

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
МОСКОВСКИЙ ФИЗИКО-ТЕХНИЧЕСКИЙ ИНСТИТУТ
(ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ)

Н. А. Кириченко

ПРИНЦИПЫ ОПТИКИ

Допущено

*Учебно-методическим объединением высших учебных заведений РФ
по образованию в области прикладных математики и физики в качестве
учебного пособия для студентов вузов, обучающихся по направлению
«Прикладные математика и физика», а также по другим математическим
и естественнонаучным направлениям и специальностям*

**МОСКВА
МФТИ
2016**

УДК 53(075)

ББК 22.3я73

К43

Рецензенты:

Доктор физико-математических наук, профессор *Г. В. Склизков*
Доктор физико-математических наук, профессор *Г. Р. Локишин*

Кириченко, Н.А.

Принципы оптики: учебное пособие / Н. А. Кириченко. —

К43 М.: МФТИ, 2016. — 308 с.

ISBN 978-5-7417-0596-4

Приведены основные определения и формулы оптики, изучаемые в курсе общей физики. Излагаются законы геометрической и волновой оптики, а также их применения для решения конкретных задач. Рассмотрено проявление законов оптики в оптических и спектральных приборах. Уделено внимание методам фурье-оптики и голографии. Рассмотрены принципы физической оптики: поляризация света, теория дисперсии, рассеяние волн. Отдельные главы посвящены распространению света в анизотропных средах, искусственной анизотропии и нелинейно-оптическим явлениям.

Пособие подготовлено на основе лекций, читавшихся автором в Московском физико-техническом институте (государственном университете), и предназначено для студентов технических вузов, изучающих данный предмет, и преподавателей, а также для специалистов в качестве справочного материала.

УДК 53(075)

ББК 22.3я73

Учебное издание

КИРИЧЕНКО Николай Александрович

Редактор *В. А. Дружинина*. Корректор *И. А. Волкова*

Подписано в печать 18.03.2016. Формат 70х100 ¹/₁₆. Усл. печ. л. 38,5.

Уч.-изд. л. 35,6. Тираж 1200 экз. Заказ № 77.

Федеральное государственное автономное образовательное учреждение высшего профессионального образования «Московский физико-технический институт (государственный университет)»

141700, Московская обл., г. Долгопрудный, Институтский пер., 9

Тел. (495) 408-58-22. E-mail: fio@nipt.ru

Отпечатано в ООО Печатный салон «Шанс» 125412, г. Москва, ул. Ижорская, д. 13, стр. 2

ISBN 978-5-7417-0596-4

© Кириченко Н. А., 2016

© Федеральное государственное автономное образовательное учреждение высшего профессионального образования «Московский физико-технический институт (государственный университет)», 2016

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ	8
Глава 1. ЭЛЕКТРОМАГНИТНЫЕ ВОЛНЫ	10
1.1. Спектральный диапазон оптического излучения	10
1.2. Волновое уравнение	11
1.3. Плоская волна	13
1.4. Принцип суперпозиции	15
1.5. Энергия, закон сохранения энергии	15
1.6. Законы отражения и преломления	16
1.7. Формулы Френеля	16
1.7.1. Теорема погашения Эвальда–Озеена (16). 1.7.2. s- и p-поляризация волны (17). 1.7.3. s-поляризованная волна (17). 1.7.4. p-поляризованная волна (18). 1.7.5. Нормальное падение волны на границу раздела сред (19). 1.7.6. Коэффициенты отражения и прохождения по энергии (20). 1.7.7. Угловая зависимость коэффициента отражения (20).	
1.8. Сферические волны	21
Глава 2. ГЕОМЕТРИЧЕСКАЯ ОПТИКА	23
2.1. Уравнение эйконала	23
2.2. Волновой фронт и лучи	25
2.3. Принцип Ферма	27
2.4. Основные законы геометрической оптики	30
2.5. Основные определения	31
2.6. Кардинальные точки оптической системы. Построение изображения	36
2.7. Формула Ньютона	37
2.8. Увеличение	37
2.8.1. Поперечное линейное увеличение (37). 2.8.2. Продольное линейное увеличение (38). 2.8.3. Угловое увеличение (38).	
2.9. Сферическая поверхность	39
2.9.1. Преломление сферической поверхностью (39). 2.9.2. Изображение предмета (42). 2.9.3. Теорема Лагранжа–Гельмгольца (43). 2.9.4. Условие синусов Аббе (44). 2.9.5. Зеркало (46).	
2.10. Тонкая линза	47
2.10.1. Линза (47). 2.10.2. Формула тонкой линзы (47). 2.10.3. Фокусное расстояние тонкой линзы (49). 2.10.4. Оптический центр тонкой линзы (49).	
2.11. Оптическая сила	52
2.12. Сложение центрированных оптических систем	53
2.12.1. Фокусные расстояния составной системы (53). 2.12.2. Оптическая сила составной системы (55).	
2.13. Оптические приборы	56

2.13.1. Луна (56). 2.13.2. Микроскоп (57). 2.13.3. Телескоп (труба Кеплера) (57). 2.13.4. Труба Галилея (59).	
2.14. Об аберрациях оптических систем	60
2.15. Основные понятия фотометрии	61
2.15.1. Поток энергии, световой поток, видность (61). 2.15.2. Сила света (62). 2.15.3. Интенсивность света (63). 2.15.4. Освещённость (63). 2.15.5. Яркость (64). 2.15.6. Светимость (65). 2.15.6. Яркость и освещённость изображения (65).	
Глава 3. ИНТЕРФЕРЕНЦИЯ ВОЛН	68
3.1. Интерференция монохроматических волн	68
3.2. Видность интерференционной картины	69
3.3. Интерференция плоских волн	69
3.4. Интерференция волн от одного источника	71
3.4.1. Схема Юнга (71). 3.4.2. Отклонение лучей клином (72). 3.4.3. Бипризма Френеля (73). 3.4.4. Интерференция волн при отражении от пластины (75). 3.4.5. Опыт Поля (77). 3.4.6. Билинза Бийе (79). 3.4.7. Бизеркала Френеля (80). 3.4.8. Интерференция при отражении от клина (80). 3.4.9. Зеркало Ллойда (81). 3.4.10. Кольца Ньютона (81).	
Глава 4. ВРЕМЕННАЯ КОГЕРЕНТНОСТЬ	84
4.1. Интерференция немонохроматических волн	84
4.1.1. Когерентность (84). 4.1.2. Время и длина когерентности (85). 4.1.3. Связь времени когерентности с шириной спектра (86). 4.1.4. Соотношение неопределённостей (87). 4.1.5. Влияние немонохроматичности на наблюдаемое число интерференционных полос (88). 4.1.6. Видность интерференционной картины (89).	
4.2. Функция временной когерентности	92
4.3. Теорема Винера–Хинчина	96
Глава 5. ПРОСТРАНСТВЕННАЯ КОГЕРЕНТНОСТЬ	100
5.1. Радиус пространственной когерентности	100
5.2. Видность интерференционной картины от протяжённого источника	102
5.2.1. Интерференционная картина в схеме Юнга (102). 5.2.2. Звёздный интерферометр Майкельсона (105).	
5.3. Теорема ван Циттерта–Цернике	106
Глава 6. ДИФРАКЦИЯ ВОЛН	111
6.1. Принцип Гюйгенса–Френеля	111
6.1.1. Принцип Гюйгенса (111). 6.1.2. О каустиках и особенностях волновых фронтов (112). 6.1.3. Принцип Гюйгенса–Френеля (113). 6.1.4. Интеграл Френеля (113).	
6.2. Задачи с осевой симметрией	115
6.2.1. Преобразование интеграла Френеля (115). 6.2.3. Зоны Френеля (115). 6.2.4. Зонные пластинки (118). 6.2.5. Поле за круглым отверстием на оси системы (120). 6.2.6. Графический метод суммирования амплитуд (123).	

6.2.7. Ещё о зонной пластинке (125).	6.2.8. Дифракция на круглом диске (128).	
6.3. Дифракция на краю экрана и на щели		130
6.3.1. Зоны Шустера (130).	6.3.2. Расчёт дифракции с помощью интеграла Френеля (132).	6.3.3. Спираль Корню (134).
6.4. Дифракция Фраунгофера		136
6.4.1. Типы дифракции (136).	6.4.2. Дифракция Фраунгофера на щели (137).	6.4.3. Дифракция Фраунгофера на круглом отверстии (139).
6.4.4. Геометрический смысл типов дифракции (141).	6.4.5. Соотношение неопределённостей (142).	
6.5. Фокусирующее действие линзы		143
6.5.1. Расчёт поля в фокусе линзы методом зон Френеля (143).	6.5.2. Поле в фокальной плоскости линзы (144).	6.5.3. Поле вблизи фокуса линзы на оптической оси (146).
6.6. Связь поля в фокальной плоскости линзы с преобразованием Фурье поля, падающего на линзу		148
6.6.1. Параксиальное приближение для интеграла Френеля (148).	6.6.2. Поле в фокальной плоскости линзы (149).	6.6.3. Фокусировка гауссова пучка излучения (151).
6.7. Связь дифракции Фраунгофера с пространственным преобразованием Фурье		152
6.8. Дифракция Фраунгофера на решетке		155
6.8.1. Дифракционная решётка (155).	6.8.2. Главные дифракционные максимумы (155).	6.8.3. Дифракция Фраунгофера на решётке (156).
6.8.4. Влияние размера щели на дифракционную картину (159).		
6.9. Дифракция рентгеновских лучей		161
6.9.1. Рентгеновское излучение (161).	6.9.2. Формула Брэгга–Вульфа (162).	6.9.3. Формулы Лауэ (163).
6.9.4. Эквивалентность формул Брэгга–Вульфа и Лауэ (164).	6.9.5. Лауэграмма (165).	
Глава 7. ОПТИЧЕСКИЕ И СПЕКТРАЛЬНЫЕ ПРИБОРЫ		167
7.1. Дифракционный предел разрешения оптических приборов		167
7.1.1. Дифракция Фраунгофера в оптических приборах (167).	7.1.2. Критерий Релея оптического разрешения (167).	7.1.3. Разрешающая способность глаза, телескопа и микроскопа (168).
7.2. Спектральные приборы		171
7.2.1. Основные характеристики (171).	7.2.2. Разрешающая способность призмы (174).	
7.3. Интерферометр Фабри–Перо		176
7.3.1. Устройство интерферометра Фабри–Перо (176).	7.3.2. Многолучевая интерференция (177).	7.3.3. Функция пропускания интерферометра (178).
7.3.4. Интерферометр как оптический резонатор (182).	7.3.5. Энергия, запасаемая в интерферометре (183).	

Глава 8. ПРИНЦИПЫ ФУРЬЕ-ОПТИКИ И ГОЛОГРАФИИ	186
8.1. Метод Релея	186
8.2. Дифракция на периодических структурах	190
8.2.1. Дифракция на синусоидальной решётке (190). 8.2.2. Эффект саморепродукции (193). 8.2.3. Функция пропускания клина (196).	
8.3. Функция пропускания линзы	197
8.4. Принципы фурье-оптики	199
8.4.1. Фурье-плоскость (199). 8.4.2. Теория Аббе формирования изображений (200). 8.4.3. Принципы пространственной фильтрации (200).	
8.5. Методы наблюдения фазовых объектов	202
8.5.1. Метод фазового контраста Цернике (202). 8.5.2. Метод тёмного поля (204).	
8.6. Элементы голографии	206
8.6.1. Схемы голографии (206). 8.6.2. Запись и восстановление плоской волны (207). 8.6.3. Голограмма с наклонным опорным пучком (209). 8.6.4. Разрешающая способность голограммы (211). 8.6.5. Ограничения на допустимую ширину спектра (212). 8.6.6. Голограмма Габора точечного источника (212). 8.6.7. Объёмные голограммы (214).	
Глава 9. ПОЛЯРИЗАЦИЯ СВЕТА	218
9.1. Типы поляризации	218
9.2. Поляризаторы	221
9.3. Поляроиды	221
9.3.1. Дихроизм (221). 9.3.2. Поляроиды (222). 9.3.3. Закон Малюса (223).	
9.4. Поляризация света при отражении	225
9.4.1. Явление Брюстера (225). 9.4.2. Стопа Столетова (227).	
Глава 10. ДИСПЕРСИЯ ВОЛН	229
10.1. Фазовая и групповая скорость	229
10.1.1. Дисперсия (229). 10.1.2. Групповая скорость (230). 10.1.3. Локализация волнового пакета (232). 10.1.4. Расплывание волнового пакета (232). 10.1.5. Формула Рэлея (235).	
10.2. Классическая теория дисперсии	236
10.2.1. Зависимость диэлектрической проницаемости среды от частоты излучения (236). 10.2.2 Нормальная и аномальная дисперсии (238). 10.2.3. Закон Бугера (240).	
10.3. Формула Лоренца–Лоренца	241
10.4. Дисперсия электромагнитных волн в плазме	242
10.4.1. Диэлектрическая проницаемость плазмы (242). 10.4.2. Закон дисперсии волн. Фазовая и групповая скорости (243). 10.4.3. Отражение электромагнитных волн от плазмы (244).	
Глава 11. РАССЕЯНИЕ СВЕТА	246
11.1. Молекулярное рассеяние	246
11.2. Сечение рассеяния	248

11.2.1. Излучение колеблющегося диполя (248).	11.2.2. Угловое распределение рассеянного излучения (249).	11.2.3. Сечение рассеяния (251).
11.3. Поглощение света малыми частицами	252	
11.4. Экстинкция. Закон Бугера–Ламберта–Бера	253	
11.5. Поляризация излучения при рассеянии	254	
11.6. Комбинационное рассеяние	256	
11.6.1. Упругое и неупругое рассеяние света (256).	11.6.2. Закономерности комбинационного рассеяния (256).	11.6.3. Классическое описание (257).
11.6.4. Квантовомеханическое описание (258).		
Глава 12. ЭЛЕМЕНТЫ КРИСТАЛЛООПТИКИ	260	
12.1. Двойное лучепреломление	260	
12.1.1. Обыкновенный и необыкновенный лучи (260).	12.1.2. Диэлектрическая проницаемость (262).	12.1.3. Показатель преломления волн в одноосных кристаллах (264).
12.1.4. Направление вектора Пойнтинга (267).		
12.2. Элементарная теория одноосных диэлектриков	268	
12.3. Кристаллические пластинки $\lambda/2$ и $\lambda/4$	269	
12.3.1. Определения (269).	12.3.2. Действие четвертьволновой пластинки (270).	12.3.3. Действие полуволновой пластинки (272).
12.4. Интерференция поляризованных лучей	272	
12.4.1. Постановка задачи (272).	12.4.2. Интерференция в параллельных лучах (273).	12.4.3. Интерференция в сходящихся лучах (274).
12.5. Способы получения линейно поляризованного света	275	
Глава 13. ИСКУССТВЕННАЯ ОПТИЧЕСКАЯ АНИЗОТРОПИЯ	277	
13.1. Пьезооптический эффект	277	
13.2. Эффект Поккельса	278	
13.3. Эффект Керра	280	
13.4. Эффект Фарадея	286	
13.4.1. Естественная оптическая активность (286).	13.4.2. Эффект Фарадея (289).	
13.5. О других способах создания оптической анизотропии	292	
Глава 14. НЕКОТОРЫЕ ВОПРОСЫ НЕЛИНЕЙНОЙ ОПТИКИ ...	294	
14.1. Генерация второй гармоники. Оптическое выпрямление	294	
14.1.1. Нелинейная поляризация (294).	14.1.2. Оптическое «выпрямление» (297).	14.1.3. Генерация второй гармоники (297).
14.2. Генерация третьей гармоники	302	
14.3. Самофокусировка излучения	304	
ЛИТЕРАТУРА	308	

ПРЕДИСЛОВИЕ

Предлагаемое пособие основано на лекциях, читавшихся автором студентам Московского физико-технического института (государственного университета) в рамках курса общей физики.

Оптика — один из важнейших разделов физики, имеющий многочисленные приложения в самых разнообразных областях современной науки и техники. Без оптических приборов сегодня невозможно представить жизнь. Развитие оптики как науки имеет долгую историю, в которой принимали участие многие выдающиеся учёные разных стран. Были в истории и заблуждения, и прозрения, была и кропотливая работа, тонкие измерения. Иногда случайные наблюдения приводили к открытию новых закономерностей.

Оптику можно условно разбить на несколько разделов. Первый из них — это геометрическая оптика, с которой чаще всего сталкиваются в повседневной жизни, принципы которой были разработаны несколько веков назад. Однако многочисленные наблюдения показали, что эти законы не универсальны, поведение света часто отклоняется от того, которое предписывается ими. Это привело к возникновению волновой оптики, объяснившей, в частности, такое явление, как дифракция. Создание данного раздела оптики в основном относится к XIX веку. Используя установленные законы, удалось разработать оптические и спектральные приборы, выяснить их предельные возможности.

Оказалось, что многие явления, изучаемые в волновой оптике, не специфичны для света, они характерны для волн самой разнообразной природы — волн в жидкостях, газах, твёрдых телах. Это позволило применять методы, разработанные в одних разделах физики, к исследованию явлений в других областях. Вместе с тем ряд явлений характерен именно для оптического излучения и обусловлен электромагнитной природой света. Сюда относятся эффекты, связанные с поляризацией и дисперсией.

Новые направления в оптике появились в связи с созданием лазеров — источников узких пучков излучения с высокой степенью монохроматичности. Эти пучки к тому же могут нести большую энергию. С последним связано рождение нелинейной оптики. Оказалось, в частности, что при распространении мощного излучения в веществе в спектре излучения могут появляться компоненты с частотами, отсутствующими в ис-

ходном излучении. Взаимодействие излучения с веществом может приводить и к таким эффектам, как самофокусировка.

В рамках одной книги трудно, даже невозможно, изложить все факты, установленные в оптике. Поэтому приходится выделять те из них, которые отражают наиболее существенные законы, знание которых позволяет ориентироваться и в других разделах. В предлагаемой книге даётся систематическое изложение общих принципов оптики и ряда их приложений. Это геометрическая оптика, основные законы теории волн, явления интерференции и дифракции, их проявления и применения в оптических и спектральных приборах. Подробно рассмотрены два основных подхода к расчёту волновых полей в разных условиях — метод интеграла Френеля и фурье-оптика. Последний подход позволяет, в частности, понять с иной точки зрения происхождение дифракции. Рассмотрено применение этого метода к голографии. В следующих главах обсуждаются явления, характерные именно для оптического излучения: поляризация света, рассеяние, поведение света в анизотропных средах. Наконец, затронуты некоторые вопросы нелинейной оптики: генерация кратных гармоник и самофокусировка излучения.

Содержание пособия отражает программу курса оптики, сложившуюся к настоящему моменту в МФТИ. Добавлено некоторое количество вопросов, прямо не отражённых в программе, но в той или иной степени затрагиваемых на семинарах и представляющих интерес для более полного понимания законов оптики. Приведено также небольшое число комментариев исторического плана, имеющих целью отразить, не претендуя на полноту, сложный путь, каким развивалась наука, некоторые важные вехи этого пути.

Для более полного изучения вопросов оптики можно воспользоваться литературой, список которой приведён в конце книги.

ГЛАВА 1

ЭЛЕКТРОМАГНИТНЫЕ ВОЛНЫ

1.1. Спектральный диапазон оптического излучения

Оптика рассматривает электромагнитное излучение в диапазоне длин волн

$$10^{-6} \text{ см} < \lambda < 0,2 \text{ см}$$

и соответственно частот

$$3 \cdot 10^{16} \text{ Гц} > \nu > 1,5 \cdot 10^{11} \text{ Гц.}$$

Этот диапазон включает как видимый свет, так и участки спектра, примыкающие к нему в области бóльших и мёньших длин волн. Видимый свет попадает в область

$$400 \text{ нм (фиол.)} < \lambda < 760 \text{ нм (красн.)}.$$

Основные диапазоны оптического излучения указаны в табл. 1.1.

Таблица 1.1

Диапазон		Частоты, Гц	Длины волн нм
Ультра- фиолетовый	Дальний (вакуумный)	$1.5 \cdot 10^{15} - 3.0 \cdot 10^{16}$	10–200
	Ближний	$7.9 \cdot 10^{14} - 1.5 \cdot 10^{15}$	200–380
Видимый		$4.0 \cdot 10^{14} - 7.9 \cdot 10^{14}$	380–740
Инфра- красный	Ближний	$1.2 \cdot 10^{14} - 4.0 \cdot 10^{14}$	740–2500
	Средний	$6.0 \cdot 10^{13} - 1.2 \cdot 10^{14}$	2500–5000
	Дальний	$1.5 \cdot 10^{11} - 6.0 \cdot 10^{13}$	5000 нм–0,2 см

В табл. 1.2 приведены длины волн и частоты света видимого диапазона спектра.

Особенность рассматриваемого в оптике спектрального диапазона состоит в том, что для такого излучения среда может считаться непрерывной, поскольку длина волны велика по сравнению с межатомными расстояниями: $\lambda \gg a$. При этом свойства среды по отношению к излучению полностью характеризуются диэлектрической и магнитной проницаемостями. Кроме того, оптический диапазон выделяют общие методы генерации, регистрации и преобразования излучения.

Таблица 1.2

Цвет	Частоты, Гц	Длины волн, нм
Фиолетовый	$6.8 \cdot 10^{14} - 7.9 \cdot 10^{14}$	380–440
Синий	$6.2 \cdot 10^{14} - 6.8 \cdot 10^{14}$	440–485
Голубой	$6.0 \cdot 10^{14} - 6.2 \cdot 10^{14}$	485–500
Зелёный	$5.3 \cdot 10^{14} - 6.0 \cdot 10^{14}$	500–565
Жёлтый	$5.1 \cdot 10^{14} - 5.3 \cdot 10^{14}$	565–590
Оранжевый	$4.8 \cdot 10^{14} - 5.1 \cdot 10^{14}$	590–625
Красный	$4.0 \cdot 10^{14} - 4.8 \cdot 10^{14}$	625–740

Вместе с тем методы, применяемые в оптике, могут в ряде случаев использоваться и для работы с электромагнитным излучением в спектральном диапазоне, существенно выходящем за пределы, указанные в таблице 1.1. Например, распространение радиоволн (то есть излучения с длиной волны от миллиметров до многих километров) в атмосфере хорошо описывается законами оптики. Наконец, многие законы, установленные применительно к оптическому излучению, справедливы для волн иной природы — в жидкостях, газах, в твёрдых телах и так далее.

1.2. Волновое уравнение

Основные свойства электромагнитных волн подробно рассматриваются в разделе «Электричество и магнетизм» курса общей физики. Поэтому мы здесь приведём только необходимые для дальнейшего результаты. Подробное изложение соответствующих вопросов содержится, например, в книгах [1–3], указанных в списке литературы.

В области пространства, в которой отсутствуют свободные заряды и токи проводимости, электромагнитное поле описывается системой уравнений Максвелла:

$$\begin{aligned} \operatorname{rot} \mathbf{E} &= -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}, \quad \operatorname{div} \mathbf{D} = 0, \\ \operatorname{rot} \mathbf{H} &= \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t}, \quad \operatorname{div} \mathbf{B} = 0, \end{aligned} \quad (1.2.1)$$

дополненной материальными уравнениями:

$$\mathbf{D} = \mathbf{E} + 4\pi \mathbf{P}, \quad \mathbf{B} = \mathbf{H} + 4\pi \mathbf{I}. \quad (1.2.2)$$

В изотропных средах в случае слишком сильных полей векторы поляризации среды $\mathbf{P}$ и намагничивания $\mathbf{I}$ даются выражениями

$$\mathbf{P} = \alpha \mathbf{E}, \quad \mathbf{I} = \kappa \mathbf{H}, \quad (1.2.3)$$

так что из (1.2.2) следуют равенства

$$\begin{aligned} \mathbf{D} &= \varepsilon \mathbf{E}, \quad \mathbf{B} = \mu \mathbf{H}; \\ \varepsilon &= 1 + 4\pi\alpha, \quad \mu = 1 + 4\pi\kappa. \end{aligned} \quad (1.2.4)$$

Этими равенствами определяются диэлектрическая (ε) и магнитная (μ) проницаемости среды.

В указанных условиях на границе раздела сред имеют место следующие соотношения (*граничные условия*):

$$\begin{aligned} D_{1n} &= D_{2n}, \quad E_{1t} = E_{2t}, \\ B_{1n} &= B_{2n}, \quad H_{1t} = H_{2t}. \end{aligned} \quad (1.2.5)$$

Здесь индексы n и t указывают соответственно на нормальную и касательную к границе раздела сред компоненты полей.

Будем предполагать, что среда однородна и изотропна, а её свойства не зависят от величины электромагнитного поля, так что диэлектрическая (ε) и магнитная (μ) проницаемости постоянны: $\varepsilon = \text{const}$, $\mu = \text{const}$. В этих условиях, исключая из (1.2.1) магнитное поле, получим *волновое уравнение* для напряжённости электрического поля:

$$\frac{1}{v^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} - \Delta \mathbf{E} = 0. \quad (1.2.6)$$

Здесь $v = c/n$ — скорость света в среде, $n = \sqrt{\varepsilon\mu}$ — показатель преломления среды. Аналогичное уравнение получается и для напряжённости магнитного поля $\mathbf{H}$:

$$\frac{1}{v^2} \frac{\partial^2 \mathbf{H}}{\partial t^2} - \Delta \mathbf{H} = 0. \quad (1.2.7)$$

Во многих случаях можно принять $\mu = 1$. Тогда $n = \sqrt{\varepsilon}$, $v = c/n$.

Если поле меняется со временем по гармоническому закону с частотой ω

$$\ddot{\mathbf{E}} = -\omega^2 \mathbf{E}, \quad \ddot{\mathbf{H}} = -\omega^2 \mathbf{H}, \quad (1.2.8)$$

то волновые уравнения (1.2.6), (1.2.7) переходят в уравнения Гельмгольца:

$$\Delta \mathbf{E} + \frac{\omega^2}{v^2} \mathbf{E} = 0, \quad \Delta \mathbf{H} + \frac{\omega^2}{v^2} \mathbf{H} = 0. \quad (1.2.9)$$

1.3. Плоская волна

Будем искать решение волнового уравнения (1.2.6), зависящее только от одной координаты x , в следующем виде:

$$\mathbf{E}(x, t) = \mathbf{E}_1 \cos(kx - \omega t + \varphi_1). \quad (1.3.1)$$

Во многих случаях более удобным является комплексное представление:

$$\mathbf{E}(x, t) = \mathbf{E}_0 \exp[i(kx - \omega t)], \quad (1.3.2)$$

в котором коэффициент $\mathbf{E}_0 = \mathbf{E}_1 \exp(i\varphi_1)$, называемый *комплексной амплитудой*, является комплексной величиной, учитывающей как амплитуду $\mathbf{E}_1$, так и слагаемое φ_1 в аргументе косинуса. Далее, если не оговаривается иное, мы будем пользоваться этим представлением.

Подстановка выражения (1.3.2) в уравнение (1.2.6) даёт связь параметров ω и k :

$$k = \frac{\omega}{v} = \frac{\omega}{c} n. \quad (1.3.3)$$

Решение (1.3.2) называется *бегущей волной*. Величина ω называется (*круговой*) *частотой* волны, величина $\nu = \omega/2\pi$ — *циклической частотой*, а число k — *волновым числом*. Волна (1.3.1) описывает гармонический колебательный процесс в каждой точке пространства, имеющий *период*

$$T = 1/\nu = 2\pi/\omega.$$

Величина

$$\varphi = kx - \omega t \quad (1.3.4)$$

называется *фазой волны*. Записывая фазу в виде

$$\varphi = k(x - vt),$$

видим, что точки волны, имеющие фиксированное значение фазы $\varphi = \text{const}$, движутся со скоростью

$$v = \omega/k = c/n. \quad (1.3.5)$$

Эта величина называется *фазовой скоростью* волны.

Длиной волны λ называется пространственный период волны, то есть расстояние между двумя ближайшими точками волны, находящимися в одной фазе колебаний. Для волны (1.3.1) эта величина равна

$$\lambda = vT = \frac{c}{nv} = \frac{2\pi c}{n\omega} = \frac{\lambda_0}{n}, \quad (1.3.6)$$

где $\lambda_0 = 2\pi c/\omega$ — длина волны той же частоты в вакууме. Длина волны связана с волновым числом соотношением

$$k = 2\pi/\lambda. \quad (1.3.7)$$

Обобщение решения (1.3.2) на трехмерный случай имеет вид

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 \exp[i(\mathbf{k}\mathbf{r} - \omega t)]. \quad (1.3.8)$$

Это же решение можно переписать в форме

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 \exp[i\mathbf{k}(\mathbf{r} - \mathbf{v}t)], \quad (1.3.9)$$

в котором явно указывается, что волна распространяется со скоростью v . Волна (1.3.2) получается отсюда как частный случай, когда вектор $\mathbf{k}$ направлен вдоль оси x . Из сравнения (1.3.8) и (1.3.9) следует

$$\omega = \mathbf{k}\mathbf{v}. \quad (1.3.10)$$

Вектор $\mathbf{k}$ называется *волновым вектором*.

Введём единичный вектор $\mathbf{s}$, $|\mathbf{s}|=1$, направленный вдоль волнового вектора $\mathbf{k}$. Тогда для $\mathbf{k}$ имеет место выражение

$$\mathbf{k} = \frac{\omega}{c} n\mathbf{s}, \quad |\mathbf{k}| = k = \frac{\omega}{c} n, \quad (1.3.11)$$

где n — показатель преломления.

Фазой волны (1.3.8) называется величина

$$\varphi = \mathbf{k}\mathbf{r} - \omega t. \quad (1.3.12)$$

Множество точек, имеющих заданное значение фазы $\varphi = \varphi_0$, представляет собой плоскость, перпендикулярную волновому вектору. По этой причине волну (1.3.8) называют *плоской*.

В плоской волне векторы $\mathbf{E}$ и $\mathbf{H}$ зависят от координат и времени одинаково:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 \exp[i(\mathbf{k}\mathbf{r} - \omega t)], \quad \mathbf{H}(\mathbf{r}, t) = \mathbf{H}_0 \exp[i(\mathbf{k}\mathbf{r} - \omega t)], \quad (1.3.13)$$

причём амплитуды полей связаны соотношением

$$\sqrt{\epsilon} E_0 = \sqrt{\mu} H_0. \quad (1.3.14)$$

Для плоских волн уравнения Максвелла (1.2.1) принимают вид

$$\mathbf{k} \times \mathbf{E} = \frac{\omega}{c} \mathbf{B}, \quad \mathbf{k} \mathbf{D} = 0, \quad (1.3.15)$$

$$\mathbf{k} \times \mathbf{H} = -\frac{\omega}{c} \mathbf{D}, \quad \mathbf{k} \mathbf{B} = 0.$$

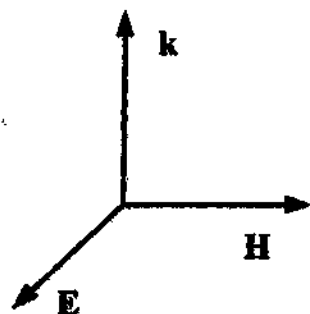


Рис. 1.3.1. Поперечный характер плоской электромагнитной волны

Из этих уравнений следует отражённое на рис. 1.3.1 свойство поперечности волн: векторы

$\{\mathbf{E}, \mathbf{H}, \mathbf{k}\}$ взаимно ортогональны.

Рассматривая только стационарные волновые процессы с определённой частотой ω в выражениях для полей вида (1.3.7) будем, как правило, опускать множитель $e^{-i\omega t}$, записывая $\mathbf{E} = \mathbf{E}_0 e^{ikr}$.

1.4. Принцип суперпозиции

Если свойства среды не меняются в зависимости от величины электромагнитного поля (например, когда средой является вакуум), то имеет место *принцип суперпозиции*, состоящий в следующем: если в некоторую область пространства попадают волны от нескольких разных источников, то напряжённости их электрического и магнитного полей складываются:

$$\mathbf{E} = \sum_i \mathbf{E}_i \quad \mathbf{H} = \sum_i \mathbf{H}_i.$$

Каждое из слагаемых в сумме не зависит от наличия или отсутствия других слагаемых, так же как и от их величин. При этом все поля, входящие в сумму, удовлетворяют тем же уравнениям, что и суммарное поле.

1.5. Энергия поля, закон сохранения энергии

Плотность энергии электромагнитного поля даётся формулой

$$w = \frac{\epsilon E^2 + \mu H^2}{8\pi}. \quad (1.5.1)$$

Плотность потока энергии в электромагнитной волне определяется вектором *Пойнтинга*:

$$\mathbf{S} = \frac{c}{4\pi} \mathbf{E} \times \mathbf{H}. \quad (1.5.2)$$

Для плоской волны имеет место соотношение:

$$S = vw, \quad v = c/n, \quad w = \epsilon E^2 / 4\pi, \quad (1.5.3)$$

вытекающее из (1.3.14), (1.5.1) и (1.5.2), а также из условия ортогональности полей $\mathbf{E}$ и $\mathbf{H}$ в плоской волне.

Закон сохранения энергии электромагнитного поля в среде имеет вид

$$\frac{\partial w}{\partial t} + \operatorname{div} \mathbf{S} = -\mathbf{j} \cdot \mathbf{E}. \quad (1.5.4)$$

Слагаемое в правой части равенства $(-\mathbf{j} \cdot \mathbf{E})$ описывает джоулевы потери в единицу времени в единице объёма. Если в среде отсутствуют свободные заряды и токи проводимости, то следует положить $\mathbf{j} \cdot \mathbf{E} \rightarrow 0$.

Среднее значение вектора Пойнтинга за период колебаний называется *интенсивностью излучения*:

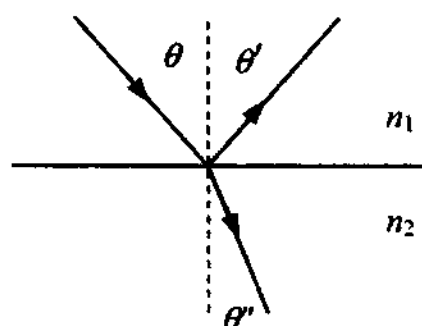
$$\mathbf{I} = \bar{\mathbf{S}} = \frac{c}{4\pi} \overline{\mathbf{E} \times \mathbf{H}}. \quad (1.5.5)$$

Для плоской волны в немагнитной среде ($\mu = 1$) с показателем преломления n интенсивность равна

$$I = \frac{c}{8\pi} n E_0^2, \quad (1.5.6)$$

где E_0 — амплитуда волны.

1.6. Законы отражения и преломления



При падении света на границу раздела сред появляются наряду с падающей волной, также волны отражённая и преломлённая (рис. 1.6.1). Если обе среды — однородные и изотропные, то направления распространения этих волн определяются двумя законами — *законом отражения*:

$$\theta = \theta' \quad (1.6.1)$$

и *законом преломления* (законом Снеллиуса):

$$n_1 \sin \theta = n_2 \sin \theta''. \quad (1.6.2)$$

Рис. 1.6.1. Падающая, отражённая и преломлённая волны на границе раздела двух сред

1.7. Формулы Френеля

1.7.1. Теорема погашения Эвальда–Озеена

Пусть на границу раздела сред падает плоская волна E . Тогда поле в среде в силу принципа суперпозиции складывается из этого поля и поля E_d , производимого колеблющимися диполями. Соответственно вектор поляризации среды можно представить в виде

$$P = N\beta(E + E_d),$$

где β — поляризуемость молекулы. Поле E распространяется со скоростью света в вакууме c , тогда как скорость распространения волны поляризации P равна c/n , где n — показатель преломления.

Имеет место *теорема погашения Эвальда–Озеена*, которая утверждает, что поле E_d содержит два слагаемых, одно из которых точно компенсирует (гасит) падающую волну, а второе описывает волну, распространяющуюся в веществе со скоростью c/n и представляющую собой наблюдаемую в среде преломлённую волну. Оказывается, погашение падающей волны осуществляется диполями, расположенными на границе среды, тогда как волна поляризации возникает вследствие взаимодействия соседних диполей в объёме среды.

Теорема погашения объясняет, каким образом в среде возникает преломлённая волна вместо волны падающей.

Данное утверждение для изотропных сред доказал шведский физик Карл Озеен в 1915 г., а для кристаллических сред — немецкий физик Пауль Эвальд в 1916 г.

1.7.2. s- и p-поляризация волны

Электромагнитная волна, распространяющаяся в изотропном пространстве, свободном от токов и зарядов, является поперечной:

$$\mathbf{E} \perp \mathbf{k}, \mathbf{H} \perp \mathbf{k}, \mathbf{E} \perp \mathbf{H}.$$

При наличии границы раздела сред в зависимости от ориентации вектора $\mathbf{E}$ по отношению к плоскости падения волны различают два типа *поляризации* (рис. 1.7.1):

1) *s-поляризация*, когда вектор $\mathbf{E}$ перпендикулярен плоскости падения,

2) *p-поляризация*, когда вектор $\mathbf{E}$ лежит в плоскости падения.

В первом случае при всех углах падения вектор $\mathbf{E}$ остаётся параллельным границе раздела сред, а во втором случае угол между вектором $\mathbf{E}$ и границей раздела меняется.

Формулы Френеля устанавливают соотношения между амплитудами падающей, отражённой и преломлённой волн в зависимости от угла падения и поляризации падающей волны.

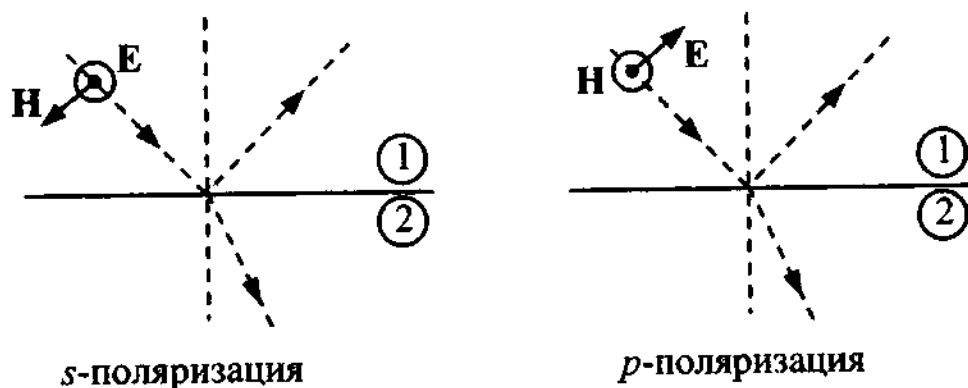


Рис. 1.7.1. Векторы $\mathbf{E}$ и $\mathbf{H}$ в падающей волне в некоторый момент времени: слева — s-поляризованная волна, справа — p-поляризованная волна. На рисунке кружок с точкой означает, что соответствующий вектор направлен «на нас»

Амплитудные коэффициенты отражения (r) и прохождения (d) волны определяются равенствами

$$r = E'_0/E_0, \quad d = E''_0/E_0, \quad (1.7.1)$$

где E_0 , E'_0 и E''_0 — амплитуды соответственно падающей (исходной), отражённой и преломлённой (прошедшей через границу раздела) волн.

1.7.3. s-поляризованная волна

Выберем условно направление вектора напряжённости электрического поля «нам навстречу» во всех трёх волнах — падающей, отражённой и прошедшей (рис. 1.7.2 слева). Вектор напряжённости магнитного поля $\mathbf{H}$ лежит в плоскости рисунка, а его направление определяется правилом винта, показанным на рис. 1.7.2 справа.

На границе раздела диэлектриков 1 и 2 выполняются условия непрерывности касательных к границе раздела компонент полей: $E_{1t} = E_{2t}$, $H_{1t} = H_{2t}$. Используя эти равенства, для амплитудных коэффициентов отражения и преломления можно получить соотношения

$$r_{\perp} = -\frac{\sin(\theta - \theta'')}{\sin(\theta + \theta'')}, \quad d_{\perp} = \frac{2 \sin \theta'' \cos \theta}{\sin(\theta + \theta'')}. \quad (1.7.2)$$

Здесь индекс « $\perp$ » указывает на s -поляризацию падающей волны (вектор $\mathbf{E}$ перпендикулярен плоскости падения).

Формулы (1.7.2) можно переписать в ином виде, если исключить угол преломления θ'' с помощью закона Снеллиуса $n_1 \sin \theta = n_2 \sin \theta''$ и ввести относительный показатель преломления $n_{21} = n_2/n_1$:

$$r_{\perp} = \frac{\cos \theta - \sqrt{n_{21}^2 - \sin^2 \theta}}{\cos \theta + \sqrt{n_{21}^2 - \sin^2 \theta}}, \quad d_{\perp} = \frac{2 \cos \theta}{\cos \theta + \sqrt{n_{21}^2 - \sin^2 \theta}}. \quad (1.7.3)$$

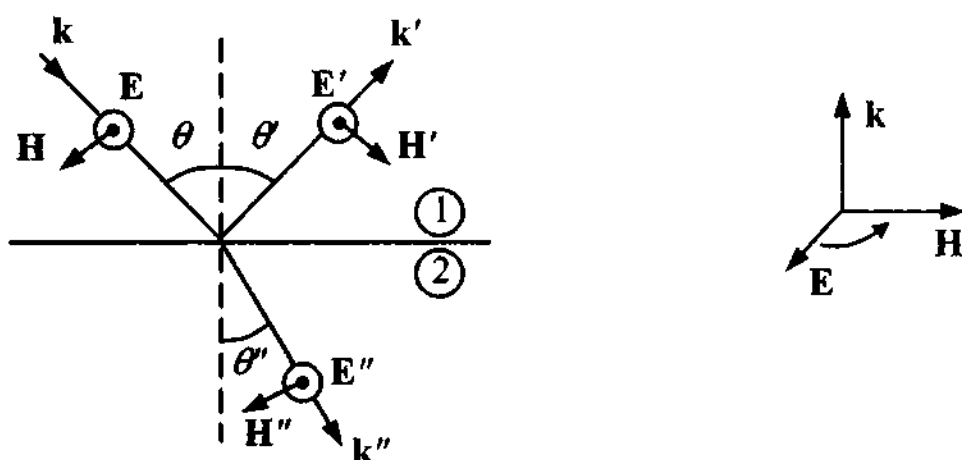


Рис. 1.7.2. Слева — отражение и преломление s -поляризованной волны на границе раздела двух диэлектриков: вектор $\mathbf{E}$ перпендикулярен плоскости рисунка, а вектор $\mathbf{H}$ лежит в плоскости рисунка; справа — правая тройка векторов $\{\mathbf{E}, \mathbf{H}, \mathbf{k}\}$. На рисунке кружок с крестом означает, что соответствующий вектор направлен «от нас»

1.7.4. p -поляризованная волна

В этом случае плоскости падения перпендикулярен вектор напряжённости магнитного поля $\mathbf{H}$. Примем условно, что векторы напряжённости электрического поля во всех волнах лежат в плоскости рисунка и направлены вправо, как показано на рис. 1.7.3 слева. Направление же вектора $\mathbf{H}$ определяется правилом буравчика (рис. 1.7.3 справа), так что в падающей и преломлённой волнах вектор $\mathbf{H}$ направлен «на нас», а в отражённой волне — «от нас». В данной ситуации амплитудные коэффициенты отражения и прохождения даются формулами

$$r_{\parallel} = -\frac{\operatorname{tg}(\theta - \theta'')}{\operatorname{tg}(\theta + \theta'')}, \quad d_{\parallel} = \frac{2 \sin \theta'' \cos \theta}{\sin(\theta + \theta'') \cos(\theta + \theta'')}. \quad (1.7.4)$$

Здесь индекс « $\parallel$ » указывает на p -поляризацию (вектор $\mathbf{E}$ параллелен плоскости падения волны).

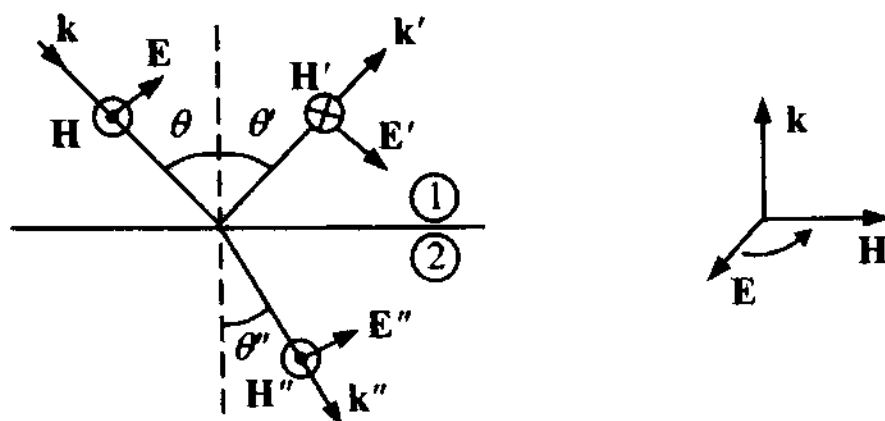


Рис. 1.7.3. Слева — отражение и преломление p -поляризованной волны на границе раздела двух диэлектриков: вектор $\mathbf{H}$ перпендикулярен плоскости рисунка, а вектор $\mathbf{E}$ лежит в плоскости рисунка; справа — правая тройка векторов $\{\mathbf{E}, \mathbf{H}, \mathbf{k}\}$

Исключив угол преломления θ'' с помощью закона Снеллиуса и введя относительный показатель преломления $n_{21} = n_2/n_1$, перепишем (1.7.4) в ином виде:

$$r_{\parallel} = -\frac{n_{21}^2 \cos \theta - \sqrt{n_{21}^2 - \sin^2 \theta}}{n_{21}^2 \cos \theta + \sqrt{n_{21}^2 - \sin^2 \theta}}, \quad d_{\parallel} = \frac{2 n_{21} \cos \theta}{n_{21}^2 \cos \theta + \sqrt{n_{21}^2 - \sin^2 \theta}}. \quad (1.7.5)$$

Соотношения (1.7.2) – (1.7.5) определяют амплитуды отражённой и преломлённой волн для двух случаев поляризации. В общем случае волну можно представить в виде суперпозиции s - и p -поляризованной компонент, и для каждой из компонент следует использовать соответствующие формулы Френеля.

Заметим, что из формул Френеля при скользющем падении волны ($\theta \rightarrow \pi/2$) независимо от поляризации следует

$$d \rightarrow 0, \quad r_{\perp} \rightarrow -1, \quad r_{\parallel} \rightarrow 1.$$

1.7.5. Нормальное падение волны на границу раздела сред

В частном случае нормального падения волны на границу раздела сред ($\theta = \theta' = 0$) исчезает различие между поляризациями волны, поскольку оба вектора ($\mathbf{E}$ и $\mathbf{H}$) параллельны границе раздела сред (рис. 1.7.4). При этом для амплитудных коэффициентов отражения и прохождения получаются формулы

$$r = \frac{n_1 - n_2}{n_1 + n_2} = \frac{1 - n_{21}}{1 + n_{21}}, \quad d = \frac{2n_1}{n_1 + n_2} = \frac{2}{1 + n_{21}}. \quad (1.7.6)$$

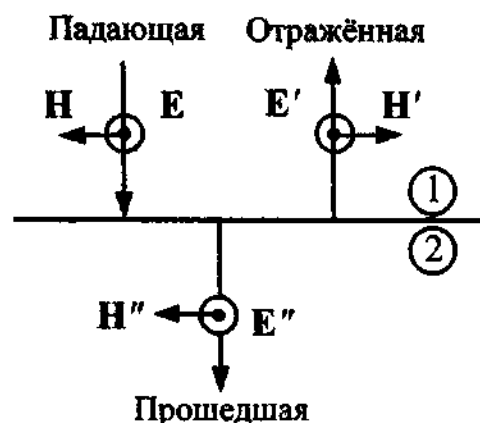


Рис. 1.7.4. Отражение и прохождение волны, падающей по нормали к границе

1.7.6. Коэффициенты отражения и прохождения по энергии

Если требуется найти зависимость интенсивности излучения, отражённого от границы раздела сред или прошедшего через эту границу, следует воспользоваться коэффициентами отражения и прохождения по энергии, определяемыми равенствами

$$R = \frac{(I_{\text{отр}})_z}{(I_{\text{пад}})_z}, \quad D = \frac{(I_{\text{прош}})_z}{(I_{\text{пад}})_z}, \quad (1.7.7)$$

где $(I_{\text{пад}})_z$, $(I_{\text{отр}})_z$ и $(I_{\text{прош}})_z$ — соответственно компоненты интенсивности падающей, отражённой и прошедшей волн, нормальные к границе раздела и согласно (1.5.5) связанные с r и d соотношениями

$$R = \frac{(cn_1 E_0'^2 / 8\pi) \cos \theta'}{(cn_1 E_0^2 / 8\pi) \cos \theta} = r^2, \quad D = \frac{(cn_2 E_0''^2 / 8\pi) \cos \theta''}{(cn_1 E_0^2 / 8\pi) \cos \theta} = \frac{n_2 \cos \theta''}{n_1 \cos \theta} d^2. \quad (1.7.8)$$

Для этих коэффициентов выполняется равенство

$$R + D = 1, \quad (1.7.9)$$

выражающее закон сохранения энергии

$$(I_{\text{пад}})_z = (I_{\text{отр}})_z + (I_{\text{прош}})_z.$$

Для случая нормального падения излучения на границу раздела среда согласно (1.7.6) имеем

$$R = r^2 = \left(\frac{n_1 - n_2}{n_1 + n_2} \right)^2, \quad D = \frac{n_2}{n_1} d^2 = \frac{4n_1 n_2}{(n_1 + n_2)^2}. \quad (1.7.10)$$

Выполнение соотношения (1.7.9) здесь очевидно.

1.7.7. Угловая зависимость коэффициента отражения

На рис. 1.7.5 показана угловая зависимость коэффициента отражения $R(\theta)$ для s -поляризованного ($R_{\perp}$) и p -поляризованного ($R_{\parallel}$) излучения.

Естественный свет является неполяризованным. Это значит, что s - и p -поляризованные компоненты присутствуют в нём с равным весом: $I_s = I_p = I/2$, где I — полная интенсивность. Интенсивность отражённого света можно записать в виде

$$I_{\text{отр}} = R_{\perp} I_{\perp} + R_{\parallel} I_{\parallel} = \frac{1}{2} (R_{\perp} + R_{\parallel}) I.$$

Обозначая $I_{\text{отр}} = R_{\text{ест}} I$, получаем коэффициент отражения для естественного света:

$$R_{\text{ест}} = \frac{1}{2} (R_{\perp} + R_{\parallel}). \quad (1.7.11)$$

Угловая зависимость $R_{\text{ест}}(\theta)$ показана на рис. 1.7.5.

В случае p -поляризованной волны существует такой угол $\theta = \theta_{\text{Б}}$, называемый *углом Брюстера*, что волна, падающая под этим углом на поверхность, не отражается обратно (рис. 1.7.5). Полагая в формулах Френеля (1.7.5) $r_{\parallel}(\theta_{\text{Б}}) = 0$, найдём формулу, определяющую этот угол:

$$\text{tg } \theta_{\text{Б}} = n_2 / n_1. \quad (1.7.12)$$

В частном случае падения волны из вакуума $n_1 = 1$ на среду с показателем преломления $n_2 = n$ оказывается $\text{tg } \theta_{\text{Б}} = n$.

При падении волны под углом Брюстера отражённая волна оказывается полностью s -поляризованной. Поэтому данный угол иногда называют *углом полной поляризации*.

Заметим, что если излучение поглощается в среде, то угол Брюстера определяется как угол, при котором коэффициент отражения $R_{\parallel}(\theta)$ достигает минимума, хотя в нуль и не обращается. При этом отражённая волна оказывается поляризованной лишь частично. Подробнее явление Брюстера обсуждается в разделе 9.4.

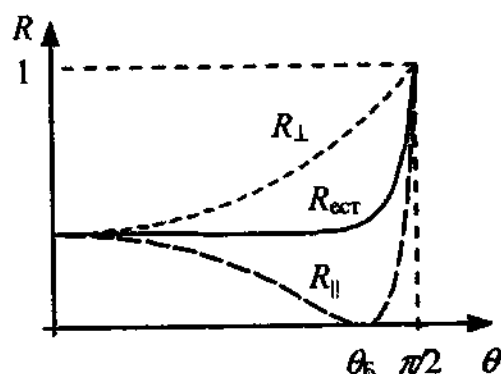


Рис. 1.7.5. Угловые зависимости коэффициентов отражения для излучения с s - и p -поляризацией, а также для естественного (неполяризованного) излучения

1.8. Сферические волны

Наряду с плоскими волнами часто нужно иметь дело с волнами сферическими: сферическая волна, *расходящаяся* из центра $r = 0$, может быть записана следующим образом:

$$A = A_0 \frac{e^{ikr - iat}}{r}. \quad (1.8.1)$$

Если волна *сходится* к центру $r = 0$, то она записывается в виде

$$A = A_0 \frac{e^{-ikr - iat}}{r}. \quad (1.8.2)$$

Строго говоря, электромагнитные волны не могут быть сферическими. Действительно, вследствие их поперечности невозможно всюду сферически симметричным способом направить векторы $\mathbf{E}$, $\mathbf{H}$, $\mathbf{k}$. Причиной этого является то, что элементарным источником электромагнитной волны является колеблющийся электрический диполь, а не сферически симметричный точечный источник. Однако использование такого приближения может быть удовлетворительным, если рассматривается поле вдали от источника в относительно небольшом диапазоне направлений.

Заметим, что электромагнитное поле диполя с дипольным моментом, меняющимся по гармоническому закону: $\mathbf{p} = \mathbf{p}_0 e^{-i\omega t}$, даётся в волновой зоне формулами

$$\mathbf{H} = k^2 (\mathbf{n} \times \mathbf{p}_0) \frac{e^{i(kr - \omega t)}}{r}, \quad \mathbf{E} = \mathbf{H} \times \mathbf{n}, \quad (1.8.3)$$

где $k = \omega/c$ — волновое число, $\mathbf{n} = \mathbf{r}/r$ — единичный вектор в направлении радиус-вектора. Иначе говоря, поле диполя подобно сферической волне, но с амплитудой, зависящей от угла между радиус-вектором точки наблюдения $\mathbf{r}$ и направлением дипольного момента $\mathbf{p}_0$.

ГЛАВА 2

ГЕОМЕТРИЧЕСКАЯ ОПТИКА

2.1. Уравнение эйконала

В общем случае решение волнового уравнения для поля, меняющегося с определённой частотой ω , не имеет вида плоской волны и может быть представлено следующим образом:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0(\mathbf{r}) \exp[i(\Phi(\mathbf{r}) - \omega t)], \quad (2.1.1)$$

где $\mathbf{E}_0$ и Φ — действительные функции. Фазой такой волны является величина $\varphi(\mathbf{r}, t) = \Phi(\mathbf{r}) - \omega t$. Функция $\Phi(\mathbf{r})$ называется *эйконалом* (от греч. *εικων* — образ, изображение, подобие). В случае плоской волны $\Phi(\mathbf{r}) = \mathbf{k} \cdot \mathbf{r}$.

Подстановка (2.1.1) в волновое уравнение (1.2.4) приводит к уравнению Гельмгольца:

$$\Delta \mathbf{E} + k^2 \mathbf{E} = 0. \quad (2.1.2)$$

Здесь

$$k = \frac{\omega}{c} n = nk_0, \quad k_0 = \frac{\omega}{c} = \frac{2\pi}{\lambda_0}, \quad (2.1.3)$$

причём для случая плоской волны величина λ_0 имеет смысл длины волны в вакууме.

Рассмотрим далее только одну компоненту поля, а соответствующую проекцию вектора $\mathbf{E}_0$ обозначим a :

$$E(\mathbf{r}, t) = a(\mathbf{r}) \exp[i(\Phi(\mathbf{r}) - \omega t)]. \quad (2.1.4)$$

Воспользуемся равенствами

$$\nabla E = [\nabla a + i a \nabla \Phi] e^{i(\Phi - \omega t)},$$

$$\Delta E = [\Delta a + 2i \nabla a \cdot \nabla \Phi - a(\nabla \Phi)^2 + i a \Delta \Phi] e^{i(\Phi - \omega t)}.$$

Подстановка найденного выражения для ΔE в уравнение Гельмгольца (2.1.2) даёт

$$-\frac{\omega^2}{c^2} n^2 a = [\Delta a - a(\nabla \Phi)^2] + i[2 \nabla a \cdot \nabla \Phi + a \Delta \Phi].$$

Отделяя здесь действительную и мнимую части, приходим к системе двух уравнений для функций a и Φ :

$$\Delta a - a(\nabla\Phi)^2 = -n^2 k_0^2 a, \quad (2.1.5)$$

$$2\nabla a \cdot \nabla\Phi + a\Delta\Phi = 0.$$

Второе уравнение в (2.1.5) можно переписать в виде

$$\operatorname{div} \mathbf{j} = 0, \quad \mathbf{j} = a^2 \nabla\Phi \equiv a^2 \mathbf{k}. \quad (2.1.6)$$

Введённый здесь вектор $\mathbf{j}$ пропорционален вектору Пойнтинга:

$\mathbf{j} = (8\pi\omega/c^2) \mathbf{S}$ и указывает направление переноса световой энергии. Соот-

ветственно равенство (2.1.6) можно интерпретировать как уравнение непрерывности для светового потока. По аналогии с гидродинамикой введём «линии тока»: касательные к ним в каждой точке совпадают с направлением вектора $\mathbf{j}$ в этих точках (рис. 2.1.1, слева). Из «линий тока» строится «трубка тока» как цилиндрическая поверхность, образующие которой начинаются с некоторого замкнутого контура Γ_0 (рис. 2.1.1, справа). Если Π — площадь поперечного сечения трубки в каком-либо месте, то уравнение (2.1.6) можно переписать в интегральной форме:

$$j\Pi = \text{const} \quad \text{или} \quad a^2 n\Pi = \text{const}. \quad (2.1.7)$$



Рис. 2.1.1. Слева — луч, «линия тока», справа — «трубка тока», начальное сечение которой образует контур Γ_0 , от точек которого исходят «линии тока»

Перепишем теперь первое уравнение в (2.1.5) в следующем виде:

$$(\nabla\Phi)^2 = k^2 + \frac{\Delta a}{a}, \quad (2.1.8)$$

где $k = nk_0 = \omega n/c$.

Введём величину $\Psi = \Phi/k_0$, называемую *приведённым эйконалом*. Тогда первое из равенств в (2.1.5) даёт уравнение

$$(\nabla\Psi)^2 = n^2 + \frac{\Delta a}{k_0^2 a}. \quad (2.1.9)$$

Последнее соотношение позволяет исследовать переход к пределу геометрической оптики, когда $\lambda \rightarrow 0$, $k_0 \rightarrow \infty$:

$$(\nabla\Psi)^2 = n^2, \quad (2.1.10)$$

или, возвращаясь к эйконалу Φ ,

$$(\nabla\Phi)^2 = k^2. \quad (2.1.11)$$

Равенство (2.1.11) называется *уравнением эйконала*. Оно применимо (в пределе $|\Delta a/a| \ll k_0^2$) для описания поля не только в однородных средах, но и тогда, когда показатель преломления меняется от точки к точке: $n \neq \text{const}$, если только скорость его изменения достаточно мала:

$$\lambda |\nabla n| \ll \delta n,$$

где λ — длина волны, δn — характерный масштаб изменения показателя преломления n в условиях задачи.

2.2. Волновой фронт и лучи

Согласно (2.1.1) электромагнитная волна в момент времени t имеет одинаковую фазу во всех точках пространства, в которых эйконал $\Phi(\mathbf{r})$ принимает одно и то же значение:

$$\Phi(\mathbf{r}) = \text{const}. \quad (2.2.1)$$

Это уравнение определяет поверхность в пространстве, называемую *волновым фронтом*. С течением времени волновой фронт смещается, то есть смещается поверхность, на которой волна имеет заданное значение фазы $\varphi_0 = \Phi_t(\mathbf{r}) - \omega t$:

$$\Phi_t(\mathbf{r}) = \varphi_0 + \omega t.$$

Вектор нормали к фронту волны $\Phi(\mathbf{r}) = \text{const}$ параллелен вектору

$$\mathbf{k} = \text{grad } \Phi(\mathbf{r}) = k_0 \text{grad } \Psi(\mathbf{r}), \quad (2.2.2)$$

имеющему смысл *волнового вектора* в точке $\mathbf{r}$ фронта (рис. 2.2.1). В самом деле, в окрестности произвольной точки $\mathbf{r}_0$ эйконал может быть представлен в виде

$$\begin{aligned} \Phi(\mathbf{r}) &\approx \Phi(\mathbf{r}_0) + (x - x_0) \left(\frac{\partial \Phi}{\partial x} \right)_{\mathbf{r}=\mathbf{r}_0} + (y - y_0) \left(\frac{\partial \Phi}{\partial y} \right)_{\mathbf{r}=\mathbf{r}_0} + (z - z_0) \left(\frac{\partial \Phi}{\partial z} \right)_{\mathbf{r}=\mathbf{r}_0} = \\ &= \Phi(\mathbf{r}_0) + \mathbf{k}(\mathbf{r} - \mathbf{r}_0), \quad \mathbf{k} = \text{grad } \Phi(\mathbf{r})|_{\mathbf{r}=\mathbf{r}_0}, \end{aligned}$$

так что фаза волны

$$\varphi(\mathbf{r}, t) = \Phi(\mathbf{r}) - \omega t \approx \varphi_0 + \mathbf{k}\mathbf{r} - \omega t, \quad \varphi_0 = \Phi(\mathbf{r}_0) + \mathbf{k}\mathbf{r}_0. \quad (2.2.3)$$

Последнее означает, что в достаточно малой окрестности точки $\mathbf{r}_0$ фронта волна совпадает с плоской волной с волновым вектором (2.2.2), направленным по нормали к фронту.

Линии, касательные к которым в каждой точке параллельны волновому вектору $\mathbf{k}$ в этой точке, называются *лучами*. Иначе говоря, лучи являются огибающими векторного поля $\mathbf{k}(\mathbf{r})$. Таким образом, лучи *ортогональны* волновому фронту и определяют направление переноса энергии излучения (рис. 2.2.2). С точки зрения упомянутой выше аналогии с гидродинамикой лучи — это линии тока.

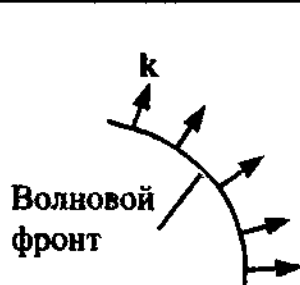


Рис. 2.2.1. Волновой фронт $\Phi(\mathbf{r}) = \text{const}$ и лучи. Стрелки указывают направления лучей, задаваемые волновым вектором $\mathbf{k} = \text{grad } \Phi(\mathbf{r})$ в точках фронта

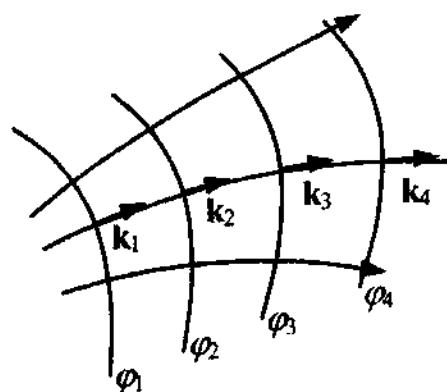


Рис. 2.2.2. Волновые фронты, соответствующие значениям фазы $\varphi_1, \varphi_2, \varphi_3, \varphi_4$, и лучи в некоторый момент времени

Покажем, что скорость движения фронта волны совпадает со скоростью света в среде. В самом деле, положение фронта (то есть поверхности с заданным значением фазы) φ_0 определяется уравнением $\Phi_t(\mathbf{r}) - \omega t = \text{const} \equiv \varphi_c$. Дифференцируя это уравнение, получим $d\Phi_t = \omega dt$. При смещении вдоль луча на $d\mathbf{r}$ эйконал получает приращение

$$d\Phi_t = d\mathbf{r} \nabla \Phi_t = \frac{\partial \Phi_t}{\partial s} ds = nk_0 ds.$$

Здесь учтено, что длина элемента луча равна $|d\mathbf{r}| = ds$, а производная от Φ_t по нормали к волновому фронту согласно уравнению эйконала равна по величине $|\nabla \Phi_t| = nk_0$. Следовательно, $nk_0 ds = \omega dt$, откуда

$$\frac{ds}{dt} = \frac{\omega}{nk_0} = \frac{c}{n} = v. \quad (2.2.4)$$

Заметим, что во многих случаях приближение эйконала неприменимо. Рассмотрим, например, суперпозицию двух одинаковых монохроматических волн, но распространяющихся в разных направлениях:

$$A(\mathbf{r}, t) = a_0 e^{i(\mathbf{k}_1 \mathbf{r} - \omega t)} + a_0 e^{i(\mathbf{k}_2 \mathbf{r} - \omega t)} = 2a_0 e^{i(\mathbf{K} \mathbf{r} - \omega t)} \cos \mathbf{q} \mathbf{r}.$$

Здесь $\mathbf{K} = (\mathbf{k}_1 + \mathbf{k}_2)/2$, $\mathbf{q} = (\mathbf{k}_1 - \mathbf{k}_2)/2$. Видно, что волновое число волны $|\mathbf{K}| = |\mathbf{k}_1 + \mathbf{k}_2|/2 = k \cos(\theta/2)$, где θ — угол между волновыми векторами, может сильно отличаться от волнового числа k исходных волн. Это связано с тем, что амплитуда волны $a(\mathbf{r}) = 2a_0 \cos \mathbf{q} \mathbf{r}$ сильно зависит от координаты: $|\Delta a/a| = q^2$. При $q \sim k$ это делает невозможным переход к пределу геометрической оптики $k \rightarrow \infty$.

Однако для описания поля колеблющегося диполя в волновой зоне (1.8.3) использование приближения эйконала допустимо: можно ввести

эйконол, обладающий сферической симметрией ($\Phi = kr$), причём в пределе геометрической оптики ($k \rightarrow \infty$) для амплитуды $\mathbf{a}(\mathbf{r}) = k^2(\mathbf{n} \times \mathbf{p}_0)/r$ выполняется неравенство $|\Delta a/a| \ll k^2$.

2.3. Принцип Ферма

Формулировка 1. Луч света распространяется в пространстве между двумя точками по тому пути, вдоль которого время его прохождения меньше, чем вдоль любого другого пути, соединяющего эти точки.

Французский математик Пьер Ферма сформулировал этот принцип в 1662 г., исходя из теологических соображений о целенаправленности действий природы, о том, что цель должна достигаться с наименьшими затратами средств. Ферма также получил ряд важных результатов в физике и математике. Ему принадлежит и формулировка (1637 г.) великой теоремы Ферма (о неразрешимости в целых числах уравнения $x^n + y^n = z^n$ при всех $n > 2$), окончательно доказанной лишь в 1994 г. Эндрю Уайлсом.

Наряду с геометрической длиной пути светового луча между точками M_0 и M

$$l(M_0, M) = \int_{M_0}^M ds \quad (2.3.1)$$

важное значение в оптике имеет величина

$$L(M_0, M) = \int_{M_0}^M n ds, \quad (2.3.2)$$

называемая *оптической длиной пути* M_0M ($ds = |d\mathbf{r}|$ — элемент длины дуги траектории луча). Поскольку скорость света в среде равна $v = c/n$, то, полагая в (2.3.2) для луча $ds = vdt$, получим

$$L(M_0, M) = \int_{M_0}^M n v dt = c \int_{M_0}^M dt = cT. \quad (2.3.2a)$$

Таким образом, оптическая длина пути пропорциональна времени распространения луча по траектории M_0M . Это соотношение позволяет дать другую формулировку принципа Ферма.

Формулировка 2. Свет распространяется между двумя точками по той траектории, оптическая длина которой минимальна.

Строго говоря, отдельно следует рассматривать ситуации, когда имеется препятствие, делающее выбор траектории неоднозначным, как это проиллюстрировано на рис. 2.3.1. Поэтому в формулировке принципа Ферма рассматриваются односвязные области пространства, то есть пред-

полагается, что выбор осуществляется между кривыми, которые могут быть непрерывным образом переведены одна в другую.

Доказательство. Докажем принцип Ферма (в формулировке 2), используя полученное выше уравнение эйконала (2.1.11). Рассмотрим инте-

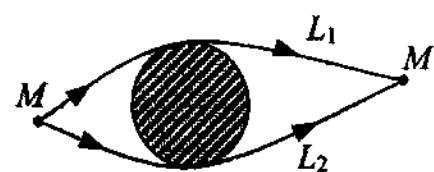


Рис. 2.3.1. Траектории лучей L_1 и L_2 , огибающие препятствие с разных сторон

грал $J = \int_{M_0}^M k dr$, вычисленный вдоль произвольной кривой, соединяющей точки M_0 и M (рис. 2.3.2). Согласно (2.2.2) $k = \text{grad } \Phi$, так что

$$J = \int_{M_0}^M \text{grad} \Phi dr = \Phi(M) - \Phi(M_0),$$

то есть этот интеграл не зависит от формы кривой, а определяется только начальной и конечной точками.

Для траектории L_0 , совпадающей с траекторией луча, векторы k и dr параллельны, то есть

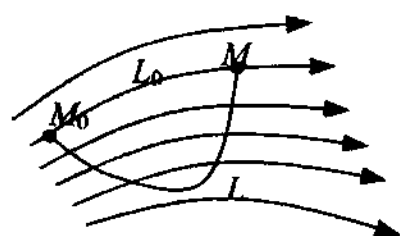


Рис. 2.3.2. К доказательству принципа Ферма

$$J = \int_{M_0}^M k ds = k_0 L_0(M_0, M),$$

где L_0 — оптическая длина участка M_0M траектории луча. Если взять какую-либо иную кривую L , соединяющую точки M_0 и M , то вдоль неё векторы k и dr не параллельны, так что $k ds \geq k dr$. Поэтому вдоль этой кривой окажется

$$k_0 L = k_0 \int_{M_0}^M n ds = \int_{M_0}^M k ds \geq \int_{M_0}^M k dr = k_0 L_0(M_0, M),$$

откуда и следует неравенство $L \geq L_0$.

Из принципа Ферма следует утверждение: если свет может идти из точки 1 в точку 2 через некоторую оптическую систему разными путями, то их оптические длины одинаковы. Говорят, что соответствующие лучи таутохронны (от греч. *ταυτο* — одинаковый и *χρονοζ* — время). Примером может служить идеальная линза, собирающая в одной точке лучи, исходящие из другой точки (рис. 2.3.3).

Примером системы, где реализуется свойство таутохронности, является также внутренняя отражающая поверхность эллипсоида вращения (рис. 2.3.4). Это свойство связано с тем, что эллипс — это фигура, сумма расстояний каждой точки которой до фокусов есть величина постоянная. Поэтому луч может распространяться от одного фокуса до другого множеством различных путей, имеющих одинаковую оптическую длину.

Следует иметь в виду, что уравнение эйконала (2.1.11) и следствия из него (в частности, принцип Ферма) справедливы, строго говоря, только в пределе геометрической оптики $\lambda \rightarrow 0$, пока не начинают проявляться волновые свойства света (чему соответствует пренебрежение слагаемым $\Delta a/a$ в уравнении (2.1.8)). В этом же пределе имеет смысл говорить и о лучах. В тех же случаях, когда длину волны нельзя считать малой, существенными могут оказаться волновые свойства света.

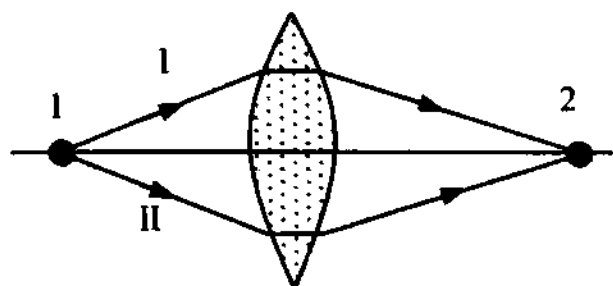


Рис. 2.3.3. Лучи, идущие из точки 1 в точку 2 через фокусирующую систему (линзу), таутохронны

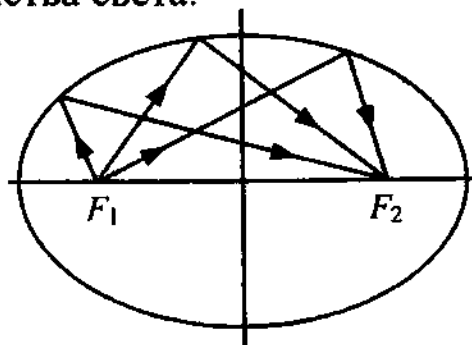


Рис. 2.3.4. Лучи, выходящие из одного фокуса эллипсоида (F_1), после отражения собираются в другом фокусе (F_2)

Пример 1. Покажем, что закон отражения (1.6.1) справедлив не только для плоских волн, но и для лучей и прямо следует из принципа Ферма.

Действительно, в этом случае луч проходит весь путь в одной среде, и геометрическая и оптическая длины отличаются лишь постоянным множителем n . В этих условиях найти кратчайший путь от точки 1 до точки 2 (рис. 2.3.5) можно простым построением. Отразим мысленно конечную точку 2 в $2'$ относительно границы раздела. Тогда видно, что кратчайший путь $1 \rightarrow 2'$ есть прямая. Вследствие симметрии точек 2 и $2'$ кратчайший путь $1 \rightarrow 2$ удовлетворяет закону: «угол падения равен углу отражения». Для сравнения на рис. 2.3.5 показаны истинная траектория луча $1 \rightarrow 2$ (сплошная линия) и одна из возможных траекторий, соединяющих точки 1 и 2 (штриховая линия).

Пример 2. Получим из принципа Ферма закон преломления света (1.6.2) на границе раздела сред с разными значениями показателя преломления (рис. 2.3.6).

Пусть точка M_1 находится на расстоянии h_1 от границы раздела сред (оси X), а точка M_2 — на расстоянии h_2 . Тогда в обозначениях рис. 2.3.6 оптическая длина пути $M_1 M_2$ равна

$$L = L_1 + L_2 = \frac{n_1 h_1}{\cos \varphi} + \frac{n_2 h_2}{\cos \psi}, \quad (2.3.3)$$

Учтём, что расстояние между точками M_1 и M_2 вдоль оси X неизменно:

$$X \equiv h_1 \operatorname{tg} \varphi + h_2 \operatorname{tg} \psi = \text{const}. \quad (2.3.4)$$

Последнее соотношение фиксирует связь между углами φ и ψ . Минимальная оптическая длина $L_{\min}$ определится из условия $dL/d\varphi = 0$, в котором дифференцирование производится с учетом связи (2.3.4):

$$\frac{dL}{d\varphi} = -n_1 h_1 \frac{\sin \varphi}{\cos^2 \varphi} - n_2 h_2 \frac{\sin \psi}{\cos^2 \psi} \frac{d\psi}{d\varphi} = 0,$$

$$\frac{dX}{d\varphi} = \frac{h_1}{\cos^2 \varphi} + \frac{h_2}{\cos^2 \psi} \frac{d\psi}{d\varphi} = 0.$$

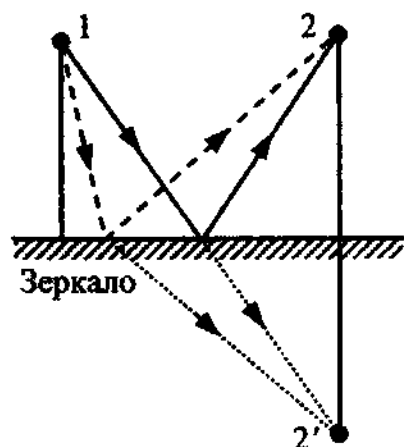


Рис. 2.3.5. К выводу закона отражения с помощью принципа Ферма

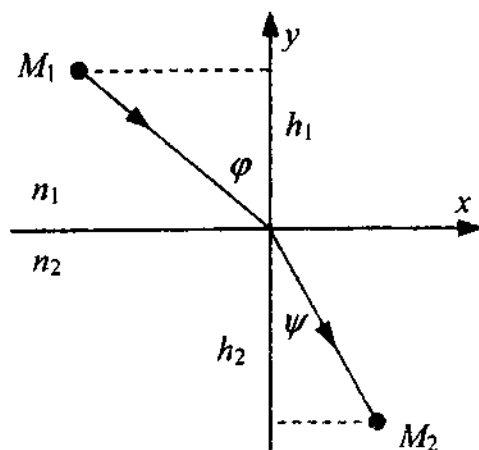


Рис. 2.3.6. К выводу закона преломления с помощью принципа Ферма

Исключая из этих равенств производную $d\psi/d\varphi$, приходим к закону Снеллиуса $n_1 \sin \varphi = n_2 \sin \psi$.

2.4. Основные законы геометрической оптики

Сформулируем основные законы, на которых базируется геометрическая оптика.

1. *Закон прямолинейного распространения света*: в оптически однородной среде ($n = \text{const}$) световые лучи распространяются прямолинейно.

2. *Закон отражения*: $\varphi = \varphi'$ (φ — угол падения, φ' — угол отражения, рис. 2.4.1).

3. *Закон преломления (закон Снеллиуса)*: $n \sin \varphi = n' \sin \varphi''$ (φ'' — угол преломления, рис. 2.4.1).

4. *Принцип независимости световых лучей*: лучи света при пересечении не возмущают друг друга. Этот принцип справедлив лишь для относительно небольших интенсивностей света, при которых изменениями свойств среды под действием электромагнитного поля можно пренебречь.

5. *Принцип обратимости хода лучей*: путь светового луча (ABC) из начальной точки A в конечную C, распространяющегося в какой-либо оп-

тической среде, заменяется на прямо противоположный путь (CBA), если свет исходит из конечной точки C в начальную A в направлении, противоположном первоначальному (рис. 2.4.2).

Сформулированный принцип связан с инвариантностью уравнений Максвелла относительно изменения знака времени: в случае непроводящих сред уравнения Максвелла не меняются при одновременной смене знаков у одной из пар величин: $(t, \mathbf{B})$ или $(t, \mathbf{E})$ (второе из полей при данном преобразовании не должно менять знак). Указанная инверсия эквивалентна смене направления распространения волны на противоположное: $\mathbf{k} \rightarrow -\mathbf{k}$, что следует из свойства поперечности электромагнитной волны (рис. 1.1.1).

2.5. Основные определения

1. Под *оптической осью* оптической системы (включающей линзы, зеркала и так далее) имеют в виду прямую линию, являющуюся осью симметрии преломляющих (отражающих) поверхностей. Она проходит перпендикулярно к этим поверхностям через их центры кривизны.

2. *Центрированной оптической системой* называют совокупность однородных преломляющих и отражающих сред, отделённых друг от друга симметричными поверхностями, центры кривизны которых находятся на одной прямой. Эту прямую называют *главной оптической осью системы*.

3. Совокупность лучей образует *пучок*.

Если лучи при своём продолжении сходятся в одной точке, то пучок называется *гомоцентрическим* (рис. 2.5.1). В оптически однородной среде гомоцентрическому пучку отвечает сферическая волновая поверхность.



Рис. 2.5.1. Слева — сходящийся гомоцентрический пучок; справа — расходящийся гомоцентрический пучок. В обоих случаях лучи стартуют от волнового фронта, показанного сплошной кривой

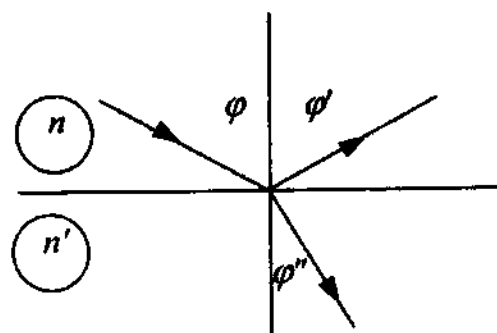


Рис. 2.4.1. Отражение и преломление лучей на границе раздела сред

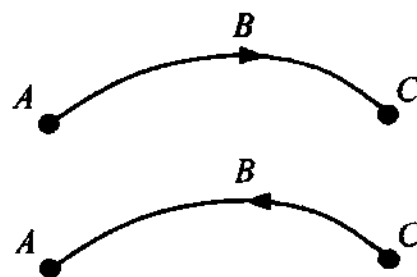
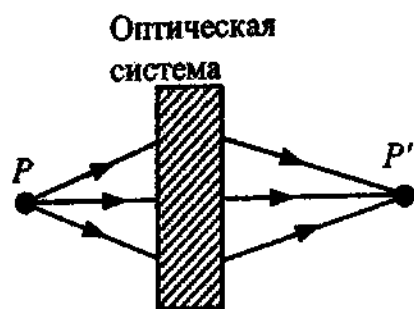


Рис. 2.4.2. Обратимость хода лучей: если возможен путь светового луча ABC (верхняя траектория), то возможен и прямо противоположный путь CBA (нижняя траектория)

Пучок называется *параксиальным*, если все образующие его лучи распространяются вдоль оси центрированной оптической системы и образуют малые углы с осью и нормальными к преломляющим и отражающим поверхностям системы.

4. Пусть оптическая система преобразует свет, не нарушая гомоцентричности пучка, причём пучок лучей, исходящий из точки P , сходится в точке P' (рис. 2.5.2). Тогда точка P' называется *изображением* точки P .



Если точка P' является источником лучей, то в силу принципа обратимости изображением окажется точка P . Поэтому точки P' и P называются *сопряжёнными* (взаимно сопряжёнными). Аналогично, сопряжёнными называются две линии (две поверхности), если одна из них является изображением другой.

Рис. 2.5.2. Лучи, идущие из точки P в точку P' через оптическую систему

5. Изображение называется *действительным* (рис. 2.5.3, слева), если световые лучи от точки P сходятся в точке P' при своем распространении. Если же в точке P' сходятся продолжения лучей в направлении, обратном направлению распространения света, то изображение называется *мнимым* (рис. 2.5.3, справа).

6. Если любая точка предмета изображается в виде точки, то изображение называется *точечным* или *стигматическим*.

7. Непрерывное множество точек, изображаемых оптической системой, называется *пространством предметов*. Совокупность точек, являющихся их изображением, называется *пространством изображений*.

8. Оптическая система, которая даёт стигматическое изображение, геометрически подобное изображаемому предмету, называется *идеальной*.

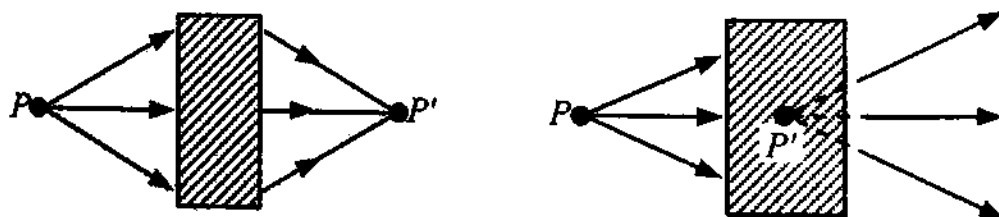


Рис. 2.5.3. Слева — оптическая система даёт действительное изображение (P') точки P ; справа — точка P' является мнимым изображением точки P

9. Пусть оптическая система превращает *параллельный* световой пучок в сходящийся (рис. 2.5.4, слева). Точка, в которой пересекаются лучи после прохождения системы, называется *фокусом* (или *действительным фокусом*). Если оптическая система превращает параллельный пучок в

расходящийся (рис. 2.5.4, справа), то точка пересечения продолжения лучей называется *мнимым фокусом*.

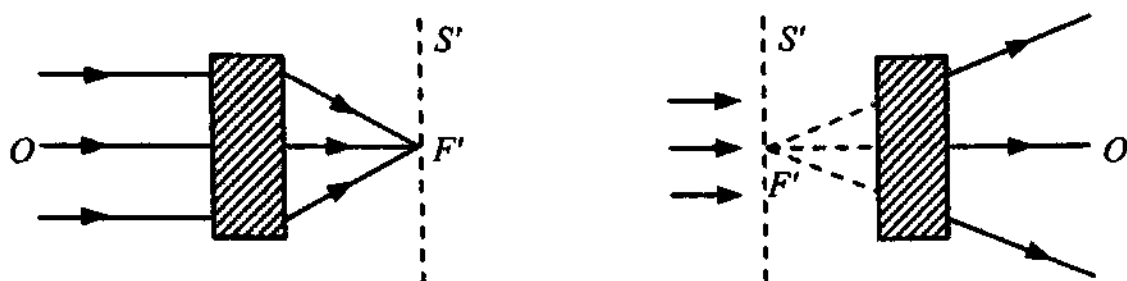


Рис. 2.5.4. Слева — параллельный пучок собирается в действительном фокусе F оптической системы. Справа — параллельный пучок преобразуется в расходящийся, причём продолжения лучей (в обратном направлении) сходятся в мнимом фокусе F' . O — главная оптическая ось, S' — задняя фокальная плоскость

10. Фокус F' называется *задним фокусом*. Плоскость S' , проходящая через задний фокус перпендикулярно главной оптической оси, называется *задней фокальной плоскостью* (рис. 2.5.4, слева).

11. *Передний фокус* — это такая точка F в пространстве предметов, что исходящий из неё пучок лучей превращается оптической системой в параллельный пучок (рис. 2.5.5).

12. Плоскость S , проходящая через передний фокус F перпендикулярно главной оптической оси, называется *передней фокальной плоскостью* (рис. 2.5.5).

13. *Линейное (или поперечное) увеличение* — это отношение линейных размеров изображения и предмета (рис. 2.5.6):

$$\beta = y'/y.$$

Отрезки y и y' считаются положительными, если они откладываются вверх от оптической оси, и отрицательными — в противоположном случае.

Увеличение считается положительным: $\beta > 0$, если величины y и y' имеют одинаковые знаки. Если же y и y' имеют противоположные знаки, то полагают $\beta < 0$. На рис. 2.5.6:

$$y > 0, y' < 0, \beta < 0.$$

Если увеличение положительное, то изображение *прямое*. В противоположном случае изображение *обратное*.

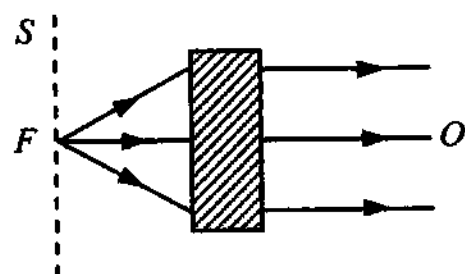


Рис. 2.5.5. Пучок лучей, исходящий из переднего фокуса F , преобразуется в параллельный пучок. O — главная оптическая ось, S — передняя фокальная плоскость

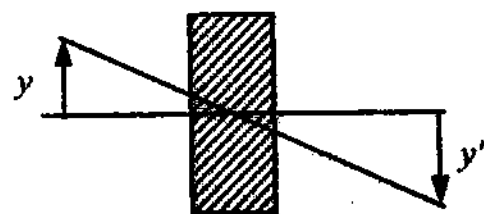


Рис. 2.5.6. К определению линейного увеличения

14. Две сопряжённые плоскости, отображающиеся с линейным увеличением $\beta = +1$, называются *главными* (рис. 2.5.7). Одна из них называется

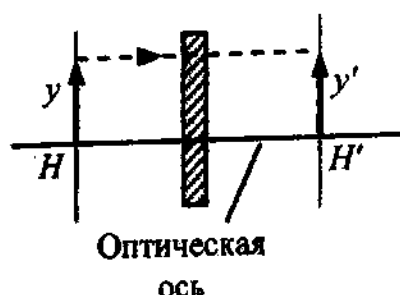


Рис. 2.5.7. Главные плоскости H и H' и главные точки оптической системы; $y = y'$

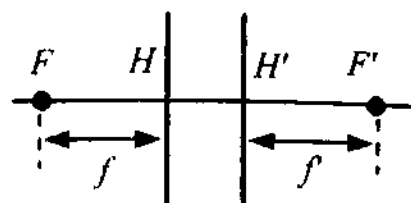


Рис. 2.5.8. Главные фокусные расстояния f и f' оптической системы

главной плоскостью пространства предметов (или *передней главной плоскостью*) H , а вторая — *главной плоскостью пространства изображений* (или *задней главной плоскостью*) H' .

Точки пересечения главных плоскостей с главной оптической осью называются *главными точками* оптической системы (рис. 2.5.7).

15. Расстояния от главных точек до соответствующих фокусов называются *главными фокусными расстояниями* (рис. 2.5.8). Фокусные расстояния равны

$$f = FH, \quad f' = F'H'.$$

Правило знаков: эти расстояния считаются положительными, если свет идёт от главной плоскости к соответствующему главному фокусу. На рис. 2.5.8

$$f < 0, \quad f' > 0.$$

16. *Угловым увеличением* называется отношение

$$\gamma = \operatorname{tg} \varphi' / \operatorname{tg} \varphi,$$

где φ' — угол наклона луча, вышедшего из оптической системы в пространство изображений, φ — угол наклона сопряжённого ему луча в пространстве предметов (рис. 2.5.9).

Правило знаков: углы φ и φ' отсчитываются в направлении от главной оси к лучу против часовой стрелки. На рис. 2.5.9 в соответствии с соглашением

$$\varphi > 0, \quad \varphi' < 0.$$

17. Сопряжённые точки, в которых $\varphi = +\varphi'$, называются *узловыми*. На рис. 2.5.10 узловые точки — это точки N и N' . Лучи, проходящие через узловые точки, попарно параллельны: $1 \parallel 1'$, $2 \parallel 2'$.

18. Рассматривают также *видимое увеличение*. Так называют отношение тангенсов углового размера предмета, наблюдаемого с помощью оп-

тической системы, к угловому размеру предмета, наблюдаемого непосредственно, без оптической системы (рис. 2.5.11):

$$\Gamma = \operatorname{tg} \varphi' / \operatorname{tg} \varphi.$$

Иногда величину Γ определяют соотношением $\Gamma = \varphi' / \varphi$. При малых углах оба определения, очевидно, совпадают.

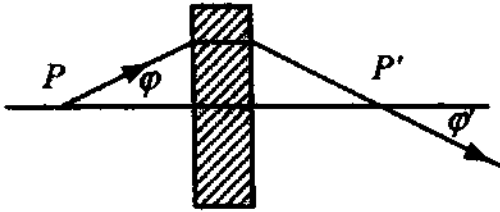


Рис. 2.5.9. К определению углового увеличения оптической системы.
 P и P' — сопряжённые точки

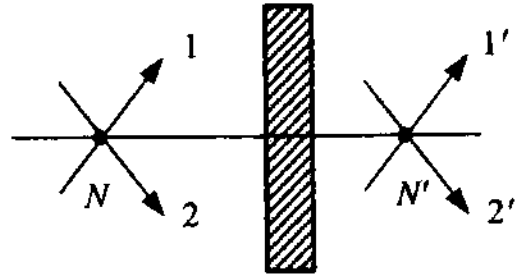


Рис. 2.5.10. Узловые точки N и N'

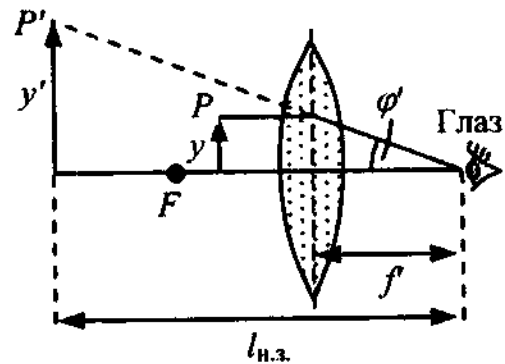
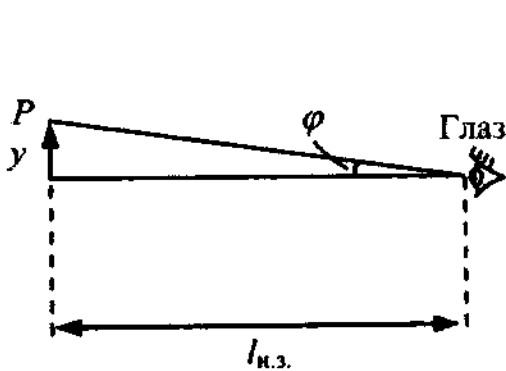


Рис. 2.5.11. К определению видимого увеличения. Слева — наблюдение предмета невооружённым глазом, справа — наблюдение того же предмета через линзу

Обычно предмет помещают за передним фокусом F (между фокусом и лупой). При этом формируется увеличенное прямое изображение предмета.

Особенностью зрения является то, что существует наименьшее расстояние, с которого лучше всего рассматривать предметы (в частности, глаз почти не утомляется). В типичных условиях это *расстояние наилучшего зрения* составляет $l_{н.з.} \approx 25$ см.

Пусть имеется тонкая линза (рис. 2.5.11). Тогда, рассматривая предмет глазом непосредственно, имеем $\operatorname{tg} \varphi = y / l_{н.з.}$. Если же с расстояния $l_{н.з.}$ рассматривается изображение, даваемое линзой, то $\operatorname{tg} \varphi' = y' / l_{н.з.}$. Отсюда находим

$$\Gamma = y' / y = \beta.$$

Это значит, что видимое увеличение совпадает поперечным увеличением.

Обычно глаз располагается вблизи заднего фокуса линзы, так что $l_{н.з.} \approx f'$. Поэтому, как видно из рис. 2.5.11, $\operatorname{tg} \varphi' \approx y / f'$. Соответственно находим

$$\Gamma \approx \frac{y/f'}{y/l_{н.з.}} = \frac{l_{н.з.}}{f'}.$$

В частности, линза с фокусным расстоянием $f = f' = 5$ см даёт увеличение $\Gamma \approx 5$.

2.6. Кардинальные точки оптической системы. Построение изображения

Фокальные и главные точки центрированной системы называются её *кардинальными точками* (рис. 2.6.1).

Центрированная оптическая система полностью характеризуется кардинальными точками. В общем случае имеется четыре различные кардинальные точки. В частных случаях это число может уменьшаться. Например, в случае тонкой линзы главные плоскости сливаются. В случае систем, содержащих одну отражающую поверхность, имеется только одна главная плоскость и один фокус.

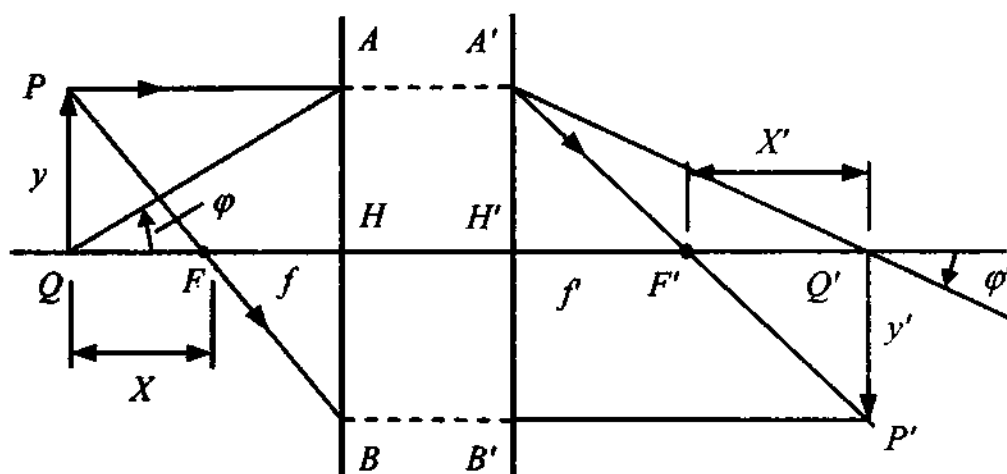


Рис. 2.6.1. Главные плоскости и фокусы оптической системы. Построение изображения $P'Q'$ предмета PQ , используя фокусы и главные плоскости оптической системы

Набор кардинальных точек может состоять не только из главных точек и фокусов. Можно, например, вместо главных точек использовать узловые. Если значения показателей преломления первой и последней сред совпадают, то узловые точки совпадают с главными. При этом оба набора кардинальных точек окажутся тождественными.

Знание кардинальных точек позволяет строить изображение произвольного предмета (рис. 2.6.1). Для этого из точек предмета проводят два луча: один — параллельно главной оптической оси, а второй — через передний фокус. Первый луч, достигнув главной плоскости, преобразуется в луч, идущий через задний фокус, а второй — в луч, идущий параллельно оптической оси. Точка пересечения вторичных лучей (или их продолже-

ний в обратном направлении) даёт изображение (действительное или мнимое) соответствующих точек предмета.

2.7. Формула Ньютона

Согласно рис. 2.6.1 для треугольников PQF и BHF имеем

$$\frac{QP}{HB} \equiv \frac{y}{-y'} = \frac{-X}{-f}. \quad (2.7.1)$$

Аналогично для треугольников $P'Q'F'$ и $A'H'F'$ находим

$$\frac{Q'P'}{H'A'} \equiv \frac{-y'}{y} = \frac{X'}{f'}. \quad (2.7.2)$$

Правило знаков: в этих формулах расстояния от переднего фокуса до предмета (X) и от заднего фокуса до изображения (X') отсчитываются от соответствующего фокуса в направлении распространения луча. Поэтому для представленной на рис. 2.6.1 ситуации

$$X < 0, X' > 0.$$

Сравнение выражений (2.7.1), (2.7.2) даёт

$$\frac{X}{f} = \frac{f'}{X'}, \text{ или } XX' = ff'. \quad (2.7.3)$$

Последнее соотношение называется *формулой Ньютона*. В частном случае, когда среды по обе стороны от системы одинаковы (одинаковы их показатели преломления), имеем $f = -f'$, и из (2.7.1) следует

$$XX' = -f^2. \quad (2.7.4)$$

2.8. Увеличение

Найдём связь увеличения оптической системы с фокусными расстояниями и положениями предмета и изображения (рис. 2.8.1).

2.8.1. Поперечное линейное увеличение

$$\beta = y'/y. \quad (2.8.1)$$

Согласно рис. 2.8.1 для прямоугольных треугольников PQF и BOF (с общей вершиной F) имеем

$$\frac{QP}{OB} \equiv \frac{y}{-y'} = \frac{-X}{-f} \Rightarrow \beta = -\frac{f}{X}. \quad (2.8.2)$$

Аналогичное соотношение записываем для прямоугольных треугольников $P'Q'F'$ и $A'O'F'$ с общей вершиной F' :

$$\frac{Q'P'}{O'A'} \equiv \frac{-y'}{y} = \frac{X'}{f'} \Rightarrow \beta = -\frac{X'}{f'}. \quad (2.8.3)$$

Последнее соотношение следует также из (2.8.2) и формулы Ньютона $XX' = ff'$.

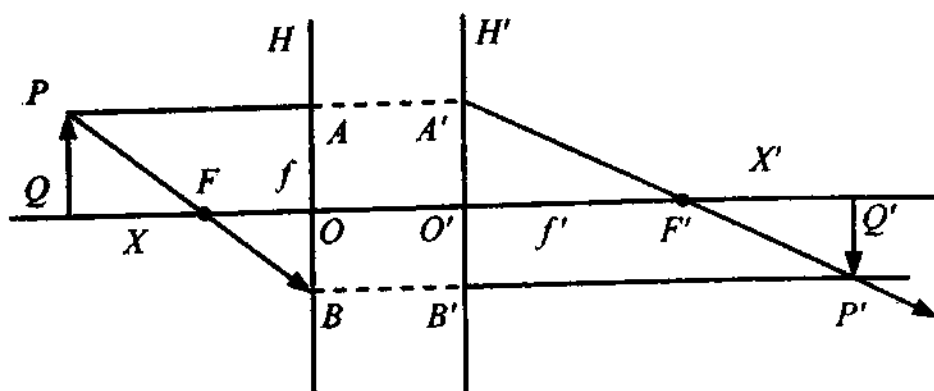


Рис. 2.8.1. К расчёту увеличения оптической системы

2.8.2. Продольное линейное увеличение

Наряду с поперечным увеличением рассматривают также увеличение *продольное*. Так называют отношение длины изображения $\delta X'$ малого отрезка к длине изображаемого предмета δX , расположенного вдоль оптической оси:

$$\alpha = \delta X' / \delta X. \quad (2.8.4)$$

Рассмотрение малых отрезков (δX и $\delta X'$) связано с тем, что данное увеличение может сильно различаться в разных точках оптической оси.

Вследствие неизменности фокусных расстояний оптической системы из формулы Ньютона находим:

$$X \delta X + X' \delta X' = 0,$$

или

$$\alpha = \frac{\delta X'}{\delta X} = -\frac{X'}{X} = -\frac{f'}{f} \beta^2. \quad (2.8.5)$$

2.8.3. Угловое увеличение

По определению угловое увеличение равно

$$\gamma = \operatorname{tg} \varphi' / \operatorname{tg} \varphi. \quad (2.8.6)$$

Согласно рис. 2.8.2 находим

$$\gamma = -\frac{\operatorname{tg}(-\varphi')}{\operatorname{tg} \varphi} = -\frac{A'O'/(f' + X')}{AO/(-f - X)} = \frac{f + X}{f' + X'}. \quad (2.8.7)$$

Исключая в этой формуле величину X' с помощью формулы Ньютона: $X' = ff'/X$, получаем

$$\gamma = \frac{X}{f'} = \frac{f}{X'}. \quad (2.8.8)$$

Если в этой формуле исключить расстояния X и X' , воспользовавшись соотношениями (2.8.2) и (2.8.3), то мы установим связь углового и поперечного увеличений:

$$\gamma = -\frac{f}{f'} \frac{1}{\beta}. \quad (2.8.9)$$

Наконец, с учётом формулы (2.8.5) получаем связь трёх типов увеличений:

$$\beta = \alpha\gamma. \quad (2.8.10)$$

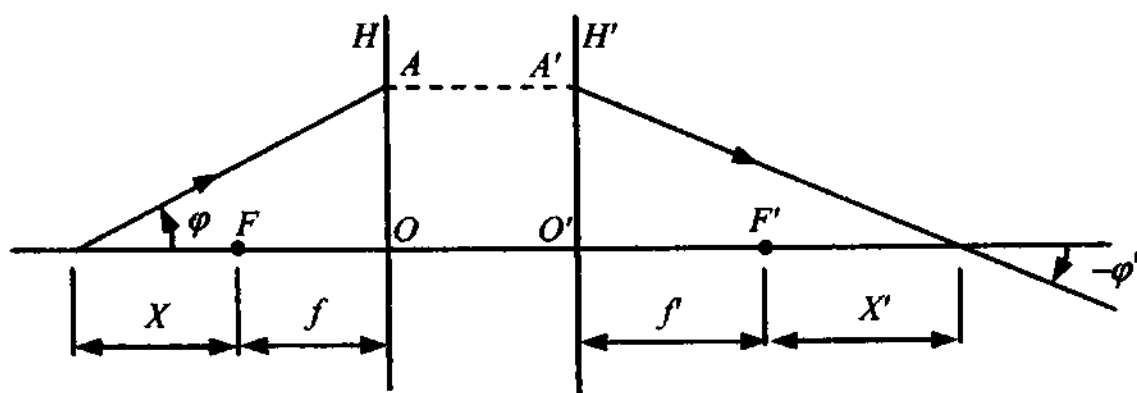


Рис. 2.8.2. К расчёту углового увеличения

Если среда с обеих сторон оптической системы одинакова (совпадают показатели преломления), то фокусные расстояния также совпадают: $f = -f'$. При этом из (2.8.5) находим

$$\alpha = \beta^2. \quad (2.8.11)$$

Соответственно из (2.8.10) или (2.8.9) вытекает связь углового и линейного увеличений:

$$\beta\gamma = 1. \quad (2.8.12)$$

Пусть предмет помещён в узловую точку. Поскольку в сопряжённых узловых точках $\gamma = +1$, то согласно (2.8.10) в этих точках продольное и поперечное увеличения совпадают: $\beta = \alpha$. Это значит, что изображение не деформируется.

2.9. Сферическая поверхность

2.9.1. Преломление сферической поверхностью

Пусть две среды с разными показателями преломления (n и n') отделены друг от друга сферической границей S радиуса R . Эта граница может рассматриваться как оптическая система, преобразующая падающее на неё излучение предмета (рис. 2.9.1).

Правило знаков: радиус сферы считается положительным, если центр находится с той стороны, куда распространяются лучи. На рис. 2.9.1

$R > 0$. Если бы выпуклость поверхности была ориентирована в другую сторону, то следовало бы положить $R < 0$.

Возьмём какую-либо точку Q на расстоянии x от поверхности и найдём положение её изображения Q' (рис. 2.9.1). Оптической осью является прямая, проходящая через точку Q и центр сферы C .

Рассмотрим луч QA , выходящий из точки Q , испытывающий преломление в точке A поверхности и пересекающий ось в точке Q' (рис. 2.9.1). Для нахождения траектории преломлённого луча проведём из центра сферы радиус CA . Этот радиус ортогонален поверхности сферы S , так что углы φ и ψ , образуемые лучом с радиусом по разные стороны поверхности, связаны законом Снеллиуса: $n \sin \varphi = n' \sin \psi$.

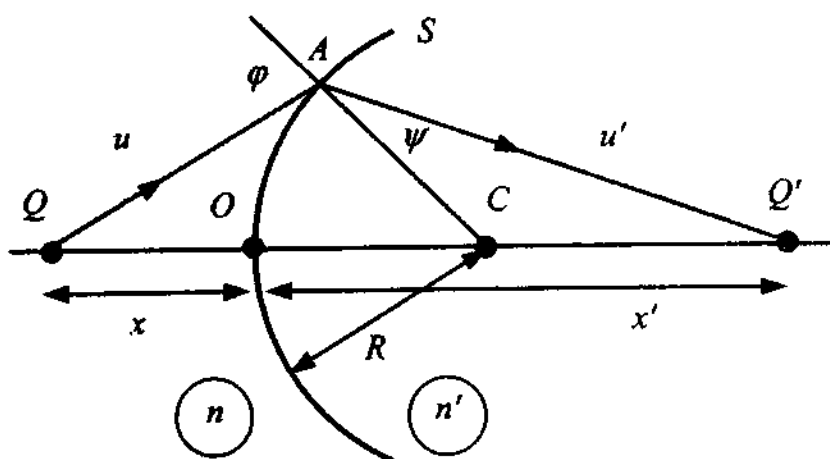


Рис. 2.9.1. К расчёту преломления на сферической поверхности S (нахождения положения изображения Q' точки Q). C — центр сферы радиуса R , n и n' — показатели преломления слева и справа от поверхности

Обозначим путь луча QA как u , а дальнейший путь AQ' как u' . Все пути отсчитываем от точки A по направлению распространения луча. Поэтому для случая, показанного на рис. 2.9.1,

$$u \equiv QA = -AQ < 0, \quad u' \equiv AQ' > 0.$$

Найдём связь путей u и u' . Из геометрических соображений имеем условия:

$$\text{площадь } QAC + \text{площадь } CAQ' = \text{площадь } QAQ',$$

$$\text{площадь } QAC = -\frac{1}{2}uR \sin \varphi \quad (u < 0),$$

$$\text{площадь } CAQ' = \frac{1}{2}u'R \sin \psi \quad (u' > 0),$$

$$\text{площадь } QAQ' = -\frac{1}{2}uu' \sin(\pi - \varphi + \psi).$$

Из этих равенств находим

$$-uR \sin \varphi + u'R \sin \psi = -uu' \sin(\varphi - \psi).$$

На основании закона преломления $\sin \varphi = \frac{n'}{n} \sin \psi$ находим

$$-u + u' \frac{n}{n'} = -\frac{uu'}{R} \left(\cos \psi - \cos \varphi \frac{n}{n'} \right),$$

или

$$\frac{n}{u} - \frac{n'}{u'} = \frac{n \cos \varphi - n' \cos \psi}{R}. \quad (2.9.1)$$

Это точное соотношение. Более употребительным является его частный случай, отвечающий *параксиальным* пучкам, для которых $|\varphi| \ll 1$, $|\psi| \ll 1$. В этом приближении

$u \approx x$ — расстояние от точки Q до поверхности ($x \equiv OQ$),

$u' \approx x'$ — расстояние от точки Q' до поверхности ($x' \equiv OQ'$).

Замечание. Выше и далее мы обозначаем заглавными буквами ($X, X', \dots$) расстояние от фокуса до предмета (изображения). Для обозначения же расстояния от преломляющих поверхностей до предмета (изображения) используем строчные буквы ($x, x', \dots$).

Правило знаков: знаки величин x и x' определяются, как и знаки величин u и u' , из того условия, что отсчёт расстояний ведётся от преломляющей поверхности *по направлению* лучей. Для случая, представленного на рис. 2.9.1, $x < 0$, $x' > 0$.

В указанных приближениях из (2.9.1) следует

$$\frac{n}{x} - \frac{n'}{x'} = \frac{n - n'}{R}. \quad (2.9.2)$$

Последнее означает, что в параксиальном приближении расстояние x' не зависит от угла, под которым распространяется луч QA . Иными словами, *приосевые* лучи, исходящие из точки Q , сходятся в одной и той же точке Q' , положение которой даётся формулой (2.9.2). Однако в случае широких пучков изображение не является точечным, оно размыто. Соответственно для увеличения резкости изображения следует сузить пучок. Это достигается применением *диафрагм*, выделяющих группу лучей, идущих от предмета с малым углом расхождения. Уменьшение же яркости изображения можно компенсировать увеличением яркости излучения предмета или применением усилителей сигнала.

Из (2.9.2) следует, что точка, в которой собирается параллельный пучок лучей ($x = -\infty$), то есть *задний фокус*, находится на расстоянии от поверхности, равном

$$f' = x' = \frac{R}{1 - n/n'}. \quad (2.9.3)$$

Положение переднего фокуса определится из условия, что исходящие из него лучи преобразуются в параллельный пучок ($x' = \infty$):

$$f = x = -\frac{n/n'}{1 - n/n'} R. \quad (2.9.4)$$

Видно, что в случае $n' > n$ фокусы расположены по разные стороны от поверхности: $f < 0$, $f' > 0$, причём

$$\frac{f}{f'} = -\frac{n}{n'}. \quad (2.9.5)$$

С учётом двух последних соотношений равенство (2.9.2) можно представить в традиционной форме:

$$\frac{f}{x} + \frac{f'}{x'} = 1. \quad (2.9.6)$$

Разрешим полученное соотношение относительно величины x' :

$$x' = \frac{f'x}{x - f}. \quad (2.9.7)$$

Имея в виду, что $x < 0$, $f < 0$, $f' > 0$, заключаем, что если предмет расположен к поверхности ближе, чем передний фокус: $x > f$, то $x' < 0$. Это значит, что изображение точки оказывается мнимым, расположенным в пространстве предметов. Если же предмет находится дальше от поверхности, чем передний фокус: $x < f$, то изображение оказывается действительным ($x' > 0$), находящимся в пространстве изображений. Рис. 2.9.1 отражает случай $x < f$. Другой случай ($x > f$) приведён в разделе 2.10 (рис. 2.10.2, слева).

2.9.2. Изображение предмета

Рассмотрим формирование изображения сферической поверхностью. Ограничимся случаем параксиальных пучков.

Построение изображения $P'Q'$ предмета PQ показано на рис. 2.9.2. Положение изображения Q' точки Q , лежащей на оптической оси, можно найти с помощью формулы (2.9.2). Можно также воспользоваться техникой, рассмотренной в разделе 2.6, где используются фокусы оптической системы. Траектория одного из лучей, соединяющих эти точки, показана штриховой линией.

Положение изображения P' крайней точки предмета P можно получить, проведя луч PP' , проходящий через центр C сферической поверхности. Этот луч не меняет направления при пересечении поверхности, падая на неё по нормали. Точка пересечения луча с нормалью, проведённой из точки Q' , даёт положение изображения P' . На рис. 2.9.2 изображение оказывается перевернутым: $y > 0$, $y' < 0$.

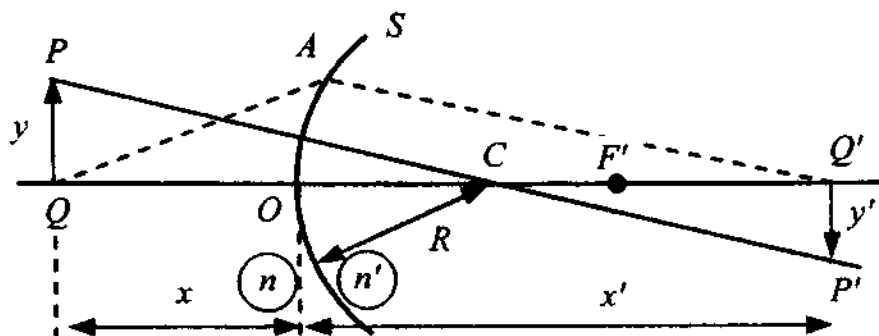


Рис. 2.9.2. Построение изображения $P'Q'$ предмета PQ , формируемого сферической преломляющей поверхностью S . C — центр сферы, $OC = R$ — радиус сферы, QQ' — главная оптическая ось, PP' — оптическая ось, F' — задний фокус ($n' > n$)

Найдём размер изображения.

В параксиальном приближении расстояние x' не зависит от угла, образованного лучом QA с оптической осью QQ' (то есть пучок лучей, испускаемых точкой Q , является гомоцентрическим).

Предварительно заметим, что для представленной на рис. 2.9.2 ситуации знаки существенных величин следующие:

$$x < 0, \quad x' > 0, \quad y > 0, \quad y' < 0.$$

Поэтому имеют место равенства

$$QC = |x| + R = -x + R, \quad Q'C = x' - R.$$

Имея это в виду, из подобия треугольников PQC и $P'Q'C$ находим соотношение размеров предмета (y) и изображения (y'):

$$\frac{y}{QC} = -\frac{y'}{Q'C} \Rightarrow \frac{y'}{y} = \frac{x' - R}{x - R}. \quad (2.9.8)$$

Исключая отсюда радиус сферы R с помощью соотношения (2.9.2):

$$\frac{n}{x} - \frac{n'}{x'} = \frac{n - n'}{R}, \quad \text{или} \quad R = \frac{(n - n')xx'}{nx' - n'x},$$

получаем искомое равенство:

$$\frac{n'y'}{ny} = \frac{x'}{x}. \quad (2.9.9)$$

2.9.3. Теорема Лагранжа–Гельмгольца

Пусть преломляющая поверхность осуществляет отображение предмета (рис. 2.9.3), причём луч QAQ' по разные стороны от поверхности образует малые углы φ и φ' с главной осью.

Поскольку $AA' = x\varphi = x'\varphi'$, или

$$x\varphi = x'\varphi', \quad (2.9.10)$$

то согласно формуле (2.9.9)

$$\frac{n'y'}{ny} = \frac{x'}{x} = \frac{\varphi}{\varphi'} \Rightarrow ny\varphi = n'y'\varphi'. \quad (2.9.11)$$

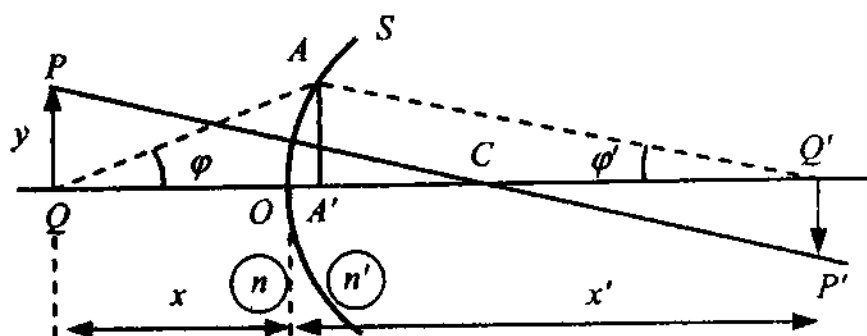


Рис. 2.9.3. К выводу теоремы Лагранжа–Гельмгольца. PQ — предмет, $P'Q'$ — изображение, S — сферическая преломляющая поверхность, C — центр сферы. Рисунок отражает случай $n' > n$

Полученное соотношение называется *теоремой Лагранжа–Гельмгольца* для преломления параксиальных пучков. Величина $ny\varphi$ не меняется при переходе через преломляющую поверхность и называется *инвариантом Лагранжа–Гельмгольца*.

2.9.4. Условие синусов Аббе

Формула (2.9.11) получена для параксиальных световых пучков. В ряде случаев приходится иметь дело с широкими пучками. Такая ситуация имеет место, например, в микроскопе, где рассматриваемый предмет помещается на относительно небольшом расстоянии от объектива.

Будем считать, что оптическая система создаёт точечное изображение предмета не только параксиальными, но и широкими пучками лучей. Мы видели (раздел 2.9.1), что в случае непараксиальных пучков изображение, создаваемое сферической поверхностью (и, вообще говоря, системой поверхностей), не является точечным. Можно, однако, в той или иной степени компенсировать связанные с этим искажения, используя, например, более сложные комбинации линз, а также линзы с несферическими поверхностями.

Имея в виду сказанное, рассмотрим точечное изображение малого предмета, создаваемое идеальной оптической системой. В качестве пред-

мета достаточно взять *малый* отрезок PQ , перпендикулярный главной оптической оси QQ' . Будем теперь считать, что данный отрезок изображается широким пучком лучей в виде отрезка $P'Q'$, перпендикулярного той же оптической оси (рис. 2.9.4). Точки отрезка и их изображения являются попарно сопряжёнными. В частности, сопряженными являются точки Q и Q' . Поэтому оптическая длина *любого* пути, соединяющего эти точки, одинакова. Аналогично, одинакова оптическая длина *любого* пути, соединяющего сопряженные точки P и P' .

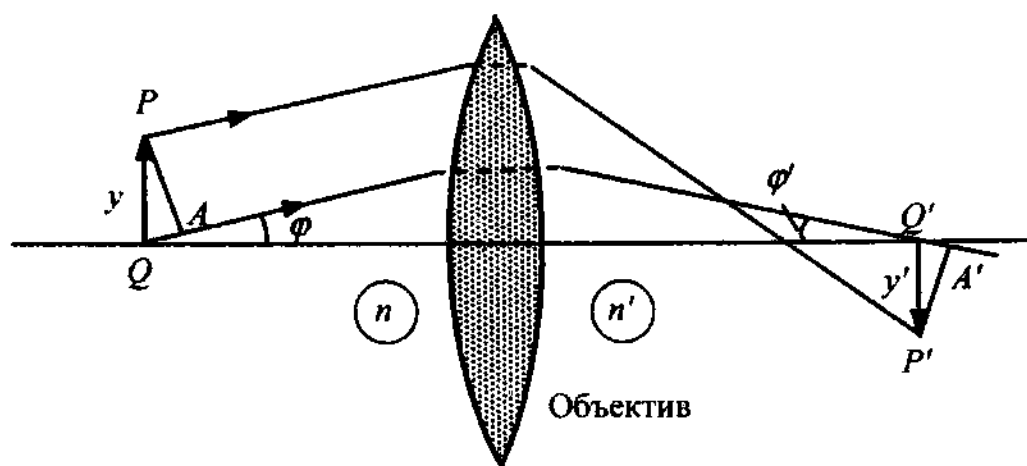


Рис. 2.9.4. К выводу условия синусов Аббе

Если бы лучи исходили под малым углом φ (то есть образовывали бы параксиальный пучок), то малый отрезок PQ практически принадлежал бы волновому фронту. При этом изображение $P'Q'$ также практически лежало бы на волновом фронте, прошедшем через объектив. Поскольку оптические длины *всех* лучей между двумя последовательными положениями волнового фронта одинаковы, то оптические длины лучей PP' и QQ' равны.

Поскольку мы считаем оптическую систему идеальной, то оптические длины лучей PP' и QQ' будут одинаковыми и для непараксиальных пучков, то есть пучков, исходящих под произвольным (немалым) углом φ от предмета.

Выберем параллельный пучок лучей, исходящих от малого предмета PQ под некоторым углом φ к оптической оси. Участок волнового фронта в этом случае изобразится отрезком PA , перпендикулярным лучам. После прохождения объектива параллельность лучей, вообще говоря, нарушится. Однако ввиду малости предмета это нарушение невелико, и лучи будут образовывать почти одинаковый угол φ' с осью. В результате волновой фронт займёт положение, в котором данный отрезок представится отрезком $P'A'$. При этом оптические длины лучей PP' и AA' будут одинаковыми (поскольку они соединяют соответственные точки двух положений волнового фронта). Таким образом,

$$L(PP') = L(QAQ'),$$

$$L(PP') = L(AQ'A').$$

Вычитая эти равенства почленно, находим

$$L(QA) = L(Q'A').$$

В соответствии с рис. 2.9.4 находим отсюда

$$ny \sin \varphi = n'y' \sin \varphi'. \quad (2.9.12)$$

Последнее соотношение называется *условием синусов Аббе* и является обобщением теоремы Лагранжа–Гельмгольца (2.9.11) на случай непараксиальных (широких) пучков.

2.9.5. Зеркало

Рассмотрим действие сферического зеркала (рис. 2.9.5). Пусть пучок лучей испускается точкой Q . Найдём положение её изображения Q' . В случае зеркала действует закон отражения «угол падения равен углу отражения». С учётом этого по аналогии с тем, как было сделано в разделе 2.9.1, записываем:

$$\text{площадь } QAQ' = \text{площадь } QAC - \text{площадь } Q'AC.$$

$$\text{площадь } QAQ' = \frac{1}{2}uu' \sin(\pi - 2\varphi).$$

$$\text{площадь } QAC = -\frac{1}{2}uR \sin \varphi.$$

$$\text{площадь } Q'AC = -\frac{1}{2}u'R \sin \varphi.$$

В этих формулах в левой части стоят положительные величины. Знаки правых частей строятся с учётом указанных выше правил: на рис. 2.9.5:

$$u < 0, u' < 0, R > 0.$$

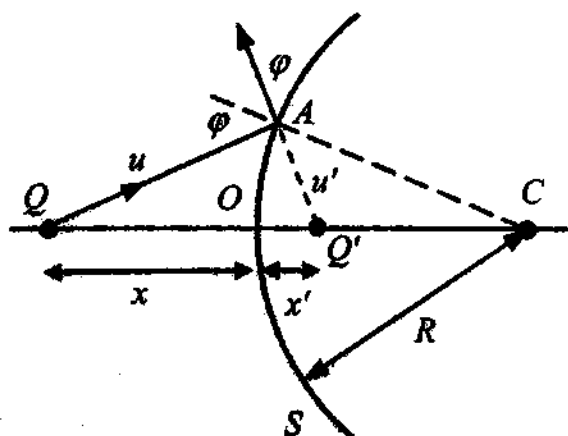


Рис. 2.9.5. К расчёту отражения сферической поверхностью S (нахождения положения изображения Q' точки Q). C — центр сферы радиуса R (на рисунке $R > 0$)

Из приведённых соотношений следует

$$\frac{1}{2}uu' \sin(\pi - 2\varphi) = -\frac{1}{2}uR \sin \varphi - \frac{1}{2}u'R \sin \varphi,$$

или

$$\frac{1}{u} + \frac{1}{u'} = -\frac{2 \cos \varphi}{R}. \quad (2.9.13)$$

Это точное соотношение. В случае параксиальных пучков ($|\varphi| \ll 1$) можно положить $u \approx x$, $u' \approx x'$, где x и x' — расстояния от точки Q и её изображения до поверхности зеркала (точки O), определяемые с учётом правила знаков. В результате приходим к формуле зеркала:

$$\frac{1}{x} + \frac{1}{x'} = -\frac{2}{R}. \quad (2.9.14)$$

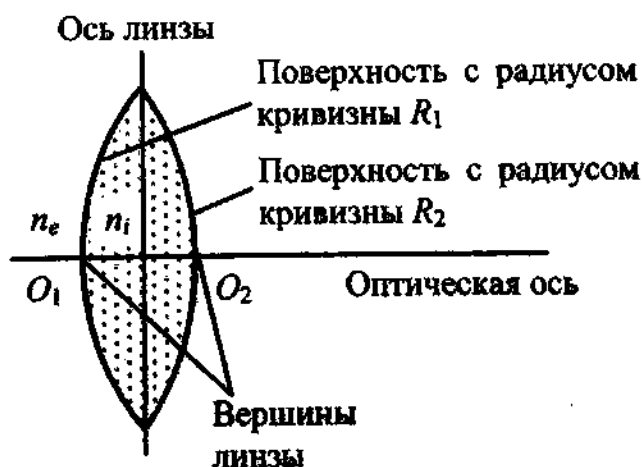
Заметим, что эта формула следует из (2.9.2), если формально положить $n=1$, $n'=-1$ и заменой $R \rightarrow -R$ (с целью сохранения прежнего определения знака R), поскольку луч не попадает по другую сторону поверхности, и центр сферы лежит против направления движения луча. Действительно, при этих условиях законы преломления и отражения приводят к одному и тому же результату для траектории луча.

2.10. Тонкая линза

2.10.1. Линза

Линза (от лат. *lens* — чечевица) — это прозрачное тело, изготовленное из прозрачного оптически однородного материала, ограниченное двумя полированными выпуклыми или вогнутыми поверхностями (одна из поверхностей может быть плоской). Обычно поверхности имеют сферическую форму, хотя применяют и *асферические* линзы, поверхности которых имеют иную, например, цилиндрическую, форму. Далее мы будем рассматривать линзы только со сферическими поверхностями (рис. 2.10.1).

Рис. 2.10.1. Характеристики линзы



Точки пересечения O и O' поверхностей линзы с оптической осью называются *вершинами линзы*. Расстояние d между вершинами называется *толщиной линзы*. Линза считается *тонкой*, если её толщина d мала по сравнению с радиусами кривизны поверхностей: $d \ll R_1$, $d \ll R_2$.

Главные плоскости тонкой линзы совпадают, поскольку вследствие малости толщины линзы луч внутри линзы практически не смещается в вертикальном направлении. Это означает, что на оси линзы (рис. 2.10.1) реализуется увеличение $\beta = +1$, определяющее главные плоскости.

2.10.2. Формула тонкой линзы

Представим линзу как совокупность двух последовательных преломляющих поверхностей. Первая поверхность (рис. 2.10.2, слева) формирует промежуточное изображение Q_1 точки Q , которое преобразуется второй поверхностью (рис. 2.10.3, справа) в окончательное изображение Q' . Будем считать, что показатель преломления вещества линзы n_e , а показатель преломления внешней среды n_i .

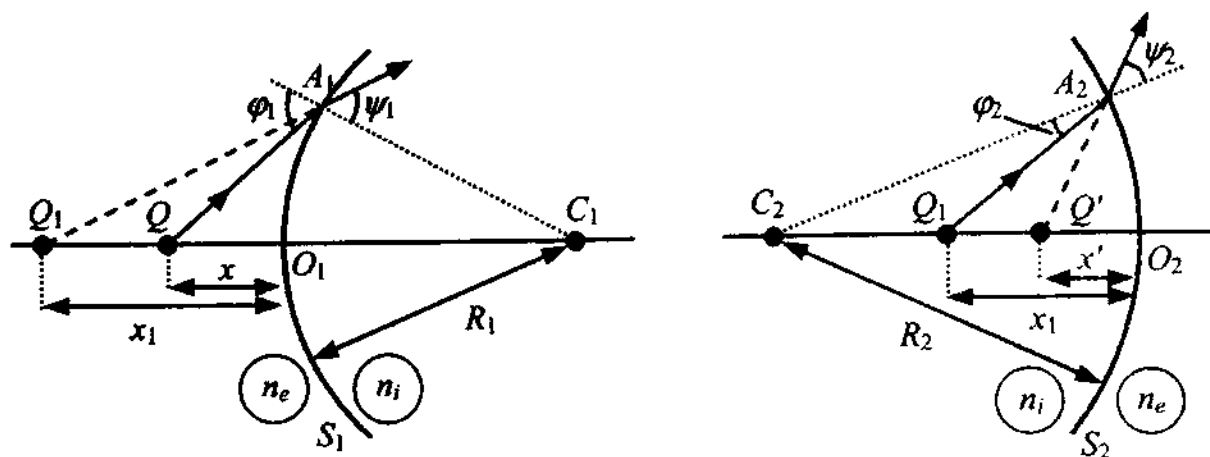


Рис. 2.10.2. К выводу формулы тонкой линзы (нахождения положения изображения Q' точки Q). C_1 и C_2 — центры сфер радиусами R_1 и R_2 , n_e и n_i — показатели преломления вне и внутри линзы соответственно, ϕ_i и ψ_i ($i = 1, 2$) — углы падения и преломления луча на поверхностях, x — положение предмета, x_1 — положение промежуточного изображения, x' — положение окончательного изображения. Штриховыми линиями показаны продолжения луча от преломляющей поверхности до точки пересечения с оптической осью. Вследствие тонкости линзы можно считать, что вершины поверхностей O_1 и O_2 совпадают

Согласно (2.9.2) для первой поверхности имеем

$$\frac{n_e}{x} - \frac{n_i}{x_1} = \frac{n_e - n_i}{R_1}. \quad (2.10.1)$$

Поскольку линза предполагается тонкой, то координата x_1 промежуточного изображения одинакова относительно обеих поверхностей. Поэтому применяя формулу (2.9.2) для луча, пересекающего вторую поверхность, можем записать

$$\frac{n_i}{x_1} - \frac{n_e}{x'} = \frac{n_i - n_e}{R_2}. \quad (2.10.2)$$

Складывая эти равенства, получаем

$$n_e \left(\frac{1}{x} - \frac{1}{x'} \right) = -(n_i - n_e) \left(\frac{1}{R_1} - \frac{1}{R_2} \right). \quad (2.10.3)$$

Заменяя в этой формуле $n = n_i/n_e$, то есть переходя к относительному показателю преломления n , перепишем последнее равенство в виде

$$\frac{1}{x} - \frac{1}{x'} = -(n-1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right). \quad (2.10.4)$$

2.10.3. Фокусное расстояние тонкой линзы

Если в (2.10.4) положить $x = -\infty$, то получим положение $x' = f'$ заднего фокуса линзы — *фокусное расстояние* линзы:

$$\frac{1}{f'} = (n-1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right). \quad (2.10.5)$$

Если положить $x' = \infty$, то из (2.10.4) найдём положение переднего фокуса:

$$x = f = -f'. \quad (2.10.6)$$

Таким образом, переднее и заднее фокусные расстояния линзы одинаковы и отличаются только знаками (вследствие принятого правила отсчёта расстояний).

Формула (2.10.5) выражает фокусное расстояние линзы через радиусы кривизны ее поверхностей. В соответствии с правилом знаков отсюда следует, что для двояковыпуклой линзы ($R_1 > 0$, $R_2 < 0$) фокусное расстояние $f' > 0$, $f < 0$, а для двояковогнутой линзы ($R_1 < 0$, $R_2 > 0$) $f' < 0$, $f > 0$.

Наконец, имея в виду выражение для фокусного расстояния (2.10.5), перепишем формулу (2.10.4) в следующем виде:

$$\frac{1}{x} - \frac{1}{x'} = \frac{1}{f}. \quad (2.10.7)$$

Это равенство называют *формулой тонкой линзы*. Подчеркнём, что в этой формуле в правой части стоит *переднее* фокусное расстояние f .

2.10.4. Оптический центр тонкой линзы

В случае тонкой линзы её главные плоскости практически совпадают. Точка пересечения главной оптической оси с этой «слившейся» главной плоскостью называется *оптическим центром* тонкой линзы (рис. 2.10.3).

Покажем, что любой луч POP' , проходящий через оптический центр O тонкой линзы, не меняет своего направления. Нужно доказать, что углы φ и φ' , образуемые лучом с главной оптической осью слева и справа от главной плоскости, равны (рис. 2.10.3).

Как видно из рис. 2.10.3, для треугольников POQ и $P'OQ'$ имеем

$$\operatorname{tg} \varphi = \frac{y}{X+f}, \quad \operatorname{tg} \varphi' = \frac{y'}{X'+f'}. \quad (2.10.8)$$

В этих формулах для ситуации, представленной на рис. 2.10.3, согласно правилам знаков

$$y > 0, y' < 0, X < 0, X' > 0, f < 0, f' > 0, \varphi < 0, \varphi' < 0.$$

Согласно методу построения изображения, проиллюстрированному на рис. 2.6.1 и 2.8.1, из точки P предмета посылают два луча: один через передний фокус, а второй параллельно главной оптической оси. Точка пересечения продолжений этих лучей даёт изображение P' точки P . Эти лучи показаны на рис. 2.10.3.

Рассматривая треугольник FOA , записываем равенство

$$y' = -yf/X.$$

Учтём также формулу Ньютона:

$$X' = ff'/X.$$

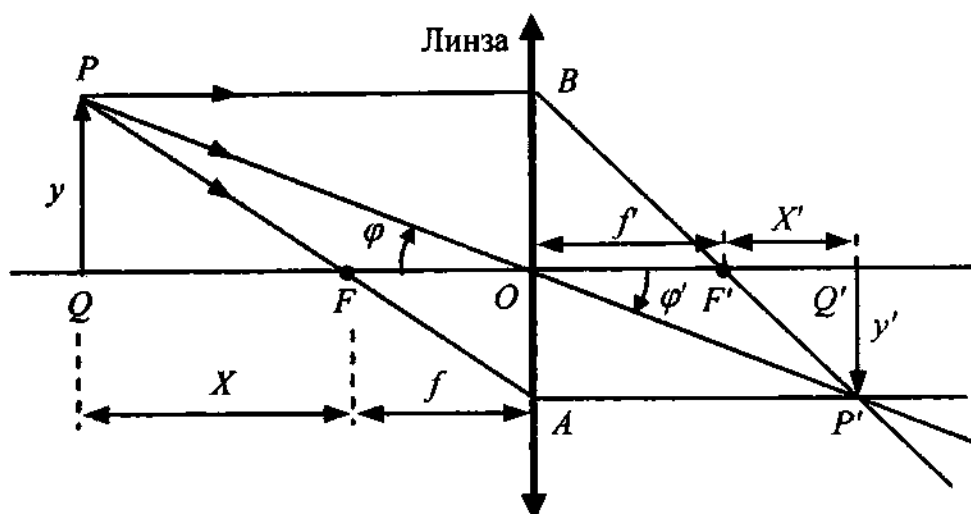


Рис. 2.10.3. Оптический центр O тонкой линзы. Построение изображения с использованием оптического центра

Подстановка этих равенств во вторую формулу (2.10.8) с учётом равенства $f = -f'$ даёт

$$\operatorname{tg} \varphi' = \frac{y'}{X'+f'} = \left(-y \frac{f}{X} \right) \frac{1}{\frac{ff'}{X} + f'} = \frac{y}{X+f} = \operatorname{tg} \varphi. \quad (2.10.9)$$

Доказанное означает, что независимо от выбора точки P луч POP' при пересечении линзы не меняет направления. Это означает также, что узловые точки тонкой линзы совпадают с оптическим центром.

Луч, проходящий через оптический центр линзы O , называют *побочной оптической осью*.

Отметим также, что из подобия треугольников POQ и $P'OQ'$ на рис. 2.10.3 следует дополнительное выражение для поперечного увеличения, даваемого тонкой линзой:

$$\beta = \frac{y'}{y} = \frac{f' + X'}{f + X}. \quad (2.10.10)$$

Эквивалентность данного представления выражению (2.8.2) $\beta = -f/X$ легко доказывается применением формулы Ньютона $XX' = ff'$ и равенства $f' = -f$.

В формуле (2.10.10) мы сохранили отдельные символы f и f' для переднего и заднего фокусных расстояний. Дело в том, что хотя они и равны по абсолютной величине, но в зависимости от кривизны поверхностей линзы (положительной или отрицательной) могут менять знак. В представлении же (2.10.10) сохраняется единообразие записи.

Наличие оптического центра позволяет находить изображение предмета, используя пару предметных лучей, из которых один проходит через центр. Второй луч может отправляться через передний фокус или параллельной оптической оси. По доказанному, все эти лучи пересекутся в точке P' изображения (рис. 2.10.3).

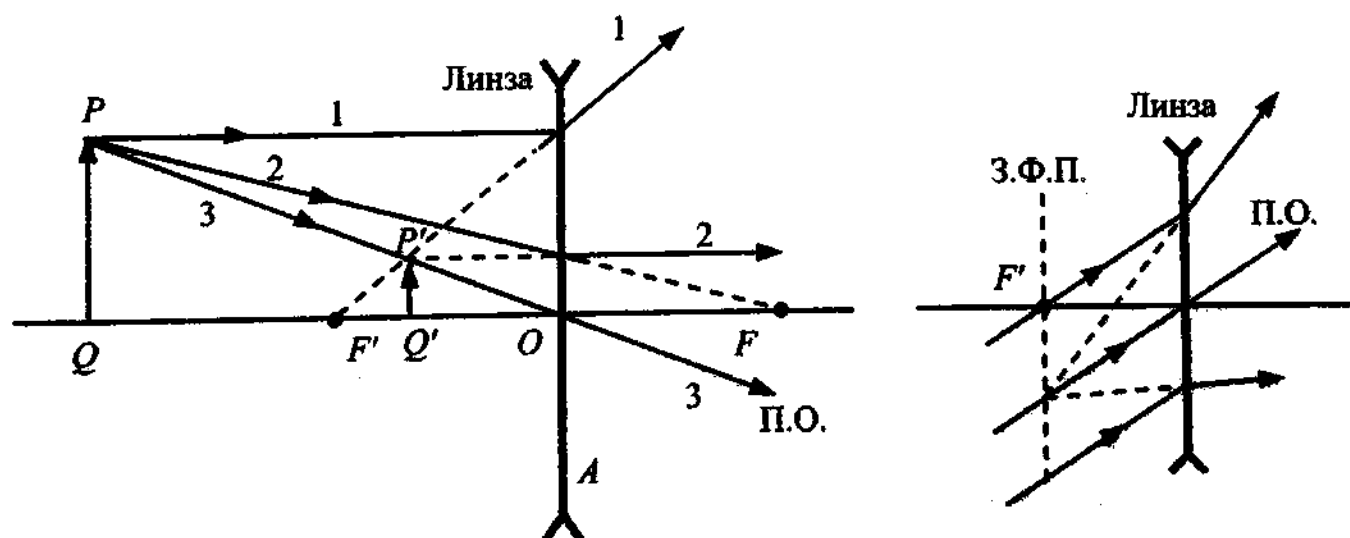


Рис. 2.10.4. Слева — построение изображения, создаваемого рассеивающей линзой; справа — преобразование параллельного пучка лучей в расходящийся. З.Ф.П. — задняя фокальная плоскость, П.О. — побочная оптическая ось

На рис. 2.10.4 (слева) показан пример построения изображения, создаваемого рассеивающей линзой. Для построения можно использовать пару фокусов. При этом передний фокус (F) оказывается справа от линзы,

а задний (F') — слева. На рис. 2.10.4 мнимое изображение $P'Q'$ оказывается на пересечении продолжений лучей 1 и 2, показанных штриховыми линиями. Возможно построение, в котором используются один из фокусов и оптический центр линзы (на рис. 2.10.4 оно реализуется парами лучей 1, 3 или 2, 3). Справа на рис. 2.10.4 показано преобразование параллельного пучка лучей, падающего на рассеивающую линзу, в расходящийся пучок. Для построения использованы побочная оптическая ось (П.О.) и задний фокус (F').

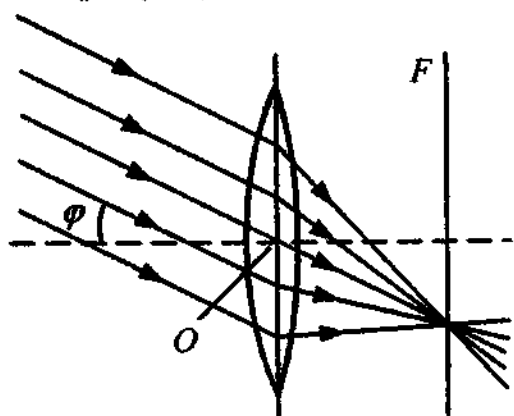


Рис. 2.10.5. Фокусировка параллельного пучка излучения в точку в фокальной плоскости тонкой линзы. O — оптический центр, F — фокальная плоскость

Любой параллельный пучок лучей, образующий произвольный, но небольшой угол φ с главной оптической осью тонкой линзы ($|\varphi| \ll 1$) собирается в точку в фокальной плоскости (рис. 2.10.5). Это осуществляется в две стадии: первая поверхность собирает лучи в точку в своей фокальной плоскости, а вторая поверхность создаёт изображение этой точки (подобно изображению предмета конечных размеров) уже в фокальной плоскости всей линзы.

Точка, в которую фокусируется параллельный пучок, находится на пересечении луча, проходящего через оптический центр линзы, с фокальной плоскостью (рис. 2.10.5).

2.11. Оптическая сила

Величина, обратно пропорциональная фокусному расстоянию пространства изображений:

$$\Phi = n'/f', \quad (2.11.1)$$

называется *оптической силой*.

Для собирающих систем оптическая сила положительна ($f' > 0$), а для рассеивающих отрицательна ($f' < 0$).

Единица измерения оптической силы — *диоптрия*. Диоптрия — это оптическая сила такой системы, фокусное расстояние которой $|f'|$ равно одному метру.

Если в формуле (2.9.2)

$$n/x - n'/x' = (n - n')/R$$

положить $x = -\infty$, то получим $x' = f'$, и мы приходим к следующему выражению для оптической силы преломляющей поверхности:

$$\Phi = \frac{n'}{f'} = \frac{n'}{x'} = -\frac{n-n'}{R}. \quad (2.11.2)$$

Знание оптической силы позволяет найти заднее фокусное расстояние:

$$f' = n'/\Phi. \quad (2.11.3)$$

Выражение для переднего фокусного расстояния следует из (2.9.2): полагая $x' = \infty$ и учитывая равенство (2.11.2) найдём

$$x = f = -n/\Phi. \quad (2.11.4)$$

Для зеркала оптическая сила может быть также определена формулой (2.11.2), если в ней положить $n = 1$, $n' = -1$:

$$\Phi = -2/R. \quad (2.11.5)$$

Для выпуклого зеркала (рис. 2.11.1, слева) $R > 0$, $\Phi < 0$. Это означает, что такое зеркало рассеивает падающее на него излучение. Если же $R < 0$ (излучение падает на вогнутую поверхность, рис. 2.11.1, справа), то зеркало фокусирует излучение ($\Phi > 0$).

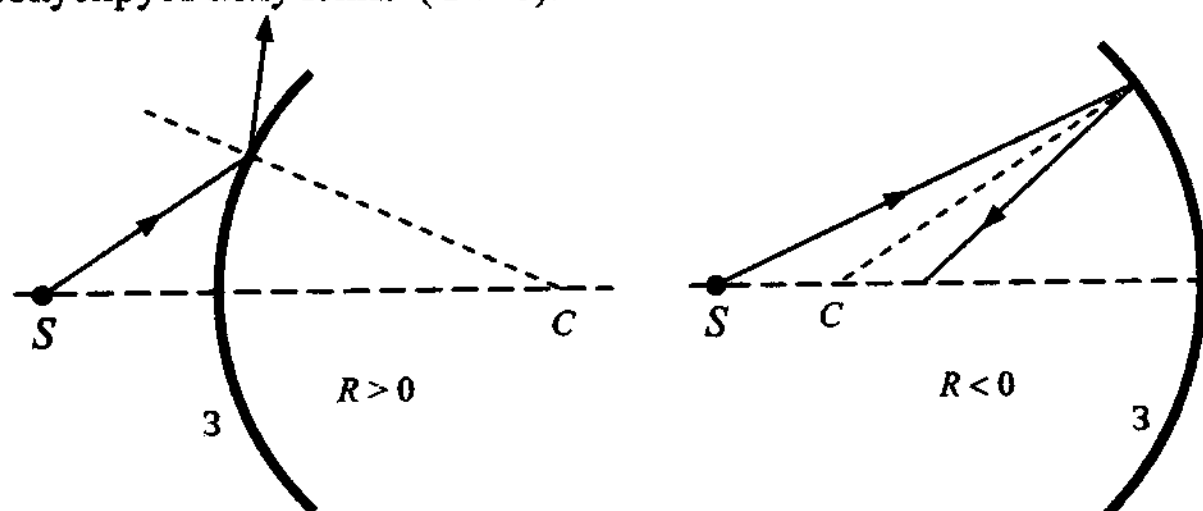


Рис. 2.11.1. Слева — отражение выпуклым зеркалом ($R > 0$); справа — отражение вогнутым зеркалом ($R < 0$). 3 — зеркало, C — центр сферической поверхности, S — источник

2.12. Сложение центрированных оптических систем

2.12.1. Фокусные расстояния составной системы

Рассмотрим две центрированные системы, главные оптические оси которых совпадают (рис. 2.12.1). Составная система при этом также оказывается центрированной.

Обозначим характеристики, относящиеся к складываемым системам, индексами 1 и 2, а характеристики, относящиеся к составной системе, теми же символами, но без индексов (рис. 2.12.1 и рис. 2.12.2).

Расстояние Δ между задним фокусом F_1' первой системы и передним фокусом F_2 второй называется *оптическим интервалом* двух складывае-

мых систем. Эта величина положительна: $\Delta > 0$, если свет идет от F'_1 к F_2 . Иначе $\Delta < 0$.

Введём расстояние d между аналогичными главными плоскостями систем. Это расстояние d также является алгебраической величиной, причём в соответствии с определением фокусных расстояний имеет место равенство

$$d = \Delta + f'_1 - f_2. \quad (2.12.1)$$

Требуется найти фокусные расстояния сложной системы, зная фокусные расстояния складываемых систем и оптический интервал между ними.

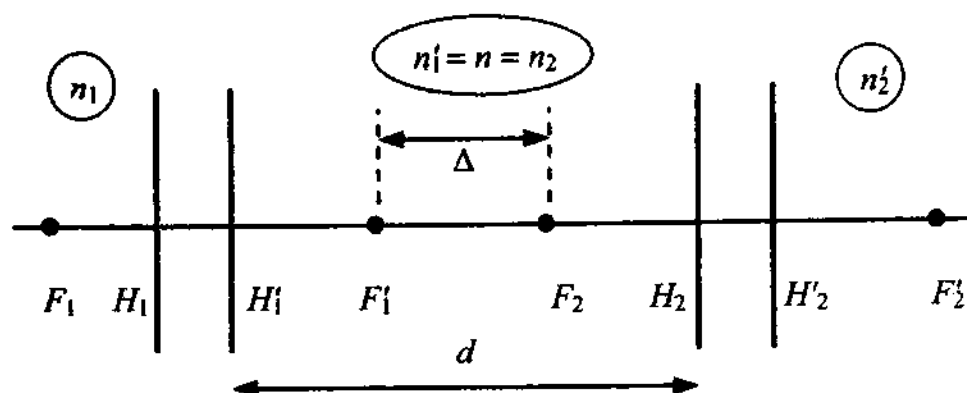


Рис. 2.12.1. Составная оптическая система, состоящая из двух систем с общей главной оптической осью. Показаны положения фокусов и главных плоскостей складываемых систем

Для расчёта воспользуемся рис. 2.12.2. Будем считать, что пространство между главными плоскостями H'_1 и H_2 складываемых оптических систем оптически однородно и имеет показатель преломления n .

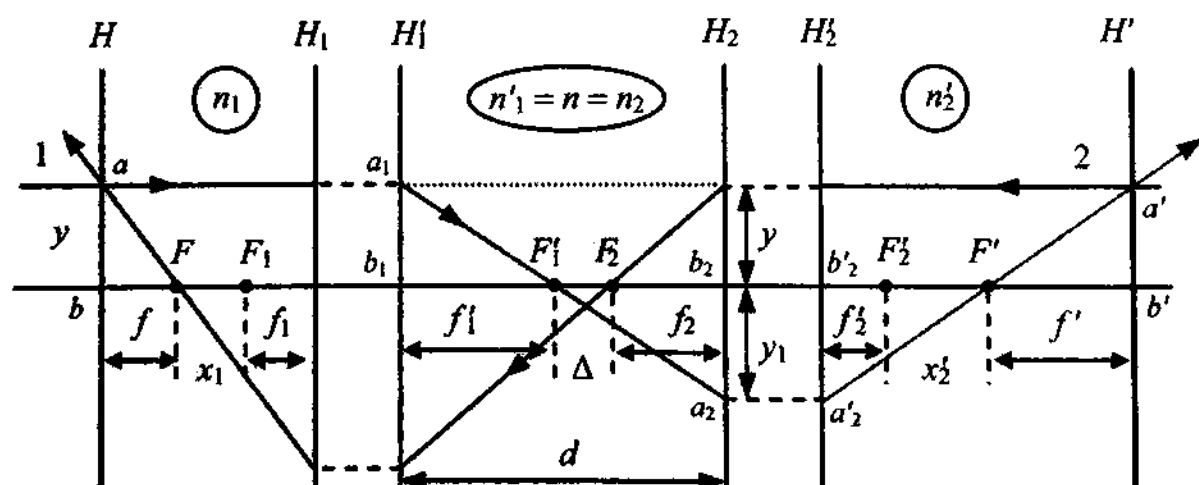


Рис. 2.12.2. К выводу закона сложения оптических систем

Пошлём луч 1 слева направо параллельно главной оптической оси. Данный луч пройдёт на расстоянии Δ от переднего фокуса F_2 системы 2 и пересечёт главную оптическую ось в точке F' , являющейся задним фоку-

сом составной системы. Это позволяет найти расстояние x'_2 от заднего фокуса F'_2 системы 2 до заднего фокуса F' составной системы. По формуле Ньютона (2.7.3) после прохождения второй системы имеем

$$(-\Delta) \cdot x'_2 = f_2 f'_2 \Rightarrow x'_2 = -f_2 f'_2 / \Delta. \quad (2.12.2)$$

Здесь и далее при выводе искомых соотношений мы предполагаем относительное расположение кардинальных точек, как показано на рис. 2.12.2. Другие случаи получаются при учёте правила знаков, сформулированного выше.

Для нахождения фокусного расстояния f' нужно найти положение главной плоскости H' составной системы. Это можно сделать следующим образом. Из подобия прямоугольных треугольников $a_1 b_1 F'_1$ и $a_2 b_2 F'_1$ (между плоскостями H'_1 и H_2 на рис. 2.12.2) следует равенство

$$\frac{y}{-y_1} = \frac{f'_1}{\Delta - f_2}. \quad (2.12.3)$$

Для прямоугольных треугольников $a' b' F'$ и $a'_2 b'_2 F'$ (между плоскостями H'_2 и H')

$$\frac{y}{-y_1} = \frac{-f'}{f'_2 + x'_2}. \quad (2.12.4)$$

Подстановка сюда выражения $x'_2 = -f_2 f'_2 / \Delta$ из (2.12.2) даёт

$$\frac{y}{-y_1} = \frac{f'}{f'_2 - \frac{f_2 f'_2}{\Delta}} = -\frac{f' \Delta}{f'_2 (\Delta - f_2)}. \quad (2.12.5)$$

Сравнение (2.12.3) и (2.12.5) приводит к формуле для заднего фокусного расстояния f' :

$$f' = -f'_1 f'_2 / \Delta. \quad (2.12.6)$$

Положение передней главной плоскости и переднее фокусное расстояние можно найти, если рассмотреть траекторию луча 2 (рис. 2.12.2), запущенного в обратном направлении. Повторяя рассуждения (с необходимыми изменениями знаков входящих в формулы величин), находим расстояние от переднего фокуса F_1 системы 1 до переднего фокуса F составной системы и суммарное переднее фокусное расстояние:

$$x_1 = f_1 f'_1 / \Delta, \quad f = f_1 f_2 / \Delta. \quad (2.12.7)$$

2.12.2. Оптическая сила составной системы

Согласно определению оптической силы (2.11.1) с учётом полученного выше соотношения (2.12.6) находим

$$\Phi = \frac{n'_2}{f'} = -\frac{n'_2 \Delta}{f'_1 f'_2}. \quad (2.12.8)$$

Преобразуем это выражение, вводя оптические силы систем 1 и 2:

$$\Phi = -\frac{\Delta}{n} \cdot \frac{n}{f'_1} \cdot \frac{n'_2}{f'_2} \equiv -\frac{\Delta}{n} \Phi_1 \Phi_2. \quad (2.12.9)$$

Здесь использованы обозначения рис. 2.12.2.

Поскольку согласно (2.12.1) $d = \Delta + f'_1 - f'_2$, то

$$\Phi = \left(-\frac{d}{n} + \frac{f'_1}{n} - \frac{f'_2}{n} \right) \Phi_1 \Phi_2.$$

Имея в виду, что

$$\frac{f'_1}{n} = \frac{1}{\Phi_1}, \quad \frac{f'_2}{n} = -\frac{1}{\Phi_2}, \quad (2.12.10)$$

получим формулу сложения оптических сил:

$$\Phi = \Phi_1 + \Phi_2 - \frac{d}{n} \Phi_1 \Phi_2. \quad (2.12.11)$$

В частном случае, когда расстояние d столь мало, что последним слагаемым можно пренебречь ($d \rightarrow 0$), получаем, что оптическая сила сложной системы равна сумме оптических сил составляющих систем:

$$\Phi = \Phi_1 + \Phi_2. \quad (2.12.12)$$

Последний результат распространяется на случай произвольного числа элементов составной оптической системы, если только эти элементы расположены достаточно близко:

$$\Phi = \sum_i \Phi_i. \quad (2.12.13)$$

Зная оптическую силу составной системы, можно найти её фокусные расстояния по формулам вида (2.11.3), (2.11.4):

$$f' = n'_2 / \Phi, \quad f = -n_1 / \Phi. \quad (2.12.14)$$

2.13. Оптические приборы

2.13.1. Лупа

Лупа — это оптическая система, состоящая из одной или нескольких линз и предназначенная для наблюдения мелких предметов, расположенных на конечном расстоянии.

Рассматривая предмет глазом непосредственно (рис. 2.13.1, слева), имеем $\text{tg } \varphi = y / l_{\text{н.з.}}$. Если же с расстояния $l_{\text{н.з.}}$ рассматривается изображение, даваемое лупой (рис. 2.13.1, справа), то $\text{tg } \varphi' = y' / l_{\text{н.з.}}$. Обычно глаз располагается вблизи заднего фокуса системы, так что $l_{\text{н.з.}} \approx f'$. Поэтому,

как видно из рис. 2.13.1 (справа), $\operatorname{tg} \varphi' \approx y/f'$. Соответственно находим видимое увеличение лупы

$$\Gamma = \frac{\operatorname{tg} \varphi'}{\operatorname{tg} \varphi} \approx \frac{y/f'}{y/l_{\text{н.з.}}} = \frac{l_{\text{н.з.}}}{f'}.$$

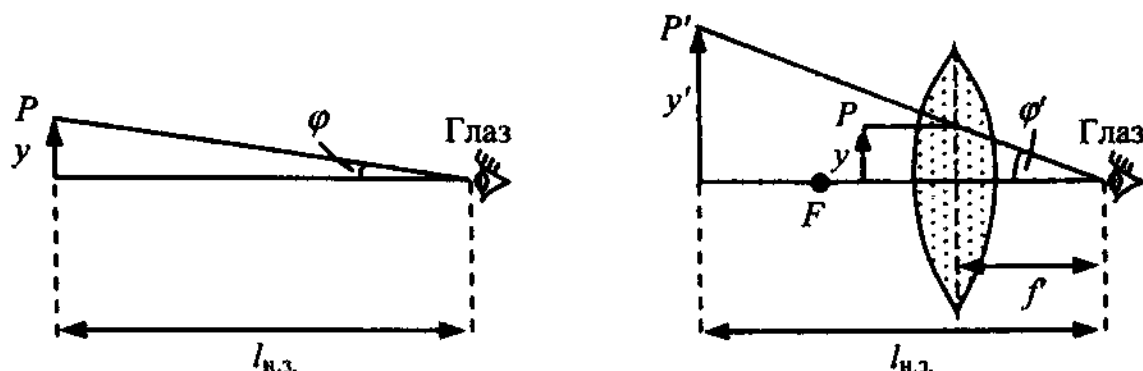


Рис. 2.13.1. Наблюдение предмета непосредственно глазом (слева) и через лупу (справа)

В частности, линза с фокусным расстоянием $f = f' = 5$ см даёт увеличение $\Gamma \approx 5\times$.

2.13.2. Микроскоп

Микроскоп предназначен для рассматривания малых предметов и включает объектив и окуляр (рис. 2.13.2).

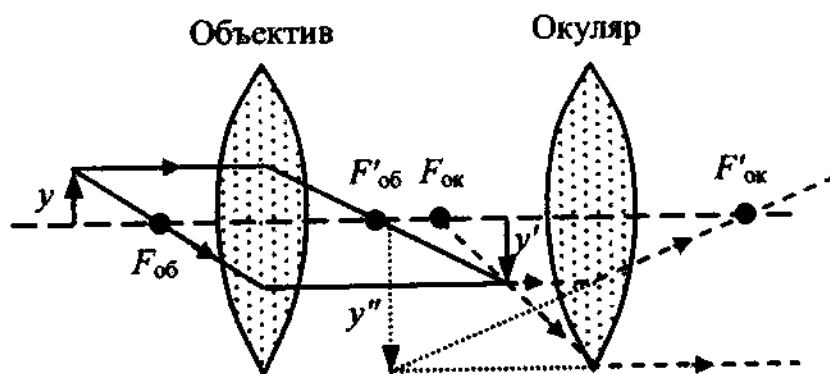


Рис. 2.13.2. Наблюдение предмета «у» через микроскоп

Предмет (y) помещается на малом расстоянии перед передним фокусом объектива, который даёт действительное перевёрнутое изображение (y'). Это изображение затем рассматривается с помощью окуляра, действующего так же, как лупа.

Увеличение микроскопа (как и ряда других оптических приборов) обусловлено тем, что конечное изображение (y'') видно в окуляр под бóльшим углом, чем при непосредственном рассматривании глазом.

2.13.3. Телескоп (труба Кеплера)

Телескопическая оптическая система — это система, включающая объектив и окуляр, причём эти части расположены так, что передний фокус окуляра совмещён с задним фокусом объектива (рис. 2.13.3). При таком расположении элементов параллельный пучок лучей, попадающий в объектив, преобразуется в параллельный пучок, выходящий из окуляра.

Объектив даёт в задней фокальной плоскости обратное уменьшенное изображение предмета. Это изображение рассматривается в окуляр, действующий как лупа.

Рис. 2.13.3 представляет ход пучка параллельных лучей в случае, когда они распространяются параллельно (рис. 2.13.3, слева) и не параллельно (рис. 2.13.3, справа) главной оптической оси.

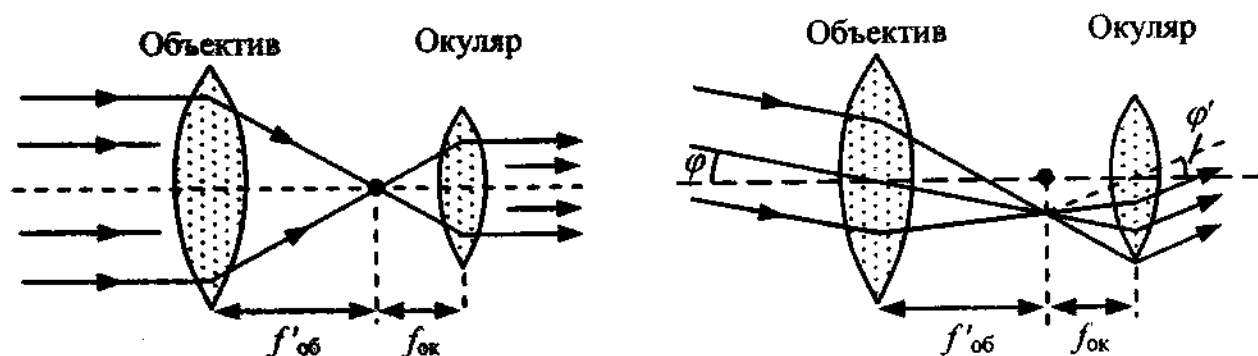


Рис. 2.13.3. Ход пучка параллельных лучей в телескопической системе. Слева — лучи параллельны главной оптической оси, справа — лучи образуют ненулевой угол φ с главной оптической осью

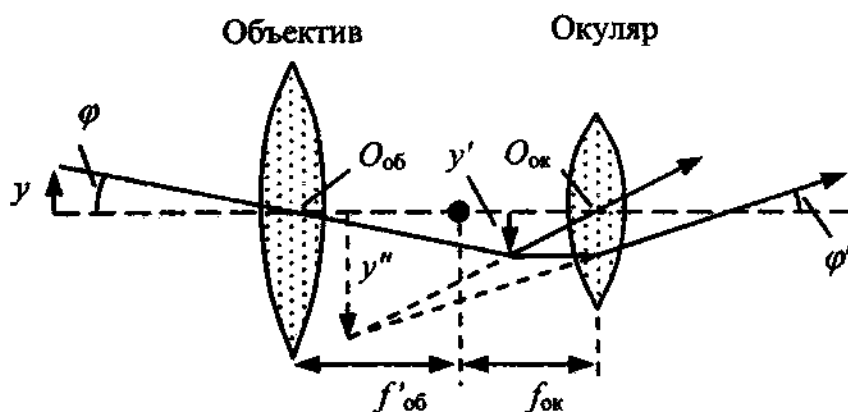


Рис. 2.13.4. К расчёту увеличения телескопа. $O_{об}$ и $O_{ок}$ — оптические центры объектива и окуляра

Для оценки увеличения, даваемого телескопом, воспользуемся рис. 2.13.4. Наблюдаемый объект y находится обычно на большом расстоянии от наблюдателя, и лучи от него распространяются под малым углом φ к оси телескопа. Поэтому промежуточное изображение y' , даваемое объективом, находится практически в общей фокальной плоскости объектива и окуляра. Далее это изображение рассматривается в окуляр и на-

блюдается как мнимое изображение y'' . Из построения, использующего лучи, проходящие через оптические центры объектива и окуляра (рис. 2.13.4), видно, что

$$\operatorname{tg} \varphi = y' / f'_{об}, \quad \operatorname{tg} \varphi' = y' / f_{ок}.$$

Отсюда находим видимое увеличение:

$$\Gamma = \frac{\operatorname{tg} \varphi'}{\operatorname{tg} \varphi} = \frac{f'_{об}}{f_{ок}}.$$

Фокусное расстояние окуляра, как правило, мало по сравнению с фокусным расстоянием объектива, что и обеспечивает значительное увеличение изображения рассматриваемого объекта.

2.13.4. Труба Галилея

Телескоп даёт увеличенное, но перевёрнутое изображение удалённого предмета. В зрительной трубе Галилея изображение получается прямым за счёт того, что в качестве окуляра используется двояковогнутая линза. При этом, как и в телескопе, задний фокус объектива совмещён с передним фокусом окуляра (рис. 2.13.5). Соответственно в окуляр (лупу) рассматривается прямое изображение.

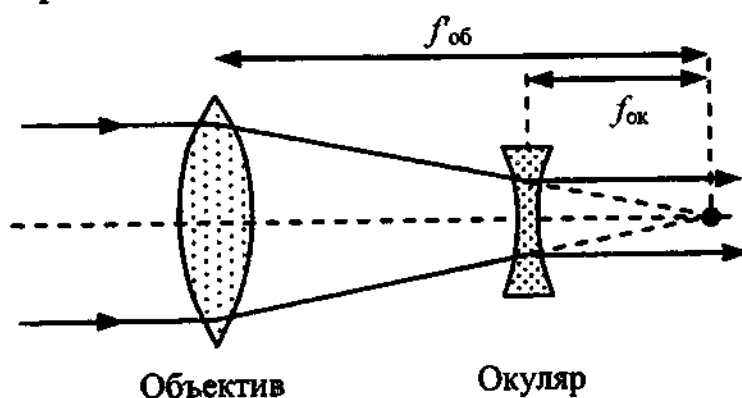


Рис. 2.13.5. Ход лучей в зрительной трубе (трубе Галилея)

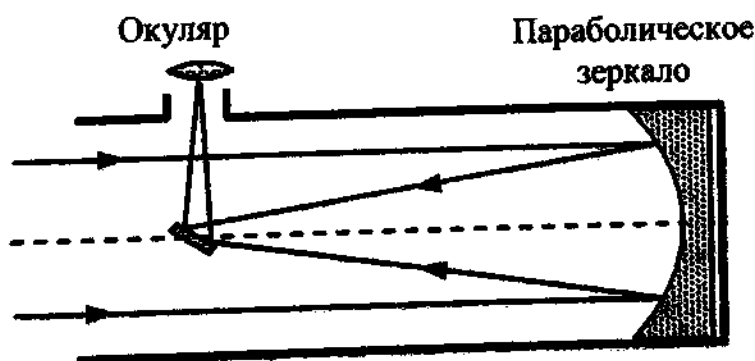


Рис. 2.13.6. Телескоп-рефлектор (схема Ньютона)

Наряду с телескопами-рефракторами, в которых объективом является линза, применяются телескопы-рефлекторы (или зеркальные телескопы) с объективом, представляющим собой вогнутое (фокусирующее) зер-

каю (рис. 2.13.6)¹. Зеркало формирует промежуточное изображение в своём фокусе. Это изображение с помощью небольшого плоского зеркала выводится из телескопа и затем рассматривается в окуляр. Пучок света может выводиться из телескопа в окуляр различными способами. На рис. 2.13.6 показана схема зеркального телескопа, предложенная Ньютоном.

Основное зеркало делают параболическим, имея в виду его свойство фокусировать широкие параллельные пучки света в фокусе. Однако если имеют дела с относительно узкими пучками света, зеркало может быть сферическим, которое проще в изготовлении.

2.14. Об аберрациях оптических систем

Аберрации оптической системы — это ошибки или дефекты изображения в реальной оптической системе, вызываемые отклонением лучей от того направления, по которому они должны были бы идти в идеальной оптической системе.

Оптические системы, как правило, не совершенны — даваемые ими изображения не абсолютно точно отображают изучаемые предметы. Например, сферическая поверхность даёт стигматическое изображение, только если оно формируется приосевыми пучками. При больших же углах, образуемых лучами с оптической осью, точки предмета перестают отображаться в сопряжённые точки изображения.

Выделяют несколько типов аберраций. Укажем только некоторые: сферическая аберрация, кома, хроматическая аберрация.

Сферическая аберрация (рис. 2.4.1, слева) возникает из-за несовпадения фокусов для лучей света, проходящих на разных расстояниях от оптической оси и приводит к нарушению гомоцентричности пучков лучей от точечного источника, без нарушения симметрии строения этих пучков.

Коматическая аберрация, или *кома* (от др.-греч. *κόμη* — волосы) — это одна из аберраций, приводящая к нарушению гомоцентричности широких световых пучков, входящих в систему под углом к оптической оси (рис. 2.4.1, справа).

Каждый участок оптической системы, удалённый от её оси на расстояние r (кольцевая зона), даёт изображение светящейся точки в виде кольца, радиус которого тем больше, чем больше r . Центры колец не совпадают, в результате чего их наложение, то есть изображение точки, даваемое системой в целом, принимает вид несимметричного пятна, похожего на комету. С этим сходством и связано название данной аберрации.

¹ «Рефрактор» (от лат. *refractio*, преломление) — оптический прибор, использующий преломляющие элементы (например, линзы). «Рефлектор» (от лат. *reflecto*, отражаю) — оптический прибор, использующий отражающие элементы (зеркала).

Хроматическая аберрация — это аберрации, возникающая вследствие зависимости показателя преломления среды от длины волны проходящего через неё излучения. Из-за этого, в частности, фокусные расстояния оптических элементов не совпадают для лучей света с разными длинами волн (лучей разных цветов).

Заметим, что хроматические аберрации минимизируются у зеркальных телескопов, поскольку отражение лучей не зависит от длины волны и определяется только углами падения на отражающую поверхность. В телескопах-рефракторах же преломляющие свойства линз и, следовательно, фокусные расстояния зависят от длины волны света.

Устранение большинства аберраций достигается соответствующим подбором комбинаций линз и зеркал в центрированной оптической системе.

Помимо аберраций, возникающих вследствие неидеальности оптических систем, существуют неустраняемые искажения изображений, связанные с дифракционными явлениями, обусловленными волновыми свойствами света.

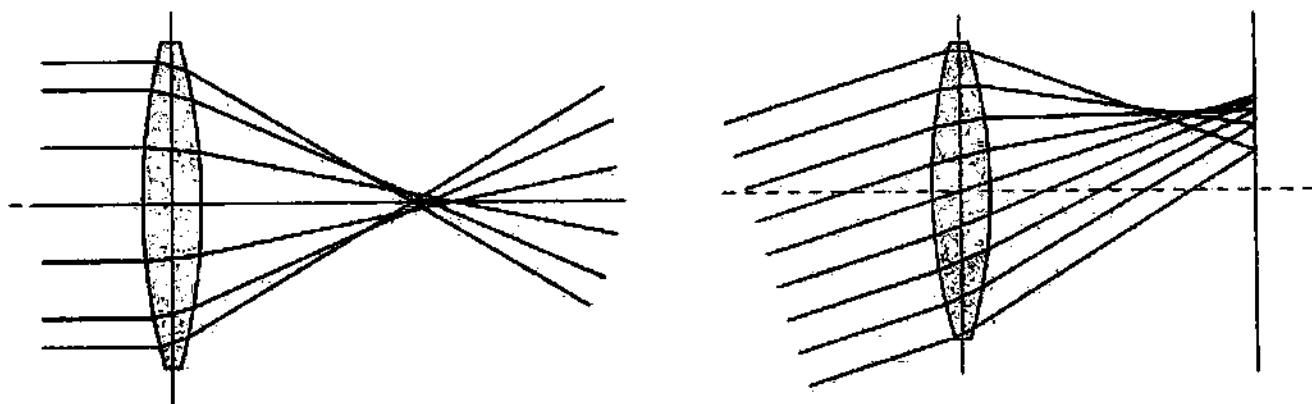


Рис. 2.14.1. Слева — сферическая аберрация, справа — аберрация «кома»

2.15. Основные понятия фотометрии

2.15.1. Поток энергии, световой поток, видность

Обозначим Φ_3 поток энергии, то есть энергию, переносимую излучением в единицу времени. Поскольку реальное излучение содержит компоненты с различной длиной волны, то полный поток может быть представлен в виде

$$\Phi_3 = \int_0^{\infty} \varphi(\lambda) d\lambda, \quad (2.15.1)$$

где $\varphi(\lambda)$ — спектральная плотность энергии (рассчитанная на единичный интервал длин волн).

Единица измерения потока энергии в системе СИ — ватт (Вт).

Вводится функция (кривая) видности $V(\lambda)$, учитывающая чувствительность глаза к излучению с различными длинами волн (рис. 2.15.1). Её называют также *относительной спектральной световой эффективностью*. Максимального значения она достигает для дневного зрения человека в зелёной части спектра при

$$\lambda = 555 \text{ нм}, \nu = 5,4 \cdot 10^{14} \text{ Гц}.$$

Для характеристики интенсивности света с учётом его способности вызывать зрительные ощущения вводится величина Φ — *световой поток*. Для интервала $d\lambda$ имеем

$$d\Phi = V(\lambda)d\Phi_{\text{э}}.$$

С учётом (2.15.1) полный световой поток может быть представлен в виде

$$\Phi = \int_0^{\infty} V(\lambda)\varphi(\lambda)d\lambda. \quad (2.15.2)$$

Единицей измерения светового потока является люмен (определение дано в следующем разделе).

2.15.2. Сила света. Фотометрические единицы измерения

Источник света, размерами которого можно пренебречь по сравнению с расстоянием от источника до места наблюдения, называется *точечным*.

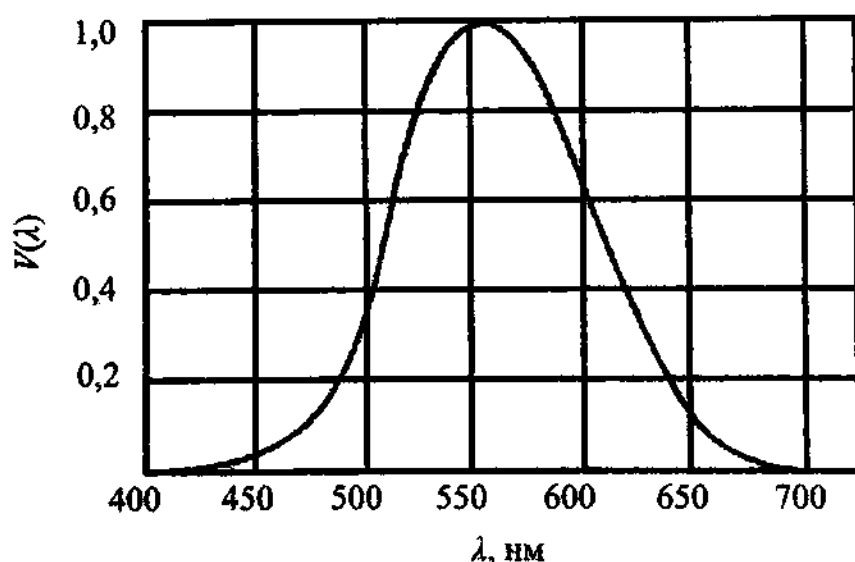


Рис. 2.15.1. Кривая (функция) видности оптического излучения

Световой поток, излучаемый в единичный телесный угол, называется *силой света*:

$$J = d\Phi/d\Omega. \quad (2.15.3)$$

Изотропный источник равномерно излучает в телесный угол $\Omega = 4\pi$ стерадиан (ср), так что полный световой поток $\Phi = 4\pi J$.

Кандела (кд) — это единица силы света J в системе СИ, основная фотометрическая единица. Она определялась так: яркость (см. ниже) полного излучателя при температуре затвердевания чистой платины ($T = 2046,6$ К) и давлении 101 325 Па равна 60 кд/см².

Эта единица реализована с помощью эталона как сила света, излучаемого с 1/60 см² эталона в направлении нормали. Данное определение действовало в период 1948 по 1979 годы. До 1948 года использовалась единица силы света свеча, равная 1,005 кд.

С 1979 года по решению Международного комитета по световым стандартам канделу определяют как силу света в заданном направлении от источника, испускающего монохроматическое излучение с частотой $540 \cdot 10^{12}$ Гц, энергетическая сила света которого $J_\Omega = d\Phi_\Omega/d\Omega$ в этом направлении равна 1/683 Вт·ср⁻¹.

Единица светового потока Φ — люмен. 1 лм — это световой поток, посылаемый источником в 1 кд в телесный угол 1 ср:

$$1 \text{ лм} = 1 \text{ кд} \cdot 1 \text{ ср}.$$

Спектральная световая эффективность излучения — это отношение светового потока на длине волны λ к соответствующему потоку излучения:

$$K(\lambda) = d\Phi/d\Phi_\Omega.$$

Размерность этой величины есть лм/Вт. Максимального значения $K_m = 683$ лм/Вт она достигает при $\lambda \approx 555$ нм. Нормированная функция $K(\lambda)$, то есть отношение

$$V(\lambda) = K(\lambda)/K_m, \quad (2.15.4)$$

есть *относительная спектральная световая эффективность*, или *видность* (рис. 2.15.1).

Величина $A = 1/K_m = 1.46 \cdot 10^{-3}$ Вт/лм называется *механическим эквивалентом света*.

2.15.3. Интенсивность света

Интенсивность света — это световой поток, пересекающий единичную площадку по нормали к ней:

$$I = d\Phi/dS_\perp. \quad (2.15.5)$$

Единица измерения интенсивности — люмен на квадратный метр, лм/м². Если точечный источник света находится в однородной прозрачной среде, то $dS_\perp = r^2 d\Omega$, так что $I = J/r^2$.

2.15.4. Освещённость

Освещённость — это световой поток, приходящийся на единицу площади освещаемой поверхности (рис. 2.15.2):

$$E = d\Phi/dS. \quad (2.15.6)$$

1 люкс (лк) — это освещённость поверхности площадью 1 м², создаваемая излучением, падающим нормально к этой поверхности, при световым потоке, равном 1 лм, 1 лк = 1 лм/1 м².

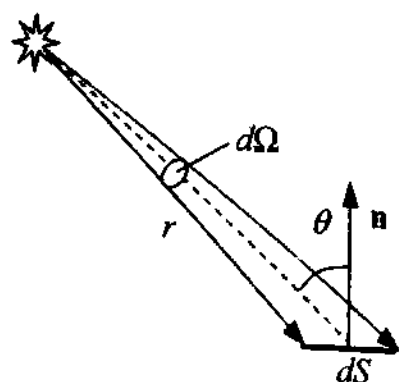


Рис. 2.15.2. К определению освещённости

Поскольку точечный источник с силой света J создает полный световой поток $4\pi J$, а на поверхности сферы радиуса r интенсивность света равна $I = 4\pi J/4\pi r^2 = J/r^2$, то люкс можно определить как такую освещённость, которую создает источник света силой в 1 кд на внутренней стороне поверхности шара радиуса 1 м, если он находится в центре и излучает изотропно.

Пусть источник света точечный. Тогда величина

$$d\Phi = Jd\Omega = J \frac{dS_{\perp}}{r^2}$$

есть поток, падающий на площадку $dS_{\perp} = dS \cos \theta$, где θ — угол падения излучения на поверхность. В результате оказывается

$$E = \frac{d\Phi}{dS} = \frac{J}{r^2} \cos \theta. \quad (2.15.7)$$

Таким образом, освещённость, создаваемая точечным источником, обратно пропорциональна квадрату расстояния до него и прямо пропорциональна косинусу угла между направлением падающих лучей и нормалью к освещаемой поверхности.

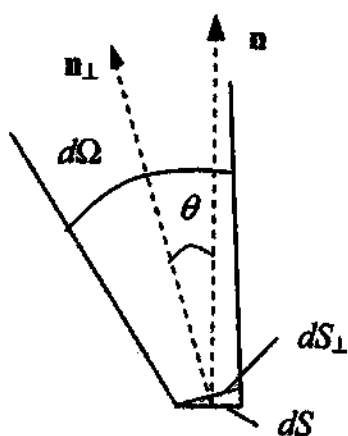


Рис. 2.15.3. К определению яркости. Рассматривается излучение с поверхности dS в телесный угол $d\Omega$. Элемент поверхности dS с единичным вектором нормали n наблюдается под углом θ к n . Видимая площадь есть $dS_{\perp} = dS \cdot \cos \theta$

2.15.5. Яркость

Это понятие характеризует светящиеся поверхности и неприменимо к точечным источникам. Световой поток, исходящий с площадки в единичный телесный угол, отнесённый к единице её видимой площади (рис. 2.15.3):

$$B = \frac{d\Phi}{dS_{\perp} d\Omega} = \frac{dJ}{dS_{\perp}} = \frac{dJ}{dS \cos \theta}. \quad (2.15.8)$$

Эту величину называют также *испуска-*

тельной (излучательной) способностью тела в заданном направлении.

Перепишем определение (2.15.8) в следующем виде:

$$d\Phi = B \cdot d\Omega \cdot dS \cos \theta. \quad (2.15.9)$$

Здесь использовано определение силы света (2.15.3). Единица измерения яркости — кандела на квадратный метр, $\text{кд}/\text{м}^2 = \text{лм}/(\text{ср} \cdot \text{м}^2)$.

Рассматривают также энергетическую яркость B_Σ . Размерность этой величины есть $[B_\Sigma] = \text{Вт}/(\text{ср} \cdot \text{м}^2)$.

Если яркость не зависит от направления ($B = \text{const}$), то источник называется *ламбертовским* (излучающим по закону Ламберта). Тогда сила света, создаваемого площадкой ΔS , будет равна

$$\Delta J = B \cdot \Delta S \cos \theta. \quad (2.15.10)$$

2.15.6. Светимость

Светимостью R протяжённого источника называется световой поток, испускаемый единичной площадкой *наружу* по всем направлениям, то есть в телесный угол 2π ($0 \leq \theta \leq \pi/2$):

$$R = \int B(\theta) \cos \theta d\Omega = 2\pi \int_0^{\pi/2} B(\theta) \cos \theta \sin \theta d\theta. \quad (2.15.11)$$

Для ламбертовского источника $B = \text{const}$, так что

$$R = \pi B.$$

2.15.6. Яркость и освещённость изображения

Пусть оптическая система создаёт некоторое изображение светящегося предмета (рис. 2.15.4). Ограничимся случаем, когда изображение формируется параксиальными пучками света.

Найдём сначала яркость изображения, рассматривая последнее как светящийся объект. Пусть исходный предмет есть квадрат со стороной y , а его изображение представляет собой квадрат со стороной y' . Пусть яркость предмета равна B . Тогда световой поток, создаваемый предметом, равен

$$\begin{aligned} d\Phi &= B(\theta) d\Omega dS \cos \theta = \\ &= 2\pi B(\theta) dS \sin \theta \cos \theta d\theta \approx 2\pi B(\theta) dS \theta d\theta \end{aligned} \quad (2.15.12)$$

(учтено, что рассматриваются только малые углы). Рассматривая изображение как светящийся предмет, запишем для него аналогичное равенство:

$$\begin{aligned} d\Phi' &= B'(\theta') d\Omega' dS' \cos \theta' = \\ &= 2\pi B'(\theta') dS' \sin \theta' \cos \theta' d\theta' \approx 2\pi B'(\theta') dS' \theta' d\theta'. \end{aligned} \quad (2.15.13)$$

Считая, что потери энергии в оптической системе малы, получаем, что от изображения идёт такой же световой поток: $d\Phi' = d\Phi$, или

$$2\pi B(\theta) dS \theta d\theta = 2\pi B'(\theta') dS' \theta' d\theta'. \quad (2.15.14)$$

Согласно теореме Лагранжа–Гельмгольца

$$ny\theta = n'y'\theta'. \quad (2.15.15)$$

Возводя это равенство почленно в квадрат и производя замену $y^2 \rightarrow dS$, $y'^2 \rightarrow dS'$, получаем

$$n^2 dS \theta^2 = n'^2 dS' \theta'^2 \Rightarrow n^2 dS \theta dq = n'^2 dS' \theta' d\theta'. \quad (2.15.16)$$

Из сравнения (2.15.14) и (2.15.16) следует

$$B'(\theta') = \left(\frac{n'}{n} \right)^2 B(\theta). \quad (2.15.17)$$

Отсюда следует, что если источник ламбертовский (его яркость не зависит от направления), то и его изображение также оказывается ламбертовским. В частном случае, когда по обе стороны от оптической системы показатели преломления среды одинаковы ($n = n'$), приходим к выводу, что яркость предмета и изображения одинаковы:

$$B' = B. \quad (2.15.18)$$

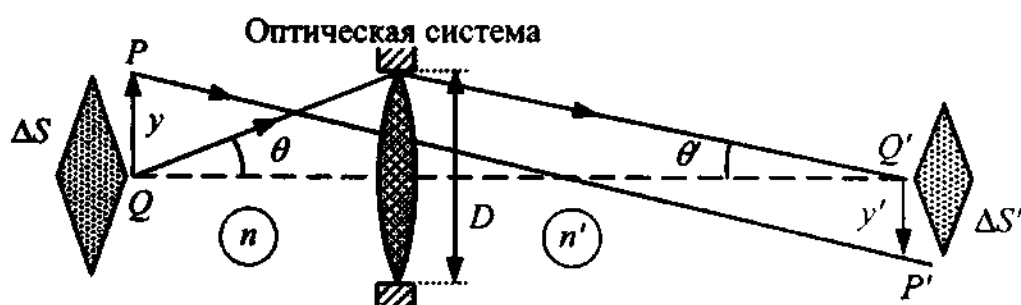


Рис. 2.15.4. К расчёту яркости и освещённости изображения $\Delta S'$ предмета ΔS

Пусть теперь изображение проектируется на экран. В этом случае мы имеем дело с освещённостью изображения. Последняя определяется соотношением (2.15.6) как световой поток, приходящийся на единицу площади освещаемой поверхности, и потому не связана с направлением излучения, исходящего от данного изображения.

Рассмотрим идеальную оптическую систему без потерь, причём будем считать, что по обе стороны от неё показатели преломления одинаковые: $n = n' = 1$. Имея в виду, что световой поток от предмета равен $\Delta\Phi = B\Delta\Omega\Delta S$, получим

$$E = \frac{\Delta\Phi}{\Delta S'} = B\Delta\Omega \frac{\Delta S}{\Delta S'}, \quad (2.15.19)$$

где ΔS и $\Delta S'$ — площади соответственно источника и изображения (рис. 2.15.4). Для параксиальных световых пучков ($|\theta| \ll 1$)

$$\Delta\Omega = \int_0^\theta 2\pi \sin \theta_1 d\theta_1 = 2\pi(1 - \cos \theta) \approx \pi\theta^2. \quad (2.15.20)$$

Учтём далее равенство

$$\Delta S'/\Delta S = \beta^2 = 1/\gamma^2, \quad \gamma = \theta'/\theta,$$

где β и γ — соответственно линейное и угловое увеличения. Согласно теореме Лагранжа–Гельмгольца

$$\theta y = \theta' y' \Rightarrow \theta^2 \Delta S = \theta'^2 \Delta S' \Rightarrow \Delta\Omega \Delta S = \Delta\Omega' \Delta S'. \quad (2.15.21)$$

Соответственно из (2.15.19) получаем:

$$E = B\Delta\Omega'. \quad (2.15.22)$$

Отсюда видно, что чем сильнее фокусируется свет от излучающего объекта, тем больше освещённость изображения. Результат достаточно очевиден: уменьшение площади пятна при одной и той же мощности излучения приводит к увеличению интенсивности в пятне.

Если предмет находится далеко от оптической системы, то его изображение оказывается практически в задней фокальной плоскости, то есть на расстоянии f' от оптической системы. Захватываемый системой световой поток определяется её диаметром D . Поэтому для величины $\Delta\Omega'$ можно записать выражение

$$\theta = \frac{D/2}{f'} \Rightarrow \Delta\Omega' = \pi\theta^2 = \frac{\pi(D/2)^2}{f'^2}. \quad (2.15.23)$$

Таким образом, получаем

$$E = \frac{\pi}{4} B \left(\frac{D}{f'} \right)^2. \quad (2.15.24)$$

Величина D/f' называется *относительным отверстием*, а $(D/f')^2$ — (геометрической) *светосилой* объектива.

ГЛАВА 3

ИНТЕРФЕРЕНЦИЯ ВОЛН

3.1. Интерференция монохроматических волн

В этом и ряде следующих разделах мы будем говорить об общих свойствах волн, в ряде случаев неспецифических для волн электромагнитных. Рассматривая какую-либо определённую физическую систему, можно конкретизировать смысл параметров, описывающих соответствующие волны. Те случаи, где природа волн существенна, оговариваются особо.

Пусть в пространстве распространяются две монохроматические волны с одной и той же частотой ω

$$A_1(\mathbf{r}, t) = a_1 \cos[\varphi_1(\mathbf{r}) - \omega t], \quad A_2(\mathbf{r}, t) = a_2 \cos[\varphi_2(\mathbf{r}) - \omega t]. \quad (3.1.1)$$

Рассмотрим суперпозицию этих волн:

$$A = A_1 + A_2.$$

Определим интенсивность волны соотношением

$$I = \overline{A^2} = \frac{1}{T} \int_0^T A^2(\mathbf{r}, t) dt, \quad (3.1.2)$$

где усреднение выполняется по периоду колебаний T .

Данное определение может отличаться численным множителем от принятого для конкретных типов физических систем и не существенно для дальнейшего рассмотрения. В частности, в случае электромагнитных волн в вакууме $I = cE^2/8\pi$.

В принятых обозначениях интенсивность суперпозиции волн равна

$$I = \overline{A^2} = \overline{(A_1 + A_2)^2} = \overline{A_1^2} + \overline{A_2^2} + 2\overline{A_1 A_2}.$$

Для рассматриваемых волн справедливы соотношения

$$\overline{A_1^2} = \frac{1}{2} a_1^2 = I_1, \quad \overline{A_2^2} = \frac{1}{2} a_2^2 = I_2,$$

$$2\overline{A_1 A_2} = a_1 a_2 \left[\overline{\cos[\varphi_1(\mathbf{r}) + \varphi_2(\mathbf{r}) - 2\omega t]} + \overline{\cos[\varphi_1(\mathbf{r}) - \varphi_2(\mathbf{r})]} \right] = a_1 a_2 \cos[\Delta\varphi(\mathbf{r})] = 2\sqrt{I_1 I_2} \cos[\Delta\varphi(\mathbf{r})].$$

Здесь введено обозначение

$$\Delta\varphi(\mathbf{r}) = \varphi_1(\mathbf{r}) - \varphi_2(\mathbf{r})$$

и учтено равенство $\overline{\cos[\varphi_1(\mathbf{r}) + \varphi_2(\mathbf{r}) - 2\omega t]} = 0$. Таким образом, получаем

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos[\Delta\varphi(\mathbf{r})]. \quad (3.1.3)$$

Видно, что интенсивность суммарной волны не равна сумме интенсивностей исходных волн.

Явление наложения волн, при котором интенсивность результирующей волны не равна всюду сумме интенсивностей складываемых волн, называется *интерференцией* (от англ. *interference* — вмешательство, помеха).

3.2. Видность интерференционной картины

Для характеристики выраженности (контрастности) интерференции вводят величину, называемую *видностью*:

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}. \quad (3.2.1)$$

Здесь $I_{\max}$ и $I_{\min}$ — максимальное и минимальное значения интенсивности в области интерференции волн. Величина V меняется в пределах $0 \leq V \leq 1$.

Интерференция отсутствует, если $I_{\max} = I_{\min}$, или $V = 0$. Интерференция выражена наиболее сильно, если $I_{\min} = 0$, или $V = 1$.

Найдём видность интерференционной картины, возникающей при интерференции двух монохроматических волн.

Согласно (3.1.3) в тех точках, где $\cos \Delta\varphi = +1$, или $\Delta\varphi = 2\pi m$, $m = 0, \pm 1, \pm 2, \dots$, оказывается

$$I = I_{\max} = (\sqrt{I_1} + \sqrt{I_2})^2. \quad (3.2.2)$$

В тех же точках, где $\cos \Delta\varphi = -1$, или $\Delta\varphi = \pi(2m + 1)$, $m = 0, \pm 1, \pm 2, \dots$, оказывается

$$I = I_{\min} = (\sqrt{I_1} - \sqrt{I_2})^2. \quad (3.2.3)$$

В соответствии с определением (3.2.1) находим видность:

$$V = \frac{2\sqrt{I_1 I_2}}{I_1 + I_2}. \quad (3.2.4)$$

Отсюда следует, что контрастность интерференционной картины максимальна: $V = 1$, если складываемые волны имеют одинаковую амплитуду.

3.3. Интерференция плоских волн

Рассмотрим частный случай, когда интерферируют две плоские волны, имеющие одинаковую частоту:

$$\begin{aligned} A_1(\mathbf{r}, t) &= a_1 \cos(\mathbf{k}_1 \mathbf{r} - \omega t + \varphi_1), \\ A_2(\mathbf{r}, t) &= a_2 \cos(\mathbf{k}_2 \mathbf{r} - \omega t + \varphi_2). \end{aligned} \quad (3.3.1)$$

Согласно (3.1.3) распределение интенсивности суммарной волны имеет вид

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos(\mathbf{K} \mathbf{r} + \delta), \quad \mathbf{K} = \mathbf{k}_1 - \mathbf{k}_2, \quad \delta = \varphi_1 - \varphi_2. \quad (3.3.2)$$

В тех точках, где $\mathbf{K} \mathbf{r} = \text{const}$, оказывается $I = \text{const}$. Но условие $\mathbf{K} \mathbf{r} = \text{const}$ определяет плоскость, перпендикулярную вектору $\mathbf{K}$. Следовательно, на экране будет наблюдаться серия полос, перпендикулярных вектору $\mathbf{K}$ (рис. 3.3.1). При этом в плоскостях, где

$$\mathbf{K} \mathbf{r} + \delta = 2\pi m, \quad m = 0, \pm 1, \pm 2, \dots, \quad (3.3.3)$$

интенсивность достигает максимума:

$$I = I_{\max} = (\sqrt{I_1} + \sqrt{I_2})^2, \quad (3.3.4)$$

а в плоскостях, где

$$\mathbf{K} \mathbf{r} + \delta_0 = \pi(2m + 1), \quad m = 0, \pm 1, \pm 2, \dots, \quad (3.3.5)$$

интенсивность минимальна:

$$I = I_{\min} = (\sqrt{I_1} - \sqrt{I_2})^2. \quad (3.3.6)$$

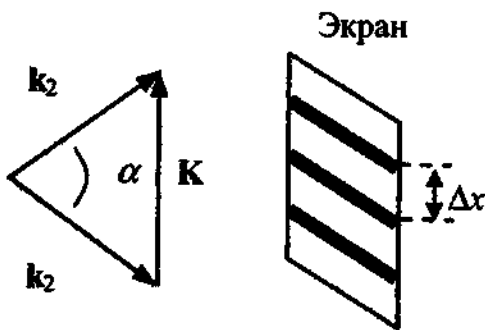


Рис. 3.3.1. Интерференция плоских волн, распространяющихся под углом α друг к другу

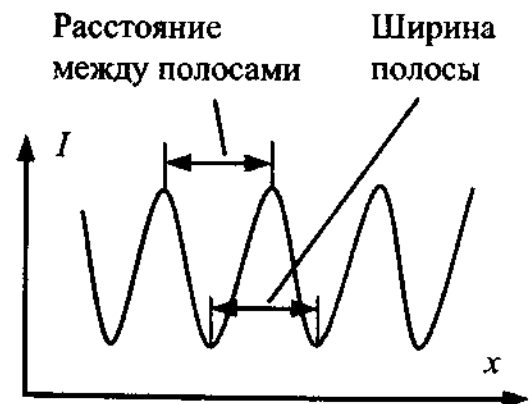


Рис. 3.3.2. Распределение интенсивности интерференционной картины на экране. Координата x вдоль экрана отсчитывается параллельно вектору $\mathbf{K}$

Расстояние между ближайшими максимумами называется *расстоянием между полосами*, а расстояние между ближайшими минимумами называется *шириной полосы* (рис. 3.3.2).

Поскольку

$$K = |\mathbf{k}_1 - \mathbf{k}_2| = 2k \sin(\alpha/2),$$

то из формулы (3.3.5) следует, что расстояние между полосами определяется из условия $K \Delta x = 2\pi$, или

$$\Delta x = \frac{2\pi}{K} = \frac{2\pi}{2k \sin(\alpha/2)} = \frac{\lambda}{2 \sin(\alpha/2)}. \quad (3.3.7)$$

Этой же формулой в рассматриваемой ситуации определяется и ширина интерференционных полос. Минимальная ширина полосы имеет место для встречных волн (то есть при $\alpha = \pi$) и составляет $\Delta x = \lambda/2$. При этом распределение интенсивности согласно (3.3.2) имеет вид

$$I(\mathbf{r}) = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos(2\mathbf{k}\mathbf{r} + \delta). \quad (3.3.8)$$

Если же волны распространяются в одном направлении ($\alpha = 0$), то $\Delta x = \infty$. В случае, когда угол α между направлениями распространения волн мал, формула (3.3.7) упрощается:

$$\Delta x = \lambda/\alpha. \quad (3.3.9)$$

3.4. Интерференция волн от одного источника

В ряде случаев для получения интерференционной картины желательно использовать единственный источник волн. В этом разделе мы рассмотрим некоторые оптические схемы, реализующие это требование.

3.4.1. Схема Юнга

Одной из важнейших интерференционных схем с одним источником волн является схема Юнга (1802 г.). Она представлена на рис. 3.4.1 и состоит в том, что лучи от источника S проходят через два отверстия: S_1 и S_2 , а затем интерферируют на экране $\mathcal{E}$. Томас Юнг ставил на пути света от источника до экрана с щелями дополнительную диафрагму, чтобы выделить достаточно узкий пучок лучей.

Данный эксперимент, выполненный английским физиком Томасом Юнгом во времена господства ньютоновской корпускулярной теории света, стал экспериментальным доказательством волновой теории. Результаты эксперимента были опубликованы в 1803 г. Юнг ввёл также термин «интерференция». Ему принадлежат исследования и в других разделах физики. Так, в теории упругости он ввёл характеристику, называемую сейчас модулем Юнга (1807 г.). Юнгу принадлежит и термин «энергия» (1807 г.).

Найдём параметры интерференционной картины в этой схеме.

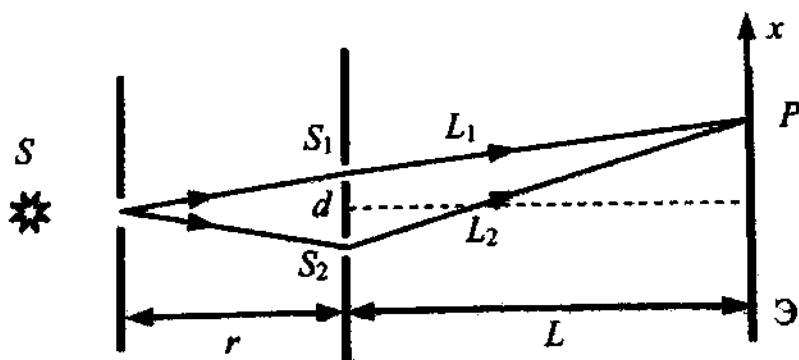


Рис. 3.4.1. Схема Юнга по наблюдению интерференционной картины от источника S

Найдём разность хода лучей от щелей S_1 и S_2 до точки на экране Э, имеющей координату x . Выбирая начало координат $x = 0$ в точке, равноудалённой (по вертикали) от обеих щелей S_1 и S_2 , на основании теоремы Пифагора имеем

$$L_1^2 = L^2 + \left(x - \frac{d}{2}\right)^2, \quad L_2^2 = L^2 + \left(x + \frac{d}{2}\right)^2.$$

Отсюда следует $L_2^2 - L_1^2 = 2xd$. Рассмотрим приосевую область интерференционной картины, в которой $d \ll L$, $|x| \ll L$. Имея в виду, что в этой области

$$L_1 + L_2 \approx 2L, \quad L_2^2 - L_1^2 \approx 2L(L_2 - L_1),$$

находим разность хода лучей:

$$\delta = L_2 - L_1 \approx xd/L.$$

Максимумы интерференционной картины наблюдаются в точках, для которых

$$\delta = m\lambda, \quad m = 0, \pm 1, \pm 2, \dots$$

При этом

$$x_m^{(\max)} = \frac{\lambda L}{d} m.$$

Минимумы интерференционной картины наблюдаются в точках, для которых

$$\delta = \left(m + \frac{1}{2}\right)\lambda, \quad m = 0, \pm 1, \pm 2, \dots$$

При этом

$$x_m^{(\min)} = \frac{\lambda L}{d} \left(m + \frac{1}{2}\right).$$

Ширина интерференционной полосы равна

$$\Delta x = x_{m+1}^{(\min)} - x_m^{(\min)} = \frac{\lambda L}{d}.$$

В рассматриваемой задаче эта величина совпадает с расстоянием между полосами:

$$\Delta x = x_{m+1}^{(\max)} - x_m^{(\max)} = \frac{\lambda L}{d}.$$

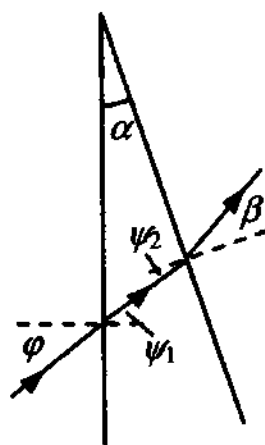
3.4.2. Отклонение лучей клином

Рассмотрим преломление лучей на клине с малым углом при вершине α и показателем преломления n (рис. 3.4.2).

Пусть луч падает на первую грань клина под малым углом φ к нормали. Тогда в соответствии с законом Снеллиуса он преломится, и внутри клина будет распространяться под углом ψ_1 (к той же нормали), таким что

$$\varphi = n\psi_1.$$

Рис. 3.4.2. Клин с малым преломляющим углом α . Луч, падающий на клин под углом φ к нормали с одной стороны, выходит под углом β к нормали с другой стороны. Штриховыми линиями показаны нормали к соответствующим сторонам клина



Поскольку угол при вершине клина равен α , то со второй нормалью этот луч будет образовывать угол

$$\psi_2 = \psi_1 - \alpha.$$

Из клина луч выйдет под углом ко второй нормали

$$\beta = n\psi_2 = n(\psi_1 - \alpha) = \varphi - n\alpha.$$

Рассматриваемый луч будет теперь образовывать с первой нормалью угол

$$\beta + \alpha = \varphi - (n - 1)\alpha.$$

В результате луч повернётся на угол

$$\varphi - [\varphi - (n - 1)\alpha] = (n - 1)\alpha.$$

Отсюда, в частности, следует, что при малых углах падения итоговый поворот луча не зависит от угла падения.

3.4.3. Бипризма Френеля

Бипризма — это оптический прибор, составленный из двух призм с малым преломляющим углом α (рис. 3.4.3). Свет, создаваемый источником S , проходит через верхнюю и нижнюю половины бипризмы, превращаясь в два пучка, распространяющихся в разных направлениях. Эти пучки можно рассматривать как производимые двумя мнимыми источниками: S_1 и S_2 . В результате в области перекрытия пучков наблюдается их интерференция.

Имея в виду сказанное в разделе 3.4.2, рассмотрим интерференцию волн, создаваемых бипризмой. Поскольку все лучи поворачиваются половинами бипризмы на один и тот же угол

$$\gamma = (n - 1)\alpha,$$

но в противоположных направлениях, то вместо одного действительного источника S возникают два мнимых: S_1 и S_2 (рис. 3.4.3). Найдём положение этих источников.

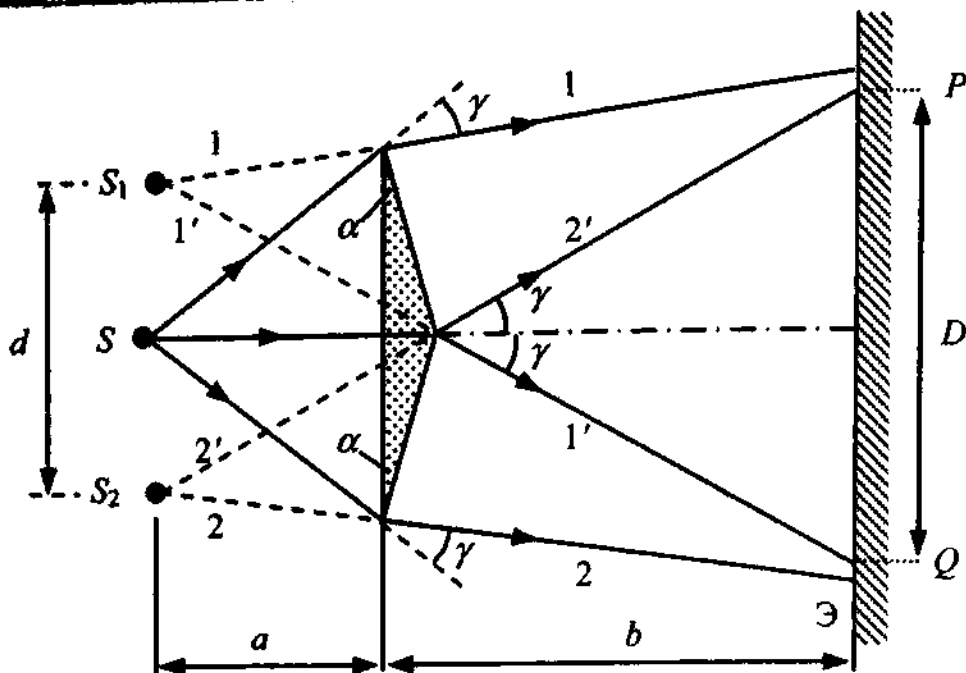


Рис. 3.4.3. Бипризма Френеля. В результате преломления света, идущего от источника S , в верхней и нижней половинах бипризмы возникают два пучка, которые можно рассматривать как создаваемые мнимыми источниками S_1 и S_2 . Интерференция этих пучков наблюдается на экране $\mathcal{E}$. Источник S находится на оси симметрии системы. Область перекрытия пучков на экране расположена между точками P и Q

Поскольку преломляющий угол призмы мал, то расстояние от всех источников (S , S_1 и S_2) до призмы приблизительно одинаково и равно a . Соответственно расстояние между мнимыми источниками составит

$$d = 2a \operatorname{tg} \gamma \approx 2a\gamma = 2a(n-1)\alpha.$$

Таким образом, задача свелась к расчёту интерференционной картины от двух источников с расстоянием d между ними, удалённых от экрана на расстояние $L = a + b$ (рис. 3.4.3, 3.4.4).

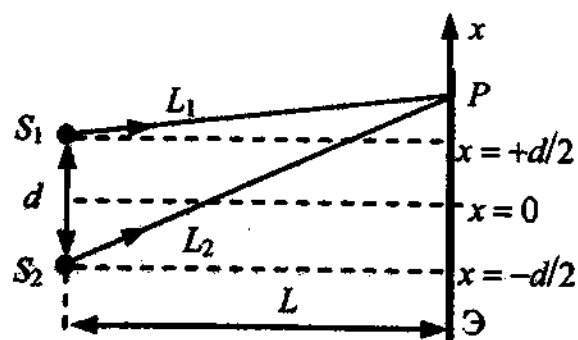


Рис. 3.4.4. К расчёту интерференции волн от двух источников (S_1 и S_2), удалённых от экрана $\mathcal{E}$ на расстояние L

Найдём положение интерференционных максимумов и минимумов. Выберем некоторую точку P с координатой x на экране. Имеем равенства:

$$L_1^2 = L^2 + (x - d/2)^2,$$

$$L_2^2 = L^2 + (x + d/2)^2.$$

Отсюда следует

$$L_2^2 - L_1^2 = 2xd.$$

Рассматривая приосевую область на экране, учтём, что расстояния L_1 и L_2 мало отличаются от L , так что $L_1 + L_2 \approx 2L$. Тогда для разности хода лучей получаем выражение

$$\delta = L_2 - L_1 = xd/L.$$

Положение интерференционных максимумов определяется из условия $\delta = m\lambda$, $m = 0, \pm 1, \pm 2, \dots$, или

$$x_m^{(\max)} = \frac{\lambda L}{d} m, m = 0, \pm 1, \pm 2, \dots$$

Положение интерференционных минимумов даётся условием $\delta = (m + 1/2)\lambda$, или

$$x_m^{(\min)} = \frac{\lambda L}{d} \left(m + \frac{1}{2} \right), m = 0, \pm 1, \pm 2, \dots$$

Таким образом, расстояние между интерференционными максимумами (как и ширина максимумов) составляет

$$\Delta x = \frac{\lambda L}{d} = \frac{(a+b)\lambda}{2a(n-1)\alpha}.$$

В случае бипризмы (рис. 3.4.3) область перекрытия пучков на экране наблюдения расположена между точками P и Q . Размер этой области составляет $D = 2b \operatorname{tg} \gamma \approx 2\beta(n-1)\alpha$. Соответственно максимальное число полос, которое могло бы наблюдаться на экране в рассматриваемых условиях, равно

$$N = \frac{D}{\Delta x} = \frac{4ab(n-1)^2 \alpha^2}{\lambda(a+b)}.$$

3.4.4. Интерференция волн при отражении от пластины

Пусть луч падает на плоскопараллельную пластину толщиной h с показателем преломления n под некоторым углом α к нормали (рис. 3.4.5). Этот луч может сразу отразиться от верхней поверхности (порождая луч 1), а может сначала попасть внутрь пластины, а затем, отразившись от нижней грани, выйти из пластины через верхнюю поверхность (порождая луч 2). В результате каждый луч расщепляется на два, которые могут интерферировать между собой. Строго говоря, возможны многократные отражения и преломления луча. Однако, поскольку фазовые соотношения для последующих лучей остаются теми же, что и для пары 1 и 2, для наших целей достаточно ограничиться рассмотрением указанной пары.

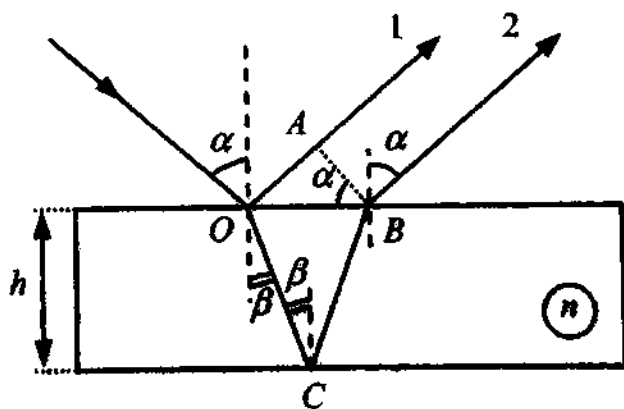


Рис. 3.4.5. К расчёту интерференции волн при отражении от пластины

Найдём оптическую разность хода лучей 1 и 2. Проведём отрезок AB перпендикулярно лучам 1 и 2. Тогда искомая длина набирается на трассах лучей до их пересечения с отрезком AB и составляет

$$\delta_1 = L_2 - L_1 = n \cdot OCB - OA.$$

Длина трассы луча 2 в пластине составляет

$$OCB = \frac{2h}{\cos \beta},$$

где β — угол преломления, даваемый законом Снеллиуса:

$$\sin \alpha = n \sin \beta.$$

Для нахождения длины отрезка OA найдём горизонтальный ход луча 2 в пластине:

$$OB = 2h \operatorname{tg} \beta.$$

Отсюда следует

$$OA = OB \sin \alpha = 2h \operatorname{tg} \beta \sin \alpha.$$

Таким образом, получаем

$$\delta_1 = n \frac{2h}{\cos \beta} - 2h \operatorname{tg} \beta \sin \alpha = \frac{2hn}{\cos \beta} \left(1 - \frac{1}{n} \sin \beta \sin \alpha \right).$$

С помощью закона Снеллиуса отсюда находим

$$\delta_1 = 2hn \cos \beta = 2h \sqrt{n^2 - \sin^2 \alpha}.$$

Помимо полученной разности хода вклад в сдвиг фаз лучей может давать и отражение от поверхностей. Согласно формулам Френеля амплитудный коэффициент отражения при нормальном падении волны равен

$$r = \frac{E_{\text{отр}}}{E_{\text{пад}}} = \frac{1-n}{1+n}$$

(формула (1.7.6)). Следовательно, фаза волны меняется на π при отражении от среды оптически более плотной (с бóльшим показателем преломления) и не меняется при отражении от среды оптически менее плотной (с меньшим показателем преломления). Такая же ситуация имеет место для s -поляризованной волны при падении под любыми углами. Для p -поляризованной волны изменение фазы на π имеет место для углов падения, не превышающих угла Брюстера: $\alpha < \theta_B = \operatorname{arctg} n$ (при котором коэффициент отражения обращается в нуль). При бóльших же углах фаза волны не меняется. Сказанное следует из формул Френеля (1.7.3), (1.7.5).

Таким образом, если угол падения не превышает угла Брюстера, то из найденной разности хода лучей δ следует вычесть поправку $\lambda/2$, обусловленную отражением луча 1 от оптически более плотной среды:

$$\delta = \delta_1 - \lambda/2$$

(λ — длина волны в вакууме).

Интерференционные максимумы интенсивности должны наблюдаться, если $\delta = m\lambda$, $m = 0, 1, 2, \dots$, или

$$2h\sqrt{n^2 - \sin^2 \alpha} = \left(m + \frac{1}{2}\right)\lambda.$$

Это равенство выделяет те углы падения лучей, при которых волны благодаря интерференции усиливаются. Если же выполняется условие

$$\delta = \left(m + \frac{1}{2}\right)\lambda, \text{ или}$$

$$2h\sqrt{n^2 - \sin^2 \alpha} = m\lambda,$$

то происходит интерференционное гашение волн.

Для наблюдения интерференционных полос на пути отражённых лучей следует поставить собирающую линзу (рис. 3.4.6).

Известно, что линза собирает параллельный пучок лучей в точку в фокальной плоскости. Эта точка (в случае тонкой линзы) находится на пересечении фокальной плоскости с лучом, параллельным пучку и проходящим через оптический центр линзы (рис. 3.4.6). В зависимости от угла падения на пластину отражённые лучи будут собираться в разных точках фокальной плоскости. Пусть падающий на пластинку свет имеет разные направления (рассеянный свет). Тогда для лучей с направлениями, отвечающими интерференционному усилению, будут наблюдаться светлые полосы. Лучи же, для которых имеет место интерференционное подавление, будут на экране формировать тёмные полосы.

Поскольку в данном эксперименте каждая из полос формируется только группой лучей с одним и тем же углом падения (наклона), то такие линии на экране называются *полосами равного наклона*.

3.4.5. Опыт Поля

С задачей, рассмотренной в предыдущем разделе, связан один из способов наблюдения интерференционной картины от точечного источника, предложенный немецким физиком Робертом Подем. Этот способ проиллюстрирован на рис. 3.4.7.

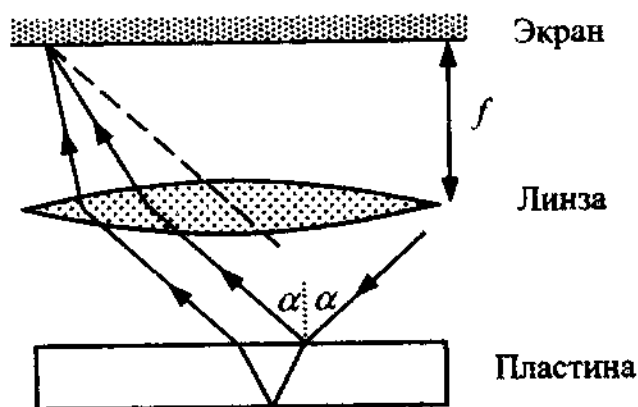


Рис. 3.4.6. Схема наблюдения интерференционных полос равного наклона при отражении света от пластины. Штриховой линией показано положение боковой оптической оси, проходящей через оптический центр линзы

В опыте используется точечный источник, свет от которого падает на стеклянную пластинку. В результате отражения от верхней и нижней грани пластины формируются системы лучей со сдвигом фаз, определяемым углом падения исходных лучей на пластину (рис. 3.4.7 слева). В этом опыте на экране наблюдается система интерференционных колец.

Вторичные лучи, идущие от пластины, можно рассматривать как создаваемые мнимыми источниками S_1 и S_2 (рис. 3.4.7 справа). Найдём положение этих источников.

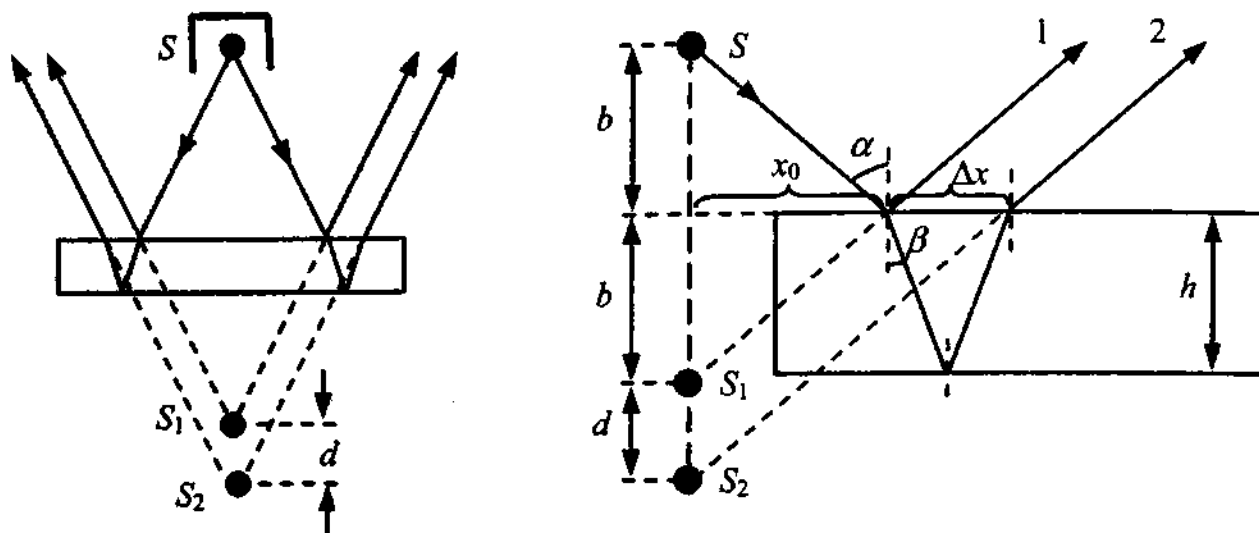


Рис. 3.4.7. Опыт Поля по наблюдению полос (колец) равного наклона. Слева — формирование двух мнимых источников S_1 и S_2 при отражении от пластины волн, идущих от точечного источника S . Справа — расчёт положения источников S_1 и S_2 . Экран сзади от источника S (на рисунке слева) выделяет конус лучей, идущих только к пластине, так что в формировании интерференционной картины участвуют только лучи, отражённые от пластины

Мнимые источники находятся на продолжении лучей, отражённых от пластины. Пусть исходный источник находится на расстоянии h от пластины. Тогда первый мнимый источник (S_1) находится симметрично ему относительно верхней грани пластины. Для нахождения положения второго мнимого источника (S_2) заметим, что луч 1 до места отражения смещается вдоль поверхности пластины на расстояние

$$x_0 = b \operatorname{tg} \alpha.$$

Горизонтальный ход луча 2 до момента выхода из пластины равен

$$\Delta x = 2h \operatorname{tg} \beta = 2h \frac{\sin \beta}{\cos \beta}.$$

По закону Снеллиуса $n \sin \beta = \sin \alpha$. Поэтому для Δx получаем выражение

$$\Delta x = \frac{2h \sin \alpha}{\sqrt{n^2 - \sin^2 \alpha}}.$$

Расстояние d между мнимыми источниками определится из равенства

$$x_0 + \Delta x = (b + d) \operatorname{tg} \alpha.$$

Используя найденные выражения для x_0 и Δx , находим отсюда

$$d = \frac{x_0 + \Delta x}{\operatorname{tg} \alpha} - b = \frac{\Delta x}{\operatorname{tg} \alpha} = \Delta x = \frac{2h \cos \alpha}{\sqrt{n^2 - \sin^2 \alpha}}.$$

Если рассматриваются пучки лучей с небольшим углом расхождения ($\alpha \ll 1$), то оказывается, что расстояние между источниками составляет

$$d = 2h/n$$

и практически не зависит от угла падения лучей на пластину.

Таким образом, в опыте Поля, как и в схеме Юнга, наблюдается интерференция волн, полученных от одного источника. Преимущество метода Поля состоит в том, что интерференцию можно продемонстрировать в больших аудиториях благодаря высокой светосиле.

3.4.6. Билинза Бийе

Билинза Бийе — это тонкая собирающая линза, разрезанная по диаметру на две части (рис. 3.4.8). Затем эти части раздвигаются на некоторое расстояние, а образовавшееся свободное пространство между половинками линзы перекрывается непрозрачной пластинкой (препятствующей прохождению прямого света от источника). Возможен также вариант, когда от каждой из половин отрезается небольшая приосевая область, а оставшиеся части сдвигаются вместе.

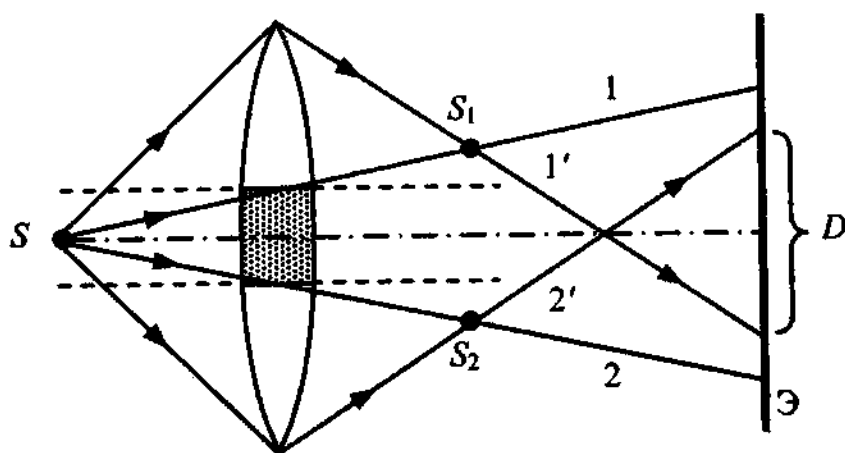


Рис. 3.4.8. Билинза Бийе. Половинки линзы раздвинуты, а образовавшаяся щель перекрыта непрозрачной пластинкой. Штриховыми линиями показаны главные оптические оси половинок линзы, а штрихпунктирной линией — главная оптическая ось (ось симметрии) системы. Половинки билинзы формируют два различных действительных изображения S_1 и S_2 точечного источника S . Интерференционная картина наблюдается на экране $\mathcal{E}$ в области D перекрытия пучков, образуемых половинками билинзы. Лучи 1 и 2 от источника S проходят через оптические центры половинок, лучи 1' и 2' проходят через крайние точки билинзы

В опытах с билинзой используется точечный источник S . Благодаря устройству билинзы источник, который находится на оптической оси системы, оказывается смещённым по отношению к главным оптическим осям половинок билинзы. Соответственно, если источник S помещён на рас-

стоянии, превышающем фокусное, то эти половинки формируют два различных действительных изображения S_1 и S_2 . При этом интерференционная картина наблюдается на экране в области D перекрытия пучков, создаваемых новыми источниками. Если же источник расположен в фокальной плоскости, то возникает пара плоских волн, распространяющихся под углом друг к другу.

3.4.7. Бизеркала Френеля

Ещё один способ расщепления волн от одного источника реализуется с помощью бизеркал Френеля (рис. 3.4.9). Эта оптическая система, предложенная Френелем в 1816 г., образована двумя зеркалами с малым углом между ними (несколько угловых минут). Цилиндрическая волна, создаваемая светящейся нитью S , расположенной перед зеркалами, отражается ими в разных направлениях. Отражённые волны можно рассматривать как волны, порождённые двумя мнимыми источниками S_1 и S_2 , что позволяет наблюдать их интерференцию.

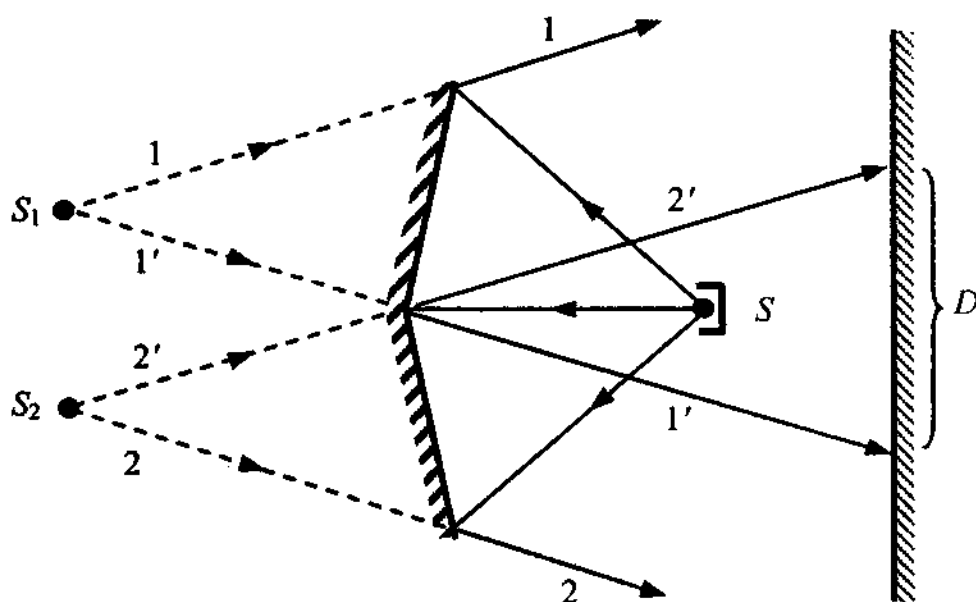


Рис. 3.4.9. Бизеркала Френеля расщепляют волну от цилиндрического источника S на две волны, идущие от мнимых источников S_1 и S_2 . Интерференция наблюдается в области D перекрытия пучков, отражённых от зеркал

3.4.8. Интерференция при отражении от клина

Интерференцию плоской волны с самой собой можно наблюдать, рассматривая её отражение от клина — системы двух тонких стеклянных пластинок, образующих малый угол друг с другом (рис. 3.4.10).

Плоская волна отражается от нижней грани клина (лучи 1 и 1'). Кроме того, пройдя через эту грань, волна отражается от верхней грани и выходит из клина через нижнюю грань (лучи 2 и 2'). Возникает пара плоских волн (представляемых лучами 1' и 2'), распространяющихся под углом 2α друг к другу. При этом возникает серия интерференционных полос, пер-

пендикулярных клину. При малом угле клина расстояние между полосами составляет $\Delta x = \lambda/2\alpha$.

Плоскую волну, падающую на клин, можно сформировать, используя точечный источник, расположенный в фокусе линзы. Таким образом, данная схема позволяет, как и в предыдущих примерах, реализовать интерференцию волн, создаваемых одним источником.

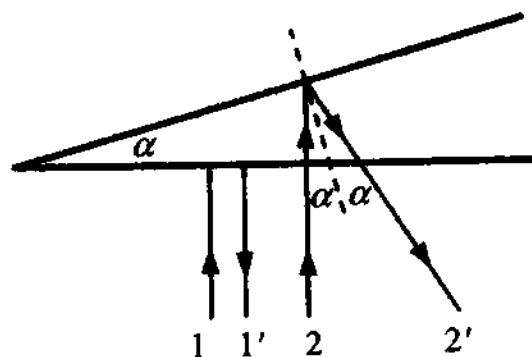


Рис. 3.4.10. Интерференция волн при отражении от клина — пары стеклянных пластинок, образующих малый угол друг с другом

3.4.9. Зеркало Ллойда

Эксперимент с зеркалом был впервые описан в 1834 году ирландским физиком Хэмфри Ллойдом. В этой схеме на экране интерферируют прямая волна от источника и волна, отражённая от зеркала (рис. 3.4.11).

В данной схеме один из лучей падает под малым углом скольжения α на отражающую поверхность. Отражённые от зеркала лучи можно рассматривать как порождённые мнимым источником S' , расположенным симметрично источнику S относительно отражающей поверхности. Таким образом, возникает пара источников, находящихся на некотором расстоянии друг от друга и порождающих интерференционную картину на экране Э.

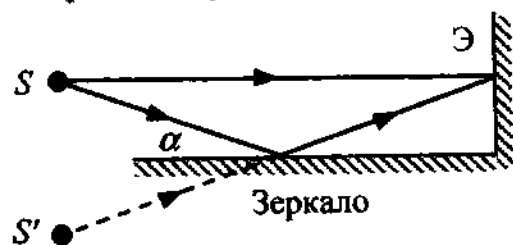


Рис. 3.4.11. Зеркало Ллойда

3.4.10. Кольца Ньютона

Опыт, ставший классическим, был описан Ньютоном в 1675 г.: свет падает на плосковыпуклую линзу с большим радиусом кривизны, лежащую на плоскопараллельной стеклянной пластинке (рис. 3.4.12). Регистрируется излучение, отражённое системой или прошедшее через неё. Оказалось, что при нормальном падении света на сферической поверхности локализована серия колец — интерференционных полос. Объяснение явлению дал Т. Юнг.

Зазор между линзой и пластинкой играет роль, аналогичную клину. Интерференция возникает между лучами, отражёнными от пластинки и от кривой поверхности линзы. Разность хода отражённых лучей зависит от толщины зазора d в месте падения луча 1 и составляет $2d$. Найдём её.

Пусть радиус кривизны линзы равен R . Тогда из теоремы Пифагора находим

$$R^2 = r^2 + (R - d)^2 \Rightarrow 2Rd = r^2.$$

Здесь, считая зазор тонким в области наблюдения, мы пренебрегли малой величиной d^2 по сравнению с удержанными слагаемыми. При отражении луча от оптически более плотной среды волна приобретает дополнительный сдвиг фазы π , соответствующий дополнительной разности хода $\lambda/2$. Поэтому находим оптическую разность хода лучей, отражённых от линзы и пластинки:

$$\Delta = 2d + \frac{\lambda}{2} = \frac{r^2}{R} + \frac{\lambda}{2}.$$

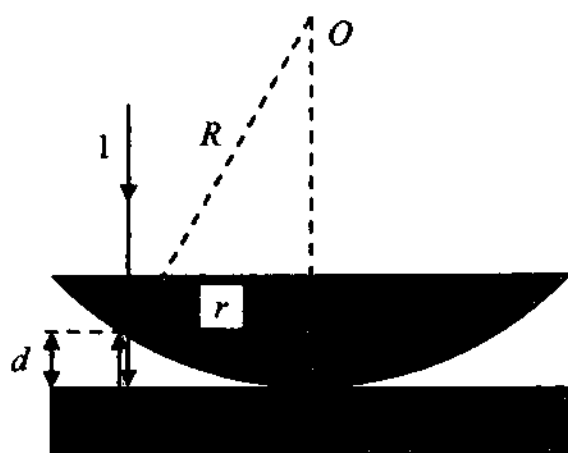


Рис. 3.4.12. Схема оптической системы, в которой наблюдаются кольца Ньютона: плоско-выпуклая линза, лежащая на поверхности плоскопараллельной пластинки



Рис. 3.4.13. Кольца Ньютона в отражённом свете

Тёмные кольца отвечают интерференционным минимумам интенсивности, наблюдающимся при условии $\Delta = (m + 1/2)\lambda$, откуда находим радиусы центров тёмных колец:

$$r_m = \sqrt{m\lambda R}, \quad m = 0, 1, 2, \dots$$

Из условия $\Delta = m\lambda$ получаем радиусы центров светлых колец:

$$r_m = \sqrt{(m - 1/2)\lambda R}, \quad m = 1, 2, \dots$$

Таким образом, в отражённом свете центр колец тёмный. Если бы мы рассматривали прошедший свет, то добавка $\lambda/2$ к разности хода не возникала вследствие того, что отражение луча от оптически более плотной среды двукратное. Соответственно тёмные и светлые кольца меняются местами, причём центр колец оказывается светлым.

В заключение данного раздела отметим следующее.

1. Наблюдающиеся кольца являются полосами *равной толщины*, поскольку их положение зависит от толщины зазора: волны, проходящие участки зазора, имеющие одну и ту же толщину, но в разных точках вокруг оси системы, образуют одно кольцо.

2. Радиусы колец зависят от длины волны. Поэтому систему тёмных и светлых колец можно наблюдать, если с помощью светофильтров выделять достаточно узкие участки спектра. Если же свет немонохроматический, то каждое кольцо будет иметь различную окраску в разных его точках: в точках, более удалённых от центра, окраска смещается в красную область спектра.

3. В расчётах мы считали, что в зазоре между линзой и пластинкой показатель преломления среды равным единице. Если это не так, то будет меняться разность хода отражённых лучей, приводя к изменению радиусов колец Ньютона.

4. Кольца Ньютона используются для измерения радиусов кривизны поверхностей, длин волн света и показателей преломления.

ГЛАВА 4

ВРЕМЕННАЯ КОГЕРЕНТНОСТЬ

4.1. Интерференция немонохроматических волн

4.1.1. Когерентность

В общем случае реальная волна, испускаемая физической системой (атомами, молекулами и т.д.), не является монохроматической. Например,

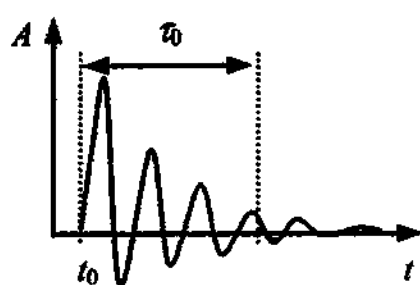


Рис. 4.1.1. Сигнал, создаваемый отдельным излучающим атомом

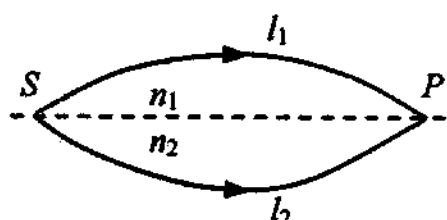


Рис. 4.1.2. Свет распространяется от источника S до точки P по двум различным путям

сигнал, испускаемый возбуждённым атомом, представляет собой затухающее колебание — *цуг*, или *волновой пакет* — с характерной длительностью τ_0 (рис. 4.1.1). Этот сигнал испускается, вообще говоря, в случайный момент времени t_0 . Поэтому при сложении волн от разных источников интерференционная картина может подавляться или вовсе пропадать из-за наложения отдельных цугов со случайными сдвигами фаз.

Можно однако попытаться получить интерференционную картину, используя единственный источник. Для этого нужно каким-либо образом «рассечь» каждый цуг волн от этого источника на два и, заставив их распространяться по разным направлениям, свести вновь в других точках

(рис. 4.1.2). Некоторые методы рассмотрены в разделе 3.4.

Тот факт, что сигнал представляет собой последовательность цугов, означает, что этот сигнал немонохроматический и модулирован как по фазе, так и по амплитуде:

$$A(t) = a(t) \cos(\omega_0 t + \alpha(t)),$$

где $a(t)$ и $\alpha(t)$ — медленно меняющиеся функции (по сравнению с $\cos \omega_0 t$). Рассмотрим, какие ограничения накладываются при этом на наблюдаемость интерференции немонохроматичностью излучения.

Две волны называются *когерентными*, если разность их фаз является постоянной. Когерентными являются две монохроматические волны, если только они имеют одинаковые частоты.

Если разность фаз волн меняется со временем, то эти волны называются *некогерентными*.

Термин «когерентность» происходит от латинского *cohaerens* — находящийся в связи.

4.1.2. Время и длина когерентности

Начнём с простейшего случая сложения двух волн с разными частотами ω_1 и ω_2 .

Временем когерентности называется такое время, в течение которого разность фаз рассматриваемых волн меняется незначительно.

В случае двух волн (рис. 4.1.3)

$$\begin{aligned} A_1 &= a_1 \cos(\omega_1 t + \alpha_1), \\ A_2 &= a_2 \cos(\omega_2 t + \alpha_2) \end{aligned} \quad (4.1.1)$$

разность фаз равна

$$\Delta\varphi = \Delta\omega \cdot t + \Delta\alpha, \quad \Delta\omega = \omega_1 - \omega_2, \quad \Delta\alpha = \alpha_1 - \alpha_2.$$

Когда сдвиг фаз составит $\Delta\omega \cdot t \sim \pi$, волны уже нельзя считать когерентными. Поэтому время когерентности можно определить условием

$$\Delta\varphi(t + t_{\text{ког}}) - \Delta\varphi(t) \sim \pi, \quad (4.1.2)$$

или

$$t_{\text{ког}} \sim \pi / \Delta\omega. \quad (4.1.3)$$

Переходя здесь от частоты к длине волны по формуле $\omega = 2\pi c / \lambda$ и считая $\Delta\omega \ll \omega$, получим

$$t_{\text{ког}} \sim \frac{\lambda^2}{2c\Delta\lambda}. \quad (4.1.4)$$

Путь, проходимый волнами за время когерентности, называется *длиной когерентности* и составляет

$$l_{\text{ког}} = ct_{\text{ког}} \sim \frac{\lambda^2}{2\Delta\lambda}. \quad (4.1.5)$$

На рис. 4.1.4 приведён график, иллюстрирующий сложение трёх волн с частотами ω_1 , $\omega_2 = \omega_1 + \Delta\omega/2$, $\omega_3 = \omega_1 + \Delta\omega$, $\Delta\omega = 0,3\omega_1$. Как видно, за время (4.1.3) в противофазе оказываются крайние волны спектра (с частотами ω_1 , ω_3), тогда как волна с частотой ω_2 дольше сохраняет когерент-

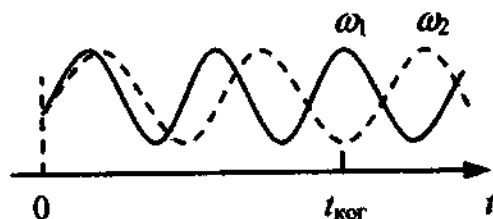


Рис. 4.1.3. Две волны с разными частотами могут считаться когерентными в течение времени $t_{\text{ког}}$, пока их разность фаз не превысит величины π

ность с другими. Соответственно время когерентности теперь более естественно определить соотношением

$$t_{\text{ког}} \sim \frac{\pi}{(\Delta\omega/2)} = \frac{2\pi}{\Delta\omega} \sim \frac{\lambda^2}{c\Delta\lambda}, \quad l_{\text{ког}} = ct_{\text{ког}} \sim \frac{\lambda^2}{\Delta\lambda},$$

отличающимся множителем «2» от (4.1.6). Более строгое обоснование данного утверждения приведено в следующем разделе.

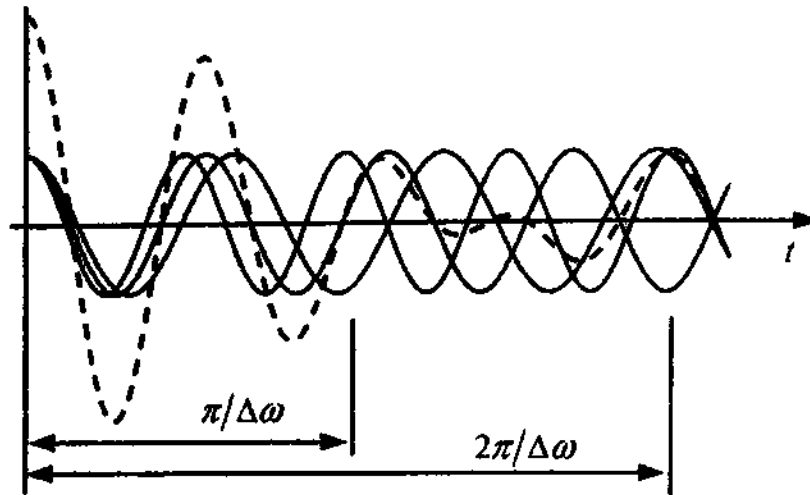


Рис. 4.1.4. Сложение трёх волн с частотами в диапазоне $\Delta\omega = 2\delta\omega$. Штриховой линией показан суммарный сигнал

Волновой пакет (пуг волн) состоит из волн с разными частотами. Соответственно с течением времени разность фаз компонент меняется. Это определяет время существования пакета: время когерентности компонент есть характерное время существования данного пуга. Как видно из примера с тремя волнами (рис. 4.1.4), формулы

$$t_{\text{ког}} \sim \frac{2\pi}{\Delta\omega} \sim \frac{\lambda^2}{c\Delta\lambda}, \quad l_{\text{ког}} = ct_{\text{ког}} \sim \frac{\lambda^2}{\Delta\lambda} \quad (4.1.6)$$

определяют время жизни волнового пакета с характерной шириной спектра $\Delta\omega$ (по частоте) и $\Delta\lambda$ (по длине волны).

Приведём типичные значения введённых величин.

Для видимого света характерная длительность отдельного пуга составляет $t_{\text{ког}} \sim 10^{-8} - 10^{-10}$ с, а длина когерентности — порядка нескольких сантиметров. Для лазерного излучения $t_{\text{ког}} \sim 10^{-4}$ с, а длина когерентности может достигать нескольких километров.

4.1.3. Связь времени когерентности с шириной спектра

Представим волновой пакет в виде суперпозиции монохроматических волн:

$$A(t) = \int_{-\infty}^{\infty} a(\omega) e^{-i\omega t} \frac{d\omega}{2\pi}, \quad a(\omega) = \int_{-\infty}^{\infty} A(t) e^{i\omega t} dt. \quad (4.1.7)$$

Пусть фурье-спектр сигнала даётся выражением

$$a(\omega) = \begin{cases} a_0, & |\omega - \omega_0| < \Delta\omega/2, \\ 0, & |\omega - \omega_0| > \Delta\omega/2. \end{cases} \quad (4.1.8)$$

Эта зависимость показана на рис. 4.1.5.

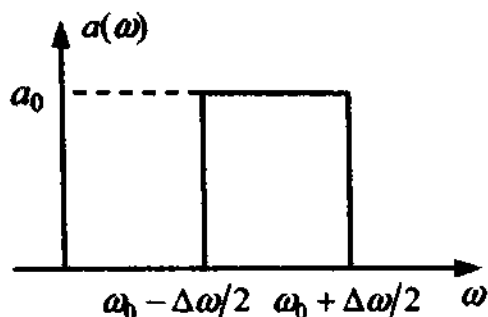


Рис. 4.1.5. Функция, описывающая фурье-спектр импульса шириной $\Delta\omega$

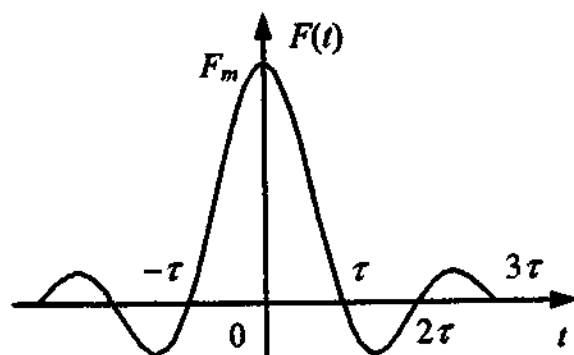


Рис. 4.1.6. Функция, описывающая импульс с фурье-спектром шириной $\Delta\omega$ и характерной длительностью $\tau \sim 2\pi/\Delta\omega$

Соответствующая временная зависимость сигнала определится по первой формуле в (4.1.7):

$$A(t) = \int_{-\infty}^{\infty} a(\omega) e^{-i\omega t} \frac{d\omega}{2\pi} = \frac{a_0}{2\pi} \int_{\omega_0 - \Delta\omega/2}^{\omega_0 + \Delta\omega/2} e^{-i\omega t} d\omega = a_0 e^{-i\omega_0 t} F(t). \quad (4.1.9)$$

Здесь введена функция

$$F(t) = \frac{1}{2\pi i t} \left(e^{i\Delta\omega t/2} - e^{-i\Delta\omega t/2} \right) = F_m \frac{\sin(\Delta\omega \cdot t/2)}{(\Delta\omega \cdot t/2)}, \quad (4.1.10)$$

где $F_m = \Delta\omega/2\pi$. График $F(t)$ показан на рис. 4.1.6.

Функция $F(t)$ обращается в первый раз в нуль при условии $\Delta\omega \cdot t/2 = \pi$, то есть в момент времени

$$t = \tau = 2\pi/\Delta\omega. \quad (4.1.11)$$

Эта величина даёт характерное время существования волнового пакета — время когерентности.

4.1.4. Соотношение неопределённостей

Поскольку ширина фурье-спектра сигнала (рис. 4.1.5 и 4.1.6) равна $\Delta\omega$, а характерная длительность есть τ , то из (4.1.11) получаем соотношение неопределённостей

$$\tau \cdot \Delta\omega = 2\pi. \quad (4.1.12)$$

Данное соотношение носит общий характер и означает следующее: произведение характерного времени существования волнового пакета на ширину соответствующего фурье-спектра есть величина порядка 2π

Обратно, если характерное время существования волнового пакета есть τ , то ширина его фурье-спектра оказывается порядка

$$\Delta\omega \sim 2\pi/\tau. \quad (4.1.13)$$

Другими словами, чем короче сигнал, тем шире его фурье-спектр.

4.1.5. Влияние немонохроматичности на наблюдаемое число интерференционных полос

Сигнал называется квазимонохроматическим, если его можно представить в виде

$$A(t) = a(t) \cos(\omega_0 t + \varphi(t)), \quad (4.1.14)$$

где амплитуда $a(t)$ и фаза $\varphi(t)$ — медленно меняющиеся функции. Последнее означает, что изменения этих функций малы за характерное время основных колебаний $T = 2\pi/\omega_0$. Квазимонохроматический сигнал можно рассматривать как гармоническое колебание с несущей частотой ω_0 , модулированное по амплитуде и частоте.

Если источник света квазимонохроматический, то интерференционную картину от него можно, тем не менее, получить. Действительно, свет, испускаемый таким источником, представляет собой суперпозицию монохроматических волн. Если каждую из этих волн расщепить на две волны, то получаемая в результате пара когерентных волн уже создаёт интерференционную картину. На другом языке можно сказать, что в рассматриваемой схеме осуществляется интерференция каждого пуга с самим собой (раздел 4.1.3). Расщепление одной волны на две можно осуществить, например, в схеме Юнга (рис. 4.1.7), в которой используется экран с двумя щелями: S_1 и S_2 .

Интерференционная картина от всего спектра получается наложением интерференционных картин от отдельных компонент спектра. Однако, поскольку период отдельных картин зависит от соответствующей длины волны, положение максимумов и минимумов различно для разных компонент спектра. Это ограничивает общее число наблюдаемых полос от источника S .

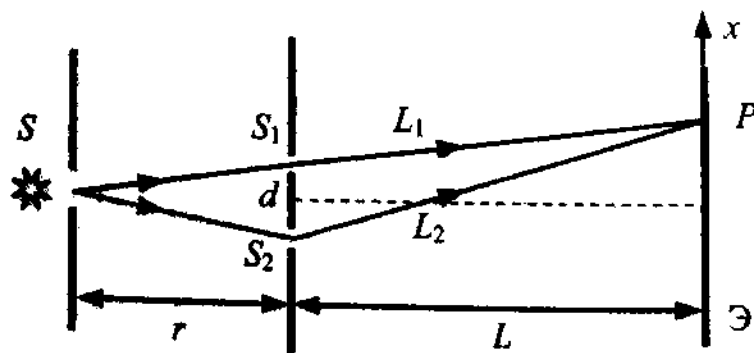


Рис. 4.1.7. Схема Юнга по наблюдению интерференционной картины от квазимонохроматического источника S . Лучи от источника проходят через два отверстия S_1 и S_2 , а затем интерферируют на экране $\mathcal{E}$

Рассмотрим компоненту спектра, имеющую длину волны λ . Как было показано в разделе 3.4.1, разность хода лучей в опыте Юнга составляет

$$\delta = L_2 - L_1 \approx xd/L. \quad (4.1.15)$$

Максимумы интерференционной картины наблюдаются в точках, для которых $\delta = m\lambda$, $m = 0, \pm 1, \pm 2, \dots$, или

$$x_m^{(\max)} = \frac{\lambda L}{d} m. \quad (4.1.16)$$

Минимумы интерференционной картины наблюдаются в точках, для которых $\delta = \left(m + \frac{1}{2}\right)\lambda$, $m = 0, \pm 1, \pm 2, \dots$, или

$$x_m^{(\min)} = \frac{\lambda L}{d} \left(m + \frac{1}{2}\right). \quad (4.1.17)$$

Ширина интерференционной полосы и расстояние между соседними полосами равно

$$\Delta x = x_{m+1}^{(\min)} - x_m^{(\min)} = x_{m+1}^{(\max)} - x_m^{(\max)} = \frac{\lambda L}{d}. \quad (4.1.18)$$

Учтём теперь, что источник S создаёт немонохроматический свет в спектральном диапазоне $\Delta\lambda$. Для такого света длина когерентности есть величина порядка

$$l_{\text{ког}} \sim \lambda^2 / \Delta\lambda.$$

Максимальная разность хода лучей δ , при которой они ещё могут считаться когерентными, не должна превышать $l_{\text{ког}}$:

$$\delta < l_{\text{ког}}.$$

Отсюда находим наибольший порядок интерференции:

$$\delta_{\max} = m_{\max} \lambda \sim \frac{\lambda^2}{\Delta\lambda} \Rightarrow m_{\max} \sim \frac{\lambda}{\Delta\lambda}. \quad (4.1.19)$$

Соответственно максимально число наблюдаемых полос интерференции составляет

$$2m_{\max} + 1 \sim \frac{2\lambda}{\Delta\lambda} \quad (4.1.20)$$

(предполагается $\Delta\lambda \ll \lambda$).

4.1.6. Видность интерференционной картины

Введём спектральную плотность интенсивности J_ω , задающую распределение интенсивности по спектру:

$$I_0 = \int dI, \quad dI = J_\omega d\omega. \quad (4.1.21)$$

Рассмотрим квазимонохроматический сигнал с шириной спектра $\Delta\omega$ и равномерным распределением интенсивности в этой полосе:

$$J_{\omega} = \begin{cases} I_0/\Delta\omega, & |\omega - \omega_0| < \Delta\omega/2, \\ 0, & |\omega - \omega_0| > \Delta\omega/2. \end{cases} \quad (4.1.22)$$

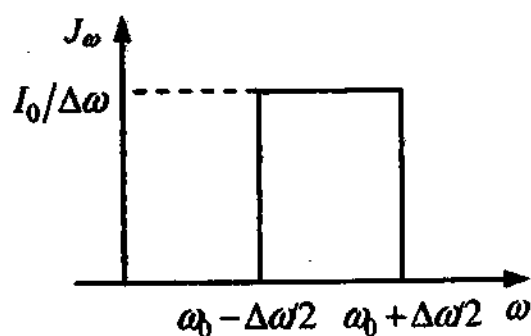


Рис. 4.1.8. Функция, описывающая импульс с шириной фурье-спектра $\Delta\omega$

Здесь I_0 — суммарная (по всему спектру) интенсивность сигнала. Эта зависимость показана на рис. 4.1.8. Поскольку различные фурье-компоненты имеют разные частоты, то они не когерентны между собой. Соответственно интерференционная картина от рассматриваемого источника получается наложением интерференционных картин от каждой из компонент.

Выберем компоненту спектра с частотой ω и интенсивностью $J_{\omega}d\omega$. Создаваемое ею распределение интенсивности на экране Э (рис. 4.1.6) даётся выражением

$$dI = 2(J_{\omega}d\omega)(1 + \cos k\delta), \quad (4.1.23)$$

где $k = \omega/c$, а δ — оптическая разность хода лучей, идущих от щелей S_1 и S_2 до точки наблюдения.

Вследствие некогерентности отдельных спектральных компонент интенсивность интерференционной картины равна сумме интенсивностей, создаваемых отдельными гармониками:

$$\begin{aligned} I(\delta) &= \int dI_{\text{интерф}} = 2 \int_{\omega_0 - \Delta\omega/2}^{\omega_0 + \Delta\omega/2} (J_{\omega}d\omega)(1 + \cos k\delta) = \\ &= \frac{2I_0}{\Delta\omega} \int_{\omega_0 - \Delta\omega/2}^{\omega_0 + \Delta\omega/2} \left[1 + \cos\left(\frac{\omega}{c}\delta\right) \right] d\omega = 2I_0 \left\{ 1 + \frac{\sin(\delta \cdot \Delta\omega/2c)}{(\delta \cdot \Delta\omega/2c)} \cos\left(\frac{\omega_0}{c}\delta\right) \right\}. \end{aligned} \quad (4.1.24)$$

В зависимости от разности хода лучей δ будут чередоваться максимумы и минимумы интерференционной картины $I_{\text{интерф}}(\delta)$. Поскольку предполагается, что $\Delta\omega \ll \omega_0$, то множитель $\cos(\omega_0\delta/c)$ при изменении δ меняется много быстрее, чем коэффициент при нём. Поэтому именно им приближённо определяется положение локальных максимумов и минимумов интенсивности:

$$I(\delta) = I_{\max} \text{ при } \cos\left(\frac{\omega_0}{c} \delta\right) = +1 \Rightarrow I_{\max} = 2I_0(1 + \varepsilon),$$

$$I(\delta) = I_{\min} \text{ при } \cos\left(\frac{\omega_0}{c} \delta\right) = -1 \Rightarrow I_{\min} = 2I_0(1 - \varepsilon).$$
(4.1.25)

Здесь

$$\varepsilon = \frac{\sin(\delta \cdot \Delta\omega/2c)}{(\delta \cdot \Delta\omega/2c)}.$$
(4.1.26)

Предполагалось, что $\varepsilon > 0$. Если же окажется, что $\varepsilon < 0$, то максимумы и минимумы поменяются местами. С учётом этого находим видность интерференционной картины:

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = |\varepsilon| = \left| \frac{\sin(\delta \cdot \Delta\omega/2c)}{(\delta \cdot \Delta\omega/2c)} \right| \equiv V(\delta).$$
(4.1.27)

Поскольку $\delta = xd/L$, то эта функция определяет относительную интенсивность интерференционных полос на экране (координата x отсчитывается вдоль экрана). График функции $V(\delta)$ показан на рис. 4.1.9.

В первый раз видность обращается в нуль: $V(\delta) = 0$ при

$$\frac{\delta \cdot \Delta\omega}{2c} = \pi,$$

или, с учётом равенств $\Delta\omega = 2\pi c \Delta\lambda / \lambda_0^2$, $\lambda_0 = 2\pi c / \omega_0$,

$$\delta = \frac{2\pi c}{\Delta\omega} = \frac{\lambda_0^2}{\Delta\lambda} \equiv \delta_{\max}.$$
(4.1.28)

Таким образом, интерференционная картина пропадает, когда разность хода лучей оказывается сравнимой с длиной когерентности. В области $\delta > \delta_{\max}$ видность оказывается, хотя и ненулевой, но малой, так что наиболее существенная часть интерференционной картины сосредоточена в области $\delta < \delta_{\max}$.

Максимальное число наблюдаемых полос равно

$$N_{\max} = 2 \frac{\delta_{\max}}{\lambda_0} = 2 \frac{2\pi c / \Delta\omega}{2\pi c / \omega_0} =$$
(4.1.29)

$$= 2 \frac{\omega_0}{\Delta\omega} = \frac{2\lambda_0}{\Delta\lambda}$$

(коэффициент «2» связан с тем, что

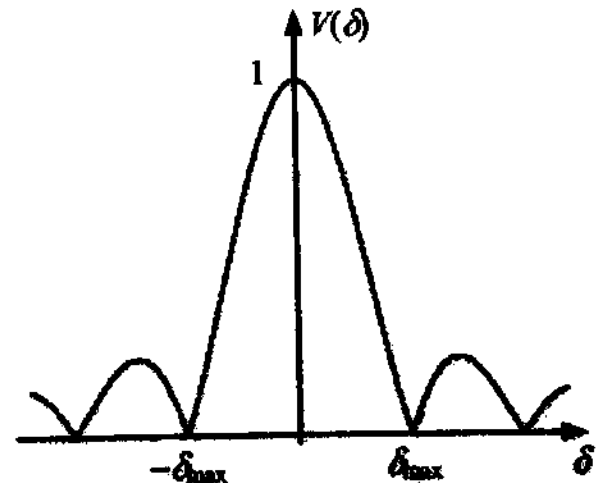


Рис. 4.1.9. Функция видности интерференционной картины при использовании квазимонохроматического источника

интерференционные полосы расположены симметрично в областях $\delta < 0$ и $\delta > 0$).

Мы пришли к тому же результату, который был ранее найден на основании оценок.

4.2. Функция временной когерентности

Определим, как и в главе 3, интенсивность волны соотношением

$$I = \overline{A^2(t)} = \frac{1}{\Delta t} \int_t^{t+\Delta t} A^2(t_1) dt_1, \quad (4.2.1)$$

где усреднение проводится по промежутку времени $\Delta t \gg T_0 = 2\pi/\omega_0$, а T_0 — характерный период колебаний волны. Время усреднения определяется инерционностью используемого прибора. Например, глаз позволяет разрешать временные промежутки ~ 0.05 с. Высокочувствительные фотоприёмники могут разрешать сигналы, следующие с интервалом $\sim 10^{-5} - 10^{-6}$ с. Эти промежутки, как правило, велики по сравнению с длительностью отдельных цугов волн, порождаемых источником излучения: $\Delta t \gg \tau_0 \sim 10^{-8} - 10^{-10}$ с.

Контраст интерференционной картины можно характеризовать видностью. Если складываются когерентные волны от двух источников, то интенсивность суммарной волны равна

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos \Delta\varphi, \quad (4.2.2)$$

где $\Delta\varphi$ — разность фаз складываемых волн в рассматриваемой точке.

Если интенсивности интерферирующих волн одинаковы: $I_1 = I_2 = I_0$, то

$$I = 2I_0(1 + \cos \Delta\varphi). \quad (4.2.3)$$

В зависимости от разности хода волн до точки наблюдения разность фаз принимает разные значения, так что интенсивность меняется в диапазоне

$$I_{\min} \leq I \leq I_{\max}, \quad I_{\min} = 0, \quad I_{\max} = 4I_0. \quad (4.2.4)$$

Видность интерференционной картины при этом составит

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = 1. \quad (4.2.5)$$

Рассмотрим интерференцию волн от одного источника, реализуемую, например, в схеме Юнга. Будем считать, что расщепляемая волна является квазимонохроматической:

$$A(t) = a(t) \cos[\omega_0 t + \varphi(t)], \quad (4.2.6)$$

представляющей собой случайную последовательность близких по форме отдельных цугов. Считаем, что характерная длительность отдельного цуга (импульса) равна τ_0 , причём $\tau_0 \gg T_0 = 2\pi/\omega_0$.

Если до точки наблюдения на экране волны доходят с задержкой τ , то интенсивность суммарной волны оказывается равной

$$I(t, \tau) = \overline{[A(t) + A(t + \tau)]^2} = \overline{A^2(t)} + \overline{A^2(t + \tau)} + 2\overline{A(t)A(t + \tau)}. \quad (4.2.7)$$

Считаем, что процесс *стационарный*. Это значит, что суммарная интенсивность $I(t, \tau)$ не зависит от момента наблюдения t , но определяется только задержкой τ во времени прибытия волн к точке наблюдения:

$$\begin{aligned} \overline{A^2(t)} = \overline{A^2(t + \tau)} &= I_0, \quad \overline{A(t)A(t + \tau)} = \Gamma(\tau), \\ I &= 2I_0 + 2\Gamma(\tau). \end{aligned} \quad (4.2.8)$$

Здесь введена величина $\Gamma(\tau)$, называемая *функцией когерентности*. Она имеет очевидное свойство:

$$\Gamma(0) = I_0. \quad (4.2.9)$$

Для дальнейшего воспользуемся комплексным представлением колебаний:

$$A(t) \rightarrow \hat{S}(t) = \hat{S}_0(t)e^{ia_0 t}, \quad \hat{S}_0(t) = a(t)e^{i\varphi(t)}. \quad (4.2.10)$$

В этих обозначениях имеем

$$\begin{aligned} I(\tau) &= \frac{1}{2} \overline{|\hat{S}(t) + \hat{S}(t + \tau)|^2} = \frac{1}{2} \left[\overline{|\hat{S}(t)|^2} + \overline{|\hat{S}(t + \tau)|^2} + 2 \operatorname{Re} \overline{\hat{S}^*(t) \hat{S}(t + \tau)} \right] = \\ &= 2I_0 + 2 \operatorname{Re} \hat{\Gamma}(\tau). \end{aligned} \quad (4.2.11)$$

Здесь использованы обозначения

$$I_0 = \frac{1}{2} \overline{|\hat{S}(t)|^2} = \frac{1}{2} \overline{a^2(t)}, \quad \hat{\Gamma}(\tau) = \frac{1}{2} \overline{\hat{S}^*(t) \hat{S}(t + \tau)}, \quad (4.2.12)$$

причём функция когерентности связана с введённой здесь комплексной величиной $\hat{\Gamma}(\tau)$ соотношением

$$\Gamma(\tau) = \operatorname{Re} \hat{\Gamma}(\tau). \quad (4.2.13)$$

Вспоминая определение (4.2.10) величины $\hat{S}(t)$, получаем выражение для комплексной функции когерентности:

$$\hat{\Gamma}(\tau) = \frac{1}{2} \overline{\hat{S}^*(t) \hat{S}(t + \tau)} = \frac{1}{2} \overline{a(t)a(t + \tau)e^{i[\varphi(t + \tau) - \varphi(t)]}} e^{ia_0 \tau}. \quad (4.2.14)$$

Выделяются два предельных случая.

1. Задержка τ между прибытиями двух компонент расщеплённой волны к точке наблюдения мала по сравнению с длительностью цуга: $\tau \ll \tau_0$.

Тогда

$$a(t) \approx a(t + \tau), \quad \varphi(t) \approx \varphi(t + \tau). \quad (4.2.15)$$

В этих условиях

$$\hat{\Gamma}(\tau) = I_0 e^{i\omega_0 \tau}. \quad (4.2.16)$$

Согласно (4.2.9) интенсивность интерференционной картины оказывается равной

$$I = 2I_0 (1 + \cos \omega_0 \tau), \quad (4.2.17)$$

а её видность

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = 1. \quad (4.2.18)$$

2. Задержка по времени прибытия волн к точке наблюдения велика по сравнению с длительностью цуга: $\tau \gg \tau_0$. В этом случае фазы интерферирующих волн могут принимать любые значения, так что

$$\overline{\exp\{i[\varphi(t + \tau) - \varphi(t)]\}} = \overline{\cos[\varphi(t + \tau) - \varphi(t)] + i \sin[\varphi(t + \tau) - \varphi(t)]} = 0.$$

Поэтому

$$\hat{\Gamma}(\tau) \Big|_{\tau \gg \tau_0} \rightarrow 0. \quad (4.2.19)$$

Видность при этом обращается в нуль: $V \Big|_{\tau \gg \tau_0} \rightarrow 0$.

Как следует из сказанного, видность интерференционной картины уменьшается благодаря дополнительному сдвигу фаз интерферирующих волн: помимо того сдвига, который связан с разностью хода волн до точки наблюдения, имеет место сдвиг фаз спектральных компонент, обусловленный немонахроматичностью излучения.

Рассмотрим теперь промежуточные значения времени задержки τ . Обозначим

$$\hat{\Gamma}(\tau) = I_0 \hat{\gamma}(\tau). \quad (4.2.20)$$

Введённая этим соотношением функция $\gamma(\tau)$ называется *степенью временной когерентности* интерферирующих волн.

Положим

$$\overline{e^{i[\varphi(t + \tau) - \varphi(t)]}} = e^{i\varphi_0(\tau)}. \quad (4.2.21)$$

Тогда для степени когерентности получаем выражение

$$\hat{\gamma}(\tau) = |\gamma(\tau)| e^{i[\omega_0 \tau + \varphi_0(\tau)]}. \quad (4.2.22)$$

С учётом сказанного перепишем формулу для интенсивности интерференционной картины (4.2.11):

$$I(\tau) = 2I_0 [1 + |\gamma(\tau)| \cos(\omega_0 \tau + \varphi_0)]. \quad (4.2.23)$$

Поскольку $|\gamma(\tau)|$ — медленно меняющаяся функция по сравнению с $\cos(\omega_0 \tau + \varphi_0)$, то максимальное и минимальное значения интенсивности оказываются равными

$$I_{\min} = 2I_0 (1 - |\gamma|), \quad I_{\max} = 2I_0 (1 + |\gamma|),$$

а видность даёт выражением

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = |\gamma(\tau)|. \quad (4.2.24)$$

Таким образом, видность совпадает со степенью когерентности излучения.

Как следует из проведённого анализа, интерференция наблюдается (видность $V \sim 1$), если время задержки волн не превышает характерной длительности цуга: $\tau < \tau_0$. Поскольку величина $c\tau_0 = l_{\text{ког}}$ — это длина когерентности, а $\delta = c\tau$ — разность хода волн, приходим к полученному ранее результату: условием наблюдения интерференции является требование, чтобы разность хода волн не превышала длины когерентности: $\delta < l_{\text{ког}}$.

Пример 1. В случае излучения с равномерным распределением энергии по спектру в диапазоне $\Delta\omega$ (рис. 4.1.7) степень когерентности и видность интерференционной картины даются формулой (4.1.27), проиллюстрированной на рис. 4.1.8.

Пример 2. Рассмотрим модельную задачу, когда цуг представляет собой синусоидальный сигнал конечной длительности (рис. 4.2.1): в течение времени τ_0 $A(t) = a \cos(\omega_0 t)$, а за пределами этого промежутка $A(t) = 0$.

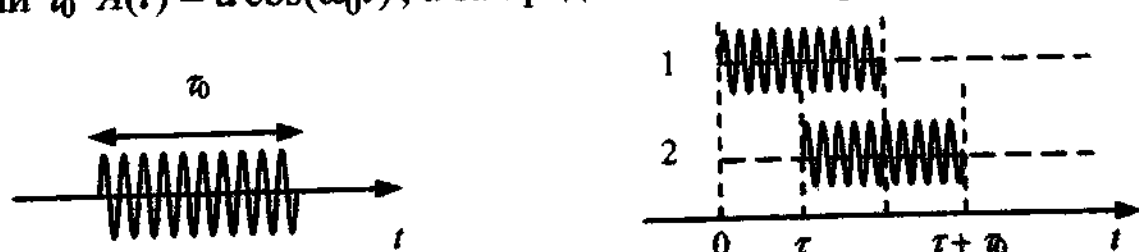


Рис. 4.2.1. Слева — квазигармонический цуг, справа — интерференция волн, полученных расщеплением цуга и доходящих до точки наблюдения с временной задержкой τ

Пусть этот цуг расщепляется на два (например, в схеме Юнга), которые интерферируют друг с другом, а до точки наблюдения образовавшиеся цуги доходят с задержкой τ . Найдём функцию когерентности и видность интерференционной картины. Первая даёт выражением

$$\Gamma(\tau) = \frac{1}{\Delta t} \int_0^{\Delta t} A(t_1) A(t_1 + \tau) dt. \quad (4.2.25)$$

Момент начала цуга принят равным $t = 0$, а время усреднения выбрано много большим, чем период колебаний волны в цуге. Для расчётов положим $\Delta t = \tau_0$.

Там, где цуги не перекрываются, интерференция отсутствует. Интенсивность волны в отдельном цуге равна

$$I_0 = \overline{A^2(t)} = \frac{1}{\tau_0} \int_0^{\tau_0} A^2(t) dt = \frac{1}{2} a^2. \quad (4.2.26)$$

Найдём функцию когерентности:

$$\begin{aligned} \Gamma(\tau) &= \frac{1}{\tau_0} \int_0^{\tau_0} A(t) A(t + \tau) dt = \frac{1}{\tau_0} \int_{\tau}^{\tau_0} A(t) A(t + \tau) dt = \\ &= \frac{1}{\tau_0} a^2 \int_{\tau}^{\tau_0} \sin(\omega_0 t) \sin[\omega_0(t + \tau)] dt. \end{aligned} \quad (4.2.27)$$

Имея в виду тождество

$$\sin(\omega_0 t) \sin[\omega_0(t + \tau)] = \frac{1}{2} [\cos(\omega_0 \tau) - \cos[\omega_0(2t + \tau)]]$$

и учитывая, что при усреднении по промежутку $\tau_0 \gg 2\pi/\omega_0$ можно положить $\overline{\cos[\omega_0(2t + \tau)]} = 0$, получаем

$$\Gamma(\tau) = \frac{1}{2\tau_0} a^2 (\tau_0 - \tau) \cos(\omega_0 \tau) = I_0 \left(1 - \frac{\tau}{\tau_0}\right) \cos(\omega_0 \tau).$$

Поскольку $\Gamma(\tau) = I_0 |\gamma(\tau)| \cos(\omega_0 \tau)$, то приходим к следующему выражению для степени когерентности волны:

$$|\gamma(\tau)| = \begin{cases} 1 - \tau/\tau_0, & \tau < \tau_0, \\ 0, & \tau > \tau_0. \end{cases} \quad (4.2.28)$$

Видность интерференционной картины согласно (4.2.24) совпадает со степенью когерентности: $V = |\gamma(\tau)|$. Зависимость видности от разности хода лучей δ вытекает из (4.2.28): для $\delta \leq l_{\text{ког}}$ находим

$$\delta = c\tau, \quad l_{\text{ког}} = c\tau_0 \Rightarrow V(\delta) = 1 - \delta/l_{\text{ког}}. \quad (4.2.29)$$

4.3. Теорема Винера–Хинчина

Рассмотрим интерференционную картину, производимую одним источником, например, в схеме Юнга. Поскольку реальные источники немо-

нохроматические, то вследствие усреднения, осуществляемого измерительным прибором, волны с разными частотами интерферировать между собой не будут — складываются интенсивности отдельных гармоник. Распределение интенсивности по спектру будем описывать с помощью спектральной плотности $J(\omega)$:

$$dI_0 = J(\omega)d\omega, \quad I_0 = \int_0^{\infty} J(\omega)d\omega. \quad (4.3.1)$$

По смыслу функции $J(\omega)$ выполняется неравенство $J(\omega) \geq 0$.

Пусть интенсивности складываемых волн в точке наблюдения одинаковы. Тогда при наложении волн, идущих по разным путям с временной задержкой τ , вклад в интенсивность интерференционной картины от участка спектра $d\omega$ равен

$$dI = 2dI_0(1 + \cos \omega\tau) = 2J(\omega)d\omega(1 + \cos \omega\tau). \quad (4.3.2)$$

Суммируя вклады от всех участков спектра, находим

$$I = 2 \int_0^{\infty} J(\omega)(1 + \cos \omega\tau) d\omega. \quad (4.3.3)$$

Вводя функцию когерентности согласно (4.2.8) и учитывая равенство (4.3.1), получаем

$$I = 2I_0 + 2\Gamma(\tau). \quad (4.3.4)$$

Последние два равенства устанавливают связь функции когерентности со спектральной плотностью интенсивности излучения:

$$\Gamma(\tau) = \int_0^{\infty} J(\omega) \cos \omega\tau d\omega. \quad (4.3.5)$$

Этому соотношению можно придать иной вид, перейдя к комплексным обозначениям:

$$\hat{\Gamma}(\tau) = \int_{-\infty}^{\infty} J(\omega) e^{i\omega\tau} d\omega, \quad \Gamma(\tau) = \text{Re} \hat{\Gamma}(\tau). \quad (4.3.6)$$

Мы распространили интегрирование по частоте до $-\infty$, имея в виду, что $J(\omega)|_{\omega < 0} = 0$.

Таким образом, функция когерентности есть (обратное) преобразование Фурье от спектральной плотности интенсивности. Данный вывод представляет собой содержание *теоремы Винера–Хинчина*. Она была получена Н. Винером (США) в 1930 г. и (в иной формулировке) А. Я. Хинчиным (СССР) в 1934 г.

Пример 1. Рассмотрим излучение с равномерным спектром интенсивности в диапазоне $\Delta\omega$ в окрестности центральной частоты ω_0 (рис. 4.3.1, слева). Вычисление по формуле (4.3.5) даёт

$$\Gamma(\tau) = \int_0^{\infty} J(\omega) \cos \omega \tau d\omega = \frac{I_0}{\Delta\omega} \int_{\omega_0 - \Delta\omega/2}^{\omega_0 + \Delta\omega/2} \cos \omega \tau d\omega = I_0 \frac{\sin(\tau \Delta\omega/2)}{(\tau \Delta\omega/2)} \cos \omega_0 \tau.$$

На рис. 4.3.1 (справа) показана найденная таким образом функция когерентности. Как видно, характерное время когерентности есть $\tau_{\max} = 2\pi/\Delta\omega = \lambda^2/\Delta\lambda$.

Пример 2. Рассмотрим излучение со следующим спектральным распределением интенсивности:

$$J(\omega) = \frac{1}{\pi} I_0 \frac{1}{(\omega - \omega_0)^2 + \varepsilon^2}. \quad (4.3.7)$$

Здесь предполагается, что спектральная линия узкая: $\varepsilon \ll \omega_0$. При этом допущении выполняется условие нормировки

$$\int_0^{\infty} J(\omega) d\omega = \frac{1}{2} I_0 \left(1 + \frac{2}{\pi} \operatorname{arctg} \frac{\omega_0}{\varepsilon} \right) \approx I_0. \quad (4.3.8)$$

Функция когерентности согласно (4.3.5) равна

$$\Gamma(\tau) = \int_0^{\infty} J(\omega) \cos \omega \tau d\omega = I_0 e^{-\varepsilon \tau} \cos(\omega_0 \tau) \quad (4.3.9)$$

(при условии $\varepsilon \ll \omega_0$ интегрирование по ω можно распространить от $-\infty$ до $+\infty$).

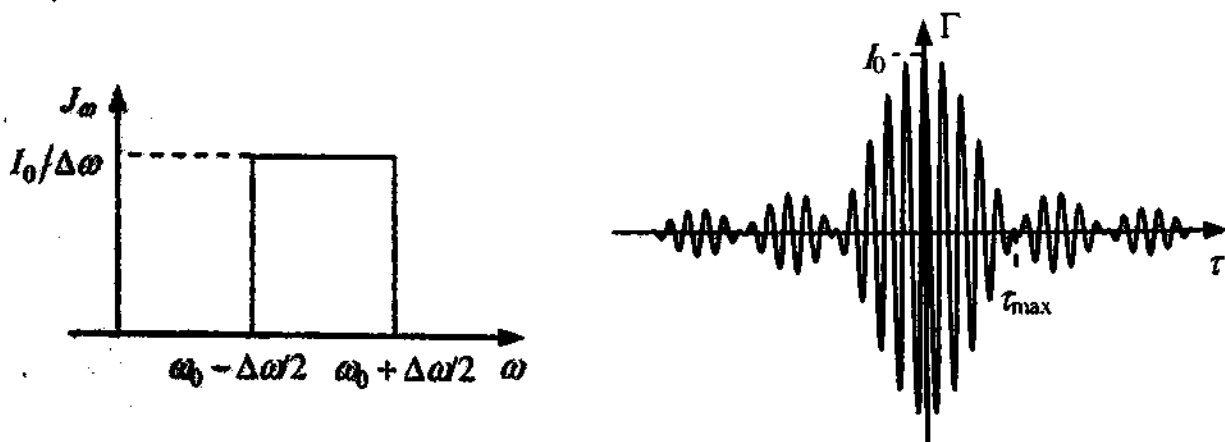


Рис. 4.3.1. Слева — равномерное распределение интенсивности излучения по спектру шириной $\Delta\omega$; справа — соответствующая функция когерентности для случая $\Delta\omega/\omega_0 = 0,1$

Из (4.3.9) находим степень когерентности и функцию видности:

$$V(\tau) = |\gamma(\tau)| = e^{-\varepsilon \tau}. \quad (4.3.10)$$

Из полученного выражения видно, что время когерентности для рассматриваемого сигнала равно $t_{\text{ког}} = 1/\varepsilon$, так что последнюю формулу можно переписать в виде

$$V(\tau) = \exp(-\tau/t_{\text{ког}}). \quad (4.3.11)$$

Данная формула аналогична соотношению (4.2.8), но точнее отражает реальную спектральную плотность интенсивности излучения, испускаемого атомом.

Заметим, что время когерентности $t_{\text{ког}}$ обратно пропорционально ширине спектра ε : $\varepsilon \cdot t_{\text{ког}} = 1$. Это соотношение представляет собой следствие соотношения неопределённостей (4.1.12) для рассмотренной задачи.

Пример 3. Найдём функцию когерентности и видность интерференционной картины, когда спектр излучения содержит две узкие спектральные линии с равной интенсивностью:

$$J(\omega) = I_0 [\delta(\omega - \omega_1) + \delta(\omega - \omega_2)] \quad (4.3.12)$$

и близкими частотами: $\Delta\omega = \omega_2 - \omega_1 \ll \omega_1$.

Для решения задачи воспользуемся формулой (4.3.6):

$$\hat{\Gamma}(\tau) = \int_{-\infty}^{\infty} J(\omega) e^{i\omega\tau} d\omega = I_0 (e^{i\omega_1\tau} + e^{i\omega_2\tau}). \quad (4.3.13)$$

Положим

$$\begin{aligned} \omega_0 &= (\omega_1 + \omega_2)/2, \quad \Delta\omega = \omega_2 - \omega_1, \\ \omega_1 &= \omega_0 - \Delta\omega/2, \quad \omega_2 = \omega_0 + \Delta\omega/2. \end{aligned} \quad (4.3.14)$$

Имея в виду, что суммарная интенсивность излучения равна $2I_0$, находим степень когерентности:

$$\gamma(\tau) = \frac{\Gamma(\tau)}{2I_0} = \frac{1}{2} (e^{-i\Delta\omega\tau/2} + e^{i\Delta\omega\tau/2}) e^{i\omega_0\tau} = \cos\left(\frac{\Delta\omega\tau}{2}\right) e^{i\omega_0\tau}. \quad (4.3.15)$$

Наконец, для функции видности с учётом условия $\Delta\omega \ll \omega_0$ получаем выражение

$$V(\tau) = |\gamma(\tau)| = \left| \cos\left(\frac{\Delta\omega\tau}{2}\right) \right|. \quad (4.3.16)$$

Отсюда следует, что видность обращается в нуль при

$$\frac{\Delta\omega\tau}{2} = \frac{\pi}{2} \Rightarrow \tau = \frac{\pi}{\Delta\omega}. \quad (4.3.17)$$

Эта величина совпадает с временем когерентности (4.1.3), найденным в разделе 4.1.2 для случая спектра, содержащего как раз две частоты (формула (4.1.3)).

ГЛАВА 5

ПРОСТРАНСТВЕННАЯ КОГЕРЕНТНОСТЬ

5.1. Радиус пространственной когерентности

Пусть в опыте Юнга (рис. 5.1.1) используется протяженный источник S , создающий излучение с длиной волны λ .

Рассмотрим сначала случай, когда источник S состоит из двух точечных источников. Эти источники называются *пространственно когерентными*, если их размеры и взаимное расположение позволяют наблюдать интерференционные полосы (например, в схеме Юнга). Если же полосы нельзя получить при любой степени монохроматичности, то источники называются *пространственно некогерентными*. Введённые определения обобщаются на случай, когда источник представляет собой протяжённое светящееся тело, если рассматривать тело как совокупность точечных источников. Иными словами, протяжённый источник будем называть *пространственно когерентным*, если он может создавать интерференционную картину в схеме Юнга или эквивалентных ей схемах.

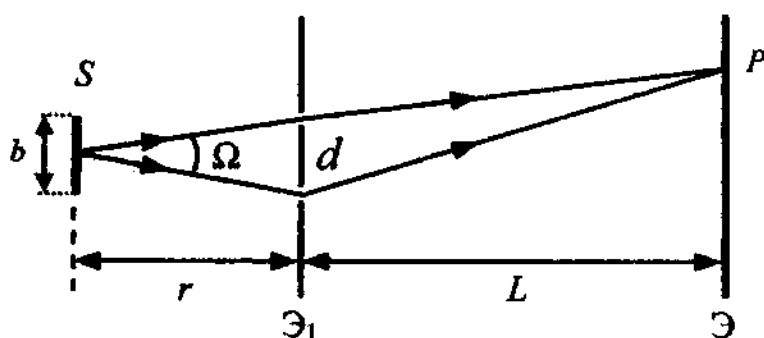


Рис. 5.1.1. Схема Юнга по наблюдению интерференционной картины от протяжённого источника. Лучи от источника S проходят через два отверстия в экране $\mathcal{E}_1$ и затем интерферируют на экране $\mathcal{E}$

Разные точки источника создают некогерентное излучение. Поэтому на экране $\mathcal{E}$ будут складываться не амплитуды волн, создаваемых разными точками, а интенсивности.

Каждая точка источника создаёт в схеме Юнга свою интерференционную картину. И все картины накладываются одна на другую, но смещены одна относительно другой (рис. 5.1.2). Если смещение невелико, то на экране $\mathcal{E}$ наблюдается такая же система интерференционных полос, как от точечного источника. В этом случае протяжённый источник S является пространственно когерентным.

Пусть средняя точка O источника S создаёт на экране $\mathcal{E}$ интерференционную картину с центром в точке P , а крайние точки O' и O'' — картины с центрами соответственно в точках P' и P'' . Выберем начало координат на экране наблюдения $\mathcal{E}$ в точке P . Если угол схождения лучей от источника к экрану $\mathcal{E}_1$ равен φ , то, как видно из рис. 5.1.2, координаты центров интерференционных картин от крайних точек источника равны

$$x' = L \frac{\varphi}{2}, \quad x'' = -L \frac{\varphi}{2}.$$

Расстояние между точками P' и P'' равно

$$\delta x = x' - x'' = L\varphi.$$

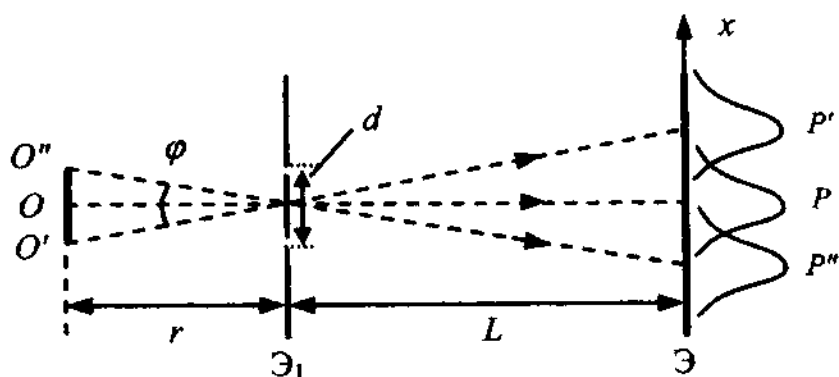


Рис. 5.1.2. Наложение интерференционных полос от разных точек протяжённого источника в схеме Юнга

Как показано в разделе 3.4.1, ширина полосы в опыте Юнга с точечным источником равна $\Delta x = \lambda L/d$, где L — расстояние от экрана с щелями $\mathcal{E}_1$ до экрана наблюдения $\mathcal{E}$, а d — расстояние между щелями. Поэтому если сдвиг всех полос меньше ширины одной полосы, то будет наблюдаться интерференционная картина — источник является пространственно когерентным. Для этого необходимо, чтобы

$$\delta x < \Delta x, \text{ или } L\varphi < \frac{\lambda L}{d}.$$

Полученное неравенство можно переписать в двух формах:

$$\varphi < \frac{\lambda}{d}, \text{ или } d < \frac{\lambda}{\varphi}.$$

Первое из неравенств указывает предельный угловой размер источника φ , от которого можно получить интерференционную картину в схеме с фиксированным размером базы d . Второе неравенство даёт максимальное значение расстояния d между щелями в интерференционной схеме, необходимое для получения интерференционной картины от источника заданного углового размера φ .

В качестве примера укажем, что Солнце имеет угловой диаметр $\varphi \approx 30'' \approx 0,01$. При длине волны излучения $\lambda = 500$ нм для получения интерференционной картины необходимо использовать экран с расстоянием между щелями

$$d < \frac{5 \cdot 10^{-5}}{0,01} \text{ см} = 0,05 \text{ мм}.$$

Поскольку угловой диаметр φ , линейный размер b источника и расстояние r от него до используемого оптического прибора (экрана с щелями) связаны соотношением $\varphi = b/r$, то условие пространственной когерентности источника можно представить в виде

$$\frac{b}{r} < \frac{\lambda}{d} \quad \text{или} \quad \frac{d}{r} < \frac{\lambda}{b}.$$

Угол $\Omega = d/r$, под которым с места расположения источника видно расстояние между щелями на экране $\mathcal{E}_1$, называется *апертурой интерференционной схемы*. Таким образом, условие пространственной когерентности источника S состоит в том, что эта апертура должна быть достаточно малой:

$$\Omega < \lambda/b.$$

Величина

$$\rho_{\text{ког}} = \frac{\lambda}{b} r = \frac{\lambda}{\varphi}$$

называется *радиусом пространственной когерентности*.

В соответствии со сказанным для наблюдения интерференционной картины необходимо, чтобы расстояние между щелями было меньше радиуса пространственной когерентности:

$$d < \frac{\lambda}{b} r = \rho_{\text{ког}}.$$

5.2. Видность интерференционной картины от протяжённого источника

5.2.1. Интерференционная картина в схеме Юнга

Каждая точка источника создаёт излучение, некогерентное с излучением от других точек источника. Поэтому в данном интерференционном опыте необходимо складывать интенсивности картин, создаваемых различными точками.

Будем считать источник светящимся диском, все точки которого создают излучение с одной и той же интенсивностью. Выберем участок ис-

точника $d\xi$. Тогда интенсивность излучения от этого участка можно представить в виде

$$dI = \frac{I_0}{b} d\xi; \int dI = \int_{-b/2}^{b/2} \frac{I_0}{b} d\xi = I_0. \quad (5.2.1)$$

После прохождения щелей (рис. 5.2.1) оставшаяся разность хода лучей стандартна для излучения от всех точек источника и не приводит к смещению частных интерференционных картин. Соответствующий сдвиг фаз волн от щелей до некоторой точки наблюдения P на экране $\mathcal{E}$ обозначим ψ_P . Существенной сейчас является разность хода лучей от источника до щелей $\delta = r_1 - r_2$ (рис. 5.2.1). С помощью теоремы Пифагора находим:

$$r_1^2 = r^2 + \left(\frac{d}{2} - \xi\right)^2, \quad r_2^2 = r^2 + \left(\frac{d}{2} + \xi\right)^2 \Rightarrow r_2^2 - r_1^2 = 2\xi d. \quad (5.2.2)$$

В малоугловом приближении $|r_1 - r_2| \ll r$ отсюда следует $r_2^2 - r_1^2 \approx 2r(r_2 - r_1)$, или

$$\delta = r_2 - r_1 \approx \xi d / r. \quad (5.2.3)$$

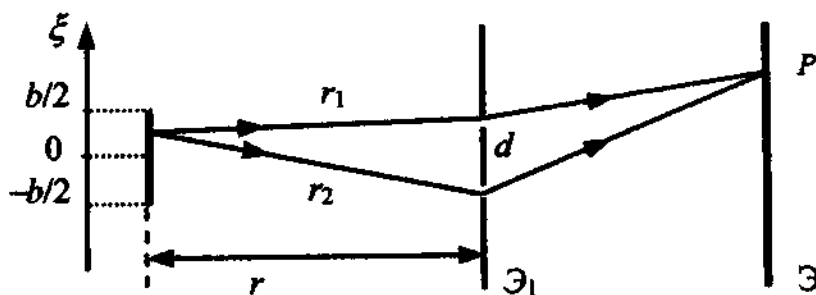


Рис. 5.2.1. К нахождению видности интерференционной картины от протяжённого источника в схеме Юнга

Далее находим распределение интенсивности излучения на экране наблюдения. Выбранный участок источника $d\xi$ создаёт на экране интерференционную картину с распределением интенсивности

$$dI = 2 \left(\frac{I_0 d\xi}{b} \right) [1 + \cos(\psi_P + k\delta)]. \quad (5.2.4)$$

Здесь $k = 2\pi/\lambda$, а величина

$$\psi_P + k\delta = \psi_P + \frac{2\pi}{\lambda} \frac{d}{r} \xi \quad (5.2.5)$$

есть полная разность фаз лучей, доходящих до точки наблюдения P от выбранной точки источника. Отсюда интегрированием по всем точкам ξ

источника получаем распределение интенсивности в наблюдаемой интерференционной картине:

$$I = 2 \frac{I_0}{b} \int_{-b/2}^{b/2} \left[1 + \cos \left(\psi_P + \frac{2\pi d}{\lambda r} \xi \right) \right] d\xi = 2I_0 \left[1 + \frac{\sin(\pi bd/\lambda r)}{(\pi bd/\lambda r)} \cos \psi_P \right]. \quad (5.2.6)$$

Фаза ψ_P зависит от положения точки наблюдения P по отношению к щелям, в зависимости от которого величина $\cos \psi_P$ меняется в пределах от -1 до $+1$. Отсюда находим видность интерференционной картины:

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = \left| \frac{\sin(\pi bd/\lambda r)}{(\pi bd/\lambda r)} \right|. \quad (5.2.7)$$

Если ввести апертуру интерференционной схемы $\Omega = d/r$, то полученное выражение можно переписать в виде

$$V = \left| \frac{\sin(\pi b\Omega/\lambda)}{(\pi b\Omega/\lambda)} \right|. \quad (5.2.8)$$

Видность убывает по мере роста параметра $\pi bd/\lambda r = \pi b\Omega/\lambda$ и обращается в первый раз в нуль при условии

$$\pi b\Omega/\lambda = \pi, \quad (5.2.9)$$

то есть при значении апертуры $\Omega = \lambda/b$. При дальнейшем росте Ω видность оказывается малой (следующий максимум почти в 5 раз меньше главного $V = 1$), и интерференционная картина выражена слабо. Следовательно, мы приходим к условию наблюдаемости интерференционной картины

$$\Omega \leq \lambda/b, \quad (5.2.10)$$

полученному ранее (в разделе 5.1) из оценочных рассуждений. Это условие эквивалентно ограничению на допустимые расстояния между щелями:

$$\frac{\pi bd}{\lambda r} < \pi, \text{ или } d < \rho_{\text{кор}} = \frac{\lambda}{b} r. \quad (5.2.11)$$

Мы получили выражение для видности интерференционной картины, считая интенсивность неизменной вдоль источника. Чтобы обобщить результат на произвольное распределение, положим, что элемент $d\xi$ источника даёт вклад в полную интенсивность на экране, равный

$$dI_0 = J(\xi) d\xi. \quad (5.2.12)$$

Тогда для интенсивности интерференционной картины вместо (5.2.6) получаем следующее выражение:

$$I = 2 \int \left[1 + \cos \left(\psi_P + \frac{kd}{r} \xi \right) \right] J(\xi) d\xi = 2I_0 + 2\Gamma, \quad (5.2.13)$$

где интегрирование распространяется по всему источнику. Здесь введена величина Γ , которая в рассматриваемом случае аналогична функции когерентности в задаче о некогерентном источнике (раздел 4.2) и даётся выражением

$$\Gamma = \text{Re} \hat{\Gamma}, \quad \hat{\Gamma} = e^{i\psi_P} \int J(\xi) e^{iq\xi} d\xi, \quad q = \frac{kd}{r}. \quad (5.2.14)$$

Положим

$$\hat{\Gamma} = I_0 e^{i\psi_P + i\varphi} |\gamma|, \quad |\gamma| = \frac{\left| \int J(\xi) e^{iq\xi} d\xi \right|}{\int J(\xi) d\xi}. \quad (5.2.15)$$

Тогда соотношение (5.2.13) примет вид

$$I = 2I_0 [1 + |\gamma| \cos(\psi_P + \varphi)]. \quad (5.2.16)$$

В зависимости от положения точки наблюдения P на экране фаза ψ_P пробегает различные значения, в результате чего интенсивность интерференционной картины меняется в диапазоне

$$2I_0(1 - |\gamma|) \leq I \leq 2I_0(1 + |\gamma|). \quad (5.2.17)$$

Соответственно находим:

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = |\gamma|. \quad (5.2.18)$$

Таким образом, видность интерференционной картины определяется преобразованием Фурье пространственного распределения интенсивности излучения по поверхности источника.

5.2.2. Звёздный интерферометр Майкельсона

В качестве одного из применений изложенной теории рассмотрим звёздный интерферометр Майкельсона (рис. 5.2.2). Свет от далёкого источника, представляющий почти параллельный пучок лучей, попадает на зеркала 3_1 и 3_2 , образующие базу интерферометра. Затем эти лучи с помощью зеркал 3_3 и 3_4 направляются через экран $\mathcal{E}_1$ с щелями (находящимися на фиксированном расстоянии друг от друга) на линзу, в фокальной плоскости которой наблюдается интерференционная картина. В этом устройстве зеркала 3_1 и

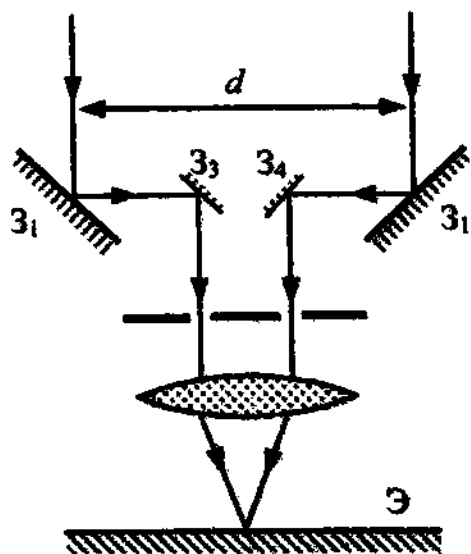


Рис. 5.2.2. Звёздный интерферометр Майкельсона

$\mathcal{Z}_2$ играют роль, аналогичную щелям в схеме Юнга. Пусть угловой диаметр звезды равен φ . Тогда согласно (5.2.7) видность интерференционной картины оказывается равной

$$V = \left| \frac{\sin(\pi\varphi d/\lambda)}{\pi\varphi d/\lambda} \right|. \quad (5.2.19)$$

Если вначале база интерферометра d мала, то есть меньше радиуса пространственной когерентности для изучаемой звезды, то на экране Э наблюдается интерференционная картина. По мере увеличения базы видность уменьшается, и, когда расстояние d достигает такого значения, что $\varphi d/\lambda = 1$, видность обращается в нуль. Зная это расстояние, можно найти угловой диаметр звезды.

Схему звёздного интерферометра предложил Альберт Абрахам Майкельсон (США) в 1890 г. Первый такой интерферометр был построен в обсерватории Маунт-Вилсон (США, Калифорния) на базе телескопа-рефлектора с диаметром зеркала в 100 дюймов (2,54 м). Отверстия в диафрагме находились на расстоянии 114 см. Максимальное расстояние между внешними зеркалами составляло 6,1 м. С помощью этого устройства были измерены угловые диаметры звёзд. Первой из них была звезда Бетельгейзе. Майкельсон нашёл (1920), что интерференционные полосы исчезали, когда расстояние между зеркалами достигало 3,065 м. Для длины волны света 575 нм отсюда можно найти, что $\varphi = 0,047''$.

Бетельгейзе — яркая звезда в созвездии Ориона (α Ориона). Это красный гигант, одна из крупнейших известных звёзд, удалённая от Земли примерно на 600 световых лет. По современным данным угловой диаметр звезды равен $\varphi = 0,055''$.

В заключение отметим, что конечность размеров источника (его отличие от точечного источника) приводит к уменьшению видности всей интерференционной картины. Число же наблюдаемых полос ограничивается немонохроматичностью излучения.

5.3. Теорема ван Циттерта–Цернике

Пусть некоторый протяжённый источник создаёт волновое поле, которое запишем в комплексной форме как

$$\hat{S}(\mathbf{r}, t) = a(\mathbf{r}, t)e^{-i\omega t}. \quad (5.3.1)$$

Это поле является в общем случае немонохроматическим и излучается некогерентно различными точками источника. Волну, приходящую в точку наблюдения, можно представить в виде суммы волн, идущих от всех точек источника:

$$a(\mathbf{r}, t) = \sum_i e^{i\varphi_i} \frac{g_i}{r_i} e^{ikr_i}. \quad (5.3.2)$$

Здесь φ_i — случайная добавка к фазе волны, производимой i -й точкой источника, связанная со случайным характером излучения, r_i — расстояние

от выбранной точки источника до точки наблюдения, g_i — амплитуда волны в точках источника.

Выберем произвольно две точки наблюдения: P_1 и P_2 , радиус-векторы которых есть соответственно $\mathbf{r}_1$ и $\mathbf{r}_2$. Волновое поле в этих точках в момент времени t равно

$$\hat{S}(\mathbf{r}_1, t) \text{ и } \hat{S}(\mathbf{r}_2, t). \quad (5.3.3)$$

Введём функцию когерентности:

$$\hat{\Gamma}_{12} = \overline{S(\mathbf{r}_1, t) S^*(\mathbf{r}_2, t)} = \overline{a(\mathbf{r}_1, t) a^*(\mathbf{r}_2, t)}, \quad (5.3.4)$$

где усреднение производится по времени, превышающему характерный период волны: $\Delta t \gg T_0 = 2\pi/\omega_0$ (как это делалось в вопросе о временной когерентности). Подставив сюда выражение (5.3.2), получим

$$\hat{\Gamma}_{12} = \sum_{i,j} \frac{g_i g_j^*}{r_{i1} r_{j2}} e^{i(\varphi_i - \varphi_j)} e^{ik(r_{i1} - r_{j2})}. \quad (5.3.5)$$

Здесь r_{i1} и r_{j2} — расстояния от соответствующих точек источника до точек P_1 и P_2 соответственно. Поскольку разные площадки источника создают некогерентные волны со случайным сдвигом фаз, то

$$\overline{\exp[i(\varphi_i - \varphi_j)]} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \quad (5.3.6)$$

Поэтому в сумме (5.3.5) остаются только слагаемые с $i = j$:

$$\hat{\Gamma}_{12} = \sum_i \frac{|g_i|^2}{r_{i1} r_{i2}} e^{ik(r_{i1} - r_{i2})}. \quad (5.3.7)$$

Перейдём от суммирования к интегрированию по всем участкам источника. Для упрощения рассмотрим сначала одномерный случай. Пусть источник плоский. Координаты его точек обозначим ξ . Будем также считать, что точки наблюдения P_1 и P_2 лежат в плоскости, параллельной плоскости источника. Координаты этих точек обозначим соответственно $\{x_1, z\}$ и $\{x_2, z\}$, где z — расстояние от плоскости источника до плоскости наблюдения (рис. 5.3.1). Кроме того, предполагаем, что расстояние z велико по сравнению с размерами источника и точками наблюдения.

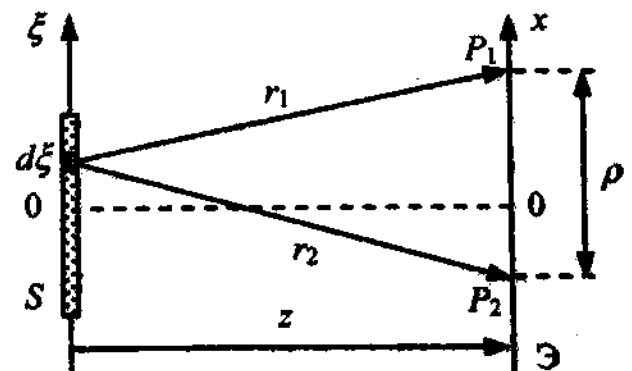


Рис. 5.3.1. Протяжённый источник S создаёт волновое поле в двух точках: P_1 и P_2 , расположенных в плоскости $\mathcal{E}$

С учётом сказанного положим

$$|g_i|^2 \rightarrow J(\xi) d\xi. \quad (5.3.8)$$

Тогда формулу (5.3.7) можно переписать в виде

$$\hat{\Gamma}_{12} = \int \frac{1}{r_1 r_2} J(\xi) e^{ik(\eta_1 - \eta_2)} d\xi. \quad (5.3.9)$$

Преобразуем это выражение, имея в виду неравенства

$$|x_1|, |x_2| \ll z, \quad |\xi| \ll z. \quad (5.3.10)$$

При данных предположениях имеем следующие приближённые соотношения:

$$r_1 = \sqrt{z^2 + (x_1 - \xi)^2} \approx z + \frac{(x_1 - \xi)^2}{2z},$$

$$r_2 = \sqrt{z^2 + (x_2 - \xi)^2} \approx z + \frac{(x_2 - \xi)^2}{2z}.$$

Отсюда находим

$$r_1 - r_2 \approx \frac{x_1^2 - x_2^2}{2z} - \frac{\xi}{z} (x_1 - x_2). \quad (5.3.11)$$

В том же приближении произведение $r_1 r_2$ в подынтегральном выражении в (5.3.9) можно заменить величиной z^2 . Таким образом, для функции когерентности получаем следующее выражение:

$$\hat{\Gamma}_{12} = \frac{1}{z^2} \exp\left(ik \frac{x_1^2 - x_2^2}{2z}\right) \int J(\xi) \exp(-iq\xi) d\xi, \quad q = \frac{k\rho}{z}. \quad (5.3.12)$$

Мы установили, что функция пространственной когерентности пропорциональна интегралу Фурье от распределения интенсивности по поверхности источника. Это утверждение называется *теоремой ван Циттерта-Цернике* по имени двух голландских физиков: Питер Хендрик ван Циттерт доказал его в 1934 г., а Фриц Цернике привёл более простой вывод в 1938 г.

Входящая в (5.3.12) величина $q = k\rho/z$ называется *пространственной частотой*.

Если точки P_1 и P_2 расположены симметрично относительно начала координат: $x_2 = -x_1$, то фазовый множитель перед интегралом в (5.3.12) обращается в единицу. Тогда нормированная функция когерентности принимает более простой вид:

$$\gamma(\rho) = \frac{\int J(\xi) \exp(-iq\xi) d\xi}{\int J(\xi) d\xi}, \quad q = \frac{k\rho}{z} \quad (5.3.13)$$

(величина $\frac{1}{z^2} \int J(\xi) d\xi$ есть интенсивность излучения в точках экрана).

Заметим, что функция (5.3.13) совпадает с видностью интерференционной картины, полученной в разделе 5.2 (формулы (5.2.15), (5.2.18)). Это не случайно. Действительно, если взять точки P_1 и P_2 в щелях экрана $\mathcal{E}_1$ в схеме Юнга (рис. 5.3.2), то на экране наблюдения $\mathcal{E}$ интерференционная картина будет даваться соотношением

$$I = \left| a_1 e^{ikL_1} + a_2 e^{ikL_2} \right|^2 = 2I_0 + 2 \operatorname{Re} \left[\overline{(a_1 a_2^*)} e^{ik(L_1 - L_2)} \right] = \quad (5.3.14)$$

$$= 2I_0 + 2 \operatorname{Re} \left(\hat{\Gamma}_{12} e^{i\psi_P} \right).$$

Здесь $\psi_P = k(L_1 - L_2)$ — разность фаз волн, доходящих от щелей до точки наблюдения P на экране $\mathcal{E}$, а $2I_0 = |a_1|^2 + |a_2|^2$ — суммарная интенсивность волн в щелях экрана $\mathcal{E}_1$. Если точки P_1 и P_2 расположены симметрично: $x_1 = -x_2$, а амплитуды интерферирующих волн одинаковы, то равенство (5.3.14) принимает вид

$$I = 2I_0 [1 + |\gamma| \cos(\psi_P + \varphi)], \quad (5.3.15)$$

в котором величина γ даётся формулой (5.3.13), φ — разность фаз интерферирующих волн в точках P_1 и P_2 . Отсюда и следует равенство $V = |\gamma|$.

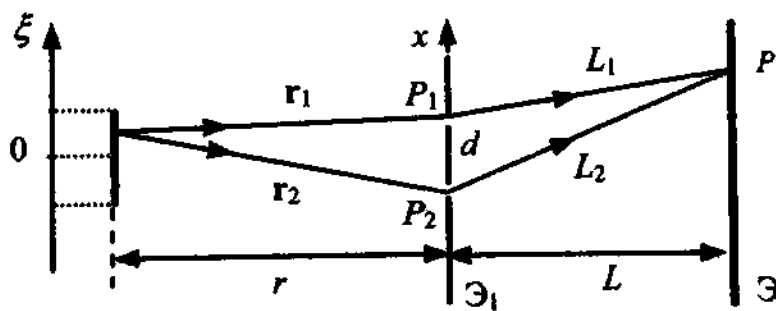


Рис. 5.3.2. К выводу связи функции пространственной когерентности с видностью интерференционной картины от протяжённого источника

Если источник имеет равномерное распределение интенсивности по поверхности, то для нормированной функции когерентности (функции видности) получается выражение (5.2.6).

Мы рассмотрели случай, когда интенсивность источника зависит только от одной координаты. Обобщение на двумерное распределение достаточно просто. Приведём окончательный результат. Пусть положение точек на источнике задаётся координатами (ξ, η) , а положение точек P_1 и P_2 на экране — координатами (x, y) . Распределение интенсивности излу-

чения по поверхности источника будем описывать функцией $J(\xi, \eta)$: $dI_0 = J(\xi, \eta)d\xi d\eta$. Тогда для нормированной функции когерентности имеет место следующее выражение:

$$\gamma(\rho_x, \rho_y) = \frac{\int J(\xi, \eta) \exp(-iq_x \xi - iq_y \eta) d\xi d\eta}{\int J(\xi, \eta) d\xi d\eta}. \quad (5.3.16)$$

Здесь использованы обозначения

$$q_x = \frac{k\rho_x}{z}, \quad q_y = \frac{k\rho_y}{z}, \quad \rho_x = x_1 - x_2, \quad \rho_y = y_1 - y_2. \quad (5.3.17)$$

Пример. Пусть источник образован двумя близко расположенными точечными источниками с интенсивностями I_1 и I_2 . Найдём функцию видности.

Выберем систему координат так, чтобы источники располагались на оси ξ симметрично относительно точки $\xi = 0$ на расстоянии b друг от друга:

$$J(\xi, \eta) = I_1 \delta(\xi + b/2) \delta(\eta) + I_2 \delta(\xi - b/2) \delta(\eta). \quad (5.3.18)$$

Пусть щели на экране в схеме Юнга также расположены вдоль оси x симметрично относительно начала координат на расстоянии d друг от друга: $x_1 = -d/2$, $x_2 = d/2$. По формулам (5.3.14), (5.3.15) находим степень когерентности:

$$\gamma = \frac{I_1 e^{iq_x b/2} + I_2 e^{-iq_x b/2}}{I_1 + I_2}, \quad q_x = \frac{kd}{z}. \quad (5.3.19)$$

Отсюда следует

$$|\gamma| = \frac{1}{I_1 + I_2} \sqrt{I_1^2 + I_2^2 + 2I_1 I_2 \cos\left(\frac{2\pi b d}{\lambda z}\right)}. \quad (5.3.20)$$

В частном случае $I_1 = I_2$ полученная формула упрощается:

$$|\gamma| = \cos\left(\frac{\pi b d}{\lambda z}\right). \quad (5.3.21)$$

Это выражение даёт также функцию видности: $V = |\gamma|$. Она обращается в нуль, если расстояние между щелями в схеме Юнга таково, что

$$\frac{\pi b d}{\lambda z} = \frac{\pi}{2} \Rightarrow d = \frac{\lambda}{2b} z = \frac{\lambda}{2\varphi}. \quad (5.3.22)$$

Здесь $\varphi = b/z$ — угловое расстояние между излучающими точками.

ГЛАВА 6

ДИФРАКЦИЯ ВОЛН

Дифракцией (от лат. *diffRACTUS* — разломанный, преломлённый) называется отклонение света от пути, предписываемого геометрической оптикой, обусловленное волновыми свойствами.

Проникновение света в область геометрической тени было известно уже в XVI–XVII веках. Однако объяснение явлению дано лишь в XIX веке. В 1800 г. Томас Юнг предположил, что дифракция обусловлена диффузией света вдоль волновых фронтов. Чередование светлых и тёмных полос на границе света и тени он считал результатом интерференции падающей волны и вторичной, цилиндрической, связанной с диффузией. Первую количественную теорию дифракции предложил французский физик Огюстен Жан Френель в 1815–1818 гг.

6.1. Принцип Гюйгенса–Френеля

6.1.1. Принцип Гюйгенса

Принцип Гюйгенса позволяет найти форму волнового фронта в различные моменты времени, если она известна в некоторый начальный момент. Этот принцип, сформулированный нидерландским учёным Христианом Гюйгенсом в 1678 г., состоит в следующем.

Каждая точка волнового фронта рассматривается как источник вторичных сферических волн. Огибающая фронтов этих волн в один и тот же момент времени даёт новое положение волнового фронта. Данный принцип проиллюстрирован на рис. 6.1.1.

Идею можно пояснить следующим экспериментом (рис. 6.1.2). Пусть на непрозрачный экран падает плоская волна. Если в экране сделано малое отверстие, то за экраном от отверстия вперёд распространяется сферическая волна. Количество отверстий можно сделать большим, и от каждого из них за экраном идёт сферическая волна.

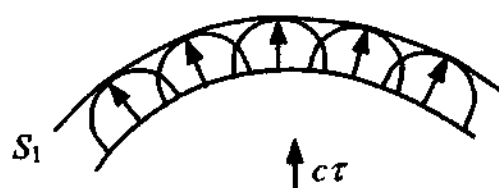


Рис. 6.1.1. Принцип Гюйгенса построения положения волнового фронта в некоторый момент t по известному положению в начальный момент. Длина стрелок равна расстоянию ct , проходимому волной за время t

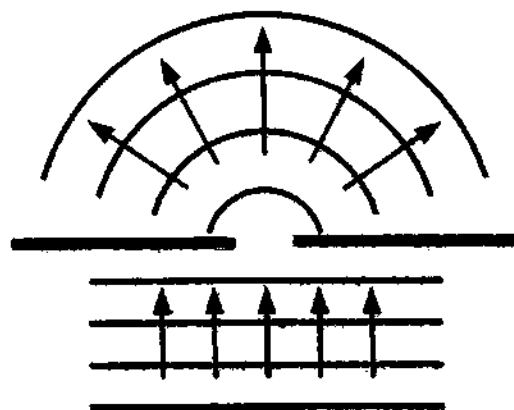


Рис. 6.1.2. Плоская волна падает на непрозрачный экран с малым отверстием. В результате за экраном возникает сферическая волна с центром, находящимся в отверстии

В результате приходим к исходному положению принципа Гюйгенса.

Принцип Гюйгенса формально справедлив только для описания распространения света в рамках геометрической оптики. Вместе с тем он позволяет объяснить многие особенности поведения световых полей в разных условиях.

6.1.2. О каустиках и особенностях волновых фронтов

Кроме описания волновых полей на языке волновых фронтов, существует ещё один, дополнительный, способ их анализа. Имеется в виду описание волновых полей на основе лучевых представлений (раздел 2.2). Подходы, основанные на представлении о волновых фронтах и лучах, очевидно, эквивалентны, поскольку, по определению, лучи — это линии, нормальные к волновым фронтам. Направления лучей есть направления движения волнового фронта.

Пусть, например, эллиптическая поверхность посылает световые лучи внутрь (рис. 6.1.3). Положение волнового фронта в начальный момент времени $t = 0$ представлено поверхностью S_0 . По принципу Гюйгенса, чтобы найти волновой фронт S в момент времени $t > 0$, нужно в каждой точке поверхности S_0 провести нормали (лучи) и на каждой из них отложить в направлении распространения волны отрезки длиной ct , где c — скорость распространения волны. Геометрическое место всех построенных данным способом точек представляет собой искомый волновой фронт S . В математике такая поверхность называется *эквидистантой* исходного фронта.

В случае, показанном на рис. 6.1.3, начальное положение волнового фронта есть S_0 . Положения в два последующих момента времени

$0 < t_1 < t_2$ обозначены на том же рисунке как S_1 и S_2 . Если фронт S_1 (в момент времени t_1) представляет собой гладкую кривую, то уже фронт S_2 (в момент времени t_2) есть кривая с двумя самопересечениями. Заметим, что если бы фронт S_0 был выпуклым (то есть посылал световой поток наружу), то лучи при $t > 0$ не пересекались бы.

Пусть волновой фронт вогнут. Тогда, как было сказано, лучи пересекаются. В этом случае можно построить огибающую системы лучей как непрерывную кривую (или поверхность),

касающуюся лучей в каждой своей точке. В математике такая кривая (поверхность) называется *эволютой*, а исходный волновой фронт — *эвольвентой* или *инволютой*. При этом оказывается, что эволюта есть геомет-

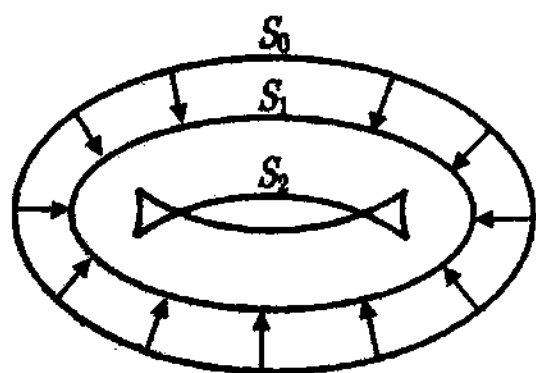


Рис. 6.1.3. Светящаяся поверхность S_0 посылает световой поток внутрь. Через некоторое время форма волнового фронта дается поверхностью S_1 , а далее принимает вид S_2 — поверхности с самопересечениями

рическое место центров кривизны эвольвенты. В физике (в частности, в оптике и в теории волн) эволюту называют *каустикой* (от греч. *καυστικός* — жгущий, палящий, воспламеняющийся). Данное название обусловлено тем, что вдоль каустики вследствие сгущения лучей интенсивность излучения существенно выше, чем вдали от неё.

Пример построения каустики при отражении плоской волны вогнутым зеркалом приведен на рис. 6.1.4 (слева). Формирование каустики сходящейся двумерной эллиптической волны показано на рис. 6.1.4 (справа).

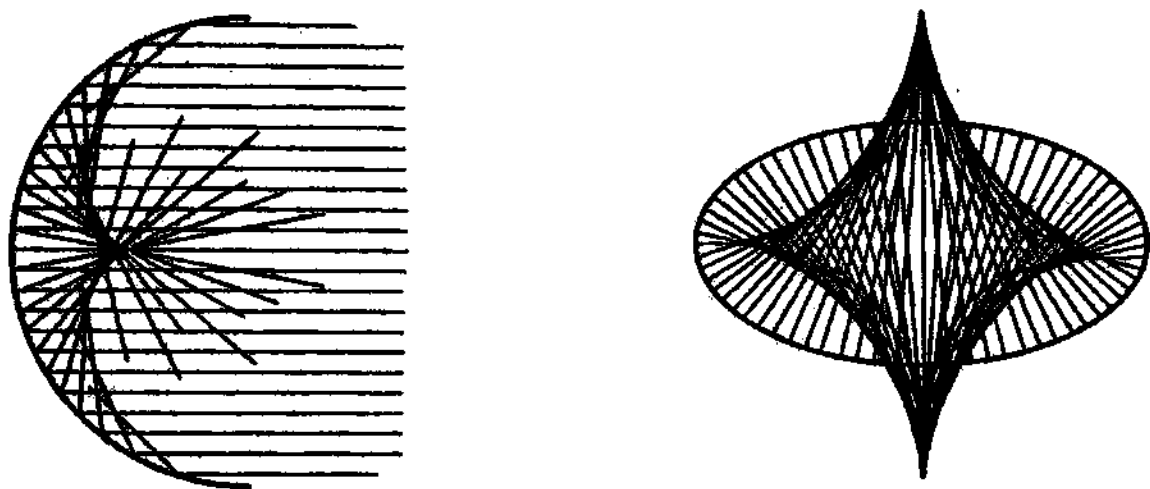


Рис. 6.1.4. Слева — каустика при отражении плоской волны от вогнутого зеркала; справа — каустика сходящейся волны, создаваемой эллиптической поверхностью

6.1.3. Принцип Гюйгенса–Френеля

В 1816 г. О. Френель дополнил принцип Гюйгенса представлением об интерференции вторичных волн:

каждый элемент волнового фронта является центром возмущения, порождающего вторичные сферические волны, в результате интерференции которых формируется результирующее световое поле в каждой точке пространства.

Данное утверждение получило название *принципа Гюйгенса–Френеля*.

6.1.4. Интеграл Френеля

Дадим математическую формулировку принципа Гюйгенса–Френеля.

Пусть имеется стационарное волновое поле

$$S(\mathbf{r}, t) = A(\mathbf{r})e^{-i\omega t}. \quad (6.1.1)$$

Заметим, что функция $A(\mathbf{r})$ удовлетворяет уравнению Гельмгольца:

$$\Delta A + \frac{\omega^2}{c^2} A = 0. \quad (6.1.2)$$

Выделим на волновом фронте Σ (рис. 6.1.5) малую площадку $d\sigma$ (вектор $d\sigma$ направлен по нормали к площадке), на которой амплитуда волны равна A_Σ . Эту площадку можно рассматривать как источник вторичной сферической волны $\sim A_\Sigma e^{ikr}/r$. Полная амплитуда волны, излучаемой площадкой, пропорциональна также площади этой площадки. Таким образом, доходящую до точки P волну можно представить в виде

$$dA_P = K(\alpha) A_\Sigma \frac{e^{ikr}}{r} d\sigma. \quad (6.1.3)$$

Множитель $K(\alpha)$ учитывает ориентацию площадки по отношению к направлению на точку наблюдения P . Для решения ряда задач дифракции можно принять, что до точки P доходит волна, пропорциональная *видимой* площади: $d\sigma_\perp = d\sigma \cos \alpha$, где α — угол между нормалью к площадке и направлением на точку наблюдения P (рис. 6.1.5). Соответственно положим

$$K(\alpha) = K_0 \cos \alpha. \quad (6.1.4)$$

Здесь $K_0 = 1/i\lambda$ (это значение будет найдено в разделе 6.2.3). Для большинства задач дифракции, рассматривающих параксиальные пучки волн, угловая зависимость $K(\alpha)$ не существенна:

$$K(\alpha) \approx K(0) = K_0.$$

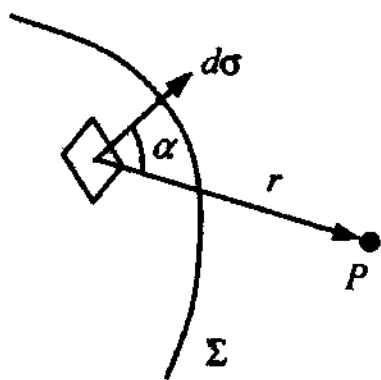


Рис. 6.1.5. К математической формулировке принципа Гюйгенса-Френеля

Результирующая волна в точке P получается суммированием волн, приходящих от всех точек волнового фронта:

$$A(P) = \int_{\Sigma} K(\alpha) A_\Sigma \frac{e^{ikr}}{r} d\sigma. \quad (6.1.5)$$

Эта формула даёт приближённое решение уравнения Гельмгольца (6.1.2) с граничным условием $A(\mathbf{r})|_{\Sigma} = A_\Sigma$ и называется *интегралом Френеля*, или *дифракционным интегралом*.

Сам Френель не конкретизировал вид функции $K(\alpha)$. Точное решение волнового уравнения для задач дифракции было получено немецким физиком Г. Р. Кирхгофом (1883) и приводит к следующему выражению: $K(\alpha) = (1 + \cos \alpha)/2i\lambda$. Для случая зависимости (6.1.4)

формула (6.1.5) называется *интегралом скалярной теории дифракции Релея–Зоммерфельда*. Она позволяет качественно описать уменьшение вклада участков волнового фронта, наблюдаемых при больших углах α .

6.2. Задачи с осевой симметрией

6.2.1. Преобразование интеграла Френеля

Рассмотрим частный, но важный случай. Пусть исходная волна создаётся точечным источником S (рис. 6.2.1). Эту волну запишем в виде

$$A_S(R) = (A_0 r_S) \frac{1}{R} e^{ikR}, \quad (6.2.1)$$

так что в точках волнового фронта Σ радиуса $R = r_S$ оказывается

$$A_\Sigma = (A_0 r_S) \frac{e^{ikr_S}}{r_S} = A_0 e^{ikr_S}. \quad (6.2.2)$$

Выделим на волновой поверхности Σ кольцевой элемент (поясок):

$$d\sigma = 2\pi r_S^2 \sin \theta d\theta \quad (6.2.3)$$

(на рис. 6.2.1 этот элемент заштрихован). Все его участки посылают в точку наблюдения P одинаковые волны.

Рассмотрим треугольник SPA . Согласно теореме косинусов имеем следующее равенство, связывающее угол θ и расстояние r :

$$r^2 = r_S^2 + (r_S + r_0)^2 - 2r_S(r_S + r_0) \cos \theta. \quad (6.2.4)$$

При фиксированном положении точек S и P неизменными остаются величины r_S и r_0 , и из (6.2.4) следует равенство для приращений θ и r :

$$r dr = r_S(r_S + r_0) \sin \theta d\theta. \quad (6.2.5)$$

Подстановка этого равенства в (6.2.4) даёт

$$d\sigma = 2\pi r_S^2 \frac{r dr}{r_S(r_S + r_0)} = 2\pi \frac{r_S r}{r_S + r_0} dr. \quad (6.2.6)$$

Соответственно интеграл Френеля (6.1.5) принимает вид

$$A(P) = \frac{2\pi r_S A_0}{r_S + r_0} e^{ikr_S} \int_{\Sigma} K(\alpha) e^{ikr} dr, \quad \alpha = \alpha(r). \quad (6.2.7)$$

Здесь учтено выражение (6.2.2) для волны A_Σ .

6.2.3. Зоны Френеля

Пусть источник S посылает сферическую волну на экран с круглым отверстием радиуса R (рис. 6.2.2). Волновой фронт, попадающий в отвер-

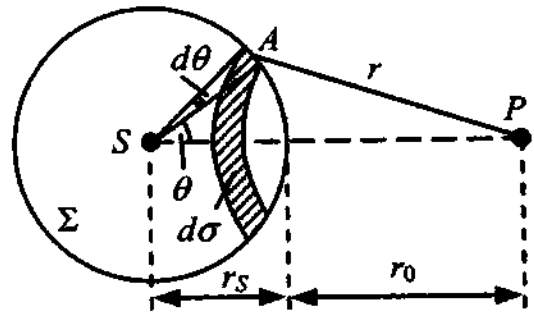


Рис. 6.2.1. Точечный источник S создаёт сферическую волну, регистрируемую в точке наблюдения P

ствие, обозначен как Σ . Проведём из точки наблюдения P , как из центра, серию сферических поверхностей. Первая из них касается фронта Σ . Её радиус обозначим r_0 . Каждая из последующих сфер имеет радиус, на $\lambda/2$ больший, чем у предыдущей (как показано на рис. 6.2.2). Иными словами, радиусы сфер равны

$$r_n = r_0 + n\lambda/2, \quad n = 0, 1, 2, \dots \quad (6.2.8)$$

Эти сферы разбивают весь волновой фронт на кольцевые области, называемые зонами Френеля.

Считаем, что $r_0 \gg \lambda$, так что при переходе из одной зоны в другую угол α меняется мало, и в пределах одной зоны функцию $K(\alpha)$ в интеграле Френеля можно считать почти постоянной: в первой зоне $K = K_1$, во второй зоне $K = K_2$ и так далее.

Имея это в виду, запишем интеграл Френеля как сумму интегралов по отдельным зонам Френеля:

$$A(P) = \frac{2\pi r_s A_0}{r_s + r_0} e^{ikr_s} \sum_{n=1}^N K_n \int_{r_n}^{r_{n+1}} e^{ikr} dr. \quad (6.2.9)$$

Здесь N — число зон, попадающих в отверстие. Входящие (6.2.9) интегралы равны

$$\begin{aligned} \int_{r_n}^{r_{n+1}} e^{ikr} dr &= \frac{1}{ik} (e^{ikr_{n+1}} - e^{ikr_n}) = \frac{1}{ik} e^{ikr_0} \left[\exp\left(i \frac{k\lambda}{2} n\right) - \exp\left(i \frac{k\lambda}{2} (n-1)\right) \right] = \\ &= \frac{1}{ik} e^{ikr_0} [(-1)^n - (-1)^{n-1}] = \frac{2}{ik} e^{ikr_0}. \end{aligned}$$

Здесь учтено тождество $e^{ik\lambda/2} = e^{i\pi} = -1$.

Таким образом, интеграл Френеля (6.2.6) представлен в виде суммы слагаемых, отвечающих отдельным зонам Френеля:

$$A(P) = \sum_{n=1}^N A_n, \quad A_n = \frac{4\pi}{ik} A_0 \frac{r_s}{r_s + r_0} e^{ik(r_s + r_0)} (-1)^n K_n. \quad (6.2.10)$$

Далее учтём, что вследствие малости размеров зон значения K_n для соседних зон мало отличаются друг от друга. Поэтому слагаемые A_n и

A_{n+1} близки по абсолютной величине, но отличаются знаком: $A_{n+1} \approx -A_n$. Вместе с тем вследствие роста угла α и соответствующего монотонного убывания функции $K(\alpha)$ можно утверждать, что

$$\lim_{n \rightarrow \infty} K_n = 0. \quad (6.2.11)$$

Имея это в виду, найдём сумму $A(P) = \sum_{n=1}^{\infty} A_n$, отвечающую случаю, когда открыты все зоны Френеля. Перегруппируем слагаемые в (6.2.10):

$$A(P) = \sum_{n=1}^{\infty} A_n = \frac{1}{2} A_1 + \left(\frac{1}{2} A_1 + A_2 + \frac{1}{2} A_3 \right) + \left(\frac{1}{2} A_3 + A_4 + \frac{1}{2} A_5 \right) + \dots \quad (6.2.12)$$

Полагая по правилу интерполяции с учётом чередования знаков слагаемых

$$A_{2k} \approx -\frac{1}{2} (A_{2k-1} + A_{2k+1}),$$

получаем, что в сумме (6.2.12) обращаются в нуль все слагаемые, кроме первого. Таким образом, в случае, когда открыт весь волновой фронт, волна в точке наблюдения оказывается равной

$$A(P) = \frac{1}{2} A_1 = -\frac{2\pi}{ik} A_0 \frac{r_S}{r_S + r_0} e^{ik(r_S + r_0)} K_1. \quad (6.2.13)$$

С другой стороны, эта же волна равна волне (6.2.1), которая доходит до точки наблюдения P непосредственно от источника S :

$$A(P) = A_S|_{R=r_S+r_0} = A_0 r_S \frac{1}{r_S + r_0} e^{ik(r_S + r_0)}. \quad (6.2.14)$$

Приравнявая выражения (6.2.13) и (6.2.14) для $A(P)$, получаем

$$K_1 = K(0) = \frac{k}{2\pi i} = \frac{1}{i\lambda}. \quad (6.2.15)$$

Таким образом, мы нашли значение коэффициента $K_0 = K(0)$ в формуле (6.1.4). Одновременно установлено, что амплитуда волны A_1 , приходящей в точку наблюдения P от первой зоны Френеля, вдвое превышает амплитуду волны $A(P)$, приходящей от всего волнового фронта (6.2.14):

$$A_1 = 2 A_S|_{R=r_S+r_0} = 2 A_0 r_S \frac{1}{r_S + r_0} e^{ik(r_S + r_0)}. \quad (6.2.16)$$

Рассмотрим характеристики зон Френеля.

Найдём сначала радиусы зон. Пусть волна от источника S падает на отверстие радиуса R . Ограничимся приосевым приближением, когда радиус отверстия R мал по сравнению с расстояниями от отверстия до источника и точки наблюдения: $R \ll r_S$, $R \ll r_P$. Как видно из рис. 6.2.2 и из

определения (6.2.8), число зон, укладываемых на открытой части волнового фронта, равно

$$m = \frac{AB}{\lambda/2}. \quad (6.2.17)$$

Длину отрезка AB найдём используя равенства (рис. 6.2.2)

$$\begin{aligned} r_S^2 &= a^2 + R^2 \Rightarrow r_S - a = \frac{R^2}{r_S + a} \approx \frac{R^2}{2a}, \\ r_P^2 &= b^2 + R^2 \Rightarrow r_P - b = \frac{R^2}{r_P + b} \approx \frac{R^2}{2b}. \end{aligned} \quad (6.2.18)$$

Здесь a и b — расстояния до экрана соответственно от источника и точки наблюдения. Учтено также, что для рассматриваемых зон $r_S + a \approx 2a$, $r_P + b \approx 2b$. Поскольку

$$AB = (r_S - a) + (r_P - b) = \frac{R^2}{2} \left(\frac{1}{a} + \frac{1}{b} \right),$$

то искомое число зон оказывается равным

$$m = \frac{R^2}{\lambda} \left(\frac{1}{a} + \frac{1}{b} \right). \quad (6.2.19)$$

Обозначим

$$\frac{1}{f} = \frac{1}{a} + \frac{1}{b}, \text{ или } f = \frac{ab}{a+b}. \quad (6.2.20)$$

Тогда для радиуса m -й зоны Френеля получаем следующее выражение:

$$R_m = \sqrt{m\lambda f}. \quad (6.2.21)$$

Найдём площади зон Френеля:

$$S_m = \pi R_m^2 - \pi R_{m-1}^2 = \pi\lambda f. \quad (6.2.22)$$

Таким образом, первые зоны Френеля имеют одинаковые площади:

$$S_m = S_1.$$

Если на отверстие падает плоская волна ($a = \infty$), а точка наблюдения отстоит от отверстия на расстояние $b = z$, то $f = z$ и

$$R_m = \sqrt{m\lambda z}, \quad S_m = \pi\lambda z. \quad (6.2.23)$$

6.2.4. Зонные пластинки

Пусть на экране полностью перекрыты все чётные (или все нечётные) зоны Френеля. Тогда все оставшиеся зоны направляют в точку наблюдения волны с одной и той же фазой — все открытые зоны усиливают действие друг друга. Такой экран называется *зонной пластинкой Френеля*

(точнее, *амплитудной зонной пластинкой*). На рис. 6.2.3 показана зонная пластинка, у которой открыты 1-я, 3-я, 5-я, ... зоны.

Фокусом зонной пластинки называется такая точка F , что для каждого из непрозрачных и прозрачных колец расстояния от его границ до точки F отличаются на $\lambda/2$. Другими словами, все прозрачные и непрозрачные кольца представляют целые зоны Френеля. В фокусе все идущие от пластинки волны складываются с одной и той же фазой и усиливают друг друга. Величина

$$f = \frac{ab}{a+b},$$

найденная в предыдущем разделе (формула (6.2.20)), называется **фокусным расстоянием** зонной пластинки.

Найдём интенсивность излучения в фокусе зонной пластинки. Пусть радиус пластинки равен R . Этому радиусу соответствует число зон, равное

$$m = R^2 / \lambda f.$$

Поскольку перекрыта половина всех зон, то оставшиеся открытыми $m/2$ зон создают в фокусе поле

$$A = \frac{m}{2} A_1 = mA_0 = \frac{R^2}{\lambda f} A_0,$$

а интенсивность излучения в фокусе оказывается равной

$$I = A^2 = \left(R^2 / \lambda f \right)^2 I_0.$$

Наряду с амплитудными зонными пластинками можно рассматривать и **фазовые пластинки**. В этом случае, например, все нечётные зоны прозрачные, а все чётные перекрыты прозрачным материалом, меняющим фазу волны на π . Тогда уже *все* зоны посылают в фокус волны с одной и той же фазой. В результате интенсивность излучения возрастает ещё в $2^2 = 4$ раза по сравнению со случаем амплитудной пластинки:

$$A = mA_1 = 2mA_0 \Rightarrow I = A^2 = 4 \left(R^2 / \lambda f \right)^2 I_0.$$

Заметим, что если зонная пластинка перекрывает большое число зон Френеля ($m \gg 1$), то последняя из них может быть неполной: она частично компенсирует вклад предпоследней зоны, тогда как основную часть поля при $m \gg 1$ создают все прочие целые зоны.

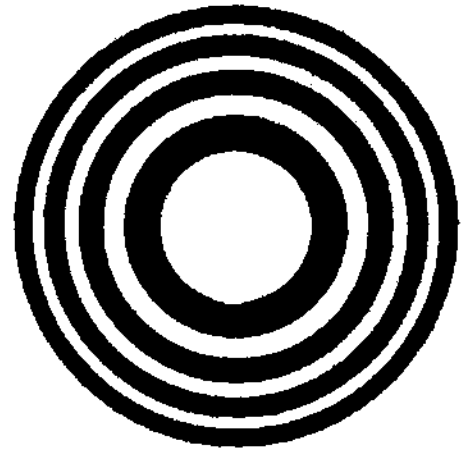


Рис. 6.2.3. Зонная пластинка с открытыми нечётными зонами и перекрытыми чётными

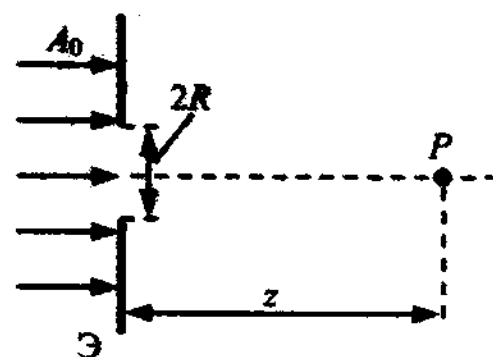
Дополнительные вопросы, относящиеся к зонной пластинке, рассмотрены в разделе 6.2.7.

6.2.5. Поле за круглым отверстием на оси системы

Пусть на отверстие радиуса R в непрозрачном экране падает по нормали плоская волна с амплитудой A_0 и интенсивностью $I_0 = |A_0|^2$. Найдём поле на оси системы на расстоянии z от экрана (рис. 6.2.4).

Будем полагать, что точка наблюдения находится на таком расстоянии, при котором угол видности α элементов волновой поверхности мал.

В рассматриваемых условиях интеграл Френеля (6.2.7) принимает вид



$$A(P) = 2\pi A_0 K(0) \int_{r_{\min}}^{r_{\max}} e^{ikr} dr = \quad (6.2.24)$$

$$= -A_0 (e^{ikr_{\max}} - e^{ikr_{\min}})$$

Рис. 6.2.4. На экран Э с круглым отверстием падает плоская волна, наблюдаемая в точке P на оси

(фазовый множитель e^{ikr_0} одинаков для всех точек волнового фронта и может быть отброшен). Минимальное и максимальное расстояния от точки P до открытых точек волнового

фронта равны

$$r_{\min} = z, \quad r_{\max} = \sqrt{z^2 + R^2}. \quad (6.2.25)$$

Для интенсивности излучения находим выражение

$$I(P) = |A(P)|^2 = 2I_0 [1 - \cos(k(r_{\max} - r_{\min}))]. \quad (6.2.26)$$

Преобразуем разность $r_{\max} - r_{\min}$ для параксиального пучка $R \ll z$:

$$r_{\max} - r_{\min} = \sqrt{z^2 + R^2} - z \approx R^2/2z. \quad (6.2.27)$$

Окончательно получаем

$$I(P) = 4I_0 \sin^2 \left(\frac{\pi R^2}{2\lambda z} \right). \quad (6.2.28)$$

Этому выражению можно придать иной вид, если заметить, что $S = \pi R^2$ — площадь отверстия, а $S_1 = \pi \lambda z$ — площадь одной зоны Френеля:

$$I(P) = 4I_0 \sin^2 \left(\frac{\pi S}{2 S_1} \right). \quad (6.2.29)$$

Отсюда видно, что интенсивность излучения в точке наблюдения максимальна, когда открыто нечётное число зон Френеля:

$$S = (2n+1)S_1, \quad n = 0, 1, 2, \dots$$

Если же открыто чётное число зон, то интенсивность обращается в нуль.

Из (6.2.28) следует, что на малых расстояниях z от отверстия ($z \ll R^2/\lambda$) интенсивность осциллирует, тогда как на больших расстояниях ($z \gg R^2/\lambda$) она монотонно убывает обратно пропорционально квадрату расстояния (подобно полю точечного источника):

$$I(P) \approx \left(\pi R^2 / \lambda z \right)^2 I_0 \sim 1/z^2. \quad (6.2.30)$$

Эта зависимость видна на рис. 6.2.5. В частности, для радиуса отверстия $R = 100\lambda$ переход от осцилляций к монотонному убыванию происходит при $z/R \sim R/\lambda = 100$.

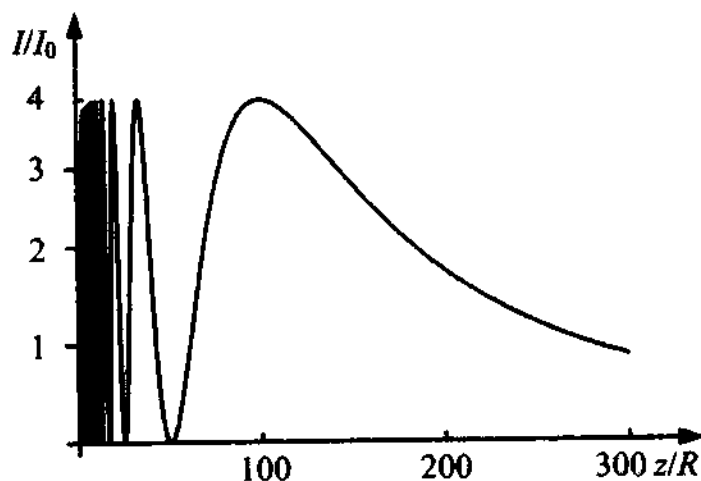


Рис. 6.2.5. Зависимость интенсивности излучения на оси отверстия от расстояния. График построен для радиуса отверстия $R = 100\lambda$

В действительности по мере приближения точки наблюдения к отверстию амплитуда осцилляций убывает, что соответствует переходу к полностью открытому волновому фронту. Укажем некоторые причины этого.

1. Спектр реального излучения всегда немонохроматичен. Чтобы оценить ограничиваемое этим число наблюдаемых максимумов, учтём, что при длине волны λ положение n -го максимума согласно (6.2.28) определяется условием

$$\frac{\pi R^2}{2z_{\max,n}\lambda} = \left(n + \frac{1}{2} \right) \pi \Rightarrow z_{\max,n} = \frac{R^2}{(2n+1)\lambda}.$$

Для длины волны λ' минимум n -го порядка попадает в точку

$$\frac{\pi R^2}{2z_{\min,n}\lambda'} = n\pi \Rightarrow z_{\min,n} = \frac{R^2}{2n\lambda'}.$$

Полагая $z_{\max,n}(\lambda) = z_{\min,n}(\lambda')$, находим

$$2n\lambda' = (2n+1)\lambda \Rightarrow \Delta\lambda \equiv \lambda' - \lambda = \lambda/2n.$$

Таким образом, максимальное число наблюдаемых максимумов интенсивности составляет

$$n = \lambda / (2\Delta\lambda). \quad (6.2.31)$$

2. К исчезновению осцилляций ведёт также неровность края отверстия δR на величину порядка ширины последней зоны Френеля. Поскольку согласно (6.2.23) радиус зон Френеля даётся формулой $R_m^2 = m\lambda z$, то изменение радиуса δR , отвечающее одной зоне, можно оценить из соотношения $2R\delta R = \lambda z$, откуда находим предельно допустимую неровность края отверстия:

$$\delta R \sim \lambda z / 2R. \quad (6.2.32)$$

В частности, если точка наблюдения находится на расстоянии от отверстия $z \sim R$, то неровность не должна превышать $\delta R \sim \lambda/2$.

Если величина δR задана, то формула (6.2.32) позволяет найти расстояния до отверстия, на которых колебания интенсивности практически исчезают:

$$z \sim 2 \frac{R}{\lambda} \delta R. \quad (6.2.32a)$$

Оценим также число наблюдаемых максимумов интенсивности вдоль оси при заданной неровности края отверстия. Используя (6.2.32) и равенство $R_n^2 = n\lambda z$, находим

$$2R\delta R \sim \lambda z \sim \frac{R^2}{n} \Rightarrow n \sim \frac{R}{2\delta R}. \quad (6.2.33)$$

3. Осцилляции могут не регистрироваться, если расстояние (вдоль оси z) между ними меньше разрешаемой измерительным прибором величины.

Пусть предел разрешения есть δz . Согласно (6.2.28) расстояние между соседними максимумами составляет (при $n \gg 1$)

$$\delta z = \frac{R^2}{\lambda} \left(\frac{1}{2n-1} - \frac{1}{2n+1} \right) \approx \frac{R^2}{2\lambda n^2}.$$

Поскольку $n \sim R^2 / \lambda z$, то из последней формулы находим расстояние, на котором осцилляции интенсивности уже не наблюдаются:

$$z_{\min} \sim R \sqrt{\delta z / \lambda}. \quad (6.2.34)$$

4. Наконец, осцилляции интенсивности могут исчезать вследствие ослабления (например, благодаря рассеянию или поглощению в среде) излучения, проходящего большие расстояния от точек экрана до места наблюдения.

6.2.6. Графический метод суммирования амплитуд

Метод зон Френеля допускает наглядное графическое представление, позволяющее относительно просто анализировать дифракцию на различных отверстиях или препятствиях.

Запишем волну, приходящую в точку наблюдения при полностью открытом волновом фронте, как A_0 . Как было установлено в разделе 6.2.3, вклад в волну от первой зоны $A_1 \approx 2A_0$, вклад от следующих зон

$$A_2 \approx -A_1 \approx -2A_0, \quad A_3 \approx -A_2 \approx 2A_0, \dots, \quad (6.2.35)$$

причём для больших номеров зон $A_n \rightarrow 0$.

Амплитуда волны есть комплексная величина, которая представляется вектором на комплексной плоскости. Тогда вклад отдельных целых зон можно представить вертикальными векторами (рис. 6.2.6). Пунктирными линиями показано убывание амплитуд с ростом номера. Соответственно поле $A(P)$, формируемое некоторым набором зон, будет даваться суммой векторов A_n , отвечающих этим зонам.

Данный подход позволяет рассчитать поле в случае, когда открыто целое число зон. Если же какие-либо зоны открыты лишь частично, требуется модифицировать процедуру.

Разобьём весь открытый волновой фронт на малые участки с площадями $\Delta\sigma$ и запишем интеграл Френеля в виде суммы:

$$A(P) = e^{ikr_1} \sum_n C_n e^{ik(r_n - r_1)} \Delta\sigma. \quad (6.2.36)$$

Здесь выделен фазовый множитель e^{ikr_1} , отвечающий внутренней границе первой открытой зоны.

С ростом номера площадки увеличивается фаза соответствующего слагаемого

$$\theta_n = k(r_n - r_1). \quad (6.2.37)$$

Выбирая за начало отсчёта фазу первого слагаемого, получаем сумму векторов на комплексной плоскости, постепенно поворачивающихся с ростом номера площадки. Фаза (6.2.37) имеет смысл угла, образуемого вектором $C_n e^{ik(r_n - r_1)} \Delta\sigma$ с осью абсцисс (рис. 6.2.7). Направление суммарного вектора поля A связано с фазой последнего элемента в сумме соотношением

$$\varphi = \theta/2. \quad (6.3.38)$$

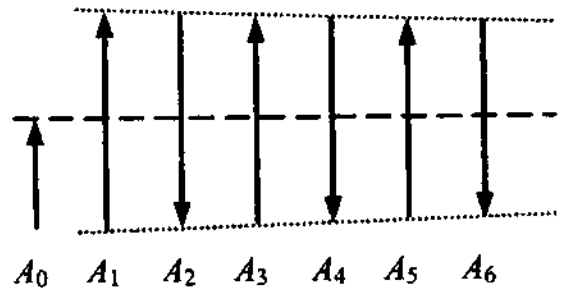


Рис. 6.2.6. Графическое (векторное) представление первых зон Френеля. Отдельно показан вектор A_0 волны от полностью открытого волнового фронта

На рис. 6.2.7 показано сложение векторов, отвечающих слагаемым в сумме (6.2.36) для первой (слева) и второй (справа) зон Френеля. В результате сложения векторов, отвечающих целой n -й зоне Френеля, получается вектор A_n , представляемый вертикальной стрелкой (рис. 6.2.6).

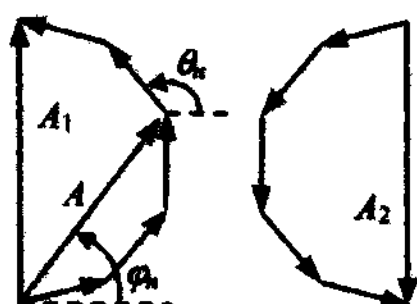


Рис. 6.2.7. Графическое суммирование амплитуд, отвечающих площадкам $\Delta\sigma$ на волновом фронте, для первой (слева) и второй (справа) зон Френеля

Переходя к разбиению волнового фронта на бесконечно малые площадки, получаем непрерывную кривую, называемую *спиралью Френеля* (рис. 6.2.8). Спираль постепенно закручивается к центру. Вектор A_0 , идущий от начальной точки O к центру, представляет волну, создаваемую в точке наблюдения полностью открытым волновым фронтом. На том же рисунке показан вектор A_1 , представляющий вклад только от первой зоны Френеля.

В действительности для длин волн, характерных свету, закручивание очень медленное вследствие условий $R \gg \lambda$, $z \gg \lambda$. Поэтому изменением радиуса можно пренебречь для большого числа витков и считать спираль окружностью радиуса A_0 .

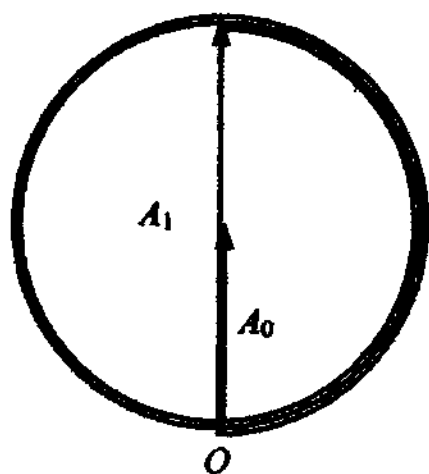


Рис. 6.2.8. Спираль Френеля для первых пяти зон. Вектор A_0 представляет волну, идущую от всего волнового фронта, а волну, идущую только от первой зоны, — вектор $A_1 = 2A_0$

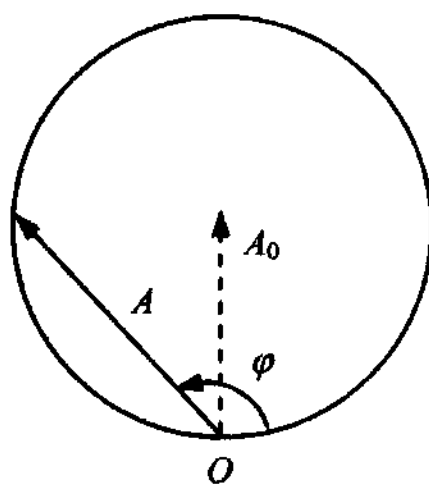


Рис. 6.2.9. Спираль Френеля для случая, когда открыты первая и частично вторая зоны. Вектор A изображает амплитуду результирующего поля, вектор A_0 представляет волну, идущую от всего волнового фронта

На рис. 6.2.9 показано нахождение амплитуды волны, когда открыты первая и частично вторая зоны. Вектор результирующей амплитуды повернут на угол φ , значение которого даётся формулами (6.2.37), (6.2.38).

Пусть отверстие имеет радиус R и находится на расстоянии z от точки наблюдения. Тогда в приосевом приближении ($R \ll z$) находим

$$\varphi = \frac{1}{2} k(r - r_1) = \frac{1}{2} k \left(\sqrt{R^2 + z^2} - z \right) \approx \frac{1}{2} \frac{2\pi}{\lambda} \frac{R^2}{2z} = \frac{\pi}{2} \frac{S}{S_1}. \quad (6.2.39)$$

Здесь $S = \pi R^2$ — площадь открытой части волнового фронта, $S_1 = \pi \lambda z$ — площадь одной зоны Френеля. Следовательно, угол φ , образуемый вектором поля с осью абсцисс, прямо пропорционален площади открытой части волнового фронта. В частности, полностью открытой первой зоне отвечает угол $\varphi = \pi/2$. Если же открыты полторы зоны, то

$$\varphi = \pi/2 + \pi/4 = 3\pi/4, \quad |A| = |2A_0 \sin \varphi| = \sqrt{2} |A_0|, \quad I = 2I_0.$$

6.2.7. Ещё о зонной пластинке

Подход, изложенный в разделе 6.2.5, позволяет рассчитать поле, создаваемое зонной пластинкой на любом расстоянии от неё.

Пусть в непрозрачном экране имеется круглое отверстие, в которое укладывается целое зон Френеля. Фокусное расстояние обозначим f . Считаем, что на пластинку по нормали падает плоская волна интенсивностью $I_0 = |A_0|^2$. Найдём поле, создаваемое одной зоной Френеля на оси в точке P , удалённой на расстояние z от плоскости отверстия и в общем случае не совпадающей с фокусом.

Согласно (6.2.24) кольцо, для которого расстояния от его краёв до точки наблюдения равны $r_{\min}$ и $r_{\max}$, создаёт поле

$$A_n(x) = -A_0 \left(e^{ikr_{\max}} - e^{ikr_{\min}} \right).$$

Обозначая радиусы внутренней и внешней границ n -й зоны как R_{n-1} и R_n , имеем

$$r_{\min} = \sqrt{z^2 + R_{n-1}^2} \approx z + R_{n-1}^2/2z, \quad r_{\max} = \sqrt{z^2 + R_n^2} \approx z + R_n^2/2z,$$

откуда следует

$$A_n = -A_0 e^{ikz} \left[\exp\left(ikR_n^2/2z\right) - \exp\left(ikR_{n-1}^2/2z\right) \right].$$

Учтём далее, что радиусы R_n и R_{n+1} равны соответственно

$$R_{n-1} = \sqrt{(n-1)\lambda f}, \quad R_n = \sqrt{n\lambda f},$$

и воспользуемся выражением для волнового числа: $k = 2\pi/\lambda$. В результате приходим к следующему соотношению:

$$\begin{aligned} A_n &= -A_0 e^{ikz} \left[\exp(i\pi n f/z) - \exp(i\pi(n-1)f/z) \right] = \\ &= -A_0 e^{ikz} \exp(i\pi n f/z) \left[1 - \exp(-i\pi f/z) \right]. \end{aligned}$$

Пусть теперь отверстие перекрыто зонной пластинкой, у которой открыты все нечётные зоны, а все чётные перекрыты, причём всего открыто

N зон. Тогда число n пробегает значения $n = 2m - 1$, $m = 1, 2, \dots, N$. Суммарное поле, производимое зонной пластинкой, равно

$$A = \sum_{m=1}^N A_{2m-1} = A_0 [1 - \exp(-i\pi f/z)] \sum_{m=1}^N \exp(i\pi(2m-1)f/z) =$$

$$= A_0 [\exp(i\pi f/z) - 1] \sum_{m=0}^{N-1} \exp(i2\pi mf/z).$$

Суммируя геометрическую прогрессию, получаем

$$A(z) = A_0 [\exp(i\pi f/z) - 1] \frac{\exp(i2\pi Nf/z) - 1}{\exp(i2\pi f/z) - 1}.$$

Наконец, находим интенсивность излучения $I(z) = |A(z)|^2$:

$$I(z) = I_0 \frac{\sin^2(\pi Nf/z)}{\cos^2(\pi f/2z)}.$$

Наибольших значений функция $I(z)$ достигает в точках z , для которых знаменатель обращается в нуль:

$$\frac{\pi f}{2z_m} = \frac{\pi}{2}(2m+1) \Rightarrow z_m = \frac{f}{2m+1}, \quad m = 0, 1, 2, \dots$$

В этих точках интенсивность равна

$$I_m = N^2 I_0.$$

Таким образом, зонная пластинка имеет главный фокус $z = f$ и серию дополнительных фокусов z_m , причём во всех этих точках интенсивность практически одинакова. Последний результат справедлив, если рассматриваются точки, не слишком близкие к экрану. Иначе начинают играть роль ограничения, рассмотренные в конце раздела 6.2.5.

В области $z > f$ также имеются локальные максимумы, интенсивность которых мала по сравнению с интенсивностью в фокусе.

В некоторых точках интенсивность обращается в нуль. В этих точках числитель равен нулю, а знаменатель отличен от нуля: $\pi Nf/z = p\pi$, или

$$z = \frac{Nf}{p}, \quad p = 0, 1, 2, \dots, \quad p \neq (2m+1)N, \quad m = 0, 1, 2, \dots$$

(значениям $p = (2m+1)N$ отвечают основной и дополнительные фокусы).

В области $z > f$ на конечном расстоянии имеется $N-1$ точек, в которых $I(z_p) = 0$. Между этими точками также имеются локальные максимумы малой амплитуды.

На большом расстоянии от зонной пластинки $z \gg \pi N f$ интенсивность монотонно убывает по закону

$$I(z) = I_0 \left(\frac{\pi N f}{z} \right)^2 \sim z^{-2}.$$

Поскольку $f = R^2 / 2\lambda N$, то неравенство $z \gg \pi N f$ принимает вид $z \gg R^2 / \lambda$.

На рис. 6.2.10 показана типичная зависимость интенсивности излучения от расстояния до зонной пластинки, отражающая отмеченные выше закономерности.

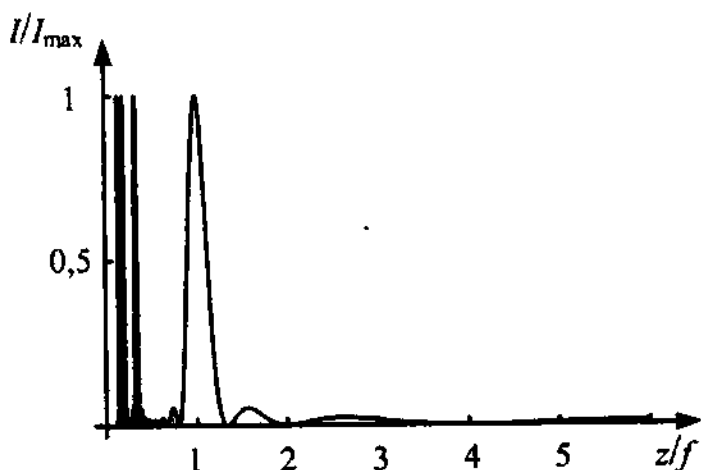


Рис. 6.2.10. Зависимость интенсивности излучения от расстояния до зонной пластинки. Число открытых зон для точки наблюдения в фокусе пластинки $N = 4$, $I_{\max} = N^2 I_0$

Проиллюстрируем полученные результаты с помощью спирали Френеля.

Пусть вначале точка наблюдения находилась на расстоянии f от экрана с отверстием, причём в отверстие помещалось ровно $2N$ зон Френеля. Перекроем отверстие зонной пластинкой, у которой открыты все нечётные зоны (тогда зона с номером $2N$ закрыта). Требуется найти поле в точке, находящейся на расстоянии $2f$ от отверстия.

Если на отверстие падает плоская волна с амплитудой A_0 и интенсивностью $I_0 = |A_0|^2$, то в фокусе пластинки ($z = f$) амплитуда равна

$$A_f = \frac{N}{2} \cdot 2A_0 = NA_0, \text{ а интенсивность } I_f = |A_f|^2 = N^2 I_0.$$

С расстояния $z = 2f$ в отверстие попадает число новых зон, равное N , поскольку площадь зон пропорциональна расстоянию до экрана: $S = \pi \lambda z$. Это значит, что каждая новая зона перекрывает две старые. Поскольку положение открытых и закрытых участков на экране не меняется, то у новых зон открытыми остаются только внешние половины (рис. 6.2.12, слева).

Пользуясь спиралью Френеля (рис. 6.2.11, справа), замечаем, что поля, создаваемые парой соседних зон Френеля ($A_1, -A_1$), равны по величине.

не: $|A_1| = \sqrt{2}|A_0|$ и противоположны по знаку. Поэтому в случае, когда в зонной пластинке имеется чётное число прозрачных участков, суммарное поле в точке $z = 2f$ обращается в нуль. Если же число открытых участков нечётное, то интенсивность окажется равной $I_1 = |A_1|^2 = 2I_0$.

Аналогично можно найти интенсивность и в других точках на оси z .

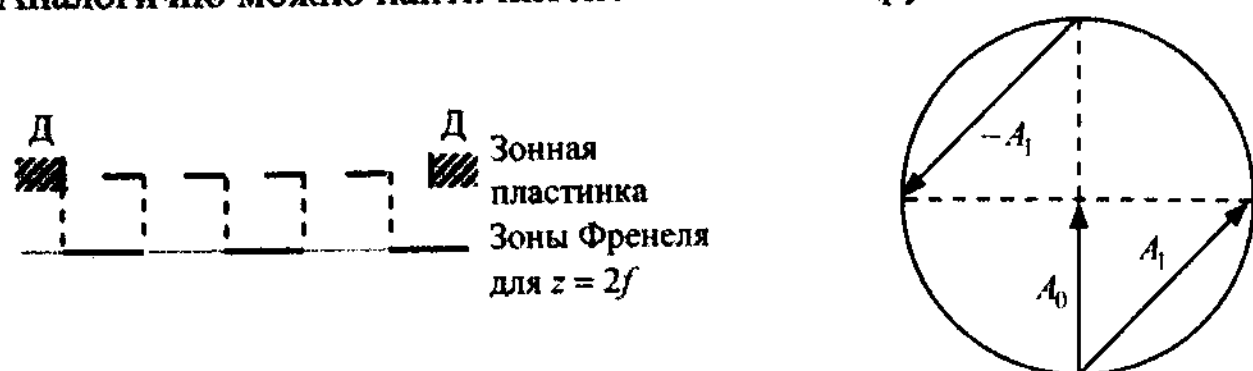


Рис. 6.2.11. Слева — зонная пластинка, помещённая в диафрагму Д (сверху) и зоны Френеля для точки Р на расстоянии $z = f$ от пластинки (снизу); справа — векторы полей, создаваемых внутренними половинами соседних зон Френеля (для точки Р)

6.2.8. Дифракция на круглом диске

Выше мы рассматривали дифракцию света на отверстиях в непрозрачном экране. Если на пути излучения источника поставлен непрозрачный диск (выше), а остальная часть волнового фронта открыта, то в области геометрической тени может наблюдаться светлое пятно (рис. 6.2.12). Его существование предсказал Симеон Дени Пуассон в 1818 г. на основании теории Френеля. Доминик Франсуа Жан Араго поставил эксперимент и подтвердил выводы теории. По имени первооткрывателей — двух французских физиков — явление получило название «*пятно Араго–Пуассона*».

Это открытие подтвердило наличие у света волновых свойств. Его следствием оказалось то, что корпускулярная теория света, выдвинутая Ньютоном в 1672 г., была предана забвению почти на столетие, главной оставалась волновая теория, заложенная в опубликованном Х. Гюйгенсом в 1678 г. «Трактате о свете». Лишь в начале XX века под давлением экспериментальных фактов идея «корпускул» — квантов света — вернулась в физику.

Отметим, что светлое пятнышко в центре тени в 1715 г. (за 100 лет до работы Френеля) наблюдал французский астроном и картограф Жозеф-Николя Делиль, но его наблюдения не привлекли внимания, так как они не были связаны с какой-либо теорией.

Описанное явление просто объясняется на основе представлений о зонах Френеля. Для этого случая построение зон удобно несколько видоизменить, чтобы исключить перекрытую область волнового фронта (рис. 6.2.13). Выберем новое начало отсчёта фазы зон — от внутреннего края первой открытой зоны. В этом представлении вектор амплитуды открытой части волнового фронта приблизительно равен A_0 и идёт от точки спирали к центру, как и на рис. 6.2.9 (штриховая линия).

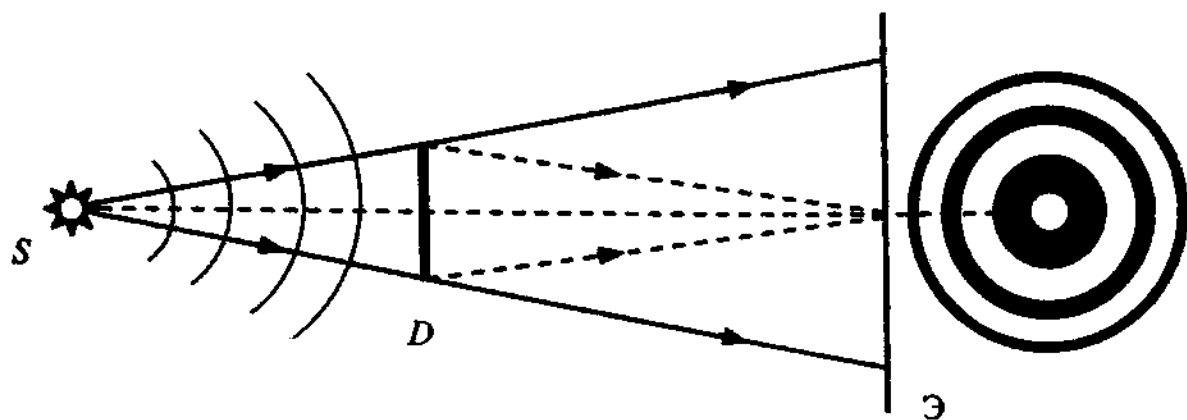


Рис. 6.2.12. Точечный источник посылает световую волну на круглый непрозрачный диск D , а на экране $\mathcal{E}$ в центре наблюдается светлое пятно Араго–Пуассона

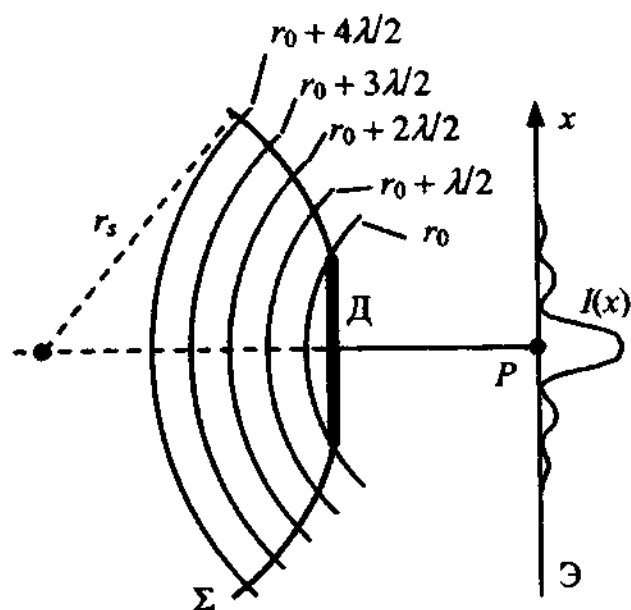


Рис. 6.2.13. Разбиение волнового фронта Σ , частично перекрытого непрозрачным диском D , на зоны Френеля для точки наблюдения P

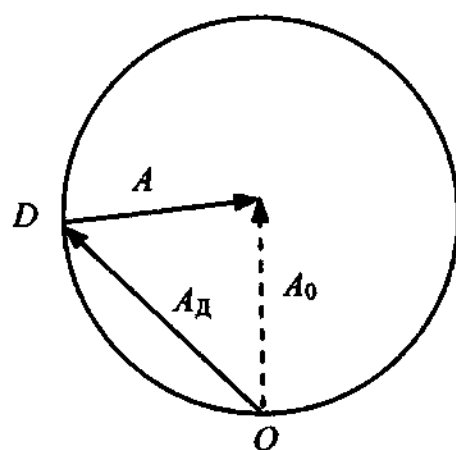


Рис. 6.2.14. Спираль Френеля для случая, когда диском перекрыта часть волнового фронта. Перекрытым зонам соответствует вектор OD . Вектор A изображает амплитуду результирующего поля

Анализ дифракции на диске можно также выполнить, используя прежнее построение зон. Дуга спирали OD отвечает перекрытой части волнового фронта, а её длина пропорциональна площади диска. Если бы диска не было, то поле равнялось бы A_0 . Но это поле можно представить как сумму полей, идущих от области диска (A_D) и от всей остальной части волнового фронта (A): $A_0 = A_D + A$. Поэтому, если перекрыть часть A_D , то поле в точке наблюдения окажется равным $A = A_0 - A_D$. Этот вектор и показан на диаграмме 6.2.14. Другими словами, в этом подходе меняется положение начала вектора амплитуды, причём результирующее поле представляется вектором A вместо A_0 (с соответственно изменённой фа-

зой). Существенно, что независимо от размера диска в центре его геометрической тени светлое пятно должно наблюдаться.

Имеются, однако, ограничения на возможность наблюдения пятна Араго-Пуассона: в частности, неровность края не должна превышать одной зоны Френеля. В противном случае в точку наблюдения приходят волны от фрагментов соседних зон, подавляющих интерференционный максимум. Поскольку

$$\Delta R = R_{m+1} - R_m = \sqrt{\lambda z} (\sqrt{m+1} - \sqrt{m}) \approx \frac{1}{2} \sqrt{\frac{\lambda z}{m}},$$

то, учитывая, что $R \approx \sqrt{m\lambda z} \Rightarrow m = R^2/\lambda z$, находим искомое ограничение на неровность края диска:

$$\Delta R < \lambda z/2R.$$

6.3. Дифракция на краю экрана и на щели

6.3.1. Зоны Шустера

Рассмотрим дифракцию плоской волны на краю экрана Э, перекрывающего половину волнового фронта (рис. 6.3.1). Точка наблюдения P расположена на экране Э₁, отстоящем на расстояние h от экрана Э. Край экрана Э считаем прямым, имеющим большую длину $L \gg h$. Найдём поле в точке P , находящейся на расстоянии x от проекции края экрана Э: при $x < 0$ точка находится в области геометрической тени, а при $x > 0$ — в области открытой части волнового фронта.

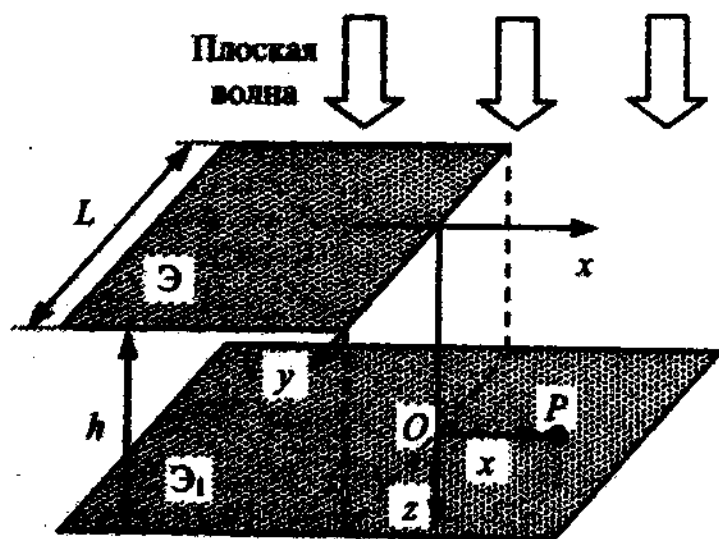


Рис. 6.3.1. Плоская волна падает по нормали на экран, перекрывающий половину волнового фронта

Начнём рассмотрение с построения, аналогичного построению зон Френеля. Для этого из точки P , как из центра, проведём серию цилиндрических поверхностей с осью, параллельной краю экрана Э (рис. 6.3.2), и радиусами

$$r_n = h + n\lambda/2, \quad n = 0, 1, 2, 3, \dots \quad (6.3.1)$$

Эти поверхности разбивают весь волновой фронт на полосы, параллельные краю, называемые зонами Шустера (по имени английского физика Артура Шустера). Значениям $n = 1, 2, 3, \dots$ отвечают внешние границы соответственно первой, второй, третьей и так далее зон.

Зоны могут находиться как справа, так и слева от точки P . На рис. 6.3.2 внешние границы зон справа обозначены как 1, 2, 3, ..., а зон слева — как 1', 2', 3' При $x > 0$ открыты все зоны справа и конечное число зон слева от точки P , а при $x < 0$ открыты только зоны справа, начиная с некоторого номера $n \geq 1$.

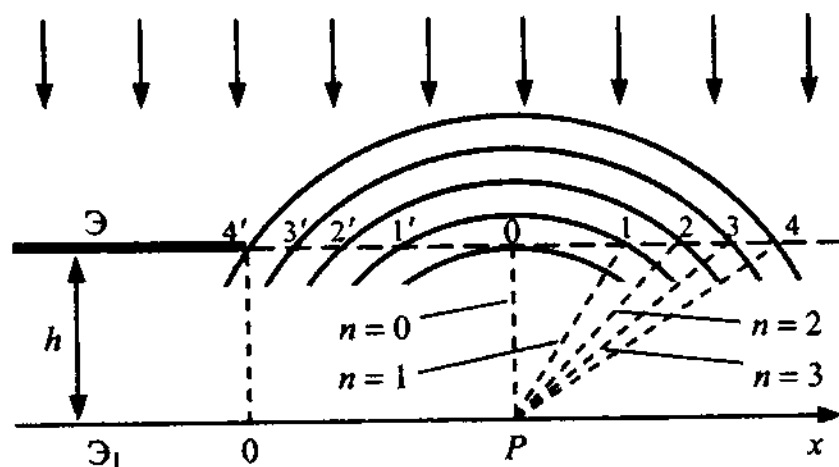


Рис. 6.3.2. К расчёту дифракции на краю экрана. Построение зон Шустера

Найдём ширину зон.

Ширина первой зоны найдётся из треугольника $01P$:

$$d_1 = \sqrt{(h + \lambda/2)^2 - h^2} = \sqrt{h\lambda + \lambda^2/4} \approx \sqrt{h\lambda}. \quad (6.3.2)$$

Здесь учтено, что $\lambda \ll h$. Суммарная ширина n зон равна

$$d_1 + d_2 + \dots + d_n = \sqrt{(h + n\lambda/2)^2 - h^2} \approx \sqrt{nh\lambda}. \quad (6.3.3)$$

Как и выше, мы рассматриваем только зоны, для которых $n\lambda \ll h$. Отсюда находим ширину n -й зоны:

$$d_n = \sqrt{nh\lambda} - \sqrt{(n-1)h\lambda} = d_1 (\sqrt{n} - \sqrt{n-1}), \quad d_1 = \sqrt{h\lambda}. \quad (6.3.4)$$

Если длина края экрана (вдоль щели) равна L , то площади зон $S_n = Ld_n$ относятся как

$$S_1 : S_2 : S_3 : S_4 \dots \approx 1 : 0,414 : 0,318 : 0,268 \dots \quad (6.3.5)$$

Таким образом, площади зон убывают с ростом номера намного быстрее, чем это имело место в случае зон Френеля. При больших номерах ($n \gg 1$) закон убывания нетрудно найти из (6.3.4):

$$S_n = S_1 (\sqrt{n} - \sqrt{n-1}) \approx S_1 \frac{1}{2\sqrt{n}}. \quad (6.3.6)$$

6.3.2. Расчёт дифракции с помощью интеграла Френеля

Пусть на экран Э падает плоская волна с амплитудой A_0 . Считая экран Э неограниченным вдоль оси y , обозначим координаты точки P в плоскости наблюдения Э₁ как $x_P = x$, $y_P = 0$. Выделим на открытой части волнового фронта участок площадью $d\sigma = dx_1 dy_1$. Тогда расстояние от него до точки P равно

$$r = \sqrt{h^2 + (x - x_1)^2 + y_1^2}. \quad (6.3.7)$$

Запишем интеграл Френеля:

$$A(P) \equiv A(x) = \int_0^\infty dx_1 \int_{-\infty}^\infty dy_1 K(\alpha) \frac{1}{r} e^{ikr}. \quad (6.3.8)$$

Поскольку размер зон убывает с ростом номера, основной вклад в поле вносит область фронта, видимая под малыми углами к нормали: $K(\alpha) \approx K(0) = k/2\pi i$. Будем также предполагать, что рассматривается область вблизи границы экрана: $|x| \ll h$. В этих предположениях можно произвести разложение в (6.3.7) по степеням координат x , x_1 , y_1 :

$$r \approx h + \frac{(x - x_1)^2 + y_1^2}{2h}. \quad (6.3.9)$$

Соответственно интеграл Френеля принимает вид

$$\begin{aligned} A(x) &= \frac{k}{2\pi i h} e^{ikh} \int_0^\infty dx_1 \int_{-\infty}^\infty dy_1 \exp\left(ik \frac{(x - x_1)^2 + y_1^2}{2h}\right) = \\ &= \frac{k}{2\pi i h} e^{ikh} \int_0^\infty \exp\left(ik \frac{(x - x_1)^2}{2h}\right) dx_1 \int_{-\infty}^\infty \exp\left(ik \frac{y_1^2}{2h}\right) dy_1. \end{aligned} \quad (6.3.10)$$

Сюда входит, во-первых, интеграл

$$\int_{-\infty}^\infty \exp(i\alpha y_1^2) dy_1 = \sqrt{i\pi/\alpha}, \quad \alpha = k/2h, \quad (6.3.11)$$

аналогичный интегралу Пуассона

$$\int_{-\infty}^\infty e^{-\alpha x^2} dx = \sqrt{\pi/\alpha}.$$

Второй интеграл в (6.3.10) имеет вид

$$\int_0^{\infty} \exp\left(ik \frac{(x-x_1)^2}{2h}\right) dx_1 = \frac{1}{\sqrt{\alpha}} \int_{-\sqrt{\alpha x}}^{\infty} \exp(i\tau^2) d\tau. \quad (6.3.12)$$

Запишем промежуточный результат:

$$A(x) = A_0 e^{ikh} \frac{1}{\sqrt{\pi i}} \int_{-\sqrt{\alpha x}}^{\infty} \exp(i\tau^2) d\tau. \quad (6.3.13)$$

Приведём данный результат к традиционной форме. Введём *интегралы Френеля*:

$$S(u) = \sqrt{\frac{2}{\pi}} \int_0^u \sin(\tau^2) d\tau, \quad C(u) = \sqrt{\frac{2}{\pi}} \int_0^u \cos(\tau^2) d\tau. \quad (6.3.14)$$

Эти интегралы обладают следующими свойствами:

$$\begin{aligned} S(-u) &= -S(u), \quad C(-u) = -C(u); \\ S(+\infty) &= 1/2, \quad C(+\infty) = 1/2. \end{aligned} \quad (6.3.15)$$

Графики данных функций показаны на рис. 6.3.3.

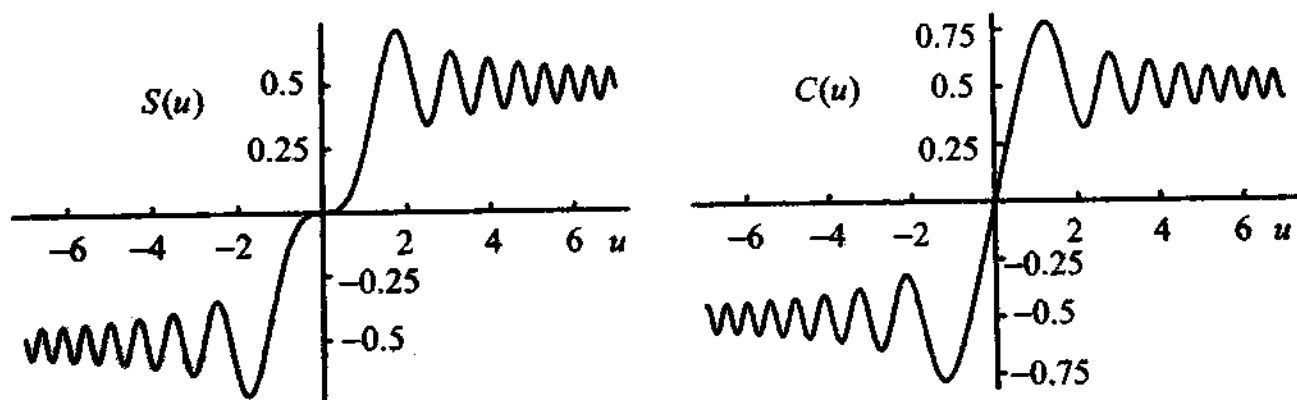


Рис. 6.3.3. Графики интегралов Френеля $S(u)$ (слева) и $C(u)$ (справа)

Поскольку

$$\begin{aligned} \int_u^{\infty} e^{i\tau^2} d\tau &= \int_0^{\infty} e^{i\tau^2} d\tau - \int_0^u e^{i\tau^2} d\tau = \frac{1}{2} \sqrt{\pi i} - \sqrt{\frac{\pi}{2}} [C(u) + iS(u)] = \\ &= \sqrt{\frac{\pi}{2}} \left\{ \left[\frac{1}{2} - C(u) \right] + i \left[\frac{1}{2} - S(u) \right] \right\}, \end{aligned} \quad (6.3.16)$$

где учтено, что $\sqrt{i} = (1+i)/\sqrt{2}$, то выражение (6.3.13) можно переписать в виде

$$A(x) = A_0 e^{ikh} \frac{1}{\sqrt{2i}} \left\{ \left[\frac{1}{2} + C(x\sqrt{k/2h}) \right] + i \left[\frac{1}{2} + S(x\sqrt{k/2h}) \right] \right\}. \quad (6.3.17)$$

Для интенсивности $I = |A|^2$ получаем

$$I = \frac{1}{2} I_0 \left\{ \left[\frac{1}{2} + C(x\sqrt{k/2h}) \right]^2 + \left[\frac{1}{2} + S(x\sqrt{k/2h}) \right]^2 \right\}. \quad (6.3.18)$$

Найденная с помощью этой формулы зависимость интенсивности излучения от положения точки P показана на рис. 6.3.4. Видно, что в области геометрической тени ($x < 0$) интенсивность монотонно возрастает по мере приближения к границе тени, достигая на самой границе значения $I = I_0/4$. В области $x > 0$ интенсивность совершает колебания с убывающей амплитудой, причём

$$I \rightarrow I_0 \text{ при } x \rightarrow +\infty. \quad (6.3.19)$$

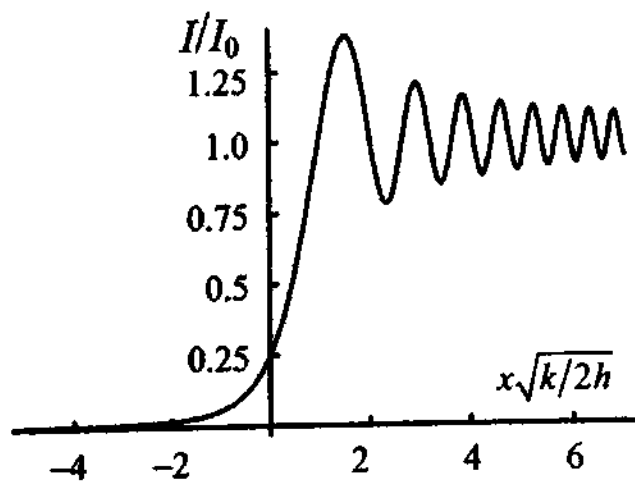


Рис. 6.3.4. График зависимости интенсивности излучения от координаты x точки наблюдения относительно края

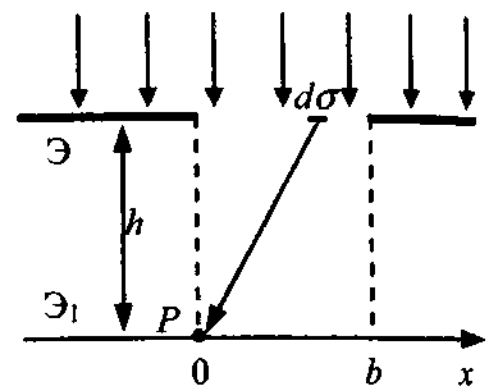


Рис. 6.3.5. Дифракция на щели шириной b в непрозрачном экране Э. Рассматривается поле на экране Э₁ под краем щели

6.3.3. Спираль Корню

Исследуем подробнее, как меняется поле в точке наблюдения в зависимости от того, какая часть волнового фронта открыта.

Пусть точка P находится точно под краем: $x = 0$ (рис. 6.3.5). Если при этом ширина открытой части волнового фронта справа от P конечна и равна b , то поле вместо (6.3.13) будет даваться формулой

$$\begin{aligned} A_1(b) &= A_0 e^{ikh} \frac{1}{\sqrt{\pi i}} \int_0^{b\sqrt{\alpha}} \exp(i\xi^2) d\xi = \\ &= \frac{1}{\sqrt{2i}} A_0 e^{ikh} \left[C(b\sqrt{\alpha}) + iS(b\sqrt{\alpha}) \right], \quad \alpha = \frac{k}{2h}. \end{aligned} \quad (6.3.20)$$

Соответственно для интенсивности волны получаем выражение

$$I_1(b) = \frac{1}{2} I_0 \left[C^2(b\sqrt{\alpha}) + S^2(b\sqrt{\alpha}) \right], \quad \alpha = \frac{k}{2h}. \quad (6.3.21)$$

График функции $I(b)$ приведён на рис. 6.3.6. Видно, что по мере открытия волнового фронта интенсивность совершает колебания с уменьшающейся амплитудой, причём $I \rightarrow I_0/4$ при $b \rightarrow \infty$.

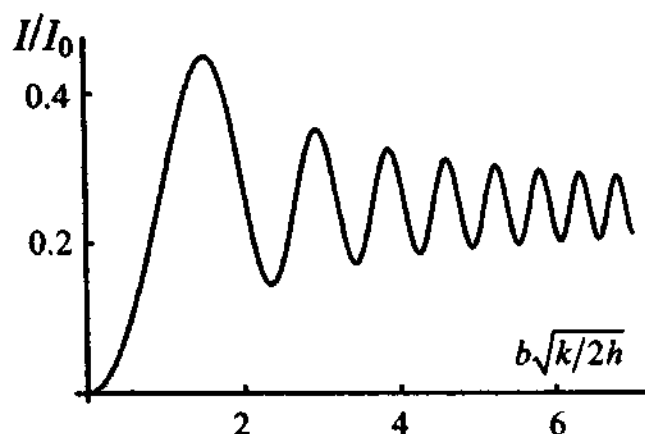


Рис. 6.3.6. График зависимости интенсивности излучения от размера b открытого справа волнового фронта

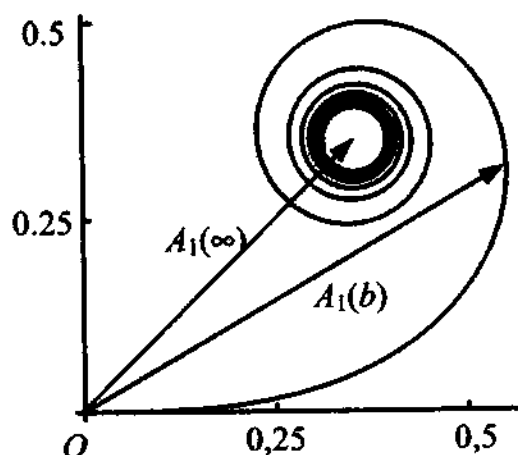


Рис. 6.3.7. Спираль Корню для случая, когда волновой фронт открывается справа от точки наблюдения

Проведённым расчётам можно дать наглядную геометрическую интерпретацию на комплексной плоскости. Введём вектор

$$A_1(b) = \frac{1}{\sqrt{2}} A_0 \left[C(b\sqrt{\alpha}) + iS(b\sqrt{\alpha}) \right], \quad (6.3.22)$$

где в отличие от (6.3.20) отброшены несущественные фазовые множители. Иначе говоря, мы выбрали начало отсчёта фазы волны от значения амплитуды при $b \rightarrow 0$. Этот вектор по мере изменения ширины b открытой части волнового фронта поворачивается против часовой стрелки, причём его конец описывает кривую — *спираль Корню* (рис. 6.3.7). Рассматриваемому случаю отвечает кривая справа от точки O . Вектор поля при ширине щели b представляется вектором $A_1(b)$. Когда $b \rightarrow +\infty$, спираль закручивается в точку с координатами $(1/\sqrt{8}, 1/\sqrt{8}) \approx (0.35, 0.35)$, которой соответствует интенсивность

$$I = \left(A_0/\sqrt{8} \right)^2 + \left(A_0/\sqrt{8} \right)^2 = I_0/4 \quad (A_1 = A_1(\infty)). \quad (6.3.23)$$

Наоборот, при $b \rightarrow -\infty$ спираль закручивается к точке с координатами

$$\left(-1/\sqrt{8}, -1/\sqrt{8} \right) \approx (-0.35, -0.35) \quad (6.3.24)$$

и той же интенсивности.

Аналогично можно рассмотреть случай, когда открывается только левая половина волнового фронта относительно точки наблюдения. Формально этому отвечают значения $b < 0$. На рис. 6.3.8 данный случай представлен спиралью слева от точки O , закручивающейся к центру (6.3.24).

Если волновой фронт попадает в щель с координатами краёв b_1 и b_2 , $b_1 < b_2$ относительно точки наблюдения ($x = 0$), то поле будет даваться формулой

$$A(b_1, b_2) = A_1(b_2) - A_1(b_1). \quad (6.3.25)$$

На рис. 6.3.9 результирующее поле представляется вектором $A(b_1, b_2)$, соединяющим точки спирали, отвечающие границам открытого фронта. Если открыт весь волновой фронт, то вектор поля идёт от центра левой спирали к центру правой и его длина равна $I = |2A_1(\infty)|^2 = I_0$.

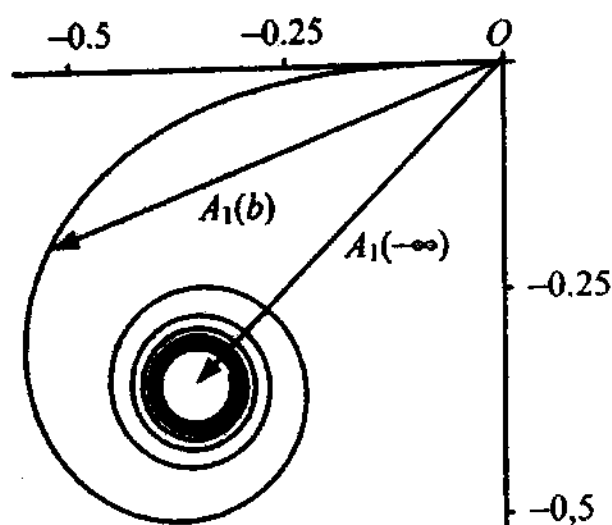


Рис. 6.3.8. Спираль Корню для случая, когда волновой фронт открывается слева от точки наблюдения. Здесь $-b > 0$ — ширина открытой области

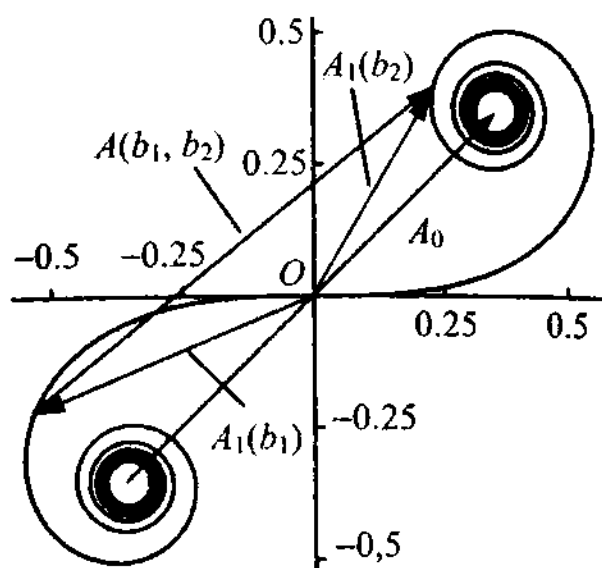
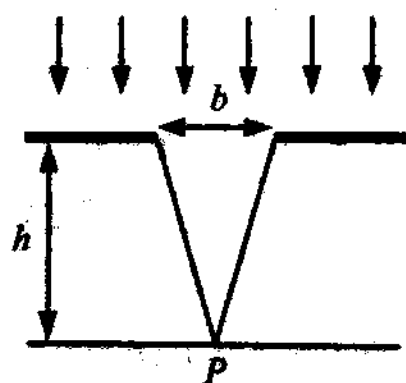


Рис. 6.3.9. Спираль Корню

6.4. Дифракция Фраунгофера

6.4.1. Типы дифракции

Пусть излучение падает на непрозрачный экран с отверстием (щелью) размером b (рис. 6.4.1). Поскольку граница n -й зоны Френеля (Шустера)



даётся формулой $r_n = \sqrt{n\lambda h}$, то по отношению к точке наблюдения P в отверстии укладывается число зон

$$n \sim b^2 / \lambda h. \quad (6.4.1)$$

Соответственно говорят, что при

$$n \ll 1 \quad (6.4.2)$$

имеет место дифракция Фраунгофера, при

$$n \sim 1 \quad (6.4.3)$$

реализуется дифракция Френеля, а при

$$n \gg 1 \quad (6.4.4)$$

Рис. 6.4.1. К определению типа дифракции на отверстии размером b

геометрическая оптика. Последний случай связан с тем, что в этом пределе открыта значительная часть волнового фронта, и роль краёв препятствия незначительна — точка наблюдения находится близко к экрану.

Величина $f = b^2/\lambda h$ называется *параметром* (или *числом*) *Френеля*. Используют также *волновой параметр*, определяемый формулой

$$p = 1/\sqrt{f} = \sqrt{\lambda h}/b. \quad (6.4.5)$$

Условие дифракции Фраунгофера можно записать в виде

$$\frac{b}{h} \ll \frac{\lambda}{b} \ll 1.$$

Это значит, что из точки наблюдения отверстие видно под малым углом: $b/h \ll 1$, а радиус когерентности значительно превышает размер отверстия: $\rho_{\text{ког}} = \lambda h/b \gg b$.

Согласно сказанному условие $f \sim 1$ означает, что в отверстие попадает около одной зоны Френеля. Обратно, расстояние, при котором отверстие имеет размер порядка одной зоны Френеля, есть

$$z_d = b^2/\lambda. \quad (6.4.6)$$

Введённая этим соотношением величина называется *дифракционной длиной светового пучка*. Часто эту величину определяют равенством $z_d = kb^2$, отличающимся множителем 2π от (6.4.6). Таким образом, дифракция Фраунгофера наблюдается на расстояниях, превышающих дифракционную длину пучка.

6.4.2. Дифракция Фраунгофера на щели

Пусть плоская волна $A = A_0 e^{ikz}$ падает по нормали на экран с щелью шириной b . Дифракцию Фраунгофера можно наблюдать в дальней зоне ($z \gg b^2/\lambda$) в виде серии полос, параллельных краю щели. Если же на пути дифрагированной волны поставить линзу, то в её фокальной плоскости возникнет серия точек (пятен), отвечающих дифракционным максимумам (рис. 6.4.2).

Согласно принципу Гюйгенса–Френеля плоский волновой фронт за щелью является источником вторичных волн. В рассматриваемом случае эти волны цилиндрические с осью вдоль отдельных полос.

Для нахождения результирующего поля вдали от щели разобьём волновой фронт на полоски шириной dx . Нужно найти сумму всех волн, идущих в определённом направлении θ (рис. 6.4.3).

Разность хода волн от полоски с координатой x и от начальной (центральной) полоски (с координатой $x = 0$) равна

$$\Delta = x \sin \theta. \quad (6.4.7)$$

Если амплитуда падающей на щель волны равна A_0 , то амплитуда волны, посылаемой полоской dx в направлении θ , есть

$$dA \sim A_0 e^{ik\Delta} \frac{dx}{b} = A_0 e^{ikx \sin \theta} \frac{dx}{b}. \quad (6.4.8)$$

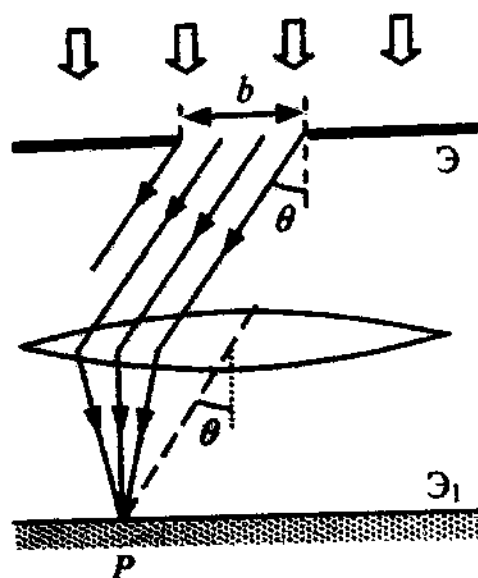


Рис. 6.4.2. Схема наблюдения дифракции Фраунгофера. Свет от щели падает на линзу. Лучи, имеющие одно и то же направление, собираются в точку в фокальной плоскости линзы (на экране $\mathcal{E}_1$)

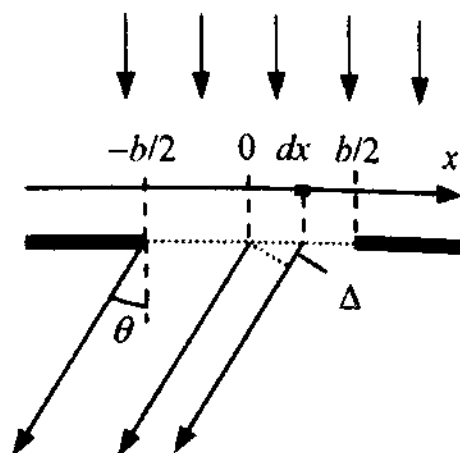


Рис. 6.4.3. Дифракция Фраунгофера на длинной щели шириной b по оси x , вытянутой вдоль оси y

Введём обозначение

$$u = k \sin \theta. \quad (6.4.9)$$

Суммируя вклады всех участков волнового фронта, найдём

$$A(\theta) = \int dA \sim \frac{A_0}{b} \int_{-b/2}^{b/2} e^{iux} dx = A_0 \frac{\sin(ub/2)}{ub/2}. \quad (6.4.10)$$

Отсюда следует выражение для интенсивности:

$$I(\theta) = I_m \left(\frac{\sin ub/2}{ub/2} \right)^2, \quad u = k \sin \theta. \quad (6.4.11)$$

Здесь $I_m = I(0)$ — интенсивность излучения в направлении падающей волны ($\theta = 0$). Зависимость (6.4.11) показана на рис. 6.4.4.

Как видно из рис. 6.4.4, имеется главный максимум, в котором сосредоточена основная часть энергии излучения. Угловой радиус θ_d этого максимума определяется из уравнения $I(\theta_d) = 0$, или

$$\frac{1}{2} kb \sin \theta_d = \pi \Rightarrow \sin \theta_d = \frac{\lambda}{b}. \quad (6.4.12)$$

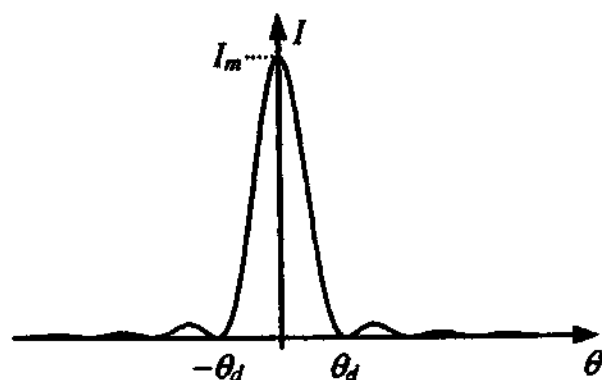


Рис. 6.4.4. Зависимость интенсивности излучения от направления при дифракции Фраунгофера на щели

Если выполняется условие $b \gg \lambda$, то для ширины максимума получаем выражение

$$2\theta_d = 2\lambda/b. \quad (6.4.13)$$

Из (6.4.12) следует, что условие первого дифракционного минимума

$$b \sin \theta = \lambda \quad (6.4.14)$$

означает, что *разность хода лучей от краёв щели $\Delta = b \sin \theta$ равна длине волны излучения.*

Мы получили картину дифракции на щели. Если бы мы рассматривали дифракцию на прямоугольном отверстии с размерами

$$-a/2 \leq x \leq a/2, \quad -b/2 \leq y \leq b/2, \quad (6.4.15)$$

то для интенсивности нашли бы следующую зависимость:

$$I(\theta) = I_m \left(\frac{\sin(u_x a/2)}{u_x a/2} \right)^2 \left(\frac{\sin(u_y b/2)}{u_y b/2} \right)^2, \quad (6.4.16)$$

$$u_x = k \sin \theta_x, \quad u_y = k \sin \theta_y.$$

Здесь θ_x и θ_y — углы, образуемые лучом с осями x и y .

6.4.3. Дифракция Фраунгофера на круглом отверстии

Во многих приложениях рассматривается дифракция на круглом отверстии. Если отверстие имеет радиус R , то решение задачи о дифракции Фраунгофера приводит к следующей зависимости интенсивности от угла:

$$I(\theta) = I_m \left(\frac{2J_1(qR)}{qR} \right)^2, \quad q = k \sin \theta, \quad I_m = I_0 \left(\frac{\pi R^2}{\lambda z} \right)^2. \quad (6.4.17)$$

Здесь $J_1(x)$ — функция Бесселя порядка 1. Она обращается в нуль при

$$x_0 = 0, \quad x_1 = 3,8317, \quad x_2 = 7,0156, \quad x_3 = 10,1735 \dots \quad (6.4.18)$$

Значение коэффициента I_m в (6.4.17), приведено в (6.2.30).

Качественный вид зависимости $I(\theta)$ в (6.4.17) аналогичен показанному на рис. 6.4.4. Угловой радиус главного максимума определится из условия $qR = x_1$, или

$$\theta = \theta_d = 0,61 \frac{\lambda}{R} = 1,22 \frac{\lambda}{D} \quad (6.4.19)$$

($D = 2R$ — диаметр отверстия). Соответственно угловой диаметр расходящегося пучка равен

$$2\theta_d = 2,44 \frac{\lambda}{D}. \quad (6.4.20)$$

Дифракционный пик $0 \leq \theta \leq \theta_d$ называют *кружком Эйри*.

Фотография картины дифракции лазерного излучения ($\lambda = 650$ нм) на отверстии ($d = 0,2$ мм) показана на рис. 6.4.5.

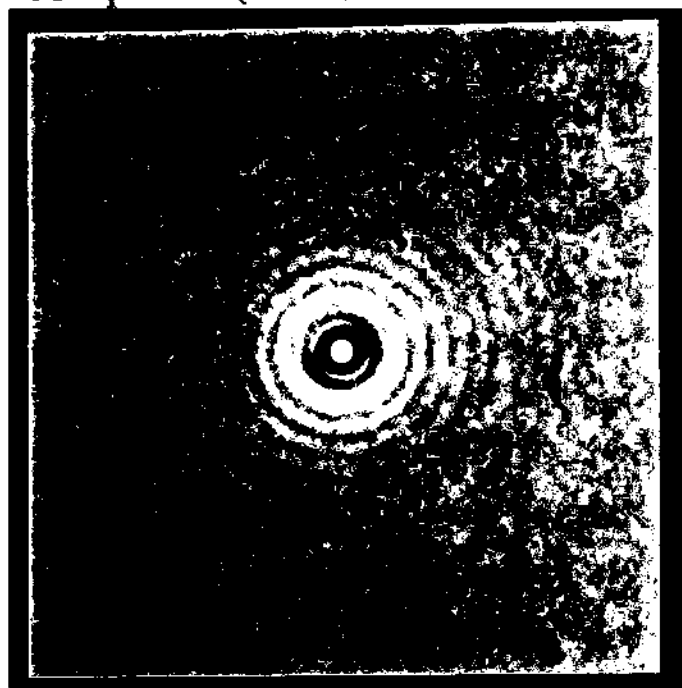


Рис. 6.4.5. Дифракция лазерного излучения с длиной волны 650 нм на отверстии диаметром 0,2 мм

Дополнение

Функции Бесселя 1-го рода $y = J_\nu(x)$ порядка ν являются решениями дифференциального уравнения Бесселя

$$x^2 y'' + x y' + (x^2 - \nu^2) y = 0 \quad (6.4.21)$$

и могут быть представлены рядом

$$J_\nu(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k! \Gamma(\nu + k + 1)} \left(\frac{x}{2}\right)^{\nu+2k}. \quad (6.4.22)$$

Второе линейно независимое решение уравнения Бесселя называется функцией Неймана и определяется равенством

$$N_\nu(x) = \frac{1}{\sin \nu \pi} [J_\nu(x) \cos \nu \pi - J_{-\nu}(x)].$$

Если индекс ν — целое число: $\nu = n$, то

$$N_n(x) = \lim_{\nu \rightarrow n} N_\nu(x).$$

Для функций Бесселя с целочисленными значениями индекса $\nu = n$ имеет место интегральное представление:

$$J_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(x \sin \varphi - n\varphi)} d\varphi. \quad (6.4.23)$$

Приведём асимптотику функций Бесселя и Неймана при $x \rightarrow \infty$:

$$J_\nu(x) \approx \sqrt{\frac{2}{\pi x}} \cos\left(x - \frac{\nu\pi}{2} - \frac{\pi}{4}\right), \quad N_\nu(x) \approx \sqrt{\frac{2}{\pi x}} \sin\left(x - \frac{\nu\pi}{2} - \frac{\pi}{4}\right). \quad (6.4.24)$$

На рис. 6.4.6 показаны графики функций $J_0(x)$ и $J_1(x)$, отражающие типичные свойства бесселевых функций с индексом $\nu \geq 0$.

Отметим также, что функции $J_0(x)$ и $J_1(x)$ связаны тождествами

$$J'_0(x) = -J_1(x), \quad \int_0^x \xi J_0(\xi) d\xi = x J_1(x). \quad (6.4.25)$$

Получим формулу (6.4.17). Поместим начало координат в центр отверстия. Точки отверстия будем нумеровать радиус-вектором ρ , а направление, в котором рассматривается дифрагированная волна, зададим единичным вектором $\mathbf{n}$. Разность хода лучей от площадки $d\sigma$ и центра отверстия определяется проекцией ρ на $\mathbf{n}$ и равна $\Delta(\rho) = \mathbf{n}\rho$ (это нетрудно увидеть из рис. 6.4.3). Зафиксируем направление наблюдения $\mathbf{n}$ (в силу цилиндрической симметрии задачи оно определяется лишь углом между вектором $\mathbf{n}$ и нормалью к плоскости отверстия). Выберем углы цилиндрической системы координат так, что

$$\rho_x = \rho \cos \varphi, \quad \rho_y = \rho \sin \varphi, \quad n_x = \sin \theta, \quad n_y = 0, \quad n_z = \cos \theta.$$

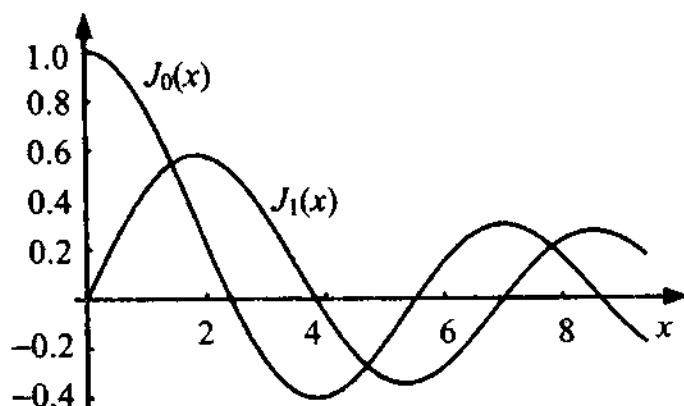


Рис. 6.4.6. Графики функций Бесселя $J_0(x)$ и $J_1(x)$

Элемент площади $d\sigma$ можно записать в цилиндрических координатах как $d\sigma = \rho d\rho d\varphi$. Величина $\Delta(\rho)$ в тех же координатах равна $\Delta(\rho) = \rho \sin \theta \cos \varphi$, а разность фаз — $k\Delta = q\rho \cos \varphi$. Здесь $q = k \sin \theta$, а θ — угол между нормалью к плоскости отверстия и направлением наблюдения (вектором $\mathbf{n}$). Суммируя вклады всех участков волнового фронта, получаем

$$\begin{aligned} A(\theta) &\sim \frac{A_0}{\pi R^2} \int e^{ik\Delta} d\sigma = \frac{A_0}{\pi R^2} \int_{\varphi=0}^{2\pi} d\varphi \int_{\rho=0}^R \exp(iq\rho \cos \varphi) \rho d\rho = \\ &= \frac{2A_0}{R^2} \int_{\rho=0}^R J_0(q\rho) \rho d\rho. \end{aligned} \quad (6.4.26)$$

Воспользуемся второй формулой в (6.4.25), которую запишем в виде

$$\int_0^R \rho J_0(q\rho) d\rho = \frac{R}{q} J_1(qR).$$

Отсюда находим

$$A(\theta) = A(0) \frac{2J_1(qR)}{qR}. \quad (6.4.27)$$

Переходя от амплитуды поля к интенсивности, приходим к формуле (6.4.16).

6.4.4. Геометрический смысл типов дифракции

Как уже было сказано в разделе 6.4.1, в зависимости от значения числа Френеля $f = b^2/\lambda h$ выделяются три характерных случая. Их классифи-

кации можно придать более наглядный смысл. Именно, учтём, что согласно (6.4.12) величина $\theta \sim \lambda/b$ есть угол дифракционной расходимости пучка излучения на отверстии (препятствии) размером $\sim b$. Соответственно на расстоянии h уширение пучка составит

$$d \sim \theta h \sim \lambda h/b. \quad (6.4.28)$$

Поскольку начальный размер пучка порядка b , то в зависимости от соотношения величин d и b выделяются три случая.

1. Пусть $d \gg b$ — дифракционное уширение пучка велико по сравнению с его начальным размером. Поскольку в этих условиях число Френеля $f = b^2/\lambda h \ll 1$, то реализуется дифракция Фраунгофера.
2. При $d \sim b$ оказывается $f \sim 1$ — это случай дифракции Френеля.
3. Если $d \ll b$, то $f \gg 1$. Это соответствует пределу геометрической оптики.

Сильное распыление пучка излучения в зоне дифракции Фраунгофера означает, что все детали изображения предмета сглаживаются, он создаёт дифракционную картину, как малый, пространственно когерентный объект. Наоборот, в зоне геометрической оптики роль дифракции мала, в изображении предмета достаточно чётко видны границы, разделяющие освещённые области и области геометрической тени.

6.4.5. Соотношение неопределённостей

Пусть на щель размером b падает плоская волна по нормали к отверстию (рис. 6.4.7). Волновой вектор волны направлен вдоль оси z :

$$\mathbf{k}^{(0)} = (k_x^{(0)}, k_z^{(0)}) = (0, k), \quad k = \omega/c. \quad (6.4.29)$$

После прохождения щели возникают волны иных направлений — волновой вектор приобретает компоненту k_x в поперечном направлении:

$$\mathbf{k} = (k_x, k_z), \quad k_x = k \sin \theta, \quad k_z = k \cos \theta. \quad (6.4.30)$$

Характерный диапазон значений k_x определяется углом дифракции θ , который даётся равенством (6.4.12):

$$\frac{1}{2} k b \sin \theta = \pi \Rightarrow k_x b = 2\pi. \quad (6.4.31)$$

В расходящемся пучке присутствуют все направления волнового вектора, локализованные главным образом в конусе $|\theta| \leq \theta_d$.

Смысл результата (6.4.31) состоит в следующем. Если мы попытаемся исследовать участок волнового фронта размером $\Delta x \sim b$, выделив его, например, диафрагмой, то это неизбежно повлечёт появление новых волн,

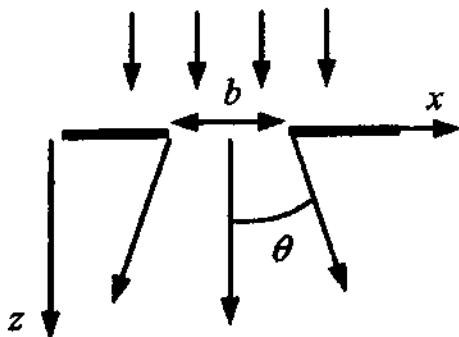


Рис. 6.4.7. К выводу соотношения неопределённостей

причём разброс их направлений, даваемый поперечной компонентой волнового вектора Δk_x , оказывается порядка $\Delta k_x \sim 2\pi/\Delta x$, или

$$\Delta k_x \Delta x \sim 2\pi. \quad (6.4.32)$$

Эта формула является общей для разнообразных явлений волновой природы и называется *соотношением неопределённостей*:

всякая попытка локализовать лучи приводит к их отклонению от первоначального направления в соответствии с соотношением (6.4.32).

6.5. Фокусирующее действие линзы

6.5.1. Расчёт поля в фокусе линзы методом зон Френеля

Пусть на идеальную линзу радиуса R падает плоская волна вдоль главной оптической оси. Если бы были применимы законы геометрической оптики, то все лучи, попадающие на линзу, собирались бы в её фокусе. Дифракция приводит к отклонению лучей и, как следствие, ограничивает интенсивность в фокусе. Оценим интенсивность излучения в фокусе.

Как известно, для линзы имеет место свойство таутохронности. Это значит, что оптические длины и, следовательно, фазы всех лучей, идущих от линзы к фокусу, одинаковы. Поэтому волновой фронт сразу после прохождения линзы оказывается сферическим.

Если бы на пути волны не было линзы, то мы могли бы осуществить разбиение волнового фронта на зоны Френеля и построить отвечающую этому разбиению спираль Френеля (рис. 6.5.1, слева).

Пусть спираль имеет радиус A_0 (отвечающий полностью открытому волновому фронту), а открыто m зон Френеля. В случае линзы радиуса R имеем $m = R^2/\lambda f$ (f — фокусное расстояние линзы).

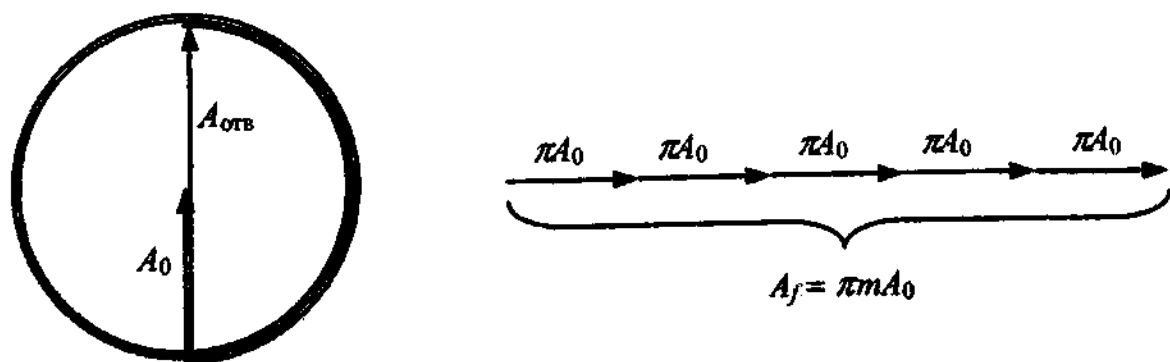


Рис. 6.5.1. Слева — спираль Френеля для отверстия с $m = 5$ открытыми зонами Френеля (соответствующее поле равно $A_{отв}$); справа — «распрямлённая» спираль Френеля длиной $A_f = mA_0$, представляющая поле в фокусе линзы

После установки линзы фазы волн, исходящих от всех частей прежних зон Френеля, оказываются уже одинаковыми. Это означает «распрямление» спирали, её превращение в прямую (рис. 6.5.1, справа): каждая зона

отображается в отрезок прямой длиной πA_0 . Соответственно полная амплитуда поля в фокусе составит $A_f = m \cdot \pi A_0$. Отсюда находим интенсивность:

$$I_f = |A_f|^2 = \pi^2 m^2 I_0 = \left(\pi R^2 / \lambda f \right)^2 I_0. \quad (6.5.1)$$

Оценим размер фокусного пятна r_f . Имея в виду, что всё излучение, попадающее на линзу, собирается в окрестности фокуса, запишем закон сохранения энергии:

$$\pi R^2 I_0 \approx \pi r_f^2 I_f \Rightarrow r_f \approx R \sqrt{\frac{I_0}{I_f}} = \frac{1}{\pi} \frac{\lambda}{R} f. \quad (6.5.2)$$

Отсюда видно, что размер пятна определяется дифракционной расходимостью на линзе, как препятствии размера $\sim R$.

6.5.2. Поле в фокальной плоскости линзы

Пусть на тонкую линзу падает плоская волна. Эту волну линза преобразует в волну со сферическим волновым фронтом (рис. 6.5.2). Будем считать, что фокус линзы находится в точке O , а радиус волнового фронта (фокусное расстояние линзы) — f . Найдём поле в точке наблюдения P , расположенной в фокальной плоскости линзы Σ .

Выберем на волновом фронте Σ площадку $d\sigma$. Согласно рис. 6.5.2 расстояние от элемента $d\sigma$ до точки P найдётся из равенства $r = f + \rho$, или

$$r^2 = f^2 + \rho^2 + 2f\rho. \quad (6.5.3)$$

Рассматривая малые отклонения от оптической оси: $\rho \ll f$, получаем

$$r \approx f + \frac{\rho^2 + 2f\rho}{2f} - \frac{(f\rho)^2}{2f^3} = f + \mathbf{n}\rho + \frac{1}{2f} [\rho^2 - (\mathbf{n}\rho)^2], \quad (6.5.4)$$

где мы ввели единичный вектор $\mathbf{n}$ в направлении от площадки $d\sigma$ к фокусу линзы O : $\mathbf{f} = \mathbf{n}f$.

Далее для расчёта поля воспользуемся интегралом Френеля, рассматривая лучи, идущие под малым углом к оптической оси:

$$A(P) = K(0) \int_{\Sigma} A_{\Sigma} \frac{e^{ikr}}{r} d\sigma, \quad K(0) = \frac{k}{2\pi i}. \quad (6.5.5)$$

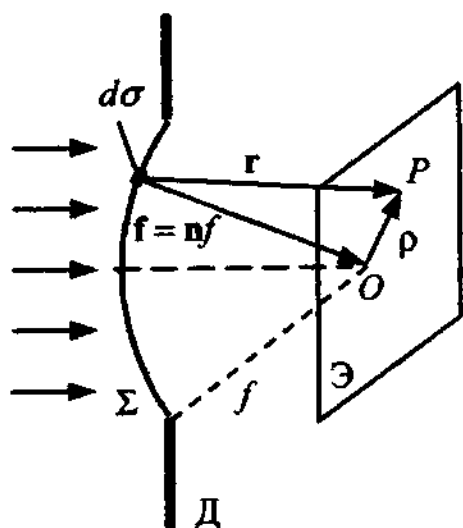


Рис. 6.5.2. Линза преобразует плоскую волну в волну со сферическим волновым фронтом Σ , которая направляется в фокус O

В принятых допущениях величину r в знаменателе в подынтегральном выражении в (6.5.5) можно приближённо положить равной $r \approx f$. Ограничиваясь сначала только первыми двумя слагаемыми в правой части (6.5.4), перепишем интеграл (6.5.5) в следующем виде:

$$A(P) = \frac{k}{2\pi i f} e^{ikf} A_0 \int_{\Sigma} e^{ik\mathbf{n}\mathbf{p}} d\sigma, \quad (6.5.6)$$

где A_0 — поле в точках линзы.

Обозначим координаты элемента волнового фронта $d\sigma$ как (ξ, η) , а координаты точки P как (x, y) . Тогда

$$d\sigma = d\xi d\eta, \quad \mathbf{p} = (x, y), \quad \mathbf{n} = \mathbf{f}/f = (-\xi/f, -\eta/f). \quad (6.5.7)$$

В результате из (6.5.6) следует

$$A(x, y) = \frac{k}{2\pi i f} e^{ikf} A_0 \int_{\Sigma} \exp[-i(q_x \xi + q_y \eta)] d\xi d\eta, \quad (6.5.8)$$

где введены *пространственные частоты*

$$q_x = kx/f, \quad q_y = ky/f. \quad (6.5.9)$$

Сначала найдём величину поля точно в фокусе: $x = 0, y = 0$. В этом случае из (6.5.8) следует

$$A_f = \frac{kA_0}{2\pi i f} e^{ikf} \int_{\Sigma} d\xi d\eta. \quad (6.5.10)$$

Входящий сюда интеграл равен площади линзы. Таким образом, для линзы радиуса R находим интенсивность:

$$I_f = |A_f|^2 = \left(kR^2/2f\right)^2 I_0 = \left(\pi R^2/\lambda f\right)^2 I_0.$$

Этот результат был получен выше методом зон Френеля (формула (6.5.1)).

Чтобы найти поле не только в фокусе, но и в его окрестности, нужно вычислять интеграл (6.5.8) при $x \neq 0, y \neq 0$.

Рассмотрим частный случай, когда линза ограничена диафрагмой, имеющей форму прямоугольника с размерами a и b вдоль осей ξ и η соответственно. Тогда из (6.5.8) находим

$$\begin{aligned} A(x, y) &= \frac{kA_0}{2\pi i f} e^{ikf} \int_{-a/2}^{a/2} e^{-iq_x \xi} d\xi \int_{-b/2}^{b/2} e^{-iq_y \eta} d\eta = \\ &= \frac{kA_0}{2\pi i f} ab \frac{\sin(q_x a/2)}{q_x a/2} \frac{\sin(q_y b/2)}{q_y b/2}. \end{aligned} \quad (6.5.11)$$

Возвращаясь согласно (6.5.9) от пространственных частот к координатам точки в фокальной плоскости, получаем

$$A(x, y) = \frac{kA_0}{2\pi if} e^{ikf} ab \frac{\sin(kax/2f)}{kax/2f} \frac{\sin(kby/2f)}{kby/2f}. \quad (6.5.12)$$

Наконец, для интенсивности излучения отсюда следует выражение

$$I(x, y) = |A(x, y)|^2 = I_{\max} \left(\frac{\sin(kax/2f)}{kax/2f} \right)^2 \left(\frac{\sin(kby/2f)}{kby/2f} \right)^2. \quad (6.5.13)$$

Здесь $I_{\max}$ — интенсивность в фокусе линзы:

$$I_{\max} = I_0 (ab/\lambda f)^2. \quad (6.5.14)$$

Последний результат прямо следует из (6.5.10) для линзы площадью ab .

На рис. 6.5.3 показано пространственное распределение интенсивности излучения в фокальной плоскости линзы квадратной формы. Хорошо видны главный и дополнительные максимумы.

Согласно (6.5.13) размеры фокусного пятна, определяемые координатами первых дифракционных минимумов:

$$I(\pm d_x/2, y) = I(x, \pm d_y/2) = 0,$$

составляют

$$d_x = 2 \frac{2\pi f}{ka} = \frac{2\lambda}{a} f, \quad d_y = 2 \frac{2\pi f}{kb} = \frac{2\lambda}{b} f. \quad (6.5.15)$$

В расчётах мы пренебрегали слагаемыми $\sim \rho^2/f$ в разложении (6.5.4) и соответствующими слагаемыми $\sim k\rho^2/f$ в фазе в подынтегральном выражении в (6.5.6). Проверим законность этого приближения: полагая согласно (6.5.15) $\rho \sim 2\lambda f/b$, находим

$$\frac{k\rho^2}{2f} \sim \frac{k}{2f} \left(\frac{2\lambda}{b} f \right)^2 = \frac{4\pi\lambda}{b^2} f.$$

Поскольку в типичных оптических приборах $b \sim f \gg \lambda$, то выполняется неравенство $k\rho^2/2f \ll 1$, оправдывающее приближение,

сделанное при расчёте поля в окрестности фокуса линзы.

6.5.3. Поле вблизи фокуса линзы на оптической оси

Представляет интерес найти, как меняется интенсивность излучения при отходе от фокуса вдоль оптической оси (рис. 6.5.4). Для решения этой задачи можно воспользоваться полученной в разделе 6.5.2 формулой (6.5.6). Действительно, при её выводе не использовалось предположение о том, что рассматривается только фокальная плоскость. Поэтому будем

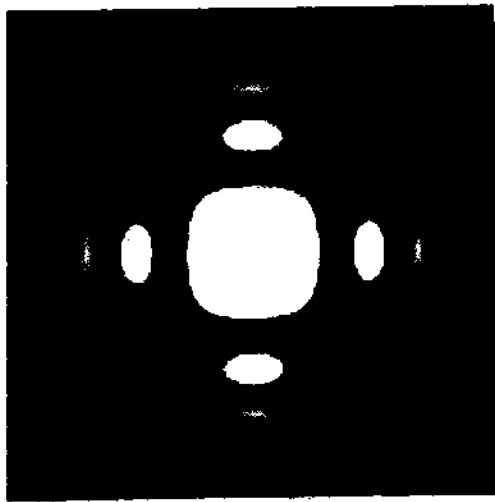


Рис. 6.5.3. Распределение интенсивности в фокальной плоскости линзы квадратной формы

теперь считать, что вектор ρ начинается в фокусе линзы и направлен вдоль оптической оси (оси z) к точке наблюдения P (на рис. 6.5.4 $\rho = \overline{OP}$).

Пусть угловой радиус линзы, наблюдаемый из точки P , равен θ ; в параксиальном приближении

$$\theta = R/(f + \rho).$$

Положим в формуле (6.5.6) $A_F = A_0 = \text{const}$:

$$A(P) = \frac{kA_0}{2\pi if} e^{ikf} \int_{\Sigma} e^{ik\rho \cos \beta} d\sigma. \quad (6.5.16)$$

При малых отходах от фокуса кольцевой элемент площади волнового фронта можно записать как

$$d\sigma = f^2 d\Omega = 2\pi f^2 \sin \beta d\beta, \quad 0 \leq \beta \leq \theta. \quad (6.5.17)$$

Учтём также, что

$$\mathbf{n}\rho = \pm \rho \cos \beta, \quad (6.5.18)$$

где знак «+» соответствует точке P справа от фокуса, а «-» — слева.

Таким образом, вместо (6.5.16) получаем следующее выражение:

$$A(\rho) = \frac{kA_0}{2\pi if} e^{ikf} 2\pi f^2 \int_0^\theta e^{\pm ik\rho \cos \beta} \sin \beta d\beta = \mp A_0 e^{ikf} \frac{f}{\rho} \left[e^{\pm ik\rho} - e^{\pm ik\rho \cos \theta} \right]. \quad (6.5.19)$$

Отсюда находим интенсивность в точке наблюдения:

$$I(\rho) = |A(\rho)|^2 = 4I_0 \left(\frac{f}{\rho} \right)^2 \sin^2 \left[\frac{1}{2} k\rho (1 - \cos \theta) \right]. \quad (6.5.20)$$

Интенсивность в фокусе линзы I_f из этой формулы получается в пределе $\rho \rightarrow 0$. Исключая угол θ в параксиальном приближении:

$$\theta \approx R/f \ll 1, \quad 1 - \cos \theta \approx \theta^2/2, \quad (6.5.21)$$

приходим к полученному ранее результату:

$$I_f = I_0 \left(kR^2/2f \right)^2 = I_0 \left(\pi R^2/\lambda f \right)^2. \quad (6.5.22)$$

Как следует из (6.5.20), по мере отхода от фокуса ($\rho \neq 0$) интенсивность осциллирует, причём по мере удаления амплитуда осцилляций убывает. График зависимости $I(\rho)$ показан на рис. 6.5.5. Интенсивность обращается в нуль на расстояниях до фокуса, удовлетворяющих условию

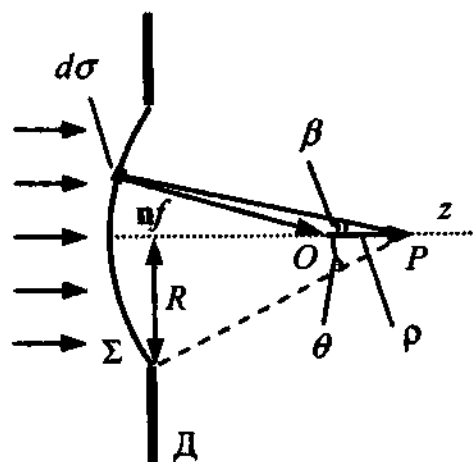


Рис. 6.5.4. К расчёту поля на оптической оси z линзы вблизи фокуса

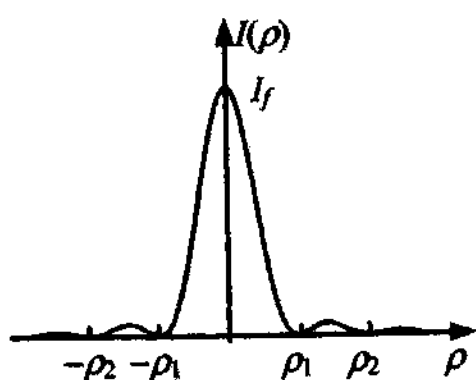


Рис. 6.5.5. Зависимость интенсивности излучения на оптической оси линзы от расстояния ρ до фокуса

$$\frac{1}{2}k\rho(1 - \cos \theta) = m\pi, \quad m = \pm 1, \pm 2, \dots,$$

или, в параксиальном приближении (6.5.21),

$$\rho_m = \frac{m\lambda}{1 - \cos \theta} \approx 2m\lambda \left(\frac{f}{R} \right)^2, \quad (6.5.23)$$

$$m = \pm 1, \pm 2, \dots$$

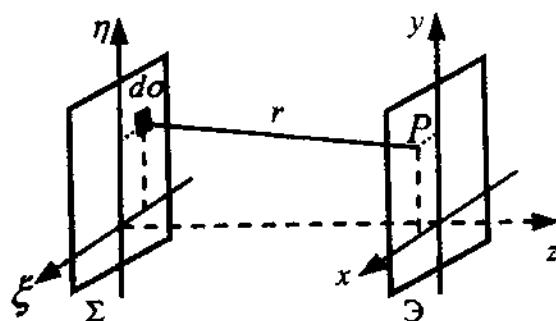
Между нулями находятся максимумы интенсивности. Таким образом, дифракция приводит к появлению дополнительных фокусов, в которых интенсивность быстро убывает по мере удаления от главного фокуса.

6.6. Связь поля в фокальной плоскости линзы с преобразованием Фурье поля, падающего на линзу

6.6.1. Параксиальное приближение для интеграла Френеля

Пусть исходное волновое поле задано на плоскости (ξ, η) (рис. 6.6.1).

Требуется найти поле в точке P на некотором расстоянии z . В параксиальном приближении в интеграле Френеля (6.1.5) можно положить $K(\alpha) \approx K(0) = 1/i\lambda$.



Исходная плоскость
($z = 0$)

Положение точки
наблюдения $P(x, y, z)$

Рис. 6.6.1. К выводу параксиального приближения интеграла Френеля

Запишем интеграл Френеля:

$$A_P(x, y) = \frac{1}{i\lambda} \int_{\Sigma} A_{\Sigma}(\xi, \eta) \frac{e^{ikr}}{r} d\sigma, \quad d\sigma = d\xi d\eta. \quad (6.6.1)$$

Расстояние от элемента площади источника $d\sigma$ до точки наблюдения P (рис. 6.6.1) равно

$$r = \sqrt{z^2 + (x - \xi)^2 + (y - \eta)^2}. \quad (6.6.2)$$

Поскольку мы сейчас имеем дело с узкими, слаборасходящимися пучками, то выполняются неравенства

$$|\xi|, |\eta|, |x|, |y| \ll z. \quad (6.6.3)$$

Это позволяет выполнить в (6.6.2) разложение

$$r \approx z + \frac{(x - \xi)^2 + (y - \eta)^2}{2z}. \quad (6.6.4)$$

Кроме того, в знаменателе подынтегрального выражения в (6.6.1) можно положить $r \approx z$. В результате мы приходим к следующему представлению:

$$A(x, y, z) = \frac{e^{ikz}}{i\lambda z} \int_{\Sigma} A_{\Sigma}(\xi, \eta) \exp \left\{ \frac{ik}{2z} [(x - \xi)^2 + (y - \eta)^2] \right\} d\xi d\eta. \quad (6.6.5)$$

Полученная формула называется *приближением Френеля дифракционного интеграла*.

6.6.2. Поле в фокальной плоскости линзы

Пусть на тонкую линзу падает волна $A_{in}(\xi, \eta)$, где координаты ξ, η перечисляют точки поверхности линзы. Пройдя линзу, лучи приобретают дополнительную разность хода

$$\Delta(\xi, \eta) = (n - 1)h(\xi, \eta), \quad (6.6.6)$$

где h — толщина линзы в месте прохождения соответствующего луча (рис. 6.6.2).

Рассмотрим сначала плосковыпуклую линзу с максимальной толщиной h_{01} и радиусом кривизны сферической поверхности R_1 (рис. 6.6.3). Выберем начало координат на оптической оси. Тогда на расстоянии

$\sqrt{\xi^2 + \eta^2}$ от оси толщина линзы составит

$$h(\xi, \eta) = h_{01} - \left[R_1 - \sqrt{R_1^2 - (\xi^2 + \eta^2)} \right]. \quad (6.6.7)$$

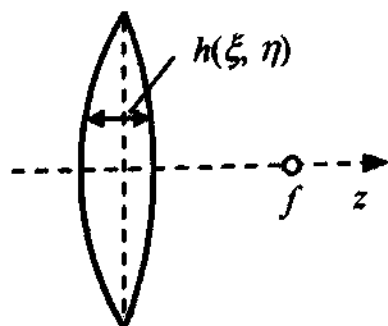


Рис. 6.6.2. Толщина линзы в зависимости от расстояния до оптической оси. f — положение фокуса

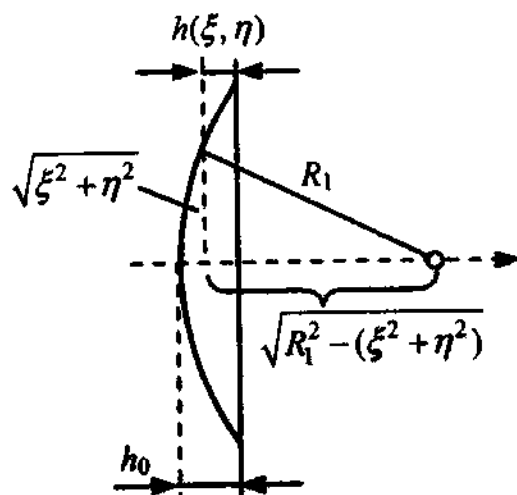


Рис. 6.6.3. Расчёт толщины плосковыпуклой линзы в зависимости от расстояния до оптической оси

В параксиальном приближении

$$\sqrt{\xi^2 + \eta^2} \ll R_1$$

из (6.6.7) следует

$$h(\xi, \eta) = h_{01} - \frac{\xi^2 + \eta^2}{2R_1}. \quad (6.6.8)$$

Аналогично можно рассмотреть случай, когда обе поверхности сферические с радиусами кривизны поверхностей, равными соответственно R_1 и R_2 :

$$h(\xi, \eta) = h_0 - \frac{1}{2}(\xi^2 + \eta^2) \left(\frac{1}{R_1} + \frac{1}{R_2} \right). \quad (6.6.9)$$

Здесь $h_0 = h_{01} + h_{02}$ — максимальная толщина линзы (достигаемая на оптической оси). Используя (6.6.6), находим дополнительный оптический ход лучей $\Delta(\xi, \eta)$ и связанный с этим дополнительный сдвиг фаз:

$$\varphi(\xi, \eta) = k\Delta(\xi, \eta) = \varphi_0 - \frac{k}{2f}(\xi^2 + \eta^2). \quad (6.6.10)$$

Здесь введены постоянная фаза $\varphi_0 = kh_0$ и фокусное расстояние линзы f по известной формуле

$$\frac{1}{f} = (n-1) \left(\frac{1}{R_1} + \frac{1}{R_2} \right).$$

Поскольку величина φ_0 не зависит от координат, её можно в дальнейшем не учитывать.

Пусть на линзу падает поле $A_{in}(\xi, \eta)$ (при $z = -0$). Сразу после прохождения линзы волна меняет фазу на $\varphi(\xi, \eta)$, так что можно положить при $z = +0$:

$$A_{\Sigma}(\xi, \eta) = A_{in}(\xi, \eta)e^{i\varphi(\xi, \eta)} = A_{in}(\xi, \eta) \exp \left[-\frac{ik}{2f}(\xi^2 + \eta^2) \right]. \quad (6.6.11)$$

Для расчёта поля в фокальной плоскости линзы $z = f > 0$ воспользуемся интегралом Френеля в форме (6.6.5):

$$\begin{aligned} A(x, y, f) &= \frac{e^{ikf}}{i\lambda f} \int_{\Sigma} A_{\Sigma}(\xi, \eta) \exp \left\{ \frac{ik}{2f} [(x-\xi)^2 + (y-\eta)^2] \right\} d\xi d\eta = \\ &= \frac{1}{i\lambda f} \exp \left[ik \left(f + \frac{x^2 + y^2}{2f} \right) \right] \int_{\Sigma} A_{in}(\xi, \eta) \exp \left\{ -\frac{ik}{f} (x\xi + y\eta) \right\} d\xi d\eta. \end{aligned} \quad (6.6.12)$$

В этой формуле интегрирование выполняется по всей области, куда падает волна $A_{in}(\xi, \eta)$. Если обозначить

$$q_x = kx/f, \quad q_y = ky/f, \quad (6.6.13)$$

то мы видим, что поле в фокальной плоскости линзы $A(x, y, f)$ пропорционально преобразованию Фурье поля $A_{in}(\xi, \eta)$, падающего на линзу:

$$A(x, y, f) \sim a(q_x, q_y) = \int_{\Sigma} A_{in}(\xi, \eta) \exp[-i(q_x \xi + q_y \eta)] d\xi d\eta. \quad (6.6.14)$$

Фазовый множитель перед интегралом в (6.6.12) не влияет на распределение интенсивности излучения:

$$I(x, y) = |A(x, y, f)|^2 = \frac{1}{\lambda^2 f^2} |a(q_x, q_y)|^2. \quad (6.6.15)$$

6.6.3. Фокусировка гауссова пучка излучения

Пусть на линзу падает излучение с гауссовым распределением

$$A_{in}(\xi, \eta) = A_0 \exp\left(-\frac{\xi^2 + \eta^2}{2R_0^2}\right). \quad (6.6.16)$$

Для интенсивности имеет место выражение

$$I_{in}(\xi, \eta) = I_0 \exp\left(-\frac{\xi^2 + \eta^2}{R_0^2}\right). \quad (6.6.17)$$

Величина R_0 называется гауссовым радиусом пучка. Полная мощность излучения равна

$$P = \int_{\eta=-\infty}^{\infty} \int_{\xi=-\infty}^{\infty} I(\xi, \eta) d\xi d\eta = \pi R_0^2 I_0. \quad (6.6.18)$$

Для нахождения поля в фокальной плоскости воспользуемся формулой (6.6.15), в которой функция $a(q_x, q_y)$ даётся выражением (6.6.14):

$$a(q_x, q_y) = A_0 \int_{\Sigma} \exp\left[-\frac{\xi^2 + \eta^2}{2R_0^2} - i(q_x \xi + q_y \eta)\right] d\xi d\eta = A_0 a_1(q_x) a_1(q_y).$$

Входящая сюда функция $a_1(q)$ равна

$$a_1(q) = \int_{-\infty}^{\infty} \exp\left(-\frac{\xi^2}{2R_0^2} - iq\xi\right) d\xi = \sqrt{2\pi} R_0 \exp\left(-\frac{1}{2} q^2 R_0^2\right),$$

в результате чего находим

$$a(q_x, q_y) = 2\pi R_0^2 \exp\left(-\frac{q_x^2 + q_y^2}{2} R_0^2\right).$$

Наконец, по формуле (6.6.15) получаем искомое распределение:

$$I(x, y) = 4I_0 \left(\frac{\pi R_0^2}{\lambda f} \right)^2 \exp \left(-\frac{k^2 R_0^2}{f^2} \rho^2 \right), \quad (6.6.19)$$

где $\rho^2 = x^2 + y^2$. Здесь учтены соотношения (6.6.13), связывающие пространственные частоты q_x, q_y с координатами x, y точки наблюдения в фокальной плоскости.

Таким образом, поле сохраняет гауссово распределение, но с иными значениями амплитуды и гауссова радиуса:

$$I_{0,f} = 4I_0 \left(\frac{\pi R_0^2}{\lambda f} \right)^2, \quad R_{0,f} = \frac{f}{kR_0} = \frac{\lambda}{2\pi R_0} f. \quad (6.6.20)$$

Можно показать, что гауссов профиль пучка сохраняется не только в фокальной плоскости, но и на произвольном расстоянии от линзы.

Нетрудно проверить выполнение закона сохранения энергии:

$$P = \pi R_0^2 I_0 = \pi R_{0,f}^2 I_{0,f}. \quad (6.6.21)$$

Заметим, что, в отличие от случая дифрагмированной плоской волны, для гауссова пучка осцилляции интенсивности в фокальной плоскости нет — имеет место монотонное убывание по мере удаления от оси.

6.7. Связь дифракции Фраунгофера с пространственным преобразованием Фурье

Дифракция Фраунгофера реализуется на больших расстояниях z на отверстиях или препятствиях (рис. 6.7.1, слева):

$$z \gg b^2 / \lambda, \quad (6.7.1)$$

где b — размер отверстия (препятствия). Далее для определённости говорим о дифракции на отверстиях.

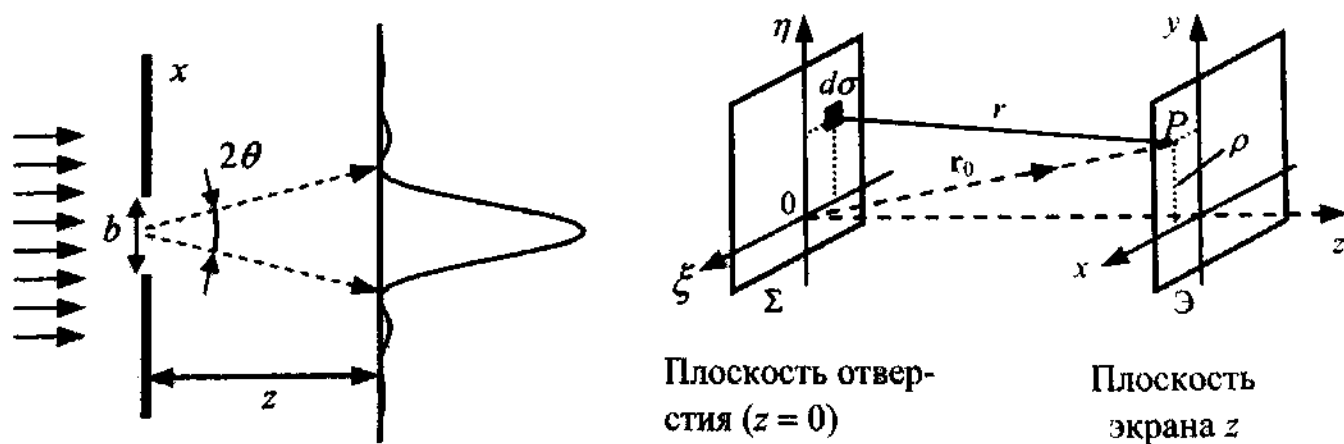


Рис. 6.7.1. Слева — дифракция Фраунгофера на отверстии, справа — обозначения в расчёте дифракции

Получим формулу, определяющую распределение поля на экране в зоне дифракции Фраунгофера (рис. 6.7.1, справа). В точках отверстия и экрана введём системы координат соответственно (ξ, η) и (x, y) . Будем считать, что эти системы имеют общее начало координат, лежащее на оси z . Запишем интеграл Френеля, выражающий поле A_P в точке P через поле A_Σ в точках отверстия:

$$A_P(x, y) = \int_{\Sigma} A_{\Sigma}(\xi, \eta) \frac{e^{ikr}}{r} K(\alpha) d\sigma, \quad d\sigma = d\xi d\eta. \quad (6.7.2)$$

Входящее в (6.7.2) расстояние r от элемента волнового фронта $d\sigma$ до точки наблюдения равно

$$r = \sqrt{z^2 + (x - \xi)^2 + (y - \eta)^2}. \quad (6.7.3)$$

Считая расстояние z большим по сравнению с характерными размерами отверстия и области наблюдения (рассматривая приосевую область дифракционной картины), имеем

$$r = \sqrt{r_0^2 + \xi^2 + \eta^2 - 2x\xi - 2y\eta} \approx r_0 + \frac{\xi^2 + \eta^2 - 2x\xi - 2y\eta}{2r_0}, \quad (6.7.4)$$

где $r_0 = \sqrt{z^2 + x^2 + y^2}$ — расстояние от центра отверстия (начала координат) до точки наблюдения, $r_0 \approx z$. В зоне дифракции Фраунгофера на отверстии с характерным размером $\sim b$ выполняется неравенство

$$\frac{k(\xi^2 + \eta^2)}{2r_0} \sim \frac{kb^2}{2r_0} \sim \pi \frac{b^2}{\lambda z} \ll 1. \quad (6.7.5)$$

Поэтому слагаемым $k(\xi^2 + \eta^2)/2r_0$ в показателе экспоненты e^{ikr} можно пренебречь. Кроме того, в предэкспоненциальном множителе можно положить $r \approx r_0 \approx z$. Функция $K(\alpha)$ в параксиальном приближении ($\alpha \ll 1$) практически постоянна:

$$K(\alpha) \approx K(0) = \frac{k}{2\pi i} = \frac{1}{i\lambda}. \quad (6.7.6)$$

С учётом сказанного интеграл Френеля принимает вид

$$A_P(x, y) = \frac{e^{ikR}}{i\lambda r_0} \int_{\Sigma} A_{\Sigma}(\xi, \eta) \exp\left(-ik \frac{x\xi + y\eta}{r_0}\right) d\xi d\eta. \quad (6.7.7)$$

Введём обозначения

$$k \frac{x}{r_0} = k \sin \theta_x = q_x, \quad k \frac{y}{r_0} = k \sin \theta_y = q_y. \quad (6.7.8)$$

Здесь θ_x и θ_y — углы, образуемые x - и y -составляющими луча $\mathbf{r}_0$ (рис. 6.7.1), с осью z . Соответственно получаем

$$A_P(x, y) = \frac{e^{ikr_0}}{i\lambda r_0} \int_{\Sigma} A_{\Sigma}(\xi, \eta) e^{-iq_x \xi - iq_y \eta} d\xi d\eta. \quad (6.7.9)$$

Функция

$$a(q_x, q_y) = \int_{\Sigma} A_{\Sigma}(\xi, \eta) e^{-iq_x \xi - iq_y \eta} d\xi d\eta \quad (6.7.10)$$

есть преобразование Фурье поля A_{Σ} в точках отверстия. И поскольку

$$q_x = kx/r_0, \quad q_y = ky/r_0, \quad (6.7.11)$$

то получаем простую связь поля в зоне дифракции Фраунгофера с преобразованием Фурье от поля в точках отверстия:

$$A_P(x, y) \sim a(kx/r_0, ky/r_0). \quad (6.7.12)$$

Таким образом, поле в зоне дифракции Фраунгофера пропорционально преобразованию Фурье поля, заданного в плоскости $z = 0$.

Пример. Пусть при $z = 0$ излучение имеет гауссово распределение:

$$A_{\Sigma}(\xi, \eta) = A_0 \exp\left(-\frac{\xi^2 + \eta^2}{2R_0^2}\right). \quad (6.7.13)$$

Найдём поле в зоне дифракции Фраунгофера. Подставив (6.7.12) в (6.7.9), получаем

$$\begin{aligned} A_P(x, y) &= \frac{e^{ikr_0}}{i\lambda r_0} A_0 \int_{\xi=-\infty}^{\infty} \int_{\eta=-\infty}^{\infty} \exp\left[-\frac{\xi^2 + \eta^2}{2R_0^2} - iq_x \xi - iq_y \eta\right] d\xi d\eta = \\ &= \frac{e^{ikr_0}}{i\lambda r_0} A_0 2\pi R_0^2 \exp\left[-\frac{1}{2}(q_x^2 + q_y^2) R_0^2\right] = \frac{kR_0^2}{ir_0} e^{ikr_0} A_0 \exp\left(-\frac{k^2 R_0^2}{2r_0^2} \rho^2\right). \end{aligned} \quad (6.7.14)$$

Здесь $\rho = \sqrt{x^2 + y^2}$. Вводя угол дифракции соотношением

$$\rho/r_0 = \sin \theta, \quad (6.7.15)$$

перепишем полученное соотношение в виде

$$A_P(\theta) = \frac{kR_0^2}{ir_0} e^{ikr_0} A_0 \exp\left(-\frac{1}{2} k^2 R_0^2 \sin^2 \theta\right). \quad (6.7.16)$$

Таким образом, излучение сохраняет гауссов профиль, причём согласно (6.7.13) его гауссов радиус растёт с расстоянием по закону

$$R(z) = \frac{r_0}{kR_0} \approx \frac{z}{kR_0} \quad (6.7.17)$$

(здесь учтено, что $z = r_0 \cos \theta \approx r_0$). При этом угловая расходимость пучка составляет

$$k^2 R_0^2 \sin^2 \theta \sim 1 \Rightarrow \theta \sim 1/kR_0 \sim \lambda/R_0. \quad (6.7.18)$$

Заметим, что в фазовом множителе e^{ikr_0} делать замену $r_0 \rightarrow z$ нельзя: возникающая вследствие отличия r_0 и z ошибка в фазе

$$k(r_0 - z) = k\left(\sqrt{z^2 + \rho^2} - z\right) \approx \frac{k\rho^2}{2z} \sim \frac{k}{2z} \left(\frac{\lambda z}{b}\right)^2 = \frac{\pi\lambda}{b^2} z$$

достигает значения $\sim \pi$ на расстоянии $z \sim b^2/\lambda$, то есть именно в зоне дифракции Фраунгофера, и растёт по мере удаления от источника.

6.8. Дифракция Фраунгофера на решетке

6.8.1. Дифракционная решётка

Дифракционная решетка — это система N периодически расположенных щелей (штрихов) в непрозрачном экране (рис. 6.8.1).

Идея дифракционной решётки была открыта шотландским математиком и астрономом Джемсом Грегори (в середине XVII века) вскоре после экспериментов Ньютона с призмой. В качестве решётки он использовал птичьи перья. Первую дифракционную решётку изготовил американский астроном, изобретатель, математик Дэвид Риттенхаус в 1786 г. Его решётка состояла из волос, натянутых между двумя винтами с тонкой резьбой.

Пусть a — ширина щели, b — расстояние между ближайшими краями двух соседних щелей (перекрытой экраном области). Величина $d = a + b$ называется *периодом решетки* (рис. 6.8.1). Длина решетки с N штрихами равна $L = Nd$.

6.8.2. Главные дифракционные максимумы

Пусть на решётку падает излучение с длиной волны λ по нормали к плоскости решётки. Вследствие дифракции лучи, прошедшие через щели в решётке, отклоняются от своего первоначального направления (рис. 6.8.2). Рассмотрим лучи, отклонившиеся на угол θ . Разность хода лучей от эквивалентных точек двух соседних штрихов равна

$$\Delta = d \sin \theta. \quad (6.8.1)$$

Если эта величина равна целому числу длин волны:

$$d \sin \theta = m\lambda, \quad m = 0, \pm 1, \pm 2, \dots, \quad (6.8.2)$$

то волны от всех штрихов взаимно усиливают друг друга.

Дифракцию Фраунгофера можно наблюдать на экране, если на пути лучей поставить линзу, а экран разместить в её фокальной плоскости (рис. 6.8.2).

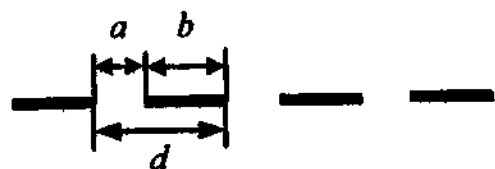


Рис. 6.8.1. Параметры дифракционной решётки

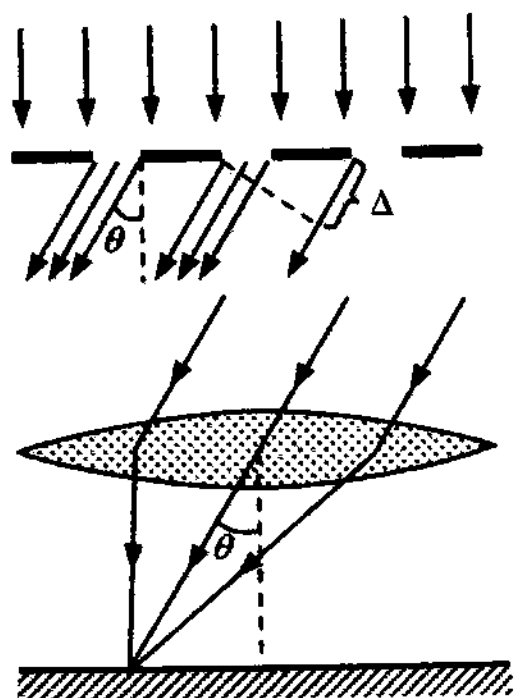


Рис. 6.8.2. Вследствие дифракции лучи, прошедшие через решётку, отклоняются от первоначального направления. Вверху: дифракция лучей на решётке; внизу: дифрагированные лучи собираются линзой на экране в её фокальной плоскости

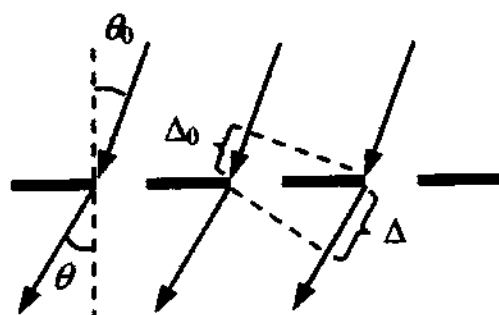


Рис. 6.8.3. При наклонном падении лучей на дифракционную решётку разность хода лучей формируется из составляющих до (Δ_0) и после (Δ) решётки

штрих, равно A_0 . Если волна падает на решётку под углом θ_0 , то разность фаз волн, дифрагировавших от двух соседних штрихов, равна

$$\delta = k(\Delta - \Delta_0) = kd(\sin \theta - \sin \theta_0). \quad (6.8.6)$$

Далее ограничимся случаем нормального падения волны на решётку: $\theta_0 = 0$.

Полагая, что фаза волны, идущей от крайнего левого штриха, равна нулю, получаем, что волна, идущая от n -го штриха, имеет фазу

$$\varphi_n = (n-1)\delta = (n-1)kd \sin \theta, \quad n = 1, 2, \dots, N. \quad (6.8.7)$$

Направления, определяемые формулой (6.8.2), задают *главные максимумы* дифракционной картины (в обе стороны от нулевого, или центрального, главного максимума, отвечающего значению $m = 0$). Полное число таких максимумов определяется из того условия, что $|\theta| \leq \pi/2$, или $|\sin \theta| \leq 1$. Соответственно имеем $|m| \leq d/\lambda$. Если период решётки $d < \lambda$, то будет наблюдаться единственный максимум, отвечающий значению $m = 0$. При $d > \lambda$ число главных максимумов не превышает

$$(1 + 2d/\lambda). \quad (6.8.3)$$

Мы рассмотрели случай, когда волна падает на решётку по нормали. Если же угол падения равен θ_0 (рис. 6.8.3), то разность хода лучей формируется как из составляющей Δ_0 до решётки, так и составляющей Δ после решётки и равна

$$\Delta - \Delta_0 = d \sin \theta - d \sin \theta_0. \quad (6.8.4)$$

Соответственно направления на главные максимумы теперь определяются условием

$$d(\sin \theta - \sin \theta_0) = m\lambda. \quad (6.8.5)$$

6.8.3. Дифракция Фраунгофера на решётке

Найдём теперь распределение интенсивности излучения в зависимости от угла дифракции. Если решётка имеет период d и содержит N штрихов, то её длина равна $L = Nd$. Пусть поле, прошедшее через один

Суммарное поле волны, распространяющейся в направлении θ , равно

$$\begin{aligned} A &= A_1 + A_2 + \dots + A_N = A_0 \left[e^{i\varphi_1} + e^{i\varphi_2} + \dots + e^{i\varphi_N} \right] = \\ &= A_0 \left[1 + e^{i\delta} + e^{2i\delta} + \dots + e^{(N-1)i\delta} \right] = A_0 \frac{e^{iN\delta} - 1}{e^{i\delta} - 1}. \end{aligned} \quad (6.8.8)$$

Имея в виду, что интенсивность $I = |A|^2$, находим

$$I = I_0 \frac{\sin^2 \left(\frac{N}{2} kd \sin \theta \right)}{\sin^2 \left(\frac{1}{2} kd \sin \theta \right)}. \quad (6.8.9)$$

Здесь $I_0 = |A_0|^2$ — интенсивность волны, идущей от одного штриха.

Главные максимумы

Направления на главные максимумы определяются из того условия, что числитель и знаменатель в (6.8.9) одновременно обращаются в нуль:

$$\frac{1}{2} kd \sin \theta = m\pi, \text{ или } d \sin \theta = m\lambda, \quad m = 0, \pm 1, \pm 2, \dots \quad (6.8.10)$$

Это совпадает со сформулированным выше условием (6.8.2).

Имея в виду предел $\lim_{x \rightarrow m\pi} \left(\frac{\sin^2 Nx}{\sin^2 x} \right) = N^2$, $x = \frac{1}{2} kd \sin \theta$, находим ин-

тенсивность излучения в главных максимумах:

$$I = N^2 I_0. \quad (6.8.11)$$

Такое возрастание интенсивности связано с интерференционным усилением волн, идущих от N штрихов.

Ширина главного максимума и положение минимумов

Для нахождения ширины главного максимума заметим, что согласно (6.8.9) интенсивность обращается в нуль при условии

$$\sin^2 \left(\frac{N}{2} kd \sin \theta \right) = 0, \quad \sin^2 \left(\frac{1}{2} kd \sin \theta \right) \neq 0. \quad (6.8.12)$$

Полагаем

$$\frac{N}{2} kd \sin \theta = Nm\pi + n\pi, \quad n = 1, 2, \dots, N-1, \quad (6.8.13)$$

откуда находим

$$d \sin \theta = \left(m + \frac{n}{N} \right) \lambda. \quad (6.8.14)$$

Эта формула определяет направления на *минимумы интенсивности* дифрагированной волны (рис. 6.8.4).

Для главного максимума нулевого порядка ($m = 0$) направление на первый минимум даётся формулой

$$d \sin \theta = \lambda / N, \text{ или } \sin \theta = \lambda / Nd. \quad (6.8.15)$$

Таким образом, направление на первый минимум определяется длиной решётки $L = Nd$:

$$L \sin \theta_{\min} = \lambda. \quad (6.8.16)$$

Иными словами, требуется, чтобы *разность хода от крайних точек решётки была равна длине волны излучения*. Подчеркнём, что величина L есть *полная* длина решётки, включающая как открытые, так и закрытые участки (рис. 6.8.1) и равная целому числу периодов.

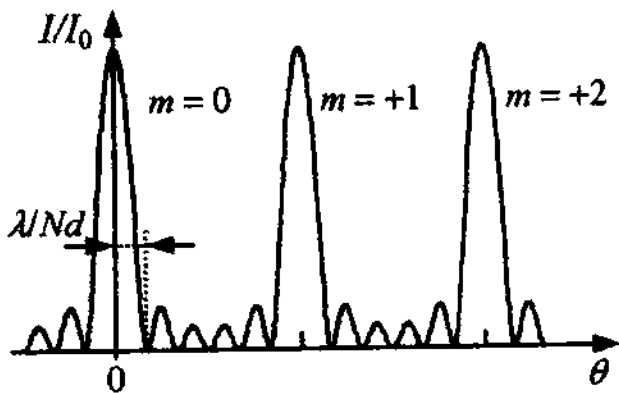


Рис. 6.8.4. Угловое распределение интенсивности излучения, прошедшего через дифракционную решётку (показана только область углов $\theta > 0$; картина для $\theta < 0$ полностью аналогична). Рисунок построен для случая, когда число щелей решётки $N = 6$

Поскольку, как правило, $\lambda \ll L$, то угловая ширина главного максимума нулевого порядка равна

$$\Delta \theta = 2\theta_{\min} = 2\lambda / Nd \ll 1. \quad (6.8.17)$$

Если период решётки достаточно велик: $d \gg \lambda$, а число штрихов $N \gg 1$, то большое число главных максимумов попадает в область углов дифракции $\theta \ll 1$. В частности, для m -го главного максимума (в этом диапазоне углов) $\theta_m = m \lambda / d$. Ближайшие к нему слева и справа минимумы имеют направления такие, что

$$\theta_{-1}^{(m)} = \left(m - \frac{1}{N} \right) \frac{\lambda}{d}, \quad \theta_{+1}^{(m)} = \left(m + \frac{1}{N} \right) \frac{\lambda}{d}. \quad (6.8.18)$$

Отсюда находим угловую ширину m -го максимума:

$$\Delta \theta = \theta_{+1}^{(m)} - \theta_{-1}^{(m)} = 2\lambda / Nd. \quad (6.8.19)$$

Это совпадает с шириной максимума 0-го порядка (6.8.17).

Дополнительные максимумы

Оценим интенсивность излучения в дополнительных максимумах. Рассмотрим главный максимум порядка $m = 0$, считая $N \gg 1$, $Nd \gg \lambda$. Тогда в его окрестности имеется большое число минимумов (рис. 6.8.4).

Для первого минимума (и первого дополнительного максимума), как видно из (6.8.17), угол дифракции мал:

$$\theta \sim \frac{\lambda}{Nd} \sim \frac{1}{Nkd} \ll \frac{1}{kd}. \quad (6.8.20)$$

При этом условии формула (6.8.9) принимает вид

$$I = I_0 \frac{\sin^2(Nkd\theta/2)}{(kd\theta/2)^2}. \quad (6.8.21)$$

Вследствие неравенства $N \gg 1$ числитель при изменении угла θ меняется существенно быстрее, чем знаменатель. Поэтому положение локальных максимумов практически совпадает с положением максимумов числителя:

$$\frac{N}{2} kd\theta_{\max} = \frac{\pi}{2} (2n+1),$$

или

$$\theta_{\max} = \frac{\lambda}{Nd} \left(n + \frac{1}{2} \right). \quad (6.8.22)$$

Здесь число n пробегает значения $n = 1, 2, \dots, N-1$ (значение $n = 0$ исключается из этого списка, так как соответствующий максимум пропадает, формально попадая в промежуток между главным максимумом и первым минимумом). Переписывая полученную формулу в виде

$$\frac{1}{2} kd\theta_{\max} = \frac{\pi}{N} \left(n + \frac{1}{2} \right), \quad (6.8.23)$$

находим из (6.8.21) интенсивность в дополнительных максимумах:

$$I^{(n)} \approx \frac{I_0}{(kd\theta_{\max}/2)^2} \approx \frac{I_0 N^2}{\pi^2 (n+1/2)^2}. \quad (6.8.24)$$

Отсюда находим

$$I^{(0)} : I^{(1)} : I^{(2)} \approx 1 : \frac{4}{9\pi^2} : \frac{4}{25\pi^2} \approx 1 : 0,045 : 0,016. \quad (6.8.25)$$

Таким образом, энергия, направляемая в дополнительные максимумы, мала по сравнению с энергией, направляемой в главный максимум.

6.8.4. Влияние размера щели на дифракционную картину

Выше мы не учитывали дифракцию волн на каждой отдельной щели дифракционной решётки. Учтём теперь, что расходимость волны зависит от ширины щели.

Как показано в разделе 6.4.2 (формула (6.4.10)), интенсивность волны за щелью шириной b даётся выражением

$$I_g(\theta) = I_0 \frac{\sin^2(kb \sin \theta/2)}{(kb \sin \theta/2)^2} \quad (6.8.26)$$

(индекс «g» означает «gap», «щель»). Соответственно вместо формулы (6.8.9) угловое распределение интенсивности теперь примет вид

$$I = I_0 \frac{\sin^2\left(\frac{1}{2}kb \sin \theta\right)}{\left(\frac{1}{2}kb \sin \theta\right)^2} \frac{\sin^2\left(\frac{1}{2}Nkd \sin \theta\right)}{\sin^2\left(\frac{1}{2}kd \sin \theta\right)}. \quad (6.8.27)$$

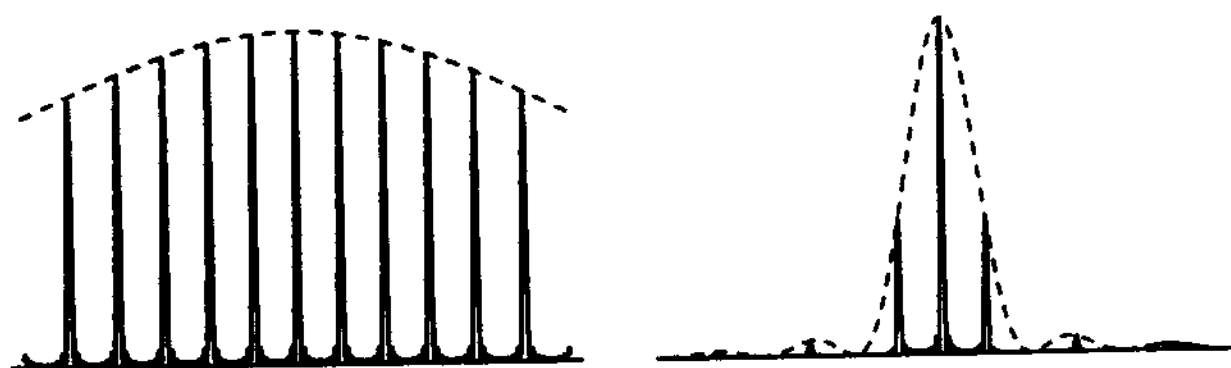


Рис. 6.8.5. Слева — угловое распределение интенсивности излучения от дифракционной решётки с параметрами: $N = 10$, $d = 10\lambda$, $b = d/20$; справа — то же для решётки с $b = d/2$ (остальные параметры те же). Диапазон углов на обоих графиках: $-\pi/5 < \theta < \pi/5$. Штриховой линией показана функция $N^2 I_g(\theta)$, пропорциональная угловому распределению интенсивности $I_g(\theta)$ волны при дифракции на отдельной щели шириной $b = d/20$ (для левого графика) и $b = d/2$ (для правого графика)

Если размер щели мал по сравнению в периодом решётки ($b \ll d$), то наблюдается большое число интерференционных максимумов — вплоть до значений $m \sim d/\lambda$. Этот случай проиллюстрирован на рис. 6.8.5 (слева). Если же ширина щели сравнима с периодом: $b \sim d$, то число наблюдаемых максимумов резко уменьшается. На рис. 6.8.5 (справа) показана дифракционная картина для случая, когда $b = d/2$. Во втором случае число наблюдаемых максимумов сокращается практически до трёх: $m = -1, 0, +1$ (остальные имеют пренебрежимо малую амплитуду).

Данный результат легко понять из следующего. Если щели широкие, то волны от них идут в узком конусе с амплитудой, быстро убывающей с ростом угла. Соответственно их вклад в дифракционные максимумы с большими номерами m (при больших углах дифракции) оказывается малым. Если же щели узкие, то от каждой из них волны близки к сферическим с амплитудой, почти не зависящей от направления. В результате их вклад в дифракционные максимумы зависит от номера m слабо.

6.9. Дифракция рентгеновских лучей

6.9.1. Рентгеновское излучение

К рентгеновскому диапазону относят электромагнитное излучение с длиной волны

$$10^{-3} \text{ нм} < \lambda < 50 \text{ нм}.$$

Иногда верхнюю границу диапазона сдвигают до 100 нм.

Исследования показали, что излучение с длиной волны $\lambda < 0.2 \text{ нм}$ — жёсткое, обладающее большой проникающей способностью. Излучение же с $\lambda > 0.2 \text{ нм}$ поглощается сильнее и проникает в вещество на меньшую глубину.

Как и свет, рентгеновское излучение обладает волновыми свойствами, проявляющимися, в частности, в явлениях интерференции и дифракции. В то же время вследствие малой длины волны такие явления наблюдать заметно сложнее.

В 1912 г. немецкий физик Макс фон Лауэ предложил для обнаружения дифракции рентгеновских лучей с длиной волны $\lambda \sim 1 \text{ \AA}$ использовать кристалл. Идея состояла в том, что межатомные расстояния в твёрдых телах именно такого порядка. Кроме того, в кристаллах атомы расположены регулярно, что позволяло ожидать наличие интерференции волн, рассеянных группой атомов. Эта идея была реализована экспериментально в том же году немецкими учёными Паулем Книппингом и Вальтером Фридрихом.

Австралийский физик У. Л. Брэгг в 1912 г. и независимо российский физик Г. В. Вульф в 1913 г. предложили простой способ расчёта картины дифракции рентгеновских лучей, известный как формула Брэгга–Вульфа.

Открытие дифракции рентгеновских лучей подтвердило их электромагнитную природу и указало место на шкале электромагнитных волн. Оно также доказало периодическую структуру кристаллов. Всё это привело к созданию рентгеноструктурного анализа — важного способа исследования строения вещества.

Рентгеновское излучение открыл в 1895 г. немецкий физик Вильгельм Конрад Рентген, за что был первым из физиков удостоен Нобелевской премии (1901). За открытие дифракции рентгеновских лучей М. Лауэ получил Нобелевскую премию (1914). В 1912 г. Уильям Лоренс Брэгг получил уравнения, описывающие дифракцию рентгеновских лучей на кристаллах (условия интерференционного отражения рентгеновских лучей от кристаллов). В 1913 г. российский учёный-кристаллограф Георгий Викторович Вульф независимо от У. Л. Брэгга пришёл к тем же результатам и первым в России начал рентгеноструктурные исследования. Формула Брэгга–Вульфа широко используется в рентгеновской спектроскопии и других разделах физики. За изучение структуры кристаллов с помощью рентгеновских лучей Нобелевскую премию (1915) получили Уильям Генри и Уильям Лоренс Брэгги (отец и сын).

6.9.2. Формула Брэгга–Вульфа

Рассмотрим схему, называемую *брэгговским отражением*, позволяющую рассчитывать дифракцию рентгеновского излучения на кристаллах. Введём *атомные плоскости*, то есть плоскости, содержащие большое число атомов (рис. 6.9.1, слева). Обозначим d расстояние между парой соседних плоскостей, а θ — угол скольжения луча.

Отражение волны от кристалла можно рассматривать как результат интерференционного усиления волн, идущих от разных атомных плоскостей. Рассмотрим две соседние параллельные плоскости (рис. 6.9.1, справа). Разность хода лучей MON и $M'O'N'$ равна

$$\Delta = PO'Q = 2d \sin \theta. \quad (6.9.1)$$

Условие интерференционного усиления этих волн имеет вид

$$2d \sin \theta = m\lambda, \quad m = \pm 1, \pm 2, \dots \quad (6.9.2)$$

Полученное соотношение называется *формулой Брэгга–Вульфа*.

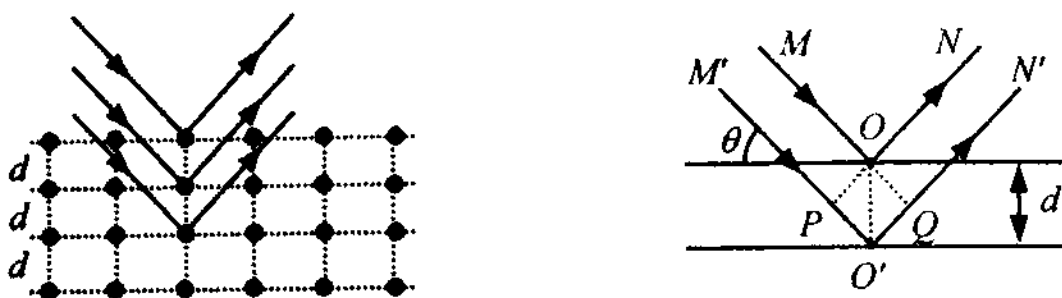


Рис. 6.9.1. Слева — рассеяние рентгеновских лучей на атомах кристаллической решётки, справа — расчёт разности хода лучей от двух соседних атомных плоскостей

Помимо дифракционных максимумов, направление на которые даётся равенством (6.9.2), имеются другие максимумы, отвечающие отражениям от других атомных плоскостей. При этом наиболее сильное отражение происходит от тех плоскостей, в которых густота атомов максимальна.

Если исследуемое вещество является монокристаллом, то при произвольном его положении условие Брэгга–Вульфа может не выполняться. Однако поворотом можно найти положение, в котором возникает дифракционный максимум, регистрируемый фотопластинкой или ионизационным детектором.

Если вещество является поликристаллом, а излучение монохроматично, то найдётся много кристалликов, для которых будет выполнено условие Брэгга–Вульфа. В результате возникает конус лучей с осью по направлению падающего пучка. Поскольку имеется дискретный набор атомных плоскостей, то возникает набор конусов дифрагировавших волн. При фотографировании на фотоплёнке будет регистрироваться серия полос. Этот способ, называемый *методом Дебая–Шеррера*, был предложен в 1926 г. П. Дебаем (Нидерланды) и П. Шеррером (Швейцария) и проиллюстриро-

ван на рис. 6.9.2. Снятая этим методом рентгенограмма называется *дебаеграммой*. На рис. 6.9.3 показан вид дебаеграммы сплава Fe—Al.

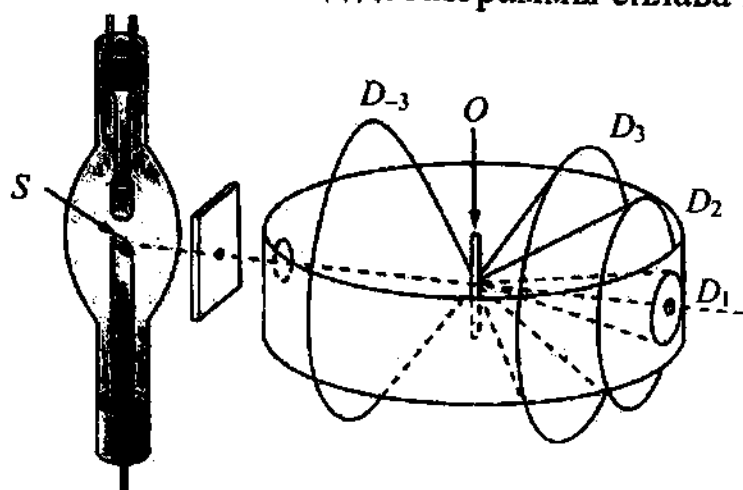


Рис. 6.9.2. Рассеяние монохроматического рентгеновского излучения на поликристаллическом образце. Излучение от источника S попадает на образец O и отражается от атомных плоскостей. Возникают конусы $D_1, D_{-1}, D_2, D_{-2}, \dots$, отвечающие различным дифракционным максимумам



Рис. 6.9.3. Дебаеграмма сплава Fe—Al

6.9.3. Формулы Лауэ

Кристалл — это трёхмерная структура, периодическая по трём направлениям. Обозначим периоды по осям кристалла 1, 2, 3 соответственно d_1, d_2, d_3 . Рассмотрим цепочку атомов 1. Пусть излучение падает на неё под углом α_0 , а наблюдается под углом α (рис. 6.9.4). Тогда разность хода лучей, дифрагировавших на соседних атомах, равна

$$\Delta = \Delta_1 - \Delta_0, \quad \Delta_1 = d_1 \cos \alpha, \quad \Delta_0 = d_1 \cos \alpha_0. \quad (6.9.3)$$

Отсюда получаем условие интерференционного усиления волн:

$$\Delta = m_1 \lambda, \quad m = 0, \pm 1, \pm 2, \dots, \quad (6.9.4)$$

или

$$d_1 (\cos \alpha - \cos \alpha_0) = m_1 \lambda. \quad (6.9.5)$$

Заметим, что в силу условия $|\cos \varphi| \leq 1$ число дифракционных максимумов по каждому направлению ограничено: $|m_k| \leq 2d_k/\lambda$, причём более одного максимума возможно, если $\lambda < 2d_k$.

Аналогичные соотношения имеют место и для дифракции на других осях:

$$d_2(\cos \gamma - \cos \gamma_0) = m_2 \lambda, \quad (6.9.6)$$

$$d_3(\cos \gamma - \cos \gamma_0) = m_3 \lambda. \quad (6.9.7)$$

Соотношения (6.9.5), (6.9.6), (6.9.7) называются формулами Лауэ.

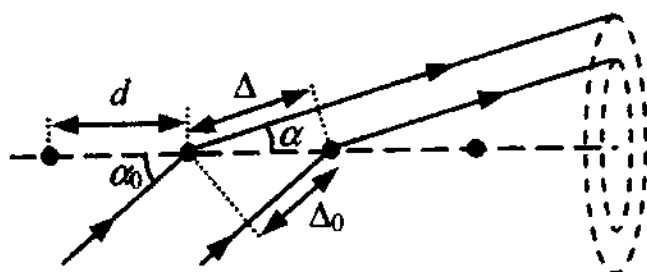


Рис. 6.9.4. К выводу формул Лауэ. Дифракция рентгеновского излучения на цепочке атомов

6.9.4. Эквивалентность формул Брэгга–Вульфа и Лауэ

Покажем, что все дифракционные максимумы, даваемые формулами Лауэ, можно рассматривать как результат отражения от каких-то атомных плоскостей.

Введём векторы рёбер элементарной ячейки: $\{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3\}$, направленные вдоль рёбер и с длиной, равной периоду кристалла в соответствующем направлении. Тогда положение атомов даётся формулой

$$\mathbf{r} = n_1 \mathbf{d}_1 + n_2 \mathbf{d}_2 + n_3 \mathbf{d}_3, \quad (6.9.8)$$

где $n_1, n_2, n_3 = 0, \pm 1, \pm 2, \dots$. Обозначим единичные векторы в направлении падающей на кристалл волны и волны дифрагированной как

$$\begin{aligned} \mathbf{n}_0 &= \{\cos \alpha_0, \cos \beta_0, \cos \gamma_0\}, \\ \mathbf{n} &= \{\cos \alpha, \cos \beta, \cos \gamma\}. \end{aligned} \quad (6.9.9)$$

Ограничимся случаем, когда элементарная ячейка прямоугольная. Ориентация вектора $\mathbf{n}$ по отношению к рёбрам показана на рис. 6.9.5. Запишем формулы Лауэ (6.9.5), (6.9.6), (6.8.7) в обозначениях (6.9.9):

$$\begin{aligned} (\mathbf{n} - \mathbf{n}_0) \mathbf{d}_1 &= m_1 \lambda, \\ (\mathbf{n} - \mathbf{n}_0) \mathbf{d}_2 &= m_2 \lambda, \\ (\mathbf{n} - \mathbf{n}_0) \mathbf{d}_3 &= m_3 \lambda. \end{aligned} \quad (6.9.10)$$

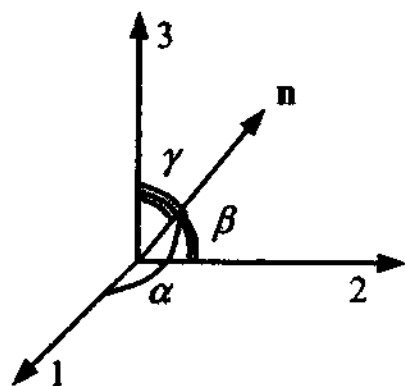


Рис. 6.9.5. Направление луча по отношению к рёбрам элементарной ячейки кристалла

Вектор $\mathbf{N} = \mathbf{n} - \mathbf{n}_0$ указывает изменение направления луча в результате дифракции (рис. 6.9.6). При этом уравнения (6.9.10) принимают вид

$$\mathbf{N} \mathbf{d}_1 = m_1 \lambda, \quad \mathbf{N} \mathbf{d}_2 = m_2 \lambda, \quad \mathbf{N} \mathbf{d}_3 = m_3 \lambda. \quad (6.9.11)$$

Проведём через атом в начале координат плоскость, перпендикулярную вектору $\mathbf{N}$. Её уравнение есть $\mathbf{N} \mathbf{r} = 0$, или, с учётом (6.9.8),

$$n_1 N d_1 + n_2 N d_2 + n_3 N d_3 = 0. \quad (6.9.12)$$

Подставляя сюда равенства (6.9.11), получим

$$n_1 m_1 + n_2 m_2 + n_3 m_3 = 0. \quad (6.9.13)$$

Отсюда видно, что для любого набора целых чисел m_1, m_2, m_3 , отвечающих максимумам дифракции на атомных цепочках, существует множество целых чисел n_1, n_2, n_3 , указывающих положения атомов кристалла, принадлежащих одной и той же плоскости, перпендикулярной вектору N . Но, как видно из рис. 6.9.6, соответствующее изменение направления луча можно интерпретировать как отражение от этой плоскости.

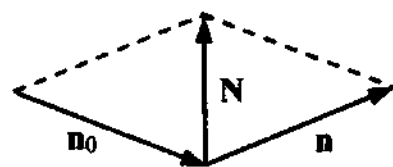


Рис. 6.9.6. Изменение направления луча в результате дифракции

6.9.5. Лауэграмма

Дифракционная картина, образованная рентгеновскими лучами, рассеянными кристаллом, называется *лауэграммой*. Используется узкий пучок излучения с непрерывным спектром, падающий на неподвижный кристалл, регистрируются лучи, прошедшие через кристалл.

Выясним структуру лауэграммы.

Первое условие

$$d_1 (\cos \alpha - \cos \alpha_0) = m_1 \lambda$$

выделяет в пространстве конус с осью 1, на котором находятся интерференционные максимумы. Второе условие

$$d_2 (\cos \beta - \cos \beta_0) = m_2 \lambda$$

выделяет конус максимумов вокруг оси 2. Пересечение этих конусов даёт набор линий, по которым конусы пересекаются (на рис. 6.9.7 это линии L_1 и L_2). Вдоль этих линий интенсивность излучения максимальна. Если учесть ещё и третье условие:

$$d_3 (\cos \gamma - \cos \gamma_0) = m_3 \lambda,$$

то есть перейти к дифракции на трёхмерной структуре, то дифрагированный пучок должен лежать уже на поверхности всех трёх конусов. Иначе интерференционного усиления не произойдёт, и рассеяние будет происходить почти равномерно во все стороны.

В общем случае три конуса не пересекаются по общим прямым — это возможно лишь при определённых условиях. Рассмотрим дифракционный максимум, выделяемый числами m_1, m_2, m_3 . Введём вектор

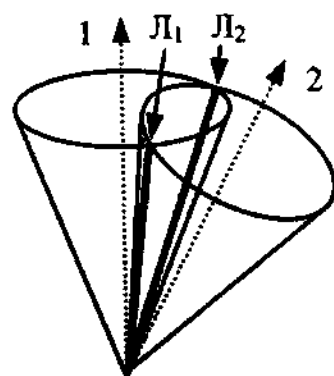


Рис. 6.9.7. Линии L_1 и L_2 пересечения конусов максимумов дифракции на осях 1 и 2

$$\mathbf{k} = \left(\frac{m_1}{d_1}, \frac{m_2}{d_2}, \frac{m_3}{d_3} \right). \quad (6.9.14)$$

Тогда условия Лауэ можно записать в виде

$$\mathbf{n} = \mathbf{n}_0 + \lambda \mathbf{k}. \quad (6.9.15)$$

Учтём очевидные тождества:

$$\begin{aligned} \mathbf{n}^2 &= \cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1, \\ \mathbf{n}_0^2 &= \cos^2 \alpha_0 + \cos^2 \beta_0 + \cos^2 \gamma_0 = 1. \end{aligned} \quad (6.9.16)$$

Тогда, возводя (6.9.15) в квадрат, получаем

$$2\mathbf{n}_0 \mathbf{k} + \lambda \mathbf{k}^2 = 0, \quad (6.9.17)$$

или

$$\lambda = -2 \frac{(m_1/d_1) \cos \alpha_0 + (m_2/d_2) \cos \beta_0 + (m_3/d_3) \cos \gamma_0}{(m_1/d_1)^2 + (m_2/d_2)^2 + (m_3/d_3)^2}. \quad (6.9.18)$$

В частном случае кристалла с простой кубической решёткой $d_1 = d_2 = d_3 \equiv d$ формула упрощается:

$$\lambda = -2 \frac{m_1 \cos \alpha_0 + m_2 \cos \beta_0 + m_3 \cos \gamma_0}{m_1^2 + m_2^2 + m_3^2} d. \quad (6.9.19)$$

Данное соотношение означает, что только для таких длин волн будут наблюдаться дифракционные максимумы. Соответственно для наблюдения дифракции на кристалле следует использовать рентгеновское излучение с широким сплошным спектром. Из всего спектра будут выделяться отдельные линии, для которых условие (6.9.18) выполняется для некоторого набора целых чисел m_1, m_2, m_3 .

При этом на фотопластинке будут регистрироваться пятна, отвечающие соответствующим дифракционным максимумам.

Лауэграмма монокристалла NaCl показана на рис. 6.9.8. Регулярное расположение пятен свидетельствует о периодической структуре исследуемого вещества.

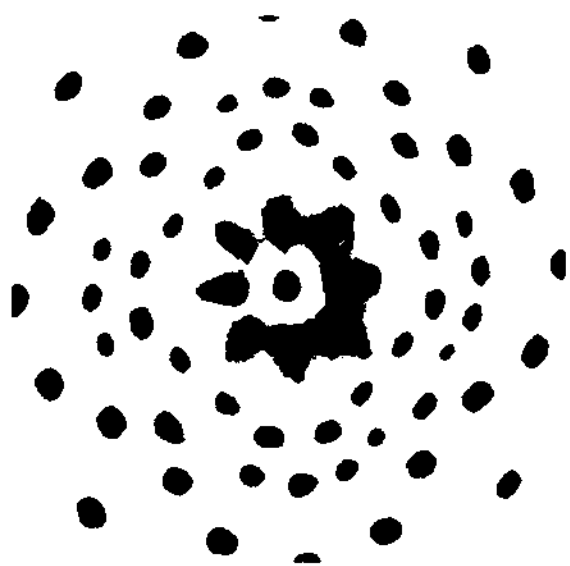


Рис. 6.9.8. Лауэграмма монокристалла NaCl

ГЛАВА 7

ОПТИЧЕСКИЕ И СПЕКТРАЛЬНЫЕ ПРИБОРЫ

7.1. Дифракционный предел разрешения оптических приборов

7.1.1. Дифракция Фраунгофера в оптических приборах

Пусть на линзу падает некоторая волна $A(x, y)$. Как было показано в разделе 6.6.2, в фокальной плоскости линзы поле пропорционально преобразованию Фурье функции $A(x, y)$. С другой стороны, если бы линзы не было, то на больших расстояниях от препятствия (отверстия) — в области дифракции Фраунгофера — поле также было бы пропорционально преобразованию Фурье поля, непосредственно идущего от препятствия (раздел 6.7). Следовательно, для наблюдения дифракции Фраунгофера достаточно на пути волны установить линзу и рассматривать дифракционную картину в её фокальной плоскости. Последняя будет подобной картине дифракции Фраунгофера. Данная идея позволяет рассчитывать световое поле в фокальной плоскости линзы, используя фраунгоферово приближение.

7.1.2. Критерий Рэлея оптического разрешения

Пусть имеются два некогерентных точечных источника излучения S_1 и S_2 (рис. 7.1.1). Интенсивность излучения в плоскости изображения Э (в фокальной плоскости линзы) равна сумме интенсивностей источников. С другой стороны, каждый из них создаёт дифракционное пятно (кружок Эйри) с полушириной (рис. 7.1.2, слева):

$$r_0 = \theta_d \cdot f = 1,22 \frac{\lambda}{D} f. \quad (7.1.1)$$

Здесь f и D — фокусное расстояние и диаметр линзы. Если пятна Эйри от источников сильно перекрываются, то изображения источников накладываются одно на другое, выглядят как одно пятно. В этом случае пара (S_1, S_2) может рассматриваться как один пространственно когерентный источник.

Критерий Рэлея определяет условие, при котором источники считаются различными. Для этого требуется, чтобы расстояние между цен-

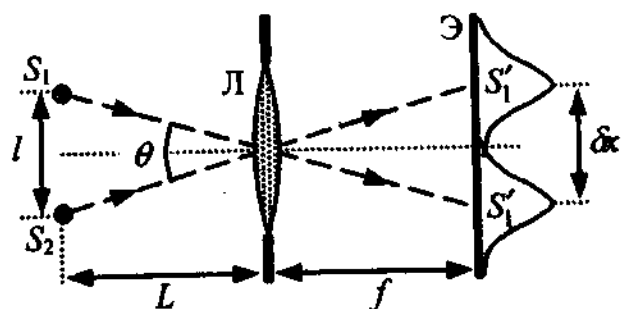


Рис. 7.1.1. Дифракционные пятна, создаваемые двумя источниками в фокальной плоскости Э линзы Л

трами дифракционных пятен было не меньше радиуса одного пятна (рис. 7.1.2, справа). Это значит, что в предельном случае максимум интенсивности одного пятна совпадает с минимумом интенсивности другого. При этом в точке провала суммарная интенсивность составляет 73,5% интенсивности в максимуме.

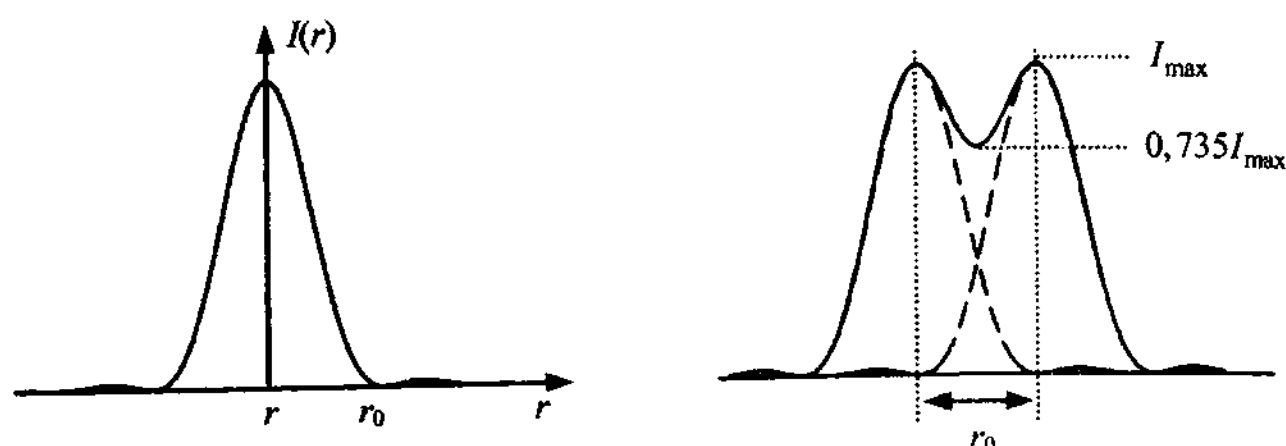


Рис. 7.1.2. Слева — дифракционное пятно от одного источника; справа — дифракционные пятна от двух источников, разрешаемых по критерию Рэлея

Критерий Рэлея указывает типичные способности глаза различать предметы. Разумеется, применение различных технических средств существенно расширяет возможности наблюдения и различения близко расположенных объектов.

Лорд Рэлей (Джон Уильям Стретт) — британский физик, являющийся одним из основоположников теории колебаний. Рэлей заложил основы теории молекулярного рассеяния света (закон Рэлея), в 1879 г. он создал теорию разрешающей способности оптических приборов (критерий Рэлея). В 1900 г. Рэлей открыл закон, в соответствии с которым энергия излучения абсолютно чёрного тела распределена по спектру в зависимости от его температуры (распределение Рэлея–Джинса). В 1904 г. за открытие аргона (совместно с Уильямом Рамзеем) он получил Нобелевскую премию по физике.

7.1.3. Разрешающая способность глаза, телескопа и микроскопа

Основным элементов всех этих приборов является линза. Поэтому достаточно рассмотреть, как для неё реализуется критерий Рэлея.

Поскольку на экране Э (рис. 7.1.1) центры изображений (точки S'_1 и S'_2) находятся друг от друга на расстоянии

$$\delta x = f\theta \quad (7.1.2)$$

(при $\theta \ll 1$), то две точки разрешаются данным оптическим прибором, если $\delta x \geq r_0$, или с учётом (7.1.1):

$$\theta f > 1,22 \frac{\lambda}{D} f \Rightarrow \theta > 1,22 \frac{\lambda}{D}. \quad (7.1.3)$$

Полученное неравенство даёт ограничение на угловое расстояние между наблюдаемыми источниками, при котором они воспринимаются как два различных (по критерию Рэлея).

Если точки S_1 и S_2 находятся на расстоянии L от линзы, то они разрешаются, если расстояние между ними

$$l = \theta L > 1,22 \frac{\lambda}{D} L. \quad (7.1.4)$$

Глаз

Средний диаметр зрачка глаза составляет $D \sim 3-4$ мм. При сильном освещении диаметр уменьшается до $D \sim 2$ мм, а при слабом — расширяется до $D \sim 8$ мм.

Принимая для оценки $D_{\text{глаз}} \sim 3$ мм, длину волны света $\lambda \sim 0,5$ мкм, получаем угловое разрешение глаза

$$\theta \sim 2 \cdot 10^{-4} \text{ рад} \approx 0,7'. \quad (7.1.5)$$

Фокусное расстояние зрачка глаза составляет $f \sim 15$ мм. Поэтому минимальное расстояние между изображениями на сетчатке двух разрешаемых глазом точек равно $\delta x \approx f \theta \sim 3 \cdot 10^{-3}$ мм. Заметим, что эта величина совпадает с характерным расстоянием между соседними светочувствительными элементами сетчатки.

Телескоп

Для телескопа с диаметром объектива $D_{\text{тел}}$ размер пятна Эйри в $D_{\text{тел}}/D_{\text{глаз}}$ меньше, чем для глаза. Поэтому угловое разрешение составит

$$\theta_{\text{тел}} = \theta_{\text{глаз}} \frac{D_{\text{глаз}}}{D_{\text{тел}}}. \quad (7.1.6)$$

В частности, при $D_{\text{тел}} = 3$ м находим

$$\theta_{\text{тел}} \sim 10^{-3} \theta_{\text{глаз}} \sim 2 \cdot 10^{-7} \text{ рад} = 0,04''. \quad (7.1.7)$$

Микроскоп

В микроскоп рассматриваются предметы, лучи от которых попадают в объектив под большим углом. При этом не все лучи, испускаемые объектом, попадают на линзу. Те из них, которые отвечают большим углом, проходят мимо и не принимают участие в формировании изображения. В результате в изображении не разрешаются мелкие детали.

Рассмотрим в качестве объекта дифракционную решётку с периодом d . Направления на главные дифракционные максимумы определяются условием

$$kd \sin \theta = 2\pi m, \quad m = 0, \pm 1, 2, \dots \quad (7.1.8)$$

Если решётка помещена в среду с показателем преломления n , то $k = (2\pi/\lambda)n$, где $\lambda = 2\pi c/\omega$ — длина волны излучения в вакууме. Соответственно направления на главные максимумы определяются из условия

$$d \sin \theta = m \frac{\lambda}{n}, \quad m = 0, \pm 1, \pm 2, \dots \quad (7.1.9)$$

Для установления характерных свойств предмета (в частности, нахождения периода решётки) нужно, чтобы в объектив, наряду с прямым светом, попали хотя бы несколько главных максимумов.

1. Нормальное падение света на решётку (рис. 7.1.3)

Предмет облучается светом, падающим по нормали. В этом случае в объектив должны попасть, по крайней мере, пучки, отвечающие дифракционным максимумам $(+1)$ -го и (-1) -го порядков. Если угловой радиус линзы, наблюдаемой из центра решётки, равен α , то максимально допустимые углы дифракции $\theta = \alpha$, и из (7.1.9) находим минимальный измеряемый период решётки:

$$d_{\min} = \frac{\lambda}{n \sin \alpha}. \quad (7.1.10)$$

Величина $n \sin \alpha$ называется *числовой апертурой* объектива.

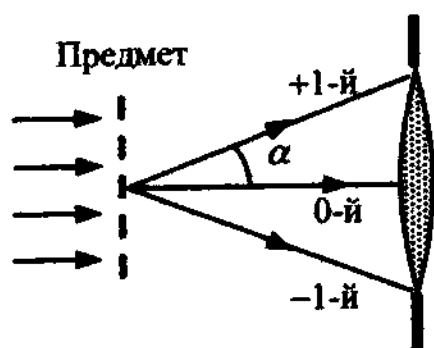


Рис. 7.1.3. Свет падает по нормали на дифракционную решётку и идёт от неё к объективу микроскопа, причём в него попадают только $(+1)$ -й, (-1) -й и 0-й главные максимумы

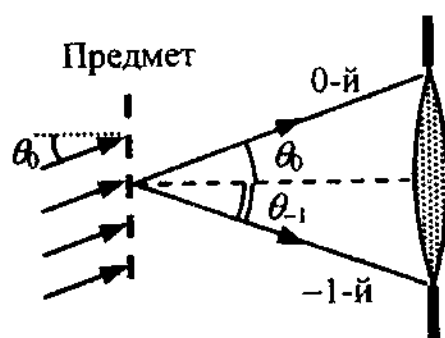


Рис. 7.1.4. Свет падает на дифракционную решётку под углом θ_0 и идёт от неё к объективу микроскопа, причём в него попадают только 0-й и (-1) -й главные максимумы

2. Наклонное падение света на решётку (рис. 7.1.4)

Для повышения разрешения микроскопа используют наклонное падение света на рассматриваемый предмет. Тогда в объектив должны попасть, по крайней мере, 0-й и 1-й дифракционные максимумы (для случая, показанного на рис. 7.2.4, это 0-й и (-1) -й максимумы).

Поскольку в данном случае направления на главные дифракционные максимумы определяются равенством

$$d(\sin \theta - \sin \theta_0) = m \frac{\lambda}{n}, \quad (7.1.11)$$

то вследствие симметрии пучков ($\theta = -\theta_0$) значению $m = -1$ отвечает условие $2d \sin \theta_0 = \lambda/n$. Полагая $\theta_0 = \alpha$, где α — угловой радиус линзы, получаем разрешение микроскопа:

$$d_{\min} = \frac{\lambda}{2nd \sin \alpha}. \quad (7.1.12)$$

Как видно из (7.1.10) и (7.1.12), заполнение пространства между предметом и объективом жидкостью (*иммерсия*) с показателем преломления $n > 1$ позволяет заметно повысить разрешение микроскопа. В качестве иммерсионной жидкости с большим значением n применяют минеральное масло ($n = 1,515$), водный раствор глицерина ($n = 1,434$), воду ($n = 1,333$), вазелиновое масло ($n = 1,503$) и другие.

7.2. Спектральные приборы

7.2.1. Основные характеристики

Как мы видели, величина дифракционного отклонения лучей, может сильно зависеть от длины волны света. Например, в случае дифракционной решётки (рис. 7.2.1) направление на главные максимумы даётся формулой

$$d \sin \theta = m\lambda, \quad (7.2.1)$$

или

$$\theta_m = \theta_m(\lambda) = \arcsin(m\lambda/d). \quad (7.2.2)$$

Отсюда следует, что чем больше длина волны, тем сильнее лучи отклоняются решёткой. При этом каждый главный максимум (кроме нулевого) будет выглядеть как полоса, окрашенная по-разному в разных своих частях (рис. 7.2.2). Это факт позволяет исследовать спектр анализируемого излучения, пространственно разделив его компоненты.

Для описания действия спектральных приборов используются различные характеристики. Рассмотрим их на примере дифракционной решётки.

1. Дисперсионная область — это максимальная ширина спектрального интервала $\Delta\lambda$, при которой спектры соседних порядков еще не пере-

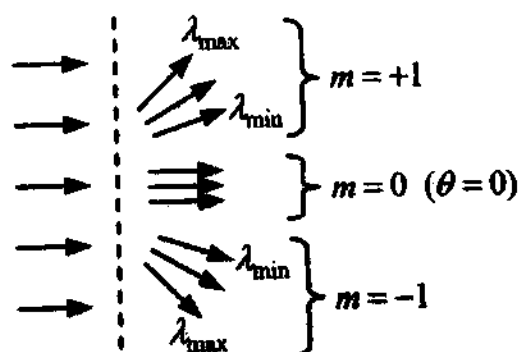


Рис. 7.2.1. Разложение света на спектральные компоненты дифракционной решёткой. Главный максимум $m = 0$ — ахроматическая полоса

крываются. Если же спектры начинают перекрываться, то данный прибор уже не позволяет исследовать данный интервал $\Delta\lambda$.

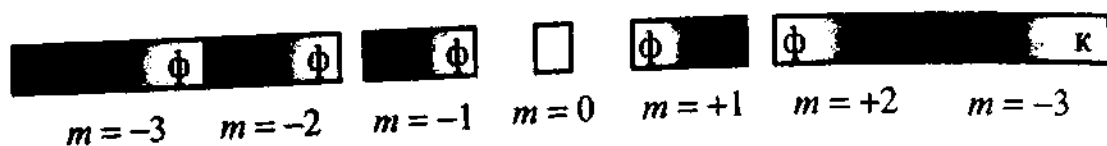


Рис. 7.2.2. Качественный вид дифракционной картины, создаваемой дифракционной решёткой, на которую падает свет с конечной шириной спектра $\Delta\lambda$. Буквами «к» и «ф» условно обозначены длинноволновая («красная») и коротковолновая («фиолетовая») границы спектрального интервала $\Delta\lambda$, число m — порядок главного максимума. В центре присутствует «белая» (ахроматическая) полоса

Найдём условие, при котором начинается перекрытие. Возьмём спектральный диапазон (λ, λ') , $\lambda' = \lambda + \Delta\lambda$. Главный максимум m -го порядка для границы спектра λ' имеет такое направление, что

$$d \sin \theta = m\lambda'. \quad (7.2.3)$$

Максимум $(m + 1)$ -го порядка для границы спектра λ должен распространяться под тем же углом θ , так что

$$d \sin \theta = (m + 1)\lambda. \quad (7.2.4)$$

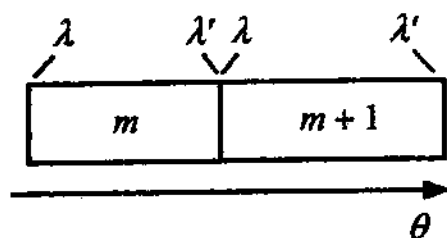
Это расположение спектров проиллюстрировано на рис. 7.2.3.

Из (7.2.3), (7.2.4) следует

$$(m + 1)\lambda = m\lambda' \Rightarrow m\Delta\lambda = \lambda \Rightarrow \Delta\lambda = \lambda/m. \quad (7.2.5)$$

Таким образом, спектры m -го и $(m + 1)$ -го порядков не перекрываются, если ширина спектрального диапазона достаточно мала:

$$\Delta\lambda \leq \lambda/m. \quad (7.2.6)$$



В частности, для видимого света:

$$400 \text{ нм} < \lambda < 760 \text{ нм}, \quad \Delta\lambda = 360 \text{ нм} \quad (7.2.7)$$

начинают перекрываться уже спектры первого и второго порядков:

$$m \leq \lambda/\Delta\lambda < 2. \quad (7.2.8)$$

2. Угловая дисперсия дифракционного прибора — это величина, определяемая равенством

$$D = d\theta/d\lambda. \quad (7.2.9)$$

Она определяет угловое расстояние $d\theta$ между спектральными линиями, отстоящими по длине волны на $d\lambda$:

$$\Delta\theta \approx D\Delta\lambda. \quad (7.3.10)$$

Для дифракционной решётки согласно (7.2.1) имеем

$$\sin \theta = m \frac{\lambda}{d} \Rightarrow D = \frac{m}{d \cos \theta}. \quad (7.2.11)$$

В частности, для малых углов дифракции отсюда следует

$$D = m/d. \quad (7.2.12)$$

Это значит, что дисперсия тем больше, чем выше порядок спектра m и чем меньше период решётки d .

3. Линейная дисперсия — это расстояние между разрешаемыми линиями спектра на экране, на котором рассматривается дифракционная картина (рис. 7.2.4). Отображая спектральные компоненты λ и $\lambda + \Delta\lambda$ в фокальной плоскости линзы, получаем линейное расстояние между ними на экране Δl . Формально линейная дисперсия определяется равенством

$$D_{\text{лин}} = dl/d\theta. \quad (7.2.13)$$

Поскольку $dl = fd\theta$, то

$$D_{\text{лин}} = f \cdot D \approx fm/d. \quad (7.2.14)$$

В последнем равенстве использовано приближённое выражение для угловой дисперсии (7.2.12), справедливое для малых углов дифракции.

4. Разрешающая способность спектрального прибора — это величина, определяемая равенством

$$R = \lambda/\Delta\lambda, \quad (7.2.15)$$

где $\Delta\lambda$ — минимальная разность длин волн двух спектральных линий, которые воспринимаются раздельно.

Найдём разрешающую способность дифракционной решётки. Воспользуемся критерием Рэлея (рис. 7.2.5). Пусть решётка имеет N штрихов. Выберем главный максимум m -го порядка для компоненты λ . Направление на первый дифракционный минимум даётся равенством

$$d \sin \theta_d = \left(m + \frac{1}{N}\right) \lambda. \quad (7.2.16)$$

Согласно критерию Рэлея компоненты считаются разрешёнными, если это же направление соответствует главному дифракционному максимуму для второй компоненты $\lambda + \Delta\lambda$:

$$d \sin \theta_d = m(\lambda + \Delta\lambda). \quad (7.2.17)$$

Из последних двух равенств находим

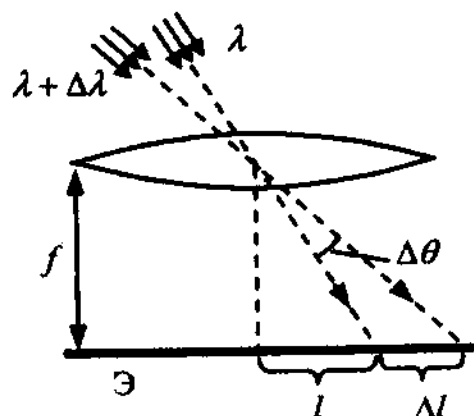


Рис. 7.2.4. Волны, отвечающие разным спектральным компонентам, отображаются с помощью линзы на экране Э в различные точки

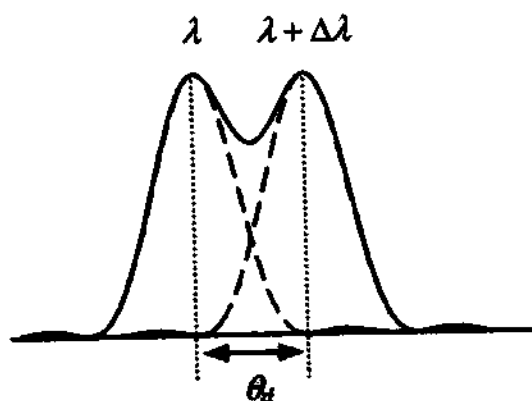


Рис. 7.2.5. Дифракционные пятна от двух компонент спектра, разрешаемых дифракционной решёткой (по критерию Рэлея). θ_d — угловой радиус главного дифракционного максимума

$$\left(m + \frac{1}{N}\right)\lambda = m(\lambda + \Delta\lambda) \Rightarrow \Delta\lambda = \frac{\lambda}{mN}. \quad (7.2.18)$$

Таким образом, приходим к следующему выражению для разрешающей способности дифракционной решётки:

$$R = \lambda/\Delta\lambda = mN. \quad (7.2.19)$$

Сравним действие решёток, отличающихся дисперсией и/или разрешающей способностью. На рис. 7.2.6 показано разрешение двух спектральных компонент тремя решётками: увеличение разрешения ведёт к уменьшению ширины линии, а увеличение дисперсии — к увеличению расстояния между линиями.

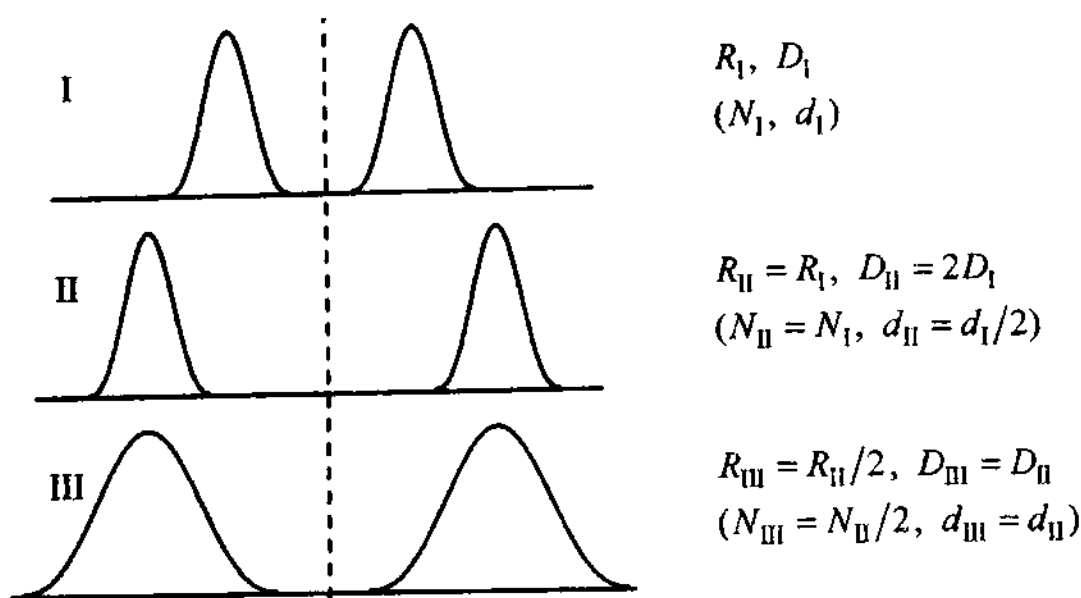


Рис. 7.2.6. Действие дифракционных решёток с различными значениями разрешающей способности R и угловой дисперсии D . Справа указано соотношение между этими параметрами для решёток I, II и III. В скобках приведены соотношения между параметрами сравниваемых решёток — числами штрихов (N) и периодами (d). $R = mN$, $D = m/d$

7.2.2. Разрешающая способность призмы

Призма разлагает падающий на неё свет в спектр — пространственно разделяет его компоненты (рис. 7.2.7). При этом чем меньше длина волны, тем сильнее отклоняются лучи: «красная» компонента отклоняется слабее, чем «фиолетовая». Поэтому призма может быть использована в качестве спектрального анализатора.



Рис. 7.2.7. Призма разлагает свет в спектр: чем меньше длина волны, тем сильнее отклонение лучей

Найдём разрешающую способность призмы. Предварительно заметим, что если бы показатель преломления вещества призмы был одинаковым для всех длин волн, то все компоненты отклонялись бы одинаково, и наблюдаемого разложения в спектр не было бы. Поэтому в данном случае существенно используется явление дисперсии — зависимости показателя преломления от частоты (длины волны) излучения.

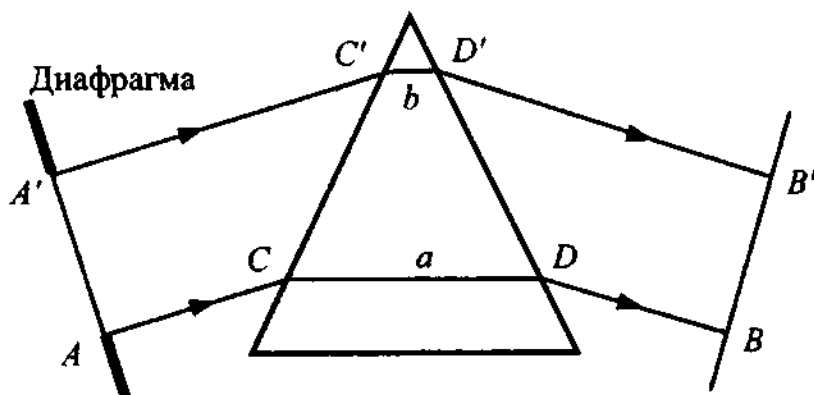


Рис. 7.2.8. Свет от диафрагмы идёт через призму. Для компоненты спектра λ' лучи $A'C'D'B'$ и $ACDB$ дают направление на главный максимум, а для компоненты λ — на дифракционный минимум

Пусть на призму падает свет с длиной волны λ' (рис. 7.2.8). Будем рассматривать симметричный ход лучей. Обозначим длину отрезка, проходящего в призме верхним лучом как b , а нижним лучом как a . Начальное положение волнового фронта представлено отрезком AA' , а положение после прохождения призмы — отрезком BB' . Это значит, что лучи перпендикулярны указанным отрезкам, а оптические длины лучей AB и $A'B'$ одинаковы:

$$\text{max: } [AC + n(\lambda')a + DB] - [A'C' + n(\lambda')b + D'B'] = 0. \quad (7.2.20)$$

Поскольку отрезок BB' — это участок волнового фронта, то данное направление можно рассматривать как направление на *дифракционный максимум* порядка $m = 0$ для длины волны λ' .

Вследствие дифракции на диафрагме появляются другие дифракционные максимумы и минимумы. Как и в случае дифракции на щели, направление на ближайший дифракционный минимум определяется тем условием, что оптическая разность хода волн крайних лучей равна длине волны.

Воспользуемся критерием Рэлея разрешения компонент λ' и λ : направление на дифракционный максимум для одной длины волны (λ') должно совпадать с направлением на *дифракционный минимум* для длины волны λ . Другими словами, для компоненты λ разность хода лучей AB и $A'B'$ теперь будет равна длине волны:

$$\min: [AC + n(\lambda)a + DB] - [A'C' + n(\lambda)b + D'B'] = \lambda. \quad (7.2.21)$$

Заметим, что отрезки AA' и BB' в последнем случае уже не представляют участков волнового фронта.

Вычитая почленно равенство (7.2.20) из (7.2.21), получим

$$[n(\lambda) - n(\lambda')]a - [n(\lambda) - n(\lambda')]b = \lambda. \quad (7.2.22)$$

Рассматривая близкие компоненты спектра ($\lambda' - \lambda \ll \lambda$) и обозначая $\Delta\lambda = \lambda' - \lambda$, находим

$$-(a - b) \frac{dn}{d\lambda} \Delta\lambda = \lambda,$$

откуда следует выражение для разрешающей способности призмы:

$$R = \frac{\lambda}{\Delta\lambda} = -(a - b) \frac{dn}{d\lambda}. \quad (7.2.23)$$

Максимальное разрешение достигается, если используется вся призма: $b = 0$, а величина a совпадает с длиной нижнего основания. Заметим также, что входящая сюда производная $dn/d\lambda$ должна быть отрицательной. Это имеет место в области *нормальной дисперсии* (подробно об этом будет сказано в разделе 10.2, посвящённом теории дисперсии).

В качестве примера укажем, что у оптического стекла — тяжёлого флинта — в жёлтой области спектра производная $dn/d\lambda \approx -956 \text{ см}^{-1}$. Поэтому призма из этого материала позволяет разрешить жёлтый дублет натрия:

$$\lambda_1 = 589,0 \text{ нм}, \lambda_2 = 589,6 \text{ нм}, \Delta\lambda = 0,6 \text{ нм},$$

если её основание равно

$$a = \frac{R}{|dn/d\lambda|} = \frac{\lambda/\Delta\lambda}{|dn/d\lambda|} = \frac{589,0 \cdot 10^{-7} \text{ см}}{(0,6 \cdot 10^{-7} \text{ см}) \cdot (956 \text{ см}^{-1})} = 1 \text{ см}.$$

7.3. Интерферометр Фабри–Перо

7.3.1. Устройство интерферометра Фабри–Перо

Интерферометр Фабри–Перо относится к числу оптических приборов, позволяющих анализировать спектры излучения с высоким разрешением и представляет собой систему двух тонких стеклянных пластинок, между которыми обычно находится воздух (рис. 7.3.1). Внутренние поверхности пластинок строго параллельны, тщательно отшлифованы и покрыты высокоотражающими слоями (материалами для покрытий служат золото, серебро, многослойные диэлектрические покрытия). Коэффициент отражения поверхностей достигает $\rho \sim 95\text{--}99\%$. Наружные поверхности пластинок обычно образуют небольшой угол с внутренними, чтобы отражён-

ный от них светлый блик не мешал наблюдению основной интерференционной картины.

Излучение от источника S , проходящее внутрь интерферометра, испытывает многократные отражения от пластин. Выходящее из интерферометра излучение собирается линзой в фокальной плоскости. В результате на экране наблюдается серия интерференционных полос (колец) равного наклона. Кольца тем уже, чем больше отражательная способность зеркал.

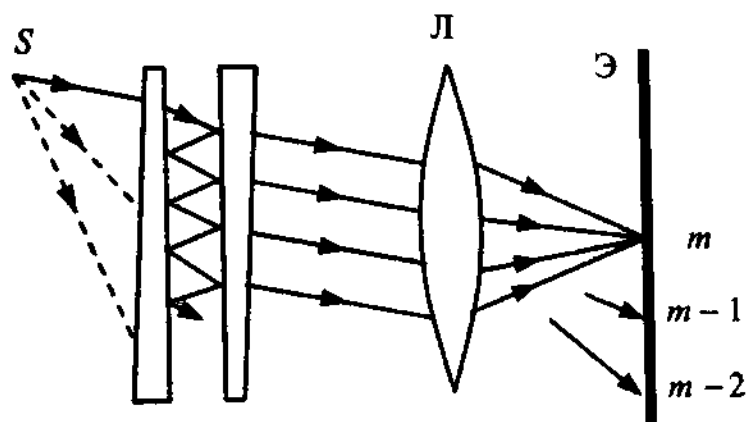


Рис. 7.3.1. Интерферометр Фабри-Перо

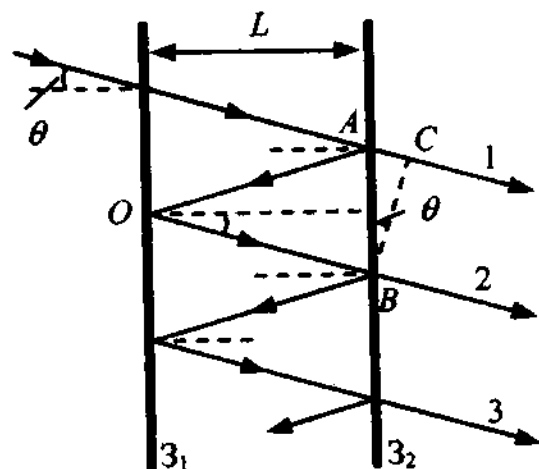


Рис. 7.3.2. К расчёту функции пропускания интерферометра Фабри-Перо

Для расчётов интерферометр можно представлять как плоскопараллельную воздушную пластинку, в которой происходят многократные отражения световых лучей (рис. 7.3.2).

7.3.2. Многолучевая интерференция

Выберем некоторый луч, падающий на интерферометр под углом θ к нормали. Как видно из рис. 7.3.2, этот луч, многократно отражаясь от зеркал, превращается в серию лучей, выходящих из интерферометра с противоположной стороны.

Разность хода лучей 1 и 2 (рис. 7.3.2) равна

$$\Delta = AOB - AC. \quad (7.3.1)$$

Поскольку, как видно из рисунка,

$$AOB = 2L/\cos \theta, \quad AB = 2L \operatorname{tg} \theta, \quad AC = AB \sin \theta = 2L \operatorname{tg} \theta \sin \theta,$$

то для разности хода получаем выражение

$$\Delta = 2L \cos \theta. \quad (7.3.2)$$

Таким образом, условие интерференционных максимумов в проходящем свете $\Delta = m\lambda$ имеет вид

$$2L \cos \theta = m\lambda, \quad m = 0, 1, 2, \dots \quad (7.3.3)$$

Если рассматривать отражённые от интерферометра волны, то условие интерференционных максимумов следует записывать в виде

$$2L \cos \theta = (m + 1/2)\lambda, \quad m = 0, 1, 2, \dots, \quad (7.3.4)$$

где учтено, что при отражении от оптически более плотной среды фаза волны меняется на π , что эквивалентно изменению разности хода на $\lambda/2$ (в проходящем свете имеет место двукратное отражение, не приводящее к изменению фазы). Соответственно условие (7.3.3) будет давать направления на интерференционные минимумы.

Таким образом, в интерферометре реализуется многолучевая интерференция.

7.3.3. Функция пропускания интерферометра

Найдём поле волны, прошедшей через интерферометр. Обозначим d и r соответственно амплитудные коэффициенты пропускания и отражения каждого из зеркал:

$$r = A_{\text{отр}} / A_{\text{пад}}, \quad d = A_{\text{прош}} / A_{\text{пад}}, \quad (7.3.5)$$

где $A_{\text{пад}}$, $A_{\text{отр}}$ и $A_{\text{прош}}$ — амплитуды падающей на зеркало, прошедшей и отражённой волн.

Пусть на интерферометр падает волна с амплитудой A_0 . В интерферометр через зеркало Z_1 войдёт волна $A_0 d$, а выйдет волна с амплитудой

$$A_{01} = A_0 d^2. \quad (7.3.6)$$

При этом часть волны, равная $r A_0 d$, отразится зеркалом Z_2 назад в интерферометр. После отражения от Z_1 волна изменит амплитуду: $A_0 r^2 d$. Пройдя затем через зеркало Z_2 , волна приобретёт амплитуду

$$A_{02} = A_0 d^2 r^2. \quad (7.3.7)$$

Третья волна, вышедшая из интерферометра через зеркало Z_2 , будет иметь амплитуду

$$A_{03} = A_0 d^2 r^4 \quad (7.3.8)$$

и так далее.

Помимо изменения амплитуды при отражениях от зеркал волны приобретают дополнительные фазы за счёт движения внутри интерферометра. Принимая фазу первой волны (A_1) за начало отсчёта, получаем следующий набор волн, вышедших из интерферометра:

$$A_1 = A_0 d^2, \quad A_2 = A_0 d^2 r^2 e^{i\delta}, \quad A_3 = A_0 d^2 r^4 e^{2i\delta}, \dots, \quad (7.3.9)$$

где

$$\delta = k\Delta = 2kL \cos \theta. \quad (7.3.10)$$

В результате суммарная волна будет даваться выражением

$$A = A_1 + A_2 + A_3 + \dots = A_0 d^2 (1 + r^2 e^{i\delta} + r^4 e^{2i\delta} + \dots). \quad (7.3.11)$$

Если угол падения θ мал, то имеется большое число N волн, вышедших из интерферометра. В этом пределе, полагая $N \rightarrow \infty$, получим

$$A = A_0 d^2 \frac{1}{1 - r^2 e^{i\delta}}. \quad (7.3.12)$$

Величина $\rho = r^2$ есть энергетический коэффициент отражения от одной пластины, а $d^2 = 1 - \rho$ — энергетический коэффициент прохождения через пластину. Имея это в виду, получаем функцию пропускания интерферометра:

$$D = \frac{I_{\text{прош}}}{I_{\text{пад}}} = \frac{|A|^2}{|A_0|^2} = \frac{(1 - \rho)^2}{1 - 2\rho \cos \delta + \rho^2}, \quad \delta = 2kL \cos \theta. \quad (7.3.13)$$

Как следует из последней формулы, пропускание максимально: $D = 1$, когда

$$\cos \delta = +1, \text{ или } \delta = 2\pi m, m = 1, 2, 3, \dots \quad (7.3.14)$$

Имея в виду определение (7.3.10) фазы δ , получаем

$$2L \cos \theta = m\lambda. \quad (7.3.14a)$$

Это совпадает с условием интерференционного усиления прошедших волн (7.3.3).

Пропускание интерферометра минимально, если

$$\cos \delta = -1, \text{ или } 2L \cos \theta_m = (m + 1/2)\lambda. \quad (7.3.15)$$

При этом

$$D_{\min} = \left(\frac{1 - \rho}{1 + \rho} \right)^2. \quad (7.3.16)$$

Если коэффициент отражения зеркал близок к единице ($1 - \rho \ll 1$), эта величина оказывается малой — интерферометр практически не прозрачен для падающего на него излучения.

Зависимость функции пропускания интерферометра от фазы δ показана на рис. 7.3.3.

Выразим фазовый сдвиг через частоту и длину волны:

$$\delta = 2kL \cos \theta = \frac{2\omega}{c} L \cos \theta = \frac{4\pi\nu}{c} L \cos \theta = \frac{4\pi L \cos \theta}{\lambda}, \quad (7.3.17)$$

где $\nu = \omega/2\pi$ — циклическая частота. Использовано выражение для длины волны $\lambda = c/\nu$. Соответственно условие резонансного пропускания интерферометра (7.3.14) можно записать в виде

$$\nu = \nu_m = \frac{c}{2L \cos \theta} m, \quad m = 1, 2, \dots \quad (7.3.18)$$

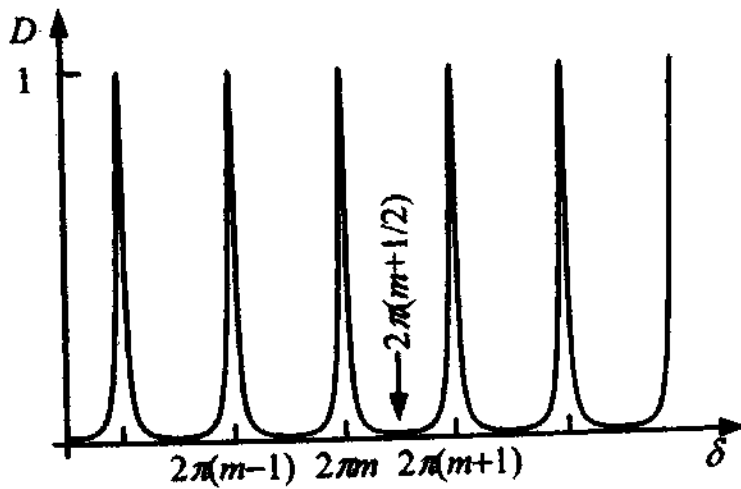


Рис. 7.3.3. Зависимость пропускания интерферометра Фабри-Перо от фазы $\delta = (4\pi L/\lambda) \cos \theta$

Для длины волны отсюда находим

$$\lambda = \lambda_m = \frac{2L \cos \theta}{m}, \quad m = 1, 2, \dots \quad (7.3.19)$$

Из формулы (7.3.18) видно, что если падающее излучение имеет широкий непрерывный спектр, то интерферометр пропускает отдельные эквидистантные (по частоте) узкие линии. Расстояние между центрами двух соседних пропускаемых линий равно

$$\nu_{m+1} - \nu_m = \frac{c}{2L \cos \theta}. \quad (7.3.20)$$

Найдём эффективную ширину отдельной линии. Ограничимся случаем $\theta = 0$. Рассмотрим m -ю полосу. В её окрестности положим

$$\delta = 2\pi m + \varepsilon \Rightarrow \cos \delta = \cos \varepsilon \approx 1 - \varepsilon^2/2, \quad |\varepsilon| \ll 1. \quad (7.3.21)$$

Тогда выражение для функции пропускания (7.3.13) принимает вид

$$D = \frac{(1 - \rho)^2}{1 - 2\rho + \rho^2 + \rho \varepsilon^2} = \frac{1}{1 + \rho \varepsilon^2 / (1 - \rho)^2}. \quad (7.3.22)$$

Выразим параметр ε через длину волны. Обозначим длину волны, отвечающую m -й линии резонансного пропускания ($\delta = 2\pi m$), как λ_m . Тогда, имея в виду равенство (7.4.14а), имеем

$$2\pi m = 4\pi L / \lambda_m. \quad (7.3.23)$$

В результате приходим к следующему выражению для параметра ε :

$$\begin{aligned} \varepsilon &= \delta - 2\pi m = \frac{4\pi L}{\lambda} - \frac{4\pi L}{\lambda_m} \approx \\ &\approx \frac{4\pi L \cos \theta}{\lambda_m^2} (\lambda_m - \lambda) = 2\pi m \frac{\lambda_m - \lambda}{\lambda_m}. \end{aligned} \quad (7.3.24)$$

Ширину линии определим по «уровню $\frac{1}{2}$ », то есть при таком значении длины волны, при котором $D(\lambda) = (1/2)D(\lambda_m) = 1/2$. Из (7.3.22) находим соответствующее условие:

$$\frac{\rho \epsilon^2}{(1-\rho)^2} = 1 \Rightarrow |\epsilon| = \frac{1-\rho}{\sqrt{\rho}}. \quad (7.3.25)$$

Отсюда и из (7.3.24) находим полуширину линии:

$$\frac{4\pi L}{\lambda_m^2}(\lambda_m - \lambda) = \frac{1-\rho}{\sqrt{\rho}} \Rightarrow \delta\lambda = \lambda - \lambda_m = \frac{\lambda_m^2}{4\pi L} \frac{1-\rho}{\sqrt{\rho}}. \quad (7.3.26)$$

Полная ширина линии вдвое больше: $\Delta\lambda = 2\delta\lambda$. Разрешающая способность интерферометра по критерию « $\frac{1}{2}$ » согласно (7.3.26) равна

$$R = \frac{\lambda}{\Delta\lambda} = \frac{\lambda}{2\delta\lambda} = \frac{2\pi L}{\lambda} \frac{\sqrt{\rho}}{1-\rho}. \quad (7.3.27)$$

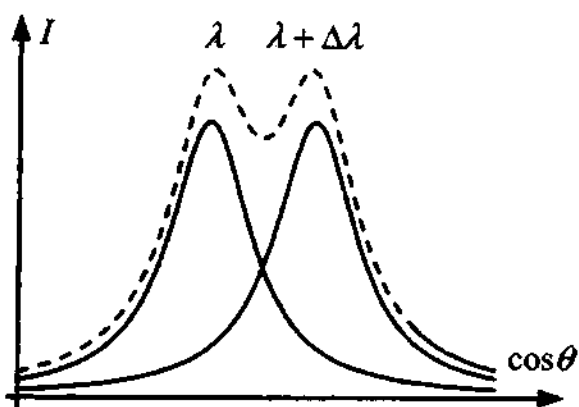


Рис. 7.3.4. Наложение интерференционных пиков от излучений с близкими длинами волн

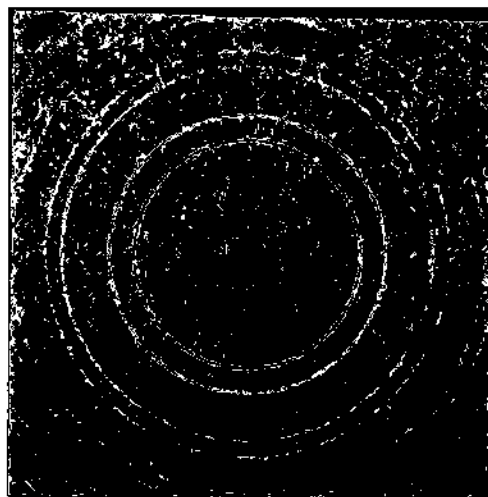


Рис. 7.3.5. Расщепление дифракционных колец, когда на интерферометр падает излучение, содержащее две близкие спектральные компоненты

Смысл данного критерия проиллюстрирован на рис. 7.3.4 и состоит в следующем. Пусть на интерферометр от источника падает излучение, содержащее две компоненты с длинами волн λ и $\lambda + \Delta\lambda$, отличающимися на полную ширину одной линии. Каждая из компонент создаёт полосы равного наклона (кольца), причём вследствие условия (7.3.14а): $\cos \theta_m = \lambda m / 2L$ направление на дифракционный максимум зависит от длины волны. Тогда в направлении на интерференционный максимум одной линии интенсивность сигнала с другой длиной волны составляет $I(\lambda_m + \Delta\lambda) \approx 0.2I_{\max}$, где $I_{\max}$ — интенсивность в максимуме одной из линий. Суммарная интенсивность посередине между линиями равна

$2(I_{\max}/2) = I_{\max}$. Следовательно, значение интенсивности в провале составляет

$$\frac{I_{\max}}{I_{\max} + 0.2I_{\max}} = 0,83 \quad (7.3.28)$$

от значения в максимуме (рис. 7.3.4).

Наблюдаемое расщепление интерференционных колец для излучения, содержащего две спектральные компоненты, показано на рис. 7.3.5.

7.3.4. Интерферометр как оптический резонатор

Резонатор — это колебательная система, устройство, способное накапливать энергию колебаний, поставляемую из внешнего источника.

Интерферометр Фабри–Перо является резонатором, способным накапливать энергию электромагнитного излучения. В качестве резонатора интерферометр входит в конструкцию лазеров, позволяя эффективно использовать энергию накачки активной среды (благодаря многократным проходам излучения). Кроме того, резонатор может применяться для выделения узких спектральных линий (рис. 7.3.6).

В этом разделе считаем, что волны распространяются перпендикулярно поверхностям, ограничивающим резонатор. Соответственно прибор прозрачен для длин волн излучения, удовлетворяющих (при нормальном падении) условию $m\lambda = 2L$ (рис. 7.3.6).

Найдём добротность резонатора, образованного двумя плоскими параллельными зеркалами с коэффициентом отражения ρ , близким к единице: $1 - \rho \ll 1$. Будем считать, что $L \gg \lambda$, где L — расстояние между зеркалами, λ — длина волны. Кроме того,

предполагаем, что потери энергии обусловлены только выходом излучения из резонатора вследствие частичной прозрачности зеркал.

Для нахождения добротности воспользуемся формулой

$$Q = 2\pi E_0 / \Delta E, \quad (7.3.29)$$

где E_0 — энергия, накопленная в резонаторе на некоторый момент времени, ΔE — энергия, теряемая за период колебаний поля.

В указанных предположениях с достаточной точностью можно считать, что в резонаторе устанавливается стоячая волна, эквивалентная суперпозиции двух одинаковых бегущих волн, распространяющихся на-

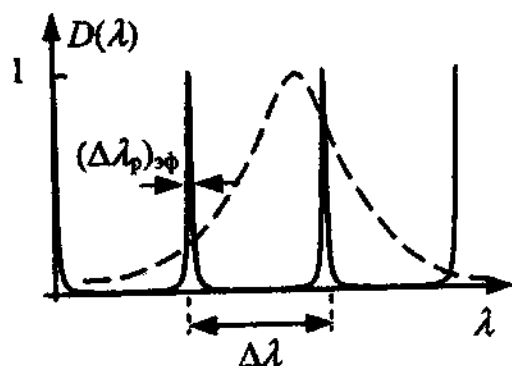


Рис. 7.3.6. Функция пропускания резонатора. Штриховой линией показан исходный контур линии, а сплошной линией — выделяемые резонатором узкие спектральные полосы

встречу друг другу. Если поток энергии в каждой из этих волн равен P , то энергия, накопленная в резонаторе, равна

$$E_0 = 2P\tau_L = 2PL/c, \quad (7.3.30)$$

где $\tau_L = L/c$ — время, за которое волна проходит расстояние L между пластинами, множитель «2» учитывает наличие двух волн. Поскольку мощность потерь составляет $(1 - \rho)2P$, то за период колебаний T теряется энергия

$$\Delta E = 2P(1 - \rho)T. \quad (7.3.31)$$

Поскольку $\lambda = cT$, то применение формулы (7.3.29) даёт

$$Q = 2\pi \frac{E_0}{\Delta E_0} = 2\pi \frac{L}{\lambda} \cdot \frac{1}{1 - \rho}. \quad (7.3.32)$$

Зная добротность, мы можем найти эффективную ширину линии излучения, выходящего из резонатора: $\Delta\lambda_{\text{эф}} = \lambda/Q$. Эта формула следует из определения добротности: $Q = \omega/\Delta\omega$, преобразованного с помощью равенства $\Delta\omega/\omega = \Delta\lambda/\lambda$, вытекающего, в свою очередь, из равенства $\omega = 2\pi c/\lambda$.

Заметим, что расстояние между пропускаемыми линиями (7.3.20), переписанное для длин волн ($\lambda = c/\nu$), имеет вид

$$\Delta\lambda = \lambda \frac{\Delta\nu}{\nu} = \lambda \frac{c}{2\nu L} = \frac{\lambda^2}{2L}. \quad (7.3.33)$$

В качестве примера укажем, что при значениях входящих в формулу (7.3.33) величин

$$L = 1 \text{ м}, \quad \lambda = 0,5 \text{ мкм}, \quad \nu = c/\lambda = 6 \cdot 10^{14} \text{ Гц}, \quad \rho = 0,95$$

добротность оказывается равной $Q = 2,6 \cdot 10^8$. Эффективная ширина линии выходящего из резонатора излучения составляет $\Delta\nu_{\text{эф}} = \nu/Q = 2,4 \cdot 10^6 \text{ Гц}$. Расстояние между соседними линиями равно $\Delta\nu = c/2L = 1,5 \cdot 10^8 \text{ Гц}$. Таким образом, выделяемые резонатором линии являются в высокой степени монохроматическими.

7.3.5. Энергия, запасаемая в интерферометре

Пусть на интерферометр с расстоянием между пластинами L падает плоская волна интенсивностью I_0 по нормали к ограничивающим поверхностям (рис. 7.3.7). Поглощением излучения в пластинах пренебрегаем.

Излучение, попадающее в интерферометр, многократно отражается от пластин и частично выходит наружу. В интерферометре распространяются

ся две встречные волны: $A_{\rightarrow}e^{ikz}$ и $A_{\leftarrow}e^{-ikz}$. Поскольку фазы этих волн зависят от координаты z , то они не когерентны. Поэтому складываться будут их энергии, а не амплитуды.

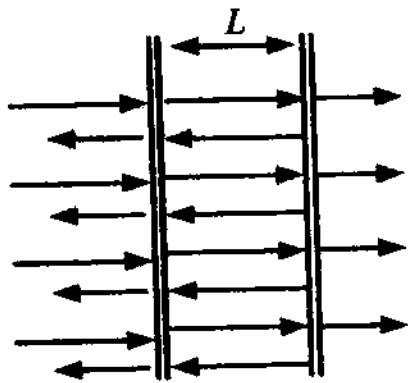


Рис. 7.3.7. На интерферометр (слева направо) падает плоская волна, которая частично отражается назад, а частично выходит справа

Найдём амплитуду волны в интерферометре, распространяющейся слева направо, в точках в непосредственной близости от *правой* пластины. Волна с амплитудой A_0 , войдя в интерферометр, приобретёт амплитуду A_0d . Дойдя до правой пластины, эта волна приобретёт фазовый множитель e^{ikL} и окажется равной $(A_{\rightarrow})_1 = A_0de^{ikL}$. После отражения от правой, а затем левой пластин волна окажется снова у правой пластины с амплитудой

$$(A_{\rightarrow})_2 = A_0de^{ikL}r^2e^{2ikL}. \quad (7.3.34)$$

Далее процесс многократно повторяется, так что полная амплитуда окажется равной

$$\begin{aligned} A_{\rightarrow} &= (A_{\rightarrow})_1 + (A_{\rightarrow})_2 + (A_{\rightarrow})_3 + \dots = \\ &= A_0de^{ikL} [1 + r^2e^{2ikL} + r^4e^{4ikL} + \dots] = \frac{A_0de^{ikL}}{1 - r^2e^{2ikL}}. \end{aligned} \quad (7.3.35)$$

Аналогично рассчитывается амплитуда волны, распространяющейся влево, в точках непосредственно у *левой* пластины. Поскольку первая из соответствующего набора волн возникает из волны A_0d после однократного отражения (от правой пластины) и двукратного прохождения по интерферометру, то получаем

$$\begin{aligned} A_{\leftarrow} &= (A_{\leftarrow})_1 + (A_{\leftarrow})_2 + (A_{\leftarrow})_3 + \dots = \\ &= A_0d [re^{2ikL} + r^3e^{4ikL} + \dots] = \frac{A_0dre^{2ikL}}{1 - r^2e^{2ikL}}. \end{aligned} \quad (7.3.36)$$

Таким образом, интенсивность волн внутри интерферометра равна

$$\begin{aligned} I_{\rightarrow} &= I_0 \left| \frac{d}{1 - r^2e^{2ikL}} \right|^2 = I_0 \frac{1 - \rho}{1 - 2\rho \cos(2kL) + \rho^2}, \\ I_{\leftarrow} &= I_0 \left| \frac{dre^{2ikL}}{1 - r^2e^{2ikL}} \right|^2 = I_0 \frac{(1 - \rho)\rho}{1 - 2\rho \cos(2kL) + \rho^2}. \end{aligned} \quad (7.3.37)$$

Здесь учтены соотношения

$$\rho = r^2, \quad d^2 = 1 - r^2 = 1 - \rho. \quad (7.3.38)$$

Полная интенсивность излучения в интерферометре равна

$$I = I_{\rightarrow} + I_{\leftarrow} = I_0 \left[\frac{1 - \rho}{1 - 2\rho \cos(2kL) + \rho^2} + \frac{(1 - \rho)\rho}{1 - 2\rho \cos(2kL) + \rho^2} \right] =$$

$$= I_0 \frac{1 - \rho^2}{1 - 2\rho \cos(2kL) + \rho^2}. \quad (7.3.39)$$

Амплитуда и интенсивность волны, вышедшей справа от интерферометра, равны

$$A_{\text{прош}} = A_{\rightarrow} d = \frac{A_0 d^2 e^{ikL}}{1 - r^2 e^{2ikL}}, \quad (7.3.40)$$

$$I_{\text{прош}} = I_0 |A_{\text{прош}}|^2 = I_0 \frac{(1 - \rho)^2}{1 - 2\rho \cos(2kL) + \rho^2}.$$

Мы пришли к выражению для функции пропускания интерферометра, полученному в разделе 7.3.3:

$$D = \frac{I_{\text{прош}}}{I_0} = \frac{(1 - \rho)^2}{1 - 2\rho \cos(2kL) + \rho^2}. \quad (7.3.41)$$

По закону сохранения энергии интенсивность излучения, отражённого от интерферометра, равна

$$I_{\text{отр}} = (1 - D)I_0 = 2\rho I_0 \frac{1 - \cos(2kL)}{1 - 2\rho \cos(2kL) + \rho^2}. \quad (7.3.42)$$

Рассмотрим частные случаи.

1) $\cos(2kL) = +1$, или $L = m\lambda/2$, $m = 1, 2, \dots$. В этом случае

$$I = I_0 \frac{1 - \rho^2}{(1 - \rho)^2} = I_0 \frac{1 + \rho}{1 - \rho}. \quad (7.3.43)$$

2) $\cos(2kL) = -1$, или $L = (2m - 1)\lambda/4$, $m = 1, 2, \dots$. В этом случае

$$I = I_0 \frac{1 - \rho^2}{(1 + \rho)^2} = I_0 \frac{1 - \rho}{1 + \rho}. \quad (7.3.44)$$

Зная интенсивность излучения, можно найти полную энергию поля в интерферометре:

$$E = Vw = VI/c,$$

где $w = I/c$ — плотность энергии, $V = SL$ — объём интерферометра, S — площадь пластин.

Таким образом, в состоянии с минимальным пропусканием интерферометр накапливает минимальное количество энергии. Если же пропускание максимально, то запасаемая в интерферометре энергия максимальна.

ГЛАВА 8

ПРИНЦИПЫ ФУРЬЕ-ОПТИКИ И ГОЛОГРАФИИ

8.1. Метод Рэлея

Пусть имеется некоторое волновое поле $S(\mathbf{r}, t)$ с частотой ω

$$S(\mathbf{r}, t) = A(\mathbf{r})e^{-i\omega t}. \quad (8.1.1)$$

Амплитуда $A(\mathbf{r})$ удовлетворяет уравнению Гельмгольца:

$$\Delta A + k^2 A = 0, \quad k^2 = \omega^2 / c^2 \quad (8.1.2)$$

(в этой главе мы ограничимся случаем, когда диэлектрическая проницаемость среды, в которой распространяются волны, равна единице).

Одной из целей волновой оптики является решение задач дифракции, расчёт волновых полей, изменённых разного рода препятствиями: экранами, отверстиями и так далее. Это значит, что нужно решить уравнение Гельмгольца (8.1.2) с граничным условием, заданным на некоторой граничной поверхности Σ :

$$A|_{\Sigma} = A_0(\mathbf{r}). \quad (8.1.3)$$

В главе 6 мы рассмотрели один из подходов к решению поставленной задачи — метод Гюйгенса–Френеля. Альтернативой ему является метод Рэлея. В ряде случаев он позволяет существенно упростить расчёты или сделать их более наглядными.

Отличие двух указанных подходов состоит в следующем. Метод Гюйгенса–Френеля основан на вычислении интеграла Френеля. При этом волновое поле представляется в виде суперпозиции *сферических волн* с центрами в точках исходного волнового фронта. В методе Рэлея волновое поле рассматривается как суперпозиция *плоских волн*. Рассмотрим подробнее суть данного метода.

Частное решение уравнения (8.1.2) есть плоская волна:

$$S(\mathbf{r}, t) = A_0 e^{i(\mathbf{k}\mathbf{r} - \omega t)}, \quad A(\mathbf{r}) = A_0 e^{i\mathbf{k}\mathbf{r}}, \quad (8.1.4)$$

причём

$$\mathbf{k}^2 = k_x^2 + k_y^2 + k_z^2 = k^2. \quad (8.1.5)$$

Для упрощения записи формул в этой главе мы будем пользоваться следующими обозначениями для компонент волнового вектора:

$$k_x = u, \quad k_y = v, \quad k_z = w, \quad k^2 = u^2 + v^2 + w^2. \quad (8.1.5a)$$

Пусть теперь поле $A(\mathbf{r})$ задано при $z = 0$:

$$A|_{z=0} = A_0 e^{i(ux+vy)} \equiv A_S(x, y). \quad (8.1.6)$$

Требуется найти это поле всюду при $z > 0$.

Согласно (8.1.4), (8.1.5) уравнению Гельмгольца (8.1.2) с граничным условием (8.1.6) удовлетворяет функция

$$A|_{z>0} = A_0 e^{iux+ivy+iwz} \equiv T_z A_S, \quad (8.1.7)$$

где величина T_z даётся выражением

$$T_z = e^{iwz}, \quad w = \sqrt{k^2 - u^2 - v^2}. \quad (8.1.8)$$

В силу теоремы единственности мы нашли решение поставленной задачи.

Введённая равенством (8.1.7) величина T_z называется *передаточной функцией* (или *частотной характеристикой*) *свободного пространства*. Она определяет изменение фазы плоской волны, прошедшей путь z от граничной поверхности $z = 0$.

Выберем систему координат так, чтобы волновой вектор имел только две компоненты:

$$\mathbf{k} = (u, w), \quad w = \sqrt{k^2 - u^2}. \quad (8.1.9)$$

Если волна распространяется под углом α к оси z , то

$$u = k \sin \alpha, \quad w = k \cos \alpha. \quad (8.1.10)$$

Следовательно, такую волну можно записать в виде

$$A(x, z) = A_0 e^{i(kx \sin \alpha + kz \cos \alpha)}. \quad (8.1.11)$$

Таким образом, зная амплитуду и фазу плоской волны на граничной поверхности $z = 0$, то есть функцию $A_S(x, y)$ в (8.1.6), мы можем, используя величину T_z , восстановить эту волну всюду в области $z > 0$.

Рассмотрим теперь более общий случай. Согласно теореме Фурье всякая интегрируемая функция представима в виде ряда или интеграла Фурье. Соответственно запишем поле $S(\mathbf{r}, t)$ как сумму фурье-гармоник:

$$S(\mathbf{r}, t) = \sum_n a_n e^{i(u_n x + v_n y + w_n z - \omega t)}, \quad (8.1.12)$$

где суммирование производится по всем допустимым значениям чисел u_n, v_n, w_n . Введём вектор $\mathbf{k}$, компонентами которого являются эти числа: $\mathbf{k} = (u, v, w)$. Если функция $S(\mathbf{r}, t)$ периодическая по координатам x, y, z , то в (8.1.12) входит сумма, в которой компоненты вектора $\mathbf{k}$ пробегают дискретный набор значений:

$$S(\mathbf{r}, t) = \sum_n a_n e^{i(\mathbf{k}_n \mathbf{r} - \omega t)}. \quad (8.1.13)$$

Если же $S(\mathbf{r}, t)$ — непериодическая функция координат, то компоненты вектора $\mathbf{k}$ пробегает непрерывный набор значений, и вместо суммы в (8.1.13) следует писать интеграл:

$$S(\mathbf{r}, t) = \int a(\mathbf{k}) e^{i(\mathbf{k}\mathbf{r} - \omega t)} \frac{d^3 k}{(2\pi)^3}. \quad (8.1.14)$$

Будем сейчас для определённости пользоваться представлением (8.1.13). Это выражение можно интерпретировать таким образом, что поле $S(\mathbf{r}, t)$ есть суперпозиция плоских волн

$$a_n e^{i(\mathbf{k}_n \mathbf{r} - \omega t)} \quad (8.1.15)$$

разных направлений, задаваемых волновыми векторами $\mathbf{k}_n$, каждая из которых удовлетворяет уравнению Гельмгольца (8.1.2). Соответственно вследствие условия $\mathbf{k}^2 = k^2$ компонента $k_z \equiv w$ волнового вектора определится из (8.1.5):

$$w_n = \pm \sqrt{k^2 - u_n^2 - v_n^2}. \quad (8.1.16)$$

Если рассматриваются только волны, распространяющиеся в область $z > 0$, то перед корнем следует выбирать знак «+».

Пусть волновое поле задано на граничной поверхности $z = 0$ и требуется найти его всюду в области $z > 0$. Иными словами, требуется решить краевую задачу

$$\begin{aligned} \Delta A + k^2 A &= 0, \quad z > 0, \\ A|_{z=0} &= A_S(x, y). \end{aligned} \quad (8.1.17)$$

Запишем функцию $A_S(x, y)$ в виде

$$A_S(x, y) = \sum_n a_n e^{i(u_n x + v_n y)}. \quad (8.1.18)$$

Тогда решение задачи (8.1.17) можно представить в виде суммы

$$A = \sum_n A_n, \quad (8.1.19)$$

слагаемые в которой являются решениями краевых задач

$$\begin{aligned} \Delta A_n + k^2 A_n &= 0, \quad z > 0, \\ A_n|_{z=0} &= a_n e^{i(u_n x + v_n y)}. \end{aligned} \quad (8.1.20)$$

В соответствии с (8.1.7) имеем

$$A_n|_{z>0} = a_n e^{i(u_n x + v_n y + w_n z)}, \quad w_n = \sqrt{k^2 - u_n^2 - v_n^2}. \quad (8.1.21)$$

В результате мы получаем решение задачи (8.1.17):

$$A = \sum_n a_n \exp \left(i u_n x + i v_n y + i \sqrt{k^2 - u_n^2 - v_n^2} z \right). \quad (8.1.22)$$

Как видно из сказанного, проблема сводится к нахождению коэффициентов Фурье a_n в разложении функции $A_S(x, y)$ — их подстановка в формулу (8.1.22) даёт искомое решение.

Выражения (8.1.18), (8.1.22) называются *пространственным разложением Фурье*.

Числа u_n, v_n в (8.1.22) называются *пространственными частотами*. Пары чисел (u_n, v_n) однозначно определяют направление распространения отдельных плоских волн, поскольку третья компонента волнового вектора (w_n) однозначно определяется по значениям первых.

Изложенный подход к нахождению волнового поля, основанный на представлении волнового поля в виде суперпозиции плоских волн, называют *методом Рэлея*. Рассмотрим его реализацию для случая, когда некоторая волна падает из области $z < 0$ на экран в плоскости $z = 0$, содержащий отверстия, непрозрачные участки и фрагменты, меняющие фазу. Требуется найти волну в области $z > 0$.

Введём *функцию пропускания* экрана D следующим образом. Пусть слева ($z < 0$) на экран падает волна, причём в точках экрана (слева от него) её амплитуда равна $A|_{z=-0} \equiv A_{\text{пад}}(x, y)$. Тогда непосредственно за экраном амплитуда прошедшей волны $A_{\text{прош}} \equiv A|_{z=+0}$ даётся выражением

$$A_S \equiv A_{\text{прош}}(x, y) = D(x, y) A_{\text{пад}}(x, y). \quad (8.1.23)$$

Данное равенство называется *граничным условием Кирхгофа*.

Согласно сказанному для нахождения волнового поля в области $z > 0$ требуется решить уравнение Гельмгольца (8.1.2) с граничным условием (8.1.23), заданным при $z = 0$.

Условие (8.1.23) подразумевает, что препятствие представляет собой *тонкий экран*, который можно считать целиком расположенным в плоскости $z = 0$.

Заметим, что функция пропускания может быть комплексной величиной, что позволяет учесть изменение не только амплитуды, но и фазы волны, преодолевающей экран.

Строго говоря, условие (8.1.23) является приближённым, поскольку учитывает только действие экрана на поле, при котором изменяются амплитуда и фаза. Однако само поле может воздействовать на экран, благодаря чему появляются дополнительные источники полей. Вместе с тем такой эффект обычно проявляется лишь на малых расстояниях от экрана,

сопоставимых с длиной волны излучения. Для расчёта же полей на больших расстояниях условие (8.1.23) работает, как правило, удовлетворительно.

В соответствии со сказанным выше для нахождения поля в области $z > 0$ достаточно найти коэффициенты Фурье a_n разложения функции $A_S(x, y) \equiv A_{\text{прош}}(x, y)$, после чего их подстановка в формулу (8.1.22) даёт искомое решение.

Слагаемое $a_n e^{ik_n r}$ в сумме (8.1.22) описывает плоскую волну, если число $w_n = \sqrt{k^2 - u_n^2 - v_n^2}$ действительное. Соответствующая волна называется *однородной*. Если же для каких-либо слагаемых окажется $w_n^2 < 0$, то такие слагаемые затухают с ростом z :

$$\begin{aligned} a_n \exp\left(iu_n x + iv_n y + i\sqrt{k^2 - u_n^2 - v_n^2} z\right) = \\ = a_n \exp(iu_n x + iv_n y) \exp\left(-\sqrt{u_n^2 + v_n^2 - k^2} z\right). \end{aligned} \quad (8.1.24)$$

Такие волны называются *неоднородными*. На больших расстояниях от граничной поверхности данные волны вследствие их затухания не участвуют в формировании поля и учитываться не должны.

Мы представили волновое поле в виде суперпозиции плоских волн, каждая из которых соответствует какому-то слагаемому в разложении Фурье поля. Используя оптические элементы (например, линзы, призмы, фильтры), можно пространственно разделять фурье-компоненты изучаемого поля (плоские волны) и, тем самым, осуществлять его преобразование, избирательно воздействуя на отдельные компоненты Фурье (то есть коэффициенты $a(k)$ в разложении (8.1.12)). Сформулированный подход называется *фурье-оптикой* и имеет многочисленные применения в задачах анализа и обработки оптических сигналов.

Идеи спектрального разложения получили широкое распространение в оптике благодаря работам Дж. Рэлея и Э. Аббе.

8.2. Дифракция на периодических структурах

8.2.1. Дифракция на синусоидальной решётке

Рассмотрим применения метода Рэлея на простейших задачах дифракции плоской волны на периодических структурах. Ограничимся сейчас случаем, когда волновые векторы всех волн лежат в плоскости (x, z) .

1. *Амплитудной синусоидальной решёткой* называется объект с функцией пропускания

$$D(x) = D_0 (1 + m \cos \Omega x). \quad (8.2.1)$$

Период этой решётки равен

$$d = 2\pi/\Omega. \quad (8.2.2)$$

Имеется естественное ограничение на значения параметров D_0 и m :

$$D_0(1+m) \leq 1, \quad (8.2.3)$$

определяемое тем, что прохождение тонкого экрана не приводит к локальному увеличению амплитуды поля.

Пусть на решётку (8.2.1) слева падает плоская волна под углом θ_0 (рис. 8.2.1):

$$A = A_0 e^{ik_0 r}, \quad u_0 = k \sin \theta_0, \quad w_0 = k \cos \theta_0. \quad (8.2.4)$$

Для нахождения поля в области $z > 0$ воспользуемся граничным условием Кирхгофа:

$$A_S = D(x) A|_{z=0} = A_0 D(x) e^{iu_0 x}. \quad (8.2.5)$$

Имея в виду выражение (8.2.1) для функции пропускания, получаем пространственное разложение Фурье функции A_S :

$$A_S = A_0 D_0 e^{iu_0 x} + \frac{m}{2} A_0 D_0 e^{i(u_0 + \Omega)x} + \frac{m}{2} A_0 D_0 e^{i(u_0 - \Omega)x}. \quad (8.2.6)$$

В соответствии со сказанным в разделе 8.1 каждое из слагаемых в этой сумме порождает плоскую волну при $z > 0$. Следовательно, получаем

$$\begin{aligned} A(x, z) &= A_0(x, z) + A_{+1}(x, z) + A_{-1}(x, z), \\ A_0(x, z) &= A_0 D_0 \exp[iu^{(0)}x + iw^{(0)}z], \\ A_{+1}(x, z) &= \frac{m}{2} A_0 D_0 \exp[iu^{(+1)}x + iw^{(+1)}z], \\ A_{-1}(x, z) &= \frac{m}{2} A_0 D_0 \exp[iu^{(-1)}x + iw^{(-1)}z]. \end{aligned} \quad (8.2.7)$$

Здесь введены обозначения

$$\begin{aligned} u^{(0)} &= u_0, \quad w^{(0)} = \sqrt{k^2 - u_0^2}, \\ u^{(+1)} &= u_0 + \Omega, \quad w^{(+1)} = \sqrt{k^2 - (u_0 + \Omega)^2}, \\ u^{(-1)} &= u_0 - \Omega, \quad w^{(-1)} = \sqrt{k^2 - (u_0 - \Omega)^2}. \end{aligned} \quad (8.2.8)$$

Пространственные частоты $u^{(0)}, u^{(+1)}, u^{(-1)}$ определяют направления распространения волн справа от решётки ($z > 0$). Полагая для этих волн $u = k \sin \theta$, записываем уравнения для определения их направлений:

$$k \sin \theta^{(0)} = k \sin \theta_0, \quad k \sin \theta^{(+1)} = k \sin \theta_0 + \Omega, \quad k \sin \theta^{(-1)} = k \sin \theta_0 - \Omega.$$

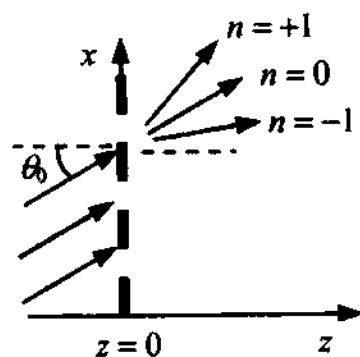


Рис. 8.2.1. Плоская волна падает под углом θ_0 на дифракционную решётку

Вспоминая равенства $k = 2\pi/\lambda$, $\Omega = 2\pi/d$, получаем

$$d(\sin \theta - \sin \theta_0) = n\lambda, \quad n = 0, \pm 1. \quad (8.2.9)$$

Мы пришли к известной формуле, указывающей направления на главные максимумы при дифракции на решётке. Заметим, что в рассматриваемом случае имеется всего три максимума. Это связано прежде всего с тем, что период решётки d и эффективная ширина щели связаны соотношением $b_{\text{eff}} = d/2$.

2. *Фазовой синусоидальной решёткой* называется объект с функцией пропускания

$$D = e^{im \cos \Omega x}, \quad \Omega = 2\pi/d. \quad (8.2.10)$$

Этот объект, не меняя амплитуды падающей на него волны, изменяет её фазу.

Пусть на решётку (8.2.10) падает слева плоская волна (рис. 8.2.1):

$$A = A_0 e^{ik_0 x}. \quad (8.2.11)$$

Найдём поле справа от решётки. Используем граничное условие

$$A|_{z=+0} = D(x) A|_{z=-0},$$

или

$$A|_{z=+0} = A_0 e^{iu_0 x + im \cos \Omega x}, \quad u_0 = k \sin \theta_0. \quad (8.2.12)$$

Обычно рассматривают фазовые решётки с малыми модуляциями фазы: $m \ll 1$. В этом случае получаем

$$\begin{aligned} A|_{z=+0} &= A_0 e^{iu_0 x} (1 + im \cos \Omega x) = \\ &= A_0 e^{iu_0 x} + \frac{im}{2} A_0 e^{i(u_0 + \Omega)x} + \frac{im}{2} A_0 e^{i(u_0 - \Omega)x}. \end{aligned} \quad (8.2.13)$$

Следовательно, от плоскости $z = +0$ исходят три волны, которым отвечают пространственные частоты

$$u = u_0 + n\Omega, \quad n = 0, \pm 1. \quad (8.2.14)$$

Полагая

$$u = k \sin \theta, \quad u_0 = k \sin \theta_0, \quad k = 2\pi/\lambda,$$

находим уравнение, определяющее направления волн:

$$d(\sin \theta - \sin \theta_0) = n\lambda, \quad n = 0, \pm 1. \quad (8.2.15)$$

Таким образом, в пределе $m \ll 1$ остаются, как и в случае амплитудной синусоидальной решётки, три волны. Разница состоит в том, что в случае фазовой решётки фазы волн, отвечающих главным максимумам $n = \pm 1$, имеют дополнительный сдвиг на $\pi/2$ (благодаря наличию множителя i во втором и третьем слагаемых в (8.2.13)).

Заметим, что если провести разложение (8.2.12) до более высоких степеней параметра m , то получим волны, распространяющиеся под углами, отвечающими значениям $n = \pm 2, \pm 3, \dots$. Ограничение на величину n при этом даётся, как и в случае обычной дифракционной решётки, условием $n \leq d/\lambda$.

8.2.2. Эффект саморепродукции

Рассмотрим экран с периодически расположенными отверстиями (например, дифракционную решетку). Будем для простоты говорить только о структуре на экране, периодичной вдоль оси x и однородной вдоль оси y . Обобщение на случай периодичности и вдоль оси y не представляет труда.

Пусть слева на экран падает плоская волна вдоль оси z (рис. 8.2.2):

$$A(x, z) = A_0 e^{ikz}, \quad z < 0. \quad (8.2.16)$$

Вследствие периодичности структуры экрана справа от него при $z = +0$ поле может быть представлено в виде ряда Фурье:

$$A_S(x) \equiv A(x, z)|_{z=+0} = D(x) A(x, z)|_{z=-0} = \sum_{n=0}^{\infty} a_n e^{iu_n x}, \quad (8.2.17)$$

где

$$u_n = n\Omega, \quad \Omega = 2\pi/d, \quad (8.2.18)$$

d — период структуры (вдоль оси x). Каждое слагаемое в сумме (8.2.2) порождает в области $z > 0$ волну

$$a_n e^{iu_n x + iw_n z}, \quad w_n = \sqrt{k^2 - u_n^2}, \quad (8.2.19)$$

так что

$$A(x, z)|_{z>0} = \sum_{n=0}^{\infty} a_n e^{iu_n x + iw_n z}. \quad (8.2.20)$$

Если период структуры $d \gg \lambda$, то есть $\Omega \ll k$, то для не слишком больших номеров гармоник n можно считать $u_n \ll k$ и записать:

$$w_n \approx k - u_n^2/2k. \quad (8.2.21)$$

Соответственно выражение для n -й волны принимает вид

$$a_n e^{ikz} \exp\left(iu_n x - i\frac{u_n^2}{2k} z\right). \quad (8.2.22)$$

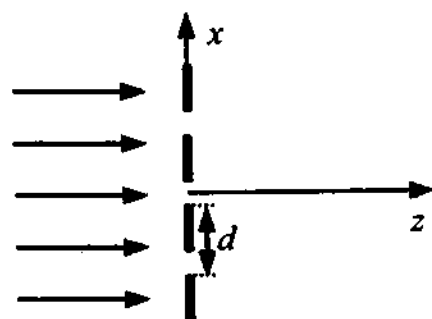


Рис. 8.2.2. Свет падает (вдоль оси z) на периодическую (с периодом d) структуру, ориентированную перпендикулярно оси z

Если на некотором расстоянии z от экрана (периодической структуры) будет выполнено условие

$$\frac{u_n^2}{2k} z = 2\pi p \quad (8.2.23)$$

(p — целое число), или

$$n^2 \frac{\Omega^2}{2k} z = 2\pi p, \quad (8.2.24)$$

то в соответствии с (8.2.22), (8.2.23) окажется

$$\begin{aligned} A(x, z) &= e^{ikz} \sum_{n=0}^{\infty} a_n \exp\left(iu_n x - i \frac{u_n^2}{2k} z\right) = \\ &= e^{ikz} \sum_{n=0}^{\infty} a_n \exp(iu_n x) = e^{ikz} A_S(x). \end{aligned} \quad (8.2.25)$$

Это значит, что с точностью до фазового множителя e^{ikz} на таком расстоянии воспроизводится структура исходного волнового поля $A_S(x)$, поскольку относительный набег фаз всех слагаемых в сумме (8.2.20) оказывается кратным 2π .

Условие (8.2.24) выполняется в точках

$$z = mz_T, \quad z_T = 2\pi \frac{2k}{\Omega^2} = \frac{2d^2}{\lambda}, \quad m = 0, 1, 2, \dots \quad (8.2.26)$$

Действительно, в этих точках

$$\frac{u_n^2}{2k} z = n^2 \frac{\Omega^2}{2k} mz_T = 2\pi mn^2 = 2\pi p, \quad (8.2.27)$$

где $p = mn^2$ — целое число. Заметим, что хотя каждое из слагаемых в сумме (8.2.20) приобретает свой набег фазы, но эти набег для всех слагаемых кратны 2π .

Таким образом, периодическая структура $A_S(x)$ периодически самовоспроизводится (повторяется) при условии

$$z = mz_T = m \frac{2d^2}{\lambda}. \quad (8.2.28)$$

Оценим максимальное число «копий» структуры, которое можно наблюдать при использовании решётки длиной D . Эффект саморепродукции возникает благодаря интерференции волн, распространяющихся от решётки под разными углами. Наименьший угол с осью z образуют 0-й и (± 1) -й главные дифракционные максимумы. Распространяясь от решетки

конечных размеров D , эти три волны перестают перекрываться на расстоянии (рис. 8.2.3):

$$L = \frac{D}{2} \operatorname{ctg} \theta, \quad (8.2.29)$$

где θ — угол дифракции, определяемый условием

$$d \sin \theta = \lambda. \quad (8.2.30)$$

Считая $\lambda \ll d$, заключаем, что $\theta \approx \lambda/d \ll 1$. Поэтому

$$L \approx \frac{D}{2\theta} \approx \frac{Dd}{2\lambda}. \quad (8.2.31)$$

На этом расстоянии число плоскостей саморепродукции составляет

$$N \sim \frac{L}{z_T} \approx \frac{Dd/2\lambda}{2d^2/\lambda} = \frac{D}{4d}. \quad (8.2.32)$$

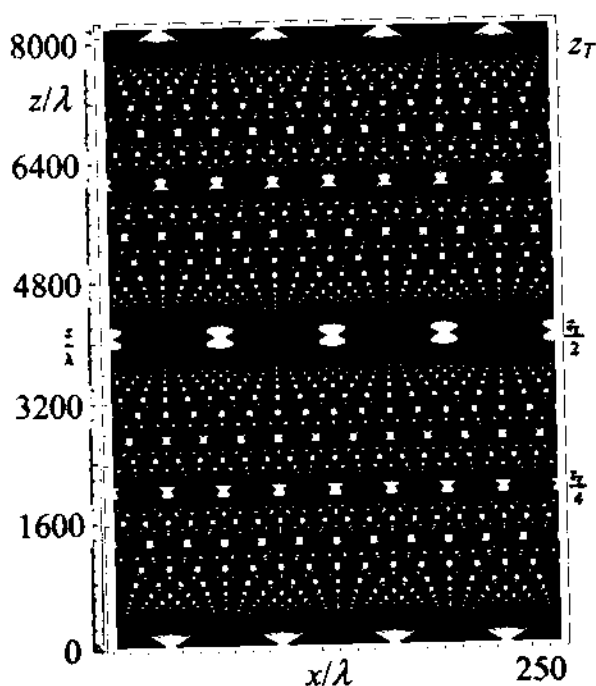
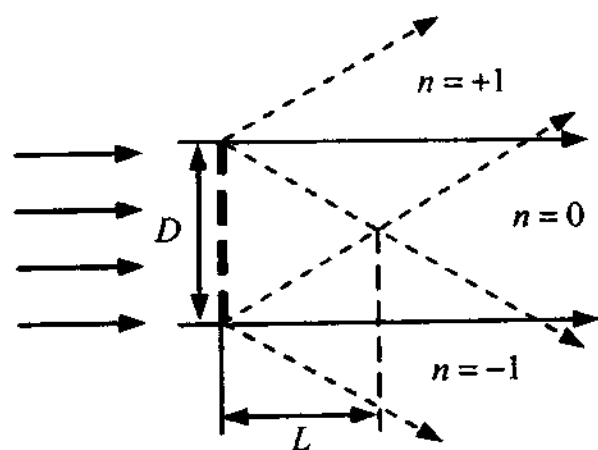


Рис. 8.2.3 (слева). Область перекрытия дифракционных максимумов от решётки длиной D
 Рис. 8.2.4 (справа). «Ковёр Талбота» — картина дифракции плоской волны на дифракционной решётке в ближней зоне $0 \leq z \leq z_T$. Координата отсчитывается снизу (от решётки) вверх до первой «копии». $z_T/\lambda = 8000$, $d/\lambda = 62,5$ (http://en.wikipedia.org/wiki/Talbot_effect)

В действительности для установления периодичности объекта по его «копии» достаточно иметь пересечение только 0-го дифракционного максимума и хотя бы одного из боковых лепестков: (+1)-го или (-1)-го. Соответственно расстояние L можно взять вдвое большее: $L_1 = D \operatorname{ctg} \theta$. Это даёт вдвое большее число наблюдаемых «копий» решётки:

$$N \sim \frac{L_1}{z_1} \approx \frac{D}{2d}. \quad (8.2.33)$$

Явление саморепродукции наблюдалось в 1836 г. английским физиком и химиком Генри Фоксом Талботом и получило название *эффекта Талбота*.

Рассмотренный эффект имеет место на расстояниях $z \geq z_T = 2d^2/\lambda$. По отношению к решётке в целом эта область попадает в зону дифракции Френеля. Действительно, все наблюдаемые «копии» находятся на расстоянии, не превышающем $L = D \operatorname{ctg} \theta \approx Dd/\lambda$. Условие дифракции Френеля имеет вид $D^2/\lambda > L$. Согласно предыдущему равенству это сводится к очевидному требованию $D > d$: размер решётки должен превышать её период.

В ближней зоне ($z < z_T$) дифракционная картина оказывается весьма сложной. На рис. 8.2.4 приведено распределение интенсивности в области $0 \leq z \leq z_T$. Эту картину называют «ковёр Талбота».

8.2.3. Функция пропускания клина

Рассмотрим фазовый объект с функцией пропускания

$$D(x) = e^{i\Omega x}. \quad (8.2.34)$$

Пусть на него падает плоская волна под углом θ_0 :

$$A = A_0 e^{i(kx \sin \theta_0 + kz \cos \theta_0)}. \quad (8.2.35)$$

Сразу после прохождения объекта находим

$$A_S(x) = D(x) A|_{z=0} = A_0 e^{i(k \sin \theta_0 + \Omega)x}. \quad (8.2.36)$$

Следовательно, в области $z > 0$ мы получаем единственную волну, направление которой определяется условием

$$k \sin \theta = k \sin \theta_0 + \Omega. \quad (8.2.37)$$

Пусть углы θ_0 и θ малы. Тогда при любом угле падения θ_0 имеем плоскую волну, изменившую направление на угол:

$$\Delta \theta = \theta - \theta_0 = \Omega/k. \quad (8.2.38)$$

В разделе 3.4.2 было показано, что в том же приближении малых углов аналогичное действие производит клин. Если преломляющий угол клина равен α , а показатель преломления материала n , то

$$\Delta \theta = (n-1)\alpha. \quad (8.2.39)$$

Сравнивая два последних равенства, находим выражение для числа Ω :

$$\Omega = k\alpha(n-1). \quad (8.2.40)$$

Результат нетрудно понять из следующего. Волна, проходя клин в том месте, где его толщина равна h , приобретает дополнительную разность хода $\Delta L = h(n-1)$. С другой стороны, если это сечение находится на рас-

стоянии x от вершины клина, то $h = \alpha x$. Поэтому $\Delta L = \alpha x(n-1)$. Соответственно получаем пропускание

$$D = e^{ik\Delta L} = e^{ik\alpha x(n-1)} \equiv e^{i\Omega x}, \quad (8.2.41)$$

что совпадает с (8.2.34), (8.2.40).

Таким образом, фазовый объект (8.2.34) отвечает клину и осуществляет простой поворот волн. В этой связи заметим, что синусоидальную решётку можно формально рассматривать как суперпозицию трёх оптических элементов: плоскопараллельной пластины и двух клиньев. Действительно, в случае амплитудной решётки функция пропускания содержала три слагаемых:

$$D = D_0 + \frac{m}{2} D_0 e^{i\Omega x} + \frac{m}{2} D_0 e^{-i\Omega x}. \quad (8.2.42)$$

Первый элемент не поворачивает лучи, тогда как второй и третий осуществляют поворот на углы, определяемые равенством

$$k \sin \theta = k \sin \theta_0 \pm \Omega, \quad (8.2.43)$$

что соответствует формуле (8.2.37).

8.3. Функция пропускания линзы

Рассмотрим на основе метода Рэлея действие линзы. Считаем линзу *тонкой* и ограничиваемся приближением *параксиальных пучков* (когда лучи образуют малые углы с главной оптической осью).

Как известно, оптические длины лучей, идущих от линзы к фокусу, одинаковы. Это значит, что плоская волна преобразуется в сферическую, сходящуюся к фокусу. Сформировавшийся сферический волновой фронт показан на рис. 8.3.1 штриховой линией. Отсюда следует, что лучи, наталкивающиеся на линзу, приобретают фазовую задержку до момента достижения сходящегося волнового фронта:

$$\Delta\varphi = k \left(f - \sqrt{f^2 + \rho^2} \right) \approx -\frac{k\rho^2}{2f}. \quad (8.3.1)$$

Здесь ρ — расстояние от точки входа луча в линзу до оптической оси. Таким образом, функцию пропускания линзы можно записать в виде

$$D = e^{i\Delta\varphi} = \exp \left(-\frac{ik\rho^2}{2f} \right). \quad (8.3.2)$$

Если линза рассеивающая, то замена $f \rightarrow -f$ даёт

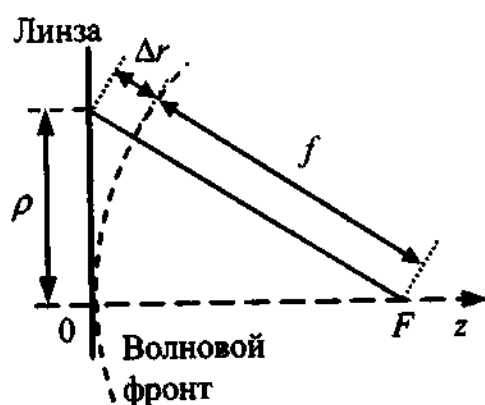


Рис. 8.3.1. К нахождению функции пропускания линзы. Фокус линзы находится в точке F

$$D = \exp\left(\frac{ik\rho^2}{2f}\right). \quad (8.3.3)$$

В качестве применения формулы (8.3.2) рассмотрим действие собирающей линзы на плоскую волну, распространяющуюся параллельно оптической оси:

$$A = A_0 e^{ikz}. \quad (8.3.4)$$

Запишем волну, прошедшую через линзу:

$$A_S(x, y) = D(x, y)A_0 = A_0 \exp\left(-\frac{k(x^2 + y^2)}{2f}\right). \quad (8.3.5)$$

Чтобы найти поле в области $z > 0$, нужно выполнить пространственное разложение Фурье функции (8.3.5):

$$a(u, v) = \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} e^{-i(ux+vy)} A_S(x, y) dx dy. \quad (8.3.6)$$

Для вычисления воспользуемся следующим равенством:

$$\int_{k_x=-\infty}^{\infty} e^{-i(ux+kx^2/2f)} dx = (1-i)\sqrt{\frac{\pi f}{k}} \exp\left(i\frac{f}{2k}u^2\right). \quad (8.3.7)$$

Аналогичное равенство имеет место и при интегрировании по dy . Соответственно получаем

$$a(u, v) = \frac{2\pi f}{ik} A_0 \exp\left[i\frac{f}{2k}(u^2 + v^2)\right]. \quad (8.3.8)$$

Чтобы восстановить поле за линзой, нужно продлить в область $z > 0$ каждую из пространственных гармоник:

$$A(x, y, z) = \int_{u=-\infty}^{\infty} \int_{v=-\infty}^{\infty} e^{iux+ivy+iz\sqrt{k^2-u^2-v^2}} a(u, v) \frac{du dv}{(2\pi)^2}. \quad (8.3.9)$$

Имея в виду параксиальное приближение: $|u| \ll k$, $|v| \ll k$, положим

$$\sqrt{k^2 - u^2 - v^2} \approx k - \frac{u^2 + v^2}{2k}. \quad (8.3.10)$$

Таким образом, задача сводится к вычислению интеграла

$$A(x, y, z) = A_0 e^{ikz} \frac{2\pi f}{ik} \int_{k_x=-\infty}^{\infty} \int_{k_y=-\infty}^{\infty} \exp\left[iux+ivy-i\frac{u^2+v^2}{2k}(z-f)\right] \frac{du dv}{(2\pi)^2}. \quad (8.3.11)$$

Расчёт с использованием равенства типа (8.3.7) даёт

$$A(x, y, z) = A_0 e^{ikz} \frac{f}{f-z} \exp\left(-i \frac{k\rho^2}{2(f-z)}\right), \quad z < f. \quad (8.3.12)$$

Заметим, что допустимость использования параксиального приближения (8.3.10) в рассматриваемой ситуации связана с быстрой сходимостью интеграла (8.3.11).

Таким образом, мы получили очевидный результат: интенсивность вдоль оптической оси меняется по закону

$$I = \frac{I_0}{(1 - z/f)^2}, \quad (8.3.13)$$

достигая бесконечного значения непосредственно в фокусе линзы $z = f$, $\rho = 0$. В рассмотренной постановке задачи линза предполагалась неограниченной. Поэтому фокусное пятно стягивается практически в точку — эффекты дифракции оказываются пренебрежимо малыми.

8.4. Принципы фурье-оптики

8.4.1. Фурье-плоскость

Как было сказано выше, в методе Рэлея волновое поле представляется в виде суперпозиции плоских волн, каждой из которых отвечает определённое направление волнового вектора. Иными словами, рассматривается пространственное преобразование Фурье. Если осуществить пространственное разделение этих компонент, то можно осуществлять преобразование поля, изменяя амплитуду или фазу отдельных фурье-гармоник. Разделение фурье-компонент производится, например, линзой в её фокальной плоскости, поскольку плоские волны, отличающиеся направлением, собираются в несовпадающие точки в фокальной плоскости (рис. 8.4.1):

$$x = f\theta \approx f \frac{k_x}{k}, \quad \text{или} \quad k_x = \frac{kx}{f}. \quad (8.4.1)$$

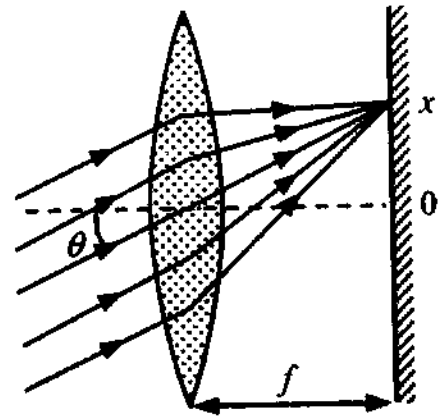


Рис. 8.4.1. Линза собирает плоскую волну в точку (дифракционное пятно) в её фокальной плоскости

Поскольку поле A_S , падающее на линзу, есть суперпозиция волн разных направлений:

$$A_S(x) = \int a(k_x) e^{ik_x x} \frac{dk_x}{2\pi}, \quad a(k_x) = \int A_S(x) e^{-ik_x x} dx, \quad (8.4.2)$$

то поле в фокальной плоскости A_f даёт представление о спектре плоских волн, образующих изучаемое поле:

$$A_f(x) \sim a(kx/f). \quad (8.4.3)$$

В этой связи фокальную плоскость называют *фурье-плоскостью*.

8.4.2. Теория Аббе формирования изображений

По теории, предложенной немецким физиком-оптиком Эрнестом Аббе (1904), формирование изображения с помощью одиночной линзы происходит в два этапа (рис. 8.4.2).

На первом этапе свет распространяется от плоскости P_1 до фурье-плоскости F , где формируется пространственный спектр предметной волны (*первая дифракция*): линза L осуществляет первое пространственное преобразование Фурье и создаёт в плоскости F *первичное изображение*.

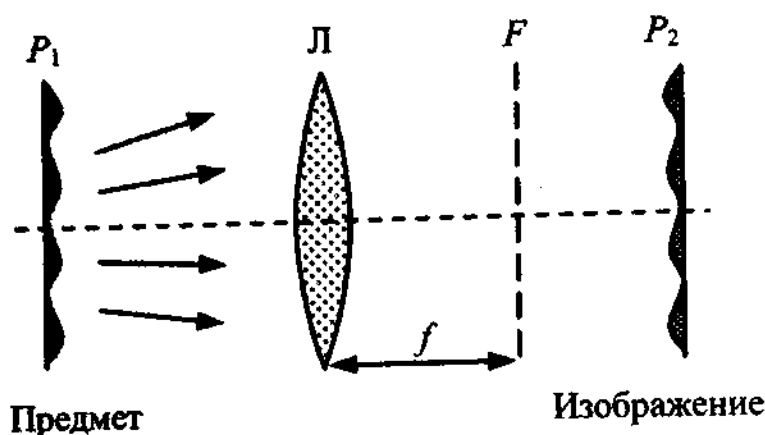


Рис. 8.4.2. Принцип двойной дифракции Аббе. Волны, исходящие от предмета, находящегося в плоскости P_1 , создают первичное изображение в фокальной плоскости F линзы L . Далее, волны, исходящие от точек первичного изображения, интерферируют, формируя вторичное изображение, находящееся в сопряженной плоскости P_2

На втором этапе свет распространяется от фурье-плоскости F до плоскости P_2 . Если плоскости P_1 и P_2 сопряженные, то в P_2 формируется изображение предмета (или *вторичное изображение*). При этом амплитудно-фазовые соотношения точек поля в плоскостях P_1 и P_2 одинаковые (может быть, с точностью до масштабного преобразования). Соответственно первичное изображение подвергается обратному преобразованию Фурье (что рассматривается как *вторая дифракция*).

Теория Аббе формирования изображения получила название *принципа двойной дифракции*.

Заметим, что вторую дифракцию можно осуществлять, поставив на пути света, идущего от фокальной плоскости, вторую линзу. Её действие аналогично действию первой линзы. Развитие этого подхода привело к иным формулировкам принципа двойной дифракции, применимым к оптическим системам, содержащим две (и более) линзы.

8.4.3. Принципы пространственной фильтрации

Эффективным инструментом преобразования волновых полей является оптическая система, показанная на рис. 8.4.3 и называемая *схемой Катрона*. Используются две одинаковые фокусирующие линзы, располо-

женные на двойном фокусном расстоянии друг от друга (то есть поставленные софокусно). Их общая фокальная плоскость F является фурье-плоскостью. Плоскости, в которых расположен предмет (P_1) и рассматривается изображение (P_2), являются сопряжёнными — в них реализуются одинаковые амплитудно-фазовые соотношения между колебаниями поля в различных (эквивалентных) точках. Очевидно, что изображение предмета является перевернутым, с линейным увеличением $\beta = -1$.

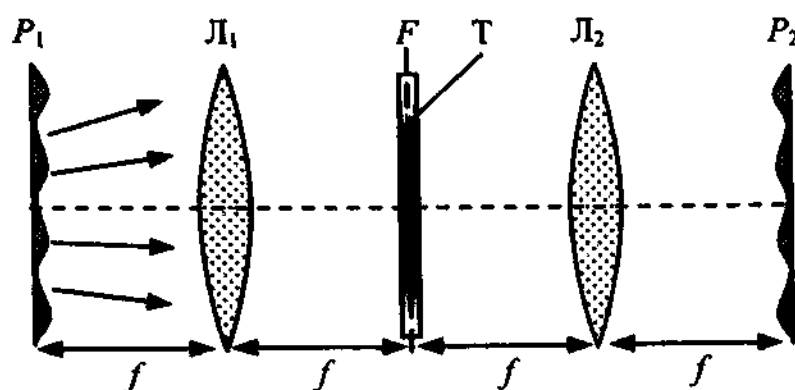


Рис. 8.4.3. Схема Катрона пространственной фильтрации

В схеме Катрона можно осуществлять избирательное воздействие на отдельные пространственные гармоники, ставя на их пути в фурье-плоскости различные элементы:

- 1) их можно удалять непрозрачными масками,
- 2) можно менять их фазы, выставя прозрачные пластинки, вносящие фазовые задержки,
- 3) можно менять амплитуду, ставя полупрозрачные экраны.

Пластину, выставляемую в фурье-плоскости и осуществляющую требуемое воздействие на фурье-спектр, называют *транспарантом* T (рис. 8.4.3).

В схеме Катрона решаются различные задачи оптической обработки информации. Такое воздействие на пространственный спектр называется *фильтрацией*, а система $(L_1 + T + L_2)$ — *фурье-фильтром*.

Пример. Пусть предметом (в плоскости P_1 на рис. 8.4.3) является дифракционная решётка с периодом d . Тогда в фурье-плоскости F возникнет серия точек, положение которых определяется направлениями на главные дифракционные максимумы.

Аббе выполнил следующий эксперимент: он поставил в фурье-плоскости вторую решётку, которая перекрывала каждый второй дифракционный максимум. В результате второй дифракции на экране P_2 возникло изображение решётки, штрихи которой расположены вдвое чаще.

Этот результат очевиден из формулы для направлений дифракционных максимумов:

$$d \sin \theta = p \lambda, \quad p = 0, \pm 1, \pm 2, \dots \quad (8.4.4)$$

Для малых углов дифракции отсюда следует

$$\theta_p = p \lambda / d. \quad (8.4.5)$$

Удаление половины (через один) дифракционных максимумов эквивалентно тому, что в предметной плоскости находится решётка с периодом $\tilde{d} = d/2$:

$$\tilde{\theta}_p = p \frac{\lambda}{\tilde{d}} = p \frac{\lambda}{d/2} = 2\theta_p. \quad (8.4.5)$$

8.5. Методы наблюдения фазовых объектов

Среди различных применений фурье-оптики большое значение получило наблюдение фазовых объектов, то есть объектов, которые не поглощают свет, но только меняют фазы прошедших через них волн. Их функцию пропускания можно записать в виде

$$D = e^{i\varphi(x, y)}. \quad (8.5.1)$$

И поскольку прошедшая через них волна не меняет своей интенсивности:

$$I|_{z=+0} = |DA_0|^2 = |A_0|^2 = I_0, \quad (8.5.2)$$

то прямое наблюдение может вызывать затруднения. Однако методами фурье-фильтрации нетрудно сделать изображение таких объектов амплитудно-модулированным, допускающим обычную регистрацию. Рассмотрим два основных подхода: метод фазового контраста и метод тёмного поля.

8.5.1. Метод фазового контраста Цернике

Данный метод предложен Ф. Цернике в 1933 г. Его суть состоит том, что фазу исходной волны, просвечивающей объект, изменяют на $\pi/2$. В схеме Катрона (рис. 8.4.3) для этого в фурье-плоскости на оси системы в качестве транспаранта Т достаточно выставить небольшую прозрачную пластинку такой толщины h , что

$$(n-1)h = (2p+1)\lambda/4, \quad p = 0, 1, 2, \dots \quad (8.5.3)$$

Пластинка должна перекрывать только прямую (просвечивающую) волну, тогда как отклонённые объектом волны через неё проходить не должны.

Рассмотрим действие данного метода сначала для случая, когда объектом является фазовая синусоидальная решётка. Считая вызываемую им модуляцию амплитуды малой, положим

$$D = e^{im \cos \Omega x} \approx 1 + im \cos \Omega x. \quad (8.5.4)$$

Пусть на решётку падает плоская волна вдоль оси z :

$$A = A_0 e^{ikz}, \quad z < 0. \quad (8.5.5)$$

Сразу после прохождения решётки волна принимает вид

$$A|_{z=+0} \equiv A_S = A_0 (1 + im \cos \Omega x), \quad (8.5.6)$$

или

$$A_S = A_0 + \frac{imA_0}{2} e^{i\Omega x} + \frac{imA_0}{2} e^{-i\Omega x}. \quad (8.5.7)$$

В области $z > 0$ от решётки идут три волны:

$$A|_{z>0} = A_0 e^{ikz} + \frac{imA_0}{2} e^{i\Omega x + iwz} + \frac{imA_0}{2} e^{-i\Omega x + iwz}, \quad (8.5.8)$$

где $w = \sqrt{k^2 - \Omega^2}$. Первая из них есть исходная (просвечивающая) плоская волна, распространяющаяся вдоль оси оптической системы. Если бы в фурье-плоскости не было никаких оптических элементов, то в сопряжённой плоскости все три волны сложились бы так, что мы получили бы в сопряжённой плоскости точное изображение исходной решётки. Пусть теперь фаза прямой волны с помощью фазовой пластинки изменяется на $\pi/2$. Наклонные же волны проходят мимо пластинки, и их фаза не меняется. В результате в плоскости изображений получаем следующую суперпозицию волн:

$$\tilde{A} = iA_0 + \frac{imA_0}{2} e^{i\Omega x} + \frac{imA_0}{2} e^{-i\Omega x} = iA_0 (1 + m \cos \Omega x). \quad (8.5.9)$$

Поскольку предметная плоскость и плоскость изображений — сопряжённые, то фазовые множители, зависящие только от z , оказываются одинаковыми для всех складываемых волн и в (8.5.9) не приведены.

Как видно из (8.5.8), в плоскости изображения мы получаем амплитудно-модулированный сигнал:

$$I = I_0 (1 + 2m \cos \Omega x). \quad (8.5.10)$$

В общем случае функция пропускания фазового объекта имеет вид (8.5.1). Считая вызванные этим объектом изменения волнового поля малыми, положим

$$D \approx 1 + i\varphi(r). \quad (8.5.11)$$

Прошедшую через такой объект волну запишем в виде

$$A = A_0 (1 + i\varphi). \quad (8.5.12)$$

Согласно идее метода фазового контраста в фурье-плоскости просвечивающая волна $A_0 e^{ikz}$ заменяется на волну $iA_0 e^{ikz}$, в результате чего в плоскости изображений мы имеем следующую суперпозицию волн:

$$\tilde{A} = iA_0 + A_0 i\varphi = iA_0 (1 + \varphi). \quad (8.5.13)$$

Таким образом, мы получаем амплитудно-модулированную волну с распределением интенсивности

$$I = |A|^2 \approx I_0(1 + 2\varphi). \quad (8.5.14)$$

Метод фазового контраста получил применения в фазово-контрастной микроскопии — методе получения изображений в оптических микроскопах, при котором фазовая модуляция световой волны преобразуется в модуляцию интенсивности.

На рис. 8.5.1 показана схема фазово-контрастного микроскопа. Свет от источника S формируется в пучок с помощью конденсора L_1 , диафрагмы D и линзы L_2 и направляется на изучаемый предмет Π . Прямой свет, не изменённый предметом (показан простыми стрелками), с помощью линзы L_3 проходит через фазовую пластинку F , находящуюся в фокусе, меняющую его фазу на $\pi/4$. Рассеянный предметом свет (показан двойными стрелками) проходит мимо пластинки. В результате интерференции этих волн в плоскости изображения возникает контрастная картина.

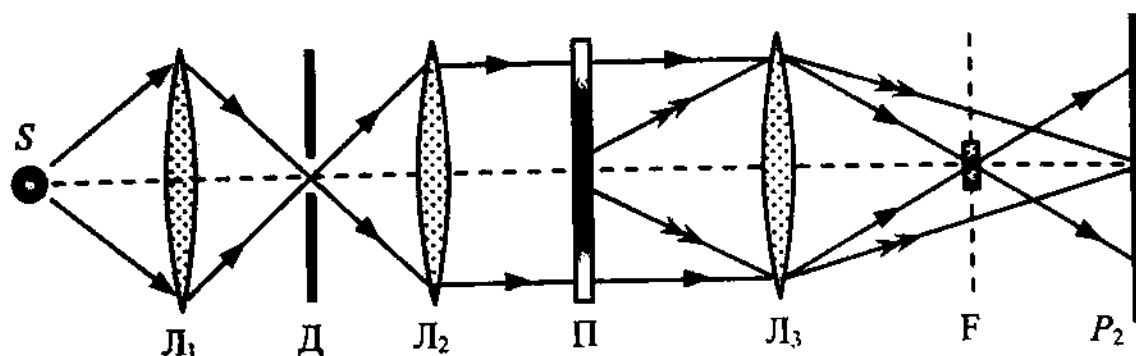


Рис. 8.5.1. Качественная схема фазово-контрастного микроскопа

«За обоснование фазово-контрастного метода, особенно за изобретение фазово-контрастного микроскопа» голландский физик Фриц Цернике был удостоен Нобелевской премии по физике (1953 г.). Фазово-контрастная микроскопия широко применяется в биологии, поскольку не требует предварительного окрашивания клеток, из-за которого они могут погибнуть.

8.5.2. Метод тёмного поля

В методе фазового контраста видность наблюдаемой картины обычно мала. Так, при $m \ll 1$ согласно (8.5.9) она составляет

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = 2m \ll 1. \quad (8.5.15)$$

В этой связи был предложен метод тёмного поля, состоящий в том, что для повышения контраста изображения следует полностью удалить прямую волну, то есть отфильтровать центральный максимум. В схеме Катрона этому отвечает установка непрозрачного экрана в фурье-плоскости на оптической оси.

Пусть, например, объектом является фазовая синусоидальная решётка, осуществляющая слабую модуляцию прошедшей через неё волны. Прошедшая через неё волна даётся формулой (6.5.6):

$$A_S = A_0 + \frac{imA_0}{2} e^{i\Omega x} + \frac{imA_0}{2} e^{-i\Omega x}.$$

В результате фильтрации в сопряжённой плоскости получаем волну

$$\tilde{A} = \frac{imA_0}{2} e^{i\Omega x} + \frac{imA_0}{2} e^{-i\Omega x} = imA_0 \cos \Omega x. \quad (8.5.16)$$

Соответственно распределение интенсивности даётся выражением

$$I = m^2 I_0 \cos^2(\Omega x) = \frac{1}{2} m^2 I_0 (1 + \cos 2\Omega x). \quad (8.5.17)$$

Видность данной картины оказывается высокой: $V=1$. Таким образом, удаётся получить высококонтрастное изображение фазового объекта. И хотя оно не точно повторяет предметную волну (в частности, отличается период), во многих случаях его достаточно, чтобы изучать свойства самого объекта.

Другим недостатком метода является то, что яркость изображения оказывается существенно ниже яркости предмета: наличие малого множителя m^2 в выражении для интенсивности (8.5.16) связано именно с тем, что вместе с центральным максимумом отфильтровывается значительная часть энергии просвечивающей волны. Этот недостаток, однако, можно компенсировать, используя более яркий свет или усилители сигнала.

Метод тёмного поля получил применение в тёмнопольной микроскопии, поскольку при регистрации только света, рассеянного изучаемым образцом, выявляются даже незначительные различия в преломляющей способности различных участков препарата.

Основы метода разработаны австро-венгерским химиком Рихардом Адольфом Зигмонди (1906), лауреатом Нобелевской премии по химии (1925) за исследования коллоидных растворов, имеющие разнообразные применения в области коллоидной химии.

Имеются различные реализации метода тёмного поля в оптической микроскопии. В одной из схем образец освещается сбоку. При таком освещении неоднородности, имеющиеся в образце, рассеивают падающий свет. С помощью этого рассеянного света и наблюдают в микроскопе изображение образца. Освещающий же световой пучок в объектив не попадает. В другой схеме тёмнопольного микроскопа используется просвечивание образца. Для наблюдения фазового объекта из полного излучения методом фурье-фильтрации удаляется просвечивающая волна. Это можно сделать, например, в схеме Катрона, рис. 8.4.3, используя в качестве транспаранта T небольшой непрозрачный диск на оси, отсекающий толь-

ко сфокусированную просвечивающую волну. Наблюдаемое изображение оказывается высококонтрастным.

8.6. Элементы голографии

8.6.1. Схемы голографии

Голография — это метод точной записи волновых полей с учётом амплитуды и фазы. Термин происходит от греч. *όλος* — полный и *γραφω* — пишу.

Для записи голограммы используется, наряду с волной от фотографируемого предмета (*предметной волной*), ещё одна волна, называемая *опорной*. На фотографии записывается результат их интерференции, в результате чего записывается информация о фазе предметной волны (относительно опорной, заданной изначально). При освещении голограммы восстанавливающей волной (которая должна быть идентичной опорной) восстанавливается предметная волна.

Качественно схема записи голограммы и восстановления изображения показана на рис. 8.6.1.

Если опорная волна падает по нормали на фотопластинку и в том же направлении, что и предметная волна, то реализуется *схема Габора*.

Если же опорная и предметная волны идут под углом к фотопластинке, то имеют дело с *внеосевой схемой*, или *схемой Лейта–Упатниекса* (США, 1962 г.).

Наконец, если предметная и опорная волны идут навстречу друг другу, то реализуется *схема Денисюка* (или *схема на встречных пучках*). Последний вариант используется при записи толстых (объёмных) голограмм (раздел 8.6.4).

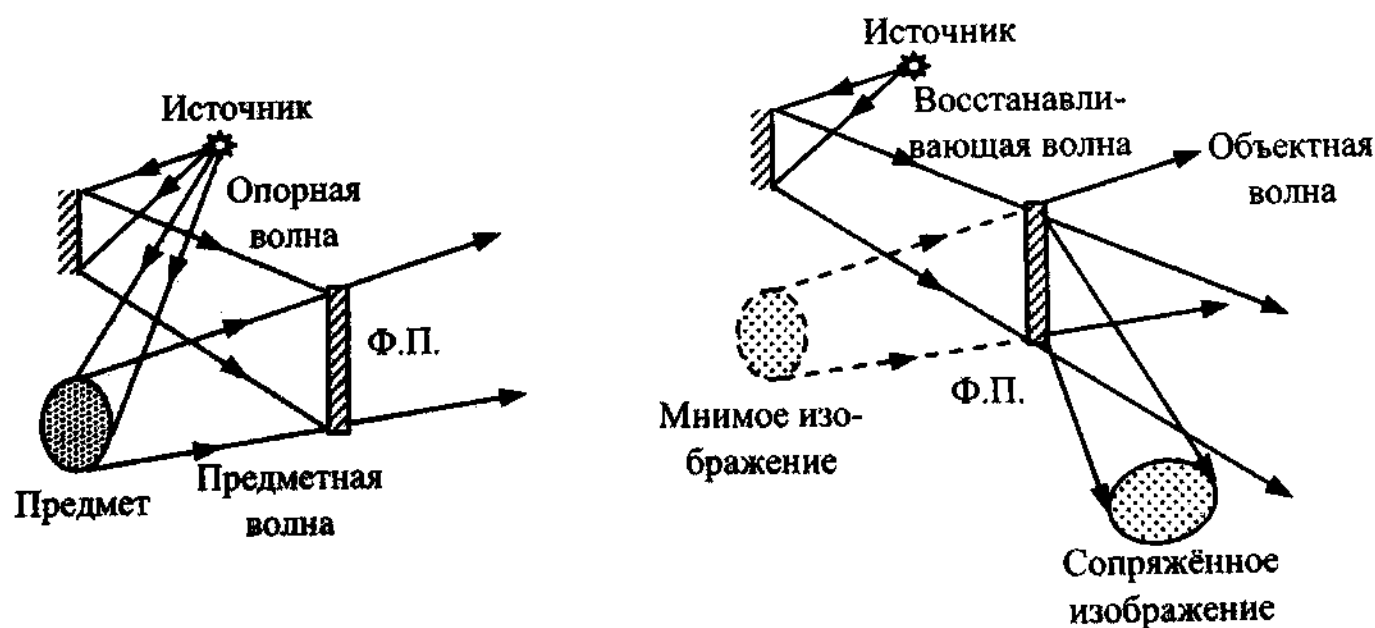


Рис. 8.6.1. Запись (слева) и восстановление (справа) голограммы. Ф.П. — фотопластинка

Идею голографии предложил венгерский физик Деннис Габор (1947). Он же предложил и сам термин «голография» (1948). «За изобретение и разработку голографического метода» ему была присуждена Нобелевская премия (1971).

Схема Габора демонстрировала лишь принципиальную возможность записи голограмм, поскольку предъявляла высокие требования к монохроматичности излучения. Поэтому первый голографический портрет был записан лишь после создания лазеров, в 1967 г.

Дальнейшее развитие голография получила в работах Юрия Николаевича Денисюка (СССР). Первые голограммы он записал в 1958 г., используя излучение ртутной лампы. В 1962 г. Денисюк изобрёл способ записи изображения в трёхмерных средах, позволяющий сохранить информацию о фазе, амплитуде и спектральном составе волны, пришедшей от объекта. Такие голограммы свободны от ряда недостатков плоских голограмм. В частности, они позволяют записать цветные изображения и могут быть восстановлены с помощью пучка обычного белого света.

8.6.2. Запись и восстановление плоской волны

Рассмотрим идею голографии на примере записи и восстановления плоской волны (рис. 8.6.2).

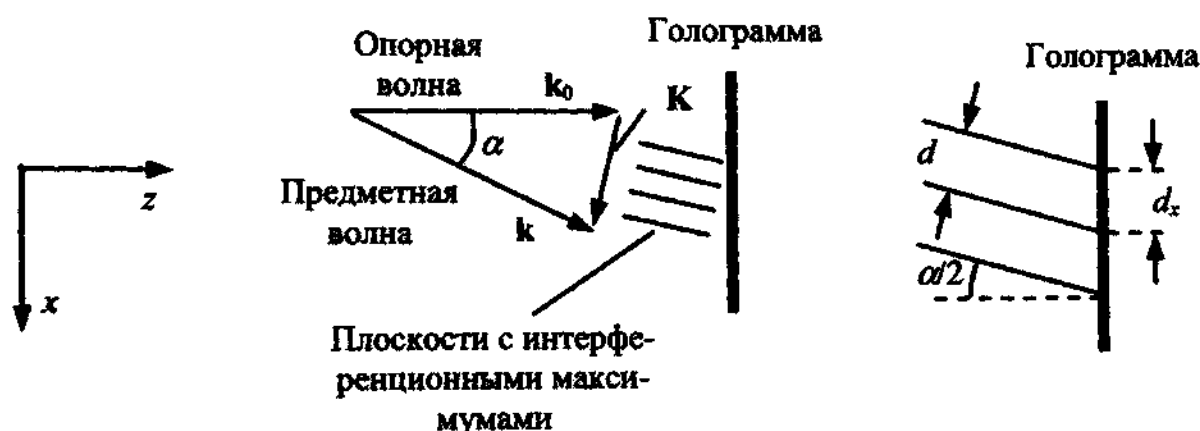


Рис. 8.6.2. Запись предметной плоской волны, падающей на фотопластинку (голограмму) под углом α , с помощью плоской волны, падающей на фотопластинку по нормали волна. Слева — интерференция предметной и опорной волн; справа — запись интерференционной картины на фотопластинку (голограмму)

Пусть на фотопластинку падает плоская волна

$$f_{\text{пр}} = f_0 e^{i k r}, \quad (8.6.1)$$

которую будем считать предметной и которую требуется записать, а затем восстановить. Считаем, что волна падает под углом α к нормали к плоскости голограммы. В качестве опорной волны используем плоскую волну, падающую по нормали к плоскости фотопластинки:

$$A = A_0 e^{i k_0 r}. \quad (8.6.2)$$

При интерференции предметной и опорной волн возникает серия плоскостей, отвечающих интерференционным максимумам, перпендикулярных вектору

$$K = k - k_0, \quad (8.6.3)$$

параллельных биссектрисе угла между векторами $\mathbf{k}$ и $\mathbf{k}_0$ и образующих угол $\alpha/2$ с плоскостью голограммы, причём в системе координат, указанной на рис. 8.6.2, компоненты векторов $\mathbf{k}$ и $\mathbf{k}_0$ следующие:

$$|\mathbf{k}| = |\mathbf{k}_0| = k, \quad k_x = k \sin \alpha, \quad k_z = k \cos \alpha, \quad k_{0,x} = 0, \quad k_{0,z} = k. \quad (8.6.4)$$

Используя эти соотношения, находим пространственное распределение интенсивности интерферирующих волн:

$$I = f_0^2 + A_0^2 + 2f_0A_0 \cos(\mathbf{K}\mathbf{r}), \quad (8.6.5)$$

$$\mathbf{K}\mathbf{r} = (\mathbf{k} - \mathbf{k}_0)\mathbf{r} = kx \sin \alpha - kz(1 - \cos \alpha).$$

В плоскости голограммы ($z = 0$) имеем

$$I(x) = f_0^2 + A_0^2 + 2f_0A_0 \cos(kx \sin \alpha). \quad (8.6.6)$$

Эта структура представляет собой амплитудную дифракционную решётку с периодом

$$d_x = \frac{2\pi}{k \sin \alpha} = \frac{\lambda}{\sin \alpha}. \quad (8.6.7)$$

Для восстановления голограммы используем волну (8.6.2):

$$A_{\text{восст.}} = A_2 e^{ik_0 \mathbf{r}}. \quad (8.6.8)$$

В соответствии с принципами фурье-оптики голограмму можно рассматривать как объект с функцией пропускания

$$\begin{aligned} D(x) &= \beta I(x) = D_1 + D_2 \cos(kx \sin \alpha) = \\ &= D_1 + \frac{1}{2} D_2 e^{ikx \sin \alpha} + \frac{1}{2} D_2 e^{-ikx \sin \alpha}. \end{aligned} \quad (8.6.9)$$

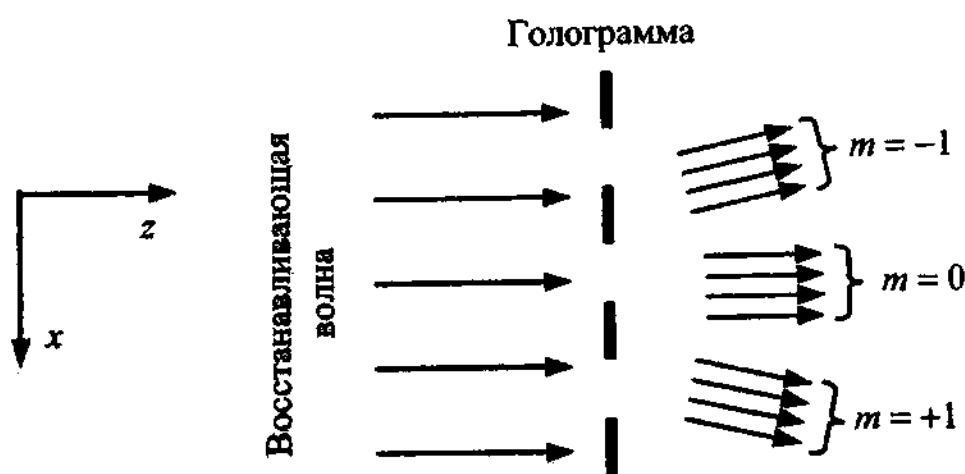


Рис. 8.6.3. Восстановление предметной волны по голограмме с помощью восстанавливающей плоской волны, падающей на фотопластинку по нормали

Таким образом, в рассматриваемой ситуации мы имеем дело с амплитудной синусоидальной решёткой с периодом, даваемым формулой (8.6.7). При дифракции плоской волны $A_{\text{восст.}}$ на такой решётке возника-

ют три волны, отвечающие главным максимумам $m = -1, 0, +1$ (рис. 8.6.3):

$$A = \underbrace{\frac{1}{2} D_2 A_2 e^{ikx \sin \alpha + ikz \cos \alpha}}_{m=+1} + \underbrace{D_1 A_2 e^{ikz}}_{m=0} + \underbrace{\frac{1}{2} D_2 A_2 e^{-ikx \sin \alpha + ikz \cos \alpha}}_{m=-1}. \quad (8.6.10)$$

Из них первый ($m = +1$) даёт волну, совпадающую с записанной предметной волной $A_{\text{пр}}$ (восстановленную волну). Максимум $m = 0$ порождает прошедшую через голограмму восстанавливающую волну. Наконец, максимуму $m = -1$ отвечает сопряжённая волна.

8.6.3. Голограмма с наклонным опорным пучком

Рассмотрим запись и восстановление произвольной предметной волны во внеосевой схеме (рис. 8.6.4).

Пусть опорная волна падает на голограмму под некоторым углом θ и описывается выражением

$$A = A_0 e^{ikr}, \quad A_S \equiv A|_{z=0} = A_0 e^{iux}, \quad (8.6.11)$$

$$u \equiv k_x = k \sin \theta, \quad k_z = \sqrt{k^2 - u^2} = k \cos \theta.$$

Запишем предметную волну в следующей форме:

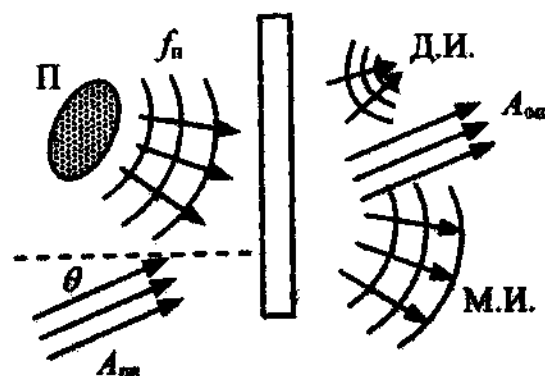
$$f_n = a(\mathbf{r}) e^{i\varphi(\mathbf{r})}. \quad (8.6.12)$$

Будем считать, что функция пропускания записанной голограммы пропорциональна интенсивности сигнала на поверхности $z = 0$:

$$D = \beta |A_S + f_{n,S}|^2 = \beta \left(|A_S|^2 + |f_{n,S}|^2 + A_S^* f_{n,S} + A_S f_{n,S}^* \right). \quad (8.6.13)$$

Индекс «S» в символах $f_{n,S}$, $f_{n,S}^*$ означает, что эти функции берутся в плоскости голограммы (то есть при $z = 0$).

Рис. 8.6.4. Запись и восстановление голограммы предмета П во внеосевой схеме. Ф.П. — фотопластинка, Д.И. — действительное изображение, М.И. — волна, идущая от мнимого изображения



Рассмотрим теперь восстановление данной голограммы, используя в качестве восстанавливающей такую же волну, что была использована при записи. Используя граничное условие Кирхгофа, запишем волну, прошедшую через голограмму:

$$\Phi_S = DA_S = \underbrace{\beta A_0^2 A_0 e^{iux}}_{\Phi_{1,S}} + \underbrace{\beta |f_{n,S}|^2 A_0 e^{iux}}_{\Phi_{2,S}} + \underbrace{\beta A_0^2 f_{n,S}}_{\Phi_{3,S}} + \underbrace{\beta A_0^2 e^{2iux} f_{n,S}^*}_{\Phi_{4,S}}. \quad (8.6.14)$$

В этом выражении отдельные слагаемые порождают следующие волны за голограммой:

1. $\Phi_{1,S} = \beta A_0^2 \cdot A_0 e^{iux}$ порождает прошедшую через голограмму восстанавливающую волну:

$$\Phi_1 \sim A_0 e^{ikr}, \quad \Phi_1|_{z=0} = \Phi_{1,S};$$

2. $\Phi_{2,S} = \beta |f_{n,S}|^2 A_0 e^{iux}$ даёт фон Φ_2 ($\Phi_2|_{z=0} = \Phi_{2,S}$) в виде амплитудно-модулированной восстанавливающей волны;

3. $\Phi_{3,S} = \beta A_0^2 f_{n,S}$ порождает искомую предметную волну (волну, идущую от мнимого изображения):

$$\Phi_3 \sim f_n, \quad \Phi_3|_{z=0} = \Phi_{3,S};$$

4. $\Phi_{4,S} = \beta A_0^2 e^{2iux} f_{n,S}^*$ даёт сопряжённое изображение Φ_4 ($\Phi_4|_{z=0} = \Phi_{4,S}$).

Учтём теперь, что множитель

$$D_2 = e^{2iux} \quad (8.6.15)$$

можно рассматривать как функцию пропускания клина (раздел 8.2.3), осуществляющего поворот волны на угол $\theta_2 = \arcsin(2u/k)$.

Таким образом, восстанавливающая, предметная и сопряжённая волны во внеосевой схеме идут в разных направлениях, что позволяет в принципе выделить записанное на голограмме изображение предмета.

Вместе с тем волны Φ_1, Φ_2, Φ_4 препятствуют получению качественного объёмного изображения. Частично уменьшить относительный вклад волны Φ_2 можно, увеличивая амплитуду опорной волны. Роль же плоской волны Φ_1 сводится к созданию равномерного светлого фона (который устраняется, например, методом тёмного поля). Наконец, наложение предметной и сопряжённой волн может заметно ограничивать углы, под которыми удаётся наблюдать записанное изображение.

Рассмотрим подробнее волну, порождаемую слагаемым $\Phi_{4,0}$ в (8.6.14).

Пусть сначала опорная (как и восстанавливающая) волна падает по нормали к голограмме: $u \equiv k_x = 0, \quad k_z = k$.

Заметим, что предметная волна $f_n = a(\mathbf{r})e^{i\varphi(\mathbf{r})}$ и комплексно-сопряжённая ей $f_n^* = a(\mathbf{r})e^{-i\varphi(\mathbf{r})}$ удовлетворяют одному и тому же уравнению Гельмгольца:

$$\Delta f_n + k^2 f_n = 0, \quad \Delta f_n^* + k^2 f_n^* = 0 \quad (8.6.16)$$

как линейному уравнению с действительными коэффициентами. Соответственно в области $z > 0$ вследствие чётности данного уравнения по координате z оно имеет два решения:

$$\Phi_4^{(1)} = a(x, y, z)e^{-i\varphi(x, y, z)}, \quad \Phi_4^{(2)} = a(x, y, -z)e^{-i\varphi(x, y, -z)}. \quad (8.6.17)$$

Предметная волна $f_n = a(x, y, z)e^{i\varphi(x, y, z)}$ идёт от предмета, тогда как волна $f_n^* = a(x, y, z)e^{-i\varphi(x, y, z)}$ распространяется в сторону предмета. Последнее становится очевидным, если рассмотреть частный случай, когда предметная волна плоская: $f_n = a_0 e^{ikz}$, $\varphi(\mathbf{r}) = kz$, и учесть, что распространяющаяся в положительном направлении волна имеет вид $f_n = a_0 e^{ikz}$. Поскольку в области $z > 0$ мы должны оставить только те волны, которые идут от предмета, то из пары (8.6.17) остаётся только волна $\Phi_4^{(2)} = a(x, y, -z)e^{-i\varphi(x, y, -z)}$. Эта волна получается зеркальным отражением предмета относительно голограммы и сходится к действительному изображению в области $z > 0$.

Пусть теперь $u \neq 0$, $k_z = \sqrt{k^2 - u^2}$. Тогда волна, порождаемая слагаемым A_4 , удовлетворяет граничному условию

$$\Phi_4|_{z=0} = \beta A_0^2 a(x, y, 0)e^{-i\varphi(x, y, 0) + 2iux}. \quad (8.6.18)$$

Множитель $D_2 = e^{2iux}$ можно рассматривать как дополнительную функцию пропускания, описывающую отклонение волны на угол $\theta_2 = \arcsin(2u/k)$ и приводящую к пространственному разделению волн от мнимого (восстановленного) и действительного (сопряжённого) изображений.

8.6.4. Разрешающая способность голограммы

Пусть на голограмме записан точечный источник. Тогда при восстановлении изображение уже не будет точкой, поскольку его размер определится диаметром дифракционного пятна:

$$l \sim \lambda L/D,$$

где D — диаметр голограммы, L — расстояние до точки наблюдения. Эта оценка относится как к действительному, так и к мнимому изображению (в последнем случае L — расстояние от предмета до голограммы).

Если расстояние между двумя точками окажется меньше Δx , то они по критерию Рэлея не разрешаются — их кружки Эйри перекрываются слишком сильно. Таким образом, на голограмме размером D можно записать детали размером $b > \lambda L/D$, которые можно различить.

8.6.5. Ограничения на допустимую ширину спектра

Поскольку на голограмме записывается интерференционная картина, то при её восстановлении необходимо, чтобы разность хода лучей от центра и крайних точек голограммы не превышала длины когерентности (рис. 8.6.5):

$$\sqrt{L^2 + (D/2)^2} - L < \frac{\lambda^2}{\Delta\lambda} \Rightarrow \frac{D^2}{8L} < \frac{\lambda^2}{\Delta\lambda} \Rightarrow \Delta\lambda < 2 \left(\frac{2\lambda}{D} \right)^2 L.$$

В противном случае волны перестают быть когерентными, интерференционная картина пропадает, и мы получаем обычную амплитудную фотографию предмета.

В качестве примера укажем, что при записи голограммы размером $D = 1$ см предмета, находящегося на расстоянии $L = 1$ м, с помощью излучения с длиной волны $\lambda = 0,5$ мкм допустимая степень некогерентности составляет

$$\Delta\lambda/\lambda < 0,04.$$

8.6.6. Голограмма точечного источника

Рассмотрим запись голограммы излучения точечного источника и её восстановление. В качестве опорной волны используем плоскую волну, падающую по нормали к фотопластинке (схема Габора, рис. 8.6.6).

Поле на фотопластинке представляет собой суперпозицию волны от источника и опорной волны:

$$A = A_s \frac{e^{ikr}}{r} + A_{\text{оп}} e^{ikz}. \quad (8.6.19)$$

На фотопластинке записывается распределение интенсивности $I = |A|^2$. Соответственно положим, что функция пропускания голограммы пропорциональна I : $D = \beta I$. Для упрощения будем предполагать, что источник находится достаточно далеко от фотопластинки, что позволяет рассматривать лишь волны, идущие от источника с малым разбросом направлений. Будем также считать, что в центре голограммы (рис. 8.6.7)

$$z = L, \quad \rho = \sqrt{x^2 + y^2} = 0$$

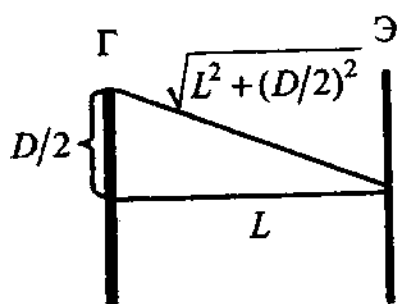


Рис. 8.6.5. К оценке разности хода лучей от центра и периферии голограммы

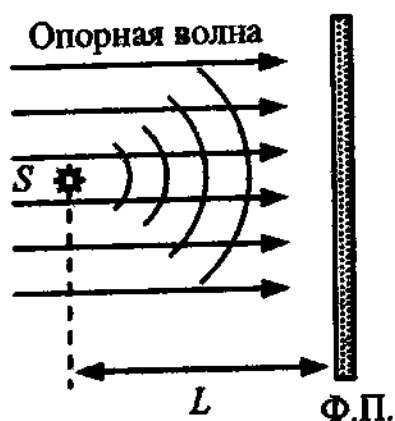


Рис. 8.6.6. Запись голограммы точечного источника S на фотопластинке Ф.П.

одинаковы амплитуды интерферирующих волн: $A_s/L = A_{\text{он}} \equiv A_0$ и их фазы. С учётом сказанного в пределах голограммы можно положить

$$A_s \frac{e^{ikr}}{r} \approx A_0 e^{ikr}. \quad (8.6.20)$$

Имея это в виду, находим распределение интенсивности на фотопластинке:

$$I = I_0 |e^{ikr} + e^{ikL}|^2 = 2I_0 \{1 + \cos[k(r-L)]\}. \quad (8.6.21)$$

В малоугловом приближении имеем

$$r - L = \sqrt{L^2 + \rho^2} - L \approx \rho^2 / 2L.$$

Таким образом, получаем

$$D = \beta I = \frac{1}{2} D_m \left[1 + \cos\left(\frac{k\rho^2}{2L}\right) \right]. \quad (8.6.22)$$

График этой функции пропускания показан на рис. 8.6.8. Кроме того, на рис. 8.6.9 показан фрагмент голограммы.

Голограмма представляет собой набор тёмных и светлых колец подобно зонной пластинке и называется *зонной пластинкой Габора*.

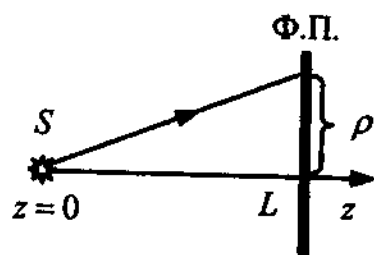


Рис. 8.6.7. К расчёту функции пропускания голограммы точечного источника

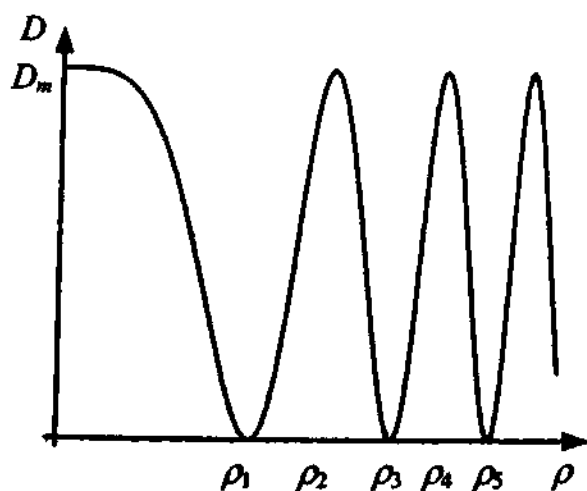


Рис. 8.6.8. Зависимость пропускания голограммы точечного источника от расстояния до центра

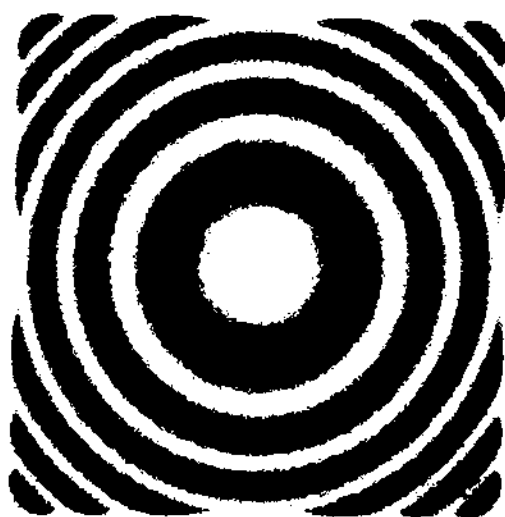


Рис. 8.6.9. Голограмма точечного источника

Радиусы центров тёмных колец даются согласно (8.6.22) формулой

$$\cos(k\rho^2/2L) = -1 \Rightarrow \rho_m = \sqrt{\frac{2L}{k}(2m-1)\pi} = \sqrt{(2m-1)\lambda L}, \quad m = 1, 2, \dots \quad (8.6.23)$$

Радиусы центров светлых колец равны

$$\cos(k\rho^2/2L) = +1 \Rightarrow \rho_m = \sqrt{\frac{2L}{k} 2m\pi} = \sqrt{2m\lambda L}, \quad m = 1, 2, \dots \quad (8.6.24)$$

Центр голограммы в рассматриваемом случае отвечает значению $m = 0$ в последней формуле.

Рассмотрим теперь восстановление найденной голограммы. Можно воспользоваться общим подходом, изложенным в предыдущем разделе. В данном случае, однако, проще поступить по-другому.

Запишем функцию пропускания (8.6.22) в комплексном виде:

$$D = \frac{1}{2} D_m + \frac{1}{4} D_m \exp\left(-\frac{ik\rho^2}{2L}\right) + \frac{1}{4} D_m \exp\left(\frac{ik\rho^2}{2L}\right). \quad (8.6.25)$$

Первое слагаемое есть константа. Это значит, что восстанавливающая волна проходит без изменения (меняется лишь её амплитуда).

Второе слагаемое совпадает с функцией пропускания фокусирующей линзы (8.3.2), имеющей фокусное расстояние $f_1 = L$. Это значит, что восстанавливающая волна собирается в действительное изображение (даёт сопряжённую волну).

Наконец, третье слагаемое совпадает с функцией пропускания рассеивающей линзы с фокусным расстоянием $f_2 = L$. При прохождении через такую линзу восстанавливающей волны формируется мнимое изображение, расположенное в точности в том месте, где находился источник, изображение которого записывалось.

8.6.7. Объёмные голограммы

До сих пор мы рассматривали голограммы, представляющие собой запись картины интерференции в тонком слое светочувствительного материала: $h \ll \Lambda$, где Λ — характерный период интерференционной картины. Такие *тонкие* голограммы не позволяют, в частности, с высокой точностью выделить записанное изображение. Данный недостаток может быть устранён, если поле интерферирующих волн (предметной и опорной) записывается в некотором *объёме* фоточувствительного материала. Такие голограммы называют *толстыми* (или *объёмными*).

Чтобы проиллюстрировать принципы записи и воспроизведения толстых голограмм, рассмотрим частный случай, когда предметная волна — плоская, падающая на фотопластинку под некоторым углом α (рис. 8.6.10):

$$f_n = f_0 e^{ikr}, \quad \mathbf{k} \equiv (u, w) = (k \sin \alpha, k \cos \alpha). \quad (8.6.26)$$

Будем считать, что опорная волна распространяется по нормали к поверхности голограммы:

$$A_{\text{оп}} = A_0 e^{ikz}. \quad (8.6.27)$$

Интенсивность суммарной волны равна

$$I = |A_0 e^{ikz} + f_0 e^{iux+iwz}|^2 = A_0^2 + f_0^2 + 2A_0 f_0 \cos[kx \sin \alpha - k(1 - \cos \alpha)z]. \quad (8.6.28)$$

Отсюда видно, что период структуры вдоль трассы опорной волны (оси z) равен (рис. 8.6.10)

$$h = \frac{2\pi}{k(1 - \cos \alpha)} = \frac{\lambda}{1 - \cos \alpha}. \quad (8.6.29)$$

Как уже говорилось, расстояние между соседними плоскостями интерференционной картины (период структуры) равно

$$d = \frac{\lambda}{2 \sin(\alpha/2)}. \quad (8.6.30)$$

Таким образом, голограмма представляет собой систему «зеркал» с периодом d , которые частично пропускают и частично отражают излучение.

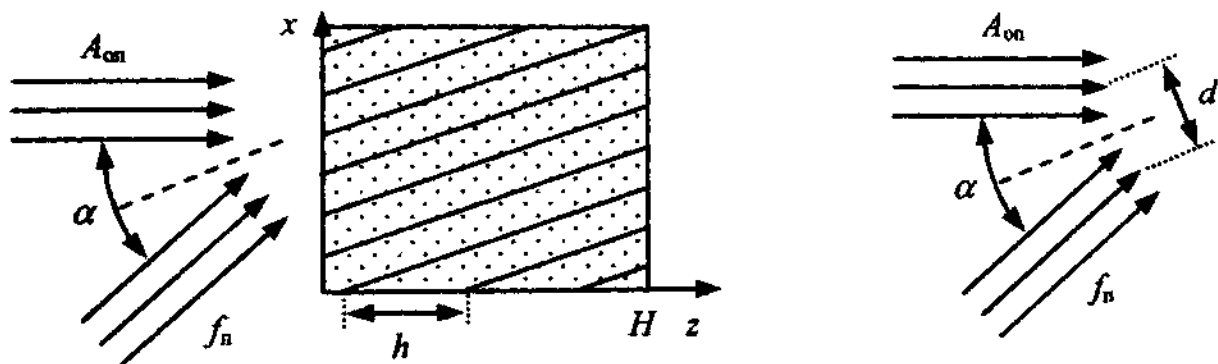


Рис. 8.6.10. Слева — запись голограммы плоской волны, падающей на фоторегистрирующую пластину под углом α , справа — интерференция скрещенных плоских волн, дающая периодическую серию интерференционных плоскостей с максимумами интенсивности на биссектрисе угла между направлениями волн

Рассмотрим теперь восстановление голограммы (8.6.28). Пусть её толщина (вдоль оси z) равна H . Восстанавливающую волну используем такую же, как и опорная (8.6.27).

Если луч пересекает в голограмме не более одного зеркала: $H < h$, то голограмма может рассматриваться как *тонкая*: её толщина почти никак не сказывается на записи и восстановлении. Если же луч пересекает много «зеркал»: $H \gg h$, толщина фоторегистрирующей среды существенна. Таким образом, условие того, что голограмма «толстая», есть

$$H \gg h = \frac{\lambda}{1 - \cos \alpha}. \quad (8.6.31)$$

В схеме Лейта–Упатниекса угол схождения лучей составляет $\alpha \leq 0,1$ рад. Соответственно период структуры для длины волны $\lambda \sim 0,5$ мкм равен

$$h \sim 2\lambda/\alpha^2 \sim 100 \text{ мкм.} \quad (8.6.32)$$

При типичной толщине фотослоя $H \sim 15 - 20$ мкм такая голограмма является «тонкой». Во встречных же пучках ($\alpha \sim \pi$, $h \sim \lambda$) для той же длины волны голограмма оказывается «толстой».

При восстановлении толстой голограммы восстанавливающая волна отражается от системы «зеркал» (рис. 8.6.11, слева). При этом отражённая волна идёт в точности в том же направлении, что и записанная предметная (как это прямо следует из закона отражения «угол падения равен углу отражения», рис. 8.6.11, справа).

Покажем, что волны, отражённые от разных «зеркал», интерференционно усиливают друг друга.

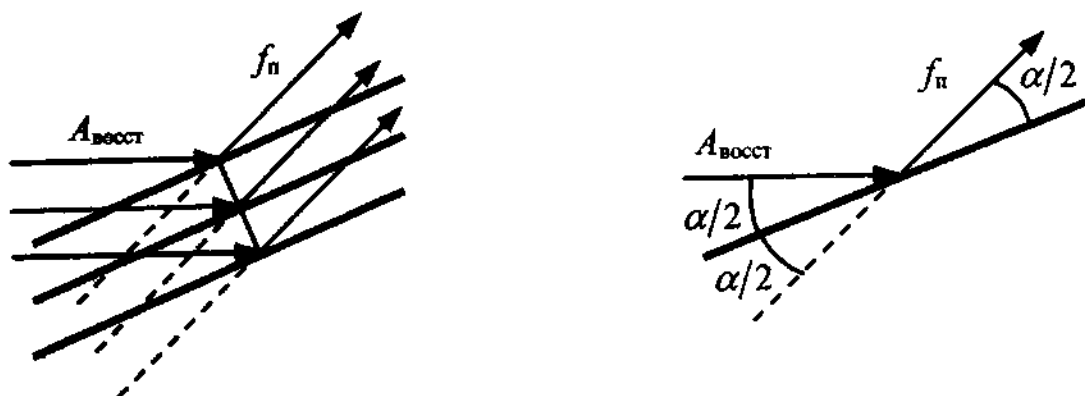


Рис. 8.6.11. Слева — восстановление голограммы (8.6.8); справа — отражение восстанавливающей волны от одной из плоскостей. Штриховой линией показано продолжение восстановленной волны (в обратном направлении)

На рис. 8.6.12 (слева) показано отражение восстанавливающей волны от периодически расположенных плоскостей, расстояние d между которыми даётся формулой (8.6.10). Для нахождения разности хода волн, отражённых от двух соседних плоскостей (рис. 8.6.12, справа), проведём отрезки OP и OQ по нормали к лучам соответственно падающей и отражённой волн. Тогда искомая разность хода оказывается равной

$$\Delta = PO'Q = 2d \cos(\pi/2 - \alpha/2) = 2d \sin(\alpha/2). \quad (8.6.33)$$

Имея в виду равенство (8.6.10), находим $\Delta = \lambda$, или, подробнее,

$$2d \sin \theta = \lambda, \quad (8.6.34)$$

где $\theta = \alpha/2$ — угол скольжения луча. Но это и означает, что отражённые волны интерференционно усиливают друг друга. Очевидно, отражение от следующих плоскостей даёт аналогичный результат: сдвиг фаз кратен 2π .

Заметим, что полученное условие совпадает с условием Брэгга-Вульфа, установленным при анализе дифракции рентгеновских лучей на кристаллах (раздел 6.9.2).

Поскольку интерференционное усиление возможно только при выполнении равенства (8.6.34), то волны, идущие под другими углами, в результате многократных отражений ослабляются. В результате при восстановлении «толстой» голограммы остаётся только записанная предметная волна — сопряжённая и восстанавливающая волны подавляются.

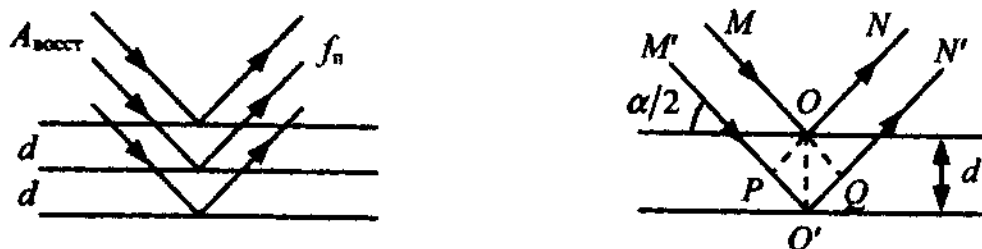


Рис. 8.6.12. Слева — отражение восстанавливающей от зеркальных плоскостей в голограмме, справа — расчёт разности хода лучей от двух соседних «зеркал»

Очевидно, при использовании немонохроматического излучения для восстановления усилятся будут только те волны, длина которых согласуется с условием (8.6.34). Это значит, что из спектра восстанавливающей волны выделяется часть, соответствующая предметной волне, записанной на голограмме. Таким образом, объёмная голограмма позволяет записывать и восстанавливать цветные изображения, используя немонохроматические опорную и предметную волны.

ГЛАВА 9

ПОЛЯРИЗАЦИЯ СВЕТА

9.1. Типы поляризации

Поляризацией называется характеристика векторных волновых полей, описывающая поведение вектора колеблющейся величины в плоскости, перпендикулярной направлению распространения волны. Термин «поляризация» применим к поперечным волнам, поскольку для продольных волн (например, волн уплотнения/разрежения среды) направление колебаний всегда совпадает с направлением распространения. Важнейшим примером является электромагнитное поле, с которым в первую очередь и связывают данное понятие.

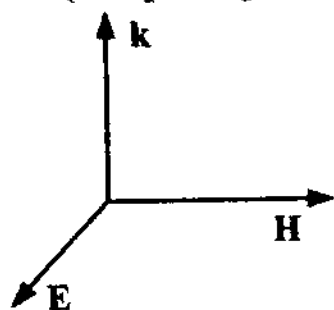


Рис. 9.1.1. Правая тройка векторов в плоской электромагнитной волне

Термин «поляризация» ввёл французский инженер Этьен-Луи Малюс, производя его от слова «полюс». Дело в том, что он был сторонником корпускулярной теории света Ньютона. По его представлениям частицы света имеют, подобно магнитам, «полюсы». В обычном свете полюсы разных частиц направлены беспорядочно. В поляризованном же свете остаются частицы только с определёнными направлениями полюсов.

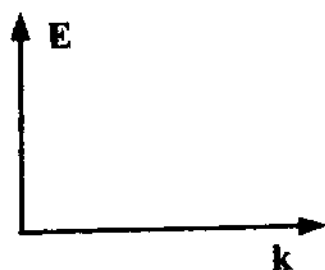


Рис. 9.1.2. Плоскость поляризации

Свет в изотропной среде представляет собой поперечные электромагнитные волны. Это значит, что векторы напряжённости электрического (E) и магнитного (H) полей совместно с волновым вектором (k) образуют правую взаимно ортогональную тройку (рис. 9.1.1).

Плоскостью поляризации называется плоскость, построенная на векторах (E, k) (рис. 9.1.2).

Плоская волна называется *линейно поляризованной*, если ориентация плоскости поляризации не меняется во времени. Иными словами, вектор напряжённости электрического поля E всегда лежит в одной и той же плоскости, содержащей вектор k . Вследствие поперечности электромагнитной волны:

$$H \perp E, H \perp k, \quad (9.1.1)$$

плоскость колебаний H также остаётся неизменной.

Из определения следует, что в линейно поляризованной волне вектор E совершает колебания

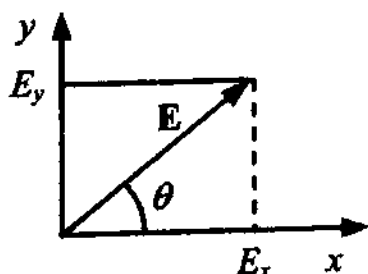


Рис. 9.1.3. Разложение вектора E на две взаимно перпендикулярные составляющие

вдоль фиксированной прямой (отсюда название поляризации «линейная»).

В естественном (неполяризованном) свете плоскость поляризации меняется случайным образом (при этом вектор $\mathbf{E}$ остаётся перпендикулярным волновому вектору $\mathbf{k}$).

Выберем систему координат таким образом, что вектор $\mathbf{k}$ направлен вдоль оси z , а вектор $\mathbf{E}$ лежит в плоскости (x, y) , как показано на рис. 9.1.3. Пусть частота колебаний поля равна ω . Тогда для компонент поля можно записать выражения

$$\begin{aligned} E_x &= E_{x0} \cos(\omega t + \varphi_x), \\ E_y &= E_{y0} \cos(\omega t + \varphi_y). \end{aligned} \quad (9.1.2)$$

Если разность фаз колебаний этих величин равна

$$\delta = \varphi_x - \varphi_y = 0, \pm \pi, \quad (9.1.3)$$

то вектор $\mathbf{E}$ совершает колебания вдоль фиксированной прямой, образующей с осью x угол θ такой, что

$$\operatorname{tg} \theta = E_y / E_x = \pm E_{y0} / E_{x0}. \quad (9.1.4)$$

Этот случай отвечает линейно поляризованной волне.

Если сдвиг фаз колебаний проекций отличается от (9.1.3), то вектор $\mathbf{E}$ совершает вращение. В этом случае говорят, что свет имеет эллиптическую поляризацию (рис. 9.1.4). Исключая время из равенств (9.1.2), получим уравнение кривой, описываемой концом вектора $\mathbf{E}$:

$$\left(\frac{E_x}{E_{x0}} \right)^2 + \left(\frac{E_y}{E_{y0}} \right)^2 - 2 \frac{E_x}{E_{x0}} \frac{E_y}{E_{y0}} \cos \delta = \sin^2 \delta. \quad (9.1.5)$$

Здесь введено обозначение $\delta = \varphi_x - \varphi_y$ для сдвига фаз колебаний проекций E_x и E_y .

Если окажется

$$\delta = \varphi_x - \varphi_y = \pm \pi/2, \quad E_{x0} = E_{y0}, \quad (9.1.6)$$

то описываемая вектором $\mathbf{E}$ кривая есть окружность. В этом случае говорят о *круговой поляризации*.

Круговая поляризация бывает двух типов — левая и правая (рис. 9.1.5). Если смотреть навстречу волновому вектору, то левой поляризации отвечает вращение вектора $\mathbf{E}$ против часовой стрелки («положительное» направление вращения). В случае правой поляризации вектор $\mathbf{E}$ вращается по часовой стрелке («отрицательное» направление вращения).

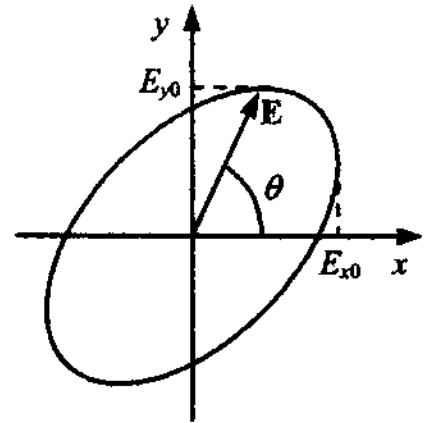


Рис. 9.1.4. В эллиптически поляризованной волне вектор $\mathbf{E}$ совершает вращение, причём его конец описывает эллипс

Наконец, отметим, что если интенсивности линейно поляризованных компонент ($I_x \sim E_{x0}^2$, $I_y \sim E_{y0}^2$) одинаковы: $I_x = I_y$, а сдвиг фаз их колебаний — случайная функция времени, то мы имеем дело с естественным — неполяризованным — светом. Из сказанного ясно, что такой свет немонохроматический и некогерентный.

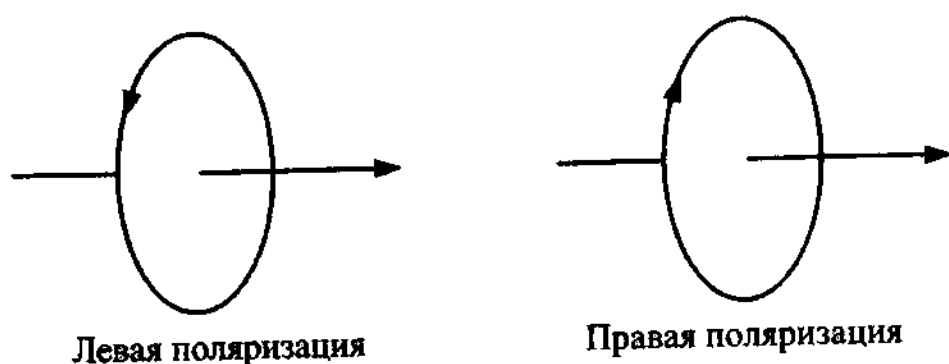


Рис. 9.1.5. Круговая левая и круговая правая поляризации электромагнитной волны

Свет с круговой поляризацией можно представлять как суперпозицию двух линейно поляризованных волн:

$$\mathbf{E} = \mathbf{E}_x + \mathbf{E}_y, \quad E_x = e_x E_0 \cos \omega t, \quad E_y = e_y E_0 \sin \omega t. \quad (9.1.7)$$

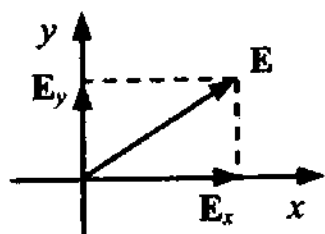


Рис. 9.1.6. Суперпозиция линейно поляризованных волн

Здесь e_x и e_y — единичные векторы в направлении соответствующих координатных осей. Векторная диаграмма сложения таких колебаний представлена на рис. 9.1.6. Если составляющие E_x и E_y осциллируют со сдвигом фаз $\pi/2$, то результирующая волна обладает круговой поляризацией.

Можно поставить обратную задачу: сконструировать линейно поляризованную волну как суперпозицию двух волн с круговой поляризацией. На векторной диаграмме рис. 9.1.7 проиллюстрировано соответствующее построение.

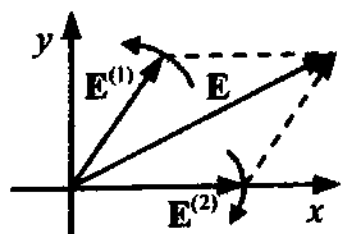


Рис. 9.1.7. Суперпозиция волн с круговой поляризацией даёт линейно поляризованную волну

Запишем две волны с круговой поляризацией, у которых вектор $\mathbf{E}$ совершает вращение в противоположных направлениях:

$$\begin{aligned} E_x^{(1)} &= a \cos(\omega t + \varphi), & E_y^{(1)} &= a \sin(\omega t + \varphi), \\ E_x^{(2)} &= a \cos(\omega t), & E_y^{(2)} &= -a \sin(\omega t). \end{aligned} \quad (9.1.8)$$

Волна $\mathbf{E}^{(1)}$ совершает вращение против часовой стрелки, а волна $\mathbf{E}^{(2)}$ — по часовой стрелке. Кроме того, у этих волн имеется сдвиг фаз φ , означающий, что в начальный момент векторы $\mathbf{E}^{(1)}$ и $\mathbf{E}^{(2)}$ образуют угол φ . Суперпозиция этих волн даёт

$$\begin{aligned} \mathbf{E} &= \mathbf{E}^{(1)} + \mathbf{E}^{(2)}, \\ E_x &= E_x^{(1)} + E_x^{(2)} = A_x \cos(\omega t + \varphi/2), \\ E_y &= E_y^{(1)} + E_y^{(2)} = A_y \cos(\omega t + \varphi/2), \\ A_x &= 2a \cos(\varphi/2), \quad A_y = 2a \sin(\varphi/2). \end{aligned} \quad (9.1.9)$$

Таким образом, мы получили линейно поляризованную волну, вектор $\mathbf{E}$ которой образует угол $\varphi/2$ с осью абсцисс (проходит по биссектрисе между векторами $\mathbf{E}^{(1)}$ и $\mathbf{E}^{(2)}$).

Говорят, что свет *частично поляризован*, если он представляет собой смесь полностью поляризованного и естественного (неполяризованного) излучения. Пусть интенсивности этих компонент равны соответственно $I_{\text{пол}}$ и $I_{\text{ест}}$. Тогда суммарная интенсивность составляет

$$I_0 = I_{\text{пол}} + I_{\text{ест}}. \quad (9.1.10)$$

Степенью поляризации излучения называется отношение

$$P = \frac{I_{\text{пол}}}{I_0} = \frac{I_{\text{пол}}}{I_{\text{пол}} + I_{\text{ест}}}. \quad (9.1.11)$$

9.2. Поляризаторы

Поляризатором называют прибор, предназначенный для получения полностью или частично поляризованного оптического излучения.

Действие поляризаторов, создающих линейно поляризованный свет, основано, в частности, на следующих явлениях:

- 1) линейный дихроизм (раздел 9.3.1),
- 2) поляризация при отражении (раздел 9.4),
- 3) поляризация при рассеянии (глава 11),
- 4) двойное лучепреломление (глава 12).

9.3. Поляроиды

9.3.1. Дихроизм

Дихроизмом называется явление, состоящее в том, что поглощение веществом света зависит от его поляризации. Иначе говоря, наиболее сильно поглощается свет с какой-то определённой поляризацией.

Термин «дихроизм» происходит от греч. *δύχρῳα* — двухцветный (*δύ* — два, *χρῳα* — цвет). Возникновение данного названия связано с тем, что вещества, обладающие свойством дихроизма, выглядят по-разному окрашенными при наблюдении под различными углами. Изменение окраски обусловлено зависимостью поглощения от длины волны света.

Например, кристалл апатита, освещаемый белым светом, кажется на просвет светло-желтым, если смотреть по направлению оптической оси, и зелёным — в перпендикулярном направлении. Окраску кристалла в этих

условиях наблюдения называют соответственно «осевой» и «базисной». При других направлениях наблюдения кристалл виден окрашенным в промежуточный цвет. Для описанного явления используется также термин *плеохроизм* (от греч. *πλεον* — больше), имея в виду, что число наблюдаемых цветов в окраске бывает большим.

Дихроизм открыл в 1809 году французский горный инженер и геолог Пьер Луи Антуан Кордье на минерале, названном в его честь «кордиерит» (с химическим составом $(\text{Mg,Fe})_2\text{Al}_3 [\text{Si}_5\text{AlO}_{18}]$), обладающем сильной анизотропией поглощения.

Выделяют:

- *линейный дихроизм*, когда различается поглощение света двух линейных взаимно перпендикулярных поляризаций,
- *круговой дихроизм*, когда различается поглощение излучения с левой и правой круговыми поляризациями.

Мерой дихроизма называют величину

$$D = \frac{K_{\max} - K_{\min}}{K_{\max} + K_{\min}}, \quad (9.3.1)$$

где $K_{\max}$, $K_{\min}$ — соответственно наибольшее и наименьшее значения коэффициента ослабления, отвечающего различным поляризациям ($K = I_{\text{прош}} / I_{\text{пад}}$).

9.3.2. Поляроиды

Один из способов получения линейно поляризованного излучения состоит в применении *поляроидов*.

Поляроиды обладают «разрешённым» направлением P (или *осью*) таким, что если вектор напряжённости электрического поля световой волны E параллелен этой оси, то эта волна свободно, практически без поглощения, проходит через поляроид. Если же $E \perp P$, то данная волна сильно поглощается веществом.

Таким образом, ставя на пути неполяризованного светового пучка поляроид P , мы получаем на выходе свет преимущественно с линейной поляризацией (рис. 9.3.1).

Действие поляроидов основано на линейном дихроизме, связанном с дихроизмом кристалликов или молекул полимера, внедрённых в прозрачную матрицу (из стекла или пластмассы) и пространственно однородно ориентированных в ней. Например, кристаллы турмалина (минерала из группы боросодержащих алюмосиликатов) обладают дихроизмом: в видимом диапазоне света они преимущественно поглощают свет с поляризацией, перпендикулярной их оптической оси (эту компоненту излучения

его кристаллы практически полностью поглощают уже при толщине около 1 мм).

Можно поставить обратную задачу: измерить степень поляризации излучения. Эта задача решается с помощью *анализатора*, в роли которого может быть использован такой же поляроид Π (рис. 9.3.1).

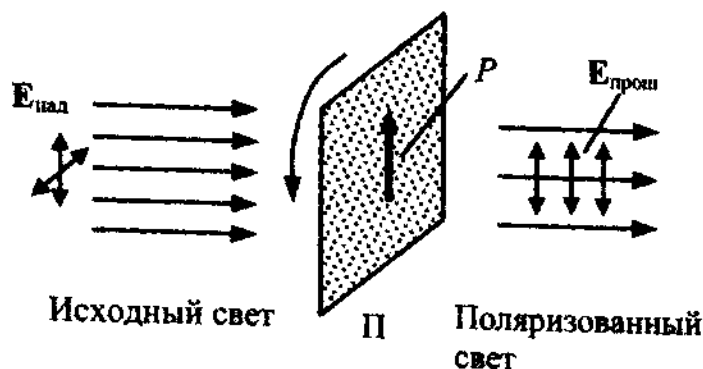


Рис. 9.3.1. Исходный свет $E_{\text{пад}}$, проходя через поляроид Π , становится поляризованным $E_{\text{прош}}$. Если исходный свет частично поляризован, то поляризатор позволяет измерить степень поляризации света. Вертикальной стрелкой показана ось поляроида P — направление, определяющее поляризацию прошедшего излучения

Пусть на такой анализатор падает частично поляризованный свет. Меняя ориентацию поляроида (поворачивая его ось), находим, когда наблюдаются максимальная (I_{max}) и минимальная (I_{min}) интенсивности проходящего излучения.

Допустим, что свет содержит неполяризованную компоненту (естественный свет) и компоненту с линейной поляризацией. Если поляроид идеальный, то он «отсекает» половину интенсивности естественного света, поскольку все поляризации в нём представлены с равным весом. В то же время линейно поляризованная компонента (в зависимости от ориентации оси поляроида) либо полностью пропускается поляроидом, либо полностью задерживается. Соответственно находим

$$I_{\text{max}} = \frac{1}{2} I_{\text{ест}} + I_{\text{пол}}, \quad I_{\text{min}} = \frac{1}{2} I_{\text{ест}}. \quad (9.3.2)$$

В результате степень поляризации исследуемого излучения определяется по формуле

$$P = \frac{I_{\text{пол}}}{I_{\text{пол}} + I_{\text{ест}}} = \frac{I_{\text{max}} - I_{\text{min}}}{I_{\text{max}} + I_{\text{min}}}. \quad (9.3.3)$$

9.3.3. Закон Малюса

Чтобы найти интенсивность волны, прошедшей через поляроид, разложим вектор E на две компоненты — вдоль и поперёк оси поляроида (рис. 9.3.2):

$$E_z = E \cos \varphi, \quad E_x = E \sin \varphi. \quad (9.3.4)$$

Представляя исходное поле как суперпозицию волн с поляризациями E_x и E_z , видим, что через поляроид проходит волна с поляризацией E_z . Другая же волна (с поляризацией E_x) полностью поглощается веществом. В итоге

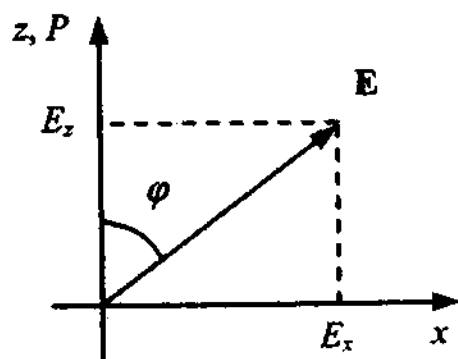
через поляроид проходит свет, поляризованный вдоль оси поляроида, причём его интенсивность ($I_{\text{прош}} \sim E_z^2$) равна

$$I_{\text{прош}} = I_0 \cos^2 \varphi, \quad (9.3.5)$$

$I_0 \sim E^2 = E_x^2 + E_z^2$ — интенсивность исходной волны, т.е. волны, падающей на поляроид.

Соотношение (9.3.5) называется *законом Малюса* по имени Э.-Л. Малюса, установившего его в 1809 г.

Рис. 9.3.2. Вектор электрического поля волны составляет угол φ с осью поляроида P (координатной осью z)



Пример. Пусть на пути неполяризованного света установлены два поляроида (поляризатор Π_1 и анализатор Π_2) с взаимно перпендикулярными осями (рис. 9.3.3). В этих условиях свет не проходит через данную оптическую систему.

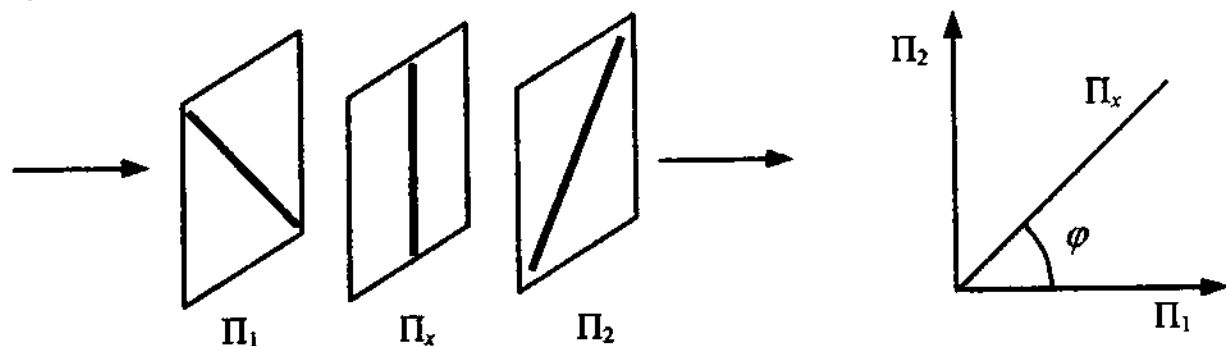


Рис. 9.3.3. Слева — свет проходит через поляризатор Π_1 и анализатор Π_2 с взаимно перпендикулярными осями. Между ними помещён поляроид Π_x ; справа — взаимное расположение осей поляризатора, анализатора и поляроида

Поместим между поляризатором и анализатором поляроид Π_x , ось которого образует угол φ с осью поляризатора Π_1 (рис. 9.3.3, справа). Найдём интенсивность волны, вышедшей через анализатор Π_2 , если интенсивность волны, падающей на систему, равна I_0 .

Через Π_1 пройдёт линейно поляризованный свет с интенсивностью $I_1 = I_0/2$. Согласно закону Малюса после промежуточного поляроида свет будет иметь интенсивность $I_x = I_1 \cos^2 \varphi$. Наконец, интенсивность света, вышедшего из анализатора Π_2 , окажется равной

$$I_2 = I_x \cos^2 \left(\frac{\pi}{2} - \varphi \right) = I_x \sin^2 \varphi = \frac{1}{8} I_0 \sin^2 2\varphi.$$

Максимум пропускания данной оптической системы достигается при $\varphi = \pi/4$, причём $(I_2/I_0)_{\max} = 1/8$.

Заметим, что система $(\Pi_1 + \Pi_2)$ позволяет изучать анизотропию пропускания объектов (Π_x) , помещённых между ними, и используется при работе с поляризованным излучением.

9.4. Поляризация света при отражении

9.4.1. Явление Брюстера

Поляризация излучения возникает при отражении света от границы раздела сред с разными оптическими свойствами.

Пусть свет падает на границу раздела двух однородных диэлектрических сред (рис. 9.4.1). В результате возникают отражённая и преломлённая волны, направления которых определяются законами отражения и преломления:

$$\theta = \theta', \quad n_1 \sin \theta = n_2 \sin \theta'' \quad (9.4.1)$$

Вместе с тем амплитуды этих волн зависят от поляризации. Как говорилось в главе 1, рассматривают s - и p -поляризации волн (рис. 9.4.2). В первом случае (s -поляризация) вектор $\mathbf{E}$ перпендикулярен плоскости падения. Во втором случае (p -поляризация) вектор $\mathbf{E}$ лежит в плоскости падения.

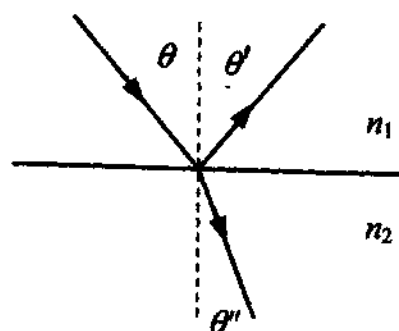
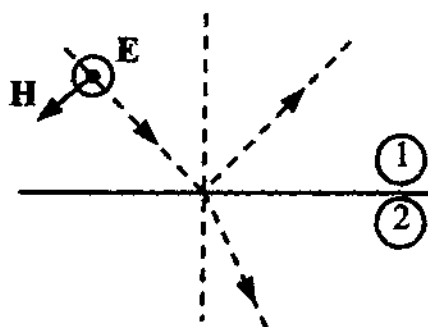
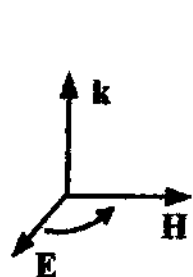
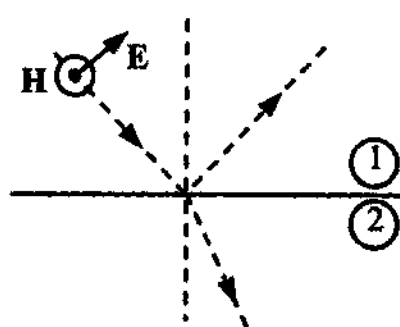


Рис. 9.4.1. Падающая, отражённая и преломлённая волны на границе раздела двух сред



s -поляризация



p -поляризация

Рис. 9.4.2. Векторы $\mathbf{E}$ и $\mathbf{H}$ в падающей волне в некоторый момент времени. Слева — s -поляризованная волна, справа — p -поляризованная волна. На рисунке кружок с точкой означает, что соответствующий вектор направлен «на нас»

Согласно формулам Френеля (1.7.2), (1.7.3) коэффициент отражения для s -поляризованной волны не обращается в нуль ни при каких углах падения.

Для p -поляризованной волны из (1.7.4), (1.7.5) следует, что коэффициент отражения обращается в нуль при угле падения $\theta = \theta_B$, удовлетворяющем условию

$$\operatorname{tg} \theta_{\text{Б}} = n, \quad (9.4.2)$$

где $n = n_2/n_1$ — относительный показатель преломления среды, на которую падает волна. Это значит, что при падении под таким углом отражённая волна оказывается полностью поляризованной, обладающей s -поляризацией. Данное явление названо по имени шотландского физика Дэвида Брюстера, открывшего его в 1815 г., а угол $\theta_{\text{Б}}$ — углом Брюстера.

На рис. 9.4.3 показана зависимость энергетических коэффициентов отражения от угла падения для случаев: s -поляризованной ($R_{\perp}$), p -поляризованной ($R_{\parallel}$) и неполяризованной ($R_{\text{ест}}$) волн.

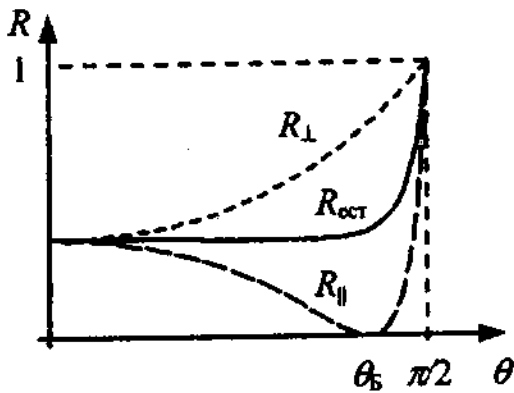


Рис. 9.4.3. Угловые зависимости коэффициентов отражения для излучения с s - и p -поляризацией, а также для неполяризованного излучения

Если на границу раздела сред падает естественный свет, то две взаимно ортогональные поляризованные компоненты представлены в нём с равным весом:

$$I_{\text{ест}} = I_{\perp} + I_{\parallel}, \quad I_{\perp} = I_{\parallel} = (1/2) I_{\text{ест}}. \quad (9.4.3)$$

Ввиду некогерентности этих компонент интенсивность отражённого света можно представить в виде

$$I_{\text{отр}} = R_{\perp} I_{\perp} + R_{\parallel} I_{\parallel} = \frac{1}{2} (R_{\perp} + R_{\parallel}) I_{\text{ест}} \equiv R_{\text{ест}} I_{\text{ест}}. \quad (9.4.4)$$

Таким образом, коэффициент отражения для естественного (неполяризованного) света равен

$$R_{\text{ест}} = (R_{\perp} + R_{\parallel})/2. \quad (9.4.5)$$

Явление Брюстера непосредственно следует из формул Френеля для p -поляризованной волны:

$$R_{\parallel} = \frac{\operatorname{tg}^2(\theta - \theta'')}{\operatorname{tg}^2(\theta + \theta'')} = \left(\frac{n^2 \cos \theta - \sqrt{n^2 - \sin^2 \theta}}{n^2 \cos \theta + \sqrt{n^2 - \sin^2 \theta}} \right)^2. \quad (9.4.6)$$

Из второго представления для коэффициента отражения в (9.4.6) при $R_{\parallel} = 0$ следует выражение для угла Брюстера (9.4.2). Из первого же выражения следует, что угол Брюстера в сумме с углом преломления составляет 90° :

$$\theta_{\text{Б}} + \theta'' = \pi/2. \quad (9.4.7)$$

Данное условие позволяет понять причину явления Брюстера.

Действительно, отражённая волна возникает вследствие того, что излучение, падающее на вещество, проникает в него и возбуждает колеба-

ния электронов. Возникающие в результате этого волны от всех электронов суммируются и формируют отражённую волну. Однако при выполнении равенства (9.4.7) преломлённый и отражённый лучи образуют прямой угол (рис. 9.4.4, слева). При этом в случае p -поляризованной волны вектор напряжённости поля преломлённой волны E' совершает колебания в направлении, параллельном волновому вектору отражённой волны. Однако, как видно из диаграммы направленности (рис. 9.4.4, справа), именно в этом направлении интенсивность излучения практически нулевая, что и приводит к отсутствию отражённого излучения при падении под углом Брюстера $\theta = \theta_B$.

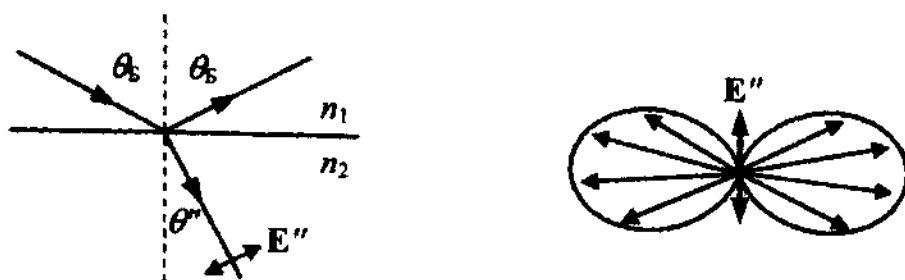


Рис. 9.4.4. Слева — падающая, отражённая и преломлённая волны на границе раздела двух сред в случае p -поляризованной волны, справа — диаграмма направленности излучения электронов, стимулированного преломлённой волной

Заметим, что при наличии поглощения излучения в среде p -поляризованная компонента частично отражается даже при падении под углом Брюстера: $R_{||}(\theta_B) \neq 0$. Дело в том, что отражённая волна формируется как результат интерференции отдельных волн, производимых осциллирующими зарядами. Однако, если излучение поглощается, интерференционное гашение оказывается неполным, поскольку амплитуды волн от отдельных электронов убывают с расстоянием не только за счёт расходимости излучения, но и за счёт поглощения. Последний фактор и приводит к тому, что отражение оказывается лишь частичным.

9.4.2. Стопа Столетова

На явлении Брюстера основан один из способов получения линейно поляризованного света.

При падении света на одну пластинку под углом Брюстера (рис. 9.4.5) интенсивность отражённого линейно поляризованного света очень мала. Например, от границы воздух–стекло ($n = 1,5$) отражается около 3,75 % интенсивности падающего луча.

Для прошедшего излучения отношение интенсивностей компонент с s - и p -поляризациями равно

$$\delta = \frac{(I_{\text{прош}})_{\perp}}{(I_{\text{прош}})_{\parallel}} = \frac{d_{\perp}^2}{d_{\parallel}^2}. \quad (9.4.8)$$

По формулам Френеля амплитудные коэффициенты прохождения $d_{\perp}$ и $d_{\parallel}$ равны

$$d_{\perp} = \frac{2 \sin \theta'' \cos \theta}{\sin(\theta + \theta'')}, \quad d_{\parallel} = \frac{2 \sin \theta'' \cos \theta}{\sin(\theta + \theta'') \cos(\theta - \theta'')}. \quad (9.4.9)$$

Следовательно, $\delta = \cos^2(\theta - \theta'')$. При падении под углом Брюстера

$$\delta = \frac{4n^2}{(1+n^2)^2} < 1. \quad (9.4.10)$$

Таким образом, в прошедшем свете доля компоненты с s -поляризацией уменьшается. В частности, при $n = 1,5$ находим $\delta = 0,852$.

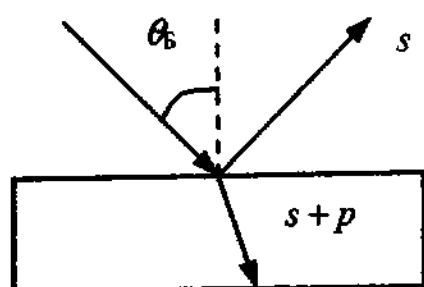


Рис. 9.4.5. Преломление света, падающего под углом Брюстера. Индексы s и p указывают поляризацию отраженного и преломленного лучей

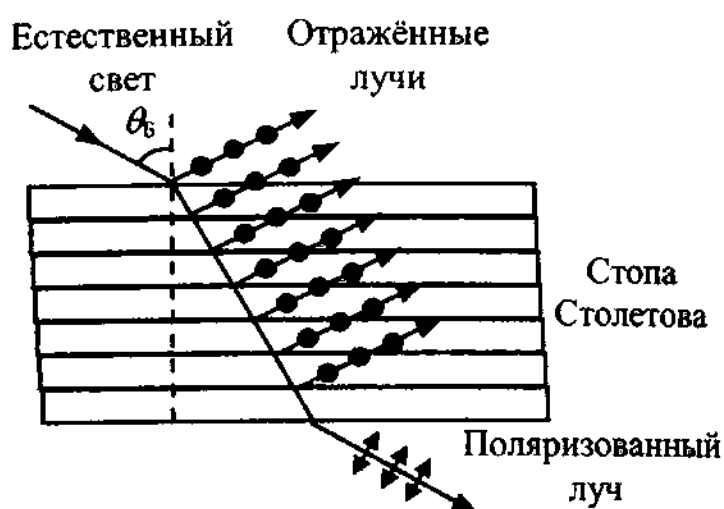


Рис. 9.4.6. Стопа Столетова, позволяющая на основе явления Брюстера делить естественный свет на две линейно поляризованные компоненты: отражённый свет преимущественно s -поляризован, а прошедший — p -поляризован

Для того чтобы увеличить степень поляризации прошедшего и отражённого излучения применяют несколько пластинок, сложенных в *стопу Столетова* (рис. 9.4.6). При её прохождении доля s -поляризованного света убывает, а p -поляризованного растёт — отражённый свет оказывается s -поляризованным, а прошедший — p -поляризованным. Например, для стопы, выполненной из $N = 16$ стеклянных пластинок (с показателем преломления $n = 1,5$) степень поляризации составляет около 99%.

ГЛАВА 10

ДИСПЕРСИЯ ВОЛН

10.1. Фазовая и групповая скорость

10.1.1. Дисперсия

До сих пор, изучая распространение волн, мы предполагали, что связь волнового числа (вектора) и частоты линейная:

$$\omega = kv, \quad (10.1.1)$$

где величина v имеет постоянное значение (не зависит от частоты волны ω) и определяется только свойствами среды, в которой распространяется волна. Эта величина, которую мы сейчас обозначим как

$$v_\phi = \omega/k, \quad (10.1.2)$$

называется *фазовой скоростью* волны. Как видно из формулы

$$u(x, t) = a \cos(kx - \omega t) = a \cos[k(x - v_\phi t)], \quad (10.1.3)$$

со скоростью v_ϕ перемещается точка волны, в которой фаза имеет фиксированное значение $\varphi(x, t) = \omega t - kx = \text{const}$.

В общем случае связь частоты и волнового числа, называемая *законом дисперсии*:

$$\omega = \omega(k), \quad (10.1.4)$$

может быть более сложной. Соответственно фазовая скорость $v_\phi = \omega/k$ оказывается зависящей от частоты волны или, что эквивалентно, от длины волны $\lambda = 2\pi/k$: $v_\phi = v_\phi(\lambda)$.

Зависимость фазовой скорости от длины волны называется *дисперсией* (от лат. *dispersio* — рассеяние). Среда, в которой имеет место дисперсия, называются *диспергирующими*.

В случае света дисперсией называют зависимость абсолютного показателя преломления вещества от частоты (или длины волны) света (частотная дисперсия). Поскольку фазовая скорость связана с показателем преломления соотношением $v_\phi = c/n$, то оба определения эквивалентны.

Дисперсию экспериментально открыл Ньютон около 1672 года, хотя теоретически она была объяснена значительно позднее. Первоначально термин «Дисперсия света» был введен для описания разложения белого света в спектр при преломлении в призме, ныне же он употребляется в более широком смысле.

10.1.2. Групповая скорость

Рассмотрим две волны, имеющие частоты ω_1 и ω_2 и соответствующие им волновые числа $k_1 = k(\omega_1)$ и $k_2 = k(\omega_2)$. Если фазовые скорости этих волн отличаются друг от друга, то результат сложения волн в разные моменты времени будет различным. Пусть волны задаются уравнениями

$$\begin{aligned} u_1(x, t) &= a \sin(k_1 x - \omega_1 t), \\ u_2(x, t) &= a \sin(k_2 x - \omega_2 t). \end{aligned} \quad (10.1.5)$$

Результирующее колебание есть сумма этих волн:

$$\begin{aligned} u(x, t) &= u_1(x, t) + u_2(x, t) = \\ &= 2a \cos\left[\frac{k_2 - k_1}{2}x - \frac{\omega_2 - \omega_1}{2}t\right] \sin\left[\frac{k_2 + k_1}{2}x - \frac{\omega_2 + \omega_1}{2}t\right]. \end{aligned} \quad (10.1.6)$$

Пусть волновые числа k_1 и k_2 , $k_2 > k_1$, близки, т.е. $k_2 - k_1 \ll k_1$. Тогда, обозначая

$$\begin{aligned} \frac{k_1 + k_2}{2} &= k, & k_2 - k_1 &= \Delta k, \\ \frac{\omega_2 + \omega_1}{2} &= \omega, & \omega_2 - \omega_1 &\approx \frac{d\omega}{dk} \Delta k, \end{aligned}$$

перепишем (10.1.6) в виде

$$u(x, t) = 2a \cos\left[\frac{\Delta k}{2}\left(x - \frac{d\omega}{dk}t\right)\right] \sin(kx - \omega t). \quad (10.1.7)$$

Полученное выражение аналогично уравнению биений: в каждой фиксированной точке наблюдения колебания имеют вид биений с периодом

$$\tau = \frac{2\pi}{|\omega_2 - \omega_1|} \approx \frac{2\pi}{\Delta k |d\omega/dk|}. \quad (10.1.8)$$

Форма волны в некоторый момент времени показана на рис. 10.1.1.

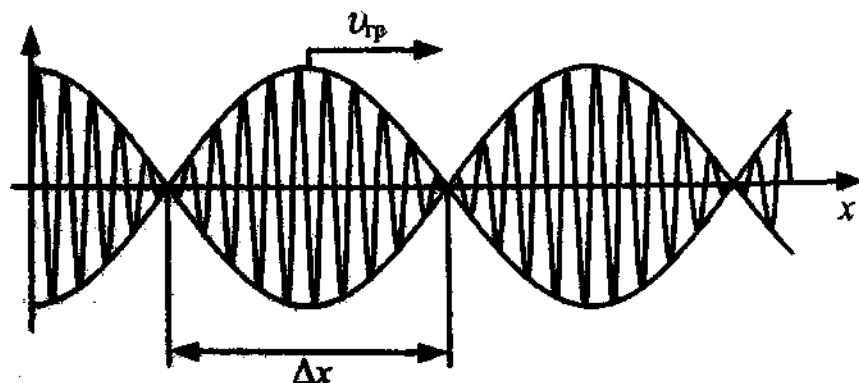


Рис. 10.1.1. Биения как результат сложения двух бегущих волн с разными частотами в диспергирующей среде

С другой стороны, фиксируя момент наблюдения, мы видим модулированную волну, у которой амплитуда модуляции даётся множителем

$$2a \cos \left[\frac{\Delta k}{2} \left(x - \frac{d\omega}{dk} t \right) \right]. \quad (10.1.9)$$

Рассматривая теперь эволюцию волны во времени, мы наблюдаем неизменный профиль, то есть огибающую высокочастотных колебаний, если движемся вдоль волны со скоростью

$$v_{\text{гр}} = d\omega/dk. \quad (10.1.10)$$

При таком движении множитель (10.1.9) сохраняет неизменное значение.

На рис. 10.1.1 показан профиль модулированной волны (10.1.7). Со скоростью $v_{\text{гр}}$ смещается огибающая высокочастотных колебаний. Величина, введенная соотношением (10.1.10), называется *групповой скоростью*.

Заметим, что период биений Δx на рис. 10.1.1 связан с величиной τ в (10.1.8) соотношением

$$\Delta x = \frac{\pi}{|\Delta k/2|} = \frac{2\pi}{|\Delta k|} \approx \frac{2\pi}{\Delta \omega |dk/d\omega|} = v_{\text{гр}} \tau.$$

Покажем, что определение групповой скорости (10.1.10) остается справедливым и в том случае, когда складывается произвольное число волн. Для этого рассмотрим суперпозицию:

$$u(x, t) = \sum_{k=k_1}^{k_2} a(k) \exp[i(kx - \omega(k)t)], \quad (10.1.11)$$

в которой будем предполагать, что суммирование производится по узкому интервалу волновых чисел: $k \in [k_1, k_2]$, $k_2 - k_1 \ll k_1$. Тогда можно разложить $\omega(k)$ в окрестности среднего значения волнового числа:

$$\omega(k) \approx \omega_0 + (k - k_0) \frac{d\omega}{dk} \Big|_{k=k_0}, \quad (10.1.12)$$

$$k_0 = \frac{k_1 + k_2}{2}, \quad \omega_0 = \omega(k_0).$$

Подставляя (10.1.12) в (10.1.11) и переходя к суммированию по переменной $q = k - k_1$, получим

$$\begin{aligned} u(x, t) &= \exp[i(k_0 x - \omega_0 t)] \sum_{q=q_1}^{q_2} a(k_0 + q) \exp[iq(x - v_{\text{гр}} t)] \equiv \\ &\equiv \exp[i(k_0 x - \omega_0 t)] F(x - v_{\text{гр}} t). \end{aligned} \quad (10.1.13)$$

Входящая сюда групповая скорость $v_{гр}$ определена равенством (10.1.10).

В формуле (10.1.13) суммирование производится по волновым числам q в диапазоне

$$q \in [q_1, q_2], \quad q_1 = -\Delta k/2, \quad q_2 = \Delta k/2, \quad \Delta k = k_2 - k_1.$$

В силу неравенства $|q| \sim \Delta k \ll k_0$ функция $F(x - v_{гр}t)$ меняется много медленнее, чем экспонента $\exp[i(k_0x - \omega_0t)]$. Поэтому суперпозицию волн $u(x, t)$ можно рассматривать как модулированную волну со средними волновым числом k_0 и частотой ω_0 , причём амплитуда модуляции даётся функцией $F(x - v_{гр}t)$. Именно с этой функцией связывают представление о *пути волн*, или *волновом пакете*. В соответствии со сказанным говорят, что волновой пакет движется с групповой скоростью, тогда как отдельные волны — с фазовой скоростью.

10.1.3. Локализация волнового пакета

Путь волн может быть локализованным в пространстве. Рассмотрим для примера набор волн, для которых в (10.1.11)

$$\begin{aligned} a(k) &= (A_0/\Delta k) dk, \quad k_1 < k < k_2, \\ k_1 &= k_0 - \Delta k/2, \quad k_2 = k_0 + \Delta k/2. \end{aligned} \quad (10.1.14)$$

Тогда, выполняя интегрирование, получим

$$\begin{aligned} u(x, t) &= A_0 \int_{k_1}^{k_2} \exp[ik(x - v_{гр}t)] \frac{dk}{\Delta k} = \\ &= A_0 \exp[i(k_0x - \omega_0t)] \frac{\sin\left[\frac{\Delta k}{2}(x - v_{гр}t)\right]}{\frac{\Delta k}{2}(x - v_{гр}t)}. \end{aligned} \quad (10.1.15)$$

Видно, что характерная пространственная ширина волнового пакета определяется из условия $\frac{\Delta k}{2}(x - v_{гр}t) \sim \pi$ и составляет

$$\Delta x \sim 2\pi/\Delta k. \quad (10.1.16)$$

10.1.4. Расплывание волнового пакета

Понятие групповой скорости имеет ограниченную область применимости, поскольку в среде с дисперсией волновой пакет с течением времени не только перемещается, но и, как правило, меняет форму.

Чтобы убедиться в этом, учтём в разложении (10.1.11) слагаемое более высокого порядка:

$$\omega(k) \approx \omega_0 + (k - k_0) \left. \frac{d\omega}{dk} \right|_{k=k_0} + \frac{1}{2} (k - k_0)^2 \left. \frac{d^2\omega}{dk^2} \right|_{k=k_0}. \quad (10.1.17)$$

Добавочный член приводит с течением времени к изменению относительных фаз волн, образующих волновой пакет, на величину $(t/2)(k - k_0)^2 \left(d^2\omega/dk^2 \right) \Big|_{k=k_0}$, так что

$$F = \sum_q a(k_0 + q) \exp \left[iq(x - v_{гр}t) - i \frac{q^2}{2} \left(\frac{d^2\omega}{dk^2} \right)_0 t \right] \equiv F(x - v_{гр}t, t) \quad (10.1.18)$$

(индекс «0» означает, что производная берётся при $k = k_0$). К моменту, когда сдвиг фаз волн достигает π :

$$\frac{t_{\max}}{2} q^2 \left. \frac{dv_{гр}}{dk} \right|_{k=k_0} \sim \pi, \quad (10.1.19)$$

волновой пакет фактически прекращает существование, разваливается. Полагая для оценки $q \sim \Delta k/2$, получим характерное время существования волнового пакета:

$$t \ll t_{\max} \sim \frac{8\pi}{(\Delta k)^2 \left| d^2\omega/dk^2 \right|_0}. \quad (10.1.20)$$

Перепишем полученную формулу в другом виде. Для этого заметим, что

$$\frac{d^2\omega}{dk^2} = \frac{dv_{гр}}{dk}.$$

Кроме того, перейдём к дифференцированию по длине волны $\lambda = 2\pi/k$:

$$\frac{dv_{гр}}{dk} = \frac{dv_{гр}/d\lambda}{dk/d\lambda} = -\frac{\lambda^2}{2\pi} \frac{dv_{гр}}{d\lambda}. \quad (10.1.21)$$

Поскольку при малой ширине спектра рассматриваемого пучка волн $\Delta k = -(2\pi/\lambda^2)\Delta\lambda$, то в результате получаем следующую оценку времени существования волнового пакета:

$$t_{\max} \sim \frac{4\lambda^2}{(\Delta\lambda)^2 \left| dv_{гр}/d\lambda \right|_0}. \quad (10.1.22)$$

В качестве примера рассмотрим волновой пакет, имеющий в начальный момент вид

$$A(x, 0) = A_0 \exp(-x^2/x_0^2). \quad (10.1.23)$$

Примем следующий закон дисперсии:

$$\omega = ck + bk^2/2. \quad (10.1.24)$$

Видно, что фазовые скорости отдельных волн зависят от их волновых чисел: $v_\phi = \omega/k = c + bk/2$. Поэтому компоненты суперпозиции будут смещаться относительно друг друга, и волновой пакет со временем расплывётся.

Чтобы воспользоваться формулой (10.1.18), найдём компоненты Фурье функции (10.1.23):

$$a(k) = \int_{-\infty}^{\infty} A(x, 0) e^{-ikx} dx = \sqrt{\pi} x_0 A_0 \exp\left(-\frac{1}{4} k^2 x_0^2\right). \quad (10.1.25)$$

Теперь временная зависимость пакета $A(x, t)$ находится следующим образом:

$$\begin{aligned} A(x, t) &= \int_{-\infty}^{\infty} a(k) e^{ikx - i\omega(k)t} \frac{dk}{2\pi} = \\ &= \frac{\sqrt{\pi}}{2\pi} x_0 A_0 \int_{-\infty}^{\infty} \exp\left(-\frac{1}{4} k^2 x_0^2\right) \exp\left[i\left(k(x - ct) - \frac{1}{2} bk^2 t\right)\right] \frac{dk}{2\pi}. \end{aligned} \quad (10.1.26)$$

Вычисление интеграла даёт

$$A(x, t) = \frac{A_0}{\sqrt{1 + 2ibt/x_0^2}} \exp\left(-\frac{(x - ct)^2}{x_0^2 + 2ibt}\right). \quad (10.1.27)$$

Отсюда для интенсивности $I = |A|^2$ получаем выражение

$$I(x, t) = \frac{I_0}{1 + (2bt/x_0^2)^2} \exp\left\{-\frac{(x - ct)^2}{x_0^2 [1 + (2bt/x_0^2)^2]}\right\}. \quad (10.1.28)$$

Таким образом, рассматриваемый волновой пакет движется с групповой скоростью

$$v_{гр} = d\omega/dk|_{k=0} = c. \quad (10.1.29)$$

При этом его эффективная ширина растёт:

$$x_{eff}(t) = x_0 \sqrt{1 + (2bt/x_0^2)^2}. \quad (10.1.30)$$

Одновременно убывает его амплитуда:

$$I(0, t) = I_0 \frac{x_0^2}{x_{eff}^2(t)}. \quad (10.1.31)$$

Заметим, что в (10.1.27) отсутствует фазовый множитель бегущей волны $\exp(ikx - i\omega t)$. Дело в том, что исходный волновой пакет (10.1.23) имеет симметричную форму: $A(x, 0) = A(-x, 0)$. Поэтому он составлен из одинакового набора волн, бегущих в положительном и отрицательном направлениях, что означает выбор $k_0 = 0$ в (10.1.17), (10.1.18). Появление же дрейфа с групповой скоростью c связано наличием слагаемого ck в законе дисперсии (10.1.24), выделяющего одно из направлений в пространстве в соответствии со знаком групповой скорости (10.1.29).

Таким образом, отход от линейной зависимости $\omega(k) = ck$ приводит к расплыванию волнового пакета. Характерное время расплывания есть

$$t \sim x_0^2 / 2b. \quad (10.1.32)$$

Это время тем больше, чем меньше коэффициент при k^2 в законе дисперсии (10.1.24).

10.1.5. Формула Рэлея

Найдем связь фазовой и групповой скоростей. Подставляя в (10.1.10) соотношение $\omega = kv_\phi$, получим

$$v_{гр} = v_\phi + k \frac{dv_\phi}{dk}. \quad (10.1.33)$$

Переходя здесь от волнового числа к длине волны $\lambda = 2\pi/k$, перепишем найденную связь в виде

$$v_{гр} = v_\phi - \lambda \frac{dv_\phi}{d\lambda}. \quad (10.1.34)$$

Это соотношение называется *формулой Рэлея*.

Как следует из (10.1.34), для сред без дисперсии ($v_\phi = \text{const}$) фазовая и групповая скорости совпадают:

$$v_{гр} = v_\phi. \quad (10.1.35)$$

На рис. 10.1.2 приведена графическая интерпретация формулы (10.1.34), предложенная П. Эренфестом. Штриховой линией показана касательная к кривой $v_\phi = v_\phi(\lambda)$. Видно, что эта прямая пересекает ось ординат в точке с координатой, равной групповой скорости $v_{гр}$.

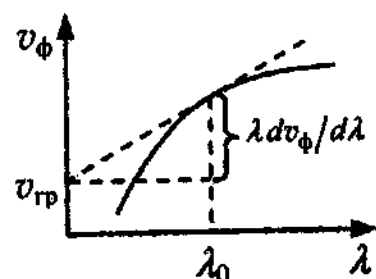


Рис. 10.1.2. Интерпретация формулы Рэлея согласно Эренфесту

В случае электромагнитных волн фазовая скорость связана с показателем преломления среды формулой

$$v_{\phi} = c/n. \quad (10.1.36)$$

Данное утверждение следует из формулы $k = \omega n/c$, то есть $v_{\phi} = \omega/k = c/n$.

Дисперсия связана с тем, что показатель преломления зависит от длины волны излучения: $n = n(\lambda)$. В связи с этим в случае световых волн *дисперсией* называют зависимость показателя преломления среды от длины волны.

Свяжем групповую скорость волны с показателем преломления. Подставим (10.1.36) в формулу Рэлея (10.1.34):

$$v_{\text{гр}} = v_{\phi} - \lambda \frac{d}{d\lambda} \left(\frac{c}{n} \right) = \frac{c}{n} + \frac{c\lambda}{n^2} \frac{dn}{d\lambda} = v_{\phi} \left(1 + \frac{\lambda}{n} \frac{dn}{d\lambda} \right). \quad (10.1.37)$$

Здесь присутствует производная от показателя преломления по длине волны. Последняя сама зависит от показателя преломления:

$$\lambda = \frac{2\pi c}{n\omega}. \quad (10.1.38)$$

Поэтому во многих случаях желательно иметь формулу, куда входит производная по частоте, которая не зависит от показателя преломления. Получим соответствующее выражение. Преобразуем сначала производную $dn/d\lambda$, используя (10.1.38):

$$\frac{dn}{d\lambda} = \frac{dn/d\omega}{d\lambda/d\omega} = \frac{dn/d\omega}{\frac{d}{d\omega} \left(\frac{2\pi c}{n\omega} \right)} = - \frac{(n\omega)^2}{2\pi c} \frac{dn/d\omega}{d(n\omega)/d\omega}. \quad (10.1.39)$$

Подстановка этого выражения в (10.1.37) даёт

$$v_{\text{гр}} = \frac{c}{n} \left(1 + \frac{\lambda}{n} \frac{dn}{d\lambda} \right) = \frac{c}{n + \omega dn/d\omega}. \quad (10.1.40)$$

10.2. Классическая теория дисперсии

10.2.1. Зависимость диэлектрической проницаемости среды от частоты излучения

Классическая теория дисперсии, предложенная Х. А. Лоренцем, основана на предположении, что среда представляет собой газ, состоящий из невзаимодействующих между собой атомов. В атоме имеется один электрон, называемый *оптическим*, квазиупруго связанный с атомом. Возвращающая сила равна $F_{\text{возвр}} = -\beta \mathbf{r}$, где $\mathbf{r}$ — радиус-вектор электрона относительно положения равновесия.

На рис. 10.2.1 показан качественный вид зависимости силы, действующей на электрон, от расстояния до центра. Положению равновесия отвечает расстояние $R = R_0$ от центра. Видно, что линейная зависимость силы от смещения имеет место только при малых отклонениях электрона от положения равновесия. В сильных полях, когда отклонения электрона оказываются не слишком малыми, зависимость оказывается нелинейной.

Уравнение, описывающее поведение электрона под действием электромагнитной волны, имеет вид

$$m \frac{dv}{dt} = F_{\text{возвр}} + F_{\text{тр}} + eE; \quad (10.2.1)$$

$$F_{\text{возвр}} = -\beta r, \quad F_{\text{тр}} = -\eta v, \quad E = E_0 e^{-i\omega t}.$$

Здесь $F_{\text{возвр}}$ — сила, возвращающая электрон в его положение равновесия, $F_{\text{тр}}$ — сила трения, учитывающая потери энергии вследствие излучения (при ускоренном движении) и передачи энергии среде. Предположив, что амплитуда колебаний электрона мала по сравнению с длиной электромагнитной волны, можно считать поле, действующее на отдельный электрон, однородным. Это позволяет считать коэффициент E_0 постоянным, не зависящим от смещения электрона.

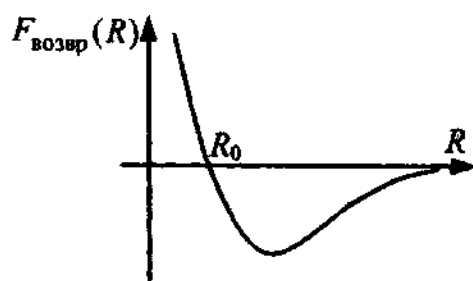


Рис. 10.2.1. Зависимость силы, действующей на электрон со стороны атома, от расстояния R до центра

Вводя собственную частоту осциллятора $\omega_0 = \sqrt{\beta/m}$ и коэффициент затухания $\gamma = \eta/2m$, перепишем уравнение (10.2.1) в виде

$$\ddot{r} + 2\gamma\dot{r} + \omega_0^2 r = \frac{eE_0}{m} e^{-i\omega t}. \quad (10.2.2)$$

Решение этого уравнения имеет вид

$$r = \frac{e/m}{\omega_0^2 - \omega^2 - 2i\gamma\omega} E. \quad (10.2.3)$$

Дипольный момент атома равен $p = er$. Заметим, что мы получили выражение для поляризуемости молекулы:

$$p = \beta E \Rightarrow \beta = \frac{e^2/m}{\omega_0^2 - \omega^2 - 2i\gamma\omega}. \quad (10.2.4)$$

Если в единице объёма среды находится N атомов, то вектор поляризации среды $P = Nr$. Поскольку что $D = E + 4\pi P = \epsilon E$, то находим диэлектрическую проницаемость:

$$\varepsilon = 1 + \frac{\omega_p^2}{\omega_0^2 - \omega^2 - 2i\gamma\omega}, \quad \omega_p^2 = \frac{4\pi Ne^2}{m}. \quad (10.2.5)$$

Величина ω_p называется *плазменной частотой*.

Если возможны несколько резонансных частот ω_{0k} , отвечающих разным связям электронов с молекулами, то формулу (10.2.5) следует записать в виде

$$\varepsilon = 1 + \frac{4\pi Ne^2}{m} \sum_k \frac{f_k}{\omega_{0k}^2 - \omega^2 - 2i\gamma_k \omega}. \quad (10.2.6)$$

Входящие сюда коэффициенты f_k , называемые *силами осцилляторов*, связаны с концентрациями электронов N_k , $\sum_k N_k = N$, отвечающих своим резонансным частотам ω_{0k} :

$$f_k = N_k / N, \quad \sum_k f_k = 1.$$

Таким образом, диэлектрическая проницаемость зависит от частоты колебаний электромагнитной волны. Она в общем случае является также комплексной величиной.

Хенрик Антон Лоренц — нидерландский физик-теоретик, лауреат Нобелевской премии (1902) за разработку теории эффекта Зеемана — расщепления спектральных линий в магнитном поле. Он создал классическую электронную теорию. Лоренц предположил, что все молекулы вещества состоят из электрически заряженных частиц, осуществив тем самым синтез идей электронного строения вещества и теории электромагнитного поля Максвелла. В 1878 г. вышла статья Лоренца «О соотношении между скоростью распространения света и плотностью и составом среды», в которой на основе электронной теории установлено соотношение между показателем преломления и плотностью среды. В этой работе Лоренц развивал теорию дисперсии среды, хотя завершена она была уже после открытия электрона (Дж. Дж. Томсон, 1897) и установления строения атома (Э. Резерфорд, 1911). Лоренцу принадлежит значительный вклад в создание теории относительности (объяснение опытов Майкельсона, вывод преобразований Лоренца, установление зависимости массы тела от его скорости, введение понятия о «местном времени»).

10.2.2. Нормальная и аномальная дисперсии

Обозначим

$$\sqrt{\varepsilon} = n + i\kappa. \quad (10.2.7)$$

Величина $n = \operatorname{Re} \sqrt{\varepsilon}$ называется *показателем преломления*, а величина $\kappa = \operatorname{Im} \sqrt{\varepsilon}$ — *показателем поглощения (затухания)*. Если ε — действительная неотрицательная величина, то введенный в (10.2.7) показатель преломления совпадает с обычным показателем преломления.

На рис. 10.2.2 показан вид зависимостей показателей преломления и поглощения от частоты, $n = n(\omega)$, $\kappa = \kappa(\omega)$.

Асимптотические значения показателя преломления следующие:

$$n_0 = n(0) = \sqrt{1 + \omega_p^2 / \omega_0^2} > 1, \quad n \rightarrow 1 \text{ при } \omega \rightarrow \infty. \quad (10.2.8)$$

Из (10.2.5) следует, что в случае слабого трения ($\gamma \ll \omega_0$) показатель поглощения достигает максимального значения при $\omega \approx \omega_0$ и составляет

$$\kappa_{\max} \approx \text{Im} \sqrt{\epsilon(\omega_0)} \approx \text{Im} \sqrt{1 + \frac{i\omega_p^2}{2\gamma\omega_0}} \approx \frac{\omega_p}{2\sqrt{\gamma\omega_0}}. \quad (10.2.9)$$

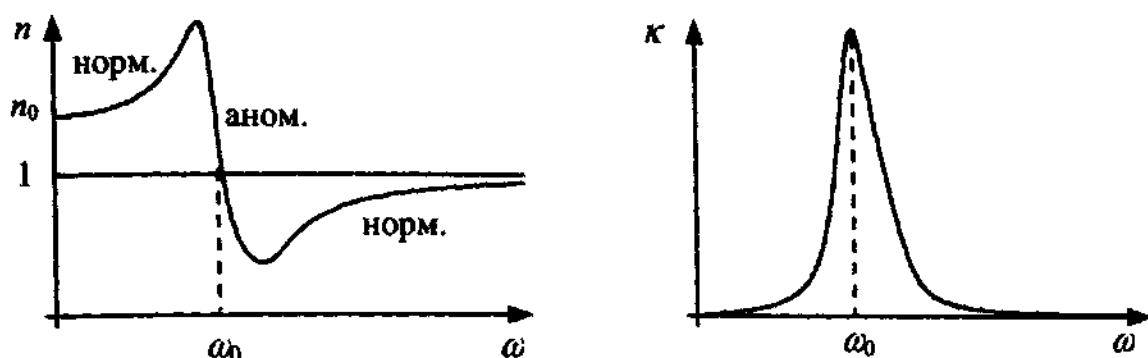


Рис. 10.2.2. Частотные зависимости показателей преломления $n(\omega)$ и поглощения $\kappa(\omega)$. Указаны области нормальной (норм.) и аномальной (аном.) дисперсий

Для случая, когда диэлектрическая проницаемость мало отличается от единицы, можно получить следующие выражения для показателей преломления и поглощения:

$$n = 1 + \frac{\omega_p^2}{2} \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + 4\gamma^2 \omega^2}, \quad \kappa = \frac{\gamma \omega_p^2 \omega}{(\omega_0^2 - \omega^2)^2 + 4\gamma^2 \omega^2}. \quad (10.2.10)$$

Из этих формул нетрудно увидеть все основные особенности поведения $n = n(\omega)$, $\kappa = \kappa(\omega)$, отражённые на рис. 10.2.2.

Область частот, где показатель преломления возрастает:

$$dn/d\omega > 0 \text{ или } dn/d\lambda < 0, \quad (10.2.11)$$

называется областью нормальной дисперсии, а область, где

$$dn/d\omega < 0 \text{ или } dn/d\lambda > 0, \quad (10.2.12)$$

областью аномальной дисперсии. Аномальная дисперсия наблюдается, если частота электромагнитной волны попадает в окрестность резонансной частоты колебаний электрона. Как видно из сопоставления зависимостей $n(\omega)$ и $\kappa(\omega)$, в области аномальной дисперсии имеет место резонансное возрастание показателя поглощения. Это проявляется в резком возрастании поглощения излучения с $\omega \approx \omega_0$, то есть в появлении полос поглощения в спектрах.

10.2.3. Закон Бугера

Поясним смысл показателя поглощения. Запишем выражение для волнового числа:

$$k = \frac{\omega}{c} \sqrt{\epsilon} = \frac{\omega}{c} n + i \frac{\omega}{c} \kappa = k_r + i \frac{\alpha}{2}. \quad (10.2.13)$$

Величину α можно представить в виде

$$\alpha = \frac{4\pi}{\lambda_0} \kappa, \quad \lambda_0 = \frac{2\pi c}{\omega}, \quad (10.2.14)$$

где λ_0 — длина волны излучения в вакууме на частоте ω . С учётом (10.2.13) для бегущей волны получается выражение

$$E \sim e^{ikx} = e^{ik_r x} e^{-\alpha x/2}. \quad (10.2.15)$$

Отсюда следует, что интенсивность излучения ($I \sim \overline{E^2}$) меняется по закону

$$I = I_0 e^{-\alpha x}. \quad (10.2.16)$$

Это соотношение называется *законом Бугера* (в более общей формулировке данный закон обсуждается в разделе 11.4). Величина α называется *коэффициентом поглощения (затухания)*. Характерная длина, на которой происходит затухание, то есть глубина, на которую излучение проникает в среду, есть

$$l_a \sim 1/\alpha. \quad (10.2.17)$$

Обратим внимание на то, что в литературе имеются некоторые расхождения в терминологии. Во многих публикациях коэффициентом поглощения называют отношение потока, прошедшего через слой, к падающему: $\beta = \Phi_{\text{прош.}} / \Phi_{\text{пад.}}$, а показателем поглощения — величину α , введённую в (10.2.6). Столь же часто встречаются иные определения: величину $\kappa = \text{Im} \sqrt{\epsilon}$ называют показателем поглощения, величину β — поглощательной способностью слоя, а величину α — коэффициентом поглощения (затухания). Именно эта последняя терминология используется в данной книге.

Строго говоря, наличие затухания ($\kappa \neq 0$) не означает, что имеет место поглощение излучения. Например, если диэлектрическая проницаемость — действительная отрицательная величина: $\epsilon < 0$, то $\kappa = \sqrt{-\epsilon} \neq 0$. Затухание волны по мере распространения в среде в этом случае означает, что происходит отражение. Поглощение же имеет место только тогда, когда происходит передача энергии от электронов к веществу. А такое происходит, если сила трения $F_{\text{тр}} = -\eta v \neq 0$, то есть при $\eta \neq 0$. Но тогда $\text{Im} \epsilon \neq 0$. Именно это последнее условие указывает на наличие поглощения.

10.3. Формула Лоренца—Лоренца

При выводе выражения для диэлектрической проницаемости мы, считывая вектор поляризации среды, учитывали действие только внешнего поля на электроны. Вместе с тем каждый электрон сам производит поле, действующее на остальные. Это становится тем более существенным, чем выше концентрация частиц. Рассмотрим данный вопрос подробнее.

В макроскопической электродинамике вводится *среднее поле*, которое определяется усреднением микрополя по физически малому объёму:

$$\mathbf{E} = \frac{1}{\Delta V} \int_{\Delta V} \mathbf{E}_{\text{микро}} dV. \quad (10.3.1)$$

Это поле создаётся всеми зарядами среды. В то же время поле $\mathbf{E}'$, действующее на отдельный заряд, отличается от $\mathbf{E}$, поскольку из поля $\mathbf{E}$ мы должны исключить часть, производимую самим рассматриваемым зарядом. Различие между $\mathbf{E}$ и $\mathbf{E}'$ обусловлено дискретностью зарядов (наличием промежутков между ними) и зависит от их концентрации.

В соответствии со сказанным мы должны найти поле $\mathbf{E}'$, называемое *эффективным*, или *действующим*.

Выделим отдельную молекулу и рассмотрим сферу вокруг неё. Считаем, что внутри сферы находится только рассматриваемая молекула. Пусть во внешнем электрическом поле среда приобретает поляризацию, описываемую вектором $\mathbf{P}$. Тогда, считая среду сплошной, мы бы заключили, что поляризация среды внутри сферы создаёт в ней поле

$$\Delta \mathbf{E} = -\frac{4\pi}{3} \mathbf{P} \quad (10.3.2)$$

(как поле внутри однородно поляризованного шара). В этой же сфере действует также поле $\mathbf{E}'$, производимое всеми прочими зарядами. Сумма двух полей даёт среднее поле:

$$\mathbf{E} = \mathbf{E}' - \frac{4\pi}{3} \mathbf{P} \Rightarrow \mathbf{E}' = \mathbf{E} + \frac{4\pi}{3} \mathbf{P}. \quad (10.3.3)$$

Мы получили выражение для действующего поля, то есть того поля, которое и производит поляризацию молекулы.

Пусть β — поляризуемость молекулы. Тогда дипольный момент можно записать в виде

$$\mathbf{p} = \beta \mathbf{E}'. \quad (10.3.4)$$

Если концентрация молекул равна N , то вектор поляризации среды оказывается равным

$$\mathbf{P} = N\beta \mathbf{E}' = N\beta \left(\mathbf{E} + \frac{4\pi}{3} \mathbf{P} \right) \Rightarrow \mathbf{P} = \frac{N\beta}{1 - \frac{4\pi}{3} N\beta} \mathbf{E}. \quad (10.3.5)$$

Наконец, вводя диэлектрическую проницаемость среды равенствами $\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P} = \epsilon\mathbf{E}$, находим

$$\epsilon = \frac{1 + \frac{8\pi}{3} N\beta}{1 - \frac{4\pi}{3} N\beta}. \quad (10.3.6)$$

Полученная формула связывает диэлектрическую проницаемость среды с поляризуемостью отдельной молекулы и концентрацией молекул. Перепишем её в следующем виде:

$$\frac{4\pi}{3} N\beta = \frac{\epsilon - 1}{\epsilon + 2}. \quad (10.3.7)$$

Это равенство называется *формулой Лоренца–Лоренца*.

Воспользовавшись выражением (10.2.4) для поляризуемости молекулы β и имея в виду соотношение $\epsilon = n^2$, получаем уточнённую зависимость показателя преломления от частоты света ω и концентрации электронов N :

$$\frac{n^2 - 1}{n^2 + 2} = \frac{4\pi Ne^2}{3m} \frac{1}{\omega_0^2 - \omega^2} \quad (10.3.8)$$

(здесь мы пренебрегли потерями энергии электронами, положив формально $\gamma \rightarrow 0$).

Интересное совпадение: формулу (10.3.7) получили независимо и почти одновременно в 1880 г. двое учёных с почти одинаковыми фамилиями: датский физик Людвиг Лоренц (L. V. Lorentz) и голландский физик Хендрик Лоренц (H. A. Lorentz).

10.4. Дисперсия электромагнитных волн в плазме

10.4.1. Диэлектрическая проницаемость плазмы

Пусть на плазму падает электромагнитная волна. Плазма представляет собой квазинейтральную смесь электронов и ионов, причём масса иона примерно в 10^4 раз больше массы электрона. Поэтому в поле внешней электромагнитной волны двигаются главным образом электроны. Ионы же можно считать покоящимися (производимый ими ток оказывается малым по сравнению с током, производимым электронами). В типичных условиях скорость движения электронов мала ($v \ll c$), так что можно пренебречь силой Лоренца, возникающей со стороны магнитного поля волны. Наконец, поскольку амплитуда колебаний электрона мала по сравнению с длиной волны излучения, можно рассматривать движение электрона как движение в однородном электрическом поле волны:

$$\mathbf{E} = \mathbf{E}_0 \cos \omega t. \quad (10.4.1)$$

Будем пренебрегать потерями энергии электрона, считая, что его энергия не передаётся среде. Соответственно в стационарном процессе отсутствует и поглощение излучения средой. В принятых условиях движение отдельного электрона описывается уравнением

$$m\ddot{r} = eE = eE_0 \cos \omega t, \quad (10.4.2)$$

решение которого имеет вид

$$r = -\frac{eE_0}{m\omega^2} \cos \omega t = -\frac{eE}{m\omega^2}. \quad (10.4.3)$$

Это смещение приводит к появлению дипольного момента электрона:

$$p = er = -e^2 E / m\omega^2. \quad (10.4.4)$$

Если концентрация электронов в плазме равна N электронов, то вектор поляризации (дипольный момент единицы объёма) оказывается равным

$$P = Np = -\frac{Ne^2}{m\omega^2} E \equiv \alpha E.$$

Имея в виду, что $D = E + 4\pi P = \epsilon E$, заключаем, что диэлектрическая проницаемость плазмы равна

$$\epsilon = 1 + 4\pi\alpha = 1 - \frac{4\pi Ne^2}{m\omega^2}.$$

Поскольку $\sqrt{4\pi Ne^2/m} = \omega_p$ — плазменная частота, то окончательно получаем

$$\epsilon = 1 - \frac{\omega_p^2}{\omega^2}. \quad (10.4.5)$$

Таким образом, в отличие от случая диэлектриков в электростатическом поле диэлектрическая проницаемость плазмы оказывается меньше единицы: $\epsilon < 1$. Более того, она отрицательна при частотах $\omega < \omega_p$.

Заметим, что аналогичной формулой описывается и диэлектрическая проницаемость металлов, поскольку электроны проводимости в металле свободные и поэтому ведут себя подобно электронам в плазме.

10.4.2. Закон дисперсии волн. Фазовая и групповая скорости

Запишем закон дисперсии:

$$k^2 = \frac{\omega^2}{c^2} \epsilon(\omega). \quad (10.4.6)$$

Подстановка сюда выражения (10.4.5) даёт

$$\omega^2 = \omega_p^2 + k^2 c^2. \quad (10.4.7)$$

Полученный закон дисперсии показан на рис. 10.4.1.

Найдём фазовую и групповую скорости волны:

$$v_{\phi} = \frac{\omega}{k} = \sqrt{\left(\frac{\omega_p}{k}\right)^2 + c^2}, \quad v_{\text{гр}} = \frac{d\omega}{dk} = \frac{kc^2}{\sqrt{\omega_p^2 + k^2c^2}} = \frac{kc^2}{\omega}. \quad (10.4.8)$$



Рис. 10.4.1. Закон дисперсии волн в плазме

Между этими скоростями имеет место соотношение

$$v_{\phi} v_{\text{гр}} = c^2. \quad (10.4.9)$$

Как видно из (10.4.9), $v_{\phi} > c$. Соответственно оказывается $v_{\text{гр}} < c$. Последнее согласуется с тем фактом, что перенос энергии происходит со скоростями, не превышающими скорость света.

10.4.3. Отражение электромагнитных волн от плазмы

Как следует из уравнений Максвелла, для плоских волн

$$E = E_0 \exp(ikx - i\omega t) \quad (10.4.10)$$

частота ω и волновое число k связаны соотношением (10.4.7). Пусть на слой плазмы падает волна с частотой $\omega < \omega_p$ (рис. 10.4.2). В этом случае

$$k^2 = \frac{\omega^2 - \omega_p^2}{c^2} < 0. \quad (10.4.11)$$

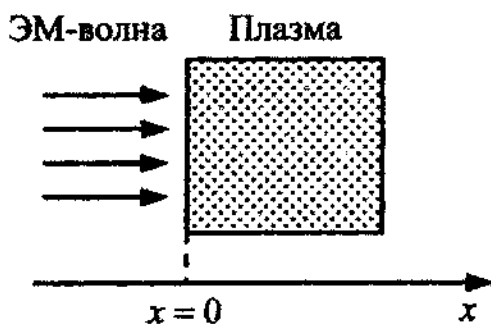


Рис. 10.4.2. Электромагнитная волна, падающая из вакуума на слой плазмы, занимающей область $x > 0$

Полагая $k = i\alpha$, перепишем выражение для волны (10.4.10) в следующем виде:

$$E = E_0 \exp(-\alpha x - i\omega t). \quad (10.4.12)$$

Отсюда видно, что амплитуда волны экспоненциально убывает — волна затухает по мере проникновения в плазму.

Таким образом, в плазме на большие расстояния могут распространяться только волны

с частотой $\omega > \omega_p$. Для волн же с частотой $\omega < \omega_p$ плазма играет роль зеркала. При этом, если в плазме энергия волны поглощается слабо, то почти вся волна отразится назад. Если же потери энергии играют существенную роль, то коэффициент отражения окажется меньше единицы.

Пусть электромагнитная волна падает на слой плазмы под некоторым углом к нормали θ . Тогда отражение может произойти и при частотах больших, чем ω_p . Действительно, при $\omega > \omega_p$ плазма представляет собой среду с показателем преломления $n = \sqrt{\epsilon} < 1$. Тогда при условии, что угол падения превышает угол полного внутреннего отражения:

$$\theta \geq \theta_{\text{в.о.}}, \sin \theta_{\text{в.о.}} = n \quad (10.4.13)$$

излучение полностью отражается, поскольку вакуум является средой оптически более плотной, чем плазма:

$$n_{\text{вак}} = 1 > n_{\text{плазма}} = \sqrt{\epsilon}. \quad (10.4.14)$$

В пределе $\omega \rightarrow \omega_p$, когда ещё $\omega < \omega_p$, угол полного внутреннего отражения $\theta_{\text{в.о.}}$ обращается в нуль, и плазма отражает волны при любых углах падения.

Пусть концентрация электронов медленно растёт в направлении оси x , как показано на рис. 10.4.3. Как правило, концентрация меняется незначительно на расстояниях порядка длины волны излучения. Поэтому можно приближённо определить локальное значение волнового числа формулой

$$k^2(x) = \frac{1}{c^2} \left[\omega^2 - \frac{4\pi e^2}{m} N(x) \right]. \quad (10.4.15)$$

Для заданной частоты волновое число обращается в нуль в точке, где концентрация электронов достигает критического значения

$$N_{\text{кр}} = m\omega^2 / 4\pi e^2. \quad (10.4.16)$$

Данное соотношение определяет положение «плазменного зеркала».

Эффект отражения волн от плазмы используется в дальней радиосвязи. Ионосфера (слой атмосферы, начинающийся на высоте примерно 50 км от поверхности Земли) представляет собой слой плазмы, формирующийся благодаря ионизирующему действию ультрафиолетовой и рентгеновской солнечной радиации, а также космического излучения. Хотя степень ионизации газа в ионосфере невысокая, концентрация электронов (до 10^3 – 10^5 электронов в 1 см^3) достаточно велика, чтобы оказалось возможным отражение длинных ($\lambda \sim$ несколько км) радиоволн. Это позволяет передавать радиосигналы в точки, находящиеся вне зоны прямой видимости (рис. 10.4.4).

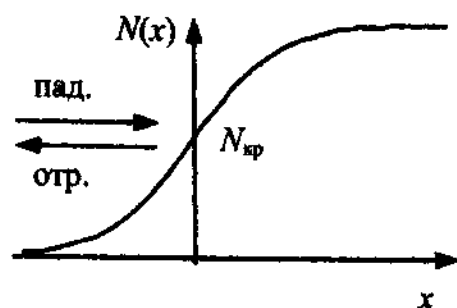


Рис. 10.4.3. Электромагнитная волна падает на слой плазмы с возрастающей концентрацией электронов

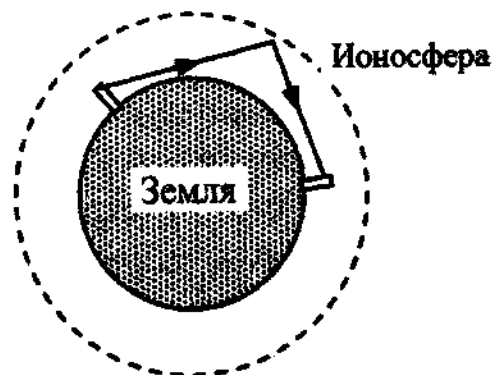


Рис. 10.4.4. Дальняя радиосвязь благодаря отражению радиоволн от ионосферы

ГЛАВА 11

РАССЕЯНИЕ СВЕТА

11.1. Молекулярное рассеяние

Рассеяние света — это преобразование света веществом, сопровождающееся изменением направления и проявляющееся как *вынужденное свечение*, обусловленное колебаниями электронов в атомах, молекулах и ионах.

И. Л. Мандельштам показал (1907), что в оптически однородной среде рассеяние может возникать, только если в среде присутствуют неоднородности. Иначе говоря, показатель преломления должен меняться от точки к точке. Дело в том, что вторичные волны, излучённые электронами, в однородной среде взаимно погашаются благодаря интерференции.

Рассеяние света в мутных средах с размерами неоднородностей $d \leq (0,1 - 0,2)\lambda$ называется *эффектом Тиндаля*.

Примерами мутных сред являются аэрозоли (дым, туман), эмульсии, стёкла с вкраплениями мелких частиц. Примеси приводят к тому, что показатель преломления среды различен в разных точках среды. Если мутная среда освещается белым светом, то при наблюдении светового пучка сбоку среда кажется голубоватой. В проходящем же свете среда выглядит красновато-оранжевой. Это означает, что вбок преимущественно рассеиваются коротковолновые спектральные компоненты света, в результате чего относительная доля длинноволновых компонент в проходящем пучке возрастает. Если размеры примесей оказываются сравнимыми с длиной волны излучения, то рассеяние становится *полихромным* — свет начинает рассеиваться почти равномерно во всём видимом диапазоне цветов от фиолетового до красного.

На явление свечения неоднородной среды при прохождении через неё света впервые обратил внимание в 1857 г. М. Фарадей (Англия). Дж. Тиндаль (Англия) подробно описал его в 1868 г., а Дж. Рэлей (1871) и Г. Ми (1908) дали ему теоретическое обоснование.

Рассеяние света наблюдается и в чистых средах без примесей. В этом случае оно называется *молекулярным рассеянием*. Возникновение неоднородностей в таких средах связано с различными флуктуациями (плотности, температура), приводящими к отклонениям показателя преломления от постоянного значения.

Если размеры неоднородностей малы по сравнению с длиной волны, то рассеяние является *упругим*: частоты исходного и рассеянного света оказываются одинаковыми.

Рассеяние света на неоднородностях, размеры которых много меньше его длины волны называется *рэлеевским*.

Объяснение данного явления следует из законов взаимодействия электромагнитных волн с электронами. Как известно, диполь с переменным дипольным моментом $p(t)$ излучает, причём мощность излучения даётся формулой

$$Q = \frac{2}{3c^3} \ddot{p}^2. \quad (11.1.1)$$

Пусть диэлектрическая проницаемость среды есть ϵ_0 , а диэлектрическая проницаемость неоднородного включения — ϵ . Если размер неоднородности a мал по сравнению с длиной волны излучения, то электрическое поле в области включения можно считать однородным. Будем считать, что эта область — шар радиуса a . В курсе электричества доказывалось, что её индуцированный дипольный момент равен

$$p = \beta E, \quad \beta = a^3 \frac{\epsilon - \epsilon_0}{\epsilon + 2\epsilon_0} \epsilon_0, \quad (11.1.2)$$

где E — электрическое поле вне шара. Пусть поле E совершает гармонические колебания: $E = E_0 \cos(\omega t + \varphi)$. Тогда

$$p = \beta E_0 \cos(\omega t + \varphi). \quad (11.1.3)$$

Подстановка этого выражения в (11.1.1) даёт после усреднения по периоду колебаний

$$Q = \frac{e^4 \beta^2 E_0^2}{3c^3} \omega^4. \quad (11.1.4)$$

Учитывая равенство $\omega = 2\pi c/\lambda$ (λ — длина волны в вакууме), получаем закон Рэлея:

$$Q = \frac{(2\pi e)^4 c \beta^2 E_0^2}{3\lambda^4} \sim \frac{1}{\lambda^4}. \quad (11.1.5)$$

Заметим, что данная зависимость выполняется, когда частоты света находятся вдали от резонансных частот вещества: $\omega \ll \omega_0$, и зависимость диэлектрической проницаемости от длины волны излучения слабая. Это обычно имеет место при рассеянии света в прозрачных средах (в частности, в воздухе).

Для неоднородностей с размерами $a > 0,2\lambda$ рэлеевское рассеяние переходит в дифракцию, причём зависимость (11.1.5) оказывается неспра-

ведливой. Если же размеры неоднородностей велики: $a \gg \lambda$, то для расчёта рассеяния применимы законы геометрической оптики, причём рассеяние приобретает полихромный характер.

Молекулярное рассеяние объясняет голубой цвет неба, поскольку именно коротковолновая компонента света Солнца сильнее отклоняется от исходного направления излучения. С этим же связано и то, что цвет Солнца на восходе и закате приобретает красновато-оранжевый оттенок (поскольку свет проходит бóльший путь, чем в случае, когда Солнце находится в зените). Поскольку вероятность флуктуаций растёт по мере увеличения температуры, то более насыщенный голубой цвет неба характерен для летнего времени по сравнению с зимним.

Первое детальное рассмотрение рассеяния света на диэлектрических сферических частицах произвольного размера было выполнено датским физиком Людвигом Лоренцем и приведено в статье, опубликованной в 1890 г. Эта работа в то время не привлекла внимание. В 1908 г. немецкий физик Густав Ми независимо выполнил детальное решение уравнений Максвелла и получил аналогичные результаты. В связи с этим теорию рассеяния света малыми частицами называют «теорией Лоренца–Ми», хотя часто используется и название «теория Ми». Само явление называют *рассеянием Ми*. Обычно этим термином обозначают случай, когда размеры частицы сопоставимы с длиной волны. Область размеров $R < 0.1\lambda$ относят к рэлеевскому рассеянию.

Подробнее с теорией рассеяния света можно познакомиться, например, по книгам: Борн К., Хафмен Д., *Поглощение и рассеяние света малыми частицами*/ пер. с англ. — М.: Мир, 1986; Борн М. Вольф Э. *Основы оптики*. Изд. 2-е/ пер. с англ. — М.: Наука, 1973.

11.2. Сечение рассеяния

11.2.1. Излучение колеблющегося диполя

Рассмотрим электромагнитную волну, создаваемую точечным диполем с периодически изменяющимся дипольным моментом в среде с $\epsilon = \mu = 1$. В волновой зоне напряжённости и потенциалы полей зависят от расстояния до диполя и времени по закону $\sim f(r - ct)/r$, а векторы напряжённости электрического и магнитного полей связаны соотношением

$$\mathbf{E} = \mathbf{H} \times \mathbf{n}, \quad (11.2.1)$$

где $\mathbf{n} = \mathbf{r}/r$ — единичный вектор в направлении распространения волны (в направлении радиус-вектора точки наблюдения относительно диполя). Картина полей $\mathbf{E}$ и $\mathbf{H}$ проиллюстрирована на рис. 11.2.1.

Пусть дипольный момент совершает колебания вдоль фиксированного направления, а его величина

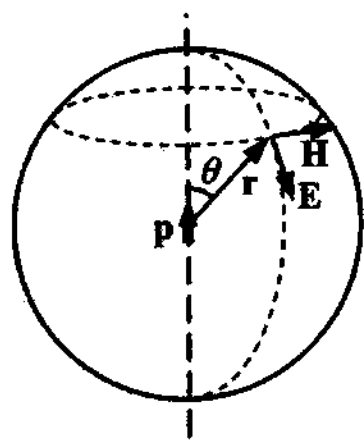


Рис. 11.2.1. Направление силовых линий электрического $\mathbf{E}$ и магнитного $\mathbf{H}$ полей излучения диполя $\mathbf{p}$ в волновой зоне

меняется по некоторому закону $p = p(t)$. Тогда электрическое и магнитное поля излучения в волновой зоне даются формулами (рис. 11.2.1)

$$E = \frac{\ddot{p}}{c^2 r} \sin \theta, \quad H = \frac{\ddot{p}}{c^2 r} \sin \theta. \quad (11.2.2)$$

Плотность потока энергии поля определяется вектором Пойнтинга:

$$S = \frac{c}{4\pi} E \times H. \quad (11.2.3)$$

Отсюда после усреднения по периоду колебаний находим интенсивность излучения:

$$I = \bar{S} = \frac{\overline{\ddot{p}^2}}{4\pi c^3 r^2} \sin^2 \theta. \quad (11.2.4)$$

Зависимость (11.2.4) даёт угловое распределение (диаграмму направленности) дипольного излучения (рис. 11.2.2). Видно, что максимум излучения идёт в направлении, перпендикулярном дипольному моменту p .

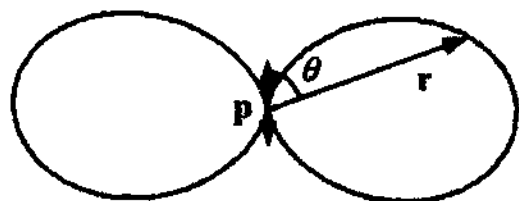


Рис. 11.2.2. Диаграмма направленности излучения точечного диполя в волновой зоне

Полная энергия, излучаемая диполем в единицу времени (мощность излучения), получается интегрированием вектора (11.2.4) по поверхности Π , окружающей диполь:

$$Q = \oint_{\Pi} S d\Pi = \frac{\ddot{p}^2}{4\pi c^3} \int_0^\pi \sin^2 \theta \cdot 2\pi \sin \theta d\theta = \frac{2\ddot{p}^2}{3c^3}. \quad (11.2.5)$$

При вычислении интеграла учтено, что вектор элементарной площадки равен

$$d\Pi = n d\Pi, \quad d\Pi = r^2 d\Omega = 2\pi r^2 \sin \theta d\theta. \quad (11.2.6)$$

11.2.2. Угловое распределение рассеянного излучения

Энергия, излучаемая диполем, зависит от направления. Обозначим как $d\Phi$ мощность, излучаемую в направлении n в телесный угол $d\Omega$. Поскольку площадь поверхности, высекаемая этим телесным углом на расстоянии r , равна $d\Pi = r^2 d\Omega$, то угловое распределение энергии излучения даётся формулой

$$d\Phi = I d\Pi = I r^2 d\Omega = \frac{\overline{\ddot{p}^2}}{4\pi c^3} \sin^2 \theta d\Omega. \quad (11.2.7)$$

Заметим, что согласно терминологии фотометрии, приведённой в разделе 2.14.2, величина $J = d\Phi/d\Omega$ есть энергетическая сила света.

Рассмотрим свободный электрон, находящийся в поле плоской линейно поляризованной волны с частотой ω . Считая амплитуду колебаний малой по сравнению с длиной волны, из уравнения движения

$$m\ddot{\mathbf{r}} = e\mathbf{E}, \quad \mathbf{E} = \mathbf{E}_0 \cos \omega t \quad (11.2.8)$$

находим дипольный момент:

$$\mathbf{p} = e\mathbf{r} = -\frac{e^2 \mathbf{E}}{m\omega^2}. \quad (11.2.9)$$

Подстановка последнего выражения в (11.2.7) даёт

$$d\Phi = \frac{e^4 E_0^2}{8\pi m^2 c^3} \sin^2 \theta d\Omega. \quad (11.2.10)$$

Выше мы предполагали, что падающая на электрон электромагнитная волна линейно поляризована. Пусть теперь волна неполяризована. Найдём угловое распределение излучения диполя для этого случая.

Входящий в формулу (11.2.10) множитель $\sin^2 \theta$ можно записать в векторной форме как

$$\sin^2 \theta = (\mathbf{n} \times \mathbf{e})^2,$$

где $\mathbf{n}$ — единичный вектор в направлении наблюдения, а $\mathbf{e}$ — единичный вектор амплитуды поля падающей на электрон волны. Требуется усреднить эту величину по всем допустимым направлениям вектора $\mathbf{e}$. Учтём, что в силу поперечности падающей волны вектор $\mathbf{e}$ направлен перпендикулярно волновому вектору $\mathbf{k}$.

Выберем оси координат следующим образом (рис. 11.2.3): направим ось z вдоль вектора $\mathbf{k}$ и для рассматриваемого направления $\mathbf{n}$ положим

$$n_x = \sin \Theta, \quad n_z = \cos \Theta,$$

где Θ — угол между векторами $\mathbf{n}$ и $\mathbf{k}$. Тогда вектор $\mathbf{e}$ будет иметь компоненты

$$e_x = \cos \varphi, \quad e_y = \sin \varphi, \quad e_z = 0.$$

Воспользуемся тождеством $(\mathbf{n} \times \mathbf{e})^2 = 1 - (\mathbf{n} \cdot \mathbf{e})^2$. Тогда в выбранной системе координат имеем

$$(\mathbf{n} \times \mathbf{e})^2 = 1 - (n_x e_x)^2 = 1 - \sin^2 \Theta \cos^2 \varphi.$$

Учтём теперь, что вектор $\mathbf{e}$ с равной вероятностью может иметь в плоскости (x, y) любое направление. Поэтому усреднение по углу φ даёт

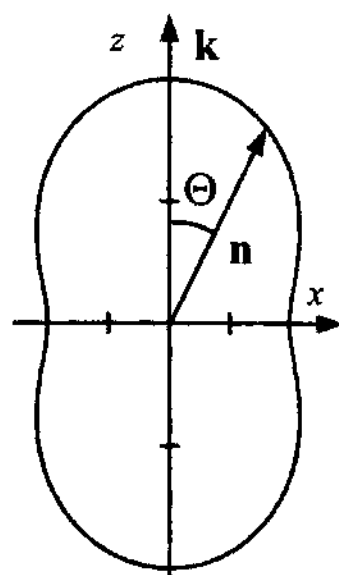


Рис. 11.2.3. Диаграмма направленности излучения электрона в поле неполяризованной электромагнитной волны

$$\overline{(\mathbf{n} \times \mathbf{e})^2} = \frac{1}{2\pi} \int_0^{2\pi} [1 - \sin^2 \Theta \cos^2 \varphi] d\varphi = 1 - \frac{1}{2} \sin^2 \Theta = \frac{1 + \cos^2 \Theta}{2}.$$

Таким образом, получаем вместо (11.2.10) следующее равенство:

$$d\Phi = \frac{e^4 E_0^2}{8\pi m^2 c^3} \frac{1 + \cos^2 \Theta}{2} d\Omega. \quad (11.2.11)$$

Диаграмма направленности излучения для рассмотренного случая показана на рис. 11.2.3.

11.2.3. Сечение рассеяния

Интенсивность падающей на заряд плоской волны равна

$$I_0 = cE_0^2/8\pi. \quad (11.2.12)$$

Введём дифференциальное сечение рассеяния волны зарядом соотношением

$$d\sigma = d\Phi/I_0. \quad (11.2.13)$$

Подстановка сюда выражений (11.2.10), (11.2.12) даёт

$$d\sigma = \left(\frac{e^2}{mc^2} \right)^2 \sin^2 \theta d\Omega. \quad (11.2.14)$$

Входящая в эту формулу величина

$$r_0 = e^2/mc^2 \quad (11.2.15)$$

называется *классическим радиусом* заряженной частицы. Для электрона она равна

$$r_0 = 2,8179 \cdot 10^{-13} \text{ см}. \quad (11.2.16)$$

Полное сечение рассеяния получается в результате интегрирования (11.2.14) по всем углам:

$$\sigma = \int d\sigma = r_0^2 \int_0^\pi \sin^2 \theta \cdot 2\pi \sin \theta d\theta = \frac{8\pi}{3} r_0^2. \quad (11.2.17)$$

Данное соотношение называется *формулой Томсона*. Заметим, что полное сечение σ не зависит от поляризации излучения, в чём можно убедиться, воспользовавшись формулой (11.2.11) вместо (11.2.10).

Результат (11.2.17) нетрудно обобщить на случай связанного электрона. Для этого достаточно в (11.2.7) использовать следующее выражение для дипольного момента электрона в поле плоской волны:

$$\mathbf{p} = \frac{e^2/m}{\omega_0^2 - \omega^2 - 2i\gamma\omega} \mathbf{E}. \quad (11.2.18)$$

Отсюда находим

$$d\sigma = r_0^2 \frac{\omega^4}{(\omega_0^2 - \omega^2)^2 + 4\gamma^2 \omega^2} \sin^2 \theta d\theta, \quad (11.2.19)$$

$$\sigma = \frac{8\pi}{3} r_0^2 \frac{\omega^4}{(\omega_0^2 - \omega^2)^2 + 4\gamma^2 \omega^2}.$$

Из полученной формулы видно, что сечение рассеяния резонансно возрастает, когда частота электромагнитной волны приближается к собственной частоте колебаний электрона, то есть при приближении к полосам поглощения вещества. В частности, при $\omega \rightarrow \omega_0$ из (11.2.19) следует

$$\sigma = \frac{8\pi}{3} r_0^2 \frac{\omega_0^2}{4\gamma^2}. \quad (11.2.20)$$

На таких частотах сечение может значительно превысить величину (11.2.17). Это явление называется *резонансной флуоресценцией* и означает, что происходит интенсивная перекачка энергии падающей на электрон волны в энергию рассеянного излучения.

11.3. Поглощение света малыми частицами

Оптическое излучение проникает в вещество на глубину $\sim \Lambda = \lambda/2\pi\kappa$, где показатель поглощения κ определяется равенством $\kappa = \text{Im} \sqrt{\epsilon}$. Величина Λ велика для многих диэлектриков, а в случае металлов оказывается порядка 10^2 нм. Это значит, что частицы с размерами 100 нм и меньше оказываются почти прозрачными для излучения — оно проходит через них лишь с малым ослаблением.

Оценим величину поглощаемой частицей энергии.

Поскольку электромагнитная волна проникает в вещество на глубину $\sim \Lambda = \lambda/2\pi\kappa$, то при диаметре частицы $2R$ поглощается доля $\sim 2R/\Lambda \approx 4\pi\kappa R/\lambda$ энергии, вошедшей в эту частицу. Если падающая на частицу мощность составляет $P_{\text{пад}}$, то поглощённая мощность оказывается равной

$$P_{\text{погл}} \sim (4\pi\kappa R/\lambda) P_{\text{пад}}. \quad (11.3.1)$$

Последовательный расчёт приводит к следующему выражению для *поглощательной способности* частицы:

$$A = \frac{P_{\text{погл}}}{P_{\text{пад}}} = \frac{8\pi R}{\lambda} \text{Im} \left(\frac{\epsilon - \epsilon_0}{\epsilon + 2\epsilon_0} \right), \quad (11.3.2)$$

где ϵ и ϵ_0 — диэлектрические проницаемости соответственно частицы и окружающей среды.

Перейдём от мощности излучения $P_{\text{пад}}$ к интенсивности I . Полагая $P_{\text{пад}} = \pi R^2 I$, заключаем, что поглощаемая частицей энергия составляет

$$P_{\text{погл}} = AP_{\text{пад}} = A_I I, \quad A_I = \frac{8\pi^2 R^3}{\lambda} \text{Im} \left(\frac{\varepsilon - \varepsilon_0}{\varepsilon + 2\varepsilon_0} \right). \quad (11.3.3)$$

11.4. Экстинкция. Закон Бугера–Ламберта–Бера

Экстинкцией называется ослабление пучка света при его распространении в веществе за счёт *поглощения* и *рассеяния*.

Пусть на среду падает пучок излучения с интенсивностью I_0 . Найдём закон $I(x)$, по которому интенсивность прямого потока убывает по мере прохождения среды.

Рассмотрим сначала ослабление потока вследствие рассеяния. Выделим слой вещества толщиной dx и площадью (перпендикулярно потоку) S . Один электрон среды за время dt приведёт к потере энергии волной:

$$\delta W = \sigma I dt. \quad (11.4.1)$$

Если концентрация электронов равна N , то в рассматриваемом слое имеется $NSdx$ электронов. Поэтому суммарная потеря энергии составит

$$(dW)_{\text{пот}} = NSdx \cdot \delta W = NSdx \cdot \sigma I dt. \quad (11.4.2)$$

В стационарном процессе количество энергии в рассматриваемом объёме не меняется со временем. Поскольку убыль энергии потока можно записать в виде $-d(SI)dt$, то из условия

$$-d(SI)dt = (dW)_{\text{пот}} \quad (11.4.3)$$

получаем уравнение

$$\left(\frac{dI}{dx} \right)_{\text{расс}} = -N\sigma I, \quad (11.4.4)$$

откуда следует

$$I = I_0 e^{-\alpha_s x}, \quad \alpha_s = N\sigma. \quad (11.4.5)$$

Величина α_s называется *коэффициентом (показателем) рассеяния* (индекс «s» — от слова *scattering*, рассеяние).

Помимо рассеяния свет в среде испытывает поглощение. Действительно, если диэлектрическая проницаемость имеет мнимую часть, связанную с потерями энергии электроном:

$$\varepsilon = \varepsilon_1 + i\varepsilon_2, \quad \sqrt{\varepsilon} = n + ik, \quad (11.4.6)$$

то волновое число оказывается комплексным:

$$k = \frac{\omega}{c} \sqrt{\varepsilon} = k_1 + ik_2, \quad k_1 = \frac{2\pi n}{\lambda}, \quad k_2 = \frac{2\pi \kappa}{\lambda}. \quad (11.4.7)$$

Благодаря поглощению ослабление потока излучения в среде происходит по закону (раздел 10.2.3)

$$\left(\frac{dI}{dx}\right)_{\text{погл}} = -\alpha_a I, \quad \alpha_a = 2k_2 = \frac{4\pi\kappa}{\lambda}. \quad (11.4.8)$$

Коэффициент поглощения¹ α_a введён ранее в разделе 10.2.3, $[\alpha_a] = \text{см}^{-1}$ (индекс «a» — от англ. *absorption*, поглощение).

Имея в виду два механизма ослабления потока, запишем уравнение

$$\frac{dI}{dx} = \left(\frac{dI}{dx}\right)_{\text{расс}} + \left(\frac{dI}{dx}\right)_{\text{погл}} = -(\alpha_s + \alpha_a)I. \quad (11.4.9)$$

Отсюда находим

$$I = I_0 e^{-\beta x}, \quad \beta = \alpha_s + \alpha_a. \quad (11.4.10)$$

Величина β называется коэффициентом экстинкции (от лат. *exstinctio* — гашение), или коэффициентом ослабления, а экспоненциальный закон (11.4.10) ослабления интенсивности — законом Бугера–Ламберта–Бера.

Зависимость (11.4.10) экспериментально установлена французским учёным П. Бугером в 1729 г., подробно рассмотрена немецким учёным И. Г. Ламбертом в 1760 г. Зависимость показателя экстинкции от концентрации рассеивающих молекул исследована экспериментально немецким учёным А. Бером в 1852 г.

11.5. Поляризация излучения при рассеянии

Рассмотрим распространение пучка света в прозрачной среде. Поскольку электромагнитная волна поперечная, то вектор напряжённости электрического поля E_0 перпендикулярен направлению распространения

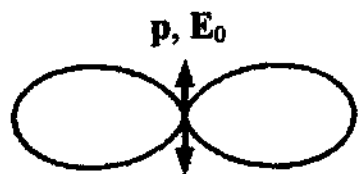


Рис. 11.5.1. Диаграмма направленности излучения диполя

волны (задаваемому волновым вектором k). Эта волна вызывает колебания электронов вдоль вектора E_0 . Рис. 11.5.1 позволяет представить возможные направления рассеяния.

Пусть волна линейно поляризована, то есть вектор E_0 совершает колебания вдоль одной прямой. Тогда поле рассеянной волны будет поляризованным (рис. 11.2.1 и 11.5.2, слева).

Если свет неполяризованный, то вектор E_0 может иметь любые направления, оставаясь перпендикулярным направлению распространения волны (волновому вектору). Однако при наблюдении рассеянного света сбоку проекция вектора E рассеянной волны вертикальна, что соответствует линейной поляризации. Отсюда следует, что сбоку от светового пуч-

¹ О терминологии см. замечание в разделе 10.2.3.

ка рассеянная волна имеет линейную поляризацию (рис. 11.5.2, справа). В случае, когда направление наблюдения образует с осью светового пучка угол, отличный от $\pi/2$, поляризация рассеянного излучения будет лишь частичной. Если исходное излучение имеет круговую поляризацию, то при наблюдении под углом, образующим с осью угол, больший, чем $\pi/2$, и поляризация рассеянного излучения оказывается эллиптической.

Пусть на мутную жидкость падает линейно поляризованная волна (рис. 11.5.3). Исходно неполяризованный пучок можно сделать поляризованным, пропустив его предварительно через поляризатор.

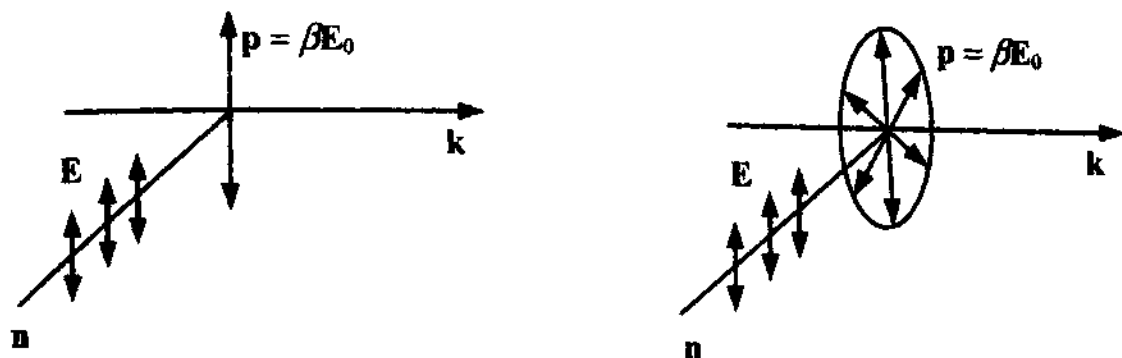


Рис. 11.5.2. Слева — наблюдаемое сбоку (в направлении n) электрическое поле рассеянной волны E , если падающая на среду волна линейно поляризована (вектор E_0). Справа — то же, если падающая волна неполяризованная: возможны различные направления вектора E_0 , перпендикулярные волновому вектору k

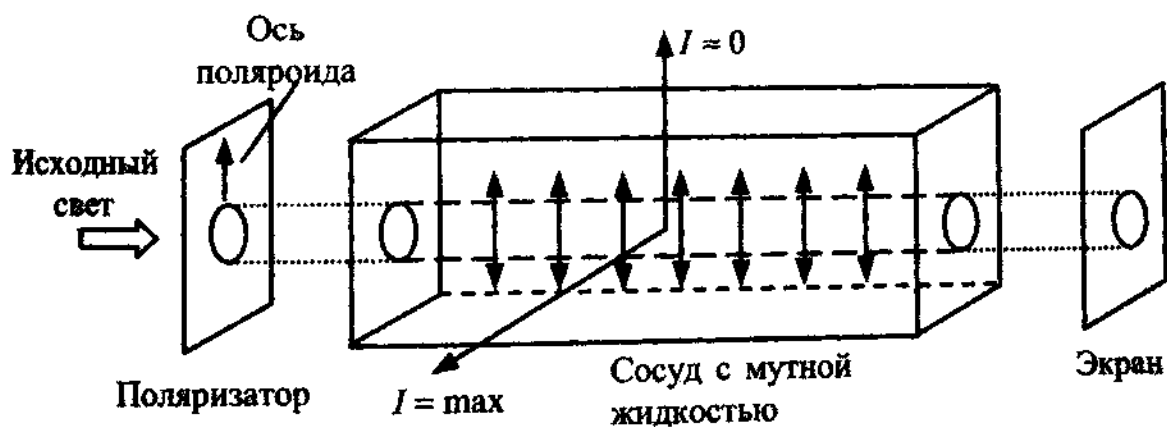


Рис. 11.5.3. Наблюдение рассеяния линейно поляризованного света в мутной среде

Дипольное излучение имеет диаграмму направленности, показанную на рис. 11.5.1: максимальная интенсивность достигается в направлении, перпендикулярном дипольному моменту p , тогда как в направлении вектора p излучение практически отсутствует. Это значит, что в экспериментальной установке на рис. 11.5.3 пучок виден сбоку (если ось поляризатора вертикальна) и остаётся практически невидимым при наблюдении кюветы снизу или сверху ($I \approx 0$).

Наблюдая кювету сбоку и поворачивая поляризатор вокруг направления исходного пучка света, мы будем видеть то усиление, то ослабление рассеянного излучения. Заметим, что на экране прошедшее излучение сохраняет ту поляризацию, которую оно приобрело после прохождения поляризатора.

11.6. Комбинационное рассеяние

11.6.1. Упругое и неупругое рассеяние света

Рассматривая выше рассеяние, мы считали, что частота света не меняется. Соответственно говорят, что имеет место *упругое рассеяние*. В самом деле, электрон под действием электромагнитной волны совершает вынужденные колебания с частотой волны. В соответствии с законами электродинамики на той же частоте происходит и излучение.

Комбинационное рассеяние — это *неупругое рассеяние* света, при котором в спектральном составе помимо частоты падающей волны ω присутствуют линии, отличающиеся от неё на величину, равную или кратную частоте внутримолекулярных колебаний. Такие новые линии называют *спутниками* (или *сателлитами*).

Данное явление называют также *эффектом Рамана*, или *рамановским рассеянием*, в честь индийского физика Чандрасекхара Венката Рамана.

Поскольку частоты спутников однозначно связаны с собственными частотами колебаний молекул, то использование комбинационного рассеяния позволило разработать *рамановскую спектроскопию* — эффективный метод изучения состава и строения вещества.

В 1921–1928 годах В. Раман с сотрудниками выполнил систематические исследования особенностей поляризации рассеянного света в жидкостях. Было установлено наличие в рассеянном свете дополнительного свечения с длиной волны, заметно отличающейся от длины волны падающего излучения. Это явление наблюдалось более чем в 100 прозрачных жидкостях, газах и твердых телах. В 1928 г. Г. С. Ландсберг и Л. И. Мандельштам (СССР) обнаружили эффект комбинационного рассеяния света в кристаллах. За работы по рассеянию света и открытие комбинационного рассеяния Раману была присуждена Нобелевская премия по физике (1930).

11.6.2. Закономерности комбинационного рассеяния

При исследовании комбинационного рассеяния были установлены следующие закономерности.

1. Спектральные линии-спутники сопровождают каждую линию первичного света.
2. Сдвиг спутников по частоте относительно первичной линии характеризует рассеивающее вещество и равен собственным частотам молекулярных колебаний.
3. Спутники образуют две группы линий, расположенных симметрично относительно возбуждающей линии. Спутники, смещённые в длинноволновую сторону относительно первоначальной линии, называются *сто-*

ксовыми (по аналогии с явлением люминесценции), а смещённые в коротковолновую — *антистоксовыми*. Интенсивность стоксовых спутников значительно выше.

4. С увеличением температуры интенсивность антистоксовых спутников возрастает.

Название спутников связано с именем английского математика и физика Джорджа Габриеля Стокса, установившего в 1852 г., что длина волны фотолюминесценции больше длины волны возбуждающего света (*правило Стокса*).

Люминесценция (от лат. *lumen* — свет и *escent* — слабое действие) — нетепловое свечение вещества, происходящее после поглощения им энергии возбуждения. Впервые люминесценция была описана в XVIII веке. Физическая природа люминесценции состоит в излучательных переходах электронов в атомах или молекулах из возбуждённого состояния в основное. Причиной первоначального их возбуждения могут служить различные факторы: внешнее излучение, температура, химические реакции и др. Важной особенностью люминесценции является то, что она способна проявляться при значительно более низких температурах, так как не использует тепловую энергию излучающей системы.

11.6.3. Классическое описание

Комбинационное рассеяние допускает качественное описание на языке классической физики.

Рассмотрим действие поля волны на электрон в молекуле. Дипольный момент, индуцируемый электромагнитной волной, равен

$$\mathbf{p} = \beta \mathbf{E}, \quad (11.6.1)$$

где β — поляризуемость молекулы. В электромагнитной волне напряжённость электрического поля совершает колебания:

$$E = E_0 \cos \omega_0 t, \quad (11.6.2)$$

что приводит к осцилляциям дипольного момента:

$$\mathbf{p} = \beta E_0 \cos \omega_0 t. \quad (11.6.3)$$

Если состояние молекулы не меняется, то поляризуемость β можно считать неизменной, в результате чего частота света не меняется при рассеянии: излучение происходит на частоте колебаний диполя. Это есть *рэлеевское рассеяние*, рассмотренное в предыдущих разделах.

В действительности поляризуемость β зависит от состояния молекулы. Если молекула колеблется с собственной частотой ω_1 , то смещение ядер $q(t)$ можно записать как

$$q = q_0 \cos(\omega_1 t). \quad (11.6.4)$$

где q_0 — амплитуда колебаний. При малых колебаниях поляризуемость β можно разложить в ряд по степеням q :

$$\beta = \beta_0 + \gamma q, \quad \gamma = (\partial \beta / \partial q)_0. \quad (11.6.5)$$

Подстановка (11.6.4), (11.6.5) в (11.6.3) даёт

$$\mathbf{p} = E_0 (\beta_0 + \gamma \cos \omega_1 t) \cos \omega_0 t. \quad (11.6.6)$$

Эту величину можно представить в виде суммы трёх слагаемых:

$$\mathbf{p} = \mathbf{p}_- + \mathbf{p}_0 + \mathbf{p}_+, \quad (11.6.7)$$

где использованы обозначения

$$\mathbf{p}_0 = \beta_0 \mathbf{E}_0 \cos \omega_0 t, \quad \mathbf{p}_{\pm} = \frac{1}{2} \gamma \mathbf{E}_0 \cos[(\omega_0 \pm \omega_1)t]. \quad (11.6.8)$$

Таким образом, в спектре колебаний диполя, помимо частоты исходной волны (ω_0), присутствуют спутники:

$$\omega_- = \omega_0 - \omega_1, \quad \omega_+ = \omega_0 + \omega_1. \quad (11.6.9)$$

Колебания $\mathbf{p}_0$ (с частотой ω_0) приводят к рэлеевскому рассеянию, тогда как колебания $\mathbf{p}_-$ и $\mathbf{p}_+$ (с частотами (11.6.9)) — к комбинационному.

1.6.4. Квантовомеханическое описание

Более наглядное объяснение комбинационного рассеяния даёт квантовомеханический подход.

Согласно квантовой механике свет с частотой ω можно представить как поток фотонов — квантов света — с энергией $\hbar\omega_0$. Молекула рассматривается как осциллятор с собственной частотой ω_1 . Изменение энергии молекулы может происходить только дискретно, порциями $\hbar\omega_1$. Это соответствует переходу между «разрешёнными» состояниями (рис. 11.6.1). С этой точки зрения комбинационное рассеяние объясняется следующим образом.

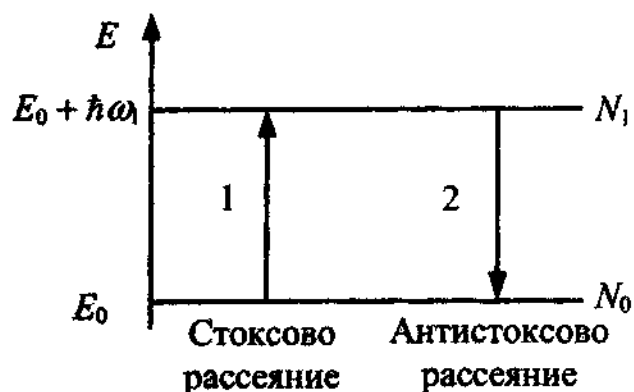


Рис. 11.6.1. Переходы между колебательными уровнями молекулы: 1 — молекула поглощает энергию $\hbar\omega_1$ (возбуждение), 2 — молекула отдаёт энергию $\hbar\omega_1$ (релаксация)

Если молекула находится в основном состоянии (с энергией E_0 , рис. 11.6.1), то в результате воздействия на неё света она может перейти в возбуждённое состояние с энергией $E_0 + \hbar\omega_1$ (стрелка 1 на рис. 11.6.1). Соответственно энергия фотона уменьшится на $\hbar\omega_1$ и составит $\hbar(\omega_0 - \omega_1)$. Но это означает, что частота рассеянного света уменьшится: $\omega_- = \omega_0 - \omega_1$ — возникает стоксова компонента. Аналогично, если молекула в момент воздействия на неё фотона находилась в возбуждённом состоянии, то она может передать фотону свою колебательную энергию $\hbar\omega_1$ и перейти в нижнее энергетическое состояние (стрелка 2 на

рис. 11.6.1) — формируется антистоксова линия. Частота фотона в этом процессе увеличится: $\omega_+ = \omega_0 + \omega_1$. Так возникают комбинационные частоты в рассеянном свете. Что касается рэлеевского рассеяния, то оно отвечает упругому взаимодействию фотона с молекулой, то есть взаимодействию, при котором состояние молекулы не меняется.

Квантовомеханический подход позволяет объяснить рост интенсивности антистоксовой линии при возрастании температуры. Действительно, число молекул, находящихся в основном и возбуждённом состояниях даётся распределением Больцмана (рис. 11.6.1):

$$N_0 = A \exp[-E_0/kT], \quad N_1 = A \exp[-(E_0 + \hbar\omega_1)/kT]. \quad (11.6.10)$$

Соответственно с ростом температуры возрастает число переходов с верхнего уровня на нижний (пропорционально N_1).

Заметим также, что интенсивность антистоксовой линии ($\omega_0 + \omega_1$) меньше, чем интенсивность стоксовой ($\omega_0 - \omega_1$). Дело в том, что число молекул на верхнем уровне (N_1) меньше, чем на нижнем (N_0):

$$N_1/N_0 = \exp(-\hbar\omega_1/kT) < 1.$$

В заключение этого раздела отметим, что если интенсивность света достаточно велика, то велика и интенсивность линий-спутников. В свою очередь, эти волны также могут действовать на молекулы, приводя к появлению в рассеянном свете компонент с частотами $\omega_0 + n\omega_1$, $n = 0, \pm 1, \pm 2, \dots$

ГЛАВА 12

ЭЛЕМЕНТЫ КРИСТАЛЛООПТИКИ

12.1. Двойное лучепреломление

12.1.1. Обыкновенный и необыкновенный лучи

Рассмотрим особенности распространения электромагнитных волн в анизотропных средах, то есть веществах, обладающих выделенными направлениями.

При падении света на некоторые кристаллы (такие, как слюда, гипс, исландский шпат) луч расщепляется на два луча. Один из них, следующий обычным законам преломления, называется *обыкновенным* (обозначается символом «*o*» = *ordinary*), а второй — *необыкновенным* (обозначается символом «*e*» = *extraordinary*) (рис. 12.1.1).

Описанное явление было обнаружено Расмусом Бартолином (Дания) в 1669 г. на кристалле исландского шпата (прозрачной разновидности кальцита CaCO_3). Нидерландский физик Х. Гюйгенс (около 1690 г.) связал явление с тем, что обыкновенный и необыкновенный лучи имеют разную поляризацию, и ввёл понятие оптической оси кристалла.

Явление расщепления одного луча на два, различающихся по свойствам, называется *двойным лучепреломлением*.

В кристаллах существуют оси, вдоль которых обыкновенный и необыкновенный лучи распространяются вместе, не разделяясь и с одинаковой скоростью. В общем случае таких осей две. Соответствующие направления называются *оптическими осями кристалла*.

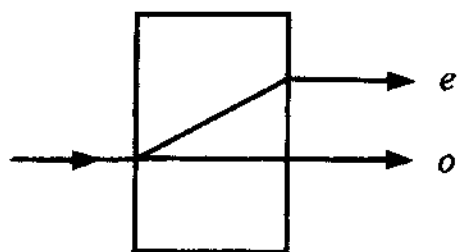


Рис. 12.1.1. Двойное лучепреломление: расщепление луча на обыкновенный (*o*) и необыкновенный (*e*) лучи

В некоторых кристаллах оптические оси совпадают, так что фактически он обладает единственной оптической осью. В этом случае кристалл называется *одноосным*.

В кристаллах с несовпадающими оптическими осями обыкновенного луча нет, все лучи являются необыкновенными.

Для обыкновенного луча показатель преломления не зависит от направления распространения, а для необыкновенного — зависит.

Для сведения приведём значения показателя преломления для двух материалов при $\lambda = 589 \text{ нм}$:

кварц: $n_o = 1,54425$, $n_e = 1,55335$;

кальцит: $n_o = 1,65835$, $n_e = 1,48640$.

Плоскость, проходящая через оптическую ось, называется *главной плоскостью* (или *главным сечением*) кристалла. Обычно выбирают главную плоскость так, чтобы она содержала не только оптическую ось, но и волновой вектор $\mathbf{k}$.

На рис. 12.1.2 показана диаграмма показателей преломления для обыкновенного («о») и необыкновенного («е») лучей в одноосном кристалле (соответствующие формулы будут получены в разделе 12.1.3). На этой диаграмме направление луча задаётся вектором $\mathbf{k}$. Показатель преломления для обыкновенного луча равен радиусу окружности, а для необыкновенного луча — расстоянию от начала координат до точки пересечения луча с эллипсовидной кривой (*обратным эллипсом*). Эта диаграмма называется *оптической индикатрисой*, или *эллипсом индексов* кристалла.

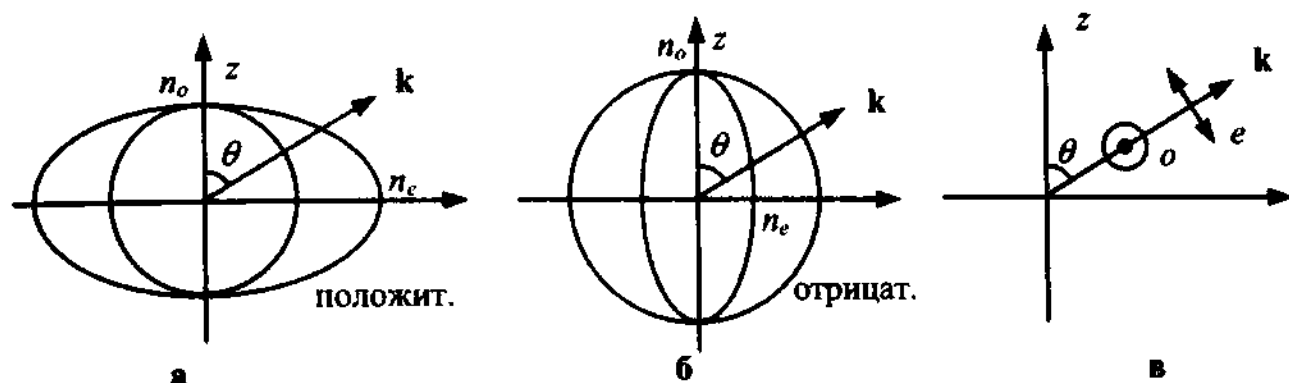


Рис. 12.1.2. Диаграмма показателей преломления (оптическая индикатриса) для обыкновенного (о) и необыкновенного (е) лучей. а — в случае положительного кристалла, б — для отрицательного кристалла; в — ориентация вектора напряжённости электрического поля $\mathbf{E}$ (поляризация) по отношению к оси кристалла z обыкновенного и необыкновенного лучей

Обыкновенный и необыкновенный лучи отличаются поляризацией: в обыкновенном луче вектор напряжённости $\mathbf{E}$ перпендикулярен главной плоскости, а в необыкновенном — лежит в главной плоскости (хотя может быть не перпендикулярным волновому вектору). Поэтому при изменении направления луча угол между оптической осью кристалла и вектором $\mathbf{E}$ в случае *обыкновенного* луча остаётся неизменным, равным $\pi/2$, и меняется в случае *необыкновенного* луча (рис. 12.1.2в). Соответственно меняются показатель преломления и скорость волны:

для обыкновенного луча: $n = n_o = \text{const}$, $v_o = c/n_o$,

для необыкновенного луча: $n = n(\theta)$, $v_e = c/n(\theta)$.

Для необыкновенного луча, распространяющегося вдоль оптической оси ($\theta = 0$), имеем

$$n(0) = n_o, \quad v_e = c/n_o = v_o.$$

Это значит, что обыкновенный и необыкновенный лучи не разделяются: идут вместе и с одинаковыми скоростями. Если же $\theta = \pi/2$, то

$$n(\pi/2) = n_e, \quad v_e = c/n_e \neq v_o.$$

В этом случае оба луча также идут вместе, но с разными скоростями.

Значение показателя преломления для необыкновенного луча в зависимости от угла θ можно найти по формуле (раздел 12.1.3)

$$\frac{1}{n^2(\theta)} = \frac{\sin^2 \theta}{n_e^2} + \frac{\cos^2 \theta}{n_o^2}, \quad \text{или} \quad n(\theta) = \frac{n_o n_e}{\sqrt{n_e^2 \cos^2 \theta + n_o^2 \sin^2 \theta}}.$$

Соответственно скорость необыкновенного луча в зависимости от направления распространения определяется соотношением

$$v^2(\theta) = \frac{c^2}{n^2(\theta)} = v_e^2 \sin^2 \theta + v_o^2 \cos^2 \theta.$$

Кристалл называется *положительным*, если

$$n_e > n_o, \quad v_e < v_o,$$

и *отрицательным* — в противоположном случае:

$$n_e < n_o, \quad v_e > v_o$$

Эффект двойного лучепреломления обусловлен тем, что скорость волны $v = c/n$ тем меньше, чем больше показатель преломления. Поэтому, например, в положительном кристалле наибольшая скорость достигается в направлении оси кристалла:

$$n_e(\theta) > n(0) = n_o \Rightarrow v_e(\theta) < v(0) = v_o.$$

И именно к оси отклоняется необыкновенный луч (рис. 12.1.3). Обыкновенный же луч следует обычным законам преломления и при нормальном падении на поверхность кристалла распространяется без отклонения.

С формальной точки зрения пространственное разделение обыкновенного и необыкновенного лучей связано с тем, что в случае необыкновенной волны (рис. 12.1.3) направления волнового вектора $\mathbf{k}$ и вектора

Пойнтинга $\mathbf{S} = \frac{c}{4\pi} [\mathbf{E}, \mathbf{H}]$, вообще говоря, не совпадают. Для обыкновен-

ной же волны эти два вектора параллельны.

12.1.2. Диэлектрическая проницаемость

В общем случае анизотропной среды вектор поляризации (дипольный момент единицы объёма) среды $\mathbf{P}$ не параллелен приложенному полю. Соответственно векторы $\mathbf{E}$ и $\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P}$ не параллельны, так что можно записать:

$$D_x = \varepsilon_{xx}E_x + \varepsilon_{xy}E_y + \varepsilon_{xz}E_z,$$

$$D_y = \varepsilon_{yx}E_x + \varepsilon_{yy}E_y + \varepsilon_{yz}E_z,$$

$$D_z = \varepsilon_{zx}E_x + \varepsilon_{zy}E_y + \varepsilon_{zz}E_z.$$

В тензорных обозначениях эти равенства имеют вид

$$\mathbf{D} = \hat{\varepsilon}\mathbf{E}, \text{ или } D_i = \sum_{k=1}^3 \varepsilon_{ik}E_k.$$

В последней формуле индексы i, k пробегают значения $i, k = 1(x), 2(y), 3(z)$.

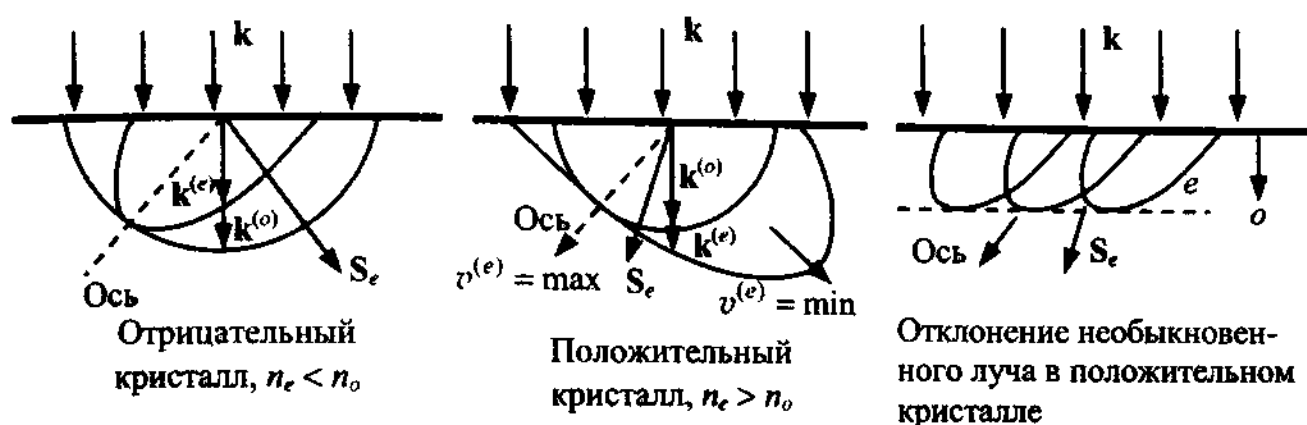


Рис. 12.1.3. Отклонение необыкновенного луча при нормальном падении света на пластину, изготовленную из отрицательного (слева) и положительного (справа) кристаллов

Набор девяти величин ε_{ik} образует *тензор диэлектрической проницаемости*, который является симметричным: $\varepsilon_{ik} = \varepsilon_{ki}$. Эти величины зависят от выбора системы координат, и существует выделенная система координат, в которой рассматриваемый тензор имеет диагональный вид:

$$\varepsilon = \begin{pmatrix} \varepsilon_{xx} & 0 & 0 \\ 0 & \varepsilon_{yy} & 0 \\ 0 & 0 & \varepsilon_{zz} \end{pmatrix}, \quad \begin{cases} D_x = \varepsilon_{xx}E_x, \\ D_y = \varepsilon_{yy}E_y, \\ D_z = \varepsilon_{zz}E_z. \end{cases}$$

Соответствующие координатные оси называются *главными*.

Если

$$\varepsilon_{xx} = \varepsilon_{yy} \neq \varepsilon_{zz},$$

то кристалл является *одноосным*, причём ось z оказывается осью кристалла. Если же все три величины $\varepsilon_{xx}, \varepsilon_{yy}, \varepsilon_{zz}$ различны, то кристалл — *двуосный*.

В случае изотропной среды (или среды с кубической симметрией) главные значения тензора $\hat{\varepsilon}$ совпадают:

$$\varepsilon_{xx} = \varepsilon_{yy} = \varepsilon_{zz} \equiv \varepsilon, \quad \mathbf{D} = \varepsilon \mathbf{E}.$$

Рассмотрим далее случай одноосных кристаллов. Для таких кристаллов вектор индукции можно разложить на проекции вдоль (индекс «||») и перпендикулярно (индекс «⊥») оси кристалла:

$$\mathbf{D} = \varepsilon_{||} \mathbf{E}_{||} + \varepsilon_{\perp} \mathbf{E}, \quad \text{или} \quad D_{||} = \varepsilon_{||} E_{||}, \quad \mathbf{D}_{\perp} = \varepsilon_{\perp} \mathbf{E}_{\perp}.$$

Для обыкновенного луча вектор $\mathbf{E}$ перпендикулярен оси кристалла, $\mathbf{E} \perp$ оси, так что

$$\varepsilon_o = \varepsilon_{\perp}$$

независимо от направления распространения луча. Соответственно показатель преломления и скорость обыкновенного луча оказываются следующими:

$$n_o = \sqrt{\varepsilon_{\perp}}, \quad v_o = c / \sqrt{\varepsilon_{\perp}}.$$

Пусть необыкновенный луч распространяется перпендикулярно оси кристалла. Тогда его вектор $\mathbf{E}$ параллелен оси: $\mathbf{E} \parallel z$, так что

$$\varepsilon_e = \varepsilon_{||}, \quad n_e = \sqrt{\varepsilon_{||}}, \quad v_e = c / \sqrt{\varepsilon_{||}}.$$

12.1.3. Показатель преломления волн в одноосных кристаллах

Найдём показатель преломления луча, распространяющегося под произвольным углом θ к оси кристалла. Ограничимся случаем немагнитного кристалла: $\mu = 1$. Предполагаем, что свободные заряды и токи проводимости отсутствуют.

Исходим из уравнений Максвелла:

$$\begin{aligned} \operatorname{rot} \mathbf{E} &= -\frac{1}{c} \frac{\partial \mathbf{H}}{\partial t}, \quad \operatorname{rot} \mathbf{H} = \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t}, \\ \operatorname{div} \mathbf{D} &= 0, \quad \operatorname{div} \mathbf{H} = 0. \end{aligned} \quad (12.1.1)$$

Рассмотрим бегущие волны. Тогда зависимость полей от координат и времени имеет вид

$$\sim \exp(i\mathbf{k}\mathbf{r} - i\omega t).$$

Соответственно уравнения Максвелла принимают вид

$$\mathbf{k} \times \mathbf{E} = \frac{\omega}{c} \mathbf{H}, \quad \mathbf{k} \times \mathbf{H} = -\frac{\omega}{c} \mathbf{D}, \quad \mathbf{k} \mathbf{D} = 0, \quad \mathbf{k} \mathbf{H} = 0. \quad (12.1.2)$$

Из этих уравнений видно, что векторы $\mathbf{D}$, $\mathbf{H}$ и $\mathbf{k}$ взаимно ортогональны (рис. 12.1.4а).

Поскольку вектор Пойнтинга равен $\mathbf{S} = \frac{c}{4\pi} [\mathbf{E}, \mathbf{H}]$, то взаимно ортогональную тройку образуют также векторы $\mathbf{E}$, $\mathbf{H}$ и $\mathbf{S}$ (рис. 12.1.4б)

Необыкновенный луч

В этом случае вектор $\mathbf{E}$ перпендикулярен вектору $\mathbf{H}$ и, следовательно, лежит в плоскости, образуемой векторами $\mathbf{D}$ и $\mathbf{k}$. Это значит, что если векторы $\mathbf{E}$ и $\mathbf{D}$ не параллельны, то направления волнового вектора $\mathbf{k}$ и вектора Пойнтинга $\mathbf{S}$ также не совпадают (рис. 12.1.4в). Поэтому луч уже может распространяться в направлении, отличном от $\mathbf{k}$.

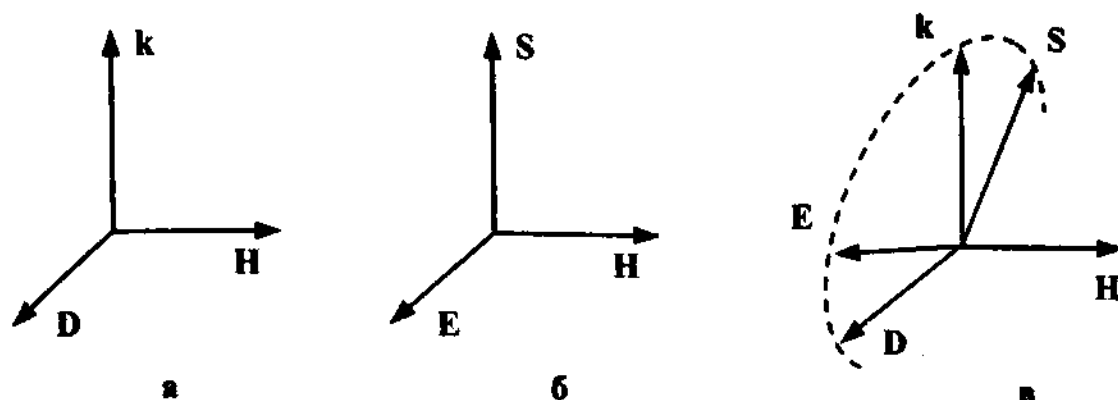


Рис. 12.1.4. Взаимное расположение векторов $\mathbf{D}$, $\mathbf{E}$, $\mathbf{H}$, $\mathbf{k}$ и $\mathbf{S}$ в одноосном кристалле: векторы $\mathbf{D}$, $\mathbf{E}$, $\mathbf{k}$ и $\mathbf{S}$ лежат в плоскости, перпендикулярной вектору $\mathbf{H}$, причём угол между $\mathbf{D}$ и $\mathbf{E}$ такой же, как и между $\mathbf{k}$ и $\mathbf{S}$

Из уравнений Максвелла в форме (12.1.2) следует, что для необыкновенной волны вектор напряжённости магнитного поля перпендикулярен главной плоскости кристалла, содержащей вектор $\mathbf{E}$ (рис. 12.1.5). В этой же плоскости лежит и вектор $\mathbf{D}$, причём $\mathbf{D} \perp \mathbf{k}$.

Рассмотрим относительное расположение векторов $\mathbf{D}$, $\mathbf{E}$, $\mathbf{k}$ и оптической оси кристалла. Обозначим угол между волновым вектором и осью как θ , а угол между векторами $\mathbf{D}$ и $\mathbf{E}$ — φ .

Из уравнения $\mathbf{k} \times \mathbf{H} = -\frac{\omega}{c} \mathbf{D}$ следует

$$H = \frac{\omega}{ck} D. \quad (12.1.3)$$

Из уравнения $\mathbf{k} \times \mathbf{E} = \frac{\omega}{c} \mathbf{H}$ находим

$$H = \frac{ck}{\omega} E \cos \varphi \quad (12.1.4)$$

(учтено, что согласно рис. 12.1.5 угол между векторами $\mathbf{E}$ и $\mathbf{k}$ равен $\pi/2 - \varphi$). Из двух последних равенств получаем уравнение, связывающее длины векторов $\mathbf{E}$ и $\mathbf{D}$ в необыкновенной волне:

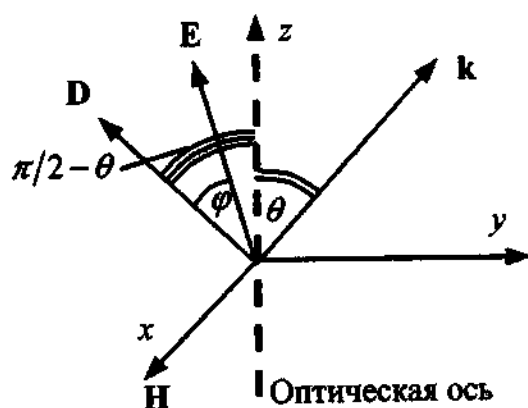


Рис. 12.1.5. Взаимное расположение векторов $\mathbf{D}$, $\mathbf{E}$, $\mathbf{H}$, $\mathbf{k}$ в одноосном кристалле для необыкновенной волны: векторы $\mathbf{D}$, $\mathbf{E}$, $\mathbf{k}$ и ось кристалла z лежат в одной плоскости, перпендикулярной вектору $\mathbf{H}$, $\mathbf{D} \perp \mathbf{k}$

$$D = \left(\frac{ck}{\omega} \right)^2 E \cos \varphi. \quad (12.1.5)$$

Умножив почленно это равенство на D , перепишем его в векторном виде:

$$D^2 = \left(\frac{ck}{\omega} \right)^2 \mathbf{D} \mathbf{E}. \quad (12.1.6)$$

Поскольку показатель преломления n определяется равенством

$$k = \omega n / c \Rightarrow ck / \omega = n, \quad (12.1.7)$$

то из (12.1.6) следует

$$\frac{1}{n^2} = \frac{\mathbf{D} \mathbf{E}}{D^2}. \quad (12.1.8)$$

Далее воспользуемся равенством

$$\mathbf{D} \mathbf{E} = D_{\parallel} E_{\parallel} + D_{\perp} E_{\perp} = \frac{D_{\parallel}^2}{\epsilon_{\parallel}} + \frac{D_{\perp}^2}{\epsilon_{\perp}}. \quad (12.1.9)$$

Наконец, учитывая, что согласно рис. 12.1.5

$$D_{\parallel} = D \sin \theta, \quad D_{\perp} = D \cos \theta, \quad (12.1.10)$$

получаем окончательно для необыкновенного луча

$$\frac{1}{n^2(\theta)} = \frac{\sin^2 \theta}{n_e^2} + \frac{\cos^2 \theta}{n_o^2}, \quad n_o = \sqrt{\epsilon_{\perp}}, \quad n_e = \sqrt{\epsilon_{\parallel}}. \quad (12.1.11)$$

Обыкновенный луч

Найдём теперь показатель преломления для обыкновенного луча. В рассматриваемом случае векторы $\mathbf{E}$ и $\mathbf{D}$ перпендикулярны главной плоскости кристалла. Ортогональность $\mathbf{E}$ этой плоскости следует из определе-

ния обыкновенного луча. Как следует из уравнения $\mathbf{k} \times \mathbf{E} = \frac{\omega}{c} \mathbf{H}$, в глав-

ной плоскости теперь лежат векторы $\mathbf{H}$ и $\mathbf{k}$. Но согласно уравнению

$\mathbf{k} \times \mathbf{H} = -\frac{\omega}{c} \mathbf{D}$ этим же векторам перпендикулярен и $\mathbf{D}$. Таким образом,

для обыкновенного луча $\mathbf{D} \parallel \mathbf{E}$. Полагая $\mathbf{D} = \epsilon_{\perp} \mathbf{E}$, из (12.1.6) находим:

$$D^2 = \left(\frac{ck}{\omega} \right)^2 D E \Rightarrow \epsilon_{\perp} = \left(\frac{ck}{\omega} \right)^2. \quad (12.1.12)$$

Поскольку в данном случае $k = n_o \omega / c$, то отсюда находим

$$\epsilon_{\perp} = n_o^2 \Rightarrow n_o = \sqrt{\epsilon_{\perp}}. \quad (12.1.13)$$

12.1.4. Направление вектора Пойнтинга

Найдём угол φ , образуемый вектором Пойнтинга $\mathbf{S}$ и волновым вектором $\mathbf{k}$ необыкновенной волны. Как видно из рис. 12.1.6, угол φ между векторами $\mathbf{S}$ и $\mathbf{k}$ равен углу между векторами $\mathbf{D}$ и $\mathbf{E}$. Соответственно искомый угол φ определится из равенства

$$\begin{aligned} \cos \varphi &= \frac{\mathbf{E} \mathbf{D}}{\sqrt{\mathbf{E}^2} \sqrt{\mathbf{D}^2}} = \frac{E_{\parallel} D_{\parallel} + E_{\perp} D_{\perp}}{\sqrt{E_{\parallel}^2 + E_{\perp}^2} \sqrt{D_{\parallel}^2 + D_{\perp}^2}} = \\ &= \frac{D_{\parallel}^2 / \epsilon_{\parallel} + D_{\perp}^2 / \epsilon_{\perp}}{\sqrt{(D_{\parallel} / \epsilon_{\parallel})^2 + (D_{\perp} / \epsilon_{\perp})^2} \sqrt{D_{\parallel}^2 + D_{\perp}^2}}. \end{aligned} \quad (12.1.14)$$

Далее заметим следующее. Поскольку $\mathbf{D} \perp \mathbf{k}$, то

$$\mathbf{D} \mathbf{k} = D_{\parallel} k_{\parallel} + D_{\perp} k_{\perp} = 0.$$

Полагая $k_{\parallel} = k \cos \theta$, $k_{\perp} = k \sin \theta$, находим

$$D_{\parallel} = D \sin \theta, \quad D_{\perp} = -D \cos \theta \quad (12.1.15)$$

(эти равенства можно сразу получить, используя рис. 12.1.6 и замечая, что вектор $\mathbf{D}$ образует с осью кристалла угол $\pi/2 - \theta$).

Имея в виду последнее равенство, перепишем (12.1.14) в следующем виде:

$$\cos \varphi = \frac{\epsilon_{\perp} \sin^2 \theta + \epsilon_{\parallel} \cos^2 \theta}{\sqrt{\epsilon_{\perp}^2 \sin^2 \theta + \epsilon_{\parallel}^2 \cos^2 \theta}}.$$

Найдём также синус угла φ :

$$\sin \varphi = \sqrt{1 - \cos^2 \varphi} = \frac{(\epsilon_{\parallel} - \epsilon_{\perp}) \sin \theta \cos \theta}{\sqrt{\epsilon_{\parallel}^2 \cos^2 \theta + \epsilon_{\perp}^2 \sin^2 \theta}}.$$

Из последних двух равенств следует:

$$\operatorname{tg} \varphi = \frac{(\epsilon_{\parallel} - \epsilon_{\perp}) \sin 2\theta}{\epsilon_{\parallel} + \epsilon_{\perp} + (\epsilon_{\parallel} - \epsilon_{\perp}) \cos 2\theta}. \quad (12.1.16)$$

Легко видеть, что в частных случаях, когда луч распространяется вдоль ($\theta = 0$) и перпендикулярно ($\theta = \pi/2$) оси кристалла, из последней формулы вытекает $\varphi = 0$. Это значит, что направления волнового вектора и вектора Пойнтинга совпадают, и луч не отклоняется от исходного направления.

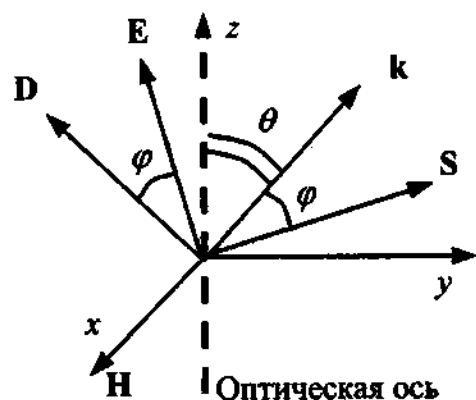


Рис. 12.1.6. Взаимное расположение векторов $\mathbf{D}$, $\mathbf{E}$, $\mathbf{H}$, $\mathbf{k}$ и $\mathbf{S}$ в одноосном кристалле

12.2. Элементарная теория одноосных диэлектриков

Рассмотрим простую модель, иллюстрирующую возникновение оптической анизотропии диэлектриков.

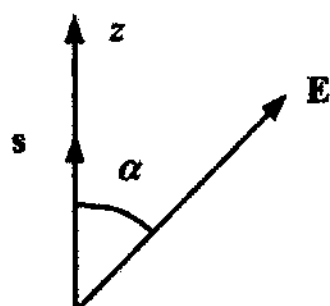


Рис. 12.2.1. Электрическое поле E волны направлено под углом α к директору — оси кристалла z

Предположим, что среда состоит из длинных молекул, ориентированных вдоль определённого направления, называемого *директором* и являющегося осью кристалла. Электроны в молекулах могут совершать колебания только вдоль этого направления (рис. 12.2.1). Считая, что выделенное направление есть ось z , запишем уравнение движения электрона:

$$m\ddot{z} = -\kappa z - \beta\dot{z} + eE_z. \quad (12.2.1)$$

Вводя стандартные обозначения:

$$\omega^2 = \kappa/m, \quad 2\gamma = \beta/m,$$

перепишем это уравнение в виде

$$\ddot{z} + 2\gamma\dot{z} + \omega^2 z = eE_z/m. \quad (12.2.2)$$

Если единичный вектор вдоль оси кристалла обозначить как s , то

$$E_z = sE = E_0 \cos \alpha \cdot e^{-i\omega t} \quad (12.2.3)$$

составляющая электрического поля волны вдоль оси кристалла (директора). Решение уравнения (12.2.2) имеет вид

$$z = g(\omega) \frac{e}{m} sE, \quad g(\omega) = \frac{1}{\omega_0^2 - \omega^2 - 2i\gamma\omega}. \quad (12.2.4)$$

Отсюда находим вектор поляризации среды:

$$\mathbf{P} = (0, 0, P_z), \quad P_z = Nez = \frac{Ne^2}{m} g(\omega) E_z. \quad (12.2.5)$$

Здесь N — число электронов в единице объёма среды. В векторном виде полученное соотношение записывается в виде

$$\mathbf{P} = sP_z = \frac{Ne^2}{m} g(\omega) (\mathbf{E}s)s. \quad (12.2.6)$$

Поскольку $\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P}$, то

$$\mathbf{D} = \mathbf{E} + \frac{4\pi Ne^2}{m} g(\omega) (\mathbf{E}s)s, \quad (12.2.7)$$

или

$$D_i = E_i + \sum_{k=1}^3 \omega_p^2 g(\omega) s_i s_k E_k = \sum_{k=1}^3 \left(\delta_{ik} + \omega_p^2 g(\omega) s_i s_k \right) E_k. \quad (12.2.8)$$

Введено обозначение для плазменной частоты: $\omega_p^2 = 4\pi Ne^2/m$. Проекции единичного вектора s таковы (рис. 12.2.1):

$$s_x = s_y = 0, \quad s_z = 1. \quad (12.2.9)$$

Полагая $D_i = \sum_{k=1}^3 \epsilon_{ik} E_k$, получаем выражение для тензора диэлектрической проницаемости:

$$\epsilon_{ik} = \delta_{ik} + \omega_p^2 g(\omega) s_i s_k. \quad (12.2.10)$$

Здесь δ_{ik} — единичный тензор. В главных осях (считая, как и раньше, ось z осью кристалла) имеем

$$\epsilon = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \epsilon_{zz} \end{pmatrix}, \quad \epsilon_{zz} = 1 + \omega_p^2 g(\omega). \quad (12.2.11)$$

Таким образом, мы имеем дело со средой, у которой

$$\epsilon_{xx} = \epsilon_{yy} = 1, \quad \epsilon_{zz} = 1 + \omega_p^2 g(\omega). \quad (12.2.12)$$

Соответственно для обыкновенного луча: $\epsilon_o = 1$, а для необыкновенного луча: $\epsilon_e = 1 + \omega_p^2 g(\omega)$.

12.3. Кристаллические пластинки $\lambda/2$ и $\lambda/4$

12.3.1. Определения

Пусть пластинка вырезана из одноосного кристалла параллельно его оси (рис. 12.3.1). В такой пластинке обыкновенный и необыкновенный лучи не разделяются, но им отвечают разные показатели преломления: n_o и n_e . Поэтому, пройдя пластинку, лучи приобретают дополнительную разность хода

$$\Delta = h(n_e - n_o). \quad (12.3.1)$$

Соответствующая разность фаз равна

$$\Delta\varphi = \frac{2\pi}{\lambda} h(n_e - n_o). \quad (12.3.2)$$

Здесь $\lambda = 2\pi c/\omega$ — длина волны света в вакууме.

Выделяются два случая.

1. Если

$$\Delta = \frac{\lambda}{4} + m\lambda, \quad m = 0, 1, 2, \dots, \quad (12.3.3)$$

пластинка называется *пластинкой в четверть волны* (или *четвертьволновой пластинкой*, *пластинкой $\lambda/4$*). Эта пластинка вносит дополнительную разность фаз обыкновенного и необыкновенного лучей

$$\Delta\varphi = \frac{\pi}{2} + 2\pi m. \quad (12.3.4)$$

2. Если

$$\Delta = \frac{\lambda}{2} + m\lambda, \quad m = 0, 1, 2, \dots, \quad (12.3.5)$$

то пластинка называется *пластинкой в полволны* (или *полуволновой пластинкой*, *пластинкой $\lambda/2$*). Эта пластинка создаёт дополнительную разность фаз

$$\Delta\varphi = \pi + 2\pi m. \quad (12.3.6)$$

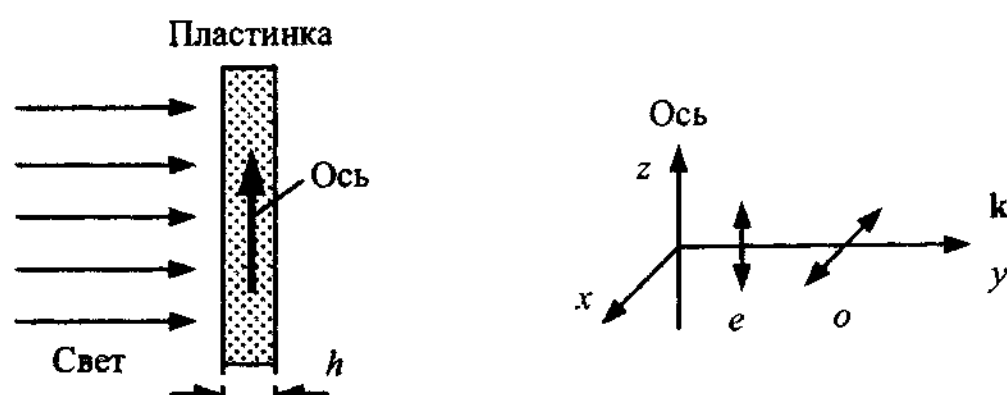


Рис. 12.3.1. Свет падает по нормали на пластинку, вырезанную вдоль оптической оси кристалла. Слева — схема эксперимента, справа — взаимная ориентация оси кристалла, волнового вектора $\mathbf{k}$ и вектора напряжённости электрического поля волны $\mathbf{E}$. Пластинка лежит в плоскости (x, y)

Отметим, что вследствие дисперсии показателя преломления обыкновенного и необыкновенного лучей зависят от длины волны излучения. Поэтому, говоря о пластинках $\lambda/4$, $\lambda/2$, ..., имеют в виду определённую спектральную компоненту.

12.3.2. Действие четвертьволновой пластинки

1. Волна с круговой поляризацией

После прохождения такой волной четвертьволновой пластинки она оказывается линейно поляризованной. Действительно, исходная волна описывается формулами

$$\begin{cases} E_x = E_0 \cos \omega t, \\ E_z = E_0 \cos \omega t. \end{cases} \quad (12.3.7)$$

Если эту волну пропустить через пластинку $\lambda/4$ (как на рис. 12.3.1), то независимо от ориентации пластинки одна компонента излучения оказывается обыкновенным лучом, а другая — необыкновенным. Соответственно после прохождения пластинки между компонентами возникает сдвиг фаз $\pi/2$, а сама уже будет описываться формулами

$$\begin{cases} E_x = E_0 \cos \omega t, \\ E_z = \pm E_0 \sin \omega t. \end{cases} \quad (12.3.8)$$

Но это отвечает круговой поляризации.

2. Линейно поляризованная волна

Если исходная волна была линейно поляризованной, то после прохождения пластинки она превращается в волну с эллиптической поляризацией. Действительно, пусть ось кристалла есть ось z , а плоскость поляризации есть $\{x, z\}$. Тогда линейно поляризованная волна, падающая на пластинку, описывается выражениями

$$\begin{cases} E_x = E_{x0} \cos \omega t, \\ E_z = E_{z0} \cos \omega t. \end{cases} \quad (12.3.9)$$

Волна, прошедшая пластинку, описывается уже следующими формулами:

$$\begin{cases} E_x = E_{x0} \cos(\omega t + \delta), \\ E_z = E_{z0} \cos\left(\omega t + \delta - \frac{\pi}{2}\right), \end{cases} \quad \text{или} \quad \begin{cases} E_x = E_{x0} \cos(\omega t + \delta), \\ E_z = E_{z0} \sin(\omega t + \delta). \end{cases} \quad (12.3.10)$$

Последние соотношения описывают эллиптически поляризованную волну с полуосями (E_{x0}, E_{z0}) . Оси эллипса (рис. 12.3.2, слева) направлены по координатным осям $\{x, z\}$. В частном случае, когда в исходной волне $E_{x0} = E_{z0} \equiv E_0$, прошедшая волна приобретает круговую поляризацию.

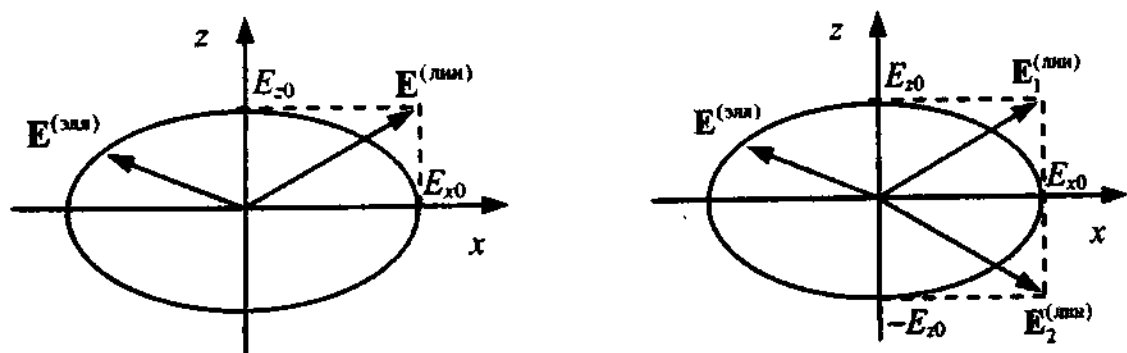


Рис. 12.3.2. Слева — преобразование волны с линейной поляризацией в эллиптически поляризованную; справа — после прохождения пластинки $\lambda/4$ эллиптически поляризованная волна превращается в линейно поляризованную, если ось пластинки направлена вдоль оси z , или вдоль оси x (порождая волны $E_1^{(\text{лин})}$ или $E_2^{(\text{лин})}$)

Можно поставить обратную задачу: пропустить полученную эллиптически поляризованную волну через пластинку $\lambda/4$ и далее, вращая пластинку, получить линейно поляризованную волну. Оказывается, мы можем получить уже две различные волны. Действительно, направляя ось кристалла вдоль осей z или x , получаем из волны (12.3.9)

$$\begin{cases} E_x = E_{x0} \cos(\omega t + \delta), \\ E_z = E_{z0} \sin(\omega t + \delta \pm \pi/2), \end{cases} \Rightarrow \begin{cases} E_x = E_{x0} \cos(\omega t + \delta), \\ E_z = \pm E_{z0} \cos(\omega t + \delta). \end{cases}$$

На рис. 12.3.2 (справа) показаны возникающие линейно поляризованные волны $E_1^{(\text{лин})}$, $E_2^{(\text{лин})}$.

12.3.3. Действие полуволновой пластинки

Если падающая волна была линейно поляризованной, то в результате прохождения пластинки $\lambda/2$ меняется плоскость поляризации, как показано на рис. 12.3.3. Этот результат объясняется тем, что возникает дополнительный сдвиг фаз обыкновенного и необыкновенного лучей в π . Это можно интерпретировать так, что просто изменяется знак одной из проекций вектора E , следствие чего видно из рис. 12.3.3.

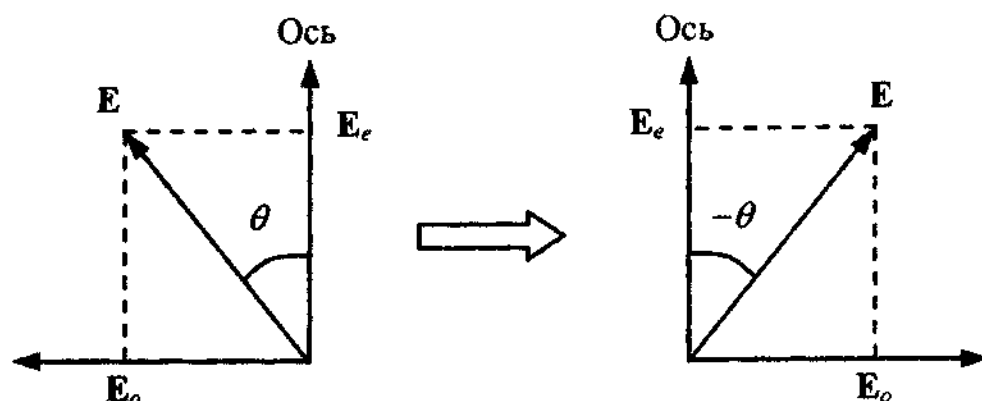


Рис. 12.3.3. Линейно поляризованная волна (слева) после прохождения полуволновой пластинки меняет направление плоскости поляризации (справа)

12.4. Интерференция поляризованных лучей

12.4.1. Постановка задачи

В одноосных кристаллах в зависимости от поляризации различаются обыкновенная и необыкновенная волны. Для них различны показатели преломления и, следовательно, фазовые скорости. В общем случае волну с произвольной поляризацией можно представить как суперпозицию волн, поляризованных вдоль и поперёк оси кристалла. Если волна распространяется вдоль оси x , то

$$E = e_1 A_1 \cos(k_1 x - \omega t) + e_2 A_2 \cos(k_2 x - \omega t + \varphi).$$

Здесь e_1 и e_2 — единичные векторы вдоль соответствующих направлений, причём $e_1 \perp e_2$. Последнее приводит к тому, что интенсивность излучения оказывается постоянной:

$$I = I_1 + I_2, \quad I_1 = \frac{v_1}{8\pi} A_1^2, \quad I_2 = \frac{v_2}{8\pi} A_2^2,$$

где v_i — скорость i -й волны. Это означает, что интерференционные явления отсутствуют: при любой разности хода первой и второй волн интенсивность остаётся неизменной (разумеется, если нет потерь энергии).

Вместе с тем сдвиг фаз, не меняя интенсивности излучения, изменяет его поляризацию. Это изменение можно зарегистрировать, направляя свет на поляроид и сводя поляризации ортогональных компонент к одной линейной, направленной вдоль оси поляроида (раздел 9.3.3). Таким способом можно наблюдать *интерференцию поляризованных лучей (волн)*.

Одна из простейших схем показана на рис. 12.4.1. С помощью николя N_1 свет превращается в линейно поляризованный и направляется на призму, изготовленную из двулучепреломляющего кристалла. В кристалле свет делится на обыкновенную и необыкновенную волны. Поскольку показатели преломления этих волн различны, то они отклоняются призмой на разные углы. Далее с помощью николя N_2 поляризации компонент сводятся к одной, и на экране Э наблюдается серия интерференционных полос. Наибольшей чувствительностью схема обладает, когда николи взаимно перпендикулярны и образуют угол 45° с осью кристалла.

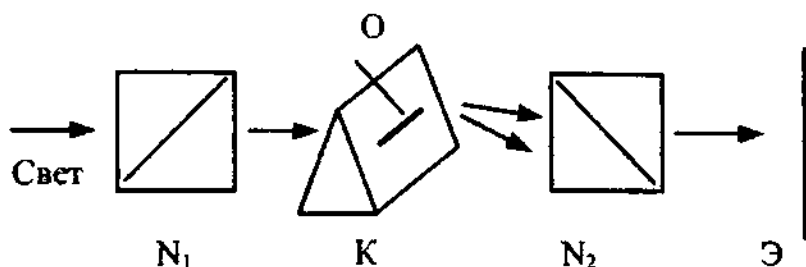


Рис. 12.4.1. Наблюдение интерференции поляризованных волн в параллельных лучах. К — кристаллическая пластинка, N_1 — поляризатор, N_2 — анализатор, О — оптическая ось кристаллической пластинки К

Укажем ещё две схемы наблюдения интерференции поляризованных лучей.

12.4.2. Интерференция в параллельных лучах

Схема установки показана на рис. 12.4.2. Из поляризатора N_1 выходит линейно поляризованная волна E . В пластинке К это волна делится на обыкновенную E_o и необыкновенную E_e , между которыми возникает сдвиг фаз. В результате линейная поляризация превращается в эллиптическую, причём в пластинке толщины d компоненты приобретают разность фаз

$$\delta = \frac{2\pi d}{\lambda}(n_e - n_o).$$

Если угол между осями поляризатора и анализатора равен 0° или 90° , то установка наиболее чувствительна к обнаружению анизотропии.

Вращая анализатор N_2 , наблюдают периодические изменения интенсивности прошедшей волны.

Если угол между осью поляризатора N_1 и кристаллической пластинки K равен 45° , то амплитуды E_e и E_o равны, и в случае сдвига фаз волн на $\delta = \pi/2$ возникает круговая поляризация.

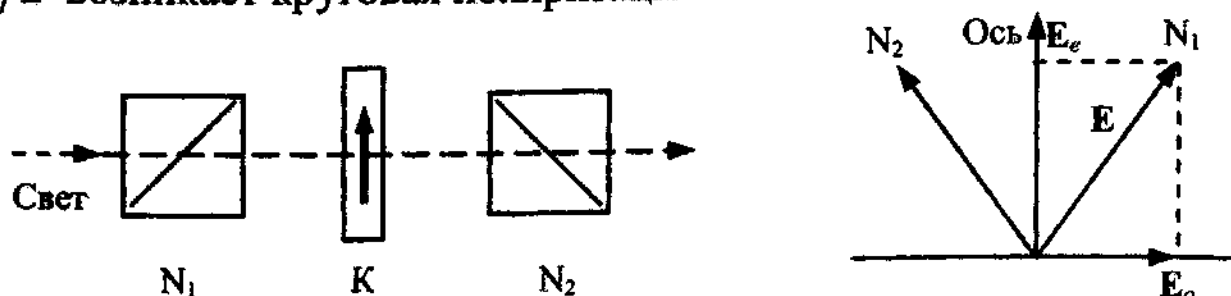


Рис. 12.4.2. Наблюдение интерференции поляризованных волн в параллельных лучах. K — кристаллическая пластинка, N_1 — поляризатор, N_2 — анализатор, Ось — оптическая ось кристаллической пластинки K

В случае пластинки переменной толщины h в зависимости от локального значения h соответствующий её участок виден в монохроматическом свете либо тёмным, либо светлым. В белом свете вследствие зависимости $\delta(\lambda)$ наблюдается различная окраска разных участков пластинки. Это явление называют *хроматической поляризацией*.

12.4.3. Интерференция в сходящихся лучах

Одна из схем показана на рис. 12.4.3. Пластинка K может быть вырезана как перпендикулярно оптической оси (как показано на рис. 12.4.3), так и параллельно ей. После прохождения поляризатора N_1 и линзы L_1 возникают линейно поляризованные волны (с одинаковой поляризацией), имеющие разный наклон к оптической оси системы. Эти лучи проходят в кристаллической пластинке разные пути и приобретают сдвиг фаз

$$\delta = \frac{2\pi d}{\lambda} \cos \psi (n_e - n_o),$$

где ψ — угол между направлением луча и нормалью к поверхности кристалла.

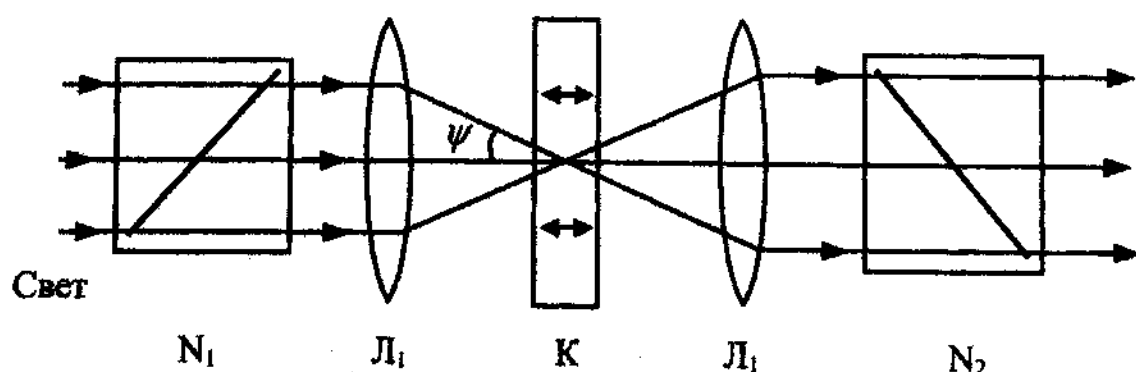


Рис. 12.4.3. Схема наблюдения интерференции поляризованных волн в сходящихся лучах. N_1 и N_2 — скрещенные поляризатор и анализатор, L_1 и L_2 — софокусные линзы, K — кристаллическая пластинка, вырезанная перпендикулярно оптической оси

Наблюдаются различные, так называемые коноскопические, фигуры. Примеры таких структур показаны на рис. 12.4.4 для случая, когда николи выставлены взаимно перпендикулярно. Если же их оси параллельны, то наблюдается дополнительная картина (белый и чёрный цвета меняются местами).

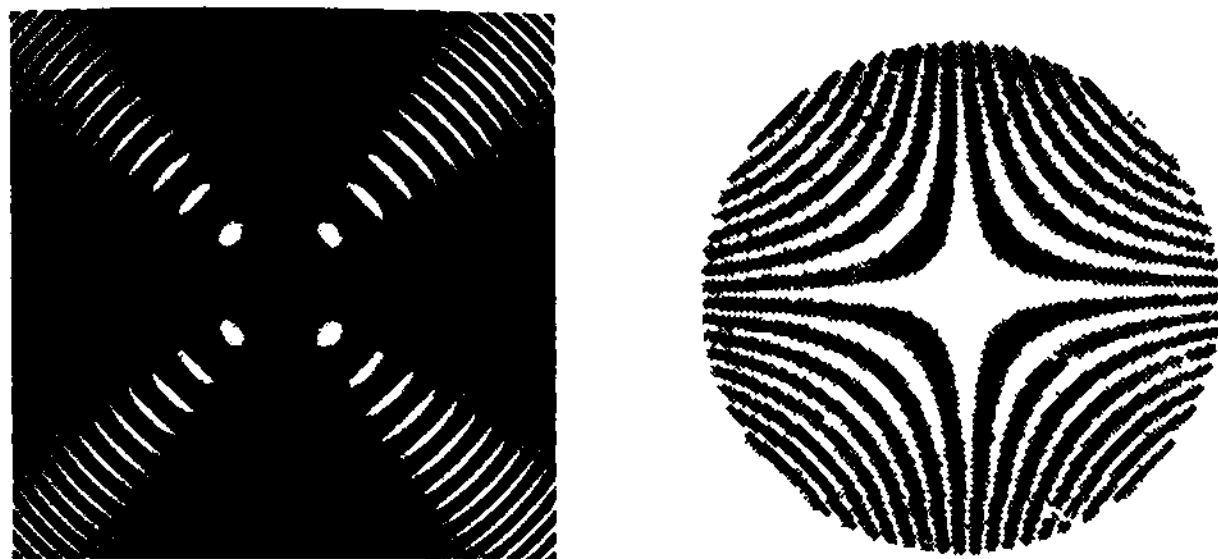


Рис. 12.4.4. Слева — интерференционные полосы в случае кристаллической пластинки, вырезанной перпендикулярно оптической оси; справа — то же для пластинки, вырезанной параллельно оптической оси

12.5. Способы получения линейно поляризованного света

Действие ряда поляризаторов основано на различии показателей преломления для обыкновенного и необыкновенного лучей ($n_e \neq n_o$), как это проиллюстрировано на рис. 12.5.1. Во многих из них используется также эффект полного внутреннего отражения, которое вследствие неравенства $n_e \neq n_o$ имеет место для обыкновенного и необыкновенного лучей при различных углах падения.

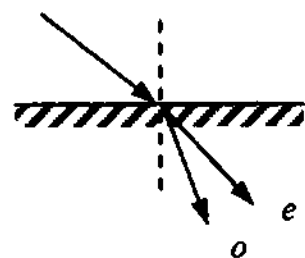


Рис. 12.5.1. Преломление обыкновенного и необыкновенного лучей, различающихся показателем преломления ($n_e < n_o$)

Остановимся подробнее на двух типах поляризаторов.

1. *Призма Николя* (1828) показана на рис. 12.5.2. Эта призма изготовлена из исландского шпата, разрезана по диагонали AB и склеена канадским бальзамом. Показатель преломления клея $n = 1,549$. В шпате $n_e < n_o$ (для необыкновенного и обыкновенного лучей), причем $n_e < n < n_o$. Обыкновенный луч, падая на границу раздела AB , испытывает полное внутреннее отражение (как от оптически менее плотной среды), а необыкновенный луч, частично отражаясь от слоя бальзама, частично проходит через этот слой. Луч, вы-

ходящий из второй призмы и параллельный исходному лучу, полностью поляризован в главной плоскости.

Призма Николя изготовлена в виде параллелепипеда со скошенными гранями. Поэтому проходящий луч испытывает параллельное смещение, и при вращении призмы вокруг входящего луча выходящий луч описывает окружность. От этого недостатка свободны другие поляризаторы, к числу которых относится призма Глана–Тейлора.

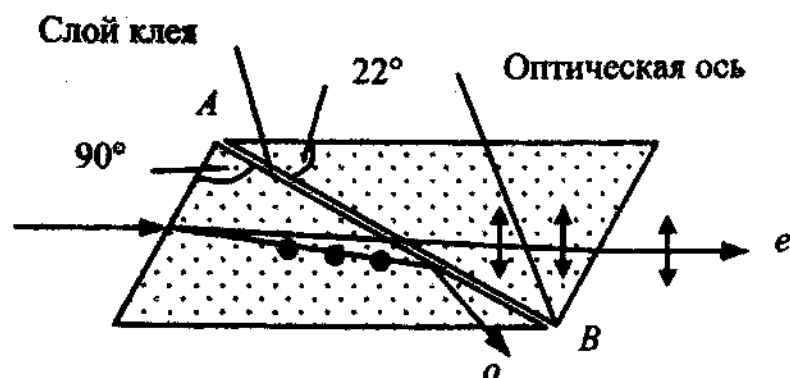


Рис. 12.5.2. Призма Николя (николь). Промежуток между треугольными призмами заполнен канадским бальзамом

2. Призма Глана–Тейлора (1948) — одна из наиболее часто используемых в настоящее время призм. Она предназначена для преобразования излучения с произвольной поляризацией в линейно поляризованное и используется для работы в УФ-области спектра, а также при работе с пучками мощного оптического излучения.

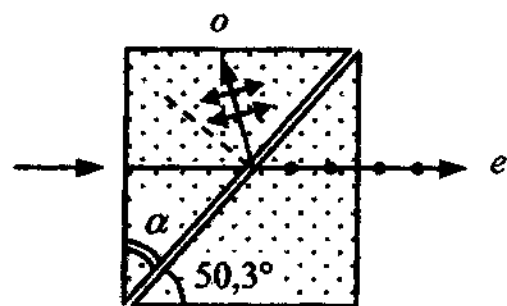


Рис. 12.5.3. Призма Глана–Тейлора. Оптическая ось обеих трёхгранных призм перпендикулярна плоскости рисунка

В данном устройстве (рис. 12.5.3) две треугольные призмы разделены воздушным зазором и вместе образуют прямоугольный параллелепипед. Призмы изготовлены из кальцита с показателями преломления для обыкновенного и необыкновенного лучей $n_o = 1,658$ и $n_e = 1,486$. Угол α подобран так, что луч падает под углом, близким к углу полного внутреннего отражения. При этом обыкновенный луч испытывает внутреннее отражение, а необыкновенный луч проходит через призму, выходя из неё без смещения.

ГЛАВА 13

ИСКУССТВЕННАЯ ОПТИЧЕСКАЯ АНИЗОТРОПИЯ

13.1. Пьезооптический эффект

В предыдущей главе мы рассмотрели распространение света в изначально анизотропных средах. Однако оптическая анизотропия может возникать и в средах исходно однородных, если их подвергнуть, например, механическому воздействию.

Это явление, открытое Томасом Иоганном Зеебеком (Германия) в 1813 г. и Д. Брюстером (Шотландия) в 1816 г., получило название «*пьезооптический эффект*» или «*фотоупругость*».

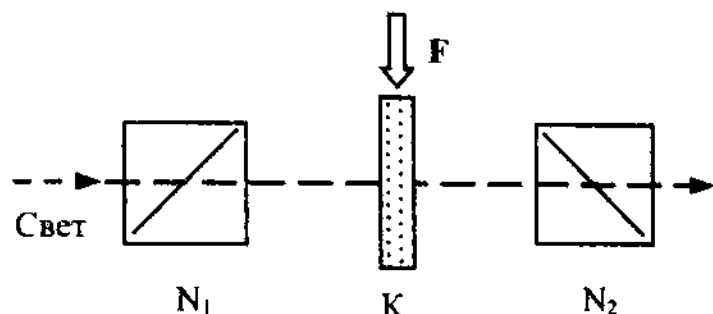


Рис. 13.1.1. Схема наблюдения пьезооптического эффекта. K — материал, деформируемый, силой F , N_1 — поляризатор, N_2 — анализатор

При одноосном растяжении или сжатии изотропное тело может приобретать свойства оптически одноосного кристалла с осью, параллельной оси растяжения или сжатия. Данное явление можно наблюдать в схеме, показанной на рис. 13.1.1. В приведённой схеме в отсутствие деформации материала свет через николю N_2 не проходит. Однако при приложении усилия F возникает разница показателей преломления для обыкновенного и необыкновенного лучей:

$$n_e - n_o = k\sigma \quad (13.1.1)$$

(σ — напряжение, k — коэффициент, зависящий от свойств материала). В результате в проходящем свете будут наблюдаться различные картины (рис. 13.1.1).

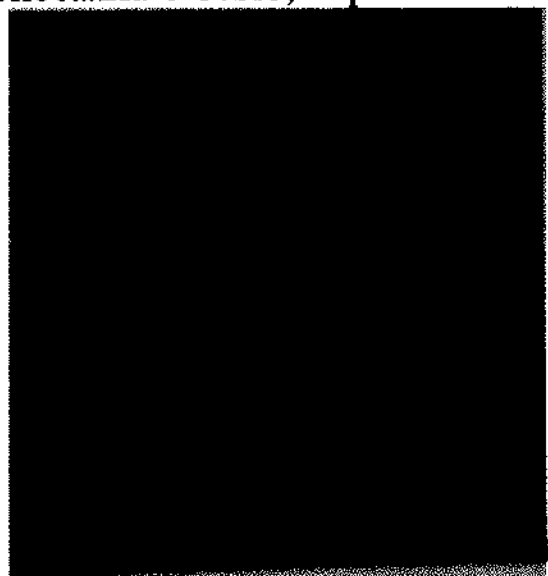


Рис. 13.1.1. Деформированный полимерный образец приобретает свойства оптически анизотропной среды

В случае произвольных деформаций вещество может приобретать свойства двуосного кристалла.

Фотоупругость обусловлена деформацией электронных оболочек атомов и молекул, ориентацией оптически анизотропных молекул, а также раскручиванием длинных молекулярных цепочек.

Явление фотоупругости имеет многочисленные применения. В частности, в материаловедении на этом основан эффективный метод исследования закономерностей разрушения материалов.

13.2. Эффект Поккельса

В некоторых средах при наложении постоянного электрического поля возникает двойное лучепреломление. При этом изменение показателя преломления пропорционально величине поля:

$$n_e - n_o = bE. \quad (13.2.1)$$

В этой связи данное явление называют также *линейным электрооптическим эффектом*.

Описанный эффект был обнаружен немецким физиком Фридрихом Поккельсом в 1893 г. и позже назван его именем — «*эффект Поккельса*». Явление длительное время не привлекало особого внимания, поскольку для заметных проявлений анизотропии обычно требуются сильные электрические поля (десятки и сотни кВ). Позже было обнаружено большое число кристаллических веществ, в которых эффект проявляется сильно даже при малых управляющих напряжениях (порядка нескольких десятков вольт).

Эффект хорошо выражен в кристаллах ниобата лития (LiNbO_3) и арсенида галлия (GaAs). Он может наблюдаться только в кристаллах, не обладающих центром симметрии. Действительно, в силу линейности зависимости (13.2.1) эффект при изменении направления поля должен менять знак, что невозможно в веществах с центральной симметрией, в которых все направления эквивалентны.

Поскольку эффект Поккельса связан с изменением электронной конфигурации под действием внешнего поля, он оказывается практически безынерционным (быстродействие порядка 10^{-10} с).

На основе эффекта созданы модуляторы света. При этом благодаря линейной зависимости показателя преломления от напряжённости электрического поля нелинейные искажения сигнала при модуляции невелики. Основным элементом модуляторов является *ячейка Поккельса*, представляющая собой кристалл, помещённый между двумя скрещенными николями (рис. 13.2.1). Николи не пропускают свет в отсутствие электрического поля, но при наложении поля вследствие возникновения двойного

лучепреломления в кристалле и вызванного этим изменения поляризации свет частично проходит.

Остановимся на теории эффекта Поккельса.

Направим ось кристалла вдоль оси x . Тогда при отклонении электрона от положения равновесия вдоль этого направления возникает возвращающая сила

$$F_{\text{возвр}} = -\kappa_{1,x}x - \kappa_2x^2. \quad (13.2.2)$$

Здесь мы учли поправку $(-\kappa_2x^2)$, которая может оказаться существенной при относительно больших смещениях x . Сила (13.1.3) различна для смещений в положительном и отрицательном направлениях, что необходимо для наблюдения эффекта Поккельса.

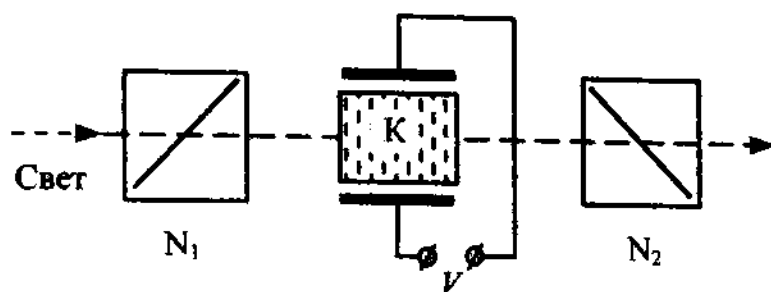


Рис. 13.2.1. Схема ячейки Поккельса. K — кристалл, помещённый во внешнее электрическое поле, N_1 — поляризатор, N_2 — анализатор

Полагаем также, что в поперечном направлении сила является нечётной функцией смещения, причём с той же точностью, что и (13.2.2), она даётся выражением

$$F_{\text{возвр}} = -\kappa_{1,y}y. \quad (13.2.3)$$

Пусть электрическое поле в конденсаторе направлено вдоль оси x . Рассмотрим сначала случай, когда вектор электрического поля волны направлен также вдоль оси x . Запишем уравнение движения под действием электрического поля световой волны (E) и поля в конденсаторе (E_0):

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2x + \beta x^2 = \frac{e}{m}(E + E_0). \quad (13.2.4)$$

Здесь введены обозначения

$$\omega_0^2 = \frac{\kappa_1}{m}, \quad \beta = \frac{\kappa_2}{m}.$$

Положим

$$x = x_0 + \xi(t), \quad (13.2.5)$$

где величина x_0 является решением уравнения

$$\omega_0^2x_0 + \beta x_0^2 = eE_0/m. \quad (13.2.6)$$

Учтём, что поправка βx_0^2 мала по сравнению с $\omega_0^2 x_0$. Тогда из последнего уравнения находим постоянную составляющую смещения электрона под действием постоянного поля E_0 в конденсаторе:

$$x_0 \approx \frac{eE_0}{m\omega_0^2}. \quad (13.2.7)$$

Дополнительное (к x_0) смещение ξ вызвано полем световой волны. Считая его малым, имеем

$$\beta x^2 = \beta(x_0 + \xi)^2 \approx \beta x_0^2 + 2\beta x_0 \xi. \quad (13.2.8)$$

Соответственно подстановка (13.2.5), (13.2.6) в (13.2.4) даёт уравнение для ξ :

$$\ddot{\xi} + 2\gamma\dot{\xi} + \tilde{\omega}_0^2 \xi = eE/m, \quad \tilde{\omega}_0^2 = \omega_0^2 + 2\beta x_0. \quad (13.2.9)$$

Таким образом, наличие электрического поля в конденсаторе приводит к изменению собственной частоты колебаний электрона:

$$\Delta\omega_0^2 = \tilde{\omega}_0^2 - \omega_0^2 = 2\beta x_0. \quad (13.2.10)$$

Полагая в (13.2.9) $E = E_1 e^{-i\omega t}$, находим смещение электрона:

$$\xi = \frac{eE/m}{\tilde{\omega}_0^2 - \omega^2 - 2i\gamma\omega} \quad (13.2.11)$$

и диэлектрическую проницаемость среды:

$$\varepsilon = 1 + \frac{\omega_p^2}{\tilde{\omega}_0^2 - \omega^2 - 2i\gamma\omega}. \quad (13.2.12)$$

Из приведённых формул следует, что индуцируемое полем конденсатора изменение показателя преломления равно

$$\Delta n = n|_{\tilde{\omega}_0^2} - n|_{\omega_0^2} = \frac{\partial n}{\partial(\omega_0^2)} \Delta(\omega_0^2) = -2\beta x_0 \frac{\partial n}{\partial(\omega^2)} = -bE_0, \quad (13.2.13)$$

где введено обозначение

$$b = \frac{\beta e}{m\omega\omega_0^2} \frac{dn}{d\omega}. \quad (13.2.14)$$

Мы рассмотрели случай, когда поляризация световой волны направлена параллельно электрическому полю в конденсаторе E_0 . Если же волна поляризована в перпендикулярном направлении, то согласно (13.2.3) поле E_0 не меняет показатель преломления.

Таким образом, в кристалле возникает эффект двойного лучепреломления.

13.3. Эффект Керра

В то время как эффект Поккельса наблюдается в кристаллах с выделенным направлением, в средах изначально изотропных также может на-

было бы возникновение двойного лучепреломления под действием внешнего электрического поля. Однако этот эффект не может быть линейным по полю вследствие центральной симметрии среды.

Явление было открыто шотландским физиком Джоном Керром в 1875 г. и названо *квадратичным электрооптическим эффектом*, или *эффектом Керра*.

Оказалось, что если до наложения поля среда была неполяризованной и изотропной, то разность показателей преломления обыкновенного и необыкновенного лучей является чётной функцией напряжённости поля E_0 . Это значит, что в разложении по степеням E_0 должны присутствовать члены лишь чётных порядков, начиная с квадратичного:

$$n_e - n_0 = B\lambda_0 E_0^2, \quad (13.3.1)$$

где λ_0 — длина волны света в вакууме. Коэффициент B зависит от свойств вещества и называется *постоянной Керра*. В этой формуле в правой части мы удержали только слагаемое с низшей степенью поля, поскольку все последующие слагаемые, как правило, малы по сравнению с ним.

Для большинства веществ $B > 0$. Это означает, что данные вещества во внешнем поле E_0 подобны положительным одноосным кристаллам. В качестве примера укажем, что при комнатной температуре на длине волны $\lambda = 589$ нм в нитробензоле ($C_6H_5NO_2$) $B = 2,2 \cdot 10^{-5}$ ед. СГСЭ, а для воды $B = 4,7 \cdot 10^{-7}$ ед. СГСЭ.

Пусть напряжение в конденсаторе $E_0 = 15\,000$ В/см = 50 ед. СГСЭ. Тогда на трассе луча $l = 5$ см разность фаз обыкновенного и необыкновенного лучей в нитробензоле $\Delta = kl(n_e - n_0) = 2\pi B l E_0^2$ составляет $\Delta\varphi = \pi/2$. Это значит, что такая ячейка может играть роль пластинки $\lambda/4$.

Эффект Керра, как и эффект Поккельса, практически безынерционный — установление анизотропии происходит за времена порядка $10^{-13} - 10^{-9}$ с.

Количественная теория эффекта Керра для газов была построена французским физиком Полем Ланжевеном в 1910 г. Предложенный им механизм эффекта состоит в следующем. Молекулы среды анизотропны, но ориентированы хаотически, вследствие чего поляризация среды отсутствует: $\bar{P} = 0$. Под действием внешнего электрического поля молекулы приобретают дипольный момент, который вследствие анизотропии, вообще говоря, не направлен по полю. Одновременно поле стремится ориентировать молекулы-диполи вдоль своего направления. При этом тепловые движения оказывают дезориентирующее действие. В результате уста-

навливается некоторая средняя ориентация молекул, определяющая степень анизотропии оптических свойств среды. Приведём количественное рассмотрение.

Как было сказано, молекулы среды анизотропны. Это значит, что дипольный момент, индуцируемый внешним полем, направлен вдоль оси молекулы (рис. 13.3.1). Найдём сначала распределение числа молекул по направлениям, устанавливающееся под действием постоянного электрического поля E_0 .

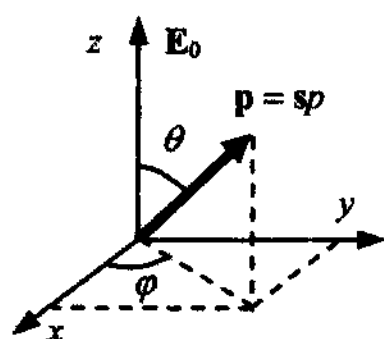


Рис. 13.3.1. Молекула во внешнем электрическом поле. Жирной стрелкой показан индуцированный полем дипольный момент молекулы

Обозначим поляризуемость молекулы как β . Тогда под действием внешнего поля она приобретает дипольный момент

$$\mathbf{p} = s\beta E_0 \cos \theta, \quad s^2 = 1. \quad (13.3.2)$$

Здесь s — единичный вектор в направлении оси молекулы. Для дальнейшего воспользуемся сферическими координатами:

$$s_x = \sin \theta \cos \varphi, \quad s_y = \sin \theta \sin \varphi, \quad s_z = \cos \theta. \quad (13.3.3)$$

Соответственно компоненты вектора $\mathbf{p}$ перепишем в виде

$$p_x = \beta E_0 s_x s_z, \quad p_y = \beta E_0 s_y s_z, \quad p_z = \beta E_0 s_z^2. \quad (13.3.4)$$

Поскольку оба направления вдоль оси молекулы эквивалентны, то проекция дипольного момента на направление поля E_0 неотрицательна: $p_z \geq 0$.

Учтём теперь, что поле не только формирует дипольный момент, но и оказывает на него ориентирующее действие. Для этого заметим, что потенциальная энергия ориентируемого квазиупругого диполя во внешнем поле равна

$$u = -\frac{1}{2} \mathbf{p} \mathbf{E}_0 = -\frac{1}{2} p_z E_0 = -\frac{1}{2} \beta E_0^2 s_z^2. \quad (13.3.5)$$

Действительно, она складывается из потенциальной энергии деформации молекулы (формирования диполя) и ориентации. Первая из них равна

$$u_{\text{деф}} = \frac{1}{2} \mathbf{p} \mathbf{E}_0.$$

Вторая энергия равна

$$u_{\text{ориент}} = -\mathbf{p} \mathbf{E}_0.$$

Сумма $u_{\text{деф}} + u_{\text{ориент}}$ даёт величину (13.3.5).

Имея это в виду, запишем распределение молекул по направлениям (распределение Больцмана):

$$dN = A \exp\left(-\frac{u}{kT}\right) d\Omega = A \exp\left(\frac{\mathbf{p}\mathbf{E}_0}{2kT}\right) d\Omega. \quad (13.3.6)$$

Здесь A — нормировочный множитель, $d\Omega = \sin\theta d\theta d\varphi$ — элемент телесного угла. Поскольку потенциальная энергия молекулы мала: $pE_0 \ll kT$, то можно разложить экспоненту по степеням u/kT :

$$dN = A \left(1 + \frac{a}{2} s_z^2\right) d\Omega. \quad (13.3.7)$$

Здесь и далее для краткости используем обозначение

$$a = \beta E_0^2 / kT, \quad a \ll 1. \quad (13.3.8)$$

Найдём нормировочный множитель A :

$$\begin{aligned} N &= A \int_{\theta=0}^{\pi} \left(1 + \frac{a}{2} \cos^2 \theta\right) \sin \theta d\theta \int_0^{2\pi} d\varphi = 4\pi A \left(1 + \frac{a}{6}\right) \Rightarrow \\ \Rightarrow A &= \frac{N}{4\pi} \left(1 - \frac{a}{6}\right). \end{aligned} \quad (13.3.9)$$

Следовательно, искомое распределение принимает вид

$$dN = N \left[1 + \frac{a}{2} \left(s_z^2 - \frac{1}{3}\right)\right] \frac{d\Omega}{4\pi}. \quad (13.3.10)$$

Теперь можно найти диэлектрическую проницаемость для электромагнитной волны. Поскольку волна имеет высокую частоту, то она не влияет на ориентацию молекул вследствие инерционности вращения. Поэтому рассматриваем поляризацию среды при заданном распределении молекул по направлениям, даваемом формулой (13.3.10).

Будем считать, что волна распространяется перпендикулярно полю в конденсаторе $\mathbf{E}_0$: $\mathbf{k} \perp \mathbf{E}_0$. Рассмотрим два случая:

- 1) волна поляризована вдоль $\mathbf{E}_0$,
- 2) волна поляризована перпендикулярно $\mathbf{E}_0$.

Пусть сначала $\mathbf{E} \parallel \mathbf{E}_0$. Это значит, что отлична от нуля лишь компонента E_z . Волна индуцирует в молекуле с осью, направленной по вектору $\mathbf{s}$, дипольный момент

$$\mathbf{p} = \beta(\mathbf{E}\mathbf{s})\mathbf{s}. \quad (13.3.11)$$

Соответственно имеем

$$p_z = \beta E_z s_z^2.$$

Вектор поляризации среды, индуцируемый полем волны, равен

$$P_z = \int p_z dN = \frac{1}{4\pi} \beta E \int_{\theta=0}^{\pi} \int_{\varphi=0}^{2\pi} \left[1 + \frac{a}{2} \left(s_z^2 - \frac{1}{3} \right) \right] s_z^2 \sin \theta d\theta d\varphi =$$

$$= \frac{1}{3} N \beta E_z \left(1 + \frac{2}{15} a \right) \equiv \alpha_{zz} E_z. \quad (13.3.12)$$

Полагая $E_z + 4\pi P_z = \epsilon_{zz} E_z$, находим диэлектрическую проницаемость:

$$\epsilon_{zz} = 1 + 4\pi \alpha_{zz} = 1 + \frac{4\pi}{3} N \beta + \frac{8\pi}{45} a. \quad (13.3.13)$$

Введём обозначения

$$\epsilon_{zz} = \epsilon_{zz}^{(0)} + \epsilon_{zz}^{(2)} E_0^2, \quad \epsilon_{zz}^{(0)} = 1 + \frac{4\pi}{3} N \beta, \quad \epsilon_{zz}^{(2)} = \frac{8\pi}{45} a. \quad (13.3.14)$$

Тогда для показателя преломления получаем выражение

$$n_z = \sqrt{\epsilon_{zz}} = n_z^{(0)} + n_z^{(2)} E_0^2, \quad (13.3.15)$$

где

$$n_z^{(0)} = \sqrt{\epsilon_{zz}^{(0)}} = 1 + \frac{2\pi}{3} N \beta, \quad n_z^{(2)} = \frac{\epsilon_{zz}^{(2)}}{2\sqrt{\epsilon_{zz}^{(0)}}} \approx \frac{4\pi}{45} a. \quad (13.3.16)$$

Пусть теперь поле волны E перпендикулярно полю в конденсаторе: $E \perp E_0$ и направлено по оси x (в обозначениях, указанных на рис. 13.3.3).

В этом случае молекулы в поле волны приобретают дипольный момент

$$\mathbf{p} = \beta(\mathbf{E}\mathbf{s})\mathbf{s} = \beta E_x s_x \mathbf{s},$$

$$p_x = \beta E_x s_x^2, \quad p_y = \beta E_x s_x s_y, \quad p_z = \beta E_x s_x s_z. \quad (13.3.17)$$

С учётом этого ищем компоненту вектора поляризации среды, параллельную напряжённости поля:

$$P_x = \int p_x dN = \int \beta E_x s_x^2 dN = N \beta E_x \int \left[1 + \frac{a}{2} \left(s_z^2 - \frac{1}{3} \right) \right] s_x^2 \frac{d\Omega}{4\pi}. \quad (13.3.18)$$

Найдём входящие сюда интегралы:

$$\int s_x^2 \frac{d\Omega}{4\pi} = \frac{1}{4\pi} \int_0^{\pi} \sin^2 \theta \cdot \sin \theta d\theta \int_0^{2\pi} \cos^2 \varphi d\varphi = \frac{1}{3},$$

$$\int s_x^2 s_z^2 \frac{d\Omega}{4\pi} = \int_0^{\pi} \sin^2 \theta \cos^2 \theta \cdot \sin \theta d\theta \int_0^{2\pi} \cos^2 \varphi d\varphi = \frac{1}{15}. \quad (13.3.19)$$

В результате получаем

$$P_x = \frac{1}{3} N \beta E_x \left(1 - \frac{a}{15} \right). \quad (13.3.20)$$

Найдём также компоненты P_y и P_z :

$$\begin{aligned} P_y &= \int p_y dN = A \int \beta E_x s_x s_y \left(1 + \frac{a}{2} s_z^2\right) d\Omega = 0, \\ P_z &= \int p_z dN = A \int \beta E_x s_x s_z \left(1 + \frac{a}{2} s_z^2\right) d\Omega = 0. \end{aligned} \quad (13.3.21)$$

Обращение в нуль этих величин связано с равенством нулю интегралов

$$\begin{aligned} \int_0^{2\pi} s_x s_y d\varphi &= \sin^2 \theta \int_0^{2\pi} \sin \varphi \cos \varphi d\varphi = 0, \\ \int_0^{2\pi} s_x s_z d\varphi &= \sin \theta \cos \theta \int_0^{2\pi} \cos \varphi d\varphi = 0. \end{aligned}$$

Полагая $E_x + 4\pi P_x = \epsilon_{xx} E_x$, из (13.3.20) находим диэлектрическую проницаемость:

$$\begin{aligned} \epsilon_{xx} &= 1 + \frac{4\pi}{3} N \beta \left(1 - \frac{a}{15}\right) = \epsilon_{xx}^{(0)} + \epsilon_{xx}^{(2)} E_0^2, \\ \epsilon_{xx}^{(0)} &= 1 + \frac{4\pi}{3} N \beta, \quad \epsilon_{xx}^{(2)} = -\frac{4\pi}{45} a. \end{aligned} \quad (13.3.22)$$

В результате приходим к следующему выражению для показателя преломления:

$$\begin{aligned} n_x &= \sqrt{\epsilon_{xx}} = n_x^{(0)} + n_x^{(2)} E_0^2, \\ n_x^{(0)} &= \sqrt{\epsilon_{xx}^{(0)}} = 1 + \frac{2\pi}{3} N \beta, \quad n_x^{(2)} = \frac{\epsilon_{xx}^{(2)}}{2\sqrt{\epsilon_{xx}^{(0)}}} \approx -\frac{2\pi}{45} a. \end{aligned} \quad (13.3.23)$$

(Далее мы увидим, что это значение n_x отвечает обыкновенному лучу.)

Нетрудно понять, что если бы вектор напряжённости поля был направлен вдоль оси y , то мы получили бы аналогичные формулы для компонент вектора поляризации:

$$P_x = 0, \quad P_y = \frac{1}{3} N \beta E_y \left(1 - \frac{a}{15}\right), \quad P_z = 0, \quad (13.3.24)$$

а для диэлектрической проницаемости — выражение

$$\epsilon_{yy} = \epsilon_{xx} = 1 + \frac{4\pi}{3} N \beta \left(1 - \frac{a}{15}\right). \quad (13.3.25)$$

Это значит, что оси выбранной системы координат являются главными, и в них тензор диэлектрической проницаемости имеет диагональный вид:

$$\epsilon = \begin{pmatrix} \epsilon_{xx} & 0 & 0 \\ 0 & \epsilon_{yy} & 0 \\ 0 & 0 & \epsilon_{zz} \end{pmatrix}. \quad (13.3.26)$$

При этом сравнение (13.3.14), (13.3.22) и (13.3.25) даёт

$$\epsilon_{xx} = \epsilon_{yy} \neq \epsilon_{zz}. \quad (13.3.27)$$

Это значит, что величина $n_x = \sqrt{\epsilon_{xx}}$ имеет смысл показателя преломления обыкновенного луча, а $n_z = \sqrt{\epsilon_{zz}}$ — необыкновенного.

Если учесть, что величина

$$n = 1 + \frac{2\pi}{3} N\beta \quad (\equiv n_x^{(0)} = n_z^{(0)}) \quad (13.3.28)$$

есть показатель преломления среды в отсутствие электрического поля E_0 , то формулы (13.3.15), (13.3.23) можно записать в виде

$$n_e - n = \frac{4\pi}{45} a, \quad n_o - n = -\frac{2\pi}{45} a. \quad (13.3.29)$$

Эти соотношения принято записывать следующим образом:

$$\frac{n_e - n}{n_o - n} = -2. \quad (13.3.30)$$

Данное равенство хорошо выполняется для большинства веществ.

Наконец, из (13.3.29) находим

$$n_e - n_o = \frac{2\pi}{15} a = \frac{n-1}{5} \frac{\beta E_0^2}{kT}. \quad (13.3.31)$$

Сравнение с (13.3.14) даёт явное выражение для постоянной Керра:

$$B = \frac{n-1}{5\lambda_0 kT} \beta. \quad (13.3.32)$$

Следует отметить, что рассмотренный выше ориентационный механизм появления анизотропии характерен лишь для газов. В жидкостях и твёрдых телах изменение поляризации часто вызывается деформацией электронного облака атомов и молекул подобно тому, как было показано в разделе 13.2, посвящённом эффекту Поккельса.

13.4. Эффект Фарадея

Помимо двойного лучепреломления анизотропные среды могут демонстрировать оптическую активность, проявляющуюся в изменении поляризации излучения по мере его распространения.

13.4.1. Естественная оптическая активность

Пусть свет линейно поляризован. Это значит, что его плоскость поляризации, то есть плоскость, содержащая векторы $\mathbf{k}$ и $\mathbf{E}$, остаётся неизменной. В 1811 г. Д. Араго, изучая прохождение света через пластинки кварца, вырезанные перпендикулярно оптической оси (рис. 13.4.1), обнаружил, что плоскость поляризации поворачивается. В 1815 г. Ж. Б. Био

(Франция) обнаружил, что плоскость поляризации могут вращать жидкости (скипидар, раствор сахара), пары органических веществ.

Вещества, способные вращать плоскость поляризации, называются *оптически активными*.

Часто используют следующую терминологию: вращение называют *положительным*, если оно происходит *по часовой стрелке* (и отвечает правой поляризации). Если же вращение происходит *против часовой стрелки* (и отвечает левой поляризации), то оно называется *отрицательным* (рис. 13.4.2).



Рис. 13.4.1. Свет распространяется в кристаллической пластинке, вырезанной перпендикулярно оптической оси

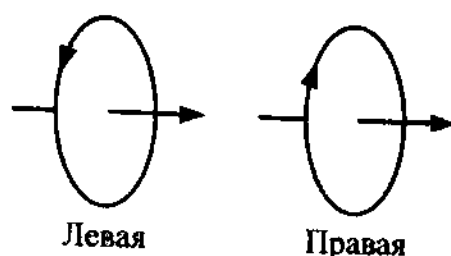


Рис. 13.4.2. Левая и правая круговые поляризации электромагнитной волны

Наблюдать вращение можно в схеме, показанной на рис. 13.4.3. В этой схеме свет от источника S проходит через николю (поляризатор) N_1 и становится линейно поляризованным. Далее он проходит через кристалл (или кювету с оптически активным веществом) K и затем пропускается через николю N_2 . Оптические оси николей взаимно перпендикулярны. Поэтому если плоскость поляризации не меняет направления, то через николю N_2 свет не проходит. Наличие оптической активности проявляется в том, что в зависимости от длины кристалла L интенсивность прошедшего через систему излучения меняется.

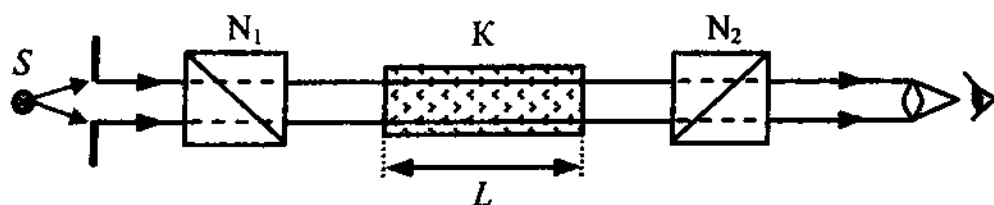


Рис. 13.4.3. Схема установки для наблюдения оптической активности

Исследования показали, что угол поворота зависит от длины пути волны L в веществе:

$$\varphi = \alpha L.$$

Величина α называется *вращательной постоянной*. Её значение различно у разных веществ. В частности, у кварца в красной области спектра

$\alpha \approx 15$ град/мм, в зелёной области $\alpha \approx 27$ град/мм и $\alpha \approx 51$ град/мм — в фиолетовой. Известны также две разновидности кварца: левовращающий и правовращающий. Для этих разновидностей постоянная α имеет противоположные знаки, но одинакова по абсолютной величине.

Направление вращения плоскости поляризации изменяется при изменении направления распространения света на обратное.

Первое объяснение вращения плоскости поляризации было предложено Френелем в 1817 г. Его идея состояла в том, что для волн с различной круговой поляризацией (рис. 13.4.2) скорость распространения различна.

Запишем волны с круговой поляризацией. Пусть волны распространяются вдоль оси z и им отвечают волновые векторы k_n и k_l (соответственно для правой и левой поляризаций). Тогда для *правой* поляризации имеем

$$\begin{aligned} E_x^{(n)} &= E_0 \cos(\omega t - k_n z), \\ E_y^{(n)} &= E_0 \cos\left(\omega t - k_n z + \frac{\pi}{2}\right), \end{aligned} \quad (13.4.1)$$

а для *левой* поляризации —

$$\begin{aligned} E_x^{(n)} &= E_0 \cos(\omega t - k_n z), \\ E_y^{(n)} &= E_0 \cos\left(\omega t - k_n z - \frac{\pi}{2}\right). \end{aligned} \quad (13.4.2)$$

Обозначим

$$\begin{aligned} k_n &= k + \alpha, \\ k_l &= k - \alpha. \end{aligned} \quad (13.4.3)$$

Суперпозиция волн (13.4.1) и (13.4.2) дает

$$\begin{aligned} E_x &= E_x^{(n)} + E_x^{(l)} = 2E_0 \cos(\alpha z) \cos(\omega t - kz), \\ E_y &= E_y^{(n)} + E_y^{(l)} = 2E_0 \sin(\alpha z) \cos(\omega t - kz). \end{aligned} \quad (13.4.4)$$

Очевидно, что в каждом сечении z это линейно поляризованная волна. Однако её плоскость поляризации на пути L поворачивается на угол

$$\varphi = \alpha L. \quad (13.4.5)$$

Свяжем угол вращения с показателями преломления для соответствующих волн. По определению фазовой скорости на основании (13.4.3) имеем

$$v_n = \frac{\omega}{k_n} = \frac{\omega}{k + \alpha}, \quad v_l = \frac{\omega}{k_l} = \frac{\omega}{k - \alpha}. \quad (13.4.6)$$

С другой стороны, фазовая скорость связана с показателем преломления соотношением $v = c/n$, так что

$$n_n = \frac{c}{v_n} = \frac{c}{\omega} (k + \alpha), \quad n_n = \frac{c}{v_n} = \frac{c}{\omega} (k - \alpha). \quad (13.4.7)$$

Следовательно,

$$n_n - n_n = \frac{2c}{\omega} \alpha, \quad (13.4.8)$$

или

$$\alpha = \frac{\omega}{2c} (n_n - n_n) = \frac{\pi}{\lambda} (n_n - n_n). \quad (13.4.9)$$

В соответствии с приведённым выше определением при $\alpha < 0$ мы имеем дело с положительным вращением, а при $\alpha > 0$ — с отрицательным.

Как видно из изложенного, оптической активностью могут обладать только среды, у которых молекулы подобны винтам (рис. 13.4.4): они должны обладать осью — выделенным направлением. Кроме того, молекула и её отражение в зеркале не могут быть совмещены.

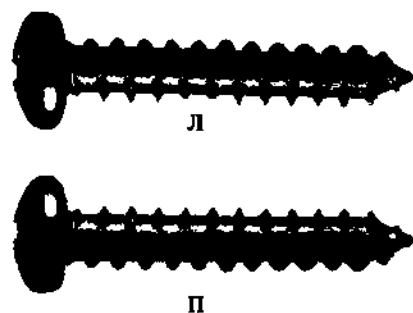


Рис. 13.4.4. Верхний винт соответствует левой поляризации, а нижний — правой

13.4.2. Эффект Фарадея

В 1845 году М. Фарадей обнаружил, что при распространении линейно поляризованного света через среду, находящуюся в магнитном поле, наблюдается вращение плоскости поляризации. В его экспериментах свет распространялся вдоль магнитного поля. В связи с этим данное явление называется также *продольным магнитооптическим эффектом*.

Схема опыта показана на рис. 13.4.5. Вещество помещено в постоянное продольное магнитное поле — в канале, просверленном в сердечнике электромагнита.

Было установлено, что угол поворота не зависит от направления распространения света и пропорционален величине поля и длине пути света в веществе

$$\varphi = VLB, \quad (13.4.10)$$

где V — характерная для вещества величина, называемая *постоянной Верде*.

Остановимся на классической теории эффекта Фарадея.

В отсутствие магнитного поля плоскость поляризации электромагнитной волны сохраняет своё положение. Это значит, что образующие её

волны с левой и правой поляризациями распространяются с одинаковой скоростью, им отвечают одинаковые показатели преломления. При наложении продольного магнитного поля симметрия нарушается, поскольку электроны среды приобретают дополнительный момент импульса с вполне определённым направлением, определяемым ориентацией магнитного поля (рис. 13.4.6).

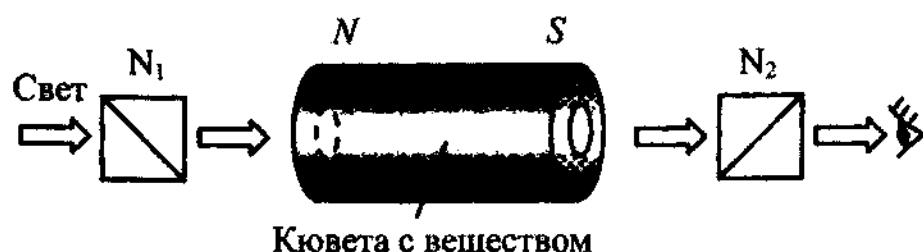


Рис. 13.4.5. Схема опыта по наблюдению эффекта Фарадея. Кювета с веществом находится в продольном магнитном поле. До и после кюветы свет пропускается через николи N_1 и N_2 с взаимно перпендикулярными осями

Запишем уравнение движения электрона в поле волны (E) и в постоянном магнитном поле (B):

$$m\ddot{\mathbf{r}} = -\kappa\mathbf{r} + e\mathbf{E} + \frac{e}{c}\dot{\mathbf{r}} \times \mathbf{B}. \quad (13.4.11)$$

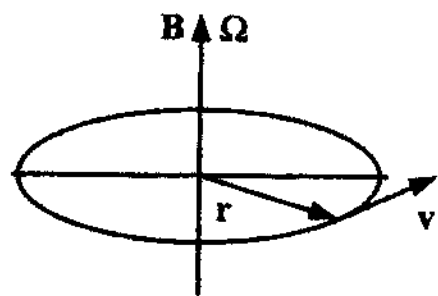


Рис. 13.4.6. Вращение электрона ($e < 0$) во внешнем магнитном поле

Мы пренебрегли трением, считая, что все характерные частоты находятся далеко от линий поглощения излучения веществом. Введём циклотронную частоту:

$$\Omega = -\frac{eB}{mc}, \quad (13.4.12)$$

дающую угловую скорость вращения электрона (рис. 13.4.6). Тогда уравнение (13.4.11) примет вид

$$\ddot{\mathbf{r}} + \Omega \times \dot{\mathbf{r}} + \omega_0^2 \mathbf{r} = e\mathbf{E}. \quad (13.4.13)$$

Пусть волна распространяется в направлении оси z , куда направлено также магнитное поле. При этом

$$\Omega_z = -\frac{eB}{mc} = \frac{|e|B}{mc} \equiv \Omega > 0. \quad (13.4.14)$$

Если волна имеет круговую поляризацию, то компоненты вектора напряжённости электрического поля равны

$$\begin{aligned} E_x &= E_0 \cos \omega t, \\ E_y &= \pm E_0 \sin \omega t. \end{aligned} \quad (13.4.15)$$

Во втором равенстве знак «+» отвечает левой поляризации, а знак «-» — правой. Таким образом, приходим к следующей системе уравнений:

$$\begin{aligned}\ddot{x} - \Omega\dot{y} + \omega_0^2 x &= E_0 \cos \omega t, \\ \ddot{y} + \Omega\dot{x} + \omega_0^2 y &= \pm E_0 \sin \omega t.\end{aligned}\quad (13.4.16)$$

Решение этой системы имеет вид

$$\begin{aligned}x &= \frac{e/m}{\omega_0^2 - \omega^2 \mp \Omega\omega} E_x, \\ y &= \frac{e/m}{\omega_0^2 - \omega^2 \mp \Omega\omega} E_y.\end{aligned}\quad (13.4.17)$$

Поскольку дипольный момент, приобретаемый электроном, равен $\mathbf{p} = e\mathbf{r}$, а вектор поляризации среды — $\mathbf{P} = N\mathbf{p} = \alpha\mathbf{E}$, где N — концентрация электронов в среде, то диэлектрическая проницаемость среды ($\epsilon = 1 + 4\pi\alpha$) оказывается равной

$$\epsilon_{\pm} = 1 + \frac{\omega_p^2}{\omega_0^2 - \omega^2 \mp \Omega\omega}. \quad (13.4.18)$$

Здесь величина $\omega_p = \sqrt{4\pi Ne^2/m}$ есть плазменная частота.

Из полученной формулы следует, что показатель преломления $n = \sqrt{\epsilon}$ зависит от того, какова поляризация волны (левая или правая).

Далее учтём, что частоты света порядка $\omega \sim 10^{15} \text{ с}^{-1}$, а циклотронная частота даже для сильных полей ($B \sim 10 \text{ Тл}$) составляет $\Omega \sim 10^{10} \text{ с}^{-1} \ll \omega$.

Поэтому $\omega^2 \pm \omega\Omega \approx (\omega \pm \Omega/2)^2$. Таким образом, для показателя преломления волн с левой и правой поляризациями в присутствии магнитного поля получаем выражение

$$n_{\pm} = \sqrt{1 + \frac{\omega_p^2}{\omega_0^2 - (\omega \pm \Omega/2)^2}}. \quad (13.4.19)$$

Пусть $n(\omega)$ — частотная зависимость показателя преломления среды. Тогда, как следует из (13.4.19),

$$n_{\text{л}} = n(\omega + \Omega/2), \quad n_{\text{п}} = n(\omega - \Omega/2), \quad n_{\text{л}} - n_{\text{п}} \approx \Omega \frac{dn}{d\omega}. \quad (13.4.20)$$

Воспользовавшись формулами (13.4.5), (13.4.9), находим угол поворота плоскости поляризации:

$$\varphi = \alpha L = \frac{\pi}{\lambda} \Omega \frac{dn}{d\omega} L = VLB, \quad (13.4.21)$$

причём для постоянной Верде получаем явное выражение:

$$V = \frac{\pi e}{\lambda mc} \frac{dn}{d\omega} = \frac{|e|}{2mc^2} \lambda \frac{dn}{d\lambda}. \quad (13.4.22)$$

13.5. О других способах создания оптической анизотропии

Помимо рассмотренных выше явлений, в которых наблюдалось возникновение оптической анизотропии, был обнаружен ещё ряд подобных явлений. Опишем кратко некоторые из них.

1. *Эффект Коттона–Мутона* — это возникновение двойного лучепреломления света в изотропном веществе, помещенном в поперечное магнитное поле (перпендикулярное световому лучу). Данное явление названо в честь французских физиков А. Коттона и Г. Мутона, подробно исследовавших его в 1907 г.

Ранее, в 1901 г., это явление обнаружили Дж. Керр (Шотландия) и (независимо) К. Майорана (Италия) в коллоидных растворах.

В отличие от эффекта Фарадея (раздел 13.4.2) при распространении света поперёк вектора индукции магнитного поля плоскость поляризации не меняется, но наблюдается обычное двойное лучепреломление. Причина возникновения анизотропии состоит в ориентации в магнитном поле поляризующихся молекул (агрегатов).

Для наблюдения эффекта Коттона–Мутона образец из изотропного вещества помещают между полюсами сильного электромагнита. Через образец пропускают монохроматический свет, линейно поляризованный в плоскости, составляющей с направлением магнитного поля угол в 45° . Вещество в магнитном поле становится оптически анизотропным с осью, параллельной магнитному полю H . В результате проходящий свет, распространяясь в виде обыкновенной и необыкновенной волн, превращается в эллиптически поляризованный. Разность показателей преломления обыкновенного n_o и необыкновенного n_e лучей равна

$$n_e - n_o = CH^2\lambda, \quad (13.5.1)$$

где λ — длина волны света.

Входящая в (13.5.1) величина C зависит от типа вещества и называется *постоянной Коттона–Мутона*. Она обратно пропорциональна абсолютной температуре T и, как правило, очень мала. Приведём значения величины C для некоторых жидкостей при $\lambda = 546$ нм и $T = 18^\circ\text{C}$:

ацетон: $C = 0,5 \cdot 10^{-13} \text{ см} \cdot \text{г} \cdot \text{с}^{-2}$,

бензол: $C = 7 \cdot 10^{-13} \text{ см} \cdot \text{г} \cdot \text{с}^{-2}$,

нитробензол: $C = 27 \cdot 10^{-13} \text{ см} \cdot \text{г} \cdot \text{с}^{-2}$.

Аномально большие значения C обнаружены в жидких кристаллах и в коллоидных растворах (от 10^{-8} до 10^{-10}).

С помощью эффекта Коттона–Мутона удаётся получить информацию о структуре молекул, образовании межмолекулярных агрегатов и подвижности молекул.

2. *Магнитооптический эффект Керра* — это эффект, состоящий в том, что при отражении линейно поляризованного света от поверхности намагниченного материала наблюдается вращение плоскости поляризации света, причём свет становится эллиптически поляризованным. Эффект линейен, то есть пропорционален намагниченности материала.

Впервые это явление наблюдал Дж. Керр в 1876 г. при отражении света от полюса железного магнита. Эффект, наблюдающийся в данной геометрии, получил название *полярного эффекта Керра* (рис. 13.5.1). Поворот плоскости поляризации отсутствует в случае непоглощающих сред.

В 1878 г. Керр обнаружил вращение плоскости поляризации при отражении от поверхности, намагниченной в плоскости распространения света. В такой геометрии, когда плоскость падения параллельна намагниченности, явление известно как *продольный эффект Керра* (рис. 13.5.1). Наиболее сильный эффект наблюдается при больших углах падения и отсутствует при нормальном падении.

В 1896 г. Питер Зееман (Нидерланды) открыл *поперечный эффект Керра*, когда вектор намагниченности перпендикулярен плоскости падения света и параллелен поверхности образца (рис. 13.5.1). Этот эффект проявляется в изменении коэффициента отражения под действием намагниченности и, как следствие, в изменении интенсивности света и сдвиге фазы линейно-поляризованного света. Данный эффект имеет место только для поглощающих материалов.

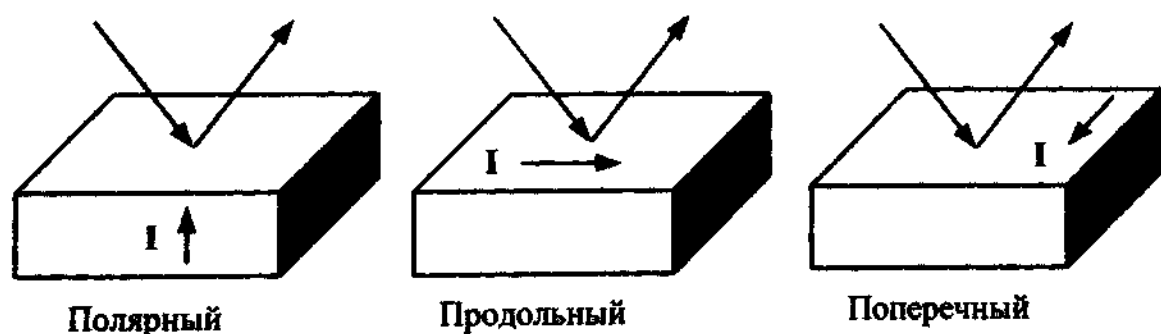


Рис. 13.5.1. Три типа эффекта Керра в зависимости от взаимной ориентации намагниченности образца (I), направления распространения световой волны и нормали к поверхности образца. Плоскость падения света совпадает с плоскостью рисунка

Заметим, что механизм эффекта Керра сходен с тем, что вызывает эффект Фарадея (изменение поляризации прошедшего через намагниченный материал света).

ГЛАВА 14

НЕКОТОРЫЕ ВОПРОСЫ НЕЛИНЕЙНОЙ ОПТИКИ

В предыдущей главе мы видели, что оптические свойства среды могут меняться под действием постоянных электрических или магнитных полей, что приводит к изменению условий распространения в этой среде электромагнитных волн. Это имеет место, например, в эффектах Керра, Поккельса и Фарадея. В этих эффектах интенсивность самих волн невелика, их влияние на свойства среды мало. Однако если интенсивность волны достаточно велика, то она сама может изменять среду, подстраивать «под себя», что приводит к ряду наблюдаемых явлений: генерации волн с частотой, отличной от частоты исходной волны, изменению траектории лучей и эффекту «линзы» — самофокусировке пучка излучения или, наоборот, возрастанию расходимости. Рассмотрим некоторые из нелинейно-оптических явлений.

14.1. Генерация второй гармоники. Оптическое выпрямление

14.1.1. Нелинейная поляризация

При распространении мощного оптического излучения в некоторых средах возможно появление волн с частотой, кратной частоте исходного излучения. Так, в 1961 г. П. Франкен с сотрудниками (США) наблюдали генерацию второй гармоники ($\lambda = 0,347$ мкм) при прохождении излучения рубинового лазера ($\lambda = 0,694$ мкм) через пластинку кварца. В этих опытах эффективность преобразования была очень низкой — составляла $\sim 10^{-8}$ от мощности исходного излучения. Позже, около 1963 года, были достигнуты высокие КПД преобразования — до 20–30%.

Рассмотрим природу данного явления. В твёрдых телах наиболее распространённым механизмом является нелинейный отклик атомного (или молекулярного) осциллятора на внешнее поле.

Под действием электромагнитной волны электроны вещества двигаются в соответствии с уравнением

$$m\ddot{r} = F_{\text{возвр}} + eE. \quad (14.1.1)$$

Здесь учтена только сила, создаваемая электрическим полем волны E , поскольку магнитное поле создаёт силу в v/c раз меньшую (v — скорость электрона). Кроме того, мы пренебрегли трением электронов, предпола-

гая, что потери энергии малы. В линейно поляризованной электромагнитной волне напряжённость электрического поля меняется со временем по закону

$$E = E_0 \cos(\omega t + \varphi_0). \quad (14.1.2)$$

Фаза φ_0 зависит от положения электрона. Ввиду малости амплитуды колебаний для каждого отдельного электрона эту фазу можно считать неизменной во времени, и мы её сейчас выписывать не будем.

Возвращающая сила стремится удержать электрон на определённом расстоянии от положения равновесия, причём $F_{\text{возвр}} = -\text{grad} U(r)$. Будем рассматривать смещение электрона вдоль одной оси. Если x — смещение x относительно равновесного положения ($x = 0$), то потенциальную энергию можно записать в виде ряда

$$U(x) = \frac{1}{2} \kappa_1 x^2 + \frac{1}{3} \kappa_2 x^3 + \dots \quad (14.1.3)$$

Мы провели разложение $U(x)$ до величин $\sim x^3$, которые начинают играть заметную роль, когда смещение электрона оказывается не слишком малым. Поскольку при этом

$$F_{\text{возвр}} = -dU/dx = -\kappa_1 x - \kappa_2 x^2, \quad (14.1.4)$$

то колебания электрона при отходе от положения $x = 0$ оказываются ангармоническими. Это имеет место не только в случае свободных колебаний, но и тогда, когда колебания возникают под действием гармонической вынуждающей силы. Второй случай реализуется, в частности, если в веществе распространяется интенсивная монохроматическая волна.

В указанных предположениях одномерное движение электрона описывается уравнением

$$\ddot{x} + \omega_0^2 x + \mu x^2 = \frac{e}{m} E_0 \cos \omega t, \quad (14.1.5)$$

где использованы обозначения

$$\omega_0^2 = \kappa_1/m, \quad \mu = \kappa_2/m. \quad (14.1.6)$$

Считая параметр нелинейности μ малым, будем решать данное уравнение методом последовательных приближений.

1-е приближение

Полагая в (14.1.5) $\mu = 0$, из соотношения

$$\ddot{x}_1 + \omega_0^2 x_1 = \frac{e}{m} E_0 \cos \omega t \quad (14.1.7)$$

находим решение низшего приближения:

$$x_1 = x_0 \cos \omega t, \quad x_0 = \frac{eE_0}{m} g(\omega), \quad g(\omega) = \frac{1}{\omega_0^2 - \omega^2}. \quad (14.1.8)$$

2-е приближение

В этом приближении полагаем

$$x = x_1 + x_2, \quad |x_2| \ll |x_1|, \quad \mu x^2 \approx \mu x_1^2. \quad (14.1.9)$$

Соответственно уравнение движения электрона (14.1.5) примет вид

$$\ddot{x}_2 + \omega_0^2 x_2 = -\mu x_1^2. \quad (14.1.10)$$

Далее учтём, что

$$x_1^2 = \frac{1}{2} x_0^2 (1 + \cos 2\omega t). \quad (14.1.11)$$

Это значит, что вынуждающая сила $(-\mu x_1^2)$ в (14.1.10) содержит два слагаемых:

$$\begin{aligned} F &= -\mu x_1^2 = F_1 + F_2, \\ F_1 &= -\frac{1}{2} \mu x_0^2, \quad F_2 = -\frac{1}{2} \mu x_0^2 \cos 2\omega t. \end{aligned} \quad (14.1.12)$$

Из них первое (F_1) является постоянным, а второе (F_2) — осциллирующим с частотой 2ω . Соответственно решение уравнения (14.1.10) можно записать в виде

$$\begin{aligned} x_2 &= x_2^{(0)} + x_2^{(2)} \cos 2\omega t, \\ x_2^{(0)} &= -\frac{1}{2} \mu x_0^2 g(0), \quad x_2^{(2)} = -\frac{1}{2} \mu x_0^2 g(2\omega). \end{aligned} \quad (14.1.13)$$

Вводя вектор поляризации среды $P = Nex$, получаем

$$P = P_{\text{л}} + P_{\text{н.л.}}, \quad (14.1.14)$$

где линейная и нелинейная поляризации равны

$$\begin{aligned} P_{\text{л}} &= Nex_1 = \alpha_1 E, \\ P_{\text{н.л.}} &= Nex_2 = P_0 + P_2, \\ P_0 &= \frac{1}{2} \alpha_0 E_0^2, \quad P_2 = \frac{1}{2} \alpha_2 E_0^2 \cos(2\omega t) \end{aligned} \quad (14.1.15)$$

(N — концентрация электронов в среде). Заметим, что линейная (α_1) и нелинейные (квадратичные: α_0, α_2) восприимчивости зависят от частоты, что учитывается множителями $g(\omega)$ и $g(2\omega)$ в (14.1.8) и (14.1.13):

$$\alpha_1 = \frac{Ne^2}{m} g(\omega), \quad \alpha_0 = -\frac{\mu Ne^3}{m^2} g^2(\omega) g(0), \quad \alpha_2 = -\frac{\mu Ne^3}{m^2} g^2(\omega) g(2\omega). \quad (14.1.16)$$

Для наблюдения квадратичных эффектов необходимо, чтобы чётные восприимчивости α_0, α_2 были отличны от нуля. Но это возможно, только если среда не изотропна. Действительно, как видно из (14.1.15), при отражении $E_0 \rightarrow -E_0$ знак составляющей $P_{н.л.}$ вектора поляризации не меняется. В изотропной же среде каждая точка может рассматриваться как центр симметрии, и изменение знака E влечёт изменение знака P . Другими словами, необходимо, чтобы существовала *полярная ось* с неэквивалентными началом и концом: « $\rightarrow$ ».

14.1.2. Оптическое «выпрямление»

Как следует из (14.1.15), слагаемое P_0 в поляризации не зависит от времени. Это означает, что при распространении излучения в нелинейном кристалле возникают неизменные во времени поляризационные заряды, приводящие к появлению постоянного электрического поля в направлении, перпендикулярном направлению распространения (волновому вектору). Это поле можно зарегистрировать, поместив кристалл между пластинами конденсатора (рис. 14.1.1). Время существования квазистационарной разности потенциалов между пластинами определяется длительностью светового импульса (но не частотой колебаний световой волны).

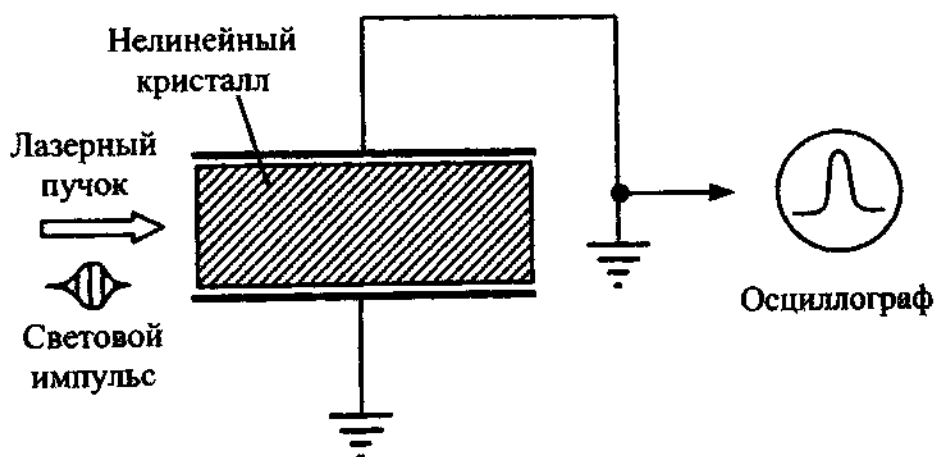


Рис. 14.1.1. Схема установки для наблюдения оптического выпрямления

14.1.3. Генерация второй гармоники

Слагаемое P_2 в (14.1.15) даёт вклад в нелинейную поляризацию, осциллирующую с частотой 2ω . Это проявляется в том, что в кристалле возникает электромагнитное поле, осциллирующее с частотой, вдвое превышающей частоту исходной волны. Описанный эффект называется *генерацией второй гармоники*.

Рассмотрим данное явление подробнее.

Запишем волновое уравнение для электрического поля, следующее из уравнений Максвелла (считая магнитную проницаемость среды $\mu = 1$):

$$\frac{1}{c^2} \frac{\partial^2 \mathbf{D}}{\partial t^2} = \Delta \mathbf{E}, \quad \mathbf{D} = \mathbf{E} + 4\pi \mathbf{P}, \quad (14.1.17)$$

или

$$\Delta \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{4\pi}{c^2} \frac{\partial^2 \mathbf{P}}{\partial t^2}. \quad (14.1.18)$$

Пусть электрическое поле волны направлено по оси кристалла, вдоль которой электроны могут совершать колебания. Волновой вектор при этом перпендикулярен данной оси (и направлен вдоль оси z). Согласно сказанному в предыдущем разделе вектор поляризации с учётом квадратичных поправок равен

$$\begin{aligned} P &= P_{\text{л}} + P_{\text{н.л.}}, \\ P_{\text{л}} &= \alpha_1 E, \quad P_{\text{н.л.}} = \alpha_2 E^2. \end{aligned} \quad (14.1.19)$$

В слагаемом $P_{\text{н.л.}}$, рассматривая лишь переменную часть E^2 , мы оставили поляризуемость α_2 (в обозначениях (14.1.15), (14.1.16)). Обозначим

$$\varepsilon_0 = 1 + 4\pi\alpha_1, \quad \varepsilon_0 = \varepsilon_0(\omega). \quad (14.1.20)$$

Это позволяет переписать равенство (14.1.17) в виде

$$\frac{\partial^2 E}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 (\varepsilon_0 E)}{\partial t^2} = \frac{4\pi}{c^2} \frac{\partial^2 P_{\text{н.л.}}}{\partial t^2}. \quad (14.1.21)$$

Необходимо иметь в виду, что коэффициент ε_0 в этом уравнении есть функция частоты: $\varepsilon_0 = \varepsilon_0(\omega)$.

Решаем это уравнение методом последовательных приближений.

1-е приближение. Отбрасывая нелинейное слагаемое в (14.1.21), имеем

$$\frac{\partial^2 E_1}{\partial z^2} - \frac{\varepsilon_0}{c^2} \frac{\partial^2 E_1}{\partial t^2} = 0 \Rightarrow E_1(z, t) = E_0 \cos(kz - \omega t), \quad (14.1.22)$$

где $k = \frac{\omega}{c} \sqrt{\varepsilon_0}$.

Решение $E_1(z, t)$ описывает волну, распространяющуюся в среде в отсутствие нелинейной поляризации.

2-е приближение. В правой части (14.1.21) учтём малость нелинейной поляризации и связанную с этим малость поправки к напряжённости поля. Полагая $E = E_1 + E_2$, $|E_2| \ll |E_1|$, имеем

$$P_{\text{н.л.}} = \alpha_2 E^2 = \alpha_2 (E_1 + E_2)^2 \approx \alpha_2 E_1^2.$$

Это даёт

$$\frac{\partial^2 E_2}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 (\epsilon_0 E_2)}{\partial t^2} = \frac{4\pi\alpha_2}{c^2} \frac{\partial^2 (E_1^2)}{\partial t^2}. \quad (14.1.23)$$

Подстановка сюда величины $E_1(z, t) = E_0 \cos(kz - \omega t)$ приводит к уравнению

$$\frac{\partial^2 E_2}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 (\epsilon_0 E_2)}{\partial t^2} = -\frac{8\pi\alpha_2 \omega^2}{c^2} E_0^2 \cos(2kz - 2\omega t). \quad (14.1.24)$$

Для дальнейшего удобно воспользоваться комплексным представлением:

$$\cos(2kz - 2\omega t) \rightarrow e^{i(2kz - 2\omega t)}.$$

Соответственно вместо (14.1.24) будем иметь дело с уравнением

$$\frac{\partial^2 E_2}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 (\epsilon_0 E_2)}{\partial t^2} = -\frac{8\pi\alpha_2 \omega^2}{c^2} E_0^2 e^{i(2kz - 2\omega t)}. \quad (14.1.25)$$

Таким образом, источником поля E_2 являются колебания поляризации $P_{нл.}$ с частотой 2ω . Последнее позволяет искать решение уравнения (14.1.25) в виде гармонической функции времени с частотой 2ω . Заменяя

$$\frac{\partial^2 E_2}{\partial t^2} = -(2\omega)^2 E_2, \quad (14.1.26)$$

получаем вместо (14.1.25) следующее равенство:

$$\frac{\partial^2 E_2}{\partial z^2} + k_2^2 E_2 = -\frac{8\pi\alpha_2 \omega^2}{c^2} E_0^2 e^{i(2kz - 2\omega t)}. \quad (14.1.27)$$

Здесь волновое число k_2 для волны с удвоенной частотой $\omega_2 = 2\omega$ определено соотношением

$$k_2 = \frac{\omega_2}{c} \sqrt{\epsilon_0(\omega_2)}. \quad (14.1.28)$$

Поскольку нелинейная поляризуемость α_2 мала, для решения уравнения (14.1.27) можно применить метод медленно меняющихся амплитуд. Если бы оказалось $\alpha_2 = 0$, то решение имело бы вид

$$E_2 \sim e^{i(k_2 z - \omega_2 t)} = e^{i(k_2 z - 2\omega t)}.$$

Соответственно при $\alpha_2 \neq 0$ положим

$$E_2(z, t) = A(z) e^{i(k_2 z - \omega_2 t)}, \quad (14.1.29)$$

где $A(z)$ — медленно меняющаяся функция (по сравнению с $e^{ik_2 z}$). Поскольку

$$\frac{\partial^2 E_2}{\partial z^2} = -k_2^2 A e^{i(k_2 z - \omega_2 t)} + 2ik_2 \frac{dA}{dz} e^{i(k_2 z - \omega_2 t)} + \frac{d^2 A}{dz^2} e^{i(k_2 z - \omega_2 t)},$$

то пренебрегая производной $d^2 A/dz^2$ по сравнению с $k_2 dA/dz$, получаем уравнение

$$2ik_2 \frac{dA}{dz} = -\frac{8\pi\alpha_2\omega^2}{c^2} E_0^2 e^{i(2k-k_2)z}. \quad (14.1.30)$$

Здесь выделяются два случая.

1. Пусть выполняются равенства

$$\varepsilon_0(\omega) = \varepsilon_0(2\omega), \quad 2k = k_2. \quad (14.1.31)$$

Сформулированные соотношения называются *условиями фазового синхронизма*. Они означают, что фазовые скорости волн с частотами ω и 2ω одинаковы:

$$v_1 = v_2, \quad v_1 = \frac{\omega}{k} = \frac{c}{\sqrt{\varepsilon_0(\omega)}}, \quad v_2 = \frac{\omega_2}{k_2} = \frac{c}{\sqrt{\varepsilon_0(2\omega)}}. \quad (14.1.32)$$

В рассматриваемой ситуации из (14.1.30) следует

$$2ik_2 \frac{dA}{dz} = -\frac{8\pi\alpha_2\omega^2}{c^2} E_0^2 \Rightarrow A(z) = \gamma z \quad (14.1.33)$$

(принято $A(0) = 0$). Здесь коэффициент γ даётся формулой

$$\gamma = i \frac{2\pi\alpha_2\omega^2}{c^2 k} E_0^2. \quad (14.1.34)$$

Таким образом, при выполнении условия (14.1.31) происходит перекачка энергии исходной волны (с частотой ω) во вторую гармонику. Напомним, что равенство (14.1.33) получено в предположении $|A| \ll E_0$. Очевидно, перекачка ограничивается тем, что во вторую гармонику может перейти лишь энергия исходной волны.

2. Пусть условия фазового синхронизма не выполняются: $k_2 \neq 2k$.

Обозначим

$$q = 2k - k_2 = \frac{2\omega}{c}(n_1 - n_2), \quad (14.1.35)$$

$$n_1 = \sqrt{\varepsilon_0(\omega)}, \quad n_2 = \sqrt{\varepsilon_0(2\omega)}.$$

Тогда из уравнения (14.1.28)

$$\frac{dA}{dz} = i \frac{4\pi\alpha_2\omega^2}{c^2 k_2} E_0^2 e^{iqz} \quad (14.1.36)$$

следует

$$A(z) = \frac{4\pi\alpha_2\omega^2}{c^2 k_2 q} E_0^2 (e^{iqz} - 1). \quad (14.1.37)$$

Это означает, что происходит периодическая перекачка энергии из первой гармоники во вторую и обратно:

$$I_2(z) = I_{2,\max} \sin^2(qz/2). \quad (14.1.38)$$

До тех пор, пока $qz \ll 1$, рост амплитуды второй гармоники происходит по закону (14.1.33), как и при выполнении условия фазового синхронизма. Расстояние, на котором $qz \sim \pi$, называется *длиной когерентности*:

$$l_{\text{ког}} \sim \frac{\pi}{q} = \frac{\pi}{2k - k_2} = \frac{\pi c}{2\omega(n_1 - n_2)} = \frac{\lambda_0}{4(n_1 - n_2)}. \quad (14.1.39)$$

Здесь $\lambda_0 = 2\pi c/\omega$ — длина волны в вакууме.

Например, для кварца на длине волны $\lambda_0 = 694,3$ нм $n_2 - n_1 \approx 0,025$. Поэтому длина когерентности оказывается малой: $l_{\text{ког}} \approx 10\lambda \approx 0,007$ мм. Условию фазового синхронизма можно удовлетворить, используя явление двойного лучепреломления, когда фазовые скорости обыкновенной и необыкновенной волн различны, а скорость необыкновенной волны зависит от направления распространения. В частности, используя одноосные кристаллы KDP (калий дигидрофосфат, KH_2PO_4), LiNbO_3 и некоторые другие, удалось преобразовать инфракрасное излучение лазера на неодимовом стекле ($\lambda = 1,064$ мкм) во вторую гармонику ($\lambda_2 = 0,532$ мкм) (линия в зелёной части спектра) с эффективностью порядка 30%.

Обсудим полученные результаты.

Согласно уравнению (14.1.25)

$$\frac{\partial^2 E_2}{\partial z^2} + k_2^2 E_2 = \frac{8\pi\alpha_2}{c^2} \frac{\partial^2 P_2}{\partial t^2}$$

вторую гармонику $E_2 = A \cos(k_2 z - 2\omega t)$ порождает волна поляризации

$$P_2 = \alpha_2 E_0^2 \cos(2kz - 2\omega t), \quad (14.1.40)$$

распространяющаяся с фазовой скоростью

$$v = \frac{2\omega}{2k} = \frac{c}{n(\omega)}.$$

При этом волны E_2 от каждого элемента среды возникают синхронно с волной поляризации, без дополнительных фазовых сдвигов. Пусть условие фазового синхронизма нарушено: $k_2 \neq 2k$. Тогда источник (волна поляризации) и индуцированные волны E_2 движутся с разными фазовыми

скоростями: $v = c/n(\omega)$ и $v_2 = c/n(2\omega)$, что приводит к появлению сдвига фаз $\Delta\varphi$ волн, возникших на расстоянии z друг от друга:

$$\Delta\varphi = (k_2 - 2k)z. \quad (14.1.41)$$

Пока этот сдвиг мал, происходит интерференционное усиление генерируемых волн, причём амплитуда второй гармоники растёт по закону $A = \gamma z$. На расстояниях же порядка $l_{\text{кор}}$, когда $\Delta\varphi \sim \pi$, сдвиг фаз возникающих волн оказывается большим, и интерференционное усиление пропадает.

14.2. Генерация третьей гармоники

В изотропных средах квадратичные восприимчивости обращаются в нуль. Однако могут оказаться ненулевыми восприимчивости нечётного порядка, в частности кубическая. В результате диэлектрическая проницаемость меняется под действием сильных полей как $\varepsilon = \varepsilon_0 + \varepsilon_2 E^2$. Возможны различные механизмы появления такой нелинейности. Одним из них является ориентационный механизм, характерный для газов и (в меньшей степени) для жидкостей, аналогичный рассмотренному в разделе 13.3, посвящённом эффекту Керра. Для твёрдых тел более характерны явления, связанные с деформацией электронного облака атомов или молекул под действием поля световой волны. Рассмотрим последний механизм подробнее.

В рассматриваемых условиях силу, возвращающую электрон в положение равновесия, можно записать в виде

$$F_{\text{возвр}} = -\kappa_1 x + \kappa_3 x^3, \quad (14.2.1)$$

где включена первая исчезающая поправка к зависимости $F_{\text{возвр}} = -\kappa_1 x$.

С учётом этого движение электрона в поле электромагнитной волны $E = E_0 \cos \omega t$ описывается уравнением, аналогичным (14.1.5):

$$\ddot{x} + \omega_0^2 x - \mu x^3 = \frac{eE_0}{m} \cos \omega t. \quad (14.2.2)$$

Здесь $\mu = \kappa_3/m$. Фазу волны, связанную с положением рассматриваемого электрона мы не выписываем. Имея в виду малость нелинейного слагаемого μx^3 , будем, как и в разделе 14.1, искать решение методом последовательных приближений.

В низшем, линейном приближении имеем

$$\ddot{x}_1 + \omega_0^2 x_1 = \frac{eE_0}{m} \cos \omega t \Rightarrow x_1 = x_{10} \cos \omega t. \quad (14.2.3)$$

Здесь использованы обозначения

$$x_{10} = \frac{eE_0}{m} g(\omega), \quad g(\omega) = \frac{1}{\omega_0^2 - \omega^2}. \quad (14.2.4)$$

Эта составляющая смещения приводит к поляризации среды $P_n = eNx_1$, пропорциональной полю:

$$P_n = \alpha_1 E_0 \cos \omega t, \quad \alpha_1 = \frac{Ne^2}{m} g(\omega). \quad (14.2.5)$$

Для получения следующего приближения положим $x = x_1 + x_3$, $\mu x^3 \approx \mu x_1^3$. Тогда из (14.2.2) получаем уравнение для поправки x_3 :

$$\ddot{x}_3 + \omega_0^2 x_3 = \mu x_1^3. \quad (14.2.6)$$

Учтём тождество

$$\cos^3 \varphi = \frac{3}{4} \cos \varphi + \frac{1}{4} \cos 3\varphi.$$

Это значит, что в правой части (14.2.6) имеются два слагаемых:

$$\mu x_1^3 = \frac{3}{4} \mu x_{10}^3 \cos \omega t + \frac{1}{4} \mu x_{10}^3 \cos 3\omega t. \quad (14.2.7)$$

Соответственно получаем

$$x_3(t) = x_3^{(1)}(t) + x_3^{(2)}, \quad (14.2.8)$$

$$x_3^{(1)} = \frac{3}{4} \mu x_{10}^3 g(\omega) \cos \omega t, \quad x_3^{(2)} = \frac{1}{4} \mu x_{10}^3 g(3\omega) \cos 3\omega t.$$

Поскольку вектор поляризации связан со смещением электрона: $P = Nex$, то из (14.2.8) следует, что в поляризации появляются составляющие двух типов:

$$P_{н.л.} = P_{н.л.}^{(1)} + P_{н.л.}^{(2)}, \quad (14.2.9)$$

$$P_{н.л.}^{(1)} = \chi_3 E_0^3 \cos \omega t, \quad P_{н.л.}^{(2)} = \frac{1}{3} \chi_3 E_0^3 \cos 3\omega t,$$

пропорциональные кубу амплитуды поля: $P_{н.л.} \sim E_0^3$.

Таким образом, в среде возникают колебания поляризации с частотой $\omega_3 = 3\omega$. В свою очередь эти колебания вызывают излучение на частоте ω_3 .

Рассмотренное явление называется *генерацией третьей гармоники*. Разумеется, для получения излучения с частотой 3ω должны выполняться требования фазового синхронизма, которые в данной ситуации выражаются равенствами

$$\omega_3 = 3\omega, \quad k_3 \approx 3k, \quad n(\omega_3) \approx n(\omega). \quad (14.2.10)$$

Практически эти требования во многих случаях существенно ограничивают эффективность преобразования излучения в третью гармонику. Тем не менее эффект был экспериментально обнаружен в 1962 г.: в излучении рубинового лазера ($\lambda = 694$ нм), прошедшего через кристаллы исландского шпата, зарегистрированы волны ультрафиолетовой части спектра с $\lambda = 231,3$ нм.

Нетрудно понять, что в подходящих средах можно ожидать генерацию высших гармоник (четвёртой, пятой и так далее). Эти гармоники действительно наблюдались. В частности, четвёртую гармонику обнаружил С. А. Ахманов (СССР) в 1974 г.

14.3. Самофокусировка излучения

В среде с кубической нелинейностью согласно сказанному в предыдущем разделе вектор поляризации в присутствии внешнего поля можно записать следующим образом:

$$P = (\chi_1 + \chi_3 E_0^2) E_0 \cos \omega t + \frac{1}{3} \chi_3 E_0^3 \cos 3\omega t. \quad (14.3.1)$$

Последнее слагаемое ответственно за генерацию третьей гармоники. Рассмотрим теперь явление, возникающее за счёт нелинейной поправки к поляризации на частоте исходного излучения. Будем считать, что перекачка энергии в волну с утроенной частотой пренебрежимо мала. Это допущение справедливо, если условие фазового синхронизма не выполняется: $n(3\omega) \neq n(\omega)$.

Введём диэлектрическую проницаемость для волны с неизменённой частотой обычным образом:

$$\begin{aligned} D &= E + 4\pi P = \varepsilon E, \\ \varepsilon &= 1 + 4\pi\chi_1 + 4\pi\chi_3 E_0^2. \end{aligned} \quad (14.3.2)$$

Поскольку нелинейная поправка $4\pi\chi_3 E_0^2$ к ε мала, то получаем показатель преломления:

$$n = \sqrt{\varepsilon} = n_0 + n_2 E_0^2, \quad (14.3.3)$$

где

$$n_0 = \sqrt{1 + 4\pi\chi_1}, \quad n_2 = 2\pi\chi_3 / n_0. \quad (14.3.4)$$

Слагаемое в (14.3.3), содержащее квадрат напряжённости поля, называют *керровской нелинейностью* (по аналогии с электрооптическим эффектом Керра).

Оказывается, что хотя слагаемое $n_2 E_0^2$ мало, но действие нелинейной поляризации может быть сильным. Проявление этой поправки зависит от знака коэффициента n_2 . Рассмотрим случай $n_2 > 0$. Это значит, что уве-

личение интенсивности $I = cn_0 E_0^2 / 8\pi$ ведёт к возрастанию показателя преломления n . В свою очередь это по закону Снеллиуса влечёт отклонение лучей в сторону большего показателя преломления и, следовательно, к дальнейшему росту n . В результате пучок должен сжиматься. Этот эффект называется *самофокусировкой*. Вместе с тем ему противодействует дифракция, вызывающая расходимость пучка. Рассмотрим данное явление подробнее.

Пусть исходный пучок выделен из плоской волны диафрагмой диаметром $2a$ (рис. 14.3.1). Показатель преломления среды вне пучка равен n_0 , а в пределах пучка (вблизи диафрагмы) $n = n_0 + n_2 E_0^2 > n_0$. Вследствие дифракции лучи расходятся в конусе с угловым радиусом

$$\theta_{\text{диф}} = 1,22 \frac{\lambda_0}{2an} \quad (14.3.5)$$

(λ_0 — длина волны в вакууме).

Крайний луч, падая на боковую поверхность пучка, испытывает преломление в соответствии с законом Снеллиуса:

$$n \sin \varphi = n_0 \sin \varphi_0. \quad (14.3.6)$$

Переходя от углов падения к углам скольжения (рис. 14.3.1), перепишем последнее равенство:

$$n \cos \theta = n_0 \cos \theta_0. \quad (14.3.7)$$

Луч не выходит из пучка, если $\theta_0 = 0$. Соответственно критический угол скольжения $\theta = \theta_{\text{кр}}$ определяется равенством

$$\cos \theta_{\text{кр}} = \frac{n_0}{n} = \frac{n_0}{n_0 + n_2 E_0^2}. \quad (14.3.8)$$

Все лучи, образующие с боковой поверхностью пучка угол $\theta < \theta_{\text{кр}}$, остаются в пучке, и в результате отражения направляются к оси.

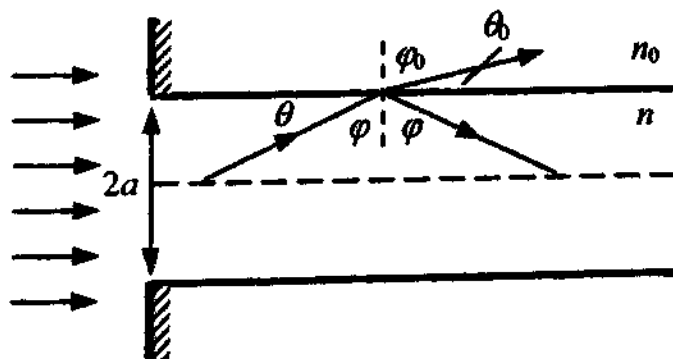


Рис. 14.3.1. К оценке порога самофокусировки. Излучение меняет показатель преломления среды, в результате чего лучи начинают отклоняться от исходного направления к оси пучка

Преобразуем выражение (14.3.8), имея в виду, что

$$n_2 E_0^2 \ll n_0, \quad \frac{n_0}{n_0 + n_2 E_0^2} \approx 1 - \frac{n_2}{n_0} E_0^2.$$

Поскольку при этом $\theta_{кр} \ll 1$, $\cos \theta_{кр} \approx 1 - \theta_{кр}^2/2$, то находим

$$\theta_{кр}^2 = 2n_2 E_0^2 / n_0. \quad (14.3.9)$$

Условие, при котором дифракционная расходимость компенсируется самофокусировкой, состоит в следующем:

$$\theta_{диф} \leq \theta_{кр}, \quad (14.3.10)$$

или

$$\left(1,22 \frac{\lambda_0}{2an}\right)^2 \leq \frac{2n_2}{n_0} E_0^2.$$

Переходя от амплитуды поля E_0 к интенсивности: $I = cn_0 E_0^2 / 8\pi$ и пренебрегая в полученном соотношении различием n и n_0 , получим критерий:

$$P = \pi a^2 I \geq P_{пор} = \frac{(1,22\lambda_0)^2 c}{64n_2}, \quad (14.3.11)$$

определяющий пороговую мощность, при превышении которой пучок начинает сжиматься. При этом дифракция уже не может остановить эффект самосжатия.

Оценка по этой формуле позволяет найти порог самофокусировки излучения рубинового лазера ($\lambda = 694,3$ нм) в сероуглероде (CS_2):

$$n_0 = 1,62, \quad n_2 = 2 \cdot 10^{-11} \text{ СГСЭ}, \quad P_{пор} \approx 20 \text{ кВт}. \quad (14.3.12)$$

В некоторых средах порог с большим значением коэффициента n_2 порог самофокусировки может снижаться на несколько порядков.

Поясним смысл критерия (14.3.11). Величина

$$L_{диф} \sim a / \theta_{диф}$$

указывает расстояния, на которых пучок становится расходящимся в конусе с угловым радиусом $\theta_{диф}$, и подобна фокусному расстоянию рассеивающей линзы. Величина же

$$L_{н.л.} = a / \theta_{кр}$$

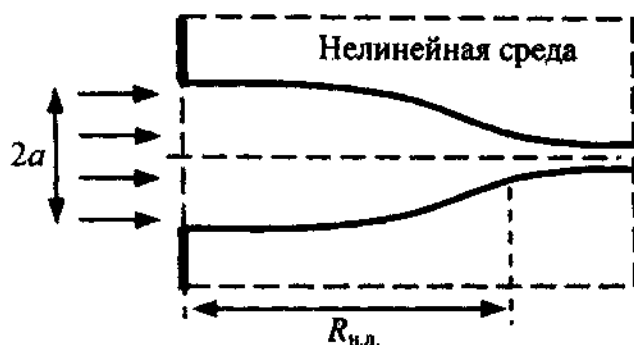
играет роль фокусного расстояния нелинейной собирающей линзы. Условие

$$L_{диф} \sim L_{н.л.} \sim a \sqrt{n_0 / (n_2 E_0^2)}$$

определяет порог самофокусировки: фокусирующее действие нелинейной среды преодолевает дефокусирующее действие дифракции.

В результате самофокусировки пучок сжимает в тонкую нить диаметром ~ 100 мкм — нелинейный диэлектрический волновод, в котором интенсивность достигает значений 4 ГВт/см^2 . Качественно явление проиллюстрировано на рис. 14.3.2. Возрастание интенсивности в области фокуса ограничивается эффектами нелинейного поглощения: многофотонным поглощением, перекачкой энергии в компоненты вынужденного рассеяния, оптическим пробоем и так далее.

Рис. 14.3.2. Самофокусировка пучка излучения и последующее распространение в канале-волноводе



Явление самофокусировки было предсказано в 1962 г. Г. А. Аскарьяном (СССР) (предложившим и сам термин «самофокусировка»). Впервые оно наблюдалось Н. Ф. Пилипецким и А. Р. Рустамовым (СССР) в 1965 г. В 1966 г. Р. Чиао, Е. Гармайр и Ч. Таунс (США) исследовали самофокусировку излучения рубинового лазера с мощностями $P = 10-100$ кВт и диаметром пучка 0,5 мм при прохождении кюветы с CS_2 . Ими была измерена пороговая мощность, оказавшаяся равной 25 кВт, причём пучок фокусировался на расстоянии $L \approx 12$ см.

При мощности излучения, значительно превышающей критическую мощность самофокусировки, пучок распадается на множество нитей с высокой плотностью энергии. Светящиеся нити были зарегистрированы в 1965 г. при самофокусировке фемтосекундного пучка излучения в жидкости и спустя 30 лет — в воздухе. Это явление получило название «филаментация».

Возникновение самофокусировки связано с возрастанием показателя преломления среды под действием поля волны. В ряде случаев, однако, показатель преломления может убывать. В твёрдых телах это зависит от знака поляризуемости χ_3 (и, следовательно, коэффициента n_2) в формулах (14.3.2) – (14.3.4). Убывание показателя преломления возможно и вследствие нагрева излучением, поскольку тепловое расширение вещества ведёт к уменьшению его плотности. Эта ситуация особенно характерна для газовых сред. В результате может наблюдаться явление, называемое *самодефокусировкой*: совместное действие нелинейной рефракции и дифракции приводит к увеличению расходимости пучка излучения.

Литература

1. Ахманов С.А., Никитин С.Ю. Физическая оптика. — М.: МГУ, 1998.
2. Сивухин Д.В. Общий курс физики. Оптика. Т. IV. — М.: Наука, 1980.
3. Кириченко Н.А. Электричество и магнетизм. — М.: МФТИ, 2011.
4. Савельев И.В. Курс общей физики. Оптика, атомная физика, физика атомного ядра и элементарных частиц. Т. III. — М.: Наука, 1971.
5. Кингсепт А.С., Локишин Г.Р., Ольхов О.А. Основы физики. Курс общей физики. Т. 1. — М.: Физматлит, 2001.
6. Максимычев А.В., Александров Д.А., Берюлёва Н.С. и др. Лабораторный практикум по общей физике: учеб пособие. Т. 2. Оптика/ под ред. А.В. Максимычева. — М.: МФТИ, 2014.
7. Локишин Г.Р. Основы радиооптики. — Долгопрудный: Издательский Дом «Интеллект», 2009.
8. Ландсберг Г.С. Оптика. — М.: Физматлит, 2003.
9. Борн М., Вольф Э. Основы оптики. — М.: Наука, 1973.
10. Борен К., Хафмен Д. Поглощение и рассеяние света малыми частицами. — М.: Мир, 1986.