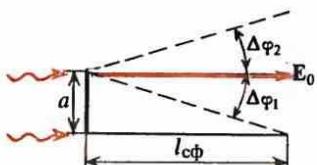
302

Возникновение фокусировки и дефокусировки в нелинейной среде:
 $n_2 < 0$ (а); $n_2 > 0$ (б)



303

К расчету длины самофокусировки



304

К расчету пороговой мощности
пучка для самофокусировки

Из-за дифракции периферийный луч пучка отклоняется от направления оси на угол $\Delta\varphi_2 \approx \lambda/(an_0)$.

(57.20)

Пороговую мощность находим из условия

$$\Delta\varphi_1 = \Delta\varphi_2, \quad (57.21)$$

принимающего с учетом (57.19) и (57.20) вид

$$a/l_{c\phi} = \lambda/(an_0). \quad (57.22)$$

Учитывая (57.13), получаем

$$a^2 = \lambda l_{c\phi}/n_0 = [\lambda a / (n_0 E_0)] \sqrt{n_0 / (2n_2)}. \quad (57.23)$$

Возводя обе части (57.23) в квадрат, находим

$$a^2 E_0^2 = \lambda^2 / (2n_0 n_2). \quad (57.24)$$

Подставляя (57.24) в (57.18), имеем

$$P_{\text{пор}} = \pi \epsilon_0 c \lambda^2 / (8n_2). \quad (57.25)$$

При $P > P_{\text{пор}}$ пучок света фокусируется, а при $P < P_{\text{пор}}$ — испытывает дифракционную расходимость. Пороговая мощность не зависит от радиуса пучка и уменьшается с уменьшением длины волны. Например для сероуглерода при облучении рубиновым лазером с $\lambda = 694,3$ нм

$$P_{\text{пор}} = 23 \text{ кВт}. \quad (57.26)$$

Основные причины возникновения нелинейности показателя преломления. 1. Как было выяснено при рассмотрении дисперсии, показатель преломления зависит от поглощения, а оно зависит от плотности излучения (см. § 51). Следовательно, показатель преломления зависит от квадрата напряженности электрического поля волны.

2. В результате электрострикции под действием электрического поля волны происходит изменение плотности диэлектрика. Давление на поверхность элемента объема диэлектрика: $p \sim p(\partial\varepsilon/\partial p)$, а связанное с соответствующим изменением плотности диэлектрика изменение показателя преломления $\Delta n \sim p(dn/dp)$ и зависит от плотности потока энергии.

3. Нагревание меняет плотность среды и, следовательно, показатель преломления. Нагревание среды зависит от плотности потока энергии в волне. Поскольку на оси пучка плотность потока больше, чем на периферии, центральные области пучка нагреваются сильнее и их показатель преломления уменьшается. Следовательно, в результате нагревания происходит дефокусировка пучка.

4. Поляризуемость молекул в различных направлениях, вообще говоря, различна. В малых внешних полях молекулы ориентированы беспорядочно и поэтому нет анизотропии в поляризационных свойствах среды. В сильных полях молекулы ориентируются определенным образом относительно поля, в результате чего поляризованность и показатель преломления становятся анизотропными, а среда в оптическом отношении превращается в одноосный кристалл. Возникает двойное лучепреломление, причем показатель преломления n необыкновенного луча зависит от направления распространения. Возникающая при этом нелинейность — ориентационной.

Инерционность. Электрострикционное изменение зависит от времени установления уплотнения. Скорость упругой волны в среде $v_3 \approx 1,5 \cdot 10^3$ м/с и, следовательно, время установления уплотнения в среде имеет порядок $t \approx a/v_3 \approx 10^{-6}$ с для $a \approx 1$ мм.

Время переориентации молекул имеет порядок $10^{-11} — 10^{-12}$ с. Вследствие этого ориентационный (керровский) механизм возникновения нелинейности для коротких лазерных импульсов продолжительностью менее 10^{-7} с является основным. Для более длительных импульсов необходимо принимать во внимание также другие механизмы возникновения нелинейности. Относительная роль различных механизмов зависит от свойств вещества и характеристик импульсов излучения.

Величина			Единица	
наименование	размерность	основное обозначение	наименование	обозначение
<i>Основные единицы</i>				
Длина	L	l	метр	м
Масса	M	m	килограмм	кг
Время	T	t	секунда	с
Сила тока	I	I	ампер	А
Термодинамическая температура	θ	T	kelvin	К
Количество вещества	N	n	моль	моль
Сила света	J	I	кандела	кд
<i>Производные единицы</i>				
Скорость фазовая	LT^{-1}	v_f, v	метр в секунду	м/с
Скорость групповая	LT^{-1}	v_g, v	метр в секунду	м/с
Скорость света в вакууме	LT^{-1}	c	метр в секунду	м/с
Ускорение	LT^{-2}	a	метр на секунду в квадрате	м/с ²
Сила	LMT^{-2}	F	ньютон	Н
Давление света	$L^{-1}MT^{-2}$	p	паскаль	Па
Энергия	L^2MT^{-2}	W, E	дюоуль	Дж
Плотность энергии, объемная	$L^{-1}MT^{-2}$	w	дюоуль на кубический метр	Дж/м ³
Мощность	L^2MT^{-3}	P	ватт	Вт
Плотность потока энергии	MT^{-3}	S	ватт на квадратный метр	Вт/м ²
Плотность потока импульса	$L^{-2}MT^{-1}$	G	килограмм-секунда на метр квадратный	кг·с/м ²
Площадь	L^2	S, σ	квадратный метр	м ²
Объем	L^3	V	кубический метр	м ³
Частота круговая	T^{-1}	ω, Ω	секунда в минус первой степени	с ⁻¹
Частота периодического процесса	T^{-1}	v	герц	Гц
Длина волны	L	λ	метр	м
			микрометр	мкм
			нанометр	нм
Время когерентности	T	τ_k	секунда	с
Продолжительность импульса	T	$\tau, \Delta t$	секунда	с
Длина когерентности	L	l_k	метр	м
Ширина когерентности	L	a_k	метр	м
Волновое число	L^{-1}	k	метр в минус первой степени	м ⁻¹
Фаза колебаний	безразмерная	ϕ	радиан	рад
Плоский угол	»	α, β, θ	радиан	рад
Телесный угол	»	Ω	стериadian	ср
Период колебаний	T	T	секунда	с
Коэффициент затухания	T^{-1}	γ	секунда в минус первой степени	с ⁻¹
Коэффициент усиления	L^{-1}	α	метр в минус первой степени	м ⁻¹
Добротность	безразмерная	Q		
Показатель преломления	»	n		
Коэффициент отражения	»	ρ		

Величина			Единица	
наименование	размерность	основное обозначение	наименование	обозначение
Коэффициент пропускания	»	τ	метр	м
Фокусное расстояние	L	f	радиан	рад
Разность фаз колебаний	безразмерная	δ	метр	м
Оптическая длина пути	L	Δ		
Показатель преломления обыкновенного луча	безразмерная	n_0		
Показатель преломления необыкновенного луча	»	n_e		
Плотность заряда объемная	$L^{-3}TI$	ρ	кулон на кубический метр	$\text{Кл}/\text{м}^3$
Абсолютная диэлектрическая проницаемость	$L^{-3}M^{-1}T^4I^2$	ϵ	фарад на метр	$\Phi/\text{м}$
Электрическая постоянная	$L^{-3}M^{-1}T^4I^2$	ϵ_0	то же	$\Phi/\text{м}$
Относительная диэлектрическая проницаемость	безразмерная	ϵ_r		
Напряженность электрического поля	$LMT^{-3}I^{-1}$	E	вольт на метр	$\text{В}/\text{м}$
Электрический момент диполя	LTI	p	кулон-метр	$\text{Кл}\cdot\text{м}$
Поляризованность	$L^{-2}TI$	P	кулон на квадратный метр	$\text{Кл}/\text{м}^2$
Электрическое смещение	$L^{-2}TI$	D	кулон на квадратный метр	$\text{Кл}/\text{м}^2$
Магнитная индукция	$MT^{-2}I^{-1}$	B	tesla	Tл
Напряженность магнитного поля	$L^2MT^{-2}I^{-1}$	H	ампер на метр	$\text{А}/\text{м}$
Абсолютная магнитная проницаемость	$LMT^{-2}I^{-2}$	μ	генри на метр	$\text{Гн}/\text{м}$
Магнитная постоянная	$LMT^{-2}I^{-2}$	μ_0	генри на метр	$\text{Гн}/\text{м}$
Относительная магнитная проницаемость	безразмерная	μ_r		
Спектральная плотность излучения по шкале круговых частот	$L^{-1}MT^{-1}$	w_{ω}	джоуль-секунда на кубический метр	$\text{Дж}\cdot\text{с}/\text{м}^3$
Спектральная плотность излучения по шкале длии волн	$L^{-2}MT^{-2}$	w_{λ}	джоуль на метр в четвертой степени	$\text{Дж}/\text{м}^4$
Сила излучения	L^2MT^{-3}	I	ватт на стерadian	$\text{Вт}/\text{ср}$
Энергетическая освещенность	MT^{-3}	E	ватт на квадратный метр	$\text{Вт}/\text{м}^2$
Энергетическая яркость	MT^{-3}	L	ватт на стерadian-квадратный метр	$\text{Вт}/(\text{ср}\cdot\text{м}^2)$
Энергетическая светимость	MT^{-3}	M	ватт на квадратный метр	$\text{Вт}/\text{м}^2$
Световой поток	J	Φ_V	люмен	лм
Освещенность	$L^{-2}J$	E_V	люкс	лк
Яркость	$L^{-2}J$	L_V	кандела на квадратный метр	$\text{кд}/\text{м}^2$
Светимость	$L^{-2}J$	M_V	люкс	лк

Заключение

Четвертым томом курса общей физики* завершается изложение классической физики. Ее достаточно краткое определение гласит: классической называется физика, в которой роль квантовых закономерностей пренебрежимо мала.. На первый взгляд кажется, что это определение бессодержательно, поскольку оно говорит не о том, чем является классическая физика, а о том, чем она не является. Однако такой взгляд обманчив, потому что существует только единая физика, которая является квантовой по своей сущности, и определение классической физики как той части единой физики, в которой роль квантовых закономерностей пренебрежимо мала, безусловно содержательно. Тем не менее целесообразно обсудить основные особенности классической физики без ссылки на квантовые закономерности.

Основными понятиями классической механики являются понятия материального тела, материальной точки, движения материальной точки по определенной траектории и силы, как причины тех или иных особенностей движения материальных тел и точек. Хотя классическая физика в современном понимании начинается с Ньютона, основные понятия и представления, на которых она базируется, зародились задолго до него. Они постепенно возникали в человеческом сознании с самых древних времен в процессе практической деятельности человека. Практическая деятельность также свидетельствовала, что все материальные тела имеют протяженность, занимают определенное место в пространстве и располагаются определенным образом друг относительно друга. Эти наиболее общие свойства материальных тел отразились в сознании человека в виде понятия пространства, а математическая формулировка этих свойств была выражена в виде системы геометрических понятий и связей между ними. Практическая деятельность человека также свидетельствовала о том, что окружающий его материальный мир находится в процессе постоянных изменений. Свойство материальных процессов иметь определенную длительность, следовать друг за другом в определенной последовательности и развиваться по этапам и стадиям отразилось в человеческом сознании в виде понятия времени.

Перечисленные выше основные понятия классической механики и понятия пространства и времени явились фундаментом, на котором покоилось развитие всей классической физики до наших дней. В процессе развития уточнялась взаимосвязь этих понятий и сами понятия, но они постоянно являлись основой классической физики. Наиболее важный результат этого развития состоит в установлении неразрывности связи пространства, времени, материи и движения. В философском плане развитие этих идей нашло свое завершение в учении диалектического материализма. Для диалектического материализма пространство и время являются формами существования материи и поэтому немыслимы без материи, а движение есть способ существования материи. Материя, пространство, время и движение всегда существуют в неразрывной связи друг с другом.

Механика Аристотеля содержала в себе основные идеи общего подхода к описанию механического движения материальных тел. Эти идеи полностью сохранили свое значение и в механике Ньютона, однако теория движения Аристотеля после примерно двухтысячелетнего господства была заменена теорией Ньютона. Аристотель считал, что все движения материальных тел можно разделить на две категории – «естественные» и «насильственные». «Естественные движения» осуществляются сами по себе, без каких-либо воздействий. Ставить вопрос о причине «естественных движений» бессмысленно. Точнее говоря, на вопрос: почему осуществляется некоторое «естественнное движение»? – всегда имеется готовый, не требующий размышлений ответ: потому что это движение естественное, происходящее именно так, а не иначе, без

* Том первый «Механика и теория относительности», том второй «Молекулярная физика» и том третий «Электричество и магнетизм» выходили в издательстве «Высшая школа» соответственно в 1976, 1981, 1983 гг.

каких-либо внешних воздействий. «Насильственные движения» сами по себе не происходят, а осуществляются под влиянием внешних воздействий, описываемых с помощью понятия силы. На вопрос: почему осуществляется некоторое «насильственное движение»? — ответ гласит: потому что на тело действует сила, под влиянием которой оно движется так, как движется. Естественным Аристотель считал движения легких тел вверх, тяжелых тел вниз и движение небесных тел по небесной сфере. Остальные движения являются насильственными. Заметим, что если тело покоятся в результате невозможности осуществить «естественное движение», то этот покой является «насильственным». Например, если тело покоятся на горизонтальном столе, то отсутствие его движения по вертикали является «насильственным» и обуславливается наличием соответствующей силы, действующей в вертикальном направлении, а отсутствие его движения по горизонтали обуславливается отсутствием силы, действующей в горизонтальном направлении. Это показывает, что закон движения не может быть положен в основу определения силы, хотя силу и можно находить из закона движения. Это замечание полностью относится и к попыткам использования второго закона Ньютона как определения силы. В механике Аристотеля сила обуславливает скорость тела, а понятие об ускорении отсутствует.

В механике Ньютона «естественным движением» в том смысле, как его понимал Аристотель, является прямолинейное равномерное движение материальной точки. В формулировке первого закона Ньютона устанавливаются условия, при которых это «естественное движение» (инерциальное) осуществляется. Он дает возможность выбрать такую систему координат, в которой такие «естественные движения» существуют. Вторым законом Ньютона устанавливается, что сила обуславливает не скорость материальной точки, а ее ускорение, причем не вообще ускорение, а ускорение в той системе координат, в которой при отсутствии силы скорость тела была бы постоянной, т. е. движение было бы «естественным». Как и в механике Аристотеля, сила учитывает влияние внешних условий на движение тела. Источниками силы являются материальные тела и, следовательно, сила является количественной мерой взаимодействия материальных тел. Третий закон Ньютона устанавливает, что сила, с которой одно из взаимодействующих тел действует на другое, равна по абсолютной величине, но направлена противоположно силе, с которой это другое тело действует на первое. Вопрос о силах в таком плане в механике Аристотеля не ставился.

Таким образом, теория движения Ньютона является принципиально новым шагом относительно теории Аристотеля. Среди главных новых моментов следует отметить все вопросы, связанные с введением систем координат, включая вопрос о принципе относительности, вопрос о свойствах взаимодействия тел, новое уравнение движения и дальнейшую разработку вопроса о пространстве и времени. Однако основные понятия механики Аристотеля и подход к проблеме движения в механике Ньютона сохранились без существенных изменений.

Следующим крупным шагом явилось создание специальной теории относительности. Ее революционный характер выразился в новом подходе к проблеме пространства и времени. В результате этого неразрывная связь пространства, времени и движения, на которую задолго до создания теории относительности указывал диалектический материализм, стала основополагающим моментом физической теории. Однако по своему содержанию специальная теория относительности полностью относится к классической физике. В результате создания общей теории относительности неразрывная связь пространства, времени, движения и материи стала основополагающим моментом наиболее общей физической теории. По своему содержанию общая теория относительности, так же как и специальная, полностью относится к классической физике. Теории, в которых существенны закономерности специальной или общей теории относительности, называют релятивистскими. Если в этих теориях несущественны квантовые закономерности, то они полностью относятся к классической физике.

Механика точки Ньютона явилась основой для построения механики совокупностей точек, составляющих материальные тела, среды и т. д. Если движение отдельных

точек описывается в соответствии с законами Ньютона, то соответствующая теория относится полностью к классической физике. Во многих случаях в механике тела или среды используется представление о сплошной среде, когда масса считается как бы непрерывно «размазанной» в пространстве, а движение элемента массы в бесконечно малом объеме описывается законами механики точки. Механика сплошных сред при этом условии относится также к классической физике. В связи с этим о механике твердого тела необходимо сделать такое замечание. Уравнения движения твердого тела включают три уравнения для координат центра масс и три уравнения моментов. Строго говоря, эти шесть уравнений не могут быть выведены только на основании трех законов Ньютона для материальной точки. Для их вывода необходимо использовать дополнительное к законам Ньютона предположение, что силы взаимодействия материальных точек, составляющие твердое тело, центральны. Однако это не изменяет принадлежности механики твердого тела к классической физике.

При анализе движения системы многих реальных точек, каждая из которых движется в соответствии с законами Ньютона, динамическое описание системы неосуществимо с технической, непригодно с теоретической и бесполезно с практической точек зрения. В системах многих частиц возникают новые закономерности движения, обусловленные наличием большого числа частиц в системе, которые называются статистическими. Статистическая физика, элементарными динамическими законами которой являются законы Ньютона, относится к классической физике и называется обычно классической статистической физикой. Следует, однако, подчеркнуть, что последовательное и полное обоснование ее возможно лишь с использованием квантовой теории.

Во всех рассмотренных выше разделах классической физики объектом исследования являлась материя в форме вещества. Другой формой материи, в исследовании которой физика достигла больших успехов, является ее полевая форма. Электрические и магнитные явления открыты очень давно, но теория этих явлений развивалась сравнительно медленно, и лишь в 60-х годах XIX столетия была завершена созданием теории Максвелла. После этого были открыты электромагнитные волны, которые существуют независимо от породивших их зарядов и токов. Это явилось экспериментальным доказательством самостоятельного существования электромагнитного поля и обосновало представление об электромагнитном поле как о форме существования материи. Движение этой формы материи описывается уравнениями Максвелла. Они представляют закон движения электромагнитного поля и описывают его порождение движущимися зарядами. Действие электромагнитного поля на заряды, носителями которых является материя в корпускулярной форме, описывается силой Лоренца. Основными понятиями, на которых основываются уравнения Максвелла, являются напряженность и индукция электромагнитного поля в точках пространства, изменяющиеся с течением времени, электромагнитное поле, порожденное зарядом, движущимся аналогично материальной точке по определенной траектории, и действующее на заряд. Это показывает, что теория, основанная на уравнениях Максвелла, относится к классической физике. Она является релятивистской инвариантной теорией и полностью относится к релятивистской классической физике.

После открытия полевой формы существования материи в виде электромагнитных волн и создания электромагнитной теории света появилась реальная возможность решить вопрос о законах взаимопревращения материи в полевой и корпускулярной форме или, другими словами, решить вопрос о взаимопревращении излучения и вещества. Казалось, что эту задачу можно успешно решить в рамках классической физики, поскольку каждая из этих форм материи хорошо описывается соответствующей классической теорией. Первое указание на недостаточность классической физики для понимания взаимоотношения этих двух форм материи было получено при анализе излучения абсолютно черного тела, когда необходимо было допустить дискретность актов испускания света. Затем были открыты корпускулярные свойства излучения и волновые свойства электронов и других частиц. Эти открытия показали, что не существует

барьера между корпускулярной и полевой формой материи, что эти формы взаимно проникают друг в друга и существуют в диалектическом единстве. Экспериментальное исследование и анализ этого диалектического единства привели к необходимости коренного пересмотра основных представлений классической физики и созданию квантовой теории.

В классической физике очень четко конкретизируются и находят свое воплощение философские категории диалектического материализма, а методологические принципы физических исследований имели большое влияние на разработку гносеологических вопросов. В ней полно и всесторонне воплощена сущность взаимного влияния и взаимопроникновения науки и философии. Это обстоятельство имеет большое мировоззренческое значение. Во всех четырех томах курса мировоззренческим вопросам уделено должное внимание. Достаточно полное освещение нашло диалектическое единство пространства, времени, движения и материи, что отразилось также и в ряде структурных особенностей курса. В частности, неприемлемо, как это часто делается, раскрывать содержание понятий пространства, времени и движения в рамках кинематики без установления органической связи между ними, а начало изложения вопроса о связи этих понятий с понятием материи откладывать до динамики, когда раскрывается понятие массы. Такой разрыв противоречит самой сущности пространства и времени, как форм существования материи, а движения — как способа ее существования. Этот разрыв ликвидируется изложением в самом начале курса физической кинематики, вводящим читателя в круг идей теории относительности, которая дает достаточно ясное воплощение в конкретной науке положения диалектического материализма о неразрывной связи пространства, времени, движения и материи. Суть этого диалектического единства прослеживается и уточняется в последующих разделах курса. Достаточно полное отражение в курсе классической физики находят вопросы всеобщей связи явлений, неуничтожаемости материи и движения, причинности и детерминизма, трактовки законов как форм выражения связи явлений и т. д. Одним словом, в классической физике воплощение в конкретном знании общих философских категорий диалектического материализма и положений марксистско-ленинской гносеологии столь полно и совершенно, что самым актуальным становится вопрос о характере незавершенности этого конкретного знания и о содержании незавершенности единства конкретного знания с общефилософскими и гносеологическими категориями. Актуальность этого вопроса обусловливается тем, что только незавершенность конкретного знания и его единства с общефилософскими и гносеологическими категориями является источником и движущей силой развития как конкретного знания, так и философских и гносеологических категорий. В рамках классической физики эта незавершенность выступает лишь в потенциальной форме и не составляет действительного отрицания завершенности. Отрицание достигнутой в классической физике завершенности знания и его единства с общефилософскими и гносеологическими категориями осуществляется лишь в рамках квантовой физики и в соответствии с диалектикой отрицания приводит не только к дальнейшему развитию физики, но и дает мощный стимул разработке общефилософских и гносеологических проблем.

Предметный указатель

- Аббе — Портера эксперимент 248
Аберрации геометрические 134
— оптических систем 134
— сферические 135
— хроматические 134, 137
Аббе уравнение 139
Акомодация 141
Амплитуда волны световой 148
— комплексная 209
Анализатор 275
Анизотропия искусственная 285
— оптическая 262
Апертура линзы 238
— объектива 243
Астигматизм наклонных пучков 137
- Биссия 33
— частота 34
Бине билинза 169
Био закон 282
Больцмана постоянная 15
Брауна и Твисса опыт 193
Брюстера явление 102
Бургера закон 310
- Величины фотометрические 44
— энергетические 44
Верде постоянная 284
Вещества аморфные
— левовращающие 282
— правовращающие 282
— оптически активные 282
Видимость 155, 246
Вина закон смещения 307
— формула 306
Винера опыты 307
Волна бегущая 33
— квазимохроматическая 68
— квазиплоская 77
— неполяризованная 82
— опорная 250
— отраженная 106
— поляризованная 82
— циркулярно 277
— частично 82
— эллиптически 273, 278
— сигнальная 250
— стоячая 35
— электромагнитная плоская 22
— — — в анизотропной среде 264
- Волновое уравнение 17, 118
— число 88, 105
Волновой вектор 21, 88
— четырехмерный 24, 32
— пакет 75, 92
— фронт 119
— — — деление 152
Волны квазимохроматические 149
— — — матрица когерентности 194
- Волны когерентные, способы получения 149
— линейно поляризованные 37
— — — суперпозиция 37
— модулированные 73
— монохроматические 88
— плоские 18
— гармонические 20
— — — инвариантность 23
Восприимчивость диэлектрическая комплексная 89, 328
— второго порядка 330
— — — линейная 330
- Вращательная способность 281
Время излучения 66
— когерентности 80
— разрешения 79
- Гаусса постоянные 127
— — — физический смысл 129
Гауссов закон 80
— пучок 27
— свет 80
- Гауссова форма линии излучения 71
Генераторы света параметрические 338
Генерации порог 312
Главные максимумы излучения 225
— — — условие 225
— скорости 266, 268
— — — лучевые 269
- Глаз 141
Глубина проникновения 105, 110
Голограмма 250
— машинная 259
— плоской волны 250
— произвольного объекта 253
— точечного объекта 251
— цифровая 259
Гологramмы толстослойные 255
Голографии задачи 250
Голографическая интерферометрия 275
— — — применения 258
Грина формула 213
Групповая скорость 76, 92
— — — в анизотропной среде 267
Гюйгенса построение 272
— принцип 162
Гюйгенса — Френеля принцип 208
- Даэленис света 28
Д'Аламбера оператор 18
Детектирование синхронное 249
Дефокусировка пучка 339
Диапазон видимый 12
— — — длины волн 12
— — — частоты волн 12
— — — эффективность зрения 15
- Диафрагма апертурная 140
— — — зрачки входной и выходной 140
Диафрагмирование 140
Диоптрия 131
Диполь электрический 291
Дипольный момент 291
Дисперсия 75
— аномальная 91
— в диэлектрике 89
— вращательная 281
— нормальная 90
— света 93
- Дифракция 78
— на непрерывных структурах 229
— на отверстии круглом 223
— — — прямоугольном 220, 232
— — — решетке 226
— — — при наклонном падении лучей 229
— — — ультразвуковых волнах 231
— — — щели 222
— Фраунгофера 219
— Френеля 219
- Длина когерентности 80
— пути оптическая 180
- Добротность 313
Добротности модулированной метод 314
- Доплера эффект 71
— — — поперечный 25
— — — продольный 25
- Единица силы света — См.: Кандела
- Жамена интерферометр 160
- Закон Брюстера 102
— ЛамBERTA 48
— преломления Снеллиуса 97, 121, 123
— сохранения энергии 109
— — — в явлениях интерференции 170
- Заселенность уровней 310
— — — инверсная 311
- Звезды двойной системы 168
- Зонная пластика 212
- Излучатель элементарный 45
Излучение лазерное 29, 312
Излучения спектральная плотность 45
— спектральный состав 64
— энергетическая сила 45
- Изменение фазы при отражении 104
— — — преломления 104
- Изображение оптическое 128

- Изображений построение 129, 241
 — пространственная фильтрация 247
- Изображения восстановление 250
- Инвариантность фазы 24
- Интенсивность интерференционной картины 172, 245
 — луча необыкновенного 274
 — — обыкновенного 274
 — — рассеяния 293
- Интервал корреляции 83
- Интерференции определение 148
 — порядок 152
- Интерференция волн
 — монохроматических 150, 151
 — — поляризованных 276
 — — света немонохроматического 154
- Интерферограмма 155
- Интерферометр звездный 167
- Интерферометра разрешающая способность 174
- Источник конечного размера 163
- Источники анизотропии 262
 — вторичные 217
- Источников плоскость 217
- Канделя 47
- Картина дифракционная 222
 — — , условие минимумов амплитуд 225
 — — интерференционная 152
 — — , видимость 155, 157, 164, 168, 173, 182, 193
 — — от белого света 158
 — — — излучения 'двойной' звезды 168
- Картинны дифракционной плоскость 217
- Керра постоянная 285
 — явление 285
- Кирхгофа закон второй 303
 — — первый 302
- Когерентности взаимной функция 190
 — длина 80, 333
 — степень 192
 — — комплексная 192, 194, 202
 — угол 165
 — ширина 165, 254
- Когерентность временная 165, 166
 — пространственная 165, 166
 — частичная 190
- Кольца интерференционные 172
- Кома 136
- Компенсатор 279
- Концентрации раствора 282
- Корни спираль 211, 233
- Котон — Мутона явление 286
- Коэффициент контрастности 251
 — преломления 88, 118
 — пропускания 108, 171, 251
 — — линзы 236
 — — рассеяния 295
 — — усиления 309, 311, 312
- Кривая дисперсионная 90
 — — каустическая 135
- Кристалл модификация левовращающая 281
 — — правовращающая 281
- Кристаллы двуосные 269
 — — одиночные 272
 — — отрицательные 272
 — — положительные 272
- Кронекера символ 262
- Кундта постоянная 284
- Лазер газодинамический 325
 — — гелий-неоновый 323
 — — на красителях 325
 — — рубиновый 321
 — — СО₂ 323, 324
- Лазера принципиальная схема 312
- Лазерные спеклы 320
- Лазерный термояд 29
- Лазеры 250, 254, 310
 — импульсные 314
 — непрерывные 314
- Линзы тонкие 131
- Линии равного наклона 181
 — — — равной толщины 183
- Ллоида зеркало 169
- Лоренцева форма линии излучения 66
 — — — — , нормированная функция 66
 — — — — поглощения 67
- Лупа 142
- Луч 96
 — необыкновенный 272, 276
 — обыкновенный 272, 276
 — света 119
- Луча направление 269
- Лучепреломление двойное 272
- Люммера — Герке пластишка 179
- Мазеры 310
- Майклельсона интерферометр 150, 152
 — — с линейными полосами 157
 — — , юстировка 158
 — — эшелон 179
- Максвелла уравнения 17, 23
- Максутова мениковая система 144
- Малиус закон 274
- Мандельштама — Бриллюзона компоненты 297
- Мандельштама — Бриллюзона формула 297
 — — явление 290
 — — — в твердых телах 298
- Материалы полупрозрачные 190
- Матрицы когерентности 194, 195
 — оптической системы 127
 — передаточная 124, 135
 — преломляющая 124
 — преобразования 127
 — — точная 134
- Матрицы определитель 198
- Матричный метод расчета 188
- Маха — Цендера интерферометр 159
- Метод темного поля 245
 — фазового контраста 245
- Микроскоп 142
- Многолистовые пленки 188
 — — диэлектрические зеркала 189
- Мод синхронизация 318
- Мода колебаний 305
- Модель излучателя классическая 63
 — элементарного рассеивателя 291
- Модуляции глубина 249
- Модуляция 73
 — — амплитуды 73
 — — фазы 73
 — — частоты 74
- Моды аксиальные (продольные) 316
 — боковые 317
 — излучения лазера 315
- Мощности излучения спектральная плотность 45
- Мощность излучения 50, 319
 — пороговая 340
- Нелинейность квадратичная 328
- Николя призма 275
- Ньютона кольца 184
- Область дисперсионная 176, 226
 — дифракции Фраунгофера 219
 — — Френеля 232
- Объектив иммерсионный 138
- Окраска тел 94, 285
- Оптическая ось 269, 273, 274
 — сила 131
- Оптической системы ахроматизация 137
 — — главные плоскости 128
 — — — точки 128
 — — фокус
- Освещенность 49
 — энергетическая 47
- Отражательная способность слоя 187
- Отражение от параллельных поверхностей 180
- Парадокс Эйнштейна — Подольского — Розена 151
- Переходы вынужденные 308
 — спонтанные 308
- Планка постоянная 12, 305
 — формула 306
- Пластиника в полволны 278
 — — цепую волну 278
 — — четверть волны 277
- Плоскости главные 128
 — поляризации вращение 281
 — — — в аморфных веществах 281
 — — — — кристаллических телах 281
 — — — — магнитном поле 284
 — — фокальные 129
- Плоскость меридианальная 136

- Плоскость падения луча 96
 — сагиттальная 137
 Плотность излучения объемная 302, 306
 — — равновесная 302
 — — спектральная 302, 306
 — импульса 28
 — потока энергии 26, 88, 107, 148, 292, 309
 — — — закон ослабления 295
 — энергии объемная 36.
 — — в диэлектрике 88
 — — — монохроматической волны 148
 Поверхности лучевые 270, 272–273
 Поглощательная способность 302
 Поглощение селективное 94
 Показателя преломления
 — нелинейность 338, 341
 Показателя преломления
 — переменный 121
 Поле зрения 141
 Полихроизм 276
 Полное отражение 104, 107
 Полосы интерференции, размытие 152
 Поляризатор 275
 Поляризация 37, 109
 — в рассеянии Ми 296
 — круговая 38
 — при двойном лучепреломлении 275
 — — рэлеевском рассеянии 293
 — частичная 194
 — эллиптическая 38
 Поляризации степень 104, 197, 198, 294
 Поляризованность линейная 328, 330
 — нелинейная 328, 330
 — статическая 332
 Поляроид 275
 Постоянная магнитная 17
 — электрическая 17
 — электродинамическая 17
 Поток световой 48
 Потока энергии мощность 27
 Предельный угол 97
 Преломление света 97
 — на сферической поверхности 123
 Приближение Кирхгофа 216
 — оптическое 216
 — параксиальное 123, 134
 — Френеля 217
 Принцип таутохронизма 120
 — Фурье-спектроскопии 155
 Призма двойкопреломляющая 273
 — поляризационная 275
 Проницаемость диэлектрическая
 — относительная 88
 — — комплексная 109
 Просветление оптики 187
 Пуассона пятно 210
 Разность фаз интерферирующих
 лучей 150, 172
 Разность хода лучей 150, 172
 Разрезающая способность 173, 175, 226, 243, 255
 — — максимальная 244.
 Распределение амплитуд 239, 240, 248
 — волн по волновым числам 92
 — интенсивности 165, 250
 — — в картине дифракционной 219
 — — — — интерференционной 172
 — — при дифракции на щели 222
 — спектральной плотности
 излучения 307
 — энергии по частотам 65
 — яркости 164
 Рассеяние комбинационное 298
 — Ми 290, 295
 — многократное 290
 — рэлеевское 290, 291
 — света 91, 290
 — — неполяризованного 294
 Расстояние фокусное 128
 — — тонкой линзы 131
 Резонатор прямоугольный 315
 — цилиндрический 318
 Решетка дифракционная 224
 — — отражательная 227
 — — — , разрешающая способность 227
 Решетки амплитудно-фазовые 228
 — фазовые 228
 Рэлея — Джинса формула 305
 — закон 293
 — интерферометр 167
 — критерий 243
- Самофокусировка пучка 339
 Самофокусировки длина 339
 Свет естественный
 — (неполяризованный) 195
 — поляризованный линейно 276, 278
 — — полностью 195
 Светимость 49
 — энергетическая 46, 302
 Световод 122
 Сигнал модулированный 249
 Силы радиометрические 29
 Силы света эталон 47
 Скорость группы волн — См.:
 — Групповая скорость
 Скорость лучевая 267
 — света 88
 — фазовая — См.: Фазовая скорость
 Спектр амплитуд 57
 — импульса изолированного
 — — прямоугольного 59
 — — импульсов пилообразных 59
 — — — прямоугольных 58
 — — — мощности 82
 — — — нормированный 86
 — — — непрерывный 57
 — — — фаз 57
 Спектра ширина 60
 Спектральная световая
- эффективность 49
 — — — относительная 49
 Спектральная плотность
 излучения 45
 — — — мощности излучения 45
 — — — светового потока 48, 50
 — — — энергетической освещенности 47
 — — — — светимости 46
 — — — — силы излучения 45
 — — — яркости 46
 — — световая эффективность 49
 — — — относительная 49
 Спектральный состав излучения 64
 Среда активная 312
 Стефана — Больцмана закон 306
 — — постоянная 306
- Тваймана — Грина интерферометр 159
 Телескоп 143
 Тело абсолютно черное 303
 — — — , спектр излучения 307
 Тензор диэлектрической
 проницаемости 262
 — — — , главные плоскости 263
 Теорема Ван — Циттерта —
 — Цернике 200
 — взаимности Гельмгольца 217
 — Винера — Хинчина 88
 — Гельмгольца — Кирхгофа 213
 — Парсеваля 62
 — Планшереля 62
 Труба зрительная 143
- Увеличение 127
 — линзы 142, 130
 — лупы 142
 — микроскопа 142
 — системы 137
 Угол Брюстера 102
 Ультрафиолетовая катастрофа 305
 Уравнение линзы 129
 — — в форме Гаусса 130
 — — — — Ньютона 130
 — Френеля 265
 — эллипсоида волновых нормалей 270
 — эйконала 110
 Условие Вульфа — Брэгга 256
 — главного максимума 225
 — излучения 215
 — квазимонохроматичности 68
 — Петцвала 137
 — пространственного
 — синхронизма 333
 — — — векторное 336
 — Рэлея 226
 — цикличности 315
 Условия стационарной
 генерации 313
 — насыщения 311
 — усиления 310

- Усреднение по периоду колебаний** 79
Устройства проекционные 144
Уширение доплеровское 70
 — неоднородное 69
 — однородное 69
 — ударное 69
Фабри — Перо интерферометр 171
 — сканирующий 177
Фаза волны
 — случайная 78
Фазовая скорость 75, 106, 111, 264
 — в анизотропной среде 267
Фарадея явление 284
Ферма принцип 120
Фильтры интерференционные 178
Флуктуации плотности потока энергии 81
Фокус 128
Форма линии поглощения 67
 — составной линии излучения 72
Формула дифракции 216
Фотоаппарат 142
Фотоны 13
 — солнечного излучения 15
 — шумовые 15
Френселя бизеркало 169
 — бипризма 169
 — зоны 208
 — интегралы 233
Френселя — Кирхгофа формула дифракции 216
 — формулы 99, 100
Функция автокогерентности 191
 — автокорреляционная 83, 191
 — зрачка 238
 — корреляционная 86, 191
 — — нормированная 86
 — толщины линзы 236
Фурье-анализ 82
 — интеграл 56
 — -оптика 236
 — -ряд 56
 — -спектрометр 156
 — -спектроскопия 156, 161
Цвета кристаллических пластинок 279
Цуга волн импульс 31
 — — энергия 31
Частота 12
 — круговая 12
Частота несмешенная 298
 — несущая 79, 249
Частоты комбинационные 330
 — отрицательные 61
Число комплексное 233
Ширина линий излучения 66, 316
 — спектра 73
Шмидта камера 144
Эйнштейна коэффициенты 308
Экспозиция световая 49
Электропроводимость среды 112
Эллипсоид волновых нормалей 269
 — лучевых скоростей 268
Энергия кванта 12
 — фотона 13
Эталон силы света 47
Эффект Покельса 286
Юнга схема 162
Яркости распределение 164
Яркость 48
 — энергетическая 46
Ячейка Керра 286
 — Покельса 287

Алексей Николаевич Матвеев

Оптика

Зав. редакцией литературы по физике

и математике Е. С. Гридасова

Редактор Г. Н. Чернышева

Мл. редакторы С. А. Доровских, Н. П. Майкова

Оформление художника Ю. Д. Федичкина

Художественный редактор

В. И. Пономаренко

Технический редактор З. А. Муслимова

Корректор Г. И. Костrikova

ИБ № 4184

Изд. № ФМ—741. Сдано в набор 05.04.84. Подп. в печать 13.02.85. Т—03075
Формат 70 × 90¹/₁₆. Бум. тип. № 3. Гарнитура Таймс. Печать офсетная.
Объем 25,74 усл. печ.л. + форзац 0,29 усл.печ.л., 52,65 усл. кр.-отт., 27,90
уч.изд.л. + форзац 0,46 уч.изд.л. Тираж 25000 экз. Зак. 289. Цена 1 р. 40 к.

Издательство «Высшая школа», 101430, Москва, ГСП-4, Наглинная ул.,
д. 29/14.

Ярославский полиграфкомбинат Союзполиграфпрома при Государственном
комитете СССР по делам издательства, полиграфии и книжной торговли.
150014, Ярославль, ул. Свободы, 97.

Поправка

Часть рисунка 125 (с. 172), выполненную черной краской, следует повернуть на 180° (ср. с рис. 124).