Федеральное агентство по образованию Российской Федерации

МОСКОВСКИЙ ИНЖЕНЕРНО-ФИЗИЧЕСКИЙ ИНСТИТУТ (ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ)

В. К. Егоров, И. Г. Зубарев, Р. С. Стариков

Лабораторный практикум ФИЗИЧЕСКИЕ ОСНОВЫ МОДУЛЯЦИИ ЛАЗЕРНОГО ИЗЛУЧЕНИЯ

Рекомендовано УМО «Ядерные физика и технологии» в качестве учебного пособия для студентов высших учебных заведений

Москва 2008

УДК 535.14(075) ББК 22.344я7 Е30

Егоров В. К., Зубарев И. Г., Стариков Р. С. Лабораторный практикум «Физические основы модуляции лазерного излучения». – М.: МИФИ, 2008. – 252 с.

Настоящее учебное пособие предназначено для студентов кафедры «Лазерная физика» МИФИ, проходящих обучение по специальности «Физика твердого тела» со специализацией в области физики лазеров при выполнении ими кафедрального лабораторного практикума, а также при изучении курсов «Физическая оптика», «Квантовая электроника» и других курсов кафедры. Учебное пособие может быть полезно также для студентов других физических специальностей при изучении ряда оптических дисциплин.

Настоящее учебное пособие подготовлено в рамках Инновационной образовательной программы.

Рецензент д. ф.-м. н., проф. А. П. Менушенков

ISBN 978-5-7262-1044-5

© Московский инженерно-физический институт (государственный университет), 2008

Содержание

ТЕОРЕТИЧЕСКАЯ ЧАСТЬ	6
Введение	6
1. Предварительные сведения	8
1.1. Сигналы в оптике	8
1.2. Преобразование Фурье в оптике	10
1.3. Линейные системы	17
2. Основные типы модуляции колебаний и волн	21
2.1. Амплитудная модуляция	23
2.2. Частотная модуляция	40
2.3. Фазовая модуляция	41
2.4. Управление поляризацией излучения	54
2.5. Изменение направления волнового вектора	55
3. Технические средства модуляции лазерного излучения	58
3.1. Фазовая и частотная модуляция	58
3.2 Амплитудная модуляция.	62
3.3. Управление поляризацией излучения	67
3.4. Изменение направления волнового вектора	70
3.5. Заключение	75
4. Электрооптический амплитудный модулятор	76
4.1. Принципиальные схемы	76
4.2. Электрооптическая фазовая модуляция света	82
4.3. Рабочие характеристики амплитудных	
электрооптических модуляторов	84
4.4. Компенсация статических и линамических	
фазовых расстроек в ЭОМ	87
4.5. Электрооптическая молуляция	
немонохроматического излучения	
4.6. Амплитудно-частотные характеристики ЭОМ	89

5. Акустооптическая модуляция света	91
5.1. Фотоупругость	91
5.2. Когерентное рассеяние света на акустических волнах.	92
5.3. Устройство акустооптического модулятора	101
5.4. Применения акустооптических модуляторов света	103
6. Пространственно-временной модулятор типа ФП-ЖК	105
6.1. Жидкие кристаллы в модуляторах света	105
6.2. Электрооптические эффекты в нематических ЖК	109
6.3. Электрооптические модуляторы света	
на основе твист-текстур ЖК	113
6.4. Модулятор света со структурой фотопроводник-ЖК	115
6.5. Заключительные замечания	117
7. Обращение волнового фронта	
в фоторефрактивных кристаллах	118
7.1. Введение	. 118
7.2. Обращение волнового фронта.	
Сопряженная (обращенная) волна	. 121
7.3. Восстановление голограммы	
сопряженной опорной волной	. 125
7.4. Динамическая голография	128
7.5. Четырехволновое взаимодействие	
в среде с нелинейной восприимчивостью	129
7.6. Формирование фазовой решетки	
в среде с кубической нелинейностью	132
7.7. Обращение волнового фронта	
при вырожденном четырехволновом взаимодействии	134
7.8. Диффузионная нелинейность фоторефрактивных	
кристаллов. Диффузионная фоторефракция	136
7.9. Формирование фазовой решетки	
в фоторефрактивном кристалле	. 142
7.10. OBФ в реальных пучках ограниченного сечения	148
7.11. Запись и считывание решеток	
в анизотропных средах	153
7.12. Динамические особенности фазовой решетки	
в фоторефрактивном кристалле	157
7.13. Заключительные замечания	162

ПРАКТИЧЕСКАЯ ЧАСТЬ	164
Лабораторная работа 1	
Изучение электрооптической модуляции	165
Лабораторная работа 2	
Изучение акустооптической модуляции	168
Лабораторная работа 3	
Пространственный модулятор типа ФП-ЖК	172
Лабораторная работа 4	
Обращение волнового фронта в фоторефрактивном	
кристалле	176
ПРИЛОЖЕНИЯ	184
1 Электрооптический эффект Поккельса	184
2. Модуляторы для ВОЛС	199
3. Фазовая модуляция света в фотоупругой среде	214
4. Современные типы модуляции оптических волн	229
Литература	251

ТЕОРЕТИЧЕСКАЯ ЧАСТЬ

ВВЕДЕНИЕ

Современная цивилизация неотделима от информационных технологий – радио, телевидения, компьютерных сетей, электронных баз данных и архивов. По мере увеличения объемов передаваемой информации растет разрыв между запросами на информацию и скоростью доступа к ней. Электрорадиосвязь по наземным проводным, кабельным и радиорелейным линиям, радиосвязь по космическим каналам уже не могут отвечать в полной мере современным требованиям. Возможности расширения их пропускной способности либо почти исчерпаны, либо сопряжены с непомерными дополнительными материальными издержками. Единственная практически приемлемая возможность расширения передаваемых потоков информации и увеличения скорости ее обработки – использование оптического диапазона. Электромагнитные волны оптического диапазона по многим причинам очень подходят для передачи и обработки с их помощью больших потоков информации. Оптические устройства передачи, хранения и обработки информации привлекательны малой в технологическом пределе энергоемкостью работы и теоретически чрезвычайно высоким быстродействием, объясняемым естественным пространственным параллелизмом оптических систем и возможной широкой полосой модуляции оптического сигнала. Например, пропускная способность оптической линии связи на один-два порядка шире, чем у линии радиодиапазона, и один оптический кабель содержит десятки и сотни волокон с исключительно высокой их защищенностью от внешних и взаимных помех, с крайне малыми потерями энергии сигналов.

Носителем информации в оптических системах служит свет. В основе оптических методов обработки информации лежит возможность преобразований амплитуды или интенсивности световой волны в оптической системе. Реальные волны, вне зависимости от их происхождения, представляют собой волновые пакеты (или группы волн), энергия которых так или иначе сосредоточена в ограниченной области и во времени, и в пространстве. Именно группы волн, и только они, могут переносить некоторую информацию и энергию из одной точки пространства в другую. Информацию несет некоторая «метка» на волне (возмущение, искажение идеализированной волны), которую называют сигналом. Простейшим сигналом может быть группа волн определенной конечной длительности, и тогда она будет нести информацию в один бит.

Для формирования сигнала в оптическом диапазоне пригодны изменения практически любого из физических параметров волны: амплитуды, частоты, фазы, направления волнового вектора и поляризации. Управление параметрами излучения собирательно называют *модуляцией*, а соответствующие устройства – *модуляторами*. Эта терминология перенесена из классической радиотехники и радиофизики, из теории и техники формирования, передачи, приема и обработки электрических сигналов.

Более точно модуляцией называют *медленное* (в сравнении с исходными колебаниями) изменение управляемого параметра, так что, по крайней мере, на протяжении нескольких периодов колебаний управляемый параметр можно считать практически неизменным. Вообще требование медленности не вытекает из каких-либо физических предпосылок, и чаще всего его накладывают для удобства математического отображения сигналов. При всем этом, в оптике и в диапазоне СВЧ скорость изменений управляемого параметра естественно ограничена техническими возможностями аппаратуры. Кроме того, скорость изменения параметра задает и ширину спектра модулированного сигнала, а на ширину спектра часто бывают наложены либо технические ограничения (например, ширина полосы пропускания линии связи), либо ограничения нормативные (запрещено излучение в соседние частотные каналы связи).

Физические эффекты, пригодные для реализации оптических модуляторов, изучены задолго до появления лазеров. Но только после появления лазеров началось широкое применение приборов на основе этих эффектов. Появились волоконно-оптические линии связи (ВОЛС), лазерные дальномеры и локаторы, оптические системы обработки радиосигналов, изображений и цифровой информации, открытые (воздушные и космические) лазерные линии связи, оптические диски CD, DVD, BD, и т. п. Поэтому исследования и разработки методов и приборов для высокоэффективного и экономичного управления лазерным, да и вообще световым излучением, поиск и отработка технологии новых материалов для них продолжаются.

1. ПРЕДВАРИТЕЛЬНЫЕ СВЕДЕНИЯ

1.1. Сигналы в оптике

В общем случае, сигнал есть знак, необходимый для представления информации при ее передаче, хранении или преобразовании. В теории сигналов и ее приложениях оперируют с сигналами, представимыми в виде функций s(u). В том случае, если сигнал известен в каждой точке u, его называют детерминированным, в противном случае – случайным.

Сигналы, для которых выполнено условие

$$\int_{\pm\infty} s^2(u) du < \infty \,,$$

относят к важнейшему для практики классу сигналов с ограниченной энергией. Сигналы конечной длительности называют финитными. Практически важными типами сигналов являются периодические сигналы, для которых верно:

$$s(u+nT) = s(u),$$

и, как их частный случай, гармонические сигналы:

$$s(u) = A\cos(2\pi f_u u + \varphi);$$

здесь T – период, $f_u = T^{-1}$ – частота, φ – фаза сигнала. Если множество определения и множество значений сигнала – непрерывные интервалы, сигнал называется аналоговым. Если сигнал определяется последовательностью значений в некоторых выбранных точках – отсчетах, то он называется дискретным. Если множество значений сигнала есть фиксированный набор точек, то сигнал называется квантованным. Дискретный квантованный сигнал называется цифровым. Цифровой сигнал, использующий только два уровня квантования – «0» и «1» – называют *бинарным* или *двоичным*. Примеры аналогового, дискретного и цифрового сигналов показаны на рис 1.1.

В радиосвязи и оптике носителем сигнала служит электромагнитная волна. Используя простейшую математическую модель, в этом случае можно говорить о сигналах вида $s(\mathbf{r}, t)$, зависящих, строго говоря, от координат \mathbf{r} и времени t. На практике для оптических и оптоэлектронных систем имеет смысл рассматривать сигналы временные и пространственные. В первом случае сигнал s(t)отождествляют с зависящими от времени амплитудой U(t) или интенсивностью I(t). Такие сигналы находят применение, главным

образом, в системах оптической связи; в настоящее время наибольшее распространение и развитие получили волоконнооптические линии связи. Основное достоинство оптического сигнала в данном случае – широкая полоса его модуляции. Информационные характеристики системы определяют характерные времена срабатывания или, что эквивалентно, полосы пропускания устройств ввода и вывода сигналов. Оптические системы, предназначенные для формирования изображения того или иного рода, оперируют с двумерными объектами в трехмерном пространстве. Фактически сигналом s(x, v) в этом случае считают амплитуду U(x, v)или интенсивность I(x, y) волны, трактуемые как «плоское» изображение.



Рис. 1.1. Аналоговый (*a*), дискретный (б) и цифровой (*в*) сигналы

В отличие от проводников и волоконно-оптических линий связи, передающих сигнал последовательно во времени, в оптической системе сигнал распространяется параллельно в пространстве. В этом случае информационные характеристики оптической системы определяются полосами пространственных частот устройств ввода информации в оптическую систему и дифракционными ограничениями разрешения самой оптической системы. Для формирования временных и пространственных сигналов используют, соответственно, временные и пространственно-временные модуляторы света (ПВМС).

1.2. Преобразование Фурье в оптике

Основными функциями систем обработки информации являются прием и передача сигналов, их хранение (запоминание) и анализ. Для анализа сигналов в современной технике используются интегральные преобразования, такие как преобразования Лапласа, Фурье, Меллина, Радона, Уолша, Адамара, Хью, Хартли, Z-преобразование, вейвлет–преобразование и многие другие. Наиболее широко используют преобразование Фурье (или ряд Фурье для периодических функций).

Преобразование Фурье определяют в одномерном (временном) случае как:

$$FT\{s(t)\} = S(f) = \int_{\pm\infty} s(t) \exp(-j2\pi ft) dt ,$$

а в двумерном (пространственном) случае как:

$$FT\{s(x, y)\} = S(f_x, f_y) = \iint_{\pm\infty} s(x, y) \exp(-j2\pi(f_x x + f_y y)) \, dxdy \, .$$

Некоторые практически важные примеры сигналов и их спектров представлены на рис. Для представления двоичных цифровых сигналов используются прямоугольные импульсы и их последовательности. На рис. 1.2 представлены функции

$$\operatorname{rect}(u) = \begin{cases} 1, |u| \le 1/2\\ 0, |u| > 1/2, \\ \operatorname{sinc}(u) = \frac{\sin \pi u}{\pi u}. \end{cases}$$

Фурье–образ прямоугольной функции rect(t) – функция sinc(f), так что $FT{rect(t)}=sinc(f)$.





Рис. 1.2. Сигналы rect(u) и sinc(u)

Рис. 1.3. Последовательность прямоугольных импульсов s(u) и ее спектр $S(\omega)$

Бесконечная последовательность прямоугольных импульсов и ее спектр Фурье показаны на рис. 1.3. «Прямоугольный» rect(x/b)rect(y/b) и «круглый» сігс(2r/b) пространственные сигналы, а также их пространственные спектры представлены на рис. 1.4. Эти два типа сигналов весьма важны в оптике, поскольку типичные апертуры оптических систем – прямоугольная или круглая, а элементарные ячейки ПВМС, используемых для формирования отсчетов изображений – пиксели – имеют прямоугольную форму.

На практике ПВМС устроены так, что изображение формируют элементарные ячейки – *пиксели*, пропускание или отражение которых соответствует значениям отсчетов изображения. Пиксели ПВМС, как правило, имеют прямоугольную или близкую к ней форму и располагаются на некотором расстоянии друг от друга, которое определяется возможностями технологии изготовления модулятора или особыми требованиями его использования. Расстояние между пикселями определяет так называемый фактор заполнения, или пространственную скважность ПВМС. Сетки с разной пространственной скважностью отверстий и соответствующие пространственные спектры представлены на рис. 1.5. Изображение рис. 1.5, реализованное с разной пространственной скважностью и соответствующие пространственные спектры представлены на рис. 1.6.



Рис. 1.4. Сигналы rect(x/b)rect(y/b) и circ(2r/b) и их спектры

Преобразование Фурье естественно для оптики. Представление немонохроматических волн суперпозицией гармонических компонент иллюстрирует эту естественность для случая разложения по временным частотам и может наблюдаться как явление дисперсии. Для случая разложения по пространственным частотам говорят о представлении некоторой монохроматической волны произвольной геометрии суперпозицией плоских волн, распространяющихся под разными углами и составляющими ее пространственный спектр. Это наблюдается при явлениях дифракции, особенно очевидно в дальней зоне. В частности, широко известно свойство линзы формировать в фокальной плоскости дифракционную картину Фраунгофера $U_f(x_f, y_f)$, с точностью до постоянного фазового множителя совпадающую с преобразованием Фурье двумерного распределения когерентного светового поля на входе:

$$U_f(x_f, y_f) = \frac{A \exp\left(i\frac{k}{2f}\left(1 - \frac{d_0}{f}\right)(x_f^2 + y_f^2)\right)}{j\lambda f} \times \\ \times \iint_{\pm\infty} t_0(x_0, y_0) \exp\left(-i\frac{2\pi}{\lambda f}(x_0x_f + y_0y_f)\right) dx_0 dy_0.$$

Здесь A – амплитуда плоской монохроматической волны, падающей на транспарант с распределением коэффициента пропускания $t_0(x_0,y_0)$, d_0 – расстояние от транспаранта до фурье-объектива, f – фокусное расстояние фурье-объектива, λ – длина волны излучения.

Распределение интенсивности в дифракционной картине Фраунгофера на *амплитудной* синусоидальной решетке с функцией пропускания:

$$t(x,y) = [1/2 + m\cos(2\pi f_0 x)/2)]\operatorname{rect}(x/l)\operatorname{rect}(y/l),$$

где m – параметр, определяющий различие максимума и минимума пропускания решетки, l – размер квадратной апертуры решетки, f_0 – частота с которой изменяется пропускание, определяется как

 $I(x,y) = (l^2/2\lambda z)^2 + \operatorname{sinc}^2(ly/2\lambda z) \{\operatorname{sinc}^2(lx/2\lambda z) +$

 $+ m^2 \operatorname{sinc}^2 [l(x+f_0 \lambda z)/2 \lambda z]/4 + m^2 \operatorname{sinc}^2 [l(x-f_0 \lambda z)/2 \lambda z]/4\};$ график его распределения вдоль оси *x* представлен на рис. 1.7.

Распределение интенсивности в дифракционной картине Фраунгофера на синусоидальной *фазовой* решетке, функция пропускания которой имеет вид:

 $t(x, y) = \exp[im\cos(2\pi f_0 x)/2)]\operatorname{rect}(x/l)\operatorname{rect}(y/l)$ определяется как:

$$I(x,y) = (l^2/\lambda z)^2 \sum_{q=\pm\infty} J_q^2 (m/2) \{ \operatorname{sinc}^2 [l(x - qf_0 \lambda z)/2 \lambda z] \} \operatorname{sinc}^2 (ly/2 \lambda z),$$

где J_q – функция Бесселя первого рода порядка q; распределение интенсивности вдоль оси x при m = 8 показано на рис. 1.8.

Методы анализа пространственного спектра изображений получили широкое распространение при решении задач распознавания изображений и многих других. Пример изображения, его рассчитанного пространственного спектра Фурье и его же пространственного спектра Фурье, сформированного в когерентном свете в фокальной плоскости линзы, изображены на рис. 1.9. Операции получения спектра Фурье проходят параллельно за один такт обработки, что позволяет получить чрезвычайно высокое быстродействие. Забегая вперед, отметим, что при модуляции спектр модулирующего сигнала смещается по шкале частот на частоту несущей, и это дает спектр модулированного сигнала.



Рис. 1.5. Сетки с разной пространственной скважностью (a) и соответствующие им пространственные спектры (δ)



Рис. 1.6. Изображение рис. 1.5, представленное с разной пространственной скважностью (слева), и соответствующие пространственные спектры (справа)



Рис.1.7. Распределение интенсивности в дифракционной картине Фраунгофера на амплитудной синусоидальной решетке



Рис.1.8. Распределение интенсивности в дифракционной картине Фраунгофера на синусоидальной фазовой решетке (*m*=8)



Рис. 1.9. Изображение (*a*), рассчитанный пространственный спектр (*б*), измеренный пространственный спектр (*в*) (представлены с точностью до масштаба)

1.3. Линейные системы

Свойство линейности дает возможность применить к оптическим системам удобный математический аппарат теории линейных систем, позволяющий однозначно определять свойства оптических систем, вводя описание их работы как линейное преобразование входного сигнала в выходной. Под системой вообще понимается любое устройство, производящее какое-либо действие над сигналом. Действие линейной системы в общем случае описывают интегралом суперпозиции:

$$S_{\rm GDIX}(u_{\rm GDIX}) = \iint_{\pm\infty} h(u_{\rm GDIX};u_{\rm GX}) S_{\rm GIX}(u_{\rm GIX}) du_{\rm GIX}$$

На практике реальные системы, строго говоря, не являются линейными, но можно говорить о той или иной степени их близости к линейным в некоторой области. В частности, как известно, при выполнении некоторых условий, в дифракционном приближении преобразование входного светового поля, осуществляемого оптической системой, характеризуется интегралом суперпозиции вида:

$$U_{\rm gblx}(x_{\rm gblx}, y_{\rm gblx}) = \iint_{\pm\infty} h(x_{\rm gblx}, y_{\rm gblx}; x_{\rm gx}, y_{\rm gx}) U_{\rm gx}(x_{\rm gx}, y_{\rm gx}) dx_{\rm gx} dy_{\rm gx}$$

в случае когерентной оптической системы, а в случае некогерентной оптической системы

$$I_{\text{\tiny GDLX}}(x_{\text{\tiny GDLX}}, y_{\text{\tiny GDLX}}) = \kappa \iint_{\pm\infty} \left| h(x_{\text{\tiny GDLX}}, y_{\text{\tiny GDLX}}; x_{\text{\tiny GX}}, y_{\text{\tiny GX}}) \right|^2 I_{\text{\tiny GX}}(x_{\text{\tiny GX}}, y_{\text{\tiny GX}}) dx_{\text{\tiny GX}} dy_{\text{\tiny GX}}$$

где U_{ax} , $_{bbx}$ - комплексные амплитуды, а $I_{ax,bbx}$ - интенсивности поля на входе и выходе соответственно, h - комплексный импульсный отклик оптической системы, κ - вещественная константа. Иначе говоря, когерентная оптическая система осуществляет линейное преобразование амплитуды, а некогерентная - линейное преобразование интенсивности. Когерентная передаточная функция:

$$H(f_x, f_y) = FT\{h\}$$

описывает действие когерентной оптической системы в пространственно-частотном представлении и совпадает с точностью до системы координат с функцией зрачка оптической системы:

$$H(f_x, f_y) = P(-\lambda d_i f_x, -\lambda d_i f_y),$$

здесь *d_i* – расстояние до изображения. Аналогичные соображения верны и для пространственно-некогерентных оптических систем,

хотя математическое описание в этом случае несколько сложнее. Таким образом, зная частотные характеристики оптической системы можно удобным образом судить о ее действии. Учитывая сказанное, следует сделать вывод как о важности учета частотных характеристик сигналов обрабатываемых оптической системой, так и о необходимости учета пространственно-частотных характеристик самой оптической системы, формируемых, в свою очередь, пространственно-частотными характеристиками оптических элементов и устройств, ее составляющих.



Рис. 1.10. Изображение, измеренное с одним и тем же разрешением (1024х1024 отсчетов) и с разным динамическим диапазоном (числом градаций серого g)

1.4. Динамический диапазон и количество информации

Во всех реальных оптико-электронных системах присутствуют шумы и помехи разного рода, искажающие, а иногда и маскирующие полезные сигналы. Точность, с которой аналоговый сигнал отображает переносимую им информацию, определяют действующие шумы; ее характеризует динамический диапазон. Под динамическим диапазоном понимают, в общем случае, отношение максимального уровня сигнала к эффективному уровню шумов. Динамический диапазон можно измерять в уровнях мощности сигнала, децибелах (логарифмическая шкала с основанием 10) или в битах (дискретная логарифмическая шкала с основанием 2).



Рис. 1.11. Изображение, измеренное с разным разрешением N × N и числом градаций серого 256, и соответствующие пространственные спектры

Количество информации, которое может содержать сигнал, определяет произведение его динамического диапазона на количество разрешимых отсчетов. Например, дискретный сигнал из 8 отсчетов, с динамическим диапазоном 256 уровней (8 бит), содержит $8 \times 8 = 64$ бита информации. Изменение информативности сигнала в зависимости от используемого динамического диапазона иллюстрирует рис. 1.10.

С другой стороны, информационная емкость сигнала может быть охарактеризована произведением его длительности на ширину полосы частот. Для двумерного пространственного случая это иллюстрирует рис. 1.11, на котором представлены одно и тоже изображение, измеренное с одинаковым числом градаций серого в отсчете, но разным пространственным разрешением, и соответствующие каждому случаю пространственные спектры.

В заключение отметим, что круг вопросов о влиянии шумов относится к области статистической оптики и статистической радиофизики; эти вопросы выходят за пределы тематики данного обзора.

2. ОСНОВНЫЕ ТИПЫ МОДУЛЯЦИИ КОЛЕБАНИЙ И ВОЛН

В радиофизике различают несколько типов модуляции электромагнитных колебаний и волн, в зависимости от управляемых параметров. Если математически представить модулированный сигнал в форме *комплексного* сигнала, в общей форме,

$$A(t) = a(t) \times \exp(i\omega_0 t),$$

и полагать a(t) комплексной функцией, то можно ввести понятия элементарных, или примитивных видов модуляции, рассматривая частные виды функции a(t).

Если a(t) вещественная, то модуляцию называют амплитудной; если $a(t) = a_0 \times \exp(jD(t) \times t)$, то модуляцию называют частотной; если $a(t) = a_0 \times \exp(j\beta(t))$, то модуляцию называют фазовой. Во всех этих выражениях величины a, D, β – вещественные. Кроме того, их полагают медленными функциями времени, то есть

 $\max(|da/dt|) \ll \max(|a|), \qquad |D(t)| \le \omega_0, \qquad |d\beta/dt| \ll \omega_0.$

Таким образом, в выбранном представлении интенсивность колебаний,

$$W = A \times A^* = a_0^2,$$

не зависит от времени при фазовой модуляции (ФМ) и при частотной модуляции (ЧМ), а при амплитудной модуляции

$$W(t) = a(t)^2$$

поскольку a(t) – переменная вещественная и медленно меняющаяся. По общим правилам, W определяют как величину, усредненную по нескольким периодам колебаний частоты ω_0 , так что W(t) – переменная положительная и такая же медленная, как и a, D, β .

В данном разделе речь будет идти о модуляции и формировании сигналов из *монохроматических* колебаний и волн, где достаточно точно определена *несущая частота* колебаний ω_0 и имеет смысл представление о *фазах* колебательных процессов. Иногда мы будем рассматривать модуляцию *интенсивности* широкополосного (шумоподобного стационарного) излучения, каждый раз особо оговаривая эти обстоятельства.

Интуитивно можно представлять себе модулированное колебание как аналогию звука с переменной громкостью и/или высотой тона. Известно, что звук с переменной громкостью («разливом») можно получить и как результат биений при сложении колебаний с близкими частотами; такие биения мы будем отличать от модуляции. Схематично эти различия показаны на рис. 2.1.

Отметим, что существуют методы модуляции, дающие в точности такой же эффект, как и простое сложение колебаний (например, сравните рис. 2.1, δ и 2.1, ϵ). Поэтому по виду сложных колебаний иногда трудно установить, образованы ли они модуляцией или сложением более простых колебаний. Чтобы отвечать на такие вопросы, нужно измерять сложные функции корреляции интересующего нас колебания, тогда как простой анализ спектра не всегда дает определенный ответ.



Рис. 2.1. Представления сложения и модуляции напряженности на комплексной плоскости (*a*), временные представления: сложения напряженностей двух гармонических колебаний с разными амплитудами (*б*); амплитудной модуляции (*в*), фазовой модуляции совместная амплитудно-фазовая модуляция (*д*)

Выбор типа модуляции, то есть самого управляемого параметра и способа управления, зависит от поставленной цели и удобства технической реализации как модуляции, так и восстановления управляющего сигнала, по существу и заключающего в себе информацию. Например, для оптической связи пригодна модуляция интенсивности световой волны. Сигналы в таком случае представляют собой группы волн длительностью порядка наносекунды и менее.

Далее рассмотрим типы колебаний и их параметры подробнее.

2.1. Амплитудная модуляция

Для монохроматических волн амплитудную модуляцию (AM) можно определить и как изменение *амплитуды* (напряженности поля) волны, а также и как изменение *интенсивности* (квадрата модуля напряженности поля) волны. Эти два определения не равносильны, так как они описывают изменения разных физических величин.

Для широкополосного излучения напряженность поля нельзя корректно определить как периодическую величину, но определена интенсивность – средний квадрат напряженности поля, или плотность потока энергии. Время усреднения здесь существенно больше условного периода колебаний, соответствующего средней частоте в спектре излучения. Соответственно, АМ широкополосного излучения определяют как изменение интенсивности.

Всякое реальное излучение в определенной количественной мере отличается от идеализированного монохроматического. Противопоставление монохроматического и широкополосного излучения на практике требует уточнений; к этому вопросу мы вернемся в конце раздела, где подробнее рассмотрена модуляция интенсивности немонохроматического излучения.

Начнем с АМ напряженности поля монохроматического излучения, где все термины и представления унаследованы от радиотехники.

АМ гармонических колебаний простейшего вида, которую обычно называют *тональной* АМ,

 $a(t) = A \times [1 + K \cos(2\pi f t + \phi)] \cos(\omega t + \psi)$ описывают следующими параметрами:

- 1) амплитуда А,
- 2) несущая частота ω,
- 3) модулирующая частота f,
- 4) коэффициент модуляции К.

Два первых параметра описывают исходные немодулированные колебания (часто их коротко называют «несущей»), а вторая пара – собственно АМ (закон модуляции). Начальные фазы ф и ψ далее

будем опускать как несущественные. Графическое отображение АМ-колебания и его основных параметров приведено на рис. 2.2.



Рис. 2.2. Тональная амплитудная модуляция: a -колебания несущей частоты ω , $\delta -$ модулирующий сигнал $\Omega = 2\pi f$, s - АМ колебание при коэффициенте модуляции K < 1, z - АМ с частичным подавлением несущей (K > 1), $\partial -$ балансная АМ (полное подавление несущей, K >> 1)

Строго говоря, для однозначности терминологии *амплитудой* следовало бы называть величину a(t), а величину A – амплитудным значением, магнитудой колебаний или иначе. Далее величину a(t) будем называть напряженностью поля, сохранив для A традици-

онное название – амплитуда. Отметим также, что приведенное выражение для напряженности a(t) относится только к модуляции монохроматического, гармонического колебания, для которого определена несущая частота ω , то есть к модуляции напряженности, но, вообще говоря, не к модуляции интенсивности.

В приведенном выражении (и далее) модулирующая частота f представлена как обычная (техническая) частота (количество периодов колебаний за единицу времени); несущая частота с представлена как циклическая (изменение фазы за единицу времени). Для такого противопоставления и выбора величин существуют видимые причины. В оптическом диапазоне несущие частоты обычно предопределены; в большинстве случаев оптические частоты не измеряют непосредственно, и возможности их изменения очень ограничены. Чаще всего оптическая частота задана, например, спектральной линией излучения лазера, то есть спектральной линией атома, молекулы или иона. В этом случае выбор представления несущей частоты практически безразличен; световые колебания и волны естественнее описывать длиной волны в вакууме, с которой частота связана формально. Потому представление несущей частоты в оптическом диапазоне как частоты циклической удобнее, так как выражение фазы проще.

Напротив, модулирующие частоты по самому своему происхождению – частоты *технические*; колебания на этих частотах генерируют электронные или электромеханические устройства. Эти частоты относятся к радиодиапазону, и их можно измерять непосредственно, пересчетом периодов колебаний за единицу времени. Поэтому естественным будет представление модулирующих частот как частот технических, формально связанных с периодом колебаний. Правда, за это каждый раз приходится платить по « 2π » в выражениях для фазы. Когда для записи математических соотношений более удобно представление модулирующей частоты *f* как циклической, мы будем использовать обозначение $\Omega \equiv 2\pi f$.

Указанное соглашение об обозначениях мы будем соблюдать и далее, применительно к прочим видам модуляции.

Рассмотрим некоторые частные особенности АМ *напряженно*сти поля электромагнитных колебаний.

Легко убедиться в том, что в спектре АМ сигнала присутствуют две боковых компоненты, расположенные симметрично относи-

тельно несущей, $\omega \pm 2\pi f$ (см. рис. 2.3). Амплитуды боковых частот в выбранном представлении равны AK/2.

При АМ монохроматической волны коэффициент модуляции K в общем случае ничем не ограничен; его принято считать положительным (K < 0 формально соответствует сдвигу начальной фазы φ на π , то есть на полупериод модулирующей частоты).



Рис. 2.3. Тональная амплитудная модуляция: слева – временное представление, справа – спектральное представление; *а* –АМ колебание при коэффициенте модуляции *K* < 1, *б* – АМ с частичным подавлением несущей (*K* > 1), *в* – балансная АМ (полное подавление несущей, *K* >> 1)

Из общего выражения для АМ колебаний видно, что устройство для АМ при |K| < 1 функционально представляет собой *двухквадрантый перемножитель*; напряженность как один сомножитель меняется на входе и на выходе, меняя знак, а модулирующий сомножитель при всех изменениях знак не меняет. В результате огибающая модулированной напряженности не проходит через ноль, и сама напряженность не меняет фазу при всех изменениях модулирующего воздействия. Если отобразить в декартовых координатах сочетание возможной напряженности поля *E* (по оси абсцисс) и модулирующего воздействия (по оси ординат), то они разместятся в двух квадрантах верхней полуплоскости.

Амплитуда несущей при K = 1 в два раза больше, чем амплитуда каждой из боковых компонент; соответственно мощность несущей в два раза больше суммарной мощности боковых компонент. При дальнейшем увеличении K > 1 доля мощности несущей уменьшается, и все большая часть мощности приходится на боковые компоненты спектра.

Режим АМ при K > 1 (рис. 2.3, б) радиотехники называют *перемодуляцией* и считают недопустимым в работе модуляторов и передатчиков радиосвязи. С более общей точки зрения такой режим нужно рассматривать как специальный; его можно использовать в технике связи, применив более сложное детектирование. Его называют АМ с *частичным подавлением несущей*. В соответствии с названием, в боковых частотах можно сосредоточить большую часть мощности передатчика; при этом на входе приемника улучшится отношение сигнал/шум.



Рис. 2.4. Отображение примитивных и сложных видов модуляции на комплексной плоскости; начальные фазы несущей частоты выбраны произвольно, чтобы разные траектории не перекрывались. *1* – АМ со слабой паразитной ФМ, *2* – то же, при *K* = 1, *3* – ФМ с небольшой паразитной АМ, *4* – ФМ с сильной паразитной АМ, *5* – балансная АМ с подавленной несущей (БАМ)

При *K* >> 1 (разумеется, при конечном *AK*) модулированное колебание можно представить как

$$a(t) \approx A' \cos(2\pi f t) \cos(\omega t)$$

с точностью до начальных фаз. Такой частный вид модуляции называют балансной модуляцией. Легко убедиться в том, что в спектре балансно-модулированного сигнала присутствуют две боковых компоненты, симметричные относительно несущей ω и отстоящие от нее на $2\pi f$, а фурье-компонента несущей частоты стремится к нулю (рис. 2.3, e); поэтому балансную модуляцию называют также *модуляцией с подавлением несущей*. Отображение некоторых простых и комбинированных видов модуляции на комплексную плоскость показано на рис. 2.4.



Рис. 2.5. Четырехквадрантное перемножение напряженностей: *a* – обобщенное представление, *б* – схема электрооптического модулятора Маха-Цендера, *в* – зависимость коэффициентов передачи модулятора по напряженности и интенсивности от управляющего напряжения *U*

Устройство для AM при |K| > 1 представляет собой *четырех-квадрантый перемножитель*; напряженность и модулирующий сомножитель могут менять знак. При переходе огибающей напряженности через ноль фаза напряженности меняется на π . Если ото-

бразить в декартовых координатах сочетание возможной напряженности поля E и модулирующего воздействия, то они разместятся в во всех четырех квадрантах (рис. 2.5).

Перемножители сигналов в радиоэлектронике основаны на свойствах нелинейных элементов и цепей, реализующих эту функцию с ограниченной точностью. В оптике речь может идти о перемножении оптического и электрического сигналов, а также о перемножении двух оптических сигналов; последний вариант менее эффективен по энергетике, но может обеспечить наивысшее быстродействие.



Техническое исполнение балансного модулятора (четырехквадрантного перемножителя) с электрическим управлением в оптике возможно на основе электрооптических *амплитудных* модуляторов типа Маха-Цендера или Поккельса; они описаны в разделе 4 и в приложениях 1, 2, 4.

При передаче сигналов модулированными колебаниями энергию собственно сигнала переносят боковые полосы спектра, а на передачу несущей приходится тратить энергию, никак не связанную с сигналом. Далее, две боковых полосы занимают полосу спектра в два раза шире, чем спектр сигнала. Напрашивается решение – убрать из спектра излучения передатчика одну боковую и частично (или полностью) подавить несущую (рис. 2.6). Так можно сосредоточить большую часть энергии передатчика в информативной боковой полосе и не расходовать энергию на передачу несущей и избыточной второй боковой. Этот вид модуляции называют *од*нополосной (ОБП = одна боковая полоса, в международной терминологии – SSB = single sideband). Встречается и название VSB (vestigial sideband, рудиментарная боковая полоса, см. рис. 2.6, δ , рис. 2.7) – частичное подавление одной боковой и несущей, как обобщение SSB.



Рис. 2.7. Однополосная тональная модуляция и ее сравнение с AM: слева – временное представление, справа – спектральное представление; *a* – однополосная модуляция с частичным подавлением несущей (VSB) при *K* < 1, *δ* – SSB с подавлением несущей, *s* – балансная двухполосная AM

Излучение с ОБП в радиотехнике применяют в разных вариантах – с частичным или полным подавлением зеркальной боковой полосы и несущей. В частности, при передаче сигналов телевещания применяют излучение ОБП с частичным подавлением несущей. В результате спектр сигнала одного канала в эфире занимает примерно 7 МГц. Расстояние между несущими соседних каналов по международным стандартам равно 8 МГц, что было бы недостижимо при двухполосном излучении с шириной спектра 14 МГц.

В радиодиапазоне применяют два способа модуляции с ОБП – фильтровый и фазовый. Фильтровый способ выделения ОБП из сигнала АМ трудно переносить в оптику, и его мы рассматривать не будем. Фазовый способ основан на суммировании пары АМ сигналов с подавленной несущей. Их формируют балансные модуляторы, т.е. перемножители сигналов:

 $a(t) = \cos(2\pi f t) \times \cos(\omega t)$ $b(t) = \sin(2\pi f t) \times \sin(\omega t)$

Из приведенных выражений видно, что на перемножитель (а) приходят напряжения несущей и модулирующий сигнал, а на перемножитель (b) – те же напряжения, но сдвинутые по фазе на $\pi/2$.

После элементарных преобразований можно получить:

 $a(t) + b(t) = \cos[(\omega + 2\pi f)t]$ – верхняя боковая полоса,

 $a(t) - b(t) = \cos[(\omega - 2\pi f)t]$ – нижняя боковая полоса.

Отсюда видно, что одну из двух боковых выбирают, изменяя на π фазу одного из двух балансно-модулированных колебаний. Так как формирование ОБП описанным способом требует главным образом операций над фазами исходных и промежуточных напряжений, то и сам способ называют фазовым. В оптике для управления сдвигом фаз на несущей частоте ω используют управляемые фазовращатели – фазовые модуляторы Поккельса или их полупроводниковые аналоги.

Можно показать (для этого школьной тригонометрии достаточно), что малые погрешности балансной модуляции приводят к появлению в спектре сигнала ОБП несущей в том же порядке малости. Малые погрешности в сдвигах фаз и малый разбаланс амплитуд a(t) и b(t) приводят к появлению второй боковой в том же порядке малости.

В оптике формирование ОБП фазовым методом в принципе возможно, но труднее, нежели в радиотехнике; требуется сложная в устройстве и настройке система. Тем не менее, фазовое формирование сигналов ОБП реализовано по крайней мере в экспериментах по оптической связи; здесь наиболее удобно применение *интегрально-оптических* устройств для балансной модуляции и управления сдвигами фаз. Возможный вариант однополосного модулятора такого типа показан на рис. 2.8; он составлен из четырех одинаковых структур модуляторов. Две структуры работают как балансные модуляторы, а две – как настраиваемые устройства управляемых задержек по фазе. Эти последние структуры не обязательны; их функции могут исполнять и балансные модуляторы при добавлении к управляющим напряжениям корректирующих постоянных составляющих. Тем не менее, показанный вариант допускает более гибкую подстройку режимов работы, при этом стоимость интегрального оптического модуля с четырьмя структурами ненамного выше стоимости упрощенного модуля.



Рис. 2.8. Блок-схема фазового формирования однополосного сигнала *a* – общая схема, *б* – оптическая схема с применением модуляторов Маха-Цендера в волноводном варианте

О применении модуляции с ОБП для оптической связи сказано в приложении 4 («Современные типы модуляции оптических волн»).

На практике часто применяют специфический для оптики способ формирования ОБП – акустооптическую модуляцию, то есть когерентное рассеяние света на бегущей ультразвуковой волне в твердых или жидких прозрачных средах. Этот метод подробнее описан в разделе 5 («Акустооптическая модуляция»).

Для приема световых АМ сигналов используют *прямое детектирование* фотодиодами, фотоэлектронными умножителями (ФЭУ) или тепловыми приемниками (болометрами). *Фотодетекторы* (фотоприемники, фотодатчики, фотоэлементы, фотоэлектрические преобразователи – термины-синонимы) по характеру протекающих в них при действии излучения физических процессов подразделяют на *квантовые* и *тепловые* приемники. В квантовых приемниках поглощенное излучение генерирует носители заряда, образующие электрический сигнал (ток или напряжение). В идеализированном квантовом приемнике каждый поглощенный в нем квант излучения создает свободный (подвижный) носитель заряда (или пару носителей заряда разных знаков). Эти явления называют фотоэффектами, и по своей природе все они квантовые. Они наиболее сильны в полупроводниках, и квантовые полупроводниковые фотоприемники превосходят все прочие по быстродействию.

В тепловых приемниках энергия падающего излучения переходит в тепло, под действием которого изменяются электрические параметры приемного элемента. Термоэлектрические эффекты проявляются и в полупроводниках, и в диэлектриках, и в сверхпроводниках, и в металлах. Быстродействие тепловых приемников (полоса пропускания от единиц до десятков герц) ограничено их теплоемкостью и иными факторами, специфичными для конкретных приборов. Основная область применения тепловых приемников излучения – оптико-физические измерения, обычно в тех диапазонах спектра, где квант излучения hv < 0,1 эВ, и уже трудно использовать полупроводниковые приёмники. В технике связи и информационной технике тепловые приемники не применяют.

Все приемники излучения, независимо от принципа их действия, представляют собой квадратичные детекторы напряженности электромагнитного поля, то есть они вырабатывают электрический сигнал, напряжение или ток которого пропорциональны падающей на приемник мощности излучения.

Покажем, как абстрактный квадратичный детектор выделяет сигнал, передаваемый с помощью АМ напряженности поля (амплитуды). Напряженность АМ-поля на детекторе

 $a(t) = A \times [1 + K \cos(2\pi ft)] \cos(\omega t).$

После возведения в квадрат и усреднения по времени на интервалах $T>>1/\omega$ (что равносильно фильтрации частот ниже несущей) получим

$$a^{2}(t) = A^{2} \times [2K\cos(2\pi ft) + 1 + K^{2}\cos(2\times 2\pi ft)]$$

Первое слагаемое – сигнал на частоте модуляции, второе – постоянная составляющая, третье – вторая гармоника сигнала (искажение при детектировании, пропорциональное квадрату коэффициента модуляции). Когда АМ применяют для передачи аналоговых сигналов, приходится учитывать искажения, вносимые квадратичным детектированием; полумерой будет снижение коэффициента модуляции до уровня $K \ll 1$, так как соотношение сигнала к шуму на входе приемника может оказаться неприемлемо низким. Сигнал на частоте модуляции f приемник может выделить, если его электрическая полоса пропускания шире (или хотя бы ненамного уже), нежели частота модуляции f. У квантовых приемников полоса пропускания может превышать 10^{10} Гц, а у тепловых приемников редко достигает 1 кГц.

Для АМ сигналов применяют и *гетеродинный* прием, то есть преобразование светового сигнала в электрический при смешении на фотодетекторе сигнальной волны с волной от вспомогательного генератора – гетеродина; так образуется электрический сигнал биений двух волн на разностной (*промежуточной*) частоте, уже в радиодиапазоне. Для гетеродинного приема в оптическом диапазоне пригодны только быстродействующие квантовые приемники.

Положим, что волна гетеродина E_R когерентна, а сигнальную волну E_S описывает спектр Фурье, плотность которого сосредоточена в окрестности частоты гетеродина ω_0 ;

$$E_{s}(t) = \int E(\omega - \omega_{0}) \exp(i\omega t) d\omega$$
$$E_{R}(t) = E_{0} \exp(i\omega_{0}t)$$

В оптическом диапазоне обычно выполнено условие

$$|\omega - \omega_0| \ll \omega_0$$

которое означает, что сигнальная волна – квазимонохроматическая.

Интенсивность поля при совместном действии волн сигнала и гетеродина

$$I(t) = (E_s + E_R)(E_s^* + E_R^*) =$$

= $f(E_0, E) + 2\int E_0 E(\omega - \omega_0) \exp(i\omega t - i\omega_0 t) d\omega =$
= $2E_0 \int E(\omega - \omega_0) \exp[(i(\omega - \omega_0)t)] d(\omega - \omega_0) + const$

Далее взаимную интенсивность усредним по многим периодам оптических колебаний; время усреднения примерно соответствует быстродействию фотоприемника. Колебания интенсивности во времени описывает интеграл Фурье

$$I(t) = 2E_0 \int E(\omega_i) \exp(i\omega_i t) d\omega_i ,$$

где $\omega_i = \omega - \omega_0$ – промежуточная частота. Этот результат означает, что спектр колебаний *интенсивности* (и сигнала фототока) повторяет форму спектра оптических колебаний *напряженности* поля сигнальной волны, но со сдвигом по частоте вниз, ровно на частоту гетеродина (рис. 2.9).



Рис. 2.9. Преобразование спектра модулированных колебаний при гетеродинном приеме:
a – преобразование к промежуточной частоте,
б – то же, при очень низкой промежуточной частоте,
в – прямое преобразование к нулевой промежуточной частоте;
*ω*_c – несущая, *ω*_i – промежуточная частота, *ω*_h – частота гетеродина

Отметим как особо важное свойство гетеродинного преобразования, видимое из последнего выражения, что при $E_{\rm R} >> E_{\rm S}$ интенсивность излучения, создающая сигнал фотоприемника на промежуточной частоте,

$$2(E_{\rm R} \times E_{\rm S}) >> E_{\rm S}^2$$

Следовательно, гетеродинное преобразование в сравнении с прямым детектированием может дать на порядки более высокую чувствительность. Ограничения чувствительности на практике вызваны влиянием шумов реального гетеродина и максимальной допустимой мощностью излучения на входе фотоприемника–смесителя.

После гетеродинного переноса спектра сигнала из оптического диапазона в радиодиапазон с ним работают средствами радиотехники; далее их рассматривать не будем, но укажем существенную черту, одинаково важную для сигналов с любым типом модуляции.

Гетеродинный прием переносит когерентность сигнальной волны в когерентность колебаний электрического сигнала, если время когерентности излучения гетеродина существенно больше времени когерентности сигнальной волны; все вариации из сигнальной волны переносятся в вариации электрического сигнала на ПЧ. Если же степень когерентности у гетеродина сравнима со степенью когерентности сигнальной волны, то к флуктуациям амплитуды и фазы в электрическом сигнале ПЧ добавятся и соответствующие флуктуации гетеродина. В спектральном представлении это можно выразить так: спектр оптического сигнала без изменений сдвигается в область спектра электрического сигнала ПЧ. а при конечной ширине спектра гетеродина спектр электрического сигнала дополнительно уширяется. Эти утверждения легко доказать и формально, если к фазам сигнала и гетеродина в вышеприведенных выкладках добавить медленные флуктуирующие добавки $\phi(t)$; в фазе электрического сигнала на ПЧ они просто сложатся и дадут флуктуации общей фазы ПЧ, $\{i\omega_i t + i\phi(t)\}$. При гетеродинном преобразовании в радиодиапазон переносится не только спектр колебаний, но и все фазовые соотношения между всеми компонентами спектра. Иначе говоря, после гетеродинного преобразования сохраняется вся информация, которая ранее присутствовала в модулированном сигнале.



Рис. 2.10. Зеркальные каналы гетеродинного приема: *a* – область несущих частот, *б* – область промежуточных частот; ω_c – несущая в основном канале, ω_i – промежуточная частота, ω_m –несущая в зеркальном канале, ω_h – частота гетеродина

Гетеродинный приемник без предварительной фильтрации сигнала перед его входом имеет два «зеркальных» частотных канала приема – выше и ниже частоты гетеродина на ω_i .(рис. 2.10, *a*). В радиодиапазоне более или менее эффективная фильтрация в спектре входных сигналов возможна, но в оптическом диапазоне она крайне затруднена или просто невозможна; чем ниже промежуточная частота ω_i , тем труднее. Эти обстоятельства нужно учитывать при гетеродинном приеме. Если частота гетеродина расположена в диапазоне спектра принимаемого сигнала, то две его боковых полосы при переносе в диапазон промежуточных частот дают два сигнала. Один из них (из боковой полосы, рассеченной часто-
той гетеродина), получается искаженным (см. рис.2.9, б). Поэтому частоту гетеродина в норме нужно устанавливать за пределами спектра сигнала, и промежуточная частота при этом получается существенно выше, чем ширина спектра принимаемого сигнала.

Отметим случаи, когда частоту гетеродина все-таки позволительно установить в пределах спектра принимаемого сигнала.

Если частота гетеродина в точности совпадает с несущей, то получается так называемый прием с прямым преобразованием (к «нулевой» промежуточной частоте). В оптике такой случай можно встретить в измерительных системах (спектрометрах, локаторах и т.п.), где генератор несущей и гетеродин – одно и то же устройство. В технике оптической связи этот режим реализовать труднее, но в принципе возможно; для этого частоту и фазу гетеродина при помощи следящих систем в приемнике привязывают к несущей частоте. Этот подход лежит в основе прогрессивных методов когерентного приема и обработки сигналов, в том числе и в оптоволоконной связи.

Если сигнал с широким спектром шумоподобен, и в нем нет выраженной (по увеличенной спектральной плотности) несущей частоты, то частоту гетеродина имеет смысл поместить вблизи от максимума спектральной плотности принимаемого сигнала. В этом случае полезным сигналом разумно считать только мощность излучения (информация о фазах может быть при этом потеряна). Такую методику применяют для детектирования широкополосного (например, теплового) излучения, модулированного по интенсивности.

Амплитудная модуляция интенсивности. Тональную модуляцию интенсивности W(t) в простейшем случае описывает выражение

 $W(t) = W_0 \times [1 + K \cos(2\pi f t + \varphi)], \quad W(t) \ge 0.$

При тональной модуляции интенсивности широкополосного излучения коэффициент модуляции K принципиально ограничен; интенсивность не может быть отрицательной, и потому 0 < K < 1. Формально это неравенство должно быть нестрогим, на практике получать K = 0 или K = 1 невозможно и обычно не нужно. В более общем случае, если |K| > 1, модулированные колебания не относят к тонально-модулированным, и считают, что спектр модулирующего сигнала сложнее, чем простой тон (то есть гармоническое колебание на частоте модуляции).

Модуляцию напряженности монохроматического колебания можно формально представить и как модуляцию интенсивности, но в этом представлении не сохранится информация о фазе колебаний несущей частоты. Представление модуляции интенсивности как перемножения сигналов сводится к одноквадрантному перемножению, независимо от технического варианта.

Модуляцию интенсивности широкополосного излучения (немонохроматического колебания) представить как модуляцию напряженности в общем случае невозможно; для этого представления не определена несущая частота. Тем не менее, всякое реальное излучение отличается от строго монохроматического, и его описывает спектр конечной ширины и некоторая средняя по спектру частота. Поэтому правомерен вопрос: насколько значима ширина спектра в условиях конкретного наблюдения или технического устройства.

Уточним смысл термина «широкополосное излучение». До сих пор мы противопоставляли его монохроматическому излучению, обращаясь только к интуитивному восприятию. При более внимательном взгляде такое противопоставление не будет очевидным. В разных применениях и разных технических устройствах количественный критерий монохроматичности будет различаться на порядки. Общий подход дает понятие *временной когерентности* излучения и ее количественные выражения через время и длину когерентности.

Если речь идет о каких-либо соотношениях фаз в напряженности поля, во времени или по трассе распространения волны, то такие соотношения фаз слабо меняются только за время когерентности или, соответственно, на длине когерентности. Следовательно, оценку степени (достаточности) монохроматичности дает сопоставление характерной длины системы или продолжительности процесса, где важны соотношения фаз, с длиной или временем когерентности (а эти величины просто связаны с шириной спектра излучения). Например, если в устройстве, использующем интерференцию двух волн, максимальная разность хода равна L, то по отношению к этому прибору излучение с длиной когерентности *l* ~ 10*L* разумно считать достаточно монохроматичным; ширина его спектра $\Delta \omega$ должна быть менее c/10L. Излучение с такой мерой когерентности (в этих конкретных обстоятельствах) называют квазимонохроматическим; в таком случае можно наблюдать интерференцию и прочие признаки колебаний, близких к монохроматическим. В частности, можно с ограниченной ($\Delta \omega$) точностью говорить о средней несущей частоте и о спектре боковых компонент, каждая из которых в этом случае уширена на $\Delta \omega$. Отметим, что даже при $\Delta \omega \sim 2\pi f$ можно говорить об уширенном спектре боковых частот (полос); при более сильном условии $\Delta \omega << 2\pi f$ это уже вполне корректно, так как спектральная область «несущих частот» не перекрывается с боковыми полосами.





Из сказанного следует, что для электрически управляемой модуляции интенсивности широкополосного излучения можно применять модуляторы с управляемым поглощением и интерференционные модуляторы с *минимальной* начальной разностью фаз – на жидких кристаллах и типа Маха-Цендера (рис. 2.11). Подробнее возможности электрооптической модуляции интенсивности немонохроматического излучения рассмотрены в разделе 4, где описан компенсированный модулятор Поккельса.

Идеальный квадратичный детектор излучения воспринимает излучение, модулированное по интенсивности, без искажения восстанавливая закон модуляции, независимо от ширины спектра «несущих частот». При этом, разумеется, возможно заметное ослабление верхних частот из-за ограниченного быстродействия детектора и/или электронного усилителя его сигнала. В информационной оптике модуляция интенсивности излучения с широким спектром встречается очень часто, и в зависимости от области применения она имеет разные частные особенности. Оставив в стороне примитивные механические и электромеханические прерыватели, приведем более современные примеры.

Для передачи сигналов по оптическим линиям связи с характерными частотами модуляции до десятков мегагерц применяют светодиодные излучатели с прямым управлением мощности излучения током питания; техника получается несложной, и специальные модуляторы интенсивности для этого не нужны. Граница применимости этой техники – 10 Mbit Fast Ethernet.

Массовое применение нашли жидкокристаллические модуляторы широкополосного излучения видимого диапазона – экраны видеомониторов, телевизоров, проекторы, индикаторы, дисплейные панели измерительных приборов и информационных табло, и многие подобные устройства. Для большинства таких применений достаточны частоты модуляции до 1 кГц (мониторы ПЭВМ, телеэкраны HD TV); здесь модулированное излучение предназначено только для восприятия человеческим глазом.

В жидкокристаллических модуляторах интенсивности пропускание зависит от разности фаз между обыкновенной и необыкновенной волнами (или сводится к ней); начальная разность фаз при этом может быть порядка π ; в таком случае длина когерентности света может быть не более нескольких длин волны. Повседневный опыт показывает, насколько эффективно работают жидкокристаллические экраны и как высок контраст воспроизводимых ими изображений. Прочие модуляторы, в основе действия которых лежит интерференция, также эффективны, если могут работать при малой начальной разности фаз. Подробнее эти вопросы изложены в разделе 4.

2.2. Частотная модуляция

Понятие «частотная модуляция» (ЧМ) применяют только к монохроматическим исходным колебаниям, частота которых имеет достаточно определенное значение; в оптике такие колебания дают только одночастотные лазеры. ЧМ – управление частотой колебаний. Наиболее прямой путь – перестройка частоты автоколебаний изменением параметров резонансной системы автогенератора. ЧМ-колебания простейшего вида

 $a(t) = A \cos[(\omega + D\cos(2\pi ft))t]$

описывают следующими параметрами:

- 1) амплитуда *A*,
- 2) несущая (частота) ω,
- 3) модулирующая частота f,
- 4) девиация D.

В оптике, в отличие от радиодиапазона, практически удается реализовать только ЧМ с девиацией и модулирующей частотой, очень малыми в сравнении с несущей. Этот тип модуляции называют узкополосной ЧМ.

На первый взгляд может показаться, что для ЧМ необходим автогенератор с управляемой частотой автоколебаний. На самом деле ЧМ можно реализовать и как иной по техническим средствам, но родственный вид модуляции – фазовую модуляцию (ФМ).

Фаза $\varphi(t)$ и частота $\omega(t)$ неразрывно связаны между собой соотношением

$$\omega(t) = d\varphi/dt.$$

Поэтому частотная и фазовая модуляция взаимосвязаны, и часто говорят о ЧМ, имея в виду ФМ, и наоборот. ФМ и ЧМ нередко собирательно называют угловой модуляцией.

2.3. Фазовая модуляция

Понятие «фазовая модуляция» применяют только к монохроматическим исходным колебаниям, *изменениям* фазы которых можно приписать достаточно определенное значение; в оптике такие колебания дают только одночастотные лазеры. Как уже было показано в предыдущем разделе, ФМ по сути родственна ЧМ. Если для ЧМ в узком собственном смысле нужен перестраиваемый автогенератор, то для ФМ в столь же узком собственном смысле нужно управляемое устройство задержки или фазовращатель. В зависимости от удобства технической реализации, в конкретных системах применяют управление либо частотой (*внутри* генератора), либо фазой (*за пределами* генератора), а управляющее воздействие (электрический сигнал) предварительно подвергают либо интегрированию, либо дифференцированию по времени. Общий вид колебаний с угловой модуляцией показан на рис. 2.12. ФМ-колебания простейшего вида (тональную ФМ)

 $a(t) = A \cos\{\omega t + \beta \cos(2\pi f t)\}$

описывают следующими параметрами:

- 1) амплитуда *A*,
- 2) несущая (частота) ω,
- 3) модулирующая частота f,
- 4) индекс модуляции $\beta \ge 0$.

Результат ЧМ и ФМ можно описывать двояко – и параметрами ЧМ, и параметрами ФМ. Например, при *тональной* (одночастотной синусоидальной) фазовой модуляции, когда модулирующая циклическая частота равна $2\pi f \equiv \Omega$ и текущее отклонение фазы равно

$$\varphi(t) = \omega t + \beta \sin(\Omega t),$$

текущую частоту описывает производная от фазы по времени $\omega(t) = \omega + \beta \Omega \cos(\Omega t).$

Тональная ЧМ эквивалентна тональной ФМ, но со сдвигом по фазе на 90° и индексом модуляции $\beta = D/\Omega$. Если управляющий сигнал гармонический, то при равенстве несущих и модулирующих частот ФМ колебания не отличаются от ЧМ, а при равенстве их девиации или индексов модуляции они тождественны; девиация

$D = \beta \Omega$.

При чистой ФМ и ЧМ интенсивность модулированной волны во времени не изменяется. Это лучше всего видно при комплексном представлении модулированного колебания, например, ФМ:

 $a(t) = A\exp[i(\omega t + \beta \cos(2\pi ft))];$

так как выражения в круглых скобках вещественные, то $a(t)a^*(t) = A^2$. Для ЧМ тот же результат получают аналогично.

Различают как предельные два случая ФМ. Первый случай – аналог узкополосной ФМ, применяемой для радиосвязи. Второй случай – аналог сдвига частоты при эффекте Доплера.

Узкополосная ФМ. В этом случае индекс модуляции гораздо менее единицы, и спектр *модулирующих* частот не очень широкий.

При тональной ФМ в спектре излучения появляются боковые компоненты, отстоящие от несущей частоты на интервалы, кратные модулирующей частоте, $n\Omega$. Амплитуда *n*-ной боковой компоненты зависит от индекса модуляции – она пропорциональна $J_n(\beta)$, значению функции Бесселя целого порядка *n* при аргументе β

(см. далее). С увеличением индекса модуляции растет и доля энергии, переносимой в боковые полосы.



Рис. 2.12. Тональная частотная/фазовая (угловая) модуляция:
 а – колебания несущей частоты ω, б – модулирующий сигнал на частоте Ω =2πf,
 в – ЧМ/ФМ колебание. Параметры модуляции для наглядности преувеличены в сравнении с типичными для систем связи:
 девиация D ~ ω/2, модулирующая частота Ω ~ ω/10, индекс модуляции β ~ 5

Покажем, что доля мощности боковых компонент в сравнении с несущей при индексе модуляции β < 1 мала, и спектр боковых компонент сосредоточен вблизи несущей.

Рассмотрим фазовую модуляцию несущей ω модулирующей частотой $2\pi f$ при индексе модуляции $\beta << 1$;

$$E(t) = \exp\{i(\omega t + \beta \cos(2\pi f t))\} =$$

 $= \exp(i\omega t) \times \exp(i\beta\cos(2\pi ft)).$

Далее потребуется (из справочника по свойствам специальных функций) разложение в ряд Фурье

$$\exp(i\beta\cos\varphi) = \sum_{n=-\infty}^{n=\infty} i^n J_n(\beta) \exp(in\varphi);$$

отсюда получим общее выражение спектра E(t) при ФМ (рис. 2.13):

$$E(t) = \exp\{i(\omega t + \beta \cos(2\pi f t))\} =$$

= $\sum_{n=-\infty}^{n=\infty} i^n J_n(\beta) \exp\{i[\omega + 2\pi f n]t\}.$

Далее (из того же справочника) возьмем представление функций Бесселя *J_n* суммой степенного ряда

$$J_{n}(\beta) \approx \left(\beta/2\right)^{n} \sum_{k=0}^{\infty} \frac{\left(-1\right)^{k}}{\Gamma(k+1)\Gamma(k+n+1)} \left(\beta/2\right)^{2k}$$

Сохраняя слагаемые не выше второго порядка малости по β, получим

$$J_0(\beta) \approx 1 - \beta^2 / 4, \quad J_1(\beta) \approx \beta / 2.$$

Пренебрегая слагаемыми выше первого порядка малости по β , получим

$$\exp(i\beta\cos\varphi) \approx$$

$$\approx J_0(\beta) + iJ_1(\beta)\exp(i\varphi) - iJ_{-1}(\beta)\exp(-i\varphi) \approx$$

$$\approx 1 + J_1(\beta)\exp(i\varphi + i\pi/2) + J_{-1}(\beta)\exp(-i\varphi - i\pi/2).$$

Теперь подставим

 $\phi = 2\pi ft$, $J_1(\beta) \approx \beta/2$, $J_{-1}(\beta) = -J_1(\beta)$, $J_0(\beta) \approx 1$ и получим

$$E(t) \approx \exp(i\omega t) +$$

+ (\beta/2) \exp{i(\omega + 2\pi f)t + i\pi/2} -
- (\beta/2) \exp{i(\omega - 2\pi f)t - i\pi/2}

В результате получено приближенное представление фазовомодулированного колебания при малом индексе модуляции. В его спектре присутствуют несущая ω и две боковые, $\omega \pm 2\pi f$, смещенные в разные стороны от несущей ω на частоту модуляции $2\pi f$. Прочие боковые спектральные компоненты имеют пренебрежимо малые амплитуды (рис. 2.13).

Следовательно, Фурье-спектр интенсивности фазовомодулированного колебания с малым индексом и несущей частотой ω_0 имеет вид

 $F(\omega) \approx \delta(\omega - \omega_0) + (\beta/2)^2 \delta(\omega - \omega_0 - 2\pi ft) + (\beta/2)^2 \delta(\omega - \omega_0 + 2\pi ft)$, Сравним ФМ-колебание с АМ-колебанием при той же несущей и модулирующей частоте, при коэффициенте модуляции K < 1:



Рис. 2.13. Тональная фазовая модуляция во временном и в спектральном представлениях. Тональная частотная модуляция отобразится так же, если девиация *D* = βΩ при той же частоте *f*

Спектр интенсивности при ФМ с малым индексом модуляции похож на спектр интенсивности при AM (рис. 2.13, δ). В обоих спектрах присутствуют несущая и две боковых на тех же расстояниях от несущей. Но ФМ и AM колебания принципиально различны в том, что между двумя этими типами модуляции различаются фазовые соотношения между боковыми частотами и несущей. В спектрах интенсивности информация о фазах полностью потеряна, но в напряженности поля фазы играют существенную роль. Позже мы покажем это на примере квадратичного детектора излучения, принимающего AM и ФМ сигналы.

При более сложном спектре ФМ спектр интенсивности получают ют аналогично: каждой спектральной компоненте приписывают собственную интенсивность, равную квадрату модуля амплитуды (без множителя 1/2), и ее собственную частоту. Действительно, интенсивность излучения можно выразить как усредненное по интервалу времени, много большему периода, произведение

$$\langle EE^* \rangle = W,$$

$$W = \left\langle \sum_k a_k \exp(i\omega_k t) \times \sum_j a_j^* \exp(-i\omega_j t) \right\rangle =$$

$$= \left\langle \sum_{k,j} \left(a_k \exp(i\omega_k t) \times a_j^* \exp(-i\omega_j t) \right) \right\rangle = \sum_k a_k a_k^*.$$

Для более строгого обоснования можно применить метод полной индукции. В приведенной цепочке равенств неявно учтено, что средние по времени

$$\langle \exp(i\omega_k t) \times \exp(-i\omega_j t) \rangle = \delta(k, j),$$

то есть между разными Фурье-компонентами корреляции нет (они ортогональны), и их произведения вклада в сумму W не дают. При шумоподобной модуляции спектр интенсивности выглядит как небольшое уширение исходного, номинально монохроматического, спектра.

<u>Небольшое отступление</u>. Отсутствие корреляции между разными Фурье-компонентами – центральное положение известной теоремы Винера-Хинчина. Это положение, однако, нельзя использовать бездумно, и вот тому пара примеров.

Пример первый. При амплитудно-модулированном колебании интенсивность совершает колебания на частоте модуляции; спектр его состоит из трех компонент – несущей и двух боковых. Если эти три компоненты взаимно не коррелированны, то регулярных периодических колебаний интенсивности быть не должно. Причина кажущегося парадокса заключена в том, что в данном случае нет никакого случайного процесса, тогда как при выводе теоремы Винера-Хинчина идет усреднение по реализациям стационарного статистического ансамбля. Иначе говоря, рассмотренный пример относится к единственной реализации; вдобавок и усреднение по времени лишь подразумевалось, а фактически методика наблюдения интенсивности АМ колебаний его выполняет только за время, много большее периода несущей, но не много большее периода модуляции. Кроме того, между двумя боковыми существует определенная корреляция: они не случайны, а имеют одинаковые амплитуды и в любой момент времени детерминировано связаны по фазам (ϕ_1, ϕ_2) так же, как и зубчатые колеса в часовом механизме, вращающиеся с разными угловыми скоростями (ω_1, ω_2) , соотношением

$$\phi_1/\omega_1 = \pm \phi_2/\omega_2 + const.$$

В конце концов, в приведенном примере выбрана не просто одна реализация из всего ансамбля, но реализация очень частная, специфическая. В этом примере статистический ансамбль образовали бы все возможные реализации разностей фаз между боковыми компонентами; каждая из реализаций в этом ансамбле выглядела бы как колебания интенсивности с различающимися магнитудами, в диапазоне от максимума до нуля. Распределение вероятностей магнитуд должно зависеть от статистических свойств ансамбля случайных фаз боковых частот.



Рис. 2.14. Отображение на комплексной плоскости амплитуд и фаз несущей и боковых компонент: *a* – AM, *б* – ФМ

Пример второй. При чистом фазово-модулированном колебании интенсивность вовсе не зависит от параметров фазовой модуляции, хотя спектр интенсивности определенно зависит от индекса модуляции и модулирующих частот. Если индекс ФМ мал, то в спектре присутствуют только несущая и две боковых частоты. Фазы этих боковых тоже сцеплены между собой, а также и с фазой несущей. Но это соотношение фаз (рис. 2.14, *a*) совсем иное, нежели при AM (рис. 2.14, *б*). В результате усреднение мощности ФМ-сигнала по большому числу периодов несущей не выявляет присутствия ФМ в этом сигнале при любом спектре модулирующих частот и при произвольном индексе модуляции. Отметим, что представление тональной ФМ как суммы несущей и двух боковых компонент – приближение, верное только при малом $\beta << 1$; расчет покажет, что в таком «усеченном» представлении при $\beta \sim 1$ и более интенсивность будто бы изменяется во времени. Можно доказать, что для удовлетворительной точности представления ФМ колебания необходимо учитывать все спектральные компоненты $\omega_0 \pm m\Omega$ примерно до $m \sim \beta$ (рис. 2.13, *б*, *в*, *г*); в таком уточненном представлении интенсивность будет изменяться пренебрежимо слабо.

Сопоставив два приведенных примера, можно убедиться в том, насколько грубо описывает модулированное колебание один только спектр интенсивности, и насколько важное значение имеют фазовые соотношения между отдельными компонентами спектра немонохроматического колебания.

На практике ФМ и ЧМ различают почти номинально, не по спектру, а по типу технических средств, применяемых для модуляции несущей или для демодуляции сигнала (восстановления формы модулирующего сигнала). Методы детектирования ЧМ сигналов описаны в следующем разделе, вместе с методами детектирования ФМ сигналов.

Между тональной ЧМ и Φ М, несмотря на родство, существует различие, которое проявляется при изменении модулирующей частоты Ω : если сохраняется индекс модуляции, то говорят о фазовой модуляции, если же сохраняется девиация $m\Omega$, то говорят о частотной модуляции.

Модуляция дискретным сигналом. Амплитудная, частотная и фазовая манипуляция. До сих пор речь шла о тональной модуляции, где спектр модулирующего сигнала примитивен. При сложном спектре модулирующего сигнала, когда m и Ω не фиксированы и изменяются во времени, различие между ЧМ и ФМ может быть более заметным. Наиболее наглядно различие между ЧМ и ФМ при быстрых, ступенчатых изменениях модулируемого параметра, обычно применяемых при передаче цифровой информации, а также в измерительной технике. Такой тип модуляции называют манипуляцией; этим термином описывают ручную работу на телеграфном ключе Морзе; в международной терминологии применяют термины «keying» и «on/off keying, OOK».

На практике манипуляцию от модуляции отличает только характер сигнала, управляющего модулятором того или иного класса.

При AM и ФМ обычно изменение модулируемого параметра мало отличается от модулирующего воздействия; при ЧМ в собственном смысле, когда модулирующий сигнал действует на управляемый элемент внутри петли обратной связи генератора, изменение частоты может заметно запаздывать от управляющего воздействия. Поэтому AM и ФМ – наиболее употребительные виды манипуляции, а ЧМ в системах скоростной связи не встречается.



Рис. 2.15. Временное представление амплитудной (*a*), частотной (б) и фазовой (в) манипуляции; *г* – колебания несущей частоты, *д* – управляющий сигнал с конечными скоростями нарастания и спада

Амплитудную, частотную и фазовую манипуляцию можно сопоставить наглядно (см. рис. 2.15 и 2.16). Резкие скачки при манипуляции – идеализации реального процесса. В действительности все переходы между логическими состояниями сигнала происходят за конечное время не только по техническим причинам, но их преднамеренно замедляют, чтобы сверх допустимого не расширять спектр модулированного излучения. Кроме того, изменения амплитуды часто происходят с некоторыми изменениями фазы. Так, на рис. 2.16 сплошными двунаправленными стрелками показаны возможные переходы между логическими битами, пунктирами – отклонения от идеализированных траекторий в реальных устройствах, вызванные паразитной ФМ, сопровождающей АМ.



Рис. 2.16. Отображение амплитудной и фазовой манипуляции на комплексной плоскости; начальная фаза несущей условно равна нулю;

- a-амплитудная манипуляция при коэффициенте модуляци
и $K\,{<}\,1,$
- a-амплитудная манипуляция при коэффициенте модуляци
иK=1

 δ – фазовая манипуляция,
 e– квадратурная фазовая манипуляция

Широкополосная ФМ или ЧМ (чирпинг). Во этом случае индекс модуляции много больше (или хотя бы порядка) единицы, но модулирующая частота довольно низкая. Если время наблюдения (измерения частоты или анализа спектра) гораздо менее полупериода частоты модуляции, то спектр модулированного колебания выглядит как исходный, но сдвинутый по отношению к нему. Знак сдвига зависит от направления изменения фазы. При последовательных наблюдениях медленная тональная ФМ выглядит как *частотная модуляция*, то есть как периодическое перемещение спектра около центральной частоты (несущей немодулированного колебания). Здесь вновь видно родство частотной и фазовой модуляции. Это легко понять, вспомнив, что частота есть производная от фазы по времени. Тогда приращение фазы, линейно зависящее от времени,

 $\Delta \varphi(t) = Dt$,

дает просто постоянный сдвиг частоты:

 $E(t) \propto \cos(\omega t + \Delta \varphi(t) + \varphi_0) = \cos((\omega + D)t + \varphi_0),$ а периодические изменения параметра *D* дают частотную модуляцию несущей ω с девиацией *D*₀ и модулирующей частотой *f*,

 $E(t) \propto \cos[(\omega + D_0 \cos(2\pi f t)t + \varphi_0]]$.

Очевидно, что дифференцирование и/или интегрирование управляющего и восстановленного сигнала дает возможность работать с ЧМ и ФМ как с единым типом модуляции. Например, в УКВ ЧМ вещании реально используют не ЧМ, а ФМ; таким способом ослабление верхних частот в восстановленном сигнале из-за ограниченной полосы пропускания приемника компенсируют увеличением доли этих верхних частот в спектре излучения передатчика. Если для формирования такого ФМ сигнала приходится использовать модуляцию частоты, то управляющий сигнал просто предварительно подвергают дифференцированию по времени (так наз. предискажению), то есть в нем намеренно поднимают верхние частоты в сравнении с нижними.

Детектирование ЧМ/ФМ сигналов. Простой квадратичный детектор (детектор интенсивности) здесь непригоден, так как при идеальной ФМ интенсивность модулированного излучения неизменна.

При тональной АМ напряженность поля

 $E(t) = [1 + K\cos(2\pi f t)]\exp(i\omega t),$

и интенсивность, на которую реагирует квадратичный детектор,

$$E(t)E^{*}(t) = [1 + K\cos(2\pi ft)]^{2} =$$

 $= 1 + 2K\cos(2\pi ft) + K^2\cos(2\times 2\pi ft)$

Для ФМ колебаний с малым индексом модуляции,

 $E(t) \approx \exp(i\omega t) +$

+
$$(\beta/2) \exp \{i(\omega + 2\pi f)t + i\pi/2\}$$

- $(\beta/2) \exp \{i(\omega - 2\pi f)t - i\pi/2\},\$

после несложных по существу, но пространных по объему расчетов получим выражение для интенсивности:

$$E(t)E^*(t) = 1;$$

этот результат точен и при произвольном индексе модуляции β , как это уже было показано ранее для общего случая ФМ. Отсюда следует, что квадратичное детектирование не может выявить ЧМ/ФМ

непосредственно, без применения специального частотного или фазового детектора (*дискриминатора*), в отличие от АМ.

Ранее было отмечено различие фазовых соотношений между несущей и боковыми в АМ и ФМ колебаниях. Это различие при квадратичном детектировании дает разные результаты для АМ и ФМ. Подчеркнем, что следует рассматривать фазовые соотношения не между какими-либо двумя спектральными компонентами модулированного колебания, а *между всеми* спектральными компонентами, включая и несущую. Нетрудно показать, что смешение излучения двух разных частот на квадратичном детекторе дает сигнал фототока на разностной частоте, и при любом соотношении между их фазами магнитуда колебаний фототока не меняется, а от разности фаз зависит только фаза колебаний фототока, обычно несущественная.

Если на квадратичном детекторе действуют *три* спектральных компоненты, то результат может быть качественно иным. Если *по-ложительные* частоты ω_{-} , ω_{0} , ω_{+} (по порядку увеличения) произвольные, то на выходе квадратичного детектора образуются колебания на частотах ($\omega_{0} - \omega_{-}$), ($\omega_{+} - \omega_{0}$), ($\omega_{+} - \omega_{-}$).

Если колебания на частотах ($\omega_0 - \omega_-$), ($\omega_+ - \omega_0$) *взаимно когерентны*, что возможно только при ($\omega_0 - \omega_-$) = ($\omega_+ - \omega_0$), то результат зависит от разности фаз между ними. В частности, если разность фаз равна π , а магнитуды одинаковы, то сигнал на частоте ($\omega_0 - \omega_-$) = ($\omega_+ - \omega_0$) на выходе детектора равен нулю. Этот случай, как нетрудно убедиться, соответствует ЧМ/ФМ.

Если же указанная разность фаз равна нулю, то сигнал на частоте ($\omega_0 - \omega_-$) = ($\omega_+ - \omega_0$) на выходе детектора максимален. Сигнал на частоте ($\omega_+ - \omega_-$), формально равной нулю, дает постоянную составляющую фототока. Произвольной, но *стабильной* разности фаз соответствует сложная совместная модуляция, АМ+ФМ. Такой же результат дает неравенство магнитуд колебаний ($\omega_0 - \omega_-$), ($\omega_+ - \omega_0$), то есть неравенство напряженностей в боковых компонентах. Отсюда следуют и два возможных вида дискриминаторов – частотный и фазовый (рис. 2.17).



Рис. 2.17. Частотные и фазовые дискриминаторы: *а* – частотный дискриминатор с резонансом пропускания, *б* – фазовый дискриминатор, *в* – векторная диаграмма для частотного дискриминатора, *г* – векторная диаграмма исходного сигнала с угловой модуляцией, *д* – векторная диаграмма для фазового дискриминатора

В частотном дискриминаторе (рис. 2.17, *a*, *в*) угловую модуляцию выявляет различие коэффициентов передачи верхней и нижней боковых частот к квадратичному детектору. В фазовом дискриминаторе (рис. 2.17, *б*, *д*) угловую модуляцию выявляет различие сдвигов фаз верхней и нижней боковых частот. На практике некоторые сдвиги фаз присутствуют в частотных дискриминаторах, а в фазовых дискриминаторах разница в коэффициенте передачи двух боковых может быть очень малой, так как обычно стараются настроить дискриминатор на резонанс с несущей частотой, как показано на рис. 2.17, *б*.

Крутизна дискриминационной характеристики, $dW/d\omega$ или ее аналог, $d\phi/d\omega$, взаимосвязаны с шириной резонанса, так или иначе формирующего частотную и фазовую характеристики дискримина-

тора. Поэтому чем больше крутизна, тем уже тот допустимый диапазон частот, в который должен уложиться спектр модулированного сигнала, чтобы его можно было восстановить в приемнике с минимальными искажениями.

Некоторые технические варианты дискриминаторов для оптического диапазона описаны далее, в разделе 3.1.

2.4. Управление поляризацией излучения

Изменение поляризации непосредственно для формирования световых сигналов в системах связи до сих пор применяют редко, но уже описаны результаты успешных испытаний волоконнооптических систем, где модуляция поляризации совмещена с другими методами модуляции для увеличения информационной нагрузки оптического сигнала. Два взаимно-ортогональных направления поляризации, как информативный признак оптического сигнала, прибавляют еще один двоичный разряд к слову-коду, переносимому интенсивностью и/или фазой светового импульса.

В амплитудных модуляторах света неволноводного типа (на монокристаллах) изменение поляризации излучения фактически присутствует как первый этап преобразования волны, вслед за которым линейный поляризатор преобразует изменения поляризации падающей световой волны в изменения амплитуды прошедшей волны. Это же обстоятельство указывает на метод детектирования изменений поляризации световой волны – первоначальное преобразование к АМ и дальнейшее детектирование АМ сигнала. Принципы действия электрооптических амплитудных модуляторов по традиции часто описывают, рассматривая именно изменения поляризации света и дальнейшее преобразование этих изменений в АМ при помощи поляризатора. Такое представление в общем верное, но не самое наглядное и удобное.

Другой вариант управления поляризацией излучения – сочетание поворота плоскости поляризации в ячейке Фарадея с двулучепреломляющей пластиной, в частности, с четвертьволновой. Кроме того, для преобразования линейной поляризации в эллиптическую можно применять ромб Френеля или иные его эквиваленты с призмами полного внутреннего отражения.

Наиболее часто модуляцию поляризации используют в измерительных системах – поляриметрах, зеемановских спектрополяриметрических устройствах и т.п.

2.5. Изменение направления волнового вектора

Этот вид модуляции обычно используют не для формирования сигналов в прямом смысле, а для изменения диаграммы направленности излучения в пространстве – в устройствах сканирования объектов световым лучом, для формирования бегущим лучом видимого изображения на экране, для перемещения луча в технологических установках по обрабатываемой детали и т.п. Соответствующие технические устройства модуляции обычно называют *дефлекторами* или *сканерами*.

Примитивный механический аналог дефлектора или сканера – вращающееся зеркало, призма или антенна радара. Их применяют там, где невысоки требования к быстродействию системы, и реальное быстродействие ограничивает не инерционность механической системы, а какие-либо иные факторы.

В радиотехнике работают и существенно более быстрые системы отклонения излучения. Их называют фазированными антенными решетками (ФАР). Такие устройства представляют собой многоэлементные одномерные или двумерные решетки диполей, к каждому из которых возбуждение подают через управляемый фазовращатель (рис. 2.18, a). В дальней зоне направление волнового вектора изменяется при изменении сдвига фаз между диполями за счет интерференции (модель Фраунгофера). Быстродействие ФАР зависит от скорости перестройки фазовращателей и на порядки превосходит быстродействие электромеханических дефлекторов.

На первый взгляд, в оптике построить аналог ФАР очень трудно. На самом деле известны чисто оптические методы эффективного решения этой задачи (рис. 2.18). Один из них – когерентное рассеяния света на ультразвуковых волнах в прозрачных материалах, в том числе и в жидкостях (рис. 2.18, *в*). Здесь работает дифракция Фраунгофера на решетке, от управляемого (частотой ультразвука) периода которой зависит угол отклонения в первый или минус первый интерференционный порядок. Примечательно то, что дефлекторы подобного типа (*акустооптические* модуляторы) дают возможность изменять не только направление световой волны в дальней зоне, но в одной и той же оптической схеме позволяют модулировать когерентный свет и по интенсивности, и по частоте.

Излучатель возбуждают сигналом y(t) на несущей Ω вида $y(t) = f(t)\sin\Omega t$. Информацию содержит функция f(t), а несущая Ω соответствует резонансной частоте излучателя. Сигнал y(t) возбуждает акустическую волну, бегущую в среде звукопровода со скоростью звука v и с амплитудой f(x - vt). Световая волна, пересекающая светозвукопровод по направлению y, в тот момент, когда весь сигнал y(t) уложился в апертуре, пространственно модулируется по фазе ультразвуковым сигналом. Индекс фазовой модуляции в данной точке примерно пропорционален амплитуде ультразвука; тогда под углом Брэгга распространяется изображение сигнала f(t).





- *а* фазированная решетка излучателей (переменная разность фаз φ_1);
 - δ дифракционная решетка (переменный угол поворота θ);
 - *в* акустооптический дефлектор (переменный период Л);
- *г* электрооптическая призма (переменный показатель преломления *n*);

Другой подход связан с применением решеток (мозаик) с одинаковыми элементами, в которых свет проходит через толстые пленки с электрически управляемым показателем преломления. Такими свойствами обладают пленки кристаллов с низкой симметрией элементарной ячейки, а особенно силен этот эффект в жидких кристаллах-нематиках. В недавнем прошлом появились и ныне привлекают к себе все большее внимание так называемые микроэлектромеханические системы (MEMC, MEMS) – механизмы с миниатюрными подвижными частями, изготовленные из материала кремниевых подложекзаготовок методами фотолитографии-травления, широко используемыми в современной технологии производства монолитных интегральных схем без существенных изменений. Минимальная масса подвижных частей MEMC позволяет достигать быстродействия большего, чем у систем с жидкими кристаллами; например, частоты переключения микрозеркал между крайними положениями могут достигать ультразвукового диапазона.

Управляемые фазовые решетки и МЕМС служат основой отдельного важного класса модуляторов света – пространственных модуляторов, которые пригодны не только для изменения направления волнового вектора примитивной уединенной волны, но и для управления формой волнового фронта в целом. В таких устройствах нуждаются системы для оптической обработки информации и адаптивная оптика (методы коррекции формы волновых фронтов и компенсации аберрации оптических систем). ПВМС для адаптивной оптики можно рассматривать как фазированные решетки с переменными параметрами, зависящими от координаты.

В особый класс пространственных модуляторов можно выделить голографические решетки, а среди них, в свою очередь, можно выделить решетки с *динамическим* управлением. Голографические решетки можно рассматривать как фазированные решетки, предназначенные для формирования диаграмм направленности, более сложных в сравнении с плоскими волнами.

Наибольший интерес представляют динамические решетки, управляемые оптическим излучением, в частности, полем интерференционной картины, образуемой при записи голограммы. Эти методы основаны на принципах нелинейной оптики; один из вариантов подробнее описан в разделе 7 («Обращение волнового фронта в фоторефрактивных кристаллах»).

В принципе для управления отклонением лучей можно использовать призмы из кристаллов с эффектом Поккельса или с эффектом фотоупругости, но диапазон углов отклонения при этом получается очень малым; такие элементы полезнее не для сканирования, а для тонкой подстройки прецизионных оптических систем.

3. ТЕХНИЧЕСКИЕ СРЕДСТВА МОДУЛЯЦИИ ЛАЗЕРНОГО ИЗЛУЧЕНИЯ

Далее для каждого из элементарных видов модуляции когерентного излучения перечислим пригодные для этого технические средства и физические эффекты. Упоминаемые далее эффекты Керра и Фарадея будем считать известными из курса общей физики. Все остальные эффекты либо кратко описаны по ходу изложения, либо описаны в соответствующих разделах данного сборника.

Начнем обзор с фазовой модуляции, так как в оптике именно фазовая модуляция наиболее важна, с одной стороны, и нуждается в минимуме технических средств, с другой стороны.

3.1. Фазовая и частотная модуляция

Фазовая модуляция представляет собой результат непосредственного воздействия единственного оптического элемента на световую волну, без иных операций и преобразований. Фазовые модуляторы входят в рассмотренные в следующем разделе амплитудные модуляторы как важнейшие и необходимые элементы. Рабочие характеристики быстродействующих амплитудных модуляторов зависят в первую очередь от качества и эффективности примененных в них фазовых модуляторов.

Укажем физические эффекты, чаще всего используемые в фазовых модуляторах.

Управление фазовой скоростью волны:

- 1. Линейный электрооптический эффект в кристаллах (эффект Поккельса).
- 2. Квадратичный электрооптический эффект (эффект Керра) в аморфных материалах, по преимуществу в жидкостях.
- 3. Фотоупругость изменение фазовой скорости света в твердом теле при его упругой деформации.
- 4. Электрооптические ориентационные эффекты в жидких кристаллах.
- 5. Эффект Фарадея в неферромагнитных материалах.
- 6. Электрооптические эффекты в полупроводниках (изменения концентрации электронов, эффект Франца-Келдыша и т.п.)
- 7. Фоторефракция изменение показателя преломления кристалла под действием света с длиной волны, меньшей в сравнении с модулируемым излучением.

Электрооптические эффекты дают максимальное быстродействие, в принципе ограниченное в кристаллах частотами оптических колебаний решетки (оптических фононов), а в полупроводниках – частотами плазменных колебаний подвижных зарядов. Чаще всего на практике используют эффект Поккельса.

Эффект Керра применяют при модуляции излучения высокой интенсивности, если модуляторы Поккельса использовать затруднительно по двум причинам:

- 1) лучевая стойкость нелинейных кристаллов невелика,
- 2) работу модулятора может нарушать фоторефрактивный эффект.

Ячейки Керра чаще всего применяют в составе амплитудных модуляторов (совместно с поляризаторами) для включения добротности внутри резонаторов лазеров и для выделения одиночного импульса из последовательности (цуга) импульсов лазерного излучения, то есть при работе с относительно мощными импульсами излучения. В ячейках Керра с жидкостями допустимы плотности потока излучения на порядки большие, нежели в электрооптических кристаллах, нет фоторефрактивных эффектов, исключены необратимые лучевые повреждения. Отметим, что в таких устройствах обычно применяют не пленочные поляризаторы, а более стойкие к излучению поляризационные призмы (предпочтительно так называемая лазерная призма Глана, или Глан-лазер).

При умеренных и малых пиковых мощностях излучения стараются использовать более удобные и эффективные модуляторы Поккельса.

Полуволновое напряжение ячейки Керра длиной в несколько сантиметров превышает 10 кВ, что на практике создает технические сложности. Так как задержка по фазе пропорциональна квадрату управляющего напряжения, его нужно изменять с большей скоростью, нежели расчетная скорость изменения задержки.

Часто к управляющему напряжению ячейки Керра добавляют постоянную составляющую (порядка 10 кВ); тогда переменная составляющая управляющего напряжения действует на более линейном и крутом участке квадратичной характеристики управления (сдвиг фаз/напряжение). При этом смягчаются требования к генератору управляющего напряжения, в частности, скорость изменения задержки по фазе мало отличается от скорости изменения управляющего напряжения. Электрооптические эффекты в полупроводниках до сих пор в практических целях применяли гораздо реже; тем не менее, в настоящее время именно это направление считают перспективным для интегральной оптоэлектроники, основанной на квантоворазмерных структурах и фотонных кристаллах. Кроме того, они уже нашли практическое применение в современной аппаратуре волоконно-оптической связи. Подробнее электрооптические эффекты в полупроводниках рассмотрены в приложении 2 («Модуляторы для ВОЛС»).

Упомянутые выше эффекты применяют для ФМ, которую мы ранее назвали узкополосной; средняя частота модулированного колебания в этом случае совпадает с исходной несущей частотой. Для сдвига оптической частоты применяют рассеяние света на ультразвуковой волне в прозрачной среде (см. раздел 5).

В измерительной и лабораторной технике ФМ/ЧМ нередко нужна не только как промежуточный этап AM, но и сама по себе.

Управление геометрической длиной оптической трассы. Для фазовой модуляции излучения вне лазера этот метод применяют крайне редко из-за невысокого быстродействия. Чаще всего системы с электромеханическим приводом зеркал используют для перестройки одночастотных лазеров; в этом случае ФМ выглядит как ЧМ.

Демодулятор ЧМ/ФМ (дискриминатор) по принципу действия может быть либо частотно-селективным фильтром, либо интерферометром, как показано на рис. 3.1. Это относится не только к оптике и СВЧ, но и к любой области радиодиапазона. При более внимательном взгляде на частотно-селективные фильтры можно заметить, что они, как правило, представляют собой либо явные или неявные аналоги двухлучевых интерферометров Майкельсона или Маха-Цендера, либо аналоги интерферометра Фабри–Перо (резонаторы).

Ранее было указано, что нарушение строгих соотношений между напряженностями боковых компонент, а также фазовых соотношений между ними и несущей при квадратичном детектировании дает сигнал на частоте модуляции, и напряжение его тем больше, чем сильнее разбалансированы напряженности и фазы. Отсюда следует общая идея детектирования ЧМ/ФМ колебаний резонансными методами: ослабление одной боковой полосы в сравнении с другой преобразует ЧМ/ФМ колебания в колебания с АМ. Для ЧМ возможно и иное, почти наглядное *представление* частотного детектирования: при помощи фильтра с шириной склона полосы пропускания несколько более девиации, характерной для принимаемого сигнала, ЧМ преобразуют в АМ, и далее используют технику детектирования АМ сигнала. В этом случае роль фильтра исполняет либо интерферометр, либо среда с поглощением, зависящим от частоты. В обоих случаях фильтрация идет в области спектра вблизи от резонанса, но в стороне от него.



Рис. 3.1. Частотные и фазовые дискриминаторы оптического диапазона: *а* – частотный дискриминатор с резонансом пропускания, *б* – частотно-фазовый дискриминатор с периодическими резонансами на шкале частот (волноводный интерферометр Маха-Цендера), *в* – частотные и фазовые характеристики пропускания дискриминатора типа (*б*) по интенсивности (сплошная линия) и по напряженности поля (пунктир), *г* – вариант дискриминатора типа (*б*) с квадратурными выходами через связанные волноводы, *д* – частотные и фазовые характеристики пропускания дискриминатора типа (*г*) по квадратурным выходам. На рис (*б*, *г*) соотношения размеров для наглядности искажены Другой способ преобразования ЧМ/ФМ в АМ использует нарушение фазовых соотношений при проходе излучения через среду с сильной дисперсией показателя преломления; этот метод по сути тоже относится к резонансным, так как сильная дисперсия возможна только в областях спектра вблизи от резонансов.

Чисто оптические методы преобразования ЧМ/ФМ в АМ можно успешно применять в ограниченном диапазоне девиаций, от единиц мегагерц до единиц гигагерц. При уменьшении девиации для сохранения эффективности преобразования нужно использовать резонаторы с очень высокой добротностью и очень точной настройкой. Увеличение девиации свыше десятков гигагерц затруднительно уже в самих модулирующих устройствах.

Для детектирования ЧМ/ФМ с малой девиацией применяют гетеродинное преобразование светового сигнала в электрический, и далее работают с электрическим сигналом одним из подходящих чисто радиотехнических методов, в той или иной форме сводящихся к аналогам резонансных; на практике встречаются и аналоги интерферометров с устройствами задержки, и резонансные контуры, и многоконтурные диспергирующие цепи. Более подробно эти вопросы изложены в курсах теории электрических цепей и сигналов, а также в теории радиоприемных устройств. Отметим, что при этом спектральная ширина несущей должна быть менее девиации (или по крайней мере одного порядка), иначе отношение сигнала к шуму может быть неприемлемо низким.

3.2. Амплитудная модуляция

Ранее, при описании параметров и свойств АМ излучения были противопоставлены модуляция напряженности поля и модуляция интенсивности. Модуляция напряженности поля – случай более сложный в сравнении с модуляцией интенсивности. Модуляция напряженности вообще требуется реже, нежели модуляция интенсивности, и для нее принципиально необходим интерферометр с фазовой модуляцией излучения внутри него. Все методы модуляции напряженности применимы и для модуляции интенсивности, но не всякая модуляция интенсивности дает модуляцию напряженности поля.

Механические и электромеханические устройства (затворы). Наиболее прямая и простая методика, пригодная только для модулирующих частот инфразвукового и звукового диапазона (до нескольких килогерц). Ее применяют часто, но почти исключительно как вспомогательное средство в измерительной и лабораторной технике. Самые примитивные амплитудные модуляторы – вращающиеся диски с отверстиями или прорезями, вибрирующие заслонки и т. п. – применяют для преобразования постоянного светового потока в периодический переменный.

Более высокое быстродействие (до ультразвукового диапазона) обеспечивают модуляторы на основе упомянутых ранее MEMC. Но для современных систем связи и иных применений необходимы модуляторы, работающие на частотах до нескольких гигагерц, что для электромеханики невозможно.

Перестройка линии резонансного поглощения за счет эффектов Зеемана или Штарка – методика, наименее удобная для технических систем. Ее применяют для преобразования изменений расстройки частоты от резонанса в амплитудную модуляцию в спектроскопии. Такое специфическое применение AM, по существу, относится к технике лабораторных исследований, а не к информационной технике.

Особо отметим квантово-размерный эффект Штарка в многослойных полупроводниковых наноструктурах. Физике и техническим применениям этого эффекта в волоконно-оптической связи посвящено приложение 2 («Модуляторы для ВОЛС»).

Нарушенное полное внутреннее отражение (НПВО) в оптических волноводах; в отличие от ранее упомянутых электрооптичсеких эффектов, проявляется только в структурированной среде, на границе материалов с разными показателями преломления. Если за счет электрооптического эффекта уменьшить показатель преломления сердцевины волновода и сделать его равным показателю преломления оболочки, то излучение беспрепятственно выходит из сердцевины (рис. 3.2).



Рис. 3.2. Полное внутреннее отражение на границе двух прозрачных сред (*a*) и нарушенное полное внутреннее отражение (*б*, *в*) при изменении разности показателей преломления

Особенно эффективной может быть модуляция интенсивности на основе НПВО в прямоугольных оптических волноводах; широ-

кое применение этого метода затруднено тем, что нужно использовать либо сочетание очень разнородных материалов, с близкими показателями преломления, но сильно различающимися электрооптическими коэффициентами, либо сложные структуры волноводов и управляющих электродов. В литературе описания современных систем такого рода почти не встречаются.

К методикам НПВО примыкают модуляторы с управляемой связью между параллельными волноводами. В них расстояние между волноводами порядка 1 мкм, и сформированы они у поверхности электрооптического кристалла (как в модуляторах Маха-Цендера на ниобате лития). Если фазовые скорости света в параллельных волноводах одинаковы, то коэффициент связи между волноводами максимален. Максимальный коэффициент связи задают скачки показателей преломления на границах волноводов и расстояние между ними, то есть эффекты НПВО. При заданном коэффициенте связи существует оптимальная длина (ее называют длиной когерентности), при которой бегущая волна из одного волновода практически полностью (за вычетом потерь на рассеяние) перетекает в другой волновод.

Если за счет электрооптического эффекта увеличить разность фазовых скоростей, то пространственный фазовый синхронизм волн, бегущих по волноводам, нарушится, и излучение из одного волновода будет перетекать в другой менее эффективно; коэффициент связи уменьшится. Математические оценки коэффициента разветвления мощности бегущей волны в связанных волноводах можно провести на основе модели связанных волн (она описана в приложении 3).

В литературе описаны *технические предложения* модуляторов – переключателей – на основе НПВО, управляемых довольно мощным излучением (~ 100 мВт). Здесь использованы *нелинейные* свойства легированных стекол. Под действием излучения с высокой напряженностью поля в сердцевине одномодового оптического волокна изменяется показатель преломления; далее эти изменения, создающие НПВО, можно использовать в нескольких технических вариантах. Изменения восприимчивости (и фазовой скорости) слабы в однородном материале сердцевины, но существенно сильнее в стекле с наноразмерными кристаллами, окруженными аморфной фазой (ситалле) и поляризованными за счет механической деформации (закручивания) волокна в процессе его вытягивания. Сообщения о реальных технических применениях этих методов и материалов пока неизвестны.

Перестройка линии пропускания интерферометра. Такая схема на первый взгляд кажется замысловатой, но именно к ней в том или ином виде чаще всего сводятся применяемые на практике быстродействующие амплитудные модуляторы. Более того, работу многих амплитудных модуляторов можно описать с единой точки зрения как действие двухлучевого интерферометра с управляемой задержкой в одном плече. Управляемая задержка (фазовая модуляция) может дать максимальное быстродействие модулятора, и физическая причина такого высокого потенциального быстродействия – отсутствие в системе макроскопических движущихся частей.

Схема Маха-Цендера: – деление волны пополам, регулируемая задержка (по существу ФМ) в одной из волн и последующее смешение этих волн. Интерференция преобразует ФМ в одной из волн в АМ. Эта принципиальная схема типична для быстродействующих модуляторов в системах оптической волоконной связи. Здесь используют не классическую схему Маха-Цендера с полупрозрачными зеркалами, а симметричные системы с разветвлениями планарных волноводов.

Модуляторы Поккельса – электрооптические модуляторы на монокристаллах с управляемым электрическим полем двулучепреломлением (эффект Поккельса) по существу представляют собой модификацию основной схемы Маха-Цендера. То же самое относится и к *модуляторам на фотоупругих материалах*, где двулучепреломление изменяется при деформации прозрачного оптического элемента. В обоих случаях волну расщепляют не в пространстве, а на две ортогонально-поляризованных волны. Управляющее воздействие изменяет разность фаз между парой волн, и далее за счет интерференции при смешении этих волн возникают изменения амплитуды суммарной волны. Чаще всего амплитудные модуляторы таких типов применяют в лабораторной и измерительной технике, где имеют дело с пучками излучения вне волноводов.

Модификацию той же схемы представляет собой *магнитоопти*ческий модулятор на основе эффекта Фарадея, где ортогональны между собой волны с круговой поляризацией разных знаков.

Модуляторы на основе нематических жидких кристаллов (НЖК) по принципу действия похожи на модуляторы Поккельса. У модуляторов на НЖК очень невысокое быстродействие (сотни герц), но их можно делать миниатюрными, и управляющие напряжения в них не превышают единиц вольт. Действительно массовое применение модуляторы на НЖК нашли в мониторах компьютеров и телевизорах; экран современного монитора (панель TFT-LCD) содержит более 3 млн. модуляторов (по три на пиксель), а его стоимость при размере диагонали экрана 19" не превышает \$150.

Между физическими эффектами в вешестве, применяемыми для АМ и ФМ в оптическом диапазоне, существует глубинная и принципиальная взаимосвязь. Одно и то же элементарное устройство например, фазовый модулятор, - в разных диапазонах длин волн может работать и как чистый фазовый модулятор (вдали от полос поглощения вещества), и как модулятор интенсивности (у края полосы поглощения). Все изменения показателя преломления вещества, в том числе и в окнах прозрачности, где поглощение пренебрежимо мало, в конечном итоге вызваны изменениями поглощения вблизи от окна прозрачности. На это указывают фундаментальные дисперсионные соотношения Крамерса-Кронига. Впрочем, из этого не следует, что эффективность АМ в полосе поглощения и эффективность ФМ в окне прозрачности у одного и того же устройства сравнимы и близки. Обычно при сравнимых мощностях электрического управления, соответствующих полуволновому напряжению ФМ, хорошие модуляторы на основе управляемого поглощения с коэффициентом модуляции около 20 дБ могут быть на порядок короче в сравнении с фазовыми модуляторами на тех же структурах.

Возможны и альтернативные схемы оптических модуляторов интенсивности – например, с применением управляемого динамического рассеяния света в жидких кристаллах или на ультразвуковых волнах в жидких и твердых прозрачных материалах (*акустооптическая модуляция*). Для современной техники связи наиболее важны быстродействующие модуляторы на основе эффекта Поккельса, квантово-размерного эффекта Штарка, а также управляемые излучением переключатели на основе эффектов нелинейной оптики. Акустооптическая модуляция по совокупности свойств и ее применениям занимает особое место в технике физических измерений. Этой теме в данном сборнике посвящены раздел 5 и лабораторная работа 2.

3.3. Управление поляризацией излучения

В простейших вариантах для изменения поляризации излучения можно использовать подвижные двулучепреломляющие пластины или клинья, но наибольшее быстродействие обеспечивают устройства с управляемым двулучепреломлением – фотоупругие, ячейки Поккельса, Керра и Фарадея. Наиболее высокое быстродействие дают модуляторы с управляемым двулучепреломлением кристаллов – ячейки Поккельса. Отметим, что поляризация – вполне адекватная характеристика как лазерного излучения, так и излучения с широким спектром, поэтому управление поляризацией равно применимо и излучению лазеров, и к белому свету.

Для управления поляризацией используют следующие физические эффекты:

- эффект Поккельса,
- электрооптические эффекты в жидких кристаллах,
- фотоупругость,
- эффект Фарадея в ферромагнетиках (в том числе в тонких ферромагнитных пленках).

В устройствах управления поляризацией, как и в большинстве упомянутых выше модуляторов, главную роль играет фазовая модуляция. Поляризация излучения изменяется при изменении разности фаз между двумя взаимно-ортогональными когерентными волнами. Вообще говоря, математически ортогональны две эллиптически поляризованные волны, у которых одинаковые по эксцентриситету эллипсы поляризации взаимно развернуты на 90°, а направления вращения вектора напряженности встречные.

На практике чаще всего пользуются разложением поляризованной волны в двух вариантах:

- разложение на взаимно ортогональные линейно поляризованные компоненты *E_x*, *E_y*; это представление удобнее для описания эффектов Керра, Поккельса, фотоупругости, нематических жидких кристаллов (ЖК);
- разложение на встречные циркулярно поляризованные компоненты, $E_x + iE_y$, $E_{x-}iE_y$; в таком представлении удобнее рассматривать эффекты Фарадея и холестерических ЖК.

При использовании управляемого двулучепреломления (эффекта Поккельса, электрооптических эффектов в нематических жидких кристаллах, фотоупругости) изменяется разность фаз между линейно поляризованными волнами, и в результате изменяется форма эллипса поляризации суммарной волны без изменения расположения его главных осей.

При использовании эффекта Фарадея изменяется разность фаз между волнами с круговой поляризацией, и в результате поворачивается эллипс поляризации суммарной волны без изменения его формы. В простейшем случае, когда падающее излучение поляризовано линейно, происходит *поворот* плоскости поляризации, без изменения степени эллиптичности поляризации.

Эффект Фарадея в тонких ферромагнитных пленках с цилиндрическими магнитными доменами имеет одну примечательную особенность. Поворот плоскости поляризации излучения, прошедшего через уединенный домен, происходит в условиях *насыщения* намагниченности ферромагнетика, поэтому плоскость поляризации прошедшего излучения фактически может принимать только два положения, слабо зависящие от напряженности внешнего магнитного поля.

При перемагничивании пленки происходит изменение соотношения объемов доменов с противоположными направлениями вектора намагниченности. В широком пучке света в дальней зоне при этом согласованно изменяются напряженности поля двух взаимнокогерентных световых волн. Сложение этих напряженностей в дальней зоне дает волну, плоскость поляризации которой поворачивается при перемагничивании пленки. Отметим, что при перемагничивании существует гистерезис, и поэтому угол поворота, вообще говоря, не следует в точности за изменениями напряженности внешнего перемагничивающего поля. Явления гистерезиса не причиняют заметных трудностей, если магнитооптическую ячейку применяют для переключения направлений поляризации света в дальней зоне.

Непрерывное (аналоговое) управление поворотом плоскости поляризации, напротив, страдает от гистерезиса в двух аспектах. При очень малой коэрцитивной силе в слабом внешнем поле пленка расслаивается на множество очень мелких цилиндрических доменов. Дополнительное рассеяние света на границах доменов (их еще называют стенками Блоха), суммарные периметры которых в этих условиях резко растут, в дальней зоне ослабляют и частично деполяризуют прошедшее излучение. В этом случае при переходе напряженности магнитного поля через ноль наблюдают небольшое ослабление прошедшего света и рост рассеянного света. Указанное рассеяние представляет собой одно из явлений дифракции.

При большой коэрцитивной силе эти эффекты могут быть пренебрежимо слабыми, так как площади сечения доменов сильно ограничены, но максимальный угол поворота поляризации несколько уменьшается, а гистерезис зависимости угла поворота от напряженности внешнего магнитного поля проявляется сильнее.

Некоторую особенность имеет изменение поляризации в так называемых твист-текстурах на основе нематических жидких кристаллов, но такие текстуры, очень широко применяемые в ЖК индикаторах, дисплеях и другой видеоаппаратуре, обычно используют не для непосредственного управления поляризацией, а для модуляции интенсивности. Твист-текстуры очень удобно применять для переключения поляризации между двумя взаимно ортогональными направлениями – поворота плоскости поляризации на 90°. Малое быстродействие ЖК-модуляторов на нематиках пока ограничивает их применение, но все более глубокие исследования ЖК-структур и кооперативных взаимодействий в них (сходных с сегнетоэлектриками) постепенно отодвигают эту границу.

Управление поляризацией можно отметить как неотьемлемую часть во многих амплитудных модуляторах, но в этих случаях изменение поляризации само по себе вряд ли верно рассматривать как первичный принцип действия модулятора. Точнее будет рассматривать изменение поляризации как альтернативное *представление* при описании амплитудного модулятора на основе управляемого двулучепреломления. Пример такого представления приведен в приложении «Электрооптический эффект Поккельса». Такое представление менее удобно для описания электрооптических амплитудных модуляторов, но более наглядно для описания *магнитооптических* модуляторов.

Во многих случаях объектом измерений служит свет с модулированной поляризацией, а измеряемой величиной – угол поворота плоскости поляризации. Известен способ увеличения малого угла поворота плоскости поляризации непосредственно перед анализатором. Для этого применяют «несовершенный» поляризатор, частично пропускающий линейно-поляризованное излучение в положении наименьшего пропускания (рис. 3.3).

Поляризатор ориентируют на наименьшее пропускание света исходной поляризации (вдоль оси *Y*, как это показано на рис. 3.3),

так что направление наилучшего пропускания A - A совпадает с направлением X. Если отношение коэффициентов наилучшего и наихудшего пропускания равно Q, то угол поворота плоскости поляризации после поляризатора $\phi \approx Q \times \phi$. Здесь угол поворота ϕ малый, но угол ϕ может быть и не малым, до $\pm \pi/3 \approx \pm 1$.



Рис. 3.3. Увеличение угла поворота плоскости поляризации при помощи несовершенного поляризатора; *А* – *А* – направление наилучшего пропускания

В описанной схеме E_y практически не изменяется при малых изменениях φ , тогда как изменения E_x повторяют изменения E_{1x} . Таким образом, можно отличить изменения сигнала, вызванные поворотом плоскости поляризации, от изменений мощности волны E_1 . Кроме этого, исключают перегрузку фотоприемника в анализаторе при большой мощности волны E_1 , а также уменьшают дробовой шум фотоприемника, связанный с большой постоянной составляющей его освещенности. Отметим, что здесь E_x – полезный сигнал, а E_y – опорный сигнал, сравнение с которым позволяет более точно отделять поворот плоскости поляризации от сопутствующих помех и искажений.

3.4. Изменение направления волнового вектора

На практике управление направлением волнового вектора (в обобщенном понимании – управление волновым фронтом оптического излучения) применяют для формирования и коррекции пучков лазерного излучения, а также в системах оптической обработки информации. Как было указано в разделе 2.5, эту задачу радикально решают решетки фазовых модуляторов. Можно применять (по крайней мере, в принципе) призмы с управляемым показателем преломления, но и этот вариант сводится к фазовой модуляции, только решетка модуляторов в таком случае не дискретная, а непрерывная. Следовательно, перечень пригодных для этого физических эффектов формально повторяет перечень из раздела 3.1.

На практике чаще всего применяют фотоупругий эффект, но в видоизмененной схеме: – решеткой фотоупругих фазовых модуляторов служит ультразвуковая волна в прозрачном материале. В такой системе (ее называют *акустооптическим* модулятором, AOM) сдвиг фаз оптических волн между элементами решетки регулируют, изменяя длину акустической волны (частоту ультразвука). Управление мощностью акустической волны позволяет регулировать мощность отклоненной световой волны. Такой же подход применяют и в интегральных оптических системах с планарными оптическими волноводами и поверхностными ультразвуковыми волнами.

Акустооптические пространственно-временные модуляторы света (ПВМС) в настоящее время наиболее успешно применяют для коммутации, спектрального анализа и параллельного ввода последовательных сигналов. Достижима скорость ввода до 0,5 Гбит/с на канал при числе каналов до 100 и высоком (до 10⁴) контрасте.

Решетку из параллельных акустооптических модуляторов можно использовать как пространственно-временной модулятор широкой плоской волны или набора узких параллельных пучков лазерного излучения. Существует несколько возможных технических путей создания пространственно-временных модуляторов света на основе акустооптического эффекта. Во-первых, при достаточной длине акустического материала в направлении распространения ультразвуковой волны область акустооптического взаимодействия может включать несколько различных дифракционных решеток цугов, последовательно сформированных одним пьезопреобразователем и переносимых бегущей волной по кристаллу. Во-вторых, можно реализовать решетку акустооптических модуляторов на общем акустооптическом кристалле при формировании на торце кристалла решетки пьезоэлектрических преобразователей, как показано на рис. 3.4. На базе первых двух методов в акустооптическом кристалле больших размеров может быть сформирован двумерный массив информации. В этом случае в каждом из независимо адресуемых каналов многоканального AOM (MAOM) возбуждают временную последовательность электрических импульсов. При достаточно малой длительности импульса считывающего света, проходящего через светозвукопровод, движущуюся последовательность акустооптических волн (цугов) можно рассматривать как неподвижную многоэлементную дифракционную решетку. Наконец, каждый элемент такой решетки может быть сформирован как на одной несущей частоте, так и мультиплексно. В такой суперпозиции каждой частоте в спектре ультразвука соответствует свой угол дифракции. Таким образом, в многоканальном многочастотном акустооптическом модуляторе (MMAOM) может быть «записан» трехмерный массив информации.



Рис. 3.4. Многоканальный акустооптический пространственно-временной модулятор света.

Уникальная особенность акустооптических ПВМС – естественное продвижение сигнала в канале.

Выпускаемые в настоящее время, акустооптические пространственно-временные модуляторы обеспечивают следующие параметры ввода:

- размерность вводимого массива до 100х1000;
- дифракционная эффективность до 10 процентов;
- время ввода кадра от единиц до десятков микросекунд;
- тактовые частоты ввода от единиц мегагерц до единиц гигагерц;
- организация ввода последовательно-параллельная;
- потребляемая мощность не более 1 Вт на канал.

В среднем ИК диапазоне перспективны *микроэлектромеханические* системы (МЕМС), на основе которых можно строить управляемые отражательные дифракционные решетки. Кроме того, решетку МЕМС-зеркал как дефлектор можно использовать и в нулевом порядке, просто как мозаичное зеркало с переменным углом наклона.

В этом же ряду упомянем пространственные фазовые модуляторы для адаптивных зеркальных систем. Здесь столь же применимы и перспективны MEMC.

В оптических системах невысокого быстродействия эффективны пространственные электрооптические модуляторы на жидких кристаллах.

Детектирование изменений волнового фронта. Управление волновым фронтом крайне редко применяют для дальней передачи сигналов (по крайней мере, в том примитивном виде, как в гелиографах XIX века). В таком простейшем случае, когда для формирования сигнала просто поворачивают диаграмму направленности отраженного света при помощи подвижного зеркала, сигнал на стороне приема восстанавливает квадратичный детектор (глаз). В более сложных вариантах изменение формы волнового фронта обнаруживают с помощью интегральных преобразований; например, преобразование Фурье переводит изменения волнового фронта в изменения распределения освещенности многоэлементной светоприемной структуры (сетчатки глаза, телевизионной камеры и т. п.). При таком подходе можно обнаруживать крупномасштабные изменения волнового фронта, но отыскать его участки с мелкомасштабными искажениями крайне трудно.

Для диагностики формы волнового фронта и выявления его локальных искажений в системах адаптивной оптики чаще всего применяют специальную телевизионную измерительную систему – анализатор Хартмана–Шака (*Hartmann–Shack wavefront sensor*); его схема показана на рис. 3.5, *a*. Он представляет собой матричный фотоприемник с двумерной решеткой раздельных датчиковячеек (рис. 3.5, δ), каждую из которых образует микролинза и пиксель матрицы ФПЗС, аналогичной фотоприемным матрицам современных телекамер и цифровых фотокамер. Микролинза фокусирует свет на свой пиксель только тогда, когда на нее падает плоский участок волнового фронта, ортогонального ее главной оптической оси. При наклоне данного участка фронта кружок размытия, в который фокусируется свет, смещается, и сигнал от данного пикселя уменьшается. Таким образом, если анализируемый волновой фронт плоский, то сигналы от всех ячеек, на которые падает свет, должны быть практически одинаковыми, и телевизионное изображение будет выглядеть как поле равномерной яркости. Искажения фронта выявляются как уменьшения яркости в отдельных участках изображения. Для обработки сигналов применяют ЭВМ со специализированными устройствами приема и обработки изображений (цифровые телевизионные системы). Типовое программное обеспечение позволяет анализировать искажения фронтов и профили лазерных пучков.



Рис. 3.5. Анализатор волнового фронта Хартмана–Шака: *a* – схема устройства; *б* – принцип действия одной ячейки; *в* – совместное действие ячеек одного ряда решетки; *г* – отображение фронтов с разными степенями искажений

3.5. Заключение

Приведенный обзор показывает, что основным принципом действия модуляторов света для большинства технических применений может быть фазовая модуляция; для когерентных оптических систем иные подходы просто непригодны. Поэтому поиск новых эффективных материалов и методов фазовой модуляции остается актуальной задачей прикладной и технической физики. Самостоятельное и важное значение имеют исследования модуляции излучения непосредственно при помощи световых сигналов, без их преобразования в электрические сигналы. Эта задача особенно актуальна для разработок высокоскоростных устройств связи, где возможности электронного формирования управляющих сигналов ограничены сверху частотами около 40 ГГц. Еще не до конца изучены возможности применения органических полимеров, молекулярных кристаллов и фрактальных олигомеров (дендримеров), более дешевых в производстве в сравнении с кристаллическими электрооптическими материалами. Основными проблемами на этом направлении будут нелинейные восприимчивости и лучевая стойкость материалов, миниатюризация и интеграция элементов за счет применения новых управляемых сред с квантоворазмерными структурами (наноструктур).

4. ЭЛЕКТРООПТИЧЕСКИЙ АМПЛИТУДНЫЙ МОДУЛЯТОР

4.1. Принципиальные схемы

Амплитудные электрооптические модуляторы (ЭОМ) служат для изменения интенсивности проходящего через них света под действием управляющего электрического напряжения.

В самом простом варианте ЭОМ содержит кристалл с изменяемым (управляемым) двулучепреломлением, помещенный между парой поляризаторов. Эти модуляторы обычно применяют при работе с лазерным излучением, которое чаще всего линейно поляризовано, и входной поляризатор не обязателен (рис. 4.1).



Рис. 4.1. Изменение разности фаз между обыкновенной и необыкновенной волнами на выходе из двулучепреломляющего кристалла

Оптическая ось кристалла направлена вдоль оси *Y*, свет распространяется вдоль оси *Z*. Плоскость поляризации падающего света составляет угол $\alpha \neq \pi n$ с оптической осью среды и угол ϕ с осью пропускания поляризатора A – A. В общем случае при такой геометрии в кристалле распространяются две линейно поляризованных волны с взаимно ортогональными плоскостями поляризации и разными фазовыми скоростями. Различие фазовых скоростей описывают n_x и n_y – показатели преломления среды для направлений вдоль осей *X* и *Y*, соответственно. Направления *X*, *Y* и *Z* взаимно ортогональны и жестко привязаны к определенным направлениям в решетке кристалла. Если кристалл одноосный (что на практике бывает чаще всего), то направление *Y* соответствует вектору поляризации обыкновенной волны, и $n_y \equiv n_o$, а направление *X* соответствует вектору ет вектору поляризации необыкновенной волны, и $n_x \equiv n_e$.

Во многих кристаллах показатели преломления n_0 , n_e изменяются при наложении внешнего электрического поля. Об этом подробнее сказано в приложении 1. На данном этапе для уяснения принципа работы модулятора достаточно только принять во внимание, что n_x и n_y – переменные, зависящие от внешнего электрического поля.

На практике выбирают угол $\alpha = 45^{\circ}$, чтобы две волны имели одинаковые амплитуды; угол ϕ выбирают равным 90° или 0°. Только при таких условиях могут быть максимальными изменения мощности прошедшего через поляризатор света (глубина модуляции). Это утверждение обосновано в приложении 1.1, где приведены общие выражения для произвольных углов α , ϕ ; далее рассмотрен только модулятор с оптимальной настройкой.

Напряженности полей двух частичных линейно поляризованных волн, прошедших через поляризатор, будут иметь вид

 $E_x \propto E_0 \cos(\omega t - 2\pi L n_x/\lambda),$ $E_y \propto E_0 \cos(\omega t - 2\pi L n_y/\lambda).$

В этом представлении напряженность поля волны, прошедшей через поляризатор, при $\phi = 0$,

 $E \propto E_x + E_y \propto E_0 \cos(\phi/2) \cos[\omega t - (\pi L/\lambda)(n_x + n_y)],$ (4.1a) а при $\phi = \pi/2$

 $E \propto E_x - E_y \propto E_0 \sin(\phi/2) \cos[\omega t - (\pi L/\lambda)(n_x + n_y)],$ (4.16) где $\phi = (2\pi L/\lambda)(n_x - n_y)$ – разность фаз, пропорциональная длине кристалла *L* при $(n_x - n_y) \neq 0.$

Фактически на выходе модулятора, *после поляризатора*, происходит интерференция двух *взаимно когерентных* волн, и напряженность поля зависит от разности фаз между ними. Выписанные формулы отображают *первичное* свойство модулятора – изменения *напряженности* поля прошедшей через него волны при изменениях $\Delta n \equiv (n_x - n_y)$.

Таким образом, напряженностью прошедшей через указанную оптическую систему волны можно управлять, изменяя разность фаз

между двумя ортогонально поляризованными взаимно когерентными волнами. В описанном ЭОМ для этого служит изменение показателей преломления n_x и n_y , то есть фазовых скоростей указанных волн под действием внешнего электрического поля E.

Изменения $\Delta n \equiv (n_x - n_y)$ под действием электрического поля происходят во всех диэлектрических материалах, но в большинстве случаев Δn зависит от напряженности поля E крайне слабо и квадратично. Гораздо более сильной и (что особо важно) *линейной* зависимость $\Delta n = a + bE$ бывает в таких кристаллах, где у элементарной ячейки кристаллической решетки *нет центра инверсии*. В общем случае напряженность поля E – вектор, а b – тензор.

Зависимость $\Delta n(E)$ бывает линейной и в кристаллах без двойного лучепреломления; тогда $\Delta n = bE$.

Подчеркнем, что $\Delta n \neq 0$ – условие недостаточное; для модуляции необходимо $dn_x/dE \neq dn_y/dE$, то есть не просто двулучепреломление, и не только зависимость $n_i(E)$.

Зависимость $n = n_0 + bE$ называют линейным электрооптическим эффектом, или эффектом Поккельса. Ряд кристаллов проявляет заметный электрооптический эффект, а технология выращивания таких кристаллов хорошо отработана. В зависимости от типа симметрии и других свойств кристалла, от ориентации его кристаллографических осей по отношению к волновому вектору света k, вектор напряженности управляющего электрического поля Е может быть либо параллелен, либо ортогонален вектору k. Электрооптический эффект называют продольным при E || k, и поперечным при $E \perp k$. Подробнее об эффекте Поккельса и его разновидностях в кристаллах разных классов симметрии сказано в приложении 1. достаточно В текущем только разделе принять зависимость $n = n_0 + bE$ как опытный факт и учесть типичный порядок величины параметра $b_{1} \sim 10^{-9}$ м/В.

Модуляцию интенсивности прошедшего света можно представлять как следствие модуляции напряженности. Интенсивность света, прошедшего через поляризатор,

$$W = W_0 \cos^2(\varphi/2) = W_0(1 - \sin\varphi), \ \phi = 0, \tag{4.2a}$$

$$W = W_0 \sin^2(\phi/2) = W_0 (1 - \cos\phi), \ \phi = \pi/2 \tag{4.26}$$

Таким образом, интенсивностью прошедшего через указанную оптическую систему света можно управлять, изменяя разность фаз между двумя ортогонально поляризованными волнами. Установив $\phi = \pi/2$, получим на выходе модулятора волну с напряженностью

 $E \propto \sin(\varphi/2) \approx \varphi/2$, если $|\varphi| < 1$.

Этот результат означает, что при такой настройке, при малых разностях фаз $|\phi|$ и при условии, что ϕ *линейно* зависит от управляющего воздействия (например, приложенного к кристаллу электрического напряжения) модулятор работает как *перемножитель* напряженности поля падающей волны и управляющего сигнала, исполняя функцию *балансного модулятора*, упомянутого во вводном обзоре.

Обратим внимание на фазу напряженности поля *E* на выходе,

 $[\omega t - (\pi L/\lambda)(n_x + n_y)]:$

ее часть, не зависящая от времени явно, изменяется при изменениях величины $(n_x + n_y)$, то есть изменения напряженности поля *E* в общем случае сопровождаются и изменениями ее фазы. В некоторых случаях сдвиги «огромной и быстрой» фазы не играют заметной роли; они равносильны изменениям длины оптической трассы на доли длины волны. В других случаях, особенно когда нужно наблюдать автокорреляцию модулированной волны в разные моменты времени, такая фазовая модуляция существенно влияет на результат. Отметим, что в современных системах оптоволоконной связи корреляцию оптических сигналов, смежных во времени, применяют для выделения кодированной информации, и поэтому возможные изменения фазы при модуляции нужно по крайней мере принимать в расчет. Отсюда же следует и другой вывод: идеальным по фазовым характеристикам можно признать модулятор напряженности, в котором приращение $\Delta(n_x + n_y) = 0$, то есть изменения n_x и n_y направлены в разные стороны и уравновешены (рис. 4.2, *a*).

Похожие свойства имеет и хорошо известный интерферометр Маха-Цендера (ИМЦ), пропускание которого изменяется при изменении разности фаз φ между двумя плечами (рис. 4.3, *a*, *б*). В ИМЦ интерферируют две волны с одинаковой поляризацией, полученные при расщеплении исходной волны на входе. Если в одном или обоих плечах ИМЦ действуют устройства управляемой задержки, то получается амплитудный модулятор Маха-Цендера (ММЦ). Классический ИМЦ обычно применяют не как модулятор в собственном смысле, а как средство измерения разно-

сти фаз в его плечах. Наибольшее распространение получили ММЦ в волноводном исполнении, применяемые для формирования сигналов в волоконных линиях связи (рис. 4.3, *в*; подробнее см. приложение 2, *«Модуляторы для ВОЛС»*).



Рис. 4.2. Изменение комплексной напряженности суммы двух взаимно когерентных волн $E_{xi} + E_{yi}$ равной интенсивности при изменении задержек этих волн по фазе на выходе из ЭОМ: a - приращения фаз $\arg(E_{x2}) - \arg(E_{x1}) = \arg(E_{y1}) - \arg(E_{y2});$ изменяется модуль напряженности, без изменения фазы $(\arg(E_1) = \arg(E_2));$ δ – неравные приращения фаз, $\arg(E_{x2}) - \arg(E_{x1}) \neq \arg(E_{y1}) - \arg(E_{y2});$ изменяются модуль и фаза суммарной напряженности $(\arg(E_1) \neq \arg(E_2))$

ММЦ обычно работают в линейно поляризованном свете (или на одной моде в волноводном исполнении), и управляемая разность фаз возникает между двумя волнами, разделенными в пространстве. Если задержка по фазе зависит от управляющего напряжения линейно, то возможен как *несимметричный* (однофазный) вариант, с управлением в одном плече, так и *антисимметричный*, с противофазным (*парафазным*) управлением в обоих плечах.

При полной симметрии структуры ММЦ она почти аналогична модулятору на двупреломляющем кристалле, с одним изменением. Так как ММЦ не содержит поляризаторов и имеет два выхода, то при разности фаз $\varphi = \pi n$ на одном выходе будет максимум пропускания, а на другом – минимум, и формулу, описывающую пропускание, придется выбирать альтернативно, в зависимости от используемого выхода. Нужно учесть, что при отражении волны от среды с меньшим показателем преломления происходит скачок фазы напряженности поля E_2 на π (потеря полуволны), как показано на рис. 4.3, δ .



Рис. 4.3. Интерферометр Маха-Цендера (*a*) и электрооптический модулятор Маха-Цендера (*б*), схема модулятора Маха-Цендера на микроволноводах в электрооптическом кристалле (*в*), сложение напряженностей интерферирующих волн в комплексном представлении (*г*)

Отметим, что в литературе термин ММЦ обычно применяют только к волноводному варианту; у него выход единственный, и он аналогичен ЭОМ на двупреломляющем кристалле при ориентации поляризатора под углом $\phi = 0$. В волноводном варианте ММЦ из-за симметрии *номинально* начальная разность фаз $\phi = 0$, но фактически $\phi \neq 0$, так как вряд ли можно выдержать симметрию структуры с точностью до нанометров. Поэтому в ММЦ несимметрию геометрии компенсируют управляющим напряжением (*смещением*), как это описано в разделе «*Рабочие характеристики амплитудных* электрооптических модуляторов».

В отличие от ММЦ-прототипа (рис. 4.3, б), где на выходном полупрозрачном зеркале световая волна расщепляется на две компоненты, и энергия света не теряется, а перераспределяется между двумя выходами, в ЭОМ на кристалле возможны два варианта:

- расщепление волны на поляризационной призме, при сохранении полной энергии света, перераспределяемой между двумя выходами, как и в ММЦ;
- 2) поглощение части света в дихроичном поляризаторе.

Второй вариант наименее пригоден как для самых слабых световых потоков (потери в поляроиде ухудшают отношение сигнал/шум), так и для очень интенсивных (поглощающий поляроид перегревается и повреждается).

Входящее в состав ММЦ устройство управляемой задержки по сути представляет собой *фазовый модулятор* света.

Амплитудный ЭОМ по принципу работы эквивалентен ММЦ с фазовыми модуляторами в обоих плечах (рис. 4.3, e), управляемыми одним и тем же напряжением, но с различающимися характеристиками ($dn_x/dE \neq dn_y/dE$). Изменение полной разности фаз между интерферирующими волнами на выходе равно разности изменений фаз в каждом плече.

Так как фазовая модуляция – неотъемлемая черта принципа работы электрооптического модулятора, далее рассмотрим основные характеристики фазового модулятора с линейным электрооптическим эффектом.

4.2. Электрооптическая фазовая модуляция света

Для фазовой модуляции излучения нужен материал с управляемыми коэффициентами преломления. Самый удобный вариант изменение n_x и/или n_y под действием внешнего электрического поля *E*.

Полагая, что в выбранном материале показатель преломления зависит от напряженности внешнего электрического поля E, получим зависимость фазы φ прошедшего через фазовый модулятор света от управляющего напряжения U. Не теряя общность, рассмотрим изменение фазы только одной волны, обыкновенной или необыкновенной, в кристалле длиной L, причем будем рассматривать только задержку по фазе $kL = (2\pi/\lambda)Ln$, опуская временную часть полной фазы.

При линейном электрооптическом эффекте для волны с выбранной поляризацией

$$n = n_0 + bE,$$

$$\varphi(E) = (2\pi/\lambda)Ln = (2\pi/\lambda)L(n_0 + bE) = \varphi_0 + \Delta\varphi(E),$$

где $\Delta \phi \propto (2\pi/\lambda)E$, $\phi_0 = (2\pi/\lambda)Ln_0$. Здесь явно показана зависимость ϕ_0 и $\Delta \phi$ от λ , чтобы напомнить, как на работу фазового модулятора влияет немонохроматичность управляемого им излучения.

Если на кристалл действует переменное поле напряженности $E = E_0 \sin\Omega t$, то фаза $\varphi(t) = \varphi_0 + \beta \sin\Omega t$, где $\beta = (2\pi Lb/\lambda)E_0$.

Величина β представляет собой индекс фазовой модуляции; ранее определенный в разделе 2. Индекс модуляции, зависящий от амплитуды управляющего напряжения, характеризует режим модуляции, но не свойства самого модулятора.

Наглядной и удобной характеристикой самого электрооптического фазового модулятора как технического устройства служит напряжение $U_{\lambda/2}$, при котором сдвиг фаз $\varphi(U_{\lambda/2}) = \pi$; его называют *полуволновым напряжением*. В литературе чаще встречается другое обозначение, $U_{\pi} \equiv U_{\lambda/2}$; его мы будем использовать далее.

При линейном электрооптическом эффекте очевидно простое соотношение между сдвигом фазы и управляющим напряжением U:

$$\varphi(U) = \pi U/U_{\pi}$$

Получим выражение полуволнового напряжения через оптические и геометрические параметры кристалла фазового модулятора.

По определению $\phi(U_{\pi}) = \pi$; следовательно,

$$(2\pi/\lambda)LbE = (2\pi/\lambda)LbU_{\pi}/h = \pi,$$

где *В* – толщина кристалла в направлении вектора напряженности управляющего электрического поля *E*.

При поперечном эффекте Поккельса В и L независимы,

$$U_{\pi} = B\lambda/(2bL).$$

При продольном эффекте Поккельса $B \equiv L$, и тогда

$$U_{\pi} = \lambda/(2b).$$

Применять модулятор тем удобнее, чем ниже его полуволновое напряжение. Заметим, что полуволновое напряжение модулятора на продольном эффекте явно не зависит от его геометрии (в видимом и ближнем ИК диапазонах $U_{\pi} \sim 10$ кВ), а у модуляторов на поперечном эффекте в микроволноводном исполнении формфактор h/L может быть менее 10^{-3} , то есть U_{π} может быть порядка единиц вольт, что хорошо согласуется с современной микроэлектронной схемотехникой.

Таким образом, элементарный фазовый модулятор можно полностью описать двумя параметрами – напряжением U_{π} и начальной

(статической) задержкой. Режим ФМ отображают либо индексом модуляции, либо отношением напряжений – управляющего и полуволнового (U/U_{π}). Начальную задержку и ее возможные вариации, не влияющие на режим собственно ФМ, приходится учитывать только в редких случаях; эти вариации равносильны малым изменениям длины трассы от источника до приемника света и выявляются только при наблюдениях интерференции.

Непосредственному применению фазовых модуляторов посвящен один из дальнейших разделов; здесь же уместно вернуться к амплитудной электрооптической модуляции, применив к ее описанию представления о фазовой модуляции и введенный параметр модулятора – полуволновое напряжение U_{π} .

4.3. Рабочие характеристики амплитудных электрооптических модуляторов

В данном разделе будут рассмотрены зависимости пропускания ЭОМ и параметров модуляции от начальной разности фаз и управляющего напряжения. Сказанное далее применимо и к волноводным ММЦ.

Вначале рассмотрим режим модуляции с $\phi_0 = 0$ и $\phi = \pi/2$. Иначе говоря, на выходе кристалла ЭОМ колебания обыкновенной и необыкновенной волн строго синфазные, но при этом на выходе поляризатора *в пропущенных им частичных волнах* колебания напряженности поля противофазные (см. рис. 4.1).

Зависимость *интенсивности* света на выходе амплитудного ЭОМ от времени при $\phi_0 = 0$ описывает выражение

 $W(t) = W_0 \sin^2(\beta \sin \Omega t/2) = (W_0/2)[1 - J_0(\beta) - 2\Sigma J_{2k}(\beta) \cos 2k\Omega t],$ где $J_k(\beta)$ – функция Бесселя порядка k, а суммирование идет от k = 1 до ∞ . В этом случае коэффициент АМ (глубина модуляции) $m \equiv (W_{max} - W_{min})/(W_{max} + W_{min})$

при любом $U_m > 0$ формально равен 100%, а от U_m зависит не коэффициент *m*, а W_{max} . Зависимость W(t) близка к синусоидальной с частотой 2Ω , но при $U_m > U_{\pi}/10$ заметный вклад в зависимость W(t)дают слагаемые с частотами 4Ω , 6Ω , и т.д. Функции Бесселя $J_k(\beta)$ немонотонны, поэтому при увеличении индекса модуляции содержание четных гармоник частоты модуляции в оптическом *сигнале* (в переменной составляющей W(t)) изменяется не самым очевидным образом. Часто режим AM характеризуют его эффективностью, определенной как отношение переменной составляющей интенсивности света на выходе к интенсивности падающего света, $\eta = W_{-}/W_{0}$. Эта величина количественно выражает долю мощности излучения, преобразованной в переменную составляющую, в заданной оптической системе и при заданном индексе модуляции β . Чаще всего имеют в виду $\beta \sim 1$ (это называют режимом большого сигнала).

При использовании продольного эффекта Поккельса

$$\eta = \sin^2(\beta/2)$$

Эффективность зависит, очевидно, только от индекса модуляции β и начальной разности фаз φ_0 ; она не зависит явно от параметров электрооптического кристалла, и ее нельзя смешивать с отношением $\Delta W/\Delta U$ при больших приращениях управляющего напряжения. Эффективность $\eta(\beta)$ немонотонна и проходит через максимум при оптимальном индексе модуляции.

Режим максимальной эффективности на практике чаще всего используют для импульсного управления излучением, когда в отсутствие управляющего напряжения на выходе модулятора мощность света должна быть как можно меньшей, а при включении напряжения – максимальной. В частности, такой режим управления применяют в системах передачи цифровой информации; во вводном обзоре он был упомянут как *амплитудная манипуляция*.

При поперечном эффекте Поккельса естественное двулучепреломление (слагаемое φ_0) может переместить рабочую точку на более крутой и линейный участок (рис. 4.4, δ). На этом участке получают максимальный коэффициент преобразования переменного управляющего напряжения в переменную составляющую интенсивности выходного излучения, dW/dU. Смещение оптимально при $\varphi_0 = \pi/2$. Такое смещение при надобности создают при помощи регулируемой постоянной добавки к управляющему напряжению.

Смещение можно создать и в тех модуляторах, где вклад естественного двулучепреломления отсутствует. Для этого либо к кристаллу дополнительно прикладывают постоянную составляющую $U_{\pi}/2$, либо на входе устанавливают четвертьволновую пластинку, вносящую сдвиг фаз $\varphi_0 = \pi/2$. В ММЦ, разумеется, используют только корректирующее напряжение.

При оптимальном смещении эффективность $\eta = \sin\beta$. Зависимость *интенсивности* прошедшего через кристалл излучения от времени на линейном участке можно представить разложением в спектр Фурье:

 $W(t) = W_0 \sin(\beta \sin \Omega t) = (W_0/2) [1 + \Sigma J_{2k+1}(\beta) \sin(2k+1)\Omega t].$

Как видно из этого выражения, в спектре переменной составляющей интенсивности присутствуют только нечетные гармоники частоты модуляции.



Рис. 4.4. Зависимости относительной интенсивности света на выходе амплитудного модулятора от разности фаз и управляющего напряжения без смещения (*a*) и со смещением (*б*) рабочей точки

При небольших амплитудах модулирующего напряжения $(U_m < U_\pi)$, изменения интенсивности света и модулирующего напряжения практически пропорциональны. В этом случае глубина модуляции и эффективность равны и определяются амплитудой модулирующего сигнала, $m = \eta = \sin\beta$. В частотном представлении этому режиму соответствует индекс модуляции $\beta << 1$ и, соответственно, $J_k(\beta) << 1$ при |k| > 1, и $J_1(\beta) \approx \beta/2$.

Выбор схемы электрооптического модулятора зависит от множества различных предъявляемых к нему требований. Например, в модуляторе на продольном эффекте полуволновое напряжение не зависит от размеров кристалла, поэтому световой диаметр модулятора в принципе не ограничен. Если изготовить его в виде тонкой широкой пластины, то однородность поля в нем будет очень хорошей, но межэлектродная емкость может оказаться неприемлемо большой, и для ее быстрой перезарядки потребуется значительная электрическая реактивная мощность. В модуляторе на поперечном эффекте можно снизить полуволновое напряжение, взяв тонкий и длинный кристалл; при этом будет ограничен световой диаметр модулятора. Отметим, что *угловая* апертура (допустимая расходимость излучения) при любой схеме модулятора на эффекте Поккельса ограничена. Это ограничение частично преодолевают в микроволноводной конструкции, где достаточно согласовать числовую апертуру волновода модулятора с апертурами прочих оптических волноводов (они редко превышают 0,2).

4.4. Компенсация статических и динамических фазовых расстроек в ЭОМ

Наличие начальной разности фаз φ_0 из-за естественного двулучепреломления приводит к существенному дрейфу рабочей точки модулятора при изменении температуры, из-за изменений разности показателей преломления ($n_x - n_y$). При изменении температуры всего на 0,01° в кристалле ADP длиной 10 см изменение разности фаз $\Delta \varphi_0 = 23^\circ$. Поэтому модуляторам на поперечном эффекте Поккельса требуется либо термостатирование, либо компенсация двулучепреломления до $\varphi_0 < 1$. Компенсацию двулучепреломления обычно называют термокомпенсацией, но на деле этот технический прием сводится к компенсации начальной разности фаз φ_0 . Для практики предпочтительна термокомпенсация, т.к. при этом не приходится существенно усложнять модулятор и увеличивать его габариты.



Рис. 4.5. Схема компенсированного электрооптического модулятора с поперечным эффектом Поккельса

В термокомпенсированном модуляторе обычно используют систему из двух идентичных кристаллов, развернутых на 90° по отношению друг к другу, при согласованных управляющих напряжениях на обеих секциях (рис. 4.5). Волна, идущая в одном кристалле как обыкновенная, в другом распространяется как необыкновенная. При этом суммируются эффекты управления и компенсируется влияние естественного двулучепреломления. Соответствующие несложные расчеты приведены в приложении 1 («Электрооптический эффект Поккельса»).

В компенсированных модуляторах зависимость ϕ_0 от температуры определяют уже не свойства материала, а погрешности изготовления, настройки и неравномерности распределения температуры по конструкции прибора. В компенсированном модуляторе начальная разность фаз ϕ_0 не компенсируется до нуля; два элемента фактически могут различаться по длине на доли микрона, и тогда $\phi_0 \sim 1$. Таким образом, термостабильность рабочей точки становится удовлетворительной, но для точной настройки на нее нужно приложить корректирующее напряжение. Иногда кристаллы ориентируют одинаково, а между ними устанавливают полуволновую пластинку.

4.5. Электрооптическая модуляция немонохроматического излучения

До сих пор речь шла о модуляции строго монохроматического излучения. Если же свет имеет широкий спектр оптических частот, то максимальная достижимая эффективность электрооптической модуляции в общем ухудшается. Все предыдущие выводы основаны на том, что две волны, на которые в кристалле расщеплена исходная волна излучения, полностью сохраняют взаимную когерентность на выходе. На самом деле между этими двумя волнами возникает задержка $L(n_x-n_y)$. В таком случае, чтобы работа модулятора не нарушилась, эта задержка должна быть существенно меньше длины когерентности исходной волны.

В хорошо скомпенсированном двухэлементном модуляторе (рис. 4.5) задержку можно довести до пренебрежимо малой управляемой задержки $(\Delta L)(n_x - n_y) < \lambda/2$, и тогда требования к длине когерентности исходного излучения смягчаются до минимума.

Попутно отметим, что в электрооптических модуляторах на жидких кристаллах, работающих по преимуществу в белом свете, нет никакой компенсации, но здесь начальная разность фаз не превышает π , а ее вариации еще на порядок меньше, поэтому в компенсации нет необходимости.

Другое полезное свойство компенсированного амплитудного ЭОМ – уменьшение нежелательной *фазовой* модуляции, сопровождающей AM в некомпенсированном модуляторе. Паразитная ФМ дополнительно расширяет спектр модулированного излучения, что крайне нежелательно в технике оптической связи в системах с частотным уплотнением каналов (WDM, DWDM).

Компенсация φ_0 уменьшает искажения спектра излучения, проходящего через ЭОМ. Компенсация ФМ основана на том, что в двух интерферирующих волнах вариации фазы имеют разные знаки, поэтому в суммарной волне, складывающейся из двух волн одинаковой амплитуды, вариация фазы отсутствует. Но не всякий компенсированный ЭОМ имеет такие свойства; нужно, чтобы ФМ в обоих плечах была правильно согласованной, то есть нужно исполнить условия $\varphi_0 << 1$, равенство (по модулю) индексов модуляции в обоих плечах при минимальном отличии *сдвига фаз между* управляющими напряжениями двух плеч от π . Этого легче всего добиться в микроволноводных ММЦ.

Компенсация в принципе не нужна в ЭОМ на продольном эффекте Поккельса в КDP, но применение таких ЭОМ ограничено изза большого (~ 9 кВ) полуволнового напряжения.

Очевидный фактор, ухудшающий максимальную достижимую эффективность электрооптического модулятора в немонохроматическом свете – обратно пропорциональная зависимость индекса модуляции от длины волны излучения *при фиксированной напря*женности управляющего поля. Так, для синего света индекс модуляции при прочих равных условиях вдвое больше, чем для ближнего ИК излучения. Из-за этого для каждой узкой спектральной области существует свой оптимум управляющего напряжения, и, кроме всего прочего, модуляция *динамически* изменяет спектральный состав прошедшего "белого" света в сравнении с исходным (например, модуляция может быть сильной в желто-зеленом свете, но слабой на фиолетовом и красном краях).

4.6. Амплитудно-частотные характеристики ЭОМ

Важнейшее свойство эффекта Поккельса – малая инерционность. В электрооптическом модуляторе Поккельса нет макроскопических движущихся частей; и это позволяет модулировать свет электрическими сигналами с частотами до десятков гигагерц. Теоретический предел частот модуляции, по-видимому, находится в дальней ИК области; практический предел – около 40 ГГц, где заканчиваются технические возможности радиоэлектронных систем. Фактическую амплитудно-частотную характеристику (АЧХ) модулятора определяют в первую очередь параметры электрических цепей управления и конструктивные особенности модулятора. Модуляторы в виде кристалла, помещенного между обкладками конденсатора электрической цепи с сосредоточенными параметрами (например, колебательного контура), могут модулировать свет на частотах до 200 МГц. Для работы на СВЧ в узкой полосе модулирующих частот кристалл помещают в объемный резонатор, а для модуляции в более широкой полосе – в СВЧ волновод. Наиболее широкую полосу частот модуляции, вплоть до субнаносекундных импульсов, получают в модуляторах на тонких монокристаллических пленках, где управляющее поле создает полосковая линия с бегущей волной напряжения. Подробнее об этом сказано в приложении 2.

На АЧХ в определенных диапазонах частот может влиять вторичный электрооптический эффект, вызванный деформациями кристалла за счет обратного пьезоэффекта. Это влияние наиболее сильно вблизи от частот электромеханических резонансов.

Кристаллы с линейными электрооптическими эффектами применяют не только в амплитудных модуляторах. Не менее важны быстродействующие *фазовые* модуляторы, для которых используют почти исключительно электрооптические кристаллы. Основные области применения фазовых модуляторов – техника связи и измерительная техника.

Очень интересны специфические применения амплитудных и фазовых модуляторов для улучшения временной когерентности лазерного излучения. Можно показать, что при некоторых, уже достаточно высоких показателях когерентности исходного излучения, флуктуации амплитуды и фазы излучения, прошедшего через соответствующим образом управляемый модулятор, удается подавить до столь низкого уровня, что полученные таким образом световые колебания станут приближаться к идеальным гармоническим колебаниям. В подобной системе фазовый модулятор корректирует (компенсирует колебаний поля флуктуации фазу фазы), а амплитудный компенсирует шумы мощности. Основные технические трудности здесь создают собственные шумы фотоприемных устройств амплитудных и фазовых компараторов, а также ограниченная полоса пропускания электронного тракта управления.

5. АКУСТООПТИЧЕСКАЯ МОДУЛЯЦИЯ СВЕТА

5.1. Фотоупругость

Изменение показателя преломления при механической деформации упругой прозрачной среды называют фотоупругостью.

В простейшем случае под действием одноосной механической нагрузки изотропная среда становится одноосной двулучепреломляющей, и направление оптической оси совпадает с направлением деформации. В такой среде при распространении света ортогонально к наведенной оптической оси на длине *L* можно получить разность фаз между обыкновенной и необыкновенной волнами

 $\varphi = (2\pi L/\lambda)(n_e - n_o) = \pi u n^3 p_{44}/\lambda \qquad (5.1)$

где u – деформация, p_{44} – упругооптический коэффициент. Фотоупругость в общем случае отображает тензор упругооптических коэффициентов p_{ii} , подробнее описанный в приложении 3.

Если упругое тело имеет форму прямоугольного параллелепипеда и подвергнуто одноосному равномерному напряжению, то

$$u = \sigma(1 + K_P)/M_Y, \tag{5.2}$$

где K_P – коэффициент Пуассона, M_Y – модуль Юнга. Пример результата расчета разности фаз φ вместе с исходными данными приведен в табл. 5.1.

Таблица 5.1

Параметр	Значение	Параметр	Значение
ребро	1 см	$\lambda =$	0,63 мкм
п	1,46	σ	1 кГс/см ²
K_P	0,17	M_Y	7,3 ⁻ 10 ¹¹ дин/см ²
p_{44}	$-7,5 \times 10^{-2}$	φ	10'

Фотоупругие свойства плавленого кварца (кубический образец)

В видимом диапазоне значительный фотоупругий эффект проявляют органические полимеры – фотоэластики, а в ИК диапазоне – халькогенидные стекла. На основе таких материалов можно сделать достаточно эффективные модуляторы. Они действуют так же, как и электрооптические, отличаются же тем, что двулучепреломление возникает и изменяется под действием механического напряжения. Их общий недостаток – низкое быстродействие, ограниченное малой скоростью звука в упругих средах.

5.2. Когерентное рассеяние света на акустических волнах

Явление фотоупругости лежит в основе методов модуляции света, основанных на когерентном рассеянии световой волны на акустической (ультразвуковой) волне как на периодической структуре, аналогичной дифракционной решетке. Периодическое изменение упругой деформации создает акустическую волну, котораясоздает изменение (модуляцию) показателя преломления в пространстве и во времени. Свет, проходящий через такую среду, претерпевает дифракцию на *фазовой* решетке, образованной периодическими пространственными изменениями показателя преломления. В результате этого исходная волна на выходе из среды расщепляется на несколько волн с различными направлениями волновых векторов. Рассеяние света на ультразвуке идет на динамической фазовой решетке, и это видоизменяет свойства рассеянных волн. Именно в этом отношении рассеяние на ультразвуке существенно отличается от рассеяния на статической решетке.

Напомним, что рассеяние на множестве отдельных рассеивателей когерентно, если в дальней зоне складываются напряженности полей отдельных рассеянных волн, но не их интенсивности. Примерами когерентного рассеяния служат зеркальное отражение, преломление, дифракция. Как пример некогерентного рассеяния назовем фотолюминесценцию. Один из признаков когерентности рассеяния – неразличимость отдельных рассеивателей в реальном и даже в воображаемом опыте.

При когерентном рассеянии света на обычной фазовой дифракционной решетке распределение световой мощности между отдельными интерференционными максимумами задает профиль штриха решетки (угол блеска). Перераспределение энергии между отдельными волнами при рассеянии на ультразвуке сходным образом зависит от параметров акустической волны. При отсутствии поглощения сумма интенсивностей рассеянных волн равна интенсивности падающей волны.

Дифракцию света на акустических волнах наблюдают в световых пучках, поперечный размер h которых во много раз больше длины акустической волны Λ . Условие $\Lambda < h$, очевидно, выполнить тем легче, чем ниже скорость звука. Практически этим условием наложено ограничение снизу на частоту ультразвука. Ограничение сверху – $\Lambda = \lambda$ – вряд ли достижимо в видимом диапазоне, но может иметь значение в дальнем ИК диапазоне.

В противоположном случае, $h < \Lambda$, свет испытывает фазовую модуляцию при периодических изменениях локального показателя преломления, и переменную периодическую рефракцию на градиенте показателя преломления. Фазовую модуляцию можно использовать в фазовых и амплитудных модуляторах невысокого быстродействия, по оптическим схемам практически не отличающихся от электрооптических. Переменную рефракцию обычно используют в устройствах отклонения световых пучков (дефлекторах), особенности которых в данной работе мы не рассматриваем.

Различают два предельных случая дифракции световой волны на фазовой акустической решетке.

В первом случае, когда толщина решетки в направлении волнового вектора падающей на нее волны мала (*насколько*?), можно считать решетку плоской и пренебречь локальной рефракцией, наблюдаемые явления называют дифракцией Рамана–Ната.

В противоположном случае, если решетку из-за большой толщины нужно рассматривать как объемную, и рефракцию на градиентах показателя преломления приходится учитывать, рассеяние света называют дифракцией Брэгга. Важно, что в этом случае решетку правильнее рассматривать не как чисто фазовую, а как комбинированную, амплитудно-фазовую.

Модель Рамана–Ната

Рассмотрим световую волну с волновым вектором $k=2\pi/\lambda$, ортогональным волновому вектору $K=2\pi/\Lambda$ акустической волны в изотропной фотоупругой среде. Если протяженность L среды в направлении вектора k мала (см. рис. 5.1, a), то можно считать, что свет распространяется прямолинейно, а локальная неоднородность показателя преломления влияет только на фазу световой волны. Так как скорость света в среде выше скорости звука на шесть порядков, за время взаимодействия волн L/c акустическую "решетку" можно считать неподвижной.

Если на среду с показателем преломления n, в которой акустическая волна создает изменение показателя преломления Δn , падает плоская световая волна $\varepsilon = \varepsilon_0 \exp[i(kx \cdot \omega t)]$, то на выходе из среды в

плоскости x = L при стоячей акустической волне распределение напряженности поля

$$\varepsilon_{x=\ell} = \varepsilon_0 \exp[ikL(n + \Delta n \cos\Omega t \cos Kz)], \qquad (5.3)$$

а при бегущей акустической волне

$$\varepsilon_{x=\ell} = \varepsilon_0 \exp[ikL(n + \Delta n \cos(Kz - \Omega t))], \qquad (5.4)$$

где $\Omega = 2\pi f_{ak}$ – частота акустических колебаний.



Рис. 5.1. Взаимодействие света с акустической волной: дифракция Рамана–Ната (*a*) и Брэгга (*б*)

Зная поле E на плоскости x = L, можно найти распределение поля во всем полупространстве x > L. Нас интересует только распределение по углам наблюдения в дальней зоне, т.е. диаграмма направленности излучения. Его описывает интегральная формула

$$\frac{s}{\lambda} \int_{-\frac{h}{2}}^{+\frac{h}{2}} \exp(ikz\sin\theta) \mathcal{E}_{x=l} dz$$

где *s* и *h* – поперечные размеры акустической решетки, θ – угол наблюдения, а $\varepsilon_{x=L}$ задано выражениями (5.3) или (5.4); в этих выражениях можно без нарушения общности положить *n* = 0, то есть исключить несущественное постоянное запаздывание по фазе $\omega L/c$. Эти формулы дают, в частности, интенсивности излучения в дифракционных максимумах и оптические частоты излучения в этих максимумах. Опустив громоздкие явные выражения, далее укажем только самые важные результаты.

Направления на дифракционный максимум порядка *т* находят из соотношения (условия Брэгга)

$$\sin\theta_m = m\lambda/\Lambda$$
.

При бегущей акустической волне интенсивность в дифракционном максимуме порядка *m* пропорциональна квадрату значения функции Бесселя порядка *m*:

$$W_m = W_0 J_m^2(\beta_0),$$

где $\beta_0 = (2\pi L/\lambda)\Delta n$ – индекс фазовой модуляции.

Частота излучения в *m*-ном максимуме равна $\omega + m\Omega$, т. е. сдвинута относительно частоты ω падающего излучения. Движение акустической решетки не изменяет интенсивности максимумов во времени, но изменяет частоты парциальных волн за счет эффекта Доплера.

При стоячей акустической волне свет испытывает когерентное рассеяние на неподвижной решетке с изменяющейся во времени величиной *n*. Это дает модуляцию интенсивностей дифракционных максимумов на частотах, кратных 2Ω ;

$$W_m = W_0 J_m^2(\beta_0 \cos \Omega t),$$

и во всех рассеянных волнах сохраняется исходная несущая частота ω , но появляются боковые частоты $\omega \pm 2m\Omega$ за счет АМ. Мощности рассеянных волн с одинаковыми |m| равны. При наклонном падении света на решетку направления рассеяния θ_m несколько изменяются в соответствии с более общими условиями Брэгга,

$$\sin\theta_{\rm m} = \sin\psi \pm m\lambda/\Lambda$$
,

но распределение мощности по ним сохраняется. Выражения для W_m представляют собой мощности фурье-компонент сигнала с тональной фазовой модуляцией, у которого ω – несущая частота, Ω – частота модуляции, β_0 – индекс модуляции.

Дифракция Рамана–Ната с точки зрения практики – лишь теоретическая модель, пригодная только для расчетов геометрии рассеяния. Модуляция излучения при рассеянии Рамана–Ната малоэффективна: излучение на выходе распределяется по многим направлениям так, что интенсивность в отдельном максимуме составляет малую долю от входной интенсивности. Даже при большой мощности ультразвука и значительном индексе модуляции мощность распределяется по множеству направлений, вдобавок неравномерно. Только для амплитудной модуляции излучения в нулевом порядке эффективность может быть приемлемой, но и для этого нужен значительный индекс модуляции, какой трудно получить в плоской решетке. Поэтому применяют ее очень редко.

Модель Брэгга

Более эффективную модуляцию получают в геометрии Брэгга. Отметим, что так называют не просто наклонное падение волны на решетку, но когерентное рассеяние волны на существенно объемной фазовой решетке. Здесь важно учитывать не только многолучевую интерференцию, которая задает направления лепестков диаграммы направленности рассеянного излучения, но и профиль решетки, который диктует характер распределения мощности рассеянного излучения по разрешенным интерференционным максимумам; при некоторых условиях такое распределение концентрирует значительную часть мощности в один лепесток диаграммы направленности.

При описании когерентного рассеяния света на объемной ультразвуковой волне полагают, что основной элементарный эффект взаимодействия волны с акустической решеткой – отражение волны от градиента показателя преломления при почти скользящем падении волны на плоскость равного показателя преломления.

На практике модулятор обычно настраивают на рассеяние излучения в интерференционные максимумы с индексами +1 или -1 (рис. 5.1, б и 5.2, в). Теория и опыт показывают, что при соответствующей настройке оптики можно сделать максимальным сечение рассеяния света именно в один из этих максимумов. Далее мы покажем это на простой модели, которая будет очень похожей на отражающую фазовую дифракционную решетку с необычным профилем штриха (см. рис. 5.1, б и рис. 5.2, в).

Будем в первом приближении полагать, что ультразвуковая волна достаточно мощная, градиенты показателя преломления в ней значительные, и поэтому отражение при наклонном падении света происходит в основном на плоскостях максимального градиента показателя преломления. Отметим, что коэффициент отражения даже на очень малом градиенте показателя преломления при почти скользящем падении света совсем не мал.

Выражение для напряженности поля рассеянного излучения в дальней зоне можно вывести, исходя из модели дифракции Фраунгофера, как это делают при анализе диаграммы направленности рассеяния на обычной дифракционной решетке.



Рис. 5.2. Дифракция света на ультразвуковой волне: схема Брэгга: *а* – дифракция на апертуре, *б* – дифракция на штрихе, *в* – интерференционные лепестки диаграммы направленности рассеянного излучения (*B*) и их дифракционная огибающая (*A*)

Ультразвуковая решетка отличается от обычной тем, что ее период и эффективный размер штриха велики по сравнению с длиной волны света. Дифракция на одном узком штрихе обычной дифракционной решетки дает широкую диаграмму направленности рассеяния. Дифракция на одном эффективно-широком штрихе ультразвуковой решетки дает узкую диаграмму направленности. Эта частичная диаграмма направленности, как обычно, задает распределение мощности рассеянного излучения по интерференционным максимумам (лепесткам диаграммы направленности) когерентного рассеяния света на решетке в целом. Несложно убедиться в том, что из-за относительной остроты максимумов дифракционной огибающей даже при малом угловом расстоянии между интерференционными максимумами, сосредоточенные в них мощности излучения могут очень сильно различаться. Простота выбранной нами модели позволяет провести все расчеты явно и до конца.

Дифракцию на одном штрихе (рис. 5.2, б) описывает диаграмма направленности

$$A = \sin\alpha/\alpha = \operatorname{sinc} (\alpha), \qquad (5.5)$$

где $\alpha = (\pi n L/\lambda)(\cos\theta - \cos\psi)$ – разность фаз между волнами, рассеянными *от края и середины одного штриха*. Ширину углового распределения можно оценить из условия A = 0, то есть $\alpha = \pm \pi/2$. Для более явного выражения ширины лепестка $A(\theta)$ нужно учесть значение $\cos(\psi)$ в конкретной оптической схеме (см. далее). Выражение (5.5) относится к штриху с резкими краями, а в ультразвуковой волне штрих решетки по эффективному профилю ближе к косинусному; поправочный коэффициент к ширине штриха α при этом порядка (несколько менее) единицы.

Многолучевую интерференцию волн, рассеянных освещенными штрихами решетки (рис. 5.2, *a*), описывает диаграмма направленности

$$B = \sin\beta/\beta \equiv \operatorname{sinc}(\beta), \qquad (5.6)$$

где $\beta = (\pi h/\lambda)(\sin\theta - \sin\psi \pm \lambda/\Lambda) - разность фаз между волнами,$ рассеянными крайним и средним освещенными штрихами решетки, для первого и минус первого порядков. Ширину такого лепестка диаграммы направленности можно оценить по аналогии с шири $ной <math>A(\theta)$, с учетом sin ψ в конкретной оптической схеме (см. далее). Здесь поправка к β , которая дает некоторое уширение *B*, может быть вызвана затуханием и расходимостью ультразвуковой волны.

В выражениях (5.5, 5.6) укажем примечательное общее свойство всех без исключения случаев когерентного рассеяния монохроматических волн: *ширина лепестка диаграммы направленности рассеяния обратно пропорциональна максимальной разности фаз между частичными рассеянными волнами*. Нетрудно показать, что этот факт равносилен соотношению неопределенностей для импульса и координаты фотона.

Множитель *А* описывает дифракцию на одном периоде (штрихе) решетки, а *B*, по Брэггу, задает направления на первый и минус первый максимумы:

$$\sin\theta_{1,-1} = \sin\psi \pm \lambda/\Lambda.$$
 (5.7)

Формально интенсивность рассеянного света пропорциональна степени перекрытия главных максимумов A и $B_{\pm 1}$, зависимость ко-

торых от угла рассеяния θ показана на рис. 5.2, *в*. С ростом частоты модуляции Ω (с уменьшением периода решетки) максимумы $B_{\pm 1}$ раздвигаются. При увеличении *L* (ей пропорциональна эффективная ширина штриха решетки) сужаются максимумы *A*. При выполнении условия

$\sin \psi = \lambda / \Lambda$

пик В.1 совпадет с пиком А. Соответствующий угол падения

$$\psi_{\rm B} = \arcsin(\lambda/2\Lambda)$$

называют углом Брэгга.

При $\psi=\psi_B$ направления на первый и минус первый максимумы дают выражения

$$\theta_{+1} = \arcsin(\lambda/2\Lambda), \quad$$
или $\theta_{+1} = \psi_{\rm b},$ (5.8)

$$\theta_{-1} = \arcsin(3\lambda/2\Lambda).$$
 (5.9)

Из (5.5)...(5.9) следует, что в геометрии Брэгга для направления θ_{-1} значение $A_{-1} = 1$, а для направления θ_{+1}

$$4_{+1} = \sin(-\pi L\lambda/\Lambda^2)/(-\pi L\lambda/\Lambda^2), \quad |A_{+1}| << 1$$
(5.10)

В геометрии Брэгга при условии $\pi L\lambda >> \Lambda^2$ рассеянный решеткой свет можно почти полностью сосредоточить в минус первом максимуме. В нулевом (неотклоненном) максимуме, тем не менее, еще может оставаться значительная доля общей мощности света.

До сих пор мы рассматривали только диаграмму направленности рассеянного излучения, но не сечение рассеяния. Для оценок сечения рассеяния света (или *дифракционной эффективности решетки*) одной лишь геометрии недостаточно, и нужно учитывать не только детали дифракции света на штрихе решетки, но и многократное рассеяние.

Можно упрощенно рассматривать ультразвуковую решетку как периодическую структуру из пластинок, на плоскостях которых свет отражается, а через сами пластинки проходит без поглощения. Но даже при таком кажущемся упрощении задача остается сложной потому, что нужно учитывать отражение при скользящем падении света на границу пластинки. Дополнительно можно упростить задачу, полагая коэффициент отражения близким к 100%.

Полное сечение рассеяния света на решетке в виде стопки отражающих плоскостей может быть на порядки величины меньше, чем апертура пучка света. Чтобы сечение рассеяния на сильном градиенте показателя преломления было близким к апертуре пучка при коэффициенте отражения 100%, в рамках нашей упрощенной модели нужно выполнить очевидное по своей сути условие:

$$L \approx \Lambda \operatorname{ctg}(\psi);$$
 (5.11)

такой должна быть минимальная ширина планок жалюзи, непроницаемых для прямого света при заданном угле падения. На самом деле это условие приближенное, оно не учитывает дифракцию света в промежутке между штрихами и возможность распространения света между отражающими плоскостями, как в волноводе. Поэтому все указанные упрощения не помогают правильно оценить сечение рассеяния. Тем не менее, даже такие примитивные соображения подсказывают, что в принципе существует некоторый оптимум продольной протяженности фронта ультразвуковой волны L, при котором перекачка мощности света в один интерференционный максимум будет наиболее эффективной.

Более точная теоретическая модель дает зависимость максимально достижимой мощности света, рассеянного в минус первый максимум, от геометрических параметров модулятора и мощности возбуждения ультразвука:

$$W_{-1} = W_0 \sin^2\left(\frac{\pi L \Delta n}{\lambda}\right) = W_0 \sin^2\left(\frac{\pi}{\lambda} \sqrt{\frac{L P_{ac} M}{2S}}\right)$$
(5.12)

где P_{ac} – акустическая мощность,

S – протяженность фронта ультразвуковой волны в направлении распространения света,

 $M = n^6 p^2 / \rho v^3 -$ коэффициент фотоупругого качества среды,

где *p* – упругооптический коэффициент,

 ρ – плотность,

v – скорость звука,

Из приведенного выражения очевидна немонотонная зависимость рассеянной мощности от мощности ультразвука; аргумент синуса в нем – уже знакомый нам индекс фазовой модуляции. О происхождении выражения (5.12) и о порождающей его модели сказано в приложении 3, где обоснована зависимость индекса модуляции от мощности акустической волны и параметров материала.

При определенном уровне акустической мощности можно почти полностью преобразовать падающий свет в минус первый дифракционный максимум, и достичь этого тем легче, чем больше показатель преломления среды и чем ниже скорость звука в ней. Отношение W_{-1}/W_0 может превышать 50% на $\lambda = 0,63$ мкм при $P_{ac} = 1$ Вт.

До сих пор мы рассматривали рассеяние плоской монохроматической волны на плоской акустической волне, ограниченной только в направлении волнового вектора света. Реальные волны любой природы всегда имеют конечную апертуру и ненулевую кривизну фронта, что необходимо учитывать при оценках расходимости рассеянных волн на выходе из АОМ. Ограничение апертуры световой волны в направлении волнового вектора акустической волны фактически уже учтено в (5.6). Дополнительные поправки, учитывающие кривизну фронтов обеих волн, можно оценить, применив подход, подробнее описанный в разделе 7 (глава «ОВФ в реальных пучках ограниченного сечения»).

5.3. Устройство акустооптического модулятора

Основу акустооптического модулятора составляют модуляционный элемент – брусок из прозрачного для света фотоупругого материала, и излучатель акустической волны – пьезоэлемент. Пьезоэлемент прочно присоединен к модуляционному элементу диффузионной сваркой или сформирован при помощи термического напыления. Рабочие оптические поверхности тщательно полируют и, как правило, наносят на них просветляющие покрытия.

Фотоупругим материалом для модуляционного элемента могут быть различные твердые вещества и жидкости. В видимом и ближних ИК и УФ диапазонах применяют кварц, сапфир, ниобат лития, молибдат свинца, парателлурит, дигидрофосфаты и др. В видимом диапазоне также пригодны жидкости: вода, ксилол, четыреххлористый углерод, сероуглерод, спирты. Вода, например, выгодна из-за низкой скорости звука в ней; твердые материалы обычно характеризует меньшее, чем у жидкостей, затухание звука, что особо важно при работе на высоких частотах.

Акустооптические модуляторы особенно эффективны в ИК диапазоне, где многие материалы прозрачны, имеют высокий показатель преломления (2,5..4) и высокую акустическую добротность. К ним относятся кремний, германий, арсенид галлия, теллур, иттриевые гранаты.

В качестве излучателей акустических волн обычно применяют пьезоэлементы из кристаллов кварца, титаната бария, ниобата лития, иных пьезокристаллов из титанатов, цирконатов, ниобатов, или пьезокерамики. Подобные излучатели хорошо работают в диапазоне до десятков мегагерц. На практике чаще всего применяют акустооптические модуляторы, позволяющие модулировать свет на частотах до 100 МГц. В принципе можно создать приборы, работоспособные в диапазонах до 1000 МГц. Для возбуждения модуляторов используют мощности ультразвука, обычно не превышающие 10 Вт.

В модуляторах можно использовать как бегущую, так и стоячую акустическую волну. Конструктивно в модуляторе со стоячей волной либо монтируют на противоположных торцах фотоупругого элемента два фазированных пьезоэлемента, либо устанавливают один пьезоэлемент и акустический отражатель. В модуляторе с бегущей волной пьезоэлемент один, а отражения от противоположного торца устраняют. Модуляторы с бегущей волной более широкополосные; модуляторы со стоячей волной могут быть более эффективными за счет использования механического резонанса в фотоупругой среде. Рабочая полоса частот возбуждения модулятора в последнем случае получится узкой, но при этом резко снижается потребляемая электрическая мощность.

В акустооптических модуляторах часто применяют двулучепреломляющие материалы. Двойное лучепреломление дает существенно лучшие результаты в устройствах отклонения световых волн (дефлекторах). Об этом подробнее сказано в приложении 3.

В реальных конструкциях акустооптических модуляторов достаточно заметно проявляются резонансные свойства пьезоэлемента. Причина этого - отражения акустической волны на стыке пьезоэлемента с фотоупругой средой и от обратной поверхности пьезоэлемента. Резонансы проявляются как неравномерность частотной характеристики модулятора: возникают пики вблизи нечетных гармоник частоты основного электромеханического резонанса, провалы вблизи четных гармоник, и в соответствии с этим существуют резонансные пики в зависимости потребляемой электрической мощности от частоты. Поэтому АОМ обычно наиболее эффективно работает в некотором диапазоне частот, ограниченном сверху примерно удвоенной частотой основного электромеханического резонанса. Напомним, что пьезоэлемент, имеющий механические резонансы на частотах Nfres, электромеханические резонансы имеет только на частотах $(2N-1) \times f_{res}$, и основной электромеханический резонанс совпадает с низшим механическим, N = 1.

5.4. Применения акустооптических модуляторов света

Способность AOM управлять несколькими параметрами световой волны стала основой их многообразных применений в технике и в лабораториях.

Амплитудную модуляцию света в акустооптических модуляторах получают при управлении акустической мощностью. В зависимости от условий применения, используют AM как в отклоненных пучках, так и в исходном пучке.

Обратим внимание на то, , что частота колебаний в рассеянной волне сдвинута на частоту модуляции по отношению к исходной. При изменении акустической (модулирующей) частоты получают частотную модуляцию световых волн в ненулевых порядках. Такую ЧМ неизбежно сопровождает некоторая паразитная АМ, и на практике это приходится учитывать.

Также обратим внимание и на то, что в акустооптическом модуляторе Брэгга для каждого из интерференционных максимумов в сущности реализуется режим однополосной модуляции (SSB, *single sideband*), причем выделению одной боковой способствуют как преимущественная перекачка света в один максимум, так и пространственная фильтрация выходной мощности (в радиотехнике SSB получают гораздо сложнее).

Наконец, обратите внимание на то, что акустооптический модулятор – это еще и управляемый частотой ультразвука *дефлектор* (устройство для отклонения) светового луча. Угол отклонения волнового вектора света монотонно зависит от частоты модуляции. При этом, в сочетании с пространственной фильтрацией одного лепестка диаграммы направленности (простым диафрагмированием), дефлектор может выполнять функцию перестраиваемого фильтра света по длине волны.

Проще говоря, AOM – это амплитудный модулятор, частотный модулятор и дефлектор света в едином устройстве.

Примечательна и практически важна возможность применения АОМ в оптико-электронном анализаторе спектра – формирователе светового поля, распределение яркости в котором воспроизводит распределение спектральной плотности в электрическом сигнале, поданном на модулятор. Такие анализаторы спектра радиосигналов, с успехом применяемые в радиоастрономии и военной радиоразведке, позволяют получать результат в реальном времени, а также анализировать спектр *одиночного* радиоимпульса, что недоступно для гетеродинных анализаторов радиодиапазона.

АОМ в качестве анализаторов спектра и перестраиваемых по длине волны фильтров особенно выгодно применять в виде интегрально-оптических модулей, где взаимодействуют не световые и ультразвуковые пучки в объемном материале, а свет в планарном волноводе и поверхностная волна ультразвука. Для возбуждения поверхностной ультразвуковой волны в таких приборах применяют пленочные планарные пьезоэлектрические структуры со встречноштрыревыми гребенчатыми электродами, наложенные на планарный оптический волновод.

АОМ широко применяют при решении ряда задач обработки информации. Пример пространственно-временного модулятора света (ПВМС) на основе многоканального АОМ был описан ранее в разделе 3.4. К сожалению, возможности миниатюризации подобных устройств ограничены из-за необходимости фильтрации (подавления) излучения в нулевом порядке дифракции.

Отметим рабочие характеристики модулятора: длительность обрабатываемых сигналов равна L/v, где L – длина модулятора вдоль оси x, v – скорость звука; на практике длина эффективного взаимодействия света с ультразвуком ограничена затуханием акустической волны. Для достижения наилучшего частотного (и пространственно-частотного) разрешения необходимо использовать среды с минимальной скоростью звука и минимальным его затуханием. Например, анизотропного кристалла TeO₂ (парателлурид) имеет для волны, возбуждаемой при определенном срезе из акустических волн скорость порядка 0,6 км/с, что позволяет обрабатывать сигнал длительностью до 20 мкс с полосой частот до 100 МГц. Другие кристаллы (например, LiNbO₃, где скорость ультразвуковых волн достигает 7 км/с), позволяют работать с сигналами длительностью до 1 мкс в полосе частот до 10 МГц.

По совокупности параметров наиболее эффективны модуляторы на таких фотоупругих материалах, как LiNbO₃, TeO₂, PbMbO₃, GaP и некоторые другие. Они обеспечивают достаточно высокую эффективность модуляции света и малое время включениявыключения оптического отклика, ~ 10⁻⁶...10⁻⁹ с.

104

6. ПРОСТРАНСТВЕННО-ВРЕМЕННОЙ МОДУЛЯТОР ТИПА ФП-ЖК

Пространственно-временные модуляторы света (ПВМС) служат ключевыми элементами информационных и вычислительных систем. Функциональная роль ПВМС весьма многогранна: отображение и визуализация информации (дисплеи), ввод оптических сигналов, преобразование по длине волны когерентности несущей, усиление яркости изображений, обработка изображений (пространственная фильтрация и др.)

В зависимости от вида управляющих сигналов, ПВМС принято подразделять на электрически управляемые (модуляторы типа «электрический сигнал – свет»), и оптически управляемые (модуляторы типа «свет – свет»). Одним из самых распространенных в настоящее время типов ПВМС являются жидкокристаллические ПВМС.

6.1. Жидкие кристаллы в модуляторах света

Жидкие кристаллы (ЖК) отличаются чрезвычайно высокой чувствительностью к управляющим воздействиям – в особенности к электрическим сигналам, достаточно хорошо выраженными пороговыми свойствами и высокой крутизной модуляционной характеристики для большинства электрооптических эффектов в ЖК, технологичностью изготовления на их основе многоэлементных приборов и устройств большого формата. В ЖК-смесях эвтектического состава температурный диапазон существования мезофазы достигает десятков градусов и включает комнатные температуры. Они способны надежно работать десятки тысяч часов.

Мезофазы в зависимости от вида и степени упорядоченности молекул подразделяют на нематическую, холестерическую и десятки смектических фаз (см. рис. 6.1).

Среднее направление ориентации осей молекул ЖК в любой точке пространства принято описывать полем единичного вектора молекулярной ориентации n(r), называемого директором. Ориентация директора в слое ЖК определяется граничными условиями на поверхностях подложек содержащей ЖК ячейки и оррелированноем молекул между собой и с внешним электрическим полем. Деформация директора в ориентированном слое ЖК

взаимосвязана с изменением его оптических и электрических свойств.



Рис. 6.1. Типы упорядочения протяженных молекул в мезофазах (ЖК):
 a – нематическая (упорядочена только ориентация, слоистой структуры нет),
 b- холестерическая (упорядочена ориентация молекул в отдельных слоях;
 в смежных слоях ориентация развернута на определенный угол),
 b – смектическая (слоистая структура с ориентацией осей молекул
 под фиксированным углом к плоскости слоя; как пример показана смектическая фаза простейшего типа А, где оси молекул
 ортогональны вертикальным слоям).

В нематической фазе молекулы стремятся ориентироваться так, чтобы их главные оси были параллельны (рис. 6.1, *a*). При этом в объеме мезофазы образуются характерные нитевидные структуры, выявляемые при наблюдениях в поляризованном свете. Корреляция ориентаций молекул очень сильна, но корреляции между продольными координатами молекул в соседних нитях слабы.

Нематическую мезофазу формируют неполярные удлиненные молекулы, которые взаимодействуют с внешним полем и друг с другом в соответствии с моделью ван дер Вальса (рис. 6.2, *a*). В отсутствие внешнего поля дипольный момент молекулы (поляризация) отсутствует. Поляризация во внешнем поле возникает как изменение распределения зарядов внутри молекулы под действием внешней возмущающей силы. Энергия взаимодействия молекулы с полем зависит от модуля проекции вектора напряженности поля на ось молекулы. Следовательно, быстрая перемена направления поля оставляет ориентацию молекулы неизменной, тогда как проекция дипольного момента на ось меняет знак. Так ведут себя все неполярные молекулы как в жидкой фазе, так и в мезофазе.

Ориентация молекул вблизи от подложки возникает вначале спонтанно, за счет притяжения спонтанно возникающих поляризационных зарядов противоположного знака в молекулах и в материале подложки (рис. 6.2, б). Поле вначале слабой спонтанной поляризации молекулы наводит поляризацию в диэлектрической подложке, эта вторичная слабая поляризация своим полем наводит дополнительную поляризацию в молекуле, и т. д. Примерно так же взаимодействуют между собой молекулы, удаленные от подложки.

В основе анализа всех ориентационных эффектов в ЖК лежит термодинамический подход: распределение директора n(r) должно соответствовать минимуму полной свободной энергии объема ЖК при заданных граничных условиях. Полная свободная энергия объема ЖК есть

$$F = F_{\rm e} + F_{\rm ce},$$

где $F_{\rm e}$ – энергия упругой деформации слоя ЖК без учета поверхностных сил, $F_{\rm ce}$ – энергия взаимодействия ЖК с электрическим полем.



Рис. 6.2. Взаимодействие ван дер Ваальса для неполярных молекул в жидкостях и мезофазах (ЖК): *a* – прилипание к диэлектрической подложке, *б* – взаимная ориентация сближенных молекул;

Преимущественная ориентация молекул возникает именно как состояние с наименьшей свободной энергией. Если не учитывать взаимодействие молекул между собой, то такая ориентация возникает под действием внешнего поля и разрушается тепловым движением при выключении поля. Но в мезофазе существенно именно взаимодействие молекул друг с другом, и оно может настолько снизить свободную энергию ориентированного ансамбля, что однажды возникшая поляризация сохраняется и в отсутствие внешних воздействий, а тепловое движение уже не может ее разрушить. Подчеркнем, что первопричина формирования мезофазы – не взаимодействие ван дер Ваальса само по себе, но участие в нем молекул со специфическим сочетанием конфигурации и поляризуемости, причем в относительно ограниченном диапазоне температур.

В холестерической мезофазе заметна слоистая структура. Внутри слоев расположение длинных молекул коррелированно практически так же, как и в нематической фазе. Ориентация молекул в смежных слоях холестерика также корррелирована – она повернута на определенный угол. Если проследить зависимость директора от координаты в направлении, ортогональном холестерическому слою, то конец вектора описывает характерный винт, шаг которого зависит как от вида молекул холестерика, так и от температуры. Такие корреляции являются причиной всех характерных для холестериков оптических эффектов. В проходящем излучении холестерики проявляют оптическую активность (вращение плоскости поляризации), в отраженном – сильную зависимость амплитуды рассеяния от угла и длины волны. Эффекты, наблюдаемые в отражённом свете, напоминают интерференцию в тонких пленках и периодических слоистых структурах, и по своей природе являются результатом интерференции при рассеянии света на слоистой структуре, период которой тождествен шагу холестерического «винта». Наиболее характерное явление – изменение окраски слоя холестерического ЖК при измененении его температуры, вызванное изменением шага винта; его часто используют для высокочувствительной индикации температурных полей.

Смектические ЖК, в которых упорядоченно-ориентированные молекулы образуют слои с расположением концов удлинённых молекул в одной плоскости, то есть с сильной корреляцией ориентации и продольного расположения молекул (рис. 6.1, *в*). Взаимное скольжение слоев происходит очень легко, что означает слабость межслоевого (продольного) взаимодействия и относительную силу внутрислоевого (поперечного). По современным данным, смектические ЖК играют важную роль в субклеточных биоструктурах. Смектики – по существу, собирательное название более чем десятка мезофаз, различающихся между собой дополнительным более сложным упорядочением и корреляциями ориентации молекул как трехмерных объектов внутри слоев. Некоторые смектические мезофазы пригодны для оптоэлектронных приборов различного назначения.

Изучаемый в данной работе ПМС использует нематический ЖК.
6.2. Электрооптические эффекты в нематических ЖК

В ЖК некоторых типов можно при помощи электрического поля управлять двулучепреломлением и на этой основе модулировать фазу и поляризацию света. В ЖК, представляющем собой текучую вязкоупругую диэлектрическую среду, изменение двулучепреломления физически обусловлено поворотом (переориентацией) самих молекул ЖК, т.е. упругой механической деформацией молекулярного слоя под действием электрического поля.

Двулучепреломление ориентированных агрегатов молекул ЖК возникает как следствие анизотропии самих молекул (например, их удлиненной формы). Этот эффект проще всего представить себе как поведение валентных электронов молекулы в трехмерной потенциальной яме. При смещении электрона вдоль длинной оси сила упругости гораздо меньше, чем при смещении поперек оси. Поэтому поляризуемость и диэлектрическая восприимчивость к продольному электрическому полю у подобных молекул больше, чем восприимчивость к поперечному полю. Аналогичный эффект наблюдают и в удлиненных молекулах красителей, когда их удается ориентировать на подложке или в растянутой полимерной пленке. Вообще молекулы нематических ЖК и красителей во многом похожи по структуре, но полезные для техники ЖК имеют полосы поглощения в сине-фиолетовой области и слабо поглощают видимый свет, тогда как полосы поглощения у разных красителей встречаются и в видимом диапазоне. Ориентированные удлиненные молекулы красителей вблизи от полос поглощения проявляют дихроизм, и потому на их основе нередко делают поляризаторы в виде растянутых пленок окрашенных полимеров. Вдали от полос поглощения в окрашенных текстурах, как и в ориентированных нематических ЖК, наблюдают двулучепреломление.

В нематическом ЖК с положительной диэлектрической анизотропией, для которого поляризуемость вдоль длинной оси молекулы больше, чем для поперечной, т.е. при

$$\Delta \varepsilon = \varepsilon_{\max} - \varepsilon_0 > 0,$$

молекулы стремятся выстроиться длинными осями вдоль силовых линий электрического поля. Поэтому первоначально молекулы таких ЖК однородно ориентируют длинными осями в плоскости подложек с нанесенными на них оптически прозрачными электродами (окись олова или индия). Для образования первичной ориентации – текстуры – проще всего использовать свойство молекул ориентироваться вдоль направления микроцарапин, образующихся при механической полировке подложки (полиимидной пленки) в одном направлении. Между двумя такими ориентирующими подложками с пленочными прозрачными электродами помещают каплю ЖК, которая под действием капиллярных сил заполняет тонкий зазор между подложками. Так образуется простейшая ЖК-ячейка. Если ориентация молекул нематика на обеих подложках одинакова, то ориентация молекул однородна по всему слою ЖК, и ее называют планарной, или *S*-текстурой. Если же направления полировки подложек ортогональны, то ориентация молекул по мере продвижения от одной подложки к другой постепенно поворачивается. Такую неоднородную ориентацию молекул называют твисттекстурой (закрученной, *T*-текстурой).



Рис. 6.3. Взаимодействие коллектива молекул в нематической мезофазе: *a* – прилипание к диэлектрической подложке с наноразмерными канавками, *б* – формирование текстуры у ориентирующей поверхности

Один из элементарных эффектов в НЖК – *S*-эффект, получивший свое название по виду деформации директора (*splay* – продольный изгиб).

При S-деформациях (рис. 6.4, a) наиболее отчетливо изменение двулучепреломления. Разность фаз Ф между обыкновенной и необыкновенной волной однозначно связана с распределением показателя преломления n(z) по толщине ячейки. Для монохроматического света с длиной волны λ разность фаз

$$\Delta \Phi = 2\pi \Delta n L/\lambda, \quad \Delta n = (1/L) \int_{L} [n(z) - n_{\perp}] dz;$$

$$n(z) = n_{\parallel} n_{\perp} \times [n_{\perp}^{2} \cos^{2}\theta(z) + n_{\parallel}^{2} \sin^{2}\theta(z)]^{-1/2},$$

где n_{\parallel} , n_{\perp} – главные значения показателя преломления (или главные оси эллипсоида оптической индикатрисы). Поместив ЖК-ячейку

между скрещенным поляризатором и анализатором, получим модуляцию интенсивности

$$I = I_0 \sin^2(2\varphi_0) \sin^2(\Delta \Phi/2),$$

где I_0 – интенсивность падающего света с учетом пропускания поляроидов, φ_0 - угол между осью поляризатора и начальной ориентацией молекул на подложках.



Рис. 6.4. Переориентация в текстурах НЖК под действием электрического поля (переход Фредерикса).
Показаны изменения директора в пространстве и при наложении поля: *1* – поле отсутствует, вид вдоль поверхности подложек; *2* – поле отсутствует, вид в плане; *3* – поле включено, вид вдоль поверхности подложек; *a* – планарная текстура (*S*), *б* – твист-текстура (*T*)

Для световой волны, ортогональной к плоскости подложек, S-текстура имеет наиболее сильно выраженное двулучепреломление; $\Delta n_{\text{max}} = n_e - n_o$. Здесь n_e и n_o обозначены соответственно показатели преломления для поляризации света, совпадающей с длинной осью молекулы, и для поляризации, ортогональной ей, т.е. показатели преломления для необыкновенной и обыкновенной волн. При достижении некоторого порогового напряжения V_{Π} на электродах (обычно единицы вольт), момент электростатических сил преодолевает момент сил упругости, удерживающих молекулы в исходном положении, и они начинают переориентацию: сначала в центре слоя, где взаимодействие с подложкой минимально, а затем и по всей толщине слоя. Этот процесс переориентации известен с 30-х годов под названием *перехода Фредерикса*. В этом процессе двулучепреломление $n_e - n_o$ уменьшается от максимума почти до нуля.

При напряжении $V > V_{\Pi}$ направления осей молекул в слое характеризуются некоторыми углами к направлению первоначальной ориентации и становятся функциями координаты слоя по толщине. Соответственно уменьшается проекция длинных осей молекул на плоскость подложки и связанное с ней значение большего показателя преломления n_e , тогда как значение n_o остается неизменным. В результате двулучепреломление ЖК ячейки определяется некоторой эффективной разностью показателей преломления $\Delta n_{эф\phi}$, которая максимальна для недеформированного слоя ЖК и приближается к нулю (ячейка почти изотропна для проходящего света) при $V > V_{\Pi}$. Установив «скрещенные» поляризаторы, под углом 45° к направлению исходной ориентации молекул ЖК, преобразуем изменение двулучепреломления через модуляцию поляризации в модуляции интенсивности проходящего света. Зависимость интенсивности света за поляризатором от $\Delta n_{эфф}$ имеет вид

$I \sim I_0 \sin^2(\pi L \Delta n_{igg}/\lambda).$

Для многих ЖК $\Delta n_{\text{max}} = 0.2 ... 0.4$, поэтому изменение двулучепреломления в электрическом поле обеспечивает максимальную глубину модуляции интенсивности видимого света уже на толщине слоя *L* около 1 мкм. В более толстых ячейках можно получить до нескольких чередующихся пиков и минимумов пропускания света (рис. 6.5, *a*), а однозначное изменение интенсивности (в пределах изменения задержки по фазе от 0 до π) можно получить при незначительном изменении ориентации молекул в части слоя ЖК. Изменение разности фаз при *S*-эффекте может превышать 10 π при толщине слоя ЖК до 10 .. 20 мкм. Поэтому *S*-текстуры наиболее критичны к управляющему напряжению, и применяют их чаще всего в специализированной аппаратуре. Возможность получить значительные фазовые сдвиги в толстой *S*-текстуре особенно ценна для адаптивных оптических систем.

Механический по сути характер переориентации молекул ЖК в вязкой среде ограничивает скорость оптического отклика на электрический сигнал, т.е. быстродействие ЖК ячейки. Для разных типов ЖК оно варьируется от секунд до микросекунд при напряжениях, соответствующих полуволновому (от десятков до долей вольта). В настоящее время интенсивно исследуют эффекты *кооперативного* взаимодействия молекул в некоторых ЖК системах, где изменения ориентации происходят за сотые доли микросекунды; подобные системы отчасти похожи на сегнетоэлектрики. Кроме того, известны молекулы, в которых ориентационные эффекты связаны не с поворотом молекулы как целого, а с изгибом молекулы; на такие процессы слабее влияет вязкость.



Рис. 6.5. Зависимость пропускания ЖК ячейки от управляющего напряжения: a – толстая *S*-текстура, возможные варианты при различной толщине слоя ЖК; δ – *T*-текстура

Возврат молекул к исходной ориентации при выключении поля происходит по действием упругих сил, причем время этого процесса (время релаксации или время памяти) обычно на 1-2 порядка больше времени электрооптического отклика. (Ныне известны молекулы ЖК, которые электрическим полем существенно более высокой частоты можно ориентировать длинной осью *поперек* поля; при этом перестройка гомеотропной текстуры в исходную планарную происходит практически с той же скоростью, как и перестройка текстуры планарной в гомеотропную.)

6.3. Электрооптические модуляторы света на основе твист-текстур ЖК

В простейших электроооптических ЖК модуляторах, как и в индикаторах и дисплеях, используют изменение поляризации проходящего света при перестройке тектуры нематика под действием электрического поля. Исходной может быть либо планарная текстура, либо твист-текстура. В обоих случаях текстуры образуются за счет ориентирующего действия подложек. В обоих случаях под действием электрического поля исходная текстура переходит в гомеотропную, не влияющую на поляризацию проходящего через нее излучения. В предыдущем разделе описано действие ЖК-ячейки с планарной текстурой.

Для практики более удобна ячейка с твист-текстурой. Получить ее не труднее, чем планарную, а применять ее в модуляторах различной конфигурации несравненно проше. В частности, гораздо легче получить монотонную зависимость пропускания ячейки от управляющего напряжения. Кроме того, при изменении ориентации лишь одного из двух поляризаторов ячейку с твист-текстурой можно перевести из режима затвора, отпираемого электрическим сигналом, в режим затвора, запираемого таким же сигналом. При использовании планарной текстуры этого добиться труднее. Причина состоит в том, что планарная текстура по оптическим свойствам эквивалентна двулучепреломляющей пластинке, и ориентировать ее ось по отношению к оси поляризатора следует под углом 45°, а по вносимой разности фаз она должна быть как можно ближе к пластинке λ/2; только при этих условиях можно получить полное просветление при скрещенных поляризаторах (и гашение света при подаче управляющего напряжения), полное гашение при параллельных поляризаторах (и просветление при подаче управляющего напряжения). Отклонения от этих условий приводят к заметному уменьшению глубины модуляции, неполному гашению или просветлению ячейки (рис. 6.5, а). Такое снижение качества модулятора особенно неприятно в индикаторах и дисплеях, где желательно, чтобы контраст активного сегмента на общем фоне был максимален, а пассивный сегмент был бы просто невидимым.

Изменение директора твист-текстуры под действием поля называют *T*-переходом (рис. 6.4, δ). При *T*-деформациях наблюдают *вращение плоскости поляризации*. Твист-текстура, в отличие от планарной, без управляющего напряжения в идеале только поворачивает на 90° эллипс поляризации проходящего через нее света, не меняя степень эллиптичности, при произвольной ориентации осей этого эллипса. Обычно направление текстуры одной из подложек параллельно, а направление второй – ортогонально оси одного из поляризаторов; ось второго поляризатора либо параллельна (*активный* сегмент виден как просветленный в темном поле, как показывает зависимость пропускания от напряжения на рис. 6.5, δ), либо ортогональна оси первого поляризатора (тогда активный сег-

мент виден как затемненный в светлом поле). При этом двулучепреломление малозаметно, так как оно появляется только при отклонениях ориентации подложек и поляризаторов от номинальных при сборке. Но даже и в этом случае небольшое остаточное двулучепреломление практически влияет не на контраст, а только на общее пропускание света в пассивной структуре. ЖК индикаторы массового применения изготавливают почти исключительно на основе нематических твист-текстур.

6.4. Модулятор света со структурой фотопроводник-ЖК

Исторически одним из первых ПМС класса "свет-свет" была структура фотопроводник-ЖК (ФП-ЖК) (рис. 6.6). В ней слой поликристаллического или аморфного ФП и слой ЖК помещены между прозрачными электродами, к которым прикладывают постоянное или переменное питающее напряжение. Считывающее излучение проходит через слой жидкого кристалла, отражается от диэлектрического зеркала и снова проходит через слой ЖК.

При освещении ФП его проводимость увеличивается, и поэтому напряжение на слое ФП падает, а на слое ЖК растет. Вызванный этим электрооптический эффект в слое ЖК создает пространственно-зависимую (пространственную) модуляцию считывающего отраженного излучения в соответствии с распределением освещенности фотопроводника. Изменение поля освещенности фотопроводника во времени вызывает пространственно-временную модуляцию излучения.

В подобных структурах следует добиваться минимального воздействия считывающего излучения на фотопроводник за счет высокого качества отражающего слоя, который может быть либо сплошным, либо дискретным (мозаика). В одном из наиболее технически совершенных приборов применено диэлектрическое зеркало из 12 чередующихся слоев сульфида цинка и фторида магния (криолита). Оно представляет собой по существу комбинацию из двух зеркал с максимумами отражения на длинах волн соответственно 488 и 632 нм, поэтому имеет коэффициент отражения не менее 90% в диапазоне 400 .. 700 нм.

Дополнительно надежную развязку каналов записи и считывания дает блокирующий (поглощающий) слой CdTe толщиной 2 мкм с поверхностным сопротивлением 10 Ом.кв. и поглощением более 10⁵ на длине волны 525 нм. Этот слой образует гетеропереход с ФП, в качестве которого обычно используют поликристаллический слой CdS толщиной 12...16 мкм.



Рис. 6.6. Структура ПВМС типа ФП-ЖК (толщина слоев показана условно, без соблюдения масштабов): *1* – считывающее излучение, *2* – просветляющее покрытие, *3* – стеклянные подложки, *4* – слой ЖК, *5* ориентирующие покрытия подложек, *6* – изолирующая прокладка, *7* – диэлектрическое зеркало, *8* – блокирующий поглощающий слой, *9* – фотопроводник,

10 – записывающее излучение, 11 – прозрачные электроды

Для питания структур ФП-ЖК с диэлектрическими слоями используют переменное напряжение. Это увеличивает срок службы прибора и улучшает его быстродействие, так как в этом случае исключены электролиз ЖК и накопление заряда на границе ФП-ЖК.

Чувствительность такой структуры при использовании эффекта динамического рассеяния в ЖК составляла 50..500 мкВт/см² (в зависимости от выбранного режима работы), время включения около 10 мс, время выключения – 30 мс, пространственное разрешение не превышало 20 лин/мм.

Наилучшие показатели пространственной модуляции света в ПВМС типа ФП-ЖК удалось получить на основе ориентационных S и В-эффектов. При этом весьма важную роль играет диэлектрическая анизотропия ЖК – при ее уменьшении разрешение улучшается. Этот факт можно объяснить уменьшением объема ЖК, в котором упругие силы компенсируют энергию взаимодействия ЖК с неоднородным электрическим полем, а эта энергия пропорциональна диэлектрической анизотропии. Пространственное разрешение структур ФП-ЖК со слоем ЖК толщиной около 1 мкм при диэлектрической анизотропии $\Delta \varepsilon \approx 0.1$ достигло 350 лин/мм.

В результате оптимизации таких структур путем согласования полных сопротивлений фоточувствительной и электроооптических удалось достигнуть чувствительности частей примерно до 2¹⁰⁶ Дж/см². Для согласования полных сопротивлений обеих частей подбирают компромисс между толщиной ЖК и его проводимостью. Для увеличения динамического диапазона изменений полного сопротивления ФП следует выбирать ФП максимально доступной толщины и снижать частоту питающего напряжения. Однако, с другой стороны, при питании ЖК ячейки переменным напряжением скорость электрооптического отклика в пределах, полагаемых амплитудой поля и вязко-упругими параметрами ЖК, очевидно, растет с частотой питающего напряжения вплоть до частот, период которых достигает времени диэлектрической релаксации ЖК. Поэтому выбирают частоту, обеспечивающую нужное быстродействие и лежащую при этом в области максимальной чувствительности структуры.

6.5. Заключительные замечания

Жидкие кристаллы были, быть может, первым примером самоорганизующихся наноструктур. ЖК среди подобных структур первыми нашли действительно широкое практическое применение. Физические механизмы действия ЖК приборов не связаны явно с квантовыми эффектами, но можно предположить, что результатом дальнейших исследований станут новые системы, где эффекты самоорганизации ЖК будут помогать упорядочиванию внедренных в них наноразмерных частиц, где квантовые эффекты играют заметную роль. В частности, некоторые смектические мезофазы могли бы выполнять функцию матриц, формирующих двумерные и трехмерные структуры с регулярным расположением нанокристаллов полупроводников или макромолекул с нужными оптическими свойствами (фотонные кристаллы). Такие подходы могут открыть новые возможности управления оптическим излучением в недорогой аппаратуре массового применения, - так же, как ныне широко применяют ЖК в видеотехнике, в дисплеях и индикаторах радиоэлектронной аппаратуры.

7. ОБРАЩЕНИЕ ВОЛНОВОГО ФРОНТА В ФОТОРЕФРАКТИВНЫХ КРИСТАЛЛАХ

7.1. Введение

Что такое обращение волнового фронта и что такое обращенная волна? Прежде, чем дать формальные определения, рассмотрим пример, где на практике могут принести пользу и свойства обращенной волны, и сам метод обращения волнового фронта.

Для сильного разогрева мишени размером менее миллиметра лазерным излучением его необходимо как можно лучше сфокусировать, чтобы весь свет попал на поверхность мишени. Для разогрева до порога термоядерных реакций нужную энергию излучения лазера невозможно получить без применения мощных усилителей. Но в усилителях волновой фронт излучения задающего генераторалазера искажается настолько, что эффективно сфокусировать его на заданной мишени невозможно. Причина искажений волнового фронта в усилителе – пространственные неоднородности накачки, усиления и температуры. Вдобавок эти неоднородности изменяются во времени. В результате пятно фокусировки вместо сотни микрон расплывается на миллиметры и более, а дополнительная оптика не помогает.

Альтернатива простейшей схеме показана на рис. 7.1, *а*. Лазерлокатор с небольшой энергией облучает мишень; рассеянное на ней излучение проходит через усилитель бегущей волны к устройству для *обращения волнового фронта*. Назначение этого устройства можно описать (пока примитивно и нестрого) как отражение каждого падающего на него светового луча точно назад на его же трассу. Тогда, вследствие обратимости световых волн во времени, отраженная (*обращенияя*) волна на обратном пути, проходя по тем же самым участкам в объеме усиливающей среды, окажется сфокусированной на поверхности мишени; все неоднородные набеги фазы в обращенной волне будут скомпенсированы.

Теоретически так же можно фокусировать излучение не только на малой, но и на удаленной мишени, попавшей в луч лазералокатора; при этом, конечно, будут компенсироваться фазовые искажения волнового фронта и в усилителе, и в турбулентной атмосфере (рис. 7.1, δ).



Рис. 7.1. Обращение волнового фронта: *а* - облучение малой мишени; *б* - облучение удаленной цели

Если устройство для обращения волны удается сделать эффективным, то иногда можно обойтись и без лазера-локатора; спонтанное излучение усилительного элемента, рассеянное на мишени, само формирует исходную волну, которая после усиления, обращения и повторного усиления фокусируется точно на освещенной поверхности мишени (рис. 7.1, *в*). По существу получается лазер с диффузным отражателем вместо одного из зеркал.

К этим примерам примыкает еще один, родственный по физической сути. Хорошо известно, что обращение волнового фронта – неотъемлемое свойство голограммы. Любая голограмма – это снимок картины полос, возникающих при интерференции предметной и опорной волн. По существу она представляет собой дифракционную решетку, но не регулярную, как в спектрометрах, а очень сложную.



 Рис. 7.2. Обращение волнового фронта при восстановлении голограммы *in situ: a* - запись попутными волнами, *δ* - восстановление встречной волной,
 в - запись встречными волнами, *г* - восстановление попутной волной

Некоторые лазеры дают излучение с очень плохим качеством волнового фронта из-за специфических оптических свойств их активной среды. Чтобы выправить волновой фронт, применяют фазокорректирующие пластинки, сформированные излучением самого лазера методами голографии (рис. 7.2). Так как предметная и опорная волны принципиально равноправны, то восстановление голограммы, проявленной непосредственно на месте съемки, излучением лазера как опорной волной, даст предметную волну; ее же при съемке пространственной фильтрацией формируют как предметную волну с хорошей формой волнового фронта, близкой к плоской или сферической. Более того, если полученную при восстановлении "хорошую" волну обратить, то есть отразить при помощи плоского зеркала строго назад, то она, в соответствии со свойствами голограммы, будет частично рассеяна назад точно в лазер. Эта рассеянная волна по отношению к исходному излучению лазера будет обращенной. Возвратив часть излучения назад в лазер, мы снижаем потери, а поэтому мощность излучения лазера возрастет, и при этом сохранится хорошее качество волнового фронта.

7.2. Обращение волнового фронта. Сопряженная (обращенная) волна

Теперь уточним основные понятия. С физической точки зрения обращённая волна – это такая волна, которая распространяется в обратном направлении по отношению к исходной волне, в каждом ее сечении имеет такое же (подобное) распределение амплитуд, а распределение фаз отличается только знаком. С математической точки зрения операция обращения волнового фронта эквивалентна операции комплексного сопряжения пространственной части волны,

$$\boldsymbol{E}_{i} = \boldsymbol{E}_{i} \exp[i(\boldsymbol{\omega}t - \boldsymbol{k}_{i}\boldsymbol{r})]; \quad \boldsymbol{E}_{i}^{*} = \boldsymbol{E}_{i} \exp[i(\boldsymbol{\omega}t + \boldsymbol{k}_{i}\boldsymbol{r})]; \quad i = 1, 2, 3$$

Поясним это на простых примерах.

Плоская волна. Пусть слева (рис. 7.3, *a*) вдоль оси z падает плоская волна

$$\boldsymbol{E} = \boldsymbol{E}_0 \exp[i(\boldsymbol{\omega} t - \boldsymbol{k} \boldsymbol{r})].$$

Поставив плоское зеркало ортогонально направлению волны, можно отразить ее точно в обратном направлении. Отражённая волна, как видно из рис. 7.3, *а*, будет обращённой по отношению к падающей.

Сферическая расходящаяся волна. Пусть из точечного источника S выходит сферическая волна

 $E = (E_0/r) \exp \left[i(\omega t - kr)\right]$

(рис. 7.3, δ). Поставив сферическое зеркало так, что его центр кривизны совпадает с точкой S, опять отразим падающую волну точно назад. При этом отражённая волна будет обращённой по отношению к падающей.

Сферическая сходящаяся волна. Здесь допустимы два варианта; первый сводится к предыдущему, если точку схождения волны (фокус) считать виртуальным источником расходящейся волны (рис. 7.3, в). Второй вариант, очевидно, приводит к установке выпуклого зеркала так, что центр его кривизны совпадет теперь с виртуальным источником (рис. 7.3, г). Отметим, что при обращении сферической *сходящейся* волны у нас гораздо шире возможности выбора кривизны зеркал.



Рис. 7.3. Обращение фронта: *a* - плоской волны, *б* - расходящейся сферической волны, *в* - сходящейся сферической волны

Получить обращение волнового фронта такими простыми средствами можно потому, что исходные волновые фронты имеют правильную геометрическую форму – плоскую или сферичаскую. Выбрав форму и расположение зеркала, мы добиваемся того, что в каждой точке пространства волна падает на зеркало по нормали. Но что делать, если фронт волны слабо отклоняется от правильной формы? Неоднородная волна. Пусть волновой фронт имеет слабые фазовые возмущения, так что амплитуда волны в каждой точке пространства постоянна (это соответствует ближней зоне относительно оптического элемента, возмущающего волновой фронт, как показано на рис. 7.4, *a*). Понятно, что именно туда, где волна имеет известную форму волнового фронта, нужно поместить зеркало, изогнутое так, чтобы в каждой точке его поверхности волна падала на него по нормали. Иначе говоря, поверхность зеркала должна совпадать с поверхностью равных фаз. Тогда при отражении мы получим волну с обращённым волновым фронтом.



Рис. 7.4. ОВФ со слабыми фазовыми неоднородностями (*a*); принцип устройства адаптивного зеркала (б)

Можно ли управлять формой отражающей поверхности? Да, можно; такие зеркала называют адаптивными (рис. 7.4, б). Берут тонкое, обычно медное зеркало, и к его задней поверхности по определённому закону прикрепляют тяги (Т). Тяги приводят в движение электродвигателями (Эд), – либо пьезоэлектрическими, либо иными.

Очевидно, что предварительно необходимо определять форму волнового фронта падающего излучения. Обычно это делают с помощью датчиков Хартмана–Шака (см. раздел 3, п. 3.4). Для обработки сигналов датчиков и управления тягами в реальном времени нужны мощные ЭВМ. быстродействия на уровне десятков герц. Вся процедура сложна; так удаётся компенсировать только крупномасштабные возмущения волнового фронта при быстродействии до десятков герц. Если же возмущения мелкомасштабные и быстро меняются по времени, то адаптивные зеркала с механическим приводом для этих целей уже непригодны.

Иной вариант – многоэлементное мозаичное зеркало. В настоящее время разработана технология формирования подвижных гибких элементов с микронными размерами на кремниевой пластине методами фотолитографии и травлениями, хорошо освоенными современной микроэлектроникой. Эти структуры можно изготавливать в виде массивов (двумерных решеток), а электронные средства управления ими даже проще, чем в ТFT-панелях, так как они органично сочетаются с технологией кремниевой микроэлектроники. Такие микроэлектромеханические системы (МЭМС, MEMS) при микронных размерах элемента отличаются высоким быстродействием; их резонансные частоты лежат в ультразвуковом диапазоне выше 100 кГц. Правда, подобные адаптивные зеркала могут компенсировать вариации фазы порядка длины волны, не более.

Еще один вариант – применение мозаик (решеток) с управляемым коэффициентом преломления вблизи от поверхности зеркала. Легче всего это делать с применением жидких кристаллов. Они могут компенсировать вариации фазы порядка десятка длин волн. По быстродействию эти структуры на порядок лучше гибких зеркал, но сильно проигрывают в сравнении с МЭМС. Управление ЖК-решетками в принципе не отличается от управления ТFT-панелями современных телевизоров и дисплеев. Результаты новейших исследований и разработок новых ЖК с более хитроумными межмолекулярными и внутримолекулярными взаимодействиями дают надежду на существенное улучшение быстродействия фазовых ЖК-модуляторов.

Более дорогие и менее технологичные, но существенно более быстрые управляемые фазовые решетки можно строить на основе отражающих мозаик из пьезоэлементов. Электрооптические кристаллы здесь применять трудно потому, что для создания фазового сдвига на полволны нужны кристаллы толщиной порядка миллиметров даже при управляющих напряжениях порядка киловольтов. Поэтому и световой поперечник одного элемента мозаики (пикселя) не может быть менее нескольких миллиметров.

Из всего сказанного в данном разделе можно заключить, что для обращения волнового фронта необходимы технические устройства,

по сути представляющие собой пространственные фазовые модуляторы излучения с микронными размерами пикселя.



Рис. 7.5. Эволюции первоначально плоской волны на различных этапах: *a* - после фазоискажающего участка трассы, *б* - после обращения фронта в идеальном устройстве,

в - после обратного прохода через фазоискажающий участок

7.3. Восстановление голограммы сопряженной опорной волной

Обычная голограмма на фотоэмульсии (рис. 7.2) при ее восстановлении *in situ* (на том же месте) волной E_2 , сопряженной с опорной волной E_1 при записи, формирует волну, сопряженную с записанной предметной (сигнальной) волной Е₃. Голограмма, записанная в геометрии Лейта-Упатниекса (рис. 7.2, *a*) при ее восстановлении (считывании) рассеивает опорную волну так, что за голограммой (по ходу опорной волны) фронт рассеянного излучения точно соответствует фронту предметной волны при записи. Этот фронт, как систему расходящихся пучков, глаз воспринимает как мнимое изображение (фантом) предмета съемки (рис. 7.2, б). Если же при восстановлении на голограмму направить волну, сопряженную с опорной волной, то и восстановленная предметная волна изменит направление; ее волновой фронт будет соответствовать системе сходящихся пучков (их глаз не воспринимает как изображение), и они сформируют действительное изображение фантомного предмета съемки. Далее мы подробнее поясним происходящие при этом явления

Направим точно навстречу волне E_1 еще одну плоскую волну, E_2 . (рис. 7.2, рис. 7.6). Теперь волна E_2 будет отражаться от синусоидальных "штрихов" решетки (когерентно рассеиваясь на градиенте показателя преломления), и отраженная плоская волна E_4 будет направлена точно навстречу плоской волне E_3 . Иначе говоря, волна E_2 отражается от плоскости "штриха" решетки в точности под углом блеска. Преобразование волны E_2 в волну E_4 эффективно, потому что в описанной геометрии выполнено условие Вульфа-Брэгга, и амплитуды отраженных парциальных плоских волн суммируются строго синфазно. Так на фазовой решетке, созданной интерференцией двух плоских волн E_1 , E_3 (они задают конфигурацию решетки) происходит обращение фронта плоской волны E_3 .



Рис. 7.6. Принципиальная геометрическая схема обращения волнового фронта при четырехволновом взаимодействии

Если волна E_3 не плоская, то ее придется разложить в интеграл Фурье по плоским волнам, и тогда тот же самый результат $E_4 \propto E_3^*$ мы получим для каждой плоской компоненты разложения. Последнее утверждение можно доказать строго математически, но мы ограничимся интуитивными доводами.

Если волна E_1 и парциальная фурье-компонента волны E_3 формируют решетку, то другая парциальная волна E_3 совместно с волной E_1 формируют такую решетку, на которой волна E_2 не может эффективно преобразовываться в волну, сопряженную с первой парциальной волной из-за нарушения условий Вульфа–Брэгта. Так как обе парциальных волны по сути равноправны, то результат сводится к тому, что каждая фурье-компонента сигнальной волны E_3 обращается только на "своей" решетке, а "чужая" решетка для нее почти "прозрачна". Если решетку формируют плоская волна E_1 и расходящаяся волна E_3 со сферическим фронтом, то сечение интерференционной картины плоскостью, ортогональной волновому вектору k, будет выглядеть как совокупность концентрических кольцевых полос. По отношению к плоской волне E_2 решетка такого рода будет работать как фазовая зонная пластинка Френеля, то есть она будет фокусировать свет в центр, от которого расходится первичная сферическая волна E_3 . Так происходит обращение волнового фронта сферической волны.

До сих пор мы различали волны E_1 и E_2 ; на самом деле они в общем случае равноправны. Поэтому фазовую решетку формирует не только интерференция волн E_3 и E_1 , но и интерференция волн E_3 и E_2 (рис. 7.6). Рассеяние волны E_1 на этой второй решетке тоже порождает обращенную волну $E_4 \sim E_3^*$. Более того, волны E_4 , порожденные на двух разных решетках, интерферируют между собой.

В условиях типового эксперимента волна E_1 все-таки выделена тем, что угол θ_{13} невелик, то есть волны E_1 , E_3 бегут в одну сторону (попутны) и поэтому формируют *редкую* решетку с относительно большим периодом,

$$\Lambda_{13} = \lambda / 2\sin(\theta_{13} / 2) \approx \lambda / \theta_{13}, \ \theta_{13} \ll 1.$$

а почти встречные волны E_2 , E_3 формируют частую решетку с малым периодом,

$$\Lambda_{23} = \lambda / 2 \sin(\theta_{23} / 2) \approx \lambda / 2, \ \theta_{23} \approx \pi.$$

Частые решетки не всегда можно формировать и использовать достаточно эффективно; препятствуют ограниченное пространственное разрешение фоточувствительных материалов и механические нестабильности оптики, в первую очередь при записи голограммы. При этом повышению чувствительности эмульсии обычно сопутствует снижение ее разрешающей способности, так что ослабление влияния механических нестабильностей за счет уменьшения времени съемки голограммы «компенсируется» ухудшением статического пространственного разрешения.

Голограмма на фотоэмульсии первоначально представляет собой *амплитудную* решетку с переменным пропусканием. Обычно голограмма после проявления проходит специальную химическую обработку (так наз. отбеливание-задубливание). В результате неравномерное почернение фотоэмульсии преобразуется в неравномерную плотность желатиновой основы эмульсии, то есть в неравномерное распределение показателя преломления. Так получают *фазовую* решетку со светопропусканием несколько менее 100%. Иногда используют толстослойные эмульсии (голограммы Денисюка), где фазовая решетка становится объемной. Она преобразует в фантомное изображение существенно большую часть мощности восстанавливающей волны в сравнении с двумерной решеткой.

Эффективность рассеяния опорной волны в восстановленную предметную волну отображает *дифракционная* эффективность, то есть отношение световых потоков в восстановленной волне и в опорной волне на этапе восстановления голограммы. Оценки дифракционной эффективности дают две модели, описывающие два предельных случая.

Для голограмм на тонких слоях эмульсии применима модель рассеяния Рамана-Ната, подробнее описанная в разделе "*Акустооптическая модуляция*". Для объемных голограмм следует применять описанную там же модель рассеяния Брэгга. В обоих случаях результаты получатся аналогичными: дифракционная эффективность монотонно и почти линейно растет при увеличении индекса фазовой модуляции от нуля до единицы. В тонкослойных голограммах индекс модуляции обычно гораздо менее единицы. В толстослойных голограммах индекс модуляции может быть сравним с единицей; увеличивать его сверх единицы нет практического смысла, так как коэффициент рассеяния в случае Брэгга зависит от индекса модуляции не линейно, а периодически (этот результат дает модель связанных волн, кратко описанная в приложении 3).

Трехмерная решетка в сравнении с двухмерной способна рассеивать назад гораздо большую часть светового потока, поэтому голограмму Денисюка можно рассматривать и в проходящем, и в отраженном свете.

7.4. Динамическая голография

Точность формирования обращенной волны и ее интенсивность в классической голографии были удовлетворительными даже для интерферометрии, но необходимость длительной мокрой химической обработки фотопленки непосредственно на месте ее установки, а также невозможность быстрого обновления снимка полностью исключали применение фотоэмульсий для адаптивного ОВФ в реальном времени. Более пригодными казались фотохромные стекла и фотоситаллы, но и в них для стирания записанной картины нужен длительный нагрев. Не подошли и фотопроводящие термопластики, для которых необходимы разделенные во времени подготовка, термическое проявление решетки и ее стирание перед повторной записью. Еще более новые и очень эффективные среды для оптической записи и хранения информации (например, бактериородопсин) по ряду причин тоже не подходят для динамического ОВФ.

Предпочтителен такой материал голограммы, в котором фазовая решетка формируется *динамически*, изменяясь вместе с интерференционной картиной. Поглощение в этом материале должно быть небольшим, чтобы легче было получить объемную решетку с большой дифракционной эффективностью. Все это означает, что в материале голограммы в идеальном случае показатель преломления должен изменяться пропорционально локальной интенсивности излучения. Чем быстрее реакция материала и релаксация в нем, тем более быстрые флуктуации волнового фронта можно компенсировать.

Следовательно, для обращения волнового фронта нужно формировать объемную фазовую динамическую голограмму обращаемой волны в материале с нелинейной восприимчивостью; показатель преломления материала должен зависеть от напряженности поля (в крайнем случае – от интенсивности) действующего на него света. Эти пожелания привлекли внимание к идеям нелинейной оптики, изучающей свойства и применения подобных материалов.

7.5. Четырехволновое взаимодействие в среде с нелинейной восприимчивостью

До появления лазеров физики и инженеры имели дело со световыми потоками относительно небольшой интенсивности, для которых коэффициент преломления оптического материала с высокой точностью можно было считать постоянной, не зависящей от присутствия света. В этой области малых напряженностей поля света восприимчивость среды называют *линейной* в том смысле, что связанная с ней поляризация среды пропорциональна напряженности поля. Потом у физиков появились лазеры, способные формировать поля излучения с напряженностями, сравнимыми с напряженностями полей химических связей в молекулах и в твердых телах. Эксперименты с сильными световыми полями показали, что поляризация вещества под действием поля световой волны в общем случае зависит от ее напряженности, а не зависящая от напряженности поляризуемость (восприимчивость) – лишь только первое слагаемое разложения восприимчивости в ряд по степеням напряженности. Тогда же появилось и название нового раздела физики – нелинейная оптика.

В начале 70-х годов стало понятно, как можно применять различные эффекты нелинейной оптики для формирования волны с обращённым волновым фронтом в реальном времени, когда фазокорректирующая самоадаптирующаяся структура формируется как динамическая голограмма, без обслуживающих ее средств измерения и управления. Наиболее важными оказались взаимодействия, связанные с нелинейной восприимчивостью третьего порядка, когда *поляризация* среды пропорциональна *третьей* степени напряженности поля (*кубичная*). При их использовании нет необходимости определять форму волнового фронта падающего излучения, и можно обратить волну практически с любым масштабом неоднородностей. Эти достоинства и привлекли внимание физиков к поиску материалов с подходящими нелинейностями.

Характерные времена изменения восприимчивости в различных средах меняются от пикосекунд (10⁻¹² с) до десятков или даже сотен секунд. Сразу же подчеркнем, что физические процессы, приводящие к изменению восприимчивости, очень многообразны, и становятся заметными они в разных материалах при разных уровнях напряженности поля.

Рассмотрим следствия, к которым приводит наличие у поляризации среды кубичной составляющей.

Направим на нелинейную среду три одинаково поляризованных волны (рис. 7.6):

$$\boldsymbol{E}_{i} = \boldsymbol{E}_{i} \exp[i(\omega t - \boldsymbol{k}_{i} \boldsymbol{r})] + \kappa.c.; \ i = 1, 2, 3.$$
(7.1)

Здесь E_1, E_2 – опорные волны; E_3 – сигнальная волна; *к.с.* – комплексно – сопряжённое выражение. В большинстве экспериментов используют плоские опорные волны, для волновых векторов которых выполнено соотношение

$$\boldsymbol{k}_1 + \boldsymbol{k}_2 = 0, \implies \boldsymbol{k}_1 = -\boldsymbol{k}_2$$

В самом общем виде диэлектрическую восприимчивость среды описывает тензор. Мы для простоты рассмотрим нелинейную вос-

приимчивость *изотропной* среды, где нелинейная восприимчивость сводится к скаляру. В этом случае поляризация принимает вид:

$$\boldsymbol{P} = \varepsilon_0 \alpha \boldsymbol{E} + \varepsilon_0 \chi^{(3)} \boldsymbol{E}_1 \boldsymbol{E}_2 \boldsymbol{E}_3 = \boldsymbol{P}^{(L)} + \boldsymbol{P}^{(NL)}$$
(7.2)

где α – линейная, а $\chi^{(3)}$ – нелинейная (кубическая) поляризуемость среды. Квадратичная поляризуемость $\chi^{(2)}$ изотропной среды равна нулю из-за требований симметрии, как и все $\chi^{(i)}$ с четными i, а слагаемые высших порядков с i > 3 опущены как малые.

Подставив выражения для полей (7.1) в нелинейную часть поляризации (7.2), с учётом приведённого выше условия получим

 $\boldsymbol{P}^{(\mathrm{NL})}(\boldsymbol{\omega}) = \varepsilon_0 \boldsymbol{\chi}^{(3)} \left| \boldsymbol{E}_1 \boldsymbol{E}_2 \right| \boldsymbol{E}_3^* \exp[i(\boldsymbol{\omega} t + \boldsymbol{k}_3 \boldsymbol{r})] + \kappa.c.$

Взаимодействие, при котором в среде с нелинейной восприимчивостью две падающих на нее волны порождают третью волну, называют *трехволновым* взаимодействием (или *смешением*, что традиционно, хотя и не очень аккуратно). Если три волны порождают четвертую, взаимодействие называют *четырехволновым* (*смешением*). Когда частоты трех исходных волн совпадают, ту же частоту имеет и порожденная ими четвертая волна; это взаимодействие называют *вырожденным четырехволновым смешением* (ВЧВС). Если частоты падающих волн различаются, то взаимодействие называют невырожденным, и частоту генерируемой четвертой волны определяет сохранение энергии.

При нелинейном взаимодействии параметры волн взаимосвязаны условиями фазового синхронизма колебаний в пространстве и во времени. Выполнение условий пространственного синхронизма принимает форму, аналогичную закону сохранения импульса (это строго верно в среде без дисперсии, а также и при вырожденном взаимодействии). Требования временного синхронизма приводят к соотношению, аналогичному сохранению энергии. Сохранение энергии и импульса при столкновениях частиц в самом общем случае в квантовой механике возникает как условия синхронизма (когерентности) временной и пространственной частей волновых функций начального и конечного состояний. При этих условиях сечение рассеяния (квадрат модуля амплитуды рассеяния) имеет резкий максимум. В частности, если в начальном состоянии строго определены энергия и импульс системы, то зависимость сечения рассеяния от импульса и энергии конечного состояния превращается в произведение δ-функций от разности энергий и импульсов. Если же неопределенности энергии и импульса начального состояния заданы соответствующими временем и длиной когерентности, то именно эти показатели когерентности и будут задавать спектральную и пространственно-частотную ширину максимума в сечении рассеяния. В среде с дисперсией сохранение импульса не нарушается, если импульсом называть вектор, пропорциональный волновому вектору, $p = \hbar k$.

7.6. Формирование фазовой решетки в среде с кубической нелинейностью

Определив зависимость *поляризации* изотропной среды от напряженности поля световой волны E, получим зависимость показателя преломления n = n(E). Для этого подставим поляризацию в выражение вектора электрической индукции D:

$$D = \varepsilon_0 E + P = \varepsilon_0 E + \varepsilon_0 \alpha E + \varepsilon_0 \chi^{(3)} E^{(3)}$$
$$D = \varepsilon_0 (1 + \alpha + \chi^{(3)} |E|^2) E = \varepsilon_0 \varepsilon E$$

С точностью до слагаемых пятой и более высоких степеней

$$n(E) = \sqrt{\varepsilon(E)} = \sqrt{\frac{D}{E}} = \sqrt{\frac{\left(\varepsilon_0 E + \varepsilon_0 \chi^{(1)} E + \varepsilon_0 \chi^{(3)} E^3\right)}{E}} = \sqrt{\varepsilon_0 \left(1 + \chi^{(1)} + \chi^{(3)} E^2\right)} \equiv \sqrt{n_0^2 + \chi^{(3)} E^2} \approx$$
$$\approx n_0 \left(1 + \frac{\chi^{(3)} E^2}{2n_0^2}\right) \equiv n_0 + n_2 E^2.$$

Отсюда видно, что зависящая от поля добавка к показателю преломления пропорциональна интенсивности поля, и она тем больше, чем больше коэффициент $\chi^{(3)}$, который обычно называют кубической (кубичной) нелинейностью. Следовательно, в среде с кубичной нелинейностью пространственная неоднородность интенсивности поля излучения создает подобную же неоднородность показателя преломления. На этом, в частности, основано и формирование динамических фазовых голограмм.

Отметим особую роль кубичной нелинейности: вообще говоря, в изотропных средах $\chi^{(3)}$ – первое отличное от нуля слагаемое в разложении *нелинейной части* поляризации по степеням напряженности поля. Слагаемое низшей степени, с $\chi^{(2)}$, не равно нулю только в анизотропной среде, а $\chi^{(1)}$ отображает обычную линейную восприимчивость. При всем этом, электростатическое поле в изотропной среде в общем случае порождает малую *наведенную* анизотропию именно за счет кубичной нелинейности.

Чтобы пояснить, как в среде с нелинейной восприимчивостью образуется динамическая голограмма, мы рассмотрим самую примитивную модель – запись плоской предметной волны с помощью плоской опорной волны. В реальных голограммах радиус кривизны фронта предметной волны на порядки больше длины волны, и поэтому на каждом участке голограммы интерферируют почти плоские волны; поэтому более сложные случаи просто сводятся к примитивному.

Если в среде действуют две плоских волны на одной частоте,

$$\boldsymbol{E}_{1} = \boldsymbol{E}_{1} \exp[i(\boldsymbol{\omega}t - \boldsymbol{k}_{1}\boldsymbol{r})] + \kappa.c.,$$

$$\boldsymbol{E}_{3} = \boldsymbol{E}_{3} \exp[i(\boldsymbol{\omega}t - \boldsymbol{k}_{3}\boldsymbol{r})] + \kappa.c.,$$

и угол между их волновыми векторами равен θ , то интенсивность суммарного поля $E = E_1 + E_3$ равна

$$I = |\mathbf{E}_{1} + \mathbf{E}_{3}|^{2} = |\mathbf{E}_{1}|^{2} + |\mathbf{E}_{3}|^{2} + (\mathbf{E}_{1}\mathbf{E}_{3}^{*} + \mathbf{E}_{1}^{*}\mathbf{E}_{3}) =$$

= $I_{1} + I_{3} + E_{1}E_{3} \{\exp[-i(\mathbf{k}_{1} - \mathbf{k}_{3})\mathbf{r}] + \exp[i(\mathbf{k}_{1} - \mathbf{k}_{3})\mathbf{r}] \} =$
= $I_{1} + I_{3} + 2E_{1}E_{3}\cos(\mathbf{k}_{1} - \mathbf{k}_{3})\mathbf{r}$

Направим ось координат *z* по биссектрисе угла θ ,

$$k_1 r = q_1 r_{\perp} + k_{1z} z;$$
 $k_3 r = q_3 r_{\perp} + k_{3z} z;$
 $q_1 = -q_3;$ $q = k \sin(\theta/2);$ $k_{1z} = k_{3z}$

где r_{\perp} – координата в плоскости волновых векторов k_1 , k_3 , выбранная ортогонально биссектрисе угла θ . Распределение интенсивности (интерференционная картина!) в этом представлении имеет вид интерференционных полос с периодом Λ ,

$$I = I_1 + I_3 + 2E_1E_3\cos(2qr_{\perp}),$$

где $q = k \sin(\theta/2), \Lambda = \pi/k \sin(\theta/2) = \lambda/2 \sin(\theta/2) = \lambda_0/2 n_0 \sin(\theta/2).$

Отсюда получим распределение показателя преломления в нелинейной среде,

$$n = n_0 + n_2 |\mathbf{E}|^2 =$$

= $n_0 + n_2 (|\mathbf{E}_1|^2 + |\mathbf{E}_3|^2) + 2n_2 E_1 E_3 \cos 2qr_\perp =$
= $\overline{n}_0 + 2n_2 E_1 E_3 \cos 2qr_\perp$,

то есть синусоидальную фазовую дифракционную решетку с периодом $\Lambda = 2k\sin(\theta/2)$. В ней поверхности равных показателей преломления ("штрихи", или рассеивающие плоскости в объемной решетке) параллельны оси *z* и ортогональны плоскости, содержащей волновые векторы k_1 , k_3 .

7.7. Обращение волнового фронта при вырожденном четырехволновом взаимодействии

Само по себе когерентное рассеяние на синусоидальной решетке показателя преломления подчинено почти тем же закономерностям, что и рассеяние на ультразвуковой волне в акустооптическом модуляторе (раздел «*Акустооптическая модуляция*» ; аналогия была бы полной при стоячей волне ультразвука, когда решетка неподвижна). Амплитуда рассеяния волны $E_2(k_2)$ на фазовой решетке с размером *l* в направлении k_2 зависит от величины

$$\beta \equiv k_2 l \times n_2 E_1 E_3$$

аналогичной *индексу фазовой модуляции* (только не во времени, а в пространстве, по координате r_{\perp}). Квадрат амплитуды рассеяния называют дифракционной эффективностью решетки.

Диаграмму направленности рассеяния определяют условия Брэгга; в данном конкретном случае, при $k_1 = -k_2$, они дают

$$\boldsymbol{E}_4(\boldsymbol{r}_\perp) \propto \boldsymbol{E}_3^*(\boldsymbol{r}_\perp).$$

До сих пор мы различали волны E_1 и E_2 ; на самом деле они в общем случае равноправны. Поэтому фазовую решетку формирует не только интерференция волн E_3 и E_1 , но и интерференция волн E_3 и E_2 (см. рис. 7.2, рис. 7.6). Рассеяние волны E_1 на этой второй решетке тоже порождает обращенную волну, $E_4 \propto E_3^*$. Более того, волны E_4 , порожденные на двух разных решетках, интерферируют между собой.

В условиях типового эксперимента волна E_1 все же выделена тем, что угол θ_{13} невелик, то есть волны E_1 , E_3 бегут в одну сторону (попутны) и поэтому формируют *редкую* решетку с относительно большим периодом,

$$\Lambda_{13} = \lambda / 2\sin(\theta_{13} / 2) \approx \lambda / \theta_{13}, \ \theta_{13} \ll 1,$$

а почти встречные волны E_2 , E_3 формируют частую решетку с малым периодом,

$$\Lambda_{23} = \lambda / 2 \sin(\theta_{23} / 2) \approx \lambda / 2, \ \theta_{23} \approx \pi.$$



 Рис. 7.7. Обращение волнового фронта при вырожденном четырехволновом взаимодействии:
 a - на «редкой» решетке, образованной попутными волнами;
 б - на «частой» решетке, образованной встречными волнами;
 в – система координат, направления волн, линейные и угловые величины, использованные при описании решетки показателя преломления

Частые решетки не всегда можно формировать и использовать достаточно эффективно; препятствуют ограниченное пространственное разрешение фоточувствительных материалов и механические нестабильности оптики, в первую очередь при записи голограммы. Более чувствительные эмульсии обычно имеют пониженную разрешающую способность (более крупное зерно), так что за ослабление влияния механических нестабильностей (при уменьшении времени съемки голограммы) расплачиваются ухудшением статического пространственного разрешения. Если нелинейная среда – разреженный газ, то тепловое движение молекул разрушает частую решетку, когда длина свободного пробега явно более половины длины волны света. Обычно в четырехволновом взаимодействии более существенную роль играет редкая решетка. Когерентное рассеяние волны E_2 на решетке с выполнением условий Брэгга лаконично отображают соотношения волновых векторов (см. рис. 7.7, β):

 $k_1 - k_3 = p = k_4 - k_2$, p = 2q, $k_1 = -k_2$, откуда следует $k_4 = -k_3$. Вектор p – вектор *обратной решетки*, соответствующей реальной решетке с периодом $\Lambda = 2k \sin(\theta/2)$.

Теперь можно кратко описать сущность физической модели обращения волнового фронта при вырожденном четырехволновом взаимодействии. Сигнальная волна при интерференции с одной из волн накачки формирует в нелинейной среде решетку показателя преломления. Вторая волна накачки (сопряженная с первой) при когерентном рассеянии на этой решетке порождает четвертую волну, сопряженную с сигнальной.

Чтобы на практике получить такой результат, нужно найти материал с подходящей и не слишком слабой нелинейностью. Один из таких материалов, описанный в следующем разделе, применен в лабораторной установке.

7.8. Диффузионная нелинейность фоторефрактивных кристаллов. Диффузионная фоторефракция

В современной оптоэлектронике часто применяют так называемые фоторефрактивные кристаллы. В них под действием света происходят изменения, в определенных условиях наблюдаемые как изменения показателя преломления, что и дало им название. Зависимость показателя преломления от освещенности можно наблюдать в различных фототропных материалах, в том числе и в фотохромных стеклах.

Описанные далее фоторефрактивные эффекты наблюдают в кристаллах, где явно выражен линейный электрооптический эффект Поккельса. В фоторефрактивном кристалле существенные роли играют подвижные и локализованные заряженные состояния. Качественно опишем в общих чертах процессы, формирующие фазовую решетку, предполагая у читателя минимум знаний по физике полупроводников.

Кристалл BSO (6Bi₂O₃·SiO₂) – полупроводник с большой шириной запрещенной зоны (на это указывает заметное поглощение в сине-фиолетовой области спектра). Из-за большой ширины запрещенной зоны $E_g >> kT$ при комнатных температурах концентрация

электронов проводимости крайне мала. В глубине запрещенной зоны расположены энергетические уровни заряженных локализованных состояний (рис. 7.8); подвижность электронов в этих состояниях на несколько порядков ниже, чем подвижность электронов в зоне проводимости. В этих состояниях внутри запрещенной зоны время жизни электронов довольно велико, поэтому их называют ловушками (или центрами прилипания).

В красной области спектра BSO поглощает на переходах с глубоких донорных уровней в зону проводимости. Концентрация доноров невелика, поэтому поглощение не очень сильное (табл. 7.1). Позже мы покажем, что в принципе существует оптимум поглощения, при котором желательные эффекты проявляются наиболее сильно.

Таблица 7.1

Параметр	BSO	BTO	Fe:LiNbO ₃
Симметрия кристалла	23	23	3m
Показатель преломления (0,63 мкм)	2,54	2,56	2,29 (n_o) 2,20 (n_e)
Поглощение (0,63 мкм), см ⁻¹	0,3	0,5	
Электрооптический коэффициент r ₄₁ , пм/В	5,0	5,7	0
Электрооптический коэффициент r_{33} , пм/В	0	0	32
Электрооптический коэффициент r ₃₁ , пм/В	0	0	10
Электрооптический коэффициент r ₂₂ , пм/В	0	0	6,8
Полуволновое напряжение, кВ	3,9	3,3	
Естественное удельное вращение, град/мм	22	6,3	нет
Диэлектрическая проницаемость (статич.)	56	46	$\frac{85 (\epsilon_{11})}{30 (\epsilon_{33})}$

Электрооптические параметры фоторефрактивных кристаллов

При локальном освещении полупроводника (независимо от механизма генерации в нем неравновесных подвижных электронов и участвующих в этом состояний), обязательно образуются пары носителей заряда разного знака. В освещенной области концентрация носителей заряда повышается, и начинается их диффузия в ближайшие к ней неосвещенные области.

Диффузия – это движение частиц с их тепловыми скоростями (по инерции). Если газ частиц разрежен, и нет стенок, то частицы разлетаются неограниченно далеко. В реальных условиях свободный разлет частиц замедляют столкновения с другими частицами. Электростатическое поле пространственного заряда сдерживает разлет зарядов разного знака, но ускоряет разлет одноименных зарядов.

Диффузию неравновесных электронов в полупроводнике в первом приближении можно рассматривать как диффузию в разреженном газе, учитывая столкновения (рассеяние) и конечное время жизни неравновесного электрона в зоне проводимости. Вместе с тем, часто уже в первом приближении нужно учитывать присутствие в полупроводнике зарядов обоих знаков и их кулоновское взаимодействие.



Рис. 7.8. Процессы генерации и рекомбинации носителей заряда в пространстве импульс-энергия

Если подвижности носителей заряда различаются, то диффузия более подвижных носителей обгоняет диффузию менее подвижных. Тогда заряды разного знака разделяются в пространстве, и образуется электростатическое поле (рис. 7.9). Это явление в полупроводниках называют эффектом Дембера. Поле разделившихся при диффузии разноименных зарядов сдерживает движение более подвижных носителей и ускоряет носители менее подвижные. При стационарном освещении поле разделившихся зарядов вызывает их дрейф по полю, так что токи диффузии и дрейфа, направленные навстречу друг другу, компенсируются.



Рис. 7.9. Процессы переноса зарядов в пространстве координата *х* – энергия *E*

В кристалле BSO подвижностью положительных зарядов можно пренебречь. Тогда диффузионный разлет зарядов из освещенной области описывают два наиболее важных параметра – коэффициент диффузии D и время жизни электрона т в зоне проводимости. Можно показать, что линейный размер области L_D , в которой при диффузии из точечного источника концентрация электронов существенно убывает за счет рекомбинации, составляет

$$L_D = \sqrt{D\tau}$$
;

эту величину называют *длиной диффузии* (или диффузионной длиной).

Время жизни неравновесных электронов проводимости ограничено их переходами в связанные состояния на ловушкахакцепторах, внутри запрещенной зоны, ближе к зоне проводимости. Время жизни этих связанных состояний иногда может быть очень заметным (секунды и более).

Очень важно то, что и доноры, и ловушки – локализованные заряженные состояния. Диффундирующие из освещенной области электроны проводимости, захваченные на ловушки в прилегающей неосвещенной области, образуют область отрицательного заряда на расстояниях порядка диффузионной длины от освещенной области. Плотность пространственного заряда в первом приближении пропорциональна освещенности. В освещенной области формируется положительный пространственный заряд незаселенных доноров. Ионизованный донор на общем фоне электронейтрального полупроводника несет положительный заряд, а занятая ловушка – отрицательный. В области перехода от освещенной к неосвещенной области пространственные заряды создают электростатическое поле. Его напряженность в первом приближении пропорциональна освещенности и обратно пропорциональна диффузионной длине; вектор напряженности направлен так, что поле объемных зарядов тормозит диффузию электронов проводимости из света в тень. Так в стационарных условиях освещения формируется стационарное распределение электрического поля. Так как концентрация собственных носителей заряда в широкозонном полупроводнике исчезающе мала, а примесные уровни энергии в нашем случае расположены в глубине запрещенной зоны, то возникший при неоднородном освещении пространственный заряд не может рассасываться с участием токов проводимости, а его поле в отсутствие свободных носителей не экранируется и релаксирует очень медленно.

Электростатическое поле может изменять оптические свойства вещества за счет электрострикции и за счет линейного электрооптического эффекта Поккельса. Электрострикция проявляется как механические напряжения и вызванные ими изменения плотности и показателя преломления. Электрострикция – эффект очень слабый, квадратично зависящий от напряженности поля, но свойственный всем полупроводникам и диэлектрикам. Напротив, эффект Поккельса линейно зависит от напряженности поля и во многих электрооптических кристаллах достаточно силен. Он проявляется как изменения поляризуемости и показателя преломления под действием электрического поля; при этом механические напряжения и деформации (пьезоэлектричество) играют второстепенную роль.

Кристалл BSO имеет кубическую элементарную ячейку и относится к классу симметрии 23, то есть имеет только оси симметрии 2-го и 3-го порядка. Среди элементов симметрии BSO *нет плоскостей симметрии*, поэтому он существует в левой и правой *киральных* формах (синонимы – *хиральные*, *энантиоморфные*). В кристалле BSO, как и во многих других киральных кристаллах, существует естественная гиротропия (оптическая активность), которая проявляется как вращение плоскости линейной поляризации проходящего излучения. В этом отношении кристалл BSO подобен кристаллу α -кварца (при T < 575°C), у которого существуют левая (L) и правая (D) модификации, различающиеся знаком угла поворота плоскости поляризации. В поликристаллических и даже в блочных кристаллах гиротропия может быть незаметной из-за одновременного присутствия L и D-фаз в равных количествах, что нередко происходит при неконтролируемом росте кристалла и образовании двойников. Киральность – необходимое, но не достаточное условие проявления гиротропии; истинная причина гиротропии – нелокальная зависимость поляризации от напряженности поля в киральной среде.

Среди элементов симметрии BSO *нет инверсии*, поэтому ему, несмотря на отсутствие естественного двойного лучепреломления, свойствен линейный электрооптический эффект Поккельса; он проявляется как двойное лучепреломление, наведенное внешним электрическим полем. В этом отношении BSO похож на более простые кубические кристаллы со структурой цинковой обманки (ZnS, GaAs), где заметен эффект Поккельса; здесь, в частности, нет симметрии по отношению к знаку координаты в направлении [111], т. к. инверсия меняет местами Zn и S.

Совместное действие фотовозбуждения, диффузии электронов проводимости, формирования пространственных зарядов и эффекта Поккельса приводят к тому, что при пространственнонеоднородном освещении в области градиента освещенности показатель преломления материала изменяется, и тем сильнее, чем больше градиент освещенности. Этот результат и называют диффузионным фоторефрактивным эффектом. Подчеркнем, что фоторефрактивный эффект – изменение показателя преломления в области *градиента* освещенности, и самые важные роли в нем играют образование электростатических полей пространственного заряда и электрооптический эффект Поккельса.

Подчеркнем, что грубой ошибкой будет утверждение, будто бы в фоторефрактивном кристалле показатель преломления *явно* зависит от освещенности. На самом деле зависимость показателя преломления непосредственно от равномерной освещенности очень слаба, но заметна зависимость показателя преломления от *градиента* освещенности. Такую взаимосвязь называют *нелокальной*. Более того, не совсем точным будет утверждение, что поляризация в фоторефрактивном кристалле зависит от третьей степени напряженности поля. Корректнее считать, что *поляризуемость* зависит от *квадрата* напряженности поля, причем *усредненного* по времени за достаточно длительный интервал (порядка миллисекунд). Иначе говоря, именно потому, что поляризуемость зависит от интенсивности света, локальная поляризация *«как бы»* зависит от третьей степени напряженности поля. На самом деле добавка (за счет нелинейности) к локальной поляризации в фоторефрактивном кристалле *пропорциональна произведению скалярной интенсивности и вектора напряженности* поля, но это не равносильно третьей степени напряженности. Вектор поляризации в фоторефрактивном кристалле совершает колебания на частоте колебаний напряженности поля, но не на ее кратных частотах.

Из всех предыдущих рассуждений следует, что диффузионный фоторефрактивный эффект – эффект нелокальный и неадиабатический. Нелокальность проявляется в том, что добавку к показателю преломления дает не локальная освещенность сама по себе, но ее градиент. Неадиабатичность проявляется в том, что приращение показателя преломления зависит не от мгновенной (текущей) освещенности, но от предыстории освещения за время порядка времени релаксации связанных объемных зарядов – глубоких доноров и ловушек. Эти времена релаксации могут превышать секунды.

Диффузионная фоторефракция ухудшает работу электрооптических модуляторов на ниобате лития и т.п. материалах при плотности потока излучения свыше единиц мВт/мм². Из-за накопления долгоживущего объемного заряда вдоль трассы светового пучка преломление кристалла искажается. Объемный заряд можно устранить отжигом кристалла, но более радикальное решение – легирование кристалла примесью, через уровни которой ускоряются переходы с акцепторных ловушек на донорные; для LiNbO₃ подходят Mg, Zn и некоторые иные двухвалентные металлы.

С другой стороны, LiNbO₃ используют и как фоторефрактивный материал; тогда его фоторефрактивные свойства преднамеренно усиливают, для чего его легируют, например, железом (см.табл. 7.1). В этом случае примесные центры Fe^{2+} действуют как глубокие доноры, а ионизованные, Fe^{3+} – как ловушки.

7.9. Формирование фазовой решетки в фоторефрактивном кристалле

Если поле освещенности представляет собой периодическое распределение интенсивности (интерференционные полосы), то в фоторефрактивном материале формируется фазовая дифракционная решетка как периодические изменения показателя преломления в пространстве (рис. 7.10). В этой решетке экстремумы показателя преломления смещены по отношению к экстремумам освещенности на полпериода решетки, то есть туда, где максимальны градиенты освещенности.



Рис. 7.10. Распределения интенсивности (*EE**), напряженности поля (*E*), показателя преломления (*n*) в фоторефрактивном кристалле (BSO) при освещении его полосами интерференции двух взаимно когерентных попутных плоских волн (формирование «редкой» решетки)

Когерентное рассеяние света на подобных решетках применяют для динамической голографии, частным случаем которой можно считать и обращение волнового фронта.

Покажем, как можно оценить коэффициент рассеяния на фазовой решетке и скорость ее формирования.

«Стационарную» напряженность электростатического поля в максимуме, $\boldsymbol{\varepsilon}$, дает условие равенства плотностей токов проводимости и диффузии при концентрации свободных электронов N,

$$eN\mu \boldsymbol{\mathcal{E}} = eD_e\nabla N_e$$

вместе с отношением коэффициента диффузии D_e электронов проводимости к их дрейфовой подвижности μ ,

$$\frac{D_e}{\mu} = \frac{k_B T}{e};$$

*k*_B - постоянная Больцмана, *T* - температура, *e* - заряд электрона. Далее,

$$\boldsymbol{\mathcal{E}} = (D_e / \mu) \times (\nabla N / N),$$
$$\nabla N / N = mp,$$

где m – относительная разность интенсивностей интерференционных полос в максимуме и в минимуме, $p = 2k\sin(\theta/2)$ – длина вектора *обратной* решетки распределения интенсивностей поля излучения.

После всех подстановок найдем напряженность поля в ее максимумах (по модулю)

$$\boldsymbol{\varepsilon} = mp \frac{k_B T}{e}$$

Из последней формулы можно сделать вывод, будто бы решетка формируется при любой, сколь угодно малой интенсивности излучения, и напряженность поля объемного заряда не зависит от интенсивности. Это, разумеется, неверно; чтобы результат был правильным, нужно учесть релаксацию заселенностей ловушек и сечение поглощения излучения. Без учета этих факторов при исчезающе малой интенсивности излучения для формирования *статического* распределения объемного заряда нужно бесконечное время. За конечное время при конечной интенсивности излучения формируется решетка с напряженностью поля

$$\boldsymbol{\varepsilon}_{R} \approx mpS\tau_{\mathrm{T}}W\frac{k_{B}T}{e},$$

где S – сечение поглощения света, τ_T - время релаксации пространственного заряда (время релаксации заселенности ловушек), W – средняя по пространству интенсивность излучения, выраженная плотностью потока фотонов.

Такое распределение далее будем называть стационарным. Приведенное выражение для напряженности поля корректно только при небольших изменениях заселенностей ловушек, когда можно пренебречь изменением локального поглощения света в кристалле. Время формирования пространственного заряда во всяком случае не меньше времени релаксации $\tau_{\rm T}$, в реальных условиях зависящего от многих физической факторов, различных по природе.

По теории, стационарное распределение объемных зарядов после мгновенного включения света формируется за время порядка

$$\tau_{\mathbf{\mathcal{E}}} = \tau_M (1 + p^2 L_D^2),$$

где $\tau_{\rm M} = \varepsilon/eN\mu$ – время максвелловской (дрейфовой) релаксации пространственного заряда свободных носителей, *N* - стационарная (темновая) концентрация электронов проводимости.

144
Максвелловская релаксация - это восстановление локальной электронейтральности проводящей среды за счет токов проводимости (дрейфа подвижных носителей заряда), вызванных электростатическим полем в окрестности избыточного заряда. В результате поля всех локальных зарядов в проводящей среде экранируются, а флуктуации концентрации подвижных зарядов рассасываются тем скорее, чем выше проводимость среды $\sigma = eN\mu$. Темновая проводимость BSO невелика из-за низкой концентрации электронов проводимости, поэтому максвелловская релаксация в нем гораздо медленнее, нежели в "классических" полупроводниках. Отметим, что в фоторефрактивном кристалле τ_M неверно рассматривать как константу; проводимость о явно зависит от концентрации электронов N, поскольку эти электроны неосновные и неравновесные. На фоне пространственно-однородной освещенности концентрация электронов *n* растет, а это ускоряет максвелловскую релаксацию и формирование пространственного заряда.

Под действием поля пространственных зарядов за счет эффекта Поккельса в кристалле формируется фазовая решетка, в которой вариация показателя преломления от максимума к минимуму

$$\delta n = \frac{n_0^3 r_{ik}}{2} \boldsymbol{\varepsilon};$$

здесь n_0 - исходный показатель преломления, r_{ik} – электрооптический коэффициент.

Подробнее об электрооптических коэффициентах сказано в приложении 1. Здесь мы только напомним, что по традиции r_{ik} образуют матрицу электрооптических коэффициентов размерностью 3×6, составленную из ненулевых коэффициентов истинного электрооптического тензора 3×3×3.

Коэффициент брэгговского рассеяния опорной волны E_2 на такой фазовой решетке удобно представить как коэффициент "отражения" (R) сигнальной волны:

$$R = \left| \frac{E_4}{E_3} \right|^2 \approx \frac{1}{4} \left(k_0 LmpSW \tau_{\rm T} \frac{n_0^3 r_{ik}}{2} \frac{k_B T}{e} \right)^2$$

Эта формула верна при малом значении $k_0 L \delta n < 1$, в котором можно распознать *индекс фазовой модуляции*. Так или иначе, на

практике в фоторефрактивных кристаллах большие индексы модуляции крайне редки.

Более общее выражение для коэффициента рассеяния дает теоретическая модель связанных волн, упрощенно описанная в приложении 3.

При таких исходных данных время формирования фазовой решетки должно быть около 4 мс. Ранее мы отметили. что в рассматриваемой геометрии предметная волна Е₃ при интерференции с волной накачки E_1 формирует редкую решетку, а при интерференции с волной E_2 – частую решетку. Период редкой решетки на данной установке $\Lambda \sim 1$ мкм; следовательно, для устойчивого наблюдения обращенной волны стабильность всех параметров установки и лазера за миллисекунды должна обеспечить стабильность формы интерференционных полос положения не хуже 0,01 мкм. И На опыте нетрудно убедиться в том, что воздействие вибраций на оптическую систему уменьшает интенсивность наблюдаемой обращенной волны, вплоть до уровня шума.

Таблица 7.2

Параметры			
Кристалл BSO		Решетка	
n ₀	2,5	Угол схождения θ_{13}	0,2
r ₄₁	5×10 ⁻¹² м/В	Волновое число k_0	$5 \times 10^5 \text{ cm}^{-1}$
ϵ_{st}	56	Период Л	~ 1 мкм
μ	$3 \times 10^{-7} \text{ m}^2/(\text{B c})$	Лазер	
τ	$5 \times 10^{-5} c$	Длина волны λ_0	0,63 мкм
l	1 см	Мощность Р	10 мВт
Δn	4×10^{-5}	Ø луча	2 мм

Параметры кристалла и лабораторной установки

Все предыдущие расчеты были проделаны в приближении заданных полей $E_1...E_3$; в частности, не учитывали поглощение света в кристалле. Между тем, только в результате поглощения происходит генерация неравновесных электронов и проявляется фоторефрактивная нелинейность.

Покажем, как можно оценить коэффициент поглощения, при котором коэффициент преобразования света в обращенную волну максимален. Очевидно, что такой максимум существует, так как коэффициент преобразования стремится к нулю и при бесконечно малом, и при очень большом поглощении. Понятно, что коэффициент поглощения не должен быть слишком велик, иначе затухание света в кристалле будет очень сильным, и мощность обращенной волны будет слишком мала. Далее будем считать поглощение света на длине кристалла небольшим, порядка десятка процентов, чтобы упростить математическую часть задачи.

Затухание *напряженности* волны в слабо поглощающей среде с коэффициентом поглощения *К* на расстоянии *х* можно описать приближенно как

$$E(x) \approx E_0(1 - Kx).$$

Полагая длину кристалла равной *L*, запишем зависимость напряженности волны *E*₄ от *K*:

$$W_4 \propto K(1 - KL/2)^4$$
;

здесь, при малом *K*, интегрирование при оценке рассеяния на фазовой решетке и затухания рассеянной волны можно заменить приближенным усреднением по трассе пучка. Решив уравнение

$$dW_4/dK = 0$$

получим

 $K_{\rm opt} \approx 1/(6L)$.

Это означает, что оптимальный коэффициент поглощения не очень мал, и на длине кристалла L при этом напряженность волны должна падать примерно на 30%. На существующей установке длина кристалла L = 1 см, то есть $K_{opt} \approx 0.16$ см⁻¹, что меньше справочного значения $K \approx 0.3$ см⁻¹ для BSO.

На практике неудобно варьировать K; проще выбрать кристалл оптимальной длины, но ее выбор тоже ограничен. Световые пучки конечного сечения пересекаются в ограниченном объеме, поэтому длину кристалла нет смысла брать большей, нежели длина перекрытия опорного и предметного пучков (волн E_1, E_3), ограниченная их поперечниками и углом схождения. На существующей установке, по приведенным выше оценкам для K, длину кристалла можно было бы увеличить, но практически она ограничена перекрытием пучков, пересекающихся под углом $\theta_{13} \approx 0,2$ на длине 1 см.

7.10. ОВФ в реальных пучках ограниченного сечения

В нашей модели ВЧВС все волны мы полагали плоскими, неограниченными в поперечном направлении. Только такие волны формируют решетку со строго определенным волновым вектором (вектором обратной решетки), p = 2q. Диаграмма направленности рассеяния на такой идеализированной решетке – б-функция; любые флуктуации волнового вектора или частоты, нарушающие строгую сопряженность волн E_1, E_2 , должны приводить к затуханию волны E_4 до нуля. В реальности поперечные сечения световых пучков не просто ограничены; часто их намеренно уменьшают при помощи коллиматоров, чтобы напряженность поля в них была по возможности большей, и нелинейности проявлялись как можно сильнее. Понятно, что в ограниченных световых пучках также будет формироваться решетка, и рассеяние на ней будет порождать четвертую волну, но соотношения между волновыми векторами уже не должны приводить к б-функциям. Ограничения апертур взаимодействующих волн порождают дифракционные искажения волновых фронтов, тем более значимые, чем меньше апертуры. В частности, интуитивно мы должны ожидать, что волна E_4 , номинально сопряженная к сигнальной волне E_{3} , в опыте будет отличаться от строго сопряженной тем сильнее, чем меньше апертура самого узкого из взаимодействующих пучков. Проще говоря, обращенный световой пучок приобретает конечную дифракционную расходимость.

Не менее важно и то, что условие взаимной сопряженности волн накачки на опыте не обязательно должно быть строгим; эффект ВЧВС можно наблюдать и при ограниченных отклонениях волновых векторов от параллельности, и при ограниченном различии их длин (то есть частот волн накачки).

Самую простую и грубую оценку допуска на различие волновых векторов дают следующие соображения: ширина интерференционного максимума тем меньше, чем больше максимальная разность фаз между частичными когерентными волнами сравнимой напряженности. Эта же разность фаз, поделенная на длину волнового вектора, есть длина когерентности. После этого ясно, что ширина лепестка угловой диаграммы направленности при обращении плоской волны приближенно равна отношению длины волны к радиусу когерентности поля на решетке. Этот радиус когерентности не превышает апертуру самого узкого из взаимодействующих пучков и дополнительно ограничен отклонением волн накачки от взаимной сопряженности. Оценку радиуса когерентности дает величина, обратная сумме волновых векторов волн накачки (при строгом сопряжении эта сумма равна нулю, то есть ее можно считать мерой точности сопряжения волн накачки).

На практике приходится выбирать компромисс между напряженностью поля, растущей при сжатии пучков, и точностью сопряжения сигнальной и обращенной волн, при сжатии пучков ухудшающейся. При сжатии пучков растет дифракционная эффективность фазовой решетки (так как растет индекс фазовой модуляции и/или коэффициент амплитудной модуляции) и, следовательно, растет амплитуда рассеянной волны. При этом же смягчаются требования к точности сопряжения волн накачки, но обращенная волна искажается сильнее, и амплитуда ее уменьшается за счет уменьшения объема нелинейной среды, где пересекаются взаимодействующие пучки и формируется фазовая решетка.

Приведенные здесь полукачественные оценки можно обосновать строго, привлекая методы фурье-анализа в пространственночастотном представлении. Мы используем иной подход, тоже основанный на фурье-анализе, но более наглядный в сравнении с формальным.

Решетка показателя преломления неподвижна в пространстве, поэтому рассеянная на ней световая волна имеет ту же частоту, что и падающая; возможные направления рассеяния наглядно отображает поверхность волновых векторов. Ее строят как сферу с радиусом, равным длине волнового вектора $|k_{\rm n}|$ падающей волны (рис. 7.11, *a*). Волновой вектор рассеянной волны $k_{\rm p}$ соединяет центр сферы **O** с ее поверхностью. Рассеяние происходит, когда разность векторов $k_{\rm n}$ и $k_{\rm p}$ равна вектору обратной решетки **p** (то есть выполнено условие Брэгга, $k_{\rm n} - k_{\rm p} = p$).

Решетку, записанную пучками конечной ширины, представляет непрерывный набор векторов обратной решетки (пространственных частот), различных по длине и направлению. Векторы этого набора заканчиваются в сфере радиуса $\approx \pi/d$, описанной вокруг конца вектора p, отвечающего точному условию Брэгга (рис. 7.11, δ). Даже плоская волна на такой решетке рассеивается в волну с определенным набором (спектром) пространственных частот (и, следовательно, с некоторой угловой расходимостью).



 $|\mathbf{k}_n| = |\mathbf{k}_n|$ $p = 2k\sin(\theta/2)$



 $|\Delta k_1| = |k_1 - p| \approx \pi/d$





Рис. 7.11. Применение модели поверхностей волновых векторов для решения задач о дифракции плоских волн на синусоидальных решетках показателя преломления:
 a - рассеяние на решетке со строго определенным вектором ;
 б - на решетке с набором пространственных спектральных компонент *в* - схема для расчета угловой селективности решетки, записанной световыми пучками конечного поперечного размера, *г* - схема для оценки спектральной селективности решетки, записанной пространственно ограниченными световыми пучками

Меру угловой селективности решетки определим как размер угла отклонения $\Delta \theta$ от точного направления пространственного синхронизма (рис. 7.11, *в*), при котором интенсивность рассеянной волны уменьшается в два раза. При таком $\Delta \theta$ конец вектора *p* удален от поверхности волновых векторов на предельное расстояние $\approx \pi/d$.

Чтобы вычислить $\Delta \theta$, рассмотрим треугольник *OBE* (рис. 6, *в*), где OB = k, $BE = p = 2k \sin(\theta/2)$, $OE = k + \pi/d$, $\angle OBE = \pi/2 - (\theta/2 - \Delta \theta)$. По «теореме косинусов»

$$(k + \pi/d)^2 = k^2 + p^2 - 2kp \cos[\pi/2 - (\theta/2 - \Delta\theta)]$$

В приближении высокой селективности, когда

 $\Delta \theta \ll 1$, $\cos \Delta \theta \approx 1$, $\sin \Delta \theta \approx \Delta \theta$,

$$(\pi/d)^2 \ll 2k\pi/d \ (\pi/d) \ll 2k$$
,

после преобразований получим

$$\Delta \theta = \frac{\lambda}{2d\sin(\theta/2)} = \frac{\lambda_0}{2n_0 d\sin(\theta/2)} = \frac{\lambda}{d}.$$

Аналогично оценим спектральную селективность решетки, то есть ее дифракционную эффективность для волны с длиной λ_2 , отличной от той длины волны λ_1 , которой записали решетку.

Концы волновых векторов k_{2n} и k_{2p} лягут на поверхность волновых векторов радиуса $|k_2|$ (см. рис. 7.11, *в*). Пусть $|k_2| > |k_1|$. Если при считывании k_2 совпадает с направлением синхронизма для волн k_1 , то направления и поверхности названных волновых векторов примут вид рис. 7.11, *в*. На векторной диаграмме видно, что длину $|k_{2n}|$ можно изменять до тех пор, пока сфера радиуса $\approx \pi/d$ с центром в конце вектора p (точка *C*) пересекается с поверхностью волновых векторов k_2 ; предельный случай – касание названной сферы и поверхности волновых векторов – соответствует уменьшению интенсивности волны k_{2p} в два раза по сравнению с точным синхронизмом.

Спектральная селективность решетки, Δλ/λ, выражается через предельно допустимую разность длин волновых векторов,

$$k_2 - k_1 = \frac{2\pi}{\lambda_2} - \frac{2\pi}{\lambda_1} = \frac{2\pi\Delta\lambda}{\lambda_2\lambda_1} = k_2 \frac{\Delta\lambda}{\lambda_1}$$

Для оценки рассмотрим треугольник OBC, где

$$OB = k_2; BC = p = 2k_1 \sin(\theta/2); OC = k_2 - \pi/d;$$

 $\angle OBC = \pi/2 - \theta/2.$

По «теореме косинусов»,

$$(k_2 - \pi/d)^2 = k_2^2 + p^2 - 2k_2p\cos(\pi/2 - \theta/2) =$$
$$= k_2^2 + 4k_1^2\sin^2(\theta/2) - 4k_2k_1\sin^2(\theta/2)$$

После учета $(\pi/d)^2 \ll 2k\pi/d$, или $\pi/d \ll 2k$, получим:

$$\frac{\Delta\lambda}{\lambda_1} = \frac{\lambda_1}{4d\sin^2(\theta/2)} = \frac{\lambda_{10}}{4n_0d\sin^2(\theta/2)}.$$

В задание работы входит измерение угловой селективности решетки показателя преломления. При постоянной длине волны лазера непосредственно измерить спектральную селективность решетки невозможно; тем не менее, косвенно можно убедиться в том, что она существует и конечна. Для этого следует изменять направление E_2 не в плоскости схождения волн E_1, E_3 при записи, а в плоскости, ортогональной к ней и проходящей через линию ОА. В этой геометрии дифракцию волны E_2 определяет проекция ее волнового вектора на плоскость схождения записывающих волн E_1, E_3 . В ортогональной плоскости все условия однородны, и лежащая в ней компонента волнового вектора ни на что не влияет. Так как проекция вектора не может быть больше его длины, то увеличение угла схождения волн в новой геометрии равносильно уменьшению длины волнового вектора, то есть увеличению длины волны. Изменение кажущейся длины волны по отношению к исходной выражается через угол (ψ) наклона волны E_2 по отношению к плоскости (E_1, E_3) .

$$k' = k \cos \psi; \ k - k' = k(1 - \cos \psi) \approx k \psi^2 / 2;$$

$$k - k' = (2\pi/\lambda) (\Delta \lambda / \lambda') = k(\Delta \lambda / \lambda') \approx k \psi^2 / 2;$$

$$\Delta \lambda / \lambda' \approx \psi^2 / 2.$$

В реальной измерительной установке угол ψ ограничен, и поэтому диапазон изменений интенсивности обращенной волны при изменении ψ невелик.

Отметим, что все сказанное в настоящем разделе относится к любым дифракционным решеткам, независимо от их происхождения. Поэтому все полученные здесь соотношения и зависимости можно применять и к другим вариантам считывания решеток, в частности, к акустооптическим модуляторам, к спектрометрам, к голограммам и пр. Наглядность примененного подхода – общее достоинство анализа волновых процессов и структур в пространстве волновых векторов (по сути – в импульсном представлении).

7.11. Запись и считывание решеток в анизотропных средах

Очень часто оптические свойства среды, в которой формируется фазовая решетка, зависят от поляризации излучения; кристаллы, в которых особенно сильны нелинейные эффекты (в том числе и фоторефрактивные), относятся либо к двулучепреломляющим, либо к гиротропным. Двулучепреломление — это зависимость фазовой скорости световой волны от направлений волнового вектора и вектора *линейной* поляризации; гиротропия — зависимость фазовой скорости от знака *круговой* (циркулярной) поляризации.

Далее мы будем рассматривать только гиротропную среду, так как на реальной установке применен именно гиротропный силленит. Двулучепреломляющие среды, где формирование и считывание фазовых решеток гораздо сложнее (а для практики этот вариант может оказаться менее удобным), здесь мы рассматривать не будем. Ограничимся тем, что отметим хорошие фоторефрактивные свойства двулучепреломляющего Fe:LiNbO₃; у него электрооптические коэффициенты гораздо больше, чем у BSO (см. табл. 1), а оптическое поглощение на переходах ловушка-зона можно оптимизировать для конкретного применения, выбирая при легировании молярную концентрацию Fe от 5×10^{-5} до 5×10^{-4} .

При распространении света в гиротропной среде от точки к точке по трассе изменяется угол поворота осей эллипса поляризации, а форма эллипса (эксцентриситет) остается неизменной; круговая поляризация при этом не изменяется. Количественной мерой гиротропии служит удельное вращение плоскости поляризации ρ , определенное как отношение угла поворота к геометрической длине пути света в среде. Феноменологически причину вращения плоскости поляризации, как показал Френель (1825 г.), можно понимать как разность фазовых скоростей волн с левой и правой круговыми поляризациями.

Если для двух круговых поляризаций ввести показатели преломления n_l и n_r (по аналогии с n_o и n_e в одноосной двулучепреломляющей среде), то разность

$$\Delta n_c \equiv n_l - n_r = \rho \lambda / \pi.$$

В изотропном гиротропном кристалле поверхность волновых векторов расщепляется на две сферы. Поскольку длина волнового вектора $k = 2\pi n/\lambda$, а длина волны в вакууме λ неизменна, то радиусы поверхностей волновых векторов можно положить равными

 n_l и n_r ; тогда векторная диаграмма будет отличаться от рассмотренных ранее только коэффициентом подобия $2\pi/\lambda$.

При записи образуются две решетки, каждая из которых сформирована интерференцией волн с одним знаком круговой поляризации. Третьей решеткой, сформированной интерференцией волн с разными знаками круговой поляризации, на практике можно пренебречь: интенсивность соответствующих интерференционных полос пропорциональна θ_{13}^2 . При $\theta_{13} = 0$ волны с противоположной круговой поляризацией не интерферируют (как ортогональные функции), а при $\theta_{13} \neq 0$ интерферируют только малые проекции одной волны на другую, и их амплитуды пропорциональны $\sin\theta_{13} \approx \theta_{13}$.



Рис. 7.12. Симметричная схема записи фазовых решеток в гиротропном кристалле; $p^{(r)}$ и $p^{(l)}$ - волновые векторы соответствующих решеток

При считывании пары решеток в гиротропном кристалле возможны четыре варианта рассеяния волн. На каждой из решеток волна может рассеяться в волну с тем же знаком (*внутримодовое* рассеяние) и в волну с противоположным знаком (*межмодовое* рассеяние) поляризации. В первом приближении векторы обратных решеток можно считать одинаковыми. Для большей строгости можно учесть, что угол θ_{13} не зафиксирован в кристалле, а задан углом схождения лучей Θ_{13} вне кристалла, так что угол $\theta_{13} \approx \Theta_{13}/n$, но это дает к длинам и направлениям векторов обратных решеток поправки того же порядка малости, что и Δn_c .

Все четыре варианта рассеяния волн двух поляризаций на двух решетках отображены на рис. 7.13. Векторные диаграммы позволяют рассчитывать геометрию рассеяния, но не дают оценок для амплитуды рассеяния. Ранее были приведены формулы для амплитуды рассеяния на решетке в изотропной среде; ими можно воспользоваться, установив значение электрооптического коэффициента для заданных поляризаций взаимодействующих волн.

В матрице электрооптических коэффициентов для кристаллов симметрии 23 отличны от нуля только $r_{41} = r_{53} = r_{63}$, и у кристалла BSO коэффициент $r_{41} = 5 \times 10^{-12}$ м/В. Так как угол θ_{13} мал, плоскости максимальной интенсивности интерференционных полос практически параллельны волновым векторам k_1 и k_3 . Следовательно, вектор созданного фоторефрактивным эффектом электрического поля ε почти ортогонален векторам k_1 и k_3 . Чтобы в такой конфигурации амплитуда рассеяния на решетке была максимальной, нужно, чтобы поперечный эффект Поккельса давал максимальный локальный сдвиг фазы на длине взаимодействия волны k_2 с решеткой. Для этого используют *косой срез* кристалла BSO (по образу и подобию KDP, применяемого в модуляторах с поперечным эффектом Поккельса): ребра параллелепипеда кристалла параллельны кристаллографическим направлениям [110], [110] и [001]. Кристалл ориентируют так, что векторы $k_1...k_3$ направлены по [110], а $p \parallel [001]$ или $p \perp [001]$ (p, разумеется, лежит в плоскости $k_1...k_3$).



Рис. 7.13. Четыре возможных процесса при считывании решеток, записанных парой волн в гиротропном кристалле в симметричной схеме

Дифракционную эффективность решетки определяет индекс фазовой модуляции, пропорциональный магнитуде изменений показателя преломления в кристалле, а также типы поляризации падающей и рассеянной волн, e_i и e_s .

Изменение показателя преломления бл можно представить в виде матрицы δn_{ii} – тензорного оператора, действующего на вектор e_i и дающего вектор e_s . (Строгое доказательство опустим).

Для двух выбранных ориентаций вектора *р* в кристалле BSO (класс 32) матрицы δn_{ii} имеют вид

~ 11

$$\boldsymbol{p} \parallel [001]: \quad \delta n_{ij} = n^3 r_{41} \mid \overline{E}_{st} \mid \begin{vmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{vmatrix},$$
$$\boldsymbol{p} \perp [001]: \quad \delta n_{ij} = n^3 r_{41} \mid \overline{E}_{st} \mid \begin{vmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{vmatrix}.$$

Единичные векторы e_i и e_s . выберем в виде

$$\mathbf{e}_{i,s}^{r} = \frac{1}{\sqrt{2}} \| \mathbf{1} \quad \mathbf{0} \quad -i \|, \quad \mathbf{e}_{i,s}^{l} = \frac{1}{\sqrt{2}} \| \mathbf{1} \quad \mathbf{0} \quad i \|,$$

В случае *р* || [001]

$$\left(\mathbf{e}_{s}^{r,l} * \delta \hat{n} \mathbf{e}_{i}^{r,l}\right) = \frac{1}{2}n^{3}r_{41} | \overline{E}_{st} |,$$

то есть амплитуды внутримодовых $(r \rightarrow r, l \rightarrow l)$ и межмодовых $(r \rightarrow l, l \rightarrow r)$ рассеяний одинаковые и ненулевые. Заметим, что при внутримодовом рассеянии волновые векторы рассеянных волн отличаются по направлению настолько слабо, что циркулярно поляризованные взаимно-когерентные рассеянные волны, не разделенные в дальней зоне, складываются в одну волну с линейной поляризацией (рис. 7.14, а).

В случае $p \perp [001]$ амплитуды внутримодовых рассеяний,

$$\begin{pmatrix} \mathbf{e}_{s}^{r} \ast \hat{\delta n} \mathbf{e}_{i}^{r} \end{pmatrix} = \begin{pmatrix} \mathbf{e}_{s}^{l} \ast \hat{\delta n} \mathbf{e}_{i}^{l} \end{pmatrix} = \mathbf{0},$$
$$\begin{pmatrix} \mathbf{e}_{s}^{r} \ast \hat{\delta n} \mathbf{e}_{i}^{l} \end{pmatrix} = \begin{pmatrix} \mathbf{e}_{s}^{l} \ast \hat{\delta n} \mathbf{e}_{i}^{r} \end{pmatrix} = i n^{3} r_{41} | \overline{E}_{st} |$$

где волны циркулярно поляризованы и разделены в дальней зоне, как показано на рис. 7.14, б. На практике удобнее наблюдать разделение волн по углам не в дальней зоне, а в фокальной плоскости подходящей линзы, как расщепление фокального пятна на части. Размеры фокальных пятен при этом отображают расходимость рассеянных пучков.



Рис. 7.14. Тонкая структура диаграммы направленности рассеяния линейно поляризованного излучения на решетке в толстом гиротропном кристалле; стрелками показаны типы поляризации рассеянных волн в дифракционных максимумах

Внутримодовое рассеяние происходит при тех же самых углах падения волны E_3 , при которых была записана решетка. Поэтому в изотропном гиротропном кристалле решетку можно записывать и при помощи линейно-поляризованных волн. При этом волны с противоположными круговыми поляризациями внутри кристалла совместно формируют по существу одну и ту же решетку показателя преломления.

7.12. Динамические особенности фазовой решетки в фоторефрактивном кристалле

Поняв, как решетка образуется, попытаемся понять, долго ли она живет, куда и как она исчезает. Немного сложнее вопрос, как одна форма решетки сменяет другую при изменении формирующих ее световых волн.

Ранее была получена оценка скорости формирования решетки показателя преломления в фоторефрактивном кристалле. Характерное время формирования решетки существенно зависит от времени максвелловской релаксации пространственного заряда, зависящего, в свою очередь, от концентрации свободных носителей. В рассматриваемом случае свободные носители – электроны проводимости – образуются под действием света; темновая концентрация свободных электронов в BSO, как в широкозонном полупроводнике, пренебрежимо мала. Отсюда следует, что время максвелловской релаксации при разных условиях освещения и даже в разных участках решетки изменяется на порядки величины; релаксация за счет токов диффузии и проводимости идет быстрее всего в сильно освещенных участках. Наиболее сильна освещенность в пучностях стоячих волн.

Ранее было отмечено, что дифракционную эффективность решетки определяет конкуренция процессов фотогенерации объемного заряда и его релаксации за счет изменения заселенностей ловушек. После выключения света неравновесные электроны рекомбинируют, и пространственный заряд малоподвижных заряженных ловушек медленно релаксирует. В процессах релаксации заметную роль могут играть только токи проводимости подвижных носителей и, конечно же, порождающие подвижные носители эффекты – термогенерация и фотовозбуждение. Опыт показывает, что в силленитах тепловая генерация носителей при комнатной температуре очень слабая, переходы с ловушек в валентную зону имеют пренебрежимо малую вероятность, и потому релаксация в темноте протекает очень медленно.

Скорость тепловой генерации носителей экспоненциально зависит от температуры и от энергии активации связанного состояния (ловушки). По данным спектроскопии поглощения, энергия активации глубоких ловушек в силленитах порядка 1 эВ, что соответствует длине волны порядка 1 мкм и температуре 10⁴ К. Кстати говоря, отжиг *нелегированных* кристаллов ниобата лития, в которых тоже может быть заметным диффузионный фоторефрактивный эффект, можно применять для восстановления рабочих характеристик электрооптических модуляторов излучения, испорченных действием очень сильного света. В нашем случае важнее отметить, что при повышении температуры, по теории, все процессы формирования и распада решетки должны ускоряться, и должна расти (но лишь линейно) дифракционная эффективность. Так или иначе, при комнатных температурах следует ожидать, что сформированная решетка далее, в темноте, может деградировать довольно медленно, за минуты или дольше. Такое свойство фоторефрактивного кристалла используют для записи информации и ее хранения в течение ограниченного времени.

Теперь ответим на вопрос: как долго можно считывать информацию, записанную в фоторефрактивном кристалле. Попутно можно поставить и такой вопрос: какую интегральную энергию оптического сигнала можно получить при считывании информации, разрушаемой в самом процессе ее считывания?

Кроме «естественных» механизмов релаксации пространственного заряда в фоторефрактивном кристалле, при четырехволновом взаимодействии существует еще один фактор – волна Е₂, которая выгоняет из теневых участков заряды, связанные на глубоких донорных уровнях. Непосредственно на опыте можно убедиться в том, что решетка, записанная при интерференции волн E_1 и E_3 , сохраняется по крайней мере в течение нескольких секунд, если эти волны перекрыты, а Е2 продолжает действовать. Если же после записи решетки перекрыть все волны и оставить кристалл в темноте, то при повторном включении одной только волны Е2 через несколько минут мощность восстановленной волны Е4 в начальный момент будет примерно такой же, как и перед перекрытием всех волн. Далее мощность волны E₄ под действием волны E₂ убывает. Эти наблюдения показывают, что темновая релаксация пространственного заряда идет гораздо медленнее, чем релаксация в условиях однородного освещения кристалла.

В экономичной по всем параметрам схеме считывания информации (решетки) на нее падает только волна E_2 . Эта волна бегущая, она генерирует носители равномерно во всем объеме, где она действует. Она стирает существующее неравномерное распределение пространственных зарядов тем быстрее, чем больше ее интенсивность. Если в процессе динамической записи/считывания интенсивность в максимуме интерференционной картины была пропорциональна ~ $4E_2^2$, то при считывании энергетическая освещенность в тех же точках пропорциональна только ~ $E_2^2/2$, то есть почти на порядок ниже. Поэтому можно ожидать, что считывание информации можно вести в течение времени на порядок большего, нежели

запись. Опыты, которые нетрудно поставить на реальной установке, эти выводы подтверждают.

Если в процессе считывания статической информации увеличить интенсивность считывающей волны, то интенсивность сигнальной волны тоже возрастет, в первом приближении пропорционально. Можно ожидать, что в первом приближении интегральная по времени энергия сигнальной волны должна слабо зависеть от интенсивности считывающей волны. Следовательно, при циклической записи и считывании информации в фоторефрактивном кристалле длительность цикла можно сократить, если увеличить интенсивность считывающей волны; при этом пиковая мощность сигнальной волны увеличится, длительность сигнала уменьшится, а энергия в первом приближении не изменится.

На первый взгляд может показаться, что волна E_2 в режиме динамической записи/считывания мешает записи, частично стирая решетку, записанную волнами E_1, E_3 , но это неверно. Волна E_2 вместе с E_1 образуют *стоячую* волну, которая записывает другую фазовую решетку, прозрачную для волн E_4 , E_3 , а в стирании может участвовать только разность E_1 и E_2 , то есть бегущая часть суммы этих волн. Отсюда следует, что выравнивание амплитуд E_1 и E_2 улучшает эффективность записи решетки и динамического обращения волнового фронта. С другой стороны, присутствие освещенности от бегущей волны должно ускорять максвелловскую релаксацию в теневых участках, то есть должно ускорять и формирование решетки. Учет названных фактов в совокупности приводит к интуитивному выводу: чем сильнее свет, тем выше дифракционная эффективность решетки и амплитуда обращенной волны. Опыт этот вывод подтверждает; но вместе с тем указывает и на явный фактор риска при работе с лазерным излучением высокой мощности. Поля объемных зарядов в фоторефрактивных кристаллах могут оказаться настолько сильными, что вызовут недопустимо высокие локальные механические напряжения, из-за которых в кристалле образуются микроскопические трещины, рассеивающие свет и нарушающие диффузию подвижных электронов проводимости. В частности, по этой причине фоторефрактивный эффект стараются подавить в кристаллах ниобата лития для модуляторов мощного лазерного излучения; с этой целью применяют легирование и/или так называемые стехиометрические кристаллы.

Отметим, что упомянутая ранее прозрачность для «чужих» волн – свойство фазовых решеток; они сами по себе даже в поглощающих средах не вносят дополнительных диссипативных потерь, и присутствие нескольких фазовых решеток влияет на рассеяние только «своих» волн. В «чужой» волне могут рассеиваться только пространственно-частотные компоненты, коррелированные с пространственно-частотным компонентами решетки; обычно такая корреляция слаба и маловероятна. Тем не менее, как было показано ранее, учет ограниченности в пространстве и немонохроматичности реальных волн превращает все δ-функции в распределения с конечной шириной и «крыльями». Эти «крылья» и дают малые вклады в рассеяние «чужих» волн; когда таких слабых рассеивателей много, их присутствие проявляется как ослабление произвольной падающей волны за счет ее диффузного рассеяния.

Свойства *амплитудных* решеток иные; здесь добавление каждой новой решетки, вообще говоря, увеличивает потери для *всех* волн, и количество сосуществующих в одной среде амплитудных решеток на практике ограничено.

Наконец, пора рассмотреть динамику замены одной решетки другой при изменениях формы фронта одной из волн, например, волны E_3 . Распределение интенсивности в полосах интерференции при сдвиге фронта E_3 полностью изменяется, начинается запись новой решетки, но прежняя решетка еще некоторое время существует. Новая запись частично стирает прежнюю. Если полосы интерференции перемещаются, например, при вибрации оптической системы, то вместо стационарного распределения вариаций показателя преломления образуется распределение «смазанное», без выраженных максимумов и минимумов. При этом индекс фазовой модуляции стремится к нулю, и снижается амплитуда рассеяния. Ранее было отмечено, что амплитуду рассеяния заметно уменьшают нестабильности полос интерференции в десятые доли микрона, если они проявляются за времена порядка миллисекунд и быстрее, то есть за характерное время записи решетки. В результате быстрой смены положений интерференционных полос обращенная волна слабеет или даже исчезает. В этом нетрудно убедиться на опыте, если придать оптической системе слабую вибрацию, легко постучав пальцем по ее платформе. Если же перекрыть лучи в секции записи (E_1 и E_3), то восстановленная волна не исчезает и от более сильных вибраций, но слабеет со временем, как отмечено выше, так как ее стирает бегущая волна E_2 .

Можно указать и другую, менее заметную, но вероятную причину разрушения обращенной волны при механических вибрациях оптики. Косвенно на этот фактор указывает практика голографической интерферометрии, когда на одну голограмму снимают две слабо различающихся сцены, а при восстановлении голограммы наблюдают интерференционные полосы, прорезающие фантом сцены и отображающие различие двух снятых сцен. (По существу так наблюдают интерференцию волн, не перекрывающихся во времени). Если между двумя сценами нет корреляции, то при восстановлении голограммы интерференционных полос не будет. Полос интерференции не будет и в том случае, когда на голограмме снято очень много сцен со слабыми и случайными отличиями (своего рода мощная выборка из статистического ансамбля). Последний случай аналогичен обращению волны при вибрациях оптики.

Все решетки записаны с одной и той же опорной волной E_1 , поэтому сопряженная волна E_2 восстанавливает все записанные волны сразу. У восстановленных волн E_4 волновые векторы могут отличаться очень мало, но временные части фазы этих волн случайные, и это делает деструктивной интерференцию взаимнокогерентных восстановленных волн. Если же решетка одна и стабильна, то флуктуации фазы E_2 не влияют на качество сигнальной волны E_4 .

7.13. Заключительные замечания

Эффекты, наблюдаемые в фоторефрактивных кристаллах класса силленитов под действием излучения, довольно разнообразны, и разнообразны механизмы релаксации, с ними связанные. Только упомянем такие эффекты, как «световое утомление» и противоположный ему эффект фотоиндуцированного поглощения. Интерпретация этих явлений основана на представлениях о довольно сложных структурах связанных состояний внутри запрещенной зоны даже нелегированных силленитов, с большим различием времен релаксации. Для более глубокого изучения этих вопросов следует обратиться к специальной литературе и современной научной периодике.

Опыт по обращению волнового фронта в фоторефрактивном кристалле, поставленный в данной лабораторной работе, правиль-

нее рассматривать как повод для первого знакомства с принципами нелинейной оптики и нелинейными характеристиками оптических материалов. В частности, технические применения фоторефрактивных кристаллов не ограничены модельными опытами; для практики наиболее интересны пространственные модуляторы излучения; в них могут быть и элементы динамической голографии.

Мы рассмотрели основные процессы формирования решетки показателя преломления в фоторефрактивном кристалле, где возникает нелокальная и неадиабатическая нелинейность. Формально, но с оговорками, ее описывают как кубичную нелинейную восприимчивость $\chi^{(3)}$.

Известно немало иных процессов, создающих нелинейность $\chi^{(3)}$ или сводимых к ней при формализованном описании. Эта тема относится к области *нелинейной оптики*. Здесь лишь только назовем еще один важный класс нелинейных взаимодействий, особенно эффективных при обращении волнового фронта – *вынужденные рассеяния*. Особенности ОВФ при вынужденном рассеянии – усиление и невырожденность (сдвиг по частоте) рассеянной волны. Примечательно, что при некоторых условиях обращенная волна при вынужденном рассеянии формируется из *единственной* падающей на среду предметной волны, без участия вспомогательных опорных волн накачки. Формирование фазовой решетки при вынужденном рассеянии происходит за время существенно меньшее, нежели в фоторефрактивном кристалле, но и этот процесс не всегда можно рассматривать как безынерционный. Заинтересованным читателям рекомендуем обратиться к соответствующей литературе. Тематика лабораторных работ по физической оптике для 7-го семестра взаимосвязана в такой степени, что многие вопросы требуют изложения сразу в нескольких методических руководствах. Авторы сочли полезным объединить все теоретические материалы и описания работ "в одной обложке", чтобы не дублировать тексты дополнений и множество рисунков. Может быть, такая организация материала поможет лучше уяснить единство идей и моделей, формально относящихся к разным разделам физической, информационной и технической оптики.

В предлагаемых лабораторных работах предметом изучения будут применения электрооптических и акустооптических эффектов для модуляции лазерного излучения. При ознакомлении с принципами работы модуляторов советуем обратить внимание на то, как в этих приборах работает фазовая модуляция излучения на основе эффектов Поккельса и фотоупругости, и как она сочетается с другими физическими эффектами и техническими приемами.

К каждой работе составлен список контрольных вопросов, предназначенных как для самопроверки студентов при подготовке к работе, так и для проверки знаний при допуске к работе и сдаче отчета. Степени трудности вопросов отмечены значками:

* – вопрос повышенной трудности;

** – вопрос особой трудности.

При подготовке к работе достаточно ответить только на основные (непомеченные) вопросы. При подготовке отчета следует изучить вопросы повышенной трудности (*).

Наиболее трудные вопросы (**) авторы предлагают для обсуждения с преподавателем при сдаче отчета.

Работы 1 и 2 проходят на одной установке; как правило, обе работы студенты выполняют в ходе одного занятия. По этой причине на схемах установок их *общие элементы* имеют *одинаковые* позиционные *обозначения*.

Лабораторная работа 1

ИЗУЧЕНИЕ ЭЛЕКТРООПТИЧЕСКОЙ МОДУЛЯЦИИ

Цель работы: изучение электрооптического модулятора на основе линейного электрооптического эффекта Поккельса, измерение его важнейших характеристик, измерение полосы пропускания фотоприемника и оптического тракта

Экспериментальная установка

В состав лабораторной установки (рис. Л1.1.) входят гелийнеоновый лазер (ЛГН-203), электрооптический модулятор МЛ-5, поляризатор, фотоприемник ФД-21КП, регулируемый источник постоянного напряжения, генератор НЧ (20 Гц...200 кГц) и осциллограф НЧ.

Осциллограф в цифровом варианте схемы наблюдения электрического сигнала виртуален; его функции выполняет ПК с платой АЦП под управлением программы общего назначения семейства LGraph или специализированных программных модулей. В альтернативном варианте применен цифровой запоминающий осциллограф, связанный с ПК по шине USB.

Лазер ЛГН-203 работает в режиме непрерывной генерации на длине волны 632,8 нм; тип колебаний ТЕМ_{00q} с линейной поляризацией, при мощности около 1мВт. Возможно применение гелийнеонового лазера ГН-02П. Возможны модификации установки,



Рис. Л1.1. Схема лабораторной установки для исследования электрооптической модуляции излучения:

1 – лазер, 2 – электрооптический модулятор МЛ-5, 4 – поляризатор,

5 - фотоприемник, 6 - осциллограф, 7 - устройство согласования,

8 – регулируемый источник постоянного напряжения 0..300 В,

9 – генератор НЧ; элемент 3 в данной конфигурации не устанавливают с заменой гелий-неонового лазера на инжекционный лазер (0,69 мкм) или неодимовый лазер с преобразованием излучения 1,06 мкм во вторую гармонику (0,53 мкм).

Электрооптический модулятор МЛ-5 работает на поперечном эффекте Поккельса; он содержит два кристалла метаниобата лития, между которыми установлена пластинка в полволны из кварца. Для нормальной работы модулятора падающий свет должен иметь горизонтальную или вертикальную линейную поляризацию и мощность не более 3 мВт при диаметре пучка от 1 до 2 мм.

Порядок выполнения работы

- 1. Установить на оптической скамье модулятор МЛ-5 на расстоянии 40..80 см от лазера.
- 2. Включить лазер, источник постоянного напряжения, НЧ генератор и осциллограф. Выходное напряжение генератора установить на ноль.
- Снять два защитных колпачка с модулятора и при помощи регулировочных винтов на юстировочной головке установить его так, чтобы лазерное излучение проходило через модулятор, не задевая краев светового отверстия. Проверить, наблюдая прошедшее через модулятор излучение на листе бумаги.
- 4. Поместить на оптическую скамью поляризатор.
- 5. Установить ориентацию поляризатора так, чтобы при изменении постоянного напряжения (от 0 до 300 В) интенсивность проходящего излучения претерпевала максимальные видимые изменения.
- Поместить фотоприемник на трассу прошедшего через модулятор излучения. Сигнал с фотоприемника подать на открытый вход [≅] осциллографа.
- Измеряя интенсивность света по вертикальному смещению луча осциллографа, снять зависимость интенсивности проходящего излучения от приложенного к модулятору постоянного управляющего напряжения в диапазоне -300 ...+300 В. Из этой зависимости получить значение полуволнового напряжения модулятора.
- 8. Одновременно с постоянным подать на модулятор переменное управляющее напряжение 15 В от генератора НЧ.
- 9. Получить на осциллографе сигнал переменной составляющей с фотоприемника при различных значениях постоянного напря-

жения на модуляторе, соответствующих положениям рабочей точки как на линейном, так и на нелинейном участках модуляционной характеристики. Зарисовать характерные осциллограммы и объяснить их.

- 10. Установить на модуляторе постоянное напряжение, соответствующее линейному участку модуляционной характеристики. При постоянной амплитуде переменного напряжения (15 В) снять АЧХ тракта модулятор-приемник (амплитудно-частотная характеристика (АЧХ) – это зависимость амплитуды переменной составляющей выходного сигнала от частоты постоянного по амплитуде управляющего сигнала).
- 11. Вывести на нуль постоянное и переменное управляющие напряжения. Выключить источник постоянного напряжения и генератор НЧ.
- 12. Закрыть защитными колпачками входное и выходное отверстия модулятора МЛ-5. Удалить с рабочей оптической скамьи на вспомогательную модулятор МЛ-5 и поляризатор.

Контрольные вопросы и задания

- 1. Что такое оптическая индикатриса?
- 2. Сравните достоинства и недостатки различных электрооптических кристаллов.
- 3. Для чего в модуляторе МЛ-5 установлена полуволновая пластинка?
- 4. Каковы особенности и возможные преимущества работы электрооптического модулятора со смещением рабочей точки на нелинейный участок модуляционной характеристики?
- 5. *Какие параметры описывают АМ, ЧМ, ФМ? Что представляют собой спектры излучения при различных видах его модуляции?
- 6. *Как реализовать АМ. ЧМ, ФМ и однополосную (SSB) модуляцию лазерного излучения при помощи электрооптических модуляторов?
- 7. *Как может быть устроен модулятор излучения на основе эффекта Фарадея? Каковы его особенности?
- *Составьте функциональную оптико-электронную схему стабилизатора мощности излучения лазера с использованием электрооптического модулятора.

Лабораторная работа 2

ИЗУЧЕНИЕ АКУСТООПТИЧЕСКОЙ МОДУЛЯЦИИ

Цель работы: изучение электрооптического модулятора на основе линейного электрооптического эффекта Поккельса, измерение его важнейших характеристик, измерение полосы пропускания фотоприемника и оптического тракта

Экспериментальная установка

В состав лабораторной установки (рис. Л2.1.) входят гелийнеоновый лазер (ЛГН-203), электрооптический модулятор МЛ-5, поляризатор, фотоприемник ФД-21КП, регулируемый источник постоянного напряжения, генератор НЧ (20 Гц...200 кГц) и осциллограф НЧ.

Осциллограф в цифровом варианте схемы наблюдения электрического сигнала виртуален; его функции выполняет ПК с платой АЦП под управлением программы общего назначения семейства LGraph или специализированных программных модулей. В альтернативном варианте применен цифровой запоминающий осциллограф, связанный с ПК по шине USB.

Лазер ЛГН-203 работает в режиме непрерывной генерации на длине волны 632,8 нм; тип колебаний ТЕМ_{00q} с линейной поляризацией, при мощности около 1мВт. Возможно применение гелийнеонового лазера ГН-02П. Возможны модификации установки, с заменой гелий-неонового лазера на инжекционный ла-



Рис. Л2.1. Схема лабораторной установки для изучения акустооптической модуляции излучения: *I* – лазер, *3* – акустооптический модулятор МЛ-201, 5 – фотоприемник, *6* – осциллограф, *10* – генератор ВЧ; элементы *2, 4, 7...9 в данной конфигурации не устанавливают* зер (0,69 мкм) или неодимовый лазер с преобразованием излучения 1,06 мкм во вторую гармонику (0,53 мкм).

Акустооптический модулятор МЛ-201 использует когерентное рассеяние света на акустической волне в геометрии Брэгга. Ультразвуковую волну в бруске из стекла ТФ-7 возбуждает пьезопреобразователь из ниобата лития. Оптимальная частота возбуждения акустической волны 80 Мгц при подведенной электрической мощности не более 4 Вт.

Порядок выполнения работы

- 1. Установить на рабочей оптической скамье акустооптический модулятор МЛ-201 на расстоянии ок. 1 м от лазера.
- 2. Включить генератор ВЧ Уровень выходной мощности генератора установить на ноль.
- 3. Рассчитать угол Брэгга для частоты модуляции 80 МГц с учетом скорости звука в стекле ТФ-7 (3км/с).
- 4. Свинтить два защитных колпачка с модулятора и при помощи регулировочных винтов на юстировочной головке установить его так, чтобы лазерное излучение проходило через модулятор, не задевая краев светового отверстия. При этом модулятор нужно установить под углом Брегга к падающему излучению. Юстировку следует проверять визуально, по проходящему и отраженному от модулятора излучению.
- 5. Установить на генераторе режим непрерывной генерации ("НГ"), частоту 80 Мгц и максимальный уровень выходной мощности (до 1 Вт).
- Наблюдая дифракционную картину в выходном излучении на светом экране, окончательно съюстировать модулятор для достижения максимальной концентрации излучения в -1 и нулевом порядках.
- Варьируя во всех доступных диапазонах выходную мощность генератора ВЧ, частоту модуляции, а также угол наклона модулятора, исследовать характер изменений дифракционной картины. Объяснить наблюдаемые явления.
- 8. Оптимизировать настройку модулятора для модулирующей частоты около 100 Мгц.
- 9. Установить генератор в режим с внутренней амплитудной модуляцией меандром на частоте 1 кГц ("ВНУТР").

- 10. Установить фотоприемник сначала в первый порядок и зарисовать осциллограмму сигнала фотоприемника. Затем проделать то же для –1 порядка. Объяснить вид осциллограмм.
- Измерить зависимость эффективности модуляции излучения от частоты в диапазонах 4 и 3 по сигналу интенсивности для нулевого и −1 порядков. При измерениях в диапазоне 3 настройку модулятора оптимизировать для 80 Мгц.
- 12. Объяснить, какие систематические погрешности возникают при проведении измерений по различным методикам (например, при неизменном положении фотоприемника в ненулевом порядке, при неизменной ориентации модулятора и т.п.).
- 13. Объяснить, какая методика измерений дает технические характеристики самого модулятора, и какая методика характеризует только конкретную оптическую систему.
- 14. Вывести на ноль уровень выходной мощности генератора ВЧ. Выключить генератор, лазер и осциллограф. Навинтить защитные колпачки на модулятор МЛ-201

Контрольные вопросы и задания

- 1. Можно ли модулировать свет на основе эффекта фотоупругости за счет квазистатических механических напряжений в упругой среде?
- 2. В чем заключается явление дифракции света на акустических волнах?
- 3. Существуют ли принципиальные ограничения в выборе частоты акустической волны для акустооптического модулятора?
- 4. Чем различается когерентное рассеяние света на стоячих и бегущих акустических волнах?
- 5. По каким параметрам различается излучение, рассеянное на акустической волне в разные интерференционные порядки?
- 6. В чем принципиально различаются режимы когерентного рассеяния Рамана–Ната и Брэгга?
- 7. Что называют рассеянием Брэгга и углом Брэгга?
- 8. Какие свойства акустооптического модулятора заданы условиями Брэгга?
- 9. Как качественно проявляются отклонения от условий Брэгга при настройке оптической схемы АОМ?

- 10. Какой параметр взаимодействия света с акустической волной называют длиной когерентности?
- 11. Укажите выражение для акустической мощности, необходимой для полной перекачки световой мощности в минус первый дифракционный порядок.
- 12. Какие факторы ограничивают перекачку мощности из исходной волны в нужный интерференционный порядок?
- 13. В каком диапазоне частот модуляции мощности ультразвука акустооптический модулятор будет эффективно работать как амплитудный модулятор? Какие исходные данные нужно учитывать при оценке эффективности?
- 14. *Составьте функциональную оптико-электронную схему стабилизатора мощности излучения лазера с использованием акустооптического модулятора.
- 15. *Как реализовать АМ, ЧМ, ФМ и однополосную (SSB) модуляцию лазерного излучения при помощи акустооптических модуляторов?
- 16. *Почему акустооптический модулятор, который очень эффективно работает при возбуждении его на частотах 80 и 240 Мгц, плохо работает на частоте 160 МГц?

Лабораторная работа 3

ПРОСТРАНСТВЕННЫЙ МОДУЛЯТОР ТИПА ФП-ЖК

Цель работы: исследование пространственно-временного модулятора света (ПВМС) на основе нематических жидких кристаллов (НЖК).

Экспериментальная установка

Блок-схема лабораторной установки для измерения некоторых характеристик ПВМС типа ФП–ЖК показана на рис. Л3.1.

Излучение ИК-светодиода *1* (0.9 мкм), сформированное линзой *2* в параллельный пучок, освещает фотоприемную часть ПВМС *3*. Элементы *1*, *2*, *3* образуют секцию записи изображения.

Излучение красного светодиода 4 (0.65 мкм), сколлимированное линзой 5, через поляризатор 6 проходит на электрооптическую часть ПВМС (3) и после отражения через поляризатор 7 идет на телекамеру 8. Объектив телекамеры настроен так, что на ее фотоприемной мишени образуется изображение слоя ЖК ПВМС. Элементы 4...8 образуют секцию считывания изображения.

Для измерения пространственного разрешения ПМС в секцию записи устанавливают другой тест-объект – стандартную штриховую миру. Разрешение определяют по непосредственной видности телевизионного изображения участка миры с наибольшим числом уверенно различимых линий.

Видеосигнал с телекамеры поступает на телемонитор 9 и на вход блока обработки видеосигнала 10. Обработанный видеосигнал наблюдают на осциллографе 11. Элементы 8...11 образуют простую телевизионную измерительную систему.

Для возбуждения ПВМС служит генератор звуковых частот 12. Для питания и управления излучением светодиодов 1 и 4 служит двухканальный регулируемый источник постоянных напряжений 13 (2..15 В). Светодиоды к нему присоединены через резисторы, ограничивающие ток (на схеме не показаны).

Тест-объект – транспарант 2 в виде чередующихся прозрачных и непрозрачных полос на фотопленке – помещают вплотную к фотоприемному входу ПМС и ориентируют так, чтобы полосы были параллельны строкам развертки изображения в телекамере (номинально они горизонтальны).

Процедура первичной обработки видеосигнала в блоке 10 включает интегрирование сигнала яркости по одной строке растра и сохранение полученного напряжения в аналоговой форме в буферном устройстве выборки–хранения (*sample–hold–reset*) в течение периода считывания и интегрирования сигнала следующей строки (около 64 мкс).



Рис. ЛЗ.1. Схема экспериментальной установки: 1 – светоизлучающий диод (0,9 мкм), 2 – тест-объект (транспарант), 3 – ПМС ФП–ЖК, 4 – светоизлучающий диод 0,6 мкм, 5 – коллимирующие линзы, 6, 7 – поляроиды, 8 – телекамера, 9 – телемонитор, 10 – блок обработки видеосигнала, 11 – осциллограф, 12 – генератор НЧ, 13 – блок питания светодиодов, 14 – ПЭВМ, 15 – монитор VGA

Обработанный сигнал яркости строк принимает аналогоцифровой преобразователь (АЦП) в составе ПЭВМ (14), работающей в операционной среде MS-DOS. Программа измерения поддерживает прием видеосигнала и его отображение на мониторе 15. Кроме того, программа поддерживает определение перепада яркости на границе света и тени, измерение частоты и амплитуды напряжения на выходе генератора 12, отображение формы этого напряжения (статически), и постоянное динамическое отображение обработанного сигнала в отдельном окне на экране 15.

Задание и порядок выполнения измерений

- 1. Включить блок питания излучателей-светодиодов. Установить ток питания светодиода в секции считывания изображения около 10 мА.
- Получить на телевизионной системе изображение тыльной поверхности ПМС. Для этого могут потребоваться перемещения телекамеры, фокусировка ее объектива и при необходимости поворот одного из поляризаторов в положение большего пропускания света.
- 3. Установить поляризаторы в скрещенное положение, чтобы излучение в зоне считывания не проходило к телекамере.
- 4. Установить в секции записи изображения ток питания ИК диода около 80 мА.
- 5. Включить питание ПМС, установить выходное напряжение генератора питания 10 В и, подбирая частоту, получить на телевизионной системе изображение тест-объекта 1 (пленка с горизонтальными полосами).
- 6. Подстроить оптическую систему в секции считывания так, чтобы изображение на видеомониторе имело наилучшую четкость.
- Получить на мониторе 15 обработанный видеосигнал. При необходимости выправить ориентацию тест-объекта так, чтобы перепады между максимумами и минимумами сигнала были наибольшими (при этом полосы тест-объекта параллельны строкам растра телевизионной системы).
- 8. По тест-объекту 1 провести измерения зависимости контраста изображения от частоты при различных значениях напряжения питания модулятора.
- 9. При оптимальной частоте измерить зависимость контраста от напряжения питания при максимальном токе питания ИК светодиода.
- 10. При оптимальном напряжении и оптимальной частоте измерить зависимость контраста от тока питания ИК светодиода.
- 11. По результатам наблюдений построить семейство зависимостей, оценить погрешности измерений и указать оптимальные режимы работы модулятора.
- 12. Установить оптимальные режимы работы модулятора для выполнения последующих измерений.

- 13. Установить тест-объект 2 (стандартная мира) и определить пространственное разрешение модулятора по изображению на телемониторе 11.
- 14. Получить у преподавателя и выполнить дополнительное задание.

Контрольные вопросы

- 1. Что такое «жидкий кристалл»?
- 2. Какие типы ЖК вы знаете?
- 3. Что происходит со светом при прохождении слоя НЖК?
- 4. Как проявляются электрооптические эффекты в НЖК?
- 5. Как устроен ПМС с электрическим управлением? Чем различаются структуры, работающие в проходящем и в отраженном свете?
- 6. Как устроен ПМС типа ФП-ЖК?
- 7. Как работает диэлектрическое зеркало?
- 8. Почему для управления электрооптическими ячейками с НЖК используют переменное напряжение?
- 9. *Как ведет себя ячейка с НЖК при управлении постоянным напряжением?
- Назовите предельные характеристики ПМС ФП–ЖК по пространственному разрешению, быстродействию, контрасту. Чем эти показатели ограничены?
- 11. **Какие приборы применяют для преобразования света в электрический сигнал в телевизионных системах?
- 12. **Какую обработку видеосигнал проходит в данной системе?
- 13. *В использованной телекамере действует система АРУ, которая поддерживает постоянную среднюю яркость телевизионного изображения при изменениях освещенности фотоприемной мишени. Как это следует учитывать при измерениях контраста?
- 14. **Почему при повышении амплитуды управляющего напряжения на структуре ФП–ЖК нарушается управляемость модулятора светом?

Лабораторная работа 4 ОБРАЩЕНИЕ ВОЛНОВОГО ФРОНТА В ФОТОРЕФРАКТИВНОМ КРИСТАЛЛЕ

Цель работы: изучение четырехволнового взаимодействия в фоторефрактивном кристалле и основных характеристик фоторефрактивного эффекта.

Экспериментальная установка

Общая схема установки показана на рис. Л4.1, оптическая схема четырехволнового взаимодействия приведена на рис. Л4.2; функциональная электрическая схема показана на рис. Л4.3.



Рис. Л4.1. Схема экспериментальной установки: ЛГ – лазер, 3_п – частично пропускающие зеркала,

3 – зеркала, Л – линзы, П – толстая стеклянная пластина,

ФП – неоднородная стеклянная фазоискажающая пластинка.

М – механический прерыватель, BSO – кристалл B₁₂SiO₂₀

Источник излучения – гелий-неоновый лазер ГН-25 (мощность до 20 мВт). Лазерный пучок зеркалами З_п...З расщеплен на три вторичных пучка:

- пучок E_1 служит опорной волной,
- пучок *E*₃ служит записываемой сигнальной волной,
- пучок *E*₂ действует как считывающая волна.

Номинальные коэффициенты отражения полупрозрачных зеркал (при их рабочей ориентации) выбраны так, чтобы интенсивности всех трех названных пучков были примерно одинаковыми.

Пучок E_4 формируется в результате рассеяния на записанной фазовой решетке как обращенная (сигнальная) волна. Ее фокусирует линза на входной срез волоконного жгута-световода (условно не показан), по которому свет идет к фотоприемнику (ФЭУ-51).

Интенсивность обращенной волны удобнее измерять как переменную составляющую сигнала фотоприемника при периодическом прерывании сигнального пучка. Электромеханический прерыватель установлен на трассе считывающей волны (E_2). Для наблюдения сигнала обращенной волны и измерения ее относительной интенсивности служит осциллограф НЧ. При проведении измерений для наблюдения сигнала функции виртуального осциллографа выполняет ПЭВМ с АЦП НЧ. ПЭВМ под управлением специализированной программы позволяет наблюдать сигнал ФЭУ-51 практически в реальном времени, останавливать отображение сигнала на экране, измерять относительную мощность сигнала с отображением результата на экран монитора ПЭВМ.

Толстая стеклянная пластина П совместно с положительной линзой Л+ изменяют угол падения считывающего пучка на кристалл при повороте пластины П; при этом область пересечения пучков $E_1...E_3$ почти не изменяется. Дополнительная отрицательная линза Л- компенсирует изменение кривизны волнового фронта и диаметра пучка E_2 линзой Л+.

Поворот пластины П толщиной *h* с показателем преломления *n* на угол ϕ смещает луч на расстояние

$$d = h \cos \varphi \left[tg \varphi - \frac{\sin \varphi}{\sqrt{n^2 - \sin^2 \varphi}} \right],$$

Угол поворота $\Delta \theta$ волнового вектора волны E_2 по отношению к первоначальному направлению $tg(\Delta \theta) = d/f$ Подставив приближенные соотношения

$$\varphi \approx \pm 10^{\circ} \approx 0,17; \sin \varphi \approx tg \varphi \approx \varphi, \cos \varphi \approx 1$$

 $d \ll f, \quad tg \Delta \theta \approx \Delta \theta, \quad n = 1,5,$
получим $\Delta \theta \approx \frac{h}{f} \frac{\varphi}{3}.$



Рис. Л4.2. Оптическая схема четырехволнового взаимодействия и геометрические параметры световых пучков в установке рис. Л4.1



Рис. Л4.3. Схема электрических соединений в установке по рис. Л4.1; М – асинхронный электродвигатель, УНЧ – усилитель низкой частоты, ПЭВМ – компьютер с АЦП, остальное – как на рис. Л4.1. Электропитание аппаратуры от сети ~ 220 В не показано

Выполнение работы

Программное обеспечение установки постоянно совершенствуется и видоизменяется в деталях пользовательского интерфейса. При выполнении всех измерений нужно использовать временную инструкцию по применению измерительной программы (на рабочем месте).

Изучение обращения волнового фронта в фоторефрактивном кристалле

- 1. Включить аппаратуру. К настройке рекомендуем приступить через 10 минут.
- 2. Запустить на ПЭВМ в режиме MS-DOS программу приема сигнала (вход через меню; F2, и далее выбрать пункт "Обращение волнового фронта".
- 3. Проверить правильность расположения опорных волн. Встречные пучки E₁, E₂ должны совпадать на трассе от зеркала З_п до линзы Л+ с наилучшей точностью, какую можно проверить на глаз. Для наблюдения пучков пользуйтесь матовой пленкой или папиросной бумагой. При явном несовпадении пучков, когда его не удается выправить настройкой зеркала З1, нужно выполнить пп.3.1 ... 3.6.

3.1. Удалить линзы Л+, Л- и пластину П.

3.2. Снова проверить расположение опорных волн. Встречные пучки E_1 , E_2 должны совпадать на трассе от зеркала 3_{π} до зеркала 3_1 . При этом переднюю грань кристалла BSO нужно ориентировать примерно ортогонально к биссектрисе угла $\angle(E_1, E_3)$.

3.3. Вернуть на места последовательно линзы Л+ и Л-. При установке линз нужно выставить их главные оптические оси на ось опорных пучков.

3.4. Установить пластину П ортогонально к пучкам.

3.5. Регулировкой зеркал 3, 3 установить положение пучка E_3 так, чтобы в объеме кристалла он максимально перекрывался с опорными пучками.

3.6. Установить матовый экран или белую бумагу у входа световода фотоприемника. При удачной настройке оптической схемы на экране должно быть видно пятно сфокусированного обращенного пучка. Плавно поворачивая пластину П вокруг вертикальной оси, добиться максимальной яркости пятна.

- 4. Подстроить положение входной части световода фотоприемника так, чтобы сигнал ФЭУ был максимальным.
- 5. Ввести в пучок E_3 фазоискажающую пластинку. На матовом экране по ходу пучка E_3 за кристаллом наблюдать диффузное световое пятно. Двигая экран по ходу пучка, убедиться в плохом качестве пучка с фазовыми искажениями.
- 6. Убрать фазоискажающую пластинку. Измерить зависимость интенсивности обращенной волны от угла поворота пластины П вокруг вертикальной оси.
- Поворотом пластины П вокруг вертикальной оси вернуть настройку схемы на максимальный сигнал. В этом положении измерить зависимость интенсивности обращенной волны от угла поворота вокруг горизонтальной оси.
- 8. Вернуть на место фазоискажающую пластинку. Измерить зависимость интенсивности обращенной волны от угла поворота пластины П вокруг вертикальной оси.
- Поворотом пластины П вокруг вертикальной оси вернуть настройку схемы на максимальный сигнал. В этом положении измерить зависимость интенсивности обращенной волны от положения фазоискажающей пластинки по ходу пучка E₃.
- Измерить и записать все важные геометрические размеры в оптической схеме на участке от кристалла до линзы Л-; при измерении диаметра пучка на грани кристалла нужно принять диаметр пучка на полувысоте по интенсивности в 2..3 раза большим, чем его видимый диаметр (для гауссова профиля). Определить угол ∠(E₁, E₃) = θ₁₃.

Изучение основных характеристик диффузионного фоторефрактивного эффекта

При проведении наблюдений и измерений нужно будет перекрывать заслонками некоторые световые пучки. Расположение заслонок показано на рис. Л4.4

- 1. Настроить установку в любой из режимов, где сигнал от обращенной волны максимален.
- 2. Измерить интенсивность обращенной волны.
- 3. Перекрыть заслонкой 1 луч *E*₁. Без задержки измерить интенсивность обращенной волны.
4. Не убирая заслонку 1, повторять измерения до тех пор, пока интенсивность не уменьшится на порядок.



Рис. Л4.4. Схема расположения заслонок для блокировки отдельных волн при изучении диффузионного фоторефрактивного эффекта в установке по рис. Л4.1

- 5. Убрать заслонку 1 и немедленно измерить интенсивность обращенной волны.
- 6. Перекрыть заслонкой 1 луч *E*₁. Без задержки измерить интенсивность обращенной волны при слабой вибрации оптической системы. Объяснить наблюдаемый эффект.
- 7. Перекрыть заслонками 1, 2 лучи E_1 и E_2 . Через одну минуту убрать заслонку 2 и немедленно измерить интенсивность обращенной волны.
- 8. Убрать заслонку 1.
- 9. Убрав все заслонки, включить подсветку кристалла синим излучением светодиода на 20 секунд.
- Выключить подсветку и наблюдать за изменением интенсивности обращенной волны. Оценить время восстановления фоторефрактивных свойств кристалла.
- 11. Дождаться восстановления максимальной интенсивности обращенной волны.
- 12. Перекрыть заслонкой 1 луч *E*₁. Накрыть кристалл темным колпаком.
- 13. Через 10 минут снять колпак и немедленно измерить интенсивность обращенной волны. Объяснить наблюдаемый эффект.

Обработка результатов

- 1. Построить зависимости интенсивности обращенной волны от $\angle(E_1, E_2)$. Сопоставить их с теоретическими зависимостями.
- 2. Оценить селективность решетки по углу схождения волн и по длине волны.
- 3. Оценить время темновой релаксации пространственного заряда в кристалле BSO.
- 4. Оценить время деградации пространственного заряда в кристалле BSO под действием освещения бегущей волной.
- 5. Оценить время жизни электронов в глубоких ловушках.
- 6. Объяснить полученные результаты.

Контрольные вопросы

- 1. Что такое ВЧВС? Какая геометрия взаимодействующих волн ему соответствует?
- 2. Опишите физический механизм формирования решетки показателя преломления в фоторефрактивном кристалле.
- 3. Приведите зависимости параметров решетки от параметров формирующих ее волн.
- 4. Какие физические причины ограничивают селективность объемной решетки показателя преломления?
- 5. Как зависит дифракционная эффективность объемной фазовой решетки от ее протяженности в направлении распространения световой волны?
- 6. Как зависит интенсивность обращенной волны от коэффициента поглощения света в фоторефрактивном кристалле?
- 7. Как зависит интенсивность обращенной волны от интенсивности предметной волны в фоторефрактивном кристалле? В чем названная схема отличается от восстановления статической голограммы?
- 8. Приведите основные зависимости селективности фазовой решетки от параметров схемы эксперимента.
- 9. Какие особенности записи и считывания фазовых решеток возникают в двупреломляющих и гиротропных кристаллах?
- 10. Как решетка показателя преломления из считывающей волны формирует обращенную волну? Какие условия для этого необходимы?

- 11. Как можно проверить, что наблюдаемая волна действительно обращенная, а не отраженная от поверхности какого-либо неудачно расположенного оптического элемента?
- 12. Назовите физические механизмы, которые приводят к кубичной зависимости поляризации от напряженности поля. В чем их сходство и различие?
- 13. Как изменится интенсивность обращенной волны в фоторефрактивном кристалле, если в два раза увеличить интенсивность только одной волны E_2 ? Как этот результат должен зависеть от интенсивности всех остальных волн (например, при изменении выходной мощности лазера) при условии, что их соотношения неизменны?
- 14. Как изменится обращенная волна в фоторефрактивном кристалле, если предметная (обращаемая) волна E_3 будет неизменной по всем параметрам, а опорный и восстанавливающий пучки (E_1 и E_2) будут на порядок расширены при сохранении их полной мощности?

1. ЭЛЕКТРООПТИЧЕСКИЙ ЭФФЕКТ ПОККЕЛЬСА

Оптически анизотропные среды

Поведение световой волны в прозрачном диэлектрике определяет в первую очередь фазовая скорость распространения волны, или показатель преломления. Во многих материалах фазовая скорость волны не зависит от направления ее распространения (от направления волнового вектора) – такие материалы изотропны. Изотропны газы, жидкости, стекла, поликристаллические материалы, кубические кристаллы, аморфные вещества. Анизотропны многие монокристаллы неорганических и органических веществ, а также неодородно напряженные или деформированные упругие материалы (стекло, ситалл, керамика, полимеры).

Анизотропия показателя преломления – следствие анизотропии диэлектрической восприимчивости (поляризуемости) вещества.

Зависимость показателя преломления материала от направления волнового вектора света геометрически отображают на *оптическую индикатрису* – в общем случае трехосный эллипсоид, уравнение которого в определенной системе координат имеет так называемый канонический вид, то есть квадратичная форма в левой его части приведена к главным осям,

$$a^{2}x_{0}^{2}+b_{2}y_{0}^{2}+c^{2}z_{0}^{2}=1,$$
 (II1.1)

где коэффициенты $a = 1/n_{x0}$, $b = 1/n_{y0}$ и $c = 1/n_{z0}$ имеют смысл главных скоростей распространения световых волн. Главные оси индикатрисы X_0 , Y_0 , Z_0 принято обозначать так, чтобы $n_{x0}, < n_{y0} < n_{z0}$.

Сечение, проходящее через центр оптической индикатрисы ортогонально оси Z_0 , есть эллипс с полуосями n_{x0} и n_{y0} . При определенном угле поворота Ψ секущей плоскости вокруг оси Y_0 форма сечения превратится в круг радиуса n_{y0} . Аналогичное преобразование будет и при повороте на угол – Ψ . Направления, ортогональные к этим круговым сечениям (см.рис. 1, δ), называют оптическими осями анизотропной двухосной среды.

Существуют и одноосные среды. У них индикатриса имеет форму эллипсоида вращения, и ось вращения совпадает с оптической осью среды. Одна из полуосей любого проходящего через центр сечения такой индикатрисы неизменна и обычно обозначается как n_o , а другая полуось, обозначаемая n_e , в разных может меняться от n_o до n_e .

В одноосной среде два из трех показателей преломления равны, поэтому уравнение ОИ

 $o^2(x_0^2+y_0^2)+e^2z_0^2=1$, где $o=1/n_o$, $e=1/n_e$, описывает эллипсоид вращения.

В изотропной среде показатель преломления один (a = b = c = 1/n), ОИ представляет собой сферу, уравнение которой ($x^2 + y^2 + z^2$)/ $n^2 = 1$.

В произвольной системе координат вид уравнения ОИ сложнее. В общем случае двухосной анизотропной среды

 $a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + 2a_{23}yz + 2a_{31}xz + 2a_{12}xy = 1$, (П1.2) где a_{ij} называют поляризационными константами среды. Для кристаллов в качестве такой произвольной системы координат обычно выбирают кристаллографические оси, которые в общем случае не совпадают с главными осями ОИ. Кристаллографические оси предопределены симметрией кристалла для каждого из 32 классов.



Рис. П1.1. Трехосный эллипсоид показателей преломления (*a*); оптические оси индикатрисы (б); сечение, ортогональное направлению распространения света (*в*)

Прохождение света через анизотропную среду

Для света в анизотропной среде важно сечение индикатрисы, ортогональное направлению распространения света. Если оно име-

ет форму эллипса, то его полуоси n_1 и n_2 определяют единственно возможные направления колебаний поля в световой волне. Волна с длиною λ , поляризованная вдоль полуоси n_1 , пройдя в среде путь L, испытает задержку по фазе $\delta_1 = (2\pi L/\lambda)n_1$, а вдоль полуоси n_2 задержка $\delta_2 = (2\pi L/\lambda)n_2$, соответственно (λ – длина волны). При прочих поляризациях свет в анизотропной среде распространяется в виде пары взаимно-ортогональных когерентных линейнополяризованных волн, каждая из которых имеет свою, отличную от другой волны, фазовую скорость.

Такое различие фазовых скоростей волн (и показателей преломления) для двух взаимно-ортогональных поляризаций называют двулучепреломлением. Только при распространении света вдоль оптической оси двулучепреломление не проявляется. В остальных случаях одна волна будет отставать по фазе от другой, и на выходе из анизотропной среды между волнами возникнет разность фаз

$$\varphi = (2\pi L/\lambda)(n_2 - n_1)$$

На практике применяют термин "оптическая разность хода",

$$\Delta L = L(n_2 - n_1)$$

При прочих равных условиях разность фаз максимальна для света, распространяющегося ортогонально к оптическим осям.

Рассмотрим плоскую монохроматическую волну $E = E_0 \sin \omega t$, падающую на одноосную слабопоглощающую анизотропную среду протяженностью L (рис. П1.1). Оптическая ось среды направлена вдоль оси Y, свет распространяется вдоль оси Z. Плоскость поляризации падающего света составляет угол α с оптической осью среды и угол ϕ с осью пропускания поляризатора A – A. Тогда амплитуды двух частичных линейно–поляризованных волн, прошедших через поляризатор, будут иметь вид

 $E_x = E_0 \sin\alpha \sin(\phi - \alpha) \cos(\omega t - 2\pi L n_x / \lambda),$ $E_y = E_0 \cos\alpha \cos(\phi - \alpha) \cos(\omega t - 2\pi L n_y / \lambda),$

где n_x и n_y – показатели преломления среды для направлений вдоль осей X и Y, соответственно. В этом представлении интенсивность прошедшего через поляризатор света

 $W \propto (E_x + E_y)^2 = W_0 [\cos^2\beta - \sin2\alpha \sin2(\phi-\alpha) \cos^2\phi/2],$ (П1.3) где $\phi = (2\pi L/\lambda)(n_x - n_y)$ – разность фаз, очевидным образом зависящая от длины кристалла *L* при $(n_x - n_y) \neq 0.$

Общее выражение (П1.1), формальное и плохо обозримое, для самых важных частных случаев существенно упрощается. На прак-

тике выбирают угол $\alpha = 45^{\circ}$, чтобы две волны имели одинаковые амплитуды; только при этом могут быть максимальными изменения мощности прошедшего через поляризатор света (глубина модуляции) при изменениях ($n_x - n_y$).

Угол ф с той же целью выбирают равным 90° или 0°. Выражение (1) для этих случаев принимает вид

$$W = W_0 \sin^2 \varphi / 2 = W_0 (1 - \cos \varphi)$$
(Π1.4)

$$W = W_0 \cos^2 \varphi / 2 = W_0 (1 - \sin \varphi) \tag{\Pi1.5}$$

соответственно.

Когерентная суперпозиция выходящих из анизотропной среды волн образует суммарную волну в общем случае с эллиптической поляризацией, и соотношение и ориентация осей эллипса зависят при этом от разности фаз. Изменение состояния поляризации света можно выявить, пропуская свет через поляризатор. Так модуляцию поляризации света (ее также иногда называют угловой модуляцией) преобразуют в амплитудную модуляцию.

В зависимости от разности фаз поляризация излучения на выходе из анизотропного материала, изменяется от линейной, совпадающей по ориентации с исходной, через непрерывно меняющуюся эллиптическую поляризацию до круговой (при $\phi = \pi/2$), и далее снова через эллиптическую до линейной (при $\phi = \pi$), но уже ортогональной по отношению к исходной (см. рис. П1.2).

φ	0	π/4	$\pi/2$	3π/4	π
$\Delta \nu L$	0	$\lambda/8$	$\lambda/4$	3λ/8	$\lambda/2$
Вид поляри- зации					
φ	π	5π/4	3π/2	7π/4	2π
$\Delta \nu L$	λ/2	5λ/8	3λ/4	7λ/8	λ
Вид поляри-					

Рис.2. Зависимость вида поляризации от ϕ и ΔnL

Линейный электрооптический эффект

Внешние электромагнитные поля и механические напряжения могут вызвать искусственную или изменить существующую оптическую анизотропию; эти изменения можно описать как деформацию и поворот оптической индикатрисы.

Для модуляции оптического излучения чаще всего используют эффект Поккельса – линейное изменение показателя преломления кристалла под действием внешнего электрического поля. Такой эффект возможен только в кристаллах, элементарная ячейка которых не имеет центра симметрии. Зависимость $\Delta n(E)$ бывает линейной и в кристаллах без двулучепреломления; тогда $\Delta n = bE$ (где в общем случае напряженность поля E – вектор, а b – тензор). Несимметрия кристалла по отношению к инверсии означает, что какие-то его наблюдаемые свойства зависят от *знака* направления. Поясним это простыми примерами.

В одномерной бесконечной цепочке из атомов

...ACBCACBCA...

видна симметрия по отношению к инверсии, то есть к перемене знака координаты атома в цепочке; началом координат, над которыми можно совершить инверсию, может быть и A, и B, в зависимости от того, как назначено элементарное звено цепочки (ячейка решетки) – ...BCACB... или ...ACBCA...

В цепочках из атомов

...АВССВАССАВ... или ...АВСВАСВАС...

есть симметрия по отношению к инверсии; в первой цепочке центр инверсии расположен между узлами С, а во второй – в узле С.

В цепочках из атомов

```
...АВССАВССАВ... или ...АВСАВСАВС...
```

симметрии по отношению к инверсии нет; перемена знака координат атомов в цепочке при любом выборе начала координат меняет хотя бы один атом на другой, ему не эквивалентный.

У всех известных кристаллов, проявляющих эффект Поккельса, при отсутствии внешних воздействий главные оси оптической индикатрисы совпадают с кристаллографическими осями. Поэтому уравнение эллипсоида ОИ имеет канонический вид. Под действием внешнего электрического поля ОИ поворачивается и деформируется, оставаясь эллипсоидом. Уравнение ОИ в таком случае сводится к зависимости (П1.2). Поляризационные константы *a_{ij}* линейно зависят от напряженности поля;

$$\begin{aligned} a_{11} &= a^2 + r_{11}E_x + r_{12}E_y + r_{13}E_z \\ a_{22} &= b^2 + r_{21}E_x + r_{22}E_y + r_{23}E_z \\ a_{33} &= c^2 + r_{31}E_x + r_{32}E_y + r_{33}E_z \\ a_{23} &= r_{41}E_x + r_{42}E_y + r_{43}E_z \\ a_{31} &= r_{51}E_x + r_{52}E_y + r_{53}E_z \\ a_{12} &= r_{61}E_x + r_{62}E_y + r_{63}E_z, \end{aligned}$$

где E_x , E_y , E_z – компоненты электростатического поля вдоль кристаллографических осей, r_{ij} – электрооптические коэффициенты среды. Матрицы r_{ij} для трех важнейших классов симметрии приведены ниже.



Рис. П1.3. Изменения индикатрисы под действием электростатического поля *E*, параллельного оптической оси *Z*, в электрооптических кристаллах трех наиболее употребительных классов. Пунктиром изображены деформированные индикатрисы

На рис. П1.3 показаны деформации оптических индикатрис под действием электрического поля E, направленного вдоль оптической оси Z; так чаще всего бывает на практике.

Отметим, что линейность a_{ij} – не закон природы, а следствие разложения зависимости $a_{ij}(E)$ в степенной ряд и сохранения в нем только линейной части приращения. Сами величины a_{ij} представляют собой коэффициенты инвариантной квадратичной формы – левой части уравнения индикатрисы, и их можно отобразить на «правильный» симметричный тензор, у которого диагональные элементы совпадают с a_{ii} , а недиагональные равны $a_{ij}/2$. Отметим также, что электрооптические коэффициенты r_{ij} не образуют тензор; матрица r_{ij} имеет размерность 6×3, а правильный электрооптический тензор – размерность 3×3×3. Тем не менее, по традиции (и чаще всего) при расчетах пользуются матрицей r_{ij} , в которую входят только 18 независимых элементов «правильного» тензора.

Из 18 коэффициентов r_{ij} от нуля отличны лишь некоторые, в соответствии с симметрией конкретного кристаллографического класса. Различных классов всего 32, среди них в 20 классах нет центра инверсии, и у представителей десяти классов (из 20) обнаружен сильно выраженный эффект Поккельса. Вид матриц r_{ij} для наиболее часто применяемых кристаллов трех классов симметрии показан ниже.

Самыми распространенными электрооптическими кристаллами долго были дигидрофосфаты с общей химической формулой MeH₂PO₄, где Me – щелочной металл или аммонийная группа, а водород часто замещен на дейтерий. Дигидрофосфаты в свое время нашли широкое применение в модуляторах для видимого диапазона, несмотря на гигроскопичность и хрупкость. Их относительно несложно выращивать гидротермальным методом (из водного раствора), и их широко применяли в пьезоэлектрических датчиках, задолго до появления лазеров.

Эти кристаллы по типу симметрии относят к классу $\overline{42m}$ (среди его элементов симметрии присутствуют зеркально-поворотные оси 4 порядка, поворотные оси второго порядка, плоскости симметрии). В этом классе отличны от нуля только три электрооптических коэффициента: $r_{41} = r_{52}$ и r_{63} . Поэтому уравнение ОИ этих кристаллов имеет вид

 $a^2(x^2 + y^2) + c^2z^2 + 2r_{41}(E_xyz + E_yxz) + 2r_{63}E_zxy = 1;$ (П1.6) внешнее поле превращает одноосный кристалл в двухосный. Возможны разные варианты взаимной ориентации векторов поля E, волнового вектора k, вектора поляризации света E и кристаллографических осей.

Рассмотрим наиболее частый на практике случай, когда вектор поля E направлен вдоль оптической оси Z кристалла. При этом уравнение ОИ имеет вид

$$a^{2}(x^{2} + y^{2}) + c^{2}z^{2} + 2r_{63}Exy = 1.$$
 (II1.7)

После поворота первоначальной системы кристаллографических координат на 45⁰ вокруг оси *Z* получится новая система координат, в которой уравнение ОИ примет канонический вид

$$(x'/n_x)^2 + (y'/n_y)^2 + (z'/n_z)^2 = 1$$

где новые главные показатели преломления линейно зависят от напряженности поля *E*:

$$n_{x'} = (a^2 - r_{63}E)^{-1/2} = n_o + r_{63}n_o^3 E/2,$$

$$n_{y'} = (a^2 + r_{63}E)^{-1/2} = n_o - r_{63}n_o^3 E/2,$$

$$n_{z'} = n_z = n_e.$$

Для наглядности рассмотрим кристалл в естественной огранке. Для класса $\overline{4}2m$ он имеет форму четырехгранной призмы с четырехгранными пирамидами, прилежащими к основанию (рис. 4, *a*).

Если свет распространяется вдоль оптической оси и вектора приложенного поля E, то этот случай (Z||E||k) называют *продольным эффектом Поккельса*. Анизотропный элемент в таком случае представляет собой пластинку, вырезанную ортогонально оптической оси кристалла (пластинка "*z*-среза"; см поз. *1* на рис. П1.4, *a*). При падении на такой элемент света, поляризованного вдоль оси X^{7} , в кристалле распространяются две взаимно-ортогонально поляризованные волны, разность фаз между которыми на выходе из кристалла

 $\varphi = (2\pi L/\lambda)(n_{z'} - n_{x'}) = (2\pi L/\lambda)r_{63}n_0^3 E = (2\pi/\lambda)r_{63}n_0^3 U$, (П1.8) где L – толщина пластинки, U – приложенное к ней электрическое напряжение. Напряжение, при котором $\varphi = \pi$, называют *полуволно-вым напряжением*, U_{π} .

Для продольного эффекта Поккельса

$$U_{\pi} = \lambda / 2r_{63} n_0^{-3} \tag{\Pi1.9}$$

Отсюда видно, что U_{π} для продольного эффекта Поккельса отображает электрооптические свойства кристалла, а не устройства, изготовленного на его основе. Поэтому часто U_{π} как параметр кристалла приводят в справочниках.

Например, при $\lambda = 0.55$ мкм и $r_{63}=10^{-9}$ см/В, $n_0 = 1.47$ (KDP) полуволновое напряжение $U_{\pi} = 9$ кВ.



Рис. П1.4. Кристалл класса 42m (дигидрофосфаты щелочных металлов и аммония): *a* – форма естественной огранки и варианты срезов кристаллов для модуляторов; *б* – деформация оптической индикатрисы под действием внешнего электрического поля

Очевидно также удобное выражение

$$\varphi = \pi U/U_{\pi} \tag{\Pi1.10}$$

Случай распространения света вдоль одной из наведенных осей, например, Y', когда волновой вектор и электрическое поле взаимно-ортогональны, называют поперечным эффектом Поккельса. Соответствующую пластинку с размерами $L \times B \times B$ для этого вырезают так, как показано на рис. 4, поз. 2. В этом случае проходящий через кристалл свет можно разложить на компоненты, поляризованные по осям X' и Z. Пройдя через кристалл, они приобретут разность фаз

$$\varphi = (2\pi L/\lambda)(n_{z'} - n_{x'}) = (2\pi L/\lambda)[n_e - (n_o + r_{63}n_0^3 U/2B)] = = \varphi_0 - \pi U/U_{\pi}, \qquad (\Pi 1.11)$$

где $\phi_0 = (2\pi L/\lambda)(n_e - n_o)$ – разность фаз, вызванная естественным двулучепреломлением, и полуволновое напряжение $U_{\pi} = \lambda B / r_{63} n_0^{-3} L$. Очевидно, что при поперечной геометрии U_{π} может быть в L/2B раз меньше, чем при продольной – при уменьшении В растет Е, а увеличение *L* теперь влияет на ϕ .



Рис. П1.5. Изменения сечения индикатрисы, ортогонального оптической оси Z кристалла класса симметрии $\overline{4}2m$ (KDP и др.), под действием электростатического поля E, параллельного оптической оси Z. Пунктиром изображено сечение, которое получится при перемене знака Е

Кроме кристаллов класса 42m, сильный линейный электрооптический эффект проявляют кристаллы кубической системы (классы 43m и 23), сегнетоэлектрики–перовскиты (классы 3m и 4mm), кристаллы тетрагональной системы (класс 6mm) и другие. Сегнетоэлектрические перовскиты в настоящее время, по-видимому, наиболее подходят для электрооптических модуляторов. Достаточно прочные механически и негигроскопичные, они позволяют получить низкие (до десятков вольт) полуволновые напряжения.

Рассмотрим эффект Поккельса в кристаллах класса 3m, к которому относятся ниобат лития LiNbO3, танталат лития LiTaO3, прустит Ag₃AsS₃ и другие. Ниобаты и танталаты прозрачны в видимой и ближней ИК частях спектра (от 0,4 до 4 мкм).

Кристаллы класса 3т одноосные и имеют восемь ненулевых электрооптических коэффициентов:

 $r_{13} = r_{32}, r_{22} = -r_{12} = -r_{61}, r_{42} = r_{51}, r_{33}.$

Практический интерес представляет случай, когда поле Е направлено вдоль оптической оси кристалла ($E_x = E_y = 0, E_z = E$). Уравнение индикатрисы в этом случае имеет вид: $(a + r_{13}E)(x^2 + y^2) + (c + r_{33}E)z^2 = 1.$



Рис. П1.5. Пространственная структура элементарной ячейки сегнетоэлектрического перовскита на примере LiNbO₃

При такой ориентации поля кристалл остается одноосным и сохраняет направление оптической оси при любом *E*. Под действием электрического поля изменяются только показатели преломления

$$n_o' = n_o - r_{13}n_o^3 E/2$$

 $n_e' = n_e - r_{33}n_e^3 E/2$

Если свет распространяется вдоль оптической оси Z, то независимо от поляризации света он приобретает фазовую задержку распространения, полное значение которой, $\delta = (2\pi L/\lambda)(n_0 - r_{13}n_o^3 E/2)$, содержит постоянную составляющую $(2\pi L/\lambda)n_o$ и управляемую электрическим полем переменную компоненту $(2\pi L/\lambda)(r_{13}n_o^3 E/2)$. Такая стереометрия оптимальна для ФМ излучения с непостоянной или не заданной заранее поляризацией; индекс модуляции в этом случае не зависит от поляризации (но поляризация излучения на выходе, вообще говоря, отличается от поляризации на входе).Если же свет распространяется ортогонально оптической оси Z, то можно получить управляемую электрическим полем разность фаз между обыкновенной и необыкновенной волнами,

 $\varphi = (2\pi L/\lambda)[(n_0 - n_e) + (r_{33}n_e^3 - r_{13}n_o^3)U/2B] = \varphi_0 + \pi U/U_{\pi},$ (П1.13) При этом полуволновое напряжение равно

$$U_{\pi} = (\lambda B) / (r_{33}n_{\rm e}^{3} - r_{13}n_{\rm o}^{3})L.$$

Для LiNbO₃ параметры $r_{13} = 8 \cdot 10^{-10} \text{ см/B}, r_{33} = 3 \cdot 10^{-9} \text{ см/B}, n_o \cong n_e \cong 2,3;$

При B = L на длине волны 0,63 мкм полуволновое напряжение $U_{\pi} \cong 2,5$ кВ; уменьшая отношение B/L, полуволновое напряжение

снижают до сотен вольт в объемных ЭОМ ($B/L \sim 0,1$) и до единиц вольт в тонкопленочных ЭОМ ($B/L \sim 10^{-2}$) для аппаратуры оптоволоконной связи.

У некоторых сегнетоэлектрических перовскитов заметно изменение показателя преломления при больших плотностях потока энергии светового пучка. При этом уменьшается глубина модуляции. Причиной этого нежелательного явления считают фотогенерацию свободных носителей заряда, их последующую диффузию и захват на уровни-ловушки, где подвижность заряда пренебрежимо мала. Поле возникающего при этом объемного заряда изменяет показатель преломления кристалла за счет того же эффекта Поккельса. Поэтому, например, для LiNbO₃ плотность световой энергии не должна превышать 1 BT/см², а для LiTaO₃ – 10 BT/см². Дозированные добавки Mg, Zn, In, Sc уменьшают этот *диффузионный фоторефрактивный эффект* до приемлемого уровня.

Связанные заряженные состояния электронов в ловушках часто имеют время жизни порядка минут и более при комнатной температуре. Следовательно, медленно релаксируют и поля объемных зарядов (ловушек), искажающие распределение управляющего электрического поля в модуляторе. Релаксацию поля объемного заряда сильно ускоряет повышение температуры кристалла или действие бегущей волны излучения; в обоих случаях электроны выбиваются из ловушек, и объемные заряды рассасываются. Подробнее диффузионный фоторефрактивный эффект и его полезные применения описаны в разделе 7.

Отметим, что стехиометрия монокристаллов сегнетоэлектриковперовскитов заметно отклоняется от их химической формулы; при всех условиях роста в кристаллах LiNbO₃ существует избыток ниобия и дефицит лития. Оптимальное сочетание электрооптических свойств имеют кристаллы, в которых избыток Nb⁺⁵ составляет примерно 1.5 ат.%; недостаток Li⁺¹ при этом достигает 6.0 ат.%.

Установлено, что зависимости электрооптических коэффициентов и фоторефракции у кристаллов LiNbO₃:Zn от количества легирующего цинка немонотонные: ход зависимостей резко меняется в окрестностях ~1.5 ат.% Zn и ~6.0 ат.% Zn. Причины этих явлений – тонкие особенности замещения ионов ниобия ионами примеси и заполнения вакансий ионами лития, выявленные прецизионными структурными исследованиями.



Рис. П1.6. Структура кристалла LiNbO3. Атомы кислорода образуют гексагональную плотноупакованную подрешетку. Атомы Nb регулярно заполняют треть октаэдрических пустот в подрешетке кислорода; еще треть регулярно заполняют атомы Li; треть октаэдрических пустот остается свободной

Термокомпенсированные ЭОМ

Покажем, как при термокомпенсации суммируются эффекты управления и компенсируется влияние естественного двулучепреломления.

Показатели преломления можно представить в виде

$$n_x(E_y) = n_y(E_y) - \Delta n(E_y) + \theta + \Upsilon,$$

где θ зависит от температуры, но не от напряженности поля *E*, а Υ – постоянная, равная разности показателей преломления при нулевой напряженности поля и при начальной температуре, где мы можем положить $\theta(T_0) = 0$. Тогда

$$n_x(-E_y) = n_y(-E_y) - \Delta n(-E_y) + \theta + \Upsilon,$$

но из-за линейности электрооптического эффекта $\Delta n(-E_y) = -\Delta n(E_y)$.

В таком случае задержка по фазе для горизонтально поляризованной компоненты света после прохода через оба кристалла

$$\delta_1 = (2\pi L/\lambda)[n_x(E_y) + n_y(-E_y) + \theta] =$$

= $(2\pi L/\lambda)[n_y(E_y) + n_y(-E_y) - \Delta n(E_y) + \theta + \Upsilon] =$
= $(2\pi L/\lambda)[-\Delta n(E_y) + \theta + \Upsilon],$

а для вертикально поляризованной компоненты

$$\delta_2 = (2\pi L/\lambda)[n_x(-E_y) + n_y(E_y) + \theta] =$$

= $(2\pi L/\lambda)[n_y(-E_y) + n_y(E_y) + \Delta n(E_y) + \theta + \Upsilon] =$
= $(2\pi L/\lambda)[\Delta n(E_y) + \theta + \Upsilon];$

общая разность фаз на выходе будет

$$\varphi = \delta_1 - \delta_2 = (2\pi L/\lambda)[-2\Delta n(E_y)] = (2\pi U/U_\pi).$$

Из окончательного выражения выпадают θ, зависящее от температуры, и Ү, создающее постоянную задержку между двумя интерферирующими волнами. Следовательно, компенсация резко уменьшает зависимость пропускания ЭОМ от температуры и выравнивает задержки интерферирующих волн по фазе, что позволяет работать не только с лазерным излучением, но и со светом, у которого длина когерентности невелика (единицы длин волн λ). Отметим также, что управляемые задержки (фазы) интерферирующих волн при изменении U получают приращения с противоположными знаками; следовательно, после поляризатора изменяется только амплитуда суммарной волны, но не приращение ее фазы, а поэтому компенсированный ЭОМ не вносит в АМ-волну сопутствующую паразитную ФМ.

Инженерно-технические особенности устройства ЭОМ

Для улучшения качества и надежности электрооптических модуляторов принимают особые технические меры. Для снижения потерь света все оптические грани тщательно полируют и наносят на них просветляющие покрытия. Хрупкие и гигроскопичные кристаллы (типа ADP, KDP) помещают в иммерсионную жидкость (кедровое масло, вазелиновое масло, глицерин, и др.). Показатель преломления жидкости выбирают близким к показателям преломления кристалла и окошек кюветы, что уменьшает потери на отражение.



Рис. П1.7. Варианты конструкции электродов у модулирующих элементов с продольным эффектом Поккельса

Для эффективного создания однородного электрического поля в электрооптических кристаллах электроды либо наносят непосредственно на кристаллы (распылением в вакууме или химическим осаждением металла), либо прижимают к ним механически. Особые требования предъявляют к электродам в модуляторах с продольным полем – они должны быть прозрачными для света. В таком случае электроды либо имеют отверстия для прохода света (рис. П1.7, а), либо имеют вид полосок, охватывающих торцы кристалла (рис. П1.7, б). Общий недостаток таких электродов – неизбежная неоднородность поля в кристалле. Более удачное решение – прозрачные проводящие пленки окислов цинка, олова или индия непосредственно на оптических поверхностях (рис. 1.7. в), особенно если нужен модулятор с большой апертурой для широких световых пучков виде тонкой пластины. Такие электроды очень прочны и в диапазоне длин волн от 0,4 до 20 мкм вносят световые потери менее 10%. Иногда, особенно при испытаниях опытных образцов или новых материалов, используют жидкостные электроды - проэлектролиты проводящие органические гелизрачные или полимеры.

Устройство электрооптических модуляторов для волоконных линий связи и устройств интегральной оптики подробнее описано в приложении 2.

2. МОДУЛЯТОРЫ ДЛЯ ВОЛС

Излучением инжекционных лазеров ВОЛС в принципе можно управлять непосредственно током накачки, и эту возможность использовали в первых, примитивных системах оптической связи; их используют и ныне при невысоких требованиях к информационной мощности линии. Локальные линии связи с применением светоизлучающих диодов также работают без специальных модуляторов. Управление током излучателя вполне удовлетворительно при частоте повторения импульсов до 1 ГГц. При более высоких частотах повторения импульсы излучения инжекционного лазера искажаются, расщепляются на части, появляется влияние предыдущего импульса на последующий; кроме того, в спектре излучения лазера могут появляться лишние компоненты, попадающие в соседние частотные каналы. Поэтому в линиях с большими потоками информации применяют быстродействующие электрооптические модуляторы, формирующие оптические сигналы из выходного излучения лазеров, работающих в режиме непрерывной генерации со стабилизацией ее параметров.



Рис. П2.1. Электрооптический модулятор типа Маха-Цендера с противофазными управляющими напряжениями в двух плечах; вариант с управлением бегущей волной напряжения; электрические вводы и переходы на оптические волокна не показаны Основная схема модулятора для ВОЛС – модулятор (интерферометр) Маха-Цендера (ММЦ) с управляемой задержкой в одном плече. На выходе интерферометра мощность излучения зависит от разности фаз (задержек) в его плечах. Для управления задержкой применяют кристаллы материалы с линейным электрооптическим эффектом Поккельса.

Интерферометр Маха-Цендера в модуляторе ВОЛС образован двумя волноводными Y-разветвителями (рис. 111), сформированными у поверхности электрооптического кристалла. Один из вариантов формирования волноводов – диффузионное легирование ниобата лития титаном (примесь внедрения) или протонно-лучевая бомбардировка (замещение подвижного лития водородом). В одноосном кристалле LiNbO₃ используют поперечный эффект Поккельса; оптическая ось ориентирована ортогонально плоскости волноводов.

На один планарный волновод нанесены пленочные электроды, к которым подводят управляющее напряжение. Толщина и ширина волноводов составляет единицы микрон, протяженность электродов – миллиметры и менее. При таких параметрах управляемого волновода для изменения пропускания модулятора от максимума до нуля достаточны напряжения в единицы вольт, поэтому для формирования управляющих напряжений пригодны полупроводниковые интегральные схемы с выходной мощностью менее 1Вт.

Вообще схему Маха-Цендера явно или неявно используют во всех быстродействующих электрооптических модуляторах, как в волноводных, так и в кристаллических. В последнем случае исходная волна расщеплена на две компоненты, поляризованные взаимно-ортогонально. Далее идет фазовая модуляция одной или обеих волн (в одном или двух виртуальных плечах) и их объединение при помощи поляризатора (равносильное объединению двух одинаково поляризованных волн на полупрозрачном зеркале).

Чтобы в отсутствии управляющего сигнала излучение из двух плеч приходило в противофазе, и мощность на выходе ММЦ была нулевой, к одному плечу ММЦ нужно приложить статическое напряжение, при подстройке которого устанавливают начальную разность фаз между плечами, близкую к 180°. Поэтому управляющие электроды связаны с формирователем сигнала только по переменной составляющей, через разделительные конденсаторы, а с источником напряжения смещения электроды гальванически связаны через индуктивности.

В более совершенных вариантах ММЦ электроды нанесены на оба планарных волновода, а управляющие напряжения к ним подводят в противофазе. При этом можно сократить длину электродов (и емкость образованного ими конденсатора), что повышает быстродействие модулятора за счет ускорения перезарядки емкости электродов при заданном ресурсе управляющих оконечных усилителей по выходному току. Самое важное достоинство двухплечевого (антисимметричного, балансного, парафазного) ММЦ – отсутствие фазовой модуляции, сопровождающей модуляцию интенсивности, то есть исключение «чирпинга». Об этом подробнее сказано в следующем разделе.

Фазовая модуляция в амплитудном ММЦ. Чирпинг

ММЦ с одним управляющим плечом (несимметричный ММЦ) модулирует не только интенсивность проходящей волны, но и ее фазу. При быстрой модуляции на фронтах управляющих импульсов возникает «чирпинг» (*chirping* – птичья трель, щебет), то есть нестационарный сдвиг частоты. Частота колебаний ω есть производная фазы φ по времени, $\omega = d\varphi/dt$, и чем быстрее меняется фаза, тем сильнее сдвигается частота при фазовой модуляции. В результате спектр прошедшего излучения дополнительно уширяется, а дисперсия на трассе вносит дополнительные искажения, тем более сильные, чем выше скорость передачи данных, протяженность трассы, дисперсия групповой скорости и ее наклон.

Чирпинг неизбежно возникает при прямой модуляции интенсивности излучения инжекционного лазера, и связан он со специфическими свойствами активной среды инжекционного лазера. По этой причине, в частности, применение модуляции излучения током накачки инжекционного лазера в оптоволоконной связи ограничено линиями с невысокой пропускной способностью.

Если в ММЦ, кроме парафазной фазовой модуляции, будет присутствовать и симметричная (синфазная) модуляция интенсивности в обоих плечах, то в выходном излучении появится дополнительный чирпинг, не связанный непосредственно с работой идеального ММЦ, но вызванный вариациями общей задержки в модуляторе; эта общая задержка будет средним арифметическим двух задержек в отдельных плечах. В соответствии с этим можно оценивать индекс паразитной фазовой модуляции и уширение спектра модулированного излучения за счет чирпинга.

Вообще влияние чирпинга на качество передачи сигналов по трассам с дисперсией существенно зависит от вида модуляции при формировании оптического сигнала в передатчике и от способа его обработки и декодирования в приемнике. С другой стороны, чирпинг можно вводить в оптический сигнал преднамеренно, для предварительной (опережающей) коррекции дисперсионного искажения сигнала на трассе; поэтому правильнее считать вредным фактором не чирпинг сам по себе, а чирпинг неуправляемый, не учитываемый на стороне приема и обработки сигнала. В двуплечем ММЦ управляемый чирпинг может быть и положительным, и отрицательным, в зависимости от соотношения управляющих напряжений в двух плечах модулятора. Ограничение его магнитуды, то есть индекса ФМ, вносят ограничения управляющих напряжений из-за конечной электрической мощности электронных устройств управления и/или ограниченного напряжения электрического пробоя структуры электрооптических ячеек модуляторов (~ 10...30 В, в зависимости от материала и толщины структуры).

Парафазный ММЦ – функциональный аналог двухкристального компенсированного электрооптического модулятора на основе эффекта Поккельса, тоже не вносящего «чирпинг» в модулируемое излучение. Описанный вариант ММЦ с *парафазным* управлением (то есть с парой управляющих напряжений, в двух секциях изменяющихся в противофазе) наиболее гибок в управлении, так как в нем можно отдельно настроить управляющие напряжения, таким образом получив нужные коэффициент модуляции и чирпинг.

Существуют антисимметричные ММЦ с иной схемой управления. Если электрооптический кристалл ориентирован так, что вектор напряженности управляющего электрического поля можно направить параллельно плоскости волноводов (например, в ниобате лития оптическая ось должна лежать в плоскости волноводов и быть ортогональной к их направлению), то для антисимметричного управления нужно подать напряжение на пару средних электродов (рис. П2.2), а пара крайних электродов должна быть «заземлена». Здесь антисимметрия возникает из—за того, что в одном плече вектор статической поляризации кристалла и вектор напряженности управляющего поля параллельны, а в другом – антипараллельны. Такую схему управления на практике применяют особенно часто, так как для нее нужен только один управляющий усилитель мощности, а на модуляцию работают оба плеча ММЦ. Коррекция чирпинга при этом, разумеется, невозможна. Однофазное управление менее гибко в сравнении с парафазным, но конфигурация электрических проводников и межсоединений при этом проще, что удешевляет структуру.



Рис. П2.2. Электрооптический модулятор типа Маха-Цендера с однофазным управляющим напряжением в двух плечах; вариант с управлением в сосредоточенном конденсаторе; электрические вводы и переходы на оптические волокна не показаны

Модуляторы, управляемые бегущей волной напряжения

Допустимая длина активной электрооптической секции модулятора ограничена протяженностью формируемого модулятором импульса; время пролета фронта импульса через активную область должно быть заметно меньше длительности самого импульса, иначе эффективность модулятора падает. Например, в модуляторах для линий со скоростью передачи 40 Гбит/с длина активной секции должна быть менее 1..2 мм.

С другой стороны, чем меньше длина активной секции, тем большее напряжение на электродах необходимо для сохранения нужного индекса фазовой модуляции. Для цифровых систем индекс модуляции β , в зависимости от выбранного формата модулированного оптического сигнала, должен быть равен $\pi/2$ или π ; соответствующее индексу $\beta = \pi/2$ напряжение между электродами, U_{π} , называют *полуволновым напряжением*.

В одном из вариантов ММЦ на ниобате лития при $U_{\pi} = 5,5$ В рабочий диапазон частот достигает 4,5 ГГц на уровне 3дБ. При таких параметрах возможна скорость передачи данных до 10 Гбит/с.

Чтобы сохранить невысокое полуволновое напряжение, применяют электрооптическую модуляцию *бегущей волной* напряжения. При правильном выборе геометрии электродов они могут работать не как конденсатор, а как микрополосковая линия с *бегущей волной напряжения*; в таком режиме, при близости скоростей света в волноводе и волны напряжения между электродами, быстродействие модулятора будет максимальным. Опыт показывает, что в модуляторах с бегущей волной напряжения индекс модуляции на частоте 40 ГГц падает не более чем на 1...2 дБ в сравнении с частотой 10 ГГц; в модуляторах с простыми электродами (простой конденсатор) индекс модуляции на частоте 10 ГГц падает на 3 дБ в сравнении с частотой 1 ГГц.

Чтобы получить режим бегущей волны, к удаленному от генератора концу микрополосковой линии нужно подключить согласованную нагрузку (поглотитель волны), а отношение ширины микрополосковой линии к толщине ее изоляции должно быть не слишком большим, чтобы ее волновое сопротивление было в удобном диапазоне 10..300 Ом. При малом волновом сопротивлении ток в микрополосковой линии будет нерационально большим. При большом волновом сопротивлении в линии вырастут потери энергии на излучение и могут возникнуть трудности при согласовании линии с выходом устройства управления. Для наносекундной и СВЧ микроэлектроники 50 Ом приняты как стандарт сопротивления всех входных и выходных цепей и линий связи, поэтому и для модуляторов ВОЛС волновое сопротивление 50 Ом предпочтительно.

Полупроводниковые модуляторы с эффектом Поккельса

Существуют ММЦ на объемных прямоугольных волноводах с микронными размерами сечения, изготовленные методами полупроводниковой микросхемотехники из полупроводников A^3B^5 . Класс симметрии этих кристаллов – 43m (кубическая ячейка), поэтому в отсутствие внешнего электрического поля они оптически изотропны, но при этом в них достаточно силен линейный электрооптический эффект Поккельса (ячейка без центра инверсии), который проявляется как наведенное электрическим полем двойное лучепреломление. Выбор материала (GaAs, InP, InAs и т. п.) зависит от рабочей длины волны.

Отметим, что при фазовой модуляции в полупроводниковых структурах проявляется паразитная АМ. Причина этого эффекта заключена в том, что электрооптический эффект в полупроводнике тем сильнее, чем ближе длина волны излучения к краю полосы поглощения; на практике, выбирая геометрию и материал модулятора при заданной длине волны излучения, с учетом технических ограничений амплитуды модулирующего напряжения, обычно идут на компромисс между эффективностью ФМ и влиянием паразитной АМ. Кроме полупроводниковых ММЦ, существуют полупроводниковые модуляторы *интенсивности*, основанные на изменении поглощения света под действием электрического поля (эффекты электропоглощения). Такие модуляторы довольно эффективны, проще в сравнении с ММЦ на электрооптических кристаллах, но они вносят чирпинг в модулируемое излучение.

Фазовая модуляция, сопутствующая модуляции интенсивности, в модуляторах интенсивности на электропоглощении обязательно порождает «чирпинг». При модуляции интенсивности чирпинг неизбежен. Поглощение в полупроводнике под действием электрического поля всегда изменяется в результате сдвига полосы поглощения или ее края. При этом обязательно меняется и показатель преломления полупроводника на рабочей длине волны, вблизи от края полосы поглощения. При коэффициенте модуляции интенсивности, близком к единице, индекс сопутствующей фазовой модуляции может быть тоже порядка единицы, а это при быстрой модуляции и порождает сильный «чирпинг». Разумеется, чирпинг можно устранить в парафазной схеме ММЦ с чистой фазовой модуляцией, не сопровождаемой симметричной модуляцией поглощения в обоих плечах.

Эффекты электропоглощения можно использовать как для модуляции интенсивности, так и для фазовой модуляции лазерного излучения. Эти два варианта различаются соотношением между рабочей длиной волны и положением края поглощения (см. далее на рис. П.2.5).

Электрооптические эффекты в полупроводниках.

В современных полупроводниковых модуляторах интенсивности на основе электропоглощения (МЭП) чаще всего используют эффект Франца-Келдыша в области объемного заряда запертого *p-n*-перехода, или на порядок более сильный квантово-размерный эффект Штарка.

Электрооптический эффект Франца-Келдыша проявляется как изменение формы и положения края полосы поглощения полупроводника, примыкающей к окну прозрачности, под действием статического электрического поля. Физический механизм эффекта Франца-Келдыша можно описывать как туннельный переход электрона внутрь запрещенной зоны (рис. П2.3), где его кинетическая энергия становится *отрицательной* (при сохранении полной энергии), а волновой вектор – мнимым (то есть волновая функция становится экспоненциально убывающей функцией координаты),

$$w \propto \exp\left(-\frac{|E-E_x|}{\Delta E}\right), \quad \Delta E = \frac{3}{2}\sqrt[3]{\frac{(e\hbar\mathcal{E})}{m^*}}$$

Край полосы поглощения при увеличении напряженности поля становится более пологим, а склон смещается в длинноволновую область (см. далее рис. П2.5). В этих условиях становится возможным поглощение фотона с энергией $hv < E_g$, и при фиксированной энергии hv сечение поглощения растет с увеличением напряженности электростатического поля, постепенно приближаясь к сечению перехода зона-зона.

Квантово-размерный эффект Штарка представляет собой уширение пика экситонного поглощения при увеличении напряженности поля. Экситоном называют связанное кулоновским притяжением состояние электрона проводимости и дырки, сходное с атомом водорода, но с существенно меньшей энергией ионизации. Поглощение света в полупроводнике может идти не только на переходах между валентной зоной и зоной проводимости, с образованием пары свободных носителей заряда, но и с образованием экситона. Узкий пик экситонного поглощения расположен в окне прозрачности и очень близок к краю поглощения на межзонных переходах.

Штарковское расщепление и вызванное этим уширение экситонной полосы поглощения свойственно всем полупроводникам, но в многослойной квантово-размерной структуре (у соседних слоев различается ширина запрещенной зоны) эти эффекты проявляются на порядок сильнее, чем в объемном полупроводнике.



- Рис. П2.3. Туннелирование электронов в сильном электростатическом поле є в собственном полупроводнике;
 - a прямой переход под треугольный барьер (ширина d, высота E_g),
 - δ «косой» переход с поглощением фотона под треугольный барьер (ширина $d^{\prime} < d$, высота $E_{e} hv$) эффект Франца-Келдыша



Рис. П2.4. Движение экситонов в электростатических полях: *a* – поле объемного заряда, разрыв связи; *б* – потенциал деформации, движение экситона в сторону сужения запрещенной зоны; *в* – потенциал деформации совместно с полем объемного заряда;

г - накопление экситонов в локальной потенциальной яме

Расщепление экситонного пика поглощения под действием электростатического поля аналогично расщеплению спектральных линий атомов и молекул (эффект Штарка). Принято считать, что существенный вклад в уширение экситонного пика дает распад экситона при увеличении напряженности поля (оно разрывает разно-именные заряды, рис. П2.4, a). Напротив, в неоднородном полупроводнике, где локально ширина запрещенной зоны зависит от координаты, экситон как целое движется в сторону сужения (рис. П2.4, δ). Такие локальные изменения ширины запрещенной зоны называют *потенциалами деформации*, так как они формируются в областях механических напряжений и деформаций. Кроме того, «потенциалы деформации» образуются и при локальных изменениях химического состава полупроводника, в гетеропереходах.



Рис. П2.5. Изменение спектра поглощения в оптическом волноводе с многослойной квантово-размерной структурой на основе InP под действием приложенного управляющего напряжения (квантово-размерный эффект Штарка). Показаны рабочие длины волн для режимов МЭП и ФМ

В области с шириной запрещенной зоны меньшей, чем в ее окрестности, экситоны могут накапливаться, как в потенциальной яме. По этой причине в квантоворазмерных ямах, в условиях ограничения поступательного движения, интенсивность экситонных пиков растет за счет накопления экситонов и увеличения времени их жизни. Не менее важно, что энергия связи в основном состоянии у двумерного экситона в 4 раза больше, чем у трехмерного. Этот вывод следует из решения задачи двух тел (водородоподобного атома) в двумерном и в трехмерном вариантах. Вообще энергия связи экситона зависит от эффективных масс электрона и дырки и от диэлектрической проницаемости полупроводника; по аналогии с атомом водорода, модифицированные «боровский» радиус и «постоянная Ридберга»

$$a_B = \frac{4\pi\varepsilon_r \hbar^2}{m_r e^2}, \quad R_y = \frac{m_r e^4}{2\hbar^2 (4\pi\varepsilon_r)^2}$$
где $m_r = \frac{m_e m_h}{m_e + m_h}; \quad m_e \ll m_h, \rightarrow m_r \approx m_e.$

В типичных полупроводниках $\varepsilon_r > 10$, а приведенная масса электрона и дырки примерно в десять раз меньше массы свободного электрона. В соответствии с этим, характерный размер экситона порядка или более 100 Å, а энергия связи $E_1 = R_y$ в тысячи раз меньше, чем в атоме водорода. Такой экситон называют экситоном Ванье-Мотта; их наблюдают при низких температурах, $kT < R_y$.

В трехмерном полупроводнике (3D): $E_{s} = -\frac{R_{y}}{n^{2}}$. В двумерной квантовой яме (3D): $E_{s} = -\frac{R_{y}}{(n-1/2)^{2}}$

В квантовых ямах во многих полупроводниках энергия связи (основное состояние экситона, $E_1 = 4R_y$) достигает 10...20 мэВ. Поэтому в квантовой яме экситонные пики заметны и при комнатной температуре (kT = 25 мэВ), тогда как, например, в объемном GaAs они видны только при температурах ниже 100 К.

При наложении электрического поля экситонный пик поглощения уширяется и смещается в область длинных волн; смещение пропорционально *квадрату* напряженности поля, как и при обычном квадратичном эффекте Штарка в спектрах атомов и молекул. Вместе с этим за счет эффекта Франца-Келдыша происходит и смещение в длинноволновую область края полосы межзонного поглощения.

Далее речь будет идти о модуляторах на основе эффекта Франца-Келдыша, но в равной мере то же самое верно и для модуляторов на основе квантоворазмерного эффекта Штарка.



Рис. П2.6. Зависимость пропускания микроволноводного модулятора интенсивности от управляющего напряжения: *a* – МЭП, *б* – ММЦ

Создать сильное электростатическое поле с напряженностью порядка 10 кВ/см, где проявляются указанные эффекты, даже в чистом собственном полупроводнике практически невозможно из-за присутствия собственных носителей заряда при ненулевой температуре. Свободные заряды под действием электрического поля создают токи проводимости и экранируют внешнее поле. Более сильно эффект Франца-Келдыша проявляется в областях обеднения полупроводника, в частности, в области электростатического поля объемного заряда в запертом *p-n*-переходе. Толщину области обеднения подбором концентрации примесей, формирующих переход, согласуют с толщиной оптического волновода модулятора. Так получают отрезок оптического волновода с управляемым пропусканием.

Уширения и сдвиги полос поглощения влекут за собой обязательные и закономерные изменения показателя преломления в окне прозрачности. Эти изменения наиболее сильны вблизи от края полосы поглощения; при всем этом изменения затухания на выбранной длине волны, на удалении от края поглощения, могут быть небольшими. Таким образом, изменения поглощения в стороне от частоты излучения (в области прозрачности) влекут за собой фазовую модуляцию излучения. Для описания этих эффектов используют дисперсионные соотношения Крамерса-Кронига, связывающие между собой действительную и мнимую части восприимчивости вещества (т.е. преломление и затухание, соответственно).

Модуляторы интенсивности на основе электропоглощения (МЭП) работают на длине волны излучения, близкой к краю поглощения. Фазовые модуляторы и ММЦ могут работать на длинах волн, удаленных от края поглощения в область прозрачности (к более длинным волнам), так что модуляция в каждом плече может быть чисто фазовой. Правда, полуволновые напряжения в последнем случае существенно выше, чем перепады напряжения, нужные для модуляции интенсивности проходящего излучения на 10...20 дБ в той (коротковолновой) области спектра, где наблюдают поглощение. Рассмотрим этот вопрос подробнее.

Эффективность ФМ и ММЦ обычно отображают полуволновым напряжением. При чистой ФМ индекс модуляции (при фиксированном управляющем напряжении) пропорционален длине области, где происходит ФМ, поэтому полуволновое напряжение для ММЦ обратно пропорционально длине такой области.

Эффективность МЭП удобно отображать как управляющее напряжение, при котором пропускание изменяется на 10 дБ (то есть от максимального, принятого за 100%, до уровня 10%). При модуляции поглощения при увеличении длины модулятора затухание света растет экспоненциально. Следовательно, при прочих равных условиях изменение проходящей мощности света на 10 дБ можно получить на меньшей длине при использовании электропоглощения, нежели в ММЦ. Чем меньше длина электродов, тем меньше их электрическая емкость; следовательно, перезарядка этой емкости при той же электрической мощности управления протекает быстрее. Поэтому в модуляторах с электропоглощением на частотах до 20 ГГц обычно нет нужды использовать электроды с бегущей волной, тогда как в ММЦ нередко такая потребность есть.

По названным причинам МЭП в современной технике конкурируют с ММЦ на электрооптических кристаллах, превосходя их по быстродействию и эффективности электрического управления. Их важнейшие (но терпимые) недостатки – более сложное сопряжение с оптоволокном и большие (свыше 10 дБ) потери, вносимые в оптический тракт. Кроме того, часто приходится мириться с некоторой паразитной ФМ, обязательно сопровождающей АМ; этот чирпинг при рациональном выборе режимов работы МЭП можно использовать как средство предварительной коррекции сигнала, компенсирующей его дисперсионные искажения на линии связи.

Полупроводниковые модуляторы излучения в интегральной оптике

Неоспоримое достоинство полупроводниковых модуляторов – возможность их естественной интеграции с лазером на одной полупроводниковой подложке и возможность формирования всей интеграции в единой последовательности технологических операций. Некоторые неудобства возникают из-за влияния модулирующих сигналов непосредственно на лазер при недостаточно хорошей электромагнитной изоляции, так как лазер и модулятор иногда разделяют лишь микронные зазоры.

Если к этому добавить интеграцию в один полупроводниковый чип электронных схем управления и общее термостатирование модуля в едином корпусе, то техническое решение в целом можно приблизить к оптимуму почти по всем показателям. Электронные схемы удобнее делать на основе кремния, но лазер на кремниевой подложке пока проблема, и возможности создания в кремнии гетеропереходов с нужными параметрами ограничены. Современные структуры инжекционных лазеров для оптоволоконной связи чаще всего выращивают на подложках InP и GaAs. В этих материалах не всегда уместны структуры типа биполярных транзисторов и даже диодов, так как в прямозонных полупроводниках рекомбинационное излучение в *p-n*-переходах может создавать нежелательные перекрестные связи между элементами схемы, нарушая ее правильную работу. Альтернативой могут быть структуры МДП, J-FET и MESFET (последние просто незаменимы в субнаносекундной микроэлектронике).

Обычно полупроводниковые модуляторы объединяют на одной подложке вместе с инжекционным лазером в общем модуле передатчика, и существенное подобие структур волноводов лазера и ММЦ в таком случае позволяет формировать их в едином технологическом процессе. Здесь для устранения «чирпинга» ММЦ тоже делают антисимметричными. Отметим, что в интегральном исполнении практически одинаковы стоимость одного ММЦ и стоимость одноплечего фазового модулятора. Поэтому ММЦ как функциональную группу можно считать базовым элементом интегральнооптических модулей для амплитудной и фазовой модуляции; ныне их все чаще используют совместно для формирования оптических сигналов со сложными типами модуляции, применяемыми в современных оптоволоконных системах связи. Из всех изложенных здесь довольно общих соображений следует, что инженерные разработки модуляторов для ВОЛС должны быть основаны на физических принципах эффективной фазовой модуляции, обеспечивающих высокое быстродействие, а работоспособность на частотах модуляции 10..20 ГГц нужно считать безусловным и минимальным требованием. При этом учитывают и особенности того типа модуляции, для которого данный модулятор предназначен.

Современная технология квантово-размерных полупроводниковых структур предлагает технике связи альтернативу – модуляторы интенсивности на основе эффектов электропоглощения (МЭП). Такие модуляторы испытаны в лабораториях и освоены, по крайней мере, в опытном производстве. Будущее покажет, в какой мере достоинства МЭП (в первую очередь – быстродействие, низкие мощности управления, простота интеграции в оптоэлектронные модули) перевесят их недостатки (большие начальные потери и чирпинг), и смогут ли МЭП вытеснить ММЦ.

3. ФАЗОВАЯ МОДУЛЯЦИЯ СВЕТА В ФОТОУПРУГОЙ СРЕДЕ

Фотоупругость. Основные термины и определения

Изменение показателя преломления при механической деформации упругой прозрачной среды называют фотоупругостью.

Если в кристалле создать механическое напряжение σ_j, то возникнет деформация:

 $U_i = s_{ij}\sigma_j$, i = 1, 2...6, j = 1, 2...6,где s_{ij} – коэффициенты сжимаемости. Здесь и далее, как это принято в тензорной алгебре, мы подразумеваем суммирование по повторяющимся индексам, опустив сам знак суммы.

Механическая деформация анизотропного материала в общем случае может привести к деформации и к изменению пространственной ориентации эллипсоида оптической индикатрисы (ОИ). Изменение поляризационных констант можно записать в виде

$$\Delta a_k = p_{kj} u_j = \pi_{kj} \sigma_j$$

где p_{kj} называют тензором упругооптических коэффициентов, а π_{kj} – тензором пьезооптических коэффициентов. В отличие от эффекта Поккельса, свойственного только пьезокристаллам, фотоупругость при определенных условиях наблюдают практически во всех прозрачных материалах, в том числе и в изотропных. Более того, фотоупругость свойственна и жидкостям, но только при быстропеременных напряжениях (звуковые и ультразвуковые колебания), а статические напряжения в жидкости, разумеется, быстро исчезают из-за ее текучасти.

Далее мы будем рассматривать фотоупругость изотропного материала. Уравнение оптической индикатрисы в нем

$$a_0(x^2+y^2+z^2) = 1.$$

Для изотропной среды в силу симметрии отличны от нуля только следующие упругооптические коэффициенты:

$$p_{11} = p_{22} = p_{33};$$

 $p_{12} = p_{21} = p_{13} = p_{31} = p_{23} = p_{32}; \ p_{44} = p_{55} = p_{66} = p_{11} - p_{12}.$

При деформации u_j уравнение ОИ принимает вид $a_1x^2 + a_2y^2 + a_3z^2 + 2a_{11}yz + 2a_5zx + 2a_6xy = 1.$ Коэффициенты *a_k* можно найти из следующего матричного уравнения:

$$\begin{vmatrix} a_{1} - a_{0} \\ a_{2} - a_{0} \\ a_{3} - a_{0} \\ a_{4} \\ a_{5} \\ a_{6} \end{vmatrix} = \begin{cases} p_{11} & p_{12} & p_{12} & 0 & 0 & 0 \\ p_{12} & p_{11} & p_{11} & 0 & 0 & 0 \\ p_{12} & p_{12} & p_{11} & 0 & 0 & 0 \\ p_{12} & p_{12} & p_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & p_{11} - p_{12} & 0 & 0 \\ 0 & 0 & 0 & 0 & p_{11} - p_{12} & 0 \\ 0 & 0 & 0 & 0 & 0 & p_{11} - p_{12} \end{vmatrix} \begin{bmatrix} u_{1} \\ u_{2} \\ u_{3} \\ u_{4} \\ u_{5} \\ u_{6} \end{bmatrix}$$

Если создана одноосная деформация в таком направлении, что отлична от нуля только одна компонента $u_1=u$, то уравнение ОИ имеет вид

$$(a_0+p_{11}u)x^2+(a_0+p_{12}u)(y^2+z^2)=1.$$

Из него следует, что под действием одноосной механической нагрузки изотропная среда становится одноосной двулучепреломляющей, и направление оптической оси совпадает с направлением деформации.

Выражения для показателей преломления

$$n_e = (a_0 + p_{11}u)^{-1/2} \approx n - p_{11}un^3/2,$$

$$n_o = (a_0 + p_{12}u)^{-1/2} \approx n - p_{12}un^3/2.$$

При распространении света ортогонально к наведенной оптической оси в такой среде на длине L можно получить разность фаз между обыкновенной и необыкновенной волнами

$$\varphi = (2\pi L/\lambda)(n_e - n_o) = (\pi Lun^3/\lambda)(p_{12} - p_{11}) = \pi Lun^3 p_{44}/\lambda.$$

Связь индекса модуляции света в ультразвуковой волне с ее механическими параметрами

Традиционная форма выражения фотоупругости как изменения диэлектрических свойств среды имеет вид

$$\Delta\left(\frac{1}{\varepsilon}\right) = \Delta\left(\frac{1}{n^2}\right) = p\sigma, \implies \Delta n = -n^3 p\sigma/2 \tag{II3.1}$$

где ε – диэлектрическая проницаемость, p – коэффициент фотоупругости, σ – механическое напряжение; все эти три величины в общем случае тензоры.

Интенсивность ультразвука пропорциональна квадрату эффективного механического напряжения σ

$$w_s = \rho v^3 \overline{\sigma}^2 / 2 \tag{\Pi3.1}$$

где *v* – фазовая скорость звука, *р* – плотность вещества.

Амплитуда эффективного механического напряжения

$$\overline{\sigma} \propto \sqrt{w_s} = \sqrt{w_s / 2\rho v^3}. \tag{II3.1}$$

Изменение длины оптического пути при изменении показателя преломления

$$\Delta n \mid L = n^3 p \overline{\sigma} / 2 = (n^3 p \sqrt{w_s / 2\rho v^3}) / 2. \qquad (\Pi 3.1)$$

Традиционно (и в какой-то мере формально) принято вводить коэффициент фотоупругого качества вещества

$$M \equiv \frac{n^6 p^2}{\rho v^3}; \tag{II3.5}$$

тогда индекс модуляции принимает традиционный вид

$$\beta = \frac{\pi L}{\lambda} \sqrt{M w_s / 2} = \frac{\pi}{\lambda} \sqrt{\frac{L M P_s}{2S}} , \qquad (\Pi 3.6)$$

где P_S – полная мощность ультразвуковой волны, S – поперечная ширина ультразвукового столба.

К сожалению, мощность ультразвука трудно измерять непосредственно; легче измерять электрическую мощность, подведенную к пьезоэлементу, и управлять ею. Поэтому (ПЗ.6) разумнее представить в виде

$$\beta = \frac{\pi}{\lambda} \sqrt{\frac{LMP_E F}{2S}}$$

где P_E – подведенная электрическая мощность, F – коэффициент преобразования, в общем случае зависящий от частоты возбуждения и иных факторов, то есть не всегда достоверно известный.

Рассеяние Брэгга: сохранение энергии и импульса

Векторная диаграмма. У световой волны, рассеянной бегущей ультразвуковой волной, частота колебаний смещена по отношению к частоте исходной волны света. Эти сдвиги частоты происходят и в режиме Раман-Ната, и в режиме Брэгга. Их можно объяснить доплеровским смещением частоты за счет движения рассеивателя, а также как результат фазовой модуляции излучения, при которой в спектре фазомодулированного колебания появляются боковые составляющие, сдвинутые по отношению к несущей на частоты, кратные модулирующей частоте. Оба представления этих эффектов по существу тождественны, так как эффект Доплера описывает
сдвиг частоты как результат прогрессирующего во времени изменения фазы колебаний, а сдвиг частоты есть производная по времени от приращения фазы колебаний.

Можно описать сдвиги частоты и изменения направлений волновых векторов рассеянных волн с единой и удобной точки зрения. Когерентное рассеяние света на ультразвуковой волне можно рассматривать как столкновение фотона с фононом (квантованным колебанием кристаллической решетки); при таком столкновении должны сохраняться энергия и импульс системы "фотон + фонон". При таком анализе энергию обычно отождествляют с частотой, а импульс – с волновым вектором. (Подумайте, на чем основана законность такой замены переменных не только для фотонов, но и для квантов звука, колебаний упругой среды).

Требования сохранения энергии и импульса в рассматриваемом нами взаимодействии приобретают форму

$$\kappa_1 = k + K, \ \omega = \omega_1 \pm \Omega. \tag{II3.7}$$

Физик-теоретик немедленно уточнит, что нужно ввести единичный вектор \mathbf{n} в направлении \mathbf{K} и записать эти требования в обобщенном виде:

для стоячей волны с длиной Λ

 $\mathbf{k}_{1} = \mathbf{k} + \mathbf{n} \times 2\pi m / \Lambda, \ \omega = \omega_{1}; \tag{II3.8a}$

для бегущей волны

$$\mathbf{k}_{1} = \mathbf{k} + \mathbf{K} + \mathbf{n} \times 2\pi m / \Lambda, \ \omega = \omega_{1} \pm m \Omega \tag{13.86}$$

но пока мы рассматриваем только порядки интерференции с $m = \pm 1$, можно использовать и укороченные выражения (ПЗ.7).

Приведенные соотношения (ПЗ.8) универсальны и в равной мере относятся ко всякому взаимодействию, где пара волн порождает третью или наоборот, одна волна разваливается на две. При этом неважно, о каких волнах идет речь; это могут быть волны де Бройля, описывающие любые реальные квантовые частицы, или волны намагниченности в кристалле ферромагнетика.

Сохранение импульса в названных взаимодействиях по сути тождественно условиям Брэгга, а сохранение энергии – синхронности колебаний поля световой волны и решетки во времени. Сохранение импульса должно выполняться с той точностью, с которой он определен в рассматриваемой системе; если волна (точнее, волновой пакет, группа волн) имеет в пространстве протяженность ΔX по какой-то из координат, то неопределенность проекции импульса $\Delta K_X \approx 1/\Delta X$. Аналогично, есть и соотношение неопределенностей для частоты и времени; при рассеянии света на ультразвуке оно связывает уширение спектральной линии рассеянного излучения со временем *когерентного* взаимодействия света и ультразвука. (Остерегайтесь грубой ошибки! Время когерентного взаимодействия – это не то краткое время, за которое свет пересекает ультразвуковой столб!)



Рис. ПЗ.1. Векторные диаграммы, отображающие рассеяние Брэгга на ультразвуке в изотропной среде: *а* – первый порядок, *б* – минус первый порядок

При взаимодействии света с ультразвуком без заметной потери точности можно считать, что $k_1 = k$, так как $\omega \gg \Omega$. Это приближение, впрочем, не стоит применять безрассудно; все-таки правильнее $k_1 = k + \delta$, но $|\delta| \ll k$. При рассеянии, вообще говоря, изменяется и частота, и направление волнового вектора света, но длину волнового вектора в первом приближении можно считать неизменной. Такое приближение позволяет решать задачи о когерентном рассеянии света на ультразвуке как задачи из школьной тригонометрии. Рассмотрим одну из важнейших для нас.

Рассеяние в первый и минус первый порядок. Зафиксированы длина волнового вектора ультразвука **К** и длина волнового вектора падающей световой волны **k**. Длина волнового вектора рассеянной волны \mathbf{k}_1 , несмотря на ее сдвиг по частоте, очень мало отличается от длины **k**. Поэтому концы волновых векторов лежат на окружно-

сти радиуса $2\pi/\lambda$ (рис. ПЗ.1), и их соединяет вектор **К**. Отметим, что **К** по смыслу есть базисный вектор одномерной обратной решетки, **n**× $2\pi/\Lambda$.

Равенство длин волновых векторов при выполнении условия Брэгга для указанных интерференционных порядков влечет за собой и равенство углов падения $\psi_{\rm b}$ и рассеяния $\theta_{1,-1}$. Говоря строже, углы $\psi_{\rm b}$ и $\theta_{1,-1}$ различаются знаками.

Другая поучительная задача, сводимая к сложению импульсов, рассмотрена в следующем разделе.

Рассеяние Брэгга в двулучепреломляющем материале

Диаграмма направленности. В анизотропной для света среде (например, в двулучепреломляющем кристалле) без потери точности уже нельзя полагать k₁ = k, так как при почти одинаковых энергиях волны с разной поляризацией для разных направлений все-таки имеют разные фазовые скорости. В изотропном материале мы не обращали внимания на поляризацию света (говоря строго, это было не совсем верно). В анизотропном материале рассеяние на ультразвуке в общем случае может идти и так, что падающая и рассеянная волны имеют разные поляризации. Тогда существенно $k_1 \neq k$, и теперь k_1 и k определены длиной волны света в вакууме и фазовыми скоростями обыкновенной и необыкновенной волн в кристалле. Для этих случаев векторные диаграммы дают особенно наглядные подсказки для расчетов. Рассмотрим один из типичных случаев – модель акустооптического дефлектора, в котором угол отклонения может быть на порядок больше угла падения (в модуляторе с изотропным материалом эти углы одинаковы).

Рассеяние в первый порядок в одноосном кристалле. Когда падающая и рассеянная волны имеют разные поляризации, то длины их волновых векторов уже неодинаковы. Рассматриваем случай, когда все волновые векторы расположены в плоскости, ортогональной оптической оси отрицательного кристалла ($n_o > n_e$). Схема сложения волновых векторов показана на рис. ПЗ.2, *а*. Из этой схемы видны соотношения

 $k_{1}\cos\theta_{1} = k\cos\psi_{B},$ $K = k\sin\psi_{B} + k_{1}\sin\theta_{1}$



Рис. ПЗ.2. Векторные диаграммы, отображающие рассеяние Брэгга на ультразвуке в одноосном отрицательном кристалле. Показаны схема сложения волновых векторов (*a*) и различие диапазонов углов падения и рассеяния, соответствующих условиям Брэгга (*б*)

Подставив сюда $K = \Omega/V$, $k = 2\pi n_o/\lambda$, $k_1 = 2\pi n_e/\lambda$, систему можно разрешить относительно синусов углов падения и рассеяния. Но прежде мы обратимся к векторной диаграмме рис.ПЗ.2, *б*. На них хорошо видно, как слабо изменяется угол падения, соответствующий условиям Брэгга, и как сильно изменяется угол рассеяния при изменениях частоты ультразвука. Отметим, что изменения угла падения нужно сравнивать с углом расходимости излучения лазера; из диаграмм рис. ПЗ.2 можно увидеть, что существует такой диапазон частот, в котором изменения угла падения не превосходят угол расходимости падающего пучка. В центре этого диапазона Ω_0 угол θ проходит через ноль, и это из диаграммы видно гораздо лучше, чем из явных аналитических выражений:

$$\psi = \arcsin\left[\frac{\lambda\Omega}{2n_oV}\left(1 + \frac{V^2(n_o^2 - n_e^2)}{(\lambda\Omega)^2}\right)\right],\$$
$$\theta = \arcsin\left[\frac{\lambda\Omega}{2n_eV}\left(1 - \frac{V^2(n_o^2 - n_e^2)}{(\lambda\Omega)^2}\right)\right]$$

Далее несложно получить

$$\Omega_0 = \frac{V}{\lambda} \sqrt{n_o^2 - n_e^2}, \quad \psi_0 = \arcsin\left(\frac{\lambda\Omega}{n_o V}\right) = \arcsin\left(\frac{\lambda_o}{\Lambda_0}\right).$$

Диапазон частот ультразвука, в которых выполнены условия Брэгга для обеих волн, *формально* ограничен приведенными ниже соотношениями:

$$\begin{aligned} \frac{\lambda\Omega}{2n_oV} & \left(1 + \frac{V^2(n_o^2 - n_e^2)}{(\lambda\Omega)^2}\right) \le 1; \\ \frac{\lambda\Omega}{2n_eV} & \left(1 - \frac{V^2(n_o^2 - n_e^2)}{(\lambda\Omega)^2}\right) \ge -1; \\ \frac{V^2(n_o^2 - n_e^2)}{(\lambda\Omega)^2} \le \frac{2n_oV}{\lambda\Omega} - 1; \\ \frac{V^2(n_o^2 - n_e^2)}{(\lambda\Omega)^2} \le 1 + \frac{2n_oV}{\lambda\Omega}. \end{aligned}$$

Фактически этот диапазон еще уже, и ограничен он более жесткими условиями – ограниченной апертурой реального устройства в реальной обстановке. Оценки таких ограничений тоже нетрудно получить при помощи векторной диаграммы, похожей на рис. ПЗ.2.

При рассеянии на фазовой решетке в анизотропной среде условия Брэгга особенно отчетливо проявляются как сильная связь волн, волновые векторы которых отличаются в точности на вектор обратной решетки, и как резкое ослабление связи падающей волны со всеми прочими волнами. Заметим, что среди всех волн с \mathbf{k}_1 , отличающимися от \mathbf{k} на вектор обратной решетки $m\mathbf{K}$, волну с $\mathbf{k}_1 = \mathbf{k} + \mathbf{K}$ как единственную выбирает закон сохранения энергии, то есть временной синхронизм колебаний. (На самом деле существует малая вероятность *многофононных* и *многофотонных* процессов, в которых и сохранение энергии видоизменяется, но мы их здесь рассматривать не станем. Эти процессы можно рассматривать и как вторичное рассеяние волны \mathbf{k}_1 на ультразвуке.)

Показанную на рис. ПЗ.3 схему на практике применяют для управляемого отклонения (сканирования) лазерного луча. В таком случае диапазон углов отклонения рассеянного излучения при малых отступлениях от условий Брэгга, когда интенсивность отклоненного луча еще заметно не изменяется, задан углом расходимости излучения лазера; чем этот угол больше, тем шире диапазон углов отклонения, но тем меньше энергетическая эффективность преобразования падающей волны в отклоненную.



Рис. ПЗ.3. Рассеяние света на акустической волне в анизотропном отрицательном кристалле; на врезке показаны экстремальные углы падения, рассеяния и диапазон допустимых волновых векторов ультразвука, при которых выполнимы условия Брэгга

Перестраиваемый акустооптический фильтр

Еще одно полезное применение дифракции Брэгга – управляемая фильтрация излучения. При изменении частоты ультразвука изменяется и условие отклонения рассеянного луча на заданный угол. В перестраиваемом акустооптическом фильтре (ПАФ) коллимированный падающий свет взаимодействует с акустической волной в сравнительно протяженной области пространства. При этом рассеянные световые волны испытывают сдвиги в соответствии с соотношением $\Omega_{\text{расс}} = \Omega_{\text{пад}} \pm \Omega_{\text{акуст}}$.

Как и при АОМ брэгговского типа, требование сохранения импульса проявляется в том, что световая волна, сдвинутая по частоте, должна испытывать и изменение направления. Для эффективного преобразования исходная и сдвинутая по частоте световые волны должны сохранять фазовый синхронизм на длине когерентности. В ПАФ пространственный фазовый синхронизм получают за счет двулучепреломления среды. При этом существует одна и только одна длина волны, для которой исходная и сдвинутая по частоте волны имеют одинаковые фазовые скорости за счет различия направлений, по которым они распространяются. Поэтому в первый интерференционный порядок будет рассеяна волна только одной вполне определенной длины. Изменяя частоту акустической волны, изменяют длину световой волны, отвечающей условию пространственного синхронизма. Таким образом, АОМ может работать как перестраиваемый узкополосный фильтр.

Подобный прибор можно использовать для разделения линий в спектрах излучения атомов или лазеров. Часто ПАФ используют для эффективного разделения неполяризованного белого света на пучки с взаимно-ортогональной линейной поляризацией, и в этом случае условия синхронизма приводят к сдвигам по частоте в разные стороны от исходной двух рассеянных волн с разной поляризацией.

Полоса пропускания подобного фильтра зависит от качества коллимации светового пучка и от длины акустооптического кристалла. Например, типичный ПАФ на основе TeO_2 обеспечивает разрешение от 1 до 2,5 нм, в зависимости от длины волны света, плавность настройки 0,1 нм и воспроизводимость настройки по длине волны лучше 0,05 нм. Для достижения таких характеристик нужно, чтобы угол расходимости света (полный раствор) был не более 5°. Парателлурид TeO_2 как материал для АОМ применяют довольно часто, так как его можно использовать в фильтрах в диапазоне длин волн от 0,3 до 4,5 мкм. Отметим, что весь указанный диапазон единственный ПАФ перекрыть не может, поскольку для этого потребовалось бы в одном приборе возбуждать акустические волны в диапазоне частот шире одной декады, что практически недостижимо.

Упрощенная модель связанных волн

При когерентном рассеянии световой волны в объемной решетке не только падающая волна \mathbf{k} порождает рассеянную \mathbf{k}_1 , но и рассеянная волна \mathbf{k}_1 , в свою очередь, снова преобразуется в волну \mathbf{k} . В таком случае систему волн \mathbf{k} и \mathbf{k}_1 называют связанными волнами или *связанными модами*. При взаимодействии таких связанных волн их энергия вовсе не делится поровну между ними, как это могло бы показаться на первый взгляд. Существенная когерентность взаимодействия между волнами требует учитывать фазовые соотношения между ними во всем объеме перекрытия их полей в пространстве и времени. В объеме ультразвуковой брэгговской решетки связь между волнами тем сильнее, чем больше периодические вариации показателя преломления в пространстве, и связь эта когерентна в той степени, в какой регулярна решетка показателя преломления и когерентна исходная (падающая) волна.

Покажем, как ведут себя амплитуды связанных волн в пространстве. Для этого мы будем мысленно двигаться вместе с фронтом падающей волны \mathbf{k} и следить за амплитудами волн \mathbf{k} и \mathbf{k}_1 во времени. Так как обе волны в среде представляют собой гармонические колебательные процессы, то можно рассматривать колебания на фронте волны просто как колебания некоего гармонического осциллятора, слабо связанного с таким же вторым осциллятором; он теряет и приобретает энергию колебаний, испытывая вынуждающую силу со стороны осциллятора-напарника.

В такой качественной постановке задача очень похожа на колебания в системе из двух слабо связанных маятников одинаковой длины, когда можно пренебречь диссипацией энергии системы в целом. Из опыта известно, что в такой системе при некоторой начальной амплитуде колебаний одного маятника и состоянии покоя второго маятника происходит периодический обмен энергиями колебаний между парой маятников. Существуют моменты времени, когда практически вся энергия системы сосредотачивается поочередно в колебаниях только одного из пары маятников.

Из-за равноценности маятников придется признать, что декремент затухания колебаний одного маятника равен инкременту нарастания колебаний другого. Период взаимного обмена энергиями тем меньше, чем сильнее связь между маятниками, и степень этой связи естественно отображает, например, декремент затухания. Период обмена энергиями можно назвать временем когерентности колебаний, подобно тому, как время когерентности диссипативнозатухающих колебаний отождествлено с характерным временем самого затухания. Отметим, что в нашей модели затухание колебаний одного маятника *не диссипативно* и в каком-то смысле обратимо, поскольку отданная маятником энергия периодически к нему возвращается.

Длительность всего процесса в целом ограничена временем длительного и слабого диссипативного затухания колебаний, то есть временем потери когерентности в системе. Можно показать, что диссипация в определенных условиях за достаточно длительное время может привести к разделению остатка энергии колебаний между маятниками почти поровну. Все это можно выразить в более строгой форме системой двух обыкновенных линейных дифференциальных уравнений, но мы предоставим это сомневающимся читателям. Здесь мы приведем укороченные уравнения для амплитуд пары связанных осцилляторов без затухания. Они укорочены в том смысле, что мы рассматриваем только медленные эволюции во времени амплитуд колебаний $a_i(t)$, полное выражение которых есть $a_i(t)cos(\omega t + \phi_i)$. Применим так называемый метод медленных амплитуд, когда мы полагаем, что $da_i(t)/dt \ll \omega$. Отметим, что эти медленные амплитуды мы рассматриваем как *знакопеременные*, чтобы в них отразились важные эволюции фаз колебаний осцилляторов.

Считая, что скорость изменения амплитуды колебаний одного осциллятора Ω пропорциональна амплитуде колебаний другого, запишем простые уравнения, считая для общности Ω величиной комплексной:

$$\frac{da_1}{dt} = \Omega a_2,$$
$$\frac{da_2}{dt} = \Omega a_1.$$

После дифференцирования по времени первого уравнения и подстановки в него второго уравнения при *вещественном* Ω получим:

$$\frac{d^2a_1}{dt^2} = \Omega^2 a_1,$$

то есть простое уравнение движения гармонического осциллятора с собственной частотой Ω . Такое же уравнение описывает и $a_2(t)$. Далее, задав начальное условие $a_1(0) = 1$, $a_2(0) = 0$, мы получим уже описанные на словах колебания с обменом энергиями.

Мы получим тот же результат, если рассмотрим не амплитуды, а энергии колебаний (квадраты амплитуд). Запишем уравнения баланса энергий двух связанных подсистем:

$$\begin{cases} \frac{dW_1}{dt} = \Omega(W_2 - W_1) \\ \frac{dW_2}{dt} = \Omega(W_1 - W_2) \end{cases} \Rightarrow \begin{cases} \frac{d^2W_1}{dt^2} = \Omega(\frac{dW_2}{dt} - \frac{dW_1}{dt}) \\ \frac{d^2W_2}{dt^2} = \Omega(\frac{dW_1}{dt} - \frac{dW_2}{dt}) \end{cases}$$

откуда следует:

$$\begin{cases} \frac{d^2 W_1}{dt^2} = 2\Omega^2 (W_1 - W_2) \\ \frac{d^2 W_2}{dt^2} = 2\Omega^2 (W_2 - W_1) \end{cases}$$

Из пары исходных уравнений очевидно сохранение энергии системы в целом:

$$\frac{d(W_1+W_2)}{dt}=0$$

Сложение пары последних уравнений дает

$$\frac{d^2(W_1-W_2)}{dt^2} = 4\Omega^2(W_1-W_2),$$

то есть уравнение гармонических колебаний разности энергий с частотой 2Ω. При сохранении их полной энергии подсистемы, очевидно, периодически обмениваются энергиями, как показано на рис. ПЗ.4. Мы пренебрегли потерями энергии системы в целом, поэтому весь процесс передачи энергии от одной части системы к другой задают начальные условия. Нас интересует система с начальным состоянием

$$W_1 = W_0, \quad W_2 = 0,$$

что соответствует возбуждению колебаний второй подсистемы от состояния покоя за счет воздействия на нее первой подсистемы.

Пара связанных волн ведет себя аналогично. Уравнения, которые описывают этот процесс, сводятся к уравнениям модели с парой маятников. Период перекачки энергии между волнами в регулярной решетке тем короче, чем больше индекс фазовой модуляции. Этому периоду в пространстве соответствует определенная протяженность волны, на которой она отдает всю свою энергию связанной с ней условием *фазового синхронизма* (когерентность рассеяния Брэгга!) парной волне. Эту протяженность называют *длиной фазового синхронизма* или *длиной когерентности*.

Заметим, что эти термины появляются везде, где взаимодействует пара связанных колебательных подсистем, и длина когерентности всегда так или иначе соответствует мере взаимосвязи этих процессов и, говоря более обобщенно, степени их коррелированности во времени и/или в пространстве. При этом существенны и энергетические, и фазовые корреляции. Так, взаимосвязь процессов разрушают и диссипация энергии, и искажения фаз. Коэффициент связи Ω может быть и комплексным, то есть обмен амплитудами колебаний происходит с некоторой их взаимной расфазировкой. В этом случае с мнимой частью Ω связано затухание; если система недиссипативная, то амплитуды $a_2(t)$ и $a_1(t)$ выравниваются при $t \to \infty$. В диссипативной системе, естественно, колебания затухают, и $a_i(\infty) \to 0$.



Рис.П3.4. Эволюции во времени амплитуд *a* и мощностей *W* пары связанных волн в фиксированной точке движущегося фронта первичной волны

Исходя из этих простых соображений, качественно легко понять, почему при заданной протяженности области взаимодействия существует оптимум акустической мощности, при котором почти вся энергия переходит из волны падающей в волну рассеянную. Столь же понятно, почему при ограниченной акустической мощности существует минимальная протяженность ультразвуковой решетки в направлении волнового вектора падающей волны, на которой ее энергия эффективно переносится в когерентно-рассеянную.

Вспомнив выражение для мощности волны, рассеянной в режиме Брэгга,

$$W_{-1} = W_0 \sin^2\left(\frac{\pi L \Delta n}{\lambda}\right),$$

и узнав в аргументе индекс фазовой модуляции, через который очевидным образом выражается длина синхронизма ($L = \lambda/(2\Delta n)$), мы можем считать, что предложенная здесь качественная модель удовлетворительна. Тем не менее, дополнительно разъясним, почему мы обращаем особое внимание на индекс модуляции как меру связи двух волн, с несколько иной, энергетической точки зрения.

Рассматривая модель Рамана–Ната, мы замечали, что перекачку энергии падающей волны в рассеянную описывает функция индекса модуляции, при небольшом аргументе качественно похожая на синусоиду. В модели Рамана–Ната локальная модуляция поля была чисто фазовой. Поэтому считать индекс модуляции мерой связи волн при фазовой модуляции вполне разумно; он описывает связь волн по напряженности поля.

В модели Брэгга мы считали решетку почти амплитуднофазовой; коэффициент отражения на градиенте показателя преломления почти очевидно связан с индексом модуляции, так как они оба пропорциональны Δn . Следовательно, и для объемной ультразвуковой решетки индекс модуляции – разумная характеристика связи волн по напряженности поля, тоже описывающая переход энергии от одной волны к другой.

Рассмотренная здесь модель не учитывала связь рассмотренной выделенной пары волн с третьими соучастниками взаимодействия, для которых условия когерентности взаимодействия (условия Брэгга) тоже выполнены. Оставаясь в рамках все тех же модельных представлений, заметим, что для всех прочих рассеянных волн, для которых не выполнено «условие блеска», эффективный индекс модуляции существенно меньше. Волновой вектор такой волны пересекает на протяжении ультразвуковой решетки несколько максимумов и минимумов плотности вещества, и чем их больше на пути, тем меньше индекс модуляции (покажите сами; это совсем просто). Следовательно, связь этих волн с исходной слабее, чем связь последней с волной, для которой «условие блеска» выполнено.

В акустооптических модуляторах индекс модуляции редко достигает $\pi/2$, но может быть очень близким к этому значению; коэффициент преобразования мощности падающей волны в рассеянную может быть около 70...80%. Далее увеличение мощности ультразвука уже не дает существенной добавки к мощности волны, рассеянной в минус первый порядок; вместо этого заметно растет мощность, рассеянная в остальные ненулевые порядки. Поэтому для улучшения эффективности АОМ полезнее уменьшать расходимость и оптимизировать апертуру падающей волны.

4. СОВРЕМЕННЫЕ ТИПЫ МОДУЛЯЦИИ ОПТИЧЕСКИХ ВОЛН^{*}

Каждый из применяемых на практике методов модуляции оптического излучения предназначен для формирования сигнала с нужными свойствами. Особое значение имеют сигналы, отображающие биты цифровой информации в современных системах связи.

Классифицировать различные виды модуляции электромагнитной волны разумно по той физической величине, в изменениях которой закодирован полезный двоичный сигнал. Таких физических величин, условно говоря, три: интенсивность, поляризация и фаза волны. Условность в этом перечислении заключена в том, что интенсивность и поляризацию можно считать информативными характеристиками даже немонохроматического излучения, а фаза, строго говоря, определена только для почти монохроматической волны. Можно причислять к информативным величинам и частоту почти периодических колебаний напряженности поля волны, но с точки зрения теории сигналов частотную модуляцию (ЧМ) можно считать частным случаем фазовой модуляции (ФМ). Приведенная классификация в существенной мере ориентирована на способ детектирования (демодуляции) сигнала, а не только на способ его формирования.

Фазовую модуляцию световых волн часто называют дифференциальной ФМ, подчеркивая самим названием то обстоятельство, что информацию могут нести только изменения фазы, но не сама фаза. Ранее было отмечено, что для выделения сигнала при ФМ необходимо получить разность фаз колебаний в два последовательных промежутка времени или/и в двух точках трассы. Иначе говоря, восстановление ФМ-сигнала основано на интерференции волны с ее же задержанной (или смещенной в пространстве) копией. Это налагает условия на время когерентности несущей волны: оно должно быть больше, чем задержка между двумя интерферирующими компонентами. Формирование копии и задержка между ними могут быть и неявными; так происходит преобразование

^{*} Использован обзор *Advanced Optical Modulation Formats*, Peter J. Winzer and Rene-Jean Essiambre, Proceedings of the IEEE, Vol. **94**, No. 5, May 2006, pp. 952–984.

ФМ/ЧМ в АМ при помощи резонаторов или диспергирующих элементов, и здесь неявно или явно присутствуют разделение волны на части, задержка между ними и суммирование, как и во всяком интерферометре.

В отличие от ФМ, при амплитудной модуляции (АМ), то есть при модуляции интенсивности, сигнал как переменную, зависящую от времени, без разделения волны на части и интерференции между ними, непосредственно выделяет элементарный квадратичный детектор. Можно также отметить, что ноль интенсивности – состояние достаточно определенное, противопоставленное «не нулю», тогда как ноль фазы световой волны в опыте невозможно корректно отличить от «не нуля»; физический смысл, как величины *наблюдаемые*, имеют только разности фаз. Эти принципиальные положения мы считали нужным повторить и подчеркнуть.

В технике оптической связи модуляцию поляризации применяют крайне редко, только в лабораторных исследованиях. Реальное применение находят только модуляция интенсивности и фазовая модуляция. Иерархию разных типов модуляции в оптической связи отображает рис. П4.1. Более подробная классификация внутри трех показанных здесь групп приведена на рис. П4.2, П4.3 и П4.4.



Рис. П4.1. Традиционная классификация форматов модуляции лазерного излучения. Здесь модуляция интенсивности излучения не противопоставлена модуляции напряженности поля

При более подробной классификации внимание обращают и на остальные изменяющиеся при модуляции характеристики. Например, АМ и ФМ могут сопровождать друг друга в одном сигнале, но сигнал относят к классу АМ или ФМ прежде всего по тому, какой способ формирования и/или детектирования для него выбран. Вы-

бор обычно не произволен и направлен на достижение наилучших качественных показателей сквозного тракта «модулятор – канал связи – детектор». Примером этому может быть ФМ, применяемая для формирования ЧМ-сигналов УКВ вещания и телевидения (звук, цветность SECAM). Так как частота есть производная фазы по времени, то применение ФМ равносильно подъему высших частот в сигнале на входе *частотного* детектора, которым компенсируется их завал в самом детекторе и далее в усилителях. Этот технический прием улучшает отношение сигнал/шум в приемнике; он относится к так называемым *предыскажениям* сигнала в передатчике, часто применяемым в самых разных системах связи, в том числе и в оптических, для компенсации искажений, возникающих на разных участках тракта передачи–приема.



Рис. П4.2. Форматы передачи данных с модуляцией интенсивности без запоминания сигнала на приемной стороне.

Во многих случаях применяют модуляцию интенсивности излучения – непосредственно или как один из компонентов сложного формата (рис. П4.2 и П4.3). Для краткости и по традиции будем называть ее амплитудной модуляцией (AM), хотя истинная амплитудная модуляция лазерного излучения в большинстве случаев действует на напряженность поля, а не на интенсивность непосредственно. Предупредим, что на рисунках и в подписях к ним для краткости «напряженностью» названа не истинная напряженность электрического поля световой волны, $E = E_0 \cos(\omega t + \varphi)$, а только множитель E_0 со знаком, так что сдвиг фазы волны на π и перемена знака E_0 эквивалентны. С учетом этого замечания AM напряженности следует понимать как изменение алгебраической величины E_0 размерной. Отметим, что AM напряженности можно рассматривать (и выполнять) как сочетание взаимно синхронных AM и Φ M.

Далее будут встречаться сокращенные названия различных форматов модуляции. Перечень этих сокращений собран в табл. П4.1 в конце данного обзора. Там же на рис. П4.15, П4.16 показаны спектральные представления сигналов с разными типами модуляции, в том числе и не упомянутыми в тексте, но указанными в табл. П4.1 и на рис. П4.2, П4.3.



Рис. П4.3. Форматы передачи данных с модуляцией интенсивности с запоминанием сигнала на приемной стороне.

Форматы модулированных оптических сигналов классифицируют также и по способу кодирования (на стороне передатчика) и восстановления (на приемной стороне) тех изменений параметров волны, которые несут информацию. Поэтому корректнее говорить не о форматах модуляции, а форматах передачи данных по оптическим каналам связи, рассматривая в комплексе формирование сигнала в передатчике, его преобразования на линии передачи, детектирование и первоначальную обработку на приемной стороне как три взаимосвязанных компонента единого процесса передачи данных.



Рис. П4.4. Форматы передачи данных с фазовой модуляцией без запоминания сигнала на приемной стороне



Рис. П4.5. Изменения мощности *W* модулированных волн при передаче AM сигналом последовательности битов 1101: *а* – манипуляция (ООК) без разделения потока битов нулями (NRZ), *б* –манипуляция (ООК) с разделением потока нулями (RZ), *в* – структура передатчика для формата RZ-OOK

Не имея возможности и намерения рассматривать все многообразие форматов модуляции оптических сигналов и их выбор для решения конкретных задач, сосредоточим внимание на важнейших характеристиках и физико-технических методах реализации наиболее выгодных форматов. В частности, важны вопросы о форме спектра модулированного сигнала и об ограничениях, затрудняющих достижение необходимых параметров. За рамками этого обзора придется оставить систематические сравнения их качественных показателей в реальных системах связи.



 Рис. П4.6. Изменения мощности W и фазы φ модулированных волн при передаче последовательности битов 1101:
а – дифференциальная фазовая манипуляция без разделения битов нулями, *δ* – то же, с разделением потока нулями, *в* – AM с подавлением несущей. Темной окраской на графиках W(t) показана фаза напряженности поля (φ)

Современные оптические системы передачи данных используют цифровое представление информации; соответственно и модуляция излучения номинально представляет собой манипуляцию (дискретные переключения амплитуд и фаз). Фактически переходы между уровнями сигналов при номинально дискретном управлении занимают конечное время; оно меньше интервала между последовательными битами информации, но иногда может достигать почти половины этого интервала. На приемной стороне выделенные сигналы чаще всего представляют собой последовательности, близкие к синусоиде с пропусками некоторых периодов (примеры см. на рис. П4.5, П4.6 и П4.7). При потоках порядка 10 Гбит/с даже для такой скорости переключения нужен модулятор с быстродействием не хуже 30 пс, а ширина спектра модулированного излучения достигает десятков гигагерц; эти параметры соответствуют и предельным возможностям электрооптики, и ширине канала систем связи с частотным мультиплексированием (100 ГГц для WDM и 50 ГГц для DWDM).

Простая периодическая последовательность световых импульсов не заключает в себе никакой информации; ее могут нести только отклонения от строгой периодичности. На стороне приемника и декодера так или иначе приходится сопоставлять текущий поток импульсов с некоторой регулярной последовательностью, которую обычно называют импульсами *синхронизации*. Представление о подобной синхронизации дает, например, прием сигналов телевидения, когда из потока сигналов нужно выделять сигналов цветности, сигналы запуска кадровой и строчной разверток изображения, сигналы синхронизации цветности. Во всех этих случаях в современных телеприемниках синхронизирующие последовательности формируют генераторы периодических колебаний, частоту и фазу которых подстраивают на совпадение с определенной компонентой принимаемого потока, выделяемой из него фильтрацией.



Рис. П4.7. Вариант структуры передатчика с подавлением несущей и разделением битов нулями при переменной скважности сигнала по мощности (формат RZ-AMI, или суррогатная балансная модуляция). Исходный АМ сигнал без разделения битов нулями (NRZ) интерферирует со своей задержанной копией и преобразуется в сигнал с чередованием фаз. Задержка Т для подавления несущей настроена на деструктивную интерференцию (антисовпадение исходного и задержанного импульсов)

По аналогии, в системах оптической связи приходится выделять тактовую последовательность (восстанавливать частоту и фазу) повторения импульсов. Располагая такой последовательностью, на приемной стороне можно решать, какой из символов кода пришел к приемнику в данный момент времени.

В общем случае восстановление синхронизирующей последовательности – задача не простая. Если передаваемый поток содержит длинный участок из одинаковых символов, то на таком участке нельзя указать критерий, по которому различаются два или более интервала времени. Такой поток битов смежных называют «не разделенным нулями» (в международной терминологии – NRZ, «non-return-to-zero»; см. пример на рис. П4.5, а). Потоки форматов семейства NRZ сложнее декодировать на приемной стороне, но они, при том же потоке битов информации, занимают спектральный диапазон примерно вдвое меньший, нежели альтернативные потоки RZ, описанные далее (пример показан на рис. П4.5, б). трудности выделения тактовой последовательности из потока NRZ преодолевают, выбирая предварительное кодирование потока битов информации таким, что внутри группы единиц или нулей определенным образом коррелированы между собой фазы и уровни мощности импульсов. Например, в формате DB (синонимы PSBT, PASS, CAPS, см. рис. П4.3) фаза импульса с логическим уровнем 1 меняется на π, если перед ним прошло нечетное число логических (периодов тактовой частоты с нулевой нулей мощностью). В формате АМІ фаза импульса с логическим уровнем 1 меняется на π независимо от количества предшествующих ему логических нулей, в том числе и при смежных логических единицах (рис. П4.7).

Если поток сигналов представляет собой регулярную последовательность импульсов одинаковой интенсивности, но различающихся по частоте или фазе (рис. П4.6, б, в), то выделять тактовую последовательность проще; для этого пригоден простой амплитудный детектор. Такой поток битов называют «разделенным нулями» (в международной терминологии – RZ, «return-to-zero»). В подобной системе логическую информацию разумно кодировать изменением фазы (или частоты) колебаний внутри импульсов, не связывая логические сигналы с интенсивностью импульсов. Для восстановления информации о фазе нужно сравнивать фазы пары последовательных импульсов; для этого выделяют электрический сигнал, который образуется при интерференции текущего и предыдущего (задержанного) импульсов. Так как сдвиг фазы при фиксированной задержке импульса во времени пропорционален несущей частоте, то появляется требование к стабильности несущей, на порядок более жесткое, нежели требования в стабильности несущих в системах связи с частотным уплотнением WDM и даже DWDM. Эти

требования указаны далее, в разделе «Дифференциальная фазовая манипуляция».

Заметим, что в оптическом диапазоне типичен подход к AM как к модуляции интенсивности, а о модуляции амплитуды в чистом виде, без потери информации о фазе, речь идет несколько реже. В некоторых случаях информация о фазе не теряется в процессе модуляции, но ее *игнорируют на стадии приема* сигналов при квадратичном детектировании. Например, в формате CSRZ (балансная AM, см. рис. П4.6, *в*) фазы в соседних импульсах отличаются на π , независимо от логического значения передаваемых битов. Если за счет дисперсии или иных искажений импульсы после прохода через линию связи расширяются, то в области перехода интенсивности через ноль они интерферируют *деструктивно*. Поэтому минимумы интенсивности на приемной стороне выражены четче, и меньше страдают от шумов и дисперсии.



Рис. П4.8. Проявления чирпинга в группе волн;

а – при нормальной дисперсии фазовой скорости на линии передачи, б – то же, при аномальной дисперсии; в – при неидеальной AM с увеличением задержки при увеличении пропускания, г – то же, при уменьшении задержки при увеличении пропускания, д – преднамеренный чирпинг для компенсации дисперсии на линии связи, е – импульс д на выходе диспергирующей линии

Часто AM, как нежелательная (*паразитная*), возникает при ФМ/ЧМ, если амплитудно-частотная характеристика некоторого участка имеет заметную неравномерность в диапазоне спектра, занимаемом сигналом. В свою очередь, в некоторых устройствах AM возникает паразитная ФМ; ее называют *«чирпингом»*. Чирпинг наглядно представлен на рис. П4.8, где показаны распределения напряженности поля E по координате x как «снимок» условно остановленной группы волн; чирпинг показан преувеличенно. (Во временном представлении, на осциллограммах, группы волн будут отображены зеркально по отношению к показанным на рисунках.) Например, чирпинг свойствен амплитудному модулятору типа Маха–Цендера (ММЦ) с ФМ в одном управляемом плече, но компенсирован в ММЦ с парафазной ФМ в обоих плечах. (ММЦ для волоконных линий связи описаны в приложении 2).

В современных высокоскоростных системах передачи данных чирпинг не рассматривают как явление однозначно вредное; при управляемых параметрах чирпинга его можно использовать как контролируемое предыскажение сигнала, компенсирующее *хроматическую дисперсию* канала связи. В оптоволоконных линиях со стандартными одномодовыми волокнами высокочастотные компоненты сигнала движутся с большей *групповой* скоростью, нежели низкочастотные. Для компенсации дисперсии (в первом порядке) нужно снижать частоту на переднем фронте импульса и повышать на заднем. Пример тому приведен на рис. П4.8, *д*, *е*, где показан импульс с чирпингом (*д*) и импульс с компенсированным чирпингом (*е*) после прохода по линии с дисперсии. Этот технический прием обычно используют для достижения минимальной длительности фемтосекундных импульсов лазерного излучения.

Чирпинг, показанный на рис. П4.8, *в*, *г*, труден для компенсации на стороне приемника. Он возникает в модуляторах без компенсации фазовых искажений. Чирпинг может возникать не только в модуляторе передатчика, но и на самой линии связи, как проявление нелинейности среды. При плотности потока излучения, свойственной линии связи с одномодовым оптоволокном, показатель преломления материала волокна изменяется при изменении интенсивности излучения, и это вносит в сигнал некоторую фазовую модуляцию, которой в исходном сигнале не было (дополнительный чирпинг). Чтобы уменьшить влияние подобной *фазовой самомодуляции* излучения, в сигнал передатчика вносят определенный корректирующий чирпинг, наподобие чирпинга для компенсации простой (линейной) дисперсии на линии передачи.

Технически для создания нормированного чирпинга нужен модулятор балансного типа, работающий в режиме с нормированной разбалансировкой двух его плеч по фазе. Следовательно, ММЦ с парафазным (раздельным) управлением в двух плечах для таких применений более удобен, нежели ММЦ с синфазным управлением (с единственным управляющим электродом). К этому добавим, что применение любых предыскажений и методов компенсации дисперсии предполагает настройку под конкретные условия работы канала связи, поэтому на практике выгоднее использовать элементы управления с большим числом степеней свободы.

Избыточное расширение спектра сигнала может сопровождать амплитудную *манипуляцию*, то есть быстрые включения и выключения мощности. Сам термин «манипуляция» указывает на аналогию режима формирования сигнала при помощи телеграфного ключа Морзе (*keying*). В режиме телеграфной манипуляции резкие включения и выключения передатчика на слух воспринимаются как резкие щелчки; а в излучении передатчика они расширяют спектр боковых частот. Для сужения спектра излучения до приемлемого уровня в радиопередатчиках сглаживают фронты нарастания и спада мощности сигнала. В системах оптической связи применяют настолько радикальное сглаживание фронтов модулированных импульсов, что форма огибающей оптического сигнала (интенсивности), как правило, слабо отличается от синусоидальной. При этом коэффициент модуляции выбирают как компромисс между шириной спектра сигнала и помехоустойчивостью связи.

Помехоустойчивость связи существенно зависит от соотношения сигнал/шум на стороне приемника. В современных системах приема данных шумовые параметры аппаратуры обычно соответствуют пределу технических возможностей; близки к этому пределу и параметры линий связи. Мощности передатчиков при этом обычно жестко ограничены сверху техническим регламентом системы связи; в оптоволоконных линиях превышение мощности передатчика искажает и сам передаваемый сигнал, и сигналы в соседних каналах WDM. В условиях таких ограничений единственный доступный инженерам способ достижения наилучшего отношения сигнал/шум – рациональный выбор формата данных.

Одно из технических решений – ограничение спектра частот, излучаемых передатчиком АМ, диапазоном от несущей частоты до границы одной боковой полосы частот (рис. П4.9). В результате вся мощность передатчика, сосредоточенная в одной боковой полосе (ОБП, SSB), расходуется *только* на передачу информации; не нагруженная никакой информацией несущая частота из спектра излучения исключена, и энергия на ее передачу не расходуется. На приемной стороне для восстановления формы сигнала нужно восстановить несущую частоту при помощи местного гетеродина; смешение несущей и принятого сигнала ОБП (на перемножителе или на квадратичном детекторе) полностью восстановит форму исходного сигнала. Отсутствие второй, подавленной боковой полосы, никак не влияет на точность восстановления информации. С этой точки зрения одна их двух боковых полос АМ сигнала *избыточна* по информации. Если учитывать влияние шумов трассы и приемника, то по мощности сигнала на входе приемника формат SSB не дает выгод в сравнении с сигналом формата CSRZ, но выгоднее формата ООК.



Рис. П4.9. Двухполосная и однополосная модуляция, спектральное представление; показаны примеры с подавлением нижней боковой полосы: *a* – двухполосная AM с частичным подавлением несущей при K < 1, *δ* – SSB с частичным подавлением несущей (VSB), *в* – SSB.

Передачу радиосигналов ОБП используют в практике DX/DXX (дальней и сверхдальней радиотелеграфии), где не требуется предельная точность восстановления несущей частоты; при расстройке гетеродина меняется только высота тона «морзянки». В оптическом диапазоне недостаточная стабильность и расстройки гетеродина могут вызвать серьезные затруднения, поэтому здесь выгоднее применять частичное подавление несущей; на передачу ослабленной несущей расходуют часть мощности передатчика, но зато в приемнике всегда присутствует «пилот-сигнал», по которому восстанавливают несущую. Этот прием используют в радиотелефонии, где точное восстановление несущей нужно для неискаженной передачи голосовых сигналов. Наиболее важная область применения радиопередачи с подавлением одной боковой и ослаблением несущей – традиционное эфирное телевещание, где сигнал яркости и синхронизирующие импульсы принимают и детектируют как АМ сигнал ОБП, а частично подавленная несущая служит опорой для автоподстройки частоты гетеродина телевизионного приемника.



Рис. П4.10. Сводка важнейших режимов управления ММЦ, дающих оптические сигналы разных форматов.

 U_{π} – полуволновое напряжение, R – поток логических сигналов (бит/с), T – период следования логических сигналов, f – тактовая частота оптических импульсов. Для форматов RZ указан коэффициент заполнения потока световой мощностью в процентах

Все описанные до сих пор форматы передачи данных можно получить в передатчике с единственным ММЦ, выбрав один из режимов его работы – статическое напряжение настройки U и динамическое модулирующее напряжение ΔU (см. рис. П4.10 и П4.11).

В технике радиосвязи практически все приемники построены по супергетеродинной схеме, с преобразованием сигнала от высокой частоты или СВЧ к относительно невысокой промежуточной частоте. Гетеродинное преобразование частоты в сравнении с прямым детектированием дает на порядки больший коэффициент передачи сигнала от антенны к цепям обработки (фильтрации и пр.) Эту методику удается полностью перенести в оптический диапазон и использовать все хорошо известные преимущества гетеродинного приема. Гетеродинное преобразование сохраняет когерентность сигнала; при этом в электрический сигнал переносится вся информация об амплитуде и фазе светового сигнала. Это, в конечном итоге, позволяет на стороне передатчика нагрузить сигнал большей информацией, а на стороне приемника – оптимизировать обработку, применяя хорошо отработанные радиотехнические решения. При обработке оптических сигналов их перенос в радиодиапазон позволяет применять достаточно развитые методы коррекции и фильтрации радиосигналов, что в конечном итоге заметно снижает вероятность ошибок при передаче информации.

Специалисты Nokia Siemens Networks, CoreOptics (Германия) и Эйндховенского технического университета (Нидерланды) успешно испытали ВОЛС со скоростью передачи сигналов 111 Гбит/с на расстояние 2400 км без усилителей мощности и ретрансляторов. На приемной стороне оптический сигнал был преобразован в электрический фотогетеродинным методом. В результате был получен большой выигрыш в отношении сигнал/шум, присущий гетеродинному приему. Самое важное преимущество испытанной системы – огромная допустимая протяженность участка линии без ретрансляторов, достигнутая в первую очередь за счет гетеродинного преобразования сигнала в приемнике. Итогом развития этих методик становится увеличение пропускной способности существующих линий связи (в том числе и с волокнами устаревших типов) при обновлении только аппаратуры терминалов.

Дифференциальная фазовая манипуляция

Применение фазовой манипуляции в оптическом диапазоне дает несколько практических преимуществ в сравнении с AM; в частности, при этом мощность передатчика используется столь же эффективно, как и при AM с ОБП. В эксперименте, описанном в предыдущем разделе, был применен когерентный прием сигнала в сочетании с дифференциальной квадратурной фазовой манипуляцией (DQPSK). Этот сложный формат ФМ ныне считают наиболее прогрессивным; использование фазы оптического сигнала как признака, несущего информацию, дает возможность в единственном импульсе оптического сигнала передавать *одно из четырех двузначных двоичных чисел*, а не один бит, как при простой AM с двоичным кодированием.



Рис. П4.11. Представление эволюций комплексной напряженности поля при важнейших типах модуляции. Масштабы диаграмм нормированы на единичную оптическую мощность, усредненную по времени. Штриховыми линиями показаны возможные варианты изменения комплексной напряженности при различных типах модуляции: *а* – модуляция интенсивности с чирпингом, *в* – дуга соответствует простому линейному фазовому модулятору, прямая – парафазному ММЦ

Фазовая модуляция и манипуляция от прочих видов модуляции отличается тем, что сама по себе фаза электромагнитной волны – величина не вполне определенная; наблюдаемые физические величины, зависящие от фазы, – периодические функции фазы. Реальный смысл имеют только разности фаз между заданными точками в пространстве и/или в заданные моменты времени. Поэтому информацию в формате DPSK переносит не фаза в импульсе однобитовой посылки, а *изменение* фазы при переходе от одного бита к следующему, как показано на рис. П4.12. Скачку фазы на π обычно приписывают *логическое* значение 1, отсутствию – значение 0. Отсюда следует, что исходные логические данные (тексты в двоичном коде) до их передачи должны быть определенным образом перекодированы. Вопросы предварительного кодирования здесь излагать нет возможности; они не относятся к физической оптике.



Рис. П4.12. Структура передатчика RZ-DPSK и действующие в нем сигналы. Для фазовой модуляции применяют либо простой фазовый модулятор (ФМ), либо модулятор Маха–Цендера (ММЦ); изменения интенсивности и фазы на выходе зависят от типа модулятора

Сигнал с фазовой модуляцией вообще, и с DPSK в том числе, квадратичный детектор обнаружить не может; для фазовой демодуляции необходима интерференция исходного сигнала с его копией, задержанной на длительность одного бита (рис. П4.13), с точностью лучше 10%. Кроме того, задержку необходимо тонко настроить так, чтобы интерференция двух одинаковых по фазе импульсов была либо строго конструктивной, либо строго деструктивной. Точность и стабильность тонкой подстройки задержки должна быть лучше 0,1 фс при длине волны 1,55 мкм, поэтому для устойчивой работы в приемниках DPSK применяют стабилизацию температуры и следящее регулирование для автоматической подстройки интерферометра. Отметим, что такой интерферометр построен по схеме, похожей на ММЦ, но с увеличенной разностью задержек, в соответствии с периодом повторения битов в потоке (100 пс при 10 Гбит/с формата DPSK), и обязательно содержит фазовый модулятор в одном плече, предназначенный для тонкой подстройки статической разности фаз.

Рассматривая особенности передачи и приема данных в формате DPSK, трудно увидеть какие-либо серьезные преимущества этой методики перед примитивными методами передачи сигналов с

простой модуляцией интенсивности. За усложнения аппаратуры и ее обслуживания на обоих концах линии связи наградой служит только улучшение отношения сигнал/шум на 3 дБ. На самом деле потоки формата DPSK обычно служат «деталями для сборки» потока более прогрессивного формата DQPSK, описанного далее.



Рис. П4.13. Балансный приемник сигнала с фазовой манипуляцией (DPSK). Для преобразования ФМ в АМ применен интерферометр DI с линией задержки Т. Выход интерферометра выполнен как симметричный направленный ответвитель

на связанных волноводах; его действие описывает модель связанных волн. При максимуме сигнала на выходе А мощность на выходе В минимальна

Дифференциальная квадратурная фазовая манипуляция

Дифференциальная квадратурная фазовая манипуляция (DQPSK) – единственный реальный формат с кодированием более одного бита в импульсе.

Сигнал формата DQPSK формируют суммированием сигналов более простого формата, DPSK, – дифференциальной фазовой манипуляции. Изменения фазы при этом виде модуляции показаны на плоскости комплексной напряженности поля на рис. П4.11, *в*. Штриховые линии показывают два возможных (предельных) варианта переходов между двумя альтернативными значениями фазы. При переходе по дуге окружности интенсивность излучения во времени не изменяется; такую чистую фазовую манипуляцию получают при помощи простого фазового модулятора. При переходе по прямой через точку [0; 0] интенсивность проходит через ноль. Такую совмещенную фазо-амплитудную модуляцию получают при помощи ММЦ с соответствующим электрическим управлением. Последний вариант применяют чаще, потому что отклонения управляющего напряжения от оптимальной формы в этом случае слабее влияют на качество модулированного сигнала, а уменьшения интенсивности при переходах от одного бита к следующему при фазовой манипуляции не играют существенной роли. Очень эффективны и переключатели фаз на основе эффектов электропоглощения в полупроводниках, особенно на квантово-размерном эффекте Штарка.



Рис. П4.14. Схема передатчика с дифференциальной квадратурной фазовой манипуляцией (DQPSK); минимальный комплект элементов.
Для фазовой манипуляции служат два модулятора Маха–Цендера (ММЦ); выходные сигналы ММЦ суммируют со взаимной задержкой по фазе π/2 (на четверть периода несущей оптической частоты).

Передатчик с дифференциальной фазовой манипуляцией обычно строят на основе интегральной пары ММЦ, управляемых как фазовые модуляторы. Излучение стабилизированного лазера вначале разводят на две ветви, в каждой из которых действует фазовый модулятор, и в одну из ветвей дополнительно введен сдвиг по фазе $\pi/2$.

Для разделения битов потока нулями, если нужен режим RZ, в передатчик дополнительно вводят прерыватель – ММЦ, управляемый как амплитудный модулятор (см. рис. П4.10, а также пример на рис. П4.12).

Использовать на практике все достоинства формата DQPSK можно только при достаточно высокой стабильности несущей частоты передатчика; эти требования, уже указанные в предыдущем

разделе, на порядок жестче, чем требования стандартов магистральной оптоволоконной связи DWDM.

Не имея возможности дать полный обзор всего многообразия форматов модуляции оптических сигналов и подробно прокомментировать их особенности и практические приложения, авторы считали своей главной задачей познакомить студентов-физиков с существом этой темы и хотя бы кратко показать ее обширность и актуальность. Главный вывод, к которому хотелось подвести читателя, таков:

Современная техника оптической связи в своем развитии уже не может довольствоваться только примитивной модуляцией интенсивности излучения, как это было в прошлом..

Между тем, даже в данном обзоре модуляция интенсивности упомянута чаще, чем иные виды модуляции. Причина этого трудно преодолимая инерция. В популярных монографиях и учебниках на эту тему редко встречается последовательное противопоставление модуляции интенсивности и модуляции напряженности поля. Техника оптической связи на ранней стадии развития с сомнениями воспринимала идеи когерентной передачи и обработки сигналов, привычные для «классической» радиотехники. Внедрение когерентных методов в оптическую связь – дело недалекого будущего, если судить о прогрессе в этой области по обилию и качеству исследовательских работ, в том числе и успешных лабораторных испытаний новых методик и аппаратуры. Читателю, заинтересованному в этой тематике, рекомендуем обзор, упомянутый в самом начале данного раздела, где разные методы формирования оптических сигналов описаны гораздо шире, и приведена необъятная библиография оригинальных работ и обзоров по теме, лишь вскользь затронутой здесь.

Международные термины и обозначения видов модуляции, форматов передачи данных и кодирования сигналов

Обозн.	Исходное название	Адекватный перевод (русск.)	
ACRZ	Alternate-chirp	Разделение битов нулями	
	return-to-zero	с чередованием чирпинга	
AMI	Alternate-mark	Двухбитное кодирование	
	inversion	с подавлением несущей (синоним DCS)	
AP	Alternate phase	Чередование фаз	
CAPS	Combined amplitude	Совмещенное кодирование	
	phase shift coding	изменением амплитуды и сдвига фазы	
C-NRZ	Chirped	без разделения битов нулями	
	non-return-to-zero	с чирпингом	
CRZ	Chirped	Разделение битов нулями	
	return-to-zero	с чирпингом	
CSRZ	Carrier-suppressed	Разделение битов нулями с подавлением	
	return-to-zero	несущей; балансная АМ	
DB	Duobinary	Двухбитное кодирование	
DCS	Duobinary carrier	Двухбитное кодирование	
	suppressed	с подавлением несущей (синоним AMI)	
DPSK	Differential phase	Дифференциальная	
	shift keying	фазовая манипуляция	
DQPSK	Differential	Дифференциальная квадратурная фазовая манипуляция	
	quadrature		
	phase shift keying		
FM	Frequency	Частотная модуляция (ЧМ)	
	modulation		
M-ASK	Multilevel amplitude	Многоуровневая	
	shift keying	амплитудная манипуляция	
NRZ	non-return-to-zero	без разделения битов нулями	
OOK	On/off keying	Телеграфная манипуляция	
PASS	Phased amplitude	Коррелированные сдвиги амплитуды	
	shift signaling		
Pol-SK	Polarization shift	Поляризационная манипуляция	
	keying		
PSBT	Phase-shaped	Коррелированная передача	
	binary transmission	двоичного кода	
RZ	return-to-zero	разделение битов нулями	
SSB	Single sideband	Однополосная модуляция	
VSB	Vestigial sideband	Частичное подавление	
		одной боковой полосы	



Рис. П4.15. Оптические спектры модулированных волн при различных типах модуляции. Шкала спектральной плотности логарифмическая, шкала частот линейная. R – частота следования битов в информационном потоке. На врезках в правых верхних углах показаны соответствующие осциллограммы интенсивности потока оптических сигналов (*eye diagram*)



Рис. П4.16 Оптические спектры модулированных волн при различных типах модуляции. Продолжение рис. П4.15

ЛИТЕРАТУРА

Основная литература

- 1. **Информационная оптика**: Учебное пособие / Евтихиев Н. Н. и др.; М.: МЭИ, 2000.
- 2. Васильев А.А., Касасент Д., Компанец И.Н., Парфенов А.В. **Пространственные модуляторы света** – М.: Радио и связь, 1987.
- 3. Магдич Л.Н., Молчанов В. Я. Акустооптические устройства и их применение. М.: Сов. Радио, 1978.
- 4. Гоноровский И.С. Радиотехнические цепи и сигналы. М.: Сов. Радио, 1971 (и более поздние издания).
- 5. Ярив А., Юх П. Оптические волны в кристаллах М.: Мир, 1987.
- 6. Зельдович Б.Я., Пилипецкий Н.Ф., Шкунов В.В. Обращение волнового фронта. – М.: Наука, 1987.
- 7. Блинов Л.М. Электро- и магнитооптика жидких кристаллов. М.: Наука, 1978.

Дополнительная литература

- 8. Winzer P.J., Essiambre R.-J. Advanced optical modulation formats. Proc. IEEE, Vol. 94, No. 5, May 2006, pp. 952–984.
- 9. Saleh B.E.A., Teich M.C. Fundamentals of photonics Wiley, 1991.
- Самарин А. В. Жидкокристаллические дисплеи. Схемотехника, конструкция и применение. Серия: Библиотека инженера – СОЛОН - Р, 2002.
- 12. Сонин А.С. Введение в физику жидких кристаллов. М.: Наука, 1983.

Владимир Константинович Егоров, Иосиф Геннадиевич Зубарев, Ростислав Сергеевич Стариков

Лабораторный практикум

ФИЗИЧЕСКИЕ ОСНОВЫ МОДУЛЯЦИИ ЛАЗЕРНОГО ИЗЛУЧЕНИЯ

Редактор Е.Н. Кочубей

Печать выполнена в соответствии с представленным автором оригиналом-макетом

Подписано в печать 15	Формат 60×84 1/16			
Объем 15,75 п.л.	Уч. изд. л. 15,75.	Тираж 120 экз.		
Изд. № 3/11.		Заказ		
Московский инженерно-физический институт				
(государственный университет). 115409, Москва, Каширское шоссе, 31.				
Типография издательства «Тровант», г. Троицк Московской обл.				