

ФИЗИЧЕСКИЕ ОСНОВЫ МИКРО- И НАНОТЕХНОЛОГИЙ

Под редакцией Ю.Б. Цветкова

*Допущено Учебно-методическим объединением вузов
по университетскому политехническому образованию
в качестве учебного пособия для студентов
высших учебных заведений, обучающихся
по машиностроительным
и приборостроительным специальностям*



Москва 2009

УДК 621.384-022.532(075.8)

ББК 32.85

Ф505

Рецензенты:

д-р техн. наук, проф. *А.М. Медведев*, д-р техн. наук,
проф. *В.В. Слепцов*, д-р техн. наук, проф. *С.Б. Нестеров*,
д-р техн. наук, проф. *В.А. Шахнов*

Ф505

Физические основы микро- и нанотехнологий: Учеб. пособие / С.П. Бычков, В.П. Михайлов, Ю.В. Панфилов, Ю.Б. Цветков; Под ред. Ю.Б. Цветкова. — М.: Изд-во МГТУ им. Н.Э. Баумана, 2009. — 176 с.: ил.

ISBN 978-5-7038-3319-3

Даны краткие сведения из отдельных разделов физики и смежных наук, составляющих физические основы микро- и нанотехнологий. Рассмотрены основные законы, понятия и определения физических явлений, процессов и сред, являющихся основой этих технологий. Приведены методы преобразования фундаментальных физических законов в прикладные технологические зависимости, расчеты физических закономерностей, определяющих параметры процессов микро- и нанотехнологий.

Для студентов 3-го и 4-го курсов, обучающихся по специальности «Электронное машиностроение».

УДК 621.384-022.532(075.8)

ББК 32.85

ПРЕДИСЛОВИЕ

В данном пособии в компактном и адаптированном виде приведены сведения из отдельных разделов физики и смежных наук, составляющие физические основы микро- и нанотехнологий.

Первая глава посвящена физике вакуума — уникальной технологической среды для реализации микро- и нанотехнологий.

Во второй главе рассмотрены физические закономерности, определяющие параметры процессов теплопередачи применительно к термическим процессам микро- и нанотехнологий.

Третья глава дает представление о методах и средствах формирования электронных, ионных, атомарных и молекулярных потоков, а в четвертой главе рассмотрено их использование как инструмента микро- и нанотехнологий.

В пятой главе рассмотрены физические свойства газоразрядной плазмы как среды и инструмента микро- и нанотехнологий.

Шестая глава посвящена анализу ключевых параметров оптического излучения и оптических процессов формирования микроизображения применительно к использованию в микролитографии.

В седьмой главе даны сведения о методах оценки качества проекционных микроизображений и основные положения фурье-преобразований в оптике для оценки качества микроизображений в проекционной микролитографии.

Изучив материал, изложенный в данном пособии, студент должен иметь представление о свойствах и возможностях физических явлений и процессов применительно к использованию в микро- и нанотехнологиях.

Студент должен знать:

- основные законы, понятия и определения физических явлений, процессов и сред, являющихся основой микро- и нанотехнологий;

- методы преобразования фундаментальных физических законов в прикладные технологические зависимости.

Студент должен уметь:

- рассчитывать и анализировать физические закономерности, определяющие параметры процессов микро- и нанотехнологий;

- моделировать физические процессы для оценки и выбора технологических параметров;

- оценивать применимость физических процессов и сред для реализации процессов микро- и нанотехнологий с заданными параметрами.

Студент должен иметь навыки:

- применения математических методов и аппарата современной физики для технологического анализа процессов, явлений и сред;

- анализа и использования современных источников научно-технической информации, освещающих применение физических процессов в микро- и нанотехнологиях.

ВВЕДЕНИЕ

Микро- и нанотехнологии представляют собой совокупность методов и технических средств (оборудования, инструментов, используемых материалов), применяемых для исследования, разработки и производства сверхминиатюрных приборов и устройств, элементы которых имеют микро- и нанометровые размеры.

Микротехнологии в общем случае включают следующие основные этапы (см. рисунок):

1) формирование на поверхности заготовки тонкого технологического слоя;

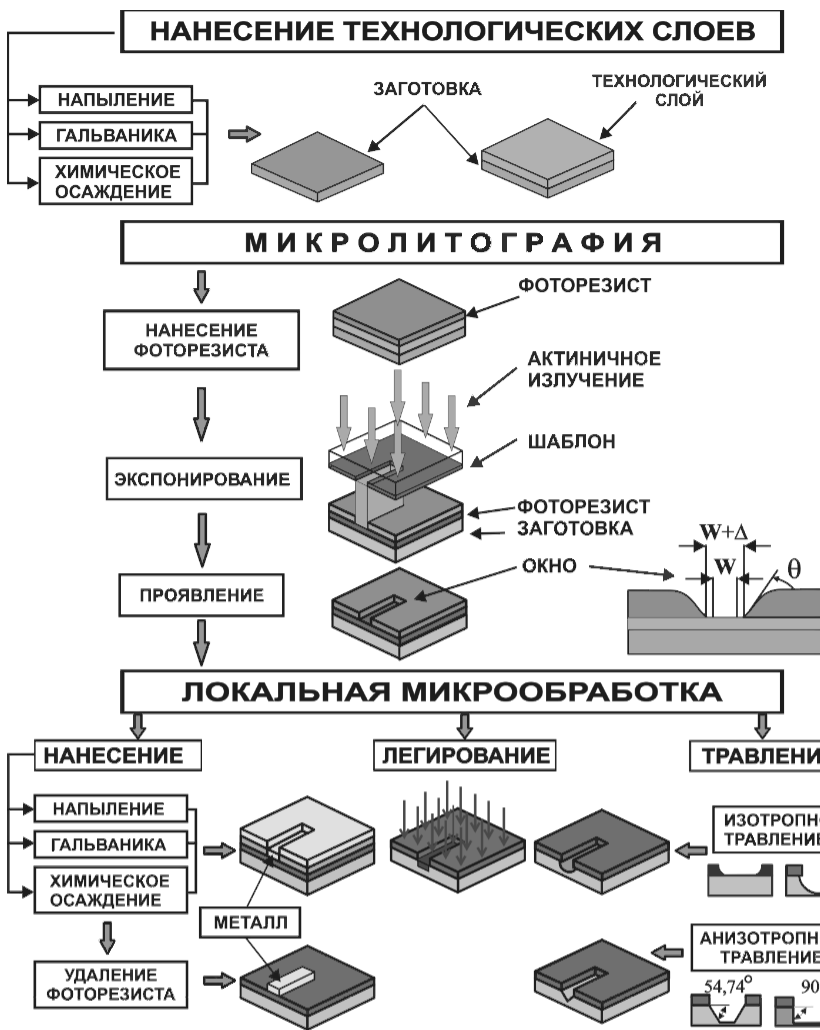
2) создание на поверхности этого слоя защитной маски с локальными окнами;

3) микрообработку технологического слоя через окна в маске.

Первый этап микротехнологий реализуется методами *элионных технологий*, позволяющих с помощью электронных, ионных, атомарных и молекулярных потоков создавать на поверхности заготовок проводящие и диэлектрические слои толщиной от нескольких нанометров до единиц микрометров.

На втором этапе используется *микролитография*, которая позволяет локализовать зоны воздействия на заготовку. Для этого на ее поверхность наносится тонкая чувствительная к актиничному излучению полимерная пленка (фоторезист), которая затем экспонируется через шаблон с требуемым рисунком (топологией).

При последующем проявлении происходит локальное удаление участков фоторезиста, образуются окна требуемых размеров и формы, через которые возможен доступ к поверхности заготовки (см. рисунок). Окна в фоторезисте, их размеры и профиль должны отвечать чрезвычайно жестким требованиям, поскольку они определяют качество третьего этапа микротехнологий.



Третий этап — это *микрообработка*, которая обеспечивает *локальное* воздействие на заготовку: нанесение материала, легирование поверхности заготовки или ее травление. При микрообработке воздействие обрабатывающей среды ограничивается (локализуется) окнами в резисте. Арсенал современных методов микрообработки весьма широк. Для *локального нанесения* могут использоваться уже упомянутые методы эллионных технологий, например вакуумное напыление и химическое осаждение из паровой фазы, а также гальваническое наращивание.

Локальное легирование проводится методами диффузии из парогазовой смеси или ионным легированием.

Локальное травление (изотропное или анизотропное) выполняется жидкостными травителями или с использованием плазменных методов, позволяющих создавать микроструктуры требуемого профиля.

Важной особенностью микротехнологий является групповой метод обработки — за один цикл экспонирования формируются миллионы окон в пределах одного модуля на заготовке, а затем также одновременно через эти окна заготовка подвергается микрообработке.

Возможность одновременного и относительно быстрого получения огромного числа элементов структур с микрометровыми и субмикрометровыми размерами оказала революционное воздействие на развитие микротехнологий.

Микротехнологии доведены до промышленного уровня, разработаны научно-технологические основы производства микроструктур на основе групповой прецизионной обработки. Основной сферой применения микротехнологий является микроэлектроника, в которой ежегодно производятся сотни миллионов изделий.

В последнее время нанотехнологии стали объектом пристального внимания мирового научного сообщества, поскольку уменьшение размеров элементов структур до нанометров может дать не только количественное, но и качественное изменение изделий — их конструкций и параметров.

Как микро-, так и нанотехнологии основываются на совершенно иных принципах, нежели технологии, имеющие дело с макротелами. Так, при обработке микроизделий функции инструмента выполняют частицы — электроны, ионы, атомы и молекулы участвующих в процессе веществ. В качестве среды обработки часто используют вакуум, парогазовые смеси, растворы реактивов. Обработка нередко ведется при высокой температуре, которая должна поддерживаться с очень высокой точностью. В связи с этим микро- и нанотехнологии строятся на основе применения в производстве современных достижений фундаментальных наук.

1. ВАКУУМ — РАБОЧАЯ СРЕДА ДЛЯ МИКРО- И НАНОТЕХНОЛОГИЙ

Изучив материал этой главы, студент должен иметь представление о физических процессах в вакууме как основной технологической среде для реализации микро- и нанотехнологий.

Студент должен знать:

- основные понятия физики вакуума для использования в микро- и нанотехнологиях (давление газа, объем газа, приходящий к единице поверхности стенки в единицу времени, длина свободного пробега молекул, степени вакуума);

- законы растворимости, диффузии газов в твердых телах, законы Фика для стационарного и нестационарного процессов диффузии, распределения концентрации растворенного газа в твердом теле.

Студент должен уметь проводить анализ физических процессов в вакууме для реализации микро- и нанотехнологий.

Студент должен иметь навыки оценки применимости физических процессов в вакуумной среде для микро- и нанотехнологий с заданными требованиями.

1.1. Основные положения физики вакуума

Вакуум — газовая среда, находящаяся под давлением ниже атмосферного ($p < p_{\text{атм}}$), которая используется в технологии производства практически всех электронных приборов (интегральных микросхем с микро- и нанометровыми размерами элементов, электронно-лучевых трубок, приемно-усилительных ламп и пр.), а также других изделий, и которая необходима для работы электронных вакуумных устройств.

В этой среде при реализации электронных микро- и нанотехнологий и при работе электронных приборов формируются потоки нейтральных и заряженных частиц, происходит управление их концентрацией, скоростью, взаимодействием с объектами, направлением движения и другими параметрами.

Согласно ГОСТ 8.417–81 давление вакуумной среды измеряется в следующих единицах: $1 \text{ Па} = 1 \text{ Н/м}^2$; $1 \text{ бар} = 10^5 \text{ Па} = 750 \text{ мм рт. ст.} = 750 \text{ тор}$; $1 \text{ тор} = 133,3 \text{ Па}$.

Вакуум обладает рядом свойств, которые используются для реализации технологических процессов:

1) предохраняет нагреваемые поверхности (нити накала, катоды и др.) от окисления, перегорания;

2) позволяет формировать потоки заряженных и нейтральных частиц с приданием им необходимой энергии и плотности;

3) благодаря ему потоки заряженных частиц могут беспрепятственно преодолевать межэлектродные расстояния (при этом они могут ускоряться, фокусироваться и отклоняться по заданной траектории);

4) позволяет сохранить чистоту обрабатываемой поверхности материала (обеспечить отсутствие сорбированных молекул газа, паров углеводорода и воды, оксидов и др.) для технологических процессов (формирования элементов топологии на полупроводниковой пластине, ионная очистка и т. п.).

В табл. 1.1, 1.2 приведены требования электронных микро- и нанотехнологий и электровакуумных приборов к вакуумной среде.

Таблица 1.1

Требования электронных технологий к вакуумной среде

Технологический процесс	Давление, Па	Название вакуума
<i>Электронно-лучевая обработка:</i>		
сварка	10^{-3}	ВВ (высокий)
размерная обработка, резание	10^{-3}	ВВ
плавка, зонная очистка	10^{-3}	ВВ
электронная литография	10^{-5}	СВВ (сверхвысокий)

Технологический процесс	Давление, Па	Название вакуума
<i>Ионная обработка:</i>		
очистка	10^{-2}	ВВ
травление	10^{-2}	ВВ
имплантация	10^{-3}	ВВ
ионная литография	10^{-5}	СВВ
<i>Выращивание монокристаллов</i>	10^{-4}	ВВ
<i>Нанесение тонких пленок:</i>		
нанесение пассивных элементов	10^{-3}	ВВ
нанесение оптических элементов	10^{-6}	СВВ
молекулярно-лучевая эпитаксия	10^{-9}	СВВ
<i>Контроль качества поверхности в вакууме:</i>		
сканирующая СВВ-туннельная микроскопия	$10^{-8} \dots 10^{-10}$	СВВ
атомно-силовая СВВ-микроскопия	$10^{-8} \dots 10^{-10}$	СВВ
<i>Сборка фотоэлектронных приборов (приборы ночного видения)</i>	10^{-10}	СВВ

Таблица 1.2

Требования электровакуумных приборов к вакуумной среде

Электровакуумные приборы	Давление, Па	Название вакуума
Газонаполненные лампы накаливания	10^4	НВ (низкий)
Газотроны, люминесцентные лампы, газовые лазеры	10^{-1}	СВ (средний)
Электронно-лучевые трубки, приемно-усилительные лампы	10^{-5}	СВВ
Лампы бегущей волны, магнетроны, клистроны	10^{-7}	СВВ
Фотоэлектронные приборы	10^{-10}	СВВ

Свойства газов при низких давлениях изучаются в разделе молекулярно-кинетической теории газов, который называется физикой вакуума. Основные постулаты, используемые в физике вакуума, можно сформулировать в следующем виде:

- 1) газ состоит из молекул, находящихся в непрерывном и беспорядочном движении;
- 2) существует устойчивое распределение молекул газа по скоростям, т. е. одинаковую скорость имеет всегда одно и то же число молекул;
- 3) газ — среда изотропная;
- 4) температура газа — мера средней кинетической энергии молекул;
- 5) взаимодействие молекул газа происходит как соударение упругих шаров (для идеального газа);
- 6) при соударении с поверхностью твердого тела молекула сорбируется, т. е. на некоторое время задерживается на стенке;
- 7) давление есть результат ударов молекул о стенку сосуда.

Распределение молекул газа по скоростям. Давление газа

Основным параметром вакуумной среды является общее давление, которое характеризуется отношением нормальной составляющей силы, действующей со стороны газа на ограничивающую поверхность, к площади этой поверхности.

Согласно молекулярно-кинетической теории газов, в стационарном состоянии (температура газа равна температуре стенки) устанавливается определенное распределение скоростей между отдельными молекулами (рис. 1.1). Для описания этого распределения используют различные виды скорости молекул.

Наиболее вероятная скорость движения молекул газа

$$v_{\text{н}} = \sqrt{\frac{2kT}{m}}, \quad (1.1)$$

где k — постоянная Больцмана; T — абсолютная температура газа; m — масса молекулы.

Средняя арифметическая скорость

$$v_{\text{а}} = \sqrt{\frac{8kT}{\pi m}} = \sum_{i=1}^t \frac{n_i v_i}{n}, \quad (1.2)$$

где n_i — количество молекул, имеющих скорость v_i ; n — концентрация молекул, $n = \sum_{i=1}^t n_i$; t — количество газов в объеме.

Средняя квадратичная скорость

$$v_k = \sqrt{\frac{3kT}{m}} = \sqrt{\frac{\sum_{i=1}^t n_i v_i^2}{n}}. \quad (1.3)$$

Соотношение скоростей (1.1)–(1.3) молекул будет определять-ся следующим образом:

$$v_H : v_a : v_k = 1 : 1,13 : 1,22.$$

Значения скоростей молекул газа для нормальной температуры $T = 293$ К приведены в табл. 1.3.

Таблица 1.3

Распределение молекул газа по скоростям

Газ	M^*	v_H , м/с	v_a , м/с	v_k , м/с
H ₂	2	1554	1752	1902
N ₂	28	417	470	510
O ₂	32	390	440	477
Воздух	29	410	462,5	501,9

* M — молекулярная масса (в атомных единицах массы).

Найдем выражение для давления p из уравнения Менделеева — Клапейрона для одного моля газа:

$$pV_m = RT,$$

где V_m — объем одного моля газа; R — универсальная газовая постоянная.

Перепишем уравнение в виде

$$p = \frac{R}{V_m} T = \frac{N_A}{V_m} \frac{R}{N_A} T = nkT, \quad (1.4)$$

где N_A — число Авогадро; n — концентрация молекул.

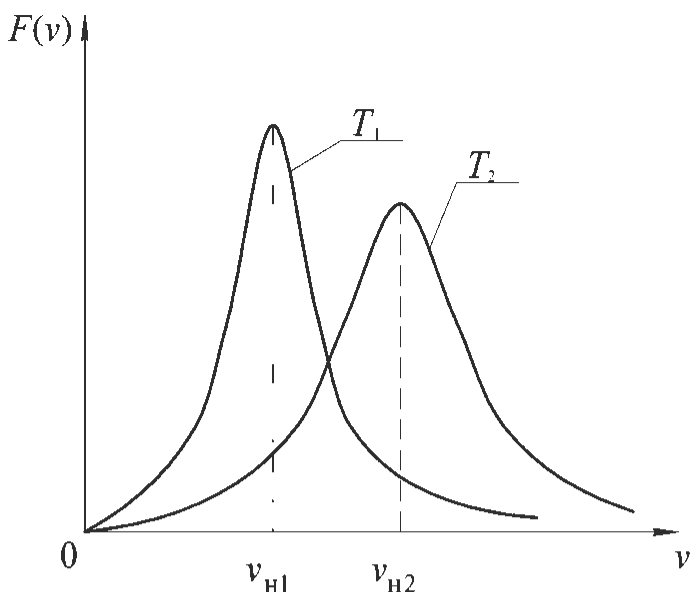


Рис. 1.1. Распределение молекул газа по скоростям:

$F(v)$ — вероятность того, что молекула имеет абсолютную скорость v ;
 v_{H1} , v_{H2} — наиболее вероятные скорости молекул для температур T_1 и T_2

Исходя из того, что средняя кинетическая энергия молекулы равна

$$\bar{E} = \frac{mv_{\text{к}}^2}{2} = \frac{3}{2}kT,$$

можно записать

$$kT = \frac{2}{3}\bar{E}. \quad (1.5)$$

Используя формулы (1.4) и (1.5), запишем основное уравнение молекулярно-кинетической теории газов:

$$p = nkT = \frac{2}{3}n\bar{E}. \quad (1.6)$$

Если в объеме находится смесь из k химически невзаимодействующих газов, то для определения общего давления газа применяют закон Дальтона:

$$p = p_1 + p_2 + \dots + p_k = \sum_{i=1}^k p_i, \quad (1.7)$$

где p_i — парциальное давление i -го газа.

Для примера ниже приведены парциальные давления газов, соответствующие формуле (1.7), для атмосферного воздуха, находящегося при температуре $T = 298$ К:

Газ	N_2	O_2	Ar	CO_2	He	H_2
Парциальное давление, Па	$7,81 \cdot 10^4$	$2,1 \cdot 10^4$	$9 \cdot 10^2$	$3 \cdot 10^2$	$5,2 \cdot 10^{-1}$	$5 \cdot 10^{-2}$

Для нахождения объема газа, ударяющегося в единицу времени о единицу поверхности, определим число соударений молекул в единицу времени о единицу поверхности твердого тела:

$$N_1 = \frac{nv_a}{4}.$$

Тогда объем газа, ударяющегося в единицу времени о единицу поверхности, с учетом формулы (1.2) можно определить как

$$V_1 = \frac{N_1}{n} = \frac{v_a}{4} = \frac{1}{4} \sqrt{\frac{8kT}{\pi m}}. \quad (1.8)$$

При нормальных условиях ($m = 29$; $T = 293$ К) $v_a = 462,5$ м/с объем газа с учетом (1.8) будет равен

$$V_1 = \frac{462,5}{4} \approx 116 \text{ м/с.}$$

Величина V_1 при постоянной температуре не зависит от давления p и определяет максимальную теоретическую быстроту действия вакуумных насосов.

Длина свободного пробега молекул газа

Важнейшей характеристикой вакуумной среды является длина свободного пробега молекул.

Определим длину свободного пробега молекул газа:

$$\lambda = \frac{v_a}{\nu},$$

где ν — число столкновений молекулы с другими частицами.

Пусть все молекулы газа неподвижны, кроме одной (нижней на рис. 1.2), движущейся со скоростью v_a . Молекула столкнется со всеми молекулами, центры которых находятся в цилиндре диаметром 2σ и высотой v_a . Тогда число столкновений

$$\nu = \pi\sigma^2 v_a n,$$

где σ — диаметр молекулы; n — концентрация молекул.

Таким образом,

$$\lambda = \frac{1}{\pi\sigma^2 n} = \frac{kT}{P\pi\sigma^2},$$

так как $n = \frac{P}{kT}$ согласно формуле (1.6).

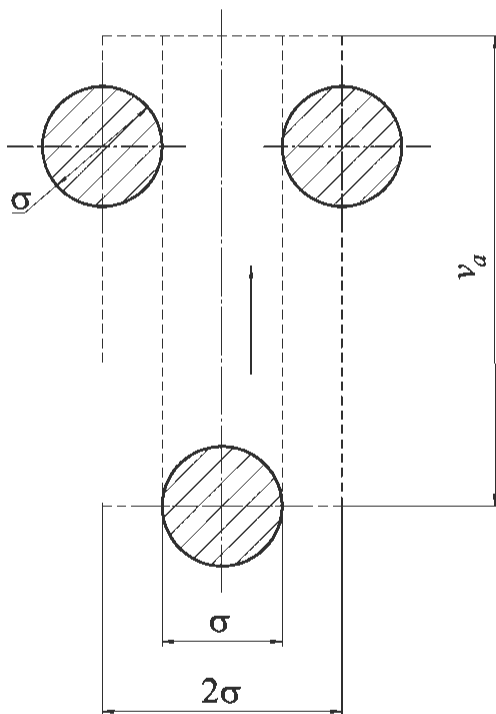


Рис. 1.2. Схема для определения длины свободного пробега молекул

В действительности все молекулы движутся, и рассматриваемая молекула движется с относительной скоростью $v_{a \text{ отн}} = \sqrt{2} v_a$. Тогда число столкновений и длина свободного пробега будут равны соответственно

$$v = \sqrt{2} \pi \sigma^2 v_a n; \quad \lambda = \frac{1}{\sqrt{2} \pi \sigma^2 n} = \frac{kT}{\sqrt{2} p \pi \sigma^2}. \quad (1.9)$$

Например, для молекулы воздуха при нормальных условиях $T = 293 \text{ К}$, $\sigma = \sigma_v = 3,7 \cdot 10^{-10} \text{ м}$, $k = 1,38 \cdot 10^{-23} \text{ Дж/К}$ с учетом (1.9) получим

$$\lambda = \frac{1,38 \cdot 10^{-23} \cdot 293}{\sqrt{2} \pi \cdot (3,7 \cdot 10^{-10})^2 p} = \frac{6,51 \cdot 10^{-3}}{p}.$$

Степени вакуума

Вакуумная среда характеризуется степенями вакуума. Когда длина свободного пробега молекул λ становится соизмеримой с размерами сосуда (вакуумной камеры, электровакуумного прибора и др.) и молекулы вместо взаимных соударений начинают сталкиваться в основном со стенками сосуда, резко меняется характер явления переноса (теплопроводности, диффузии, вязкости, движения потоков заряженных частиц — электронов и ионов).

Таким образом, степень вакуума, определяемая значением λ , существенно влияет на процессы, происходящие в вакуумированном объеме.

Низкий вакуум — характеризуется давлением газа p , при котором $\lambda \ll d$, где d — характерный размер вакуумного сосуда.

Средний вакуум — характеризуется давлением газа p , при котором λ соизмерима с d ($\lambda \approx d$).

Высокий вакуум — характеризуется давлением газа p , при котором $\lambda \gg d$.

Сверхвысокий вакуум — характеризуется давлением газа p , при котором за время рабочего процесса не происходит заметного изменения свойств поверхности, первоначально свободной от адсорбированного газа.

Таким образом, степень вакуума определяется двумя параметрами: давлением p (и соответственно длиной свободного пробега λ) и характерным размером вакуумного сосуда d (например, диаметром). Условные границы степеней вакуума при $d = 16 \dots 250 \text{ мм}$ приведены в табл. 1.4.

Условные границы степеней вакуума

Критерий степени вакуума	Давление, Па	Название вакуума
$pd \geq 1,2 \text{ м}\cdot\text{Па}$	$10^5 \dots 10^2$	Низкий
$0,004 < pd < 1,2 \text{ м}\cdot\text{Па}$	$10^2 \dots 10^{-1}$	Средний
$pd \leq 0,004 \text{ м}\cdot\text{Па}$	$10^{-1} \dots 10^{-4}$	Высокий
$\theta < 1, p < 4 \cdot 10^{-4}/t_0$	Менее 10^{-4}	Сверхвысокий

Условия существования сверхвысокого вакуума определяются его взаимодействием с поверхностью материала и, в частности, коэффициентом заполнения поверхности $\theta = N'_{\text{пов}}/N'_n$, где $N'_{\text{пов}}$ — число сорбированных молекул на единице поверхности; N'_n — число мест на единице поверхности. Если за время проведения рабочего процесса t_0 сохраняется хотя бы малая доля чистой (ювенильной) поверхности, т. е. $\theta < 1$, то граница сверхвысокого вакуума определяется как $p \leq 4 \cdot 10^{-4}/t_0$. С учетом того, что $t_0 = 1 \text{ с}$, можно принять $p < 10^{-4} \text{ Па}$.

1.2. Физические процессы при вакуумной откачке

Для получения вакуумной среды используются вакуумные системы, примеры которых показаны на рис. 1.3. Для создания низкого и среднего вакуума ($p \approx 10^3 \dots 10^{-1} \text{ Па}$) применяются, как правило, механические насосы (рис. 1.3, а). Для получения высокого вакуума ($p \approx 10^{-3} \text{ Па}$) используются последовательно соединенные форвакуумный (механический) и высоковакуумный (например, диффузионный) насосы (рис. 1.3, б).

Вакуумные насосы можно классифицировать следующим образом:

1) по назначению — сверхвысоковакуумные, высоковакуумные и низковакуумные (форвакуумные);

2) по принципу действия — механические, сорбционные (для низкого, среднего и высокого вакуума), диффузионные, магнитные электроразрядные, геттерно-ионные, крионасосы (для высокого и сверхвысокого вакуума).

Диапазон значений давления, используемого в вакуумных и электронных технологиях, равен $10^5 \dots 10^{-12} \text{ Па}$.

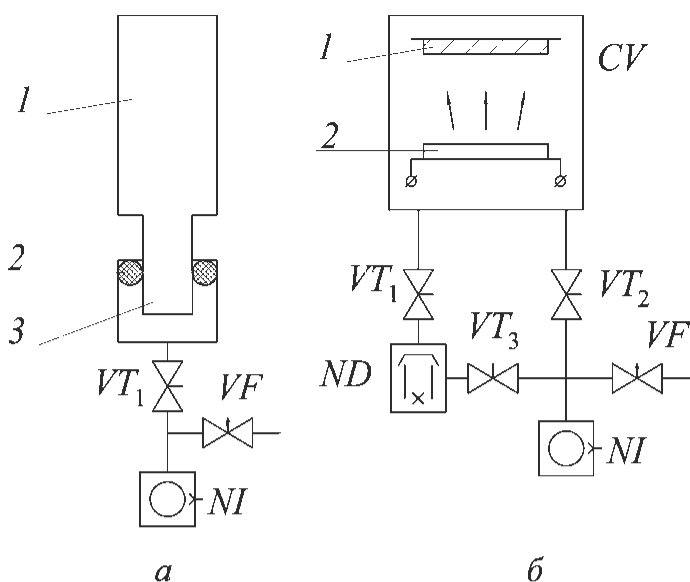


Рис. 1.3. Принципиальные схемы вакуумных систем:

а — низковакуумная пневматическая откачка прибора ($p \approx 10^2$ Па): 1 — корпус прибора; 2 — герметизирующий элемент; 3 — пневматический элемент;

б — высоковакуумная откачка ($p \approx 10^{-3}$ Па) установки для нанесения тонких пленок: 1 — подложка; 2 — испаритель; VT_1, VT_2, VT_3 — вентили тарельчатые; VF — натекатель; ND, NI — насосы диффузионный и механический; CV — колпак вакуумный

Полное давление измеряют вакуумметрами, состоящими из преобразователя давления (ПД) и измерительного блока (ИБ). Вакуумметры могут быть классифицированы по методу измерения и принципу действия. По методу измерения различают вакуумметры прямого и косвенного измерения, по принципу действия — жидкостные, деформационные, тепловые, электронные, магнитные и др.

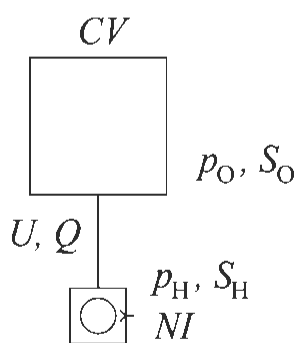


Рис. 1.4. Расчетная принципиальная схема вакуумной системы

Основными параметрами вакуумной системы (рис. 1.4) являются быстрота откачки объекта S_0 , быстрота действия насоса S_n , производительность насоса Q и проводимость трубопровода U .

Быстрота откачки объекта $S_0, \text{ м}^3/\text{с}$, — объем газа, поступающий в единицу времени из откачиваемого объекта в трубопровод при давлении p_0 .

Быстрота действия насоса $S_{\text{н}}$, м³/с, — объем газа, удаляемый насосом в единицу времени через впускной патрубок при давлении $p_{\text{н}}$.

Производительность насоса Q , м³·Па/с, — поток газа, проходящий через впускной патрубок насоса.

Проводимость трубопровода U , м³/с, — количество газа, протекающего через трубопровод в единицу времени при разности давлений на концах трубопровода, равной единице:

$$U = \frac{Q}{p_0 - p_{\text{н}}}.$$

Основное уравнение вакуумной техники

Для стационарного потока ($Q = \text{const}$) при откачке вакуумной камеры выполняется условие сплошности газовой среды:

$$Q = P_0 S_0 = P_{\text{н}} S_{\text{н}} = \text{const}.$$

Установим взаимосвязь между основными параметрами вакуумной системы S_0 , $S_{\text{н}}$, U :

$$S_{\text{н}} = \frac{Q}{p_{\text{н}}} = \frac{U(p_0 - p_{\text{н}})}{p_{\text{н}}}; \quad S_0 = \frac{Q}{p_0} = \frac{U(p_0 - p_{\text{н}})}{p_0};$$

$$\frac{1}{S_{\text{н}}} = \frac{p_{\text{н}}}{U(p_0 - p_{\text{н}})}; \quad \frac{1}{S_0} = \frac{p_0}{U(p_0 - p_{\text{н}})};$$

$$\frac{1}{S_0} - \frac{1}{S_{\text{н}}} = \frac{p_0 - p_{\text{н}}}{U(p_0 - p_{\text{н}})} = \frac{1}{U}$$

или

$$\frac{1}{S_0} - \frac{1}{S_{\text{н}}} = \frac{1}{U}. \quad (1.10)$$

Уравнение (1.10) называется *основным уравнением вакуумной техники*. Согласно этому уравнению при увеличении проводимости трубопровода до бесконечности быстрота откачки камеры стремится к быстроте действия насоса. При закрытом трубопроводе его проводимость стремится к нулю и быстрота откачки камеры также равна нулю.

При вакуумной откачке объектов обеспечивается получение требуемого давления и состава рабочей среды в вакуумированном объеме и сохранение этих параметров в течение регламентированного периода времени.

К этапам вакуумной откачки относят:

- 1) удаление газов и паров из вакуумированного объема;
- 2) обезгаживающий прогрев материалов, находящихся в вакууме.

Расчет времени откачки идеальной вакуумной системы

Для первого этапа вакуумной откачки рассмотрим методику расчета времени откачки идеальной вакуумной системы. Под идеальной вакуумной системой понимается полностью герметичная система, в которой отсутствует газовыделение с поверхности материала.

Дано: V — объем камеры (рис. 1.5); S_0 — быстрота откачки (принимается $S_0 = \text{const}$); p_0, p_k — начальное и конечное давление.

Найти: время откачки t_0 от давления p_0 до давления p_k .

1. Определим количество газа dG , откачанное за время dt :

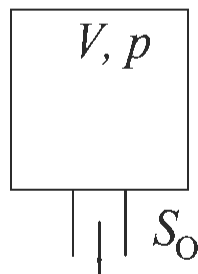
$$dG = S_0 p dt = -V dp; \quad \frac{dp}{p} = -\frac{S_0}{V} dt.$$

После интегрирования получим

$$\int_{p_0}^{p_k} \frac{dp}{p} = -\int_0^{t_0} \frac{S_0}{V} dt; \quad \ln P \Big|_{p_0}^{p_k} = -\frac{S_0}{V} t_0$$

или

$$\ln \frac{p_0}{p_k} = \frac{S_0}{V} t_0.$$



2. Найдем время откачки t_0 :
для идеальной системы (конечное давление p_k стремится к нулю при бесконечном времени t_0)

$$t_0 = \frac{V}{S_0} \ln \frac{p_0}{p_k};$$

Рис. 1.5.
Схема
идеальной
вакуумной
системы

для реальной системы (при бесконечном времени t_0 давление p_k стремится к предельному минимальному давлению системы p')

$$t_0 = \frac{V}{S_0} \ln \frac{p_0}{p_k - p'}, \quad (1.11)$$

где p' — предельное давление системы, $p' = Q_{\text{ост}} / S_0$ ($Q_{\text{ост}}$ — остаточное газовыделение в системе).

3. Запишем уравнения кривых откачки для идеальной и реальной систем:

$$p = p_0 \exp\left(-\frac{S_0}{V} t\right) \text{ — для идеальной системы;}$$

$$p = p' + p_0 \exp\left(-\frac{S_0}{V} t\right) \text{ — для реальной системы.}$$

4. Перейдем к десятичным логарифмам:

$$\ln x = 2,3 \lg x,$$

тогда формулу (1.11) можно переписать:

$$t_0 = 2,3 \frac{V}{S_0} \lg \frac{p_0}{p_k - p'}. \quad (1.12)$$

Пример. Дано: объем камеры 10 л (10^{-2} м³), $S_0 = 5$ л/с = 0,005 м³/с (для насоса 2НВР-5ДМ при условии $S_0 = S_H$), $p_0 = 10^5$ Па, $p_k = 10^1$ Па, $p' = 1,3 \cdot 10^{-2}$ Па.

Найти: время откачки t_0 .

Решение. Согласно формуле (1.12) имеем

$$t_0 = 2,3 \frac{V}{S_0} \lg \frac{p_0}{p_k - p'} = 2,3 \frac{0,01}{0,005} \lg \frac{10^5}{10^1 - 1,3 \cdot 10^{-2}} = 18,4 \text{ с.}$$

Растворимость, диффузия газа в твердом теле. Первый закон Фика для стационарного процесса диффузии

Основное количество газа в вакуумной камере находится в приповерхностном слое материала камеры и вакуумной арматуры

в растворенном состоянии. При высоковакуумной откачке именно удаление растворенного в материале газа при обезгаживающем прогреве занимает большую часть времени. Рассмотрим процессы растворимости, диффузии и проницаемости газа в твердом теле.

Сорбция — процесс поглощения газов материалами.

Адсорбция — явление прилипания молекул и атомов газа к поверхности тела.

Абсорбция — явление проникания молекул и атомов газа внутрь твердого или жидкого тела, т. е. растворение газа в объеме материала.

Физическая адсорбция обусловлена действием сил Ван-дер-Ваальса между молекулами газа и адсорбента (притяжением молекул с постоянными и индуцируемыми диполями). Врем адсорбции молекул воздуха $\tau_c \approx 10^{-10}$ с при температуре $t = 20$ °С.

Хемосорбция — поверхностное поглощение газов за счет химического взаимодействия (валентных сил).

Сорбент — поглощающее вещество.

Сорбат — поглощаемое вещество.

Десорбция (газовыделение) — явление, обратное сорбции, состоящее в освобождении сорбированных газов с поверхности или из толщи материала.

Растворимость — свойство газа растворяться в определенном объеме твердого тела до его насыщения. Растворимость зависит от свойств газа и материала. В неметаллах газ растворяется обычно в молекулярном состоянии и объемная концентрация газа C пропорциональна давлению p : $C = Sp$, где S — коэффициент растворимости. В металлах молекулы газа сначала диссоциируют на атомы, а потом растворяются. Поэтому для двухатомных газов объемная концентрация газа C пропорциональна $p^{1/2}$: $C = Sp^{1/2}$.

В общем случае объемная концентрация газа в твердом теле определяется по закону Фрейндлиха:

$$C = Sp^{1/j}, \quad (1.13)$$

где j — число атомов в молекуле газа.

Коэффициент растворимости

$$S = S_0 \exp\left(\pm \frac{Q_s}{jRT}\right),$$

где S_0 — константа растворимости; j — число атомов в молекуле газа, для металла $j = 2$, а для неметалла $j = 1$; Q_s — энергия активации растворимости (знак плюс применяется для газов, образующих с материалом химические соединения (H_2 в Ti)); знак минус — для газов, образующих истинные растворы (H_2 в Fe, Cu, Ni, газы в металлах)).

Газ, растворенный в твердом теле, диффундирует из области с большей концентрацией в область с меньшей концентрацией за счет явления диффузии.

Стационарный процесс диффузии имеет место, когда концентрация газа C неизменна во времени в каждой точке твердого тела ($C(t) = \text{const}$). Этот процесс описывается первым законом Фика, согласно которому удельный поток газа через единичную площадку составляет

$$q' = -\frac{\partial}{\partial x}(DC), \quad (1.14)$$

где C — концентрация газа в точке на расстоянии X от начала координат; D — коэффициент диффузии, m^2/s .

Если D не зависит от X , то формулу (1.14) можно переписать:

$$q' = -D \frac{\partial C}{\partial x},$$

где $\partial C / \partial x$ — градиент концентрации газа в твердом теле.

Газопроницаемость твердого тела

На основании первого закона Фика получим выражение для удельного потока газа через стенку (рис. 1.6).

Удельный поток газа через тонкую стенку

$$q' = -D \frac{\partial C}{\partial x} = -D \frac{C_{\text{атм}} - C_{\text{вак}}}{z},$$

где $C_{\text{атм}}$ — объемная концентрация газа в материале у поверхности, соприкасающейся с атмосферой; $C_{\text{вак}}$ — объемная концентрация газа в материале у поверхности, соприкасающейся с вакуумом; z — толщина стенки.

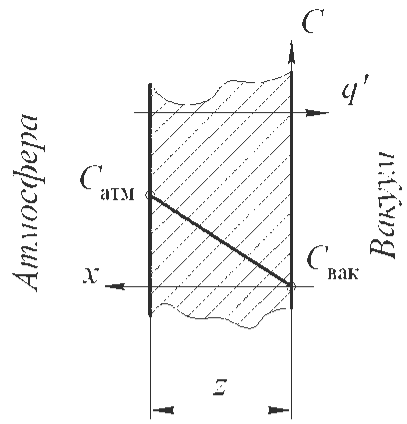


Рис. 1.6. Распределение объемной концентрации газа в тонкой стенке

Перейдя от концентрации C к коэффициенту растворимости S (см. формулу (1.13)), получим

$$q' = -DS \frac{p_{\text{атм}}^{1/j} - p_{\text{вак}}^{1/j}}{z}. \quad (1.15)$$

Если $p_{\text{вак}} \ll p_{\text{атм}}$, то (1.15) можно переписать в виде

$$q' = -DS \frac{p_{\text{атм}}^{1/j}}{z}.$$

Коэффициент диффузии D определяется как

$$D = D_0 \exp\left(-\frac{Q_D}{jRT}\right),$$

где Q_D — энергия активации диффузии.

Таким образом,

$$q' = -\Pi \frac{p_{\text{атм}}^{1/j}}{z},$$

где Π — коэффициент проницаемости, $\Pi = \Pi_0 \exp\left(-\frac{Q_{\Pi}}{jRT}\right)$; Π_0 —

константа проницаемости, $\Pi_0 = D_0 S_0$; Q_{Π} — энергия активации проницаемости ($Q_{\Pi} = Q_D \pm Q_S$, где знак плюс применяется для химического соединения газа и материала, знак минус — для истинного раствора).

Второй закон Фика для нестационарного процесса диффузии

Нестационарный процесс диффузии наблюдается при обезгаживании материала или, напротив, при поглощении им газа и описывается вторым законом Фика:

$$\frac{\partial C}{\partial t} = \frac{\partial}{\partial x} \left(D \frac{\partial C}{\partial x} \right), \quad (1.16)$$

или, если D в (1.16) не зависит от X ,

$$\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial x^2}.$$

Рассмотрим распределения объемной концентрации (рис. 1.7) в твердом плоском полубесконечном теле для различных значений времени $C = f(x, t)$, полученные при решении (1.16). Распределения приведены для противоположных процессов: газопоглощения (рис. 1.7, а) и газовыделения (рис. 1.7, б). В течение времени $t = t_1; t_2; t_3$ приповерхностный слой твердого тела насыщается газом (см. рис. 1.7, а) или обезгаживается (см. рис. 1.7, б).

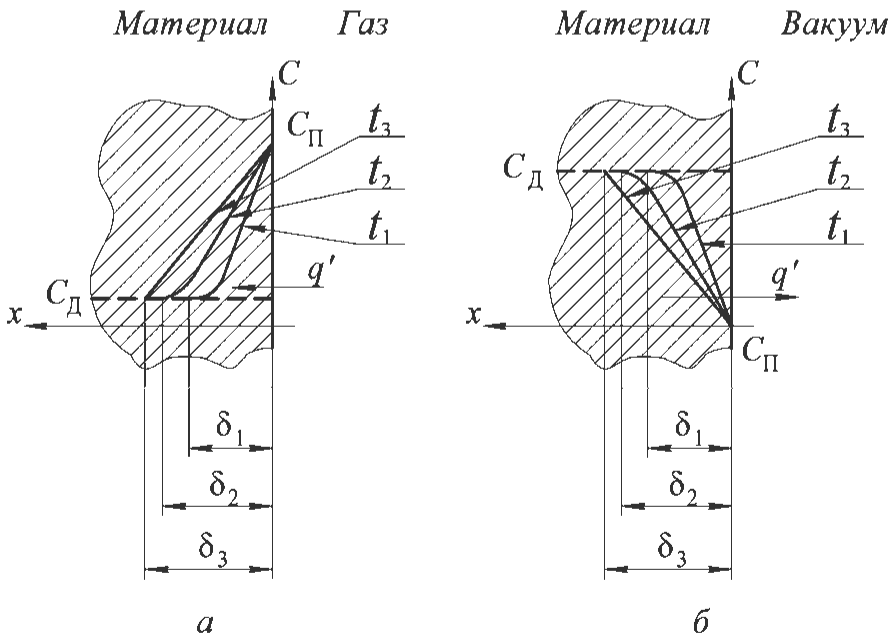


Рис. 1.7. Распределения объемной концентрации газа в твердом теле: а — газопоглощение; б — газовыделение

Толщину слоя δ твердого тела, насыщенного или обедненного газом, можно определить как

$$\delta = \sqrt{\pi Dt}. \quad (1.17)$$

Формулу (1.17) можно использовать для упрощенного расчета удельного потока газопоглощения или газовыделения при нестационарном процессе диффузии газа в твердом теле:

$$q' = D \frac{C_0 - C_1}{\sqrt{\pi Dt}},$$

где C_0, C_1 — значения максимальной и минимальной установившейся объемной концентрации растворенного газа в твердом теле (например, для процесса газовыделения на рис. 1.7, б $C_0 = C_d$; $C_1 = C_n$).

Таким образом, рассмотренные в этой главе закономерности процессов в вакуумной среде и на границе «вакуум — твердое тело» позволяют рассчитывать основные параметры вакуумных систем: давление, скорость и длину свободного пробега молекул газа, быстроту откачки рабочей камеры, определять критерии оценки степеней вакуума.

Контрольные вопросы и задания

1. Какие единицы измерения давления p газа применяются в вакуумных технологиях?
2. Чем определяется максимальная теоретическая быстрота действия вакуумных насосов?
3. Каким образом длина свободного пробега молекул зависит от температуры и давления газа?
4. Назовите критерии степеней вакуума.
5. Как зависит коэффициент растворимости S газа в твердом теле от температуры T ?
6. В каком состоянии — атомарном или молекулярном — растворяется газ в металлах и неметаллах?
7. Как зависит объемная концентрация C газа в твердом теле от температуры T и давления p ?

8. Как взаимосвязаны основные параметры вакуумной системы: быстрота откачки S_0 вакуумной камеры, быстрота действия S_n насоса и проводимость U вакуумного трубопровода?

9. Как зависит удельный поток газа q' через тонкую стенку от температуры t и давления газа p ?

10. Каким образом определить толщину слоя твердого тела, насыщенного или обедненного газом?

2. ТЕПЛОПЕРЕДАЧА

Изучив материал этой главы, студент должен иметь представление о видах, свойствах и возможностях процессов теплопередачи применительно к их использованию в электронных технологиях.

Студент должен знать:

- основные законы теплопередачи;
- понятия и определения применяемых физических величин;
- методы преобразования фундаментальных физических законов в выражения, используемые для расчета теплопередачи в конкретных расчетных схемах.

Студент должен уметь:

- выделять и анализировать физические закономерности, определяющие параметры процессов теплопередачи для исследуемых электронных технологий;
- моделировать требуемые процессы теплопередачи для оценки и выбора технологических параметров необходимых режимов нагрева, выдержки и охлаждения;
- оценивать применимость различных нагревательных устройств для реализации заданных термических процессов.

Студент должен иметь навыки:

- применения математических методов и аппарата современной физики для анализа процессов теплопередачи в электронных технологиях;
- анализа и использования современных источников научно-технической информации, освещающих применение процессов теплопередачи в электронных технологиях.

Значимую роль в производстве изделий электронной техники играют термические процессы, т. е. процессы, при которых происходит заданное изменение температуры обрабатываемого изделия во времени.

Для электровакуумных приборов прежде всего следует выделить технологический процесс термовакуумной обработки, при котором обезгаживание стеклооболочки осуществляется при температуре около 300...450 °С, обезгаживание металлических деталей внутренней арматуры — при нагреве до температуры 800...900 °С. Активирование оксидного катода требует разогрева до температуры 1100...1200 °С, формирование мультищелочных катодов проводят на входных экранах электронно-оптических преобразователей, нагретых до температуры 180...200 °С.

Для полупроводникового и микроэлектронного производства термические процессы также являются неотъемлемой частью технологии. Например, диффузионное легирование материалов или термическое окисление кремния предполагает нагрев до температуры около 1200 °С, эпитаксиальное наращивание слоев кремния, германия, соединений $A_{III}B_V$ — нагрев до температуры 1250 °С, процессы осаждения металлических (W, Mo, Ni, Al), полупроводниковых и изолирующих (SiO_2 , Si_3N_4 , Al_2O_3 и др.) слоев из газовой фазы проводятся при нагреве до температуры 300...700 °С, отжиг дефектов кристаллических структур происходит при температуре 400...1300 °С. Планарная технология изготовления полупроводниковых структур, по существу, состоит из чередования термических и фотолитографических процессов.

Большое количество термических процессов используется в качестве фоново-вспомогательных, обеспечивающих требуемое качество изделия. К ним относятся нагрев подложек при вакуумном напылении и вжигании металлических контактов, отжиг для снятия внутренних напряжений в стекле, отжиг для снятия радиационных дефектов после имплантации, различные прецизионные пайки, термообработка при герметизации готовых приборов компаундами и стеклами, низкотемпературные (150...200 °С) процессы сушки, в том числе процессы сушки пленок фоторезистов.

Для расчета или оценки параметров нагревательного устройства, обеспечивающих заданный режим термообработки объекта, необходимо знать физические основы теплопередачи.

Теплопередача (теплообмен) — учение о самопроизвольных необратимых процессах распространения теплоты в пространстве.

Процесс распространения теплоты — процесс обмена внутренней энергией между отдельными элементами тела или областями рассматриваемой системы тел.

Существуют три основных способа теплопередачи:

- *теплопередача теплопроводностью* — процесс молекулярного переноса теплоты в телах или между ними за счет разности температур в рассматриваемой системе тел (*контактная теплопередача*);

- *конвекция* — процесс переноса теплоты при перемещении объемов жидкости или газа в пространстве, неразрывно связанный с переносом самой этой жидкости или газа из области с одной температурой в область с другой температурой (*теплопередача с переносом среды*);

- *тепловое излучение* — процесс переноса теплоты с помощью электромагнитных волн, обусловленный температурой и оптическими свойствами излучающего тела (*бесконтактная теплопередача*). Теплообмен излучением осуществляется одновременным последовательным протеканием трех процессов: превращение внутренней энергии в энергию излучения, перенос излучения, поглощение излучения веществом.

При анализе теплопередачи в процессах термической обработки выделим две ее составляющие:

- 1) распространение теплоты внутри нагреваемого образца (материала, заготовки, детали, узла) — в этом случае для расчета изменения температуры во времени в заданной точке на поверхности или внутри обрабатываемого объекта базовой является *теория теплопроводности*;

- 2) передача теплоты от источника (нагревательного элемента) к нагреваемому объекту — при термообработке в условиях вакуума чаще всего используются различные инфракрасные нагреватели (спирали, металлические панели, трубчатые электронагреватели (ТЭН), пленочные нагреватели, металлические, металлокерамические, карборундовые электронагреватели, трубчатые кварцевые термоизлучатели, дуговые ксеноновые термоизлучатели (ДКсТ), галогеновые лампы накаливания (ГЛН) и др.), при рас-

чете теплоотдачи от которых следует использовать *теорию теплообмена излучением*.

2.1. Теплопередача теплопроводностью

При описании процессов теплопередачи теплопроводностью для вывода соотношения между характеристическими параметрами используется феноменологический подход, при котором вещество рассматривается как сплошная среда, игнорируется его микроскопическая структура и устанавливаются общие связи между параметрами, характеризующими исследуемое явление в целом.

Температурное поле

Аналитическое исследование теплопроводности сводится к изучению пространственно-временного изменения температуры в рассматриваемой системе тел, т. е. к нахождению зависимости типа

$$t = f(x, y, z, \tau).$$

Уравнение такого рода представляет собой математическое описание температурного поля рассматриваемой системы тел.

Температурное поле — это совокупность значений температуры во всех точках рассматриваемой системы тел для любого момента времени.

Следует различать стационарное и нестационарное температурные поля.

Математическое описание *нестационарного* температурного поля (температура изменяется с течением времени от одной точки к другой) характеризует *неустановившийся* тепловой режим и является наиболее общим случаем температурного поля.

В случае, если температура изменяется от одной точки к другой, но не меняется с течением времени, имеет место *стационарное* температурное поле, характеризующее *установившийся* тепловой режим. Математическое описание пространственного стационарного температурного поля имеет вид

$$t = f_1(x, y, z); \quad dt/d\tau = 0.$$

Температурный градиент

Изотермической называется поверхность, соединяющая точки тела, имеющие равную температуру (геометрическое место точек, имеющих одинаковую температуру, в температурном поле — рис. 2.1).

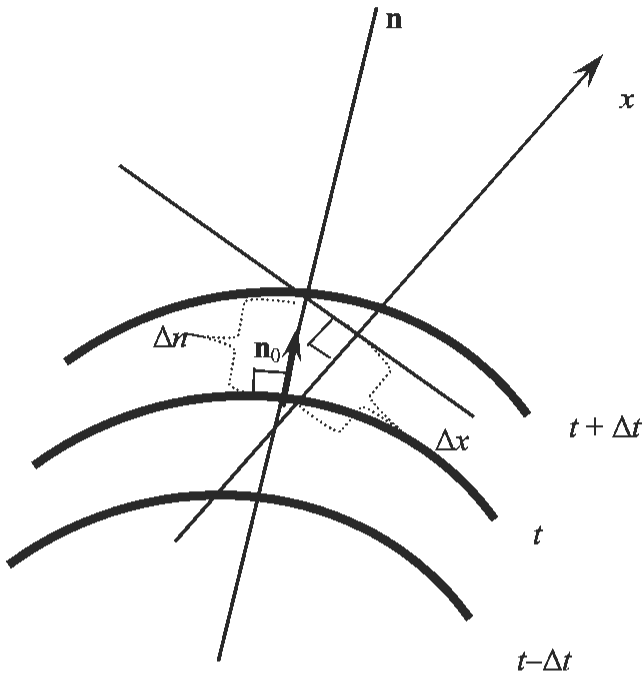


Рис. 2.1. Изотермы

Температурным градиентом называется вектор, направленный по нормали к изотермической поверхности в сторону увеличения температуры и численно равный производной температуры по этому направлению:

$$\mathbf{grad} t = \mathbf{n}_0 \frac{\partial t}{\partial n},$$

где \mathbf{n}_0 — единичный вектор, нормальный к изотермической поверхности и направленный в сторону увеличения температуры.

Проекция вектора градиента температуры на ось x в декартовой системе координат равна

$$\mathbf{grad} t_x = \frac{\partial t}{\partial n} \cos(\widehat{n, x}) = \frac{\partial t}{\partial x}.$$

Аналогичные выражения имеют место для осей y и z .

Скалярная величина температурного градиента $\partial t/\partial n$ также называется температурным градиентом.

Закон Фурье

Введем следующие обозначения:

dF — элемент изотермической поверхности, м^2 ;

$d\tau$ — элементарный промежуток времени, с ;

dQ_τ — количество теплоты, проходящее через dF за $d\tau$, Вт .

Тогда согласно гипотезе Фурье

$$dQ_\tau = -\lambda \frac{\partial t}{\partial n} dF d\tau,$$

где λ — коэффициент пропорциональности — физический параметр вещества, который характеризует способность вещества проводить теплоту и называется коэффициентом теплопроводности. Значения коэффициента теплопроводности для некоторых веществ приведены в табл. 2.1.

Количество теплоты, проходящей в единицу времени через единицу площади изотермической поверхности, называется плотностью теплового потока, $\text{Вт}/\text{м}^2$:

$$q = \frac{dQ_\tau}{dF d\tau}.$$

Из гипотезы Фурье следует, что

$$\mathbf{q} = -\mathbf{n}_0 \lambda \frac{\partial t}{\partial n}.$$

Для скалярных величин

$$q = -\lambda \frac{\partial t}{\partial n}.$$

Эти уравнения являются математическими записями основного закона теплопроводности — *закона Фурье*: *плотность теплового потока пропорциональна градиенту температуры*.

Значения коэффициента теплопроводности λ

Наименование вещества	Температура, °С	Коэффициент теплопроводности λ , Вт/(м·К)
Газы:		0,006 – 0,600
воздух (1 бар)	0	0,0244
	200	0,0393
	1200	0,0915
азот (0,5 – 500 бар)	0 – 300	0,0242
кислород (0,5 – 500 бар)	535	0,0245
диоксид углерода (0,5 – 300 бар)	< 200	0,0149
гелий (0,5 – 1000 бар)	-268 – 600	0,142
водород (1 – 500 бар)	20 – 200	0,172
Капельные жидкости:		0,07 – 0,70
вода	0	0,56
	40	0,64
	100	0,68
моторное масло	27	0,15
Твердые тела:		
теплоизоляционные материалы	—	< 0,25
обыкновенное стекло	20	0,74
кварцевое стекло	400	1,76
	800	2,40
	1200	3,05
медь чистая	0	396
алюминий	27	204
сплавы	27	16 – 130
пластики	27	0,2 – 0,3
Стекланная вата	88	0,051 – 0,059

Количество теплоты Q , Вт, проходящее в единицу времени через изотермическую поверхность F , называется *тепловым потоком*:

$$Q = \int_F q dF = \int_F \lambda \frac{dt}{dn} dF.$$

Количество теплоты Q_τ , Дж, проходящее за время τ через изотермическую поверхность F , называется *полным количеством теплоты*:

$$Q_\tau = \int_0^\tau \int_F q dF d\tau = \int_0^\tau \int_F \lambda \frac{dt}{dn} dF d\tau.$$

Дифференциальное уравнение теплопроводности

Решение задач теплопередачи теплопроводностью обычно заключается в нахождении температурного поля и тепловых потоков в рассматриваемой системе тел посредством интегрирования соответствующего *дифференциального уравнения теплопроводности* при заданных конкретных *условиях однозначности*. Это уравнение устанавливает связь между временным и пространственным изменениями температуры t в любой точке тела, в котором происходит теплопередача теплопроводностью.

При выводе *дифференциального уравнения теплопроводности* используются следующие допущения:

- тело однородно и изотропно;
- физические параметры тела постоянны;
- внутренние источники теплоты, характеризующиеся *мощностью внутренних источников теплоты* $q_v = (x, y, z, \tau)$, Вт/м³, распределены равномерно;
- температурная деформация объема незначительна по сравнению с самим этим объемом.

Вывод уравнения проводят на базе закона сохранения энергии для элементарного объема $dv = dx dy dz$ за элементарный промежуток времени $d\tau$.

Для твердых тел, в которых теплопередача описывается законом Фурье, уравнение имеет вид

$$\frac{\partial t}{\partial \tau} = \frac{1}{c\rho} \left[\frac{\partial}{\partial x} \left(\lambda \frac{\partial t}{\partial x} \right) + \frac{\partial}{\partial y} \left(\lambda \frac{\partial t}{\partial y} \right) + \frac{\partial}{\partial z} \left(\lambda \frac{\partial t}{\partial z} \right) \right] + \frac{q_v}{c\rho}$$

или

$$\frac{\partial t}{\partial \tau} = \frac{1}{c\rho} \operatorname{div}(\lambda \operatorname{grad} t) + \frac{q_v}{c\rho},$$

где λ — коэффициент теплопроводности, Вт/(м·К); c — теплоемкость, Дж/(кг·К); ρ — плотность материала твердого тела, кг/м³.

Если, как и предполагалось при выводе, принять $\lambda = \text{const}$, получим уравнение

$$\frac{\partial t}{\partial \tau} = \frac{\lambda}{c\rho} \left(\frac{\partial^2 t}{\partial x^2} + \frac{\partial^2 t}{\partial y^2} + \frac{\partial^2 t}{\partial z^2} \right) + \frac{q_v}{c\rho}.$$

Введем обозначения:

$$\frac{\partial^2 t}{\partial x^2} + \frac{\partial^2 t}{\partial y^2} + \frac{\partial^2 t}{\partial z^2} = \nabla^2 t; \quad \frac{\lambda}{c\rho} = a,$$

где ∇^2 — оператор Лапласа (в декартовых координатах); a — коэффициент температуропроводности, м²/с (мера теплоинерционных свойств тела, теплофизический параметр вещества).

В новых обозначениях уравнение приобретает вид

$$\frac{\partial t}{\partial \tau} = a \nabla^2 t + \frac{q_v}{c\rho}.$$

Частные случаи:

1) при отсутствии внутренних источников ($q_v = 0$) — уравнение Фурье:

$$\frac{\partial t}{\partial \tau} = a \nabla^2 t;$$

2) стационарное температурное поле ($\partial t / \partial \tau = 0$) — уравнение Пуассона:

$$\frac{\partial^2 t}{\partial x^2} + \frac{\partial^2 t}{\partial y^2} + \frac{\partial^2 t}{\partial z^2} + \frac{q_v}{\lambda} = 0;$$

3) стационарное температурное поле при отсутствии внутренних источников — уравнение Лапласа:

$$\frac{\partial^2 t}{\partial x^2} + \frac{\partial^2 t}{\partial y^2} + \frac{\partial^2 t}{\partial z^2} = 0.$$

Условия однозначности

Для задания полного математического описания конкретного процесса теплопроводности кроме дифференциального уравнения требуется задать частные особенности анализируемого процесса. Для тепловых процессов частные условия — это *условия однозначности*.

Геометрические условия — задание формы и размеров тела или тел в рассматриваемой системе.

Физические условия — определение значимых в рамках решаемой задачи теплофизических параметров среды и тел (λ , c , ρ , мощности внутренних источников теплоты как функции координат и времени $q_v = f(x, y, z, \tau)$, и др.).

Временные (начальные) условия — задание распределения температуры в изучаемом теле в начальный момент времени $t_0 = f(x, y, z)$ при $\tau = 0$ (например, $t_0 = \text{const}$).

Граничные (краевые) условия — описание особенностей протекания процесса на граничных поверхностях рассматриваемого тела:

а) граничные условия первого рода — известно распределение температуры на поверхности тела для каждого момента времени $t_c = f(x, y, z, \tau)$, например $t_c = \text{const}$ (использование индекса «с» является общепринятым для обозначения температуры поверхности);

б) граничные условия второго рода — задаются значения теплового потока для любой точки поверхности тела для каждого момента времени $q_c = f(x, y, z, \tau)$, простейший случай $q_c = \text{const}$;

в) граничные условия третьего рода — задаются температура окружающей среды и закон теплообмена между поверхностью тела и окружающей средой.

Процесс теплообмена между поверхностью и средой описывается *законом Ньютона — Рихмана*, согласно которому количество теплоты, отдаваемое единицей поверхности тела в окружающую среду в единицу времени (рис. 2.2), пропорционально разности температур поверхности тела t_c и окружающей среды t_α :

$$q = \alpha(t_c - t_\alpha).$$

Коэффициент пропорциональности α — это *коэффициент теплоотдачи*, характеризующий интенсивность теплообмена, Вт/(м²·К).

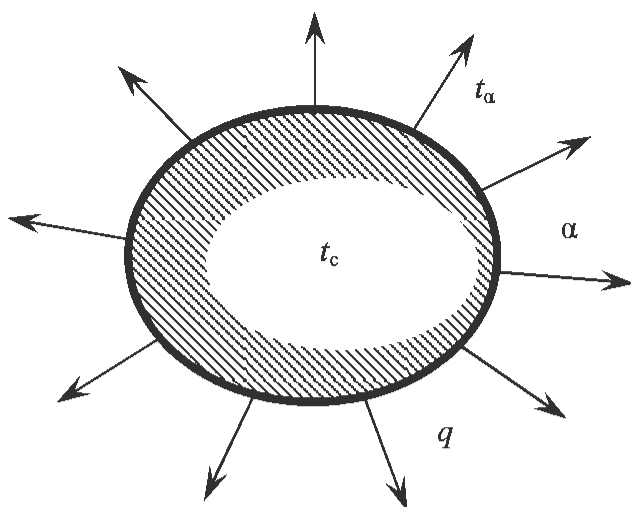


Рис. 2.2. Тепловое взаимодействие тела и окружающей среды

Согласно закону сохранения энергии количество теплоты, отводимое единицей поверхности тела вследствие теплообмена, равно количеству теплоты, подводимому к единице поверхности в единицу времени из внутренних объемов тела вследствие теплопроводности:

$$\alpha(t_c - t_a) = -\lambda \left. \frac{\partial t}{\partial n} \right|_c.$$

Решение задач теплопроводности заключается в определении аналитического описания температурного поля для рассматриваемой системы тел, зная которое, можно определить и тепловые потоки между телами или элементами этой системы тел. Получение такого описания возможно для простых расчетных схем. Для решения более сложных задач детально разработаны численные методы.

2.2. Теплообмен излучением

Виды потоков излучения

Энергия излучения, испускаемая произвольной поверхностью F по всевозможным направлениям полупространства по всем длинам волн в единицу времени, называется *интегральным (или полным) потоком излучения* Q , Вт.

Энергия излучения, испускаемая произвольной поверхностью F по всевозможным направлениям полупространства в интервале длин волн от λ до $\lambda + d\lambda$ в единицу времени, называется *поток* *спектрального (монохроматического) излучения* Q_λ , Вт/м.

Интегральный поток излучения, испускаемый с единицы поверхности, называется *поверхностной интегральной плотностью потока излучения* E , Вт/м²:

$$E = \frac{dQ}{dF},$$

где dQ — интегральный поток, испускаемый элементарной площадкой dF .

Поверхностная интегральная плотность потока излучения, испускаемого в интервале длин волн от λ до $\lambda + d\lambda$, отнесенная к величине этого интервала длин волн, называется *спектральной плотностью потока излучения* E_λ , Вт/м³:

$$E_\lambda = \frac{dE}{d\lambda}.$$

Энергия интегрального излучения, испускаемая единицей элементарной площадки dF в единицу времени в пределах единичного элементарного телесного угла $d\omega$ в направлении \mathbf{l} , определяемом углом ψ между этим направлением и направлением нормали \mathbf{n} к поверхности (рис. 2.3), называется *угловой плотностью интегрального излучения* I_ψ , Вт/(м²·ср):

$$I_\psi = \frac{d^2Q}{dFd\omega} = \frac{dE}{d\omega}.$$

Энергия спектрального (монохроматического) излучения, испускаемая единицей элементарной площадки dF в единицу времени в пределах единичного элементарного телесного угла $d\omega$ в направлении \mathbf{l} , определяемом углом ψ между этим направлением и направлением нормали \mathbf{n} к поверхности, называется *угловой плотностью спектрального (монохроматического) излучения* $I_{\psi\lambda}$, Вт/(м³·ср):

$$I_{\psi\lambda} = \frac{d^2Q_\lambda}{dFd\omega} = \frac{dE_\lambda}{d\omega}.$$

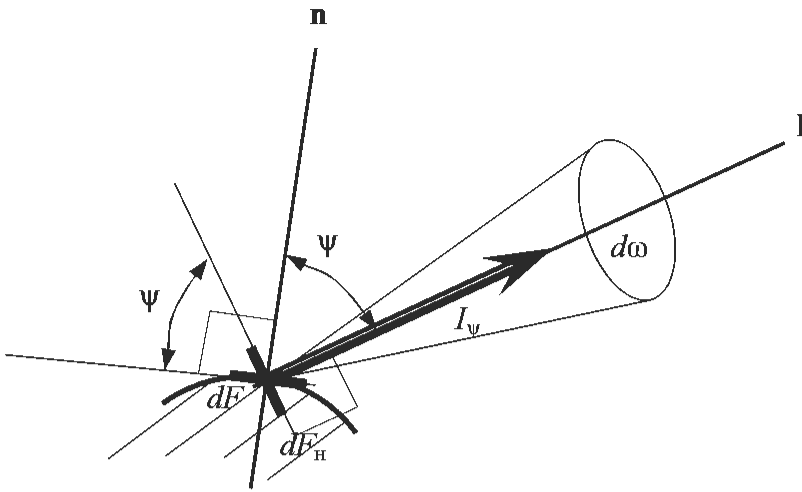


Рис. 2.3. К определению угловой плотности и интенсивности излучения

Энергия интегрального излучения, испускаемая элементарной площадкой в единицу времени в пределах единичного элементарного телесного угла $d\omega$, отнесенная к площади dF_n проекции этой площадки на плоскость, ортогональную к направлению излучения в направлении l , определяемом углом ψ между этим направлением и направлением нормали n к поверхности, называется *интенсивностью (яркостью) интегрального излучения* I , Вт/(м²·ср):

$$I = \frac{d^2Q}{dF_n d\omega} = \frac{dE}{\cos \psi d\omega} = \frac{I_\psi}{\cos \psi}.$$

Энергия спектрального (монохроматического) излучения, испускаемая элементарной площадкой в единицу времени в пределах единичного элементарного телесного угла $d\omega$, отнесенная к площади dF_n проекции этой площадки на плоскость, ортогональную к направлению излучения в направлении l , определяемом углом ψ между этим направлением и направлением нормали n к поверхности, называется *интенсивностью (яркостью) спектрального (монохроматического) излучения* I_λ , Вт/(м²·ср):

$$I_\lambda = \frac{d^2Q_\lambda}{dF_n d\omega} = \frac{dE_\lambda}{\cos \psi d\omega} = \frac{I_{\psi\lambda}}{\cos \psi}.$$

Установим взаимосвязь между интенсивностью (яркостью) и плотностью потока полусферического излучения. Плотность пото-

ка излучения в данном направлении из последнего выражения можно представить как

$$dE = I \cos \psi d\omega.$$

Телесный угол $d\omega$ — это угол, под которым из какой-либо точки элементарной площадки одного тела видна элементарная площадка другого тела:

$$d\omega = \frac{dF}{r^2},$$

где dF — элементарная площадка, вырезаемая телесным углом $d\omega$ на поверхности сферы радиусом r (рис. 2.4). Площадь поверхности этой площадки может быть представлена произведением двух элементарных дуг $r d\psi$ и $r \sin \psi d\theta$, где ψ — угол, дополнительный к углу широты, а θ — угол долготы. Отсюда получаем

$$d\omega = \sin \psi d\psi d\theta.$$

Подставив полученное соотношение в выражение для E и проинтегрировав, получим значение плотности теплового потока излучения:

$$E = I \int_0^{2\pi} d\theta \int_0^{\pi/2} \sin \psi \cos \psi d\psi = I\pi.$$

Отсюда

$$I = \frac{E}{\pi},$$

где E — плотность потока интегрального полусферического излучения; I — интенсивность (яркость) излучения.

Полученное соотношение показывает, что для излучения, подчиняющегося *закону косинусов Ламберта*, интенсивность (яркость) излучения в направлении нормали к поверхности излучения в π раз меньше плотности потока интегрального излучения. Данная зависимость верна и для спектрального излучения, подчиняющегося *закону косинусов*.

Изотропное излучение — это излучение, характеризующееся одинаковой по всем направлениям интенсивностью.

Идеально диффузное излучение — это изотропное излучение, испускаемое поверхностью твердого тела.

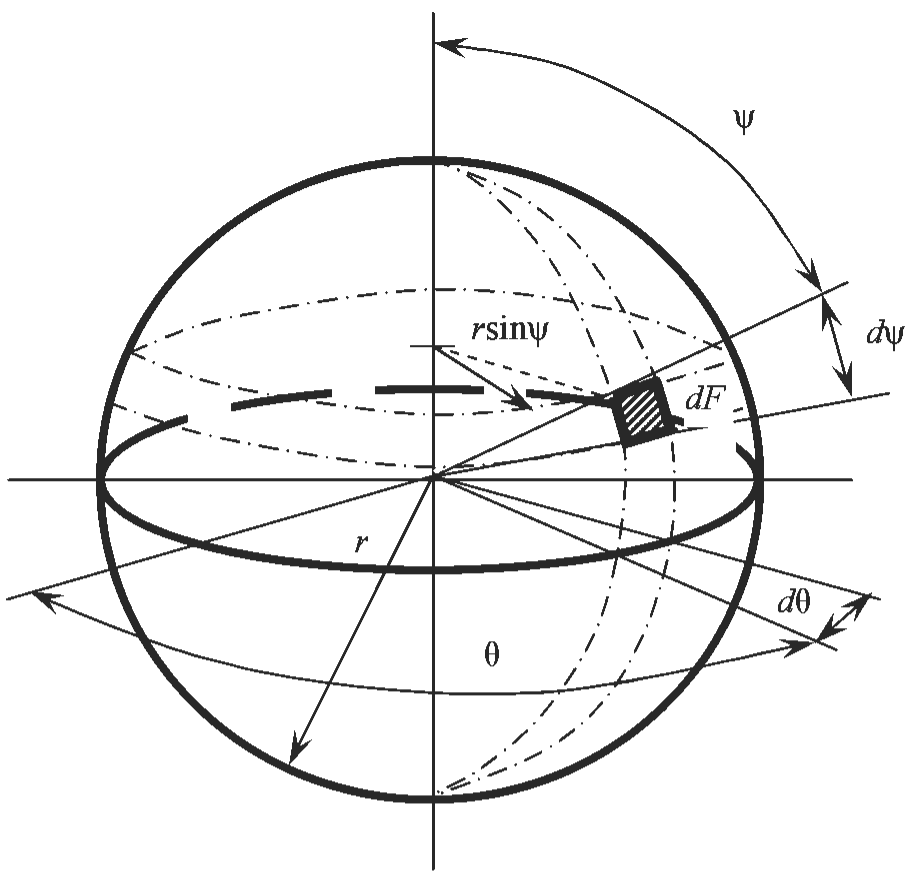


Рис. 2.4. К определению взаимосвязи между интенсивностью и плотностью потока полусферического излучения

Поток собственного излучения (Q , E) — это излучение, которое определяется только природой данного тела и его температурой.

Поток падающего излучения ($Q_{\text{пад}}$, $E_{\text{пад}}$) — это поток энергии излучения других тел в рассматриваемой системе тел, попадающий на поверхность данного тела. Этот поток, попадая на поверхность данного тела, частично поглощается, частично отражается, частично пропускается ею.

Поток поглощенного излучения — это часть падающего излучения, поглощенная данным телом; плотность потока поглощенного излучения $E_{\text{погл}}$, Вт/м²:

$$E_{\text{погл}} = AE_{\text{пад}},$$

где A — коэффициент поглощения тела.

Тела, поглощающие всю падающую на них энергию, называются *абсолютно черными* ($A = 1$).

Спектральный коэффициент поглощения (A_λ) в общем случае зависит от длины волны излучения. Тела, для которых спектральная поглощательная способность не зависит от длины волны падающего излучения, называются *серыми телами*. Для серых тел $A_\lambda = \text{const} < 1$, т. е. часть падающего потока излучения отражается и (или) пропускается этими телами.

Поток отраженного излучения — часть падающего потока излучения, которую поверхность данного тела отражает обратно окружающим его телам; плотность потока отраженного излучения $E_{\text{отр}}$, Вт/м²:

$$E_{\text{отр}} = RE_{\text{пад}},$$

где R — коэффициент отражения тела.

Поверхность, от которой отражается весь поток падающего излучения ($R = 1$) и отражение на которой происходит по законам геометрической оптики, называется *абсолютно зеркальной* (*блестящей*). Поверхность, от которой отражается весь поток падающего излучения ($R = 1$) при идеально диффузном отражении, равномерном по всем направлениям независимо от направления падающего на поверхность излучения, называется *абсолютно белой*.

Поток пропускаемого излучения — часть падающего потока излучения, проходящая сквозь тело; плотность потока пропускаемого излучения $E_{\text{проп}}$, Вт/м²:

$$E_{\text{проп}} = DE_{\text{пад}},$$

где D — коэффициент пропускания тела.

Тела, пропускающие весь поток падающего излучения ($D = 1$), называются *абсолютно прозрачными* или *диатермичными*. Тела, для которых $0 < D < 1$, называются *полупрозрачными*. Тела, совсем не пропускающие поток падающего излучения ($D = 0$), называются *непрозрачными*.

Совместный процесс взаимного испускания, поглощения, отражения и пропускания излучения в рассматриваемой системе тел называется *теплообменом излучением*.

Для тела, участвующего в теплообмене излучением с другими телами, по закону сохранения энергии имеет место уравнение теплового баланса

$$E_{\text{погл}} + E_{\text{отр}} + E_{\text{проп}} = E_{\text{пад}},$$

или

$$A + R + D = 1.$$

При измерении потока излучения от данного тела реально воспринимается не собственное, а так называемое *эффективное* излучение.

Плотностью потока эффективного излучения $E_{\text{эф}}$ называют суммарную плотность потоков собственного излучения и излучения, отраженного его поверхностью:

$$E_{\text{эф}} = E + E_{\text{отр}} = E + RE_{\text{пад}}.$$

Эффективное излучение зависит не только от физических свойств и температуры данного тела, но и от физических свойств и температуры окружающих его тел, а также от формы, размеров и относительного расположения тел в пространстве. Вследствие этого физические свойства эффективного и собственного излучений, в том числе и их спектры, различны.

Виды потоков излучения, характеризующих взаимодействие рассматриваемой поверхности с окружающей средой, представлены на рис. 2.5.

Направление теплообмена излучением в рассматриваемой системе тел определяется *потокom результирующего излучения*.

Значение потока результирующего излучения определяется разностью значений потоков получаемого и испускаемого данным телом излучений. Эта разность определяется в зависимости от расположения условной расчетной поверхности относительно данного тела.

При рассмотрении теплового баланса относительно поверхности, расположенной внутри тела вблизи его поверхности,

$$E_{\text{рез}} = E_{\text{погл}} - E = AE_{\text{пад}} - E. \quad (2.1)$$

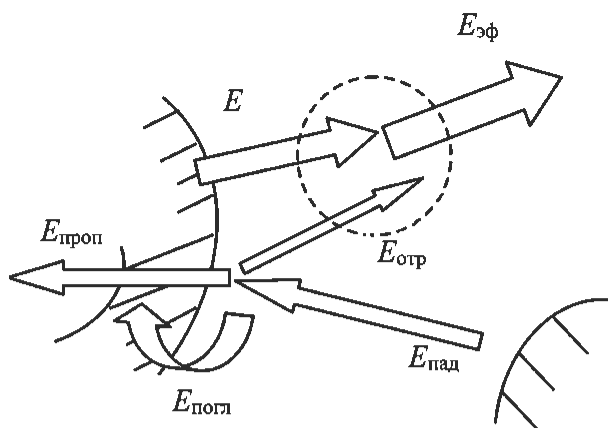


Рис. 2.5. Поток излучения, испускаемые телом

Баланс относительно поверхности, расположенной вне тела вблизи его поверхности, дает выражение

$$E_{\text{рез}} = E_{\text{пад}} - E_{\text{эф}}. \quad (2.2)$$

При рассмотрении теплообмена излучением относительно условной поверхности, находящейся на некотором удалении от поверхности данного тела, получаем наиболее общее выражение для плотности потока результирующего излучения:

$$E_{\text{рез}} = E_{\text{пад1}} - E_{\text{пад2}}.$$

Положительное значение потока результирующего излучения для данного тела означает его нагрев, отрицательное — остывание, равенство нулю — тепловое равновесие в рассматриваемой системе тел.

Можно установить зависимость между плотностями потоков результирующего и эффективного излучений, а именно:

– выражаем $E_{\text{эф}}$ из уравнения (2.2):

$$E_{\text{эф}} = E_{\text{пад}} - E_{\text{рез}}; \quad (2.3)$$

– выражаем $E_{\text{пад}}$ из уравнения (2.1):

$$E_{\text{пад}} = \frac{E + E_{\text{рез}}}{A}; \quad (2.4)$$

– подставляем выражение (2.4) в формулу (2.3):

$$E_{\text{эф}} = E_{\text{рез}} \left(\frac{1}{A} - 1 \right) + \frac{E}{A}. \quad (2.5)$$

В общем случае из выражения (2.5) получаем для потоков излучения

$$Q_{\text{эф}} = Q_{\text{рез}} \left(\frac{1}{A} - 1 \right) + \frac{EF}{A} = Q_{\text{рез}} \left(\frac{1}{A} - 1 \right) + E_0 F. \quad (2.6)$$

Частный случай для непрозрачных тел ($D = 0, A + R = 1$):

$$E_{\text{эф}} = \frac{E - RE_{\text{рез}}}{A} = E_0 - \frac{R}{A} E_{\text{рез}};$$

для потоков излучения:

$$Q_{\text{эф}} = Q_0 - \frac{R}{A} Q_{\text{рез}} = E_0 F - \frac{R}{A} Q_{\text{рез}}.$$

Законы теплового излучения

При длинах волн $\lambda = 0,4 \dots 800$ мкм излучение обладает свойствами теплового излучения.

Законы теплового излучения получены для абсолютно черных тел (АЧТ) в условиях термодинамического равновесия.

Закон Планка (1900 г.). Этот закон устанавливает зависимость спектральной плотности потока собственного излучения АЧТ от температуры и длины волны излучения:

$$E_{0\lambda} = \frac{2\pi c_1 / \lambda^5}{\exp[c_2 / (\lambda T)] - 1},$$

где λ — длина волны, м; $c_1 = 5,944 \cdot 10^{-17}$ — первая константа излучения, Вт·м²; $c_2 = 1,4388 \cdot 10^{-2}$ — вторая константа излучения, м·К; \exp — основание натурального логарифма; T — абсолютная температура тела, К.

Таким образом, испускание энергии происходит неравномерно по длинам волн и зависит от температуры. Каждой длине волны соответствует свое значение $E_{0\lambda}$. Каждой температуре отвечает своя кривая распределения $E_{0\lambda}(\lambda)$, имеющая ярко выраженный мак-

симум, причем при $\lambda > 0$ и $\lambda < \infty$ значения $E_{0\lambda} > 0$. Повышение температуры ведет к значительному увеличению энергии излучения.

Закон Планка имеет два предельных случая, описываемых законом Рэлея — Джинса и законом Вина.

Закон Рэлея — Джинса. Имеет место как частный случай закона Планка при условии, что произведение λT велико по сравнению со значением постоянной c_2 .

При этом ограничиваются двумя слагаемыми разложения экспоненциальной функции в ряд по степеням $c_2/(\lambda T)$:

$$\exp\left(\frac{c_2}{\lambda T}\right) = 1 + \frac{1}{1!} \frac{c_2}{\lambda T} + \frac{1}{2!} \frac{c_2^2}{(\lambda T)^2} + \dots$$

Получаем следующее выражение вместо выражения закона Планка:

$$E_{0\lambda} = \frac{2\pi c_1 T}{c_2 \lambda^4}.$$

Закон Вина. Частный случай закона Планка, который соответствует малому значению λT по сравнению с постоянной c_2 . При этом можно пренебречь единицей в выражении закона Планка, тогда оно переходит в зависимость, называемую законом Вина:

$$E_{0\lambda} = \frac{2\pi c_1}{\lambda^5} \exp\left(\frac{-c_2}{\lambda T}\right).$$

Закон смещения Вина. Этот закон описывает положение максимума спектральной плотности потока излучения.

Значение длины волны, соответствующее максимуму излучения, получают как экстремум функции спектральной плотности потока излучения из закона Планка. Для этого находят производную функции по длине волны, приравнявая которую к нулю, получают трансцендентное уравнение

$$\exp\left(\frac{-c_2}{\lambda_{\max} T}\right) + \frac{c_2}{5\lambda_{\max} T} - 1 = 0.$$

Решение этого уравнения

$$\lambda_{\max} T = 2,8978 \cdot 10^{-3} \text{ м}\cdot\text{К},$$

где λ_{\max} — длина волны, которой соответствует максимальная спектральная плотность потока излучения.

Таким образом, закон смещения Вина гласит, что *максимальное значение спектральной плотности потока излучения с повышением температуры смещается в сторону более коротких длин волн.*

Закон Стефана — Больцмана. Этот закон устанавливает зависимость поверхностной интегральной плотности потока излучения от температуры. Впервые эта зависимость экспериментально была установлена Стефаном в 1879 г., затем теоретически получена Больцманом в 1884 г. Выражение закона Стефана — Больцмана может быть также получено интегрированием значения спектральной плотности потока излучения из закона Планка по всем длинам волн.

Поверхностная интегральная плотность потока излучения как функция температуры описывается следующим выражением:

$$E_0 = \int_0^{\infty} E_{\lambda} d\lambda = \sigma_0 T^4,$$

где $\sigma_0 = 5,6703 \cdot 10^{-8}$ Вт/(м²·К⁴) — *постоянная Стефана — Больцмана.*

На практике это уравнение используют в более удобном виде:

$$E_0 = c_0 (T/100)^4,$$

где $c_0 = 5,67$ Вт/(м²·К⁴) — *излучательная способность АЧТ.*

Таким образом, *поверхностная интегральная плотность потока излучения изменяется пропорционально четвертой степени абсолютной температуры.*

Применительно к серым телам закон Стефана — Больцмана выглядит следующим образом:

$$E = \epsilon E_0 = \epsilon c_0 (T/100)^4 = c (T/100)^4,$$

где $\epsilon = E/E_0 = c/c_0$ — *коэффициент теплового излучения (коэффициент черноты); c — излучательная способность серого тела.*

Отсюда понятен физический смысл коэффициента теплового излучения, определяемого как отношение поверхностной интегральной плотности потока собственного излучения серого тела

к его значению для абсолютно черного тела при той же температуре.

Закон Кирхгофа. Этот закон связывает между собой излучательную и поглощательную способности любого тела по отношению к излучению АЧТ. Для интегрального излучения твердых тел математическая формулировка закона Кирхгофа выглядит следующим образом:

$$E_1/A_1 = E_2/A_2 = \dots = E_n/A_n = E_0/1$$

В данном выражении A_i — коэффициенты поглощения серых тел; E_i — плотности потоков интегрального излучения этих серых тел; 1 и E_0 — коэффициент поглощения и плотность потока интегрального излучения АЧТ при одной и той же температуре.

Используя выражения для плотности потока интегрального излучения из закона Стефана — Больцмана

– для серых тел

$$E_i = \varepsilon_i c_0 (T/100)^4,$$

– для АЧТ

$$E_0 = c_0 (T/100)^4,$$

получаем после подстановки другую форму записи закона Кирхгофа:

$$c_1/A_1 = c_2/A_2 = \dots = c_n/A_n = c_0/1$$

или

$$\varepsilon_1/A_1 = \varepsilon_2/A_2 = \dots = \varepsilon_n/A_n = 1.$$

Последнее выражение означает, что для любого серого тела коэффициент теплового излучения (*коэффициент черноты*) и коэффициент поглощения равны между собой.

Закон косинусов Ламберта. Этот закон (1760 г.) позволяет определить энергию излучения, испускаемую телом по отдельным направлениям полупространства.

Закон косинусов Ламберта гласит: *поток излучения абсолютно черного тела в данном направлении пропорционален потоку излучения в направлении нормали к поверхности данного тела и косинусу угла между данным направлением и направлением нормали.*

При использовании угловой плотности потока излучения закон Ламберта выглядит следующим образом:

$$I_{\psi} = I_n \cos \psi,$$

где I_{ψ} и I_n — угловые плотности потоков интегрального излучения соответственно в направлении, определяемом углом ψ , и в направлении нормали к поверхности.

При переходе к интенсивности (яркости) излучения, для которой

$$I = I_{\psi} / \cos \psi,$$

получаем соотношение

$$I = I_{\psi} / \cos \psi = I_n = \text{idem},$$

из которого следует, что *если излучение подчиняется закону Ламберта, то интенсивность (яркость) данного потока излучения не зависит от направления и является постоянной величиной.*

Теплообмен излучением между двумя телами, произвольно расположенными в пространстве

Теплообмен излучением между двумя элементарными черными поверхностями. Найдем значение потока результирующего излучения при теплообмене излучением между черными элементарными площадками dF_1 и dF_2 , положение в пространстве которых характеризуется нормальными \mathbf{n}_1 и \mathbf{n}_2 (рис. 2.6), с учетом того, что плотности потоков их собственного излучения I_{ψ_1} и I_{ψ_2} являются постоянными вдоль поверхности каждого тела, температуры поверхностей T_1 и T_2 , поверхности расположены на расстоянии r друг от друга.

Для этого рассчитаем потоки излучения $d^2 Q_{\text{пад}1}$, падающего с поверхности dF_1 на поверхность dF_2 , и излучения $d^2 Q_{\text{пад}2}$, падающего с поверхности dF_2 на поверхность dF_1 , а результирующий поток излучения найдем как разность этих потоков.

Поток с площадки dF_1 на dF_2 :

$$d^2 Q_{\text{пад}1} = I_{\psi_1} d\omega_1 dF_1 = I_1 \cos \psi_1 d\omega_1 dF_1;$$

$$d\omega_1 = \frac{dF_2 \cos \psi_2}{r^2};$$

$$I_1 = \frac{E_{01}}{\pi};$$

$$d^2Q_{\text{пад1}} = E_{01} \frac{\cos \psi_1 \cos \psi_2}{\pi r^2} dF_1 dF_2.$$

Поток с площадки с dF_2 на dF_1 :

$$d^2Q_{\text{пад2}} = I_{\psi_2} d\omega_2 dF_2 = I_2 \cos \psi_2 d\omega_2 dF_2;$$

$$d\omega_2 = \frac{dF_1 \cos \psi_1}{r^2};$$

$$I_2 = \frac{E_{02}}{\pi};$$

$$d^2Q_{\text{пад2}} = E_{02} \frac{\cos \psi_1 \cos \psi_2}{\pi r^2} dF_1 dF_2.$$

Пусть для определенности $T_1 > T_2$, тогда

$$d^2Q_{\text{рез}} = d^2Q_{\text{пад1}} - d^2Q_{\text{пад2}};$$

$$d^2Q_{\text{рез}} = \sigma_0 \frac{\cos \psi_1 \cos \psi_2}{\pi r^2} dF_1 dF_2 (T_1^4 - T_2^4).$$

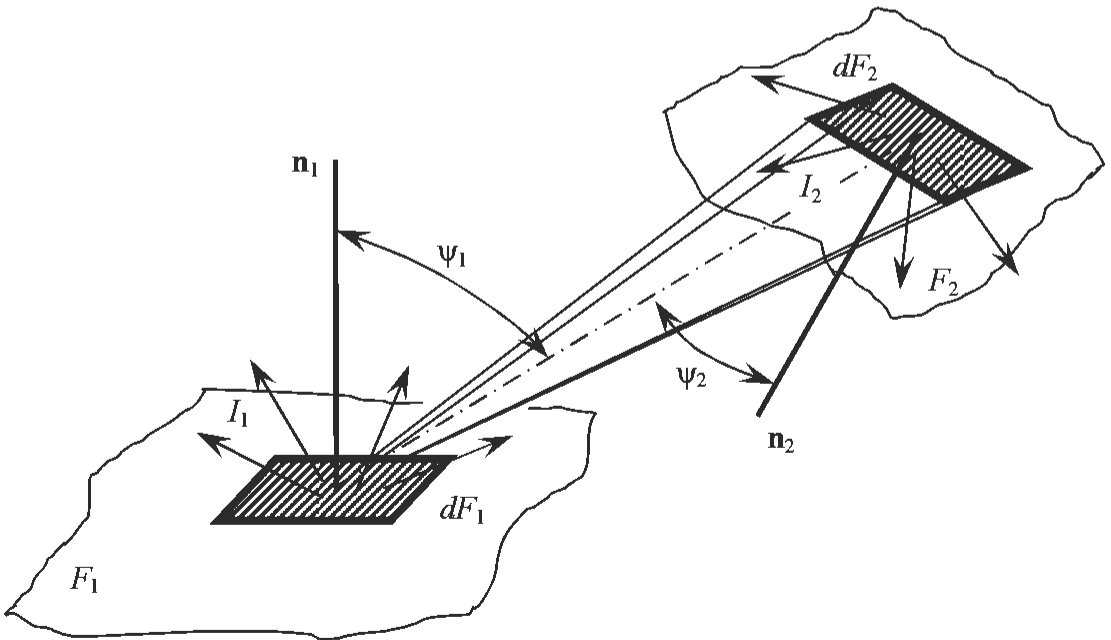


Рис. 2.6. Теплообмен излучением между двумя элементарными поверхностями

Введем обозначения *элементарных угловых коэффициентов*:

$$d\varphi_{1,2} = \frac{\cos \psi_1 \cos \psi_2}{\pi r^2} dF_2;$$

$$d\varphi_{2,1} = \frac{\cos \psi_1 \cos \psi_2}{\pi r^2} dF_1.$$

Тогда

$$d^2 Q_{\text{пад1}} = E_{01} dF_1 d\varphi_{1,2} = dQ_1 d\varphi_{1,2};$$

$$d^2 Q_{\text{пад2}} = E_{02} dF_2 d\varphi_{2,1} = dQ_2 d\varphi_{2,1}.$$

Здесь dQ_1 и dQ_2 — *полные элементарные потоки* собственного излучения, испускаемые площадками dF_1 и dF_2 во всех направле-

ниях полупространства; $d\varphi_{1,2} = \frac{d^2 Q_{\text{пад1}}}{dQ_1}$, $d\varphi_{2,1} = \frac{d^2 Q_{\text{пад2}}}{dQ_2}$ — *эле-*

ментарные угловые коэффициенты, определяющие долю энергии излучения, которая попадает с элементарной площадки одного тела на элементарную площадку другого тела, по отношению к полному потоку собственного излучения, испускаемому первой элементарной площадкой.

Теплообмен излучением между большой и малой черными поверхностями. Найдем значения потоков излучения элементарных площадок dF_1 и dF_2 соответственно на поверхностях конечных размеров F_1 и F_2 с учетом того, что плотности потоков собственного излучения черных тел при $T = \text{const}$ являются постоянными вдоль поверхности каждого тела:

$$dQ_{\text{пад1}} = \int_{F_2} d^2 Q_{\text{пад1}} = E_{01} dF_1 \int_{F_2} d\varphi_{1,2} = E_{01} dF_1 \varphi_{1,2};$$

$$dQ_{\text{пад2}} = \int_{F_1} d^2 Q_{\text{пад2}} = E_{02} dF_2 \int_{F_1} d\varphi_{2,1} = E_{02} dF_2 \varphi_{2,1};$$

$$\varphi_{1,2} = \int_{F_2} d\varphi_{1,2} = \int_{F_2} \frac{\cos \psi_1 \cos \psi_2}{\pi r^2} dF_2;$$

$$\varphi_{2,1} = \int_{F_1} d\varphi_{2,1} = \int_{F_1} \frac{\cos \psi_1 \cos \psi_2}{\pi r^2} dF_1.$$

Здесь $\varphi_{1,2}$, $\varphi_{2,1}$ — *местные угловые коэффициенты излучения*, которые характеризуют долю энергии, испускаемую элементарной площадкой одного тела на конечную поверхность другого тела, по отношению к полному потоку собственного излучения, испускаемому элементарной площадкой первого тела:

$$\varphi_{1,2} = \frac{dQ_{\text{пад1}}}{dQ_1}; \quad \varphi_{2,1} = \frac{dQ_{\text{пад2}}}{dQ_2}.$$

Теплообмен излучением между двумя черными поверхностями конечных размеров. Найдем теперь результирующий поток излучения между поверхностями F_1 и F_2 :

$$Q_{\text{пад1}} = \int_{F_1} dQ_{\text{пад1}} = E_{01} \int_{F_1} \varphi_{1,2} dF_1 = E_{01} F_1 \left(\frac{\int_{F_1} \varphi_{1,2} dF_1}{F_1} \right) = E_{01} F_1 \varphi_{\text{ср1,2}} = Q_1 \varphi_{\text{ср1,2}};$$

$$Q_{\text{пад2}} = \int_{F_2} dQ_{\text{пад2}} = E_{02} \int_{F_2} \varphi_{2,1} dF_2 = E_{02} F_2 \left(\frac{\int_{F_2} \varphi_{2,1} dF_2}{F_2} \right) = E_{02} F_2 \varphi_{\text{ср2,1}} = Q_2 \varphi_{\text{ср2,1}}.$$

Здесь $\varphi_{\text{ср1,2}}$, $\varphi_{\text{ср2,1}}$ — *интегральные угловые коэффициенты излучения*, характеризующие долю энергии излучения, которая попадает с одного тела, имеющего конечную площадь поверхности, на другое тело с конечной площадью поверхности, по отношению к полному потоку собственного излучения первого тела:

$$\varphi_{\text{ср1,2}} = \frac{Q_{\text{пад1}}}{Q_1}; \quad \varphi_{\text{ср2,1}} = \frac{Q_{\text{пад2}}}{Q_2}.$$

Результирующий поток излучения

$$\begin{aligned} Q_{\text{рез}} &= Q_{\text{пад1}} - Q_{\text{пад2}} = \varphi_{\text{ср1,2}} E_{01} F_1 - \varphi_{\text{ср2,1}} E_{02} F_2 = \\ &= (E_{01} - E_{02}) F_1 \varphi_{\text{ср1,2}} = (E_{01} - E_{02}) F_2 \varphi_{\text{ср2,1}}. \end{aligned}$$

Следует отметить, что все угловые коэффициенты являются геометрическими характеристиками рассматриваемой излучающей системы тел, так как определяются только геометрической формой и взаимным расположением тел.

Теплообмен излучением между двумя серыми телами, произвольно расположенными в пространстве. Для системы, состоящей из двух серых тел, поток результирующего излучения определяется как разность соответствующих долей *не собственных, а эффективных излучений* каждого тела:

$$Q_{\text{рез}} = Q_{\text{пад1}} - Q_{\text{пад2}} = \varphi_{\text{ср1,2}} Q_{\text{эф1}} - \varphi_{\text{ср2,1}} Q_{\text{эф2}}.$$

При этом с учетом

$$Q_{\text{эф}} = E_{\text{эф}} F = EF/A + (1/A - 1) Q_{\text{рез}}$$

получаем

$$Q_{\text{рез}} = \varphi_{\text{ср1,2}} [E_1 F_1 / A_1 + (1/A_1 - 1)(-Q_{\text{рез}})] - \\ - \varphi_{\text{ср2,1}} [E_2 F_2 / A_2 + (1 - 1/A_2) Q_{\text{рез}}],$$

или

$$Q_{\text{рез}} = [1 + \varphi_{\text{ср1,2}}(1/A_1 - 1) + \varphi_{\text{ср2,1}}(1/A_2 - 1)] = \\ = \varphi_{\text{ср1,2}} F_1 (E_1 / A_1) - \varphi_{\text{ср2,1}} F_2 (E_2 / A_2) = \varphi_{\text{ср1,2}} F_1 E_{01} - \varphi_{\text{ср2,1}} F_2 E_{02},$$

откуда

$$Q_{\text{рез}} = \frac{\sigma_0 \varphi_{\text{ср1,2}} F_1 [(T_1 / 100)^4 - (T_2 / 100)^4]}{1 + \varphi_{\text{ср1,2}}(1/A_1 - 1) + \varphi_{\text{ср2,1}}(1/A_2 - 1)}.$$

Алгебраический метод определения угловых коэффициентов (метод поточной алгебры Поляка)

Метод поточной алгебры, разработанный Г.Л. Поляком, позволяет в ряде случаев решать задачи по определению угловых коэффициентов без применения интегрирования, ограничиваясь только простыми алгебраическими уравнениями.

Этот метод основан на ряде свойств угловых коэффициентов и *взаимных поверхностей* тел в системе, где происходит теплообмен излучением, причем плотность потока излучения принимается постоянной в пределах поверхностей этих тел.

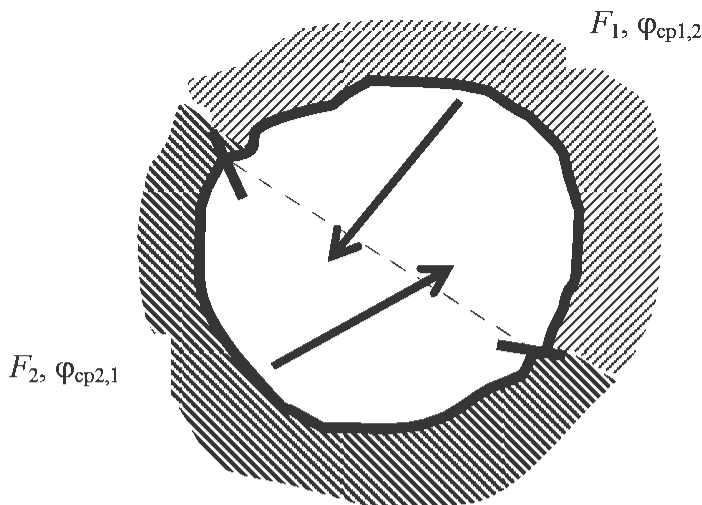


Рис. 2.7. Взаимные поверхности при теплообмене излучением

Взаимными поверхностями называются произведения значений площадей излучающих теплообменивающихся поверхностей и соответствующих значений угловых коэффициентов, например: $d\varphi_{1,2}dF_1$ и $d\varphi_{2,1}dF_2$ — соответствующие элементарные взаимные поверхности; $\varphi_{ср1,2}F_1$ и $\varphi_{ср2,1}F_2$ — соответствующие средние взаимные поверхности.

Свойство 1. Взаимные поверхности равны.

Если два тела находятся во взаимном теплообмене излучением (рис. 2.7), то их взаимные поверхности равны, а угловые коэффициенты обратно пропорциональны площадям поверхностей этих тел:

$$\varphi_{ср1,2}F_1 = \varphi_{ср2,1}F_2;$$

$$\varphi_{ср1,2}/\varphi_{ср2,1} = F_2/F_1.$$

Свойство 2. Вогнутая поверхность частично излучает «на себя».

Если теплообмен излучением происходит между телом с вогнутой поверхностью F и окружающей средой, то угловой коэффициент переноса энергии излучения этой поверхности в окружающую среду $\varphi_{F,окр}$ меньше единицы на значение углового коэффициента $\varphi_{F,F}$ переноса энергии излучения этой поверхности «на себя» (рис. 2.8):

$$\varphi_{F,окр} = 1 - \varphi_{F,F}.$$

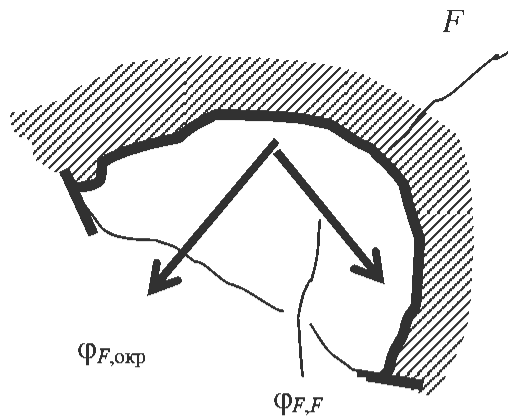


Рис. 2.8. Излучение вогнутой поверхности

Для взаимных поверхностей

$$F\Phi_{F,окр} = F - F\Phi_{F,F}.$$

Баланс потоков излучения, испускаемых поверхностью F :

$$Q_{F,окр} = EF\Phi_{F,окр} = EF - EF\Phi_{F,F}.$$

Для тел с плоской или выпуклой поверхностью

$$\Phi_{F,F} = 0, \Phi_{F,окр} = 1.$$

Свойство 3. Площадь поверхности тела равна сумме его взаимных поверхностей.

Если происходит взаимный теплообмен излучением между одним из тел замкнутой системы тел и всеми остальными телами (рис. 2.9), то сумма угловых коэффициентов переноса энергии излучения этого тела на все остальные тела, образующие замкнутую систему тел, равна единице, а сумма всех взаимных поверхностей этого тела равна площади его поверхности:

$$\varphi_{1,1} + \varphi_{1,2} + \varphi_{1,3} + \dots + \varphi_{1,n} = 1;$$

$$F_1\varphi_{1,1} + F_1\varphi_{1,2} + F_1\varphi_{1,3} + \dots + F_1\varphi_{1,n} = F_1;$$

$$Q_1 = E_1F_1 = E_1F_1\varphi_{1,1} + E_1F_1\varphi_{1,2} + E_1F_1\varphi_{1,3} + \dots + E_1F_1\varphi_{1,n}.$$

Свойство 4. Значение взаимной поверхности не зависит от формы и протяженности поверхности в замкнутой системе тел.

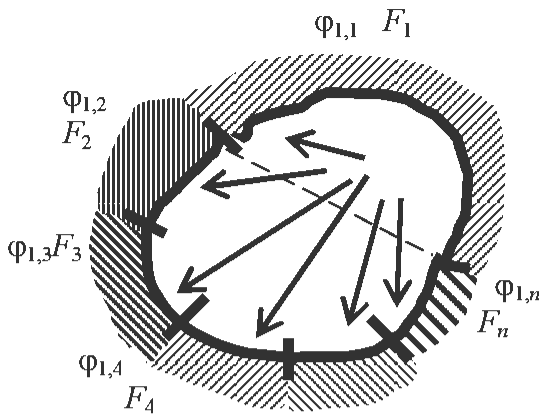


Рис. 2.9. Взаимный теплообмен излучением, при котором площадь поверхности тела равна сумме его взаимных поверхностей

Если между телом с поверхностью F_1 и телом с поверхностью F_2 происходит теплообмен излучением и эти два тела образуют замкнутую систему тел (рис. 2.10), то угловой коэффициент переноса энергии излучения с поверхности F_1 на поверхность F_2 , а также значение взаимной поверхности не зависят от формы и протяженности поверхности F_2 :

$$\varphi_{1,2} = \varphi_{1,2'} = \varphi_{1,2''} = \varphi_{1,2'''};$$

$$F_1 \varphi_{1,2} = F_1 \varphi_{1,2'} = F_1 \varphi_{1,2''} = F_1 \varphi_{1,2'''}.$$

Это свойство является следствием свойства 2.

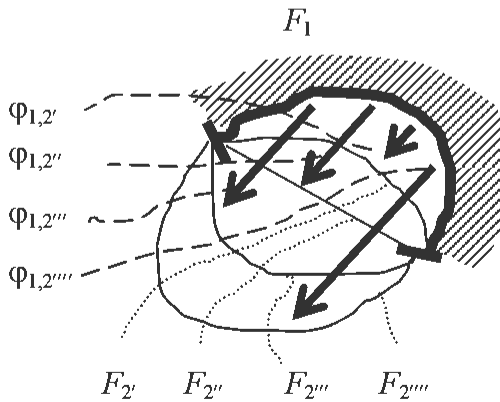


Рис. 2.10. Взаимодействие тел с поверхностями различных размеров

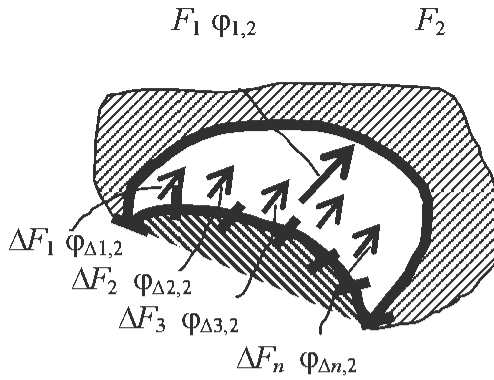


Рис. 2.11. Угловые коэффициенты переноса энергии

Свойство 5. Средние и местные (локальные) угловые коэффициенты равны.

Если значения локальных угловых коэффициентов переноса энергии излучения с любого элемента ΔF_i данной поверхности F_1 на другую поверхность F_2 равны между собой, то значение среднего углового коэффициента переноса энергии излучения с данной поверхности F_1 на другую поверхность F_2 равно значениям этих локальных угловых коэффициентов.

Доказательство этого свойства базируется на следующем рассуждении: если поверхность F_1 произвольным образом разделить на элементы ΔF_i (рис. 2.11) так, что

$$F_1 = \Delta F_1 + \Delta F_2 + \Delta F_3 + \dots + \Delta F_n,$$

и при этом

$$\varphi_{\Delta 1,2} = \varphi_{\Delta 2,2} = \varphi_{\Delta 3,2} = \dots = \varphi_{\Delta n,2},$$

то тепловой поток, излучаемый первым телом на второе

$$Q_{1 \rightarrow 2} = E_1 F_1 \varphi_{1,2} = E_1 \Delta F_1 \varphi_{\Delta 1,2} + E_1 \Delta F_2 \varphi_{\Delta 2,2} + \dots$$

$$\dots + E_1 \Delta F_n \varphi_{\Delta n,2} = E_1 [\Delta F_1 + \Delta F_2 + \Delta F_3 + \dots + \Delta F_n] \varphi_{\Delta n,2} = E_1 F_1 \varphi_{\Delta i,2}.$$

Отсюда для угловых коэффициентов можно записать

$$\varphi_{1,2} = \varphi_{\Delta i,2},$$

при этом для взаимных поверхностей справедливо равенство

$$F_1 \varphi_{1,2} = F_1 \varphi_{\Delta i,2}.$$

Свойство 6. Экранирование поверхностей.

Если тело с поверхностью F_i полностью закрывается другими телами от тела с поверхностью F_k (рис. 2.12), то

$$\varphi_{k,i} = \varphi_{i,k} = 0.$$

Используя приведенные выше уравнения, можно значительно упростить решение некоторых задач по определению угловых коэффициентов теплообмена излучением между телами и последующие расчеты результирующих потоков излучения в системе серых поверхностей. Рассмотрим несколько важных вариантов.

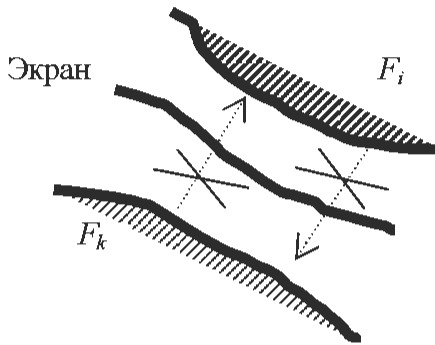


Рис. 2.12. Экранирование поверхностей

Вариант 1. Две поверхности образуют замкнутую систему, причем одна из поверхностей не имеет вогнутостей (рис. 2.13).

На основании свойства 1 угловых коэффициентов можем написать следующее соотношение:

$$\varphi_{1,2}/\varphi_{2,1} = F_2/F_1.$$

Поверхность F_1 не имеет вогнутости, и поэтому согласно определению и свойству 2

$$\varphi_{1,2} = 1,0.$$

Тогда угловой коэффициент $\varphi_{2,1}$ равен

$$\varphi_{2,1} = F_1/F_2.$$

Согласно свойству 3 угловой коэффициент $\varphi_{2,2}$ равен

$$\varphi_{2,2} = 1 - \varphi_{2,1} = 1 - F_1/F_2.$$

Вариант 2. Две вогнутые поверхности образуют замкнутую систему (рис. 2.14).

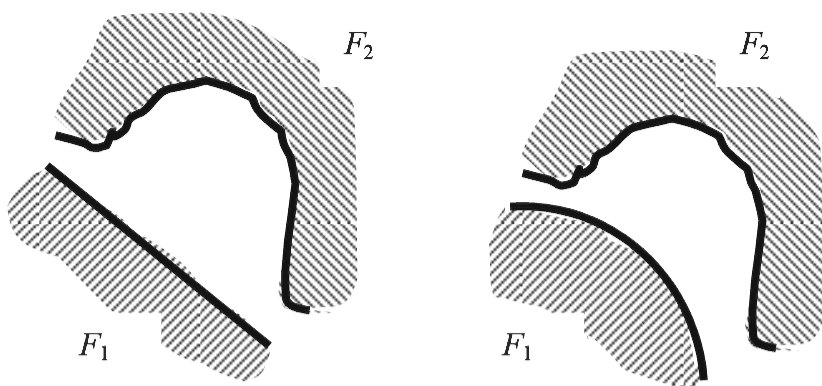


Рис. 2.13. Замкнутая система двух поверхностей, одна из которых не имеет вогнутостей

Между поверхностями проведем разделительную (вспомогательную) поверхность F_0 , не имеющую вогнутостей в сторону поверхности F_2 . Тогда согласно свойству 4

$$\varphi_{2,1} = \varphi_{2,0}.$$

Угловой коэффициент $\varphi_{2,0}$ по аналогии с $\varphi_{2,1}$ для варианта 1 равен

$$\varphi_{2,0} = \varphi_{2,1} = F_0/F_2.$$

Согласно свойству 1 с учетом формулы для $\varphi_{2,2}$ для варианта 1 получим

$$\varphi_{1,2} = \varphi_{2,1}F_2/F_1 = (F_0/F_2)(F_2/F_1) = F_0/F_1.$$

Угловые коэффициенты $\varphi_{2,2}$ и $\varphi_{1,1}$ могут быть определены согласно свойству 3 следующим образом:

$$\varphi_{2,2} = 1 - \varphi_{2,1} = 1 - F_0/F_2;$$

$$\varphi_{1,1} = 1 - \varphi_{1,2} = 1 - F_0/F_1.$$

Вариант 3. Одно тело с поверхностью F_1 , не имеющее вогнутостей, находится внутри другого тела (система «тело без вогнутостей — оболочка», рис. 2.15).

Согласно свойству 1 и с учетом $\varphi_{1,1} = 0$ и $\varphi_{1,2} = 1,0$ имеем

$$\varphi_{2,1} = F_1/F_2.$$

Согласно свойству 3

$$\varphi_{2,2} = 1 - \varphi_{2,1} = 1 - F_1/F_2.$$

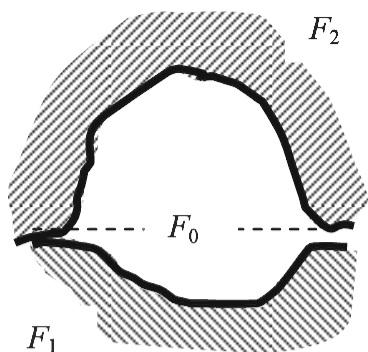


Рис. 2.14. Замкнутая система двух вогнутых поверхностей

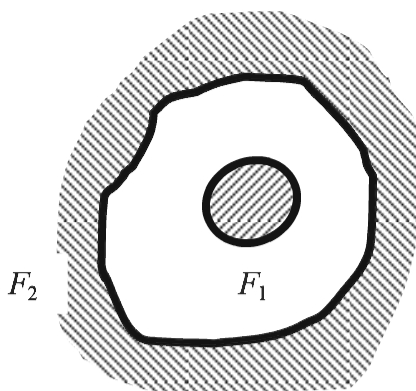


Рис. 2.15. Замкнутая система «тело без вогнутостей — оболочка»

Вариант 4. Одно тело с поверхностью F_1 , имеющее вогнутости, находится внутри другого тела (система «тело — оболочка», рис. 2.16).

На поверхность F_1 «натянем» вспомогательную поверхность F_0 (поверхность-делитель), которая не имеет вогнутостей в сторону поверхности F_2 . Тогда по аналогии с рассмотренным вариантом 3 и учитывая (согласно свойству 4), что $\varphi_{2,1} = \varphi_{2,0}$, можем записать

$$\varphi_{2,1} = \varphi_{2,0} = F_0/F_2.$$

В соответствии со свойством 1

$$\varphi_{1,2} = \varphi_{2,1}F_2/F_1 = F_0/F_1.$$

Согласно свойству 3,

$$\varphi_{2,2} = 1 - \varphi_{2,1} = 1 - F_0/F_2;$$

$$\varphi_{1,1} = 1 - \varphi_{1,2} = 1 - F_0/F_1.$$

Вариант 5. Две бесконечные плоскопараллельные поверхности (рис. 2.17).

Согласно свойству 2 для плоских поверхностей

$$\varphi_{2,2} = 0; \quad \varphi_{1,1} = 0.$$

Следовательно,

$$\varphi_{1,2} = 1 - \varphi_{1,1} = 1,0;$$

$$\varphi_{2,1} = 1 - \varphi_{2,2} = 1,0.$$

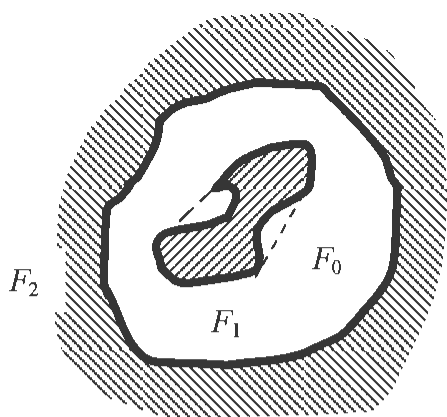


Рис. 2.16. Замкнутая система «тело — оболочка»

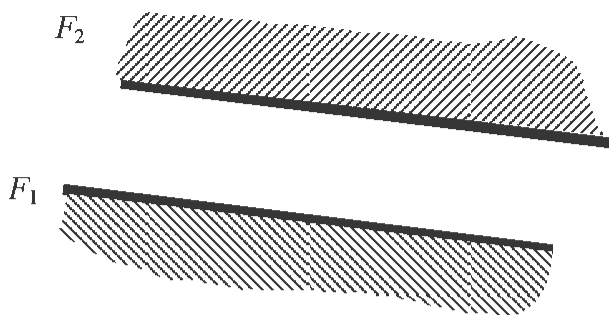


Рис. 2.17. Две плоскопараллельные поверхности

2.3. Теория теплопередачи и ее использование для анализа физических процессов различной природы

Изложенные в этой главе физические основы и методы расчета, а также приведенные расчетные зависимости позволяют выполнять приближенные расчеты при анализе процессов теплопередачи. При этом следует иметь в виду, что данная теоретическая база является лишь вводным материалом в содержании курса теплопередачи.

При необходимости проведения точных и детальных расчетов по определению температурного поля или расчетов распределения тепловых потоков в рассматриваемой системе тел следует решать поставленную задачу с требуемой степенью точности, при этом сложность решения находится в обратной зависимости от степени упрощения. Для многих задач, особенно задач теплопроводности, имеются аналитические решения высокой степени точности, однако чаще всего существующая теория дает лишь средства для решения, а конкретные вычисления приходится делать самому. Более того, сложность реальных процессов теплопередачи такова, что в большинстве случаев расчеты проводят с использованием численных методов (например, метода конечных разностей или метода конечных элементов), полуэкспериментальных и экспериментальных зависимостей.

Очень мощный специфический аппарат решения задач теплопроводности разработан в рамках теории подобия, где на основании анализа исходной системы дифференциальных уравнений, довольно точно описывающих условия конкретного рассматриваемого процесса, получают систему критериальных уравнений, для которой в процессе экспериментальных исследований находят так называемые полуэмпирические расчетные зависимости, используемые для расчета требуемых величин.

В экспериментальных исследованиях широко используется метод аналогий, базирующийся на понятии аналогичных процессов как процессов, имеющих разную физическую природу, но описываемых одинаковыми по форме записи системами дифференциальных уравнений. Например, многие физические процессы описываются в общем виде уравнениями типа

$$\mathbf{q} = -k \frac{dW}{dL} \mathbf{l}_0,$$

где \mathbf{q} — векторная величина, поток; k — коэффициент пропорциональности; dW/dL — градиент определяющей переменной W по направлению L , определяемому единичным вектором \mathbf{l}_0 .

В качестве примеров можно привести следующие зависимости:
– уравнение теплопроводности Фурье

$$\mathbf{q}_m = -\lambda \frac{dt}{dn} \mathbf{n}_0,$$

где \mathbf{q}_m — поток теплоты; λ — коэффициент теплопроводности; t — температура;

– уравнение диффузии Фика

$$\mathbf{q}_M = -D \frac{dc}{dn} \mathbf{n}_0,$$

где \mathbf{q}_M — поток массы; D — коэффициент диффузии; c — концентрация;

– закон Ома

$$\mathbf{j} = -\chi \frac{du}{dn} \mathbf{n}_0,$$

где \mathbf{j} — плотность тока; χ — удельная проводимость; u — потенциал.

Математически эти уравнения являются абсолютно одинаковыми, отличаясь только буквами, обозначающими переменные, и поэтому любой из этих процессов может быть моделью другого. Другими словами, результаты, полученные для одного процесса, могут быть использованы для оценки результатов другого процесса. Или, иначе, зависимости $q(W)$ имеют одинаковый характер. Это позволяет использовать широко применяемую на практике *электротепловую аналогию*, когда моделирование как стационарных, так и нестационарных тепловых потоков проводят с помощью электрических токов в электрической цепи, аналогичной рассматриваемой тепловой схеме.

Таким образом, материал, изложенный в этой главе, может рассматриваться не только как введение в курс теплопередачи, но и служить демонстрацией типичных подходов, используемых для анализа физических процессов различной природы.

Контрольные вопросы и задания

1. Какие виды теплопередачи вы знаете?
2. Что такое температурный градиент?
3. Что такое температурное поле?
4. Каково содержание закона Фурье?
5. Выведите дифференциальное уравнение теплопроводности.
6. Что такое условия однозначности? Какие условия однозначности вы знаете?
7. Каковы интегральные и спектральные характеристики излучения?
8. Что такое угловая плотность и интенсивность излучения?
9. Каковы составляющие падающего потока излучения при взаимодействии с поверхностью тела?
10. Что такое поток эффективного излучения?
11. Что такое поток результирующего излучения?
12. Какова зависимость между потоками эффективного и результирующего излучения?
13. Сформулируйте закон Планка, его частные случаи и закон смещения Вина.
14. Сформулируйте закон Стефана — Больцмана.
15. Сформулируйте закон косинусов Ламберта.
16. Сформулируйте закон Кирхгофа.

17. Как рассчитать поток результирующего излучения при теплообмене излучением между двумя элементарными черными поверхностями? Что такое элементарный угловой коэффициент?

18. Как рассчитать поток результирующего излучения при теплообмене излучением между двумя черными поверхностями конечных размеров? Что такое интегральный угловой коэффициент?

19. Как рассчитать поток результирующего излучения при теплообмене излучением между двумя серыми поверхностями конечных размеров?

20. Какие свойства угловых коэффициентов используются в методе поточной алгебры?

21. Определите угловые коэффициенты для замкнутой системы двух поверхностей.

3. ФОРМИРОВАНИЕ ПОТОКОВ ЧАСТИЦ

Изучив материал этой главы, студент должен иметь представление о методах и средствах формирования электронных, ионных, атомарных и молекулярных потоков для использования их в микро- и нанотехнологиях.

Студент должен знать:

- зависимость плотности тока термоэлектронной эмиссии от температуры и работы выхода электронов;
- способы управления скоростью испарения материалов в вакууме;
- особенности испарения сплавов;
- физическую сущность распыления материалов ионной бомбардировкой.

Студент должен уметь:

- рассчитывать скорость ионного распыления материалов в вакууме;
- различать и анализировать методы формирования ионных пучков с горячим и холодным катодом;
- оценивать применимость процесса распыления материалов ионной бомбардировкой в микро- и нанотехнологиях.

Студент должен иметь навыки работы с источниками технической информации, используемые для поиска значений парамет-

ров, необходимых для расчета электронных, ионных, атомарных и молекулярных потоков.

В электронных технологиях в качестве инструмента используются остросфокусированные электронные, ионные, атомарные, оптические, рентгеновские пучки, газоразрядная плазма, а также мощные электрические и магнитные поля, с помощью которых обрабатывают изделия, измеряют их размеры, контролируют свойства, диагностируют значения параметров, управляют технологическими процессом и оборудованием.

В табл. 3.1 приведены параметры пучков, характерные для электронных, ионных и плазменных технологий: для фотонов с энергией 1,6...3,5 эВ, соответствующей излучению в видимом диапазоне, минимальный размер обработки составляет приблизительно 1000 нм, а в области ультрафиолетового и длинноволнового рентгеновского излучения (диапазон значений энергии фотонов составляет 5...1000 эВ) размер обработки можно уменьшить до 100 нм.

Таблица 3.1

Характеристики пучков атомных частиц

Тип пучка	Длина волны, нм	Энергия, эВ	Минимальный размер, нм	Удельная мощность, Вт/см ²
Оптический	200 – 400	1,6 – 3,5	1000	$10^{-2} - 10^{10}$
Рентгеновский	0,2 – 2,0	5 – 1000	100	$10^{-3} - 1$
Электронный	0,01	$10^2 - 10^5$	10	$10^{-2} - 10^{10}$
Ионный	0,001	$10^2 - 10^7$	1	$10^{-4} - 10^5$
Атомарный	0,001	$0,1 - 10^{4*}$	1*	$10^{-6} - 10^{2*}$

* Максимальные значения энергии и удельной мощности атомарного пучка, а также минимальный размер обработки относятся к ионному пучку с нейтральным зарядом.

В обычно используемом диапазоне значений энергии электронов 0,1...100,0 кэВ можно получить разрешение, сравнимое с размерами атомов (порядка 0,1 нм). Ограничения на минимальный размер пучка электронов (порядка 10 нм) связаны с их рассеянием. Ионные и атомарные пучки характеризуются отсутствием ограничений на дебройлевскую длину волны даже при малых энергиях ионов. Доля рассеянных ионов обычно очень мала, поскольку их

размеры соизмеримы с периодом кристаллических решеток материалов, с которыми они взаимодействуют. Минимальный размер пучка ионов или атомов может достигать примерно 1 нм.

Для формирования потоков частиц и управления этими потоками (например, при электронно-лучевой обработке) формирующие, ускоряющие, фокусирующие и отклоняющие электроды выполняют как единый блок — электронно-оптическую систему (ЭОС).

3.1. Формирование электронных потоков

Оптические пучки, используемые в элионных технологиях для засветки фоторезиста, испарения материалов при нанесении тонких пленок, подгонки тонкопленочных резисторов и выполнения других операций, формируются с помощью твердотельных и газовых лазеров. Рентгеновское излучение используется в рентгенолитографическом и аналитическом оборудовании и формируется с помощью рентгеновских трубок или синхротрона.

Электронные пушки

Формирование потоков электронов осуществляется с помощью электронных пушек (рис. 3.1). Пушка состоит из источника электронов (эмиттера) и двух или более электродов с различными электрическими потенциалами. Обычно заземленный анод разгоняет электроны до конечной скорости при движении от катода, имеющего отрицательный потенциал, соответствующий требуемой энергии носителей заряда. Однако возможен случай, когда катод заземлен, а на анод подается положительный потенциал 10...100 кВ. Назначением электрода Венельта, потенциал которого близок к потенциалу эмиттера, является придание силовым линиям (особенно у заостренного конца эмиттера) такой формы, чтобы обеспечить требуемую интенсивность поля и направление движения электронного пучка.

Важнейшей характеристикой источника электронов является его яркость β , выражение для которой имеет вид

$$\beta = \frac{I}{\Omega S},$$

где I — ток эмиссии; Ω — телесный угол, в который происходит эмиссия; S — площадь эмитирующей поверхности.

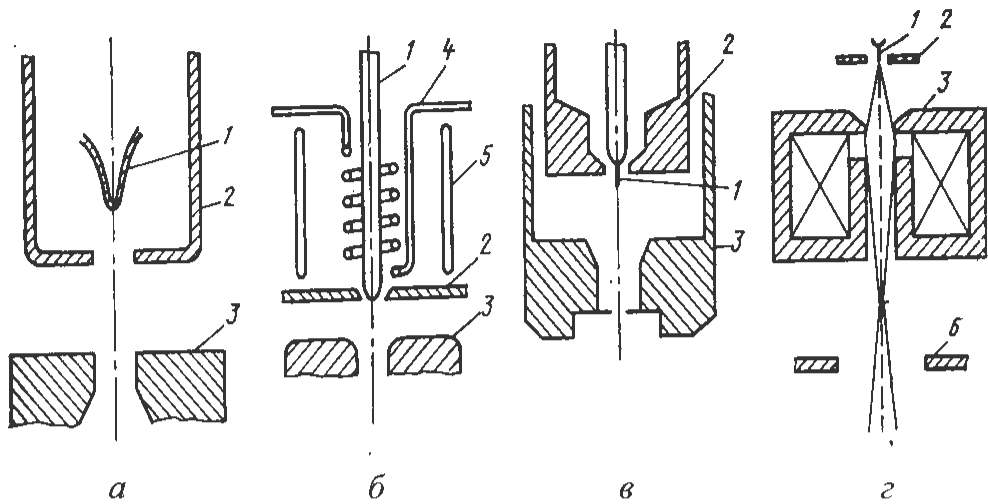


Рис. 3.1. Схемы электронных пушек:

а — с термоэмиссионным катодом из вольфрамовой проволоки; *б* — типа Броерса с катодом из гексаборида лантана с косвенным подогревом; *в* — автоэмиссионный источник с электростатической фокусирующей системой; *г* — автоэмиссионный источник с магнитной линзой и двумя анодами; 1 — катод; 2 — управляющий электрод; 3 — анод; 4 — нить подогрева; 5 — экран; б — второй анод

Плотность тока термоэмиссии

Сила тока зависит от материала катода и механизма эмиссии. Плотность тока термоэмиссии j_0 можно рассчитать по формуле Ричардсона — Дэшмана:

$$j_0 = AT^2(1-r) \exp \left[-\frac{e}{kT} (\varphi - \sqrt{eE}) \right],$$

где $A = 120 \text{ A}/(\text{см}\cdot\text{К})^2$ — константа Ричардсона; T — температура эмиттера, К; r — коэффициент отражения от потенциального барьера; e — заряд электрона; $e\varphi$ — работа выхода электрона, эВ; k — постоянная Больцмана, Дж/К; E — напряженность электрического поля, В/м.

Для вольфрамового катода $e\varphi = 4,5 \text{ эВ}$ при $T = 2873 \text{ К}$, $E = 10^5 \text{ В/м}$ (ускоряющее напряжение равно $U = 25 \text{ кВ}$), $A(1-r) = 60 \text{ A}/(\text{см}\cdot\text{К})^2$, плотность тока термоэмиссии составляет $j_0 =$

$= 6,3 \text{ A/cm}^2$. Для гекaborид-лантанового катода расчетное значение плотности тока при $A(1 - r) = 40 \text{ A/(cm}\cdot\text{K)}^2$, $e\phi = 2,6 \text{ эВ}$, $T = 1873 \text{ К}$, $U = 25 \text{ кВ}$, $E = 10^6 \text{ В/м}$ равно $j_0 = 49 \text{ A/cm}^2$.

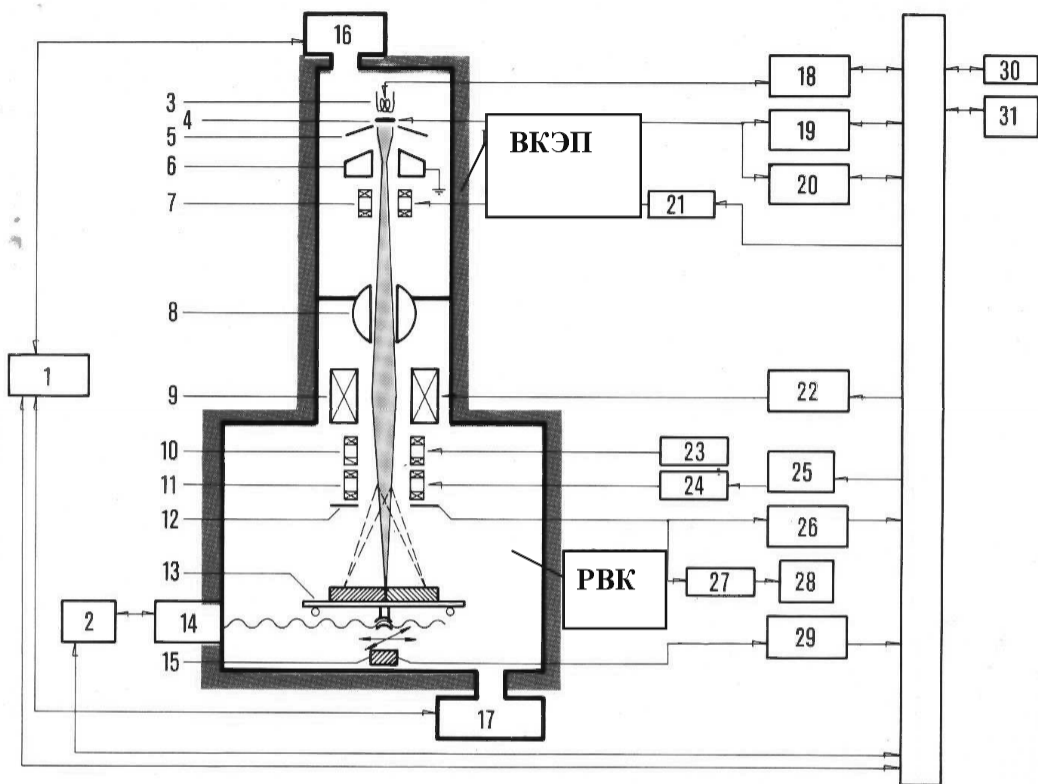


Рис. 3.2. Схема установки для электронно-лучевой сварки:

ВКЭП — вакуумная камера электронной пушки; РВК — рабочая вакуумная камера; 1 — контроллер управления системами откачки 16 и 17; 2 — контроллер управления приводом 14 координатного стола 13; 3 — нагреватель катода 4; 5–7 — элементы электронной пушки; 8–12 — элементы управления электронным пучком; 15 — датчики перемещения координатного стола; 18–21 — контроллеры управления электронной пушкой; 22–28 — контроллеры управления электронным пучком; 29 — контроллер управления перемещением координатного стола; 30, 31 — персональные компьютеры системы автоматического управления двух уровней

Для придания электронному пучку необходимой энергии (скорости) и формы служат формирующие и апертурные диафрагмы, фокусирующие и отклоняющие линзы, стигматоры, сканирующие системы и другие элементы ЭОС (рис. 3.2).

3.2. Формирование ионных потоков

Формирование пучков ионов осуществляется с помощью источников ионов, которые состоят из разрядной (ионизационной) камеры, источника электронов (нагреваемого термокатода, холодного или полого катода), анода, магнитной системы (для повышения эффективности ионизации), экстрагирующего электрода и электрода первичной фокусировки, системы подачи рабочего газа или пара. Ионные источники должны генерировать однородные высокоинтенсивные пучки со стабильными во времени параметрами и многозарядные ионные пучки, обеспечивать предельную плотность ионного тока при максимально низких экстрагирующих напряжениях, ионизировать газообразные (Ar , N_2 , O_2 , BF_3 , PCl_3 , AsF_3 , AlCl_3 и т. п.) и твердые (B , As , Sb , Al , Ti и т. п.) вещества, формировать ионный пучок заданного сечения, быть простыми в управлении и обладать большим временем наработки на отказ катода и всего источника.

Поскольку обеспечить в одном источнике все перечисленные требования невозможно, применяют различные типы источников, основными из которых являются автономные источники ионов с горячим, холодным или полым катодом, дуоплазмотроны, источники с высокочастотным и сверхвысокочастотным возбуждением, источники с поверхностной ионизацией и др.

Ионные источники с горячим катодом

В источниках ионов с горячим катодом (рис. 3.3, а) в качестве источника электронов используются катоды прямого накала или с косвенным подогревом. Экстракция ионов в этих источниках осуществляется вдоль оси разрядной камеры коническим пучком (аксиальная экстракция) или перпендикулярно оси разрядной камеры через отверстие прямоугольной формы пучком ионов в виде узкой ленты (радиальная экстракция). Основным достоинством таких источников является возможность получения высокоинтенсивных пучков ионов, а главным недостатком — быстрое разрушение термокатода при использовании химически активных рабочих газов.

Ионные источники с холодным катодом и разрядом Пеннинга

Автономные источники ионов с холодным катодом и разрядом Пеннинга (рис. 3.3, б) содержат цилиндрический кольцевой анод, размещенный между параллельно расположенными катодом и антикатодом. Ионы вытягиваются из плазмы самостоятельного газового разряда через отверстие в антикатоде и ускоряются системой экстракции. Благодаря наличию соленоида источник обладает высокой эффективностью ионизации (до 10^{14} ион/см³), срок службы катода превышает 1000 ч. Недостатком такой конструкции является низкая сила тока пучка ионов (до 100 мкА), большая пульсация тока и возможность использования только газообразных рабочих веществ.

Типовой жидкометаллический источник ионов состоит из двух электродов: острия-анода и заземленного экстрактора, между которыми приложено напряжение около 4...10 кВ. Острие выполнено в виде тонкой вольфрамовой трубки с капиллярным отверстием диаметром 50 мкм, через которое поступает жидкий металл. При давлении внутри трубки порядка 1 МПа и наличии электрического поля высокой напряженности ($\sim 10^{10}$ В/м) на конце капилляра образуется выступ жидкого металла в виде конуса, с вершины которого эмитируется расходящийся пучок положительных ионов.

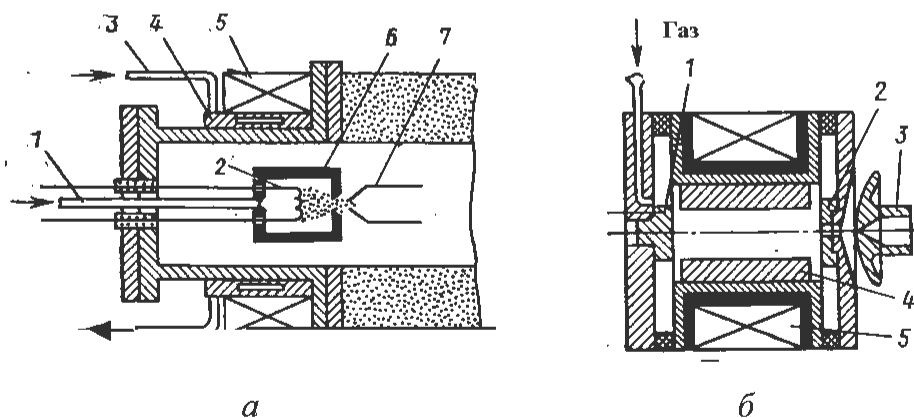


Рис. 3.3. Схемы автономных источников ионов:

а — с горячим катодом: 1 — ввод газа; 2 — катод; 3 — ввод охлаждающей жидкости; 4 — держатель; 5 — электромагнит; 6 — разрядная камера; 7 — экстрактор; б — с холодным катодом: 1 — катод; 2 — антикатод; 3 — экстрактор; 4 — цилиндрический кольцевой анод; 5 — соленоид

3.3. Формирование атомарных и молекулярных потоков

Основными способами формирования атомарных и молекулярных пучков являются: термическое и дуговое испарение, испарение взрывом, ионное распыление, а также получение ионных пучков из газовой фазы с последующей нейтрализацией.

С помощью испарения можно формировать потоки атомов и молекул металлов, сплавов, полупроводников и диэлектриков.

Уравнение Клаузиуса — Клапейрона и закон Герца — Кнудсена

Термическое испарение (рис. 3.4) с точки зрения термодинамики описывается уравнением Клаузиуса — Клапейрона:

$$\frac{dp_{\text{нас}}}{dT} = \frac{H_{\text{г}} - H_{\text{ж}}}{T(V_{\text{г}} - V_{\text{ж}})} = \frac{\Delta H}{TV_{\text{г}}},$$

где $p_{\text{нас}}$ — давление насыщенного пара испаряемого материала, Па; T — температура материала, К; $H_{\text{г}}$, $H_{\text{ж}}$ — энтальпия газа и жидкости соответственно, ккал/кмоль; $V_{\text{г}}$, $V_{\text{ж}}$ — объем газа и жидкости соответственно, м³ ($V_{\text{г}} \gg V_{\text{ж}}$); ΔH — теплота испарения, ккал/кмоль.

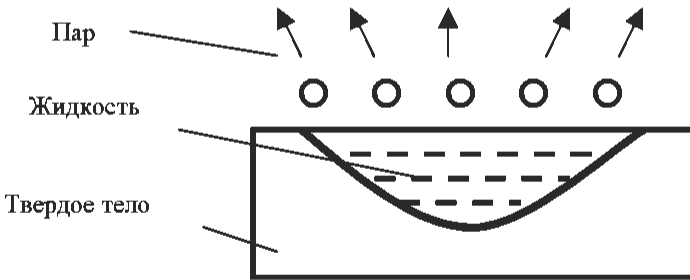


Рис. 3.4. Схема процесса термического испарения

Так как

$$V_{\text{г}} = \frac{R_0 T}{p},$$

где R_0 — универсальная газовая постоянная, Дж/(кмоль К); p — давление пара, Па, то

$$\frac{dp_{\text{нас}}}{p} = \frac{\Delta H}{T^2 R_0} dT; \quad \lg p_{\text{нас}} = -\frac{\Delta H}{R_0 T} + C; \quad \lg p_{\text{нас}} = A - \frac{B}{T},$$

где C — постоянная интегрирования; A и B — константы (табличные данные).

Согласно молекулярно-кинетической теории газов, термическое испарение подчиняется закону Герца — Кнудсена:

$$\frac{dN_{\text{и}}}{dt \cdot A} = \frac{p_{\text{нас}} \alpha_{\text{и}}}{\sqrt{2\pi m k T_{\text{и}}}},$$

где $N_{\text{и}}$ — количество испаренных атомов или молекул; t — время, с; A — площадь испарения, м^2 ; $\alpha_{\text{и}}$ — коэффициент испарения (для чистых материалов $\alpha_{\text{и}} = 1$); m — масса испаренного атома или молекулы, кг; $T_{\text{и}}$ — температура испарения, К.

Скорость испарения $v_{\text{и}}$, $\text{кг}/(\text{м}^2 \cdot \text{с})$, рассчитывают по формуле

$$v_{\text{и}} = m \frac{dN_{\text{и}}}{dt \cdot A} = m \frac{p_{\text{нас}} \alpha_{\text{и}}}{\sqrt{2\pi m k T_{\text{и}}}} = p_{\text{нас}} \alpha_{\text{и}} \sqrt{\frac{m}{2\pi k T_{\text{и}}}}$$

или

$$v_{\text{и}} = 5,83 \cdot 10^{-3} p_{\text{нас}} \sqrt{\frac{M}{T_{\text{и}}}},$$

где M — молекулярная масса испаряемого вещества, $\text{кг}/\text{кмоль}$.

Испарение сплавов описывается законом Рауля:

$$\frac{p_{\text{нас.}A_{\text{раств}}}}{p_{\text{нас.}A}} = \frac{x_A}{x_A + (100 - x_A)M_A / M_B},$$

где $p_{\text{нас.}A_{\text{раств}}}$ — давление насыщенного пара материала A в растворе; $p_{\text{нас.}A}$ — давление насыщенного пара однокомпонентного материала A ; x_A — содержание материала A в растворе, мас. долей, %; M_A , M_B — молекулярные массы материалов сплава A и B , $\text{кг}/\text{кмоль}$.

Испарение диэлектриков и полупроводников может происходить следующим образом: без диссоциации (SiO , MgF_2); с диссоциацией (при $T > 1800$ К практически все вещества диссоциируют,

а при $T > T_{\text{н}} + 200 \dots 400$ К без диссоциации испаряются MgO , Al_2O_3 , BeO , SiO_2 , ThO_2); с разложением, когда химический состав пара не соответствует испаряемому веществу (Ag_2S , CuI , WC , CrN , Cr_2O_3 , Fe_2O_3 , $\text{A}_{\text{III}}\text{B}_{\text{V}}$).

Ионное распыление материалов

Ионное распыление материала имеет место при взаимодействии («бомбардировке») ускоренных до $0,5 \dots 5,0$ кэВ ионов с веществом, находящимся в твердом или жидком состоянии. Сущность метода заключается в механическом выбивании атомов или молекул материала мишени путем передачи им кинетической энергии ускоренных ионов инертного газа (рис. 3.5). Основным показателем эффективности данного процесса является коэффициент распыления, который можно рассчитать по следующей формуле:

$$S = \frac{N_A z q_e \Delta m}{M J_i t},$$

где N_A — число Авогадро, атом/кмоль; $z q_e$ — заряд иона (z — кратность ионизации), Кл; Δm и M — масса, кг, и молекулярная масса, кг/кмоль, распыляемого вещества; J_i — ионный ток, А; t — время, с.

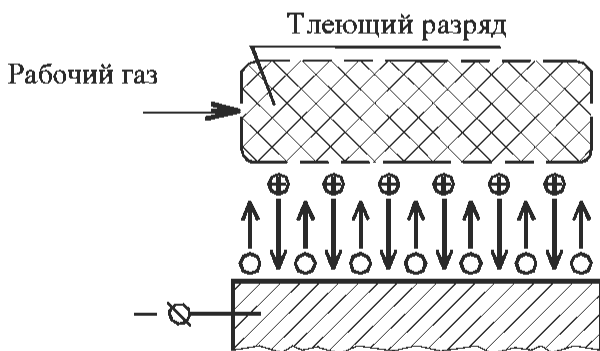


Рис. 3.5. Схема процесса ионного распыления

Скорость ионного распыления рассчитывают по следующим формулам ($z = 1$):

$$v_p = S \frac{j_i M}{q_e N_A} \quad \text{или} \quad v_p = S \frac{j_i M \cdot 10^9}{q_e N_A \rho},$$

где j_i — плотность ионного тока, А/м^2 ; ρ — плотность распыляемого материала, кг/м^3 (в первом из двух выражений v_p измеряется в $\text{кг}/(\text{м}^2 \cdot \text{с})$, во втором — в нм/с).

Поток атомов или молекул, сформированный из газовой фазы

Поток атомов или молекул, сформированный из газовой фазы, характеризуется небольшой энергией ($E_T = kT_T = 0,1 \dots 0,2$ эВ, где T_T — температура газа, К) и гибким регулированием интенсивности N_T или скорости v_T с помощью изменения давления газа p_T в широком диапазоне $10^{-5} \dots 10^5$ Па:

$$N_T = \frac{p_T}{\sqrt{2\pi m_T k T_T}}; \quad v_T = 5,83 \cdot 10^{-3} p_T \sqrt{M/T_T},$$

где N_T измеряется в $\text{атом}/(\text{м}^2 \cdot \text{с})$, а v_T — $\text{кг}/(\text{м}^2 \cdot \text{с})$

При ионно-лучевой обработке диэлектрических материалов возникает проблема нейтрализации образующегося поверхностного заряда. Одним из решений этой проблемы является использование атомарных пучков с энергией $1 \dots 10$ кэВ, получаемых с помощью ионно-оптических систем, снабженных устройством компенсации заряда пучка ионов.

Нейтрализацией ионного пучка можно получать высокоэнергетические потоки атомов или молекул: $E_a \approx E_i = q_e z U$, где U — ускоряющее напряжение, В.

Поток атомов или молекул N_a , $\text{атом}/(\text{м}^2 \cdot \text{с})$,

$$N_a = \frac{j_i k_n}{z q_e},$$

где k_n — коэффициент нейтрализации, $\text{атом}/\text{ион}$.

Скорость (интенсивность) потока v_a , $\text{кг}/(\text{м}^2 \cdot \text{с})$, можно представить как

$$v_a = \frac{M j_i k_n}{N_A z q_e}.$$

Характеристики электронных, ионных, атомарных и молекулярных потоков формируются в зависимости от требуемых условий взаимодействия этих потоков с обрабатываемым или измеряемым материалом.

Контрольные вопросы и задания

1. Как зависит плотность тока j_e термоэлектронной эмиссии от температуры T и работы выхода $e\phi$ (закон Ричардсона — Дэшмана)?
2. В чем различие методов формирования ионных пучков с горячим и холодным катодом?
3. Как управлять скоростью испарения материалов в вакууме?
4. Объясните особенности испарения сплавов в вакууме (закон Рауля).
5. В чем заключается физическая сущность ионного распыления материалов в вакууме?
6. Как можно управлять скоростью ионного распыления материалов в вакууме?
7. В чем заключается физический смысл коэффициента распыления материалов ионной бомбардировкой?

4. ВЗАИМОДЕЙСТВИЕ ПОТОКОВ ЧАСТИЦ С МАТЕРИАЛАМИ

Изучив материал этой главы, студент должен иметь представление о свойствах и возможностях электронных, ионных, атомарных и молекулярных потоков как инструментов микро- и нанотехнологий.

Студент должен знать:

- физическую сущность модели Арчарда, описывающей пробег электрона в твердом теле;
- основы представления пробега иона в твердом теле, описываемого эффектом парных соударений;
- сущность эффекта каналирования при проникании ускоренного иона в кристаллический материал.

Студент должен уметь:

- рассчитывать глубину проникания ускоренных электронов в твердое тело;
- моделировать процесс изменения температуры поверхности твердого тела вследствие его нагрева при взаимодействии с электронным пучком;

- рассчитывать концентрацию имплантированных ионов в твердом теле.

Студент должен иметь навыки работы с источниками технической информации, используемыми для выявления доминирующих факторов взаимодействия электронных, ионных, атомарных и молекулярных потоков с обрабатываемыми материалами.

4.1. Взаимодействие электронных и ионных потоков с материалами

Взаимодействие электронных и ионных потоков с газообразными и твердыми материалами характеризуется большим разнообразием эффектов, а следовательно, и многообразием областей их применения в электронных технологиях.

Поверхность твердого тела отличается от объемного материала геометрическим расположением атомов, структурой электронных связей, химическим соединением и т. п. Все поверхности, соприкасающиеся с атмосферой, покрыты слоями адсорбированных атомов и молекул. Так, если при давлении 10^{-4} Па получить ювенильную поверхность, то примерно через 1 с она оказывается покрытой монослоем ($\approx 10^{15}$ атом/см²) адсорбированных газов. Чтобы удержать поверхность на атомарно-чистом уровне для проведения, например, операции молекулярно-лучевой эпитаксии, необходимо обеспечить сверхвысокий вакуум с давлением не выше 10^{-8} Па.

Эффекты взаимодействия

Эффекты, возникающие при взаимодействии электронного пучка с веществом (рис. 4.1), реализуются как дискретные события, сопровождающиеся ионизацией и возбуждением остаточных газов и паров (рис. 4.1, *а*), появлением вторичных электронов (рис. 4.1, *б*), возбуждением колебаний плотности плазмы, ионизацией на внутренних электронных оболочках, вызывающих рентгеновское излучение и эмиссию Оже-электронов (рис. 4.1, *б*, *в*) и атомных частиц (рис. 4.1, *г*), рождением электронно-дырочных пар (рис. 4.1, *д*, *к*) с последующим световым излучением, переходным излучением и возбуждением упругих колебаний кристаллической решетки — воз-

буждением фононов (рис. 4.1, *е*), нагревом (рис. 4.1, *жс*), образованием дислокаций и радиационных дефектов (рис. 4.1, *з*) и химическими реакциями (рис. 4.1, *и*) и другими явлениями.

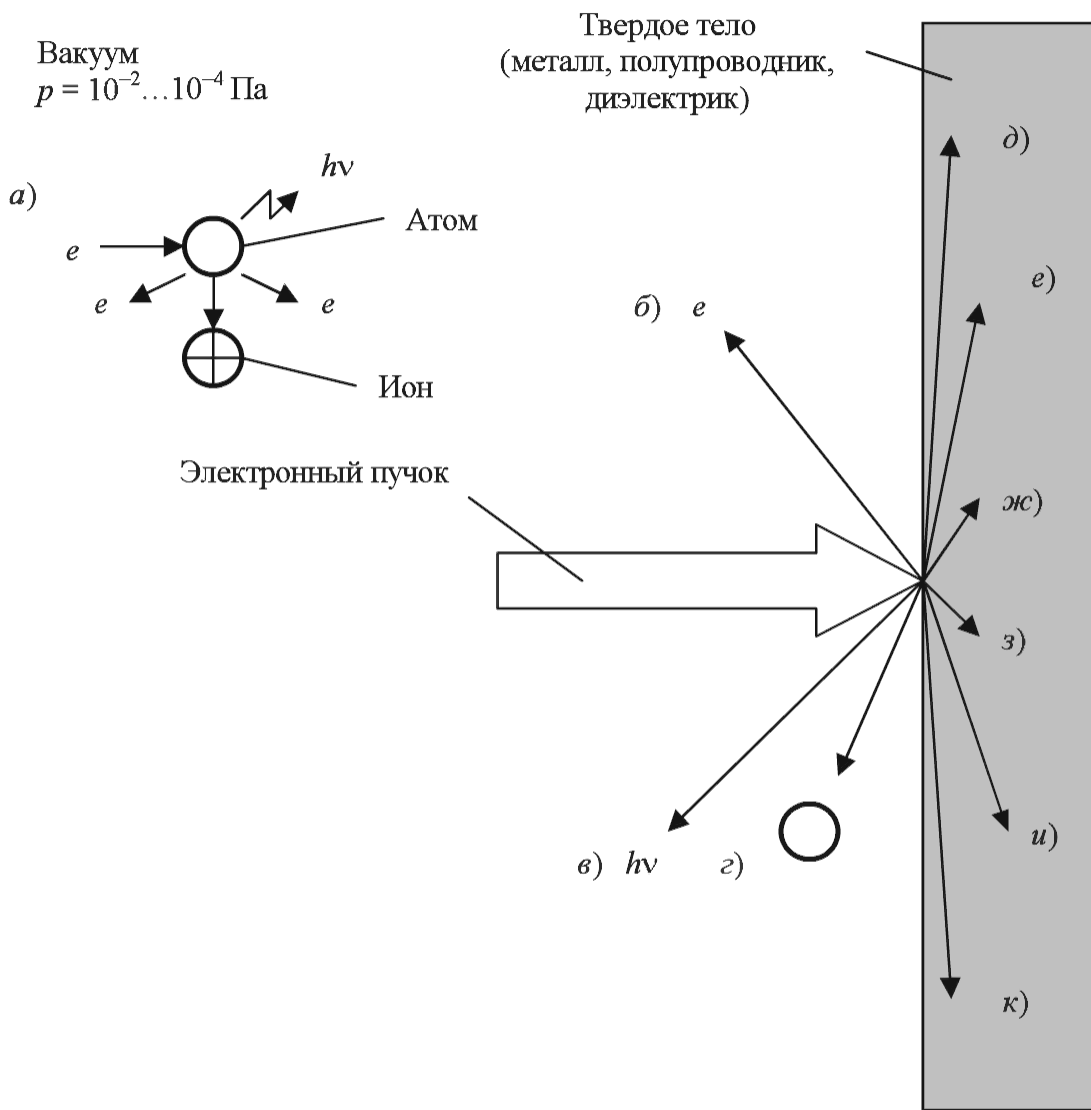


Рис. 4.1. Эффекты взаимодействия электронного пучка с остаточным газом и с твердым телом:

а — ионизация и возбуждение атомов и молекул остаточных газов и паров электронным ударом; *б* — эмиссия электронов e ; *в* — эмиссия фотонов $h\nu$; *г* — эмиссия атомных частиц; *д* — ионизация атомов; *е* — возбуждение фононных колебаний; *жс* — нагрев; *з* — образование дислокаций и радиационных дефектов; *и* — стимулирование химических реакций; *к* — увеличение проводимости полупроводников и диэлектриков

При столкновении ускоренных электронов с атомами или молекулами остаточных или рабочих газов и паров происходит их ионизация (образование и поддержание газоразрядной плазмы) и испускание фотонов (свечение плазмы). Метод получения плазмы электронным ударом используется во многих источниках ионов и в плазменных установках.

В зависимости от параметров электронного пучка и свойств материала мишени (обрабатываемой детали) может иметь место упругое и неупругое отражение электронов от поверхности твердого тела, что дает возможность наблюдать за поверхностью с помощью электронного микроскопа, получать новые электронные пучки, анализировать химический состав материала мишени и многое другое. Электронные пучки могут вызывать эмиссию атомов и молекул с облучаемой поверхности за счет испарения или сублимации вещества, термо- и электроностимулированной десорбции адсорбированных газов, разложения химических соединений и т. п.

Эффекты, возникающие при проникании электронного пучка в вещество на глубину x , определяются характером и значением потерь энергии электронов E_e в твердом теле:

$$\frac{dE_e}{dx} = \left(\frac{N_A q_e^4}{2\pi\epsilon_0^2} \right) \left(\frac{Z\rho}{M} \right) \left(\frac{1}{E_e} \ln \frac{1,66E_e}{J} \right),$$

где N_A — число Авогадро, атом/кмоль; q_e — заряд электрона, Кл; ϵ_0 — диэлектрическая проницаемость вакуума, Ф/м; Z — атомный номер; ρ — плотность вещества, кг/м³; M — молекулярная масса, кг/кмоль; $J = 13,5Z$ — потенциал ионизации, эВ.

Упругие и неупругие столкновения

В зависимости от энергии электрона имеют место:

а) упругие столкновения с атомами вещества (энергия и направление движения частицы мало изменяются), когда скорость электрона $v_e > v_{орб}$, где $v_{орб}$ — орбитальная скорость электронов атома, равная $2,2 \cdot 10^6$ м/с, а средняя энергия E , передаваемая электроном атому с массой m_a , равна

$$E = E_e \frac{4m_e m_a}{(m_e + m_a)^2} \sin^2 \frac{\Theta}{2},$$

где Θ — угол рассеяния электрона ($\approx 1^\circ$), при этом $E/E_e \approx 10^{-9} \dots 10^{-10}$;

б) неупругие столкновения (изотропное или диффузное рассеяние), когда $v_e < v_{орб}$ и потеря энергии происходит дискретно на межатомном расстоянии $0,3 \dots 0,4$ нм, что вызывает возбуждение, ионизацию, диссоциацию, ассоциацию атомов и молекул вещества, колебания кристаллической решетки и т. п.

Модель Арчарда

Согласно модели Арчарда (рис. 4.2), траекторный пробег электронов R_Σ (сумма пробегов электрона от атома к атому (рис. 4.2, а)) приблизительно равен

$$R_\Sigma \approx \frac{32\epsilon_0 E_e^2}{Nq_e^4 Z},$$

где N — плотность атомов ($10^{28} \dots 10^{29} \text{ м}^{-3}$), а проекция пробега R на ось x составляет

$$R \approx \frac{E_e^2}{\rho b},$$

где b — константа торможения, равная примерно $5 \cdot 10^5 \text{ кэВ}^2 \cdot \text{см}^2/\text{г}$ (E_e измеряется в кэВ; ρ — в $\text{г}/\text{см}^3$).

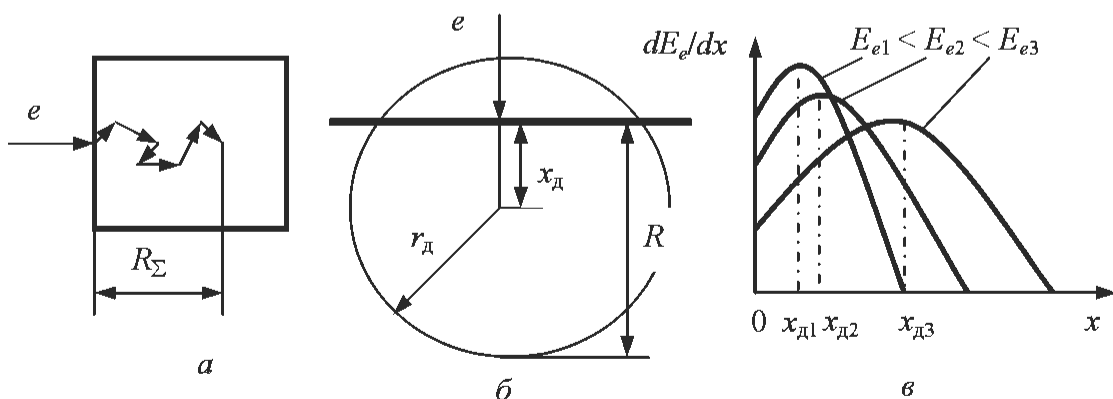


Рис. 4.2. Модель Арчарда:

а — траектория электрона; б — схема модели Арчарда; в — потери энергии электронов по глубине материала в зависимости от значения этой энергии

Таким образом, модель Арчарда позволяет рассчитать глубину проникания электрона x_d , на которой выделяется максимум энергии, а также радиус r_d диффузного рассеяния энергии электронов (рис. 4.2 б, в):

$$x_d \approx \frac{12R}{Z+8}; \quad r_d \approx R - x_d = \frac{R(Z-4)}{Z+8}.$$

Из рис. 4.2, в следует, что с увеличением энергии электронов E_e увеличивается x_d и уменьшается максимум выделения энергии dE_e/dx .

Параметры электронных пучков

Физические эффекты в объеме твердого тела, возникающие в результате проникания электронного пучка в глубь материала, делятся на нетермические и термические. К первым относятся ионизация атомов мишени, возбуждение фононных колебаний, образование дислокаций и радиационных дефектов, активация химических реакций и увеличение проводимости полупроводников и диэлектриков; ко вторым — плавление, испарение, сварка и размерная обработка (прошивка отверстий, пазов, профилирование и т. п.).

Из нетермических эффектов наиболее часто применяют изменение структуры и свойств материала электронорезиста при облучении его остросфокусированным электронным пучком с энергией 20...30 кэВ. В электронно-лучевой литографии в качестве резиста используют органические материалы, которые при взаимодействии с ускоренными электронами либо полимеризуются (образуются более крупные молекулы), либо подвергаются деструкции (распадению на более мелкие молекулы).

Схемы различных эффектов взаимодействия ускоренных ионов с веществом показаны на рис. 4.3.

Доза облучения электронорезиста $K = j\tau$, Кл/м², где j — плотность тока, А/м²; τ — время экспонирования, с. Доза, необходимая для начала химической реакции, называется пороговой $K_{\text{пор}}$ и составляет $10^{-5} \dots 10^{-7}$ Кл/см². Контрастность электронорезиста $\gamma = [\lg(K_0/K_{\text{пор}})]^{-1}$, где K_0 — доза при 100 %-ном проявлении.

В негативном электронорезисте под действием электронного пучка создаются перекрестные связи в углерод-водородных цепоч-

ках и материал становится нерастворимым и стойким к нагреву, т. е. после проявления на подвергнутых облучению участках остается рисунок из неудаленной полимерной маски. В другой группе полимеров, при облучении которых происходит разрыв основной цепочки молекул и тем самым увеличивается растворимость материала, рисунок образуется на оставшейся части пленки — позитивного электронорезиста.

Нагрев вещества при облучении его электронами имеет несколько особенностей: максимум поглощенной энергии находится в объеме твердого тела (см. рис. 4.2, в), а не на поверхности, как, например, при лазерной обработке; с помощью термического воздействия электронов можно вызывать в веществе структурные фазовые переходы, отжиг дефектов, диффузию примесей, рекристаллизацию, плавление материала, десорбцию и испарение с поверхности атомных частиц.

Тепловое воздействие электронного пучка с энергией E_0 на твердое тело характеризуется распределением температуры по его поверхности во времени:

$$T(r, t) = \frac{E_0}{\rho c (4\pi\sigma_T t)^{3/2}} \exp\left(\frac{-r^2}{4\sigma_T t}\right),$$

где r — расстояние от центра электронного пучка до рассматриваемой точки, м; $\sigma_T = \lambda_T / (\rho c)$ — температуропроводность материала, м²/с; λ_T — коэффициент теплопроводности, Вт/(м·К); ρ — плотность материала, кг/м³; c — теплоемкость, Дж/(кг·К).

Установившаяся температура может быть определена как

$$T(r) \approx \frac{r_0^2}{2r} \frac{P_e}{\lambda_T},$$

где r_0 — радиус электронного пучка, м; P_e — удельная мощность электронного пучка, Вт/м².

Изменение температуры поверхности при воздействии электронного пучка можно рассчитать по формуле

$$\Delta T = \frac{3}{2} \frac{E_0 I_e}{\pi q_e \lambda_T R},$$

где I_e — ток электронного пучка, А; q_e — заряд электрона, Кл.

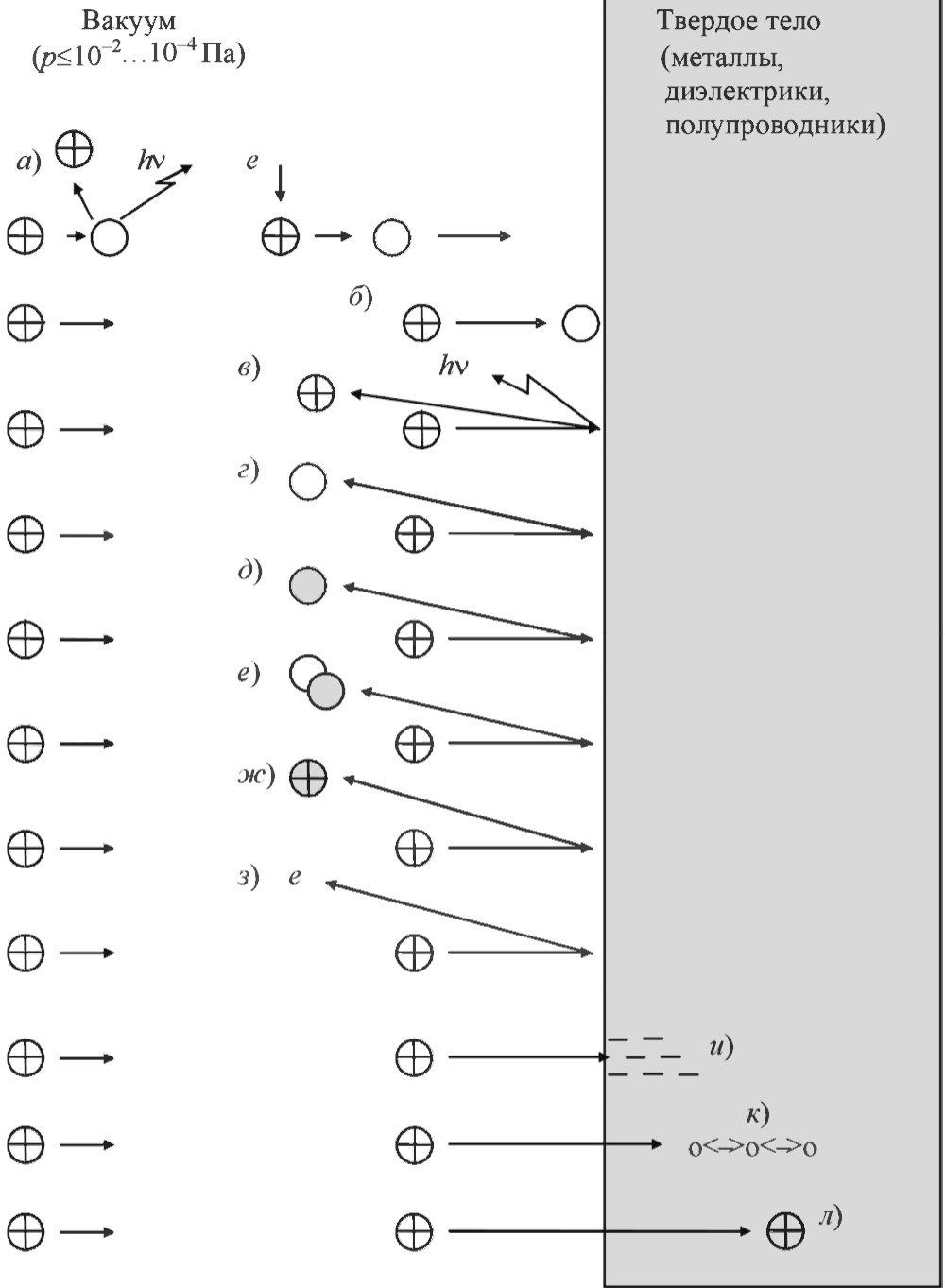


Рис. 4.3. Эффекты взаимодействия ускоренных ионов с веществом:
а — ионизация газа; *б* — осаждение ионов на поверхность; *в* — отражение иона;
г — десорбция атома или молекулы; *д* — физическое распыление атомов; *е* — химическое распыление атомов; *ж* — вторичная ионная эмиссия; *з* — эмиссия электронов; *и* — нагрев вещества; *к* — химические реакции; *л* — ионная имплантация

Скорость испарения материала

$$v_{\text{и}} = 5,38 \cdot 10^{-3} p_s \sqrt{M/T},$$

где p_s — упругость пара, Па, материала при температуре T , К; M — молекулярная масса испаряемого материала, кг/моль.

Параметры электронных пучков можно регулировать в широком диапазоне значений для проведения различных видов обработки изделий (табл. 4.1).

Несмотря на сложности выполнения электронно-лучевой обработки, связанные с необходимостью помещать объекты обработки в вакуум, она успешно конкурирует с другими методами благодаря следующим преимуществам:

- универсальности (можно обрабатывать практически любые материалы, причем не только изменять геометрические размеры деталей и свойства их поверхности, но и проводить различные измерения);

- экологической чистоте (процесс протекает в высоком вакууме, электронный пучок не вносит загрязнений и не подвержен изнашиванию, контроль с его помощью, как правило, является неразрушающим;

- управляемости (можно легко регулировать энергию, фокусировку, модуляцию и отклонение электронных пучков, причем малая инерционность позволяет быстро перемещать пучок с одного участка на другой и обеспечивать высокую скорость обработки и локальность воздействия).

Таблица 4.1

Параметры электронных пучков

Вид электронно-лучевой обработки	Энергия электронов E_e , кэВ	Диаметр пучка $d_{\text{п}}$, мм	Удельная мощность P_e , Вт/см ²
Нетермическое воздействие	20 – 250	$10^{-5} - 10^{-2}$	$10^{-2} - 10^5$
Химическое воздействие	20 – 5000	10 – 300	$10 - 3 \cdot 10^3$
Плавка	15 – 40	10 – 50	$10^3 - 10^4$
Испарение	10 – 40	2 – 30	$2 \cdot 10^3 - 2 \cdot 10^4$
Сварка	15 – 175	1 – 6	$8 \cdot 10^4 - 10^7$
Резание	20 – 150	$5 \cdot 10^{-3} - 10^{-1}$	$10^5 - 10^{10}$

Одной из наиболее важных областей использования электронно-лучевой обработки в машиностроении является сварка. Электронным пучком можно сваривать стали с высоким содержанием углерода; металлы с высокой теплопроводностью (медь и ее сплавы, золото, серебро, платину); тугоплавкие металлы (вольфрам, тантал, молибден, ниобий, титан и его сплавы); химически активные металлы (бериллий, ванадий). Наиболее характерными объектами для электронно-лучевой сварки являются элементы зубчатых зацеплений, узлы двигателей самолетов, элементы электродвигателей, электронные устройства и приборы, сильфоны и мембраны из коррозионно-стойкой стали, элементы прецизионных устройств и механизмов.

4.2. Взаимодействие ионных потоков с материалами

В зависимости от материала, формы, энергии и удельной мощности ионного пучка при взаимодействии его с веществом происходят процессы направленного изменения геометрических размеров и свойств твердого тела, а также их контроля. Ионный пучок может быть сформирован как из положительных, так и из отрицательных ионов, однако наиболее часто используют ионы, несущие положительный заряд, в связи с тем, что получить эти ионы и управлять параметрами их пучков намного проще. В качестве материала для создания ионного пучка могут быть использованы атомы практически всех элементов таблицы Менделеева, включая ионы молекул, что при возможности варьирования кратностью заряда делает ионные пучки потенциально более гибким средством воздействия на обрабатываемый или исследуемый объект по сравнению с электронными пучками.

Основные эффекты, возникающие при взаимодействии ускоренных до энергии 100 эВ ... 10 МэВ ионов с веществом и наиболее часто используемые в ионных и плазменных технологиях, можно проиллюстрировать схемой, приведенной на рис. 4.3.

Упругие и неупругие столкновения. Коэффициент аккомодации

Двигаясь в вакууме к облучаемому образцу, ионы могут сталкиваться с атомами и молекулами остаточных или рабочих газов и паров. Как и в случае с электронными пучками, может происхо-

дуть рассеяние и потеря энергии первичных частиц, однако кроме ионизации попавшегося на пути атома и испускания при этом фотона возможно явление перезарядки ионного пучка. Оно сводится к тому, что движущийся ион захватывает электрон у атома, с которым он столкнулся, и продолжает движение практически с той же скоростью, но уже в нейтральном состоянии. Эффект перезарядки ионного пучка можно осуществить и с помощью направленного на него потока электронов, нейтрализующих заряд и тем самым создают пучок ускоренных атомов, которые используются, например, для обработки диэлектриков.

Ударяющиеся о поверхность образца ионы могут оказаться связанными с ней за счет адсорбции или хемосорбции, что может быть использовано для получения тонкопленочных слоев. Вероятность того, что ударившийся о поверхность ион останется на ней, рассчитывается с помощью коэффициента аккомодации

$$\alpha = (E_i - E'_d)/(E_i - E''_d) \leq 1,$$

где E_i — кинетическая энергия иона; E'_d — энергия десорбированного атома до наступления термодинамического равновесия с подложкой; E''_d — энергия десорбированного атома после установления термодинамического равновесия с подложкой.

При столкновении с поверхностью твердого тела ионный пучок может рассеиваться атомами или группой атомов облучаемого образца, при этом происходит отклонение траектории ионов от первоначального направления и обмен энергиями между ионами и атомами мишени. В зависимости от материала взаимодействующих частиц и энергии ионного пучка обмен энергиями может быть упругим (с ядрами) и неупругим (с электронами) (рис. 4.4).

В первом случае потеря энергии иона по глубине проникания dE_n/dx , эВ/нм, рассчитывается по формуле

$$\frac{dE_n}{dx} = 0,278 \frac{Z_1 Z_2}{\sqrt{Z_1^{2/3} + Z_2^{2/3}}} \frac{M_1}{M_1 + M_2} N,$$

где Z_1 и M_1 , Z_2 и M_2 — атомный номер и молекулярная масса, кг/кмоль, ускоренного иона и атома материала облучаемого образца соответственно; N — плотность атомов, нм⁻³.

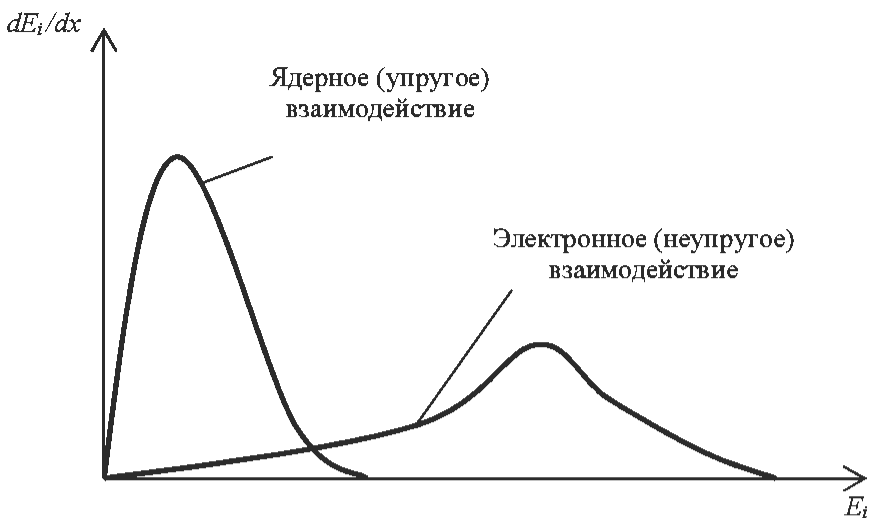


Рис. 4.4. Виды взаимодействия иона с твердым телом в зависимости от энергии иона

Максимальная энергия E_{\max} , передаваемая атому материала ионом, обладающим энергией E_i ,

$$E_{\max} = \frac{4M_1M_2}{(M_1 + M_2)^2} E_i,$$

а пробег иона R_i^n , нм, в твердом теле

$$R_i^n = \frac{3,6E_i(Z_1^{2/3} + Z_2^{2/3})(M_1 + M_2)}{NZ_1Z_2M_1}.$$

Во втором случае потерю энергии иона E_e по глубине проникания x рассчитывают по формуле

$$\frac{dE_e}{dx} = 3,28 \cdot 10^{-3} (Z_1 + Z_2) N \sqrt{\frac{E_i}{M_1}},$$

а пробег иона — по формуле

$$R_i^e = R_i^n \left[1 + \frac{2}{3} R_i^n \cdot 3,28 \cdot 10^{-3} (Z_1 + Z_2) \frac{N}{\sqrt{M_1}} \right].$$

При неупругих соударениях энергия может выделяться в виде квантов излучения — свечения (ионолюминесценция), характеристического или тормозного рентгеновского излучения. Эти эффекты используются в ионной микро- и спектроскопии. Если импульс иона достаточно велик, то он может сместить поверхностный атом из положения, где он слабо связан с кристаллической структурой образца, в положение, где связь оказывается сильнее (атомная дислокация). Ионы, обладающие более высокими энергиями, могут вызывать внутренние дислокации в толще образца.

Упругоотраженные ионы могут вызывать десорбцию слабо связанных с поверхностью атомов и молекул газов и паров, а также различных видов загрязнений. Этот эффект широко используется для очистки поверхностей деталей перед многими операциями, особенно перед процессами эпитаксии и осаждения тонких пленок.

Коэффициент распыления

Если бомбардирующие поверхность образца ионы передают настолько большой импульс, что полностью освобождают от связей один или несколько атомов, то происходит физическое распыление. Это явление наиболее эффективно проявляется при энергии иона 0,5...5 кэВ, а показателем эффективности является коэффициент распыления $S = N_a/N_i$, где N_a — количество распыленных атомов; N_i — число падающих на поверхность мишени ионов.

Коэффициент распыления зависит от максимальной энергии E_{\max} , передаваемой атому материала падающим на поверхность мишени ионом (обладающим энергией E_i), коэффициента α , энергии межатомных связей распыляемого материала U_0 . Его рассчитывают по формуле

$$S(E) = \frac{3}{4\pi^2} \alpha \frac{E_{\max}}{U_0}.$$

Здесь коэффициент α зависит от соотношения масс распыленных атомов и падающих ионов (при $M_2/M_1 = 0,01$ $\alpha = 0,17$, при $M_2/M_1 = 1,0$ $\alpha = 0,23$, при $M_2/M_1 = 10,0$ $\alpha = 5,0$); $U_0 \approx 4H_{\text{субл}}$ ($H_{\text{субл}}$ — энергия сублимации распыляемого материала).

Коэффициент распыления S зависит также от угла φ падения иона: $S(\varphi) = S(0)/\cos \varphi$, где $S(0)$ — коэффициент распыления при падении иона перпендикулярно поверхности мишени, т. е. при $\varphi = 0$. Увеличение коэффициента распыления при отклонении направления падения иона от перпендикуляра к поверхности объясняется гипотезой о парных упругих столкновениях (рис. 4.5).

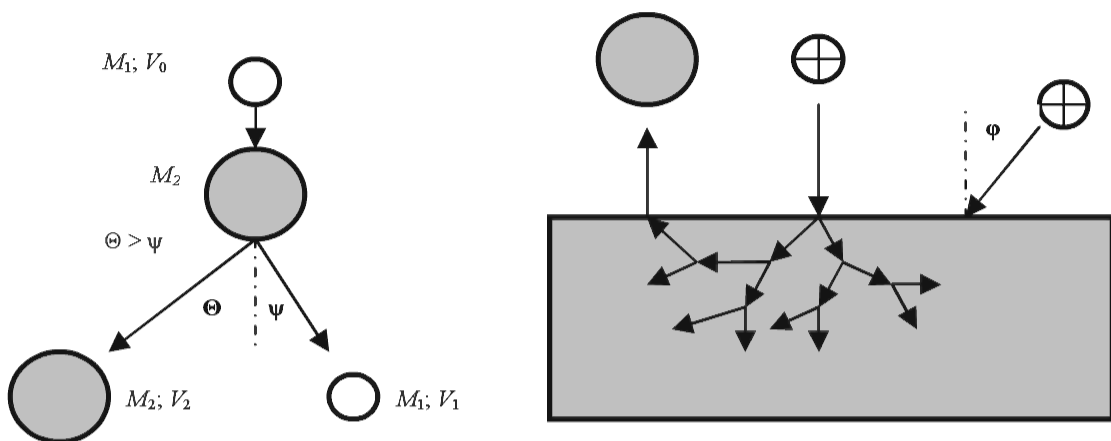


Рис. 4.5. Эффект парных соударений

Эффект парных соударений

Согласно гипотезе парных соударений, при столкновении ускоренного иона, обладающего молекулярной массой M_1 , с атомом материала мишени, молекулярная масса которого M_2 , угол отклонения иона Ψ меньше угла отклонения атома Θ :

$$\Theta = \frac{1}{2} \left(\pi - \Psi - \arcsin \frac{\sin \Psi}{\mu} \right);$$

$$\cos \Psi = \frac{1}{2} \left[(1 - \mu) \frac{V_0}{V_1} + (1 + \mu) \frac{V_1}{V_0} \right];$$

$$\mu = \frac{M_2}{M_1}.$$

Таким образом, при $M_1 < M_2$ вероятность выхода иона увеличивается за счет увеличения углов Θ и Ψ ; при $M_1 > M_2$ вероят-

ность выхода иона меньше, поскольку движение атомов зарождается на большей глубине. При этом атомы или группы атомов в приповерхностном слое толщиной порядка 1 нм движутся в направлении границы раздела «твердое тело — вакуум» и выходят из мишени. Вылетающие атомы или кластеры (группы атомов) могут находиться в нейтральном, заряженном или возбужденном состоянии.

Явление ионного распыления широко используется для строго дозированного удаления вещества с определенных участков обрабатываемого изделия (ионное фрезерование), нанесения тонких пленок из большой номенклатуры материалов в вакууме, анализа распределения распыленных частиц по массам.

Пучок химически активных ионов на основе H^+ , O^+ , N^+ , Cl^+ , F^+ и других элементов может вступать в химическую реакцию с атомами облучаемого образца и образовывать на его поверхности новые химические соединения, в том числе и газообразные. Этот эффект называется химическим ионным распылением и широко используется в операциях ионно-химического травления материалов и ионно-химического осаждения тонких пленок.

Если при ионном распылении поверхностные атомы возбуждаются до ионизированных состояний и покидают образец, то имеет место вторичная ионная эмиссия. Сила тока вторичных ионов моноизотопа распыленного вещества равна

$$I_{i2} = \eta kCS,$$

где η — чувствительность датчика; k — коэффициент ионизации ($10^{-5} \dots 10^{-1}$); C — концентрация i -го элемента в мишени; S — коэффициент распыления i -го элемента (0,1...10,0). Этот эффект лежит в основе принципа действия анализатора химического состава вещества.

При ионной бомбардировке металлических поверхностей возникает вторичная электронная эмиссия, которая является естественным средством поддержания самостоятельного газового разряда. Это явление используется также в ионной микроскопии и спектроскопии.

Если энергия, переданная пучком ионов атомам мишени, недостаточна для их распыления, то атомы начинают колебаться около положения равновесия, растрачивая энергию на нагрев ок-

ружающей среды, что приводит к нагреву мишени. При достаточной мощности, выделяемой на единицу поверхности мишени, возникают процессы (плавление и испарение вещества мишени, термоэмиссия электронов, термоизлучение и т. п.), аналогичные процессам, происходящим при облучении мишени электронами. Примером термической обработки с помощью ионных пучков может служить протонный отжиг полупроводниковых материалов. Однако они, как правило, не выдерживают конкуренции с нагревом электронными пучками, получение которых значительно проще.

Глубина проникания ионов

Ионные пучки подобно электронным могут изменять химический состав вещества, в которое они проникают. Если в качестве материала использовать ионорезист, а ионный пучок сфокусировать в пятно малого диаметра, то можно получить изображение, размеры которого составят доли микрометра.

При энергии ионов более 30 кэВ они могут проникать в глубину образца. На этом явлении основан процесс, который называется ионной имплантацией или ионным легированием (рис. 4.6 и 4.7). Независимо от материала мишени в нее можно внедрить атомы практически любых элементов таблицы Менделеева, что широко используется для изменения механических, электрических, химических, оптических, эмиссионных и других свойств вещества. Концентрация имплантированных ионов определяется дозой легирования $D = j_i t$, где D измеряется в Кл/м²; j_i — плотность ионного тока, А/м²; t — время легирования, с. На практике принята следующая градация процесса по дозам легирования: $D < 10^{10}$ Кл/см² — малая доза, $D = 10^{14}$ Кл/см² — средняя доза, $D > 10^{17}$ Кл/см² — большая доза.

Глубина проникания ионов зависит от их энергии и атомного номера, а также от кристаллографической ориентации атомов мишени по отношению к направлению движения ионов. Так, с увеличением энергии увеличивается глубина проникания ионов (рис. 4.8, а, б), а более легкие ионы (например, ионы бора), проникают глубже, чем тяжелые (например, ионы сурьмы). При определенном расположении атомов кристаллической решетки наблюдается эффект

каналирования (рис. 4.9), т. е. глубокого проникания ионов в материал мишени. Критический угол падения ионов, при котором имеет место этот эффект, рассчитывают по формуле

$$\Psi_{кр} = \sqrt{\frac{2Z_1 Z_2 q_e^2}{E_i d}},$$

где $q_e = 1$ — заряд электрона; E_i — энергия иона, кэВ; d — расстояние между атомами, нм.

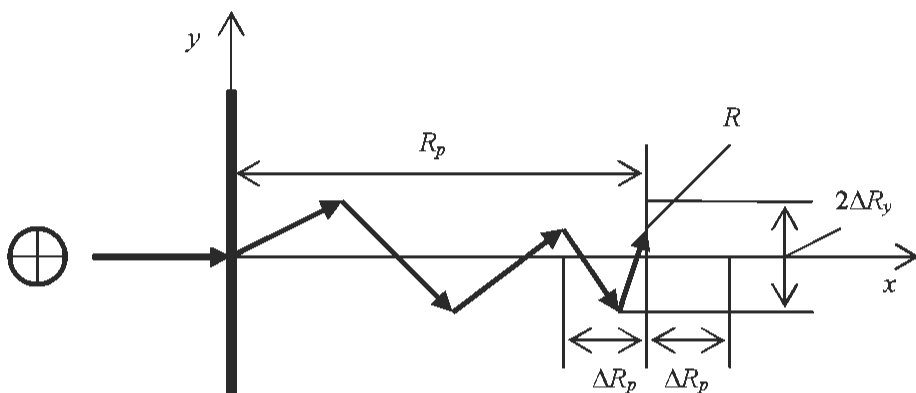


Рис. 4.6. Схема пробега имплантированного иона при нормальном законе распределения:

R — суммарная длина пробега иона; R_p — проекционный пробег иона (средняя глубина); ΔR_p — разброс значений проекционного пробега иона; ΔR_y — разброс значений пробега иона по оси y

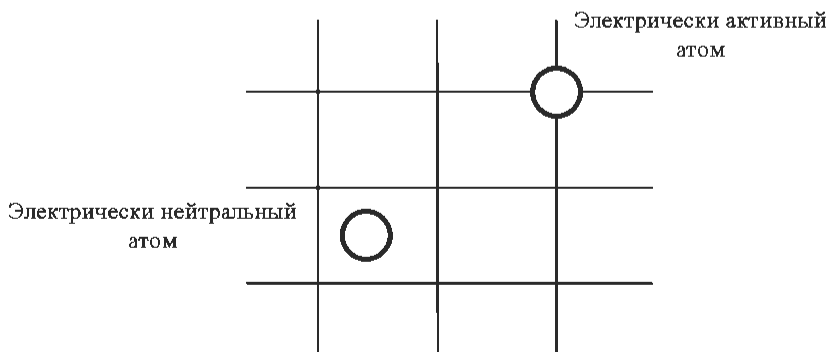


Рис. 4.7. Схема размещения имплантированного иона в кристаллической решетке

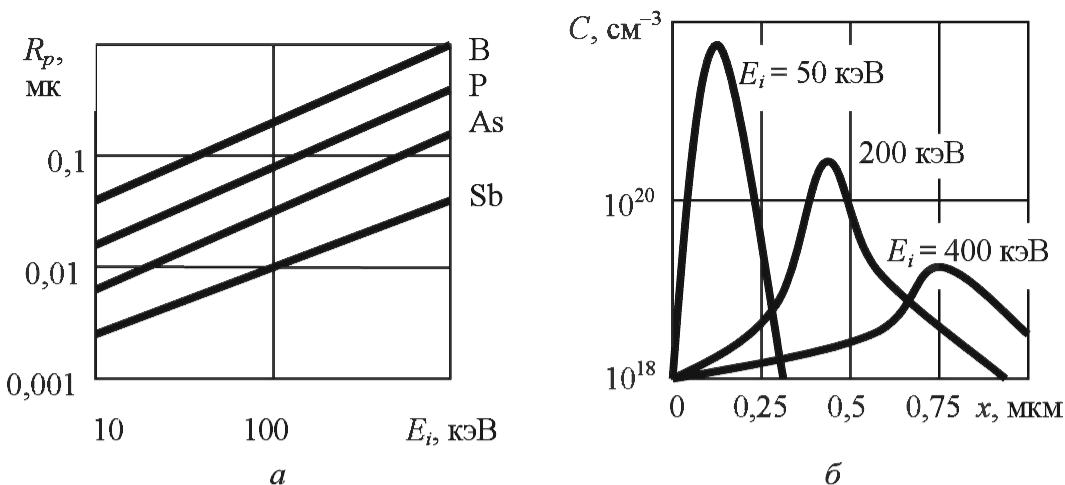


Рис. 4.8. Глубина легирования кремния ионами бора, фосфора, мышьяка и сурьмы (а) и концентрационный профиль (б) в зависимости от энергии ионов

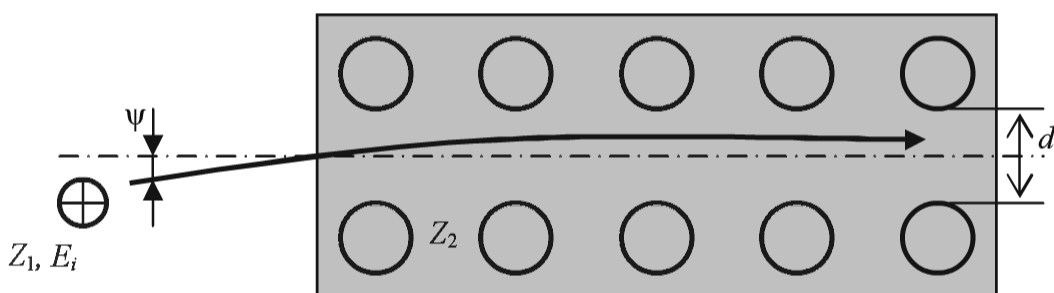


Рис. 4.9. Эффект каналирования

Так как имплантация ионов — процесс термодинамически неравновесный, то можно создавать соединения, которые принципиально невозможно получить диффузией или металлургическим способом, а также достичь концентрации имплантированного материала, существенно превышающей предел растворимости данной примеси в материале мишени.

Этот эффект используется в микроэлектронике для получения p - n -переходов в полупроводниковых материалах: в машиностроении — для легирования сталей, в других областях — для изменения в широких пределах свойств материалов.

Концентрация имплантированных в материал атомов примеси распределяется по нормальному закону:

$$C(x, y) = \frac{D}{\sqrt{2\pi} \cdot \Delta R_p} \exp \left[-\left(\frac{x - R_p}{\Delta R_p} \right)^2 \right] \exp \left(-\frac{y^2}{2\Delta R_y} \right),$$

где x — координата, перпендикулярная поверхности мишени, м; y — координата, параллельная поверхности мишени, м; D — доза имплантации, ион/м²; ΔR_p — средний квадратичный разброс значений глубины проникания ионов (по координате x), м; R_p — математическое ожидание глубины проникания ионов (проекционный пробег), м; ΔR_y — средний квадратичный разброс значений пробега ионов по координате y , м.

В микроэлектронике ионное легирование осуществляется через резистивную маску, что приводит к изменению концентрации имплантированных атомов вблизи маски (рис. 4.10) по следующему закону:

$$C(x, y) = \frac{C(x)}{2} \left[\operatorname{erfc} \left(\frac{y - a}{\Delta R_y \sqrt{2}} \right) - \operatorname{erfc} \left(\frac{y + a}{\Delta R_y \sqrt{2}} \right) \right],$$

где erfc — стандартная функция ошибок.

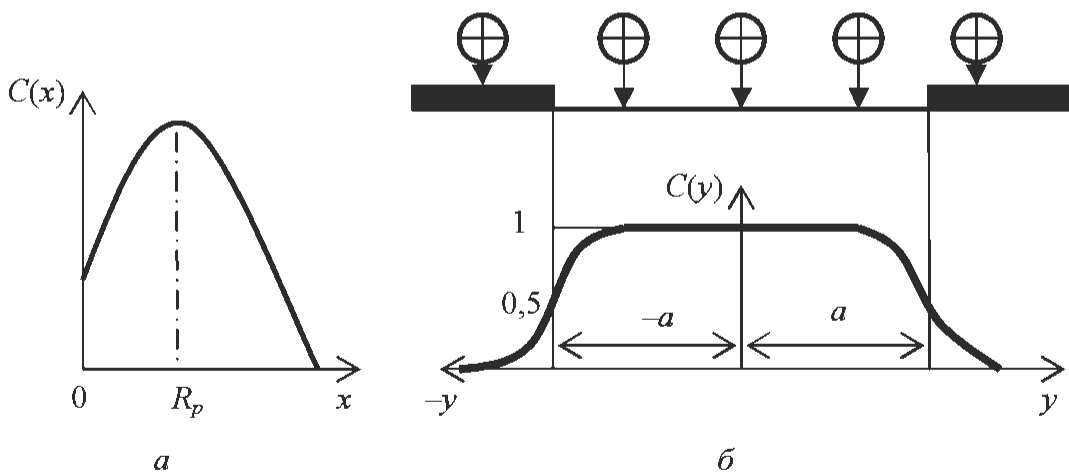


Рис. 4.10. Схема изменения концентрации легирующей примеси на линии маски:

a — распределение концентрации ионов $C(x)$ в глубь подложки; b — схема процесса имплантации ионов в подложку через окно маски шириной $2a$ и распределение концентрации ионов $C(y)$ по оси y

Таким образом, в зависимости от энергии пучков электронов и ионов меняется характер проникания в материал ускоренных электронов и ионов, а также состояние приповерхностных слоев материала, что служит основой для выбора метода и расчета режимов технологического процесса.

Контрольные вопросы и задания

1. От каких факторов зависит глубина проникания ускоренных электронов в твердое тело? Опишите модель Арчарда.
2. Как распределяется температура по поверхности твердого тела при его нагреве электронным пучком?
3. От каких факторов зависит глубина проникания ускоренных ионов в твердое тело?
4. Расскажите об эффекте парных соударений при ионном распылении материалов.
5. Чем полная глубина пробега иона в твердом теле отличается от проекционного пробега?
6. Объясните сущность эффекта каналирования при проникании ускоренного иона в кристаллический материал.
7. По какому закону распределяются в твердом теле имплантированные атомы?

5. ФОРМИРОВАНИЕ ГАЗОРАЗРЯДНОЙ ПЛАЗМЫ И ЕЕ ВЗАИМОДЕЙСТВИЕ С МАТЕРИАЛАМИ

Изучив материал этой главы, студент должен иметь представление о свойствах и возможностях газоразрядной плазмы как инструмента микро- и нанотехнологий.

Студент должен знать:

- состав газоразрядной плазмы и условия, необходимые для ее «зажигания»;
- причины возникновения темного катодного пространства в газоразрядной плазме;
- способы ионизации вещества.

Студент должен уметь:

- анализировать сущность явления ионизации электронным ударом;
- применять на практике закон Пашена, например для определения минимального напряжения зажигания газового разряда;
- оценивать применимость высокочастотной плазмы в технологических процессах.

Студент должен иметь навыки работы с источниками технической информации, используемыми для поиска значений основных параметров газоразрядной плазмы при применении ее в микро- и нанотехнологиях.

С помощью газоразрядной плазмы можно осаждать металлические и диэлектрические пленки, стимулировать осаждение из газовой фазы диэлектрических пленок, пленок переходных металлов и их силицидов, вытравливать материал через резистивную маску после операций микролитографии, а также получать ионные и электронные пучки большой интенсивности. Плазменная обработка, заменившая жидкостное травление, получила название «сухое травление».

Плазменная обработка осуществляется при давлении ниже атмосферного и поэтому совместима с другими «вакуумными» процессами — электронно- и ионно-лучевыми, лазерными, рентгеновскими и другими операциями. Микротопология или микрорельеф на обрабатываемых изделиях формируется повторением цикла, включающего три группы операций: 1) получение, обработка и легирование тонких пленок и слоев; 2) микролитография (фото-, электроно-, ионо- и рентгенолитография); 3) травление топологического рисунка или микрорельефа. Благодаря использованию сухого травления геометрические размеры рисунка могут быть получены с погрешностью менее 0,1 мкм. Для определения момента окончания травления, контроля протекания химических и физических процессов в плазме, измерения скорости осаждения пленок и других параметров применяются различные методы диагностики плазмы.

Газоразрядная плазма (рис. 5.1), состоящая из электронов, ионов и электрически нейтральных атомов, молекул и радикалов, генерирующая различные виды излучений, также может служить инструментом для обработки материалов.

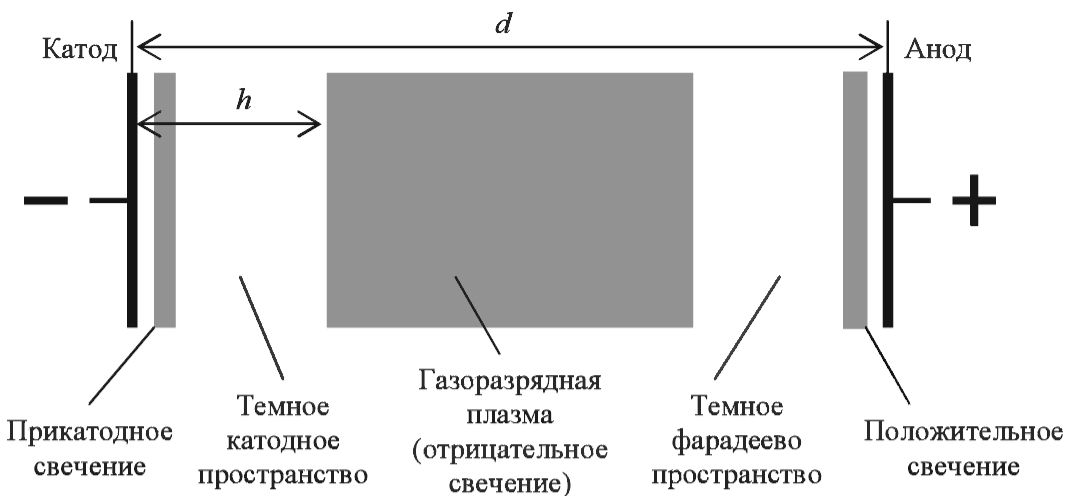


Рис. 5.1. Низкотемпературная газоразрядная плазма

5.1. Параметры газоразрядной плазмы

Параметрами газоразрядной плазмы являются состав и концентрация частиц, температура электронов и ионов, плазменное давление и др. В плазменных технологиях в качестве рабочего газа для обработки материалов чаще других используются Ar , O_2 , N_2 , H_2 , CF_4 , CCl_4 , SiH_4 , различные углеводородные соединения C_xH_y при давлении $0,65 \dots 250$ Па; концентрация ионов в плазме составляет порядка 10^{10} ион/см³, а электронов — $10^8 \dots 10^{10}$ эл/см³; энергия электронов может составлять $1,2 \dots 30,0$ эВ, а частота ВЧ-плазмы может изменяться в пределах $3,5 \dots 27,0$ МГц.

Различные виды плазменной обработки материалов зависят от энергетических характеристик плазмы и доминирующего влияния одного из эффектов в пространстве между областью газового разряда и электродами (рис. 5.2).

Важным параметром плазмы является энергия E_e (температура kT_e) электронов, которая имеет нормальное распределение (рис. 5.3, а) и от которой зависит вероятность образования ионов — коэффициент ионизации (рис. 5.3, б).

Длина свободного пробега электрона в плазме

$$\lambda = \frac{1}{S_e N},$$

где S_e — сечение столкновения электрона с атомами плазмообразующего газа, м^2 (для аргона при $E_e = 2$ эВ, $S_e = 3 \cdot 10^{-16} \text{ см}^2$); N — концентрация атомов или молекул газа, равная $N = p/(kT) = 2,5 \times 10^{16} \text{ атом/см}^3$ (при давлении аргона $p = 100$ Па и температуре стенок вакуумной камеры $T = 293$ К). При этих параметрах плазмы $\lambda = 0,13 \text{ см}$.

Среднюю скорость электронов в плазме рассчитывают по формуле

$$v_e = \sqrt{\frac{8kT_e}{\pi m_e B}},$$

где m_e — масса электрона при скорости v_e ; $B = 6,24 \cdot 10^{11} \text{ эВ} \cdot \text{с}^2 / (\text{г} \cdot \text{см}^2)$ — коэффициент (при $E_e = kT_e = 2$ эВ $v_e = 10^8 \text{ см/с} = 1000 \text{ км/с}$). Частота столкновений электронов с молекулами газа в плазме $\nu = v_e S_e N$ и составляет $7,5 \cdot 10^8 \text{ с}^{-1}$ или 750 МГц при приведенных выше значениях параметров плазмы.

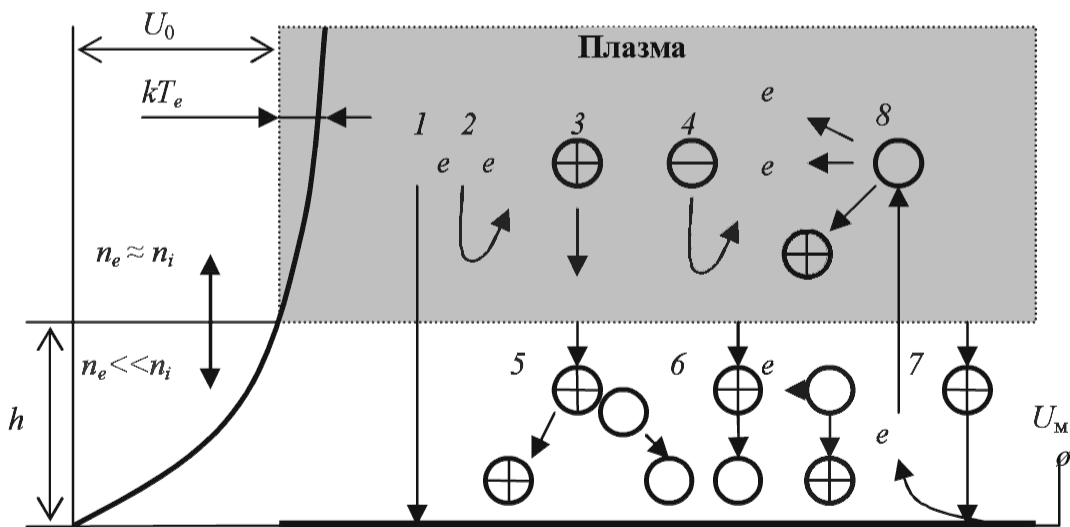


Рис. 5.2. Явления в газоразрядной плазме и ее энергетические характеристики: 1 — уход быстрого электрона; 2 — отражение медленного электрона; 3 — инжекция иона; 4 — отражение отрицательного иона; 5 — рассеяние на нейтральной частице; 6 — обмен заряда иона с нейтральной частицей; 7 — эмиссия вторичного электрона; 8 — ионизация электронным ударом; U_0 — прикатодный потенциал; h — толщина прикатодной области — темного катодного пространства; n_e и n_i — концентрация соответственно электронов и ионов; kT_e — энергия электрона; U_M — потенциал мишени

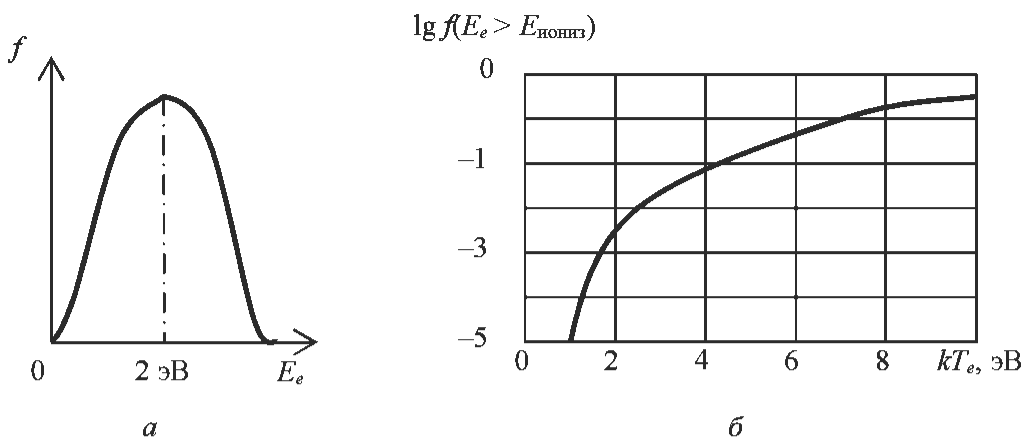


Рис. 5.3. Распределение электронов в газоразрядной плазме по энергиям (а) и вероятность ионизации атомов аргона (б)

Нейтральные частицы в газоразрядной плазме характеризуются скоростью образования радикалов (электрон-молекулярного взаимодействия)

$$R = Kn_e N,$$

где K — константа скорости диссоциации (для HCl при $kT_e = 2$ эВ $K = 4 \cdot 10^{-10}$ см³/с). При $p_{cl} = 100$ Па и $n_e = 10^{10}$ см⁻³ $N = 2,5 \cdot 10^{16}$ см⁻³, а $R = 7,5 \cdot 10^{16}$ шт/(см³·с). Длина, м, свободного пробега атомов или молекул

$$\lambda = \frac{6,51 \cdot 10^{-3}}{p},$$

где p — давление газа, Па.

Число атомов или молекул, ударяющихся о единицу поверхности в единицу времени,

$$N_1 = \frac{p}{\sqrt{2\pi kTM}},$$

где T — температура газа, К; M — молекулярная масса газа, кг/кмоль.

5.2. Кривые Пашена

Напряжение зажигания U_z самостоятельного газового разряда зависит от рода газа и произведения давления p и расстояния d

между электродами. Эта зависимость иллюстрируется кривыми Пашена (рис. 5.4).

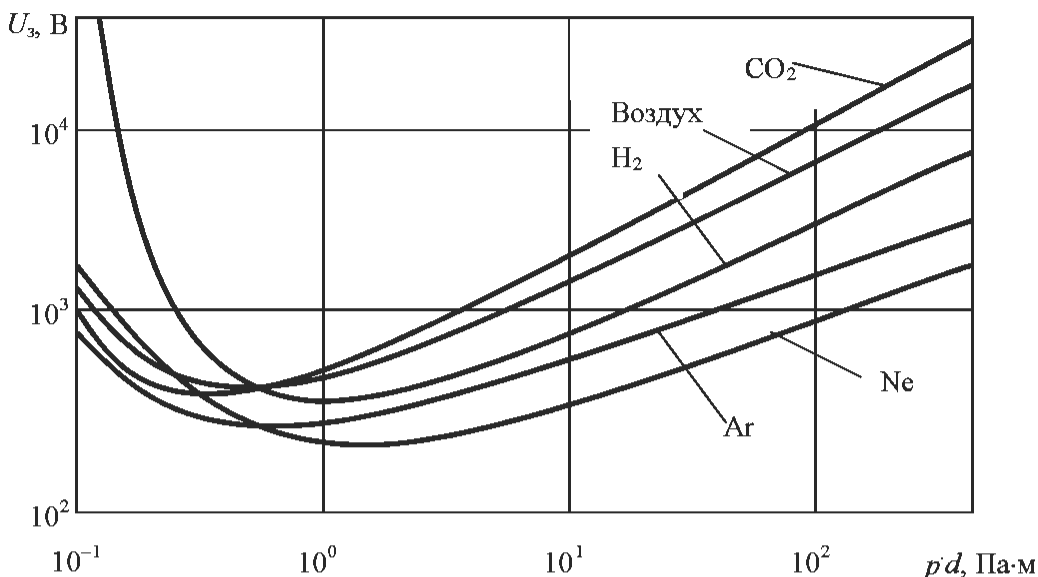


Рис. 5.4. Зависимость потенциала зажигания газовой разрядной лампы $U_з$ от произведения pd для различных газов (кривые Пашена)

Поток ионов Q_i , ион/($м^2 \cdot с$), в катодное пространство приблизительно равен

$$Q_i \approx n_i \sqrt{\frac{kT_i}{M_i}},$$

где kT_i — энергия иона, Дж; M_i — молекулярная масса иона, кг/кмоль.

Плотность ионного тока $j_i = Q_i q_e$, А/ $м^2$, а толщина прикатодной области h , см, составляет приблизительно

$$h \approx \sqrt{\frac{5,47 \cdot 10^{-8} \cdot U_M^{3/2}}{M_i^{1/2} j_i}},$$

где U_M — потенциал мишени, В.

Параметры газоразрядной плазмы сильно зависят от давления плазмообразующего газа (рис. 5.5).

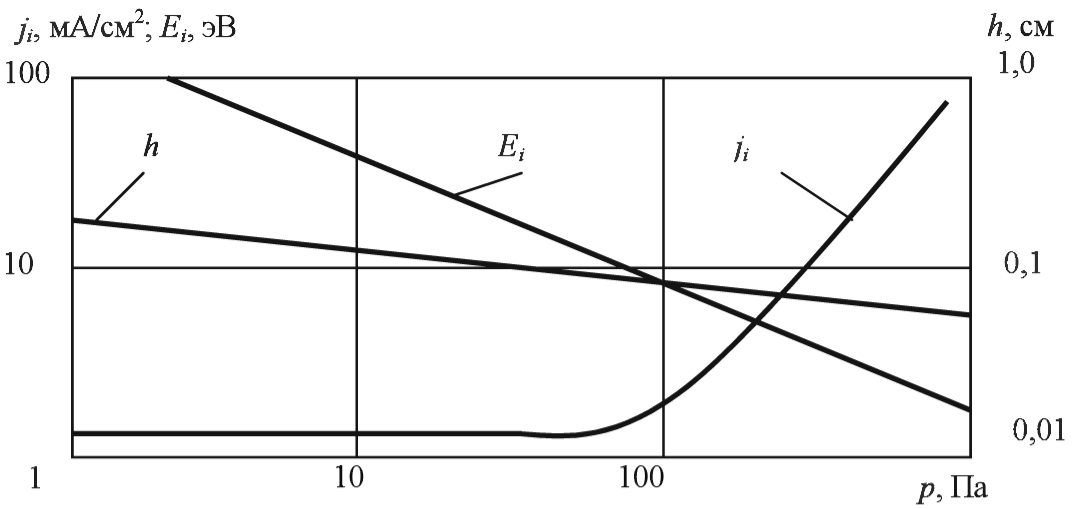


Рис. 5.5. Зависимость параметров газоразрядной плазмы (удельного ионного тока j_i , энергии ионов E_i и толщины прикатодной области h) от давления плазмообразующего газа p

5.3. Высокочастотная плазма

Высокочастотная плазма формируется при условии, что пробег электронов равен расстоянию d между электродами (рис. 5.6), а пробег значительно более тяжелых ионов намного меньше. Необходимую частоту изменения полярности на электродах определяют с учетом неравенства $f > 1/t_e$, где время пробега электронов $t_e = d/v_e$ (при $v_e = 10^6$ м/с и $d = 0,1$ м $t_e = 10^{-7}$ с, а $f > 10^7$ Гц). Наиболее распространена стандартная частота $f = 13,56$ МГц, при которой время t_i пробега ионами расстояния d равно $t_i = d/v_i$ (при $v_i \approx 500$ м/с $t_i = 2 \cdot 10^{-4}$ с), а пробег ионов за t_i составляет $d_i = v_i/f$ ($d_i = 3,7 \cdot 10^{-5}$ м, или 37 мкм).

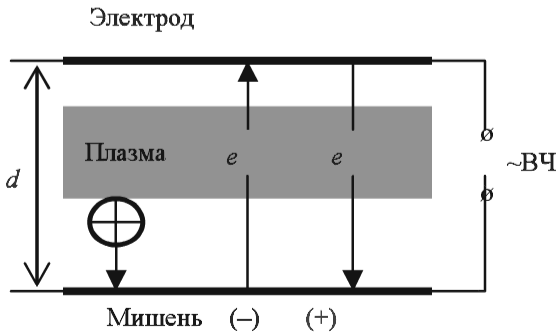


Рис. 5.6. Схема высокочастотной плазмы

Физические явления в газоразрядной плазме намного разнообразнее, чем представленные в этой главе, и изучаются специальной научной дисциплиной «Физика плазмы». Те немногие приведенные здесь сведения необходимы лишь для понимания сути плазменных технологий, применяемых при производстве электронных приборов.

Контрольные вопросы и задания

1. Из каких частиц состоит газоразрядная плазма? Каковы условия, необходимые для ее «зажигания»?
2. Перечислите причины возникновения темного катодного пространства в газоразрядной плазме.
3. Поясните сущность явления ионизации электронным ударом. Какие еще существуют способы ионизации вещества?
4. От чего зависит минимальное напряжение зажигания газового разряда (закон Пашена)?
5. Объясните сущность, способ получения, основные характеристики высокочастотной плазмы.

6. ФОРМИРОВАНИЕ ОПТИЧЕСКОГО ИЗЛУЧЕНИЯ

Изучив материал этой главы, студент должен иметь представление о свойствах и параметрах оптического излучения применительно к использованию его в микролитографии.

Студент должен знать:

- законы, понятия и определения скалярной теории дифракции, являющиеся основой формирования микроизображений при микролитографии;
- методы преобразования фундаментальных законов оптики в прикладные технологические зависимости.

Студент должен уметь:

- анализировать физические закономерности, определяющие процесс микролитографии в ближнем поле;
- математически описывать оптические процессы формирования микроизображения;

- рассчитывать профили распределения пространственной интенсивности излучения при контактной микролитографии и литографии с микрозазором;
- оценивать применимость оптических систем для реализации процессов микролитографии с заданными параметрами.

6.1. Формирование микрорельефа в резисте

Как показано во введении, микролитография, выполняя роль универсального метода локализации групповой микрообработки, является ключевым процессом микротехнологий. Возможность одновременного и относительно быстрого получения огромного числа элементов структур с микрометровыми и субмикрометровыми размерами оказала революционное воздействие на развитие микротехнологий. Поэтому микролитография считается важнейшей технологией современного микроэлектронного производства, ее стоимость достигает 35 % от общей стоимости производства.

Процесс микролитографии основан на изменении свойств резиста при экспонировании его соответствующим излучением. Формирование микрорельефа в резисте осуществляется поэтапно (рис. 6.1).

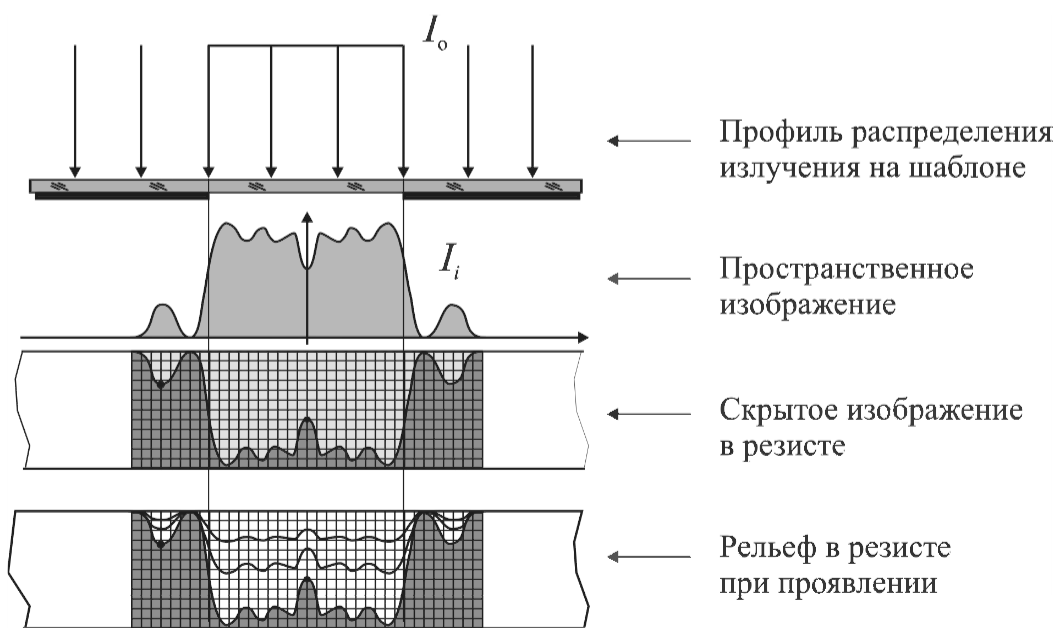


Рис. 6.1. Этапы формирования микрорельефа в резисте

При экспонировании актиничное излучение проходит через фотошаблон и оптическую систему, при этом формирующийся на поверхности фоторезиста профиль распределения интенсивности пространственного изображения может существенно отличаться от заданного на фотошаблоне.

Излучение поглощается резистом, в нем происходят фотохимические реакции (сшивка или деструкция), которые вызывают необратимые химические изменения и приводят к образованию скрытого изображения в виде участков с модифицированными свойствами.

При проявлении происходит поверхностное растворение, т. е. локальное удаление, участков резиста и образование микрорельефа.

Таким образом, в процессе микролитографии можно выделить следующие этапы, основанные на принципиально различающихся физико-химических явлениях:

- формирование пространственного оптического изображения;
- деструкция резиста;
- проявление резиста.

Первый этап определяет пространственное распределение интенсивности излучения в плоскости резиста и служит определяющим при достижении требуемых параметров микрорельефа. Именно поэтому последующие параграфы посвящены рассмотрению методов расчета распределения интенсивности излучения, которое существенно зависит от вида и системы экспонирования.

6.2. Системы экспонирования

Как показано в предыдущем параграфе, для получения в резисте микрорельефа с заданной шириной и наклоном краев профиля необходимо прежде всего сформировать оптическое пространственное микроизображение, затем передать его в резист и далее превратить в микрорельеф.

Излучение, экспонирующее резист, называется актиничным. Оно является своеобразным инструментом, который формирует микрорельефы, сопоставимые по размерам с длиной волны самого излучения. Поэтому постоянным стремлением разработчиков систем микролитографии является уменьшение длины волны используемого излучения.

К настоящему времени разработаны различные варианты микролитографии: фотолитография, рентгенолитография, электронолитография, ионолитография. В каждом из этих процессов используются соответствующие источники актиничного излучения и резисты.

В фотолитографии применяется ультрафиолетовое излучение (УФ-излучение, UV) — электромагнитное излучение, занимающее диапазон между видимым и рентгеновским излучением.

Диапазон УФ-излучения имеет несколько градаций длин волн (обозначения диапазонов в скобках соответствуют обозначениям, принятым в англоязычной технической литературе):

- ближнее (*Near UV — NUV*): 435...330 нм;
- среднее (*Mid UV — MUV*): 330...280 нм;
- дальнее или вакуумное (*Deep UV — DUUV*): 100...200 нм;
- экстремальное (*Extreme UV — EUV*): 5...100 нм.

Традиционные источники ближнего УФ-излучения — ртутные лампы высокого давления. Характерными линиями спектра таких ламп являются: *g*-линия (длина волны 435 нм), *h*-линия (405 нм), *i*-линия (365 нм).

Источниками дальнего УФ-излучения являются эксимерные лазеры. В зависимости от состава рабочей смеси они имеют следующие длины волн: KrF — 248 нм, ArF — 193 нм, F₂ — 157 нм. Ультрафиолетовое излучение со столь малой длиной волны сильно поглощается как воздухом, так и стеклом, поэтому оптические элементы таких источников делают из кварца или фтористого кальция CaF₂.

Экстремальное УФ-излучение близко к мягкому рентгеновскому излучению. Его основным источником является инициированная мощным лазером плазма.

Рентгенолитография основана на применении рентгеновского излучения с длиной волны 1...5 нм, получаемого при воздействии электронным пучком на вращающуюся мишень. Оно также может быть получено из плазменных источников или на выходе накопительных колец синхротронов.

В электронолитографии используются сфокусированные потоки электронов или ионов с длиной волны $10^{-3}...10^{-5}$ нм. Источники — электронные или ионные оптические системы — должны обеспечивать рабочее напряжение 25...100 кэВ.

Среди перечисленных видов микролитографии доминирующее положение занимает фотолитография. Она является основным методом промышленного производства микроструктур с размерами элементов от десятков микрометров (печатные платы) до десятков нанометров (твердотельные интегральные микросхемы). Поэтому основное внимание при последующем рассмотрении будет уделено именно этому виду микролитографии.

По способу передачи изображения шаблона на поверхность фоторезиста можно выделить контактное и проекционное экспонирование (рис. 6.2).

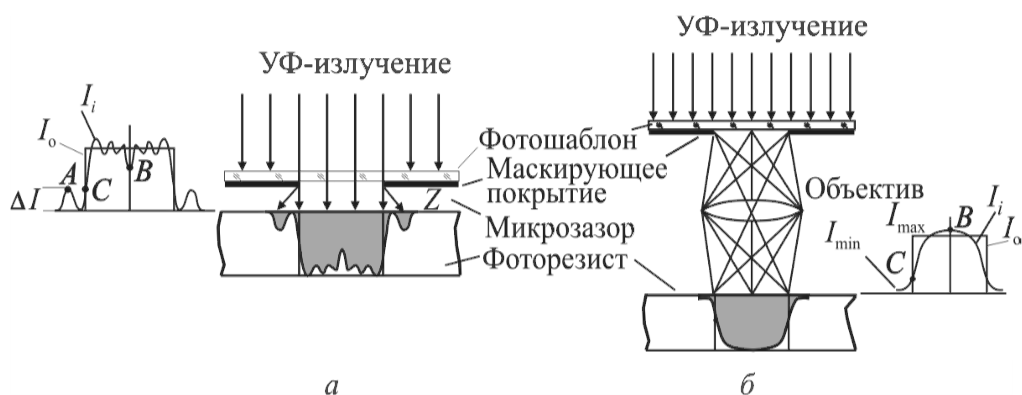


Рис. 6.2. Схемы контактного (а) и проекционного (б) экспонирования

Контактное экспонирование выполняется при прижме фотошаблона к плоской заготовке — пластине. При этом необходимо учитывать, что в разных областях рабочего поля между ними остаются микрозазоры. В зависимости от силы прижима микрозазор составляет от 0,2 мкм при жестком контакте до 15 мкм при мягком контакте и до 25 мкм при экспонировании с гарантированным микрозазором.

Наличие остаточного микрозазора ведет к дифракционным явлениям, которые размывают изображение. В результате номинальный прямоугольный профиль распределения интенсивности на объект (шаблоне) I_0 превращается в сложный, трудно прогнозируемый профиль изображения I_i на поверхности фоторезиста (рис. 6.2, а).

В распределении интенсивности можно выделить некоторые характерные точки, такие, как точка A — максимум интенсивности в затененной области, точка B — наименее освещенная точка свет-

лого поля, точка C , отвечающая краю геометрической тени. При завышенном времени экспонирования экспозиция в характерной точке A может оказаться достаточной для полного удаления фоторезиста при проявлении. В этом случае в области тени, там, где требуется сохранить маскирующую пленку фоторезиста, появятся сквозные полосы обнаженной подложки (двойной край).

При заниженном времени экспонирования экспозиция в точке B может оказаться недостаточной и в освещенной области изображения сохранятся валики не удаленного при проявлении фоторезиста.

Проекционное экспонирование исключает контакт фотошаблона с пластиной, так как они находятся на значительном расстоянии друг от друга, а изображение проецируется в плоскость фоторезиста оптическим объективом. Однако оптические явления, прежде всего дифракционные, искажают профиль распределения интенсивности изображения (рис. 6.2, б).

Соотношения значений интенсивности в центре освещенной области I_{\max} и в затененной области I_{\min} определяют контраст изображения. Как и в контактной фотолитографии, по этому параметру оценивают возможность получения рельефа с требуемым наклоном края профиля.

Для оценки разрешения проекционной фотолитографии необходимо учитывать параметры оптической системы, с помощью которой формируется микроизображение. Важнейшим из них является числовая апертура A (рис. 6.3), определяемая выражением

$$A = n \sin \alpha, \quad (6.1)$$

где n — коэффициент преломления среды; α — половина угла при вершине конуса лучей, попадающих в точку изображения на оптической оси.

Разрешающая способность оптической системы обычно определяется расстоянием W между двумя соседними точками изображения, при котором максимум дифракционной картины одной точки совпадает с минимумом дифракционной картины другой точки (критерий Рэлея). В этом случае

$$W = \frac{0,61\lambda}{A}, \quad (6.2)$$

где λ — длина волны экспонирующего излучения; A — числовая апертура проекционного объектива.

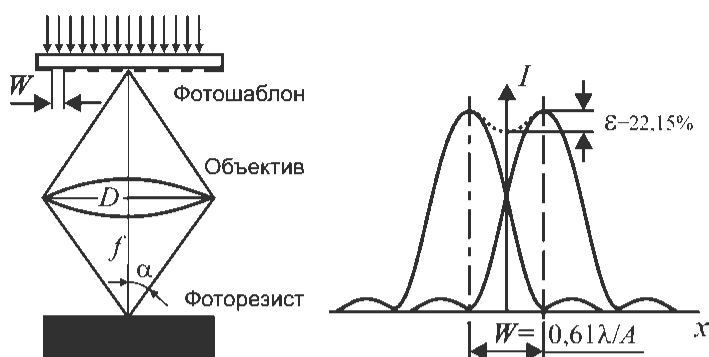


Рис. 6.3. Проекционное экспонирование:

a — схема; *б* — к определению разрешающей способности

Критерий Рэля обеспечивает различимость двух соседних точек при визуальном наблюдении изображения с помощью микроскопа. При этом освещенность в области между точками ослабляется на величину $\varepsilon = 22,15\%$ по сравнению с максимумами интенсивности излучения. При использовании высокочувствительных фотоэлементов можно различить изображения точек, расположенных вблизи абсолютного предела разрешения, когда ослабление освещенности между ними составляет $2...3\%$. В этом случае значение W рассчитывают по формуле

$$W = \frac{0,5\lambda}{A}. \quad (6.3)$$

При регистрации изображения на светочувствительных слоях (фотоэмульсии, фоторезисте) предельное разрешение будет зависеть также от свойств этих материалов и способа их фотохимической обработки. При этом, как правило, нельзя реализовать предельное разрешение объектива и результаты получаются хуже, чем при визуальной оценке. Поэтому в микролитографии обычно используют эмпирическую зависимость

$$W = \frac{k\lambda}{A}, \quad (6.4)$$

где $k = 0,6$ — для лабораторных условий; $k = 0,8$ — для производственных условий.

Глубина резкости ΔZ определяет возможное отклонение плоскости резкого изображения от идеальной плоскости изображения оптической системы и также является ее важным параметром.

Для оценки ΔZ используют второй критерий Рэля, который формулируется следующим образом: в оптической системе, создающей сферическую волновую поверхность Σ с центром в точке M , можно допустить расфокусировку $\Delta Z = MM_1$ при условии, что расстояние между волновой поверхностью Σ и сферой сравнения S с центром в точке M_1 , измеренное по краю выходного зрачка, не превышает величину $\lambda/4$ (рис. 6.4).

Для этого необходимо, чтобы предельное значение ΔZ составляло

$$\Delta Z = \pm \frac{\lambda}{2A^2}. \quad (6.5)$$

Из формул (6.4) и (6.5) следует, что повышение разрешения оптической системы за счет увеличения апертуры ведет к резкому уменьшению глубины резкости объектива. Часто этот параметр не превышает 1...3 мкм, поэтому для компенсации неплоскостности пластин необходимо применять системы автофокусировки.

Следует подчеркнуть, что разрешающая способность и глубина резкости, рассчитанные по формулам (6.4), (6.5), описывают теоретические пределы этих параметров.

Качество реальной оптической системы характеризуется степенью соответствия основных параметров объекта и его изображения, сформированного системой.

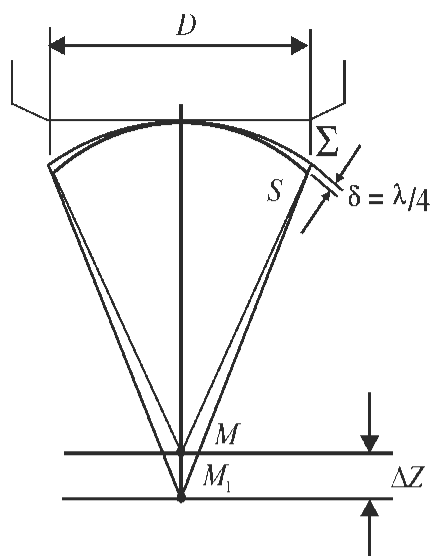


Рис. 6.4. Глубина резкости микроизображения

6.3. Основы теории формирования микроизображений

Формирование микроизображений является основным этапом микролитографии, на котором закладываются размер формируемого микрорельефа и его профиль.

Как для контактного, так и для проекционного способа экспонирования основной причиной искажений и размытия изображения является дифракция.

Излучение, проходящее через фотошаблон, дифрагирует на границах его прозрачных и непрозрачных участков, что ведет к изменению размеров элементов и формы изображения.

Для описания и оценки возникающих погрешностей необходим расчет дифракционного распределения интенсивности излучения на поверхности фоторезиста. Традиционно такой расчет выполняют на основе скалярной теории дифракции Френеля — Кирхгофа. Эта теория построена с учетом волновой природы света и основана на использовании принципа Гюйгенса, объясняющего прохождение световых волн через преграду.

Волновые процессы в оптике

Принцип Гюйгенса может быть сформулирован следующим образом: *каждая точка волнового фронта рассматривается как источник вторичных элементарных волн, которые распространяются во все стороны. Вторичный волновой фронт определяется огибающей всех вторичных волн* (рис. 6.5, а).

Если волновой фронт встречает на пути экран с отверстием, размер которого соизмерим с длиной волны света, часть фронта в отверстии можно считать единичным вторичным источником, излучающим равномерно во все стороны (рис. 6.5, б).

Когда размер отверстия превышает длину волны, полагают, что часть волнового фронта в отверстии состоит из большого числа вторичных источников с одинаковыми амплитудами и фазами (рис. 6.5, в), дающими в точке P суммарную интенсивность (рис. 6.5, г).

Когда источники излучения находятся в бесконечности, волновые фронты можно считать плоскими и параллельными (рис. 6.5, д). Если при этом размер отверстия в экране достаточно велик по

сравнению с длиной волны, вторичные источники образуют плоские перемещающиеся прямолинейно вторичные волновые фронты. В данном случае дифракционные эффекты проявляются лишь на краях отверстия.

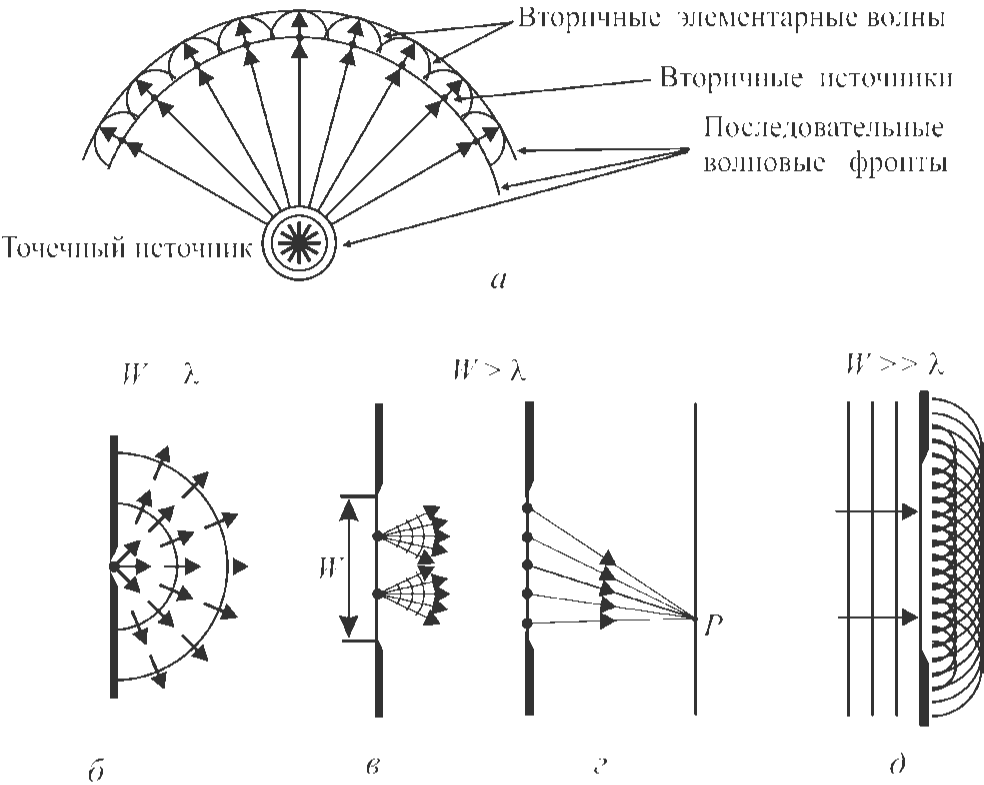


Рис. 6.5. Принцип Гюйгенса и дифракция света на отверстиях

Таким образом, принцип Гюйгенса позволяет объяснить явление дифракции света в общем виде. Чтобы дать количественную оценку интенсивности распространяющихся за преградой световых волн, Френель дополнил принцип Гюйгенса. Он ввел понятие об амплитуде и о фазе колебаний элементарных волн с учетом их интерференции.

Представление волн в векторном и комплексном виде

Начнем рассмотрение дифракции с повторения основных понятий волновых процессов, таких как гармонические колебания, векторное и комплексное представление волн.

При анализе фотолитографических задач мы будем базироваться в основном на волновых свойствах света и лишь частично, когда речь пойдет о практической реализации литографических процессов, коснемся способа описания света как потока фотонов.

Световые электромагнитные волны представляют собой возмущения электромагнитного поля, которые распространяются в вакууме со скоростью света ($c = 2,998 \cdot 10^8$ м/с). Векторы напряженности электрического и магнитного полей взаимно перпендикулярны и их значения гармонически изменяются во времени.

Гармонические колебания определяются функцией, которая описывает проекцию радиус-вектора точки, движущейся по окружности с постоянной скоростью, на диаметр окружности (рис. 6.6, а). Эта проекция определяется соотношениями:

$$U = A \sin\left(\frac{2\pi}{T}t + \varphi\right); \quad U = A \cos\left(\frac{2\pi}{T}t + \varphi^*\right), \quad (6.6)$$

где U — отклонение (возмущение) поля; A — максимальное отклонение или амплитуда; T — период колебаний; φ , φ^* — начальная фаза (отклонение в момент времени $t = 0$).

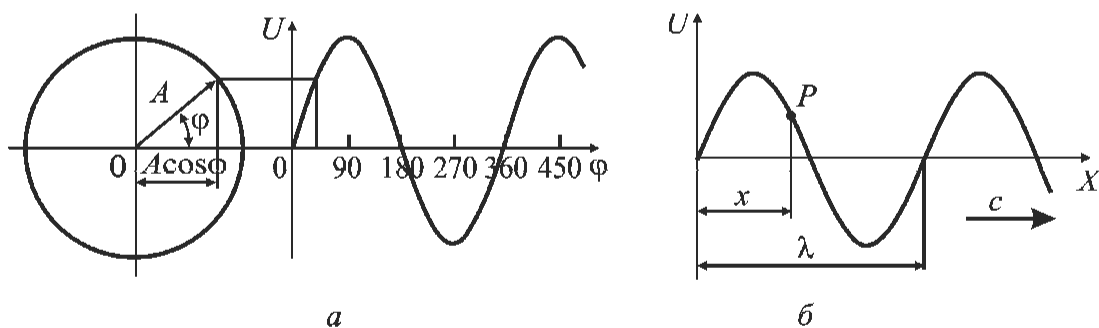


Рис. 6.6. Волновые процессы:

а — построение синусоиды; *б* — параметры поперечной волны

Колебания такого вида совершают векторы напряженности электрического и магнитного полей, однако обычно рассматривают электрический вектор, так как глаз человека реагирует именно на эту составляющую электромагнитного поля.

Уравнения (6.6) описывают гармоническое изменение вектора напряженности во времени. Однако электромагнитные возмущения помимо колебаний во времени еще и распространяются в пространстве, создавая синусоидальную кривую (рис. 6.6, б). В связи с этим уравнение (6.6) должно быть скорректировано. Дело в том, что точки волны, разнесенные вдоль оси X , совершают колебания, взаимно сдвинутые по фазе. Волна дойдет до точки с координатой x за время $t_0 = x/c$, при этом волна изменит фазу на $(2\pi/T)t_0$, и формула колебания этой точки будет иметь вид

$$U = A \sin \left[\frac{2\pi}{T}(t - t_0) + \varphi \right] = A \sin \left[\left(\frac{2\pi}{T}t - \frac{2\pi}{Tc}x \right) + \varphi \right]. \quad (6.7)$$

Расстояние между двумя ближайшими точками, колеблющимися в одной фазе, называется длиной волны λ . Время, необходимое для преодоления этого расстояния, — это период T , а величина, обратная периоду, — частота f . Отметим, что длина волны λ связана с периодом $T = 1/f$ и скоростью распространения волны очевидными соотношениями $\lambda = cT = c/f$, поэтому уравнение (6.7) можно записать таким образом:

$$U = A \sin \left[\left(\frac{2\pi}{T}t - \frac{2\pi}{\lambda}x \right) + \varphi \right]. \quad (6.8)$$

Выражение (6.8) описывает гармоническое колебание в момент времени t в точке с координатой x при распространении возмущения вдоль оси X . Когда начальные фазы всех волн совпадают или мы имеем дело с одной волной, можно положить $\varphi = 0$.

Обратим внимание на двоякую пространственно-временную периодичность волнового процесса. Если мы зафиксируем на оси точку P с координатой x и проследим за ее поведением во времени, то увидим, что она совершает колебания с частотой f (периодом T), амплитудой A и фазой $\varphi(P) = -2\pi x/\lambda$.

Если же мы зафиксируем время, придав ему какое-то определенное значение t , и будем рассматривать всю волну, то обнаружим, что она приняла форму синусоиды, образовав в пространстве правильно чередующиеся гребни и впадины. Расстояние между соседними гребнями равно длине волны λ .

Запишем уравнение (6.8) в более компактном и общем виде, введя для этого пространственный вектор $\mathbf{x} = (x, y, z)$ проекции x, y и z на соответствующие координатные оси:

$$U(\mathbf{x}, t) = A \sin(\omega t - k\mathbf{x} + \varphi), \quad (6.9)$$

где $\omega = 2\pi/T = 2\pi f$ — круговая частота; $k = 2\pi/\lambda$ — волновое число.

В оптике вместо тригонометрических функций часто вводят экспоненциальные, что весьма упрощает математическое описание волновых процессов. Воспользуемся для этого формулой Эйлера:

$$e^{j\varphi} = \cos \varphi + j \sin \varphi; \quad e^{-j\varphi} = \cos \varphi - j \sin \varphi. \quad (6.10)$$

Действительная Re и мнимая Im части выражения (6.10) представляют собой тригонометрические функции $\cos \varphi$ и $\sin \varphi$ соответственно. Большинство математических операций легче выполнять с показательными функциями, чем с тригонометрическими. Поэтому вычисления рационально делать следующим образом: введя вместо косинуса или синуса показательную функцию, выполнить с ней все необходимые вычисления и в конце вернуться, если это желательно, к тригонометрическим функциям, взяв соответственно действительную или мнимую часть.

Используя показательную функцию, можно записать выражение (6.9) в виде

$$\begin{aligned} U(\mathbf{x}, t) &= A \exp[j(\omega t - k\mathbf{x} + \varphi)] = \\ &= A \exp j(-k\mathbf{x} + \varphi) \exp(j\omega t) = U(\mathbf{x}) \exp(j\omega t). \end{aligned} \quad (6.11)$$

Комплексную функцию $U(\mathbf{x}) = A \exp j(-k\mathbf{x} + \varphi)$, зависящую только от пространственных координат точки $\mathbf{x} = (x, y, z)$, называют *комплексной амплитудой (фазором)* и часто используют как самостоятельный параметр волны.

Взяв действительную или мнимую часть выражения (6.11), легко получить выражения для косинусоидального или синусоидального представления волны. Например, выражение $U = \text{Re}[U(\mathbf{x}) \times \exp(j\omega t)]$ эквивалентно выражению $U = A \cos(\omega t - k\mathbf{x} + \varphi)$.

Как следует из рис. 6.6, a , волну можно характеризовать вектором A , образующим угол φ с горизонтальной осью.

Этот вектор можно также представить в комплексной плоскости (рис. 6.7).

Начало вектора лежит в начале координат, а конец — в точке, представляющей комплексное число

$$Ae^{j\varphi} = A \cos \varphi + jA \sin \varphi. \quad (6.12)$$

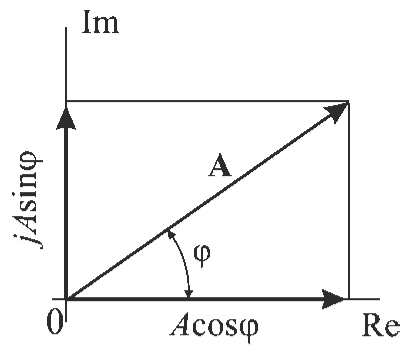


Рис. 6.7. Комплексное представление волн

Преимущества комплексного представления волн выявляются при выполнении математических операций. Например, для сложения двух волн, характеризуемых комплексными числами, их напряженности удобно представить в виде сумм их действительных и мнимых частей: $U_1 = a + jb$, $U_2 = c + jd$. Выражение для суммы волн записывается так:

$$U_1 + U_2 = (a + c) + j(b + d).$$

Для перемножения двух волн представим их в виде модулей с фазовыми множителями: $U_1 = Ae^{j\varphi_1}$, $U_2 = Be^{j\varphi_2}$. Здесь A и B — действительные числа. Тогда произведение волн будет иметь вид

$$U_1 U_2 = AB e^{j(\varphi_1 + \varphi_2)}. \quad (6.13)$$

Правило умножения комплексных чисел имеет ряд интересных следствий. Рассмотрим, например, число $e^{j(\pi/2)}$. Согласно формуле (6.12) имеем

$$e^{j(\pi/2)} = \cos \frac{\pi}{2} + j \sin \frac{\pi}{2} = j. \quad (6.14)$$

Из уравнения (6.13) следует, что умножение комплексного числа на $e^{j(\pi/2)} = j$ увеличивает его фазу на $\pi/2$, т. е. поворачивает вектор, соответствующий этому числу, на 90° против хода часовой стрелки.

Аналогичный результат можно получить, если представить исходное число в виде суммы действительной и мнимой частей.

Пусть $U = a + jb$ — напряженность волны, которая на комплексной плоскости представлена модулем $|U|$ и проекциями a и jb , при этом результат умножения U на j имеет вид $jU = -b + ja$.

Напомним, что модуль комплексного числа определяется выражением

$$|U| = \sqrt{a^2 + b^2}. \quad (6.15)$$

Векторы U и jU имеют одинаковый модуль, однако умножение вектора U на j приводит к его повороту на 90° . Напомним также, что два комплексных числа, отличающихся друг от друга только знаком мнимой части, называются сопряженными. Число, сопряженное с комплексным числом $U = a + jb$, обозначается как $U^* = a - jb$. Произведение UU^* равно квадрату модуля комплексного числа и не зависит от аргумента (т. е. от фазы), так как

$$UU^* = (a + jb)(a - jb) = a^2 + b^2 = |U|^2. \quad (6.16)$$

Формулу (6.16) часто используют для записи соотношения между комплексной амплитудой гармонического колебания (например, световой волны) и интенсивностью электромагнитного поля (последняя выражает яркость объекта или освещенность изображения). Интенсивность поля, т. е. средняя по времени энергия, протекающая в единицу времени через единицу площади, пропорциональна квадрату амплитуды световой волны: $I(x) \sim U^2$. Тогда

$$I(x) \sim Ue^{ip}Ue^{-ip} = U^2. \quad (6.17)$$

6.4. Скалярная теория дифракции

Уравнение Гельмгольца

Зависимость между возмущением $U(\mathbf{x}, t)$ точки P , ее координатами (x, y, z) и временем t , выраженная в дифференциальной форме, называется *волновым уравнением*.

Для его получения следует найти частные производные второго порядка от возмущения $U(\mathbf{x}, t)$ по времени t и координате \mathbf{x} и сравнить их между собой:

$$\left[\frac{\partial^2}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right] U(\mathbf{x}, t) = 0. \quad (6.18)$$

Применив обозначение $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ оператора Лапласа, запишем скалярное волновое уравнение в виде

$$\Delta U(\mathbf{x}, t) - \frac{1}{c^2} \frac{\partial^2 U(\mathbf{x}, t)}{\partial t^2} = 0. \quad (6.19)$$

В большинстве физико-оптических явлений частота колебаний остается неизменной ($\omega = \text{const}$), поэтому выражение $\exp(j\omega t)$ при расчетах обычно опускают. В таком случае волновое уравнение для комплексной амплитуды предельно упрощается:

$$(\Delta + k^2)U(\mathbf{x}) = 0. \quad (6.20)$$

Волновое уравнение в виде выражения (6.20) называется *уравнением Гельмгольца*. Этому уравнению должна подчиняться комплексная амплитуда любого монохроматического возмущения, распространяющегося в свободном пространстве.

Теорема Грина

Учтем, что в соответствии с принципом Гюйгенса — Френеля волновой фронт распространяется от источника излучения во все стороны и вторичные источники образуются во всем объеме, окружающем рассматриваемую точку (см. рис. 6.5, а).

Таким образом, возмущение среды в точке \mathbf{x}_i будет определяться суммарным действием колебаний, исходящих из каждой точки объема, с учетом их амплитуд и фаз. Однако даже если точно известны распределения амплитуд и фаз вторичных источников, интегрирование по объему даже простейшего вида представляет собой чрезвычайно трудоемкую задачу. Поэтому необходимо эквивалентно заменить действие объема, в котором находится рассматриваемая точка, действием поверхности, окружающей этот объем. Это позволит перейти от интегрирования по объему к интегрированию по по-

верхности. Такой переход возможен с помощью теоремы Грина, которая формулируется следующим образом.

Пусть $U(\mathbf{x})$ и $G(\mathbf{x})$ — две произвольные комплексные функции координат, а S — замкнутая поверхность, ограничивающая объем V . Если функции U и G , их первые и вторые производные однозначны и непрерывны внутри объема, ограниченного поверхностью S , и на самой этой поверхности, то

$$\iiint_V (G\Delta U - U\Delta G)dV = \iint_S \left(G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n} \right) dS, \quad (6.21)$$

где $\partial/\partial n$ — частная производная в каждой точке поверхности S , взятая по направлению внешней нормали к этой поверхности.

Теорема Грина является основным звеном теории дифракции. Ее суть состоит в следующем.

Полагаем, что функция $U(\mathbf{x})$, выражающая возмущение светового поля в каждой точке поверхности S , известна. Если для той же точки мы сможем подобрать вспомогательную функцию $G(\mathbf{x})$, обладающую перечисленными выше свойствами, то по теореме Грина можно провести эквивалентную замену интеграла по *объему* интегралом по *поверхности*.

Замена объемного интеграла поверхностным дает огромную экономию вычислительных ресурсов. Поэтому включение в расчеты некоей вспомогательной функции, не имеющей прямого физического смысла, представляется вполне оправданным.

Выбор функции Грина G и замкнутой поверхности S существенно влияет на конечный результат. В различных курсах по физической оптике приводятся конечные формулы, полученные при различных значениях этой функции.

Для выяснения сходства и различий этих формул проведем их дальнейший анализ.

Интегральная теорема Гельмгольца — Кирхгофа

Рассмотрим прежде всего выбор функции Грина, сделанный Кирхгофом, и интегральную теорему, следующую из этого вывода.

По существу, нам необходимо выразить оптическое возмущение в точке \mathbf{x} , через значение этого возмущения на поверхности S

(рис. 6.8, *a*). Следуя Кирхгофу, выберем в качестве функции Грина $G(\mathbf{x})$ сферическую волну единичной амплитуды, распространяющуюся из точки \mathbf{x}_i . Такая функция G (так называемая функция Грина свободного пространства) для произвольной точки \mathbf{x}_o поверхности S имеет вид

$$G(\mathbf{x}_o) = \frac{\exp(jkr_i)}{r_i}, \quad (6.22)$$

где $r_i = |\mathbf{x}_i - \mathbf{x}_o|$ — длина вектора, направленного из точки \mathbf{x}_i в точку \mathbf{x}_o .

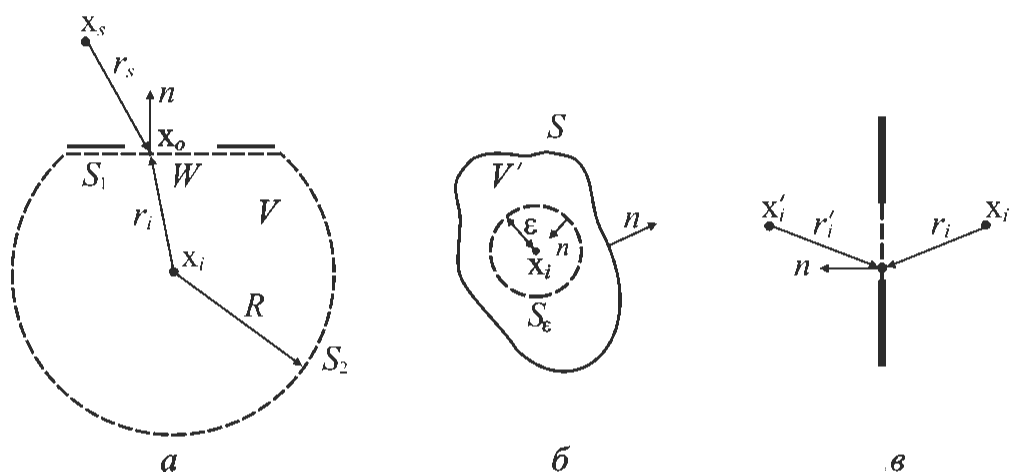


Рис. 6.8. Дифракция на плоском экране

a — трактовка Кирхгофа; *б* — поверхность интегрирования; *в* — трактовка Зоммерфельда

Для использования теоремы Грина необходимо, чтобы функция $G(\mathbf{x})$, ее первая и вторая производные были непрерывными в объеме V , ограниченном поверхностью S . Поскольку точка \mathbf{x}_i является точкой разрыва (при \mathbf{x}_i функция $G(\mathbf{x})$ имеет особенность), ее необходимо исключить из области интегрирования. Для этого окружим \mathbf{x}_i небольшой сферической поверхностью S_ϵ радиусом ϵ (рис. 6.8, *б*). Затем применим теорему Грина, причем интегрирование будем вести по объему V' , заключенному между поверхностями S и S_ϵ .

Поверхностью интегрирования будет в данном случае поверхность $S' = S + S_\epsilon$. Внутри объема V' возмущение $G(\mathbf{x})$, представ-

ляющее собой расходящуюся сферическую волну, удовлетворяет уравнению Гельмгольца

$$(\Delta + k^2)G(\mathbf{x}) = 0. \quad (6.23)$$

Используя оба уравнения Гельмгольца (6.20) и (6.23) для преобразования левой части формулы Грина, получаем

$$\iiint_{V'} (G\Delta U - U\Delta G)dV = -\iiint_{V'} (Gk^2U - Uk^2G)dV = 0. \quad (6.24)$$

Тогда теорема (6.21) преобразуется к виду

$$\iint_{S'} \left(G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n} \right) dS = 0,$$

или

$$\iint_S \left(G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n} \right) dS = -\iint_{S_e} \left(G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n} \right) dS. \quad (6.25)$$

В формуле (6.25) функция $U(\mathbf{x})$ характеризует возмущение поля на поверхности S' в некоторой точке \mathbf{x}_o , а функция $G(\mathbf{x})$ — воздействие на эту точку сферической волны единичной амплитуды, расходящейся из точки \mathbf{x}_i , т. е. функции Грина. Поэтому для этой точки

$$G(\mathbf{x}_o) = \frac{\exp(jkr_i)}{r_i}. \quad (6.26)$$

Производная по нормали функции Грина приобретает вид

$$\begin{aligned} \frac{\partial G(\mathbf{x}_o)}{\partial n} &= \cos(n, r_i) \frac{\partial G(\mathbf{x}_o)}{\partial r} = \\ &= \cos(n, r_i) \left[(-1) \frac{1}{r_i^2} \exp(jkr_i) + \frac{1}{r_i} jk \exp(jkr_i) \right] = \\ &= \cos(n, r_i) \frac{\exp(jkr_i)}{r_i} \left(jk - \frac{1}{r_i} \right), \end{aligned} \quad (6.27)$$

где $\cos(n, r_i)$ — косинус угла между направлением внешней нормали n и вектором r_i , соединяющим точки \mathbf{x}_i и \mathbf{x}_o .

Для частного случая, когда точка \mathbf{x}_o лежит на поверхности S_ε , имеем $\cos(n, r_i) = -1$ и выражения (6.26) и (6.27) принимают вид

$$G(\mathbf{x}_o) = \frac{\exp(jk\varepsilon)}{\varepsilon}; \quad \frac{\partial G(\mathbf{x}_o)}{\partial n} = \frac{\exp(jk\varepsilon)}{\varepsilon} \left(\frac{1}{\varepsilon} - jk \right). \quad (6.28)$$

Если $\varepsilon \rightarrow 0$, то в силу непрерывности функции U (и ее производных) в точке \mathbf{x}_i можно записать

$$\begin{aligned} \iint_{S_\varepsilon} \left(G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n} \right) dS = \\ = 4\pi\varepsilon^2 \left[\frac{\exp(jk\varepsilon)}{\varepsilon} \frac{\partial U(\mathbf{x}_i)}{\partial n} - U(\mathbf{x}_i) \frac{\exp(jk\varepsilon)}{\varepsilon} \left(\frac{1}{\varepsilon} - jk \right) \right]_{\varepsilon \rightarrow 0} = \\ = -4\pi U(\mathbf{x}_i). \end{aligned} \quad (6.29)$$

При этом полагалось, что интеграл $\iint_{S_\varepsilon} \left(G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n} \right) dS$ — это объем, опирающийся на площадь S_ε . При $\varepsilon \rightarrow 0$ площадь $S_\varepsilon = 4\pi\varepsilon^2$, значение $U \rightarrow U(\mathbf{x}_i)$, а объем равен произведению значений площади основания и функции в точке \mathbf{x}_i . Подстановка полученного результата в (6.25) дает

$$U(\mathbf{x}_i) = \frac{1}{4\pi} \iint_S \left\{ \frac{\partial U}{\partial n} \left(\frac{\exp(jkr_i)}{r_i} \right) - U \frac{\partial}{\partial n} \left(\frac{\exp(jkr_i)}{r_i} \right) \right\} dS. \quad (6.30)$$

Соотношение (6.30) представляет собой математическую запись интегральной теоремы Гельмгольца — Кирхгофа. Она позволяет выразить параметры поля в любой точке \mathbf{x}_i через граничные параметры волны на любой замкнутой поверхности, окружающей эту точку.

Применение интегральной теоремы

Рассмотрим задачу о дифракции на отверстии в бесконечном непрозрачном экране. Предполагается, что возмущение от точечного источника с координатой \mathbf{x}_s падает сверху на непрозрачный

экран с отверстием, как показано на рис. 6.8, *a*. Необходимо рассчитать параметры поля в точке \mathbf{x}_i за отверстием.

Воспользуемся интегральной теоремой Гельмгольца — Кирхгофа, выбрав соответствующим образом поверхность интегрирования. Следуя Кирхгофу, возьмем замкнутую поверхность S таким образом, чтобы она состояла из двух частей (см. рис. 6.8, *a*). Пусть плоская поверхность S_1 , лежащая сразу за дифракционным экраном, замыкается большим сферическим колпаком S_2 радиусом R с центром в рассматриваемой точке \mathbf{x}_i . Полная замкнутая поверхность S образована поверхностями S_1 и S_2 , поэтому интеграл (6.21) имеет вид

$$U(\mathbf{x}_i) = \frac{1}{4\pi} \iint_{S_1+S_2} \left(G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n} \right) dS, \quad (6.31)$$

где, как и прежде, $G = \frac{\exp(jkr_i)}{r_i}$.

При увеличении R поверхность S_2 принимает форму полусферической оболочки. Можно показать, что интеграл по S_2 не будет давать вклада в общий интеграл.

Граничные условия Кирхгофа

Устранив интегрирование по поверхности S_2 , можно выразить возмущение в точке \mathbf{x}_i через возмущение и его производную по нормали, взятые на бесконечной плоскости S_1 , расположенной непосредственно за экраном:

$$U(\mathbf{x}_i) = \frac{1}{4\pi} \iint_{S_1} \left(G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n} \right) dS. \quad (6.32)$$

Экран непрозрачен везде, кроме открытого отверстия, которое мы обозначим W . Кирхгоф принял следующие предположения.

1. На отверстии W распределение поля U и его производная $\partial U / \partial n$ имеют точно такие же значения, какие они имели бы в отсутствие экрана.

2. На той части поверхности S_1 , которая лежит в области геометрической тени экрана, распределение поля и его производная тождественно равны нулю.

Эти граничные условия Кирхгофа лежат в основе скалярной теории дифракции. Первое условие позволяет определить возмущение, падающее на отверстие, пренебрегая наличием экрана, а второе дает возможность пренебречь интегрированием по всей поверхности, за исключением самого отверстия.

Формула дифракции Френеля — Кирхгофа

С учетом условий Кирхгофа выражение (6.30) запишем в виде

$$U(\mathbf{x}_i) = \frac{1}{4\pi} \iint_W \left(G \frac{\partial U}{\partial n} - U \frac{\partial G}{\partial n} \right) dS. \quad (6.33)$$

Обычно полагают, что расстояние r_i от отверстия до точки наблюдения во много раз больше длины волны. При этом выражение (6.27) можно преобразовать к следующему виду:

$$\begin{aligned} \frac{\partial G(\mathbf{x}_o)}{\partial n} &= \cos(n, r_i) \frac{\partial G(\mathbf{x}_o)}{\partial r} = \\ &= \cos(n, r_i) \frac{\exp(jkr_i)}{r_i} \left(jk - \frac{1}{r_i} \right) \approx jk \cos(n, r_i) \frac{\exp(jkr_i)}{r_i}. \end{aligned} \quad (6.34)$$

Подставляя приближенное выражение (6.34) и выражение (6.26) для G в (6.30), находим, что

$$U(\mathbf{x}_i) = \frac{1}{4\pi} \iint_W \frac{\exp(jkr_i)}{r_i} \left[\frac{\partial U}{\partial n} - jkU \cos(n, r_i) \right] dS. \quad (6.35)$$

Предположим, что на отверстие падает сферическая волна

$$U(\mathbf{x}_o) = A_s \frac{\exp(jkr_s)}{r_s}, \quad (6.36)$$

исходящая из одиночного источника в точке \mathbf{x}_s , расположенного на расстоянии r_s от \mathbf{x}_o (см. рис. 6.8, а). Если расстояние r_s во много раз больше длины волны, то (6.35) сразу можно упростить:

$$U(\mathbf{x}_i) = \frac{A_s}{j\lambda} \iint_W \frac{\exp[jk(r_s + r_i)]}{r_s r_i} \left[\frac{\cos(n, r_i) - \cos(n, r_s)}{2} \right] dS. \quad (6.37)$$

Этот результат, справедливый для случая, когда отверстие освещается одиночным точечным источником, известен как *формула дифракции Френеля — Кирхгофа*.

Рассмотрим одну полезную интерпретацию формулы (6.37). Перепишем ее следующим образом:

$$U(\mathbf{x}_i) = \iint U'(\mathbf{x}_o) \frac{\exp(jkr_i)}{r_i} dS; \quad (6.38)$$

$$U'(\mathbf{x}_o) = \frac{1}{j\lambda} \left[\frac{A_s \exp(jkr_s)}{r_s} \right] \left[\frac{\cos(n, r_i) - \cos(n, r_s)}{2} \right]. \quad (6.39)$$

Исходя из выражения (6.39), можем считать, что поле в точке \mathbf{x}_i создается бесконечным множеством вторичных точечных источников, расположенных в пределах самого отверстия. Амплитуда $U'(\mathbf{x}_o)$ вторичного источника, расположенного в точке \mathbf{x}_o , пропорциональна амплитуде $[A_s \exp(jkr_s)]/r_s$ волны, исходящей из точки \mathbf{x}_s .

Однако, во-первых, амплитуда вторичного источника отличается от амплитуды падающей волн множителем λ^{-1} , т. е. амплитуда колебания вторичного источника обратно пропорциональна длине волны. Во-вторых, эта амплитуда уменьшается за счет *коэффициента наклона* $1/2[\cos(n, r_i) - \cos(n, r_s)]$, который никогда не превышает единицы и всегда положителен. В результате каждому вторичному источнику соответствует анизотропная «картина направленности». В-третьих, фаза излучения вторичного источника в точке $U(\mathbf{x}_o)$ отличается от фазы падающей волны на 90° , что следует из наличия в выражении (6.39) множителя $1/j$.

Эти любопытные свойства вторичных источников были, по существу, предсказаны Френелем. Математический вывод Кирхгофа показал, что эти свойства объясняются волновой природой света, при этом были уточнены некоторые интуитивные предположения Френеля.

Формула дифракции Рэлея — Зоммерфельда

Отметим, что приведенные выше краевые значения Кирхгофа математически противоречивы. Из теории потенциала известно, что если на некотором участке замкнутой поверхности функция и

ее производные по нормали равны нулю, то они должны быть равны нулю и на всей замкнутой поверхности, которой является поверхность S . В этом случае выражение (6.39) вступает в противоречие с реальной физической ситуацией.

Для того чтобы выйти из этого положения, Зоммерфельд предложил в формуле (6.33) так изменить функцию Грина G , чтобы эта функция обращалась в нуль на всей поверхности W , но вывод приведенного выше выражения все же оставался справедливым.

Для этого новая функция Грина G_- должна подчиняться следующим условиям:

- 1) удовлетворять волновому уравнению;
- 2) быть равной нулю на поверхности W ;
- 3) совпадать с функцией G при $r_i = 0$, т. е. в точке x_i .

В этом случае отпадает необходимость наложения граничных условий одновременно на U и $\partial U / \partial n$, тем самым устраняются противоречия теории Кирхгофа. Введение этой вспомогательной функции ничего не меняет по существу, но облегчает математическое решение. При выполнении названных условий в выражении (6.33) устраняется математическое противоречие, а само оно значительно упрощается, так из него исключается первый член подынтегрального выражения.

Согласно Зоммерфельду, предположим, что функция G_- создается не только точечным источником, помещенным в точку x_i , но и вторым точечным источником в точке x'_i .

Точка x'_i представляет собой зеркальное изображение точки x_i и лежит по другую сторону экрана (см. рис. 6.8, в). Пусть оба источника имеют одинаковую длину волны, а излучения этих источников сдвинуты по фазе на 180° . Функция Грина в этом случае имеет вид

$$G_-(\mathbf{x}) = \frac{\exp(jkr_i)}{r_i} - \frac{\exp(jkr'_j)}{r'_j}, \quad (6.40)$$

где r'_j — расстояние между точками x'_i и x_o .

Функция (6.40) удовлетворяет всем условиям функции Грина. Она обращается в нуль на поверхности плоского экрана и переходит в функцию G в точке x_i . В этой точке функция $G = G_- = \infty$, так же как и функция G . Соответствующая производная от функции G_-

$$\frac{\partial G_-(\mathbf{x}_o)}{\partial n} = \cos(n, r_i) \left(jk - \frac{1}{r_i} \right) \frac{\exp(jkr_i)}{r_i} - \cos(n, r'_i) \left(jk - \frac{1}{r'_i} \right) \frac{\exp(jkr'_i)}{r'_i}.$$

Для точки \mathbf{x}_o на поверхности S_1 имеем $r_i = r'_i$; $\cos(n, r_i) = -\cos(n, r'_i)$. Следовательно, на поверхности S_1 функция $G_-(\mathbf{x}_o) = 0$, тогда

$$\frac{\partial G_-(\mathbf{x}_o)}{\partial n} = 2 \cos(n, r_i) \left(jk - \frac{1}{r_i} \right) \frac{\exp(jkr_i)}{r_i}. \quad (6.41)$$

Таким образом, на всей поверхности S_1 функция Грина G_- в виде выражения (6.40) обращается в нуль, что устраняет противоречия теории Кирхгофа. Подставив функцию (6.41) в (6.33), получим

$$U(\mathbf{x}_i) = \frac{1}{\lambda} \iint_W U(\mathbf{x}_o) \left(\frac{1}{kr_i} - j \right) \frac{\exp(jkr_i)}{r_i} \cos(n, r_i) dS. \quad (6.42)$$

Выражение (6.42) — это *дифракционное уравнение Рэлея — Зоммерфельда*. Для упрощения расчетов обычно вводят некоторые ограничения на соотношения входящих в уравнение (6.42) параметров.

Приближение Кирхгофа

Приближение Кирхгофа применимо в тех случаях, когда расстояние r_i от отверстия до точки наблюдения во много раз больше длины волны λ . Это условие уже использовалось нами при выводе формулы дифракции Френеля — Кирхгофа (6.37). Для контактной фотолитографии это условие означает, что зазор между фотошаблоном и пластиной z во много раз превышает длину волны экспонирования, т. е. из условия $z \gg \lambda$ следует, что $r_i \gg \lambda$. В этом случае $k \gg 1/r_i$ и выражение (6.42) можно записать как

$$U(\mathbf{x}_i) = \frac{1}{j\lambda} \iint_W U(\mathbf{x}_o) \frac{\exp(jkr_i)}{r_i} \cos(n, r_i) dS. \quad (6.43)$$

Выражение (6.43) является *интегралом Кирхгофа*. Приведем его к виду, сравнимому с результатами, полученными ранее (см. формулу (6.37)). Пусть, как и прежде, отверстие освещается сфе-

рической волной из точечного источника, расположенного в точке \mathbf{x}_s (см. рис. 6.8, *a*). Тогда

$$U(\mathbf{x}_o) = A_s \frac{\exp(jkr_s)}{r_s};$$

$$U(\mathbf{x}_i) = \frac{A_s}{j\lambda} \iint_W \frac{\exp[jk(r_s + r_i)]}{r_s r_i} \cos(n, r_i) dS. \quad (6.44)$$

Уравнение (6.44) отличается от аналогичного уравнения (6.37) Френеля — Кирхгофа только значением коэффициента наклона.

Отметим, что формула (6.43) представляет собой выражение принципа Гюйгенса — Френеля в виде интеграла суперпозиции, который можно записать следующим образом:

$$U(\mathbf{x}_i) = \iint_W h(\mathbf{x}_o, \mathbf{x}_i) U(\mathbf{x}_o) dS, \quad (6.45)$$

где весовая функция $h(\mathbf{x}_o, \mathbf{x}_i)$ определяется выражением

$$h(\mathbf{x}_o, \mathbf{x}_i) = \frac{1}{j\lambda} \frac{\exp(jkr_i)}{r_i} \cos(n, r_i). \quad (6.46)$$

В выражении (6.46) член уравнения $\exp(jkr_i)/r_i$ описывает сферическую волну, расходящуюся из точки $(0, 0, 0)$, а коэффициент наклона $\cos(n, r_i) = z/r_i$ (см. рис. 6.8, *a*). Физический смысл параметров j и λ рассмотрен при анализе уравнения (6.37).

Таким образом, функция $h(\mathbf{x}_o, \mathbf{x}_i)$ представляет собой сферическую волну, распространяющуюся из точки \mathbf{x}_o и умноженную на коэффициент наклона. При этом каждая точка \mathbf{x}_o отверстия W служит источником таких волн, которые суммируются в точке \mathbf{x}_i .

Приближение Френеля

Дальнейшие упрощения можно получить, принимая некоторые приближения для величины r_i . Следуя Френелю, будем полагать, что расстояние z между экраном (объектом) и плоскостью наблюдения (изображением) значительно превышает максимальный линейный размер окна W (рис. 6.9).

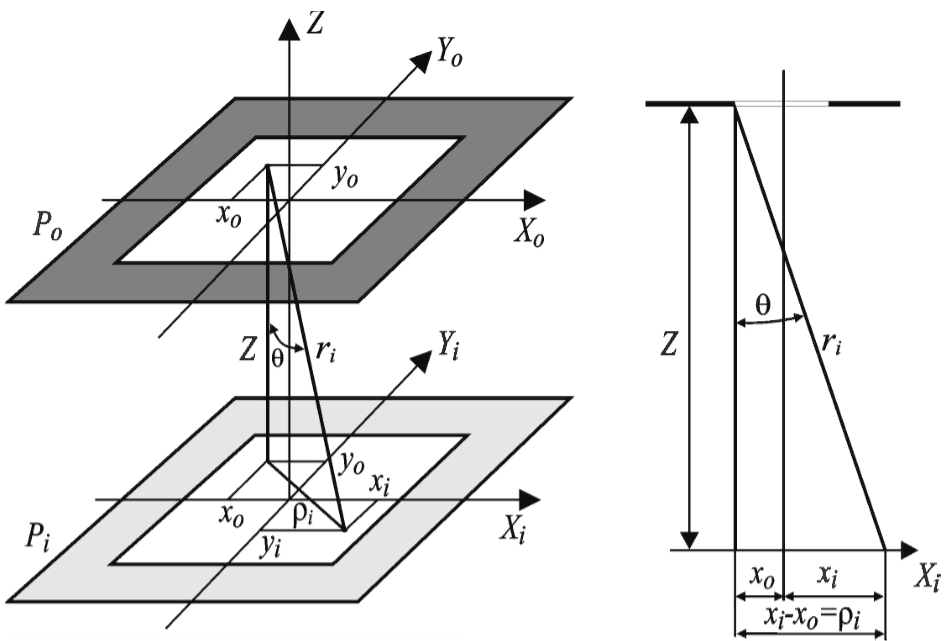


Рис. 6.9. Формирование изображения в приближении Френеля

Кроме того, будем предполагать, что в плоскости наблюдения рассматривается только конечная область вблизи оси Z и что расстояние z намного больше максимального размера этой области, т. е.

$$|\mathbf{x}_i - \mathbf{x}_o| \ll z.$$

Выражение (6.46) для функции $h(\mathbf{x}_o, \mathbf{x}_i)$ принимает вид

$$h(\mathbf{x}_o, \mathbf{x}_i) = \frac{1}{j\lambda} \left(\frac{z}{r_i} \right) \frac{\exp(jkr_i)}{r_i},$$

т. е.

$$h(\mathbf{x}_o, \mathbf{x}_i) = A(\mathbf{x}_o, \mathbf{x}_i) \exp(jkr_i),$$

где

$$A(\mathbf{x}_o, \mathbf{x}_i) = \frac{1}{j\lambda} \frac{z}{r_i^2}.$$

Точная формула для расстояния r_i (см. рис. 6.9) будет выглядеть так:

$$r_i = \sqrt{z^2 + (x_i - x_o)^2 + (y_i - y_o)^2} = z \sqrt{1 + \left(\frac{x_i - x_o}{z} \right)^2 + \left(\frac{y_i - y_o}{z} \right)^2}. \quad (6.47)$$

Разложение Тейлора для квадратного корня дает следующую аппроксимацию его первыми двумя членами разложения:

$$\begin{aligned} \sqrt{1+b} \approx 1 + \frac{1}{2}b &\Rightarrow z\sqrt{1 + \left(\frac{x_i - x_o}{z}\right)^2 + \left(\frac{y_i - y_o}{z}\right)^2} \approx \\ &\approx z \left[1 + \frac{1}{2} \left(\frac{x_i - x_o}{z}\right)^2 + \frac{1}{2} \left(\frac{y_i - y_o}{z}\right)^2 \right]. \end{aligned} \quad (6.48)$$

С учетом этого приближения, которое называют *приближением Френеля*, в выражении для функции h можно сделать следующие упрощения:

- для амплитудного члена выражения провести аппроксимацию первого порядка:

$$r_i \approx z \Rightarrow A(\mathbf{x}_o, \mathbf{x}_i) \approx \frac{1}{j\lambda z};$$

- для фазового члена выражения провести аппроксимацию второго порядка:

$$\begin{aligned} r_i \approx z + \frac{(x_i - x_o)^2 + (y_i - y_o)^2}{2z} &\Rightarrow \\ \Rightarrow \exp(jkr_i) \approx \exp(jkz) \exp\left\{ j \frac{k}{2z} [(x_i - x_o)^2 + (y_i - y_o)^2] \right\}. \end{aligned}$$

В результате весовая функция в приближении Френеля будет иметь вид

$$h(x_i, y_i) = \frac{\exp(jkz)}{j\lambda z} \exp\left\{ j \frac{k}{2z} [(x_i - x_o)^2 + (y_i - y_o)^2] \right\}. \quad (6.49)$$

Таким образом, когда расстояние z достаточно велико по сравнению с размерами объекта и изображения, можно использовать приближения Френеля. При этом сферическая волна вторичного источника заменяется параболической, а коэффициент наклона $\cos(n, r_i) \approx 1$.

Вернемся теперь к выражению (6.43) и перепишем его как интеграл суперпозиции с бесконечными пределами. При этом поло-

жим, что в соответствии с граничными условиями Кирхгофа функция $U(x_o, y_o)$ за пределами отверстия W равна нулю. В результате выражение (6.43) в приближении Френеля примет вид

$$U^*(x_i, y_i) = \frac{\exp(jkz)}{j\lambda z} \times \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} U(x_o, y_o) \exp\left\{j \frac{k}{2z} [(x_i - x_o)^2 + (y_i - y_o)^2]\right\} dx_o dy_o. \quad (6.50)$$

Дифракция при контактной фотолитографии

В большинстве случаев для оценки профиля распределения интенсивности при контактной фотолитографии используется дифракционное уравнение в приближении Френеля, так как оно легко преобразуется к одномерному виду. Это существенно снижает трудоемкость расчетов.

Предположим, что прозрачное квадратное окно на фотошаблоне со стороны W равномерно освещено нормально падающей плоской монохроматической волной единичной амплитуды. Для данного случая формулу (6.48) можно переписать в виде

$$U(x_i, y_i) = \frac{\exp(jkz)}{j\lambda z} \int_{-W/2}^{+W/2} \int_{-W/2}^{+W/2} \exp\left\{j \frac{k}{2z} [(x_i - x_o)^2 + (y_i - y_o)^2]\right\} dx_o dy_o. \quad (6.51)$$

Выражение (6.51) можно представить в виде произведения двух интегралов:

$$U(x_i, y_i) = \frac{\exp(jkz)}{j\lambda z} A(x_i) B(y_i), \quad (6.52)$$

где

$$\begin{aligned} A(x_i) &= \int_{-W/2}^{+W/2} \exp\left[j \frac{k}{2z} (x_i - x_o)^2\right] dx_o; \\ B(y_i) &= \int_{-W/2}^{+W/2} \exp\left[j \frac{k}{2z} (y_i - y_o)^2\right] dy_o. \end{aligned} \quad (6.53)$$

После замены переменных

$$\xi = \sqrt{\frac{k}{\pi z}}(x_o - x_i); \quad \eta = \sqrt{\frac{k}{\pi z}}(y_o - y_i)$$

интегралы (6.53) существенно упрощаются:

$$A(x_i) = \sqrt{\frac{\pi z}{k}} \int_{\xi_1}^{\xi_2} \exp\left(j \frac{\pi}{2} \xi^2\right) d\xi; \quad (6.54)$$

$$B(y_i) = \sqrt{\frac{\pi z}{k}} \int_{\eta_1}^{\eta_2} \exp\left(j \frac{\pi}{2} \eta^2\right) d\eta.$$

Пределы интегрирования определяются соотношениями

$$\xi_1 = -\sqrt{\frac{k}{\pi z}}\left(x_i + \frac{W}{2}\right); \quad \xi_2 = \sqrt{\frac{k}{\pi z}}\left(-x_i + \frac{W}{2}\right); \quad (6.55)$$

$$\eta_1 = -\sqrt{\frac{k}{\pi z}}\left(y_i + \frac{W}{2}\right); \quad \eta_2 = \sqrt{\frac{k}{\pi z}}\left(-y_i + \frac{W}{2}\right).$$

Традиционно интегралы $A(x_i)$ и $B(y_i)$ выражают через интегралы Френеля:

$$C(\alpha) = \int_0^{\alpha} \cos \frac{\pi t^2}{2} dt; \quad S(\alpha) = \int_0^{\alpha} \sin \frac{\pi t^2}{2} dt. \quad (6.56)$$

Учитывая, что

$$\int_{\xi_1}^{\xi_2} \exp\left(j \frac{\pi}{2} \xi^2\right) d\xi = \int_0^{\xi_2} \exp\left(j \frac{\pi}{2} \xi^2\right) d\xi - \int_0^{\xi_1} \exp\left(j \frac{\pi}{2} \xi^2\right) d\xi,$$

определяем:

$$A(x_i) = \sqrt{\frac{\pi z}{k}} \left\{ [C(\xi_2) - C(\xi_1)] + j[S(\xi_2) - S(\xi_1)] \right\}; \quad (6.57)$$

$$B(y_i) = \sqrt{\frac{\pi z}{k}} \left\{ [C(\eta_2) - C(\eta_1)] + j[S(\eta_2) - S(\eta_1)] \right\}.$$

Подставив (6.57) в (6.52), можно получить распределение комплексной амплитуды:

$$U(x_i, y_i) = \frac{\exp(jkz)}{2j} \{ [C(\xi_2) - C(\xi_1)] + j[S(\xi_2) - S(\xi_1)] \} \times \\ \times \{ [C(\eta_2) - C(\eta_1)] + j[S(\eta_2) - S(\eta_1)] \}. \quad (6.58)$$

Соответствующее распределение интенсивности излучения имеет вид

$$I(x_i, y_i) = \frac{1}{4} \{ [C(\xi_2) - C(\xi_1)]^2 + [S(\xi_2) - S(\xi_1)]^2 \} \times \\ \times \{ [C(\eta_2) - C(\eta_1)]^2 + [S(\eta_2) - S(\eta_1)]^2 \}. \quad (6.59)$$

Уравнение (6.59) применяют для расчета двумерных дифракционных распределений, например на углах топологических элементов. Его также можно развить для случая, когда маскирующее покрытие фотошаблона не является полностью непрозрачным.

Выражение (6.59) легко преобразуется для одномерного случая, когда окно на шаблоне представляет собой узкую длинную щель, т. е. когда $W_y \rightarrow \pm\infty$. Учитывая, что $C(\infty) = S(\infty) = 1/2$ и $C(-\infty) = S(-\infty) = -1/2$, получаем для сечения щели:

$$I(x_i) = \frac{1}{2} \{ [C(\xi_2) - C(\xi_1)]^2 + [S(\xi_2) - S(\xi_1)]^2 \}. \quad (6.60)$$

Уравнение (6.60) при соответствующем выборе пределов интегрирования пригодно для расчета дифракции на наборе полос, т. е. на дифракционной решетке.

Расчет дифракционного распределения для элемента в виде одномерного длинного окна наиболее прост и может быть выполнен даже вручную с использованием таблиц функций Френеля. В этом случае расчетная схема (см. рис. 6.9) упрощается и приводится к виду, показанному на рис. 6.10. Учитывая симметричность получаемой кривой распределения относительно центра окна, расчет можно проводить лишь для половины кривой.

Положительную область изображения — от нуля вправо — следует разделить на n расчетных интервалов ($n \geq 20$). Нулевая точка и n точек в начале интервалов будут иметь координаты $x_{ik} = k(W/n)$, где $k = 0 \dots n$.

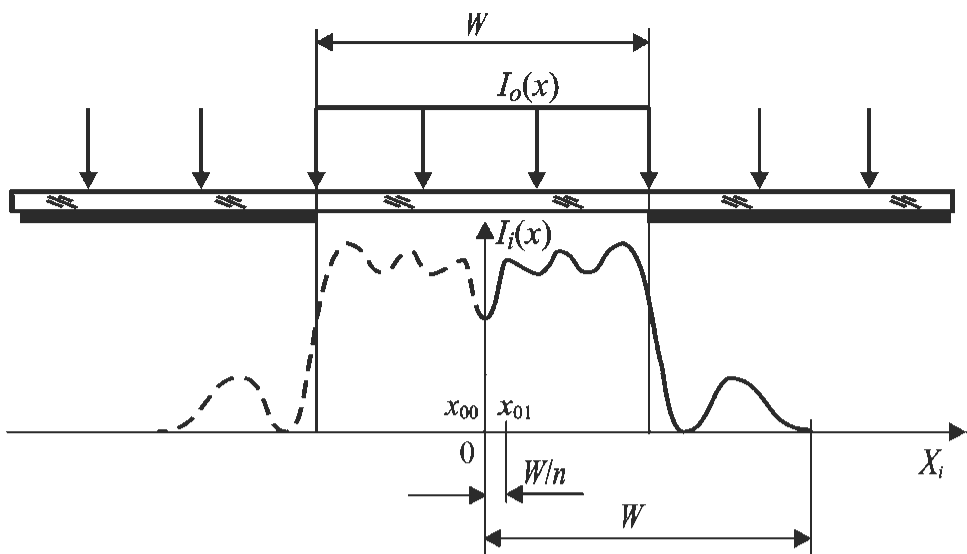


Рис. 6.10. Расчетная схема дифракции при контактной фотолитографии

Для каждой из этих точек вычисляют значения пределов интегрирования ξ_1, ξ_2 по формуле (6.55). Далее по известным таблицам находят интегралы Френеля, соответствующие вычисленным значениям $C(\xi_1), C(\xi_2), S(\xi_1), S(\xi_2)$, и рассчитывают значения интенсивности излучения в i -й точке по формуле (6.60).

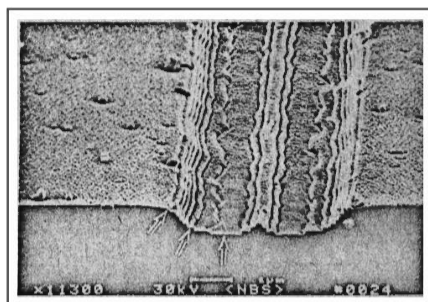
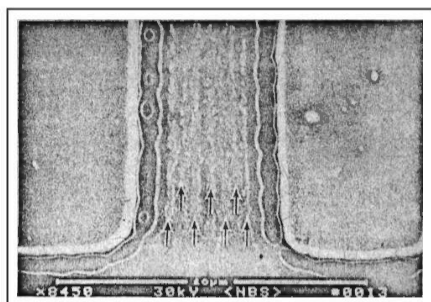
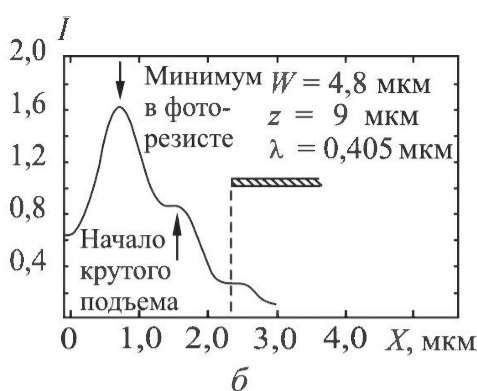
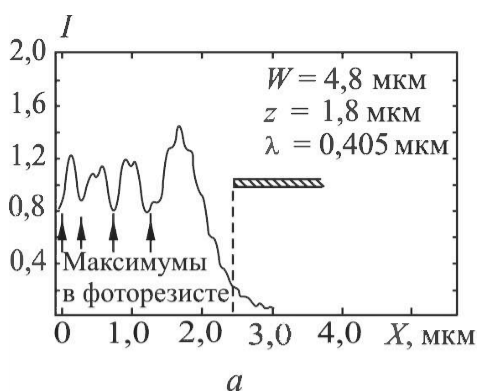
Аналогично находят все $n + 1$ точки изображения. В табл. 6.1 приведен фрагмент такого расчета. По полученным значениям строят график распределения интенсивности. Примеры таких графиков показаны на рис. 6.11, а, б.

Таблица 6.1

Расчет распределения интенсивности

k	x_{ik}	ξ_1	ξ_2	$C(\xi_1)$	$C(\xi_2)$	$S(\xi_1)$	$S(\xi_2)$	$I(x_i)$
0	0	-5,60	5,60	-0,449	0,445	-0,475	0,484	0,860
1	2,5	6,16	5,04	0,521	0,558	0,547	0,525	1,157
...
n	50,0	16,80	-5,60	-0,517	-0,449	-0,491	-0,475	0,003

Отметим, что введенные при выводе формул Френеля ограничения на соотношения размера окна, микрозазора и длины волны в реальных процессах микролитографии часто нарушаются. Тем не менее получаемые расчетные распределения весьма близки к экспериментальным результатам.



в

z

Рис. 6.11. Экспериментальная оценка дифракционного уравнения: *а, б* — расчетные кривые; *в, z* — микропрофили в фоторезисте

В качестве примера на рис. 6.11, *в, z* представлены два микропрофиля в толстом фоторезисте, полученные специалистами Национального бюро стандартов США при экспонировании длинного окна шириной 4,8 мкм. Использовалось излучение с длиной волны 0,405 мкм, микрозазоры составляли 1,8 и 9 мкм соответственно.

Расчетные профили интенсивности в обоих случаях точно воспроизведены в фоторезисте. Так, при зазоре 1,8 мкм расчетный профиль имеет семь локальных минимумов (см. рис. 6.11, *а*), которые легко распознаются на микрофотографии (см. рис. 6.11, *в*). Не менее явно видно соответствие между расчетной кривой при микрозазоре 9 мкм (см. рис. 6.11, *б*) и соответствующим профилем в фоторезисте (см. рис. 6.11, *z*).

Эти результаты показывают, что даже при нарушении некоторых ограничений Френеля (например, микрозазор 1,8 мкм незначительно превышает длину волны излучения, он также сопоставим

с размером окна) полученные зависимости дают вполне приемлемую сходимость с экспериментом.

Контрольные вопросы и задания

1. Получите волновое уравнение, дифференцируя уравнение волны в синусоидальной, косинусоидальной и комплексной формах.
2. Перейдите от волнового уравнения в общем виде для $u(x, t)$ к уравнению для фазора $U(x)$.
3. Выведите формулу дифракции Френеля — Кирхгофа.
4. Докажите, что, переходя к приближению Френеля, в фазовом члене уравнения нельзя ограничиться аппроксимацией первого порядка.
5. Выполните переход от формулы Френеля для амплитуды к формуле для интенсивности излучения.

7. ПРОЕКЦИОННОЕ ФОРМИРОВАНИЕ МИКРОИЗОБРАЖЕНИЙ

Изучив материал этой главы, студент должен иметь представление об особенностях оценки качества проекционных микроизображений применительно к использованию в микролитографии.

Студент должен знать:

- основные положения фурье-преобразований в оптике;
- методы оценки качества формируемых микроизображений с использованием понятий оптической передаточной функции, зрачковой функции, автокорреляции этих функций.

Студент должен уметь:

- анализировать физические закономерности, определяющие процесс микролитографии при проекции микроизображений;
- математически описывать оптические процессы формирования проекционного микроизображения;
- рассчитывать профили распределения пространственной интенсивности излучения при проекционной микролитографии;
- оценивать применимость проекционных оптических систем для реализации процессов микролитографии с заданными параметрами.

Студент должен иметь навыки применения математических методов фурье-преобразований для технологического анализа процессов на примере микролитографии.

7.1. Качество проекционного изображения

В проекционных системах фотолитографии изображение фотошаблона в плоскости фоторезиста формируется с помощью объектива. Наличие оптической системы предусматривает применение иных, нежели при контактном экспонировании, методов расчета формируемых микроизображений и параметров для оценки качества.

Для оценки предельного разрешения проекционной системы используется критерий Рэля, учитывающий основной параметр объектива — его апертуру.

На практике разрешающая способность (разрешение) часто оценивают с помощью тестовых решеток с одинаковыми прозрачными и непрозрачными полосами (линиями) (рис. 7.1, а). В этом случае разрешение эквивалентно предельному числу пар линий, воспроизведенных на 1 мм длины изображения, или пространственной частоте ν решетки с шагом P :

$$\nu = \frac{1}{2W} = \frac{1}{P}. \quad (7.1)$$

Качество пространственного изображения фотошаблона после прохождения оптической системы существенно зависит от апертуры. Чем меньше апертура объектива, тем более размытым и менее четким получается формируемое им изображение. Распределение интенсивности в изображении является уже не прямоугольным, а скорее колоколообразным. В области прозрачного окна фотошаблона интенсивность, как правило, не равна номинально заданной, а в области тени она не равна нулю (рис. 7.1, б).

Для количественной оценки качества пространственного изображения в оптике используют понятие *модуляции*, т. е. отношения амплитуды распределения интенсивности I_a к среднему значению I_m (рис. 7.1, в). Часто применяется также эквивалентное понятие *контраста*, который выражается через максимальное I_{\max} и минимальное I_{\min} значения интенсивности (освещенности) изображения (см. рис. 7.1, в):

$$T = \frac{I_a}{I_m} = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}. \quad (7.2)$$

Контраст и модуляция измеряются в долях единицы или в процентах.

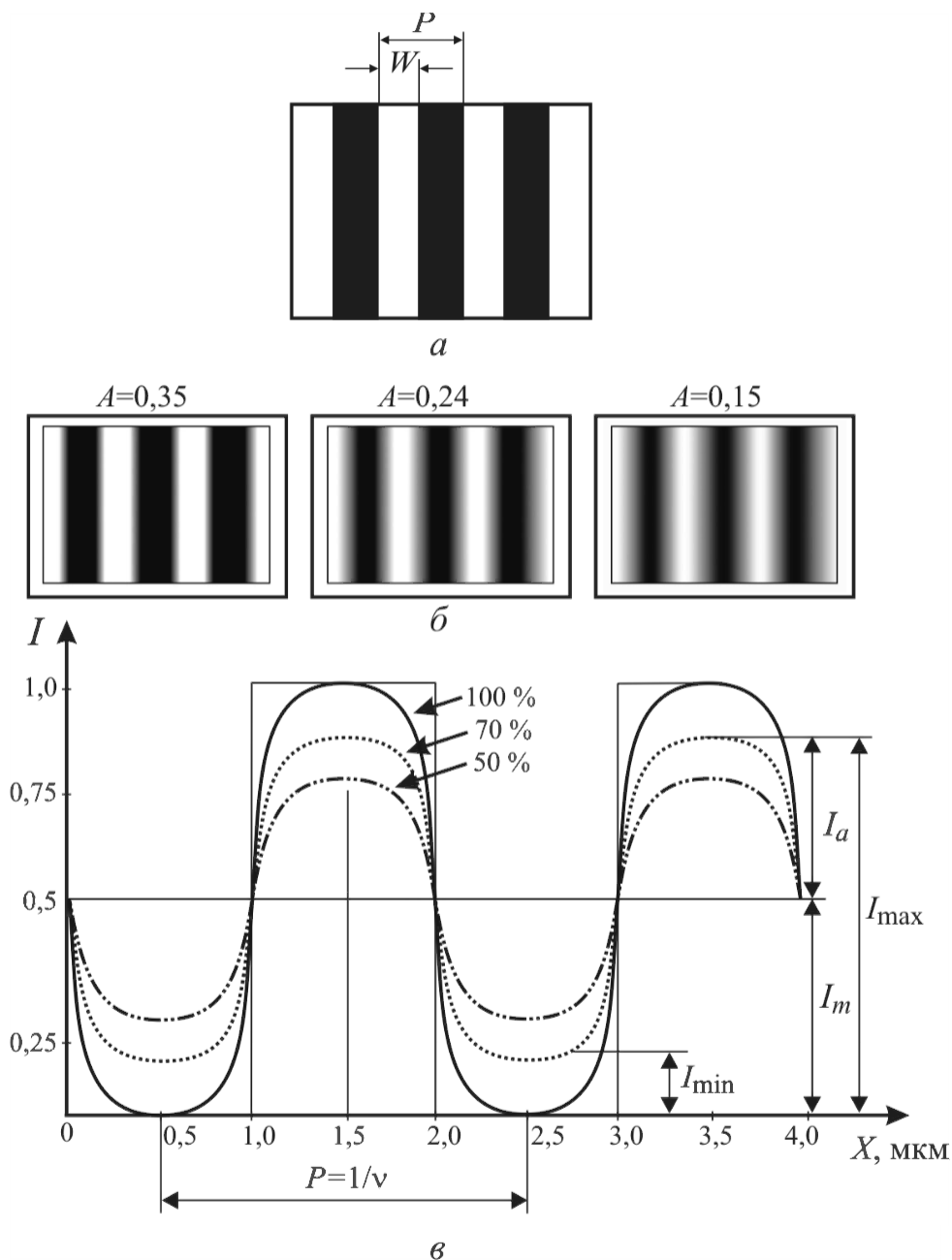


Рис. 7.1. Пространственное изображение элементов топологии: *a* — на фотошаблоне; *б* — после прохождения оптической системы; *в* — модуляция изображения

7.2. Понятие изображающей системы

Оптическая система должна преобразовывать и регистрировать максимальное количество информации, получаемой от объекта — фотошаблона. Необходимо также, чтобы основные параметры изображения — контраст отдельных элементов, их число на определенном участке и взаимное расположение — передавались в плоскость изображения с минимальными искажениями.

Чем меньше изображение отличается от объекта по распределению световой энергии, тем выше качество оптической системы. Двумерный объект можно характеризовать значениями освещенности I_o в каждой точке некоторого участка плоскости, т. е. математически представить объект в виде функции $I_o(x_o, y_o)$ двух независимых переменных.

Представим соответствующее изображение как функцию освещенности $I_i(x_i, y_i)$ на плоскости, а изображающую систему, в частности оптическую, в виде математического оператора, который показывает, как нужно подействовать на функцию объекта $I_o(x_o, y_o)$, чтобы получить функцию освещенности изображения $I_i(x_i, y_i)$. Такое воздействие оптической системы можно описать по крайней мере двумя различными способами.

Первый способ основан на представлении изображения состоящим из множества точек. Освещенность каждой из них зависит от значений яркости сопряженной точки объекта и соседних с нею точек.

Второй способ базируется на теории Фурье и заключается в представлении функции любого объекта и его изображения в виде сумм составляющих специального вида, а именно синусоидальных гармонических составляющих с различными частотами, амплитудами и начальными фазами. В таком случае воздействие оптической системы можно описать с помощью так называемых передаточных функций.

При обоих способах предполагают, что оптическая система, создающая изображение объекта, должна удовлетворять определенным требованиям, т. е. быть линейной и изопланатичной.

Свойство линейности означает следующее: если объектам I_{o1} и I_{o2} соответствуют изображения I_{i1} и I_{i2} , то объекту $I_o = aI_{o1} + bI_{o2}$ должно соответствовать изображение $I_i = aI_{i1} + bI_{i2}$ при любых

значениях множителей a и b . Следовательно, значение освещенности изображения, получаемое при суммарном воздействии нескольких значений яркости объекта, равно сумме значений освещенности, которые получаются от каждого из этих значений в отдельности.

Изопланатичность системы означает, что в ней форма функции рассеяния точки и линии сохраняется в пределах изучаемого участка изображения. Оптические системы обеспечивают выполнение этого условия не по всей площади изображения, а лишь в пределах ограниченных участков поля зрения, называемых *изопланатическими зонами*.

7.3. Связь между объектом и изображением

Изображение, построенное оптической системой, всегда отличается от объекта тем, что переходы между соседними частями получаются в изображении более плавными, чем в объекте, а границы — как бы размытыми.

Для учета действия оптической системы (т. е. для установления связи между объектом и изображением) объект обычно представляют в виде совокупности простых по форме элементов, причем таких, чтобы действие оптической системы на каждый элемент в отдельности можно было бы полностью изучить, а также описать математически. Если такие элементы в сумме составляют объект, то, суммируя действие оптической системы на отдельные элементы объекта (что допустимо ввиду линейности системы), можно полностью описать элементы изображения.

Примем два допущения, не искажающие анализа качества изображения.

1. Будем считать масштаб изображения равным единице и обозначим координаты в плоскостях объекта и изображения одинаково (x, y), так как их числовые значения совпадают (при всех расчетах должно учитываться конкретное значение масштаба).

2. Не будем учитывать снижение освещенности изображения в зависимости от эффективного относительного отверстия системы, а две величины — яркость объекта и освещенность изображения — назовем интенсивностями применительно к рассматриваемому случаю некогерентного освещения.

В качестве элементов простой формы, на которые можно разложить объект, обычно выбирают светящуюся точку и бесконечно длинную светящуюся линию (рис. 7.2).

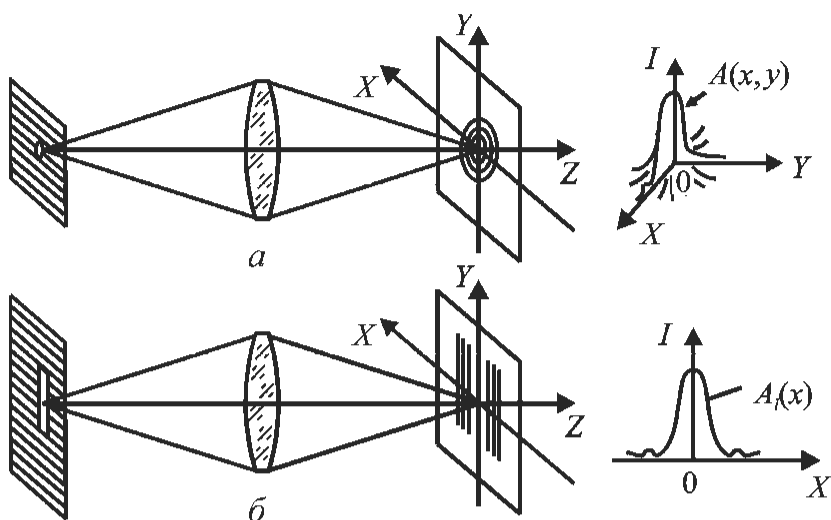


Рис. 7.2. Получение изображений простейших объектов:
a — точки, *б* — линии

Изображение отдельной светящейся точки объекта в действительности никогда не может быть бесконечно малой точкой. Даже если оптическая система не имеет aberrаций, изображение точки представляет собой протяженное пятно рассеяния из-за дифракции на диафрагме, ограничивающей апертуру оптической системы. В частности, при круглой форме зрачка оптической системы пятно рассеяния имеет вид известной дифракционной картины с центральным диском Эри, окруженным кольцами убывающей интенсивности. Математически распределение интенсивности в изображении изолированной точки описывается функцией рассеяния точки $A(x, y)$, график которой приведен в правой части рис. 7.2, *a*.

Поскольку представляет интерес лишь относительное распределение интенсивности I , удобнее нормировать функцию рассеяния, приняв объем под ее графиком за единицу:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} A(x, y) dx dy = 1. \quad (7.3)$$

Изображение бесконечно длинной светящейся линии можно получить суммированием бесконечного количества светящихся точек, расположенных вдоль нее. Поэтому математическое описание изображения, называемое функцией рассеяния линии (или щели) $A_l(x)$ (рис. 7.2, б), имеет вид

$$A_l(x) = \int_{-\infty}^{+\infty} A(x, y) dy. \quad (7.4)$$

Из формул (7.3) и (7.4) следует нормировка функции рассеяния линии:

$$\int_{-\infty}^{+\infty} A_l(x) dx = 1. \quad (7.5)$$

Функцию рассеяния точки (и линии) в литературе иногда называют *импульсным откликом* (или *импульсной реакцией*). Это название пришло из радиотехники и теории связи, где оно применяется для описания формы выходного сигнала, если на вход устройства подан бесконечно короткий импульс. Понятно, что в случае оптических систем аналогами такого импульса являются бесконечно малая (не имеющая протяженности) точка и бесконечно узкая щель (см. формулы (7.3) – (7.5)).

7.4. Свертка

Простые по форме элементы объекта, такие, как точка или линия, изображаются проекционной оптической системой в виде функций рассеяния. Чтобы определить интенсивность $I_l(x)$ в некоторой точке x изображения, нужно просуммировать ординаты функций рассеяния, полученные от точки объекта с той же координатой x и от соседних точек объекта, функции рассеяния которых вносят свой вклад в интенсивность в точке x .

Математическая операция, позволяющая провести такое построение, называется *сверткой*. Для пояснения смысла операции свертки воспользуемся свойствами прямоугольной (единичной импульсной) функции.

Рассмотрим последовательность прямоугольных импульсов $f(x)$ (рис. 7.3, а), у которых с ростом номера уменьшается ширина,

но увеличивается высота, в связи с чем площадь импульса остается постоянной и равной единице.

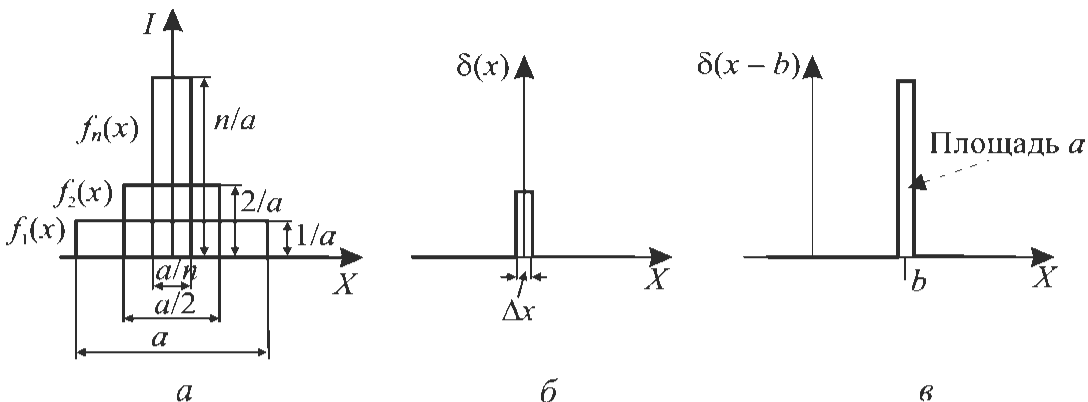


Рис. 7.3. Прямоугольные функции:

a — последовательность прямоугольных импульсов; b — единичный импульс; v — сдвиг и масштабирование единичного импульса

Выберем узкую длинную прямоугольную функцию δx , определенную в интервале от $-\Delta\xi/2$ до $+\Delta\xi/2$ и имеющую высоту $1/\Delta\xi$ (рис. 7.3, б). При умножении функции δx на постоянную (например, a) площадь под графиком функции также увеличится в a раз. Смещение графика этой функции вдоль оси X на расстояние b приведет к вычитанию b из переменной x : $\delta(x - b)$ (рис. 7.3, в).

Суммируя импульсные функции, масштабированные различными постоянными и смещенные на разные расстояния вдоль оси X , можно получить такие совокупности импульсов, которые будут описывать любые произвольные функции.

Рассмотрим последнее утверждение более подробно.

Согласно определению отклик пространственно инвариантной системы на сумму импульсов будет равен сумме откликов системы на каждый отдельный импульс. По известному отклику системы на единичный импульс можно найти отклик системы на единичный импульс, увеличенный в a раз и смещенный на расстояние b . Другими словами, мы можем найти отклик на каждый импульс из совокупности, составляющей $I(x)$, т. е. отклик всей функции $I(x)$.

Математически приведенное выше положение сформулируем следующим образом.

Во-первых, представим $I(x)$ в виде суммы импульсных функций δx (рис. 7.4, а) шириной $\Delta\xi$ и высотой $1/\Delta\xi$. Для каждого конкретного значения x , например $x = \xi$ (рис. 7.4, б), высота единичной функции равна $I(x)$ и эта функция сдвинута на ξ от центра.

Соотнесем между собой значения функции $I(x)$ и прямоугольной функции δx . При $x = \xi$ высота $I(\xi)$ в $I(\xi)\Delta(\xi)$ раз больше высоты функции δx , так как $I(\xi)/\delta x = I(\xi)\Delta(\xi)$. Следовательно, прямоугольную функцию при $x = \xi$ можно представить в виде

$$I_o(x) = I(\xi)\Delta\xi\delta(x - \xi). \quad (7.6)$$

Обратим внимание на то, что $\delta(x - \xi)$ — это сдвинутая прямоугольная функция, а $I(x)$ — это, по существу, масштабный коэффициент — аналог a (см. рис. 7.3, в).

Изменяя значения ξ , получаем набор прямоугольных функций, сумма которых описывает входную функцию, т. е. распределение интенсивности на объекте $I_o(x)$:

$$I_o(x) = \sum_{\xi=-\infty}^{\xi=+\infty} I(\xi)\delta(x - \xi)\Delta\xi. \quad (7.7)$$

Во-вторых, найдем отклик системы на отдельный импульс. Пусть прямоугольной функции x на входе оптической системы соответствует функция $A_I(x)$ на выходе (рис. 7.4, в).

Из пространственной инвариантности системы следует, что смещение входной функции на ξ вызывает такое же смещение выходной функции $A_I(x)$, не изменяя ее вида. Таким образом, входная прямоугольная функция (7.6) размывается оптической системой (рис. 7.4, г) и на выходе описывается выражением

$$I_i = I(\xi)A_I(x - \xi)\Delta\xi. \quad (7.8)$$

В-третьих, объединим отклики на входные импульсы (см. рис. 7.4, г). Совокупности импульсных функций, описывающих распределение интенсивности на объекте $I_o(x)$, соответствует сумма выходных функций, описывающих распределения интенсивности в изображении $I_i(x)$:

$$I_i(x) = \sum_{\xi=-\infty}^{\xi=+\infty} I(\xi)A_I(x - \xi)\Delta\xi. \quad (7.9)$$

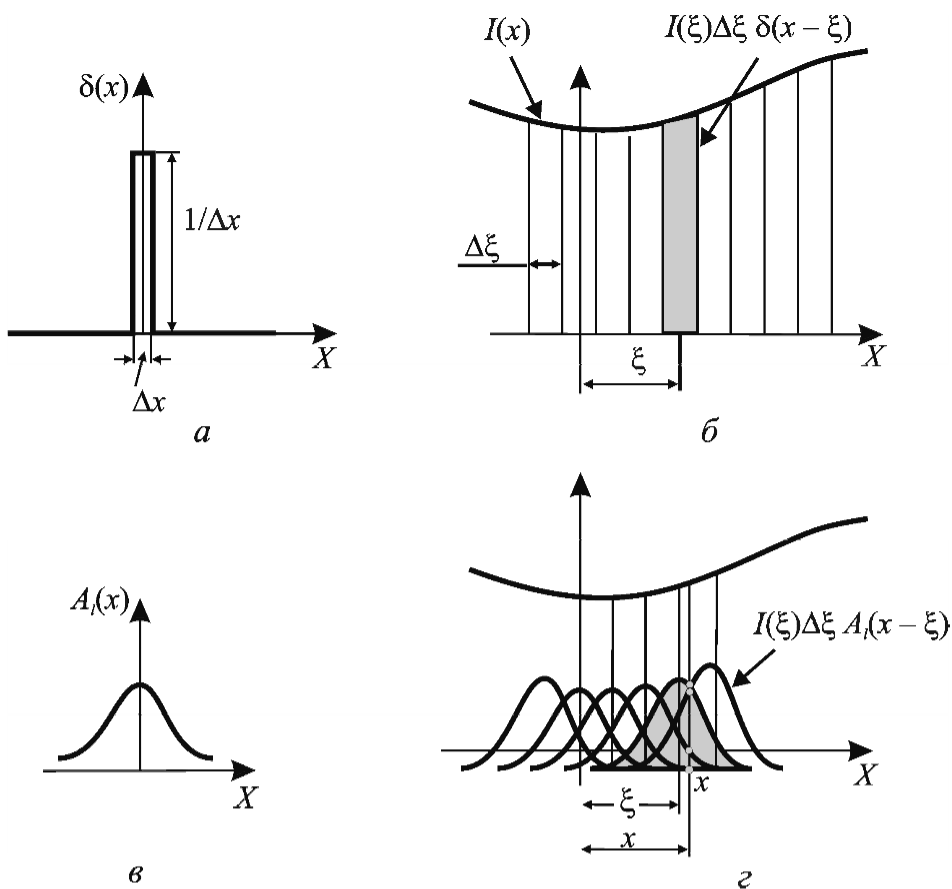


Рис. 7.4. Свертка функции с откликом системы:

a — прямоугольная функция; *б* — входная функция в виде совокупности прямоугольных функций; *в* — отклик системы (функция рассеяния); *г* — формирование свертки в точке с координатой x

Заменим теперь конечное приращение $\Delta\xi$ бесконечно малым $d\xi$. При этом сумма переходит в интеграл:

$$I_i(x) = \int_{\xi=-\infty}^{\xi=+\infty} I(\xi)A_i(x-\xi)\Delta\xi. \quad (7.10)$$

Выражение (7.10) представляет собой интеграл свертки входной функции и отклика на узкую импульсную функцию.

Важным свойством интеграла свертки является коммутативность, которая в исследуемом случае проявляется следующим образом. Отметим, что импульсная функция, отстоящая от точки x на величину ξ , т. е. в точке $(x - \xi)$, имеет ординату $I(x - \xi)$. Очевидно, что соответствующая функция рассеяния этого импульса на рас-

стоянии ξ , т. е. в точке x , имеет ординату $A_I(\xi)$, умноженную на значение функции в точке $(x - \xi)$: $I(x - \xi)A_I(\xi)$.

Объединив функции рассеяния (отклики на входные импульсные функции), получим

$$I_i(x) = \int_{\xi=-\infty}^{\xi=+\infty} I(x - \xi)A_I(\xi)d\xi. \quad (7.11)$$

Уравнение (7.11) эквивалентно уравнению (7.10), т.е. мы показали, что интеграл свертки действительно обладает свойством коммутативности.

Расчет распределения интенсивности в изображении может быть существенно упрощен при использовании преобразований Фурье.

7.5. Фурье-преобразования в оптике

Понятие пространственной частоты

В параграфе 7.2 мы рассматривали объект как совокупность элементов простой формы: точек или линий. Но можно представить произвольный объект как совокупность элементарных пространственных составляющих (гармоник), интенсивность которых изменяется по синусоидальному или косинусоидальному закону и которые отличаются друг от друга по частоте, амплитуде и фазе.

Применительно к решетке пространственная частота в соответствии с формулой (7.1) равна количеству ее периодов, размещающихся на единице длины. Так, n -я пространственная гармоника периода P имеет период P/n , что соответствует пространственной частоте n/P , которая представляет собой число повторений на единице длины изображения.

Ряды Фурье

Периодическую функцию $f(x)$ аргумента x , имеющую период P , т. е. пространственную частоту $\nu = 1/P$, можно представить в виде суммы синусоид или косинусоид, имеющих частоты $\nu, 2\nu, 3\nu, \dots, n\nu$ и периоды $P, P/2, P/3, \dots, P/n$:

$$f(x) = \frac{a_0}{2} + a_1 \cos \frac{2\pi x}{P} + b_1 \sin \frac{2\pi x}{P} + \dots + a_n \cos n \frac{2\pi x}{P} + b_n \sin n \frac{2\pi x}{P} + \dots =$$

$$= \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos \frac{2\pi n x}{P} + \sum_{n=1}^{\infty} b_n \sin \frac{2\pi n x}{P}, \quad (7.12)$$

где P измеряется в миллиметрах, а v — в единицах на миллиметр.

Коэффициенты такого ряда определяются по формулам

$$a_0 = \frac{2}{P} \int_P f(x) dx; \quad a_n = \frac{2}{P} \int_P f(x) \cos \frac{2\pi n x}{P} dx; \quad b_n = \frac{2}{P} \int_P f(x) \sin \frac{2\pi n x}{P} dx. \quad (7.13)$$

Легко убедиться, что для четной функции все $b_n = 0$; для нечетной функции все $a_n = 0$.

Тестовые решетки с одинаковыми прозрачными и непрозрачными полосами (см. рис. 7.1, а), имеющие прямоугольное (П-образное) пропускание, могут быть описаны рядами Фурье:

$$f(x) = 0,5 + \frac{2}{\pi} \cos \frac{2\pi x}{P} - \frac{2}{3\pi} \cos 3 \frac{2\pi x}{P} + \frac{2}{5\pi} \cos 5 \frac{2\pi x}{P} - \dots; \quad (7.14)$$

$$f(x) = 0,5 + \frac{2}{\pi} \sin \frac{2\pi x}{P} + \frac{2}{3\pi} \sin 3 \frac{2\pi x}{P} + \frac{2}{5\pi} \sin 5 \frac{2\pi x}{P} + \dots \quad (7.15)$$

Уравнение (7.14) соответствует расположению начала координат в центре окна (рис. 7.5, а), а уравнение (7.15) — случаю, когда начало координат совпадает с границей окна (рис. 7.5, б).

На рис. 7.5, в в пределах одного периода показаны частные суммы двух и трех членов ряда Фурье. Увеличение числа членов ряда Фурье делает суммарную функцию все более похожей на исходное прямоугольное распределение.

Ряд Фурье в комплексной форме

Разложение Фурье может быть записано в более простой форме с помощью комплексных экспоненциальных функций. Используем для этого экспоненциальные выражения для косинуса и синуса, являющиеся следствием формулы Эйлера:

$$\cos \theta = (e^{j\theta} + e^{-j\theta})/2; \quad \sin \theta = (e^{j\theta} - e^{-j\theta})/(2j). \quad (7.16)$$

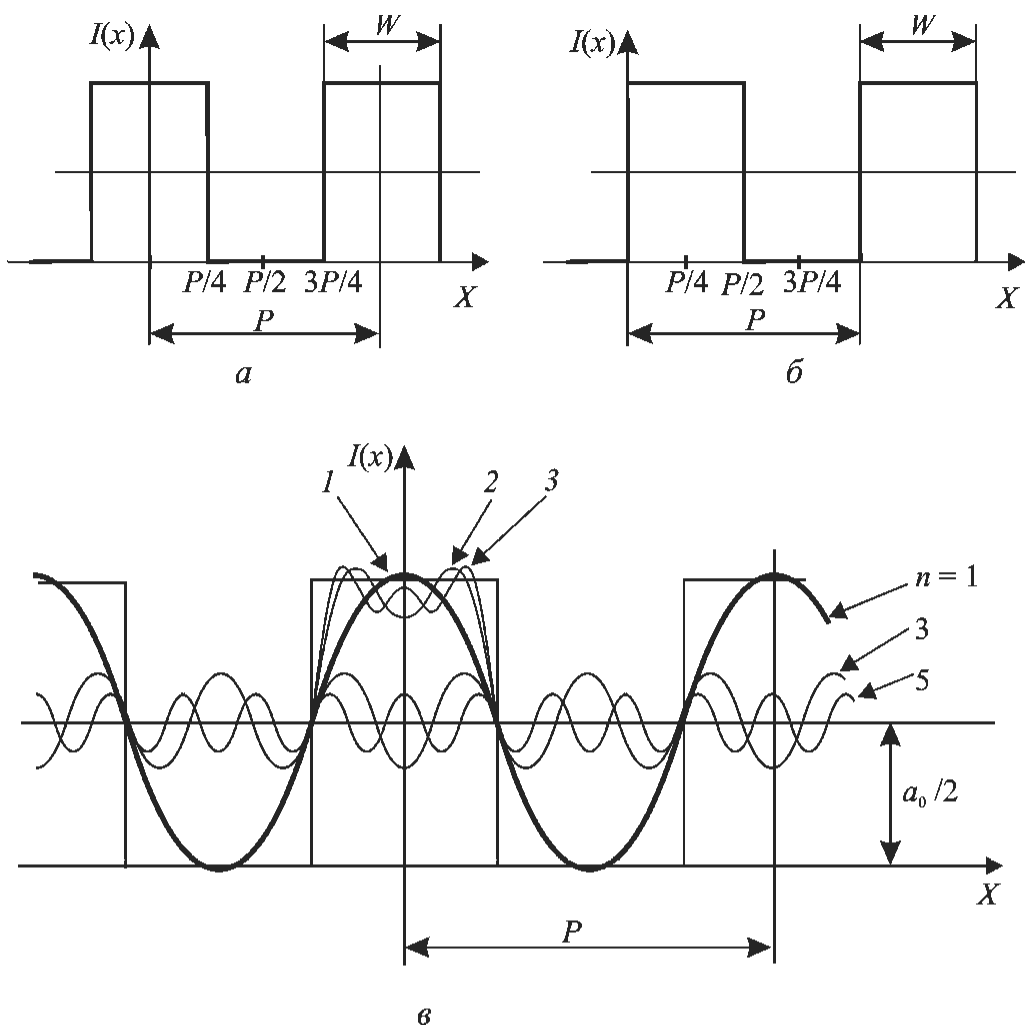


Рис. 7.5. Представление функции рядом Фурье:

1 — первый член ряда; 2 — сумма двух членов ряда Фурье; 3 — сумма трех членов ряда Фурье

Представим общий член выражения (7.13) в виде

$$\begin{aligned}
 a_n \cos \frac{2\pi n x}{P} + b_n \sin \frac{2\pi n x}{P} &= a_n \frac{e^{j\frac{2\pi n x}{P}} + e^{-j\frac{2\pi n x}{P}}}{2} + b_n \frac{e^{j\frac{2\pi n x}{P}} - e^{-j\frac{2\pi n x}{P}}}{2j} = \\
 &= \frac{a_n - j b_n}{2} e^{j\frac{2\pi n x}{P}} + \frac{a_n + j b_n}{2} e^{-j\frac{2\pi n x}{P}}. \quad (7.17)
 \end{aligned}$$

Если мы обозначим

$$C_n = \frac{a_n - j b_n}{2}; \quad C_{-n} = \frac{a_n + j b_n}{2}; \quad C_0 = \frac{a_0}{2}, \quad (7.18)$$

то формула (7.13) примет вид

$$f(x) = \sum_{n=-\infty}^{n=+\infty} C_n e^{j \frac{2\pi n x}{P}}. \quad (7.19)$$

Отметим, что суммирование ведется по целым (как положительным, так и отрицательным) значениям n , включая также и нуль. При этом комплексные коэффициенты $C_{n<0} = C_{n>0}^*$ и поэтому $|C_{n<0}| = |C_{n>0}^*|$, где звездочка означает комплексное сопряжение.

Формула (7.19) имеет очень простой вид. Покажем, что существует также очень простая формула для определения коэффициентов C по заданной функции $f(x)$. Подставим для этого значения a_0, a_n, b_n из формул (7.14) в формулы (7.18) и получим:

$$\begin{aligned} C_n &= \frac{a_n - j b_n}{2} = \frac{1}{P} \int_P f(x) \left(\frac{\cos 2\pi n x}{P} - j \frac{\sin 2\pi n x}{P} \right) dx = \\ &= \frac{1}{P} \int_P f(x) \exp \left(-j \frac{2\pi n x}{P} \right) dx; \quad (7.20) \end{aligned}$$

$$\begin{aligned} C_{-n} &= \frac{a_n + j b_n}{2} = \int_P f(x) \left(\frac{\cos 2\pi n x}{P} + j \frac{\sin 2\pi n x}{P} \right) dx = \\ &= \frac{1}{P} \int_P f(x) \exp \left(j \frac{2\pi n x}{P} \right) dx; \quad (7.21) \end{aligned}$$

$$C_0 = \frac{a_0}{2} = \frac{1}{P} \int_P f(x) dx. \quad (7.22)$$

Формулы (7.20) – (7.22) можно объединить в одну формулу

$$C_n = \frac{1}{P} \int_P f(x) \exp \left(-j \frac{2\pi n x}{P} \right) dx, \quad (7.23)$$

где n принимает все положительные и отрицательные целые значения, включая нуль. Таким образом, в комплексной форме разложение Фурье имеет вид

$$f(x) = \sum_{n=-\infty}^{n=+\infty} C_n \exp\left(j \frac{2\pi n x}{P}\right); \quad (7.24)$$

$$C_n = \frac{1}{P} \int_P f(x) \exp\left(-j \frac{2\pi n x}{P}\right) dx. \quad (7.25)$$

Интеграл Фурье

При увеличении периода P пространственные частоты становятся ближе друг к другу. Это означает, что в предельном случае непериодическая функция может содержать все частоты. Разложение таких функций осуществляется с помощью интеграла Фурье. Для разложения такой функции в выражение (7.19) подставим значение C_n из (7.23):

$$f(x) = \sum_{n=-\infty}^{\infty} \exp\left(j \frac{2\pi n x}{P}\right) \frac{1}{P} \int_P f(x) \exp\left(-j \frac{2\pi n x}{P}\right) dx. \quad (7.26)$$

Непериодическую функцию можно рассматривать как предельный случай периодической функции при стремлении периода к бесконечности, т. е. когда $P \rightarrow \infty$. В рассматриваемой формуле множители n/P можно принять за дискретные значения

$$v_1 = \frac{1}{P}; \quad v_2 = \frac{2}{P}; \quad v_3 = \frac{3}{P}; \quad \dots; \quad v_n = \frac{n}{P}$$

переменной v_n , непрерывно меняющейся от $-\infty$ до $+\infty$.

Приращение переменной $\Delta v_n = \Delta v_{n+1} - \Delta v_n = 1/P$. При $P \rightarrow \infty$ можно ввести замену $\Delta v \rightarrow dv = 1/P$. С учетом этой замены сумма (7.26) переходит в интеграл и мы получаем выражение интеграла Фурье в комплексной форме:

$$f(x) = \int_{-\infty}^{+\infty} F(v) \exp(j2\pi v x) dv, \quad (7.27)$$

причем функция $F(v)$ определяется через $f(x)$ формулой

$$F(v) = \int_{-\infty}^{+\infty} f(x) \exp(-j2\pi v x) dv. \quad (7.28)$$

Фурье-преобразование

Формулы (7.27) и (7.28) выражают так называемое фурье-преобразование функции $f(x)$. Отметим, что комплексная экспонента в формулах имеет разные знаки.

Для одномерного объекта с распределением интенсивности $f(x)$ прямое фурье-преобразование позволяет найти частотную или спектральную характеристику. Это значит, что интенсивность объекта представляется в виде бесконечно большого набора гармонических составляющих всех пространственных частот ν , а модуль $|F(\nu)|$ выражает ту долю, которая приходится на каждое выбранное значение частоты, т. е. выражает спектральную плотность.

Обратное фурье-преобразование позволяет восстановить распределение интенсивности по спектральной характеристике объекта.

Фурье-преобразования позволяют упростить расчеты распределения интенсивности в изображении, если известны распределение интенсивности в объекте и функция рассеяния системы.

Воспользуемся известной теоремой, которая формулируется так: фурье-преобразование некоторой функции, являющейся сверткой других функций, равно произведению фурье-преобразований функций, подвергаемых свертке.

Докажем эту теорему применительно к рассмотренному ранее примеру свертки, а именно к выражению интенсивности $I_i(x)$ в точке изображения. Для формулы (7.11) напишем фурье-преобразование обеих частей равенства:

$$\int_{-\infty}^{+\infty} I_i(x) e^{-j2\pi\nu x} dx = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} A_l(\xi) I(x - \xi) d\xi \right] e^{-j2\pi\nu x} dx. \quad (7.29)$$

Обозначим $I_i(\nu)$ фурье-преобразование функции $I_i(x)$, а в правой части равенства (7.29) выберем следующий порядок интегрирования:

$$I_i(\nu) = \int_{-\infty}^{+\infty} A_l(\xi) d\xi \int_{-\infty}^{+\infty} I(x - \xi) e^{-j2\pi\nu x} dx. \quad (7.30)$$

Во внутреннем интеграле правой части (7.30) введем новую переменную $x' = x - \xi$. Тогда

$$e^{-j2\pi nx} = e^{-j2\pi nx'} e^{-j2\pi n\xi}$$

и равенство (7.30) приобретет вид

$$I_i(v) = \int_{-\infty}^{+\infty} A(\xi) e^{-j2\pi n\xi} d\xi \int_{-\infty}^{+\infty} I(x') e^{-j2\pi nx'} dx', \quad (7.31)$$

т. е. действительно $I_i(v)$ равно произведению фурье-преобразований функций, связанных операцией свертки (в данном случае функции распределения интенсивности в объекте и функции рассеяния).

Фурье-преобразование изображения

Результат, полученный выше, позволяет существенно упростить нашу основную задачу — вычисление интенсивности $I_i(x)$ в плоскости изображения проекционной системы.

Действительно, теперь для решения этой задачи нет необходимости вычислять свертку распределения интенсивности в объекте $I_o(x)$ с функцией рассеяния линии $A_l(x)$. Значительно проще сделать это через фурье-преобразование, так как фурье-преобразование изображения равно произведению фурье-преобразований объекта $I_o(x)$ и изображения изолированной линии $A_l(x)$. Таким образом, если от распределения интенсивности в объекте $I_o(x)$ перейти к фурье-преобразованию, т. е. к спектру пространственных частот объекта,

$$I_o(v) = \int_{-\infty}^{+\infty} I(x') e^{-j2\pi vx'} dx', \quad (7.32)$$

а от функции рассеяния линии $A_l(\xi)$ — к фурье-преобразованию

$$A(v) = \int_{-\infty}^{+\infty} A(\xi) e^{-j2\pi v\xi} d\xi, \quad (7.33)$$

то спектр пространственных частот изображения $I_i(\nu)$, т. е. фурье-преобразование от распределения интенсивности в изображении $I_i(x)$, будет иметь вид

$$I_i(\nu) = I_o(\nu)A(\nu). \quad (7.34)$$

7.6. Оптическая передаточная функция

Из предыдущего параграфа следует очень важный вывод.

Если представить объект в виде суммы гармоник различных пространственных частот, можно оценить, какими они станут после прохождения оптической системы, т. е. в изображении. Для этого входные параметры каждой гармоники следует преобразовать с помощью соответствующего коэффициента $A(\nu)$. Просуммировав преобразованные гармоники, получим представление объекта после прохождения оптической системы, т. е. изображение.

Функция $A(\nu)$, определяющая, каким образом каждая частотная составляющая передается оптической системой с учетом дифракции, аберраций, ошибок изготовления оптической системы, называется *оптической передаточной функцией* (ОПФ) системы.

ОПФ, являющаяся, согласно выражению (7.33), преобразованием Фурье от функции рассеяния импульсной функции, представляет собой в общем случае комплексную функцию, которую можно записать в показательной форме через модуль комплексной функции $T(\nu)$ и аргумент $\varphi(\nu)$:

$$A(\nu) = \int_{-\infty}^{+\infty} A(\xi)e^{-j2\pi n\xi} d\xi = T(\nu)e^{j\varphi(\nu)}. \quad (7.35)$$

Естественно, ОПФ можно записать и в тригонометрической форме. Используя формулу Эйлера, имеем

$$\begin{aligned} A(\nu) &= \int_{-\infty}^{+\infty} A(\xi)e^{-j2\pi n\xi} d\xi = \int_{-\infty}^{+\infty} A(\xi)[\cos 2\pi n\xi - j \sin 2\pi n\xi] d\xi = \\ &= \int_{-\infty}^{+\infty} A(\xi) \cos 2\pi n\xi d\xi - j \int_{-\infty}^{+\infty} A(\xi) \sin 2\pi n\xi d\xi = C(\nu) - jS(\nu). \end{aligned} \quad (7.36)$$

Интегралы, обозначенные через C и S , называются соответственно косинус-преобразованием и синус-преобразованием функции рассеяния и представляют собою действительную и мнимую части комплексной функции $A(v)$. Поэтому модуль комплексной функции

$$T(v) = \sqrt{[C(v)]^2 + [S(v)]^2} = \sqrt{\left[\int_{-\infty}^{+\infty} A(\xi) \cos 2\pi n \xi d\xi \right]^2 + \left[\int_{-\infty}^{+\infty} A(\xi) \sin 2\pi n \xi d\xi \right]^2}, \quad (7.37)$$

а аргумент может быть определен из соотношений

$$\sin \varphi(v) = \frac{S(v)}{\sqrt{[C(v)]^2 + [S(v)]^2}} = \frac{S(v)}{T(v)} = \frac{\int_{-\infty}^{+\infty} A(\xi) \sin 2\pi n \xi d\xi}{T(v)};$$

$$\cos \varphi(v) = \frac{C(v)}{\sqrt{[C(v)]^2 + [S(v)]^2}} = \frac{C(v)}{T(v)} = \frac{\int_{-\infty}^{+\infty} A(\xi) \cos 2\pi n \xi d\xi}{T(v)}.$$
(7.38)

Для уяснения смысла модуля $T(v)$ и аргумента $\varphi(v)$ комплексной ОПФ рассмотрим более подробно, как изображается отдельная частотная составляющая объекта системой, имеющей известную функцию рассеяния.

Возьмем объект, имеющий косинусоидальное распределение интенсивности (рис. 7.6, а):

$$I_o(x) = I_0 + I_1 \cos 2\pi v x, \quad (7.39)$$

где I_0 — постоянная составляющая.

Если функцию рассеяния линии (при текущей переменной ξ) обозначить через $A_l(\xi)$, то интенсивность изображения можно записать в виде

$$I_i(x) = \int_{-\infty}^{+\infty} A_l(\xi) I(x - \xi) d\xi = \int_{-\infty}^{+\infty} A_l(\xi) [I_0 + I_1 \cos 2\pi v(x - \xi)] d\xi =$$

$$= I_0 \int_{-\infty}^{+\infty} A_l(\xi) d\xi + I_1 \int_{-\infty}^{+\infty} A_l(\xi) \cos 2\pi\nu(x - \xi) d\xi. \quad (7.40)$$

Интеграл в первом слагаемом представляет собой нормированную функцию рассеяния линии и равен единице. В подынтегральном выражении второго слагаемого применим формулу косинуса разности двух углов:

$$\begin{aligned} I_i(x) &= I_0 + I_1 \int_{-\infty}^{+\infty} A_l(\xi) [\cos 2\pi\nu x \cos 2\pi\nu\xi + \sin 2\pi\nu x \sin 2\pi\nu\xi] d\xi = \\ &= I_0 + I_1 \int_{-\infty}^{+\infty} A_l(\xi) \cos 2\pi\nu x \cos 2\pi\nu\xi d\xi + I_1 \int_{-\infty}^{+\infty} A_l(\xi) \sin 2\pi\nu x \sin 2\pi\nu\xi d\xi = \\ &= I_0 + I_1 \left[\cos 2\pi\nu x \int_{-\infty}^{+\infty} A_l(\xi) \cos 2\pi\nu\xi d\xi + \sin 2\pi\nu x \int_{-\infty}^{+\infty} A_l(\xi) \sin 2\pi\nu\xi d\xi \right]. \end{aligned} \quad (7.41)$$

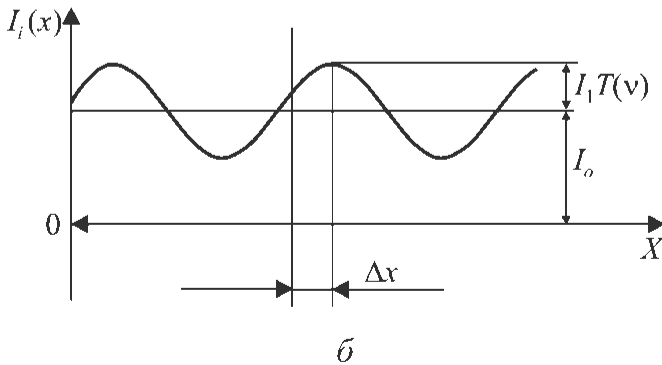
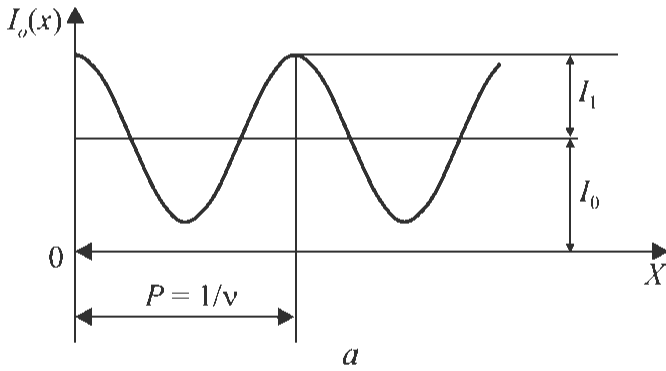


Рис. 7.6. Косинусоидальный объект (а) и его изображение (б)

Два интеграла в (7.41) — это знакомые нам по формуле (7.36) косинус-преобразование C и синус-преобразование S функции рассеяния. Следовательно, умножив и поделив на $\sqrt{[C(v)]^2 + [S(v)]^2}$, получим

$$\begin{aligned}
 I_i(x) &= I_0 + I_1 [C(v) \cos 2\pi vx + S(v) \sin 2\pi vx] = \\
 &= I_0 + I_1 \sqrt{[C(v)]^2 + [S(v)]^2} \times \\
 &\quad \times \left[\frac{C(v) \cos 2\pi vx}{\sqrt{[C(v)]^2 + [S(v)]^2}} + \frac{S(v) \sin 2\pi vx}{\sqrt{[C(v)]^2 + [S(v)]^2}} \right]. \quad (7.42)
 \end{aligned}$$

Дроби $C(v)$ и $S(v)$ согласно (7.38) равны

$$\begin{aligned}
 \frac{C(v)}{\sqrt{[C(v)]^2 + [S(v)]^2}} &= \cos \varphi(v); \\
 \frac{S(v)}{\sqrt{[C(v)]^2 + [S(v)]^2}} &= \sin \varphi(v). \quad (7.43)
 \end{aligned}$$

Учитывая также выражение (7.37), получим

$$\begin{aligned}
 I_0 + I_1 \sqrt{[C(v)]^2 + [S(v)]^2} \cdot [\cos \varphi(v) \cos 2\pi vx + \sin \varphi(v) \sin 2\pi vx] = \\
 = I_0 + I_1 T(v) \cos [2\pi vx - \varphi(v)]. \quad (7.44)
 \end{aligned}$$

Таким образом, изображение косинусоидального объекта (рис. 7.6, б) остается косинусоидальным и имеет такую же пространственную частоту, как объект. Тем не менее изображение (7.44) отличается от объекта (7.39) двумя особенностями.

Первая особенность состоит в том, что модуляция (отношение амплитуды переменной составляющей распределения к среднему значению — к постоянной составляющей) для изображения меньше, чем для объекта. В объекте $m_o = I_1/I_0$, а в изображении $m_i = I_1 T(v)/I_0$, т. е.

$$T(v) = \frac{m_i}{m_o}. \quad (7.45)$$

Таким образом, значение модуля $T(v)$ оптической передаточной функции для каждой пространственной частоты равно отношению модуляции гармонической составляющей в изображении к модуляции этой составляющей в объекте и называется *коэффициентом передачи модуляции* (КПМ) системы.

Совокупность значений КПМ для различных пространственных частот составляет функцию передачи модуляции (ФПМ) системы.

Следует отметить, что для значения пространственной частоты $v = 0$ значение $T(v) = 1$, что легко проверить подстановкой $v = 0$ в выражение (7.37). Примерная форма ФПМ показана на рис. 7.7, а.

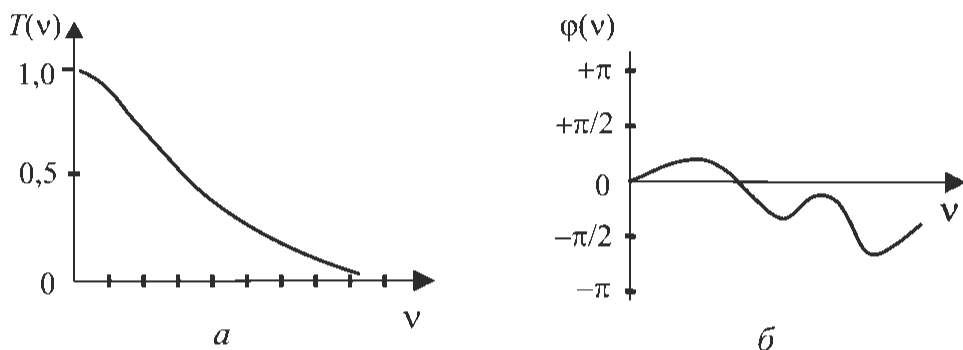


Рис. 7.7. Вид функции передачи модуляции (а) и фазы (б)

Вторая особенность заключается в том, что распределение интенсивности в изображении отличается от распределения в объекте еще и сдвигом косинусоиды на $\varphi(v)$ (в угловой мере).

Совокупность значений сдвига (смещения) фазы $\varphi(v)$ для различных пространственных частот составляет функцию передачи фазы (ФПФ) системы. Линейное смещение косинусоиды Δx (рис. 7.7, б) должно составлять, очевидно, такую же часть от периода, т. е. от $1/v$, какую фазовый угол сдвига $\varphi(v)$ составляет от 2π , т. е. $\Delta x = \varphi(v) / 2\pi$.

Для $v = 0$ угол $\varphi(v) = 0$, что легко проверить подстановкой $v = 0$ в формулу (7.38).

Форма ФПФ существенно зависит от симметричности функции рассеяния $A(\xi)$ относительно оси ординат.

В случае симметрии, т. е. если $A(\xi) = A(-\xi)$, функция рассеяния является четной. Тогда произведение $A(\xi) \sin 2\pi v x$ — нечетная функция (ввиду нечетности синуса), а синус-преобразование $S(v)$ в

формуле (7.37) оказывается равным нулю для всех значений (как интеграл от нечетной функции в симметричных пределах), т. е. функция передачи фазы равна нулю или π . Отметим, что при симметричной функции рассеяния $A(\xi)$

$$T(\nu) = C(\nu) = \int_{-\infty}^{+\infty} A_l(\xi) \cos 2\pi\nu\xi d\xi = \int_{-\infty}^{+\infty} A_l(\xi) e^{-j2\pi\nu\xi} d\xi. \quad (7.46)$$

При асимметричной функции рассеяния (для оптических систем это может быть на краю поля зрения, при дефектах центрирования и др.) функция передачи фазы для $\nu \neq 0$ может принимать значения между $+\pi$ и $-\pi$ (см. рис. 7.7, б).

7.7. Зрачковая функция и ее связь с оптической передаточной функцией

Введенные понятия передаточных функций дают возможность связать оптическую передаточную функцию системы с так называемой зрачковой функцией.

Эта функция описывает, как изменяется под действием оптической системы структура световой волны, исходящей из произвольной точки объекта M (рис. 7.8) и попадающей на входной зрачок 1 оптической системы. Эта световая волна является сферической, так как все лучи исходят из одной точки M . Поэтому сфера с центром в точке M , проходящая через центр O входного зрачка системы, представляет собой волновую поверхность постоянной фазы.

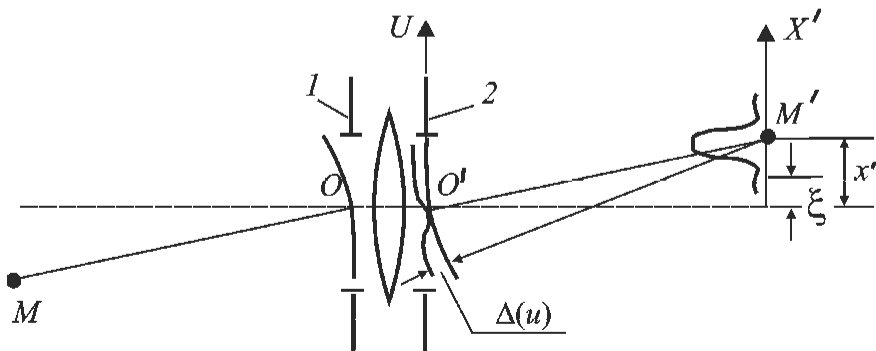


Рис. 7.8. Волновая aberrация оптической системы:
1 — входной зрачок; 2 — выходной зрачок

На выходе из оптической системы, имеющей абберации, форма этой волны отстает от сферической, поэтому лучи, вышедшие из системы и являющиеся нормальными к реальной волновой поверхности, после оптической системы уже не сходятся в одну точку M' .

Зададимся целью определить, насколько поверхность вышедшей световой волны, рассматриваемой вблизи выходного зрачка 2 системы, отстает от идеальной сферической. образуем для этого так называемую сферу сравнения с центром в точке M' , проходящую через центр O' выходного зрачка.

Текущие координаты точек на сфере сравнения обозначим U и V (ось U — в плоскости рис. 7.9, а ось V перпендикулярна к ней и на рисунке не показана).

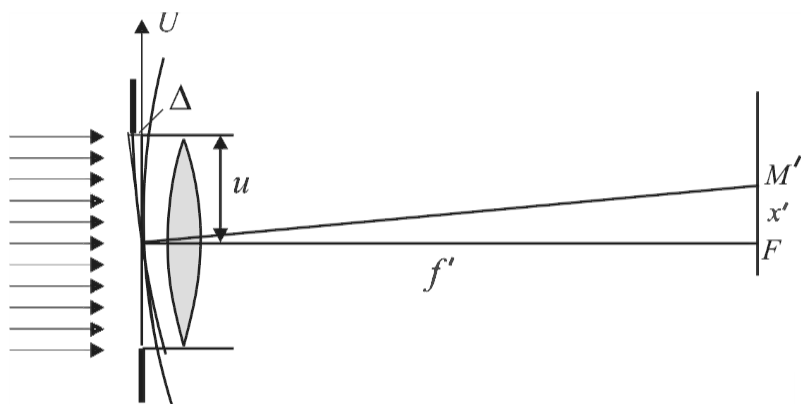


Рис. 7.9. Дифракция плоской волны на апертуре оптической системы

Отметим отрезки между реальной волновой поверхностью и сферой сравнения, отсчитывая их вдоль того луча, который проходит через интересующую нас точку (u, v) сферы сравнения. Эти отрезки (на рис. 7.9 они обозначены как Δ) могут быть различными для разных точек (u, v) ; их называют волновой абберацией лучей, проходящих через систему.

Волновая абберация $\Delta(u, v)$ дает сдвиг фазы волны $2\pi\Delta(u, v)/\lambda$. Подчеркнем, что такие сдвиги определяют фазу именно на сфере сравнения. Действительно, на реальной волновой поверхности фаза (по определению такой поверхности) постоянна. Добавляя отрезки $\Delta(u, v)$, можно перейти от этой поверхности к сфере сравнения. Так как сфера сравнения располагается у выход-

ного зрачка оптической системы, то «работает» лишь участок сферы, соответствующий размерам выходного зрачка.

Зрачковая функция учитывает не только волновую aberrацию, но и коэффициент пропускания, который в общем случае может быть различным для разных точек сферы (u, v) . Если коэффициент пропускания по интенсивности обозначить $\tau(u, v)$, то понятно, что амплитудный коэффициент пропускания равен $\sqrt{\tau(u, v)}$. Как правило, пропускание света приблизительно постоянно по всей площади выходного зрачка, и, считая его стопроцентным, можно положить $\tau(u, v) = 1$ в пределах контура выходного зрачка, а вне этого контура зрачковая функция равна нулю, т. е. принимается, что за пределами зрачка $\tau(u, v) = 0$.

В общем случае зрачковая функция $P(u, v)$ может быть представлена как комплексная:

$$P(u, v) = \sqrt{\tau(u, v)} e^{j(2\pi/\lambda)\Delta(u, v)}, \quad (7.47)$$

причем модуль $\sqrt{\tau(u, v)}$, т. е. коэффициент пропускания по площади зрачка, определяет уменьшение амплитуды пропускаемой световой волны, а аргумент — сдвиг фазы волны, вызванный волновой aberrацией $\Delta(u, v)$.

7.8. Связь комплексной амплитуды изображения со зрачковой функцией

Чтобы определить значение комплексной амплитуды на выходе из оптической системы, т. е. на сфере сравнения у выходного зрачка, нужно комплексную амплитуду на входе системы умножить на $\sqrt{\tau(u, v)}$ (для учета коэффициента пропускания системы) и на $e^{j(2\pi/\lambda)\Delta(u, v)}$ (для учета фазовых сдвигов, вызываемых волновой aberrацией системы), т. е. просто умножить на $P(u, v)$ из формулы (7.47). Это значит, что комплексная амплитуда волны на выходе из оптической системы пропорциональна зрачковой функции $P(u, v)$ системы.

Напомним, что нас прежде всего интересует распределение комплексной амплитуды $U(x)$ в плоскости изображения. Рассмотрим

рим, как оно связано с распределением комплексной амплитуды на выходе из системы, пропорциональным зрачковой функции. Для упрощения возьмем одномерный случай и приведем функцию $P(u, v)$ к $P(u)$, но результаты можно будет распространить на случай $U(x, y)$ и $P(u, v)$.

При изучении связи $U(x)$ с $P(u)$ учтем, что размеры зрачков оптической системы всегда ограничены и из-за дифракции изображение точки получается не точечным, а в виде дифракционного пятна рассеяния (см. рис. 7.8). Это происходит даже при отсутствии aberrаций, т. е. если реальная волновая поверхность на выходе из системы имеет сферическую форму.

Рассмотрим оптическую систему, эквивалентную бесконечно тонкой линзе. Пусть в плоскости входного зрачка этой линзы расположена прямоугольная (щелевая) апертура, на которую падает параллельный поток (плоская волна) излучения (см. рис. 7.9).

Поскольку это одномерный случай, будем считать, что в направлении, перпендикулярном плоскости рис. 7.9, все параметры потока постоянны, а длина апертуры не ограничена. При обозначении координат в плоскости апертуры будем использовать u и v , сохраняя традиционные обозначения x и y для плоскости объектов и x' и y' для плоскости изображений.

Параллельный пучок должен фокусироваться линзой в фокусе F плоскости изображения (x', y') , находящейся на фокусном расстоянии f от апертуры (см. рис. 7.9). На апертуре происходит дифракция, поэтому свет попадает не только в фокус F , но и в произвольные точки M' вблизи фокуса. Излучение от каждой точки плоской волны до точки M' будет доходить с некоторой разностью хода и соответствующим сдвигом фаз, нежели до фокуса F .

Суммарная амплитуда потока, попадающего в точку M' , может быть получена суммированием вкладов от элементов шириной du волнового фронта.

В области апертуры можно представить некую поверхность, расстояние от каждой точки которой до точки M будет одинаковым. Если расположить эту поверхность до входного зрачка линзы, то она будет плоскостью, расположенной под углом к падающей плоской волне. Возникающую вследствие этого оптическую разность хода до точки M' от соответствующих точек падающей

волны и от аппроксимирующей плоскости обозначим Δ , тогда фазовый сдвиг будет равен $2\pi\Delta/\lambda$.

Каждый элемент шириной du вносит в точке M' вклад в комплексную амплитуду величиной $(du)\cos 2\pi\Delta/\lambda$ или $(du)e^{j2\pi\Delta/\lambda}$. Как следует из рис. 7.9, $\Delta/u = x'/f'$, откуда

$$\Delta = ux'/f'. \quad (7.48)$$

Комплексная амплитуда $U(x')$ в точке изображения определяется следующим интегралом:

$$U(x') = \int_{-\infty}^{+\infty} P(u)e^{j2\pi u \left(\frac{x'}{\lambda f'} \right)} du, \quad (7.49)$$

где $P(u)$ — распределение амплитуды на входном зрачке.

Для придания формуле (7.49) большей наглядности пронормируем входящие в нее значения. Положим фокусное расстояние равным единице длины ($f' = 1$), а разность хода Δ будем измерять числом длин волн λ . В результате получим $\lambda f' = 1$, а формулу (7.49) приведем к виду

$$U(x') = \int_{-\infty}^{+\infty} P(u)e^{j2\pi ux'} du. \quad (7.50)$$

Подчеркнем, что формула (7.50) представляет собой фурье-преобразование зрачковой функции $P(u)$.

Будем полагать оптическую систему линейной и не поглощающей излучения. С учетом этого можно с помощью принципа обратимости восстановить распределение амплитуды в апертуре, имея распределение амплитуды в изображении. С учетом (7.50) получим

$$P(u) = \int_{-\infty}^{+\infty} U(x')e^{-j2\pi ux'} dx', \quad (7.51)$$

где знак минус появился в результате изменения направления хода лучей.

Интегралы (7.50) и (7.51) имеют вид пары фурье-преобразований. При комплексной зрачковой функции, т. е. при

наличии волновой aberrации $\Delta(u)$, подставив (7.47) в (7.50), найдем, что $\Delta(u)$ дает дополнительный к величине $2\pi ux'$ фазовый сдвиг $(2\pi/\lambda)\Delta(u)$. Для двумерного случая эти зависимости можно записать как

$$U(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(u, v) e^{j2\pi(ux+vy)} dudv; \quad (7.52)$$

$$P(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} U(x, y) e^{-j2\pi(ux+vy)} dx dy. \quad (7.53)$$

Таким образом, распределение амплитуд в апертуре оптической системы является фурье-преобразованием распределения амплитуд в плоскости изображения, и наоборот.

7.9. Оптическая передаточная функция как автокорреляция зрачковой функции

Обратимся к рис. 7.8, где показано образование пятна рассеяния вокруг точки M' , которая является геометрическим изображением сопряженной точки M объекта. Ясно, что рассмотрение дифракции на апертуре оптической системы принципиально не изменится, если оптическую ось системы заменить линией, соединяющей точки M и M' , т. е. считать, что оптическая ось направлена по оси OM' и что точка M' — начало оси координат x' . Координаты, отсчитываемые от точки M' , будем обозначать ξ , поэтому формулы (7.52) и (7.53) примут следующий вид:

$$U(\xi) = \int_{-\infty}^{+\infty} P(u) e^{j2\pi u \xi} du; \quad (7.54)$$

$$P(u) = \int_{-\infty}^{+\infty} U(\xi) e^{-j2\pi u \xi} d\xi. \quad (7.55)$$

Для двумерного случая эти зависимости можно записать так:

$$U(\xi, \eta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(u, v) e^{j2\pi(u\xi+v\eta)} dudv; \quad (7.56)$$

$$P(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} U(\xi, \eta) e^{-j2\pi(u\xi + v\eta)} d\xi d\eta. \quad (7.57)$$

Отметим, что по своему физическому смыслу распределение интенсивности точки (или линии) в изображении описывается функцией рассеяния точки (линии). Таким образом, распределение световой волны $U(\xi, \eta)$ по формуле (7.57) эквивалентно амплитудной функции рассеяния точки. Аналогично для одномерного случая $U(\xi)$ (см. формулу (7.55)) эквивалентна функции рассеяния линии для рассматриваемого одномерного случая.

Амплитудная функция рассеяния связана со знакомой нам (см. формулу (7.3)) функцией рассеяния точки $A(x, y)$, выражающей распределение интенсивности следующим соотношением:

$$A(x, y) = U(x, y)U^*(x, y). \quad (7.58)$$

Аналогично для функции рассеяния линии (см. формулу (7.4)) можно записать

$$A(x) = U(x)U^*(x). \quad (7.59)$$

Из определения оптической передаточной функции ее модуль для случая симметричной функции (см. формулу (7.46)) составляет

$$\begin{aligned} T(v) = C(v) &= \int_{-\infty}^{+\infty} A(\xi) e^{-j2\pi v\xi} d\xi = \int_{-\infty}^{+\infty} U(\xi)U^*(\xi) e^{-j2\pi v\xi} d\xi = \\ &= \int_{-\infty}^{+\infty} U(\xi) \left[\int_{-\infty}^{+\infty} P^*(U) e^{-j2\pi v\xi} du \right] e^{-j2\pi v\xi} d\xi. \end{aligned} \quad (7.60)$$

Интеграл в квадратных скобках равен $U^*(\xi)$ по известному свойству фурье-преобразования от сопряженной комплексной функции:

$$U^*(\xi) = \int_{-\infty}^{+\infty} P^*(U) e^{-j2\pi v\xi} du. \quad (7.61)$$

Для доказательства этого свойства следует подставить в выражение (7.54) комплексную функцию, представленную в виде $P(u) = a(u) + jb(u)$, и показательную функцию, преобразованную по фор-

муле Эйлера. Далее остается сравнить полученный результат с расчетом фурье-преобразования от комплексной сопряженной функции $P^*(u) = a(u) - jb(u)$. Для этого расчета также надо преобразовать показательную функцию $\exp(-j2\pi u\xi)$ по формуле Эйлера.

Меняя порядок интегрирования, продолжим преобразования:

$$T(v) = \int_{-\infty}^{+\infty} P^*(U) \left[\int_{-\infty}^{+\infty} U(\xi) e^{-j2\pi(u+\nu)\xi} d\xi \right] e^{-j2\pi\nu\xi} du. \quad (7.62)$$

Интеграл в квадратных скобках в соответствии с выражением (7.55) равен $P(u+\nu)$, поэтому ОПФ оптической системы

$$T(v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P^*(u) P(u+\nu) du. \quad (7.63)$$

Такой интеграл от произведения какой-либо функции и сопряженной и сдвинутой по аргументу, называется автокорреляционной функцией. Этот интеграл равен нулю при тех значениях u , при которых хотя бы одна из функций $P^*(u)$ или $P(u+\nu)$ равна нулю.

Для выяснения физического смысла сдвига ν вернемся к формуле (7.50) и вспомним условие $\lambda f' = 1$, принятое при выводе формулы (7.51). Отказавшись от этого ограничения, заметим, что зависящий от переменной фазовый член в формуле (7.50) имеет вид $2\pi\Delta/\lambda = 2\pi(ux'/f')/\lambda$.

Отношение x'/λ показывает, во сколько раз Δ (оптическая разность хода) меньше координаты u : $\Delta = u(x'/f')$. Выразим теперь u числом длин волн λ , соответствующих этой координате. Получим

$$\Delta/\lambda = u(x'/f')/\lambda. \quad (7.64)$$

Таким образом, фазовый член в случае произвольных λ и f' имеет вид $2\pi x(u/(\lambda f'))$. Фаза волны — величина безразмерная. В нашем случае она определяется произведением координаты x' на член $u/(\lambda f')$, имеющий размерность пространственной частоты (1/длина).

Учтем реальную размерность фазового члена, переписав интеграл (7.63) в следующем виде:

$$T(\nu) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P^*(u)P(u + \nu\lambda f') du = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P^*(u)P(u + u_0) du. \quad (7.65)$$

Так как интеграл (7.65) является функцией u_0 (при этом u — переменная интегрирования), то значению пространственной частоты ν соответствует $\nu = u_0 / (\lambda f')$, т. е.

$$u_0 = \lambda f' \nu. \quad (7.66)$$

При этом аргумент u и сдвиг u_0 выражаются в единицах длины.

Таким образом, ОПФ оптической системы можно определить, зная параметры только светового поля на сфере на выходе из системы, созданного малым источником — объектом.

Для двумерного случая выражение (7.65) можно переписать в виде

$$T(\nu_x, \nu_y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P^*(u, \nu)P(u + \lambda f' \nu_x, \nu + \lambda f' \nu_y) du d\nu. \quad (7.67)$$

Выполнив замену переменных $u' = u + \lambda f' \nu_x / 2$ и $\nu' = \nu + \lambda f' \nu_y / 2$, представим выражение (7.67) в симметричной форме:

$$T(\nu_x, \nu_y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P\left(u' + \frac{\lambda f' \nu_x}{2}, \nu' + \frac{\lambda f' \nu_y}{2}\right) \times \\ \times P^*\left(u' - \frac{\lambda f' \nu_x}{2}, \nu' - \frac{\lambda f' \nu_y}{2}\right) du' d\nu'. \quad (7.68)$$

Как следует из выражений (7.67) и (7.68), автокорреляционная функция отлична от нуля только в области перекрытия двух зрачков, смещенных один относительно другого на величину, пропорциональную пространственной частоте.

Эта область графически представлена на рис. 7.10, на котором видно, что пределы интегрирования по площади в (7.67) и (7.68) можно ограничить заштрихованной областью ΔG .

Ясно, что для пространственной частоты, равной нулю, сдвиг $\lambda f' \nu = 0$, т. е. область перекрытия в этом случае максимальна и равна площади зрачка G . Так как для нулевой пространственной час-

тоты значение ОПФ равно единице, автокорреляционную функцию по (7.67) и (7.68) можно нормировать делением на ее наибольшее значение, равное $\iint |P(u, v)|^2 dudv$.

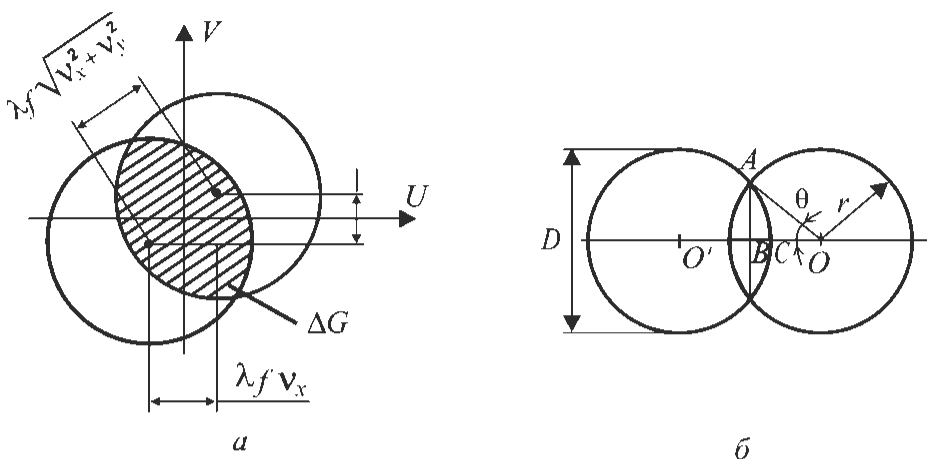


Рис. 7.10. Автокорреляция зрачковой функции:
a — область перекрытия зрачков; *б* — расчетная схема

Формулы (7.67) и (7.68) с переменными u, v после нормирования приобретают вид

$$\begin{aligned}
 T(\mathbf{v}_x, \mathbf{v}_y) &= \frac{A_0(\mathbf{v}_x, \mathbf{v}_y)}{\iint_G |P(u, v)|^2 dudv} = \frac{\iint_{\Delta G} P^*(u, v) P(u + \lambda f' v_x, v + \lambda f' v_y) dudv}{\iint_G |P(u, v)|^2 dudv} = \\
 &= \frac{\iint_{\Delta G} P\left(u' + \frac{\lambda f' v_x}{2}, v' + \frac{\lambda f' v_y}{2}\right) P^*\left(u' - \frac{\lambda f' v_x}{2}, v' - \frac{\lambda f' v_y}{2}\right) du' dv'}{\iint_G |P(u, v)|^2 dudv}. \quad (7.69)
 \end{aligned}$$

7.10. Системы дифракционного качества с постоянным пропусканием по площади зрачка

Формула (7.69) позволяет наглядно представить геометрический образ для ОПФ объектива (см. рис. 7.10). Используем этот результат для получения аналитического выражения ОПФ безабберационных систем. При рассмотрении оптических систем вы-

сокого качества часто полагают, что их остаточными оптическими искажениями (абберациями) можно пренебречь. В этом случае основные искажения изображения происходят из-за дифракции. Именно такие оптические системы называются системами дифракционного качества.

Если предположить, что пропускание по площади зрачка постоянно, то для систем дифракционного качества зрачковая функция будет действительной и, как следует из формулы (7.47), $P(u, v) = \sqrt{\tau(u, v)} = \text{const} = 1$. Следовательно, в выражении (7.69) функцию $P(u, v)$ можно вынести за знак интеграла:

$$T(v_x, v_y) = \frac{\iint dudv}{\iint_G dudv}. \quad (7.70)$$

Таким образом, значение оптической передаточной функции для безабберационной системы сводится к значению функции передачи модуляции. Это значение равно отношению площади области ΔG , зависящей от пространственной частоты, к площади зрачка G .

Площадь пересечения двух одинаковых кругов с известными значениями радиусов легко подсчитать.

Из рассмотрения рис. 7.10, б следует, что если θ — угол, образованный радиусом, соединяющим точку A пересечения дуг двух окружностей с центром O одной из них, то $AB = r \sin \theta$ и $BO = r \cos \theta$.

Площадь сектора, образованного углом θ , равна $S_1 = \theta r^2$, площадь треугольника ABO равна

$$S_2 = AB \cdot \frac{BO}{2} = \sin \theta \cos \theta \frac{r^2}{2}. \quad (7.71)$$

Площадь затемненной на рис. 7.10, б области перекрытия кругов

$$T_0 = 4 (S_1 - S_2) = 2 (\theta - \sin \theta \cos \theta) r^2,$$

а отношение ее к площади круга

$$T = \frac{T_0}{\pi r^2} = \frac{2}{\pi} (\theta - \sin \theta \cos \theta). \quad (7.72)$$

Сдвиг одного круга относительно другого равен $OO' = 2r \cos \theta$, максимальное его значение, при котором область перекрытия кругов исчезает, равно диаметру $D = 2r$, т. е. относительный сдвиг, изменяющийся от единицы до нуля, составляет

$$\frac{OO'}{D} = \frac{2r \cos \theta}{2r} = \cos \theta. \quad (7.73)$$

Наибольший сдвиг зрачка, связанный с нахождением автокорреляционной функции, является важной характеристикой: он определяет наибольшую предельную пространственную частоту ν_{lim} , пропускаемую оптической системой. При еще больших частотах коэффициент передачи модуляции системы становится равным нулю.

Полагая $u_0 = D$ (D — диаметр зрачка оптической системы), получаем $D = \lambda f' \nu_{\text{lim}}$. Отсюда

$$\nu_{\text{lim}} = \frac{D}{\lambda f'} = \frac{1}{\lambda f' / D}, \quad (7.74)$$

где ν_{lim} выражается числом периодов на единицу длины (например, 1/мм).

Отметим, что для большинства оптических систем соотношение D/f' можно заменить на $2 \sin \alpha$, где α — апертурный угол в пространстве изображений. В этом случае уравнение (7.74) принимает вид

$$\nu_{\text{lim}} = \frac{2A}{\lambda}. \quad (7.75)$$

Для представления кривой функции передачи модуляции безабберационного объектива удобно выразить пространственную частоту в относительных единицах, приняв предельное ее значение за единицу.

Частота в относительных единицах $R = \nu/\nu_{\text{lim}}$ связана с ν очевидным соотношением

$$R = \nu \frac{\lambda}{2A}, \quad (7.76)$$

а из (7.73) следует, что $R = \cos \theta$.

Таким образом, $\theta = \arccos R$ и уравнение функции передачи модуляции оптической системы дифракционного качества после подстановки этого значения θ в (7.72) примет вид

$$T(v) = \frac{2}{\pi} \left(\arccos R - R\sqrt{1-R^2} \right). \quad (7.77)$$

График этой функции (рис. 7.11) в области низких частот представляет собой практически прямую линию и лишь при $T(v) < 0,4$ искривляется по направлению к точке $R = 1$.

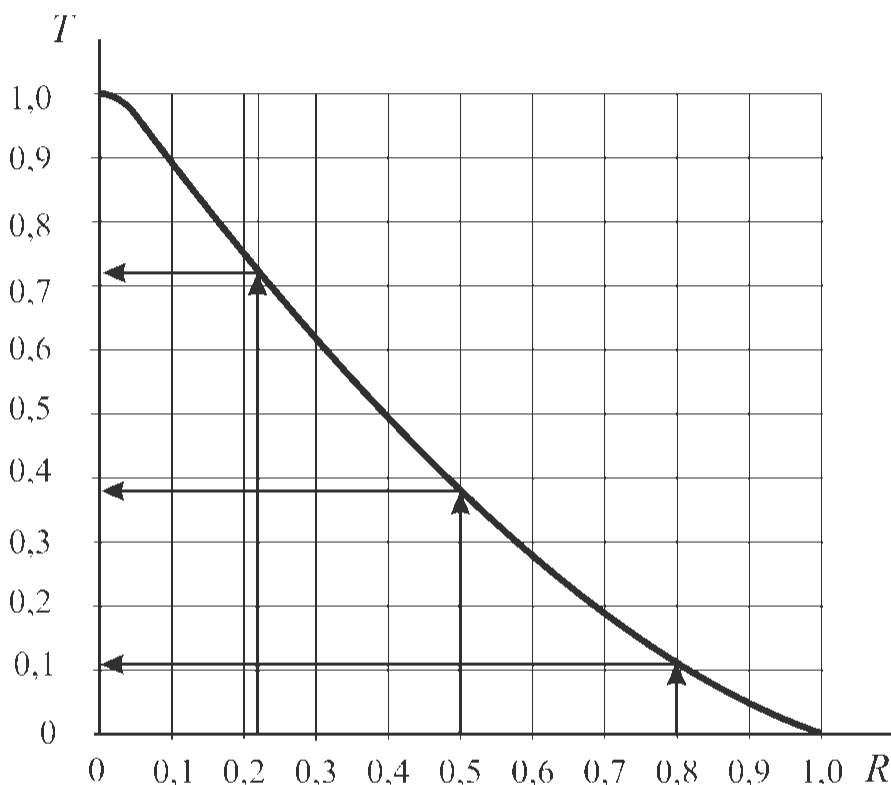


Рис. 7.11. Функция передачи модуляции

Функция $T(v)$ имеет ключевое значение для расчета пространственного распределения интенсивности, поскольку определяет уменьшение амплитуды всех составляющих разложения Фурье.

7.11. Учет распределения интенсивности в изображении

Распределение интенсивности в изображении, создаваемом оптической системой, можно найти при использовании фурье-преобразований функций, входящих в интеграл свертки. В резуль-

тате получают соответственно частотное распределение интенсивности в объекте и оптическая передаточная функция. Простое перемножение этих функций позволяет получить частотное распределение интенсивности в изображении. Применяв к нему обратное фурье-преобразование, можно найти искомое распределение интенсивности в изображении.

Каждая косинусоидальная составляющая пространственного распределения в объекте преобразуется в изображение без изменений ее частоты, меняются только амплитуда и фаза этой составляющей.

Для случая дифракционно-ограниченной системы, когда волновые aberrации отсутствуют, может быть предложена следующая процедура.

1. Для известного (заданного) пространственного распределения интенсивности $I_o(x)$ в объекте находят частотное распределение объекта $I_o(v)$ с помощью фурье-преобразования:

$$I_o(v) = F_T [I_o(x)]. \quad (7.78)$$

Результатом является частотный спектр объекта, то есть набор значений коэффициентов Фурье для различных гармоник, описывающих заданный профиль распределения интенсивности в объекте.

2. Рассчитывают частотное распределение изображения $I_i(v)$. Для этого частотное распределение объекта умножают на коэффициент передачи модуляции:

$$I_i(v) = T(v) [I_o(v)]. \quad (7.79)$$

Предварительно для каждой частоты рассчитывают соответствующее значение $T(v)$, далее умножают его на значение коэффициента Фурье соответствующей гармоники.

3. Находят пространственное распределение интенсивности в изображении за счет обратного фурье-преобразования:

$$I_i(x) = F_T^{-1} [I_i(v)]. \quad (7.80)$$

Полученные в (7.79) новые значения коэффициентов Фурье используются для воссоздания соответствующих гармоник разложения. Суммирование значений гармоник для текущей координа-

ты профиля позволяет получить значение интенсивности в изображении для данной координаты.

4. Оценивают распределение интенсивности в изображении с технологической точки зрения. Дифракционное размытие изображения приводит к тому, что на участки фоторезиста, лежащие в области геометрической тени, попадает часть экспонирующего излучения. Получаемая этими участками доза экспозиции может стать недостаточной для утонения фоторезиста, в результате он не сможет служить защитной маской при последующих операциях.

Деструкция фоторезиста в затененных участках будет зависеть не только от степени размытия изображения, но и от контраста самого фоторезиста, который часто выражается через его модуляцию.

Для получения микрорельефа в фоторезисте необходимо, чтобы модуляция оптического изображения превышала модуляцию фоторезиста, т. е. чтобы выполнялось условие $T > T_f$. У фоторезистов, согласованных с источниками излучения в диапазоне линий g , h , i , контраст составляет $\gamma = 2...3$, при этом согласно формуле (7.2) модуляция равна $T_f = 0,36...0,70$. Поэтому на значение оптического контраста накладывается ограничение $T \geq 0,6$.

Как следует из рис. 7.12, изображение с высоким оптическим контрастом имеет меньший клин проявления. Показанное пунктиром уменьшение разности между E_1 и E_0 (дозы начала и полного проявления фоторезиста) также делает изображение более резким.

Таким образом, контраст (модуляция) оптического изображения и контраст фоторезиста определяют разрешение всего литографического процесса. Поэтому для полученного распределения интенсивности в изображении необходимо оценить его контраст, сравнить с требуемым контрастом и вынести решение о возможности получения заданного изображения в данной литографической системе.

Контрольные вопросы и задания

1. Опишите основные свойства оптической передаточной функции $T(v)$. Объясните их математическое происхождение со ссылками на соответствующие формулы. Дайте графическую иллюстрацию функции $T(v)$.

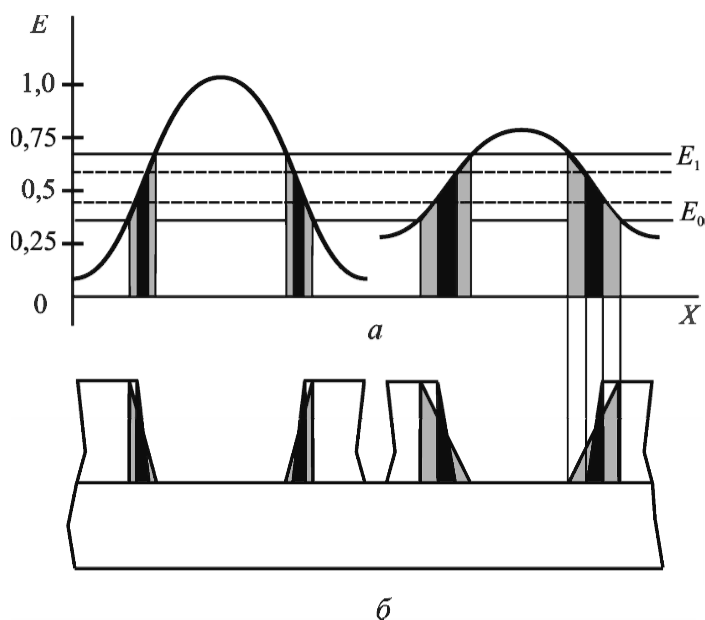


Рис. 7.12. Влияние параметров изображения и фоторезиста на качество микрорельефа:
a — контрасты изображения и фоторезиста; *б* — клин проявления

2. Оптическая передаточная функция $T(\nu)$ определена для объекта с косинусоидальным распределением интенсивности. Что можно сказать про $T(\nu)$ для объекта с прямоугольным распределением интенсивности?

3. Разложите в ряд Фурье заданную функцию распределения интенсивности. Разложение выполните в комплексной и тригонометрической форме. Сравните результаты.

4. Поясните этапы выполнения расчета интенсивности в проекционном изображении:

- нахождение частотного распределения объекта;
- нахождение частотного распределения изображения (определение значения $T(\nu)$ для каждой частоты);
- нахождение пространственного распределения изображения (выполнение обратного фурье-преобразования).

Каждый этап этого вопроса проиллюстрируйте графически.

5. По рассчитанному распределению интенсивности оцените применимость литографической системы для формирования заданного микроизображения.

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

Балекин В.И., Иванов А.Н. Основы расчета и конструирования электронных пушек: Конспект лекций. Л.: ЛЭТИ, 1979.

Березин Г.Н., Никитин А.С., Сурис Р.А. Оптические основы контактной фотолитографии. М.: Радио и связь, 1982.

Блох А.Г., Журавлев Ю.А., Рыжков Л.Н. Теплообмен излучением: Справ. / М.: Энергоатомиздат, 1991.

Броудай И., Мерей Дж. Физические основы микротехнологии: Пер. с англ. М.: Мир, 1987.

Вакуумная техника: Справ. / Под ред. Е.С. Фролова, В.Е. Минайчева. М.: Машиностроение, 1992.

Введение в фотолитографию / Под ред. В.П. Лаврищева. М.: Энергия, 1977.

Виноградов Ю.В. Основы электроники и полупроводниковой техники: Учеб. для вузов. М.: Энергия, 1972.

Гудмен Дж. Введение в фурье-оптику: Пер. с англ. М.: Мир, 1970.

Дулин В.Н. Электронные приборы: Учеб. для вузов. М.: Энергия, 1977.

Ефимов И.Е., Козырь И.Я., Горбунов Ю.И. Микроэлектроника. Физические и технологические основы, надежность: Учеб. пособие для вузов. М.: Высш. шк., 1986.

Жигарев А.А. Электронная оптика и электронно-лучевые приборы: Учеб. для вузов. М.: Высш. шк., 1972.

Исаченко В.П., Осипова В.А., Сукомел А.С. Теплопередача: Учеб. для вузов. М.: Энергоиздат, 1981.

Ключников А.Д., Иванцов Г.П. Теплопередача излучением в огнетехнических установках. М.: Энергия, 1970.

Котлецов Б.Н. Микроизображения: Оптические методы получения и контроля. Л.: Машиностроение. Ленингр. отд-ние, 1987.

Лыков А.В. Теория теплопроводности. М.: Высш. шк., 1967.

Моро У. Микролитография: В 2 ч.: Пер. с англ. М.: Мир, 1990.

Пипко А.И., Плисковский В.Я., Пенченко Е.А. Конструирование и расчет вакуумных систем. М.: Энергия, 1979.

Пресс В.П. Фотолитография в производстве полупроводниковых приборов. М.: Энергия, 1968.

Розанов Л.Н. Вакуумная техника. Учеб. для высш. школы. 2-е изд. С-Пб.: Унивак, 2000.

Теоретические основы теплотехники. Теплотехнический эксперимент. Справ. / Под ред. В.А. Григорьева, В.М. Зорина. М.: Энергоатомиздат, 1988.

Теплотехника: Учеб. для вузов / Под ред. В.Н. Луканина. М.: Высш. шк., 2000.

Фотолитография и оптика / Под ред. Я.А. Федотова. М.: Сов. радио, 1977.

Электронные, квантовые приборы и микроэлектроника: Учеб. пособие для вузов / Под ред. Н.Д. Федорова. М.: Радио и связь, 1998.

Modelling projection printing of positive photoresisys / F.N. Dill, A.R. Neureuther, J.A. Tuttle, E.J. Walker // IEEE Transactions on Electron Devices. 1977. V. ED-22, N 7. P. 456–467.

Novotny D.B. Measurement of the Separation Distance in Contact and Proximity Lithography // *Journal of Electrochemical Society: Solid State Science and Technology.* 1986. V. 136, N 12. P. 2600–2607.

ОГЛАВЛЕНИЕ

Предисловие.....	3
Введение.....	5
1. Вакуум — рабочая среда для микро- и нанотехнологий.....	8
1.1. Основные положения физики вакуума	8
1.2. Физические процессы при вакуумной откачке	17
Контрольные вопросы и задания	26
2. Теплопередача.....	27
2.1. Теплопередача теплопроводностью	30
2.2. Теплообмен излучением.....	37
2.3. Теория теплопередачи и ее использование для анализа физических процессов различной природы	61
Контрольные вопросы и задания	63
3. Формирование потоков частиц	64
3.1. Формирование электронных потоков.....	66
3.2. Формирование ионных потоков	69
3.3. Формирование атомарных и молекулярных потоков	71
Контрольные вопросы и задания	75
4. Взаимодействие потоков частиц с материалами	75
4.1. Взаимодействие электронных и ионных потоков с материалами ..	76
4.2. Взаимодействие ионных потоков с материалами	84
Контрольные вопросы и задания	94
5. Формирование газоразрядной плазмы и ее взаимодействие с материалами.....	94
5.1. Параметры газоразрядной плазмы	96
5.2. Кривые Пашена	98
5.3. Высокочастотная плазма	100
Контрольные вопросы и задания	101
6. Формирование оптического излучения	101
6.1. Формирование микрорельефа в резонансе.....	102
6.2. Системы экспонирования.....	103

6.3. Основы теории формирования микроизображений.....	109
6.4. Скалярная теория дифракции.....	115
Контрольные вопросы и задания	134
7. Проекционное формирование микроизображений.....	134
7.1. Качество проекционного изображения	135
7.2. Понятие изображающей системы.....	137
7.3. Связь между объектом и изображением	138
7.4. Свертка	140
7.5. Фурье-преобразования в оптике.....	144
7.6. Оптическая передаточная функция.....	151
7.7. Зрачковая функция и ее связь с оптической передаточной функцией.....	156
7.8. Связь комплексной амплитуды изображения со зрачковой функцией.....	158
7.9. Оптическая передаточная функция как автокорреляция зрачковой функции	161
7.10. Системы дифракционного качества с постоянным пропусканием по площади зрачка.....	165
7.11. Учет распределения интенсивности в изображении.....	168
Контрольные вопросы и задания	170
Список рекомендуемой литературы	172

Учебное издание

**Бычков Сергей Павлович
Михайлов Валерий Павлович
Панфилов Юрий Васильевич
Цветков Юрий Борисович**

ФИЗИЧЕСКИЕ ОСНОВЫ МИКРО- И НАНОТЕХНОЛОГИЙ

Редактор *С.А. Серебрякова*
Корректор *Е.В. Авалова*
Компьютерная верстка *С.А. Серебряковой*
Художник *С.С. Водичиц*

Подписано в печать 30.06.2009. Формат 60×84/16.
Усл. печ. л. 10,17. Изд. № 11. Тираж 300 экз. Заказ

Издательство МГТУ им. Н.Э. Баумана.
Тннография МГТУ им. Н.Э. Баумана.
105005, Москва, 2-я Бауманская ул., 5.