

ОСНОВЫ голографии

Кафедра «Физика»

Лекционный курс

Автор

Илясов В.В.



Аннотация

Лекционный курс предназначен для студентов специальности 210102 Светотехника и источники света.

Автор

**Илясов Виктор Васильевич –
ДОКТОР ТЕХНИЧЕСКИХ НАУК, ДОЦЕНТ.**

ОГЛАВЛЕНИЕ

ОГЛАВЛЕНИЕ	2
Лекция № 1	7
1.1. Физические основы квантовой электроники	7
1.2. Спонтанное и вынужденное излучения	7
1.3. Спектральные линии	10
Лекция № 2	12
2.1. Поглощение и усиление	12
2.2. Инверсия	14
2.3. Процесс оптической накачки	15
Лекция № 3	17
Принципы работы лазера	17
Лекция № 4	22
Пассивные элементы лазера. Оптические резонаторы	22
4.1. Активное вещество	26
Лекция № 5	29
5.1. Селекция мод	29
Лекция № 6	33
Режимы работы лазеров	33
6.1. Стационарный режим генерации	33
6.2. Работа лазеров в режиме модуляции добротности	36
6.3. Модуляция добротности резонаторов	38
Лекция № 7	40
Синхронизация мод в лазерах	40
7.1. Синхронизация мод в лазерах	42
Лекция № 8	45
Типы лазеров	45

Основы голографии

8.1. Твердотельные лазеры	45
8.2. Четырехуровневые лазеры	46
Лекция № 9	50
Полупроводниковые лазеры.....	50
Лекция № 10	54
Лазеры на основе гетероструктур.....	54
10.1. Низкоразмерные полупроводниковые гетероструктуры	55
10.2. Фотодиоды	56
Лекция № 11. Физические основы формирования голограмм.....	59
11.1. Основы формирования голограмм. Типы голограмм	59
11.2. Восстановление волнового фронта. Основное уравнение голографии.....	60
Лекция № 12. Геометрия формирования голограмм	63
12.1. Схема регистрации голограмм Габора	64
12.2. Устранение проблемы второго изображения.....	65
12.3. Схемы записи объемных голограмм. Схема Ю.Н. Денисюка	66
Лекция № 13. Получение голограмм с падающим под углом опорным пучком	68
13.1. Понятие о спектре пространственных частот	68
13.2. Дифракция на периодических структурах	70
13.3. Постановка общей задачи о дифракции.....	73
Лекция № 14. Получение голограмм с падающим под углом опорным пучком (продолжение)	75
14.1. Получение голограмм с наклонным опорным пучком	75
при недиффузном объектном пучке	75
14.2. Голографирование с наклонным опорным пучком при диффузном освещении объекта	80
Лекция № 15. Материалы для записи голограмм.....	82
15.1. Определения плоских и объемных голограмм	82
15.2. Амплитудные и фазовые голограммы. Среды для их регистрации.....	85
15.3. Максимально достижимая дифракционная эффективность тонких голограмм .	90

Основы голографии

15.4. Максимально достижимая дифракционная эффективность объемных голограмм.	96
Лекция № 16. Интеграл суперпозиции	100
16.1. Линейность оптических систем и интеграл суперпозиции.	100
16.2. Геометрический анализ голограмм точечного источника.	103
Лекция 17. Изображающие свойства голограмм точечного источника	112
17.1. Изображающие свойства голограмм в первом паракиальном приближении	112
17.2. Влияние конечных размеров голограммы. Структура изображения точечного объекта, формируемого голограммой в отсутствии aberrаций	119
17.3. Монохроматические aberrации третьего порядка	124
Лекция 18. Факторы, ограничивающие качество изображения, формируемого методом голографии	130
18.1. Интермодуляционные помехи	130
18.2. Шумы регистрирующей среды	131
18.3. Помехи, обусловленные нелинейностью характеристической кривой регистрирующей среды	133
18.4. Влияние неравномерности толщины регистрирующей среды на качество восстановленного изображения	138
Лекция 19. Системы голографической памяти	140
19.1. Системы голографической памяти	140
Лекция 20. Голографическая интерференция с цифровым восстановлением и обработкой данных	146
Лекция 21. Голографические мультиплекторы для систем передачи данных со спектральным уплотнением	152
Лекция 22. Голографические метки идентификации	154
Лекция 23. Солнечные энергетические установки с голограммными оптическими элементами в системах концентрации энергии	158
Лекция 24. Голографические проекционные системы	160
Лекция 25. Голографические системы кодирования информации	162
Лекция 26. Голографические системы кодирования информации (продолжение)	165
26.1. Изготовленные методом голографии фотонные полимеры и их применение .	165

Дополнительная литература169

Лекция № 1

1.1. Физические основы квантовой электроники

Квантовая электроника – область электроники, в которой исследуются явления генерации и усиления электромагнитных колебаний на основе эффектов вынужденного излучения и нелинейного взаимодействия излучения с веществом, как основы при создании квантовых генераторов (лазеров) и др. устройств.

Для описания электромагнитного излучения оптического диапазона используют традиционно электродинамику (на базе уравнений Максвелла) или квантово-механический подход (с использованием уравнения Шредингера).

Электромагнитное излучение (электромагнитные волны) — распространяющееся в пространстве возмущение (изменение состояния) электромагнитного поля (то есть, взаимодействующих друг с другом электрического и магнитного полей). Основными характеристиками электромагнитного излучения принято считать частоту, длину волны и поляризацию. Видимое, инфракрасное и ультрафиолетовое излучения составляет так называемую оптическую область спектра в широком смысле этого слова.

1.2. Спонтанное и вынужденное излучения

В квантовой электронике используются физические явления, в которых основное участие принимают *связанные электроны*, входящие в состав атомов, молекул вещества (во всех агрегатных состояниях). Энергия электрона (в соответствии с законами квантовой механики) имеет ряд дискретных значений (энергетические уровни). Весь набор разрешённых квантовой механикой уровней образует энергетический спектр атома. Основным называется уровень E_0 с наименьшей энергией, все остальные уровни – возбуждёнными. Переход связанных электронов с одного уровня на другой сопровождается излучением или поглощением электромагнитной энергии

$$h\nu_{i,j} = |E_i - E_j|, \quad (1.1)$$

где h – постоянная Планка, $\nu_{i,j}$ – частота излучения (поглощения) при квантовом переходе с уровня E_i на уровень E_j .

Различают *спонтанные* и *вынужденные переходы*. При *спонтанном квантовом переходе* испускание фотона происходит вне зависимости от воздействий на атомную систему. *Вынужденный квантовый переход* происходит под воздействием внешнего излучения частоты ν , удовлетворяющей условию (1.1). При вынужденном излучении фотон имеет тоже направление излучения и поляризацию, что и фотон стимулирующего излучения. Частота испущенного фотона в точности совпадает с частотой вынужденного излучения (рис.1.1).

Атом находится в возбуждённом состоянии некоторое время τ , после чего переходит в основное состояние. Пусть время жизни на возбуждённом уровне составляет τ_i , то при большом числе возбуждённых частиц N_i их количество убывает по закону

Основы голографии

$$N_i = N_0 \exp\left(-\frac{t}{\tau_i}\right), \quad (1.2)$$

где N_0 - общее число возбужденных частиц в начальный момент времени, t - текущее время, τ_i - время, за которое число возбужденных частиц N_i уменьшается в e раз.

Частицы, потерявшие энергию и испустившие фотон, могут вновь возбуждаться и переходить на уровень E_i . Тем меньше τ_i , тем чаще будут испускаться фотоны.

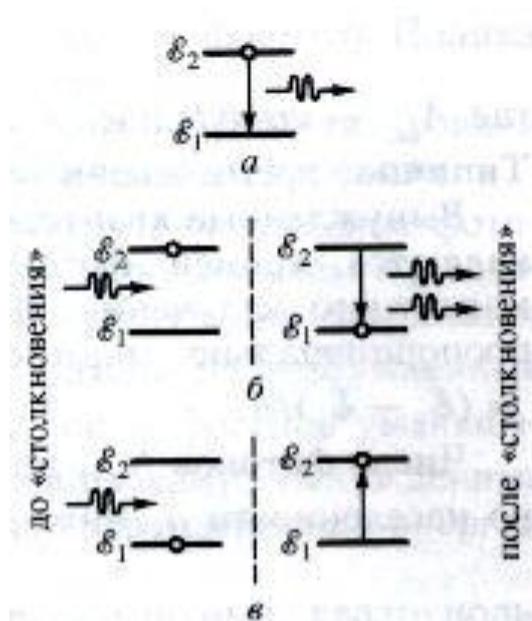


Рис. 1.1. Квантовые переходы: *a* — спонтанный переход между уровнями ϵ_1 и ϵ_2 с генерацией фотона; *b* — вынужденный переход вниз с увеличением числа фотонов; *в* — вынужденный переход вверх с поглощением фотона

Величина $A = 1/\tau_i$ называется *вероятностью спонтанного испускания* с уровня E_i и определяет среднее число фотонов, испускаемых одной частицей за 1 с.

При спонтанном переходе с уровня E_2 на уровень E_1 , величина A_{21} называется *вероятностью перехода*. Полная вероятность A_i спонтанного испускания с уровня E_i на любой другой уровень равна сумме вероятностей отдельных спонтанных переходов

Основы голографии

$$A_i = \sum A_{ik}, \quad (1.3)$$

где A_{ik} - коэффициент Эйнштейна для спонтанного испускания. Типичное время жизни возбуждённых атомов составляет 10^{-8} с.

Вынужденные квантовые переходы генерируют фотоны, которые являются копией фотонов, стимулирующих усиление электромагнитного излучения. Число вынужденных квантовых переходов пропорционально плотности излучения ρ_ν на частоте $\nu = (E_i - E_j)/h$.

Число фотонов N_{ki} , поглощённых в 1 см^3 за 1 с, пропорционально населённости n_k нижнего уровня E_k и плотности излучения ρ_ν :

$$N_{ki} = B_{ki} n_k \rho_\nu, \quad (1.4)$$

где B_{ki} - коэффициент Эйнштейна для поглощения в квантовой системе. Этот коэффициент характеризует вероятность поглощения и равен числу фотонов, поглощаемых одной частицей за 1 с, при приведённой плотности излучения ρ_ν . Вынужденное излучение характеризуется числом фотонов N_{ik}^* , испущенных в 1 см^3 за 1 с,

$$N_{ik}^* = B_{ik} n_k \rho_\nu, \quad (1.5)$$

где B_{ik} - коэффициент Эйнштейна для вынужденного испускания, который характеризует вероятность вынужденного испускания. Этот коэффициент определяется числом фотонов, испускаемых в среднем одной частицей под действием излучения плотности $\rho_\nu = 1$ за 1 с.

Эйнштейн А. рассмотрел равновесную систему, в которой число фотонов, испускаемых в переходе $E_i \rightarrow E_k$ с частотой $\nu = |E_i - E_k|/h$, равно числу фотонов той же частоты, поглощаемых при обратном квантовом переходе $E_k \rightarrow E_i$. Тогда

$$\begin{aligned} N_{ik} + N_{ik}^* &= N_{ki}, \\ A_{ik} n_i + B_{ik} n_i \rho_\nu &= B_{ki} n_k \rho_\nu. \end{aligned}$$

Равновесная плотность излучения ρ_ν связана с коэффициентами Эйнштейна следующими соотношениями:

$$\begin{aligned} g_k B_{ki} &= g_i B_{ik}, \\ A_{ik} &= \frac{8\pi h \nu^3}{c^3} \frac{g_k}{g_i}, \\ B_{ki} &= \frac{8\pi h \nu^3}{c^3} B_{ik}, \end{aligned} \quad (1.6)$$

где g_i и g_k - степени вырождения уровней E_i и E_k соответственно, c - скорость света, коэффициент $8\pi h \nu^3 / c^3$ входит в формулу Планка для плотности энергии равновесного излучения.

Вероятность излучаемых переходов зависит от свойств уровней E_i и E_k , между которыми осуществляются квантовые переходы.

Основы голографии

В квантовой системе имеет место резонансное поглощение фотонов. Суть этого явления заключается в том, что если атом находится на нижнем уровне E_k и электромагнитное излучение содержит n фотонов частоты ν_{ki} , то возможен переход атома на возбуждённый уровень E_i . При этом поглощается фотон, число фотонов уменьшается и становится равным $(n-1)$, т.е. вынужденные переходы вверх приводят к поглощению электромагнитного излучения веществом.

Различают *излучательные* (дипольные, магнитные, квадрупольные) и *безызлучательные* квантовые переходы. Излучательные переходы сопровождаются изменением дипольного момента P_{ik} , магнитного момента M_{ik} , квадрупольного момента G_{ik} , которые связаны с коэффициентами Эйнштейна. При безызлучательных переходах изменение энергии квантовой системы связано с её взаимодействием с другими квантовыми системами.

Среди квантовых переходов различают также *разрешённые* и *запрещённые* переходы. Если в какой-либо момент вероятность перехода отлична от нуля, то квантовый переход возможен. Если же вероятность перехода равна нулю, то квантовый переход запрещён. В случае, когда квантовый переход с некоторого возбуждённого уровня на более низкие уровни запрещён, то возбуждённые состояния на таком уровне могут существовать довольно длительное время. Такой уровень называется метастабильным. Например, время жизни возбуждённых атомов хрома в рубиновом лазере на метастабильном уровне составляет $\approx 10^{-3}$ с. Наличие долгоживущего метастабильного уровня позволяет создать на нём высокую населённость возбуждённых состояний.

1.3. Спектральные линии

В идеальном случае возбуждённые атомы должны излучать строго монохроматическое излучение одной частоты. На практике излучение образует спектральную линию определённой ширины и формы. Если Δt - время жизни атома в возбуждённом состоянии, а ΔE - значение его энергии состояния, то они связаны соотношением неопределённости Гейзенберга

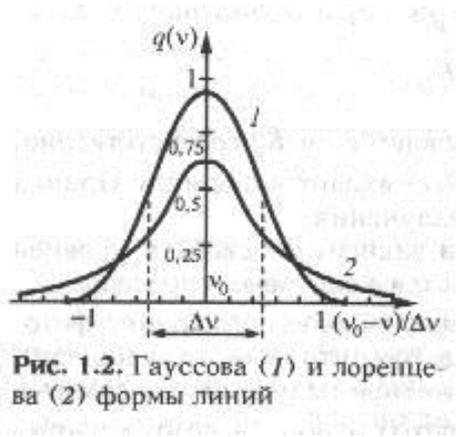
$$\Delta E \cdot \Delta t \approx \hbar.$$

Неопределённость или «размытие» уровня обратно пропорционально времени жизни τ_0 частицы в основном состоянии. Поскольку время жизни свободной частицы на энергетическом уровне всегда конечно, то существует определённая естественная ширина спектральной линии.

Спектральное распределение квантов спонтанного излучения определяет ширину уровня $\Delta E = \hbar/\tau_0$.

Контур линии спонтанного излучения имеет *лоренцеву* форму (рис. 1.2) с шириной линии $\Delta \nu_{\lambda} = \Delta E/\hbar = 1/(2\pi\tau_0)$.

Основы голографии



Шириной линии называют интервал частот между точками, для которых интенсивность излучения (или поглощения) падает в два раза.

Лоренцева форма линии имеет вид резонансной кривой с максимумом на частоте ν_0 и описывается так называемым форм-фактором

$$q(\nu) = \frac{1}{2\pi} \frac{\Delta \nu_L}{(\nu - \nu_0) + \Delta \nu_L^2/4}$$

В реальных условиях спектральные линии более размыты и представляют собой полосы излучения и поглощения. Причиной этого являются различные физические явления. Например, уширение линии происходит вследствие эффекта Доплера, вызывающего смещение частоты движущихся частиц. Доплеровски уширенная линия описывается функцией Гаусса и симметрична относительно частоты ν_0 (рис. 1.2). Форм-фактор этой линии имеет вид

$$q(\nu) = \frac{1}{\nu_\tau \sqrt{\pi}} \exp \left[- \left(\frac{\nu - \nu_0}{\nu_\tau} \right)^2 \right],$$

где $\nu_\tau = \nu_0(u_0/c)$ - доплеровский сдвиг частоты при средней тепловой скорости движения излучающей частицы, u_0 - средняя тепловая скорость, c - скорость света. С увеличением частоты роль доплеровского уширения линии возрастает.

В твёрдых телах уширение спектральных линий и даже их расщепление возможно вследствие влияния электрических и магнитных полей (эффект Штарка, эффект Зеемана).

Лекция № 2

2.1. Поглощение и усиление

В естественных условиях при равновесии между средой и веществом нижние уровни энергии заселены более плотно, чем верхние. Существует фундаментальный закон распределения частиц по энергиям (распределение Больцмана)

$$N_n = C \exp\left(-\frac{E_n}{kT}\right), \quad (2.1)$$

где C - константа, зависящая от полного числа частиц в единице объёма, k - постоянная Больцмана.

Основная проблема, возникающая при создании квантовых усилителей и генераторов, состоит в поиске способов нарушения теплового равновесия рабочего вещества так, чтобы населённость верхних уровней была существенно выше населённости нижних уровней.

Система квантовых частиц, в которой хотя бы для двух уровней энергии более высоко расположенный уровень населён значительно больше нижнего, называют *системой с инверсионной населённостью*.

В соответствии с уравнением (1.7) отношение населённостей верхнего уровня N_2 к нижнему N_1 определяется соотношением

$$\frac{N_2}{N_1} = \exp\left(-\frac{E_2 - E_1}{kT}\right) = \exp\left(-\frac{h\nu_{21}}{kT}\right), \quad (2.2)$$

где $\nu_{21} > 0$ - частота перехода.

Температуру вещества можно определить как

$$T = \frac{h\nu_{21}}{k \ln(N_1/N_2)}. \quad (2.3)$$

При $N_2 < N_1$ - обычный тепловой режим и $T > 0$. При инверсионной населённости, когда $N_2 > N_1$, температура перехода становится отрицательной величиной ($T < 0$). Отметим, что здесь отрицательная температура является условной ртематической величиной, характеризующей физический процесс инверсии населённостей для данной пары уровней энергии в квантовой системе.

Поглощаемое квантовой системой частиц излучение нарушает в ней тепловое равновесие. В случае, когда вероятность переходов по влиянием поля накачки становится сравнимой с вероятностью релаксационных переходов, равновесное распределение населённостей изменяется. Доля энергии, поглощаемой квантовой системой, уменьшается и возникает эффект насыщения. В этой ситуации коэффициент поглощения α падает:

$$\alpha = -\frac{dI}{I dx} = -\frac{d\rho}{c\rho dt}.$$

Здесь I - плотность мощности излучения, x - направление распространения волны, ρ - плотность излучения.

Основы голографии

Динамика изменения числа частиц на верхнем уровне определяется уравнением:

$$\frac{dn_2}{dt} = -\left(\omega_{12} + \frac{1}{\tau_0}\right)n_2 + \omega_{12}n_1 - W_{21}n_2 + W_{12}n_1, \quad (2.4)$$

где первый член соответствует спонтанному переходу и вероятности релаксации ω_{12} частиц с верхнего уровня на нижний; второй член соответствует релаксационному заселению второго уровня с вероятностью ω_{12} ; третий и четвёртый члены описывают индуцированные переходы соответственно на уровни 1 и 2. Известно, что

$$W_{21} = B_{21}\rho \text{ и } W_{12} = B_{12}\rho,$$

где $\rho = \frac{8\pi\nu^3}{c^3} \left[\frac{h\nu}{\exp(h\nu/kT - 1)} \right]$ - плотность энергии поля излучения. Известно

также, что $g_1B_{12} = g_2B_{21}$.

Тогда получаем

$$\frac{dn_2}{dt} = -\left(\frac{1}{\tau} + \frac{g_1 + g_2}{g_2} \frac{2B_{12}\rho}{\pi\Delta\nu_l}\right)n_2 + \left(\omega_{12} + \frac{2B_{12}\rho}{\pi\Delta\nu_l}\right)n_1, \quad (2.4,a)$$

где g_1, g_2 - кратность вырождения соответствующих уровней; $n = n_1 + n_2$, а

$$\tau = \frac{1}{\omega_{21} + \omega_{12} + 1/\tau_0}.$$

В стационарных условиях $dn_2/dt = 0$, и при $\rho \rightarrow \infty$ имеем

$$n_2 = \frac{g_2 n}{g_1 + g_2}; \quad n_1 = \frac{g_1 n}{g_1 + g_2}.$$

В результате населённости верхнего и нижнего уровней выравниваются, и происходит полное насыщение.

Величину резонансного усиления (поглощения) определим как

$$\alpha = \left(\frac{n_1}{g_1} - \frac{n_2}{g_2}\right) \left(\frac{2g_1B_{12}h\nu}{c\pi\Delta\nu_l}\right),$$

или разность населённостей

$$z = z_0 \left[1 + \frac{I(g_1 + g_2)}{2g_2I_s} \right]^{-1}, \quad (2.5)$$

где I_s - уровень насыщения, z_0 - разность населённостей в отсутствие внешнего поля.

Условие, при котором активная среда позволяет усиливать проходящее через нее электромагнитное излучение, можно записать в виде

$$I = I_0 \exp(\alpha - \beta_n),$$

где β_n - коэффициент потерь, служащий количественной характеристикой потерь в активной среде.

При помещении активной среды в резонатор (с высокой добротностью) усиление может превысить суммарные потери усилителя в цепи обратной связи. Произойдёт самовозбуждение усилителя, и он превратится в генератор. Условием

Основы голографии

самовозбуждения лазера с резонатором типа Фабри_Перо, состоящего из двух зеркал, расположенных на расстоянии L , с коэффициентами отражения R_1 и R_2 , будет неравенство

$$\alpha \geq \beta_n + 2L \left(\ln \frac{1}{R_1 R_2} \right)^{-1}. \tag{2.6}$$

Другими словами, коэффициент усиления равен сумме коэффициентов потерь активной среды и коэффициентов потерь на зеркалах.

2.2. Инверсия

В квантовой электронике обычно имеют дело с большим числом атомов, представляющим целый ансамбль частиц. Достаточно вспомнить, что одном кубическом сантиметре газа при нормальных условиях (температуре $t = 0 \text{ }^\circ\text{C}$ и давлении, равном 133.3 Па) находится $2.7 \cdot 10^{19} \text{ см}^{-3}$ атомов (число Лошмидта), а в твердом теле - около 10^{22} см^{-3} .

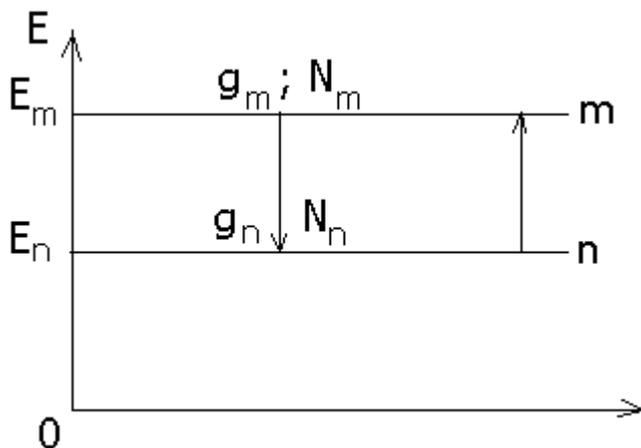


Рис. 2.1. Диаграмма энергетических уровней ансамбля молекул

Частицы, входящие в состав ансамбля, непрерывно взаимодействуют между собой, что приводит к вполне определенному распределению их по энергиям. При термодинамическом равновесии распределение частиц по энергиям подчиняется закону Больцмана, согласно которому вероятность пребывания атома в состоянии i с энергией E_i равна

$$P_i = C e^{-\frac{E_i}{kT}}, \tag{2.7}$$

Где C – константа; T – абсолютная температура; k – постоянная Больцмана, $k = 1.38 \cdot 10^{-23} \text{ Дж / } ^\circ\text{C}$.

Основы голографии

Таким образом, из общего числа частиц ансамбля N_0 на уровне i находятся

$$N_i = N_0 C e^{-\frac{E_i}{kT}} \quad (2.8)$$

Отношение числа частиц, пребывающих в состоянии m , с энергией E_m к числу частиц с энергией E_n будет

$$N_m/N_n = e^{-(E_m - E_n)/kT} \quad (2.9)$$

В случае, если одному энергетическому уровню атома соответствует несколько различных его устойчивых состояний, то данный энергетический уровень вырожден. Число g различных независимых состояний с одинаковым значением энергии определяет кратность вырождения или статистический вес уровня.

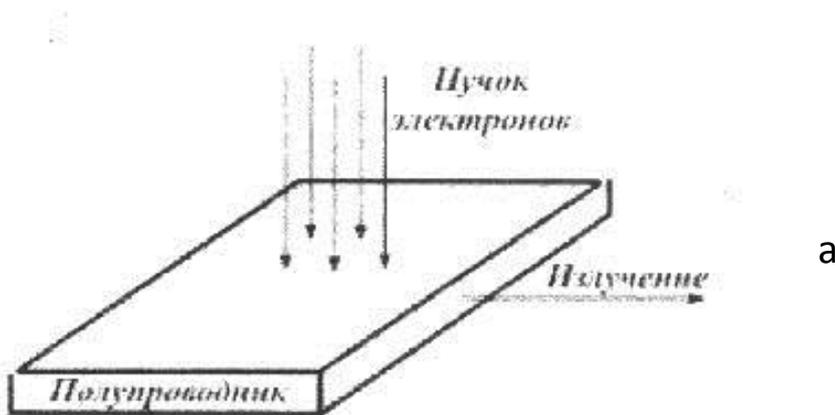
Отношение населенностей двух вырожденных уровней m и n при их статистических весах g_m и g_n в отличие от (1.7) определяется соотношением

$$N_m/N_n = (g_m/g_n) e^{-(E_m - E_n)/kT} \quad (2.10)$$

Из (2.9) и (2.10) следует, что при термодинамическом равновесии число частиц, находящихся на верхнем энергетическом уровне, оказывается меньше, чем на нижнем. Причем. Чем выше располагается энергетический уровень, тем меньше находится на нем частиц. На рис.2.1 представлена диаграмма энергетических уровней. Длина горизонтальной черты (на рисунке) определяет населенность того или иного уровня.

2.3. Процесс оптической накачки

Различают поперечную и продольную накачку полупроводниковых лазеров.



Основы голографии



Рис. 2. . Продольная (а) и поперечная (б) накачки электронным пучком

Накачка электронным пучком вызывает сильный нагрев полупроводника, поэтому требуется хорошее охлаждение. Например, для GaAs (арсенида галлия) принято охлаждать до 80 К.

Полупроводниковые слои, способные к генерации лазерного излучения, реализуются в разных формах:

- как гомоструктура;
- как гетероструктура;
- в виде квантовых барьеров (квантовых ям) и квантовых точек.

Лекция № 3

Принципы работы лазера

Лазер представляет собой источник когерентного электромагнитного излучения оптического излучения, действие которого основано на использовании вынужденного излучения атомов или молекул.

Слово лазер происходит от английского Laser (Light Amplification by Stimulated Emission of Radiation – усиление света вынужденным излучением). Если в традиционных источниках света используется спонтанное излучение системы возбуждённых атомов, то в лазерах все атомы когерентно излучают кванты света, тождественно равные между собой по частоте, направлению распространения и поляризации.

В основе работы лазеров лежат явления вынужденного излучения под действием внешнего электромагнитного поля, усиления и формирования потока излучения.

Возбуждение квантовой системы осуществляется путём накачки – импульсного или постоянного воздействия на активную среду электромагнитным излучением определённой частоты. Возбуждение активной среды может осуществляться по трёх- или четырёхуровневой схеме (рис. 3.1).

Примером трёхуровневой схемы является лазер на рубине ($Al_2O_3 + Cr^{3+}$). Содержание атомов хрома составляет около 0,05%. Красный цвет кристаллов обусловлен расположением энергетических уровней атомов хрома в рубине. Возбуждение атомов хрома в рубине осуществляется за счёт оптической накачки. Возбуждённые атомы хрома переходят в полосу поглощения E_3 и затем через время 10^{-8} с спонтанно переходят на метастабильный уровень E_2 ; их время жизни там составляет 10^{-3} с. В атомных масштабах соотношение времён 10^{-8} с и 10^{-3} с является вечностью, и поэтому на метастабильном уровне накапливается большое число (больше половины всех) возбуждённых атомов. Метастабильный уровень становится населённым, и в квантовой системе возникает инверсия населённостей уровней по отношению к основному уровню E_1 . Накачка рубина осуществляется лампой, обеспечивающей вспышку белого света. Энергетические полосы хрома в рубине позволяют использовать до 15% света лампы-вспышки. При содержании атомов хрома в рубине 0,05% концентрация хрома составляет 10^{19} см⁻³. Обычно возбуждается около 50 % атомов хрома. Наличие фотона с частотой $\nu = (E_2 - E_1)/h$ позволяет стимулировать лазерную генерацию, заключающуюся в вынужденном излучении при переходы ансамбля возбуждённых атомов хрома с уровня E_2 на уровень E_1 .

Лазерный переход $E_2 \rightarrow E_1$ позволяет генерировать излучение с частотой

$$\nu = \frac{E_2 - E_1}{h}. \quad (3.1)$$

Основы голографии

Лазер на стекле с примесью Nd^{3+} или на основе кристалла алюмоиттриевого граната с ионами Nd^{3+} является примером четырёхуровневой системы (рис. 3.1,б). В энергетическом спектре такого типа лазеров между метастабильным уровнем E_3 и основным уровнем E_1 имеется уровень E_2 с небольшой заселённостью возбуждёнными атомами. Лазерный переход между уровнями E_3 и E_2 позволяет генерировать излучение с частотой $\nu = (E_3 - E_2)/h$. По четырёхуровневой схеме работает также ряд лазеров на газовых средах.



Рис. 2.1. Процессы инверсии населенности и генерации в трех- (а) и четырехуровневых (б) системах

Положительная обратная связь

осуществляется с помощью двух зеркал, образующих интерферометр Фабри-Перо, который является оптическим открытым резонатором.

В резонаторе могут возбуждаться колебания только определённых длин волн и определённой структуры. Частоты этих колебаний называются резонансными и соответствуют собственным частотам резонатора. Возникающие колебания имеют определённую структуру (или моду), которая соответствует собственному типу колебаний

резонатора (от латинского *modus* –мера, способ). Мода может быть интерпретирована как стационарная конфигурация электромагнитного поля, электрическая составляющая которой может быть записана в виде

$$\vec{E}(r,t) = \vec{E}_0 U(r) \exp(i\omega t), \tag{3.2}$$

где ω - собственная частота резонатора.

Резонатор характеризуется добротностью Q , которая представляет собой отношение запасённой в резонаторе энергии к энергии, теряемой в резонаторе за период колебаний. Добротность можно выразить через эффективное время жизни фотона τ_{eff} в моде

$$Q = \omega \cdot \tau_{eff}.$$

Электромагнитные волны, распространяющиеся вдоль оптической оси резонатора, отражаются и интерферируют между собой. Коэффициент отражения зеркал достаточно высок, и поэтому в пространстве резонатора могут существовать электромагнитные колебания строго определённой длины волны. На поверхности идеально отражающих зеркал амплитуда световых колебаний

Основы голографии

обратной ширине спектра генерируемых мод, а в режиме синхронизации мод может достигать значений $10^{-12} \dots 10^{-13}$ с, что позволяет резко повысить пиковую мощность излучения.

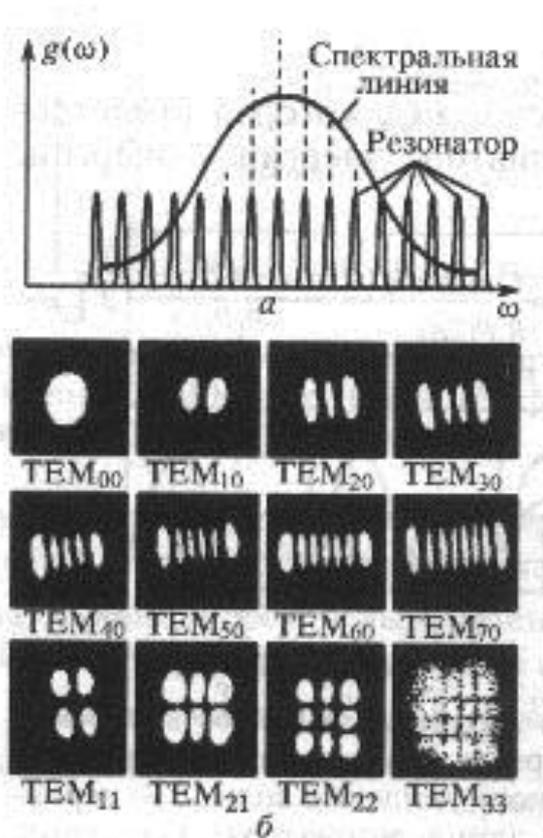


Рис. 2.3. Продольные собственные частоты на фоне спектральной линии (а), трансверсальные моды лазера (б)

Лазерные зеркала обеспечивают возможность многократного прохода волны, если волновой вектор направлен по оси резонатора. Многократное прохождение в резонаторе световой волны обеспечивает её усиление путём многократного «опустошения» метастабильного уровня и генерации фотонов. Инвертированная активная среда при каждом проходе будет усиливать волну и повышать плотность фотонов.

Если $P(0)$ - мощность аксиальной моды первоначального спонтанного излучения, а $P(2L)$ - мощность после двойного прохода резонатора, то при коэффициентах отражения зеркал R_1 и R_2 и коэффициенте усиления среды β и коэффициенте поглощения α можно записать

$$P(2L) = P(0) \exp[2(\beta - \alpha)R_1R_2].$$

Самовозбуждение наступает при условии $P(2L) > P(0)$, или

$$\beta > \alpha + \frac{1}{2L} \ln \frac{1}{R_1R_2}.$$

Обычно одно зеркало делают глухим, т.е. $R_1 \approx 1$, а второе $R_2 \approx 0,95$. Накопленное в резонаторе излучение, преодолев определённый порог интенсивности, выходит из зеркала с коэффициентом отражения R_2 .

Сформированный таким образом пучок лазерного излучения является когерентным излучением с высокой спектральной плотностью излучения.

Гармоническое колебание называется **монохроматическим**, если оно может быть описано выражением

$$A(t) = A_0 \sin(2\pi\nu_0 t + \varphi_0),$$

где $A(t)$, A_0 - текущее и максимальное значения амплитуды соответственно, $2\pi\nu_0$ - круговая частота, φ_0 - начальная фаза.

Ширина спектра $\Delta\nu$ излучения определяется степенью монохроматичности излучения

Основы голографии

$$\mu = \frac{\Delta \nu}{\nu_0},$$

где ν_0 - центральная частота.

При $\mu \ll 1$ излучение называют *квазимонохроматическим*. Лазеры позволяют получать излучение со значением $\mu \approx 10^{-10}$ при достаточно большой мощности.

Понятие монохроматичности тесно связано с понятием когерентности. *Когерентность* света представляет собой взаимную согласованность протекания во времени световых колебаний в разных точках пространства и (или) времени, которая характеризует их способность к интерференции.

Различают *пространственную* и *временную* когерентность. Пространственная когерентность связывается с корреляцией фазы колебаний в разных точках пространства в один и тот же момент времени. При сложении когерентных колебаний возникает устойчивая интерференционная картина.

Корреляцию колебаний в определённой точке пространства можно наблюдать только в определённом интервале времени, который принято называть *временем когерентности*. Время когерентности обычно принимается за время жизни излучённого колебания τ . Расстояние, проходимое светом за время когерентности, называют *длиной когерентности* L . При $\tau \approx 10^{-8} \text{ с}$ длина когерентности $L = c\tau = 300 \text{ см}$. Известно, что ширина спектральной линии связана с временем жизни как $\Delta \nu \approx 1/\tau$. В этом случае длина когерентности связана с шириной спектральной линии как $L \approx c/\Delta \nu$. Таким образом, чем уже частотный спектр излучения, тем больше время когерентности, выше степень временной когерентности и лучше монохроматичность излучения.

Лекция № 4

Пассивные элементы лазера. Оптические резонаторы

Одним из основных элементов оптического лазера является резонатор. Термин «резонанс» предполагает нарастание амплитуды электромагнитных волн вследствие многократного отражения их некоторыми отражающими поверхностями.

Колебательная система, представляющая собой некоторый объем диэлектрика, ограниченного замкнутой окружающей поверхностью, носит название объемного резонатора. Объемный резонатор имеет бесчисленное множество собственных резонансных частот. Задача определения параметров объемных резонаторов и их собственных частот сводится к решению уравнений Максвелла при соответствующих граничных условиях на стенках. Задача значительно упрощается, если использовать результаты, полученные для других электродинамических систем.

Обычно используемые в электродинамических системах замкнутые резонаторы имеют размеры, соизмеримые с длиной рабочей волны. При этом спектр собственных частот в рабочем диапазоне сильно разрежен. В лазерной технике замкнутые резонаторы используют лишь в крайней длинноволновой части оптического диапазона. В ИК области и в видимой части спектра ввиду весьма малых λ используют резонаторы с размерами значительно превосходящими длину волны. При этом собственные типы колебаний имеют высокий порядок. В замкнутых объемных резонаторах спектр собственных частот резонатора становится весьма плотным, т.е. разница между двумя соседними частотами в рабочем диапазоне становится малой. Число колебаний ΔN , приходящихся на интервал частот $\Delta\omega$, равно

$$\Delta N = \frac{V}{2\pi^2 c^3} \omega^2 \Delta\omega, \quad (2.56)$$

где V - объем резонатора; c – скорость света.

Из (2.56) следует, что для высоких частот (когда длина волны мала по сравнению с размерами резонатора) спектр не зависит от формы резонатора и определяется его объемом. Применение закрытых резонаторов в оптическом диапазоне не представляется возможным. Использование резонаторов в оптическом диапазоне становится возможным, если сделать спектры их собственных частот более редкими и уменьшить потери в резонаторе, т.е. увеличить его добротность. Это достигается удалением боковых стенок резонатора. При этом в резонаторе остаются лишь те типы колебаний, которые образуются путем отражения волн от торцевых стенок. Все другие типы волн исчезают, в результате спектр собственных частот становится разреженным по сравнению со спектром аналогичного замкнутого резонатора. Простейшим видом резонатора является резонатор, состоящий из двух плоских зеркал. Резонатор в лазере может быть образован двумя сферическими зеркалами, одним плоским и другим сферическим зеркалами. Если система зеркал обеспечивает циркуляцию луча по замкнутому контуру, то такой резонатор называют кольцевым.

Основы голографии

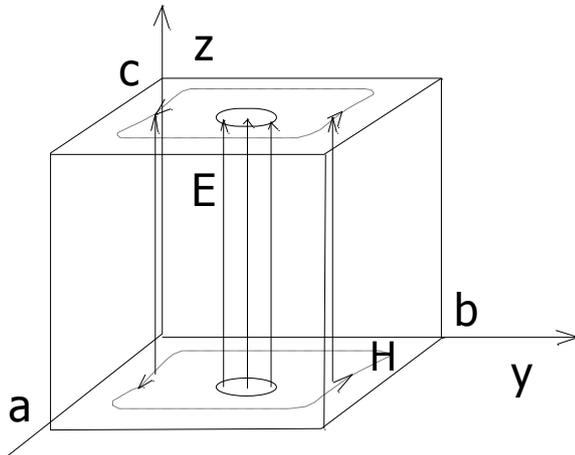


Рис. 2.11. Прямоугольный объемный закрытый резонатор

Рассмотрим основные свойства прямоугольного объемного закрытого резонатора (Рис. 2.11). Определим собственные колебания резонатора, полагая, что его стенки обладают идеальной проводимостью и что внутри резонатора среда с $\epsilon=1$ и $\mu=1$. Волновое уравнение для вектора \vec{E} имеет вид

$$\Delta E + k^2 \vec{E} = 0.$$

Решая это уравнение методом разделения переменных и принимая во внимание граничные условия $E_t = 0$, а также, что внутри резонатора нет свободных зарядов ($div \vec{E} = 0$), найдем

$$\begin{aligned} E_x &= A \cos \frac{m\pi}{a} x \sin \frac{n\pi}{b} y \sin \frac{p\pi}{c} z ; \\ E_y &= B \sin \frac{m\pi}{a} x \cos \frac{n\pi}{b} y \sin \frac{p\pi}{c} z ; \\ E_z &= C \sin \frac{m\pi}{a} x \sin \frac{n\pi}{b} y \cos \frac{p\pi}{c} z , \end{aligned} \quad (2.57)$$

где m, n и p - целые числа, каждое из которых может принимать значения $0, 1, 2, \dots$; A, B и C - комплексные постоянные числа. Величины $m\pi/a$, $n\pi/b$ и $p\pi/c$ представляют собой составляющие волнового вектора \vec{k} по осям x, y, z , т.е. $k \cos \alpha = m\pi/a = k_x$; $k \cos \beta = n\pi/b = k_y$; $k \cos \gamma = p\pi/c = k_z$, где α, β, γ - углы между направлением вектора \vec{k} и соответствующими осями координат. Волновое число k определяется из уравнения

$$k^2 = k_x^2 + k_y^2 + k_z^2 = \left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2 + \left(\frac{p\pi}{c}\right)^2. \quad (2.58)$$

Составляющие магнитного поля \vec{H} можно найти из уравнения

$$\vec{H} = \frac{1}{jk} \text{rot} \vec{E}.$$

Отсюда имеем:

$$\begin{aligned} H_x &= \frac{\pi}{jk} \left(\frac{n}{b} C - \frac{p}{c} B \right) \sin \frac{m\pi x}{a} \cos \frac{n\pi y}{b} \cos \frac{p\pi z}{c} ; \\ H_y &= \frac{\pi}{jk} \left(\frac{p}{c} A - \frac{m}{a} C \right) \cos \frac{m\pi x}{a} \sin \frac{n\pi y}{b} \cos \frac{p\pi z}{c} ; \\ H_z &= \frac{\pi}{jk} \left(\frac{a}{m} B - \frac{n}{b} A \right) \cos \frac{m\pi x}{a} \cos \frac{n\pi y}{b} \sin \frac{p\pi z}{c} . \end{aligned} \quad (2.59)$$

Основы голографии

Таким образом, каждой тройке числе m, n, p соответствует некоторое поле в резонаторе, определенное уравнениями (2.57) и (2.59) с волновым числом k , которое может быть найдено из (2.58). Поле имеет длину волны

$$\lambda = \frac{2}{\sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2 + \left(\frac{p}{c}\right)^2}}$$

и частоту

$$\nu = \frac{c}{2} \sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2 + \left(\frac{p}{c}\right)^2}.$$

Каждой тройке чисел соответствует волна с определенной частотой, имеющая свое распределение поля и свое определенное направление распространения.

Простейшее колебание в полости будет при $m=1, n=1, p=0$ (если два индекса будут равны нулю, то поле вообще исчезает). Оно имеет частоту

$$\nu = \frac{c}{2} \sqrt{\frac{1}{a^2} + \frac{1}{b^2}}.$$

Это так называемая основная собственная частота колебаний резонатора. При этом электрическое поле имеет только одну составляющую по оси z , т.е. $E_x = E_y = 0; E_z \neq 0$; магнитное поле по оси z не имеет составляющей, а отличными от нуля будут составляющие по осям x и y .

Заметим, что множитель $\pm j$, стоящий в формулах для магнитного поля, показывает, что магнитное поле сдвинуто по фазе на $\pi/2$ относительно электрического. При стоячих волнах в резонаторе бывает момент времени, когда вся энергия сосредоточена в магнитном поле. Электрическое поле в этот момент времени отсутствует, и бывают моменты, когда вся энергия сосредоточена в электрическом поле.

На примере рассмотренного резонатора видно, что резонаторы имеют дискретные резонансные частоты колебаний. Это означает, что поля правильной формы в резонаторе возникают лишь в случае, если частота возбуждения будет в точности равна резонансной частоте. Этот результат получен в предположении, что нет потерь энергии в стенках резонатора и в диэлектрике, заполняющей резонатор. В действительности проводимость стенок имеет вполне определенное значение, не исключается возможность также потерь в диэлектрике. В резонаторе частота колебаний ω_g отлична от частоты ω незаполненного резонатора $\omega_g = \omega / \sqrt{\varepsilon \mu}$. Если ε и μ - комплексные величины, то частота оказывается также величиной комплексной: $\omega_g = \omega' - j\omega''$. Мнимая часть ω_g определяет затухание колебаний. И то и другое приводит к потере энергии при колебаниях. Затухающие колебания, возникающие в резонаторе при отсутствии источника энергии, представляют собой суперпозицию частот, расположенных около резонансной частоты ω_0 .

Основы голографии

Таким образом, имеется некоторый узкий интервал частот, внутри которого возможно возбуждение резонатора. Кривая зависимости интенсивности возникающих в резонаторе колебаний от частоты воздействующего на него колебания (резонансная кривая) имеет вполне определенную форму и полуширину. Острота резонансной кривой определяется относительными потерями в резонаторе за один проход, т.е. его добротностью

$$Q = \frac{\omega_0 \text{ Запасенная энергия}}{2\pi \text{ Мощность потерь}} .$$

Если запасенная в резонаторе энергия будет W , то мощность потерь равна взятой с обратным знаком производной по времени от W :

$$\frac{dW}{dt} = -\frac{\omega_0}{2\pi} \frac{W}{Q} ,$$

отсюда

$$W_t = W_0 e^{-\frac{\omega_0}{2\pi Q} t} . \tag{2.60}$$

Таким образом, запасенная в начальный момент времени энергия W_0 , экспоненциально убывает. Коэффициент затухания оказывается обратно пропорциональным добротности резонатора.

Из (2.60) следует, что напряженность электрического поля в резонаторе изменяется в течение времени по закону

$$E_t = E_0 e^{-\frac{\omega_0}{4\pi Q} t} e^{-j\omega_0 t} .$$

Отсюда, представляя E_t через интеграл Фурье, находим

$$E_t = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} E(\omega) e^{-j\omega t} d\omega ,$$

где

$$E(\omega) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} E_0 e^{-\frac{\omega_0 t}{4\pi Q} j(\omega - \omega_0) t} dt .$$

Произведя соответствующее интегрирование, найдем следующее распределение энергии по частоте в резонаторе:

$$|E(\nu)| \approx \frac{1}{(\nu - \nu_0)^2 + [\nu_0 / (2Q)]^2} ,$$

где произведена соответствующая замена ω на $2\pi\nu$. Таким образом, резонансная кривая имеет лоренцову форму (Рис.2.12). На высоте, равной половине максимального значения энергии, ширина спектральной линии резонатора равняется ν_0 / Q . Половина этой величины называется полушириной резонансной кривой. Таким образом, полуширина спектральной линии резонатора обратно пропорциональна его добротности, или, так как добротность обратно пропорциональна потерям, резонансная кривая имеет размытость, прямо пропорциональную потерям. Острая кривая получается только в резонаторе, обладающем весьма малыми потерями.

Основы голографии

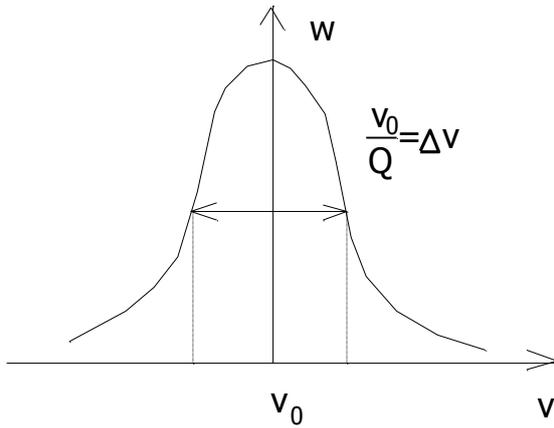


Рис. 2.12. Резонансная кривая резонатора

Из (2.60) следует, что при наличии затухания энергия в резонаторе уменьшается в e раз за время $\tau = 2\pi Q / \omega_0 = Q / \nu_0$. Так как $\nu_0 = 1/T_0$, где T_0 - время, требуемое для одного колебания, то $Q = \tau/T$. Таким образом, добротность резонатора определяет число свободных колебаний, необходимое для уменьшения энергии в нем до 0.368 начального значения, а величина τ может быть рассмотрена как время «звучания» резонатора.

Поскольку добротность определяется потерями энергии, а последние зависят от частоты, то и добротность резонатора является функцией частоты и имеет различные значения для различных собственных частот резонатора. Наибольшей добротностью резонаторы обладают при самых низких типах колебаний (которые в них возникают), т.е. при основных частотах. Это обстоятельство наряду с необходимостью разрежения спектра собственных частот приводит к тому, что размеры закрытых резонаторов выбирают соизмеримыми с рабочей длиной волны.

В оптическом диапазоне необходимо использовать открытые резонаторы, обладающие не только сильно разреженным спектром собственных частот, но и достаточно высокой добротностью на возникших в них колебаниях высших типов. Высокая добротность в открытых системах осуществляется благодаря одному из трех физических явлений, происходящих в них:

- отражения электромагнитных волн от краев резонатора;
- образования каустических поверхностей;
- явления полного внутреннего отражения.

4.1. Активное вещество

Необходимыми элементами любого лазера является ансамбль молекул, представляющих собою рабочее вещество, в котором может быть осуществлена инверсия, т.е. распределение по энергиям, несвойственное термодинамическому равновесию, удовлетворяющее условию $N_m > N_n g_m / g_n$.

В лазерах в качестве рабочего вещества используются газы или твердое тело. В связи с этим при изучении лазеров необходима их классификация.

Основу активной среды твердотельных лазеров составляет матрица. Генерирующими центрами являются частицы, введенные в матрицу. В качестве таких центров используются ионы переходных металлов, редкоземельных элементов или актинидов, энергетические уровни которых используются для создания инверсионной населенности. Относительная концентрация таких активных частиц - активаторов – обычно невелика и составляет единицы или доли процентов от полного числа атомов в среде. Матрица непосредственно не

Основы голографии

участвует в физических процессах, приводящих к генерации, но ее вещество определяет основные технические свойства активной среды.

Вещество матрицы должно удовлетворять следующим требованиям: допускать введение и равномерное распределение активных ионов, на переходах которых осуществляется генерация; быть химически стойким, оптически и механически прочным; прозрачным для излучения накачки и генерации; выдерживать значительный нагрев (возникающий в процессе создания инверсии и генерации излучения); обладать хорошей проводимостью и высокой фотохимической стойкостью; быть технологичным.

В настоящее время в качестве матриц применяются диэлектрические кристаллы с упорядоченной структурой, твердые растворы в виде смешанных разупорядоченных систем и стекла. В качестве кристаллических активных сред широко используются оксидные лазерные кристаллы: рубин $\text{Al}_2\text{O}_3:\text{Cr}^{3+}$, иттриево-алюминиевый гранат (ИФГ) $\text{Y}_3\text{Al}_5\text{O}_{12}:\text{Nd}^{3+}$, александрит $\text{BeAl}_2\text{O}_4:\text{Cr}^{3+}$, гадолиний-скандий-галлиевый гранат (ГСГГ) $\text{Gd}_3\text{Sc}_2\text{Ga}_3\text{O}_{12}:\text{Nd}^{3+}$ и др. – фтористые кристаллы: CaF_2 , BaF_2 , KZnF_3 и др. В качестве активаторов используются главным образом ионы группы железа (Cr^{3+} , V^{2+} , Ni^{2+} , Co^{2+} и др.), а также ионы редкоземельных элементов (Nd^{3+} , Pr^{3+} , Eu^{3+} и др.).

В кристаллах с разупорядоченной структурой типа $\text{CaF}_2-\text{SrF}_2$, BaF_2-YF_3 , $\text{LiCd}(\text{MoO}_4)_2$, $(\text{Nd},\text{Sc})\text{P}_5\text{O}_{14}$ приместные ионы (в основном неодима) входят в состав многих отличающихся по структуре активных центров, в результате чего спектры поглощения этих веществ состоят из широких полос, что позволяет увеличить эффективность накачки.

В качестве аморфных диэлектриков применяют стекла различных составов. В качестве активаторов в стеклянных матрицах обычно используется двух- и трехзарядные ионы редкоземельных элементов неодима, иттербия, гадолиния и др. В отличие от кристаллов ионы активатора входят в стекло как компоненты стекла. При производстве лазерных стекол к ним предъявляются повышенные требования к чистоте исходных материалов, однородности и спектральным свойствам.

Основным достоинством стеклянных активных сред является простота изготовления АЭ любой формы и размеров, высокая оптическая однородность, изотропность свойств, возможность введения активатора в необходимых концентрациях (с равномерным распределением его по АЭ). Однако, по ряду параметров стекло уступает кристаллам. Спектральные линии излучения в стекле расширены, что является следствием отсутствия определенной кристаллической структуры. Стекла имеют более низкий квантовый выход излучения, худшие теплофизические свойства (малую теплопроводность и высокий коэффициент термического расширения); ограниченную область прозрачности.

В жидкостных лазера инверсия осуществляется при помощи оптической накачки.

Основы голографии

В полупроводниковых лазерах рабочим телом является кристалл полупроводника. Генерация осуществляется при рекомбинации неравновесных электронов и дырок при соответствующих переходах носителей. Полупроводниковые лазеры подразделяются на две основные подгруппы.

Лазеры, в которых рабочим телом являются кристаллы-полупроводники, изготовленные таким образом, что одна их часть обладает электронной проводимостью (n – тип), а другая – проводимостью, обусловленной наличием дырок (p – тип). В этих условиях на границе, называемой p - n – переходом, возникает большой градиент концентрации как электронов, так и дырок. Инверсия в данном случае осуществляется при инжекции носителей тока через p - n - переход под влиянием прилагаемой к кристаллу разности потенциалов. Полупроводниковые лазеры, относящиеся к данной подгруппе, получили название инжекционных ПКГ, типичным представителем их является лазер на p - n – переходе в арсениде галлия.

Ко второй подгруппе относятся лазеры, в которых рабочим телом являются кристаллы полупроводников одного типа. В этом типе лазеров возбуждение осуществляется при бомбардировке кристалла полупроводника электронным пучком или достигается оптической накачкой. В качестве рабочих тел используются, например, полупроводники-кристаллы сульфида кадмия CdS , арсенида галлия $GaAs$, теллура свинца $PbTe$ и др.

В последние годы большое внимание уделяется химической накачке, когда инверсия возникает при той или другой химической реакции. Химическая реакция может быть осуществлена либо при фотодиссоциации молекул, как, например, в лазере на атомарном газе, либо в электрическом разряде в смеси газов, например, $Ne - O_2$, $Ar - O_2$. В результате этих процессов образуются атомы или молекулы в возбужденном состоянии.

Лекция № 5

5.1. Селекция мод

Уже отмечалось, что в резонаторах может одновременно возбуждаться большое число собственных колебаний. Число колебаний в интервале рабочей частоты, возникающих в открытых резонаторах, значительно меньше, чем в равноценных по размерам закрытых резонаторах. Однако это число оказывается еще очень большим. Ввиду того, что линия люминесценции активного вещества значительно шире интервала частот между отдельными колебаниями в резонаторе, излучение лазера представляет собой совокупность различных частот, лежащих внутри линии люминесценции. Многомодовый режим работы генератора приводит к значительному ухудшению таких основных свойств лазерного излучения, как монохроматичность, когерентность, угловая расходимость. При применении лазера иногда необходимо иметь одномодовый, одночастотный лазер. В связи с этим применяют специальные методы подавления нежелательных колебаний высших порядков, т.е. осуществляется селекция колебаний. При подавлении колебаний высших порядков общая энергия излучения заметно не изменяется, но она при этом сосредотачивается в одном типе колебаний. При осуществлении одномодового одночастотного режима работы лазера, очевидно, должны быть приняты меры к селекции как поперечных, так и продольных колебаний.

Селекция поперечных колебаний заключается в подавлении высших колебаний, т.е. с индексами m и n , отличными от нуля. В результате селекции в резонаторе остаются только продольные колебания вида TEM_{00q} , при этом мощность продольных колебаний может увеличиваться за счет увеличения числа активных частиц, которые ранее участвовали в формировании поперечных мод.

Существует много различных методов селекции поперечных колебаний, рассмотрим лишь наиболее часто встречающиеся.

Наиболее распространенным и простым способом подавления поперечных колебаний является использование диафрагмы, которая располагается между рабочим веществом и одним из зеркал. Поскольку поле в резонаторе, соответствующее основному типу TEM_{00q} , сосредотачивается в основном вблизи оси, а поле более высоких типов проецируется на значительно большие расстояния от оси, диафрагма, расположенная на оси, будет оказывать более сильное влияние на распространение высоких типов. Подбирая диафрагмы, можно осуществить подавление всех типов колебаний, кроме основного TEM_{00q} . Однако и для основного типа колебаний диафрагма вносит значительные потери, поэтому данный метод селекции имеет существенный недостаток, заключающийся в уменьшении выходной мощности при переходе на одномодовый режим.

Основы голографии

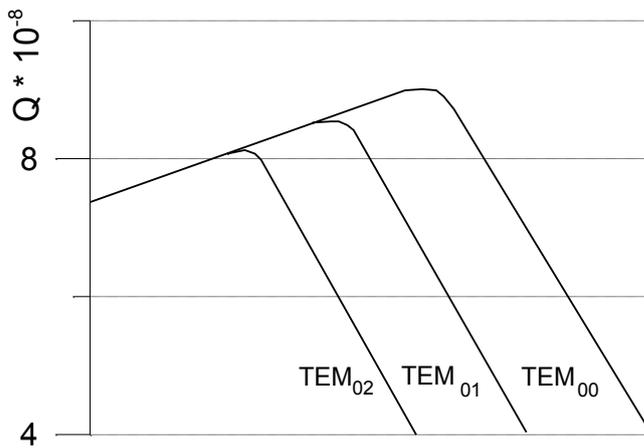


Рис. 2.26. Зависимость добротности резонатора от числа зон Френеля N для различных типов колебаний

Значительно более эффективным является метод, основанный на увеличении дифракционных потерь колебаний высших порядков при сравнительно меньшем их увеличении для колебаний основного типа. Потери в резонаторе зависят от числа Френеля, причем потери различны для различных типов. При уменьшении числа Френеля потери быстро растут, при этом увеличивается и разность между потерями колебаний основного типа и потерями колебаний более высоких

порядков. При $N \ll 1$ эта разность достигает таких значений, что оказывается возможным сравнительно просто осуществить селекцию. Для данной цели особенно пригоден резонатор, состоящий из сферического и плоского зеркал, расположенных на расстоянии, приблизительно равном радиусу кривизны сферического зеркала (плоско-сферический резонатор). Дифракционные потери в таком резонаторе определяются в основном числом Фарадея для сферического зеркала

$$N_{cф} = \frac{a_2^2}{\lambda \sqrt{bR}} \sqrt{1 - \frac{b}{R}} \quad (2.79)$$

Из (2.79) следует, что при b/R , близком к единице, $N_{cф}$ может иметь настолько малые значения, при которых разница в дифракционных потерях различных мод оказывается вполне достаточной для осуществления селекции. Зависимость добротности резонатора от числа зон Френеля N для различных типов колебаний представлена на Рис. 2.26. Селекция поперечных мод может быть осуществлена путем использования сильной зависимости коэффициента отражения от угла вблизи угла полного внутреннего отражения. В этом случае используется резонатор, имеющий в качестве одного из отражателей призму, в которой могут происходить многократные внутренние отражения. Призма обладает высоким отражением только в направлении, составляющем весьма малый угол с направлением выбранной оси (около 1°). В призме, угол которой θ_{np} равен критическому углу полного внутреннего отражения $\theta_{кр}$, коэффициенты отражения по амплитуде для лучей с параллельной и перпендикулярной поляризацией по отношению к плоскости падения при условии, что угол $\delta = (\theta_{кр} - \theta_{пад}) \ll 1$, соответственно равны:

$$\rho_{\Pi} = 1 - \left[(2N)^{3/2} / (1 - 1/N^2) \right]^{1/4} \delta^{1/2} ;$$

Основы голографии

$$\rho_{\perp} = 1 - \left[2^{3/2} / \left(1 - 1/N^2 \right)^{1/4} N^{1/2} \right] \delta^{1/2}$$

Легко найти относительную интенсивность отраженного света для призмного селектора мод. Например, при угле призмы θ_{np} несколько больше $\theta_{кр}$, коэффициент отражения в диапазоне углов $2(\theta_{np} - \theta_{кр})$ примерно равен 1.0.

Угловая селекция может быть осуществлена также с помощью интерферометра Фабри - Перо, устанавливаемого внутри резонатора. Так как интерферометр пропускает излучение не только в зависимости от длины волны, но и от направления распространения волны, боковые моды испытывают большие потери.

Рассмотрим основные методы подавления продольных мод. Использование вышеуказанных методов позволяет добиться излучения, содержащего только одну поперечную моду TEM_{00q} . При этом угловая расходимость луча лазера значительно уменьшается, приближаясь к значению, определяемому дифракционной расходимостью. Однако при этом в пределах контура усиления активной среды будет содержаться значительное число продольных колебаний, соответствующих различным значениям индекса. Интервал между соседними продольными типами колебаний будет зависеть от длины резонатора и определяться величиной $c/2L$. В ряде случаев возникает необходимость получения одночастотного излучения - колебания лишь на одной продольной волне. Для этого могут быть использованы различные методы. Наиболее эффективным из них является метод сложных резонаторов, представляющих собой связанные резонаторы, позволяющие реализовать продольные типы колебаний различной интенсивности. Изменяя коэффициент связи между резонаторами и их длины, можно осуществить возбуждение одного или нескольких типов колебаний. Простейшим видом такого резонатора является линейный трехзеркальный резонатор, представленный на рис.2.27. Допустим, что зеркала плоские, бесконечных размеров.

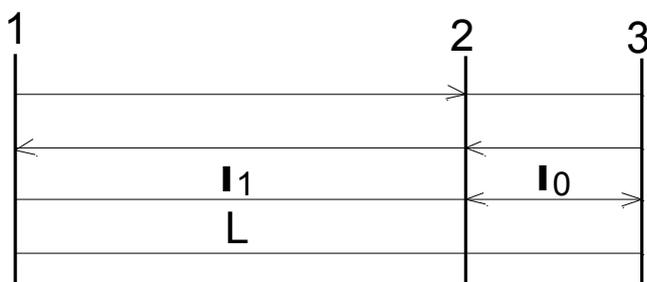


Рис. 2.27. Селектор продольных мод в виде двух связанных резонаторов

Основы голографии

Считая, что в такой системе плоские электромагнитные волны распространяются в направлении нормали к зеркалам, соотношения между комплексными амплитудами отраженных волн от соответствующих зеркал будут иметь следующий вид:

$$\begin{aligned} A_1 e^{\gamma l_1} &= A_2 \rho_1; & A_3 e^{\gamma l_0} &= A_1 \tau_2 + A_4 \rho_2; \\ A_2 e^{\gamma l_1} &= A_1 \rho_2 + A_4 \tau_2; & A_4 e^{\gamma l_0} &= A_3 \rho_3 \end{aligned}, \quad (2.80)$$

где γ - постоянная распространения волн в среде, заполняющей резонаторы; ρ_i и τ_i - коэффициенты отражения и пропускания по полю.

Решение данной системы уравнений находят при приравнении нулю ее определителя. Раскрывая определитель, получим уравнение

$$e^{2\gamma(l_1+l_0)} - \rho_2 \rho_3 e^{2\gamma l_1} - \rho_1 \rho_2 e^{2\gamma l_0} + \rho_1 \rho_3 (2\rho_2^2 - 1) = 0. \quad (2.81)$$

Полученные из этого уравнения значения γ позволяют определить собственные частоты резонатора и их потери.

Так, если бы среднее зеркало отсутствовало ($\rho_2 = 0$), то рассчитанный по (2.81) спектр собственных частот был бы эквидистантным с интервалом между частотами, равным $\Delta \nu_q = c/(2L)$ и представленным на рис. 2.28,а.

В случае, когда среднее зеркало имеет коэффициент отражения $\rho_2 = 0.65$, а крайние зеркала $\rho_1 = \rho_3 = 0.99$, то частоты, соответствующие корням уравнения (2.81), будут равны (при $l_1 = 3l_0$):

$$\begin{aligned} \nu_{q1} &= \frac{c}{8l_0} 4q; & \nu_{q3} &= \frac{c}{8l_0} (4q + 2); \\ \nu_{q2} &= \frac{c}{8l_0} (4q + 1.14); & \nu_{q4} &= \frac{c}{8l_0} (4q + 2.86) \end{aligned}.$$

При этом амплитуды колебаний в относительных единицах будут соответственно равны 1.2; 0.76; 0.24; 0.76 (см. Рис.2.27,б).

Таким образом, в системе связанных резонаторов добротность для разных типов колебаний существенно различна. Максимальную добротность в сложных резонаторах будут иметь частоты, совпадающие для различных его ветвей. Эффективность данного метода селекции может быть повышена путем использования нескольких связанных резонаторов.

Лекция № 6

Режимы работы лазеров

Динамика работы лазеров и, как следствие, его основные параметры (энергия и длительность импульса, средняя и пиковая мощность, частота следования импульсов и т.п.) в значительной мере определяют области его применения. Знание динамики процессов в лазере является необходимым условием решения практической проблемы управления параметрами лазерного излучения. В существующих лазерах реализованы как импульсные, так и непрерывные режимы работы. Они позволяют получать импульсы длительностью до долей пикосекунд; дают возможность работать как в режиме одиночных импульсов, так и их последовательностей. Каждый из режимов работы имеет свои особенности. В режиме свободной генерации обеспечиваются наиболее высокие уровни энергии излучения при наибольших значениях КПД, в режиме модулированной добротности – наиболее высокие уровни импульсной мощности излучения при несколько более низких КПД и т.д.

6.1. Стационарный режим генерации

Основным критерием для классификации режимов является соотношение между длительностью стационарной генерации $\tau_{\text{стац}}$, когда мощность излучения практически неизменна, и длительностью развития генерации и гашения, которые в первом приближении определяются постоянной времени резонатора $\tau_{\text{рез}}$, характеризующей переходный процесс в резонаторе. Очевидно, это то время, за которое плотность энергии в резонаторе уменьшается до уровня $1/e$ от исходного при «отключении» поля:

$$\tau_{\text{рез}} = Q/(2\pi\nu) = 2L\ln/(c\sum\alpha_{\text{рез}}),$$

где $\sum\alpha_{\text{рез}}$ - суммарные потери резонатора за один проход.

Режим работы лазера считается стационарным, если выполняется условие $\tau_{\text{стац}} \gg \tau_{\text{рез}}$. Теоретически стационарный режим работы может быть исследован, если в балансных уравнениях принять, что накачка является стационарной и действует неограниченное время. В системе также отсутствуют другие факторы, приводящие к изменению населенности энергетических уровней, т.е. нужно положить $dN/dt = 0$. Таким образом, стационарная генерация есть не более как идеализация, пригодная для описания отдельных реальных ситуаций лишь в некотором приближении.

Реализуемые на практике режимы работы лазеров, по существу, являются нестационарными, что приводит к флуктуациям энергетических и пространственных характеристик излучения. Причин нестационарности лазерной

Основы голографии

генерации довольно много. Их условно можно разделить на два типа: имеющие техническую природу и физическую.

К факторам, имеющим техническую природу, относятся нестабильность геометрических параметров резонаторов, активной среды, системы накачки и т.п. Все источники технических флуктуаций лазерного излучения оказывают свое влияние на генерацию через два основных параметра лазера: коэффициент усиления активной среды и коэффициент потерь излучения в резонаторе. Оба эти параметра под воздействием источников нестабильностей изменяются во времени, и эти флуктуации приводят к нестабильности процесса генерации.

Основными причинами, приводящими к нестабильности излучения, являются механические вибрации зеркал резонатора, активного элемента и других внутренних элементов резонатора. К появлению вибрации приводят различного рода механические колебания помещения, работа системы охлаждения и элементов управления излучения. Источником технических нестабильностей лазерного излучения является также неравномерность нагрева активного элемента в процессе накачки, что вызывает флуктуации температуры активного элемента, приводящие к изменению его плотности, показателя преломления и влияющие на населенность нижнего рабочего уровня. Тепловые эффекты приводят также к перестройке мод, смещению и изменению ширины линии флюоресценции.

К факторам, имеющим физическую природу, прежде всего следует отнести многомодовость излучения лазера. Разным модам соответствует разное распределение параметров поля в резонаторе, разное значение коэффициентов усиления и потерь. Стабильность кинетики генерации сильно зависит от количества одновременно возбуждаемых мод. Изменение числа или индекса генерируемых одновременно мод сразу вызывает нестабильность генерируемого излучения.

Кроме того, нужно иметь ввиду, что первопричиной лазерного излучения являются случайные процессы спонтанного испускания. И хотя в конечном итоге эти процессы упорядочиваются в пространстве и во времени в результате специфических процессов, происходящих в резонаторе лазера, этот процесс не является полным.

Расчеты энергетических параметров генерации обычно проводят для усредненной вдоль длины резонатора плотности энергии. Такое приближение позволяет существенно упростить решение задачи о стационарной генерации. Для описания потоков энергии в этом случае используют уравнения переноса (базирующиеся на законах лучевой оптики). Для двух потоков (распространяющихся в противоположных направлениях) они имеют вид:

$$\begin{aligned} dP_1/dz &= (\chi - \chi_{nom})P_1; \\ dP_2/dz &= -(\chi - \chi_{nom})P_2 \end{aligned} \quad (1.1)$$

Здесь P_1 и P_2 - потоки энергии через единицу поверхности поперечного сечения активной среды в единицу времени (Рис. 1.1,а); χ - средний по длине коэффициент усиления активной среды; χ_{nom} -средний по длине коэффициент вредных потерь. Дополним уравнения переноса граничными условиями (на

Основы голографии

торцах активной среды), предполагая, что активная среда занимает весь объем резонатора:

$$P_1|_{z=0} = \rho_1 P_2; \quad P_2|_{z=l} = \rho_2 P_1,$$

где ρ_1 и ρ_2 - коэффициенты отражения зеркал резонатора; l - длина активной среды.

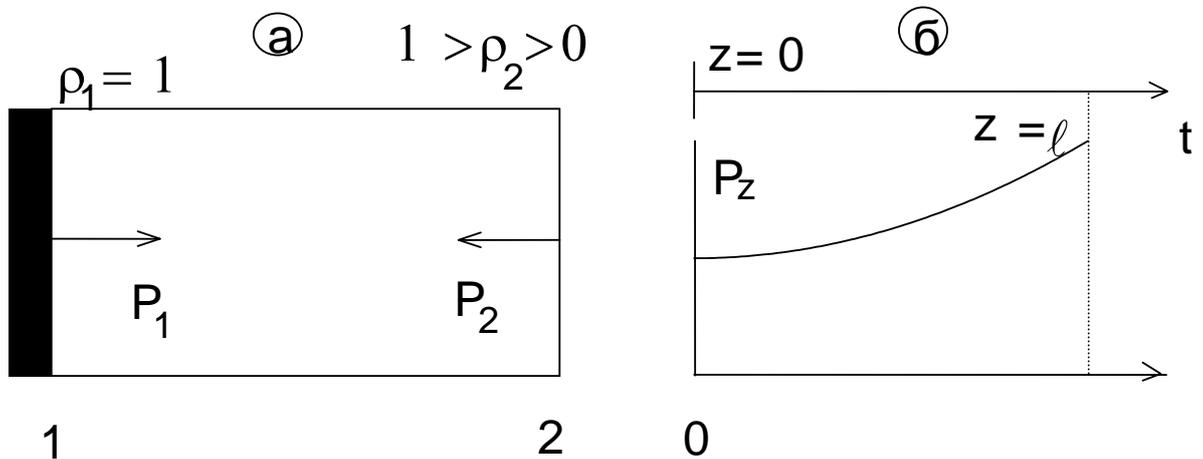


Рис.1.1. Схема распространения потоков энергии (а) и распределения плотности энергии (б) в резонаторе

Можно показать, что в режиме генерации коэффициент усиления не зависит от параметров дифференциальных уравнений (1.1) и однозначно определяется граничными условиями.

В стационарном режиме среднее значение коэффициента усиления зависит только от среднего значения коэффициента вредных потерь $\chi_{ном}$, длины стержня l и коэффициентов отражения ρ_1 и ρ_2 :

$$\chi = \chi_{пот} + \frac{1}{l} \ln \frac{1}{\sqrt{\rho_1 \rho_2}}. \quad (1.2)$$

По физическому смыслу второе слагаемое в (1.2) определяет полезные потери, связанные с выводом излучения. Поэтому его обычно называют коэффициентом полезных потерь $\chi_{н.п.}$.

Зная коэффициент полезных потерь и среднее значение коэффициента вредных потерь можно найти количество энергии, выходящей из резонатора в единицу времени

$$P = \frac{vls}{\alpha} [\chi_0 - \chi_{пот} - \chi_{н.п.}] \frac{\chi_{н.п.}}{\chi_{пот} + \chi_{н.п.}},$$

где $ls = V$ - объем активного элемента; v - скорость распространения излучения; χ_0 - начальный коэффициент усиления; α - параметр нелинейности (который характеризует нелинейность взаимодействия частиц с внешним излучением, обусловленную насыщением распределения частиц по уровням).

6.2. Работа лазеров в режиме модуляции добротности

На практике свободное излучение лазера представляет собой, как правило, последовательность относительно коротких импульсов - «пичков». В условиях свободной генерации лазер испускает световые импульсы длительностью порядка 100 нс. Однако этого недостаточно для ряда применений. Требуется моноимпульсы более короткой длительности и более высокой мощности.

Задача генерации коротких одиночных импульсов обычно решается путем использования резонаторов с управляемой добротностью. Этот режим еще называют режимом генерации гигантских импульсов. Мощность получаемых импульсов достигает несколько гигаватт, а длительность имеет порядок 10-100 нс (может быть понижена до 1-3 нс).

Принцип действия лазеров с управляемой добротностью основан на сокращении времени излучения посредством накопления активных центров на метастабильном уровне. Этого достигают, помещая в резонатор лазера модулятор, управляемый внешним сигналом.

Управляя добротностью резонатора, обеспечивают высокий уровень вредных потерь. Это приводит к увеличению порога генерации и позволяет создать большую инверсионную населенность в активной среде (что возможно в случае большого времени жизни верхнего уровня). В этих условиях формируется единый короткий световой импульс большой мощности.

Режим модуляции добротности лазера характеризуется следующими основными параметрами: временем задержки генерации $t_{зад}$, длительностью гигантского импульса τ_u , пиковой мощностью P^{max} и энергией W излучения импульса.

В общем виде балансные уравнения для нестационарного процесса были сформулированы Статцем и де Марсом. Искомые функциями в них являются плотность инверсионной заселенности $N(t)$ и плотность числа фотонов на частоте генерации $M(t)$. Система балансных уравнений имеет вид

$$dM(t)/dt = B^1 M(t)N(t) - M(t)f(t)/\tau_{pez};$$

$$dN(t)/dt = -\beta B^1 M(t)n(t),$$

где $B^1 = B\hbar\omega$ - коэффициент Эйнштейна для вынужденных переходов в канале генерации, помноженной на энергию фотона; τ_{pez} - время жизни фотона в резонаторе, определяемом совокупностью вредных и полезных потерь $Q = \omega/\tau_{pez}$; β - целое число, описывающее изменение разности заселенностей рабочих уровней при излучении одного фотона; $f(t)$ - некоторая функция от времени, изменяющаяся от значений $f(0) = Q/Q_{min}$ до 1. Плотность числа фотонов на частоте генерации легко выражается через плотность светового потока $S(t)$, скорость распространения излучения и энергию фотона

$$M(t) = S(t)/(v\hbar\omega).$$

Основы голографии

Развитие гигантского импульса рассматривают, выделяя три основных временных этапа, которые обычно характеризуют изменением во времени мощности генерируемого излучения $P(t)$, добротности резонатора $Q(t)$, плотности инверсионной заселенности $N(t)$ и пороговой плотности инверсионной заселенности $N(t)_{пор}$, которые представлены на Рис. 1.2.

Под действием импульса накачки в активном элементе возрастает плотность инверсионной заселенности $N(t)$. В какой-то момент происходит включение добротности, которая начинает быстро расти, что приводит к понижению порога генерации. В момент времени, когда пороговое значение $N(t)_{пор}$ сравняется с плотностью инверсионной заселенности $N(t)$, начнется процесс генерации. Это время обычно выбирают в качестве начального момента.

На следующем этапе наблюдается медленное развитие генерации. Для него характерно практически постоянное значение плотности инверсионной заселенности. На линейном этапе мощность генерируемого излучения увеличивается незначительно, и быстрое ее возрастание наблюдается лишь в конце этапа.

Этап формирования гигантского импульса существенно меньше этапа линейного развития. В течение короткого промежутка времени плотность инверсионной заселенности $N(t)$ падает до минимального значения N_{min} (которое оказывается меньше конечного порогового значения $N_{пор}$). Практически вся энергия, сосредоточенная в гигантском импульсе, высвечивается на данном этапе.

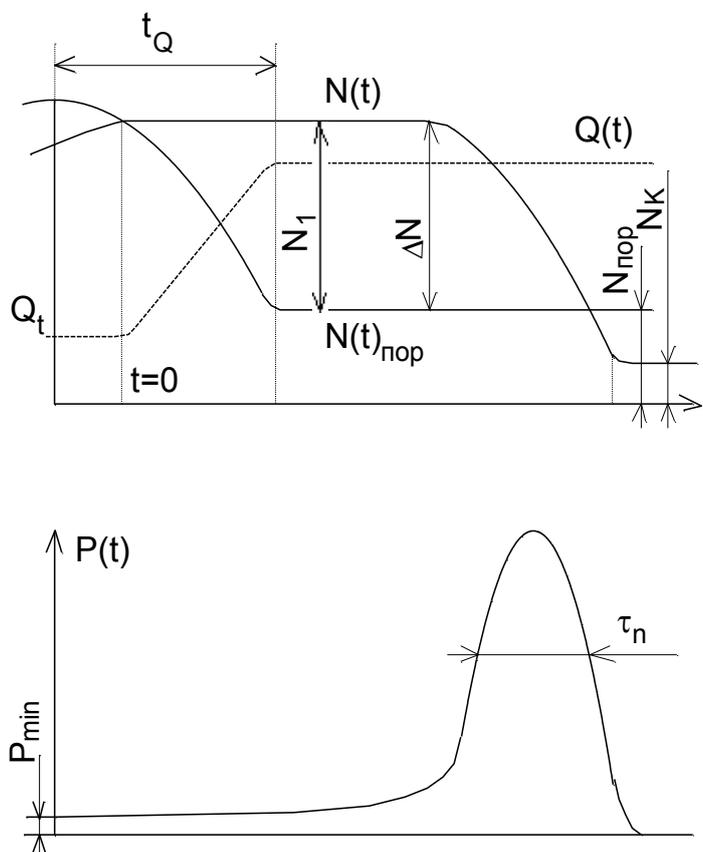


Рис.1.2. Схема формирования гигантского импульса

Для увеличения энергии гигантского импульса необходимо увеличивать превышение начального значения плотности инверсионной заселенности (к моменту генерации) по отношению к пороговому значению плотности инверсионной заселенности, т.е. увеличивать разность $\Delta N = N_1 - N_{nop}$. Для этого обеспечивают глубокую модуляцию добротности, высокую интенсивность накачки и согласовывают момент времени включения добротности с импульсом накачки. Добротность должна включаться в тот момент времени, когда плотность инверсионной заселенности приблизится к своему предельному значению.

Мощность P генерируемого излучения можно представить в виде

$$P = \hbar\omega v \chi_{n.n.} MV,$$

где V - объем, участвующий в генерации; v - скорость распространения излучения; $\chi_{n.n.}$ - коэффициент полезных потерь. Оптимальное значение добротности резонатора (которое позволяет получить максимальную пиковую выходную мощность при фиксированном N_1) соответствует случаю $N_1/N_{nop} = 3.5$.

Оценим энергию гигантского импульса. В процессе излучения импульса плотность инверсионной заселенности изменяется от N_1 до N_k . Это соответствует рождению в единице объема активной среды $(1/\beta)(N_1 - N_k)$ фотонов. Умножив это число на энергию фотона, генерирующий объем активной среды V и отношение $\chi_{n.n.}/(\chi_{nom} + \chi_{n.n.})$, описывающее долю генерируемой световой энергии, выходящей через зеркало резонатора в виде полезного излучения, получим

$$W = (\hbar\omega/\beta)V(N_1 - N_k)\chi_{n.n.}/(\chi_{nom} + \chi_{n.n.}).$$

Длительность импульса можно получить на основе следующего выражения

$$\tau_{n.n.} \approx W/P_{max}.$$

6.3. Модуляция добротности резонаторов

Все известные методы модуляции добротности резонаторов подразделяются на активные и пассивные. К активным относятся методы, в которых модулирующие устройства меняют значение потерь по заданному закону. Пассивные модулирующие элементы управляются непосредственно полем излучения, имеющимся в резонаторе лазера.

Модулирующие устройства можно разделить по характеру физических эффектов (на которых основано их действие) на следующие: оптико-механические, электрооптические и акустические. При оптико-механической модуляции добротности модулирующее устройство, как правило, представляет собой вращающуюся призму-крышу, выполняющую функции одного из зеркал резонатора (Рис.1.3,а). Время включения добротности в этом случае зависит от частоты вращения призмы. При частотах вращения призмы 20000-30000 об/мин время переключения добротности имеет порядок 10^{-7} с. Время коммутации в этом

Основы голографии

случае достаточно велико. Для увеличения скорости переключения в резонатор можно ввести дополнительные неподвижные отражатели (Рис.1.3,б).

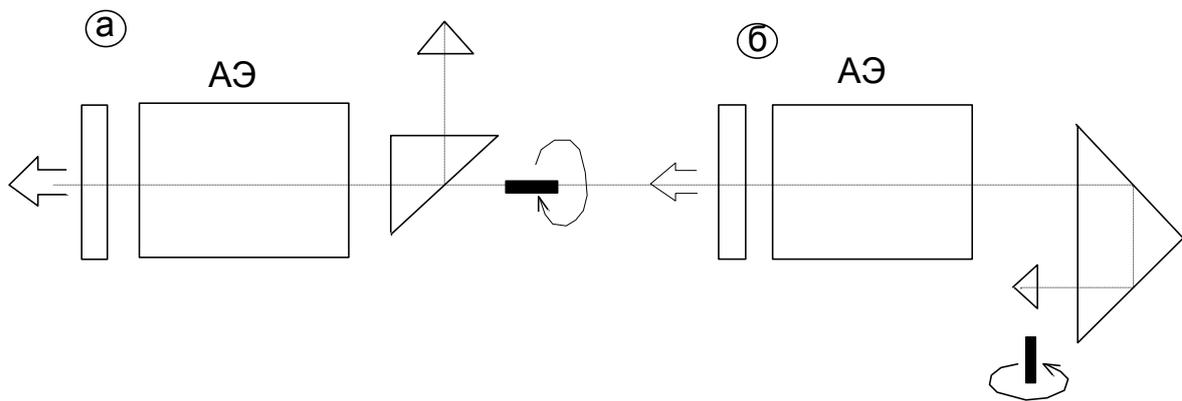


Рис. 1.3. Оптические схемы с оптически-механической модуляцией добротности (АЭ – активный элемент)

Оптико-механические затворы с вращающей призмой, несмотря на колебания оси излучателя, низкий КПД, значительные шумы и вибрацию двигателя, нашли широкое применение в лазерах с управляемой генерацией благодаря простоте конструкции, надежности, а также возможности работы в широком температурном и спектральном диапазонах.

Лекция № 7

Синхронизация мод в лазерах

В условиях свободной генерации зависимость интенсивности выходного излучения лазера от времени, как правило, имеет вид нерегулярных пиков, одной из причин этого является генерирование большого числа различных поперечных и продольных мод. Наличие внутри резонатора лазера различных оптических элементов обычно приводит к возникновению локальных резонаторов, препятствующих возбуждению некоторых типов колебаний за счет интерференционной селекции мод. Каждый такой резонатор обладает определенной полосой пропускания, поэтому совместное действие всех резонаторов может существенно ограничить число мод, возбуждаемых в системе. Для устранения этого явления используют так называемые неселективные резонаторы, в которых зеркальные покрытия наносят на клиновидные подложки 1 и 2, а торцы активного элемента (АЭ) срезают под углом Брюстера (Рис.1.4).

В такой системе возникает многочастотная генерация со спектральной шириной, близкой к действительной ширине эмиссионной линии данного лазерного перехода. Это удобно для синхронизации мод, так как эффективность данного режима возрастает с увеличением числа генерируемых продольных мод. Под синхронизацией мод понимают процесс, в результате которого разности фаз между соседними модами остаются постоянными. В результате интерференции продольных мод появляются флуктуации интенсивности и происходит генерация весьма коротких световых импульсов очень большой мощности.

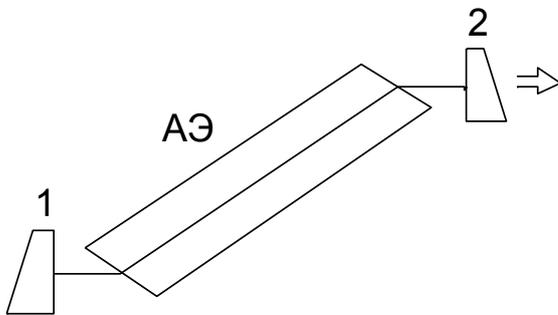


Рис. 1.4. Схема лазера с неселективным резонатором

Синхронизация фаз продольных мод лазера позволяет получать импульсы лазерного излучения длительностью 10^{-10} - 10^{-12} с пиковой мощностью до 10^{12} Вт.

Идею синхронизации продольных мод с точки зрения их интерференции можно объяснить следующим образом. Пусть в резонаторе все моды генерируют одновременно, амплитуды их одинаковы, а фазы синхронизированы. Пусть число таких мод равно $2N+1$, амплитуды мод равны E_0 , а разность фаз $\Phi_K - \Phi_{K-1} = \alpha = const$.

Тогда результирующее электрическое поле $E(t)$ электромагнитной волны можно представить в виде

Основы голографии

$$E(t) = \sum_{l=-N}^N E_0 e^{i[(\omega_0 + l\Delta\omega)t + l\alpha]}, \quad (1.3)$$

где ω_0 - центральная частота моды; $\Delta\omega = \pi c/L$ - расстояние между соседними модами. Для простоты предполагаем, что фаза центральной моды равна нулю. Выполняя суммирование в выражении (1.3), найдем

$$E(t) = A(t) e^{i\omega_0 t},$$

где

$$A(t) = \frac{\sin [(2N + 1)(\Delta\omega t + \alpha)/2]}{\sin [(\Delta\omega t + \alpha)/2]} E_0. \quad (1.4)$$

Таким образом, $E(t)$ представляет собой несущую синусоидальную волну с центральной частотой ω_0 и амплитудой $A(t)$. Выходная мощность излучения пропорциональна $A^2(t)$. Временная зависимость интенсивности излучения при генерации $2N + 1 = 7$ и $2N + 1 = 15$ модам приведена на Рис.1.5.

Таким образом, поле в резонаторе оказывается промодулированным по амплитуде функцией $A(t)$. Выполнение условий синхронизации мод приводит к их взаимной интерференции, и генерация излучения принимает вид последовательности коротких световых импульсов. Максимумы амплитуды огибающей последовательности импульсов соответствуют тем моментам времени, когда знаменатель выражения (1.4) обращается в ноль. Следовательно, временной интервал между соседними импульсами равен

$$\tau = 2\pi/\Delta\omega = 2L/c.$$

Это время есть так называемый период модуляции и соответствует полному проходу резонатора. Следовательно, генерацию можно представить как импульс излучения, распространяющийся в резонаторе туда и обратно. Временной интервал $\Delta\tau$ между пиком импульса и первым нулевым уровнем дается выражением

$$\Delta\tau = 1/\Delta\nu_{ген},$$

где $\Delta\nu_{ген} = (2T + 1)\Delta\omega/2\pi$ - полная ширина спектра генерируемого излучения. Длительность импульса τ_u , измеренная на полувывоте, также приблизительно равна $\Delta\tau$.

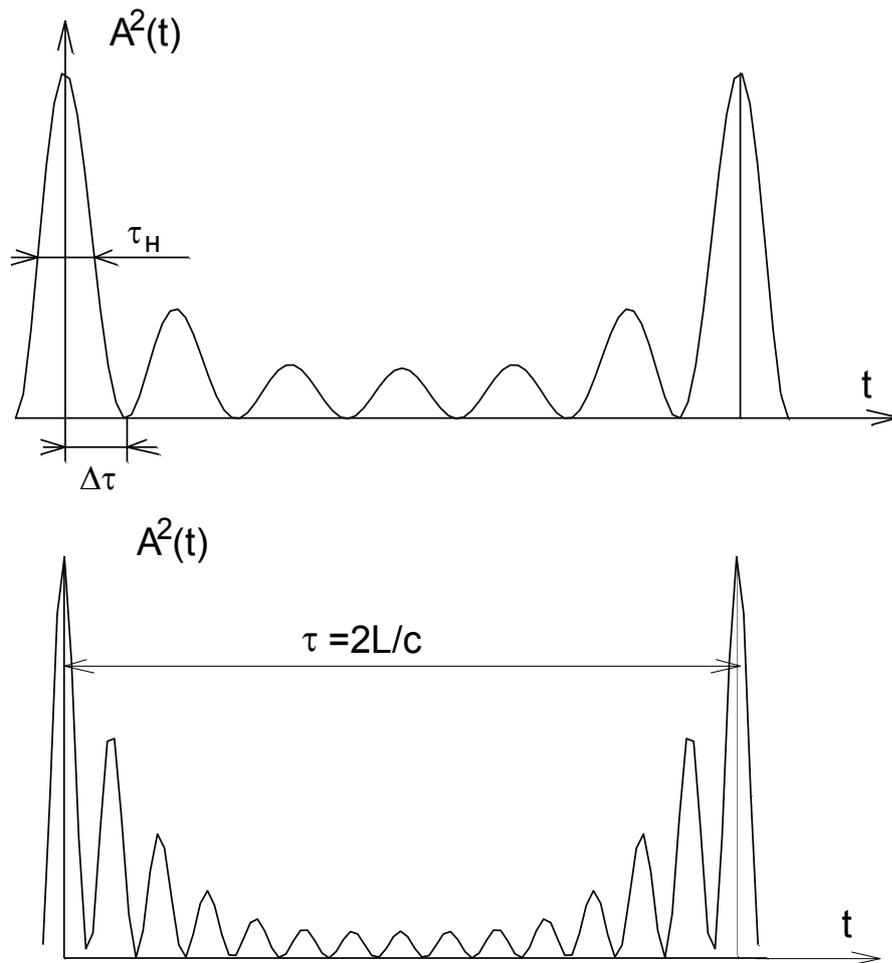


Рис. 1.5. Временная зависимость мощности излучения в случае генерации семи (а) и 15-ти (б) продольных мод

Таким образом, для получения коротких световых импульсов требуются активные материалы с большой шириной контура линии усиления лазера. Пиковая мощность излучения в режиме синхронизации мод пропорциональна $(2N + 1)^2 A^2 t$, тогда как, если фазы мод случайны, мощность излучения пропорциональна сумме мощностей отдельных мод. Отсюда следует, что благодаря синхронизации мод пиковая мощность увеличивается пропорционально числу синхронизированных мод, в то время как средняя мощность практически не зависит от синхронизации мод.

7.1. Синхронизация мод в лазерах

В методах синхронизации мод различают активную и пассивную синхронизацию. Первая осуществляется под действием внешнего сигнала, управляющего работой модулятора, вторая - под влиянием поля излучения спонтанно с помощью нелинейной оптической среды.

Основы голографии

Для иллюстрации активной синхронизации мод предположим, что осуществляется периодическая модуляция параметров резонатора с частотой Ω , равной или кратной разности частот соседних мод. Модуляция амплитуды центральной моды имеет вид

$$E(t) = E_0 [1 + \beta \cos(\Omega t)] e^{i(\omega_0 t + \varphi_0)},$$

где β - глубина модуляции ($0 \leq \beta \leq 1$). Используя формулу Эйлера для представления \cos , получим

$$E(t) = E_0 e^{i(\omega_0 t + \varphi_0)} + (\beta E_0 / 2) e^{i(\omega_0 + \Omega)t + i\varphi_0} + (\beta E_0 / 2) e^{i(\omega_0 - \Omega)t + i\varphi_0}.$$

Таким образом, периодическая амплитудная модуляция с частотой Ω приводит к превращению моды на несущей частоте ω_0 в три синфазные моды на частотах ω_0 , $\omega_0 + \Omega$, $\omega_0 - \Omega$, что вызовет обмен энергией с двумя соседними модами. Поля боковых составляющих соответствуют собственным частотам резонатора и лежат в пределах контура линии усиления лазерного перехода, что приводит к их усилению. Они тоже подвергаются модуляции с частотой Ω , в результате чего возникают боковые составляющие с частотой $\omega_0 \pm 2\Omega$. Этот процесс продолжается до тех пор, пока все аксиальные моды внутри области генерации не окажутся синхронизованными.

Лазер с активной синхронизацией мод содержит амплитудный или фазовый модулятор, размещаемый вблизи одного из зеркал резонатора. Амплитудная модуляция осуществляется акустооптическим или электрооптическим модулятором, применяемым совместно со скрещенными поляризаторами.

Пассивная синхронизация мод технически осуществляется гораздо проще. В этом случае периодическая модуляция параметров резонатора может быть выполнена с помощью находящегося внутри резонатора нелинейного оптического элемента, например, просветляющего фильтра.

С точки зрения интерференционных представлений каждая пара продольных мод, возбуждаемых в лазере с просветляющимся фильтром, будет периодически модулировать коэффициент резонансного поглощения фильтра с частотой, равной разности частот рассматриваемых мод. В результате этого может реализоваться процесс синхронизации между фазами различных мод.

Синхронизация продольных мод в лазере с просветляющимся фильтром может быть рассмотрена на основе использования флуктуационных представлений. Под действием импульса накачки происходит возбуждение активной среды и начинается спонтанное излучение на частоте рабочего перехода. Спонтанное излучение является по своей природе флуктуационным процессом и его интенсивность флуктуирует с характерным временем корреляции $\tau_{кор} \approx 1/\Delta\omega_{люм}$. К тому моменту, когда вклад усиленных спонтанных шумов начинает превышать вклад прямого спонтанного излучения, появляется периодическая корреляция поля в моменты t и $t+T$, где T - полное время прохода излучения по резонатору. На спектральном языке наличие

Основы голографии

периодической корреляции означает формирование отдельных дискретных мод, составляющих эквидистантный набор частот $\omega_k = \omega_0 + k(2\pi/T)$. Важно. Что образование квазипериодической флуктуационной картины происходит еще до достижения лазером порога. Эти исходные флуктуации оказываются существенными для формирования окончательной временной картины излучения.

Коэффициент усиления активной среды на некоторой частоте пропорционален произведению разности заселенностей рабочих уровней на функцию $G(\omega_K)$, описывающую спектральный профиль линии люминесценции. Поэтому в процессе роста инверсии под действием накачки пороговое условие выполняется сначала для центральной частоты, затем постепенно входят в генерацию и другие частоты, причем процесс регенеративного усиления шумов непрерывно переходит в процесс генерации.

Таким образом, к моменту, когда начинается просветление фильтра. Излучение представляет собой квазипериодическую шумовую картину с шириной спектра $\Delta\omega_0$, много меньшей исходной ширины линии $\Delta\omega_{\text{люм}}$.

В процессе просветления фильтра происходит преобразование профиля поля, установившегося к началу просветления. Преобразование поля состоит в преимущественном росте тех точек профиля, где интенсивность максимальна. В результате многократных проходов излучения через просветляющий фильтр каждый отдельный выброс интенсивности сокращается по длительности. Одновременно с этим происходит подчеркивание наиболее интенсивных выбросов: их рост происходит быстрее, чем рост менее интенсивных выбросов.

Сужение импульсов и выделение наиболее интенсивных импульсов происходит одновременно, в едином процессе взаимодействия поля с фильтром. Однако лишь второй эффект равнозначен фазировке мод исходного спектра. Именно этот эффект может приводить к формированию регулярной временной картины излучения из первоначально хаотической временной структуры. Эффект же сужения импульсов означает появление новых дополнительных спектральных компонент, а к фазировке исходных компонент прямого отношения не имеет.

Условия и характер фазового согласования зависят от оптической схемы и параметров лазера. Большое значение имеет положение фототропного затвора относительно зеркал резонатора, которое влияет на расстояние между максимумами импульсов и их интенсивностями. Размещение затвора в положениях $x = L/m$ (L - база резонатора; m - целое число; x - расстояние затвора от 100% - го зеркала) приводит к увеличению частоты следования импульсов. Расположение затвора вплотную к глухому зеркалу наиболее удобно с точки зрения практического использования, так как позволяет получить цуг импульсов с максимальной с максимальной пиковой мощностью и наибольшей скважностью.

Лекция № 8

Типы лазеров

Все известные типы лазеров разделяют на группы: твердотельные лазеры, полупроводниковые лазеры, газовые лазеры, эксимерные лазеры и лазеры на свободных электронах.

8.1. Твердотельные лазеры

Представляют собой лазер, в которой активной средой служат диэлектрические кристаллы или стекла, активированные ионами редкоземельных элементов или ионами группы железа. Примером данного типа служат лазеры на рубине ($Al_2O_3 + Cr^{3+}$), на стекле с примесью Nd^{3+} или на основе кристалла алюмоиттриевого граната с ионами Nd^{3+} , которые рассмотрены выше.

На рис. 8.1 приведена типичная функциональная схема твердотельного лазера. Часто в резонатор встраивают устройство управления, позволяющее формировать лазерное излучение с нужными характеристиками. КПД таких лазеров составляет несколько процентов. Пороговая мощность излучения достигает несколько десятков джоулей в импульсном и порядка ста милливатт в непрерывном режимах.

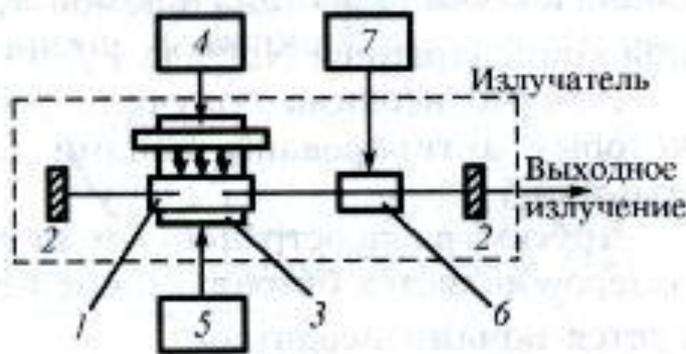


Рис. 8.1. Функциональная схема твердотельного лазера: 1 – активный элемент; 2 – зеркала резонатора; 3 – осветитель; 4 – блок питания накачки; 5 – блок охлаждения; 6 – устройство управления; 7 – блок управления

Обратную связь в оптическом резонаторе можно включать и выключать с помощью устройства управления. Обычно используют электрооптический затвор, например в виде ячейки Керра. Обратная связь включается на время 10^{-8} с. Это время и определяет длительность импульса. Таким образом, накопленная на метастабильном уровне энергия возбужденных ионов излучает энергию за очень короткий промежуток времени

10^{-8} с, что позволяет развить гигантскую мощность. Подобные конструкции получили название лазеров с модулированной добротностью. Гигантские импульсы можно увеличить по мощности, используя каскад лазерных усилителей.

8.2. Четырехуровневые лазеры

Среди твердотельных лазеров, работающих по четырехуровневой схеме, широкое распространение получили лазеры, у которых в качестве матрицы используется стекло и кристалл АИГ с добавками неодима.

Спектр атомов редкоземельных элементов, включенных в монокристалл, очень близок к спектру изолированных атомов. Это объясняется тем, что незаполненная оболочка 4f у редкоземельных элементов расположена ближе к ядру, чем оболочка 3d элементов группы железа, и в значительной степени экранируется от электрического поля решетки. Поэтому положение энергетических уровней редкоземельных ионов слабо зависит от типа кристалла, и для всех них характерно наличие узких линий флюоресценции.

Упрощенная схема энергетических уровней иона неодима в кристалле граната представлена на рис. 8.2. Спектр поглощения иона неодима состоит из большого числа сравнительно узких полос. Однако основной энергетически вклад в поглощение света накачки дают пять полос, расположенных в области 520, 580, 740, 800 и 900 нм. Наиболее интенсивная люминисценция соответствует длине волны 1.06 мкм. Время перехода иона из состояний, соответствующих уровням накачки, на метастабильный уровень существенно меньше, чем время жизни метастабильного уровня. Поэтому можно считать, что существует не множество уровней накачки, а лишь один условный уровень. Время жизни на уровне накачки менее 10^{-8} с. Метастабильный уровень также состоит из большого числа подуровней, но, поскольку время теплового равновесия между ними значительно меньше времени основного перехода, его также можно считать одиночным. Время жизни метастабильного уровня лежит в пределах $10^{-3} \div 10^{-4}$ с. Энергетических уровней (условно соответствующих второму по схеме четырехуровневому лазеру) также несколько, но каждый из них можно считать единственным, полагая, что его населенность задается по закону Больцмана. Нижний лазерный уровень ${}^4I_{11/2}$ достаточно далеко отстоит от основного состояния ${}^4I_{9/2}$ и (вследствие быстрой безызлучательной релаксации) имеет малое время жизни, порядка $6 \cdot 10^{-8}$ с. Поэтому он практически не заселен при комнатной температуре, и это состояние сохраняется даже при очень больших уровнях накачки. Четырехуровневая схема возбуждения в сочетании с высоким коэффициентом усиления обеспечивает низкий порог генерации лазерных сред, активированных неодимом.

Основы голографии

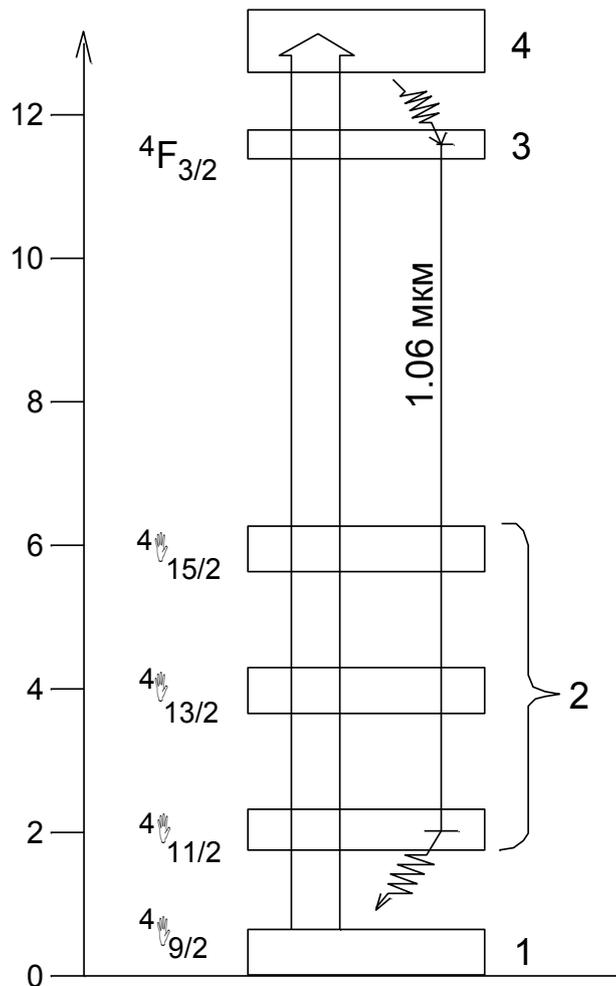


Рис. 8.2. Диаграмма энергетических уровней 1- 4 ионов неодима в кристалле граната

С учетом такого приближения возможно описание кинетики населенностей этих эффективных лазерных уровней.

Основными упрощающими приближениями (вытекающими из соотношения вероятностей излучательных переходов между уровнями) являются следующие:

1. Все ионы, возбужденные под действием накачки. Быстро переходят безызлучательным образом на метастабильный уровень 3.
2. Вероятность безызлучательных переходов с метастабильного уровня вниз пренебрежимо мала по сравнению с излучательными.
3. Время безызлучательных переходов ионов между уровнями $4 \rightarrow 3$ и $2 \rightarrow 1$ пренебрежимо мало по сравнению с временем излучательного перехода между уровнями $3 \rightarrow 2$.

Таким образом мощность генерации для лазеров описывается формулой:

$$P_{\text{ген}} = I_s N h \nu_{\text{ген}} [\eta B u_{\text{нак}} (1 - \delta) - (P_{31} + P_{32}) \delta].$$

Основы голографии

Если выполняются условия

$$P_{12} \ll P_{21}; \quad \frac{g_2}{g_1} e^{-h\nu_{21}/kT} \ll 1,$$

то практически все частицы сосредотачиваются на уровнях 1 и 3.

Пороговую скорость накачки находят из формулы

$$\eta B u_{\text{нак}}^{\text{пор}} = (P_{31} + P_{32}) \frac{\delta}{1 - \delta}$$

при выполнении условия

$$\frac{g_2}{g_1} e^{-h\nu_{21}/(kT)} \ll \delta.$$

При $h\nu_{21} \gg kT$, небольших δ и прочих равных условиях пороговая скорость накачки четырехуровневого генератора значительно ниже, чем в трехуровневом генераторе.

Лазерные кристаллы АИГ- Nd^{3+} получают на основе чистого кристалла АИГ, в который в процессе его выращивания вводят в качестве примеси положительные ионы неодима. Чистый кристалл АИГ является бесцветным прозрачным кристаллом. По своим оптическим свойствам – изотропным. Показатель преломления $n = 1.82$. Кристалл оптически прозрачен в диапазоне 240-6000 нм, имеет высокую твердость (около 8.5 ед. по Моосу) и хорошие теплопроводящие свойства, которые на порядок выше, чем у неодимового стекла, и уступают лишь рубину.

Современная технология позволяет выращивать достаточно большие и совершенные кристаллические АЭ АИГ- Nd^{3+} длиной 50-250 мм и диаметром 4-15 мм, что позволяет получать в режиме свободной модуляции до десятков джоулей при КПД > 1%, а в непрерывном режиме до сотен ватт при КПД около 2.5-3 %.

Лазеры на АИГ: Nd^{3+} применяются в технологических установках для сварки, сверления, резки, скрайбирования, лазерного отжига и др.

Другим веществом, широко используемым в качестве матрицы АЭ, является стекло. Ионы активатора аморфной матрицы не внедряются в узлы решетки, а входят в стекло в качестве его компонента, что допускает большое введение активных центров (до 6%, что соответствует $N \approx 10^{21} \text{ см}^{-3}$). Структура энергетических уровней иона Nd^{3+} в стекле примерно такая же, как и в кристалле, но полосы поглощения и люминесценции шире.

Наибольшее время жизни метастабильного состояния и наибольший квантовый выход люминесценции имеют силикатные стекла. Неодимовое стекло в отличие от кристаллов в силу больших объемов и более высокой концентрации активатора хорошо накапливает энергию. Поэтому его наиболее целесообразно применять в качестве активной среды импульсных лазеров высокой энергии. Достигнуты значения импульсной энергии излучения в десятки килоджоулей при КПД до единиц процентов. Стекло является отличным оптическим материалом. Главным достоинством стекла является его высокая оптическая однородность.

Основы голографии

Градиент показателя преломления хорошего лазерного стекла составляет $dn/d\lambda = \pm(0,5 \div 2) \times 10^{-8} \text{ см}^{-1}$. По этой причине расходимость излучения лазеров на стекле меньше, чем у рубиновых и АИГ- Nd^{3+} лазеров.

К сожалению, стекло имеет и ряд недостатков. Оно не обладает высокой термостойкостью и хорошей проводимостью. Излучение накачки неравномерно нагревает АЭ, и это (в силу низкой теплопроводности) приводит к неоднородности температурного поля, что является причиной термоупругих искажений, вызывающих оптические искажения. Другим недостатком является фотохимическая неустойчивость стекла, приводящая к постепенному снижению характеристик генерации.

Существует и ряд факторов, ограничивающих мощность генерируемого излучения. Это процессы поглощения собственного излучения на микропримесях и эффекты самофокусировки.

Наибольшей лучевой стойкостью обладают неодимовые стекла, не содержащие примесей металлов (платины, железа), имеющих технологическое происхождение. В режиме свободной генерации порог разрушения составляет $10^3 \div 10^4 \text{ Дж/см}^2$. В режиме модулированной добротности лучшие стекла имеют пороговые плотности энергии излучения (разрушающей торцы АЭ) в несколько сотен джоулей на квадратный сантиметр. Для сравнения, рубины и гранаты в режиме модулированной добротности разрушаются при $10 \div 30 \text{ Дж/см}^2$.

В режиме модуляции добротности импульсные мощности лазеров на стекле могут достигать значений 10^{14} Вт при наносекундной длительности импульсов. В режиме синхронизации мод возможно получение импульсов длительностью менее 10^{-14} с .

Лекция № 9

Полупроводниковые лазеры

Полупроводниковые лазеры представляют собой лазер на основе полупроводниковой активной среды, в которой используются разрешённые квантовые переходы. Полупроводниковые слои, способные к генерации лазерного излучения, реализуются в разных формах: как гомоструктура; как гетероструктура; в виде квантовых барьеров (квантовых ям) и квантовых точек.

При поглощении фотона с энергией $h\nu \geq E_g$ происходит переход электрона из валентной зоны в зону проводимости. При этом в валентной зоне образуется дырка.

Одновременно с генерацией электронно-дырочных пар в полупроводнике происходит процесс спонтанного излучения и вынужденная рекомбинация. Для создания условий усиления света необходимо создать вырожденных полупроводник, в котором нарушено тепловое равновесие. Однако, в чистых полупроводниках создать состояние с одновременным вырождением электронов и дырок трудно. Обычно используют два полупроводника n- и p- типов, в каждом из которых электроны и дырки вырождены. Работа лазера основана на возникновении инверсии населённости в области p-n перехода при инжекции носителей заряда.

Энергетические зоны полупроводникового лазера приведены на рис.9.1 для изолированных (а) и контактируемых (в) полупроводников. В p-n-переходе может возникнуть условие инверсии населённости только при условии прямого напряжения. Через p-n-переход текут токи, состоящие из электронов и дырок. В тонком слое p-n-перехода они рекомбинируют, излучая фотоны. Встречные потоки электронов и дырок будут поддерживать в p-n-переходе концентрацию, достаточную для их вырождения.

Энергетические уровни непрерывных зон почти до уровня Ферми F_c и F_v заняты электронами и дырками соответственно. Положительные дырки «р» и подвижные электроны «е» формируются акцепторными и, соответственно, донорными атомами с концентрацией $n \sim 10^{18} \text{ см}^{-3}$.

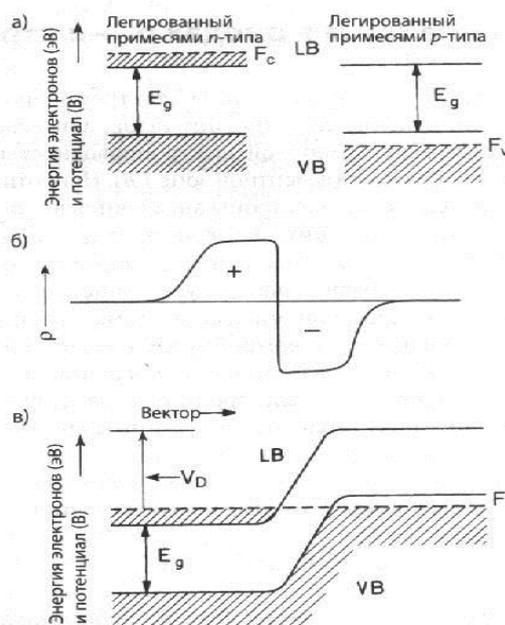


Рис.9.1. Энергетические диаграммы p-n-перехода для изолированных (а) и контактируемых (в) полупроводников, и объёмная плотность заряда (б).

При контакте проводников р- и n-типа получается диод – основной элемент полупроводникового диодного лазера.

Электроны до тех пор диффундируют в зону с дырочной электропроводностью, а дырки – в зону с электронной электропроводностью, пока образующийся объёмный заряд V_D не достигнут уровня совпадения энергий Ферми в проводниках р- и n-типа.

$$V_D = \frac{kT}{e} \ln \left(\frac{N_A N_D}{n_1^2} \right)$$

где N_A, N_D - плотность акцепторов и доноров в проводнике с дырочной и электронной проводимостью соответственно; n_1 - относительно низкая, термически генерируемая плотность электронов в нелегированном полупроводнике при температуре T ; k - постоянная Больцмана, e - элементарный заряд.

Пространственная зависимость потенциала V и плотность заряда ρ вытекают из уравнения Пуассона

$$-\frac{\partial^2 V}{\partial x^2} = \frac{\rho}{\epsilon_0}.$$

На рис.9.1 представлены:

а) Энергетические уровни полупроводников n- и р-типа с сильным легированием: E_g - ширина запрещённой зоны, F_c и F_v – энергии Ферми полупроводника n- и р-типа, L_V, V_D – зоны валентности и проводимости.

б) если n- и р-зоны находятся в контакте, то электроны диффундируют в зону с дырочной проводимостью, а дырки – в зону с электронной проводимостью.

В результате возникает объёмная плотность заряда ρ и происходит повышение энергии электронов в зоне р-типа с разностью потенциалов V_D , ограничивающей диффузию (F - полная энергия Ферми).

В результате дрейфа электронов в зону с дырочной электропроводностью и положительных дырок в валентной зоне – в зону с электронной электропроводностью образуется узкая зона с **инверсией населённости**. Толщина d определяется через постоянную диффузии D электронов в слое, легированном примесями р-типа, и временем рекомбинации τ . Например, для полупроводника GaAs (при $D = 10 \frac{cm^2}{c}$ и $\tau = 10^{-9} c$) получают толщину $d = 1$ мкм. В этой узкой полосе возможны усиление света и генерация лазерного излучения.

Чем выше электрическое поле в р-n переходе, тем больший ток через него протекает. Минимальный ток, при котором вынужденное излучение превышает поглощение, называется **пороговым**. На пороге генерации должно выполняться условие $kR \geq 1$, где k - коэффициент усиления на длине активной среды между зеркалами, R - коэффициент отражения зеркал резонатора. При токе выше порогового р-n переход является усиливающей средой. Если ввести положительную обратную связь в виде резонатора, то из усилителя света можно

Основы голографии

получить генератор. В качестве резонатора используют гладкие грани полупроводникового кристалла.

Базовая структура лазера с р-п переходом представлена на рис.9.2. Две боковые грани структуры скалываются или полируются перпендикулярно плоскости перехода. Две другие грани делаются шероховатыми для того, чтобы исключить излучение в направлениях, не совпадающих с главным. Такая структура называется резонатором Фабри-Перо.

Смещение лазерного диода в прямом направлении вызывает протекание тока. Вначале, при низких значениях тока, возникает спонтанное излучение, распространяющееся во всех направлениях. При увеличении смещения ток достигает порогового значения, при котором создаются условия для стимулированного излучения, и р-п переход испускает монохроматичный строго направленный луч света.

Принцип действия светодиода основан на излучательной рекомбинации инжектированных носителей в прямосмещенном р-п переходе, где L_p – диффузионная длина (~250 мкм).

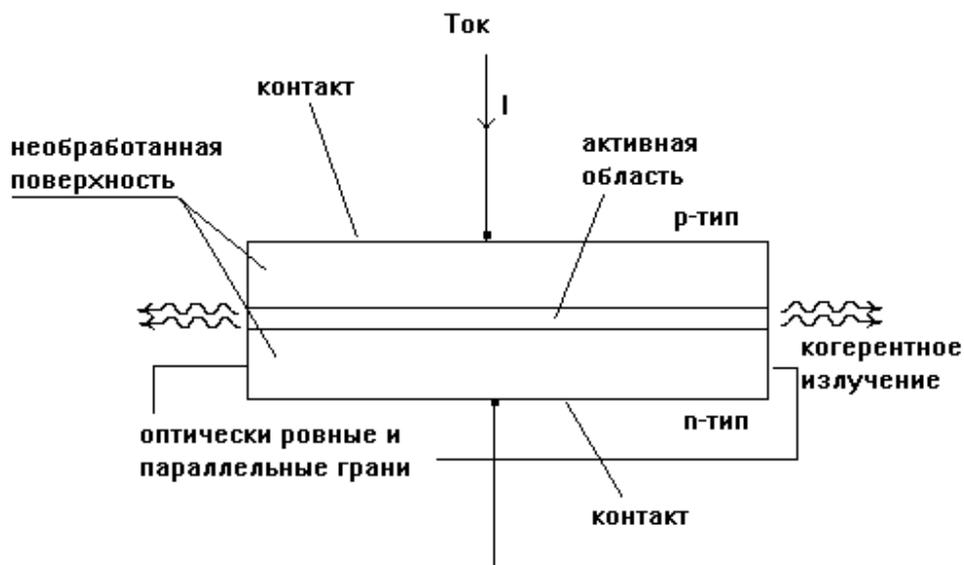


Рис.9.2. Структура лазера с р-п-переходом

Конструкции светодиодов представлены на рис.9.3. Среди светодиодных структур основной является структура с плоской геометрией. Обычно прямозонные светодиоды (красное излучение) формируются на подложках GaAs (а), тогда как непрямоzonные (оранжевое, жёлтое и зелёное излучения) - на подложках GaP (б). При использовании подложки GaAs на неё наращивается переходный слой $GaAs_{1-x}P_x$ переменного состава с x , изменяющимся в пределах $0 \div 0.4$, а затем слой $GaAs_{1-x}P_x$ с постоянным составом.

Основы голографии

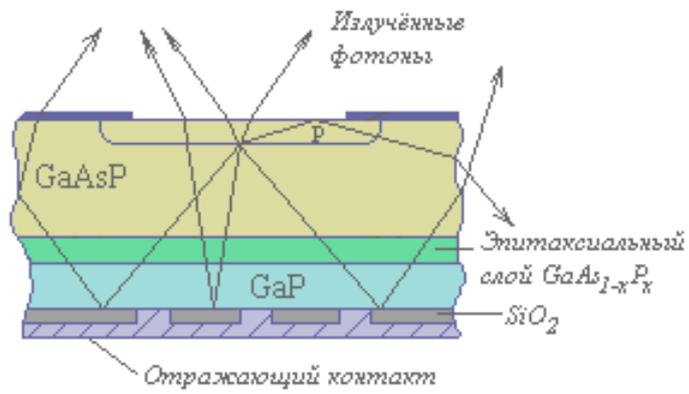


Рис.9.3. Конструкция светодиода на подложке GaP.

б

Лекция № 10

Лазеры на основе гетероструктур

Гетероструктурой называется полупроводниковая структура с несколькими гетеропереходами.

Гетеропереходом называется контакт двух различных по химическому составу полупроводников (типа А и В).

В 1963 году Ж.И. Алфёров и Г. Кремер сформулировали концепцию полупроводниковых лазеров на основе двойной гетероструктуры.

Уменьшение толщины слоя В приводит к появлению уровней размерного квантования:

E_c и E_v – края зоны проводимости и валентной зоны; E_e и E_h – уровни размерного квантования для электронов и дырок.

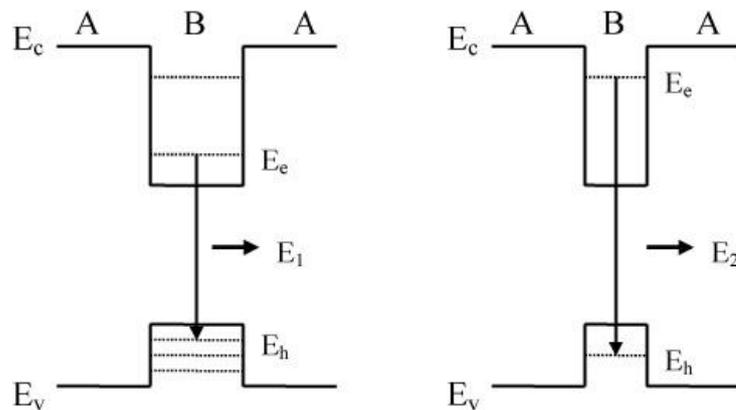


Рис. 10.1. Зонная диаграмма (энергетическая схема) двойной гетероструктуры

Энергия испускаемого при рекомбинации электрона и дырки фотона (E_1 и E_2) определяется не только ширинами запрещённых зон материалов, но и шириной слоя (потенциальной ямы). Поэтому $E_2 > E_1$.

Узкозонный полупроводник, помещённый между широкозонными, представляет собой потенциальную яму.

В соответствии с квантовой теорией энергия попавшей в эту яму частицы (например, электрона или дырки) квантуется, т.е. принимает набор дискретных значений в зависимости от ширины потенциальной ямы, высоты потенциального барьера и массы частицы. Это определяет и свойства излучения: длину волны, ширину линии, число линий.

В обычной гетероструктуре активный слой имеет толщину 300–500 нм, и в нём наблюдаются эффекты, присущие объёмному кристаллу.

Основы голографии

Формируя узкую потенциальную яму (тонкий, в один-два десятка нанометров, слой узкозонного полупроводника, сопоставимый с дебройлевской длиной волны электрона или дырки), обеспечивают довольно узкую линию излучения и низкий пороговый ток.

Действительно, число разрешённых состояний в активной области квантово-размерной структуры значительно меньше, чем в обычной, и при этом облегчается создание инверсной заселённости.

10.1. Низкоразмерные полупроводниковые гетероструктуры

Были выращены структуры с ультратонкими (десятые доли нанометра) слоями (а это означает, кроме всего прочего, ещё более узкую линию излучения). Как это ни удивительно, следы присутствия такого слоя можно увидеть... невооружённым глазом: в спектре излучения доминирует линия квантовой ямы, и, скажем, вместо прежнего зелёного света структура светит жёлтым или красным (в зависимости от толщины).

Созданы также *квантовые проволоки* (или *квантовые нити* – квазиодномерные структуры, в которых носители заряда могут свободно перемещаться только в одном направлении, вдоль нити) и, наконец, *квантовые точки* (или *квантовые ящики* – квазинуль-мерные структуры, т.е. трёхмерные потенциальные ямы, где квазичастицы локализованы во всех трёх направлениях).

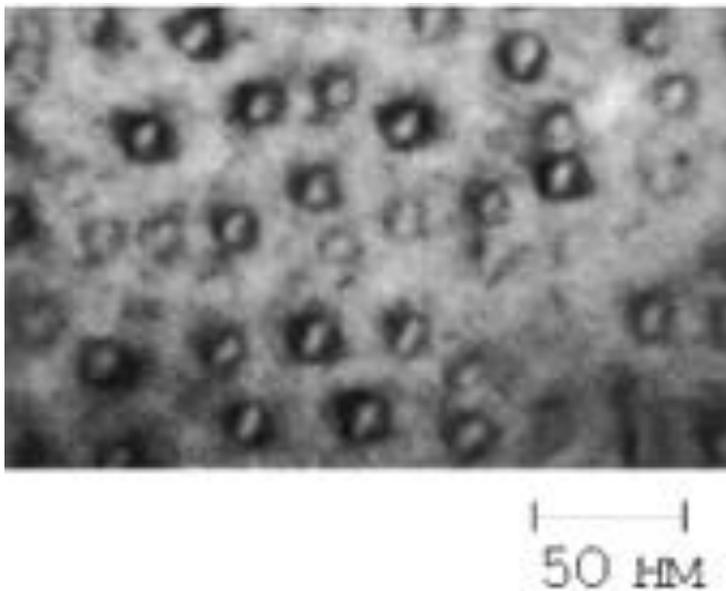


Рис. 10.2.

На фото, полученном с помощью просвечивающего электронного микроскопа, показаны квантовые точки InAs в матрице GaAs. Характерные размеры такого объекта – можно назвать его *точкой* (с точки зрения человека) или *ящиком* (с точки зрения электрона) – те же несколько нанометров. Полная – во всех направле-ниях – локализация приводит к тому, что энергетический спектр

подобной структуры по-настоящему дискретен, поэтому квантовые точки иногда называют *искусственными атомами*.

10.2. Фотодиоды

Фотогальванический эффект возникает при поглощении света полупроводником при одновременной генерации подвижных носителей – электронов и дырок.

Эти носители разделяются в пространстве. Причиной разделения носителей может быть электрическое поле приэлектродного барьера Шоттки на контакте металл-полупроводник, поле $p-n$ -перехода или гетероперехода.

Схема процессов, происходящих в $p-n$ -переходе, на который падает излучение с энергией фотонов $\hbar\omega > E_g$ (где E_g - ширина запрещённой зоны) приведена на рис.5.3.

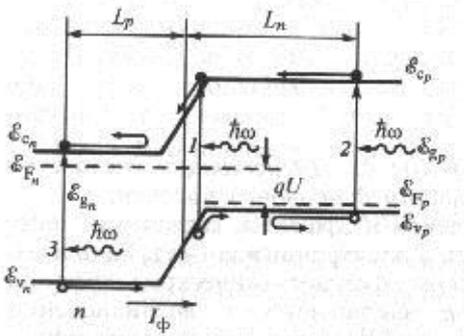


Рис. 10.3. Схема процессов, происходящих в $p-n$ -переходе под воздействием света

Основная часть носителей генерируется в областях, примыкающих к $p-n$ -переходу (области 2 и 3 на рис.10.3). Если генерация происходит на расстояниях меньше диффузионной длины электронов L_n и дырок L_p , то возникшие носители дойдут до $p-n$ -перехода, не рекомбинируя с основными носителями этих областей. В области объёмного заряда избыточные носители накапливаются в разных областях, что приводит к возникновению фотоЭДС и фототока при наличии внешней цепи.

Фототок через $p-n$ -переход протекает независимо от приложенного напряжения и определяется уравнением

$$I = I_s \exp\left(\frac{qU}{kT} - 1\right) - I_F, \quad (10.1)$$

где I_s - ток насыщения, создаваемый свободными носителями заряда; I_F - фототок, пропорциональный скорости генерации избыточных электронно-дырочных пар в области диффузионных длин неосновных носителей;

$qU = E_{Fn} - E_{Fp}$. Вольт-амперная характеристика фотодиода представлена на рис. 10.4.

Значения фотоЭДС можно определить, положив $I = 0$. Тогда

$$E_F = \frac{kT}{q} \ln\left(1 + \frac{I_F}{I_s}\right).$$

Основы голографии

Если к $P-n$ -переходу приложить обратное напряжение превышающее величину пробоя, то возможно усиление фотоЭДС за счёт лавинного умножения генерированных светом носителей. Лавинные фотодиоды перспективны при обнаружении слабых оптических сигналов.

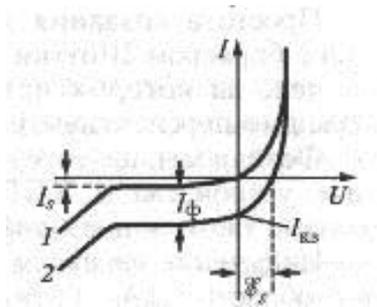


Рис.10.4. Вольт-амперная характеристика фотодиода:

I_s - ток насыщения; I_{Φ} - фототок; $I_{кз}$ - ток короткого замыкания; E_s - фотоЭДС; освещение есть (1), нет(2)

Фотодиоды Шоттки со структурой металл-полупроводник позволяют повысить быстродействие приёмников излучения до 10^{-10} с. В таких структурах граница спектральной характеристики сдвигается в сторону более длинных волн. Фотодиоды Шоттки имеют малое сопротивление базы фотодиода. Инерционность таких приборов определяется временем пролёта фотоносителей через область объёмного заряда ($10^{-10} - 10^{-11}$ с). Простота создания выпрямляющих фоточувствительных структур с барьером Шоттки на различного типа полупроводниках открывает большие перспективы их использования.

Большое распространение получили $p-i-n$ -диоды, в которых i -область состоит из слабо легированного полупроводника. Электрическое поле в i -области ускоряет транспорт носителей и снижает барьерную емкость фотодиода. Последнее обстоятельство позволяет снизить величину ёмкости на входе предусилителя и уменьшить время переходных процессов. Малые рабочие напряжения в фотодиодном режиме позволяют обеспечить совместимость $p-i-n$ -диодов с интегральными схемами.

Фотодиоды Шоттки со структурой металл-полупроводник позволяют повысить быстродействие приёмников излучения до 10^{-10} с. В таких структурах граница спектральной характеристики сдвигается в сторону более длинных волн. Фотодиоды Шоттки имеют малое сопротивление базы фотодиода. Инерционность таких приборов определяется временем пролёта фотоносителей через область объёмного заряда ($10^{-10} - 10^{-11}$ с). Простота создания выпрямляющих фоточувствительных структур с барьером Шоттки на различного типа полупроводниках открывает большие перспективы их использования.

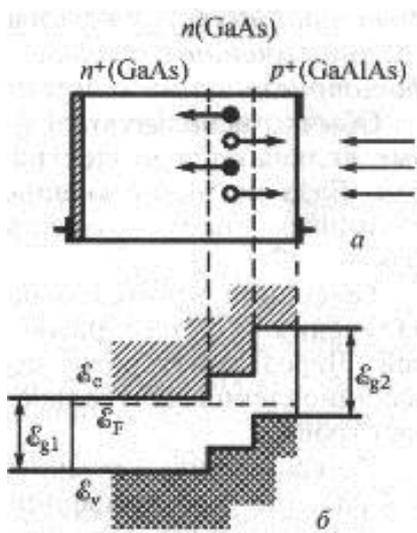


Рис.10.5. Схема фотодиода с гетероструктурой (а) и его зонная диаграмма (б)

Основы голографии

Фотодиоды на гетероструктурах позволяют создать фотоприёмные устройства с КПД, близким к 100%. На рис. 10.5 показано устройство и зонная диаграмма гетерофотодиода.

Излучение вводится через слой GaAlAs. Поглощение происходит в n -области GaAs. Разница в ширине запрещённой зоны по обеим сторонам от гетероперехода составляет около 0,4 эВ. Генерируемые в n -области дырки переносятся в p^+ -область. Ширина активной n -области выбирается такой, что происходит полное поглощение излучения. Структура работает при небольших напряжениях. Выбирая соответствующие пары полупроводников, можно создать фотодиоды для любой части оптического спектра. Отметим, что в гетероструктурах длина волны определяется разницей ширины запрещённых зон и не связана со спектральной характеристикой поглощаемого излучения.

Лекция № 11.

Физические основы формирования голограмм

В последние годы метод голографии находит все большее применение в науке и технике. Голограммы широко используются в оптическом приборостроении, в информационных системах, устройствах памяти и т.д. Развитие техники обуславливает прогресс в голографических технологиях и ведет к дальнейшему расширению областей практического применения метода.

Термин «голография» включает в себя два греческих слова, означающих «полный» и «отображаю» и был предложен венгерским ученым Д. Габором.

11.1. Основы формирования голограмм. Типы голограмм

Голография – это интерференционный метод регистрации световых волн, дифрагировавших на объекте, освещенном когерентным светом. При этом дифрагированные волны должны проинтерферировать с согласованной с ним по фазе опорной волной.

Если волны обладают достаточной степенью когерентности, то распределение разности их фаз в пространстве остается постоянным в течение времени, необходимого для регистрации голограммы. Возникает стационарная интерференционная картина с определенным распределением интенсивности. Поле, соответствующее этой картине, носит название голографического поля. Отображение этой картины на каком-либо носителе называется голограммой. Голограмма содержит информацию о фазе и амплитуде дифрагированных на объекте волн, благодаря чему обеспечивается возможность их точного восстановления при освещении голограммы восстанавливающей волной, подобной опорной волне.

Сохранение воспроизводимой информации о фазе является уникальной особенностью метода голографии. Отметим, что фотография позволяет сохранить только пространственное распределение интенсивности света, формирующее изображение объекта. Интенсивность же описывается квадратом модуля комплексной амплитуды световой волны и, следовательно, не содержит информацию о ее фазовом распределении. В методе голографии информация об амплитуде и фазе регистрируемой волны, идущей от объекта, кодируется с помощью опорной волны еще до регистрации голограммы.

Голографический метод применим ко всем волнам: световым, рентгеновским, акустическим и т.д. при условии наличия когерентных источников этих волн, пригодных для формирования соответствующих голографических полей. В настоящее время наибольшее распространение получила оптическая голография, что объясняется доступностью лазеров – источников когерентного излучения, и средств регистрации и наблюдения восстановленных изображений.

Ведутся работы по рентгеновской голографии, основанной на использовании ондуляторов – синхротронных источников когерентного рентгеновского излучения, и цифрового восстановления голограмм в виртуальном

Основы голографии

пространстве компьютера. Однако, широкого распространения эти методы не получили.

Пока акустическая и сейсмическая голография практически не развиваются, поскольку не могут конкурировать по информативности с методами компьютерной томографии.

**11.2. Восстановление волнового фронта.
Основное уравнение голографии.**

Формализуем, описанный выше принцип голографической записи волновых фронтов. Обратимся к рис. 11.1, на котором представлены два предмета M и N.

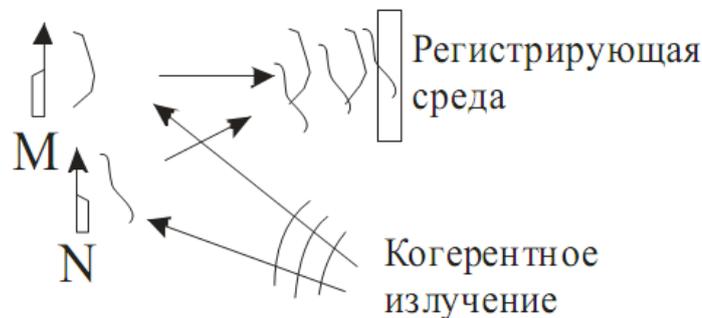


Рис. 11.1. Общая схема получения голограммы

Они освещаются когерентным светом от одного и того же источника. Отраженное от предметов излучение интерферирует в области пространства, где расположена регистрирующая среда. Поскольку время регистрации и восстановления голограммы значительно превышает период осцилляции используемого излучения, то при описании процессов получения голограммы и восстановления волнового фронта нас будут интересовать только комплексные амплитуды света непосредственно вблизи регистрирующей среды. Комплексную амплитуду света, падающего на пластинку от предмета M, можно записать в виде:

$$O = o \exp(i\varphi_o), \tag{11.1}$$

Где o и φ_o являются функциями пространственных координат плоскости регистрации голограммы, описывающими действительную амплитуду и фазу волны с комплексной амплитудой O . Аналогично этому комплексная амплитуда света, достигающего регистрирующей среды и отраженного от предмета N, записывается как:

$$R = r \exp(i\varphi_r). \tag{11.2}$$

Основы голографии

Здесь r и φ_r являются функциями пространственных координат плоскости регистрации голограммы, описывающими действительную амплитуду и фазу волны с комплексной амплитудой R .

Обозначим через O^* и R^* величины, комплексно-сопряженные комплексным амплитудам O и R . Тогда амплитуду голографического поля в плоскости регистрации голограммы – U , мы сможем описать с помощью следующего выражения:

$$U = O + R. \tag{11.3}$$

При этом интенсивность голографического поля будет описываться следующим образом

$$I = (O + R)(O + R)^*. \tag{11.4}$$

Раскрывая последнее выражение получим:

$$I = OO^* + RR^* + OR^* + RO^* = I_O + I_R + OR^* + O^* R, \tag{11.5}$$

где I_O и I_R - интенсивности опорной и объектной волн.

Далее положим, что используемая нами регистрирующая среда является линейным детектором интенсивности голографического поля. Используем для регистрации голограммы галогенидосеребряную фотоэмульсию. Ее использование будет означать, что экспозиция и обработка фотоматериала выбраны так, чтобы почернение фотоэмульсии (определяющее амплитудный коэффициент пропускания голограммы) было пропорционально интенсивности голографического поля, описываемой формулой (11.5). Осуществляемая при этом запись голограммы называется линейной. Из формулы (11.5) получим выражение, описывающее комплексную амплитуду поля A , формируемого в плоскости голограммы C :

$$A = kCI = k\{C(I_O + I_R) + COR^* + CO^* R\}. \tag{11.6}$$

Каждое из слагаемых в (11.6) описывает комплексную амплитуду одной из волн, формируемых голограммой при дифракции на ее структуре восстанавливающего излучения. Если в качестве восстанавливающего излучения использовать излучение, отраженное одним из указанных на рис. 11.1 объектов (например объектом R , т.е. положить, что $C=R$), то выражение (11.6) можно переписать в виде:

$$A = k\{R(I_O + I_R) + RR^* O + R^2 O^*\} = k\{R(I_O + I_R) + I_R O + R^2 O^*\}. \tag{11.7}$$

Проанализируем полученное выражение. Его *первое слагаемое* описывает комплексную амплитуду волны, распространяющейся в направлении распространения восстанавливающейся волны, т.е. волны отраженной от предмета N .

Второе слагаемое описывает волну, идущую в направлении распространения излучения, отраженного от предмета M , комплексная амплитуда которой пропорциональна амплитуде O . Эта волна формирует восстановленное с помощью голограммы мнимое изображение предмета M .

Основы голографии

Третье слагаемое описывает комплексную амплитуду волны, сходящейся в искаженное множителем R^2 действительное изображение предмета M .

Одну из формирующих голограммную структуру волн, обычно ту, которая потом используется в качестве восстанавливающей волны, принято называть опорной волной, а вторую – объектной волной.

Волну, описываемую первым, вторым и третьим слагаемыми (11.7), обычно называют нулевым, -1 и +1 порядками дифракции, соответственно.

Пространственное разделение этих порядков обычно обеспечивается соответствующим выбором углов падения опорной и объектной волн на плоскость регистрации голограммы.

Лекция № 12.

Геометрия формирования голограмм

Влияние геометрии получения голограмм на ее дифракционные свойства рассмотрим на примере сечения голографического поля, формируемого при интерференции когерентных волн, испускаемых точечными источниками S_1 S_2 . Сечение голографического поля представлено на рис. 12.1.

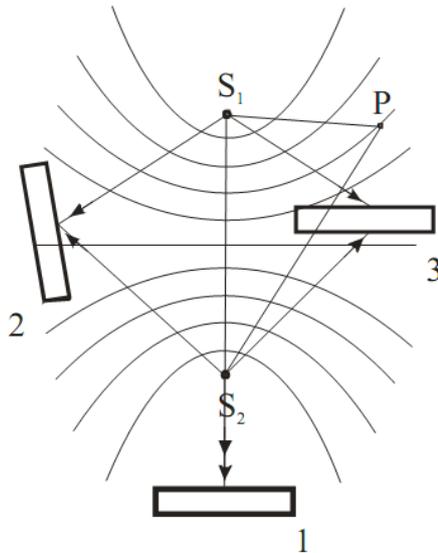


Рис. 12.1. Поперечное сечение семейства пучностей в голографическом поле

В этом поперечном сечении, которое содержит оба точечных источника, следами поверхностей максимумов интенсивности голографического поля (интерференционной картины) являются параболы.

В трехмерном пространстве это семейство параболоидов вращения, которые можно получить, поворачивая плоскость рисунка вокруг оси симметрии, в качестве которой выступает прямая, соединяющая оба точечных источника. Указанные параболоиды в случае синфазности волн, испускаемых обоими источниками, описываются следующим геометрическим выражением:

$$S_1P - S_2P = m\lambda, \quad (12.1)$$

Где P – некоторая точка голографического поля, характеризующаяся максимальной интенсивностью; λ – длина волны излучения, формирующего голографическое поле; $m = 0, 1, 2, \dots$. Цифрами 1, 2, 3 на рисунке обозначены некоторые характерные положения фотопластинки, либо другой регистрирующей среды, используемой для записи голограммы.

Д. Габор, не имеющий в своем распоряжении лазера и вынужденный максимально использовать свет от источников с низкой степенью когерентности, помещал пластинку в положение 1.

12.1. Схема регистрации голограмм Габора

В положении 1 (рис.12.1) главные лучи объектного и опорного пучков распространяются по одному направлению. Полученные таким образом голограммы называются осевыми голограммами или голограммами Габора. При их записи разность хода объектной и опорной волн в пределах поверхности пластинки минимальна по сравнению со всеми другими возможными положениями. Это позволяет использовать ее для формирования голографического поля источниками излучения с низкой степенью когерентности. Относительно большое расстояние между соседними поверхностями максимумов снижает требования к разрешающей способности регистрирующей среды.

Принципиальная схема записи голограмм Габора представлена на рис. 12.2. Здесь S – источник когерентного излучения, T – транспарант с изображением объекта, H – голограмма.

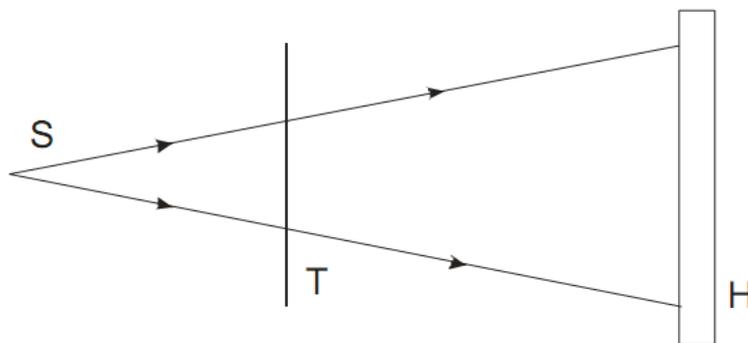


Рис. 12.2. Принципиальная схема записи голограмм Габора

В соответствии с приведенной схемой общую комплексную амплитуду U света, падающего на светочувствительную среду, в плоскости регистрации голограммы можно представить в виде суммы комплексной амплитуды недифрагировавшей на структуре объекта фоновой или опорной волны R и комплексной амплитуды волны O , дифрагировавшей на объекте.

$$U = R + O , \quad (12.2)$$

$$\text{где: } U = u_0 \exp^{i\varphi_u} ;$$

$$R = r_0 \exp^{i\varphi_r} ;$$

$$A = o_0 \exp^{i\varphi_o} .$$

Отсюда интенсивность излучения I в плоскости регистрации голограммы может быть описана следующим образом:

$$I = UU^* = (R+O)(R+O)^* = r_0^2 + o_0^2 + r_0 o_0 \exp[i(\varphi_r - \varphi_o)] + r_0 o_0 \exp[i(\varphi_o - \varphi_r)] . \quad (12.3)$$

При линейной обработке и ее восстановлении опорной волной с комплексной амплитудой R амплитуда поля в плоскости голограммы, непосредственно за ней – A , может быть описана с точностью до коэффициента пропорциональности следующим образом:

$$A = RI = r_0^2 (R + O) + r_0^2 \frac{O_0^2}{r_0} \exp(i\varphi_r) + r_0^2 \exp(i2\varphi_r) O^* =$$

$$r_0^2 \left\{ (R + O) + \frac{O_0^2}{r_0} \exp(i\varphi_r) + \exp(i2\varphi_r) O^* \right\},$$

(12.4)

где $RR^* = r_0^2$ и $R = r_0 \exp(i\varphi_r)$.

Если амплитуда опорной волны одинакова по всей плоскости голограммы, то первый член правой части выражения (12.4) описывает волновой фронт, комплексная амплитуда которого пропорциональна амплитуде исходной волны U в выражении (12.2).

Далее, если амплитуда опорной волны настолько велика, что $O_0^2/r_0 \ll 1$, то вторым слагаемым выражения (12.4) можно пренебречь. Наконец для опорной волны, фаза которой практически не изменяется на всей плоскости голограммы, третий член пропорционален величине, сопряженной комплексной амплитуде объектной волны. Он создает второе, сопряженное изображение объекта. В случае использования приведенной выше схемы сопряженное изображение является действительным.

Таким образом, при разглядывании восстанавливающей голограмму источника сквозь голограмму будут видны этот источник, мнимое изображение объекта и сопряженное действительное изображение объекта.

Если наблюдатель сфокусирует глаз на мнимом изображении, то действительное изображение окажется расфокусированным. Наоборот, если поместить экран в той плоскости, где дифрагированные на структуре голограммы лучи создают действительное изображение, на нем будет присутствовать паразитная засветка от расфокусированного мнимого изображения объекта. Эти взаимные световые помехи от изображения двойника в направлении наблюдения являются основным, наиболее значимым недостатком голограмм, получаемых по осевой схеме Д.Габор.

12.2. Устранение проблемы второго изображения

Вновь обратимся к рис. 12.1. Если поместить регистрирующую среду в положение, обозначенное цифрой 2, реализуется схема записи голограмм, предложенная Е. Лейтом. Полученные таким образом голограммы называются внеосевыми голограммами. Они характеризуются более высокими, по сравнению с осевыми голограммами, требованиями к разрешающей способности регистрирующих сред и степени когерентности излучения, используемого для

Основы голографии

формирования голографического поля. Внеосевое расположение регистрирующей среды требует использования лазерных источников излучения для формирования голографического поля и позволяет пространственно разделить порядки дифракции голограммы, т.е. позволяет избежать преналожения изображений, восстанавливаемых в +1 и -1 порядках дифракции.

Преимущества внеосевых схем дали возможность применять голографию для любых прозрачных или отражающих объектов, например, полутонковых транспорантов. При использовании внеосевого метода, в отличие от осевой голографии, интенсивность опорного пучка не обязательно должна быть намного больше интенсивности объектной волны. Однако интенсивность опорного пучка все же должна быть достаточно большой по сравнению с амплитудой его модуляции, обусловленной объектной волной. Это необходимо, чтобы результирующая экспозиция на выходила за пределы линейного участка характеристической кривой регистрирующей среды, т.е. чтобы запись голографического поля оставалась линейной.

12.3. Схемы записи объемных голограмм. Схема Ю.Н. Денисюка

Схема Лейта (положение 2 на рис. 12.1) позволяет в зависимости от толщины регистрирующей среды и ширины спектра пространственных частот регистрировать как тонкие, так и объемные голограммы.

Если толщина регистрирующей среды мала по сравнению с расстоянием между поверхностями максимумов интенсивности регистрируемого голографического поля, то голограмма, полученная в этом положении, действует как двумерная дифракционная решетка. При этом голограмма, по существу, представляет собой систему полос на поверхности регистрирующей среды. То же самое справедливо и для габоровских осевых голограмм.

Голограммы, получаемые на относительно толстых (по сравнению с периодом следования максимумов интенсивности голографического поля) регистрирующих средах представляют собой объемную дифракционную решетку, состоящую из последовательности частично отражающих поверхностей. Такая решетка обладает селективным (зависящем от угла падения и длины волны возмущающей волны) откликом, описываемым законом Брэгга. Голограммы, обладающие такими свойствами, называются объемными или брэгговскими.

Если толстая по сравнению с периодом следования максимумов голографического поля регистрирующая среда устанавливается в положение 3, то опорная и объектная сферические волны падают на нее с разных сторон. В этом случае расстояние между поверхностями максимумов интенсивности голографического поля составляет примерно половину длины волны регистрирующего излучения и эти поверхности близки к плоскостям, параллельным поверхности регистрирующей среды. Такая схема регистрации голограммы предложена Ю.Н. Денисюком. При регистрации голограммы в такой схеме в объеме регистрирующей среды образуется большое количество частично

Основы голографии

отражающих излучение поверхностей, называемых стратами, действующих подобно отражательному интерференционному фильтру.

Даже для толщин регистрирующей среды 10-12 мкм число этих страт может быть более 50. Большое число содержащихся в голограмме частично отражающих поверхностей обуславливает их высокую спектральную селективность, позволяющую восстанавливать записанное на них изображение в белом свете. Такие голограммы называются голограммами Ю.Н. Денисюка или отражательными объемными голограммами.

Лекция № 13.

Получение голограмм с падающим под углом опорным пучком

13.1. Понятие о спектре пространственных частот

Введем понятие плоской волны. Волна называется плоской, если ее амплитуда и фаза в любой момент времени постоянны по всей плоскости, уравнение которой имеет вид:

$$\vec{r} \cdot \vec{n} = const, \quad (13.1)$$

Где r – радиус-вектор точки в пространстве, а n – единичный вектор, нормальный к рассматриваемой плоскости.

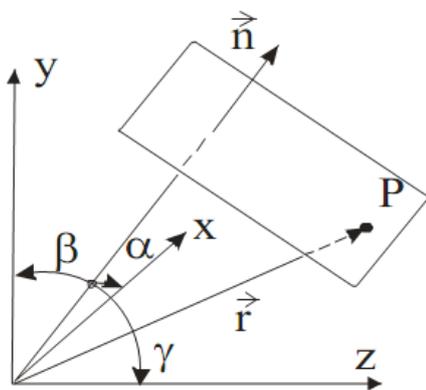


Рис. 13.1. Плоская волна в прямоугольной системе координат

Положим, что величина электрического поля V может быть описана следующим выражением

$$V(x, y, z, t) = a_1 \exp(-ik\vec{r} \cdot \vec{n}) \exp(i2\pi ft), \quad (13.2)$$

где a_1 – постоянная амплитуда волны, k – константа.

Тогда для конкретных значений $r = r_1$ и $t = t_1$ фаза волны будет равна $\varphi_1(r_1, t_1)$:

$$\varphi_1(r_1, t_1) = 2\pi ft_1 - k\vec{r}_1 \cdot \vec{n}. \quad (13.3)$$

В более поздний момент времени $t_2 > t_1$ тоже значение φ_1 фаза будет иметь на большем расстоянии от начала координат, в то время как на прежнем расстоянии она возрастет. Отсюда следует, что плоскости постоянных фаз перемещаются в пространстве. Следовательно, приведенное выражение (13.2) описывает плоскую волну. Направление вектора n , нормального к плоскости постоянной фазы, является направлением распространения волны. Если $\cos\alpha$, $\cos\beta$, $\cos\gamma$ – направляющие косинусы вектора n , то выражение (13.2) можно переписать в виде:

Основы голографии

$$V(x, y, z) = a_1 \exp[-ik(x \cos \alpha + y \cos \beta + z \cos \gamma)] \exp(i2\pi ft), \quad (13.4)$$

где x, y, z – компоненты вектора r в декартовых координатах.

Подстановка (13.4) в волновое уравнение

$$\nabla^2 V(x, y, z, t) = \frac{1}{c^2} \frac{\partial^2 V(x, y, z, t)}{\partial t^2}, \quad (13.5)$$

дает:

$$-k^2 (\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma) = -\frac{4\pi^2 f^2}{c^2} = -\frac{4\pi^2}{\lambda^2}, \quad (13.6)$$

где λ – длина волны излучения. Поскольку направляющие косинусы удовлетворяют соотношению:

$$\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1, \quad (13.7)$$

то (13.4) является решением волнового уравнения при условии

$$k = \frac{2\pi}{\lambda}. \quad (13.8)$$

Величина k называется волновым числом.

С учетом этого соотношение (13.4) мы можем переписать в виде:

$$\begin{aligned} V(x, y, z, t) &= a_1 \exp\left[-2\pi i \left(x \frac{\cos \alpha}{\lambda} + y \frac{\cos \beta}{\lambda} + z \frac{\cos \gamma}{\lambda}\right)\right] \exp(i2\pi ft) = \\ &= a_1 \exp[-2\pi i(\xi x + \eta y + \zeta z)] \exp(i2\pi ft) = A(x, y, z) \exp(i2\pi ft). \end{aligned} \quad (13.9)$$

В данном курсе мы будем рассматривать лишь монохроматическое излучение. Тогда множитель $\exp(i2\pi ft)$ можно опустить и для описания электромагнитной волны пользоваться только ее комплексной амплитудой $A(x, y, z)$. Величины ξ, η и ζ определяемые равенствами

$$\xi = \frac{\cos \alpha}{\lambda}; \quad \eta = \frac{\cos \beta}{\lambda}; \quad \zeta = \frac{\cos \gamma}{\lambda} \quad (13.10)$$

Называются пространственными частотами. Они обратны пространственным периодам волны, измеренным по осям x, y, z , соответственно. Пространственные частоты измеряются в обратных миллиметрах (мм^{-1}).

Пространственные частоты могут принимать как положительные, так и отрицательные значения. Воспользуемся следующим правилом знаков для пространственных частот и фаз волн. Если направление распространения волны составляет с соответствующей осью угол меньше 90 градусов, то пространственная частоты положительна, если больше 90 градусов, то она отрицательна.

Основы голографии

Если ориентировать систему координат так, чтобы, например, ось z совпала с направлением распространения волны ($\xi = \eta = 0, \zeta = 1/\lambda$), то легко видеть, что в (13.9) фаза волны в фиксированный момент времени уменьшается с увеличением расстояния от источника.

Величины ξ, η и ζ являются независимыми, их связь можно получить из (13.7). при подстановке (13.10) в (13.7) получим:

$$\lambda^2 \xi^2 + \lambda^2 \eta^2 + \lambda^2 \zeta^2 = 1, \tag{13.11}$$

Или

$$\zeta = \pm \frac{1}{\lambda} (1 - \lambda^2 \xi^2 - \lambda^2 \eta^2)^{1/2}, \tag{13.12}$$

где знак определяется направлением распространения волны в соответствии с ранее принятым правилом знаков. Теперь мы можем записать комплексную амплитуду $A(x, y, z)$ плоской волны в виде:

$$\begin{aligned} A(x, y, z) &= a_1 \exp \left[-2\pi i \left(x \frac{\cos \alpha}{\lambda} + y \frac{\cos \beta}{\lambda} + z \frac{\cos \gamma}{\lambda} \right) \right] = \\ &= a_1 \exp \left[-2\pi i (\xi x + \eta y) \right] \exp \left[- \left(\frac{2\pi i}{\lambda} \right) z (1 - \lambda^2 \xi^2 - \lambda^2 \eta^2)^{1/2} \right] = \\ &= A(x, y, 0) \exp \left[-i \left(\frac{2\pi}{\lambda} \right) z (1 - \lambda^2 \xi^2 - \lambda^2 \eta^2)^{1/2} \right]. \end{aligned} \tag{13.13}$$

Из (13.13) видно, что величина комплексной амплитуды плоской волны на произвольном расстоянии z равна произведению комплексной амплитуды волны при $z = 0$ и экспоненты, убывающей при увеличении z .

Рассмотрим теперь поведение световой волны, встречающей на своем пути препятствие. Чтобы получить точное решение задачи дифракции, необходимо решить волновое уравнение при граничных условиях, соответствующих выбранному препятствию. Такой прямой подход, к сожалению, позволяет получить решение задачи дифракции лишь для препятствий простой формы. Представляющие практический интерес задачи дифракции обычно решают приближенными методами. Рассмотрим один из них.

13.2. Дифракция на периодических структурах

Рассмотрим плоскую волну с амплитудой a_1 , распространяющуюся в направлении положительной полуоси z и падающую на прозрачный предмет (транспарант), находящийся в плоскости $z = 0$. Пусть транспарант имеет пропускание t , описываемое периодической функцией

$$t(x, y) = t_0 + t_1 \cos 2\pi \eta y. \tag{13.14}$$

Основы голографии

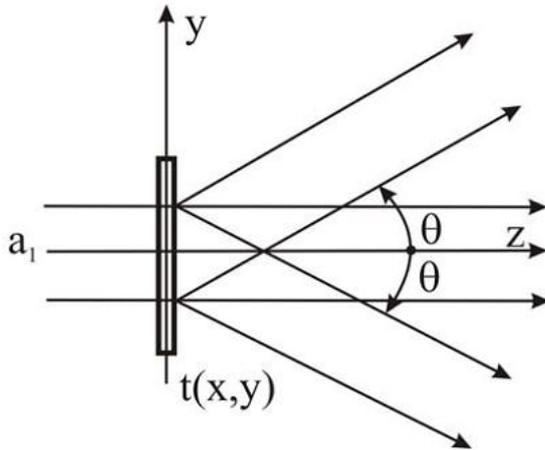


Рис. 13.2. Дифракция плоской волны на прозрачном материале с периодической структурой

Здесь t_0 и t_1 вещественные постоянные, т.е. предполагается, что транспарант чисто амплитудный и вносит фазового сдвига в проходящую через него волну; η – пространственная частота амплитудного коэффициента пропускания транспаранта.

Понятие пространственной частоты применимо к синусоидальной составляющей пространственного распределения любой физической величины, а не только к распределению комплексных амплитуд света.

Непосредственно за транспарантом амплитуды волны $A(x,y,0)$ равна произведению амплитуды падающего света a_1 и пропускания t

$$\begin{aligned}
 A(x,y,0) &= a_1 t(x,y) = a_1 t_0 + a_1 t_1 \cos 2\pi \eta y = \\
 &= a_1 t_0 + \frac{1}{2} a_1 t_1 \exp(2\pi i \eta y) + \frac{1}{2} a_1 t_1 \exp(-2\pi i \eta y)
 \end{aligned}
 \tag{13.15}$$

Заметим, что *второе слагаемое* (13.15) имеет тот же вид, что и решение волнового уравнения (13.9), если в него подставить $\xi = 0$ и $\eta < 0$. Отсюда можно сделать вывод о том, что это второе слагаемое (13.15) описывает плоскую волну, которая распространяется параллельно плоскости yz и характеризуется отрицательной пространственной частотой. При этом обозначенный на рис. 13.2 угол θ , который составляет (определяет) направление распространения этой волны с осью z и определяется как $90 - \beta$, является отрицательным углом, поскольку для него должно выполняться равенство

$$\sin \theta_2 = \lambda \eta$$

Аналогично *третье слагаемое* (13.15) описывает плоскую волну, которая также распространяется параллельно плоскости yz с положительной пространственной частотой по оси y , образуя при этом с осью z положительный угол θ .

Первое слагаемое в (13.15) не зависит ни от x , ни от y и описывает плоскую волну, распространяющуюся в направлении оси z .

Итак, при падении плоской волны, распространяющейся вдоль оси z , на транспарант с синусоидальным в направлении y амплитудным пропусканием за транспарантом возникают три плоские волны:

Основы голографии

первая, с амплитудой $a_1 t_0$, распространяется вдоль оси z (недифрагированная волна, т.е. волна, не изменившая направление своего распространения);

вторая, с амплитудой $a_1 t_1/2$, распространяется в плоскости yz , образуя с осью z угол по модулю $\theta = \arcsin(\lambda\eta)$ (дифрагированная волна -1 порядка);

третья, с амплитудой $a_1 t_1/2$, распространяется в плоскости yz вверх от оси z , образуя с ней такой же по модулю угол θ (дифрагированная волна +1 порядка).

Мы рассмотрели один из важнейших случаев дифракции. Транспаранты с периодическим распределением амплитудного пропускания называются дифракционными решетками. В большинстве случаев голограмму можно рассматривать как транспарант с периодически модулированным амплитудным пропусканием. Поэтому можно ожидать, что голограмма будет воздействовать на падающий свет примерно так же, как обычная дифракционная решетка.

Определим комплексную амплитуду света в плоскости xy при $z = d$ для дифракции плоской волны на помещенном в плоскости $z = 0$ транспаранте с синусоидальным амплитудным пропусканием $t(x,y)$. За транспарантом возникают три плоские волны, комплексные амплитуды которых в плоскости $z = 0$ описываются выражением (13.15). С помощью ранее полученного выражения (13.13) определим комплексные амплитуды этих волн в плоскости $z = d$. Результирующая комплексная амплитуда при $z = d$ является их суммой и имеет вид:

$$\begin{aligned}
 A(x, y, d) = & a_1 t_0 \exp\left(-i \frac{2\pi d}{\lambda}\right) + \\
 & + \frac{1}{2} a_1 t_1 \exp(i2\pi\eta y) \exp\left[-i \frac{2\pi d}{\lambda} (1 - \lambda^2 \eta^2)^{1/2}\right] + \\
 & + \frac{1}{2} a_1 t_1 \exp(-i2\pi\eta y) \exp\left[-i \frac{2\pi d}{\lambda} (1 - \lambda^2 \eta^2)^{1/2}\right].
 \end{aligned}
 \tag{13.16}$$

Поскольку зависящие от z показатели экспонент в (13.16) являются мнимыми, каждое из трех слагаемых этого выражения описывает распространяющуюся волну. Однако для некоторых длин волн λ показатели экспонент становятся вещественными. При $\lambda\eta \rightarrow 1$ угол дифракции $\theta = \arcsin(\lambda\eta)$ увеличивается, приближаясь к 90 градусам. Для больших значений длин волн, удовлетворяющих неравенству $\lambda^2 \eta^2 > 1$, выражение $(1 - \lambda^2 \eta^2)^{1/2}$ становится мнимым. В этом случае второе и третье слагаемые (13.16), соответствующие первым порядкам дифракции, будут описывать поверхностные волны, т.е. волны, распространяющиеся вдоль поверхности транспаранта и затухающие по экспоненте с увеличением расстояния от нее. Если приведенное выше условие записать в виде $\lambda > 1/\eta$, то видно, что поверхностные волны возникают при падении на решетку излучения, длина волны которого больше периода решетки, равного $1/\eta$. Если записать условие существования затухающих волн в виде $\eta > 1/\lambda$, то будет видно, что в распределении поля на расстоянии $d \gg \lambda$ от транспаранта не содержится никакой информации о его пространственных частотах, превышающих $1/\lambda$.

13.3. Постановка общей задачи о дифракции

Если функция амплитудного пропускания $t(x,y)$ не синусоидальная, но периодическая, то ее можно привести к сумме синусоидальных функций, разложив в ряд Фурье:

$$t(x,y) = \sum_l \sum_k t_{lk} \exp(-i2\pi\xi_l x) \exp(-i2\pi\eta_k y). \quad (13.17)$$

Если же $t(x,y)$ непериодическая, то ряд Фурье нужно заменить на интеграл Фурье.

При этом коэффициенты t_{lk} заменяются произведением

$T(\xi, \eta)d\xi d\eta$, где $T(\xi, \eta)$ - Фурье образ функции $t(x,y)$, определяемый с помощью следующего выражения:

$$T(\xi, \eta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t(x,y) \exp(2\pi i \xi x) \exp(2\pi i \eta y) dx dy. \quad (13.18)$$

Тогда выражение, описывающее комплексную амплитуду поля за транспарантом, на расстоянии d от него, имеет вид:

$$A_2(x,y,d) = a_1 \iint \left[T(\xi, \eta) \exp\left(-i \frac{2\pi d}{\lambda} (1 - \lambda^2 \xi^2 - \lambda^2 \eta^2)^{1/2}\right) \right] \times \exp(-i2\pi\xi x) \exp(-i2\pi\eta y) d\xi d\eta, \quad (13.19)$$

где интегрирование производится по всем ξ и η , удовлетворяющим неравенству $(\xi^2 + \eta^2) \leq 1/\lambda^2$.

Анализ уравнения (13.19) приводит к следующему заключению: если плоская волна с амплитудой a_1 , распространяется в направлении оси z , падает на помещенный в плоскости $z = 0$ транспарант с амплитудным пропусканием $t(x,y)$, то спектр $A_2(\xi, \eta)$ комплексной амплитуды волны в плоскости $z = d$ имеет вид

$$A_2(\xi, \eta)|_{z=d} = a_1 T(\xi, \eta) \exp\left[-i \frac{2\pi d}{\lambda} (1 - \lambda^2 \xi^2 - \lambda^2 \eta^2)^{1/2}\right]. \quad (13.20)$$

Если ограничиться рассмотрением низкочастотных составляющих спектра, т.е. допустить, что $\xi, \eta \ll 1/\lambda$, то квадратный корень в (13.20) можно записать в виде

$$(1 - \lambda^2 \xi^2 - \lambda^2 \eta^2)^{1/2} \approx 1 - \frac{1}{2} \lambda^2 \xi^2 - \frac{1}{2} \lambda^2 \eta^2, \quad (13.21)$$

а (13.20) заменить приближенным выражением

$$A_2(\xi, \eta)|_{z=d} \approx a_1 T(\xi, \eta) \exp\left[i\pi\lambda d (\xi^2 + \eta^2)\right]. \quad (13.22)$$

Отметим, что фазовый множитель $\exp(-i2\pi d/\lambda)$ в (13.22) опущен, поскольку он не зависит от пространственных частот. Отбрасывание фазового множителя, постоянного по всей плоскости, эквивалентно сдвигу начала отсчета времени.

Основы голографии

Понятие пространственных частот и решение простейших задач дифракции подготовило нас к детальному анализу основной схемы получения голограмм – схемы получения голограмм с наклонным опорным пучком.

В начале рассмотрим случай недиффузного объектного пучка.

Лекция № 14.

Получение голограмм с падающим под углом опорным пучком (продолжение)

14.1. Получение голограмм с наклонным опорным пучком при недиффузном объектном пучке

Получение голограммы с помощью опорной волны, падающей на плоскость регистрирующей среды под углом, отличным от угла падения объектной волны – есть эффективный метод пространственного разделения восстановленных изображений. Пространственно-частотный анализ этого метода основан на принятии несущей (опорной) волны, пространственная частота которой модулируется информацией о предмете. Таким образом, выражение «голограмма с несущей частотой» эквивалентно выражению «внеосевая голограмма». При использовании метода несущей частоты отпадает необходимость получения опорной волны за счет света, прошедшего через предмет. Поэтому, при применении внеосевых голограмм, в противоположность габоровским голограммам, нет необходимости ограничиваться транспарантами с большими прозрачными участками. На рис. 14.1 показан простой способ деления волнового фронта, позволяющий освещать прозрачный транспарант когерентной плоской волной и получать наклонную плоскую волну от того же источника.

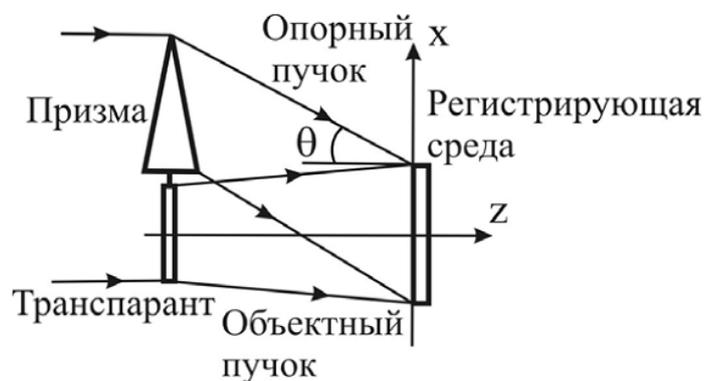


Рис. 14.1. Схема получения голограмм с внеосевым опорным пучком

В качестве объекта можно взять полутонный транспарант. Пусть $O((x,y))$ – комплексная амплитуда объектной волны в плоскости голограммы, $R = r \exp(2\pi i \xi_r x) = r \exp(ikx \sin\theta)$ – комплексная амплитуда плоской опорной волны. Из сравнения этих выражений проводимого с учетом правила знаков для фазы, гласящего, что фаза волны обратно пропорциональна пройденному оптическому пути, получим выражение для пространственной

Основы голографии

частоты опорной волны, представленной на рис. 14.1. Пространственная частота опорной волны

$$\xi_{опорн} = -\xi_r = -(\sin\theta)/\lambda$$

соответствует волновому вектору опорной волны, направленному вниз от оси z, где θ – угол, образованный им в плоскости xz с осью z.

Пусть после записи интерференционной картины, образованной волновыми фронтами $O(x,y)$ и $R(x,y)$ мы получили голограмму с амплитудным пропусканием, пропорциональным интенсивности зарегистрированного голографического поля I , описываемого следующим выражением:

$$I = OO^* + RR^* + OR^* + O^*R = OO^* + r^2 + Or \exp(-2\pi i \xi_r x) + O^*r \exp(2\pi i \xi_r x), \quad (14.1)$$

т.е. ее амплитудное пропускание $t(x,y)$ равно:

$$t = t_0 - kI, \quad (14.2)$$

где t_0 – начальный уровень пропускания, k – коэффициент пропорциональности.

Если на стадии восстановления голограмма освещается исходной опорной волной, то для комплексной амплитуды поля сразу за голограммой мы можем написать:

$$A(x,y) = Rt = t_0 r \exp(2\pi i \xi_r x) - k \left[OO^* r \exp(2\pi i \xi_r x) + r^3 \exp(2\pi i \xi_r x) + Or^2 + O^* r^2 \exp(4\pi i \xi_r x) \right]. \quad (14.3)$$

Ранее мы утверждали (без доказательства), что при выборе соответствующего направления распространения опорной волны можно отделить нужную восстановленную волну от остальных волн, дифрагированных голограммой. Для доказательства этого и определения связи условия пространственного разделения дифрагированных волн с максимальной пространственной частотой пропускания объекта, проведем пространственно-частотный анализ выражения (14.3).

Пусть регистрируемый на голограмме транспарант имеет пропускание $S(x,y)$ и спектр $S(\xi,\eta)$, где $S(\xi,\eta)$ представляет собой Фурье-образ $S(x,y)$. Протяженность спектра $S(\xi,\eta)$ лежит в пределах от $-\xi_{\max}$ до ξ_{\max} и от $-\eta_{\max}$ до η_{\max} .

При освещении объекта распространяющейся вдоль оси z плоской волной комплексная амплитуда объектной волны, падающей на голограмму, равна $O(x,y)$. Соответствующий этой функции спектр определяется выражением:

$$O(\xi,\eta) = o_1 S(\xi,\eta) \exp \left[-i \frac{2\pi d}{\lambda} (1 - \lambda^2 \xi^2 - \lambda^2 \eta^2)^{1/2} \right], \quad (14.4)$$

где o_1 – постоянная амплитуда плоской волны, падающей на транспарант, d – расстояние между транспарантом и голограммой.

Основы голографии

Отметим, что максимальная протяженность (ширина) спектра $O(\xi, \eta)$ в плоскости пространственных частот определяется интервалом, в котором функция $S(\xi, \eta)$ не равна 0. Выражение (14.3) содержит не только $O(x, y)$, но и комплексно-сопряженную ей величину $O^*(x, y)$ со спектром:

$$O'(\xi, \eta) = O^*(-\xi, -\eta) = o_1 S^*(-\xi, -\eta) \exp \left[+i \frac{2\pi d}{\lambda} (1 - \lambda^2 \xi^2 - \lambda^2 \eta^2)^{1/2} \right]. \quad (14.5)$$

Теперь с помощью (14.4) и (14.5) найдем абсолютное значение спектра функции $A(x, y)$, определяемой выражением (14.3). Первый член в правой части (14.3), $t_0 \exp(2\pi i \xi_r x)$, описывает плоскую волну, представляющую собой недифрагированный свет, распространяющийся в направлении восстанавливающей волны. Ее спектр представляет собой δ -функцию, располагаемую в точке спектральной плоскости с координатами $(-\xi_r, 0)$. Изобразим ее для наглядности большой вертикальной стрелкой на рис. 14.2.

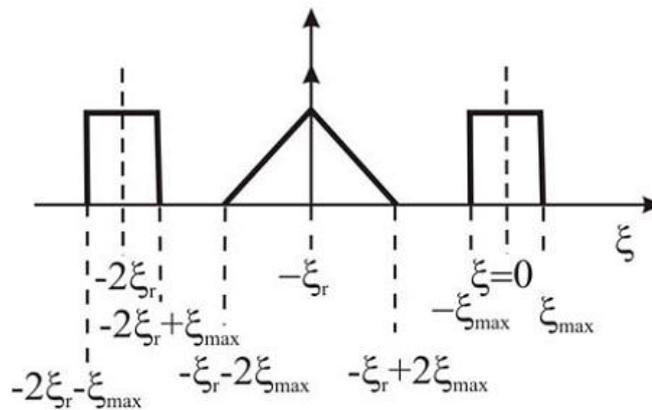


Рис. 14.2. Пространственный спектр восстановленного голограммой излучения

Второе слагаемое в (14.3) (первое в скобках) - $kOO^* \exp(2\pi i \xi_r x)$, преобразуется в частотном пространстве в функцию автокорреляции спектра O . Прежде всего надо знать максимальную протяженность функции автокорреляции в спектре пространственных частот! Эта величина является главным фактором, определяющим пространственную частоту, соответствующую требуемому углу падения опорной волны. Для ее нахождения напомним, что интеграл корреляции подобно интегралу свертки, не равен нулю лишь в пределах области перекрытия функций при сканировании одной из них с помощью другой. Интервал значений переменных, в котором интеграл не равен нулю, определяется суммой ширин обеих функций; в случае автокорреляции максимальная протяженность результирующей функции в два раза больше ширины функции, подвергаемой операции автокорреляции. Опущкая постоянные множители, получаем Фурье-образ второго слагаемого (14.3):

Основы голографии

$$\begin{aligned}
 OO^* \exp(2\pi i \xi_r x) &\supset O^*(\xi, \eta) * O(\xi + \xi_r, \eta) = \\
 &= \left[S^*(\xi, \eta) \exp\left(i \frac{2\pi d}{\lambda} (1 - \lambda^2 \xi^2 - \lambda^2 \eta^2)^{1/2} \right) \right] * \\
 &* \left[S(\xi + \xi_r, \eta) \exp\left(-i \frac{2\pi d}{\lambda} (1 - \lambda^2 (\xi + \xi_r)^2 - \lambda^2 \eta^2)^{1/2} \right) \right].
 \end{aligned} \tag{14.6}$$

Ранее указывалось, что ширина спектра $O(\xi, \eta)$ определяется его амплитудным распределением и, следовательно, шириной спектра $S(\xi, \eta)$.

Таким образом, функция автокорреляции в два раза шире функции $S(\xi, \eta)$: ее ширина составляет

$$2[\xi_{\max} - (-\xi_{\max})] = 4\xi_{\max} \text{ по оси } \xi$$

и $4\eta_{\max}$ по оси η . Ее центр соответствует центру функции $S(\xi + \xi_r, \eta)$ и совпадает с пространственной частотой, соответствующей углу падения восстанавливающей волны. Симметричное размытие дифрагированного света вокруг направления распространения восстанавливающей волны называют интермодуляцией, имея в виду, что оно обусловлено взаимной интерференцией точек объекта.

Второе слагаемое в скобках(14.3), $-kr^3 \exp(2\pi i \xi_r x)$, аналогично первому экспоненциальному слагаемому в правой части и преобразуется в δ -функцию в точке $(\xi = -\xi_r, 0)$.

Все рассмотренные три слагаемые выражения (14.3) называются членами нулевого порядка дифракции. Они описывают световые волны, не испытавшие отклонения, т.е. распространяющиеся за голограммой в том же направлении, что и падающая на нее волна.

Третий член в квадратных скобках в (14.3) пропорционален исходной волне $O(x, y)$, которая падала на плоскость регистрации голограммы при ее записи и называлась нами объектной. Абсолютная величина ее спектра $|O(\xi, \eta)|$ пропорциональна $|S(\xi, \eta)|$. Положим, что $|S(\xi, \eta)|$ описывает симметричное распределение вокруг центральной пространственной частоты

$\xi = 0, \eta = 0$ в интервале $\pm \xi_{\max}$ и $\pm \eta_{\max}$ и, следовательно, то же частотное распределение соответствует функции $O(x, y)$ в изображении спектра.

Последнее слагаемое в (14.3), $-kO^* r^2 \exp(4\pi i \xi_r x)$, описывает волновой фронт, несущий информацию о комплексно сопряженной комплексной амплитуде объектной волны в плоскости записи и восстановления голограммы. Его пространственный спектр определяется соотношением:

$$-kO^* r^2 \exp(4\pi i \xi_r x) \supset -kr^2 O^* [-(\xi + 2\xi_r), -\eta] \cdot \tag{14.7}$$

Абсолютная величина этого спектра пропорциональна

$$|S^* [-(\xi + 2\xi_r), -\eta]|,$$

Основы голографии

т.е. это частотное распределение подобно распределению спектра объектной волны, но является его зеркальным отражением и смещено по оси ξ на величину ξ_r , а граничные значения частот равны

$$\xi = -2 \xi_r \pm \xi_{\max} \text{ И } \pm \eta_{\max}.$$

На рис. 14.2 построены абсолютные значения спектров волн на выходе из голограммы. Из него видно, что использование опорной волны с соответствующей высокой пространственной частотой (т.е. большим углом падения θ) обеспечивает разделение волн за голограммой.

Как следует из этого рисунка наложения волн нулевого порядка дифракции и образующих изображения волн первых порядков дифракции можно избежать, если пространственная частота опорной волны удовлетворяет соотношению:

$$|\xi_{\text{опорн}} - \xi_0| = |-\xi_r - \xi_0| = \xi_r + \xi_0 \geq 3\xi_{\max}, \quad (14.8)$$

где ξ_0 – центральная пространственная частота спектра предмета, которую мы при анализе считали равной нулю. Выражение (14.8) представляет собой условие возможности пространственного разделения дифракционных порядков голограммы. Физически, оно сводится к необходимости выбора таких углов падения опорной и восстанавливающей волн, при которых величины углов, составляемых направлениями распространения объектной и каждой из вышеназванных волн, превышали бы полуторный угловой размер объекта. Отсюда следует, что для выполнения условия минимального углового разделения волн, светочувствительный материал, используемый для записи голограммы, должен обладать высокой разрешающей способностью в направлении оси x .

Заменим в основном уравнении голографии (14.1) комплексную амплитуду объектной волны $O(x,y)$ в плоскости голограммы ее компонентой в направлении x с наивысшей пространственной частотой $\xi_0 + \xi_{\max}$,

$$A(\xi_0 + \xi_{\max}) \exp[-2\pi i(\xi_0 + \xi_{\max})x]. \quad (14.9)$$

Тогда (14.1) можно будет представить в виде:

$$\begin{aligned} I &= A^2 + r^2 + A(\xi_0 + \xi_{\max})r \exp[-2\pi i(\xi_0 + \xi_{\max} + \xi_r)x] + \\ &+ A(\xi_0 + \xi_{\max})r \exp[2\pi i(\xi_0 + \xi_{\max} + \xi_r)x] = \\ &= I_0 + (const) \cos[2\pi i(\xi_0 + \xi_{\max} + \xi_r)x]. \end{aligned} \quad (14.10)$$

В аргументе косинуса содержится частота полос $\xi_r + \xi_0 + \xi_{\max}$, которая должна быть зарегистрирована на светочувствительной среде.

Если учесть условие (14.8), то оказывается, что подлежащая регистрации пространственная частота должна в 4 раза превосходить наивысшую пространственную частоту объекта.

Требование высокой разрешающей способности регистрирующей среды при получении голограмм с наклонным опорным пучком, является своего рода

платой за пространственное разделение пучков, восстановленных с помощью голограммы.

Отметим, что выполнение соотношения (14.8) обеспечивает лишь условие углового разделения пучков нулевого и первых порядков и не гарантирует их пространственного разделения во всем пространстве за голограммой. Тем не менее, его выполнение свидетельствует о том, что в пространстве за голограммой существует такая область, в которой пучки будут пространственно изолированы друг от друга.

14.2. Голографирование с наклонным опорным пучком при диффузном освещении объекта

Рассмотренный нами ранее метод освещения частично освещенного прозрачного транспаранта плоской волной обладает рядом недостатков:

- Трудность наблюдения восстановленного мнимого изображения, заключающаяся в необходимости сканирования зрачком наблюдателя всей плоскости голограммы;
- Сильная неравномерность интенсивности объектной волны в плоскости регистрации голограммы, затрудняющая выбор интенсивности опорной волны.

Эти недостатки можно устранить, если использовать диффузное освещение голографируемого транспаранта. Для этого между лазерным источником и транспарантом помещают диффузный экран, например матовое стекло. Так как диффузный экран рассеивает свет в широком телесном угле, то наблюдателю не нужно сканировать зрачком всю поверхность голограммы, чтобы увидеть все изображение транспаранта. Однако фаза света, рассеянного диффузным экраном и прошедшего через объект, представляет собой быстро меняющуюся пространственную функцию координат в плоскости голограммы. Свет в этой плоскости может сохранять когерентные свойства. Это происходит, если:

- Исходная волна, освещающая диффузный экран, пространственно когерентна по всей площади экрана;
- Максимальная длина пути света от источника до голограммы через диффузный экран отличается от длины пути опорного пучка не больше, чем на длину когерентности;
- Экран остается неподвижным.

Голограмма, полученная при диффузном освещении, обладает рядом замечательных свойств. Дело в том, что диффузный экран имеет более широкий спектр пространственных частот, чем голографический транспарант; он рассеивает свет в широком телесном угле так, что каждая точка апертуры голограммы получает свет от всех точек транспаранта. Вследствие этого на стадии восстановления через любую часть голограммы можно наблюдать все мнимое изображение объекта.

При смещении направления наблюдения изображение видно с другой стороны. Если мы имеем голограмму двумерного транспаранта и хотим

Основы голографии

наблюдать его изображение, то сможем восстановить его целиком, даже для частично поврежденной голограммы.

Диффузная подсветка объекта, помимо преимуществ, имеет существенные недостатки:

Это зернистая, спекловая структура изображений, восстанавливаемых с помощью таких голограмм. Восстановленные изображения состоят из отдельных светлых пятен, разделенных абсолютно темными промежутками. Размер пятен находится на пределе разрешающей способности голограммы, а их контраст – V , определяется как отношение разности максимальной и минимальной интенсивностей элементов изображения к их сумме,

$$V = \frac{I_{\text{макс}} - I_{\text{мин}}}{I_{\text{макс}} + I_{\text{мин}}}, \quad (14.11)$$

равен 1.

Причина появления спеклов кроется в невозможности регистрации всего поля, рассеянного диффузором. Потеря и не регистрация на голограмме части поля объекта приводит к перераспределению интенсивности восстановленного изображения, имеющего вид пятнистой структуры.

Наличие спеклов ограничивает область практического использования голограмм с диффузной подсветкой объекта. На сегодня не существует радикального метода борьбы со спеклами

Единственное, что предлагается делать, это использовать метод последовательной регистрации множества реализаций одного и того же восстановленного изображения, характеризующихся различными картинками спеклов.

Практически этот метод реализуется путем установки вращающегося рассеивателя в восстанавливающий пучок лучей. Наличие вращающегося рассеивателя позволяет усреднить во времени различные реализации картин спеклов и свести их к постоянному по плоскости изображения шуму.

Вместе с тем рассеиватель обуславливает изменение структуры восстанавливающего пучка и, тем самым, приводит к снижению разрешения в восстановленном изображении.

Лекция № 15.**Материалы для записи голограмм****15.1. Определения плоских и объемных голограмм**

Ранее, рассматривая различные схемы регистрации голограмм, мы упоминали о двух типах голограмм, разделяющихся по величине отношения толщины регистрирующей среды к периоду следования полос в регистрируемом голографическом поле. При этом отмечали, что голограмма, обладающая толщиной, малой, по сравнению с расстоянием между поверхностями максимумов интенсивности регистрируемого голографического поля, действует подобно двумерной дифракционной решетке и, по существу, представляет собой систему полос на поверхности. Такие голограммы называются плоскими или тонкослойными. Голограммы же, получаемые на относительно толстых, по сравнению с периодом следования максимумов интенсивности голографического поля, регистрирующих средах, представляют собой объемную дифракционную решетку, состоящую из последовательности частично отражающих поверхностей и обладающую спектрально селективным откликом, описываемым законом Брэгга. Такие голограммы называются объемными или брэгговскими. В 1966 году Клейн определил нижний предел толщины объемных голограмм. Для этого он ввел параметр Q , определяемый с помощью следующего выражения:

$$Q = \frac{2\pi\lambda T}{nd^2} , \quad (15.1)$$

Основы голографии

где λ – длина волны излучения в воздухе, T – толщина голограммы, d – период следования максимумов интенсивности зарегистрированного голографического поля, n – средний показатель преломления регистрирующей среды. По Клейну голограмму можно считать объемной при значениях $Q \geq 10$. Рассмотрим типичный случай: $T = 15$ мкм, $\lambda = 0,633$ мкм – излучение He-Ne лазера, $n = 1,52$ – желатина. Положив $Q = 10$, найдем, что голограмму можно будет считать объемной при расстоянии между интерференционными поверхностями - d , равном 1,98 мкм. Отметим, что здесь мы не акцентируем внимание на вопросе, где измеряется период следования интерференционных поверхностей: в воздухе или в материале голограммы. Это происходит из-за равенства этих периодов. Покажем это на примере голограммы двух параллельных пучков, симметрично падающих под углами $\theta/2$ на плоскость регистрации голограммы (см. рис. 2.1). На этом рисунке обозначены: ML - граница раздела воздух - регистрирующая среда (регистрирующая среда со средним показателем преломления n располагается внизу под границей раздела), AB – расстояние между

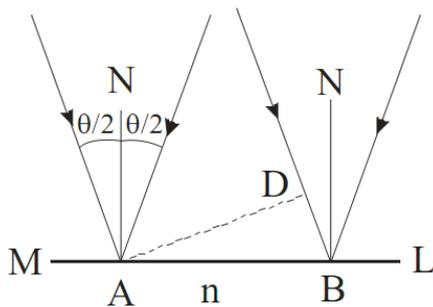


Рис.2.1. Регистрация голограммы двух параллельных пучков.

Основы голографии

соседними поверхностями максимумов интенсивности голографического поля, формируемого двумя параллельными пучками. N – нормали к поверхности регистрирующей среды, восстановленные из точек A и B , AD – нормаль, опущенная из точки A на луч параллельного пучка, падающий в точку B регистрирующей среды. Дальнейшие рассуждения проведем для воздушной среды, располагаемой вблизи границы раздела воздух – регистрирующая среда. Если положить, что в точке A опорный и объектный пучки лучей синфазны, то в точке B разность их фаз должна составлять 2π . Отсюда следует, что расстояние AB будет соответствовать расстоянию между соседними поверхностями максимумов интенсивности голографического поля лишь в случае выполнения условия:

$$BD = AB \sin\left(\frac{\theta}{2}\right) = \frac{\lambda}{2} . \quad (2.2)$$

Из (2.2) получим выражение, описывающее пространственный период – d голограммы двух плоских волн, симметрично падающих на плоскость ее регистрации:

$$d = AB = \frac{\lambda}{2 \sin\left(\frac{\theta}{2}\right)} . \quad (2.3)$$

Отсюда получим выражение, описывающее пространственную частоту рассматриваемой голограммной структуры – ν :

$$\nu = \frac{1}{d} = \frac{2 \sin\left(\frac{\theta}{2}\right)}{\lambda} . \quad (2.4)$$

Выражение (2.4) описывает пространственную частоту голографического поля вблизи поверхности регистрирующей среды. Для описания пространственной частоты голографического поля внутри регистрирующей среды необходимо учесть следующее. Во-первых, обусловленное законом Снеллиуса уменьшение синуса угла $\theta/2$ в n раз. Во-вторых, уменьшение рабочей длины волны излучения в регистрирующей среде, по сравнению с длиной волны в воздухе в n раз. Пропорциональное уменьшение числителя и знаменателя выражения (2.4) при переходе из воздуха в регистрирующую среду свидетельствует о равенстве пространственных частот голографического поля в воздухе и регистрирующей среде.

15.2. Амплитудные и фазовые голограммы. Среда для их регистрации.

Голограмма, как это уже отмечалось выше, представляет собой транспарант, обладающий амплитудным пропусканием, в той или иной степени соответствующим, в идеальном случае пропорциональным, распределению интенсивности голографического поля, сформированного объектной и опорной волнами на стадии регистрации голограммы. Очевидно, что изготовление такого транспаранта предопределяет использование той или иной светочувствительной регистрирующей среды. Ее экспонирование и, если необходимо, последующая фотохимическая обработка должны приводить к изменению ее оптических свойств. Напишем выражение, описывающее комплексную амплитуду плоской волны F , прошедшей расстояние T в некой оптически прозрачной среде с коэффициентом амплитудного поглощения α

$$F = A \exp\left(-i \frac{2\pi n_0 T}{\lambda}\right) \exp(-\alpha T), \quad (2.5)$$

Основы голографии

где A – комплексная амплитуда волны на входе в среду, n_0 – показатель преломления среды, λ – длина волны в воздухе. Из (2.5) следует, что у среды, пригодной для записи голограмм, при экспонировании и фотохимической обработке должен изменяться, по крайней мере, один из параметров α , n_0 или T . Обычно меняется существенно лишь один из этих параметров. Исходя из вышеизложенного, все обычно используемые регистрирующие среды для голографии делят на две основные группы. К первой из них относят среды с амплитудной модуляцией, т.е. среды с модулируемым коэффициентом амплитудного поглощения α . Наиболее яркими и часто встречающимися представителями таких сред являются галогенидосеребряные фотоэмульсии и фотохромные органические и неорганические материалы. Ко второй группе относятся среды с фазовой модуляцией, т.е. среды, у которых в процессе экспонирования или фотохимической обработки изменяется либо толщина, либо показатель преломления. Наиболее характерными представителями таких сред являются также галогенидосеребряные фотоэмульсии, органические и неорганические фоторезисты, бихромированная желатина, фотополимеры.

Основы голографии

Галогенидосеребряные фотоэмульсии являются универсальными регистрирующими средами, поскольку модулируемый с их помощью параметр полностью определяется характером их послезекспозиционной фотохимической обработки. Так, фотохимическая обработка, включающая в себя лишь проведение двух основных операций: проявления и фиксирования, приводит к образованию амплитудных голограмм. Если к проявлению и фиксированию добавить еще и отбеливание, т.е. перевод восстановленных при проявлении зерен металлического серебра в прозрачные для света соли серебра, то такая обработка приведет к образованию фазовых голограмм с модуляцией показателя преломления. Добавление к вышеперечисленным операциям проявления, фиксирования и отбеливания еще одной операции фиксирования, приводящей к вымыванию из желатиновой основы фотоэмульсии прозрачных солей серебра, обуславливает формирование чисто фазовых голограмм с модуляцией показателя преломления желатиновой матрицы фотослоя и рельефно-фазовых голограмм, характеризующихся модуляцией толщины слоя прошедшей фотохимическую обработку фотоэмульсии. Фотоэмульсии в голографии используются в виде слоев толщиной от десятых долей микрометра до десятка микрометров, политых на жесткие стеклянные, либо гибкие ацетатные подложки. Бесспорными достоинствами галогенидосеребряных фотоэмульсий являются: их универсальность, возможность очувствления практически к любому участку спектра оптического излучения, высокая светочувствительность, составляющая $10^{-3} - 10^{-5}$ Дж/см², а также высокая разрешающая способность, превышающая 3000 мм⁻¹. Столь же очевидны и их недостатки.

Основными являются повышенный уровень шумов из-за рассеяния излучения на зернах серебра, необходимость их нанесения на подложки из раствора, затрудняющая получение пленок высокой степени равнотолщинности, а также принципиально "мокрая", т.е. проводимая в растворах при набухании слоя фотохимическая обработка, обуславливающая деформацию структуры голограммы из-за усадки толщины слоя фотоэмульсии при его сушке.

Основы голографии

Бихромированная желатина применяется в голографии, в основном, для записи объемных пропускающих и отражательных голограмм. Толщина слоев желатины может достигать десятых долей миллиметра, что позволяет использовать ее для записи множества голограмм на одном и том же участке регистрирующей среды. В качестве достоинств бихромированной желатины как среды для записи голограмм можно назвать высокую разрешающую способность, превышающую 3000 мм^{-1} , возможность использования слоев большой толщины, низкий уровень поглощения излучения, низкий уровень светорассеяния. Она светочувствительна в диапазоне длин волн $0,25 - 0,5 \text{ мкм}$. Основными недостатками желатины являются: относительно низкая светочувствительность, составляющая 10^{-2} Дж/см^2 , нанесение на подложки из раствора, фотохимическая обработка с набуханием слоя и последующей усадкой, а также необходимость защиты изготовленной на желатине голограммы от влаги, приводящей к исчезновению голограммной структуры.

В настоящее время запись большинства объемных голограмм осуществляется на фотополимерах. Из них наибольшее распространение и известность получили фотополимеры фирмы Du Pont. Они выпускаются в промышленных масштабах и широко используются для изготовления защитных голографических меток, например голограмм на кредитных картах, банкнотах и т.д. Фотополимеры могут быть очувствлены практически в любом диапазоне видимой области спектра. Их разрешающая способность также превышает 3000 мм^{-1} , что позволяет использовать эти среды для регистрации отражательных голограмм по схеме Ю.Н. Денисюка. Их светочувствительность составляет десятки мДж/см^2 . В качестве основных достоинств фотополимеров можно отметить низкий уровень шума, простоту послеэкспозиционной обработки. Недостатком этих сред является сложность их нанесения на подложку в виде равнотолщинной пленки.

Основы голографии

Неорганические и органические фоторезисты в голографии, как правило, используются в виде тонких пленок, толщиной до 1 мкм. Такие среды используются для получения рельефно-фазовых голограмм, т.е. фазовых голограмм с модуляцией толщины регистрирующей среды. Они светочувствительны, как правило, в коротковолновой области спектра (менее 0,5 мкм) и обладают светочувствительностью 10^{-2} Дж/см². В голографии они обычно используются в виде нанесенных на подложки слоев толщиной менее 1 мкм. Нанесение органических резистов на подложки осуществляют методом центрифугирования, что при тщательном соблюдении технологии позволяет получать исключительно равнотолщинные пленки. Еще лучше обстоят дела с неорганическими фоторезистами, или, как их еще называют, халькогенидными стеклообразными полупроводниками. Эти среды наносят на подложки методом испарения в вакууме. Послеэкспозиционная обработка фоторезистов – мокрая, без набухания. При этом неорганические фоторезисты допускают сухую послеэкспозиционную обработку в газовой среде. В голографии фоторезисты, в основном, используются для изготовления голограммных оптических элементов и мастер – голограмм, являющихся основой для производства никелевых матриц для тиражирования голографических этикеток, наклеек и т.д. При использовании неорганических фоторезистов следует иметь в виду то, что их нанесение на подложки методом термического испарения в вакууме, позволяющим получать исключительно равнотолщинные пленки, имеет и обратную сторону, заключающуюся в кластерной природе формируемой таким образом пленки резиста. Пленка, получаемая методом термического испарения халькогенидного стекла в вакууме, может быть условно разбита на отдельные участки – кластеры, каждый из которых характеризуется постоянной, но отличной от соседних участков пленки, скоростью травления в проявляющем растворе. Отметим, что средний размер кластера, как правило составляет 400 – 600 нм. Кластерность пленки, в конечном счете, приводит к паразитной модуляции поверхностного рельефа получаемых на ней голограммных структур и обуславливает наличие коротковолновой границы применимости голограмм, получаемых на халькогенидных стеклообразных полупроводниках. Для голограмм, не проходящих апостериорную обработку поверхности и характеризующихся максимальной дифракционной эффективностью на рабочей длине волны

- λ , эта граница, определяемая из условия допустимого уровня светорассеяния $\sigma \leq \lambda/100$, где σ - среднеквадратичная шероховатость поверхности голограммы, соответствует 80 мкм. Однако, несмотря на указанные особенности халькогенидных стеклообразных полупроводников, они являются одними из основных претендентов на роль регистрирующей среды для получения прецизионных голограммных оптических элементов, предназначенных для работы в коротковолновой области спектра.

15.3. Максимально достижимая дифракционная эффективность тонких голограмм

Ранее было показано, что в поле, восстановленном с помощью голограммы, помимо основной восстановленной волны, соответствующей изображению объекта, присутствуют еще и дополнительные волны, например, нулевой и сопряженный основному восстановленному изображению порядка дифракции. Оценку энергетического коэффициента полезного действия голограммы принято

осуществлять с помощью специального параметра, называемого дифракционной эффективностью голограмм. Численно она определяется как отношение энергии излучения, дифрагировавшего в восстановленное изображение, к общей энергии излучения, падающего на голограмму при ее восстановлении. Дифракционная эффективность определяется толщиной используемой регистрирующей среды, типом реализуемой в ней модуляции оптических параметров, а также общими характеристиками среды, такими, как среднее поглощение, светорассеяние и т.д.

Рассмотрим максимально достижимую дифракционную эффективность тонких амплитудных и фазовых голограмм, т.е. дифракционную эффективность тонких голограмм, получаемых на идеальных средах, не обладающих светорассеянием и исходной оптической плотностью. Анализ проведем для случая голограмм, зарегистрированных с помощью плоской, наклонно падающей на регистрирующую среду опорной волны с комплексной амплитудой $R = r \exp(2\pi i \xi x)$ и аксиальной немодулированной плоской объектной волной с амплитудой o .

Основы голографии

Начнем с рассмотрения амплитудной голограммы. Предположим так же, как мы это делали раньше, что зависящая от экспозиции часть амплитудного коэффициента пропускания проявленной голограммы t_E пропорциональна интенсивности регистрируемой интерференционной картины. Тогда:

$$t_E = [o + r \exp(2\pi i \xi x)][o + r \exp(-2\pi i \xi x)] = o^2 + r^2 + 2or \cos(2\pi \xi x) = t_{E_0} + t_1 \cos(2\pi \xi x). \quad (2.6)$$

Для тонкой амплитудной голограммы полное амплитудное пропускание t определяется как $t = t_0 - t_E$, где t_0 – пропускание неэкспонированной регистрирующей среды. Максимальный диапазон изменения коэффициента пропускания лежит в пределах от 0 до 1. В нашем случае он будет реализован при $t_0 = 1$ и $t_{E_0} = t_1 = 1/2$, поскольку функция косинус меняется от -1 до 1. Отсюда можно сделать вывод о том, что максимальная модуляция амплитудного пропускания голограммы обеспечивается при равных друг другу амплитудах, а, следовательно, и интенсивностях опорной и объектной волн. При указанных выше условиях (2.6.) можно переписать в виде:

$$t = t_0 - t_E = 1 - \frac{1}{2} - \frac{1}{2} \cos(2\pi \xi x) = \frac{1}{2} - \frac{1}{4} \exp(2\pi i \xi x) - \frac{1}{4} \exp(-2\pi i \xi x). \quad (2.7)$$

Пусть голограмма восстанавливается плоской волной единичной

Основы голографии

амплитуды, падающей на ее поверхность по нормали. Тогда амплитуда волны, распространяющейся за голограммой, также может быть описана с помощью выражения (2.7). Из (2.7) следует, что восстанавливающее излучение, прошедшее через рассматриваемую нами амплитудную голограмму, дифрагирует только в нулевой, плюс первый и минус первый порядки дифракции. При этом, поскольку амплитуда дифрагированной волны первого порядка составляет $\frac{1}{4}$ равной единице амплитуды восстанавливающей волны, то дифракционная эффективность такой голограммы, определяемая по энергии, равна $\frac{1}{16}$, т.е. равна 6,25%. Практически используемые регистрирующие среды обладают остаточным поглощением и светорассеянием, поэтому рассчитанная нами дифракционная эффективность на практике обычно не достигается. Несколько большей дифракционной эффективностью могут обладать голограммы, пропускание которых описывается не косинусоидальной, а прямоугольной периодической функцией пространственных координат. Таким пропусканием могут обладать голограммы, синтезированные на компьютере. В этом случае в первой половине периода пропускание равно 0, а во второй оно равно 1. Первые два члена разложения прямоугольной функции в ряд Фурье имеют вид:

$$t_E = \frac{1}{2} + \frac{2}{\pi} \cos(2\pi\xi x) - \dots + \dots = \frac{1}{2} + \frac{1}{\pi} [\exp(2\pi i \xi x) + \exp(-2\pi i \xi x)] - \dots \quad (2.8)$$

Основы голографии

При $t = 1 - t_E$ дифракционная эффективность амплитудных голограмм с прямоугольным профилем функции амплитудного пропускания равна $1/\pi^2 = 10,1\%$. Отметим, что многоточие в (2.8) свидетельствует о наличии высших порядков в изображении, восстановленном с помощью рассматриваемых голограмм. Низкая дифракционная эффективность тонких амплитудных голограмм обуславливает их крайне редкое использование в практической деятельности.

Теперь перейдем к рассмотрению максимальной дифракционной эффективности фазовых тонких голограмм. При этом будем считать, что рассматриваемая нами фазовая регистрирующая среда не поглощает свет и обладает чисто фазовым коэффициентом амплитудного пропускания t :

$$t = \exp[i\varphi(x)]. \quad (2.9)$$

Из (2.9) следует, что линейная запись на тонких фазовых голограммах возможна лишь при малых значениях φ . Для нахождения максимально возможной эффективности голограммы снимем с φ это ограничение. Положим, что регистрирующая среда экспонируется и обрабатывается таким образом, что фазовый сдвиг $\varphi(x)$, приобретаемый плоской волной

Основы голографии

при прохождении обработанной голограммы, пропорционален интенсивности света, действовавшей при экспозиции, т.е.:

$$\varphi(x) = \varphi_0 + \varphi_1 \cos(2\pi\xi x). \quad (2.10)$$

Тогда амплитудное пропускание голограммы t будет иметь вид:

$$t = \exp(i\varphi_0) \exp[i\varphi_1 \cos(2\pi\xi x)]. \quad (2.11)$$

Если опустить постоянный фазовый множитель, то (2.11) можно представить в виде ряда Фурье:

$$t = \exp[i\varphi_1 \cos(2\pi\xi x)] = \sum_{n=-\infty}^{+\infty} i^n J_n(\varphi_1) \exp(in2\pi\xi x), \quad (2.12)$$

где J_n - функция Бесселя первого рода n -го порядка. Если голограмма с амплитудным пропусканием t освещается аксиальной плоской волной единичной амплитуды, то амплитуда дифрагированной волны $+1$ -го порядка описывается функцией $J_1(\varphi_1)$, вид которой представлен на рис. 2.2. Ее максимальное значение равно 0,582, а максимальная эффективность составляет 33,9%.

Основы голографии

Несколько большей дифракционной эффективностью обладает голограмма с фазовой модуляцией коэффициента амплитудного пропускания, описываемой прямоугольной функцией. Такая голограмма обладает максимальной эффективностью, равной 40,4%.

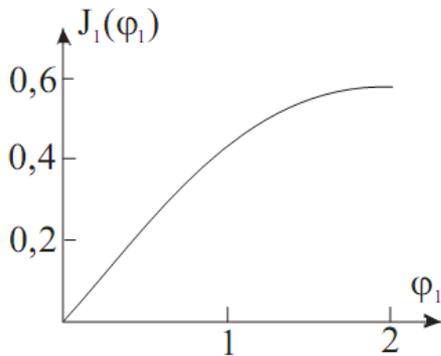


Рис. 2.2. График функции Бесселя первого рода первого порядка

Заканчивая рассмотрение дифракционной эффективности тонких фазовых голограмм, отметим, что примененное нами выше скалярное приближение теории дифракции справедливо лишь для случая малых величин отношения амплитуды фазовой модуляции голограмм к пространственному периоду их структуры. При больших амплитудах фазовой модуляции существенно более точные результаты дает строгое векторное решение задачи дифракции излучения на структуре

голограммы. В соответствии с ним, дифракционная эффективность тонкой фазовой, как правило, рельефно-фазовой, голограммы помимо амплитуды фазовой модуляции существенно зависит от угла падения на голограмму восстанавливающего излучения и состояния его поляризации. При этом максимальное значение дифракционной эффективности таких голограмм может достигать 80 - 90 %.

15.4. Максимально достижимая дифракционная эффективность объемных голограмм.

Рассмотрим поперечное сечение элементарной голограммы, зарегистрированной в объемной регистрирующей среде (см. рис.2.3.)

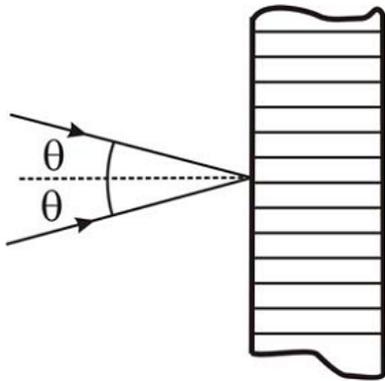


Рис.2.3. Элементарная объемная голограмма двух плоских волн

Представленные на рис 2.3. горизонтальные линии в толще регистрирующей среды представляют собой следы пересечения плоскости чертежа со стратами, т.е. с плоскостями зарегистрированных пучностей голографического поля, существовавшего в регистрирующей среде во время экспозиции. Предположим, что толщина регистрирующей среды равна 15 мкм, длина волны регистрирующего и восстанавливающего голограмму излучения равна 0,63 мкм, а углы падения опорного и объектного пучков на поверхность регистрирующей среды при записи голограммы составляли 15° . Можно сосчитать, что при указанных параметрах период следования страт составляет 1,22 мкм. Столь малая величина пространственного периода голограммной структуры приводит к многократному взаимодействию восстанавливающего пучка со стратами в процессе восстановления голограммы. Поэтому при анализе энергетических характеристик таких голограмм необходимо учитывать последовательное рассеяние восстанавливающего пучка лучей на всех встречающихся на его пути стратах голограммной структуры. Чтобы амплитуда результирующей дифрагированной волны была максимальной, волны, рассеянные последовательно расположенными стратами, должны быть синфазны.

Основы голографии

Для этого необходимо, чтобы выполнялось определенное соотношение между длиной волны λ , углом θ и периодом следования страт d . Это соотношение представляет собой закон Брэгга, который можно описать в виде следующего выражения:

$$2d \sin \theta = \lambda , \quad (2.13)$$

где λ – длина волны в воздухе. Закон Брэгга определяет однозначную связь между периодом следования страт, длиной волны излучения и углом падения восстанавливающего пучка лучей. Отсюда следует, что голограммы, свойства которых описываются законом Брэгга, являются селективными по отношению к параметрам восстанавливающего пучка лучей.

Отметим, что в общем случае страты располагаются параллельно биссектрисе угла между опорным и объектным пучками. Относительно этой же биссектрисы и определяется, в общем случае угол θ , а период следования страт определяется вдоль опущенной на их поверхность нормали.

Дифракционная эффективность объемных голограмм не может быть рассчитана в рамках примененного нами ранее Фурье – анализа структуры голограммы, поскольку он не учитывает последовательное многократное рассеяние излучения на стратах. К настоящему времени

Основы голографии

разработано несколько подходов к исследованию энергетических и селективных свойств объемных голограмм. Среди них подходы, основанные на разбиении объемной голограммы на тонкие слои, на модовой теории излучения, на теории слоистых сред. Наиболее простой и эффективной в использовании в настоящее время считается разработанная Когельником теория связанных волн, основанная на учете постоянного обмена энергией между восстанавливающей и восстановленной волной в процессе их прохождения через объем голограммы. Эта теория не только предсказывает селективный отклик объемных голограмм, но также правильно описывает их высокую, достигающую до 100%, дифракционную эффективность. Столь высокая дифракционная эффективность объемных голограмм может быть объяснена следующим образом. Ранее мы показали, что максимальная эффективность тонкой фазовой голограммы может достигать 40,4%. Теперь представим объемную фазовую голограмму в виде стопы последовательно расположенных тонких фазовых голограмм, причем расположенных так, что обеспечивается синфазное сложение восстановленных пучков лучей. Именно синфазность пучков, восстановленных отдельными слоями объемной голограммы, и обеспечивает ее высокую дифракционную эффективность. В рамках настоящего курса мы не будем подробно останавливаться на теории

связанных волн. Отметим лишь то, что она предсказывает возможность достижения 100% дифракционной эффективности у объемных фазовых пропускающих и отражательных голограмм без потерь. При этом предсказанные теорией Когельника максимальные величины дифракционной эффективности адсорбционных пропускающих и отражательных голограмм не превышают 3,7% и 7,2% , соответственно. Естественно, максимальные величины дифракционной эффективности объемных голограмм реализуются при точном выполнении закона Брэгга и при оптимальных для каждого типа голограмм значениях амплитуды модуляции показателя преломления или коэффициента поглощения. Говоря о селективности объемных голограмм, следует отметить существенную зависимость ее характера от типа голограммы. Так, объемные пропускающие голограммы, регистрируемые в попутных пучках, т.е. в пучках лучей, падающих на регистрирующую среду с одной и той же стороны, обладают угловой селективностью и не обладают спектральной селективностью. Говоря иначе, их дифракционная эффективность существенно зависит от величины угла падения восстанавливающего пучка лучей и практически не зависит от его длины волны. То есть в случае пропускающих объемных голограмм

Основы голографии

для любой из достаточно широкого спектрального диапазона длины волны восстанавливающего излучения могут быть найдены такие условия восстановления голограммы, при которых будет реализована ее максимальная дифракционная эффективность, примерно равная эффективности этой же голограммы при ее восстановлении опорным пучком лучей при выполнении закона Брэгга. В отличие от пропускающих голограмм, отражательные объемные голограммы обладают ярко выраженной спектральной селективностью. Это позволяет осуществлять их восстановление немонахроматическим излучением белого света. Из всего спектра падающего на голограмму излучения она выделит и использует для формирования восстановленного изображения лишь излучение узкого спектрального диапазона, удовлетворяющего закону Брэгга.

Лекция № 16.**Интеграл суперпозиции**

Ранее мы с вами рассмотрели энергетические характеристики голограмм, т.е. определили максимально возможную дифракционную эффективность и селективные, с точки зрения энергетики, свойства голограмм различного типа. Настоящий раздел курса посвятим анализу изображающих свойств голограмм, т.е. анализу их аберрационных характеристик. При этом сразу отметим, что толщина регистрирующей среды, в случае отсутствия ошибок структуры голограммы, т.е. ошибок в положении и деформаций страт, не оказывает никакого влияния на качество формируемого с ее помощью изображения. Следовательно, изображающие свойства голограмм, т.е. их свойства как оптической системы, не зависят от толщины регистрирующей среды и полностью определяются картиной выхода страт на поверхность. В связи с этим в рамках данного курса мы ограничимся рассмотрением пропускающих тонкослойных голограмм точечных источников. Сначала обоснуем правомочность ограничения видов рассматриваемых голограмм голограммами точечных источников, основанную на возможности прогнозирования изображающих свойств голограмм любых объектов по результатам анализа голограмм точечных объектов. Для этого представим голограмму в виде некоторой оптической системы и рассмотрим ее основные свойства.

16.1. Линейность оптических систем и интеграл суперпозиции.

Голограмма, как и любая другая оптическая система, как правило, может рассматриваться как детерминированная система, в которой определенный входной сигнал, восстанавливающий пучок лучей, преобразуется в единственный выходной сигнал. При этом совершенно необязательно, чтобы входной сигнал приводил бы к формированию выходного сигнала. Так, восстановление пропускающей голограммы на стеклянной подложке излучением дальней УФ области спектра, лежащей вне пределов спектральной полосы пропускания стекла, не приводит к формированию какого-либо выходного сигнала на выходе системы, но от этого оптическая система не перестает быть детерминированной.

Для описания системы воспользуемся математическим оператором $\Phi\{\}$, описывающим связь между входными и выходными функциями. Отсюда, если функция $g_1(x_1, y_1)$ описывает входной сигнал системы, а

Основы голографии

$g_2(x_2, y_2)$ описывает соответствующий выходной сигнал, то при указанном выше определении оператора $\Phi\{\}$ эти две функции оказываются связанными соотношением:

$$g_2(x_2, y_2) = \Phi\{g_1(x_1, y_1)\} . \quad (3.1)$$

Если для всех входных функций t и s и для всех комплексных постоянных a и b выполняется свойство суперпозиции, т.е.

$$\Phi\{as(x_1, y_1) + bt(x_1, y_1)\} = a\Phi\{s(x_1, y_1)\} + b\Phi\{t(x_1, y_1)\} , \quad (3.2)$$

то такая система называется линейной. Наиболее значимая особенность линейных систем состоит в том, что их отклик на произвольный входной сигнал можно выразить через отклики на определенные элементарные сигналы, на которые следует разложить входной сигнал. В качестве такого элементарного сигнала выберем сигнал, описываемый δ – функцией. Тогда входной сигнал $g_1(x_1, y_1)$ мы можем представить в виде:

$$g_1(x_1, y_1) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g_1(\xi, \eta) \delta(x_1 - \xi, y_1 - \eta) d\xi d\eta . \quad (3.3)$$

Основы голографии

Соотношение (3.3.) является представлением g_1 в виде линейной комбинации взвешенных и смещенных δ – функцией. Чтобы найти отклик системы на входной сигнал g_1 , подставим (3.3) в (3.1):

$$g_2(x_2, y_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g_1(\xi, \eta) \Phi\{\delta(x_1 - \xi, y_1 - \eta)\} d\xi d\eta . \quad (3.4)$$

Из (3.4) получим выражение, описывающее отклик системы в точке с координатами (x_2, y_2) на сигнал вида δ – функции в точке с координатами (ξ, η) на входе

$$h(x_2, y_2; \xi, \eta) = \Phi\{\delta(x_1 - \xi, y_1 - \eta)\} . \quad (3.5)$$

Функция h называется импульсным откликом системы. Пользуясь понятием импульсного отклика системы, мы можем довольно просто описать связь ее входного и выходного сигналов:

$$g_2(x_2, y_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g_1(\xi, \eta) h(x_2, y_2; \xi, \eta) d\xi, d\eta . \quad (3.6)$$

Выражение (3.6), известное как интеграл суперпозиции, выражает тот очень важный факт, что линейная система полностью характеризуется ее откликом на входные импульсы. Чтобы полностью определить выходной сигнал, необходимо, вообще говоря, знать эти отклики для любых координат в плоскости входного сигнала.

В случае линейных оптических систем, а именно такими и являются голограммы, этот вывод означает то, что изображающие свойства оптической системы (в том числе и голограммы) полностью определяются изображениями точечных источников, располагаемых по всему полю объектов. Следовательно, зная изображающие свойства голограмм точечных источников, мы можем прогнозировать свойства голограмм любых, сколь угодно сложных объектов.

16.2. Геометрический анализ голограмм точечного источника.

Рассмотрим схему записи голограммы, представленную на рис. 3.1.

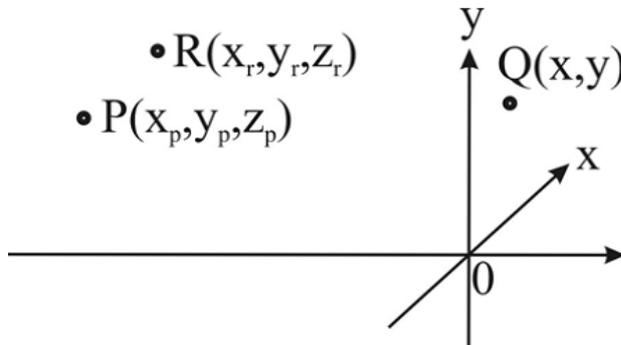


Рис. 3.1. Схема регистрации голограммы

Здесь P - объектный точечный источник, x, y – плоскость расположения регистрирующей среды, R – опорный точечный источник. Опишем реализуемую в приведенной выше схеме разность фаз между предметной и опорной волнами на поверхности голограммы. Для этого будем считать, что объектная, опорная и восстанавливающая волны в любой точке Q в плоскости голограммы характеризуются фазой в этой точке по отношению к фазе в фиксированном начале координат O . При этом будем считать, что амплитуды сферических волн, исходящих из двух вышеназванных источников P, R и восстанавливающего источника, постоянны в плоскости голограммы. Пусть пространство по обе стороны от голограммы имеет одинаковый показатель преломления; тогда относительные фазы можно рассчитать по геометрическим разностям хода световых лучей. Для этого воспользуемся методом Мейера, справедливым в параксиальном приближении.

Основы голографии

Пусть $A = a_0 \exp(i\varphi_0)$ – комплексная амплитуда порожденной источником P объектной волны в плоскости регистрации голограммы. $R = r_0 \exp(i\varphi_r)$ – комплексная амплитуда порожденной источником R опорной волны в плоскости регистрации голограммы. Тогда интенсивность голографического поля I в плоскости регистрации голограммы в соответствии с основным уравнением голографии будет описываться следующим выражением:

$$I = a_0^2 + r_0^2 + RA^* + R^*A. \quad (3.7)$$

Наибольший интерес в (3.7) для нас представляют интерференционные члены

$$RA^* + R^*A = 2a_0r_0 \cos(\varphi_r - \varphi_a), \quad (3.8)$$

которые описывают периодические пространственные вариации интенсивности голографического поля, т.е. интерференционные полосы. Пространственная частота полос на голограмме определяется скоростью изменения аргумента косинуса в выражении (3.8), описывающего разность фаз $\varphi_r - \varphi_a$. Вернемся к рис. 3.1. Объектный источник P расположен в плоскости, отстоящей от начала координат O на расстояние $z_p = -d$. Начало координат O лежит в плоскости голограммы x, y . Опорный точечный источник R расположен в точке с координатами (x_r, y_r) на расстоянии z_r от плоскости голограммы. Если R находится слева от голограммы и опорная волна исходит из R , то $z_r < 0$; если R

справа от плоскости голограммы и опорная волна сходится в точку R , то $z_r > 0$. Наша задача состоит в расчете величины $\varphi_r - \varphi_a$ в произвольной точке Q в плоскости регистрации голограммы. Чтобы определить число периодов изменения интенсивности на единицу расстояния вдоль координатных осей, т.е. пространственную частоту, продифференцируем величину $(1/2\pi)(\varphi_r - \varphi_a)$ по пространственной координате. Тогда мы узнаем, сколько интерференционных полос на единицу длины в данном направлении должно быть зарегистрировано в зависимости от взаимного расположения источников. Начальные фазы волн, исходящих из P и R совершенно произвольны. Пусть они выбраны так, что в точке O в плоскости регистрации голограммы фазы обеих волн одинаковы. Мы можем считать эти значения фаз равными нулю. Поскольку P и R – точечные источники, каждый из них излучает сферическую волну, фаза которой в любой точке пространства пропорциональна радиальному расстоянию от этой точки до источника.

Тогда, вычислив разность хода $PQ - PO$, мы получим фазу φ_a световой волны, приходящей в точку Q из точки P . Аналогично можно рассчитать

фазу φ_r для волны, приходящей в точку Q из источника R . Следует обратить внимание на знак фазы в точке Q , для чего рассмотрим рис.3.2.

Модуль величины разности фаз φ_a , соответствующей разности хода $PQ - PO$, равен $|\varphi_a| = (2\pi/\lambda)|PQ - PO|$, где λ – длина волны. Если P – реальный точечный источник, испускающий расходящуюся сферическую волну, и если $PQ > PO$, то волновой фронт, пришедший в точку Q , испущен источником раньше, чем волновой фронт, одновременно достигший точки O . Поэтому фаза волнового фронта в точке Q должна быть

меньше фазы в O (предполагается, что фаза возрастает со временем) и, следовательно, $\varphi_a = - (2\pi/\lambda)(PQ - PO)$. Если R представляет собой точку, в которой фокусируется опорная волна, то она должна находиться по другую сторону от голограммы по отношению к точке P . Для $RQ > RO$ фаза волнового фронта в Q больше, чем в O , поскольку волновой фронт, достигший точки Q , был испущен позднее. Таким образом, для сходящейся опорной волны $\varphi_r = + (2\pi/\lambda)(RQ - RO)$, в то время как в обычном случае расходящейся опорной волны $\varphi_r = - (2\pi/\lambda)(RQ - RO)$.

Основы голографии

Теперь вернемся к вычислению разности фаз $\varphi_r - \varphi_a$ для случая, когда и P и R являются источниками расходящихся сферических волн, расположенными по одну сторону от голограммы. Для фазы объектной волны в Q получаем

$$\varphi_a = -\frac{2\pi}{\lambda}(PQ - PO) = -\frac{2\pi}{\lambda} \left\{ \left[(x - x_p)^2 + (y - y_p)^2 + z_p^2 \right]^{1/2} - \left[x_p^2 + y_p^2 + z_p^2 \right]^{1/2} \right\} =$$

$$-\frac{2\pi}{\lambda} z_p \left\{ \left[1 + \frac{(x - x_p)^2 + (y - y_p)^2}{z_p^2} \right]^{1/2} - \left[1 + \frac{x_p^2 + y_p^2}{z_p^2} \right]^{1/2} \right\}, \quad (3.9)$$

где предполагается, что значение z_p отрицательно, так, что знак φ_a остается отрицательным. Если P и Q расположены не слишком далеко от оси z , и если z_p достаточно велико, то (3.9) можно, в первом приближении, переписать в виде:

$$\varphi_a \approx \frac{2\pi}{\lambda} \left[\frac{1}{2z_p} (x^2 + y^2 - 2xx_p - 2yy_p) \right]. \quad (3.10)$$

Такое приближение вполне удовлетворительно описывает структуру голограммы. Вместе с тем, его совершенно недостаточно для описания изображающих свойств голограмм точечного источника, в первую очередь их aberrаций. В дальнейшем мы обсудим этот вопрос более

подробно.

Фаза $\varphi_r(x, y)$ опорной волны в точке Q может быть рассчитана аналогичным способом. В результате получим:

$$\varphi_r \approx \frac{2\pi}{\lambda} \left[\frac{1}{2z_p} (x^2 + y^2 - 2xx_r - 2yy_r) \right]. \quad (3.11)$$

Тогда разность фаз объектной и опорной волны в точке Q будет описываться выражением:

$$\begin{aligned} \varphi_r - \varphi_a &\approx \frac{2\pi}{\lambda} \left[(x^2 + y^2) \left(\frac{1}{2z_r} - \frac{1}{2z_p} \right) - x \left(\frac{x_r}{z_r} - \frac{x_p}{z_p} \right) - y \left(\frac{y_r}{z_r} - \frac{y_p}{z_p} \right) \right] = \\ &= \frac{2\pi}{\lambda} \Delta l. \end{aligned} \quad (3.12)$$

Величина в квадратных скобках в (3.12) представляет собой разность хода между световыми волнами, пришедшими в точку Q из точек P и R .

16.2.1. Осевая голограмма

В этом случае объектный и опорный источники находятся на оси z , так, что в (3.12) x_p, y_p, x_r, y_r равны нулю. Светлые полосы в интерференционной картине образуются при $\Delta l = n\lambda$, где n – целое число. Поскольку разность хода Δl симметрична относительно начала координат, полосы в данном случае имеют вид окружностей и интерференционная картина представляет собой зонную структуру, описываемую выражением:

$$\Delta l = \frac{x^2 + y^2}{2} \left(\frac{1}{z_r} - \frac{1}{z_p} \right) = \frac{\rho^2}{2} \left(\frac{1}{z_r} - \frac{1}{z_p} \right), \quad (3.13)$$

где ρ – отсчитываемое в плоскости голограммы расстояние от начала координат до исследуемой точки. В соответствии с (3.8) интенсивность интерференционной картины изменяется в пространстве по косинусоидальному закону. Имеющая место в общем случае

 Основы голографии

нелинейность зависимости Δl от пространственных координат свидетельствует о наличии зависимости пространственной частоты от координат на поверхности голограммы. При этом в каждой точке поверхности голограммы ее пространственная частота может быть охарактеризована с помощью параметра, называемого локальной

пространственной частотой – $\nu(\rho)$. Локальную пространственную частоту можно определить как деленную на 2π пространственную скорость изменения фазы интерференционной картины в точке Q:

$$\nu(\rho) = \frac{\partial(\varphi_r - \varphi_a)}{\partial\rho} \frac{1}{2\pi} = \frac{\partial}{\partial\rho} \left(\frac{\Delta l}{\lambda} \right). \quad (3.14)$$

Подставив (3.13) в (3.14) получим:

$$\nu(\rho) = \frac{\rho}{\lambda} \left(\frac{1}{z_r} - \frac{1}{z_p} \right). \quad (3.15)$$

Из (3.15) следует, что спектр пространственных частот осевой голограммы полностью определяется длиной волны записи голограммы, ее апертурой и расстояниями от объектного и опорного источников до плоскости регистрации голограммы. Чем ближе расположены опорный и объектный источники друг к другу, тем уже спектр пространственных частот в формируемом ими голографическом поле. Отметим, что выражение (3.15) позволяет нам оценить необходимую разрешающую способность среды выбираемой нами для регистрации голограммы. Разрешающая способность регистрирующей среды должна быть больше максимальной локальной пространственной частоты регистрируемой голограммы.

16.2.2. Внеосевая голограмма

Подставляя в (3.13) $\Delta l = n\lambda$, получим:

$$\Delta l = (x^2 + y^2) \left(\frac{1}{2} \right) \left(\frac{1}{z_r} - \frac{1}{z_p} \right) - x \left(\frac{x_r}{z_r} - \frac{x_p}{z_p} \right) - y \left(\frac{y_r}{z_r} - \frac{y_p}{z_p} \right) = n\lambda, \quad (3.16)$$

т.е. уравнение окружности с координатами центра:

$$x_j = \frac{z_p x_r - z_r x_p}{z_p - z_r}, \quad y_j = \frac{z_p y_r - z_r y_p}{z_p - z_r} \quad (3.17)$$

и радиусом кривизны, описываемым выражением:

$$\rho^2 = \left(\frac{z_p x_r - z_r x_p}{z_p - z_r} \right)^2 + \left(\frac{z_p y_r - z_r y_p}{z_p - z_r} \right)^2 + \frac{2n\lambda z_p z_r}{z_p - z_r}. \quad (3.18)$$

Локальные пространственные частоты внеосевой голограммной структуры, вдоль какой либо координаты на поверхности голограммы (x или y), также, как и в случае осевой голограммы, можно определить с помощью выражения (3.16), разделив его на λ и взяв частную производную по интересующей нас координате. Так распределение локальных пространственных частот по координате x - ξ может быть описано с помощью следующего выражения:

$$\xi = \left[x \left(\frac{1}{z_r} - \frac{1}{z_p} \right) - \left(\frac{x_r}{z_r} - \frac{x_p}{z_p} \right) \right] \frac{1}{\lambda}. \quad (3.19)$$

Из (3.19) следует, что пространственная частота внеосевой голограммы в ее центре при $x=0$ не равна 0 и определяется величинами углов падения опорного и объектного пучков на поверхность регистрации голограммы. Зависимость пространственной частоты от пространственной координаты апертуры голограммы, так же, как и в случае осевой голограммы, линейна.

Основы голографии

В качестве примера внеосевой голограммы рассмотрим голограмму, образованную при интерференции плоской волны, нормально падающей на поверхность регистрации голограммы ($x_r = y_r = 0, z_r = \infty$), со сферической объектной волной, исходящей из точки ($x_p, y_p = 0, z_p$). Координаты центра системы круговых интерференционных полос, радиусы которых соответствуют целым значениям n в (3.18), определяются выражением (3.17) и равны $x_j = x_p$ и $y_j = 0$. Таким образом, центр интерференционной картины, имеющей вид зонной структуры, является в рассматриваемом случае основанием перпендикуляра, опущенного из P на плоскость регистрации голограммы. Если центр регистрирующей среды находится в точке O , то будет зарегистрирована внеосевая часть интерференционной картины, что соответствует схеме получения голограмм по Лейту и Упатниексу. Частоту интерференционных полос ξ в направлении x можно найти, дифференцируя величину $\Delta l / \lambda$, определяемую выражением (3.16), при условии $x_r = y_r = y_p = 0$ и $z_r = \infty$. В результате получаем:

$$\xi = -\frac{x}{z_p \lambda} + \frac{x_p}{z_p \lambda} . \quad (3.20)$$

Предполагая, что центр голограммы находится в O , сравним полученное выражение для пространственной частоты (3.20) с соответствующим выражением, полученным для случая осевой голограммы. В центре осевой голограммы, как мы помним, пространственная частота полос равна 0, в то время как для внеосевой схемы частота в центре равна $x_p / z_p \lambda$. По мере удаления от центра голограммы вдоль оси x в одну из сторон частота растет, а в другую сторону убывает. Краям голограммы соответствуют максимальное и минимальное значения пространственных частот.

Основы голографии

Завершая анализ структуры внеосевых голограмм точечных источников, необходимо сделать следующее замечание, касающееся вида соответствующих им интерференционных полос. Сделанный нами ранее вывод о том, что интерференционные полосы на внеосевой голограмме имеют вид окружностей, явился следствием сделанных нами ограничений в числе членов разложения радикала в степенной ряд. На самом деле интерференционные полосы на внеосевой голограмме представляют собой проекцию на плоскость регистрации голограммы интерференционных полос в виде окружностей, формируемых опорным и объектным источниками на плоскости, проходящей через плоскость регистрации голограммы и ориентированной перпендикулярно прямой, соединяющей точечные опорный и объектный источники. Т.е. реально интерференционные полосы внеосевой голограммы точечных источников имеют форму эллипсов.

Итак, мы рассмотрели структуру голограмм точечных источников. Теперь обратимся к анализу их изображающих свойств, под которыми будем понимать основные закономерности формирования изображения, восстанавливаемого с помощью голограмм.

Лекция 17.

Изображающие свойства голограмм точечного источника

17.1. Изображающие свойства голограмм в первом параксиальном приближении

Воспользуемся основным уравнением голографии и перепишем его в более удобном для дальнейших преобразований виде:

$$H = C(|O|^2 + |R|^2 + O * R + OR^*). \quad (3.21)$$

Здесь H – описываемая с точностью до постоянного множителя комплексная амплитуда восстановленного волнового фронта, сформированного голограммой, полученной с помощью объектного волнового фронта O и опорного волнового фронта R , в плоскости самой голограммы, непосредственно за ней, при ее восстановлении волновым фронтом C . Напомним, что два последних слагаемых соответствуют волновым фронтам, которые расходятся от мнимого изображения объекта или сходятся к его действительному изображению, в то время как два первых слагаемых не несут никакой полезной информации об объекте. Мы будем рассматривать случай внеосевой голограммы Лейта и Упатниекса, при котором используются опорная и объектная волны, падающие под разными углами на плоскость регистрации голограммы, причем падающие под углами, позволяющими пространственно разделить все три порядка дифракции голограммы. Поэтому в выражении (3.21) нас будут интересовать лишь два слагаемых, соответствующих действительному и мнимому восстановленным изображениям, соответственно:

$$H_r = CO * R \quad \text{и} \quad H_v = COR^*. \quad (3.22)$$

Значки r и v введены и будут использоваться в дальнейшем для обозначения выражений, соответствующих действительному и мнимому изображениям, соответственно.

Основы голографии

Напомним, что целью проводимого нами исследования является изучение изображающих свойств голограммы точечных источников. Следовательно, с точностью до постоянного множителя, мы можем вместо комплексных амплитуд излучения этих источников в плоскости голограммы рассматривать соответствующие им распределения фаз: φ_r , φ_v , φ_o , φ_R , φ_C . Ранее, рассматривая структуру голограммы точечных источников, мы уже выводили выражение, описывающее распределение фазы излучения точечного источника $P(x_P, y_P)$ на удаленной от него на расстояние Z_P плоскости xu (см. рис. 3.3).

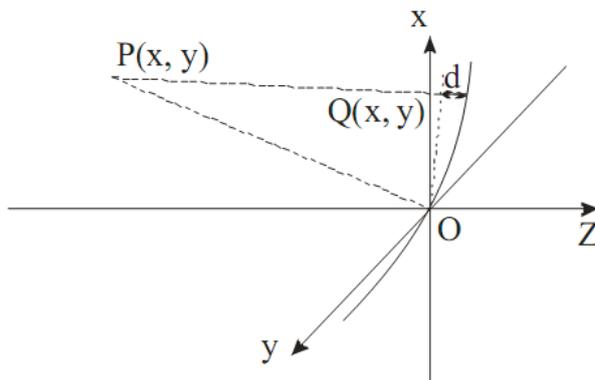


Рис. 3.3. К расчету распределения фазы

Основы голографии

Это выражение с точностью до значения фазы в начале координат может быть представлено следующим образом:

$$\begin{aligned} \varphi_P(x, y) &= -\frac{2\pi}{\lambda_P} d = -\frac{2\pi}{\lambda_P} (PQ - PO) = \\ &= -\frac{2\pi}{\lambda_P} \left\{ \sqrt{(x - x_P)^2 + (y - y_P)^2 + z_P^2} - \sqrt{x_P^2 + y_P^2 + z_P^2} \right\} = \\ &= \frac{2\pi}{\lambda_P} z_P \left\{ \sqrt{1 + \frac{[(x - x_P)^2 + (y - y_P)^2]}{z_P^2}} - \sqrt{1 + \frac{(x_P^2 + y_P^2)}{z_P^2}} \right\}. \end{aligned} \quad (3.23)$$

Со слагаемыми в (3.23), содержащими знаки радикала, мы поступим точно так же, как поступали с ними ранее при анализе структуры голограммы, т.е. разложим в степенной ряд. При этом в разложении ограничимся не двумя, как раньше, а тремя первыми членами ряда. Тогда из (3.23) получим:

$$\varphi_P(x, y) \approx \frac{2\pi}{\lambda_P} \left\{ \frac{x^2 + y^2 - 2xx_P - 2yy_P}{2z_P} - \frac{1}{8z_P^3} \left(x^4 + y^4 + 2x^2y^2 - 4x^3x_P - 4y^3y_P - 4x^2yy_P - 4xy^2x_P + 6x^2x_P^2 + 6y^2y_P^2 + 2x^2y_P^2 + 2y^2x_P^2 + 8xxy_Py_P - 4xx_P^3 - 4xx_Py_P^2 - 4yy_P^3 - 4xx_P^2y_P \right) \right\}. \quad (3.24)$$

Далее, если считать плоскость $xу$ плоскостью голограммы, располагаемой так, что ее геометрический центр совпадает с началом координат, то можно заключить, что выражение (3.24) будет справедливо для всех трех точечных источников, участвующих в процессах регистрации и восстановления голограммы. Тогда (3.22) мы можем переписать в виде:

$$\varphi_v = \varphi_C + \varphi_O - \varphi_R ; \quad \varphi_r = \varphi_C - \varphi_O + \varphi_R . \quad (3.25)$$

Из анализа выражений (3.25) можно сделать следующие выводы. Если распределения фаз опорного и восстанавливающего источников в плоскости голограммы идентичны, т.е. выполняется равенство $\varphi_C = \varphi_R$, то восстановленное с помощью голограммы распределение фазы φ_v будет абсолютно идентично распределению фазы объектной волны в процессе записи голограммы. Если же амплитуда восстанавливающей волны будет комплексно сопряжена амплитуде опорной волны, т.е.

будет выполняться равенство $\varphi_C = -\varphi_R$, то восстановленное распределение фазы φ_r будет равно $-\varphi_O$, т.е. произойдет восстановление обращенной объектной волны и будет сформировано неискаженное действительное изображение объекта. Второй из рассмотренных случаев называется случаем восстановления голограммы с использованием обращения волнового фронта.

Теперь рассмотрим общие случаи восстановления голограммы. Для этого положим, что распределения фаз φ_R и φ_v , соответствующие восстановленным в ± 1 порядках дифракции изображениям и реализуемые в плоскости голограммы непосредственно за ней, можно представить в виде суммы фазовых распределений, соответствующих гомоцентрическому пучку лучей, порожденному или сходящемуся в точечный источник сферической волны единичной амплитуды $I(x_I, y_I, z_I)$, и некоторой волновой аберрации, т.е. фазовой добавки $(2\pi/\lambda_C)W(x, y)$:

$$\varphi_{v,r}(x, y) \approx \frac{2\pi}{\lambda_C} \left\{ \begin{array}{l} \frac{x^2 + y^2 - 2xx_I - 2yy_I}{2z_I} - \\ \frac{1}{8z_I^3} \left(\begin{array}{l} x^4 + y^4 + 2x^2y^2 - 4x^3x_I - 4y^3y_I - 4x^2yy_I - \\ - 4xy^2x_I + 6x^2x_I^2 + 6y^2y_I^2 + 2x^2y_I^2 + 2y^2x_I^2 + \\ + 8xyx_Iy_I - 4xx_I^3 - 4xx_Iy_I^2 - 4yy_I^3 - 4xx_I^2y_I \end{array} \right) + \\ + W(x, y) \end{array} \right\}. \quad (3.26)$$

Подставим в (3.25) выражения, подобные (3.24), описывающие распределения фаз опорной, объектной и восстанавливающей волн и отличающиеся от (3.24) заменой индекса p на индексы R , O и C , соответственно, и приравняем полученное выражение к (3.26). При этом учтем, что, в общем случае, длина волны восстанавливающего источника λ_C может не совпадать с длиной волны излучения опорного и объектного источников λ_R . Длины же волн опорного и объектного источников должны быть, бесспорно, равны друг другу, поскольку источники должны быть когерентны. Возможное неравенство длин волн восстанавливающего и регистрирующих голограмму источников учтем путем введения коэффициента $\mu = \lambda_C / \lambda_R$.

Анализ полученного равенства, описывающего изображающие свойства голограммы, проведем применительно к двум случаям. Во-первых, рассмотрим, так называемое, первое параксиальное приближение, т.е. случай, когда $x, y, x_R, x_P, x_I, x_C, y_R, y_P, y_C, y_I$ много меньше z_P, z_R, z_C и z_I . При этом степенные ряды, ранее использовавшиеся нами, могут быть ограничены не тремя первыми членами, как это было при выводе выражений (3.24) и (3.26), а лишь двумя первыми членами, как мы это делали при анализе структуры голограммы. В этом случае анализ полученного нами равенства позволяет сделать вывод о существовании соотношений между длинами волн и координатами восстановленного, восстанавливающего, объектного и опорного источников, позволяющих обеспечить равенство нулю фазовой добавки $W(x,y)$, т.е. позволяющих обеспечить восстановление безабберационного изображения точечного источника I . Эти соотношения описываются выражениями, подобными хорошо известным формулам тонкой линзы:

$$\begin{aligned} \frac{1}{Z_I} &= \frac{1}{Z_C} \pm \mu \left(\frac{1}{Z_O} - \frac{1}{Z_R} \right) \\ \frac{X_I}{Z_I} &= \frac{X_C}{Z_C} \pm \mu \left(\frac{X_O}{Z_O} - \frac{X_R}{Z_R} \right) \\ \frac{Y_I}{Z_I} &= \frac{Y_C}{Z_C} \pm \mu \left(\frac{Y_O}{Z_O} - \frac{Y_R}{Z_R} \right). \end{aligned} \tag{3.27}$$

Основы голографии

В (3.27) верхний знак соответствует изображению, восстанавливаемому в -1 порядке дифракции и характеризующемуся распределением фазы φ_v , а нижний знак соответствует изображению, восстанавливаемому в +1 порядке дифракции и соответствующему распределению фазы φ_r .

Выражения (3.27) описывают изображающие свойства голограммы точечного объекта в первом параксиальном приближении. Первое из них описывает зависимость расстояния от голограммы до восстановленного изображения от расстояний от голограммы до опорного, объектного и восстанавливающего источников. Второе и третье выражения позволяют по известным x и y координатам объектного, опорного и восстанавливающего источников определить соответствующие координаты восстановленного изображения. Из первого уравнения (3.27) найдем выражение, описывающее фокусное расстояние голограммы - f_h . Для этого подставим в него $z_C = \infty$. Отсюда получим:

$$\frac{1}{f_h} = \pm \mu \left(\frac{1}{Z_O} - \frac{1}{Z_R} \right). \quad (3.28)$$

Если подставить (3.28) в первое уравнение (3.27), то можно получить:

$$\frac{1}{Z_I} = \frac{1}{Z_C} + \frac{1}{f_h}. \quad (3.29)$$

Сравнение (3.29) с известной формулой тонкой линзы свидетельствует об их идентичности. Отличаются от случая тонкой линзы лишь второе и третье уравнения, описывающие координаты изображения в плоскости его локализации. И это отличие обусловлено внеосевой схемой записи рассматриваемой нами обобщенной голограммы. Покажем это, рассмотрев линейное увеличение голограммы. Определим его, как отношение приращений dx_I и dy_I к вызвавшим их приращениям dx_C и dy_C . Из (3.27) получим:

$$\frac{dx_I}{dx_C} = \frac{z_I}{z_C}; \quad \frac{dy_I}{dy_C} = \frac{z_I}{z_C}. \quad (3.30)$$

Сравнение (3.29) и (3.30) с формулами тонкой линзы свидетельствует о подобии их импульсных откликов, а, следовательно, и изображающих свойств в первом приближении. Это позволяет нам считать голограмму точечного источника силовым голограммным оптическим элементом.

Сформулированные нами условия формирования голограммой безабберационного изображения в первом параксиальном приближении справедливы лишь для случая монохроматического излучения. Дело в том, что в (3.27) и в (3.28) входит параметр μ , зависящий от соотношения длин волн записи и восстановления голограммы.

Зависимость фокусного расстояния голограммы от длины волны приводит к тому, что она в случае использования немонохроматического восстанавливающего излучения формирует элементарные монохроматические изображения объекта, располагаемые на различных расстояниях от плоскости голограммы, что приводит к искажению формируемого ею цветного изображения объекта. Такое искажение изображения представляет собой хроматическую aberrацию, называемую хроматизмом положения. Численно она принимается равной расстоянию между двумя сформированными линзой элементарными монохроматическими изображениями, соответствующими границам рабочего спектрального диапазона голограммы. В волновой мере такая aberrация выражена быть не может.

17.2. Влияние конечных размеров голограммы. Структура изображения точечного объекта, формируемого голограммой в отсутствии аберраций

До настоящего времени при проведении анализа изображающих свойств голограммы точечного источника мы ограничивались сравнениями фазовых распределений объектного, опорного, восстанавливающего и восстановленного пучков лучей в плоскости голограммы. При этом мы не учитывали влияния конечных размеров апертуры голограммы на структуру формируемого с ее помощью изображения точечного объекта. Для учета влияния конечных размеров голограммы на структуру формируемого ею изображения точечного источника воспользуемся полученными ранее соотношениями и запишем выражение, описывающее в первом приближении реализуемую в плоскости голограммы комплексную амплитуду восстановленных с ее помощью изображений H_V и H_R :

$$H_{V,R}(x, y) \approx \exp \left\{ j \frac{2\pi}{\lambda_c} \left\{ \frac{1}{2} (x^2 + y^2) \left[\frac{1}{z_c} \pm \mu \left(\frac{1}{z_o} - \frac{1}{z_r} \right) \right] \right\} \right\} \times \\ \times \exp \left\{ -j \frac{2\pi}{\lambda_c} \left\{ x \left[\frac{x_c}{z_c} \pm \mu \left(\frac{x_o}{z_o} - \frac{x_r}{z_r} \right) \right] + y \left[\frac{y_c}{z_c} \pm \mu \left(\frac{y_o}{z_o} - \frac{y_r}{z_r} \right) \right] \right\} \right\}. \quad (3.31)$$

Для учета апертурного ограничения голограммы введем в (3.31) функцию $M(x, y)$, равную 1 в пределах светового размера голограммы, т.е. в пределах ее линейной апертуры, и равную 0 в остальных точках плоскости (x, y) . Тогда комплексную амплитуду восстановленного изображения в плоскости голограммы, сразу за голограммой, в первом параксиальном приближении с учетом апертурных ограничений $H'_{V,R}(x, y)$ мы можем описать в виде:

$$H'_{V,R}(x, y) = M(x, y) H_{V,R}(x, y). \quad (3.32)$$

Основы голографии

Для нахождения импульсного отклика тонкой линзы применим к выражению (3.32) преобразование Френеля, позволяющее в параксиальном приближении по комплексной амплитуде поля в плоскости голограммы найти распределение амплитуды в плоскости изображений, расположенной от голограммы на расстоянии z_I . Поскольку (3.32) описывает комплексную амплитуду прошедшего голограмму излучения точечного источника, то его преобразование Френеля будет с точностью до постоянной фазы описывать импульсный

отклик голограммы точечного источника $h(x_I, y_I; x_c, y_c)$:

$$h(x_I, y_I; x_c, y_c) = H_{V,R}(x_I, y_I) \approx \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H'_{V,R}(x, y) \exp\left\{-j \frac{2\pi}{\lambda_p} \frac{(x^2 + y^2)}{2z_I}\right\} \exp\left[j \frac{2\pi}{\lambda_p z_I} (xx_I + yy_I)\right] dx dy \quad (3.33)$$

Подставим (3.32) в (3.33):

$$h(x_I, y_I; x_c, y_c) \approx \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} M(x, y) \exp\left\{j \frac{2\pi}{\lambda_c} \frac{(x^2 + y^2)}{2} \left(\frac{1}{z_c} \pm \mu \left(\frac{1}{z_o} - \frac{1}{z_r}\right) - \frac{1}{z_I}\right)\right\} \times \exp\left\{j \frac{2\pi}{\lambda_c} \left[\left(\frac{x_I}{z_I} - \left(\frac{x_c}{z_c} \pm \mu \left(\frac{x_o}{z_o} - \frac{x_r}{z_r}\right)\right)\right)x + \left(\frac{y_I}{z_I} - \left(\frac{y_c}{z_c} \pm \mu \left(\frac{y_o}{z_o} - \frac{y_r}{z_r}\right)\right)\right)y\right]\right\} dx dy \quad (3.34)$$

Если в (3.34) подставить формулу голограммы (3.27), то можно получить выражение, описывающее импульсный отклик в плоскости изображения (x_I, y_I) :

Основы голографии

$$h(x_I, y_I; x_c, y_c) \approx \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} M(x, y) \exp \left\{ j \frac{2\pi}{\lambda_c} \left[\left(\frac{x_I}{z_I} - \left(\frac{x_c}{z_c} \pm \mu \left(\frac{x_o}{z_o} - \frac{x_r}{z_r} \right) \right) x + \left(\frac{y_I}{z_I} - \left(\frac{y_c}{z_c} \pm \mu \left(\frac{y_o}{z_o} - \frac{y_r}{z_r} \right) \right) y \right] \right\} dx dy . \quad (3.35)$$

Из (3.35) следует, что комплексная амплитуда формируемого голограммой изображения точечного источника, т.е. импульсный отклик голограммы, соответствует картине дифракции Фраунгофера на апертуре голограммы, причем картине, имеющей координаты центра:

$$x_I = \left[\frac{x_c}{z_c} \pm \mu \left(\frac{x_o}{z_o} - \frac{x_r}{z_r} \right) \right] z_I ; \quad (3.36)$$

$$y_I = \left[\frac{y_c}{z_c} \pm \mu \left(\frac{y_o}{z_o} - \frac{y_r}{z_r} \right) \right] z_I . \quad (3.37)$$

Если апертуру голограммы ограничить круглой апертурной диафрагмой, центр которой совпадает с центром поверхности голограммы, то распределение интенсивности в картине дифракции

излучения на диафрагме такой формы, как известно, будет описываться функцией вида:

$$I(r_0) = A \left[\frac{J_1(kDr_0/2z)}{kDr_0/2z} \right]^2 ; \quad (3.38)$$

где A – коэффициент пропорциональности, J_1 – функция Бесселя первого рода первого порядка, $k = 2\pi/\lambda_p$ – волновое число, z – расстояние от апертурной диафрагмы до плоскости наблюдения (в нашем случае оно равно z_I), D – диаметр апертурной диафрагмы, r_0 – радиус-вектор в плоскости наблюдения. В нашем случае он отсчитывается в плоскости изображений от точки с координатами (x_I, y_I) , т.е.

$$r_0^2 = \left(x_I - \frac{x_p}{z_p} z_I \right)^2 + \left(y_I - \frac{y_p}{z_p} z_I \right)^2 . \quad (3.39)$$

Отметим, что функция $J_1(x)/x$ при $x = 0$ равна 0,5.

Картину дифракции, описываемую соотношением (6.8), в оптике обычно называют картиной Эйри по имени ее первого исследователя. Распределение интенсивности в ней представлено на рис. 3.4, а ее общий

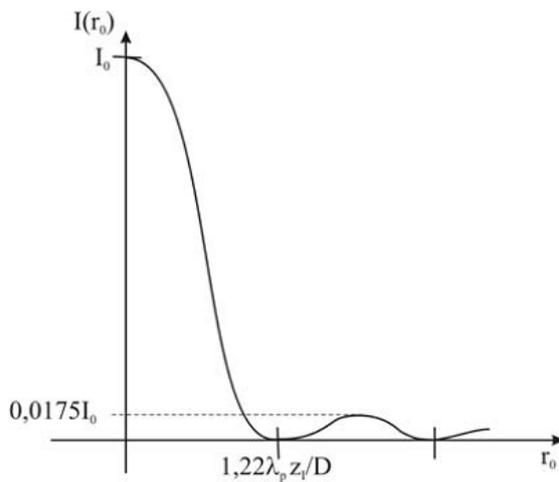


Рис. 3.4. Распределение интенсивности в картине Эйри

вид - на рис. 3.5. Таким образом, изображение точечного объекта, формируемое голограммой, ограниченной круглой дифракмой, имеет вид светлого пятна, окруженного concentрическими чередующимися темными и светлыми кольцами убывающей интенсивности. При этом в центральном пятне сосредоточено 83,8% всей энергии, формирующей изображение точечного источника, в первом светлом кольце – 7,2%, во втором кольце сосредоточено 2,8% энергии, а в третьем – менее 1,5%.



Рис. 3.5. Картина Эйри

Отметим, что наличие у голограммы, как и у любой другой оптической системы, aberrаций приводит, при их малых величинах, к перераспределению энергии в картине Эйри, а по мере увеличения aberrаций и к разрушению самой картины. При этом при величинах волновых aberrаций менее $\lambda_p/4$ согласно критерию, сформулированному Рэлеем, картина Эйри, формируемая оптической системой с aberrациями, полностью соответствует картине, формируемой безaberrационной системой.

17.3. Монохроматические aberrации третьего порядка

В предыдущем подразделе мы показали в первом параксиальном приближении способность голограмм формировать в монохроматическом излучении безабберационные изображения объектов. Напомним, что параксиальное приближение соответствует участку апертуры голограммы, расположенному вблизи ее центра. Вклад в изображение излучения, прошедшего через более удаленные от центра голограммы участки ее апертуры, приводит, в общем случае, к появлению aberrаций, т.е. к появлению искажений в изображениях, формируемых с помощью голограммы. Покажем это с помощью выражений (3.24) - (3.27), полученных при учете первых трех членов степенных рядов, описывающих выражения, стоящие под знаками радикала в (3.23). Для этого из (3.26) найдем выражение, описывающее aberrации $(2\pi/\lambda_p)W(x, y)$, реализуемые в плоскости голограммы, формирующей изображение точечного источника $I(x_I, y_I)$.

$$\frac{2\pi}{\lambda_c}W(x, y) = \frac{2\pi}{\lambda_c} \left\{ \begin{array}{l} -\frac{1}{8}(x^2 + y^2)^2 S + \frac{1}{2}(x^2 + y^2)(xC_x + yC_y) - \\ -\frac{1}{2}(x^2 A_x + 2xyA_{xy} + y^2 A_y) - \\ -\frac{1}{4}(x^2 + y^2)F + \frac{1}{2}(xD_x + yD_y) \end{array} \right\}, \quad (3.40)$$

где

$$S = \frac{1}{Z_C^3} \pm \mu \left(\frac{1}{Z_O^3} - \frac{1}{Z_R^3} \right) - \frac{1}{Z_I^3}; \quad (3.41)$$

$$C_x = \frac{X_C}{Z_C^3} \pm \mu \left(\frac{X_O}{Z_O^3} - \frac{X_R}{Z_R^3} \right) - \frac{X_I}{Z_I^3}; \quad C_y = \frac{Y_C}{Z_C^3} \pm \mu \left(\frac{Y_O}{Z_O^3} - \frac{Y_R}{Z_R^3} \right) - \frac{Y_I}{Z_I^3}; \quad (3.42)$$

$$A_x = \frac{X_C^2}{Z_C^3} \pm \mu \left(\frac{X_O^2}{Z_O^3} - \frac{X_R^2}{Z_R^3} \right) - \frac{X_I^2}{Z_I^3}; \quad A_y = \frac{Y_C^2}{Z_C^3} \pm \mu \left(\frac{Y_O^2}{Z_O^3} - \frac{Y_R^2}{Z_R^3} \right) - \frac{Y_I^2}{Z_I^3};$$

$$A_{xy} = \frac{X_C Y_C}{Z_C^3} \pm \mu \left(\frac{X_O Y_O}{Z_O^3} - \frac{X_R Y_R}{Z_R^3} \right) - \frac{X_I Y_I}{Z_I^3}; \quad (3.43)$$

$$F = \frac{X_C^2 + Y_C^2}{Z_C^3} \pm \mu \left(\frac{X_O^2 + Y_O^2}{Z_O^3} - \frac{X_R^2 + Y_R^2}{Z_R^3} \right) - \frac{X_I^2 + Y_I^2}{Z_I^3}; \quad (3.44)$$

$$D_x = \frac{X_C^3 + X_C Y_C^2}{Z_C^3} \pm \mu \left(\frac{X_O^3 + X_O Y_O^2}{Z_O^3} - \frac{X_R^3 + X_R Y_R^2}{Z_R^3} \right) - \frac{X_I^3 + X_I Y_I^2}{Z_I^3};$$

$$D_y = \frac{Y_C^3 + Y_C X_C^2}{Z_C^3} \pm \mu \left(\frac{Y_O^3 + Y_O X_O^2}{Z_O^3} - \frac{Y_R^3 + Y_R X_R^2}{Z_R^3} \right) - \frac{Y_I^3 + Y_I X_I^2}{Z_I^3}. \quad (3.45)$$

Проанализируем выражение (3.40) более детально. Оно состоит из пяти слагаемых, соответствующих пяти известным из курса оптики монохроматическим абберациям третьего порядка. При этом первое слагаемое (3.40), пропорциональное четвертой степени расстояния от центра голограммы до исследуемой точки ее апертуры, описывает сферическую абберацию. Соответствующий ей коэффициент S , зависящий от геометрических параметров схем записи и восстановления голограммы, описывается выражением (3.41). Из (3.41) следует, что сферическая абберация голограммы определяется расстояниями от центра голограммы до опорного, объектного и восстанавливающего источников и не зависит от их поперечных координат x, y . Поэтому она имеет место даже в случае осевой голограммы Габора. Физически сферическая абберация проявляется в перекачке энергии из центрального кружка пятна рассеяния в его периферийную область, что

Основы голографии

приводит к росту интенсивности дифракционных колец, окружающих главный максимум картины Эйри. Реализуемые в пятне рассеяния при наличии сферической aberrации распределения интенсивности представлены на рис. 3.6.

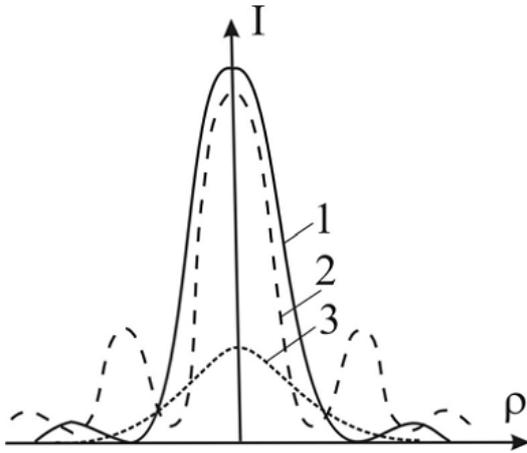


Рис. 3.6. Распределение интенсивности в пятне рассеяния при наличии сферической aberrации (1- aberrация отсутствует; 2 и 3 – изменение распределения интенсивности с ростом сферической aberrации)

Второе слагаемое в выражении (3.40) описывает aberrацию “кома”. Ее коэффициенты C_x и C_y описываются выражениями (3.42). Из (3.40) и (3.42) следует, что кома, в отличие от сферической aberrации, не является симметричной aberrацией относительно координат в плоскости голограммы, а также то, что ее величина зависит не только от расстояний от центра голограммы до опорного, объектного и восстанавливающего источников, но еще и от поперечных координат этих источников. Физически кома выражается, при небольших величинах, в несимметричной интенсивности дифракционных колец картины рассеяния. Большие величины aberrации “кома” приводят к преобразованию центрального круга пятна Эйри в запятую, отсюда и название “coma” (см. рис. 3.7).

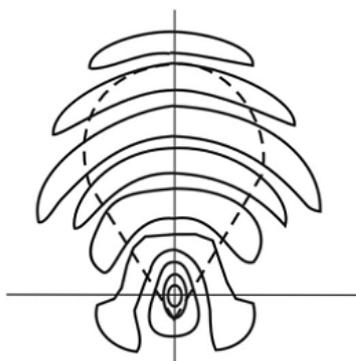


Рис. 3.7. Пятно рассеяния при наличии aberrации кома

Основы голографии

Третье слагаемое выражения (3.40) описывает aberrацию астигматизма. Ее коэффициенты A_x , A_y , A_{xy} описываются выражениями (3.43). Физически астигматизм выражается в преобразовании кружка рассеяния в два взаимно перпендикулярных штриха, располагаемых на

различных расстояниях от линзы. Структура пучка лучей, искаженного aberrацией астигматизм, представлена на рис. 3.8.

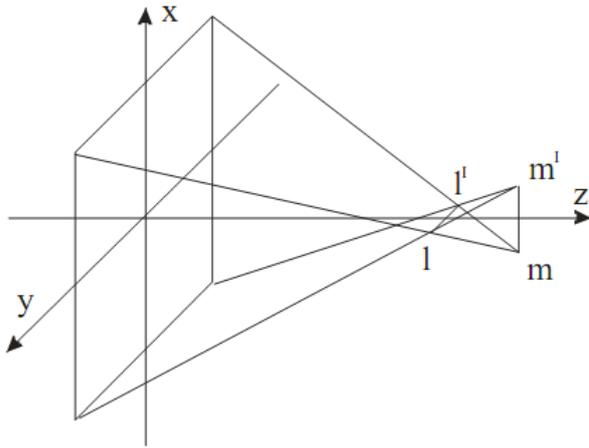


Рис. 3.8. Структура астигматического пучка лучей

Здесь ll' и mm' – взаимно перпендикулярные фокальные линии, в которые выродилось центральное ядро пятна Эйри. Вид пятна рассеяния, реализуемый в плоскости, расположенной посередине между фокальными штрихами, представлен на рис. 3.9.

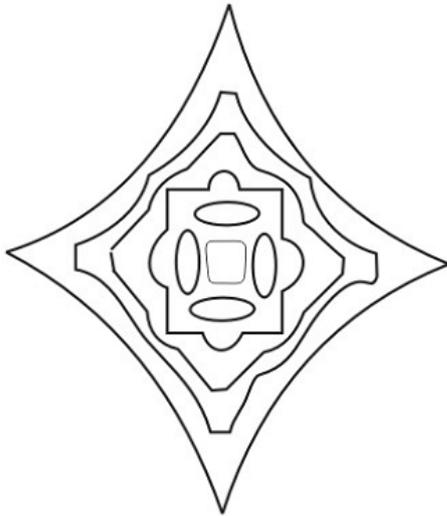


Рис. 3.9. Пятно рассеяния при наличии астигматизма в плоскости наилучшей фокусировки

Четвертое слагаемое в (3.40) описывает aberrацию кривизны поля. Коэффициент этой aberrации F описан выражением (3.44). Физически кривизна поля выражается в зависимости расстояния от центра голограммы до плоскости восстановленного изображения от координат объекта, а, следовательно, от координат изображения. Формируемое при этом пятно рассеяния остается неискаженным и полностью соответствует картине Эйри.

Последнее слагаемое в (3.40) описывает aberrацию дисторсии. Ее коэффициенты описываются выражениями (3.45). Физически она выражается в искривлении восстановленного изображения объекта.

Из анализа выражений, описывающих aberrации голограмм

точечного объекта, можно сделать следующие основные выводы об их отображающих свойствах.

1. Изображающие свойства голограмм полностью описываются длинами волн записи и восстановления и положениями опорного, объектного и восстанавливающего точечных источников.
2. Голограмма, восстанавливаемая на длине волны ее записи, является безaberrационным оптическим элементом для двух пар сопряженных точек, совпадающих по положению с опорным и объектным восстанавливающими источниками, использовавшимися при записи голограммы.

Основы голографии

3. Дисторсия голограмм, в отличие от тонких линз, тождественно не равна 0. Это объясняется отсутствием у голограммы, в общем случае, оптической оси. В частном случае осевой голограммы, т.е. при $x_O/z_O = x_R/z_R$ и $y_O/z_O = y_R/z_R$, дисторсия голограммы, так же, как и дисторсия тонкой линзы, тождественно равна 0.

4. Голограмма обладает много большими, по сравнению с тонкой линзой, хроматическими аберрациями. Об этом свидетельствует присутствующий во всех вышеприведенных выражениях коэффициент μ .

Приведенные выше основные особенности отображающих свойств голограмм обуславливают преимущественное использование элементов голограммной оптики в оптических системах, предназначенных для использования в монохроматическом излучении при малых полях зрения.

Лекция 18.

Факторы, ограничивающие качество изображения, формируемого методом голографии

Качество изображений, восстанавливаемых с помощью голограмм, помимо aberrаций, определяется еще рядом существенных факторов. Среди них наиболее значимыми являются: интермодуляционные помехи, шумы регистрирующей среды и помехи, обусловленные нелинейностью ее характеристической кривой, деформация восстанавливающего и восстановленного волновых фронтов, обусловленная фазовыми неоднородностями и деформацией регистрирующей среды, помехи, обусловленные дискретным характером регистрации голографического поля.

18.1. Интермодуляционные помехи

Интермодуляционные помехи обусловлены взаимной интерференцией точек объекта между собой. Природу этих помех и метод устранения их влияния на восстановленное изображение мы рассматривали ранее при анализе внеосевой схемы записи голограмм. Напомним вкратце суть вопроса. Из основного уравнения голографии следует, что комплексная амплитуда поля, восстанавливаемая с помощью линейно зарегистрированной голограммы в плоскости самой голограммы, сразу за ней, может быть описано с помощью выражения, содержащего четыре слагаемых. Из них два последних слагаемых соответствуют волновым фронтам, которые расходятся от мнимого изображения объекта или сходятся к его действительному изображению, в то время как два первых слагаемых не несут никакой полезной информации об объекте. Они описывают волны, распространяющиеся по направлению распространения восстанавливающей волны в нулевом порядке дифракции голограммы. Ширина полосы занимаемых этими волнами пространственных частот равна удвоенной полосе пространственных частот объектной и опорной волн, соответственно. Волны, описываемые двумя первыми слагаемыми основного уравнения голографии, как раз и являются источниками интермодуляционных помех. Отметим, что в осевой схеме голографии Габора волны, соответствующие 0 и ± 1 порядкам дифракции голограммы, распространяются в одном и том же направлении, что и восстанавливающая волна, что приводит к невозможности устранения

интермодуляционных помех из восстановленного изображения объекта. Условие пространственного разделения порядков дифракции голограммы одновременно и является условием отсутствия интермодуляционных помех. Ранее мы с вами для наиболее распространенного случая плоских опорной и восстанавливающей волн сформулировали и обосновали условие возможности пространственного разделения дифракционных порядков голограммы. Оно сводилось к необходимости выбора таких углов падения опорной и восстанавливающей волн, при которых величины углов, составляемых направлениями распространения объектной и каждой из вышеназванных волн, превышали бы полуторный угловой размер объекта. Мы не зря говорим здесь о возможности разделения, поскольку сам факт пространственного разделения порядков дифракции голограммы в той или иной плоскости локализации восстановленного изображения определяется еще и расстоянием от голограммы до восстановленного изображения.

18.2. Шумы регистрирующей среды

Прежде чем говорить о шумах регистрирующей среды уточним само понятие шума и его отличия от помехи. Помехи, в частности интермодуляционные, о которых мы говорили выше, детерминированы. Распределение их комплексной амплитуды определяется структурами опорного и объектного пучков лучей и не зависит от конкретной реализации голограммы. Т.е. если вы в одной и той же схеме, на одной и той же регистрирующей среде при соблюдении одних и тех же технологических режимов зарегистрируете несколько голограмм, то все они будут характеризоваться одним и тем же распределением интермодуляционных помех в плоскости восстановленного изображения. Иное дело шум. Он имеет случайную природу и может быть описан лишь статистически. В частности, каждая реализация голограммы будет описываться своим, характерным только для нее, распределением шума.

Одним из наиболее значимых источников шумов в восстановленных изображениях является структура регистрирующей среды. Из всех применяемых в настоящее время голографических регистрирующих сред наиболее шумящими являются галогенидосеребряные регистрирующие среды, обрабатываемые без отбеливания, либо с отбеливанием без последующего фиксирования отбеленных голограмм. Голограммная структура в обработанных таким образом фотоматериалах состоит либо из зерен металлического серебра в амплитудных голограммах, полученных без отбеливания, либо из

Основы голографии

зерен солей серебра в голограммах, полученных методом отбеливания, располагаемых в желатиновой матрице. Наличие этих зерен и их случайное расположение в желатиновой матрице обуславливают случайные вариации коэффициента пропускания голограммы и, как следствие этого, рассеяние проходящего через голограмму излучения. Значительно меньшим уровнем светорассеяния, а, следовательно, и шума, характеризуются галогенидосеребряные регистрирующие среды, прошедшие фотохимическую обработку по методу "желатины, сенсibilизированной серебром". При обработке по этому методу зерна солей серебра, полученных при отбеливании голограмм, удаляются из желатиновой матрицы в процессе второго фиксирования отбеленного фотослоя, образующиеся в желатиновой матрице пустоты "схлопываются" при сушке слоя. Сама же голограммная структура сохраняется в среде в виде вариаций показателя преломления желатиновой матрицы, возникающих в процессе проявления и отбеливания фотослоя вблизи зерен металлического серебра. Еще меньшим уровнем шума обладают голограммы, получаемые на неструктурированных регистрирующих средах, таких как бихромированная желатина и фотополимерные материалы. Спектр пространственных частот шума регистрирующих сред обычно простирается вплоть до граничных значений регистрируемых пространственных частот, что не позволяет осуществлять пространственное разделение шума и восстановленного изображения. Следовательно, единственным эффективным способом подавления шума в восстановленном изображении является грамотный выбор регистрирующей среды. Помимо самой среды, источниками шума, накладывающегося на восстановленное изображение, могут являться царапины, загрязнения или дефекты подложки голограммы. Шумы, обусловленные этими причинами, как правило, сосредоточены в области низких пространственных частот.

18.3. Помехи, обусловленные нелинейностью характеристической кривой регистрирующей среды

При рассмотрении свойств голограмм мы обычно постулировали линейность зависимости коэффициента пропускания прошедшей фотохимическую обработку регистрирующей среды от интенсивности зарегистрированного голографического поля. В действительности, характеристическая кривая реальных регистрирующих сред, описывающая зависимость их амплитудного коэффициента пропускания от экспозиции, линейна лишь на небольшом участке в области средних экспозиций и существенно нелинейна в областях малых и больших экспозиций. Типичный вид характеристической кривой представлен на рис. 4.2.

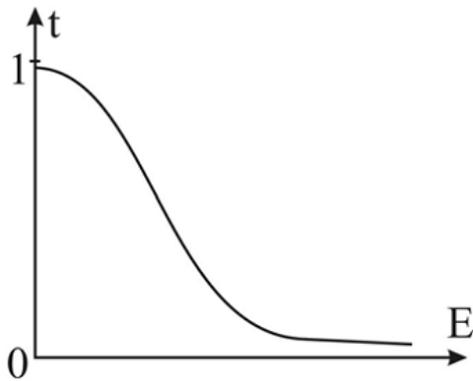


Рис. 4.1. Типичный вид характеристической кривой негативной регистрирующей среды

Такая кривая может быть достаточно точно описана полиномом третьей степени.

$$t = c_0 + c_1 E + c_2 E^2 + c_3 E^3. \quad (4.1)$$

Первый член в правой части (4.1) не зависит от экспозиции, второй член соответствует линейному соотношению между пропусканием и экспозицией. Третий и четвертый члены характеризуют нелинейность записи. Для обычных областей экспозиций наиболее важны нелинейные эффекты, обусловленные квадратичным членом. Рассмотрим его вклад в структуру изображения, восстанавливаемого с помощью голограммы. Для этого выделим из коэффициента пропускания голограммы t его

Основы голографии

составляющую t_2 , зависящую от квадрата экспозиции:

$$t_2 = c_2 E^2 . \quad (4.2)$$

Отметим, что вклад в восстановленное с помощью голограммы изображение второго члена (4.1), линейно зависящего от экспозиции, мы уже рассматривали ранее при анализе основного уравнения голографии. Предположим, что экспонирование голограммы осуществлялось в течение времени τ в голографическом поле, сформированном с помощью объектной и опорной волн с комплексными амплитудами на поверхности регистрирующей среды A и R , соответственно. Экспозиция, полученная регистрирующей средой при этих условиях, может быть описана с помощью следующего выражения:

$$E = kI\tau = k(AA^* + RR^* + AR^* + A^*R)\tau , \quad (4.3)$$

где k – коэффициент пропорциональности, I – интенсивность

Основы голографии

голографического поля. Подставляя (4.3) в (4.2) получаем:

$$t_2 = c_2 k^2 (AA^* + RR^* + AR^* + A^* R)^2 \tau^2. \quad (4.4)$$

Для упрощения анализа предположим, что опорная волна не модулирована, т.е. произведение RR^* мы можем считать постоянным в пределах апертуры голограммы. Отделим теперь от (4.4) все члены, пропорциональные первым степеням каких – либо членов в равенстве (4.3), т.е. члены, описывающие уже известные нам 0 и ± 1 порядки дифракции голограммы, и запишем.

$$t_2 = \text{линейные члены} + c_2 k^2 \tau^2 [(AA^*)^2 + A^2 R^{*2} + A^{*2} R^2 + 2AA^* AR^* + 2AA^* A^* R]. \quad (4.5)$$

Кубический член в (4.1) приводит к появлению членов функции амплитудного пропускания голограммы, часть из которых имеет такую же форму, как и слагаемые, входящие в (4.5). При количественных расчетах вклад этих слагаемых следовало бы включить в квадратичный нелинейный эффект. Однако мы будем проводить лишь качественный анализ и этой добавкой пренебрежем. Предположим, что в процессе восстановления голограмма освещается исходной опорной волной R , и рассмотрим лишь ту часть дифрагированного на голограмме света, которая обусловлена членом пропускания t_2 (линейными членами t_2 мы при этом пренебрежем). При этом будем считать опорную волну плоской, аксиально падающей волной единичной амплитуды, т.е. положим $R=1$.

Из (4.5) следует, что квадратичные члены функции пропускания голограммы приводят к появлению в дифрагированном на голограмме излучении пяти дополнительных волн. Кратко рассмотрим их. Для этого перепишем (4.5) в приближенном виде, описывающем комплексные амплитуды восстановленных волн с точностью до постоянной,

$$R[(AA^*)^2 + A^2 R^{*2} + A^{*2} R^2 + 2AA^* AR^* + 2AA^* A^* R] = \\ = (AA^*)^2 + A^2 + A^{*2} + 2AA^* A + 2AA^* A^*. \quad (4.6)$$

Основы голографии

Для оценки ширины пространственного спектра волны, соответствующей тому или иному слагаемому (4.6), необходимо найти Фурье-образ этого слагаемого. При этом необходимо иметь в виду, что в частотном пространстве произведению двух функций соответствует свертка их Фурье - образов, а произведению функции на такую же, но комплексно сопряженную функцию соответствует автокорреляция Фурье - образов этой функции.

Тогда оказывается, что спектр AA^* имеет в два раза более широкую полосу частот, чем спектр объектной волны A , а квадрат этого произведения - $(AA^*)^2$, занимает в 4 раза более широкую полосу пространственных частот, чем спектр объектной волны, располагающейся симметрично относительно нулевой пространственной частоты $\xi=0$, соответствующей частоте опорной волны. Таким образом, нелинейная запись голограмм приводит к двукратному расширению спектра интермодуляционных помех, о которых мы говорили выше. Это, в свою очередь, приводит к двукратному, по сравнению со случаем линейной записи, увеличению минимального угла падения опорной волны, обеспечивающего пространственное разделение нулевого и первого порядков дифракции голограммы.

Второе слагаемое в правой части (4.6) представляет собой квадрат комплексной амплитуды опорной волны. Оно приводит к формированию в восстановленном поле искаженного абберациями изображения объекта, располагаемого на вдвое меньшем расстоянии от голограммы, чем основное изображение. При этом прямая, соединяющая центр голограммы с центром этого искаженного изображения, составляет с направлением распространения опорной волны угол, вдвое больший, чем угол между направлениями распространения опорной и объектной волн при регистрации голограммы. Спектр этого изображения в два раза шире спектра объектной волны, а его центр располагается на удвоенной средней пространственной частоте

Основы голографии

объектной волны. Аналогичным образом третье слагаемое (4.6) можно рассматривать как искаженное абберациями действительное изображение объекта, восстановленного на вдвое меньшем расстоянии от голограммы, по сравнению с расстоянием от голограммы до объекта при записи голограммы.

Четвертое слагаемое (4.6) описывает волну, распространяющуюся в направлении основного восстановленного изображения объекта и накладывающуюся на него. Ее пространственный спектр в три раза превышает спектр объекта, а его середина совпадает с его средней пространственной частотой. В случае диффузно рассеивающего объекта он приводит к появлению гало – ореола вокруг восстановленного изображения. В случае зеркального объекта он приводит к появлению ложных изображений. При этом наиболее заметны ложные изображения крупных деталей объекта, характеризующихся относительно высокой интенсивностью излучения в узкой полосе пространственных частот.

Пятое слагаемое (4.6) описывает аналогичную помеху, локализуемую вокруг сопряженного изображения.

Мы рассмотрели только квадратичные нелинейные члены. Аналогичным образом можно проанализировать нелинейность третьего и более высоких порядков. Она также приводит к формированию ложных изображений и помех, в том числе и тех, которые локализируются вокруг основных восстановленных изображений.

Единственный способ борьбы с помехами, обусловленными нелинейностью характеристической кривой регистрирующей среды, состоит в выборе условий экспонирования среды, обеспечивающих регистрацию голограммы в пределах линейного участка характеристической кривой.

18.4. Влияние неравномерности толщины регистрирующей среды на качество восстановленного изображения

Неравномерность толщины регистрирующей среды по апертуре голограммы приводит к фазовым ошибкам опорной и объектной волн, интерферирующих в толще регистрирующей среды. При этом, поскольку изменения толщины регистрирующей среды обычно плавные, то для случаев регистрации низкочастотных тонкослойных голограмм мы можем считать фазовые набеги, обусловленные вариациями толщины среды, одинаковыми как для опорной, так и для объектной волны. Опишем изменение толщины регистрирующей среды функцией $z(x, y)$. Тогда фазовый набег $\varphi(x, y)$, вносимый регистрирующей средой в проходящую через нее волну, может быть описан следующим образом:

$$\varphi(x, y) = k[z_0 + (n - 1)z(x, y)], \quad (4.7)$$

где k – волновое число, z_0 – максимальная толщина регистрирующей среды, n – показатель преломления регистрирующей среды. Наличие вариаций толщины регистрирующей среды приведет к появлению фазовых добавок в выражения, описывающие комплексные амплитуды опорной и объектной волн, формирующих голографическое поле внутри объема регистрирующей среды. С учетом этих добавок комплексные амплитуды объектной и опорной волн будут иметь вид:

$$O'(x, y) = O(x, y) \exp[i\varphi(x, y)] \quad \text{и} \quad (4.8)$$

$$R'(x, y) = R(x, y) \exp[i\varphi(x, y)] \quad , \quad (4.9)$$

соответственно. Воспользуемся (4.8) и (4.9) и получим выражение, описывающее распределение интенсивности в голографическом поле,

Основы голографии

формируемом объектной и опорной волнами внутри регистрирующей среды $I(x,y)$

$$I(x,y) = |O(x,y) + R(x,y)|^2 = |O(x,y)|^2 + |R(x,y)|^2 + O(x,y)R^*(x,y) + O^*(x,y)R(x,y). \quad (4.10)$$

Выражение (4.10) не содержит $\varphi(x,y)$, что свидетельствует об отсутствии влияния неоднородности толщины регистрирующей среды на структуру регистрируемой голограммы. Будем считать регистрирующую среду линейной. Тогда, как известно, с точностью до коэффициента пропорциональности k мы можем описать комплексную амплитуду волны, восстановленной в -1 порядке дифракции, в виде произведения $C'(x,y)O(x,y)R^*(x,y)$, где $C'(x,y) = C(x,y)\exp[i\varphi(x,y)]$ – комплексная амплитуда восстанавливающей волны в объеме голограммы при ее восстановлении. Если в качестве восстанавливающей волны используется волна, использовавшаяся при записи голограммы в качестве опорной волны, то комплексная амплитуда восстановленной волны будет иметь вид $O(x,y)\exp[i\varphi(x,y)]$. Таким образом, в

рассмотренном нами случае тонкослойной низкочастотной голограммы вариации толщины регистрирующей среды не приводят к каким либо изменениям структуры регистрируемой голограммы, однако они обуславливают фазовые искажения в восстановленном изображении объекта из-за искажений восстанавливающей волны. Прделанный нами анализ справедлив для случая низкочастотных голограмм. При большой разнице углов падения на плоскость регистрации голограммы опорного и объектного пучков нарушается условие равенства фазовых задержек, вносимых регистрирующей средой в эти пучки, что приводит к искажениям как восстановленного изображения, так и структуры голограммы. Подробно этот случай мы рассмотрим в рамках курса "Голографические оптические элементы и устройства".

Лекция 19.

Системы голографической памяти

С каждым годом голография все больше входит в нашу повседневную практику. В этой заключительной лекции курса "Основы голографии и голограммной оптики" мы рассмотрим ряд практических применений голографии. Среди них: системы голографической памяти, голографическая интерферометрия с цифровым восстановлением и обработкой данных, голографические мультиплексоры для систем передачи данных со спектральным уплотнением, голографические метки идентификации, проекционные системы концентрации энергии солнца и системы кодирования информации.

19.1. Системы голографической памяти

Насущная необходимость увеличения объема хранимой информации и увеличения скорости адресного доступа к ней обуславливают перманентное совершенствование запоминающих устройств, в том числе и наиболее распространенных из них - оптических дисковых систем памяти. Так до середины 90-х годов на рынке безраздельно господствовали действующие в ближнем ИК диапазоне системы памяти на CD-RW дисках, позволяющие на диске диаметром 120 мм записать до 1 Гбайта информации. Им на смену пришли системы на дисках DVD-RAM и DVD-RW, работающие практически в том же диапазоне длин волн и позволяющие за счет более совершенных оптической системы и алгоритма обработки сигнала увеличить объем хранимой информации до 4 Гбайт. В настоящее время рынок завоевывает третья генерация оптических дисковых систем памяти, называемая DVD BR "blue – ray" (голубой луч). Она отличается от предыдущего поколения существенно более короткой длиной волны используемого излучения, составляющей 405 нм, что позволяет, сохранив все преимущества используемых в DVD системах алгоритмов обработки сигналов, довести объем информации, сохраняемой на одностороннем диске диаметром 120 мм, до 100 Гбайт. Конкурирующей с DVD BR разновидностью этой генерации систем оптической памяти является HD-DVD (DVD система повышенной плотности), позволяющая на одном оптическом диске стандартного размера хранить до 45 Гбайт. На этом практически исчерпываются все возможности повышения плотности записи информации в однослойном и многослойном рельефно-фазовом виде. Правда фирма Iomega недавно объявила о создании 800-гигабайтного формата оптических дисков, названного ими

Основы голографии

"Articulated Optical DVD", однако образцов использующих этот формат систем памяти пока представлено не было, не были опубликованы и принципы построения подобных систем.

Израильская компания Mempile разработала технологию TeraDisc, которая позволяет записать на диск, изготовленный из полимера, сходного с плексигласом, до терабайта данных. Об этом пишет издание Computerworld. На новых дисках можно будет хранить до 250 тысяч песен в формате mp3, 40 фильмов в HD- или 115 в DVD-качестве. Для работы с информацией в технологии TeraDisc используются два лазера, один из которых отслеживает местонахождение нужных данных, а другой производит чтение и запись. Ожидается, что дисководы TeraDisc появятся в продаже в ближайшее время и будут стоить от трех до четырех тысяч долларов. Они смогут читать 500-гигабайтные болванки. Болванка объемом в 700 гигабайт будет стоить 30 долларов. Еще через несколько лет разработчики обещают уместить на диске 5 терабайт данных. Разработать новые, гораздо более вместительные носители информации пытаются не только в Mempile. Японская компания TDK работает над технологией, которая позволит уместить на болванке до 200 гигабайт данных. Британская Plasmon активно разрабатывает технологию UDO, благодаря которой на диски помещается 60 гигабайт.

Дальнейший прогресс в разработке дисковых систем оптической памяти разработчики связывают с использованием метода оптической голографии, позволяющего использовать не только поверхность, но и весь объем диска. По теоретическим оценкам использование объемной голографической записи позволит довести плотность хранения информации на дисках до величины в 1 Тбит/см^3 .

Идея использования метода голографии для создания систем хранения информации была опубликована Питером ван Херденом, сотрудником компании Поляроид, еще в 1963 году. В качестве физического предела плотности хранения информации в голографических системах им была определена плотность в $1 \text{ бит}/\lambda^3$.

Голографические запоминающие устройства (ГЗУ) записывают и считывают информацию, представляемую в виде двумерных растров битов – масок с отверстиями, называемых страницами. Последние представляют собой высококачественные бинарные оптические изображения. Такой параллельный способ представления информации позволяет реализовывать высокие скорости при ее обмене, достигающие гигабита в секунду, при малом времени доступа, не превышающем 1 мс.

Основы голографии

Данные в ГЗУ, как правило, распределяются по объему регистрирующей среды, в то время, как в магнитных и простых оптических системах памяти они хранятся на поверхности носителя. Потенциально достижимая при использовании всего объема регистрирующей среды плотность хранения информации в ГЗУ превышает 10^{11} бит/см³.

Принцип действия ГЗУ состоит в следующем. Битовая информация, т.е. 0 или 1, с помощью пространственного модулятора света вводится в объектный пучок, падающий на поверхность регистрирующей среды, на которую направляется также и опорный пучок. Так регистрируется голограмма. При ее освещении параллельным пучком на обращении она формирует копию объектного пучка, распространяющегося в обратном направлении, и, тем самым, восстанавливает изображение зарегистрированной страницы информации. Воспроизведенная таким образом цифровая информация считывается с помощью матрицы фотоприемников, располагаемых в плоскости фокусировки страницы. При этом высокая угловая селективность объемных голограмм, достигающая для сред толщиной в 1 см величины порядка $0,001^0$, позволяет осуществлять их мультиплексирование, т.е. позволяет записывать множество независимо восстанавливаемых голограмм в одном объеме регистрирующей среды.

Смена страниц информации может осуществляться различными способами, например, изменением угла падения восстанавливающего пучка лучей или изменением положения этого пучка на поверхности регистрирующей среды. Комбинированные способы смены страниц, содержащие оба указанных варианта, обычно используются в системах памяти, выполняемых на основе трехмерных голографических дисков.

Основным сдерживающим фактором в развитии голографических систем памяти долгие годы являлось отсутствие регистрирующей среды, пригодной для мультиплексной объемной голографической записи.

Ситуация изменилась в конце девяностых годов с разработкой объемных фотополимерных сред для объемной голографической записи. Наиболее перспективной из них в настоящее время считается фотополимер Aprilis, разработанный Aprilis Inc., США. Фотополимерные регистрирующие среды фирмы Aprilis в настоящее время являются практически единственными средами, пригодными для промышленного внедрения в системах голографической памяти. Более 15 компаний в мире вели до последнего времени разработку

Основы голографии

голографических систем памяти, базирующихся на этих фотополимерах. На сегодняшний день победителями этой гонки стали американская компания InPhase Technologies и японская Optware. Так, американская компания InPhase Technologies сообщила о совместной разработке с японской фирмой Hitachi Maxell Ltd голографического диска, размером с диск DVD, способного хранить 300 Гбайт информации, что примерно в 60 раз превышает емкость диска DVD, и обладающего скоростью доступа к информации в 10 раз превышающую скорость, реализуемую в DVD системах. Скорость записи информации составляет 23 Мб в секунду при времени экспонирования 1 страницы 2,7 мс. При этом

плотность хранения информации достигает 515 Гбит на квадратный дюйм. Диск может содержать 320 страниц информации, содержащих 1,3 миллиона бит каждая, что соответствует 20 художественным фильмам, записанным на одном и том же диске. Информация записывается во всей толщине диска, составляющей примерно 1,5 мм, методом углового мультиплексирования. Шаг мультиплексирования составляет 0,067 градуса. Рабочая длина волны записи и считывания информации составляет 407 нм. Запись и считывание информации в системе осуществляются двумя разными лазерными чипами с одинаковыми характеристиками. Модуляция опорной волны в устройстве осуществляется с помощью мегапиксельного микроэлектромеханического зеркального устройства. InPhase Technologies обещает выпустить диски и соответствующие системы памяти в свободную продажу. По словам Ron Tarasoff, вице-президента фирмы Turner Entertainment – дистрибьютора голографических систем памяти, розничная цена дисков не будет превышать 100 долларов. Сначала предполагается выпуск в продажу дисков емкостью 400 Гб. В дальнейшем, в течение ближайших 5 лет, компания предполагает выпускать линейку систем голографической дисковой памяти с емкостями от 800 Гб до 1,6 Тб при скоростях доступа к информации 120 Мбит в секунду.

Рассмотрим поподробнее используемую в указанных системах дисковой памяти оригинальную коллинеарную голографическую технологию записи и считывания информации, основанную на использовании традиционного для систем оптической памяти предварительно отформатированного отражающего слоя и пространственного мультиплексирования информации.

Основы голографии

Предназначенные для использования в системе диски выполняются на прозрачной пластиковой подложке, на поверхность которой спирально нанесена дорожка пит, используемых для адресации точно так же, как это делается в обычных CD и DVD дисках. Поверхность дорожки пит покрыта отражающим слоем из алюминия. Поверх слоя алюминия на диск нанесен прозрачный слой полимера, полностью закрывающий питы и имеющий плоскую верхнюю поверхность. На нее нанесено многослойное диэлектрическое зеркало, отражающее излучение синезеленой области спектра и пропускающее излучение красной области спектра. На диэлектрическое покрытие нанесен тонкий защитный слой полимера. Поверх него нанесен слой голографической фотополимерной регистрирующей среды, защищенный сверху еще одним защитным полимерным слоем. Сервоинформация с диска считывается на отражение через регистрирующую среду и многослойное диэлектрическое зеркало с помощью сфокусированного излучения лазерного диода, излучающего в красной области спектра. Реализуемые

в системе принципы адресации и фокусировки оптической головки совершенно идентичны тем, что реализуются в CD системах, и поэтому не требуют каких-либо устройств дополнительной юстировки и защиты от вибраций.

Запись информации на диске осуществляется с помощью излучения лазерного диода, работающего в синей области спектра. Это излучение проходит через расширитель пучка, после чего направляется на работающий на отражение микрзеркальный пространственный модулятор света. В центральной области апертуры модулятора отображается фазовое изображение страницы записываемой информации, а на ее периферии – изображение кольцевой зоны радиальной дифракционной решетки, предназначенной для формирования опорного пучка лучей. При этом каждая из страниц информации отображается в виде совокупности субстраниц, пространственное расположение которых и объем хранящейся в них информации определяется аберрационными свойствами основного высокоапертурного объектива оптической головки. Сформированные с помощью пространственного модулятора света опорный и объектный пучки лучей пропускаются через четвертьволновую пластинку, изменяющую состояние их поляризации с линейной на круговую. Затем они проходят через основной объектив головки, отражаются от находящегося на диске многослойного диэлектрического зеркала и фокусируются в объеме регистрирующей среды. Там они формируют объемную пропускающую голограмму страницы информации.

Основы голографии

Считывание информации осуществляется аналогичным образом. Отличие процедур считывания и записи информации состоит лишь в том, что на стадии считывания информации на пространственном модуляторе света отображается лишь кольцеобразная радиальная дифракционная структура, формирующая восстанавливающий пучок лучей. Считывание восстановленных страниц информации осуществляется с помощью высокоскоростного CMOS приемника. При этом число пикселей фотоприемного устройства выбирается примерно в три раза большим, чем число пикселей пространственного модулятора света. Это позволяет существенно упростить алгоритм распознавания бит информации. Изменение пространственной частоты дифракционной структуры, формирующей опорный пучок лучей, позволяет осуществлять мультиплексирование регистрируемых страниц информации. При этом динамический диапазон используемой в системе регистрирующей среды позволяет осуществлять запись до 20 страниц информации на одном и том же ее участке.

Такой вариант технологии голографической записи информации назван разработчиками коллинеарной технологией, поскольку в ней все пучки лучей, используемые на стадиях адресации, записи и считывании

информации практически сосны и фокусируются в толщу диска одним и тем же объективом.

По заключению разработчиков, технология коллинеарной голографической записи информации будет принята в качестве основной при создании следующих поколений относительно простых и недорогих систем голографической дисковой памяти.

Лекция 20.**Голографическая интерференция с цифровым восстановлением и обработкой данных**

Традиционно при практической реализации методов голографической интерферометрии для регистрации голограмм использовались галогенидосеребряные фотоэмульсии и фототермопластические пленки. Получаемые на этих средах дифракционные структуры восстанавливались излучением лазера и сформированные с их помощью интерференционные картины оцифровывались и вводились в ЭВМ для последующей цифровой обработки. Современный уровень развития вычислительной техники и электронных средств регистрации изображения, таких как камеры с ПЗС матрицами, позволил отказаться от использования твердых носителей голограммной структуры и физического восстановления голограмм. Мощности персональных компьютеров, высокое разрешение и малый размер пикселей современных ПЗС матриц позволяют оцифровывать не только восстановленное изображение, но и саму голограмму и проводить в цифровом компьютерном пространстве все дальнейшие операции, в том числе восстановление голограмм, формирование интерференционных картин и их последующую обработку.

В настоящее время указанная техника применяется, в основном, при регистрации следующих двух основных типов голограмм: голограмм сфокусированного изображения и голограмм Френеля. Цифровое восстановление голограмм первого типа осуществляется в рамках Фурье преобразования. Восстановление голограмм второго типа требует для своей реализации математического моделирования процесса дифракции излучения на структуре голограммы. При этом ограничения, накладываемые размером и числом пикселей ПЗС матриц, на предельную величину регистрируемой пространственной частоты голографического поля обуславливают перспективность использования в этом процессе схем записи голограмм, характеризующихся малыми величинами несущей пространственной частоты, в том числе и осевой схемы Габора. При этом цифровое восстановление голограмм позволяет успешно разрешить проблему переналожения изображений, восстанавливаемых в различных порядках дифракции.

В частности, в литературе сообщается о применении осевой схемы Габора и цифровой регистрации голограмм для исследования вибраций и деформации металлических пластин методом двухэкспозиционной голографической интерферометрии. Запись голограмм здесь осуществлялась в излучении импульсного рубинового лазера с помощью ПЗС матрицы с размером пикселя 11 мкм и числом пикселей 512x512. В

работе P. Picart, J. Leval, D. Mounier, S. Gougeon. *Time-averaged digital holography*. *Optics Letters*, 2003, V.28, №20, P.1900-1902 сообщается об использовании осевой схемы записи голограмм сфокусированного изображения для исследования модовой структуры вибраций акустических преобразователей. Формирование голографического поля здесь осуществлялось в излучении He-Ne лазера а его регистрация проводилась с помощью 12 битовой цифровой ПЗС камеры с размером пикселя 4,65 мкм и общим числом пикселей 1024x1360. Такие параметры камеры позволяли использовать ее для регистрации голограммных структур с пространственными частотами до 70 мм^{-1} . Использование для практически аналогичных целей, т.е. для исследования вибраций и механических напряжений строительных конструкций внеосевых цифровых голограмм сфокусированного изображения описано, например, в работе J. Chambard, V. Chalvidan, X. Carniel, J. Pascal. *Pulsed TV-holography recording for vibration analysis applications*. *Optics and Lasers in Engineering*, 2002, V.38, P.131-143. В этой работе для записи голограмм использовалось излучение второй гармоники Nd:YAG импульсного лазера, работающего в режиме генерации парных импульсов и ПЗС камера Sony модель XC-8500CE с разрешением 580x786 пикселей. Отметим, что данная работа выполнялась в рамках международной научной программы BRITE-EURAM. Возможность совмещения в одном устройстве двух указанных выше схем цифровой записи голограмм рассмотрена в работе S.R. Guntaka, B. Bowe, V. Toal, S.Martin. *Holographic Optical Element for Combined Holographic and Digital Speckle Pattern Interferometry*. *SPIE Proc.*, 2003, V.4933, P.239-245, в которой предложен голограммный оптический элемент, позволяющий

Основы голографии

создать относительно простой универсальный интерферометр для голографической записи с электронной регистрацией, обладающий возможностью реализации обеих типов голограмм.

На базе рассматриваемой нами голографии с электронной регистрацией и цифровым восстановлением голограмм возникло даже новое направление в микроскопии, называемое цифровой голографической микроскопией. Ее суть состоит в использовании микроскопа и цифровой камеры для регистрации голограмм объектов микроскопического масштаба и в последующем цифровом восстановлении зарегистрированных изображений. Впервые эта техника была предложена в работах Haddad et al в 1992 году и получила

существенное дальнейшее развитие в настоящее время. Столь пристальный интерес исследователей к цифровой голографической микроскопии объясняется, прежде всего, ее широкими возможностями в части получения объемного изображения за одну экспозицию, что обуславливает перспективность использования данного метода при исследовании движущихся объектов. Кроме того метод позволяет одновременно получать информацию, как об амплитуде, так и о фазе исследуемого объекта.

С целью увеличения достигаемого разрешения в восстановленных методом цифровой голографической микроскопии изображениях в настоящее время стараются запись голограмм проводить в коротковолновых УФ диапазонах спектра. Так в работе Morlens A.S., Gauntier J. и др. сообщается о достижении ими 0,8 мкм разрешения в изображениях, восстановленных с помощью цифровых голограмм, зарегистрированных на длине волны 32 нм по осевой схеме Д. Габора с помощью микроскопа с увеличением 110^{\times} . Столь невысокое значение достигнутой ими разрешающей способности объясняется, в первую очередь, характерной для таких длин волн малой, равной 0,025, числовой апертурой используемой ими оптической системы. Исследователи из Института технической оптики (Штутгарт, Германия) Pedrini G., Zhang F. и Osten W. провели исследования, направленные на повышение разрешающей способности метода голографической микроскопии. В своей работе они пошли по пути поиска компромисса между стремлением минимизации рабочей длины волны и возможностью создания оптических систем с приемлемыми величинами числовой апертуры. В качестве рабочего они использовали излучение с длиной волны 193 нм. Свой выбор исследователи объясняют относительно широкой доступностью работающих на этой длине волны источников когерентного излучения – эксимерных лазеров и

Основы голографии

относительной легкостью изготовления предназначенных для работы на этой длине волны пропускающих и отражательных оптических элементов и систем. В созданном ими голографическом микроскопе использовался эксимерный лазер с длиной когерентности 100мкм, что позволило регистрировать голограммы областей объектов, площадью $0,5 \times 0,5 \text{ мм}^2$. Оптическая схема микроскопа позволяла исследовать как прозрачные, так и отражающие излучение объекты. При этом имелась возможность изменения состояния поляризации излучения, поступающего в объектное плечо микроскопа. В составе опорного плеча разработанного ими устройства имелась еще и регулируемая линия задержки, позволявшая изменять длину объектного плеча голографического микроскопа. Наблюдавшаяся нестабильность амплитудно – фазового распределения лазерного излучения от импульса к импульсу не позволяла использовать технику постоянного фазового сдвига и обусловила построение микроскопа по схеме с внеосевым опорным пучком. В качестве регистратора голографического поля разработчиками была выбрана ПЗС матрица PCO Sensicam em680,

Основы голографии

обладающая прекрасной светочувствительностью на рабочей длине волны. Ее квантовый выход составлял 20% на длине волны 193нм. Число элементов разрешения матрицы равно 1280 x 1024 при размере пикселя 8 x 8 мкм². Угол падения опорной волны был выбран исходя из необходимости пространственного разделения нулевого и плюс-минус первого порядков дифракции. Кривизна опорного волнового фронта выбиралась примерно равной кривизне фронта объектной волны, что обеспечивало регистрацию безлинзовых Фурье – голограмм. Выбор именно такой схемы регистрации голограмм, по мнению разработчиков микроскопа, позволил существенно сократить число ложных восстановленных изображений, обусловленных дискретным характером записи голографического поля. Равная 0,34 числовая апертура объектного пучка обеспечивала в случае полной компенсации aberrаций микроскопа его разрешающую способность в 0,35 мкм. Компенсация aberrаций представляла собой сложную научно-техническую задачу. Дело в том, что низкое качество пучка, излучаемого эксимерным лазером, обуславливало необходимость использования пинхола – маленькой диафрагмы, диаметром 0,5 мкм, для формирования расходящегося опорного пучка. При этом высокая плотность энергии опорного пучка, составляющая в плоскости диафрагмы около 4 Дж/мм², приводила к прогоранию диафрагмы и, как следствие этого, к нарушению структуры пучка. Использование опорного пучка с недетерминированным амплитудно-фазовым распределением, в свою очередь, приводило к появлению aberrаций в восстанавливаемом изображении, обусловленных отличием параметров схемы записи голограммы от условий ее восстановления. Эти aberrации устранялись в виртуальном компьютерном пространстве в процессе цифрового восстановления и обработки изображений. Для этого, сначала записывалась голограмма покровного стекла без исследуемого образца, и с ее помощью считывалось распределение фазовых задержек в оптической системе микроскопа, соответствующих aberrациям его оптических элементов. Затем на основе известного распределения aberrаций микроскопа формировалось математическое описание корректора, которое использовалось на стадии восстановления изображений исследуемого. Испытания микроскопа, проведенные в ходе регистрации и восстановления изображений субмикронной структуры процессора компьютера, показали достижение реальной разрешающей способности в 0,5 мкм в пределах пространства изображений 0,4 x 0,4 мм².

Основы голографии

К ограничениям рассмотренного метода цифровой голографической микроскопии следует отнести дифракционный предел разрешения, относительно низкую разрешающую способность современных ПЗС матриц и пространственное переналожение изображений, восстанавливаемых в различных порядках дифракции. Причем, в ряде случаев эти ограничения имеют различную направленность. Так, стремление уменьшения дифракционного предела разрешения обуславливает необходимость уменьшения расстояния между объектом и плоскостью регистрации голограммы. Последнее сопровождается расширением интервала регистрируемых пространственных частот и требованием увеличения необходимого разрешения ПЗС матрицы. Устранение переналожения изображений в процессе цифрового восстановления голограмм приводит, как правило, к потере части фазовой информации о восстанавливаемом изображении и поэтому не всегда приемлемо. Все вышеизложенное свидетельствует о чрезвычайной важности оптимального выбора геометрических параметров схем записи в методе цифровой голографической микроскопии.

Подводя итог рассмотрению метода голографии с электронной регистрацией и цифровым восстановлением голограмм можно сделать вывод о том, что синтез оптической голографии с современными цифровыми технологиями существенно упростил технологию голографического процесса и, тем самым, обеспечил возможность более широкого внедрения метода голографии в практику.

Лекция 21.

Голографические мультиплекторы для систем передачи данных со спектральным уплотнением

Потребность общества в передаче все больших объемов информации по волоконно-оптическим линиям связи приводит к необходимости постоянного повышения скорости передачи данных. Увеличение скорости не является единственно возможным путем повышения объемов передаваемой информации. Существует и другой, активно развиваемый в настоящее время, подход к этой проблеме, основанный на спектральном уплотнении передаваемой информации. Его суть сводится к созданию в волоконно-оптической линии связи нескольких эффективных каналов, каждый из которых работает на своей длине волны.

Реализация линий связи со спектральным уплотнением включает в себя спектральное мультиплексирование оптических каналов, осуществляемое путем сведения пучков излучения ряда пространственно разнесенных полупроводниковых лазеров,

Основы голографии

генерирующих на различных частотах и модулируемых независимым образом, в одно общее волокно, и демультимплексирование, проводимое на выходе из волокна и заключающееся в пространственном разнесении и соответствующей фокусировке излучения каждого из каналов.

В волоконно-оптических линиях связи со спектральным уплотнением информации голограммные элементы используются, как правило, лишь в устройствах пространственного совмещения и разнесения спектральных каналов. Все известные к настоящему времени схемы мультиплексоров и демультимплексоров, предназначенных для спектрального уплотнения информации, основаны на использовании голограммных (дифракционных) оптических элементов. По типу применяемых голограммных элементов описанные в литературе схемы мультиплексоров можно условно разделить на две основные группы.

Первую группу составляют устройства, использующие брэгговские голограммы с полным внутренним отражением. Характерными примерами таких устройств являются многокаскадные устройства, выполненные на базе планарных волноводов с голографическими системами ввода и вывода сигнала. В литературе описан демультимплексор, содержащий объемную брэгговскую отражательную решетку, записанную в толще оптического волокна так, чтобы ее страты располагаются перпендикулярно оси волокна. В другом устройстве решетка, вытравленная на боковой поверхности волокна, используется для вывода излучения из волокна с одновременным его спектральным демультимплексированием.

Вторую группу составляют работающие в свободном пространстве мультиплексоры на дискретных элементах. Характерным примером такого мультиплексора является устройство, предназначенное для совместной работы с многомодовым волокном и основанное на использовании плоской голографической отражательной рельефной решетки с частотой 1800 мм^{-1} для пространственного совмещения, либо разделения спектральных каналов. Применение плоской решетки предопределило необходимость дополнения устройства двумя Фурье-преобразующими объективами, устанавливаемыми на его входе и выходе.

Во всех рассмотренных типах спектральных мультиплексоров голограммные элементы используются, как правило, лишь в качестве спектрального селектора. Вместе с тем использование принципов голографии позволяет, в общем случае, создать интегральный элемент, объединяющий в себе спектральный селектор, Фурье-объектив и устройство ввода (вывода) излучения в волокно.

Лекция 22.

Голографические метки идентификации

Применение голограммных структур в системах идентификации является одним из наиболее значимых и известных применений голографии. Голограммы, наносимые на банкноты, паспорта и кредитные карты, как правило, сочетают в себе все наиболее значимые элементы (степени) защиты. Среди них: наличие кинетических эффектов, проявляющихся в изменении изображения при изменении угла наблюдения голограммы, формирование трехмерных изображений зарегистрированных объектов, радужная окраска изображений и т.д. Потребности в совершенствовании защитных технологий приводят не только к широкому использованию всех современных достижений голографии в производстве меток идентификации, но к стимулированию и постановке работ, направленных на разработку специальных типов голограмм, предназначенных исключительно для целей аутентификации документов и изделий. Среди таких специальных типов голограмм можно выделить защищенные торговыми марками Кинеграмму (Kinegram) и Экселграмму (Exelgram).

Кинеграмма была разработана специалистами швейцарской фирмы Landis & Gyg около двадцати лет назад. Это была первая голограмма, обладавшая кинетическим эффектом, позволявшим наблюдать изменяющееся двумерное изображение зарегистрированного объекта при изменении угла наблюдения голограммы. Отметим, что изображение, восстанавливаемое обычной физически зарегистрированной голограммой, трехмерное и, следовательно, реализуемая с ее помощью возможность наблюдения различных ракурсов объекта не может быть отнесена к кинетическому эффекту.

Экселграмма была разработана специалистами фирмы CSIRO. По своей сути, она представляет собой синтезированную на компьютере голограмму, отображенную на носителе с помощью электронно-лучевого генератора изображений. Такие голограммы обладают высокой разрешающей способностью и прекрасным качеством восстановленных изображений, но, к сожалению, характеризуются относительно высокой стоимостью изготовления оригиналов.

Основы голографии

Все перечисленные выше типы голограмм являются рельефно-фазовыми голограммами, что обеспечивает относительно низкую стоимость их массового тиражирования. Более высокая стоимость производства и тиражирования отражательных объемных голограмм, зарегистрированных по методу Ю.Н. Денисюка, уже перестает быть сдерживающим фактором их использования в идентификационных документах. В качестве доказательств этого положения можно привести Идентиграмы (Identigram) – объемные отражательные голограммы на фотополимерной пленке, клеиваемые в немецкие паспорта и голограммы "Правильного изображения" (True Image) – цветные объемные отражательные голограммы, разработанные японской фирмой Dai Nippon для защиты торговых марок.

Устойчивость защитного сегмента мирового голографического рынка во многом определяется стойкостью голографических меток идентификации, т.е. сложностью их подделки, и конкурентоспособностью голографических методов защиты документов и изделий, по сравнению с другими современными защитными технологиями.

Помимо голограмм в защитных технологиях широкое применение находят специальные бумага, чернила и техника печати. Защита бумаги, как правило, осуществляется путем введения в нее водяных знаков, цветных волокон и других включений. Средства защиты, вводимые в чернила, обычно заключаются в изменении цвета при изменении угла наблюдения рисунка или при изменении спектрального состава используемого освещения. Защитные возможности печати обычно состоят в использовании микропечати, невозпроизводимой обычными средствами тиражирования документов, и в скрытой фоновой печати, позволяющей скрыть текстовую информацию в рисунке, фотографии и т.д. Голографические технологии долгие годы успешно конкурировали с указанными способами защиты документов и изделий.

Основы голографии

В настоящее же время ситуация сильно изменилась. Во-первых, существенно выросли знания, умение и возможности нелегальных производителей голограмм. Во-вторых, появились новые современные технологии, ничем не уступающие, а порой даже превосходящие по стойкости, используемые на сегодняшний день голографические способы идентификации. Среди них наиболее эффективными являются: внедрение в документ или на поверхность товара RFID чипов, несущих в себе большой объем закодированной информации, и "track & trace" технология, характеризующаяся отслеживанием всех перемещений товара или документа в процессе их реализации и использования. Основные преимущества этих технологий и их привлекательность для потребителей состоят в том, что они: позволяют наносить на документ или изделие индивидуальную информацию о серийном номере изделия, дате и месте его изготовления и т.д.; допускают считывание хранящейся на метках информации без использования каких – либо специальных устройств, за исключением обычных магнитных и оптических считывателей штрихкода; обеспечивают автоматическую проверку подлинности.

Для успешной конкуренции с этими технологиями голографических меток идентификации разработчики защитных голограмм должны обратить серьезное внимание на решение следующих задач:

1. Введение в голографические метки персонифицированной информации, т.е. кодов изделий, или их серийных, заводских номеров. Пусть даже объем этой информации будет несколько ниже, чем у чипов.
2. Стандартизацию и унификацию защитных голограмм, выпускаемых различными изготовителями.
3. Разработку и выпуск единых автоматизированных стандартизованных устройств проверки подлинности голографических меток, выпускаемых различными потребителями.
4. Обеспечение сравнимости стоимости изготовления поддельных меток идентификации с величиной ущерба, наносимого при этом легальным производителям защищаемой продукции.

Основы голографии

Решение этих основных задач позволит голографическим меткам идентификации вновь занять ключевые позиции на рынке защитных технологий. Основной же проблемой голографических меток идентификации, как и, в целом, всех защитных технологий, является обеспечение разумного компромисса между обеспечиваемой степенью защиты документов или изделий и стоимостью изготовления меток.

Лекция 23.

Солнечные энергетические установки с голограммными оптическими элементами в системах концентрации энергии

Одна из основных проблем использования солнечной энергии заключена в высокой стоимости кристаллического кремния, составляющего основу солнечных энергетических установок. С целью уменьшения размеров используемых полупроводниковых панелей в таких установках обычно применяются линзовые или зеркальные концентраторы энергии, позволяющие существенно уменьшить площадь кремниевых панелей без снижения объемов производимой электрической энергии. Такие концентраторы громоздки и их использование существенно ухудшает внешний вид крыш домов, использующих солнечные энергетические установки. С целью устранения этого недостатка солнечных энергетических установок американская компания Prism Solar Technologies, NY разработала и испытала солнечную батарею с голографическим концентратором, стоимость которой составляет лишь 25% от стоимости традиционных солнечных энергетических установок, что позволяет ей успешно конкурировать с электрическими генераторами на жидком топливе.

Помимо этого, голографическая технология позволила выполнить энергетическую установку в виде сверкающих на солнце плоских радужных панелей легко монтируемых на крышах домов или на их окнах. По сообщению президента компании Rick Lewandowski созданная ими энергетическая установка при той же электрической мощности, что и у традиционных установок на зеркальных или линзовых концентраторах, использует кремниевые панели, площадь которых составляет от 25 до 85% площади панелей в традиционных установках. Такой эффект достигается за счет использования голографических концентраторов солнечной энергии, фокусирующих излучение в несколько достаточно узких фокальных линий. Указанные голографические концентраторы представляют собой голограммные цилиндрические линзы, нанесенные на плоские стеклянные подложки. При этом голограммная структура одновременно выполняет функцию устройства ввода излучения в своеобразный планарный волновод, функции которого выполняет подложка голограммы. Фокусировка

Основы голографии

солнечного излучения в устройстве осуществляется в толще стеклянной подложки голограммы при многократных переотражениях света от ее поверхностей. За счет использования голографических концентраторов

солнечной энергии разработчикам удалось снизить стоимость 1 Вт получаемой электроэнергии с 4,5\$, характерных для традиционных солнечных панелей, до 1,5\$. К концу 2006 года компания планирует

привлечь еще 6млн долларов инвестиций и организовать массовое производство и продажу солнечных установок с голографическими концентраторами энергии. Ориентировочная розничная цена установок будет составлять 2,4 доллара за 1 Вт мощности.

Говоря о перспективности использования голограммных оптических элементов в концентраторах солнечной энергии, следует иметь в виду следующее. Действительно, голографические системы характеризуются меньшей, примерно равной 10, концентрирующей способностью, по сравнению с линзовыми или зеркальными системами, у которых данный параметр достигает 100. При этом этот недостаток голографических систем одновременно является и их достоинством, поскольку позволяет отказаться от включения в состав солнечной энергетической установки систем охлаждения кремниевых светочувствительных панелей, что также снижает стоимость установок.

Кроме того, меньшая чувствительность голографических концентраторов к положению Солнца относительно горизонта позволяет отказаться и от систем слежения за Солнцем, являющихся непременным атрибутом традиционных солнечных энергетических установок. Немаловажное преимущество голограммных оптических элементов, используемых в качестве концентраторов солнечной энергии, состоит в возможности использования их спектральной селективности для выделения наиболее эффективного для преобразования в электрическую энергию диапазона солнечного излучения и отсеечения теплового излучения, обуславливающего необходимость отвода тепла от кремниевых фотоприемных панелей.

Следующий этап модернизации солнечной установки с голографическим концентратором, по сообщению Lewandowski, будет заключаться в переходе на концентраторы, состоящие из двух голограммных структур, наносимых на каждую из сторон подложки, что позволит еще больше снизить стоимость установки и довести ее до 1,5 долларов за 1 Вт в течение нескольких ближайших лет.

Лекция 24.

Голографические проекционные системы

Сообщается о выполненной недавно созданной английской компанией Light Blue Optics в сотрудничестве с Cambridge University Engineering Department, разработке компактного голографического проектора двумерных изображений, предназначенного для встраивания в ноутбуки, мобильные телефоны и прочие портативные цифровые устройства. По информации специалистов компании им удалось преодолеть основные недостатки голографических проекционных систем, заключающиеся в их чрезвычайной сложности и, соответственно, высокой стоимости. Разработанное ими устройство состоит из микропроцессора, выполненного в виде одной микросхемы, несущей на своей поверхности работающий на отражение малогабаритный жидкокристаллический пространственный модулятор света с электронной адресацией и полупроводникового лазера. С помощью микропроцессора по поступающему на его вход стандартному композитному видеосигналу осуществляется расчет и отображение синтезированной голограммы, восстановление которой лазерным излучением приводит к формированию требуемого двумерного изображения на любом внешнем экране. Проекционный объектив в устройстве не используется, поскольку синтез голограммы осуществляется с учетом расстояния от проектора до экрана. Быстродействие процессора и пространственного модулятора света позволяет отображать не только статические изображения, но и видеофильмы со стандартной частотой смены кадров. Действующий макет описанного голографического проектора был недавно продемонстрирован в St John's Innovation Centre, Cambridge, UK. По оценкам присутствующих на демонстрации проектора экспертов формируемые с помощью проектора изображения отличались высокой четкостью и яркостью и вполне были пригодны для практического использования. Промышленный выпуск таких проекторов Light Blue Optics планирует начать в ближайшие годы. В последующем предполагается на основе этого устройства разработка и выпуск проекционных систем для трехмерного цветного телевидения.

Основы голографии

Применение голограммных оптических элементов в бытовой проекционной технике позволило фирме CLARO создать проекционную телевизионную систему с экраном, выполненным в виде фазового голографического рассеивателя. На вид рассеиватель представляет собой прозрачную стеклянную пластину размером 1,5 x 1,0 м² и толщиной 10 мм. В ее центре, в зоне с размерами 814 x 610 мм², нанесена фазовая объемная голограмма. Наблюдаемое на голографическом экране цветное двумерное изображение проецируется на него с помощью обычного цифрового проектора, устанавливаемого на расстоянии 1,5 метров за экраном. Взаимное расположение экрана и проектора выбрано таким, чтобы за счет угловой селективности голограммной структуры экрана обеспечить отсутствие в поле зрения зрителей нежелательных бликов и отражений. Голографическая телевизионная система Claro Holographic TV, состоящая из голографического экрана, проектора и специально разработанной акустической системы с прозрачными стеклянными рупорами коммерчески доступна и уже появилась на рынке. Ее цена составляет 25000 английских фунтов.

Другой, трехмерный, вариант голографической телевизионной системы разрабатывается в настоящее время немецкой компанией SeeReal. На выставке SID 2007 специалисты этой компании продемонстрировали прототип трехмерной голографической телевизионной системы с экраном 20 дюймов. Выбранные ими технические решения позволили обойти принципиальные трудности, обычно сопровождающие создателей голографических трехмерных дисплеев, а именно: чрезвычайно большое требуемое число и малый размер обрабатываемых пикселей, формирующих изображение голограммы, и большой объем требуемых вычислений. Специалисты

компании SeeReal пошли по пути создания системы с вынесенным зрачком, что позволило обойти указанные выше проблемы за счет сведения до минимума числа возможных зрителей. В их устройстве, предназначенном для одного зрителя, каждая точка наблюдаемого трехмерного пространства формируется субголограммой, содержащей 8x8 пикселей. Чрезвычайно малый выходной зрачок устройства обусловил необходимость введения в его состав специальной сервисной системы, отслеживающей положение глаз зрителя. По прогнозам представителя компании Mark Thorsen в ближайшие годы они расширят число возможных зрителей их голографической системы и выйдут с нею на рынок. Необходимые для этого технические решения ими уже найдены.

Лекция 25.

Голографические системы кодирования информации

В настоящее время большинство голографических систем ограничения доступа и идентификации строится на основе рельефно-фазовых голограмм, изготавливаемых методом тиснения. При этом интенсивное развитие цифровых технологий сканирования и печати обуславливает неуклонное снижение стойкости этих голограмм к копированию и фальсификации. Такое положение дел приводит к необходимости использования в рассматриваемых системах других типов голограмм, обеспечивающих принципиально более высокую степень защиты от подделок. Например, ими могут быть голограммы с протяженным опорным либо объектным источником, т.е. голограммы, получаемые с помощью пучков со случайным распределением фазы. В

частности, В. Javidi и J.L. Horner предложили для устройств ограничения доступа и идентификации оптические системы, в которых кодирование информации осуществляется белым шумом с помощью двух матовых стекол, одно из которых устанавливается в плоскости входного транспаранта, а другое - в его Фурье плоскости. Процедура расшифровки закодированных таким способом изображений невозможна без использования обоих матовых стекол, использовавшихся на стадии кодирования. Закодированная подобным способом информация может быть зарегистрирована любым подходящим способом, в том числе и в виде голограммы.

В рассмотренных системах кодировке подвергается непосредственно само изображение шифруемого объекта. J.F. Neaue с коллегами предложил кодировать вместо самого изображения его голограмму, подобно тому, как это делается в голографических системах памяти с фазовым мультиплексированием информации. Суть его идеи заключается в фазовом кодировании опорного пучка лучей, осуществляемом с помощью комбинации случайного фазового рассеивателя и транспарантов с функциями пропускания, описываемыми ортогональными кодами. Такой способ шифрования изображений сохраняет высокую стойкость, характерную для случайного фазового кодирования информации, а также обеспечивает свойственные ортогональным кодам низкий уровень перекрестных помех и малый минимально необходимый объем библиотеки ключей. В

Основы голографии

качестве основного недостатка способа, существенно ограничивающего возможность его практической реализации, можно отметить чрезвычайно высокие, составляющие единицы микрометров, требования к взаимному расположению голограммы и матовых стекол на стадии восстановления голограммы.

Указанная проблема решена исследователями из Южной Кореи J. Kim (Chungbuk National University), J. Choi, J. An, N. Kim и K. Lee, (Sunchon National University). Им впервые удалось создать работоспособное, удобное в эксплуатации цифровое устройство ограничения доступа и идентификации, сочетающее в себе электронные биометрические технологии с полностью цифровым методом голографического хранения данных, закодированных путем использования опорного пучка со случайным распределением фазы.

Устройство представляет собой карту, на которой в цифровой форме хранятся все данные о человеке: документы, фотография, отпечатки пальцев. Причем хранятся в виде голограмм, мультиплексированных методом фазового кодирования опорного пучка. Используемая для кодирования фазовая маска – диффузный рассеиватель располагается вне устройства в установке чтения карт, что делает невозможным несанкционированное считывание и копирование информации с карт вне специализированной установки. Запись информации на карте осуществляется с помощью коммерчески доступных ПЗС камеры и оптоэлектронного модуля для снятия отпечатков пальцев. Полученная с этих устройств цифровая информация первоначально кодируется с помощью цифрового электронного кода и затем преобразуется в распределение коэффициента пропускания пространственного модулятора света с электронной адресацией. Проходящий через этот модулятор параллельный пучок лазерного излучения считывает и несет на себе отображенную модулятором информацию. Последняя записывается на фотополимерной регистрирующей среде в виде голограмм Фурье с помощью опорного пучка, предварительно прошедшего через диффузный рассеиватель, изготовленный методом литографии. При этом текстовый документ, фотография и отпечатки пальцев регистрируются последовательно на одном и том же участке

фотополимера с использованием техники углового мультиплексирования. Считывание информации, записанной на карте, осуществляется в специальном устройстве путем освещения карты восстанавливающим пучком, прошедшим через рассеиватель, полностью идентичный рассеивателю, использовавшемуся на стадии

Основы голографии

записи голограммы. Восстановленный с помощью голограммы пучок лучей направляется на светочувствительную поверхность CMOS камеры, преобразуется с ее помощью в цифровой сигнал, обрабатывается и визуализируется на экране монитора. Применение в устройстве цифровых технологий записи и обработки информации позволяет осуществлять сравнение восстановленных с помощью карты изображений с оригиналами документов путем оценки уровня битовых ошибок, что существенно упрощает процедуру идентификации личности по фотографии и рисунку папиллярных линий.

В качестве источника излучения в устройствах записи и считывания карт используется Nd : YAG лазер производства фирмы Coherent, излучающий на длине волны 532 нм. Регистрация голограмм осуществляется на фотополимере HRF 150-38. Модуляция амплитуды объектного пучка в устройстве записи информации осуществляется с помощью жидкокристаллического пространственного модулятора света фирмы Erson, содержащего 640 x 480 пикселей, размером 42 x 42 мкм² каждый. Регистрация информации в устройстве восстановления голограмм осуществляется с помощью CMOS камеры, состоящей из 800 x 600 пикселей, размером 8 x 8 мкм² каждый. Запись голограмм

осуществляется объектным и опорным пучками лучей, симметрично падающими на поверхность регистрирующей среды. Угол схождения этих пучков выбран равным 6 угловым градусам, что соответствует несущей пространственной частоте примерно в 200 мм⁻¹. Размер голограммы составляет, 2 x 2 мм². Реализуемая в устройстве плотность записи информации равна 52 x 39 бит/мм². При плотностях энергии в плоскости регистрации голограммы в опорном и объектном пучках, равных 1 и 0,4 Вт/см², соответственно, время экспонирования не превышает 4 с. Фиксирование зарегистрированных голограмм осуществляется путем их послеэкспозиционной обработки равномерным потоком УФ излучения мощностью 8 Вт в течение 2 минут.

Используемая в устройствах записи и восстановления карт фазовая маска со случайным распределением фазовой составляющей ее коэффициента пропускания изготовлена методом литографии, на слое фоторезиста, нанесенного на стеклянную подложку, и характеризуется бинарными фазовыми сдвигами 0 и π радиан. Она состоит из 64 x 64 пикселей, каждый из которых имеет размеры 200 x 200 мкм². Структура маски такова, что она не может быть ни скопирована, ни зарегистрирована с помощью ПЗС камер. Именно это ее свойство позволяет использовать фазовую маску в качестве ключа при кодировании и декодировании информации.

Лекция 26.

Голографические системы кодирования информации (продолжение)

26.1. Изготовленные методом голографии фотонные полимеры и их применение

Под фотонными полимерами обычно понимают полимерные материалы, обладающие периодически, с периодом соизмеримым с длиной волны излучения, изменяющейся в пространстве величиной показателя преломления. Такие полимеры находят широкое применение в оптоэлектронике. В частности, на их основе создаются переключатели информационных каналов, оптические аттенюаторы, узкополосные и перестраиваемые спектрально-селективные фильтры, модуляторы и т.д.

Использование в указанных устройствах именно полимерных материалов объясняется, в первую очередь, тем, что полимеры состоят из большого числа циклически повторяющихся цепей мономеров. Они позволяют объединять в едином материале множество сред, обладающих различными механическими и оптическими характеристиками, они легко интегрируются в уже существующие устройства, что позволяет создавать элементную базу оптоэлектроники нового типа. Кроме того, фотонные полимеры являются единственным, на сегодняшний день, материалом, пригодным для создания интегральных многофункциональных информационных каналов большой протяженности.

Наиболее распространенный и относительно недорогой метод получения фотонных полимеров включает в себя химический синтез светочувствительной однородной по показателю преломления фотополимерной среды и последующую пространственную модуляцию ее показателя преломления путем воздействия на среду специальным образом сформированным голографическим полем. При этом характер получаемой модуляции показателя преломления полимера обычно соответствует закону распределения интенсивности излучения в голографическом поле. Голографическое поле для изготовления фотонных полимеров может быть сформировано, как традиционными методами физической голографии, так и с помощью синтезированных голограмм. Отметим, что литографические методы изготовления фотонных полимеров, по своей сути, являются голографическими, поскольку требуют для своей реализации наличия фотошаблона, являющегося разновидностью синтезированной голограммы.

Основы голографии

По реализуемой в фотонных полимерах амплитуде модуляции показателя преломления их условно можно разделить на две большие группы. К первой группе можно отнести полимеры с амплитудой мене 0,1. Наиболее яркими представителями этой группы являются бихромированная желатина и фотополимеры. На их основе обычно изготавливаются различного вида маршрутизаторы и направленные ответвители информационных каналов. Ко второй группе относятся материалы, характеризующиеся, как правило, двумерной или трехмерной модуляцией показателя преломления с амплитудой, превышающей или равной 0,2. Они обычно используются для создания оптоэлектронных устройств с запрещенными фотонными зонами, т.е. устройств, подобных полупроводниковым кристаллам с их запрещенными энергетическими уровнями. Для создания таких материалов обычные фотополимеры, к сожалению, не подходят, поскольку обладают весьма малой предельной величиной амплитуды модуляции показателя преломления, не превосходящей 0,08. Указанное ограничение может быть преодолено различными путями. В частности,

специалисты американской Air Force Research Laboratory совместно с учеными Crystal Research, Inc. предложили создавать фотонные полимеры второй группы путем внедрения наночастиц жидких кристаллов в матрицу оптических полимеров. Технология получения таких материалов включает в себя диспергирование жидких кристаллов в полимерную матрицу. Экспонирование системы жидкий кристалл-полимер УФ излучением с пространственно неоднородным распределением интенсивности приводит к перераспределению жидких кристаллов по объему полимера. Они "уходят" из освещенных областей и скапливаются в темных, неэкспонированных участках полимера.

Пространственная модуляция показателя преломления системы жидкий кристалл – полимер при этом образуется за счет различия исходных показателей преломления составляющих ее материалов. Амплитуда такой модуляции может варьироваться в широких пределах и может достигать до 0,5, что вполне достаточно для фотонных полимеров второй группы. Изменение амплитуды модуляции рассматриваемых фотонных полимеров предлагается осуществлять путем соответствующего выбора компонентов системы. Отметим, что показатели преломления полимеров лежат в пределах от 1,3 до 1,9, а показатели преломления жидких кристаллов лежат в пределах от 1,5 до 1,9.

Основы голографии

Изготовленные таким образом фотонные полимеры нашли применение в ряде оптоэлектронных устройств. В качестве примера рассмотрим два из них: линию задержки сигнала радиочастоты и систему передачи сигнала синхронизации. Оптоэлектронная широкополосная линия задержки предназначена для имитации вращения неподвижных антенн РЛС путем изменения фазовой задержки между соседними элементами их раскрыва. Ее основное преимущество перед чисто электронными аналогами заключается в возможности одновременной обработки всего спектра широкополосного сигнала радиочастоты. Отметим, что в оптоэлектронной линии задержки сигнал передается в виде потока монохроматического излучения модулированного по амплитуде сигналом радиочастоты. Конструктивно такая линия выполняется в виде кремниевой подложки, несущей в себе несколько линеек фотоприемников. На поверхность подложки нанесена длинная спиралевидная дорожка из полимера, включающая в себя вкрапления фотонного полимера. При этом полимерная дорожка выполняет функцию оптического волновода, а вкрапления фотонного полимера обеспечивают частичный вывод излучения из волновода, играя роль голографических направленных ответвителей. Каждый ответвитель расположен над одним из указанных выше фотоприемников так, что извлекаемая с его помощью из волновода часть светового потока направляется на светочувствительную площадку этого фотоприемника. Обеспечиваемое устройством время задержки лежит в пределах от 1 пс до 50 нс, что соответствует времени распространения излучения по волноводу от входа устройства до первого и последнего ответвителя, соответственно. Дифракционная эффективность голографических направленных ответвителей варьируется в зависимости от требований к конкретной реализации устройства и обычно лежит в пределах от 0,5% до 7%. Устройство обеспечивает работу без существенных дисперсионных искажений сигнала в пределах полосы частот радиосигнала 60 ГГц, определяемой характеристиками фотоприемников. Спектральная ширина излучения не должна превышать при этом 300 ГГц, что соответствует диапазону длин волн в 1064 ± 1 нм. Отметим, что сечение волноводов устройства составляет 5×5 мкм² при их общей длине 10 м. Изготовлены они методом фотолитографии путем травления полимера через резистивную маску. Потери излучения в волноводе не превышают 0,02 дБ/см в рабочем диапазоне длин волн.

Основы голографии

Применение фотонных полимеров не ограничивается линиями задержки. Они с успехом могут использоваться, например, при создании систем передачи сигнала синхронизации в электронных устройствах. Указанные системы представляют собой оптические разветвители информационных каналов веерного типа, отличающиеся постоянным временем передачи сигнала от входа до каждого чипа – потребителя. Рассмотрим одну из реализаций таких систем, обеспечивающую передачу одного сигнала синхронизации 42 чипам выполненную на основе стеклянного планарного волновода. На одну из его поверхностей нанесен двумерный растр, состоящий из 43 скрещенных голограммных решеток, обеспечивающих доставку сигнала синхронизации потребителям и вывод указанного сигнала из планарного волновода. Такая система позволяет успешно передавать 25 ГГц сигнал синхронизации 42 потребителям.

В заключение следует отметить, что фотонные полимеры служат прекрасной средой для создания фотонных кристаллов, отличающихся трехмерной решеткой, и устройств на их основе. В частности, в литературе описан демультиплексор, выполненный в слое нанесенного на стеклянную подложку полимера, содержащего две треугольные фотонно-кристаллические призмы, выполненные из фотонного полимера. Сравнение фотонных полимеров с другими материалами фотоники позволяет сделать вывод о перспективности и эффективности их использования для создания коммерчески конкурентоспособных многофункциональных интегральных устройств передачи и обработки информации.

Дополнительная литература

1. Кольер Р., Беркхард К., Лин Л. Оптическая голография –М.:Мир, 1973, 686с.
2. Оптическая голография /Под ред. Г. Колфилда. –М.:Мир, 1982, 735с.
3. Миллер М. Голография. - Л.: Машиностроение, 1979, 140с.
4. Ландсберг Г.С. Оптика — М.: ФИЗМАТЛИТ, 2003, 848 с.
5. Фризер Х. Фотографическая регистрация информации - М.:Мир, 1978, 670с.
6. Ахманов С.А., Никитин С.Ю. Физическая оптика — М.: Издательство МГУ им. М. В. Ломоносова: Наука, 2004, 654 с.
7. Saxby G. Practical holography. – Taylor&Francis, 2003, 482p.
8. Янг М. Оптика и лазеры, включая волоконную оптику и оптические волноводы— М.: Мир, 2005 , 542 с.
9. Гудмен Дж. введение в Фурье-оптику. - М.: Мир, 1970, 364с.
10. Palmer K. Diffraction grating handbook. – Richardson grating laboratory, 2000, 144 p.
11. Horimai H., Tan X., Li J. Collinear holography. Applied optics, 2005, v.44, №13, p.2575-2579.
12. Ludman J., Caulfield H.J., Riccobono J. Holography for the new millennium. – Springer, 2002, 344p.
13. Pedrini G., Zhang F., Osten W. Digital holographic microscopy in the deep ultraviolet. Applied Optics, 2007, v. 46, №32, p. 7630 – 7635.
14. Корешев С.Н. Основы голографии и голограммной оптики. Учеб.пос. ИТМО, Санкт-Петербург, 2009. – 97с.