

А.С. Перин, В.М. Шандаров

Основы физической и квантовой оптики

Учебное пособие

Томск
2018

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования

Томский государственный университет
систем управления и радиоэлектроники

А.С. Перин, В.М. Шандаров

Основы физической и квантовой оптики

Учебное пособие для студентов, обучающихся по направлению подготовки
11.03.02 «Инфокоммуникационные технологии и системы связи»,
11.04.02 «Инфокоммуникационные технологии и системы связи»,
12.04.03 «Фотоника и оптоинформатика»

Томск
2018

УДК 621.391.1.519.8(075.8)

ББК 32.88-01я73

П 274

Перин, Антон Сергеевич

П 274 Основы физической и квантовой оптики: учеб. пособие / А.С. Перин, В.М. Шандаров – Томск: Томск. гос. ун-т систем упр. и радиоэлектроники, 2018. – 195 с.

Учебное пособие включает лекционный материал, относящийся к избранным областям волновой оптики, квантовой электроники и нелинейной оптики. Изложены основные положения электромагнитной теории света, теории плоских электромагнитных волн в безграничной среде, рассмотрены элементы теории гауссовых световых пучков и элементы Фурье-оптики, основы электромагнитной теории диэлектрических волноводов и элементы квантовой электроники. В пособие включен также ряд вопросов, относящихся к современной оптике и рассматриваемых, как правило, в специальной научной литературе. Это элементы феноменологической теории электрооптического, фоторефрактивного эффектов и некоторые положения нелинейной оптики. Для облегчения усвоения материала в приложении приведены основные соотношения векторного анализа.

Предназначено для студентов всех форм обучения, обучающихся по направлению подготовки 11.03.02 «Инфокоммуникационные технологии и системы связи», 11.04.02 «Инфокоммуникационные технологии и системы связи», 12.04.03 «Фотоника и оптоинформатика».

УДК 621.391.1.519.8(075.8)

ББК 32.88-01я73

© А.С. Перин, В.М. Шандаров, 2018

© Томск. гос. ун-т систем упр.

и радиоэлектроники, 2018

Содержание

Введение	3
1. Основные положения электромагнитной теории света	4
1.1. Историческая справка об эволюции взглядов на природу света	5
1.2. Способы описания электромагнитного излучения	7
1.3. Уравнения Максвелла	9
1.3.1. Уравнения Максвелла в интегральной форме.....	9
1.3.2. Уравнения Максвелла в дифференциальной форме.....	11
1.3.3. Уравнение непрерывности полного тока.....	13
1.3.4. Материальные уравнения.....	14
1.4. Уравнения граничных условий для векторов электромагнитного поля	17
1.4.1. Граничные условия для векторов электрического поля....	18
1.4.2. Граничные условия для векторов магнитного поля.....	20
1.5. Баланс энергии электромагнитного поля	22
1.5.1. Уравнение баланса энергии (формула Умова-Пойнтинга)	22
1.5.2. Электрическая и магнитная энергия электромагнитного поля.....	24
1.5.3. Перенос энергии и групповая скорость в электромагнитном поле.....	24
1.6. Символический метод (метод комплексных амплитуд).....	25
1.7. Комплексные проницаемости	27
1.8. Уравнения Максвелла в символической форме.....	28
1.9. Вектор Пойнтинга в символической форме.....	28
1.10. Выражение для вектора Пойнтинга в символической форме...	29
1.11. Энергетические характеристики монохроматического ЭМП...	30
1.12. Формула для баланса энергии монохроматического ЭМП.....	31
2. Плоские электромагнитные волны в безграничной среде.....	33
2.1. Волновое уравнение для безграничной среды.....	33
2.2. Решение волнового уравнения. Плоские волны.....	34
2.3. Гармонические волны	35
2.4. Распространение плоской волны в произвольном направлении...	37
2.5. Структура поля электромагнитных волн.....	38
2.6. Поляризация плоских электромагнитных волн.....	41
2.7. Вектор Джонса.....	43
2.8. Элементы для преобразования состояния поляризации света....	44
2.9. Отражение и преломление плоских световых волн на плоской границе раздела.....	45
2.9.1. Законы отражения и преломления электромагнитных волн	46
2.9.2. Формулы Френеля.....	49
2.9.3. Явление полного внутреннего отражения.....	

2.10. Когерентность световых волн.....	52
2.10.1. Понятие полной когерентности.....	52
2.10.2. Временная когерентность.....	54
2.10.3. Функция взаимной когерентности.....	54
2.10.4. Временная и пространственная когерентность.....	55
3. Оптика ограниченных световых пучков	57
3.1. Угловой спектр плоских волн.....	57
3.2. Параболическое уравнение.....	60
3.3. Круговой гауссов пучок.....	61
3.4. Угловой спектр гауссова пучка.....	64
3.5. Гауссовы пучки высших порядков.....	65
3.6. Элементы Фурье – оптики.....	66
3.6.1. Преобразование Фурье в когерентной оптической системе.	67
3.6.2. Некоторые математические операции, реализуемые в оптической системе.....	71
3.6.3. Обратное преобразование Фурье оптического сигнала.....	72
3.6.4. Пространственная фильтрация в когерентной оптической системе.....	73
4. Распространение оптических волн в материальных средах.....	74
4.1. Дисперсия показателя преломления диэлектрической среды.....	74
4.2. Распространение света в направляющих структурах.....	79
4.2.1. Электромагнитная теория планарного волновода.....	79
4.2.2. Дисперсионное уравнение планарного волновода.....	83
4.2.3. Анализ дисперсионного уравнения.....	86
5. Элементы квантовой оптики.....	88
5.1. Постоянная Планка.....	88
5.2. Постулаты Бора.....	89
5.3. Корпускулярно – волновой дуализм.....	89
5.4. Физическая интерпретация волн де Бройля.....	91
5.5. Соотношения неопределенностей.....	92
5.6. Уравнение Шредингера.....	93
5.7. Операторы.....	96
5.8. Гармонический осциллятор.....	101
5.9. Взаимодействие излучения с атомными системами.....	103
5.9.1. Энергетические уровни.....	103
5.9.2. Квантовые переходы.....	104
5.9.3. Соотношения между коэффициентами Эйнштейна.....	107
5.10. Форма и ширина спектральной линии.....	108
5.10.1. Естественная ширина спектральной линии.....	108
5.10.2. Уширение спектральной линии из-за столкновений.....	110
5.10.3. Допплеровское уширение спектральной линии.....	110 111

5.10.4. Другие механизмы уширения спектральных линий.....	112
5.11. Отрицательная температура.....	115
5.12. Взаимодействие бегущих электромагнитных волн с активной средой.....	117
5.13. Методы достижения инверсии населенностей.....	119
6. Принципы построения лазеров.....	119
6.1. Открытые оптические резонаторы.....	119
6.1.1. Элементарная теория открытых резонаторов.....	121
6.1.2. Типы волн в реальных открытых резонаторах.....	122
6.1.3. Добротность типов колебаний открытого резонатора.....	124
6.2. Условия самовозбуждения лазера.....	125
6.3. Газовые лазеры.....	127
6.3.1. Лазер на смеси гелия и неона (He – Ne лазер).....	129
6.3.2. Ионный аргоновый лазер.....	130
6.3.3. Молекулярные газовые лазеры.....	131
6.4. Лазеры на ионных кристаллах и стеклах.....	131
6.4.1. Элементы накачки твердотельных лазеров.....	132
6.4.2. Рубиновый лазер.....	135
7. Взаимодействия света с физическими полями.....	135
7.1. Электрооптический эффект.....	136
7.1.1. Феноменологическое описание электрооптического эффекта.....	139
7.1.2. Электрооптическая модуляция фазы световой волны.....	141
7.1.3. Пример электрооптического модулятора интенсивности света.....	142
7.2. Акустооптический эффект.....	144
7.2.1. Феноменологическое описание акустооптического эффекта	145
7.2.2. Типы акустических волн в твердом теле.....	147
7.2.3. Основные представления об акустооптическом взаимодействии.....	
7.2.4. Корпускулярная (квантово-механическая) интерпретация акустооптического взаимодействия.....	149
7.2.5. Иллюстрация векторными треугольниками дифракции Рамана – Ната и брэгга.....	151
7.2.6. Эффективность дифракционных процессов.....	152
7.2.7. Акустооптический модулятор.....	154
7.3. Фоторефрактивный эффект.....	156
7.3.1. Механизмы пространственного перераспределения носителей электрического заряда.....	157
7.3.2. Модель фоторефракции.....	159

7.3.3. Схема формирования элементарных фоторефрактивных решеток.....	160
7.3.4. Стирание (релаксация) решетки.....	162
7.3.5. Некоторые характеристики ФРЭ и фоторефрактивных материалов.....	163
8. Элементы нелинейной оптики.....	
8.1. Элементы теории нелинейно – оптических явлений.....	164
8.1.1. Квадратичная и кубичная оптические нелинейности.....	167
8.1.2. Возможные типы нелинейно–оптических эффектов в средах с квадратичной и кубичной нелинейностью.....	167
8.2. Нелинейное волновое уравнение.....	171
8.3. Генерация второй гармоники.....	
8.4. Условия фазового синхронизма при генерации второй гармоники.....	173
8.5. Распространение светового пучка в нелинейно – оптической среде.....	178
8.5.1. Пространственные оптические солитоны в среде с керровской нелинейностью.....	179
8.5.2. Пространственные оптические солитоны в среде с фоторефрактивной нелинейностью.....	180
8.6. Другие нелинейно – оптические эффекты.....	182
Рекомендуемая литература	183
Содержание.....	

ВВЕДЕНИЕ

Учебное пособие отражает основные разделы лекционного курса по дисциплине. Цель пособия – изложить основные положения современной оптики и квантовой электроники, знание и глубокое понимание которых необходимо как специалистам – разработчикам, так и специалистам по эксплуатации оборудования и систем оптической связи. Материал курса разбит на 8 тематических разделов, относящихся к областям физической оптики, квантовой электроники (оптики) и нелинейной оптики. Это «Основные положения электромагнитной теории света», «Плоские электромагнитные волны в безграничной среде», «Оптика ограниченных пучков», «Распространение оптических волн в материальных средах», «Элементы квантовой оптики», «Принципы построения лазеров», «Взаимодействие света с физическими полями», «Элементы нелинейной оптики». Учебное пособие может быть полезно студентам, обучающимся по направлению подготовки 11.03.02 «Инфокоммуникационные технологии и системы связи», 11.04.02 «Инфокоммуникационные технологии и системы связи», 12.04.03 «Фотоника и оптоинформатика»

1.1. 1. ОСНОВНЫЕ ПОЛОЖЕНИЯ ЭЛЕКТРОМАГНИТНОЙ ТЕОРИИ СВЕТА

Оптические методы и приборы широко используются в современной науке и технике как для исследования вещества и всевозможных физических явлений, так и в промышленном производстве, либо при эксплуатации машин и систем различного назначения.

Оптические явления - это явления взаимодействия света с веществом. При рассмотрении такого взаимодействия можно использовать три подхода: а) классический; б) полуклассический; в) квантовый.

Классический подход подразумевает представление оптического излучения в виде световых лучей или электромагнитных волн соответствующего диапазона частот. Вещество при этом рассматривается как сплошная среда с некоторыми макроскопическими электродинамическими параметрами ϵ , μ , σ . Как свет, так и вещество при таком подходе рассматриваются в рамках классической физики.

При **полуклассическом** подходе учитывается структура вещества, т.е. структура энергетических уровней атомов и молекул, энергетических зон кристаллов, статистика заполнения различных квантовых состояний. В этом случае используется принцип квантования вещества при классической трактовке света, представляемого в виде электромагнитных волн.

При **квантовом** подходе осуществляется квантование и вещества, и излучения. Такой подход используется в квантовой электродинамике. Здесь, наряду с дискретностью структуры вещества, учитываются корпускулярные свойства света. Это соответствует переходу от классической оптики к оптике, которую можно назвать квантовой оптикой. Одним из основных ее понятий является **фотон** – квант светового поля.

Название данного курса – “**Основы физической и квантовой оптики**”. Нужно заметить, что перечень вопросов, объединяемых термином “квантовая оптика”, не определен строгими рамками. Принято относить сюда вопросы квантования электромагнитного поля, анализа различных состояний квантованного поля, рассмотрения когерентных свойств света. В данном курсе, наряду с рядом явлений, имеющих непосредственное отношение к квантовой оптике (взаимодействие света с квантовыми системами, квантовые переходы, лазерные эффекты), рассматривается ряд явлений, при анализе которых возможно использование полуклассического и классического подходов (эту часть материала можно отнести к разделу физической оптики).

1.2. Историческая справка об эволюции взглядов на природу света

Вопрос о природе света волновал человека в течение многих веков, с самых древних времен. Представления о физической сущности света изменялись, и лишь в XX веке развитие квантовой физики, кажется, расставило все точки над *i*. Однако не надо забывать об ограниченной глубине степени познания реальных физических явлений человеческим разумом в каждый конкретный момент времени. Поэтому вполне возможно, что через какое – то время представление о природе света может обогатиться новыми знаниями и совершить еще один шаг в направлении той бесконечной глубины, которая никогда не будет достигнута, но всегда будет манить...

Итак, что же такое свет?

Часть мыслителей древности ассоциировала ощущение света с истечением из глаз человека неких лучей, которые как бы ощупывают наблюдаемый объект, и, отражаясь от него, улавливаются глазом и дают разуму определенное представление об объекте (в современной технике это называется активной радиолокацией!). К числу сторонников этой теории относились последователи Пифагора и такие философы, как Аристотель.

Существовало также представление о свете, как о результате воздействия на глаз мелких “атомов” (корпускул), испускаемых светящимся телом (Демокрит, Лукреций, и их последователи). Естественно, что на том уровне знаний было сложно обосновать подобные гипотезы о природе света экспериментальными результатами и развитыми теоретическими моделями. Мыслители древности исходили, в основном, из интуитивных умозаключений, основанных на субъективном восприятии реального мира.

Значительный прогресс в развитии представлений о природе света достигнут в XVII веке, что было обусловлено бурным развитием науки (в частности, механики) в этот период. “Отец” классической механики Исаак Ньютон последовательно развивал корпускулярную теорию света, считая, что свет – это поток корпускул, которые либо выбрасываются, либо отражаются телами и летят с огромной скоростью. Они предполагались разных размеров и массы: наибольшие размеры и масса, как считалось, соответствуют темно – красным лучам, наименьшие – фиолетовым. Ньютон попытался вести в корпускулярную модель элементы волнового движения. Он допустил, что световые корпускулы движутся в светоносном эфире, возбуждая эфирные волны, распространяющиеся значительно быстрее самих корпускул. Возможно, в этом проявилось гениальное предвидение корпускулярно – волнового дуализма света?

Оппоненты Ньютона (Роберт Гук, Христиан Гюйгенс, Михаил Ломоносов) считали, что свет – это некие волны или импульсы, бегущие в

упругом эфире, заполняющем все пространство. Их аргументы сводились к следующим вопросам:

- Почему не взаимодействуют два пучка корпускул при их пересечении?
- Как могут материальные тела (корпускулы) двигаться с такой огромной скоростью (скорость света была измерена Рёмером в 1676 г.)?
- Если свет - это поток корпускул, то почему мы не замечаем уменьшения массы Солнца?

По Гюйгенсу свет – это передача импульсов возбуждения в эфире, заполняющем все пространство. Постулируемая несжимаемость эфира объясняла высокую скорость распространения света и свободное прохождение одного светового пучка через другой.

Конечно, обе теории были механистическими. Это было обусловлено успехами классической механики, и попытками использовать ее стройный математический аппарат для интерпретации множества физических явлений.

В XIX веке волновая концепция, казалось, окончательно победила, благодаря результатам исследований Августина Жака Френеля и Томаса Юнга, заложивших основы волновой оптики.

Юнг открыл явление интерференции, что объясняло цвета тонких пленок, цвета побежалости на поверхности металлов, эффект колец Ньютона и т.д. Френель уточнил волновой принцип Гюйгенса, дополнив его принципом интерференции Юнга (принцип Гюйгенса – Френеля). На основе этого принципа он обстоятельно исследовал дифракцию света.

Особый интерес представляли эксперименты по интерференции поляризованного света (А.Ж.Френель, Д.Ф. Араго – 1816 г.), в результате которых Френель пришел к выводу о поперечности световых волн.

Наконец, благодаря работам М. Фарадея и некоторых других ученых, Джеймс Кларк Максвелл дал убедительное математическое обоснование электромагнитной природы света, записав в 1864 году свои знаменитые уравнения. В 1888 г. Г.Герц экспериментально подтвердил факт существования электромагнитных волн, получив волны метрового диапазона.

Однако победа волновой концепции природы света не являлась безоговорочной, поскольку существовали некоторые проблемы, которые не могли быть разрешены с позиций волновой теории.

Действительно, в физике в конце XIX - начале XX веков классический электродинамический подход не позволял разрешить следующие проблемы:

- проблему теоретического описания излучения абсолютно черного тела;
- проблему фотоэлектрического эффекта, в частности существования красной границы фотоэффекта;
- проблему объяснения стабильности и размера атомов.

Основная причина невозможности разрешения данных вопросов с позиций классической физики заключалась в том, что **классическая электродинамика успешно объясняет лишь те оптические явления, где**

несущественна структура элементарных процессов взаимодействия излучения с веществом.

Попытки решения указанных проблем, в конце концов, увенчались успехом и привели к кардинальному углублению понимания природы света и рождению новой области физики – квантовой физики. Фундаментальным положением квантовой физики и квантовой оптики является положение о корпускулярно – волновом дуализме, согласно которому, в частности, свет может проявлять в разных ситуациях как корпускулярные, так и волновые свойства. Целый ряд оптических явлений может быть объяснен с точки зрения волновой природы света. Все эти явления принято относить к области физической оптики. Рассмотрение некоторых вопросов физической оптики составляет первую часть данного курса.

1.2. Способы описания электромагнитного излучения

Оптическое излучение, строго говоря, следует рассматривать как электромагнитное поле, подчиняющееся квантовым закономерностям. При рассмотрении различных оптических явлений можно ограничиться приближенными представлениями, описывающими световое поле в виде:

- а) световых лучей;
- б) световых волн;
- в) фотонных коллективов.

Свет с позиций геометрической оптики.

Световые лучи или совокупности лучей, образующие световые пучки, соответствуют приближению геометрической оптики. В этом случае не принимается во внимание конечность длины световой волны и считается, что световая энергия распространяется вдоль определенных кривых, называемых лучами. В однородной среде лучи представляют собой прямые линии. Световые лучи не взаимодействуют друг с другом и после пересечения распространяются независимо.

Основу геометрической оптики составляет принцип Ферма или принцип наименьшего оптического пути (принцип наименьшего времени). Его суть заключается в том, что оптическая длина пути (ОДП), по которому распространяется свет от точки A до точки B , короче оптического пути любой другой кривой, соединяющей эти точки.

Величина ОДП вдоль отрезка длиной dl в среде с показателем преломления n определяется как $n \cdot dl$, поэтому для некоторой кривой AB ее можно найти из соотношения: $ОДП = \int_A^B n \cdot dl$. Если среда оптически однородна, то $n = const$ и $ОДП = n \cdot l_{AB}$, где l_{AB} – длина кривой. Поскольку

$dl = \frac{c}{n} dt$, то $ОДП = \int_A^B c \cdot dt = c \cdot \int_A^B dt$. Последнее выражение поясняет суть определения “принцип наименьшего времени”. Из него, в частности, вытекает факт прямолинейности световых лучей в оптически однородной среде.

Свет с позиций волновой теории.

Явления дифракции и интерференции света не могут быть объяснены в рамках геометрической оптики и свидетельствуют о его волновой природе. Основой математического аппарата волновой оптики являются уравнения Максвелла, использующие феноменологический подход. Свойства вещества при этом характеризуются диэлектрической проницаемостью ϵ , магнитной проницаемостью μ , удельной проводимостью σ . Эти характеристики отражают реакцию среды на внешнее электромагнитное возмущение. К основным свойствам световых волн можно отнести следующие:

- а) в вакууме они распространяются со скоростью $c = (\epsilon_0 \mu_0)^{-1/2}$;
- б) световые волны имеют поперечную структуру;
- в) они переносят энергию, плотность потока и направление движения которой определяются вектором $\bar{P} = [\bar{E}, \bar{H}]$;
- г) свет может быть поляризованным или неполяризованным;
- д) все известные приемники оптического излучения реагируют на интенсивность (величина, численно равная энергии, переносимой волной за единицу времени через единичную площадку, перпендикулярную направлению распространения света), вследствие очень высокой частоты световых волн.

Свет с позиций квантовой теории.

Световое поле – это совокупность квантов электромагнитного излучения (фотонов). К основным свойствам фотонов можно отнести следующие:

- а) фотон обладает энергией $E = h\omega$ и импульсом $\bar{p} = h\bar{k}$;
- б) масса покоя фотона равна нулю, а его фазовая скорость равна скорости света;
- в) фотон не имеет электрического заряда;
- г) спин фотона равен единице;
- д) фотон характеризуется некоторым состоянием поляризации;
- е) фотоны относятся к Бозе – частицам, они подчиняются статистике Бозе – Эйнштейна, поэтому неограниченное число фотонов может находиться в одинаковых квантовых состояниях.

1.3. Уравнения Максвелла

1.3.1. Уравнения Максвелла в интегральной форме

В классической теории электромагнитного поля учитываются только макроскопические свойства вещества. В этом случае все расстояния и размеры элементарных областей пространства при рассмотрении считаются гораздо большими, чем размеры атомов или величина межатомных расстояний в твердом теле. Материальные среды рассматриваются как сплошные, без учета их микроструктуры. Фактически это значит, что вместо реальных значений величин, характеризующих электромагнитное поле в некоторой точке пространства, мы используем некоторые их значения, полученные путем усреднения по элементарному объему, включающему достаточно большое число атомов вещества. Отметим, что в данном курсе для представления физических величин используется международная система единиц (СИ), что также определяет и форму записи уравнений Максвелла и некоторых других соотношений. Будем далее использовать термины “световое поле, световые волны”, имея в виду электромагнитные поля и волны соответствующего диапазона частот в диэлектрических средах.

Электромагнитное поле (ЭМП) характеризуется четырьмя векторными величинами. Это:

- \bar{E} - напряженность электрического поля (единица измерения - В/м);
- \bar{D} - электрическая индукция (единица измерения - Кл/м²);
- \bar{H} - напряженность магнитного поля (единица измерения - А/м);
- \bar{B} - магнитная индукция (единица измерения - Вб/м²).

Материальность ЭМП проявляется в его силовом воздействии на пробный электрический заряд. Как известно, сила, действующая на такой заряд, определяется соотношением:

$$\bar{F} = q \cdot \bar{E} + q \cdot [\bar{v}, \bar{B}] \quad (1.1),$$

где q – величина электрического заряда; \bar{v} - вектор его скорости. Как следует из (1.1), магнитная составляющая полного ЭМП может быть обнаружена лишь с помощью движущихся электрических зарядов. Сила кулоновского взаимодействия, обусловленная электрической составляющей ЭМП, не зависит от состояния пробного заряда (неподвижный или движущийся). Определить ЭМП в некоторой области – значит указать величину и направление всех векторов поля в любой ее точке. Для описания ЭМП используется система уравнений Максвелла, которые являются результатом обобщения накопленных в течение продолжительного времени знаний в области электромагнитных явлений.

Система уравнений Максвелла в интегральной форме имеет вид:

Первое уравнение:
$$\oint_l \vec{H} \cdot d\vec{l} = I_{\Sigma} \quad (1.2),$$

т.е. циркуляция напряженности магнитного поля по замкнутому контуру равна сумме токов, пронизывающих этот контур. Это закон полного тока, теоретически сформулированный Ампером на основе экспериментов Эрстеда по обнаружению магнитного поля вблизи проводника с электрическим током. Здесь $d\vec{l}$ – направленный элемент длины; $I_{\Sigma} = I_{\text{пр}} + I_{\text{см}}$, где $I_{\text{пр}}$ – ток проводимости, а $I_{\text{см}}$ – ток смещения. Таким образом, магнитное поле может порождаться как токами проводимости, так и токами смещения.

Второе уравнение:
$$\oint_l \vec{E} \cdot d\vec{l} = -\frac{\partial \Phi}{\partial t} \quad (1.3)$$

- циркуляция напряженности электрического поля по замкнутому контуру (а это не что иное, как ЭДС электромагнитной индукции) равна производной по времени от потока магнитной индукции, который пронизывает контур. Это закон электромагнитной индукции Фарадея, который в 1831 г. экспериментально обнаружил эффект появления электрического напряжения на зажимах катушки, помещенной в переменное магнитное поле. Он показывает, что всякое изменение магнитного поля приводит к возникновению электрического поля.

Третье уравнение:
$$\oint_S \vec{D} \cdot d\vec{S} = \sum q \quad (1.4)$$

- поток вектора электрической индукции через замкнутую поверхность равен алгебраической сумме электрических зарядов, охватываемых данной поверхностью – закон Гаусса, найденный экспериментально. Соотношение (1.4) указывает, что силовые линии электрического поля могут начинаться или заканчиваться на электрических зарядах (истоках и стоках). Если зарядов во всей области поля нет, то силовые линии замкнуты или уходят в бесконечность. Иллюстрацию этих положений представляет рис. 1.1.

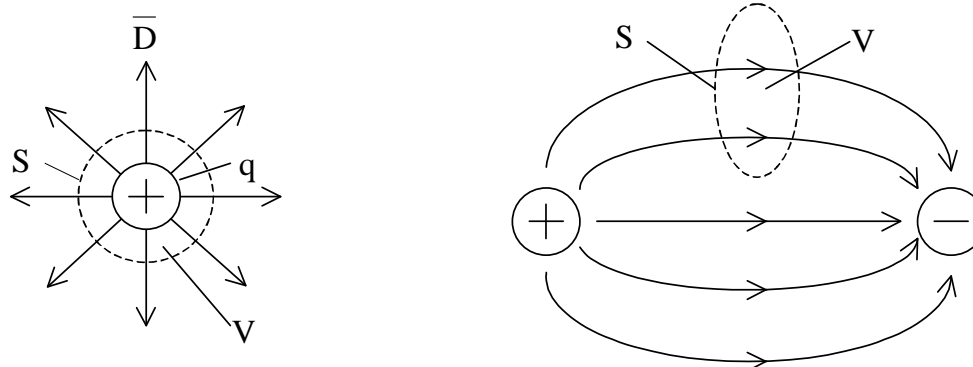


Рис. 1.1.

Четвертое уравнение:
$$\oint_S \vec{B} \cdot d\vec{S} = 0 \quad (1.5)$$

указывает на отсутствие в природе магнитных зарядов (поток вектора магнитной индукции через любую замкнутую поверхность всегда равен нулю, а магнитные силовые линии всегда замкнуты). Это иллюстрирует рис.

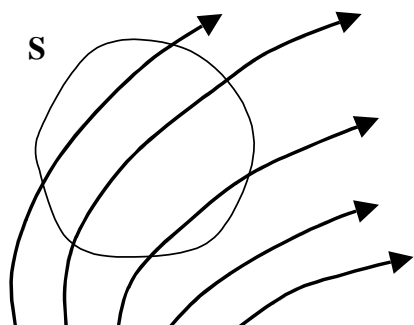


Рис. 1.2.

1.2, где число силовых линий магнитного поля, входящих в область, ограниченную поверхностью S , равно числу выходящих из нее линий.

От данных уравнений можно легко перейти к их дифференциальной форме. В таком виде они описывают локальные характеристики поля в любой произвольной точке пространства и потому более удобны для описания пространственных зависимостей векторов ЭМП. При переходе используются

теоремы Остроградского - Гаусса и Стокса.

1.3.2. Уравнения Максвелла в дифференциальной форме

При переходе к дифференциальной форме уравнений Максвелла используем следующие соотношения между дифференциальными и интегральными характеристиками ЭМП.

А) Величина тока I , протекающего через некоторое сечение, представляет собой поток вектора плотности тока $\vec{\delta}$ через это сечение:

$$I = \int_S \vec{\delta} \cdot d\vec{S} \quad (1.6)$$

Б) Величина электрического заряда q в некоторой области пространства может быть выражена в виде объемного интеграла от объемной плотности заряда ρ в данной области:

$$q = \int_V \rho \cdot dV \quad (1.7)$$

В) Величина магнитного потока Φ через площадку, ограниченную некоторым контуром, представляется потоком вектора магнитной индукции \vec{B} через эту площадку:

$$\Phi = \int_S \vec{B} \cdot d\vec{S} \quad (1.8)$$

Представим (1.2) в форме:

$$\oint_l \vec{H} \cdot d\vec{l} = I_\Sigma = \int_S \vec{\delta}_\Sigma \cdot d\vec{S}, \quad \text{где } \vec{\delta}_\Sigma - \text{плотность полного тока, а } S - \text{поверхность,}$$

стягиваемая контуром l . Согласно теореме Стокса, левая часть этого соотношения преобразуется к виду:

$\oint_l \bar{H} \cdot d\bar{l} = \int_S \text{rot} \bar{H} \cdot d\bar{S}$, т.е. $\int_S \text{rot} \bar{H} \cdot d\bar{S} = \int_S \bar{\delta}_\Sigma \cdot d\bar{S}$. Последнее равенство выполняется для произвольной поверхности S и из него вытекает равенство подынтегральных выражений:

$$\text{rot} \bar{H} = \bar{\delta}_\Sigma \quad (1.9).$$

Здесь $\bar{\delta}_\Sigma$ выражает сумму плотностей всех токов. В дополнение к току проводимости, Максвелл ввел в рассмотрение ток смещения, не связанный с движением электрических зарядов, а характеризующий прохождение переменного тока по цепи, содержащей емкости. При этом $\bar{\delta}_{см} = \frac{\partial \bar{D}}{\partial t}$.

Действительно, если вокруг одной из обкладок конденсатора построим воображаемую поверхность (рис. 1.3), то, согласно закону Гаусса:

$\oint_S \bar{D} \cdot d\bar{S} = q$, где q – заряд на обкладке. Ток в рассматриваемой цепи

$$I = \frac{dq}{dt} = \frac{d}{dt} \oint_S \bar{D} \cdot d\bar{S} = \oint_S \frac{d\bar{D}}{dt} \cdot d\bar{S}, \text{ но } I = \int_S \bar{\delta} \cdot d\bar{S}, \text{ поэтому величина } \frac{d\bar{D}}{dt} \text{ имеет}$$

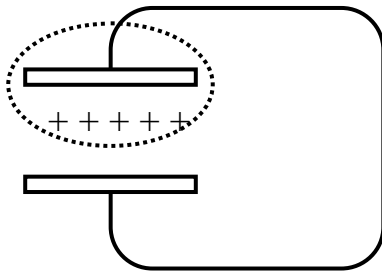


Рис. 1.3.

размерность плотности тока, который Максвелл назвал током смещения. Таким образом, (1.9) принимает вид:

$$\text{rot} \bar{H} = \frac{\partial \bar{D}}{\partial t} + \bar{\delta} \quad (1.10).$$

Аналогично, (1.3) приведем к виду:

$$\text{rot} \bar{E} = -\frac{\partial \bar{B}}{\partial t} \quad (1.11),$$

используя соотношения: $\Phi = \int_S \bar{B} \cdot d\bar{S}$ и $\frac{\partial \Phi}{\partial t} = \int_S \frac{\partial \bar{B}}{\partial t} \cdot d\bar{S}$.

Третье и четвертое уравнения Максвелла (1.4) и (1.5) преобразуем, используя формулу Остроградского - Гаусса:

$\oint_S \bar{D} \cdot d\bar{S} = \int_V \text{div} \bar{D} \cdot dV$, а $\sum q = \int_V \rho \cdot dV$, где ρ – объемная плотность

электрического заряда. В итоге получаем:

$$\text{div} \bar{D} = \rho \quad (1.12).$$

И, наконец, из (1.5) получим:

$$\text{div} \bar{B} = 0 \quad (1.13).$$

В случае светового поля, как отмечалось, среда предполагается диэлектрической, поэтому токи проводимости и электрические заряды

отсутствуют. Соответственно, уравнения Максвелла в дифференциальной форме для светового поля имеют вид:

$$\operatorname{rot} \bar{H} = \frac{\partial \bar{D}}{\partial t} \quad (1.14),$$

$$\operatorname{rot} \bar{E} = -\frac{\partial \bar{B}}{\partial t} \quad (1.15),$$

$$\operatorname{div} \bar{D} = 0 \quad (1.16),$$

$$\operatorname{div} \bar{B} = 0 \quad (1.17).$$

1.3.3. Уравнение непрерывности полного тока

Важным дополнением к системе уравнений Максвелла является уравнение непрерывности, выражающее, по сути, закон сохранения электрического заряда. Его легко получить из первого уравнения Максвелла в дифференциальной форме. Применим к обеим его частям операцию дивергенции, что с учетом соотношения $\operatorname{div}(\operatorname{rot} \bar{a}) \equiv 0$ дает:

$$\operatorname{div} \left(\frac{\partial \bar{D}}{\partial t} + \bar{\delta} \right) = 0 \quad (1.18).$$

Меняя далее порядок операций дифференцирования по пространственным координатам и времени, а также используя соотношение $\operatorname{div} \bar{D} = \rho$, получим уравнение непрерывности в форме:

$$\operatorname{div} \bar{\delta} = -\frac{d\rho}{dt} \quad (1.19).$$

Его соответствие закону сохранения заряда становится очевидным, если проинтегрировать его обе части по объему:

$$\int_V \operatorname{div} \bar{\delta} \cdot dV = -\int_V \frac{d\rho}{dt} \cdot dV = -\frac{d}{dt} \int_V \rho \cdot dV.$$

С другой стороны, используя формулу Остроградского – Гаусса и выражение для электрического заряда через его объемную плотность, получим:

$$\int_V \operatorname{div} \bar{\delta} \cdot dV = \oint_S \bar{\delta} \cdot \bar{dS} = I, \quad \text{а} \quad -\frac{d}{dt} \int_V \rho \cdot dV = -\frac{dq}{dt}.$$

Таким образом, получено соотношение: $I = -dq/dt$, которое и выражает закон сохранения заряда. Видим, что любое изменение величины электрических зарядов в выбранном объеме сопровождается возникновением тока проводимости, втекающего в данный объем или вытекающего из него, в зависимости от того, увеличивается или уменьшается величина заряда.

1.3.4. Материальные уравнения

Материальные уравнения учитывают влияние среды на связь между векторами поля. Их вывод должен учитывать ее микроскопические свойства, т.е. атомно – молекулярную структуру. Однако строгое определение такой связи представляет собой слишком сложную, в большинстве случаев неразрешимую, задачу. Поэтому для описания влияния среды используется феноменологический подход, т.е. вводятся некоторые макроскопические параметры, определяющие связь между характеристиками ЭМП в веществе. К данным параметрам относятся диэлектрическая (ϵ) и магнитная (μ) проницаемости, а также удельная проводимость среды (σ). Основные свойства этих параметров зависят от типа материальной среды, а их численные значения, как правило, находятся экспериментальным путем.

При определении основных свойств материальных параметров используют идеализированные модели материальной среды, выделяя следующие типы:

а) *Линейные и нелинейные среды.* К линейным относят среды, параметры которых не зависят от амплитуды поля. Соответственно, в нелинейных средах их параметры ϵ , μ , σ являются функциями поля. Т.е. $\epsilon = \epsilon(E)$, $\mu = \mu(H)$, $\sigma = \sigma(E)$. Строго говоря, все материальные среды являются нелинейными, однако их нелинейность проявляется только при достаточно сильных полях и лишь в некоторых специальных случаях возможно ее наблюдение в слабых полях. В случае оптически нелинейных сред от амплитуды светового поля зависит диэлектрическая проницаемость. Отметим, что наблюдение и практическое использование нелинейно – оптических эффектов стало возможным после появления лазеров.

б) *Однородные и неоднородные среды.* Под однородной понимается среда, материальные параметры которой не зависят от пространственных координат. Если же такая зависимость наблюдается хотя бы для одного из параметров ϵ , μ , σ , то среда является неоднородной.

в) *Изотропные и анизотропные среды.* Изотропными называются среды, материальные параметры которых не зависят от направления, в противном случае они называются анизотропными.

г) *Среды с пространственной и временной дисперсией.* О дисперсии говорят в случае нелокального отклика среды на воздействие. Под нелокальным откликом понимается ситуация, когда значение вектора \bar{D} (или \bar{B}) в данной точке пространства в данный момент времени может зависеть от значения \bar{E} (или \bar{H}) в предыдущие моменты времени либо в других точках пространства.

Рассмотрим соотношения между \bar{D} и \bar{E} , \bar{B} и \bar{H} для разных типов материальных сред.

А. Вакуум.

В этом случае для определения ЭМП достаточно двух векторов: \vec{E} и \vec{B} , а \vec{D} и \vec{H} задаются соотношениями:

$$\vec{D} = \epsilon_0 \cdot \vec{E} \quad (1.20),$$

$$\vec{B} = \mu_0 \cdot \vec{H} \quad (1.21),$$

где ϵ_0 - универсальная электрическая постоянная (или диэлектрическая проницаемость вакуума) - $\epsilon_0 = 8,85 \cdot 10^{-12} = (1/36\pi) \cdot 10^{-9}$ Ф/м; μ_0 - универсальная магнитная постоянная (магнитная проницаемость вакуума) - $\mu_0 = 4\pi \cdot 10^{-7}$ Гн/м.

Б. Изотропная среда.

Это наиболее простой случай. Материальные уравнения для изотропной среды имеют вид:

$$\vec{D} = \epsilon_0 \cdot \vec{E} + \vec{P} = \epsilon_0 \vec{E} + \epsilon_0 \chi \vec{E} = \epsilon_0 (1 + \chi) \vec{E} = \epsilon \cdot \vec{E} \quad (1.22),$$

$$\vec{B} = \mu_0 \vec{H} + \vec{M} = \mu_0 \vec{H} + \mu_0 \chi_M \vec{H} = \mu_0 (1 + \chi_M) \vec{H} = \mu \cdot \vec{H} \quad (1.23),$$

где \vec{P} - вектор поляризации среды; \vec{M} - намагниченность среды; χ - диэлектрическая восприимчивость; χ_M - магнитная восприимчивость среды; ϵ , μ - абсолютные диэлектрическая и магнитная проницаемости. Смысл этих соотношений сводится к следующему. В отсутствие электрического поля все атомы и молекулы вещества электрически нейтральны. Под действием поля \vec{E} происходит деформация орбит электронов (с позиций классической физики), т.е. центры “тяжести” положительного и отрицательного зарядов в атоме смещаются в направлении поля в разные стороны. В итоге атом может быть представлен в виде электрического диполя. То же можно сказать о молекулах вещества, не имеющих собственного дипольного момента. Если молекулы вещества являются полярными, т.е. имеют некоторый собственный дипольный момент, то в отсутствие поля они ориентированы в пространстве хаотично. Под действием поля они ориентируются вдоль его направления. Таким образом, в электрическом поле атом или молекула могут характеризоваться дипольным моментом $\vec{p} = q \cdot l \cdot \vec{i}$, где q - величина электрического заряда, l - расстояние между “центрами тяжести” разноименных зарядов, \vec{i} - единичный вектор в направлении от отрицательного заряда к положительному. Степень поляризованности вещества определяется вектором поляризации \vec{P} , который численно равен дипольному моменту единицы объема вещества: $\vec{P} = N \cdot \vec{p} = N \cdot q \cdot l \cdot \vec{i}$, где N - число диполей в единице объема. При не очень сильных полях $\vec{P} \sim \vec{E}$, т.е. это соотношение может быть представлено в виде: $\vec{P} = \epsilon_0 \chi \cdot \vec{E}$, где величину χ называют диэлектрической восприимчивостью. Соответственно, поле в веществе характеризуют вектором \vec{D} , определяемым выражением (1.22).

Аналогичным образом вводится и соотношение (1.23), при выводе которого учитывается, что атомы или молекулы вещества могут иметь собственный магнитный момент.

Из соотношений (1.22), (1.23) видим, что в изотропных средах векторы \bar{D} и \bar{E} параллельны, так же как \bar{B} и \bar{H} .

В. Анизотропная среда.

Анизотропия физических свойств вещества является следствием упорядоченности его микроструктуры и существованием, в связи с этим, некоторых выделенных направлений. Связь между векторами ЭМП в этом случае более сложна, т.к. физические свойства вещества не являются скалярными величинами, а описываются тензорами. В анизотропной среде векторы \bar{D} и \bar{E} (или \bar{B} и \bar{H}) могут быть непараллельными. Так, для компонент вектора \bar{D} можно записать следующие соотношения:

$$D_x = \epsilon_{xx} E_x + \epsilon_{xy} E_y + \epsilon_{xz} E_z \quad (1.24),$$

$$D_y = \epsilon_{yx} E_x + \epsilon_{yy} E_y + \epsilon_{yz} E_z \quad (1.25),$$

$$D_z = \epsilon_{zx} E_x + \epsilon_{zy} E_y + \epsilon_{zz} E_z \quad (1.26).$$

Здесь ϵ_{ij} - компоненты тензора диэлектрической проницаемости (тензор второго ранга). В анизотропной среде этот тензор может быть приведен к диагональному виду, когда отличны от нуля лишь его компоненты ϵ_{xx} , ϵ_{yy} , ϵ_{zz} . Соответствующая система координат называется главными осями тензора.

Аналогичные соотношения можно записать для компонент вектора \bar{B} в магнитно - анизотропной среде:

$$B_x = \mu_{xx} H_x + \mu_{xy} H_y + \mu_{xz} H_z \quad (1.27),$$

$$B_y = \mu_{yx} H_x + \mu_{yy} H_y + \mu_{yz} H_z \quad (1.28),$$

$$B_z = \mu_{zx} H_x + \mu_{zy} H_y + \mu_{zz} H_z \quad (1.29),$$

где μ_{ij} - компоненты тензора абсолютной магнитной проницаемости.

Г. Среда с пространственной и временной дисперсией.

В этом случае связь между векторами поля оказывается еще более сложной. Для изотропной среды учет нелокальности и инерционности процесса поляризации может быть выражен через соотношение:

$$\bar{D}(\bar{r}, t) = \int \int_{V-\infty}^t Q(|\bar{r} - \bar{r}'|, t - t') \bar{E}(\bar{r}', t) dt' dV \quad (1.30),$$

где ядро преобразования Q определяется свойствами среды. Здесь связь между \bar{D} и \bar{E} зависит от геометрии тела и вида процесса в предыдущие моменты времени.

Д. Деление сред по диэлектрическим и магнитным свойствам.

К числу материальных уравнений можно отнести и закон Ома в дифференциальной форме:

$$\vec{\delta} = \sigma \cdot \vec{E} \quad (1.31),$$

где σ - удельная проводимость среды. В зависимости от величины σ все среды делятся на проводники, полупроводники и диэлектрики. Величина относительной диэлектрической проницаемости $\epsilon > 1$. Для сегнетоэлектриков ϵ может достигать нескольких тысяч.

В качестве примера в таблице представлены значения удельной проводимости некоторых материалов.

Материал	Уд. Проводимость (сим/м)
Серебро	$6,14 \cdot 10^7$
Свинец	$0,48 \cdot 10^7$
Плавленый кварц	$2 \cdot 10^{-17}$
Мрамор	$10^{-7} - 10^{-9}$
Стекло	10^{-12}

По магнитным свойствам все материалы делят на диамагнетики ($\mu < 1$), парамагнетики ($\mu > 1$) и ферромагнетики ($\mu \gg 1$).

1.4. Уравнения граничных условий для векторов электромагнитного поля

Уравнения Максвелла совместно с материальными уравнениями описывают ЭМП в безграничной среде. В реальных случаях области определения поля содержат границы раздела сред с различными электродинамическими параметрами. Поэтому для выяснения конкретного вида решений электродинамических задач необходимо знать законы изменения векторов поля на границах раздела. Такие законы задаются уравнениями граничных условий. Рассмотрим их отдельно для нормальных и тангенциальных составляющих векторов поля на границе, принимая во внимание возможность представления любого вектора в виде суммы двух (или произвольного числа) других векторов. Так, если в некоторой точке на границе раздела S задан вектор \vec{a} , он может быть представлен в виде:

$$\vec{a} = \vec{n}_0 \cdot a_n + \vec{\tau}_0 \cdot a_\tau \quad (1.32),$$

где \vec{n}_0 - единичный вектор нормали к границе; $\vec{\tau}_0$ - ортогональный ему и касательный к границе раздела, единичный вектор, лежащий в одной

плоскости с векторами \bar{a} и \bar{n}_0 ; a_n и a_τ - проекции вектора \bar{a} на направления, задаваемые векторами \bar{n}_0 и $\bar{\tau}_0$.

Обоснованность такого подхода подтверждается, как будет видно далее, различием вида уравнений граничных условий для нормальных и тангенциальных составляющих векторов поля.

1.4.1. Граничные условия для векторов электрического поля

Уравнения граничных условий как для электрического, так и магнитного векторов, могут быть получены непосредственно из уравнений Максвелла.

А. Граничные условия для нормальных компонент.

Пусть задана граница раздела двух сред S . Выберем на ней малый элемент ΔS и построим на нем цилиндр высотой Δh (рис. 1.4). Под малостью элемента ΔS понимаем, что его можно считать плоским, а поле вблизи него - однородным. В общем случае на границе может находиться электрический заряд, располагающийся вблизи нее очень тонким слоем.

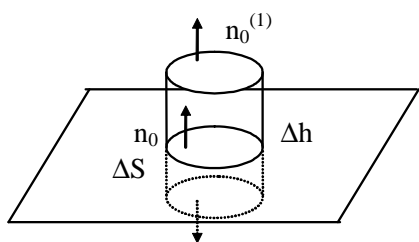


Рис. 1.4

К построенной области применим теорему Гаусса. Ввиду однородности поля поток вектора \bar{D} через верхнее и нижнее основания цилиндра находится умножением скалярного произведения вектора \bar{D} и внешней единичной нормали на площадь ΔS . Заряд внутри цилиндра равен $\xi \cdot \Delta S$, где ξ - поверхностная плотность заряда на границе. Итак, получим соотношение:

$$\bar{D}_1 \cdot \bar{n}_0^{(1)} \cdot \Delta S + \bar{D}_2 \cdot \bar{n}_0^{(2)} \cdot \Delta S + \Phi_b = \xi \cdot \Delta S \quad (1.33),$$

где Φ_b - поток вектора электрической индукции через боковую поверхность цилиндра. Уменьшая неограниченно высоту цилиндра Δh , придем к слиянию оснований цилиндра на границе раздела. При этом поток Φ_b через его боковую поверхность исчезает ($\Phi_b=0$), а единичные векторы нормалей к основаниям цилиндра оказываются параллельными, но противоположно направленными ($\bar{n}_0^{(1)} = \bar{n}_0$; $\bar{n}_0^{(2)} = -\bar{n}_0$, где \bar{n}_0 - нормаль к границе раздела, направленная в первую среду). В результате получим из (1.33):

$$(\bar{D}_1 - \bar{D}_2) \cdot \bar{n}_0 = \xi \quad \text{или} \quad D_{1n} - D_{2n} = \xi \quad (1.34).$$

Таким образом, нормальная компонента вектора \bar{D} при переходе через границу изменяется на величину поверхностной плотности электрического

заряда (если граница несет заряд). В случае $\xi=0$ нормальная компонента вектора \bar{D} непрерывна. Поскольку $\bar{D} = \epsilon \cdot \bar{E}$, то для нормальных компонент вектора \bar{E} , в случае свободной от заряда границы, справедливо соотношение:

$$\frac{E_{1n}}{E_{2n}} = \frac{\epsilon_2}{\epsilon_1} \quad (1.35).$$

В. Граничные условия для тангенциальных компонент.

Вообразим, что заданная граница раздела пересекается некоторой плоскостью P , которую можно считать перпендикулярной выделенному

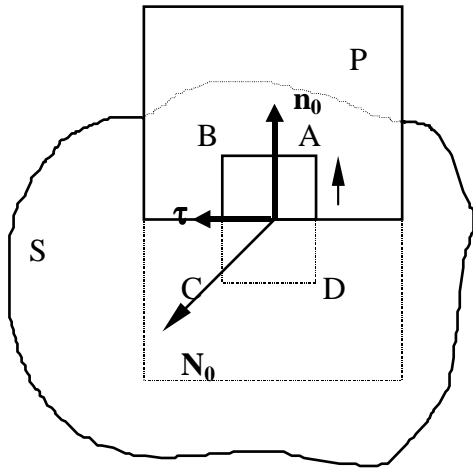


Рис. 1.5.

малому элементу поверхности (малому в упоминавшемся выше смысле). Выделим в плоскости P прямоугольный контур $ABCD$, пересекающий границу S , не выходя за пределы выделенного малого элемента (рис. 1.5). Пусть $AB=CD=\Delta l$; $BC=AD=\Delta h$. Боковые стороны контура параллельны нормали к границе, определяемой единичным вектором \bar{n}_0 . Единичный вектор, совпадающий с линией пересечения плоскости P и границы S , обозначим $\bar{\tau}$. Его направление выбрано так, чтобы выполнялось соотношение $\bar{\tau} = [\bar{N}_0, \bar{n}_0]$, где \bar{N}_0 - единичная нормаль к плоскости P , образующая с

направлением обхода контура $ABCD$ правовинтовую систему.

Применим к данному контуру уравнение Максвелла

$$\oint_L \bar{E} \cdot d\bar{l} = -\frac{\partial \Phi}{\partial t}.$$

Тогда получим:

$$\bar{E}_1 \cdot \bar{\tau} \cdot \Delta l - \bar{E}_2 \cdot \bar{\tau} \cdot \Delta l + C_b = -\frac{\partial \bar{B}}{\partial t} \cdot \bar{N}_0 \cdot \Delta l \cdot \Delta h$$

Здесь два первых слагаемых соответствуют вкладу сторон AB и CD в циркуляцию вектора \bar{E} по этому контуру, а третье выражает долю боковых сторон.

При неограниченном уменьшении величины Δh стороны AB и CD сливаются на границе, а правая часть последнего равенства и C_b исчезают. В результате имеем:

$$(\bar{E}_1 - \bar{E}_2) \cdot \bar{\tau}_0 = 0 \quad \text{или} \quad E_{1\tau} = E_{2\tau} \quad (1.36).$$

Из условия $E_\tau = D_\tau / \epsilon$ получим :

$$\frac{D_{1\tau}}{D_{2\tau}} = \frac{\epsilon_1}{\epsilon_2} \quad (1.37).$$

Таким образом, тангенциальные составляющие вектора \vec{E} на границе раздела непрерывны, а вектора \vec{D} - испытывают разрыв.

Итак, для случая свободной от электрического заряда границы получены условия:

$$\begin{aligned} D_{2n} &= D_{1n}; & D_{2\tau} &= (\epsilon_2/\epsilon_1)D_{1\tau} \\ E_{1\tau} &= E_{2\tau}; & E_{2n} &= (\epsilon_1/\epsilon_2)E_{1n} \end{aligned}$$

Сравнивая эти соотношения, видим, что линии векторов \vec{D} и \vec{E} на границе раздела преломляются. Действительно, для 1-ой среды:

$$\operatorname{tg}\alpha_1 = E_{\tau 1}/E_{n 1}$$

$$\operatorname{tg}\alpha_2 = E_{\tau 2}/E_{n 2}$$

$$\text{и } (\operatorname{tg}\alpha_1 / \operatorname{tg}\alpha_2) = E_{n 2} / E_{n 1} = (\epsilon_1 / \epsilon_2)$$

Для вектора \vec{D} в изотропной среде:

$$D_{1\tau} / D_{1n} = \operatorname{tg}\alpha_1;$$

$$D_{2\tau} / D_{2n} = \operatorname{tg}\alpha_2;$$

$$(\operatorname{tg}\alpha_1 / \operatorname{tg}\alpha_2) = (D_{1\tau} / D_{2\tau}) \cdot (D_{2n} / D_{1n}) = (D_{1\tau} / D_{2\tau}) = (\epsilon_1 / \epsilon_2) \quad (1.38)$$

Т.е. в изотропной среде при наличии границы раздела, векторы \vec{D} и \vec{E} остаются параллельными.

1.4.2. Граничные условия для векторов магнитного поля

Уравнения граничных условий для векторов магнитного поля можно найти, используя тот же подход, что и в случае электрического поля.

А. Нормальные компоненты магнитных векторов.

Используем, как и в случае граничных условий для нормальных компонент вектора \vec{D} , элемент объема цилиндрической формы, построенный на границе раздела двух сред. Найдем поток вектора \vec{B} через поверхность цилиндра, исходя из уравнения:

$$\oint_S \vec{B} \cdot d\vec{S} = 0$$

или, с учетом сделанных предположений:

$$\vec{B}_1 \cdot \vec{n}_0^{(1)} \cdot \Delta S + \vec{B}_2 \cdot \vec{n}_0^{(2)} \cdot \Delta S + \Phi_b = 0 \quad (1.39).$$

Опять переходя к пределу при $\Delta h \Rightarrow 0$, имеем: $(\vec{B}_1 - \vec{B}_2) \cdot \vec{n}_0 = 0$, т.е. нормальная компонента вектора \vec{B} всегда непрерывна. Соответственно, получаем следующие соотношения:

$$B_{1n} = B_{2n}; \quad H_{1n} = (\mu_2/\mu_1) \cdot H_{2n} \quad (1.40).$$

Б. Тангенциальные компоненты магнитных векторов

В случае тангенциальных компонент магнитного поля необходимо различать две возможных ситуации: когда на границе раздела отсутствуют любые токи и когда на границе имеется поверхностный ток. Аналогично случаю тангенциальных компонент электрического поля, найдем циркуляцию вектора \vec{H} по контуру $ABCD$ (рис. 1.5). Получим следующее соотношение:

$$\vec{H}_1 \cdot \vec{\tau} \cdot \Delta l - \vec{H}_2 \cdot \vec{\tau} \cdot \Delta l + C_b = \vec{N}_0 \cdot \left(\frac{\partial \vec{D}}{\partial t} + \vec{\delta} \right) \cdot \Delta l \cdot \Delta h \quad (1.41).$$

При отсутствии поверхностных токов правая часть этого соотношения обращается в ноль при $\Delta h \Rightarrow 0$, что дает:

$$(H_1 - H_2) \cdot \tau_0 = 0 \quad \text{или} \quad H_{1\tau} - H_{2\tau} = 0 \quad \text{или} \quad H_{1\tau} = H_{2\tau} \quad (1.42).$$

Соответственно, $B_{1\tau} / \mu_1 = B_{2\tau} / \mu_2$ или

$$B_{2\tau} = (\mu_2 / \mu_1) \cdot B_{1\tau} \quad (1.43).$$

Случай поверхностного тока.

Это ток, текущий по границе раздела, не занимая объема. Его плотность определяется соотношением:

$$\vec{\eta} = \vec{i}_0 \lim_{\Delta l \rightarrow 0} \frac{\Delta I}{\Delta l} \quad (1.44),$$

где \vec{i}_0 - единичный вектор вдоль направления движения положительных зарядов в данной точке; Δl - элемент линии, перпендикулярной вектору \vec{i}_0 , а ΔI - величина тока, протекающего через данный элемент (рис. 1.6).

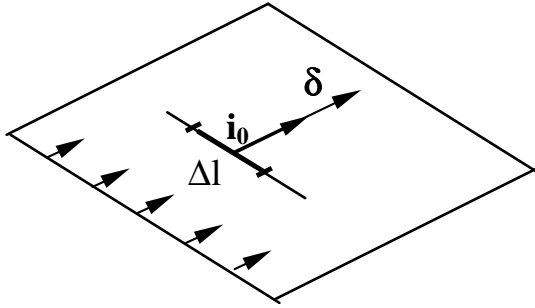


Рис. 1.6.

Поверхностный ток – это лишь математическая абстракция, однако близкая к такой ситуация возможна для границы раздела диэлектрика с очень хорошим проводником. Если

$\vec{\eta} \neq 0$, то правая часть (1.41) при $\Delta h \Rightarrow 0$ примет вид:

$$\vec{N}_0 \cdot \vec{\delta} \cdot \Delta S = \vec{N}_0 \cdot \vec{\delta} \cdot \Delta l \cdot \Delta h = \vec{N}_0 \cdot \vec{i}_0 \left[\frac{\Delta I}{(\Delta l \cdot \Delta h)} \right] \cdot \Delta l \cdot \Delta h = \vec{N}_0 \cdot \vec{\eta} \cdot \Delta l \quad (1.45).$$

Таким образом, получим:

$$(\vec{H}_1 - \vec{H}_2) \cdot \vec{\tau} = \vec{N}_0 \cdot \vec{\eta} \quad \text{или} \quad H_{1\tau} - H_{2\tau} = \eta_n \quad (1.46).$$

Относительно поведения векторных линий магнитного поля при их переходе через границу можно сделать такой же вывод, как и для

электрического поля: векторные линии магнитного поля преломляются, а векторы \vec{B} и \vec{H} остаются параллельными в изотропной среде.

1.5. Баланс энергии электромагнитного поля

Вследствие материальности ЭМП, оно является носителем энергии. При взаимодействии поля с веществом или его воздействии на измерительные приборы, происходит преобразование энергии ЭМП в другие виды. Так, мы ощущаем тепло и видим свет, можем наблюдать с помощью различных приборов и экспериментов результат ее электромеханического, электрохимического либо фотоэлектрического преобразования. Особый интерес представляют процессы распространения ЭМП в пространстве и переноса его энергии.

1.5.1. Уравнение баланса энергии (формула Умова-Пойнтинга).

Плотность электрической мощности в окрестности некоторой точки пространства можно выразить в форме:

$$p = \vec{\delta} \cdot \vec{E} \quad (1.47),$$

$$\text{где } p = \lim_{\Delta V \rightarrow 0} \frac{\Delta P}{\Delta V} \quad (1.48),$$

а ΔP - количество энергии ЭМП, переходящее за 1 с в энергию иного вида внутри элементарного объема ΔV .

Возьмем уравнения Максвелла в дифференциальной форме и умножим скалярно первое - на \vec{E} , второе - на \vec{H} , а затем вычтем из второго первое:

$$\vec{H} \cdot \text{rot} \vec{E} = -\vec{H} \cdot \frac{\partial \vec{B}}{\partial t}$$

$$\vec{E} \cdot \text{rot} \vec{H} = \vec{E} \cdot \frac{\partial \vec{D}}{\partial t} + \vec{E} \cdot \vec{\delta}, \text{ что в результате даст:}$$

$$\vec{H} \cdot \text{rot} \vec{E} - \vec{E} \cdot \text{rot} \vec{H} = -\left(\vec{H} \cdot \frac{\partial \vec{B}}{\partial t} + \vec{E} \cdot \frac{\partial \vec{D}}{\partial t} \right) - \vec{\delta} \cdot \vec{E} \quad (1.49).$$

Выражение в левой части (1.49), согласно одному из тождеств векторного анализа, может быть представлено как $\text{div}[\vec{E}, \vec{H}]$, тогда для (1.49) получим:

$$\text{div}[\vec{E}, \vec{H}] = -\left(\vec{H} \cdot \frac{\partial \vec{B}}{\partial t} + \vec{E} \cdot \frac{\partial \vec{D}}{\partial t} \right) - \vec{\delta} \cdot \vec{E} \quad (1.50).$$

Проинтегрируем последнее соотношение по рассматриваемому объему V и применим к его левой части теорему Остроградского - Гаусса:

$$\int_V \operatorname{div}[\bar{E}, \bar{H}] dV = \oint_S [\bar{E}, \bar{H}] \cdot \bar{dS} = - \int_V \left(\bar{H} \cdot \frac{\partial \bar{B}}{\partial t} + \bar{E} \cdot \frac{\partial \bar{D}}{\partial t} \right) \cdot dV - \int_V \bar{\delta} \cdot \bar{E} \cdot dV \quad (1.51),$$

где S - замкнутая поверхность, охватывающая объем V . Последнее слагаемое в (1.51) имеет размерность мощности, поэтому все члены (1.51) должны иметь размерность мощности.

Выясним физический смысл величины $\int_V \left(\bar{H} \cdot \frac{\partial \bar{B}}{\partial t} + \bar{E} \cdot \frac{\partial \bar{D}}{\partial t} \right) \cdot dV$. Для этого представим, что к поверхности S изнутри примыкает другая, непроницаемая для ЭМП, поверхность S^l . В этом случае на поверхности S поле отсутствует, т.е. внутренняя область является изолированной и: $\oint_S [\bar{E}, \bar{H}] \cdot \bar{dS} = 0$. Соответственно, имеем:

$$P = - \int_V \left(\bar{H} \cdot \frac{\partial \bar{B}}{\partial t} + \bar{E} \cdot \frac{\partial \bar{D}}{\partial t} \right) \cdot dV \quad (1.52).$$

Поскольку область пространства V не сообщается с окружающей средой, то P - это скорость изменения энергии W , сосредоточенной внутри объема V , т.е. $P = -(dW/dt)$. Знак (-) выбран здесь из соображений убывания энергии при наличии потерь ($P > 0$). Таким образом:

$$\int_V \left(\bar{H} \cdot \frac{\partial \bar{B}}{\partial t} + \bar{E} \cdot \frac{\partial \bar{D}}{\partial t} \right) \cdot dV = \frac{dW}{dt} \quad (1.53).$$

Введем обозначение $\bar{\Pi} = [\bar{E}, \bar{H}]$, тогда, (1.51) примет вид:

$$\oint_S \bar{\Pi} \cdot \bar{dS} + \frac{dW}{dt} + P_l = 0 \quad (1.54).$$

Это уравнение баланса энергии или интегральная формула Умова - Пойнтинга. Вектор $\bar{\Pi} = [\bar{E}, \bar{H}]$ называют вектором Пойнтинга. Он характеризует плотность потока мощности ЭМП и измеряется в Вт/м². Положительный поток вектора $\bar{\Pi}$ через границу соответствует излучению энергии из выделенной области, отрицательный - ее поступлению в данную область из окружающего пространства.

1.5.2. Электрическая и магнитная энергия электромагнитного поля

Итак, скорость изменения энергии ЭМП в некоторой области V определяется соотношением:

$$\frac{dW}{dt} = \int_V \left(\bar{H} \cdot \frac{\partial \bar{B}}{\partial t} + \bar{E} \cdot \frac{\partial \bar{D}}{\partial t} \right) \cdot dV.$$

Для безынерционной среды (где μ и ϵ не зависят от времени) члены $\bar{H} \cdot \frac{\partial \bar{B}}{\partial t}$ и $\bar{E} \cdot \frac{\partial \bar{D}}{\partial t}$ можно представить как: $\bar{H} \cdot \frac{\partial \bar{B}}{\partial t} = \mu \bar{H} \cdot \frac{\partial \bar{H}}{\partial t} = \frac{\partial}{\partial t} \left(\frac{\mu H^2}{2} \right)$ и $\bar{E} \cdot \frac{\partial \bar{D}}{\partial t} = \frac{\partial}{\partial t} \left(\frac{\epsilon E^2}{2} \right)$. Соответственно:

$$\frac{dW}{dt} = \frac{\partial}{\partial t} \int_V \left(\frac{\mu H^2}{2} + \frac{\epsilon E^2}{2} \right) \cdot dV \quad (1.55)$$

и энергия W в данном объеме будет равна:

$$W = \int_V \left(\frac{\mu H^2}{2} + \frac{\epsilon E^2}{2} \right) \cdot dV \quad (1.56).$$

Видно, что этот интеграл состоит из двух частей, одна из которых зависит только от магнитного, а вторая - только от электрического полей. Т.е.:

$W = W^E + W^M$, где $W^E = \frac{1}{2} \int_V \epsilon E^2 dV$, а $W^M = \frac{1}{2} \int_V \mu H^2 dV$ - это электрическая и магнитная энергии ЭМП.

1.5.3. Перенос энергии и групповая скорость в электромагнитном поле.

Если ΔW - энергия ЭМП в элементарном объеме ΔV , то величина $w = \lim_{\Delta V \rightarrow 0} \frac{\Delta W}{\Delta V}$ представляет собой объемную плотность энергии ЭМП.

Тогда из соотношения $W = \frac{1}{2} \int_V (\mu H^2 + \epsilon E^2) \cdot dV$ вытекает, что

$$w = \frac{1}{2} (\mu H^2 + \epsilon E^2) \quad (1.57),$$

причем слагаемые $w^M = \mu H^2/2$ и $w^E = \epsilon E^2/2$ - это объемные плотности магнитной и электрической энергии, соответственно. Их единица измерения - Дж/м³. Очевидно, что соотношение (1.57) характеризует распределение энергии ЭМП в пространстве. С учетом данного соотношения и введенных ранее обозначений, выражение

$\text{div}[\bar{E}, \bar{H}] = - \left(\bar{H} \cdot \frac{\partial \bar{B}}{\partial t} + \bar{E} \cdot \frac{\partial \bar{D}}{\partial t} \right) - \bar{\delta} \cdot \bar{E}$ может быть переписано в виде:

$$\text{div} \bar{\Pi} + \frac{dw}{dt} + p^l = 0 \quad (1.58).$$

Это дифференциальная форма теоремы Умова - Пойнтинга. Она показывает, что изменение плотности энергии ЭМП в каждой точке пространства определяется характером ее преобразования (этому соответствует локальное значение плотности мощности p^j) и движения (которое характеризуется вектором Пойнтинга $\bar{\Pi}$).

Зная величины w и $\bar{\Pi}$, можно найти скорость движения энергии v_e , т.е. групповую скорость ЭМП. Так, выделим в потоке энергии ЭМП цилиндрический объем с поперечным сечением ΔS . Количество энергии, проходящее за 1 секунду через некоторую условную границу l_1 , равно $W_{(\Delta t=1)} = |\bar{\Pi}| \cdot \Delta S$. Оно заполняет цилиндр до l_2 , причем расстояние Δl между l_1 и l_2 численно равно скорости движения энергии $\Delta l = v_e$. Для определения этой величины необходимо $W_{(dt=1)}$ разделить на количество энергии, приходящееся на единицу длины цилиндра, которое равно: $W_{(\Delta l=1)} = w \cdot \Delta S$. Таким образом:

$$v_e = \frac{W_{(\Delta t=1)}}{W_{(\Delta l=1)}} = \frac{|\bar{\Pi}| \cdot \Delta S}{w \cdot \Delta S} = \frac{|\bar{\Pi}|}{w} \quad (1.59).$$

Или, учитывая одинаковое направление векторов \bar{v}_e и $\bar{\Pi}$,

$$\bar{v}_e = \frac{\bar{\Pi}}{w} \quad (1.60).$$

1.6. Символический метод (метод комплексных амплитуд)

Световые волны представляют собой электромагнитное поле с очень высокими частотами. Как уже отмечено, система уравнений Максвелла для описания электромагнитного поля в диэлектрической среде имеет вид:

$$\begin{aligned} \text{rot} \bar{H} &= \frac{\partial \bar{D}}{\partial t} & \text{div} \bar{D} &= 0 \\ \text{rot} \bar{E} &= -\frac{\partial \bar{B}}{\partial t} & \text{div} \bar{B} &= 0 \end{aligned}$$

Будем рассматривать установившиеся во времени процессы, которым свойственны гармонические колебания. Такой подход возможен и продуктивен в силу применимости принципа суперпозиции к описанию электромагнитных процессов в линейных средах. Очевидно, что в данном случае любая временная зависимость может быть представлена в виде суммы (конечной либо бесконечной) отдельных гармонических составляющих.

Гармонические колебания величины U описываются известным соотношением:

$$U(t) = U_m \cos(\omega t + \varphi) \quad (1.61),$$

где U_m - амплитуда; ω - круговая частота; φ - начальная фаза; $(\omega t + \varphi)$ - полная фаза колебания. Гармонические колебания - это периодический процесс, т.е. рассматриваемая величина в любой момент времени удовлетворяет соотношению:

$$U(t) = U(t+T),$$

где T - период колебаний - $T = \frac{2\pi}{\omega} = \frac{1}{f}$.

Скалярные и векторные гармонические функции координат и времени (амплитуды и фазы в общем случае зависят от положения точки) выражаются соотношениями:

$$U(x, y, z, t) = U_m(x, y, z) \cdot \cos[\omega t + \varphi(x, y, z)] = U_m(\bar{r}) \cos[\omega t + \varphi(\bar{r})] \quad (1.62),$$

$$\begin{aligned} \bar{V}(\bar{r}, t) = \bar{x}_0 \cdot V_{mx}(\bar{r}) \cdot \cos[\omega t + \varphi_x(\bar{r})] + \bar{y}_0 \cdot V_{my}(\bar{r}) \cdot \cos[\omega t + \varphi_y(\bar{r})] + \\ + \bar{z}_0 \cdot V_{mz}(\bar{r}) \cdot \cos[\omega t + \varphi_z(\bar{r})] \end{aligned} \quad (1.63).$$

При $\varphi_x(\bar{r}) = \varphi_y(\bar{r}) = \varphi_z(\bar{r})$ можно записать: $\bar{V}(\bar{r}, t) = \bar{V}_m(\bar{r}) \cdot \cos[\omega t + \varphi(\bar{r})]$,

где $\bar{V}_m = \bar{x}_0 \cdot V_{mx} + \bar{y}_0 \cdot V_{my} + \bar{z}_0 \cdot V_{mz}$ - векторная амплитуда.

Основой символического метода является формула Эйлера, которая выражает тригонометрические функции через экспоненциальные:

$$\exp[i(\omega t + \varphi)] = \cos(\omega t + \varphi) + i \cdot \sin(\omega t + \varphi) \quad (1.64).$$

Анализ гармонических процессов существенно упрощается при использовании метода комплексных амплитуд. Его основная идея состоит в том, что действительная гармоническая функция является, согласно формуле Эйлера, вещественной частью экспоненциальной функции от мнимого аргумента. Практический смысл данного подхода обусловлен тем, что многие алгебраические преобразования выполняются проще для экспоненциальной функции, чем для тригонометрических функций. Итак, выражение $U = U_m \cos(\omega t + \varphi)$ может быть представлено в виде:

$$U = \operatorname{Re} \dot{U} = \operatorname{Re} \dot{U}_m \cdot \exp(i\omega t) \quad (1.65),$$

где:

$$\dot{U} = U_m \cdot \exp[i(\omega t + \varphi)] = U_m \cdot \exp(i\varphi) \cdot \exp(i\omega t) = \dot{U}_m \cdot \exp(i\omega t)$$

а $\dot{U}_m = U_m \cdot \exp(i\varphi)$ - комплексная амплитуда.

Величину \dot{U} называют комплексом величины U . Таким образом, для представления скалярных и векторных величин можно использовать соотношения $\varphi = \operatorname{Re} \dot{\varphi}$; $\bar{V} = \operatorname{Re} \dot{\bar{V}}$. Очевидно, что вещественная гармоническая величина может быть выражена через ее комплекс и комплексно - сопряженную величину этого комплекса в виде:

$$U = \frac{1}{2}(\dot{U} + \dot{U}^*) \quad (1.66)$$

Если комплекс $\dot{\vec{V}}$ удовлетворяет некоторому линейному дифференциальному уравнению, то ему удовлетворяют и обе его части - вещественная и мнимая. Поэтому, если нужно найти решение этого уравнения в виде \bar{V} , его можно искать в виде $\dot{\vec{V}}$, а потом найти $Re \dot{\vec{V}}$. Будем использовать метод комплексных амплитуд (символический метод) при изучении гармонических во времени процессов, т. к. он ко всему еще и упрощает уравнения электромагнитного поля, освобождая их от временной зависимости. Действительно, дифференцирование комплексов по времени эквивалентно простому умножению начального выражения на $i\omega$, т.е.:

$$\frac{d}{dt} \dot{\vec{V}} = i\omega \dot{\vec{V}}.$$

Так, например, уравнение $rot \dot{\vec{E}} = -\frac{\partial \dot{\vec{B}}}{\partial t}$ можно трансформировать к виду:

$$rot \dot{\vec{E}} = -\frac{\partial \dot{\vec{B}}}{\partial t} \rightarrow rot(\dot{\vec{E}}_m \cdot exp(i\omega t)) = -\frac{\partial(\dot{\vec{B}}_m \cdot exp(i\omega t))}{\partial t} \rightarrow$$

$$exp(i\omega t) \cdot rot \dot{\vec{E}}_m = -i\omega \cdot \dot{\vec{B}}_m \cdot exp(i\omega t) \rightarrow rot \dot{\vec{E}}_m = -i\omega \dot{\vec{B}}_m$$

Очевидно, что во многих случаях такой подход может существенно упростить решение электродинамических задач.

1.7. Комплексные проницаемости

В комплексной форме первое уравнение Максвелла имеет вид:

$$rot \dot{\vec{Z}} = i\omega \epsilon \dot{\vec{E}} + \dot{\vec{\sigma}},$$

что с учетом закона Ома в дифференциальной форме дает:

$$rot \dot{\vec{H}} = i\omega \epsilon \dot{\vec{E}} + \dot{\vec{\sigma}} = i\omega \left(\epsilon + \frac{\dot{\vec{\sigma}}}{i\omega} \right) \cdot \dot{\vec{E}} = i\omega \left(\epsilon - i \frac{\dot{\vec{\sigma}}}{\omega} \right) \cdot \dot{\vec{E}} \quad (1.67).$$

Полученное уравнение справедливо для случая проводящих сред и имеет форму, аналогичную форме первого уравнения Максвелла для диэлектрической среды, в котором роль диэлектрической проницаемости играет комплексная величина:

$$\dot{\epsilon} = \left(\epsilon - i \frac{\dot{\vec{\sigma}}}{\omega} \right) \quad (1.68).$$

Ее называют комплексной диэлектрической проницаемостью и представляют в форме: $\dot{\epsilon} = \epsilon' - i\epsilon''$, где $\epsilon' = \epsilon$; $\epsilon'' = \frac{\dot{\vec{\sigma}}}{\omega}$.

Очевидно, что соотношение величин мнимой и вещественной частей комплексной диэлектрической проницаемости определяет, насколько проводящей является рассматриваемая среда. В качестве характеристики

проводящих свойств среды используют отношение ϵ''/ϵ' , которое называют тангенсом угла диэлектрических потерь:

$$\operatorname{tg}D = \frac{\epsilon''}{\epsilon'} = \frac{y}{\omega\epsilon}$$

Физически эта характеристика определяет соотношение между амплитудами токов проводимости и смещения в данной среде. При наличии в среде токов проводимости появляются тепловые потери энергии электромагнитного поля. Таким образом, комплексный характер диэлектрической проницаемости является признаком потерь энергии поля, обусловленных неидеальными диэлектрическими свойствами среды. Аналогичным образом можно ввести понятие комплексной магнитной проницаемости, комплексный характер которой свидетельствует о магнитных потерях электромагнитной энергии вследствие явления магнитного гистерезиса в данной среде.

1.8. Уравнения Максвелла в символической форме

Использование метода комплексных амплитуд и понятия комплексных проницаемостей позволяет записать систему уравнений Максвелла для монохроматического поля в форме:

$$\begin{aligned} \operatorname{rot} \dot{\vec{Z}} &= i\omega\epsilon \dot{\vec{E}} \\ \operatorname{rot} \dot{\vec{E}} &= -i\omega\mu \dot{\vec{Z}} \end{aligned} \quad (1.69).$$

Эти уравнения образуют полную систему, достаточную для описания монохроматических полей, т.к. два других уравнения являются следствиями записанных.

1.9. Вектор Пойнтинга в символической форме.

Метод комплексных амплитуд не может непосредственно применяться к вычислению квадратичных величин. Действительно, пусть есть два комплексных числа $\dot{a} = a' + ia''$ и $\dot{b} = b' + ib''$, вещественные части которых $\operatorname{Re}(\dot{a}) = a'$ и $\operatorname{Re}(\dot{b}) = b'$. Находя вещественную часть произведения этих чисел, видим, что она не равна произведению их вещественных частей:

$$\dot{a} \cdot \dot{b} = a'b' - a''b'' + i(a''b' + a'b'')$$

$$\operatorname{Re}(\dot{a}\dot{b}) = a'b' - a''b'' \neq \operatorname{Re}(\dot{a}) \cdot \operatorname{Re}(\dot{b}).$$

Таким образом, механический подход замены вещественных гармонических функций их комплексами при возврате к вещественным функциям после решения задачи, в случае выполнения нелинейных операций, некорректен, т.к. приводит к неверному результату.

Тем не менее, символический метод может использоваться и при выполнении нелинейных операций. В этом случае представление

вещественных скалярных или векторных функций их комплексами осуществляется по правилу:

$$\bar{V} = \frac{1}{2}(\dot{V} + \dot{V}^*)$$

1.10. Выражение для вектора Пойнтинга в символической форме

Используя последнее соотношение, выражение для мгновенных значений вектора Пойнтинга в монохроматическом электромагнитном поле можно представить в виде:

$$\bar{P} = [\bar{E}, \bar{Z}] = \frac{1}{4} \left[(\dot{E} + \dot{E}^*) (\dot{Z} + \dot{Z}^*) \right] \quad (1.70).$$

В теории гармонических процессов важную роль играют квадратичные величины, средние по времени. При их вычислении рассматривают промежуток времени τ , значительно превышающий период процесса T : $\tau \gg T$. В связи с этим, понятия средней величины и средней за период совпадают. Найдем среднее значение вектора Пойнтинга, отмечая факт усреднения по времени волнистой чертой над знаком вектора:

$$\tilde{\bar{P}} = \frac{1}{\Phi_0} \int_0^{\Phi} \bar{P} \cdot dt$$

Из (1.70) вытекает, что :

$$\begin{aligned} \bar{P} &= \frac{1}{4} \left\{ [\dot{E}, \dot{H}] + [\dot{E}^*, \dot{H}^*] + [\dot{E}, \dot{H}^*] + [\dot{E}^*, \dot{H}] \right\} = \\ &= \frac{1}{4} \left\{ [\dot{E}, \dot{H}] + [\dot{E}, \dot{H}]^* \right\} + \frac{1}{4} \left\{ [\dot{E}, \dot{H}^*] + [\dot{E}, \dot{H}^*]^* \right\} = \frac{1}{2} \operatorname{Re}[\dot{E}, \dot{H}] + \frac{1}{2} \operatorname{Re}[\dot{E}, \dot{H}^*] \end{aligned} \quad (1.71).$$

Так как $\dot{E} = \dot{E}_m \exp[i(\omega t + \varphi_E)]$, а $\dot{H} = \dot{H}_m \exp[i(\omega t + \varphi_H)]$, где φ_E и φ_H – начальные фазы электрического и магнитного векторов, то:

$$[\dot{E}, \dot{H}] = [\dot{E}_m, \dot{H}_m] \exp[i(2\omega t + \varphi_E + \varphi_H)] = [\dot{E}_m, \dot{H}_m] \exp[i(\varphi_E + \varphi_H)] \exp(i2\omega t),$$

$$\text{а } \operatorname{Re}[\dot{E}, \dot{H}] = [\dot{E}_m, \dot{H}_m] \cdot \cos(2\omega t + \varphi_E + \varphi_H).$$

Тогда усреднение этого выражения по времени дает:

$$\frac{1}{\Phi_0} \int_0^{\Phi} \operatorname{Re}[\dot{E}, \dot{H}] dt = \frac{1}{\Phi} [\dot{E}_m, \dot{H}_m] \cdot \int_0^{\Phi} \exp[i(2\omega t + \varphi_E + \varphi_H)] dt = 0$$

Второе слагаемое в (1.70) дает:

$$[\dot{E}, \dot{H}^*] = [\dot{E}_m, \dot{H}_m] \exp[i(\omega t + \varphi_E)] \cdot \exp[-i(\omega t + \varphi_H)] =$$

$$= [\dot{E}_m, \dot{H}_m] \exp[i(\varphi_E - \varphi_H)]$$

и $Re[\bar{E}, \bar{H}^*] = [\bar{E}_m, \bar{H}_m] \cdot \cos(\varphi_E - \varphi_H)$, т.е. здесь временная зависимость исчезает. Тогда усреднение по времени данного выражения, очевидно, дает просто саму его величину:

$$\frac{1}{\Phi_0} \int_{\Phi_0} Re[\bar{E}, \bar{H}^*] dt = Re[\bar{E}, \bar{H}^*] \cdot \frac{1}{\Phi_0} \int_{\Phi_0} dt = Re[\bar{E}, \bar{H}^*]$$

Таким образом:

$$\tilde{P} = \frac{1}{T} \int_0^T \frac{1}{2} Re[\dot{E}, \dot{H}^*] dt = \frac{1}{2} Re[\dot{E}, \dot{H}^*] = \frac{1}{2} Re[\dot{E}^*, \dot{H}] \quad (1.72).$$

Видим, что среднее значение вектора Пойнтинга можно получить как реальную часть вектора:

$$\dot{P} = \frac{1}{2} [\dot{E}, \dot{H}^*] \quad (1.73),$$

который называется комплексным вектором Пойнтинга $\tilde{P} = Re \dot{P}$.

1.11. Энергетические характеристики монохроматического ЭМП

1. Средние значения важных величин:

Средние значения векторов поля, как следует из записанных выше выражений, определяются соотношениями:

$$\tilde{E}^2 = \frac{1}{2} \bar{E}_m^2 = \frac{1}{2} E_m^2; \quad \tilde{H}^2 = \frac{1}{2} \bar{H}_m^2 = \frac{1}{2} H_m^2 \quad (1.74),$$

где значки векторов в правой части можно опустить. Среднее значение объемной плотности электромагнитной энергии найдем, исходя из общего выражения:

$w = \frac{1}{2} (m\bar{H}^2 + e\bar{E}^2)$ и, учитывая монохроматичность электромагнитного поля:

$$\tilde{w} = \frac{1}{2} \left(m \cdot \frac{1}{2} \bar{H}_m^2 + e \cdot \frac{1}{2} \bar{E}_m^2 \right) = \frac{1}{4} (m\bar{H}_m^2 + e\bar{E}_m^2) = \frac{1}{4} (mH_m^2 + eE_m^2) \quad (1.75).$$

Выражение для усредненной по времени плотности мощности электромагнитного поля определим аналогичным образом. Переходя от действительных величин к комплексам, получим для мгновенного значения плотности мощности соотношение:

$$p = \bar{\partial} \cdot \bar{E} = \frac{1}{4} (\dot{\partial} + \dot{\partial}^*) \cdot (\dot{E} + \dot{E}^*) = \frac{1}{4} (\dot{\partial} \cdot \dot{E} + \dot{\partial}^* \cdot \dot{E}^* + \dot{\partial} \cdot \dot{E}^* + \dot{\partial}^* \cdot \dot{E}),$$

из которого опять следует, что для первых двух слагаемых усреднение по времени приводит к нулевому результату. В итоге получим:

$$\tilde{p} = \frac{1}{2} Re(\dot{\partial}, \dot{E}^*) = Re \left[\frac{1}{2} \dot{\partial} \cdot \dot{E}^* \right] = Re \dot{p},$$

где $\dot{p} = \frac{I}{2} \dot{\delta} \cdot \dot{E}^*$ - комплексная плотность мощности. В среде без потерь векторы $\dot{\delta}$ и \dot{E} коллинеарны и совпадают по фазе, поэтому для такой среды:

$$\tilde{p} = \text{Re } \dot{p} = \frac{I}{2} \delta_m E_m \quad \text{при } \varphi_\delta = \varphi_E.$$

1.12. Формула для баланса энергии монохроматического ЭМП

Как и в общем случае, исходим из системы уравнений Максвелла, но записанной для комплексов векторов поля. При этом первое уравнение запишем для комплексно – сопряженных величин:

$$\text{rot} \dot{E} = -i\omega \dot{Z} \quad (1.76),$$

$$\text{rot} \dot{Z}^* = -i\omega \dot{E}^* + \dot{\delta}^* \quad (1.77).$$

Равенство (1.75) домножим на \dot{Z}^* , равенство (1.77) - на \dot{E} , затем из первого равенства вычтем второе, и получим:

$$\dot{Z}^* \cdot \text{rot} \dot{E} - \dot{E} \cdot \text{rot} \dot{Z}^* = -i\omega \dot{Z} \cdot \dot{Z}^* + i\omega \dot{E} \cdot \dot{E}^* - \dot{E} \cdot \dot{\delta}^*$$

Используя опять известное из векторного анализа соотношение, получим:

$$\text{div} [\dot{E}, \dot{Z}^*] = -i\omega (\dot{Z} \cdot \dot{Z}^* - \dot{E} \cdot \dot{E}^*) - \dot{E} \cdot \dot{\delta}^*$$

Домножив обе части равенства на $\frac{1}{2}$, можно переписать его в виде:

$$\text{div} \dot{P} = -i\omega \left(\frac{\dot{Z} \cdot \dot{Z}^*}{2} - \frac{\dot{E} \cdot \dot{E}^*}{2} \right) - \frac{1}{2} \dot{E} \cdot \dot{\delta}^*,$$

но $\dot{Z} \cdot \dot{Z}^* = \overline{Z}_m \cdot \overline{Z}_m^* = \overline{Z}_m^2$, $\dot{E} \cdot \dot{E}^* = \overline{E}_m \cdot \overline{E}_m^* = \overline{E}_m^2$, поэтому :

$$\text{div} \dot{P} = -2i\omega \left(\frac{\dot{Z}_m^2}{4} - \frac{\dot{E}_m^2}{4} \right) - \frac{1}{2} \dot{E} \cdot \dot{\delta}^*$$

$$\text{или } \Rightarrow \text{div} \dot{P} = -2i\omega (\tilde{w}^M - \tilde{w}^\varnothing) - \dot{p} \quad (1.78).$$

Интегрируя последнее соотношение по объёму, и применяя формулу Остроградского, получим:

$$\oint_S \dot{P} dS = -2i\omega (\tilde{W}^M - \tilde{W}^\varnothing) - \dot{C} \quad (1.79).$$

Здесь:

$$W^M = \int_V w^M dV = \frac{I}{4} \int_V m Z_m^2 dV$$

$$W^\varnothing = \int_V w^\varnothing dV = \frac{I}{4} \int_V e E_m^2 dV$$

$\dot{P} = \int_V \dot{p} dV$ - комплексная мощность.

Уравнения (1.78) и (1.79) - это уравнения баланса энергии монохроматического ЭМП в комплексной форме.

2. ПЛОСКИЕ ЭЛЕКТРОМАГНИТНЫЕ ВОЛНЫ В БЕЗГРАНИЧНОЙ СРЕДЕ

2.1. Волновое уравнение для безграничной среды

Из уравнений Максвелла следует, что переменное электрическое (магнитное) поле приводит к возникновению изменяющегося в пространстве магнитного (электрического) поля. Очевидно, что данный процесс приведет к распространению электромагнитного возмущения в пространстве. Рассмотрим основные законы и особенности такого распространения. Начнем с простейшего случая непроводящей, безграничной, однородной изотропной среды без сторонних токов и зарядов ($\sigma = 0$; $\bar{\delta} = 0$; $\rho = 0$). Данная ситуация приводит к замкнутой системе уравнений относительно векторов поля \bar{E} и \bar{H} :

$$\text{rot}\bar{H} = \varepsilon \frac{\partial \bar{E}}{\partial t} \quad (2.1a),$$

$$\text{rot}\bar{E} = -\mu \frac{\partial \bar{H}}{\partial t} \quad (2.1б),$$

$$\text{div}\bar{E} = 0 \quad (2.1в),$$

$$\text{div}\bar{H} = 0 \quad (2.1г).$$

Для ее решения возьмем, например, второе уравнение и применим к обеим его частям операцию rot . Порядок следования операций дифференцирования по времени и пространственным координатам для функций, непрерывных в пространстве и времени, может меняться, поэтому в результате получим:

$$\text{rot}(\text{rot}\bar{E}) = -\mu \frac{\partial}{\partial t} \text{rot}\bar{H}$$

Используя первое уравнение Максвелла, и полагая среду безынерционной, приводим его к виду:

$$\text{rot}(\text{rot}\bar{E}) = -\mu \frac{\partial}{\partial t} \left(\varepsilon \frac{\partial \bar{E}}{\partial t} \right) = -\mu\varepsilon \frac{\partial^2 \bar{E}}{\partial t^2} \quad (2.2).$$

Таким образом, получено уравнение относительно вектора поля \bar{E} , связывающее его пространственные и временные производные. Это уравнение называется волновым, оно описывает распространение электромагнитного поля в безграничной изотропной среде. Преобразуем его, учитывая известное соотношение векторного анализа, и условие $\text{div}\bar{E} = 0$:

$$\text{rot}(\text{rot}\bar{E}) = \text{grad}(\text{div}\bar{E}) - \nabla^2 \bar{E} = -\nabla^2 \bar{E}$$

В итоге получим другую форму волнового уравнения:

$$\nabla^2 \bar{E} - \mu\varepsilon \frac{\partial^2 \bar{E}}{\partial t^2} = 0 \quad (2.3).$$

Аналогичным путем можно получить волновое уравнение для вектора \bar{H} :

$$\nabla^2 \bar{H} - \mu \varepsilon \frac{\partial^2 \bar{H}}{\partial t^2} = 0 \quad (2.4).$$

2.2. Решение волнового уравнения. Плоские волны.

Найдем простейшее решение волнового уравнения. Пусть поле \bar{E} зависит только от одной пространственной координаты z . Тогда:

$$\frac{\partial \bar{E}}{\partial x} = \frac{\partial \bar{E}}{\partial y} = 0, \text{ а } \nabla^2 \bar{E} = \frac{\partial^2 \bar{E}}{\partial z^2}.$$

В этом случае приходим к одномерному волновому уравнению:

$$\frac{\partial^2 \bar{E}}{\partial z^2} - \varepsilon \mu \frac{\partial^2 \bar{E}}{\partial t^2} = 0 \quad (2.5).$$

Так как $\rho = 0$, то из условия $\operatorname{div} \bar{D} = 0$ получаем: $\frac{dE_z}{dz} = 0$, что приводит к решению вида $E_z = \text{const}$. Но это соответствует электростатическому случаю, который не удовлетворяет рассматриваемым условиям переменного поля. Таким образом, полагаем $E_z = 0$ и решение волнового уравнения может иметь только поперечную (относительно направления распространения) компоненту поля \bar{E} .

Пусть для простоты положим $E_y = 0$, а $E_x \neq 0$. Соответственно, (2.5) принимает вид **скалярного** одномерного волнового уравнения:

$$\frac{\partial^2 E_x}{\partial z^2} - \mu \varepsilon \frac{\partial^2 E_x}{\partial t^2} = 0 \quad (2.6).$$

Вид этого уравнения подсказывает конструкцию его решения. Очевидно, что переменные z и t должны входить в выражение для поля E_x одинаковым образом. Поэтому решение можно представить в виде плоской скалярной волны:

$$E_x(t, z) = E_{x1} \left(t - \frac{z}{v} \right) + E_{x2} \left(t + \frac{z}{v} \right) \quad (2.7).$$

Здесь $v = \frac{1}{\sqrt{\mu \varepsilon}}$ - скорость распространения волны в среде, а первое и второе

слагаемые соответствуют волнам, бегущим в направлениях $+z$ и $-z$, соответственно.

Но $m \cdot e = m_0 \cdot e_0 \cdot m_r \cdot e_r$. Тогда $v = \frac{1}{\sqrt{\mu_0 \cdot \varepsilon_0 \cdot \mu_r \cdot \varepsilon_r}} = \frac{c}{n}$, где $\mu_r = \frac{\mu}{\mu_0}$ и $\varepsilon_r = \frac{\varepsilon}{\varepsilon_0}$

- относительные магнитная и диэлектрическая проницаемости среды;

$n = \sqrt{\mu_r \cdot \varepsilon_r}$ - показатель преломления среды.

Постоянные $\mu_0 = 4\pi \cdot 10^{-7} \text{ Гн/м}$; $\epsilon_0 = \frac{1}{36\pi} \cdot 10^{-9} \text{ Ф/м}$; $c = 3 \cdot 10^8 \text{ м/с}$.

2.3. Гармонические волны

Если при $z=0$ имеется возмущение $E(t) = E_m \cdot \cos(\omega t + \varphi)$, то, согласно (2.7), получим:

$$E_1(z, t) = E_{m1} \cdot \cos\left[\omega\left(t - \frac{z}{v}\right) + \varphi\right] \quad (2.8).$$

$$E_2(z, t) = E_{m2} \cdot \cos\left[\omega\left(t + \frac{z}{v}\right) + \varphi\right]$$

То есть данному возмущению соответствуют две гармонических плоских волны, которые бегут в направлениях $+z$ и $-z$. Мгновенное значение возмущения в некоторой точке определяется амплитудой E_m волны и ее фазой $[\omega(t \mp \frac{z}{v}) + \varphi]$. Выражение для фазы волны принято записывать в

форме: $[\omega t \mp k \cdot z + \varphi]$, где $k = \frac{\omega}{v}$ - волновое число. Если E_m не зависит от поперечных координат, то волна называется однородной.

Геометрическое место точек, в которых фаза волны ($\omega t \mp kz + \varphi = \text{const}$) одинакова, называется волновым или фазовым фронтом.

В некоторый фиксированный момент времени $t=t_0$ фаза плоской волны ($\omega t \mp kz + \varphi = \text{const}$ при фиксированном значении z , то есть волновой фронт представляет собой плоскость, нормальную к оси z . Поэтому волна называется плоской.

При изменении времени на величину Δt волновой фронт в пространстве смещается на Δz . При этом $(\omega \cdot \Delta t - k \cdot \Delta z) = 0$, так как фаза волны определяется выбранным волновым фронтом. Отсюда:

$$\frac{\omega}{k} = \frac{\Delta z}{\Delta t} = v_\phi \quad (2.9),$$

где v_ϕ - фазовая скорость волны. В пространстве изменение фазы волны $\Delta\varphi = 2\pi$ соответствует расстоянию, равному длине волны λ . Но:

$$\Delta\varphi = k \cdot \lambda = 2\pi .$$

Отсюда $k = \frac{2\pi}{\lambda}$. Таким образом, смысл волнового числа заключается в том, что оно определяет число длин волн, укладываемых на отрезке длиной 2π .

2.4. Распространение плоской волны в произвольном

направлении

При распространении плоской волны в произвольном направлении, которое не совпадает с какой – либо координатной осью, все компоненты оператора ∇ в заданной декартовой системе координат могут быть отличны от нуля. Однако и в этом случае плоская волна остается поперечной, и мы используем для ее описания уравнения:

$$\nabla^2 E_\tau - \mu\epsilon \frac{\partial^2 E_\tau}{\partial t^2} = 0 \quad (2.10a)$$

или

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \cdot E_\tau - \mu\epsilon \cdot \frac{\partial^2}{\partial t^2} E_\tau = 0 \quad (2.10b),$$

где E_τ - поперечные компоненты вектора \vec{E} . Решением данного уравнения является гармоническая плоская волна:

$$E(\vec{r}, t) = E_m \cdot \cos(\omega t - \vec{k} \cdot \vec{r}) \quad (2.11).$$

Здесь мы для простоты полагаем, что начальная фаза $\phi = 0$, а вектор \vec{k} - это волновой вектор, параллельный единичному вектору нормали к фазовому фронту \vec{n} . Величина и направление волнового вектора \vec{k} определяются соотношением:

$$\vec{k} = \vec{n} \cdot \frac{\omega}{v} = \vec{n} \cdot \omega \cdot \sqrt{\mu\epsilon} = \frac{\omega}{v} (\bar{x}_0 n_x + \bar{y}_0 n_y + \bar{z}_0 n_z),$$

где n_x, n_y и n_z - декартовы координаты единичного вектора \vec{n} .

Представим себе некоторую новую систему координат $x'y'z'$, введенную таким образом, что плоская волна распространяется вдоль оси OZ' . Тогда фаза волны определяется величиной $(\omega t - kz')$. Удобно ввести вектор

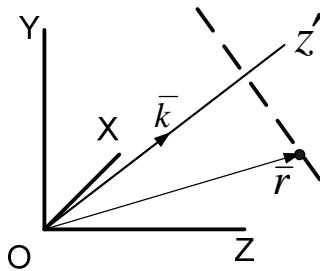


Рис. 2.1.

$\vec{k} = \vec{n} \cdot k$. По абсолютной величине он равен волновому числу k и направлен по движению волны. Положение точки пространства с координатами x, y, z (в старой системе координат), для которой необходимо записать выражение для поля плоской волны, определяется радиус – вектором $\vec{r} = \bar{x}_0 x + \bar{y}_0 y + \bar{z}_0 z$. Эта точка лежит на плоскости равной фазы $z' = \text{const}$. Тогда $kz' = \vec{k} \cdot \vec{r}$.

Чтобы перейти к основным координатам (x, y, z) , учтем, что: $\vec{k} = k\vec{z}'_0 = k(\bar{x}_0 \cos \alpha + \bar{y}_0 \cos \beta + \bar{z}_0 \cos \gamma)$, где α, β, γ - углы между единичным вектором нормали к волновому фронту волны \vec{z}'_0 и осями x, y, z старой системы координат. Итак:

$$kz' = \vec{k} \cdot \vec{r} = k(x \cdot \cos \alpha + y \cdot \cos \beta + z \cdot \cos \gamma) \quad (2.12).$$

В результате мы и получаем записанное ранее выражение для поля плоской волны, направление распространения которой определяется волновым вектором \bar{k} .

2.5. Структура поля электромагнитных волн

Используя свойства оператора «набла» $\bar{\nabla}$, систему уравнений Максвелла в дифференциальной форме для случая непроводящей среды можно представить в виде:

$$\begin{aligned} \text{rot}\bar{H} &= \frac{\partial\bar{D}}{\partial t} \Rightarrow \bar{\nabla} \times \bar{H} = i\omega\bar{D} \\ \text{rot}\bar{E} &= -\frac{\partial\bar{B}}{\partial t} \Rightarrow \bar{\nabla} \times \bar{E} = -i\omega\bar{B} \\ \text{div}\bar{D} &= 0 \Rightarrow \bar{\nabla} \cdot \bar{D} = 0 \\ \text{div}\bar{B} &= 0 \Rightarrow \bar{\nabla} \cdot \bar{B} = 0 \end{aligned} \quad (2.13).$$

Рассматривая плоскую гармоническую электромагнитную волну с полем $\dot{\bar{E}}(x, y, z, t) = \dot{\bar{E}}_m \cdot \exp[i(\omega t - k\bar{z})]$, видим, что воздействие на нее оператора $\bar{\nabla}$ эквивалентно скалярному или векторному умножению величины $-i\bar{k}$ на выражение для $\dot{\bar{E}}$:

$$\bar{\nabla} \cdot \dot{\bar{E}} \rightarrow -i\bar{k} \cdot \dot{\bar{E}} \quad (2.14),$$

$$\bar{\nabla} \times \dot{\bar{E}} \rightarrow -i\bar{k} \times \dot{\bar{E}} \quad (2.15).$$

С учетом этого система уравнений (2.12) примет вид:

$$\begin{aligned} -i\bar{k} \times \dot{\bar{H}} &= i\omega\dot{\bar{D}} & \text{или} & & \bar{k} \times \dot{\bar{H}} &= -\epsilon\omega\dot{\bar{E}} \\ -i\bar{k} \times \dot{\bar{E}} &= -i\omega\dot{\bar{B}} & & & \bar{k} \times \dot{\bar{E}} &= \mu\omega\dot{\bar{H}} \\ -i\bar{k} \times \dot{\bar{D}} &= 0 & & & \bar{k} \cdot \dot{\bar{D}} &= 0 \\ -i\bar{k} \times \dot{\bar{B}} &= 0 & & & \bar{k} \cdot \dot{\bar{B}} &= 0 \end{aligned} \quad (2.16).$$

Анализ данных соотношений позволяет легко получить наглядное представление о структуре поля плоской электромагнитной волны в однородной изотропной среде:

- из первого уравнения вытекает, что $\bar{E} \perp \bar{k}$ и $\bar{E} \perp \bar{H}$;
- из второго уравнения следует, что $\bar{H} \perp \bar{k}$ и $\bar{H} \perp \bar{E}$;
- из третьего и четвертого соотношений следует, что $\bar{D} \perp \bar{k}$, $\bar{B} \perp \bar{k}$.

Таким образом, в плоской электромагнитной волне, распространяющейся в однородной изотропной среде, вектора \bar{E} и \bar{H} ортогональны друг другу и ортогональны волновому вектору. Аналогично, вектора \bar{D} и \bar{B} также ортогональны волновому вектору. Соответственно, все эти вектора лежат в плоскости, перпендикулярной к волновому вектору.

Еще один важный вывод о параметрах плоской волны следует из рассмотрения первого и второго уравнений системы (2.16). Так, рассматривая первое из них, получим:

$$[\bar{k}, \bar{H}] = \left[\frac{\omega}{v} \bar{n}, \bar{H} \right] = \omega \sqrt{\epsilon \mu} \cdot [\bar{n}, \bar{H}].$$

Находя модули правой и левой частей, видим:

$$|\bar{k} \times \bar{H}| = \omega \sqrt{\epsilon \mu} \cdot H_m = e \omega \epsilon E_m \quad \text{и тогда}$$

$$\frac{E_m}{H_m} = W = \frac{\omega \sqrt{\epsilon \mu}}{\omega \epsilon} = \sqrt{\frac{\mu}{\epsilon}} \quad (2.17).$$

Таким образом, отношение амплитуд электрического и магнитного векторов в плоской волне определяется только материальными параметрами среды.

Величина этого отношения $W = \sqrt{\frac{\mu}{\epsilon}}$ имеет размерность [Ом] и называется волновым, или характеристическим, сопротивлением среды. Следует отметить, что термин «сопротивление» используется исключительно в силу размерности, но никак не связан с какими – либо тепловыми потерями электромагнитной энергии. Для вакуума величина волнового сопротивления составляет:

$$W_0 = \sqrt{\frac{\mu_0}{\epsilon_0}} = 120\pi \text{ [Ом]} \quad (2.18).$$

Итак, расположение векторов в плоской электромагнитной волне соответствует рисунку 2.2.

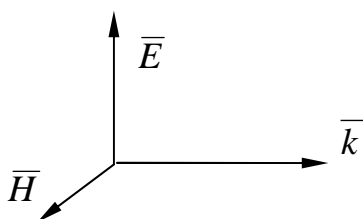


Рис. 2.2.

В общем случае, в зависимости от поляризации волны (линейная, эллиптическая, круговая) векторы \bar{E} и \bar{H} могут синхронно изменять свое положение в пространстве.

2.6. Поляризация плоских электромагнитных волн

Электромагнитная волна с векторами \bar{E} и \bar{H} , направление которых может быть однозначно определено в любой момент времени, называется поляризованной.

При случайных положениях векторов \bar{E} и \bar{H} в пространстве поле является неполяризованным.

Плоскость поляризации – плоскость, проходящая через вектор \bar{E} и направление распространения волны. В зависимости от того, какую фигуру описывает конец вектора \bar{E} в пространстве при распространении волны, различают линейную, эллиптическую и круговую поляризации.

Математически волну с произвольным видом поляризации можно представить в виде двух составляющих:

$$\begin{aligned}\bar{E}_x &= \bar{x}_0 E_{1m} \cos(\omega t - kz) \\ \bar{E}_y &= \bar{y}_0 E_{2m} \cos(\omega t - kz - \varphi)\end{aligned}\quad (2.19).$$

В общем случае эти составляющие в плоскости, ортогональной волновому вектору, имеют разные амплитуды и сдвинуты по фазе друг относительно друга.

Рассмотрим плоскость $z=0$, для которой эти выражения принимают вид:

$$\frac{E_x}{E_{1m}} = \cos(\omega t) \quad (2.20),$$

$$\frac{E_y}{E_{2m}} = \cos(\omega t) \cdot \cos \varphi + \sin(\omega t) \cdot \sin \varphi \quad (2.21).$$

Посмотрим, как изменяется положение вектора \bar{E} в плоскости XOY. Для этого исключим из данных уравнений временной множитель, полагая:

$$\cos(\omega t) = \frac{E_x}{E_{1m}}; \quad \sin(\omega t) = \sqrt{1 - \frac{E_x^2}{E_{1m}^2}}.$$

В результате из второго уравнения получим:

$$\left(\frac{E_x}{E_{1m}}\right)^2 + \left(\frac{E_y}{E_{2m}}\right)^2 - 2 \frac{E_x}{E_{1m}} \cdot \frac{E_y}{E_{2m}} \cdot \cos \varphi = \sin^2 \varphi \quad (2.22).$$

Рассмотрим характерные виды поляризации плоской волны, соответствующие различным фазовым сдвигам между составляющими вектора \bar{E} вдоль осей x и y.

1. $\varphi = 0$.

При этом $\left(\frac{E_x}{E_{1m}} - \frac{E_y}{E_{2m}}\right)^2 = 0$, откуда получим:

$$\frac{E_x}{E_{1m}} = \frac{E_y}{E_{2m}} \quad \rightarrow \quad E_y = \frac{E_{2m}}{E_{1m}} E_x \quad (2.23).$$

Видим, что это уравнение прямой с наклоном к оси X , определяемым отношением $\frac{E_{2m}}{E_{1m}}$. Действительно, при синфазном изменении составляющих

$E_x \sim \cos \omega t, E_y \sim \cos \omega t$, результирующий вектор \bar{E} изменяет только свою величину, а угол наклона плоскости, в которой он лежит, относительно поперечных координатных осей не изменяется. Очевидно, что поляризация будет линейной при выполнении условий: $\varphi = n\pi, (n = 0, \pm 1, \dots)$. Выражение для поля плоской волны с линейной поляризацией в общем случае можно записать в форме:

$$\bar{E} = (\bar{x}_0 E_{1m} + \bar{y}_0 E_{2m}) \cdot \cos(\omega t - kz) = E_0 (\bar{x}_0 \cos \alpha + \bar{y}_0 \sin \alpha) \cdot \cos(\omega t - kz) \quad (2.24),$$
 где $\alpha = \arctg(E_{2m}/E_{1m})$. В частных случаях, при поляризации в плоскостях XOZ и YOZ получим, соответственно:

$$\bar{E} = E_0 \bar{x}_0 \cdot \cos(\omega t - kz), \quad \bar{E} = E_0 \bar{y}_0 \cdot \cos(\omega t - kz).$$

2. Фазовый сдвиг $\varphi = 90^\circ$. При этом получим:

$$\left(\frac{E_x}{E_{1m}}\right)^2 + \left(\frac{E_y}{E_{2m}}\right)^2 = 1 \quad (2.25).$$

Это – уравнение эллипса с большой и малой полуосями, ориентированными по осям x и y . Направление вращения вектора \bar{E} определяется знаком φ . При $\varphi = 90^\circ$ из (2.19) следует: $E_x = E_0 \cos(\omega t)$, а $E_y = E_0 \cos(\omega t - 90^\circ) = E_0 \sin(\omega t)$.

Вращение вектора \bar{E} в этом случае происходит по часовой стрелке, если смотреть вдоль направления распространения волны. Такую поляризацию называют левой эллиптической поляризацией. Для $\varphi = -90^\circ$ вектор \bar{E} вращается в противоположном направлении – это правое вращение. Если выполняется условие $E_{1m} = E_{2m}$, то эллипс превращается в окружность, а поляризацию называют круговой. В случае круговой поляризации выражения для поля плоской волны могут быть записаны в виде:

$$\bar{E} = E_0 [\bar{x}_0 \cos(\omega t - kz) + \bar{y}_0 \sin(\omega t - kz)] \quad (2.26),$$

$$\dot{\bar{E}} = E_0 (\bar{x}_0 \pm i\bar{y}_0) \cdot \exp[i(\omega t - kz)] \quad (2.27).$$

Волна с круговой поляризацией представляется суммой двух линейно поляризованных волн с одинаковыми частотами и фазовым сдвигом $(\pi/2 \pm m\pi)$. В свою очередь, линейно поляризованная волна может быть представлена в виде суммы волн правой и левой круговой поляризации. Действительно, взяв для определенности волну с линейной поляризацией в плоскости XOZ , представим ее поле в виде:

$$\begin{aligned} \dot{\vec{E}} &= E_0 \bar{x}_0 \cdot \exp[i(\omega t - kz)] = \frac{E_0}{2} [(\bar{x}_0 + i\bar{y}_0) + (\bar{x}_0 - i\bar{y}_0)] \exp[i(\omega t - kz)] = \\ &= \frac{E_0}{2} \cdot (\bar{x}_0 + i\bar{y}_0) \cdot \exp[i(\omega t - kz)] + \frac{E_0}{2} \cdot (\bar{x}_0 - i\bar{y}_0) \cdot \exp[i(\omega t - kz)] \end{aligned} \quad (2.28).$$

3. Произвольный фазовый сдвиг φ . В этом случае поляризация электромагнитных волн также эллиптическая, но направления главных осей эллипса поляризации не совпадают с координатными осями X и Y . Эллипс вписан в прямоугольник с размерами сторон $2E_{m1}$ и $2E_{m2}$ (рис. 2.3). Угол Ψ между направлением главной оси эллипса поляризации и осью X можно выразить через E_{m1} , E_{m2} и фазовый сдвиг φ следующим образом:

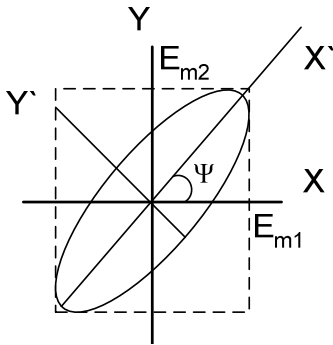


Рис. 2. 3.

$$\operatorname{tg} 2\Psi = \frac{2E_{1m}E_{2m}}{E_{1m}^2 - E_{2m}^2} \cdot \cos \varphi \quad (2.29).$$

Выражение для поля плоских волн при эллиптической поляризации, в комплексной форме принимает вид:

$$\dot{\vec{E}} = [\bar{x}_0 E_{1m} + \bar{y}_0 E_{2m} \exp(-i\varphi)] \cdot \exp[i(\omega t - kz)] \quad (2.30).$$

2.7. Вектор Джонса

В 1941 г. Кларком Джонсом предложен метод векторного описания поляризации световых волн. По имени “изобретателя” эти векторы называют векторами Джонса. Вектор Джонса имеет вид:

$$\vec{J} = \begin{bmatrix} E_x \exp(i\phi_x) \\ E_y \exp(i\phi_y) \end{bmatrix} \quad (2.31).$$

Из (2.31) видно, что данный вектор является комплексным, т.е. его компоненты задаются комплексными числами. Следует также отметить, что \vec{J} представляет собой вектор в абстрактном математическом, а не в реальном физическом пространстве.

Вектор Джонса содержит полную информацию об амплитудах и фазах компонент вектора электрического поля. При описании поляризации электромагнитных волн, как правило, используется нормированный вектор Джонса, удовлетворяющий условию:

$$\vec{J} \cdot \vec{J}^* = 1 \quad (2.32).$$

Соответственно, линейно поляризованная световая волна с заданным направлением вектора \vec{E} представляется вектором Джонса:

$$\begin{pmatrix} \cos \psi \\ \sin \psi \end{pmatrix} \quad (2.33),$$

где ψ - азимутальный угол между направлением вектора \bar{E} и осью X . Для направления поляризации, ортогонального заданному соотношением (2.33), заменив угол ψ на $\psi + \pi/2$, получим вектор Джонса в виде:

$$\begin{pmatrix} -\sin \psi \\ \cos \psi \end{pmatrix} \quad (2.34).$$

В случае поляризации волны вдоль осей X и Y получим:

$$\bar{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad \bar{y} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (2.35).$$

Для световых волн с правой и левой круговой поляризацией векторы Джонса записываются в форме:

$$\bar{R} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}; \quad \bar{L} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \quad (2.36).$$

Эти состояния круговой поляризации являются взаимно ортогональными, поскольку:

$$\bar{R}^* \cdot \bar{L} = 0 \quad (2.37).$$

Очевидно, что любая пара взаимно ортогональных векторов Джонса может играть роль базиса в пространстве всех векторов Джонса. Таким образом, любая поляризация может быть представлена суперпозицией двух взаимно ортогональных поляризаций \bar{x} и \bar{y} либо \bar{R} и \bar{L} . Так, как отмечалось ранее, базисные линейные поляризации можно разложить на две круговые поляризации и наоборот. Эти разложения имеют вид:

$$\bar{R} = \frac{1}{\sqrt{2}}(\bar{x} - i \cdot \bar{y})$$

$$\bar{L} = \frac{1}{\sqrt{2}}(\bar{x} + i \cdot \bar{y})$$

$$\bar{x} = \frac{1}{\sqrt{2}}(\bar{R} + \bar{L})$$

$$\bar{y} = \frac{1}{\sqrt{2}}(\bar{R} - \bar{L})$$

Как в частных случаях простых поляризаций, можно записать вектор Джонса и для более общего случая эллиптической поляризации света:

$$\bar{J}(\psi, \delta) = \begin{pmatrix} \cos \psi \\ \exp(i\delta) \cdot \sin \psi \end{pmatrix} \quad (2.38).$$

Отметим, что векторы Джонса находят широкое применение при вычислениях состояния поляризации света в случае его прохождения через

произвольную последовательность оптически анизотропных элементов и поляризаторов.

2.8. Элементы для преобразования состояния поляризации света.

Элементы, преобразующие состояние поляризации света и называемые обычно поляризаторами и поляроидами, играют важную роль в оптике и оптическом приборостроении. По роду используемых физических эффектов их можно разделить на элементы с анизотропией оптического поглощения (или оптическим дихроизмом) и с анизотропией показателя преломления.

Классическим примером материала с анизотропией оптического поглощения является кристалл турмалина, для которого поглощение обыкновенно поляризованного света значительно выше, чем для необыкновенно поляризованного. Однако он редко применяется для создания поляризационных элементов, т.к. в видимом диапазоне и поглощение необыкновенной волны оказывается существенным. Достаточно хорошими свойствами обладают синтетические органические пленочные материалы, в которых молекулы имеют вид длинных цепочек и ориентированы преимущественно в одном направлении, благодаря специальной обработке. Примером являются пленки поливинилового спирта с добавками йода. Они могут пропускать до 80% света, поляризованного в одном направлении, и менее 1% света, поляризованного в ортогональном направлении. Стоимость такого рода пленочных поляризаторов оказывается значительно более низкой по сравнению с таковой для монокристаллических поляризаторов. Однако существенным недостатком пленочных поляризаторов является возможность их разрушения при очень высокой интенсивности проходящего излучения.

Поляризующие элементы из природного исландского шпата (кальцит, CaCO_3) обладают высоким оптическим качеством, прозрачны в диапазоне длин волн от 0,2 до 2,2 мкм, устойчивы к воздействию интенсивного лазерного излучения. Существует несколько типов таких элементов. Это призмы Николя, Глана, Волластона, Рошона и т.д. Все они используют эффект анизотропии показателя преломления. Призмы Николя и Глана пропускают излучение лишь одной поляризации, призмы Волластона и Рошона на выходе имеют два ортогонально поляризованных световых луча, распространяющихся под некоторым углом относительно направления падающего излучения.

На рис. 2.4 показана геометрия призмы Николя. Элемент состоит из двух частей с оптически полированными гранями, склеенных специальным оптическим клеем – канадским бальзамом. Показатели преломления кальцита составляют 1,6584 (обыкновенный) и 1,4864 (необыкновенный) на длине волны $\lambda=0,63$ мкм. Показатель преломления канадского бальзама – 1,526.

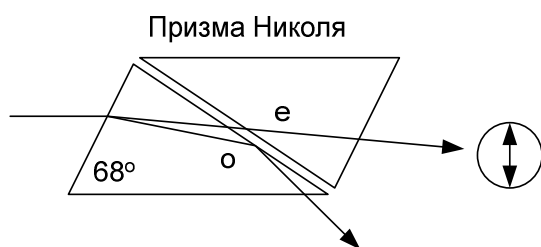


Рис. 2. 4.

Благодаря геометрии, обыкновенный луч испытывает полное внутреннее отражение от склейки и не выходит через выходную грань. Только излучение с необыкновенной поляризацией проходит через призму Николя. Призмы Глана также выполняются из кальцита, но вместо промежуточного слоя из канадского бальзама они имеют воздушную

прослойку (так называемую посадку на оптический контакт).

2.9. Отражение и преломление плоских световых волн на плоской границе раздела

При решении многих электродинамических задач электромагнитные волны распространяются в среде с резко или плавно изменяющимися электродинамическими параметрами. Типичный пример, с которым мы ежедневно встречаемся – прохождение световых волн через границы раздела двух прозрачных сред. При этом, как известно, возникают отраженные и преломленные волны, направления распространения и амплитуды которых зависят от соотношения между параметрами граничащих сред.

Рассмотрим основные явления при прохождении плоской электромагнитной волны через плоскую границу раздела и получим некоторые соотношения, описывающие законы отражения и преломления. Кроме того, получим выражения для амплитудных и фазовых соотношений между падающей и отраженной и преломленной волнами. Изучая наклонное падение плоской волны на границу, рассмотрим два качественно различных случая. В первом из них вектор \vec{E} параллелен граничной плоскости и перпендикулярен плоскости падения волны (плоскость, в которой лежат нормаль к границе раздела и волновой вектор падающей волны). Это случай горизонтальной поляризации. Во втором случае вектор \vec{E} лежит в плоскости падения - это случай вертикальной поляризации. Как будет видно из дальнейшего анализа, выражения для коэффициентов отражения горизонтально и вертикально поляризованных волн различны, что и обуславливает такой порядок рассмотрения. Кроме того, любая линейно поляризованная волна может быть разложена на компоненты с горизонтальной и вертикальной поляризацией, которые могут рассматриваться отдельно, что обосновывает корректность подобного подхода.

2.9.1. Законы отражения и преломления электромагнитных

ВОЛН

Ранее было записано выражение для плоской электромагнитной волны, распространяющейся в произвольном направлении, определяемом радиус – вектором \bar{r} , в виде:

$$\bar{E}(\bar{r}, t) = \bar{E}_m \cdot \cos(\omega t - \bar{k} \cdot \bar{r}) \quad (2.39).$$

Используя комплексные амплитуды, это выражение можно переписать в форме:

$$\dot{\bar{E}}(\bar{r}, t) = \dot{\bar{E}}_m \cdot \exp[-ik(x \cdot \cos\alpha + y \cdot \cos\beta + z \cdot \cos\gamma)] \quad (2.40).$$

Здесь временной множитель учтен в комплексной амплитуде волны, а $\cos(\alpha)$, $\cos(\beta)$, $\cos(\gamma)$ - направляющие косинусы нормали к волновому фронту.

Для рассмотрения отражения электромагнитных волн выберем следующую систему координат (рис. 2.5). Пусть граница раздела сред

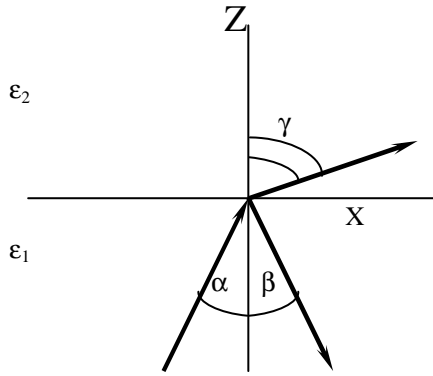


Рис. 2.5

совпадает с плоскостью XOY . Положим для определённости, что в падающей волне нормаль \bar{n} лежит в плоскости XZ . В то же время никаких ограничений на направление нормалей \bar{n}_1 (в отраженной волне) и \bar{n}_2 (в преломленной волне) не накладываем. В таком случае поля падающей, отраженной и преломленной волн, соответственно, будут описываться

соотношениями:

$$\dot{\bar{E}}_0(\bar{r}, t) = \dot{\bar{E}}_{m0} \cdot \exp[-ik_0(x \cdot \cos\alpha_0 + z \cdot \cos\gamma_0)] \quad (2.41),$$

$$\dot{\bar{E}}_1(\bar{r}, t) = \dot{\bar{E}}_{m1} \cdot \exp[-ik_1(x \cdot \cos\alpha_1 + y \cdot \cos\beta_1 + z \cdot \cos\gamma_1)] \quad (2.42),$$

$$\dot{\bar{E}}_2(\bar{r}, t) = \dot{\bar{E}}_{m2} \cdot \exp[-ik_2(x \cdot \cos\alpha_2 + y \cdot \cos\beta_2 + z \cdot \cos\gamma_2)] \quad (2.43).$$

Запишем теперь граничное условие для тангенциальных компонент вектора \bar{E} при $z=0$: $E_{\tau 0} + E_{\tau 1} = E_{\tau 2}$. Это условие должно выполняться в любой момент времени t и в любой точке на границе раздела, т.е. при любых координатах x , y . Для простоты рассмотрим случай горизонтальной поляризации. Подставляя в уравнение для граничного условия выражения для полей падающей, отраженной и преломленной волн, получим:

$$\begin{aligned} & E_{m0} \cdot \exp(-i\omega_0 t) \cdot \exp[-ik_0(x \cdot \cos\alpha_0)] + \\ & + E_{m1} \cdot \exp(-i\omega_1 t) \cdot \exp[-ik_1(x \cdot \cos\alpha_1 + y \cdot \cos\beta_1)] = \\ & = E_{m2} \cdot \exp(-i\omega_2 t) \cdot \exp[-ik_2(x \cdot \cos\alpha_2 + y \cdot \cos\beta_2)] \end{aligned} \quad (2.44).$$

Но это тождество справедливо лишь при выполнении следующих условий:

1. $\omega_0 = \omega_1 = \omega_2$, т.е. частота электромагнитной волны не изменяется при отражении от неподвижной границы.
2. $\cos \beta_1 = \cos \beta_2 = 0$ т.е. волновые векторы отраженной и преломленной волн лежат в плоскости падения, как и волновой вектор падающей волны.
3. $k_0 \cdot \cos \alpha_0 = k_1 \cdot \cos \alpha_1 = k_2 \cdot \cos \alpha_2$

или, с учетом равенства частот этих волн:

$$\frac{\cos \alpha_0}{v_0} = \frac{\cos \alpha_1}{v_1} = \frac{\cos \alpha_2}{v_2}$$

Отсюда вытекает, во-первых, известный закон, гласящий о том, что угол падения равен углу отражения: $\alpha_0 = \alpha_1$ (т.к. $v_0 = v_1$). Второй важный вывод вытекает из второй части приведенного соотношения, которую можно записать в виде: $n_1 \cdot \cos \alpha_1 = n_2 \cdot \cos \alpha_2$. Углы, обозначенные в нашем случае как α_i , обычно называют углами скольжения. Но в подобных задачах принято оперировать углами падения (угол между направлением волнового вектора волны и нормалью к границе раздела), обозначаемыми здесь как γ_i . Учитывая, что $\alpha + \gamma = \pi/2$, последнее соотношение запишем в виде:

$$n_1 \cdot \sin \gamma_1 = n_2 \cdot \sin \gamma_2 \quad (2.45),$$

а это не что иное, как известный закон Снеллиуса.

Справедливость полученных законов многократно подтверждена экспериментально, а сами законы хорошо известны и изучаются уже в школьном курсе физики. Здесь продемонстрировано, что они логически вытекают непосредственно из решений уравнений Максвелла и соответствующих граничных условий.

2.9.2. Формулы Френеля

Получим соотношения, определяющие амплитуды отраженных и преломленных электромагнитных волн - формулы Френеля. При этом будем рассматривать, как уже отмечалось, вертикально и горизонтально поляризованные волны отдельно. Избавимся от временных зависимостей и будем иметь дело только с проекциями амплитуд векторов \vec{E} и \vec{H} (в силу линейности ситуации).

Волна с горизонтальной поляризацией

Рассматриваем ситуацию, когда граница раздела двух сред совпадает с плоскостью XU , а плоскость падения световой волны – с плоскостью XZ . Для случая горизонтальной поляризации вектор \vec{E} параллелен границе раздела, а вектор \vec{H} лежит в плоскости падения. Граничные условия для тангенциальных компонент векторов \vec{E} и \vec{H} имеют в данном случае вид:

$$E_{m0} + E_{m1} = E_{m2} \quad (2.46),$$

$$H_{m0} \cos \theta_0 - H_{m1} \cos \theta_1 = H_{m2} \cos \theta_2 \quad (2.47).$$

Используя понятие волнового сопротивления среды, можно записать:

$H_m = \frac{E_m}{W}$, тогда второе из этих соотношений примет вид:

$$\frac{E_{m0} \cos \theta_0 - E_{m1} \cos \theta_1}{W_1} = \frac{E_{m2} \cos \theta_2}{W_2} \quad (2.48).$$

Теперь из данного соотношения с учетом (8.8) можно легко найти выражения для коэффициентов Френеля, связывающие амплитуды отраженной и прошедшей волн с падающей. Так, для отраженной волны получим:

$$\frac{E_{m1}}{E_{m0}} = R_{\perp} = \frac{W_2 \cos \theta_0 - W_1 \cos \theta_2}{W_2 \cos \theta_0 + W_1 \cos \theta_2} \quad (2.49),$$

где индекс \perp означает, что вектор \vec{E} перпендикулярен плоскости падения, т.е. волна имеет горизонтальную поляризацию. В данном выражении учтено, что $W_0 = W_1$ и $\theta_0 = \theta_1$. Аналогичным образом получим формулу для коэффициента Френеля (коэффициента пропускания), определяющего амплитуду волны, преломленной во вторую среду:

$$\frac{E_{m2}}{E_{m0}} = T_{\perp} = \frac{2W_2 \cos \theta_0}{W_2 \cos \theta_0 + W_1 \cos \theta_2} \quad (2.50).$$

Волна с вертикальной поляризацией

Повторяя ту же процедуру, получим выражения для коэффициентов Френеля при вертикальной поляризации электромагнитных волн (вектор \vec{H} параллелен границе раздела, а вектор \vec{E} лежит в плоскости падения). Граничные условия в данном случае, имеют вид:

$$H_{m0} + H_{m1} = H_{m2} \quad (2.51),$$

$$E_{m0} \cdot \cos \theta_0 - E_{m1} \cdot \cos \theta_1 = E_{m2} \cdot \cos \theta_2 \quad (2.52).$$

Выразим теперь амплитуды E_{mi} во втором уравнении через H_{mi} и найдем соответствующие коэффициенты Френеля:

$$W_0 \cdot H_{m0} \cdot \cos \theta_0 - W_1 \cdot H_{m1} \cdot \cos \theta_1 = W_2 \cdot (H_{m0} + H_{m1}) \cdot \cos \theta_2$$

$$H_{m0} \cdot (W_0 \cdot \cos \theta_0 - W_2 \cdot \cos \theta_2) = H_{m1} \cdot (W_1 \cdot \cos \theta_1 + W_2 \cdot \cos \theta_2)$$

Коэффициент Френеля для отраженной волны:

$$R_{\parallel} = \frac{H_{m1}}{H_{m0}} = \frac{W_1 \cdot \cos \theta_0 - W_2 \cdot \cos \theta_2}{W_1 \cdot \cos \theta_0 + W_2 \cdot \cos \theta_2} \quad (2.53),$$

и для преломленной волны:

$$T_{\parallel} = \frac{H_{2m}}{H_{0m}} = \frac{2W_2 \cos \theta_2}{W_1 \cos \theta_0 + W_2 \cos \theta_2} \quad (2.54).$$

Здесь индекс \parallel указывает на то, что вектор \vec{E} лежит в плоскости падения. Для диэлектрической среды выражения для коэффициентов Френеля обычно записывают в несколько другой форме, учитывая что в этом случае $\mu=1$, а $W = 1/\sqrt{\epsilon}$:

$$R_{\perp} = \frac{\sqrt{\epsilon_1} \cos \theta_0 - \sqrt{\epsilon_2} \cos \theta_2}{\sqrt{\epsilon_1} \cos \theta_0 + \sqrt{\epsilon_2} \cos \theta_2} \quad (2.55),$$

$$T_{\perp} = \frac{2\sqrt{\epsilon_1} \cos \theta_0}{\sqrt{\epsilon_1} \cos \theta_0 + \sqrt{\epsilon_2} \cos \theta_2} \quad (2.55),$$

$$R_{\parallel} = \frac{\sqrt{\epsilon_2} \cdot \cos \theta_0 - \sqrt{\epsilon_1} \cdot \cos \theta_2}{\sqrt{\epsilon_2} \cdot \cos \theta_0 + \sqrt{\epsilon_1} \cdot \cos \theta_2} \quad (2.56),$$

$$T_{\parallel} = \frac{2\sqrt{\epsilon_1} \cos \theta_2}{\sqrt{\epsilon_1} \cos \theta_0 + \sqrt{\epsilon_1} \cos \theta_2} \quad (2.57).$$

Из этих соотношений можно видеть, что в случае вертикальной поляризации коэффициент Френеля R_{\parallel} и, соответственно, амплитуда отраженной волны при некотором угле падения могут быть равны нулю. Действительно, рассматривая для определенности среды с соотношением параметров $\epsilon_2 > \epsilon_1$ и находя значения коэффициентов R_{\perp} и R_{\parallel} при углах падения волны на границу раздела $\theta=0$ и $\theta=90^\circ$, получим:

$$R_{\perp}(0) = \frac{\sqrt{\epsilon_1} - \sqrt{\epsilon_2}}{\sqrt{\epsilon_1} + \sqrt{\epsilon_2}} < 0 ; \quad R_{\perp}(90^\circ) = \frac{-\sqrt{\epsilon_2} \cos \theta_2}{\sqrt{\epsilon_2} \cos \theta_2} = -1 \quad (2.58).$$

Видим, что в данном диапазоне углов падения $R_{\perp} < 0$. Для R_{\parallel} имеем:

$$R_{\parallel}(0) = \frac{\sqrt{\epsilon_2} - \sqrt{\epsilon_1}}{\sqrt{\epsilon_2} + \sqrt{\epsilon_1}} > 0 ; \quad R_{\parallel}(90^\circ) = \frac{-\sqrt{\epsilon_1} \cdot \cos \theta_2}{\sqrt{\epsilon_1} \cdot \cos \theta_2} = -1 \quad (2.59),$$

т.е. знак коэффициента отражения при изменении угла падения в данном диапазоне меняется с положительного на отрицательный для волны с вертикальной поляризацией. А это значит, что при некотором угле падения отраженная волна с такой поляризацией исчезает. Данный угол, называемый углом Брюстера, можно найти из условия:

$$n_2 \cdot \cos \theta_1 - n_1 \cdot \cos \theta_2 = 0 \quad (2.60).$$

Используя закон Снеллиуса $n_1 \cdot \sin \theta_1 = n_2 \cdot \sin \theta_2$, выразим n_2 в виде:

$$n_2 = n_1 \cdot \frac{\sin \theta_1}{\sin \theta_2}$$

и подставим его в указанное условие (2.60). В результате придем к соотношению: $\sin \theta_1 \cdot \cos \theta_1 = \sin \theta_2 \cdot \cos \theta_2$. Но, так как $\theta_1 \neq \theta_2$, то это соотношение может выполняться лишь в случае, когда $\theta_2 = (90^\circ - \theta_1)$. При этом $\sin \theta_1 = \cos \theta_2$. Поэтому условие (2.60), или $R_{\perp} = 0$, примет вид: $n_2 \cdot \cos \theta_1 - n_1 \cdot \sin \theta_1 = 0$ или

отсюда:

$$\operatorname{tg}\theta_1 = \operatorname{tg}\theta_B = \frac{n_2}{n_1} \quad (2.61).$$

Эффект Брюстера или эффект полного прохождения вертикально поляризованной электромагнитной волны через границу раздела диэлектриков широко используется на практике. Одним из примеров являются газоразрядные трубки газовых лазеров. Их выходные окна наклонены к оптической оси трубки под углом Брюстера для выделения излучения с линейной поляризацией и минимизации потерь оптического резонатора для данного излучения.

2.9.3. Явление полного внутреннего отражения

Рассмотрим случай, когда электромагнитная волна падает на границу раздела из оптически более плотной среды ($n_1 > n_2$). Из закона Снеллиуса следует, что при этом угол преломления больше, чем угол падения: $\theta_2 > \theta_1$. Таким образом, при величине угла падения, отвечающей условию:

$$\sin\theta_1 = \frac{n_2}{n_1} \quad (2.62)$$

преломленная волна должна распространяться вдоль границы раздела ($\sin\theta_2=1$). Угол $\theta_{1\text{кр}} = \arcsin(n_2/n_1)$ называют критическим углом полного внутреннего отражения. Очевидно, что при углах падения больше критического электромагнитная волна во второй среде уже не может уходить от границы. Выясним, какова структура поля электромагнитной волны в первой и во второй средах при таких углах падения. В этом случае выражение для величины $\cos\theta_2$, необходимой для определения коэффициентов Френеля, можно записать в форме:

$$\cos\theta_2 = \sqrt{1 - \sin^2\theta_2} = \pm i \cdot \sqrt{\frac{n_1^2}{n_2^2} \cdot \sin^2\theta_1 - 1} = \pm i \cdot \sqrt{\frac{\epsilon_1}{\epsilon_2} \cdot \sin^2\theta_1 - 1} \quad (2.63).$$

Тогда выражения для коэффициентов отражения при горизонтальной и вертикальной поляризации принимают вид:

$$R_{\perp} = \frac{\sqrt{\epsilon_1} \cdot \cos\theta_1 - i\sqrt{\epsilon_2} \cdot \sqrt{\frac{\epsilon_1}{\epsilon_2} \sin^2\theta_1 - 1}}{\sqrt{\epsilon_1} \cdot \cos\theta_1 + i\sqrt{\epsilon_2} \cdot \sqrt{\frac{\epsilon_1}{\epsilon_2} \sin^2\theta_1 - 1}} \quad (2.64),$$

$$\begin{aligned}
R_{//} &= \frac{\sqrt{\varepsilon_2} \cdot \cos \theta_1 - i\sqrt{\varepsilon_1} \cdot \sqrt{\frac{\varepsilon_1 \sin^2 \theta_1 - 1}{\varepsilon_2}}}{\sqrt{\varepsilon_1} \cdot \cos \theta_1 + i\sqrt{\varepsilon_1} \cdot \sqrt{\frac{\varepsilon_1 \sin^2 \theta_1 - 1}{\varepsilon_2}}} = \\
&= \frac{\sqrt{\varepsilon_2} \cdot \cos \theta_1 - i\sqrt{\frac{\varepsilon_1}{\varepsilon_2}} \cdot \sqrt{\varepsilon_1 \cdot \sin^2 \theta_1 - \varepsilon_2}}{\sqrt{\varepsilon_1} \cdot \cos \theta_1 + i\sqrt{\frac{\varepsilon_1}{\varepsilon_2}} \cdot \sqrt{\varepsilon_1 \cdot \sin^2 \theta_1 - \varepsilon_2}}
\end{aligned} \tag{2.65}$$

Таким образом, при $\theta_1 > \theta_{1\text{кр}}$ коэффициенты отражения являются комплексными величинами: $R = |R| \cdot \exp(i\varphi)$, причем $|R_{//}| = |R_{\perp}| = 1$. А это значит, что потоки энергии, связанные с падающей и отраженной волнами, одинаковы.

Поле (E_{Σ}) в первой среде найдем, учитывая наличие в ней как падающей (E_i), так и отраженной (E_r) волн: $E_{\Sigma} = E_i + E_r$. Выражения для комплексных амплитуд падающей и отраженной волн, очевидно, можно записать в виде:

$$E_i = E_{m0} \cdot \exp[-i(k \cdot \sin \theta \cdot x + k \cdot \cos \theta \cdot z)] \tag{2.66}$$

$$E_r = E_{m0} \cdot |R| \cdot \exp(i \cdot \varphi) \cdot \exp[-i(k \cdot \sin \theta \cdot x - k \cdot \cos \theta \cdot z)] \tag{2.67}$$

Здесь φ - фаза коэффициента отражения. Домножая E_i на единицу, представленную в форме $\exp(-i\varphi/2) \cdot \exp(i\varphi/2)$, и, проведя простые преобразования, получим:

$$\begin{aligned}
E_{\Sigma} &= E_{m0} \cdot \exp(-ik \cdot \sin \theta \cdot x) \cdot \exp(-ik \cdot \cos \theta \cdot z) \cdot \exp\left(i\frac{\varphi}{2}\right) \cdot \exp\left(-i\frac{\varphi}{2}\right) + \\
&+ E_{m0} \cdot \exp(-ik \cdot \sin \theta \cdot x) \cdot \exp(ik \cdot \cos \theta \cdot z) \cdot \exp\left(i\frac{\varphi}{2}\right) \cdot \exp\left(i\frac{\varphi}{2}\right) = \\
&= E_{m0} \cdot \exp(-ik \cdot \sin \theta \cdot x) \cdot \exp\left(i\frac{\varphi}{2}\right) \cdot \left[\exp(-ik \cdot \cos \theta \cdot z) \cdot \exp\left(-i\frac{\varphi}{2}\right) + \right. \\
&\left. + \exp(ik \cdot \cos \theta \cdot z) \cdot \exp\left(i\frac{\varphi}{2}\right) \right] = \\
&= 2E_{m0} \cdot \cos\left(k \cdot \cos \theta \cdot z + \frac{\varphi}{2}\right) \cdot \exp\left[-i\left(k \sin \theta \cdot x - \frac{\varphi}{2}\right)\right]
\end{aligned} \tag{2.68}$$

Таким образом, суммарное поле в первой среде представляет собой волну, бегущую вдоль x с фазовой скоростью, определяемой соотношением:

$$k_x = \frac{\omega}{v_x} = k \cdot \sin \theta, \text{ откуда } v_x = \frac{v}{\sin \theta}.$$

Распределение амплитуды поля данной волны в направлении оси Z представляет собой стоячую волну с волновым числом $k_z = k \cdot \cos \theta$. Волна такого рода, у которой амплитуда возмущения в плоскости волнового фронта изменяется, называется неоднородной плоской волной, направляемой границей раздела сред.

Основные свойства неоднородной плоской волны:

- 1) Плоскости равных фаз ($x=const$) и равных амплитуд ($z=const$) ортогональны.
- 2) В отличие от однородной, неоднородная волна имеет продольные компоненты полей (поля H - при горизонтальной поляризации и E - при вертикальной).
- 3) Фазовая скорость неоднородной волны больше, чем фазовая скорость однородной волны в первой среде $v_x = v/\sin \theta$, но меньше, чем такая скорость во второй среде (т.к. величина θ ограничена величиной критического угла).

Посмотрим теперь, каково распределение поля во второй среде. В общем случае оно должно быть записано в форме:

$$E_2 = E_{m2} \cdot \exp(i\omega t) \cdot \exp[-i(k_2 \cdot \cos \alpha_2 \cdot x + k_2 \cdot \cos \theta_2 \cdot z)]$$

Но $k_2 \cdot \cos \alpha_2 = k_2 \cdot \sin \theta_2 = k_1 \cdot \sin \theta_1$, как вытекает из законов отражения. Поэтому, с учетом выражения для $\cos \theta_2$ получим:

$$E_2 = E_{m2} \cdot \exp(i\omega t) \cdot \exp[-ik_2 \cdot x \cdot \sin \theta_1] \cdot \exp\left[-k_2 \sqrt{\frac{n_1^2}{n_2^2} \sin^2 \theta_1 - 1} \cdot z\right] \quad (2.69).$$

Видим, что поле во второй среде также представляет собой неоднородную плоскую волну, бегущую вдоль границы раздела. Но в данном случае амплитуда поля очень быстро уменьшается по экспоненциальному закону с удалением от границы (поле становится пренебрежимо малым на расстояниях $z \approx \lambda$). Тем не менее, оно может быть обнаружено в физических экспериментах даже для света видимого диапазона длин волн.

Рассмотрим фазовые соотношения при полном внутреннем отражении (как уже отмечено, коэффициенты отражения являются в этом случае комплексными величинами). Из вида выражений для R_{\perp} , R_{\parallel} следует, что они могут быть представлены в форме: $R=Z/Z^*$, где Z^* - комплексно сопряженная величина к Z . Фазы коэффициентов отражения для волн горизонтальной и вертикальной поляризации различны и являются функциями угла падения. Из представления $R=Z/Z^*$ и с учетом соотношения $Z=|Z| \cdot \exp(i\phi)$, вытекает:

$$R = |R| \cdot \exp(i\varphi) = \frac{|Z| \cdot \exp(i\alpha)}{|Z| \cdot \exp(-i\alpha)} = \exp(2i\alpha)$$

Поэтому отсюда очень просто находим фазу $\varphi/2$:

$$\operatorname{tg}(\varphi_{\perp}/2) = \operatorname{tg}\alpha_{\perp} = -\frac{\sqrt{n_1^2 \cdot \sin^2 \theta_1 - n_2^2}}{n_1 \cdot \cos \theta_1} \quad (2.70).$$

Аналогично:

$$\operatorname{tg}(\varphi_{\parallel}/2) = \operatorname{tg}\alpha_{\parallel} = \frac{-\left(\frac{n_1}{n_2}\right) \sqrt{n_1^2 \cdot \sin^2 \theta_1 - n_2^2}}{n_2 \cdot \cos \theta_1} \quad (2.71)$$

И действительно, видим, что величины фазовых сдвигов при полном внутреннем отражении от границы раздела ЭМВ с разной поляризацией различны.

Явление полного внутреннего отражения электромагнитных волн от границы раздела двух сред находит широкое применение в современной технике. В случае радиодиапазона оно является основой для построения диэлектрических волноводных элементов. Особую роль данный эффект и подобные волноводные элементы играют в оптике. Волоконные световоды, каналные и планарные оптические волноводы представляют собой основу элементной базы современных оптических систем связи и обработки информации.

2.10. Когерентность световых волн

Под понятием когерентности волновых процессов подразумевается согласованность, корреляция каких – либо характеристик этих процессов, рассматриваемых в разных точках пространства в разные моменты времени.

2.10.1. Понятие полной когерентности

Рассмотрим следующий пример. Пусть точечный источник света испускает монохроматическую сферическую волну. С помощью линзы часть светового поля преобразуется в плоскую волну, для которой поле определяется выражением:

$$\dot{E}_x(z, t) = E_{mx} \cdot \cos(\omega t - kz + \varphi),$$

где φ – фаза волны в точке $z=0$ в момент времени $t=0$. Для определенности считаем плоскую волну поляризованной линейно. Пусть теперь мы имеем две плоские световые волны $\dot{E}_{1x}(z, t)$ и $\dot{E}_{2x}(z, t)$ с одинаковыми амплитудами и частотами, но разными начальными фазами, от двух точечных источников. Суммарная интенсивность света в некоторой точке z задается соотношением:

$$I(z, t) = \frac{1}{T} \int_0^T [E_{1x}(z, t) + E_{2x}(z, t)]^2 dt \quad (2.72).$$

Для сокращения записей операцию усреднения по времени будем представлять также в форме:

$$I(z, t) = \left\langle [E_{1x}(z, t) + E_{2x}(z, t)]^2 \right\rangle \quad (2.73).$$

Необходимость усреднения по времени при вычислении энергетических характеристик светового поля, как уже отмечалось, обусловлена инерционностью любых фотоприемных элементов, что не позволяет измерить мгновенное значение вектора поля. Подставляя в (2.72) выражения для $E_{1x}(z, t)$ и $E_{2x}(z, t)$, получим:

$$\begin{aligned} I(z, t) &= \frac{1}{T} \int_0^T [E_{1m} \cdot \cos(\omega t - kz + \varphi_1) + E_{2m} \cdot \cos(\omega t - kz + \varphi_2)]^2 dt = \\ &= \frac{1}{T} \int_0^T E_{1m}^2 \cdot \cos^2(\omega t - kz + \varphi_1) dt + \frac{1}{T} \int_0^T E_{2m}^2 \cdot \cos^2(\omega t - kz + \varphi_2) dt + \\ &+ \frac{1}{T} \int_0^T 2 \cdot E_{1m} \cdot E_{2m} \cdot \cos(\omega t - kz + \varphi_1) \cos(\omega t - kz + \varphi_2) dt = \\ &= \frac{1}{2T} \int_0^T E_{1m}^2 [1 + \cos(2\omega t - 2kz + 2\varphi_1)] dt + \frac{1}{2T} \int_0^T E_{2m}^2 [1 + \cos(2\omega t - 2kz + 2\varphi_2)] dt + \\ &+ \frac{1}{T} \int_0^T E_{1m} \cdot E_{2m} \cdot [\cos(2\omega t - 2kz + \varphi_1 + \varphi_2) + \cos(\varphi_1 - \varphi_2)] dt = \\ &= \frac{E_{1m}^2}{2} + \frac{E_{2m}^2}{2} + E_{1m} E_{2m} \frac{1}{T} \int_0^T \cos(\varphi_1 - \varphi_2) dt = I_1 + I_2 + 2\sqrt{I_1 I_2} \frac{1}{T} \int_0^T \cos(\varphi_1 - \varphi_2) dt. \end{aligned}$$

Таким образом,

$$I(z, t) = I_1 + I_2 + 2\sqrt{I_1 I_2} \frac{1}{T} \int_0^T \cos(\varphi_1 - \varphi_2) dt \quad (2.74).$$

Здесь I_1 и I_2 – интенсивности света в точке z , создаваемые первым и вторым источниками, соответственно. Очевидно, что суммарная интенсивность поля в данной точке существенным образом зависит от усредненного по времени значения разности фаз возмущений в этой точке, создаваемых разными источниками. При $I_1 = I_2 = I_0$ величина I в точке z может изменяться от 0 до $4I_0$ в зависимости от разности фаз $(\varphi_1 - \varphi_2)$, если эта разность фаз постоянна во времени. Если же происходит хаотичное изменение значения $(\varphi_1 - \varphi_2)$, то $\langle \cos(\varphi_1 - \varphi_2) \rangle = 0$ и тогда $I = I_1 + I_2$. Первый из этих случаев соответствует **полной когерентности** двух волн, а второй – их **полной некогерентности**.

Рассмотренная ситуация является идеализированной, поскольку мы говорили о точечных источниках и строго монохроматических полях. Реальные источники всегда имеют конечные размеры, а реальные световые

волны не могут быть строго монохроматическими. Поэтому попытаемся уточнить понятие когерентности.

2.10.2. Временная когерентность

Будем считать источник излучения точечным, но само излучение полагаем не строго гармоническим, а имеющим форму цуга:

$$E = E_0 \cdot \exp(i\omega t - t/\tau) \quad (2.75),$$

где τ – некоторая постоянная. Пусть цуг расщепляется на два с помощью полупрозрачного зеркала, а глухое зеркало изменяет направление распространения излучения, так что оба цуга встречаются в точке C (рис.2.6). Если разность путей цугов L такова, что $L/c \ll \tau$, то в точке C , очевидно, встречаются почти одни и те же части исходного цуга. Если же $L/c > \tau$, то в

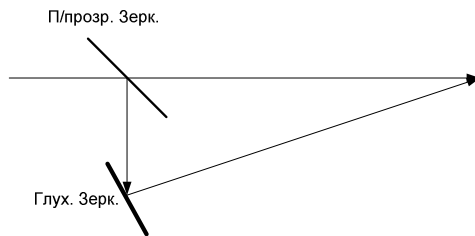


Рис. 2.6.

данной точке встречаются уже разные цуги. Произведение τc определяет длину когерентности, а сама рассмотренная ситуация иллюстрирует понятие временной когерентности. Постоянную времени τ называют временем когерентности.

2.10.3. Функция взаимной когерентности

При рассмотрении источников излучения конечных размеров используем более общий подход, суть которого иллюстрируется рисунком 2.7. Здесь излучение источника конечных размеров проходит через экран с двумя щелями, и мы находим интенсивность света в некоторой точке M плоскости P . Очевидно, она определяется соотношением:

$$I = \langle E(t) \cdot E^*(t) \rangle = \langle [E_1(t - \tau_1) + E_2(t - \tau_2)] \times [E_1^*(t - \tau_1) + E_2^*(t - \tau_2)] \rangle \quad (2.76).$$

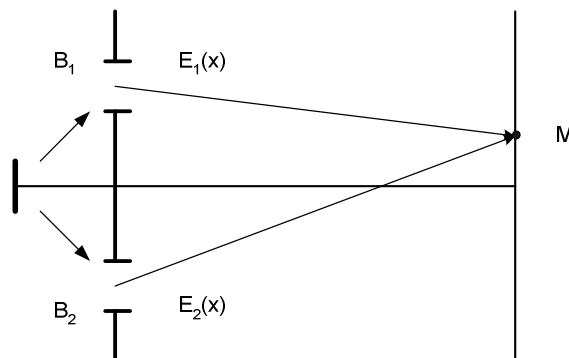


Рис. 2.7.

Введем следующие обозначения:

$$\langle E_1(t - \tau_1) \cdot E_1^*(t - \tau_1) \rangle = \langle E_1(t) \cdot E_1^*(t) \rangle = I_1$$

$$\langle E_2(t - \tau_2) \cdot E_2^*(t - \tau_2) \rangle = \langle E_2(t) \cdot E_2^*(t) \rangle = I_2$$

Здесь I_1 – интенсивность света в точке M , создаваемая при закрытом отверстии B_2 ; а I_2 – интенсивность света в этой точке, создаваемая при закрытом отверстии B_1 . Далее:

$$\begin{aligned} & \langle E_1(t - \tau_1) \cdot E_2^*(t - \tau_2) + E_2(t - \tau_2) \cdot E_1^*(t - \tau_1) \rangle = \\ & = \langle E_1(t - \tau_1 + \tau_2) \cdot E_2^*(t) + E_2(t) \cdot E_1^*(t - \tau_1 + \tau_2) \rangle = 2 \operatorname{Re} \langle E_1(t + \tau) \cdot E_2^*(t) \rangle, \end{aligned}$$

где $\tau = \tau_2 - \tau_1$.

Тогда выражение для интенсивности света в точке M можно записать в виде:

$$I_{\Sigma} = I_1 + I_2 + 2 \operatorname{Re}[\langle E_1(t + \tau) E_2^*(t) \rangle] \quad (2.77).$$

Последнее слагаемое в (2.77) описывает эффект корреляции световых возмущений, приходящих в точку M от отверстий B_1 и B_2 . Величину $\Gamma_{12}(\tau) = \langle E_1(t + \tau) E_2^*(t) \rangle$ называют функцией взаимной когерентности.

Однако наиболее часто используется нормированная функция взаимной когерентности γ_{12} , определяемая соотношением:

$$\gamma_{12}(\tau) = \frac{\Gamma_{12}(\tau)}{\sqrt{I_1 I_2}} \quad (2.78).$$

Таким образом, для интенсивности света в точке M окончательно имеем:

$$I_{\Sigma} = I_1 + I_2 + 2\sqrt{I_1 I_2} \operatorname{Re}[\gamma_{12}(\tau)] \quad (2.79).$$

Очевидно, что величина γ_{12} изменяется в пределах от 0 до 1.

2.10.4. Временная и пространственная когерентность

Рассмотрим модифицированную схему эксперимента (рис. 2.8). В этом случае в точку M приходят волновые возмущения только от отверстия B_1 .

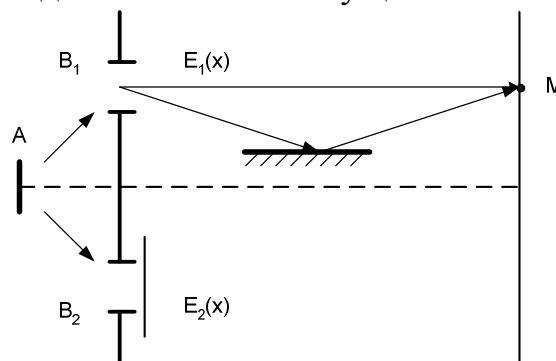


Рис. 2.8.

Соответственно, выражение для нормированной функции взаимной когерентности принимает вид:

$$\gamma_{11}(\tau) = \langle E_1(t + \tau)E_1^*(t) \rangle / I_1 \quad (2.80).$$

Очевидно, что функция $\gamma_{11}(\tau)$ описывает корреляцию световых колебаний в некоторой точке пространства в разные моменты времени, т.е. она характеризует временную когерентность светового поля.

Второй частный случай соответствует равенству оптических путей волновых возмущений от отверстий B_1 и B_2 до точки M (рис. 2.9). Тогда $B_1M=B_2M$ и $\tau=0$. Из соотношения (2.78) для данного случая получим:

$$\gamma_{12}(0) = \langle E_1(t)E_2^*(t) \rangle / \sqrt{I_1I_2}.$$

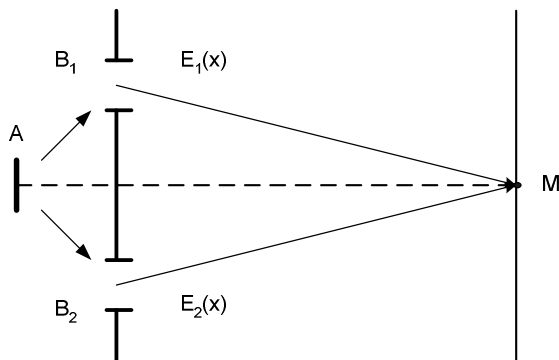


Рис. 2.9.

Это соотношение, очевидно, характеризует корреляцию световых колебаний в двух точках пространства в один и тот же момент времени, т.е. пространственную когерентность светового поля.

3. ОПТИКА ОГРАНИЧЕННЫХ СВЕТОВЫХ ПУЧКОВ

Плоская однородная световая волна – это всего лишь математическая абстракция. Реально в оптике имеют дело с ограниченными световыми пучками. В случае монохроматического лазерного излучения подобные пучки иногда могут рассматриваться как плоские световые волны. Однако во многих случаях их поведение существенно отличается от поведения плоских волн. Это обусловлено эффектами дифракции. Основной задачей теории дифракции является определение амплитуды, фазы и поляризации (в случае векторных волн) волнового возмущения в любой точке пространства при наличии препятствий на пути распространения поля. Ввиду сложности дифракционных задач, важную роль играют различные приближенные методы их решения.

3.1. Угловой спектр плоских волн

Представление световых полей в виде углового спектра плоских волн часто используется при приближенных решениях дифракционных задач. Данный подход, как и Фурье – анализ сигналов во временной области, основан на применимости принципа суперпозиции к линейным системам.

Пусть плоская монохроматическая волна распространяется вдоль оси Z и проходит через транспарант, помещенный в плоскости $z=0$. Комплексная амплитуда светового поля в данной плоскости после транспаранта имеет вид $E_0(x, y, 0)$. Необходимо определить амплитуду поля в некоторой точке $M(x, y, z)$ за экраном. Для наглядности упростим ситуацию, рассматривая лишь одну поперечную координату. Положим, что световое поле однородно в направлении оси Y , т.е. может быть представлено в виде $E_0(x, 0)$. Очевидно, что аналогично функциям во времени, распределение $E_0(x)$ может быть представлено в виде суммы (конечной или бесконечной) гармонических составляющих с амплитудами и фазами, определяемыми соотношениями Фурье:

$$E_0(x) = \int_{-\infty}^{\infty} F(\omega_x) \exp(-i\omega_x \cdot x) d\omega_x \quad (3.1),$$

$$F(\omega_x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} E_0(x) \exp(i\omega_x \cdot x) dx \quad (3.2).$$

Здесь ω_x указывает на число периодов гармонической функции на единице длины. Гармоническое распределение амплитуды светового поля по пространственной координате x можно представить как результат падения на эту плоскость плоской волны, распространяющейся под некоторым углом к оси Z . Таким образом, заданное распределение $E_0(x)$ можно рассматривать как результат суперпозиции совокупности плоских световых волн, распространяющихся под разными углами к оси Z . Характеризуя плоские

волны волновыми векторами, а их углы относительно оси Z – проекциями волновых векторов на ось X , соотношения (3.1, 3.2) представим в виде:

$$E_0(x) = \int_{-\infty}^{\infty} F_0(k_x) \exp(-ik_x \cdot x) dk_x \quad (3.3),$$

$$F_0(k_x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} E_0(x) \exp(ik_x \cdot x) dx \quad (3.4).$$

Функция $F_0(k_x)$ называется угловым спектром распределения $E_0(x)$ и представляет собой не что иное, как указанную совокупность плоских световых волн. Действительно, выражение $F_0(k_x) \exp(-ik_x \cdot x)$ определяет комплексную амплитуду плоской световой волны в плоскости $z=0$ с компонентами волнового вектора k_x и $k_z = \sqrt{k^2 - k_x^2}$, и с амплитудой $F_0(k_x)$, зависящей от направления распространения. Временной множитель $\exp(i\omega t)$ здесь опущен. Отметим, что $k_x = k \cdot \sin \alpha$, $k_z = k \cdot \cos \alpha$, где α – угол между волновым вектором \vec{k} и осью Z .

Итак, выражение (3.3) представляет собой разложение комплексной амплитуды поля $E_0(x, 0)$ в угловой спектр плоских волн.

Если бы угловой спектр светового поля был известен в любой плоскости, параллельной плоскости транспаранта, т.е. при любом z , то распределение поля $E(x)$ в этой плоскости могло быть определено соотношением:

$$E(x, z) = \int_{-\infty}^{\infty} F(k_x, z) \exp(-ik_x \cdot x) dk_x \quad (3.5).$$

В случае однородной среды световое поле $E(x, y, z)$ удовлетворяет уравнению Гельмгольца $\nabla^2 \bar{E} + k^2 \bar{E} = 0$. В силу линейности ситуации каждая плоская волна углового спектра и весь угловой спектр также удовлетворяют этому уравнению. Подставляя (3.5) в уравнение Гельмгольца, легко получить следующее уравнение:

$$\frac{\partial^2 F(k_x, z)}{\partial x^2} + (k^2 - k_x^2) \cdot F(k_x, z) = 0 \quad (3.6).$$

Действительно, для одной из компонент углового спектра имеем:

$$\begin{aligned} & \nabla^2 [F(k_x, z) \cdot \exp(-ik_x x)] + k^2 \cdot F(k_x, z) \cdot \exp(-ik_x x) = \\ & = \frac{\partial^2}{\partial x^2} [F(k_x, z) \cdot \exp(-ik_x x)] + \frac{\partial^2}{\partial z^2} [F(k_x, z) \cdot \exp(-ik_x x)] + \\ & + k^2 F(k_x, z) \cdot \exp(-ik_x x) = -k_x^2 F(k_x, z) \cdot \exp(-ik_x x) + \\ & + \frac{\partial^2 F(k_x, z)}{\partial z^2} \cdot \exp(-ik_x x) + k^2 F(k_x, z) \cdot \exp(-ik_x x) = \\ & = \frac{\partial^2 F(k_x, z)}{\partial z^2} \cdot \exp(-ik_x x) + (k^2 - k_x^2) F(k_x, z) \cdot \exp(-ik_x x) \end{aligned} \quad (3.7).$$

Очевидно, решение уравнения (3.7) имеет вид:

$$F(k_x, z) = A \cdot \exp[i(k^2 - k_x^2)^{1/2} z] + B \cdot \exp[-i(k^2 - k_x^2)^{1/2} z] \quad (3.8).$$

Используем начальное условие $F(k_x, 0) = F_0(k_x)$ и учтем, что нас интересуют составляющие, распространяющиеся в направлении $+z$. В этом случае окончательно имеем:

$$F(k_x, z) = F(k_x, 0) \cdot \exp[-i(k^2 - k_x^2)^{1/2} z] \quad (3.9).$$

Видим, что по мере удаления от плоскости $z=0$ угловой спектр изменяется. Это связано с изменением фазовых сдвигов между его отдельными компонентами, распространяющимися под разными углами к оси Z . Скорость изменения углового спектра зависит от области значений k_x , играющей в этом спектре существенную роль, т.е. от его ширины.

Все полученные результаты легко обобщить на случай светового поля, зависящего от обеих поперечных координат. При этом соотношения (3.3, 3.4, 3.7) принимают форму:

$$E_0(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_0(k_x, k_y) \exp[-i(k_x \cdot x + k_y \cdot y)] dk_x dk_y \quad (3.10)$$

$$F_0(k_x, k_y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E_0(x, y) \exp[i(k_x \cdot x + k_y \cdot y)] dx dy \quad (3.11)$$

$$F(k_x, k_y, z) = F(k_x, k_y, 0) \cdot \exp[-i(k^2 - k_x^2 - k_y^2)^{1/2} z] \quad (3.12).$$

Таким образом, сформулированная ранее задача определения светового возмущения в точке $M(x, z)$, при заданном распределении поля $E(x, 0)$ в плоскости $z=0$, сводится к последовательности следующих операций:

- А) с помощью соотношения (3.11) находим угловой спектр заданного светового поля в плоскости $z=0$;
- Б) используя (3.12), находим угловой спектр в плоскости $z=z_0$;
- В) по найденному угловому спектру $F(k_x, z_0)$ из (3.10) находим распределение светового поля в плоскости z , и, соответственно, в точке $M(x, z)$.

Понятие углового спектра дает столь же наглядную интерпретацию причины дифракции ограниченных пучков, как и принцип Гюйгенса – Френеля. Действительно, амплитуды плоских волн, составляющих угловой спектр, не меняются в пространстве. Поэтому для бездифракционного распространения светового поля необходимо лишь сохранение фазовых соотношений между компонентами углового спектра. Однако эти соотношения не могут сохраняться в случае оптически однородной среды, т.к. изменение фазы для каждой компоненты определяется множителем $\exp[-i(k^2 - k_x^2)^{1/2} z]$, зависящим от величины k_x .

Связь углового спектра за экраном с его исходным видом и видом отверстия

Если световое поле $E_n(x)$ проходит через экран (транспарант) с комплексной функцией пропускания $g(x)$, то поле за экраном имеет вид:

$$E_0(x,0) = E_n(x) \cdot g(x) \quad (3.13).$$

Угловой спектр поля $E_0(x,0)$ представляет собой в этом случае свертку угловых спектров падающего светового поля $F_i(k_x)$ и транспаранта $\Phi(k_x)$:

$$F_0(k_x) = F_i(k_x) \otimes \Phi(k_x) = \int_{-\infty}^{\infty} F(\xi) \cdot \Phi(k_x - \xi) \cdot d\xi \quad (3.14).$$

В качестве примера рассмотрим случай нормального падения плоской световой волны с единичной амплитудой на непрозрачный экран с отверстием шириной a в направлении оси x . Функция пропускания экрана при этом определяется соотношением:

$$g(x) = \begin{cases} 1 \rightarrow \text{при } |x| < a \\ 0 \rightarrow \text{при } |x| > a \end{cases}.$$

Угловой спектр падающей волны представляет собой дельта – функцию:

$F_i(k_x) = \delta(k_x)$, а угловой спектр поля за экраном имеет вид:

$$F_0(k_x) = \delta(k_x) \otimes \Phi(k_x) = \Phi(k_x) = \frac{1}{2\pi} \int_{-a/2}^{a/2} \exp(ik_x x) = \frac{\sin(k_x a / 2)}{\pi k_x} \quad (3.15).$$

Для оценки ширины углового спектра используют условие равенства нулю функции $\Phi(k_x)$. При этом из соотношения $\sin(k_x a / 2) = 0$ получим $k_x = 2\pi/a$ или $\sin\theta = \lambda/a$, где θ - угол отклонения волнового вектора \vec{k} от оси z . Таким образом, ширина углового спектра оценивается по величине отношения длины волны света к размеру отверстия.

3.2. Параболическое уравнение

Распространение световых волн в пространстве описывается волновым уравнением: $\nabla^2 \bar{E} + k^2 \bar{E} = 0$, где $\bar{E} = \bar{E}(x, y, z, t)$. Рассматривая световые пучки, принимаем во внимание, что их поле сконцентрировано вблизи оси пучка и быстро спадает до нуля в поперечных направлениях. Поэтому для них нет необходимости в рассмотрении точек пространства, далеко отстоящих от оси пучка. В то же время, вследствие эффектов дифракции, амплитуда поля в пучке, в отличие от случая плоских волн, изменяется в продольном направлении. Для учета эффектов дифракции ищем поле пучка (в скалярной форме) в виде:

$$E(x, y, z, t) = A(x, y, z) \cdot \exp[i(\omega t - kz)] \quad (3.16),$$

где $A(x, y, z)$ – медленно уменьшающаяся с ростом z комплексная функция.

Производные поля E по пространственным координатам имеют вид:

$$\frac{\partial^2 E}{\partial x^2} = \frac{\partial^2 A}{\partial x^2} \cdot \exp[i(\omega t - kz)] , \quad \frac{\partial^2 E}{\partial y^2} = \frac{\partial^2 A}{\partial y^2} \cdot \exp[i(\omega t - kz)] ,$$

$$\frac{\partial^2 E}{\partial z^2} = \frac{\partial^2 A}{\partial z^2} \cdot \exp[i(\omega t - kz)] - 2ik \cdot \frac{\partial A}{\partial z} \cdot \exp[i(\omega t - kz)] - k^2 A \cdot \exp[i(\omega t - kz)] .$$

Подставляя эти соотношения в волновое уравнение, получим:

$$\frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2} + \frac{\partial^2 A}{\partial z^2} - 2ik \cdot \frac{\partial A}{\partial z} = 0 .$$

Как правило, поперечные размеры световых пучков значительно больше длины волны света, поэтому амплитуда пучка незначительно изменяется на расстоянии в длину волны. Это позволяет пренебречь слагаемым $\frac{\partial^2 A}{\partial z^2}$ по сравнению с $\frac{\partial A}{\partial z}$ и последнее уравнение принимает вид:

$$\frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2} - 2ik \cdot \frac{\partial A}{\partial z} = 0 \quad (3.17).$$

Это и есть параболическое уравнение теории дифракции.

3.3. Круговой гауссов пучок

Большинство лазерных источников излучает свет в виде слабо расходящихся световых пучков с поперечным распределением интенсивности, описываемым функцией Гаусса $\exp(-r^2/h^2)$, где r – расстояние от центра пучка, а h – полуширина пучка. Такие световые пучки, называемые гауссовыми, имеют целый ряд интересных свойств. Так, поперечный профиль интенсивности в пучке остается гауссовым в любом сечении, а его угловой спектр также имеет гауссово распределение. Используя стандартный подход [3, 4], получим выражения для поля кругового гауссова пучка, как решения параболического уравнения:

$$\frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2} - 2ik \cdot \frac{\partial A}{\partial z} = 0 .$$

Решение ищут в форме:

$$A = A_0 \exp\left\{-i\left[P + \frac{k}{2q}(x^2 + y^2)\right]\right\} \quad (3.18),$$

где P и q – комплексные функции продольной координаты z . Параметр $P(z)$ определяет комплексное изменение фазы, а $q(z)$ характеризует распределения амплитуды в поперечных плоскостях и кривизну волнового фронта в пучке. Дифференцируя (3.18) по пространственным координатам, получим:

$$\frac{\partial^2 A}{\partial x^2} = A \cdot \left[-i \frac{k}{q} - \frac{k^2}{q^2} \cdot x^2 \right] \quad \frac{\partial^2 A}{\partial y^2} = A \cdot \left[-i \frac{k}{q} - \frac{k^2}{q^2} \cdot y^2 \right]$$

$$\frac{\partial A}{\partial z} = A \cdot \left\{ -i \left[\frac{\partial P}{\partial z} - \frac{k}{2q^2} \cdot \frac{\partial q}{\partial z} \cdot (x^2 + y^2) \right] \right\}$$

Подставим эти соотношения в (3.17) и получим:

$$\left\{ (x^2 + y^2) \cdot \left(-\frac{k^2}{q^2} + \frac{k^2}{q^2} \cdot \frac{\partial q}{\partial z} \right) - k \left(\frac{2i}{q} + 2 \frac{\partial P}{\partial z} \right) \right\} = 0 \quad (3.19).$$

Условие (3.19) должно выполняться при любых x, y, z . Но это возможно лишь при одновременном выполнении условий:

$$\frac{\partial q}{\partial z} - 1 = 0 \quad \text{и} \quad \frac{i}{q} + \frac{\partial P}{\partial z} = 0. \quad \text{Отсюда находим:}$$

$$q = z + z_0 \quad (3.20),$$

тогда $\frac{\partial P}{\partial z} = -\frac{i}{z + z_0}$. Наконец,

$$P(z) = -i \cdot [\ln(z + z_0) + \ln C] \quad (3.21).$$

Подставляя (3.20) и (3.21) в (3.18), запишем поле пучка в виде:

$$A = A_0 \exp \left\{ -i \left[-i \cdot \ln(z + z_0) + \frac{k}{2(z + z_0)} \cdot (x^2 + y^2) \right] \right\} \quad (3.22).$$

Постоянная интегрирования z_0 может быть найдена из граничных условий. Удобно полагать, что при $z=0$ фазовый фронт пучка является плоским, а распределение амплитуды имеет вид:

$$\exp \left(-\frac{x^2 + y^2}{w_0^2} \right), \quad \text{где } w_0 \text{ – размер перетяжки пучка.}$$

$$\text{Тогда } \exp \left(-ik \frac{x^2 + y^2}{2(z + z_0)} \right) = \exp \left(-\frac{x^2 + y^2}{w_0^2} \right)$$

Или $-\frac{ik}{2z_0} = -\frac{1}{w_0^2}$, откуда

$$z_0 = \frac{ikw_0^2}{2} \quad (3.23).$$

Для удобства будем записывать выражение для q в форме $q = z + iz_0$.

Предположение плоского фазового фронта означает, что $P(0)=0$:

$$P(0) = -i \cdot \left[\ln \left(\frac{ikw_0^2}{2} \right) + \ln(c) \right] = 0, \quad \text{т.е. } \ln(c) = -\ln \left(\frac{ikw_0^2}{2} \right). \quad \text{Тогда имеем:}$$

$$iP(z) = [\ln(iz_0 + z) - \ln(iz_0)] = \ln \left(1 - i \frac{z}{z_0} \right) \quad (3.24).$$

Соответственно,

$$\exp[-iP(z)] = \exp[-\ln\left(1 - i\frac{z}{z_0}\right)] = \frac{1}{1 - i\frac{z}{z_0}} = \left[1 + \frac{z^2}{z_0^2}\right]^{1/2} \cdot \exp(i\Phi) \quad (3.25),$$

где $\Phi = \arctg\left(-\frac{z}{z_0}\right)$. Далее, из соотношения для $q(z)$ найдем $1/q$:

$$\frac{1}{q} = \frac{1}{z + iz_0} = \frac{1}{z\left(1 + \frac{z_0^2}{z^2}\right)} - i\frac{2}{kw_0^2\left(1 + \frac{z^2}{z_0^2}\right)} \quad (3.26).$$

Введем обозначения:

$$R(z) = z\left(1 + \frac{z_0^2}{z^2}\right) \quad (3.27),$$

$$w^2(z) = w_0^2\left(1 + \frac{z^2}{z_0^2}\right) \quad (3.28).$$

С учетом этого $\frac{1}{q} = \frac{1}{R} - i\frac{\lambda}{\pi w^2(z)}$. Наконец, получим выражение для поля гауссова пучка в форме:

$$\begin{aligned} A &= A_0 \exp\left\{-i\left[P + \frac{k}{2q}(x^2 + y^2)\right]\right\} = \\ &= A_0 \cdot \left[1 + \frac{z^2}{z_0^2}\right]^{-1/2} \exp\left(-\frac{x^2 + y^2}{w^2}\right) \cdot \exp\left[i\Phi - i\left(\frac{k}{2R}(x^2 + y^2)\right)\right] \end{aligned} \quad (3.29).$$

Если заданы длина волны света λ и ширина пучка w_0 , то его параметры оказываются полностью определенными. Добавка $i\Phi$ в (3.29) описывает отличие фазы пучка от фазы плоской волны. Амплитудный множитель перед экспонентой в (3.29) описывает изменение амплитуды пучка вследствие эффекта дифракции. Поле пучка в полной форме включает еще множитель $\exp(-ikz)$. Тогда волновой фронт пучка определяется соотношением:

$$kz + \frac{k}{2R}(x^2 + y^2) - \Phi = kz_I \quad (3.30).$$

Пренебрегая величиной Φ , которая, как правило, невелика, получим соотношение:

$$\frac{k}{2R}(x^2 + y^2) = k \cdot (z_I - z) \text{ или}$$

$$\frac{(x^2 + y^2)}{2R} = z_I - z \quad (3.31).$$

Последнее соотношение представляет собой уравнение поверхности вращения. Параметр R является радиусом кривизны этой поверхности для точек на оси пучка.

Пространственная форма пучка определяется изменением его амплитуды. Она может быть найдена из условия:

$$\frac{x^2 + y^2}{w^2} = 1 \text{ или:}$$

$$x^2 + y^2 = w_0^2 \cdot \left(1 + \frac{z^2}{z_0^2} \right) \quad (3.32).$$

Выражение (3.32) является уравнением однополостного гиперболоида вращения. Угол дифракционной расходимости пучка – это угол между асимптотами гиперболоида и осью пучка:

$$\theta = \frac{\lambda}{\pi w_0} \quad (3.33).$$

3.4. Угловой спектр гауссова пучка

Пусть распределение амплитуды светового поля в плоскости $z=0$ задано соотношением:

$E(x) = E_0 \cdot \exp\left(-\frac{x^2}{w^2}\right)$, что соответствует полю в перетяжке одномерного гауссова пучка. Найдем угловой спектр такого пучка [5]:

$$\begin{aligned} F(k_x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} E(x) \exp(-ik_x x) dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{w_0^2}\right) \exp(-ik_x x) dx = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left[-\frac{x^2}{w_0^2} - ik_x x - \left(\frac{ik_x w_0}{2}\right)^2 + \left(\frac{ik_x w_0}{2}\right)^2\right] dx = \\ &= \frac{1}{2\pi} \cdot \exp\left(-\frac{k_x^2 w_0^2}{4}\right) \cdot \int_{-\infty}^{\infty} \exp\left[-\left(\frac{x}{w_0} + \frac{ik_x w_0}{2}\right)^2\right] dx = \frac{w_0}{2\sqrt{\pi}} \cdot \exp\left[-k_x^2 \cdot \frac{w_0^2}{4}\right] \quad (3.34). \end{aligned}$$

Таким образом, видим, что угловой спектр гауссова пучка также имеет гауссову форму.

3.5. Гауссовы пучки высших порядков

Выше рассматривалось лишь одно аналитическое решение уравнения (3.17), а именно фундаментальный гауссов пучок, который называют также

основной модой свободного пространства. Но кроме этого существуют и другие решения параболического уравнения (3.17), определяющие поля пучков с неизменной формой распределения амплитуды по поперечному сечению. Данные решения называют высшими модами свободного пространства. Все подобные решения (3.17) образуют полную ортогональную систему функций, поэтому любое произвольное монохроматическое световое поле может быть представлено в виде суперпозиции мод свободного пространства.

В декартовой системе координат решение уравнения (3.17) может быть записано в виде:

$$A = A_0 \cdot g\left(\frac{x}{w}\right) \cdot h\left(\frac{y}{w}\right) \cdot \exp\left\{i\left[P + \frac{k}{2q}(x^2 + y^2)\right]\right\} \quad (3.35),$$

где g – функция пространственных координат x и z , а h – функция y и z . Для действительных функций g и h это решение описывает моды, поперечное распределение поля которых определяется радиусом гауссова пучка $w(z)$. Подставляя (3.35) в (3.17), можно выяснить, что функции g и h удовлетворяют тому же дифференциальному уравнению, что и полиномы Эрмита $H_n(t)$

$$\frac{d^2 H_n}{dt^2} - 2t \cdot \frac{dH_n}{dt} + 2nH_n = 0 \quad (3.36),$$

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

$$g \cdot h = H_m\left(\sqrt{2} \frac{x}{w}\right) \cdot H_n\left(\sqrt{2} \frac{y}{w}\right) \quad (3.37),$$

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

$$H_0(t) = 1 \quad (3.38),$$

$$H_1(t) = 2t \quad (3.39),$$

$$H_2(t) = 4t^2 - 2 \quad (3.40),$$

$$H_3(t) = 8t^3 - 12t \quad (3.41).$$

Для математического описания мод более высоких порядков можно использовать выражение (3.37), если в правую его часть вставить произведение $g \cdot h$. Распределение поля в модах свободного пространства будет определяться, таким образом, произведением функций Эрмита и

Гаусса

$$A(x, y, z) = \frac{w_0}{w} H_m\left(\sqrt{2} \frac{x}{w}\right) H_n\left(\sqrt{2} \frac{y}{w}\right) \cdot \exp\left\{i(kz - \Phi) - (x^2 + y^2) \left[\frac{1}{w^2} - \frac{ik}{2R}\right]\right\} \quad (3.42).$$

В частном случае $m=0$, $n=0$ мы имеем фундаментальный гауссов пучок - основную моду свободного пространства. Параметр $R(z)$ в (3.42) для всех мод одинаков. Это означает, что кривизна волнового фронта одинакова для всех мод и закон его изменения один и тот же. Однако фазовый сдвиг Φ зависит от поперечного индекса. Можно найти, что

$$\Phi(m, n, z) = (m + n + 1) \operatorname{arctg}\left(\frac{\lambda z}{\pi w_0^2}\right) \quad (3.43).$$

Из выражения (3.42) видно, что фазовая скорость с ростом поперечного индекса увеличивается.

3.6. Элементы Фурье - оптики

Появление лазеров явилось существенным стимулом для развития методов когерентной оптики и голографии. Подход к световому полю, как к носителю информации, породил целую область, стоящую на стыке радиотехники и оптики. Ее называют радиооптикой, а также оптической обработкой информации [9, 10]. Основу методов и устройств оптической обработки информации составляют явления интерференции и дифракции, в которых наиболее явно проявляются волновые свойства света.

Основная идея методов оптической обработки информации заключается в переносе информации, подлежащей обработке, на оптический транспарант, в виде функции его пропускания от пространственных координат или пространственной зависимости его показателя преломления. Затем данный транспарант зондируется когерентным или некогерентным пучком света. Анализ светового поля, полученного в результате освещения транспаранта, и преобразованного оптической системой, дает информацию о параметрах исходного сигнала.

К основным достоинствам оптических методов обработки информации можно отнести следующее.

А) Большую информационную емкость. Действительно, световое поле представляет собой функцию, как минимум, двух пространственных координат. А электрический сигнал зависит лишь от одной переменной – времени. Поэтому объем информации при передаче некоторого оптического сигнала в виде изображения за некоторое время Δt значительно больше ее объема, передаваемого за то же время с помощью электрического сигнала. Так, телевизионный кадр выводится на экран электронно – лучевой трубки за

40 мс. В оптической системе за это же время можно передать огромное число подобных кадров, если изображение быстро изменяется.

Б) Многоканальность (параллельность передачи и обработки информации). Поскольку световое поле зависит от нескольких пространственных координат, то одна из них может рассматриваться как независимая переменная, а другая – как параметр, определяющий номер канала. В этом случае обработка информации может вестись параллельно по многим каналам.

В) Высокое быстродействие.

Потенциальное быстродействие оптических систем определяется скоростью света. Например, операция двумерного преобразования Фурье осуществляется сферической линзой за время распространения света от оптического транспаранта до задней фокальной плоскости линзы. При фокусном расстоянии линзы $F=30$ см оценка этого времени дает $\tau \approx 10^{-9}$ с. Правда, это лишь потенциальное быстродействие. Реально оно ограничивается скоростью ввода информации в систему и скоростью ее вывода. Во многих случаях оно определяется быстродействием электронного “обрамления” системы.

3.6.1. Преобразование Фурье в когерентной оптической системе

Рассмотрим простую оптическую систему, состоящую из одиночной тонкой сферической линзы (Рис. 3.1). Считаем, что эта система не имеет aberrаций и в ней нет эффектов поглощения, рассеяния и отражения света.

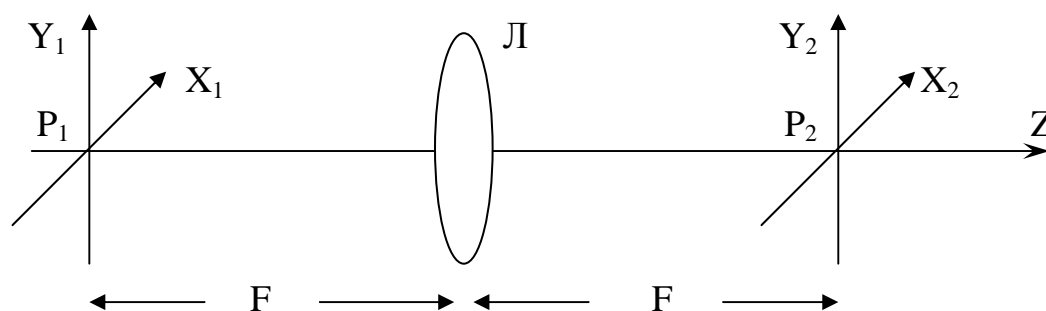


Рис. 3.1. Простейшая оптическая система.

В передней фокальной плоскости P_1 (при $z=0$) линзы с фокусным расстоянием F расположен носитель обрабатываемой информации – транспарант. Примером амплитудного транспаранта с пространственной модуляцией функции пропускания является фотопленка или фотопластинка. Типичный пример фазового транспаранта – пространственный электрооптический модулятор света или акустооптический модулятор. В

общем случае функция пропускания транспаранта может быть и комплексной:

$$T(x_1, y_1) = S(x_1, y_1) \cdot \exp[i\phi(x_1, y_1)] \quad (3.44).$$

Плоскость P_2 ($z=2F$) является выходной плоскостью системы. Это задняя фокальная плоскость линзы.

Пусть в направлении z слева от транспаранта в системе распространяется плоская однородная монохроматическая световая волна с частотой ω и волновым числом k :

$$\bar{E} = \bar{E}_0 \cdot \exp[i(\omega t - kz)] \quad (3.45).$$

Тогда на выходе транспаранта поле этой волны принимает вид:

$$\bar{E}_1(x_1, y_1, t) = \bar{E} \cdot T(x_1, y_1) = \bar{E}_0 \cdot S(x_1, y_1) \cdot \exp[i(\omega t + \phi(x_1, y_1))] \quad (3.46).$$

Видим, что это соотношение описывает операцию умножения постоянной величины \bar{E}_0 на функцию $T(x_1, y_1)$. Если непосредственно за первым поместить второй транспарант с функцией пропускания $Q(x_1, y_1)$, то после него получим световое поле вида:

$$\bar{E}_1(x_1, y_1) = \bar{E} \cdot T(x_1, y_1) \cdot Q(x_1, y_1) \quad (3.47).$$

Этот пример показывает, как легко реализуется в оптической системе операция перемножения двух или произвольного числа двумерных функций, представленных функциями пропускания транспарантов.

Найдем распределение комплексной амплитуды поля на выходе оптической системы. Для этого воспользуемся принципом Гюйгенса – Френеля, согласно которому каждая точка волнового фронта светового поля является источником вторичной сферической волны. Найдем сумму вкладов всех подобных источников вторичных волн в плоскости P_1 в напряженность светового поля в точке с координатами x_2, y_2 плоскости P_2 . Для небольших углов между направлением лучей и оптической осью системы получим:

$$E(x_2, y_2) = A \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\exp(ikr_{12})}{r_{12}} \cdot E_1(x_1, y_1) dx_1 dy_1 \quad (3.48),$$

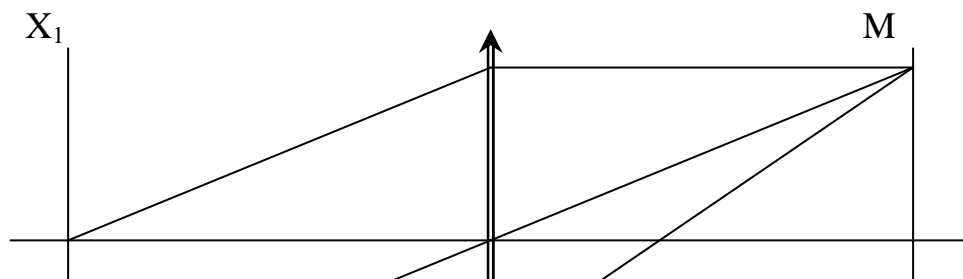
где A – коэффициент пропорциональности, а r_{12} – расстояние между точкой $2(x_2, y_2)$ и текущими точками плоскости P_1 :

$$r_{12} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + z^2} \quad (3.49).$$

Если $r_{12} \gg x_2 - x_1, y_2 - y_1$, то величину r_{12} в знаменателе (5) можно вынести из – под знака интеграла. В этом случае:

$$E(x_2, y_2) = \frac{A}{r_{12}} \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(ikr_{12}) \cdot E_1(x_1, y_1) dx_1 dy_1 \quad (3.50).$$

Выразим r_{12} через x_1, x_2, y_1, y_2 . Для простоты рассмотрим одномерный случай (Рис. 3.2).



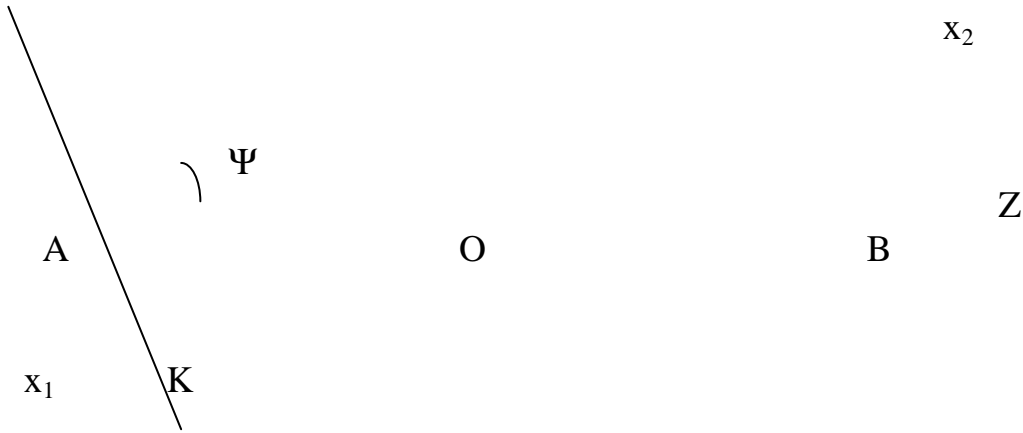


Рис. 3. 2.

Поскольку плоскость P_2 совпадает с задней фокальной плоскостью линзы L , то в рассматриваемую точку $M(x_2)$ будут собираться только параллельные лучи, распространяющиеся в области I в некотором направлении, задаваемом углом Ψ . Построим плоскость V , проходящую через начало координат системы и перпендикулярную направлению Ψ . Из курса оптики известно, что оптическая длина пути между точкой M с координатой x_2 и любой точкой на плоскости V является постоянной величиной. Обозначим эту величину как ρ . Найдем ее как сумму двух отрезков KO и OM :

$$\rho = KO + OM = \sqrt{F^2 - x_1^2 \cos^2 \Psi} + \sqrt{F^2 + x_2^2} \quad (3.51).$$

Если $x_1, x_2 \ll F$, то:

$$\rho = F \left(1 - \frac{1}{2} \frac{x_1^2 \cos^2 \Psi}{F^2} + \dots \right) + F \left(1 + \frac{1}{2} \frac{x_2^2}{F^2} + \dots \right) \quad (3.52).$$

Поскольку треугольники Ax_1O и OBM равны, то $x_1 = x_2$, а если еще $\Psi \ll 1$, то $\rho \approx 2F$ и величина ρ не зависит ни от x_1 , ни от x_2 . Расстояние r_{12} отличается от ρ на величину отрезка x_1K , которая зависит от положения точки x_1 на плоскости P_1 :

$$x_1K = -x_1 \cdot \sin(\Psi) \quad (3.53).$$

Здесь знак (-) учитывает направление оси X_1 . Учитывая, что $\sin(\Psi) = \frac{x_2}{F}$, получим:

$$r_{12} = \rho + x_1K = 2F - \frac{x_1x_2}{F} \quad (3.54).$$

Тогда:

$$E_2 = \frac{A}{r_{12}} \exp(ik \cdot 2F) \cdot \int_{-\infty}^{\infty} \exp(-ik \frac{x_1x_2}{F}) E_1(x_1) dx_1 \quad (3.55).$$

Введем обозначения:

$$B = \frac{A}{r_{12}} \exp(i \frac{4\pi}{\lambda} F) ; \quad \omega_{x2} = \frac{2\pi}{\lambda F} x_2.$$

Тогда для E_2 получим:

$$E_2(x_2) = B \cdot \int_{-\infty}^{\infty} E_1(x_1) \cdot \exp(-i\omega_{x2}x_1) dx_1 \quad (3.56).$$

Теперь вспомним выражения для прямого и обратного преобразований Фурье для радиосигнала $f(t)$ и его спектра $F(\omega)$:

$$F(\omega) = \int_{-\infty}^{\infty} f(t) \cdot \exp(-i\omega t) dt \quad (3.57),$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) \cdot \exp(i\omega t) d\omega \quad (3.58).$$

Сравнивая выражения для $E_2(x_2)$, полученные выше, и для $F(\omega)$, легко видеть, что они полностью аналогичны по форме. Таким образом, $E_2(x_2)$ является прямым преобразованием Фурье от распределения $E_1(x_1)$. При этом роль времени играет координата x_1 , а роль временной частоты ω - величина ω_{x2} , которую называют пространственной частотой, поскольку она является функцией координаты плоскости наблюдения x_2 .

Как отмечено, переход к одномерному случаю имел целью лишь упрощение рассмотрения. Можно просто записать аналогичное выражение и для двумерного случая:

$$E_2(x_2, y_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp[-i(\omega_{x2}x_1 + \omega_{y2}y_1)] \cdot E_1(x_1, y_1) dx_1 dy_1 \quad (3.59).$$

Рассмотрим физический смысл последнего выражения. При падении плоской однородной световой волны на транспарант происходит ее дифракция на транспаранте. Дифрагированное поле представляет собой суперпозицию плоских световых волн, распространяющихся в пространстве во всех возможных направлениях. Амплитуда каждой из этих волн определяется видом функции $E_1(x_1, y_1)$. Таким образом, уже на этапе дифракции светового поля на транспаранте возникают элементы преобразования Фурье. Роль временных частот здесь играют углы или направления распространения плоских волн. Поэтому говорят, что на данном этапе происходит разложение светового поля в угловой спектр плоских волн. Поскольку фокусирующая линза собирает все лучи, распространяющиеся в одном направлении ($x_2/F, y_2/F$) в одну точку (x_2, y_2) в ее задней фокальной плоскости, то она ставит в соответствие каждой плоской волне точку на плоскости P_2 . Теперь роль частоты спектра Фурье будет играть не направление распространения плоской волны, а координаты в фокальной плоскости P_2 . Комплексная амплитуда поля в этой точке соответствует амплитуде спектральной составляющей Фурье.

Таким образом, в когерентной оптической системе распределения напряженности светового поля в фокальных плоскостях линзы связаны двумерным преобразованием Фурье. Функция $E_2(x_2, y_2) = S(\omega_{x2}, \omega_{y2})$ – это пространственный спектр сигнала $E_1(x_1, y_1)$. В связи с этим плоскость P_1 называют сигнальной плоскостью или плоскостью изображения, а плоскость P_2 – спектральной плоскостью.

3.6.2. Некоторые математические операции, реализуемые в оптической системе

Рассмотрим несколько примеров, иллюстрирующих достоинства оптических методов при выполнении аналоговых вычислений.

Вычисление постоянной составляющей в световом поле

Пусть световое поле в сигнальной плоскости имеет вид $E_1(x) = E_0$. Распределение поля в спектральной плоскости определяется из (3.56):

$E_2(x_2) = E_0 \int_{-\infty}^{\infty} \exp(-i\omega_{x2}x_1) dx_1$. В этом случае для любого значения $\omega_{x2} \neq 0$ величина указанного интеграла будет равна нулю. Только в случае $\omega_{x2} = 0$, т.е. при $x_2 = 0$, получим $E_2(0) = E_0 \int_{-\infty}^{\infty} dx_1 = \infty$. Таким образом, для

вычисления постоянной составляющей светового поля необходимо поместить точечный фотоприемник в плоскости P_2 в точку с координатой $x_2=0$. Реально апертура линзы определяет размер фокальной точки и ширину спектра пространственных частот.

Обратная ситуация, когда **в сигнальной плоскости – точечный источник**.

Условия имеют вид:

$\bar{E}_1(0) = E_0$; $\bar{E}_1(x) = E_0$ при $x_1 \neq 0$. Тогда поле в спектральной плоскости будет равно: $E_2(x_2) = \int_{-\infty}^{\infty} E_1(x_1) \cdot \exp(-i\omega_{x2}x_1) dx_1 = E_0 \cdot x_1 = 0$

Выполнение операции интегрирования.

Это простейшее интегральное преобразование, т.е. вычисление определенного интеграла от распределения поля в плоскости изображения. Полагая $\omega_{x2} = 0$, т.е. рассматривая точку, соответствующую фокусу линзы в спектральной плоскости, получим:

$$E_2(0,0) = \int_{-\infty}^{\infty} E_1(x_1) dx_1.$$

Таким образом, для реализации операции интегрирования необходимо в задней фокальной плоскости линзы, в точке фокуса, поместить точечный фотоприемник. Величина отклика фотоприемника будет численно равна определенному интегралу (в бесконечных пределах) от распределения поля в плоскости изображения.

3.6.3. Обратное преобразование Фурье оптического сигнала

В когерентной оптической системе достаточно просто может быть выполнена и операция обратного преобразования Фурье. Такая возможность очевидна, поскольку, как отмечено, световые поля в передней и задней плоскостях сферической линзы связаны преобразованием Фурье. Т.е. если в плоскости изображения само изображение является пространственным спектром некоторого светового поля, то в спектральной плоскости мы должны получить копию этого светового поля. Этот факт имеет строгое математическое подтверждение. На рис. 3.3 представлена схема установки, которая позволяет выполнить двукратное двумерное преобразование Фурье, т.е. сначала получить пространственный спектр некоторого светового поля, затем по спектру восстановить само поле.

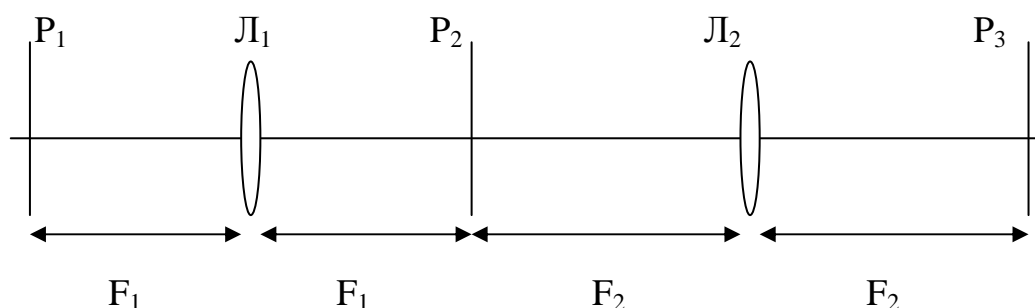


Рис. 3.3. Схема оптической системы, выполняющей двукратное преобразование Фурье.

Следует отметить, что такая система формирует перевернутое изображение, поэтому при равенстве фокусных расстояний обеих линз световое поле в некоторой точке плоскости P_3 отвечает соотношению:

$$E_3(x_3) = E_1(-x_3).$$

Покажем, что последнее соотношение действительно справедливо. Для этого найдем поле в некоторой точке в плоскости P_3 :

$$E_3(x_3) = \int_{-\infty}^{\infty} E_2(x_2) \exp\left[-i\left(\frac{2\pi}{\lambda F} x_3 \cdot x_2\right)\right] dx_2 \quad (3.60).$$

Но для E_2 имеем:

$$E_2(x_2) = \int_{-\infty}^{\infty} E_1(x_1) \exp[-i(\frac{2\pi}{\lambda F} x_2 \cdot x_1)] dx_1.$$

Подставив последнее выражение в (3.60), получим:

$$\begin{aligned} E_3(x_3) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E_1(x_1) \exp[-i(\frac{2\pi}{\lambda F} x_2 \cdot x_1)] \cdot \exp[-i(\frac{2\pi}{\lambda F} x_3 \cdot x_2)] dx_1 dx_2 = \\ &= \int_{-\infty}^{\infty} E_1(x_1) \left\{ \int_{-\infty}^{\infty} \exp[-i(\frac{2\pi}{\lambda F} (x_1 + x_3) x_2)] dx_2 \right\} dx_1 = \\ &= \int_{-\infty}^{\infty} E_1(x_1) \left\{ \lim_{D \rightarrow \infty} \int_{-D/2}^{D/2} \exp[-i(\frac{2\pi}{\lambda F} (x_1 + x_3) x_2)] dx_2 \right\} dx_1 = \\ &= \int_{-\infty}^{\infty} E_1(x_1) \cdot \delta(x_1 + x_3) dx_1 \end{aligned} \quad (3.61).$$

$$\text{Но } \int_{-\infty}^{\infty} E_1(x_1) \cdot \delta(x_1 + x_3) dx_1 = E_1(-x_3).$$

Таким образом, видим, что в данной системе поля в передней фокальной плоскости первой линзы и в задней фокальной области второй линзы связаны соотношением:

$$E_3(x_3) = E_1(-x_3) \quad (3.62).$$

3.6.4. Пространственная фильтрация в когерентной оптической системе

Операция фильтрации требуемых спектральных составляющих в оптической системе может быть выполнена достаточно просто. Для этого в спектральной плоскости достаточно поместить некоторый транспарант – пространственный фильтр. Если функция пропускания транспаранта отвечает соотношению $H(\omega_{x2}, \omega_{y2})$, то распределение поля в этой плоскости после транспаранта будет иметь вид:

$$E_2^\Phi(x_2, y_2) = S(\omega_{x2}, \omega_{y2}) \cdot H(\omega_{x2}, \omega_{y2}) \quad (3.63).$$

Примером простейшего пространственного фильтра является фильтр с прямоугольной амплитудной характеристикой и постоянной фазовой характеристикой. Он представляет собой непрозрачный экран с окнами нужной формы и размеров.

Наиболее часто пространственные фильтры в оптических системах используются для:

- А) подавления постоянной составляющей;
- Б) выделения нужных пространственных частот;
- В) подавления некоторых заданных пространственных частот.

4. РАСПРОСТРАНЕНИЕ ОПТИЧЕСКИХ ВОЛН В МАТЕРИАЛЬНЫХ СРЕДАХ

4.1. Дисперсия показателя преломления диэлектрической среды

Из курса электродинамики известно, что при распространении электромагнитных волн в поглощающей среде их характеристики зависят от частоты (например, фазовая скорость, затухание и длина волны). Такая зависимость называется дисперсией. Дисперсионные эффекты проявляются при распространении электромагнитных, в том числе световых, волн очень часто. Это объясняется тем, что диэлектрическая проницаемость или показатель преломления вещества являются результатом дискретности его структуры [2, 5].

Классическая модель механизма фазовой задержки электромагнитной волны в среде.

Пусть имеем цепочку из N атомов, расположенных один за другим. Если в среде существует (или на среду падает) световая волна, то под влиянием высокочастотного электрического поля атом поляризуется. В переменном поле атом превращается в электрический диполь, дипольный момент которого осциллирует с частотой внешнего электромагнитного поля, переизлучая световую волну с той же частотой. Фаза переизлученной волны, как и фаза колебаний диполя, определяется возвращающей силой (силой взаимодействия электрона с ядром атома). Таким образом, между фазой падающей и переизлученной волн возникает некоторая разница. В результате их интерференции фаза результирующего поля оказывается промежуточной. Это повторяется для всех последующих атомов, а результирующая фазовая задержка между фазой световой волны, прошедшей через среду, и фазой волны такой же частоты, прошедшей такое же расстояние в вакууме, трактуется как изменение скорости световой волны при ее распространении в среде. Действительно, добавка ϕ к фазе волны, прошедшей расстояние L в среде, определяется соотношением:

$$\phi = \frac{\omega}{v} L \quad (4.1),$$

где v – скорость волны. При неизменных расстоянии и частоте, очевидно, такие добавки могут различаться лишь в случае различия скоростей распространения света в разных средах.

Таким образом, световая волна в среде распространяется медленнее, чем в вакууме. Величина отношения этих скоростей называется абсолютным показателем преломления среды. Естественно, что каждый из упоминаемых

диполей излучает во всех направлениях. Почему же тогда отсутствует излучение в стороны и назад? Дело в том, что фаза излучения каждого диполя определяется фазой возмущающей волны. Поэтому их излучение когерентно и, интерферируя, складывается в фазе лишь в прямом направлении. В данной простой модели предполагалось, что свойства диполей одинаковы. Если их свойства различаются, переизлученные волны можно наблюдать и в направлениях, отличных от прямого. Это явление называется рассеянием излучения.

Эффекты рассеяния существенно проявляются в материалах, содержащих большое количество примесей. В очень чистых материалах такое рассеяние минимально. До сих пор предполагалось, что вся энергия, затраченная на возбуждение колебаний диполя, переизлучается в виде электромагнитной волны. Реально, даже в простейшем случае, колебания под воздействием внешнего поля совершает не только электрон, но и ядро атома. Кроме того, существуют и силы межатомного взаимодействия. При определенных частотах часть энергии падающей волны не переизлучается в виде электромагнитной волны, а тратится на колебания атомов в решетке, что приводит к нагреванию вещества. Таким образом, часть световой энергии может поглощаться.

Для более строгого (хотя и только качественного) описания данного эффекта будем рассматривать электрон, совершающий колебания малой амплитуды около положения равновесия. Уравнение движения электрона может быть записано в виде [5, 13]:

$$m \frac{d^2 \bar{r}}{dt^2} + 2\gamma m \frac{d\bar{r}}{dt} + m\omega_0^2 \bar{r} = -e\bar{E} \quad (4.2).$$

Здесь \bar{r} – смещение электрона относительно положения равновесия; γ – коэффициент, учитывающий затухание колебаний вследствие наличия “сил трения”; ω_0 – собственная частота колебаний электрона; m – масса электрона; e – его заряд; \bar{E} – напряженность внешнего электрического поля.

В записи (4.1а) в уравнении первое слагаемое – это ньютоновская сила, второе учитывает наличие некоторой “силы трения”, третье – это возвращающая “упругая” сила. В правой части – сила, действующая на электрон со стороны внешнего электромагнитного поля. Ограничимся анализом данного одномерного уравнения. Считаем, что внешнее поле следует гармоническому закону:

$$\bar{E} = \dot{\bar{E}}(\omega) \cdot \exp(-i\omega t) + c.c. \quad (4.3).$$

Здесь комплексная амплитуда поля включает фазу:

$$\dot{\bar{E}}(\omega) = \bar{E}_0 \cdot \exp(i\phi) \quad (4.4).$$

Подставим (4.3) в (4.2), полагая, что $r = r_m \cdot \exp(-i\omega t) + c.c.$ и учтем, что

$$\frac{d\bar{r}}{dt} = -i\omega\bar{r} + c.c.; \quad \frac{d^2\bar{r}}{dt^2} = -\omega^2\bar{r} + c.c.: \quad -\omega^2\bar{r} - 2i\gamma\omega\bar{r} + \omega_0^2\bar{r} = -\frac{e}{m}\bar{E}(\omega)e^{-i\omega t} + c.c..$$

Отсюда находим \bar{r} :

$$\bar{r} = -\frac{e}{m}\bar{E}(\omega)\frac{1}{\omega_0^2 - 2i\gamma\omega - \omega^2} + k.c. \quad (4.5).$$

Дипольный момент атома $\bar{p} = -e \cdot \bar{r}$ и, если в единице объема концентрация электронов равна N , то вектор поляризации \bar{P} , определяемый как дипольный момент единицы объема, равен:

$$\bar{P} = Ne\bar{r} \quad (4.6),$$

или

$$\bar{P} = N\frac{e^2}{m} \cdot \frac{1}{\omega_0^2 - 2i\gamma\omega - \omega^2} \cdot \bar{E}(\omega)e^{-i\omega t} + k.c. \quad (4.7).$$

Введем обозначение

$$\kappa(\omega) = \frac{Ne^2}{\epsilon_0 m} \cdot \frac{1}{\omega^2 - 2i\gamma\omega - \omega_0^2} \quad (4.8).$$

Тогда:

$$P = \epsilon_0 \kappa(\omega) E(\omega) \cdot \exp(-i\omega t) + c.c. \quad (4.9).$$

Из (4.9) видим, что наведенная поляризация действительно пропорциональна амплитуде внешнего электрического поля и имеет ту же частоту. Для выявления характера частотной зависимости параметров световых волн начнем с уравнений Максвелла:

$$\begin{aligned} \text{rot}\bar{H} &= \frac{\partial\bar{D}}{\partial t} + \bar{\delta} \\ \text{rot}\bar{E} &= -\frac{\partial}{\partial t}(\mu\bar{H}), \end{aligned}$$

где

$$\begin{aligned} \bar{D} &= \epsilon_0\bar{E} + \bar{P} = \epsilon_0(1 + \kappa)\bar{E} = \epsilon\bar{E} \\ \bar{P} &= \epsilon_0\kappa\bar{E} \end{aligned}$$

Первое из уравнений перепишем в виде:

$$\text{rot}\bar{H} = \sigma\bar{E} + \frac{\partial}{\partial t}(\epsilon\bar{E})$$

Считая среду непроводящей и немагнитной ($\sigma=0$; $\mu=\mu_0$), получим волновое уравнение обычным способом:

$$\text{rot rot } \bar{E} = -\mu_0 \frac{\partial}{\partial t} \text{rot } \bar{H} = -\mu_0 \frac{\partial}{\partial t} \cdot \frac{\partial\bar{D}}{\partial t}$$

или

$$\text{grad div } \bar{E} - \nabla^2 \bar{E} = -\mu_0 \frac{\partial^2 \bar{D}}{\partial t^2}$$

и, поскольку $\text{div}\bar{E} = 0$, то данное уравнение запишется в форме:

$$\nabla^2 E - \mu_0 \frac{\partial^2 \bar{D}}{\partial t^2} = 0 \quad (4.10).$$

Полагаем, что гармоническая световая волна распространяется вдоль z и что она плоская и однородная ($\partial/\partial x = \partial/\partial y = 0$). Волновое уравнение в таком случае принимает вид:

$$\frac{\partial^2 \bar{E}}{\partial z^2} - \mu_0 \epsilon \frac{\partial^2 \bar{E}}{\partial t^2} = 0 \quad (4.11).$$

Подставляем в (4.11) решение в виде плоской волны, бегущей вдоль оси z : $E(z, t) = E_0 \cdot \exp[i(\omega t - kz)]$, после чего приходим к соотношению:

$$-k^2 E + \mu_0 \epsilon \omega^2 E = 0 \quad (4.12).$$

Далее, учитывая, что

$$k = \frac{\omega}{v} = \frac{\omega}{c} \cdot n, \text{ получим: } \mu_0 \epsilon \omega^2 = \frac{\omega^2}{c^2} n^2 \text{ или } n^2 = \mu_0 \epsilon c^2 = \frac{\mu_0 \epsilon_0 (1 + \kappa)}{\mu_0 \epsilon_0} = (1 + \kappa).$$

Итак, имеем для n^2 соотношение:

$$n^2 = 1 + \frac{Ne^2}{m} \cdot \frac{1}{\omega_0^2 - i\gamma\omega - \omega^2} \quad (4.13).$$

Очевидно, что в случае отсутствия потерь ($\gamma = 0$) показатель преломления является вещественным. Его частотная зависимость представляет собой разрывную функцию (разрыв при $\omega = \omega_0$) (рис. 4.1).

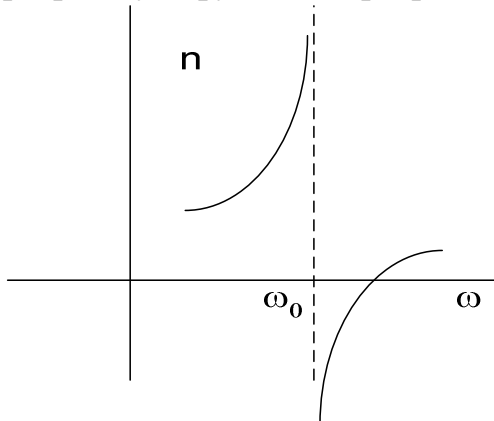


Рис. 4.1.

В случае, когда $\gamma \neq 0$, n становится комплексной величиной. Вид частотных зависимостей действительной и мнимой частей представлен на рис. 4.2. Здесь сплошной линией отмечена зависимость действительной части $n(\omega)$, а пунктиром — зависимость мнимой части. Ясно, что мнимая часть \underline{n} указывает на величину потерь на поглощение. Эти потери максимальны при $\omega = \omega_0$. Области, соответствующие частотам $\omega < \omega_0$, $\omega > \omega_0$ составляют область нормальной дисперсии.

Участок кривой вблизи ω_0 с отрицательным наклоном кривой $n(\omega)$ соответствует области аномальной дисперсии.

Большинство материалов, естественно, имеет несколько частот собственных колебаний ω_{0n} и, следовательно, несколько линий поглощения и областей аномальной дисперсии показателя преломления.

Кривые дисперсии (зависимости $n(\omega)$) всегда имеют характерный вид. В промежутке между двумя линиями поглощения тангенс угла наклона кривой дисперсии всегда положителен.

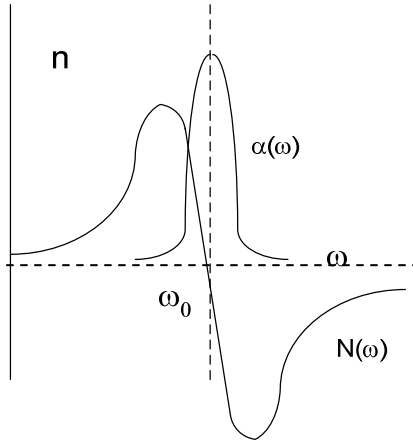


Рис. 4.2.

У высокочастотной границы этого промежутка величина показателя преломления самая высокая. При уменьшении частоты она сначала резко уменьшается, затем скорость спада существенно уменьшается. При достижении следующей полосы поглощения величина n опять резко уменьшается. Внутри полосы поглощения наклон кривой скачком изменяется на обратный. Это соответствует аномальной дисперсии.

Величина n в этой области резко возрастает. На

низкочастотном крае полосы поглощения наклон кривой $n(\omega)$ опять становится положительным. И т.д.

Используя последнее соотношение (4.13), легко видеть, что в пределе низких частот ($\omega \rightarrow 0$) величина $n > 1$:

$$n^2 = 1 + \frac{Ne^2}{m\omega_0^2} \quad (4.14).$$

При очень высоких частотах ($\omega \gg \omega_0$) величина $n \rightarrow 1$. Это справедливо практически для всех веществ. Физическая причина такого поведения – инерционность электрона. При очень высоких частотах электрон (в одноэлектронной модели атома) не успевает за внешним полем, поэтому дипольный момент атома \bar{p} незначителен и при $\omega \rightarrow \infty$ $\bar{p} \rightarrow 0$.

Итак, как следует из рис. 4.2, зависимость $n(\omega)$ наиболее сильна вблизи резонансных частот, когда частота светового поля совпадает или достаточно близка к собственной частоте колебаний атомов или молекул. Это значит, что дисперсия показателя преломления наиболее сильна вблизи максимумов поглощения среды. Уже из этого факта вытекает, что между вещественной и мнимой частями показателя преломления существует сильная связь. Она имеет глубокую физическую основу и может быть выражена из принципа причинности. Соответствующее соотношение называют соотношением Крамерса – Кронига:

$$n(\omega) - 1 = \frac{2}{\pi} \int_0^{\infty} \frac{\omega'}{(\omega')^2 - \omega^2} \kappa(\omega') d\omega' \quad (4.15).$$

Это соотношение носит фундаментальный характер. Оно показывает, что величина $n(\omega)$ полностью определяется зависимостью $\kappa(\omega)$ и наоборот.

4.2. Распространение света в направляющих структурах

Как известно, узкий световой пучок не может распространяться в однородной среде без изменения своего поперечного размера вследствие дифракции. Сохранение профиля пучка возможно при распространении света в направляющих структурах – волоконных, канальных и планарных оптических волноводах. Эффект направления световой энергии в таких волноводах основан на явлении полного внутреннего отражения света на границе раздела диэлектрических сред. Первые работы по диэлектрическим оптическим волноводам выполнены в конце 60-х годов прошлого столетия. В настоящее время в этой области, включающей волоконную и интегральную оптику, выполнен огромный объем исследований и реализован целый ряд идей по созданию как отдельных функциональных элементов, так и более сложных интегрально – оптических схем для целей оптической обработки информации и оптической связи [3, 11].

В данном подразделе рассматриваются элементы электромагнитной теории планарного оптического волновода, являющегося базовым компонентом интегральной оптики, и кратко обсуждаются основные типы волноводно – оптических структур.

4.2.1. Электромагнитная теория планарного волновода

Рассмотрим планарный диэлектрический волновод, представляющий собой трехслойную структуру с размерами, не ограниченными в поперечных

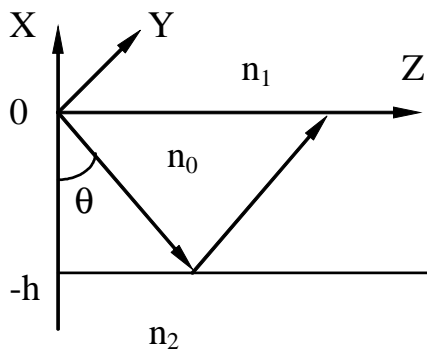


Рис. 4.3. Планарный волновод

направлениях. Средний слой с толщиной h имеет более высокий показатель преломления, в сравнении с двумя другими слоями (рис. 4.3). Примером подобных волноводов служит диэлектрическая пленка с высоким показателем преломления, нанесенная на подложку с более низким показателем преломления (например, пленка Ta_2O_5 на стеклянной подложке).

Проведем анализ планарного оптического волновода, используя волновой подход [3, 11]. Рассмотрим наиболее простой случай, соответствующий изотропной и непроводящей среде в пределах каждой из отдельных областей структуры. Считаем, что свет распространяется в направлении оси z , а в направлении оси y световое поле является однородным. Показатели преломления в пределах слоев структуры удовлетворяют соотношению: $n_2 < n_1 < n_0$. Ищем поле направляемых волн в виде:

$$\bar{E}(x, z) = \bar{E}_m(x) \cdot \exp[i(\omega t - \beta z)] \quad (4.16).$$

При анализе будем исходить из уравнений Максвелла:

$$\begin{cases} \text{rot} \bar{H} = \frac{\partial \bar{D}}{\partial t} \\ \text{rot} \bar{E} = -\frac{\partial \bar{B}}{\partial t} \end{cases} \quad (4.17),$$

где подразумевается, что $\rho = 0$; $\bar{\delta} = 0$.

Учтем, что $\bar{D} = \varepsilon \cdot \bar{E}$ и $\bar{B} = \mu \bar{H}$, тогда в декартовой системе координат эти уравнения принимают вид:

$$\begin{aligned} \frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z} &= \varepsilon \frac{\partial E_x}{\partial t} & \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} &= -\mu \frac{\partial H_x}{\partial t} \\ \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x} &= \varepsilon \frac{\partial E_y}{\partial t} & \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} &= -\mu \frac{\partial H_y}{\partial t} \\ \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} &= \varepsilon \frac{\partial E_z}{\partial t} & \frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} &= -\mu \frac{\partial H_z}{\partial t} \end{aligned}$$

Далее, учтем, что при заданном виде поля операция дифференцирования по времени эквивалентна умножению соответствующей компоненты на $i\omega$, а операция дифференцирования по z эквивалентна ее умножению на $-ik$.

Учтем также, что $\frac{\partial}{\partial y} = 0$, тогда получим:

$$\begin{aligned} \beta H_y &= \omega \varepsilon E_x & \beta E_y &= -\omega \mu H_x \\ -i\beta H_x - \frac{\partial H_z}{\partial x} &= i\omega \varepsilon E_y & -i\beta_x - \frac{\partial \bar{E}_z}{\partial x} &= -i\omega \mu H_y \\ \frac{\partial H_y}{\partial x} &= i\omega \varepsilon E_z & \frac{\partial E_y}{\partial x} &= -i\omega \mu H_z \end{aligned} \quad (4.18).$$

Видим, что все эти уравнения делятся на две независимых системы:

$$1) \begin{cases} -i\beta H_x - \frac{\partial H_z}{\partial x} = i\omega \varepsilon E_y \\ \beta E_y = -\omega \mu H_x \\ \frac{\partial E_y}{\partial x} = -i\omega \mu H_z \end{cases} \quad 2) \begin{cases} \beta H_y = \omega \varepsilon E_x \\ \frac{\partial H_y}{\partial x} = i\omega \varepsilon E_z \\ i\beta E_x + \frac{\partial E_z}{\partial x} = i\omega \mu H_y \end{cases} \quad (4.19)$$

Первая система связывает компоненты поля E_y , H_x и H_z , а вторая система – компоненты H_y , E_x , E_z . Это поперечно – электрические или *TE* и поперечно – магнитные или *TM* волны. Им соответствуют ориентации векторов поля в световой волне внутри слоя, показанные на рис. 4.4.

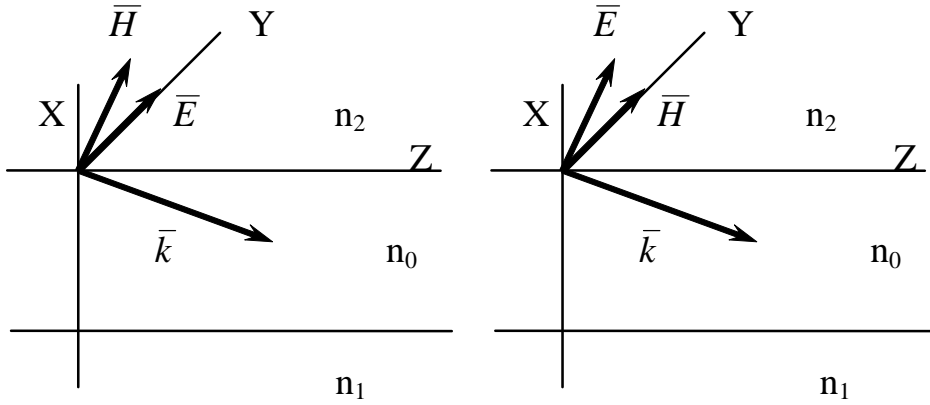


Рис. 4.4. Расположение векторов \vec{E} и \vec{H} в ТЕ волне (слева) и в ТМ волне (справа).

Обратим внимание на следующие 2 уравнения из систем (1) и (2):

$$\frac{\partial E_y}{\partial x} = -i\omega\mu H_z \quad \text{и} \quad \frac{\partial H_y}{\partial x} = i\omega\epsilon E_z \quad (4.19).$$

Их можно записать в форме:

$$H_z = -\frac{1}{i\omega\mu} \frac{\partial E_y}{\partial x} = \frac{i}{\omega\mu} \frac{\partial E_y}{\partial x} \quad (4.20)$$

и

$$E_z = \frac{1}{i\omega\epsilon} \frac{\partial H_y}{\partial x} = -\frac{i}{\omega\epsilon} \frac{\partial H_y}{\partial x} \quad (4.21).$$

Далее эти соотношения понадобятся при записи граничных условий, необходимых для получения дисперсионного уравнения волновода.

Как можно видеть, системы (4.17) являются замкнутыми, т.е. они позволяют получить соответствующие уравнения (волновые уравнения) для каждой из компонент поля. Так, подставляя в первое уравнение системы (4.17) выражения для H_x и H_z , получим:

$$-i\beta \left(-\frac{\beta}{\omega\mu} \right) E_y - \frac{\partial}{\partial x} \left(\frac{i}{\omega\mu} \frac{\partial E_y}{\partial x} \right) = i\omega\epsilon E_y \quad \text{или}$$

$$\frac{\beta^2}{\omega\mu} E_y - \frac{1}{\omega\mu} \frac{\partial^2 E_y}{\partial x^2} - \omega\epsilon E_y = 0 \quad \text{или}$$

$$\frac{\partial^2 E_y}{\partial x^2} - \beta^2 E_y + \omega^2 \epsilon \mu E_y = 0,$$

но $\omega^2 \epsilon \mu = k^2$, поэтому окончательно волновое уравнение для компоненты E_y принимает вид:

$$\frac{\partial^2 E_y}{\partial x^2} + (k^2 - \beta^2) \cdot E_y = 0 \quad (4.22).$$

Аналогично получим:

$$i\beta \frac{\beta}{\omega\epsilon} H_y + \frac{\partial}{\partial x} \left(-\frac{i}{\omega\epsilon} \frac{\partial H_y}{\partial x} \right) - i\omega\mu H_y = 0 \quad \text{или}$$

$$\frac{\beta^2}{\omega\epsilon} H_y - \frac{1}{\omega} \frac{\partial}{\partial x} \left(\frac{1}{\epsilon} \frac{\partial H_y}{\partial x} \right) - \omega\mu H_y = 0$$

Однако величину $(1/\epsilon)$ мы в общем случае не можем вынести из под знака дифференцирования, поскольку ϵ может меняться скачком (в рассматриваемом случае) на границах раздела $x=0$ и $x=-h$, или непрерывно в пределах световодного слоя в градиентных волноводах.

$$\text{Т.е. } \frac{\partial}{\partial x} \left(\frac{1}{\epsilon} \frac{\partial H_y}{\partial x} \right) = \frac{\partial}{\partial x} \left(\frac{1}{\epsilon} \right) \cdot \frac{\partial H_y}{\partial x} + \frac{1}{\epsilon} \frac{\partial^2 H_y}{\partial x^2} \quad (4.23).$$

Тогда:

$$\beta^2 H_y - \epsilon \left[\frac{\partial}{\partial x} \left(\frac{1}{\epsilon} \right) \frac{\partial H_y}{\partial x} + \frac{1}{\epsilon} \frac{\partial^2 H_y}{\partial x^2} \right] - \omega^2 \mu \epsilon H_y = 0 \quad (4.24)$$

В пределах каждого из слоев поле описывается уравнением:

$$\frac{\partial^2 H_y}{\partial x^2} + (k^2 - \beta^2) H_y = 0 \quad (4.25).$$

Подобным же образом можно получить и другие волновые уравнения из системы (4.17) для компонент, например, H_z и E_z .

Решения волновых уравнений.

Получим решения для волнового уравнения (4.21), т.е. для TE-мод. Вид распределений $E_y(x)$ определяется решением уравнения:

$$\frac{\partial^2 E_y}{\partial x^2} + (k^2 - \beta^2) E_y = 0,$$

где $k = nk_0$ – волновое число света в среде; $k_0 = \frac{2\pi}{\lambda}$ – волновое число в вакууме; n – показатель преломления среды. В случае, когда выполняется условие полного внутреннего отражения для световой волны в центральном слое с показателем преломления n_0 , величина $(k^2 - \beta^2) > 0$ для этого слоя, но $(k^2 - \beta^2) < 0$ для прилегающих сред.

Действительно, $k = k_0 n_0$ при $0 > x > -h$ и $k = n_1 k_0$ при $x < -h$; $k = n_2 k_0$ при $x > 0$. С другой стороны $\beta = n_0 k_0 \sin \theta$ и если $\theta > \theta_{кр}$, то $n_0 \sin \theta > n_1, n_2$.

$$\begin{aligned} \text{Соответственно, } (k_0 n_0)^2 - \beta^2 &> 0, \\ (k_0 n_1)^2 - \beta^2 &< 0, \\ (k_0 n_2)^2 - \beta^2 &< 0. \end{aligned}$$

Таким образом, решения волнового уравнения (4.21) для различных областей структуры имеют вид:

а) Волноводный слой ($n=n_0$):

$$E_y(x) = A_1 \exp[i\kappa x] + A_2 \exp[-i\kappa x] \quad (4.26),$$

где $\kappa = \sqrt{(k_0 n_0)^2 - \beta^2}$.

б) Подложка и покровный слой ($n=n_1, n_2$):

$$E_y(x) = B_1 \exp[i\gamma_{1,2} x] + B_2 \exp[-i\gamma_{1,2} x] \quad (4.27),$$

где $\gamma_{1,2} = \sqrt{\beta^2 - k_0^2 n_{1,2}^2}$.

Из условий конечности поля $E_y(\pm\infty)$ следует:

$$E_y(x)^{(1)} : E_y(-\infty)^{(1)} = 0 \rightarrow B_2^{(1)} = 0$$

$$E_y(x)^{(2)} : E_y(\infty)^{(2)} = 0 \rightarrow B_1^{(2)} = 0$$

Таким образом, окончательно имеем:

$$\begin{cases} E_y^{(2)}(x) = B \cdot \exp(-\gamma_2 x) \\ E_y^{(0)}(x) = A_1 \cdot \exp(i\kappa x) + A_2 \cdot \exp(-i\kappa x) \\ E_y^{(1)} = C \cdot \exp[\gamma_1(x+h)] \end{cases} \quad (4.28).$$

Аналогичные соотношения можно получить и для *TM* волн.

Для конкретизации распределений полей мы должны определить соотношения между постоянными A , B , C , что может быть сделано с использованием граничных условий.

4.2.2. Дисперсионное уравнение планарного волновода

Граничные условия позволяют получить дисперсионное уравнение, связывающее характеристики волноводных мод с частотой и параметрами направляющей структуры.

Исходим из соотношений (4.27) и стандартных граничных условий для тангенциальных компонент векторов \vec{E} и \vec{H} при $x=0$ и $x=-h$:

$$\begin{aligned} E_y^{(0)}(0) &= E_y^{(2)}(0) & H_z^{(0)} &= H_z^{(2)}(0) \\ E_y^{(0)}(-h) &= E_y^{(1)}(-h) & H_z^{(0)}(-h) &= H_z^{(1)}(-h) \end{aligned} \quad (4.29)$$

С учетом соотношения (4.18), выражения для H_z заменим на аналогичные

соотношения для $\frac{\partial E_y}{\partial x}$. Из (4.28) имеем:

$$\begin{aligned}
E_y^{(2)}(0) &= B \\
E_y^0(0) &= A_1 + A_2 \\
E_y^{(0)}(-h) &= A_1 \exp(-i\kappa h) + A_2 \exp(i\kappa h) \\
E_y^{(1)}(-h) &= C
\end{aligned}$$

$$\left. \frac{\partial E_y^{(2)}}{\partial x} \right|_{x=0} = -\gamma_2 \qquad \left. \frac{\partial E_y^{(0)}}{\partial x} \right|_{x=0} = i\kappa A_1 - i\kappa A_2$$

$$\begin{aligned}
\left. \frac{\partial E_y^{(0)}}{\partial x} \right|_{x=-h} &= i\kappa A_1 \exp(-i\kappa h) - i\kappa A_2 \exp(i\kappa h) \\
\left. \frac{\partial E_y^{(1)}}{\partial x} \right|_{x=-h} &= -\gamma_1 C
\end{aligned}$$

Для упрощения запишем распределения полей (13) в виде :

$$\begin{aligned}
E_y^{(2)}(x) &= B \cdot \exp(-\gamma_2 x) \\
E_y^{(0)}(x) &= A_1 \cdot \cos(\kappa x) + A_2 \cdot \sin(\kappa x) \\
E_y^{(1)}(x) &= C \exp[\gamma_1(x+h)]
\end{aligned} \tag{4.30}$$

И тогда:

$$E_y^{(2)}(0) = B \qquad \left. \frac{\partial E_y^{(2)}}{\partial x} \right|_{x=0} = -\gamma_2$$

$$E_y^{(0)}(0) = A_1 + A_2 \qquad \left. \frac{\partial E_y^{(0)}}{\partial x} \right|_{x=0} = \kappa A_2$$

$$E_y^{(1)}(-h) = C$$

$$\left. \frac{\partial E_y^{(0)}}{\partial x} \right|_{x=-h} = \kappa A_1 \cdot \sin(\kappa h) + \kappa A_2 \cdot \cos(\kappa h)$$

$$\left. \frac{\partial E_y^{(1)}}{\partial x} \right|_{x=-h} = -\gamma_1 C$$

В итоге получим:

$$1) B = A_1$$

$$2) -\gamma_2 B = \kappa A_2 \rightarrow A_2 = -\frac{\gamma_2}{\kappa} \cdot B$$

$$3) A_1 \cdot \cos(\kappa h) - A_2 \cdot \sin(\kappa h) = C$$

$$4) \kappa A_1 \sin(\kappa h) + \kappa A_2 \cos(\kappa h) = \gamma_1 C$$

Выразим все коэффициенты через B .

Выражение для полей с учетом найденных коэффициентов:

$$E_y^{(2)}(x) = B \exp(-\gamma_2 x)$$

$$E_y^{(0)}(x) = B \left[\cos \kappa h - \frac{\gamma_2}{\kappa} \sin(\kappa x) \right] \quad (4.31).$$

$$E_y^{(1)}(x) = B \left[\cos \kappa h + \frac{\gamma_2}{\kappa} \sin(\kappa h) \right] \exp[\gamma_1(x+h)] E$$

Однако в этих выражениях не определены величины κ , γ_1 , γ_2 , которые зависят от угла θ или от постоянной распространения β . Их можно определить, используя оставшееся соотношение 4):

$$4) \kappa A_1 \sin(\kappa h) + \kappa A_2 \cos(\kappa h) = \gamma_1 C$$

$$A_1 = B ; A_2 = -\frac{\gamma_2}{\kappa} \cdot B$$

$$C = B \cos(\kappa h) + \frac{\gamma_2}{\kappa} B \sin(\kappa h) = B \left[\cos(\kappa h) + \frac{\gamma_2}{\kappa} \sin(\kappa h) \right]$$

С учетом этих соотношений мы можем переписать выражения (4.31) для поперечного распределения полей в пределах каждого из слоев структуры. А из 4) найдем дисперсионное уравнение:

$$\sin(\kappa h) \left[\kappa - \frac{\gamma_1 \gamma_2}{\kappa} \right] = \cos(\kappa h) [\gamma_1 + \gamma_2] \quad \text{или}$$

$$\operatorname{tg}(\kappa h) = \frac{\vartheta_1 + \vartheta_2}{\kappa - \frac{\vartheta_1 \vartheta_2}{\kappa}} = \frac{\frac{\gamma_1}{\kappa} + \frac{\gamma_2}{\kappa}}{1 - \frac{\gamma_1 \gamma_2}{\kappa^2}}$$

Сравнивая это выражение с известным соотношением из тригонометрии

$$\operatorname{tg}(\alpha + \beta) = \frac{\operatorname{tg}\alpha + \operatorname{tg}\beta}{1 - \operatorname{tg}\alpha \operatorname{tg}\beta}, \quad \text{придем к соотношению:}$$

$$\kappa h = \alpha + \beta, \quad \text{где } \alpha = \operatorname{arctg}\left(\frac{\gamma_1}{\kappa}\right)$$

$$\beta = \operatorname{arctg}\left(\frac{\gamma_2}{\kappa}\right)$$

Или, с учетом периодичности функции tg ,

$$\kappa h = \operatorname{arctg}\left(\frac{\gamma_1}{\kappa}\right) + \operatorname{arctg}\left(\frac{\gamma_2}{\kappa}\right) + \pi m \quad (4.32).$$

Это и есть дисперсионное уравнение планарного диэлектрического волновода. Запишем его в несколько другом виде, учитывая что:

$$\kappa = \sqrt{(k_0 n_0)^2 - \beta^2} = k_0 \sqrt{n_0^2 - n_0^2 \sin^2 \theta}$$

$n_0 \sin \theta = N_m$ - эффективный показатель преломления.

$$\gamma_1 = \sqrt{\beta^2 - k_0^2 n_1^2} = k_0 \sqrt{N_m^2 - n_1^2}$$

$$\gamma_2 = k_0 \sqrt{N_m^2 - n_2^2}$$

Тогда, наконец, получаем:

$$k_0 h \sqrt{n_0^2 - N_m^2} = \arctg \sqrt{\frac{N_m^2 - n_1^2}{n_0^2 - N_m^2}} + \arctg \sqrt{\frac{N_m^2 - n_2^2}{n_0^2 - N_m^2}} + m\pi \quad (4.33).$$

Аналогичным образом можно получить дисперсионное уравнение и для ТМ мод тонкопленочного волновода.

4.2.3. Анализ дисперсионного уравнения

На рис. 4.5 представлен пример семейства дисперсионных кривых для ТЕ мод трехслойного планарного волновода со ступенчатым изменением

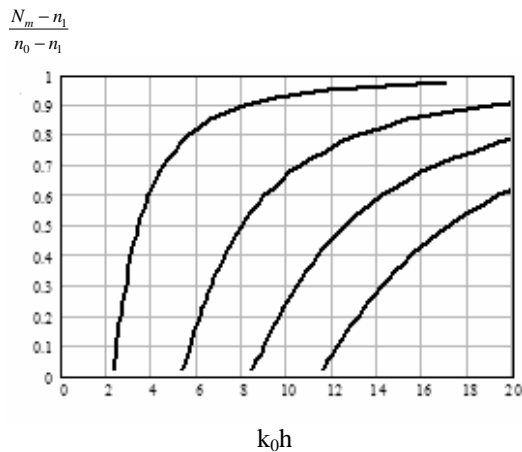


Рис. 4.5.

показателя преломления при переходе от слоя к слою, полученных путем численного решения уравнения (4.33). Из анализа данного уравнения можно сделать следующие выводы, часть из которых наглядно иллюстрируется характером дисперсионных кривых.

1) Это уравнение задает условия существования в структуре волноводной моды. Видим, что при заданных параметрах h , k_0 , n_0 , n_1 , n_2 и m существование моды возможно лишь при строго определенной величине N , т.е. угла падения θ на

границы раздела.

2) Величина N_m для той же структуры зависит от m -номера моды. Разные моды обозначаются как TE_0 , TE_1 TE_n , TM_0 , TM_1 ... TM_n . Номер m определяет число нулей в поперечном распределении поля.

3) При заданных параметрах структуры в ней может существовать ограниченное число мод, т.к. величина N_m изменяется в пределах $n_1 < N_m < n_0$.

4) Условие $N_m = n_1$ соответствует отсечке волноводной моды.

5) В случае асимметричного волновода ($n_1 \neq n_2$) для каждой моды существует критическая толщина пленки, соответствующая отсечке

волноводной моды ($N_m = n_1$). Ее величину можно найти из дисперсионного уравнения:

$$k_0 h_{kp} = \frac{m\pi + \operatorname{arctg} \sqrt{\frac{n_1^2 - n_2^2}{n_0^2 - n_1^2}}}{\sqrt{n_0^2 - n_1^2}}$$

6) Минимальная критическая толщина соответствует TE_0 -моду, для которой $m=0$:

$$k_0 h_{min} = \frac{\operatorname{arctg} \sqrt{\frac{n_1^2 - n_2^2}{n_0^2 - n_1^2}}}{\sqrt{n_0^2 - n_1^2}}$$

7) Для симметричного волновода ($n_1 = n_2$) не существует отсечки для TE_0 и TM_0 мод даже при $h=0$!

8) При увеличении толщины пленки число мод, поддерживаемых пленкой увеличивается.

9) Для конкретной волноводной структуры с ростом номера моды m уменьшается как N_m так и угол θ_m .

5. ЭЛЕМЕНТЫ КВАНТОВОЙ ОПТИКИ

5.1. Постоянная Планка

Введение Планком знаменитой постоянной открыло новый этап в современной физике и оптике. Одной из проблем физики в конце XIX – начале XX века являлась проблема теоретического описания эффекта теплового излучения. Классический подход оказался бессильным.

В 1900 г. Макс Планк получил свое выражение для распределения мощности излучения абсолютно черного тела по частотам, которое показало хорошее согласие теоретических результатов с экспериментальными как для низких, так и для высоких частот [7, 8]. Для самого ученого было поразительным, что его решение могло быть получено только в предположении, что электромагнитное излучение испускается не непрерывно, а в виде порций энергии – квантов энергии.

Энергия кванта W по Планку пропорциональна частоте излучения:

$$W = h\omega \quad (5.1),$$

где $h=1,05 \cdot 10^{-34}$ Дж·с – постоянная Планка. Сообщение Планка о своей работе на съезде Немецкого физического общества в Берлине 14 декабря 1900 года, по сути, знаменовало день рождения квантовой физики.

Однако Планк считал, что энергия только излучается квантами, а поглощается непрерывно. Он не решился сделать вывод о том, что само излучение не непрерывно, а состоит из отдельных квантов. Этот шаг был сделан в 1905 году А. Эйнштейном. Эйнштейн пришел к выводу, что монохроматическое излучение состоит из взаимно независимых квантов энергии, исходя из результата Планка. Для обоснования этой гипотезы Эйнштейн провел, в частности, анализ фотоэффекта (который в экспериментах впервые наблюдался Столетовым). Представление о квантах энергии объясняло явление “красной границы фотоэффекта”. Метод задерживающего потенциала позволил ранее выявить зависимость энергии фотоэлектронов от частоты излучения, попадающего на фотокатод вакуумной трубки. Резкое исчезновение фототока при превышении длины волны излучения некоторой величины, зависящей от материала фотокатода, и получило название эффекта “красной границы”. Эйнштейн постулировал, что излучение не только испускается, но и поглощается в виде квантов. Результатом стала знаменитая формула Эйнштейна для фотоэффекта:

$$W_k = h \cdot \omega - A \quad (5.2),$$

где W_k – кинетическая энергия фотоэлектрона; A – работа выхода. Отсюда частота, соответствующая «красной границе», определяется соотношением:

$$\omega_k = A/h \quad (5.3).$$

Эйнштейн постулировал также, что каждый квант обладает импульсом \bar{p} :

$$\bar{p} = h \cdot \bar{k} \quad (5.4),$$

где \bar{k} – волновой вектор. Фактически, это означало, что свет (или его элементарные частицы – фотоны, энергия которых определяется величиной кванта) в определенных ситуациях может проявлять свойства механических частиц.

5.2. Постулаты Бора

Представления о квантах энергии сыграли важную роль в развитии теории строения атома. В начале XX века уже было известно, что спектры излучения атомов не являются сплошными, а представляют собой набор отдельных линий. Попытки объяснения такого характера излучения атомов привели к появлению разных моделей строения атома. Одна из них – планетарная модель. Она предполагала движение электронов вокруг ядра по круговым орбитам разного радиуса, а период вращения соответствовал разным частотам излучения. Однако такое вращение является неравномерным движением и подразумевает постоянное излучение энергии, что делает невозможным существование стабильных орбит электронов. В этом случае классический подход также оказался бессильным.

В 1913 году датский физик Нильс Бор предложил теорию атома водорода. В ее основе лежали два постулата:

1. Наличие разрешенных орбит, находясь на которых, электроны, вопреки требованиям классической электродинамики, не испускают излучения.
2. Испускание или поглощение энергии происходит при переходе электрона с одной разрешенной орбиты на другую, или, иначе говоря, при переходе атома с одного энергетического уровня на другой. Эти переходы называют квантовыми. При каждом таком переходе испускается или поглощается один фотон, если речь идет о поглощении или испускании света. Его энергия $h\omega$ равна разности энергий соответствующих энергетических уровней E_n и E_m :

$$h\omega = E_n - E_m \quad (5.5).$$

Это соотношение названо правилом частот Бора.

Радиус орбиты электрона r_n по Бору определяется выражением:

$$r_n = \frac{h^2 \cdot n}{m \cdot e^2} \quad (5.6),$$

где m и e – масса и заряд электрона; $n=1, 2, \dots$ – целое число.

Теория Бора явилась важным шагом в развитии квантовой физики. Одной из ее ошибок было постулирование определенных (фиксированных) орбит электронов. В последовательной квантовой теории таких орбит нет, но есть энергетические уровни и правило частот Бора.

5.3. Корпускулярно – волновой дуализм

В 1924 году Луи де Бройль (видимо, под впечатлением положения о двойственности природы света) высказал идею о том, что квантовые соотношения Планка и Эйнштейна $W = h\omega$ и $\bar{P} = h \cdot \bar{k}$, характеризующие свет, справедливы и для всех материальных частиц. Это означало, что в определенных ситуациях частицы могут проявлять волновые свойства. Соотношение $\bar{P} = h \cdot \bar{k}$ обычно записывается в форме:

$$p = \frac{2\pi}{\lambda} \cdot h \quad (5.7),$$

где величина λ получила название «длина волны де Бройля». В соответствии с этой идеей де Бройль связал движение свободной частицы с волновой функцией:

$$\psi = A \cdot \exp[i(\omega t - \bar{k} \cdot \bar{r})] = A \cdot \exp[i(\omega t - k \cdot \xi)] \quad (5.8).$$

Это выражение можно записать в виде:

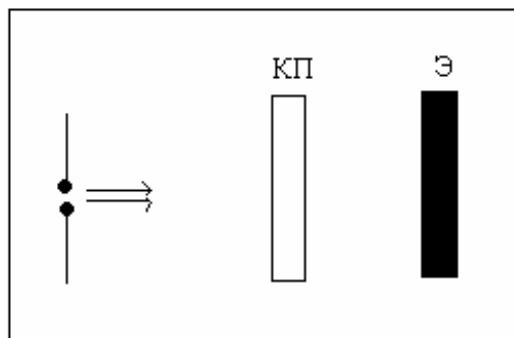
$$\psi = A \cdot \exp\left[i\left(\frac{W}{h} \cdot t - \frac{p}{h} \cdot \xi\right)\right] \quad (5.9),$$

где W – энергия частицы, p – величина импульса ($p=m \cdot V$, где V – скорость движения частицы). Это выражение соответствует плоской волне и называется волной де Бройля. Таким образом, движению свободной частицы соответствует волновой процесс с длиной волны $\lambda=2\pi \cdot h/p$ и частотой колебаний $\omega=W/h$. Т.е. волновые характеристики частицы связаны с ее динамическими характеристиками W и p . Конечно, необходимо отметить, что в общем случае движение частиц описывается более сложными функциями, точно так же, как реальные электромагнитные поля описываются более сложными выражениями, по сравнению с полем плоской волны.

Идея де Бройля получила подтверждение в 1927 году, в эксперименте по дифракции электронов на кристаллической решетке.

Схема эксперимента по дифракции электронов

Схема эксперимента по дифракции электронов иллюстрируется рисунком 5.1. Здесь после прохождения пучка электронов через



кристаллическую пластину (КП) на экране Э наблюдалась картина дифракции. Длина волны, которую можно было ассоциировать с соответствующими электронам волнами, найденная по интерференционной картине, соответствовала расчетной. Позднее были проведены другие эксперименты в схеме,

Рис. 5.1.

аналогичной схеме интерферометра Юнга. Оказалось, что интерференционная картина формируется даже в случае пролета через щель одиночных электронов (пучок очень малой интенсивности). Это стало окончательным подтверждением дуализма свойств частиц.

Корпускулярно – волновой дуализм понимается в настоящее время как потенциальная способность микрообъектов проявлять различные свойства в зависимости от условий (в частности от условий наблюдения).

5.4. Физическая интерпретация волн де Бройля

В свое время предпринимались многочисленные попытки наглядно интерпретировать волны де Бройля, т.е. представить микрообъект в виде симбиоза корпускулы и волны. По одной из гипотез, волны де Бройля представлялись в виде волн материи. По другой, частицы пытались представить в виде пакетов волн. В конце концов, стало ясно, что волновые функции де Бройля имеют лишь статистический, вероятностный смысл. Такая интерпретация была предложена немецким физиком Максом Борном. Согласно ей, квадрат модуля волновой функции $|\psi|^2$ определяет вероятность нахождения частицы в момент времени t в данной точке пространства.

Пусть dV - бесконечно малый элемент пространства, в котором движется частица. Согласно М.Борну, вероятность локализации частицы в момент t в элементе dV :

$$dF = |\Psi(x, y, z, t)|^2 dV \quad (5.10).$$

Здесь x, y, z – координаты точки, принадлежащей объему dV . Для объема V конечных размеров вероятность нахождения частиц в нем:

$$F = \int_V |\Psi(x, y, z, t)|^2 dV \quad (5.11).$$

Если V стремится к бесконечности, то

$$\lim_{V \rightarrow \infty} F(V, t) = \int_V |\Psi(x, y, z, t)|^2 dV = 1 \quad (5.12),$$

т.к. это вероятность достоверного события (частица обязательно где – то находится). Кстати, последнее соотношение для волны де Бройля не выполняется (интеграл расходится). Данный результат свидетельствует о том, что эта волна является математической абстракцией.

В заключение посмотрим, чему равны фазовая и групповая скорости волн де Бройля. Пусть частицы и соответствующая волна движутся вдоль оси X . В этом случае:

$$\psi(x, t) = A \cdot \exp[i(\omega t - kx)] = A \cdot \exp\left[i\left(\frac{W}{h} \cdot t - \frac{p}{h} \cdot x\right)\right] \quad (5.13).$$

По определению $v_\phi = \frac{\omega}{k}$, и для рассматриваемой задачи $\omega = \frac{W}{h}$. Как известно из теории относительности, энергия движущейся частицы (при $v^2 \ll c^2$) определяется соотношением:

$$E = \sqrt{m_n^2 \cdot c^4 + p^2 \cdot c^2} = m_n c^2 \cdot \left(1 + \frac{p^2 c^2}{2 \cdot m_n^2 \cdot c^4} \right) + \dots = m_n c^2 + \frac{p^2}{2m_n} + \dots$$

Тогда:

$$\omega = \frac{E}{h} = \frac{m_n c^2}{h} + \frac{p^2}{2m_n h} = \frac{m_n c^2}{h} + \frac{k^2 h}{2m_n}$$

Соответственно, получим: $v_\phi = \frac{\omega}{k} = \frac{m_n c^2}{kh} + \frac{kh}{2m_n}$. Видим, что величина

фазовой скорости зависит от k , т.е. от длины волны. Таким образом, для волн де Бройля характерна дисперсия фазовой скорости. Важность этого факта заключается в невозможности интерпретации волновых свойств частицы ее представлением в виде пакета волн де Бройля. Очевидно, вследствие дисперсии такой пакет не может сохранять свою форму в пространстве и времени, он расплывается.

Групповая скорость определяется соотношением:

$$v_g = \frac{d\omega}{dk} = \frac{kh}{m_n} = \frac{p}{m_n} = \frac{m_n v}{m_n} = v$$

Таким образом, групповая скорость волны де Бройля равна скорости самой частицы.

5.5. Соотношения неопределенностей

Вероятностная трактовка физического смысла волновой функции ψ показывает, что координаты движущейся частицы для наблюдателя оказываются неопределенными. Можно говорить лишь о вероятности нахождения частицы в некоторой области пространства.

В экспериментах положение частицы можно определить и точно, например, при ее попадании на сцинтиллирующий экран. Однако при этом импульс частицы изменяется и возникает неопределенность импульса. При точном определении импульса (например, по интерференционной картине при дифракции электронов на двух щелях) неопределенной становится координата частицы.

Возможная точность одновременного измерения координаты и импульса частицы определяется соотношением Гейзенберга:

$$\Delta x \cdot \Delta p_x \approx h \quad (5.14),$$

которое является одним из соотношений неопределенностей. Оно указывает,

что, чем выше точность определения координаты, тем ниже точность определения импульса. Неопределенность значения импульса можно выразить соотношением $\Delta p_x = m \cdot \Delta v_x$, поэтому:

$$\Delta x \cdot \Delta v_x \approx \frac{h}{m} \quad (5.15).$$

Последнее соотношение объясняет, почему не возникает проблем при одновременном измерении координат и импульса макроскопических частиц или объектов в классической физике. Очевидно, что при большой массе m произведение $\Delta x \cdot \Delta v_x$ мало, и траектория частицы может быть определена сколь угодно точно.

Аналогичное соотношение имеет место для энергии – времени:

$$\Delta E \cdot \Delta t \geq h \quad (5.16),$$

где ΔE и Δt – неопределенность значения энергии и времени нахождения частицы в состоянии с такой энергией. Действительно, следуя формуле Эйнштейна, для определения точного значения энергии необходимо точно задать частоту. Но в случае точного задания момента времени частота в принципе не может быть определена, т.к. определение частоты требует наблюдения процесса в течение некоторого конечного интервала времени.

Время Δt , определяемое соотношением (5.16), характеризует время жизни частицы на определенном энергетическом уровне E_i . Если $E_i = const$, т.е. $\Delta E_i \rightarrow 0$, то $\Delta t \rightarrow \infty$, т.е. частица находится в стационарном состоянии. Если время жизни ограничено величиной $\Delta \tau$, то энергия частицы имеет разброс в пределах $\Delta E_i \approx h \cdot \Delta \tau^{-1}$. Величину ΔE_i называют “шириной энергетического уровня”.

Общий смысл соотношений неопределенностей заключается в том, что наблюдаемые в микромире явления носят неопределенный, случайный характер и подчиняются при этом законам теории вероятностей. С одной из форм соотношений неопределенности мы часто встречаемся в традиционной радиотехнике. Это:

$$\Delta \omega \cdot \Delta \tau \approx 1 \quad (5.17).$$

Оно показывает, что частоту ω даже просто синусоидального пакета можно определить более – менее точно лишь при очень большой длительности пакета. В случае малой длительности, пакет обладает достаточно широким спектром и ему невозможно поставить в соответствие какую – либо строго определенную частоту.

5.6. Уравнение Шредингера

Важную роль в квантовой физике играет волновая теория квантовых явлений в реальном микромире, сформулированная немецким физиком Эрвином Шредингером в 1926 году. Это феноменологическая теория, в

основу которой положены два достаточно сильных допущения. **Во – первых**, предполагается, что в рассматриваемых явлениях не происходит рождение или исчезновение частиц, т.е. в любом процессе число частиц данного типа остается постоянным. **Во – вторых**, в нерелятивистской теории Шредингера полагается, что скорость частиц много меньше, чем скорость света. Таким образом, данная теория применима лишь к частицам, имеющим массу покоя. Для фотонов, не имеющих массы покоя, необходимы специальные рассмотрения. Кроме того, такие важные эффекты, как усиление и генерация света в квантовых системах, как раз основаны на эффектах поглощения и излучения (рождения) фотонов и, казалось бы, находятся вне поля применимости данной теории. Тем не менее, теория Шредингера используется при решении многих задач квантовой физики и, в принципе, позволяет описать достаточно сложные квантовые системы. Ключевым соотношением в этой теории является уравнение Шредингера, по значимости эквивалентное волновому уравнению в классической электродинамике. Один из путей, которым можно прийти к уравнению Шредингера, основан на формальной аналогии между видом волновой функции для волн де Бройля и выражением для поля плоской электромагнитной волны [8]. Воспользуемся данным упрощенным подходом, хотя, конечно, данный результат может быть получен и другим, «более квантово – физическим», путем.

Как известно, в классической электродинамике волновые решения вида $E = A \cdot \exp[i(\omega t - kx)]$, отвечающие плоским электромагнитным волнам, удовлетворяют волновым уравнениям вида:

$$\frac{d^2 E}{dx^2} + k^2 E = 0 \quad (5.18).$$

Волновая функция в квантовой физике, описывающая поведение свободно движущейся микрочастицы, по виду совпадает с выражением для поля плоской электромагнитной волны:

$$\psi = A \cdot \exp[i(\frac{W}{h}t - \bar{k} \cdot \bar{r})] = A \cdot \exp[i(\frac{W}{h}t - k_x x - k_y y - k_z z)] \quad (5.19).$$

Следовательно, ей можно поставить в соответствие уравнение:

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} + k^2 \psi = 0 \quad (5.20),$$

которое должно играть по отношению к волновой функции ψ такую же роль, как уравнение (5.18) по отношению к выражению для поля плоской электромагнитной волны. Используя формализм оператора «набла», (5.20) может быть записано в виде:

$$\nabla^2 \Psi + k^2 \Psi = 0 \quad (5.21).$$

В (5.20) и (5.21) $k^2 = k_x^2 + k_y^2 + k_z^2$. Для свободной частицы, как отмечалось,

$k^2 = p^2/h^2$. Учитывая, что $p = m \cdot v$, а $p^2 = m^2 v^2 = 2m \cdot \frac{mv^2}{2} = 2m \cdot W_k$,

где W_k – кинетическая энергия частицы, уравнение (5.21) можно представить в форме:

$$\nabla^2 \psi + \frac{2m}{h^2} W_k \cdot \psi = 0 \quad (5.22).$$

Это уравнение характеризует волновые свойства свободной частицы с массой m , т.е. это – волновое уравнение частицы. Важный случай соответствует движению микрочастицы под действием внешних сил (частица движется в потенциальном поле внешних сил). В этом случае ее кинетическая энергия равна:

$$W_k = W - U \quad (5.23),$$

где W – полная энергия, а $U = U(x, y, z)$ – потенциальная энергия частицы. Соответственно, в случае движения частицы в некотором силовом поле, ее волновое уравнение имеет вид:

$$\nabla^2 \psi + \frac{2m}{h^2} (W - U) \psi = 0 \quad (5.24).$$

Это уравнение записано для случая гармонической временной зависимости или для состояния частицы с заданной энергией. Обращаясь к классической электродинамике или теории сигналов, вспомним, что любой произвольно изменяющийся во времени процесс может быть представлен в виде набора гармонических составляющих, и ему нельзя поставить в соответствие какую – то определенную частоту. Точно так же поведение реальной квантовой системы может быть описано с помощью набора элементарных волн де Бройля и ее волновой функции не может соответствовать точно заданная энергия, которая также изменяется при изменении состояния квантовой системы. Вместе с тем, волновая функция такой системы должна определять ее состояние в любой момент времени, а для этого необходимо, чтобы между значениями волновой функции и ее производной существовала однозначная связь. Таким образом, уравнение (5.24) желательно бы преобразовать к форме, в которой не фигурирует величина энергии системы, но присутствуют как волновая функция, так и ее производная. Для этого выражение для волновой функции в виде:

$$\psi(\vec{r}, t) = \exp(i \frac{W}{h} t) \cdot \psi_0(\vec{r}) \quad (5.25)$$

продифференцируем по времени t :

$$\frac{\partial \psi}{\partial t} = i \frac{W}{h} \cdot \exp(i \frac{W}{h} t) \cdot \psi_0(\vec{r}) = i \frac{W}{h} \cdot \psi(\vec{r}) \quad (5.26).$$

Отсюда получим:

$$W\psi = -i\hbar \frac{\partial \psi}{\partial t} \quad (5.27).$$

Подставляя выражение для $W\psi$ в (5.24), приходим к уравнению, в которое не входит энергия системы, но присутствуют как сама волновая функция, так и ее производная по времени:

$$\left[\frac{\hbar^2}{2m} \nabla^2 - U \right] \psi = i\hbar \frac{\partial \psi}{\partial t} \quad (5.28).$$

Это соотношение справедливо и при произвольной зависимости $\psi(t)$. Уравнения (5.24) и (5.28) – это уравнения Шредингера. Уравнение (5.24) описывает стационарные состояния системы, т.е. состояния с точно заданной энергией. Уравнение (5.28) является более общим, оно описывает, в том числе, и системы с произвольно изменяющейся во времени волновой функцией.

Волновая функция ψ , являющаяся решением уравнения Шредингера, определяет возможные состояния физической системы (т.е. частицы или группы частиц). Соответственно, ее называют также функцией состояния. Например, волна де Бройля описывает состояние частицы, движущейся в пространстве с постоянной скоростью. Состояния, представляемые волновыми функциями вида (5.25) $\psi(\vec{r}, t) = \exp(i \frac{W}{\hbar} t) \cdot \psi_0(\vec{r})$, в квантовой механике называются стационарными.

Уравнение Шредингера – это линейное дифференциальное уравнение. Соответственно, для систем, описываемых им, справедлив принцип суперпозиции. Т.е., если ψ_1 и ψ_2 – волновые функции, соответствующие двум состояниям системы, то $a_1\psi_1 + a_2\psi_2$ – также волновая функция этой системы.

Из физических соображений ясно, что волновая функция должна быть однозначной, конечной и непрерывной функцией пространственных координат и времени. Действительно, $|\psi|^2$ определяет плотность вероятности локализации частицы в заданной точке пространства в заданный момент времени t . Если ψ не удовлетворяет перечисленным выше требованиям, то величина $|\psi|^2$ теряет физический смысл. Например, при неоднозначности ψ вероятность нахождения частицы в заданной точке при одних и тех же условиях имеет несколько значений.

5.7. Операторы.

В квантовой физике многие физические величины могут квантоваться, т.е. в некоторых случаях они принимают лишь дискретные значения. Для математического описания таких величин не пригодны обычные

непрерывные функции, используемые в классической теории. Квантовая теория использует более общие математические методы, основой которых являются операторы [8].

Оператор – это математический символ, определяющий совокупность действий, которые надо провести над заданной функцией U для получения некоторой другой функции V . В общем случае будем обозначать оператор символом \hat{L} , изображая его действие на некоторую функцию U в виде произведения $\hat{L}U$. В итоге оператор \hat{L} определяется соотношением:

$$\hat{L}U = V \quad (5.29).$$

Вообще говоря, с операторами мы встречаемся достаточно часто и в классической физике и математике. Типичными примерами являются символы, предписывающие выполнить операции дифференцирования (d/dx), интегрирования (\int), извлечения квадратного корня, логарифмирования и т.д.

Наибольший практический интерес представляют линейные операторы, удовлетворяющие условию:

$$\hat{L}[C_1U_1(x) + C_2U_2(x)] = C_1\hat{L}U_1(x) + C_2\hat{L}U_2(x) \quad (5.30).$$

Здесь C_1 и C_2 – произвольные постоянные. К линейным относятся операторы дифференцирования, интегрирования, операторы Гамильтона, Лапласа и некоторые другие. Условию линейности не удовлетворяют, например, операторы логарифмирования, извлечения корней, возведения в степень.

Важное место в квантовой физике занимают самосопряженные или эрмитовы операторы. Самосопряженный линейный оператор \hat{L} должен удовлетворять соотношению:

$$\int U_1(x)\hat{L}U_2(x)dx = \int U_2(x)\hat{L}^*U_1(x)dx \quad (5.31).$$

Здесь \hat{L}^* – оператор, комплексно – сопряженный с \hat{L} (как и в случае обычных комплексно – сопряженных величин или функций, он получается из оператора \hat{L} также заменой i на $-i$).

В качестве примера покажем, что оператор $\hat{L} = i \frac{\partial}{\partial x}$ является самосопряженным. Для этого положим, что некоторые функции $U_1(x)$ и $U_2(x)$ определены на промежутке $-\infty < x < \infty$, причем $U_1(\pm\infty) = U_2(\pm\infty) = 0$. В этом случае левая часть соотношения (5.31) принимает вид:

$$\int_{-\infty}^{+\infty} U_1(x) i \frac{\partial U_2(x)}{\partial x} dx = i \cdot \int_{-\infty}^{+\infty} U_1(x) \cdot d[U_2(x)] \quad (5.32).$$

Проведя интегрирование по частям, получим:

$$\begin{aligned}
& i \cdot \int_{-\infty}^{+\infty} U_1(x) \cdot d[U_2(x)] = i \cdot U_1(x)U_2(x) \Big|_{-\infty}^{+\infty} - i \cdot \int_{-\infty}^{+\infty} U_2(x) \cdot d[U_1(x)] = \\
& = 0 - i \cdot \int_{-\infty}^{+\infty} U_2(x) \cdot \frac{\partial U_1(x)}{\partial x} dx = \int_{-\infty}^{+\infty} U_2(x) \cdot (-i) \frac{\partial U_1(x)}{\partial x} dx = \\
& = \int_{-\infty}^{+\infty} U_2(x) \cdot (-i \frac{\partial}{\partial x}) U_1(x) dx = \int_{-\infty}^{+\infty} U_2(x) \cdot \hat{L}^* U_1(x) dx
\end{aligned} \tag{5.33}$$

Таким образом, условие самосопряженности для оператора $i(\partial/\partial x)$ действительно выполняется. В то же время легко, например, показать, что оператор $\partial/\partial x$ не является самосопряженным.

Операторы можно складывать, вычитать и перемножать по правилам обычной алгебры. Однако при перемножении нельзя менять порядок сомножителей.

Собственное значение оператора

В ряде случаев воздействие оператора \hat{L} на некоторую функцию $U(x)$ эквивалентно умножению этой функции на постоянную L . Эта постоянная и называется собственным значением оператора \hat{L} :

$$\hat{L}U(x) = L \cdot U(x) \tag{5.34}$$

Аналогично, функция $U(x)$, удовлетворяющая этому соотношению, называется собственной функцией оператора \hat{L} . В случае линейного дифференциального оператора уравнение (5.34) оказывается линейным дифференциальным уравнением. Из курса математики известно, что такое уравнение для заданных граничных условий имеет ненулевые решения лишь при определенных значениях L . Эти постоянные и являются собственными значениями оператора. Дифференциальный оператор, как правило, имеет множество собственных значений и собственных функций. Совокупность всех собственных значений образует спектр, который может быть как сплошным, так и дискретным.

С примером дискретного спектра встречаемся, решая, например, задачу о движении частицы между двумя отражающими плоскостями. Для простоты считаем, что частица движется вдоль оси x между плоскостями $x=0$ и $x=a$. Полагаем, что потенциальная энергия U частицы при $0 < x < a$ равна нулю. При $x=0$ и $x=a$ она становится бесконечной ($U \rightarrow \infty$). При этих условиях частица может находиться лишь внутри промежутка между плоскостями, поскольку она не в состоянии преодолеть бесконечно высокий потенциальный барьер. Т.е. данные плоскости для частицы являются идеально отражающими. В этом случае говорят, что частица находится в бесконечно глубокой потенциальной яме. Движение частицы при этом описывается уравнением Шредингера в форме:

$$\frac{d^2\psi}{dx^2} + \frac{2m}{h^2}W\psi = 0 \text{ при } (0 \leq x < a) \quad (5.35).$$

В точках $x=0$; $x=a$ функция ψ обращается в 0 (поскольку стенки – идеально отражающие, то вероятность нахождения на них частицы равна нулю). Решение этого уравнения ищем в виде:

$$\psi = [A \cdot \cos(qx) + B \cdot \sin(qx)] \cdot \exp\left(i \frac{W}{h} t\right) \quad (5.36),$$

где $q = \sqrt{\frac{2m}{h^2}W}$. Исходя из граничных условий $\psi = 0|_{x=0, x=a}$, находим:

$$A = 0; \quad q = \frac{n\pi}{a}, \text{ где } n=1, 2, 3, \dots$$

Тогда:

$$W = \frac{q^2 h^2}{2m} = \frac{n^2 \pi^2 h^2}{2a^2 m} \quad (5.37).$$

Таким образом, полная энергия частицы в рассматриваемой задаче может принимать лишь дискретные значения W_n , соответствующие различным значениям n . Эти значения образуют бесконечный ряд дискретных энергетических уровней, характеризующих различные состояния частицы. При переходе частицы из одного состояния в другое, ее энергия должна меняться скачком, т.е. в данной ситуации имеет место квантование энергии. Поскольку уравнение (5.35) может быть представлено в форме:

$$\frac{d^2}{dx^2}\psi = -\frac{2m}{h^2}W\psi, \text{ то постоянная } -\frac{2m}{h^2}W, \text{ очевидно, представляет собой}$$

собственное значение оператора $\frac{d^2}{dx^2}$. Таким образом, мы пришли к выводу о дискретности спектра собственных значений данного оператора.

В качестве примера сплошного спектра собственных значений оператора рассмотрим оператор d/dx . Дифференциальное уравнение, определяющее его собственные значения, имеет вид:

$$\frac{dU}{dx} - LU = 0 \quad (5.38).$$

Решение данного уравнения, очевидно, легко записать:

$$U = A \cdot \exp(Lx) \quad (5.39).$$

Для того чтобы функция U имела физический смысл, она должна быть ограничена на бесконечности. Это возможно лишь в случае, если $L = \pm i|L|$. Таким образом, собственными значениями оператора d/dx является сплошной спектр мнимых чисел.

Отметим, что первый из приведенных примеров имеет близкую физическую аналогию в курсе электродинамики. Он соответствует

распространению электромагнитных волн в волноводных структурах, где собственные значения указанного оператора определяют дискретный спектр значений постоянных распространения направляемых волн.

Из курса математики известно, что собственные значения самосопряженного оператора всегда вещественны. В первом из примеров оператор d^2/dx^2 является самосопряженным, и его собственные значения действительно оказались вещественными. Напротив, оператор d/dx не удовлетворяет условию самосопряженности (это можно проверить самостоятельно) и его собственные значения являются мнимыми числами.

Важным свойством собственных функций линейных самосопряженных операторов является их ортогональность. Условие ортогональности определяется соотношением:

$$\int_{\nu} U_m^* U_n dv = 0 \quad \text{при } m \neq n \quad (5.40).$$

Для нормированных собственных функций это условие представляют в форме:

$$\int_{\nu} U_m^* U_n dv = \begin{cases} 1 & \text{при } m = n \\ 0 & \text{при } m \neq n \end{cases} \quad (5.41).$$

Совокупность собственных функций оператора U_n ($n = 1, 2, 3, \dots$) образует полную систему функций. Свойство полноты характеризуется тем, что любую функцию φ , заданную на том же интервале, что и U_n , и удовлетворяющую тем же граничным условиям, можно разложить в ряд по данным собственным функциям:

$$\varphi = \sum_n a_n U_n \quad (5.42),$$

где a_n – коэффициенты разложения, определяемые формулой:

$$a_n = \int_{\nu} \varphi U_n^* dv \quad (5.43).$$

Данный ряд будет сходящимся, если $\int_{\nu} |\varphi|^2 dv$ существует.

Операторы и физические величины в квантовой механике

Для аналитического описания физических величин в квантовой механике используются соответствующие этим величинам операторы. При этом постулируется, что значения физической величины, точно определяемой при опыте, должны совпадать с собственными значениями соответствующего оператора. Поскольку величины, имеющие физический смысл, должны представляться вещественными числами, то операторы физических величин должны быть обязательно самосопряженными. Рассмотрим некоторые наиболее важные операторы квантовой механики.

1. Оператор полной энергии.

Перепишем уравнение Шредингера для стационарных состояний (5.24) в форме:

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + U(x, y, z) \right] \psi = W\psi \quad (5.44).$$

Отсюда видим, что воздействие на волновую функцию ψ оператора $\left[-\frac{\hbar^2}{2m} \nabla^2 + U \right]$ равносильно умножению ψ на W , т.е. величина W в этом соотношении играет роль собственного значения данного оператора. Таким образом, $\left[-\frac{\hbar^2}{2m} \nabla^2 + U(x, y, z) \right]$ – не что иное, как оператор полной энергии.

Он называется гамильтонианом и обозначается \hat{H} :

$$\hat{H} = \left[-\frac{\hbar^2}{2m} \nabla^2 + U(x, y, z) \right] \quad (5.45).$$

Функцию $U(x, y, z)$ мы также должны заменить на оператор потенциальной энергии, но, поскольку эта энергия может принимать любые значения, данный оператор должен просто совпадать с $U(x, y, z)$. Точно так же величины x, y, z заменяем на их операторы, которые совпадают с x, y, z . Таким образом, уравнение Шредингера для стационарных состояний можно окончательно записать в виде:

$$\hat{H}\psi = W\psi \quad (5.46).$$

А для общего случая:

$$-i\hbar \frac{\partial \psi}{\partial t} = \hat{H}\psi \quad (5.47).$$

Т.к. $W = W_k + U(x, y, z)$, то величину $-\frac{\hbar^2}{2m} \nabla^2$ следует рассматривать, как оператор кинетической энергии

$$\hat{W}_k = -\frac{\hbar^2}{2m} \nabla^2 = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \quad (5.48).$$

Но $W_k = \frac{mV^2}{2} = \frac{m^2V^2}{2m} = \frac{p^2}{2m}$, следовательно:

$$\hat{W}_k = \frac{\hat{p}^2}{2m} \quad (5.49),$$

где \hat{p} – оператор импульса.

Сравнивая соотношения $\hat{W}_k = -\frac{\hbar^2}{2m} \nabla^2$ и $\hat{W}_k = \frac{\hat{p}^2}{2m}$ видим, что

$\hat{p}^2 = -\hbar^2 \nabla^2$. Это соотношение будет выполняться при условии:

$$\hat{p} = i\hbar\nabla \quad (5.50),$$

где \hat{p} – это оператор импульса частицы. Проекции оператора \hat{p} на оси x, y, z :

$$\hat{p}_x = i\hbar \frac{\partial}{\partial x}; \quad \hat{p}_y = i\hbar \frac{\partial}{\partial y}; \quad \hat{p}_z = i\hbar \frac{\partial}{\partial z} \quad (5.51).$$

5.8. Гармонический осциллятор

Рассмотрим частицу с массой m , смещенную на некоторое малое расстояние относительно положения устойчивого равновесия. На частицу действует возвращающая упругая сила. Считая, что частица движется вдоль прямой, уравнение ее движения можно записать в виде:

$$m \frac{d^2 x}{dt^2} = F = -\alpha \cdot x \quad (5.52).$$

Здесь x – отклонение частицы от положения равновесия; $F = -\alpha \cdot x$ – упругая сила, возвращающая ее в это положение; α – величина, называемая коэффициентом упругости. Очевидно, что решение данного уравнения

$$\frac{d^2 x}{dt^2} + \frac{\alpha}{m} x = 0 \quad (5.53)$$

в случае $\alpha = \text{const}$ находится очень просто и в вещественной форме имеет вид:

$$x = A \cdot \cos\left(\sqrt{\frac{\alpha}{m}} \cdot t + \varphi\right) \quad (5.54).$$

Видим, что в указанных условиях частица совершает около положения равновесия гармонические колебания с частотой, определяемой только физическими параметрами m и α :

$$\omega = \sqrt{\frac{\alpha}{m}} \quad (5.55).$$

Такая идеализированная система называется гармоническим осциллятором.

Модель гармонического осциллятора применима к любой системе, совершающей гармонические колебания с малой амплитудой вблизи состояния устойчивого равновесия (атом в молекуле, электрон в атоме, математический маятник и т.д.). Поскольку многие явления, рассматриваемые в рамках данного курса, связаны с изменением состояния электронов, атомов и других квантовых систем, проведем анализ некоторых свойств гармонического осциллятора.

Используя известное выражение для импульса частицы:

$$p = mv = m \frac{dx}{dt} \quad (5.56),$$

преобразуем соотношение (5.52):

$$m \frac{d^2 x}{dt^2} = -\alpha \cdot x \Rightarrow m \frac{dv}{dt} = -\alpha \cdot x \Rightarrow \frac{dp}{dt} = -\alpha \cdot x \quad (5.57).$$

Таким образом, имеем два соотношения:

$\frac{dp}{dt} = -\alpha \cdot x$ и $\frac{dx}{dt} = \frac{p}{m}$. Дифференцируя первое по времени, получим аналог

уравнения (5.52) для импульса частицы:

$$\frac{d^2 p}{dt^2} = -\alpha \frac{p}{m} \quad (5.58).$$

Полная энергия осциллятора, находящегося в силовом поле, как отмечалось, равна:

$$W = W_k + U, \text{ где } W_k = \frac{mv^2}{2} = \frac{P^2}{2m}.$$

Потенциальная энергия U связана с действующей на частицу силой F соотношением:

$$\vec{F} = -\text{grad}U \quad (5.59).$$

В нашем одномерном случае:

$$F = -\frac{dU}{dx} = -\alpha \cdot x \quad (5.60).$$

Интегрируя (5.60), найдем U :

$$U = \int \alpha \cdot x \cdot dx = \frac{\alpha \cdot x^2}{2} + C \quad (5.61).$$

Полагая $C=0$ (величина константы интегрирования в каждой конкретной задаче определяется начальными условиями), получим:

$$U = \frac{\alpha \cdot x^2}{2} \quad (5.62).$$

Далее, найдем полную энергию осциллятора $W = W_k + U$:

$$W = \frac{p^2}{2m} + \frac{\alpha \cdot x^2}{2} = \frac{mv^2}{2} + \frac{\alpha \cdot x^2}{2} = \frac{m}{2} \left[\left(\frac{dx}{dt} \right)^2 + \frac{\alpha}{m} x^2 \right] \quad (5.63),$$

или:

$$W = \frac{m}{2} \left[\left(\frac{dx}{dt} \right)^2 + \omega^2 x^2 \right] \quad (5.64).$$

Состояние гармонического осциллятора с точки зрения квантовой теории характеризуется волновой функцией ψ , удовлетворяющей уравнению Шредингера:

$$\frac{d^2\psi}{dx^2} + \frac{2m}{h^2} \left[W - \frac{\alpha x^2}{2} \right] \cdot \psi = 0 \quad (5.65)$$

Решение этого уравнения хорошо исследовано. Оно является конечным и однозначным (как отмечалось выше, волновая функция должна отвечать таким условиям) на интервале $-\infty < x < \infty$ при дискретных значениях постоянной W :

$$W = W_n = \left(n + \frac{1}{2} \right) \cdot h\omega \quad (5.66),$$

где $n = 0, 1, 2, \dots$ – любое целое число.

Таким образом, энергия гармонического осциллятора, находящегося в поле потенциальных сил, может принимать только дискретные значения и при изменении его состояния изменяется скачком на величину, кратную энергии кванта $h\omega$. Наименьшая величина энергии гармонического осциллятора равна:

$$W_0 = \frac{1}{2} h\omega \quad (5.67).$$

Она называется нулевой энергией гармонического осциллятора.

Важность полученного результата заключается в том, что любая квантовая система при наличии каких – либо сил, внутренних или внешних, во многих случаях проявляет свойства дискретности ее энергии, т.е. ее

энергия квантуется.

5.9. Взаимодействие излучения с атомными системами

Как отмечалось ранее, при строгом подходе к рассмотрению эффектов взаимодействия электромагнитного поля с веществом, процессы должны рассматриваться как взаимодействие двух квантовых систем: фотонов, образующих поле, и атомов вещества. Поведение и фотонов, и атомов описывается квантовыми законами. Однако во многих случаях возможно использование полуклассического подхода, предполагающего рассмотрение электромагнитного поля с позиций классической физики. В данном разделе будет использоваться именно такой подход.

5.9.1. Энергетические уровни

Энергия свободной частицы, как известно, может принимать любые значения. При движении же частицы в некоторой системе, как мы видели, ее энергия может принимать только дискретные значения, т.е. квантуется. Возможные дискретные значения энергии квантовой системы (система микрочастиц, атомы, молекулы и т.д.), как уже отмечалось, называют энергетическими уровнями. В свободных атомах квантуется энергия электронов и имеется система энергетических уровней электронов, или система электронных уровней атома. Если рассматривать более крупные квантовые системы – молекулы, то можно отметить следующие движения частиц:

- А) движение электронов в атоме;
- Б) колебания атомов в молекуле;
- В) вращение и поступательное движение молекул.

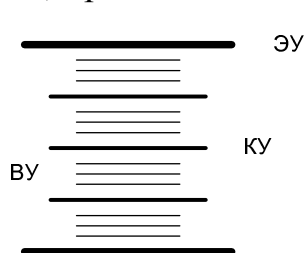


Рис. 5.2.

Из этих видов движения квантуется энергия, связанная с движением электронов внутри атома, колебаниями атомов в молекулах и вращательным движением молекул. В результате квантования образуется система энергетических уровней свободной молекулы (рис. 5.2). Основой системы являются электронные уровни, которые отстоят друг от друга на 1 – 10 эВ. Между электронными уровнями (ЭУ) располагаются колебательные уровни атомов (КУ) с расстоянием $\sim 0,1$ эВ. Между колебательными уровнями находятся вращательные уровни (ВУ) с расстоянием между ними 10^{-3} эВ и менее.

Переходы между электронными уровнями соответствуют излучению в видимом и ультрафиолетовом диапазонах, между колебательными уровнями

– инфракрасному диапазону, а между вращательными уровнями – СВЧ диапазону. Такой вид имеет система энергетических уровней свободных и слабо связанных молекул и атомов. В твердых телах взаимодействие частиц является очень сильным, что приводит к расщеплению отдельных энергетических уровней и даже к образованию практически сплошных зон с очень близко расположенными уровнями. Между разрешенными зонами имеются зоны с запрещенными значениями энергии – запрещенные зоны.

При рассмотрении систем энергетических уровней частицы самый нижний уровень, соответствующий минимальной внутренней энергии, называют основным. Остальные энергетические уровни, соответствующие более высокой внутренней энергии частицы, называют возбужденными. Когда говорят, что частица перешла с одного энергетического уровня на другой, то имеется в виду, что внутренняя энергия частицы изменилась на величину, равную разности энергий данных энергетических уровней. При переходе на более высокий уровень частица поглощает энергию, а при переходе на более низкий – отдает.

Переходы могут быть излучательными (с поглощением или излучением кванта энергии) или безызлучательными. В случае безызлучательных переходов с уменьшением энергии, квант энергии выделяется в виде тепла в среде. При «безызлучательных» переходах с увеличением энергии частица получает эту энергию не от электромагнитного поля, а в результате взаимодействия с другими частицами либо с решеткой в случае твердого тела.

Состояние, из которого запрещены все излучательные переходы в более низкие энергетические состояния, называется метастабильным. Такие состояния играют очень важную роль в квантовых приборах.

5.9.2. Квантовые переходы

При переходах частицы из одного энергетического состояния в другое говорят о спонтанных и вынужденных, или индуцированных, переходах [7].

А). Спонтанные переходы.

Это самопроизвольные квантовые переходы частицы из верхнего энергетического состояния в одно из нижних состояний. Спонтанные переходы могут сопровождаться электромагнитным излучением, частота которого определяется из постулата Бора $h\omega_{ji} = E_j - E_i$, где E_j и E_i – энергия верхнего и нижнего уровней. Очевидно, что:

$$\omega_{ji} = \frac{E_j - E_i}{h} \quad (5.68),$$

и эта частота называется частотой квантового перехода. Число частиц с

одинаковой энергией в единице объема называется населенностью соответствующего энергетического уровня.

Рассмотрим двухуровневую систему и положим $i=1$; $j=2$. Обозначим населенности данных уровней как N_1 и N_2 . При спонтанных переходах происходит уменьшение числа частиц N_2 на верхнем уровне и увеличение их числа N_1 на нижнем уровне. Уменьшение населенности верхнего уровня за время dt пропорционально N_2 и dt :

$$\Delta N_2 = -A_{21}N_2 \cdot dt \quad (5.69),$$

где постоянную A_{21} называют коэффициентом Эйнштейна для спонтанных переходов. Этот коэффициент определяет вероятность спонтанного перехода в 1с. Решение (5.69) имеет вид:

$$N_2(t) = N_2(0) \cdot \exp(-A_{21}t) \quad (5.70),$$

где $N_2(0)$ – населенность верхнего уровня в момент $t=0$. Поскольку общее число частиц в объеме неизменно, то при уменьшении величины N_2 число N_1 возрастет на такую же величину. Из (5.70) следует, что через время $t=1/A_{21}$ населенность N_2 уменьшится в e раз ($e=2,718$) по сравнению с $N_2(0)$. Величина $\tau=1/A_{21}$ характеризует время жизни частицы в возбужденном состоянии и называется временем жизни на уровне. Коэффициент Эйнштейна A_{21} определяет среднее число спонтанных переходов в единице объема в 1 с., т.е. среднее число частиц, совершивших самопроизвольный переход из верхнего состояния в нижнее. Он имеет размерность и может быть любым по величине в отличие от математической вероятности, изменяющейся от 0 до 1.

Число спонтанных переходов в единице объема за 1 с.:

$$n_{21} = A_{21}N_2 \quad (5.71)$$

С учетом (5.70) при спонтанных переходах в системе за 1 с выделяется энергия:

$$P_{21} = n_{21} \cdot h\omega_{21} = A_{21} \cdot N_2 \cdot h\omega_{21} \quad (5.72).$$

Спонтанные переходы случайны, что означает неодновременность и независимость актов излучения при таких переходах. Фазы, направления распространения, поляризация и частота электромагнитных волн при этом не совпадают и спонтанное излучение не когерентно.

В системе частиц с несколькими энергетическими уровнями возможны спонтанные переходы с данного уровня на все нижние уровни. Полная вероятность A_j спонтанного перехода с уровня j на нижние уровни i равна сумме вероятностей отдельных переходов A_{ji} :

$$A_j = \sum_{i=1}^{j-1} A_{ji} = A_{j1} + A_{j2} + \dots + A_{j(j-1)} \quad (5.73).$$

Время жизни частицы на уровне j при этом определяется соотношением:

$$\tau_j = \frac{1}{A_j} \quad (5.74).$$

Б). Индукцированные переходы.

Это квантовые переходы частиц под действием внешнего электромагнитного поля, частота которого совпадает или близка к частоте перехода. При этом возможен как переход с верхнего уровня на нижний, так и с нижнего на верхний.

В первом случае происходит вынужденное испускание кванта с энергией $h\omega_{21}$. Особенность вынужденного испускания заключается в полной идентичности появившегося фотона фотону внешнего поля. Вынужденное излучение имеет ту же частоту, фазу, направление распространения и поляризацию, что и вынуждающее излучение. Таким образом, начальная энергия поля с частотой ω_{21} может увеличиться, что указывает на возможность реализации квантовых усилителей и генераторов. Необходимо отметить, что на вынужденный переход с излучением не затрачивается энергия внешнего поля. Напротив, для перевода частицы в возбужденное состояние 2 необходима энергия внешнего поля $h\omega_{21}$.

Вынужденные переходы, как и спонтанные, имеют статистический характер. Для их описания также вводятся вероятностные коэффициенты:

А) W_{21} – вероятность вынужденного перехода сверху вниз.

Б) W_{12} – вероятность такого перехода снизу вверх.

Эти коэффициенты также определяют вероятность перехода частицы с одного уровня на другой за одну секунду. Вероятности индуцированных переходов пропорциональны объемной плотности энергии внешнего поля U_ω в единичном спектральном интервале на частоте перехода и определяются соотношениями:

$$W_{21} = B_{21}U_\omega; \quad W_{12} = B_{12}U_\omega \quad (5.75).$$

Здесь B_{21} и B_{12} – коэффициенты Эйнштейна для вынужденных переходов с излучением и поглощением энергии, соответственно. Они имеют смысл вероятности вынужденных переходов в 1 с при единичной объемной плотности энергии внешнего поля ($U_\nu = \frac{1 \text{ Дж} \cdot \text{см}^{-3}}{c}$).

Число вынужденных переходов сверху вниз с излучением энергии в единицу времени в единице объема пропорционально величине W_{21} и населенности N_2 :

$$n_{21(\epsilon)} = W_{21}N_2 = B_{21}U_\omega N_2 \quad (5.76).$$

Аналогично, число вынужденных переходов с поглощением энергии в единицу времени в единичном объеме определяется соотношением:

$$n_{12(\epsilon)} = W_{12}N_1 = B_{12}U_\omega N_1 \quad (5.77).$$

Поскольку в квантовых системах при наличии внешнего электромагнитного поля резонансной частоты происходят вынужденные переходы как с излучением, так и с поглощением энергии, то в таких

системах в принципе возможно как ослабление, так и усиление внешнего поля. Условия, необходимые для реализации эффекта квантового усиления, будут рассмотрены далее.

5.9.3. Соотношения между коэффициентами Эйнштейна

Эйнштейн получил соотношения связи между коэффициентами A_{ij} и B_{ij} , рассматривая систему с двумя уровнями в состоянии термодинамического равновесия [7]. При этом число излученных квантов в системе должно быть равно числу поглощенных, т.е. должно выполняться условие $n_{21}=n_{12}$.

В состоянии термодинамического равновесия в системе имеется равновесная плотность поля излучения U_ω , определяющая число вынужденных переходов. Полное число переходов сверху вниз n_{21} :

$$n_{21} = n_{21(c)} + n_{21(e)} = (A_{21} + B_{21}U_\omega)N_2 \quad (5.78).$$

Число переходов n_{12} с нижнего уровня на верхний (только вынужденные):

$$n_{12} = n_{12(e)} = B_{12}U_\omega N_1 \quad (5.79).$$

Приравнявая n_{21} и n_{12} , получим:

$$B_{12}U_\omega N_1 = (A_{21} + B_{21}U_\omega)N_2 \quad (5.80).$$

Отсюда находим U_ω :

$$U_\omega = \frac{\frac{A_{21}}{B_{21}}}{\frac{B_{12}N_1}{B_{21}N_2} - 1} \quad (5.81).$$

В состоянии термодинамического равновесия соотношение населенностей уровней определяется законом Больцмана:

$$\frac{N_2}{N_1} = \frac{g_2}{g_1} \cdot \exp\left(-\frac{E_2 - E_1}{kT}\right) \quad (5.82),$$

где g_2, g_1 – статистические веса энергетических уровней; k – постоянная Больцмана; T – абсолютная температура. В случае отсутствия вырождения энергетических уровней $g_2=g_1=1$, тогда получаем:

$$\frac{N_2}{N_1} = \exp\left(-\frac{E_2 - E_1}{kT}\right) \quad (5.83).$$

Подставляя (5.83) в (5.81), с учетом равенства $E_2 - E_1 = h\omega_{21}$, получим:

$$U_\nu = \frac{\frac{A_{21}}{B_{21}}}{\frac{B_{12}}{B_{21}} \cdot \exp\left(\frac{h\omega_{21}}{kT}\right) - 1} \quad (5.84).$$

Эйнштейн постулировал, что равновесная спектральная плотность энергии должна быть равна ее значению, найденному по формуле Планка для равновесного излучения абсолютно черного тела:

$$U_\nu = \frac{8\pi\nu^2}{c^3} \frac{1}{\exp\left(\frac{h\nu}{kT}\right) - 1} \quad (5.85).$$

Отсюда следует:

$$B_{12} = B_{21} \quad (5.86),$$

$$A_{21} = \frac{8\pi h \nu_{21}^3}{c^3} \cdot B_{21} \quad (5.87).$$

Необходимо отметить сильную частотную зависимость коэффициента Эйнштейна по спонтанным переходам $A_{21} \sim \nu^3$, которая указывает на существенную роль спонтанного излучения в оптическом диапазоне.

5.10. Форма и ширина спектральной линии

5.10.1. Естественная ширина спектральной линии.

До сих пор, записывая правило Бора для переходов квантовой системы из одного энергетического состояния в другое, мы полагали, что энергетические уровни являются бесконечно узкими, т.е. таким переходам соответствуют строго определенные частоты ω_{ij} . Однако на самом деле ширина уровней конечна [7]. Реально излучение для перехода не является строго монохроматическим, а может быть представлено некоторым спектром частот.

Спектральной линией излучения называют зависимость интенсивности излучения (или поглощения) от частоты при переходе квантовой системы из одного состояния в другое. Графическое изображение этой зависимости называют контуром спектральной линии (рис. 5.3). Шириной спектральной линии считают разность частот, на которых интенсивность излучения или поглощения падает вдвое по сравнению с ее максимальным значением. Частота перехода (центральная) это частота, соответствующая максимуму спектральной линии.

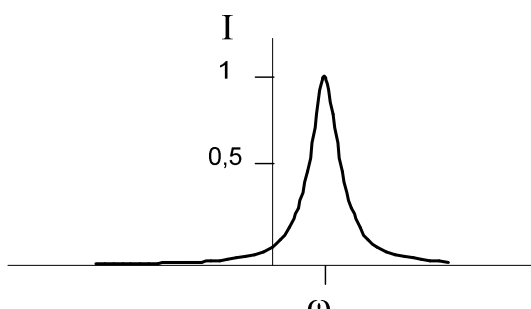


Рис. 5.3. Контур спектральной линии

Уже отмечено, что минимальная ширина уровня (и, естественно, спектральной линии) определяется соотношением неопределенностей Гейзенберга в форме:

$$\Delta E \cdot \Delta t \geq h \quad (5.88),$$

где ΔE и Δt – неопределенности энергии и времени. Величина Δt

считается равной τ - времени жизни электрона в возбужденном состоянии. Таким образом, $\Delta E_j \geq h/\tau_j$, где время жизни частицы в возбужденном состоянии τ_j определяется вероятностью спонтанных переходов с данного уровня на все нижние. Соответственно, поскольку частица или система, находясь на основном энергетическом уровне, не может за счет спонтанных переходов перейти на любой другой, то этот уровень имеет бесконечно малую ширину.

Неопределенность частоты перехода между двумя размытыми уровнями j и i ($j>i$) определяется соотношением:

$$\delta\omega_{\omega_{max}-\omega_{min}} = \frac{\Delta E_j + \Delta E_i}{h} \quad (5.89),$$

где ΔE – ширина отдельных уровней.

Под естественной шириной спектральной линии понимают ширину спектральной линии изолированной и неподвижной квантовой системы. Она минимальна и определяется временем жизни по спонтанному излучению.

Можно с помощью достаточно простых рассуждений получить выражение для формы спектральной линии. Для этого рассмотрим двухуровневую квантовую систему, в которой в начальный момент времени на верхнем уровне находится N_{02} частиц. Число частиц, переходящих с уровня E_2 на уровень E_1 , убывает во времени по закону $N_2(t)=N_{02} \cdot \exp(-A_{21} \cdot t)$. Спонтанный переход каждой частицы на нижний уровень сопровождается излучением кванта энергии, поэтому энергия излучения убывает во времени по этому же закону, т.е.:

$$P_2(t) = P_0 \cdot \exp(-A_{21} \cdot t) \quad (5.90).$$

Напряженность поля излучения изменяется при этом в соответствии с выражением:

$$E(t) = (2P_2)^{1/2} \cdot \cos(\omega_0 t) = E_0 \cdot \exp(-A_{21} t) \cdot \cos(\omega_0 t) \quad (5.91),$$

где $E_0 = (2P_2)^{1/2}$. Таким образом, излучение может быть представлено в виде затухающего гармонического колебания $E(t)$, амплитудный спектр которого $g(\omega)$ можно найти с помощью интеграла Фурье:

$$g(\omega) = \int_{-\infty}^{\infty} E(t) \cdot \exp(-i\omega t) dt \quad (5.92).$$

Подставляя сюда (5.91) и проводя интегрирование, получим:

$$g(\omega) = \frac{E_0}{\sqrt{(\omega - \omega_0)^2 + (A_{21}/2)^2}} \quad (5.93).$$

Спектр мощности определяется как:

$$G(\omega) = \frac{2P_0}{(\omega - \omega_0)^2 + (A_{21}/2)^2} \quad (5.94).$$

Контур линии может трактоваться, как контур плотности вероятности поглощения или излучения фотона с энергией в интервале от $h\omega$ до $h(\omega + d\omega)$. В этом случае должно выполняться условие нормировки:

$$\int_0^{\infty} G(\omega) d\omega = 1 \quad (5.95).$$

Отсюда можно получить окончательное выражение для контура спектральной линии:

$$G(\omega) = \frac{A_{21}/2}{(\omega - \omega_0)^2 + (A_{21}/2)^2} \quad (5.96).$$

Как видим, данное выражение аналогично описывающему частотную характеристику резонансного контура. Это кривая Лоренца.

5.10.2. Уширение спектральной линии из-за столкновений.

Реальные спектральные линии имеют ширину больше естественной. Причины этого уширения различны. Первая – упругие и неупругие столкновения частиц в газе (о газе говорим, поскольку его атомы или молекулы в первом приближении могут считаться уединенными и не взаимодействующими). Упругими называют процессы взаимодействия, в которых суммарная кинетическая энергия сталкивающихся частиц не изменяется (частицы не обмениваются внутренней энергией и не переходят на другие уровни). В процессе неупругих столкновений суммарная кинетическая энергия сталкивающихся частиц либо возрастает (прямые соударения или соударения первого рода), либо убывает (обратные соударения или соударения второго рода). В данном случае изменяется внутренняя энергия сталкивающихся частиц, связанная с электронными, колебательными и вращательными уровнями.

В результате неупругих столкновений происходит сокращение времени жизни частиц в данном энергетическом состоянии до времени среднего пробега между двумя столкновениями. Результат – уширение спектральной линии. Вероятность столкновения частиц и, соответственно, вероятность квантовых переходов одинаковы для всех частиц газа. Поэтому форма и ширина спектральной линии всего газа и каждой частицы одинаковы (и больше естественной ширины). Такой процесс называют однородным уширением спектральной линии.

Однородное уширение наблюдается и в случае столкновения частиц со стенками, ограничивающими объем газа. Эти столкновения существенно влияют на форму линии при малом давлении газа, когда вероятность столкновения частиц мала.

5.10.3. Допплеровское уширение спектральной линии

Линия называется неоднородно уширенной, если она представляет собой суперпозицию нескольких неразрешенных спектральных линий. Пример – доплеровское уширение. Оно наблюдается в газовых средах. Вследствие хаотического движения частиц газа, излучение частицы, движущейся в направлении глаза наблюдателя со скоростью v , имеет доплеровское смещение частоты на величину $\omega_0 \cdot (v/c)$, где ω_0 – частота излучения покоящейся частицы (в нашем случае она совпадает с частотой перехода).

Поскольку линия излучения имеет естественное уширение, излучение движущейся частицы представляет собой спектральную линию, сдвинутую по частоте на $\omega_0 \cdot (v/c)$. В газе находятся частицы с различными скоростями движения, поэтому сдвиги по частоте, испытываемые излучением каждой частицы, будут различными. При этом форма линии излучения газа в целом определяется функцией распределения излучающих частиц по скоростям $f(v)$, так что форма спектральной линии будет иметь вид:

$$g(\omega)d\omega = f(v)dv \quad (5.97).$$

Частота излучателя, движущегося со скоростью v в направлении наблюдателя, равна $\omega_0 + \omega_0 \cdot (v/c)$. Выразим отсюда величины v и dv через ω и ω_0 :

$$v = \frac{\omega - \omega_0}{\omega_0} \cdot c, \quad dv = c \cdot \frac{d\omega}{\omega_0}$$

Подставляя эти выражения в (5.94), получим:

$$g\omega \cdot d\omega = f\left[\frac{\omega - \omega_0}{\omega_0} \cdot c\right] \cdot c \cdot \frac{d\omega}{\omega_0} \quad (5.98).$$

В газе при термодинамическом равновесии распределение частиц по скорости подчиняется закону Максвелла - Больцмана:

$$f(v) \cdot dv = \frac{I}{\sqrt{\pi}} \cdot \exp\left[-\left(\frac{v}{v_0}\right)^2\right] \cdot \frac{dv}{v_0} \quad \text{где } v_0 = \sqrt{\frac{2kT}{m}} \text{ – среднее значение скорости}$$

частиц в газе. Соответственно,

$$f(v) \cdot dv = \frac{I}{\sqrt{\pi}} \cdot \exp\left[-\left(\frac{\omega - \omega_0}{\omega_0}\right)^2\right] \cdot \frac{d\omega}{\Delta\omega_D} \quad (5.99),$$

где $\Delta\omega_D = \omega_0 \cdot \frac{v}{c}$.

Линия с таким профилем называется доплеровски уширенной линией. Ее форма симметрична относительно частоты $\omega = \omega_0$. Ширина этой линии определяется параметром $\Delta\omega_D$. При $(\omega - \omega_0) = \Delta\omega_D$ интенсивность линии

уменьшается в e раз по сравнению с максимальным значением. Полуширина доплеровски уширенной линии:

$$\Delta\omega_D / 2 = (\ln 2)^2 \cdot \Delta\omega_D \quad (5.97).$$

5.10.4. Другие механизмы уширения спектральных линий

В твердых веществах уширение спектральных линий обусловлено иными механизмами, чем в газах. Так, колебания кристаллической решетки модулируют внутрикристаллическое электрическое поле в месте расположения активного иона. Это приводит к модуляции энергетических уровней иона и уширению спектральной линии. Ширина линии увеличивается также из-за тепловых колебаний самих ионов. Пространственная неоднородность физических параметров среды или неоднородности электрического и магнитного полей также могут приводить к уширению спектральных линий в твердотельных средах.

5.11. Отрицательная температура

Возможность усиления и генерации электромагнитных колебаний в квантовых системах определяется возможностью индуцированных переходов в среде под воздействием электромагнитного поля. Как уже отмечалось, испускаемые в процессе индуцированных переходов фотоны полностью идентичны индуцирующим, что и может привести к усилению индуцирующего электромагнитного излучения. Однако при взаимодействии квантовой системы с электромагнитным полем происходят и переходы с поглощением энергии, а также возможны релаксационные, безызлучательные переходы. Вынужденные переходы с излучением и поглощением энергии имеют резонансный характер, т.к. происходят при совпадении частоты внешнего поля с частотой квантового перехода (точнее, когда частота поля находится в пределах спектральной линии квантовой системы).

Рассмотрим обмен энергии между полем и простейшей двухуровневой квантовой системой. Пусть двум энергетическим уровням E_1 и E_2 соответствуют населенности N_1 и N_2 , а частота внешнего поля $\omega_{\text{вн}}$ удовлетворяет условию $\omega_{\text{вн}} = \omega_{12}$, где ω_{12} – центральная частота спектральной линии.

Число вынужденных переходов с излучением энергии в единице объема, как отмечалось, равно $n_{21} = B \cdot U_{\nu} \cdot N_2$. Мощность электромагнитного излучения $P_{\text{выд}}$, выделяемая при этом в единице объема, равна:

$$P_{\text{выд}} = n_{21} h \omega_{12} = B \cdot U_{\nu} \cdot N_2 h \omega_{12}.$$

Аналогично, число вынужденных переходов с поглощением энергии и поглощаемая в единице объема мощность внешнего электромагнитного поля $P_{\text{погл}}$ равны:

$$\begin{aligned} n_{12} &= B \cdot U_{\nu} \cdot N_1 \\ P_{\text{погл}} n_{21} \cdot h\omega_{12} &= BU_{\nu} N_1 h\omega_{12} \end{aligned}$$

Очевидно, что изменение мощности внешнего электромагнитного поля определяется разностью выделяемой и поглощаемой мощностей:

$$P = P_{\text{выд}} - P_{\text{погл}} = BU_{\nu} h\omega_{12} (N_2 - N_1) \quad (5.98).$$

Величину P принято называть мощностью взаимодействия. Если $P > 0$, то в рассматриваемой квантовой системе происходит усиление внешнего электромагнитного поля, а отрицательная ее величина $P < 0$ говорит об ослаблении этого поля. Следовательно, условием усиления являются соотношения:

$$(N_2 - N_1) > 0 \quad \text{или} \quad \frac{N_2}{N_1} > 1 \quad (5.99).$$

В обычном состоянии термодинамического равновесия в соответствии с законом Больцмана $N_m = N_0 \cdot \exp\left(-\frac{E}{kT}\right)$ выполняется условие $N_2 < N_1$ или:

$$\frac{N_2}{N_1} = \exp\left[-\frac{(E_2 - E_1)}{kT}\right] \quad (5.100).$$

Это выражение определяет соотношение населенностей энергетических уровней, однако формально оно может рассматриваться и как соотношение, определяющее значение абсолютной температуры по величинам населенностей разных энергетических уровней системы, находящейся в состоянии термодинамического равновесия. Несложными преобразованиями приведем его к виду:

$$T_n = \frac{E_2 - E_1}{k \cdot \ln(N_1 / N_2)} = \frac{h\omega_{21}}{k \cdot \ln(N_1 / N_2)} \quad (5.101),$$

где величину T_n называют температурой перехода. В состоянии термодинамического равновесия $N_1 > N_2$, т.е. $T_n > 0$. Состояние с $N_2 > N_1$ является обратным, т.е. инверсным. Поэтому состояние с $N_2 > N_1$, когда в квантовой системе возможно усиление электромагнитного излучения, называется состоянием с инверсией населенности или просто инверсией населенности. При $N_2 > N_1$, из (5.101) видим, что $T_n < 0$. Поэтому состояние с инверсией населенностей уровней иногда называют состоянием с отрицательной температурой. Таким образом, условие усиления в квантовой системе можно выразить и в форме: $T_n < 0$. Среда, в которой возможна инверсия населенности или состояние с отрицательной температурой, называется также активной средой.

Следует отметить, что температура перехода совпадает с реальной температурой только в состоянии термодинамического равновесия. Условие $N_2 > N_1$, при котором $T_n < 0$, конечно, вовсе не означает, что абсолютная температура среды действительно меньше нуля.

В состоянии термодинамического равновесия $N_1 > N_2$, поэтому $n_{12} > n_{21}$. Соответственно, величина N_1 при наличии резонансного внешнего поля убывает, а N_2 – возрастает. При достаточно большом значении U_v может произойти выравнивание населенностей уровней, т.е. может быть достигнуто условие $N_2 = N_1$. При этом число переходов частиц с излучением и поглощением энергии уравнивается ($n_{12} = n_{21}$) и наступает динамическое равновесие. Данный режим называется насыщением перехода. Очевидно, что дальнейшее увеличение плотности мощности электромагнитного поля, т.е. величины U_v , все равно не может привести к инверсии населенностей.

Итак, при воздействии на двухуровневую квантовую систему электромагнитного поля можно достичь насыщения перехода, но не инверсии населенностей.

Для более строгого анализа взаимодействия электромагнитного поля с двухуровневой квантовой системой необходимо, как отмечено, учитывать также наличие спонтанных и безызлучательных переходов. Величины населенностей уровней при этом могут быть найдены из решения кинетических уравнений. Для двухуровневой системы эти уравнения имеют вид:

$$\frac{dN_1}{dt} = -N_1 B U_v - N_1 w_{12} + N_2 B U_v + N_2 A_{21} + N_2 w_{21} \quad (5.102)$$

$$\frac{dN_2}{dt} = N_1 B U_v + N_1 w_{12} - N_2 B U_v - N_2 A_{21} - N_2 w_{21} \quad (5.103)$$

$$N_1 + N_2 = N \quad (5.104),$$

где N – полное число частиц; w_{12} , w_{21} – вероятности безызлучательных переходов с первого уровня на второй и со второго на первый. Очевидно, что при сохранении полного числа частиц N для двухуровневой системы

$$\frac{dN_1}{dt} = -\frac{dN_2}{dt} \quad (5.105).$$

Для стационарного состояния системы имеем: $\frac{dN_1}{dt} = \frac{dN_2}{dt} = 0$ и тогда из

(5.103, 5.104) следует:

$$\begin{cases} N_1 (B U_v + w_{12}) - N_2 (B U_v + A_{21} + w_{21}) = 0 \\ N_1 + N_2 = N \end{cases} \quad (5.106).$$

Решая систему двух этих уравнений, можно получить:

$$N_1 (B U_v + w_{12}) + N_2 (B U_v + A_{21} + w_{21}) - N (B U_v + A_{21} + w_{21}) = 0$$

или

$$N_1 = \frac{N(BU_v + A_{21} + w_{21})}{2BU_v + A_{21} + w_{12} + w_{21}} = \frac{N}{A_{21} + w_{12} + w_{21}} \cdot \frac{A_{21} + w_{21} + BU_v}{I + \delta_{12} \cdot U_v} \quad (5.107).$$

где $\delta_{12} = \frac{2B}{A_{21} + w_{12} + w_{21}}$.

Аналогично:

$$N_2 = \frac{N}{A_{21} + w_{12} + w_{21}} \cdot \frac{w_{12} + BU_v}{I + \delta_{12} \cdot U_v} \quad (5.108).$$

Разность и отношение этих величин даются выражениями:

$$N_2 - N_1 = \frac{N \cdot [w_{12} - (A_{21} + w_{21})]}{A_{21} + w_{12} + w_{21}} \cdot \frac{I}{I + \delta_{12} \cdot U_v} \quad (5.109).$$

$$\frac{N_1}{N_2} = \frac{A_{21} + w_{21} + BU_v}{w_{12} + BU_v} \quad (5.110).$$

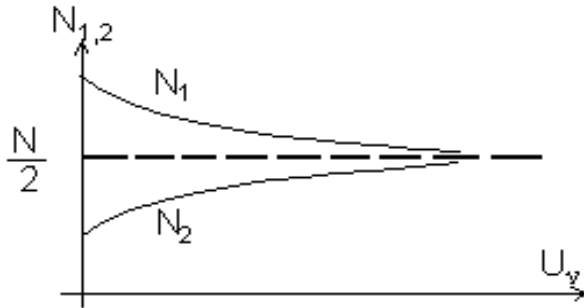


Рис. 5.4. Зависимости $N_{1,2}(U_v)$ при начальном термодинамическом равновесии системы.

Из приведенных выше соотношений вытекает, что при малых U_v величины N_1 и N_2 изменяются по линейному закону, при больших – система стремится к насыщению. Это иллюстрируют кривые на рис. 5.4.

Используя полученное соотношение (5.109), мощность взаимодействия, введенная ранее, может быть выражена в

виде:

$$P = h\omega_{21} \cdot BU_v (N_2 - N_1) = h\omega_{21} \cdot B \cdot N \cdot \frac{w_{12} - (A_{21} + w_{21})}{A_{21} + w_{12} + w_{21}} \cdot \frac{U_v}{I + \delta_{12} \cdot U_v} \quad (5.111).$$

Очевидно, что вид зависимости $P(U_v)$ определяется множителем $U_v \cdot (1 + \delta_{12} \cdot U_v)^{-1}$. Предельное значение этой мощности $P_{пред}$ найдем из условия $P_{пред} = P(U_v = \infty)$:

$$P_{пред} = 0,5h\omega_{21}N[w_{12} - (A_{21} + w_{21})] \quad (5.112).$$

Таким образом, в режиме насыщения от электромагнитного поля отбирается мощность $P_{пред}$. Она необходима для поддержания равенства населенностей уровней, которое постоянно стремится нарушиться из-за спонтанных и безызлучательных переходов с вероятностями A_{21} , w_{12} , w_{21} .

5.12. Взаимодействие бегущих электромагнитных волн с активной средой

Рассмотрим случай, когда плоская электромагнитная волна падает на активную среду и распространяется в ней в направлении оси Z (рис. 5.5).

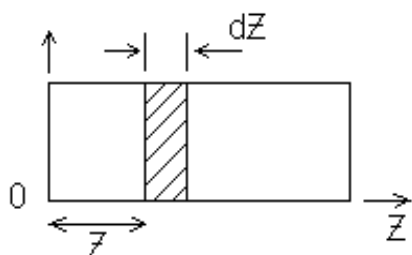


Рис. 5.5.

Пусть S – поток энергии через единичное поперечное сечение образца. Величина S изменяется при прохождении излучения через среду, т.к. в ней происходят вынужденные переходы с поглощением и выделением энергии. Изменение dS в слое dz за 1с, т.е. изменение мощности, определяется соотношением:

$$dS = Bh\omega_{21}U_v(N_2 - N_1)dz \quad (5.113).$$

Поток энергии связан с объемной плотностью энергии и групповой скоростью соотношением $S = v_g U_v$. Тогда из (5.113) следует, что

$$dS = Bh\omega_{21} \cdot \frac{S}{v_g} \cdot (N_2 - N_1)dz \quad \text{или} \quad \frac{dS}{S} = \frac{Bh\omega_{21}}{v_g} \cdot (N_2 - N_1)dz. \quad \text{Вводя}$$

обозначение: $-\kappa = \frac{Bh\omega_{21}}{v_g}(N_2 - N_1)$, приходим к уравнению вида

$$\frac{dS}{S} = -\kappa \cdot dz. \quad \text{Это уравнение известно в оптике как дифференциальный закон}$$

Бугера. Коэффициент κ называется коэффициентом поглощения.

Сначала рассмотрим случай, когда величина U_v очень мала. Это позволяет пренебречь изменением N_1 и N_2 . Тогда $(N_2 - N_1)$ в приведенных соотношениях не зависит от dz и для S получим:

$$S(z) = S(0) \cdot \exp(-\kappa z) \quad (5.114),$$

где $S(0)$ – плотность потока энергии в начале образца. Формула (5.114) представляет собой интегральный закон Бугера.

При $N_1 - N_2 > 0$ (при отсутствии инверсии населенностей) $\kappa > 0$ и энергия затухает, а κ представляет собой коэффициент поглощения. При $N_1 - N_2 < 0$ (при инверсии населенностей) $\kappa < 0$, поэтому (5.114) описывает усиление потока энергии. Введем для этого случая величину κ_a , характеризующую коэффициент усиления активной среды:

$$\kappa_a = -\kappa = \frac{Bh\omega_{21}}{v_g} \cdot (N_2 - N_1) \quad (5.115).$$

Тогда закон Бугера можно записать в виде:

$$S(z) = S(0) \cdot \exp(\kappa_a \cdot z)$$

Для учета потерь энергии в среде, связанных с поглощением и рассеянием на неоднородностях, введем показатель потерь α (считая, что

потери распределены в среде однородно): $\frac{dS(z)}{S} = [\kappa_a(\omega_0) - \alpha] \cdot S(z)$. Для малой величины U_v решение этого уравнения опять можно записать в виде $S(z) = S(0) \cdot \exp[(\kappa_a - \alpha)z]$. При большой мощности U_v величины N_1, N_2 и, следовательно, κ_a зависят от z . Поэтому, с учетом (5.109):

$$\kappa_a = \frac{Bh\omega_{21}N}{v_z} \cdot \frac{w_{12} - (A_{21} + w_{21})}{w_{12} + A_{21} + w_{21}} \cdot \frac{I}{I + \delta_{12} \cdot U_v} \quad (5.116).$$

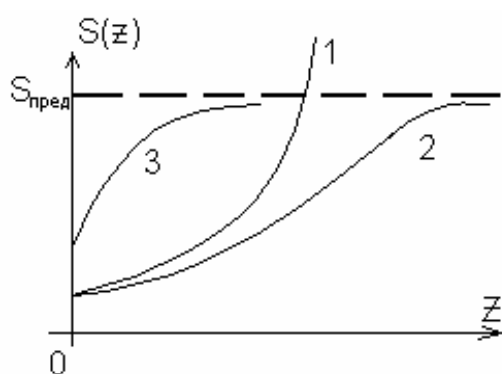
Обозначив $\kappa_a(U_v = 0) = \kappa_{a0}$, получим $\kappa_a = \frac{\kappa_{a0}}{I + \delta_{12} \cdot U_v}$. Как уже отмечалось,

$$U_v = \frac{S}{v_z}, \text{ тогда } \kappa_a = \frac{\kappa_{a0}}{I + \delta_{12} \cdot \frac{S(z)}{v_z}}.$$

В активной среде S растет с увеличением длины пути (координаты z), следовательно, κ_a при этом уменьшается. Тогда при некотором z может быть выполнено условие $\kappa_a = \alpha$, т.е. величины коэффициента усиления и показателя потерь сравниваются. В этом случае $\frac{dS}{dz} = 0$, т.е. мощность ЭМ волны достигает своего предельного значения. Найдем $S_{пред}$ из условия:

$$\kappa_a = \frac{\kappa_{a0}}{I + \delta_{12} \cdot U_v} = \alpha:$$

$$S_{пред} = \frac{\kappa_{a0} - \alpha}{\alpha \cdot \frac{\delta_{12}}{v_z}} = \frac{v_z (\kappa_{a0} - \alpha)}{\alpha \cdot \delta_{12}} \quad (5.117).$$



Величина $S_{пред}$ существенно зависит от α . При $\alpha \rightarrow 0$ $S_{пред} \rightarrow \infty$. Из соотношения (5.117) следует, что $S_{пред}$ не зависит от величины входной мощности $S(0)$. Это означает возможность создания квантовых генераторов излучения. При достаточно большой протяженности активной среды даже очень слабое начальное излучение приведет к достижению предельной мощности. Рис.

5.6 демонстрирует поведение $S(z)$ при разной начальной величине $S(0)$. Кривая 1 соответствует случаю $\alpha = 0$, т.е. случаю отсутствия потерь. Кривая 2 характеризует случай $\alpha \neq 0$ при той же величине $S(0)$, как для зависимости 1. Зависимость 3 соответствует большей величине $S(0)$. Видим, что в этом случае $S_{пред}$ достигается на меньшей длине.

Рис. 5.6.

5.13. Методы достижения инверсии населенностей

В результате анализа в п. 5.11 показано, что в двухуровневой квантовой системе, при ее взаимодействии с ЭМП, увеличение плотности энергии электромагнитного поля в стационарном режиме может привести лишь к режиму насыщения, но не к состоянию с инверсией населенностей. Для достижения такого состояния предложено и используется несколько методов. Суть их, при необходимости получения стационарного состояния с инверсией населенностей, заключается в использовании трех- и четырехуровневых квантовых систем. Поясним ее на примере системы с тремя уровнями (рис. 5.7).

Пусть трем уровням соответствуют энергии W_1, W_2, W_3 ($W_1 < W_2 < W_3$) и населенности N_1, N_2 и N_3 ($N_1 > N_2 > N_3$), соответственно. Для достижения инверсии населенностей используется источник накачки (вспомогательного излучения). Частота поля накачки соответствует переходам частиц между уровнями 1 – 3 и определяется условием $\omega_{13} = \frac{W_3 - W_1}{h}$. В состоянии термодинамического равновесия соотношение населенностей определяется законом Больцмана: $\frac{N_i}{N_j} = \exp\left(-\frac{W_i - W_j}{kT}\right)$. Рассматриваем радиодиапазон, для которого можно положить $(W_i - W_j)/kT \ll 1$. В этом случае, разлагая экспоненту в ряд, получим следующие соотношения для населенностей уровней в состоянии термодинамического равновесия:

$$\frac{N_{02}}{N_{01}} = 1 - \frac{W_2 - W_1}{kT} \quad (5.118),$$

$$\frac{N_{03}}{N_{01}} = 1 - \frac{W_3 - W_1}{kT} \quad (5.119).$$

При наличии накачки, в условиях насыщения населенности уровней 1 и 3 выравниваются и тогда имеем:

$$N_1 = N_3 = N_{10} \left(1 - \frac{W_3 - W_1}{2kT}\right) \quad (5.120).$$

При этом возможно достижение условий $N_2 > N_1$ или $N_3 > N_2$, в зависимости от начальной населенности уровня 2, что иллюстрируется диаграммами на рисунке 5.7.

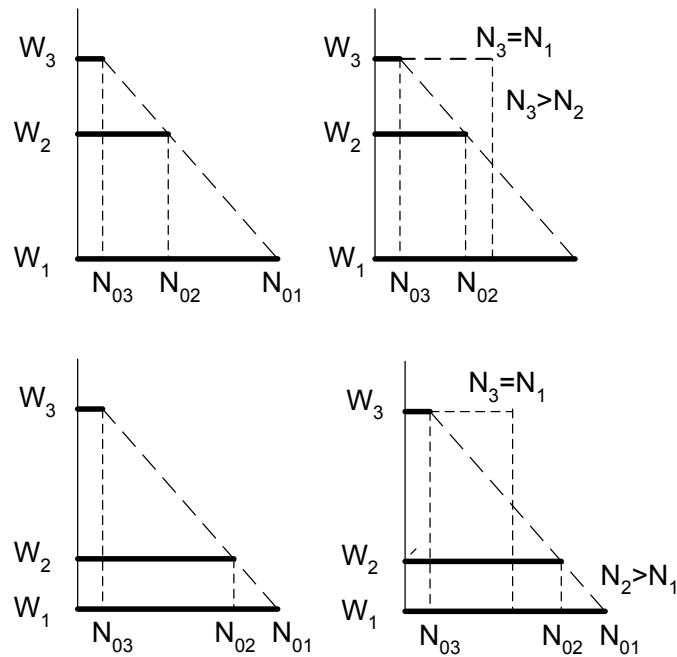


Рис. 5.7. Иллюстрация возможности достижения состояния с инверсией населенностей в трехуровневой системе. Верхний ряд: $W_2 > (W_1 + W_3)/2$; нижний ряд: $W_2 < (W_1 + W_3)/2$.

В четырехуровневой схеме накачки реализуется практически та же самая идея, в том числе и на оптических частотах. Из других методов создания инверсии населенностей можно отметить метод пространственного разделения частиц с разной энергией, реализованный в первом квантовом генераторе на молекулах аммиака - мазере. Следует также отметить, что импульсная накачка позволяет достичь состояния с инверсией населенностей даже в двухуровневой системе.

6. ПРИНЦИПЫ ПОСТРОЕНИЯ ЛАЗЕРОВ

6.1. ОТКРЫТЫЕ ОПТИЧЕСКИЕ РЕЗОНАТОРЫ

Процессы взаимодействия электромагнитного излучения с квантовыми системами на оптических частотах не отличаются принципиально от таковых для СВЧ диапазона. Однако значительно более высокие частоты оптического излучения приводят к ряду существенных особенностей [7, 8].

1. В системе энергетических уровней, разделенных оптическими частотами, согласно закону Больцмана, в условиях теплового равновесия практически все частицы находятся на основном уровне. В трехуровневой системе $N_{20}=N_{30}=0$; $N_{10}=N$. Очевидно, в подобном случае можно легко достичь инверсии населенностей для пары уровней 3 – 2, если накачка происходит с первого уровня на третий. Состояние инверсии населенностей для пары уровней 1 – 2 получить значительно сложнее.

2. На оптических частотах возрастает роль спонтанных переходов и безызлучательных переходов. Действительно, как отмечалось, величина коэффициента Эйнштейна для спонтанных переходов пропорциональна кубу частоты перехода. Соотношение между вероятностями таких переходов зависит как от вещества, так и от рассматриваемых энергетических уровней.

3. В оптическом диапазоне нет мощных монохроматических источников накачки. Поэтому используются либо немонахроматические источники света, либо другие методы накачки. Подробнее этот вопрос будет рассмотрен при рассмотрении различных типов лазеров.

6.1.1. Элементарная теория открытых резонаторов

Генераторы электромагнитного излучения оптического диапазона, использующие эффект индуцированного излучения, называются лазерами. Слово LASER – это аббревиатура словосочетания **L**ight **A**mplification by **S**timulated **E**mission of **R**adiation.

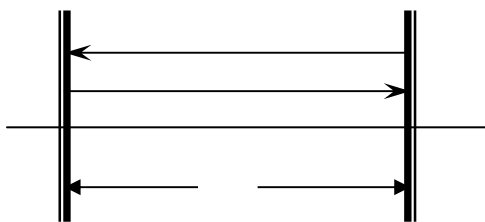
Важнейшими элементами лазеров, как и любых других генераторов электромагнитных колебаний, являются резонаторы. В радиодиапазоне, как известно, в качестве резонаторов используются замкнутые полости пространства. Резонаторы лазеров не имеют боковых поверхностей и называются открытыми.

В радиодиапазоне объемные резонаторы имеют размеры, сравнимые с длиной волны излучения. На более низких частотах вместо резонаторов используют колебательные системы на сосредоточенных элементах. Использование объемных резонаторов в области высоких частот ограничивается миллиметровым диапазоном длин волн, т.к. дальнейшее

увеличение частот приводит к технологическим сложностям реализации объемных резонаторов. Кроме того, пропорциональное уменьшение длины волны и размеров резонатора приводит к существенному уменьшению его добротности.

С другой стороны, использование объемных резонаторов с размерами, много большими длины волны, в оптическом диапазоне приводит к сильному сгущению спектра собственных частот резонатора и перекрытию резонансных кривых отдельных типов колебаний. В результате резонатор теряет свои резонансные свойства. Выходом явилось использование открытых резонаторов. Их размеры много больше длины волны, а спектр собственных частот достаточно разрежен.

Простейший открытый резонатор (резонатор Фабри – Перо) состоит из двух плоских параллельных зеркал, расположенных на некотором расстоянии друг от друга перпендикулярно оси, соединяющей их.



Рассмотрим основные особенности поля в таком резонаторе [7], для чего предполагаем зеркала бесконечно протяженными.

Собственные типы колебаний (моды) открытого резонатора могут рассматриваться как результат интерференции световых волн, распространяющихся от одного зеркала к другому. В результате такой интерференции в резонаторе возникают стоячие волны.

Если световые волны распространяются вдоль оси резонатора z , то стоячая волна образуется, когда на расстоянии L между зеркалами укладывается целое число полуволн. Это условие можно записать в виде:

$$L = q \frac{\lambda}{2} \quad (6.1),$$

где q – целое число, а λ – длина волны. Очевидно, что собственная частота такой стоячей волны равна:

$$\nu_q = \frac{c}{\lambda} = \frac{c}{2L} q \quad (6.2).$$

Расстояние по частоте между соседними модами:

$$\Delta\nu = \nu_q - \nu_{q+1} = \frac{c}{2L} q - \frac{c}{2L} (q-1) = \frac{c}{2L} \quad (6.3).$$

Стоячие волны или моды резонатора, образованные волнами, распространяющимися вдоль его оптической оси, называют **продольными** или **аксиальными** модами.

Некоторые численные оценки:

Пусть $L = 30$ см; $\lambda = 0,6$ мкм. Тогда
$$q = \frac{2L}{\lambda} = \frac{60}{0,6 \cdot 10^{-4}} = 10^6.$$

Таким образом, на длине резонатора укладывается $\sim 10^6$ полуволн. В радиодиапазоне это число составляет $1 \div 3$.

Расстояние между модами:
$$\Delta\nu = \frac{c}{2L} = \frac{3 \cdot 10^{10}}{60} = 5 \cdot 10^8 = 500 \text{ МГц}$$

Для разных типов лазеров величина L меняется от долей миллиметра (полупроводниковые лазеры) до ~ 2 м (газовые лазеры). Соответственно, изменяются и указанные параметры.

Стоячие волны в открытом резонаторе могут образоваться и при распространении света под некоторым углом к оси резонатора. В этом случае условие образования стоячей волны:

$$L \cdot \cos\theta = q \frac{\lambda}{2} \quad (6.4),$$

где θ – угол между направлением распространения плоской световой волны, формирующей моду резонатора, и осью z . Соответственно, для частоты света такой моды имеем соотношение:

$$\nu_{q\theta} = \frac{c}{2L \cdot \cos\theta} \cdot q \quad (6.5).$$

(Для рассматриваемого резонатора с бесконечно протяженными зеркалами угол θ может быть любым). Такие стоячие волны называют **поперечными модами** открытого резонатора.

Из последнего соотношения видим, что при заданном L одну и ту же длину волны (одной и той же частоты) могут иметь много типов колебаний, отличающихся величиной θ . Таким образом, имеет место вырождение по волновому вектору (по направлению).

6.1.2. Типы волн в реальных открытых резонаторах

Ранее рассматривалась двумерная конфигурация резонатора с зеркалами бесконечной протяженности. Рассмотрим случай конечных размеров зеркал резонатора. Пусть они имеют прямоугольную форму.

Для описания собственных волн такого резонатора полагаем его сначала закрытым. Из электродинамики известно, что в нем существуют электрические (E_{mnq}) и магнитные (H_{mnq}) типы колебаний. Однако в теории открытых резонаторов типы колебаний обозначают как TEM_{mnq} . Их делят на продольные и поперечные.

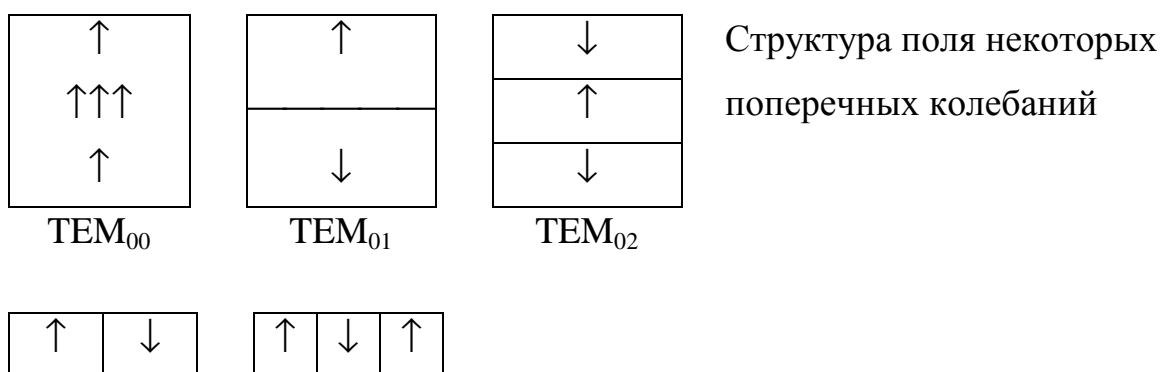
Индекс q в обозначении типа часто опускают ($q \gg m, n$ для открытых резонаторов и, кроме того, разность частот $\Delta\nu \ll \nu$).

Каждый поперечный тип колебаний характеризуется определенной структурой поля в поперечном сечении резонатора и определяется заданием чисел m и n . Для зеркал прямоугольной формы m и n характеризуют число изменений направления поля вдоль осей x и y . Для круглых зеркал: m характеризует число изменений направления поля по радиусу, а n – по азимуту.

Таким образом, в резонаторе существуют чисто продольные ($m=n=0$) типы колебаний и поперечные моды, образованные волнами, распространяющимися под углом к оси резонатора.

При переходе от закрытого резонатора к открытому появляется новый фактор – дифракционные потери на краях зеркал. Для мод с малыми величинами m и n такие потери сравнительно малы (поле сконцентрировано в центральной области и быстро спадает к краям зеркал). Поэтому удаление стенок не приводит к значительным дифракционным потерям. В то же время для мод с большими индексами m и n дифракционные потери очень велики и данные моды не могут существовать в открытом резонаторе.

Таким образом, дифракционные потери играют в открытом резонаторе важную роль, уменьшая число существующих в нем типов колебаний.



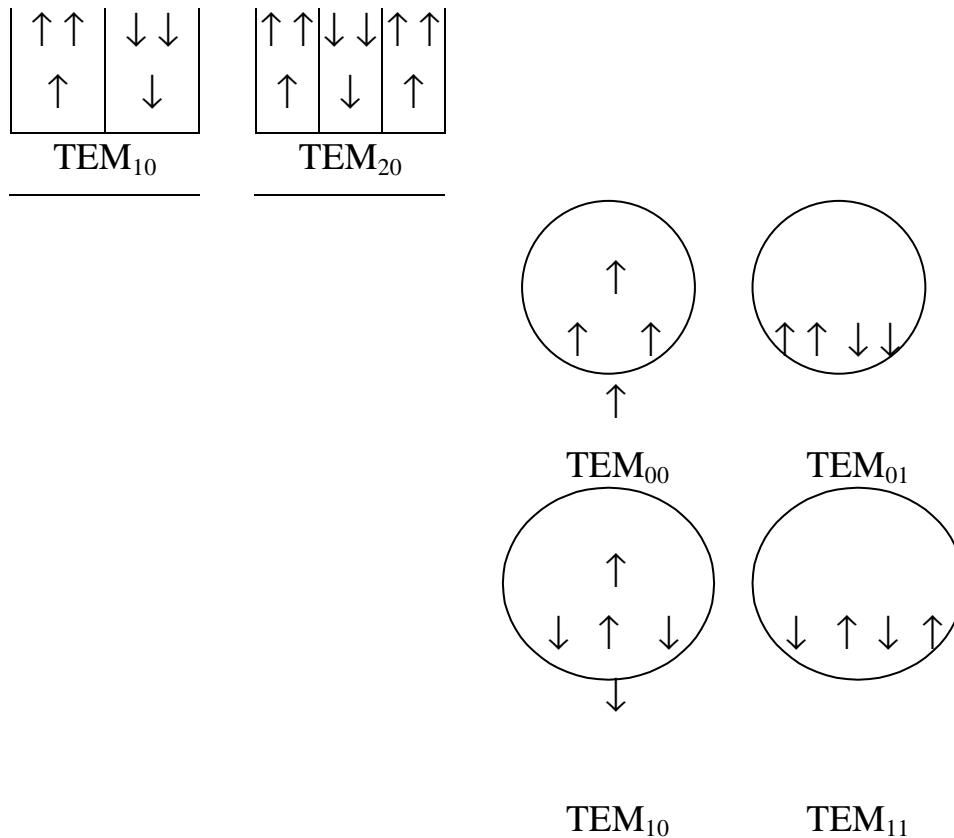


Рис. 6.2.

6.1.3. Добротность типов колебаний открытого резонатора

Важной характеристикой открытых резонаторов является их добротность (как и для других резонансных систем). Пусть Q – добротность типа колебаний резонатора; $U_{\text{тк}}$ – энергия, запасенная в типе колебаний; а ω – его собственная частота. Тогда:

$$U_{\text{тк}} = U_{\text{тк}}^0 \cdot \exp\left(-\frac{\omega}{Q}t\right) \quad (6.6),$$

где $U_{\text{тк}}^0$ – начальная энергия, запасенная в типе колебаний. Дифференцируя это соотношение, получим величину изменения энергии за время dt :

$$dU_{\text{тк}} = -\frac{\omega}{Q}U_{\text{тк}}^0 \exp\left(-\frac{\omega}{Q}t\right)dt = -\frac{\omega}{Q}U_{\text{тк}}dt \quad (6.7).$$

Сначала найдем добротность типа колебаний за счет выхода энергии через зеркала с конечным коэффициентом отражения $r_{\text{отр}}$ (коэффициент

пропускания $-1-r_{отр}$). Поглощением в зеркалах пренебрегаем. Уменьшение энергии типа колебаний при однократном отображении для волны с энергией $U_{тк}/2$:

$$\Delta U_{мк} = -\frac{U_{мк}}{2}(1-r_{отр}) \quad (6.8).$$

Эта энергия теряется за период времени $\Delta t = \frac{L}{c}$ – время прохода волны через резонатор. Поэтому в среднем за единицу времени волна теряет энергию:

$$\Delta U_{мк(t=1)} = -\frac{U_{мк}(1-r_{отр})}{2\Delta t} = -\frac{U_{мк}(1-r_{отр}) \cdot c}{2L} \quad (6.9).$$

Мода (тип колебаний) образуется двумя волнами, бегущими навстречу, поэтому энергия, теряемая типом колебаний в единицу времени, в два раза больше:

$$\Delta U_{мк} = -\frac{U_{мк}(1-r_{отр})}{L} \cdot c \quad (6.10).$$

А энергия, теряемая типом колебания за время dt :

$$dU_{мк} = -\frac{U_{мк}(1-r_{отр})}{L} \cdot c \cdot dt \quad (6.11).$$

Сравним это соотношение с соотношением (6.10):

$$dU_{тк} = -\frac{\omega}{Q} \cdot U_{мк} \cdot dt = -\frac{U_{мк}(1-r_{отр})}{L} \cdot c \cdot dt$$

Отсюда:

$$Q = \frac{\omega L}{(1-r_{отр})c} = \frac{2\pi \cdot L}{\lambda(1-r_{отр})} \quad (6.12).$$

Оценим величину добротности Q для параметров резонатора: $L=30$ см: $\lambda=0,6$ мкм: $r_{отр}=0,9$:

$$Q = \frac{2\pi \cdot 30}{0,6 \cdot 10^{-4} \cdot 0,1} \approx 3 \cdot 10^7$$

Это существенно выше, чем добротность объемных резонаторов в радиодиапазоне.

При проведенном рассмотрении не учитывались такие факторы, как шероховатость зеркал, дифракционные потери, непараллельность установки зеркал и т.д. Их учет может существенно изменить результат оценки.

Следует отметить, что в открытых резонаторах возбуждается много типов колебаний. Ранее оценка расстояния по частоте между соседними аксиальными модами дала величину 500 МГц для $L = 30$ см и $\lambda = 0,6$ мкм. Ширина линии люминесценции рабочей среды обычно значительно больше. Поэтому в пределах линии могут укладываться сотни продольных мод, между которыми располагается еще множество поперечных мод.

Типы открытых резонаторов

В ОКГ наряду с резонаторами Фабри-Перо широко используются резонаторы со сферическими зеркалами. Они характеризуются соотношением длины L , радиуса кривизны зеркал, и их поперечных размеров. Такие резонаторы могут в некоторых случаях иметь меньшие потери, чем резонатор Фабри - Перо. Они также менее критичны к неточности настройки. Ясно также, что в сферических резонаторах возможны поперечные типы колебаний более высокого порядка, чем в резонаторах с плоскими зеркалами.

6.2. Условия самовозбуждения лазера

Условия самовозбуждения лазера, как и любого другого генератора, включают в себя условия баланса фаз и баланса амплитуд (мощностей).

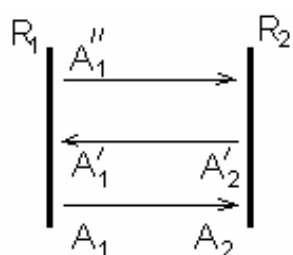


Рис. 6.3.

Рассмотрим резонатор, заполненный активной средой (рис. 6.3). Пусть \dot{R}_1 и \dot{R}_2 – комплексные коэффициенты отражения зеркал, а L – длина резонатора. Положим, что у зеркала 1 возникает спонтанное излучение, которое в виде плоской световой волны A_1 распространяется в направлении к зеркалу 2. Вследствие взаимодействия с активной средой, у зеркала 2 это поле будет иметь вид:

$$A_2 = A_1 \cdot \exp(\kappa_a \cdot L) \cdot \exp(-\alpha \cdot L) \cdot \exp(-ikL) \quad (6.13).$$

Здесь κ_a – коэффициент усиления; α – коэффициент потерь в активной среде, обусловленных рассеянием света на неоднородностях; k – волновое число. После отражения от зеркала 2, поле отраженной волны A_2' принимает вид:

$$A_2' = A_1 \cdot \dot{R}_2 \cdot (1 - \xi) \exp(\kappa_a \cdot L) \cdot \exp(-\alpha \cdot L) \cdot \exp(-ikL) \quad (6.14),$$

где ξ – коэффициент, учитывающий дифракционные потери при отражении от зеркала. При обратном проходе к зеркалу 1 поле A_2' снова усиливается и после отражения от него имеет вид:

$$A_1'' = A_1 \cdot \dot{R}_1 \cdot \dot{R}_2 (1 - \xi)^2 \exp(2\kappa_a \cdot L) \cdot \exp(-2\alpha \cdot L) \cdot \exp(-2ikL) \quad (6.15).$$

Очевидно, что для получения стационарных колебаний (или стоячей волны в резонаторе) должно выполняться условие:

$A_1'' = A_1$, откуда получаем:

$$\dot{R}_1 \cdot \dot{R}_2 (1 - \xi)^2 \exp(2\kappa_a \cdot L) \cdot \exp(-2\alpha \cdot L) \cdot \exp(-2ikL) = 1 \quad (6.16).$$

Представим коэффициенты \dot{R} в виде: $\dot{R} = |R| \cdot e^{i \cdot \Delta\phi}$, где $|R|$ - модуль коэффициента отражения, а $\Delta\phi$ учитывает изменение фазы волны при отражении от зеркала. Учтем также, что $k = 2\pi/\lambda$:

$$|\dot{R}_1| \cdot |\dot{R}_2| (1 - \xi)^2 \exp(2(\kappa_a - \alpha) \cdot L) \cdot \exp\left(-i\left(\Delta\phi_1 + \Delta\phi_2 + 2\frac{2\pi}{\lambda}L\right)\right) = 1 \quad (6.17).$$

Отсюда получим уравнение баланса амплитуд в форме:

$$|\dot{R}_1| \cdot |\dot{R}_2| (1 - \xi)^2 \exp(2(\kappa_a - \alpha) \cdot L) = 1 \quad (6.18)$$

и уравнение баланса фаз:

$$\Delta\phi + \Delta\phi + 2\frac{2\pi}{\lambda}L = 2\pi \cdot m \quad (6.19).$$

Уравнение баланса фаз означает наличие конструктивной интерференции, т.е. положительной обратной связи в резонаторе. Оно определяет частоту лазерного излучения и выполняется на любой резонансной частоте открытого резонатора. Следовательно, на каждой из этих частот может быть достигнута генерация, если выполняется условие баланса амплитуд.

Согласно (6.18), колебания в резонаторе не будут затухать, если усиление в активной среде компенсирует все потери в резонаторе. Логарифмируя (6.18), получим:

$$\begin{aligned} \ln(R_1 \cdot R_2 \cdot (1 - \xi)^2) + 2(\kappa_a - \alpha)L &= 0 \quad \text{или} \\ 2\kappa_a \cdot L &= 2\alpha L - \ln(R_1 \cdot R_2 \cdot (1 - \xi)^2) \quad \text{или} \\ \kappa_a &= \alpha - \frac{1}{2L} \ln(R_1 \cdot R_2 \cdot (1 - \xi)^2) \quad \text{и, окончательно:} \end{aligned}$$

$$\kappa_a = \alpha + \frac{1}{L} \ln \sqrt{\frac{1}{R_1 \cdot R_2 \cdot (1 - \xi)^2}} \quad (6.20).$$

В (6.20) первое слагаемое справа характеризует потери в активной среде, а второе – потери в зеркалах.

Величина $\kappa_a = \alpha + \frac{1}{L} \ln \sqrt{\frac{1}{R_1 \cdot R_2 \cdot (1 - \xi)^2}}$ представляет собой пороговый коэффициент усиления. Генерация возможна, когда $\kappa_a > \alpha$.

6.3. Газовые лазеры

В газовых лазерах активной средой является газ, смесь нескольких

газов, или смесь газов с парами металла. Агрегатное состояние активной среды обуславливает ряд особенностей характеристик лазерного излучения. Среди них можно выделить, в первую очередь, следующие:

а) Рабочие уровни в газе – это уровни почти изолированных частиц (атомов, ионов, молекул). Взаимодействие между частицами в газе значительно слабее, чем в твердых телах и жидкостях, поэтому линии рабочих переходов могут быть очень узкими.

б) Газ оптически более однороден, чем другие среды, поэтому потери на рассеяние и дифракционные потери в таких средах минимальны.

в) Вследствие низкой концентрации активных частиц газовые среды имеют значительно более низкий коэффициент усиления. Поэтому для получения генерации нужны существенные размеры. Эти размеры изменяются от сантиметров до сотни метров.

Область длин волн излучения газовых лазеров: различные газовые лазеры генерируют оптическое излучение в диапазоне от ультрафиолетового ($\sim 0,2$ мкм) до далекого инфракрасного (более 400 мкм).

Коэффициент полезного действия (КПД) газовых лазеров во многих случаях невысок, однако для отдельных типов таких приборов он достигает достаточно больших величин, например – до 30% для лазера на CO_2 .

Газовые лазеры делят на атомарные, ионные и молекулярные. Соответственно, рабочими уровнями активных сред являются энергетические уровни нейтральных атомов, ионов или молекул газов. Рабочие области спектра для данных лазеров составляют:

- для атомарных – как правило, от видимого до инфракрасного диапазона (до 25 мкм);
- для ионных – в основном, они лежат в видимом и ультрафиолетовом диапазонах;
- молекулярные лазеры работают, обычно, на длинах волн более 5 мкм.

Процессы, используемые в газовых лазерах для достижения инверсии населенностей

Для газовых активных сред использование прямой оптической накачки нерационально, поскольку для этого потребовались бы когерентные узкополосные источники излучения, т.е. практически другие лазеры. Наиболее часто для получения инверсии населенностей используется электрический разряд в газе. Возбуждение осуществляется за счет столкновений атомов с электронами, а также столкновений между возбужденными и невозбужденными атомами. Кроме того, состояние инверсии населенностей может быть получено в результате химических реакций, при которых образуются атомы или радикалы в возбужденных состояниях, а также при диссоциации молекул с образованием атомов в

возбужденных состояниях. Однако в большинстве случаев используется все – таки электрический разряд в газовой среде. Рассмотрим особенности нескольких типов газоразрядных лазеров.

6.3.1. Лазер на смеси гелия и неона (He – Ne лазер).

Это первый газовый лазер, он создан в конце 1960 года. Несмотря на громадные успехи квантовой электроники, такие лазеры и в настоящее время являются одними из наиболее широко используемых типов оптических квантовых генераторов. Это обусловлено хорошо отработанной конструкцией и технологией изготовления таких приборов, а также щадящими условиями в тлеющем разряде внутри газоразрядной трубки.

Для получения инверсии населенностей в He – Ne лазерах используется механизм столкновений второго рода. Это процесс, реализуемый в смеси двух газов. Атомы одного из них (A) в разряде за счет столкновений первого рода (неупругие столкновения между электронами газового разряда и атомами, находящимися в невозбужденном состоянии $\rightarrow A + e \Rightarrow A^* + e$). Возбужденное состояние

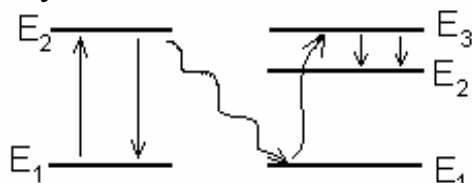


Рис. 6.3.

A^* является метастабильным, т.е. атомы газа A могут долго находиться в возбужденном состоянии. Энергия возбужденного состояния атома A должна быть близка к энергии верхнего возбужденного уровня атома другого газа (рис. 6.3).

При неупругих столкновениях атома A^* с атомом B (в основном состоянии) происходит эффективная резонансная передача энергии от A^* к B . В результате атом B переходит на уровень 3, а атом A^* переходит в основное состояние: $A^* + B \Rightarrow A + B^*$. Вероятность такого перехода велика, если $\omega_{12A} = \omega_{13B}$. В результате в системе B может быть получено состояние инверсии населенностей для уровней 3 – 2. В такой ситуации газ A – вспомогательный, а B – основной (рабочий). В He – Ne лазере в качестве рабочего газа выступает Ne, а вспомогательным является He. Диаграмма энергетических уровней смеси He и Ne схематично изображены на рис. 6.4.

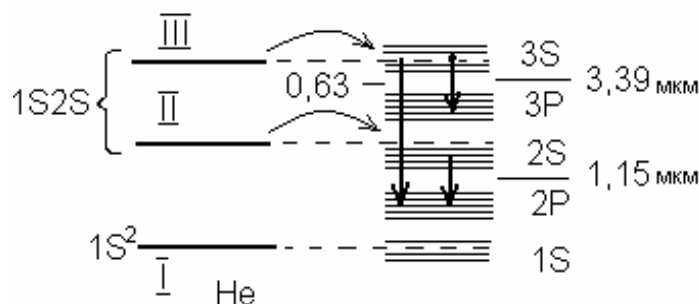


Рис. 6.4.

Основное состояние атома *He*, имеющего 2 электрона – $1S^2$. Ему соответствуют квантовые числа $n = 1; l = 0; S = 1/2; j = 1/2$. Поскольку оболочка атома заполнена полностью, то все суммарные моменты количества движения $\bar{S}, \bar{L}, \bar{j}$ равны нулю. Основные состояния атомов обозначены цифрой I.

При разряде из – за неупругих соударений со свободными электронами происходит возмущение (возбуждение) атомов *He*, которые могут переходить на метастабильные уровни II и III. Энергия возбужденных атомов *He* передается к атомам *Ne*.

Энергетические уровни *He* II и III близки к уровням $2S$ и $3S$ атома *Ne*. Поэтому при неупругих соударениях возбужденных атомов *He* с невозбужденными атомами *Ne* происходит эффективная передача избыточной энергии атомов *He* к атомам *Ne*. Атомы *Ne* переходят в возбужденное состояние $2S$ или $3S$, а атомы *He* – в основное состояние I.

Очевидно, что существует и обратный процесс передачи энергии от атомов *Ne* к атомам *He*. Для того, чтобы преобладал прямой процесс, необходимо, чтобы концентрация *He* была выше концентрации *Ne*. Их отношение находится обычно в пределах от 5 до 15.

В результате неупругих соударений атомов *He* и *Ne* населенности уровней $2S$ и $3S$ возрастают, а уровней $2P$ и $3P$ остаются малыми. Время жизни уровней $2S$ и $3S$ у *Ne* больше, чем уровней $2P$ и $3P$. Поэтому при некотором токе разряда возникает инверсия населенностей в переходах $3S - 3P$, $2S - 2P$ и $3S - 2P$.

Этим переходам соответствуют следующие длины волн излучения:

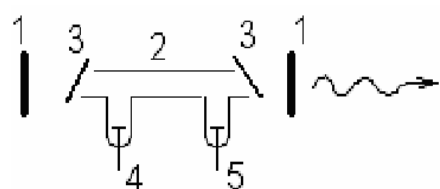
$$3S - 3P - 3,39 \text{ мкм}$$

$$2S - 2P - 1,15 \text{ мкм}$$

$$3S - 2P - 0,6328 \text{ мкм}$$

Инверсия для всех указанных переходов достигается при одинаковых условиях разряда, поэтому для выделения той или иной линии излучения используют селективные элементы. Селекция может осуществляться, например, за счет сильной частотной зависимости коэффициента отражения зеркал резонатора. Использование сменных зеркал (глухих) на головке револьверного типа позволяет реализовать трехчастотные лазеры (ЛГ–126, ЛГН–113).

Конструкция типичного *He - Ne* лазера иллюстрируется схемой на рисунке 6.5. Активный элемент состоит из газоразрядной трубки, наполненной смесью данных газов, и оптического резонатора.



1 - зеркала

2 - трубка

3 - выходные окна

4,5 - анод и катод

Выходная мощность лазера зависит от тока разряда, общего давления

газовой смеси в трубке, соотношения парциальных давлений газов, диаметра трубки.

Рис. 6.5.

Зависимость выходной мощности от тока разряда имеет максимум (рис. 6.6.). Сначала при росте тока разряда мощность излучения возрастает. Однако при некоторой величине этого тока начинает проявляться электронное возбуждение уровней $2P$ и $3P$ Ne с метастабильного уровня $1S$. Вследствие этого населенности нижних рабочих уровней возрастают, а инверсия и выходная мощность падают вплоть до полного срыва генерации.

Увеличение общего давления в смеси повышает концентрацию активных атомов He и Ne , поэтому населенность возбужденных уровней и мощность излучения сначала возрастают. Однако при слишком высоком давлении уменьшается длина свободного пробега электронов и, соответственно, их энергия в ускоряющем электрическом поле на этом пути. Это приводит к снижению энергии, передаваемой атомами He и к уменьшению выходной мощности. Оптимальная величина давления составляет ~ 100 Па.

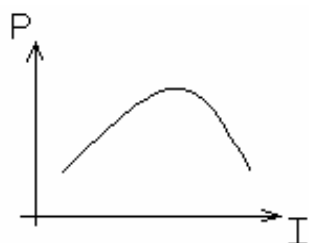


Рис. 6.6.

Выходная мощность $He - Ne$ лазеров составляет до 100 мВт ($\lambda = 0,63$ мкм) и до 1 Вт ($\lambda = 1,15$ мкм). Благодаря отработанной конструкции срок службы их может превышать 20000 часов. Рабочий ток разряда зависит от типа и выходной мощности и меняется от $0,02$ А до $0,1$ А.

6.3.2. Ионный аргоновый лазер

В газовых ионных лазерах рабочими переходами являются переходы между энергетическими уровнями ионов. К существенным отличиям активной среды в этом случае можно отнести следующее:

- а) Расстояние между рабочими уровнями ионов больше, чем у нейтральных атомов, поэтому длина волны излучения этих лазеров соответствует видимому и ближнему ультрафиолетовому диапазонам.
- б) Вероятность переходов между рабочими уровнями ионов больше, чем у нейтральных атомов.

Рабочими средами ионных лазеров являются чистые инертные газы: аргон (Ar), криптон (Kr), ксенон (Xe), Ne , а также ионизированные пары различных химических элементов (Se, Zn, I, Cd).

Энергетические уровни основного и возбужденного состояний ионов расположены значительно выше уровней возбужденных состояний

нейтральных атомов. Поэтому вероятность их прямого возбуждения очень мала. Установлено, что заселение верхнего, рабочего уровня происходит в результате ступенчатого возбуждения при соударениях электронов с ионами, находящимися в основном состоянии. Для эффективной работы лазера плазма газового разряда должна быть высокоионизированной, а кинетическая энергия электронов должна быть существенно выше, чем в случае тлеющего разряда. Поэтому в таких лазерах применяется не тлеющий, а дуговой разряд с плотностью тока от 100 А/см^2 в непрерывном режиме до 1000 А/см^2 в импульсном режиме.

Наиболее распространен ионный *Ar* лазер. Квантовые переходы соответствуют длинам волн от 0,45 до 0,51 мкм (сине – зеленая область). Вследствие высокой плотности тока все *Ar* лазеры имеют водяное охлаждение. Газоразрядные трубки изготавливаются для повышения термической прочности из бериллиевой керамики. Для предотвращения бомбардировки стенок трубки ионами и электронами используется продольное магнитное поле, создаваемое соленоидом. Оно удерживает заряженные частицы вблизи оси трубки. Использование магнитного поля позволяет повысить выходную мощность в несколько раз (за счет повышения плотности тока). Коэффициент полезного действия таких лазеров также очень низок и может достигать лишь 0,01..0,1%. Мощность излучения в непрерывном режиме достигает 175 Вт.

6.3.3. Молекулярные газовые лазеры

Для атомарных и ионных лазеров характерен низкий к.п.д. Это связано с тем, что верхний рабочий уровень лежит высоко над основным уровнем. Поскольку инверсия населенностей создается при электронном возбуждении из основного состояния, то только небольшая часть электронов (с высокой кинетической энергией) участвует в процессе возбуждения. Для *He – Ne* лазеров оценка числа таких электронов дает $5 \div 6\%$.

К.п.д. лазеров можно увеличить, используя в качестве рабочих энергетические уровни, находящиеся вблизи основного. Например, это колебательные уровни молекул. Соответственно, такие лазеры излучают в инфракрасном диапазоне. Для возбуждения молекул используют различные методы, например – электрический разряд, химические реакции, оптическую и тепловую накачку. Наиболее распространен лазер на молекулах двуокиси углерода (*CO₂* – лазер). Здесь используется смесь рабочих газов: *CO₂*, *N₂*. Для увеличения теплопроводности смеси и отвода тепла к стенкам газоразрядной трубки в нее добавляют также *He*.

Мощность излучения при изменении диаметра газоразрядной трубки в таких лазерах меняется мало. Поэтому этот диаметр может достигать $15 \div 30$ см, а длина – нескольких метров. Особенность данного типа лазеров –

необходимость прокачки CO_2 через трубку, т.к. в рабочей смеси происходит диссоциация молекул на CO и O_2 : $2CO_2 = 2CO + O_2$.

Типичные характеристики излучения CO_2 – лазеров:

$\lambda = 10,6\text{мкм}$ (есть и другие линии, например $\lambda = 9,6\text{мкм}$);

$P_{\text{вых}} = 9\text{кВт}$ (непрерывный);

К.П.Д. – 30%;

$P_{\text{имп. max}} = 100\text{МВт}$.

6.4. Твердотельные лазеры

К семейству твердотельных лазеров можно отнести лазеры на ионных кристаллах и стеклах, а также полупроводниковые и волоконно – оптические лазеры. Однако здесь речь пойдет лишь о лазерах на основе ионных кристаллов. Активными средами таких оптических генераторов являются кристаллические диэлектрики с введенными в них активными ионами. Это, как правило, ионы хрома (Cr^{3+}) или неодима (Nd^{3+}). К основным отличительным особенностям твердотельных лазеров по сравнению с газовыми можно отнести следующее.

а) Концентрация активных ионов в кристалле может составлять от сотых долей процента (если активная примесь специально вводится в материал) до десятков процентов (если такая примесь входит в состав кристалла). В любом случае эта концентрация значительно выше, чем в газовых средах. Поэтому в твердых активных средах можно достичь значительно более высоких населенностей на рабочих уровнях, и, соответственно, более высоких коэффициентов усиления и мощности излучения на единицу объема. Таким образом, возможно достижение очень высоких световых мощностей при очень малых размерах активных элементов.

б) Для создания инверсии населенностей в твердотельных лазерах на ионных кристаллах, как правило, используется оптическая накачка.

в) Энергетические уровни свободного иона и того же иона, окруженного другими ионами в кристалле, существенно различаются. Число таких уровней у иона в кристалле больше, а их положение по шкале энергии (или частот) может отличаться от такового для свободного иона. Влияние сильных внутрикристаллических электрических полей приводит также к уширению энергетических уровней активных ионов и даже к возникновению вместо узких уровней широких энергетических полос – зон. В то же время для получения лазерной генерации используются переходы электронов между энергетическими уровнями незаполненных внутренних оболочек иона, для которых влияние внутрикристаллического поля сказывается слабо. В итоге эти уровни оказываются достаточно узкими. Таким образом, спектр поглощения подобных лазерных кристаллов состоит как из отдельных

достаточно узких линий, так и может иметь широкие полосы поглощения, что повышает эффективность оптической накачки.

Первым твердотельным лазером был лазер на рубине, длина волны излучения которого составляет 0,69 мкм, созданный почти одновременно с гелий – неоновым лазером. В течение долгого времени подобные лазеры находили широкое применение, однако их существенным недостатком являлась сложность достижения непрерывного режима генерации. В настоящее время они практически не применяются, уступив место лазерам на основе кристаллов иттрий – алюминиевого граната, легированных неодимом ($Y_3Al_5O_{12}:Nd^{3+}$, или $YAG:Nd^{3+}$). Последние обеспечивают генерацию излучения с длиной волны 1,06 мкм (ближний инфракрасный диапазон) как в непрерывном, так и в импульсном режиме. При использовании удвоителей частоты лазеры на $YAG:Nd^{3+}$ успешно конкурируют с ионными аргоновыми лазерами как по мощности излучения в зеленой области видимого спектра, так и по когерентности излучения, значительно превосходя их по к.п.д. Последние достижения в области разработки твердотельных лазеров связаны с созданием новых лазерных кристаллов и отработкой методов диодной накачки, что позволило улучшить характеристики излучения таких приборов как в ближнем ИК, так и в видимом диапазоне длин волн и существенно повысить их к.п.д.

Рассмотрим некоторые особенности построения и характеристики современных твердотельных лазеров. Во – первых, остановимся на элементах оптической накачки.

6.4.1. Элементы накачки твердотельных лазеров

Для реализации оптической накачки необходимы мощные источники света, у которых основная доля энергии должна попадать в полосы поглощения активного вещества. Наиболее подходящими до недавнего времени являлись газоразрядные лампы (например, ксеноновые или криптоновые). Они обладают достаточно высоким к.п.д. (до 50%). В последнее время наряду с газоразрядными лампами все шире используются мощные светодиодные линейки или матрицы.

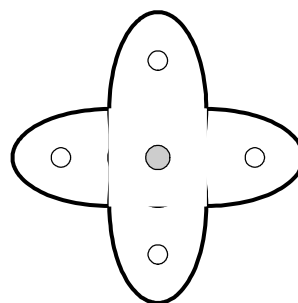
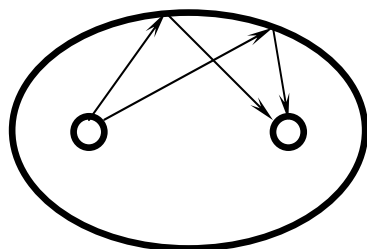
Для передачи как можно большей части излучаемой энергии в активное вещество используют различные отражающие и фокусирующие устройства.

а) В первых экспериментальных образцах твердотельных лазеров использовались цилиндрические отражатели кругового сечения и газоразрядные лампы спиральной формы. Однако их эффективность оказалась недостаточно высокой, поэтому в последующих реально работающих лазерах стали применять отражатели эллиптической формы.

б) Эллиптический отражатель имеет в поперечном сечении форму эллипса, а центры активного элемента в форме стержня и газоразрядной

лампы совпадают с положением главных фокусов эллипса. Очевидно, что такая геометрия позволяет наиболее эффективно доставить излучение лампы накачки к активному элементу. Установлено, что при хорошей юстировке в активный элемент попадает до 75% энергии излучения лампы.

Кроме



в)
простого

Рис. 6.7. Примеры форм отражательных систем

эллиптического отражателя иногда используются также многоэллипсные рефлекторы (рис. 6.7), представляющие собой комбинацию нескольких эллиптических отражателей, у которых активный элемент располагается в одном из фокусов, а в остальных – лампы накачки. Эффективность такой системы не превышает эффективности простого эллиптического отражателя, однако использование нескольких ламп накачки позволяет повысить суммарную мощность излучения лазера. Ясно, что конструкция отражательной системы в данном случае оказывается значительно более сложной, что является ее существенным недостатком.

г) В случае использования полупроводниковых лазерных диодов для оптической накачки активных элементов твердотельных лазеров возможны схемы с боковой и торцевой накачкой. Боковая накачка, т.е. ввод излучения в активный элемент через его боковую поверхность, используется в мощных лазерах. В приборах малой и средней мощности, как правило, используются схемы торцевой накачки, т.е. ввод излучения накачки с торца активного элемента, через одно из зеркал оптического резонатора.

6.4.2. Лазер на иттрий – алюминиевом гранате с неодимом

Лазеры на гранате с неодимом - самые популярные из твердотельных лазеров. Активным элементом здесь является цилиндрический образец из монокристалла $Y_3Al_5O_{12}$, в котором часть ионов Y^{3+} замещена ионами Nd^{3+} (концентрация активной примеси, как правило, не превышает 3 ат. %).

На рис.6.8 представлена упрощенная схема энергетических уровней иона Nd^{3+} в YAG и возможных переходов трех электронов внутренней 4f - оболочки этого иона. Как отмечено выше, на их энергетические уровни кристаллическое поле влияет очень слабо из-за экранирования электронами

внешних оболочек. Поэтому данные энергетические уровни достаточно узки. Длинам волн 0,73 и 0,8 мкм соответствуют полосы поглощения, используемые для оптической накачки. Эти полосы связаны с уровнем ${}^4F_{3/2}$ (E_3) безызлучательными переходами с временем релаксации $\sim 10^{-7}$ с. Из всех возможных переходов наиболее медленный – это переход ${}^4F_{3/2} - {}^4I_{11/2}$. Поэтому на уровне ${}^4F_{3/2}$ (E_3) может запасаться большая доля энергии накачки, и он удовлетворяет требованиям, предъявляемым к верхнему лазерному рабочему уровню. Переходы с уровня ${}^4I_{11/2}$ в основное состояние ${}^4I_{9/2}$ являются безызлучательными и характеризуются временем релаксации $\tau=10^{-9}$ с, поэтому уровень ${}^4I_{11/2}$ можно считать пустым, то есть этот уровень хорошо подходит на роль нижнего лазерного уровня (E_2). Таким образом, переход ${}^4F_{3/2}$ (E_3) - ${}^4I_{11/2}$ (E_2) является перспективным для получения лазерной генерации в четырехуровневой схеме. Данному переходу соответствует длина волны излучения $\lambda=1,064$ мкм (ближний ИК- диапазон).

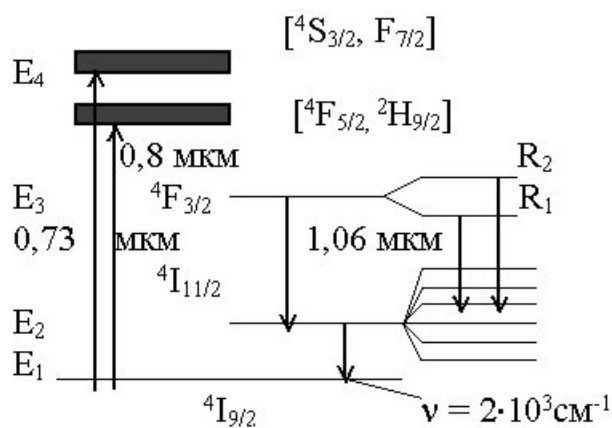


Рис. 6.8. Упрощенная схема энергетических уровней кристалла YAG:Nd^{3+}

Большое время жизни верхнего лазерного уровня (E_3) с характерным временем $\tau=0,23$ мс позволяет успешно работать в режиме модулированной добротности. YAG:Nd^{3+} – лазеры могут генерировать излучение как в непрерывном, так и в импульсном режиме. При работе лазера в импульсном режиме для накачки используются ксеноновые лампы, в непрерывном – криптоновые. Размеры активных элементов достигают $0,5 \times 12$ см. Характеристики излучения:

- 1) в непрерывном многомодовом режиме мощность излучения может достигать 500 Вт;
- 2) в импульсном режиме с частотой повторения импульсов 50 Гц средняя мощность может достигать величину до 200 Вт;
- 3) в режиме модулированной добротности мгновенная мощность в импульсе достигает до 50 МВт;

4) коэффициент полезного действия лазеров с ламповой накачкой составляет до 3 %.

Лазеры на YAG:Nd^{3+} нашли применение в науке, в медицине, а наиболее широко используются в обработке материалов (резка, сверление, сварка, осаждение металлов и т.д.).

6.4.3. Твердотельные лазеры с диодной накачкой

Использование полупроводниковых лазеров в качестве источников накачки твердотельных лазеров позволило существенно уменьшить их размеры и повысить к.п.д. В качестве активной примеси в лазерных материалах в этом случае также наиболее широко используется неодим. Как отмечалось, для его ионов характерно наличие полосы поглощения вблизи длины волны 0,8 мкм, что хорошо согласуется с длиной волны излучения AlGaAs лазеров. В то же время, особенности характеристик излучения полупроводниковых лазеров предъявляют специфические требования к характеристикам лазерных материалов. Так, большая расходимость излучения полупроводниковых лазеров накладывает ограничения на длину активных элементов при торцевой накачке, что требует увеличения поглощения излучения накачки. В результате, требуется более высокая концентрация ионов Nd^{3+} в таких материалах.

В современных твердотельных лазерах в качестве активных сред применяются Nd:YAG , Nd:YVO_4 (ванадат иттрия), Nd:LSB (скандоборат лантана). Типичная конструкция лазера с диодной накачкой, генерирующего излучение с длиной волны 0,53 мкм, представлена схематически на рис. 6.9. Здесь излучение лазерного диода (ЛД) с длиной волны 0,81 мкм с помощью фокусирующей оптической системы (ФО) вводится через торец в активный кристалл (YVO_4). Зеркало (З) прозрачно для излучения накачки, но обладает высоким коэффициентом отражения для излучения с длиной волны 1,06 мкм. В оптический резонатор, образованный двумя зеркалами (З), помещен нелинейный кристалл КТР (титанил - фосфат калия), осуществляющий удвоение частоты (эффект удвоения частоты или генерации второй гармоники рассматривается в Главе 8). Второе зеркало резонатора является прозрачным для излучения с $\lambda=532$ нм и хорошо отражает на основной гармонике. Выведенное из резонатора излучение с удвоенной частотой коллимируется оптической системой (К). Термохолодильник (ТХ), в качестве которого используется элемент Пелтье, необходим для стабилизации длины волны излучения полупроводникового лазера.

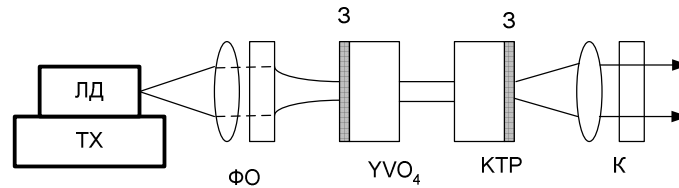


Рис. 6. 9. Схема конструкции твердотельного лазера с диодной накачкой

В рассмотренной схеме в оптическом резонаторе размещены два кристалла, один из которых является активной средой, а второй играет роль нелинейного элемента для удвоения частоты. Дальнейшего улучшения характеристик подобных лазеров и упрощения их конструкции можно достичь, объединяя функции активной среды и удвоителя частоты в одном кристалле. Известны два кристалла, удовлетворяющие подобным требованиям. Это $\text{Nd:YAl}_3(\text{BO}_3)_4$ – алюмоиттриевый борат Nd:YAB , а также Nd:CSB – скандоборат церия.

7. ВЗАИМОДЕЙСТВИЯ СВЕТА С ФИЗИЧЕСКИМИ ПОЛЯМИ

При распространении световых волн в среде, на которую воздействуют какие – либо возмущения (или физические поля), возможно изменение характеристик оптического излучения. В таких случаях говорят о взаимодействии света с различными физическими полями в среде. В зависимости от природы возмущений можно говорить об электрооптическом, акустооптическом, магнитооптическом, фоторефрактивном, термооптическом эффектах и т.д. Рассмотрим некоторые из них, имеющие значительное практическое значение.

7.1. Электрооптический эффект

Электрооптический (ЭО) эффект – это изменение показателя преломления среды под действием приложенного постоянного или переменного электрического поля [3]. В некоторых кристаллах, не имеющих центра симметрии, величина изменения показателя преломления Δn пропорциональна напряженности электрического поля \bar{E} : ($\Delta n \sim \bar{E}$). Это линейный ЭО эффект или эффект Поккельса, который впервые был обнаружен в 1893 году. Известен также квадратичный ЭО эффект, или эффект Керра, который может наблюдаться в любой среде, включая жидкости и газы. Величина Δn в этом случае пропорциональна квадрату напряженности поля ($\Delta n \sim E^2$). Этот эффект обнаружен и исследован Керром впервые в 1875 году.

Качественно механизм электрооптического эффекта может быть объяснен следующим образом. Под действием электрического поля световой волны “центры тяжести” положительного и отрицательного зарядов в нейтральных атомах смещаются, превращая эти атомы в электрические диполи, осциллирующие с частотой поля. Величина наведенной в среде поляризации определяется дипольным моментом единицы объема, и, в конечном итоге, характеризует величину диэлектрической проницаемости (или показателя преломления) среды. В свою очередь, дипольный момент атома определяется величиной малого смещения \bar{r} электрона относительно состояния устойчивого равновесия. Данное смещение \bar{r} может быть найдено из уравнения:

$$\frac{d^2\bar{r}}{dt^2} + 2\gamma \cdot \frac{d\bar{r}}{dt} + \omega_0^2 \cdot \bar{r} = -\frac{e}{m} \bar{E} ,$$

где ω_0 – собственная частота осциллятора; γ - коэффициент, учитывающий возможную диссипацию энергии, т.е. наличие сил “трения”; e и m – заряд и масса электрона; \bar{E} – напряженность внешнего электрического поля. Под

полем \bar{E} понимается высокочастотное поле световой волны. В случае воздействия на среду внешнего низкочастотного (или постоянного) дополнительного поля \bar{E}_0 это уравнение принимает вид:

$$\frac{d^2\bar{r}}{dt^2} + \gamma \cdot \frac{d\bar{r}}{dt} + \omega_0^2 \cdot \bar{r} = -\frac{e}{m}(\bar{E} + \bar{E}_0)$$

Очевидно, что при одном и том же поле \bar{E} , величина наведенной поляризации на оптической частоте будет различной в случаях наличия и отсутствия низкочастотного поля. С другой стороны, наличие “начальной” поляризации среды, наведенной под действием низкочастотного электрического поля, может интерпретироваться и как изменение собственной частоты осциллятора ω_0 . А это приводит к изменению величины диэлектрической проницаемости среды на заданной частоте светового поля.

В принципе, для описания электрооптических явлений может использоваться строгий квантовый подход. Расчеты электрооптических свойств некоторых неорганических кристаллов показали их допустимое согласие с экспериментальными результатами. Однако на практике для описания электрооптического эффекта используется феноменологический подход, использующий понятие ЭО тензора [3, 13].

7.1.1. Феноменологическое описание электрооптического эффекта

Как известно, диэлектрическая проницаемость среды является тензором второго ранга. Т.е. связь между векторами \bar{E} и \bar{D} имеет вид:

$$\begin{cases} D_x = \epsilon_{xx} \cdot E_x + \epsilon_{xy} \cdot E_y + \epsilon_{xz} \cdot E_z \\ D_y = \epsilon_{yx} \cdot E_x + \epsilon_{yy} \cdot E_y + \epsilon_{yz} \cdot E_z \\ D_z = \epsilon_{zx} \cdot E_x + \epsilon_{zy} \cdot E_y + \epsilon_{zz} \cdot E_z \end{cases} \quad (7.1).$$

Эти соотношения принято записывать в сокращенной форме:

$$D_i = \epsilon_{ij} \cdot E_j \quad (7.2),$$

где индекс i принимает поочередно все возможные значения (x, y, z в приведенных соотношениях), а по повторяющемуся индексу j проводится суммирование.

Для представления оптических свойств или показателя преломления анизотропного кристалла, используется понятие оптической индикатрисы [2, 5, 6], которой в общем случае соответствует поверхность второго порядка. Уравнение индикатрисы можно представить в виде:

$$B_{ij} \cdot x_i \cdot x_j = 1 \quad (7.3),$$

где B_{ij} – тензор относительной диэлектрической непроницаемости среды. Компоненты данного тензора определяются соотношением:

$$B_{ij} = \epsilon_0 \cdot \left(\frac{\partial E_i}{\partial D_j} \right) \quad (7.4).$$

Главные значения тензора диэлектрической непроницаемости выражаются через главные показатели преломления, если тензор B_{ij} приведен к главным осям (т.е. только диагональные его компоненты B_{ii} отличны от нуля, а $B_{12}=B_{13}=B_{23}=0$). Тогда:

$$B_{11} = \frac{1}{n_x^2}; \quad B_{22} = \frac{1}{n_y^2}; \quad B_{33} = \frac{1}{n_z^2} \quad (7.5).$$

При таком подходе небольшое изменение показателя преломления под действием электрического поля (либо других возмущений) соответствует изменению формы, размера и ориентации оптической индикатрисы. Эти изменения наиболее удобно описывать через малые изменения ΔB_{ij} коэффициентов B_{ij} :

$$\Delta B_{ij} = r_{ijk} \cdot E_k + S_{ijkl} \cdot E_k \cdot E_l \quad (7.6).$$

Здесь первое слагаемое в правой части соответствует линейному ЭО эффекту, а второе – квадратичному, т.е. эффекту Керра. Соответственно, r_{ijk} – тензор коэффициентов линейного ЭО эффекта. ЭО коэффициент определяет изменение диэлектрической непроницаемости среды в зависимости от приложенного электрического поля. Он связывает тензор 2-го ранга с вектором (или тензором 1-го ранга), поэтому также является тензором, но уже третьего ранга.

В общем случае тензор 3-го ранга имеет 27 независимых компонент. Однако тензор B_{ij} симметричен (как и ϵ_{ij}) и имеет только 6 независимых компонент, поэтому тензор r_{ijk} имеет только 18 независимых элементов. Кроме того, симметрия среды накладывает требования инвариантности ее физических свойств и, соответственно, тензора r_{ijk} относительно операций симметрии, преобразующих кристалл в самого себя. В итоге многие из коэффициентов r_{ijk} оказываются равными нулю. В частности, в centrosymmetric среде эффект должен быть одинаков как для поля E , так и для поля $-E$. А это возможно лишь при нулевом тензоре r_{ijk} . Т.е. в среде с центром симметрии линейный электрооптический эффект должен отсутствовать.

Симметрия тензора ϵ_{ij} (и, соответственно, B_{ij}) и наличие только 18 независимых компонент у тензора r_{ijk} позволяет упростить запись соотношений типа (7.6). Это достигается уменьшением числа индексов в записи соответствующих тензоров. Так, для B_{ij} проводят замену индексов по схеме: (11) \Rightarrow (1); (22) \Rightarrow (2); (33) \Rightarrow (3); (23) \Rightarrow (4); (13) \Rightarrow (5); (12) \Rightarrow (6). Алгоритм замены пары индексов на единственный поясняется диаграммой (рис. 7.1).

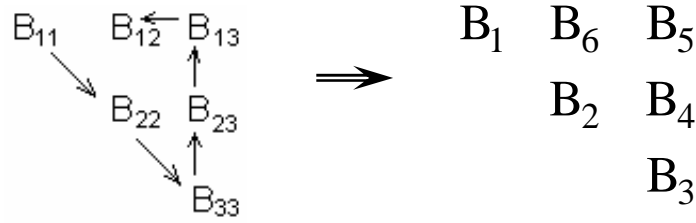


Рис. 7.1.

По такой же схеме проводят замену первых двух индексов у r_{ijk} . В итоге соотношение, определяющее связь между E и ΔB , запишется в матричной форме: $\Delta B_i = r_{ij} \cdot E_j$. Здесь i пробегает значения от 1 до 6; j – от 1 до 3. Таким образом, компоненты тензора r_{ij} могут быть представлены в виде матрицы из 3-х столбцов с 6 строками.

Изменение главных показателей преломления найдем из соотношения:

$\Delta B_i = \Delta \left(\frac{1}{n^2} \right)_i$. Представляя его в виде: $\Delta B = \frac{dB}{dn} \cdot \Delta n$, получим:

$\Delta B = -2 \cdot \frac{1}{n^3} \cdot \Delta n$. Таким образом:

$$\Delta n_i = -\frac{1}{2} n^3 \cdot \Delta B_i = -\frac{1}{2} n^3 \cdot r_{ij} \cdot E_j \quad (7.7).$$

Пример 1. Тензор r_{ijk} для кристалла ниобата лития (LiNbO_3):

$$\begin{bmatrix} 0 & -r_{22} & r_{13} \\ 0 & r_{22} & r_{13} \\ 0 & 0 & r_{33} \\ 0 & r_{51} & 0 \\ r_{51} & 0 & 0 \\ -r_{22} & 0 & 0 \end{bmatrix}$$

В случае внешнего электрического поля, имеющего составляющую лишь вдоль оптической оси кристалла: $E = E_3$; $\Delta B_i = r_{i3} \cdot E_3$. Видим, что для LiNbO_3 из r_{i3} отличны от нуля лишь $r_{13}=r_{23}=r_{13}$; $r_{33}=r_{33}$, а $r_{43}=r_{53}=r_{63}=0$. Соответственно, при обыкновенной поляризации света (вектор \vec{E} перпендикулярен оси z) $\Delta n_1 = \Delta n_2 = -\frac{1}{2} n_0^3 \cdot r_{13} \cdot E_3$; а при необыкновенной –

$$\Delta n_3 = -\frac{1}{2} n_e^3 \cdot r_{33} \cdot E_3.$$

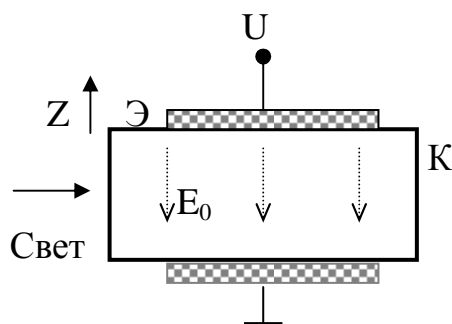
Величины коэффициентов r_{13} и r_{33} для LiNbO_3 существенно

различаются. Так, $r_{13}=9,6 \cdot 10^{-10}$ см/В; $r_{33}=30,9 \cdot 10^{-10}$ см/В. Очевидно, что для получения эффекта одинаковой величины, использование компоненты r_{33} требует примерно в три раза меньших электрических полей. В таблице 7.1 приведены для справки величины некоторых линейных электрооптических коэффициентов наиболее известных электрооптических кристаллов.

Материал	λ , мкм	n	Таблица 7.1.	
			r_{ij} , 10^{-12}	см/В
LiNbO ₃	0,633	$n_e=2,2$ $n_o=2,286$	$r_{13}=8,6$	
			$r_{22}=3,4$	
			$r_{33}=30,8$	
			$r_{51}=28$	
LiTaO ₃	0,633	$n_e=2,180$ $n_o=2,176$	$r_{13}=7,5$	
			$r_{22}=1$	
			$r_{33}=33$	
			$r_{51}=20$	
KH ₂ PO ₄ (KDP)	0,633	$n_e=1,467$ $n_o=1,507$	$r_{41}=8$	
			$r_{61}=11$	
SBN:75	0,633	$n_e=2,299$ $n_o=2,312$	$r_{13}=67$	
			$r_{33}=1340$	
KNbO ₃	0,633	$n_1=2,28$ $n_2=2,329$ $n_3=2,169$	$r_{13}=28$	
			$r_{42}=380$	
			$r_{51}=105$	

7.1.2. Электрооптическая модуляция фазы световой волны

Простейшие элементы электрооптической модуляции света представляют собой модуляторы фазы световой волны с продольным и поперечным электрическим полем. В первом из них направления распространения света и внешнего электрического поля совпадают, во втором они ортогональны. Элемент с продольным полем требует прозрачных электродов на оптических гранях кристалла, отчасти поэтому геометрия с поперечным электрическим полем используется наиболее часто [3].



Э

Рис. 7.2. Схема электрооптического модулятора с поперечным электрическим полем. Э – электроды, К – кристалл.

Конструкция модулятора фазы световой волны с поперечным полем поясняется рисунком 7.2. Здесь к электрооптическому кристаллу (например, к кристаллическому образцу из ниобата лития) прикладывается внешнее электрическое поле, перпендикулярное к направлению распространения света. Дополнительная фазовая задержка линейно поляризованной световой волны с необыкновенной поляризацией (вектор \vec{E} параллелен направлению оптической оси кристалла) определяется в этом случае соотношением:

$$\Delta\Phi_e = k_0 \cdot \Delta n \cdot z = -\frac{\pi}{\lambda} n_e^3 r_{33} E_0 \cdot z = -\frac{\pi}{\lambda} n_e^3 r_{33} \frac{U}{d} \cdot L \quad (7.8),$$

где k_0 – волновое число света в вакууме; $E_0 = (U/d)$; U – величина напряжения, приложенного к кристаллу; d – размер кристалла в направлении внешнего электрического поля; L – длина электродов в направлении распространения света. Соответственно, для обыкновенно поляризованной волны величина фазовой задержки равна:

$$\Delta\Phi_o = -\frac{\pi}{\lambda} n_o^3 r_{13} \frac{U}{d} \cdot L \quad (7.9).$$

Достоинством электрооптических модуляторов с поперечным электрическим полем является возможность увеличения электрически индуцированного фазового сдвига за счет увеличения отношения длины электродов к толщине кристалла L/d . Эффективность электрооптических модуляторов разной конфигурации характеризуется величиной полуволнового напряжения, которое определяется как напряжение, необходимое для достижения изменения фазы световой волны π . В данном случае для необыкновенной волны оно определяется соотношением:

$$\frac{\pi}{\lambda} n_e^3 r_{33} \frac{L}{d} \cdot U_{\lambda/2} = \pi \text{ или } U_{\lambda/2} = \frac{\lambda}{n_e^3 r_{33}} \frac{d}{L} \quad (7.10).$$

Для обыкновенной волны получим:

$$U_{\lambda/2} = \frac{\lambda}{n_o^3 r_{13}} \frac{d}{L} \quad (7.11).$$

Следует отметить, что для достижения чисто фазовой модуляции в модуляторе с поперечным полем необходимо, чтобы свет был линейно поляризован, а плоскость поляризации должна быть параллельной либо перпендикулярной оптической оси кристалла.

7.1.3. Пример электрооптического модулятора интенсивности света

Фазовая модуляция световой волны может быть достаточно просто преобразована в модуляцию интенсивности, частоты, поляризации. Известно большое число различных конфигураций электрооптических модуляторов света. В качестве примера, на рис. 7.3 представлена схема электрооптического модулятора интенсивности света в волноводно – оптическом исполнении [11]. Его основа – интерферометр Маха – Цендера на канальной волноводной структуре (КС) в ниобате лития с реализацией

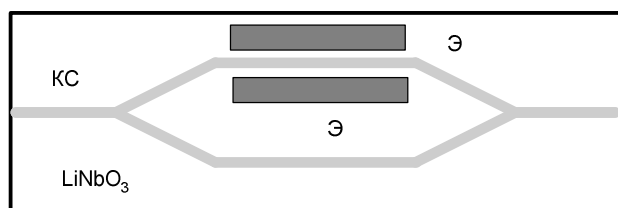


Рис. 7.3.

фазовой модуляции света поперечным электрическим полем в одном из плеч интерферометра. Для формирования КС используется метод термической диффузии

титана в LiNbO_3 при температуре 1000°C . Топология КС

обеспечивается методом

фотолитографии. Свет распространяется по одиночному каналу, затем разветвляется в два плеча интерферометра. Волноводный канал в одном из плеч проходит в зазоре между электродами, к которым приложено электрическое управляющее напряжение. Оно модулирует фазу светового поля в данном канале. После объединения световых потоков во втором Y – разветвителе, модуляция фазы преобразуется в модуляцию интенсивности света. Типичный поперечный размер одномодовых канальных волноводов в ниобате лития для видимого диапазона длин волн составляет около 5 мкм, поэтому глубина модуляции около 100% достигается при величинах управляющего напряжения 5 – 7 В. Для использования максимального электрооптического коэффициента r_{33} направление управляющего электрического поля должно совпадать с направлением оси z кристалла.

7.2. Акустооптический эффект

Акустооптический (АО) эффект – это явление взаимодействия световых волн с акустическими возмущениями, распространяющимися в упругой среде [3, 10]. Физической причиной взаимодействия этих волн является упругооптический (или фотоупругий) эффект, заключающийся в изменении показателя преломления прозрачного материала под действием упругих деформаций (или напряжений).

Очевидно, что по сути фотоупругий эффект сводится также к изменению поляризации среды на частоте световой волны при наличии упругих деформаций кристаллической решетки, т.е. изменения межатомных расстояний или формы элементарной ячейки в кристаллической среде.

При взаимодействии света с акустическими волнами основной интерес представляет эффект дифракции света на акустических возмущениях среды. Этот эффект впервые предсказан Бриллюэном в 1922 г., а обнаружен экспериментально Дебаем и Сирсом в 1932 г. в США и Люка и Бикаром примерно в то же время во Франции. Однако основные работы в этой области выполнены уже после создания лазеров.

АО эффект представляет собой уникальный инструмент для анализа упругих полей в твердых телах и для управления лазерным излучением. Значительный интерес к акустооптическим явлениям и вопросам реализации акустооптических элементов и устройств обусловлен также возможностями их применения в системах оптической обработки информации, в которых подобные элементы осуществляют ввод в систему радиосигнальной информации.

7.2.1. Феноменологическое описание акустооптического эффекта

Фотоупругий эффект выражается в изменении показателя преломления среды при наличии упругих деформаций. Он проявляется при любом состоянии вещества – в газах, жидкостях, твердых телах. Феноменологически он, как и электрооптический эффект, описывается также изменением компонент тензора диэлектрической непроницаемости B_{ij} при наличии упругих деформаций S_{kl} :

$$\Delta B_{ij} = \Delta \left(\frac{1}{n^2} \right)_{ij} = p_{ijkl} S_{kl} \quad (7.12),$$

где S_{kl} – компоненты тензора упругих деформаций, а p_{ijkl} – компоненты фотоупругого тензора или фотоупругие коэффициенты. Таким образом, при наличии поля упругих деформаций, уравнение оптической индикатрисы кристалла определяется выражением:

$$(B_{ij} + p_{ijkl} S_{kl}) x_i x_j = 1 \quad (7.13).$$

Вследствие симметрии тензоров B_{ij} и S_{kl} , возможно также сокращение числа индексов, как и в случае электрооптического эффекта. Соотношение (7.12) при этом примет вид:

$$\Delta \left(\frac{1}{n^2} \right)_i = P_{ij} S_j \quad (7.14),$$

где $i, j=1,2,\dots,6$. Здесь значения индексов $i=1, 2, 3$ соответствуют деформации полуосей эллипсоида (индикатрисы), а $i = 4, 5, 6$ – повороту индикатрисы.

Последний случай, т.е. поворот оптической индикатрисы, можно интерпретировать как возможность изменения поляризации световой волны при АО взаимодействии. Аналогично случаю электрооптического эффекта, можно показать, что выражение для возмущения показателя преломления среды вследствие фотоупругого эффекта в случае отсутствия поворота оптической индикатрисы определяются соотношением:

$$\Delta n_i = -\frac{1}{2} n_i^3 P_{ij} S_j \quad (7.15).$$

Следует отметить, что величины фотоупругих коэффициентов для большинства материалов отличаются незначительно. Они изменяются в диапазоне $p \approx 0,1 \div 0,3$. Так, в таблице 7.2 приведены значения фотоупругих коэффициентов для некоторых изотропных материалов, а в таблице 7.3 для некоторых кристаллов.

Таблица 7.2.

Материал	λ , мкм	p_{11}	p_{12}
Вода	0,63	$\pm 0,31$	$\pm 0,31$
Плавленый кварц (SiO ₂)	0,63	0,121	0,27
Стекло As ₂ S ₃	1,15	0,308	0,299

Таблица 7.3.

Материал	λ , мкм	p_{11}	p_{12}	p_{13}	p_{14}	p_{31}	p_{33}	p_{41}	p_{44}
Al ₂ O ₃	0,644	-0,23	-0,03	0,02	0	-0,04	-0,2	0,01	-0,1
LiNbO ₃	0,633	-0,026	0,09	0,133	-0,075	0,179	0,071	-0,151	0,146
LiTaO ₃	0,633	-0,081	0,081	0,093	-0,026	0,089	-0,044	-0,085	0,028
α – SiO ₂	0,589	0,16	0,27	0,27	-0,03	0,29	0,1	-0,047	-0,079
SBN:75	0,633	0,16	0,1	0,08		0,11	0,47		
TeO ₂	0,633	0,0074	0,187	0,34		0,091	0,24		-0,17

7.2.2. Типы акустических волн в твердом теле

Акустические волны представляют собой распространение упругих возмущений в среде. Они обусловлены смещениями частиц вещества и реакцией на эти смещения окружающих частиц, вследствие наличия сил упругости. Акустические волны могут существовать только в материальной среде. По виду смещений или наличия тех или иных компонент упругих деформаций акустические волны делят на следующие типы.

Продольные волны (волны сжатия).

Смещение частиц в случае продольных волн параллельно направлению распространения волны. Для них характерны деформации сжатия (и

растяжения). Вектор поляризации волны параллелен направлению распространения. Деформации в продольной волне описываются диагональными компонентами S_{ii} .

Поперечные или сдвиговые волны.

Смещение частиц в сдвиговой волне перпендикулярно волновому вектору. Деформации сдвига соответствуют недиагональные компоненты S_{ij} тензора упругих деформаций ($i \neq j$).

Чисто продольные и чисто сдвиговые акустические волны существуют лишь в изотропных средах. При рассмотрении анизотропной среды в общем случае можно говорить лишь о квазипродольных и квазисдвиговых волнах, т.е. вектор упругого смещения в акустической волне, распространяющейся в произвольном направлении, имеет составляющие как в направлении ее волнового вектора, так и в ортогональном к нему направлении. При распространении в среде, обладающей пьезоэффектом, упругие волны могут сопровождаться полем электрического потенциала.

Сказанное относится к упругим волнам в безграничной среде. В полуограниченной среде, вблизи поверхности твердого тела, возможно существование поверхностных волн. Наибольший интерес представляют следующие типы поверхностных волн.

А) Волны Рэлея. В простейшем случае (изотропная среда) для рэлеевской волны характерно как продольное, так и поперечное смещения частиц (с фазовым сдвигом между этими составляющими в 90°). В результате каждая точка в поле такой волны совершает движение по эллиптической траектории. Амплитуда смещения убывает с удалением от поверхности. В случае изотропной среды смещение происходит в сагиттальной плоскости (плоскость, в которой лежат волновой вектор и нормаль к поверхности полупространства). В более общем случае кристаллических сред плоскость эллипса поляризации не обязательно лежит в сагиттальной плоскости, а для пьезокристаллов рэлеевская волна может сопровождаться полем электрического потенциала.

Б) Волны Гуляева-Блюстейна (чисто сдвиговые поверхностные волны). Для данного типа волн смещение частиц параллельно поверхности полупространства и ортогонально направлению распространения волны. Такие волны не существуют на поверхности изотропного твердого тела. Необходимое условие существования чисто сдвиговых поверхностных волн – наличие пьезоэффекта. Подобные решения найдены для некоторых кристаллов кубической и гексагональной симметрии. Особенностью их является более слабая локализация энергии вблизи поверхности, чем для рэлеевских поверхностных волн.

В) Другие типы поверхностных акустических волн могут существовать в тонких пластинах и в системах типа “полупространство – слой”.

7.2.3. Основные представления об акустоопическом взаимодействии.

Пусть в диэлектрической среде распространяется акустическая волна. Тогда за счет фотоупругого эффекта показатель преломления материала оказывается гармонически модулированным в пространстве и времени:

$$n(z, t) = n_s + \Delta n \cdot \cos(\Omega t - Kz) \quad (7.16),$$

где n_s – невозмущенное значение показателя преломления; Δn – амплитудное значение изменения показателя преломления; Ω – частота акустической волны; K – ее волновое число. Таким образом, в рассматриваемой среде существует фазовая решетка с периодом $\Lambda = \frac{2\pi}{K}$, движущаяся вдоль оси Z со

скоростью акустической волны $v_A = \frac{\Omega}{K}$. Поскольку типичная величина скорости акустических волн примерно на 5 порядков меньше скорости света, то для света данная решетка является практически неподвижной. Плоская электромагнитная волна проходит через область среды, возмущенную акустической волной, испытывая дифракцию на периодической фазовой решетке. В результате на выходе данной области световое поле уже не может быть представлено в виде просто плоской волны. Оно состоит из целого набора дифрагированных волн. При анализе дифракции света на акустических волнах выделяют два режима: дифракцию Рамана – Ната и дифракцию Брэгга.

А) Дифракция Рамана-Ната.

Режим дифракции Рамана – Ната наблюдается при малой ширине акустического пучка в направлении распространения световой волны и соответствует дифракции света на тонкой фазовой решетке. В результате фаза прошедшего светового пучка оказывается синусоидально модулированной по его сечению. Соответственно, прошедшее световое поле может быть представлено в виде набора дифрагированных плоских световых волн. Направления распространения дифрагированных волн θ_m определяются соотношением:

$$\sin \theta_m = \sin \theta_0 \pm m \frac{\lambda}{\Lambda} \quad (7.17),$$

где $m=0, \pm 1, \pm 2, \dots$ и т.д.; θ_0 – угол падения световой волны относительно акустического пучка (угол между волновым вектором световой волны и нормалью к волновому вектору акустической волны); λ – длина волны света.

Отметим, что термин «тонкая решетка» можно интерпретировать возможностью пересечения световым лучом только одного фазового фронта

акустической волны. Режим дифракции Рамана – Ната наблюдается, как правило, при низких частотах акустических волн. Поскольку мощность дифрагированного света распределена в нескольких дифракционных максимумах, то такой режим не очень выгоден с энергетической точки зрения.

Б) Дифракция Брэгга.

Этот случай соответствует дифракции света на «толстых решетках», т.е. на акустических пучках, у которых несколько фазовых фронтов пересекаются падающим световым лучом. В этом случае уже нельзя говорить о синусоидальной модуляции фазового фронта светового пучка, прошедшего через акустический пучок. Процесс дифракции света на толстой решетке (широком акустическом пучке) качественно интерпретируется как процесс частичного отражения света от волновых фронтов акустической волны или параллельных слоев, соответствующих максимальной амплитуде акустического поля. Вследствие интерференции многих отраженных слабых пучков, при некотором угле падения света относительно акустического пучка возможна их конструктивная интерференция лишь для одного направления. При этом в прошедшем световом поле присутствуют лишь два дифракционных максимума – максимум нулевого порядка (положение которого соответствует направлению падающего светового пучка) и дифракционный максимум $+1$ или -1 порядка. Угол падения световой волны на акустический пучок (угол между волновым вектором падающей световой волны и фазовым фронтом акустической волны) определяется условием, аналогичным известному условию Брэгга для дифракции рентгеновских лучей на кристаллической решетке:

$$\sin \theta_B = \frac{\lambda}{2\Lambda \cdot n} \quad (7.18)$$

где n – показатель преломления материала. Благодаря этой аналогии, этот угол называют углом Брэгга, а данный режим дифракции – дифракцией Брэгга. Рис. 7.4 иллюстрирует геометрию брэгговского отражения плоской световой волны от структуры с периодом Λ . Очевидно, что отраженные от разных фронтов лучи могут конструктивно интерферировать в том случае, когда их разность хода, определяемая суммой отрезков $AO+OB$, равна длине волны света λ . Таким образом, $AO + OB = 2\Lambda \cdot \sin \theta = \lambda$, откуда получим (7.18). Показатель преломления n в (7.18) учитывает изменение длины волны света в среде.

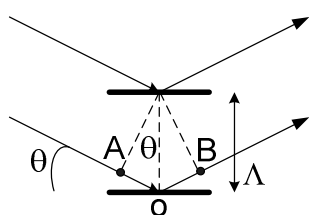


Рис. 7.4.

Дифракция Брэгга наблюдается при высоких частотах акустических волн (определение высоких частот можно истолковать здесь как значения частот, для которых световой луч пересекает хотя бы несколько волновых фронтов акустической волны при

падении на акустический пучок под углом Брэгга). Очевидно, что данный режим более выгоден в энергетическом отношении, чем режим дифракции Рамана – Ната, т.к. дифрагированная волна здесь лишь одна.

7.2.4. Корпускулярная интерпретация акустооптического взаимодействия

Дифракция света на акустической волне может рассматриваться как процесс столкновения фотона (кванта света) с фоном (акустическим квантом). В результате взаимодействия рождается новый фотон, а в процессе взаимодействия должны выполняться законы сохранения энергии и квазиимпульса. Эти законы выражаются соотношениями:

$$\omega_i \pm \Omega = \omega_d \quad (7.19),$$

$$\bar{k}_i \pm \bar{K} = \bar{k}_d \quad (7.20),$$

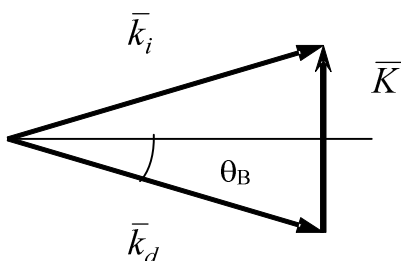
где индексы (i) и (d) соответствуют падающей и дифрагированной световым волнам. Первое из этих выражений показывает, что частота дифрагированной волны отличается от частоты падающей на частоту акустической волны вследствие эффекта Доплера. Второе условие указывает на замкнутость векторного треугольника, построенного из волновых векторов падающей, дифрагированной и акустической волн. Использование данных соотношений и векторных диаграмм дает наглядную интерпретацию возможности разных режимов и особенностей акустооптического взаимодействия в разных материальных средах. Рассмотрим несколько подобных примеров.

Брэгговское акустооптическое взаимодействие плоских волн в оптически изотропной среде.

Для оптически изотропной среды, учитывая, что $\Omega \ll \omega$, можно записать следующее соотношение:

$$|\bar{k}_i| = \frac{\omega_i}{c} \cdot n \approx \frac{\omega_d}{c} \cdot n = |\bar{k}_d| \quad (7.21).$$

Таким образом, векторный треугольник, выражающий закон сохранения квазиимпульса, для оптически изотропной среды является равнобедренным (по крайней мере, при реально рассматриваемых частотах акустических волн, когда различие в частоте падающего и дифрагированного световых пучков не приводит к существенному различию их скоростей в среде вследствие материальной дисперсии). Поэтому легко получить соотношение для угла Брэгга прямо из данного треугольника (рис. 7.5):



$$\sin \theta_B = \frac{|\bar{K}|}{2|\bar{k}_i|} = \frac{2\pi\lambda}{2\Lambda 2\pi n} = \frac{\lambda}{2\Lambda n} = \frac{\lambda}{2v_A n} \cdot f_A \quad (7.22),$$

где f_A – частота акустической волны.

Брэгговское акустооптическое взаимодействие в оптически анизотропной среде

В случае оптически анизотропной среды падающая и дифрагированная световые волны могут иметь разную поляризацию и тогда $|\bar{k}_i| \neq |\bar{k}_d|$. Соответственно, для взаимодействия такого типа векторный треугольник равнобедренным не является, что позволяет наблюдать в оптически анизотропной среде дифракционные процессы, недопустимые для изотропной среды. К таким процессам относится, например, коллинеарная брэгговская дифракция, когда волновые векторы падающей и дифрагированной волн коллинеарны. Из условия (7.16) можно получить выражение для частоты акустических волн, соответствующей такому типу дифракции:

$$\frac{2\pi}{\lambda} \cdot n_i \pm \frac{2\pi}{v_A} \cdot f_A = \frac{2\pi}{\lambda} \cdot n_d \Rightarrow f_A = \frac{v_A}{\lambda} |(n_i - n_d)| \quad (7.23).$$

Векторные диаграммы на рис. 7.6 иллюстрируют геометрию коллинеарного акустооптического взаимодействия и широкополосного акустооптического взаимодействия в оптически анизотропной среде. Для последнего случая при некоторой частоте акустической волны волновой вектор дифрагированной волны оказывается ортогональным волновому вектору акустической волны. Частота, для которой выполняется это условие, выбирается в качестве центральной частоты акустооптических модуляторов и дефлекторов.

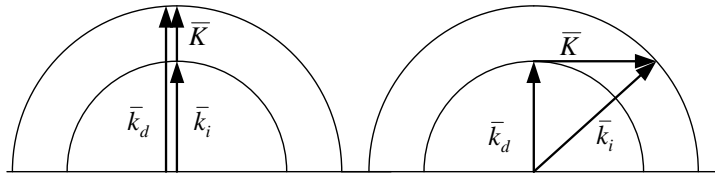


Рис. 7. 6.

7.2.5. Иллюстрация векторными треугольниками

дифракции Рамана – Ната и Брэгга

Выше отмечена роль ширины акустического пучка в направлении распространения света при разных режимах дифракции. Этот параметр определяет дифракционную расходимость акустического пучка на рассматриваемой частоте. При малой ширине пучка дифракционная расходимость велика, т.е. велика неопределенность направления акустического волнового вектора. Соответственно, становятся возможными процессы многократной последовательной дифракции света на акустической волне, которым удовлетворяют допустимые направления акустического волнового вектора, диапазон которых определяется расходимостью акустического пучка. Количественно для этого необходимо, чтобы угловая расходимость акустического пучка была больше, чем угол между отдельными дифрагированными пучками:

$$\frac{\Lambda}{L} > \frac{\lambda}{2n\Lambda} \quad (7.24).$$

Это условие принято записывать в следующем виде:

$$\frac{\lambda L}{2n\Lambda^2} < 1 \quad (7.25).$$

При большой ширине акустического пучка, когда его дифракционная расходимость значительно меньше величины угла Брэгга, процессы многократной дифракции не могут происходить. Это соответствует режиму брэгговской дифракции, для которого должно выполняться условие:

$$\frac{\lambda L}{2n\Lambda^2} \gg 1 \quad (7.26).$$

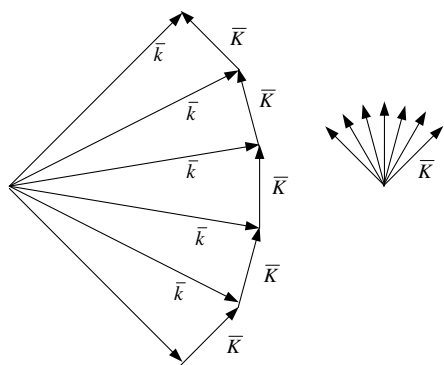


Рис. 7.7.

необходимую компоненту в угловом спектре акустического поля для очередного акта дифракции.

Очевидно, что граница между данными режимами дифракции не может быть резкой, существует переходной режим, для которого вероятность многократной дифракции отлична от нуля. Возможность возникновения многих дифракционных максимумов при дифракции света на узком акустическом пучке иллюстрирует схема на рис. 7.6. Здесь при заданном направлении падающей световой волны, при ее дифракции на акустических волнах каждая дифрагированная волна “находит”

7.2.6. Эффективность дифракционных процессов.

В случае дифракции Рамана-Ната, как отмечалось, световой пучок, прошедший через область, занятую акустической волной, имеет синусоидальную фазовую модуляцию амплитудного фронта. В результате, величина дифракционной эффективности (отношение интенсивности дифрагированного пучка к интенсивности падающего) для максимума порядка m может быть найдена из соотношения:

$$\eta_m = J_m^2(\delta) = J_m^2\left(\frac{2\pi L \cdot \Delta n_0}{\lambda \cdot \cos \theta}\right) \quad (7.27),$$

где J_m – функция Бесселя.

В случае дифракции Брэгга, дифракционная эффективность определяется соотношением:

$$\eta = \frac{I_d}{I_0} = \sin^2 \xi \quad (7.28),$$

где
$$\xi^2 = \left(\frac{\pi \cdot \cos \theta \cdot L}{\lambda_0}\right)^2 \frac{M_2}{2} P_a \quad (7.29).$$

В (7.25) P_a – акустическая мощность, L – ширина акустического пучка, а M_2 – постоянная, называемая коэффициентом акустооптического качества материала. Величина коэффициента M_2 определяется физическими свойствами упругой среды:

$$M_2 = \frac{n^6 p^2}{\rho V^3} \quad (7.30),$$

где ρ – плотность вещества, V – скорость упругой волны в заданном направлении. Для кристаллических материалов их физические свойства могут изменяться достаточно сильно при изменении ориентации и поляризации взаимодействующих световых и акустических волн. В результате и величина коэффициента M_2 обладает сильной анизотропией и существенно отличается для разных материалов. В таблице 7.4 приведены значения физических характеристик и коэффициента M_2 для некоторых ориентаций наиболее интересных акустооптических материалов. Приведенные значения коэффициента M_2 нормированы относительно данной величины для плавленого кварца, поэтому для них не указана размерность. Можно видеть, что некоторые материалы “лучше” кварца по акустооптическим свойствам почти в 1000 раз (например, TeO_2). Это обусловлено, в первую очередь, очень низкой скоростью сдвиговой акустической волны в некоторых направлениях в данном кристалле.

Таблица 7.4

Материал	λ , мкм	n	ρ , г/см ³	Поляр. и напр. ак. волны	V , 10 ³ м/с	Поляр. и напр. света	M_2
----------	-----------------	-----	-------------------------------	--------------------------------	------------------------------	----------------------------	-------

LiNbO ₃	0,63	2,2	4,7	Прод. в [1120]	6,57		6,99
As ₂ S ₃	0,63	2,61	3,2	Продольная	2,6	433	433
TeO ₂	0,63		5,99	[$\bar{1}$ 10] в [110]	0,617	Произв. в [001]	793
SBN:75	0,63	2,3		Прод. в [001]		[001] поляр.	38,6

7.2.7. Акустооптический модулятор.

АО эффект широко используется в элементах управления параметрами лазерного излучения. Базовый АО элемент называют акустооптическим модулятором, хотя он может выполнять функции дефлектора излучения, а также рассматриваться как преобразователь частоты света. Не останавливаясь на особенностях различных АО элементов и их характеристиках, рассмотрим основную схему АО модулятора (рис. 7.8).

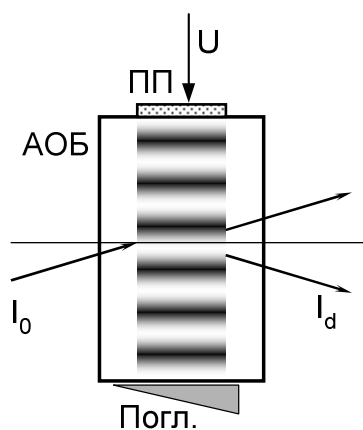


Рис. 7. 8.

Здесь на блоке из акустооптического материала (АОБ) размещен пьезопреобразователь (ПП), на который подается электрический сигнал, преобразуемый в бегущую акустическую волну. На противоположной поверхности блока размещен акустический поглотитель (Погл.), препятствующий возникновению в АОБ стоячих акустических волн. Световой пучок с интенсивностью I_0 падает на акустический пучок под углом Брэгга. Световое поле на выходе АО модулятора состоит из недифрагированного и дифрагированного световых пучков. Величина угла дифракции

$$\sin \theta = \frac{\lambda}{v} \cdot f$$

определяется соотношением $\sin \theta = \frac{\lambda}{v} \cdot f$, т.е. при изменении частоты управляющего сигнала АО модулятор играет роль дефлектора. Частота дифрагированного излучения отличается от частоты падающего на величину f , что используется в устройствах сдвига частоты света.

7.3. Фоторефрактивный эффект

Фоторефрактивный эффект (ФРЭ) или эффект фоторефракции заключается в изменении показателя преломления диэлектрических или полупроводниковых материалов под действием света. Впервые он был обнаружен в 1965 году в электрооптических кристаллах (в ниобате лития)

при проведении экспериментов в области нелинейной оптики. В подобных кристаллах с течением времени наблюдался эффект искажения пространственной структуры световых пучков, и значительно возростала интенсивность рассеяния излучения. Было ясно, что причиной данных явлений служили индуцированные световым полем в кристалле оптические неоднородности. Сначала данный эффект рассматривался как паразитное явление, что послужило поводом для относящегося к нему термина “optical damage” – “оптическое повреждение”. Затем было выявлено, что эффект наведенных оптических неоднородностей обладает свойством реверсивности, поскольку оптическая однородность кристаллов восстанавливалась в результате нагревания до температуры около 170°C либо их экспозиции в течение некоторого времени при однородном освещении. Такая реверсивность указывала на возможность использования эффекта фоторефракции в качестве механизма голографической записи в кристаллах ниобата лития и некоторых других электрооптических кристаллах. Дальнейшие исследования показали, что фоторефрактивный эффект проявляется, если в рассматриваемой среде развиваются следующие элементарные процессы [12].

1. Возбуждение свободных носителей электрического заряда с донорных примесных центров в запрещенной зоне под действием света, что приводит к появлению электронов в зоне проводимости, либо фотовозбуждения валентных электронов с их захватом на примесные акцепторные уровни, что приводит к появлению дырок в валентной зоне. Для излучения ультрафиолетового диапазона возможны также прямые межзонные переходы электронов вследствие фотоионизации.
2. Пространственное перераспределение носителей заряда вследствие тепловой диффузии, либо их дрейфа в приложенном к кристаллу внешнем электрическом поле. Дрейф носителей может быть обусловлен также фотогальваническим (фотовольтаическим) эффектом, который ранее ассоциировался с наличием в кристалле неких внутренних электрических полей.
3. Захват носителей заряда в неосвещенных областях кристалла ловушками (в качестве которых могут выступать, например, ионизированные донорные центры) и формирование, вследствие этого, неоднородного поля пространственного электрического заряда E_{sc} .
4. Модуляция показателя преломления кристалла полем E_{sc} вследствие линейного электрооптического эффекта.

Схема, иллюстрирующая процесс разделения зарядов при фотовозбуждении электрона с донорного уровня (D) с последующим его транспортом в зоне проводимости и захватом на ловушку (D^+) представлена на рис. 7.9.

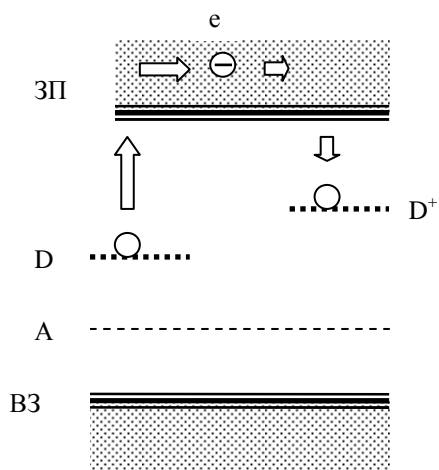


Рис. 7.9.

Фоторефрактивные свойства электрооптических кристаллов существенно зависят от наличия в них различных легирующих примесей. К подобным активным примесям, в первую очередь, относятся ионы некоторых переходных элементов, склонные к перезарядке (*Fe, Cr, Cu, Rh* и т.д.). В номинально чистых кристаллах проявление ФРЭ может быть обусловлено присутствием неконтролируемых примесей, либо дефектами кристаллической решетки.

7.3.1. Механизмы пространственного перераспределения носителей электрического заряда

Как отмечено, пространственное перераспределение носителей электрического заряда может быть обусловлено процессами тепловой диффузии, дрейфа носителей во внешнем электрическом поле и фотовольтаическим эффектом [12]. Во многих случаях в суммарный эффект вносят вклад все или несколько из этих механизмов. В общем случае, плотность тока проводимости, обусловленного движением носителей заряда, определяется соотношением:

$$\bar{\delta}_{\Sigma} = \bar{\delta}_{drift} + \bar{\delta}_{dif} + \bar{\delta}_{phv} \quad (7.31),$$

где отдельные слагаемые соответствуют вкладам дрейфового ($\bar{\delta}_{drift}$), диффузионного ($\bar{\delta}_{dif}$) и фотовольтаического ($\bar{\delta}_{phv}$) механизмов.

а) Дрейфовый механизм.

В случае дрейфа носителей электрического заряда (электронов или дырок) под действием приложенного к кристаллу внешнего электрического поля \bar{E}_0 , свободные носители перемещаются из освещенных областей кристалла в неосвещенные, где захватываются глубокими ловушечными центрами. Плотность дрейфового тока определяется законом Ома:

$$\bar{\delta}_{drift} = \hat{\sigma} \cdot \bar{E} \quad (7.32),$$

$$\hat{\sigma} = e \cdot \hat{\mu} \cdot n_e \quad (7.33),$$

где $\hat{\sigma}$ - тензор проводимости; \bar{E} - локальное электрическое поле внутри кристалла, учитывающее вклад внешнего поля \bar{E}_0 и поля пространственного

заряда \bar{E}_{SC} ; $\hat{\mu}$ - тензор подвижности носителей заряда; n_e – концентрация подвижных носителей. В общем случае подвижность носителей различна для направлений разных осей кристаллов, поэтому при непараллельности внешнего электрического поля и кристаллофизических осей возможно возникновение компонент дрейфового тока, ортогональных направлению поля \bar{E} .

Перераспределение носителей продолжается до тех пор, пока наведенное поле пространственного заряда \bar{E}_{SC} не скомпенсирует в освещенных областях приложенное внешнее поле \bar{E}_0 . Предельное изменение показателя преломления при этом определяется величиной внешнего поля E_0 . Для большинства известных фоторефрактивных кристаллов величина внешнего поля может достигать $(2\div 3) \cdot 10^4$ В/см.

Для дрейфового механизма пространственного перераспределения носителей заряда характерно формирование несмещенных фоторефрактивных решеток. В этом случае отклик фоторефрактивной среды является локальным, т.е. пространственное распределение изменения показателя преломления повторяет картину пространственного распределения интенсивности светового поля.

б) Тепловая диффузия носителей заряда.

Распределение концентрации фотовозбужденных носителей в пространстве является неоднородным при неоднородности интенсивности возбуждающего светового поля. Вследствие эффекта тепловой диффузии возникает диффузионный ток, локальная величина которого пропорциональна градиенту концентрации носителей в данной точке пространства. Диффузия приводит к переносу зарядов из освещенных областей в неосвещенные. Носители заряда захватываются ловушками, что приводит к возникновению наведенного поля пространственного заряда E_{sc} . Стационарное состояние достигается, когда диффузионный ток компенсируется встречным током свободных носителей заряда, вызываемым наведенным полем E_{sc} .

Плотность диффузионного тока определяется соотношением:

$\delta_{dif} = eD \cdot \frac{\partial n}{\partial z}$, где D – величина коэффициента диффузии носителей заряда, а

e – заряд электрона. Диффузионный ток движения электронов в области их неоднородного распределения обусловлен тепловой энергией среды, его можно представить как результат движения электронов в некотором эффективном электрическом поле, величина которого определяется соотношением:

$$E_D(z) = \frac{k_B T}{e} \cdot \frac{1}{n(z)} \cdot \frac{dn(z)}{dz} \quad (7.34).$$

Здесь k_B – постоянная Больцмана, T – температура, $n(z)$ – локальная концентрация электронов, $dn(z)/dz$ – градиент концентрации электронов. Наведенное поле \bar{E}_{sc} и поле \bar{E}_D направлены навстречу друг другу. При выполнении условия $|\bar{E}_{sc}| = |\bar{E}_D|$ достигается стационарный режим.

в) Фотовольтаический эффект

Еще один механизм транспорта электрических зарядов обусловлен фотогальваническим (фотовольтаическим) эффектом. Суть данного эффекта заключается в том, что в некоторых кристаллах без центра симметрии, при их однородном освещении, вдоль полярной оси кристалла возникает стационарный ток, плотность которого пропорциональна интенсивности света.

$$|\bar{\delta}_{phv}| = G \alpha I ,$$

где G – постоянная Гласса; α - оптическое поглощение, I – интенсивность света.

При неоднородном освещении плотность фотовольтаического тока зависит от пространственной координаты, а величина наведенного поля

пространственного заряда определяется соотношением: $E_{sc} = \frac{|\bar{\delta}_{phv}|}{\sigma}$,

где σ - проводимость кристалла, в которую вносят вклад как темновая проводимость, так и фотопроводимость среды ($\sigma = \sigma_d + \sigma_{ph} \cdot I$); σ_{ph} – постоянная фотопроводимости среды. Считается, что механизм возникновения фотовольтаического тока связан с асимметрией элементарных процессов фотовозбуждения и рекомбинации носителей заряда. В таких материалах, как $LiNbO_3:Fe$, $LiNbO_3:Cu$ фотовольтаический эффект является доминирующим, а величина поля E_{sc} , обусловленного данным эффектом, может достигать значений до $\approx 10^5$ В/см.

7.3.2. Модель фоторефракции

Наиболее популярная модель фоторефрактивного эффекта представляет кристалл в виде компенсированного полупроводника. Она предполагает наличие донорных и акцепторных уровней в запрещенной зоне кристалла. Донорные уровни заполнены лишь частично. Ионизированные доноры с концентрацией N_D^+ играют роль центров захвата (ловушек). Постулируется наличие акцепторных центров с концентрацией $N_A = N_D^+$. В фотопереходах они не участвуют (они полностью заполнены электронами).

Для описания ФРЭ используется следующая система уравнений.

1. Кинетическое уравнение, определяющее баланс заряженных частиц при фотовозбуждении электронов с донорных центров:

$$\frac{\partial N_D^+}{\partial t} = (\beta + s \cdot I_{em}) \cdot (N_D - N_D^+) - \gamma \cdot n \cdot N_D^+ \quad (7.35).$$

2. Уравнение Пуассона:

$$\frac{\partial}{\partial z} (\epsilon_0 \epsilon_r E_z) = \rho \quad (7.36).$$

3. Уравнение непрерывности:

$$\frac{\partial J_z}{\partial z} + \frac{\partial \rho}{\partial t} = 0 \quad (7.37).$$

4. Соотношение для объемной плотности электрического заряда:

$$\rho = e \cdot (N_D^+ - N_A - n) \quad (7.38).$$

5. Суммарная плотность тока в кристалле:

$$J_z = e \mu n \cdot E_z + \mu k_B T \frac{\partial n}{\partial z} + \beta_{ph} (N_D - N_D^+) \cdot I_{em} \quad (7.39).$$

Здесь β – коэффициент термической генерации электронов; s – сечение фотоионизации; I_{em} – интенсивность света; N_D – концентрация доноров; γ – коэффициент рекомбинации; n – концентрация электронов в зоне проводимости; μ – подвижность электронов; β_{ph} – фотовольтаический коэффициент; J_z , E_z – составляющие векторов плотности тока и напряженности электрического поля внутри кристалла вдоль оси z . Здесь рассматривается одномерный случай, когда все переменные зависят только от этой координаты.

В результате решения данной системы можно найти величину поля пространственного заряда E_{sc} и, соответственно, величину изменения показателя преломления среды, индуцированного вследствие линейного электрооптического эффекта. Однако в большинстве случаев решение удается получить лишь при упрощающих предположениях: $n \ll N_A \ll N_D$.

7.3.3. Схема формирования элементарных фоторефрактивных решеток.

Рассмотрим кинетику записи и релаксации элементарных фоторефрактивных решеток в простейшем случае, когда учитывается только

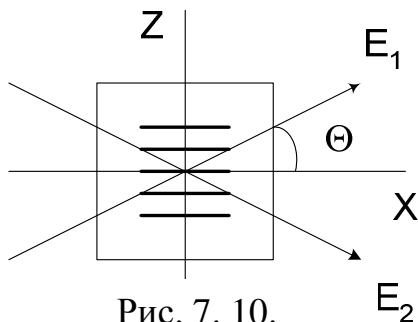


Рис. 7. 10.

один механизм пространственного перераспределения носителей заряда. Такая ситуация может быть реализована в материалах с выраженным фотовольтаическим эффектом при отсутствии приложенного внешнего электрического поля и при периодах интерференционной картины более 5 мкм.

Пусть две плоские световые волны распространяются в фоторефрактивном кристалле под малыми углами θ к оси X (рис. 7. 10). При поляризации волн в плоскости XOZ и малых углах пересечения можно записать их поля в скалярном приближении в виде:

$$\dot{E}_1 = E_{01} \cdot \exp[i(\omega t - k \cdot \cos \theta \cdot x - k \cdot \sin \theta \cdot z)] \quad (7.40),$$

$$\dot{E}_2 = E_{02} \cdot \exp[i(\omega t - k \cdot \cos \theta \cdot x + k \cdot \sin \theta \cdot z)] \quad (7.41).$$

Пересекаясь, световые волны формируют интерференционную картину, поле которой находим в виде суммы:

$$\dot{E}_\Sigma = \dot{E}_1 + \dot{E}_2$$

Распределение интенсивности в данной картине:

$$\begin{aligned} I_\Sigma &= (\dot{E}_1 + \dot{E}_2) \cdot (\dot{E}_1^* + \dot{E}_2^*) = \dot{E}_1 \dot{E}_1^* + \dot{E}_2 \dot{E}_2^* + \dot{E}_1 \dot{E}_2^* + \dot{E}_1^* \dot{E}_2 = \\ &= I_1 + I_2 + 2\sqrt{I_1 \cdot I_2} \cdot \cos(2k \cdot \sin \theta \cdot z) \end{aligned}$$

или $I_\Sigma = I_1 + I_2 + 2\sqrt{I_1 \cdot I_2} \cdot \cos(K \cdot z)$, где $K = \frac{2\pi}{\Lambda} = 2k \cdot \sin \theta$,

а Λ - период решетки.

В области интерференционной картины, вследствие неоднородного распределения интенсивности света $I(z)$, возникает фотовольтаический ток с плотностью:

$$\delta_{ph} = G \cdot \alpha \cdot I(z) \quad (7.42).$$

Это приводит к пространственному перераспределению электрических зарядов и возникновению электрического поля E_{sc} . Наличие этого поля, в свою очередь, приводит к появлению тока проводимости с локальной плотностью $\delta_c = \sigma \cdot E_{sc}$ и тока смещения $\delta_s = \varepsilon \frac{\partial E_{sc}}{\partial t}$.

Вкладом механизма диффузии, как отмечалось, пренебрегаем, так как его следует учитывать при $\Lambda < 5 \div 2$ мкм. Это обусловлено обратной пропорциональной зависимостью величины диффузионного поля и пространственного периода световой интерференционной картины, что иллюстрируется соотношением (7.34).

При формировании фоторефрактивных решеток возможны два режима:

а) короткозамкнутый кристалл;

б) разомкнутый кристалл.

Рассмотрим второй случай, который ближе к реальности. Для разомкнутого кристалла ток проводимости в цепи отсутствует, т.е. выполняется условие: $\delta_{полн} = 0$, что приводит к дифференциальному уравнению:

$$\sigma E_{sc} + \varepsilon \frac{\partial E_{sc}}{\partial t} + G\alpha I = 0, \text{ или:}$$

$$\frac{dE_{sc}}{dt} + \frac{\sigma}{\varepsilon} E_{sc} = -G\alpha \frac{I}{\varepsilon} \quad (7.43).$$

Общее решение однородного уравнения $\frac{dE_{sc}}{dt} + \frac{\sigma}{\varepsilon} E_{sc} = 0$ имеет вид:

$$E_{sc} = A \cdot \exp\left(-\frac{\sigma}{\varepsilon} t\right) + C \quad (7.44).$$

Для определения констант интегрирования A и C используем условия, определяющие величину поля E_{sc} в начальный момент времени $t=0$ после «включения» светового поля, и значение E_{sc} в стационарном режиме. Первое из них – $E_{sc}(0)=0$ дает: $A + C = 0$, откуда получаем: $A = -C$. Итак, решение для E_{sc} теперь имеет вид:

$$E_{sc} = C \left[1 - \exp\left(\frac{\sigma}{\varepsilon} t\right) \right] \quad (7.45).$$

Стационарное значение E_{sc}^{st} достигается при условии $\frac{\partial}{\partial t} = 0$. При этом:

$$\frac{\sigma}{\varepsilon} E_{sc}^{st} = -G\alpha \frac{I}{\varepsilon} \quad \text{или} \quad E_{sc}^{st} = -\frac{G\alpha I}{\sigma}. \quad \text{Очевидно, что } E_{sc}^{st} = E(\infty) = C. \quad \text{Итак,}$$

окончательно получаем выражение для величины поля E_{sc} в виде:

$$E_{sc} = -\frac{G\alpha I}{\sigma} \cdot \left[1 - \exp\left(-\frac{t}{\tau}\right) \right] \quad (7.46),$$

где $\tau = \frac{\varepsilon}{\sigma}$ – время Максвелловской релаксации электрического заряда в среде.

Данное соотношение описывает временную эволюцию наведенного поля пространственного заряда E_{sc} в процессе записи фоторефрактивной решетки в среде с фотовольтаическим механизмом транспорта носителей заряда при сделанных предположениях. Как уже отмечено, удельная проводимость среды характеризуется двумя составляющими – темновой проводимостью σ_d и фотопроводимостью $\sigma_{ph} \cdot I$. Соответственно, можно видеть, что стационарное значение поля E_{sc} зависит от соотношения этих составляющих. При высокой темновой проводимости ($\sigma_d \gg \sigma_{ph} I$) из (7.46) получим:

$$E_{sc} = -\frac{G\alpha I}{\sigma_d} \quad (7.47).$$

При пренебрежимо малой величине темновой проводимости ($\sigma_d \ll \sigma_{ph} I$) имеем:

$$E_{sc} = -\frac{G\alpha}{\sigma_{ph}} \quad (7.48),$$

т.е. в этом случае стационарное значение поля пространственного заряда не зависит от величины интенсивности света.

7.3.4. Стирание (релаксация) решетки

Представляет интерес временная зависимость величины поля пространственного заряда фоторефрактивной решетки в процессе ее темновой релаксации, поскольку это определяет потенциально достижимое время хранения голографической информации в фоторефрактивных голографических средах. Очевидно, релаксация поля E_{sc} описывается тем же уравнением (7.43), но без правой части, т.е. при $I=0$. Как отмечалось, общее решение данного однородного уравнения имеет вид:

$$E_{sc}(t) = A \cdot \exp\left[-\frac{\sigma}{\varepsilon}t\right] + C, \text{ где для определения постоянных } A \text{ и } C \text{ используем}$$

условия:

$E(0) = E_0$; $E(\infty) = 0$, где E_0 – величина поля E_{sc} в момент «выключения» света. Из второго условия вытекает соотношение $C=0$, а первое дает: $A=E_0$. Таким образом, процесс темновой релаксации решетки определяется зависимостью:

$$E_{sc}(t) = E_0 \cdot \exp\left(-\frac{t}{\tau}\right) \quad (7.49).$$

Как отмечалось, изменение показателя преломления среды под действием света при фоторефракции обусловлено линейным электрооптическим эффектом. Соответственно, временная зависимость амплитуды элементарной фоторефрактивной решетки $n(t, z)$ может быть представлена в форме:

$$\begin{aligned} n(t, z) &= n_0 - \frac{1}{2}n^3 r E_{sc} = n_0 - \frac{1}{2}n^3 r G \alpha \cdot \frac{I}{\sigma} \left[1 - \exp\left(-\frac{t}{\tau}\right)\right] = \\ &= n_0 - \Delta n_{sc} \left[I_1 + I_2 + 2\sqrt{I_1 I_2} \cdot \cos(Kz) \right] \cdot \left[1 - \exp\left(-\frac{t}{\tau}\right)\right] = \\ &= n_0 - \Delta n_{sc} (I_1 + I_2) \left[1 - \exp\left(-\frac{t}{\tau}\right)\right] + \\ &+ \Delta n_{sc} \cdot 2\sqrt{I_1 I_2} \cos(Kz) \left[1 - \exp\left(-\frac{t}{\tau}\right)\right] \end{aligned} \quad (7.50).$$

Из приведенного соотношения видно, что возмущенная часть показателя преломления среды состоит из двух составляющих: постоянной, не зависящей от координаты z , и переменной, с гармонической зависимостью от этой координаты.

7.3.5. Некоторые характеристики ФРЭ и фоторефрактивных материалов

Важными характеристиками фоторефрактивных кристаллов являются фоторефрактивная чувствительность, дифракционная эффективность и время записи – релаксации ФР решеток, определяющее, в конечном итоге, время хранения голографической информации.

Понятие **чувствительности** характеризует величину световой энергии, необходимой для записи информации в фоторефрактивном кристалле. Она определяется следующим образом:

$$S_h = \sqrt{\eta} / W_l m \quad (7.51),$$

где W_l – полная энергия светового поля, необходимая для записи одиночной решетки на площади в 1 см^2 ; m – видность (контраст) интерференционной картины. Голографическая чувствительность измеряется в $\text{см}^2/\text{Дж}$.

Дифракционная эффективность фоторефрактивных решеток определяется формулой Когельника:

$$\eta = \sin^2 \left(\frac{\pi L \Delta n}{\lambda \cos \theta} \right) \quad (7.52),$$

где L – толщина решетки; Δn – максимальное изменение показателя преломления.

Время записи – стирания ФР решеток является также очень важной характеристикой, поскольку оно определяет возможности материалов для их использования в качестве среды для долговременного хранения голографической информации или для нужд динамической голографии, где требуется как можно более быстрый отклик на изменения светового поля.

Некоторые характеристики для наиболее перспективных фоторефрактивных материалов представлены в таблице 7.5.

Таблица 7.5.

Материал	λ , мкм	r , пм/В	n	τ , с
LiNbO ₃	0,6	$r_{33}=31$	$n_e=2,2$	0,1 ÷ годы
LiTaO ₃	0,6	$r_{33}=31$	$n_e=2,2$	1 ÷ годы
BaTiO ₃	0,5	$r_{42}=1640$	$n_e=2,4$	$2 \cdot 10^{-3}$
SBN	0,5	$r_{33}=1340$	$n_e=2,3$	$5 \cdot 10^{-3}$
GaAs	1,1	$r_{12}=1,3$	$n_e=3,4$	$30 \cdot 10^{-6}$
BSO	0,6	$r_{41}=5$	$n_e=2,54$	$2 \cdot 10^{-3}$
KNbO ₃	0,6	$r_{42}=380$	$n=2,3$	

8. ЭЛЕМЕНТЫ НЕЛИНЕЙНОЙ ОПТИКИ

Как известно, материальная среда проявляет нелинейные свойства в отношении некоторого воздействия, если ее физические характеристики зависят от величины этого воздействия. Соответственно, оптическая нелинейность наблюдается при зависимости диэлектрической проницаемости материалов от интенсивности света. В долазерную эпоху нелинейно–оптические эффекты в экспериментах практически не наблюдались (если не считать линейный и квадратичный электрооптические эффекты, которые, как будет отмечено далее, могут быть отнесены к нелинейно–оптическим явлениям). Было известно, что световые пучки в вакууме проходят друг сквозь друга, без какого – либо взаимодействия. Этот факт рассматривался как бесспорная иллюстрация применимости принципа суперпозиции, являющегося следствием линейности уравнений классической оптики, к рассмотрению световых волн. С квантовых позиций отсутствие взаимодействия световых пучков свидетельствовало о невозможности непосредственного взаимодействия фотонов. Предполагалось, что принцип суперпозиции остается справедливым и при распространении света в материальной среде.

Линейность оптических явлений в долазерный период закономерна, поскольку для известных некогерентных источников света величина напряженности электрического поля не превышает $10^2 \div 10^3$ В/м. Это значительно меньше напряженности внутриатомных электрических полей ($10^9 \div 10^{12}$ В/м) и недостаточно для заметного изменения поляризуемости атомов вещества полем световой волны. Для когерентного лазерного излучения величина напряженности электрического поля достигает $10^9 - 10^{10}$ В/м, что уже сопоставимо с напряженностью внутриатомных полей. А это означает, что при прохождении световой волны такой амплитуды через вещество, его поляризуемость и, соответственно, высокочастотная диэлектрическая проницаемость, могут изменяться – они зависят от амплитуды воздействующего светового поля. Среда является в этом случае нелинейной, а уравнения электромагнитного поля также становятся нелинейными [3, 5-8, 13].

Появление мощных лазеров позволило наблюдать и изучить целый ряд нелинейно–оптических явлений и стимулировало развитие нелинейной оптики. Таким образом, возможность проявления нелинейно–оптических свойств материала зависит от величины интенсивности света. Одна и та же среда может рассматриваться как линейная при малых интенсивностях света и как нелинейная – при больших.

8.1. Элементы теории нелинейно – оптических явлений

8.1.1. Квадратичная и кубичная оптические нелинейности.

Рассмотрим диэлектрическую восприимчивость среды χ . Из курса электродинамики известно, что она связывает величину поляризации среды, индуцированной полем электромагнитной волны, с величиной напряженности этого поля. В случае изотропного материала эта связь определяется соотношением:

$$\bar{P} = \chi \cdot \bar{E} \quad (8.1),$$

где χ – постоянная скалярная величина. В общем случае оптически анизотропной среды коэффициент χ связывает между собой два вектора, поэтому он является тензором второго ранга. Для понимания различия между нелинейностями разного типа в рамках данного описания тензорный характер χ не принимается во внимание и далее, если это не приводит к принципиальным ошибкам, математические выражения записываются для простоты в скалярной форме [5, 13].

Соотношение (8.1) – одно из классических материальных уравнений. Если коэффициент χ не зависит от напряженности светового поля, то рассматриваемая среда является линейной. Для лазерного излучения, в случае высокой его интенсивности, невозможно ограничиться только линейным приближением. Выражение для диэлектрической восприимчивости χ должно быть представлено в этом случае в более общем виде, учитывающем ее зависимость от напряженности электрического поля световой волны:

$$\chi(E) = \xi_0 + \eta \cdot E + \theta \cdot E^2 + \dots \quad (8.2),$$

где $\xi_0, \eta, \theta, \dots$ – параметры среды, характеризующие вид функциональной зависимости ее поляризуемости от напряженности поля световой волны. При этом соотношение (8.1) принимает форму:

$$P(E) = \chi(E) \cdot E = \xi_0 \cdot E + \eta \cdot E^2 + \theta \cdot E^3 + \dots \quad (8.3).$$

Таким образом, отклик среды (величина P) на поле световой волны (E) оказывается нелинейным. Коэффициенты η и θ называют нелинейными восприимчивостями (квадратичной и кубичной, соответственно), а среды, для которых соответствующие слагаемые в (8.3) доминируют, называют средами с квадратичной или кубичной оптической нелинейностью. Нелинейностями более высокого порядка в разложении (8.2), как правило, пренебрегают, в силу малости каждого последующего члена по сравнению с предыдущим. Важно отметить, что квадратичная оптическая нелинейность может проявляться лишь в средах без центра симметрии. Действительно, в

центросимметричной среде диэлектрическая восприимчивость $\chi(E)$ не должна измениться при замене направления поля \vec{E} на противоположное, т.е. при изменении знака E . Поскольку $\chi(E) = \xi_0 + \eta \cdot E + \theta \cdot E^2$, то выполнение последнего условия возможно лишь при $\eta = 0$. Таким образом, в центросимметричной среде низшим типом оптической нелинейности является кубичная нелинейность.

Следует отметить, что под термином «нелинейная оптика» понимается совокупность явлений, в основе которых лежат обратимые процессы изменения характеристик среды. Подобные изменения исчезают после «выключения» света. Таким образом, эффекты необратимых изменений среды, проявляющихся также в сверхсильных световых полях (например, оптический пробой) к области нелинейной оптики не относят, а рассматривают в отдельных разделах физики.

8.1.2. Возможные типы нелинейно–оптических эффектов в средах с квадратичной и кубичной нелинейностью

А. Среда с квадратичной нелинейностью.

В этом случае $P(E) = \xi_0 \cdot E + \eta \cdot E^2$. Положим, что в такой среде распространяется плоская монохроматическая световая волна с полем:

$$E = E_0 \cos(\omega t - kz) \quad (8.4).$$

Поляризация, наводимая световой волной в среде, определяется соотношением:

$$P(E) = \xi_0 E_0 \cos(\omega t - kz) + \eta \cdot E_0^2 \cos^2(\omega t - kz) = \xi_0 E_0 \cos(\omega t - kz) + \eta \cdot \frac{E_0^2}{2} + \eta \cdot \frac{E_0^2}{2} \cdot \cos(2\omega t - 2kz) \quad (8.5).$$

Таким образом, квадратичная нелинейность приводит к возникновению составляющих поляризации на удвоенной частоте светового поля, а также на нулевой частоте. Постоянная поляризация дает возможность оптического детектирования светового поля, а поляризация на удвоенной частоте – возможность генерации его второй гармоники. Типы возможных нелинейно – оптических эффектов в среде с квадратичной нелинейностью иллюстрирует рис. 8.1, где представлены частотные составляющие поляризации квадратично – нелинейной среды при воздействии гармонического светового поля.

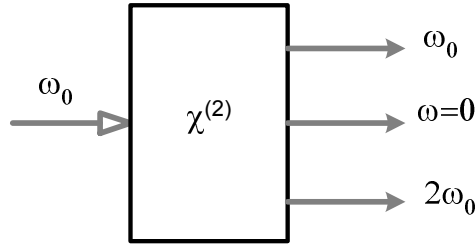


Рис. 8.1.

Если в материале с квадратичной нелинейностью распространяются две световые волны с разными частотами, то появляется возможность генерации световых полей с комбинационными частотами – суммарной и разностной. Действительно, пусть имеем $E_1 = E_{01} \cdot \cos(\omega_1 t - k_1 z)$, $E_2 = E_{02} \cdot \cos(\omega_2 t - k_2 z)$. Тогда выражение для наведенной поляризации получим в форме:

$$\begin{aligned}
 P(E_1 + E_2) &= \xi_0 \cdot [E_{01} \cos(\omega_1 t - k_1 z) + E_{02} \cos(\omega_2 t - k_2 z)] + \\
 &\eta \cdot [E_{01} \cos(\omega_1 t - k_1 z) + E_{02} \cos(\omega_2 t - k_2 z)]^2 = \\
 &= \xi_0 \cdot E_{01} \cos(\omega_1 t - k_1 z) + \xi_0 \cdot E_{02} \cos(\omega_2 t - k_2 z) + \\
 &\eta \cdot \left(\frac{E_{01}^2 + E_{02}^2}{2} \right) + \eta \cdot \frac{E_{01}^2}{2} \cos(2\omega_1 t - 2k_1 z) + \eta \cdot \frac{E_{02}^2}{2} \cos(2\omega_2 t - 2k_2 z) + \\
 &+ \eta \cdot E_{01} E_{02} \cos[(\omega_1 - \omega_2)t - (k_1 - k_2)z] + \\
 &+ \eta \cdot E_{01} E_{02} \cos[(\omega_1 + \omega_2)t - (k_1 + k_2)z]
 \end{aligned} \tag{8.6}$$

Отсюда видно, что индуцированная поляризация состоит из линейной и нелинейной частей, последняя из которых включает постоянную составляющую и компоненты с частотами $2\omega_1$, $2\omega_2$, $(\omega_1 - \omega_2)$, $(\omega_1 + \omega_2)$.

Необходимо отметить, что случай комбинации световой волны и постоянного или низкочастотного переменного электрического поля в материале с квадратичной оптической нелинейностью соответствует линейному электрооптическому эффекту. Действительно, при $\omega_2=0$ из (8.6) следует, что в индуцированной нелинейной поляризации присутствует компонента с частотой ω_1 , причем ее величина пропорциональна произведению $E_{01} \cdot E_{02}$. Таким образом, обусловленное наличием постоянного электрического поля изменение диэлектрической проницаемости (и, соответственно, показателя преломления) среды пропорционально напряженности данного постоянного поля.

Б. Среда с кубической нелинейностью.

Выражение для наведенной поляризации в среде с кубической нелинейностью имеет вид: $P(E) = \xi_0 \cdot E + \theta \cdot E^3$. Подставляя сюда выражение для поля световой волны (8.4), получим:

$$\begin{aligned}
 P(E) &= \xi_0 \cdot E_0 \cos(\omega t - kz) + \theta \cdot E_0^3 \cos^3(\omega t - kz) = \\
 &= \xi_0 \cdot E_0 \cos(\omega t - kz) + \theta \cdot E_0^3 \cdot [\cos(\omega t - kz) \cdot \cos^2(\omega t - kz)] = \\
 &= \left(\xi_0 \cdot E_0 + \frac{3}{4} \theta \cdot E_0^3 \right) \cdot \cos(\omega t - kz) + \frac{\theta \cdot E_0^3}{4} \cdot \cos(3\omega t - 3kz)
 \end{aligned} \quad (8.7).$$

Таким образом, при распространении монохроматической световой волны в среде с кубической нелинейностью, возникают составляющие поляризации на той же самой частоте и на ее третьей гармонике (рис. 8.2).

Нелинейная добавка к поляризации среды на частоте ω с амплитудой $\frac{3}{4} \theta \cdot E_0^3$ означает изменение диэлектрической проницаемости (или показателя преломления) среды, что является причиной эффектов самовоздействия световых полей. Составляющая поляризации на утроенной частоте

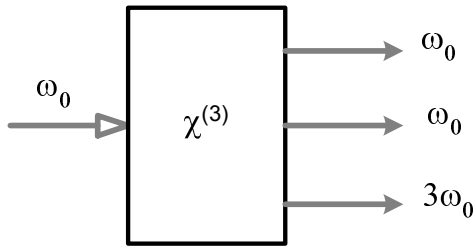


Рис. 8.2.

обуславливает возможность генерации третьей гармоники основной волны. Часто среды с кубической оптической нелинейностью называют керровскими средами. Такое название связано, в частности, с тем, что квадратичный электрооптический эффект (эффект Керра) наблюдается в centrosymmetric средах, т.е. в средах с кубической нелинейностью. Как

проделано с квадратично – нелинейной средой, можно показать, что при воздействии на среду с кубической нелинейностью гармонического светового и постоянного электрического полей, при произвольной малой амплитуде светового поля, в среде появляется наведенная поляризация на частоте световой волны, величина которой пропорциональна квадрату напряженности постоянного поля. Таким образом, квадратичный электрооптический эффект можно также отнести к семейству нелинейно – оптических эффектов.

Следует отметить, что в среде с квадратичной нелинейностью при двухчастотном воздействии возможно наблюдать нелинейно – оптические эффекты, характерные для кубично – нелинейной среды. Так, если $\omega_2 = 2\omega_1$, то среди составляющих нелинейной поляризации мы находим слагаемое:

$$\eta \cdot E_{01} E_{02} \cos[(\omega_1 - \omega_2)t - (k_1 - k_2)z] = \eta \cdot E_{01} E_{02} \cos[(\omega_1)t - (k_1 - k_2)z],$$

т.е. в этом случае возникает нелинейная добавка к показателю преломления материала на частоте ω_1 , обусловленная распространением в среде световой волны с частотой ω_2 . Соответственно, при таких условиях в среде с квадратичной нелинейностью могут наблюдаться эффекты пространственного самовоздействия световых пучков, характерные для сред

с кубической нелинейностью. Вдобавок, при тех же условиях составляющая нелинейной поляризации на суммарной частоте имеет частоту $\omega_1 + \omega_2 = 3\omega_1$. Таким образом, в рассматриваемой ситуации квадратично – нелинейная среда ведет себя как среда и с квадратичной, и с кубической нелинейностями.

8.2. Нелинейное волновое уравнение

При анализе нелинейно – оптических взаимодействий световых волн используют нелинейное волновое уравнение. Получим его из стандартных уравнений Максвелла для диэлектрической среды:

$$\operatorname{rot}\bar{H} = \frac{\partial\bar{D}}{\partial t}; \quad \operatorname{rot}\bar{E} = -\frac{\partial\bar{B}}{\partial t}.$$

В материальном уравнении для связи векторов \bar{D} и \bar{E} представим вектор поляризации среды в виде суммы двух слагаемых \bar{P}_l и \bar{P}_{nl} , учитывающих линейный и нелинейный вклады:

$\bar{D} = \epsilon_0\bar{E} + \bar{P}_l + \bar{P}_{nl}$. Тогда, поскольку $\epsilon_0\bar{E} + \bar{P}_l = \epsilon\bar{E}$, получим:

$$\bar{D} = \epsilon\bar{E} + \bar{P}_{nl} \quad (8.8).$$

Теперь обычным путем придем к волновому уравнению:

$$\begin{aligned} \operatorname{rot}\operatorname{rot}\bar{E} &= -\mu\frac{\partial}{\partial t}\operatorname{rot}\bar{H} = -\mu\frac{\partial^2\bar{D}}{\partial t^2} = -\mu\frac{\partial^2(\epsilon\bar{E})}{\partial t^2} - \mu\frac{\partial^2\bar{P}_{nl}}{\partial t^2} \Rightarrow \\ \Rightarrow \operatorname{grad}\operatorname{div}\bar{E} - \nabla^2\bar{E} + \mu\epsilon\frac{\partial^2\bar{E}}{\partial t^2} &= -\mu\frac{\partial^2\bar{P}_{nl}}{\partial t^2} \Rightarrow \\ \Rightarrow \nabla^2\bar{E} - \mu\epsilon\frac{\partial^2\bar{E}}{\partial t^2} &= \mu\frac{\partial^2\bar{P}_{nl}}{\partial t^2} \end{aligned} \quad (8.9).$$

Правая часть нелинейного волнового уравнения зависит от нелинейной поляризации среды, которая приводит к изменению спектрального состава светового поля в среде, вследствие оптической нелинейности.

8.3. Генерация второй гармоники

Генерация второй гармоники или удвоение частоты света представляет собой один из наиболее важных с практической точки зрения нелинейно – оптических эффектов. Впервые генерация второй гармоники излучения рубинового лазера экспериментально наблюдалась Франкеном в 1961 году. Лазерный луч с длиной волны $\lambda = 694,3$ нм фокусировался на поверхность пластины из кристаллического кварца. Спектральный состав прошедшего излучения анализировался спектрометром. Было обнаружено, что прошедшее излучение содержит компоненту с удвоенной частотой (т.е. с $\lambda=347,15$ нм). Эффективность преобразования в излучение с удвоенной частотой в первых экспериментах составляла лишь около 10^{-8} . Использование более

эффективных материалов, увеличение мощности лазера, отработанные методики по обеспечению условий фазового синхронизма позволили в последующие годы существенно улучшить эффективность преобразования, доведя ее почти до единицы.

Рассмотрим в качестве примера случай падения двух плоских световых волн с частотами ω_1 и ω_2 на плоскую границу непроводящей немагнитной среды. Будем считать волны распространяющимися вдоль оси Z . Входная граница совпадает с плоскостью $z=0$. Поля волн представим в виде:

$$E_i^{\omega_1}(t) = \text{Re}[E_i^{\omega_1} \cdot \exp(i\omega_1 t)] = \frac{1}{2}[E_i^{\omega_1} \cdot \exp(i\omega_1 t) + \text{к.с.}] \quad (8.10),$$

$$E_k^{\omega_2}(t) = \text{Re}[E_k^{\omega_2} \cdot \exp(i\omega_2 t)] = \frac{1}{2}[E_k^{\omega_2} \cdot \exp(i\omega_2 t) + \text{к.с.}] \quad (8.11).$$

Первое из них поляризовано вдоль оси \bar{i} , а второе – вдоль оси \bar{k} . Представим компоненты вектора поляризации среды на частоте $\omega_3 = \omega_1 + \omega_2$ вдоль \bar{i} – го направления в виде:

$P_i^{\omega_3 = \omega_1 + \omega_2}(t) = \text{Re}[P_i^{\omega_3} \cdot \exp(i\omega_3 t)]$ и найдем выражение для поляризации среды с учетом ее квадратичной нелинейности. Используем в данном случае стандартную форму записи соотношения для вектора поляризации, принятую в большинстве работ по нелинейной оптике []:

$$P_i = \epsilon_0 \chi_{ij} E_j + 2d_{ijk} E_j E_k + 4\chi_{ijkl} E_j E_k E_l + \dots \quad (8.12),$$

где d_{ijk} и χ_{ijkl} – компоненты тензоров квадратичной и кубичной нелинейных восприимчивостей среды. Вклад квадратичной нелинейности определяется членом:

$$P_i = 2d_{ijk} E_j E_k \quad (8.13).$$

Тогда для поляризации среды имеем:

$$P_i(t) = 2d_{ijk} \left[\frac{1}{2} E_j^{\omega_1} \cdot \exp(i\omega_1 t) + \frac{1}{2} E_j^{\omega_2} \cdot \exp(i\omega_2 t) + \text{с.с.} \right] \times \\ \times \left[\frac{1}{2} E_k^{\omega_1} \cdot \exp(i\omega_1 t) + \frac{1}{2} E_k^{\omega_2} \cdot \exp(i\omega_2 t) + \text{с.с.} \right] \quad (8.14).$$

Рассмотрим из этого соотношения только член, соответствующий суммарной частоте.

$$P_i^{\omega_1 + \omega_2}(t) = \frac{1}{2} \left[d_{ijk} E_j^{\omega_1} E_k^{\omega_2} \exp[i(\omega_1 + \omega_2)t] + d_{ikj} E_k^{\omega_2} E_j^{\omega_1} \exp[i(\omega_2 + \omega_1)t] + \text{с.с.} \right] \quad (8.15).$$

Здесь использовано правило суммирования по повторяющимся индексам. В среде без потерь, обладающей мгновенным откликом, $d_{ijk} = d_{ikj}$, тогда получим:

$$P_i^{\omega_1+\omega_2}(t) = \frac{1}{2} P_i^{\omega_3=\omega_1+\omega_2} \exp[i(\omega_1 + \omega_2)t] + c.c. =$$

$$= [d_{ijk} E_j^{\omega_1} E_k^{\omega_2} \exp[i(\omega_1 + \omega_2)t] + c.c.]$$

или:

$$P_i^{\omega_3=\omega_1+\omega_2} = 2d_{ijk} E_j^{\omega_1} E_k^{\omega_2}$$

Отметим, что коэффициенты d_{ijk} , как правило, определяются из результатов экспериментов. В большинстве случаев это эксперименты по генерации второй гармоники, когда $\omega_1 = \omega_2 = \omega$. В этом случае:

$$P_i^{2\omega} = d_{ijk} E_j^{\omega} E_k^{\omega}.$$

Рассмотрим теперь простейший одномерный случай, считая что световое поле распространяется вдоль оси Z , а $\frac{\partial}{\partial x} = \frac{\partial}{\partial y} = 0$. Полагаем, что имеется три бегущих вдоль Z плоских световых волны с частотами $\omega_1, \omega_2, \omega_3$ с полями:

$$E_i^{(\omega_1)}(z, t) = \frac{1}{2} [E_{1i}(z) \cdot \exp[i(\omega_1 t - k_1 z)] + c.c.]$$

$$E_k^{(\omega_2)}(z, t) = \frac{1}{2} [E_{2k}(z) \cdot \exp[i(\omega_2 t - k_2 z)] + c.c.]$$

$$E_j^{(\omega_3)}(z, t) = \frac{1}{2} [E_{3j}(z) \cdot \exp[i(\omega_3 t - k_3 z)] + c.c.] \quad (8.16),$$

где индексы i, j, k означают декартовы координаты, и каждый из индексов может принимать значения x или y .

Подставим соотношения (8.13) в нелинейное волновое уравнение (8.9). Очевидно, что нужно провести дифференцирование полей по координате z , поскольку они однородны вдоль поперечных координат ($\frac{\partial}{\partial x} = \frac{\partial}{\partial y} = 0$).

Тогда, взяв i -ю компоненту, получим:

$$\nabla^2 E_i^{(\omega_1)}(z, t) = \frac{\partial^2}{\partial z^2} E_i^{(\omega_1)}(z, t) = \frac{1}{2} \frac{\partial^2}{\partial z^2} \{ E_{1i}(z) \cdot \exp[i(\omega_1 t - k_1 z)] + c.c. \}.$$

Проведя дифференцирование, учтем, что амплитуды взаимодействующих волн слабо изменяются на расстояниях порядка длины волны. Соответственно, можно пренебречь второй производной поля по z :

$$\frac{\partial^2 E_{1i}}{\partial z^2} \ll k_1 \frac{\partial E_{1i}}{\partial z}. \text{ В итоге получим:}$$

$$\nabla^2 E_i^{(\omega_1)}(z, t) = -\frac{1}{2} [k_1^2 E_{1i}(z) + 2ik_1 \frac{\partial E_{1i}(z)}{\partial z}] \cdot \exp[i(\omega_1 t - k_1 z)] + c.c. \quad (8.17).$$

Аналогичные соотношения получим для других составляющих светового поля. Учтем, что нелинейное волновое уравнение должно удовлетворяться

для каждой частотной компоненты, запишем соответствующее соотношение для поля $E_i^{(\omega_i)}(z, t)$:

$$\begin{aligned} & \left[\frac{k_1^2}{2} E_{1i} + ik_1 \frac{\partial E_{1i}}{\partial z} \right] \cdot \exp[i(\omega_1 t - k_1 z)] + c.c. = \\ & = \left[\frac{1}{2} \omega_1^2 \mu_0 \varepsilon E_{1i} \cdot \exp[i(\omega_1 t - k_1 z)] + c.c. - \mu_0 \frac{\partial^2}{\partial t^2} [P_{NL}(z, t)]_i \right] \end{aligned} \quad (8.18).$$

Подставим в это уравнение выражение для нелинейной поляризации, учтем, что $\omega_1^2 \mu_0 \varepsilon = k_1^2$, и, проведя несложные преобразования, получим три уравнения:

$$\frac{\partial E_{1i}}{\partial z} = -i\omega_1 \sqrt{\frac{\mu_0}{\varepsilon_1}} d_{ijk} E_{3j} E_{2k}^* \cdot \exp[-i(k_3 - k_2 - k_1)z] \quad (8.19),$$

$$\frac{\partial E_{2k}^*}{\partial z} = i\omega_2 \sqrt{\frac{\mu_0}{\varepsilon_2}} d_{kij} E_{1i} E_{3j}^* \cdot \exp[-i(k_1 - k_3 + k_2)z] \quad (8.20)$$

$$\frac{\partial E_{3j}}{\partial z} = -i\omega_3 \sqrt{\frac{\mu_0}{\varepsilon_3}} d_{jik} E_{1i} E_{2k} \cdot \exp[-i(k_1 + k_2 - k_3)z] \quad (8.21).$$

Полученные уравнения могут использоваться для анализа различных взаимодействий в среде с квадратичной оптической нелинейностью. Остановимся лишь на случае генерации второй гармоники. При этом $\omega_1 = \omega_2$, а $\omega_3 = 2\omega_1$. Соответственно, нам необходимы лишь два из полученных уравнений – первое или второе и последнее. При анализе считаем взаимодействие слабым, чему соответствует условие $dE_{1i}/dz = 0$, т.е. пренебрегаем изменением амплитуды волны основной частоты в процессе генерации второй гармоники. Поэтому мы можем рассматривать лишь последнее из этих трех уравнений - уравнение (8.21):

$$\frac{\partial E_{3j}}{\partial z} = -i\omega \sqrt{\frac{\mu_0}{\varepsilon}} d_{jik} E_{1i} E_{1k} \cdot \exp[i \cdot \Delta k \cdot z] \quad (8.22),$$

где $\omega = \omega_1 = \omega_3/2$; $\varepsilon = \varepsilon_3$; $\Delta \bar{k} = \bar{k}_3^j - \bar{k}_1^i - \bar{k}_1^k$. Здесь k_1^i - волновое число для волны с частотой ω_1 при поляризации света вдоль направления i .

Положим, что i, j, k могут принимать значения x и y вдоль направления поляризации собственных мод в кристаллической среде. Световая волна с частотой ω_1 представляет собой, в общем случае, суперпозицию собственных мод, поэтому может иметь и x , и y составляющие. Величина волнового числа k зависит от состояния поляризации, что отмечено верхними индексами.

Легко получить решение уравнения (8.22) для случая отсутствия второй гармоники на входе нелинейной среды, при $E_{3j}(0) = 0$. Для кристалла длиной L это решение имеет вид:

$$E_{3j}(L) = -i\omega \sqrt{\frac{\mu_0}{\epsilon}} d_{jik} E_{li} E_{lk} \cdot \frac{\exp[i \cdot \Delta k \cdot L] - 1}{i\Delta k} \quad (8.23).$$

Интенсивность поля второй гармоники на выходе среды найдем в виде:

$$E_{3j}^*(L) \cdot E_{3j}(L) = -\omega^2 \frac{4\mu_0}{\epsilon} (d_{jik})^2 E_{li}^2 E_{lk}^2 L^2 \cdot \frac{\sin^2[\frac{1}{2}\Delta k \cdot L]}{[\frac{1}{2}\Delta k \cdot L]^2} \quad (8.24).$$

Выражение для выходной мощности найдем, используя соотношение:

$$\frac{P^{(2\omega)}}{S} = \frac{1}{2} \sqrt{\frac{\epsilon}{\mu_0}} E_{3j} E_{3j}^* \quad (8.25),$$

где S – площадь поперечного сечения пучка. Подставляя в это выражение (8.25), получим:

$$\frac{P^{(2\omega)}}{S} = 2 \sqrt{\frac{\mu_0}{\epsilon}} \cdot \omega^2 (d_{jik})^2 E_{li}^2 E_{lk}^2 L^2 \cdot \frac{\sin^2[\frac{1}{2}\Delta k \cdot L]}{[\frac{1}{2}\Delta k \cdot L]^2} \quad (8.26).$$

Запишем это выражение через мощность первой гармоники:

$$\frac{P^{(2\omega)}}{S} = 8 \left(\frac{\mu_0}{\epsilon} \right)^{3/2} \cdot \frac{\omega^2 (d_{jik})^2 L^2}{n^3} \cdot \frac{P_i^{(\omega)} P_k^{(\omega)}}{S^2} \cdot \frac{\sin^2[\frac{1}{2}\Delta k \cdot L]}{[\frac{1}{2}\Delta k \cdot L]^2} \quad (8.27).$$

Или, при $P_i^{(\omega)} = P_k^{(\omega)} = P^{(\omega)} / 2$ оно примет вид:

$$\frac{P^{(2\omega)}}{P^{(\omega)}} = 2 \left(\frac{\mu_0}{\epsilon} \right)^{3/2} \cdot \frac{\omega^2 (d_{jik})^2 L^2}{n^3} \cdot \frac{P^{(\omega)}}{S} \cdot \frac{\sin^2[\frac{1}{2}\Delta k \cdot L]}{[\frac{1}{2}\Delta k \cdot L]^2} \quad (8.28).$$

8.4. Условия фазового синхронизма при генерации второй гармоники

Из (8.28) вытекает, что эффективная генерация второй гармоники требует выполнения условия фазового синхронизма $\Delta k=0$. Поскольку $\omega_3 = 2\omega$, а $\omega_1 = \omega_2 = \omega$, то требование $\Delta k=0$ эквивалентно условию $k^{(2\omega)} = 2k^{(\omega)}$. Действительно, причиной возникновения излучения с удвоенной частотой является нелинейная поляризация, наводимая в среде световой волной с частотой ω . Волна наведенной поляризации распространяется в среде со скоростью распространения индуцирующего излучения. Если $k^{(2\omega)} \neq 2k^{(\omega)}$, то между излучением с удвоенной частотой и наводимой поляризацией возникает фазовый сдвиг, нарастающий с

расстоянием, так что в некоторой точке он может достигать величины π . Расстояние, на котором фазовый сдвиг изменяется от нуля до π , называется "когерентной длиной":

$$l_c = \frac{2\pi}{\Delta k} = \frac{2\pi}{k^{(2\omega)} - 2k^{(\omega)}} \quad (8.29).$$

Ее величина служит оценкой максимальной длины кристалла, которую можно эффективно использовать для генерации второй гармоники. В области нормальной дисперсии показатель преломления растет с увеличением частоты, поэтому:

$$\Delta k = k^{(2\omega)} - 2k^{(\omega)} = \frac{2\omega}{c}(n^{(2\omega)} - n^{(\omega)}) \quad (8.30).$$

Здесь использовано соотношение $k = \omega n/c$. Тогда когерентная длина может быть выражена в форме:

$$l_c = \frac{\pi c}{\omega \cdot (n^{(2\omega)} - n^{(\omega)})} = \frac{\lambda}{2(n^{(2\omega)} - n^{(\omega)})} \quad (8.31),$$

где λ – длина волны падающего света.

Для обеспечения условий фазового синхронизма используется оптическая анизотропия кристаллов. Из (8.30) видим, что для этого необходимо выполнить условие $n^{(2\omega)} = n^{(\omega)}$, т.е. условие равенства показателей преломления среды на основной частоте и на частоте второй гармоники. В области нормальной дисперсии показатели преломления как обыкновенной, так и необыкновенной волн, распространяющихся в заданном направлении, увеличиваются с ростом частоты. Таким образом, выполнить условие равенства показателей преломления невозможно, если волны с частотами ω и 2ω относятся к одному типу (обыкновенные или необыкновенные). Однако условие фазового синхронизма может быть выполнено при использовании волн разных типов.

В качестве примера рассмотрим зависимости величин показателей преломления необыкновенной и обыкновенной волн в одноосном, отрицательном кристалле (пример такого кристалла – кристалл ниобата

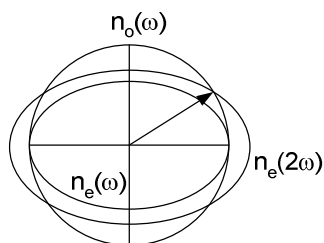


Рис. 8.3.

лития), от угла θ между направлением распространения и оптической осью (осью Z). Эти зависимости схематично изображены на рис. 8.3. Можно видеть, что для обыкновенной волны основной частоты и необыкновенной волны с удвоенной частотой имеются направления распространения света, удовлетворяющие условию фазового синхронизма (одно из них отмечено стрелкой).

Величина необыкновенного показателя преломления одноосного кристалла в зависимости от угла в плоскости, содержащей оптическую ось, определяется соотношением:

$$\frac{1}{n_e^2(\theta)} = \frac{\cos^2 \theta}{n_o^2} + \frac{\sin^2 \theta}{n_{e0}^2} \quad (8.32),$$

где θ – угол между направлением распространения света и оптической осью; n_o и n_{e0} – главные значения показателей преломления кристалла. Таким образом, направления синхронизма могут быть найдены из условия:

$$\frac{1}{n_o^2(\omega)} = \frac{\cos^2 \theta_s}{n_o^2(2\omega)} + \frac{\sin^2 \theta_s}{n_{e0}^2(2\omega)} \quad (8.33),$$

откуда получим:

$$\sin^2 \theta_s = \frac{n_o^{-2}(\omega) - n_o^{-2}(2\omega)}{n_e^{-2}(2\omega) - n_o^{-2}(2\omega)} \quad (8.34).$$

8.5. Распространение светового пучка в нелинейно – оптической среде

При анализе нелинейной поляризации, наводимой световым полем в среде с кубичной нелинейностью, мы выяснили, что возникает составляющая поляризации на частоте падающей световой волны, величина которой пропорциональна квадрату амплитуды светового поля. Это значит, что световое поле с неоднородным распределением интенсивности наводит в такой среде оптические неоднородности, которые, в свою очередь, могут изменять пространственную структуру светового поля. В этом случае говорят о пространственном самовоздействии светового поля или пучка. Рассмотрим основные особенности такого самовоздействия и некоторые его результаты.

Пусть световой пучок падает на границу нелинейно – оптической среды. Полагаем его поле в виде:

$$E(x, z, t) = A(x, z) \cdot \exp(ikz - i\omega t) \quad (8.35),$$

где x и z – поперечная и продольная координаты. В данном случае предполагается, что размер пучка в направлении x значительно превышает длину волны света, а в направлении z его поле однородно. Соответственно, можно использовать параксиальное приближение, т.е. считать, что амплитуда $A(x, z)$ изменяется в направлении z намного медленнее, чем в направлении x . Уравнение, описывающее поведение пучка в нелинейной диэлектрической среде, получим, исходя из стандартной системы уравнений Максвелла:

$$\begin{aligned} \operatorname{rot} \bar{H} &= \frac{\partial \bar{D}}{\partial t} \\ \operatorname{rot} \bar{E} &= -\frac{\partial \bar{B}}{\partial t} \end{aligned}$$

Для этого используем обычную процедуру, применяя к обеим частям второго из этих уравнений операцию rot :

$$rot(rot\bar{E}) = -\frac{\partial}{\partial t} rot\bar{B} = -\mu \frac{\partial}{\partial t} rot\bar{H} = -\mu \frac{\partial^2 \bar{D}}{\partial t^2},$$

или $grad(div\bar{E}) - \nabla^2 \bar{E} = -\mu \frac{\partial^2 \bar{D}}{\partial t^2}$. Поскольку для диэлектрической среды $grad(div\bar{E}) = 0$, то получаем:

$$\nabla^2 \bar{E} = \mu \frac{\partial^2 \bar{D}}{\partial t^2} = -\mu \epsilon \omega^2 \bar{E} \quad (8.36).$$

С учетом (8.35) выражения для пространственных производных поля \bar{E} имеют вид:

$$\frac{\partial^2 E}{\partial x^2} = \frac{\partial^2 A}{\partial x^2} \cdot \exp(ikz - i\omega t) \quad (8.37a),$$

$$\frac{\partial^2 E}{\partial z^2} = \frac{\partial^2 A}{\partial z^2} \exp(ikz - i\omega t) + 2ik \frac{\partial A}{\partial z} \exp(ikz - i\omega t) - k^2 A \cdot \exp(ikz - i\omega t) \quad (8.37b).$$

В параксиальном приближении можно положить $\frac{\partial^2 A}{\partial z^2} \ll \frac{\partial A}{\partial z}$ и пренебречь

второй производной $\frac{\partial^2 A}{\partial z^2}$, тогда (8.36) примет вид:

$$\frac{\partial^2 A}{\partial x^2} + 2ik \frac{\partial A}{\partial z} - k^2 A = -\mu \epsilon \omega^2 A \quad (8.38).$$

Здесь мы ограничились скалярным приближением, считая световой пучок линейно поляризованным. Поскольку рассматриваемая среда является оптически нелинейной, представим выражение для ее диэлектрической проницаемости ϵ в форме:

$$\epsilon = \epsilon_l + \epsilon_{nl}(|E|^2) \quad (8.39),$$

где ϵ_l и ϵ_{nl} – линейная и нелинейная составляющие диэлектрической проницаемости. Но $\mu \epsilon_l \omega^2 = k^2$, поэтому: $\frac{\partial^2 A}{\partial x^2} + 2ik \frac{\partial A}{\partial z} = -\mu \epsilon_{nl} \omega^2 A$ или:

$$\frac{\partial A}{\partial z} - \frac{i}{2k} \cdot \frac{\partial^2 A}{\partial x^2} = \frac{i\mu \epsilon_{nl} \omega^2}{2k} A \quad (8.40).$$

Правую часть этого уравнения запишем в форме:

$$\frac{i\mu \epsilon_{nl} \omega^2 \epsilon_l}{2k \epsilon_l} = \frac{ik^2 \epsilon_{nl}}{2k \epsilon_l} = \frac{ik}{2} \cdot \frac{\epsilon_{nl}}{\epsilon_l} \quad (8.41).$$

Но ϵ_{nl} можно представить как $\epsilon_{nl} = \epsilon - \epsilon_l = \Delta \epsilon$. Т.к. $\epsilon = n^2$, то: $\Delta \epsilon = (n^2)' = 2n \cdot \Delta n$ и

$\frac{\epsilon_{nl}}{\epsilon_l} = \frac{2n \cdot \Delta n_{nl}}{n^2} = \frac{2\Delta n_{nl}}{n}$. Таким образом, (8.40) примет вид:

$$\frac{\partial A}{\partial z} - \frac{i}{2k} \cdot \frac{\partial^2 A}{\partial x^2} = \frac{ik\Delta n_{nl}}{n} A \quad (8.42).$$

Это уравнение для амплитуды светового пучка, распространяющегося в нелинейно – оптической среде. В правой части данного уравнения величина нелинейной добавки к показателю преломления среды Δn_{nl} является функцией пространственных координат и зависит от локальной интенсивности светового поля. Рассмотрим различные механизмы оптической нелинейности, приводящие к изменениям показателя преломления среды на частоте распространяющегося в ней светового поля.

1. Керровская оптическая нелинейность, обусловленная нелинейной связью наведенной поляризации в среде и напряженности электрического поля в световой волне. Это изменение показателя преломления на частоте падающей волны в среде с кубической нелинейностью (см. 8.7). Величина изменения показателя преломления определяется соотношением:

$$\Delta n_{nl}^{(\kappa)} = n_{(2)} \cdot I(x, z) \quad (8.43),$$

где $n_{(2)}$ – нелинейный показатель преломления среды. Для обычных материалов (газы, жидкости, стекло, кристаллы) величина $n_{(2)}$ очень мала, так что заметные изменения показателя преломления, приводящие к эффектам пространственного самовоздействия световых пучков, наблюдаются при интенсивностях света в сотни МВт/см^2 и более. Данный механизм характеризуется очень высокой скоростью отклика (она ограничивается лишь инерционностью электронов).

Отметим, что знак нелинейного показателя преломления $n_{(2)}$ определяет знак нелинейности среды. В случае $n_{(2)} > 0$ среда называется самофокусирующей, при $n_{(2)} < 0$ ее называют самодефокусирующей.

2. Термооптическая нелинейность – изменение показателя преломления при изменении температуры среды, обусловленном поглощением света. Величина изменения показателя преломления в этом случае определяется соотношением:

$$\Delta n_{nl}^{(TO)} = \frac{\partial n}{\partial T} \cdot \delta T[I(x, z)] \quad (8.44),$$

где $\frac{\partial n}{\partial T}$ – температурный коэффициент показателя преломления, а $\delta T[I(x, z)]$ – локальное изменение температуры среды вследствие поглощения света. Термооптическая нелинейность может быть значительно более сильной, чем керровская, однако она и значительно более медленная.

3. Оптическая нелинейность, обусловленная фоторефрактивным эффектом. Как отмечалось в п. 7.3, такая нелинейность обусловлена изменением показателя преломления вследствие линейного электрооптического эффекта при условии индуцирования светом в фоторефрактивном кристалле электрического поля пространственного

заряда. Она может быть поистине гигантской, приводя к значительным возмущениям показателя преломления даже при микроваттных интенсивностях света. Однако она также значительно медленнее в сравнении с керровской и во многих случаях медленнее термооптической. Изменения показателя преломления для разных механизмов фоторефракции определяются выражениями:

Для фотовольтаического механизма:

$$\Delta n_{nl}^{(phv)} = A \cdot I_d^{-1} \cdot \frac{I(x, z)/I_d}{1 + I(x, z)/I_d} \quad (8.45),$$

где A – константа, зависящая от физических характеристик и электрооптических коэффициентов материала; I_d – так называемая темновая интенсивность, определяемая условием $\sigma_d = \sigma_{ph} I$, т.е. интенсивность света, при которой величина фотопроводимости материала равна его темновой проводимости.

Для дрейфового механизма:

$$\Delta n_{nl}^{(drift)} = B \cdot \frac{I_{bg}}{I_{bg} + I(x, z)} \cdot E_{ext} \quad (8.46),$$

где I_{bg} – интенсивность фоновой подсветки; E_{ext} – напряженность внешнего электрического поля, приложенного к кристаллу; B – постоянная, зависящая от физических характеристик и электрооптических коэффициентов материала.

Для диффузионного механизма:

$$\Delta n_{nl}^{dif} = -\frac{1}{2} n_l^3 r \cdot E_D(z) = -\frac{1}{2} n_l^3 r \cdot \frac{k_B T}{e} \frac{1}{n(x)} \cdot \frac{dn(x)}{dx} \quad (8.47).$$

Отметим, что в случае диффузионного механизма транспорта электрических зарядов изменение показателя преломления носит нелокальный характер, т.е. оно определяется величиной градиента концентрации фотовозбужденных носителей заряда. Соответственно, эффекты самовоздействия световых пучков проявляются в данном случае не в самофокусировке или самодефокусировке пучков, а в их самоискривлении.

Уравнение (8.42) позволяет проследить, как изменяется профиль светового пучка в нелинейной среде при различных механизмах и величине оптической нелинейности, а также различных параметрах светового пучка. Его решение может быть получено в аналитической форме лишь в ряде частных случаев. В большинстве же ситуаций возможно только его численное решение.

Особый интерес представляют решения уравнения, отвечающие режимам сохранения поперечного профиля светового пучка, называемым режимами пространственных оптических солитонов. В таком случае решение (8.42) ищем в виде:

$$A(x, z) = U(x) \cdot \exp(i\gamma z) \quad (8.48).$$

Тогда:

$$\frac{\partial A(x, z)}{\partial z} = i\gamma \cdot A(x, z) = i\gamma \cdot U(x) \cdot \exp(i\gamma z),$$

$$\frac{\partial^2 A(x, z)}{\partial x^2} = \frac{\partial^2 U}{\partial x^2} \cdot \exp(i\gamma z).$$

Подставляя эти соотношения в (8.42), получим:

$$i\gamma U \cdot \exp(i\gamma z) - \frac{i}{2k} \cdot \left[\frac{\partial^2 U}{\partial x^2} \cdot \exp(i\gamma z) \right] = \frac{ik\Delta n_{nl}}{n} \cdot U(x) \cdot \exp(i\gamma z),$$

что после упрощения дает:

$$U'' - 2k\gamma U + \frac{2k^2\Delta n_{nl}}{n} \cdot U = 0 \quad \text{или}$$

$$U'' + \left(\frac{2k^2\Delta n_{nl}}{n} - 2k\gamma \right) \cdot U = 0 \quad (8.49),$$

где $U'' = \frac{\partial^2 U}{\partial x^2}$.

Это уравнение для пространственных солитонов, которое называют также нелинейным уравнением Шредингера. Действительно, в данной форме записи аналогия с соответствующим уравнением очевидна. В то же время это уравнение совпадает по форме с волновым уравнением для градиентного диэлектрического волновода. Соответственно, его решения представляют собой собственные моды оптического волновода, индуцированного в нелинейной среде самим световым пучком. Следует отметить, что нелинейная добавка к показателю преломления среды Δn_{nl} может быть как положительной, так и отрицательной. Соответственно, нелинейная добавка γ к постоянной распространения положительна либо отрицательна. В первом случае ($\Delta n_{nl} > 0$) среда является самофокусирующей, а решение уравнения (8.49) представляет собой **светлый солитон**, т.е. световой пучок, распространяющийся без дифракционного расплывания в нелинейно – оптической среде. Противоположный случай соответствует самодефокусирующей среде. Собственное решение (8.49) при этом называют темным солитоном, под которым понимается как бы бездифракционное поведение неосвещенной области в световом поле.

Рассмотрим отдельно решения нелинейного уравнения Шредингера (или солитонного уравнения) для различных механизмов оптической нелинейности среды.

8.5.1. Пространственные оптические солитоны в среде с

керровской нелинейностью

В данном случае нелинейное возмущение показателя преломления среды определяется нелинейной восприимчивостью третьего порядка и, как уже отмечалось, может быть записано в форме:

$$\Delta n_{nl} = n_{(2)} \cdot I(x, z) = n_{(2)} \cdot U^2.$$

Солитонное уравнение (8.49) примет вид:

$$U'' + \left(\frac{2k^2 n_{(2)}}{n} \cdot U^2 - 2k\gamma \right) \cdot U = 0 \quad (8.50).$$

Далее, обозначая $\frac{kn_2}{n} = \alpha$ получим окончательно:

$$U - \frac{1}{2k\gamma} \cdot U'' = \frac{\alpha}{\gamma} \cdot U^3 \quad (8.51).$$

Коэффициент α может быть как положительным, так и отрицательным. В первом случае, для самофокусирующей среды, решение для светлого солитона имеет вид:

$$U(\xi) = \left(\frac{2\gamma}{\alpha} \right)^{1/2} \operatorname{sech}(\xi) \quad (8.52),$$

где $\xi = (2k\gamma)^{1/2} x$ – нормированная поперечная координата.

В случае самодефокусирующей среды решение дает профиль темного солитона в виде:

$$U(\xi) = \left(\frac{\gamma}{\alpha} \right)^{1/2} \tanh(\xi / \sqrt{2}) \quad (8.53),$$

где $\xi = (-2k\gamma)^{1/2} x$.

Таким образом, в случае среды с керровской нелинейностью профили световых пучков, соответствующих светлым и темным пространственным солитонам, выражаются в аналитической форме и определяются соотношениями (8.52) и (8.53). Рисунок 5.4 иллюстрирует поперечные профили интенсивности таких солитонов.

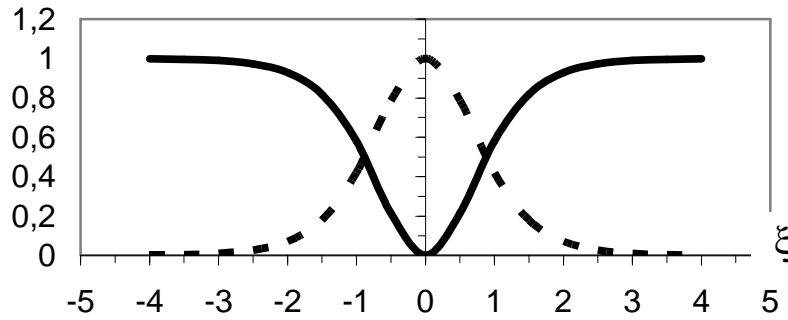


Рис. 5.4. Профили интенсивности светлого (пунктир) и темного (сплошная) пространственных солитонов.

8.5.2. Пространственные оптические солитоны в среде с фоторефрактивной нелинейностью

В случае фотовольтаических солитонов, подставляя в (8.49) выражение (8.45) для Δn_{nl} , получим:

$$U - \frac{1}{2k\gamma} \cdot \frac{\partial^2 U}{\partial x^2} = \frac{k}{n\gamma} \cdot A \cdot \frac{U^2}{1+U^2} \cdot U \quad \text{или:}$$

$$U - \frac{1}{2k\gamma} \cdot U'' = \frac{\beta}{\gamma} \cdot \frac{U^3}{1+U^2} \quad (8.54),$$

где $\beta = \frac{kA}{n}$.

И, наконец, для фоторефрактивных солитонов при дрейфовом механизме фоторефракции можно получить подобное уравнение в виде:

$$U - \frac{1}{2k\gamma} \cdot \frac{\partial^2 U}{\partial x^2} = \frac{k}{n\gamma} \cdot D \cdot \frac{I_{bg}}{I_{bg} + I(x,z)} \cdot U \cdot E_{ext} \quad \text{или}$$

$$U - \frac{1}{2k\gamma} \cdot U'' = \frac{F}{\gamma} \cdot \frac{U}{1+U^2} \quad (8.55).$$

Таким образом, получено два похожих уравнения, соответствующих режимам пространственных оптических солитонов в средах с фотовольтаическим и дрейфовым механизмами фоторефракции. Можно видеть, что в случае фотовольтаических солитонов, при малой интенсивности световых пучков, (8.54) приводится к виду солитонного уравнения для сред с керровской нелинейностью. Отметим, что во всех случаях правая часть уравнений может быть как положительной, так и отрицательной. Это соответствует случаям самофокусирующей и самодиффузирующей нелинейностей. Для фотовольтаического эффекта знак нелинейности определяется направлением фотовольтаического тока. Для дрейфового

механизма фоторефракции это соответствует прямой или обратной полярности приложенного к кристаллу внешнего электрического поля. В результате решения данных уравнений дают режимы светлых и темных пространственных солитонов.

8.6. Другие нелинейно – оптические эффекты

Кроме рассмотренных выше, существует еще целый ряд нелинейно–оптических эффектов, к которым относятся:

- а) вынужденное комбинационное рассеяние света и вынужденное рассеяние Мандельштама–Бриллюэна;
- б) параметрическая генерация света;
- в) оптическое выпрямление;
- г) исчезновение красной границы фотоэффекта;
- д) оптическое просветление и оптическое затемнение среды.

Остановимся кратко на сути некоторых из них.

А. Вынужденное рассеяние света

Эффект комбинационного рассеяния света, называемого также рамановским рассеянием, представляет собой процесс неупругого рассеяния света на оптических колебаниях молекул или внутримолекулярных колебаниях. При таком рассеянии частота рассеянного света отличается от частоты падающей волны.

При комбинационном рассеянии спектр прошедшего через вещество излучения содержит комбинационные частоты $\omega_0 \pm m\Omega$, где ω_0 - частота падающей волны; Ω - частота молекулярных колебаний, m - целое.

Составляющие, имеющие частоты, меньше чем частота падающей волны, называют стоксовыми компонентами. Составляющие с более высокой частотой – антистоксовы компоненты. Результат рассеяния зависит от исходного состояния молекулы. Если молекула находится в основном (невозбужденном) состоянии, то при поглощении кванта с энергией $h\omega_0$ она переходит в возбужденное состояние с энергией $h\omega_0 + h\Omega$. Время жизни в таком состоянии должно быть очень маленьким, если эта энергия не соответствует одному из разрешенных состояний. При переходе на разрешенный уровень молекула может оказаться как в основном состоянии, так и в возбужденном. В первом случае изменение частоты излучения не происходит, во втором частота рассеянного излучения оказывается ниже частоты падающего. Это стоксова линия.

Если молекула находится в возбужденном состоянии, то она может перейти в невозбужденное состояние в результате рассеяния. Энергия рассеянного кванта в этом случае больше, чем энергия падающего. Это антистоксова линия. Интенсивность антистоксовых линий всегда ниже, чем

стоксовых. Это обусловлено большей населенностью основного состояния в условиях термодинамического равновесия.

При вынужденном комбинационном рассеянии (ВКР) световая волна возбуждает молекулярные колебания, на которых потом происходит рассеяние этой волны. При ВКР резко возрастает интенсивность антистоксовой компоненты. ВКР, в отличие от спонтанного КР, возникает при интенсивностях света, превышающих некоторую пороговую величину. Эффективность преобразования падающей волны в рассеянные при ВКР очень сильно возрастает и теоретически может достичь 100%.

О рассеянии Мандельштама – Бриллюэна говорят, имея в виду дифракцию света на акустических колебаниях решетки или среды (в жидкостях). Этот эффект с квантовой точки зрения интерпретируется как эффект трехчастичного взаимодействия. В результате таких взаимодействий фотон может как исчезать, так и рождаться. При большой интенсивности света возможно усиление акустических колебаний, что, в свою очередь, ведет к увеличению интенсивности рассеяния. Этот эффект и есть вынужденное рассеяние Мандельштама - Бриллюэна. С позиций классической физики процесс генерации акустических волн может рассматриваться как результат электрострикции, т.е. возникновения упругих деформаций в электрическом поле световой волны. Возрастание мощности акустических колебаний наблюдается, если приток энергии от поля световой волны компенсирует потери на затухание звука.

Пороговая плотность мощности для вынужденного рассеяния Мандельштама – Бриллюэна составляет $10^7 - 10^9$ Вт/см². Частотные сдвиги при ВРМБ значительно меньше, чем при ВКР.

Б. Исчезновение красной границы фотоэффекта.

Суть данного эффекта заключается в исчезновении порогового характера появления фототока в экспериментах по фотоэффекту в случае использования очень мощного светового излучения. Это является результатом двух- или многофотонного поглощения, когда атом практически одновременно поглощает по крайней мере два фотона. Вероятность таких актов поглощения зависит от интенсивности света и может быть заметной в сильных световых полях.

Список обозначений

- n – показатель преломления среды
 c – скорость света
 ϵ - абсолютная диэлектрическая проницаемость среды
 ϵ_0 - универсальная диэлектрическая постоянная или диэлектрическая проницаемость вакуума
 μ - абсолютная магнитная проницаемость среды
 μ_0 - магнитная проницаемость вакуума
 σ - удельная проводимость среды
 $h=1,05 \cdot 10^{-34}$ Дж·с - постоянная Планка
 ω - круговая частота
 E – энергия фотона
 \vec{p} - импульс частицы
 \vec{E} - вектор напряженности электрического поля
 \vec{D} - электрическая индукция
 \vec{H} - напряженность магнитного поля
 \vec{B} - магнитная индукция
 q – величина электрического заряда
 I – электрический ток
 \vec{j} - плотность электрического тока
 Φ - магнитный поток
 ρ - объемная плотность электрического заряда
 \vec{P} - вектор поляризации среды
 χ - диэлектрическая восприимчивость среды
 ϵ_{ij} - компоненты тензора диэлектрической проницаемости
 \vec{P} - вектор Пойнтинга
 w – объемная плотность энергии электромагнитного поля
 T – период колебаний
 $\hat{\epsilon}$ - комплексная диэлектрическая проницаемость
 k – волновое число
 k_0 – волновое число света в вакууме
 λ - длина волны
 \vec{r} - радиус – вектор точки
 W_0 - волновое сопротивление вакуума
 $\vec{x}_0, \vec{y}_0, \vec{z}_0$ - орты декартовой системы координат
 $F(k_x, k_y, z)$ - угловой спектр плоских волн в плоскости z
 w_0 – полуширина перетяжки гауссова пучка
 m – масса электрона
 e – заряд электрона

β - постоянная распространения

U – потенциальная энергия

$\hat{\mathcal{E}}$ - оператор

A_{21} - коэффициент Эйнштейна по спонтанным переходам

B_{12}, B_{21} - коэффициенты Эйнштейна по вынужденным переходам

N_i - населенность уровня

T – абсолютная температура

w_{12}, w_{21} - вероятности безызлучательных переходов частиц в двухуровневой квантовой системе

κ_a - коэффициент усиления активной среды

Δn - изменение показателя преломления

B_{ij} - компоненты тензора относительной диэлектрической непроницаемости

r_{ijk} - электрооптические коэффициенты

p_{ijkl} - компоненты тензора фотоупругости

S_{kl} - компоненты тензора упругих деформаций

E_{sc} - напряженность поля пространственного электрического заряда

D - коэффициент диффузии электронов в зоне проводимости

E_D - величина диффузионного поля

G - постоянная Гласа

σ_d - темновая проводимость диэлектрика

σ_{ph} - коэффициент фотопроводимости

N_D - концентрация донорных центров

N_D^+ - концентрация ионизированных донорных центров

s - сечение фотоионизации

γ - коэффициент рекомбинации

n_e – концентрация электронов в зоне проводимости

μ - подвижность электронов

β_{ph} - фотовольтаический коэффициент

СПИСОК ЛИТЕРАТУРЫ

1. Никольский В.В. Электродинамика и распространение радиоволн / В.В. Никольский, Т.И. Никольская.- М.: Наука, 1989.- ISBN 5-02-014033-3.
2. Борн М. Э. Основы оптики / М. Борн, Э. Вольф. - М.: Наука, 1973. - 720 с.
3. Ярив А. Оптические волны в кристаллах / А.Ярив, П.Юх. - М.: Мир, 1987. - 616 с.
4. Гончаренко А.М.. Гауссовы пучки света / А.М. Гончаренко. – Минск: Наука и техника, 1977. – 144 с.
5. Виноградова М.Б. / М.Б. Виноградова М.Б., О.В. Руденко, А.П. Сухоруков. Теория волн. - М.: Наука. Гл. ред. физ. - мат. лит, 1979. - 384 с.
6. Акаев А. Оптические методы обработки информации / А.Акаев, А.Майоров.- М.: Высшая школа. - 1988.
7. Информационная оптика / Под ред. Н.Н.Евтихиева. – М.: Изд-во МЭИ, 2000. – 612 с. - ISBN 5-7046-0584-2.
8. Ландау Л.Д. Электродинамика сплошных сред / Л.Д. Ландау, Е.М. Лифшиц. - М.: Наука, 1982.- 624 с.
9. Интегральная оптика / Под ред. Т. Тамира. – М.:Мир, 1978. – 344 с.
10. Страховский Г.М. Основы квантовой электроники / Г.М. Страховский, А.В. Успенский. - М.: Высшая школа, 1973.- 312 с.
11. Гольдштейн Л.Д. Электромагнитные поля и волны / Л.Д.Гольдштейн, Н.В.Зернов. М.: Сов. Радио, 1971.
12. Вихман Э. Квантовая физика / Э.Вихман. - М.: Наука: Гл. ред. физ. - мат. лит, 1977. – 416 с.
13. Тарасов Л.В. Введение в квантовую оптику / Л.В. Тарасов – М.: Высшая школа, 1987.
14. Ярив А. Квантовая электроника / А. Ярив. - М.: Сов. Радио, 1980.
15. Петров М.П. Фоторефрактивные кристаллы в когерентной оптике / М.П.Петров, С.И.Степанов, А.В.Хоменко. - М.: Наука, 1992. – 320 с. ISBN 5-02-024627-1.
16. Цернике Ф. Прикладная нелинейная оптика / Ф. Цернике, Дж. Мидвинтер. - М.: Мир, 1976.