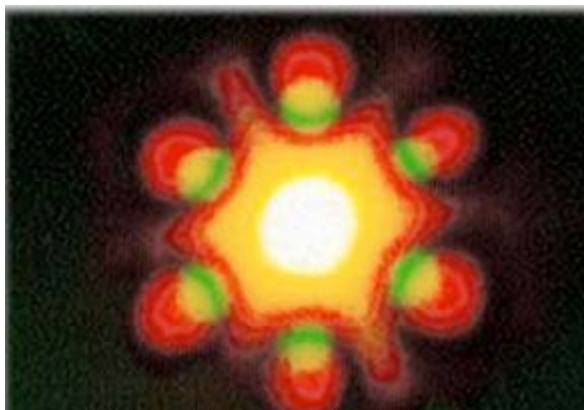
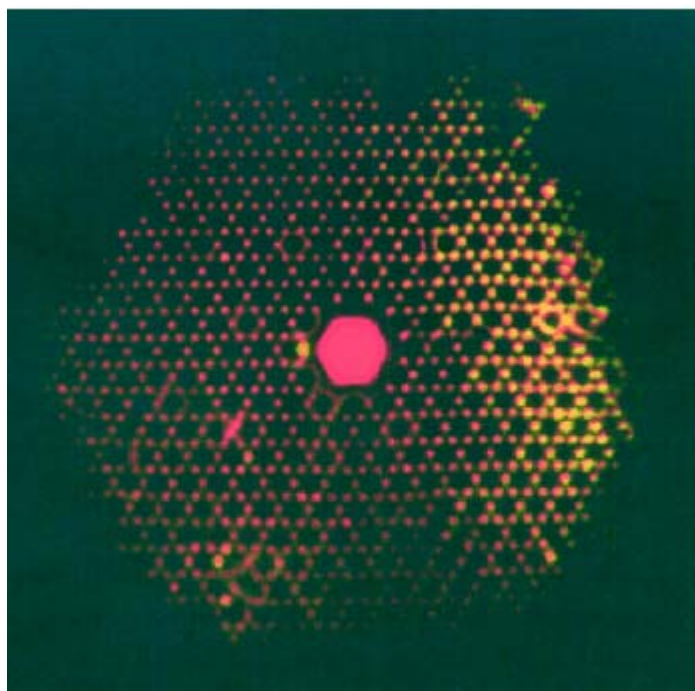


МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
КРАСНОЯРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
КАФЕДРА ОБЩЕЙ ФИЗИКИ
ИНСТИТУТ ФИЗИКИ им. Л.В. КИРЕНСКОГО СО РАН



В.Г.Архипкин, Г.С.Патрин

ЛЕКЦИИ ПО ОПТИКЕ



Красноярск 2006

СОДЕРЖАНИЕ

ЛЕКЦИЯ №1	4
1.1. Введение: предмет оптики. Краткий исторический обзор.	4
1.2. Распространение световые волн в вакууме.	8
ЛЕКЦИЯ №2	16
2.1. Поток энергии, интенсивность света.	16
2.2. Поляризация света.	20
ЛЕКЦИЯ №3	25
3.1. Квазимонохроматические и квазиплоские волны.	25
3.2. Спектральное разложение светового поля.	28
3.3. Световые волны в веществе	31
ЛЕКЦИЯ №4	36
4.1. Основные понятия фотометрии.	36
4.2. Световые пучки и импульсы: энергия, мощность, интенсивность	39
4.3. Геометрическая оптика.	40
ЛЕКЦИЯ №5	48
5.1. Свойства центрированных оптических систем.	48
5.2. Элементы матричной оптики.	52
5.3. Распространение светового луча в оптически неоднородной среде.	57
ЛЕКЦИЯ №6	59
6.1. Интерференция света.	59
6.2. Методы реализации интерференционной картины.	64
ЛЕКЦИЯ №7	70
7.1. Многолучевая интерференция, интерферометр Фабри-Перо.	70
7.2. Интерференция квазимонохроматического света.	72
8. ЛЕКЦИЯ №8	82
8.1. Дифракция как проявление волновой природы света.	82
8.2. Дифракция на круглом отверстии.	86
8.3. Дифракция на краю непрозрачного экрана.	90
9. ЛЕКЦИЯ №9	94
9.1. Приближение Френеля в теории дифракции.	94
9.2. Дифракция Фраунгофера как пространственное преобразование Фурье.	97
9.3. Дифракционные решетки.	101
9.4. Дифракция рентгеновских лучей	104
ЛЕКЦИЯ №10	106
10.1. Классическая электронная теория дисперсии Лоренца: показатель преломления среды, дисперсия и поглощение света в линейной изотропной среде.	106
10.2. Распространение светового импульса в диспергирующей среде, групповая скорость.	111
10.3. Рассеяние света в мутных средах.	113
ЛЕКЦИЯ №11	116
11.1. Отражение и преломление поляризованного света на границе раздела	116
11.2. Отражение света от поверхности металла.	122
ЛЕКЦИЯ №12	127

12.1. Оптическая анизотропия и основные эффекты кристаллооптики.	127
12.2. Структура световой волны в анизотропном кристалле.	128
12.3. Одноосный кристалл.	133
ЛЕКЦИЯ №13	138
13.1. Двойное лучепреломление света на границе с анизотропной средой.	138
13.2. Получение и анализ поляризованного света.	140
13.3. Интерференция поляризованных лучей	143
13.4. Наведенная анизотропия: электрооптические и магнитооптические эффекты.	144
13.5. Естественное вращение плоскости поляризации света.	147
ЛЕКЦИЯ №14	149
14.1. Линейная и нелинейная оптика.	149
14.2. Нелинейная поляризованность среды и нелинейно-оптические явления.	152
ЛЕКЦИЯ №15	158
15.1. Пространственное накопление нелинейно-оптических явлений, фазовое согласование в анизотропных кристаллах.	158
15.2. Самофокусировка света.	161
15.3. Параметрическая генерация света.	162

ЛЕКЦИЯ №1

Введение: предмет оптики. Краткий исторический обзор. Распространение световых волн в вакууме: электромагнитная природа света, уравнения Максвелла, волновое уравнение; плоская волна, сферическая волна; принцип суперпозиции; поперечность световой волны.

1.1. Введение: предмет оптики. Краткий исторический обзор. *Оптика* – это раздел физики, изучающий свойства и физическую природу света, его распространение в различных средах и взаимодействие с веществом, а также способы генерации и регистрации света. Это широко разветвленная область исследований, взаимодействующая со многими разделами физики, например, электродинамика, квантовая механика, физика твердого тела и др. Термин оптика происходит от греческого слова "optike" – наука о зрительных восприятиях, а само это слово происходит от греческого "optos" – видимый, зримый.

Оптическое излучение или *свет* представляет собой электромагнитные волны. Поэтому оптика – это часть общего учения об электромагнитном поле, которая называется *электродинамикой*. Под светом в настоящее время понимают не только видимое излучение, но и примыкающие к нему широкие области инфракрасного (ИК) и ультрафиолетового (УФ) излучения. В таком понимании оптический спектр занимает диапазон от границы ИК излучения $\lambda = 2$ мм ($\nu = 1,5 \times 10^{11}$ Гц) до коротковолновой границы УФ $\lambda = 10^{-6}$ см ($\nu = 3 \times 10^{16}$ Гц). Выделение указанной области обусловлено не столько близостью соответствующих участков спектра, но и, главным образом, сходством методов и приборов, применяемых для ее исследования и разработанных исторически, в основном, при изучении видимого света (линзы, зеркала, призмы, дифракционные решетки, интерференционные приборы и т.д.). *Собственно видимое излучение* занимает диапазон $\lambda = 400 - 760$ нм, УФ – $\lambda = 10 - 400$ нм, ИК – $\lambda = 760$ нм – 2 мм. Указанные границы достаточно условны.

Современная оптика подразделяется на:

- **геометрическую оптику**, которая, не рассматривая вопроса о природе света, исходит из *эмпирических законов его распространения* и использует представление о распространяющихся *независимо друг от друга световых лучах, преломляющихся и отражающихся на границе сред с разными оптическими свойствами и прямолинейных в оптически однородных средах*. Наиболее важное значение геометрическая оптика имеет для расчета и конструирования оптических приборов.

- **Физическую оптику**, которая рассматривает проблемы, связанные с природой света и световых явлений. Совокупность явлений, в которых проявляется волновая природа света, изучается в разделе физической оптики, называемый *волновой оптикой*, в рамках которой объясняются такие явления как дифракция, интерференция и поляризация света, распространение света в анизотропных средах и т.п. Ее основанием служат *уравнения Максвелла*. Свойства среды при этом характеризуются *диэлектрической ϵ и магнитной μ проницаемостями*, которые в общем случае зависят от *частоты света и ее волнового вектора* (частотная и пространственная дисперсия). Эти величины однозначно определяют *показатель преломления среды $n = \sqrt{\epsilon\mu}$* . Феноменологически *волновая оптика объясняет все законы геометрической оптики и границы ее применимости*. В отличие от геометрической оптики, она позволяет рассматривать процессы распространения света и тогда, когда *размеры систем, формирующих и рассеивающих световые пучки, меньше длины волны*.

• **Фотометрию** – раздел физической оптики, в котором рассматриваются *энергетические характеристики светового излучения в процессах его испускания, распространения и взаимодействия с веществом*. Ряд задач фотометрии решается с учетом особенностей и закономерностей восприятия света человеческим глазом (интересно отметить, что 90% информации человек получает с помощью глаз, т.е. зрения). Изучение закономерностей восприятия света человеческим глазом занимается *физиологическая оптика*, которая исследует механизмы зрения, и, в конечном счете, смыкается с биофизикой и психологией.

• **Кристалло- и металлооптику**, изучающую *связь диэлектрической ϵ и магнитной μ проницаемостей с молекулярной и кристаллической структурой вещества*. Здесь находят объяснение оптические явления в рассеивающих и анизотропных средах, вблизи границы разделов сред с разными оптическими свойствами, зависимость оптических свойств вещества от длины волны света (*дисперсия*), влияния на них давления, температуры, звука, электрических и магнитных полей и др.

• **Нелинейная оптика**, которая изучает *зависимость оптических явлений от интенсивности световых потоков*. Отметим, что изобретение лазеров, значительно раздвинула границы применимости оптики.

• Хорошо описывая распространение света в материальных средах, волновая оптика не смогла удовлетворительно объяснить процессы его испускания и поглощения. Явления, в которых при взаимодействии света и вещества проявляются *квантовые свойства атомов и молекул*, а также *квантовая природа света*, изучаются **квантовой оптикой**, с элементами которой вы познакомитесь в курсе *квантовой и атомной физики*. Здесь лишь отметим, что использование *квантовых свойств микрообъектов* привело к созданию *мазеров и лазеров* – генераторов электромагнитного излучения с уникальными свойствами, которые произвели революцию практически во всех областях физики, в том числе и оптики, а также во многих областях техники и технологии.

Все разделы оптики имеют многочисленные практические применения, существенно расширившиеся с изобретением лазеров. Перечислим лишь некоторые из них, которые развиваются на основе использования достижений лазерной техники: *лазерная биология и медицина, лазерная фотохимия, лазерное разделение изотопов, лазерный управляемый термоядерный синтез, оптоэлектроника, интегральная оптика, оптические компьютеры, адаптивная оптика, кабельная оптическая видеосвязь, оптическая память, оптическая локация, дальнометрия, дистанционный контроль окружающей среды, лазерная обработка материалов (резка, сварка, упрочение), лазерные микротехнологии* и др.

Кратко остановимся на **истории развития оптики**. Оптика – одна из самых древнейших наук, тесно связанная с потребностями практики. Она восходит к самым древним цивилизациям. Например, *прямолинейное распространение света* было известно в Месопотамии (5 тыс. лет до н.э.) и использовалось еще в древнем Египте при строительных работах. **Пифагор** (6 век до н.э.) высказал мысль, что тела видны благодаря испускаемым ими частицами. **Аристотель** (5 век до н.э.) считал, что свет есть возбуждение среды, находящейся между объектом и глазом. Он также занимался атмосферной оптикой, а радугу объяснял отражением света каплями воды. Школа **Платона** сформулировала два важнейших закона геометрической оптики – *прямолинейность лучей света* и *равенство углов падения и отражения*. **Евклид** (3 век до н.э.) в своих трактатах рассматривал возникновение изображений при отражении от зеркал.

Главный вклад греков, явившийся первым шагом в развитии оптики как науки, состоит не в их гипотезах и догадках о природе света, а в том, что они нашли законы его прямолинейного распространения и отражения (*катоптрика*) и умели ими пользоваться.

Следующий важный шаг – понимание законов преломления (*диоптрика*), был сделан много веков спустя. В средние века хорошо были известны эмпирические правила построения изображений, даваемых линзами, начало развиваться искусство изготовления линз. В XIII веке были изобретены очки, около 1590 г. – двухлинзовый микроскоп (**З.Янсен**). В 1609 г. **Галилей** построил первый телескоп, с помощью которого сделал ряд астрономических открытий. Точные законы преломления света были экспериментально установлены лишь около 1620 г. (**В.Снелль**, **Р.Декарт**). **Декарт** в 1637 г. опубликовал свою знаменитую "Диоптрику". Около 1660 г. был сформулирован **П.Ферма принцип Ферма**. Этим был завершен фундамент построения геометрической оптики.

Дальнейшее развитие оптики связано с открытием явлений *дифракции* (1665 г., **Ф.Гримальди**) и *интерференции* света, и *двойного лучепреломления* (1669 г., **Э.Бартоллини**), которое не поддавалось истолкованию в рамках геометрической оптики, а также с работами **И.Ньютона**, **Р.Гука** и **Х.Гюйгенса**.

И.Ньютон (1643-1727) обращал большое внимание на периодичность световых явлений и допускал возможность интерференции, но отдавал предпочтение *корпускулярной концепции света*, согласно которой свет – это поток частиц, действующий на *эфир*, введенный Декартом (механическая среда – переносчик света) и вызывающих в нем колебания. Движением световых частиц через эфир переменной плотности (из-за колебаний) и их взаимодействием с материальными телами по Ньютону обусловлены отражение и преломление света, цвета тонких пленок, дифракция света и его дисперсия. Ньютон подробно изучил эти явления. При этом он не считал возможным рассматривать свет как колебания самого эфира, т.к. в то время на этом пути не удавалось удовлетворительно объяснить прямолинейность световых лучей и *поляризацию света*, впервые осознанную именно Ньютоном и следовавшую из классических опытов Гюйгенса по двойному лучепреломлению. По Ньютону, поляризация – изначальное свойство света, объясняемое определенной ориентацией световых частиц по отношению к образуемому ими лучу.

Х.Гюйгенс (1629-1695), следовал идеям Леонардо да Винчи и развивал работы Гримальди и Гука, исходя из аналогии между акустическими и оптическими явлениями. Он считал, что световое возбуждение – это импульсы упругих колебаний эфира, распространяющиеся с *большой, но конечной скоростью*. Заметим, что И.Кеплер и Декарт считали скорость света бесконечной, Гук и Ньютон – конечной. Первое экспериментальное измерение скорости света произвел **О.Ремер**. Огромным вкладом Гюйгенса в оптику является установление им принципа, согласно которому каждая точка фронта волнового возбуждения может рассматриваться как источник вторичных сферических волн, огибающая которых представляет собой фронт реальной распространяющейся волны в последующие моменты времени (*принцип Гюйгенса-Френеля*). Опираясь на этот принцип, Гюйгенс объяснил законы отражения и преломления света. Причем из его теории следовало правильное выражение для показателей преломления: $n_{21} = v_1/v_2$, где $v_{1,2}$ – скорости света в первой и во второй среде. Он также объяснил двойное лучепреломление. Однако, сформулировав фундаментальный принцип, Гюйгенс не разработал последовательной волновой теории, которая могла бы противостоять воззрениям Ньютона. Поэтому корпускулярная

теория истечения света сохранила свое господствующее положение до начала XIX века.

Победа волновой оптики связана с работами **Т.Юнга** и **О.Френеля**. Юнг ввел *принцип интерференции* (1801) и на этой основе объяснил цвета тонких пленок. Френель, используя принцип Гюйгенса, дал удовлетворительное волновое объяснение прямолинейности распространения света и многочисленным дифракционным явлениям. В опытах Френеля и **Д.Араго** было установлено, что *волны, поляризованные перпендикулярно друг другу, не интерферируют*. Это позволило Юнгу и независимо от него Френелю высказать важную идею о *поперечности световых колебаний*, исходя из которой последний построил теорию кристалло-оптических явлений.

Т.о. все известные к тому времени оптические явления получили волновую интерпретацию. Вместе с тем представления о поперечности колебаний эфира привели к необходимости наделения его противоречивыми свойствами, в частности, твердого тела, в котором свободно перемещаются тела. Эти трудности были преодолены лишь при последовательном развитии учения об электромагнитном поле Дж.Максвеллом.

Основываясь на открытиях М.Фарадея, **Дж.Максвелл** пришел к выводу, что *свет есть электромагнитное поле*, а не упругая волна эфира. Первым указанием на непосредственную связь электромагнетизма со светом было открытое Фарадеем явление вращения плоскости поляризации света в магнитном поле (1848). Далее было установлено, что отношение электромагнитной и электростатической единиц силы тока по абсолютной величине и размерности совпадают со скоростью света c (**В.Вебер**, **Ф.Кольрауш**, 1856). Максвелл теоретически показал (1865), а **Г.Герц** подтвердил экспериментально, что электромагнитное поле распространяется в вакууме с этой скоростью. В прозрачной среде скорость света $v = c/n = c/\sqrt{\epsilon\mu}$. **Ф.Леру** (1862) обнаружил явление *аномальной дисперсии*, т.е. рост показателя преломления с увеличением длины волны света. **А.Кундт** показал, что это явление связано с поглощением и присуще многим веществам. Эти и последующие открытия привели к представлению о веществе как совокупности осцилляторов, с которыми взаимодействует свет (**В.Зельмейер**, 1872).

В 90-х годах XIX века **П.Друде**, **Г.Гельмгольц** и, в особенности, **Х.Лоренц** (1896) при создании электронной теории строения вещества объединили идею об осцилляторах и электромагнитную теорию света. Плодотворное представление об электронах, которые входят в состав атомов и молекул и способны совершать в них колебания, позволило описать многие оптические явления, в т.ч. *нормальную и аномальную дисперсию*. Рэлей (1890) создал основы *статистической оптики* (излучение ансамбля осцилляторов, рассеяние света). **Д.С.Рожественский** (1912) выполнил наиболее точные работы по аномальной дисперсии света, хорошо согласующиеся с электронной теорией. **П.Зееман** (1886) наблюдал, а Лоренц (1897) объяснил действие магнитного поля на частоты излучения и поглощения атомов. В полном согласии с теорией Максвелла оказалась и величина *светового давления* (впервые высказал Кеплер), измеренная **П.Н.Лебедевым** (1899).

Т.о. *построение электромагнитной теории света и дополнение ее электронной теорией взаимодействия света и вещества явилось крупнейшим этапом в развитии оптики*.

Электромагнитная теория света стала отправным пунктом при создании *теории относительности*. Данные оптических опытов с движущимися средами наблюдателями относительно источника излучения противоречили теоретическим представлениям того времени. **Юнг** (1804) для объяснения аберрации (искажения изображения) неподвижного, неувлекаемого Землей эфира. **Френель** (1818), напротив, для независимости показателя преломления тел от движения (наблюдения Араго), необходимо чтобы тела частично увлекали эфир. Этот вывод был подкреплён опытом **Физо**. Электродинамика движущихся сред, развитая **Лоренцем** (1896) в рамках электронной теории, также приводила к частичному увлечению эфира. Классический опыт **Майкельсона** (1881), а также опыт **Майкельсона-Морли** (1887) не обнаружили такого увлечения (эфирного ветра). Этот и ряд других опытов, противоречивших представлениям о среде-переносчике электромагнитных колебаний нашли объяснение в созданной **А.Эйнштейном** специальной (частной) теории относительности (1905), приведшей к кардинальному пересмотру классической физики и устранению необходимости в эфире и других гипотетических сред-переносчиках света.

Плодотворность классической электродинамической теории света Максвелла-Лоренца неоднократно подтверждалась в дальнейшем, например, в истолковании **И.Таммом** и **И.Франком** (1937) природы излучения Вавилова -Черенкова (1934), в выдвижении **Д.Габором** (1947) идеи голографии (полная запись волнового поля в одной плоскости) и разработке метода голографии в трехмерной среде, начало которому положили работы **Ю.Денисюка**.

Не менее интересной и поучительной, но быть может более драматичной, представляется история развития квантовых представлений о природе света и вещества. Однако в данном курсе мы не будем затрагивать эти вопросы.

Совершенно новый этап в развитии оптики начался с изобретением лазеров – источников когерентного излучения, которые привели к качественным изменениям не только в оптике, но и в других областях физики и техники.

Более подробно с историей развития оптики можно познакомиться, например, в следующих книгах:

1. Лансберг Г.С. Оптика. М.: Наука, 1980.
2. Сивухин Л.В. Общий курс физики. Оптика. М.: Наука, 1980.
3. Борн М., Вольф Э. Основы оптики. М.: Наука, 1980.

1.2. Распространение световые волн в вакууме. Решающие и наиболее поразительные периоды развития физики — это периоды великих обобщений, когда явления, ранее казавшиеся *разобщиенными*, неожиданно становятся всего лишь *разными аспектами одного и того же процесса*. История физики и, в частности, оптики — это история таких обобщений, а в основе успеха физической науки лежит главным образом наша *способность к синтезу*. По-видимому, самым знаменательным событием XIX столетия считается тот день, когда Дж.Максвелл в 1860 г. сопоставил законы электричества и магнетизма с законами поведения света. В результате были частично объяснены свойства света — старой и тонкой субстанции, настолько загадочной и важной, что в свое время при написании главы о сотворении мира сочли нужным отвести для него отдельный акт творения ("Да будет свет"). Создание электромагнитной теории света — одно из величайших достижений физики. Не даром, говоря современным языком, в рейтинге величайших ученых "всех времен и народов" создатель этой теории Дж.Максвелл постоянно находится на одном из первых мест (обычно после А.Эйнштейна).

Электромагнитная природа света. Вопрос о природе света представляет собой одну из центральных проблем физической оптики. Многие крупные открытия в области физики так или иначе связаны с попытками понять, что такое свет. Во второй половине XIX века Максвелл предсказал существование электромагнитного поля и доказал, что свет – это электромагнитные волны, распространяющиеся со скоростью примерно 300 000 км/с.

Измерение скорости света – важная задача в физике. История измерений скорости света – это история весьма трудоемких, а зачастую уникальных для своего времени экспериментов. Вместе с тем, это и история блестящих идей и находок, история создания замечательных оптических инструментов, методов особо точных измерений, нашедших в последствии важные применения. К концу XIX века физика располагала набором фактов, свидетельствующих в пользу *электромагнитной природы света*. К их числу относятся опыты Фарадея, в которых наблюдалось влияние магнитного поля на распространение света в веществе, опыты Лебедева, в которых было измерено световое давление, опыты Герца, в которых было доказано существование электромагнитных волн, эксперименты по взаимодействию света с веществом.

Одним из первых аргументов в пользу электромагнитной природы света было *совпадение скорости распространения электромагнитных волн, вычисленной Максвеллом, со скоростью света*. В 1849 г. Физо измерил скорость света и получил значение 315000 км/с. В 1857 г. Вебер и Кольрауш измерили электродинамическую постоянную c , равную отношению электромагнитной и электростатической единиц заряда, и получили значение $c = 310800$ км/с. В 1861 г. Максвелл вывел систему уравнений для электромагнитного поля, из которой вытекала возможность существования электромагнитных волн, причем скорость распространения волны определялась значением электродинамической постоянной c . Он также обратил внимание на то, что найденное Вебером и Кольраушем значение c весьма близко к скорости света, измеренной Физо. Это позволило ему заключить, что *свет представляет собой электромагнитную волну*. Предсказанное Максвеллом существование электромагнитных волн экспериментально подтвердил Г.Герц, который первым осуществил генерацию и прием электромагнитных волн и исследовал их свойства. Чтобы доказать единую сущность световых и электромагнитных волн, Герц продемонстрировал *отражение, преломление и поляризацию* электромагнитных волн. Таким образом, единая сущность света и электричества была подтверждена экспериментально. Оптика могла быть теперь включена в электродинамику, так же как акустика давно уже вошла в механику.

Трудность измерения скорости света связана с ее чрезвычайно большой величиной. История измерений скорости света весьма интересна и с ней можно познакомиться, например, в книге С.А.Ахманова и С.Ю.Никитина "Физическая оптика", 1998 (дополнение 1). *Скорость света в вакууме c – фундаментальная (мировая) физическая постоянная*, имеющая огромное значение не только в оптике, но и в физике в целом, в астрономии, астрофизике. Поэтому абсолютные и относительные измерения скорости света привлекают внимание физиков и инженеров на протяжении более трехсот лет. За это время точность измерений величины c возросла от $\Delta c/c = 0.3$ до $\Delta c/c \simeq 3 \times 10^{-9}$. Лазерная физика и техника открыли совершенно новые возможности. Наиболее точное определение c , выполненное в лазерных экспериментах в 90-х годах прошлого столетия, дало значение $c = 299\,792\,456,2$ м/с.

Уравнения Максвелла в вакууме. Система уравнений для электромагнитного поля получена Максвеллом в середине XIX в путем обобщения опытных данных с

электрическими зарядами, токами и магнитами. Уравнения Максвелла имеют очень глубокое физическое содержание, далеко выходящее за рамки тех фактов и представлений, на основе которых они были получены. Эти уравнения хорошо описывают быстропеременное электромагнитное поле, включая световые волны, и составляют основу теории излучения электромагнитных волн движущимися зарядами и теории взаимодействия света и вещества.

Электромагнитное поле имеет две компоненты – *электрическую* и *магнитную*. Первая описывается *вектором электрической напряженности*, вторая – *вектором магнитной напряженности*. В удобной для оптики гауссовой системе единиц *уравнения Максвелла для электромагнитного поля в вакууме* имеют вид

$$\operatorname{rot} \vec{E} = -\frac{1}{c} \frac{\partial \vec{H}}{\partial t} \quad \operatorname{rot} \vec{H} = \frac{1}{c} \frac{\partial \vec{E}}{\partial t} \quad (1.1)$$

$$\operatorname{div} \vec{E} = 0 \quad \operatorname{div} \vec{H} = 0 \quad (1.2)$$

Здесь $\vec{E}$ и $\vec{H}$ – напряженности электрического и магнитного полей, c – скорость света в вакууме. Первое уравнение (1.1) представляет собой математическую формулировку закона *электромагнитной индукции*, а второе – показывает, что *магнитное поле порождается переменным электрическим полем*. Первое уравнение (1.2) выражает факт *отсутствия статического электрического поля в вакууме*, а второе – постулирует *отсутствие магнитных зарядов*.

Волновое уравнение. Уравнения (1.1)-(1.2) позволяют вывести замкнутые уравнения для полей $\vec{E}$ и $\vec{H}$, которые называют *волновым уравнением*.

Дифференцируя второе уравнение (1.1) по времени и меняя порядок следования временной и пространственных производных, имеем

$$\operatorname{rot} \left(\frac{\partial \vec{H}}{\partial t} \right) = \frac{1}{c} \frac{\partial^2 \vec{E}}{\partial t^2} \quad (1.3)$$

Воспользовавшись первым уравнением (1.1), получим

$$\operatorname{rot} \operatorname{rot} \vec{E} = -\frac{1}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} \quad (1.4)$$

Применяя известное из векторного анализа соотношение для дифференциальных операторов, преобразуем левую часть последнего уравнения к виду

$$\operatorname{rot} \operatorname{rot} \vec{E} = \operatorname{grad} \operatorname{div} \vec{E} - \Delta \vec{E} \quad (1.5)$$

Δ – оператор Лапласа, который в декартовых координатах x, y, z имеет вид

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad (1.6)$$

Поскольку в вакууме свободные заряды отсутствуют, т. е. $\operatorname{div} \vec{E} = 0$, для вектора напряженности электрического поля получаем следующее уравнение

$$\Delta \vec{E} - \frac{1}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} = 0 \quad (1.7)$$

Аналогичным образом получается уравнение для $\vec{H}$

$$\Delta \vec{H} - \frac{1}{c^2} \frac{\partial^2 \vec{H}}{\partial t^2} = 0 \quad (1.8)$$

Уравнения (1.7) и (1.8) линейны по полю. Поэтому они эквивалентны совокупности скалярных уравнений того же самого вида, в каждое из которых входит только одна компонента напряженности электрического или магнитного поля. Действительно, запишем векторы $\vec{E}$ и $\vec{H}$ через декартовы компоненты $\vec{E}_{x,y,z}$ и $\vec{H}_{x,y,z}$, соответственно:

$$\vec{E} = \vec{x}_0 E_x + \vec{y}_0 E_y + \vec{z}_0 E_z, \quad \vec{H} = \vec{x}_0 H_x + \vec{y}_0 H_y + \vec{z}_0 H_z, \quad (1.9)$$

$\vec{x}_0, \vec{y}_0, \vec{z}_0$ — единичные векторы ("орты"), направленные вдоль осей x, y, z декартовой системы координат.

Умножая скалярно уравнения (1.7) и (1.8) последовательно на $\vec{x}_0, \vec{y}_0, \vec{z}_0$, получаем, что каждая из компонент полей E_α или H_α ($\alpha = x, y, z$) удовлетворяет скалярному уравнению

$$\Delta E_\alpha - \frac{1}{c^2} \frac{\partial^2 E_\alpha}{\partial t^2} = 0, \quad \Delta H_\alpha - \frac{1}{c^2} \frac{\partial^2 H_\alpha}{\partial t^2} = 0, \quad (1.10)$$

Уравнения (1.7), (1.8) и (1.10) называются *волновыми уравнениями*. Они описывают распространения полей $\vec{E}$ и $\vec{H}$ в пространстве и времени. Их решения имеют характер распространяющихся волн.

Рассмотрим свойства световых волн на примере наиболее простых ("эталлонных") волн. К числу таких волн относят *плоские* и *сферические волны*. Подчеркнем, что эти волны являются *идеализациями* и в природе их не существует, но они позволяют, как будет показано в дальнейшем, рассматривать процесс распространения любых световых волн.

Плоская волна. Предположим, что произвольная компонента поля f (например, E_α или H_α) зависит лишь от одной пространственной координаты z и времени t , т.е. $f = f(z, t)$. Тогда уравнения (1.10) упрощаются и принимают вид:

$$\frac{\partial^2 f}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2 f}{\partial t^2} = 0. \quad (1.11)$$

Можно показать, что решение уравнения (1.11) имеет вид:

$$f(z, t) = f_1(t - z/c) + f_2(t + z/c), \quad (1.12)$$

где $f_{1,2}$ — произвольные (дифференцируемые) функции своих аргументов.

Формула (1.12) есть *общее решение уравнения (1.11)*. Данное решение представляет *суперпозицию* двух волн. Первая из них распространяется вдоль направления оси z , вторая — против. Скорости обеих волн одинаковы и равны c .

Действительно, возмущение f_1 , находившееся в момент времени t_1 в точке z_1 , в момент времени t_2 приходит в точку z_2 , определяемую соотношением $t_1 - z_1/c = t_2 - z_2/c$. Отсюда при $t_2 > t_1$ имеем $z_2 > z_1$ и скорость распространения волнового возмущения равна $v = (z_2 - z_1)/(t_2 - t_1) = c$, т.е. скорости света.

Функции $f_{1,2} = f_{1,2}(z, t)$ описывают *плоские волны*, так как *волновое возмущение имеет одно и то же значение во всех точках бесконечной плоскости, перпендикулярной направлению распространения*. Конкретный вид функций $f_{1,2}$ определяется граничными и начальными условиями задачи.

Плоская гармоническая волна. Для определенности рассмотрим декартову компоненту поля $E_x(z, t) = E(z, t)$. Пусть при $z = 0$ напряженность светового поля изменяется по гармоническому закону $E(0, t) = A \cos \omega t$. Тогда в соответствии с формулой (1.12) в области $z \geq 0$ будет распространяться плоская гармоническая волна,

которая имеет вид:

$$E(z, t) = A \cos [\omega(t - z/c)] = A \cos (\omega t - kz) = A \cos [\omega t - \varphi(z)] \quad (1.13)$$

Здесь A – амплитуда волны, $\omega = 2\pi\nu$ – круговая частота, $\nu = 1/T$ – частота колебаний, T – период, $k = \omega/c = 2\pi\nu/c = 2\pi/\lambda$ – волновое число, $\lambda = cT = c/\nu$ – длина волны. Величина $\Phi = \omega t - kz$ называется *полной фазой волны* и зависит от t и z . Фазу $\varphi(z) = kz$, связанную изменением пути z , пройденного волной, называют *набегом фазы* или *фазовым сдвигом*.

Геометрическое место одинаковых значений фаз называют *волновым фронтом*. В плоской гармонической волне это плоскость, перпендикулярная направлению распространения. Волну типа (1.13) называют *монохроматической* плоской волной. Монохроматический значит "одноцветный". Происхождение этого термина связано с тем, что в видимом диапазоне глаз регистрирует изменение частоты излучения как изменение цвета, т.е. разным длинам волн соответствуют разные цвета.

Для плоских волн поверхность постоянной фазы совпадает с поверхностями постоянной амплитуды. Такие волны называют *однородными*.

В общем случае необходимо учитывать и начальную фазу ψ_0 . Тогда выражение для плоской волны принимает вид:

$$E(z, t) = A \cos (\omega t - kz + \psi_0), \quad (1.14)$$

Вид функции в (1.13) или (1.14) показывает, что она **периодична** по времени с периодом T , и по координате z с *пространственным периодом* λ .

В общем случае, когда плоская волна распространяется в произвольном направлении относительно декартовой системы координат, она описывается следующим уравнением:

$$E(\vec{r}, t) = A \cos (\omega t - \vec{k}\vec{r} + \psi_0), \quad (1.15)$$

где $\vec{r}$ – радиус вектор точки с координатами x, y, z , $\vec{k} = \vec{n}k$ – волновой вектор, $\vec{n}$ – единичный вектор вдоль волнового вектора.

Отметим, что применительно к монохроматической волне, ее *скорость распространения это скорость, с которой передается от точки к точке фаза колебания*. Действительно, скорость распространения фазы определяется из требования $\omega(t - z/c) = \text{const}$, что означает неизменность фазы в плоскости z . Отсюда следует, что $dz/dt = c$, c называют *фазовой скоростью монохроматической волны*. Опыт показывает, что *в вакууме фазовая скорость не зависит от частоты*.

Часто вместо тригонометрических функций удобно использовать экспоненциальные. В основе этого лежит формула Эйлера: $\exp(i\psi) = \cos \psi + i \sin \psi$. Так как большинство математических операций легче выполнять с показательной функцией, чем с тригонометрическими, то поступают следующим образом: вместо косинуса или синуса вводят показательную функцию, производят с ней все необходимые вычисления и в конце возвращаются к тригонометрическим функциям, взяв действительную или мнимую части. Пользуясь показательной функцией выражение (1.14) можно записать в виде

$$E(z, t) = \frac{1}{2} E_0 \exp [i(\omega t - kz)] + \text{K.C.}, \quad (1.16)$$

где $E_0 = A \exp(i\psi_0)$ – *комплексная амплитуда*.

Выражение $\mathcal{E}(z, t) = 1/2 E_0 \exp [i(\omega t - kz)]$ называют *комплексным представлением монохроматической волны*. Видно, что модуль комплексной амплитуды равен действительной амплитуде, а аргумент – фазе световой волны. На комплексной

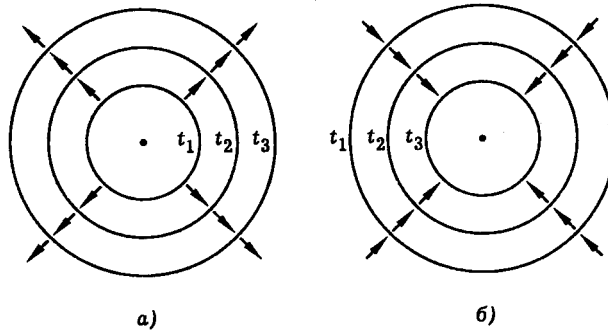


Рис. 1.1. *Расходящаяся* (а) и *сходящаяся* (б) сферические волны. Показаны положения фронтов в последовательные моменты времени $t_1 < t_2 < t_3$

плоскости комплексная амплитуда изображается в виде вектора, длина которого равна действительной амплитуде A и направленный под углом φ к оси абсцисс. Удобство комплексной записи заключается в том, что при дифференцировании (1.16) по времени и координате имеем

$$\frac{\partial \mathcal{E}(z, t)}{\partial t} = i\omega \mathcal{E}(z, t), \quad \frac{\partial \mathcal{E}(z, t)}{\partial z} = -ik \mathcal{E}(z, t). \quad (1.17)$$

Сферическая волна. Несложно убедиться, что волновому уравнению (1.11) удовлетворяют и волны вида $E_\alpha = E_\alpha(t, r)$ и $H_\alpha = H_\alpha(t, r)$, у которых напряженности полей зависят только от *одной пространственной координаты* – модуля радиус-вектора $r = \sqrt{x^2 + y^2 + z^2}$. Такие волны называют *сферическими*. Наряду с плоской, сферическая волна также является *эталонной волной*, имеющей большое значение для оптики. Сферическая волна представляет решение скалярного волнового уравнения (1.11) и имеет вид:

$$f(r, t) = \frac{F_1(t - r/c)}{r} + \frac{F_2(t + r/c)}{r} \quad (1.18)$$

Формула (1.18) описывает две сферические волны, *расходящуюся* от начала координат и *сходящуюся* к нему, как показано на **Рис.1.1**.

Гармоническая сферическая волна. Если на сфере радиуса r_0 задать гармоническое возмущение, синфазное во всех точках сферы $f(r_0, t) = (A_0/r_0) \cos [\omega(t - r_0/c)]$, то возбуждаемая таким источником расходящаяся волна при $r > r_0$ может быть представлена в виде

$$f(r, t) = (A_0/r_0) \cos [\omega(t - r_0/c)]. \quad (1.19)$$

Здесь в отличие от плоской волны амплитуда зависит от координаты, а фазовый и амплитудный фронты представляют собой сферы. Строго говоря, сферическая волна испускается точечным источником света, т.е. представляет абстракцию, как и плоская волна. На практике для многих задач можно считать волну сферической, если расстояние r превосходит линейные размеры источника в десять раз и более.

Принцип суперпозиции. Принцип суперпозиции утверждает, что световые волны разных частот и разных направлений распространяются в вакууме независимо друг от друга. На это указывают самые простые эксперименты. Например, через одно и то же отверстие два наблюдателя могут видеть разные объекты; при этом наблюдаемые ими картины, вообще говоря, никак не связаны между собой. Опыт подтверждает принцип суперпозиции в широких пределах: свету далекой звезды,

идушей к нам из космоса, не мешает распространяться свет других звезд или свет горячей поблизости лампочки. Они беспрепятственно проходят друг через друга, не искажая и не "замечая" друг друга.

Математически *принцип суперпозиции является следствием линейности волнового уравнения*, описывающего распространение световых волн в вакуум. В самом деле, если поля $E_{1,2,\dots,n}$ являются решениями волнового уравнения, то его решением оказывается и сумма полей $E = \sum_n E_n$.

В этом можно убедиться, подставляя, например, в уравнения Максвелла плоские волны. При этом волновое уравнение распадается на независимые уравнения для отдельных волн. Отметим, что принцип суперпозиции, как правило, справедлив для случая слабых световых полей. Смысл этого равенства состоит в том, распространяющиеся одновременно волны не влияют друг на друга, поэтому результирующее возмущение есть просто сумма неискажающих друг друга возмущений.

Поперечность световой волны. Покажем, что плоская монохроматическая световая волна, распространяющаяся в вакууме, является *поперечной*, т. е. компоненты полей в направлении распространения волны отсутствуют. Здесь удобно воспользоваться комплексным представлением плоской гармонической волны:

$$\vec{E} = \frac{1}{2}\vec{\mathcal{E}} \exp[i(\omega t - \vec{k}\vec{r})] + K.C., \quad \vec{H} = \frac{1}{2}\vec{\mathcal{H}} \exp[i(\omega t - \vec{k}\vec{r})] + K.C. \quad (1.20)$$

где $\vec{\mathcal{E}}$ и $\vec{\mathcal{H}}$ — комплексные амплитуды напряженностей электрической и магнитной составляющих волны, соответственно; $\vec{k} = \vec{n}\omega/c$ — волновой вектор вдоль направления распространения) $\vec{k}$; $\vec{n}$ — единичный вектор вдоль направления волнового вектора.

Вычислим $\text{rot } \vec{E}$ и $\text{div } \vec{E}$, которые входят в уравнения Максвелла. Так как операция $\text{rot } \vec{E}$ сводится к действию дифференциальных операторов $\partial/\partial\alpha$, где $\alpha = x, y, z$, то их действие на экспоненциальный множитель $\exp[i(\omega t - \vec{k}\vec{r})]$, сводится к умножению на соответствующие декартовы компоненты волнового вектора

$$\frac{\partial \exp[i(\omega t - \vec{k}\vec{r})]}{\partial \alpha} = -ik_\alpha \exp[i(\omega t - \vec{k}\vec{r})]. \quad (1.21)$$

С учетом (1.21) для ротора и дивергенции вектора $\vec{E}$ получаем

$$\text{rot } \vec{E} = \frac{1}{2}[-i\vec{k}, \vec{\mathcal{E}}] \exp[i(\omega t - \vec{k}\vec{r})] + K.C., \quad (1.22)$$

$$\text{div } \vec{E} = \frac{1}{2}(-i\vec{k}, \vec{\mathcal{E}}) \exp[i(\omega t - \vec{k}\vec{r})] + K.C. \quad (1.23)$$

Здесь $[\cdot, \cdot]$ и $(\cdot, \cdot)$ означает векторное и скалярное произведение соответствующих векторов. Также вычисляется и производная по времени от $\vec{E}$.

$$\frac{\partial \vec{E}}{\partial t} = \frac{1}{2}i\omega \vec{\mathcal{E}} \exp[i(\omega t - \vec{k}\vec{r})] + K.C. \quad (1.24)$$

Аналогичные формулы имеют место для вектора $\vec{H}$. Пользуясь ими, систему уравнений Максвелла можно записать в виде:

$$[\vec{k}, \vec{\mathcal{E}}] = \frac{\omega}{c}\vec{\mathcal{H}}, \quad [\vec{k}, \vec{\mathcal{H}}] = -\frac{\omega}{c}\vec{\mathcal{E}}, \quad (\vec{k}, \vec{\mathcal{E}}) = 0, \quad (\vec{k}, \vec{\mathcal{H}}) = 0. \quad (1.25)$$

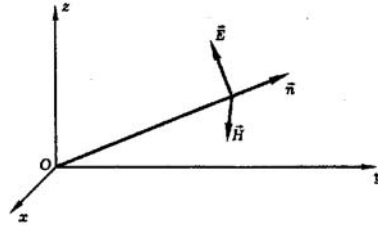


Рис. 1.2. Структура плоской монохроматической световой волны

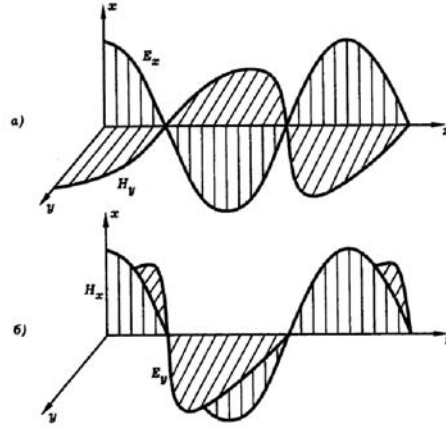


Рис. 1.3. Плоские гармонические волны E_x, H_y (а) и E_y, H_x (б). Здесь показана "мгновенная фотография—распределение полей $\vec{E}$ и $\vec{H}$ как функция координаты z в некоторый момент времени t для волн, бегущих слева направо.

Учитывая, что $\vec{k} = \vec{n}k = \vec{n}\omega/c$ ($\vec{n} = \vec{k}/k$ — единичный вектор), перепишем (1.25) в виде

$$[\vec{n}, \vec{\mathcal{E}}] = \vec{\mathcal{H}}, \quad [\vec{n}, \vec{\mathcal{H}}] = -\vec{\mathcal{E}}, \quad (\vec{n}, \vec{\mathcal{E}}) = 0, \quad (\vec{n}, \vec{\mathcal{H}}) = 0. \quad (1.26)$$

Из (1.26) и (1.20) получаем векторные соотношения, определяющие структуру поля плоской монохроматической световой волны:

$$\vec{E} = -[\vec{n}, \vec{H}], \quad (\vec{n}, \vec{E}) = 0, \quad \vec{H} = [\vec{n}, \vec{E}], \quad (\vec{n}, \vec{H}) = 0. \quad (1.27)$$

Из соотношений $(\vec{n}, \vec{E}) = 0$ и $(\vec{n}, \vec{H}) = 0$ следует, что $E_z = H_z = 0$, т.е. *плоская световая волна является поперечной*. В такой волне *колебания векторов $\vec{E}$ и $\vec{H}$ происходят в плоскости, перпендикулярной направлению распространения волны*, а векторы $\vec{n}$, $\vec{E}$ и $\vec{H}$ образуют *правую тройку векторов (Рис.1.2)*. Из соотношения $\vec{H} = [\vec{n}, \vec{E}]$ можно заключить, что векторы $\vec{E}$ и $\vec{H}$ изменяются с одной фазой, т.е. *синфазно*, как показано на **Рис.1.3** и $|\vec{E}| = |\vec{H}|$.

ЛЕКЦИЯ №2

Поток энергии, интенсивность света. Импульс и давление света. Поляризация света: эллиптическая, круговая, линейная поляризации; естественно поляризованный свет.

2.1. Поток энергии, интенсивность света. Способность переносить энергию — одно из основных свойств света. При распространении световых волн происходит перенос энергии поля в пространстве даже в вакууме. Действительно, пусть световая волна распространяется в направлении оси z . Уравнения Максвелла в вакууме имеют вид

$$\operatorname{rot} \vec{H} = \frac{1}{c} \frac{\partial \vec{E}}{\partial t}, \quad \operatorname{rot} \vec{E} = -\frac{1}{c} \frac{\partial \vec{H}}{\partial t}. \quad (2.1)$$

Умножим первое уравнение скалярно на $\vec{E}$, а второе — на $\vec{H}$ и вычтем из первого второе. Получим

$$\frac{1}{c} \left(\vec{E} \frac{\partial \vec{E}}{\partial t} + \vec{H} \frac{\partial \vec{H}}{\partial t} \right) = \vec{E} \operatorname{rot} \vec{H} - \vec{H} \operatorname{rot} \vec{E} \quad (2.2)$$

Используя формулу

$$\operatorname{div}[\vec{E} \vec{H}] = \vec{H} \operatorname{rot} \vec{E} - \vec{E} \operatorname{rot} \vec{H}, \quad (2.3)$$

и вводя обозначения

$$w = \frac{1}{8\pi} (E^2 + H^2), \quad \vec{S} = \frac{c}{4\pi} [\vec{E}, \vec{H}] \quad (2.4)$$

соотношение (2.2) перепишем в виде

$$\frac{\partial w}{\partial t} = -\operatorname{div} \vec{S}. \quad (2.5)$$

Величины w и $\vec{S}$ называют *объемной плотностью энергии поля* и *вектором потока энергии* (или *вектором Умова-Пойнтинга*), соответственно.

Уравнение (2.5) является следствием уравнений Максвелла и выражает *закон сохранения энергии* для электромагнитного поля. Проинтегрируем обе части (2.5) по некоторому объему V , ограниченному замкнутой поверхностью Σ (**Рис. 2.1**). Интеграл по объему от $\operatorname{div} \vec{S}$ в правой части преобразуем, используя теорему Остроградского-Гаусса, в интеграл по поверхности, ограничивающий этот объем :

$$\int_V \operatorname{div} \vec{S} dV = \int_{\Sigma} \vec{S} \vec{l}_n d\sigma, \quad (2.6)$$

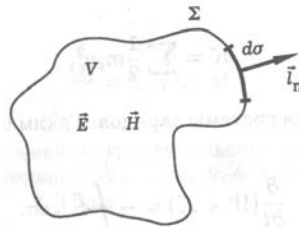


Рис. 2.1. К выводу закона сохранения энергии для системы заряженных частиц в световом поле

$\vec{l}_n$ — единичный вектор нормали к элементу $d\sigma$ поверхности Σ , ограничивающей объем V .

$$\frac{\partial W}{\partial t} = - \int_{\Sigma} \vec{S} \vec{l}_n d\sigma, \quad (2.7)$$

где $W = \int_V w dv$ — полная энергия в объеме, а в правой части интеграл представляет поток энергии через поверхность.

Уравнение (2.7) показывает, что скорость изменения энергии электромагнитного поля в некотором объеме, не содержащем зарядов и токов, равна потоку энергии поля через поверхность этого объема.

Так как $\vec{H} = [\vec{n}, \vec{E}]$ и $\vec{E} = -[\vec{n}, \vec{H}]$, то для плоской монохроматической волны можно записать

$$\vec{S} = \frac{c}{4\pi} [\vec{E}, \vec{H}] = \frac{c}{4\pi} [\vec{E}, [\vec{n}, \vec{E}]] = \frac{c}{4\pi} E^2 \vec{n}, \quad |\vec{S}| = S = \frac{c}{4\pi} E^2. \quad (2.8)$$

Отсюда с учетом формулы для плотности энергии (2.4) получаем

$$\vec{S} = cw\vec{n}, \quad S = cw, \quad (2.9)$$

где $\vec{n}$ — единичный вектор вдоль направления распространения волны $\vec{k}$. Таким образом, в вакууме световая энергия переносится в направлении распространения плоской волны со скоростью света c , т.е. направление переноса энергии совпадает с направлением волнового вектора.

Интенсивность света. Реальное (нелазерное) световое излучение представляет случайный или статистический процесс, свойства которого определяются вероятностными законами, обусловленные фундаментальными физическими свойствами источников оптического излучения.

Прежде всего, всякий реальный источник света — это совокупность огромного числа атомов. Каждый из атомов может стать элементарным источником излучения, если благодаря какому-либо механизму (столкновение с другим атомом, электронный удар и т. д.) он будет переведен в возбужденное состояние. Случайный, статистический характер механизма возбуждения определяет статистическую природу излучения. Любой из атомов может начать высвечиваться в произвольный, случайный момент времени (мы не можем предвидеть момент высвечивания атома).

Кроме того, возбужденный атом излучает конечное время. С классической точки зрения атом можно рассматривать как колебательную систему (осциллятор) с определенной добротностью. Излучение такого осциллятора представляет собой процесс свободных затухающих колебаний, как показано на **Рис. 2.2**. Такую волну называют **цугом**. Атомные системы обладают чрезвычайно малым затуханием, т.е. имеют высокую добротность. Поэтому длительность цуга τ_0 существенно больше периода оптических колебаний T : $\tau_0 \gg T$. Характерные длительности цугов для оптического диапазона излучения $\tau_0 \sim 10^{-8} - 10^{-7}$ с, что на много порядков превышает период светового колебания $T \simeq 10^{-15}$ с. Возбужденный атом, испуская цуг, постепенно

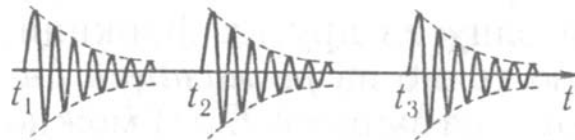


Рис. 2.2. Последовательность цугов, излучаемых атомом какого-либо источника света

теряет энергию и по прошествии времени порядка τ_0 переходит в основное, невозбужденное состояние. Поэтому длительность цуга можно рассматривать как *время жизни атома в возбужденном состоянии*. Последовательность цугов, излучаемых атомом какого-либо источника света, показана на **Рис.2.2**. Статистический характер механизма возбуждения приводит к тому, что моменты возникновения излучаемых цугов $t_1, t_2, \dots, t_n$ также являются случайными независимыми величинами.

Наконец, ясно, что *в излучении одновременно принимает участие большое число атомов*. Обозначая последовательность цугов, излучаемых каким-либо n -м атомом через $e_n(t)$, суммарное излучение, согласно принципу суперпозиции, можно записать в виде:

$$E(t) = \sum_n e_n(t) \quad (2.10)$$

Причем сумма (2.10) содержит в каждый момент времени огромное число отличных от нуля слагаемых. Такое излучение можно представить в виде

$$E(t) = A(t) \cos(\omega_0 t - \varphi(t)), \quad (2.11)$$

где амплитуда $A(t)$ и фаза $\varphi(t)$ — *медленно и хаотически* меняющиеся функции: они остаются практически неизменными на интервалах времени Δt , содержащих большое число периодов светового колебания $T = 2\pi/\omega_0$. Именно в этом смысле понимается медленность их изменения, а колебание (2.11) называют квазимонохроматическим. Характерный временной масштаб их изменения τ_0 называют *временем корреляции* или *временем когерентности*. Говорят, что значения $A(t)$ и $A(t + \tau)$ *коррелированы*, т.е. изменяются *согласовано*, при $\tau < \tau_0$ и *некоррелированы* (изменяются *не согласовано*) при $\tau > \tau_0$. Это же относится и к значениям фазы $\varphi(t)$ и $\varphi(t + \tau)$.

Квазимонохроматическое колебание (2.11) изображается на векторной диаграмме в виде вектора с *медленно и хаотически* изменяющимися длиной и направлением. В комплексном представлении колебанию $E(t)$ (2.11) соответствует комплексная функция

$$V(t) = A(t) \exp[-i(\omega t - \varphi(t))] = f(t) \exp(-i\omega t), \quad (2.12)$$

где $f(t) = A(t) \exp[i\varphi(t)]$ — комплексная амплитуда.

Подчеркнем, что в выражении (2.12) *статистический (случайный)* характер имеет первый сомножитель — комплексная амплитуда $f(t)$, второй же сомножитель $\exp(-i\omega t)$ описывает гармоническое колебание с частотой ω . Функцию $f(t)$ называют *огibaющей* медленно и хаотически меняющегося случайного процесса $V(t)$, т.е. функция $f(t)$ описывает случайную амплитудно-фазовую модуляцию гармонического процесса. Реальное возмущение связано с $V(t)$ формулой $E(t) = 1/2 \operatorname{Re} V(t) = 1/2 \operatorname{Re} [f(t) \exp(-i\omega t)]$.

Установим соотношение между комплексной функцией $V(t)$ и результатами экспериментальных наблюдений. Но сначала сделаем несколько замечаний о свойствах фотоприемника, с помощью которого проводится анализ светового излучения. Таким фотоприемником может быть глаз, фотопластинка, фотоэлемент и т. д. Все эти устройства регистрируют *поток энергии, усредненный за некоторое, характерное для данного фотоприемника, время ΔT* , которое называют *временем регистрации* (или постоянной времени фотоприемника).

Постоянные времени разных приемников сильно различаются: у сетчатки глаза время экспозиции (регистрации) $\Delta T \simeq 10^{-1}$ с (именно поэтому глаз не успевает

замечать мелькания кадров на экране кино и телевизора). Время экспозиции фотоматериалов обычно $\Delta T \simeq 10^{-2} - 10^{-4}$ с, хотя может быть и минуты и даже часы. У некоторых типов фотоэлементов время регистрации $\Delta T \simeq 10^{-6} - 10^{-8}$ с. *Результат измерения существенно зависит от соотношения между временем экспозиции и характерным временным масштабом регистрируемых процессов.*

Квазимонохроматические колебания характеризуются двумя временными масштабами: периодом несущего колебания $T = 2\pi/\omega_0$ и временем когерентности τ_0 . Для видимого света, как отмечалось, $T = 2\pi/\omega_0 \simeq 10^{-15}$ с, а время когерентности даже для самых узких спектральных линий не превышает значений $\tau_0 \simeq 10^{-7} - 10^{-8}$ с (излучение когерентных источников света — лазеров — имеет время когерентности на несколько порядков больше). Таким образом, в оптике всегда $\Delta T \gg T$, т.е. *не существует фоторегистрирующих устройств, способных следить за отдельным периодом светового колебания.*

При сравнении ΔT и τ_0 можно выделить два случая:

1. $\Delta T \lesssim \tau_0$. Такой фотоприемник может следить за случайными изменениями светового потока, обусловленными конечной длительностью цугов излучения атомов.

2. $\Delta T \gtrsim \tau_0$. Фотоприемник регистрирует *средний световой поток*. Этот способ регистрации называют *инерционным*. Этот случай реализуется в большинстве оптических экспериментов, и его мы будем иметь в виду в дальнейшем.

Поскольку световой поток (вектор Пойтинга) пропорционален квадрату напряженности поля E^2 в волне, то регистрируется средняя за время ΔT величина

$$\bar{S} = \frac{c}{4\pi} \frac{1}{\Delta T} \int_{t-\Delta T/2}^{t+\Delta T/2} E^2(t) dt \equiv I. \quad (2.13)$$

Здесь черта сверху означает усреднение за время $\Delta T \gg T$. Этот *усредненный поток энергии* называют *интенсивностью* света.

Используя (2.11) перепишем (2.13) в виде

$$I = \frac{c}{4\pi} \frac{1}{\Delta T} \int_{t-\Delta T/2}^{t+\Delta T/2} A^2(t) \cos^2[\omega t - \varphi(t)] dt. \quad (2.14)$$

Выражая квадрат косинуса с помощью равенства $\cos^2 \alpha = (1 + \cos 2\alpha)/2$, получим:

$$I = \frac{1}{2\Delta T} \int_{t-\Delta T/2}^{t+\Delta T/2} A^2 dt + \frac{1}{2\Delta T} \int_{t-\Delta T/2}^{t+\Delta T/2} A^2(t) \cos[2\omega t - 2\varphi(t)] dt. \quad (2.15)$$

За время наблюдения ΔT подынтегральная функция во втором слагаемом, осциллируя с удвоенной световой частотой, многократно меняет знак, поэтому второй интеграл пренебрежимо мал в сравнении с первым, поэтому

$$I = \frac{c}{8\pi} \overline{A^2(t)}, \quad \overline{A^2(t)} = \frac{1}{\Delta T} \int_{t-\Delta T/2}^{t+\Delta T/2} A^2(t) dt. \quad (2.16)$$

С другой стороны, используя комплексное представление (2.12), имеем:

$$\overline{f(t)f^*(t)} = \overline{A(t)e^{i\varphi(t)}A^*(t)e^{-i\varphi(t)}} = \overline{A^2(t)} \quad (2.17)$$

Таким образом, наблюдаемая интенсивность выражается через комплексную амплитуду $V(t)$ равенством

$$I = 1/2 \overline{f(t)f^*(t)} = 1/2 \overline{V(t)V^*(t)} \quad (2.18)$$

В этом представлении выражение для интенсивности принимает вид:

$$I = \frac{c}{8\pi} |\vec{f}|^2. \quad (2.19)$$

Данные формулы применимы и для плоской монохроматической волны в вакууме. Размерность интенсивности, как и потока энергии, есть эрг/(см²с). В системе СИ размерность I Вт/м². Часто используется внесистемная единица измерения интенсивности Вт/см² = 10⁷ эрг/(см²с).

2.2. Поляризация света. Плоская волна, распространяющаяся вдоль оси z , может колебаться либо вдоль оси x (E_x, H_y), либо вдоль y (E_y, H_x). Эти две поперечные волны отличаются друг от друга направлениями векторов $\vec{E}$ и $\vec{H}$ или, как говорят, направлениями *поляризации*. Такие волны называются *линейно поляризованными*, так как при фиксированном значении z вектор $\vec{E}$ все время находится в одной плоскости и движется по прямой линии. *Плоскость, в которой лежат вектор $\vec{E}$ и единичный вектор $\vec{n}$, характеризующий направление распространения волны называют плоскостью поляризации.* Линейно поляризованные волны также называют *плоско поляризованными*.

Далее мы, в основном, будем говорить об *электрической компоненте световой волны*, потому что при взаимодействии света с веществом сила, действующая со стороны светового поля на электрический заряд в v/c (v — скорость движения заряда) раз превосходит действие магнитного поля.

В общем случае плоская гармоническая волна $\vec{E}$ может иметь обе отличные от нуля компоненты E_x , и E_y . Тогда $\vec{E}$ можно записать в виде суперпозиции двух волн

$$\vec{E}(t, z) = \vec{x}_0 E_x(t, z) + \vec{y}_0 E_y(t, z), \quad (2.20)$$

где $\vec{x}_0, \vec{y}_0$ — единичные орты, направленные вдоль соответствующих осей. Пусть

$$E_x(t, z) = A_1 \cos(\omega t - kz + \varphi_1), \quad E_y(t, z) = A_2 \cos(\omega t - kz + \varphi_2). \quad (2.21)$$

Найдем уравнение траектории, по которой движется конец вектора $\vec{E}(t, z)$ в плоскости $z = \text{const}$. Вводя обозначение $\psi = \omega t - kz$ и, учитывая, что $\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta$, перепишем (2.21) следующим образом:

$$E_x = A_1(\cos \psi \cos \varphi_1 - \sin \psi \sin \varphi_1), \quad E_y = A_1(\cos \psi \cos \varphi_2 - \sin \psi \sin \varphi_2) \quad (2.22)$$

Отсюда, путем несложных преобразований, (2.22) можно переписать в виде

$$\frac{E_x}{A_1} \sin \varphi_2 - \frac{E_y}{A_2} \sin \varphi_1 = \cos \psi \sin(\varphi_2 - \varphi_1) \quad (2.23)$$

$$\frac{E_x}{A_1} \cos \varphi_2 - \frac{E_y}{A_2} \cos \varphi_1 = \sin \psi \sin(\varphi_2 - \varphi_1) \quad (2.24)$$

Возводя в квадрат правые и левые части этих уравнений и складывая их, найдем

$$\left(\frac{E_x}{A_1}\right)^2 + \left(\frac{E_y}{A_2}\right)^2 - 2 \frac{E_x E_y}{A_1 A_2} \cos(\varphi_2 - \varphi_1) = \sin^2(\varphi_2 - \varphi_1) \quad (2.25)$$

Уравнение (2.25) является уравнением эллипса, вписанным в прямоугольник, стороны которого параллельны осям x, y и имеют длины $2A_1$ и $2A_2$ **Рис.2.3** а. Ориентация эллипса зависит от величины разности фаз φ (**Рис.2.3** г).

Таким образом, в общем случае при распространении плоской монохроматической световой волны конец вектора $\vec{E}$ в плоскости $z = \text{const}$ описывает эллипс. Аналогично ведет себя и вектор напряженности магнитного поля. Такая волна называется

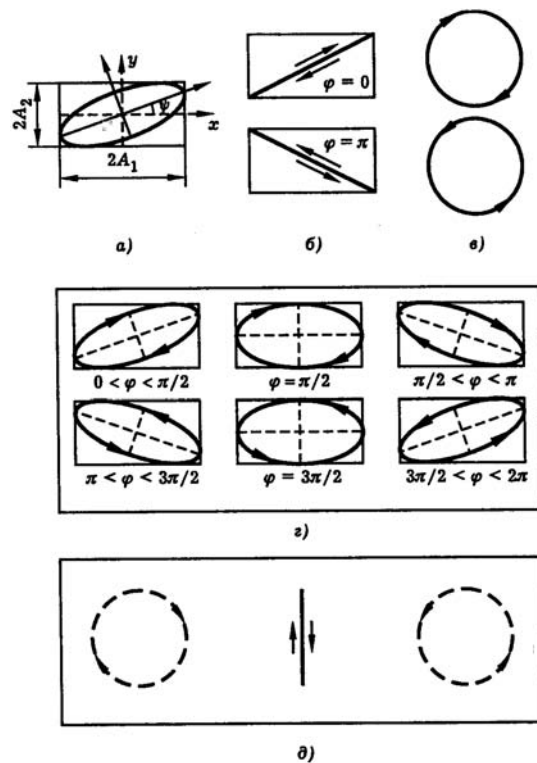


Рис. 2.3. Состояние поляризации плоской гармонической волны

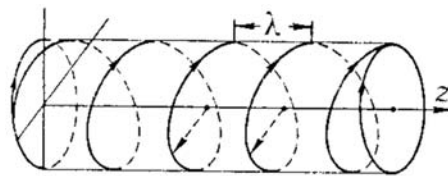


Рис. 2.4. Пространственная структура эллиптически поляризованной волны

эллиптически поляризованной. Двигаясь по эллипсу в плоскости $z = 0$, конец вектора $\vec{E}$ может вращаться по часовой или против часовой стрелки. Поэтому говорят о *правой поляризации* (для наблюдателя, смотрящего навстречу световому лучу, вращение $\vec{E}$ происходит по часовой стрелке) и *левой поляризации* (вращение $\vec{E}$ происходит в противоположном направлении). Направление вращения зависит от знака разности фаз $\varphi = (\varphi_2 - \varphi_1)$: для право поляризованной волны $0 < \varphi < \pi$, а для лево поляризованной — $-\pi < \varphi < 0$.

На **Рис.2.4** показана пространственная структура эллиптически поляризованной световой волны с правой круговой поляризацией. Винтовая линия, изображенная на рисунке, есть геометрическое место точек концов вектора $\vec{E}$ в различных плоскостях в фиксированный момент времени. Шаг винта равен длине волны λ . С течением времени изображенная винтовая линия перемещается, не деформируясь, в направлении распространения волны, вдоль оси z . В каждой фиксированной плоскости $z = 0$ вектор $\vec{E}$ вращается по часовой стрелке, если смотреть навстречу лучу.

Линейная поляризация. Если разность фаз $\varphi = (\varphi_2 - \varphi_1) = m\pi$, где $m = 0, \pm 1, \pm 2, \dots$ то эллипс переходит в прямую, описываемую уравнением

$$\frac{E_x}{E_y} = (-1)^m \frac{A_1}{A_2} \quad (2.26)$$

В этом случае волна является *линейно* или *плоско поляризованной*. На **Рис.2.3** б показаны два возможных направления поляризации в плоско поляризованной волне, соответствующие $\varphi = 0$ и $\varphi = \pi$.

Круговая поляризация. Если $A_1 = A_2 = A$ и $\varphi = m\pi/2$, где $m = \pm 1, \pm 3, \pm 5, \dots$ то одна из компонент вектора $\vec{E}$ проходит через максимум в тот момент, когда другая обращается в нуль. В этом случае эллипс вырождается в окружность, которая описывается уравнением

$$E_x^2 + E_y^2 = A^2, \quad (2.27)$$

т.е., конец вектора $\vec{E}$ движется по окружности, вращаясь по часовой или против часовой стрелки. Такое состояние поляризации волны называют *круговой* или *циркулярной поляризацией* (**Рис.2.3** в). Различают *правую* и *левую круговые поляризации*. Для *правой поляризации*

$$\varphi = \pi/2 + 2m\pi, \quad \vec{E}^+ = \vec{x}_0 E_x^+ + \vec{y}_0 E_y^+, \quad (2.28)$$

где

$$\vec{E}_x^+ = A \cos(\psi + \varphi_1), \quad \vec{E}_y^+ = -A \sin(\psi + \varphi_1) \quad (2.29)$$

Для *левой поляризации*

$$\varphi = -\pi/2 + 2m\pi, \quad \vec{E}^- = \vec{x}_0 E_x^- + \vec{y}_0 E_y^-, \quad (2.30)$$

где

$$\vec{E}_x^- = A \cos(\psi + \varphi_1), \quad \vec{E}_y^- = A \sin(\psi + \varphi_1) \quad (2.31)$$

Из формул (2.28)–(2.31) следует, что

$$\vec{E}_x = \vec{E}_x^+ + \vec{E}_y^- = \vec{x}_0 2A \cos(\psi + \varphi_1), \quad (2.32)$$

т.е., *сумма право- и лево-поляризованных волн дает линейно поляризованную волну*.

Естественно поляризованный свет. В плоской монохроматической световой волне напряженность электрического поля есть регулярная функция координат и времени. Такую волну называют *полностью поляризованной* или просто *поляризованной*. В общем случае волна поляризована эллиптически, характеристики эллипса поляризации определяются амплитудами и фазами ортогональных компонент светового поля E_x, E_y .

Конечная апертура реальных световых пучков и немонохроматичность света приводят к отличиям от этой идеальной картины. Если свет лазера бывает близок по своей структуре к поляризованной волне, то поляризация излучения нелазерного источника света, как правило, испытывает быстрые хаотические изменения во времени. Поле немонохроматической световой волны естественно рассматривать как случайный процесс. Для такой волны направление вектора $\vec{E}$ в плоскости фронта волны случайным образом меняется с течением времени. Если при этом все направления $\vec{E}$ оказываются *равновероятными*, то свет называется *неполяризованным* или *естественно поляризованным*. Таков, например, солнечный свет или свет лампы накаливания. Если же существует преимущественное направление вектора $\vec{E}$, то говорят, что свет *частично поляризован*.

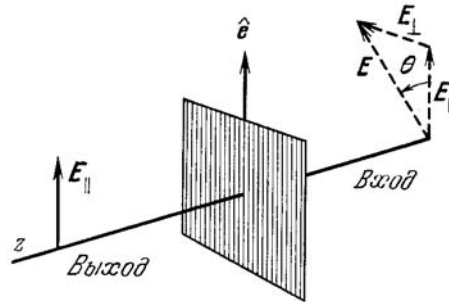


Рис. 2.5. Прохождение света через поляроид. Ось пропускания направлена по $\vec{e}$. $\vec{E}_{\parallel}$ -составляющая $\vec{E}$, параллельная $\vec{e}$, проходит. Вторая составляющая $\vec{E}_{\perp}$, перпендикулярная $\vec{e}$, полностью поглощается.

Как получить поляризованное излучение? Существует много разных способов получения поляризованного света. Устройства, которые позволяют получать поляризованное излучение, называют *поляризаторами*. Простейшим примером поляризатора является *поляроид*. Он состоит из тонкого слоя маленьких кристаллов герпатита (соль йода и хинина), выстроенных своими осями параллельно друг другу. Эти кристаллы *поглощают* свет, когда *колебания происходят вдоль оси*, в направлении которой выстроены кристаллы, и *почти не поглощают света*, когда колебания совершаются в *перпендикулярном* направлении. Направление, перпендикулярное направлению, вдоль которого выстроены молекулы, называют *осью свободного пропускания* поляроида. Если поле $\vec{E}$ направлено вдоль этой оси, то свет проходит практически без поглощения.

Направим на поляроид свет, поляризованный под углом θ к оси свободного пропускания. Какая интенсивность будет у пучка, прошедшего через поляроид? Разложим наш пучок света на две компоненты, как показано на **Рис.2.5**: одну на направление, колебания вдоль которого поглощаются (она пропорциональна $\sin \theta$), и вторую — ей перпендикулярную (продольная компонента), пропорциональную $\cos \theta$. Через поляроид пройдет только часть, пропорциональная $\cos \theta$; компонента $E_{\perp}$ пропорциональная $\sin \theta$, поглотится. Амплитуда света, прошедшего через поляроид $E_{\parallel}$, меньше амплитуды падающего света, так как $E_{\parallel} = E \cos \theta$. Таким образом, интенсивность прошедшего через поляроид света I_e связана с интенсивностью падающего света $I_i \sim E_{\parallel}^2$ следующим соотношением:

$$I_e = I_i \cos^2 \theta, \quad (2.33)$$

которое называют *законом Малюса*.

Таким образом, поляроид можно использовать для получения поляризованного света, так как только компонента, параллельная оси свободного пропускания, проходит через него. Поляроид также используют как *анализатор* для того, чтобы определить, (1) *является ли свет поляризованным* и (2) *какова плоскость его поляризации*. Если свет поляризован, то максимальное пропускание будет в случае, когда плоскость поляризации параллельна оси поляроида, а минимальное, когда перпендикулярна ей. В случае неполяризованного света интенсивность прошедшего света $I_e = 1/2 I_i$ при любом положении оси поляроида (докажите самостоятельно).

Частично поляризованный свет количественно характеризуют величиной, называемой *степенью поляризации*, которая определяется как

$$P = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}, \quad (2.34)$$

где I_{max} и I_{min} — максимальное и минимальное пропускание анализатора, соответственно.

Очевидно, что два поляроида с осями, расположенными перпендикулярно друг другу, не пропускают света. Но если между ними поместить третий поляроид, ось которого направлена под углом 45° к осям двух других, часть света пройдет через такую систему, хотя известно, что поляроид только поглощает свет, создавать свет он не может. Тем не менее, поставив третий поляроид под углом 45° , мы увеличиваем количество прошедшего света. Это похоже на фокус. В чем здесь дело? Проанализируйте это явление самостоятельно.

ЛЕКЦИЯ №3

Квазимонохроматические и квазиплоские волны. Спектральное разложение светового поля. Световые волны в веществе: уравнения Максвелла в среде; материальные уравнения, волновое уравнение для света в среде; плоская монохроматическая световая волна в линейной однородной и изотропной среде, комплексная диэлектрическая проницаемость, линейная оптическая восприимчивость и комплексный показатель преломления среды.

3.1. Квазимонохроматические и квазиплоские волны. Реальные источники света генерируют излучение, которое, на первый взгляд, имеет мало общего с рассмотренными плоскими и сферическими волнами — *эталонными волнами*, являющимися *идеальными моделями* или *математической абстракцией*. Это в первую очередь относится к характеру изменения волнового возмущения в пространстве. Так, световой пучок гелий-неонового лазера — это, конечно, не плоская и не сферическая волна. Для многих лазерных и, в особенности, нелазерных источников модель гармонического изменения во времени напряженностей электрического и магнитного полей или *монохроматической волны* также далека от *реальности*. В случае светового импульса, когда длительность световой волны *ограничена* во времени (например импульс рубинового или неодимового лазера), мы имеем *негармонический* (немонохроматический) волновой процесс, имеющий конечную длительность.

Как описать реальные световые пучки и импульсы и каковы закономерности их распространения в пространстве? Оказывается, что в основу описания реальных световых волн могут быть положены плоская и сферическая гармонические волны.

Квазигармонические волны. На практике, как правило, мы имеем дело со световыми волнами, *амплитуда и фаза которых изменяется во времени* или имеют вид *импульсов* — излучение с ограниченной длительностью, а не с гармонической волной. Говорят, что *реальные световые источники излучают волны, модулированные по времени*.

Такие световые поля в плоскости $z = 0$ можно представить в виде:

$$E(t, 0) = A(t) \cos \omega t, \quad (3.1)$$

если амплитуда $A(t)$ — *медленно меняющаяся функция времени* (по сравнению с косинусом). Это означает, что на масштабе периода световых колебаний $T = 2\pi/\omega$ амплитуда изменяется незначительно. Этот источник возбудит плоскую *амплитудно модулированную волну*

$$E(t, z) = A(t - z/c) \cos[\omega(t - z/c)], \quad (3.2)$$

являющуюся решением одномерного волнового уравнения (1.10).

В более общем случае во времени изменяется не только амплитуда, но и фаза колебаний источника.

$$E(t, 0) = A(t) \cos[\omega t - \varphi(t)]. \quad (3.3)$$

Таким образом, в пространстве возбуждается волна, имеющая *амплитудную и фазовую модуляцию* по времени.

Плоскую немонахроматическую волну в вакууме можно записать в виде

$$E(t, z) = A(t - z/c) \cos[\omega(t - z/c) + \varphi(t - z/c)], \quad (3.4)$$

если амплитуда $A(t)$ и фаза $\varphi(t)$ — *медленно меняющиеся функции времени* (по сравнению с косинусом). Это означает, что на масштабе периода световых колебаний $T = 2\pi/\omega$ амплитуда изменяется незначительно.

Изменения функций $A(t)$ и $\varphi(t)$ могут быть достаточно быстрыми. Например, современные лазеры способны генерировать световые импульсы длительностью $\tau \sim 10^{-12} - 10^{-14}$ с (*пико- и фемтосекундные импульсы*). Тем не менее в подавляющем большинстве случаев соответствующие изменения амплитуды и фазы остаются медленными в масштабе периода световых колебаний $T = 2\pi/\omega$. В видимом диапазоне $T \simeq 10^{-13} - 10^{-15}$ с, и даже для пикосекундных импульсов ($1 \text{ пс} = 10^{-12}$ с) выполняется неравенство $\tau \gg T$. Свойства таких *медленно модулированных во времени волн близки к свойствам гармонической волны*. Поэтому их называют *квазигармоническими* или *квазимонохроматическими*.

Прямой подстановкой можно проверить, что в рассматриваемом приближении (3.4) удовлетворяют волновому уравнению (1.10). Вид решения (3.4) показывает, что *форма временной модуляции плоской волны остается неизменной при ее распространении в свободном пространстве*.

Колебания типа (3.3) на комплексной плоскости изображаются в виде вектора с медленно изменяющимися длиной и направлением. Этому вектору соответствует комплексная функция $\mathcal{E} = 1/2 A(t) \exp[i\varphi(t)]$, которая является аналогом комплексной амплитудой гармонического процесса. Ее называют *комплексной амплитудой квазимонохроматического процесса*.

Квазиплоские волны. На практике невозможно создать источник света, генерирующий волны с идеально плоским или сферическим волновым фронтом. Причина состоит в том, что *конечная пространственная протяженность* или, как говорят, *конечная апертура источника* приводит (из-за дифракции) к тому, что *амплитуда световых колебаний изменяется в плоскости, перпендикулярной направлению распространения света*. Другими словами, возникает *пространственно модулированная волна*, у которой значения амплитуды и фазы зависят от поперечных координат, т. е. имеет место ситуация, принципиально отличающаяся от плоской волны. Тем не менее, достаточно часто плоскую волну можно рассматривать не только как математическую абстракцию.

Действительно, из физических соображений можно заключить, что если поперечный размер (апертура) источника света равна d , то волновой процесс будет обладать свойствами плоской волны не только при $d \rightarrow \infty$, но и для конечных, но достаточно больших апертур. "Большой апертурой" считается апертура, у которой *линейный размер много больше длины волны* ($d \gg \lambda$). Применительно к световому излучению это условие практически выполняется всегда, так как, например, в видимом диапазоне $\lambda \simeq 10^{-4}$ см, а диаметр типичного лазерного пучка $d \simeq 0.1$ см и более. Следовательно, отношение $d/\lambda \approx 10^3 \gg 1$. В таких широкоапертурных пучках *амплитуда и фаза медленно (на масштабе длины световой волны) меняются в плоскости, перпендикулярной направлению распространения света, а волновые фронты на значительных расстояниях мало отличаются от плоской*. Поэтому можно с большой степенью точности считать что, мы имеем дело с *волнами близкими к плоским* или, как говорят, с *квазиплоскими волнами*, т.е. почти плоскими.

В случае гармонической квазиплоской волны поле на выходной апертуре источника можно записать в виде

$$E(t, x, y) = A(x, y) \cos \omega t, \quad (3.5)$$

где волновой фронт не зависит от z (плоский), а *амплитуда волны медленно меняется в плоскости x, y* с характерным масштабом изменения d . Можно ожидать, что если $d \gg \lambda$, то существует область *не слишком малых расстояний z* ($z \gg \lambda, z > d$), для которых волна, возбуждаемая источником вида (3.5), остается *квазиплоской* (почти плоской):

$$E(t, x, y, z) = A(x, y) \cos(\omega t - kz). \quad (3.6)$$

Более точный критерий применимости модели (3.6) удастся установить, лишь обращаясь к решению волнового уравнения (1.10) с граничным условием (3.5).

Для примера рассмотрим распространение так называемого *гауссова светового пучка*. Пусть при $z = 0$ волновой фронт плоский, а распределение амплитуды описывается гауссовой кривой:

$$E(t, x, y, z = 0) = A_0 \exp[-(r^2/d^2)] \cos \omega t, \quad (3.7)$$

где $r^2 = x^2 + y^2$, d — радиус пучка.

Как можно показать (см., например, С.А.Ахманов, С.Ю.Никитин "Физическая оптика"), приближенное решение волнового уравнения с таким граничным условием можно представить в виде:

$$E(t, r, z = 0) = \frac{1}{2} A_k(r, z) \exp[i(\omega t - kz)] + K.C., \quad (3.8)$$

где $A_k(r, z)$ — комплексная амплитуда волны, определяемая формулой:

$$A_k(r, z) = \frac{A_0}{(1 - iz/z_d)} \exp\left[-\frac{r^2}{(1 - iz/z_d)d^2}\right]. \quad (3.9)$$

$$z_d = kd^2/2 \quad (3.10)$$

Выражение (3.9) удобно переписать как $A_k(r, z) = A \exp(-i\varphi)$, где A и φ — вещественные амплитуда и фаза волны.

$$A(r, z) = \frac{A_0}{\sqrt{1 + (z/z_d)^2}} \exp\left[-\frac{r^2}{1 + (z/z_d)^2}\right], \quad (3.11)$$

$$\varphi(r, z) = \frac{r^2}{d^2} \frac{z/z_d}{1 + (z/z_d)^2} - \arctg(z/z_d). \quad (3.12)$$

Согласно (3.11) и (3.12) *распределение амплитуды и фазы* в поперечном сечении пучка определяется только одним параметром — отношением пути z , пройденного волной, к характерной длине $z_d = kd^2/2 = \pi d^2/\lambda$, который называют *дифракционной длиной* светового пучка. При $z \ll z_d$ выражение для светового поля упрощается и приобретает вид:

$$E(t, r, z) = A_0 \exp(-r^2/d^2) \cos(\omega t - kz), \quad (3.13)$$

т.е. мы имеем волну вида (3.6), у которой амплитуда имеет гауссовское распределение в поперечном сечении.

Чем больше величина z_d , т. е. чем больше отношение d/λ , тем больше расстояние z , на котором сохраняется плоский фазовый фронт. Например, для излучения гелий-неонового лазера с длиной волны $\lambda = 0,63$ мкм и радиусом пучка $d = 0.1$ см дифракционная длина $z_d = \pi d^2/\lambda = 55$ м. Расширяя пучок $d = 1$ см, получим $z_d = 500$ м.

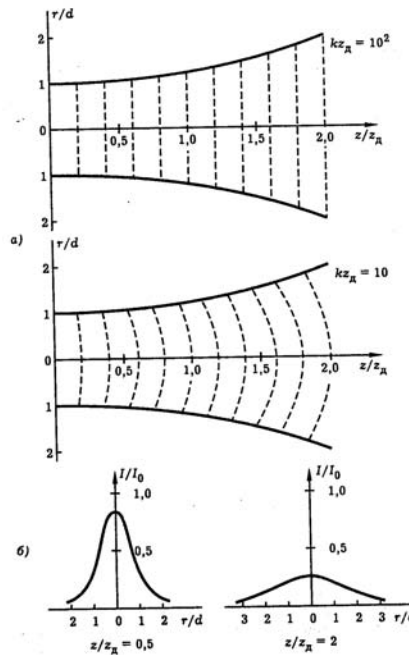


Рис. 3.1. Распространение гауссова пучка. Контур пучка в продольном сечении (сплошные кривые) и волновые фронты (пунктир) (а); поперечный профиль интенсивности излучения (б)

Замечательное свойство гауссова пучка состоит в том, что *он сохраняет свою форму* и при $z > z_d$. При распространении гауссовского пучка изменяется лишь ширина пучка. В этой области, однако, пучок приобретает свойства сферической волны. Из (3.11) можно показать, что при $z \gg z_d$ амплитуда волны убывает пропорционально пройденному расстоянию $A(r, z) \simeq A_0 z_d / z \sim 1/z$, а фазовый фронт становится сферическим $\varphi(r, z) \sim r^2/d^2$. Указанные свойства гауссова пучка иллюстрирует **Рис. 3.1**, на котором показаны изменение поперечного размера и искривление волнового фронта пучка при его распространении в свободном пространстве.

С какой степенью точности можно приблизиться на практике к идеальной монохроматической волне? На первый взгляд может показаться, что речь идет о технической проблеме. Действительно, казалось бы повышение стабильности параметров, например, одночастотного лазера непрерывного действия может привести к генерации оптических колебаний, сколь угодно близких к гармоническим. На самом же деле имеется *принципиальный предел монохроматичности*, определяемый квантовыми флуктуациями в излучающих атомах и молекулах, которые и оказываются причиной неустраняемых амплитудной и фазовой модуляций волны. Хотя связанная с квантовыми флуктуациями немонохроматичность излучения относительно невелика — отклонение частоты от средней не превышает 10^2 Гц. Но с ними приходится считаться в квантовых стандартах частоты и времени, прецизионных оптических экспериментах. Обращаясь же к принципиальной стороне дела, можно сказать, что *идеальная плоская волна есть такая же абстракция, как и идеальная монохроматическая волна. Идеальная монохроматическая плоская волна не осуществима.*

3.2. Спектральное разложение светового поля. Произвольную волну можно представить в виде суперпозиции эталонных волн, или, как говорят, *разложить его в спектр*, т.е. выполнить *спектральное разложение*. Особое значение в оптике

имеет *разложение волновых пучков и импульсов по плоским гармоническим волнам*. Такое разложение оказывается не только удобной математической операцией, оно **фактически осуществляется** в реальном оптическом эксперименте.

Один из классических примеров такого рода — знаменитый опыт Ньютона, в котором наблюдалось *разложение света в спектр* с помощью стеклянной призмы. Заключение Ньютона о том, что "солнечный свет состоит из лучей различной преломляемости", на современном математическом языке спектральных разложений означает, что поле *плоской немонохроматической волны* $E(t, z)$ можно представить в виде *суперпозиции плоских мономонохроматических волн с разными частотами*.

Действительно, из математики известно, что произвольную *непериодическую функцию* $f(t)$ можно записать в виде *интеграла Фурье*:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) \exp(i\omega t) d\omega, \quad (3.14)$$

т.е. разложить в спектр по гармоническим колебаниям или, как говорят, в *частотный спектр* $F_\omega = F(\omega) \exp(i\omega t)$. Формулу (3.14) также называют *спектральным разложением*.

Формула (3.14) "говорит", что функцию $f(t)$ можно представить как суперпозицию (сумму) большого числа колебаний, амплитуда которых определяется *спектральной амплитудой* $F(\omega)$ или *фурье-компонентой*, определяемой обратным Фурье-преобразованием

$$F(\omega) = \int_{-\infty}^{\infty} f(t) \exp(-i\omega t) dt, \quad (3.15)$$

Так как $f(t)$ — действительная функция, то нетрудно показать, что $F(-\omega) = F^*(\omega)$.

Таким образом можно сказать, что каждая гармоническая компонента заданного при $z = 0$ возмущения $E(t)$ возбуждает мономонохроматическую световую волну

$$E_\omega(t, z) = A(\omega) \exp[i(\omega t - kz - \varphi(\omega))]. \quad (3.16)$$

Формула (3.16) наряду с (3.4) дает еще один способ описания *немонохроматической плоской волны* — в виде *суперпозиции плоских мономонохроматических волн*. Принято говорить, что выражение (3.4) описывает *немонохроматическую волну во временном представлении*, а (3.16) — в *спектральном представлении*.

Периодическую функцию $f(t) = f(t + T)$ можно разложить в ряд Фурье

$$f(t) = \sum_{-\infty}^{\infty} F(n) \exp(in\omega t), \quad (3.17)$$

где $\omega = +2\pi/T$, а коэффициенты Фурье $F(n)$ определяются

$$F(n) = \frac{1}{T} \int_{-\infty}^{\infty} f(t) \exp(-in\omega t) dt. \quad (3.18)$$

В этом случае, в отличие от (3.15), спектр является *дискретным*.

В общем случае спектральная амплитуда $F(\omega)$ является комплексной функцией частоты: $F(\omega) = |F(\omega)| \exp[i\psi(\omega)]$. Модуль комплексной спектральной амплитуды представляет собой действительную амплитуду гармоники с частотой ω в спектре $f(t)$, а аргумент $\psi(\omega)$ — действительную фазу этого колебания. Комплексность спектральной амплитуды связана с тем, что разные гармоники, образующие в совокупности процесс $f(t)$, имеют, вообще говоря, различные фазы. Часто такая полная спектральная информация об оптическом процессе бывает не нужна.

Более того, в оптике ее трудно экспериментально получить. На практике обычно измеряют величину $S(\omega) = |F(\omega)|^2$, которую называют *спектральной плотностью* или *контуром спектральной линии*. В этом выражении информация о фазах гармонических колебаний, составляющих процесс $f(t)$, утрачена. Тем не менее, для оптики спектральная плотность является важной характеристикой процессов, поскольку именно она обычно измеряется в экспериментах. Доказано, имеет место следующее соотношение

$$\int_{-\infty}^{\infty} f^2(t)dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega)d\omega, \quad (3.19)$$

которое называют *равенством Парсеваля*. Это равенство отражает взаимосвязь между временным и спектральным представлениями.

Используя равенство Парсеваля, можно записать

$$\int_{-\infty}^{\infty} E^2(t)dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} |f(\omega)|^2 d\omega = \frac{1}{\pi} \int_0^{\infty} |f(\omega)|^2 d\omega. \quad (3.20)$$

Формула (3.20) отражает тот факт, что *полная энергия немонахроматических световых волн выражается через интеграл по положительным частотам от ее спектральной плотности $|E(\omega)|^2$* , которая характеризует распределение энергии света по спектру. Именно эта характеристика света наблюдается и может быть **экспериментально** измерена. Она оказывается пропорциональной энергии светового импульса, прошедшей через площадку единичной площади в окрестности данной точки.

Подчеркнем различие между $f(\omega)$ и $|f(\omega)|^2$: знание $f(\omega)$ полностью определяет функцию $E(t)$, описывающую немонахроматическую волну, тогда как знание $|E(\omega)|^2$ еще не позволяет восстановить $E(t)$, так как в ней не содержится информация о фазах монахроматических составляющих. Однако именно последней можно сопоставить реальный *физический процесс* — экспериментальное измерение спектра такой волна с помощью соответствующего спектрального прибора.

Для примера рассмотрим спектральный состав немонахроматической световой волны, представляющей отрезок косинусоиды длительностью τ (прямоугольный световой импульс):

$$E(t) = \begin{cases} E_0 \cos \omega_0 t, & -\tau/2 < t < \tau/2; \\ 0, & |t| \geq \tau/2. \end{cases} \quad (3.21)$$

Используя формулу (3.14), получаем:

$$E(\omega) = \frac{\tau}{2} E_0 \left[\frac{\sin(\omega_0 - \omega)\tau/2}{(\omega_0 - \omega)\tau/2} + \frac{\sin(\omega_0 + \omega)\tau/2}{(\omega_0 + \omega)\tau/2} \right]. \quad (3.22)$$

Вторым слагаемым в (3.22) можно пренебречь, так как $|\omega_0 - \omega| \ll |\omega_0 + \omega|$. На **Рис.3.2** показан график функции (3.22). Ее *главный максимум* расположен при $x = (\omega_0 - \omega)\tau/2 = 0$, где она имеет значение, равное 1, и она обращается в нуль при $x = \pm\pi, \pm 2\pi, \dots$ В промежутках она имеет второстепенные максимумы и минимумы. Таким образом, для прямоугольного импульса света фурье-компоненты в основном сосредоточены в главном максимуме вблизи значения ω_0 . *Контур спектральной линии* имеет вид

$$|E(\omega)|^2 = \frac{\tau^2}{4} E_0^2 \frac{\sin^2(\omega_0 - \omega)\tau/2}{[(\omega - \omega_0)\tau/2]^2}. \quad (3.23)$$

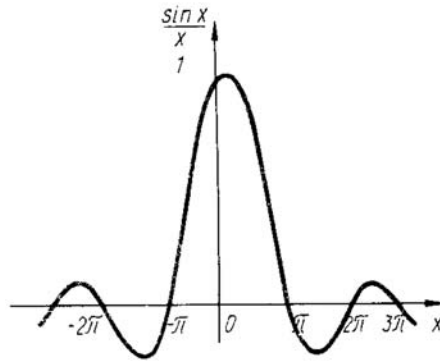


Рис. 3.2. График функции $\sin x/x$

Контур спектральной линии характеризуют *шириной спектральной линии*, которая определяется как интервал частот $\Delta\omega = \omega_1 - \omega_2$, где $\omega_{1,2}$ ($\omega_1 > \omega_2$) значения частот, при которых спектральная плотность равна половине максимальной плотности — *полуширина* спектральной линии. Можно показать, что между *длительностью импульса* и *шириной его спектра* имеет место следующее соотношение:

$$\Delta\omega \cdot \tau \approx 2\pi. \quad (3.24)$$

Оно имеет *универсальный* характер, так как справедливо для произвольного импульса независимо от его конкретного вида. Из формулы (3.24) следует, что *ширина спектральной линии* $\Delta\omega$ *приблизительно равна обратной длительности импульса* τ^{-1} . Такие импульсы называют *спектрально ограниченными*. Для квазимонохроматической волны $\Delta\omega \ll \omega_0$, где ω_0 несущая частота.

Подчеркнем, что *чем короче импульс, тем шире его частотный спектр*. Применительно к оптике это означает, что чем *плотнее* сконцентрирована энергия светового поля по времени, тем *шире* она распределена по спектру. В частности, *предельно* короткие импульсы света, длительность которых соизмерима с периодом световых колебаний, обладают частотным спектром, ширина которого соизмерима с несущей частотой световой волны. Такие импульсы нельзя считать квазимонохроматическими.

Спектральные разложения естественным образом обобщаются и на волновые пучки — пространственно модулированные волны. Волновой пучок тоже можно представить в виде суперпозиции плоских волн, но теперь речь идет о *разложении по волнам, распространяющимся в различных направлениях*. Различные спектральные компоненты в таком разложении характеризуются углами между направлением распространения волны и координатным осями. Поэтому здесь говорят об *угловом спектре пространственно модулированной волны*. Отметим, что разложение в угловой спектр физически происходит в очень простых по постановке опытах. В дальнейшем мы увидим, что форма углового спектра светового пучка определяется распределением освещенности на экране, расположенном на достаточно большом удалении от источника света ($z \gg z_d$), либо в фокальной плоскости линзы. В таких опытах свободное пространство и линза выполняют такую же операцию фурье-разложения по отношению к угловому спектру, что призма по отношению к частотному.

3.3. Световые волны в веществе. Уравнения Максвелла в среде. Как известно, вещество состоит из атомов. В обычных условиях число атомов столь велико, и они

расположены настолько близко друг к другу, что дискретная структура среды, как правило, не проявляется. Поэтому *вещество* можно рассматривать как *сплошную (непрерывную) среду*. Простейшая *физическая картина взаимодействия света с веществом* представляется следующим образом. Падающий на среду свет раскачивает колебания электронов в атомах, которые становятся источниками *вторичных световых волн*. Полное световое поле формируется в результате сложения падающего излучения и испускаемых вторичных волн всеми атомами.

Теория взаимодействия света и вещества строится на основе *уравнений Максвелла и материальные уравнений*, которые описывают свойства вещества. Система уравнений Максвелла в среде имеет вид

$$\operatorname{rot} \vec{E} = -\frac{1}{c} \frac{\partial \vec{B}}{\partial t}, \quad \operatorname{div} \vec{D} = 4\pi\rho \quad (3.25)$$

$$\operatorname{rot} \vec{H} = \frac{1}{c} \frac{\partial \vec{D}}{\partial t} + \frac{4\pi}{c} \vec{j}, \quad \operatorname{div} \vec{B} = 0. \quad (3.26)$$

$$\vec{P} = \vec{P}(\vec{E}), \quad \vec{M} = \vec{M}(\vec{H}), \quad \vec{D} = \vec{E} + 4\pi\vec{P}, \quad \vec{B} = \vec{H} + 4\pi\vec{M}, \quad \vec{j} = \vec{j}(\vec{E}), \quad (3.27)$$

где $\vec{E}$ и $\vec{D}$ — *напряженность и индукция электрического поля*; $\vec{H}$ и $\vec{B}$ — *напряженность и индукция магнитного поля*; $\vec{P}$, $\vec{M}$ — *поляризованность* (дипольный момент единицы объема среды) и *намагниченность* (магнитный момент единицы объема), ρ и j — *плотность заряда и тока* в среде, соответственно, c — *скорость света в вакууме*. Уравнения записаны в гауссовой системе единиц.

Материальные уравнения. Уравнения (3.27) называют *материальными уравнениями*. Они описывают *связь характеристик вещества с электромагнитным полем*. Их физический смысл состоит в том, что они описывают *отклик среды на электромагнитное поле*, а именно, возникновение *электрического дипольного момента* единицы объема $\vec{P}$, *магнитного момента* единицы объема $\vec{M}$, *тока проводимости* $\vec{j}$ (если среда проводящая).

Хотя уравнения Максвелла в среде имеют вид, подобный уравнениям в вакууме, тем не менее имеются определенные различия, которые состоят в следующем. Физические величины, входящие в уравнения Максвелла в среде, — это *средние значения*. Усреднение производится по элементам объема, содержащим *макроскопически большое число* атомов или молекул, т.е. большим по сравнению со средним расстоянием между частицами. В то же время *линейный размер этих элементов должен быть много меньше характерного размера макроскопических неоднородностей*, мерой которых служит длина волны электромагнитной волны. Такие элементы объема называют *физически бесконечно малыми*. Для видимого света это порядка 10^{-7} м. В объеме $(10^{-7}\text{м})^3 = 10^{-21}$ м³ содержится примерно 10^8 атомов в конденсированном веществе и $3 \cdot 10^4$ молекул любого газа при нормальных условиях. Именно в *таком приближении вещество представляется непрерывным и однородным*.

Как ведут себя поляризованность, намагниченность и ток проводимости в среде, возбуждаемые световой волной? Это *центральный вопрос* физики взаимодействия света с веществом. Разработаны различные модели, позволяющие определить материальные уравнения. Отметим, что много принципиальных результатов было получено

здесь в последние сорок лет или, как говорят, в "лазерную эпоху". Физика взаимодействия света с веществом и по сей день остается одним из наиболее динамичных разделов физической оптики.

Волновое уравнение для света в среде. Для простоты рассмотрим *диэлектрические, нейтральные и немагнитные* среды, в которых $\vec{j} = 0$, $\rho = 0$, $\vec{B} = \vec{H}$. Рассматриваемый класс сред довольно широк. К нему относятся, например, воздух, вода, кристаллы, стекла, пластмассы и т. д. Для таких сред материальные уравнения имеют вид:

$$\vec{D} = \vec{E} + 4\pi\vec{P}, \quad \vec{P} = \vec{P}(\vec{E}) \quad (3.28)$$

а уравнения Максвелла

$$\text{rot } \vec{E} = -\frac{1}{c} \frac{\partial \vec{H}}{\partial t}, \quad \text{div } \vec{D} = 0 \quad (3.29)$$

$$\text{rot } \vec{H} = \frac{1}{c} \frac{\partial \vec{D}}{\partial t}, \quad \text{div } \vec{H} = 0. \quad (3.30)$$

Из (3.28) и (3.29) нетрудно вывести уравнение

$$\text{rot rot } \vec{E} + \frac{1}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} = -\frac{4\pi}{c^2} \frac{\partial^2 \vec{P}}{\partial t^2}, \quad (3.31)$$

которое называют *волновым уравнением для света в веществе*. Здесь, в отличие от вакуума, правая часть не равна нулю. Это означает, что *оптическая поляризованность среды является источником светового поля*.

Возникновение поляризованности можно понять следующим образом. При воздействии светового поля на электроны в веществе последние смещаются из положения равновесия, приобретая дипольный момент $\vec{p}$. Полный дипольный момент $\vec{P}$, называемый поляризованностью, равен сумме всех диполей $\vec{p}$ в единице объема вещества. Поскольку поле периодическое, движение зарядов тоже периодическое и ускоренное, то $\vec{P}$ и $\vec{p}$ излучают. Общее поле $\vec{E}$ есть сумма поля падающей волны и полей, испускаемых всеми диполями. Поляризованность $\vec{P}$ по фазе может отличаться от $\vec{E}$.

Уравнение (3.31) и второе уравнение в (3.28) образуют в совокупности *замкнутую систему уравнений, которая, будучи дополненной соответствующими начальными и граничными условиями, полностью определяет процесс распространения света в диэлектрической нейтральной немагнитной среде*.

Используя $\text{div } \vec{D} = 0$ (а значит и $\text{div } \vec{E} = 0$) для изотропных сред нетрудно показать, что

$$\Delta \vec{E} - \frac{4\pi}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} = -\frac{1}{c^2} \frac{\partial^2 \vec{P}}{\partial t^2}, \quad (3.32)$$

Мы в основном будем рассматривать *однородные, линейные и изотропные среды*. Для однородной среды ее характеристики для всей среды одинаковы. т.е. не зависят от координат.

Линейными называются среды, в которых связь между поляризованностью и напряженностью электрического поля является линейной функцией. Такое приближение справедливо, когда напряженность поля световой волны много меньше внутриаомных полей. В противном случае среды называют *нелинейными* и для них зависимость $\vec{P}$ от $\vec{E}$ имеет нелинейный характер.

Изотропными называют среды, свойства которых не зависят от направления распространения света. В таких средах *поляризованность ориентирована параллельно*

электрическому полю ($\vec{P} \parallel \vec{E}$). Среда считается *анизотропной*, когда вектор поляризованности не параллелен вектору поля. Свойства такой среды зависят от направления распространения волны.

Плоская монохроматическая световая волна в линейной однородной и изотропной среде. Пусть в некоторой линейной среде распространяется плоская монохроматическая световая волна

$$\vec{E} = \frac{1}{2} \vec{\mathcal{E}} \exp[i(\omega t - \vec{k}\vec{r})] + K.C \quad (3.33)$$

Здесь $\vec{\mathcal{E}}$ — комплексная амплитуда волны, ω — частота, $\vec{k}$ — волновой вектор. В силу линейности уравнений Максвелла и материальных уравнений, магнитное поле $\vec{H}$, электрическая индукция $\vec{D}$ и оптическая поляризованность $\vec{P}$ также будут иметь вид, подобный (3.33). Обозначим их комплексные амплитуды через $\vec{\mathcal{H}}$, $\vec{\mathcal{D}}$ и $\vec{\mathcal{P}}$, соответственно.

В комплексном представлении уравнения Максвелла принимают вид:

$$[\vec{k}, \vec{\mathcal{E}}] = \frac{\omega}{c} \vec{\mathcal{H}}, \quad (\vec{k}, \vec{\mathcal{D}}) = 0 \quad [\vec{k}, \vec{\mathcal{H}}] = -\frac{\omega}{c} \vec{\mathcal{D}}, \quad (\vec{k}, \vec{\mathcal{H}}) = 0 \quad (3.34)$$

Исключим из этих уравнений $\vec{\mathcal{H}}$. Для этого умножим первое уравнение векторно на $\vec{k}$ и, используя третье уравнение, получим:

$$[\vec{k}[\vec{k}, \vec{\mathcal{E}}]] + \frac{\omega^2}{c^2} \vec{\mathcal{D}} = 0. \quad (3.35)$$

Раскрывая двойное векторное произведение, приходим к уравнению

$$(\vec{k}, \vec{\mathcal{E}}) \vec{k} - \vec{k}^2 \vec{\mathcal{E}} + \frac{\omega^2}{c^2} \vec{\mathcal{D}} = 0. \quad (3.36)$$

Данное уравнение справедливо для *любых* линейных сред: как изотропных, так и анизотропных. Если же световая волна распространяется в вакууме, то $\vec{\mathcal{E}} = \vec{\mathcal{D}}$ и из (3.36) следует, что $k = \omega/c$. Очевидно, что соотношения (3.34) и связанные с ними справедливы и для действительных величин $\vec{E}$, $\vec{H}$, $\vec{D}$ и $\vec{P}$.

Комплексная диэлектрическая проницаемость, линейная оптическая восприимчивость и комплексный показатель преломления среды. В изотропной однородной среде векторы $\vec{E}$, $\vec{P}$ и $\vec{D}$ должны быть параллельны, но тогда должны быть параллельны и векторы комплексных амплитуд $\vec{\mathcal{E}}$, $\vec{\mathcal{P}}$ и $\vec{\mathcal{D}}$, которые не зависят от времени (для плоских волн). В общем случае для линейной, изотропной и немагнитной среды связь между $\vec{\mathcal{P}}$, $\vec{\mathcal{D}}$ и $\vec{\mathcal{E}}$ можно записать в виде:

$$\vec{\mathcal{P}} = \chi(\omega) \vec{\mathcal{E}}, \quad \vec{\mathcal{D}} = \varepsilon(\omega) \vec{\mathcal{E}}, \quad (3.37)$$

где $\chi(\omega)$ и $\varepsilon(\omega)$ — скалярные величины, которые, в общем случае являются комплексными величинами и зависят от частоты света (дисперсия). Зависимость $\chi(\omega)$ и $\varepsilon(\omega)$ от частоты обусловлена *инерционностью* среды, т.е. запаздыванием поляризации относительно поля, что и приводит к *временной* или *частотной дисперсии*. Величина $\chi(\omega)$ называется *линейной оптической восприимчивостью среды*, а $\varepsilon(\omega)$ — *комплексной диэлектрической проницаемостью*, называется .

Подставив первое уравнение из (3.37) в уравнение $(\vec{k}, \vec{\mathcal{D}}) = 0$, получим $(\vec{k}, \vec{\mathcal{E}}) = 0$. С учетом этого уравнение (3.36) принимает вид

$$k^2 = \frac{\omega^2}{c^2} \varepsilon(\omega), \quad (3.38)$$

или

$$k = \frac{\omega}{c} n(\omega), \quad (3.39)$$

где $n(\omega) = \sqrt{\varepsilon(\omega)} = \sqrt{1 + 4\pi\chi(\omega)}$ — комплексный показатель преломления среды.

Из формулы (3.39) видно, что k может быть комплексной $k = k' - ik''$, если $\varepsilon(\omega)$ комплексная или отрицательная.

Если среда является магнитной, то аналогичным образом вводится магнитная проницаемость $\mu(\omega)$, определяемая соотношением $\vec{B} = \mu(\omega)\vec{H}$. В этом случае можно показать, что показатель преломления принимает вид: $n(\omega) = \sqrt{\varepsilon(\omega)\mu(\omega)}$.

Уравнение (3.39), связывающее между собой частоту ω и волновое число k световой волны, называется *дисперсионным уравнением*. Данное уравнение справедливо для *линейных изотропных сред*. Если же световая волна распространяется в *вакууме*, то $k = \omega/c$. Формула (3.39) позволяет записать волновой вектор $\vec{k}$ для световой волны в среде в виде

$$\vec{k} = \frac{\omega}{c} n(\omega) \vec{m} = \frac{\omega}{v} \vec{m}, \quad (3.40)$$

где $\vec{m} = \vec{k}/k$ — единичный вектор, направленный вдоль $\vec{k}$, а величину $v = c/n(\omega)$ называют *фазовой скоростью* света в веществе.

Интенсивность световой волны в прозрачной, линейной и изотропной среде определяется теми же соотношениями, что и в вакууме, но в них надо скорость света c заменить на cn :

$$I = \frac{cn}{8\pi} |\vec{\mathcal{E}}|^2 \quad (3.41)$$

ЛЕКЦИЯ №4

Основные понятия фотометрии. Световые пучки и импульсы: энергия, мощность, интенсивность. Геометрическая оптика: эйконал; принцип Ферма; зеркальное отражение; преломление, закон Снеллиуса; преломление и отражение на сферической поверхности.

4.1. Основные понятия фотометрии. Фотометрия – раздел физической оптики, в котором рассматриваются *энергетические характеристики светового излучения в процессах его испускания, распространения и взаимодействия с веществом*. Основной задача в фотометрии — измерение *энергии*, приносимой световой волной, а также измерение величин, которые так или иначе связаны с ней. Здесь используются два типа величин: *энергетические величины*, характеризующие энергетические параметры безотносительно к его действию на приемники излучения, т.е. не зависящее от частоты (неселективное) действие, и *световые*, которые характеризуют *физиологическое* действие света, оцениваемое по воздействию на глаз и другие селективные приемники света (с учетом спектральной чувствительности).

Для описания энергетических характеристик света обычных (нелазерных) источников используют следующие фотометрические понятия и величины: *поток излучения, сила света, освещенность поверхности, энергетическая светимость поверхности, яркость источника света*.

Поток излучения или *поток лучистой энергии* $d\Phi$ есть *количество лучистой энергии dW , протекающей через площадку σ в единицу времени*:

$$d\Phi = dW/dt \quad (4.1)$$

или это *мощность* сквозь площадку σ . В СГС эта величина имеет размерность $\text{эрг} \cdot \text{с}^{-1}$. Он связан с интенсивностью очевидным соотношением $\Phi = \int_{\sigma} I d\sigma$.

Когда излучение распространяется в виде сферической волны от точечного источника, для характеристики пространственно-угловой плотности потока используют величину, называемую *силой света*. Сила света J в направлении, задаваемом азимутальным и орбитальным углами θ и φ в сферической системе координат, определяется как *энергетический поток в единичном телесном угле* или как отношение потока энергии $d\Phi$ к телесному углу $d\Omega$, в котором распространяется излучение:

$$J(\theta, \varphi) = d\Phi/d\Omega \quad (4.2)$$

Телесный угол — это угол внутри конуса с вершиной в точке, где расположен источник света и основанием с площадью σ , находящемся на расстоянии R от вершины:

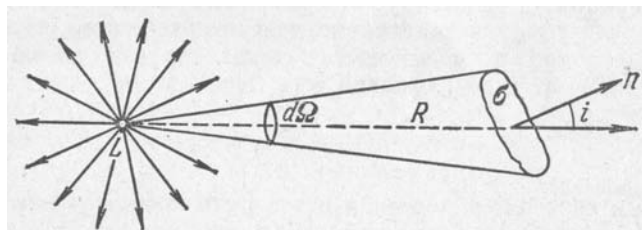


Рис. 4.1. К определению понятия потока излучения

$d\Omega = \sigma \cos i / R^2$ (**Рис. 4.1**), i — угол между осью этого конуса и нормалью к основанию конуса. В сферических координатах $d\Omega = \sin \theta d\theta d\varphi$, где θ и φ — азимутальный (широта) и меридиональный (долгота) углы. Размерность яркости — эрг·с⁻¹·стерад⁻¹.

Полный поток излучения определяется интегралом от силы света по всем направлениям:

$$\Phi = \int J d\Omega = \int_0^{2\pi} d\varphi \int_0^\pi J(\theta, \varphi) \sin \theta d\theta \quad (4.3)$$

Если источник излучает *равномерно во всех направлениях*, то сила света не зависит от θ и φ и $\Phi = 4\pi J$.

Полный поток излучения характеризует данный источник, его *нельзя изменить никакими оптическими системами*. Оптические системы лишь *перераспределяют поток в пространстве*, т.е. создают большую или меньшую концентрацию света по некоторым направлениям. При этом сила света возрастает по одним направлениям и уменьшается по другим. Если сила излучения $J(\theta)$, характеризующая данный элемент поверхности $d\sigma$ протяженного источника пропорциональна видимой части по данному направлению площади этого элемента $d\sigma \cos \theta$, то говорят, что такой источник удовлетворяет *закону Ламберта*.

Освещенность поверхности определяется как *отношение потока излучения E к площади $d\sigma$ облучаемого элемента поверхности*:

$$E = d\Phi/d\sigma = J d\Omega/d\sigma = J \cos i / R^2 \quad (4.4)$$

Размерность освещенности есть эрг·с⁻¹·см⁻². Когда излучение падает перпендикулярно поверхности, ее освещенность равна интенсивности.

Для характеристики протяженного источника вводят яркость и светимость.

Яркость источника света B определяется как *отношение силы света J к площади проекции $ds_\perp$ светящегося элемента поверхности на плоскость, перпендикулярную направлению наблюдения*:

$$B = J/d\sigma_\perp = d\Phi/d\sigma \cos i d\Omega, \quad (4.5)$$

где $d\sigma_\perp = d\sigma \cos i$ — площадь элемента поверхности источника, i — угол между нормалью к светящейся поверхности и направлением на точку наблюдения. Размерность яркости есть эрг·с⁻¹·стерад⁻¹·см⁻². Для ламбертова источника яркость одинакова во всех направлениях.

Энергетическая светимость R поверхности — это *поток излучения $d\Phi$ с единичной площади излучающей поверхности*:

$$R = d\Phi/d\sigma, \quad (4.6)$$

где $d\sigma$ элементарная площадь поверхности. Размерность светимости есть эрг·с⁻¹·см⁻².

Освещенность и светимость характеризуют *мощность света, который распространяется вблизи определенной точки пространства во всех направлениях*. Интегральной характеристикой является поток — мощность, переносимая во всех направлениях через всю заданную поверхность.

Сила света, освещенность и энергетический поток связаны с яркостью: следующими соотношениями:

$$J = \int B_i \cos i d\sigma, \quad E = \int B_i \cos i d\Omega, \quad \Phi = \int \int B_i \cos i d\sigma d\Omega. \quad (4.7)$$

Светимость связана с яркостью излучающей поверхностью следующим образом:

$$R = \int B(\theta) \cos \theta d\Omega = \int_0^{2\pi} d\varphi \int_0^{\pi/2} B(\theta) \cos \theta \sin \theta d\theta \quad (4.8)$$

Для ламбертова источника $B(\theta) = B(0) = B$. Нетрудно показать, что $R = \pi B$.

В оптических измерениях часто используются приемники излучения, которые чувствительны к длине волны света. Их называют *селективными* или *избирательными*. К ним относят фотопластинки, фотоэлементы и особенно человеческий глаз. Для каждого из них своя *спектральная кривая чувствительности*. Они характеризуются *световыми величинами*. Чувствительность глаза характеризуют *кривой видности*, отражающей относительную чувствительность глаза от длины волн света. Максимальная чувствительность соответствует длинам волн в области 550-560 нм (она принимается за единицу). Световые характеристики субъективны и используются только для видимого света.

Сила света измеряется в *канделах* (кд): 1 кд это сила света в заданном направлении монохроматического излучения с частотой $5.4 \cdot 10^{14}$ Гц, которая составляет $1/683$ Вт·ср⁻¹.

Световой поток, испускаемый точечным источником света силой 1 кд в телесной угол 1 ср равен в *1 люмену* (лм) (1 лм = 1 кд · 1 ср).

Освещенность и *светимость* измеряются в *люксах* (1лк = 1лм/1 м²).

Установлено, что на длине волны $\lambda = 555$ нм световому потоку 1 лм соответствует поток энергии $A = 0.00160$ Вт/лм. Коэффициент A называют *механическим эквивалентом света*. Связь между световым потоком Φ_e и потоком энергии Φ определяется соотношением

$$\Phi = \frac{\Phi_e A}{V(\lambda)}, \quad (4.9)$$

где Φ_e измеряется в ваттах, а Φ — в люменах; $V(\lambda)$ — кривая относительной чувствительности глаза.

Сравнительные характеристики лазерного и солнечного света. Применительно к лазерному пучку яркость B можно определить как $B = J/\Omega$, где Ω — телесный угол, в котором он распространяется. Лазерный пучок характеризуют *углом расходимости* θ , определяемый как показано на **Рис. 4.2**. Тогда $\Omega = \pi\theta^2$ (для лазерных пучков $\theta \ll 1$). Так как интенсивность $I = P/\pi\rho_0^2$ (P — мощность), то получаем

$$B = \frac{P}{\pi^2 \rho_0^2 \theta^2}. \quad (4.10)$$

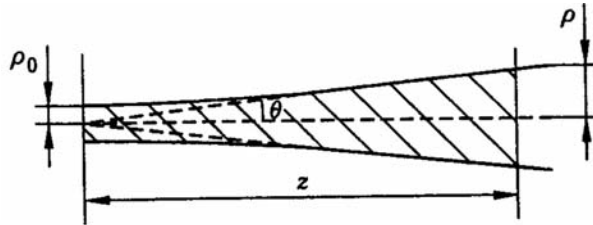


Рис. 4.2. К расчету яркости лазерного излучения для лазерного пучка. ρ_0 — радиус пучка в самом узком месте, которое называют перетяжкой; ρ — радиус пучка на расстоянии z от перетяжки; θ — угол расходимости лазерного пучка.

Для лазерного пучка $\theta = \lambda/2\pi\rho_0$, следовательно $B = 4P/\lambda^2$. Например, для гелий-неонового лазера с параметрами $\lambda = 0.63$ мкм, $P = 1$ мВт получаем $B = 10^6$ Вт·см⁻²·стерад⁻¹.

Яркость Солнца вычислим по формуле $B = P/(4\pi R)^2$, P — полная мощность излучения Солнца, R — радиус Солнца. Полагая $P = 3.83 \cdot 10^{26}$ Вт, $R = 6.96 \cdot 10^{10}$ см, получим $B = 5 \cdot 10^2$ Вт·см⁻²·стерад⁻¹. Таким образом, яркость лазера на несколько порядков превышает яркость Солнца.

4.2. Световые пучки и импульсы: энергия, мощность, интенсивность. Реальный световой пучок имеет конечный поперечный размер и его можно описать квазиплоской волной, амплитуда которой зависит от координат в плоскости перпендикулярной направлению распространения:

$$E = A(x, y) \cos(\omega t - kz) \quad (4.11)$$

Аналогичным образом поле светового импульса описывается квазиплоской квазигармонической волной, амплитуда которой зависит не только от координат, но и от времени:

$$E = A(x, y, t) \cos(\omega t - kz) \quad (4.12)$$

Энергетические характеристики световых пучков и импульсов характеризуют понятиями *интенсивности света*, *полной мощности пучка* и *полной энергии импульса*. Выражения для интенсивности света в моделях квазиплоской волны и квазиплоской квазигармонической волны получим, подставив выражения (4.11), (4.12) в формулу для интенсивности (2.19). При этом оказывается, что интенсивность света в пучке зависит от координат в плоскости поперечного сечения:

$$I(x, y) = \frac{c}{8\pi} A^2(x, y) \quad (4.13)$$

Интенсивность импульсного излучения зависит, кроме того, и от времени

$$I(x, y, t) = \frac{c}{8\pi} A^2(x, y, t) \quad (4.14)$$

Полная мощность светового пучка определяется как интеграл от интенсивности по поперечному сечению пучка:

$$P = \int \int_{-\infty}^{\infty} I(x, y) dx dy \quad (4.15)$$

Аналогичным образом полная энергия импульса есть

$$W = \int \int \int_{-\infty}^{\infty} I(x, y, t) dx dy dt \quad (4.16)$$

Например, для осесимметричного пучка с гауссовским распределением интенсивности

$$I(r) = I_0 \exp(-r^2/r_0^2) \quad (4.17)$$

полная мощность есть

$$P = \int_0^{\infty} I_0 2\pi r dr = \pi r_0^2 I_0 \quad (4.18)$$

где I_0 — интенсивность в центре пучка, r_0 — его радиус.

Для измерения энергетических характеристик света используют его *тепловое действие*, а также явление *фотоэффекта*. Располагая измерительной аппаратурой с достаточным временным и пространственным разрешением, можно измерить поперечные распределения интенсивности непрерывного $I(x, y)$ и импульсного излучений.

Для оценок удобно использовать такие параметры как *эффективная интенсивность света* I_{ef} , *эффективная мощность импульса* P_{ef} , *эффективная напряженность* электрического поля световой волны E_{ef} . Эффективная интенсивность непрерывного излучения определяется как отношение мощности пучка P к площади его поперечного сечения s

$$I_{ef} = P/s \quad (4.19)$$

Для светового импульса длительностью T и энергией W

$$P_{ef} = W/T, \quad I_{ef} = W/(Ts), \quad (4.20)$$

Эффективная напряженность светового поля определяется как

$$E_{ef} = \sqrt{8\pi I_{ef}/c} \quad (4.21)$$

Рассмотрим несколько примеров.

Солнечный свет. Интенсивность солнечного света вблизи поверхности Земли составляет около $0,1 \text{ Вт/см}^2$. Радиус земной орбиты $R = 1,5 \cdot 10^8 \text{ км}$. Полная мощность равна $4\pi R^2 = 3 \cdot 10^{26} \text{ Вт}$.

Гелий-неоновый лазер. Это газовый лазер непрерывного действия, излучающий свет с длиной волны $\lambda = 0,63 \text{ мкм}$. Типичная мощность излучения составляет $P = 10^{-2} \text{ Вт}$. При радиусе пучка $r = 0,2 \text{ см}$ его эффективная интенсивность $I_{ef} = P/(\pi r^2) = 0,1 \text{ Вт/см}^2$. Это сравнительно небольшая интенсивность, она близка к интенсивности солнечного света на поверхности Земли. Эффективная напряженность светового поля согласно формуле (4.21) $E_{ef} = 3 \cdot 10^{-2} \text{ В/см}$.

Лазер на углекислом газе. Этот лазер генерирует инфракрасное излучение на длине волны $\lambda = 10,6 \text{ мкм}$. В непрерывном режиме мощность генерации составляет обычно $P = 10 - 100 \text{ Вт}$. В импульсном режиме лазер генерирует импульсы с энергией $W = 0,1 \text{ Дж}$ при длительности импульса $T = 10^{-7} \text{ с}$ и обладает эффективной мощностью $P_{ef} = 10^6 \text{ Вт}$.

Лазер на гранате с неодимом. Твердотельный лазер на кристалле алюмоиттриевого граната, активированного ионами неодима, излучает в ближнем инфракрасном диапазоне на длине волны $\lambda = 1,06 \text{ мкм}$. В непрерывном режиме с возбуждаемого излучением газоразрядной лампы активного элемента длиной $6 - 8 \text{ см}$ и диаметром около 1 см удается получить мощность $P = 1 - 10 \text{ Вт}$. В импульсном режиме лазер генерирует импульс длительностью $T = 10^{-3} \text{ с}$ и энергией $W = 1 \text{ Дж}$ (режим свободной генерации), при этом эффективная мощность излучения $P_{ef} = 10^3 \text{ Вт}$. В режиме модулированной добротности лазер генерирует импульс длительностью $T = 10^{-8} \text{ с}$ и энергией $W = 0,1 \text{ Дж}$. Мощность такого ("гигантского") импульса составляет 10^7 Вт .

4.3. Геометрическая оптика. *Геометрическая оптика* описывает распространение света, используя понятие *лучей* или *пучков*, отвлекаясь от его волновой природы. Она представляет собой *предельный случай волновой оптики* при $\lambda \rightarrow 0$. В действительности достаточно, чтобы длина волн была много меньше характерных для данной задачи линейных размеров. Она также изучает поведение световых лучей (пучков) в оптических инструментах, которые состоят из различных отражающих и преломляющих поверхностей. Лежащие в основе геометрической оптики *законы отражения* и *преломления* могут быть выведены на основе уравнений Максвелла в предельном случае $\lambda \rightarrow 0$. Геометрическая оптика хотя и является *приближением*, она представляет огромный интерес с технической и исторической точек зрения.

В геометрической оптике основным понятием является *луч* — *геометрическая линия, вдоль которой распространяется световая энергия*. Основные свойства лучей: *они независимы друг от друга, т.е. не взаимодействуют друг с другом, и в однородной среде распространяются прямолинейно* (если ничто им не препятствует). *Поверхность нормальная к лучам называется волновой поверхностью*.

Если перед точечным источником света поместить экран с отверстием, то отверстие выделит в пространстве за экраном некоторый объем, внутри которого распространяется световая энергия, называемый — *световым пучком*. Он имеет вид конуса с углом раствора α , который определяется расстоянием от источника до экрана и размером отверстия (апертурная диафрагма). При $\alpha = 0$ говорят о *параллельном световом пучке*. Представления о *параллельном пучке* используются широко, но следует помнить, что *в действительности такие пучки не существуют* — это *полезная идеализация*. Наблюдать можно лишь световые пучки, но не лучи, потому что *световые лучи* — это тоже *идеализация*, удобная для различных геометрических построений и расчетов. С понятием луча (пучка) связывают как энергетические характеристики, например, световой поток, так и частоту, длину волны, поляризацию и другие.

С точки зрения *лучевой оптики* каждая светящаяся точка рассматривается как *вершина расходящегося пучка лучей*, который называют *гомоцентрическим*, т.е. имеющим общий центр. Если после отражения и преломления этот пучок превращается в пучок, *сходящийся в одну точку*, то последний также является *гомоцентрическим*, а его центр есть *изображение* светящейся точки. При *сохранении гомоцентричности каждая точка источника дает только одну точку изображения*. Такие изображения называются *точечными* или *стигматическими*. Волновая поверхность гомоцентрического пучка в однородной и изотропной среде является *сферической поверхностью*.

Если в результате отражения и преломления пучка гомоцентричность нарушается, то волновая поверхность перестает быть сферой. При этом стигматичность теряется и изображение точка уже не является точкой. Это явление называют *астигматизмом*.

Основные законы геометрической оптики: закон зеркального отражения и закон преломления Снеллиуса.

Зеркальное отражение. Зеркальное отражение возникает всякий раз, когда плоская волна падает на гладкую плоскую поверхность, например, зеркало: свет не проходит через зеркало, а уходит от него по прямой линии. Отражение удовлетворяет следующими условиями:

- а) *отраженный луч лежит в плоскости падения, образованной падающим лучом и нормалью к поверхности;*
- б) *угол отражения θ_0 равен углу падения θ_1 ($\theta_0 = \theta_1$ на Рис.4.3, оба угла измеряются от нормали) .*

Преломление, закон Снеллиуса. Пучок света определенного цвета *преломляется*, т.е. *отклоняется от прямолинейного распространения* всякий раз, когда на его пути встречается *поверхность раздела двух сред с разными показателями преломления n* . Величина отклонения зависит от отношения коэффициентов преломления n_2/n_1 (n_1 — показатель преломления среды 1, из которой пучок выходит, n_2 — показатель преломления среды 2, в которую он входит) и от *угла падения* — угла между *преломленным пучком и нормалью к поверхности θ_2* , на которую падает

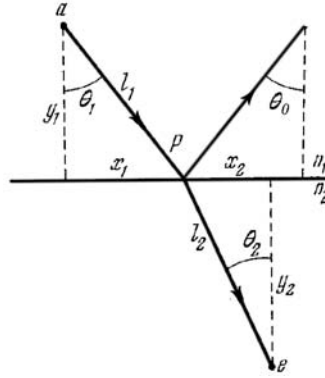


Рис. 4.3. Преломление и отражение света на плоской поверхности

пучок (**Рис.4.3**). Ниже мы покажем, что

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \quad \text{или} \quad \frac{\sin \theta_1}{\sin \theta_2} = \frac{n_2}{n_1} = n, \quad (4.22)$$

где n — относительный показатель преломления. Соотношение (4.22) называется *законом преломления Снеллиуса*.

При явлениях преломления и отражения имеет место закон *взаимности* или *обратимости* световых лучей, который означает, что при преломлении и отражении на границе двух сред лучи остаются *взаимными*, т.е. при изменении направления лучей на обратное их взаимное расположение не меняется.

Эйконал. В плоской монохроматической волне, распространяющейся в изотропной однородной среде с показателем преломления n , зависимость напряженности поля от координат и времени имеет вид:

$$E(\vec{r}, t) = E_0 \exp i(k_0 n \vec{s} \vec{r} - \omega t), \quad (4.23)$$

где $\vec{s} = \vec{k}/k$ — единичный вектор, называемый *лучевым вектором*, указывает направление распространения плоской волны; $\vec{k} = \vec{s}k = k_0 n \vec{s}$, $k_0 = \omega/c = 2\pi/\lambda_0$ — волновое число. В *неоднородной среде* показатель преломления есть функция координат $n = n(\vec{r})$, и выражение (4.23) уже не является решением уравнений Максвелла. В том случае решение можно искать в виде:

$$E(\vec{r}, t) = E_0(\vec{r}) \exp [ik_0 S(\vec{r}) - \omega t]. \quad (4.24)$$

Величину $S(\vec{r})$ называют *эйконалом*. Она представляет собой вещественную скалярную функцию координат. Предполагая, что $S(\vec{r})$ и амплитуда $E_0(\vec{r})$ *медленно изменяются от координаты* $\vec{r}$ рассматриваемой точки, можно показать, что (4.24) является *приближенным* решением уравнений Максвелла в случае *больших* k_0 (или малых длин волн λ_0) при условии, что $S(\vec{r})$ удовлетворяет так называемому *уравнению эйконала*. Найдем его.

Напряженность поля $E(\vec{r}, t)$ при $\lambda_0 \rightarrow 0$ испытывает быстрые осцилляции в пространстве, но амплитуда $E_0(\vec{r})$ и эйконал $S(\vec{r})$ изменяются медленно, оставаясь конечными по величине. В малых участках пространства $S(\vec{r})$ вблизи $\vec{r}_0$ можно разложить в ряд, ограничиваясь членами первого порядка:

$$S(\vec{r}) = S(\vec{r}_0) + (\vec{r} - \vec{r}_0)(\partial S / \partial \vec{r})_{\vec{r}_0} = S(\vec{r}_0) + (\vec{r} - \vec{r}_0) \nabla S, \quad (4.25)$$

где $\nabla = \vec{i} \partial / \partial x + \vec{j} \partial / \partial y + \vec{k} \partial / \partial z$ — дифференциальный оператор "набла".

Перепишем (4.24) с учетом (4.25):

$$E(\vec{r}, t) = E_0 \exp [ik_0 \nabla S \cdot \vec{r} - \omega t], \quad E_0 = E_0(\vec{r}_0) \exp ik_0 [S(\vec{r}_0) - \vec{r}_0 \cdot \nabla S]. \quad (4.26)$$

Таким образом, в каждом малом участке пространства, в пределах которого выполняется (4.26), произвольную волну можно рассматривать как плоскую, при условии, что $S(\vec{r})$ удовлетворяет уравнению

$$\nabla S(\vec{r}) = n\vec{s}, \quad (\nabla S)^2 = n^2, \quad (4.27)$$

которое называют *уравнением эйконала*. Волновые поверхности монохроматической волны (поверхности равных фаз) (4.24) определяются уравнением:

$$k_0 S(\vec{r}) - \omega t = 0. \quad (4.28)$$

Уравнения (4.27) показывают, что *лучи, т.е. линии, касательные к которым в каждой точке распространения волны в направлении единичного вектора $\vec{s}$, ортогональны к волновым поверхностям*. Поверхности равных фаз перемещаются в направлении луча $\vec{s}$ со скоростью $v = c/n$. В общем случае при показателе преломления $n(\vec{r})$, меняющемся от точки к точке, лучи будут искривлены. Поэтому *когда говорят о прямолинейном распространении света, всегда речь идет об однородной среде*.

Получим **уравнение для лучей в случае оптически неоднородных сред**. Радиус-вектор $\vec{r}$ точки P , лежащей на луче, будем рассматривать как функцию длины дуги l , т.е. $d\vec{r}/dl = \vec{S}$. Используя (4.27), можно записать, что $n d\vec{r}/dl = \nabla S$. Дифференцируя это уравнение по l , получаем: $d/dl(\nabla S) = \nabla dS/dl = \nabla n$ (здесь использовали, что $dS/dl = \nabla S \cdot \vec{S} = n$).

$$\frac{d}{dl} \left(n \frac{d\vec{r}}{dl} \right) = \nabla n \quad (4.29)$$

Отсюда нетрудно увидеть, что в однородной среде ($n = \text{const}$) $\nabla n = 0$ и (4.29) принимает вид: $d^2\vec{r}/dl^2 = 0$. Решение этого уравнения хорошо известно: $\vec{r} = \vec{a}l + \vec{b}$, где $\vec{a}$ и $\vec{b}$ постоянные векторы. Это прямая линия, направленная по $\vec{a}$ и проходящая через точку $\vec{r} = \vec{b}$. Таким образом, в однородной среде световые лучи прямолинейны. В неоднородной среде этот закон нарушается.

Принцип Ферма. По мере развития науки нам хочется получить нечто большее, чем просто формулу. Сначала мы наблюдаем явления, затем с помощью измерений получаем числа и, наконец, находим закон, связывающий эти числа. Но истинное величие и ценность науки состоит в том, что мы можем найти такой способ рассуждения, при котором закон становится очевидным. Роль таких достаточно общих рассуждений играют принципы, которые и составляют основу любой науки.

Впервые общий принцип, объясняющий поведение света, был предложен П.Ферма, известный теперь как **принцип Ферма** или **принцип наименьшего времени**: *луч света распространяется в пространстве между двумя точками по такому пути, вдоль которого время его прохождения наименьшее, чем вдоль любого другого, соединяющего эти точки*. Иногда его формулируют и так: *луч света распространяется между двумя точками по такой траектории, оптическая длина (произведение геометрической длины на показатель преломления), которой минимальна*.

Принцип Ферма можно обосновать, используя уравнение эйконала (4.27). Рассмотрим некоторую область с показателем преломления $n(\vec{r})$, через каждую точку которого проходит только один луч (лучи нигде не пересекаются). Пусть точки A и

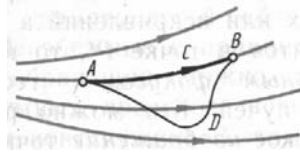


Рис. 4.4. К обоснованию принципа Ферма.

B (**Пус.4.4**) лежат на одном луче. Используя (4.27), $n\vec{s} = \nabla S(\vec{r})$, вычислим интеграл вдоль произвольной кривой, соединяющей точки A и B :

$$\int_{AB} n\vec{s}d\vec{r} = \int_{AB} \nabla S d\vec{r} = S(B) - S(A). \quad (4.30)$$

Этот интеграл равен разности значений эйконала в точках A и B и, следовательно, не зависит от пути интегрирования. Так как $\vec{s}d\vec{r} = dl \cos(\vec{s}, \widehat{d\vec{r}}) \leq dl$, то $S(B) - S(A) = \int_{AB} n\vec{s}d\vec{r} \leq \int_{AB} n dl$, причем знак равенства имеет место только в том случае, когда направления векторов $d\vec{r}$ и $\vec{s}$ совпадают в каждой точке рассматриваемой кривой, т.е. когда она представляет реальный луч, т.е. луч ACB на Рис.4.4. Для любой другой кривой, например, ADB (**Пус.4.4**) *оптическая длина пути* оказывается больше, чем для реального луча. Поскольку оптическая длина пути равна произведению скорости света на время, требуемое для прохождения вдоль этой кривой, то ясно, что свет между точками A и B распространяется по тому, который требует наименьшего времени.

В качестве примера получим, используя принцип Ферма, закон преломления света на границе раздела сред с разными показателями преломления (**Пус.4.3**). Пусть точка a находится на расстоянии y_1 от границы раздела (оси X), точка b — на расстоянии y_2 . В обозначениях, указанных на **Пус.4.3** оптическая длина пути $ab = l = l_1 + l_2$ равна

$$l = \frac{n_1 y_1}{\cos \theta_1} + \frac{n_2 y_2}{\cos \theta_2}. \quad (4.31)$$

Расстояние X между точками a и b фиксированно, т.е.

$$X = y_1 \tan \theta_1 + y_2 \tan \theta_2 = \text{const}, \quad (4.32)$$

Это соотношение задает связь между углами θ_1 и θ_2 . Минимальная оптическая длина l_{min} определяется из условия $dl/d\theta_1 = 0$, в котором дифференцирование производится с учетом связи (4.32):

$$\begin{aligned} \frac{dl}{d\theta_1} &= n_1 y_1 \frac{\sin \theta_1}{\cos^2 \theta_1} + n_2 y_2 \frac{\sin \theta_2}{\cos^2 \theta_2} \frac{d\theta_2}{d\theta_1} = 0 \\ \frac{dX}{d\theta_1} &= \frac{y_1}{\cos^2 \theta_1} + \frac{y_2}{\cos^2 \theta_2} \frac{d\theta_2}{d\theta_1} = 0. \end{aligned} \quad (4.33)$$

Исключая из этих равенств производную $d\theta_1/d\theta_2$, приходим к закону Снеллиуса $n_1 \sin \theta_1 = n_2 \sin \theta_2$. Аналогичным образом можно получить закон отражения.

Интересно отметить, что сам Ферма сформулировал его, исходя из теологических соображений о целенаправленности действий природы, о том, что цель должна достигаться с наименьшими затратами средств.

Принцип Ферма относится к числу так называемых *экстремальных* или *вариационных принципов*, которые играют важную роль при описании явлений природы. С

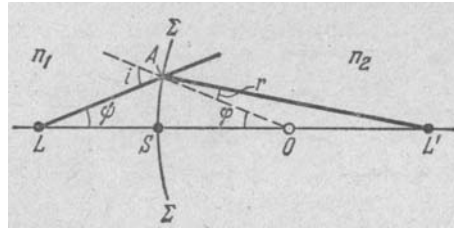


Рис. 4.5. Преломление на сферической поверхности. LA — падающий луч, $L'A$ — преломленный луч, i — угол падения, r — угол преломления, $AO = R$ — радиус сферической поверхности, $|LS| = a_1$, $|SL'| = a_2$

помощью принципа Ферма можно объяснить все явления, изучаемые в геометрической оптике, хотя не всегда это просто сделать.

Преломление и отражение на сферической поверхности. Предположим, что две среды с показателями преломления n_1 и n_2 разделяются сферической поверхностью Σ (Рис.4.5). На линии LL' проходящей через центр нашей сферы O , поместим гомоцентрический точечный источник света L . Рассмотрим узкий гомоцентрический конус лучей, падающий на поверхность раздела двух сред. Мы предполагаем пучок настолько узким, т.е. угол ψ таким малым, что практически можно считать, что

$$LS \approx LA, \quad L'S \approx L'A. \quad (4.34)$$

Такой пучок называют *параксиальным*, т.е. приосевым, а соотношения (4.34) — *условием параксиальности пучка*.

Построим для луча LA , падающего на поверхность Σ под углом i , сопряженный ему преломленный луч AL' (угол преломления r) и найдем положение точки, в которой преломленный луч пересечет ось системы. Так как $LO/LA = \sin i / \sin \varphi$, $AL'/OL' = \sin \varphi / \sin r$, то

$$\frac{LO}{LA} \frac{AL'}{OL'} = \frac{\sin i}{\sin r} = \frac{n_2}{n_1}. \quad (4.35)$$

Все отрезки вдоль оси будем отсчитывать от точки S , считая положительными отрезки, откладываемые от S вправо (в направлении распространяющегося света), и отрицательными — отрезки, откладываемые влево. Тогда $AL \approx SL = -a_1$, $AL' \approx S'L' = a_2$, $AO = SO = R$ — радиус сферы. Учитывая, что $LO = -a_1 + R$, $OL' = a_2 - R$, из формулы (4.35) получим

$$n_1 \left(\frac{1}{a_1} - \frac{1}{R} \right) = n_2 \left(\frac{1}{a_2} - \frac{1}{R} \right) = Q, \quad (4.36)$$

т.е. величина произведения $n \left(\frac{1}{a} - \frac{1}{R} \right)$ при преломлении сохраняется. Его называют *нулевым инвариантом Аббе*. Формулу (4.36) удобно представить в виде:

$$\frac{n_1}{a_1} - \frac{n_2}{a_2} = \frac{n_1 - n_2}{R}. \quad (4.37)$$

Соотношение (4.37) позволяет найти положение точки L' по заданной L . Оно справедливо для любого луча параксиального пучка. Из этой формулы также видно, что a_2 зависит только от a_1 при заданных параметрах задачи n_1, n_2, R . Величину $\Phi = (n_1 - n_2)/R$ называют *оптической силой* сферической преломляющей поверхности.

Пользуясь установленным выше правилом знаков, можно разобрать все случаи выпуклой ($R > 0$) или вогнутой ($R < 0$) поверхности. Точно так же в зависимости от того, будут ли a_1 и a_2 иметь разные знаки или одинаковые, мы будем иметь случаи, когда изображение располагается с противоположной по сравнению с источником стороны преломляющей поверхности или лежит по одну сторону с ним. В первом случае ($a_2 > 0$) точка, именуемая *изображением*, есть действительно точка пересечения преломленных лучей. Такое изображение называется *действительным*. Во втором случае ($a_2 < 0$), очевидно, преломленные лучи, идущие во второй среде, остаются расходящимися и реально не пересекаются. В этом случае название *изображения* относится к той воображаемой точке, которая представляет собой место пересечения предполагаемого продолжения преломленных лучей. Такое изображение называется *мнимым*. Таким образом, гомоцентрический пучок после преломления направлен так, что его лучи или пересекаются в одной точке (действительное изображение), или могут быть представлены как пересекающиеся в одной точке (мнимое изображение). Именно в этом смысле он и остается *гомоцентрическим*. Поэтому при всех построениях мы одинаково можем пользоваться как действительным, так и мнимым изображением.

Также отметим, что из формулы (4.37) следует, что если бы источник был помещен в точку L' , то изображение расположилось бы в L . Это свойство называют *взаимностью*.

Из уравнения (4.37) также следует, что при $a_1 = -\infty$ (луч идет параллельно оптической оси)

$$a_2 = \frac{n_2 R}{n_2 - n_1} = -\frac{n_2}{\Phi} = f_2, \quad (4.38)$$

при $a_2 = \infty$

$$a_1 = -\frac{n_1 R}{n_2 - n_1} = \frac{n_1}{\Phi} = f_1, \quad (4.39)$$

т. е. f_1 и f_2 зависят только от *радиуса кривизны поверхности* R и *показателей преломления* n_1, n_2 обеих сред. Величины f_1 и f_2 характеризуют преломляющую поверхность, их называют *фокусными расстояниями*: f_1 — переднее фокусное расстояние (точка F_1 — передний фокус); f_2 — заднее фокусное расстояние (точка F_2 — задний фокус).

Таким образом, в *фокусе сферической поверхности* сходятся после преломления *параллельные лучи* (т. е. лучи, идущие из бесконечно удаленной точки).

Фокусы, также как и изображения, могут быть *действительными* и *мнимыми*, т. е. представлять точку пересечения преломленных лучей (бывших до преломления параллельными) или их предполагаемых продолжений. Так, если вогнутая сторона

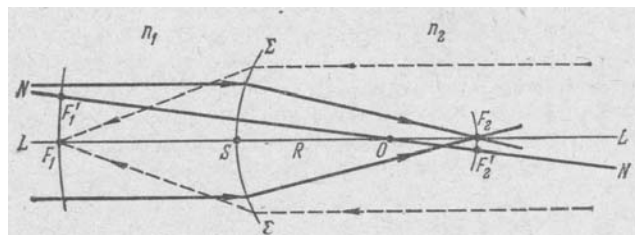


Рис. 4.6. Передний F_1 и задний F_2 фокусы сферической поверхности Σ . $SO = R$ — радиус сферы. F_1F_1', F_2F_2' — передняя и задняя фокальная поверхности.

поверхности раздела обращена к среде, имеющей меньший показатель преломления, то оба фокуса будут мнимыми, как это следует из формул (4.38) и (4.39).

Параллельные лучи, идущие справа налево вдоль NO (**Рис. 4.6**), сойдутся в фокусе F'_1 расположенном на линии NO и лежащем также на расстоянии $|f_1|$ от преломляющей поверхности. Геометрическое место точек F'_1 образует сферическую поверхность с радиусом $|R - f_1|$, concentрическую с преломляющей сферой (с центром в точке O). Эту поверхность называют *передней фокальной поверхностью*. Аналогично строится *задняя фокальная поверхность* радиуса $|f_2 - R|$. Малые участки этих поверхностей (для параксиальной области) могут быть приняты за плоскости — *фокальные плоскости*. Фокусные расстояния сферической поверхности различны по знаку и не равны между собой по абсолютной величине ибо $n_1 \neq n_2$. Нетрудно проверить, $f_2/f_1 = -n_2/n_1$.

Сферическое зеркало. Формулу (4.37) можно применить и к случаю отражения, если положить $n_2 = -n_1$. Тогда получаем известную формулу сферического зеркала:

$$\frac{1}{a_1} + \frac{1}{a_2} = \frac{2}{R}. \quad (4.40)$$

Фокусное расстояние такого зеркала определится по формуле (4.38), откуда получаем $f = R/2$. Формула зеркала принимает вид

$$\frac{1}{a_1} + \frac{1}{a_2} = \frac{1}{f}. \quad (4.41)$$

В случае зеркала изображение действительное, если оно лежит по одну сторону с источником, и мнимое, если расположено за зеркалом.

Случаи вогнутого и выпуклого зеркала отличаются лишь знаком R : фокус вогнутого зеркала — действительный, а фокус выпуклого зеркала — мнимый.

Чтобы получить законы плоского зеркала, достаточно положить $R = \infty$. В этом случае найдем $a_1 = -a_2$, т.е. изображение точки в плоском зеркале мнимое и симметрично расположенное (оно "находится" по другую зеркалу).

ЛЕКЦИЯ №5

Свойства центрированных оптических систем. Элементы матричной оптики. Распространение светового луча в оптически неоднородной среде.

5.1. Свойства центрированных оптических систем. Большинство оптических систем содержит по крайней мере две преломляющие поверхности (простейший пример — линза) или большее их число. Система сферических поверхностей называется *центрированной*, если центры всех поверхностей лежат на одной прямой, которая называется *главной оптической осью* системы. Теория таких систем особенно проста, когда распространяющиеся в них пучки лучей являются *параксиальными*.

Образование изображений в параксиальном приближении систематически исследовано Гауссом, поэтому теорию центрированных оптических систем в параксиальном приближении называют *гауссовой оптикой* или *теорией идеальной оптической системы*, т. е. системы, в которой сохраняется *гомоцентричность пучков* и *изображение геометрически подобно предмету*. Согласно этому определению *всякой точке пространства объектов соответствует точка пространства изображений*; эти точки называют *сопряженными*. Точно так же каждой прямой или плоскости пространства объектов должна соответствовать сопряженная прямая или плоскость пространства изображений. Таким образом, теория *идеальной оптической системы* есть чисто *геометрическая теория*, устанавливающая соотношение между точками, линиями, плоскостями. Идеальная оптическая система может быть осуществлена с достаточным приближением в виде *центрированной оптической системы*, если параксиальными пучками, т. е. областью вблизи оси симметрии. В теории Гаусса *требование "тонкости" системы отпадает, но лучи предполагаются параксиальными*. Линия, соединяющая центры сферических поверхностей, представляет собой *ось симметрии центрированной системы* и называется *главной оптической осью* системы. Теория Гаусса устанавливает ряд так называемых *кардинальных точек и плоскостей*, задание которых полностью описывает все свойства оптической системы и позволяет пользоваться ею, не рассматривая реального хода лучей в системе. К ним относятся: *фокусы и фокальные плоскости; главные точки и главные плоскости; узлы и узловые плоскости*.

Точка, в которой пересекаются параллельные лучи называются фокусом. Плоскости, перпендикулярные оптической оси в точках фокуса — *фокальные плоскости*. Их построение показано на **Рис. 5.1**. Пусть MM и NN — крайние сферические поверхности, ограничивающие оптическую систему, O_1O_2 — ее главная ось.

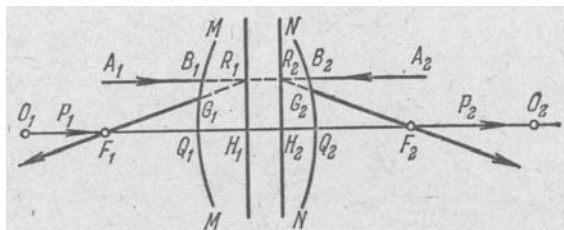


Рис. 5.1. Фокусы F_1 и F_2 и главные плоскости H_1R_1 и H_2R_2 центрированной оптической системы, образованной сферическими поверхностями MM и NN . O_1O_2 — оптическая ось, P_1, P_2 — сопряженные точки, $F_1H_1 = f_1, F_2H_2 = f_2$ — фокусные расстояния, $P_1H_1 = a_1, P_2H_2 = a_2$.

Проведем луч A_1B_1 , параллельный O_1O_2 ; точка B_1 — место входа этого луча в систему. Согласно свойству идеальной системы лучу A_1B_1 соответствует в пространстве изображений сопряженный луч G_2F_2 , выходящий из системы в точке G_2 . Как идет луч внутри системы, нас не интересует. Вторым луч P_1Q_1 выберем вдоль главной оси. Сопряженный ему луч Q_2P_2 будет также идти вдоль главной оси. Точка пересечения F_2 двух лучей G_2F_2 и Q_1P_2 есть изображение точки, в которой пересекаются параллельные лучи A_1B_1 и P_1Q_1 , сопряженные с G_2F_2 и Q_2P_2 . Точку F_2 называют вторым или *задним фокусом* системы. Плоскость, проходящая через фокус перпендикулярно к оси, носит название *фокальной*. Аналогичным образом определяется *передний фокус* F_1 системы, в котором пересекаются параллельные лучи A_2B_2 и P_2Q_2 .

Продолжим теперь F_1G_1 и F_2G_2 до пересечения с продолжениями A_1B_1 и A_2B_2 и отметим точки пересечения $R_{1,2}$. Эти точки являются *сопряженными*. Действительно, R_1 есть точка пересечения лучей $A_1B_1R_1$ и $F_1Q_1R_1$, которым сопряжены соответственно лучи $R_2G_2F_2$ и $R_2B_2A_2$, пересекающиеся в R_2 . Из построения видно, что R_1 и R_2 лежат на одинаковом расстоянии от главной оси ($H_1R_1 = H_2R_2$).

Можно показать что любая точка линии H_1R_1 сопряжена с точкой линии H_2R_2 , лежащей на такой же высоте от оптической оси, как и выбранная. То же относится и к плоскостям, проведенным через H_1R_1 и H_2R_2 перпендикулярно к главной оси, ибо вся система симметрична относительно оси. Плоскости H_1R_1 и H_2R_2 называются *главными плоскостями*. Точки $H_{1,2}$ пересечения главных плоскостей с осью называют *главными точками* системы. Расстояния от главных точек до фокусов называются *фокусными расстояниями* системы: $f_1 = H_1F_1$ и $f_2 = H_2F_2$.

Обозначим расстояния от главных плоскостей до сопряженных точек $P_{1,2}$ через $a_{1,2}$. Можно показать, что имеют место следующие соотношения (смотри, например, Г.С.Лансберг, "Оптика", М., Наука, 1976, задача 106 на стр. 883):

$$f_1/a_1 + f_2/a_2 = 1, \quad x_1x_2 = f_1f_2, \quad f_1/f_2 = -n_1/n_2, \quad V = -x_2/f_2 = -f_1/x_1, \quad (5.1)$$

где $x_{1,2} = a_{1,2} - f_{1,2}$ — расстояния сопряженных точек $P_{1,2}$ от соответствующих фокусов. Для $n_1 = n_2$ (источник и его изображение лежат в одной среде, например, в воздухе), имеем:

$$a_1^{-1} + a_2^{-1} = f^{-1}, \quad x_1x_2 = f^2, \quad f_2 = -f_1 = f_1. \quad (5.2)$$

Пользуясь правилом знаков, согласно которому все отрезки вдоль оси отсчитываемые от системы, считаются положительными, а влево — отрицательными, с помощью этих соотношений можно описать все свойства как собирающих, так и рассеивающих систем, ввести понятие мнимых точек и мнимых изображений. Отметим, что главные плоскости и главные точки могут лежать как внутри, так и вне системы совершенно несимметрично относительно поверхностей, ограничивающих систему.

Всякая оптическая система характеризуется *линейным* или *поперечным увеличением*, которое определяется как отношение $V = y_2/y_1$, где $y_{1,2}$ линейный (поперечный) размер ("высота") предмета и изображения, соответственно. В *параксиальном приближении* $V = y_2/y_1 = n_1a_2/n_2a_1$ ($a_{1,2}$ — расстояние от предмета (изображения) до оптической системы). Плоскости, в которых $V = 1$, называются *сопряженными*.

Кроме линейного увеличения, систему также характеризуют *угловым увеличением* W , определяемое соотношением:

$$W = \frac{\operatorname{tg} u_2}{\operatorname{tg} u_1}, \quad (5.3)$$

где $u_{1,2}$ — углы между лучами A_1M_1 , A_2M_2 и оптической осью (**Пис.5.2**). Так как

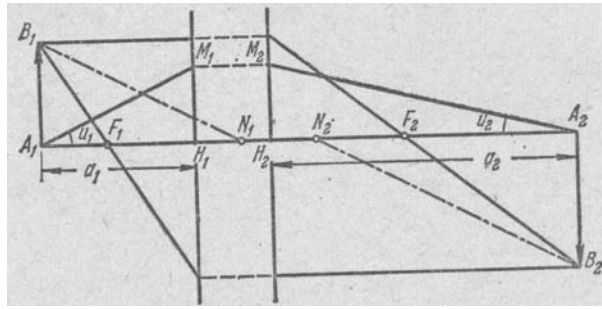


Рис. 5.2. К определению углового увеличения системы. A_1M_1 , A_2M_2 — сопряженные лучи, $F_{1,2}$ — фокусы, $H_{1,2}$ — главные точки, H_1M_1 , H_2M_2 — главные плоскости, $N_{1,2}$ — узлы.

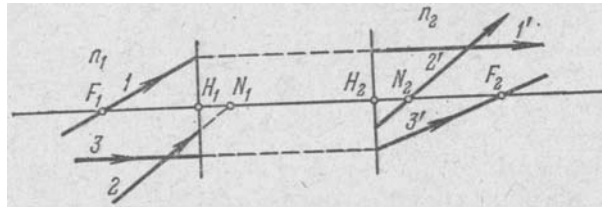


Рис. 5.3. Кардинальные точки и плоскости систем: $F_{1,2}$ — главные фокусы, $N_{1,2}$ — узлы, $H_{1,2}$ — главные точки.

$H_1M_1 = H_2M_2$ (**Рис.5.2**), то $W = a_1/a_2$. Очевидно, что $WV = n_1/n_2$ или $WV = 1$ при $n_1 = n_2$.

Точки, в которых угловое увеличение $W = 1$, называются *узлами* или *узловыми точками*. Для них характерно, что *сопряженные лучи, проходящие через узлы, параллельны друг другу*, так как $u_1 = u_2$. Можно показать, что в каждой системе такой парой точек будут точки N_1 и N_2 (**Рис.5.3**), отстоящие от первого и второго фокусов соответственно на расстояния, равные второму и первому фокусным расстояниям, т.е. $x_1 = F_1N_1 = f_2$ и $x_2 = F_2N_2 = f_1$. Точки $N_{1,2}$ являются *сопряженными*, так как их координаты удовлетворяют уравнению $x_1x_2 = f_1f_2$. Их расстояния относительно главных плоскостей равны соответственно $H_1N_1 = a_1 = f_2 - f_1$ и $H_2N_2 = a_2 = f_2 - f_1$, и следовательно, для этих точек $W = a_1/a_2 = 1$. *Плоскости, проходящие через узлы перпендикулярно к оптической оси, называются узловыми плоскостями*.

Шесть плоскостей (две фокальные, две главные и две узловые) и шесть точек главной оси, им соответствующие (фокусы, главные точки, узлы), называются *кардинальными* (**Рис.5.3**). Когда по обе стороны системы располагается одна и та же среда, фокусные расстояния равны ($f_1 = -f_2$). Узловые точки теперь сливаются с главными, так как $F_1N_1 = F_2N_2 = f_2$, и система характеризуется положением всего лишь четырех точек и плоскостей.

Зная свойства кардинальных плоскостей и точек, можно построить изображение в любой системе, пользуясь двумя лучами, исходящими из одной точки. При этом для линз отпадает требование тонкости. **Рис.5.4** показывает, как можно построить изображение в толстой линзе, если дано расположение ее главных плоскостей и фокусов. Здесь проведены лучи, построение которых особенно просто

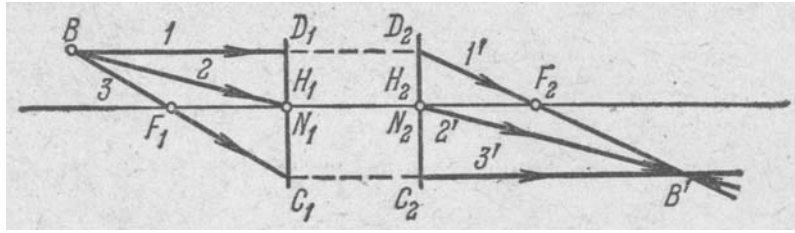


Рис. 5.4. Построение изображений в системе с использованием кардинальных точек. $F_{1,2}$ — фокусы, $H_{1,2}$ — главные точки, $H_{1,2}D_{1,2}$ — главные плоскости, $N_{1,2}$ — узлы. Здесь главные точки и узлы совпадают, так как показатели преломления слева и справа от оптической системы одинаковы.

определяет положение точки B' , сопряженной с точкой B . В силу гомоцентричности пучка любой другой луч из B пройдет через B' . Луч 1, проведенный параллельно главной оси, имеет в качестве сопряженного луч 1', пересекающий вторую главную плоскость на высоте $H_2D_2 = H_1D_1$ и проходящий через фокус F_2 . Луч 2 идущий через узел N_1 имеет сопряженный луч 2', проходящий через второй узел параллельно лучу 2. Луч 3, проходящий через фокус F_1 и пересекающий главную плоскость на высоте $H_1C_1 = H_1D_1$, пройдет на той же высоте ($H_1C_1 = H_2C_2$) через вторую главную плоскость и пойдет параллельно главной оси. Для построения изображения можно ограничиться двумя лучами из трех.

Тонкая линза может рассматриваться как частный случай толстой линзы, в которой точки $H_{1,2}$ совпадают и главные плоскости сливаются. Узловые точки, совмещенные с $H_{1,2}$, также совпадут, образуя оптический центр линзы. Построение изображения производится также при помощи каких-либо двух простейших лучей.

Часто приходится иметь дело с изображением *пространственных* предметов, отдельные точки которых лежат на разных расстояниях от главной плоскости. Поэтому вводят *продольное увеличение* U , показывающее отношение длины изображения Δx_2 к длине изображаемого малого отрезка Δx_1 , если последний расположен вдоль оси. Отметим, что речь идет об увеличении малых по длине отрезков, ибо продольное увеличение для разных точек оси различается очень значительно. Используя выражение формулы (Ф) для V можно получить следующее выражение:

$$U = \frac{n_2}{n_1} V^2, \quad (5.4)$$

где $V = y_2/y_1$ — линейное увеличение ($y_{1,2}$ — поперечный линейный размер предмета и его изображения, соответственно). Продольное увеличение характеризует резкость изображения пространственного объекта на экран ("глубину оптической системы"). Между U , W и V существует связь

$$UW = V \quad (5.5)$$

Теория идеальной оптической системы носит совершенно *общий характер*, т.е. применима к аксиально симметричным системам произвольной конструкции. Система оказывается полностью заданной, если известно взаимное расположение четырех кардинальных точек. Положение этих точек в каждой конкретной системе, разумеется, зависит от ее конструкции (от кривизны преломляющих и отражающих поверхностей, их расположения, показателя преломления и т. п.). Существует несколько методов анализа преобразования оптических лучей системе и нахождения

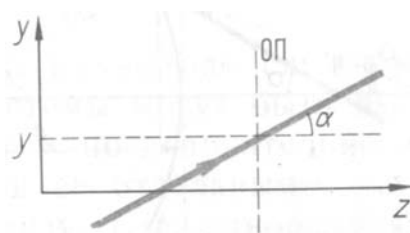


Рис. 5.5. Параметры y, α меридионального луча: z — оптическая ось, ОП — опорная плоскость

кардинальных точек. На практике широко используется изящный метод, основанный на использовании специальных матриц.

5.2. Элементы матричной оптики. Преобразование луча в оптической системе удобно описывать с помощью матриц. В оптической системе сферические и плоские поверхности служат границами раздела различных однородных сред. Траектория луча состоит из отрезков прямых линий — векторов. По определению, центрированные оптические системы обладают *цилиндрической симметрией*, по крайней мере на отдельных участках. Это позволяет существенно упростить трехмерную задачу преобразования вектора-луча, сведя ее к двумерной: рассматриваются лучи, лежащие в плоскости в сечения, проходящей через главную оптическую ось. Такая *плоскость* называется *меридиональной*, а *лучи*, лежащие в этой плоскости, — *меридиональными*.

Будем рассматривать только меридиональные лучи, лежащие в одной плоскости yz с главной оптической осью z . Выберем некоторую плоскость, перпендикулярную оптической оси z — *опорная плоскость (ОП)*. Любой меридиональный луч можно определить заданием двух параметров: *координаты y* точки его пересечения с опорной плоскостью и *угла α* , который он составляет с осью z (**Рис.5.5**). В дальнейшем для характеристики направления луча удобно вместо α использовать параметр $V = n\alpha$ (произведение показателя преломления среды на угол α).

Преобразование параметров y и V луча при переходе от одной опорной плоскости ОП₁ к другой ОП₂ в *параксиальном приближении* является *линейным*. Для любой пары опорных плоскостей оно имеет вид:

$$y_2 = Ay_1 + BV_1, \quad V_2 = Cy_1 + DV_1. \quad (5.6)$$

Это преобразование можно записать в матричной форме:

$$\begin{pmatrix} y_2 \\ V_2 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} y_1 \\ V_1 \end{pmatrix} \quad (5.7)$$

Опорные плоскости можно выбирать в разных местах оптической системы. Для данной пары плоскостей ОП₁ и ОП₂ преобразование параметров любого параксиального луча описывается одной и той же матрицей, сопоставляемой промежутку между ОП₁ и ОП₂. Ее элементы A, B, C, D зависят от свойств этого промежутка, т.е. от того какие преломляющие поверхности и какие среды находятся между этими опорными поверхностями. Матрица, описывающая преобразование лучей всей оптической системой, получается перемножением матриц, сопоставляемых отдельным промежуткам.

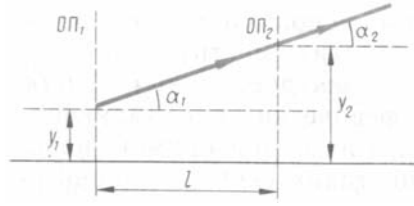


Рис. 5.6. К выводу матрицы преобразования для оптического промежутка

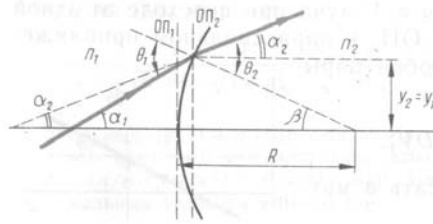


Рис. 5.7. К выводу матрицы преломления на сферической поверхности

Для описания поведения параксиального луча в центрированной оптической системе достаточно знать матрицы преобразования для трех основных элементов: оптического промежутка (т. е. участка однородной среды), преломляющей и отражающей поверхностей.

Оптический промежуток между ОП₁ и ОП₂ (Рис. 5.6) характеризуется толщиной l и показателем преломления n . Преобразование параметра y можно найти из Рис.5.6: $y_2 = y_1 + l \tan \alpha_1$. В параксиальном приближении углы наклона лучей считаются малыми. Поэтому $y_2 \approx y_1 + l\alpha_1$. Перейдем от α_1 к $V_1 = n\alpha_1$: уравнение для y_2 принимает вид $y_2 = y_1 + LV_1$, $L = l/n$ — *приведенная толщина* оптического промежутка. Так как наклон луча при переходе от ОП₁ и ОП₂ не изменяется, то $V_1 = V_2$, и преобразование параметров луча оптическим промежутком можно описать с помощью следующей матрицы F :

$$\begin{pmatrix} y_2 \\ V_2 \end{pmatrix} = \begin{pmatrix} 1 & L \\ 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ V_1 \end{pmatrix}, \quad F = \begin{pmatrix} 1 & L \\ 0 & 1 \end{pmatrix} \quad (5.8)$$

Сферическая преломляющая поверхность, разделяющая две среды с показателями преломления n_1 и n_2 (Рис.5.7), характеризуется радиусом кривизны R . Чтобы одни и те же формулы были справедливы для выпуклой и для вогнутой поверхностей, значение R считают положительным, если центр кривизны лежит на оси Z справа от границы, и отрицательным в противном случае. Для нахождения матрицы преломления выберем опорные плоскости ОП₁ и ОП₂ по обе стороны в непосредственной близости от преломляющей поверхности. Расстояние между ними $R(1 - \cos \beta)$ в параксиальном приближении пренебрежимо мало. Поэтому можно считать, что луч пересекает их на одном и том же расстоянии от оси, т. е. $y_1 \approx y_2$. Чтобы найти, как преобразуется параметр $V = n\alpha$, воспользуемся законом преломления, который для параксиальных лучей сводится к $n_1\theta_1 = n_2\theta_2$. Так как $\theta_1 = \alpha_1 + \beta$ и $\theta_2 = \alpha_2 + \beta$ (Рис.5.7), то умножая первое из этих равенств на n_1 , а второе — на n_2 и используя закон преломления, получаем, что $n_2\alpha_2 = n_1\alpha_1 + \beta(n_1 - n_2)$. Перепишем его, вводя $V_2 = n_2\alpha_2$, $V_1 = n_1\alpha_1$ и $\beta = y_1/R$: $V_2 = V_1 - Py_1$, где $P = (n_1 - n_2)/R$ — *оптическая*

сила преломляющей поверхности. Учитывая, что $y_1 = y_2$, получаем закон преобразования параметров луча при преломлении, описываемый следующей матрицей преломления T :

$$\begin{pmatrix} y_2 \\ V_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -P & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ V_1 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 \\ -P & 1 \end{pmatrix} \quad (5.9)$$

В случае плоской поверхности $R = \infty$, $P = 0$ и преломление описывается единичной матрицей, так как $V_1 = V_2$.

Чтобы включить в рассмотрение отражающие поверхности, вводят следующее правило: когда луч распространяется в отрицательном направлении оси z , показатель преломления среды, через которую он проходит, считается отрицательным ($-n$). Тогда закон отражения $\theta_2 = -\theta_1$ формально можно рассматривать как частный случай закона преломления при $n_1 = -n_1$. Матрица преобразования параметров луча при отражении от сферической поверхности имеет точно такой же вид как матрица преломления (5.9), если в выражении для оптической P заменить n_2 на $-n_1$: $P = -2n_1/R$. Для выпуклого зеркала $R > 0$ и оптическая сила отрицательна ($P < 0$), для вогнутого — положительна ($P > 0$).

Толщина оптического промежутка $l = z_2 - z_1$, между опорными плоскостями, проходящими через z_1 и z_2 положительна при $z_2 > z_1$ и отрицательна при $z_2 < z_1$, т. е. $l < 0$ для лучей, распространяющихся влево. Так как для таких лучей и показатель преломления нужно считать отрицательным, приведенная толщина $L = l/n$ положительной. Поэтому матрица F (5.8) для оптического промежутка не зависит от направления лучей. Параметры луча на выходе сложной оптической системы могут быть выражены через параметры входного луча путем последовательного применения преобразований этих параметров ее отдельными элементами.

Рассмотрим в качестве примера линзу, т. е. однородную с показателем преломления n (например, стекло), ограниченную сферическими поверхностями с радиусами кривизны $R_{1,2}$ (Рис. 5.8). Толщина линзы, т.е. расстояние между поверхностями вдоль оптической оси равно l . Обозначим: $K_1(y_1, V_1)$ — параметры некоторого луча на ОП₁ (перед входом в линзу), K_2 — на ОП₂ (после преломления на первой поверхности), K_3 — на ОП₃ (после прохождения оптического промежутка приведенной толщины $L = l/n$), $K_4(y_4, V_4)$ на задней поверхности. Пусть F — матрица оптического промежутка (5.8), $T_{1,2}$ — матрицы преломления (5.9) соответственно для первой и второй поверхностях, соответственно. В них $P_{1,2} = (n - 1)/R_{1,2}$ (n — относительный показатель преломления). Матрица преобразования M луча в этой системе равна произведению матриц для ее отдельных элементов взятых в обратном порядке: $M = T_2 F T_1$ или $K_4 = M K_1 = T_2 F T_1 K_1$. Выполняя перемножение матриц,



Рис. 5.8. Толстая линза

получаем

$$M = \begin{pmatrix} 1 & 0 \\ -P_2 & 1 \end{pmatrix} \begin{pmatrix} 1 & L \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -P_1 & 1 \end{pmatrix} = \begin{pmatrix} 1 - P_1 L & L \\ -(P_1 + P_2 - P_1 P_2 L) & 1 - P_2 L \end{pmatrix} \quad (5.10)$$

В случае *тонкой линзы*, когда толщину оптического промежутка l между преломляющими поверхностями можно считать пренебрежимо малой ($l \rightarrow 0$), матрица T вырождается в единичную и полная матрица M имеет такой же вид, как и матрица F преломления на одной поверхности, но с оптической силой $P = P_1 + P_2 = (n - 1)(1/R_1 - 1/R_2)$.

Таким способом можно найти полную матрицу M преобразования параметров параксиального луча для произвольной центрированной оптической системы, если известны кривизна и взаимное расположение ее преломляющих и отражающих поверхностей и значения показателей преломления. Введем обозначения для ее элементов A, B, C, D :

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \quad (5.11)$$

Она связывает параметры входящего луча (y_1, V_1) с параметрами выходящего (y_2, V_2) :

$$\begin{pmatrix} y_2 \\ V_2 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} y_1 \\ V_1 \end{pmatrix} \quad (5.12)$$

Матричные элементы удовлетворяют соотношению $\det M = AD - BC = 1$ и связаны с *фокусным расстоянием* системы F и координатами ее *главных плоскостей* $H_{1,2}$:

$$F = -\frac{1}{C}, \quad H_1 = \frac{D - 1}{C}, \quad H_2 = \frac{A - 1}{C} \quad (5.13)$$

Для примера рассмотрим систему, изображенную на **Рис.5.9**. Пусть луч, входящий в оптическую систему параллельно оптической оси на некоторой высоте y_1 . Для него $\alpha_1 = 0$ и, следовательно, $V_1 = 0$. На выходе из системы луч имеет параметры $y_2 = Ay_1$, $V_2 = CV_1$. Угол его наклона к оптической $\alpha_2 = V_2/n_2$, поэтому луч (или его продолжение) пересечет оптическую ось в точке F_2 , отстоящей от последней преломляющей поверхности ОП₂ на расстоянии $t_2 = -y_2/\alpha_2 = -n_2 y_2/V_2$. Подставляя сюда y_2 , V_2 , получаем $t_2 = -n_2 A/C$. Расстояние t_2 не зависит от y_1 , т. е. все лучи, входящие в систему параллельно оптической оси, проходят (в параксиальном приближении) через одну и ту же точку F_2 , которую называют *задней фокальной точкой* или *задним главным фокусом* оптической системы.

Пересечение продолжений входящего параллельно оптической оси луча и выходящего луча происходит в плоскости H_2 — *главной плоскости*. Определим фокусное

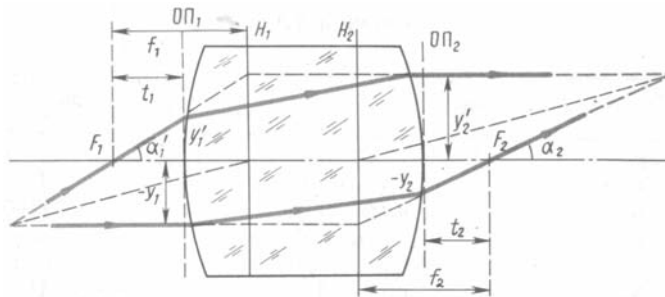


Рис. 5.9. К нахождению кардинальных точек оптической системы

расстояние F_2 как смещение вдоль оси от H_2 до F_2 . Тогда $f_2 = -y_2/\alpha_2 = -n_2 y_1/V_2$ (рис.5.8). Подставляя $V_2 = Cy_1$ получаем $f_2 = -n_2/C$, т. е. фокусное расстояние определяется элементом C матрицы M оптической системы (7.20). Чтобы найти положение передней фокальной точки F_1 , рассмотрим луч, идущий через нее под некоторым углом α'_1 . На выходе из системы он должен быть параллелен оптической оси, т. е. для него $V'_1 = 0$. Поэтому $V'_2 = Cy'_1 + DV'_1 = 0$. Подставив сюда $V'_1 = n_1 \alpha'_1$, найдем $y'_1 = -dn_1 \alpha'_1/C$. Из **Рис.5.9** видно, что $t_1 = -y'_1/\alpha'_1 = n_1 D/C$. Мы получили, что расстояние t_1 не зависит от α'_1 , т. е. все лучи из F_1 после прохождения через систему будут параллельны оптической оси. Рассматривая продолжения падающего и выходящего лучей, определяем положение передней главной плоскости H_1 и переднее фокусное расстояние, отсчитываемое от H_1 (при этом учтено, что для матрицы $\det M = 1$). Когда показатели преломления сред по обе стороны от системы одинаковы $f_2 = -f_1$, т.е. ее переднее и заднее фокусные расстояния равны модулю, но противоположны по знаку.

Фокусы $F_{1,2}$ и точки пересечения главных плоскостей $H_{1,2}$ с оптической осью называются *кардинальными точками*. Их положение полностью определяет преобразование любого параксиального луча оптической системой. Если оно известно, можно построить выходящий из системы луч, не рассматривая реального хода лучей в системе. Можно показать, что они полностью определяются элементами матрицы оптической системы.

Фокусное расстояние оптической системы полностью определяется элементом C матрицы преобразования лучей: $f_2 = -1/C$. Как и у тонкой линзы, этот элемент, взятый с противоположным знаком, называется *оптической силой* системы $P = -C$. Для толстой линзы, как видно из матрицы (5.10), $P = P_1 + P_2 - P_1 P_2 L$. Подставляя выражения для $P_{1,2}$ и L , можно получить

$$P = (n - 1)[1/R_1 - 1/R_2 + (n - 1)l/nR_1 R_2] \quad (5.14)$$

При $P > 0$ линза называется *собирающей*, при $P < 0$ — *рассеивающей*. Диагональные элементы A и D матрицы M (18) позволяют найти положение главных плоскостей толстой линзы: $h_1 = (D - 1)/C$ и $h_2 = (A - 1)/C$ (расстояния от входной и выходной плоскости)

Все полученные результаты справедливы при следующих условиях:

- свет поступает в систему в виде параксиальных пучков;
- пучки составляют небольшие углы с главной осью системы;
- показатель преломления постоянен для всех лучей. т.е. среда не имеет дисперсии или свет монохроматичен.

Все три условия не соблюдаются в практической оптике. Мы, как правило, имеем дело с немонохроматическим светом и должны учитывать зависимость показателя преломления от длины волны (дисперсия). Ограничение пучками, слабо наклоненными к оси, означало бы отказ от получения изображения точек, лежащих в стороне от главной оси системы, применение лишь параксиальных пучков вело к использованию незначительных световых потоков. К тому же реальная линза является толстой. Все это приводит к тому, что реальные оптические системы обнаруживают различного вида *абберации* или искажения, которые приходится устранять, используя различные ухищрения, правда этого невозможно делать до бесконечности.

Необходимо помнить, что *геометрическая оптика* — всего лишь некоторое *приближение*. В этом смысле принцип Ферма, в отличие, например, от закона сохранения энергии или импульса, тоже приближение.

5.3. Распространение светового луча в оптически неоднородной среде. Распространение светового луча в оптически неоднородной среде может быть описано, используя уравнения для эйконала. Для неоднородной среды $n = n(\vec{r})$ и уравнение (4.29) можно привести к виду: $(dn/dl)\vec{s} + n(d\vec{s}/dl) = \nabla n$. Так как $dn/dl = \nabla n \vec{s}$, то получаем уравнение на $\vec{s}$:

$$\frac{d\vec{s}}{dl} = \frac{1}{n}[\nabla n - \vec{s}(\nabla n \vec{s})]. \quad (5.15)$$

Производная единичного вектора $\vec{s}$ по длине луча l характеризует кривизну луча. Так как $d\vec{s}/dl = \vec{N}/R$, где $\vec{N}$ — единичный вектор главной нормали к лучу, R — радиус его кривизны. Умножая скалярно обе части (5.15) на $\vec{N}$ и учитывая, что $\vec{N}\vec{s} = 0$, получаем выражение для радиуса кривизны луча:

$$\frac{1}{R} = \vec{N} \frac{\nabla n}{n}. \quad (5.16)$$

Отсюда следует, что при $\vec{N}\nabla n > 0$ (угол между $\vec{N}$ и ∇n острый) — луч изгибается в область с большим показателем преломления.

Чтобы физически понять, что в неоднородной среде луч света искривляется и даже установить в какую сторону, рассмотрим простую ситуацию. Представим, что среда состоит оптически однородных горизонтальных слоев, показатель преломления которых меняется от слоя к слою (**Рис. 5.10**). Будем считать, что внутри слоя он остается одинаковым и $n_1 < n_2 < n_3$. Используя закон преломления $\sin \alpha_1 / \sin \alpha_2 = n_2 / n_1$ к нашей слоистой среде, нетрудно увидеть, что траектория луча должна иметь вид ломаной линии, как показано на (**Рис. 5.10** а). В действительности показатель преломления меняется не скачками, а непрерывно, поэтому луч представляет не ломаную, а кривую линию (**Рис. 5.10** б), причем искривление выпуклостью обращено в сторону уменьшения показателя преломления.

Определим радиус кривизны пучка. Пусть световой пучок диаметра D входит в среду, показатель преломления которой убывает в направлении вниз вверх (**Рис. 5.11**). В точке A , соответствующей верхнему краю, он равен $n - \Delta n$, в точке B , отвечающей нижнему краю — n . Для простоты будем считать $n \gg \Delta n$. Обозначим радиус кривизны пучка $R \gg D$. Угол ψ считаем малым, чтобы полагать, что

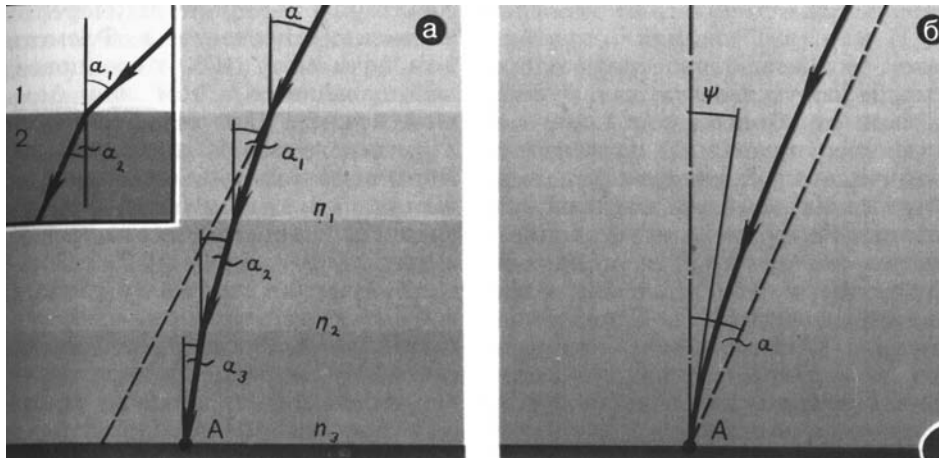


Рис. 5.10. Искривление светового луча при распространении в оптически неоднородной среде

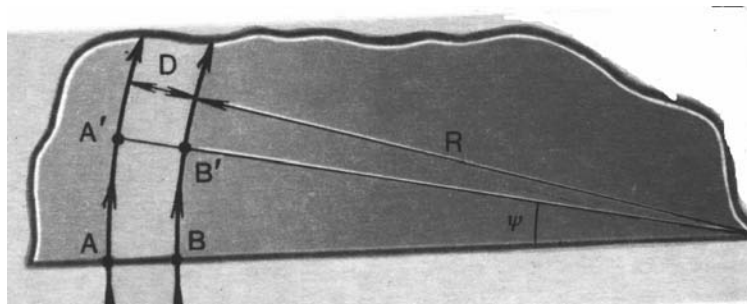


Рис. 5.11. К расчету радиуса кривизны светового луча

показатель преломления в A' такой же как и в A . Дугу AA' луч проходит за время $\psi(R + D)/v_A$, а BB' — за $\psi R/v_B$, которое, очевидно, равно: $\psi(R + D)/v_A = \psi R/v_B$. Так как $v_A = c/(n - \Delta n)$ $v_B = c/n$, то получаем следующее уравнение на радиус кривизны $Rn = (R + D)(n - \Delta n)$, откуда получаем $R = n/(\Delta n/D)$. Отношение $\Delta n/D$ характеризует быстроту изменения показателя преломления с расстоянием — градиент показателя преломления. Видно, что чем больше градиент показателя преломления, тем круче изгибается световой пучок, причем в сторону увеличения показателя преломления.

Это явление ответственно за *рефракцию света в атмосфере* — *искривление световых лучей при прохождении в атмосфере, вызванное оптической неоднородностью атмосферного воздуха*. Различают *астрономическую* и *земную рефракцию*. В первом случае рассматривается искривление световых лучей, приходящих от небесных тел (Солнца, звезд и т.п.), а во втором — от земных объектов. В обоих случаях *вследствие искривления лучей наблюдатель видит объект не в том направлении, какое соответствует действительности; объект может представляться искаженным*. Возможно наблюдение объекта даже тогда, когда тот фактически находится за линией горизонта.

Явление рефракции также позволяет объяснить ряд явлений, наблюдаемых, например, при закате солнца: сплюснутость солнечного диска, появление слепой полосы, кажущееся увеличение солнечного диска. Искривлением лучей света в неравномерно нагретом воздухе объясняются и такие интересные явления как *миражи*, которые бывают разных видов. Например, в жаркой пустыне возникает иллюзия находящейся на горизонте водной глади или на асфальте, нагретого солнечными лучами видны "лужи", исчезающие при приближении к ним.

ЛЕКЦИЯ №6

Интерференция света: интерференционные явления в оптике, когерентность; интерференция монохроматических волн. Методы реализации интерференционной картины; интерферометр Майкельсона.

Волновые свойства света наиболее отчетливо проявляют себя в явлениях *интерференции* и *дифракции*. Эти явления характерны для волн любой природы, но для световых волн их можно наблюдать лишь при *определенных условиях*.

6.1. Интерференция света. Интерференционные явления в оптике, когерентность. Если при наложении двух световых пучков интенсивность света на экране оказывается не равной сумме интенсивностей, а периодически меняется от точки к точке, образуя систему темных и светлых полос, то говорят, что имеет место *интерференция света*. Подобный опыт наиболее просто осуществить, используя пучки лазерного света. Образно говорят, что при интерференции **свет + свет = темнота** (Рис.6.1).

Интерференция света весьма тонкое физическое явление, которое не часто наблюдается в обыденной жизни, но в физике играет важную роль. Из повседневного опыта мы знаем, что если комната освещается двумя лампочками, то освещенность в любой точке равна сумме освещенностей, создаваемых каждой лампочкой отдельно. Этот экспериментальный факт, иногда называемый *законом сложения интенсивностей*, имеет место для *независимых источников света* (ламп, свечей и т. п.). Но возможны ситуации, когда *закон сложения интенсивностей нарушается*. Так бывает, если накладываются пучки света, исходящие из одного и того же источника, но прошедшие разные оптические пути. Опыт показывает, что **при определенных условиях** наложение таких пучков дает *распределение интенсивности света в виде чередующихся темных и светлых полос*, которое называют *интерференционной картиной* или *интерферограммой*.

Если допустить, что *свет представляет собой волну*, то явление *интерференции* легко объясняется. Действительно, при наложении двух *одинаковых волн* в некоторых точках пространства колебания *усиливают* друг друга, если они находятся "в

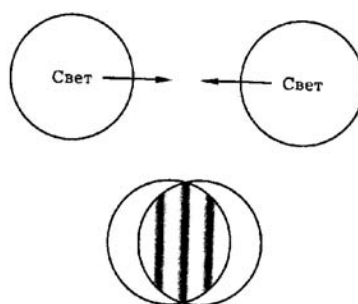


Рис. 6.1. Явление интерференции. При наложение двух световых пучков образуется интерференционная картина — чередование темных и светлых полос, т.е. происходит перераспределение света по пространству: темные полосы соответствуют отсутствию света, светлые — усиленному свету. В пределах темных полос на экране два световых пучка практически полностью гасят друг друга.

фазе", т.е. встречаются два гребня (два максимума) или две впадины (два минимума), в то время как в других точках колебания *взаимно гасятся*, если гребень одной волны встречается с впадиной другой (волны находятся "в противофазе"). При этом в тех областях, где *освещенность максимальна*, его *интенсивность* не в два, а в *четыре раза* больше той, которая была бы при наличии одной волны. В других точках, где *освещенность минимальна*, — *интенсивность* падает до нуля.

Отметим, что максимумы и минимумы освещенности, наблюдаемые в интерференционных картинах, не связаны с какими-либо превращениями энергии света — *в местах минимумов световая энергия не переходит в другие форма энергии*, происходит лишь **перераспределение** светового потока, в результате чего максимумы освещенности в одних местах *компенсируются* минимумами в других. Закон сохранения энергии при этом не нарушается.

В оптике явление интерференции впервые наблюдалось в ставшем классическим опыте Юнга (1801 г) — **Рис.6.2**. Здесь свет от источника S проходит сначала через *маленькое отверстие* в экране Q_1 , а затем падает на другой экран Q_2 с *двумя маленькими отверстиями*, разнесенными на некоторое расстояние. Прошедший через отверстия свет падает на экран Q_3 , где и наблюдается интерференционная картина. Опыт Юнга был первым убедительным доказательством того, что наложение света может образовать темноту, а наблюдение интерференции в опыте Юнга явилось экспериментальным доказательством *волновой природы* света.

При объяснении явления интерференции важным понятием является понятие **когерентности света**. Исторически оно возникло в связи с интерференционными опытами. Появление интерференционной картины в опыте Юнга (и в других опытах) зависит от того, какой свет падает на экран Q_2 с двумя точечными отверстиями. Если это свет *точечного* источника, каковым является маленькое отверстие в экране Q_1 (**Рис.6.2**), то интерференция есть. Если же это свет от протяженного источника или свет, рассеянный матовой пластинкой, то интерференции нет.

Способность света давать интерференционную картину называют *когерентностью*. Когерентность связана со *структурой* света: *когерентный свет* — это свет, структура которого близка к *плоской или сферической гармонической волне*. Про такой свет говорят, что он имеет *высоко упорядоченную структуру*. Понятию когерентности соответствует понятиям "согласование", "корреляция". В противоположность этому *некогерентный свет* — это свет, не способный давать интерференцию и подчиняющийся закону сложения интенсивностей. Такой свет представляет

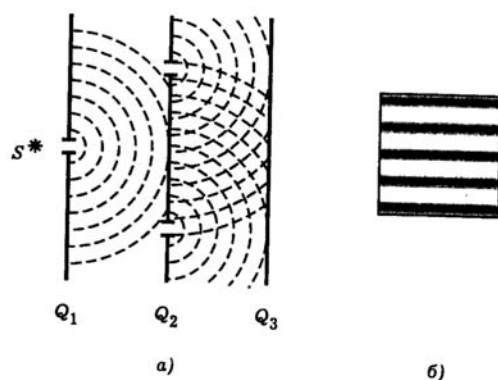


Рис. 6.2. Опыт Юнга. Схема опыта (а) и фрагмент наблюдаемой интерференционной картины (б)

собой случайно модулированную волну, т.е. волну, у которой амплитуда и фаза описываются случайными функциями.

Рассмотрим несколько характерных случаев интерференции на примере *монохроматических* волн. Монохроматические волны считаются *когерентными*, если они имеют *одинаковые частоты*, а *разность фаз между ними остается неизменной с течением времени*. Такие волны интерферируют. Подчеркнем, что *интерференция имеет место для волн одинаковой поляризации*.

Интерференция монохроматических волн. Интерференция сферических монохроматических волн. Рассмотрим интерференцию от *двух одинаковых синфазных монохроматических точечных источников* S_1 и S_2 , находящихся на расстоянии d друг от друга (**Рис.6.3**). Это соответствует опыту Юнга (**Рис.6.2**), где интерферируют *сферические волны*. В настоящее время постановка этого опыта существенно упрощается при использовании в качестве источника лазера. При этом отпадает необходимость экрана Q_1 с отверстием S .

Если расстояние l от плоскости Q_2 до экрана, где наблюдают интерференционную картину, удовлетворяет условию $l \gg d$, то можно считать амплитуды обеих волн в точке наблюдения практически *одинаковыми*. Согласно принципу суперпозиции для напряженности поля в точке P запишем

$$E = E_0 \cos(\omega t - kr_1) + E_0 \cos(\omega t - kr_2) = 2E_0 \cos\left[\frac{1}{2}k(r_2 - r_1)\right] \cos\left[\omega t - \frac{1}{2}k(r_1 + r_2)\right] \quad (6.1)$$

где $r_{1,2}$ — расстояния от источников до точки наблюдения (**Рис. 6.3**). Величину $\Delta = r_2 - r_1$ называют *разностью хода* интерферирующих волн.

Так как интенсивность результирующей волны пропорциональна квадрату амплитуды $A = 2E_0 \cos[k(r_2 - r_1)/2]$, то

$$I = \frac{c}{8\pi} A^2 = 4I_0 \cos^2\left[\frac{1}{2}k(r_2 - r_1)\right] = 2I_0[1 + \cos(k \cdot \Delta)], \quad (6.2)$$

$I_0 = (c/8\pi)E^2$ — интенсивность волны, испускаемой одним источником. Видно, что освещенность экрана *периодически* меняется в пространстве — в одних местах она максимальна, в других — минимальна.

Положение **максимумов**, соответствующее $\cos(k \cdot \Delta) = 1$, определяется условием

$$k \cdot \Delta = \pm 2\pi m, \quad (6.3)$$

где целое число $m = 0, 1, 2, \dots$ — *порядок интерференции*. Так как $k = 2\pi/\lambda$, то условие (6.3) принимает вид $\Delta = m\lambda$, т.е. *разность хода равна целому числу длин*

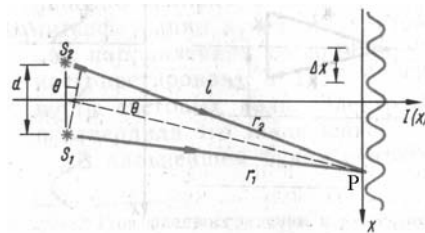


Рис. 6.3. К интерференции волн, испускаемых точечными источниками S_1 и S_2 .

волн. Когда волны распространяются в среде с показателем преломления n геометрическую разность хода заменяют на *оптическую разность хода* интерферирующих волн: $\Delta = n(r_2 - r_1)$.

Положение **минимумов** соответствует $\cos k\Delta = -1$ или

$$k \cdot \Delta = (2m + 1)\pi, \quad (6.4)$$

т.е. *разности хода, равной нечетному числу полуволн* $\Delta = (2m + 1)\lambda/2$.

В областях минимума освещенность равна нулю, а в максимумах — в четыре раза больше интенсивности исходных волн. *Светлые полосы* также называют областями *конструктивной интерференции* (здесь интерференционное слагаемое положительно и имеет место *усиление света*), а *темные* — области *деструктивной интерференции* (интерференционный член отрицательный — *гашение* или *ослабление света*).

Если начальная разность фаз $\Delta\varphi = \varphi_1 - \varphi_2 \neq 0$, картина будет той же самой, но темные и светлые полосы принимают некоторое промежуточное положение, зависящее от $\Delta\varphi$. Действительно, в этом случае условие, например, максимума интенсивности в интерференционной картине имеет вид $\Delta/\lambda + \Delta\varphi/2\pi = m$. Следовательно, отличие $\Delta\varphi$ от нуля эквивалентно тому, что Δ/λ не равно целому числу, как это было при $\Delta\varphi = 0$.

Чтобы получить *зависимость освещенности экрана от координаты*, выразим разность хода Δ через координату x точки наблюдения. Для этого введем угол θ , образуемый направлением на точку P с перпендикуляром к линии, соединяющей источники (**Рис. 6.3**). При $\theta \ll 1$ ($d \ll l$) для разности хода можно записать $\Delta \simeq d \cdot \theta$. Учитывая, что $\theta \simeq x/l$ имеем $\Delta \simeq x \cdot d/l$. Подставляя Δ в (6.2), получаем

$$I = 2I_0 \left[1 + \cos \left(\frac{kdx}{l} \right) \right] \quad (6.5)$$

В $x = 0$ расположен *максимум*, соответствующий нулевой разности хода. Для него порядок интерференции $m = 0$. Расстояние между *соседними максимумами* (или минимумами) Δx определяется из условия $kd\Delta x/l = 2\pi$, откуда $\Delta x = 2\pi l/kd = \lambda l/d$. Введем угол схождения лучей, т.е. угол, под которым видны источники из точки наблюдения, $\alpha \simeq d/l$, для Δx получаем: $\Delta x \simeq \lambda/\alpha$.

Интерференция плоских монохроматических волн. Пусть в пространстве распространяются две *монохроматические плоские волны* с *одинаковой поляризацией*, *частотой* ω и *одинаковыми начальными фазами*, которые положим равными нулю. В комплексном представлении плоские монохроматические волны имеют вид: $E_{1,2} = F_{1,2}(\vec{r})e^{-i\omega t} + \text{к.с.}$, где $F_{1,2}(\vec{r}) = A_{1,2}(\vec{r})e^{i\varphi_{1,2}(\vec{r})}$ — комплексные амплитуды, а $\varphi_{1,2}(\vec{r}) = \vec{k}_{1,2}\vec{r}$ — фазы каждой волны в точке наблюдения, соответственно. Предполагаем, что волновые векторы $\vec{k}_1, \vec{k}_2$ имеют углы $\pm\alpha$ с осью z и лежат

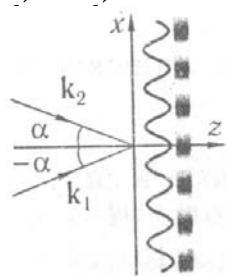


Рис. 6.4. Интерференция плоских волн. Волновые векторы $\vec{k}_{1,2}$ плоских волн лежат в плоскости xz , под углом $\pm\alpha$ к оси z .

в плоскости xz (**Рис.6.4**). Проекции векторов $\vec{k}_1$ и $\vec{k}_2$ на ось y (перпендикулярно плоскости рисунка) равны нулю.

Согласно принципу суперпозиции колебательный процесс в любой точке наблюдения есть сумма колебаний, поэтому комплексная амплитуда $F(\vec{r})$ результирующей волны $E = F(\vec{r})e^{-i\omega t}$ равна сумме комплексных амплитуд $F_1(\vec{r})$ и $F_2(\vec{r})$: $F(\vec{r}) = F_1(\vec{r}) + F_2(\vec{r})$, а ее интенсивность $I \sim |F(\vec{r})|^2$ равна

$$|F(\vec{r})|^2 = |F_1(\vec{r})|^2 + |F_2(\vec{r})|^2 + 2\text{Re } F_1^*(\vec{r})F_2(\vec{r}), \quad (6.6)$$

или переходя к интенсивности $I = (c/8\pi)|F|^2$, перепишем (6.6) в виде:

$$I(\vec{r}) = I_1(\vec{r}) + I_2(\vec{r}) + 2\sqrt{I_1 I_2} \cos(\Delta\varphi) \quad (6.7)$$

где $\Delta\varphi = \varphi_2(\vec{r}) - \varphi_1(\vec{r}) = \vec{K}\vec{r}$ — разность фаз слагаемых колебаний в точке наблюдения; $\vec{K} = \vec{k}_2 - \vec{k}_1$, $I_{1,2}$ — интенсивности исходных волн.

Из (6.7) видно, что интенсивность результирующей волны складывается из интенсивностей источников (слагаемые $I_{1,2}$) плюс *интерференционное* слагаемое. Таким образом, *интенсивность результирующей волны не равна сумме интенсивностей слагаемых волн*. В пространстве, где волны перекрываются, происходит *перераспределение* потоков энергий: в некоторых точках пространства результирующая интенсивность оказывается больше суммы интенсивностей слагаемых волн, а в других, наоборот, меньше. Результат зависит от третьего слагаемого $2\sqrt{I_1 I_2} \cos(\Delta\varphi)$, т.е. от *разности фаз $\Delta\varphi$ колебаний в данной точке наблюдения*. Оно описывает *интерференцию*.

Разность фаз обусловлена тем, что волны проходят разные пути до точки наблюдения или, как говорят, *разностью хода*. При $\Delta\varphi = 2\pi m$, соответствующей разности хода равной целому числу длин волн, $\cos(\Delta\varphi) = 1$ имеется *интерференционный максимум m -го порядка* (светлые полосы), а при $\Delta\varphi = (2m + 1)\pi$, соответственно — *интерференционный минимум*. При $I_1 = I_2$ в областях *минимума* света действительно нет, что соответствует темным полосам.

Поскольку проекции волновых векторов $\vec{k}_{1,2}$ на оси x, y, z соответственно равны $k_{1x} = k \sin \alpha$, $k_{1y} = k_{2y} = 0$, $k_{2x} = -k \sin \alpha$, $k_{1z} = k_{2z} = k \cos \alpha$, то с учетом этого перепишем (6.7) в виде:

$$I(x) = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos(2kx \sin \alpha). \quad (6.8)$$

Видно, что результирующая интенсивность не зависит от z , т.е. одинакова в любой плоскости наблюдения $z = \text{const}$, и поскольку она не зависит также от y , то в любой плоскости $z = \text{const}$ она имеет вид чередующихся светлых и темных полос, параллельных y , т.е. перпендикулярных плоскости рисунка (**Рис.6.4**).

Расстояние между двумя соседними максимумами интенсивности (интерференционными полосами) или между двумя соседними минимумами называют *шириной интерференционной полосы*. Очевидно, что переход от одной светлой полосы к другой соответствует изменению аргумента синуса в (6.8) на 2π : $2kx \sin \alpha \Delta x = 2\pi$. Отсюда находим

$$\Delta x = l = \frac{\pi}{2 \sin(\beta/2)} \simeq \frac{\lambda}{\beta}, \quad \beta \ll 1, \quad \beta = 2\alpha. \quad (6.9)$$

Контраст интерференционной картины характеризуют величиной, называемой *видностью*, которая определяется как

$$V = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} = \frac{2\sqrt{I_1 I_2}}{I_1 + I_2} \quad (6.10)$$

Видность максимальна и равна единице при равных интенсивностях интерферирующих волн $I_1 = I_2$ ($I_{\min} = 0$) и близка к нулю, когда интенсивность одной и интерферирующих волн существенно больше другой.

6.2. Методы реализации интерференционной картины. Рассмотренные выше примеры интерференции являются *идеализированными*, так при интерпретации предполагалось, что источник света является точечным и излучает *монохроматические сферические* или *плоские волны*. Поэтому эти результаты имеют *ограниченную применимость*, так как реальные источники света не являются таковыми. *Реальные* источники испускают *немонохроматический свет*, состоящий из отдельных *цугов*, соотношение фаз между которыми имеют случайный характер. Длительность цугов τ_0 , как правило, не превышает 10^{-7} с, а его протяженность $\tau_0 c$ (c — скорость света) менее одного метра даже в случае узких спектральных линий излучения газоразрядных источников низкого давления. Только *лазеры* позволяют получать *когерентные* световые волны, близкие по свойствам к монохроматическим плоским или сферическим волнам.

Излучение обычных (нелазерных) источников света представляют собой наложение огромного числа *несогласованных* между собой цугов волн, т.е. беспорядочные *некогерентные* колебания, которые *не могут интерферировать*. Действительно, каждая частотная компонента немонхроматического излучения создает свою интерференционную картину (полосы) в плоскости наблюдения. Эти полосы накладываются друг на друга, причем максимумы одной картины могут совпадать с минимумами другой. В результате освещенность экрана оказывается однородной, т.е. интерференция исчезает. Поэтому для наблюдения интерференции с использованием

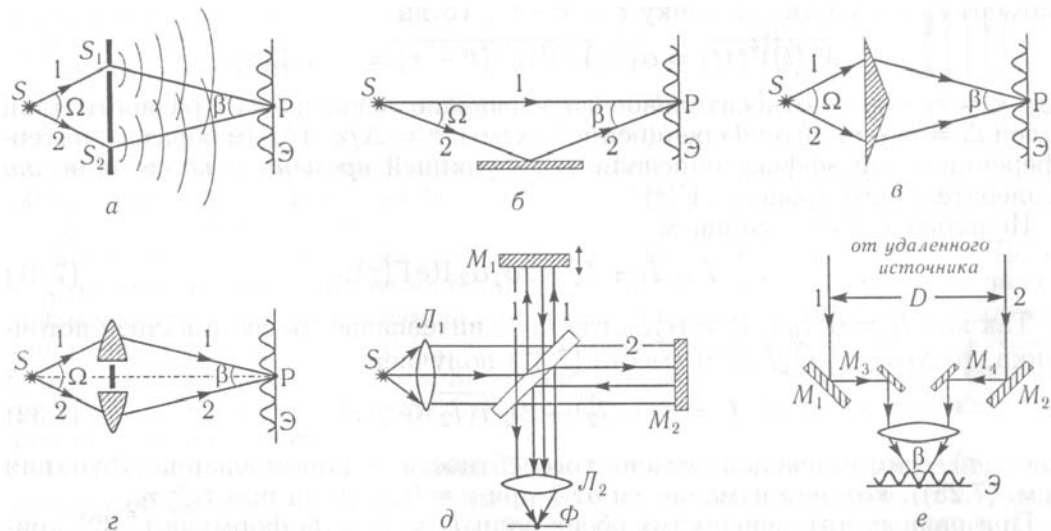


Рис. 6.5. Некоторые интерференционные схемы, широко применяемые в оптике: а — схема Юнга, б — схема Ллойда, в — бипризма Френеля, г — билинза Бийе, д — интерферометр Майкельсона, е — звездный интерферометр; все лучи 1 и 2 идут от удаленного источника

немонохроматического света приходится прибегать к различным ухищрениям: применять спектральные фильтры, располагать источники таким образом, чтобы их можно было считать точечными и другим.

Наблюдать интерференцию света от некогерентных источников можно, если разделить излучение на *два* или *несколько пучков*, а затем свести их вместе. Хотя в каждом из пучков за время наблюдения фазовые соотношения между цугами хаотически изменяются эти изменения одинаковы в разных пучках. Интерференционная картина будет наблюдаться, если *разность хода между пучками не превышает длины отдельного цуга*.

Способы разделения волны от первичного источника на две когерентные между собой волны можно разбить на две группы: *деление волнового фронта* и *деление амплитуды*. На **Рис.6.5** приведены примеры нескольких интерференционных схем, широко применяемых на практике. Схемы на **Рис. 6.5** а-г относятся к первой группе. Во всех схемах излучение источника S попадает на экран по двум различным путям (плечам) интерференционной схемы, отмеченным на рисунках индексами 1 и 2. В любом случае вместо первичного источника можно рассматривать два когерентных источника S_1 и S_2 , находящихся на некотором расстоянии d . Для определения положения и формы интерференционных полос излучение можно считать монохроматическим. Наиболее важными характеристиками любой интерференционной схемы являются: *угол схождения* волны β в точке наблюдения и угол Ω между лучами, исходящими из источника S , которые каким-либо способом сводятся далее в точку наблюдения. Угол Ω называется *апертурой интерференции*.

Две когерентные световые волны можно получить в результате *отражения света* от двух поверхностей плоскопараллельной прозрачной пластинки (пленки) (**Рис.6.6** а). В каждую точку наблюдения приходят волны, разность хода которых будет такой же, как от источников S_1 и S_2 — изображений S , создаваемых верхней и нижней поверхностями. Интерференционные полосы на экране B будут иметь вид концентрических колец с центром в точке O .

При *точечных* источниках света будут наблюдаться резкие интерференционные картины при любом положении экрана, пересекающего интерференционные полосы. Такие полосы интерференции называют *нелокализованными*.

У *протяженного источника* каждый элемент можно рассматривать как точечный источник, излучение которого некогерентно с излучением других элементов источника. Интерференционные картины, создаваемые разными элементами, просто складываются, и если полосы в них смещены одна относительно другой, то результирующая картина оказывается размытой. Совпадение интерференционных полос в отдельных картинах будет только при определенных условиях наблюдения (или при расположении экрана в определенном месте), поэтому интерференционные полосы в

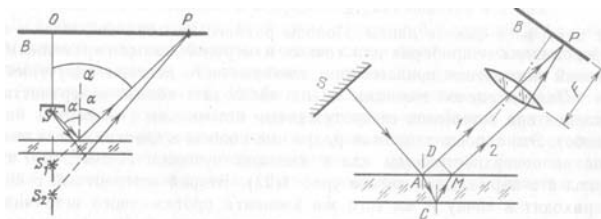


Рис. 6.6. а — Интерференция света при отражении двумя параллельными плоскостями; б — схема наблюдения полос равного наклона.

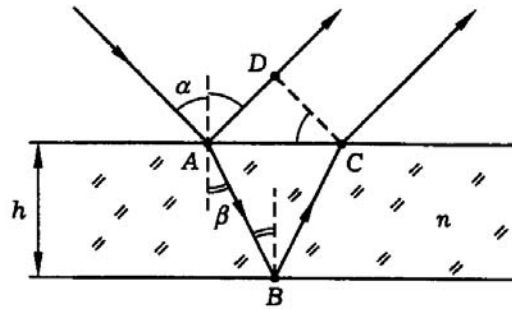


Рис. 6.7. Интерференция света на плоскопараллельной пластинке

случае протяженного источника оказываются *локализованными*. Различают *полосы равного наклона* и *полосы равной толщины*.

Полосы равного наклона. Полосы равного наклона можно наблюдать в фокальной плоскости линзы, поставленной на пути излучения от протяженного источника, отраженного от плоскопараллельной пластинки (**Рис. 6.6 б**). Говорят, что они локализованы в бесконечности. В этом случае в каждую точку P экрана B , приходят лучи, которые до линзы были параллельны друг другу. От точки A разделения лучей луч 2 проходит дополнительно расстояние $|AC| + |CM|$ в пластинке, а в воздухе его путь меньше, чем у луча 1 на отрезок $|AD|$. Разность хода не лучей 1 и 2 не зависит от положения точки A падения первичного луча на пластинку и результат интерференции в точке P будет одинаковым для всех элементов протяженного источника. Полосы равного наклона также наблюдают в *интерферометре Майкельсона* (**Рис. 6.5 д**) и в звездном интерферометре — **Рис. 6.5 е**.

Пусть *плоская монохроматическая световая волна* падает на плоскопараллельную пластинку, обе грани которой частично отражают свет, как показано на (**Рис. 6.7**). Подсчитаем оптическую разность хода лучей Δ , отраженных передней и задней гранями пластинки.

Обозначим толщину пластинки h , угол падения света α , угол преломления β , показатель преломления материала пластинки n . В силу закона Снеллиуса $\sin \alpha = n \cos \beta$. Так как фазовый набег световой волны в среде пропорционален показателю преломления n , искомая разность хода лучей равна $\Delta = 2ABn - AD$, $AD = AC \sin \alpha$, $AC = 2AB \sin \beta$, $AB = h / \cos \alpha$. Отсюда получаем

$$\Delta = 2hn \cos \beta \quad (6.11)$$

Для разности фаз интерферирующих волн получаем:

$$\delta = k\Delta = \frac{4\pi hn}{\lambda} \cos \beta = \frac{4\pi h}{\lambda} \sqrt{n^2 - \sin^2 \alpha}. \quad (6.12)$$

Здесь λ — длина волны в вакууме. Формулы (6.11) и (6.12) записаны без учета потери полуволны при отражении от границы раздела.

Таким образом, *разность хода лучей, отраженных гранями плоскопараллельной пластинки, зависит от угла падения света на пластинку ("угла наклона") и не зависит от положения источника*, т.е. если источник переместить, разность фаз останется прежней. Отсюда следует, что при использовании протяженного источника света интерференционные полосы также будут отчетливыми, как и в случае

с точечным источником света. Это означает, что каждый элемент протяженного источника создает в фокальной плоскости линзы свою интерференционную картину, причем положение полос во всех этих картинах совершенно одинаково. Про такие полосы говорят, что они *локализованы*, в данном случае либо в бесконечности, либо в фокальной плоскости линзы.

С другой стороны, в соответствии с формулой (6.12) светлые полосы расположены в местах, для которых $\delta = m\lambda$ (m — целое число или *порядок интерференции*). Полоса, соответствующая данному порядку интерференции, обусловлена светом, падающим на пластинку под вполне определенным углом α . Поэтому такие полосы называют *интерференционными полосами равного наклона*. Если ось линзы расположена перпендикулярно пластинке, полосы имеют вид концентрических окружностей с центром в фокусе, причем в центр картины порядок интерференции максимален (докажите это).

Отметим, что аналогичные полосы равного наклона, в принципе, можно наблюдать и в прошедшем свете. В этом случае один из лучей проходит прямо а другой — после двух отражений на поверхности пластики.

Полосы равной толщины. Полосы равной толщины возникают в результате интерференции от мест одинаковой толщины. Их можно наблюдать вблизи поверхности тонкой пленки или клина при освещении протяженным источником света. Они *локализованы* вблизи поверхности пленки или клина. Это часто наблюдаемые *радужные полосы* в пленках масла или бензина на поверхности воды или в мыльных пузырях.

На **Рис. 6.8** показана картина отражения световых лучей, вышедших из одной точки источника света и отраженных от поверхностей пленки. Луч 1 образует угол θ с нормалью к поверхности пленки, второй луч 2 приходит в точку P после отражения от нижней поверхности пленки в точке B . Обычно толщина пленки меняется мало, поэтому разность хода лучей 1 и 2 приближенно равна $\Delta = n(|AB| + |BP|) - |CP| \simeq 2h\sqrt{n^2 - \sin^2 \theta}$. Так как в глаз попадают только те лучи из точки P , для которых угол θ лежит в узком интервале значений, разность хода практически одинакова для всех пар интерферирующих лучей, создающих изображение точки P на сетчатке глаза. Контур каждой интерференционной полосы проходит по тем местам пленки, где ее толщина имеет одно и то же значение и интервал изменений $\sin \theta$ мал. Для тонкой пленки эти полосы локализуются вблизи отражающей поверхности. Получающаяся интерференционная картина получила название *полосы равной толщины*.

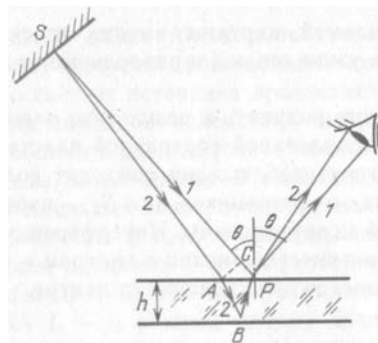


Рис. 6.8. Схема наблюдения полос равной толщины.

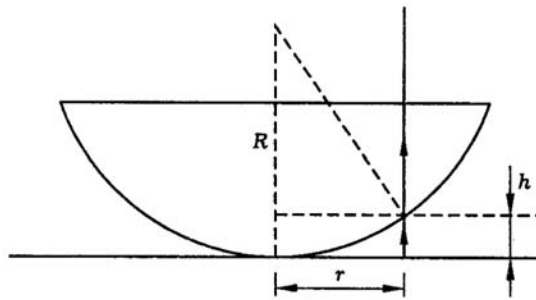


Рис. 6.9. К расчету радиусов колец Ньютона

Отметим, что положение максимумов и для полос равного наклона и для полос равной толщины зависит от длины волны λ . Поэтому темные и светлые полосы получаются при освещении монохроматическим светом. При наблюдении в белом свете получается совокупность смещенных друг относительно друга полос, образованных лучами разных длин волн, и интерференционная картина приобретает радужную окраску.

Кольца Ньютона. Полосы равной толщины можно наблюдать с помощью стеклянной пластинки и линзы. Если прижать выпуклую поверхность линзы к пластинке и направить на нее параллельный пучок света, то лучи, отраженные пластинкой и поверхностью линзы будут интерферировать, образуя систему интерференционных линий в виде концентрических темных и светлых колец, которые называют *кольцами Ньютона*. Вычислим радиусы интерференционных колец. В данном случае разность интерферирующих лучей есть $\Delta = 2h$, где h — расстояние от поверхности линзы до пластинки. Из Рис. 6.9 видно, что $h = R - \sqrt{R^2 - r^2}$, где R — радиус кривизны поверхности линзы, r — расстояние от точки касания линзы и пластинки до светового луча. Полагая $r \ll R$, получим приближенное выражение $\Delta = r^2/R$. Радиусы светлых колец определяются условием $\Delta = m\lambda$, λ — длина световой волны, $m = 0, 1, 2, \dots$. Таким образом, для радиусов колец Ньютона получаем выражение:

$$r_m = \sqrt{mR\lambda} \quad (6.13)$$

Интерферометр Майкельсона широко используется на практике. Его схема показана на Рис. 6.10. Пучок света от монохроматического источника направляется на светоделительную пластину, где делится на два — прошедший и отраженный, примерно одинаковой интенсивности. Пройдя некоторые расстояния, эти пучки попадают на зеркала, отражаются ими в обратных направлениях и вновь падают на делительную пластину. Пластина снова частично отражает и частично пропускает свет, в результате чего образуется пучок света, представляющий собой смесь пучков, прошедших через разные плечи интерферометра. Этот пучок света наблюдается на экране. Двигая одно из зеркал интерферометра в направлении падающего на него светового пучка, можно изменять оптическую разность хода лучей и наблюдать изменение интерференционной картины.

Пусть источник света попадающий в интерферометр Майкельсона представляет плоскую монохроматическую световую $E_0 = A_0 \cos(\omega t - kz)$. Обозначим через r и t коэффициенты отражения и пропускания света по (по амплитуде световой волны) делительной пластинкой, предполагая, что поглощением света в пластике можно пренебречь ($r^2 + t^2 = 1$). На пути от источника света до плоскости наблюдения волна делится на две волны, причем каждая по одному разу пропускается и отражается

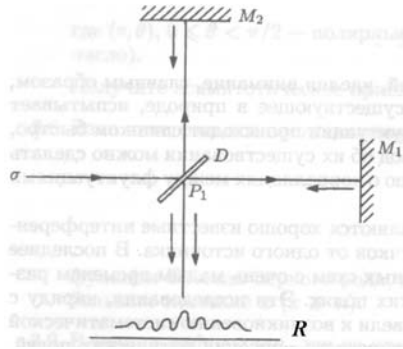


Рис. 6.10. Схема интерферометра Майкельсона. σ — источник света; D — делитель пучка; $M_{1,2}$ — зеркала; R — плоскость наблюдения.

делительной пластинкой. На выходе интерферометра они могут интерферировать, образуя интерференционную картину в плоскости наблюдения, описываемую формулой (6.7), которая для равных интенсивностей ($r = t = 1/2$) принимает вид.

$$I = \frac{1}{2} I_0 [1 + \cos(\Delta\varphi)]. \quad (6.14)$$

Здесь величина $\Delta\varphi = k \cdot \Delta$ — *разность фаз* волн, прошедших разные плечи интерферометра, $\Delta = |l_1 - l_2|$ — разность хода, $k = 2\pi/\lambda$. Также предполагается, что светоделительная пластинка имеет бесконечно тонкую толщину и не дает вклада в разность хода.

Из формуле (6.14) видно, что интенсивность света на выходе интерферометра осциллирует от нуля при $\Delta\varphi = (2m + 1)\pi$ до I_0 (при $\Delta\varphi = 2m\pi$), если одно из зеркал интерферометра смещается. Так как пропускание зависит от длины волны, то *интерферометр Майкельсона может служить фильтром, или селективным по длинам волн отражателем света*. В этой последней роли он часто используется для *селекции частот в лазерах*.

Отметим, что интерференционная картина, наблюдаемая с помощью интерферометра Майкельсона, весьма чувствительна к различным малым возмущениям, таким как механическое смещение зеркал, изменение состава, температуры и плотности среды, через которую проходят световые лучи. Поэтому интерферометр можно использовать для различных *точных измерений*. Например, с помощью интерферометра можно измерить механическое смещение порядка длины световой волны, т. е. около $\sim 10^{-4}$ см. При необходимости точность измерения может быть повышена еще на несколько порядков. Такие точные измерения используют, например, в сейсмологии для регистрации движений земной коры.

Интерферометр Майкельсона используют для *абсолютного* измерения длин световых волн. Для этого нужно подсчитать число максимумов, возникающих на экране наблюдения при смещении подвижного зеркала интерферометра на известное расстояние Δz . Тогда длина волны $\lambda = 2\Delta z/N$. Эта техника использовалась для очень точных измерений длин волн излучения лазеров.

ЛЕКЦИЯ №7

Многолучевая интерференция, интерферометр Фабри-Перо. Интерференция квазимонохроматического света: временная когерентность и время когерентности; функция когерентности; пространственная когерентность и площадь когерентности.

7.1. Многолучевая интерференция, интерферометр Фабри-Перо. До сих пор мы рассматривали интерференцию двух перекрывающихся волн. Для увеличения резкости интерференционной картины часто используют принцип *многолучевой интерферометрии*, когда в образовании интерференции участвуют сразу *несколько* (более двух) световых волн. Рассмотрим основные особенности многолучевой интерференции на примере плоско параллельной пластинки, обе грани которой хорошо отражают свет (**Рис. 7.1**), либо двух зеркал, установленных параллельно. Первый из них называют *эталоном Фабри-Перо*, а второй — *интерферометром Фабри-Перо*.

Пусть плоская монохроматическая световая волна падает на эталон Фабри-Перо. Свет, частично проникая внутрь, многократно отражается от граней, как показано на **Рис. 7.1**. Для наблюдения интерференции на выходе интерферометра ставят линзу, в фокальной плоскости которой наблюдают полосы равного наклона. При каждом отражении часть излучения выходит наружу, образуя набор волн, которые интерферируют между собой. В прошедшем свете интерференционная картина имеет вид узких светлых полос на почти совершенно темном фоне (**Рис. 7.2**). (Подумайте, что будет в отраженном свете.)

Рассмотрим основные характеристики возникающей при этом интерференционной картины. Пусть ρ, τ — коэффициент отражения и пропускания света по амплитуде волны для каждой из граней пластинки, причем $\rho^2 + \tau^2 = 1$ (поглощение не учитываем); $\rho^2 = R, \tau^2 = T$ — коэффициенты отражения и пропускания по интенсивности. Обозначим амплитуду падающей волны A_0 , а *прошедшей* через пластинку — A ,

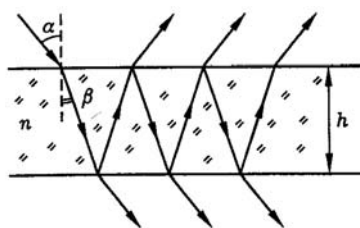


Рис. 7.1. Схема многолучевой интерферометрии света

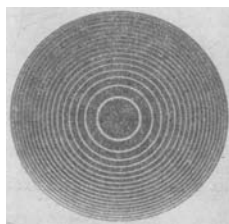


Рис. 7.2. Интерференционная картина в эталоне Фабри-Перо (линии равного наклона).

которую можно записать как суперпозицию (сумму) всех прошедших лучей:

$$A = A_0\tau^2 + A_0\tau^2\rho^2e^{i\delta} + A_0\tau^2\rho^4e^{2i\delta} + \dots = A_0\tau^2(1 + \rho^2e^{i\delta} + \rho^4e^{2i\delta} + \dots), \quad (7.1)$$

где $\delta = k \cdot \Delta$ — набег фазы световой волны за двойной проход через пластинку; $\Delta = 2hn \cos \beta$ — оптическая длина пути, n — показатель преломления материала пластинки, h — ее толщина, β — угол преломления света в пластинке, связанный с углом падения α формулой $\sin \alpha = n \sin \beta$. В формуле (7.1) учтено, что первый прошедший луч дважды проходит через границы пластинки, второй — дважды проходит и дважды отражается на границах, третий — дважды проходит и четырежды отражается и т. д.

Для набег фазы можно записать

$$\delta = \frac{4\pi}{\lambda} hn \cos \beta. \quad (7.2)$$

В круглых скобках (7.1) стоит сумма бесконечно убывающей геометрической прогрессии со знаменателем $q = \rho^2e^{i\delta}$, модуль которого меньше единицы. Используя формулу $\sum_{n=0}^{\infty} q^n = (1 - q)^{-1}$, находим

$$A = A_0 \frac{\tau^2}{1 - \rho^2e^{i\delta}}. \quad (7.3)$$

В результате для коэффициента пропускания света пластинкой $T(\delta) = I/I_0 = |A/A_0|^2$ получаем следующее выражение:

$$T(\delta) = \frac{\tau^4}{|1 - Re^{i\delta}|^2} = \frac{1}{1 + F \sin^2 \delta/2}, \quad (7.4)$$

где величина $F = 4R/(1 - R)^2$ называется *фактором резкости*, который характеризует резкость интерференционной картины.

На **Рис. 7.3** показана зависимость $T(\delta)$ для двух разных значений параметра $F = 0.1$ и $F = 10$. Видно, что *чем F больше, тем полосы пропускания интерферометра уже* (интерференционная картина резче выражена). Сравнение интерференционных картин, наблюдаемые с помощью интерферометра Майкельсона и интерферометра Фабри-Перо, показывает, что *многолучевая интерференция значительно увеличивает резкость интерференционной картины*, т.е. интерференционные полосы в интерферометре Фабри-Перо значительно уже, чем в двухлучевых интерферометрах. Фактор резкости F тем больше, чем выше коэффициент отражения на гранях пластинки. Для зеркальной поверхности R может достигать величины

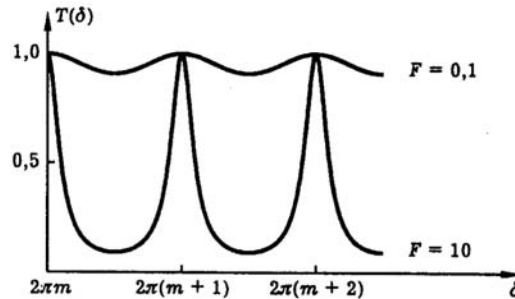


Рис. 7.3. Зависимость функции $T(\delta)$ при разных значениях фактора резкости F

$R = 0.99$ и выше. В этом случае $F \geq 4 \cdot 10^4$. Высокая резкость интерференции в интерферометре Фабри-Перо позволяет использовать его как *спектральный прибор высокой разрешающей способности*.

7.2. Интерференция квазимонохроматического света. Уже отмечалось, что обычные (нелазерные) источники света являются немонахроматическими, поэтому их излучение не может интерферировать. Однако во многих случаях их можно считать квазимонохроматическими. При описании интерференции **квазимонохроматического** света важную роль играет понятие **когерентности**. Различают *временную* и *пространственную когерентность*.

Временная когерентность и время когерентности. Понятие временной когерентности рассмотрим на примере интерферометра Майкельсона (**Рис. 6.5** д). Предположим, что свет является квазимонохроматическим, т.е. его ширина частот $\Delta\omega$ мала по сравнению со средней частотой ω_0 этой полосы ($\Delta\omega \ll \omega_0$). С помощью полупрозрачной пластинки световой пучок разделяется в интерферометре Майкельсона на два пучка, которые снова соединяются, приобретая *разность хода* $\Delta = c \cdot \Delta t$ (c — скорость света, Δt — время задержки между пучками света). Если разность хода Δ достаточно мала, в плоскости наблюдения образуются интерференционные полосы. Говорят, что возникновение *интерференционных полос* является проявлением *временной когерентности* между двумя пучками, поскольку *контраст* между полосами зависит от времени задержки Δt . Экспериментально установлено, что интерференционные полосы наблюдаются только тогда, когда выполняется условие

$$\Delta\omega \cdot \Delta t \leq 2\pi. \quad (7.5)$$

Время задержки $\Delta t \simeq 2\pi/\Delta\omega = \tau_c$ называется *временем когерентности* света, а соответствующая разность хода $\Delta = l_c = c \cdot \tau_c = 2\pi c/\Delta\omega$ — *продольной длиной когерентности света* или (коротко) *длиной когерентности*. Поскольку $\omega = 2\pi c/\lambda$, где λ — длина волны, $\Delta\omega = 2\pi c \cdot \Delta\lambda/\lambda_0^2$, выражение для длины когерентности можно записать в виде

$$l_c = \frac{\lambda_0}{\Delta\lambda} \lambda_0, \quad (7.6)$$

где λ_0 — средняя длина волны света.

Приближенно это явление можно интерпретировать следующим образом. Каждая частотная компонента, присутствующая в спектре света, создает в пространстве свою интерференционную картину — периодические полосы. Результирующую картину в плоскости наблюдения можно рассматривать как результат сложения этих распределений. Распределения, созданные различными частотными компонентами, будут иметь различную пространственную периодичность. Поэтому с *увеличением времени задержки* между двумя пучками интерференционная картина будет становиться все менее и менее различимой, так как минимумы одного распределения будут налагаться на максимумы другого. В результате картина окажется однородной, так как в ней не будет выраженных максимумов и минимумов интенсивности, т.е. интерферограмма не формируется. Это происходит, *когда время задержки Δt становится больше времени когерентности τ_c ($\Delta t > \tau_c$), и интерференционные полосы исчезают*.

Математически это явление описывается, используя понятие *корреляций*. Мы уже отмечали, что квазимонохроматический свет можно рассматривать как *стационарный случайный процесс*, и представлять его в виде последовательности медленно

модулированных волновых цугов со средней частотой, равной средней частоте света ω_0 , и длительностью порядка обратной ширины полосы частот света ($\Delta\omega^{-1}$), т.е. порядка времени когерентности τ_c . Делитель пучка интерферометра Майкельсона расщепляет каждый волновой цуг на два одинаковых цуга. Волновые цуги двух пучков вновь складываются в плоскости наблюдения. При этом цуги, порожденные одним падающим волновым цугом, оказываются сдвинутыми друг относительно друга, поскольку между пучками возникает временная задержка. Ясно, что корреляция (согласование) между двумя пучками будет иметь место в плоскости наблюдения в том случае, если время задержки мало по сравнению со временем когерентности ($\Delta t \ll \tau_c$), т.е. когда *налагаются части одного и того же цуга*. Если же время задержки Δt значительно больше, чем время когерентности τ_c , то корреляция будет отсутствовать, так как *налагаются разные цуги*. Таким образом, *наличие или отсутствие интерференционных полос* в плоскости наблюдения связано с *корреляцией или отсутствием корреляции* между двумя световыми пучками, достигающих плоскость наблюдения, соответственно. *Для наблюдения интерференционной картины разность хода не должна превышать длину когерентности, используемого света.*

Для примера, дающего представление о типичных порядках величин времени и длины когерентности, рассмотрим белый свет, занимающий видимый диапазон от $\lambda_1 = 0.4$ мкм до $\lambda_2 = 0.8$ мкм с шириной полосы $\Delta\nu = \nu_1 - \nu_2 = 1/\lambda_1 - 1/\lambda_2 \approx 10^4$ см $^{-1}$ и излучение лазера. В первом случае длина когерентности составляет $l_c = 10^{-4}$ см. Чтобы увеличить длину когерентности, используют спектральные фильтры для выделения более узкой полосы спектра. Для излучения гелий-неонового лазера $\Delta\nu = 0.04$ см $^{-1}$, соответственно, длина когерентности $l_c = 25$ см.

Рассмотрим теперь *количественное описание временной когерентности*. Пусть колебание в некоторой точке наблюдения создано суперпозицией двух *квазимонохроматических волн* $\mathcal{E}_{1,2}(t) = E_{1,2}(t) \exp(i\omega_0 t) = A_{1,2}(t) e^{i\varphi_{1,2}(t)} \exp(i\omega_0 t)$ с одинаковой несущей частотой ω_0 . Тогда суммарное поле в этой точке равно $\mathcal{E}(t) = \mathcal{E}_1(t) + \mathcal{E}_2(t)$. Регистрируемая приемником интенсивность есть $I \sim |\mathcal{E}(t)|^2 = |\mathcal{E}_1(t) + \mathcal{E}_2(t)|^2$, поэтому интенсивность результирующего колебания можно представить в виде

$$I = I_1 + I_2 + 2\text{Re} \overline{\mathcal{E}_1 \mathcal{E}_2^*} = I_1 + I_2 + 2\text{Re} \overline{E_1 E_2^*}, \quad (7.7)$$

где $I_{1,2} = \overline{|E_{1,2}|^2}$ — интенсивности слагаемых волн.

Формула (7.7) представляет собой наиболее общее выражение, справедливое для любой двухлучевой интерференционной схемы. Третье (*интерференционное*) слагаемое отражает тот факт, что интенсивность результирующего колебания может отличаться от суммы интенсивностей, т.е. описывает явление интерференции.

Пусть волны $\mathcal{E}_{1,2}(t)$ вышли из *одного источника*, но в точку наблюдения пришли по *разным путям*, например, как в интерферометре Майкельсона. Тогда в точке наблюдения их можно представить в виде: $\mathcal{E}_1(t) = \alpha_1 \mathcal{E}(t) = \alpha_1 E(t) \exp(i\omega_0 t)$, $\mathcal{E}_2(t) = \alpha_2 \mathcal{E}(t + \tau) = \alpha_2 E(t + \tau) \exp(i\omega_0(t + \tau))$, так как одна волна *запаздывает* по отношению к другой на время τ , определяемое разностью хода. Коэффициенты пропорциональности $\alpha_{1,2}$ учитывают ослабление волн из-за сферической расходимости или поглощения.

С учетом этого для среднего значения $\overline{\mathcal{E}_1(t) \mathcal{E}_2^*(t)}$ можно записать

$$\overline{\mathcal{E}_1(t) \mathcal{E}_2(t)} = \alpha_1 \alpha_2 \overline{\mathcal{E}(t) \mathcal{E}^*(t + \tau)} = \alpha_1 \alpha_2 \Gamma(\tau). \quad (7.8)$$

Функция

$$\Gamma(\tau) = \overline{\mathcal{E}(t)\mathcal{E}^*(t+\tau)} = \overline{E(t)E^*(t+\tau)}e^{i\omega_0\tau} = \frac{1}{\Delta T} \int_{t-\Delta T/2}^{t+\Delta T/2} E(t)E^*(t+\tau)dt e^{i\omega_0\tau} = \overline{A(t)A^*(t+\tau)}e^{i\Delta\varphi}e^{i\omega_0\tau} \quad (7.9)$$

называется *функцией корреляции* или *функцией временной когерентности* квази-монохроматической волны $\mathcal{E}(t)$ со случайной амплитудой $A(t)$ и фазой $\varphi(t)$). Здесь черта сверху означает усреднение по времени на интервале ΔT , $\Delta\varphi = \varphi(t) - \varphi(t+\tau)$.

При $\tau = 0$ функция $\Gamma(\tau)$ равна *интенсивности света*

$$\Gamma(0) = I = \frac{1}{\Delta T} \int_{t-\Delta T/2}^{t+\Delta T/2} E(t)E^*(t)dt. \quad (7.10)$$

Заметим, что $\Gamma(0)$ равна сумме действительных положительных чисел $E(t_i)E^*(t_i) = A^2(t_i)\Delta t_i$, т. е. изображающие отдельные колебания вектора на векторной диаграмме выстроены коллинеарно вдоль действительной оси. Очевидно, что суммарный вектор $\Gamma(0)$ имеет при этом максимально возможную длину, т.е. $\Gamma(0) \geq \Gamma(\tau)$. Таким образом, *функция $\Gamma(\tau)$ при $\tau = 0$ имеет максимально возможное значение, равное интенсивности света.*

Типичный вид функции когерентности показан на **Рис. 7.4**. Характерное время спада функции когерентности называется *временем когерентности* τ_c . Пока $\tau < \tau_c$ аргумент комплексной функции $E(t)E^*(t+\tau) = A(t)A^*(t+\tau)e^{i[\varphi(t)-\varphi(t+\tau)]}$, равный $\Delta\varphi = \varphi(t) - \varphi(t+\tau)$, близок к нулю при любом t на интервале усреднения ΔT , так как $\varphi(t+\tau)$ "помнит" значение фазы $\varphi(t)$ в предшествующий момент времени и повторяет изменения почти в точности. В этом случае суммарное колебание $\mathcal{E}(t)$ состоит почти из той же совокупности цугов, что и $\mathcal{E}(t+\tau)$. Отсюда следует, что при $\tau \ll \tau_c$ разность $\Delta\varphi \simeq 0$, $A(t) \simeq A(t+\tau)$, поэтому $E(t)E^*(t+\tau) \simeq A^2(t) = I_0$ и $\Gamma(\tau) \simeq I_0 e^{-i\omega\tau}$, т.е. $\Gamma(\tau)$ имеет *модуль, мало отличающийся от интенсивности и аргумент, близкий к $\omega\tau$.*

При $\tau \geq \tau_c$ значения фаз $\varphi(t)$ и $\varphi(t+\tau)$ не коррелированы, т. е. при заданном $\varphi(t)$ фаза $\varphi(t+\tau)$ может с равной вероятностью принять любое значение от нуля до 2π , следовательно и разность фаз $\Delta\varphi(t) = \varphi(t) - \varphi(t+\tau)$ может быть любой, принимая с равной вероятностью (за время наблюдения $\Delta T \gg \tau_c$) все возможные значения. Величина $\overline{\exp(i\Delta\varphi)}$ оказывается близкой к нулю. Таким образом заключаем, что $\Gamma(\tau) = \overline{E(t)E^*(t+\tau)}e^{-i\omega\tau} \simeq 0$, при $\tau \geq \tau_c$.

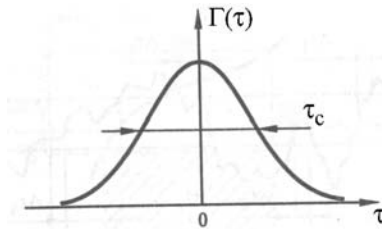


Рис. 7.4. Типичный вид функции когерентности.

Отметим, что вообще говоря величина $\Gamma(\tau)$ зависит от времени усреднения ΔT детектора. Если $\Delta T > \tau_c$, то $\Gamma(\tau)$ будет близка к нулю и интерференция не наблюдается. Поэтому можно сказать, что некогерентность есть результат несовершенства измерительных приборов. Для наблюдения интерференции, необходимо, чтобы $\Delta T \leq \tau_c$.

Итак, функция $\Gamma(\tau)$ является количественной мерой **коррелированности** (согласованности) или **мерой когерентности** значений квазимонохроматической волны $\mathcal{E}(t)$ в два момента времени разделенных промежутком τ в одной точке пространства: при $\tau < \tau_c$ $\Gamma(\tau) \neq 0$ (имеет место корреляция или когерентность), при $\tau \geq \tau_c$ ($\Gamma(\tau) \simeq 0$) — корреляции нет.

Таким образом, интерференционный эффект определяется функцией временной когерентности $\Gamma(\tau)$ колебательного процесса $\mathcal{E}$ и имеет место, когда $\Gamma(\tau) \neq 0$. Для некогерентных волн $\Gamma(\tau) = 0$, и интерференция отсутствует, т.е. *некогерентные волны не могут интерферировать*. При наложении некогерентных волн имеет место закон сложения интенсивностей. В частности, если волны $\mathcal{E}_1$ и $\mathcal{E}_2$ испущены независимыми источниками, то они некогерентны и, таким образом, не могут давать интерференционных эффектов.

Интерференция может наблюдаться лишь в том случае, если волны, по крайней мере, частично когерентны, т.е. $\Gamma(\tau) \neq 0$. Поэтому в интерференционных схемах всегда используют один источник света, а две (или несколько) интерферирующих волн получают путем деления (расщепления) одной волны, излучаемой источником.

Перепишем (7.7) в виде

$$I = I_1 + I_2 + 2\alpha_1\alpha_2\text{Re } \Gamma(\tau) = I_1 + I_2 + 2\sqrt{I_1 I_2}\text{Re } \gamma(\tau), \quad (7.11)$$

где $\gamma(\tau) = \Gamma(\tau)/\sqrt{I_1 I_2}$ — нормированная комплексная степень когерентности, которая изменяется от 1 при $\tau = 0$ до 0 при $\tau \geq \tau_c$. Здесь использовалось, что $I_1 = \alpha_1 I_0$ и $I_2 = \alpha_2 I_0$, где I_0 — интенсивность света вблизи источника.

При равных интенсивностях $I_1 = I_2 = I_0$ обеих волн формула (7.11) принимает вид

$$I = 2I_0[1 + \text{Re } \gamma(\tau)]. \quad (7.12)$$

Запишем $\gamma(\tau)$ в виде $\gamma(\tau) = |\gamma(\tau)|e^{i[\omega_0\tau + \theta(\tau)]}$, где $\theta(\tau) = \overline{\varphi(t) - \varphi(t + \tau)}$ — аргумент комплексной функции $\overline{E(t)E^*(t + \tau)}$.

С учетом этого получаем

$$I = 2I_0[1 + \text{Re } (|\gamma(\tau)|e^{i[\omega_0\tau + \theta(\tau)]})]. \quad (7.13)$$

Отсюда следует, что $I_{max} = 2I_0[1 + |\gamma(\tau)|]$, $I_{min} = 2I_0[1 - |\gamma(\tau)|]$ и для **видности** имеем

$$V = |\gamma(\tau)|, \quad (7.14)$$

т.е. *модуль нормированной степени когерентности есть видность интерференционной картины*, которая измеряется экспериментально. Из (7.13) следуют общие важные выводы:

1. Пусть временное запаздывание одной волны относительно другой превышает время когерентности ($\tau \geq \tau_c$), соответственно разность хода волн $\Delta \geq c\tau_c$. Тогда $\gamma(\tau) \simeq 0$ и $I = I_1 + I_2$. Таким образом, в этом случае *волны перестают быть когерентными и не интерферируют* (хотя эти волны и созданы одним источником света).

2. Пусть $\tau < \tau_c$. Тогда $\gamma(\tau) \neq 0$ и $I \neq I_1 + I_2$, т.е. волны *частично когерентны* и наблюдается интерференция.

3. Если $\tau \ll \tau_c$, то $\gamma(\tau) \simeq 1$, т.е. волны *когерентны*, и вместо (7.13) получаем

$$I = 2I_0(1 + \cos \omega_0 \tau), \quad (7.15)$$

где $\tau = \Delta/c$ — временное запаздывание, Δ — разность хода. Аналогичное выражение имеет место для строго гармонических волн. Мы получили, таким образом, ответ на вопрос о применимости синусоидальной идеализации: *при сложении когерентных волн все происходит так, как если бы волны были строго гармоническими.*

Таким образом, общий характер зависимости интенсивности от $\tau = \Delta/c$ в интерференционной картине, т.е. при изменении положения точки наблюдения, таков: *с ростом τ постепенно уменьшается видность (контраст) полос.* При $\tau \geq \tau_c$ полосы вовсе исчезают и мы приходим к закону сложения интенсивностей.

Анализ показывает, что *максимальная разность хода, при которой еще видна интерференция*, определяется как $\Delta_{max} \simeq c\tau_c$, а максимальный порядок интерференции — $m_{max} \simeq \Delta_{max}/\lambda = c\tau_c = \tau_c/T_0$. Величину $l_c = c\tau_c$ называют *длиной когерентности* или *длиной цуга* (это пространственная протяженность цуга излучения длительности τ_c). Таким образом, максимально допустимая разность хода ограничена длиной цуга, максимальный порядок интерференции — числом периодов колебаний, содержащихся в цуге излучения. Изучая распределение интенсивности в интерференционной картине, можно определить функцию когерентности квазимонохроматической волны, излучаемой источником света.

Пространственная когерентность и площадь когерентности. Понятие "точечный источник" — это такая же идеализация как и монохроматическая волна. Все реальные источники являются *протяженными*, а это приводит к уменьшению контраста интерференционной картины. Опыт показывает, что *при увеличении размеров источника видность интерференционной картины постепенно уменьшается и при достаточно больших размерах интерференционные полосы полностью исчезают.* Качественно характер изменения видности полос при увеличении размеров источника можно понять, если предположить, что протяженный источник состоит из *независимых* излучателей (**Рис.7.5**). Действительно, весь источник света можно "разбить" на малые площадки ΔS_i с линейным размером много меньше длины волны, состоящие из атомов, излучающих независимо от атомов другой площадки ΔS_j ($i \neq j$), поэтому волны, излучаемые этими площадками, являются некогерентными и интерференции между ними не будет.

При описании интерференции света, испускаемого протяженным источником, важным является понятие **пространственной когерентности**. Рассмотрим это понятие на примере интерференционного опыта Юнга, в котором используется квазимонохроматический свет от *протяженного* источника S (**Рис.7.6**). Пусть источник S ,

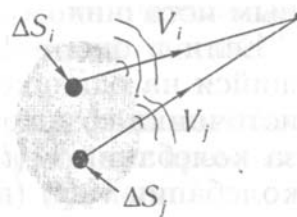


Рис. 7.5. Интерференция от протяженного источника

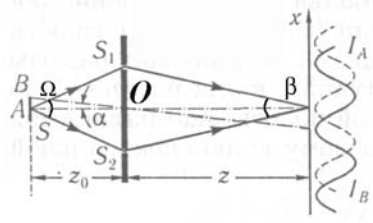


Рис. 7.6. К определению пространственной когерентности в опыте Юнга. I_B , I_A — интерференционные картины, создаваемые светом, пришедшим из точки B и A источника, соответственно; $S_1 S_2 = d$.

плоскость с отверстиями S_1 и S_2 и плоскость наблюдения удалены друг от друга настолько, что волны, приходящие от каждого малого элемента источника к отверстиям S_1 и S_2 , а также волны, приходящие к плоскости наблюдения от этих отверстий, можно считать *плоскими*. Тогда, очевидно, что интерференционная картина, создаваемая крайней точкой B источника (сплошная кривая I_B на **Рис.7.6**) смещена относительно интерференционной картины от центральной точки A (пунктирная кривая I_A на расстояние $L = \alpha z$. Если это смещение меньше ширины интерференционной полосы $l \simeq d/\beta$, где $\beta = d/z$ — угол схождения интерферирующих волн, d — расстояние между отверстиями, то интерференционная картина наблюдается. В противном случае (при $\alpha > \lambda/d$) интерференционная картина не наблюдается. Это связано с тем что на темную полосу (минимум интенсивности от одной точки источника) накладывается светлая полоса (максимум интенсивности от другой точки), и в результате получается равномерная освещенность плоскости наблюдения. Таким образом, для максимально допустимой ширины источника b , при которой еще интерференция наблюдается, можно записать следующую оценку:

$$b \leq \frac{z_0}{d} \lambda = \lambda/\Omega, \quad (7.16)$$

или, при заданной ширине источника, максимально допустимое расстояние между отверстиями d , которое и определяет апертуру интерференции $\Omega \simeq d/z_0$.

Возникновение полос в *опыте Юнга* есть проявление *пространственной когерентности* между двумя пучками, приходящими в точку наблюдения от двух отверстий $S_{1,2}$, поскольку контраст между полосами зависит от расстояния между этими отверстиями. Для наблюдения интерференции необходимо, чтобы отверстия $S_{1,2}$ располагались в пределах области с центром в точке O и площадью $\Delta S \sim (z_0 \cdot \Omega/2)^2 \sim z_0^2 \lambda^2 / \Delta \sigma$, где $\Delta \sigma \sim b^2$ — площадь источника. ΔS называют *площадью когерентности* света вблизи точки O , а квадратный корень от площади когерентности называют *поперечной длиной когерентности*. Площадь когерентности будет тем больше, чем больше z_0 . Введем телесный угол, независимый от расстояния z_0 ,

$$\Delta \Omega = \Delta S / z_0^2 = \lambda^2 / \Delta \sigma, \quad (7.17)$$

под которым площадь когерентности видна из центра источника. Удобно выразить площадь когерентности через другой телесный угол $\Delta \Omega' = \Delta \sigma / z_0^2$, под которым источник света виден из точки O . Поскольку $\Delta \sigma = z_0^2 \cdot \Delta \Omega'$, для площади когерентности получаем $\Delta S \sim \lambda^2 / \Delta \Omega'$.

Для иллюстрации пространственной когерентности рассмотрим несколько примеров. Предположим, что линейный размер теплового источника $b = 1$ мм и что источник испускает квазимонохроматический свет со средней длиной волны

$\lambda = 5000 \text{ \AA}$. Пусть плоскость, в которой расположены отверстия, находится на расстоянии $z_0 = 2 \text{ м}$ от источника. Площадь когерентности в этой плоскости равна $\Delta S = (1 \cdot 10^2 / 10^{-1})^2 \cdot (5 \cdot 10^{-5})^2 = 1 \text{ мм}^2$, т.е. ее линейные размеры будут порядка 1 мм. В качестве второго примера оценим площадь когерентности пучка солнечного света, освещающего поверхность Земли. Чтобы удовлетворить нашему предположению о квазимонохроматичности света, мы должны предварительно пропустить солнечный свет через фильтр с узкой полосой пропускания, скажем вблизи длины волны $\lambda \sim 5000 \text{ \AA}$. Угловой радиус Солнца составляет примерно $\alpha = 0^\circ 16' \sim 0.00465$ радиан. Следовательно, телесный угол $\Delta\Omega'$, под которым солнечный диск виден с поверхности Земли, равен $\Delta\Omega \approx \pi\alpha^2 = 6.6 \cdot 10^{-5}$ стер и площадь когерентности равна $\Delta A = 3.7 \cdot 10^{-3} \text{ мм}^2$. Таким образом, линейный размер площади когерентности на поверхности земли для фильтрованного солнечного света составляет величину порядка 0.061 мм.

Сравним площадь когерентности солнечного света на поверхности земли с площадью когерентности света от более удаленных звезд. Заметим, что площадь когерентности обратно пропорциональна величине телесного угла, под которым источник виден из центральной точки Q плоскости, в которой производится оценка. При наблюдении с поверхности земли угловой диаметр звезды обычно на много порядков меньше углового диаметра солнца. Следовательно, площадь когерентности света звезды на поверхности Земли должна намного превышать площадь когерентности солнечного света. Для примера рассмотрим Бетельгейзе (α -Ориона), которая фактически была первой звездой, чей угловой диаметр был установлен с помощью интерференционной техники и составил величину $2\alpha \sim 0.047$ секунд или $2.3 \cdot 10^{-7}$ радиан. Телесный угол, под которым эта звезда видна с поверхности земли, равен, соответственно, $\Delta\Omega' \approx 4.15 \cdot 10^{-14}$ стер. Таким образом, площадь когерентности света Бетельгейзе, пропущенного через узкополосный фильтр на длине волны $\lambda = 5000 \text{ \AA}$, на поверхности земли равна $\Delta A = 6 \text{ м}^2$.

Таким образом, имеются значительные корреляции между световыми колебаниями, достигающими поверхности земли от Бетельгейзе, в двух точках, максимальное расстояние между которыми составляет величину порядка 2.45 м. Существует множество звезд, чей угловой диаметр значительно меньше углового диаметра Бетельгейзе, так что высокая степень корреляции в свете от этих звезд имеет место на гораздо больших площадях. Таким образом, свет звезд, попадающий в апертуру телескопа, высоко коррелирован в пределах площадей, которые в общем случае значительно превышают площадь этой апертуры.

Для количественной оценки видности интерференционной картины используется функция *пространственная когерентность*, которая описывает свойства волн, излучаемых протяженными источниками света.

Если в опыте Юнга используется точечный источник света S , находящийся на одинаковом расстоянии от отверстий S_1 и S_2 , то волны от этого источника создают на отверстиях колебания с одинаковой амплитудой и фазой в любой момент времени, т.е. $\mathcal{E}_1(t) = \mathcal{E}_2(t) = \mathcal{E}(t)$, при этом $\overline{\mathcal{E}_1(t)\mathcal{E}_2^*(t)} = \overline{\mathcal{E}(t)\mathcal{E}^*(t)} = I$. Таким образом, колебания $\mathcal{E}_1(t)$ и $\mathcal{E}_2(t)$ полностью когерентные. В случае протяженного источника колебания в точках S_1 и S_2 создаются всеми точками источника, причем лишь центральная точка источника дает на отверстиях одинаковые по амплитуде и фазе колебания. Любая другая точка посылает волну, создающую на отверстиях разные по фазе колебания так как она находится на разных расстояниях от отверстий. Поэтому суммарные колебания в области отверстий отличаются по фазе (и по амплитуде),

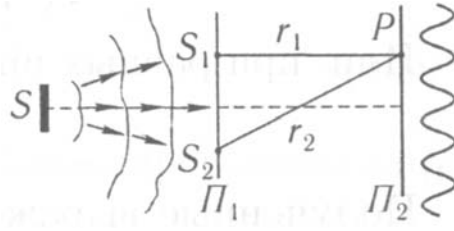


Рис. 7.7. К определению функции пространственной когерентности

что и приводит, как будет показано, к уменьшению контраста (видности) интерференционной картины, а значит и к уменьшению пространственной когерентности.

Пусть $\mathcal{E}_1(t)$ и $\mathcal{E}_2(t)$ — колебания, созданные протяженным квазимонохроматическим источником S в точках S_1 и S_2 (**Рис. 7.7**). Чтобы экспериментально определить, когерентны ли эти колебания установим в плоскости Π_1 непрозрачный экран, проделав два малых отверстия в точках S_1 и S_2 . Волны света, проходя через отверстия, идут в разных направлениях (из-за дифракции) и, перекрываясь, создают (или не создают) интерференционную картину в плоскости наблюдения Π_2 , которая, как будет показано, дает информацию о пространственной когерентности.

Колебание в точке наблюдения P есть результат наложения волн, пришедших в нее от S_1 и S_2 . Колебания, созданные волной, пришедшей от отверстия S , запишем в виде $\alpha_1 \mathcal{E}_1(t)$, α_1 учитывает ослабление волны из-за расходимости. Аналогично волна от отверстия S_2 создает колебание $\alpha_2 \mathcal{E}_2(t + \tau)$, где τ — время запаздывания по отношению к первой волне. Результирующее колебание есть $\mathcal{E}(t) = \alpha_1 \mathcal{E}_1(t) + \alpha_2 \mathcal{E}_2(t + \tau)$. Для интенсивности получаем:

$$I = |\mathcal{E}(t)|^2 = I_1 + I_2 + 2\text{Re} \overline{\mathcal{E}_1(t) \mathcal{E}_2^*(t + \tau)}. \quad (7.18)$$

Будем считать, что для всей интересующей нас области наблюдения запаздывание τ одной волны относительно другой мало по сравнению с временем когерентности τ_c (но на много больше периода светового колебания T_0 , $\tau \ll \tau_c$). Это означает, что за время $\tau = \Delta/c$ ($\Delta = r_2 - r_1$) амплитуда и фаза колебания $\mathcal{E}_2(t)$ не успевает заметно измениться, т. е. колебание в момент времени $t + \tau$ состоит из той же совокупности цугов, что и в момент времени t . Тогда

$$E_2(t + \tau) \simeq E_2(t), \quad \mathcal{E}_2(t + \tau) = E_2(t + \tau) e^{i\omega_0(t+\tau)} \simeq E_2(t) e^{i\omega_0(t+\tau)} \quad (7.19)$$

Учитывая (7.19), (7.18) можно переписать в виде

$$I = I_1 + I_2 + 2\text{Re} \overline{E_1(t) E_2^*(t)} e^{i\omega_0 \tau} = I_1 + I_2 + 2\text{Re} \Gamma_{12}(0) e^{i\omega_0 \tau}. \quad (7.20)$$

Функцию $\Gamma_{12}(0) = \overline{E_1(t) E_2^*(t)}$ называют *функцией пространственной когерентности*. Она описывает корреляцию колебаний в разных точках пространства в один и тот же момент времени.

Из (7.20) следует, что интенсивность света во всех максимумах одинакова, так как $\Gamma_{12}(0)$ не зависит от положения точки наблюдения, и определяется лишь апертурой интерференции (в опыте Юнга — угловым расстоянием между отверстиями S_1 и S_2) или, при заданном расстоянии, размерами источника. Таким образом, видность интерференционной картины одинакова во всех точках.

Вводя нормированную функцию $\gamma_{12}(0) = \Gamma_{12}(0)/\sqrt{I_1 I_2}$, называемую *степенью пространственной когерентности*, выражение для интенсивности запишем в виде:

$$I = I_1 + I_2 + \sqrt{I_1 I_2} \operatorname{Re} \gamma_{12}(0) e^{i\omega_0 \tau}. \quad (7.21)$$

При равных интенсивностях $I_1 = I_2 = I_0$

$$I = |\mathcal{E}(t)|^2 = 2I_0[1 + |\gamma_{12}(0)| \cos(\omega_0 \tau + \theta)]. \quad (7.22)$$

θ — аргумент функции $\gamma_{12}(0)$. Выражения (7.21) и (7.22), определяющие интерференционную картину при использовании протяженного источника, справедливы для любой двухлучевой интерференционной схемы. Видность V выражается через $\gamma_{12}(0)$ равенством S_1

$$V = |\gamma_{12}(0)|. \quad (7.23)$$

Степень пространственной когерентности связана с размерами источника света. В частном случае, когда источник представляет собой равномерно светящуюся полосу шириной b , $\gamma_{12}(0)$ имеет вид:

$$V = |\gamma_{12}| = \left| \frac{\sin\left(\frac{\pi d}{\lambda z_0/b}\right)}{\frac{\pi d}{\lambda z_0/b}} \right|, \quad (7.24)$$

где z_0 — расстояние от источника до экрана с двумя отверстиями, расположенными на расстоянии d друг от друга.

Зависимость этой функции от расстояния d показана на **Рис.7.8**. Видно, что степень когерентности колебаний в двух точках, разнесенных на расстояние d уменьшается от 1 при $d = 0$ до нуля при $d = \rho_0 = 2\pi z/kb = \lambda z/b$. Если расстояние между точками S_1 и S_2 $d < \rho_0$, то колебания в этих точках называются частично когерентными: при этом $\gamma_{12}(0) \neq 0$. При $d > \rho_0$ колебания в точках S_1 и S_2 считаются некогерентными и $\gamma_{12}(0) = 0$.

Величину $\rho_0 = \lambda z/b$ называют *радиусом пространственной когерентности*, а площадку $\Delta S = \pi \rho_0^2$ — *площадью когерентности*. Внутри площади когерентности колебания в любых точках считаются частично когерентными, вне — некогерентными.

Вводя угловой размер источника $\varphi = b/z_0$, можно записать

$$\rho_0 = \lambda/\varphi \quad (7.25)$$

Таким образом, если расстояние между отверстиями S_1 и S_2 меньше радиуса когерентности, то интерференция наблюдается или, что аналогично, апертура интерференции должна быть меньше углового размера площади когерентности.

Если зафиксировать расстояние между точками S_1 и S_2 и приближать плоскость, в которой они находятся, к источнику, то степень когерентности колебаний в этих точках уменьшается. При стремлении z_0 к нулю колебания в любых двух точках,

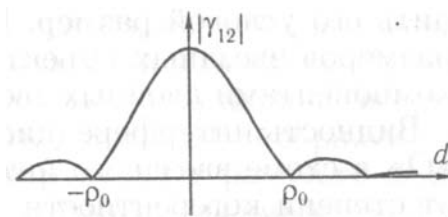


Рис. 7.8. Нормированная функция пространственной когерентности

как угодно близко расположенных, некогерентны, так как любые две точки источника излучают некогерентно. Наоборот, с удалением этой плоскости колебания в точках S_1 и S_2 становятся все более когерентными и при достаточном удалении когерентность колебаний становится полной. *Размер площади когерентности по мере удаления от источника постепенно (пропорционально z) увеличивается.* Таким образом, первоначально некогерентное поле излучения в процессе распространения приобретает частичную когерентность.

8. ЛЕКЦИЯ №8

Дифракция как проявление волновой природы света: принцип Гюйгенса-Френеля; дифракционный интеграл Френеля; зоны Френеля. Дифракция на круглом отверстии: ближняя и дальняя зоны дифракции. Дифракция на краю экрана, дифракция на диске и пятно Пуассона.

8.1. Дифракция как проявление волновой природы света. Понятие "дифракция" в оптике связывается с *нарушением прямолинейности распространения света*. В широком смысле слова *дифракцию* определяют как *любое отклонение распространения света от прямолинейного, не связанное с отражением или преломлением*. В более узком смысле *дифракцией* называют явление *огибания волной препятствия*. Такие явления хорошо известны для длинных волн, например звуковых волн или волн на поверхности воды. В оптике этому соответствует проникновение света в область геометрической тени. В теории волн под *дифракцией* понимают всю совокупность явлений в волновом поле, возникающих при наличии препятствий распространению волн. Наконец, используя понятие интерференции света, можно сказать, что *дифракция — это интерференция в ограниченных световых пучках*. Принципиальное значение дифракции состоит в том, что она, как и интерференция, доказывает волновую природу света. Фундаментальный смысл дифракции состоит в том, что она *ограничивает* возможности концентрации света в пространстве, кладет предел разрешающей способности оптических и спектральных приборов, влияет на формирование оптического изображения и т. п.

Первое сообщение о наблюдении дифракции света было сделано Гримальди. Он установил, что *переход от света к тени происходит постепенно*, а не резко. Этот результат не мог найти удовлетворительного объяснения в рамках корпускулярной теории света, которой в то время придерживались и согласно которой свет должен распространяться прямолинейно, а изображение отверстия в плоскости наблюдения должно иметь резкую границу.

Принцип Гюйгенса-Френеля. Понимание природы дифракционных явлений связано с развитием представлений о свете как о волне. Первый шаг на этом пути сделал в конце XVII в. (1678) голландский ученый Христиан Гюйгенс. Основываясь на догадке о том, что свет это волна, он выдвинул идею, раскрывающую механизм распространения света. Гюйгенс полагал, что свет распространяется от источника подобно волне на поверхности воды. На фронте светового возмущения каждая

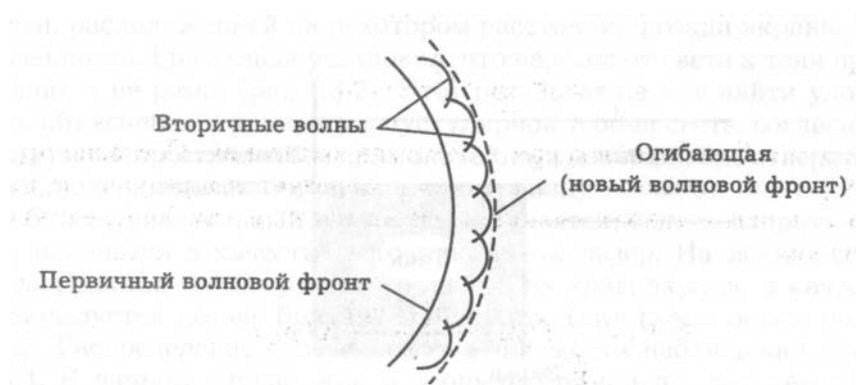


Рис. 8.1. Построение огибающей световой волны по Гюйгенсу

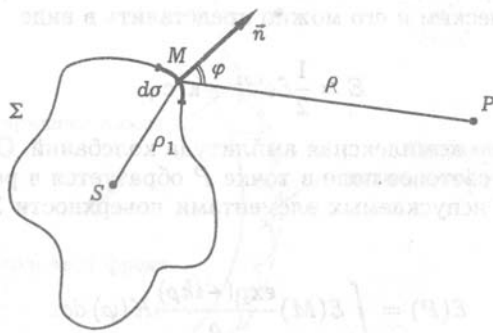


Рис. 8.2. К выводу Интеграла Гюйгенса-Френеля: S — источник света, P — точка наблюдения, $SM = \rho_1$, $SP = \rho$, $\vec{n}$ — нормаль к поверхности Σ в точке M .

точка есть источник вторичной сферической волны. Положение волнового фронта в следующий момент времени определяется огибающей вторичных волн.

Принцип Гюйгенса иллюстрирует **Рис. 8.1**, на котором показаны волновой фронт светового возмущения, элементарные вторичные волны, огибающая вторичных волн. Пользуясь этим принципом, можно объяснить такие явления, как распространение света от точечного источника, распространение светового пучка, отражение и преломление света. О.Френель ((1818)) дополнил принцип Гюйгенса представлением о том, что *вторичные световые волны могут как усиливать, так и ослаблять друг друга*. Другими словами, они могут *интерферировать*. Световое поле есть *результат интерференции элементарных вторичных волн, испускаемых каждым элементом некоторой волновой поверхности*. Это утверждение называют *принципом Гюйгенса-Френеля*. Основываясь на этом принципе, Френель смог с большой точностью рассчитать распределение света в дифракционных картинах.

Дифракционный интеграл Френеля. Принцип Гюйгенса-Френеля позволяет построить элементарную теорию дифракции света. Пусть имеется точечный источник света S . Требуется найти световое поле в некоторой точке P , если между точками S и P расположено препятствие распространению света, например экран с отверстием или непрозрачный диск.

Сначала рассмотрим математическую формулировку принципа Гюйгенса-Френеля. Введем некоторую замкнутую поверхность Σ , охватывающую источник света, и будем считать каждый элемент $d\sigma$ этой поверхности источником вторичной сферической световой волны (**Рис. 8.2**). Рассмотрим некоторую точку M на поверхности $d\sigma$. Считая источник света S точечным, обозначим расстояние от S до M через ρ_1 , а расстояние от M до точки наблюдения P через ρ . Введем также угол φ между нормалью $\vec{n}$ к поверхности Σ в точке M и направлением на точку наблюдения MP . Для простоты будем считать, что источник света испускает монохроматическую волну.

Принцип Гюйгенса-Френеля утверждает, что световое поле в точке P есть результат *наложения (сложения) световых волн, испускаемых всеми элементами поверхности Σ* . Волну, испускаемую элементом поверхности $d\sigma$, можно считать сферической. Поэтому можно записать

$$E(P) = \int_{\Sigma} E(M) \frac{\exp(-ik\rho)}{\rho} K(\varphi) d\sigma. \quad (8.1)$$

Здесь $E(P)$ и $E(M)$ — комплексные амплитуды поля в точках P и M , ω и $k = \omega/c = 2\pi/\lambda$ — частота и волновое число световой волны, $K(\varphi)$ — "коэффициент наклона", монотонно убывающий от некоторого начального значения $K(0)$ до нуля при изменении угла φ от нуля до $\pi/2$. Он учитывает то обстоятельство, что вклад элемента $d\sigma$ в результирующее поле зависит от ориентации данного элемента поверхности по отношению к направлению на точку наблюдения.

Интеграл (8.1) называют *интегралом Гюйгенса-Френеля*. Формула (8.1) получена на основе *качественных* физических соображений. Множитель $\exp(-ik\rho)/\rho$ в подынтегральном выражении описывает распространение элементарной вторичной сферической световой волны. Наиболее существенно то, что интеграл Гюйгенса-Френеля учитывает фазы элементарных вторичных волн, приходящих в точку P от различных элементов поверхности Σ , т.е. принимается во внимание *интерференция вторичных волн*.

Зоны Френеля. Френель предложил хотя и *приближенный*, но изящный способ расчета дифракционных картин, основанный на представлении о так называемых *полуволновых зонах* или *зонах Френеля*, которые вводятся следующим образом. Выберем поверхность Σ в виде сферы с центром в точке S . Согласно принципу Гюйгенса-Френеля, данную поверхность можно рассматривать как источник вторичных световых волн. Выделим на сфере кольцевые зоны так, чтобы расстояния от границ зоны до точки наблюдения отличались на *половину длины световой волны*. Обозначив границы зон буквами $M_0, M_1, M_2, \dots$, получим

$$\begin{aligned} M_0P &= OP + \lambda/2 \\ M_1P &= M_0P + \lambda/2 \\ &\dots\dots\dots \\ M_nP &= M_{n-1}P + \lambda/2 \end{aligned} \quad (8.2)$$

где λ — длина световой волны, P — точка наблюдения поля, O — центр нулевой зоны Френеля (**Рис.8.3**).

Формулы (8.2) определяют положение границ зон Френеля. Смысл разбиения поверхности Σ на зоны состоит в том, что *разность фаз элементарных вторичных волн, приходящих в точку наблюдения от данной зоны, не превышает π* . Сложение таких волн приводит к их взаимному усилению. Поэтому каждую зону Френеля

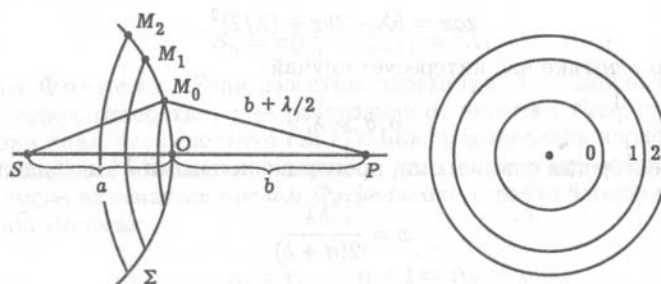


Рис. 8.3. Зоны Френеля. S — источник света, P — точка наблюдения. Цифры 0, 1, 2 на правом рисунке соответствуют нулевой, первой и второй зонам Френеля.

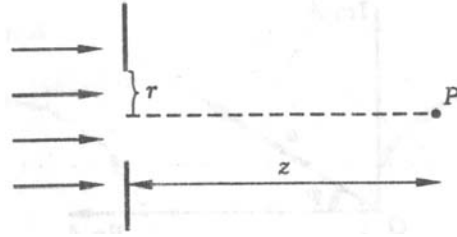


Рис. 8.4. Схема дифракции на круглом отверстии. Плоская монохроматическая волна падает слева на круглое отверстие радиуса r . Штриховая линия — ось пучка. Точка наблюдения P находится на расстоянии z от отверстия.

можно рассматривать как источник вторичных волн, имеющих определенную фазу. Причем две соседние зоны Френеля действуют как источники, колеблющиеся в противофазе, а, значит, вторичные волны приходят в точку наблюдения также в противофазе. Подчеркнем, что положение границ френелевских зон зависит от выбора точки наблюдения.

Для того, чтобы оценить относительный вклад френелевских зон в интеграл Гюйгенса-Френеля, оценим радиусы зон и их площади (**Рис.8.3**). Пусть a — радиус сферы Σ , b — кратчайшее расстояние от точки P до сферы, r_0 — радиус нулевой зоны Френеля. Из **Рис.8.3** видно, что

$$r_0^2 = a^2 - (a - x)^2 = 2ax - \lambda^2, \quad (8.3)$$

С другой стороны,

$$r_0^2 = (b + \lambda/2)^2 - (b + x)^2. \quad (8.4)$$

Из (8.3), (8.4) следует, что

$$2ax = b\lambda - 2bx + (\lambda/2)^2. \quad (8.5)$$

Обычно в оптике $a, b \gg \lambda, x$. Поэтому для x можно записать

$$x = \frac{b\lambda}{2(a + b)} \quad (8.6)$$

и при $r_0^2 \approx 2ax$ получаем

$$r_0 = \sqrt{\frac{\lambda ab}{a + b}}. \quad (8.7)$$

Формула (8.7) дает радиус нулевой зоны Френеля. Аналогичным образом для n -й зоны Френеля можно получить:

$$r_n = \sqrt{(n + 1)\lambda \frac{ab}{a + b}}. \quad (8.8)$$

8.2. Дифракция на круглом отверстии. Для примера рассмотрим случай дифракции монохроматической плоской волны на круглом отверстии (**Рис.8.4**). Выберем некоторую точку P на оси пучка и определим, как меняется интенсивность света в данной точке при изменении радиуса отверстия. Пусть z расстояние от точки наблюдения до экрана с отверстием, а r — радиус отверстия. В качестве поверхности Σ — источника вторичных волн — введем круг радиуса r , лежащий в плоскости экрана и совпадающий с отверстием. Разобьем поверхность Σ на зоны Френеля. В данном случае они представляют собой кольца на плоскости. Их радиусы можно подсчитать по формуле (8.7), полагая $a \rightarrow \infty, b = z$. Получим

$$r_n = \sqrt{(n+1)\lambda z}. \quad (8.9)$$

Из формулы (8.9) следует, что *зоны Френеля имеют одинаковые площади*, определяемые формулой

$$S_n = \pi(r_n^2 - r_{n-1}^2) = \pi\lambda z. \quad (8.10)$$

Отметим, что полученный результат, как можно показать, справедлив для зон Френеля любой формы.

Если известны длина световой волны, радиус отверстия и расстояние от экрана с отверстием до точки наблюдения поля, то, используя (8.9), можно вычислить число зон Френеля N_F , попадающих в пределы отверстия, или *число открытых френелевских зон*, видимых из точки P . Это число называется *числом Френеля*, оно играет важную роль в теории дифракции. Полагая $r_n = r, n+1 = N_F$, получим из (8.9)

$$N_F = \frac{r^2}{\lambda z}. \quad (8.11)$$

Введение зон Френеля позволяет графически анализировать дифракционные явления. Вычисление результирующего светового поля, описываемого интегралом Гюйгенса-Френеля (8.1), по сути дела сводится к суммированию световых колебаний, возбуждаемых элементарными вторичными источниками. С математической точки зрения задача сводится к суммированию гармонических колебаний, имеющих одну и ту же частоту, но разные амплитуды и фазы. Это можно сделать графическим способом с помощью построения *векторной диаграммы*.

Как известно, гармоническое колебание с амплитудой a и фазой φ можно охарактеризовать комплексной амплитудой $A = a \exp(i\varphi)$, которая изображается вектором на плоскости переменных $\text{Re } A, \text{Im } A$, причем длина вектора равна a , а угол наклона к оси $\text{Re } A$ равен $\varphi = \arctan(\text{Im } A / \text{Re } A)$. Сумма нескольких гармонических колебаний с произвольными амплитудами и фазами есть также гармоническое колебание

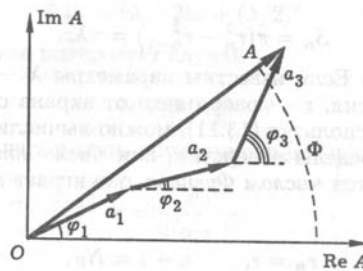


Рис. 8.5. Сложение гармонических колебаний методом векторной диаграммы

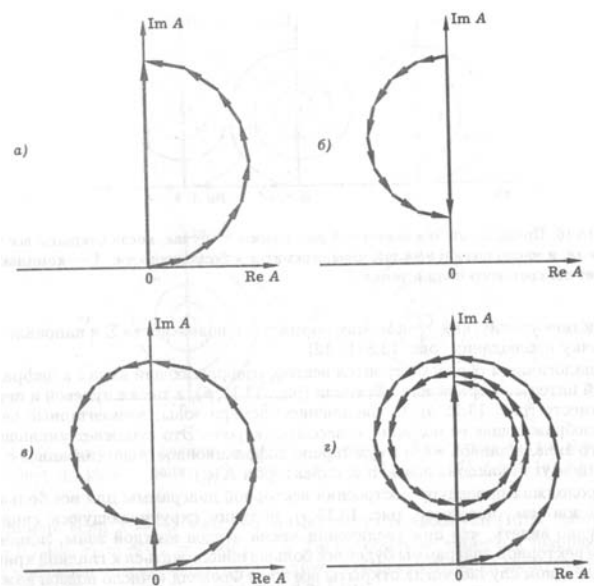


Рис. 8.6. Векторные диаграммы для расчета амплитуды дифракционного светового поля: а) открыта нулевая зона Френеля (все остальные закрыты); б) открыта первая зона Френеля; в) открыты нулевая и первая зоны Френеля; г) открыты нулевая, первая, вторая и третья зоны Френеля. Короткие векторы обозначают комплексные амплитуды полей, приходящих в точку наблюдения от отдельных подзон Френеля. Длинные вектора изображает результирующее световое поля в точке наблюдения.

на той же самой частоте. Амплитуду A и фазу Φ результирующего колебания можно найти, складывая по правилу сложения векторов векторы, изображающие колебания-слагаемые. Каждый такой вектор имеет длину, равную амплитуде колебания и угол наклона к оси абсцисс, равный фазе данного колебания. После построения векторной суммы, амплитуда результирующего колебания находится как длина полученного вектора-суммы, а фаза результирующего колебания — как угол наклона этого вектора к оси абсцисс, как показано на **Рис.8.5**.

Применим этот метод для расчета дифракционного интеграла Гюйгенса-Френеля. Сначала вычислим вклад в дифракционный интеграл какой-нибудь одной, например нулевой зоны Френеля. Соответствующее построение показано на **Рис.8.6** а. Оно выполняется следующим образом. Разбиваем зону Френеля на множество concentric колец (подзон). Очевидно, разбиение можно произвести таким образом, чтобы площади подзон были примерно одинаковы, а число подзон было достаточно большим. В этом случае вклады подзон изображаются векторами, которые имеют одинаковую длину, но разные углы наклона к оси абсцисс. Первый и последний векторы повернуты друг относительно друга на угол π — в соответствии с определением зоны Френеля. По мере увеличения радиуса вклада подзоны (и, следовательно, длина соответствующего вектора) немного уменьшается в следствие увеличения угла между нормалью к поверхности Σ и направлением на точку наблюдения. Аналогичным образом строится вектор, изображающий вклад в дифракционный интеграл первой

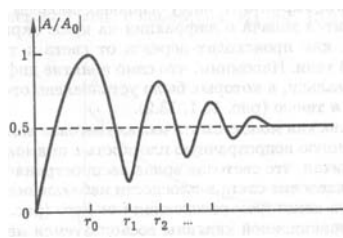


Рис. 8.7. Зависимость амплитуды света в точке наблюдения от радиуса отверстия. $r_{0,1,2}$, — радиусы френелевских зон.

зоны Френеля (**Рис.8.6 б**). С увеличением номера зоны, элементарные векторы, изображающие ее подзоны, становятся короче. Это отражает уменьшение общего вклада данной зоны в суммарное дифракционное поле, связанное с увеличением угла наклона зоны, т. е. с фактором $K(\varphi)$. Продолжая процедуру построения векторной диаграммы для все большего числа зон, получим скручивающуюся *спираль*. (**Рис.8.6 г**) Отметим, что при увеличении числа подзон каждой зоны, ломанная линия векторной диаграммы будет все больше приближаться к гладкой кривой. В предельном случае, когда открыты все зоны Френеля и число подзон каждой зоны стремится к бесконечности, получим векторную диаграмму, которая имеет вид гладкой скручивающейся спирали, которую и называют *спиралью Френеля*.

Таким образом, результирующее поле E_P в точке наблюдения можно представить в виде суммы $A_P = A_0 - A_1 + A_2 - A_3 + \dots + A_n$, где A_n — амплитуда поля, создаваемое соответствующей зоной Френеля ($n = 0, 1, 2, \dots$ — номер зоны Френеля, причем $A_0 > A_1 > A_2 > A_3$ и т.д. Отсюда следует, что амплитуда света в точке наблюдения немонотонно (осциллирующее) зависит от радиуса отверстия, как показано на **Рис.8.7**. Причем максимальная интенсивность света в точке наблюдения достигается на расстоянии, при котором отверстие совпадает с нулевой зоной Френеля. В этом случае амплитуда поля в два раза (а интенсивность света в 4 раза) выше, чем в отсутствие экрана.

Из определения френелевских зон следует, что по отношению к данной точке наблюдения поля две соседние зоны действуют как источники света, колеблющиеся в противофазе — посылаемые ими световые волны в значительной степени гасят друг друга за счет деструктивной интерференции. Все *четные* зоны Френеля дают вклад в результирующее поле одного знака. Все *нечетные* — противоположного знака. Поэтому если каким-либо образом "выключать" все четные либо все нечетные зоны, можно получить значительное увеличение интенсивности света в точке наблюдения. Такая пластинка называется *зонной пластинкой Френеля*. Можно сказать, что зонная пластинка *перераспределяет* световое поле в пространстве или *фокусирует* свет аналогично как это делает линза. Из формулы (8.9) следует, что фокусное расстояние F связано с радиусами френелевских зон r_n следующим соотношением

$$r_n = \sqrt{(n+1)\lambda F}. \quad (8.12)$$

Ближняя и дальняя зоны дифракции. Рассмотрим теперь, как меняется интенсивность света I на оси отверстия по мере увеличения расстояния z от экрана с отверстием. Зафиксируем радиус отверстия r . По мере удаления от экрана число зон Френеля на отверстии уменьшается, пока, наконец, в пределах отверстия не останется одна нулевая зона Френеля. В этот момент интенсивность света I в точке



Рис. 8.8. Зависимость интенсивности света на оси отверстия от расстояния до экрана. z_d — дифракционная длина светового пучка

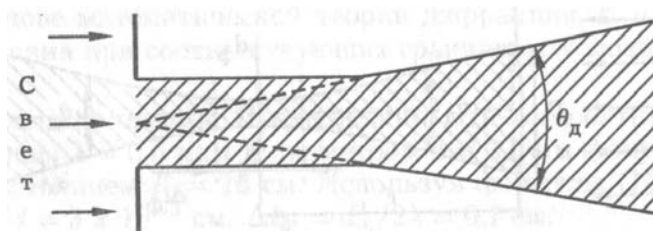


Рис. 8.9. Дифракция светового пучка и угол дифракционной расходимости θ_d

наблюдения достигает максимума (**Рис.8.8**), после чего монотонно убывает с ростом расстояния z . Расстояние z_d , при котором отверстие совпадает с нулевой зоной Френеля, называют *дифракционной длиной* светового пучка. Она определяет границу между двумя различными зонами. Зона, для которой $z \ll z_d$, называется *ближней зоной* дифракции. В этой зоне световой пучок сохраняет структуру, заданную формой отверстия, а интенсивность света на оси пучка примерно равна интенсивности исходной световой волны. Для точек ближней зоны в пределах отверстия помещается множество зон Френеля, и поперечный профиль пучка поддерживается постоянным за счет интерференции элементарных вторичных волн, идущих от разных зон Френеля и его можно считать параллельным. Зона, для которой $z \gg z_d$ называется *дальней зоной* дифракции. В этой зоне интенсивность света на оси пучка много меньше интенсивности исходной волны, следовательно, световой пучок *расширяется*. Для точек дальней зоны в пределах отверстия помещается только центральная часть нулевой зоны Френеля. Интерференция элементарных вторичных волн выражена слабее. Она уже не в состоянии поддерживать исходный поперечный профиль пучка, пучок становится расходящимся. Характер изменения поперечного размера светового пучка в процессе дифракции показан на **Рис.8.9**. Используя определение дифракционной длины z_d и формулу (8.9), получаем, что $z_d = r^2/\lambda$ (r — радиус пучка). Дифракционная длина связана с числом Френеля $N_F = r^2/\lambda z$ соотношением $N_F = z_d/z$. Откуда следует, что в ближней зоне $N_F \gg 1$, а в дальней — $N_F \ll 1$.

Оценим дифракционную расходимость пучка θ_d (**Рис.8.9**), исходя из представлений об интерференции элементарных вторичных волн. Полагая, что положение границы светового пучка определяется деструктивной интерференцией лучей, приходящих от противоположных границ отверстия, т.е. условием $\Delta = \lambda/2$, где Δ — разность хода (**Рис.8.10**). Отсюда видно, что $\Delta \approx d \sin \theta_d/2$, где d — диаметр отверстия. Как правило, дифракционная расходимость невелика ($\theta_d \ll 1$, поэтому можно

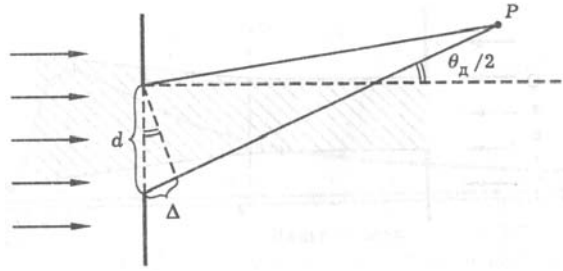


Рис. 8.10. К расчету дифракционной расходимости светового пучка

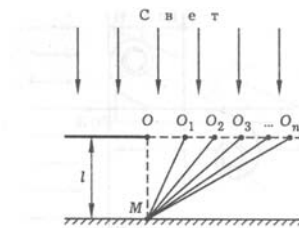


Рис. 8.11. Построение зон Френеля при дифракции на краю экрана.

приблизенно записать $\Delta \approx d\theta_d/2$, из которого следует, что

$$\theta_d = \lambda/d. \quad (8.13)$$

Итак, *дифракционная расходимость светового пучка в дальней зоне* определяется отношением длины волны к начальному диаметру пучка d : *дифракционная расходимость пучка тем больше, чем меньше его начальный размер*. Диаметр пучка в дальней зоне выражается формулой $d(z) = z\theta_d = z\lambda/d$, z — координата, отсчитываемая вдоль пучка от экрана с отверстием. Оценим дифракционную длину z_d и угловую расходимость θ_d для пучка гелий-неонового лазера: для $d = 2$ мм, $\lambda = 0.6$ мкм получим $z_d = 1.5$ м, $\theta_d = 3 \cdot 10^{-3}$ рад.

8.3. Дифракция на краю непрозрачного экрана. Эта задача относится к числу основных проблем теории дифракции или, так именно ее решение дает объяснение как происходит переход от света к тени на границе геометрической тени. Предположим, что плоская монохроматическая световая волна встречается на своем пути полубесконечную непрозрачную плоскость с прямолинейной границей ("край экрана"). Считая, что световая волна распространяется по нормали к экрану, найдем распределение света в плоскости наблюдения, параллельной экрану и находящейся на некотором расстоянии l от него (**Рис.8.11**). Для построения дифракционной картины воспользуемся методом зон Френеля. В качестве поверхности Σ , излучающей вторичные волны, выберем плоскость, являющуюся продолжением экрана; эта поверхность совпадает с волновым фронтом световой волны. Введем френелевские зоны для точки M , лежащей в плоскости наблюдения точно под краем экрана (**Рис.8.11**).

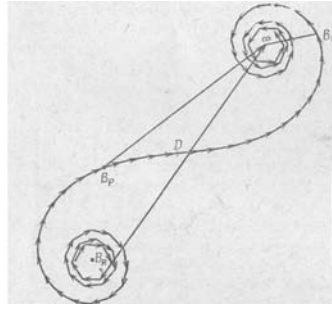


Рис. 8.12. Симметричная спираль Френеля

В данном случае зоны Френеля имеют вид плоских полос, параллельных краю экрана. Обозначая границы зон буквами $O_1, O_2, O_3, \dots$, можно записать:

$$\begin{aligned} OM &= l, \\ O_1M &= l + \lambda/2, \\ O_2M &= O_1M + \lambda/2, \\ &\dots\dots\dots \end{aligned} \quad (8.14)$$

где λ — длина световой волны. Обозначим через d_n расстояние от края экрана до начала френелевской зоны с номером n . Очевидно, что

$$d_n^2 = OO_n^2 = (l + n\lambda/2)^2 - l^2 = nl\lambda + (n\lambda/2)^2. \quad (8.15)$$

Учитывая, что $l \gg \lambda$ и пренебрегая последним слагаемым, получим приближенно

$$d_n = \sqrt{nl\lambda}. \quad (8.16)$$

Используя эту формулу, нетрудно подсчитать площадь n -й зоны Френеля:

$$S_n = (d_{n+1} - d_n)L = (\sqrt{n+1} - \sqrt{n})L\sqrt{\lambda l} = \frac{L\sqrt{\lambda l}}{\sqrt{n+1} + \sqrt{n}} \approx \frac{L}{2} \sqrt{\frac{\lambda l}{n}}. \quad (8.17)$$

Здесь L — длина края экрана. Таким образом, в данном случае площади френелевских зон уменьшаются с ростом номера зоны n как $1/\sqrt{n}$.

Теперь разделим каждую зону Френеля на большое (в пределе — бесконечно большое) число подзон. Для определения дифракционного светового поля нужно просуммировать световые колебания, создаваемые в точке наблюдения элементарными вторичными волнами, приходящими от всех открытых зон и подзон. Суммирование можно провести методом векторной диаграммы. При этом следует учесть, что поскольку при удалении от края экрана ширина френелевских зон уменьшается, длина вектора, изображающего вклад отдельной подзоны, будет тем меньше, чем дальше расположена зона от края экрана. В итоге векторная диаграмма приобретает вид, показанный на **Рис. 8.12** (верхняя часть).

Предположим теперь, что при фиксированной точке наблюдения поля M (**Рис. 8.11**) край экрана начинает отодвигаться влево. В этом случае будут "открываться" зоны Френеля, расположенные слева от первоначального положения края экрана. Картина расположения границ зон Френеля справа и слева от точки O симметрична, поэтому симметричной будет и соответствующая спираль Френеля (**Рис. 8.12**). Пользуясь спиралью Френеля, можно построить полную картину дифракции света на краю экрана.

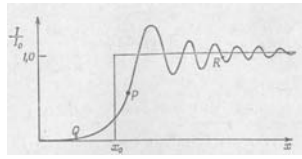


Рис. 8.13. Ход интенсивности вблизи края тени. Геометрический край тени находится в точке x_0 .

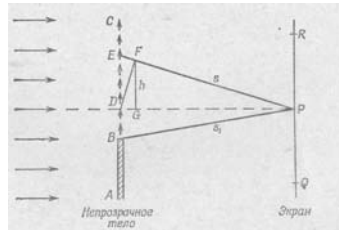


Рис. 8.14. Далекий источник отбрасывает тень от непрозрачного предмета на экран.

Пусть требуется определить интенсивность в точке P (**Рис. 8.14**). Сложим волны, от точки D вверх до бесконечности и вниз от D до точки B_P на **Рис. 8.12**. Весь вклад над областью B_P до этой точки дается спиральной кривой. Если бы суммирование заканчивалось в некоторой точке, то полная амплитуда представилась бы вектором от B_P до этой точки. В нашем случае суммирование ведется до бесконечности, так что искомая амплитуда есть вектор $\vec{B}_{P\infty}$. Точка на кривой, соответствующая точке B_P , зависит от положения точки P , потому что точка D кривой всегда относится к выбранной точке P . Следовательно, в зависимости от положения P над B начальная точка, откуда проводится вектор, попадает в разные места спирали, и результирующий вектор $\vec{B}_{P\infty}$ имеет многочисленные максимумы и минимумы, т.е. зависимость интенсивности на втором экране имеет осциллирующий вид, как показано на **Рис. 8.13**. Но если мы находимся в точке Q , по другую сторону от P , то нам понадобится только верхний конец спирали, т.е. начальной точкой результирующего вектора будет не D , а B_Q , и, следовательно, книзу от P интенсивность должна непрерывно падать при удалении Q в область тени. Нетрудно убедиться, что интенсивность в точке, лежащей прямо против края, равна $1/4$ от интенсивности падающего света, потому что для этой точки мы имеем только половину спирали в отличие от целой спирали, если бы точки лежали далеко в освещенной области. Если точка R расположена высоко, результирующий вектор проводится от центра одной спирали до центра другой, а для точки на краю тени амплитуда равна половине этого вектора. Следовательно отношение интенсивностей получается $1/4$.

Дифракция на диске и пятно Пуассона. Предположим, что плоская монохроматическая световая волна падает по нормали на круглый непрозрачный диск радиуса r , а наблюдение поля ведется в некоторой точке P , расположенной в области геометрической тени на оси диска. Действуя как и в предыдущем случае, можно показать, что зависимость амплитуды световых колебаний A в точке наблюдения от радиуса диска r является монотонно убывающей, т.е. чем больше радиус диска, тем меньше интенсивность света в точке наблюдения. Действительно, если диск закрывает m

первых зон Френеля, то амплитуда результирующего поля в точке P равна

$$A = A_{m+1} - A_{m+2} + A_{m+3} - \dots = \frac{A_{m+1}}{2} + \left(\frac{A_{m+1}}{2} - A_{m+2} + \frac{A_{m+3}}{2} \right) + \dots = \frac{A_{m+1}}{2},$$

т.е. даже в случае достаточно большого экрана, закрывающего сразу несколько зон Френеля, интенсивность света в центре геометрической тени отлична от нуля.

Таким образом, теория Френеля предсказывает проникновение света в центр геометрической тени диска, установленного на пути плоской монохроматической световой волны. В свое время этот результат рассматривался как аргумент против теории Френеля. Однако эксперименты, выполненные Д.Араго (1786-1853), показали, что при освещении непрозрачного диска светом точечного источника в центре области геометрической тени действительно существует маленькое светлое пятно! Это пятно получило название "пятно Пуассона" — по имени автора идеи эксперимента.

9. ЛЕКЦИЯ №9

Теория дифракции Кирхгофа. Приближение Френеля в теории дифракции; дифракция Френеля на одномерных структурах. Дифракция Фраунгофера как пространственное преобразование Фурье. Предел разрешения оптических приборов. Дифракционные решетки.

9.1. Приближение Френеля в теории дифракции. Интеграл Гюйгенса-Френеля (8.1), полученный на основе качественных физических соображений, получил математическое обоснование в работах Кирхгофа на основе электродинамической теории света. Было показано, что его можно рассматривать как *приближенное* решение волнового уравнения с соответствующими граничными условиями, причем коэффициент $K(\varphi) = (i/2\lambda)(1 + \cos \varphi)$. При малых углах φ получаем $K(\varphi) \approx i/\lambda$. С учетом этого интеграл Гюйгенса-Френеля принимает вид:

$$E(P) = \frac{i}{\lambda} \int_{\Sigma} E(M) \frac{\exp(-ik\rho)}{\rho} d\sigma. \quad (9.1)$$

здесь $E(P)$ — комплексная амплитуда светового поля, Σ — поверхность, стягивающая отверстие, P — точка наблюдения поля, M — некоторая точка на поверхности Σ , ρ — расстояние между точками M и P , λ — длина световой волны, $k = 2\pi/\lambda$ — волновое число.

В такой форме интеграл Гюйгенса-Френеля часто используется для количественных расчетов дифракционных картин. При расчетах дифракционных картин широко применяются два основных приближения теории дифракции, *приближение Френеля* и *приближение Фраунгофера*. Первое из них описывает дифракцию *слаборасходящихся* ("параксиальных") пучков света, а второе — дифракцию в *дальней зоне*.

Сначала рассмотрим *френелевское приближение* на примере **дифракции плоской монохроматической световой волны на отверстии**. Общее решение задачи дифракции дается в этом случае интегралом (9.1).

Введем координаты x, y в плоскости экрана с отверстием и координаты x_0, y_0 в плоскости наблюдения, находящейся на расстоянии z от экрана с отверстием и параллельной ему (**Рис.9.1**). В этом случае дифракционный интеграл принимает вид:

$$E(x_0, y_0, z) = \frac{i}{\lambda} \iint_{-\infty}^{\infty} E_0(x, y) \frac{\exp(-ik\rho)}{\rho} dx dy, \quad (9.2)$$

$$\rho = \sqrt{z^2 + (x - x_0)^2 + (y - y_0)^2}.$$

В оптике часто имеют дело с узкими *слабо расходящимися пучками света*, когда выполняются неравенства $z \gg x, y, x_0, y_0$. С учетом этих неравенств для ρ можно

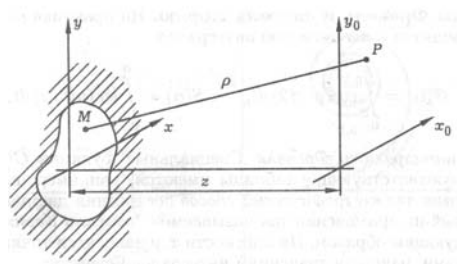


Рис. 9.1. К постановке задачи о дифракции

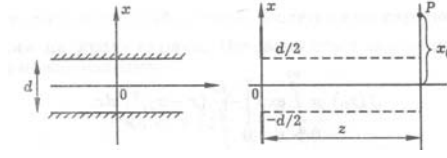


Рис. 9.2. К расчету дифракции плоской волны на щели.

приблизенно записать

$$\rho = z + \frac{(x - x_0)^2 + (y - y_0)^2}{2z}. \quad (9.3)$$

Подставляя (9.3) в (9.2) и пренебрегая отличием ρ от z в знаменателе подынтегрального выражения, получим

$$E(x_0, y_0, z) = \frac{i}{\lambda z} e^{-ikz} \iint_{-\infty}^{\infty} E_0(x, y) \exp \left[-ik \frac{(x - x_0)^2 + (y - y_0)^2}{2z} \right] dx dy. \quad (9.4)$$

Формула (9.4) дает решение задачи дифракции в *приближении Френеля*.

Формула (9.3) с физической точки зрения означает замену сферических волновых фронтов вторичных волн Гюйгенса параболическими поверхностями. Такое приближение накладывает определенные ограничения на допустимые размеры отверстия и положение точки наблюдения поля. Приближение Френеля исключает из рассмотрения некоторые области пространства: это область вблизи экрана с отверстием, а также точки пространства, лежащие на больших расстояниях от оси пучка. Однако из физических соображений ясно, что в первой из указанных областей световое поле почти такое же, как в падающей волне, а во вторую область свет почти не проникает. Поэтому этими областями можно пренебречь.

Дифракция Френеля на щели. Если начальное поле $E(x, y)$ зависит лишь от одной пространственной переменной, например, как в случае дифракции на одной щели, $E_0(x, y) = E_0(x)$, говорят, что имеет место *дифракция на одномерной структуре*. В этом случае, вычисляя интеграл по y , для (9.4) можно записать

$$E(x_0, z) = \frac{1 + i}{\sqrt{2\lambda z}} e^{-ikz} \int_{-\infty}^{\infty} E_0(x) \exp \left[-\frac{ik(x - x_0)^2}{2z} \right] dx. \quad (9.5)$$

При выводе (9.5) воспользовались известным интегралом

$$\int_{-\infty}^{\infty} \exp \left(-\frac{iky^2}{2z} \right) dy = \sqrt{\frac{2\pi z}{ik}} = (1 - i) \sqrt{\frac{\pi z}{k}}. \quad (9.6)$$

Как пример, рассмотрим **дифракцию на щели**. Обозначим ширину щели d и направим ось x перпендикулярно щели (Рис.9.2). Поле на щели зададим в виде

$$E_0(x) = \begin{cases} 1, & |x| \leq d/2; \\ 0, & |x| > d/2. \end{cases} \quad (9.7)$$

Подставив (9.7) в (9.5), получим

$$E(x_0, z) = E(x_0) \frac{1 + i}{\sqrt{2\lambda z}} e^{-ikz} F(x_0), \quad F(x_0) = \int_{-d/2}^{d/2} \exp \left[-\frac{ik(x - x_0)^2}{2z} \right] dx, \quad (9.8)$$

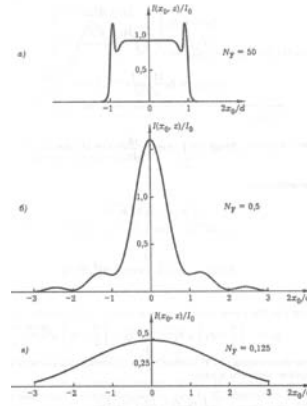


Рис. 9.3. Дифракция плоской волны на щели

Введем переменную $\xi = (x - x_0)\sqrt{k/z\pi}$. Тогда

$$F(x_0) = \sqrt{\frac{\pi z}{k}} \int_{\xi_1}^{\xi_2} \exp\left(-\frac{i\pi\xi^2}{2}\right) d\xi, \quad (9.9)$$

где $\xi_1 = -(d/2 + x_0)\sqrt{k/\pi z}$, $\xi_2 = (d/2 - x_0)\sqrt{k/\pi z}$

Распределение интенсивности света в дифракционной картине запишем в виде

$$I(x_0, z) = \frac{1}{2} I_0 \left| \int_{\xi_1}^{\xi_2} \exp\left(-\frac{i\pi\xi^2}{2}\right) d\xi \right|^2, \quad (9.10)$$

где I_0 — интенсивность падающей волны; $\xi_1 = -\alpha(1 + p)$, $\xi_2 = \alpha(1 - p)$, $\alpha = \sqrt{kd^2/4\pi z} = \sqrt{2N_F}$ (N_F — число Френеля), $p = 2x_0/d$.

Интеграл в (9.10) удобно записать через интегралы Френеля $C(\xi) = \int_0^\xi \cos(\pi t^2/2) dt$ и $S(\xi) = \int_0^\xi \sin(\pi t^2/2) dt$, которые играют важную роль в теории дифракции:

$$\int_{\xi_1}^{\xi_2} \exp\left(-\frac{i\pi\xi^2}{2}\right) d\xi = [S(\xi_1) - S(\xi_2)] + [C(\xi_1) - C(\xi_2)]^2. \quad (9.11)$$

Функции $S(\xi)$ и $C(\xi)$ — это специальные функции, которые затабулированы, а таблицы имеются в соответствующих справочниках.

Вид дифракционной картины существенно зависит от *числа Френеля*. При $N_F \gg 1$ дифракция почти не проявляется (**Рис.9.3 а**): профиль интенсивности излучения остается почти прямоугольным, ширина пучка остается равной ширине щели, а интенсивность света оси пучка совпадает с интенсивностью падающей волны. Влияние дифракции заметно лишь вблизи границы области геометрической тени, где наблюдаются осцилляции интенсивности и свет слегка проникает в область геометрической тени. По мере удаления точки наблюдения от экрана со щелью, число Френеля уменьшается. В области, где $N_F \simeq 1$ (**Рис.9.3 б**), интенсивность света на оси пучка испытывает значительные осцилляции, появляются боковые максимумы интенсивности, однако ширина светового пучка все еще примерно равна ширине щели. В дальней зоне, где $N_F \ll 1$, световой пучок сильно уширяется и поперечный профиль пучка уже не имеет ничего общего с исходным профилем (**Рис.9.3 в**).

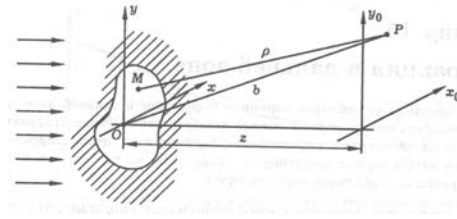


Рис. 9.4. К постановке задачи дифракции Фраунгофера

9.2. Дифракция Фраунгофера как пространственное преобразование Фурье.

Дифракция Фраунгофера или дифракция в параллельных лучах наблюдается, если источник света и точка наблюдения бесконечно удалены от препятствия, вызвавшего дифракцию. Схема наблюдения фраунгоферовской дифракции показана на **Рис.9.4**.

Опыты по дифракции световых пучков показывают, что на достаточно большом расстоянии, или, как говорят, в *дальней зоне* угловое распределение интенсивности излучения перестает зависеть от координаты z , отсчитываемой вдоль оси пучка. Картина дифракции приобретает устойчивую структуру, вид которой зависит только от распределения поля в начальном сечении. Дифракцию в *дальней зоне* называют *дифракцией Фраунгофера*.

Пусть плоская монохроматическая световая волна нормально падает на экран с отверстием, расположенный в плоскости $z = 0$ (**Рис.9.4**). Вычислим распределение интенсивности излучения в некоторой плоскости x_0, y_0 , параллельной экрану с отверстием и расположенной на достаточно большом расстоянии z от него. Используем формулу Гюйгенса-Френеля (9.1).

Пусть O — некоторая точка в плоскости экрана с отверстием, которую мы примем за начало отсчета, а $b = \sqrt{z^2 + x_0^2 + y_0^2}$ — расстояние от точки O до точки наблюдения поля P . В параксиальном приближении, когда $z \gg x, y, x_0, y_0$, для ρ можно записать:

$$\rho = b + \frac{x^2 + y^2}{2b} - \frac{xx_0 + yy_0}{b} \quad (9.12)$$

Подставляя (9.12) в интеграл Френеля (9.1), получим

$$E(x_0, y_0, z) = \frac{i}{\lambda b} e^{-ikb} \iint_{-\infty}^{\infty} E_0(x, y) \exp \left[-\frac{ik(x^2 + y^2)}{2b} \right] \exp \left[\frac{ik(xx_0 + yy_0)}{b} \right] dx dy \quad (9.13)$$

В одномерном случае

$$E(x_0, y_0, z) = \frac{i + 1}{\sqrt{2\lambda b}} e^{-ikb} \int_{-\infty}^{\infty} E_0(x) \exp \left(-\frac{ikx^2}{2b} \right) \exp (ikx \sin \theta) dx, \quad (9.14)$$

где $\sin \theta = x_0/b$ — имеет смысл *угловой координаты* точки наблюдения поля.

Формулы (9.14) соответствуют френелевскому приближению. Из нее следует, что угловое распределение поля в дифракционной картине меняется по мере изменения расстояния z . Однако в области больших z это изменение становится все более и более слабым и при

$$kd^2/2b \ll 1 \quad \text{или} \quad z \gg z_d = kd^2/2, \quad (9.15)$$

где d — максимальный поперечный размер отверстия в экране, а z_d — *дифракционная длина пучка*, устанавливается устойчивое угловое распределение поля, определяемое формулой

$$E(x_0, y_0, z) = \frac{i+1}{\sqrt{2\lambda b}} e^{-ikb} \int_{-\infty}^{\infty} E_0(x) \exp(ikx \sin \theta) dx, \quad (9.16)$$

Область пространства, определяемая условием (9.15), называется *дальней зоной* дифракции или *зоной Фраунгофера*. Таким образом, в *дальней зоне* формируется *устойчивое угловое распределение поля*, не меняющееся при дальнейшем распространении светового пучка.

Выражение (9.16) называют *дифракционным интегралом в приближении Фраунгофера*. Оно справедливо в *дальней дифракционной зоне*.

С математической точки зрения выражение (9.16) представляет собой *пространственный интеграл Фурье*. По аналогии с интегралом Фурье по времени величину $k_x = k \sin \theta$ называют *пространственной частотой*. По физическому смыслу это есть *поперечная компонента волнового вектора*, направленного в точку наблюдения поля. Причем между *пространственной частотой k_x* и *угловой координатой θ* точки наблюдения поля имеется взаимно однозначное соответствие. Это позволяет записать комплексную амплитуду поля в точке наблюдения следующим образом:

$$E(P) = \frac{i+1}{\sqrt{2\lambda b}} e^{-ikb} E_0(k_x) \exp(ikx \sin \theta) dx, \quad (9.17)$$

$$E_0(k_x) = \int_{-\infty}^{\infty} E_0(x) \exp(ikx \sin \theta) dx,$$

где $E_0(k_x)$ — *пространственная спектральная амплитуда*, соответствующая распределению поля $E_0(x)$.

Таким образом, *дифракционное поле в дальней зоне пропорционально пространственной фурье-амплитуде исходного пучка*. Используя (9.17), запишем выражение для интенсивности излучения в *дальней зоне*:

$$I(P) = \frac{c}{8\pi} \frac{1}{\lambda b} S_0(k_x), \quad (9.18)$$

где $S_0(k_x) = |E_0(k_x)|^2$ — *пространственная спектральная плотность*, или *угловой спектр излучения*.

Угловое распределение интенсивности излучения в дальней зоне повторяет форму углового спектра светового пучка. Отсюда следует физический смысл фраунгоферовой дифракции как *пространственного разложения ограниченного светового пучка на плоские волны*. Согласно спектральным представлениям, поперечная компонента волнового вектора возникает вследствие ограничения апертуры (т.е. поперечных размеров) пучка отверстием. Представление ограниченного пучка в виде набора плоских волн, распространяющихся в разных направлениях, аналогично представлению импульса конечной длительности в виде суммы гармонических колебаний разных частот.

Дифракция Фраунгофера на щели. Рассмотрим *фраунгоферовскую дифракцию плоской волны на щели* шириной d (Рис.9.5). Полагая, что

$$E_0(x) = \begin{cases} 1, & |x| \leq d/2; \\ 0, & |x| > d/2, \end{cases} \quad (9.19)$$

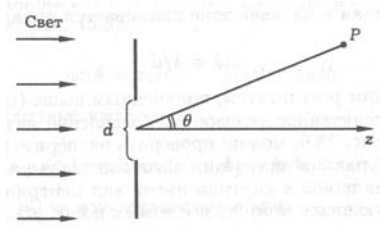


Рис. 9.5. К расчету фраунгоферовской дифракции на щели

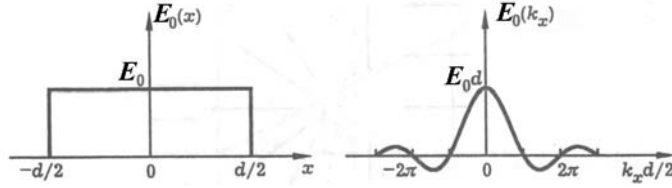


Рис. 9.6. Начальное распределение амплитуды поля $E_0(x)$ и пространственная спектральная амплитуда $E_0(k_x)$ при дифракции плоской монохроматической волны на щели шириной d .

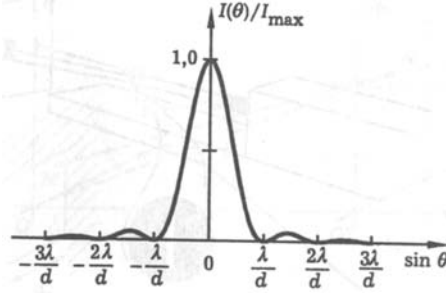


Рис. 9.7. Угловое распределение интенсивности света в дальней зоне при дифракции на щели: d — ширина щели, λ — длина световой волны

по формуле (9.17) получим

$$E(k_x) = E_0 \int_{-d/2}^{d/2} E_0(x) e^{ik_x x} \sin \theta dx = E_0 d \cdot \text{sinc}(k_x d/2), \quad \text{sinc}(x) = \sin x/x. \quad (9.20)$$

Графики функций E_0 и E_{k_x} показаны на **Рис. 9.6**. Используя (9.17), и (9.18), находим амплитуду поля и интенсивность света в точке P :

$$E(P) = E_0 \frac{i+1}{\sqrt{2\lambda b}} e^{-ikb} \text{sinc}(k_x d/2), \quad I(P) = I_{\max} \text{sinc}^2(k_x d/2), \quad (9.21)$$

где $I_{\max} = I_0 d^2 / \lambda b$, $I_0 = (c/8\pi) |E_0|^2$ — интенсивность падающей волны. Подставляя $k_x = k \sin \theta$ в выражение для интенсивности, находим угловое распределение интенсивности излучения в дальней зоне

$$I(P) = I_{\max} \text{sinc}^2(kd \sin \theta/2) = I_{\max} \text{sinc}^2\left(\frac{\pi d \sin \theta}{\lambda}\right). \quad (9.22)$$

График распределения угловой интенсивности показан на **Рис. 9.7**. Отсюда можно определить дифракционную расходимость пучка $\Delta\theta$ в дальней зоне, исходя из условия $I(\theta) = 0$. Отсюда следует, что При $\theta \ll 1$ угол отклонения θ , при котором $I(\theta) = 0$, $\theta = \lambda/d$, а полный угол расходимости $\theta_d = 2\theta = 2\lambda/d \sim \lambda/d$.

Дифракция Фраунгофера на круглом отверстии. Аналогичным образом можно рассмотреть дифракцию на *круглом отверстии*. В этом случае угловое распределение интенсивности в дальней зоне описывается функцией Эйри:

$$I(\theta) = I_0 \left[\frac{J_1(u)}{u} \right]^2, \quad (9.23)$$

где I_0 — интенсивность падающей волны; $J_1(u)$ — функция Бесселя первого порядка; $u = 2\pi\theta r/\lambda$; r — радиус отверстия.

График функции Эйри аналогичен графику **Рис.9.8**. Нетрудно показать, что на центральный максимум, называемый *диск Эйри*, приходится 84% светового потока. Полная угловая ширина (дифракционная расходимость) равна

$$\theta_d = 1.22 \frac{\lambda}{d}, \quad (9.24)$$

т.е. при дифракции плоской волны на круглом отверстии она становится расходящейся с углом расходимости θ_d , причем, чем больше d , меньше θ_d .

Предел разрешения оптических приборов. Способность оптических приборов, например линзы, давать *раздельные изображения* двух близко расположенных точечных объектов, называется *разрешением*. Разрешение линзы тем больше, чем более близко лежащие точки различимы, как отдельные изображения. Два *принципиальных* фактора ограничивают разрешение линзы. Первый — это наличие *аббераций*. Вторым фактором, ограничивающим разрешение, является *дифракция*, которая *принципиально не устранима*, так как она есть результат волновой природы света. Именно дифракция ответственна за существование *предела разрешения* линзы, даже при *отсутствии аббераций*. Поэтому обсудим этот фактор подробнее.

Любая линза имеет *конечную апертуру* (поперечный размер). При формировании изображения точечного объекта с помощью линзы, имеет место дифракция на ее апертуре. Фактически изображение представляет собой дифракционную картину. Для круглого отверстия изображение точки представляет круглый центральный пик (диск Эйри), окруженный чередующимися темными и светлыми интерференционными кольцами. Центральный пик имеет угловую ширину $\theta_d = 1.22\lambda/d$, где d — диаметр. Если два точечных объекта очень близки, то их изображения перекрываются.

Рэлей предложил простой критерий, согласно которому два изображения *разрешимы*, когда *центр дифракционного диска одного изображения совпадает с первым минимумом дифракционной картины другого* — **критерий Рэля (Рис.9.8)**.

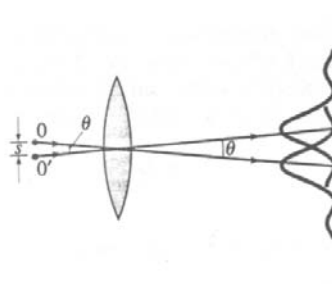


Рис. 9.8. Иллюстрация критерия Рэля. O и O' — точечные объекты, угловое расстояние между которыми θ .

Другими словами, это означает, что два объекта разрешимы, если они разделены углом θ_d . Это и есть **предел разрешения**, допускаемый волновой природой света из-за дифракции.

9.3. Дифракционные решетки. Дифракционная решетка представляет собой пространственную периодическую структуру, период которой **соизмерим** с длиной световой волны. Различают пропускательные, отражательные, амплитудные и фазовые решетки. Пропускательные решетки работают на пропускание света, отражательные — на отражение. Амплитудные решетки пространственно модулируют амплитуду, а фазовые — фазу световой волны. Простейшая амплитудная пропускательная решетка представляет собой систему равноотстоящих параллельных щелей в непрозрачном экране. Отражательную амплитудную решетку изготавливают путем нанесения штрихов на плоское или вогнутое зеркало. Фазовая решетка может представлять собой профилированную стеклянную пластину или профилированное зеркало.

Рассмотрим дифракцию плоской монохроматической световой волны на периодической системе щелей, в непрозрачном экране (**Рис.9.9**). Падающая на решетку световая волна создает в щелях когерентные (сфазированные) источники вторичных световых волн. Результирующее световое поле образуется в результате интерференции этих волн. Определим распределение интенсивности света в *фраунгоферовой дифракционной картине*, наблюдаемой на *бесконечности* либо в *фокальной плоскости линзы* при падении на решетку плоской монохроматической волны. В каждой точке наблюдения происходит многолучевая интерференция когерентных пучков света одинаковой интенсивности, дифрагировавших на отдельных одинаковых щелях (**Рис.9.9**). Обозначим через $E_1 = E_0 \sin u/u$ поле в точке наблюдения, излучаемое первой щелью (см.(9.22)). Поля, излучаемые остальными щелями, имеют вид

$$E_2 = E_1 e^{i\delta}, E_3 = E_1 e^{2i\delta}, \dots, E_N = E_1 e^{i(N-1)\delta}, \quad (9.25)$$

где δ — разность фаз между соседними световыми пучками.

Полная напряженность в точке P от всех N периодических элементов решетки есть сумма всех полей

$$E_P = E_1 [1 + e^{i\delta} + e^{2i\delta} + \dots + e^{i(N-1)\delta}] = E_1 \frac{1 - e^{iN\delta}}{1 - e^{i\delta}}, \quad (9.26)$$

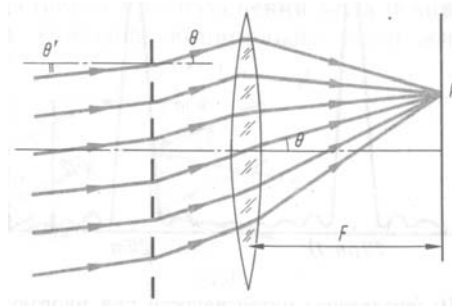


Рис. 9.9. Дифракция Фраунгофера на системе одинаковых параллельных щелей

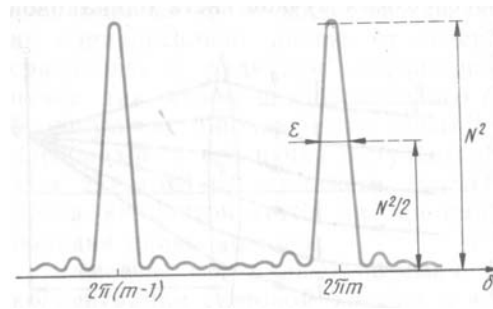


Рис. 9.10. Распределение интенсивности при фраунгоферовской дифракции на решетке, состоящей из одинаковых щелей $N = 8$

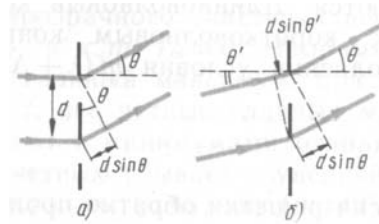


Рис. 9.11.

которая является суммой геометрической прогрессии. Для интенсивности $I_P \sim |E_P|^2$ получаем

$$I_P = I_1 \frac{\sin^2(N\delta/2)}{\sin^2(\delta/2)}. \quad (9.27)$$

Здесь I_1 — интенсивность в точке наблюдения P от одной щели. В результате интерференции всех когерентных вторичных волн происходит перераспределение светового потока по направлениям и получается существенно отличающееся от I_1 распределение интенсивности в фраунгоферовой дифракционной картине.

В тех направлениях θ_m , для которых $\delta/2 = m\pi$ ($m = \pm 1, \pm 2, \dots$), второй множитель в (9.27) принимает значение N^2 , и интенсивность в N^2 раз больше, чем от одной щели. Так происходит потому, что разность хода Δ вторичных волн от соседних щелей для направлений θ_m равна целому числу длин волн ($\Delta = m\lambda$), и все они приходят в точку наблюдения в одинаковой фазе. Таким образом, в этом направлении будет наблюдаться *максимум интенсивности*, если только оно не совпадает с минимумом распределения интенсивности от одной щели. Такие максимумы называются *главными*, а целое число m — *порядком главного максимума* или *порядком спектра*.

Интенсивность в (9.27) обращается в нуль каждый раз, когда $\sin(N\delta/2) = 0$, но $\sin(\delta/2) \neq 0$. В соответствующих таким значениям δ направлениях лежат *минимумы*, интенсивность света в которых равна нулю. Между ними находятся *побочные* (или *дополнительные*) максимумы. Между двумя соседними главными максимумами расположены $N - 1$ побочных максимумов.

График второго множителя в (9.27) приведен на **Рис. 9.10**. Разность фаз δ между соседними интерферирующими пучками, зависит от направления θ и от *постоянной решетки* d — пространственного периода. Для щелевой решетки (**Рис. 9.11**) d — это есть суммарная ширина прозрачного и непрозрачного участков.

При наклонном падении плоской волны под углом θ' (**Рис.9.11 б**) разность хода соседних пучков Δ и положение главных максимумов определяется условием

$$\Delta = d(\sin \theta_m - \sin \theta') = m\lambda. \quad (9.28)$$

При нормальном падении монохроматической волны на решетку (**Рис.9.10 а**) $\theta' = 0$, т.е. $\Delta = d \sin \theta$ и направление на главный максимум m -го порядка определяется условием:

$$d \sin \theta_m = m\lambda. \quad (9.29)$$

Формулы (9.28) и (9.29) называют *уравнением дифракционной решетки*. Оно позволяет сделать важные выводы. Из (9.29) следует, что решетка будет давать заметную дифракцию, т.е. значительные углы отклонения θ , только в том случае, если *период решетки соизмерим с длиной световой волны* — $d \simeq \lambda \simeq 10^{-4}$ см. Следовательно, оптическая дифракционная решетка должна иметь число штрихов (щелей) на миллиметр порядка $10^2 - 10^3$. Изготовить такую решетку — сложная техническая задача. Поэтому неслучайно, что первые дифракционные решетки хорошего качества появились лишь в XIX в. Решетка с мелким периодом должна отклонять лучи сильнее, чем решетка с крупным периодом. Вместе с тем, решетка со слишком мелким периодом (период меньше длины волны) вообще не будет давать дифракции, так как согласно через такую решетку может проходить только не отклоненная волна ($m = 0, \theta = 0$).

Из (9.28) также видно, что при скользящем падении, когда угол $\theta' \approx \pi/2$, даже грубая решетка может давать заметную дифракцию. Действительно, при $d \gg \lambda$ $\sin \theta \sin \theta' = 2 \sin [(\theta - \theta')/2] \cos [(\theta + \theta')/2] \approx (\theta - \theta') \cos \theta'$. Тогда уравнение решетки (9.28) принимает вид

$$\theta - \theta' = m\lambda/d_e, \quad d_e = d \cos \theta'. \quad (9.30)$$

Так как $\theta' \approx \pi/2$, то $d_e \ll d$, т.е. эффективный период d_e может быть много меньше d .

Из уравнения (9.28) следует, что *положение главных максимумов дифракционной картины зависит от длины волны*. Поэтому, если направить на решетку пучок немонохроматического излучения, то разные спектральные составляющие излучения будут отклоняться решеткой на разные углы. Отсюда вытекает возможность использовать дифракционную решетку как *спектральный прибор*, который, подобно призме, осуществляет *пространственное разложение немонохроматического излучения по длинам волн*.

При большом числе интерферирующих пучков (в современных дифракционных решетках N достигает 200 тыс.) получаются очень узкие и резкие главные максимумы. Ширина ε главного максимума на половине высоты определяется из условия $\sin^2(N\varepsilon/4)/(\varepsilon/4)^2 = N^2/2$, откуда следует, что $\varepsilon = 5.6/N$. *Резкость* многолучевой интерференционной картины в монохроматическом свете (т. е. отношение расстояния между главными максимумами соседних порядков к их ширине) определяется полным числом штрихов: $F = 2\pi/\varepsilon = 1.13N$.

Важным параметром любого спектрального прибора является *разрешающая способность* $R = \lambda/\delta\lambda$, где $\delta\lambda = |\lambda_2 - \lambda_1|$ — абсолютное значение минимальной разности длин волн двух соседних спектральных линий с длинами волн $\lambda_{1,2}$, при которой эти линии регистрируются отдельно. Определим разрешающую способность для дифракционной решетки. Согласно критерию Рэлея $\delta\lambda$ определяется из условия, что интерференционный максимум m -го порядка для λ_2 должен совпадать с минимумом для λ_1 . Отсюда следует, что $m\lambda_2 = \lambda_1 + \lambda_1/N$. Откуда получаем, что $R = \lambda/\delta\lambda = mN$.

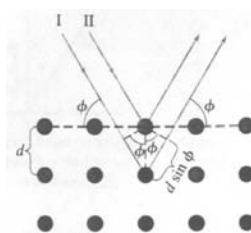


Рис. 9.12. Дифракция рентгеновских лучей на кристалле. d — расстояние между атомными плоскостями.

Другим важным параметром является *угловая дисперсия решетки* $D_\theta = d\theta/d\lambda$, которая характеризует изменение положения главных максимумов при изменении длины волны. Дифференцируя (9.28) и предполагая нормальное падение света на решетку, получаем, что $d\theta/d\lambda = \tan \theta/\lambda$, т.е. дисперсия растет с увеличением угла дифракции θ . Таким образом, для спектра определенного порядка m дисперсия тем больше, чем меньше период решетки.

Формула (9.28) также позволяет определить максимальную спектральную ширину $\Delta\lambda$ исследуемого излучения, при которой спектры соседних порядков не перекрываются. Это соответствует условию $m(\lambda + \Delta\lambda) = (m + 1)\lambda$, откуда следует, что $\Delta\lambda = \lambda/N$. Величину $\Delta\lambda$ называют *свободной областью дисперсии* дифракционной решетки.

Мы рассмотрели простейшие примеры дифракции на одномерных периодических структурах. Еще более интересные явления наблюдаются при дифракции на двумерных и трехмерных периодических структурах. Последнее имеет место, например, при дифракции рентгеновских лучей в кристаллах.

9.4. Дифракция рентгеновских лучей. Электромагнитное излучение с длиной волны единицы нанометров и меньше называют *рентгеновскими*. Они были открыты в 1895 году В.Рентгеном и названы им *X-лучами*. Одним из самых замечательных свойств рентгеновского излучения считается большая проникающая способность — они хорошо проходят через непрозрачные для обычного света вещества. Так как их длина волны сравнима с периодом кристаллической решетки, то они эффективно дифрагируют на них. Предположим, что параллельный монохроматический пучок рентгеновских лучей с длиной волны λ падает под углом ϕ на поверхность кристалла, с кубической решеткой, где атомы расположены на расстоянии d , как показано на **Рис.9.12**. Два луча, отраженных от соседних атомных плоскостей конструктивно интерферируют, если выполняются следующие условия:

$$2d \sin \phi = m\lambda, \quad m = 1, 2, 3, \dots \quad (9.31)$$

где $m = 1, 2, 3, \dots$. Формулу (9.31) называют условием Вульфа-Брэгга.

Таким образом, в направлении угла ϕ наблюдается дифракционный максимум. Если известна λ , то измеряя ϕ и m , можно найти d — расстояние между кристаллографическими плоскостями, т.е. можно определять структуру кристаллов — *рентгеноструктурный анализ*. Если известны d , то измеряя ϕ и m , можно найти длину волны падающего излучения — *рентгеновская спектроскопия*.

В последние годы огромный интерес получили исследования распространения света в средах с периодически модулированной диэлектрической проницаемостью с периодом, сравнимым с длиной волны света. Оказалось, что такие структуры обладают

очень удивительными и интересными свойствами. В частности, здесь возникают полностью запрещенные зоны в спектре собственных колебаний состояний кристалла, подобно электронным спектрам в твердых телах. Такие структуры получили название *фотонных кристаллов*. Указанное свойство отличает фотонные кристаллы от дифракционных решеток. Оно означает, что в данном спектральном диапазоне свет не может войти в образец или выйти из него в любом направлении. Это уникальное свойство фотонного кристалла с которым связывают возможные революционные преобразования в технике оптической связи, физике лазеров и оптической компьютерной технологии.

ЛЕКЦИЯ №10

Классическая электронная теория дисперсии Лоренца: показатель преломления среды, дисперсия и поглощение света в линейной изотропной среде. Распространение светового импульса в диспергирующей среде, групповая скорость. Рассеяние света в мутных средах.

10.1. Классическая электронная теория дисперсии Лоренца: показатель преломления среды, дисперсия и поглощение света в линейной изотропной среде. При рассмотрении распространения света в веществе были *феноменологически* введены *материальные уравнения*, связывающие *поляризованность* среды $\vec{P}$ (или электрическую индукцию $\vec{D}$) с напряженностью $\vec{E}$ электрического поля монохроматической волны, что позволило описать свойства монохроматических плоских волн в веществе. В линейной изотропной среде материальные уравнения имеют вид (см. лекцию №3):

$$\vec{P} = \chi(\omega)\vec{E}, \quad \vec{D} = \varepsilon(\omega)\vec{E}, \quad \varepsilon(\omega) = 1 + 4\pi\chi(\omega) \quad (10.1)$$

Для обоснования феноменологических материальных уравнений и нахождения явного вида входящей в них диэлектрической восприимчивости χ (или диэлектрической проницаемости ε) в той или иной среде необходимо рассматривать *микроскопическую теорию взаимодействия световой волны с веществом*.

Наиболее правильная и последовательная теория взаимодействия света с веществом должна строиться на основе квантовой механики. Однако основные закономерности распространения света в веществе можно понять, используя *классическую модель среды*, согласно которой, среда представляет собой *группу невзаимодействующих между собой атомов* (по крайней мере это справедливо для газообразной среды). В простейшем случае предполагается, что в атоме имеется один электрон, называемый *оптическим*, связанный с ядром, которое считается неподвижным. В первом приближении предполагается, что сила f , действующая на этот электрон со стороны ядра линейно зависит от смещения x электрона от положения равновесия ($\vec{f} = -k\vec{x}$, k — коэффициент пропорциональности). Поэтому можно считать, что под действием света электрон ведет себя как будто он закреплен на "пружинке", т.е. представляет собой *линейный осциллятор* (колебательная система) с массой m и резонансной частотой ω_0 , а его поведение описывается законами классической механики. Такая модель называется *классической осцилляторной моделью атома* или *моделью Лоренца*.

Отметим что в действительности это не так, но, тем не менее, основные результаты данной модели подтверждаются экспериментом и теорией, основанной на квантовой механике. Классическое описание не противоречит квантовому, а является его предельным случаем.

Далее будем считать вещество *линейным, изотропным и однородным*. Оптическую поляризованность среды $\vec{P}$, имеющую смысл дипольного момента единицы объема, можно представить в виде

$$\vec{P} = N\vec{p}, \quad \vec{p} = e\vec{x}, \quad (10.2)$$

где N — число атомов в единице объема, $\vec{p}$ — *дипольный момент* отдельного атома, e — заряд электрона, $\vec{x}$ — смещение электрона относительно равновесного положения. Так как смещение электрона относительно ядра происходит под действием

световой волны, то $\vec{x} = \vec{x}(\vec{E})$. Зависимость $\vec{x}(\vec{E})$ можно определить, используя классическую осцилляторную модель атома. В этой модели атомы представляют собой гармонические осцилляторы, колебания которых описываются на основе классических представлений и законов Ньютона.

Пусть световая волна является плоской волной

$$\vec{\mathcal{E}} = (1/2)\vec{E} \exp[i(\omega t - \vec{k}\vec{r})] + \text{к.с.} \quad (10.3)$$

Уравнение движения осциллятора под действием световой волны согласно второго закона Ньютона имеет вид:

$$\ddot{\vec{x}} + \gamma\dot{\vec{x}} + \omega_0^2\vec{x} = \frac{e}{m}\vec{\mathcal{E}} \quad (10.4)$$

Здесь m — масса электрона, $\omega_0 = \sqrt{k/m}$ — собственная частота колебаний электрона в атоме; параметр γ описывает затухание колебаний, обусловленное тем, что колеблющийся электрон сам становится источником излучения и поэтому его свободные колебания затухают.

С учетом (10.3) уравнение (10.4) принимает вид:

$$\ddot{\vec{x}} + \gamma\dot{\vec{x}} + \omega_0^2\vec{x} = \frac{1}{2}\frac{e}{m}\vec{E} \exp[i(\omega t - \vec{k}\vec{r})] + \text{к.с.} \quad (10.5)$$

Здесь $\vec{r}$ имеет смысл радиус-вектора, который описывает положение электрона. Поскольку размер атома a много меньше длины световой волны λ , то $\vec{k}\vec{r} \approx ka = 2\pi a/\lambda \ll 1$. Поэтому приближенно атом можно считать *точечным* и не учитывать изменение поля в пределах атома.

Ищем решение уравнения (10.5) в виде

$$\vec{x} = \frac{1}{2}\vec{x}_0 \exp[i(\omega t - \vec{k}\vec{r})] + \text{к.с.}, \quad (10.6)$$

где $\vec{x}_0$ — комплексная амплитуда смещения электрона относительно атомного ядра в состоянии равновесия, т.е. в отсутствие света. Подставляя (10.6) в (10.5) находим

$$\vec{x}_0 = \frac{e}{m} \frac{1}{\omega_0^2 - \omega^2 + i\omega\gamma}. \quad (10.7)$$

Из (10.2) и (10.6), для поляризации среды получаем

$$\vec{P} = Ne\vec{x} = \frac{1}{2}Ne\vec{x}_0 \exp[i(\omega t - \vec{k}\vec{r})] + \text{к.с.} \quad (10.8)$$

В комплексном представлении $\vec{P} = Ne\vec{x} = (1/2)\vec{P}_0 \exp[i(\omega t - \vec{k}\vec{r})] + \text{к.с.}$, из (10.8) следует, что

$$\vec{P}_0 = \frac{Ne^2}{m} \frac{1}{\omega_0^2 - \omega^2 + i\omega\gamma} \vec{E} \quad (10.9)$$

Учитывая, что $\vec{P}_0 = \varkappa(\omega)\vec{E}$, находим выражение для *линейной оптической восприимчивости* среды в модели Лоренца

$$\varkappa(\omega) = \frac{Ne^2}{m} \frac{1}{\omega_0^2 - \omega^2 + i\omega\gamma}, \quad (10.10)$$

и для диэлектрической проницаемости $\varepsilon(\omega) = 1 + 4\pi\varkappa(\omega)$

$$\varepsilon(\omega) = 1 + \frac{4\pi Ne^2}{m} \frac{1}{\omega_0^2 - \omega^2 + i\omega\gamma} = 1 + \frac{\omega_p^2}{\omega_0^2 - \omega^2 + i\omega\gamma}. \quad (10.11)$$

Параметр $\omega_p = \sqrt{4\pi Ne^2/m}$ имеет размерность частоты и его часто называют "*плазменной частотой*".

Оптические свойства среды также характеризуют *комплексным показателем преломления* n , который для немагнитной среды связан с диэлектрической проницаемостью соотношением $n = \sqrt{\varepsilon(\omega)} = n' - in''$. Реальную часть $n' = \text{Re } n$ называют — *вещественным показателем преломления*, а мнимая часть $n'' = \text{Im } n$, как будет показано далее, ответственна за *поглощения (затухание)* среды световой волны в среде.

Пусть среда является достаточно разреженной, т.е. плотность атомов N мала. В этом случае ε мало отличается от единицы ($\varepsilon = 1 + \alpha$, $\alpha \ll 1$), и для $n = \sqrt{\varepsilon} \simeq 1 + \alpha/2$ можно записать

$$n = 1 + \frac{1}{2} \frac{\omega_p^2}{\omega_0^2 - \omega^2 + i\omega\gamma} = 1 + \frac{\omega_p^2}{2} \frac{\omega_0^2 - \omega^2 - i\omega\gamma}{(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2} \quad (10.12)$$

Формула дает не только показатель преломления, через атомные постоянные, но и показывает как он меняется с частотой света ω . Зависимость показателя преломления от частоты называют *дисперсией*, так как именно из-за дисперсии свет разлагается ("диспергирует") призмой в спектр.

Из (10.12) для n' и n'' нетрудно получить следующие выражения

$$n'(\omega) = 1 + \frac{\omega_p^2}{2} \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2} = 1 + \Delta n, \quad (10.13)$$

$$n''(\omega) = \frac{\omega_p^2}{2} \frac{\omega\gamma}{(\omega_0^2 - \omega^2)^2 + \omega^2\gamma^2}. \quad (10.14)$$

Дисперсия и поглощение света в линейной изотропной среде. Взаимодействие света с линейной изотропной средой приводит к двум основным физическим эффектам: *уменьшению фазовой скорости света в среде* и его *затуханию* за счет поглощения. Так как *волновое число* $k = \omega n/c$, то, k тоже комплексное число $k = k' - ik''$, где $k' = \omega n'/c$, а $k'' = \omega n''/c$.

Покажем, что *мнимая часть* волнового числа k'' , а значит и мнимая часть показателя преломления n'' , *определяет затухание световой волны* ($k'' > 0$), а *действительная часть* — *фазовую скорость* распространения света в среде.

Пусть плоская световая волна падает нормально на границу среды. В среде ее можно записать в виде

$$\mathcal{E} = \frac{1}{2} E_0 \exp [i(\omega t - kz)], \quad (10.15)$$

где E_0 — амплитуда поля на границе среды $z = 0$. Учитывая, что $k = k' - ik''$, получаем

$$\mathcal{E} = \frac{1}{2} E_0 e^{-k''z} \exp [i(\omega t - k'z)]. \quad (10.16)$$

Вычисляя интенсивность, получаем

$$I = I_0 \exp (-\alpha z), \quad (10.17)$$

где I_0 — интенсивность волны на границе среды; $\alpha = 2k'' = 2\omega n''/c$ — *показатель поглощения среды*. Формула (10.17) отражает *закон поглощения света в среде — закон Бугера*.

Таким образом, *амплитуда волны по мере распространения в среде уменьшается по экспоненциальному закону* — световая волна затухает. Говорят, что среда *поглощает* часть волны, и из среды она выходит, потеряв часть своей энергии. Потеря энергии связана с *затуханием* осцилляторов.

Именно *поглощение* дает в солнечном спектре темные линии. Свет, излучаемый поверхностью Солнца, проходит через солнечную атмосферу (и атмосферу Земли), и часть его на частотах, равных резонансным частотам атомов в атмосфере Солнца, сильно поглощается. Наблюдение подобных спектральных линий солнечного света позволило установить резонансные частоты атомов, и, следовательно, химический состав солнечной атмосферы. С помощью этих методов обнаружили, что *химические элементы на Солнце и звездах ничем не отличаются от земных*.

Так как фазовая скорость световой волны $v_p = \omega/k'$, то учитывая, что $k' = n'(\omega)\omega/c$, получаем $v_p = c/n'(\omega)$, т.е. *фазовая скорость света в среде отличается от скорости света в вакууме и зависит от частоты*.

Перепишем формулу (10.16), подставив в нее выражение для $k' = k_0(1 + \Delta n)$ ($k_0 = \omega/c$ — волновое число в вакууме)

$$\mathcal{E} = \frac{1}{2} E e^{-k''z} e^{-i\Delta n k_0 z} \exp[i(\omega t - k_0 z)]. \quad (10.18)$$

Отсюда следует, что *среда приводит не только к затуханию амплитуды света, но и к задержке по фазе* падающего на нее светового поля. Фаза запаздывает на величину $\Delta n k_0 z$, поскольку в экспоненте стоит знак минус.

Такой же вывод вытекает и из того факта, что *поляризованность* среды является *комплексной величиной* (см. формулы (10.6) и (10.7)). Это означает, что *колебания поляризованности происходят со сдвигом (задержкой) фазы, зависящей от частоты*. Колеблющиеся диполи вещества сами становятся источниками излучения на той же частоте ω , но со сдвигом фазы. Переизлученные волны, складываясь (интерферируя) с первичной волной, образуют волну, распространяющуюся в среде. *Отставание по фазе вторичных волн по отношению к первичной волне* (зависящее от частоты), приводит к тому, что *фазовая скорость волны в среде отличается от скорости света в вакууме и становится зависящей от частоты*.

Вблизи резонанса при $|\omega_0 - \omega| \ll \omega_0$, и, следовательно, $\omega_0^2 - \omega^2 \approx 2\omega_0(\omega_0 - \omega)$, выражения для n' , n'' (10.13) и (10.14) упрощаются

$$n'(\omega) = 1 + \frac{\omega_p^2}{4\omega_0} \frac{(\omega_0 - \omega)}{(\omega_0 - \omega)^2 + \gamma^2/4}. \quad (10.19)$$

$$n''(\omega) = \frac{\omega_p^2}{4\omega_0} \frac{\gamma}{(\omega_0 - \omega)^2 + \gamma^2/4}. \quad (10.20)$$

Формулы (10.13), (10.14), (10.19), (10.20), полученные с помощью модели Лоренца, описывают *дисперсию и поглощение света в разреженной линейной изотропной среде*.

Как правило в оптике $\gamma \ll \omega_0$. Типичные зависимости $n'(\omega)$ и $\alpha(\omega)$ показаны на **Рис.10.1**. Зависимость показателя поглощения от частоты называют *спектральным контуром линии поглощения*. В рассматриваемой модели среды, состоящей из неподвижных затухающих атомных осцилляторов, она имеет *лоренцевскую форму* с максимумом при $\omega = \omega_0$. Нетрудно убедиться, что спектральная линия поглощения имеет ширину (на полувысоте), равную γ .

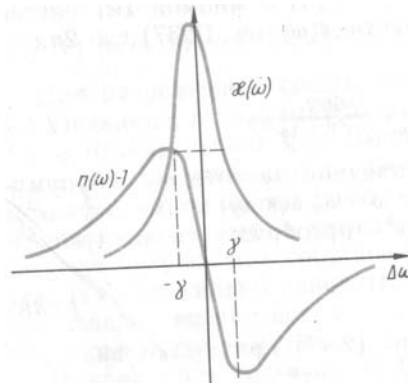


Рис. 10.1. Дисперсионная кривая и контур поглощения

График зависимости показателя преломления от частоты называют *дисперсионной кривой*. Показатель преломления n' заметно отличается от единицы лишь вблизи полосы поглощения, причем максимальное и минимальное значения он достигает на частоте $\omega = \omega_0 \pm \gamma/2$, соответственно.

В зависимости от знака производной $\partial n/\partial \omega$ выделяют две области частот: область *нормальной дисперсии*, где показатель преломления возрастает с ростом частоты света ($\partial n/\partial \omega > 0$), и область *аномальной дисперсии*, где показатель преломления уменьшается с ростом частоты ($\partial n/\partial \omega < 0$). Как видно из **Рис.10.1**, область *нормальной дисперсии* расположена за пределами полосы поглощения, т.е. в области *прозрачности вещества*. Здесь $|\omega_0 - \omega| \gg \gamma$, и для показателя преломления, как это следует из (10.13), можно записать

$$n = 1 + \frac{1}{2} \frac{\omega_p^2}{\omega_0^2 - \omega^2} \quad (10.21)$$

Значительно более узкая область *аномальной дисперсии* расположена *вблизи центра линии поглощения света*.

Обратим внимание, что при $\omega > \omega_0$ показатель преломления больше единицы, и в этом случае $v_p < c$. При $\omega > \omega_0$ *показатель преломления становится меньше единицы* ($n < 1$), а значит *скорость распространения света в веществе больше скорости света в вакууме* c . Но это не противоречит теории относительности, согласно которой скорость любых тел или сигналов не может превышать скорости света в вакууме, что не имеет никакого отношения к фазовой скорости монохроматической волны. Физическая причина того, что фазовые скорости могут быть больше c , состоит в следующем. Как показывает анализ, при $\omega > \omega_0$ смещение заряда находится в противофазе с внешней силой $e\vec{E}(t)$, а поле, образованное смещением заряда, усиливает внешнее поле $\vec{E}(t)$, что в свою очередь приводит к увеличению возвращающей силы, и фазовая скорость будет больше скорости света в вакууме.

Изложенные результаты справедливы не только для *электронов*, но и для *ионов*. В последнем случае под m , e и N следует понимать массу, заряд и концентрацию *ионов*. Отметим, что заряд и масса ионов могут существенно отличаться от заряда и массы электрона (особенно масса иона), поэтому резонансные частот также могут сильно отличаться от резонансных частот атомов.

Опыт показывает, что обычно наблюдается не одна, а несколько линий поглощения даже для одноэлектронных атомов. Чтобы это учесть, в классической модели принимается, что вещество построено из частиц различного типа — электронов и

ионов, которые ведут себя как затухающие осцилляторы с различными *собственными частотами*. В газах их взаимодействием можно пренебречь. В этом случае формула (10.11) принимает вид

$$\epsilon(\omega) = 1 + \frac{4\pi Ne^2}{m} \sum_k \frac{f_k}{\omega_0^2 - \omega^2 + i\omega\gamma}. \quad (10.22)$$

Величины f_k , называемые *силами осцилляторов*, удовлетворяют *правилу сумм* $\sum_k f_k = 1$.

10.2. Распространение светового импульса в диспергирующей среде, групповая скорость. До сих пор, рассматривая распространение света в среде, мы обсуждали плоские монохроматические волны с определенной частотой и волновым вектором. Однако в большинстве случаев излучение не является монохроматическим и имеет вид *импульсов*. Ограниченная длительность импульса приводит к существованию некоторой конечной полосы частот или длин волн. Всякая среда обычно характеризуется дисперсией, т.е. монохроматические волны *распространяются* с различными фазовыми скоростями, зависящими от частоты. В результате различные спектральные компоненты при распространении через среду приобретают разные фазовые набеги. Так как результат суммирования колебаний зависит от распределения фаз, то, очевидно, что форма импульса на выходе из среды, вообще говоря, отличается от формы импульса на входе. Другим словами, в процессе распространения в диспергирующей среде световой импульс *деформируется*. Кроме того скорость переноса энергии световым импульсом, распространяющимся в диспергирующей среде, существенно может отличаться от фазовой скорости.

Если *дисперсия не велика*, то деформация формы импульса происходит медленно, и мы можем следить за перемещением отдельной амплитуды поля в импульсе, например, максимальной амплитуды поля, и определить скорость, с которой она переносится. Определенную таким образом скорость, согласно Рэлею, называют *скоростью перемещения импульса* или *групповой скоростью*. Таким образом, *групповая скорость — это скорость движения огибающей волнового пакета или группы волн в диспергирующей среде*. В общем случае она отличается от фазовой скорости любой из составляющих его монохроматических спектральных компонент.

Рассмотрим понятие групповой скорости на примере двух близких по частоте плоских монохроматических волн, распространяющихся в диспергирующей среде вдоль оси z

$$E_{1,2}(z, t) = A \sin(\omega_{1,2}t - k_{1,2}z). \quad (10.23)$$

Здесь $\omega_{1,2}$ и $k_{1,2} = k_{1,2}(\omega_{1,2})$ — частоты и соответствующие им волновые векторы; Суммарное колебание имеет вид

$$E(z, t) = E_1(z, t) + E_2(z, t) = 2A \cos(\Delta\omega t - \Delta k z) \sin(\omega_0 t - k_0 z), \quad (10.24)$$

где $\omega_0 = (\omega_2 + \omega_1)/2$; $k_0 = (k_1 + k_2)/2$; $\Delta k = (k_1 - k_2)/2$; $\Delta\omega = (\omega_2 - \omega_1)/2$; $|\Delta k| \ll k_0$, $|\Delta\omega| \ll \omega_0$.

Волна, описываемая формулой (10.24) показана на **Рис.10.2**. Она представляет собой колебания на средней (несущей) частоте ω_0 , промодулированные с частотой $\Delta\omega \ll \omega_0$. Его амплитуда $a = 2A \cos(\Delta\omega t - \Delta k z)$ (или огибающая) меняется во времени и в пространстве *медленно* по сравнению с несущей частотой ω_0 . Такие колебания называются *биениями*.

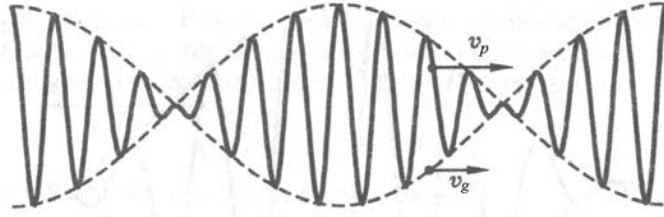


Рис. 10.2. Группа волн, представляющих суперпозицию двух близких по частоте монохроматических волн

Для нахождения *групповой скорости* v_g запишем условие постоянства амплитуды

$$\Delta\omega t - \Delta k z = \text{const} \quad (10.25)$$

Дифференцируем по t : $\Delta\omega - \Delta k \cdot dz/dt = 0$. Откуда для групповой скорости $v_g = dz/dt$ получаем следующее выражение

$$v_g = \frac{\Delta\omega}{\Delta k} = \frac{d\omega}{dk} \quad (10.26)$$

Используя связь волнового числа с показателем преломления $k(\omega) = (\omega/c)n(\omega)$, для групповой скорости можно записать

$$v_g = \frac{v_p}{1 + \frac{\omega}{n} \frac{dn}{d\omega}}. \quad (10.27)$$

Поскольку понятие групповой скорости применимо к областям нормальной дисперсии, где $dn/d\omega > 0$, то из (10.27) следует, что $v_g < v_f$. Это означает, что модуляция световой волны перемещается медленнее, чем несущая. Различие между фазовой и групповой скоростью тем больше, чем больше дисперсия $dn/d\omega$. В отсутствие дисперсии $v_f = v_g$. Это строго выполняется лишь для вакуума, где свет на любой частоте распространяется со скоростью c .

Понятие групповой скорости, введенное на примере двух волн, остается справедливым и в случае, когда мы имеем дело с импульсом, состоящим из большого числа спектральных составляющих. Действительно, всякий импульс $E(z, t)$ можно представить в виде Фурье разложения

$$E(z, t) = \int_{-\infty}^{\infty} A(k) e^{i[\omega(k)t - kz]} dk, \quad (10.28)$$

где $A(k)$ — амплитуда плосковолновой составляющей с волновым числом k (Фурье-образ $E(z, t)$). Согласно принципу суперпозиции, интеграл (10.28) удовлетворяет уравнениям Максвелла, так как плоские волны являются их решением. Параметры ω и k связаны дисперсионным соотношением $k = n(\omega)\omega/c$. В изотропной среде $\omega(-k) = \omega(k)$.

Световой импульс характеризуют центральной частотой ω_0 (или волновым числом k_0) и шириной полосы частот $\Delta\omega$ с центром ω_0 . Рассмотрим эволюцию такого импульса во времени. Разложим $\omega(k)$ в ряд Тейлора в окрестности k_0

$$\omega(k) = \omega_0 + \left(\frac{d\omega}{dk} \right)_0 (k - k_0) + \dots \quad (10.29)$$

Здесь мы ограничились *линейным* членом разложения, предполагая, что импульс является узкополосным, т.е. $\Delta\omega \ll \omega_0$. На *временном* языке это означает, что длительность импульса много больше периода световых колебаний.

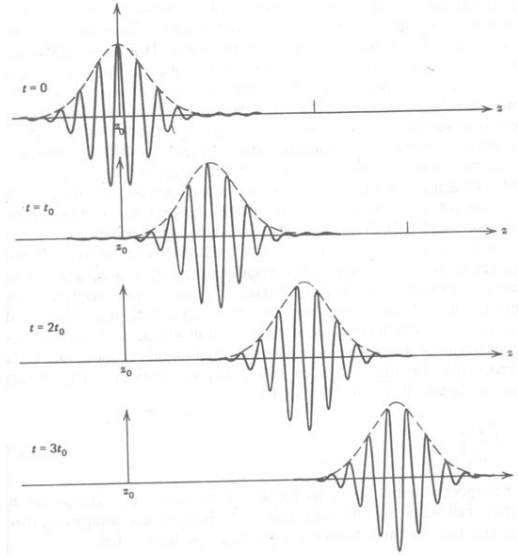


Рис. 10.3.

Подставляя (10.29) в (10.28), получаем

$$E(z, t) = e^{i(\omega_0 t - k_0 z)} \int_{-\infty}^{\infty} A(k) \exp \left[i \left[\left(\frac{d\omega}{dk} \right)_0 t - z \right] (k - k_0) \right] dk = \quad (10.30)$$

$$e^{i(\omega_0 t - k_0 z)} E[z - (d\omega/dk)_0 t].$$

Здесь $E[z - (d\omega/dk)_0 t]$ — огибающая импульса.

Отсюда следует, что импульс распространяется с *сохранением своей формы* (Рис.10.3) с групповой скоростью $v_g = (d\omega/dk)_0$. Это приближение имеет место, когда $A(k)$ имеет ярко выраженный пик при k_0 , а частота является медленно меняющейся функцией k . В противном случае необходимо учитывать следующие члены разложения в (10.29), но при этом форма импульса при распространении уже не будет сохраняться.

Итак, *огибающая импульса перемещается с групповой скоростью*, при этом с течением времени ее *профиль остается неизменным*. Отметим, что понятие групповой скорости не является универсальным. Оно имеет смысл лишь для узких световых сигналов, распространяющихся в прозрачных средах. В резонансных условиях, когда импульс сильно искажается, понятие групповой скорости не применимо. Более тонкий анализ показывает, что в отличие от фазовой скорости волны, характеризующей скорость смещения точки волны, имеющей фиксированное значение фазы, *групповая скорость отражает скорость переноса энергии в волне*, т.е. скорость передачи взаимодействия.

10.3. Рассеяние света в мутных средах. Известно, что в вакууме свет распространяется прямолинейно. Это связано с взаимным интерференционным гашением вторичных сферических волн, идущих от разных точек пространства, согласно принципа Гюйгенса. Такое гашение имеет место для всех направлений, кроме прямого, что и обуславливает прямолинейность светового пучка в вакууме. Аналогичным образом в однородной среде сферические световые волны, испускаемые отдельными атомами, интерферируют так, что световой луч оказывается прямолинейным.

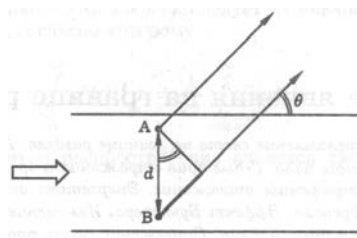


Рис. 10.4. К объяснению механизма прямолинейного распространения света в однородной среде

Механизм подавления бокового свечения поясняет **Рис.10.4**. Пусть А некоторый атом среды расположен в пределах светового пучка. Под действием падающей световой волны оптический электрон атома начинает совершать вынужденные колебания. в результате чего атом становится источником волн. Рассмотрим излучение атома, идущее под углом θ к оси светового пучка. Как видно из рисунка, для любого угла θ кроме $\theta = 0$, найдется такой атом В, который в том же самом направлении испускает волну, находящуюся в противофазе к волне первого атома. В самом деле, если, например, падающее излучение имеет структуру плоской волны, то атом В находится в том же самом поперечном сечении пучка на расстоянии d от атома А, определяемом условием $d \sin \theta = \lambda/2$, где λ — длина световой волны. Таким образом, атомы А и В взаимно гасят излучение друг друга в направлении θ . Поскольку аналогичную пару атомов можно подобрать для любого угла $\theta \neq 0$, приходим к выводу, что свет распространяется прямолинейно. Такое уничтожение возможно лишь при *одинаковой* амплитуде всех вторичных волн. А это достигается в *однородной* с точностью до λ среде.

Можно ожидать, что в *неоднородной* среде луч света не будет прямолинейным, т.е. будет переизлучение света или, как говорят, *рассеяние* назад и в стороны, так как амплитуды и фазы вторичных волн в неоднородной среде различны. Это возможно в *мутных* средах с размерами частиц $d > \lambda$ и в чистых с размером *неоднородности плотности* с такими же размерами. Действительно, опыт показывает, что для *неоднородных сред характерно сильное рассеяние света*, т.е. свет попав в такую среду переизлучается в разные стороны. Именно по этой причине, например, задымленный воздух и туман выглядят "мутными".

Можно показать, что интенсивность рассеянного света прямо пропорциональна четвертой степени частоты внешнего воздействия или обратно пропорциональна четвертой степени длины волны.

$$P \sim \omega^4 \sim \frac{1}{\lambda^4} \quad (10.31)$$

Выражение (10.31) называется законом Рэлея.

Именно этим законом объясняется голубой цвет неба. Солнечный свет проходит сквозь воздух, и, когда мы смотрим в сторону от Солнца, например, перпендикулярно падающему лучу, мы видим свет голубой окраски. Так как $\lambda_b/\lambda_r = 4500/6500 = 1.44$, то интенсивность рассеиваемого синего света примерно в 4 раза больше, чем красного. Кроме того, солнечное излучение, рассеянное молекулами воздуха, *поляризовано*. В этом можно убедиться, если посмотреть в солнечный день на голубое небо через поляриод. Вращая поляриод, вы увидите как возникают темные полосы, возникновение которых и свидетельствует поляризации рассеянного света.

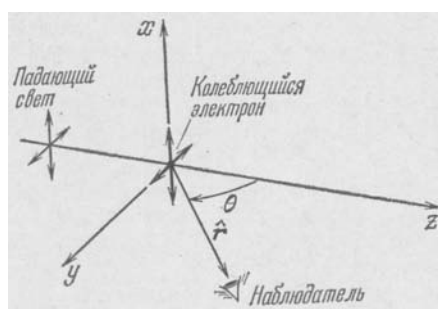


Рис. 10.5. Поляризация света при рассеянии. Наблюдатель видит полную проекцию движения электрона на ось $\vec{x}$ и лишь часть проекции на ось $\vec{y}$, пропорциональную $\cos \theta$. При $\theta = \pi/2$ рассеянное излучение полностью поляризовано по оси $\vec{x}$.

Объяснение поляризации голубого неба заключается в следующем. Пусть $\vec{z}$ — направление распространения света от Солнца до данной молекулы воздуха (**Рис.10.5**). Электрическое поле в солнечном излучении не поляризовано. Электроны в молекуле ведут себя подобно осциллятору, находящемуся под действием падающего света. Поэтому их колебания представляют суперпозицию движений $\vec{x}$ и $\vec{y}$. Колеблющиеся электроны излучают во всех направлениях, но их излучение неодинаково для разных направлений. Амплитуда и направление поляризации электрического поля, излучаемого отдельным колеблющимся точечным зарядом, зависят от проекции амплитуды движения колеблющегося заряда, представляющую собой амплитуду той компоненты движения электрона, которая перпендикулярна направлению распространения $\vec{z}$, т.е. направлению от электрона к наблюдателю. Если вектор $\vec{r}$ направлен по $\vec{y}$, то наблюдатель видит только x -компоненту движения электрона. Поэтому наблюдаемое им излучение будет полностью поляризованным $\vec{x}$. В этом случае интенсивность излучения равна половине интенсивности излучения в направлении $\vec{z}$. Подчеркнем, что линейно-поляризованное излучение, наблюдаемое под соответствующим углом, имеет место при единичном рассеянии. При многократном рассеянии происходит деполяризация излучения, т.е. рассеянный становится неполяризованным.

Интересно отметить, что пчелы реагируют на поляризацию излучения. Информацию о поляризации неба они используют для ориентации.

ЛЕКЦИЯ №11

Оптические явления на границе раздела сред: отражение и преломление поляризованного света на границе раздела; полное внутреннее отражение; формулы Френеля; эффект Брюстера. Отражение света от поверхности металла.

11.1. Отражение и преломление поляризованного света на границе раздела. Хорошо известно, что луч света, падающий на границу раздела двух различных оптических сред, разделяется на два луча: *отраженный* и *преломленный* (**Рис.11.1**), направления которых хорошо описываются в терминах геометрической оптикой. Однако такие свойства как *поляризация, интенсивность отраженного и преломленного света* можно понять только на основе электромагнитной теории света. А именно эти свойства отражения и преломления находят широкое применение в различных устройствах — линзах, призмах, зеркалах, которые позволяют фокусировать свет и формировать оптические изображения, осуществлять спектральное разложение света, получать полностью поляризованный свет и т. п. Явление отражения света используется в лазерных резонаторах и интерферометрах.

Количественная теория отражения и преломления света строится на основе уравнений Максвелла и граничных условий для электромагнитного поля. Как известно из электродинамики, на *границе раздела диэлектриков* должны быть **непрерывны тангенциальные** (т. е. параллельные поверхности раздела) *компоненты напряженности электрического и магнитного поля*, т.е.

$$\mathcal{E}_{1t} = \mathcal{E}_{2t}, \quad \mathcal{H}_{1t} = \mathcal{H}_{2t} \quad (11.1)$$

Здесь индекс "t" обозначает тангенциальную компоненту, цифры "1" и "2" обозначают среды по разные стороны границы раздела.

Пусть на плоскую границу раздела двух сред (**Рис.11.1**) падает плоская монохроматическая линейно поляризованная световая волна

$$\mathcal{E}_1 = \frac{1}{2} E_1 \exp [i(\omega_1 t - \vec{k}_1 \vec{r})] + \text{к.с} \quad (11.2)$$

Световые поля *отраженной* и *преломленной* волн в общем случае имеют вид

$$\mathcal{E}_{0,2} = \frac{1}{2} E_{0,2} \exp [i(\omega_{0,2} t - \vec{k}_{0,2} \vec{r})] + \text{к.с}, \quad (11.3)$$

где индексы 0 и 1 соответствуют отраженной и преломленной волне, соответственно. Угол θ_1 между нормалью к границе раздела $\vec{n}$ и волновым вектором $\vec{k}$ падающей

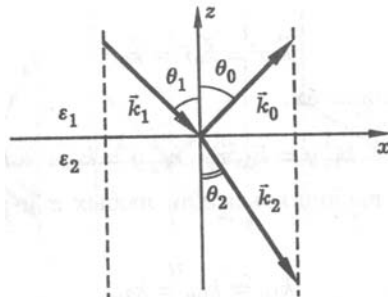


Рис. 11.1. Отражение и преломление плоской волны на границе раздела двух сред

называют *углом падения* (**Рис.11.1**); углы $\theta_{0,2}$ — *углом отражения* и *углом преломления*; $\varepsilon_{1,2}$ — комплексные диэлектрические проницаемости сред.

Направим ось z по нормали к границе раздела, ось x вдоль границы раздела, а y перпендикулярно плоскости рисунка. Граничное условие (11.1) в плоскости границы раздела $z = 0$, запишем в виде:

$$E_{1t} \exp[i(\omega t - \vec{k}_1 \vec{r})] + E_{0t} \exp[i(\omega_0 t - \vec{k}_0 \vec{r})] = E_{2t} \exp[i(\omega_2 t - \vec{k}_2 \vec{r})], \quad (11.4)$$

Соотношение (11.4) должно выполняться в *любой* момент времени t и для *всех* точек $\vec{r}$, лежащих на границе раздела. Поэтому частоты отраженной и преломленной волн должны быть равны частоте ω падающей волны. Отсюда следует, что

$$\omega t - \vec{k}_1 \vec{r} = \omega t - \vec{k}_0 \vec{r} = \omega t - \vec{k}_2 \vec{r} \quad \text{или} \quad \vec{k}_1 \vec{r} = \vec{k}_0 \vec{r} = \vec{k}_2 \vec{r}. \quad (11.5)$$

Переходя к декартовым координатам, перепишем второе соотношение (11.5)

$$k_{1x}x + k_{1y}y = k_{0x}x + k_{0y}y = k_{2x}x + k_{2y}y. \quad (11.6)$$

Эти условия должны выполняться для *любых* x и y . В частности, полагая $x = 0$ и $y = 0$, получим, соответственно

$$k_{1y} = k_{0y} = k_{2y}, \quad k_{1x} = k_{0x} = k_{2x} \quad (11.7)$$

Отсюда следует, что *волновые векторы отраженной и преломленной волн лежат в одной плоскости, проведенной через волновой вектор падающей волны и нормалью к границе раздела*, называемую *плоскостью падения*.

Пусть плоскость xz совпадает с плоскостью падения (**Рис.11.1**). Тогда

$$k_{1x} = k_{0x} = k_{2x}, \quad k_{1y} = k_{0y} = k_{2y} = 0, \quad (11.8)$$

где

$$k_{1x} = k_1 \sin \theta_1, \quad k_{0x} = k_0 \sin \theta_0, \quad k_{2x} = k_2 \sin \theta_2, \quad (11.9)$$

$$k_1 = k_0 = \frac{\omega}{c} n_1, \quad k_2 = \frac{\omega}{c} n_2, \quad (11.10)$$

$n_{1,2}$ — комплексные показатели преломления первой и второй сред.

Так как $k_{1x} = k_{0x}$, то из первых двух формул (11.9) следует, что

$$\theta_0 = \theta_1, \quad (11.11)$$

т.е. *угол падения равен углу отражения* — **закон отражения** света.

Предположим теперь, что обе среды *прозрачны*, т.е. показатели преломления n_1 и n_2 действительны. Подставляя k_1 и k_2 из (11.10) в выражения для k_{1x} и k_{2x} (11.9) и учитывая, что $k_{1x} = k_{2x}$, получаем

$$n_1 \sin \theta_1 = n_2 \sin \theta_2 \quad \text{или} \quad \frac{\sin \theta_1}{\sin \theta_2} = \frac{n_2}{n_1} \quad (11.12)$$

Эти формулы связывает между собой угол падения и преломления света. Они выражают *закон преломления*, или *закон Снеллиуса*.

Подчеркнем, что *отклонение луча света есть следствие изменения скорости света в разных материалах*. Этот вывод поясняет **Рис.11.2**, где показаны ряд последовательных максимумов в амплитуде волны, падающей из вакуума на стекло. Стрелки, перпендикулярные указанным максимумам, отмечают направление распространения волны. Всюду в волне происходят колебания с одной и той же частотой. Поэтому расстояния между максимумами по обе стороны совпадают вдоль самой

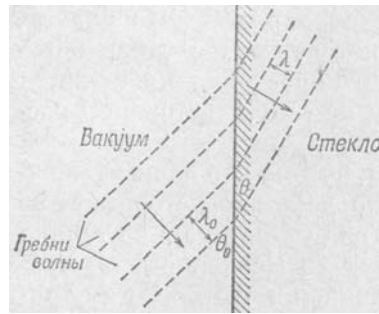


Рис. 11.2. Связь между преломлением и изменением скорости света

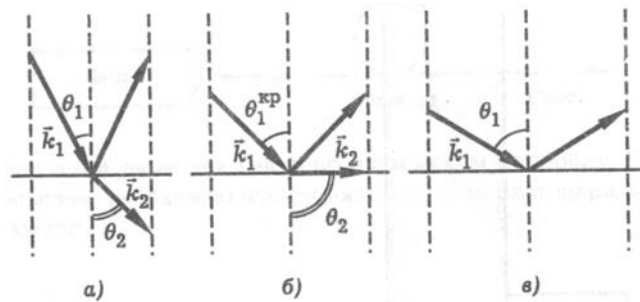


Рис. 11.3. Отражение и преломление света при переходе луча из оптически более плотной среды в оптически менее плотную ($n_1 > n_2$). а) Угол падения меньше критического, б) угол падения равен критическому, в) угол падения больше критического. В последнем случае преломленный луч отсутствует и имеет место эффект полного внутреннего отражения света

поверхности, так как волны здесь должны быть согласованы и заряд на поверхности колеблется с той же самой частотой. Наименьшее расстояние между гребнями есть длина волны λ . В вакууме $\lambda_0 = 2\pi c/\omega$, а в среде с показателем преломления $\lambda = 2\pi v/\omega$ или $\lambda = 2\pi c/\omega n$, где $v = c/n$ — скорость световой волны в среде. Чтобы "сшить" волны на границе волна в среде должна изменить свое направление. Причем это возможно при условии, что $\lambda_0/\sin \theta_0 = \lambda/\sin \theta$ или $\sin \theta_0/\sin \theta = n$, а это и есть закон Снеллиуса.

Полное внутреннее отражение. Рассмотрим случай, когда световой луч пересекает границу раздела, переходя из *оптически более плотной среды в менее плотную*, т.е. $n_1 > n_2$. Согласно закону Снеллиуса, угол преломления будет больше угла падения $\theta_2 > \theta_1$. Картина преломления для этого случая показана на **Рис.11.3 а**. Если угол падения составляет величину

$$\theta_1^{cr} = \arcsin(n_2/n_1), \quad (11.13)$$

то по формуле (11.12) получаем $\theta_2 = \pi/2$ и, следовательно, преломленный луч не проникает во вторую среду, а распространяется вдоль границы раздела (**Рис.11.3 б**). Если же $\theta_1 > \theta_1^{cr}$, т.е. угол падения превышает критический, то преломленный луч отсутствует, и свет полностью отражается от границы раздела (**Рис.11.3 в**). Этот эффект называется *полным внутренним отражением света*. Угол падения, определяемый формулой (11.13), называется *предельным углом полного отражения*.

Например, для границы стекло-воздух ($n_1 = 1.5$, $n_2 = 1$) этот угол составляет около 40° .

Полное внутреннее отражение позволяет передавать свет на большие расстояния по оптическому волокну.

Формулы Френеля. Определим теперь соотношение между энергиями отраженного и преломленного лучей. Для этого необходимо вычислить *амплитуды отраженной и преломленной волн*. Будем считать падающую, отраженную и преломленную световые волны плоскими и монохроматическими (смотри формулы (11.2)-(11.3), где надо считать $\omega = \omega_0 = \omega_2$). Среда, образующие границу раздела, также будем считать линейными и изотропными и характеризовать комплексными диэлектрическими проницаемостями $\varepsilon_{1,2}$.

Пусть падающая, отраженная и преломленная волны являются плоскими монохроматическими волнами

$$\mathcal{E} = \frac{1}{2} \vec{E} \exp[i(\omega t - \vec{k} \vec{r})] + \text{к.с.}, \quad \mathcal{H} = \frac{1}{2} \vec{H} \exp[i(\omega t - \vec{k} \vec{r})] + \text{к.с.}, \quad (11.14)$$

Уравнения Максвелла имеют вид:

$$[\vec{k}, \vec{E}] = \frac{\omega}{c} \vec{H}, \quad [\vec{k}, \vec{H}] = -\frac{\omega}{c} \varepsilon \vec{E}, \quad (11.15)$$

где $\vec{E}$, $\vec{H}$ — комплексные амплитуды электрического и магнитного поля.

Величина тангенциальной компоненты поля зависит от состояния поляризации падающей световой волны. Рассмотрим отдельно случаи, когда вектор $\vec{E}$ *перпендикулярен* и *параллелен* плоскости падения, поскольку волну с произвольной поляризацией можно представить в виде суперпозиции двух линейно поляризованных волн указанных типов $\vec{E} = \vec{E}_\perp + \vec{E}_\parallel$.

Пусть вектор $\vec{E}_1$ падающей волны *перпендикулярен* плоскости падения ($\vec{E}_1 = \vec{E}_{1\perp}$). Такую поляризацию волны называют "s" поляризацией. В этом случае *отраженная и преломленная волны* будут также *поляризованы перпендикулярно плоскости падения*. Векторы $\vec{H}$ для *магнитного поля* всех трех волн будут лежать в *плоскости падения*. Используя систему координат, показанную на **Рис.11.1**, можно записать

$$\mathcal{E}_t = \mathcal{E}_y = \mathcal{E}, \quad \mathcal{H}_t = \mathcal{H}_x. \quad (11.16)$$

В принятых обозначениях граничные условия (11.1) принимают вид

$$\mathcal{E}_1 + \mathcal{E}_0 = \mathcal{E}_2, \quad \mathcal{H}_{1x} + \mathcal{H}_{0x} = \mathcal{H}_{2x}. \quad (11.17)$$

Подставив (11.14) в (11.17), получим следующие уравнения для комплексных амплитуд

$$E_1 + E_0 = E_2, \quad H_{1x} + H_{0x} = H_{2x}. \quad (11.18)$$

Выразим H_{0x} , H_{1x} через $E_{0,1}$, соответственно, а H_{2x} через E_2 , используя первое соотношение в (11.15)

$$H_{0,1x} = -\frac{c}{\omega} k_z E_{0,1}, \quad H_{2x} = -\frac{c}{\omega} k_z E_2. \quad (11.19)$$

Подставляя (11.19) во вторую формулу (11.18) и учитывая, что $k_{0z} = -k_{1z}$, получаем

$$k_{1z}(E_1 - E_0) = k_{2z}E_2 \quad (11.20)$$

Введем *комплексные коэффициенты отражения* ($r_\perp$) и *преломления* ($t_\perp$) света

$$r_\perp = E_0/E_1, \quad t_\perp = E_2/E_1 \quad (11.21)$$

Используя (11.21), первое соотношение (11.18) и (11.20) можно переписать в виде

$$1 + r_{\perp} = t_{\perp}, \quad k_{1z}(1 - r_{\perp}) = k_{2z}t_{\perp} \quad (11.22)$$

Решая алгебраическую систему уравнений (11.22), получаем

$$r_{\perp} = \frac{k_{1z} - k_{2z}}{k_{1z} + k_{2z}}, \quad t_{\perp} = \frac{2k_{1z}}{k_{1z} + k_{2z}} \quad (11.23)$$

Формулы (11.23) устанавливают связь между амплитудами падающей, отраженной и преломленной световых волн.

Аналогичным образом рассматривается случай, когда вектор падающей световой волны *параллелен* плоскости падения $\vec{E} = \vec{E}_{\parallel}$ — ”*p*” поляризация. В этом случае граничные условия (11.1) имеют вид

$$H_1 + H_0 = H_2, \quad E_{1x} + E_{0x} = E_{2x} \quad (11.24)$$

Выразим E_x через H , используя второе соотношение в (11.15)

$$E_x = \frac{c}{\omega} \frac{k_z}{\varepsilon} H. \quad (11.25)$$

Подставляя (11.25) во второе соотношение (11.24), получим уравнение

$$\frac{k_{1z}}{\varepsilon_1}(H_1 - H_0) = \frac{k_{2z}}{\varepsilon_2} H_2 \quad (11.26)$$

Определим *коэффициенты отражения и преломления* формулами $r_{\parallel} = H_0/H_1$ и $t_{\parallel} = H_2/H_1$. Используя это, первое соотношение (11.24) и (11.26) можно представить в виде

$$1 + r_{\parallel} = t_{\parallel}, \quad 1 - r_{\parallel} = \frac{k_{2z}\varepsilon_1}{k_{1z}\varepsilon_2} t_{\parallel}. \quad (11.27)$$

Из (11.27) несложно определить искомые величины

$$r_{\parallel} = \frac{(k_{1z}/\varepsilon_1) - (k_{2z}/\varepsilon_2)}{(k_{1z}/\varepsilon_1) + (k_{2z}/\varepsilon_2)}, \quad t_{\parallel} = \frac{2k_{1z}/\varepsilon_1}{(k_{1z}/\varepsilon_1) + (k_{2z}/\varepsilon_2)} \quad (11.28)$$

Предположим теперь, что обе среды *прозрачны*, т.е. проницаемости $\varepsilon_{1,2}$, а также показатели преломления сред $n_{1,2} = \sqrt{\varepsilon_{1,2}}$ действительные величины. Нормальные компоненты волновых векторов падающей и преломленной волн выражаются формулами (**Рис.11.1**): $k_{1z} = -k_1 \cos \theta_1$, $k_{2z} = -k_2 \cos \theta_2$, где волновые числа $k_{1,2} = (\omega/c)n_{1,2}$. Используя это для коэффициентов отражения $r_{\perp}$ и $r_{\parallel}$ можно записать

$$r_{\perp} = \frac{n_1 \cos \theta_1 - n_2 \cos \theta_2}{n_1 \cos \theta_1 + n_2 \cos \theta_2}, \quad r_{\parallel} = \frac{n_2 \cos \theta_1 - n_1 \cos \theta_2}{n_2 \cos \theta_1 + n_1 \cos \theta_2}. \quad (11.29)$$

Принимая во внимание закон Снеллиуса (11.12), эти выражения нетрудно преобразовать к виду

$$r_{\perp} = \frac{\sin(\theta_1 - \theta_2)}{\sin(\theta_1 + \theta_2)}, \quad r_{\parallel} = \frac{\operatorname{tg}(\theta_1 - \theta_2)}{\operatorname{tg}(\theta_1 + \theta_2)} \quad (11.30)$$

Формулы (11.30) определяют *амплитудные коэффициенты отражения света* на границе раздела линейных изотропных *прозрачных* сред. Их называются *формулами Френеля*.

Из закона сохранения энергии с очевидностью следует, что

$$r_{\perp}^2 + t_{\perp}^2 = 1, \quad r_{\parallel}^2 + t_{\parallel}^2 = 1 \quad (11.31)$$

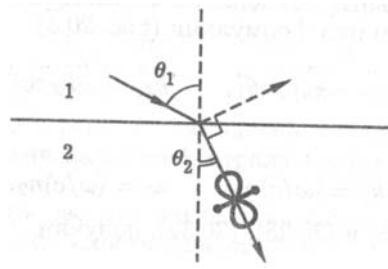


Рис. 11.4. К пояснению эффекта Брюстера

Эффект Брюстера. Из второй формулы (11.30) следует, что при определенных условиях коэффициент отражения $r_{\parallel}$ обращается в ноль, когда знаменатель стремится к бесконечности. Это имеет место, если сумма углов падения и преломления удовлетворяет условию:

$$\theta_1 + \theta_2 = \pi/2 \quad (11.32)$$

Таким образом, если *падающая световая волна поляризована в плоскости падения* и выполняется соотношение (11.32), то *отраженная волна отсутствует*. Этот эффект называют *эффектом Брюстера*. Угол, при котором это происходит, называют *углом Брюстера*. Используя (11.32) и закон преломления, нетрудно вычислить величину этого угла.

$$\theta_B = \arctan(n_2/n_1) \quad (11.33)$$

Например, для границы раздела "воздух-стекло" $\theta_B = 56^\circ 40'$.

Брюстеровский угол называют также *углом полной поляризации*. Действительно, если падающий под этим углом свет неполяризован, то отраженный пучок света линейно поляризован перпендикулярно плоскости падения. Таким образом, *эффект Брюстера можно использовать для получения линейно поляризованного света*. Это явление объясняется следующим образом. Отраженный луч света представляет собой переизлучение диполей второй среды. Из **Рис.11.4** видно, что при выполнении условия (11.32) отраженный и преломленный лучи оказываются взаимно перпендикулярными. При этом диполи второй среды, колеблющиеся параллельно вектору $\vec{E}$ преломленной волны и, следовательно, перпендикулярно преломленному лучу, не испускают свет в направлении отраженного луча, так как это направление совпадает с направлением колебаний диполей. В результате отраженный луч отсутствует, и вся энергия света передается преломленному лучу. Отсюда следует, что эффект Брюстера возможен лишь при поляризации падающего луча в плоскости падения. Если же падающий луч поляризован перпендикулярно плоскости падения, то отраженный луч должен наблюдаться при любом угле падения.

Изменение фазы световой волны при отражении и преломлении. Из формул для коэффициентов отражения и пропускания (11.23) и (11.28) следует, что если обе среды прозрачны, то амплитудные коэффициенты пропускания действительные положительные числа. Это значит, что фаза преломленной волны совпадает с фазой падающей волны на границе раздела.

Амплитудные коэффициенты отражения также действительны, но могут быть как *положительными*, так и *отрицательными*. Следовательно, в зависимости от условий, фаза отраженной волны может либо *совпадать* с фазой падающей волны на границе раздела, либо *отличаться* от нее на π .

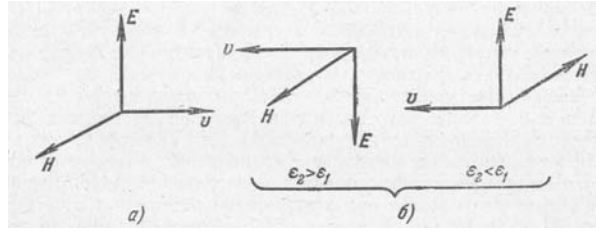


Рис. 11.5. Расположение векторов $\vec{E}$, $\vec{H}$ и $\vec{v}$ в падающей (а) и отраженной волнах

Если $\theta_1 = 0$, то несложно получить, что

$$r_{\perp} = -\frac{n_2 - n_1}{n_2 + n_1}, \quad r_{\parallel} = \frac{n_2 - n_1}{n_2 + n_1}, \quad (11.34)$$

т.е., при нормальном падении $r_{\perp} = -r_{\parallel}$. Смысл этого результата состоит в том, что при отражении назад один из векторов $\vec{E}$ или $\vec{H}$ световой волны должен изменить свое направление на обратное; только в этом случае векторы $\vec{E}$, $\vec{H}$ и $\vec{k}$ как в падающей, так и в отраженной волне образуют правую тройку векторов в соответствии с требованием уравнений Максвелла. При $n_2 > n_1$ можно показать, что $r_{\perp} < 0$, $r_{\parallel} > 0$ и, следовательно, при отражении вектор $\vec{H}$ сохраняет свое направление, а вектор $\vec{E}$ меняет направление а противоположное. т.е. его фаза меняется на π . При отражении от менее плотной среды ($n_2 < n_1$) все происходит наоборот, т.е. фаза E -волны сохраняется, а H -волны меняется на π (**Рис.11.5**). Об этом обычно говорят как о *потере полуволны при отражении*. Для преломленной волны фазы обеих волн не меняются.

Отношение потока энергии отраженной волны к потоку падающей называют *энергетическим коэффициентом отражения*. Так как энергия пропорциональна квадрату амплитуды, то, очевидно, что последний связан с амплитудными коэффициентами как $R_{\perp, \parallel} = r_{\perp, \parallel}^2$. Используя формулы для $r_{\perp}, r_{\parallel}$ (11.29), нетрудно записать выражения для $R_{\perp, \parallel}$

$$R_{\perp} = \left(\frac{n_1 \cos \theta_1 - n_2 \cos \theta_2}{n_1 \cos \theta_1 + n_2 \cos \theta_2} \right)^2, \quad R_{\parallel} = \left(\frac{n_2 \cos \theta_1 - n_1 \cos \theta_2}{n_2 \cos \theta_1 + n_1 \cos \theta_2} \right)^2. \quad (11.35)$$

Если падающий пучок света линейно поляризован, а вектор E составляет угол ψ с плоскостью падения, коэффициент отражения будет

$$R = R_{\perp} \cos^2 \psi + R_{\parallel} \sin^2 \psi. \quad (11.36)$$

Для неполяризованного света (11.36) следует усреднить по всем углам ψ .

$$R = (R_{\perp} + R_{\parallel})/2. \quad (11.37)$$

11.2. Отражение света от поверхности металла. Формулы Френеля, описывающие отражение и преломление света от диэлектриков, справедливы и для металлов. Однако имеются существенные отличия при отражении света от поверхности металлов по сравнению с диэлектриками. Они обусловлены *дисперсионными* свойствами диэлектрической проницаемости металлов.

В металлах некоторые из электронов не связаны с каким-либо определенным атомом. Такие электроны называют "свободными". Они ответственны за электрическую *проводимость* металла. В отличие от оптических электронов в атомах диэлектрика на свободные электроны не действует "квазиупругая" сила, удерживающая их около

какого-то атома, но сила “трения” характеризующая сопротивление движению электрона, остается. Следовательно уравнение (10.4) классической теории дисперсии и все следствия из него можно применить к свободным электронам, положив обусловленную квазиупругой силой собственную частоту ω_0 равной нулю. Поэтому формула (10.11) для диэлектрической проницаемости применима и для металлов, но в ней необходимо положить $\omega_0 = 0$:

$$\varepsilon(\omega) = n^2 = 1 - \frac{\omega_p^2}{\omega^2 + i\omega\gamma}. \quad (11.38)$$

Здесь плазменная частота $\omega_p = \sqrt{4\pi N e^2 / m}$ имеет такой же вид как и в формуле (10.11), но под N понимают концентрацию свободных электронов.

Формула (11.38) для показателя преломления в металлах предсказывает совершенно *разный характер* распространения волн в металлах в областях *низких* и *высоких* частот. В случае *высоких частот*, удовлетворяющих неравенству $\omega \gg \gamma$, в формуле (11.38) можно пренебречь мнимым слагаемым $i\omega\gamma$, и для диэлектрической проницаемости получается вещественное выражение

$$\varepsilon(\omega) = n^2 = 1 - \frac{\omega_p^2}{\omega^2} \quad (11.39)$$

Из этой формулы видно, что плазменная частота ω_p , имеет смысл своего рода *критической частоты*. При $\omega < \omega_p$ диэлектрическая проницаемость отрицательна, а показатель преломления чисто мнимый. Это значит, что волны с $\omega < \omega_p$ (но $\omega > \gamma$) не могут распространяться в металле из-за сильного затухания, причем это затухание не связано с поглощением (т.е. диссипацией) энергии. Физически это означает, что происходит *полное отражение* падающей волны от среды.

Чтобы показать это, воспользуемся формулами (11.12), (11.28) и (11.29), которые справедливы и в случае комплексных показателей преломления. При нормальном падении

$$r_{\perp} = -r_{\parallel} = -\frac{n_2 - n_1}{n_2 + n_1} = -\frac{n - 1}{n + 1} = -\frac{(n' - 1) - in''}{(n' + 1) + in''}, \quad (11.40)$$

где $n = n_2/n_1 = n' - in''$ — *относительный показатель преломления*. Для энергетического коэффициента отражения получаем

$$R = \left| -\frac{(n' - 1) - in''}{(n' + 1) + in''} \right|^2 = \frac{(n' - 1)^2 + n''^2}{(n' + 1)^2 + n''^2} \quad (11.41)$$

Отсюда видно, что при чисто мнимом показателе преломления ($n' = 0$) коэффициент отражения равен единице.

При $\omega > \omega_p$ показатель преломления становится вещественным, а металл — прозрачным для излучения. Обычно плазменная частота у металлов попадает в область рентгеновских лучей, но для некоторых металлов область прозрачности начинается с ультрафиолетовых лучей. Например, у натрия длина волны, соответствующая граничной частоте ω_p , составляет 210 нм. Прозрачность щелочных металлов в ультрафиолетовой области спектра была обнаружена на опыте Вудом в 1933 г.

При *низких частотах*, когда $\omega \ll \gamma$, и, следовательно, $\omega^2 \ll 2\gamma\omega$, из (11.38) получаем

$$\varepsilon(\omega) = n^2 = 1 + i\frac{\omega_p^2}{2\gamma\omega}. \quad (11.42)$$

Концентрация свободных электронов в металлах такова, что $\omega_p \gg \Gamma$, поэтому $\omega_p^2/2\gamma\omega \gg \omega_p^2/2\gamma^2 \gg 1$, т.е. мнимая часть много больше действительной. Для показателя преломления получаем

$$n \simeq \sqrt{i \frac{\omega_p^2}{2\gamma\omega}} = \frac{1+i}{2} \frac{\omega_p}{\sqrt{2\gamma\omega}}. \quad (11.43)$$

Отсюда следует, что показатель преломления является комплексным с одинаковыми вещественной и мнимой частями $n' = n'' = \omega_p/\sqrt{2\Gamma\omega} \gg 1$.

Таким образом, амплитуда волны уменьшается по мере проникновения в металл. Глубина проникновения $z \simeq 1/\alpha''$, где α — показатель поглощения. Такие волны проникают вглубь металла на расстояние, которое много меньше длины волны в вакууме (скин-эффект). При нормальном падении коэффициент отражения R , определяемый действительной частью показателя преломления для них близок к единице: $R = [(n' - 1)/(n' + 1)]^2 \gg 1$ и волна практически полностью отражается от поверхности.

Для промежуточных частот $\omega \approx \omega_p$ нужно пользоваться полным выражением (11.38), а не его предельными формами. В этом случае у показателя преломления отличны от нуля зависящие от частоты вещественная и мнимая части. Это значит, что волны разных частот при распространении в металле по-разному затухают. Очень тонкие слои металла прозрачны даже для видимого света. Например, тонкий слой золота, полученный напылением в вакууме на стеклянную подложку, пропускает видимый свет, но сильно поглощает инфракрасное излучение.

Отметим, что в общем случае при отражении света от металла коэффициенты отражения $r_\perp$ и $r_\parallel$, вообще говоря, *комплексны*, т.е. имеют место *скачки фаз*. Они, как правило, различны для $\vec{E}_\perp$ и $\vec{E}_\parallel$. *Если падающий свет поляризован под углом к плоскости падения, то отраженный свет будет поляризован эллиптически.* Исследуя эллиптическую поляризацию отраженного света, можно определять оптические постоянные металла n' и n'' .

Уравнения (11.38) или (11.39) применимы также и к свободным электронам в плазме, например в ионосферной плазме. Их можно использовать для объяснения характера распространения радиоволн в ионосфере Земли. Граничная частота здесь попадает в радиодиапазон, поэтому волны длиной порядка 10 м и более отражаются ионосферой, что широко используется для радиосвязи, тогда как ультракороткие (УКВ) свободно проходят сквозь нее. Это обстоятельство открывает возможность радиолокации Луны и планет и важно для радиоастрономии, использующей технику ультракоротких волн. Исследование частотной зависимости отражения радиоволн дает хороший метод изучения ионосферы.

Стоячие электромагнитные волны. Мы уже отмечали, что в процессе отражения может иметь место изменение фазы волны. Поэтому, если уравнение *падающей* волны есть

$$E_1 = A \sin(\omega t - kz), \quad (11.44)$$

то для волны, *отраженной* в точке $z = 0$, имеем

$$E_2 = A \sin(\omega t + kz + \delta), \quad (11.45)$$

где ω и k — частота и волновое число. Перемена знака при z соответствует изменению *направления* распространения, а δ означает изменение фазы при отражении.

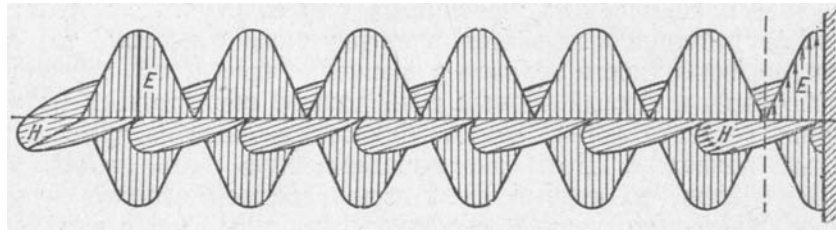


Рис. 11.6. Стоячая электромагнитная волна

Результирующая волна записывается в виде

$$E = E_1 + E_2 = 2A \cos(kz + \delta/2) \sin(\omega t + \delta/2). \quad (11.46)$$

Формула (11.46) показывает, что амплитуда колебаний равна $2A \cos(kz + \delta/2)$ меняясь от точки к точке по простому гармоническому закону. Множитель же, выражающий периодическое изменение во времени, $\sin(\omega t + \delta/2)$, не зависит от координаты. То обстоятельство, что амплитуда выражается гармонической функцией, показывает, что знак амплитуды остается неизменным в пределах полуволны и меняется на противоположный при изменении z на $\lambda/2$, т. е. при переходе от одной полуволны к другой. Другими словами, когда в пределах одной полуволны все E положительны, то в пределах соседней они отрицательны. Если считать амплитуду положительной величиной, то указанное обстоятельство можно выразить утверждением, что фаза колебания остается постоянной в пределах полуволны и меняется на π при переходе от одной полуволны к другой. Такую волну называют *стоячей волной*.

Из формулы (11.46) также следует, что в стоячей волне имеется ряд точек, которым соответствует амплитуда, равная нулю. Эти точки определяются из условия $kz + \delta/2 = n\pi/2$, где $n = 1, 3, 5, \dots$ — нечетные числа. Точки эти расположены, очевидно, на расстоянии полуволны одна от другой и называются *узловыми точками* или *узлами* стоячей волны. Посредине между ними расположены места, соответствующие максимальным значениям амплитуды, равным $2A$. Эти точки называются *пучностями*. Они определяются из условия $kz + \delta/2 = n\pi/2$, где $n = 0, 2, 4, \dots$ где — четные числа. Величина δ , определяющая изменение фазы при отражении, зависит от соотношения диэлектрических проницаемостей сред $\varepsilon_{1,2}$, на границе которых происходит отражение: если $\varepsilon_2 > \varepsilon_1$, то $\delta = 0$ для магнитной составляющей и $\delta = \pi$ — для электрической компоненты. В этом случае происходит потеря полуволны для электрического вектора. При $\varepsilon_2 < \varepsilon_1$ — наоборот электрический вектор сохраняет свою фазу неизменной, а магнитный теряет полуволну. Это различие в δ ведет к тому, что узлы одного из векторов совпадают с пучностями другого, что показано на **Рис. 11.6**. Это отличается от случая бегущих волн, электрический и магнитный векторы колеблются синфазно.

Из рассмотрения члена $\sin(\omega t + \delta/2)$ нетрудно видеть, что моменты прохождения через максимум вектора E и вектора H также отличаются друг от друга на четверть периода. Эти особенности *стоячей волны* приводят к тому, что в ней *нет непрерывного движения энергии в направлении распространения волны*, как в бегущих волнах; *энергия стоячей волны локализована* и переходит от области пучности E (где она имеет форму электрической) к области пучности H (т. е. обращается в магнитную) и обратно. Таким образом, вместо течения энергии мы имеем дело с колебаниями ее, сопровождающими переход энергии из одной формы в другую. Это обстоятельство и повело к появлению термина “стоячая волна”.

Сстоячие волны можно наблюдать не только при отражении волн, но и всякий раз, когда навстречу друг другу идут две когерентные волны одинаковой амплитуды. Простейший практический прием реализации этого условия есть отражение волны. Стоячие волны возникают, например, в резонаторе Фабри-Перо, образованным двумя плоскими зеркалами, установленными параллельно.

Из изложенного следует, что в зависимости от условий опыта можно заранее предвидеть, где расположатся узлы электрического и магнитного векторов. Этим обстоятельством можно воспользоваться, чтобы на опыте решить вопрос о том, какой из двух векторов, составляющих световую волну, электрический или магнитный, производит непосредственное действие на большинство приборов, предназначенных для обнаружения света (глаз, фотографическая пластинка, флуоресцирующий экран, фотоэлемент и т. д.).

ЛЕКЦИЯ №12

Оптическая анизотропия и основные эффекты кристаллооптики. Структура световой волны в анизотропном кристалле. Одноосный кристалл.

12.1. Оптическая анизотропия и основные эффекты кристаллооптики. Некоторые кристаллы обладают необычными оптическими свойствами. Например, кристалл исландского шпата (кальцит) *преломляет* свет по-разному в зависимости от того, с какой стороны падает свет на кристалл. В этом кристалле есть *одно направление*, вдоль которого при *нормальном* падении луч света проходит прямолинейно (**Рис.12.1 а**). Его называют *обыкновенным лучом (o)*. В других направлениях луч, проходя через кристалл, *раздваивается* и образуется два пучка примерно одинаковой интенсивности (**Рис.12.1 б**). Луч, который проходит прямо называют *обыкновенным лучом (o)*, а другой — *необыкновенным (e)*. При повороте кристалла относительно оси падающего пучка пятно, соответствующее обыкновенному лучу, на экране остается неподвижным, а второе, соответствующее необыкновенному, поворачивается вокруг первого синхронно с поворотом кристалла (**Рис.12.2 б**). Анализ поляризации лучей с помощью поляроида показывает, что оба луча, вышедших из кристалла, *линейно поляризованы*, причем направления поляризации в них взаимно ортогональны (**Рис.12.2 в**). Это явление называется *двойным лучепреломлением*, а зависимость оптических свойств, например, показателя преломления, от направления в кристалле называют *оптической анизотропией*.

Причиной оптической анизотропии является **анизотропия структуры среды**. Анизотропия среды может быть обусловлена как анизотропией составляющих ее

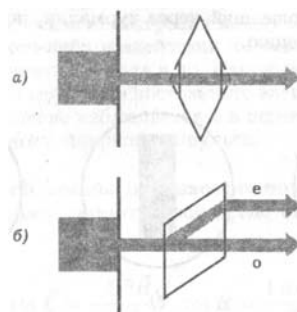


Рис. 12.1. Прохождение света вдоль оси исландского шпата (а). Двойное лучепреломление света, падающего нормально к естественной грани кристалла исландского шпата (б)

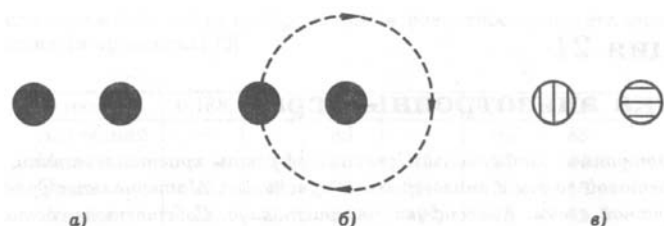


Рис. 12.2. Картины, наблюдаемые на экране в опытах с кристаллом исландского шпата (а, б). Поляризация лучей (в)

частиц, так и характером их расположения. В частности, изотропная среда может быть построена из анизотропных частиц, а анизотропная – из частиц изотропных.

Оптическая анизотропия среды характеризуется различной по разным направлениям способностью реагировать на действие падающего света. Реакция эта состоит в смещении электрических зарядов под действием поля световой волны. *В анизотропных средах величина смещения заряда в поле данной напряженности зависит от ее направления.* Это приводит к тому, что *диэлектрическая проницаемость*, а, значит, и *показатель преломления* среды *различны* для разных направлений электрического вектора световой волны. Таким образом, в анизотропных кристаллах *показатель преломления среды и, соответственно, скорость света в среде зависят от направления распространения световой волны и плоскости ее поляризации.* Поэтому для анизотропной среды волновая поверхность отлична от сферической.

Анизотропные свойства наблюдаются у *кристаллов*, но отсутствуют у газов, жидкостей (кроме жидких кристаллов), пластмасс, стекол. В некоторых случаях оптическая анизотропия может возникнуть и в изотропной среде в результате какого-либо воздействия на нее, например, механического напряжения, внешнего электрического поля и другие. В частности, двойное лучепреломление может возникать под действием постоянного электрического поля (эффект Керра). Аналогичное явление наблюдается и в переменном электрическом поле, а также в поле мощного лазерного импульса.

Еще одним проявлением анизотропии является *дихроизм* или *анизотропное поглощение*. Это явление состоит в том, что в некоторых кристаллах световая волна с определенной ориентацией вектора $\vec{E}$ поглощается сильнее, чем волны с другой ориентацией электрического поля. Например, кристалл турмалина толщиной около 1 мм практически полностью поглощает обыкновенную волну и почти не поглощает необыкновенную. Такими же свойствами обладают поляроидные пленки, у которых сильный дихроизм проявляется уже при толщине порядка 0.1 мм.

Механизм анизотропного поглощения можно пояснить следующим образом. Анизотропия структуры турмалина приводит к тому, что *электроны имеют возможность двигаться преимущественно в одном направлении относительно кристалла*. Если поляризация падающей световой волны совпадает с этим направлением, то световое поле вызывает сильную раскачку электронов и передает им свою энергию, а те, в свою очередь, передают энергию кристаллической решетке. В результате световая волна поглощается. Если же поляризация падающей волны перпендикулярна направлению возможного движения электронов в кристалле, то колебания электронов практически не возбуждаются, либо электроны колеблются с небольшой амплитудой, отдавая свою энергию вторичному излучению, а не решетке кристалла. В этом случае световая волна испытывает лишь незначительное поглощение. Именно поэтому при облучении неполяризованным (естественным) светом на выходе из кристалла образуется линейно поляризованный свет: турмалин пропускает свет лишь той поляризации, которая ортогональна направлению возможного движения электронов в кристалле.

12.2. Структура световой волны в анизотропном кристалле. Рассмотрим особенности распространения плоской монохроматической световой волны в анизотропном кристалле. Ее структура в линейной анизотропной, однородной и прозрачной среде определяется уравнениями Максвелла.

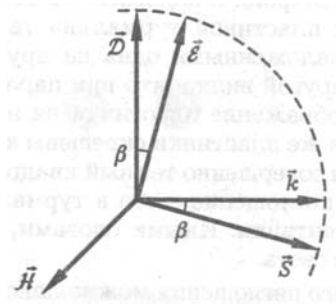


Рис. 12.3. Взаимная ориентация векторов $\vec{E}$, $\vec{H}$, $\vec{D}$ и $\vec{k}$ в анизотропном кристалле. Так как $\vec{H}$ перпендикулярен всем остальным векторам (их концы которых соединены пунктирной линией), то последние лежат в одной плоскости, перпендикулярной ему.

Для плоских монохроматических волн в комплексном представлении эти уравнения имеют вид

$$[\vec{k}, \vec{E}] = \frac{\omega}{c} \vec{H}, \quad [\vec{k}, \vec{H}] = -\frac{\omega}{c} \vec{D}, \quad (\vec{k}, \vec{D}) = 0, \quad (\vec{k}, \vec{H}) = 0, \quad (12.1)$$

где $\vec{E}$, $\vec{H}$ и $\vec{D}$ — комплексные амплитуды электрической и магнитной составляющих световой волны, и электрической индукции, соответственно; ω — частота, $\vec{k}$ — волновой вектор.

Из (12.1) видно, что векторы $\vec{H}$, $\vec{D}$, $\vec{k}$ образуют правую тройку взаимно перпендикулярных векторов, как показано на **Рис.12.3**. Вектор $\vec{E}$ перпендикулярен вектору $\vec{H}$ и лежит в плоскости векторов $\vec{D}$ и $\vec{k}$, но вообще говоря, не параллелен вектору $\vec{D}$.

Направление переноса энергии в световой волне определяется вектором Пойнтинга $\vec{S} = (c/4\pi)[\vec{E}, \vec{H}]$. Для характеристики этого направления введем *единичный вектор* $\vec{s} = \vec{S}/|\vec{S}|$, ориентированный вдоль вектора $\vec{S}$. Его называют *лучевым вектором*, так как направление переноса энергии — это и есть направление лучей. Ясно, что вектор Пойнтинга $\vec{S}$ перпендикулярен векторам $\vec{E}$ и $\vec{H}$ и, следовательно, лежит в плоскости векторов $\vec{k}$, $\vec{E}$, $\vec{D}$, но он не параллелен вектору $\vec{k}$ (**Рис.12.3**). Это означает, *направление лучевого вектора $\vec{s}$ в анизотропном кристалле может не совпадать с направлением $\vec{k}$, т.е с направлением нормали к волновому фронту световой волны*. Угол β между векторами $\vec{k}$ и $\vec{S}$ равен углу между векторами $\vec{D}$ и $\vec{E}$ и называется *углом анизотропии*. Плоскость равных фаз перемещается вдоль вектора $\vec{k}$ с фазовой скоростью v . Скорость u перемещения этой плоскости вдоль луча $\vec{s}$ называется *лучевой скоростью*. Когда $\vec{k}$ и $\vec{s}$ не совпадают, *лучевая и фазовая скорости не равны* и связаны, как видно из **Рис.12.3**, соотношением $v = u \cos \beta = u(\vec{k}\vec{s})$.

Оказывается, что именно **непараллельность** векторов $\vec{D}$ и $\vec{E}$ обуславливает своеобразные оптические свойства анизотропных кристаллов. Причина этого явления состоит в следующем. При распространении световой волны на электроны действует внешняя сила, направленная вдоль вектора $\vec{E}$ волны. Однако смещение электронов происходит не в направлении этой внешней силы, а в направлении, определяемом **структурой** кристалла. В результате вектор поляризации среды $\vec{P}$

оказывается *не параллельным* вектору $\vec{E}$. В свою очередь вектор электрической индукции $\vec{D} = \vec{E} + 4\pi\vec{P}$ также *не параллелен* вектору $\vec{E}$.

Важнейшим свойством анизотропной среды является то, что *подвижность оптических электронов по отношению к различным направлениям в кристалле неодинакова*, т.е. в одних направлениях они могут смещаться под действием света довольно свободно, а в других нет. Поэтому связь между *поляризованностью* и *напряженностью* световой волны имеет вид

$$P_\alpha = \varkappa_{\alpha x}E_x + \varkappa_{\alpha y}E_y + \varkappa_{\alpha z}E_z \quad (12.2)$$

где индекс $\alpha = (x, y, z)$.

Уравнения (12.2) принято записывать в сокращенной форме

$$P_\alpha = \varkappa_{\alpha\beta}E_\beta \quad (12.3)$$

Здесь $\alpha, \beta = (x, y, z)$, а по повторяющимся индексам подразумевается суммирование; $\varkappa_{\alpha\beta}(\omega)$ является *тензором*. В анизотропной среде связь между *поляризованностью* и *полем* имеет *тензорный характер*, а *восприимчивость* представляет собой *тензор*.

Физически (12.2) означает, что x -компонента вектора $\vec{P}$ зависит в общем случае как от x -компоненты вектора $\vec{E}$, так и от y и z -компонент. То же самое касается и других проекций вектора $\vec{P}$.

Так как $\vec{D} = \vec{E} + 4\pi\vec{P}$ или в декартовых координатах $D_\alpha = E_\alpha + 4\pi P_\alpha$, то используя (12.2), получаем $D_\alpha = E_\alpha + 4\pi\varkappa_{\alpha\beta}(\omega)E_\beta$ или

$$D_\alpha = \varepsilon_{\alpha\beta}(\omega)E_\beta, \quad (12.4)$$

где $\varepsilon_{\alpha\beta}(\omega) = \delta_{\alpha\beta} + 4\pi\varkappa_{\alpha\beta}(\omega)$ — *тензор диэлектрической проницаемости* линейной анизотропной среды, $\delta_{\alpha\beta}$ — символ Кронекера.

Можно показать, что в *прозрачной среде тензоры линейной оптической восприимчивости и диэлектрической проницаемости симметричны*, т.е.

$$\varkappa_{\alpha\beta} = \varkappa_{\beta\alpha}, \quad \varepsilon_{\alpha\beta} = \varepsilon_{\beta\alpha}, \quad \beta, \alpha = x, y, z. \quad (12.5)$$

Тензор $\varepsilon_{\alpha\beta}$ (или $\varkappa_{\alpha\beta}$) означает собой совокупность девяти чисел (матрицу)

$$\begin{pmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{xy} & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_{zz} \end{pmatrix} \quad (12.6)$$

Не всякая матрица является тензором, а лишь такая, которая связывает два вектора по правилу (12.3), которое представляет краткую запись полного соотношения (12.2). Компоненты тензора $\varepsilon_{\alpha\beta}$ определяются структурой конкретного кристалла, а также зависят от выбора системы координат.

Можно выбрать направления осей декартовой системы координат так, что симметричный тензор $\varepsilon_{\alpha\beta}$ будет диагональным. В этой системе координат тензор имеем вид диагональной матрицы

$$\begin{pmatrix} \varepsilon_{xx} & 0 & 0 \\ 0 & \varepsilon_{yy} & 0 \\ 0 & 0 & \varepsilon_{zz} \end{pmatrix}$$

Эта матрица соответствует следующей связи между $\vec{D}$ и $\vec{E}$

$$D_x = \varepsilon_{xx}E_x, \quad D_y = \varepsilon_{yy}E_y, \quad D_z = \varepsilon_{zz}E_z. \quad (12.7)$$

Формулы (12.7) "говорят" о том, что существуют три направления в кристалле, такие, что если вектор $\vec{E}$ совпадает с одним из этих направлений, то вектор $\vec{D}$ (и $\vec{P}$) будет ему *коллинеарен*. Направления x, y, z в кристалле, для которых выполняются равенства (12.7), называются *главными направлениями* или *главными осями координат*, а величины $\varepsilon_{xx}, \varepsilon_{yy}, \varepsilon_{zz}$ — *главными диэлектрическими проницаемостями* анизотропного кристалла. В общем случае они не равны друг другу.

В прозрачной среде им соответствуют три главных показателя преломления $n_x = \sqrt{\varepsilon_{xx}}, n_y = \sqrt{\varepsilon_{yy}}$ и $n_z = \sqrt{\varepsilon_{zz}}$. Это значит, что волна, в которой $\vec{E}$ параллелен оси x , распространяется с фазовой скоростью $v_x = c/n_x$, соответственно, волны, в которых колебание поля происходит параллельно осям y и z , имеют фазовые скорости $v_y = c/n_y$ и $v_z = c/n_z$. Кристаллы, у которых три главных показателя преломления различны, т.е. $n_x \neq n_y \neq n_z$, называются *двуосными*. В таких кристаллах существует два направления, при распространении вдоль которых световой луч не раздваивается. Эти направления называют *оптическими осями*.

Если одинаковы два из трех главных показателей, т.е. $n_x = n_y \neq n_z$, то кристалл называется *одноосным*. В этом случае ось z называют *оптической осью кристалла*. Подчеркнем, что оптическая ось — это не фиксированная прямая в кристалле, а *направление*, т.е. любая прямая параллельная оси z , является оптической осью кристалла. Показатель преломления n_z обычно обозначают n_e , а показатели $n_x = n_y = n_o$. Кристаллы, у которых $n_e > n_o$ называют *положительными* (кварц), а при $n_e < n_o$ — *отрицательными* (исланский шпат).

Отметим, что приведение тензора $\varepsilon_{\alpha\beta}$ к диагональному виду означает *поворот системы координат*. Такому тензору может быть поставлена в соответствие поверхность второго порядка

$$\frac{x^2}{\varepsilon_{xx}} + \frac{y^2}{\varepsilon_{yy}} + \frac{z^2}{\varepsilon_{zz}} = 1, \quad (12.8)$$

которую называют *эллипсоидом диэлектрической проницаемости* или с учетом $\varepsilon = n^2$

$$\frac{x^2}{n_x^2} + \frac{y^2}{n_y^2} + \frac{z^2}{n_z^2} = 1, \quad (12.9)$$

— *эллипсоид показателя преломления*. Здесь n_x, n_y и n_z — *главные значения показателя преломления*. В сечении этого эллипсоида плоскостью, перпендикулярной направлению распространения света, получим эллипс. *Главные полуоси этого эллипса и будут значением показателя преломления для двух взаимно перпендикулярно поляризованных компонент*.

Собственные состояния поляризации световой волны в анизотропном кристалле. Основная особенность распространения света в анизотропном кристалле состоит в том, что *световая волна с произвольным состоянием поляризации распадается на две линейно поляризованные волны с ортогональными поляризациями, бегущие с разными фазовыми скоростями*. Рассмотрим сначала простые частные случаи.

Пусть плоская монохроматическая световая волна распространяется вдоль одной из *главных осей* кристалла, например оси z (**Рис.12.4**). Обозначим орты главных осей кристалла $\vec{x}_0, \vec{y}_0, \vec{z}_0$. В этом случае волновой вектор волны $\vec{k} \parallel \vec{z}_0$. Из поперечности световой волны $(\vec{k}, \vec{D}) = 0$, следует, что вектор $\vec{D}$ лежит в плоскости x, y .

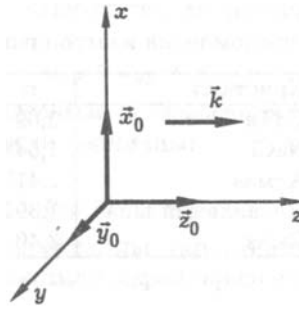


Рис. 12.4. К анализу собственных состояний поляризации световой волны в анизотропном кристалле. Показаны главные оси кристалла x , y , z и волновой вектор световой волны $\vec{k}$

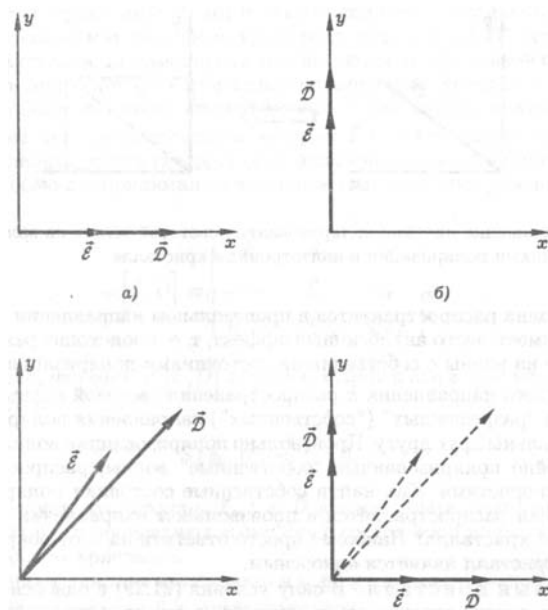


Рис. 12.5. Собственные состояния поляризации в анизотропном кристалле

а) Предположим теперь, что $\vec{D} \parallel \vec{x}_0$, т. е. волна поляризована вдоль главной оси x кристалла. Тогда $D_y = D_z = 0$, и по формулам (12.7) получаем $E_y = E_z = 0$, $D_x = \epsilon_{xx}E_x$, откуда следует, что вектор $\vec{D}$ параллелен вектору $\vec{E}$. Это означает, что по отношению к данной волне среда подобна изотропной, и волна бежит со скоростью $v_1 = c/\sqrt{\epsilon_{xx}}$. Поляризация волны устойчива (**Рис.12.5 а**).

б) Пусть теперь $\vec{D} \parallel \vec{y}_0$, т.е. световая волна параллельна главной оси y кристалла. Рассуждая аналогичным образом, нетрудно показать, что $\vec{D} \parallel \vec{E} \parallel \vec{y}_0$, но скорость волны будет теперь $v_2 = c/\sqrt{\epsilon_{yy}}$. Поляризация волны устойчива (**Рис.12.5б**).

в) Пусть теперь $\vec{D} \nparallel \vec{x}_0$, $\vec{D} \nparallel \vec{y}_0$. Используя формулы (12.7) и учитывая, что $\epsilon_{xx} \neq \epsilon_{yy}$, приходим к выводу, что векторы $\vec{D}$ и $\vec{E}$ лежат в плоскости x, y , но не параллельны друг другу (**Рис.12.5 в**). Однако в этом случае векторы $\vec{D}, \vec{E}, \vec{k}$ не компланарны и, следовательно, волна с такой структурой не может распространяться в кристалле в силу уравнений Максвелла. Таким образом, в анизотропном кристалле волна распадается на две волны с ортогональными поляризациями, бегущие

с разными скоростями, т.е. $\vec{D} = \vec{x}_0 D_x + \vec{y}_0 D_y$ (**Рис.12.5** г). По мере распространения в кристалле разность фаз между ортогональными компонентами поля будет меняться. Поэтому волна в кристалле будет иметь эллиптическую поляризацию, причем параметры эллипса поляризации будут меняться по мере изменения расстояния, пройденной волной в кристалле.

На частном примере показано, что в анизотропном кристалле произвольная световая волна неустойчива — она распадается на две линейно поляризованные волны с ортогональными направлениями поляризации, которые бегут в одном и том же направлении, но с разными фазовыми скоростями. В рассмотренном примере, когда световая волна распространяется вдоль главной оси z кристалла, есть только два устойчивых состояния поляризации волны: линейная поляризация вдоль оси x и линейная поляризация вдоль оси y . Эти состояния поляризации называют собственными состояниями поляризации световой волны в анизотропном кристалле.

Можно показать, что и в общем случае, когда волна распространяется в произвольном направлении относительно оси кристалла, также происходит разложение световой волны на волны с собственными состояниями поляризации.

Итак, для каждого направления $\vec{k}$ распространения световой волны в кристалле имеются два "разрешенных" ("собственных") направления поляризации, которые ортогональны друг другу. Произвольно поляризованная волна распадается на две линейно поляризованные "собственные" волны, распространяющиеся с разными скоростями. Найдем эти собственные состояния поляризации, если световая волна распространяется в произвольном направлении в одноосном кристалле.

12.3. Одноосный кристалл. В одноосном кристалле $n_{xx} = n_{yy} = n_0 \neq n_{zz} = n_e$, поэтому главная ось z физически выделена. Ее называют *оптической осью кристалла*. Любая световая волна, распространяющаяся вдоль оптической оси, сохраняет свою поляризацию, при этом скорость волны не зависит от ее поляризации. В одноосном кристалле такое направление только одно.

Пусть плоская монохроматическая световая волна распространяется в одноосном анизотропном кристалле в некотором направлении, характеризуемом волновым вектором $\vec{k}$. Допустимые состояния поляризации волны определяются следующими обстоятельствами. С одной стороны, из-за анизотропии кристалла векторы $\vec{D}$ и $\vec{E}$ световой волны, вообще говоря, не параллельны друг другу. С другой стороны, в силу уравнений Максвелла векторы $\vec{D}$, $\vec{E}$ и $\vec{k}$ должны лежать в одной плоскости, т.е. должны быть *компланарными*.

Первое из этих условий (*неколлинеарность* векторов $\vec{D}$ и $\vec{E}$) математически выражается формулами (12.7), которые для одноосного кристалла могут быть записаны в виде

$$D_x = n_0^2 E_x, \quad D_y = n_0^2 E_y, \quad D_z = n_e^2 E_z. \quad (12.10)$$

Условие компланарности векторов $\vec{D}$, $\vec{E}$ и $\vec{k}$ можно представить в виде равенства нулю смешанного произведения векторов

$$(\vec{k}, [\vec{E}, \vec{D}]) = 0 \quad (12.11)$$

Подчеркнем, что уравнения (12.10) и (12.11) являются *независимыми*: (12.10) представляют собой материальные уравнения среды, а (12.11) есть следствие уравнений Максвелла. Все вместе они могут выполняться не для всякой поляризации

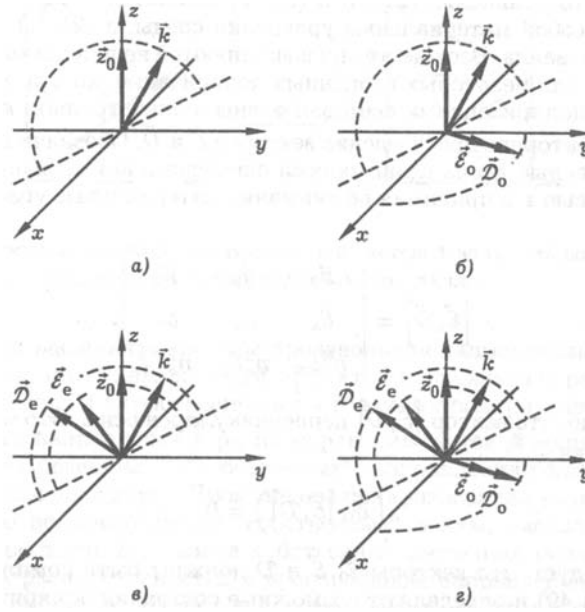


Рис. 12.6. Поляризация обыкновенной и необыкновенной световых волн в одноосном анизотропном кристалле. Главная плоскость (а), поляризация обыкновенной волны (б), поляризация необыкновенной волны (в), взаимная ориентация обыкновенной и необыкновенной волн (г). Пунктирной линией соединены концы векторов, лежащих в одной плоскости

волны, а лишь для некоторых избранных поляризаций, которые и называются *собственными поляризациями* световой волны в анизотропном кристалле.

Так как вектор $[\vec{E}, \vec{D}]$ перпендикулярен оси z , то смешанное произведение

$$(\vec{z}_0, [\vec{E}, \vec{D}]) = 0, \quad (12.12)$$

т.е. векторы $\vec{z}_0$, $\vec{D}$ и $\vec{E}$ должны быть компланарны. Условия (12.11), (12.12) определяют возможные состояния поляризации световой волны в одноосном кристалле.

Оба указанных условия выполняются в двух случаях. Во-первых, если вектор $\vec{D} \parallel \vec{E}$ и, следовательно,

$$[\vec{E}, \vec{D}] = 0. \quad (12.13)$$

Во-вторых, если вектор $\vec{D}$ лежит в плоскости векторов $\vec{z}_0$ и $\vec{k}$, что математически можно записать в виде равенства

$$(\vec{z}_0, [\vec{k}, \vec{D}]) = 0, \quad (12.14)$$

Оба полученных условия удобно выразить, если ввести понятие *главной плоскости* — плоскость, в которой лежат волновой вектор световой волны $\vec{k}$ и оптическая ось $\vec{z}_0$ одноосного анизотропного кристалла.

Условие (12.13) выполняется для волны, поляризованной перпендикулярно главной плоскости. Такую волну называют *обыкновенной* (о-волна). Вектор поляризации этой волны перпендикулярен оптической оси кристалла, векторы $\vec{D}_0$ и $\vec{E}_0 = \vec{D}_0/n_0^2$ параллельны, скорость обыкновенной волны $\vec{v}_0$ не зависит от направления распространения в кристалле. Обыкновенные волны в кристаллах ведут

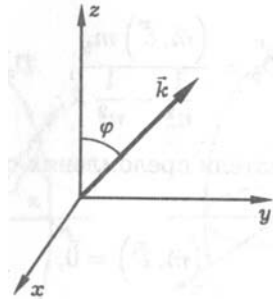


Рис. 12.7. К расчету скорости необыкновенной волны в одноосном анизотропном кристалле: $\vec{k}$ — волновой вектор световой волны, z — оптическая ось кристалла

себя, как волны в изотропных средах. Поэтому волновые поверхности излучения точечного источника являются сферами. Отсюда и название — обыкновенная волна.

Условие (12.14) выполняется для волны, поляризованной в главной плоскости. Такая волна называется *необыкновенной* (е-волна). Вектор индукции этой волны не перпендикулярен оптической оси кристалла, а векторы $\vec{D}_e$ и $\vec{E}_e = \vec{D}_0/n_e^2$ не параллельны друг другу (Рис.12.6), скорость распространения необыкновенной волны $\vec{v}_e$ зависит от направления в кристалле. Векторы $\vec{D}_0$ и $\vec{D}_e$ взаимно перпендикулярны.

Итак, в одноосном анизотропном кристалле для любого направления $\vec{k}$ существует два разрешенных ("собственных") направления поляризации световой волны. Одно из них перпендикулярно главной плоскости, другое ей параллельно. Волна с произвольным состоянием поляризации распадается в кристалле на две линейно поляризованные волны со взаимно ортогональными ("собственными") направлениями поляризации. Скорости распространения этих волн различны. Скорость обыкновенной волны не зависит от направления распространения и равна $\vec{v}_0 = c/n_0$. Скорость необыкновенной волны зависит от направления распространения в кристалле и лежит в диапазоне между c/n_0 и c/n_e . Последнее означает, что показатель преломления для необыкновенной волны зависит от направления распространения.

Для луча, распространяющегося вдоль оптической оси $n_0 = n_e$, $v_0 = v_e$, т.е. вдоль оптической оси существует только одна скорость распространения света. Различие между v_0 и v_e для всех направлений, кроме направления оптической оси, обуславливает явление *двойного лучепреломления света* в одноосных кристаллах.

Определим скорость необыкновенной волны для заданного направления распространения в кристалле. Предположим, что необыкновенная волна распространяется в некотором направлении $\vec{k}$, не совпадающим ни с одной из главных осей одноосного анизотропного кристалла (Рис.12.7). Из уравнений Максвелла (12.1) для плоских волн, запишем дисперсионное уравнение

$$[\vec{k}, [\vec{k}, \vec{E}]] + \frac{\omega^2}{c^2} \vec{D} = 0, \quad \text{или} \quad (\vec{k}, \vec{E})\vec{k} - k^2 \vec{E} + \frac{\omega^2}{c^2} \vec{D} = 0 \quad (12.15)$$

Так как для обыкновенной волны имеем $\vec{D} = n_0^2 \vec{E}$ и $(\vec{k}, \vec{E}) = 0$, то согласно уравнению (12.15), имеем $k = (\omega/c)n_0$, $v_0 = c/n_0$ вне зависимости от направления $\vec{k}$.

Введем единичный вектор $\vec{m} = \vec{k}/k$ и показатель преломления для необыкновенной n , связанный с ее волновым числом формулой $k = (\omega/c)n$. С учетом этого второе

уравнение (12.15) запишем в виде:

$$\vec{E} - \frac{1}{n^2} \vec{D} = (\vec{m}, \vec{E}) \vec{m}. \quad (12.16)$$

Для одноосного анизотропного кристалла компоненты вектора индукции $\vec{D}$ связаны с компонентами вектора напряженности электрического поля световой волны формулами (12.10). Переходя к декартовым компонентам и учитывая (12.10), получим три скалярных уравнения

$$D_x = \frac{(\vec{m}, \vec{E}) m_x}{\frac{1}{n_0^2} - \frac{1}{n^2}}, \quad D_y = \frac{(\vec{m}, \vec{E}) m_y}{\frac{1}{n_0^2} - \frac{1}{n^2}}, \quad D_z = \frac{(\vec{m}, \vec{E}) m_z}{\frac{1}{n_e^2} - \frac{1}{n^2}} \quad (12.17)$$

где n_0 и n_e — главные показатели преломления одноосного кристалла. Поскольку $(\vec{m}, \vec{D}) = 0$ или $m_x D_x + m_y D_y + m_z D_z = 0$, то (12.17) можно записать в виде

$$\frac{m_x^2 + m_y^2}{\frac{1}{n_0^2} - \frac{1}{n^2}} + \frac{m_z^2}{\frac{1}{n_e^2} - \frac{1}{n^2}} = 0 \quad (12.18)$$

Это уравнение называется *уравнением нормалей Френеля*, оно позволяет рассчитать показатель преломления необыкновенной волны n для любого направления $\vec{m}$ в кристалле.

Уравнение Френеля удобно записать в более компактном виде, если направление распространения световой волны в кристалле характеризовать не вектором $\vec{m}$, а углом φ между волновым вектором и оптической осью кристалла z (**Рис.12.7**). Учитывая, что $|\vec{m}|^2 = 1$ и, следовательно, $m_x^2 + m_y^2 + m_z^2 = 1$, где $m_z = \cos \varphi$, $m_x^2 + m_y^2 = \sin^2 \varphi$, уравнение (12.18) можно записать в виде

$$\frac{\sin^2 \varphi}{\frac{1}{n_0^2} - \frac{1}{n^2}} + \frac{\cos^2 \varphi}{\frac{1}{n_e^2} - \frac{1}{n^2}} = 0 \quad \text{или} \quad (12.19)$$

$$\frac{\sin^2 \varphi}{n_e^2} + \frac{\cos^2 \varphi}{n_0^2} = \frac{1}{n^2}$$

Из последнего уравнения получаем, что

$$n(\varphi) = \frac{n_0 n_e}{\sqrt{n_e^2 \cos^2 \varphi + n_0^2 \sin^2 \varphi}}. \quad (12.20)$$

Таким образом, зная главные показатели преломления одноосного кристалла n_0 и n_e , и направление распространения можно вычислить показатель преломления необыкновенной волны и ее *скорость* для любого направления в кристалле.

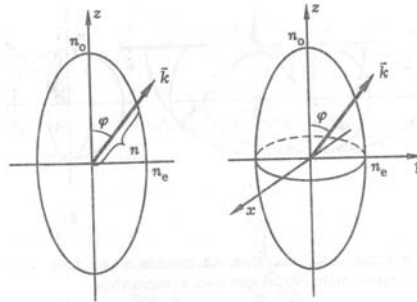


Рис. 12.8. Эллипс и эллипсоид показателя преломления. Оптическая ось совпадает с направлением оси z .

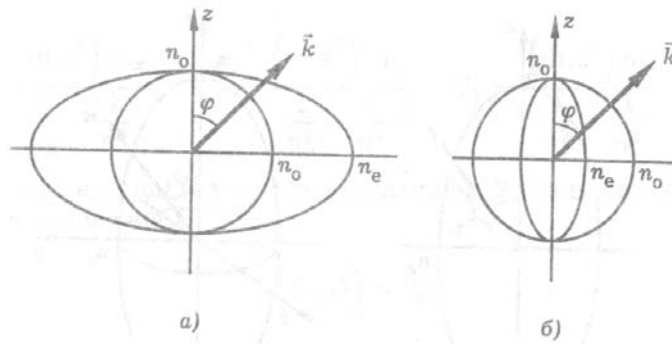


Рис. 12.9. Сфера и эллипсоид показателя преломления для положительного (а) и отрицательного (б) одноосных анизотропных кристаллов

В плоскости главного сечения зависимость $n(\varphi)$ имеет вид, показанный на **Рис.12.8**. Она имеет вид эллипса с полуосями n_0 и n_e . Значение показателя преломления необыкновенной волны n , распространяющейся в некотором направлении $\vec{k}$, равно длине отрезка, выделенного жирной линией. Из-за осевой симметрии одноосного кристалла поверхность показателя преломления в трехмерном пространстве имеет вид эллипсоида вращения с полуосями n_0 и n_e (**Рис.12.8 б**). Его называют *эллипсоидом показателя преломления*.

Отметим, что поверхность показателя преломления для обыкновенной волны представляет собой сферу радиуса n_0 . Сфера и эллипсоид показателей преломления соприкасаются в направлении оптической оси кристалла. При этом для положительного одноосного кристалла ($n_0 < n_e$) сфера лежит внутри эллипсоида, а для отрицательного кристалла ($n_0 > n_e$), наоборот, эллипсоид лежит внутри сферы (**Рис.12.9**).

При распространении света вдоль оптической оси (ось z) обыкновенная и необыкновенная волны распространяются с одинаковой скоростью $v_0 = c/n_0$. При распространении в направлении, перпендикулярном к оптической оси, о-волна имеет по-прежнему скорость $v_0 = c/n_0$, а необыкновенная — $v_e = c/n_e$. Для всех других направлений скорость распространения е-волны зависит от угла между направлением распространения волны и оптической осью, при этом направления фазовой и лучевой скоростей также не совпадают.

ЛЕКЦИЯ №13

Двойное лучепреломление света на границе с анизотропной средой. Получение и анализ поляризованного света. Наведенная анизотропия: электрооптические и магнитооптические эффекты. Естественное вращение плоскости поляризации света.

13.1. Двойное лучепреломление света на границе с анизотропной средой. Мы уже отмечали, что преломление света на границе анизотропного кристалла имеет ряд особенностей. По-первых, попадая в кристалл, световой луч раздваивается — эффект *двойного лучепреломления*. Во-вторых, независимо от состояния поляризации падающего света оба преломленных луча оказываются линейно поляризованными во взаимно перпендикулярных направлениях, причем один из преломленных лучей может не лежать в плоскости падения. В-третьих, преломление луча имеет место даже при нормальном падении света на границу анизотропного кристалла.

Для объяснения этих особенностей рассмотрим преломление света на границе одноосного анизотропного кристалла (**Рис.13.1**). Граничные условия для электромагнитного поля требуют непрерывности тангенциальных компонент электрического и магнитного полей: $E_{1t} = E_{2t}$ и $H_{1t} = H_{2t}$. Как и в случае изотропных сред, отсюда следует равенство тангенциальных компонент волновых векторов для падающей, отраженной и преломленной волн, и, как следствие, закон отражения и закон преломления Снеллиуса.

Так как показатели преломления для обыкновенной и необыкновенной волн в анизотропном кристалле, вообще говоря, *различны*, закон преломления следует записывать отдельно для обыкновенной и необыкновенной волн:

$$n_1 \sin \theta_1 = n_2^o \sin \theta_2^o, \quad n_1 \sin \theta_1 = n_2^e \sin \theta_2^e, \quad (13.1)$$

причем углы преломления $\theta_2^{o,e}$ для обыкновенной и необыкновенной волн также различны (**Рис.13.1**).

Для необыкновенной волны направление луча, задаваемое направлением вектора Пойнтинга $\vec{S} = (c/8\pi)[\vec{E}, \vec{H}]$, в общем случае также не совпадает с направлением волнового вектора. Это обстоятельство и объясняет эффект преломления света в анизотропном кристалле при нормальном падении (**Рис.13.2**) — *аномальное преломление*.

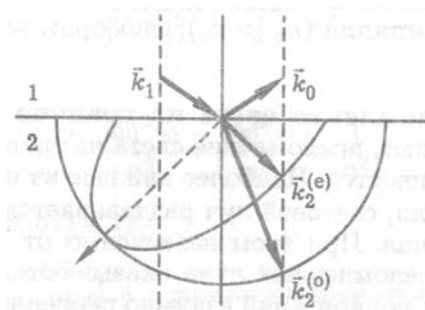


Рис. 13.1. Преломление света на границе одноосного анизотропного кристалла. Оптическая ось лежит в плоскости падения (показана пунктиром), $n_e < n_o$, $k_{1,0}$ — волновые вектора падающей и отраженной волн; $k_2^{o,e}$ — волновые вектора преломленной обыкновенной и необыкновенной волн.

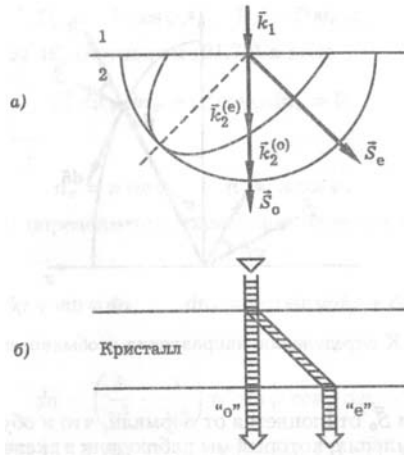


Рис. 13.2. Картина двойного лучепреломления для обыкновенной и необыкновенной волн. Волновые векторы (а), световые пучки (б)

Пусть плоская монохроматическая световая волна с волновым вектором $\vec{k}_1$ падает нормально на границу одноосного анизотропного кристалла так, что оптическая ось кристалла, показанная на рисунке пунктиром, лежит в плоскости падения. Так как при *нормальном* падении $k_{1x} = k_{0x} = k_{2x}$ (ось x направлена вдоль границы раздела), то волновые векторы преломленных волн в кристалле направлены также как и вектор падающей волны, по нормали к границе раздела. Такое же направление имеет луч $\vec{S}_o$ (вектор Пойтинга) обыкновенной волны. В то же время *необыкновенный* луч $\vec{S}_e$ отклоняется от нормали, приводя к эффекту *двойного лучепреломления*, который мы наблюдают в эксперименте.

Определим направление необыкновенного луча. Поскольку луч $\vec{S}$ ортогонален вектору напряженности электрического поля световой волны $\vec{E}$, достаточно показать, что вектор $\vec{E}$ необыкновенной волны параллелен касательной к эллипсоиду показателя преломления. Обозначим через $\vec{n}$ вектор, проведенный из центра эллипсоида показателя преломления в точку пересечения эллипсоида с волновым вектором $\vec{k}$ необыкновенной волны (**Рис.13.3**). Модуль вектора $\vec{n}$ равен показателю преломления n для необыкновенной волны, распространяющейся в данном направлении $\vec{k}$. Направление касательной к эллипсоиду показателя преломления будет определяться вектором $d\vec{n}$, который имеет смысл приращения вектора $\vec{n}$, возникающего при увеличении угла φ на величину $d\varphi$.

Покажем, что вектор $\vec{E}$ параллелен вектору $d\vec{n}$, т.е.

$$[\vec{E}, d\vec{n}] = y_0(E_z dn_x - E_x dn_z) = 0. \quad (13.2)$$

Ради простоты будем считать, что плоскость главного сечения, показанная на **Рис.13.3**, перпендикулярна одной из главных осей кристалла, например оси y . Из этого рисунка также видно, $D_x = -D \cos \varphi = n_0^2 E_x$ и $D_z = D \sin \varphi = n_e^2 E_z$. С учетом этого перепишем (13.2) в виде

$$n_0^2 \sin \varphi dn_x + n_e^2 \cos \varphi dn_z = 0. \quad (13.3)$$

Так как $n_x = n \sin \varphi$, $n_z = n \cos \varphi$, то $dn_x = \sin \varphi dn + n \cos \varphi d\varphi$, $dn_z = -n \sin \varphi d\varphi + \cos \varphi dn$, где зависимость $n(\varphi)$ определяется уравнением Френеля (12.19). Отсюда

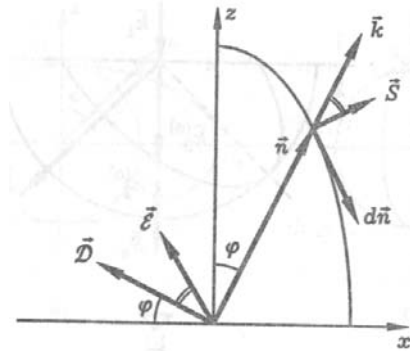


Рис. 13.3. К определению направления необыкновенного луча

следует, что

$$dn_x = n \cos \varphi d\varphi + \sin \varphi dn, \quad dn_z = -n \sin \varphi d\varphi + \cos \varphi dn, \quad (13.4)$$

где

$$dn = \left(\frac{1}{n_0^2} - \frac{1}{n_e^2} \right) n^3 \sin \varphi \cos \varphi d\varphi. \quad (13.5)$$

Подставив (13.4), (13.5) в (13.3) и учитывая (12.19), получаем тождество. Это означает, что соотношение (13.2) выполняется и, следовательно, *вектор Пойтинга направлен по нормали к эллипсоиду показателя преломления*.

В рассмотренных примерах предполагалось, что оптическая ось одноосного анизотропного кристалла лежит в плоскости падения световой волны на границу кристалла. В этом случае преломление света происходит в плоскости падения. Если же плоскость падения и плоскость главного сечения не совпадают, то преломленный необыкновенный луч не будет лежать в плоскости падения, так как отклонение этого луча от волнового вектора происходит в главной плоскости, т. е. в направлении оптической оси кристалла. Таким образом, преломление света в анизотропном кристалле происходит не обязательно в плоскости падения.

13.2. Получение и анализ поляризованного света. Для получения и анализа поляризованного света можно использовать любое физическое явление, *чувствительное к поляризации света*. К таким явлениям относятся: *анизотропное отражение, анизотропное поглощение, анизотропное преломление*.

Анизотропное отражение. Как известно, отражение света на границе раздела двух изотропных сред существенно зависит от поляризации падающей световой волны. Если падающий свет линейно поляризован, то при определенных условиях отражение может быть полностью подавлено. Для этого нужно, чтобы световая волна была поляризована в плоскости падения и падала на границу раздела под углом Брюстера. Если же направить на стеклянную пластинку под углом Брюстера неполяризованный (естественный) свет, то отраженный пластинкой свет будет линейно поляризован перпендикулярно плоскости падения.

Анизотропное поглощение. Существуют анизотропные кристаллы, обладающие *различным поглощением* по отношению к обыкновенной и необыкновенной волнам. Например, в кристалле турмалина сильно поглощается обыкновенная волна. Если



Рис. 13.4. Анизотропное поглощение. Получение (а) и анализ света с помощью кристалла турмалина

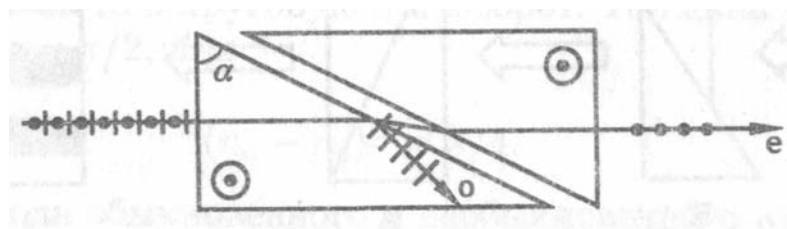


Рис. 13.5. Призма Глана

направить на кристалл неполяризованный свет, то при определенной толщине пластинки можно получить на выходе линейно поляризованный свет (необыкновенная волна проходит через кристалл, обыкновенная поглощается). В этом случае кристалл работает как поляризатор света (**Рис.13.4**).

Если же направить на кристалл линейно поляризованный свет, то пропускание света будет зависеть от взаимной ориентации кристалла и направления поляризации света. Вращая кристалл относительно оси светового пучка и наблюдая изменение интенсивности прошедшего света, можно установить сам факт линейной поляризации, а также ее направление. В этом случае кристалл работает как анализатор (**Рис.13.4** б). Аналогичным образом работают пленочные поляроиды, использующие синтетические анизотропные материалы.

Анизотропное преломление. Для получения поляризованного света можно использовать явление двойного лучепреломления. В этом случае оба вышедшие из кристалла световых луча (обыкновенный и необыкновенный) линейно поляризованы.

Призма Глана. Для получения и анализа поляризованного света на практике широко применяют призмные поляроиды (призма Глана, призма Николя и др.). На **Рис.13.5** показана схема *призмы Глана*. Призма состоит из двух кристаллов кальцита, разделенных воздушным промежутком. Главные показатели преломления кальцита $n_0 = 1,658$ $n_e = 1,486$. Оптическая ось кристалла перпендикулярна плоскости рисунка. Исходный пучок света падает нормально на входную грань призмы. Так как этот пучок перпендикулярен оптической оси кристалла, двойного лучепреломления не происходит. Световой пучок в кристалле не раздваивается и сохраняет направление исходного пучка, однако световая волна распадается на обыкновенную и необыкновенную волны, причем разность показателей преломления для этих волн достигает максимальной возможной величины $\Delta n = n_0 - n_e$. При этом угол α подбирается таким образом, чтобы свет падал на границу раздела кристалл-воздух под углом, близким к предельному углу полного внутреннего отражения. Так как на этой границе скачок показателя преломления больше для обыкновенной волны,

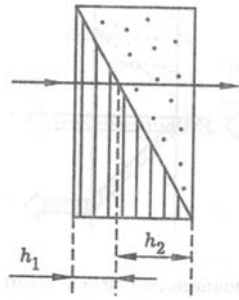


Рис. 13.6. Компенсатор Бабинe

эта волна испытывает полное внутреннее отражение. Необыкновенная волна, не испытывая полного внутреннего отражения, проходит через воздушный промежуток во второй кристалл и образует на выходе линейно поляризованный пучок света. В таком варианте призма работает как поляризатор света. Разумеется, призма Глана может работать и как анализатор. Коэффициент пропускания призмой линейно поляризованного света зависит от взаимной ориентации вектора поляризации волны и оптической оси кристаллов.

Компенсатор. Компенсаторами называются устройства, позволяющие плавно изменять разность фаз ортогональных колебаний светового поля. С помощью компенсатора можно преобразовать эллиптически поляризованный свет в линейно поляризованный, а также измерить параметры эллипса поляризации. На **Рис.13.6** показана схема *компенсатора Бабинe*. Компенсатор представляет собой пару кварцевых клиньев, образующих вместе плоскопараллельную пластину. Главные показатели преломления кварца $n_o = 1,544$ и $n_e = 1,553$. Оптические оси клиньев направлены перпендикулярно друг другу и перпендикулярно световому пучку. В такой конфигурации не происходит двойного лучепреломления, однако пучок распадается на обыкновенную и необыкновенную волны, бегущие с разными скоростями. Разность фаз ортогональных колебаний поля, вносимая компенсатором, определяется выражением

$$\Delta\varphi = \frac{2\pi}{\lambda}(n_e - n_o)(h_1 - h_2), \quad (13.6)$$

где λ — длина световой волны, $n_{e,o}$ — главные показатели преломления кристалла, $h_{1,2}$ — расстояния, пройденные световым лучом в первом и во втором клиньях. Эта разность фаз плавно меняется при перемещении компенсатора поперек светового пучка.

Управление поляризацией света. Физические явления, происходящие при распространении света в анизотропном кристалле, можно использовать для управления поляризацией света. Так, с помощью компенсатора можно управлять параметрами эллипса поляризации поляризованного светового пучка. На практике часто бывает нужно преобразовать линейную поляризацию в круговую и наоборот, а также менять направление линейной поляризации или направление вращения в циркулярно-поляризованной волне. Для этих целей служат специальные пластинки из анизотропных кристаллов — так называемые *четвертьволновые* и *полуволновые* пластинки.

Пусть плоская монохроматическая световая волна нормально падает на кристаллическую пластинку, вырезанную из одноосного анизотропного кристалла параллельно его оптической оси — оптическая ось кристалла параллельна грани пластинки, на которую падает свет. Попадая в кристалл, световая волна распадается на две линейно поляризованные волны с ортогональными направлениями поляризации — обыкновенную и необыкновенную. Так как показатели преломления кристалла для

обыкновенной и необыкновенной волн различны, то при распространении в кристалле между обыкновенной и необыкновенной волнами возникает фазовый набег

$$\Delta\varphi = \frac{2\pi}{\lambda}(n_e - n_o)d, \quad (13.7)$$

d — толщина пластинки. Изменение разности фаз ортогональных компонент светового поля изменяет состояние поляризации световой волны. На этом основано действие пластинок, управляющих поляризацией света.

Рассмотрим *четвертьволновую пластинку*. Такая пластинка преобразует *линейную поляризацию света в эллиптическую* и наоборот. Толщина пластинки d подбирается такой, чтобы $\Delta\varphi = \pi/2$ или $(n_e - n_o)d = \lambda/4$, т.е. разность хода обыкновенного и необыкновенного лучей в пластинке равна четверти длины волны. Когда *оптическая ось кристалла составляет угол в $\pi/4$ с направлением поляризации* падающей на нее световой волны линейная поляризация света преобразуется в *круговую*, так как при этом на входе пластинки ортогональные колебания светового поля в обыкновенной и необыкновенной волнах синфазны и одинаковы по амплитуде. На выходе пластинки эти колебания оказываются сдвинутыми по фазе на $\pi/2$ и, следовательно, выходной пучок имеет круговую поляризацию. Поворотом пластинки на $\pi/2$ относительно оси светового пучка можно изменить направление круговой поляризации на обратное. Аналогичным образом можно осуществить и обратную операцию — преобразования круговой поляризации света в линейную.

Полуволновая пластинка. Такая пластинка поворачивает плоскость поляризации света на $\pi/2$. В этом случае $\Delta\varphi = \pi$ и $(n_e - n_o)d = \lambda/2$

13.3. Интерференция поляризованных лучей. Явление интерференции, при котором две световые волны способны взаимно погасить друг друга, возможно, лишь при условии, что *волны имеют одинаковое направление поляризации*. Если же волны поляризованы во взаимно перпендикулярных плоскостях, то они не могут погасить друг друга ни при какой разности фаз и, следовательно, не могут давать интерференционную картину. Более того, *интенсивность света вообще не зависит от разности фаз ортогональных компонент поля*.

Отсюда следует, что обыкновенная и необыкновенная волны, распространяющиеся в анизотропном кристалле, не могут интерферировать между собой. Действительно, наблюдения показывают, что при прохождении светового пучка через пластину прозрачного анизотропного кристалла его интенсивность не меняется, т. е. интенсивность света на выходе равна интенсивности входного пучка. При этом сохранение интенсивности света имеет место независимо от разности фаз, возникающей между обыкновенной и необыкновенной волнами в кристалле.

Фазовый набег, возникающий между обыкновенной и необыкновенной волнами в анизотропном кристалле влияет не на интенсивность, а на *поляризацию света*.

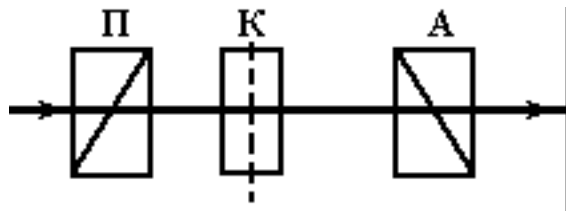


Рис. 13.7. Схема для наблюдения интерференции поляризованных лучей: П — поляризатор, К — кристалл, А — анализатор.

Изменение поляризации можно трансформировать в изменение интенсивности, если пропустить световой луч через поляроид. Таким образом, *используя поляризационные приспособления, можно наблюдать зависимость интенсивности света от разности фаз ортогональных компонент поля световой волны*. Такое явление называют *интерференцией поляризованных лучей*.

Схема наблюдения интерференции поляризованных лучей показана на **Рис.13.7**. Ее основными элементами являются источник света, поляризатор и анализатор. Обычно поляризатор и анализатор устанавливают в "скрещенном" положении, т.е. так, что система в целом не пропускала свет. Если теперь в световой пучок между поляризатором и анализатором внести пластинку из анизотропного кристалла, то система начнет пропускать свет.

Если положения поляризатора и анализатора фиксированы, то интенсивность света на экране зависит от взаимной ориентации поляризатора и кристалла, а также от фазового набегу $\Delta\varphi$, возникающего между обыкновенной и необыкновенной волнами в кристалле. В частности, если оптическая ось кристалла составляет угол $\pi/4$ с плоскостью поляризации падающего на него света и $\Delta\varphi = \pi$, т.е. кристаллическая пластинка работает как *полуволновая*, то плоскость поляризации света поворачивается в пластинке на $\Delta\varphi = \pi/2$ и вышедший из пластинки свет полностью пропускается анализатором. В этом случае свет на экране имеет максимальную интенсивность, равную интенсивности входного светового пучка. При этом же положении кристалла, но при $\Delta\varphi = 2\pi$ вращения плоскости поляризации света не происходит и свет на экране не появляется.

Фазовый набег, возникающий между обыкновенной и необыкновенной волнами в кристалле, зависит от материала пластинки, ее толщины, а также от длины световой волны. Поэтому, если в системе, показанной на **Рис.13.7**, используется белый свет, волны разных частот приобретают в кристалле разные фазовые набегу $\Delta\varphi$, по-разному изменяют состояние поляризации и, следовательно, по-разному пропускаются анализатором. Таким образом, коэффициент пропускания системы зависит от длины световой волны. Это приводит к тому, что свет, прошедший через систему, приобретает окраску.

13.4. Наведенная анизотропия: электрооптические и магнитооптические эффекты. Свойства анизотропной среды могут приобретать и изотропные материалы, если они подвергаются "анизотропному" внешнему воздействию. Жидкости, пластмассы, стекла демонстрируют оптическую анизотропию под воздействием механического напряжения, внешнего электрического поля или мощного поляризованного лазерного импульса. Эти явления называют *наведенной* или *искусственной анизотропией*.

Эффект Керра. Некоторые вещества, состоящие из *анизотропных молекул* и не обладающие собственным дипольным моментом, приобретают анизотропию под действием внешнего *электрического поля*. Этот эффект называют *эффектом Керра*. Анизотропия этих молекул проявляется в том, что дипольный момент, который приобретает молекула во внешнем электрическом поле, зависит от направления поля: в некотором направлении молекула поляризуется внешним полем легче, чем в других. Это направление называют *осью наибольшей поляризации*.

Хаотическая ориентация анизотропных молекул в пространстве (в жидкостях и газах) приводит к тому, что оптическая анизотропия не проявляется: ни одна из ориентации молекул не является преимущественной, поэтому макроскопически (в объеме, содержащем большое число молекул) среда оказывается оптически изотропной, и характер распространения света не зависит от его направления и поляризации. Однако, если такое вещество поместить во внешнее электрическое поле, то возникающий при этом момент силы, действующей на индуцированный полем диполь, заставляет молекулу повернуться таким образом, чтобы ось наибольшей поляризуемости сориентировалась преимущественно вдоль поля. В результате хаотическая ориентация осей наибольшей поляризуемости молекул сменяется частично упорядоченной ориентацией, и среда становится *оптически анизотропной*, подобно одноосному кристаллу, в котором направление оптической оси совпадает с направлением поля.

Световой волне, в которой вектор $\vec{E}$ совпадает направлением ориентирующего внешнего поля (т. е. необыкновенной волне), соответствует показатель преломления n_e и фазовая скорость $v_e = c/n_e$, а волне, в которой вектор $\vec{E}$ перпендикулярен направлению внешнего поля (обыкновенная волна), соответствует показатель преломления n_o , т.е. среда во внешнем поле приобретает свойства положительного кристалла. Анизотропия, возникающая во внешнем электрическом поле, была открыта Керром в 1875 г. Опыт показывает, что разность фаз обыкновенной и необыкновенной волн в среде равна

$$\Delta\varphi = \frac{2\pi}{\lambda}(n_e - n_o)d = 2\pi B d E^2, \quad (13.8)$$

где d — путь проходимый лучом света в среде. Квадратичная зависимость разности фаз от напряженности поляризующего поля указывает на независимость знака разности $n_e - n_o$ от направления поля. Константа B называется *постоянной Керра*.

Особенно заметно эффект Керра проявляется в нитробензоле, который обладает большим значением B : на пути $d = 10$ см и напряженности поля $E = 1,5 \cdot 10^6$ В/м разность фаз близка к значению $\pi/2$ (луч света распространяется в среде перпендикулярно приложенному полю).

На **Рис.13.8** показана схема наблюдения эффекта Керра. Между двумя *скрещенными* поляризаторами P_1 и P_2 помещается “ячейка Керра”, представляющая собой конденсатор, заполненный исследуемым веществом. При отсутствии напряжения на конденсаторе, луч света не проходит через систему, что указывает на то, что среда при отсутствии поля изотропна. При наложении внешнего поля (приложении разности потенциалов между пластинами конденсатора) среда становится оптически анизотропной, поэтому линейно-поляризованная первым поляризатором P_1 волна, пройдя через среду, становится эллиптически поляризованной и частично проходит через второй поляризатор.

Отметим исключительно малую *инерционность* эффекта Керра: при включении поля анизотропия вещества проявляется уже через время порядка 10^{-12} с. Подавая

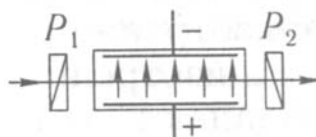


Рис. 13.8. Схема наблюдения эффекта Керра.

на обкладки конденсатора переменное напряжение, можно использовать устройство на **Рис.13.8** для модуляции лазерного пучка вплоть до частот порядка 10^9 Гц. На этом принципе основаны высокочастотные оптические затворы.

Эффект Поккельса. Электрическое поле может изменять не только оптические характеристики жидкостей и газов, оптически изотропных в отсутствие поля, но также и оптические свойства кристаллов. Если поместить одноосный кристалл в электрическое поле и направить вдоль оптической оси z луч света, в котором поле $\vec{E}$ может иметь только x и y -компоненты, то фазовые скорости волн E_x и E_y становятся различными (в отсутствие поля обе компоненты имеют равные скорости $v_x = v_y = c/n_o$). Это явление называется *эффектом Поккельса*. Малая инерционность позволяет использовать эффект для высокочастотной модуляции света. В отличие от эффекта Керра, разность фаз волн E_x и E_y в кристалле оказывается пропорциональной первой степени напряженности поля. Изменение показателя преломления (Δn) при эффекте Поккельса происходит за счет смещения электронов в кристалле в сторону того или иного иона, при этом за счет изменения электронной конфигурации меняется поляризуемость, а с ней — и показатель преломления. В силу линейности эффект Поккельса может наблюдаться лишь в кристаллах, не обладающих центром симметрии. Дело в том, что при изменении направления ориентирующего поля на противоположное, должен измениться на противоположный и знак Δn . Это невозможно в кристаллах с центром симметрии, поскольку для них оба взаимно противоположных направления поля физически эквивалентны.

Оба эффекта — Керра и Поккельса — могут проявляться как в изменении свойств оптически анизотропных сред, так и в наведении искусственной анизотропии в среде, изотропной в отсутствие внешнего воздействия.

Эффект Фарадея. При прохождении света через вещество, помещенное в магнитное поле, наблюдается *поворот плоскости поляризации линейно-поляризованной волны*. Этот эффект называется *эффектом Фарадея*. Угол поворота

$$\varphi = \rho l \vec{H}, \quad (13.9)$$

где ρ — постоянная Верде, l — длина среды, $\vec{H}$ — напряженность магнитного поля.

Рассмотрим качественное объяснение этого эффекта в рамках классического подхода. *Линейно-поляризованную волну* можно представить в виде суперпозиции двух волн с *левой* $\vec{E}_-$ и *правой* $\vec{E}_+$ *круговой поляризацией*. В отсутствие магнитного поля эти волны распространяются в среде с *одинаковой скоростью*. В этом случае при вращении векторов $\vec{E}_-$ и $\vec{E}_+$ их сумма $\vec{E}_- + \vec{E}_+$ все время остается на линии

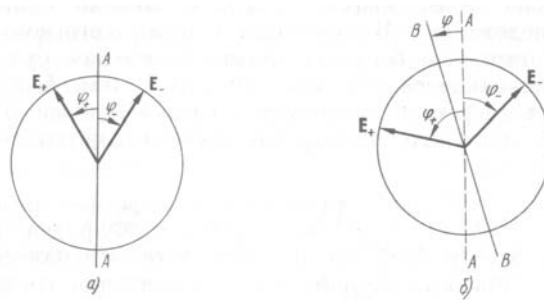


Рис. 13.9. Направление поляризации в точке $z = 0$ (AA) и в точке $z = l$ (BB)

АА (**Рис.13.9** а), так как их суперпозиция дает линейно поляризованную волну, проходящей через линию АА.

Магнитное поле приводит к анизотропии показателей преломления $n_{-,+}$ для левой и правой поляризованных волн, т.е. $n_- \neq n_+$. Пройдя в такой среде путь l , волны $\vec{E}_-$ и $\vec{E}_+$ приобретают набег фазы, равный $\varphi_{\pm} = kl = kn_{\pm}$, который и определяет положение векторов $\vec{E}_-$ и $\vec{E}_+$ в момент времени t . Как видно из **Рис.13.9** б, суммарный вектор $\vec{E}$ повернут относительно первоначальной вертикальной ориентации на угол

$$\varphi = \frac{1}{2}(\varphi_+ - \varphi_-) = \frac{\omega}{2c}(n_+ - n_-)l, \quad (13.10)$$

и колебания происходят вдоль направления ВВ.

В чем причина анизотропии показателя преломления для левой и правой круговой поляризации? С точки зрения классической физики молекулы и атомы — это системы с круговым движением электронов, создающих магнитный момент с частотой обращения ω_0 . Известно, что в магнитном поле электронная орбита прецессирует с частотой $\Omega \ll \omega_0$. При этом возможна прецессия как совпадающая с направлением движения электронов, так и противоположная. Поэтому в спектре атома, помещенного в магнитное поле, появляются две характерные частоты $\omega^+ = \omega_0 + \Omega$ и $\omega^- = \omega_0 - \Omega$. Частота ω^+ соответствует право-, а ω^- - левополяризованному кругу излучению. Это приводит к отличию показателей преломления n^+ и n^- на указанных частотах (дисперсия), т.е. $n^+ - n^- = \Delta n$.

Для наблюдения эффекта Фарадея исследуемое вещество помещается между полюсами электромагнита, расположенный между двумя скрещенными поляризаторами. В сердечнике магнита просверлен канал, через который пропускается пучок линейно-поляризованного света, распространяющийся *вдоль направления магнитного поля*. При отсутствии магнитного поля (ток в обмотке электромагнита выключен) через второй поляризатор свет не проходит. При включении магнитного поля происходит поворот плоскости поляризации на угол φ , пропорциональный пройденному светом пути и величине магнитного поля. В результате свет частично проходит через второй поляризатор. Поместив исследуемое вещество между двумя зеркалами, можно существенно увеличить проходимый светом путь (за счет многократных отражений) и, тем самым увеличит угол поворота.

13.5. Естественное вращение плоскости поляризации света. Существуют среды, при распространении в которых *линейно поляризованного света* происходит *поворот его плоскости поляризации* в отсутствие внешнего магнитного поля. Это явление называют *естественным вращением плоскости поляризации* или *естественной активностью*, или *гиротропией*. Вещества, в которых это происходит, называют *гиротропными* или *естественно оптически активными (киральными)*. Естественная активность была открыта Араго в 1811 г. для кристаллов кварца. Позднее Био обнаружил поворот плоскости поляризации света в растворах сахара.

Для наблюдения этого эффекта вещество помещают между скрещенными поляризаторами. Проходящий свет можно погасить поворотом одного из поляризаторов. Это значит, что свет остается линейно поляризованным, но его направление поляризации повернуто на некоторый угол относительно первоначального. Угол поворота зависит от длины волны света и пропорционален толщине слоя l активного вещества: $\varphi = \alpha_0(\lambda)l$, α_0 — постоянная вращения данного вещества, зависящая от длины волны.

Опыт показывает, что *при изменении направления распространения света на противоположное поворот плоскости поляризации происходит в обратную сторону*. Этим естественное вращение отличается от эффекта Фарадея. Активную среду называют *правовращающей*, если для наблюдателя, смотрящего навстречу световому лучу поворот происходит по часовой стрелке, и *левовращающей* в противоположном случае.

Поворот плоскости поляризации объясняется тем, что в *оптически активной среде волны с правой и левой круговой поляризацией имеют различные фазовые скорости*, т.е. различные показатели преломления, как и в случае эффекта Фарадея. Поэтому для угла поворота можно записать

$$\varphi = \frac{1}{2}(\varphi_- - \varphi_+) = \frac{\omega}{2c}(n_- - n_+)l, \quad (13.11)$$

Знак φ в (13.11) определен так, что $\varphi > 0$ в случае правого вращения.

Направление вращения плоскости поляризации у различных веществ зависит от природы вещества. Как правило, оптически активные вещества существуют в одной из двух модификаций: *право-* или *левовращающие*. Но есть вещества, обладающие обеими модификациями одновременно. Например, кварц встречается в природе в виде право- и левовращающего. Обе модификации кристалла отличаются друг от друга внешней формой и внутренней кристаллической структурой. По симметрии они отличаются примерно как левая спираль от правой, т.е. они не могут быть совмещены друг с другом, но зеркальное изображение одного из них совмещается с другим. Другими словами, они обладают *зеркальной симметрией*. Для оптически активных жидкостей эффект вращения обусловлен асимметричным строением молекул. Оптическая активность возникает тогда, когда в веществе преобладают молекулы только одной модификации.

Отметим что свойством оптической активности обладают все органические молекулы живых организмов, причем они существуют в одной из двух модификаций: либо правые, либо левые. Например, все молекулы ДНК имеют свойства правой спирали. Почему это так, объяснения о сих пор еще нет. Считается, что ответ на этот вопрос является ключом к разгадке эволюции жизни на нашей планете.

Теоретическое объяснение различия фазовых скоростей для левой и правой круговых поляризаций основано на учете *структуры и конечного размера молекул*. При этом существенно, что действующее на электроны поле световой волны $\vec{E}(\vec{r}, t)$ для каждого момента времени t в разных точках *протяженной молекулы* различно, т.е. необходимо учитывать *пространственную неоднородность светового поля*, действующего на молекулы или, другими словами, необходимо учитывать *пространственную дисперсию*. Влияние неоднородности поля волны на индуцированный дипольный момент молекулы с макроскопической точки зрения означает, что поляризованность среды в каждой точке зависит не только от значения поля в данной точке, но и от значений поля в соседних точках в области молекулярных размеров. Это означает, что связь между поляризованностью и полем имеет *нелокальный* характер. Таким образом, оптическая активность есть проявление пространственной дисперсии в оптике.

ЛЕКЦИЯ №14

Линейная и нелинейная оптика: нелинейно-оптические явления; специфика нелинейных явлений в оптическом диапазоне. Нелинейная поляризованность среды и нелинейно-оптические явления: классическая модель нелинейной среды — ансамбль нелинейных осцилляторов.

Хорошо известно, что оптические свойства среды зависят от частоты излучения. А зависят ли они от *интенсивности излучения*? До появления лазеров ответ на этот вопрос был отрицательный. Действительно, напряженность поля световой волны от существующих тепловых источников света не превышает 10^3 В/см, тогда как *внутриатомные поля* характеризуются напряженностями приблизительно от 10^7 В/см (для полупроводников) и до 10^9 В/см (для диэлектриков). Интенсивности света от тепловых источников *недостаточно* для того, чтобы изменить характеристики облучаемой среды.

С появлением лазеров ситуация кардинально изменилась. *Высокая пространственная когерентность лазерного излучения* позволяет осуществлять значительную *концентрацию световой энергии* и получать световые пучки с интенсивностями порядка 10^8 — 10^{12} Вт/см², что соответствует напряженности поля порядка 10^5 — 10^9 В/см. Теперь поле световой волны становится сопоставимым с внутриатомными полями. Поэтому на поставленный выше вопрос следует дать положительный ответ.

Если раньше, в “долазерной” оптике, шкала частот (или длин волн) являлась основной шкалой, при помощи которой классифицировали эффекты взаимодействия излучения с веществом, то с появлением лазеров возникла необходимость учета *зависимости эффектов взаимодействия от интенсивности излучения*. Подчеркнем, что речь идет не о каких-либо малых поправках, а о четко наблюдаемых макроявлениях, которые существенно изменяют поведение светового пучка в среде. Изучение зависимости характера оптических эффектов от интенсивности излучения привело к возникновению нового раздела оптики — *нелинейной оптики*.

14.1. Линейная и нелинейная оптика. В рамках классического подхода распространение света в среде описывают уравнениями Максвелла, дополненными материальными уравнениями. При этом предполагается, что *материальные уравнения линейны*; следовательно, система уравнений Максвелла является линейной в целом. Линейность системы уравнений означает, что световые волны с разными характеристиками (например, с разными частотами), распространяются в среде независимо друг от друга, т. е. выполняется принцип суперпозиции световых волн. Такая картина соответствует *линейной оптике*.

Появление лазеров, способных давать мощные световые пучки, привело к необходимости учитывать *зависимость оптических характеристик среды от напряженности поля световой волны*. Это означает, что *материальные уравнения становятся нелинейными*. Наведенная поляризованность есть функция напряженности поля световой волны $\vec{P} = P(\vec{E})$, которая определяется *структурой и свойствами среды*. В общем случае это *нелинейная функция*.

Таким образом, *при больших мощностях световых волн вещество становится нелинейным*. Как правило, во многих случаях функцию $P(\vec{E})$ можно представить в виде разложения в ряд по степеням напряженности поля волны (ряд Тейлора). В простейшем случае для нелинейной изотропной среды связь между поляризованностью

и напряженностью световой волны E можно записать в виде

$$P = \chi^{(1)}E + \chi^{(2)}E^2 + \chi^{(3)}E^3 + \chi^{(4)}E^4 + \chi^{(5)}E^5 \dots = P_l + P_{nl}, \quad (14.1)$$

где $\chi^{(1)}$ — линейная восприимчивость среды, а дополнительные оптические характеристики $\chi^{(2)}$, $\chi^{(3)}$ и т.д. называют *оптическими нелинейными восприимчивостями*: $\chi^{(2)}$ — *квадратичная нелинейная восприимчивость* или *нелинейность второго порядка*, $\chi^{(3)}$ — *кубическая нелинейная восприимчивость* или *нелинейность третьего порядка* и т.д. В общем случае эти восприимчивости зависят от частоты светового поля (дисперсия).

Поляризованность $P_l = \chi^{(1)}E$ называют *линейной поляризованностью среды*, а $P_{nl} = \chi^{(2)}E^2 + \chi^{(3)}E^3 + \chi^{(4)}E^4 + \chi^{(5)}E^5 \dots$ — *нелинейной*. В *изотропных средах* (газы и жидкости) *восприимчивости четного порядка отсутствуют*, т.е. равны нулю из-за свойств симметрии. В *анизотропных средах нелинейные восприимчивости имеют тензорный характер*, т.е. являются тензорами.

Формула (14.1) подразумевает, что выполняются условия

$$\chi^{(2)}/\chi^{(1)} \sim 1/E_a, \quad \chi^{(3)}/\chi^{(2)} \sim 1/E_a, \quad (14.2)$$

где E_a — напряженность внутриатомного (кулоновского) поля. Соотношения (14.2) означают, что здесь каждое последующее слагаемое меньше предыдущего.

При $E \ll E_a$, как правило, в разложении *сохраняется только первый член* — это и есть *приближение линейной оптики*. Когда же E становится сравнимым с E_a ($E < E_a$), необходимо учитывать последующие члены в разложении (14.1). Причем чем больше интенсивность светового поля, тем сильнее проявляются слагаемые более высокого порядка. В настоящее время регистрируются нелинейные эффекты 100-200 порядков.

Итак, в общем случае *материальное уравнение является нелинейным*, а *среды считаются нелинейными*. Нелинейность материальных уравнений приводит к *нелинейности всей системы уравнений Максвелла*, описывающих распространение световых волн в среде. *Основным следствием этой нелинейности является отказ от принципа суперпозиции волн*: в нелинейных средах *световые волны, распространяясь в среде, взаимодействуют друг с другом и обмениваются энергией*. В результате одни волны могут ослабляться, а другие усиливаться. Такая картина соответствует **нелинейной оптике**.

Таким образом, в основе нелинейной оптики лежат нелинейные материальные уравнения (типа (14.1)), приводящие к ряду специфических нелинейно-оптических явлений, которые наблюдаются при распространении в среде достаточно интенсивного светового излучения. Определяя *предмет нелинейной оптики*, можно сказать, что она *рассматривает изменения оптических свойств вещества в зависимости от напряженности поля световой волны*, а также *обусловленные этими изменениями нелинейно-оптические явления*, отражающие специфику распространения в средах мощных световых пучков, а также различные эффекты взаимодействия световых волн.

Нелинейно-оптические явления. Перечислим некоторые нелинейно-оптические явления, которые можно наблюдать при распространении в среде достаточно мощного светового пучка.

Оптическое детектирование. В среде под воздействием световой волны возбуждается *постоянное* (не зависящее от времени) электрическое поле, напряженность которого пропорциональна интенсивности световой волны.

Генерация оптических гармоник. За счет части энергии исходной световой волны с частотой ω в среде возбуждаются световые волны с частотами 2ω (*вторая гармоника*), 3ω (*третья гармоника*) и т. д. Более общим случаем является нелинейно-оптическое сложение и вычитание частот двух или нескольких частот лазерного излучения: $\omega_{s,d} = \omega_1 \pm \omega_2$, $\omega_{s,d} = \omega_1 \pm \omega_2 \pm \omega_3$ и т.д.

Параметрическая генерация света. За счет части энергии исходной волны с частотой ω в среде возбуждаются две дополнительные световые волны — одна с частотой ω_1 , другая — с частотой ω_2 , причем $\omega = \omega_1 + \omega_2$. По желанию экспериментатора частоту ω_1 можно непрерывно изменять или, как говорят, плавно перестраивать в пределах от нуля до ω .

Самофокусировка света. По достижении некоторого порогового значения интенсивности световой волны наблюдается *стягивание светового пучка* в тонкую “нить”. Это явление особенно эффектно, поскольку при обычных интенсивностях световые пучки в среде расширяются (за счет дифракции), “размазываются”, ослабевают. Здесь же световой пучок как бы фокусируется; его энергия не диссипирует, а, напротив, концентрируется в пределах достаточно узкого “светового канала”.

Вынужденное рассеяние света. Таких явлений несколько. Выделим два: *вынужденное комбинационное рассеяние* (ВКР) и *вынужденное рассеяние Мандельштама—Бриллюэна* (ВРМБ). По достижении некоторого порогового значения интенсивности исходной световой волны с частотой ω в среде возбуждаются интенсивные внутримолекулярные колебания (в случае ВКР) либо интенсивные акустические волны (в случае ВРМБ). В результате взаимодействия световой волны с этими типами движений возникают дополнительные световые волны (сателлиты) с частотами $\omega \pm n\omega_1$, где n — целые числа, а ω_1 — характерная частота либо внутримолекулярных колебаний (в случае ВКР), либо спектра частот нормальных колебаний решетки (в случае ВРМБ). Обычно наблюдают несколько сателлитов, интенсивности которых могут быть сопоставимы с интенсивностью исходной волны.

Исчезновение красной границы фотоэффекта. При достаточно большой интенсивности световой волны *внешний фотоэффект* наблюдают при частотах, которые в несколько раз меньше граничной частоты ω_0 , соответствующей потенциалу ионизации.

Затемнение среды. Среда, оптически прозрачная для относительно слабого излучения, становится *непрозрачной для интенсивного излучения*.

Просветление среды. Среда, непрозрачная для относительно слабого излучения, становится *прозрачной для интенсивного излучения* или, как говорят *просветляется*.

Специфика нелинейных явлений в оптическом диапазоне. Подчеркнем *специфичность* нелинейных явлений в *оптическом* диапазоне в сопоставлении с традиционным радиодиапазоном.

Во-первых, в оптическом диапазоне осуществляется *взаимодействие самих световых волн*, распространяющихся в нелинейной среде. Например, при *параметрической генерации света* производится *перестройка частоты непосредственно световых волн*: волна накачки с частотой ω передает часть своей энергии волнам с частотами ω_1 и ω_2 . В радиодиапазоне же обмен энергией между разными частотами происходит не на “уровне полей”, а на “уровне электрических токов”, являющихся источниками полей; при параметрической генерации производится перестройка частоты тока в соответствующей схеме посредством изменения одной из емкостей данного устройства.

Во-вторых, в оптическом диапазоне нелинейные явления (включая и сам процесс оптической генерации) происходят в *естественных средах* — нелинейных кристаллах, жидкостях, газах. В случае же радиодиапазона понятие “нелинейная среда” практически отсутствует. Для наблюдения нелинейных явлений в радиодиапазоне используют специальные элементы и устройства; эти явления наблюдаются в колебательных контурах, электронных лампах и т. п.

В-третьих, в оптическом диапазоне могут ярко проявляться *корпускулярные (квантовые)* свойства излучения. Эти свойства обуславливают целый ряд нелинейных явлений, не имеющих аналогий в радиодиапазоне, например явления исчезновения красной границы фотоэффекта, затемнения и просветления среды.

Для нелинейно-оптических явлений существенна *нелинейность среды*, которая проявляется лишь при прохождении сквозь нее интенсивного оптического излучения. Именно поэтому фактическое рождение нелинейной оптики связывают с появлением источников такого излучения — *лазеров*. На примере нелинейной оптики видно, какую поистине глубокую “революцию” в оптике произвело появление лазеров. Оказалось, что нелинейные явления для оптического диапазона так же характерны, как и для радиодиапазона, но они во многих отношениях специфичны и что общая картина нелинейных явлений в оптическом диапазоне представляется более многогранной, нежели в радиодиапазоне.

14.2. Нелинейная поляризованность среды и нелинейно-оптические явления.

Причиной нелинейно-оптических явлений является *нелинейная поляризованность среды*. Рассмотрим к каким нелинейно-оптическим явлениям она может приводить. Ради простоты будем считать световые волны плоскими и монохроматическими, т.е. идеально когерентными, и распространяющимися вдоль оси z :

$$E(t) = A \cos(\omega t - kz), \quad (14.3)$$

где A — вещественная амплитуда, а ω — частота, k — волновое число.

Пусть поляризованность среды описывается выражением (14.1). Подставляя (14.3) в (14.1), получаем

$$P(t) = \chi^{(1)} A \cos(\omega t - kz) + \chi^{(2)} A^2 \cos^2(\omega t - kz) + \chi^{(3)} A^3 \cos^3(\omega t - kz). \quad (14.4)$$

Здесь мы ограничились лишь первыми тремя слагаемыми в (14.1). Напомню, что второе слагаемое отсутствует в изотропных средах, оно отлично от нуля в анизотропных средах. Преобразуем (14.4) к виду

$$P(t) = \frac{1}{2} \chi^{(2)} A^2 + \left(\chi^{(1)} A + \frac{3}{4} \chi^{(3)} A^3 \right) \cos(\omega t - kz) + \frac{1}{2} \chi^{(2)} A^2 \cos(2\omega t - 2kz) + \frac{1}{4} \chi^{(3)} A^3 \cos(3\omega t - 3kz). \quad (14.5)$$

Результат (14.5) означает, что при прохождении световой волны частоты ω через нелинейную среду должно наблюдаться *переизлучение*, включающее в себя:

- постоянное (статическое) электрическое поле, обусловленное первым слагаемым в (14.5) (*эффект выпрямления*),
- световую волну частоты ω (второе слагаемое),
- световую волну частоты 2ω — *вторая гармоника* (третье слагаемое),
- световую волну частоты 3ω — *третья гармоника* (четвертое слагаемое).

Учет последующих членов в разложении (14.1) приводит к наличию переизлучения оптических гармоник более высоких порядков (четвертого, пятого и т.д.).

Поляризованность на частоте ω согласно (14.2):

$$P_1(\omega) = \left(\chi^{(1)} A + \frac{3}{4} \chi^{(3)} A^3 \right) \cos(\omega t - kz) \quad (14.6)$$

приводит к *переизлучению* на этой частоте.

С учетом (14.6) для *электрической индукции* можно записать

$$D_1(\omega) = (1 + 4\pi\chi^{(1)} + 3\pi\chi^{(3)}A^2) E(t), \quad (14.7)$$

и для *диэлектрической проницаемости* получаем

$$\varepsilon = \varepsilon_0 + \varepsilon_2 A^2. \quad (14.8)$$

Здесь $\varepsilon_0 = 1 + 4\pi\chi^{(1)}$, $\varepsilon_2 = 3\pi\chi^{(3)}$. Второе слагаемое в (14.8) описывает *нелинейную добавку к диэлектрической проницаемости на частоте ω* . Так как показатель преломления $n = \sqrt{\varepsilon}$, то

$$n = \sqrt{\varepsilon_0 + \varepsilon_2 A^2} \simeq \sqrt{\varepsilon_0} \left(1 + \frac{\varepsilon_2}{2\varepsilon_0} A^2 \right) = n_0 + n_2 A^2, \quad (14.9)$$

где $n_2 = n_0 \varepsilon_2 / 2\varepsilon_0 = 3n_0 \pi \chi^{(3)} / 2(1 + 4\pi\chi^{(1)})$. Величина n_2 имеет размерность, обратную интенсивности света, она является удобной характеристикой кубичной нелинейной среды. Например, для кристалла кварца $n_2 = 3 \cdot 10^{-16} \text{ см}^2/\text{Вт}$.

Отсюда видно, что *световая волна в нелинейной среде изменяет показатель преломления среды*, причем добавка к показателю преломления пропорциональна интенсивности волны. Как будет показано в далее, она приводит, например, к явлению *самофокусировки* света.

Оценим по величине линейную и нелинейные оптические восприимчивости среды, например кристалла. Линейная восприимчивость $\chi^{(1)}$ связана с показателем преломления среды n формулой $n^2 = 1 + 4\pi\chi^{(1)}$. Типичное значение показателя преломления для кристалла составляет $n = 1.5$. Исходя из этой цифры, получаем оценку $\chi^{(1)} \sim 0.1$.

Для оценки *квадратичной восприимчивости* $\chi^{(2)}$ воспользуемся материальным уравнением (14.1). Согласно этому уравнению, размерность $\chi^{(2)}$ определяется формулой $\chi^{(2)} = \chi^{(1)} / E$. Воспользуемся этой формулой для количественной оценки. Поскольку восприимчивость есть характеристика вещества, в качестве напряженности электрического поля E следует подставить некоторую величину, характерную для среды. Характерным масштабом поля в среде является *напряженность внутриатомного электрического*, которую можно оценить по формуле $E_a = e/a^2$, где e — заряд электрона, a — размер атома. Полагая $e = 4.7 \cdot 10^{-10}$ СГСЭ, $a = 0.5 \cdot 10^{-8}$ см (боровский радиус), получим $E_a = 2 \cdot 10^7$ СГСЭ. Таким образом, для χ_2 получаем следующую оценку (по порядку величины) $\chi^{(2)} = 5 \cdot 10^{-9}$ СГСЭ. Для кристалла KDP, широко применяемого в нелинейной оптике, $\chi^{(2)} = 5.3 \cdot 10^{-9}$ СГСЭ.

Аналогичным образом можно оценить кубичную нелинейную восприимчивость по формуле $\chi^{(3)} = \chi^{(2)} / E_a^2$. Численная оценка дает $\chi^{(3)} = 2.5 \cdot 10^{-16}$ СГСЭ = $2.5 \cdot 10^{-16} \text{ см}^3/\text{эрг}$. Реальная величина χ_3 , например для кристалла кварца, равна $10^{-14} \text{ см}^3/\text{эрг}$. Имеются, однако, материалы, обладающие значительно более высокой кубической восприимчивостью.

Интересно сравнить между собой величины линейной и нелинейной поляризованностей среды (восприимчивости разных порядков сравнить нельзя, так как они имеют разные размерности). Например, ограничиваясь нелинейностью второго порядка,

можно записать

$$P = P_l + P_{nl}^{(2)} = \chi^{(1)} E + \chi^{(2)} E^2, \quad (14.10)$$

откуда можно записать

$$\frac{P_{nl}^{(2)}}{P_l} = \frac{\chi^{(2)} E^2}{\chi^{(1)} E} = \frac{\chi^{(2)} E}{\chi^{(1)}}. \quad (14.11)$$

Используя для оценки формулу $\chi^{(2)} = \chi^{(1)}/E_a$, получим

$$\frac{P_{nl}^{(2)}}{P_l} = \frac{E}{E_a}. \quad (14.12)$$

Таким образом, отношение квадратичной поляризации среды к линейной равно отношению напряженности электрического поля световой волны к напряженности внутриатомного поля.

Аналогичным образом можно показать, что

$$\frac{P_{nl}^{(3)}}{P_l} = \left(\frac{E}{E_a} \right)^2. \quad (14.13)$$

Оценим теперь величину E/E_a . Интенсивность света связана с напряженностью поля E световой волны формулой $I = cE^2/8\pi$. Отсюда $E = \sqrt{8\pi I/c}$. Полагая $I = 10^9$ Вт/см², получаем $E = 3 \cdot 10^3$ СГСЭ. Отсюда $E/E_a = 10^{-4}$. Таким образом, для данной интенсивности света относительная величина нелинейной поляризованности оказывается весьма мала. Тем не менее, нелинейный эффект может быть сильным благодаря тому, что он может *накапливаться* в процессе распространения световой волны.

Классическая модель нелинейной среды — ансамбль нелинейных осцилляторов. Механизмы оптической нелинейности весьма разнообразны. Однако наиболее универсальным из них можно, по-видимому, считать механизм, связанный с *нелинейностью элементарного осциллятора среды* — атома или молекулы. Для примера вычислим нелинейную *квадратичную поляризованность* и *квадратичную восприимчивость* среды, рассматривая ее как ансамбль нелинейных (ангармонических) осцилляторов.

Используя второй закон Ньютона, уравнение движения осциллятора запишем в виде

$$m\ddot{x} = F_v + eE \quad (14.14)$$

Здесь m , e — масса и заряд электрона, x — смещение центра электронного облака относительно атомного ядра, $E = A \exp[i(\omega t - kz)] + k.c.$ — напряженность электрического поля световой волны, F_v — возвращающая сила, обусловленная притяжением электрона к ядру. Она связана с потенциальной энергией $U(x)$ электрона в поле ядра соотношением

$$F_v = -\frac{\partial U}{\partial x}. \quad (14.15)$$

В окрестности положения равновесия электрона ($x = 0$) потенциальную энергию $U(x)$ можно представить в виде разложения по степеням x

$$U(x) = \frac{1}{2}\alpha x^2 + \frac{1}{3}\beta x^3 + \dots \quad (14.16)$$

Первое слагаемое в (14.16) соответствует параболическому приближению. Остальные слагаемые описывают отличие формы реальной потенциальной ямы от параболической. Учет этих слагаемых важен, если амплитуда колебаний электрона достаточно велика. Последнее может иметь место в поле световой волны большой интенсивности. Подставив (14.16) в (14.15), получим

$$F_v = -(\alpha x + \beta x^2 + \dots) \quad (14.17)$$

Таким образом, *возвращающая сила оказывается нелинейной функцией смещения*. Подставляя (14.17) в (14.14) и ограничиваясь учетом первой нелинейной поправки, получим уравнение

$$\ddot{x} + \omega_0^2 x + \gamma x^2 = \frac{e}{m} E \quad (14.18)$$

где $\omega_0 = \sqrt{\alpha/m}$ — собственная частота колебаний осциллятора, $\gamma = \beta/m$ — параметр нелинейности. Затухание электронных колебаний учтем, добавляя в левую часть уравнения (14.18) слагаемое $\Gamma \dot{x}$. В итоге получим уравнение

$$\ddot{x} + \Gamma \dot{x} + \omega_0^2 x + \gamma x^2 = \frac{e}{m} E \quad (14.19)$$

Уравнение (14.19) описывает колебания атомного осциллятора под действием поля световой волны. Оно учитывает нелинейность осциллятора, которая становится существенной, если амплитуда колебаний достаточно велика, поэтому его называют *ангармоническим осциллятором*. Ангармонизм элементарного осциллятора приводит к появлению нелинейной поляризованности среды.

Используя уравнение (14.19) вычислим поляризованность среды P , которая определяется как дипольный момент единицы объема. Считая среду однородной, запишем $P = Nr = Nex$, где N — число атомов в единице объема, $r = ex$ — дипольный момент элементарного осциллятора (атома), e — заряд электрона, x — смещение электрона относительно положения равновесия, определяемое уравнением (14.19). Таким образом, для вычисления поляризации среды необходимо решить уравнение (14.19).

Точное решение этого уравнения невозможно. Это *типичная* для нелинейной оптики ситуация, когда уравнение, описывающее нелинейный эффект, не имеет точного решения или это решение настолько сложно, что практически им трудно воспользоваться. В этих условиях используют различные приближения, причем выбор конкретного метода для каждой задачи, вообще говоря, индивидуален.

Воспользуемся *методом возмущений* (метод последовательных приближений), который является одним из наиболее универсальных методов анализа нелинейных систем. Основная идея этого метода состоит в том, что сначала описывают движение системы в *линейном приближении*, а затем рассматривают нелинейный эффект как *малую поправку*. Предположим, что амплитуда колебаний осциллятора настолько мала, что в любой момент времени нелинейный член в уравнении (14.19) много меньше линейных слагаемых, т.е. $\gamma x^2 \ll |\omega_0^2 x|$ или $|\gamma x| \ll \omega_0^2$. В этом случае решение уравнения (14.19) можно представить в виде

$$x = x_l + x_{nl}, \quad (14.20)$$

где x_l — решение линейного уравнения

$$\ddot{x}_l + \Gamma \dot{x}_l + \omega_0^2 x_l = \frac{e}{m} E, \quad (14.21)$$

а x_{nl} — нелинейная поправка, малая по сравнению с x_l :

$$x_{nl} \ll x_l. \quad (14.22)$$

В случае плоских монохроматических волн решение уравнения (14.21) имеет вид

$$x_l = \frac{1}{2} \tilde{x}_l \exp[i(\omega t - kz)] + \text{к.с.} = \frac{1}{2} \alpha(\omega) A \exp[i(\omega t - kz)] + \text{к.с.}, \quad (14.23)$$

где $\tilde{x}_l$ — амплитуда,

$$\alpha(\omega) = \frac{e}{m} \frac{1}{\omega_0^2 - \omega^2 + i\omega\Gamma}. \quad (14.24)$$

Подставим (14.20) в (14.19), получим уравнение

$$\ddot{x}_l + \ddot{x}_{ln} + \Gamma \dot{x}_l + \Gamma \dot{x}_l + \omega_0^2 x_l + \omega_0^2 x_{nl} + \gamma(x_l + x_{ln})^2 = \frac{e}{m} E. \quad (14.25)$$

Используя (14.21) и пренебрегая x_{nl} в круглых скобках, согласно (14.22), уравнение (14.25) преобразуется к виду

$$\ddot{x}_{ln} + \Gamma \dot{x}_l + \omega_0^2 x_{nl} = -\gamma x_l^2. \quad (14.26)$$

Таким образом мы произвели *линеаризацию* исходного нелинейного уравнения (14.18) по малому параметру x_{nl} . Уравнение (14.26) представляет собой линейное уравнение вынужденных колебаний, в котором роль вынуждающей силы играет член γx_l^2 , определяемый решением уравнения движения электрона в линейном приближении (14.23). Найдём решение уравнения (14.26). Согласно (14.23) для x_l^2 можно записать

$$x_l^2 = (1/4) \tilde{x}_l^2 e^{i(2\omega t - 2kz)} + (1/2) |\tilde{x}_l|^2 + \text{к.с.} \quad (14.27)$$

т.е. вынуждающая сила содержит постоянную (не зависящую от времени) составляющую и переменную компоненту, осциллирующую на частоте второй гармоники 2ω . Так как уравнения (14.26) линейные, то такую же структуру будет иметь его решение. Поэтому решение ищем в виде

$$x_{nl} = \frac{1}{2} x_2 e^{i(\omega t - 2kz)} + \frac{1}{2} x_0 + \text{к.с.} \quad (14.28)$$

Подставляя (14.28), (14.27) в (14.26), получим

$$\frac{1}{2} (\omega_0^2 - 4\omega^2 + 2i\omega\Gamma) x_2 e^{i(\omega t - 2kz)} + \frac{1}{2} \omega_0^2 x_0 + \text{к.с.} = -\frac{1}{4} \tilde{x}_l^2 e^{i(2\omega t - 2kz)} - \frac{1}{4} \gamma |\tilde{x}_l|^2 + \text{к.с.} \quad (14.29)$$

Приравнявая коэффициенты при одинаковых экспонентах, получим нелинейные поправки $x_{0,2}$

$$x_0 = \frac{\gamma}{2\omega_0^2} |\tilde{x}_l|^2, \quad x_2 = -\frac{\gamma}{2(\omega_0^2 - 4\omega^2 + 2i\omega\Gamma)} \tilde{x}_l^2 \quad (14.30)$$

Вычислить теперь поляризованность среды, представив ее в виде $\mathcal{P} = \mathcal{P}_l + \mathcal{P}_{nl}$, где

$$\mathcal{P}_l = \frac{1}{2} P_l e^{i(\omega t - kz)}, \quad P_l = Ne \tilde{x}_l, \quad (14.31)$$

$$\mathcal{P}_{nl} = -\frac{1}{2} [P_2 e^{i(2\omega t - 2kz)} + P_0], \quad P_2 = Ne x_2, \quad P_0 = Ne x_0. \quad (14.32)$$

Здесь P_l линейная, а P_{nl} нелинейная поляризованности; P_2 и P_0 — квадратичные нелинейные поляризованности на частоте 2ω и нулевой частоте, соответственно.

Используя выражения для $x_{2,0}$ (14.30), для $P_{2,0}$ получаем

$$P_2 = -\frac{1}{2}\gamma Ne \frac{\alpha^2(\omega)}{\omega_0^2 - 4\omega^2 + 2i\omega\Gamma} A^2, \quad P_0 = -\frac{1}{2}\gamma Ne \frac{\alpha^2(\omega)}{\omega_0^2} A^2, \quad (14.33)$$

где $\alpha(\omega)$ определяемая формулой (14.24).

Таким образом поляризованность квадратичной среды содержит три спектральные компоненты: компоненту на частоте возбуждающей световой волны (линейная поляризованность), компоненту на частоте 2ω и постоянную составляющую (на нулевой частоте).

Компонента поляризованности на частоте 2ω ответственна за генерацию второй оптической гармоники.

Под действием мощной световой волны в квадратично-нелинейной среде также возникает *статическая поляризованность* (на нулевой частоте), величина которой пропорциональна интенсивности света. Этот эффект называется *оптическим детектированием* или *выпрямлением света*. Она, в свою очередь, приводит к появлению постоянного электрического поля в среде, которое может быть зарегистрировано и измерено, если нелинейную среду поместить в конденсатор.

Квадратичная нелинейная восприимчивость $\chi^{(2)}(2\omega)$ на частоте второй гармоники 2ω определяется как коэффициент пропорциональности между комплексной амплитудой поляризованности среды на этой частоте и квадратом комплексной амплитуды поля (14.32). Из (14.33) получаем

$$\chi^{(2)}(2\omega) = -\frac{1}{2}\gamma Ne \frac{\alpha^2(\omega)}{\omega_0^2 - 4\omega^2 + 2i\omega\Gamma} A^2. \quad (14.34)$$

Аналогично для квадратичной восприимчивости среды на нулевой частоте получаем

$$\chi^{(2)}(0) = -\frac{1}{2}\gamma Ne \frac{\alpha^2(\omega)}{\omega_0^2}. \quad (14.35)$$

Полученные формулы (14.34) и (14.35) показывают, что нелинейная восприимчивость среды зависит, во-первых, от ангармоничности элементарного осциллятора, характеризуемой параметром γ , а во-вторых, от величины $\alpha(\omega)$, которая пропорциональна линейной поляризуемости атома. Поэтому нелинейные среды нужно искать, прежде всего, среди сред с большими показателями преломления. Кроме того нелинейная восприимчивость среды возрастает в *резонансных условиях*, когда частота возбуждающей световой волны ω либо частота второй гармоники 2ω близка к собственной частоте колебаний элементарного осциллятора ω_0 . Последнее более предпочтительнее, так как в первом случае возрастает поглощение возбуждающей волны.

ЛЕКЦИЯ №15

Пространственное накопление нелинейно-оптических явлений, фазовое согласование в анизотропных кристаллах. Самофокусировка света. Параметрическая генерация света.

15.1. Пространственное накопление нелинейно-оптических явлений, фазовое согласование в анизотропных кристаллах. Рассмотрим какие условия необходимы для эффективной генерации гармоник. Во-первых, в среде должна наводиться соответствующая нелинейная поляризованность. Но этого недостаточно. В оптическом диапазоне длина волны излучения $\lambda \ll L$ (L — линейные размеры среды). Это означает, что в любом нелинейном образце совершается множество *локальных* нелинейно-оптических актов (процессов) переизлучения (например, на частоте второй гармоники), обусловленные большим количеством атомов. *Вторичные световые волны, возникшие в результате переизлучения атомов среды, представляют собой результат интерференции.* Поэтому важно знать, каковы условия *конструктивной интерференции*, которая и приводит к *накоплению нелинейно-оптических явлений в пространстве*. В качестве примера рассмотрим *генерацию второй гармоники*.

Пусть в квадратичную нелинейную входит плоская монохроматическая волна с частотой ω_1 и волновым вектором k_1 в направлении z

$$E_1(z, t) = E_1 \cos(\omega_1 t - k_1 z). \quad (15.1)$$

Фазовая скорость этой волны $v_1 = \omega_1/k_1 = c/n_1(\omega_1)$, $n(\omega_1)$ — показатель преломления. Она возбуждает в каждой точке среды колебания *поляризованности* $P_2(z, t) = (1/2)E_1^2 \chi^{(2)}(2\omega_1) \cos(2\omega_1 t - 2k_1 z)$ с частотой $2\omega_1$, которая также распространяется с той же фазовой скоростью, что и исходная волна $E_1(z, t)$. *Распространение волны поляризованности можно рассматривать как бегущую в среде антенну, которая излучает волну на удвоенной частоте $2\omega_1$ (вторая гармоника)* $E_2(z, t) = E_2 \cos(2\omega_1 t - k_2 z)$ с волновым вектором $k_2 \neq k_1$.

Из-за дисперсии фазовая скорость второй гармоники $v_2 = 2\omega_1/k_2 = c/n_2(2\omega_1)$ не равна фазовой скорости v_1 исходной волны, так как $n_1 \neq n_2$, а значит, и фазовой скорости волны поляризованности $P_2(z, t)$. Это означает, что между исходной волной, а также волной поляризованности, и волной, излучаемой “антенной” (квадратичной поляризованностью) в разные моменты времени (и в разных местах) возникает фазовый сдвиг $\Delta\varphi = (k_2 - 2k_1)z$. Расстояние, на котором фазовый сдвиг составляет π называется *длиной когерентности*. Ее можно определить как

$$z = L_c \simeq \frac{\pi}{k_2 - 2k_1} = \frac{\pi c}{2\omega_1(n_2 - n_1)}. \quad (15.2)$$

Из (15.2) видно, что L_c определяется разностью показателей преломления на частоте второй гармоники и исходной волн. Обычно длина когерентности мала. Например, для кварца разность показателей преломления $(n_2 - n_1) \simeq 0.025$ (для длины волны рубинового лазера $\lambda = 694,3$ нм). Согласно (15.2), длина когерентности оказывается чрезвычайно малой: $L_{coh} \simeq 10^{-2} \text{ мм} \simeq 10\lambda$.

Для того чтобы происходила *эффективная передача энергии* от исходной световой волны с частотой ω к переизлученной волне на удвоенной частоте, на длине пути z в среде фазовый сдвиг $\Delta\varphi$ между волной поляризованности и переизлученной

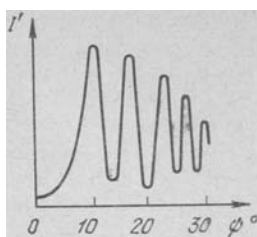


Рис. 15.1. Экспериментальная зависимость интенсивности второй гармоники от угла поворота кварцевой пластинки

волной должен быть существенно меньше π , т.е.

$$\Delta\varphi = (k_2 - 2k_1)z \ll \pi \quad \text{или} \quad z \ll L_c. \quad (15.3)$$

В этом случае все вторичные волны, излученные бегущей нелинейной поляризованностью, складываются *синфазно*, поэтому амплитуда суммарной волны нарастает или, как говорят, происходит *накопление нелинейных явлений в пространстве*. При длине среды $z > L_c$ становится существенной расфазировка вторичных волн, излученных волной поляризованности $P_2(z, t)$. В результате амплитуда волны второй гармоники начинает уменьшаться, так как происходит обратная “перекачка” энергии второй гармоники в исходную волну по мере распространения в нелинейной среде. Волна второй гармоники периодически (с периодом $2L_c$) то нарастает, то уменьшается. На **Рис.15.1** показана экспериментальная зависимость интенсивности второй гармоники от угла поворота кварцевой пластинки, играющей роль нелинейной среды, относительно лазерного пучка. Изменение угла поворота приводило к изменению длины пути, проходимого световым пучком внутри пластинки.

Для эффективного преобразования излучения исходной волны во вторую гармонику необходимо иметь большую длину когерентности L_c . Для этого должно выполняться условие $k_2 = 2k_1$ или $n_2 = n_1$, которое называют *условием фазового или волнового согласования*.

Физически *условие фазового согласования* означает, что волны с частотой ω_1 и $2\omega_1$ распространяются с одинаковыми фазовыми скоростями. Это условие можно выполнить для световых волн, распространяющихся в строго определенных направлениях внутри нелинейного *двухлучепреломляющего кристалла*, используя взаимодействие волн с разной поляризацией — *обыкновенной и необыкновенной* волн. Эти направления называют *направлениями волнового согласования*.

Возможность получения фазового согласования в анизотропных кристаллах показана на **Рис.15.2**, где изображены *индикатриссы показателей преломления* на частотах ω_1 и $2\omega_1$ в плоскости главного сечения одноосного отрицательного кристалла (например, KDP). Если из точки О провести некоторый луч, то расстояние от О до точки пересечения луча с индикатриссой даст значение показателя преломления для волны, распространяющейся внутри кристалла в направлении данного луча. Кривые 1 и 2 — показатели преломления обыкновенной и необыкновенной волн на частоты ω_1 , соответственно; кривые 3 и 4 — показатели преломления обыкновенной и необыкновенной волн частоты $2\omega_1$. Для обыкновенных волн эти кривые (тонкие линии) являются окружностями, а для необыкновенных (жирные) — эллипсами. Направление OO_1 есть оптическая ось кристалла. Из рисунка видно, что если выбрать направление распространения световой волны (ось Oz) вдоль линии OA ,

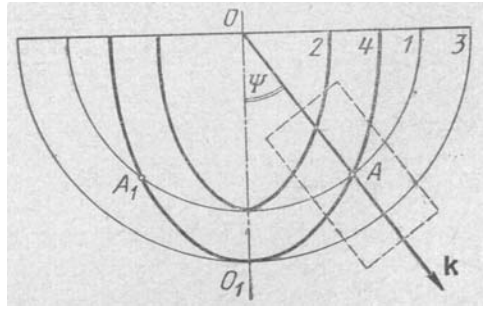


Рис. 15.2. Индикатриссы показателей преломления на частотах ω_1 и $2\omega_1$ в плоскости главного сечения одноосного отрицательного кристалла.

проходящей через точку A пересечения кривых 1 и 4, то в этом случае будем иметь $n_0(\omega_1) = n_e(2\omega_1)$.

Вообще говоря, существует бесконечно много направлений синхронизма. В случае, изображенном на **Рис.15.2**, эти направления образуют коническую поверхность с вершиной в точке O и с углом раствора конуса Ψ . Отметим, что *направления фазового согласования существенно зависят не только от выбора кристалла, но и от частоты волны*, подлежащей удвоению.

Пусть исходная волна в кристалле является обыкновенной, вектор $\vec{E}_1$ в этой волне перпендикулярен оптической оси z (его проекции на эту ось равны нулю). Так как связь квадратичной поляризованности P_2 с напряженностью волны накачки E_1 имеет тензорный характер — $P_{2i}^{(2)} = \chi_{ijk}^{(2)} E_{1j} E_{1k}$ ($i, j, k = x, y, z$, по повторяющимся индексам подразумевается суммирование), то ее проекция на оптическую ось z оказывается отличной от нуля, так как она содержит слагаемые, пропорциональные величинам $E_x(\omega)E_y(\omega)$, $E_x^2(\omega)$ и $E_y^2(\omega)$. Другими словами, благодаря нелинейности обыкновенная волна частоты ω_1 порождает колебания поляризованности вдоль оптической оси на удвоенной частоте. Эти колебания поляризованности, параллельные оптической оси, порождают вторую гармонику, в которой колебания вектора $\vec{E}_2$ также параллельны оптической оси, т. е. *вторая гармоника в кристалле является необыкновенной волной*. Если теперь выбрать направление распространения, в котором фазовая скорость волны частоты $2\omega_1$ оказывается равной фазовой скорости обыкновенной волны ω_1 , получим эффективную генерацию второй гармоники. В частности, при генерации второй гармоники с длиной волны ($\lambda_2 = 0.53$ мкм) излучения неодимового лазера ($\lambda_1 = 1.06$ мкм) в кристалла KDP эффективность преобразования может достигать несколько десятков процентов (40-50% и даже больше) при длине кристалла в 3-5 см. Для этого удобно вырезать из кристалла образец с осью, направленной вдоль одного из направлений синхронизма, как показано на **Рис.15.2** штриховым прямоугольником, а исходная световая волна должна иметь поляризацию, совпадающую с поляризацией обыкновенной волны.

К общему случаю, если в нелинейный кристалл направить две световые волны с частотами $\omega_{1,2}$, то, как можно показать, на выходе можно получить излучение не только с частотами $2\omega_1, 2\omega_2$, но и на суммарной $\omega_s = \omega_1 + \omega_2$ (процесс сложения частот) и разностной частотах $\omega_d = \omega_1 - \omega_2$ (процесс вычитания). Эти процессы называют также трехволновым смещением частот. В этом случае условия фазового согласования имеют вид: $\vec{k}_{s,d} = \vec{k}_1 \pm \vec{k}_2$.

Аналогичным образом можно показать, что в кубичных средах возможно генерация третьей гармоники. В этом случае фазовое согласование имеет вид: $k_3 = 3k_1$, где k_3 — волновое число третьей гармоники. Кроме того, в таких средах изменяется показатель преломления и на основной частоте, причем нелинейная добавка пропорциональна интенсивности, которая может приводить к явлению *самофокусировки* света.

15.2. Самофокусировка света. Самофокусировка света обусловлена изменением показателя преломления среды под действием поля интенсивной световой волны. Пусть в нелинейную среду с показателем преломления n_0 входит мощный световой пучок амплитуды E . Внутри пучка среда становится *нелинейной*, так как при этом показатель преломления среды внутри пучка увеличивается ($n = n_0 + n_2 E^2$), т.е. под воздействием мощного светового пучка среда внутри пучка становится оптически более плотной. В лазерных пучках интенсивность света обычно максимальна у оси пучка и спадает к краям. Соответственно и показатель преломления n максимален на оси, постепенно уменьшается к краям, т.е. среда становится подобной собирающей (фокусирующей) линзе. В результате диаметр пучка по мере распространения уменьшается, так как если бы на его пути находилась фокусирующая линза. Этот эффект и называют *самофокусировкой*.

Фокусирующему действию среды препятствует дифракционная расходимость. Угловая расходимость, определяемая дифракцией, равна $\theta_d \simeq \lambda/n_0 D$, где D — диаметр пучка. Это означает, что пучок света ограниченного диаметра можно представить в виде суперпозиции плоских волн разных направлений, причем в этой сумме заметную амплитуду имеют лишь плоские волны, направления которых ограничены углом θ_d .

Так как внутри пучка среда оптически более плотная, то для лучей, падающих на граничную поверхность пучка изнутри, возможен эффект полного внутреннего отражения. Критический угол полного внутреннего отражения (между направлением луча и осью пучка) определяется соотношением $\cos \theta_0 = n_0/(n_0 + n_2 E^2)$, т.е.

$$\theta_0 = \arccos(n_0/(n_0 + n_2 E^2)) \quad (15.4)$$

Очевидно, что световые лучи, распространяющиеся под углом к оси пучка $\theta < \theta_0$, отражаясь от боковой поверхности пучка, должны отклоняться к его оси — *нелинейная рефракция*, т.е. среда становится подобной собирающей линзе. Поэтому диаметр пучка по мере распространения в среде должен уменьшаться. В действительности результат зависит соотношения между углом дифракционной расходимости и углом полного внутреннего отражения.

Когда $\theta_d > \theta_0$, дифракция преобладает над нелинейной рефракцией и пучок света будет расходиться, хотя и медленнее, чем в линейной среде.

При $\theta_d = \theta_0$ дифракция и нелинейная рефракция *компенсируют* друг друга, и в среде распространяется нерасходящийся цилиндрический пучок света — режим *самоканализации*. Образно можно сказать, что световой пучок создал в среде своеобразный *световод* и распространяется в нем. Когда $\theta_d < \theta_0$ происходит *самофокусировка*.

Пороговую мощность, при которой начинается самофокусировка, можно оценить из следующих соображений. Ситуация, когда $\theta_d = \theta_0$, является критической — ей и соответствует пороговое значение мощности. Это условие перепишем в виде

$$\arccos[n_0/(n_0 + n_2 E_p^2)] \simeq \lambda/n_0 D \quad (15.5)$$

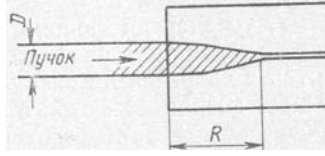


Рис. 15.3. Явление самофокусировки

Обычно $n_2 E_p^2 \ll n_0$, поэтому приближенно можно записать $\arccos(n_0/n_0 + n_2 E_p^2) \approx \sqrt{1 - n_0/(n_0 + n_2 E_p^2)} \approx \lambda/n_0 D$. Откуда получаем, что $E_p^2 \approx \lambda^2/(n_0 n_2 D^2)$. Эта формула определяет пороговое значение напряженности поля в световой волне, при которой пучок распространяется в среде без расходимости и его поперечное сечение остается неизменным. Этот эффект называют *самоканализацией*. Пороговая мощность, при которой начинается самофокусировка, равна $W_p = (cn_0/8\pi)E_p^2 \pi D^2 = c\lambda^2/8n_2$. Видно, что пороговая мощность световой волны уменьшается с увеличением частоты излучения и нелинейной кубической восприимчивости. Например, пороговая мощность для рубинового лазера в сероуглероде равна приблизительно $W_p = 20$ кВт.

При мощности $W > W_p$ световой пучок диаметра D *самофокусируется* на расстоянии R в среде и распространяется далее в виде световой нити с диаметром $d \ll D$, как показано на (**Рис.15.3**).

В действительности исследования показывают более сложную картину самофокусировки. При самофокусировке световая нить имеет более тонкую структуру— она распадается на ряд более тонких нитей диаметром 2-5 мкм, механизм возникновения которых до конца не ясен.

15.3. Параметрическая генерация света. Идея *параметрического усиления света* предложена С.А Ахмановым и Р.В.Хохловым в 1962 году. Пусть в квадратичной среде распространяются волны с частотами $\omega_{0,1}$. Волна с частотой ω_0 имеет большую амплитуду, достаточную для наведения квадратичной нелинейности. Эту волну называют *волной накачки*, а слабую волну с частотой ω_1 называют *сигнальной волной*. В результате нелинейного взаимодействия этих волн в среде возникает нелинейная (квадратичная) поляризованность, осциллирующая на частоте $\omega_2 = \omega_0 - \omega_1$, которая переизлучает в среде вторичные волны на этой частоте. Эту волну называют *холостой*. При выполнении условия синхронизма холостая и сигнальная волна будут эффективно усиливаться за счет перекачки энергии из волны накачки.

Действительно, пусть эти волна накачки, сигнальная и холостая имеют вид

$$E_{0,1,2}(z, t) = E_{0,1,2} \cos(\omega_{0,1,2}t - k_{0,1,2}z), \quad (15.6)$$

причем

$$\omega_0 = \omega_1 + \omega_2. \quad (15.7)$$

Квадратичную поляризованность среды на частотах сигнальной ω_1 и холостой $\omega_2 = \omega_0 - \omega_1$ волн можно записать в виде

$$P_1^{(2)} = \chi^{(2)} E_0 E_2 \cos[\omega_1 t - (k_0 - k_2)z], \quad P_2^{(2)} = \chi^{(2)} E_0 E_1 \cos[\omega_2 t - (k_0 - k_1)z]. \quad (15.8)$$

Поляризованность $P_2^{(2)}$ наводится в процессе взаимодействия накачки и сигнальной волн и приводит к переизлучению на частоте ω_2 . Соответственно $P_1^{(2)}$ возникает

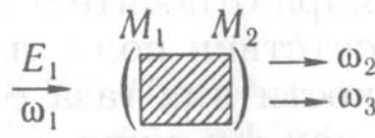


Рис. 15.4. Упрощенная схема параметрического генератора света.

при взаимодействии волн E_0 и E_1 , приводя к переизлучению на частоте ω_1 . При распространении волны $E_{1,2}$ *усиливаются* за счет энергии волны накачки E_0 , которая ослабляется

Для эффективного взаимодействия волны поляризации P_1 с переизлученной волной на частоте ω_2 необходимо выполнение условия $k_0 - k_2 = k_1$. Аналогично для другой волны $k_0 - k_1 = k_2$. Оба эти условия совпадают. Их принято записывать в виде

$$k_1 + k_2 = k_0. \quad (15.9)$$

Это условие есть условие *волнового согласования* для рассматриваемого случая. При выполнении условия волнового согласования (15.9) энергия волны накачки E_0 эффективно передается волнам E_1 и E_2 . Это явление называют *параметрической генерацией света*.

Качественно эффект возникновения комбинационных частот можно понять следующим образом. Мощная волна накачки с частотой ω_0 при распространении в среде приводит к изменению оптических свойств среды на частоте ω_0 , т.е. под действием этой волны возникает модуляция показателя преломления, а следовательно, модуляция фазы второй волны. Известно, что спектр колебаний, имеющего несущую частоту ω_2 , фаза которого промодулирована с частотой ω_0 , содержит компоненты с боковыми частотами $\omega_0 \mp \omega_1$.

Отметим, что частоту ω_1 , а следовательно, $\omega_2 = \omega_0 - \omega_1$ можно изменять по желанию экспериментатора, тем самым осуществляя параметрическую генерацию света.

Для получения параметрической генерации света достаточно в нелинейный кристалл направить волну накачки E_0 . Излучения $E_{1,2}$ возникнут и усилятся из шумов. Такой источник света называют *параметрическим генератором света (ПГС)*. Для того чтобы не использовать длинный кристалл, его помещают в резонатор, который представляет собой, в простейшем случае, два плоских зеркала, установленных параллельно. Слабое излучение с частотами $\omega_{1,2}$ существуют в резонаторе благодаря колебаниям осцилляторов, возбужденных тепловым движением.

Схема ПГС приведена на **Рис.15.4**. Нелинейный кристалл помещается внутри резонатора, образованного зеркалами $M_{1,2}$. Он ориентируется таким образом, чтобы для волны E_0 , а для некоторых волн $E_{1,2}$, распространяющихся вдоль оси резонатора перпендикулярно поверхности зеркал $M_{1,2}$, частоты которых удовлетворяют условию резонанса (15.7), выполнялось условие фазового согласования (15.9). В этом случае возникает усиление первоначально слабого спонтанного излучения на частотах $\omega_{1,2}$. Если, при достаточно мощной волне накачки, усиление волн $E_{1,2}$ превосходит потери за один проход в резонаторе, происходит параметрическая генерация. Меняя ориентацию кристалла, условие фазового согласования выполняется для другой пары частот, и таким образом, происходит плавная перестройка частоты.

Отметим, что условие $\omega_0 = \omega_1 + \omega_2$ эквивалентно $\hbar\omega_0 = \hbar\omega_1 + \hbar\omega_2$ ($\hbar$ — постоянная Планка). Так как $\hbar\omega$ — энергия фотона, то последнее условие отражает закон

сохранения энергии. Условие 15.9 можно переписать в виде $\hbar k_1 + \hbar k_2 = \hbar k_0$, которое можно трактовать как сохранение импульса для фотонов в среде.

В заключение подчеркнем роль когерентности света в нелинейно-оптических процессах, которая, как правило, имеют место при использовании именно лазерного излучения. *Высокая концентрация мощности, требуемая для проявления нелинейных свойств среды, возможна лишь при условии достаточно высокой пространственной когерентности излучения.* В этом смысле все нелинейно-оптические процессы требуют использования когерентного света. Но для эффективной генерации второй гармоники и подобных ей процессов световая волна накачки должна не только быть пространственно когерентной, но и иметь *высокую временную когерентность*. Последнее обусловлено чувствительностью волнового согласования к выбору направлению распространения и частот взаимодействующих волн. Исследования показывают существенную зависимость рассмотренных нелинейных процессов от степени когерентности оптического излучения.