

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ

В.А. Зверев

**МЕТОДОЛОГИЯ И ИНСТРУМЕНТАРИЙ
В СОВРЕМЕННОЙ ОПТОТЕХНИКЕ**

Учебное пособие



Санкт-Петербург

2010

Зверев В.А. Методология и инструментарий в современной оптотехнике. Учебное пособие. – СПб: СПбГУ ИТМО, 2010. – 367 с.

Методически материал пособия представлен в виде двух разделов. В первом разделе пособия рассмотрены вопросы, составляющие теоретическую основу общей оптотехники. Рассматривая теорию световых явлений как естественное развитие феноменологической теории электромагнитного поля, получены уравнение волны и основное уравнение геометрической оптики – уравнение эйканала, рассмотрены свойства волны и электромагнитного поля на границе раздела двух однородных изотропных сред. Получены основные соотношения геометрической оптики, рассмотрены теоретические основы интерференции света и типовые схемы интерферометров, получены соотношения, определяющие когерентность света; описаны понятия и приведены основные соотношения фотометрии.

Во втором разделе пособия рассмотрены вопросы применения теоретических основ оптотехники для решения прикладных проблем построения оптических систем и принципиальных схем визуальных оптических приборов. В геометрическом приближении, положив в основу понятие и свойства световой трубки, получены соотношения, определяющие световой поток, проходящий через оптическую систему, и освещённость образованного ею изображения. На основании представлений о функции рассеяния в изображении точки показано, что оптическую систему можно рассматривать как фильтр пространственных частот. Изложены основы теории дифракции и рассмотрена дифракционная и абберационная структуры изображения точки. Рассмотрены вопросы построения принципиальных схем визуальных оптических приборов, приведены основные соотношения, определяющие их функциональное назначение и эффективность применения.

Материал пособия ориентирован на магистрантов оптических специальностей и, прежде всего, по направлению подготовки «Оптотехника», а также может быть полезным для инженерно-технических работников оптического приборостроения.

Учебное пособие подготовлено на кафедре Прикладной и компьютерной оптики Санкт-Петербургского государственного университета информационных технологий, механики и оптики.

Рецензент:



В 2009 году Университет стал победителем многоэтапного конкурса, в результате которого определены 12 ведущих университетов России, которым присвоена категория «Национальный исследовательский университет». Министерством образования и науки Российской Федерации была утверждена Программа развития государственного образовательного учреждения высшего профессионального образования «Санкт-Петербургский государственный университет информационных технологий, механики и оптики» на 2009–2018 годы.

© Санкт-Петербургский государственный университет информационных технологий, механики и оптики, 2010
© В.А. Зверев, 2010

СОДЕРЖАНИЕ

Раздел I. ОСНОВЫ ТЕОРЕТИЧЕСКОЙ ОПТОТЕХНИКИ.....	6
1. Физические явления – экспериментальная основа	
электромагнитной теории света.....	6
1.1. Электромагнитное поле.....	6
1.2. Волновое уравнение	10
1.3. Скалярная форма решения волнового уравнения	12
1.4. Векторная форма решения волнового уравнения	16
1.5. Поляризация световых колебаний	18
1.6. Закон сохранения энергии электромагнитного поля	30
1.7. Электромагнитное поле на границе раздела двух сред	33
1.8. Плоская волна на границе раздела двух однородных изотропных сред.....	37
1.9. Формулы Френеля	40
1.10. Коэффициенты отражения и пропускания света на границе раздела двух сред.....	45
1.11. Поляризация света на поверхности раздела двух сред.....	46
1.12. Полное внутреннее отражение	53
1.13. Двойное лучепреломление света.....	57
2. Основы геометрической оптики.....	62
2.1. Основное уравнение геометрической оптики.....	62
2.2. Световые лучи	64
2.3. Закон преломления	70
2.4. Принцип Ферма.....	73
2.5. Интегральный инвариант Лагранжа-Пуанкаре.....	76
2.6. Конгруэнции лучей.....	77
2.7. Теорема Малюса	78
3. Интерференция света. Интерферометры.....	82
3.1. Элементы теории интерференции света.....	82
3.2. Интерферометр Майкельсона.....	86
3.3. Интерферометр Тваймана-Грина	90
3.4. Интерферометр Физо.....	92
4. Когерентность квазимонохроматического света	97
4.1. Квазимонохроматический свет.....	97
4.2. Корреляционные функции световых пучков лучей	99
4.3. Интерференция квазимонохроматического света	104
5. Основные понятия и соотношения фотометрии	116
5.1. Интенсивность излучения некогерентного источника конечных размеров в геометрическом приближении	117
5.2. Оптическое излучение. Световой поток. Энергетические и эффективные характеристики оптического излучения.....	120
5.3. Фотометрические величины и единицы их измерения.....	125

5.3.1. Сила света	125
5.3.2. Световой поток	127
5.3.3. Освещенность. Светимость	129
5.3.4. Яркость	130
5.3.5. Световая энергия. Экспозиция	132
5.4. Энергетические величины	134
5.4.1. Связь между световыми и энергетическими величинами	134
5.4.2. Световая эффективность	135
5.5. Фотометрические свойства тел	138
5.5.1. Общие соотношения	138
5.5.2. Гладкая поверхность раздела двух диэлектриков	142
5.5.3. Отражение света от гладкой поверхности металла	145
5.5.4. Поглощение света в веществе	158
Раздел II. ОСНОВЫ ПРИКЛАДНОЙ ОПТОТЕХНИКИ.....	162
1. Основы теории геометрического построения изображения предмета системой оптических поверхностей.....	162
1.1. Параксиальная оптика	162
1.2. Оптика нулевых лучей	165
1.3. Кардинальные точки оптической системы	166
1.4. Геометрическое построение изображения и основные оптические формулы	169
2. Светотехника оптических систем	174
2.1. Световая трубка	174
2.2. Инвариант Штраубеля	176
2.3. Ограничение световых пучков лучей в оптической системе	180
2.4. Световой поток, проходящий через оптическую систему от излучающего элемента поверхности, расположенного на оптической оси перпендикулярно к ней	185
2.5. Освещенность в осевой точке изображения	188
2.6. Световой поток, проходящий через оптическую систему от излучающего элемента поверхности, расположенного вне оптической оси	194
2.7. Освещенность элемента изображения, расположенного вне оптической оси системы	198
3. Структура изображения, образованного оптической системой, и оценка его качества	200
3.1. Оптическая система как фильтр пространственных частот	204
3.2. Структура изображения точечного предмета, образованного оптической системой	211

3.2.1. Введение в дифракционную теорию образования изображения	211
3.2.2. Интегральная теорема Гельмгольца-Кирхгофа	220
3.2.3. Теория дифракции Кирхгофа	227
3.2.4. Дифракция Фраунгофера и Френеля	233
3.2.5. Дифракция Фраунгофера на отверстиях разной формы	238
3.2.6. Дифракция Фраунгофера на системе одинаковых и одинаково ориентированных отверстий	249
3.2.7. Дифракционный интеграл при наличии aberrаций	257
4. Геометрическая теория aberrаций	267
4.1. Уравнение луча. Волновая и поперечная aberrации	267
4.2. Функция волновой aberrации. Первичные aberrации	270
4.3. Aberrационные свойства сферической поверхности	277
5. Опотехника приборов визуального наблюдения	287
5.1. Опотехника микроскопа	290
5.1.1. Лупа, простой микроскоп	290
5.1.2. Сложный микроскоп	292
5.1.3. Разрешающая способность оптической системы микроскопа и полезное увеличение образованного ею изображения	297
5.1.4. Анализ влияния освещения предмета в микроскопе на его изображение	299
5.2. Опотехника телескопических приборов	312
5.2.1. Увеличительное действие телескопических систем	312
5.2.2. Светотехника телескопических систем	314
5.2.3. Проницающая сила телескопа	316
Приложение. Математический инструментарий в опотехнике	321
А. Векторное исчисление	321
Б. Разложение в ряд произвольных функций	335
В. Преобразование Фурье	350
Литература	363

Раздел I. ОСНОВЫ ТЕОРЕТИЧЕСКОЙ ОПТОТЕХНИКИ

1. Физические явления – экспериментальная основа электромагнитной теории света

1.1. Электромагнитное поле

Принято считать, что в пространстве задано поле некоторой величины, если в каждой точке пространства (или в некоторой его части) определено значение этой величины. Поле может быть скалярным или векторным в зависимости от характера исследуемой величины, стационарным (установившимся), если оно не меняется с течением времени в каждой точке пространства, или нестационарным (неустановившимся) в противном случае.

Физический смысл уравнений Максвелла можно понять, следуя исторической последовательности развития электродинамики. Французский физик Шарль Огюстен Кулон в 1785 году установил один из основных законов электростатики, определяющий силу взаимодействия между двумя «точечными» электрическими зарядами, в виде:

$$F = k \frac{q_1 q_2}{r^2}, \quad (1.1)$$

где F – сила взаимодействия зарядов; q_1, q_2 – величины взаимодействующих зарядов; r – расстояние между зарядами; k – коэффициент пропорциональности, величина которого определяется выбором системы единиц физических величин, входящих в это выражение; так, например, в системе единиц измерения физических величин СГС (сантиметр, грамм, секунда) коэффициент k полагается равным единице. В соответствии с законом Кулона электрическое поле будет вполне определено, если в каждой точке этого поля определена сила, действующая на помещенный в ней единичный положительный заряд. Эта сила называется напряженностью электрического поля и обычно обозначается буквой \mathbf{E} . При $q_1 = q$, а $q_2 = 1$ имеем

$$\mathbf{E} = k \frac{q}{r^3} \mathbf{r}, \quad (1.2)$$

где \mathbf{r} – радиус-вектор, проведенный из заряда q в рассматриваемую точку поля. Для описания электрических процессов в материальной среде вводится вектор электрической индукции $\mathbf{D} = \mathbf{D}(\mathbf{E})$. Для большинства изотропных сред

$$\mathbf{D} = \varepsilon \mathbf{E}, \quad (1.3)$$

где $\varepsilon = \varepsilon(x, y, z)$ – диэлектрическая проницаемость (для вакуума в системе СГС $\varepsilon = 1$). При этом поток вектора электрической индукции через

произвольную замкнутую поверхность S определяется электрическим зарядом, находящимся в объеме V , ограниченном этой поверхностью, т.е.:

$$\oint_S \mathbf{D} \cdot d\mathbf{S} = 4\pi \int_V \rho dV. \quad (1.4)$$

Здесь $d\mathbf{S} = \mathbf{n}dS$, где \mathbf{n} – единичный вектор внешней нормали к граничной поверхности, а ρ – плотность заряда в пространстве (заряд в единице объема).

Поднеся компас к проводнику с током, датский физик Ханс Кристиан Эрстед в 1820 году обнаружил, что магнитная стрелка устанавливается перпендикулярно проводнику. Из опыта Эрстеда следует, что магнитное поле имеет направленный характер и должно характеризоваться векторной величиной. Основную силовую характеристику магнитного поля назвали магнитной индукцией, которую принято обозначать буквой \mathbf{B} . Для описания магнитных процессов в среде вводится вектор магнитной напряженности $\mathbf{H} = \mathbf{H}(\mathbf{B})$. Для большинства изотропных сред

$$\mathbf{B} = \mu\mathbf{H}, \quad (1.5)$$

где $\mu = \mu(x, y, z)$ – магнитная проницаемость среды (для вакуума в системе СГС $\mu = 1$). Из опыта Эрстеда следует, что силовые линии магнитного поля замкнуты сами на себя. Поэтому для любого магнитного поля поток вектора магнитной индукции через любую замкнутую поверхность равен нулю, т.е.

$$\oint_S \mathbf{B} \cdot d\mathbf{S} = 0. \quad (1.6)$$

Из опыта известно, что для магнитного поля, как и для электрического, справедлив принцип суперпозиции: поле \mathbf{B} , порождаемое несколькими токами, равно векторной сумме полей \mathbf{B}_i , порождаемых каждым током в отдельности, т.е. $\mathbf{B} = \sum_i \mathbf{B}_i$. Пусть замкнутый контур

лежит в плоскости, перпендикулярной току. При этом циркуляция вектора \mathbf{B} по контуру L равна

$$\oint_L \mathbf{B} \cdot d\mathbf{l} = \frac{4\pi}{c} I,$$

где под I следует понимать ток, охватываемый контуром, при этом вектор $d\mathbf{l}$ совпадает с касательной в каждой точке контура в направлении обхода. Если некоторый контур охватывает несколько проводов с током, то в силу принципа суперпозиции имеем

$$\oint_L \mathbf{B} \cdot d\mathbf{l} = \oint_L \left(\sum_K \mathbf{B}_K \right) d\mathbf{l} = \sum_K \oint_L \mathbf{B}_K \cdot d\mathbf{l} = \frac{4\pi}{c} \sum_K I_K.$$

Здесь I_K – алгебраическая величина. Если токи текут во всем пространстве, то $\sum_K I_K = \int_S \mathbf{j} \cdot \mathbf{n} dS$. Этот интеграл берется по произвольной поверхности S , натянутой на контур. Вектор \mathbf{j} определяет плотность тока в той точке, где расположена площадка dS ; \mathbf{n} – орт положительной нормали к этой площадке (т.е. нормали, образующей с направлением обхода контура при вычислении циркуляции правовинтовую систему). При этом

$$\oint_L \mathbf{B} \cdot d\mathbf{l} = \frac{4\pi}{c} \int_S \mathbf{j} \cdot d\mathbf{S}.$$

Для большинства изотропных сред

$$\mathbf{j} = \sigma \mathbf{E} + \mathbf{j}_{СТР}, \quad (1.7)$$

где $\sigma = \sigma(x, y, z)$ – удельная электропроводимость; $\mathbf{j}_{СТР}$ – плотность так называемых сторонних токов, т.е. токов, поддерживаемых любыми силами, кроме сил электрического поля.

Согласно гипотезе Д.К. Максвелла магнитное поле порождается не только токами, текущими в проводнике, но и переменными электрическими полями в диэлектриках или в вакууме. Величина, пропорциональная скорости изменения электрического поля во времени, была названа Максвеллом током смещения. При этом циркуляция вектора магнитной напряженности вдоль замкнутого контура L определяется полным током через произвольную поверхность S , ограниченную рассматриваемым контуром, в виде

$$\oint_L \mathbf{H} \cdot d\mathbf{l} = \frac{4\pi}{c} \int_S \mathbf{j} \cdot d\mathbf{S} + \frac{1}{c} \frac{d}{dt} \int_S \mathbf{D} \cdot d\mathbf{S}. \quad (1.8)$$

В 1826 году немецкий физик Георг Симон Ом установил закон, согласно которому сопротивление R контура определяется выражением

$R = \frac{E}{I}$, где E – электродвижущая сила (ЭДС), равная работе, совершаемой над единичным зарядом при обходе им замкнутого контура; I – сила тока:

$$I = -\frac{dq}{dt}.$$

В 1831 году выдающийся английский физик Майкл Фарадей, основоположник учения об электромагнитном поле, описал ставший теперь классическим эксперимент, открывший новую главу электродинамики. Много раз, повторяя свои опыты в различных вариантах, М. Фарадей пришел к выводу, что при всяком пересечении проводником линий магнитной индукции в последнем появляется ток, названный им индукционным, при этом протекший заряд Δq

пропорционален числу пересеченных силовых линий $\Delta\Phi$ и обратно пропорционален электрическому сопротивлению проводника R , т.е.

$R\Delta q = \frac{1}{c}\Delta\Phi$. Д.К. Максвелл ввел связанный с контуром магнитный поток

$\Phi = \int_S \mathbf{B} \cdot \mathbf{n} dS$, где S – натянутая на контур поверхность. Отождествив

величину $\Delta\Phi$ с приращением потока Φ и учитывая закон Ома, Д.К. Максвелл придал закону электромагнитной индукции М. Фарадея следующий вид:

$$\mathbf{E} = R\mathbf{I} = -R \frac{dq}{dt} = -\frac{1}{c} \frac{d\Phi}{dt} \quad \text{или} \quad \mathbf{E} = -\frac{1}{c} \int_S \frac{\partial \mathbf{B}}{\partial t} \cdot \mathbf{n} dS.$$

Вполне очевидно, что напряженность электрического поля, определяющая индукционный ток в опыте М. Фарадея, неэлектростатического происхождения. Циркуляция этого поля по контуру дает величину ЭДС, индуцируемой в контуре: $\mathbf{E} = \oint_L \mathbf{E} \cdot d\mathbf{l}$.

$$\text{При этом} \quad \oint_L \mathbf{E} \cdot d\mathbf{l} = -\frac{1}{c} \frac{d}{dt} \int_S \mathbf{B} \cdot d\mathbf{S}, \quad (1.9)$$

т.е. циркуляция вектора напряженности электрического поля вдоль замкнутого контура L (ЭДС индукции) определяется скоростью изменения потока вектора магнитной индукции через поверхность S , ограниченную рассматриваемым контуром. Знак «минус» соответствует правилу Ленца для направления индуцированного тока.

Если считать, что вектора электромагнитного поля \mathbf{E} , \mathbf{B} , \mathbf{D} и \mathbf{H} являются непрерывными функциями координат, то, рассматривая потоки векторов \mathbf{B} и \mathbf{D} через поверхности, ограничивающие бесконечно малые объемы, и циркуляции векторов \mathbf{E} и \mathbf{H} по бесконечно малым контурам, можно от интегральных уравнений (1.4), (1.6), (1.8) и (1.9) перейти к системе дифференциальных уравнений Максвелла, характеризующих поле в каждой точке пространства. В соответствии с определением соответствующих функций дифференциальные уравнения Максвелла в рассматриваемом случае принимают вид:

$$\operatorname{div} \mathbf{D} = 4\pi\rho \quad (1.10)$$

$$\operatorname{div} \mathbf{B} = 0 \quad (1.11)$$

$$\operatorname{rot} \mathbf{H} = \frac{4\pi}{c} \mathbf{j} + \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} \quad (1.12)$$

$$\operatorname{rot} \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}. \quad (1.13)$$

Уравнения Максвелла не образуют полной замкнутой системы, позволяющей рассчитывать электромагнитные процессы при наличии

материальной среды. Поэтому их необходимо дополнить соотношениями (1.3), (1.5) и (1.7), которые называют уравнениями состояния или материальными уравнениями.

1.2. Волновое уравнение

Английский математик У. Гамильтон ввел символ

$$\nabla = \mathbf{i} \frac{\partial}{\partial x} + \mathbf{j} \frac{\partial}{\partial y} + \mathbf{k} \frac{\partial}{\partial z},$$

называемый «набла». При своем применении он сохраняет черты как вектора, так и оператора дифференцирования. «Умножение» (т.е. действие) оператора Гамильтона на скаляр f и вектор \mathbf{A} (точнее на скалярное и векторное поля соответственно) производится по следующим естественным правилам

$$\nabla f = \left(\mathbf{i} \frac{\partial}{\partial x} + \mathbf{j} \frac{\partial}{\partial y} + \mathbf{k} \frac{\partial}{\partial z} \right) f = \mathbf{i} \frac{\partial f}{\partial x} + \mathbf{j} \frac{\partial f}{\partial y} + \mathbf{k} \frac{\partial f}{\partial z} = \text{grad} f,$$

$$\nabla \mathbf{A} = \left(\mathbf{i} \frac{\partial}{\partial x} + \mathbf{j} \frac{\partial}{\partial y} + \mathbf{k} \frac{\partial}{\partial z} \right) (\mathbf{i} A_x + \mathbf{j} A_y + \mathbf{k} A_z) = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} = \text{div} \mathbf{A},$$

$$\begin{aligned} \nabla \times \mathbf{A} &= \left(\mathbf{i} \frac{\partial}{\partial x} + \mathbf{j} \frac{\partial}{\partial y} + \mathbf{k} \frac{\partial}{\partial z} \right) \times (\mathbf{i} A_x + \mathbf{j} A_y + \mathbf{k} A_z) = \\ &= \mathbf{i} \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) + \mathbf{j} \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) + \mathbf{k} \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) = \\ &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix} = \text{rot} \mathbf{A}. \end{aligned}$$

Рассмотрим электромагнитное поле в однородной непроводящей ($\sigma = 0$) среде в области поля, не содержащей ни зарядов ($\rho = 0$), ни токов ($j = 0$). При этих условиях, используя оператор набла, уравнения (1.10)–(1.13) удобно представить в виде:

$$\nabla \times \mathbf{H} - \frac{1}{c} \dot{\mathbf{D}} = 0 \quad (1.16)$$

$$\nabla \times \mathbf{E} + \frac{1}{c} \dot{\mathbf{B}} = 0 \quad (1.17)$$

$$\nabla \mathbf{B} = 0 \quad (1.15)$$

$$\nabla \mathbf{D} = 0. \quad (1.14)$$

Точка над буквой означает дифференцирование по времени.

Дифференциальное уравнение, которому должен удовлетворять каждый из векторов поля в отдельности, можно получить путем исключения другого. Подставив соотношение (1.5) в уравнение (1.17), разделив его на μ и векторно умножив на оператор набла, получаем:

$$\nabla \times \left(\frac{1}{\mu} \nabla \times \mathbf{E} \right) + \frac{1}{c} \nabla \times \dot{\mathbf{H}} = 0. \quad (1.18)$$

Дифференцируя по времени уравнение (1.16), имеем

$$\nabla \times \dot{\mathbf{H}} - \frac{1}{c} \ddot{\mathbf{D}} = 0. \quad (1.19)$$

Выражение (1.19) с учетом материального уравнения (1.3) позволяет выражение (1.18) представить в виде

$$\nabla \times \left(\frac{1}{\mu} \nabla \times \mathbf{E} \right) + \frac{\varepsilon}{c^2} \dot{\mathbf{E}} = 0. \quad (1.20)$$

В результате последующих преобразований выражение (1.20) можно привести к виду:

$$\nabla^2 \mathbf{E} - \frac{\varepsilon \mu}{c^2} \ddot{\mathbf{E}} + \nabla(\ln \mu) \times (\nabla \times \mathbf{E}) + \nabla[\mathbf{E} \nabla(\ln \varepsilon)] = 0. \quad (1.21)$$

Аналогично находим

$$\nabla^2 \mathbf{H} - \frac{\varepsilon \mu}{c^2} \ddot{\mathbf{H}} + \nabla(\ln \varepsilon) \times (\nabla \times \mathbf{H}) + \nabla[\mathbf{H} \nabla(\ln \mu)] = 0. \quad (1.22)$$

Здесь $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$. В рассматриваемой части пространства

среду будем считать однородной. В этом случае $\nabla(\ln \varepsilon) = 0$ и $\nabla(\ln \mu) = 0$.

При этом уравнения (1.21) и (1.22) принимают вид:

$$\nabla^2 \mathbf{E} - \frac{\varepsilon \mu}{c^2} \ddot{\mathbf{E}} = 0 \quad (1.23)$$

$$\nabla^2 \mathbf{H} - \frac{\varepsilon \mu}{c^2} \ddot{\mathbf{H}} = 0. \quad (1.24)$$

Уравнение вида (1.23) или (1.24) называется уравнением д'Аламбера или волновым уравнением. Эти уравнения волнового движения определяют существование электромагнитных волн, распространяющихся

со скоростью $V = \frac{c}{\sqrt{\varepsilon \mu}}$. Впервые постоянная c была определена

немецкими физиками Фридрихом Кольраушем и Вильгельмом Вебером в 1856 году из отношения значений емкости конденсатора, измеренных в электростатических и электромагнитных единицах. Оказалось, что она совпадает со скоростью света в вакууме. Используя этот результат, Максвелл развил свою электромагнитную теорию света,

предсказывающую существование электромагнитных волн. Правильность его предсказания была подтверждена знаменитыми экспериментами немецкого физика Генриха Герца в 1888 году и русским физиком Петром Николаевичем Лебедевым, который в 1895 году впервые получил и исследовал миллиметровые электромагнитные волны.

1.3. Скалярная форма решения волнового уравнения

В однородной среде в областях, где отсутствуют заряды и токи, любой из декартовых компонентов $A(\mathbf{r}, t)$ векторов поля \mathbf{E} или \mathbf{H} в соответствии с уравнениями (1.23) и (1.24) удовлетворяет однородному волновому уравнению

$$\frac{\partial^2 A}{\partial t^2} - \frac{c^2}{\epsilon\mu} \nabla^2 A = 0. \quad (1.25)$$

Пусть $\mathbf{r}(x, y, z)$ – радиус-вектор некоторой точки P , а $\mathbf{s}(s_x, s_y, s_z)$ – единичный вектор с фиксированным направлением. При этом любое решение уравнения (1.25) в виде $A = A(\mathbf{r} \cdot \mathbf{s}, t)$ определяет плоскую волну, так как в каждый момент времени величина A постоянна в плоскостях $\mathbf{r} \cdot \mathbf{s} = \text{const}$, которые перпендикулярны к единичному вектору \mathbf{s} .

Удобно выбрать такое положение осей декартовой системы координат ξ, ζ, η , чтобы ось $O\xi$ была бы направлена вдоль орта \mathbf{s} . Тогда в соответствии с рис. 1.1 имеем

$$\mathbf{r} \cdot \mathbf{s} = \zeta$$

$$\text{или } xs_x + ys_y + zs_z = \zeta.$$

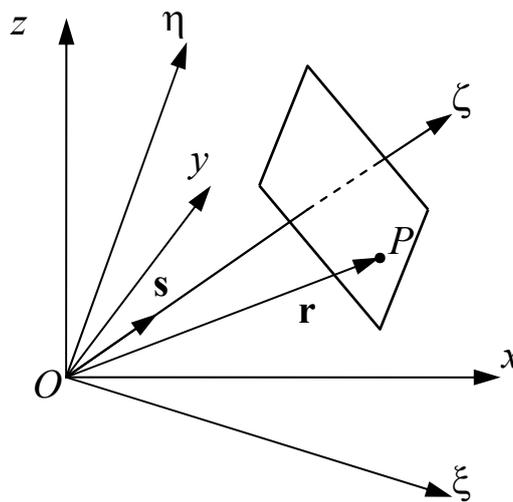


Рис. 1.1. Распространение плоской волны

В системе координат x, y, z имеем

$$\nabla^2 A = \frac{\partial^2 A}{\partial x^2} + \frac{\partial^2 A}{\partial y^2} + \frac{\partial^2 A}{\partial z^2}.$$

Переходя к новой системе координат, получаем

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial \zeta} \frac{\partial \zeta}{\partial x} = \frac{\partial}{\partial \zeta} s_x; \quad \frac{\partial}{\partial y} = \frac{\partial}{\partial \zeta} s_y; \quad \frac{\partial}{\partial z} = \frac{\partial}{\partial \zeta} s_z.$$

При этом

$$\nabla^2 A = \frac{\partial^2 A}{\partial \zeta^2} (s_x^2 + s_y^2 + s_z^2) = \frac{\partial^2 A}{\partial \zeta^2}.$$

Таким образом, в системе координат ξ, ζ, η уравнение (1.25) принимает вид:

$$\frac{\partial^2 A}{\partial t^2} - \frac{c^2}{\epsilon\mu} \frac{\partial^2 A}{\partial \zeta^2} = 0. \quad (1.26)$$

Для решения этого уравнения перепишем его в виде

$$\left(\frac{\partial}{\partial t} - \frac{c}{\sqrt{\epsilon\mu}} \frac{\partial}{\partial \zeta} \right) \left(\frac{\partial}{\partial t} + \frac{c}{\sqrt{\epsilon\mu}} \frac{\partial}{\partial \zeta} \right) A = 0.$$

Введем новые переменные

$$t - \frac{\sqrt{\epsilon\mu}}{c} \zeta = p; \quad t + \frac{\sqrt{\epsilon\mu}}{c} \zeta = q.$$

Отсюда находим

$$t = \frac{1}{2}(q + p), \quad \zeta = \frac{1}{2} \frac{c}{\sqrt{\epsilon\mu}} (q - p).$$

При этом

$$\frac{\partial}{\partial p} = \frac{\partial}{\partial t} \frac{\partial t}{\partial p} + \frac{\partial}{\partial \zeta} \frac{\partial \zeta}{\partial p} = \frac{1}{2} \left(\frac{\partial}{\partial t} - \frac{c}{\sqrt{\epsilon\mu}} \frac{\partial}{\partial \zeta} \right),$$

$$\frac{\partial}{\partial q} = \frac{\partial}{\partial t} \frac{\partial t}{\partial q} + \frac{\partial}{\partial \zeta} \frac{\partial \zeta}{\partial q} = \frac{1}{2} \left(\frac{\partial}{\partial t} + \frac{c}{\sqrt{\epsilon\mu}} \frac{\partial}{\partial \zeta} \right).$$

Применив эти соотношения к функции A , получаем

$$\frac{\partial^2 A}{\partial p \partial q} = \frac{1}{4} \left(\frac{\partial^2 A}{\partial t^2} - \frac{c^2}{\epsilon\mu} \frac{\partial^2 A}{\partial \zeta^2} \right) = 0.$$

Отсюда следует, что общее решение уравнения (1.26) имеет вид:

$$A = A_1(p) + A_2(q),$$

где A_1 и A_2 – произвольные функции, т.е. любая функция от аргумента

$t \pm \frac{\sqrt{\epsilon\mu}}{c} \zeta$ будет решением уравнения (1.26).

Пусть, например, $A_2 = 0$. При этом $A = A_1 \left(t - \frac{\sqrt{\epsilon\mu}}{c} \zeta \right)$. Легко видеть,

что в каждой плоскости $\zeta = const$ при изменении времени изменяется и поле. С другой стороны, в каждый момент времени поле различно для различных значений координаты ζ . Вполне очевидно, что значение поля будет одним и тем же для сочетаний координаты ζ и момента времени t , удовлетворяющих условию:

$$\frac{c}{\sqrt{\epsilon\mu}} t - \zeta = const. \quad (1.27)$$

Положив в этом условии $t = 0$, получаем $-\zeta = const = -\zeta_0$. Тогда условие (1.27) можно записать в виде:

$$\frac{c}{\sqrt{\epsilon\mu}} t = \zeta - \zeta_0.$$

Это значит, что если при $t = 0$ в некоторой точке $\zeta = \zeta_0$ пространства поле имело некоторое фиксированное значение, то через промежуток времени t то же самое значение поля будет на расстоянии, равном $\frac{c}{\sqrt{\epsilon\mu}} t$

от начальной точки в направлении оси $O\zeta$. Отсюда следует, что электромагнитное поле распространяется в пространстве, например, вдоль оси $O\zeta$, со скоростью $V = \frac{c}{\sqrt{\epsilon\mu}} = \frac{c}{n}$, где $\sqrt{\epsilon\mu} = n$ – показатель

преломления среды. Таким образом, функция $A_1 \left(t - \frac{1}{V} \zeta \right)$ определяет плоскую волну, бегущую в положительном направлении оси $O\zeta$.

Очевидно, что функция $A_2 \left(t + \frac{1}{V} \zeta \right)$ определяет плоскую волну, бегущую в противоположном, отрицательном, направлении оси $O\zeta$. В точке $\mathbf{r} = \mathbf{r}_0$ пространства возмущение, вызванное волной, зависит только от времени: $A(\mathbf{r}_0, t) = f(t)$.

Важным частным случаем решения волнового уравнения (1.26) является простая периодическая функция времени. Волна, описываемая такой функцией, называется монохроматической. Зависимость компонентов поля от времени в монохроматической волне определяется множителем: $\cos(\omega t + \delta)$, где ω – угловая частота волны. При этом функцию $f(t)$ можно представить в виде:

$$f(t) = a \cos(\omega t + \delta). \quad (1.28)$$

Величина a ($a > 0$) называется амплитудой, а аргумент косинуса, т.е. $\omega t + \delta$, называется фазой светового колебания. Величина $v = \frac{\omega}{2\pi} = \frac{1}{T}$ называется частотой колебаний. При замене величины t на $t + T$ значение функции $f(t)$ остается неизменным. Отсюда следует, что промежуток времени $t = T$ определяет период колебаний. Волновые функции в форме (1.28) называют гармоническими функциями относительно времени.

Заменим в выражении (1.28) величину t величиной $t - \frac{\mathbf{r} \cdot \mathbf{s}}{V}$. При этом получаем

$$A(\mathbf{r}, t) = a \cos \left[\omega \left(t - \frac{\mathbf{r} \cdot \mathbf{s}}{V} \right) + \delta \right]. \quad (1.29)$$

Полученная волновая функция определяет гармоническую плоскую волну, распространяющуюся в направлении, заданном единичным вектором \mathbf{s} . Расстояние, на которое перемещается поверхность волны за время T , равное периоду одного колебания, определяет период изменения напряженности поля в пространстве, равный $\lambda = VT = V \frac{2\pi}{\omega}$. Расстояние, равное λ , называется длиной волны, при этом приведенная к вакууму длина волны $\lambda_0 = cT = n\lambda$. Эта длина волны соответствует длине распространяющейся в вакууме гармонической волны той же частоты. Легко убедиться, что величина $A(\mathbf{r}, t)$, определяемая выражением (1.29), не изменится, если величину $\mathbf{r} \cdot \mathbf{s}$ заменить величиной $\mathbf{r} \cdot \mathbf{s} + \lambda$. Выражение (1.29) можно представить в виде:

$$A(\mathbf{r}, t) = a \cos[\omega t - \mathbf{k} \cdot \mathbf{r} + \delta], \quad (1.30)$$

где $\mathbf{k} = k\mathbf{s}$ называется волновым вектором или вектором распространения волны в среде, при этом волновое число $k = \frac{\omega}{V} = \frac{2\pi}{\lambda} = n \frac{2\pi}{\lambda_0} = nk_0$.

Амплитуда a и начальная фаза δ плоской монохроматической волны не зависят от \mathbf{r} и t , т.е. одинаковы во всем пространстве во все моменты времени, а, следовательно, выражение (1.30) определяет однородную волну. Никакие реальные волны этим свойством не обладают. Плоскую монохроматическую волну можно рассматривать как частный случай гармонических волн более сложной формы, в общем виде определяемых уравнением

$$A(\mathbf{r}, t) = a(\mathbf{r}) \cos[\omega t - g(\mathbf{r})], \quad (1.31)$$

где $a > 0$ и g – вещественные скалярные функции положения. Поверхности $g(\mathbf{r}) = const$ называют поверхностями постоянной фазы или волновыми поверхностями.

Расчеты, связанные с гармоническими волнами, упрощаются, если использовать экспоненциальные функции вместо тригонометрических. При этом уравнение (1.31) можно записать в виде

$$\begin{aligned} A(\mathbf{r}, t) &= \operatorname{Re}\{U(\mathbf{r})\exp(-i\omega t)\} = \\ &= \frac{1}{2}[U(\mathbf{r})\exp(-i\omega t) + U^*(\mathbf{r})\exp(i\omega t)], \end{aligned} \quad (1.32)$$

где $U(\mathbf{r})$ – комплексная функция вида: $U(\mathbf{r}) = a(\mathbf{r})\exp[ig(\mathbf{r})]$, а символ Re означает, что берется вещественная часть. Если операции, производимые над функцией A , линейны, то в выражении (1.32) символ Re можно опустить и оперировать непосредственно с комплексной функцией вида:

$$A(\mathbf{r}, t) = U(\mathbf{r})\exp(-i\omega t). \quad (1.33)$$

Используя выражение (1.33), находим, что $\frac{\partial^2 A}{\partial t^2} = -\omega^2 U(\mathbf{r})\exp(-i\omega t)$,

а $\nabla^2 A = \exp(-i\omega t)\nabla^2 U(\mathbf{r})$. При этом уравнение (1.25) принимает вид:

$$\nabla^2 U + n^2 k_0^2 U = 0. \quad (1.34)$$

Полученное уравнение называют уравнением Гельмгольца. Величину U называют комплексной амплитудой волны. В частности, для плоской волны имеем

$$g(\mathbf{r}) = \omega \frac{\mathbf{r} \cdot \mathbf{s}}{V} - \delta = k \mathbf{r} \cdot \mathbf{s} - \delta = \mathbf{k} \cdot \mathbf{r} - \delta.$$

При этом $U(\mathbf{r}) = a(\mathbf{r})\exp(i\mathbf{k} \cdot \mathbf{r})$, где $a(\mathbf{r}) = a_0(\mathbf{r})\exp(-i\delta)$.

1.4. Векторная форма решения волнового уравнения

Простейшим электромагнитным полем является поле плоской волны. Электромагнитную плоскую волну общего вида определим векторами

$$\mathbf{E} = \mathbf{E}(\psi), \quad \mathbf{H} = \mathbf{H}(\psi), \quad \text{где } \psi = t - \frac{\mathbf{r} \cdot \mathbf{s}}{V}.$$

При этом

$$\begin{aligned} \dot{\mathbf{E}} &= \frac{\partial \mathbf{E}}{\partial t} = \frac{\partial \mathbf{E}}{\partial \psi} \frac{\partial \psi}{\partial t} = \frac{\partial \mathbf{E}}{\partial \psi}, \\ \nabla \mathbf{E} &= \mathbf{i} \cdot \frac{\partial \mathbf{E}}{\partial \psi} \frac{\partial \psi}{\partial x} + \mathbf{j} \cdot \frac{\partial \mathbf{E}}{\partial \psi} \frac{\partial \psi}{\partial y} + \mathbf{k} \cdot \frac{\partial \mathbf{E}}{\partial \psi} \frac{\partial \psi}{\partial z} = \\ &= -\frac{1}{V} (\mathbf{i}s_x + \mathbf{j}s_y + \mathbf{k}s_z) \frac{\partial \mathbf{E}}{\partial \psi} = -\frac{1}{V} \mathbf{s} \cdot \frac{\partial \mathbf{E}}{\partial \psi}, \end{aligned}$$

$$\begin{aligned}
(\nabla \times \mathbf{E})_x &= \frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = \frac{\partial E_z}{\partial \psi} \frac{\partial \psi}{\partial y} - \frac{\partial E_y}{\partial \psi} \frac{\partial \psi}{\partial z} = \\
&= -\frac{1}{V} \frac{\partial E_z}{\partial \psi} s_y + \frac{1}{V} \frac{\partial E_y}{\partial \psi} s_z = \frac{1}{V} \left(\frac{\partial \mathbf{E}}{\partial \psi} \times \mathbf{s} \right)_x;
\end{aligned}$$

вполне очевидно, что $(\nabla \times \mathbf{E})_y = \frac{1}{V} \left(\frac{\partial \mathbf{E}}{\partial \psi} \times \mathbf{s} \right)_y$, $(\nabla \times \mathbf{E})_z = \frac{1}{V} \left(\frac{\partial \mathbf{E}}{\partial \psi} \times \mathbf{s} \right)_z$;

при этом $\nabla \times \mathbf{E} = \frac{1}{V} \frac{\partial \mathbf{E}}{\partial \psi} \times \mathbf{s}$. Аналогично получаем

$$\dot{\mathbf{H}} = \frac{\partial \mathbf{H}}{\partial \psi}, \quad \nabla \mathbf{H} = -\frac{1}{V} \mathbf{s} \cdot \frac{\partial \mathbf{H}}{\partial \psi}, \quad \nabla \times \mathbf{H} = \frac{1}{V} \frac{\partial \mathbf{H}}{\partial \psi} \times \mathbf{s}.$$

Подставив полученные выражения в уравнения Максвелла (1.10)–(1.13), с учетом материальных уравнений (1.3) и (1.5) преобразуем их к виду:

$$\mathbf{s} \cdot \frac{\partial \mathbf{E}}{\partial \psi} = 0, \tag{1.35}$$

$$\mathbf{s} \cdot \frac{\partial \mathbf{H}}{\partial \psi} = 0, \tag{1.36}$$

$$\frac{\partial \mathbf{H}}{\partial \psi} \times \mathbf{s} - \varepsilon \frac{V}{c} \frac{\partial \mathbf{E}}{\partial \psi} = 0, \tag{1.37}$$

$$\frac{\partial \mathbf{E}}{\partial \psi} \times \mathbf{s} + \mu \frac{V}{c} \frac{\partial \mathbf{H}}{\partial \psi} = 0. \tag{1.38}$$

Полагая постоянную интегрирования равной нулю (т.е. пренебрегая постоянным полем во всем пространстве) и учитывая, что $\frac{c}{V} = \sqrt{\varepsilon \mu} = n$, в результате интегрирования выражений (1.35)–(1.38) получаем

$$\mathbf{s} \cdot \mathbf{E} = 0, \tag{1.39}$$

$$\mathbf{s} \cdot \mathbf{H} = 0, \tag{1.40}$$

$$\mathbf{E} = \sqrt{\frac{\mu}{\varepsilon}} (\mathbf{H} \times \mathbf{s}), \tag{1.41}$$

$$\mathbf{H} = -\sqrt{\frac{\varepsilon}{\mu}} (\mathbf{E} \times \mathbf{s}). \tag{1.42}$$

Из вида равенств (1.39) и (1.40) следует «поперечность» электромагнитного поля, поскольку векторы электрической и магнитной напряженностей поля лежат в плоскости, перпендикулярной к направлению распространения поля. Как следует из выражений (1.41) и (1.42), векторы \mathbf{E} и \mathbf{H} ортогональны друг другу и образуют вместе с ортом \mathbf{s} правую систему ортогональных векторов, как показано на рис. 1.2.

Кроме того, из выражений (1.41) и (1.42) следует, что $E = \sqrt{\frac{\mu}{\epsilon}} H$,

$$H = \sqrt{\frac{\epsilon}{\mu}} E.$$

В результате получаем

$$\sqrt{\epsilon} E = \sqrt{\mu} H. \quad (1.43)$$

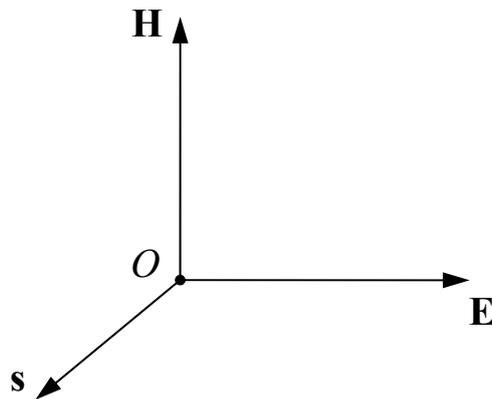


Рис.1.2. Взаимное расположение векторов напряженности электрического **E** и магнитного **H** полей и вектора **s** направления распространения электромагнитной волны

1.5. Поляризация световых колебаний

Направим ось z системы координат x, y, z вдоль волнового вектора $\mathbf{k} = ks$. Тогда у векторов **E** и **H** могут быть отличны от нуля только проекции на оси x и y . Уравнения Максвелла допускают, в частности, такое решение, при котором у вектора **E** во всех точках в пространстве и во все моменты времени отлична от нуля только одна проекция, например, $E_x(z, t)$.

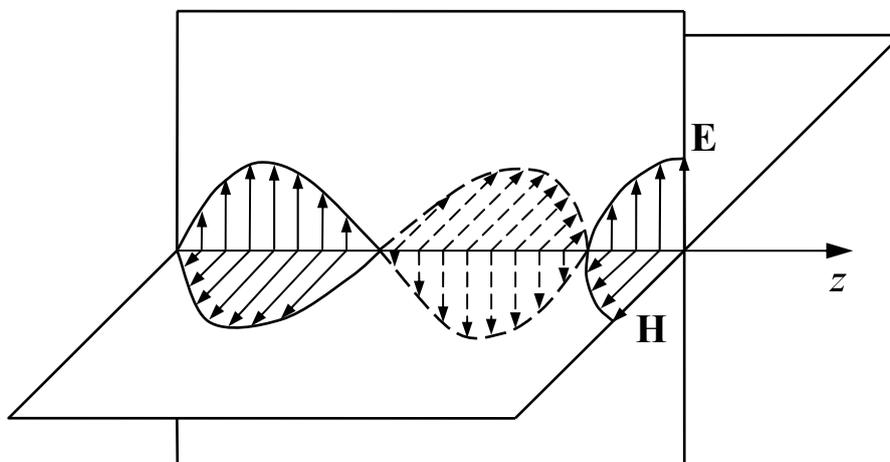


Рис. 1.3. Электромагнитная волна

Вследствие свойства поперечности у вектора \mathbf{H} при этом отлична от нуля только проекция на ось y , т.е. $H_y(z, t)$. Эти проекции связаны между собой соотношением $\sqrt{\varepsilon}E_x = \sqrt{\mu}H_y$. Мгновенный «снимок» такой волны, показывающий векторы \mathbf{E} и \mathbf{H} в разных точках оси z в один момент времени, показан на рис. 1.3. В этом случае говорят, что волна имеет линейную (или плоскую) поляризацию. Плоскость, в которой лежит вектор напряженности электрического поля волны \mathbf{E} и волновой вектор \mathbf{k} , называют плоскостью поляризации или плоскостью колебаний. Чтобы представить себе изменение электрического и магнитного полей с течением времени, можно считать, что вся система векторов, показанная на рис. 1.3, движется как целое вдоль оси z со скоростью V (в вакууме со скоростью c).

Свет, в котором представлены все направления колебаний вектора \mathbf{E} в плоскости, перпендикулярной к направлению распространения, называется естественным светом. Излучение обычных источников света не поляризовано. Линейно поляризованный свет получают, пропуская естественный через оптические поляризаторы, действие которых основано на различных физических принципах. С их помощью можно не только получить линейно поляризованный свет, но и выяснить, имеет ли исследуемое излучение линейную поляризацию. Выполняющее такую роль поляризационное устройство называют анализатором. В отличие от обычных источников света излучение газового лазера, окна разрядной трубки которого наклонены на некоторый угол к ее оптической оси (угол Брюстера), обладает линейной поляризацией.

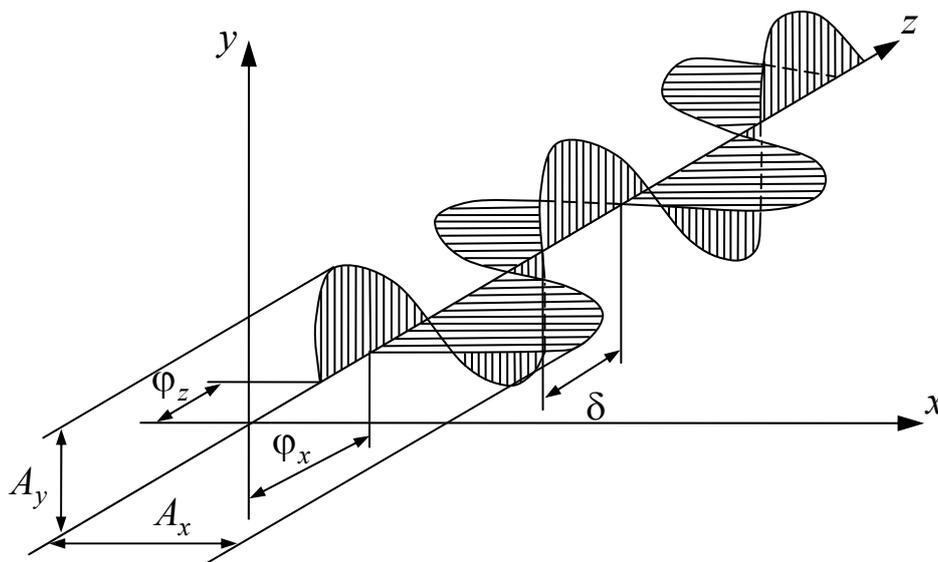


Рис.1.4. Сложение двух электромагнитных волн

В общем случае вектор \mathbf{E} может иметь любые направления в плоскости, перпендикулярной к направлению распространения волны. Пусть наряду с волной, поляризованной в плоскости xOz , в том же

направлении (в направлении оси z) распространяется другая волна той же частоты ω , но поляризованная в плоскости yOz , как показано на рис. 1.4. Вследствие линейности уравнений Максвелла такое наложение (суперпозиция) волн также является решением уравнений. В зависимости от разности фаз складываемых линейно поляризованных волн результирующая волна может иметь различную поляризацию.

Рассмотрим электрическое поле $\mathbf{E}(z, t)$ волны, возникающей при сложении двух волн одинаковой частоты с ортогональными направлениями линейной поляризации:

$$E_x(z, t) = E_{0x} \exp[i(kz - \omega t)] = A_x \exp(-i\varphi_x) \exp[i(kz - \omega t)] \quad (1.44)$$

$$E_y(z, t) = E_{0y} \exp[i(kz - \omega t)] = A_y \exp(-i\varphi_y) \exp[i(kz - \omega t)]. \quad (1.45)$$

При одинаковых (или отличающихся на $n\pi$, где n – целое число) фазах φ_x и φ_y комплексных амплитуд E_{0x} и E_{0y} в каждой точке происходит сложение взаимно перпендикулярных колебаний в одной фазе, что дает колебание в новом направлении. Результирующая волна будет линейно поляризованной. Направление ее поляризации зависит от отношения амплитуд A_x и A_y , как показано на рис. 1.5.

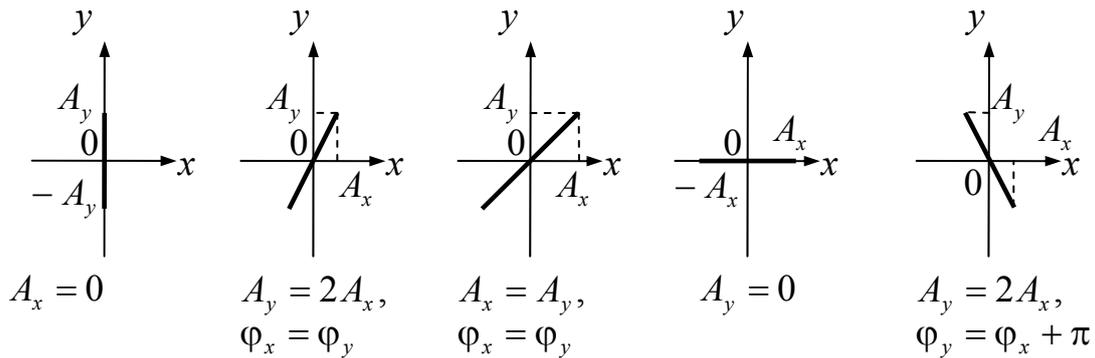


Рис. 1.5. Зависимость направления линейно поляризованной волны от отношения амплитуд складываемых волн

Пусть волна, поляризованная в плоскости yOz , отстает по фазе на $\frac{1}{2}\pi$ от волны, поляризованной в плоскости xOz , т.е. $\varphi_x - \varphi_y = \frac{1}{2}\pi$. Если амплитуды этих волн одинаковы ($A_x = A_y = A$), то вектор \mathbf{E} в любой точке z будет вращаться в плоскости xOy против часовой стрелки, оставаясь неизменным по модулю. Например, в точке $z = 0$ имеем

$$E_x(t) = A \cos(\omega t + \varphi_x) \quad (1.46)$$

$$E_y(t) = A \cos\left(\omega t + \varphi_x - \frac{1}{2}\pi\right) = A \sin(\omega t + \varphi_x). \quad (1.47)$$

Такую волну называют поляризованной по кругу или циркулярно поляризованной. Когда при наблюдении навстречу волне вращение

вектора \mathbf{E} в фиксированной плоскости происходит так, как в рассмотренном примере при $z = 0$, т.е. против часовой стрелки, то говорят о волне левой круговой поляризации. Правая круговая поляризация соответствует вращению вектора в фиксированной плоскости $z = const$ в направлении часовой стрелки:

$$E_x(t) = A \cos(\omega t + \varphi_x) \quad (1.48)$$

$$E_y(t) = A \cos\left(\omega t + \varphi_x + \frac{1}{2}\pi\right) = -A \sin(\omega t + \varphi_x), \quad (1.49)$$

т.е. волна $E_y(t)$ опережает по фазе волну $E_x(t)$ на $\frac{1}{2}\pi$.

На практике для превращения линейно поляризованного света в свет с круговой поляризацией используют анизотропные кристаллические пластинки, в которых две волны с ортогональными направлениями линейной поляризации имеют различные фазовые скорости. Подбором толщины пластинки можно получить на выходе заданную разность фаз этих волн и тем самым требуемое состояние поляризации.

При использовании комплексной записи в виде (1.44) и (1.45) для складываемых волн с ортогональными направлениями линейной поляризации результирующая волна имеет линейную поляризацию, если отношение комплексных амплитуд E_{0y} к E_{0x} выражается вещественным числом. Результирующая волна имеет круговую поляризацию, если отношение комплексных амплитуд определяется мнимым числом, по

модулю равным единице. При $\frac{E_{0y}}{E_{0x}} = \exp\left(i\frac{\pi}{2}\right) = i$, волна будет иметь

правую круговую поляризацию, а при $\frac{E_{0y}}{E_{0x}} = -i$ – левую. В общем случае

при наложении световых колебаний во взаимно перпендикулярных плоскостях кривая, которую описывает конец вектора \mathbf{E} в произвольной фиксированной плоскости $z = const$, является геометрическим местом точек, координаты которых определяются уравнениями вида:

$$E_x = A_x \cos(\tau + \varphi_x) \quad (1.50)$$

$$E_y = A_y \cos(\tau + \varphi_y) \quad (1.51)$$

$$E_z = 0. \quad (1.52)$$

Здесь $\tau = kz - \omega t$.

Для того, чтобы исключить τ из уравнений (1.50) и (1.51), перепишем их в виде:

$$\frac{E_x}{A_x} = \cos \tau \cos \varphi_x - \sin \tau \sin \varphi_x \quad (1.53)$$

$$\frac{E_y}{A_y} = \cos \tau \cos \varphi_y - \sin \tau \sin \varphi_y. \quad (1.54)$$

При этом

$$\frac{E_x}{A_x} \sin \varphi_y - \frac{E_y}{A_y} \sin \varphi_x = \cos \tau \sin(\varphi_y - \varphi_x),$$

$$\frac{E_x}{A_x} \cos \varphi_y - \frac{E_y}{A_y} \cos \varphi_x = \sin \tau \sin(\varphi_y - \varphi_x).$$

Возведя в квадрат левую и правую части этих уравнений и сложив полученные выражения соответствующих частей, имеем

$$\left(\frac{E_x}{A_x}\right)^2 + \left(\frac{E_y}{A_y}\right)^2 - 2\frac{E_x E_y}{A_x A_y} \cos \delta = \sin^2 \delta, \quad (1.55)$$

где $\delta = \varphi_y - \varphi_x$. Выражение (1.55) представляет собой уравнение конического сечения. Известно, что квадратичная форма $Ax^2 + 2Bxy + Cy^2 = H$ будет эллиптической, если дискриминант формы $D = AC - B^2 > 0$. В рассматриваемом случае соответствующий дискриминант равен

$$D = \frac{1}{A_x^2 A_y^2} - \frac{\cos^2 \delta}{A_x^2 A_y^2} = \frac{\sin^2 \delta}{A_x^2 A_y^2} > 0.$$

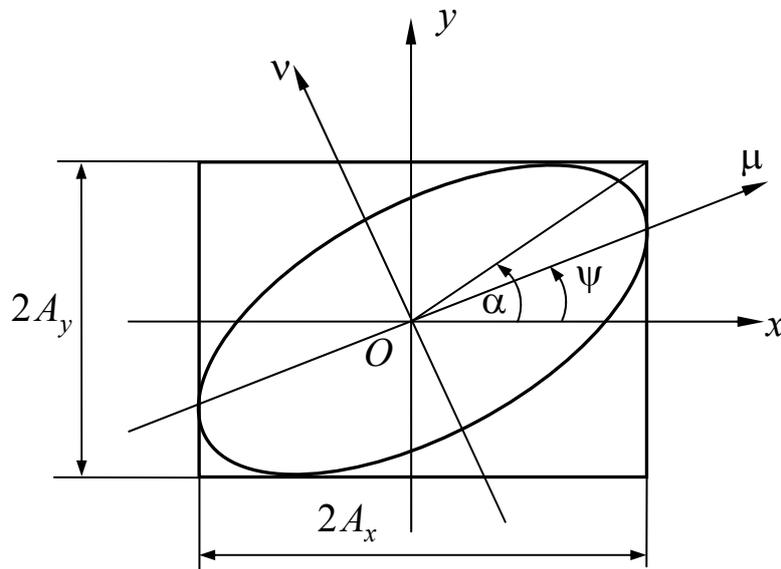


Рис. 1.6. Эллиптически поляризованная волна

Таким образом, уравнение (1.55) описывает эллипс, вписанный в прямоугольник, стороны которого параллельны осям координат Ox и Oy и имеют длины $2A_x$ и $2A_y$, как показано на рис. 1.6. Положив в уравнении

(1.55) величину $E_x = A_x$, получаем $E_y = A_y \cos \delta$, а при $E_y = A_y$ получаем $E_x = A_x \cos \delta$, т.е. эллипс касается сторон прямоугольника в точках $(\pm A_x, \pm A_y \cos \delta)$ и $(\pm A_x \cos \delta, \pm A_y)$. В этом случае говорят, что волна, описываемая уравнениями (1.50)–(1.52), эллиптически поляризована. В общем случае оси эллипса не параллельны осям Ox и Oy .

Пусть $O\mu$ и $O\nu$ – новые оси, совпадающие с осями эллипса, а ψ ($0 \leq \psi < \pi$) – угол между осью Ox и направлением главной оси $O\mu$, как показано на рисунке. При этом компоненты E_μ и E_ν связаны с компонентами E_x и E_y соотношениями

$$E_\mu = E_x \cos \psi + E_y \sin \psi \quad (1.56)$$

$$E_\nu = -E_x \sin \psi + E_y \cos \psi. \quad (1.57)$$

Если $2a$ и $2b$ ($a \geq b$) – длины осей эллипса, то уравнение эллипса относительно осей $O\mu$ и $O\nu$ можно записать в виде:

$$E_\mu = a \cos(\tau + \varphi_0) \quad (1.58)$$

$$E_\nu = b \sin(\tau + \varphi_0). \quad (1.59)$$

Чтобы определить амплитуды a и b , приравняем в развернутом виде правые части уравнений (1.56) и (1.58) и уравнений (1.57) и (1.59). При этом, учитывая выражения (1.50) и (1.51), получаем

$$a(\cos \tau \cos \varphi_0 - \sin \tau \sin \varphi_0) = \cos \tau (A_x \cos \varphi_x \cos \psi + A_y \cos \varphi_y \sin \psi) - \sin \tau (A_x \sin \varphi_x \cos \psi + A_y \sin \varphi_y \sin \psi),$$

$$b(\sin \tau \cos \varphi_0 + \cos \tau \sin \varphi_0) = \cos \tau (A_y \cos \varphi_y \cos \psi - A_x \cos \varphi_x \sin \psi) + \sin \tau (A_x \sin \varphi_x \sin \psi - A_y \sin \varphi_y \cos \psi).$$

Приравнивая коэффициенты при $\cos \tau$ и при $\sin \tau$, имеем

$$a \cos \varphi_0 = A_x \cos \varphi_x \cos \psi + A_y \cos \varphi_y \sin \psi, \quad (1.60)$$

$$a \sin \varphi_0 = A_x \sin \varphi_x \cos \psi + A_y \sin \varphi_y \sin \psi; \quad (1.61)$$

$$b \cos \varphi_0 = A_x \sin \varphi_x \sin \psi - A_y \sin \varphi_y \cos \psi, \quad (1.62)$$

$$b \sin \varphi_0 = -A_x \cos \varphi_x \sin \psi + A_y \cos \varphi_y \cos \psi. \quad (1.63)$$

Возведя левые и правые части этих уравнений в квадрат и последовательно складывая полученные части уравнений (1.60) и (1.61) и уравнений (1.62) и (1.63), находим, что

$$a^2 = A_x^2 \cos^2 \psi + A_y^2 \sin^2 \psi - A_x A_y \sin 2\psi \cos \delta,$$

$$b^2 = A_x^2 \sin^2 \psi + A_y^2 \cos^2 \psi - A_x A_y \sin 2\psi \cos \delta,$$

где $\delta = \varphi_y - \varphi_x$. Отсюда следует, что

$$a^2 + b^2 = A_x^2 + A_y^2. \quad (1.64)$$

Перемножив левые и правые части уравнений (1.60) и (1.62) и уравнений (1.61) и (1.63) и сложив соответствующие части полученных выражений, имеем

$$\pm ab = A_x A_y \sin \delta. \quad (1.65)$$

Наличие двух знаков в уравнении (1.65) определяет возможность двух направлений вращения вектора \mathbf{E} , конец которого описывает эллипс.

Поделив левые и правые части уравнений (1.62) и (1.63) соответственно на левые и правые части уравнений (1.60) и (1.61), получаем

$$\begin{aligned} \pm \frac{b}{a} &= \frac{A_x \sin \varphi_x \sin \psi - A_y \sin \varphi_y \cos \psi}{A_x \cos \varphi_x \cos \psi + A_y \cos \varphi_y \sin \psi} = \\ &= \frac{-A_x \cos \varphi_x \sin \psi + A_y \cos \varphi_y \cos \psi}{A_x \sin \varphi_x \cos \psi + A_y \sin \varphi_y \sin \psi}. \end{aligned}$$

Отсюда следует, что

$$\begin{aligned} &A_x^2 \sin^2 \varphi_x \sin \psi \cos \psi + A_x A_y \sin \varphi_x \sin \varphi_y \sin^2 \psi - \\ &- A_x A_y \sin \varphi_x \sin \varphi_y \cos^2 \psi - A_y^2 \sin^2 \varphi_y \sin \psi \cos \psi = \\ &= -A_x^2 \cos^2 \varphi_x \sin \psi \cos \psi + A_x A_y \cos \varphi_x \cos \varphi_y \cos^2 \psi - \\ &- A_x A_y \cos \varphi_x \cos \varphi_y \sin^2 \psi + A_y^2 \cos^2 \varphi_y \sin \psi \cos \psi. \end{aligned}$$

Преобразовав это выражение, получаем

$$(A_x^2 - A_y^2) \sin 2\psi = 2A_x A_y \cos \delta \cos 2\psi.$$

Отсюда следует, что

$$\operatorname{tg} 2\psi = \frac{2A_x A_y}{A_x^2 - A_y^2} \cos \delta. \quad (1.66)$$

Обозначим $\frac{A_y}{A_x} = \operatorname{tg} \alpha$, где $0 \leq \alpha \leq \frac{\pi}{2}$.

При этом уравнение (1.66) можно представить в виде:

$$\operatorname{tg} 2\psi = \frac{2 \operatorname{tg} \alpha}{1 - \operatorname{tg}^2 \alpha} \cos \delta = \frac{2 \sin \alpha \cos \alpha}{\cos^2 \alpha - \sin^2 \alpha} \cos \delta$$

или

$$\operatorname{tg} 2\psi = \operatorname{tg} 2\alpha \cos \delta. \quad (1.67)$$

Из выражений (1.64) и (1.65) следует, что

$$\pm \frac{2ab}{a^2 + b^2} = \frac{2A_x A_y}{A_x^2 + A_y^2} \sin \delta = \frac{2 \operatorname{tg} \alpha}{1 + \operatorname{tg}^2 \alpha} \sin \delta = 2 \sin \alpha \cos \alpha \sin \delta$$

или

$$\pm \frac{2ab}{a^2 + b^2} = \sin 2\alpha \sin \delta. \quad (1.68)$$

Обозначим $\pm \frac{b}{a} = \operatorname{tg} \chi$, где $\chi \left(-\frac{\pi}{4} \leq \chi \leq \frac{\pi}{4} \right)$ – вспомогательный угол, при этом численное значение $\operatorname{tg} \chi$ определяет величину отношения осей эллипса, а знак при χ характеризует два варианта, которые можно использовать при описании эллипса. При этом выражение (1.68) можно представить в виде:

$$\sin 2\chi = \sin 2\alpha \sin \delta. \quad (1.69)$$

Итак, если заданы величины A_x и A_y , относящиеся к произвольному положению осей, разность фаз δ и если $0 \leq \alpha \leq \frac{\pi}{2}$, то главные полуоси эллипса a и b и угол ψ ($0 \leq \psi < \pi$), который большая ось эллипса образует с осью Ox , определяются формулами:

$$a^2 + b^2 = A_x^2 + A_y^2 \quad (1.70)$$

$$\operatorname{tg} 2\psi = \operatorname{tg} 2\alpha \cos \delta \quad (1.71)$$

$$\sin 2\chi = \sin 2\alpha \sin \delta. \quad (1.72)$$

С другой стороны, если известны длины осей a и b и ориентация эллипса (т.е. заданы величины a , b и ψ), то эти формулы позволяют найти амплитуды A_x и A_y и разность фаз δ . Существуют оптические устройства, которые позволяют определять эти величины прямым способом.

В предельных случаях эллипс поляризации вырождается либо в прямую, либо в окружность. Согласно выражению (1.55) эллипс вырождается в прямую при $\delta = \varphi_y - \varphi_x = m\pi$, где $m = 0, \pm 1, \pm 2, \dots$. При

$$\text{этом } \frac{E_y}{E_x} = (-1)^m \frac{A_y}{A_x}.$$

В этом случае имеем линейно поляризованную волну. В том случае, когда эллипс вырождается в окружность, имеем круговую поляризацию волны. Вполне очевидно, что условием этого является преобразование описанного прямоугольника в квадрат, что эквивалентно выравниванию амплитуд: $A_x = A_y = A$. Кроме того, в этом случае один из компонентов вектора \mathbf{E} должен равняться нулю, когда другой достигает максимального значения. Последнее условие эквивалентно равенству:

$\delta = \varphi_y - \varphi_x = \pm(2m-1)\frac{\pi}{2}$, где $m = 1, 2, 3, \dots$. При соблюдении названных условий уравнение (1.55) принимает вид уравнения окружности:

$$E_x^2 + E_y^2 = A^2.$$

В случае правой поляризации $\sin \delta > 0$.

При этом

$$\delta = \frac{\pi}{2} + 2m\pi, \text{ где } m = 0, \pm 1, \pm 2, \dots;$$

$$E_x = a \cos(\tau + \varphi_x),$$

$$E_y = a \cos\left(\tau + \varphi_x + \frac{\pi}{2} + 2m\pi\right) = -a \sin(\tau + \varphi_x).$$

При левой поляризации $\sin \delta < 0$.

При этом

$$\delta = -\frac{\pi}{2} + 2m\pi, \text{ где } m = 0, \pm 1, \pm 2, \dots;$$

$$E_x = a \cos(\tau + \varphi_x),$$

$$E_y = a \cos\left(\tau + \varphi_x - \frac{\pi}{2} + 2m\pi\right) = a \sin(\tau + \varphi_x).$$

Если вместо вещественного представления компонентов вектора \mathbf{E} использовать комплексное:

$$E_x = A_x \exp[i(\tau + \varphi_x)],$$

$$E_y = A_y \exp[i(\tau + \varphi_y)], \text{ то}$$

$$\frac{E_y}{E_x} = \frac{A_y}{A_x} \exp[i(\tau + \varphi_y - \tau - \varphi_x)] = \frac{A_y}{A_x} \exp(i\delta)$$

или

$$\frac{E_y}{E_x} = \frac{A_y}{A_x} \exp(i\delta).$$

Значение этого отношения позволяет сразу же определить характер поляризации:

– линейная поляризация электромагнитной волны, если $\delta = \pm m\pi$, где $m = 0, 1, 2, \dots$; при этом

$$\frac{E_y}{E_x} = (-1)^m \frac{A_y}{A_x};$$

– правая круговая поляризация электромагнитной волны, если $A_x = A_y$, $\delta = \frac{\pi}{2} + 2m\pi$, где $m = 0, \pm 1, \pm 2, \dots$; при этом

$$\frac{E_x}{E_y} = \exp\left[i\left(\frac{\pi}{2} + 2m\pi\right)\right] = \exp\left(i\frac{\pi}{2}\right) = i;$$

– левая круговая поляризация электромагнитной волны, если

$A_x = A_y$, $\delta = -\frac{\pi}{2} + 2m\pi$, где $m = 0, \pm 1, \pm 2, \dots$; при этом

$$\frac{E_x}{E_y} = \exp\left[i\left(-\frac{\pi}{2} + 2m\pi\right)\right] = \exp\left(-i\frac{\pi}{2}\right) = -i.$$

Можно показать, что в общем случае для правой эллиптической поляризации мнимая часть отношения $\frac{E_y}{E_x}$ положительна, тогда как для левой эллиптической поляризации она отрицательна. На рис. 1.7 показаны эллипсы поляризации при разных значениях δ .

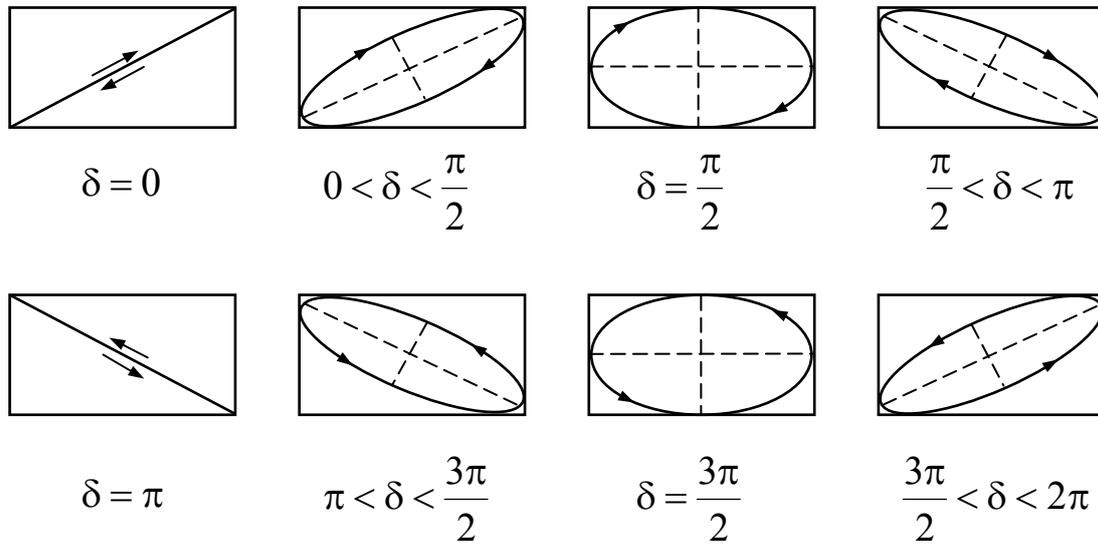


Рис. 1.7. Эллиптическая поляризация при различных значениях разности фаз δ

Для характеристики эллипса поляризации необходимы три независимые величины, к которым из уже рассмотренных можно отнести амплитуды A_x и A_y и разность фаз δ или малую и большую полуоси эллипса a и b и угол ψ , характеризующий ориентацию эллипса. Для практических целей состояние поляризации удобно характеризовать параметрами, обладающими одинаковой физической размерностью. Такие параметры были введены английским физиком и математиком Джорджем Габриелем Стоксом в 1852 году при его исследованиях частично поляризованного света. Параметрами Стокса для плоской монохроматической волны служат следующие четыре величины:

$$s_0 = A_x^2 + A_y^2 \tag{1.73a}$$

$$s_1 = A_x^2 - A_y^2 \tag{1.73б}$$

$$s_2 = 2A_x A_y \cos \delta \tag{1.73в}$$

$$s_3 = 2A_x A_y \sin \delta. \tag{1.73г}$$

Лишь три из этих величин независимы, так как справедливо тождество

$$s_0^2 = s_1^2 + s_2^2 + s_3^2. \quad (1.74)$$

Из выражения (1.73) следует, что параметр s_0 пропорционален интенсивности волны.

Напомним, что $\frac{A_y}{A_x} = \operatorname{tg} \alpha$. При этом из соотношения (1.73) имеем

$$s_0 = A_x^2 \left(1 + \frac{A_y^2}{A_x^2} \right) = A_x^2 (1 + \operatorname{tg}^2 \alpha) = \frac{A_x^2}{\cos^2 \alpha}.$$

Полученное соотношение позволяет выражение (1.72) преобразовать к виду:

$$\sin 2\chi = 2 \cos^2 \alpha \operatorname{tg} \alpha \sin \delta = 2 \cos^2 \alpha \frac{A_x A_y}{A_x^2} \sin \delta = 2 \frac{A_x A_y \sin \delta}{s_0}.$$

Учитывая при этом выражение (1.73г), получаем

$$s_3 = s_0 \sin 2\chi. \quad (1.75)$$

В результате подстановки соотношений (1.73б) и (1.73в) в выражение (1.66) имеем

$$s_2 = s_1 \operatorname{tg} 2\psi. \quad (1.76)$$

Подставим соотношения (1.75) и (1.76) в выражение (1.74):

$$s_0^2 = s_1^2 + s_1^2 \operatorname{tg}^2 2\psi + s_0^2 \sin^2 2\chi.$$

Отсюда находим, что

$$s_1 = s_0 \cos 2\chi \cos 2\psi. \quad (1.77)$$

При этом выражение (1.76) можно представить в виде:

$$s_2 = s_0 \cos 2\chi \sin 2\psi.$$

Итак,

$$s_1 = s_0 \cos 2\chi \cos 2\psi \quad (1.78a)$$

$$s_2 = s_0 \cos 2\chi \sin 2\psi \quad (1.78б)$$

$$s_3 = s_0 \sin 2\chi. \quad (1.78в)$$

Отсюда следует, что параметры s_1 , s_2 и s_3 достаточно простыми соотношениями связаны с углом $\chi \left(-\frac{\pi}{4} \leq \chi \leq \frac{\pi}{4} \right)$, характеризующим эллиптичность поляризации и направление вращения электрического вектора, и с углом $\psi \left(0 \leq \psi < \pi \right)$, характеризующим ориентацию эллипса.

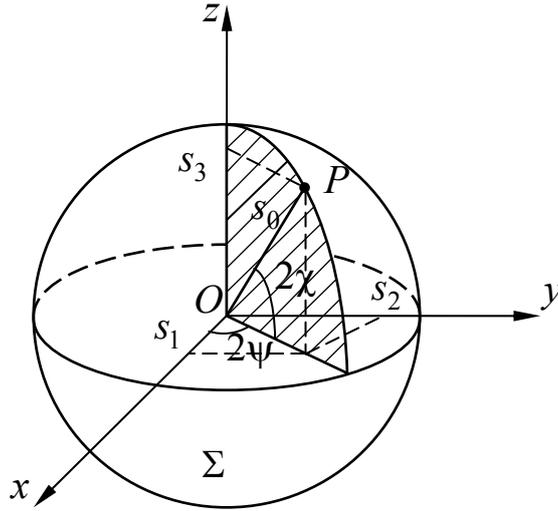


Рис. 1.8. Представление состояния поляризации монохроматической волны по Пуанкаре (сфера Пуанкаре)

Полученные выражения (1.78а)–(1.78в) определяют возможность простого геометрического представления различных состояний поляризации: параметры Стокса s_1 , s_2 и s_3 можно рассматривать как декартовы координаты точки P на сфере Σ радиуса s_0 , причем углы 2χ и 2ψ являются сферическими угловыми координатами этой точки, как показано на рис. 1.8. Таким образом, каждому возможному состоянию поляризации плоской монохроматической волны заданной интенсивности ($s_0 = const$) соответствует одна точка на сфере Σ и наоборот. Так как знак угла χ определяет направление вращения вектора, то, как следует из выражения (1.78в), правая поляризация представляется точками на поверхности Σ , лежащими выше экваториальной плоскости (плоскости xy), а левая поляризация – точками на поверхности Σ , лежащими ниже этой плоскости.

Для линейно поляризованного света разность фаз δ равна нулю или числу, кратному π . В соответствии с выражением (1.78в) параметр Стокса s_3 равен нулю, если $\chi = m\pi$, где $m = 0, \pm 1, \pm 2, \dots$. Таким образом, линейная поляризация представляется точками окружности в экваториальной плоскости. Для круговой поляризации $A_x = A_y$ и при правой поляризации угол $\delta = \frac{\pi}{2}$, а при левой поляризации угол $\delta = -\frac{\pi}{2}$. При этом в соответствии с соотношением (1.72) $\sin 2\chi = \pm 1$. Следовательно, правая круговая поляризация представляется северным полюсом ($s_1 = s_2 = 0, s_3 = s_0$), а левая круговая поляризация – южным полюсом ($s_1 = s_2 = 0, s_3 = -s_0$).

Рассмотренное геометрическое представление различных состояний поляризации точками на сфере было предложено французским математиком, физиком и философом Жюлем Анри Пуанкаре. Оно чрезвычайно полезно в кристаллооптике для определения влияния кристаллических сред на состояние поляризации проходящего через них света. Сфера Σ называется сферой Пуанкаре.

1.6. Закон сохранения энергии электромагнитного поля

Умножим уравнение (1.12) на \mathbf{E} , а уравнение (1.13) на \mathbf{H} и образуем разность полученных произведений:

$$\mathbf{E} \cdot \text{rot} \mathbf{H} - \mathbf{H} \cdot \text{rot} \mathbf{E} = \frac{4\pi}{c} \mathbf{j} \cdot \mathbf{E} + \frac{1}{c} \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \frac{1}{c} \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t}. \quad (1.79)$$

Заметим, что

$$\begin{aligned} \text{div}(\mathbf{E} \times \mathbf{H}) &= \mathbf{i} \left(\frac{\partial \mathbf{E}}{\partial x} \times \mathbf{H} + \mathbf{E} \times \frac{\partial \mathbf{H}}{\partial x} \right) + \mathbf{j} \left(\frac{\partial \mathbf{E}}{\partial y} \times \mathbf{H} + \mathbf{E} \times \frac{\partial \mathbf{H}}{\partial y} \right) + \\ &+ \mathbf{k} \left(\frac{\partial \mathbf{E}}{\partial z} \times \mathbf{H} + \mathbf{E} \times \frac{\partial \mathbf{H}}{\partial z} \right). \end{aligned}$$

Из свойств смешанного произведения векторов следует, что

$$\begin{aligned} \mathbf{i} \left(\frac{\partial \mathbf{E}}{\partial x} \times \mathbf{H} \right) &= \mathbf{H} \left(\mathbf{i} \times \frac{\partial \mathbf{E}}{\partial x} \right), \\ \mathbf{i} \left(\mathbf{E} \times \frac{\partial \mathbf{H}}{\partial x} \right) &= -\mathbf{i} \left(\frac{\partial \mathbf{H}}{\partial x} \times \mathbf{E} \right) = -\mathbf{E} \left(\mathbf{i} \times \frac{\partial \mathbf{H}}{\partial x} \right) \text{ и так далее.} \end{aligned}$$

В результате получаем

$$\begin{aligned} \text{div}(\mathbf{E} \times \mathbf{H}) &= \mathbf{H} \left(\mathbf{i} \times \frac{\partial \mathbf{E}}{\partial x} + \mathbf{j} \times \frac{\partial \mathbf{E}}{\partial y} + \mathbf{k} \times \frac{\partial \mathbf{E}}{\partial z} \right) - \\ &- \mathbf{E} \left(\mathbf{i} \times \frac{\partial \mathbf{H}}{\partial x} + \mathbf{j} \times \frac{\partial \mathbf{H}}{\partial y} + \mathbf{k} \times \frac{\partial \mathbf{H}}{\partial z} \right) = \mathbf{H} \cdot \text{rot} \mathbf{E} - \mathbf{E} \cdot \text{rot} \mathbf{H}. \end{aligned}$$

Полученное векторное тождество позволяет выражение (1.79) представить в виде

$$\frac{1}{c} \left(\mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} \right) + \frac{4\pi}{c} \mathbf{j} \cdot \mathbf{E} + \text{div}(\mathbf{E} \times \mathbf{H}) = 0. \quad (1.80)$$

Умножив это выражение на $\frac{c}{4\pi}$ и проинтегрировав по произвольному объему, получаем

$$\frac{1}{4\pi} \int_V \left(\mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} \right) dV + \int_V \mathbf{j} \cdot \mathbf{E} dV + \frac{c}{4\pi} \int_V \text{div}(\mathbf{E} \times \mathbf{H}) dV = 0.$$

В соответствии с теоремой Остроградского-Гаусса имеем

$$\int_V \operatorname{div}(\mathbf{E} \times \mathbf{H}) dV = \oint_S (\mathbf{E} \times \mathbf{H}) d\mathbf{S},$$

где S – замкнутая поверхность, ограничивающая объем V . При этом полученное выражение можно представить в виде

$$\begin{aligned} & \frac{1}{4\pi} \int_V \left(\mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} + \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} \right) dV + \int_V \mathbf{j} \cdot \mathbf{E} dV + \\ & + \frac{c}{4\pi} \int_S (\mathbf{E} \times \mathbf{H}) d\mathbf{S} = 0. \end{aligned} \quad (1.81)$$

Используя материальные уравнения (1.3) и (1.5), получаем

$$\mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} = \mathbf{E} \cdot \frac{\partial}{\partial t} (\varepsilon \mathbf{E}) = \frac{1}{2} \frac{\partial}{\partial t} (\varepsilon \mathbf{E}^2) = \frac{1}{2} \frac{\partial}{\partial t} (\mathbf{E} \cdot \mathbf{D}),$$

$$\mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} = \mathbf{H} \cdot \frac{\partial}{\partial t} (\mu \mathbf{H}) = \frac{1}{2} \frac{\partial}{\partial t} (\mathbf{H} \cdot \mathbf{B}).$$

Обозначим

$$\frac{1}{8\pi} \mathbf{E} \cdot \mathbf{D} = w_e, \quad \frac{1}{8\pi} \mathbf{H} \cdot \mathbf{B} = w_m, \quad \int_V (w_e + w_m) dV = W.$$

При этом выражение (1.100) принимает вид:

$$\frac{dW}{dt} + \int_V \mathbf{j} \cdot \mathbf{E} dV + \frac{c}{4\pi} \int_S (\mathbf{E} \times \mathbf{H}) \mathbf{n} dS = 0. \quad (1.82)$$

Плотность тока \mathbf{j} в уравнениях Максвелла можно представить в виде суммы токов

$$\mathbf{j} = \mathbf{j}_c + \mathbf{j}_V, \quad (1.83)$$

где плотность тока проводимости

$$\mathbf{j}_c = \sigma \mathbf{E},$$

а плотность конвекционного тока

$$\mathbf{j}_V = \rho \mathbf{V},$$

где \mathbf{V} – скорость конвекционного движения зарядов.

Обозначив скорость движения заряда q_k через \mathbf{V}_k , будем считать, что $\mathbf{V}_k \ll c$, где c – скорость света в вакууме. Сила, действующая со стороны поля (\mathbf{E}, \mathbf{B}) на заряд q , движущийся со скоростью \mathbf{V} , определяется так называемым законом Лорентца:

$$\mathbf{F} = q \left[\mathbf{E} + \frac{1}{c} (\mathbf{V} \times \mathbf{B}) \right],$$

основанным на экспериментальных данных. Отсюда следует, что если все заряды q_k за время δt смещаются на расстояние $\delta \mathbf{x} = \mathbf{V}_k \delta t$ ($k = 1, 2, \dots$), то полная работа, выполненная при этом, равна

$$\delta A = \sum_k \mathbf{F}_k \cdot \delta \mathbf{x}_k = \sum_k q_k \mathbf{E}_k \cdot \mathbf{V}_k \delta t,$$

так как $\mathbf{V}_k (\mathbf{V}_k \times \mathbf{B}) = 0$.

Если число заряженных частиц велико, то распределение зарядов можно считать непрерывным. При этом $q_k = \rho dV$, где ρ – плотность зарядов. Тогда последнее равенство можно представить в виде

$$\delta A = \delta t \int_V \rho \mathbf{V} \cdot \mathbf{E} dV, \quad (1.84)$$

при этом интегрирование выполняется по произвольному объему. Выражение (1.84) удобно представить в виде

$$\frac{\delta A}{\delta t} = \int_V \mathbf{j}_V \cdot \mathbf{E} dV. \quad (1.85)$$

Определим вектор \mathbf{G} и скаляр Q соотношениями

$$\mathbf{G} = \frac{c}{4\pi} \mathbf{E} \times \mathbf{H} \quad (1.86)$$

$$Q = \int_V \mathbf{j}_c \cdot \mathbf{E} dV = \int_V \sigma \mathbf{E}^2 dV. \quad (1.87)$$

Равенство (1.81) и соотношения (1.85)–(1.87) позволяют представить выражение (1.82) в виде

$$\frac{dW}{dt} = -\frac{\partial A}{\partial t} - Q - \int_S \mathbf{G} \cdot \mathbf{n} dS. \quad (1.88)$$

Если замкнутая поверхность S представляет собой граничную поверхность полного поля, то потоком энергии через граничную поверхность можно пренебречь, что эквивалентно равенству $\int_S \mathbf{G} \cdot \mathbf{n} dS = 0$.

Для непроводящей среды ($\sigma = 0$) величина $Q = 0$. При этом выражение (1.88) в результате интегрирования принимает вид: $W + A = 0$. Следовательно, в замкнутой системе в непроводящей среде (т.е. при отсутствии потерь энергии) увеличение энергии W в единицу времени определяется работой по переносу зарядов в течение этого времени. При $\sigma \neq 0$ величина Q , называемая джоулевым теплом, определяет количество теплоты, выделяемой в единицу времени в проводнике из-за наличия электрического сопротивления. Если замкнутая поверхность не охватывает собой полного поля, то поле достигает этой поверхности и в соответствии с выражением (1.88) происходит дополнительное уменьшение энергии, т.е. поверхностный интеграл в выражении (1.88) определяет поток энергии через эту поверхность. Таким образом, величина W определяет полную энергию, заключенную внутри объема.

В непроводящей среде, свободной от зарядов и токов, в соответствии с выражением (1.88) имеем

$$\frac{dW}{dt} = \frac{d}{dt} \int_V (w_e + w_m) dV = \frac{d}{dt} \int_V w dV = - \int_S \mathbf{G} \cdot \mathbf{n} dS.$$

Следовательно, величину w_e можно отождествить с плотностью энергии электрического поля, а w_m – с плотностью энергии магнитного поля.

Плотность потока энергии электромагнитного поля определяется так называемым вектором Умова-Пойнтинга, равным

$$\mathbf{G} = \frac{c}{4\pi} \mathbf{E} \times \mathbf{H}. \quad (1.86)$$

Заметим, что впервые вектор плотности потока любого вида энергии был введен в 1874 году русским физиком Николаем Алексеевичем Умовым.

Направление вектора Умова-Пойнтинга перпендикулярно векторам \mathbf{E} и \mathbf{H} и совпадает с направлением распространения электромагнитной энергии, а его величина равна энергии, переносимой в единицу времени через единичную площадку, перпендикулярную вектору \mathbf{G} .

В оптике наибольший интерес представляет усредненный вектор Умова-Пойнтинга. Величина его служит мерой интенсивности света, а направление указывает направление распространения света.

1.7. Электромагнитное поле на границе раздела двух сред

Любая оптическая система представляет собой систему поверхностей раздела двух сред с различными свойствами, характеризующимися величинами ε и μ . Уравнения электромагнитного поля можно применить к каждой среде с учетом значений электрических и магнитных параметров, характеризующих каждую среду. Однако, необходимо иметь систему граничных условий, которые связывали бы на границе раздела двух сред два решения уравнений электромагнитного поля, каждое из которых справедливо в отдельности по одну сторону от границы раздела. Выведем соотношения, описывающие переход электромагнитного поля через такую поверхность раздела.

Заменим поверхность раздела сред S тонким переходным слоем, внутри которого ε и μ быстро, но непрерывно меняются от значений, характеризующих среду с одной стороны поверхности, до значений с другой ее стороны. Внутри этого слоя построим цилиндр, ограниченный с боков частокотлом нормалей к поверхности S в пределах малых площадок δS_1 и δS_2 , параллельных поверхности S и служащих основаниями цилиндра с каждой ее стороны, как показано на рис. 1.9. Поскольку во всем цилиндре вектор \mathbf{V} и его производные непрерывны, мы можем применить теорему Остроградского-Гаусса к интегралу от $div\mathbf{V}$, взятому по объему цилиндра, в виде:

$$\oint_V \operatorname{div} \mathbf{B} dV = \oint_S \mathbf{B} \cdot \mathbf{n} dS = 0. \quad (1.89)$$

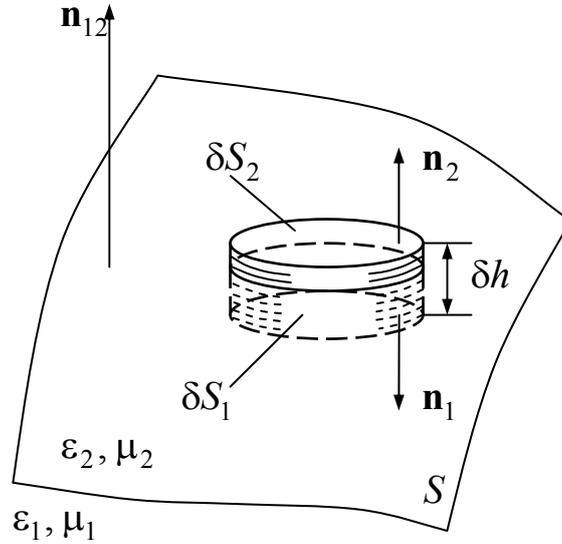


Рис. 1.9. К выводу граничных условий для нормальных компонентов векторов \mathbf{B} и \mathbf{D}

Здесь \mathbf{n} – единичный вектор внешней нормали. Заметим, что второй интеграл берется по всей поверхности цилиндра. Так как площадки dS_1 и dS_2 предполагаются малыми, можно считать, что на них вектор магнитной индукции \mathbf{B} принимает постоянные значения \mathbf{B}_1 и \mathbf{B}_2 . Тогда выражение (1.89) можно заменить выражением вида:

$$\mathbf{B}_1 \cdot \mathbf{n}_1 \delta S_1 + \mathbf{B}_2 \cdot \mathbf{n}_2 \delta S_2 + \text{вклад от стенок} = 0. \quad (1.90)$$

Если высота цилиндра dh стремится к нулю, то переходный слой переходит в поверхность, а вклад от стенок цилиндра исчезает при условии, что отсутствует поверхностный поток магнитной индукции. Такой поток никогда не наблюдается и, следовательно, в пределе имеем

$$(\mathbf{B}_1 \cdot \mathbf{n}_1 + \mathbf{B}_2 \cdot \mathbf{n}_2) \delta S = 0, \quad (1.91)$$

где δS – площадь поверхности пересечения рассматриваемого цилиндра с поверхностью S . Если \mathbf{n}_{12} – единичный вектор нормали, направленный из первой среды во вторую, то $\mathbf{n}_1 = -\mathbf{n}_{12}$, а $\mathbf{n}_2 = \mathbf{n}_{12}$. При этом соотношение (1.91) принимает вид:

$$\mathbf{n}_{12} (\mathbf{B}_2 - \mathbf{B}_1) = 0, \quad (1.92)$$

т.е. нормальная составляющая вектора магнитной индукции непрерывна на поверхности раздела двух сред.

Подобным образом можно исследовать электрическое поле на границе раздела двух сред. Однако, в этом случае при наличии зарядов следует учесть выражение (1.10). При этом аналогично выражению (1.89) получаем

$$\oint_V \operatorname{div} \mathbf{D} dV = \oint_S \mathbf{D} \cdot \mathbf{n} dS = 4\pi \oint_V \rho dV. \quad (1.93)$$

При слиянии площадок δS_1 и δS_2 объемная плотность зарядов ρ переходит в поверхностную плотность зарядов $\tilde{\rho}$, определяемую соотношением

$$\lim_{\delta h \rightarrow 0} \int_V \rho dV = \int_S \tilde{\rho} dS. \quad (1.94)$$

Аналогично объемная плотность тока \mathbf{j} переходит в поверхностную плотность тока $\tilde{\mathbf{j}}$:

$$\lim_{\delta h \rightarrow 0} \int_V \mathbf{j} dV = \int_S \tilde{\mathbf{j}} dS. \quad (1.95)$$

Если площадку δS и высоту δh выбрать достаточно малыми, то выражение (1.93) можно заменить выражением вида:

$$\mathbf{D}_1 \cdot \mathbf{n}_1 \delta S_1 + \mathbf{D}_2 \cdot \mathbf{n}_2 \delta S_2 + \text{вклад от стенок} = 4\pi \tilde{\rho} \delta S.$$

Вклад от стенок при уменьшении высоты δh уменьшается, приближаясь к нулю и поэтому в пределе при $\delta h \rightarrow 0$ получаем

$$\mathbf{n}_{12} (\mathbf{D}_2 - \mathbf{D}_1) = 4\pi \tilde{\rho}, \quad (1.96)$$

т.е. при наличии на поверхности раздела сред слоя с поверхностной плотностью заряда $\tilde{\rho}$ нормальный компонент вектора электрического смещения при переходе через эту поверхность испытывает скачок, равный $4\pi \tilde{\rho}$.

Исследуем теперь поведение тангенциальной составляющей векторов электрического и магнитного полей. Заменяем поверхность раздела сред переходным слоем, а цилиндр, показанный на рис. 1.9, прямоугольной площадкой, стороны которой параллельны и перпендикулярны поверхности S , как показано на рис. 1.10.

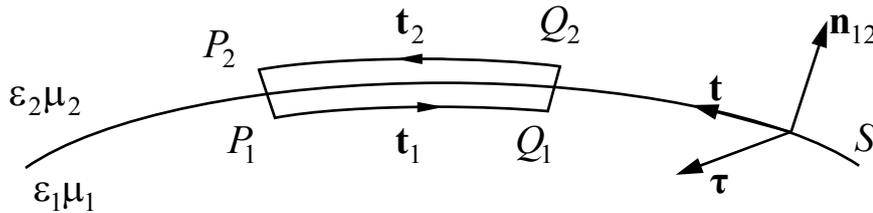


Рис. 1.10. К выводу граничных условий для тангенциальных компонентов векторов \mathbf{E} и \mathbf{H}

Пусть $\boldsymbol{\tau}$ – единичный вектор, перпендикулярный плоскости рассматриваемого прямоугольника. Используя уравнение (1.13), в соответствии с теоремой Стокса имеем

$$\int_S \text{rot} \mathbf{E} \cdot \boldsymbol{\tau} dS = \oint_L \mathbf{E} \cdot d\mathbf{l} = -\frac{1}{c} \int_S \dot{\mathbf{B}} \cdot \boldsymbol{\tau} dS. \quad (1.97)$$

Здесь первый и третий интегралы берутся по площади прямоугольника, а второй – вдоль его границ. Если длины $P_1Q_1 = \delta l_1$ и $P_2Q_2 = \delta l_2$ малы, то на каждой из этих сторон вектор \mathbf{E} можно заменить

постоянными векторами \mathbf{E}_1 и \mathbf{E}_2 . При этом и вектор $\dot{\mathbf{B}}$ можно считать постоянным. Тогда выражение (1.97) можно заменить выражением вида:

$$\mathbf{E}_1 \cdot \mathbf{t}_1 \delta l_1 + \mathbf{E}_2 \cdot \mathbf{t}_2 \delta l_2 + \text{вклад от концов} = -\frac{1}{c} \dot{\mathbf{B}} \cdot \boldsymbol{\tau} \delta l \delta h, \quad (1.98)$$

где δl – линейный элемент, по которому прямоугольник пересекается с поверхностью раздела; \mathbf{t} – единичный вектор касательной к поверхности.

Если теперь постепенно уменьшать высоту прямоугольника, то вклад от концов $P_1 P_2$ и $Q_1 Q_2$ будет уменьшаться, приближаясь к нулю. Предположим также, что вектор $\dot{\mathbf{B}}$ при этом остается конечным. Тогда в пределе при $\delta h \rightarrow 0$ получим

$$(\mathbf{E}_1 \cdot \mathbf{t}_1 + \mathbf{E}_2 \cdot \mathbf{t}_2) \delta l = 0. \quad (1.99)$$

В соответствии с рисунком

$$\mathbf{t}_1 = -\mathbf{t} = -\boldsymbol{\tau} \times \mathbf{n}_{12}, \quad \mathbf{t}_2 = \mathbf{t} = \boldsymbol{\tau} \times \mathbf{n}_{12}.$$

При этом выражение (1.99) принимает вид:

$$\boldsymbol{\tau} [\mathbf{n}_{12} \times (\mathbf{E}_2 - \mathbf{E}_1)] = 0.$$

Так как ориентация прямоугольника, а, следовательно, и единичного вектора $\boldsymbol{\tau}$, произвольна, то

$$\mathbf{n}_{12} \times (\mathbf{E}_2 - \mathbf{E}_1) = 0, \quad (1.100)$$

т.е. тангенциальная составляющая электрического вектора непрерывна на поверхности раздела двух сред.

Аналогичный анализ поведения тангенциальной составляющей магнитного вектора позволяет получить выражение, аналогичное выражению (1.98), вида:

$$\begin{aligned} & \mathbf{H}_1 \cdot \mathbf{t}_1 \delta l_1 + \mathbf{H}_2 \cdot \mathbf{t}_2 \delta l_2 + \text{вклад от концов} = \\ & = \frac{1}{c} \dot{\mathbf{D}} \cdot \boldsymbol{\tau} \delta l \delta h + \frac{4\pi}{c} \tilde{\mathbf{j}} \cdot \boldsymbol{\tau} \delta l. \end{aligned} \quad (1.101)$$

В результате предельного перехода ($\delta h \rightarrow 0$) получаем

$$\mathbf{n}_{12} \times (\mathbf{H}_2 - \mathbf{H}_1) = \frac{4\pi}{c} \tilde{\mathbf{j}}. \quad (1.102)$$

Итак, из полученных соотношений следует, что векторы напряженности электрического \mathbf{E} и магнитного \mathbf{H} полей и векторы электрической \mathbf{D} и магнитной \mathbf{B} индукции удовлетворяют следующим граничным условиям:

1. Тангенциальные составляющие напряженности электрического поля непрерывны, т.е. $\mathbf{E}_{t_1} = \mathbf{E}_{t_2}$, на границе раздела двух сред, где физические свойства среды, характеризуемые значениями ϵ и μ , изменяются скачком.

2. Нормальные компоненты вектора электрической индукции непрерывны, т.е. $\mathbf{D}_{n_1} = \mathbf{D}_{n_2}$, при отсутствии поверхностных зарядов на границе раздела сред.
3. Тангенциальные составляющие вектора напряженности магнитного поля непрерывны, т.е. $\mathbf{H}_{t_1} = \mathbf{H}_{t_2}$, при отсутствии поверхностных токов на границе раздела сред.
4. Нормальные составляющие вектора магнитной индукции непрерывны, т.е. $\mathbf{B}_{n_1} = \mathbf{B}_{n_2}$.

1.8. Плоская волна на границе раздела двух однородных изотропных сред

Применим соотношения, которым удовлетворяют векторы электромагнитного поля на поверхностях, где физические свойства среды претерпевают разрыв, к исследованию распространения плоской волны, падающей на плоскую поверхность, разделяющую две однородные изотропные среды. Если на границу раздела двух однородных сред с разными оптическими свойствами падает плоская волна, то она разделяется на две волны: проходящую во вторую среду и отраженную. Существование двух волн определяется необходимостью соблюдения граничных условий прохождения электромагнитным полем поверхности раздела двух сред. Предположим, что эти волны также будут плоскими, для которых следует определить значения амплитуд и направлений распространения. Индексами i , r и t будем обозначать падающую, отраженную и проходящую (преломленную) волны.

Плоская волна, распространяющаяся в направлении единичного вектора $\mathbf{s}^{(i)}$, полностью определена, если известно поведение возмущения во времени в одной точке пространства, поскольку, если $\mathbf{F}(t)$ представляет зависимость возмущения от времени в какой-то одной точке, то эта зависимость в другой точке, отстоящей от первой на расстоянии,

определяемом вектором \mathbf{r} , будет $\mathbf{F}\left(t - \frac{\mathbf{r} \cdot \mathbf{s}^{(i)}}{V}\right)$. Уравнение, определяющее

напряженность $\mathbf{E}(\mathbf{r}, t)$ электрического поля или индукцию $\mathbf{B}(\mathbf{r}, t)$ магнитного поля, представим в виде: $\mathbf{A}(\mathbf{r}, t) = \mathbf{A}_0 \exp[-i(\omega t - \mathbf{k} \cdot \mathbf{r})]$, где $\mathbf{A}_0 = \mathbf{a} \exp(i\varphi)$, при этом переменная часть фазового множителя в выражении, определяющем гармоническую электромагнитную плоскую волну, равна

$$\tau = \omega t - \mathbf{k} \cdot \mathbf{r} = \omega \left(t - \frac{\mathbf{r} \cdot \mathbf{s}}{V} \right),$$

где \mathbf{s} – единичный вектор направления распространения волны; V – скорость (фазовая) распространения волны.

На границе двух сред вторичные поля будут так же изменяться во времени, как и первичное поле падающей волны. Следовательно, если $\mathbf{s}^{(r)}$ и $\mathbf{s}^{(t)}$ – единичные векторы в направлении распространения отраженной и падающей волн, то, приравнявая аргументы трех волновых функций в произвольной точке $\mathbf{r}(x, y, z)$ на поверхности раздела двух сред, получаем

$$\omega_i \left(t - \frac{\mathbf{r} \cdot \mathbf{s}^{(i)}}{V_1} \right) = \omega_r \left(t - \frac{\mathbf{r} \cdot \mathbf{s}^{(r)}}{V_1} \right) = \omega_t \left(t - \frac{\mathbf{r} \cdot \mathbf{s}^{(t)}}{V_2} \right),$$

где V_1 и V_2 – скорости распространения волн в первой и во второй средах. Чтобы граничные условия соблюдались в любой момент времени, коэффициенты при t в переменной части фазового множителя должны быть одинаковыми для всех трех волн. Поэтому частоты отраженной и преломленной волн равны частоте ω_i падающей волны. При этом

$$t - \frac{\mathbf{r} \cdot \mathbf{s}^{(i)}}{V_1} = t - \frac{\mathbf{r} \cdot \mathbf{s}^{(r)}}{V_1} = t - \frac{\mathbf{r} \cdot \mathbf{s}^{(t)}}{V_2}. \quad (1.103)$$

Выбрав в качестве поверхности раздела сред плоскость $z = 0$, соотношения (1.103) можно представить в виде:

$$\frac{x s_x^{(i)} + y s_y^{(i)}}{V_1} = \frac{x s_x^{(r)} + y s_y^{(r)}}{V_1} = \frac{x s_x^{(t)} + y s_y^{(t)}}{V_2}. \quad (1.104)$$

Равенства (1.104) должны выполняться для любых значений x и y на поверхности раздела, а поэтому

$$\frac{s_x^{(i)}}{V_1} = \frac{s_x^{(r)}}{V_1} = \frac{s_x^{(t)}}{V_2}, \quad \frac{s_y^{(i)}}{V_1} = \frac{s_y^{(r)}}{V_1} = \frac{s_y^{(t)}}{V_2}. \quad (1.105)$$

Плоскость, определяемая вектором $\mathbf{s}^{(i)}$ и нормалью к поверхности раздела, называется плоскостью падения. Из соотношений (1.105) следует, что и векторы $\mathbf{s}^{(r)}$ и $\mathbf{s}^{(t)}$ лежат в этой плоскости. Считая плоскость xz плоскостью падения и обозначая через θ_i , θ_r и θ_t углы, которые векторы $\mathbf{s}^{(i)}$, $\mathbf{s}^{(r)}$ и $\mathbf{s}^{(t)}$ образует с осью Oz , в соответствии с рис. 1.11 имеем

$$\left. \begin{aligned} s_x^{(i)} &= \sin \theta_i, & s_y^{(i)} &= 0, & s_z^{(i)} &= \cos \theta_i; \\ s_x^{(r)} &= \sin \theta_r, & s_y^{(r)} &= 0, & s_z^{(r)} &= \cos \theta_r; \\ s_x^{(t)} &= \sin \theta_t, & s_y^{(t)} &= 0, & s_z^{(t)} &= \cos \theta_t. \end{aligned} \right\} \quad (1.106)$$

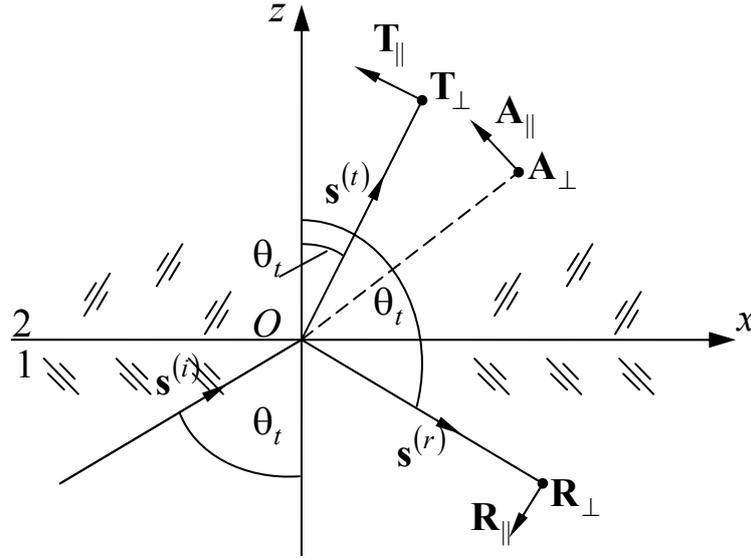


Рис. 1.11. Плоская волна на границе раздела двух однородных изотропных сред

Если волна распространяется из первой среды во вторую, то проекция вектора \mathbf{s} на ось Oz положительна, а если в противоположном направлении – то отрицательна. Таким образом,

$$s_z^{(i)} = \cos \theta_i \geq 0, \quad s_z^{(r)} = \cos \theta_r \leq 0, \quad s_z^{(t)} = \cos \theta_t \geq 0. \quad (1.107)$$

Подставив соотношения (1.106) в первую систему равенств (1.105), получаем

$$\frac{\sin \theta_i}{V_1} = \frac{\sin \theta_r}{V_1} = \frac{\sin \theta_t}{V_2}. \quad (1.108)$$

Отсюда следует, что $\sin \theta_r = \sin \theta_i$, при этом, как следует из соотношений (1.107), $\cos \theta_r = -\cos \theta_i$. Следовательно, $\theta_r = \pi - \theta_i$. Последнее соотношение вместе с утверждением, что нормаль $\mathbf{s}^{(r)}$ к отраженной волне лежит в плоскости падения, соответствует закон отражения плоской волны от плоской поверхности раздела двух сред.

Напомним, что в среде, характеризуемой величинами ϵ и μ , скорость распространения электромагнитной волны $V = \frac{c}{\sqrt{\epsilon\mu}}$, где c – скорость

распространения волны в вакууме. Используя соотношение Максвелла, согласно которому показатель преломления среды $n = \sqrt{\epsilon\mu}$, из соответствующего равенства (1.108) находим, что

$$\frac{\sin \theta_i}{\sin \theta_t} = \frac{V_1}{V_2} = \sqrt{\frac{\epsilon_2 \mu_2}{\epsilon_1 \mu_1}} = \frac{n_2}{n_1} = n_{12}. \quad (1.109)$$

Следующее отсюда соотношение

$$n_1 \sin \theta_i = n_2 \sin \theta_t$$

вместе с утверждением, что нормаль $\mathbf{s}^{(t)}$ к прошедшей волне лежит в плоскости падения, составляет закон преломления (закон Декарта-Снеллиуса) плоской волны на границе раздела двух сред.

Если $n_2 > n_1$, то $n_{12} > 1$. При этом говорят, что оптическая плотность второй среды больше, чем первой. В этом случае, учитывая выражение (1.109), имеем

$$\sin \theta_t = \frac{1}{n_{12}} \sin \theta_i \leq \sin \theta_i.$$

Однако, если вторая среда оптически менее плотная, чем первая (т.е. если $n_{12} < 1$), то угол θ_t приобретает вещественное значение при том условии, что

$$\sin \theta_t = \frac{\sin \theta_i}{n_{12}} \leq 1.$$

$$\text{При } \frac{\sin \theta_i}{n_{12}} > 1: \cos \theta_t = \pm i \sqrt{\frac{\sin^2 \theta_i}{n_{12}^2} - 1}.$$

Таким образом, при $n_{12} < 1$ угол θ_t приобретает вещественное значение лишь для таких углов падения θ_i , для которых $\sin \theta_i \leq n_{12}$. При $\sin \theta_i > n_{12}$ наблюдается так называемое полное внутреннее отражение.

1.9. Формулы Френеля

Определив направления распространения отраженной и преломленной волн, определим их амплитуды. Предположим, что обе среды (однородные и изотропные) обладают нулевой проводимостью и, следовательно, совершенно прозрачны (нет выделения джоулева тепла, а, следовательно, нет и потерь электромагнитной энергии). Их магнитные проницаемости фактически будут отличаться от единицы на пренебрежимо малые величины, а поэтому будем считать, что $\mu_1 = \mu_2 = 1$. Пусть $\mathbf{E}_i = A \exp(-i\tau_i)$, где A – амплитуда вектора электрического поля падающей волны, которую будем считать комплексной величиной с фазой, равной постоянной части аргумента волновой функции. Переменная ее часть имеет вид:

$$\tau_i = \omega \left(t - \frac{\mathbf{r} \cdot \mathbf{s}^{(i)}}{V_1} \right) = \omega \left(t - \frac{x \sin \theta_i + z \cos \theta_i}{V_1} \right). \quad (1.110)$$

Разложим каждый вектор на составляющие компоненты: параллельный (с индексом \parallel) и перпендикулярный (с индексом \perp) плоскости падения. Выбор положительных направлений для параллельных и перпендикулярных составляющих компонентов вектора поля показан на рис. 1.11. При этом составляющие вектора электрического поля падающей волны в левой системе координат можно записать в виде:

$$\left. \begin{aligned} E_x^{(i)} &= -A_{\parallel} \cos \theta_i \exp(-i\tau_i), \\ E_y^{(i)} &= A_{\perp} \exp(-i\tau_i), \\ E_z^{(i)} &= A_{\parallel} \sin \theta_i \exp(-i\tau_i). \end{aligned} \right\} \quad (1.111)$$

Вектор магнитного поля определяется соотношением (1.42), которое при $\mu = 1$ можно представить в виде:

$$\mathbf{H} = \sqrt{\varepsilon} \mathbf{s} \times \mathbf{E}.$$

Выразив вектора \mathbf{E} и \mathbf{H} через их проекции на координатные оси, получаем

$$H_x \mathbf{i} + H_y \mathbf{j} + H_z \mathbf{k} = \sqrt{\varepsilon} (s_x \mathbf{i} + s_y \mathbf{j} + s_z \mathbf{k}) \times (E_x \mathbf{i} + E_y \mathbf{j} + E_z \mathbf{k}).$$

При этом

$$H_x^{(i)} = \sqrt{\varepsilon_1} (s_y^{(i)} E_z^{(i)} - s_z^{(i)} E_y^{(i)}),$$

$$H_y^{(i)} = \sqrt{\varepsilon_1} (s_z^{(i)} E_x^{(i)} - s_x^{(i)} E_z^{(i)}),$$

$$H_z^{(i)} = \sqrt{\varepsilon_1} (s_x^{(i)} E_y^{(i)} - s_y^{(i)} E_x^{(i)}).$$

С учетом соотношений (1.106) и (1.111) находим, что составляющие вектора магнитного поля падающей волны в левой системе координат равны

$$H_x^{(i)} = -A_{\perp} \cos \theta_i \sqrt{\varepsilon_1} \exp(-i\tau_i),$$

$$H_y^{(i)} = -A_{\parallel} \sqrt{\varepsilon_1} \exp(-i\tau_i),$$

$$H_z^{(i)} = A_{\perp} \sin \theta_i \sqrt{\varepsilon_1} \exp(-i\tau_i).$$

Аналогично, если T и R – комплексные амплитуды прошедшей и отраженной волн, то соответствующие составляющие векторов электрического и магнитного полей равны следующим величинам.

Поле прошедшей волны

$$\left. \begin{aligned} E_x^{(t)} &= -T_{\parallel} \cos \theta_t \exp(-i\tau_t), \\ E_y^{(t)} &= T_{\perp} \exp(-i\tau_t), \\ E_z^{(t)} &= T_{\parallel} \sin \theta_t \exp(-i\tau_t); \\ H_x^{(t)} &= -T_{\perp} \cos \theta_t \sqrt{\varepsilon_2} \exp(-i\tau_t), \\ H_y^{(t)} &= -T_{\parallel} \sqrt{\varepsilon_2} \exp(-i\tau_t), \\ H_z^{(t)} &= T_{\perp} \sin \theta_t \sqrt{\varepsilon_2} \exp(-i\tau_t), \end{aligned} \right\} \quad (1.112)$$

$$\text{где } \tau_t = \omega \left(t - \frac{\mathbf{r} \cdot \mathbf{s}^{(t)}}{V_2} \right) = \omega \left(t - \frac{x \sin \theta_t + z \sin \theta_t}{V_2} \right). \quad (1.113)$$

Поле отраженной волны

$$\left. \begin{aligned} E_x^{(r)} &= R_{\parallel} \cos \theta_r \exp(-i\tau_r), \\ E_y^{(r)} &= R_{\perp} \exp(-i\tau_r), \\ E_z^{(r)} &= -R_{\parallel} \sin \theta_r \exp(-i\tau_r); \\ H_x^{(r)} &= -R_{\perp} \cos \theta_r \sqrt{\varepsilon_1} \exp(-i\tau_r), \\ H_y^{(r)} &= R_{\parallel} \sqrt{\varepsilon_1} \exp(-i\tau_r), \\ H_z^{(r)} &= R_{\perp} \sin \theta_r \sqrt{\varepsilon_1} \exp(-i\tau_r), \end{aligned} \right\} \quad (1.114)$$

$$\text{где } \tau_r = \omega \left(t - \frac{\mathbf{r} \cdot \mathbf{s}^{(r)}}{V_1} \right) = \omega \left(t - \frac{x \sin \theta_r + z \sin \theta_r}{V_1} \right). \quad (1.115)$$

Граничные условия (1.100) и (1.102) требуют, чтобы на границе раздела двух сред тангенциальные составляющие векторов \mathbf{E} и \mathbf{H} были непрерывны. Следовательно, должны выполняться соотношения

$$\left. \begin{aligned} E_x^{(i)} + E_x^{(r)} &= E_x^{(t)}, & H_x^{(i)} + H_x^{(r)} &= H_x^{(t)}, \\ E_y^{(i)} + E_y^{(r)} &= E_y^{(t)}, & H_y^{(i)} + H_y^{(r)} &= H_y^{(t)}, \end{aligned} \right\} \quad (1.116)$$

при этом условия для нормальных составляющих векторов \mathbf{B} и \mathbf{D} , определяемые выражениями (1.92) и (1.96), естественным образом выполняются. Подставив в соотношения (1.116) соответствующие значения всех составляющих векторов \mathbf{E} и \mathbf{H} и учитывая тот факт, что $\cos \theta_r = \cos(\pi - \theta_i) = -\cos \theta_i$, при $\tau_i = \tau_r = \tau_t$ получаем следующие четыре соотношения

$$\left. \begin{aligned} (A_{\parallel} + R_{\parallel}) \cos \theta_i &= T_{\parallel} \cos \theta_t, \\ A_{\perp} + R_{\perp} &= T_{\perp}, \\ \sqrt{\varepsilon_1} (A_{\perp} - R_{\perp}) \cos \theta_i &= \sqrt{\varepsilon_2} T_{\perp} \cos \theta_t, \\ \sqrt{\varepsilon_1} (A_{\parallel} - R_{\parallel}) &= \sqrt{\varepsilon_2} T_{\parallel}. \end{aligned} \right\} \quad (1.117)$$

Каждое из уравнений (1.117) содержит составляющие компоненты векторов поля одного типа. Следовательно, волны рассматриваемых двух типов независимы друг от друга. Выразим из уравнений (1.117) составляющие компоненты векторов поля отраженной и прошедшей волн через составляющие компоненты векторов поля падающей волны.

Используя соотношения Максвелла $n = \sqrt{\varepsilon}$ ($\mu = 1$), получаем

$$\left. \begin{aligned} T_{\parallel} &= \frac{2n_1 \cos \theta_i}{n_2 \cos \theta_i + n_1 \cos \theta_t} A_{\parallel}, \\ T_{\perp} &= \frac{2n_1 \cos \theta_i}{n_1 \cos \theta_i + n_2 \cos \theta_t} A_{\perp}. \end{aligned} \right\} \quad (1.118)$$

$$\left. \begin{aligned} R_{\parallel} &= -\frac{n_2 \cos \theta_i - n_1 \cos \theta_t}{n_2 \cos \theta_i + n_1 \cos \theta_t} A_{\parallel}, \\ R_{\perp} &= \frac{n_1 \cos \theta_i - n_2 \cos \theta_t}{n_1 \cos \theta_i + n_2 \cos \theta_t} A_{\perp}. \end{aligned} \right\} \quad (1.119)$$

Используя полученные соотношения и материальные уравнения, легко убедиться в том, что граничные условия (1.92) и (1.96) для нормальных составляющих векторов поля соблюдаются. Соотношения (1.118) и (1.119) называются формулами Френеля. В несколько ином менее общем виде эти формулы были получены Френелем в 1823 году.

Умножим числитель и знаменатель первого из уравнений (1.118) на $\sin \theta_t$:

$$\begin{aligned} T_{\parallel} &= \frac{2n_1 \cos \theta_i \sin \theta_t}{n_2 \sin \theta_t \cos \theta_i + n_1 \sin \theta_t \cos \theta_t} A_{\parallel} = \frac{2 \sin \theta_t \cos \theta_i}{\sin \theta_i \cos \theta_i + \sin \theta_t \cos \theta_t} A_{\parallel} = \\ &= \frac{2 \sin \theta_t \cos \theta_i}{\sin \theta_t \sin \theta_i (\sin \theta_i \cos \theta_t + \cos \theta_i \sin \theta_t) + \cos \theta_t \cos \theta_i (\sin \theta_i \cos \theta_t + \cos \theta_i \sin \theta_t)} \times \\ &\times A_{\parallel} = \frac{2 \sin \theta_t \cos \theta_i}{\sin(\theta_i + \theta_t) \cos(\theta_i - \theta_t)} A_{\parallel}. \end{aligned}$$

Аналогично преобразуется второе из уравнений (1.118).

Теперь умножим на $\sin \theta_t$ числитель и знаменатель первого из уравнений (1.119):

$$\begin{aligned}
R_{\parallel} &= \frac{n_2 \sin \theta_t \cos \theta_i - n_1 \sin \theta_t \cos \theta_t}{n_2 \sin \theta_t \cos \theta_i + n_1 \sin \theta_t \cos \theta_t} A_{\parallel} = -\frac{\sin \theta_t \cos \theta_i - \sin \theta_t \cos \theta_t}{\sin \theta_t \cos \theta_i + \sin \theta_t \cos \theta_t} A_{\parallel} = \\
&= \frac{\sin \theta_t \sin \theta_i (\sin \theta_t \cos \theta_i - \sin \theta_t \cos \theta_t) + \cos \theta_t \cos \theta_i (\cos \theta_t \sin \theta_i - \sin \theta_t \cos \theta_i)}{\sin(\theta_i + \theta_t) \cos(\theta_i - \theta_t)} \times A_{\perp} \\
&\times A_{\parallel} = -\frac{\sin(\theta_i - \theta_t) \cos(\theta_i + \theta_t)}{\sin(\theta_i + \theta_t) \cos(\theta_i - \theta_t)} A_{\parallel} = -\frac{\operatorname{tg}(\theta_i - \theta_t)}{\operatorname{tg}(\theta_i + \theta_t)} A_{\parallel}.
\end{aligned}$$

алогично преобразуется второе из уравнений (1.119). В результате преобразований уравнения (1.118) и (1.119) приобретают обычно принятую форму:

$$\left. \begin{aligned}
T_{\parallel} &= \frac{2 \sin \theta_t \cos \theta_i}{\sin(\theta_i + \theta_t) \cos(\theta_i - \theta_t)} A_{\parallel}, \\
T_{\perp} &= \frac{2 \sin \theta_t \cos \theta_i}{\sin(\theta_i + \theta_t)} A_{\perp};
\end{aligned} \right\} \quad (1.120)$$

$$\left. \begin{aligned}
R_{\parallel} &= -\frac{\operatorname{tg}(\theta_i - \theta_t)}{\operatorname{tg}(\theta_i + \theta_t)} A_{\parallel}, \\
R_{\perp} &= -\frac{\sin(\theta_i - \theta_t)}{\sin(\theta_i + \theta_t)} A_{\perp}.
\end{aligned} \right\} \quad (1.121)$$

Так как углы θ_i и θ_t вещественны (случай полного внутреннего отражения пока исключаем), то тригонометрические функции, стоящие в правой части уравнений (1.120) и (1.121), также вещественны. Следовательно, фаза каждого компонента отраженной и прошедшей волн либо равна фазе соответствующего компонента падающей волны, либо отличается от нее на π . Так как знаки T_{\parallel} и T_{\perp} совпадают со знаками A_{\parallel} и A_{\perp} , фаза прошедшей волны равна фазе падающей. При нормальном падении волны угол $\theta_i = 0$, а, следовательно, и угол $\theta_t = 0$. При этом соотношения (1.20 и (1.121) принимают вид:

$$\left. \begin{aligned}
T_{\parallel} &= \frac{2}{n+1} A_{\parallel}, \\
T_{\perp} &= \frac{2}{n+1} A_{\perp},
\end{aligned} \right\} \quad (1.122)$$

$$\left. \begin{aligned}
R_{\parallel} &= -\frac{n-1}{n+1} A_{\parallel}, \\
R_{\perp} &= -\frac{n-1}{n+1} A_{\perp},
\end{aligned} \right\} \quad (1.123)$$

где $n = \frac{n_2}{n_1}$. При этом различие между параллельным и перпендикулярным компонентами исчезает. Из соотношений (1.113) следует, что фаза

отраженной волны отличается от падающей на π . С другой, стороны, в случае отраженной волны фаза зависит от относительных значений углов θ_i и θ_t . Если оптическая плотность второй среды больше, чем первой (т.е. $\varepsilon_2 > \varepsilon_1$), то $\theta_t < \theta_i$. При этом, согласно соотношениям (1.121), знаки R_{\parallel} и R_{\perp} отличны от знаков A_{\parallel} и A_{\perp} , т.е. фазы падающей и отраженной волн отличаются на π . При тех же обстоятельствах значение $\operatorname{tg}(\theta_i - \theta_t)$ положительно, но $\operatorname{tg}(\theta_i + \theta_t)$ в знаменателе выражения, определяющего величину R_{\parallel} , при $\theta_i + \theta_t > \frac{\pi}{2}$ становится отрицательным; в этом случае фазы компонентов R_{\parallel} и A_{\parallel} совпадают. Заметим, что при $\theta_i + \theta_t = \frac{\pi}{2}$ величина $R_{\parallel} = 0 \cdot A_{\parallel} = 0$.

1.10. Коэффициенты отражения и пропускания света на границе раздела двух сред

Интенсивность света (плотность потока энергии электромагнитного поля) в соответствии с выражением (1.86) равна

$$G = \frac{c}{4\pi} EH = \frac{c}{4\pi} \sqrt{\frac{\varepsilon}{\mu}} E^2 = \frac{cn}{4\pi\mu} E^2.$$

Предположим, что обе среды (однородные и изотропные) обладают нулевой проводимостью и, следовательно, совершенно прозрачны. Их магнитные проницаемости практически будут отличаться от единицы на пренебрежимо малые величины, что позволяет принять $\mu_1 = \mu_2 = 1$. При этом интенсивность света равна

$$G = \frac{cn}{4\pi} E^2. \quad (1.124)$$

Поэтому световой поток падающей волны, который падает на единицу площади поверхности раздела сред за одну секунду, равен

$$J^{(i)} = G^{(i)} \cos \theta_i = \frac{cn_1}{4\pi} |A|^2 \cos \theta_i.$$

Световой поток отраженной и преломленной волн, покидающей единицу площади поверхности раздела за одну секунду, определяется подобными же выражениями, а именно:

$$J^{(r)} = G^{(r)} \cos \theta_i = \frac{cn_1}{4\pi} |R|^2 \cos \theta_i;$$

$$J^{(t)} = G^{(t)} \cos \theta_t = \frac{cn_2}{4\pi} |T|^2 \cos \theta_t.$$

Отношения

$$\rho = \frac{J^{(r)}}{J^{(i)}} = \frac{|R|^2}{|A|^2} \quad (1.125)$$

и

$$\tau = \frac{J^{(t)}}{J^{(i)}} = \frac{n_2 \cos \theta_t}{n_1 \cos \theta_i} \frac{|T|^2}{|A|^2} \quad (1.126)$$

называют коэффициентами отражения и пропускания соответственно.

Здесь $|A|^2 = |A_{\parallel}|^2 + |A_{\perp}|^2$, $|R|^2 = |R_{\parallel}|^2 + |R_{\perp}|^2$, $|T|^2 = |T_{\parallel}|^2 + |T_{\perp}|^2$.

Используя выражения (1.120) и (1.121), получаем

$$\begin{aligned} |A|^2(\rho + \tau) &= |R|^2 + \frac{n_2 \cos \theta_t}{n_1 \cos \theta_i} |T|^2 = \frac{\operatorname{tg}^2(\theta_i - \theta_t)}{\operatorname{tg}^2(\theta_i + \theta_t)} |A_{\parallel}|^2 + \\ &+ \frac{\sin^2(\theta_i - \theta_t)}{\sin^2(\theta_i + \theta_t)} |A_{\perp}|^2 + 4 \frac{n_2 \cos \theta_t}{n_1 \cos \theta_i} \frac{\sin^2 \theta_t \cos^2 \theta_i}{\sin^2(\theta_i + \theta_t) \cos^2(\theta_i - \theta_t)} |A_{\parallel}|^2 + \\ &+ 4 \frac{n_2 \cos \theta_t}{n_1 \cos \theta_i} \frac{\sin^2 \theta_t \cos^2 \theta_i}{\sin^2(\theta_i + \theta_t)} |A_{\perp}|^2 = |A_{\parallel}|^2 + |A_{\perp}|^2. \end{aligned}$$

Отсюда следует, что $\rho + \tau = 1$. Этот результат непосредственно следует и из закона сохранения энергии.

1.11. Поляризация света на поверхности раздела двух сред

Отражение и пропускание света на поверхности раздела двух сред зависит от поляризации падающей волны. Определим коэффициенты отражения и пропускания света, поляризованного параллельно и перпендикулярно плоскости падения.

Пусть вектор \mathbf{E} падающей волны образует с плоскостью падения угол α_i . Тогда

$$A_{\parallel} = A \cos \alpha_i, \quad A_{\perp} = A \sin \alpha_i.$$

Пусть, далее,

$$J_{\parallel}^{(i)} = \frac{cn_1}{4\pi} |A_{\parallel}|^2 \cos \theta_i = J^{(i)} \cos^2 \alpha_i,$$

$$J_{\perp}^{(i)} = \frac{cn_1}{4\pi} |A_{\perp}|^2 \cos \theta_i = J^{(i)} \sin^2 \alpha_i;$$

$$J_{\parallel}^{(r)} = \frac{cn_1}{4\pi} |R_{\parallel}|^2 \cos \theta_i,$$

$$J_{\perp}^{(r)} = \frac{cn_1}{4\pi} |R_{\perp}|^2 \cos \theta_i.$$

Тогда

$$\begin{aligned}\rho &= \frac{J^{(r)}}{J^{(i)}} = \frac{J_{\parallel}^{(r)} + J_{\perp}^{(r)}}{J^{(i)}} = \frac{J_{\parallel}^{(r)}}{J_{\parallel}^{(i)}} \cos^2 \alpha_i + \frac{J_{\perp}^{(r)}}{J_{\perp}^{(i)}} \sin^2 \alpha_i = \\ &= \rho_{\parallel} \cos^2 \alpha_i + \rho_{\perp}^2 \sin^2 \alpha_i,\end{aligned}\tag{1.127}$$

где

$$\left. \begin{aligned}\rho_{\parallel} &= \frac{J_{\parallel}^{(r)}}{J_{\parallel}^{(i)}} = \frac{|R_{\parallel}|^2}{|A_{\parallel}|^2} = \frac{\operatorname{tg}^2(\theta_i - \theta_t)}{\operatorname{tg}^2(\theta_i + \theta_t)} \\ \rho_{\perp} &= \frac{J_{\perp}^{(r)}}{J_{\perp}^{(i)}} = \frac{|R_{\perp}|^2}{|A_{\perp}|^2} = \frac{\sin^2(\theta_i - \theta_t)}{\sin^2(\theta_i + \theta_t)}.\end{aligned}\right\}\tag{1.128}$$

Подобным же образом получаем

$$\tau = \frac{J^{(t)}}{J^{(i)}} = \tau_{\parallel} \cos^2 \alpha_i + \tau_{\perp} \sin^2 \alpha_i,\tag{1.129}$$

где

$$\left. \begin{aligned}\tau_{\parallel} &= \frac{J_{\parallel}^{(t)}}{J_{\parallel}^{(i)}} = \frac{n_2 \cos \theta_t}{n_1 \cos \theta_i} \frac{|T_{\parallel}|^2}{|A_{\parallel}|^2} = \frac{\sin 2\theta_i \sin 2\theta_t}{\sin^2(\theta_i + \theta_t) \cos^2(\theta_i - \theta_t)} \\ \tau_{\perp} &= \frac{J_{\perp}^{(t)}}{J_{\perp}^{(i)}} = \frac{n_2 \cos \theta_t}{n_1 \cos \theta_i} \frac{|T_{\perp}|^2}{|A_{\perp}|^2} = \frac{\sin 2\theta_i \sin 2\theta_t}{\sin^2(\theta_i + \theta_t)}.\end{aligned}\right\}\tag{1.130}$$

Легко убедиться, что и в этом случае

$$\rho_{\parallel} + \tau_{\parallel} = 1, \quad \rho_{\perp} + \tau_{\perp} = 1.$$

При нормальном падении волны различие между параллельным и перпендикулярным компонентами исчезает, при этом, используя соотношения (1.122), (1.123), (1.124) и (1.126), получаем

$$\left. \begin{aligned}\rho &= \left(\frac{n-1}{n+1}\right)^2 \\ \tau &= \frac{4n}{(n+1)^2}.\end{aligned}\right\}\tag{1.131}$$

Заметим, что $\lim_{n \rightarrow 1} \rho = 0$, $\lim_{n \rightarrow 1} \tau = 1$. Вполне очевидно, что аналогичные результаты получаются и для предельных значений R_{\parallel} , T_{\parallel} и R_{\perp} , T_{\perp} .

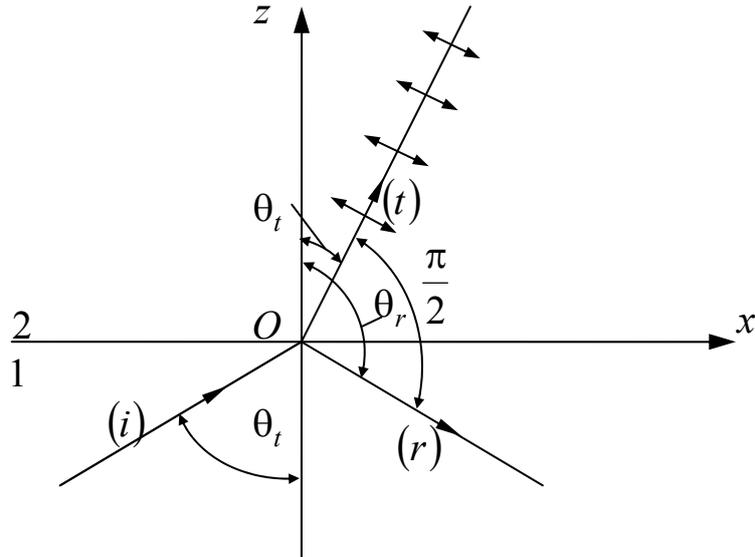


Рис. 1.12. К определению угла полной поляризации (угла Брюстера)

Весьма важное явление наблюдается при $\theta_i + \theta_t = \frac{\pi}{2}$. В этом случае $\text{tg}(\theta_i + \theta_t) = \infty$, а, следовательно, в соответствии с первым из выражений (1.128), величина $\rho_{\parallel} = 0$. Как следует из рис. 1.12, при $\theta_i + \theta_t = \frac{\pi}{2}$ отраженный и преломленный лучи перпендикулярны друг другу. Из закона преломления следует, что

$$n_2 \sin \theta_t = n_2 \sin \left(\frac{\pi}{2} - \theta_i \right) = n_2 \cos \theta_i = n_1 \sin \theta_i.$$

При этом

$$\text{tg} \theta_i = n. \quad (1.132)$$

Угол θ_i , определяемый этим выражением, называется углом полной поляризации или углом Брюстера. Его важность впервые была отмечена Давидом Брюстером в 1815 году. Если свет падает под этим углом, электрический вектор отраженной волны не имеет составляющей в плоскости падения.

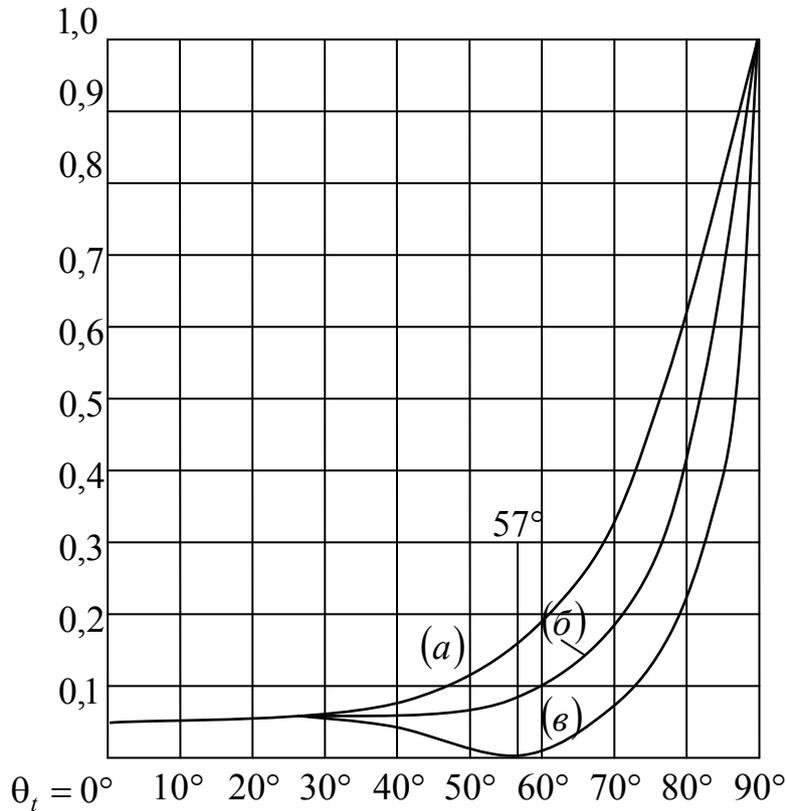


Рис. 1.13. Зависимость отражательной способности от угла падения:

$$(a) - \rho_{\perp}; (б) - \frac{1}{2}(\rho_{\parallel} + \rho_{\perp}); (в) - \rho_{\parallel}$$

Зависимость коэффициента отражения света поверхностью стекла (стекло марки БК6 с показателем преломления $n_d = 1,53998$) от угла падения θ_i показана на рис. 1.13. Здесь представлены кривые: (а) – $\rho_{\perp} = \rho_{\perp}(\theta_i)$; (б) – $\frac{1}{2}(\rho_{\parallel} + \rho_{\perp}) = \rho_{\parallel+\perp}(\theta_i)$; (в) – $\rho_{\parallel} = \rho_{\parallel}(\theta_i)$. Значения соответствующих величин, вычисленные с использованием формул (1.128), приведены в таблице:

θ_i	0	20°	45°	57°	70°	80°	90°
Кривая (а)	0,045	0,053	0,101	0,165	0,316	0,554	1,000
Кривая (б)	0,045	0,045	0,056	0,083	0,178	0,394	1,000
Кривая (в)	0,045	0,038	0,010	0	0,041	0,234	1,000

Для кривой (б) имеем

$$\rho = \frac{1}{2}(\rho_{\parallel} + \rho_{\perp}) = \rho_{\parallel} \cos^2 \alpha_i + \rho_{\perp} \sin^2 \alpha_i = \rho_{\parallel} + (\rho_{\perp} - \rho_{\parallel}) \sin^2 \alpha_i.$$

Отсюда следует, что

$$\frac{1}{2}(\rho_{\perp} - \rho_{\parallel}) = (\rho_{\perp} - \rho_{\parallel}) \sin^2 \alpha_i.$$

При этом $\sin^2 \alpha_i = \frac{1}{2}$. Таким образом, кривая (б) на рис. 1.13 соответствует углу $\alpha_i = 45^\circ$.

Направление колебаний в естественном свете быстро изменяется беспорядочным, случайным образом. Соответствующее значение коэффициента отражения ρ можно получить путем усреднения по всем направлениям. Среднее значение функции $\sin^2 \alpha$ равно:

$$\begin{aligned} \sin^2 \alpha_{cp} &= \frac{2}{\pi} \int_0^{\frac{\pi}{2}} \sin^2 \alpha d\alpha = -\frac{2}{\pi} \int_0^{\frac{\pi}{2}} \sqrt{1 - \cos^2 \alpha} d(\cos \alpha) = \frac{2}{\pi} \int_0^1 \sqrt{1 - x^2} dx = \\ &= \frac{2}{\pi} \left(\frac{1}{2} x \sqrt{1 - x^2} + \frac{1}{2} \arcsin x \right) \Big|_0^1 = \frac{2}{\pi} \left(0 + \frac{1}{2} \cdot \frac{\pi}{2} \right) = \frac{1}{2}. \end{aligned}$$

Среднее значение функции $\cos^2 \alpha$ равно: $\cos^2 \alpha_{cp} = 1 - \sin^2 \alpha_{cp} = \frac{1}{2}$.

Так как средние значения $\sin^2 \alpha$ и $\cos^2 \alpha$ равны $\frac{1}{2}$, то для средних значений $J_{\parallel}^{(i)}$ и $J_{\perp}^{(i)}$ получаем

$$J_{\parallel}^{(i)} = J_{\perp}^{(i)} = \frac{1}{2} J^{(i)}. \quad (1.133)$$

Однако, для отраженного света оба компонента в общем случае неодинаковы. Действительно, используя соотношения (1.133), имеем

$$\begin{aligned} J_{\parallel}^{(r)} &= \frac{1}{2} \frac{J_{\parallel}^{(r)}}{J_{\parallel}^{(i)}} J^{(i)} = \frac{1}{2} \rho_{\parallel} J^{(i)}, \\ J_{\perp}^{(r)} &= \frac{1}{2} \frac{J_{\perp}^{(r)}}{J_{\perp}^{(i)}} J^{(i)} = \frac{1}{2} \rho_{\perp} J^{(i)}. \end{aligned}$$

При этом коэффициент отражения определится выражением вида:

$$\rho = \frac{J^{(r)}}{J^{(i)}} = \frac{J_{\parallel}^{(r)} + J_{\perp}^{(r)}}{J^{(i)}} = \frac{1}{2} (\rho_{\parallel} + \rho_{\perp}), \quad (1.134)$$

т.е. и в этом случае зависимость коэффициента отражения от угла θ_i будет описываться кривой (б) на рис. 1.13.

В общем случае отраженный свет частично поляризован. Степень его поляризации P можно определить выражением вида:

$$P = \left| \frac{\rho_{\parallel} - \rho_{\perp}}{\rho_{\parallel} + \rho_{\perp}} \right|.$$

Учитывая формулу (1.134), это выражение можно представить в виде:

$$P = \frac{1}{2\rho} \{ |\rho_{\parallel} - \rho_{\perp}| \}.$$

Выражением в фигурных скобках определяют иногда поляризованную часть отраженного света. Аналогичные соотношения можно получить и для проходящего света. Можно показать, что для естественного света также справедливо равенство: $\rho + \tau = 1$. В случае линейно поляризованного падающего света отраженный и прошедший свет остается линейно поляризованным, так как их фазы либо не изменяются, либо изменяются на π . Однако, направления колебаний в отраженном и проходящем свете изменяются относительно направления колебаний в падающем свете в противоположные стороны. Покажем это.

Угол α между плоскостью колебаний и плоскостью падения, называемый азимутом колебания, будем считать положительным, если плоскость колебаний поворачивается по часовой стрелке вокруг направления распространения, как показано на рис. 1.14. Можно предположить, что азимут колебания изменяется в пределах от $-\frac{\pi}{2}$ до $\frac{\pi}{2}$.

Для падающей, отраженной и прошедшей волн электрического поля имеем

$$\operatorname{tg} \alpha_i = \frac{A_{\perp}}{A_{\parallel}}, \operatorname{tg} \alpha_r = \frac{R_{\perp}}{R_{\parallel}}, \operatorname{tg} \alpha_t = \frac{T_{\perp}}{T_{\parallel}}.$$

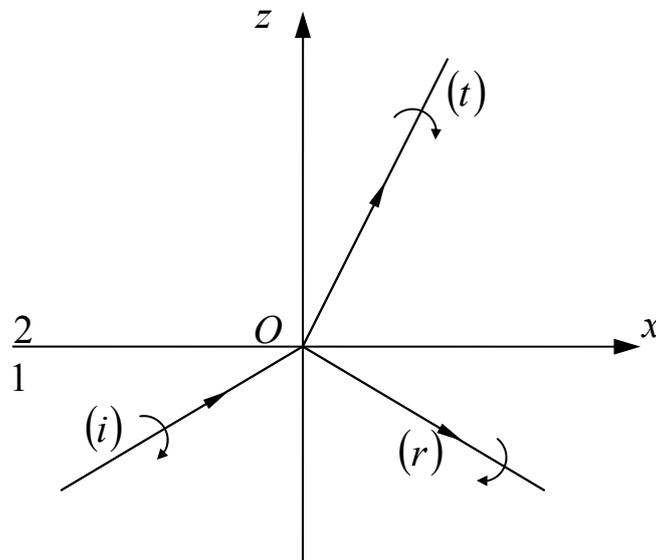


Рис. 1.14. К определению знаков азимутальных углов

Используя формулы Френеля (1.120) и (1.121), получаем

$$\operatorname{tg} \alpha_r = \frac{\cos(\theta_i - \theta_t)}{\cos(\theta_i + \theta_t)} \operatorname{tg} \alpha_i, \quad (1.135)$$

$$\operatorname{tg} \alpha_t = \cos(\theta_i - \theta_t) \operatorname{tg} \alpha_i. \quad (1.136)$$

Так как $0 \leq \theta_i \leq \frac{\pi}{2}$, $0 \leq \theta_t < \frac{\pi}{2}$ (при $n > 1$), то

$$|tg\alpha_r| \geq |tg\alpha_i|, \quad (1.137)$$

$$|tg\alpha_t| \leq |tg\alpha_i|. \quad (1.138)$$

Знак равенства в соотношении (1.137) справедлив лишь при нормальном или скользящем падении света ($\theta_i = \theta_t = 0$ или $\theta_i = \frac{\pi}{2}$), а в соотношении (1.138) – лишь при нормальном падении. Из неравенств (1.137) и (1.138) следует, что при отражении угол между плоскостью колебаний и плоскостью падения увеличивается, а при преломлении он уменьшается.

В том случае, когда угол падения равен углу Брюстера, т.е. при $\theta_i + \theta_t = \frac{\pi}{2}$, в соответствии с формулой (1.135) $tg\alpha_r = \infty$ при любом значении угла α_i . Из закона Брюстера следует, что свет поляризуется при падении его на границу раздела двух сред под углом Брюстера. Однако, доля отраженного при этом света сравнительно мала. Поляризовать падающий свет без изменения направления его распространения можно с помощью стопы тонких плоскопараллельных пластинок.

Используя выражения (1.130), находим, что отношение интенсивности двух компонент волны после прохождения через обе поверхности пластинки определяется формулой вида:

$$\left(\frac{\tau_{\perp}}{\tau_{\parallel}} \right)^2 = \cos^4(\theta_i - \theta_t) < 1.$$

Таким образом, при выходе из пластинки параллельный компонент преобладает над перпендикулярным, причем степень поляризации тем больше чем больше угол θ_i . Если угол θ_i равен углу Брюстера, то

$\theta_t = \frac{\pi}{2} - \theta_i$, а $tg\theta_i = n$; при этом получаем

$$\left(\frac{\tau_{\perp}}{\tau_{\parallel}} \right)^2 = \sin^4 2\theta_i = \left(\frac{2tg\theta_i}{1+tg^2\theta_i} \right)^4 = \left(\frac{2n}{1+n^2} \right)^4.$$

При $n_d = 1,56889$ (стекло марки БК10) получаем $\left(\frac{\tau_{\perp}}{\tau_{\parallel}} \right)^2 \cong 0,675$.

Следовательно, если свет проходит, например, через пять пластинок, то это отношение равно $0,657^5 \cong 0,14$.

1.12. Полное внутреннее отражение

При распространении света из оптически более плотной среды в среду с меньшей оптической плотностью, т.е. когда

$$n = n_{12} = \frac{n_2}{n_1} = \sqrt{\frac{\varepsilon_2 \mu_2}{\varepsilon_1 \mu_1}} < 1,$$

$\sin \theta_t = 1$ при $\sin \theta_i = n_{12}$, при этом $\theta_t = 90^\circ$, а направление распространения света совпадает с касательной к поверхности раздела. Если величина угла θ_i превышает предельное значение ($\sin \theta_i > n_{12}$), то свет не входит во вторую среду. Весь падающий свет отражается обратно в первую среду, т.е. в этом случае наблюдается полное внутреннее отражение. Таким образом, поток энергии через границу раздела двух сред отсутствует, однако, тем не менее, электромагнитное поле во второй среде не равно нулю. Действительно, если в фазовом множителе прошедшей волны, определяемом выражением (1.113), положить

$$\sin \theta_t = \frac{1}{n} \sin \theta_i, \quad \cos \theta_t = \pm i \sqrt{\frac{\sin^2 \theta_i}{n^2} - 1}, \quad (1.139)$$

то получим

$$\exp(-i\tau_t) = \exp\left[-i\omega\left(t - \frac{x \sin \theta_i}{nV_2}\right)\right] \exp\left[\mp \frac{\omega z}{V_2} \sqrt{\frac{\sin^2 \theta_i}{n^2} - 1}\right],$$

где $n = n_{12} = \frac{n_2}{n_1}$. В результате имеем

$$E^{(t)} = T \exp(-i\tau_t) = \tilde{T} \exp\left[-i\omega\left(t - \frac{x \sin \theta_i}{nV_2}\right)\right], \quad (1.140)$$

где $\tilde{T} = T \exp\left[\mp \frac{\omega z}{V_2} \sqrt{\frac{\sin^2 \theta_i}{n^2} - 1}\right]$ при $n < 1$.

Выражение (1.140) описывает неоднородную волну, которая распространяется вдоль поверхности раздела сред в плоскости падения (т.е. в направлении оси x) и меняется экспоненциально с изменением расстояния z от этой поверхности. Напомним, что волна называется неоднородной, если поверхности постоянной амплитуды не совпадают с поверхностями постоянной фазы.

При положительном знаке перед корнем в выражении (1.140) при увеличении расстояния z от поверхности амплитуда росла бы неограниченно, что противоречит опыту. При отрицательном знаке амплитуда очень быстро уменьшается с глубиной проникновения z ,

причем $\frac{\omega z}{V_2} = 2\pi\nu\frac{n}{c}z = 2\pi\frac{n}{\lambda_0}z = 2\pi\frac{z}{\lambda}$, т.е. эффективная глубина проникновения порядка длины волны.

Экспериментальная проверка наличия возмущения во второй (менее плотной) среде представляет довольно трудную задачу, поскольку любое устройство, используемое для обнаружения возмущения, будет нарушать граничные условия. Грубое подтверждение можно получить, если поместить вторую преломляющую среду на расстоянии около четверти длины волны от поверхности раздела, на которой происходит полное внутреннее отражение, и наблюдать проникновение излучения во вторую среду. И, тем не менее, проникновение поля во вторую (менее плотную) среду наблюдалось в очень тонких опытах Л.И. Мандельштама еще 1914 году.

Чтобы применить формулы Френеля (1.121) к случаю полного внутреннего отражения, перепишем их в виде:

$$\left. \begin{aligned} R_{\parallel} &= -\frac{\sin\theta_i \cos\theta_i - \sin\theta_t \cos\theta_t}{\sin\theta_i \cos\theta_i + \sin\theta_t \cos\theta_t} A_{\parallel}, \\ R_{\perp} &= -\frac{\sin\theta_i \cos\theta_t - \sin\theta_t \cos\theta_i}{\sin\theta_i \cos\theta_t + \sin\theta_t \cos\theta_i} A_{\perp}. \end{aligned} \right\} \quad (1.141)$$

Подставив в эти выражения значения величин, определяемых формулами (1.139) (при положительном знаке перед квадратным корнем), получаем

$$\left. \begin{aligned} R_{\parallel} &= -\frac{n^2 \cos\theta_i - i\sqrt{\sin^2\theta_i - n^2}}{n^2 \cos\theta_i + i\sqrt{\sin^2\theta_i - n^2}} A_{\parallel}, \\ R_{\perp} &= \frac{\cos\theta_i - i\sqrt{\sin^2\theta_i - n^2}}{\cos\theta_i + i\sqrt{\sin^2\theta_i - n^2}} A_{\perp}. \end{aligned} \right\} \quad (1.142)$$

Следовательно,

$$|R_{\parallel}| = |A_{\parallel}|, \quad |R_{\perp}| = |A_{\perp}|, \quad (1.143)$$

т.е. для каждого компонента интенсивность света, отраженного при полном внутреннем отражении, равна интенсивности падающего света. Таким образом, хотя во второй среде поле отлично от нуля, но поток энергии через поверхность отсутствует.

Можно показать, что хотя в общем случае компонент вектора Умова-Пойнтинга в направлении, нормальном к границе, конечен, его значение, усредненное по времени, равно нулю. Это означает, что не существует постоянного потока во вторую среду, а энергия течет туда и обратно.

И, наконец, определим изменение фаз компонент отраженной и падающей волн. На основании соотношений (1.143) можно принять, что

$$\frac{R_{\parallel}}{A_{\parallel}} = \exp(i\delta_{\parallel}), \quad \frac{R_{\perp}}{A_{\perp}} = \exp(i\delta_{\perp}).$$

Согласно соотношениям (1.142) каждое из отношений $\frac{R_{\parallel}}{A_{\parallel}}$ и $\frac{R_{\perp}}{A_{\perp}}$

определяется выражением, имеющим форму $\psi(\psi^*)^{-1}$. Следовательно, если α – аргумент ψ (т.е. $\psi = a \exp(i\alpha)$, где a и α вещественны), то $\exp(i\delta) = \psi(\psi^*)^{-1} = \exp(2i\alpha)$, т.е. $\operatorname{tg} \frac{\delta}{2} = \operatorname{tg} \alpha$, а поэтому

$$\left. \begin{aligned} \operatorname{tg} \frac{1}{2} \delta_{\parallel} &= -\frac{\sqrt{\sin^2 \theta_i - n^2}}{n^2 \cos \theta_i}, \\ \operatorname{tg} \frac{1}{2} \delta_{\perp} &= -\frac{\sqrt{\sin^2 \theta_i - n^2}}{\cos \theta_i}. \end{aligned} \right\} \quad (1.144)$$

Отсюда следует, что оба компонента испытывают скачок фазы разной величины. Вследствие этого линейно поляризованный свет при полном внутреннем отражении оказывается поляризованным эллиптически. При этом для относительной разности фаз $\delta = \delta_{\perp} - \delta_{\parallel}$ имеем

$$\operatorname{tg} \frac{1}{2} \delta = \frac{\operatorname{tg} \frac{1}{2} \delta_{\perp} - \operatorname{tg} \frac{1}{2} \delta_{\parallel}}{1 + \operatorname{tg} \frac{1}{2} \delta_{\perp} \operatorname{tg} \frac{1}{2} \delta_{\parallel}}.$$

Подставив в это выражение соотношения (1.144) и преобразовав, получаем

$$\operatorname{tg} \frac{1}{2} \delta = \frac{\cos \theta_i \sqrt{\sin^2 \theta_i - n^2}}{\sin^2 \theta_i}. \quad (1.145)$$

Это выражение обращается в нуль при скольльзящем падении $\left(\theta_i = \frac{\pi}{2} \right)$

и при падении света под критическим углом $\tilde{\theta}_i$ ($\sin \tilde{\theta}_i = n$). Между этими двумя углами лежит угол, соответствующий максимуму относительной разности фаз. Он определяется уравнением:

$$\frac{d}{d\theta_i} \left(\operatorname{tg} \frac{1}{2} \delta \right) = \frac{2n^2 - (1+n^2) \sin^2 \theta_i}{\sin^3 \theta_i \sqrt{\sin^2 \theta_i - n^2}} = 0.$$

Положив числитель дроби равным нулю, получаем

$$\sin^2 \theta_i = \frac{2n^2}{1+n^2}. \quad (1.146)$$

Подставив это соотношение в выражение (1.145), получаем формулу, определяющую максимальную величину δ_m относительной разности фаз, в виде:

$$\operatorname{tg} \frac{1}{2} \delta_m = \frac{1-n^2}{2n}. \quad (1.147)$$

Еще Френелем было показано, что при полном внутреннем отражении можно из линейно поляризованного света получить свет, поляризованный по кругу. Выбирая направление поляризации падающей волны под углом 45° к плоскости падения (т.е. $\alpha_i = 45^\circ$), получаем равные амплитуды компонент: $|A_{\parallel}| = |A_{\perp}|$. Тогда в соответствии с равенствами (1.143) имеем: $|R_{\parallel}| = |R_{\perp}|$. Затем подбираем величины n и θ_i таким образом, чтобы относительная разность фаз δ была равна 90° . Для получения такого значения δ при одном отражении необходимо, чтобы в соответствии с соотношением (1.147)

$$\operatorname{tg} \frac{\pi}{4} = 1 = \frac{1-n^2}{2n}.$$

При этом $n = \sqrt{2} - 1 = 0,414$. Отсюда следует, что показатель преломления $n_{21} = \frac{1}{n}$ более плотной среды относительно менее плотной должен быть, по крайней мере, не менее 2,41. Это значение показателя преломления довольно велико, хотя непоглощающие вещества, показатель преломления которых приближается или даже превышает его, существуют.

Френель использовал два полных внутренних отражения в стекле. Когда $n_{21} = 1,51$, то, согласно соотношениям (1.146) и (1.147), максимальная относительная разность фаз $\delta_m = 45^\circ 56'$ получается при угле падения θ_i , равном $51^\circ 20'$. Поэтому в соответствии с уравнением (1.145) значение $\delta = 45^\circ$ можно получить при одном из следующих углов падения:

$$\theta_i = 48^\circ 37', \quad \theta_i = 54^\circ 37'.$$

При этом разность фаз в 90° возникает в результате двух последовательных полных внутренних отражений при любом из приведенных углов. Для получения такой разности фаз применяется стеклянный блок, известный как ромб Френеля, главное сечение которого показано на рис. 1.15.

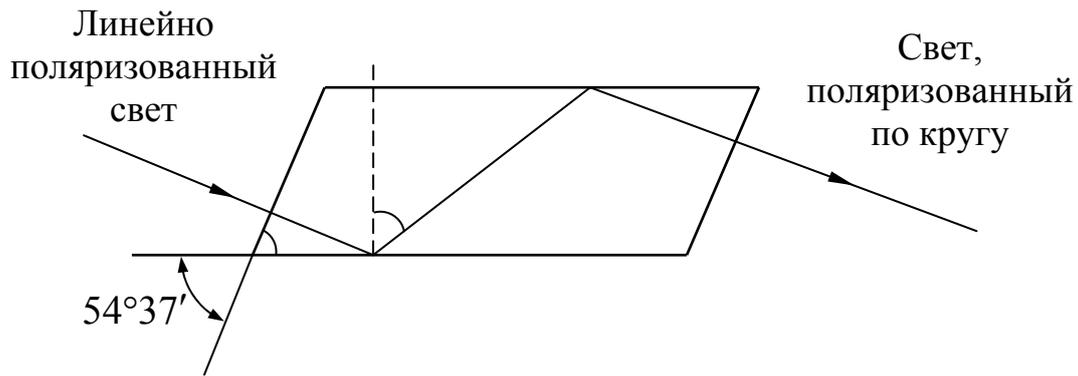


Рис. 1.15. Ромб Френеля

Вполне очевидно, что ромб Френеля можно использовать для получения эллиптически поляризованного света. В этом случае азимут падающего (линейно поляризованного) света должен отличаться от 45° . Можно также с помощью ромба Френеля получить линейно поляризованный свет из света, поляризованного эллиптически.

Измерение предельного (критического) значения угла θ_i позволяет удобно и достаточно точно определять показатель преломления, используя равенство $n = \sin \theta_i$. Приборы, используемые для этой цели, называются рефрактометрами.

1.13. Двойное лучепреломление света

При падении световой волны на границу раздела двух изотропных сред от границы во второй среде распространяется одна волна, волновой вектор которой лежит в плоскости падения и составляет с нормалью к границе угол φ_2 , определяемый законом преломления: $n_1 \sin \varphi_1 = n_2 \sin \varphi_2$. Если среда анизотропна, то в ней в общем случае возникают две волны, распространяющиеся от границы в разных направлениях и с разными скоростями. Это явление называется двойным лучепреломлением. В 1669 году датчанин Эразм Бартолин опубликовал работу «*Experimenta crystalli islandici disdiaclastici, quibus mira et insolita refraction detegitur*» («Опыты с кристаллами исландского известкового шпата, которые обнаруживают удивительное и странное преломление»). В этой работе, как известно, речь идет о том, что если луч света падает на поверхность исландского шпата, то он, преломляясь, раздваивается. Бартолин, кроме того, заметил, что преломление одного из двух лучей («обыкновенного» луча) в точке падения подчиняется закону преломления, тогда как преломление второго луча («необыкновенного» луча) этому закону не подчиняется. Датский ученый открыл также, что в кристалле исландского шпата существует направление, установленное им практически точно, вдоль которого луч не раздваивается. Явление двойного лучепреломления было впоследствии обстоятельно исследовано Гюйгенсом. Гюйгенс дал объяснение двойного лучепреломления на основе гипотезы о том, что падающая на границу

волна порождает в кристалле элементарные вторичные волны двух видов: сферические (обыкновенные) и эллиптические (необыкновенные), скорость которых зависит от направления.

При распространении световой волны в анизотропном кристалле на электроны действует внешняя сила, направленная вдоль вектора \mathbf{E} волны. Однако, смещение электронов происходит не в направлении этой внешней силы, а в направлении, определяемом структурой кристалла. В результате вектор электрической поляризации среды \mathbf{P} оказывается не параллельным вектору \mathbf{E} . В свою очередь вектор электрической индукции $\mathbf{D}=\mathbf{E}+4\pi\mathbf{P}$ также не параллелен вектору \mathbf{E} . Рассматривая вещество как некое непрерывное распределение, взаимодействующее с полем, для достаточно слабых полей можно принять, что величина \mathbf{P} пропорциональна величине \mathbf{E} , т.е. $\mathbf{P}=\eta\mathbf{E}$. Множитель η называют диэлектрической восприимчивостью. В результате получаем, что диэлектрическая проницаемость связана с диэлектрической восприимчивостью соотношением $\varepsilon=1+4\pi\eta$. При этом $\mathbf{D}=\varepsilon\mathbf{E}$. Простейшее предположение, служащее первым шагом учета атомной структуры вещества, заключается в рассмотрении вещества как совокупности определенных физических объектов – молекул, которые могут поляризоваться и, следовательно, приобретать под действием внешнего поля электрический и магнитный моменты. Поскольку оптические материалы, как правило, представляют собой немагнитные вещества, их магнитную проницаемость принято считать равной $\mu=1$, т.е. принято считать вектор магнитной индукции $\mathbf{B}=\mathbf{H}$. В первом приближении принято считать, что компоненты вектора электрической индукции в электрически анизотропной среде являются линейными функциями компонент поля:

$$\left. \begin{aligned} D_x &= \varepsilon_{xx}E_x + \varepsilon_{xy}E_y + \varepsilon_{xz}E_z, \\ D_y &= \varepsilon_{yx}E_x + \varepsilon_{yy}E_y + \varepsilon_{yz}E_z, \\ D_z &= \varepsilon_{zx}E_x + \varepsilon_{zy}E_y + \varepsilon_{zz}E_z. \end{aligned} \right\} \quad (1.148)$$

Девять величин $\varepsilon_{xx}, \varepsilon_{yy}, \dots$ являются постоянными среды и составляют тензор диэлектрической проницаемости. Следовательно, вектор \mathbf{D} равен произведению этого тензора на вектор \mathbf{E} . Заметим, что из условия симметричности тензора диэлектрической проницаемости следует теорема, выражающая закон сохранения энергии в дифференциальной форме. При этом из девяти компонент тензора только шесть независимы.

Симметричность тензора ε позволяет привести выражение, определяющее плотность электрической энергии w_e , к такой форме, при которой сохраняются лишь члены, содержащие компоненты в суммарной степени, не выше второй. Рассмотрим в пространстве x, y, z уравнение поверхности второго порядка вида:

$$\varepsilon_{xx}x^2 + \varepsilon_{yy}y^2 + \varepsilon_{zz}z^2 + 2\varepsilon_{yz}yz + 2\varepsilon_{xz}xz + 2\varepsilon_{xy}xy = const. \quad (1.149)$$

Левая часть уравнения (1.149) должна иметь положительную квадратичную форму, потому что при замене координат x , y и z компонентами вектора \mathbf{E} она становится равной $8\pi w_e$, а плотность энергии w_e должна быть положительной для любого значения вектора поля. Поэтому уравнение (1.149) описывает эллипсоид, который всегда можно привести к его главным осям. Таким образом, существует система координат, связанная с кристаллом, в которой уравнение эллипсоида имеет вид:

$$\varepsilon_x x^2 + \varepsilon_y y^2 + \varepsilon_z z^2 = const. \quad (1.150)$$

В этой системе главных диэлектрических осей материальные уравнения принимают простую форму:

$$D_x = \varepsilon_x E_x, \quad D_y = \varepsilon_y E_y, \quad D_z = \varepsilon_z E_z. \quad (1.151)$$

Величины $\varepsilon_x, \varepsilon_y, \varepsilon_z$ называются главными диэлектрическими проницаемостями. Из приведенных формул непосредственно следует, что векторы \mathbf{D} и \mathbf{E} всегда имеют различные направления, если только направление вектора \mathbf{E} не совпадает с одной из главных осей или если все главные диэлектрические проницаемости не равны друг другу. В последнем случае ($\varepsilon_x = \varepsilon_y = \varepsilon_z$) эллипсоид вырождается в сферу.

В зависимости от соотношения между главными диэлектрическими проницаемостями $\varepsilon_x, \varepsilon_y, \varepsilon_z$ все кристаллы делятся на три группы: изотропные, одноосные и двуосные. Изотропными называются кристаллы, у которых все три главные диэлектрические проницаемости одинаковы:

$$\varepsilon_x = \varepsilon_y = \varepsilon_z. \quad (1.152)$$

Если одинаковы две из трех главных диэлектрических проницаемостей, т.е. если

$$\varepsilon_x = \varepsilon_y \neq \varepsilon_z, \quad (1.153)$$

то кристалл называется одноосным. У двуосного кристалла все три главные диэлектрические проницаемости различны:

$$\varepsilon_x \neq \varepsilon_y \neq \varepsilon_z. \quad (1.154)$$

В области прозрачности (спектрального пропускания) кристаллы характеризуют также показателями преломления. Изотропный кристалл характеризуется одним показателем преломления: $n = \sqrt{\varepsilon}$.

Одноосный кристалл имеет два главных показателя преломления: $n_0 = \sqrt{\varepsilon_x} = \sqrt{\varepsilon_y}$, $n_e = \sqrt{\varepsilon_z}$. И, наконец, двуосный кристалл имеет три главных показателя преломления: $n_x = \sqrt{\varepsilon_x}$, $n_y = \sqrt{\varepsilon_y}$, $n_z = \sqrt{\varepsilon_z}$.

В зависимости от соотношения между главными показателями преломления одноосные кристаллы делятся на положительные и

отрицательные. Положительными принято называть кристаллы, у которых $n_o < n_e$, а отрицательными – кристаллы, у которых $n_o > n_e$. В табл. 1.1–1.3 приведены показатели преломления некоторых кристаллов трех названных групп.

Таблица 1.1. Показатели преломления изотропных кристаллов

Кристалл	n
CdTe	2,69
NaCl	1,544
Алмаз	2,417
Плавиновый шпат	1,392
GaAs	3,40

Таблица 1.2. Показатели преломления одноосных кристаллов

Кристалл	n_o	n_e
<i>положительные</i>		
Лед	1,309	1,310
Кварц	1,544	1,553
BeO	1,717	1,732
Циркон	1,923	1,968
Рутил	2,616	2,903
ZnS	2,354	2,358
<i>отрицательные</i>		
(NH ₄)H ₂ PO ₄ (ADP)	1,552	1,478
Берилл	1,598	1,590
KH ₂ PO ₄ (KDP)	1,507	1,467
NaNO ₃	1,587	1,336
Кальцит	1,658	1,486
Турмалин	1,638	1,618
LiNbO ₃	2,300	2,208
BaTiO ₃	2,416	2,364
Прустит	3,019	2,739

Таблица 1.3. Показатели преломления двуосных кристаллов

Кристалл	n_x	n_y	n_z
Гипс	1,520	1,523	1,530
Полевой шпат	1,522	1,526	1,530
Слюда	1,552	1,582	1,588
Топаз	1,619	1,620	1,627
NaNO ₂	1,344	1,411	1,651
SbSi	2,7	3,2	3,8
YAlO ₃	1,923	1,938	1,947

Основная особенность распространения света в анизотропном кристалле состоит в том, что световая волна с произвольным состоянием поляризации распадается на две линейно поляризованные волны с ортогональными направлениями поляризации, бегущие с разными фазовыми скоростями.

Особенности преломления света на границе анизотропной среды можно использовать для получения поляризованного света. Вместо отдельного кристалла для этой цели более удобными оказываются их комбинации, называемые поляризационными призмами. Наиболее подходящий материал для их изготовления – исландский шпат, у которого сравнительно велико различие между обыкновенным и необыкновенным показателями преломления ($n_o = 1,658$; $n_e = 1,486$), реже применяют кварц ($n_o = 1,544$; $n_e = 1,553$). Первая поляризационная призма была изобретена В. Николем в 1828 году. В настоящее время существует много модификаций призмы Николя. Однако, призма Николя стала настолько популярной, что сегодня «николь» и «поляризационная призма» – синонимы.

В настоящее время для получения поляризованного света широко применяют так называемые поляроидные пленки. Их действие основано на свойстве, известном как дихроизм. Вещества, обладающие этим свойством, имеют различные коэффициенты поглощения для света, поляризованного в различных направлениях. Например, можно изготовить пленки из поливинилового спирта с внедренным йодом, которые пропускают почти 80% света, поляризованного в одной плоскости, и менее 1% света, поляризованного в перпендикулярном направлении.

2. Основы геометрической оптики

2.1. Основное уравнение геометрической оптики

Рассмотрим гармоническое электромагнитное поле общего вида

$$\left. \begin{aligned} \mathbf{E}(\mathbf{r}, t) &= \mathbf{E}_0(\mathbf{r}) \exp(-i\omega t), \\ \mathbf{H}(\mathbf{r}, t) &= \mathbf{H}_0(\mathbf{r}) \exp(-i\omega t), \end{aligned} \right\} \quad (2.1)$$

где вектора $\mathbf{E}_0(\mathbf{r})$ и $\mathbf{H}_0(\mathbf{r})$ определяются комплексными векторными функциями положения, при этом вещественные части стоящих справа выражений описывают физические поля. Подставив выражения (2.1) в уравнения Максвелла (1.14)–(1.17), в среде, свободной от зарядов и токов, получаем

$$\operatorname{rot} \mathbf{E}_0 - ik_0 \mu \mathbf{H}_0 = 0, \quad (2.2)$$

$$\operatorname{rot} \mathbf{H}_0 + ik_0 \varepsilon \mathbf{E}_0 = 0, \quad (2.3)$$

$$\operatorname{div} \varepsilon \mathbf{E}_0 = 0, \quad (2.4)$$

$$\operatorname{div} \mu \mathbf{H}_0 = 0. \quad (2.5)$$

При выводе этих уравнений были использованы материальные уравнения $\mathbf{D} = \varepsilon \mathbf{E}$, $\mathbf{B} = \mu \mathbf{H}$ и соотношения $k_0 = \frac{\omega}{c} = \frac{2\pi}{\lambda_0}$, где λ_0 – длина

волны в вакууме. Однородную плоскую волну, распространяющуюся в среде с показателем преломления $n = \sqrt{\varepsilon \mu}$ в направлении, определяемом единичным вектором \mathbf{s} , можно представить следующими векторными функциями:

$$\left. \begin{aligned} \mathbf{E}_0(\mathbf{r}) &= \mathbf{e} \exp(ik_0 n \mathbf{s} \cdot \mathbf{r}), \\ \mathbf{H}_0(\mathbf{r}) &= \mathbf{h} \exp(ik_0 n \mathbf{s} \cdot \mathbf{r}), \end{aligned} \right\} \quad (2.6)$$

где \mathbf{e} и \mathbf{h} – постоянные векторы, в общем случае комплексные. На достаточно больших расстояниях от малого источника поля $\mathbf{s} \cdot \mathbf{r} \approx r$. Поэтому разумно предположить, что в этих случаях поле общего типа можно представить в виде:

$$\left. \begin{aligned} \mathbf{E}_0(\mathbf{r}) &= \mathbf{e}(\mathbf{r}) \exp[ik_0 L(\mathbf{r})], \\ \mathbf{H}_0(\mathbf{r}) &= \mathbf{h}(\mathbf{r}) \exp[ik_0 L(\mathbf{r})], \end{aligned} \right\} \quad (2.7)$$

где $L(\mathbf{r})$ – «оптический путь» – вещественная скалярная функция положения; $\mathbf{e}(\mathbf{r})$ и $\mathbf{h}(\mathbf{r})$ – векторные функции положения, обязательно комплексные, если необходимо учесть все возможные состояния поляризации излучения.

Применив к выражениям (2.7) хорошо известные векторные тождества, получаем

$$\operatorname{rot} \mathbf{E}_0 = (\operatorname{rot} \mathbf{e} + ik_0 \operatorname{grad} L \times \mathbf{e}) \exp(ik_0 L), \quad (2.8)$$

$$\operatorname{div} \varepsilon \mathbf{E}_0 = (\varepsilon \operatorname{div} \mathbf{e} + \mathbf{e} \operatorname{grad} \varepsilon + ik_0 \varepsilon \operatorname{grad} L) \exp(ik_0 L). \quad (2.9)$$

Аналогичные выражения получаем для $rot \mathbf{H}_0$ и для $div \mu \mathbf{H}_0$. Полученные выражения позволяют преобразовать уравнения (2.2)–(2.5) к виду:

$$grad L \times \mathbf{e} - \mu \mathbf{h} = -\frac{1}{ik_0} rot \mathbf{e}, \quad (2.10)$$

$$grad L \times \mathbf{h} + \varepsilon \mathbf{e} = -\frac{1}{ik_0} rot \mathbf{h}, \quad (2.11)$$

$$\mathbf{e} grad L = -\frac{1}{ik_0} (\mathbf{e} grad \ln \varepsilon + div \mathbf{e}), \quad (2.12)$$

$$\mathbf{h} grad L = -\frac{1}{ik_0} (\mathbf{h} grad \ln \mu + div \mathbf{h}). \quad (2.13)$$

Пусть $\lambda_0 \rightarrow 0$. При этом $k_0 \rightarrow \infty$. Если в этом случае выражения, которые умножаются на $\frac{1}{ik_0}$, не будут чрезвычайно большими, то правыми частями уравнений (2.10)–(2.13) можно пренебречь и записать их в виде:

$$grad L \times \mathbf{e} - \mu \mathbf{h} = 0, \quad (2.14)$$

$$grad L \times \mathbf{h} + \varepsilon \mathbf{e} = 0, \quad (2.15)$$

$$\mathbf{e} grad L = 0, \quad (2.16)$$

$$\mathbf{h} grad L = 0. \quad (2.17)$$

Легко видеть, что уравнения (2.16) и (2.17) получаются из уравнений (2.14) и (2.15), если умножить их скалярно на $grad L$.

Из уравнения (2.14) следует, что $\mathbf{h} = \frac{1}{\mu} grad L \times \mathbf{e}$. При этом уравнение (2.15) можно представить в виде:

$$\frac{1}{\mu} [(\mathbf{e} grad L) grad L - \mathbf{e} (grad L)^2] + \varepsilon \mathbf{e} = 0. \quad (2.18)$$

В соответствии с уравнением (2.16) первый член полученного уравнения обращается в нуль. При этом, поскольку вектор \mathbf{e} не равен нулю во всем рассматриваемом пространстве, получаем:

$$(grad L)^2 = n^2, \quad (2.19)$$

где $n = \sqrt{\varepsilon \mu}$ – показатель преломления среды.

Функцию $L(\mathbf{r})$ называют эйконалом, а уравнение (2.19) – уравнением эйконала.

Важно обратить внимание на то, что уравнение (2.19) получено в предположении, что $\lambda_0 \rightarrow 0$. Раздел оптики, в котором пренебрегают конечностью длин волн, что соответствует предельному переходу при $\lambda_0 \rightarrow 0$, называется геометрической оптикой, поскольку в этом

приближении оптические законы можно сформулировать на языке геометрии. Уравнение (2.19) называют основным уравнением геометрической оптики. Поверхности $L(\mathbf{r}) = const$ называют геометрическими волновыми поверхностями или геометрическими волновыми фронтами.

Формулу (2.19) можно представить в виде:

$$\left(\frac{\partial L}{\partial x}\right)^2 + \left(\frac{\partial L}{\partial y}\right)^2 + \left(\frac{\partial L}{\partial z}\right)^2 = n^2. \quad (2.20)$$

Формулы (2.19) и (2.20) справедливы лишь тогда, когда изменения $a(\mathbf{r})$ на расстояниях, сравнимых с длиной волны, малы по сравнению с самой величиной $a(\mathbf{r})$. Поэтому нельзя ожидать, что геометрическая оптика даст правильное описание поля, например, на границе тени, так как там интенсивность света резко меняется, или вблизи точек, где происходит резкая концентрация света (например, вблизи резкого изображения точки). Ответ на вопрос о распространении интенсивности поля в таких областях дает теория дифракции.

2.2. Световые лучи

Энергия, протекающая за время t через основание цилиндра, ось которого параллельна орту \mathbf{s} , а площадь поперечного сечения равна единице, равна энергии, содержащейся в цилиндре длиной Vt , где V – скорость распространения электромагнитного поля в пространстве. Следовательно, поток энергии, равный энергии, протекающей через основание цилиндра в единицу времени, равен Vw , где w – плотность энергии электромагнитного поля. Плотности электрической и магнитной энергии соответственно равны

$$w_e = \frac{1}{8\pi} \mathbf{E} \cdot \mathbf{D}, \quad w_m = \frac{1}{8\pi} \mathbf{H} \cdot \mathbf{B},$$

при этом плотность энергии электромагнитного поля равна $w = w_e + w_m$.

Учитывая соотношение (1.43), получаем

$$w = \frac{\varepsilon}{4\pi} E^2 = \frac{\mu}{4\pi} H^2. \quad (2.21)$$

С другой стороны, вектор Умова-Пойнтинга в соответствии с выражением (1.105) равен

$$\mathbf{G} = \frac{c}{4\pi} \mathbf{E} \times \mathbf{H} = \frac{c}{4\pi} EH\mathbf{s} = \frac{c}{4\pi} \sqrt{\frac{\varepsilon}{\mu}} E^2 \mathbf{s} = \frac{c}{4\pi} \sqrt{\frac{\mu}{\varepsilon}} H^2 \mathbf{s}. \quad (2.22)$$

Сравнивая выражения (2.21) и (2.22), получаем

$$\mathbf{G} = \frac{c}{\sqrt{\varepsilon\mu}} w\mathbf{s} = Vw\mathbf{s}. \quad (2.23)$$

Отсюда следует, что вектор Умова-Пойнтинга определяет поток энергии и по величине, и по направлению распространения.

Векторы электрической и магнитной напряженности гармонического электромагнитного поля можно представить выражениями вида:

$$\begin{aligned} \mathbf{E}(\mathbf{r}, t) &= \operatorname{Re}\{\mathbf{E}_0(\mathbf{r})\exp(-i\omega t)\} = \\ &= \frac{1}{2}[\mathbf{E}_0(\mathbf{r})\exp(-i\omega t) + \mathbf{E}_0^*(\mathbf{r})\exp(i\omega t)], \end{aligned} \quad (2.24)$$

$$\begin{aligned} \mathbf{H}(\mathbf{r}, t) &= \operatorname{Re}\{\mathbf{H}_0(\mathbf{r})\exp(-i\omega t)\} = \\ &= \frac{1}{2}[\mathbf{H}_0(\mathbf{r})\exp(-i\omega t) + \mathbf{H}_0^*(\mathbf{r})\exp(i\omega t)], \end{aligned} \quad (2.25)$$

где \mathbf{E}_0 и \mathbf{H}_0 – комплексные векторные функции координат.

Так как частоты колебаний в оптическом диапазоне излучения очень велики (ω достигает величины порядка 10^{15} c^{-1}), непосредственно наблюдать мгновенные значения ни одной из таких быстро осциллирующих величин невозможно. Можно говорить лишь об их значениях, усредненных за некоторый интервал времени $-\tau \leq t \leq \tau$, который велик по сравнению с периодом колебаний $T = \frac{2\pi}{\omega}$. В частности,

средняя по времени плотность энергии электрического поля равна

$$\begin{aligned} \langle w_e \rangle &= \frac{1}{2\tau} \int_{-\tau}^{\tau} \frac{\varepsilon}{8\pi} \mathbf{E}^2 dt = \frac{\varepsilon}{16\pi\tau} \int_{-\tau}^{\tau} \frac{1}{4} [\mathbf{E}^2 \exp(-2i\omega t) + 2\mathbf{E}_0 \cdot \mathbf{E}_0^* + \\ &+ \mathbf{E}_0^{*2} \exp(2i\omega t)] dt. \end{aligned}$$

Здесь

$$\begin{aligned} \frac{1}{2\tau} \int_{-\tau}^{\tau} \exp(-2i\omega t) dt &= -\frac{1}{4i\omega\tau} \exp(-2i\omega t) \Big|_{-\tau}^{\tau} = \\ &= -\frac{1}{4i\omega\tau} [\exp(-2i\omega\tau) - \exp(2i\omega\tau)] = \frac{1}{4\pi} \frac{T}{\tau} \sin 2\omega\tau. \end{aligned}$$

Поскольку время τ велико по сравнению с периодом T , величина $\frac{T}{\tau}$ мала по сравнению с единицей, а поэтому интегралом, содержащим $\exp(-2i\omega t)$, можно пренебречь. Подобным же образом можно показать,

что $\frac{1}{2\tau} \int_{-\tau}^{\tau} \exp(2i\omega t) dt = \frac{1}{4\pi} \frac{T}{\tau} \sin 2\omega\tau$, а, следовательно, и интегралом,

содержащим $\exp(2i\omega t)$ можно пренебречь. В результате окончательно получаем

$$\langle w_e \rangle \approx \frac{\varepsilon}{16\pi} \mathbf{E}_0 \cdot \mathbf{E}_0^*. \quad (2.26)$$

Аналогично находим, что средняя по времени плотность энергии магнитного поля равна

$$\langle w_m \rangle \approx \frac{h}{16\pi} \mathbf{H}_0 \cdot \mathbf{H}_0^*. \quad (2.27)$$

Среднее по времени значение вектора Умова-Пойнтинга равно

$$\begin{aligned} \langle \mathbf{G} \rangle &= \frac{1}{2\tau} \int_{-\tau}^{\tau} \frac{c}{4\pi} (\mathbf{E} \times \mathbf{H}) dt = \frac{c}{4\pi} \frac{1}{2\tau} \int_{-\tau}^{\tau} \frac{1}{4} [\mathbf{E}_0 \times \mathbf{H}_0 \exp(-2i\omega t) + \\ &+ \mathbf{E}_0 \times \mathbf{H}_0^* + \mathbf{E}_0^* \times \mathbf{H}_0 + \mathbf{E}_0^* \times \mathbf{H}_0^* \exp(2i\omega t)] dt \approx \\ &\approx \frac{c}{16\pi} (\mathbf{E}_0 \times \mathbf{H}_0^* + \mathbf{E}_0^* \times \mathbf{H}_0) = \frac{c}{8\pi} \operatorname{Re} \{ \mathbf{E}_0 \times \mathbf{H}_0^* \}. \end{aligned} \quad (2.28)$$

Подставив выражения (2.7) в выражения (2.26) и (2.27), получаем

$$\langle w_e \rangle = \frac{\varepsilon}{16\pi} \mathbf{e} \cdot \mathbf{e}^*, \quad \langle w_m \rangle = \frac{\mu}{16\pi} \mathbf{h} \cdot \mathbf{h}^*. \quad (2.29)$$

Из выражений (2.14) и (2.15) следует, что

$$\mathbf{h} = \frac{1}{\mu} \operatorname{grad} L \times \mathbf{e} \quad (2.30)$$

$$\mathbf{e}^* = -\frac{1}{\varepsilon} \operatorname{grad} L \times \mathbf{h}^*. \quad (2.31)$$

При этом выражения (2.29) принимают вид

$$\begin{aligned} \langle w_e \rangle &= -\frac{1}{16\pi} \mathbf{e} (\operatorname{grad} L \times \mathbf{h}^*) = \frac{1}{16\pi} (\mathbf{e} \times \mathbf{h}^*) \operatorname{grad} L, \\ \langle w_m \rangle &= \frac{1}{16\pi} (\operatorname{grad} L \times \mathbf{e}) \mathbf{h}^* = \frac{1}{16\pi} (\mathbf{e} \times \mathbf{h}^*) \operatorname{grad} L. \end{aligned}$$

В результате имеем

$$\langle w_e \rangle = \langle w_m \rangle = \frac{1}{16\pi} (\mathbf{e} \times \mathbf{h}^*) \operatorname{grad} L.$$

Следовательно, в приближении геометрической оптики усредненные по времени плотности электрического и магнитного полей равны друг другу.

Подставив выражения (2.7) в выражение (2.28), получаем

$$\langle \mathbf{G} \rangle = \frac{c}{8\pi} \operatorname{Re} \{ \mathbf{e} \times \mathbf{h}^* \}.$$

Используя соотношение (2.30), полученное выражение можно представить в виде:

$$\langle \mathbf{G} \rangle = \frac{c}{8\pi\mu} \operatorname{Re} \{ \mathbf{e} \times (\operatorname{grad} L \times \mathbf{e}^*) \} = \frac{c}{8\pi\mu} \operatorname{Re} \{ (\mathbf{e} \cdot \mathbf{e}^*) \operatorname{grad} L - (\mathbf{e} \operatorname{grad} L) \mathbf{e}^* \}.$$

В соответствии с уравнением (2.16): $\mathbf{e} \text{ grad} L = 0$.

При этом

$$\langle \mathbf{G} \rangle = \frac{c}{8\pi\mu} \text{Re} \left\{ (\mathbf{e} \cdot \mathbf{e}^*) \text{grad} L \right\} = \frac{c}{8\pi\mu} (\mathbf{e} \cdot \mathbf{e}^*) \text{grad} L.$$

Первое из выражений (2.29) позволяет представить полученное выражение в виде

$$\langle \mathbf{G} \rangle = 2 \frac{c}{\varepsilon\mu} \langle w_e \rangle \text{grad} L. \quad (2.32)$$

Средняя по времени плотность полной энергии $\langle w \rangle = \langle w_e \rangle + \langle w_m \rangle = 2\langle w_e \rangle$. При $\varepsilon\mu = n^2$ отношение $\frac{c}{\sqrt{\varepsilon\mu}} = V$, а отношение

$\frac{\text{grad} L}{n}$ в соответствии с уравнением (2.19) эйконала определяет некоторый единичный вектор \bar{s} , равный

$$\mathbf{s} = \frac{\text{grad} L}{n} = \frac{\text{grad} L}{|\text{grad} L|}. \quad (2.33)$$

В результате получаем, что

$$\langle \mathbf{G} \rangle = V \langle w \rangle \mathbf{s}. \quad (2.34)$$

Отсюда следует, что направление усредненного по времени вектора Умова-Пойнтинга совпадает с нормалью к геометрическому волновому фронту, а абсолютная его величина равна произведению средней плотности энергии на скорость $V = \frac{c}{n}$. Выражение (2.34) свидетельствует о том, что в приближении геометрической оптики средняя плотность полной энергии электромагнитного поля распространяется со скоростью $V = \frac{c}{n}$.

Полученные результаты позволяют ввести понятие геометрических световых лучей.

Геометрические световые лучи можно определить как траектории перемещения плотности световой энергии, ортогональные геометрическим волновым фронтам $L = \text{const}$, при этом направление перемещения в каждой точке траектории совпадает с направлением усредненного вектора Умова-Пойнтинга.

Если радиус-вектор $\mathbf{r}(s)$ точки P , расположенной на луче, рассматривать как функцию длины s дуги луча, то в соответствии с рис. 2.1 отношение $\frac{d\mathbf{r}}{ds} = \mathbf{s}$, при этом уравнении луча в соответствии с выражением (2.33) можно представить в виде:

$$n \frac{d\mathbf{r}}{ds} = \text{grad} L(\mathbf{r}). \quad (2.35)$$

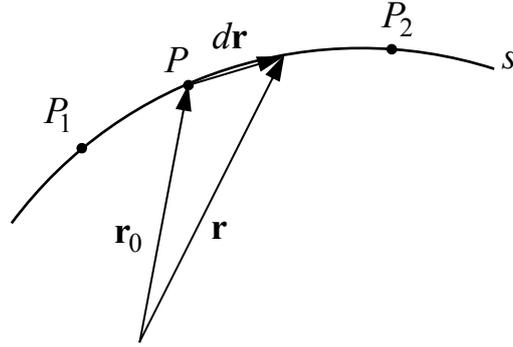


Рис. 2.1. Радиус-вектор произвольной точки траектории луча

Умножив уравнение (2.35) скалярно на орт \mathbf{s} , получаем

$$n = \frac{d\mathbf{r}}{ds} \text{grad}L(\mathbf{r}).$$

Но

$$d\mathbf{r} \text{grad}L(\mathbf{r}) = \frac{\partial L(x, y, x)}{\partial x} dx + \frac{\partial L(x, y, x)}{\partial y} dy + \frac{\partial L(x, y, x)}{\partial z} dz = dL(\mathbf{r}).$$

При этом

$$n = \frac{dL(\mathbf{r})}{ds}. \quad (2.36)$$

Дифференцируя выражение (2.35) по s , получаем

$$\frac{d}{ds} \left(n \frac{d\mathbf{r}}{ds} \right) = \frac{d}{ds} \text{grad}L = \frac{d}{d\mathbf{r}} \frac{d\mathbf{r}}{ds} \text{grad}L = \frac{d}{d\mathbf{r}} \frac{dL}{ds} = \frac{dn}{d\mathbf{r}} = \frac{d\mathbf{r}}{d\mathbf{r}} \text{grad}n,$$

то есть

$$\frac{d}{ds} \left(n \frac{d\mathbf{r}}{ds} \right) = \text{grad}n. \quad (2.37)$$

В однородной среде $n = \text{const}$, при этом уравнение (2.37) принимает вид: $\frac{d^2\mathbf{r}}{ds^2} = 0$. Очевидное решение этого уравнения можно представить в

виде:

$$\mathbf{r} = s\mathbf{a} + \mathbf{b},$$

где \mathbf{a} и \mathbf{b} – постоянные векторы. Это решение представляет собой уравнение прямой линии, направленной вдоль вектора \mathbf{a} и проходящей через точку $\mathbf{r} = \mathbf{b}$. Следовательно, в однородной среде световые лучи есть прямые линии.

Рассмотрим два волновых фронта: $L_1 = L = \text{const}$ и $L_2 = L + dL = \text{const}$, как показано на рис. 2.2. Из выражения (2.36) следует, что приращение эйконала $dL(\mathbf{r})$ определяется произведением расстояния между волновыми фронтами на показатель преломления среды. Интеграл $\int n ds$ вдоль кривой,

определяющей траекторию луча, называется оптической длиной этой кривой.

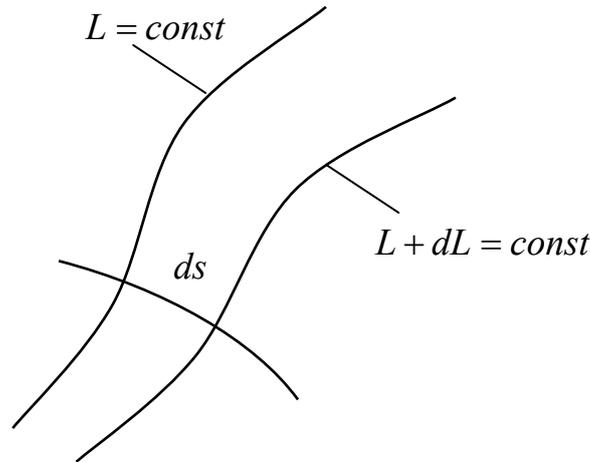


Рис. 2.2. Смещение волнового фронта на расстояние ds

Представим себе, что два фиксированных положения волнового фронта связаны между собой множеством световых лучей, проходящих на своем пути оптические среды, разделяющие рассматриваемые волновые фронты. Обозначая квадратными скобками оптическую длину произвольного луча, проходящего через соответствующие точки P_1 и P_2 волнового фронта в первом и во втором положениях (смотри, например, рис. 2.1), получаем:

$$[P_1P_2] = \int_{P_1}^{P_2} n ds = L(P_2) - L(P_1). \quad (2.38)$$

Так как значение эйконала L при каждом положении волнового фронта постоянно, то постоянна и разность эйконалов при двух фиксированных положениях волнового фронта. При этом из выражения (2.38) следует, что

$$\int_{P_1}^{P_2} n ds = \text{const}. \quad (2.39)$$

С другой стороны, $n ds = \frac{c}{V} ds = c dt$, где dt – время прохождения световой волной расстояния ds вдоль луча (время распространения поля или переноса световой энергии на расстояние ds вдоль луча). Следовательно, время распространения световой волны от точки P_1 до точки P_2 равно

$$T = \int_{T_1}^{T_2} dt = \frac{1}{c} \int_{P_1}^{P_2} n ds = \text{const}.$$

Основываясь на этой физической интерпретации выражения (2.39), его называют законом таухронизма (законом одновременности), однако, формулируют так: оптическая длина световых лучей между двумя фиксированными положениями волнового фронта постоянна.

2.3. Закон преломления

Учитывая тождество $\text{rot grad} = 0$, находим, что в соответствии с уравнением (2.35) вектор $n\mathbf{s} = n\frac{d\mathbf{r}}{ds}$, называемый лучевым вектором, удовлетворяет соотношению:

$$\text{rot } n\mathbf{s} = 0. \quad (2.40)$$

Рассмотрим прохождение луча через поверхность раздела двух сред. Заменим поверхность раздела сред S тонким переходным слоем, внутри которого величины ε и μ быстро, но непрерывно меняются от значений, характеризующих среду с одной стороны поверхности, до значений – с другой ее стороны. Построим элементарную прямоугольную площадку, стороны которой P_1Q_1 и P_2Q_2 параллельны, а стороны P_1P_2 и Q_1Q_2 перпендикулярны к поверхности S , как показано на рис. 2.3.

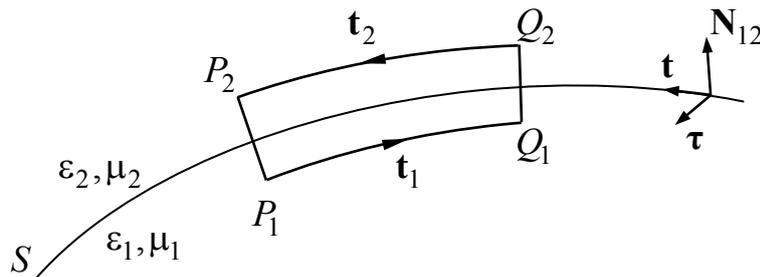


Рис. 2.3. К выводу закона преломления

Пусть \mathbf{t} – единичный вектор касательной к поверхности, а \mathbf{N}_{12} – единичный вектор нормали, направленный из первой среды во вторую. Если обозначить через $\boldsymbol{\tau}$ единичный вектор нормали к элементарной площадке, то, интегрируя выражение (2.40) по площади элемента и используя теорему Стокса, получаем:

$$\int_{\Sigma} (\text{rot } n\mathbf{s}) \boldsymbol{\tau} d\Sigma = \oint_C n\mathbf{s} \cdot d\mathbf{r}, \quad (2.41)$$

где второй интеграл берется по ограничивающему элементарную площадку $d\Sigma$ контуру $P_1Q_1Q_2P_2P_1$. Если длина отрезков $P_1Q_1 = \delta l_1$, $P_2Q_2 = \delta l_2$ и $P_1P_2 = Q_1Q_2 = \delta h$ мала, то на каждой из этих сторон вектор $n\mathbf{s}$ можно заменить постоянными векторами $n_1\mathbf{s}_1$ и $n_2\mathbf{s}_2$. Тогда этот интеграл можно представить в виде суммы:

$$n_1 \mathbf{s}_1 \cdot \mathbf{t}_1 \delta l_1 + n_2 \mathbf{s}_2 \cdot \mathbf{t}_2 \delta l_2 + \frac{1}{2} (n_1 \mathbf{s}_1 + n_2 \mathbf{s}_2) (\mathbf{N}_{12} + \mathbf{N}_{21}) \delta h = 0. \quad (2.42)$$

При $\delta h \rightarrow 0$ справедливо равенство $\delta l_1 = \delta l_2 = \delta l$. При этом выражение (2.42) принимает вид:

$$(n_1 \mathbf{s}_1 \cdot \mathbf{t}_1 + n_2 \mathbf{s}_2 \cdot \mathbf{t}_2) \delta l = 0. \quad (2.43)$$

В соответствии с принятым на рис. 2.3 направлением векторов имеем: $\mathbf{t}_1 = -\mathbf{t} = -\boldsymbol{\tau} \times \mathbf{N}_{12}$; $\mathbf{t}_2 = \mathbf{t} = \boldsymbol{\tau} \times \mathbf{N}_{12}$.

Подставив эти соотношения в выражение (2.43), получаем:

$$(n_2 \mathbf{s}_2 - n_1 \mathbf{s}_1) (\boldsymbol{\tau} \times \mathbf{N}_{12}) = \boldsymbol{\tau} [\mathbf{N}_{12} \times (n_2 \mathbf{s}_2 - n_1 \mathbf{s}_1)] = 0,$$

или

$$\mathbf{N}_{12} \times (n_2 \mathbf{s}_2 - n_1 \mathbf{s}_1) = 0. \quad (2.44)$$

Отсюда следует, что тангенциальная составляющая лучевого вектора $n\mathbf{s}$ непрерывна при переходе через поверхность раздела двух сред или, что то же самое, вектор $\tilde{\mathbf{N}}_{12} = n_2 \mathbf{s}_2 - n_1 \mathbf{s}_1$ перпендикулярен к этой поверхности в точке падения луча, при этом $\tilde{\mathbf{N}}_{12} = \mu \mathbf{N}_{12}$, где μ – скалярный множитель, называемый астигматической постоянной или постоянной отклонения.

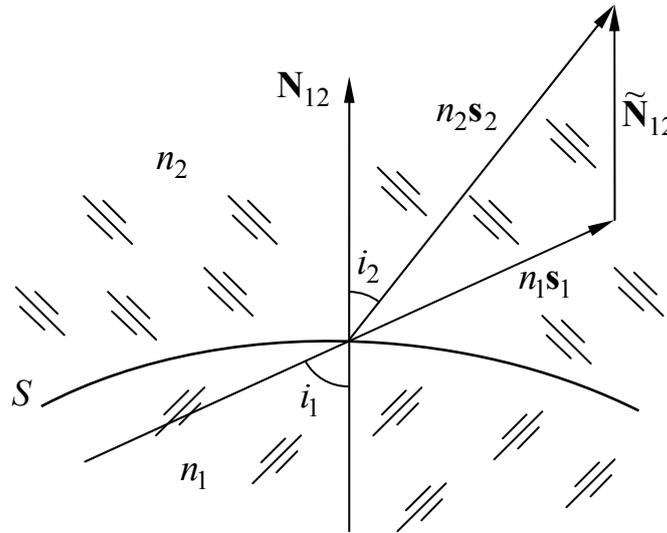


Рис. 2.4. К выводу закона преломления

Пусть i_1 и i_2 – углы, образованные падающим и преломленным лучами с нормалью к поверхности раздела двух сред в точке падения, т.е. с нормалью \mathbf{N}_{12} , как показано на рис. 2.4. Тогда в соответствии с выражением (2.44) имеем:

$$n_2 (\mathbf{N}_{12} \times \mathbf{s}_2) = n_1 (\mathbf{N}_{12} \times \mathbf{s}_1), \quad (2.45)$$

при этом:

$$n_2 \sin i_2 = n_1 \sin i_1. \quad (2.46)$$

Итак, в соответствии с выражением (2.45) луч падающий, нормаль к поверхности раздела сред в точке падения луча и луч преломленный лежат

в одной плоскости. При этом в соответствии с выражением (2.46) отношение синуса угла падения луча к синусу угла преломления равно отношению показателей преломления сред $\frac{n_2}{n_1}$. Эти два положения, выражаемые формулами (2.45) и (2.46), определяют закон преломления Декарта-Снелиуса.

Обратимся вновь к выражению (2.44) и покажем применение его для построения хода преломленного луча. Умножив это выражение на некоторый масштабный множитель M , получаем

$$(Mn's' - Mns) \times \mathbf{N} = 0. \quad (2.47)$$

Пусть на преломляющую поверхность произвольной формы в точку Q падает луч, направление которого определяется ортом \mathbf{s} , как показано на рис. 2.5. Отложим от точки Q в направлении орта \mathbf{s} отрезок QA_1 , равный nM единиц линейной меры (например, $100n$ мм). Из точки A_1 в направлении орта \mathbf{N} проведем линию. На этой линии определим положение точки A_2 , удаленной от точки Q на расстояние $QA_2 = n'M$ тех же единиц линейной меры, например, с помощью циркуля. Через точки Q и A_2 проведем линию, направление которой определим ортом \mathbf{s}' . Поскольку построения на рис. 2.5 выполнены в соответствии с формулой (2.47), орт \mathbf{s}' определяет направление хода преломленного луча.

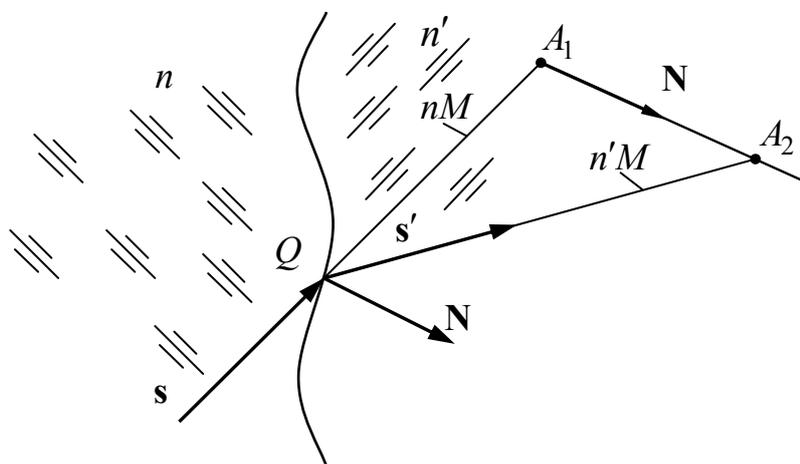


Рис. 2.5. Преломление луча на поверхности раздела двух сред

С другой стороны, из выражения (2.44) следует, что

$$n's' = ns + \mu \mathbf{N}. \quad (2.48)$$

Эта формула является основной в методе расчета хода луча через систему сферических, а также несферических поверхностей, предложенном Иваном Васильевичем Лебедевым в 1938 году (журнал «Оптико-механическая промышленность», 1938, №7). Умножив скалярно обе части выражения (2.48) на вектор \mathbf{N} , получаем

$$\mu = n' \cos i' - n \cos i. \quad (2.49)$$

Используя закон преломления в форме (2.46), выражение (2.49) легко преобразовать к виду:

$$\mu = \sqrt{n'^2 - n^2 + (n\mathbf{N} \cdot \mathbf{s})^2} - n\mathbf{N} \cdot \mathbf{s}. \quad (2.50)$$

Для луча, отраженного поверхностью, имеем: $n' = n$, а $i' = \pi - i$. При этом

$$\mu = -2n \cos i = -2n\mathbf{N} \cdot \mathbf{s}.$$

Подставив это соотношение в выражение (2.48), при $n' = n$ получаем

$$\mathbf{s}' = \mathbf{s} - 2\mathbf{N}(\mathbf{N} \cdot \mathbf{s}). \quad (2.51)$$

Формула (2.51) является одной из основных, применяемых при решении задач юстировки зеркальных и зеркально-призмных систем.

Выразим орты \mathbf{s} , \mathbf{N} и \mathbf{s}' через их проекции на оси декартовых координат:

$$\mathbf{s} = s_x \mathbf{i} + s_y \mathbf{j} + s_z \mathbf{k},$$

$$\mathbf{N} = N_x \mathbf{i} + N_y \mathbf{j} + N_z \mathbf{k},$$

$$\mathbf{s}' = s'_x \mathbf{i} + s'_y \mathbf{j} + s'_z \mathbf{k}.$$

Подставив эти выражения в формулу (2.51), преобразуем ее к виду:

$$\begin{pmatrix} s'_x \\ s'_y \\ s'_z \end{pmatrix} = m' \begin{pmatrix} s_x \\ s_y \\ s_z \end{pmatrix}, \quad (2.52)$$

где матрица m' имеет вид:

$$m' = \begin{pmatrix} 1 - 2N_x^2 & -2N_x N_y & -2N_x N_z \\ -2N_x N_y & 1 - 2N_y^2 & -2N_y N_z \\ -2N_x N_z & -2N_y N_z & 1 - 2N_z^2 \end{pmatrix}. \quad (2.53)$$

2.4. Принцип Ферма

Принцип Ферма, известный теперь как принцип наименьшего оптического пути, утверждает, что оптическая длина реального луча между любыми двумя оптически несопряженными точками P_1 и P_2 ,

равная $\int_{P_1}^{P_2} n ds$, короче оптической длины любой другой кривой,

соединяющей эти точки и лежащей в некоторой регулярной окрестности луча. Под регулярной окрестностью понимается область, которую можно заменить лучами таким образом, что через каждую ее точку будет проходить один (и только один) луч. Покажем справедливость принципа Ферма в случае прохождения света через одну преломляющую поверхность произвольной формы, разделяющую две однородные среды с показателями преломления n и n' .

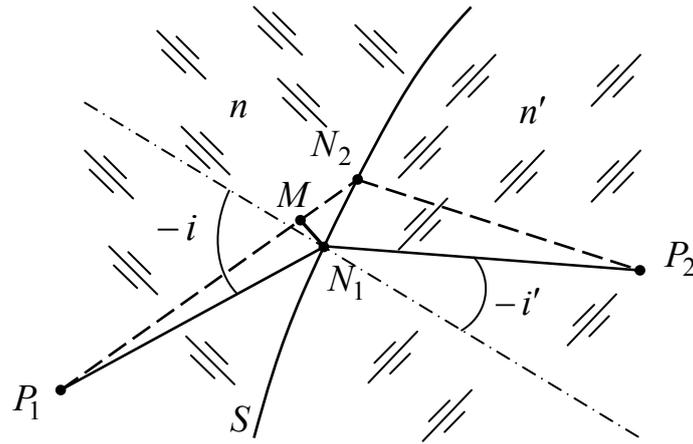


Рис.2.6. Принцип Ферма и закон преломления

Обратимся к рис. 2.6, на котором показан путь луча, проходящего через оптически несопряженные точки P_1 и P_2 и падающего в точку N_1 преломляющей поверхности S , который образует с нормалью к поверхности в точке падения луча углы i и i' . Выберем на преломляющей поверхности точку N_2 , бесконечно близкую к N_1 . При этом отрезок кривой N_1N_2 будем считать величиной первого порядка малости, что позволяет заменить его дифференциалом: $N_1N_2 = dS$. Соединив точку N_2 с точками P_1 и P_2 , получим путь $P_1N_2P_2$, соседний с путем $P_1N_1P_2$ луча. Оптическая длина хода луча $P_1N_1P_2$ может быть представлена в виде: $[P_1P_2] = ns + n's'$, где $s = P_1N_1$, а $s' = N_1P_2$. При переходе от траектории луча $P_1N_1P_2$ к соседней траектории $P_1N_2P_2$ оптическая длина хода луча получит приращение, равное:

$$d[P_1P_2] = nds + n'ds'. \quad (2.54)$$

Чтобы определить величину этого приращения, восстановим в точке N_1 перпендикуляр N_1M к лучу P_1N_2 . Из-за малости угла $N_1P_1N_2$ можно принять $P_1M \approx P_1N_1$. Поэтому отрезок MN_2 можно рассматривать как приращение ds отрезка $P_1N_1 = s$ при замене пути луча P_1N_1 путем P_1N_2 . Пренебрегая величинами выше первого порядка малости, из треугольника MN_1N_2 находим, что $ds = -dS \sin i$. При этом приращение отрезка $N_1P_2 = s'$ равно $ds' = dS \sin i'$. Полученные выражения позволяют равенство (2.54) представить в виде:

$$d[P_1P_2] = (n' \sin i' - n \sin i) dS. \quad (2.55)$$

Отсюда получаем

$$\frac{d[P_1P_2]}{dS} = n' \sin i' - n \sin i.$$

При соблюдении закона преломления имеем:

$$\frac{d[P_1P_2]}{dS} = 0, \quad (2.56)$$

т.е. при соблюдении закона преломления длина оптического пути $[P_1P_2]$ имеет экстремальное значение, что свидетельствует о справедливости принципа Ферма. С другой стороны, можно считать, что при соблюдении принципа Ферма равенство $n' \sin i' = n \sin i$ определяет закон преломления.

Если каждому числу x из чисел некоторого класса сопоставлено другое число y , то, как известно, мы имеем дело с функцией $y = y(x)$. Если каждой функции $y(x)$ из некоторого класса функций сопоставлено некоторое число F , то говорят, что задан функционал $F[y(x)]$. Вариацию функционала обозначают символом $\delta F[y(x)]$. Вариация функционала является аналогом дифференциала функции. Для достижения функционалом $F[y(x)]$ экстремального значения необходимо, чтобы вариация (если она существует) при некоторой функции $y = \tilde{y}(x)$ обращалась в ноль. Учитывая изложенное, принцип Ферма в обозначениях аппарата вариационного исчисления можно записать в виде:

$$\delta \int_{P_1}^{P_2} n ds = 0. \quad (2.57)$$

Поскольку $n = \frac{c}{V}$, где $c = const$, выражение (2.57) можно представить эквивалентным выражением вида:

$$\delta \int_{P_1}^{P_2} \frac{ds}{V} = 0, \quad (2.58)$$

т.е. $\delta \int_{T_1}^{T_2} dt = 0.$

С другой стороны, $k = nk_0$, т.е. $\lambda_0 = n\lambda$, что позволяет выражение (2.57) представить в виде:

$$\delta \int_{P_1}^{P_2} \frac{ds}{\lambda} = 0. \quad (2.59)$$

Принцип Ферма содержит в себе и обобщает следующие положения лучевой оптики:

1. Закон преломления распространения света в однородной среде, т.е. в среде с постоянным показателем преломления.
2. Законы преломления и отражения лучей на поверхностях раздела сред, т.е. на тех поверхностях, где показатель преломления меняется скачком.
3. Позволяет рассчитать путь света в среде с показателем преломления, непрерывно изменяющимся вдоль пути, для которого справедливо уравнение (2.57).

4. Устанавливает закон обратимости светового пути, в соответствии с которым линия, представляющая собой возможный путь течения световой энергии, распространяющейся в одном направлении, есть также возможный путь ее течения в обратном направлении.

Действительно, если вариация интеграла (2.57) равна нулю, когда он берется в пределах от P_1 до P_2 , то она также равна нулю при перемене пределов интегрирования местами (т.е. в пределах от P_2 до P_1).

Следует помнить, что все приведенные выше положения применимы только в пределах тех условий, в которых справедливы законы геометрической (лучевой) оптики.

2.5. Интегральный инвариант Лагранжа-Пуанкаре

Предположим, что показатель преломления n является непрерывной функцией координат. Тогда, применив теорему Стокса к интегралу от нормального компонента $rot\ n\mathbf{s}$, взятому по любой открытой поверхности, получаем:

$$\oint_C n\mathbf{s} \cdot d\mathbf{r} = 0. \quad (2.60)$$

Интегрирование выполняется по замкнутому контуру C , ограничивающему указанную поверхность. Полученное выражение называется интегральным инвариантом Лагранжа-Пуанкаре и означает,

что интеграл $\int_{P_1}^{P_2} n\mathbf{s} \cdot d\mathbf{r}$, взятый между любыми двумя точками поля P_1 и P_2 , не зависит от пути интегрирования.

Можно показать, что формула (2.60) остается справедливой и в том случае, когда контур C пересекает поверхность, разделяющую две однородные среды с разными показателями преломления. Для доказательства этого будем считать, что контур C делится на части C_1 и C_2 , расположенными по разные стороны от преломляющей поверхности S , а точки пересечения контура C с поверхностью S соединены другой кривой K , лежащей на этой поверхности, как показано на рис. 2.7.

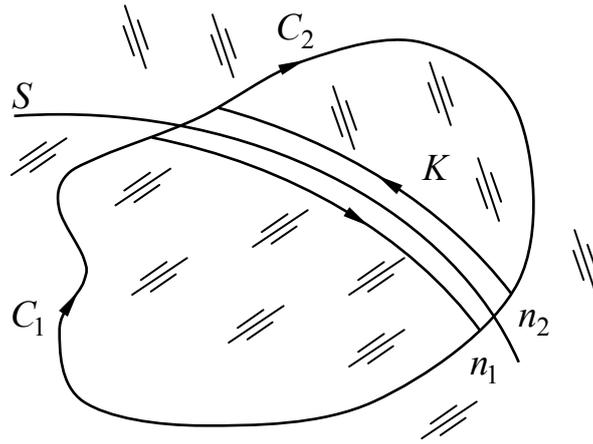


Рис. 2.7. К выводу инварианта Лагранжа-Пуанкаре при наличии поверхности раздела двух однородных сред

Применяя формулу (2.60) к обоим контурам C_1K и C_2K и складывая полученные выражения, имеем:

$$\int_{C_1} n_1 \mathbf{s}_1 \cdot d\mathbf{r} + \int_{C_2} n_2 \mathbf{s}_2 \cdot d\mathbf{r} + \int_K (n_2 \mathbf{s}_2 - n_1 \mathbf{s}_1) d\mathbf{r} = 0. \quad (2.61)$$

Интеграл вдоль кривой K равен нулю, поскольку, как было показано при выводе закона преломления, вектор $\tilde{\mathbf{N}}_{12} = n_2 \mathbf{s}_2 - n_1 \mathbf{s}_1$, перпендикулярен к поверхности S в любой точке кривой K . При этом выражение (2.61) сводится к выражению (2.60), что и требовалось доказать.

2.6. Конгруэнции лучей

Уравнение произвольной прямой в пространстве можно определить уравнением вида:

$$x = a_{11} + a_{12}z, \quad y = a_{21} + a_{22}z. \quad (2.62)$$

Задавая независимым друг от друга величинам коэффициентов a_{11}, \dots, a_{22} всевозможные значения, получим уравнения всех прямых, заполняющих пространство. Связывая величины коэффициентов одной зависимости $F_1(a_{11}, a_{12}, a_{21}, a_{22}) = 0$, получаем комплекс прямых. Вполне очевидно, что в общем случае через каждую точку пространства проходит конус лучей с вершиной в этой точке. Добавляя еще одну зависимость $F_2(a_{11}, a_{12}, a_{21}, a_{22}) = 0$, мы выделяем из комплекса конгруэнцию (от лат. *congruens* (*congruentis*) – соразмерный, соответствующий, совпадающий) прямых.

Пусть через каждую точку некоторой поверхности $z = f(x, y)$ проходит прямая, определяемая коэффициентами $a_{11}, a_{12}, a_{21}, a_{22}$, которые являются функциями от x и y :

$$\begin{aligned}
a_{11} &= a_{11}(x, y), \\
a_{12} &= a_{12}(x, y), \\
a_{21} &= a_{21}(x, y), \\
a_{22} &= a_{22}(x, y).
\end{aligned}$$

В результате получаем конгруэнцию прямых, так как из написанных четырех равенств можно исключить x и y и получить два уравнения, связывающих величины a_{11}, \dots, a_{22} . Конгруэнцией прямых, например, является совокупность исходящих из светящейся точки лучей после выхода из оптической системы при условии, что последняя среда однородна.

Соотношение (2.40), а, именно, $rots = 0$, определяет все системы лучей, которые могут существовать в неоднородной среде, и выделяет их из более общих семейств кривых. В однородной среде, т.е. при $n = const$, это соотношение принимает вид:

$$rots = 0. \quad (2.63)$$

Левую часть соотношения (2.40) можно заменить тождеством

$$rots = nrots + (gradn) \times s. \quad (2.64)$$

Умножив при этом соотношение (2.64) скалярно на вектор s , получаем, что система лучей в любой среде должна удовлетворять соотношению:

$$s \cdot rots = 0. \quad (2.65)$$

Система кривых, заполняющих некоторую часть пространства так, что через каждую точку этой части пространства в общем случае проходит одна кривая, называется конгруэнцией. Если существует семейство поверхностей, пересекающих каждую кривую под прямым углом, то рассматриваемую систему кривых называют нормальной конгруэнцией. Если такого семейства поверхностей нет, то говорят о косо́й конгруэнции. В обычной (световой) геометрической оптике рассматривают только нормальные конгруэнции лучей. Косые конгруэнции играют важную роль в электронной оптике.

Если все лучи, составляющие конгруэнцию, имеют вид прямых, то такая конгруэнция называется прямолинейной конгруэнцией лучей. Формулы (2.63) и (2.65) определяют необходимые и достаточные условия, соблюдение которых свидетельствует о том, что конгруэнция лучей является соответственно нормальной и нормальной прямоугольной.

2.7. Теорема Малюса

Если все лучи в однородной среде имеют общую точку, например, исходят из точечного источника, то говорят, что лучи образуют гомоцентрический пучок. Такой пучок образует нормальную конгруэнцию, поскольку каждый луч пучка пересекает под прямым углом сферические

поверхности (волновые фронты), центр которых расположен в точке пересечения лучей.

В 1808 году французский физик Этьен Луи Малюс (1775–1812) показал, что если гомоцентрический пучок прямолинейных лучей преломляется или отражается какой-либо поверхностью, то получающийся после этого пучок лучей (в общем случае уже не гомоцентрический) тоже образует нормальную конгруэнцию. Позднее Дюпин (1816 г.), Кветеле (1825 г.) и Жергонн (1825 г.) обобщили результат Малюса. Работы этих ученых позволили сформулировать следующую теорему, называемую теоремой Малюса: нормальная прямолинейная конгруэнция остается нормальной после любого числа преломлений и отражений.

Рассмотрим нормальную прямолинейную конгруэнцию лучей в однородной среде с показателем преломления n_1 , которые преломляются на поверхность S , отделяющей эту среду от другой однородной среды с показателем преломления n_2 , как показано на рис. 2.8.

Пусть L_1 – один из волновых фронтов в первой среде, A_1 и P – точки пересечения произвольного луча в первой среде с поверхностями L_1 и S соответственно, а A_2 – точка на преломленном луче. Если точку A_1 сместить в другую точку B_1 на том же волновом фронте, то точка P на поверхности преломления сместится в точку Q . Теперь на преломленном в точке Q луче выберем такую точку B_2 , чтобы оптический путь от B_1 до B_2 равнялся оптическому пути от A_1 до A_2 , т.е. чтобы соблюдалось равенство:

$$[A_1PA_2] = [B_1QB_2]. \quad (2.66)$$

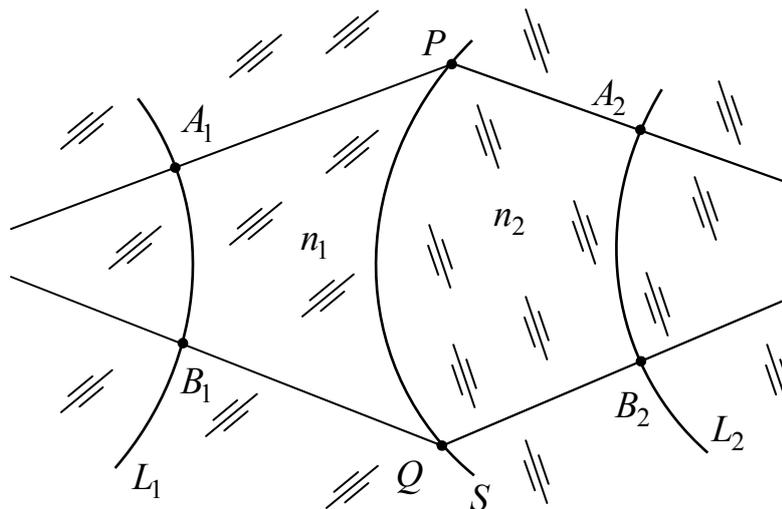


Рис. 2.8. К доказательству теоремы Малюса

Если перемещать точку B_1 по всей поверхности L_1 , то точка B_2 при своем перемещении заполнит поверхность L_2 . Покажем, что преломленный луч QB_2 перпендикулярен к этой поверхности.

Вычисляя интегральный инвариант Лагранжа-Пуанкаре по замкнутому пути $A_1PA_2B_2QB_1A_1$, получаем:

$$\int_{A_1PA_2} nds + \int_{A_2B_2} ns \cdot d\mathbf{r} + \int_{B_2QB_1} nds + \int_{B_1A_1} ns \cdot d\mathbf{r} = 0 \quad (2.67)$$

на основании равенства (2.66) имеем

$$\int_{A_1PA_2} nds + \int_{B_2QB_1} nds = 0.$$

Кроме того, поскольку на волновом фронте L_1 единичный вектор \mathbf{s} в каждой точке перпендикулярен к нему, интеграл $\int_{B_1A_1} ns \cdot d\mathbf{r} = 0$. В результате

выражение (2.67) принимает вид:

$$\int_{A_2B_2} ns \cdot d\mathbf{r} = 0. \quad (2.68)$$

Полученное соотношение должно выполняться на поверхности L_2 для любого отрезка кривой. Это возможно только в том случае, если $\mathbf{s} \cdot d\mathbf{r} = 0$ для каждого линейного элемента $d\mathbf{r}$ поверхности L_2 , т.е. если преломленные лучи перпендикулярны к ней или, другими словами, если преломленные лучи образуют нормальную конгруэнцию, что и требовалось доказать.

Поскольку точка B_1 на волновом фронте L_1 выбрана произвольно, то равенство $[A_1PA_2] = [B_1QB_2]$ позволяет утверждать, что оптическая длина пути между любыми двумя волновыми фронтами одинакова для всех лучей. Вполне очевидно, что этот результат остается справедливым для случая последовательных преломлений (и отражений) на нескольких поверхностях, а так же, как это непосредственно следует из выражения (2.38), в случае распространения лучей в среде с непрерывно изменяющимся показателем преломления. Эта теорема, доказанная с помощью интегрального инварианта Лагранжа-Пуанкаре, называется принципом равного оптического пути. Из этой теоремы следует, что геометрические волновые фронты нормальной конгруэнции лучей или совокупности нормальных конгруэнций, образованных в результате последовательных преломлений (и отражений), «оптически параллельны» друг другу.

С последней теоремой невольно ассоциируется гипотетическая теорема, впервые выдвинутая в 1690 году Х. Гюйгенсом, согласно которой каждую точку волнового фронта можно считать центром вторичного светового возмущения, порождающего элементарные сферические волны,

при этом волновым фронтом в любой последующий момент времени служит огибающая этих вторичных сферических волн. Принцип Гюйгенса может служить правилом для построения поверхностей, «оптически параллельных» друг другу.

Дополнив построение Гюйгенса утверждением, что вторичные волны интерферируют между собой, О. Френель в 1818 году впервые смог объяснить явление дифракции. Это сочетание построения Гюйгенса с принципом интерференции называется принципом Гюйгенса-Френеля. Позднее (в 1882 г.) Г. Кирхгоф придал исследованиям Френеля строго математическое обоснование, и с этого времени началось широкое изучение дифракции. Тем не менее, в большинстве практически важных случаев из-за математических трудностей приходится прибегать к приближенным методам и тут принцип Гюйгенса-Френеля служит мощным инструментом, позволяющим решать большинство вопросов, встречающихся в прикладной оптике.

3. Интерференция света. Интерферометры

Интерферометрами называют измерительные приборы, основанные на использовании интерференции света. Рассмотрим кратко основные понятия интерференции света.

3.1. Элементы теории интерференции света

В соответствии с волновой теорией световое возмещение u в некоторой точке пространства на расстоянии r от источника света описывается уравнением

$$u = a \sin 2\pi \left(\frac{t}{T} - \frac{r}{\lambda} \right) = a \sin \frac{2\pi}{T} \left(t - \frac{r}{V} \right), \quad (3.1)$$

где a – амплитуда; T – период колебания; t – время; λ – длина волны света, а V – скорость его распространения.

Величину $\nu = \frac{1}{T}$ называют частотой колебаний, а величину

$\omega = \frac{2\pi}{T} = 2\pi\nu$ называют круговой, или циклической, частотой.

Произведение расстояния r , проходимого лучом в данной среде, на показатель преломления n этой среды называют оптической длиной пути l , т.е. $l = rn$. Если луч проходит k сред с показателями преломления n_1, n_2, \dots, n_k , то суммарная оптическая длина пути

$$L = \sum_{i=1}^{i=k} r_i n_i. \quad (3.2)$$

При этом уравнение (2.1) можно написать в виде

$$u = a \sin 2\pi \left(\frac{t}{T} - \frac{L}{\lambda} \right).$$

Под разностью хода двух лучей понимают разность оптических длин L_1 и L_2 , проходимых этими лучами, т.е.

$$\Delta = L_2 - L_1. \quad (3.3)$$

Отношение разности хода Δ к длине волны λ называют порядком интерференции m :

$$m = \frac{\Delta}{\lambda}. \quad (3.4)$$

Число m может быть как целым, так и дробным. При этом разность фаз двух колебаний равна

$$\delta = \varphi_2 - \varphi_1 = 2\pi \frac{\Delta}{\lambda} = 2\pi m. \quad (3.5)$$

При сложении двух одинаково направленных колебаний равного периода $u_1 = a_1 \sin \omega t$ и $u_2 = a_2 \sin(\omega t + \delta)$, имеющих амплитуды a_1 и a_2 и

разность фаз δ , возникает колебание того же периода $u^2 = (u_1 + u_2)^2$, при этом

$$a^2 = a_1^2 + a_2^2 + 2a_1a_2 \cos \delta \quad (3.6)$$

$$\operatorname{tg} \psi = \frac{a_2 \sin \delta}{a_1 + a_2 \cos \delta}. \quad (3.7)$$

Интенсивность света (плотность потока энергии электромагнитного поля) в соответствии с выражением (1.86) равна

$$G = \frac{c}{4\pi} EH = \frac{c}{4\pi} \sqrt{\frac{\varepsilon}{\mu}} E^2 = \frac{cn}{4\pi\mu} E^2 = \frac{c}{4\pi} \sqrt{\frac{\mu}{\varepsilon}} H^2 = \frac{cn}{4\pi\varepsilon} H^2.$$

Под интенсивностью колебания I понимают величину, равную квадрату амплитуды: $I = a^2$. В соответствии с выражением (3.6) интенсивность результирующего колебания I связана с интенсивностями I_1 и I_2 формулой

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos \delta. \quad (3.8)$$

Эта формула является основной для теории двухлучевых интерферометров. Из этой формулы следует, что величина I принимает максимальное значение, равное I_{\max} при $\delta = 0; \pm 2\pi; \pm 4\pi; \dots$ или при $\Delta = 0; \pm \lambda; \pm 2\lambda; \dots$, а минимальное значение I_{\min} – при $\delta = \pm \pi; \pm 3\pi; \dots$ или при $\Delta = \frac{\lambda}{2}; \frac{3}{2}\lambda; \dots$. При этом

$$\left. \begin{aligned} I_{\max} &= I_1 + I_2 + 2\sqrt{I_1 I_2} = (a_1 + a_2)^2 \\ I_{\min} &= I_1 + I_2 - 2\sqrt{I_1 I_2} = (a_1 - a_2)^2. \end{aligned} \right\} \quad (3.9)$$

При равенстве интенсивностей складываемых колебаний, т.е. при $I_1 = I_2 = I_0$, имеем

$$I = 2I_0(1 + \cos \delta) = 4I_0 \cos^2 \frac{\delta}{2} = 4I_0 \cos^2 \pi \frac{\Delta}{\lambda}. \quad (3.10)$$

Из выражения (3.10) следует, что при изменении δ интенсивность I изменяется по синусоидальному закону, при этом $I_{\max} = 4I_0$, а $I_{\min} = 0$. Обычно в разных точках поля интерференции, т.е. в плоскости, в которой наблюдают интерференционную картину, разности фаз δ имеют различные значения. Геометрическое место точек поля, лежащих на одной линии, в которых величины δ одинаковы, называют интерференционной полосой, а расстояние b между серединами двух соседних полос – шириной полосы. Переходу от одной полосы к соседней полосе соответствует изменение δ , равное 2π (или изменение Δ , равное λ) независимо от ширины полос. При изменении разности хода Δ интерференционные полосы перемещаются в направлении, обратном направлению увеличения разности хода перпендикулярно к их длине.

Изменение Δ на длину волны λ вызывает перемещение полос на расстояние, равное ширине полосы. Практически нетрудно измерить смещение полос с точностью до 0,05–0,1 ширины полосы, т.е. измерить изменение Δ с погрешностью 0,05 – 0,1 λ .

Контрастность интерференционной картины определяется отношением

$$K = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}, \quad (3.11)$$

где I_{\max} и I_{\min} – максимальное и минимальное значения интенсивности света в светлой и темной полосах соответственно. В идеальном случае, т.е. при $I_1 = I_2$, величина $I_{\min} = 0$, при этом контрастность $K = 1$. В общем случае $I_1 \neq I_2$. Пусть $I_2 = qI_1$. Тогда формула (3.8), определяющая интенсивность результирующего колебания, принимает вид

$$I = (1 + 2\sqrt{q} \cos \delta + q)I_1. \quad (3.12)$$

При этом из выражения (3.11) находим

$$K = \frac{2\sqrt{q}}{1 + q}. \quad (3.13)$$

Заметим, что контрастность интерференционной картины одна и та же при $q_2 = \frac{1}{q_1}$. Легко убедиться, что даже при $q = 5$ контрастность интерференционной картины остается хорошей ($K = 0,75$).

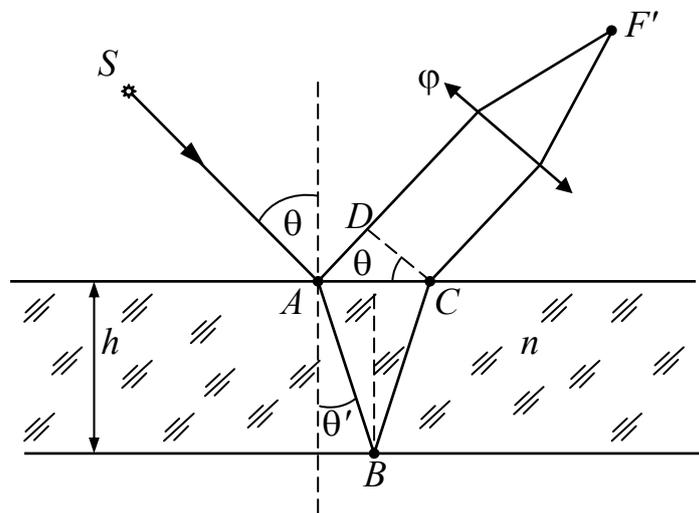


Рис. 3.1. Интерференционные полосы локализованные в бесконечности, возникающие при отражении и преломлении света на поверхностях плоскопараллельной пластинки

Среди множества других особенно важен частный случай интерференции света, отраженного от двух поверхностей плоскопараллельной пластинки, как показано на рис. 3.1. На рисунке луч света от источника S падает на первую (переднюю) поверхность

пластинки в точке A . Частично отражаясь от этой поверхности, луч после преломления в точке A проходит в точку B на второй (задней) поверхности. Частично отражаясь от этой поверхности в точке B , луч приходит в точку C на первой поверхности, преломляясь в которой, выходит из плоскопараллельной пластинки параллельно лучу, отраженному в точке A . Разность оптических путей этих лучей на линии DC , перпендикулярной лучам, как показано на рисунке, равна

$$\Delta = n(|AB| + |BC|) - |AD|.$$

Здесь n – показатель преломления материала пластинки, при этом предполагается, что источник S расположен в воздухе ($n_1 \approx 1$). Обозначив угол падения луча буквой θ , а толщину пластинки – буквой h , получаем

$$\Delta = 2nh \cos \theta', \quad (3.14)$$

при этом $\sin \theta' = \frac{1}{n} \sin \theta$. Поскольку при отражении волны от первой поверхности пластинки в соответствии с формулами Френеля ее фаза изменяется на π , то разность фаз δ рассматриваемых волн равна

$$\delta = 4\pi n \frac{h}{\lambda_0} \cos \theta' \pm \pi = 4\pi \frac{h}{\lambda_0} \sqrt{n^2 - \sin^2 \theta'} \pm \pi, \quad (3.15)$$

где λ_0 – длина волны света в вакууме. Результат сложения двух колебаний можно наблюдать либо глазом, аккомодированным на бесконечность, либо в точке F' на экране, расположенном в фокальной плоскости объектива φ . Из формулы (3.15) следует, что разность фаз δ определяется углом θ падения луча на поверхность пластинки и не зависит от положения источника света S . Отсюда вытекает, что при использовании протяженного источника интерференционные полосы будут столь же отчетливыми, как и с точечным источником: каждый элемент протяженного источника создает в фокальной плоскости объектива φ свою интерференционную картину, причем положение полос во всех этих картинах совершенно одинаково. Так как это справедливо для одной вполне определенной плоскости наблюдения, то про такие полосы говорят, что они локализованы. В данном случае они локализованы в бесконечности (или в фокальной плоскости объектива). В соответствии с формулой (3.14) светлые полосы расположены под углами θ , для которых

$2nh \cos \theta' \pm \frac{1}{2} \lambda_0 = m \lambda_0$, где m – целое число, называемое порядком интерференции. Полоса, соответствующая некоторому порядку интерференции, обусловлена светом, падающим на пластинку под вполне определенным углом θ . Поэтому такие полосы называют интерференционными полосами равного наклона. Если ось объектива расположена перпендикулярно пластинке, полосы имеют вид

концентрических колец с центром в фокусе. В центре картины порядок интерференции максимален.

3.2. Интерферометр Майкельсона

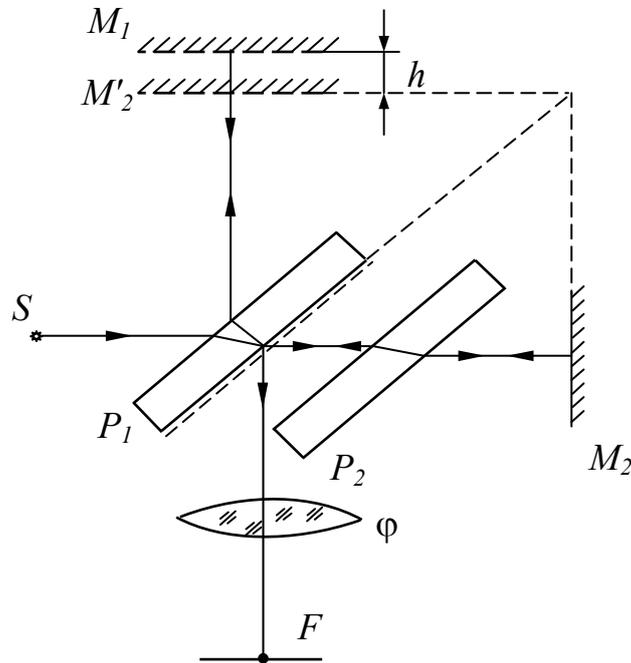


Рис. 3.2. Схема интерферометра Майкельсона

Для наблюдения полос равного наклона вместо плоскопараллельной пластинки удобно использовать интерферометр Майкельсона. Его упрощенная схема показана на рис. 3.2. Свет от протяженного источника S падает на пластинку P_1 , задняя сторона которой покрыта тонким полупрозрачным отражающим слоем серебра или алюминия. Здесь световой пучок лучей разделяется на два взаимно перпендикулярных пучка. Отраженный пластинкой P_1 пучок падает на зеркало M_1 , отражается от него, вновь падает на пластинку P_1 , где снова разделяется на два пучка. Один из них идет к источнику S и интереса не представляет, а другой попадает в зрительную трубу, установленную на бесконечность, или на объектив ϕ , в фокальной плоскости которого расположен экран для наблюдения интерференции. Прошедший сквозь пластинку P_1 световой пучок лучей падает на зеркало M_2 , отражаясь от него, возвращается на пластинку P_1 , на задней поверхности которой вновь делится на два пучка, при этом отраженный пучок направляется в объектив ϕ . Таким образом, от одного источника S получают два пучка примерно одинаковой интенсивности, которые распространяются после разделения пластинкой P_1 в разных «плечах» интерферометра, затем снова встречаются и создают интерференционную картину в фокальной плоскости объектива ϕ . Зеркало M_2 неподвижно, а зеркало M_1 можно передвигать с помощью

микроскопического винта в направлении к пластинке P_1 и от нее. Световой пучок, отраженный от зеркала M_1 , на пути к объективу φ трижды проходит сквозь пластинку P_1 , а пучок, отраженный зеркалом M_2 проходит сквозь пластинку P_1 только один раз. Чтобы устранить это различие в оптическом пути, не позволяющее работать с полосами в белом свете, между пластинкой P_1 и зеркалом M_2 помещена компенсирующая пластинка P_2 из такого же материала и такой же толщины, как у пластинки P_1 , и параллельная ей. На рис. 3.2 показано изображение M'_2 зеркала M_2 , образованное отражающей поверхностью пластинки P_1 . Оптическая длина пути от источника до точки наблюдения для луча, отразившегося от зеркала M_2 , будет такой же, как и для воображаемого луча, отразившегося от мнимой отражающей поверхности M'_2 . Поэтому можно считать, что интерференционная картина, наблюдаемая в фокальной плоскости объектива φ , возникает из-за воздушного слоя между отражающей поверхностью M_1 и мнимой отражающей поверхностью M'_2 . При параллельных поверхностях M_1 и M'_2 полосы имеют вид концентрических окружностей с центром в фокусе F' объектива.

Если после разделительной пластинки P_1 пучки имеют одинаковую интенсивность, то распределение интенсивности в фокальной плоскости объектива описывается формулой (3.10), где разность хода волновых фронтов Δ , как и в случае плоскопараллельной пластинки, в соответствии с формулой (3.14) равна

$$\Delta = 2h \cos \theta.$$

Разность хода при заданном расстоянии h между зеркалом M_1 и изображением M'_2 зеркала M_2 , т.е. при фиксированном положении подвижного зеркала, зависит только от угла θ наклона лучей по отношению к оптической оси. Конкретному значению угла θ соответствует кольцо в фокальной плоскости объектива, радиус которого равен $f'g\theta$, где f' – фокусное расстояние объектива. Поэтому положение и размер светлых и темных колец не зависит от положения источника S , т.е. можно использовать протяженный источник. При этом интерференционная картина (интерференционные кольца или полосы) локализована в фокальной плоскости объектива φ . Центру интерференционной картины ($\theta = 0$) соответствует максимальная разность хода $\Delta_{\max} = 2h$, равная удвоенному расстоянию между зеркалом M_1 и изображением M'_2 зеркала M_2 .

До сих пор предполагалось, что источник S излучает монохроматический свет. Определим требование, которому должен удовлетворять источник излучения в интерференционном устройстве.

Предположим, что источник излучает две очень узкие, близкие друг к другу спектральные линии с частотами ω_1 и ω_2 . При этом можно считать, что вместо одного имеется два расположенных в одном месте источника, независимо друг от друга излучающих волны с частотами ω_1 и ω_2 . В интерферометре с таким источником каждая из волн создает свою интерференционную картину и эти картины просто налагаются одна на другую. Полная интенсивность при наложении двух интерференционных картин определится равенством вида

$$I(\Delta) = I_1(\Delta) + I_2(\Delta) = 2I_{10}(1 + \cos k_1\Delta) + 2I_{20}(1 + \cos k_2\Delta), \quad (3.16)$$

где $k = \frac{\omega}{c} = \frac{2\pi}{\lambda}$ – волновое число, а I_0 – равномерная интенсивность волны.

При $I_{10} = I_{20} = I_0$ равенство (3.16) принимает вид:

$$I(\Delta) = 4I_0 \left[1 + \cos \frac{\delta k}{2} \Delta \cos k\Delta \right], \quad (3.17)$$

где $k = \frac{1}{2}(k_1 + k_2)$; $\delta k = k_2 - k_1$.

Из формулы (3.17) следует, что при разности хода волнового фронта, равной Δ , ближайший максимум определяется условием $\cos k\Delta = \pm 1$ при $\cos \frac{1}{2} \delta k \Delta \cos k\Delta > 0$, а ближайший минимум определяется тем же условием

при $\cos \frac{1}{2} \delta k \Delta \cos k\Delta < 0$, т.е. в соответствии с формулой (3.17)

$$I_{\max} = 4I_0 \left[1 + \left| \cos \frac{1}{2} \delta k \Delta \right| \right], \quad \text{а} \quad I_{\min} = 4I_0 \left[1 - \left| \cos \frac{1}{2} \delta k \Delta \right| \right]. \quad \text{При этом в}$$

соответствии с формулой (3.11) получаем

$$K = \left| \cos \frac{1}{2} \delta k \Delta \right|.$$

Контрастность интерференционной картины принимает свое максимальное значение, равное единице, при $\Delta = 0$, т.е. вблизи полосы нулевого порядка, где освещенность в минимумах приближается к нулю. С увеличением разности хода контрастность уменьшается и обращается в

нуль при $\frac{1}{2} \delta k \Delta = \frac{\pi}{2}$. При этом при малой разности частот

$\delta\omega = \omega_1 - \omega_2 = c\delta k$ получаем, что светлые полосы интерференционной картины излучения одной спектральной линии совпадают с темными

полосами излучения другой линии при $\Delta = \frac{\lambda^2}{2\delta\lambda}$. Введение разности хода

между волновыми фронтами эквивалентно задержке одного из них во

времени. Поэтому способность световых колебаний в одной точке исходного волнового фронта к интерференции после его разделения на два фронта и последующего их соединения с некоторой разностью хода называется временной когерентностью. Максимальная разность хода, при которой возможна интерференция, называется длиной когерентности излучения $l_{\text{ког}}$, а соответствующее ей запаздывание – временем когерентности $\tau_{\text{ког}} = \frac{l_{\text{ког}}}{c}$. Уменьшение когерентности световых колебаний

с увеличением временной задержки, т.е. уменьшение видности интерференционных полос при возрастании разности хода, связано с конечной шириной спектральной линии источника квазимонохроматического света. Такое излучение можно рассматривать как совокупность не скоррелированных между собой отдельных монохроматических волн, частоты которых сплошь заполняют некоторый интервал $\delta\omega$, малый по сравнению со средней частотой ω . Каждая монохроматическая волна из этой совокупности создает в интерферометре свою картину полос, при этом полное распределение освещенности определяется простым наложением этих картин.

Будем считать, что отдельные монохроматические компоненты равномерно заполняют спектральный интервал шириной $\delta\omega$ («прямоугольный» контур линии). Разделим мысленно этот интервал на две равные части. При этом вполне очевидно, что любой из монохроматических компонент первой части найдется монохроматическая компонента из второй части с разностью частот, равной $\frac{1}{2}\delta\omega$. Распределение освещенности в интерференционной картине от каждой пары определяется формулой (3.17) при $\delta k = \frac{1}{2c}\delta\omega \cong \pi \frac{\delta\lambda}{\lambda^2}$. Контрастность полос полной интерференционной картины в этом случае становится равной нулю при такой разности хода $\Delta_{\text{max}} = l_{\text{ког}}$, когда $\frac{1}{2}\delta k \Delta_{\text{max}} = \frac{\pi}{2}$. В результате получаем

$$l_{\text{ког}} = \frac{\lambda^2}{\delta\lambda}. \quad (3.18)$$

Длине когерентности $l_{\text{ког}}$ соответствует максимально возможный порядок интерференции

$$m_{\text{max}} = \frac{l_{\text{ког}}}{\lambda} = \frac{\lambda}{\delta\lambda}.$$

Для белого света (солнце, лампа накаливания, дуга с угольными электродами) и визуального наблюдения эффективный диапазон длин волн $\delta\lambda$ простирается приблизительно от 400 до 700 нм, т.е. $\delta\lambda = 300$ нм при

$\lambda \approx 600$ нм. Однако, учитывая селективную чувствительность глаза, можно считать, что возможный порядок интерференции

$$m > \frac{\lambda}{\delta\lambda} \approx 2.$$

Используя излучение узких спектральных линий газоразрядных источников низкого давления, можно наблюдать интерференцию при оптической разности хода, достигающей нескольких десятков сантиметров. Основная причина, ограничивающая в этом случае длину когерентности, – хаотическое тепловое движение излучающих атомов, приводящее к доплеровскому уширению спектральных линий. Так, например, для излучения оранжевой линии $\lambda = 605,78$ нм стандартной криптоновой лампы, используемой в современном эталоне длины (по определению, 1 м равен 1650763,73 длин волн этой линии), длина когерентности достигает 0,8 м. Исключительно высокая монохроматичность излучения стабилизированных по частоте газовых лазеров позволяет получить интерференцию при разности хода в несколько километров.

Условие временной когерентности световых колебаний можно записать в виде

$$\Delta < l_{\text{ког}} = c\tau_{\text{ког}}.$$

3.3. Интерферометр Тваймана-Грина

Заменим в схеме, представленной на рис. 3.2, протяженный источник излучения S точечным, расположенным в переднем фокусе объектива F_k , как показано на рис. 3.3. Интерферометр, типовая схема которого представлена на рис. 3.3, предложен и запатентован Твайманом и Грином в 1916 году. Коллимированный пучок лучей образует плоский волновой фронт, соответствующий осевому на рис. 3.2, при этом угол $\theta = 0$. Поэтому кольцевые полосы в интерференционной картине будут отсутствовать и все интерференционное поле будет иметь равномерную освещенность. Возникающие при перемещении зеркала M_1 изменения интенсивности расположенного в его заднем фокусе F'_n приемника излучения P . При равных амплитудах колебаний складываемых после светоделителя P монохроматических волновых фронтов интенсивность света определится формулой (3.10) в виде

$$I_{\sigma}(\Delta) = 2I_{0\sigma}(1 + \cos 2\pi\sigma\Delta), \quad (3.19)$$

где $\sigma = \frac{1}{\lambda}$; в спектроскопии величину σ принято называть волновым числом. Как было показано, характер изменения контрастности интерференционных полос при изменении разности хода (например, путем перемещения зеркала M_1 на рис. 3.3 в направлении оптической оси)

определяется спектральным составом света. Следовательно, при спектральной плотности лучистого потока $P(\sigma)$ в параллельном пучке лучей за объективом Φ_k интенсивность света, падающего на объектив Φ_{Π} приемного устройства, без учета потерь в оптической системе интерферометра определяется очевидным выражением

$$I(\Delta) = \int_0^{\infty} P(\sigma)(1 + \cos 2\pi\sigma\Delta) d\sigma. \quad (3.20)$$

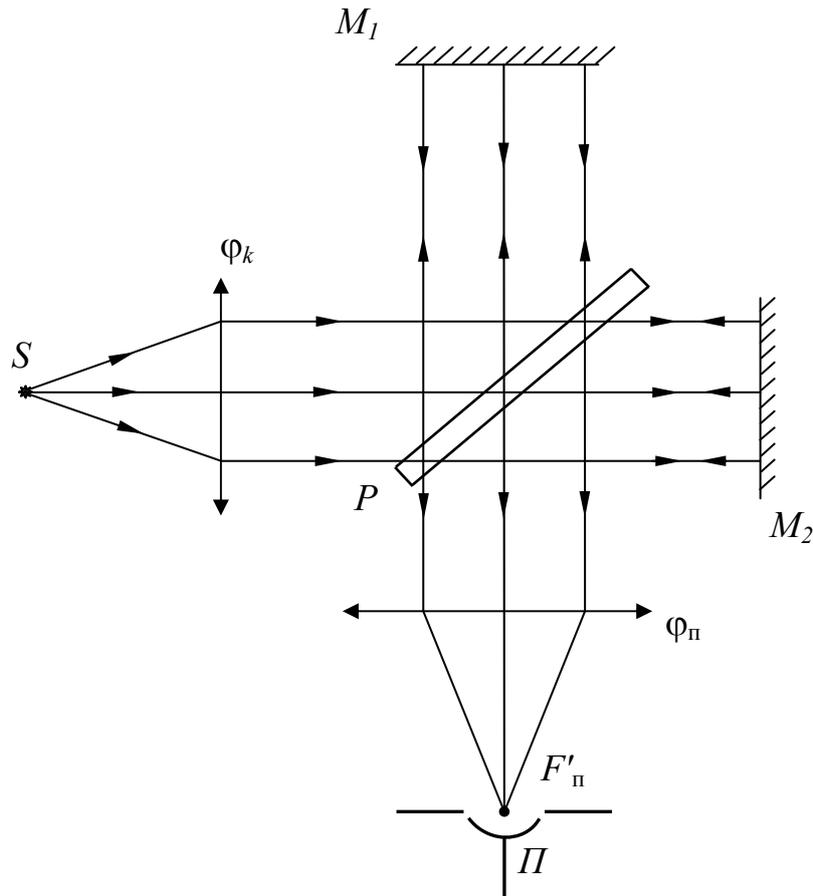


Рис. 3.3. Схема интерферометра Тваймана-Грина

Этот интеграл можно записать в виде

$$I(\Delta) = \int_0^{\infty} P(\sigma) d\sigma + \int_0^{\infty} P(\sigma) \cos(2\pi\sigma\Delta) d\sigma. \quad (3.21)$$

Когда два зеркала расположены на равном оптическом расстоянии от разделителя пучка, то $\Delta = 0$ и $I(\Delta) = I(0) = 2 \int_0^{\infty} P(\sigma) d\sigma$. При этом имеем

$$I(\Delta) - \frac{1}{2} I(0) = \int_0^{\infty} P(\sigma) \cos(2\pi\sigma\Delta) d\sigma. \quad (3.22)$$

Левая часть выражения (3.22) определена при $\sigma > 0$, а при $\sigma < 0$ можно считать ее равной нулю. При этих условиях выражение (3.22) представляет собой обратное косинус-преобразование Фурье для зависимости $I(\Delta) - \frac{1}{2}I(0)$. Тогда косинус-преобразование Фурье этой зависимости принимает вид

$$P(\sigma) = \int_0^{\infty} \left[I(\Delta) - \frac{1}{2}I(0) \right] \cos(2\pi\sigma\Delta) d\Delta. \quad (3.23)$$

Уравнение (3.23) часто называют основным уравнением Фурье-спектрологии. Оно дает возможность вычислить для каждого конкретного значения σ , а, следовательно, для каждого конкретного значения длины волны λ , спектральную плотность $P(\sigma)$ светового потока, излучаемого источником S , путем интегрирования выражения в правой части.

Изменение Δ достигается при перемещении зеркала. Точность, с которой можно контролировать величину перемещения, ставит предел точности прибора в целом. В результате рассматриваемый метод измерения спектральной плотности излучения применим главным образом для работы на длинных волнах и широко использовался на протяжении многих лет в инфракрасной области спектра. Однако в настоящее время, благодаря всестороннему техническому совершенствованию область длин волн расширена в видимый и даже ультрафиолетовый диапазоны.

3.4. Интерферометр Физо

Обратимся к рис. 2.4, где представлено главное сечение прозрачной пленки с плоскими поверхностями, образующими клин, угол которого достаточно мал и равен α . Из точечного источника S световой луч падает на первую поверхность пленки в точку A под углом θ к нормали к поверхности. Преломленный в точке A луч в результате френелева отражения от второй поверхности в точке B выходит из пленки в точке C на первой поверхности. Вторым луч, исходящий из точки S , падает на поверхность пленки в некоторую точку D . Отраженный в результате френелева отражения в этой точке луч встречается с первым лучом в некоторой точке P . При этом разность оптических путей этих лучей между точками S и P равна:

$$\Delta = nSA + n'(AB + BC) + nCP - n(SD + DP).$$

Представляет интерес частный случай, когда точка D совпадает с точкой C , т.е. когда второй луч падает в точку C . При этом вполне очевидно, что точка P также будет совпадать с точкой C , как показано на рис. 3.4. В этом случае разность оптических путей для рассматриваемых лучей равна:

$$\Delta = nSA + n'(AB + BP) - nSP. \quad (3.24)$$

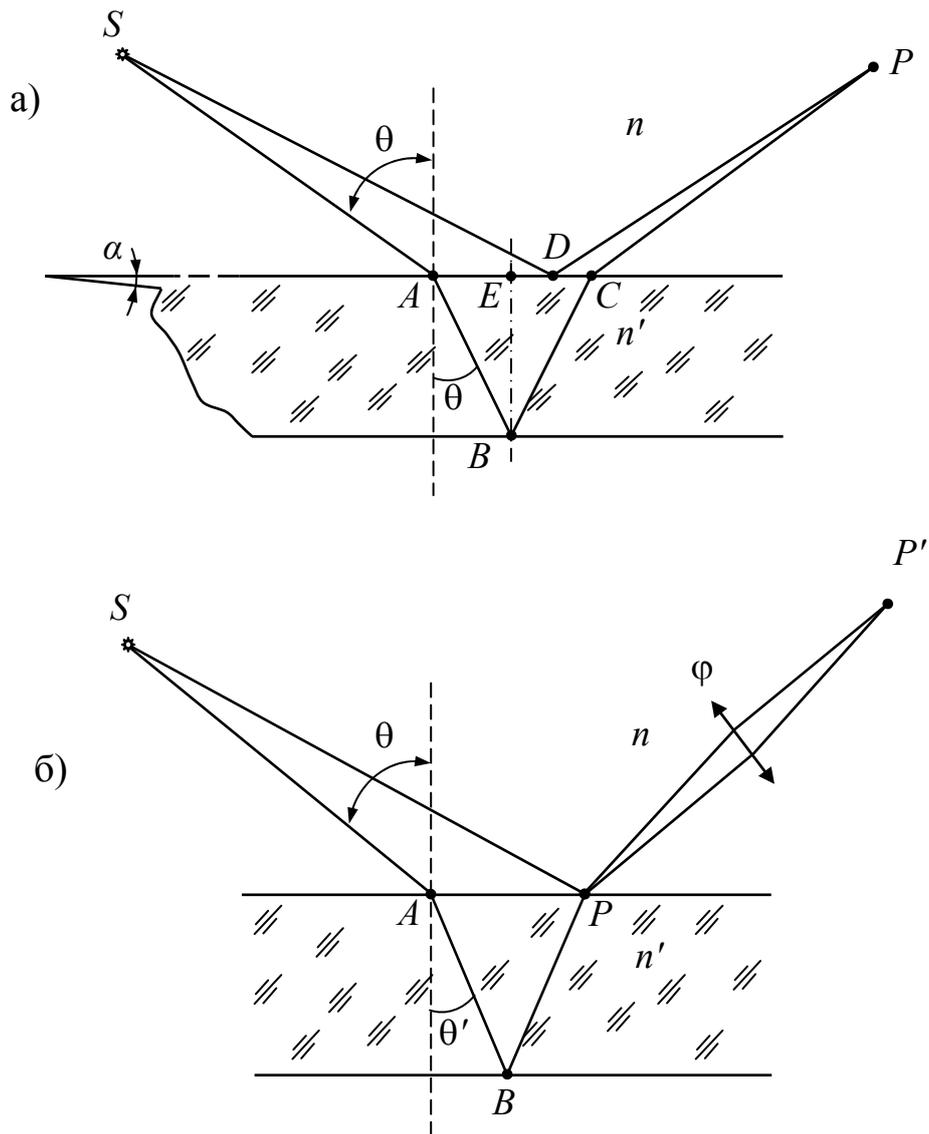


Рис. 3.4. Локализация интерференционных полос, возникающая после отражения и преломления света на поверхностях клина

Полагая угол α и толщину пленки $h = BE$ вблизи точки B измеренную по нормали к первой поверхности в точке B , достаточно малыми, равенство (3.24) заменим приближенным равенством вида:

$$\begin{aligned} \Delta &\approx nSA + n'(AB + BP) - n(SP + AP \sin \theta) = \\ &= n'(AB + BP) - n AP \sin \theta \end{aligned}$$

или

$$\Delta = n'(AB + BP - AP \sin \theta'). \quad (3.25)$$

Но

$$AB + BP \approx \frac{2h}{\cos \theta'}, \text{ а } AP \approx 2htg\theta'.$$

Полученные приближенные равенства позволяют преобразовать равенство (3.25) к виду

$$\Delta \approx 2n'h \left(\frac{1}{\cos \theta'} - \frac{\sin^2 \theta'}{\cos \theta'} \right) = 2n'h \cos \theta'. \quad (3.26)$$

При этом разность фаз световых колебаний в точке P равна

$$\delta = \frac{4\pi}{\lambda_0} n'h \cos \theta'. \quad (3.27)$$

При малой клиновидности (малом изменении толщины) пленки величина h практически одинакова для всех пар лучей от протяженного источника, приходящих в точку P' , сопряженную с P , и различие величин δ в точке P' определяется главным образом различием значений $\cos \theta'$. Если интервал изменений $\cos \theta'$ достаточно мал, то интервал изменения величины δ в точке P' много меньше π даже с источником значительных размеров и полосы отчетливо видны. Очевидно, что они локализованы на первой поверхности пленки. Практически условие малости интервала изменений $\cos \theta'$ можно выполнить при наблюдении в направлении, близком к нормальному, или при ограничении входного зрачка. Последнее условие вполне выполняется при наблюдении интерференционной картины невооруженным глазом, поскольку зрачок глаза и сам по себе может быть достаточно мал.

Учитывая изменение фазы на π при отражении на одной из поверхностей пленки, из выражений (3.9) и (3.27) находим, что максимумы интенсивности в P' (и, очевидно, в P) определяются из условия

$$2n'h \overline{\cos \theta'} \pm \frac{\lambda_0}{2} = (m-1)\lambda_0, \quad (3.28a)$$

а минимумы интенсивности – из условия

$$2n'h \overline{\cos \theta'} \pm \frac{\lambda_0}{2} = \frac{2m-1}{2} \lambda_0, \quad (3.28b)$$

где $m = 1, 2, 3, \dots$; $\overline{\cos \theta'}$ – среднее значение $\cos \theta'$ для точек источника, из которых свет приходит в точку P' . Из соотношений (3.28) следует, что если $\overline{\cos \theta'}$ достаточно постоянен, то интерференционные полосы соответствуют совокупности мест пленки, где оптические толщины одинаковы. По этой причине такие полосы часто называют полосами равной толщины. Их можно наблюдать в тонкой воздушной прослойке между отражающими поверхностями двух прозрачных пластинок. Вблизи нормального падения лучей условия (3.28) для темной полосы при $\overline{\cos \theta'} = 1$ и длине волны в воздухе $\lambda = \frac{\lambda_0}{n'}$ имеют вид

$$h = \frac{m\lambda}{2}, \text{ где } m = 0, 1, 2, \dots \quad (3.29)$$

Таким образом, интерференционные полосы определяют контуры слоев равной толщины, причем расстояние между полосами соответствует изменению толщины на $\frac{\lambda}{2}$.

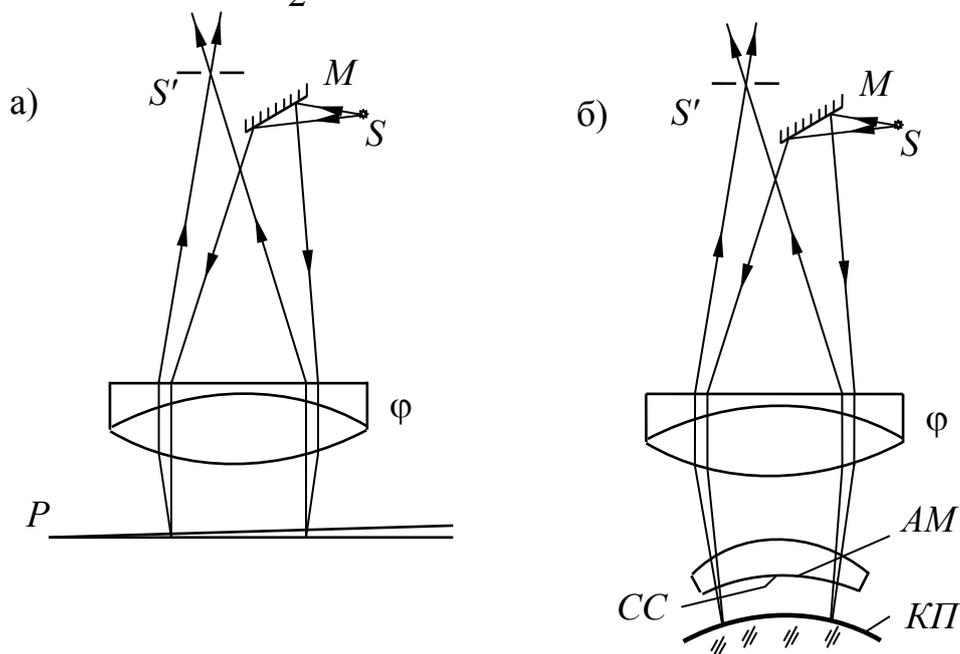


Рис. 3.5. Схема интерферометра Физо

Итак, необходимым условием требуемого контраста (четкости) интерференционных полос служит ограниченность диапазона изменения величины $\cos \theta'$, соответствующего каждой точке пленки, при этом при $\cos \theta' \approx 1$ полосы определяют контуры слоев равной оптической толщины. Оба эти условия одновременно выполняются на большой площади пленки в интерферометре Физо, схема которого представлена на рис. 3.5. Расходящийся пучок света из квазимонохроматического источника S после отражения от диагонального зеркала M преобразуется объективом φ в параллельный пучок, который падает по нормали на пленку P. Свет, отраженный поверхностями пленки, снова проходит через объектив φ и собирается в отверстии S' в фокальной плоскости объектива φ. Глаз, помещенный сразу за S' и аккомодированный на пленку, видит полосы, определяющие контуры слоев оптической толщины на всей поверхности пленки, освещенной объективом φ. Эти полосы часто называют полосами Физо. Полосы Физо можно получить и в толстых пленках при условии, что источник света достаточно мал.

Такие интерферометры находят применение в оптическом производстве для проверки постоянства оптических толщин плоскопараллельных прозрачных пластинок. Интерферометры Физо, варианты модифицированной схемы которых показан на рис. 3.5, находят применение и для контроля формы оптических поверхностей. На этой

схеме AM – апланатический мениск; CC – сфера сравнения (эталонная поверхность); $KП$ – контролируемая сферическая поверхность.

Схема интерферометра Тваймана-Грина, представленная на рис. 3.3, эквивалентна схеме интерферометра Физо, но в отличие от последней здесь пути световых пучков полностью разделены. Оптическая разность хода между выходящими после делителя лучами равна

$$\Delta = 2nh, \quad (3.30)$$

где h – разность расстояний от осевой точки деления пучков до зеркал M_1 и M_2 . Соответствующая разность фаз равна

$$\delta = \frac{4\pi}{\lambda_0} nh. \quad (3.31)$$

Заметим, что интерференционные полосы, являющиеся геометрическим местом точек, для которых величина δ постоянна, образуются в любой плоскости той области, где встречаются оба луча из S . В этом случае принято говорить, что такие полосы не локализованы. Они всегда наблюдаются с точечным источником и их контрастность (их видность) зависит только от относительной интенсивности обеих волн.

В соответствии с соотношениями (3.31) и (3.9) находим, что глаз, помещенный в фокальную плоскость объектива $\varphi_{\text{п}}$, увидит в некоторой точке P светлую полосу, если

$$nh = \frac{m-1}{2} \lambda_0, \quad (3.32a)$$

или темную полосу, если

$$nh = \frac{2m-1}{4} \lambda_0, \quad (3.32b)$$

где $m = 1, 2, \dots$

Если одно из зеркал в интерферометре Тваймана-Грина представляет собой сферу со стрелкой шарового сегмента h , то интерференционная картина будет выглядеть в виде системы чередующихся светлых и темных колец, при этом порядок интерференции будет равен $m = \frac{2nh}{\lambda_0}$. Различные

модификации интерферометра Тваймана-Грина находят широкое применение в оптическом производстве при контроле оптических поверхностей и деталей.

4. Когерентность квазимонохроматического света

4.1. Квазимонохроматический свет

В монохроматическом волновом поле амплитуда колебаний в любой точке P постоянна, тогда как фаза линейно меняется со временем. Свет от реального физического источника никогда не бывает строго монохроматическим, так как даже самая узкая спектральная линия обладает конечной шириной. Кроме того, физический источник света имеет конечные размеры и состоит из огромного числа элементарных излучателей. Таким образом, свет спектральной линии недостаточно монохроматичен и состоит из цугов волн конечной длины, большое количество которых проходит за любой интервал времени, необходимый для наблюдения.

Пусть $F(t)$ – световое возмущение в некоторой точке в момент времени t , вызванное одиночным волновым цугом, и пусть F равно нулю для $|t| \geq t_0$. Представим эту функцию в виде интеграла Фурье:

$$F(t) = \int_{-\infty}^{\infty} f(\nu) \exp(-2\pi i \nu t) d\nu, \quad (4.1)$$

где, согласно обратному преобразованию Фурье,

$$f(\nu) = \int_{-\infty}^{\infty} F(t) \exp(2\pi i \nu t) dt. \quad (4.2)$$

Если за время наблюдения эту точку минует N таких волновых цугов, то полное световое возмущение можно определить выражением вида:

$$V(t) = \sum_{n=1}^N F(t - t_n), \quad (4.3)$$

где величина t_n обозначает момент времени прихода соответствующего волнового цуга. Средняя интенсивность света за временной интервал 2τ , необходимый для наблюдения, равна

$$I = \frac{1}{2\tau} \int_{-\tau}^{\tau} |V(t)|^2 dt \approx \frac{1}{2\tau} \int_{-\infty}^{\infty} |V(t)|^2 dt, \quad (4.4)$$

если τ велико по сравнению с t_0 (половиной длительности волнового цуга).

Используя соотношения (4.1) и (4.3), получаем

$$V(t) = \int_{-\infty}^{\infty} \upsilon(\nu) \exp(-2\pi i \nu t) d\nu, \quad (4.5)$$

где $\upsilon(\nu) = f(\nu) \sum_{n=1}^N \exp(2\pi i \nu t_n)$. (4.6)

Следовательно, в соответствии с теоремой Парсеваля имеем

$$\int_{-\infty}^{\infty} |V(t)|^2 dt = \int_{-\infty}^{\infty} |v(v)|^2 dv = \int_{-\infty}^{\infty} |f(v)|^2 \sum_{n=1}^N \sum_{m=1}^N \exp[2\pi i v(t_n - t_m)] dv. \quad (4.7)$$

$$\begin{aligned} \text{Здесь } \sum_{n=1}^N \sum_{m=1}^N \exp[2\pi i v(t_n - t_m)] &= N + \sum_{n \neq m} \exp[2\pi i v(t_n - t_m)] = \\ &= N + 2 \sum_{n < m} \cos 2\pi v(t_n - t_m). \end{aligned} \quad (4.8)$$

Поскольку значения величины t_n распределяются случайным образом, вероятности того, что слагаемые в правой части выражения (4.8) будут положительными или отрицательными, одинаковы. Следовательно, средняя величина двойной суммы в выражении (4.8), а, следовательно, и в выражении (4.7) при большом числе таких реализаций равна N . При этом, используя выражения (4.4) и (4.7), получаем, что средняя интенсивность равна

$$I = \frac{N}{2\tau} \int_{-\infty}^{\infty} |f(v)|^2 dv, \quad (4.9)$$

т.е. пропорциональна интегралу от интенсивностей $i(v) = |f(v)|^2$ (некогерентная суперпозиция) монохроматических компонент, составляющих одиночные волновые цуги.

Предположим, что длительность всех волновых цугов равна Δt и в течение этого времени $F(t)$ есть простая периодическая функция с частотой ν_0 , т.е.

$$\left. \begin{aligned} F(t) &= f_0 \exp[-2\pi i \nu_0 t], & \text{если } |t| \leq \frac{\Delta t}{2} \\ F(t) &= 0, & \text{если } |t| > \frac{\Delta t}{2}, \end{aligned} \right\} \quad (4.10)$$

где f_0 – постоянная величина.

Тогда, учитывая выражения (4.2) и (4.10), получаем

$$f(v) = f_0 \int_{-\frac{\Delta t}{2}}^{\frac{\Delta t}{2}} \exp[2\pi i(v - \nu_0)t] dt = f_0 \cdot \Delta t \left\{ \frac{\sin[\pi(v - \nu_0)\Delta t]}{\pi(v - \nu_0)\Delta t} \right\}. \quad (4.11)$$

Заметим, что функция $\left\{ \frac{\sin[\pi(v - \nu_0)\Delta t]}{\pi(v - \nu_0)\Delta t} \right\}^2$, определяющая

распределение интенсивностей фурье-компонент в выражении (4.10), принимает первое нулевое значение при $\Delta v \cdot \Delta t = \pm 1$, при этом промежутки

времени Δt равен длительности одного волнового цуга. При $\Delta t \rightarrow \infty$ частотный интервал $\Delta \nu \rightarrow 0$, т.е. при этом имеем монохроматический свет.

Полученные результаты дают идеализированное представление о свете реальных источников, в излучении которых волновые цуги беспорядочно изменяются. И тем не менее, для любого светового возмущения $V(t)$ или его фурье-образа $v(\nu)$ можно определить величины Δt и $\Delta \nu$, которые мы вправе рассматривать соответственно как среднюю длительность волновых цугов, составляющих V , и как эффективный частотный диапазон фурье-спектра. Можно показать, что эти величины всегда удовлетворяют соотношению:

$$\Delta t \cdot \Delta \nu \geq \frac{1}{4\pi}. \quad (4.12)$$

Следует отметить, что в большинстве практически важных случаев знак неравенства в соотношении (4.12) можно заменить знаком, обозначающим порядок величины.

Промежуток времени Δt , входящий в соотношение (4.12), известен как время когерентности света. Если $\bar{\lambda}_0$ – средняя длина волны, то длина Δl , определяемая соотношением

$$\Delta l = c\Delta t \approx \frac{c}{\Delta \nu} = \frac{(\bar{\lambda}_0)^2}{\Delta \lambda_0}, \quad (4.13)$$

известна как длина когерентности. Свет, удовлетворяющий условию $\frac{\Delta \lambda_0}{\lambda_0} \ll 1$, называют квазимонохроматическим светом.

4.2. Корреляционные функции световых пучков лучей

Рассмотрим волновое поле, образованное протяженным полихроматическим источником S . Световое возмущение в некоторой точке P будем считать вещественной скалярной функцией положения и времени

$$V^{(r)}(P, t) = \int_0^{\infty} a(\nu) \cos[\varphi(\nu) - 2\pi\nu t] d\nu. \quad (4.14)$$

Свяжем с $V^{(r)}$ комплексную функцию

$$V(P, t) = \int_0^{\infty} a(\nu) \exp\{i[\varphi(\nu) - 2\pi\nu t]\} d\nu. \quad (4.15)$$

При этом имеем

$$V(P, t) = V^{(r)}(P, t) + iV^{(i)}(P, t), \quad (4.16)$$

где $V^{(i)}(P, t) = \int_0^{\infty} a(\nu) \sin[\varphi(\nu) - 2\pi\nu t] d\nu. \quad (4.17)$

Эти величины с огромной частотой изменяются во времени, а поэтому наблюдаемая интенсивность $I(P)$ пропорциональна среднему значению квадрата модуля светового возмущения:

$$I(P) = 2 \langle (V^{(r)}(P, t))^2 \rangle = \langle V(P, t) \cdot V^*(P, t) \rangle. \quad (4.18)$$

Рассмотрим теперь две точки P_1 и P_2 в волновом поле. Можно экспериментально определить не только $I(P_1)$ и $I(P_2)$, но и интерференционные эффекты, возникающие при суперпозиции колебаний, исходящих из этих точек. Представим себе, что в исследуемое поле помещен непрозрачный экран E_p с небольшими отверстиями в точках P_1 и P_2 , и рассмотрим распределение интенсивности на втором экране E_q , находящемся на некотором расстоянии от экрана E_p в направлении, противоположном направлению на источник, как показано на рис. 4.1.

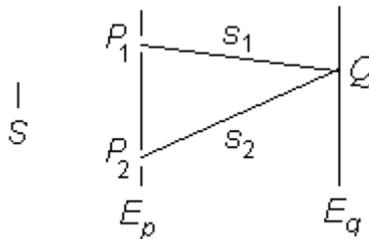


Рис. 4.1. Интерференция полихроматического света от протяженного источника света S

Не нарушая общности вывода, будем считать, что показатель преломления среды между двумя экранами равен единице. Пусть s_1 и s_2 – расстояния от произвольной точки Q экрана E_q до точек P_1 и P_2 соответственно. Точки P_1 и P_2 можно считать центрами вторичных возмущений, так что комплексное возмущение в точке Q можно определить выражением вида:

$$V(Q, t) = K_1 V(P_1, t - t_1) + K_2 V(P_2, t - t_2). \quad (4.19)$$

Здесь t_1 и t_2 – время распространения света от точек P_1 и P_2 до точки Q соответственно, т.е. $t_1 = \frac{s_1}{c}$, $t_2 = \frac{s_2}{c}$, где c – скорость света в вакууме.

Коэффициенты K_1 и K_2 обратно пропорциональны отрезкам s_1 и s_2 и, кроме того, зависят от размера отверстий и геометрии всего устройства (углов падения и углов дифракции в точках P_1 и P_2). Из теории дифракции света следует, что фаза вторичных волн, распространяющихся из точек P_1 и P_2 , отличается от фазы первичной волны на четверть периода, а, следовательно, коэффициенты K_1 и K_2 являются чисто мнимыми величинами.

Из выражений (4.18) и (4.19) следует, что интенсивность света в точке Q равна

$$I(Q) = K_1 K_1^* \langle V_1(t-t_1) \cdot V_1^*(t-t_1) \rangle + K_2 K_2^* \langle V_2(t-t_2) \cdot V_2^*(t-t_2) \rangle + \\ + K_1 K_2^* \langle V_1(t-t_1) \cdot V_2^*(t-t_2) \rangle + K_2 K_1^* \langle V_2(t-t_2) \cdot V_1^*(t-t_1) \rangle. \quad (4.20)$$

Здесь и далее вместо $V(P_i, t)$ для краткости будем писать $V_i(t)$. Поскольку предполагается, что поле стационарно, то $\langle V_i(t) \rangle$ не зависит от выбора начала отсчета времени, и корреляционные функции $\Gamma_{ij}(t', t'') = \langle V_i(t+t') \cdot V_j(t+t'') \rangle$ зависят лишь от разности $t'-t''$. При этом, например,

$$\langle V_1(t-t_1) \cdot V_1^*(t-t_1) \rangle = \langle V_1(t) \cdot V_1^*(t) \rangle = I_1. \quad (4.21)$$

Изложенные соображения позволяют выражение (4.20) представить в виде:

$$I(Q) = |K_1|^2 I_1 + |K_2|^2 I_2 + 2 \cdot |K_1 K_2| \Gamma_{12}^{(r)}(\tau), \quad (4.22)$$

где $\Gamma_{12}^{(r)}(\tau)$ – вещественная часть функции

$$\Gamma_{12}(\tau) = \langle V_1(t+\tau) \cdot V_2^*(t) \rangle, \quad (4.23)$$

$$\tau = \frac{s_2 - s_1}{c}.$$

Понятие, определяемое выражением (4.23), служит основным в теории частичной когерентности. Его называют взаимной когерентностью световых колебаний в точках P_1 и P_2 , причем колебания в точке P_1 рассматриваются в момент времени, запаздывающий на величину τ по сравнению с моментом времени колебаний в точке P_2 . Функцию $\Gamma_{12}(\tau)$ называют взаимной функцией когерентности волнового поля. Если точка P_2 совпадает с точкой P_1 , получаем

$$\Gamma_{11}(\tau) = \langle V_1(t+\tau) \cdot V_1^*(t) \rangle. \quad (4.24)$$

В этом случае говорят об автокогерентности световых колебаний в точке P_1 . В общей теории стационарных случайных процессов $\Gamma_{12}(\tau)$ называется взаимной корреляционной функцией величин $V_1(t)$ и $V_2(t)$, а $\Gamma_{11}(\tau)$ – автокорреляционной функцией величины $V_1(t)$.

При $\tau=0$ соотношение (3.24) сводится к выражению для обычной интенсивности: $\Gamma_{11}(0)=I_1$, $\Gamma_{22}(0)=I_2$. Вполне очевидно, что член $|K_1|^2 I_1$ в выражении (3.22) определяет интенсивность света, которая наблюдалась бы в точке Q , если бы открытым было бы лишь отверстие в точке P_1 ($K_2=0$). Аналогичный смысл имеет и член $|K_2|^2 I_2$. Обозначим эти значения интенсивностей соответственно через $I^{(1)}(Q)$ и $I^{(2)}(Q)$, т.е.

$$\left. \begin{aligned} I^{(1)}(Q) &= |K_1|^2 I_1 = |K_1|^2 \cdot \Gamma_{11}(0), \\ I^{(2)}(Q) &= |K_2|^2 I_2 = |K_2|^2 \cdot \Gamma_{22}(0), \end{aligned} \right\} \quad (4.25)$$

а также нормируем $\Gamma_{12}(\tau)$, положив

$$\gamma_{12}(\tau) = \frac{\Gamma_{12}(\tau)}{\sqrt{\Gamma_{11}(0)}\sqrt{\Gamma_{22}(0)}} = \frac{\Gamma_{12}(\tau)}{\sqrt{I_1}\sqrt{I_2}}. \quad (4.26)$$

Величину $\gamma_{12}(\tau)$ называют комплексной степенью когерентности световых колебаний. Выражения (4.25) и (4.26) позволяют формулу (4.22) представить в виде:

$$I(Q) = I^{(1)}(Q) + I^{(2)}(Q) + 2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)} \cdot \gamma_{12}^{(r)}(\tau), \quad (4.27)$$

где $\gamma_{12}^{(r)}$ – вещественная часть γ . Соотношение (4.27) выражает общий закон интерференции для стационарных оптических полей. Выражение (4.27) остается справедливым при условии, что $s_2 - s_1 = c \cdot \tau$ заменено на разность хода $[P_2Q] - [P_1Q]$, если свет от точек P_1 и P_2 попадает в точку Q не прямо, а через промежуточную оптическую систему, причем $[P_iQ] = \int_{P_i}^Q ndS$,

и если можно пренебречь эффектами дисперсии. При таком обобщении формула (4.27) выполняется и в том случае, когда два интерферирующих пучка световых лучей получаются из первичного пучка путем «деления амплитуды», например, как в интерферометре Майкельсона. В этом случае в соотношении (4.27) вместо $\gamma_{12}^{(r)}(\tau)$ будет входить $\gamma_{11}^{(r)}(\tau)$.

В отличие от возмущения $V^{(r)}$ корреляционные функции $\gamma_{12}^{(r)}$ и $\Gamma_{12}^{(r)}$ представляют собой величины, которые можно определить из эксперимента. Для этого, используя схему, показанную на рис. 4.1, измеряют интенсивность $I(Q)$ в некоторой точке Q позади экрана, для которой $P_2Q - P_1Q = c \cdot \tau$. Затем отдельно измеряют интенсивности $I^{(1)}(Q)$ и $I^{(2)}(Q)$ света, прошедшего через каждое отверстие. Выразив величину $\gamma_{12}^{(r)}$ через три найденные величины, в соответствии с выражением (4.27) получим

$$\gamma_{12}^{(r)} = \frac{I(Q) - I^{(1)}(Q) - I^{(2)}(Q)}{2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)}}. \quad (4.28)$$

Для определения $\Gamma_{12}^{(r)}$ следует также измерить интенсивности $I(P_1)$ и $I(P_2)$ в каждом отверстии. Тогда в соответствии с соотношениями (4.26) и (4.28) получаем

$$\Gamma_{12}^{(r)} = \sqrt{I(P_1)}\sqrt{I(P_2)} \cdot \gamma_{12}^{(r)} = \frac{1}{2} \sqrt{\frac{I(P_1)I(P_2)}{I^{(1)}(Q)I^{(2)}(Q)}} [I(Q) - I^{(1)}(Q) - I^{(2)}(Q)].$$

Смысл γ_{12} легче всего понять, если выражение (4.27) представить в несколько иной форме.

Обозначим через $\bar{\nu}$ среднюю частоту света и запишем

$$\gamma_{12}(\tau) = |\gamma_{12}(\tau)| \exp\{i[\alpha_{12}(\tau) - 2\pi\bar{\nu}\tau]\}, \quad (4.29)$$

где $\alpha_{12}(\tau) = 2\pi\bar{\nu}\tau + \arg \gamma_{12}(\tau)$.

Тогда выражение (4.27) можно представить в виде:

$$I(Q) = I^{(1)}(Q) + I^{(2)}(Q) + 2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)} \cdot |\gamma_{12}(\tau)| \cos[\alpha_{12}(\tau) - \delta], \quad (4.30)$$

где параметр τ и разность фаз δ равны:

$$\tau = \frac{s_2 - s_1}{c}, \quad \delta = 2\pi\bar{\nu}\tau = \frac{2\pi}{\lambda}(s_2 - s_1), \quad (4.31)$$

а $\bar{\lambda}$ – средняя длина волны. Если $|\gamma_{12}(\tau)|$ достигает своего максимального значения, равного единице, то интенсивность в точке Q будет совпадать с интенсивностью, которая получилась бы при строго монохроматическом свете с длиной волны $\bar{\lambda}$ и разности фаз между колебаниями в точках P_1 и P_2 , равной $\alpha_{12}(\tau)$. В таком случае можно сказать, что колебания в точках P_1 и P_2 (при соответствующем времени задержки τ между ними) когерентны. Если $\gamma_{12}(\tau)$ имеет другое экстремальное значение, а именно, нулевое, то последний член в выражении (4.30) исчезает. Тогда не возникает никаких видимых интерференционных эффектов от этих пучков и можно сказать, что колебания некогерентны. Если $|\gamma_{12}(\tau)|$ не совпадает ни с одним из двух экстремальных значений, т.е. если $0 < |\gamma_{12}(\tau)| < 1$, то говорят, что колебания частично когерентны, причем модуль $|\gamma_{12}(\tau)|$ определяет степень их когерентности.

При любом значении $|\gamma_{12}(\tau)|$ выражение (4.30) можно представить в виде:

$$I(Q) = |\gamma_{12}(\tau)| \cdot \left\{ I^{(1)}(Q) + I^{(2)}(Q) + 2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)} \times \right. \quad (4.32)$$

$$\left. \times \cos[\alpha_{12}(\tau) - \delta] \right\} + [1 - |\gamma_{12}(\tau)|] \cdot [I^{(1)}(Q) + I^{(2)}(Q)].$$

Можно считать, что члены в фигурных скобках определяют результат суперпозиции двух когерентных световых пучков с интенсивностями $|\gamma_{12}(\tau)|I^{(1)}(Q)$ и $|\gamma_{12}(\tau)|I^{(2)}(Q)$ и относительной разностью фаз $\alpha_{12}(\tau) - \delta$; остальные члены выражения (4.32) определяют результат суперпозиции двух некогерентных световых пучков с интенсивностями $[1 - |\gamma_{12}(\tau)|]I^{(1)}(Q)$ и $[1 - |\gamma_{12}(\tau)|]I^{(2)}(Q)$. Таким образом, свет от обоих отверстий, достигающий точки Q , можно считать как бы состоящим из смеси когерентной и некогерентной частей с отношением интенсивностей

$$\frac{I_{\text{ког}}}{I_{\text{неког}}} = \frac{|\gamma_{12}(\tau)|}{1 - |\gamma_{12}(\tau)|} \quad (4.33)$$

ИЛИ

$$\frac{I_{\text{ког}}}{I_{\text{полн}}} = |\gamma_{12}(\tau)|, \quad (4.34)$$

где $I_{\text{полн}} = I_{\text{ког}} + I_{\text{неког}}$.

Теория частичной когерентности привлекательна тем, что она оперирует с величинами (а именно, с корреляционными функциями и с усредненными по времени интенсивностями), которые в принципе можно определить экспериментально.

4.3. Интерференция квазимонохроматического света

Возвращаясь к схеме, показанной на рис. 4.1, заметим, что интенсивность света в точке Q интерференционной картины определяется выражениями (4.30) и (4.31). При этом максимумы и минимумы интенсивности вблизи точки Q с хорошей точностью определяются выражениями:

$$\begin{aligned} I_{\text{max}} &= I^{(1)}(Q) + I^{(2)}(Q) + 2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)}|\gamma_{12}(\tau)| \\ I_{\text{min}} &= I^{(1)}(Q) + I^{(2)}(Q) - 2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)}|\gamma_{12}(\tau)| \end{aligned} \quad (4.35)$$

Следовательно, видность (контраст) интерференционных полос в точке Q равна:

$$k(Q) = \frac{I_{\text{max}} - I_{\text{min}}}{I_{\text{max}} + I_{\text{min}}} = \frac{2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)}}{I^{(1)}(Q) + I^{(2)}(Q)}|\gamma_{12}(\tau)|. \quad (4.36)$$

Если $I^{(1)}(Q) = I^{(2)}(Q)$, то соотношение (4.36) принимает вид:

$$k(Q) = |\gamma_{11}(\tau)|, \quad (4.37)$$

т.е. в этом случае видность полос равна степени когерентности.

Согласно выражению (4.30) и соотношениям (4.31) положение максимумов интенсивности вблизи точки Q определяется выражением:

$$\frac{2\pi}{\lambda}(s_2 - s_1) - \alpha_{12}(\tau) = 2m\pi \quad (m=0, \pm 1, \pm 2, \dots).$$

Таким же выражением определялось бы положение максимумов при освещении отверстий строго монохроматическим светом с длиной волны $\bar{\lambda}$, если бы фаза колебаний в точке P_1 отставала относительно фазы в точке P_2 на $\alpha_{12}(\tau)$. Отставание фазы на величину 2π соответствует смещению интерференционной картины в направлении, параллельном P_1P_2 , на величину, равную $\frac{a}{d}\bar{\lambda}$, где d – расстояние между точками P_1 и P_2 , а a – расстояние между экранами E_p и E_q на рис. 4.1. Следовательно, интерференционные полосы, полученные с квазимонохроматическим светом, смещены относительно полос, которые образовались бы при синфазном освещении точек P_1 и P_2 монохроматическим светом, на величину

$$x = \frac{\bar{\lambda}}{2\pi} \frac{a}{d} \alpha_{12}(\tau) \quad (4.38)$$

в направлении, параллельном линии, соединяющей отверстия.

Итак, из измерений видности (контраста) и положения интерференционных полос можно определить амплитуду и фазу комплексной степени когерентности квазимонохроматических пучков света. Можно показать, что

$$\Gamma_{12}(\tau) = \langle V(P_1, t + \tau) V^*(P_2, t) \rangle = 4 \int_0^{\infty} G_{12}(\nu) \exp(-2\pi i \nu \tau) d\nu, \quad (4.39)$$

где функцию $G_{12}(\nu)$ можно назвать взаимной спектральной плотностью световых колебаний в точках P_1 и P_2 . Из уравнений (4.26) и (4.39) следует, что

$$\gamma_{11}(\tau) = \frac{\int_0^{\infty} G(\nu) \exp(-2\pi i \nu \tau) d\nu}{\int_0^{\infty} G(\nu) d\nu}.$$

Следовательно, согласно обратной теореме Фурье, функция $G(\nu)$ пропорциональна фурье-образу величины $\gamma_{11}(\tau)$. Но модуль $\gamma_{11}(\tau)$ равен видности полос, образующихся в соответствующем интерференционном эксперименте, а фаза $\gamma_{11}(\tau)$ связана с их положением простым соотношением.

На практике время задержки τ одного интерференционного пучка относительно другого сравнительно мало, что позволяет упростить полученные формулы. Согласно уравнениям (4.26), (4.29) и (4.39) имеем

$$\begin{aligned} |\Gamma_{12}(\tau)| \exp[i\alpha_{12}(\tau)] &= \sqrt{I_1} \sqrt{I_2} |\gamma_{12}(\tau)| \exp[i\alpha_{12}(\tau)] = \\ &= 4 \int_0^{\infty} G_{12}(\nu) \exp[-2\pi i(\nu - \bar{\nu})\tau] d\nu \end{aligned} \quad (4.40)$$

Если модуль $|\tau|$ так мал, что $|(\nu - \bar{\nu})\tau| \ll 1$ для всех частот, при которых $|G_{12}(\nu)|$ имеет заметную величину, т.е. если

$$|\tau| \ll \frac{1}{\Delta\nu}, \quad (3.41)$$

то с достаточно малой погрешностью можем заменить экспоненциальный член подынтегрального выражения (4.40) единицей. Условие (4.41) означает, что, согласно соотношению $\Delta\nu \approx \frac{1}{\Delta t}$, модуль $|\tau|$ должен быть мал по сравнению со временем когерентности света. При этом условии модули

$|\Gamma_{12}(\tau)|$, $|\gamma_{12}(\tau)|$ и фаза $\alpha_{12}(\tau)$ незначительно отличаются от $|\Gamma_{12}(0)|$, $|\gamma_{12}(0)|$ и $\alpha_{12}(0)$ соответственно.

Удобно положить:

$$J_{12} = \Gamma_{12}(0) = \langle V_1(t) \cdot V_2^*(t) \rangle, \quad (4.42)$$

$$\mu_{12} = \gamma_{12}(0) = \frac{\Gamma_{12}(0)}{\sqrt{\Gamma_{11}(0)}\sqrt{\Gamma_{22}(0)}} = \frac{J_{12}}{\sqrt{J_{11}}\sqrt{J_{22}}} = \frac{J_{12}}{\sqrt{I_1}\sqrt{I_2}} \quad (4.43)$$

$$\beta_{12} = \alpha_{12}(0) = \arg \gamma_{12}(0) = \arg \mu_{12}. \quad (4.44)$$

Здесь $J_{12} = J(P_1, P_2)$ и т.д. Используя уравнения (4.26) и (4.29), при выполнении условия (4.41) получаем:

$$\gamma_{12}(\tau) \approx |\mu_{12}| \exp[i(\beta_{12} - 2\pi\nu\tau)] = \mu_{12} \exp(-2\pi i\nu\tau), \quad (4.45)$$

$$\Gamma_{12}(\tau) \approx |J_{12}| \exp[i(\beta_{12} - 2\pi\nu\tau)] = J_{12} \exp(-2\pi i\nu\tau). \quad (4.46)$$

Таким образом, при выполнении условия (4.41) во всех полученных формулах можно заменить величины $\gamma_{12}(\tau)$ и $\Gamma_{12}(\tau)$ на величины, стоящие в правых частях выражений (4.45) и (4.46) соответственно. В частности, закон интерференции (4.30) при этом примет вид:

$$I(Q) \approx I^{(1)}(Q) + I^{(2)}(Q) + 2\sqrt{I^{(1)}(Q)}\sqrt{I^{(2)}(Q)} \cdot |\mu_{12}| \cos(\beta_{12} - \delta). \quad (4.47)$$

Этот закон будет выполняться до тех пор, пока разность хода $|s_2 - s_1| = c|\tau|$ между интерферирующими пучками будет мала по сравнению с длиной когерентности $\frac{c}{\Delta\nu}$, т.е. до тех пор, пока

$$|\Delta L| = |s_2 - s_1| = \frac{\bar{\lambda}}{2\pi} \delta \ll \frac{\bar{\lambda}^2}{\Delta\lambda}, \quad (4.48)$$

где использовано соотношение $\frac{c}{\Delta\nu} = \frac{\lambda^2}{\Delta\lambda}$.

Уравнение (4.47) является основной формулой элементарной (квазимонохроматической) теории частичной когерентности.

Если справедливо уравнение (4.47) (т.е. соблюдены неравенства (4.41) или (4.48)), то корреляция между колебаниями в любых двух точках P_1 и P_2 волнового поля характеризуется не $\Gamma_{12}(\tau)$, а J_{12} , т.е. величиной, которая зависит не от разности времени τ , а от положения этих точек. Как следует из выражения (4.45), в пределах применимости элементарной теории имеем

$$|\gamma_{12}(\tau)| \approx |\mu_{12}|. \quad (4.49)$$

Таким образом, величина $|\mu_{12}|$ при $0 \leq |\mu_{12}| \leq 1$ определяет степень когерентности колебаний в точках P_1 и P_2 . Из уравнения (4.47) следует, что фаза β_{12} величины μ_{12} представляет собой эффективную разность фаз этих колебаний. Величину μ_{12} (так же, как и $\gamma_{12}(\tau)$, частным случаем которой она

является) обычно называют комплексной степенью когерентности (иногда комплексным коэффициентом когерентности), а величину J_{12} – взаимной интенсивностью.

Определим взаимную интенсивность J_{12} и комплексную степень когерентности μ_{12} для точек P_1 и P_2 экрана E_p , освещаемого протяженным квазимонохроматическим первичным источником S . Источник излучения будем считать плоским, параллельным экрану E_p , а среду между источником и экраном – однородной. Будем также считать, что малы как линейные размеры источника S по сравнению с расстоянием OO_p между источником и экраном, как показано на рис. 4.2, так и углы между линией OO_p и линиями, соединяющими произвольную точку G источника с точками P_1 и P_2 .

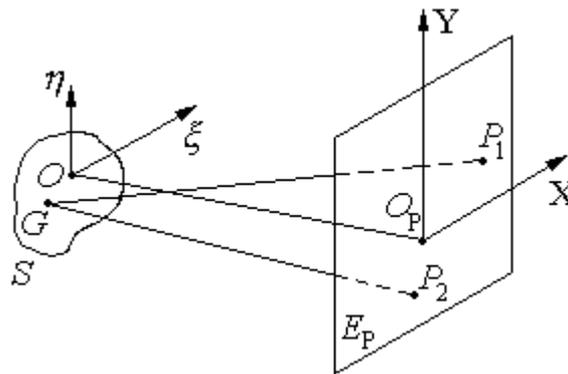


Рис. 4.2. К теореме Ван-Циттерта-Цернике

Мысленно разделим источник излучения на элементы dS_1, dS_2, \dots с линейными размерами, малыми по сравнению со средней длиной волны $\bar{\lambda}$, и центрами, находящимися в точках G_1, G_2, \dots . Если $V_{m1}(t)$ и $V_{m2}(t)$ – комплексные световые возмущения в точках P_1 и P_2 , возбуждаемые элементом dS_m , то общее возмущение в этих точках равно

$$V_1(t) = \sum_m V_{m1}(t), \quad V_2(t) = \sum_m V_{m2}(t). \quad (4.50)$$

Следовательно,

$$J(P_1, P_2) = \langle V_1(t)V_2^*(t) \rangle = \sum_m \langle V_{m1}(t)V_{m2}^*(t) \rangle + \sum_{m \neq n} \langle V_{m1}(t)V_{n2}^*(t) \rangle. \quad (4.51)$$

Световые колебания, создаваемые различными элементами источника, можно считать статистически независимыми (взаимно некогерентными), т.е. можно считать, что среднее значение поля, создаваемого ими, равно нулю:

$$\langle V_{m1}(t)V_{n2}^*(t) \rangle = \langle V_{m1}(t) \rangle \langle V_{n2}^*(t) \rangle = 0, \quad (4.52)$$

когда $m \neq n$.

Пусть R_{m1} и R_{m2} – расстояния от элемента dS_m источника до точек P_1 и P_2 соответственно. Тогда

$$\left. \begin{aligned} V_{m1}(t) &= A_m \left(t - \frac{R_{m1}}{v} \right) \cdot \frac{\exp \left[-2\pi i \bar{v} \cdot \left(t - \frac{R_{m1}}{v} \right) \right]}{R_{m1}}, \\ V_{m2}(t) &= A_m \left(t - \frac{R_{m2}}{v} \right) \cdot \frac{\exp \left[-2\pi i \bar{v} \cdot \left(t - \frac{R_{m2}}{v} \right) \right]}{R_{m2}}, \end{aligned} \right\} \quad (4.53)$$

где величина $|A_m|$ характеризует силу, а $\arg A_m$ – фазу излучения m -го элемента; v – скорость света в среде между источником и экраном.

Следовательно,

$$\begin{aligned} \langle V_{m1}(t) V_{m2}^*(t) \rangle &= \langle A_m \left(t - \frac{R_{m1}}{v} \right) A_m^* \left(t - \frac{R_{m2}}{v} \right) \rangle \times \\ &\times \frac{\exp \frac{2\pi i \bar{v} \cdot (R_{m1} - R_{m2})}{v}}{R_{m1} R_{m2}} = \langle A_m(t) A_m^* \left(t - \frac{R_{m2} - R_{m1}}{v} \right) \rangle \times \\ &\times \frac{\exp \frac{2\pi i \bar{v} \cdot (R_{m1} - R_{m2})}{v}}{R_{m1} R_{m2}}. \end{aligned} \quad (4.54)$$

Если разность хода $R_{m2} - R_{m1}$ мала по сравнению с длиной когерентности света, в аргументе A_m^* можно пренебречь запаздыванием, равным $\frac{R_{m2} - R_{m1}}{v}$. Тогда из соотношений (4.51), (4.52) и (4.54) получаем

$$J(P_1, P_2) = \sum_m \langle A_m(t) A_m^*(t) \rangle \cdot \frac{\exp \frac{2\pi i \bar{v} \cdot (R_{m1} - R_{m2})}{v}}{R_{m1} R_{m2}}. \quad (4.55)$$

Величина $\langle A_m(t) A_m^*(t) \rangle$ определяет интенсивность излучения элемента источника dS_m . В любом практически интересном случае число элементов источника можно считать настолько большим, что их распределение можно считать непрерывным.

Обозначая через $I(G)$ интенсивность света на единицу площади источника и полагая при этом, что $I(G)dS$ – энергия, излученная элементом dS источника, т.е. $I(G_m)dS_m = \langle A_m(t) A_m^*(t) \rangle$, вместо выражения (4.55) получаем

$$J(P_1, P_2) = \int_S I(G) \cdot \frac{\exp \left[i \bar{k} (R_1 - R_2) \right]}{R_1 R_2} dS, \quad (4.56)$$

где R_1 и R_2 – расстояния между произвольной точкой G источника и точками P_1 и P_2 , а $\bar{k} = \frac{2\pi\bar{v}}{v} = 2\pi\bar{\lambda}$ – волновое число в среде между

источником излучения и экраном. Комплексная степень когерентности $\mu(P_1, P_2)$ в соответствии с выражениями (4.56) и (4.43) равна

$$\mu(P_1, P_2) = \frac{1}{\sqrt{I(P_1)}\sqrt{I(P_2)}} \int_S I(G) \cdot \frac{\exp[i\bar{k}(R_1 - R_2)]}{R_1 R_2} dS, \quad (4.57)$$

где $I(P_1)$ и $I(P_2)$ – интенсивности в точках P_1 и P_2 соответственно, при этом

$$\left. \begin{aligned} I(P_1) = J(P_1, P_1) &= \int_S \frac{I(G)}{R_1^2} dS, \\ I(P_2) = J(P_2, P_2) &= \int_S \frac{I(G)}{R_2^2} dS. \end{aligned} \right\} \quad (4.58)$$

Интеграл (4.57) подобен интегралу, который получается при вычислении в соответствии с принципом Гюйгенса-Френеля комплексного светового возмущения в дифракционной картине, возникающей при дифракции сферической волны на отверстии в непрозрачном экране. Точнее, из формулы (4.57) следует, что комплексная степень когерентности, которая описывает корреляцию колебаний в фиксированной точке P_2 и переменной точке P_1 плоскости, освещенной протяженным квазимонохроматическим первичным источником излучения, равна нормированной комплексной амплитуде в соответствующей точке P_1 некоторой дифракционной картины с центром в точке P_2 . Эта картина получится, если заменить источник излучения дифракционным отверстием такого же размера и формы и заполнить его сферической волной, сходящейся в точке P_2 , причем распределение амплитуд по волновому фронту в отверстии должно быть пропорциональным распределению интенсивности по источнику излучения. Этот результат впервые был получен Ван-Циттертом, а позднее, более простым способом, Цернике. Его принято называть теоремой Ван-Циттерта-Цернике.

В большинстве приложений можно считать, что интенсивность $I(G)$ не зависит от положения точки G на поверхности S , т.е. интенсивность света, излучаемая каждой точкой поверхности, одна и та же. Тогда соответствующая проблема определения светового возмущения в плоскости экрана E_p совпадает с проблемой дифракции сферической волны постоянной амплитуды на отверстии такого же размера и формы, как и источник.

Пусть (ξ, η) – координаты произвольной точки G источника в системе координат с началом в точке O и пусть (X_1, Y_1) и (X_2, Y_2) – координаты точек P_1 и P_2 в системе координат с началом в точке O_p и осями,

параллельными осям первой системы, как показано на рис. 4.2. Тогда, если R – расстояние между точками O и O_p , то

$$R_1^2 = (X_1 - \xi)^2 + (Y_1 - \eta)^2 + R^2 \text{ или}$$

$$R_1 \approx R + \frac{(X_1 - \xi)^2 + (Y_1 - \eta)^2}{2R}. \quad (4.59)$$

Для расстояния R_2 справедливо точно такое же выражение, а, следовательно,

$$R_1 - R_2 \approx \frac{(X_1^2 + Y_1^2) - (X_2^2 + Y_2^2)}{2R} - \frac{(X_1 - X_2)\xi + (Y_1 - Y_2)\eta}{R}. \quad (4.60)$$

В знаменателе подынтегральных выражений (4.56) и (4.57) расстояния R_1 и R_2 с достаточно хорошей точностью можно заменить расстоянием R . Обозначим

$$\frac{X_1 - X_2}{R} = p, \quad \frac{Y_1 - Y_2}{R} = q, \quad (4.61)$$

$$\psi = \frac{\bar{k}[(X_1^2 + Y_1^2) - (X_2^2 + Y_2^2)]}{2R}. \quad (4.62)$$

Тогда выражение (4.57) можно представить в виде:

$$\mu_{12} = \frac{\exp(i\psi) \iint_S I(\xi, \eta) \exp[-i\bar{k}(p\xi + q\eta)] d\xi d\eta}{\iint_S I(\xi, \eta) d\xi d\eta}. \quad (4.63)$$

Следовательно, если линейные размеры источника и расстояние между точками P_1 и P_2 малы по сравнению с расстоянием от этих точек до источника, то степень когерентности $|\mu_{12}|$ равна абсолютному значению нормированного преобразования Фурье от функции, описывающей интенсивность источника.

Для однородного источника излучения в виде круга радиуса ρ с центром в точке O комплексная степень когерентности равна:

$$\mu_{12} = |\mu_{12}| \exp(i\beta_{12}) = \left(\frac{2J_1(v)}{v} \right) e^{i\psi}, \quad (4.64)$$

где $v = \bar{k}\rho\sqrt{p^2 + q^2} = \frac{2\pi}{\lambda} \frac{\rho}{R} \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2}$; J_1 – функция Бесселя первого рода первого порядка.

Величина ψ , определяемая соотношением (4.62), представляет собой разность фаз, равную $\frac{2\pi}{\lambda}(OP_1 - OP_2)$, величиной которой можно пренебречь, когда $OP_1 - OP_2 \ll \bar{\lambda}$.

Величина $\left| \frac{2J_1(v)}{v} \right|$ монотонно уменьшается по мере увеличения

величины v от значения, равного единице при $v=0$, до нулевого значения при $v=3,83$. Таким образом, степень когерентности монотонно уменьшается, когда точки P_1 и P_2 удаляются друг от друга. При удалении точек P_1 и P_2 на расстояние

$$P_1P_2 = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2} = \frac{0,61R\bar{\lambda}}{\rho} \quad (4.65)$$

достигается полная некогерентность. При дальнейшем увеличении v вновь возникает небольшая когерентность, но степень ее остается меньше 0,14. Полная некогерентность вновь наступает при $v=7,02$. Проходя через нуль, функция $J_1(v)$ каждый раз меняет знак, т.е. фаза $\beta_{12}=\arg\mu_{12}$ при этом изменяется на π . Следовательно, после каждого исчезновения полос яркие и темные полосы меняются местами.

Функция $\left| \frac{2J_1(v)}{v} \right|$ монотонно уменьшается от значения, равного единице при $v=0$, до 0,88 при $v=1$, т.е. при

$$P_1P_2 = \frac{0,16R\bar{\lambda}}{\rho}.$$

Считая отклонение в 12% максимально допустимым отклонением от идеального значения, равного единице, получаем, что квазимонохроматический однородный источник углового радиуса $\alpha = \frac{\rho}{R}$ почти когерентно освещает площадку в виде круга диаметром $0,16 \frac{\bar{\lambda}}{\alpha}$. Этот результат полезен при оценке размеров источника, требуемых при проведении экспериментов по интерференции и дифракции света.

В качестве примера рассмотрим размеры «области когерентности» вокруг произвольной точки экрана, освещаемого непосредственно Солнцем. Угол 2α , под которым солнечный диск виден на поверхности Земли, составляет около $0^\circ 32' \approx 0,00931$ рад. Следовательно, если пренебречь изменением яркости по поверхности солнечного диска, диаметр d области когерентности приблизительно равен $2 \frac{0,16\bar{\lambda}}{0,00931} \approx 34\bar{\lambda}$.

Для средней длины волны $\bar{\lambda}=5,5 \cdot 10^{-4}$ мм получаем $d \approx 0,02$ мм.

При выводе формулы Ван-Циттерта-Цернике было принято, что среда между источником S и точками P_1 и P_2 однородна. Обобщим эту формулу, например, на случай среды, состоящей из ряда однородных областей с разными показателями преломления. Оптическую систему, состоящую из ряда однородных сред с различными показателями преломления, можно

охарактеризовать с помощью функции пропускания среды $K(G, P, \nu)$, которая определяется как комплексная амплитуда светового возмущения в точке P , обусловленного возмущением частоты ν с единичной амплитудой и нулевой фазой, испускаемым точечным источником излучения, расположенным в точке G элемента dS . Для однородной среды из принципа Гюйгенса-Френеля следует, что $K(G, P, \nu) = -i \frac{\exp(i\bar{k}R)}{\bar{\lambda}R}$, где R – расстояние GP ; при этом предполагается, что угол между линией GP и нормалью к элементу dS достаточно мал. В более общем случае множитель $\frac{\exp(i\bar{k}R_{mj})}{R_{mj}}$ необходимо заменить на $i\bar{\lambda}K(G_m, P_j, \bar{\nu})$ и тогда, переходя к непрерывному распределению точек G_m , вместо выражения (4.56) получаем следующее соотношение:

$$J(P_1, P_2) = \bar{\lambda}^2 \int_S I(G) K(G, P_1, \bar{\nu}) K^*(G, P_2, \bar{\nu}) dS. \quad (4.66)$$

В соответствии с выражениями (4.66) и (4.43) имеем

$$\mu(P_1, P_2) = \frac{\bar{\lambda}^2}{\sqrt{I(P_1)}\sqrt{I(P_2)}} \int_S I(G) K(G, P_1, \bar{\nu}) K^*(G, P_2, \bar{\nu}) dS, \quad (4.67)$$

где интенсивности в точках P_1 и P_2 соответственно равны

$$I(P_1) = J(P_1, P_1), \quad I(P_2) = J(P_2, P_2).$$

Будем считать, что

$$\left. \begin{aligned} i\bar{\lambda}K(G, P_1, \bar{\nu})\sqrt{I(G)} &= U(G, P_1), \\ i\bar{\lambda}K(G, P_2, \bar{\nu})\sqrt{I(G)} &= U(G, P_2). \end{aligned} \right\} \quad (4.68)$$

При этом формулы (4.66) и (4.67) принимают вид:

$$J(P_1, P_2) = \int U(G, P_1) U^*(G, P_2) dS, \quad (4.69)$$

$$\mu(P_1, P_2) = \frac{1}{\sqrt{I(P_1)}\sqrt{I(P_2)}} \int U(G, P_1) U^*(G, P_2) dS. \quad (4.70)$$

Заметим, что определяемая соотношением (4.68) величина $U(G, P)$ пропорциональна возмущению, которое возникало бы в точке P от строго монохроматического точечного источника, испускающего излучение частотой $\bar{\nu}$ (с силой $\sqrt{I(G)}$ и нулевой фазой) и расположенного в точке G .

Таким образом, формулы (4.69) и (4.70) можно рассматривать как соотношения, выражающие взаимную интенсивность $J(P_1, P_2)$ и комплексную степень когерентности $\mu(P_1, P_2)$, обусловленные протяженным квазимоно-хроматическим источником излучения, через возмущения, создаваемые в точках P_1 и P_2 каждой точкой ассоциированного монохроматического источника.

Полезное при решении проблем когерентности в инструментальной оптике выражение (4.70) было впервые предложено Г.Г. Гопкинсом, исходящим из эвристических предположений. Основная ценность этой формулы состоит в том, что так же, как и теорема Ван-Циттерта-Цернике, она позволяет рассчитать комплексную степень когерентности света, излучаемого некогерентным источником, без явного использования процесса усреднения.

Предположим, что взаимная интенсивность известна для любой пары точек на воображаемой поверхности A , пересекающей пучок, испускаемый протяженным первичным источником S квазимонохроматического света. Покажем, что в этом случае можно определить взаимную интенсивность для каждой пары точек на любой другой поверхности B , освещаемой светом от поверхности A либо непосредственно, либо через оптическую систему.

Будем считать, что среда между поверхностями A и B однородна и ее показатель преломления равен единице. Пусть $U(S, Q_1)$ и $U(S, Q_2)$ – световые возмущения в точках Q_1 и Q_2 на поверхности B , как показано на рис. 4.3, обусловленные произвольной точкой G ассоциированного монохроматического источника излучения S .

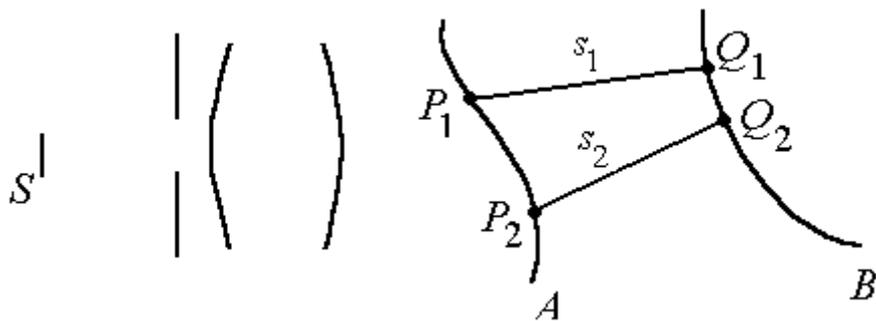


Рис. 4.3. Распространение взаимной интенсивности

Тогда, согласно выражению (4.69), взаимная интенсивность $J(Q_1, Q_2)$ определяется выражением

$$J(Q_1, Q_2) = \int_S U(G, Q_1) U^*(G, Q_2) dS. \quad (4.71)$$

Величины $U(G, Q_1)$ и $U(G, Q_2)$ на основании принципа Гюйгенса-Френеля можно выразить через возмущения во всех точках поверхности A в виде:

$$U(G, Q_1) = \int_A U(G, P_1) \frac{\exp(i\bar{k}s_1)}{s_1} \cdot x_1 dP_1. \quad (4.72)$$

Здесь отрезок s_1 – расстояние между точкой Q_1 и произвольной точкой P_1 на поверхности A , x_1 – коэффициент наклона в точке P_1 , описывающий изменение амплитуды вторичных волн в зависимости от направления;

$\bar{k} = 2\pi \frac{\bar{v}}{c}$ – среднее волновое число. Для небольших отклонений от

нормали к поверхности A в точке P_1 можно принять $x_1 \approx -\frac{i}{\lambda}$. Используя

выражение (4.72) и аналогичное выражение для $U(G, Q_2)$, получаем

$$U(G, Q_1)U^*(G, Q_2) = \iint_{AA} U(G, P_1)U^*(G, P_2) \times \frac{\exp[i\bar{k}(s_1 - s_2)]}{s_1 s_2} x_1 x_2^* dP_1 dP_2, \quad (4.73)$$

при этом точки P_1 и P_2 независимо друг от друга пробегают всю поверхность интегрирования A . Подставив выражение (4.73) в выражение (4.71) и учитывая соотношение (4.69), получаем:

$$J(Q_1, Q_2) = \iint_{AA} J(P_1, P_2) \frac{\exp[i\bar{k}(s_1 - s_2)]}{s_1 s_2} x_1 x_2^* dP_1 dP_2. \quad (4.74)$$

Полученная формула, предложенная Цернике, описывает распространение взаимной интенсивности. При ее выводе неявно предполагалось, что свет от каждой точки поверхности A достигает точек Q_1 и Q_2 . В том случае, когда точки Q_1 и Q_2 совпадают, выражение (4.74) принимает вид:

$$I(Q) = \iint_{AA} \sqrt{I(P_1)} \sqrt{I(P_2)} \mu(P_1, P_2) \frac{\exp[i\bar{k}(s_1 - s_2)]}{s_1 s_2} x_1 x_2^* dP_1 dP_2. \quad (4.75)$$

В этой формуле интенсивность в точке Q представлена в виде суммы вкладов от каждой пары элементов dP_1 и dP_2 произвольной поверхности A , пересекающей световой пучок, как показано на рис. 4.4.

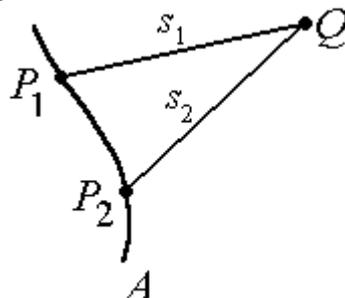


Рис. 4.4. Определение интенсивности света в точке Q

Вклад от каждой пары элементов зависит от интенсивности в точках P_1 и P_2 и содержит в качестве весового множителя соответствующее значение комплексной степени когерентности $\mu(P_1, P_2)$. Соотношение (4.75) можно считать выражением принципа Гюйгенса-Френеля для распространения интенсивности в частично когерентном поле.

Если на пути света от поверхности A до поверхности B находится оптическая система, то, очевидно множитель $x \frac{\exp(i\bar{k}s)}{s}$ следует заменить соответствующей функцией пропускания $K(P, Q)$. Тогда вместо выражения (4.74) получим более общую формулу, а именно:

$$J(Q_1, Q_2) = \int \int_{AA} J(P_1, P_2) K(P_1, Q_1) K^*(P_2, Q_2) dP_1 dP_2. \quad (4.76)$$

5. Основные понятия и соотношения фотометрии

Слово «фотометрия» составлено из двух греческих слов: «фос» (φως) – свет и «метрео» (μετρεω) – измеряю, т.е. в переводе на русский язык, его следует понимать как «световые измерения». В бытовом понимании «свет» – это ощущение, которое возникает у человека под влиянием падающего в его глаз электромагнитного излучения с длинами волн, лежащими в пределах от 380 нм до 780 нм. В технике под словом «свет» понимают то излучение, которое вызывает зрительное ощущение. Измерение этого излучения и составляет задачу фотометрии.

Первой фотометрической работой в истории науки было разделение звезд по их блеску (по освещенности от этих звезд) на шесть классов – звездных величин, выполненное Гиппархом во II веке до Р.Х. В звездном каталоге Птолемея, приложенном к его «Альмагесту», содержатся оценки звездных величин свыше 1000 звезд. Сопоставляя их с современными объективными измерениями освещенности от этих же звезд, установили, что когда одна звезда имеет звездную величину на единицу больше, чем другая, то освещенность от первой в 2,5 раза меньше, чем от второй. Постепенное расширение общего интереса к измерениям света можно проследить по литературным памятникам и отметить его у Данте (XIII в.), у Леонардо да Винчи (XV в.) и у Галилея (XVII в.).

Начало практического применения некоторых видов световых измерений можно, по-видимому, отнести к концу XVI и началу XVII веков и особенно к XVIII веку, когда были опубликованы книги основоположников фотометрии французского ученого Пьера Бугера (1698–1758) (Bouguer P. Essai d'optique sur la gradation de la lumière. – Paris, 1729) и немецкого ученого Иоганна Генриха Ламберта (1728–1777) (Lambert J.H. Photometria, sive de mensura et gradibus luminis, colorum et umbrae – Augsburg, 1760). С тех пор методы световых измерений непрерывно совершенствуются, следуя за ускоряющимся темпом технического прогресса, предъявляющего все более высокие требования к различным видам измерений лучистой энергии.

С древнейших времен и до XIX века единственная возможность замечать и оценивать излучение была связана со зрением человека. Естественно поэтому, что все фотометрические законы и соотношения развивались только в связи с воздействием излучения на глаз наблюдателя и что световые измерения могли осуществляться только в пределах видимого спектра. С появлением приемников, чувствительных к ультрафиолетовым и инфракрасным лучам, содержание фотометрии стало расширяться и в настоящее время ее можно определить как совокупность методов и теории, охватывающих энергетику процессов излучения, распространения и превращения (в частности, поглощения) лучистой энергии в любой части электромагнитного спектра. Однако, чаще всего

фотометрические соотношения применяются к ультрафиолетовому, видимому и инфракрасному излучениям, объединяемым в общем понятии оптического излучения.

5.1. Интенсивность излучения некогерентного источника конечных размеров в геометрическом приближении

Плотность потока энергии электромагнитного поля определяется вектором Умова-Пойнтинга

$$\mathbf{G} = \frac{c}{4\pi} \mathbf{E} \times \mathbf{H} = \frac{c}{4\pi} E\mathbf{H}s,$$

где \mathbf{E} и \mathbf{H} – векторы напряженности электрического и магнитного полей соответственно.

Направление вектора Умова-Пойнтинга перпендикулярно векторам \mathbf{E} и \mathbf{H} и совпадает с направлением распространения электромагнитной волны, а его величина равна энергии, переносимой в единицу времени через единичную площадку, перпендикулярную вектору \mathbf{G} . Таким образом, вектор $\mathbf{G} = Vws$, где w – плотность энергии, а V – скорость распространения электромагнитных волн.

Строго монохроматическое поле, которое можно рассматривать как фурье-компоненту произвольного поля, создает гармонический осциллятор или набор подобных осцилляторов с одинаковой частотой излучения. В оптике обычно имеют дело с источниками, излучающими свет в узком, но конечном диапазоне частот. Такой источник можно рассматривать как набор большого числа гармонических осцилляторов, частоты которых попадают в указанный диапазон.

Для вычисления интенсивности света в какой-либо точке P пространства необходимо просуммировать все поля, созданные каждым осциллятором (элементом источника), т.е.

$$\mathbf{E} = \sum_n \mathbf{E}_n;$$

$$\mathbf{H} = \sum_n \mathbf{H}_n.$$

Так как частоты колебаний в оптическом диапазоне излучения очень велики, непосредственно наблюдать мгновенные значения ни одной из таких быстро осциллирующих величин невозможно. Можно говорить лишь об их значениях, усредненных за некоторый интервал времени, который велик по сравнению с периодом колебаний. Величина среднего по времени вектора Умова-Пойнтинга служит мерой интенсивности света в некоторой точке P , т.е.

$$\begin{aligned}
I(P) &= \langle \mathbf{G} \rangle = \frac{c}{4\pi} \langle \mathbf{E} \times \mathbf{H} \rangle = \frac{c}{4\pi} \left| \sum_{n,m} \langle \mathbf{E}_n \times \mathbf{H}_m \rangle \right| = \\
&= \frac{c}{4\pi} \left| \sum_n \langle \mathbf{E}_n \times \mathbf{H}_n \rangle + \sum_{n \neq m} \langle \mathbf{E}_n \times \mathbf{H}_m \rangle \right|.
\end{aligned} \tag{5.1}$$

Световые колебания, создаваемые различными элементами источника, можно считать независимыми (взаимно некогерентными), причем среднее значение поля равно нулю. Тогда

$$\sum_{n \neq m} \langle \mathbf{E}_n \times \mathbf{H}_m \rangle = 0.$$

Поэтому при решении многих оптических задач можно принять, что вторая сумма в выражении (5.1) равна нулю (в этих случаях говорят, что поля некогерентны).

Тогда

$$I(P) = \frac{c}{4\pi} \left| \sum_n \langle \mathbf{E}_n \times \mathbf{H}_n \rangle \right| = \left| \sum_n \langle \mathbf{G}_n \rangle \right|, \tag{5.2}$$

где \mathbf{G}_n – вектор Умова-Пойнтинга, соответствующий n -му элементу источника.

Направление усредненного по времени вектора Умова-Пойнтинга совпадает с нормалью к геометрическому волновому фронту, а абсолютная его величина равна произведению средней плотности энергии $\langle w \rangle$ на скорость распространения электромагнитных волн V . Волновой фронт излучения точечного источника в однородной среде имеет сферическую форму. Источники излучения конечного размера принято считать точечными, если их угловые размеры малы, при этом в геометрической оптике источник излучения считается точечным, если его угловой размер стремится к нулю.

Свет от источника распространяется, как правило, во все стороны, заполняя все окружающее пространство. Для анализа пространственного распределения излучаемой источником энергии используют представление о телесном угле.

Телесным углом называется часть пространства, заключенная внутри одной полости некоторой конической поверхности (с замкнутой направляющей), порождаемой движением прямой линии (образующей), проходящей через неподвижную точку (вершину конической поверхности), как показано на рис. 5.1, при этом всякая (не проходящая через вершину) линия, которую образующая пересекает при своем движении, называется направляющей. Если вершина телесного угла расположена в центре сферы произвольного радиуса r , то коническая поверхность, ограничивающая телесный угол, вырезает на сфере часть ее поверхности, площадь S которой пропорциональна квадрату радиуса r ,

т.е. $S = \omega r^2$. Коэффициент ω является мерой телесного угла. Единицей телесного угла является стерadian (ср). Телесный угол, равный одному стерadianу, соответствует части поверхности сферы, площадь которой равна площади квадрата со стороной, равной радиусу сферы.

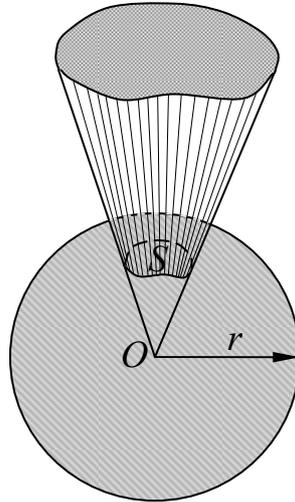


Рис. 5.1. Телесный угол

Коническая поверхность телесного угла может принимать различные формы. Если это поверхность прямого кругового конуса с плоским углом 2α при вершине, то телесный угол (в стерadianах) $\omega = 2\pi(1 - \cos \alpha) = 4\pi \sin^2 \frac{\alpha}{2}$. Если конус телесного угла разворачивается в плоскость, то телесный угол, соответствующий полупространству, оказывается равным 2π ср и, наконец, если площадь S охватывает всю поверхность сферы, то полный телесный угол в каждой точке пространства равен 4π ср. Когда коническая поверхность сжимается около какого-то направления и площадь вырезаемой конусом поверхности сферы становится бесконечно малой, телесный угол также становится бесконечно малым и равным $d\omega = \frac{dS}{r^2}$.

Пусть δS – элемент поверхности волнового фронта, соответствующего определенной точке поверхности источника излучения. Через поверхность δS проходят конические пучки лучей, исходящих из каждой точки источника. Осевые лучи этих пучков заполняют конус с телесным углом $d\omega$, как показано на рис. 5.2. Если угол раствора конуса достаточно мал, то можно пренебречь зависимостью проекции вектора \mathbf{G}_n от направления и записать выражение (5.2) в виде:

$$I(P) = \sum_n \langle \mathbf{G}_n \rangle = \sum_n I_n. \quad (5.3)$$

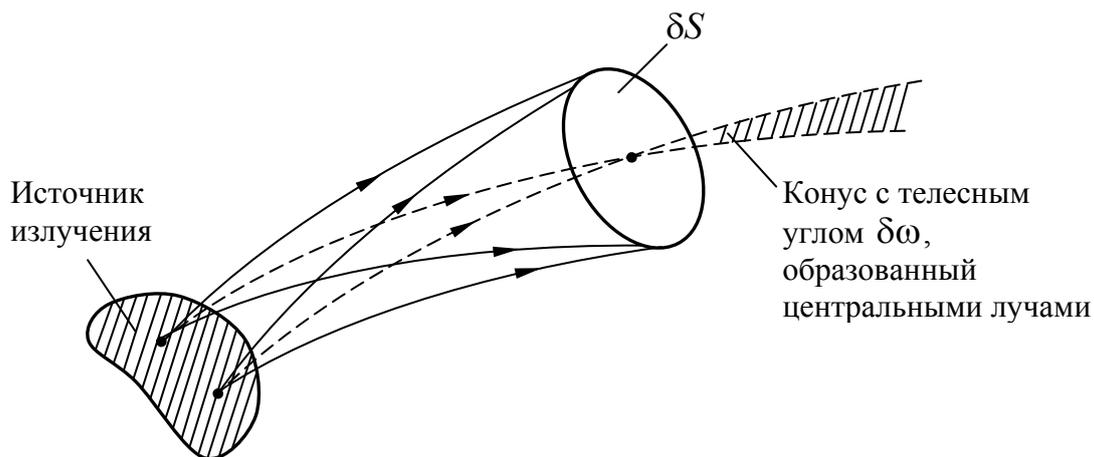


Рис. 5.2. К выводу закона интенсивности в геометрической оптике для некогерентного источника излучения конечных размеров

Предположим, что число точечных источников (осцилляторов) настолько велико, что их распределение без существенной ошибки можно считать непрерывным. Вклад от каждого точечного источника бесконечно мал, однако, суммарный эффект конечен. В этом случае сумма (интеграл) пропорциональна телесному углу $\delta\omega$, т.е.

$$I(P) = L\delta\omega,$$

а полный (усредненный по времени) поток энергии $\delta\Phi_e$, проходящий через элемент δS в единицу времени, равен

$$\delta\Phi_e = L\delta\omega\delta S. \quad (5.4)$$

Формула (5.4) играет важную роль в фотометрии.

5.2. Оптическое излучение. Световой поток. Энергетические и эффективные характеристики оптического излучения

Одним из видов энергии (греч. *energeia* – действие) является излучение, испускаемое телами естественного и искусственного происхождения. Такие тела называют источниками излучения. Известными каждому естественными источниками излучения являются Солнце, Луна, Земля, планеты и созвездия, небо, облака, полярные сияния. Искусственные источники оптического излучения могут быть тепловыми, люминесцентными и газоразрядными. Сюда же относятся светодиоды и лазеры.

Приемниками излучения можно считать все тела живой и неживой природы, поглощающие излучение. По типу действия излучения различают приемники тепловые (болометры, термоэлементы, терморезисторы и др.), фотоэлектрические (фотоэлементы, фотоумножители, фотосопротивления), фотографические (фотографические слои). Среди приемников живой природы, прежде всего,

можно назвать глаза (сетчатку глаз) и кожу человека и листья растений. Приемники излучения обладают различным механизмом преобразования энергии излучения. Исследование этого механизма помогает наиболее эффективно использовать энергию излучения.

Как известно, любое материальное тело, имеющее температуру выше нуля абсолютной шкалы, излучает энергию в окружающее пространство. Следовательно, все тела, с которыми нам приходится иметь дело в жизни, непрерывно обмениваются энергией, так как любой поток излучения переносит энергию от излучающего тела к поглощающему. При тепловом излучении световое излучение обусловлено спонтанными переходами электронов с высоких уровней энергии на более низкие; инфракрасное излучение происходит за счет изменения колебательного и вращательного движений атомов; при люминесцентном излучении атомы и электроны спонтанно переходят с высоких уровней энергии на более низкие, а возбуждаются электромагнитным полем. В газоразрядном источнике излучение возникает в результате электрического разряда в атмосфере инертных газов, паров металла или их смесей. Принцип действия излучающих полупроводниковых диодов (светодиодов) основан на явлении электролюминесценции при протекании тока в структурах с *p-n*-переходом. Устройство, генерирующее когерентные электромагнитные волны за счет вынужденного испускания или вынужденного рассеяния света активной средой, находящейся в оптическом резонаторе, называется лазером. Слово «лазер» представляет собой аббревиатуру слов английского выражения: «Light Amplification by Stimulated Emission of Radiation» – усиление света вынужденным излучением.

Непрерывный спектр электромагнитного излучения распространяется от γ -лучей с минимальной длиной волны 10^{-7} мкм, возникающих при распаде радиоактивных элементов, до длинноволновых радиоизлучений и излучений генераторов переменного тока промышленной частоты с длиной волны 6000 км. Средняя область спектра электромагнитного излучения, охватывающая инфракрасное излучение с длиной волн от 1 мм до 0,78 мкм, видимое излучение – от 0,78 мкм до 0,38 мкм и ультрафиолетовое излучение с длинами волн от 0,38 мкм до 0,01 мкм, носит название оптической области спектра. Излучение этой области спектра называется оптическим излучением. Объединение излучений этих спектральных областей в одну группу объясняется как единством принципов возбуждения оптического излучения, так и общностью методов их преобразования и использования.

Полная энергия излучения любого спектрального состава, переносимая электромагнитными волнами за единицу времени, значительно превышающую периоды световых колебаний, через какую-либо поверхность, определяет мощность излучения, которую в оптике

принято называть потоком излучения или лучистым потоком. Количественной характеристикой излучения является его мощность (поток излучения) Φ_e , а качественной характеристикой – его спектральный состав, при этом $\Phi_e = \Phi_e(\lambda)$, где λ – длина волны монохроматического излучения. Спектральный состав излучения определяется спектральной плотностью потока излучения, равной

$$\varphi_{e\lambda} = \frac{d\Phi_{e\lambda}}{d\lambda}.$$

Спектральная плотность потока излучения $\varphi_{e\lambda}$ характеризует распределение энергии в спектре. При заданной (или известной) зависимости $\varphi_{e\lambda}$ от λ полный поток излучения в области спектра от λ_1 до λ_2 определяется очевидным выражением вида:

$$\Phi_e = \int_{\lambda_1}^{\lambda_2} \varphi_{e\lambda} d\lambda. \quad (5.5)$$

Спектр излучения называют сплошным, если спектральная плотность потока излучения $\varphi_{e\lambda}$ – непрерывная функция λ , отличная от нуля в широком интервале длин волн. Сплошной спектр имеет, например, свет, излучаемый раскаленными твердыми телами и жидкостями.

Спектр излучения называют линейчатым, если $\varphi_{e\lambda}$ практически отлична от нуля лишь в узких дискретных интервалах длин волн $\lambda_i \pm \frac{1}{2} \Delta\lambda_i$ ($\Delta\lambda_i \ll \lambda_i$), каждому из которых соответствует своя спектральная линия. Линейчатый спектр имеют, например, газоразрядные источники излучения. Полный поток сложного излучения, имеющего линейчатый спектр, определяется суммированием:

$$\Phi_e = \Phi_{e\lambda_1} + \Phi_{e\lambda_2} + \dots + \Phi_{e\lambda_n} = \sum_{k=1}^{k=n} \Phi_{e\lambda_k}.$$

Спектр излучения называют полосатым, если соответствующие ему спектральные линии образуют дискретные группы – полосы, которые состоят из множества тесно расположенных линий. Непрерывную полосу имеет, например, спектр излучения люминофора в люминесцентном источнике излучения.

Для обнаружения излучения и количественной оценки (измерения) потока излучения применяют приемники излучения (приемники лучистой энергии), которые, как правило, неодинаково реагируют на потоки излучения различного спектрального состава. Такие приемники называют селективными приемниками. Следовательно, реакция селективного приемника при воздействии на него потока излучения будет зависеть не только от величины последнего, но и от спектрального состава излучения.

В общем виде этой зависимости можно придать следующую математическую форму:

$$dX_{\lambda} = S_{\text{абс}}(\lambda)d\Phi_{e\lambda}, \quad (5.6)$$

где $d\Phi_{e\lambda}$ – поток излучения в пределах малого спектрального диапазона $d\Phi_{e\lambda} = \varphi_{e\lambda}d\lambda$; dX_{λ} – некоторая величина, характеризующая изменение реакции приемника; $S_{\text{абс}}(\lambda)$ – абсолютная спектральная характеристика чувствительности приемника излучения.

Спектральная характеристика приемника излучения определяет спектральный диапазон его работы.

Чувствительностью приемника излучения называют отношение изменения измеряемой реакции приемника (например, тока или напряжения), вызванного падающим на приемник излучением, к количественной характеристике этого излучения в заданных эксплуатационных условиях. Абсолютной спектральной характеристикой чувствительности приемника излучения называют зависимость монохроматической чувствительности $S_{\text{абс}}(\lambda)$, измеренной в абсолютных единицах, от длины волны падающего на приемник потока излучения.

Если приемником излучения является глаз человека, то величиной dX_{λ} принято считать световой поток $d\Phi_{\lambda}$, пропорциональный потоку излучения $d\Phi_{e\lambda}$ и соответствующий вызываемому им зрительному ощущению в глазу, при этом коэффициент пропорциональности принято называть коэффициентом видности K_{λ} . В этом случае формула (5.6) имеет вид:

$$d\Phi_{\lambda} = K_{\lambda}d\Phi_{e\lambda}. \quad (5.7)$$

В большинстве случаев спектральные характеристики приемников излучения имеют вид плавных кривых с одним максимумом при $\lambda = \lambda_{\text{max}}$, при этом абсолютная спектральная характеристика однотипных приемников изменяется от приемника к приемнику. Однако, относительная спектральная характеристика чувствительности приемника

излучения $S(\lambda) = \frac{S_{\text{абс}}(\lambda)}{S_{\lambda_{\text{max}}}}$, определяющая зависимость его

монохроматической чувствительности, отнесенной к значению максимальной чувствительности, от длины волны регистрируемого излучения, остается практически неизменной для приемников излучения одного и того же типа, но зависит от материала чувствительного слоя приемника и для известных применяемых материалов приводится в справочниках. Зная из справочника $S(\lambda)$ и измерив $S_{\lambda_{\text{max}}}$, можно легко определить $S_{\text{абс}}(\lambda)$. Вполне очевидно, что относительная спектральная характеристика чувствительности приемника излучения $S(\lambda)$ может

изменяться в пределах $0 \leq S(\lambda) \leq 1$. Учитывая изложенное, выражение (5.6) можно представить в виде:

$$dX_\lambda = S_{\text{abc}}(\lambda)\varphi_\lambda d\lambda = S_{\lambda_{\text{max}}}S(\lambda)\varphi_\lambda d\lambda. \quad (5.8)$$

Чтобы получить полную реакцию приемника от всего потока излучения, необходимо проинтегрировать выражение (5.8):

$$X = S_{\lambda_{\text{max}}} \int_0^\infty S(\lambda)\varphi_\lambda d\lambda = S_{\lambda_{\text{max}}} \Phi_{\text{эф}}.$$

Таким образом, эффективным потоком излучения $\Phi_{\text{эф}}$ источника для данного приемника излучения называют поток, который при чувствительности приемника, постоянной по всему спектру и равной максимальному значению $S_{\lambda_{\text{max}}}$, вызвал бы такую же реакцию приемника, какую вызывает весь падающий реальный поток при реальной спектральной чувствительности. Интегральный поток от источника равен

$$\Phi_e = \int_0^\infty \varphi_{e\lambda} d\lambda.$$

Следовательно, можно определить интегральную чувствительность

$$S = \frac{X}{\Phi_e} = S_{\lambda_{\text{max}}} \frac{\int_0^\infty S(\lambda)\varphi_{e\lambda} d\lambda}{\int_0^\infty \varphi_{e\lambda} d\lambda} = S_{\lambda_{\text{max}}}\chi. \quad (5.9)$$

Отношение интегралов в полученном выражении показывает, какую долю в сложном потоке, падающем на приемник излучения, составляет эффективный поток $\Phi_{\text{эф}}$ для данных приемника излучения и источника. Это отношение обозначают буквой χ и называют коэффициентом использования потока излучения данным приемником излучения или спектральным коэффициентом полезного действия (КПД) приемника.

Если приемником излучения является глаз человека, то относительную спектральную характеристику глаза называют функцией (или коэффициентом) относительной спектральной видности (или чувствительности глаза) V_λ , определяемой отношением

$$V_\lambda = \frac{K_\lambda}{K_{\lambda_{\text{max}}}}. \quad (5.10)$$

Как показали многочисленные экспериментальные исследования, коэффициент видности K_λ (при дневном зрении) достигает своего максимального значения $K_{\lambda_{\text{max}}}$ при длине волны излучения $\lambda_{\text{max}} = 0,555 \text{ мкм}$, при этом при $\lambda = \lambda_{\text{max}}$ функция относительной видности V_λ достигает своего максимального значения, равного единице. Из выражения (5.10) следует, что $0 \leq V_\lambda \leq 1$.

Итак, как следует из формул (5.7) и (5.10), монохроматическому потоку излучения $d\Phi_{e\lambda}$ длины волны λ соответствует световой поток, равный

$$d\Phi_{\lambda} = K_{\lambda} d\Phi_{e\lambda} = K_{\lambda_{\max}} V_{\lambda} d\Phi_{e\lambda}. \quad (5.11)$$

При сплошном спектре излучения распределение светового потока определяется спектральной плотностью светового потока

$$f_{\lambda} = \frac{d\Phi_{\lambda}}{d\lambda} = K_{\lambda} \frac{d\Phi_{e\lambda}}{d\lambda} = K_{\lambda} \varphi_{e\lambda},$$

где $\varphi_{e\lambda}$ – спектральная плотность потока излучения.

Для светового потока в пределах малого спектрального диапазона (для элементарного потока) имеем

$$d\Phi_{\lambda} = f_{\lambda} d\lambda = K_{\lambda} \varphi_{e\lambda} d\lambda. \quad (5.12)$$

При этом световой поток сложного излучения со сплошным спектром равен

$$\Phi = \int_{\lambda_1}^{\lambda_2} K_{\lambda} \varphi_{e\lambda} d\lambda = K_{\lambda_{\max}} \int_{\lambda_1}^{\lambda_2} V_{\lambda} \varphi_{e\lambda} d\lambda. \quad (5.13)$$

Здесь λ_1 и λ_2 – длины волн, определяющие границы спектральной чувствительности глаза. За пределами видимой области спектра подынтегральная функция равна нулю.

5.3. Фотометрические величины и единицы их измерения

5.3.1. Сила света

На некоторой ступени развития культуры и науки человек, получивший в свое распоряжение разнообразные источники искусственного света, начал интересоваться возможностью их сопоставительной количественной оценки. Одна из первых задач в области световых измерений состояла в том, чтобы определить какой из двух источников светит сильнее и во сколько раз.

В 1604 году Иоганн Кеплер сформулировал один из главных законов фотометрии, согласно которому освещение, производимое источником света, изменяется обратно пропорционально квадрату расстояния до него. Если два источника освещают поверхность, расположенную на расстоянии r_1 и r_2 от каждого из них, то при одинаковом освещении каждым из них в соответствии с законом, сформулированным И. Кеплером, имеем

$$\frac{I_1}{r_1^2} = \frac{I_2}{r_2^2}, \quad (5.14)$$

где I_1 и I_2 – величины, характеризующие свечение сравниваемых источников в направлении на освещаемую поверхность. Эти величины получили название сил света источников. Выражение (5.14) позволяет

рассчитать во сколько раз сила света одного источника больше силы света другого.

Единица силы света, которая всегда была исходной для практики, воспроизводилась в разное время по-разному. Первыми эталонами силы света были свечи. Сохранились сведения об английской спермацетовой свече и о немецких парафиновых и стеариновых свечах, которые изготавливались особо тщательно.

Во Франции около 1800 года свеча была заменена лампой Карселя, к фитилю которой подводилось сурепное масло. В Англии в 1887 году была предложена пентановая лампа, в бесфитильной горелке которой сгорала смесь паров пентана и подогретого воздуха. Из эталонов, питаемых жидким топливом, наиболее практичной оказалась лампа Гефнера, в тщательно стандартизованный корпус которой наливался химически чистый амилацетат. Лампа Гефнера служила эталоном в Германии, Австрии, Швейцарии и некоторых других странах до 1948 года.

Значительно более удобными в работе были эталоны, использующие свечение твердых тел. В 1879 году француз Виоль предложил считать единицей ту силу света, которую дает в направлении, перпендикулярном к поверхности, квадратный сантиметр чистой платины при температуре ее затвердевания. Эта единица (единица Виоля) была примерно в 20 раз больше пентановой свечи. Однако, этот эталон оказался нестабильным и от него вскоре пришлось отказаться.

В 1909 году метрологические учреждения Англии, Франции и США согласились поддерживать некоторую новую единицу силы света, близкую к пентановой свече, с помощью группы тщательно изготовленных и постоянно контролируемых ламп накаливания. Первоначально групповой эталон состоял из угольных ламп, а в дальнейшем были созданы аналогичные группы из вольфрамовых ламп. Зафиксированной таким образом единице силы света было присвоено название «международная свеча». С 1924 года эта единица силы света была принята в СССР. Ею пользовались до 1948 года, когда она повсеместно была заменена канделой (кд) (от английского candle – свеча).

Еще в 1947 году канделу (в то время ее называли новой свечой) было принято воспроизводить с помощью эталона, использующего свечение абсолютно черного тела, яркость которого при температуре затвердевания чистой платины было решено считать равной $60 \text{ кд} \cdot \text{см}^{-2}$. С 1 января 1948 года эта единица была повсеместно принята за основу измерения всех фотометрических величин. Важно отметить, что единица силы света является одной из семи основных единиц Международной системы – СИ.

Сила света источника в общем случае зависит от направления излучения и в разных направлениях бывает весьма различной. Если источник имеет форму цилиндра (например, прямая газосветная лампа), то его наибольшая сила света направлена перпендикулярно к оси цилиндра, а

наименьшая – вдоль оси. Нередко можно встретить графическое изображение распределения силы света источника в виде кривой в полярной системе координат, радиусы-векторы которой пропорциональны силам света источника в соответствующих направлениях. Такая кривая дает наглядное представление о распределении света, излучаемого источником или светильником, т.е. источником, помещенным в некоторую арматуру.

5.3.2. Световой поток

Предположим, что источник dS_u излучает свет в окружающее пространство в пределах телесного угла $d\omega_u$, вершина которого совпадает с источником dS_u . Будем считать, что в некотором направлении сила света (в канделах) в пределах телесного угла $d\omega_u$ равна dI . С центром в вершине телесного угла построим сферы радиусов r_1, r_2 и r_3 , как показано на рис. 5.3, и обозначим через dS_1, dS_2 и dS_3 элементарные площадки, вырезанные конусом телесного угла $d\omega_u$ на соответствующих сферах. Согласно закону обратной пропорциональности квадрату расстояния освещенность площадок пропорциональна отношениям $\frac{dI}{r_1^2}, \frac{dI}{r_2^2}, \frac{dI}{r_3^2}$. За

единицу времени на площадки dS_1, dS_2 и dS_3 упадет одна и та же энергия, которая переносится рассматриваемым пучком лучей в пределах угла $d\omega_u$.

Составив произведения $\frac{dI}{r_1^2} dS_1, \frac{dI}{r_2^2} dS_2$ и $\frac{dI}{r_3^2} dS_3$, замечаем, что каждое из

них равно $dI d\omega_u$, так как $\frac{dS_1}{r_1^2} = \frac{dS_2}{r_2^2} = \frac{dS_3}{r_3^2} = d\omega_u$.

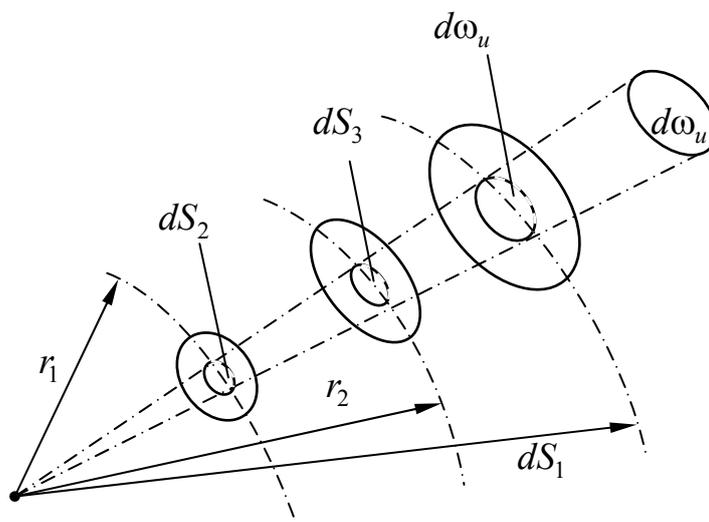


Рис. 5.3. Световой поток источника dS_u внутри телесного угла $d\omega_u$

Величина

$$d^2\Phi = dId\omega_u, \quad (5.15)$$

равная световой мощности, переносимой пучком лучей, распространяющимся внутри телесного угла $d\omega_u$, получила название светового потока.

Световой поток, определяемый соотношением (5.15), представляет собой ту часть полного светового потока, излучаемого источником dS_u , которая приходится на телесный угол $d\omega_u$. Если сила света источника зависит от направления распространения излучения, то полный световой поток $d\Phi$, излучаемый источником dS_u в окружающее пространство, равен

$$d\Phi = \int_{\omega_u} dId\omega_u, \quad (5.16)$$

где $\omega_u \leq 4\pi$.

Единица светового потока – люмен (лм) (от латинского lumen – свет, лампа), при этом световой поток, распространяющийся внутри телесного угла в 1 ср при силе света источника, расположенного в его вершине, равной 1 кд, равен 1 лм, т.е. лм=кд·ср. В светотехнической практике часто пользуются представлением о средней сферической силе света I_θ источника излучения как о силе света такого условного источника S_u , который, излучая одинаково во все стороны, излучает такой же световой поток, что и реальный источник. Вполне очевидно, что полный световой поток Φ источника связан с его средней сферической силой света I_θ простым соотношением:

$$\Phi = 4\pi I_\theta.$$

При определении понятий силы света и светового потока предполагалось, что источник света расположен в вершине телесного угла, т.е. в точке. Однако, источник излучения точкой физически быть не может. Следует обратить внимание на то обстоятельство, что понятие силы света источника непосредственно связано с направлением излучения, при этом лучи, исходящие из разных точек источника и падающие на освещаемый предмет, должны располагаться внутри малого телесного угла. В равной мере и расстояние от источника до освещаемой поверхности приобретает определенность только тогда, когда размер источника мал по сравнению с расстоянием, на котором определяется создаваемое им освещение. Прекрасным примером точечного источника могут служить звезды, каждая из которых имеет огромные размеры. Однако, расстояние до звезд столь велико, что их размеры не мешают точному определению направления на них. Таким образом, чем больше расстояние, на котором наблюдается световое действие источника, тем с большим основанием можно считать его точечным. Так, например, если источником излучения служит равномерно светящийся диск,

то погрешность определения силы света не превышает 1 %, если расстояние до источника превосходит его диаметр всего лишь в пять раз.

5.3.3. Освещенность. Светимость

Величина

$$dE_n = \frac{dI}{r^2}, \quad (5.17)$$

которая определяет величину освещения на поверхности, расположенной на расстоянии r от источника излучения dS_u с силой света dI , при нормальном падении лучей, как показано на рис. 5.4, называется освещенностью поверхности. За единицу освещенности принимается такая освещенность, которую создает источник излучения силой света в 1 кд на поверхности, расположенной нормально к направлению излучения на расстоянии 1 м от источника. Единица освещенности – люкс (лк) (от латинского lux – свет солнца, освещение).

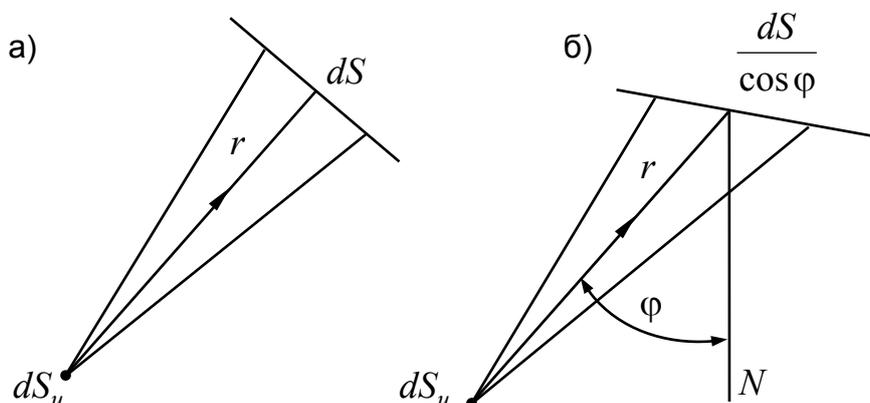


Рис. 5.4. Лучи источника излучения dS_u , падающие на поверхность dS по нормали к ней (а) и под углом φ к нормали (б)

Умножим числитель и знаменатель отношения (5.17) на бесконечно малый телесный угол $d\omega_u$, в котором сила света источника dS_u равна dI . В результате получаем

$$dE_n = \frac{dI d\omega_u}{r^2 d\omega_u} = \frac{d^2\Phi}{dS_0},$$

где $d^2\Phi = dI d\omega_u$ – световой поток, распространяющийся от источника dS_u в телесный угол $d\omega_u$, а $dS_0 = r^2 d\omega_u$ – площадь элемента поверхности, нормальной к направлению излучения, на который падает этот поток. Отсюда следует, что освещенность поверхности представляет собой поверхностную плотность падающего на нее светового потока. Таким образом, единица освещенности люкс определяет такую поверхностную плотность светового потока, при которой световой поток в 1 лм покрывает поверхность площадью, равной 1 м^2 . Если лучи от источника падают на

поверхность под углом φ к нормали N к поверхности, как показано на рис. 5.4, то тот же световой поток $d^2\Phi$ падает на площадку $dS = \frac{dS_0}{\cos\varphi}$. В

этом случае

$$dE_\varphi = \frac{dI}{r^2} \cos\varphi = dE_n \cos\varphi. \quad (5.18)$$

Каждый элемент dS_m излучающей или освещенной поверхности излучает или отражает в окружающее пространство (точнее, в полупространство, измеряемое телесным углом 2π) некоторый световой поток $d^2\Phi_{\text{изл}}$. Поверхностная плотность излучаемого или отраженного светового потока

$$dM = \frac{d^2\Phi_{\text{изл}}}{dS_m} \quad (5.19)$$

называется светимостью. Единица светимости – люмен на квадратный метр ($\text{лм}\cdot\text{м}^{-2}$).

5.3.4. Яркость

Предположим, что бесконечно малый элемент dS_u излучающей поверхности освещает бесконечно малый элемент dS некоторой поверхности, расположенной на расстоянии r от dS_u . Будем считать, что из каждой точки элемента dS_u лучи проходят через принятую в качестве осевой точку элемента dS , образуя телесный угол $d\omega$. При этом световой поток, падающий на поверхность элемента dS , в соответствии с выражением (5.4) равен

$$d^2\Phi = L d\omega dS_0 = L d\omega dS \cos\varphi,$$

где φ – угол между осью телесного угла $d\omega$ и нормалью к элементу dS . Коэффициент пропорциональности L называется яркостью излучающего элемента dS_u в направлении к освещаемому элементу dS . Телесный угол в этом уравнении

$$d\omega = \frac{dS_u}{r^2} \cos\varphi_u,$$

где φ_u – угол между осью телесного угла $d\omega$ и нормалью к элементу dS_u . При этом

$$d^2\Phi = L \frac{dS_u \cos\varphi_u dS \cos\varphi}{r^2}. \quad (5.20)$$

Заметим, что отношение $\frac{dS \cos\varphi}{r^2}$ определяет телесный угол $d\omega_u$ с вершиной в любой точке элемента dS_u .

В соответствии с выражением (5.15) световой поток $d^2\Phi = dId\omega_u$. При этом из выражения (5.20) следует, что

$$L = \frac{d^2\Phi}{dS_u \cos \varphi_u d\omega_u} = \frac{dI}{dS_u \cos \varphi_u}. \quad (5.21)$$

Выражение (5.21) определяет яркость поверхности S_u в любой ее точке и в любом направлении (при любом значении угла φ_u) отношением силы света элемента поверхности в выбранном направлении к площади проекции этого элемента на плоскость, перпендикулярную направлению силы света. Отсюда следует, что за единицу яркости следует принять яркость такой поверхности, для которой сила света, выраженная в канделах, равна выраженной в квадратных метрах площади ее проекции на плоскость, перпендикулярную к направлению силы света. Таким образом, единица яркости – кандела на квадратный метр ($\text{кд}\cdot\text{м}^{-2}$).

Выражение (5.20) можно представить в виде:

$$d^2\Phi = LdS_0 \cos \varphi_u d\omega_u = LdS \cos \varphi d\omega. \quad (5.22)$$

В этом выражении в среднюю часть входят величины, относящиеся только к площадке излучателя dS_u , а в правую часть – те же величины, но относящиеся к площадке dS . Если площадку dS перемещать в направлении пучка лучей, сохраняя ее величину и угол падения φ , то как телесный угол $d\omega_u$, так и телесный угол $d\omega$, а, соответственно и световой поток $d^2\Phi$, будут изменяться пропорционально изменению обратной величины квадрата расстояния между площадками. Следовательно, яркость L должна оставаться постоянной в направлении рассматриваемого пучка лучей и равной яркости элемента dS_u в этом направлении. Это справедливо только в тех случаях, когда среда, через которую распространяется пучок лучей, не поглощает и не рассеивает свет.

Изложенные соображения позволяют распространять понятие яркости поверхности на пучок световых лучей, при этом представление о яркости светового пучка следует считать более общим. Понятие яркости пучка лучей, введенное в трудах академика Владимира Александровича Фока, Андрея Александровича Гершуна, Сергея Осиповича Майзеля и др., оказалось весьма удобным при исследовании объемных источников или рассеивающих излучение сред, например, безоблачного неба, где невозможно указать поверхность, с которой можно было бы связать определение яркости, т.е. в подобных случаях понятие яркости поверхности вообще теряет смысл. Единственным объектом, которому можно приписать, например, яркость безоблачного неба, является сам пучок световых лучей, пришедший к нам с того или иного направления.

Обратимся вновь к выражению (5.22), из которого находим, что

$$L = \frac{d^2\Phi}{dS \cos \varphi d\omega} = \frac{dE_\varphi}{\cos \varphi d\omega} = \frac{dE_n}{d\omega},$$

т.е. яркость элементарного пучка лучей пропорциональна освещенности, которую он создает на перпендикулярной к нему поверхности и обратно пропорциональна телесному углу, который он заполняет.

Из выражения (5.21) следует, что

$$dI = LdS_u \cos \varphi_u. \quad (5.23)$$

Для элемента источника, излучающего по закону Ламберта, яркость излучающей поверхности одинакова по всем направлениям. Если для всех направлений (т.е. для всех значений угла φ_u) $L(\varphi_u) = L = const$, то

$$dI = dI_0 \cos \varphi_u,$$

где $dI_0 = LdS_u$ – сила света излучающего элемента dS_u в нормальном к нему направлении (т.е. при $\varphi_u = 0$). Если для всех точек светящейся поверхности S_u справедлив закон Ламберта, то $I = I_0 \cos \varphi_u$.

Если поверхность dS_u излучает равномерно во всех направлениях, то полный световой поток, излучаемый элементом dS_u по одну сторону от себя, равен

$$d\Phi = LdS_u \int_0^{2\pi} \cos \varphi_u d\omega_u.$$

Телесный угол $\omega_u = 2\pi(1 - \cos \varphi_u)$, и, соответственно, $d\omega_u = 2\pi \sin \varphi_u d\varphi_u$. Тогда

$$d\Phi = \pi LdS_u \int_{\varphi_u=0}^{\varphi_u=\frac{\pi}{2}} 2 \sin \varphi_u \cos \varphi_u d\varphi_u = \pi LdS_u.$$

При этом светимость поверхности

$$M = \frac{d\Phi}{dS_u} = \pi L. \quad (5.24)$$

Пусть освещенность поверхности равна E . Светимость идеально рассеивающей поверхности $M = E$, а яркость ее в соответствии с соотношением (5.24) равна $L = \frac{E}{\pi}$. Значение идеального рассеивателя

состоит в том, что с его предельными свойствами удобно сравнивать свойства всех реальных материалов.

5.3.5. Световая энергия. Экспозиция

Классическая фотометрия, основы которой были заложены трудами П. Бугера и И. Ламберта в XVIII веке, предполагает применение непрерывно излучающих источников, отвечающих требованиям стабильности излучения в течение длительного времени, которое само по

себе значения не имело. Однако, уже давно стали выявляться случаи, в которых длительность свечения играла существенную роль. Примером такого случая может служить применение импульсных источников света, которые с некоторых пор начали приобретать все более широкое распространение. Длительность свечения импульсных источников излучения измеряется иногда тысячными и даже миллионными долями секунды, что очень существенно, например, для фотографии. Излучение импульсного источника обладает переменной мощностью, которая, возрастая от нуля, достигает за малые доли секунды своего максимального значения, а затем падает до нуля. Поэтому излучение импульсного источника характеризуют или длительностью импульса и максимальной мощностью излучения или энергией импульса, которая определяется интегралом от мощности по времени излучения.

Одним из видов навигационного ограждения пути корабля в море, осуществляемому с помощью плавающих буев или бакенов, являются проблесковые огни. В темное время суток эти огни, вспыхивающие периодически на десятые доли секунды, позволяют мореплавателю определять свое местоположение. Темновые паузы имеют значительно большую длительность. Инерционные свойства глаза заставляют принимать во внимание не только силу света проблескового огня, но и длительность свечения.

При фотометрических процессах (фотография, фотосинтез) результат также зависит не только от освещенности светочувствительного слоя, но и от времени освещения.

В связи с тем, что число таких случаев непрерывно растет и они приобретают все большее значение, оказалось необходимым пополнить фотометрическую систему рядом новых величин, из которых в первую очередь надо отметить следующие:

1. Световая энергия Q в соответствии с определением равна

$$Q = \int_0^{\tau} \Phi dt \quad (5.25)$$

и измеряется в люмен-секундах (лм·с).

2. Экспозиция (или количество освещения) H определяется интегралом

$$H = \int_0^{\tau} E dt \quad (5.26)$$

и измеряется в люкс-секундах (лк·с).

Но $\Phi = \int_S E dS$. Тогда $dQ = \int_{\tau} E dS dt = HdS$. При этом

$$H = \frac{dQ}{dS}. \quad (5.27)$$

Из последнего выражения следует, что экспозиция H представляет собой поверхностную плотность световой энергии.

5.4. Энергетические величины

Итак, распределение светового потока по телесному углу определяется силой света, по поверхности – освещенностью или светимостью, а одновременно по телесному углу и по поверхности – яркостью. Представление о световом потоке и обо всех связанных с ним величинах формировалось в процессе развития визуальной фотометрии. Существенными факторами, определяющими формирование этих представлений, во-первых, было то, что невидимые излучения еще не были известны, и, во-вторых, то, что глаз человека был единственным средством обнаружения и оценки излучения любого источника. Таким образом, световой поток оказывался естественной мерой всякого излучения.

5.4.1. Связь между световыми и энергетическими величинами

К видимому свету, свойства которого были изучены еще И. Ньютоном, добавились открытые в начале XIX века инфракрасное и ультрафиолетовое излучения. Позже была установлена их общая электромагнитная природа и общая скорость распространения в вакууме. Появились разнообразные приемники излучения, чувствительные к излучению различного спектра, а соответственно расширился и спектральный диапазон применяемого излучения. Одновременно с этим во все большей степени ощущалась необходимость оценки общей мощности переносимого электромагнитного излучения независимо от его спектрального состава, которая получила название потока излучения (или лучистого потока). Поток излучения принято обозначать буквой P или Φ_e и, как всякую мощность, измерять в ваттах. Как естественное следствие того, что поток излучения становится основой радиационной энергетики, возникла потребность во всех видах производных величин, применяемых для характеристики его пространственного распределения, т.е. появилась необходимость в системе энергетических величин, построенных аналогично системе световых (визуальных) величин. Отличие этих новых величин от прежних состоит только в том, что вместо светового потока в их определении входит поток излучения.

Наиболее естественным и удобным способом наименования этих новых энергетических величин принят такой, при котором за каждой энергетической величиной сохраняется название ее светового аналога с добавлением слова «энергетический». Что касается обозначения, то за каждой энергетической величиной сохраняется обозначение, используемое для аналогичного светового понятия, с добавлением внизу индекса e . Таким образом, плотность потока излучения по телесному углу получает название энергетической силы света или силы излучения источника и

определяется соотношением $I_e = \frac{d\Phi_e}{d\omega}$, где $d\Phi_e$ – бесконечно малый поток излучения, распространяющийся от источника внутри бесконечно малого телесного угла $d\omega$. Энергетическая сила света измеряется в ваттах на стерадиан ($\text{Вт}\cdot\text{ср}^{-1}$). Поверхностная плотность потока излучения имеет следующие названия: энергетическая освещенность, равная $E_e = \frac{d\Phi_{e\text{пад}}}{dS}$, и

энергетическая светимость, равная $M_e = \frac{d\Phi_{e\text{изл}}}{dS_u}$, где $d\Phi_{e\text{пад}}$ и $d\Phi_{e\text{изл}}$ –

бесконечно малые потоки, падающий на элемент поверхности и излучаемый элементом поверхности соответственно. Эти величины измеряют в ваттах на квадратный метр ($\text{Вт}\cdot\text{м}^{-2}$). Аналогичным образом вводится и понятие об энергетической яркости как о поверхностной плотности энергетической силы света, равной

$$L_e = \frac{dI_e}{dS_u \cos \varphi},$$

т.е. об энергетической силе света, отнесенной к площади проекции элемента поверхности на плоскость, перпендикулярную к направлению энергетической силы света. Эта величина измеряется в ваттах на стерадиан и на квадратный метр ($\text{Вт}\cdot\text{ср}^{-1}\cdot\text{м}^{-2}$).

5.4.2. Световая эффективность

При сплошном спектре полный поток излучения в соответствии с формулой (5.5) равен

$$\Phi_e = \int_0^{\infty} d\Phi_{e\lambda} = \int_0^{\infty} \varphi_{e\lambda} d\lambda. \quad (5.28)$$

Если поток излучения $d\Phi_{e\lambda}$ приходится на видимую часть спектра, то при этом возможно световое воздействие излучения на глаз и ощущение некоторого светового потока $d\Phi_\lambda$, в соответствии с выражением (5.11) равного

$$d\Phi_\lambda = K_{\lambda_{\max}} V_\lambda \varphi_{e\lambda} d\lambda, \quad (5.29)$$

где $K_{\lambda_{\max}}$ – коэффициент пропорциональности, который называют максимальной световой эффективностью излучения. Если мощность измеряется в ваттах, а световой поток – в люменах, то коэффициент $K_{\lambda_{\max}}$ равен числу люменов светового потока, вызванного потоком излучения в один ватт при той длине волны излучения, при которой $V_\lambda = 1$, т.е. при $\lambda_0 = 0,555$ мкм.

Развитие фотометрии как теоретической, так и экспериментальной, позволило Международному комитету мер и весов принять в 1977 году,

что 1 Вт излучения частоты $\nu_0 = 540 \cdot 10^{12}$ Гц ($\lambda_0 = \frac{c}{\nu_0} = 0,555$ мкм),

обладающего наибольшей световой эффективностью при дневном освещении, эквивалентен 683 лм светового потока.

Световой эффективностью излучения называют отношение светового потока к потоку вызвавшего его излучения. Вполне очевидно, что световая эффективность излучения зависят от спектрального состава излучения и равна нулю в тех случаях, когда излучение не содержит потока в видимой части спектра. Световая эффективность солнечного излучения близка к $100 \text{ лм} \cdot \text{Вт}^{-1}$, а для монохроматических потоков световая эффективность $K_\lambda = V_\lambda K_m$, где $K_m = 683 \text{ лм} \cdot \text{Вт}^{-1}$, а V_λ – относительная спектральная чувствительность глаза.

Световой поток источника излучения сложного спектрального состава может быть определен как сумма элементарных световых потоков всех его монохроматических составляющих:

$$\Phi = K_m \int_{\lambda} V_\lambda d\Phi_{e\lambda} = K_m \int_{\lambda} V_\lambda \varphi_{e\lambda} d\lambda. \quad (5.30)$$

Это выражение определяет понятие светового потока, откуда следует, что световой поток является величиной, пропорциональной потоку излучения, оцененному в соответствии с относительной спектральной чувствительностью среднего глаза.

Иногда световой поток выражают в ваттах, условно называемых световыми ваттами, число которых считается равным интегралу $\int_{\lambda} V_\lambda d\Phi_{e\lambda}$,

если поток излучения $d\Phi_{e\lambda} = \varphi_{e\lambda} d\lambda$ выражен в ваттах. Важно обратить внимание на то, что один световой ватт эквивалентен 683 лм независимо от спектрального состава излучения.

Отношение светового потока в световых ваттах к потоку излучения в ваттах представляет собой безразмерную величину, равную

$$\eta = \frac{\int_{\lambda} V_\lambda \varphi_{e\lambda} d\lambda}{\int_{\lambda} \varphi_{e\lambda} d\lambda} \quad (5.31)$$

которую называют световым коэффициентом полезного действия (КПД) излучения. Величина этого коэффициента может достигать единицы, если речь идет о монохроматическом излучении с длиной волны $\lambda = 0,555$ мкм. Во всех других случаях световой КПД излучения меньше единицы и падает до нуля, если излучение не захватывает видимой области спектра.

Заметим, что произведение $K_m \eta = 683 \eta \text{ лм} \cdot \text{Вт}^{-1}$ определяет световую эффективность излучения, которая, естественно, зависит от его спектрального состава и может принимать значения от $683 \text{ лм} \cdot \text{Вт}^{-1}$ до нуля.

Следует различать световую эффективность излучения и световую эффективность источника, которая равна отношению светового потока к мощности, потребляемой им от источника энергии, например, от электрической сети. Световая эффективность источников света колеблется от $8\text{--}9 \text{ лм}\cdot\text{Вт}^{-1}$ для низкотемпературных вакуумных ламп накаливания (у вакуумных ламп давление остаточных газов в колбе после откачки $10^{-5}\text{--}10^{-7}$ мм рт.ст.) до $9\text{--}13,5 \text{ лм}\cdot\text{Вт}^{-1}$ для высокотемпературных газонаполненных ламп накаливания (у газонаполненных ламп после откачки колба заполняется смесью инертных газов, давление которых у холодной лампы приближенно равно 600 мм рт.ст.). Световая эффективность ламп накаливания с йодным циклом достигает $25\text{--}30 \text{ лм}\cdot\text{Вт}^{-1}$ при среднем сроке службы до 2000 ч.

Большая заслуга в создании электрических ламп накаливания принадлежит выдающемуся русскому ученому Александру Николаевичу Лодыгину (1847–1923), который разработал конструкцию ламп накаливания с угольной нитью в 1872 году и создал первые образцы ламп, пригодных для работы. Американский изобретатель и предприниматель Томас Алва Эдисон (1847–1931) усовершенствовал лампу накаливания, предложив в 1879 году конструкцию лампы, основные принципиальные элементы которой сохранились до настоящего времени.

Выдающаяся роль в деле создания газоразрядных источников света принадлежит русским электротехникам. Академик Петербургской АН Василий Владимирович Петров (1761–1834) в 1802 году первым получил устойчивую электрическую дугу в воздухе и указал на возможность ее использования в качестве источника света. Однако, лишь спустя 70 лет удалось осуществить на этом принципе источник, приемлемый для целей освещения. Заслуга эта принадлежит известному русскому изобретателю Петру Николаевичу Яблочкову (1847–1894), создавшему источник, нашедший применение как в России, так и за рубежом, и получивший название «свеча Яблочкова». Световая эффективность современных газоразрядных ламп равна: натриевых ламп – до $100 \text{ лм}\cdot\text{Вт}^{-1}$, люминесцентных ламп – до $70\text{--}80 \text{ лм}\cdot\text{Вт}^{-1}$, ртутных ламп высокого давления (ВД) и сверхвысокого давления (СВД) – до $60 \text{ лм}\cdot\text{Вт}^{-1}$, а газовых ламп ВД и СВД – до $45\text{--}50 \text{ лм}\cdot\text{Вт}^{-1}$. При значительно более высокой световой эффективности (световой отдаче), чем у ламп накаливания, газоразрядные лампы имеют и значительно больший срок службы. Если средний срок службы осветительной лампы накаливания равен 1000 ч. и не может быть увеличен без снижения и без того низкой световой эффективности, то срок службы современных люминесцентных ламп достигает 14000 ч.

Легко видеть, что равенства, подобные равенству (5.30), можно составить не только для светового потока, но и для любой другой

фотометрической величины. Например, для яркости, выраженной в канделах на квадратный метр, справедливо равенство:

$$L = K_m \int_{L_e} V_\lambda dL_e = K_m \int_{\lambda} V_\lambda L_{e\lambda} d\lambda, \quad (5.32)$$

где V_λ имеет прежнее значение, а $dL_e = L_{e\lambda} d\lambda$ – энергетическая яркость, приходящаяся на спектральный интервал $\lambda, \lambda + d\lambda$, Вт·ср⁻¹·м⁻². Заметим, что яркость излучения капиллярных ртутно-кварцевых ламп СВД с водяным охлаждением достигает $(1,5–1,8) \cdot 10^9$ кд·м⁻² при средней визуальной яркости Солнца $L = 1,95 \cdot 10^9$ кд·м⁻².

5.5. Фотометрические свойства тел

Фотометрические свойства окружающих нас предметов определяют яркость и цветность (окраску), которые они принимают под влиянием падающего излучения.

5.5.1. Общие соотношения

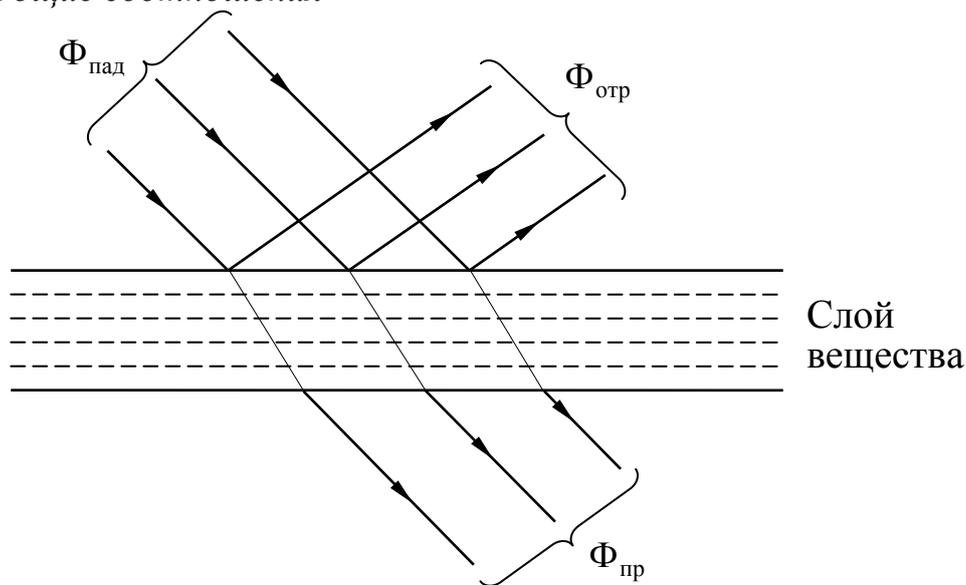


Рис. 5.5. Разделение светового потока, падающего на слой вещества

Для определенности будем считать, что освещенный предмет представляет собой плоский слой прозрачного вещества, как показано на рис. 5.5, толщина которого мала по сравнению с линейными размерами его освещаемой поверхности. Пусть на поверхность слоя из пространства, расположенного перед ним, падает произвольно распределенный по направлениям световой поток $\Phi_{\text{пад}}$. Освещенный слой разделит упавший на него световой поток на три части: поток $\Phi_{\text{отр}}$, отраженный от слоя в ту полусферу, из которой пришел падающий поток; поток $\Phi_{\text{пр}}$, прошедший через слой и вышедший в другую полусферу, и поглощенный поток $\Phi_{\text{погл}}$,

который в веществе слоя превратится в теплоту или другую форму энергии.

В соответствии с законом сохранения энергии сумма отраженного, прошедшего и поглощенного световых потоков равна падающему потоку:

$$\Phi_{\text{отр}} + \Phi_{\text{пр}} + \Phi_{\text{погл}} = \Phi_{\text{пад}}$$

или

$$\frac{\Phi_{\text{отр}}}{\Phi_{\text{пад}}} + \frac{\Phi_{\text{пр}}}{\Phi_{\text{пад}}} + \frac{\Phi_{\text{погл}}}{\Phi_{\text{пад}}} = 1. \quad (5.33)$$

Отношение светового потока, отраженного от слоя, к световому потоку, падающему на слой, называется коэффициентом отражения слоя и обозначается буквой ρ .

Отношение светового потока, прошедшего через слой, к световому потоку, падающему на слой, называется коэффициентом пропускания слоя и обозначается буквой τ .

Отношение светового потока, поглощенного в слое, к световому потоку, падающему на слой, называется коэффициентом поглощения слоя и обозначается буквой α .

При этом в соответствии с формулой (5.33) имеем

$$\rho + \tau + \alpha = 1. \quad (5.34)$$

Для окрашенных веществ эти коэффициенты зависят от спектрального состава падающего излучения. Для монохроматического излучения с длиной волны $\lambda \pm \frac{1}{2}\delta\lambda$ их принято обозначать $\rho(\lambda)$, $\tau(\lambda)$ и $\alpha(\lambda)$.

Спектральную зависимость коэффициентов $\rho(\lambda)$, $\tau(\lambda)$ и $\alpha(\lambda)$ удобно изображать графически, откладывая по оси абсцисс длину волны λ , а по оси ординат – один из коэффициентов. Пример такого графика показан на рис. 5.6.

Поток излучения любого спектрального состава всегда можно рассматривать как сумму большого числа монохроматических составляющих с мощностями $d\Phi_{e\lambda} = \varphi_{e\lambda} d\lambda$, каждая из которых, попав на слой, претерпевает изменения, зависящие от его спектральных свойств. При этом мощности отраженной, пропущенной и поглощенной монохроматической составляющей излучения могут быть представлены в виде произведений:

$$\rho(\lambda)\Phi_{e\lambda}, \tau(\lambda)\Phi_{e\lambda} \text{ и } \alpha(\lambda)\Phi_{e\lambda}.$$

Световой поток падающего излучения можно определить выражением (5.30) в виде:

$$\Phi_{\text{пад}} = K_m \int_{\lambda} V_{\lambda} d\Phi_{e\lambda} = K_m \int_{\lambda} V_{\lambda} \varphi_{e\lambda} d\lambda.$$

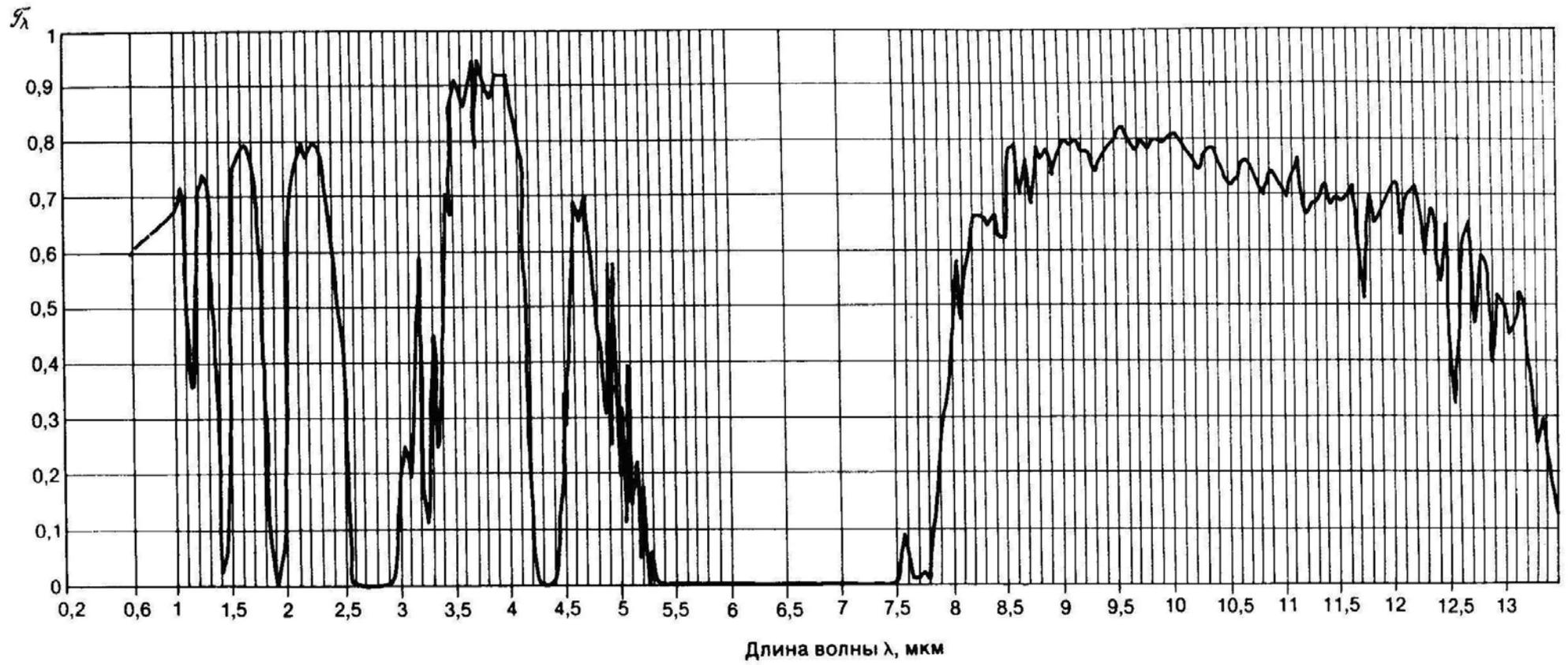


Рис. 5.6. Спектральный коэффициент пропускания атмосферы

Вполне очевидно, что световые потоки: отраженный от слоя, прошедший через него и поглощенный в нем, определяются подобными интегралами вида:

$$\Phi_{\text{отр}} = K_m \int_{\lambda} \rho(\lambda) V_{\lambda} \Phi_{e\lambda} d\lambda;$$

$$\Phi_{\text{пр}} = K_m \int_{\lambda} \tau(\lambda) V_{\lambda} \Phi_{e\lambda} d\lambda;$$

$$\Phi_{\text{погл}} = K_m \int_{\lambda} \alpha(\lambda) V_{\lambda} \Phi_{e\lambda} d\lambda.$$

В соответствии с этим визуальные коэффициенты отражения, пропускания и поглощения можно определить выражениями в такой форме:

$$\rho = \frac{\int_{\lambda} \rho(\lambda) V_{\lambda} \Phi_{e\lambda} d\lambda}{\int_{\lambda} V_{\lambda} \Phi_{e\lambda} d\lambda}, \quad (5.35)$$

$$\tau = \frac{\int_{\lambda} \tau(\lambda) V_{\lambda} \Phi_{e\lambda} d\lambda}{\int_{\lambda} V_{\lambda} \Phi_{e\lambda} d\lambda}, \quad (5.36)$$

$$\alpha = \frac{\int_{\lambda} \alpha(\lambda) V_{\lambda} \Phi_{e\lambda} d\lambda}{\int_{\lambda} V_{\lambda} \Phi_{e\lambda} d\lambda}, \quad (5.37)$$

где все интегралы берутся в пределах видимого спектра.

Если в выражениях (5.35), (5.36) и (5.37) формально положить множитель $V_{\lambda} \equiv 1$, то они примут вид выражений, определяющих энергетические коэффициенты отражения ρ_e , пропускания τ_e и поглощения α_e . Таким образом,

$$\rho_e = \frac{\int_{\lambda} \rho(\lambda) \Phi_{e\lambda} d\lambda}{\int_{\lambda} \Phi_{e\lambda} d\lambda}, \quad (5.38)$$

$$\tau_e = \frac{\int_{\lambda} \tau(\lambda) \Phi_{e\lambda} d\lambda}{\int_{\lambda} \Phi_{e\lambda} d\lambda}, \quad (5.39)$$

$$\alpha_e = \frac{\int_{\lambda} \alpha(\lambda) \Phi_{e\lambda} d\lambda}{\int_{\lambda} \Phi_{e\lambda} d\lambda}, \quad (5.40)$$

где все интегралы распространены на весь спектр излучения источника.

Отсюда следует, что энергетические коэффициенты отражения, пропускания и поглощения характеризуют фотометрические свойства слоя для нейтрального приемника. Действие всех известных нейтральных приемников излучения основано на нагревании, возникающем при поглощении падающего на них излучения. Поэтому все нейтральные приемники имеют черную поверхность, причем хороший приемник должен поглощать излучение в невидимых областях спектра такую же часть падающего на него излучения, как и в видимой области. Поверхность современного нейтрального приемника имеет высокий и практически одинаковый коэффициент поглощения в спектральном диапазоне излучения от 0,3 мкм до 40 мкм.

В соответствии с выражениями (5.35)–(5.40) величина коэффициентов отражения, пропускания и поглощения зависит не только от спектральных свойств материалов, определяемых функциями $\rho(\lambda)$, $\tau(\lambda)$ и $\alpha(\lambda)$, но также от вида функции $\varphi_{e\lambda}$. Влияние спектральной чувствительности глаза на коэффициенты ρ , τ и α подобно влиянию спектрального состава падающего излучения, при этом изменение спектральной чувствительности глаза, наступающее, например, при переходе к сумеречному зрению, влечет за собой изменение величины этих коэффициентов. Только для нейтральных материалов, спектральные характеристики которых в пределах рассматриваемого интервала длин волн не зависят от длины волны, коэффициенты отражения, пропускания и поглощения не зависят ни от спектрального состава падающего излучения, ни от спектральной чувствительности глаза и совпадают с их значениями для любого монохроматического излучения.

5.5.2. Гладкая поверхность раздела двух диэлектриков

Падающий на гладкую поверхность раздела двух диэлектриков луч делится на два луча: отраженный и преломленный, при этом все лучи и нормаль к поверхности раздела в точке падения луча лежат в одной плоскости.

Коэффициент отражения ρ гладкой поверхности, разделяющей два диэлектрика, зависит от состояния поляризации падающего луча. Электрические колебания в падающем луче всегда можно разложить на две составляющие: параллельные плоскости падения и перпендикулярные к ней. Таким образом, каждый луч можно считать состоящим из двух поляризованных лучей с колебаниями, происходящими в двух взаимно перпендикулярных плоскостях. Коэффициенты отражения каждого из этих лучей определяются формулами Френеля, причем для луча с колебаниями, перпендикулярными плоскости падения, коэффициент отражения

$$\rho_{\perp} = \frac{\sin^2(\theta_i - \theta_t)}{\sin^2(\theta_i + \theta_t)}, \quad (5.41)$$

а для луча с колебаниями, происходящими в плоскости падения, коэффициент отражения

$$\rho_{\parallel} = \frac{\operatorname{tg}^2(\theta_i - \theta_t)}{\operatorname{tg}^2(\theta_i + \theta_t)}. \quad (5.42)$$

При нормальном падении луча, т.е. при $\theta_i = \theta_t = 0$, выражения (5.41) и (5.42) принимают вид:

$$\rho_0 = \left(\frac{n_2 - n_1}{n_2 + n_1} \right)^2. \quad (5.43)$$

Используя это выражение, находим, что если первой средой является воздух ($n_1 = 1$), то для оптического стекла с показателем преломления $n_2 = n_e = 1,51829$ (крон К8) коэффициент отражения $\rho_0 = 0,0424$, для оптического стекла с показателем преломления $n_2 = n_e = 1,76171$ (тяжелый флинт ТФ5) коэффициент $\rho_0 = 0,0761$, а для воды при $n_2 = 1,33$ величина $\rho_0 = 0,02$. Представим соотношение (5.43) в виде:

$$\rho_0 = \left(1 - \frac{2n_1}{n_2 + n_1} \right)^2.$$

При этом вполне очевидно, что при $n_1 = 1$ чем больше величина показателя преломления n_2 , тем больше значение коэффициента отражения ρ_0 . В оптических устройствах, рассчитанных на применение излучения инфракрасной области спектра, применяют такие материалы, как кремний (*Si*) и германий (*Ge*), пропускающие инфракрасное излучение. Для излучения длины волны $\lambda = 7$ мкм показатели преломления $n_2(\text{Si}) = 3,42$, а $n_2(\text{Ge}) = 4,006$. В соответствии с выражением (5.43) при $n_1 = 1$ коэффициенты отражения $\rho_0(\text{Si}) = 0,30$, а $\rho_0(\text{Ge}) = 0,36$.

Если на гладкую поверхность диэлектрика падает естественный луч, в котором мощности обеих составляющих одинаковы, то для такого луча коэффициент отражения

$$\rho = \frac{1}{2} \left[\frac{\sin^2(\theta_i - \theta_t)}{\sin^2(\theta_i + \theta_t)} + \frac{\operatorname{tg}^2(\theta_i - \theta_t)}{\operatorname{tg}^2(\theta_i + \theta_t)} \right]. \quad (5.44)$$

Из формул Френеля следует, что после отражения от гладкой поверхности диэлектрика естественный свет становится частично поляризованным со степенью поляризации

$$P(\rho) = \frac{\rho_{\perp} - \rho_{\parallel}}{\rho_{\perp} + \rho_{\parallel}} = \frac{\cos^2(\theta_i - \theta_t) - \cos^2(\theta_i + \theta_t)}{\cos^2(\theta_i - \theta_t) + \cos^2(\theta_i + \theta_t)}, \quad (5.45)$$

причем наиболее сильное электрическое колебание в отраженном луче происходит в направлении, перпендикулярном плоскости падения. Из выражения (5.45) следует, что при нормальном ($\theta_i = 0$) и при скользющем ($\theta_i = \frac{\pi}{2}$) падении луча отраженный луч будет естественным ($P = 0$). При промежуточных значениях угла θ_i степень поляризации $P > 0$ и достигает максимума ($P = 1$), если $\theta_i + \theta_t = \frac{\pi}{2}$. Таким образом, если пучок лучей естественного света падает на гладкую поверхность диэлектрика под углом Брюстера, то отражается от нее только полностью поляризованная часть. В этом случае коэффициент отражения

$$\rho_B = \frac{1}{2} \frac{(n_2^2 - n_1^2)^2}{(n_2^2 + n_1^2)^2}.$$

Отсюда следует, что если менее плотной средой является воздух ($n_1 = 1$), то чем больше показатель преломления n_2 диэлектрика, тем большая часть падающего под углом Брюстера света отражается в виде полностью поляризованного света.

Изменение коэффициента отражения при изменении длины волны света определяется спектральной зависимостью показателя преломления, называемой дисперсией света. Дифференцируя функциональную зависимость (5.43) по n_2 в предположении, что $n_1 = 1$, получаем

$$\frac{d\rho_0}{\rho_0} = \frac{4dn}{n^2 - 1}, \quad (5.46)$$

где $n_2 = n$. В качестве основного принимаем показатель преломления n_e стекла для длины волны света $\lambda = \lambda_e = 546,07$ нм. Для стекла марки К8 ($n_e = 1,51829$) разность показателей преломления в пределах спектрального диапазона излучения от $\lambda = \lambda_h = 404,66$ нм до $\lambda = \lambda_r = 706,52$ нм равна 0,01736, а для тяжелого флинта марки ТФ5 ($n_e = 1,76171$) эта разность равна 0,06307. При этом коэффициент отражения крона меняется примерно на 0,05 от его основного значения, а для тяжелого флинта – на 0,12. Таким изменением коэффициента отражения во многих случаях можно пренебречь.

Формулы Френеля определяют также и коэффициент пропускания τ гладкой поверхности раздела двух диэлектриков. В том случае, когда электрические колебания перпендикулярны плоскости падения, коэффициент пропускания определяется выражением

$$\tau_{\perp} = \frac{\sin 2\theta_i \sin 2\theta_t}{\sin^2(\theta_i + \theta_t)}, \quad (5.47)$$

а для случаев, когда колебания параллельны плоскости падения,

$$\tau_{\parallel} = \frac{\sin 2\theta_i \sin 2\theta_t}{\sin^2(\theta_i + \theta_t) \cos^2(\theta_i - \theta_t)}. \quad (5.48)$$

Легко убедиться в том, что $\rho_{\perp} + \tau_{\perp} = \rho_{\parallel} + \tau_{\parallel} = 1$, как и должно быть в соответствии с законом сохранения энергии.

Коэффициент пропускания τ той же поверхности для естественного луча определяется выражением вида:

$$\tau = \frac{\sin 2\theta_i \sin 2\theta_t}{2 \sin^2(\theta_i + \theta_t)} \frac{1 + \cos^2(\theta_i - \theta_t)}{\cos^2(\theta_i - \theta_t)}, \quad (5.49)$$

при этом степень поляризации луча, прошедшего во вторую среду, равна

$$P(\tau) = \frac{\tau_{\parallel} - \tau_{\perp}}{\tau_{\parallel} + \tau_{\perp}} = \frac{1 - \cos^2(\theta_i - \theta_t)}{1 + \cos^2(\theta_i - \theta_t)}. \quad (5.50)$$

Из выражения (5.49) следует, что при нормальном падении луча ($\theta_i = 0$) это выражение принимает вид:

$$\tau_0 = \frac{4n_1 n_2}{(n_1 + n_2)^2}, \quad (5.51)$$

при этом степень поляризации луча во второй среде $P(\tau_0) = 0$.

5.5.3. Отражение света от гладкой поверхности металла

В случае однородной изотропной среды с диэлектрической проницаемостью ε , магнитной проницаемостью μ и проводимостью σ при соблюдении материальных уравнений $\mathbf{j} = \sigma \mathbf{E}$, $\mathbf{D} = \varepsilon \mathbf{E}$ и $\mathbf{B} = \mu \mathbf{H}$ уравнения Максвелла можно записать в виде:

$$\text{rot} \mathbf{H} - \frac{\varepsilon}{c} \dot{\mathbf{E}} = \frac{4\pi}{c} \sigma \mathbf{E}, \quad (5.52)$$

$$\text{rot} \mathbf{E} + \frac{\mu}{c} \dot{\mathbf{H}} = 0, \quad (5.53)$$

$$\text{div} \mathbf{E} = \frac{4\pi}{\varepsilon} \rho, \quad (5.54)$$

$$\text{div} \mathbf{H} = 0. \quad (5.55)$$

Применив операцию дивергенции к уравнению (5.52) и используя уравнение (5.54), получаем

$$-\frac{\varepsilon}{c} \text{div} \dot{\mathbf{E}} = \frac{4\pi}{c} \sigma \frac{4\pi}{\varepsilon} \rho.$$

Дифференцируя уравнение (5.54) по времени, имеем

$$\text{div} \dot{\mathbf{E}} = \frac{4\pi}{\varepsilon} \dot{\rho}.$$

Исключив $\text{div} \dot{\mathbf{E}}$ из двух последних уравнений, получаем дифференциальное уравнение

$$\dot{\rho} + \frac{4\pi\sigma}{\varepsilon}\rho = 0, \quad (5.56)$$

очевидное решение которого можно записать в виде:

$$\rho = \rho_0 \exp\left(-\frac{t}{\tau}\right), \quad (5.57)$$

где $\tau = \frac{\varepsilon}{4\pi\sigma}$. Время релаксации τ чрезвычайно мало для любой среды, обладающей заметной проводимостью. Для металлов это время значительно меньше периода световой волны. Так, например, для света оранжевой области спектра период колебаний волны равен $2 \cdot 10^{-15}$ сек, тогда как для меди отношение $\frac{\tau}{\varepsilon}$ порядка $2 \cdot 10^{-19}$ сек. Для любого реально мыслимого значения ε время τ так мало по сравнению с периодом световой волны, что в металле величина ρ всегда практически равна нулю. При этом уравнение (5.54) можно переписать в виде:

$$\operatorname{div}\mathbf{E} = 0. \quad (5.58)$$

Исключив из уравнений (5.52) и (5.53) величину \mathbf{H} и используя уравнение (5.58), получаем, что величина \mathbf{E} удовлетворяет волновому уравнению вида:

$$\nabla^2\mathbf{E} = \frac{\mu\varepsilon}{c^2}\ddot{\mathbf{E}} + \frac{4\pi\mu\sigma}{c^2}\dot{\mathbf{E}}. \quad (5.59)$$

Наличие члена с $\dot{\mathbf{E}}$ означает затухание волны, т.е. при распространении через среду волна постепенно ослабевает.

Если поле строго монохроматично и обладает круговой частотой ω , т.е. если \mathbf{E} и \mathbf{H} имеют вид $\mathbf{E} = \mathbf{E}_0 \exp(-i\omega t)$, то производная $\dot{\mathbf{E}} = -i\omega\mathbf{E}$, а $\ddot{\mathbf{E}} = -\omega^2\mathbf{E}$. При этом уравнение (5.59) примет вид:

$$\nabla^2\mathbf{E} + \tilde{k}^2\mathbf{E} = 0, \quad (5.60)$$

$$\text{где } \tilde{k}^2 = \frac{\omega^2\mu}{c^2}\left(\varepsilon + i\frac{4\pi\sigma}{\omega}\right). \quad (5.61)$$

Если ввести величину $\tilde{\varepsilon} = \varepsilon + i\frac{4\pi\sigma}{\omega}$, то уравнение (5.60) формально станет идентичным с уравнением Гельмгольца для непроводящих сред, где фигурирует вещественная диэлектрическая проницаемость ε . Аналогия с непроводящими средами еще больше усилится, если, кроме комплексного волнового числа \tilde{k} и комплексной диэлектрической проницаемости $\tilde{\varepsilon}$, ввести также комплексную фазовую скорость \tilde{V} и комплексный показатель преломления \tilde{n} , которые определяются соотношениями:

$$\tilde{V} = \frac{c}{\sqrt{\mu\tilde{\varepsilon}}}, \quad (5.62)$$

$$\tilde{n} = \frac{c}{\tilde{V}} = \sqrt{\mu\tilde{\epsilon}} = \frac{c}{\omega} \tilde{k}. \quad (5.63)$$

$$\text{Пусть } \tilde{n} = n(1 + i\chi), \quad (5.64)$$

где n и χ вещественны; величину χ назовем показателем затухания (называют также коэффициентом экстинкции). Величины n и χ легко выразить через материальные постоянные ϵ , μ и σ .

Возведя в квадрат обе части соотношения (5.63), получаем

$$\tilde{n}^2 = n^2(1 + 2i\chi - \chi^2).$$

Кроме того, из соотношений (5.62) и (5.63) имеем

$$\tilde{n}^2 = \mu\tilde{\epsilon} = \mu \left(\epsilon + i \frac{4\pi\sigma}{\omega} \right).$$

Приравнявая вещественные и мнимые части в этих двух выражениях для \tilde{n}^2 , получаем

$$n^2 = (1 - \chi^2) = \mu\epsilon, \quad (5.65)$$

$$n^2\chi = \frac{2\pi\mu\sigma}{\omega} = \frac{\mu\sigma}{\nu}. \quad (5.66)$$

Отсюда следует, что

$$n^2 = \frac{1}{2} \left(\sqrt{\mu^2\epsilon^2 + \frac{4\mu^2\sigma^2}{\nu^2}} + \mu\epsilon \right) \quad (5.67)$$

$$n^2\chi^2 = \frac{1}{2} \left(\sqrt{\mu^2\epsilon^2 + \frac{4\mu^2\sigma^2}{\nu^2}} - \mu\epsilon \right). \quad (5.68)$$

Здесь выбран положительный знак перед квадратным корнем, так как n и χ вещественны, а, следовательно, n^2 и $n^2\chi^2$ должны быть положительными.

Уравнение (5.60) формально идентично волновому уравнению для непроводящей среды, но в этом уравнении волновое число комплексное. Простейшим решением уравнения (5.60) является плоская гармоническая во времени волна

$$\mathbf{E} = \mathbf{E}_0 \exp\{i[\tilde{k}(\mathbf{r} \cdot \mathbf{s}) - \omega t]\}. \quad (5.69)$$

В соответствии с соотношениями (5.63) и (5.64) находим, что $\tilde{k} = \frac{\omega}{c} \tilde{n} = \frac{\omega}{c} n(1 + i\chi)$. При этом выражение (5.69) можно представить в виде:

$$\mathbf{E} = \mathbf{E}_0 \exp\left[-\frac{\omega}{c} n\chi(\mathbf{r} \cdot \mathbf{s})\right] \exp\left\{i\omega\left[\frac{n}{c}(\mathbf{r} \cdot \mathbf{s}) - t\right]\right\}.$$

Вещественная часть этого выражения

$$\operatorname{Re} \mathbf{E} = \mathbf{E}_0 \exp\left[-\frac{\omega}{c} n \chi(\mathbf{r} \cdot \mathbf{s})\right] \cos\left\{\omega\left[\frac{n}{c}(\mathbf{r} \cdot \mathbf{s}) - t\right]\right\}, \quad (5.70)$$

представляющая собой электрический вектор, является плоской волной длины $\lambda = 2\pi \frac{c}{\omega n}$, затухание которой определяется экспоненциальным членом. Так как плотность энергии w волны пропорциональна среднему по времени от \mathbf{E}^2 , то вполне очевидно, что w будет уменьшаться в соответствии с законом

$$w = w_0 \exp[-\aleph(\mathbf{r} \cdot \mathbf{s})], \quad (5.71)$$

$$\text{где } \aleph = \frac{2\omega}{c} n \chi = \frac{4\pi\nu}{c} n \chi = \frac{4\pi}{\lambda_0} n \chi = \frac{4\pi}{\lambda} \chi. \quad (5.72)$$

Здесь λ_0 – длина волны в вакууме, λ – длина волны в среде. Постоянная \aleph называется коэффициентом поглощения. Плотность энергии падает в e раз на расстоянии d ,

$$\text{где } d = \frac{1}{\aleph} = \frac{\lambda_0}{4\pi n \chi} = \frac{\lambda}{4\pi \chi}. \quad (5.73)$$

«Глубина проникновения» d излучения различных длин волн λ_0 в медь ($\sigma \approx 5,14 \cdot 10^{17} \text{ сек}^{-1}$, $\mu = 1$) представлена в табл. 5.1. Обычно эта величина составляет очень малую долю длины волны излучения.

Таблица 5.1. «Глубина проникновения» излучения различных длин волн в медь

	Ультрафиолетовая область	Инфракрасная область	СВЧ	Длинные радиоволны
λ_0	1000 Å = 10^{-5} см	10 мк = 10^{-3} см	10 см	1000 м = 10^5 см
d	$6,2 \cdot 10^{-8}$ см	$6,2 \cdot 10^{-7}$ см	$6,2 \cdot 10^{-5}$ см	$6,2 \cdot 10^{-3}$ см

Заметим, что при $\sigma = 0$ уравнение (5.67) переходит в соотношение Максвелла: $n^2 = \mu\epsilon$, при этом из уравнения (5.66) следует, что в этом случае $\chi = 0$. Идеальный проводник характеризуется бесконечной проводимостью. Из соотношений (5.65) и (5.66) находим, что $\frac{\epsilon}{\sigma} = \frac{1 - \chi^2}{\nu \chi}$.

Отсюда следует, что при $\sigma \rightarrow \infty$ величина $\chi^2 \rightarrow 1$, при этом величина $n \rightarrow \infty$. Такой проводник вообще не позволял бы электромагнитной волне проникать на какую-либо глубину и отражал бы весь падающий свет.

Итак, основные уравнения, описывающие распространение плоской гармонической волны в проводящей среде, отличаются от соответствующих уравнений для прозрачного диэлектрика лишь тем, что вещественные постоянные ϵ и k заменяются на комплексные $\tilde{\epsilon}$ и \tilde{k} . Поскольку формулы, полученные для прозрачного диэлектрика, содержат лишь линейные соотношения между компонентами векторов поля плоских

монохроматических волн, то они применимы и в рассматриваемом случае. В частности, остаются справедливыми граничные условия для распространения волны через поверхность раздела, а, следовательно, и формулы, относящиеся к преломлению и отражению.

Рассмотрим распространение плоской волны из диэлектрика в проводник, причем обе среды будем считать бесконечными, а за поверхность раздела между ними выберем плоскость $z = 0$. При этом закон преломления луча на поверхности раздела сред можно записать в виде:

$$\sin \theta_t = \frac{1}{\tilde{n}} \sin \theta_i. \quad (5.74)$$

Так как показатель преломления \tilde{n} – величина комплексная, то комплексным будет и угол θ_t . Таким образом, эта величина уже не имеет простого смысла угла преломления.

Выберем в качестве плоскости падения плоскость xz . Тогда зависящая от координат часть фазы волны в проводнике определится выражением $\tilde{k}(\mathbf{r} \cdot \mathbf{s}^{(t)})$, где

$$s_x^{(t)} = \sin \theta_t, \quad s_y^{(t)} = 0, \quad s_z^{(t)} = \cos \theta_t. \quad (5.75)$$

Из соотношений (5.74), (5.75) и (5.64) имеем

$$s_x^{(t)} = \sin \theta_t = \frac{\sin \theta_i}{n(1+i\chi)} = \frac{1-i\chi}{1+\chi^2} \frac{\sin \theta_i}{n} \quad (5.76)$$

$$\begin{aligned} s_z^{(t)} &= \cos \theta_t = \sqrt{1 - \sin^2 \theta_t} = \\ &= \sqrt{1 - \frac{1-\chi^2}{(1+\chi^2)^2} \frac{\sin^2 \theta_i}{n^2} + i \frac{2\chi}{(1+\chi^2)^2} \frac{\sin^2 \theta_i}{n^2}}. \end{aligned} \quad (5.77)$$

Величину $s_z^{(t)}$ удобно представить в форме:

$$s_z^{(t)} = \cos \theta_t = q e^{i\gamma}, \quad (5.78)$$

где q и γ – вещественные величины, которые можно выразить через n , χ и $\sin \theta_i$, если возвести в квадрат соотношения (5.77) и (5.78) и приравнять отдельно вещественные и мнимые части. Это дает

$$q^2 \cos 2\gamma = 1 - \frac{1-\chi^2}{(1+\chi^2)^2} \frac{\sin^2 \theta_i}{n^2},$$

$$q^2 \sin 2\gamma = \frac{2\chi}{(1+\chi^2)^2} \frac{\sin^2 \theta_i}{n^2}.$$

Тогда, используя соотношения (5.76) и (5.78), получаем

$$\begin{aligned}
\tilde{k}(\mathbf{r} \cdot \mathbf{s}^{(t)}) &= \frac{\omega}{c} n(1+i\chi)(x s_x^{(t)} + z s_z^{(t)}) = \\
&= \frac{\omega}{c} n(1+i\chi) \left[\frac{x(1-i\chi)}{n(1+\chi^2)} \sin \theta_i + z(q \cos \gamma + iq \sin \gamma) \right] = \\
&= \frac{\omega}{c} [x \sin \theta_i + znq(\cos \gamma - \chi \sin \gamma) + iznq(\chi \cos \gamma + \sin \gamma)].
\end{aligned} \tag{5.79}$$

Отсюда следует, что величины q и γ входят в выражение (5.79) только в сочетании с координатой z , т.е. поверхности постоянной амплитуды определяются уравнением

$$z = const, \tag{5.80}$$

а, следовательно, являются плоскостями, параллельными поверхности раздела. Поверхности постоянной вещественной фазы определяются уравнением

$$x \sin \theta_i + znq(\cos \gamma - \chi \sin \gamma) = const \tag{5.81}$$

и являются плоскостями, нормали к которым образуют угол θ'_t с нормалью к границе раздела сред, причем

$$\left. \begin{aligned}
\sin \theta'_t &= \frac{\sin \theta_i}{\sqrt{\sin^2 \theta_i + n^2 q^2 (\cos \gamma - \chi \sin \gamma)^2}} \\
\cos \theta'_t &= \frac{nq(\cos \gamma - \chi \sin \gamma)}{\sqrt{\sin^2 \theta_i + n^2 q^2 (\cos \gamma - \chi \sin \gamma)^2}}.
\end{aligned} \right\} \tag{5.82}$$

В общем случае поверхности постоянной амплитуды и поверхности постоянной фазы не совпадают друг с другом, а поэтому волна в металле оказывается неоднородной.

Если квадратный корень в соотношениях (5.82) обозначить через n' , то уравнение для $\sin \theta'_t$ можно переписать в виде закона преломления:

$$\sin \theta'_t = \frac{\sin \theta_i}{n'}.$$

Однако, здесь n' зависит не только от величин, характеризующих среду, но и от угла падения θ_i .

Так как, по предположению, первой средой служит диэлектрик, то отраженная волна будет обычной (однородной) волной с вещественной фазой. Как и в соотношениях (1.21), компоненты амплитуды падающей волны A_{\parallel} и A_{\perp} и соответствующие компоненты отраженной волны R_{\parallel} и R_{\perp} связаны соотношениями

$$\left. \begin{aligned}
R_{\parallel} &= -\frac{\operatorname{tg}(\theta_i - \theta_t)}{\operatorname{tg}(\theta_i + \theta_t)} A_{\parallel}, \\
R_{\perp} &= -\frac{\sin(\theta_i - \theta_t)}{\sin(\theta_i + \theta_t)} A_{\perp}.
\end{aligned} \right\} \tag{5.83}$$

Поскольку угол θ_t – комплексный, то комплексны и отношения $\frac{R_{\parallel}}{A_{\parallel}}$ и $\frac{R_{\perp}}{A_{\perp}}$, т.е. при отражении происходят характерные изменения фазы. Таким образом, падающий линейно поляризованный свет при отражении от поверхности металла в общем случае становится эллиптически поляризованным.

Пусть φ_{\parallel} и φ_{\perp} – фазы, а $r_{0\parallel}$ и $r_{0\perp}$ – абсолютные значения коэффициентов отражения, т.е.

$$r_{\parallel} = \frac{R_{\parallel}}{A_{\parallel}} = \Re_{\parallel} \exp(i\varphi_{\parallel}), \quad r_{\perp} = \frac{R_{\perp}}{A_{\perp}} = \Re_{\perp} \exp(i\varphi_{\perp}). \quad (5.84)$$

Предположим, что падающий свет линейно поляризован и азимутальный угол равен α_i , т.е.

$$\operatorname{tg}\alpha_i = \frac{A_{\perp}}{A_{\parallel}}, \quad (5.85)$$

и пусть α_r – азимутальный угол (обычно комплексный) для отраженного света, т.е.

$$\operatorname{tg}\alpha_r = \frac{R_{\perp}}{R_{\parallel}} = \frac{\cos(\theta_i - \theta_t)}{\cos(\theta_i + \theta_t)} \operatorname{tg}\alpha_i = p e^{i\Delta} \operatorname{tg}\alpha_i, \quad (5.86)$$

$$\text{где } p = \frac{\Re_{\perp}}{\Re_{\parallel}}, \quad \Delta = \varphi_{\perp} - \varphi_{\parallel}. \quad (5.87)$$

Заметим, что величина α_r вещественна в следующих двух случаях:

1. При нормальном падении плоской волны ($\theta_i = 0$); тогда $P = 1$ и $\Delta = 0$, так что $\operatorname{tg}\alpha_r = \operatorname{tg}\alpha_i$.

2. При скользющем падении плоской волны $\left(\theta_i = \frac{\pi}{2}\right)$; тогда $P = 1$ и $\Delta = \pi$, так что $\operatorname{tg}\alpha_r = -\operatorname{tg}\alpha_i$.

Между двумя только что рассмотренными экстремальными углами существует так называемый главный угол падения θ_i , для которого $\Delta = \frac{\pi}{2}$.

При этом угле линейно поляризованный свет в общем случае превращается после отражения в эллиптически поляризованный свет. Если, кроме того, $p \operatorname{tg}\alpha_i = 1$, то согласно соотношению (5.86) $\operatorname{tg}\alpha_r = i$, при этом отраженный свет поляризован по кругу.

Предположим, что на металл падает линейно поляризованный свет, и между R_{\parallel} и R_{\perp} вводится с помощью подходящего компенсатора дополнительная разность фаз Δ . Если полная разность фаз равна нулю, то

отраженный свет будет линейно поляризован с азимутом α'_r , в соответствии с выражением (5.86) определяемым из соотношения

$$tg\alpha'_r = ptg\alpha_i. \quad (5.88)$$

Вполне естественно, что угол α'_r называется углом восстановленной поляризации, хотя его обычно определяют только для падающего света, линейно поляризованного с азимутом $\alpha_i = 45^\circ$.

Введем угол ψ , определяемый из соотношения

$$tg\psi = p. \quad (5.89)$$

Используя соотношения (5.74) и (5.83) и зная постоянные металла n и k , легко найти зависимость величин p и Δ от θ_i . Однако, проблема заключается не в том, чтобы найти ψ и Δ по известным значениям n и χ , а чтобы определить n и χ по экспериментально наблюдаемым значениям амплитуды и фазы света, отраженного от металла.

Так как все величины R_{\parallel} , R_{\perp} , Φ_{\parallel} , Φ_{\perp} , ψ и Δ являются функциями угла θ_i и величин n и χ , то измерение любых двух из них для какого-то значения угла падения θ_i позволит, вообще говоря, найти значения n и χ . Во многих экспериментах определяют последние две из этого ряда величин, а поэтому представляет интерес получить зависимости n и χ от ψ и Δ .

Из выражений (5.74) и (5.86) имеем

$$\frac{1 + p \exp(i\Delta)}{1 - p \exp(i\Delta)} = -\frac{\cos\theta_i \cos\theta_t}{\sin\theta_i \sin\theta_t} = -\frac{\sqrt{\tilde{n}^2 - \sin^2\theta_i}}{\sin\theta_i tg\theta_i}. \quad (5.90)$$

Так как $p = tg\psi$, левую часть уравнения (5.90) можно преобразовать к виду:

$$\frac{1 + p \exp(i\Delta)}{1 - p \exp(i\Delta)} = \frac{1 + \exp(i\Delta)tg\psi}{1 - \exp(i\Delta)tg\psi} = \frac{\cos 2\psi + i \sin 2\psi \sin \Delta}{1 - \sin 2\psi \cos \Delta}. \quad (5.91)$$

Подставив это выражение в выражение (5.90), получаем

$$\frac{\sqrt{\tilde{n}^2 - \sin^2\theta_i}}{\sin\theta_i tg\theta_i} = -\frac{\cos 2\psi + i \sin 2\psi \sin \Delta}{1 - \sin 2\psi \cos \Delta}. \quad (5.92)$$

В видимой области спектра обычно

$$n^2(1 + \chi^2) \gg 1,$$

а поэтому величиной $\sin^2\theta$ по сравнению с \tilde{n}^2 можно пренебречь. Тогда

$$\frac{\tilde{n}}{\sin\theta_i tg\theta_i} = \frac{n(1 + i\chi)}{\sin\theta_i tg\theta_i} \approx -\frac{\cos 2\psi + i \sin 2\psi \sin \Delta}{1 - \sin 2\psi \cos \Delta}.$$

Приравнявая вещественные части, находим, что

$$n \approx -\frac{\sin\theta_i tg\theta_i \cos 2\psi}{1 - \sin 2\psi \cos \Delta}. \quad (5.93)$$

Приравнивая мнимые части и используя найденное соотношение (5.93), получаем

$$\chi \approx \operatorname{tg} 2\psi \sin \Delta. \quad (5.94)$$

Иногда оказываются полезными и другие соотношения для n и χ . Возведем в квадрат левую и правую части соотношения (5.92):

$$\frac{\tilde{n}^2 - \sin^2 \theta_i}{\sin^2 \theta_i \operatorname{tg}^2 \theta_i} = - \frac{\cos^2 2\psi - \sin^2 2\psi \sin^2 \Delta + i \sin 4\psi \sin \Delta}{(1 - \sin 2\psi \cos \Delta)^2}.$$

Если учесть, что $\tilde{n}^2 = n^2(1 - \chi^2) + 2in^2\chi$, и приравнять отдельно вещественные и мнимые части, то получим

$$n^2(1 - \chi^2) = \sin^2 \theta_i \left[1 + \operatorname{tg}^2 \theta_i \frac{\cos^2 2\psi - \sin^2 2\psi \sin^2 \Delta}{(1 - \sin 2\psi \cos \Delta)^2} \right] \quad (5.95)$$

$$2n^2\chi = \frac{\sin^2 \theta_i \operatorname{tg}^2 \theta_i \sin 4\psi \sin \Delta}{(1 - \sin 2\psi \cos \Delta)^2}. \quad (5.96)$$

Формулы (5.95) и (5.96) определяют не n и χ , а комбинации $n^2(1 - \chi^2)$ и $n^2\chi$, которые имеют простой физический смысл. При $\mu = 1$ (что всегда справедливо в оптическом диапазоне) величина $n^2(1 - \chi^2)$ в соответствии с формулой (5.65) определяет диэлектрическую проницаемость, а величина $n^2\chi$ в соответствии с формулой (5.66) определяет отношение проводимости к частоте. Из значений этих величин и, в частности, из их дисперсии можно получить информацию о структуре металлов.

Рассмотрим интенсивность света, отраженного гладкой поверхностью металла и прошедшего сквозь нее. Удобно положить

$$\tilde{n}_2 \cos \theta_2 = u_2 + iv_2, \quad (5.97)$$

где u_2 и v_2 – вещественные величины. Возведя выражение (4.97) в квадрат и используя закон преломления в виде: $\tilde{n}_2 \sin \theta_2 = n_1 \sin \theta_1$, получаем

$$(u_2 + iv_2)^2 = \tilde{n}_2^2 - n_1^2 \sin^2 \theta_1. \quad (5.98)$$

Приравнивая отдельно вещественную и мнимую части выражения (5.98), находим

$$\left. \begin{aligned} u_2^2 - v_2^2 &= n_2^2(1 - \chi_2^2) - n_1^2 \sin^2 \theta_1, \\ u_2 v_2 &= n_2^2 \chi_2. \end{aligned} \right\} \quad (5.99)$$

Отсюда следует, что

$$\left. \begin{aligned} 2u_2^2 &= n_2^2(1 - \chi_2^2) - n_1^2 \sin^2 \theta_1 + \\ &+ \sqrt{[n_2^2(1 - \chi_2^2) - n_1^2 \sin^2 \theta_1]^2 + 4n_2^4 \chi_2^2}, \\ 2v_2^2 &= -[n_2^2(1 - \chi_2^2) - n_1^2 \sin^2 \theta_1] + \\ &+ \sqrt{[n_2^2(1 - \chi_2^2) - n_1^2 \sin^2 \theta_1]^2 + 4n_2^4 \chi_2^2}. \end{aligned} \right\} \quad (5.100)$$

При определении коэффициентов отражения и пропускания света на поверхности раздела прозрачного диэлектрика и проводника рассмотрим отдельно случаи, когда электрический вектор падающей волны перпендикулярен и параллелен плоскости падения.

Электрический вектор перпендикулярен плоскости падения (ТЕ волна)

В этом случае при замене в соответствующем выражении (1.119) величины $n_2 \cos \theta_i$ на $n_2 \cos \theta_2 = u_2 + iv_2$ получаем

$$r_{\perp} = \mathfrak{R}_{\perp} \exp(i\varphi_{\perp}) = \frac{n_1 \cos \theta_1 - (u_2 + iv_2)}{n_1 \cos \theta_1 + (u_2 + iv_2)}. \quad (5.101)$$

Здесь $\theta_1 = \theta_i$. Из соотношения (5.101) находим, что

$$\rho_{\perp} = \mathfrak{R}_{\perp}^2 = \frac{(n_1 \cos \theta_1 - u_2)^2 + v_2^2}{(n_1 \cos \theta_1 + u_2)^2 + v_2^2} \quad (5.102)$$

$$\operatorname{tg} \varphi_{\perp} = \frac{2v_2 n_1 \cos \theta_1}{n_1^2 \cos^2 \theta_1 - u_2^2 - v_2^2}. \quad (5.103)$$

В соответствии с формулой (1.137) имеем

$$t_{\perp} = \mathfrak{T}_{\perp} \exp(i\aleph_{\perp}) = \frac{2n_1 \cos \theta_1}{n_1 \cos \theta_1 + u_2 + iv_2}. \quad (5.104)$$

В соответствии с формулой (1.130) коэффициент пропускания $\tau = \frac{n_2 \cos \theta_2}{n_1 \cos \theta_1} |t|^2$. Здесь все величины вещественные. Поэтому, учитывая

физический смысл отношения $\frac{n_2 \cos \theta_2}{n_1 \cos \theta_1}$ в выражении (1.130), в

рассматриваемом случае можем написать

$$\tau = \frac{\sqrt{n_2^2 - n_1^2 \sin^2 \theta_1}}{n_1 \cos \theta_1} |t|^2 = \frac{\sqrt{n_2^2 - n_1^2 \sin^2 \theta_1}}{n_1 \cos \theta_1} \mathfrak{T}^2.$$

При этом

$$\tau_{\perp} = \frac{4n_1 \cos \theta_1 \sqrt{n_2^2 - n_1^2 \sin^2 \theta_1}}{(n_1 \cos \theta_1 + u_2)^2 + v_2^2} \quad (5.105)$$

$$\operatorname{tg} \aleph_{\perp} = \frac{v_2}{n_1 \cos \theta_1 + u_2}. \quad (5.106)$$

Электрический вектор параллелен плоскости падения (ТМ волна)

Используя соответствующую формулу (1.119), получаем

$$\begin{aligned}
r_{\parallel} &= \Re_{\parallel} \exp(i\varphi_{\parallel}) = -\frac{\frac{1}{n_1} \cos \theta_1 - \frac{1}{\tilde{n}_2} \cos \theta_2}{\frac{1}{n_1} \cos \theta_1 + \frac{1}{\tilde{n}_2} \cos \theta_2} = \\
&= -\frac{\tilde{n}_2^2 \cos \theta_1 - n_1 \tilde{n}_2 \cos \theta_2}{\tilde{n}_2^2 \cos \theta_1 + n_1 \tilde{n}_2 \cos \theta_2} = \\
&= -\frac{[n_2^2(1-\chi_2^2) + 2in_2^2\chi_2] \cos \theta_1 - n_1(u_2 + iv_2)}{[n_2^2(1-\chi_2^2) + 2in_2^2\chi_2] \cos \theta_1 + n_1(u_2 + iv_2)}.
\end{aligned} \tag{5.107}$$

Отсюда находим

$$\rho_{\parallel} = \Re_{\parallel}^2 = \frac{[n_2^2(1-\chi_2^2)\cos\theta_1 - n_1u_2]^2 + [2n_2^2\chi_2\cos\theta_1 - n_1v_2]^2}{[n_2^2(1-\chi_2^2)\cos\theta_1 + n_1u_2]^2 + [2n_2^2\chi_2\cos\theta_1 + n_1v_2]^2} \tag{5.108}$$

$$\operatorname{tg} \varphi_{\parallel} = 2n_1n_2^2 \cos \theta_1 \frac{(1-\chi_2^2)v_2 - 2\chi_2u_2}{n_2^4(1+\chi_2^2)^2 \cos^2 \theta_1 - n_1^2(u_2^2 + v_2^2)}. \tag{5.109}$$

Используя соответствующую формулу (1.118), получаем

$$\begin{aligned}
t_{\parallel} &= \Im_{\parallel} \exp(i\mathfrak{N}_{\parallel}) = \frac{2n_1\tilde{n}_2 \cos \theta_1}{\tilde{n}_2^2 \cos \theta_1 + n_1\tilde{n}_2 \cos \theta_2} = \\
&= \frac{2(1+i\chi_2) n_1n_2 \cos \theta_1}{[n_2^2(1-\chi_2^2) + 2in_2^2\chi_2] \cos \theta_1 + n_1(u_2 + iv_2)}.
\end{aligned} \tag{5.110}$$

При этом

$$\tau_{\parallel} = \Im_{\parallel}^2 = \frac{4n_1 \cos \theta_1 (1+\chi_2^2) n_2^2 \sqrt{n_2^2 - n_1^2 \sin^2 \theta_1}}{[n_2^2(1-\chi_2^2)\cos\theta_1 + n_1u_2]^2 + (2n_2^2\chi_2\cos\theta_1 + n_1v_2)^2} \tag{5.111}$$

$$\operatorname{tg} \mathfrak{N}_{\parallel} = \frac{n_2^2(1+\chi_2^2)\chi_2 \cos \theta_1 - n_1(\chi_2u_2 - v_2)}{n_2^2(1+\chi_2^2) \cos \theta_1 + n_1(u_2 + \chi_2v_2)}. \tag{5.112}$$

При нормальном падении света на поверхность раздела сред угол $\theta_1 = 0$, при этом различие между R_{\parallel} и R_{\perp} исчезает, плоскость падения становится неопределенной, а величина u_2 и v_2 в соответствии с выражениями (5.100) становятся равными $u_2 = n_2$, $v_2 = n_2\chi_2$. Подставив эти величины в выражения (5.102) и (5.108), получаем

$$\rho = \left| \frac{R_{\parallel}}{A_{\parallel}} \right|^2 = \left| \frac{R_{\perp}}{A_{\perp}} \right|^2 = \frac{(n_2 - n_1)^2 + n_2^2\chi_2^2}{(n_2 + n_1)^2 + n_2^2\chi_2^2} = \left| \frac{\tilde{n}_2 - n_1}{\tilde{n}_2 + n_1} \right|^2 \tag{5.113}$$

или

$$\rho = \left| \frac{\tilde{n} - 1}{\tilde{n} + 1} \right|^2 = \frac{(n-1)^2 + n^2\chi^2}{(n+1)^2 + n^2\chi^2}, \quad (5.114)$$

где $\tilde{n} = \frac{\tilde{n}_2}{n_1}$, $n = \frac{n_2}{n_1}$, $\chi = \chi_2$.

Таблица 5.2. Оптические постоянные некоторых металлов для $\lambda = 0,589$ мкм

Металл	χ	n	ρ
Алюминий	5,23	1,44	0,83
Ртуть	4,80	1,60	0,77
Сурьма	4,94	3,04	0,70
Платина	4,26	2,06	0,70
Серебро	3,67	0,18	0,95
Висмут	3,66	1,90	0,65
Золото	2,82	0,37	0,85
Медь	2,62	0,64	0,73
Свинец	3,48	2,01	0,62
Вольфрам	3,25	3,46	0,54
Никель	3,42	1,58	0,66

Оптические постоянные ряда чистых металлов, полученные экспериментально, приведены в табл. 5.2. Постоянные n и χ , а вместе с ними и коэффициенты отражения металлов, зависят от длины волны света. Эта зависимость в области длин волн от 0,2 до 0,7 мкм показана на рис. 5.7 для некоторых часто используемых металлов. Наименее селективным (наиболее нейтральным) из них является алюминий, а наибольшим коэффициентом отражения в видимой области спектра обладает серебро.

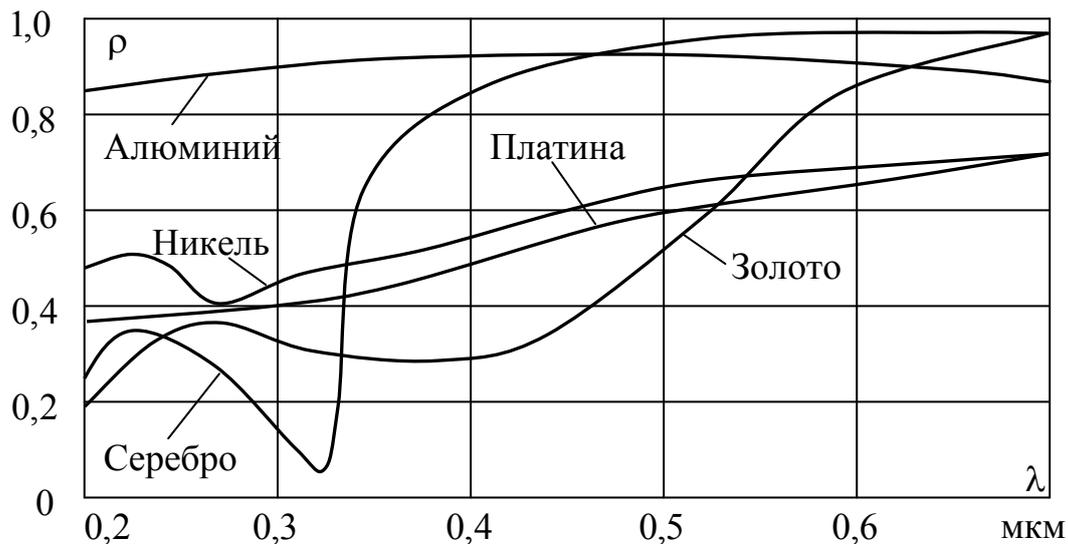


Рис. 5.7. Коэффициенты отражения металлов в видимых и ультрафиолетовых лучах

Кривая коэффициента отражения серебра имеет резко выраженный минимум при длине волны излучения, равной 0,316 мкм. На этом участке его коэффициент отражения такой же, как у стекла. Слой серебра толщиной 0,1 мкм пропускает при этой длине волны около половины падающего потока, в то время как при длине волны 0,4 мкм он пропускает около 0,01 %, а при $\lambda = 0,5$ мкм – не больше 0,001 %. При смещении длины волны падающего излучения в инфракрасную область коэффициенты его отражения от поверхности металлов становятся больше. Это следует и из рис. 5.8, на котором показаны кривые коэффициентов отражения от слоев серебра, золота и алюминия, нанесенных испарением наиболее чистых материалов в сверхвысоком вакууме ($10^{-9} - 10^{-10}$) Па. Высокие коэффициенты отражения от алюминиевых зеркал (до 92 %) можно получить и в ультрафиолетовой области спектра (до 220 нм), если испарять чистый металл (чистота 99,99 %) с очень большой скоростью (примерно 300 Å/с и выше).

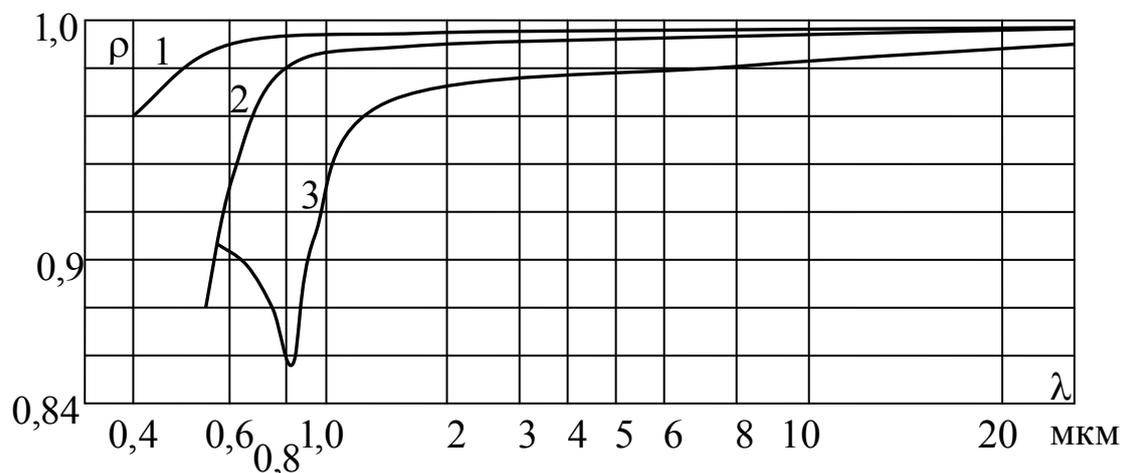


Рис. 5.8. Коэффициенты отражения чистых серебра (1), золота (2) и алюминия (3) в видимых и инфракрасных лучах

5.5.4. Поглощение света в веществе

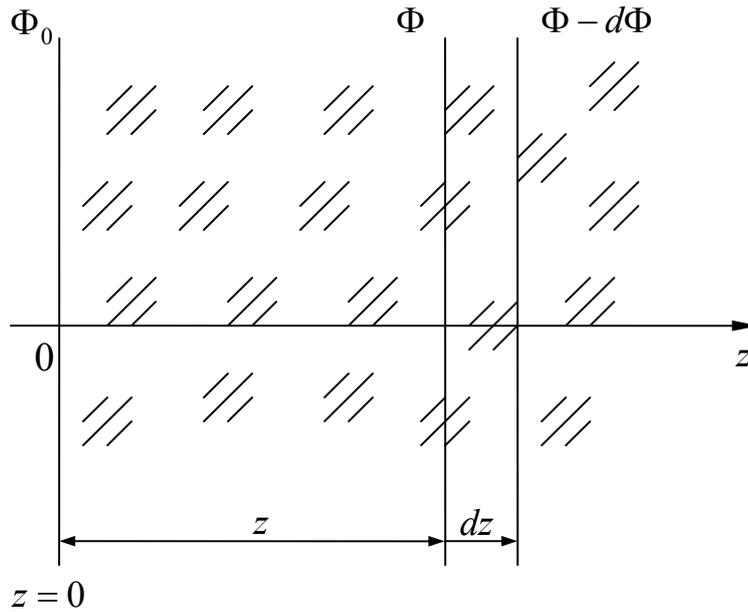


Рис. 5.9. Поглощение света в веществе

Предположим, что через толщу однородного прозрачного (т.е. нерассеивающего), но поглощающего вещества распространяется пучок лучей излучения, длины волн которого приходятся на узкий спектральный интервал $\lambda \pm \Delta\lambda$. Пучок лучей заполняет малый телесный угол в направлении оси z , а поэтому будем считать его пучком параллельных лучей. Когда этот пучок лучей проходит через плоскость, отстоящую на расстоянии z от начала отсчета, как показано на рис. 5.9, то переносимый им световой поток будем считать равным Φ . В плоскости с координатой $z+dz$, т.е. после прохождения слоя вещества толщиной dz , поток уменьшается на величину $d\Phi$. В области малых изменений величин естественно положить

$$-d\Phi = \alpha' \Phi dz, \quad (5.115)$$

где α' – натуральный показатель поглощения рассматриваемого вещества для излучения длины волны λ , который следует считать величиной постоянной. Поскольку при увеличении расстояния z (dz положительно) световой поток Φ уменьшается, величина $d\Phi$ должна быть отрицательной, что и определено знаком минус в выражении (5.115). Из соотношения (5.115) следует, что показатель поглощения α' имеет размерность, обратную размерности длины.

Дифференциальное уравнение (5.115) легко решить путем разделения переменных. Вводя постоянную интегрирования C , получаем

$$\int \frac{d\Phi}{\Phi} = \ln \Phi = -\alpha' z + C$$

или

$$\Phi = \exp(-\alpha' z + C). \quad (5.116)$$

Полагая значение светового потока при $z=0$ равным Φ_0 , имеем $\Phi_0 = \exp(C)$. При этом выражение (5.116) можно записать в виде

$$\Phi = \Phi_0 \exp(-\alpha'z).$$

Отсюда следует, что коэффициент пропускания слоя толщиной z равен

$$\tau = \frac{\Phi}{\Phi_0} = \exp(-\alpha'z). \quad (5.117)$$

Соотношение (5.117) определяет закон Бугера-Ламберта. Закон Бугера был установлен в 1729 году, а в 1760 году Ламберт придал ему удобную математическую форму, которой пользуются и в настоящее время. В начале XX века С.И. Вавилов экспериментально показал, что в пределах очень широких изменений плотности излучения (в 10^{20} раз) показатель поглощения многих веществ остается неизменным.

Итак, прохождение параллельного пучка лучей через поглощающую среду связано с поглощением излучения в объеме. Если считать, что площадь поперечного сечения пучка равна S , то объем, в котором поглощается поток $d\Phi$, равен Sdz . Разделив выражение (5.115) на $dV = Sdz$ и отбросив знак минус, получаем, что объемная плотность поглощенного потока равна

$$\frac{d\Phi}{dV} = \frac{\alpha'\Phi}{S} = \alpha'E_n,$$

где E_n – освещенность плоскости, перпендикулярной к параллельному пучку лучей.

В практических расчетах чаще пользуются не натуральным, а десятичным показателем поглощения a , выражая коэффициент пропускания соотношениями вида:

$$\tau = 10^{-0,434\alpha'z} = 10^{-az}, \quad (5.118)$$

где десятичный показатель поглощения

$$a = 0,434\alpha' = -\frac{\lg \tau}{z}, \quad (5.119)$$

$\lg \tau$ – десятичный логарифм τ .

Из выражения (4.118) следует, что если $z = \frac{1}{a}$, то коэффициент пропускания $\tau = 10^{-1}$, т.е. длина $\frac{1}{a}$ представляет собой такую толщину слоя, после прохождения которой пучок ослабляется в 10 раз.

Величину

$$D = \lg \frac{1}{\tau} = -\lg \tau = az \quad (5.120)$$

принято называть оптической плотностью слоя поглощающего вещества. Из выражения (5.120) следует, что десятичный показатель поглощения равен оптической плотности слоя, толщина которого равна единице.

Таблица 5.3. Коэффициент пропускания τ и оптическая плотность D

Коэффициент пропускания τ	0	1	2	3	4	5	6	7	8	9
	Оптическая плотность D									
0,1	1,000	0,959	0,921	0,996	0,854	0,824	0,796	0,770	0,745	0,721
0,2	0,699	0,678	0,658	0,638	0,620	0,602	0,585	0,569	0,553	0,538
0,3	0,523	0,509	0,495	0,481	0,469	0,456	0,444	0,432	0,420	0,409
0,4	0,398	0,387	0,377	0,367	0,357	0,347	0,337	0,328	0,319	0,310
0,5	0,301	0,292	0,284	0,276	0,268	0,260	0,252	0,244	0,237	0,229
0,6	0,222	0,215	0,208	0,201	0,194	0,187	0,180	0,174	0,168	0,161
0,7	0,155	0,149	0,143	0,137	0,131	0,125	0,119	0,114	0,108	0,102
0,8	0,097	0,092	0,086	0,081	0,076	0,071	0,065	0,060	0,056	0,051
0,9	0,046	0,041	0,036	0,032	0,027	0,022	0,018	0,013	0,009	0,004
1,0	0,000									

В табл. 5.3 даны оптические плотности, соответствующие коэффициенту пропускания τ от 0,1 до 1,0 через 0,01. Первый столбец этой таблицы содержит десятые доли коэффициента пропускания, а его сотые доли приведены в первой строке. Таким образом, оптическая плотность, соответствующая любому двухзначному коэффициенту пропускания, может быть найдена на пересечении соответствующей строки и нужного столбца. Если коэффициент пропускания оказывается меньше 0,1 (например, 0,0035), то соответствующую ему плотность можно найти, передвинув запятую вправо до первой значащей цифры и увеличив затем найденную плотность на столько единиц, на сколько знаков пришлось предварительно передвинуть запятую. Из табл. 5.3 следует, что при $\tau = 0,35$ плотность $D = 0,456$. Следовательно, при $\tau = 0,0035$ плотность $D = 2,456$.

Оптической плотностью пользуются очень широко и часто в расчетах предпочитают ее коэффициенту пропускания или поглощения. На это есть две причины. Первая, которая имеет существенное значение для химиков, состоит в том, что во многих случаях оптическая плотность раствора пропорциональна количеству растворенного вещества. Часто можно считать, что растворитель практически не поглощает проходящего излучения. Тогда справедлив закон Бэра, установленный в 1852 году, согласно которому показатель поглощения a раствора пропорционален его концентрации C . В таком случае можно написать, что $a = kC$, где k – удельный показатель поглощения растворенного вещества, рассчитанный на единичную концентрацию. При этом

$$D = ax = kCz, \quad (5.121)$$

т.е. оптическая плотность раствора пропорциональна его концентрации.

Другая причина, которая заставляет отдавать предпочтение оптической плотности, состоит в том, что, как следует из выражения (5.120), она пропорциональна толщине x поглощающего слоя, в то время как коэффициент пропускания слоя является показательной функцией его толщины. Поэтому рассчитывать оптическую плотность слоя во многих случаях гораздо проще.

Раздел II. ОСНОВЫ ПРИКЛАДНОЙ ОПТОТЕХНИКИ

1. Основы теории геометрического построения изображения предмета системой оптических поверхностей

1.1. Параксиальная оптика

Теоретические соотношения, полученные в предположении, что лучи, проходящие через оптически сопряженные точки в пространстве предметов и изображений, расположены в непосредственной близости от оптической оси, принято называть параксиальной оптикой (от греч. *para* – возле, мимо, вне и лат. *axis* – ось) или оптикой Гаусса.

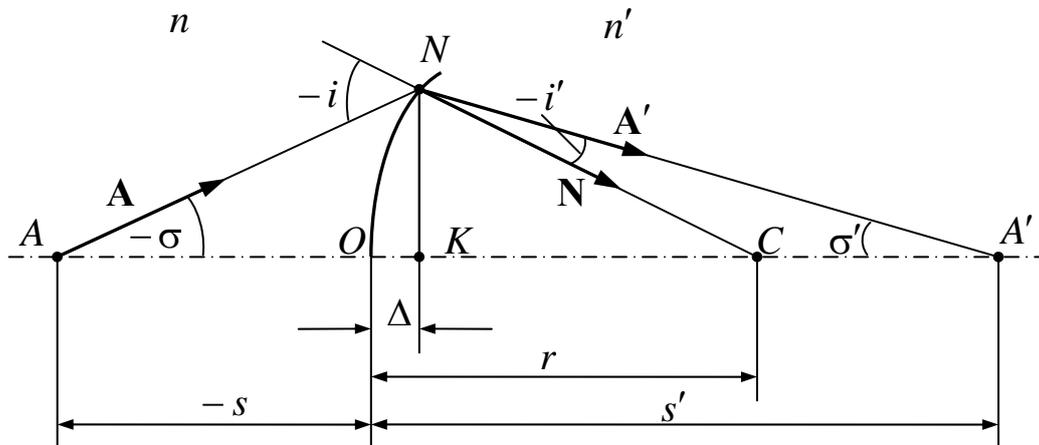


Рис.1.1. Преломление луча осевого пучка на сферической поверхности раздела двух однородных сред

Рассмотрим ход луча через одну преломляющую поверхность сферической формы, центр кривизны которой расположен в точке C , разделяющую однородные среды с показателями преломления n и n' , как показано на рис. 1.1. Следует заметить, что в случае одной преломляющей поверхности сферической формы любая прямая, проведенная через центр ее кривизны, может быть принята за оптическую ось. Однако, будем считать, что рассматриваемая преломляющая поверхность входит в состав некоторой центрированной системы оптических поверхностей, обладающей вполне определенным положением оптической оси, под которой понимаем линию, проходящую через центры кривизны всех поверхностей. Точку пересечения преломляющей поверхности с оптической осью (точку O) называют вершиной этой поверхности. Расстояние от вершины O поверхности до центра C ее кривизны равно радиусу r кривизны поверхности. Если центр кривизны поверхности расположен справа от ее вершины, то величина радиуса считается положительной, в противном случае – отрицательной.

Пусть в среде с показателем преломления n на оптической оси расположена некоторая светящаяся точка A . Рассмотрим ход луча, исходящего из точки A под углом $-\sigma$ к оптической оси и падающего на

преломляющую поверхность в точке N . Нормалью к преломляющей поверхности в точке падения луча служит прямая NC , а ее направление от точки падения определяется ортом \mathbf{N} . Орт нормали \mathbf{N} направлен из среды с показателем преломления n в среду с показателем преломления n' . Направление луча AN определяется ортом \mathbf{A} . Этот луч с нормалью к поверхности в точке падения образует угол падения $-i$. Преломленный в точке N поверхности луч проходит в среде с показателем преломления n' и пересекает оптическую ось в точке A' под углом σ' к ней. Направление луча NA' определяется ортом \mathbf{A}' . Этот луч с нормалью к поверхности в точке падения образует угол преломления $-i'$.

Угол падения i будем считать положительным, если для совмещения орта \mathbf{A} падающего луча с ортом нормали \mathbf{N} орт \mathbf{A} следует повернуть вокруг точки падения луча (вокруг точки N) против часовой стрелки; в противном случае угол будем считать отрицательным. Аналогично угол преломления луча i' будем считать положительным, если для совмещения орта \mathbf{A}' преломленного луча с ортом нормали \mathbf{N} орт \mathbf{A}' необходимо повернуть вокруг точки N против часовой стрелки; в противном случае – отрицательным.

Положительным направлением оптической оси системы поверхностей принято считать направление слева направо. При этом угол σ будем считать положительным, если для совмещения орта \mathbf{A} с оптической осью его следует повернуть вокруг точки A против часовой стрелки; в противном случае – отрицательным. Аналогично угол σ' будем считать положительным, если для совмещения орта \mathbf{A}' с оптической осью его следует повернуть вокруг точки A' против часовой стрелки, а если по часовой стрелке, то отрицательным.

Обозначим отрезки $OA = -s$ и $OA' = s'$, при этом начало отрезков s и s' будем считать расположенным в вершине O преломляющей поверхности. В соответствии с рис. 1.1 имеем

$$\frac{r}{\sin \sigma} = \frac{r-s}{\sin i}, \quad \frac{r}{\sin \sigma'} = \frac{r-s'}{\sin i'}.$$

Учитывая закон преломления в форме (I.2.46), из этих соотношений получаем

$$\frac{r-s'}{r-s} = \frac{\sin i' \sin \sigma}{\sin i \sin \sigma'} = \frac{n \sin \sigma}{n' \sin \sigma'}. \quad (1.1)$$

Из точки N на оптическую ось опустим перпендикуляр $NK = m$. Тогда в соответствии с рисунком имеем

$$\sin \sigma = -\frac{m}{NA} = -\frac{m}{\sqrt{m^2 + (\Delta - s)^2}},$$

$$\sin \sigma' = \frac{m}{NA'} = \frac{m}{\sqrt{m^2 + (s' - \Delta)^2}}.$$

При этом

$$\frac{\sin \sigma}{\sin \sigma'} = - \frac{\sqrt{m^2 + (s' - \Delta)^2}}{\sqrt{m^2 + (\Delta - s)^2}}. \quad (1.2)$$

Из рисунка следует, что $\Delta = r - \sqrt{r^2 - m^2}$. Вполне очевидно, что $\lim_{m \rightarrow 0} \Delta = 0$.

Будем считать, что в общем случае отрезки $s \neq 0$ и $s' \neq 0$. Тогда

$$\lim_{m \rightarrow 0} \sin \sigma = \lim_{m \rightarrow 0} \sigma = 0, \quad \lim_{m \rightarrow 0} \sin \sigma' = \lim_{m \rightarrow 0} \sigma' = 0.$$

При этом из соотношения (1.2) следует, что

$$\lim_{m \rightarrow 0} \frac{\sin \sigma}{\sin \sigma'} = - \lim_{m \rightarrow 0} \frac{\sqrt{m^2 + (s' - \Delta)^2}}{\sqrt{m^2 + (\Delta - s)^2}} = \frac{s'_0}{s_0}.$$

Полученный результат позволяет соотношение (1.1) представить в виде:

$$\frac{r - s'_0}{r - s_0} = \frac{ns'_0}{n's_0}. \quad (1.3)$$

Это соотношение легко преобразовать к виду:

$$n' \left(\frac{1}{s'_0} - \frac{1}{r} \right) = n \left(\frac{1}{s_0} - \frac{1}{r} \right). \quad (1.4)$$

В результате получили формулу, известную под названием формулы Аббе, хотя она была выведена еще Ньютоном.

Формула (1.4) получена в предположении, что луч, выходящий из осевой точки A , пересекает преломляющую поверхность на высоте $m \rightarrow 0$. Такой луч принято называть параксиальным лучом. Вполне очевидно, что при этом формула (1.4) справедлива для любой преломляющей поверхности вращения несферической формы, если радиус кривизны в вершине поверхности $r_0 \neq 0$.

Если осевая точка предмета расположена на бесконечно большом расстоянии от вершины O преломляющей поверхности, т.е. если $s_0 = \infty$, то осевая точка изображения A'_0 займет положение, называемое задним фокусом поверхности и обозначаемое через F'_0 , а отрезок s'_0 , называемый задним фокусным расстоянием f'_0 , в соответствии с формулой (1.4) определится соотношением

$$f'_0 = \frac{n'r}{n' - n}. \quad (1.5)$$

Вполне очевидно, что при $s'_0 = \infty$ точка A_0 займет положение переднего фокуса F_0 , при этом переднее фокусное расстояние f_0 определится соотношением

$$f_0 = -\frac{nr}{n' - n}. \quad (1.6)$$

Из выражений (1.5) и (1.6) следует, что

$$\frac{f'_0}{f_0} = -\frac{n'}{n}. \quad (1.7)$$

1.2. Оптика нулевых лучей

Восстановим в вершине O преломляющей поверхности перпендикуляр к оптической оси. На этом перпендикуляре, на произвольном расстоянии от оптической оси, равном h , обозначим точку P . Проведем из точки P в точки A_0 и A'_0 линии, которые с оптической осью образуют углы $-\alpha$ и α' , как показано на рис. 1.2, при этом отрезки OA_0 и OA'_0 соответственно равны $-s_0$ и s'_0 .

Перепишем формулу (1.4) в виде:

$$\frac{n'}{s'_0} - \frac{n}{s_0} = \frac{n' - n}{r}.$$

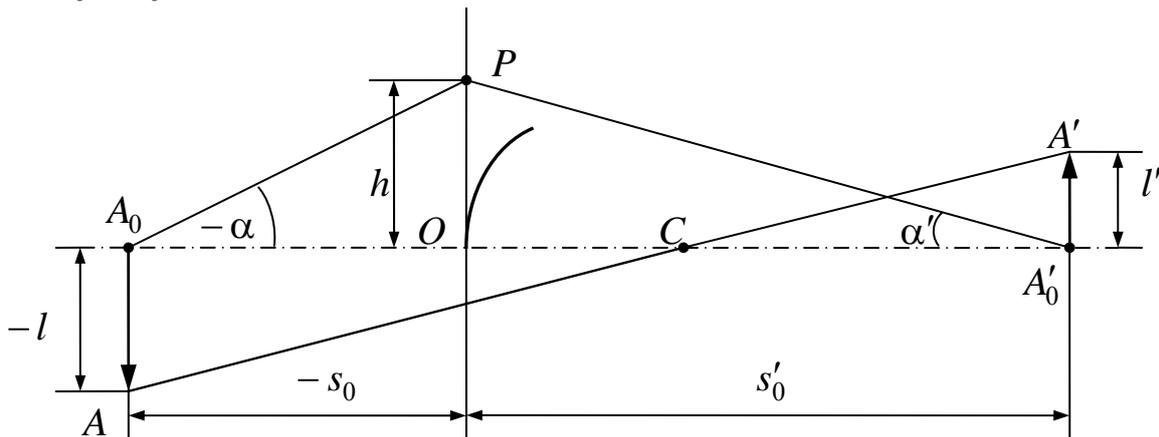


Рис. 1.2. Построение хода нулевого луча

Умножим левую и правую части этой формулы на h . Тогда в соответствии с рис. 1.2 получаем

$$n' \operatorname{tg} \alpha' - n \operatorname{tg} \alpha = h \frac{n' - n}{r}, \quad (1.8)$$

где $\operatorname{tg} \alpha = \frac{h}{s_0}$, $\operatorname{tg} \alpha' = \frac{h}{s'_0}$.

В вычислительной практике тангенсы углов принято обозначать просто углами. При этом полученное выражение можно представить в виде рекуррентной формулы для i -ой преломляющей поверхности:

$$n_{i+1} \alpha_{i+1} - n_i \alpha_i = h_i \frac{n_{i+1} - n_i}{r_i}. \quad (1.9)$$

Эта формула справедлива для любой преломляющей поверхности, при этом для $i + 1$ поверхности имеем

$$h_{i+1} = h_i - \alpha_{i+1}d_i, \quad (1.10)$$

где d_i – расстояние между вершинами i -ой и $i + 1$ -ой поверхностей.

Плоскость, перпендикулярную оптической оси и проходящую через вершину поверхности будем называть главной плоскостью преломляющей поверхности. Линию $A_0HA'_0$ называют первым нулевым лучом. По сути дела правильнее было бы называть этот луч осевым фиктивным (фр. fictive < лат. fictio – вымысел) или осевым виртуальным лучом.

В плоскости, перпендикулярной оптической оси и проходящей через точку A_0 , т.е. в плоскости предмета, отложим отрезок $A_0A = -l$, как показано на рис. 1.2. Луч, выходящий из точки A предмета, проходит через центр кривизны преломляющей поверхности без изменения направления и пересекает плоскость изображения в точке A' на расстоянии $A'_0A' = l'$ от оптической оси. При этом поперечное увеличение изображения, образованного преломляющей поверхностью, определяется отношением

$$V_0 = \frac{l'}{l} = \frac{s'_0 - r}{s_0 - r}. \quad (1.11)$$

Сопоставив левую и правую части соответственно выражений (1.3) и (1.11), получаем

$$V_0 = \frac{ns'_0}{n's_0}. \quad (1.12)$$

Но $\frac{s'_0}{s_0} = \frac{hs'_0}{hs_0} = \frac{\alpha}{\alpha'}$. Тогда из выражений (1.11) и (1.12) следует, что

$$nl\alpha = n'l'\alpha'.$$

Вполне очевидно, что это выражение справедливо для любой произвольной последовательности преломляющих поверхностей вращения. Записанное в виде:

$$n_i l_i \alpha_i = n_{i+1} l_{i+1} \alpha_{i+1}, \quad (1.13)$$

это выражение называется инвариантом Лагранжа-Гельмгольца.

Рассмотрим применение полученных соотношений для решения частных задач.

1.3. Кардинальные точки оптической системы

Оптическую систему из k сферических преломляющих поверхностей принято записывать в виде:

$$\begin{array}{r}
n_1 = \\
r_1 = \\
d_1 = \quad n_2 = \\
r_2 = \\
\text{-----} \\
r_{k-1} = \\
d_{k-1} = \quad n_k = \\
r_k = \quad n_{k+1} =
\end{array}$$

Из формулы (1.9) следует, что

$$\alpha_{i+1} = \frac{n_i}{n_{i+1}} \alpha_i + h_i \frac{n_{i+1} - n_i}{n_{i+1} r_i}. \quad (1.14)$$

Пусть расстояние от вершины первой поверхности до осевой точки предмета равно s_1 . Тогда высота $h_1 = s_1 \alpha_1$. Выбрав произвольно значение угла α_1 (или высоты h_1) и подставив в формулу (1.14), находим значение угла α_2 :

$$\alpha_2 = \frac{n_1}{n_2} \alpha_1 + \frac{n_2 - n_1}{n_2 r_1} s_1 \alpha_1.$$

Затем находим высоту нулевого луча на главной плоскости второй поверхности:

$$h_2 = h_1 - \alpha_2 d_1.$$

Полученные значения угла α_2 и высоты h_2 позволяют вычислить значения угла α_3 и высоты h_3 . Продолжая подобные вычисления, находим высоту h_k на главной плоскости последней в системе поверхности и угол α_{k+1} . Положение изображения, образованного рассматриваемой оптической системой, определится расстоянием s' от вершины k -ой поверхности до точки пересечения нулевого луча с оптической осью, равным

$$s' = \frac{h_k}{\alpha_{k+1}}.$$

Из инварианта (1.13) следует, что поперечное увеличение изображения, образованного системой оптических поверхностей, равно

$$V_0 = \frac{l_{k+1}}{l_1} = \frac{n_1 \alpha_1}{n_{k+1} \alpha_{k+1}}.$$

Таким образом, при известных значениях радиусов кривизны оптической системы, расстояний между вершинами поверхностей и показателей преломления разделяемых поверхностями сред расчет первого

нулевого (осевого фиктивного) луча позволяет определить положение изображения и его поперечное увеличение.

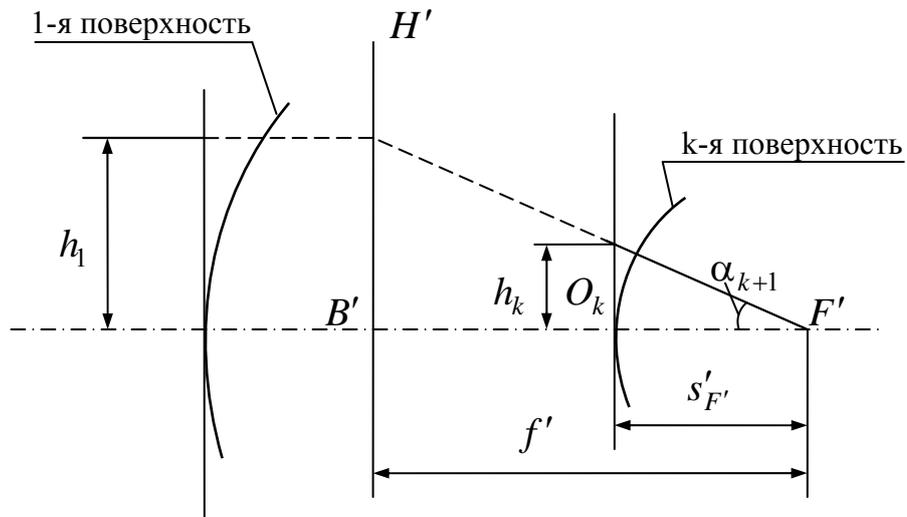


Рис. 1.3. Кардинальные точки оптической системы

Если предмет расположен на бесконечно большом расстоянии от оптической системы, то в этом случае угол $\alpha_1 = 0$. Выбрав при этом любое значение высоты h_1 , в результате расчета хода нулевого луча находим высоту h_k на главной плоскости последней поверхности и последний угол α_{k+1} , образованный нулевым лучом с осью, как показано на рис. 1.3. Точка пересечения луча с осью определяет положение изображения бесконечно удаленной осевой точки и называется задним фокусом F' оптической системы. Плоскость, перпендикулярная к оптической оси и проведенная через точку F' , называется задней фокальной плоскостью. Положение фокуса F' относительно последней поверхности системы определяется задним фокальным отрезком $s'_{F'}$, в соответствии с рис. 1.3 равным:

$$s'_{F'} = \frac{h_k}{\alpha_{k+1}}. \quad (1.15)$$

Плоскость H' , нормальная к оптической оси и содержащая точку пересечения падающего луча с продолжением выходящего, называется задней главной плоскостью; точка пересечения главной плоскости с оптической осью (точка B') называется задней главной точкой. При этом заднее фокусное расстояние, равное расстоянию от задней главной точки до заднего фокуса определяется соотношением:

$$f' = \frac{h_1}{\alpha_{k+1}}. \quad (1.16)$$

При некотором положении осевой точки предмета ее изображение, образованное оптической системой, расположено на бесконечно большом расстоянии. Эта точка F называется передним фокусом оптической

системы, а плоскость, нормальная к оптической оси и проходящая через передний фокус, называется передней фокальной плоскостью. Положение переднего фокуса относительно вершины первой поверхности определяется передним фокальным отрезком s_F .

Плоскость H , нормальная к оптической оси и содержащая точку пересечения продолжения луча, проходящего через передний фокус F оптической системы, с продолжением выходящего из оптической системы луча при $\alpha_{k+1} = 0$, называется передней главной плоскостью, а точка пересечения передней главной плоскости с оптической осью (точка B) называется передней главной точкой. Расстояние от передней главной точки до переднего фокуса называется передним фокусным расстоянием. Переднее фокусное расстояние и передний фокальный отрезок можно вычислить в результате расчета обратного хода нулевого луча.

Главные точки B и B' и фокусы F и F' называются кардинальными точками оптической системы.

1.4. Геометрическое построение изображения и основные оптические формулы

Обратимся к рис. 1.4, на котором оптическая система представлена главными плоскостями H и H' и, соответственно, главными точками B и B' , и определено положение фокусов F и F' . Кроме того, на рисунке показано некоторое положение предмета $l = A_0A$. Требуется построить изображение этого предмета. Для этого достаточно построить изображение A' одной лишь точки A . Опустив из точки A' перпендикуляр на оптическую ось, получим изображение A'_0A' всего отрезка A_0A . Чтобы найти изображение точки A , достаточно определить ход двух лучей, исходящих из точки A , и найти точку их пересечения в пространстве изображений. Один луч, исходящий из точки A , направим параллельно оптической оси. При этом выходящий из оптической системы луч должен пройти через задний фокус F' и пересечься с лучом падающим в точке P'_1 на задней главной плоскости на расстоянии $B'P'_1 = l$ от оптической оси. Таким образом, положение луча P'_1F' определено.

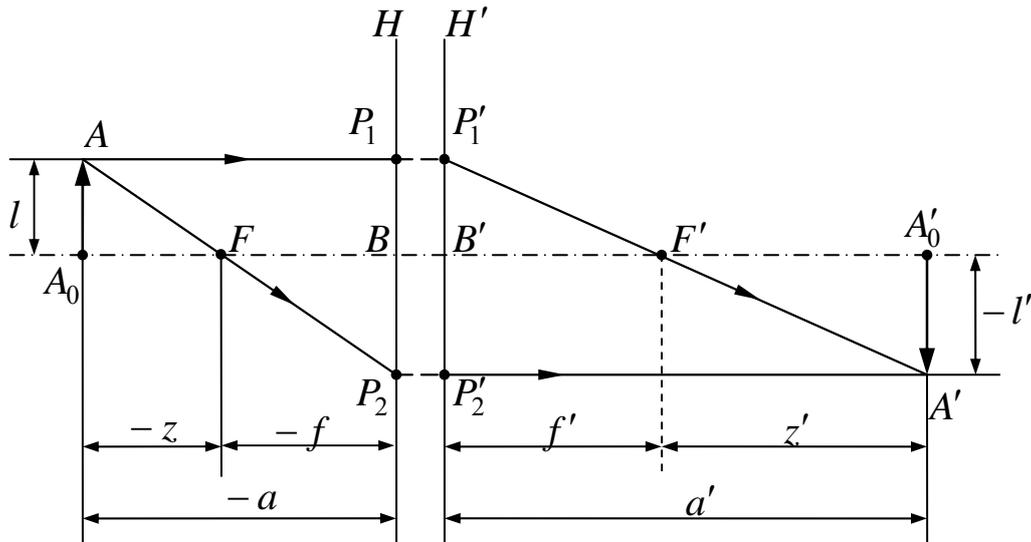


Рис. 1.4. Геометрическое построение изображения

Второй луч, исходящий из точки A , направим в передний фокус F оптической системы. Этот луч должен пересечься с выходящим из оптической системы параллельно ее оптической оси лучом в точке P_2 на передней главной плоскости на расстоянии $BP_2 = -l'$ от нее. Построенные в пространстве изображений лучи P_1F' и P_2A' пересекаются в точке A' , которая, следовательно, оптически сопряжена с точкой A . Перпендикуляр $A'A_0$, опущенный на оптическую ось из точки A' , определяет изображение отрезка A_0A .

Для получения основных аналитических соотношений, необходимых в практике расчета оптических систем, обратимся вновь к рис. 1.4. На рисунке введены обозначения: $l = A_0A$, $l' = -A'_0A'$, $f = -BF$, $f' = B'F'$. В оптике принято считать, что свет распространяется слева направо. Для отрезков, лежащих на оптической оси, принято следующее правило знаков: если направление отсчета отрезка совпадает с направлением движения света (слева направо), отрезок считается положительным, в противном случае – отрицательным.

Для отрезков, перпендикулярных к оптической оси, один конец которых лежит на оси, действует следующее правило знаков: если отрезок лежит выше оси, он положителен (например, отрезок A_0A), в противном случае он считается отрицательным (например, отрезок A'_0A').

Для последующего вывода необходимых соотношений введем отрезки на оптической оси, начала которых лежат в точках F и F' , а концы – в оптически сопряженных точках A_0 и A'_0 соответственно: $-z = FA_0$ и $z' = F'A'_0$. Из подобия треугольников FA_0A и $F'BP_2$ получаем:

$$-\frac{l'}{l} = \frac{-f}{-z}. \quad (1.17)$$

Отсюда следует, что поперечное увеличение изображения, образованного оптической системой, равно

$$V_0 = \frac{l'}{l} = -\frac{f}{z}. \quad (1.18)$$

Из подобия треугольников $F'A_0A'$ и $F'B'P_1'$ получаем

$$\frac{-l'}{l} = \frac{z'}{f'}. \quad (1.19)$$

В этом случае поперечное увеличение изображения равно

$$V_0 = \frac{l'}{l} = -\frac{z'}{f'}. \quad (1.20)$$

Приравнявая правые части соотношений (1.18) и (1.20), получаем формулу Ньютона в виде:

$$zz' = ff'. \quad (1.21)$$

Заметим, что положение главной точки B в оптической системе определяется отрезком $z = -f$. При этом в соответствии с формулой Ньютона отрезок $z' = -f'$, что соответствует положению главной точки B' , т.е. задняя главная плоскость является изображением передней главной плоскости и, следовательно, точки P_1 и P_1' и P_2 и P_2' являются оптически сопряженными. Кроме того, как следует из выражений (1.18) и (1.20), при $z = -f$ и $z' = -f'$ поперечное увеличение в изображении главных плоскостей $V_0 = 1^X$. Таким образом, произведение отрезков z и z' для оптической системы постоянно и равно произведению ее фокусных расстояний.

Введем отрезки $-a = BA_0$ и $a' = B'A_0'$. Начало этих отрезков считаем лежащими в точках B и B' . В соответствии с рисунком имеем

$$\left. \begin{aligned} z &= a - f, \\ z' &= a' - f'. \end{aligned} \right\} \quad (1.22)$$

Подставив эти значения величин z и z' в формулу Ньютона, получаем

$$f'a + fa' = aa'.$$

Поделив это выражение почленно на aa' , приводим его к виду:

$$\frac{f'}{a'} + \frac{f}{a} = 1. \quad (1.23)$$

Эту формулу называют формулой отрезков или оптической формулой. Используя равенства (1.22), находим:

$$\frac{a'}{a} = \frac{f' + z'}{f + z}. \quad (1.24)$$

Из формулы (1.21) следует, что

$$z' = \frac{ff'}{z}.$$

Подставив это значение z' в выражение (1.24), получаем

$$\frac{a'}{a} = \frac{f'}{z} \frac{z+f}{f+z} = \frac{f'}{z}. \quad (1.25)$$

В соответствии с формулой Ньютона имеем:

$$\frac{f'}{z} = \frac{z'}{f}. \quad (1.26)$$

При этом

$$\frac{a'}{a} = \frac{f'}{z} = \frac{z'}{f}. \quad (1.27)$$

Умножив эти соотношения на $-\frac{f}{f'}$ и учитывая соотношения (1.18) и

(1.20), получаем

$$V_0 = -\frac{f}{f'} \frac{a'}{a}. \quad (1.28)$$

Заметим, что $\frac{a'}{a} = \frac{ha'}{ha} = \frac{\alpha}{\alpha'}$. При этом

$$V_0 = \frac{l'}{l} = -\frac{f}{f'} \frac{\alpha}{\alpha'}. \quad (1.29)$$

Отсюда следует, что

$$fal = -f'\alpha'l'. \quad (1.30)$$

Применив инвариант Лагранжа-Гельмгольца в форме (1.13) последовательно ко всем поверхностям системы, получаем

$$n_1\alpha_1l_1 = n_{k+1}\alpha_{k+1}l_{k+1}$$

или

$$nal = n'\alpha'l'. \quad (1.31)$$

Сопоставив соотношения (1.29) и (1.30), имеем

$$\frac{f'}{f} = -\frac{n'}{n}. \quad (1.32)$$

Таким образом, для оптической системы в воздухе, т.е. при $n = n' = 1$, формулы (1.21), (1.23) и (1.29) принимают вид:

$$zz' = -f'^2 \quad (1.33)$$

$$\frac{1}{a'} - \frac{1}{a} = \frac{1}{f'} \quad (1.34)$$

$$V_0 = \frac{\alpha}{\alpha'}. \quad (1.35)$$

Умножив формулу отрезков (1.34) на h и учитывая, что $\frac{h}{a} = \alpha$, а $\frac{h'}{a'} = \alpha'$, получаем

$$\alpha' - \alpha = h\varphi, \quad (1.36)$$

где φ – оптическая сила системы поверхностей, при этом $\varphi = \frac{1}{f'}$.

2. Светотехника оптических систем

2.1. Световая трубка

Пусть $P(\xi, \eta)$ – произвольная точка на некоторой излучающей поверхности S , отнесенная к некоторой криволинейной системе координат на этой поверхности. При этом световой поток (усредненный по времени), излучаемый площадкой dS поверхности S в телесный угол $d\omega$ в направлении, определяемом полярными углами (α, β) , равен

$$d^2\Phi = L d\omega dS \cos \varepsilon, \quad (2.1)$$

где L – фотометрическая яркость излучения площадки dS в точке (ξ, η) в направлении (α, β) , т.е. в общем случае $L = L(\xi, \eta; \alpha, \beta)$; ε – угол между направлением (α, β) и нормалью к элементу поверхности, как показано на рис. 2.1. Множитель $\cos \varepsilon$ в выражении (2.1) определяет тот факт, что физический смысл имеет не сам элемент поверхности dS , а его проекция на плоскость, перпендикулярную к направлению (α, β) .

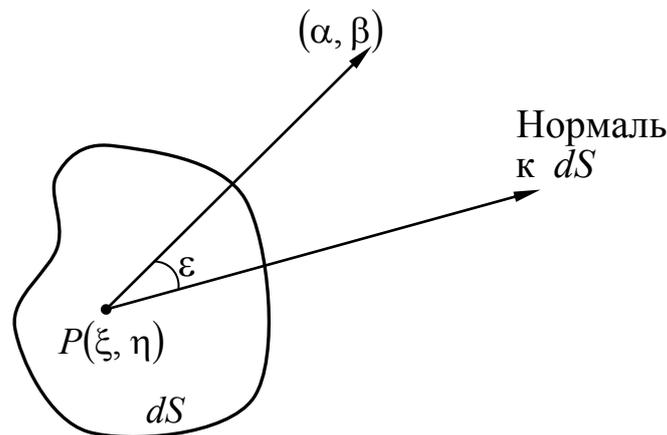


Рис. 2.1. Световой поток, излучаемый площадкой dS

Пусть сечение телесного угла $d\omega$ наклонной плоскостью образует площадку dS_v на некотором расстоянии R от площадки dS вдоль оси телесного угла $d\omega$, нормаль к которой образует угол ε_v , как показано на рис. 2.2. При этом

$$d^2\Phi = L \frac{dS_v \cos \varepsilon_v}{R^2} dS \cos \varepsilon = L d\omega_v dS_v \cos \varepsilon_v. \quad (2.2)$$

С другой стороны, световой поток, проходящий через площадку dS_v , равен

$$d^2\Phi_v = L_v d\omega_v dS_v \cos \varepsilon_v. \quad (2.3)$$

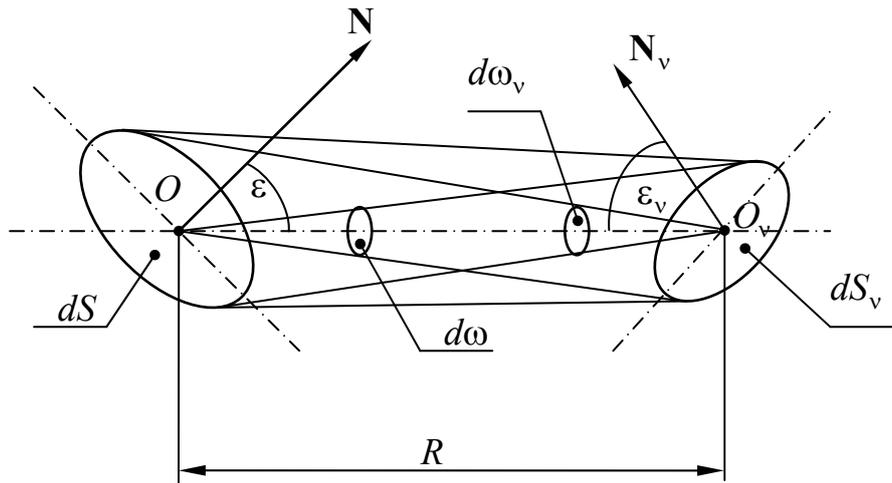


Рис. 2.2. Формирование световой трубки

В общем случае $d^2\Phi_v = \tau d^2\Phi$,

где τ – коэффициент пропускания среды, разделяющей площадки dS и dS_v .

Поскольку в рассматриваемом случае нас интересуют лишь соотношения геометрических величин, среду будем считать абсолютно прозрачной, что эквивалентно равенству $\tau = 1$. При этом из сопоставления выражений (2.2) и (2.3) следует, что $L = L_v$. Полученное равенство позволяет интерпретировать величину L не только как яркость излучающей поверхности, но и как яркость излучения в плоском сечении светового пучка.

Совокупность геометрических лучей, заполняющих телесный угол $d\omega$, образует гомоцентрический пучок лучей, исходящих из точки излучающего элемента, опирающийся на освещаемый элемент. Совокупность геометрических лучей, проходящих через две произвольно расположенные площадки (диафрагмы), размеры которых значительно меньше расстояния между ними, образует совокупность геометрических пучков лучей, называемую физическим пучком, при этом поверхность, ограничивающую поперечные размеры физического пучка, принято называть световой трубкой.

Из сопоставления выражений (2.1) и (2.2) следует, что

$$d\omega dS \cos \varepsilon = d\omega_v dS_v \cos \varepsilon_v. \quad (2.4)$$

Величину $d^2G = d\omega dS \cos \varepsilon$ называют геометрическим фактором пучка световых лучей. Инвариантность геометрического фактора d^2G относительно площадок dS и dS_v физического пучка, определяемая выражением (2.4), означает, что он является мерой множества геометрических лучей, составляющих физический пучок, не зависящей от того, какая из площадок является излучающей. Вполне очевидно, что геометрический фактор элементарного физического пучка, ограниченного

площадкой dS и площадкой dS_v , образованной сечением телесного угла $d\omega$ плоскостью не зависит от выбора расстояния R между ними, а, следовательно, не зависит от выбора расстояния R (при $\tau = 1$) и величина светового потока, проходящего через площадки dS и dS_v , контуры которых образуют контуры световой трубки. При этом яркость излучения в каждом сечении телесного угла плоскостью остается неизменной. Поэтому яркость $L_v = L$ можно считать яркостью элементарного физического пучка. Понятие яркости пучка лучей, введенное в трудах академика Владимира Александровича Фока, Андрея Александровича Гершуна, Сергея Осиповича Майзеля и других, оказалось весьма удобным при исследовании объемных источников излучения или рассеивающих излучение сред, например, неба, где исходное, и вполне естественное, определение яркости, строго говоря, вообще теряет смысл.

2.2. Инвариант Штраубеля

Пусть через некоторую элементарную площадку dS_{v1} в пределах телесного угла $d\omega_{v1}$ под углом ε_{v1} к нормали к площадке dS_{v1} проходит некоторый поток излучения. Будем считать, что площадка dS_v , образованная сечением телесного угла $d\omega_{v1}$ плоскостью, расположена на поверхности S_v , разделяющей среды, показатели преломления которых равны $n_v = n_{v1}$ и $n'_v = n_{v2}$. Ось телесного угла $d\omega_v$, вершина которого расположена в точке в пределах площадки dS_v , образует с нормалью к последней угол ε_v , а после преломления на границе раздела двух сред – угол ε'_v , как показано на рис. 2.3. В соответствии с законом преломления

$$n_v \sin \varepsilon_v = n'_v \sin \varepsilon'_v.$$

Возведя в квадрат обе части этого равенства и дифференцируя, получаем

$$n_v^2 \sin \varepsilon_v \cos \varepsilon_v d\varepsilon_v = n_v'^2 \sin \varepsilon'_v \cos \varepsilon'_v d\varepsilon'_v. \quad (2.5)$$

Построим элементарный телесный угол $d\omega_v$, ограничив его четырьмя гранями пирамиды (частный случай конической поверхности), а, именно, двумя меридиональными плоскостями XOM и XON , образующими элементарный двугранный угол $d\varphi$, и двумя плоскостями, проходящими через центр O сферы радиуса R и через элементы дуг малых кругов на поверхности сферы, как показано на рис. 2.4. Если элементарную площадку, вырезанную на поверхности сферы, принять в виду малости длин элементарных дуг меридиональных сечений, равных $Rd\varepsilon_v$, за прямоугольник, то площадь этого элемента поверхности определится произведением основания прямоугольника, длина которого равна

$R \sin \varepsilon_v d\varphi$, на его высоту, равную $R d\varepsilon_v$. При этом элементарный телесный угол $d\omega_v$, равный отношению $\frac{dS}{R^2}$, определится выражением

$$d\omega_v = \frac{dS}{R^2} = \frac{R \sin \varepsilon_v d\varphi R d\varepsilon_v}{R^2} = \sin \varepsilon_v d\varepsilon_v d\varphi. \quad (2.6)$$

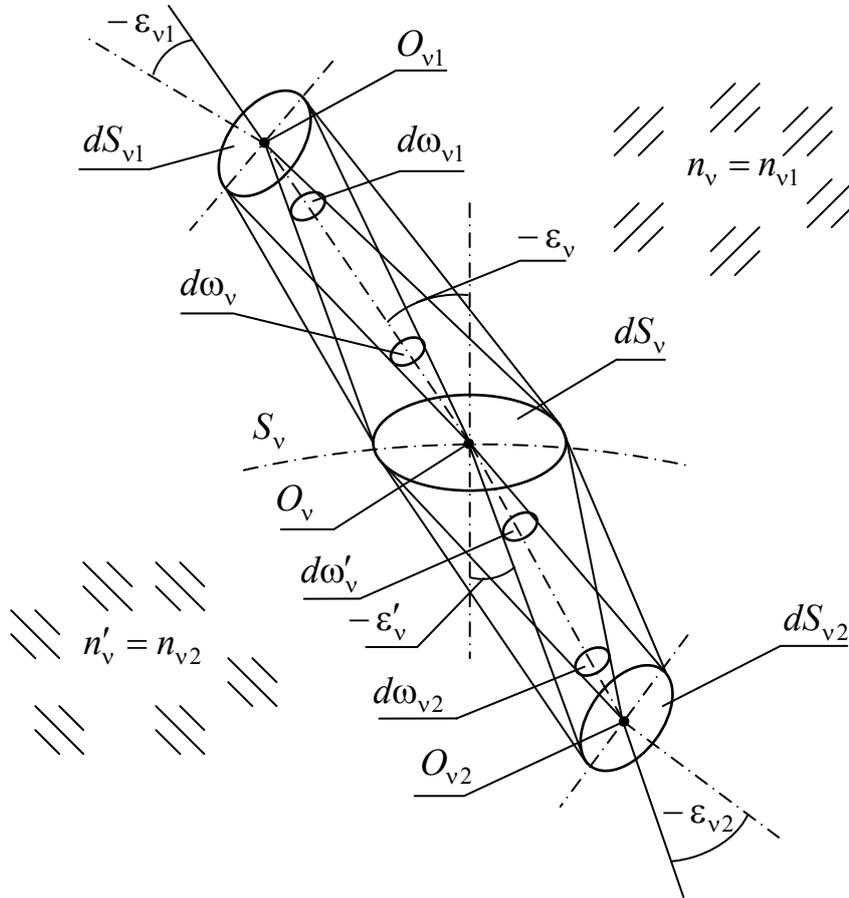


Рис. 2.3. Преломление световой трубки на границе раздела двух сред

После преломления пучка лучей на границе раздела двух сред телесный угол $d\omega'_v$ определится аналогичным выражением вида:

$$d\omega'_v = \sin \varepsilon'_v d\varepsilon'_v d\varphi, \quad (2.7)$$

где $d\varphi'$ равен углу $d\varphi$, так как согласно первой части формулировки закона преломления луч падающий, луч преломленный и нормаль к поверхности раздела двух сред в точке падения луча лежат в одной плоскости.

Это соотношение определяет основной инвариант для световой трубки: произведение квадрата показателя преломления, проекции элементарной площадки, контур которой определяет контур световой трубки, на плоскость, перпендикулярную оси трубки, и элементарного телесного угла с вершиной в точке этой площадки остается инвариантным для элементарной световой трубки, претерпевающей преломления на каком угодно числе преломляющих поверхностей:

$$\begin{aligned} n_1^2 d\omega_1 dS_1 \cos \varepsilon_1 = \dots = n_v^2 d\omega_v dS_v \cos \varepsilon_v = \dots = \\ = n_p^2 d\omega_p dS_p \cos \varepsilon_p = d^2 J. \end{aligned} \quad (2.12)$$

Эта формула была строго доказана в 1902 году Штраубелем и называется теоремой, или инвариантом, Штраубеля.

При абсолютной прозрачности сред, разделяемых преломляющими поверхностями, в соответствии с законом сохранения энергии световой поток, проходящий через световую трубку, претерпевающую какое угодно число преломлений, в соответствии с выражением (2.3) равен

$$\begin{aligned} d^2 \Phi = L_1 d\omega_1 dS_1 \cos \varepsilon_1 = \dots = L_v d\omega_v dS_v \cos \varepsilon_v = \dots = \\ = L_p d\omega_p dS_p \cos \varepsilon_p. \end{aligned}$$

Разделив эти равенства почленно на инвариант Штраубеля, получаем

$$\frac{L_1}{n_1^2} = \dots = \frac{L_v}{n_v^2} = \dots = \frac{L_p}{n_p^2} = L_0,$$

где L_0 – так называемая редуцированная (или приведенная к вакууму) яркость.

В соответствии с рис. 2.3 имеем

$$d\omega_v = \frac{dS_{v1} \cos \varepsilon_{v1}}{R_{v1}^2}; \quad d\omega'_v = \frac{dS_{v2} \cos \varepsilon_{v2}}{R_{v2}^2},$$

где R_{v1} и R_{v2} – расстояния от осевой точки O_v площадки dS_v до осевых точек O_{v1} и O_{v2} площадок dS_{v1} и dS_{v2} соответственно.

Эти соотношения позволяют выражение (2.9) представить в виде:

$$\frac{n_v^2 dS_v dS_{v1} \cos \varepsilon_v \cos \varepsilon_{v1}}{R_{v1}^2} = \frac{n_v'^2 dS_v dS_{v2} \cos \varepsilon'_v \cos \varepsilon_{v2}}{R_{v2}^2}. \quad (2.13)$$

Пусть в выражении (2.13) углы $\varepsilon_v = \varepsilon_{v1} = \varepsilon'_v = \varepsilon_{v2} = 0$, а площадки dS_v , dS_{v1} и dS_{v2} световых трубок имеют круглую форму. При этом площадь этих элементарных площадок равна

$$dS_{v1} = \pi(dl_{v1})^2; \quad dS_{v2} = \pi(dl_{v2})^2; \quad dS_v = \pi(dm_v)^2,$$

где dl_{v1} , dl_{v2} , dm_v – радиусы соответствующих площадок. Обозначим

$$\frac{dm_v}{R_{v1}} = d\sigma_{v1}; \quad \frac{dm_v}{R_{v2}} = d\sigma_{v2}.$$

Эти соотношения и принятые обозначения позволяют выражение (2.13) представить в виде:

$$n_v^2(dl_{v1})^2(d\sigma_{v1})^2 = n_v'^2(dl_{v2})^2(d\sigma_{v2})^2.$$

Применив полученное выражение к последовательности из p преломляющих поверхностей, получаем

$$n_1 d\sigma_1 dl_1 = n_2 d\sigma_2 dl_2 = \dots = n_v d\sigma_v dl_v = \dots = n_p d\sigma_p dl_p. \quad (2.14)$$

Для оптически сопряженных отрезков, т.е. в том случае, когда отрезок dl_{v+1} является изображением отрезка dl_v , это выражение называется инвариантом Лагранжа-Гельмгольца и записывается в виде

$$n\alpha l = n'\alpha' l', \quad (2.15)$$

где l , l' – оптически сопряженные малые отрезки, расположенные перпендикулярно оптической оси; α , α' – малые углы, образованные лучами светового пучка с оптической осью в пространстве предметов и изображений соответственно.

Отношение малых отрезков определяет величину поперечного увеличения:

$$V = \frac{l'}{l} = \frac{n\alpha}{n'\alpha'}. \quad (2.16)$$

Из малой величины поперечных размеров световой трубки следует, что инвариант Лагранжа-Гельмгольца справедлив для малых отрезков l и l' и для лучей, образующих малые углы α и α' с осью световой трубки, т.е. в параксиальной (от греч. *para'* – возле, мимо, вне и лат. *axis* – ось) области. Инвариант Лагранжа-Гельмгольца является одним из основных соотношений параксиальной оптики.

2.3. Ограничение световых пучков лучей в оптической системе

Любая оптическая система состоит из элементов, имеющих ограниченные размеры, вследствие чего из всего светового потока, излучаемого каждой изображаемой точкой (точнее, малым элементом поверхности), через оптическую систему проходит только часть его. Вместе с тем ограничена и величина предмета (угловая или линейная), из каждой точки которого световой поток попадает в оптическую систему. Первое ограничение определяет освещенность изображения, а второе – угловую или линейную величину изображаемого предмета, называемую угловым или линейным полем системы.

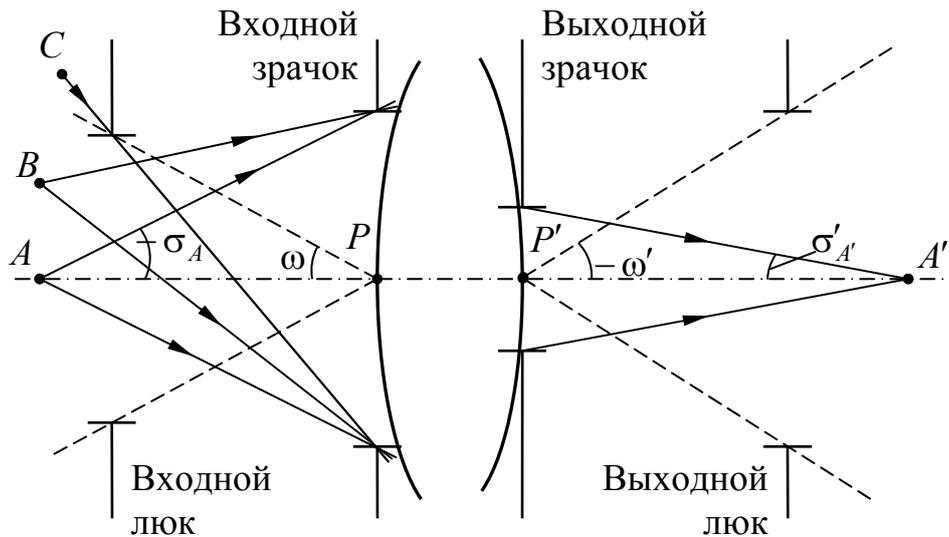


Рис. 2.5. Ограничение световых пучков лучей в оптической системе

Оптическая система, как правило, содержит последовательность диафрагм, ограничивающих световые пучки лучей, формирующих изображение. Под диафрагмой будем понимать не только отверстие в непрозрачном экране, но и края линз и зеркал или края их оправ. Будем считать, что диафрагмы имеют круглую форму с центром на оптической оси системы. Найдем параксиальное изображение каждой диафрагмы в пространстве предметов, образованное предшествующей ей оптической системой. Изображение диафрагмы (в частном случае непосредственно диафрагма), которое видно из осевой точки предмета A под наименьшим углом $2\sigma_A$, как показано на рис. 2.5, определяет максимальную угловую величину светового пучка лучей, проходящих через оптическую систему, и называется входным зрачком оптической системы. Реальная диафрагма, изображение которой (или сама диафрагма) является входным зрачком, называется апертурной диафрагмой. Изображение входного зрачка, образованное всей системой поверхностей в пространстве изображений, называется выходным зрачком оптической системы. Угол σ_A , образованный с оптической осью лучом, выходящим из осевой точки A предмета в край входного зрачка, называется апертурным углом в пространстве предметов или передней угловой апертурой. Соответствующий угол $\sigma'_{A'}$ называется апертурным углом в пространстве изображений или задней угловой апертурой. Произведение синуса апертурного угла на показатель преломления среды соответствующего пространства называется числовой апертурой (соответственно передней или задней). Числовая апертура определяет предельное разрешение в дифракционной картине изображения двух точек, образованного идеальной оптической системой. Изображение диафрагмы (в частном случае сама диафрагма), которое из центра входного зрачка P видно под

наименьшим углом 2ω , как показано на рис. 2.5, ограничивает величину изображаемого предмета и называется входным люком, при этом угол 2ω называется угловым полем оптической системы в пространстве предметов, а угол ω – углом поля. Реальная диафрагма, изображение которой является входным люком, называется полевой диафрагмой. Изображение входного люка, образованное оптической системой в целом в пространстве изображений, называется выходным люком.

Как следует из рисунка, верхний крайний луч наклонного светового пучка, выходящего из точки B в плоскости предмета, проходит через верхние крайние точки входного люка и входного зрачка. Нижний крайний луч наклонного светового пучка, выходящего из точки C в плоскости предмета, проходит через верхнюю крайнюю точку входного люка и нижнюю крайнюю точку входного зрачка, т.е. световой пучок полностью перекрывается экраном входного люка. Таким образом, для точек в плоскости предмета в промежутке между точками B и C происходит нарастающее срезание световых пучков от нулевого в точке B до полного в точке C . Такое срезание световых пучков называется виньетированием. Вполне очевидно, что если входной люк расположен в плоскости предмета (полевая диафрагма при этом может быть расположена в плоскости промежуточного или конечного изображения), то виньетирование световых пучков лучей отсутствует. Иногда при конструировании оптических устройств (при разработке оптических систем) эффект виньетирования применяют для устранения части пучка лучей, недопустимо нарушающих его гомоцентричность. Кроме того, с помощью дополнительных диафрагм можно существенно уменьшить влияние на качество изображения света, рассеянного на торцах линз и поверхностях оправ линз, на стенках корпусов оптических устройств и т.п.

Пусть $2y_A$ – линейное поле оптической системы в пространстве предметов, а $2y'_A$ – в пространстве изображений. При этом луч, выходящий из внеосевой точки предмета в направлении центра P входного зрачка, образует полевой угол ω , тангенс которого равен

$$\operatorname{tg} \omega = \frac{y}{a_p - a},$$

где a – расстояние от передней главной точки H оптической системы до осевой точки A предмета; a_p – расстояние от передней главной точки до осевой точки P (центра) входного зрачка.

В пространстве изображений имеем аналогичное соотношение вида:

$$\operatorname{tg} \omega' = \frac{y'}{a'_{p'} - a'}.$$

Угловое увеличение W изображения, образованного оптической системой, определяется отношением

$$W = \frac{tg\omega'}{tg\omega}.$$

Луч, выходящий из внеосевой точки предмета и проходящий через центр входного зрачка, называется главным лучом. Если апертурная диафрагма расположена в задней фокальной плоскости той части оптической системы, которая предшествует этой диафрагме, то входной зрачок будет находиться в бесконечности и все главные лучи в пространстве предметов будут параллельны оптической оси. Такая оптическая система называется системой с телецентрическим ходом главных лучей в пространстве предметов или телецентрической со стороны предмета. Если апертурная диафрагма расположена в передней фокальной плоскости той части системы, которая следует за этой диафрагмой, то выходной зрачок будет находиться в бесконечности и все главные лучи в пространстве изображений будут параллельны оптической оси. В этом случае оптическая система называется телецентрической со стороны изображения. Плоскость, в которой расположены оптическая ось и главный луч, называется меридиональной плоскостью. Плоскость, перпендикулярная к меридиональной плоскости по главному лучу (в которой расположен главный луч), называется сагиттальной плоскостью.

При анализе соотношений параксиальной оптики луч, проведенный из осевой точки предмета под произвольным углом α к оптической оси, будем называть нулевым лучом осевого пучка или просто осевым нулевым лучом; луч, проведенный из внеосевой точки предмета под углом β к оптической оси будем называть нулевым лучом внеосевого пучка или просто внеосевым нулевым лучом. Внеосевой нулевой луч, проведенный через центр входного зрачка, будем называть главным нулевым лучом.

Отношение тангенсов углов β'_p и β_p , образованных главным нулевым лучом с оптической осью в пространстве изображений и предметов соответственно, называется угловым увеличением изображения, образованного оптической системой, которое в принятых обозначениях в параксиальной оптике (в оптике Гаусса) определяется выражением

$$W = \frac{\beta'_p}{\beta_p}, \quad (2.17)$$

при этом поперечное (линейное) увеличение изображения входного зрачка (поперечное увеличение выходного зрачка) в соответствии с инвариантом Лагранжа-Гельмгольца определяется выражением

$$V_P = \frac{n\beta_p}{n'\beta'_p}. \quad (2.18)$$

Пусть точки пересечения главного нулевого луча с главными плоскостями оптической системы расположены на расстоянии H от

оптической оси. При $\beta_p = \frac{H}{a_p}$, а $\beta'_p = \frac{H}{a'_p}$ поперечное увеличение изображения входного зрачка равно

$$V_p = \frac{na'_p}{n'a_p}. \quad (2.19)$$

Как следует из выражений (2.17) и (2.18), взаимосвязь углового увеличения изображения предмета и поперечного увеличения изображения входного зрачка (т.е. выходного зрачка) определяется соотношением

$$W = \frac{n}{n'} \frac{1}{V_p}. \quad (2.20)$$

Осевые точки оптически сопряженных зрачков называются узловыми точками, если угловое увеличение изображения $W = 1^X$. Из выражения (2.20) следует, что если центр входного зрачка расположен в передней узловой точке, а центр выходного зрачка – в задней узловой точке, то поперечное увеличение в изображении зрачка $V_p = \frac{n}{n'}$. При этом в соответствии с выражением (2.19) расстояния от главных точек оптической системы до соответствующих узловых точек равны друг другу, т.е. $a_p = a'_p$. Используя формулу отрезков, получаем

$$a_p = a'_p = \frac{n' - n}{n'} f'. \quad (2.21)$$

Отсюда следует, что если показатель преломления среды в пространстве предметов и в пространстве изображений один и тот же ($n = n'$), то узловые точки совпадают с главными точками оптической системы.

Фокусы, главные и узловые точки называются кардинальными точками оптической системы. В 1841 году в труде «Диоптрические исследования» немецкий математик Карл Фридрих Гаусс (1777–1855) показал, что соотношения параксиальной оптики могут быть получены, если предположить пропорциональность малых углов падения и преломления лучей на преломляющую поверхность. Им впервые были введены понятия кардинальных точек центрированной оптической системы и получена формула отрезков, названная впоследствии его именем.

Заметим, что величины α и β введены в оптику в 1911 году в ту пору сотрудником фирмы Герца в Германии М. Ланге.

2.4. Световой поток, проходящий через оптическую систему от излучающего элемента поверхности, расположенного на оптической оси перпендикулярно к ней

Элементарный световой поток $d^2\Phi$, излучаемый элементом dS , расположенным на оптической оси перпендикулярно к ней, и проходящий через элементарную площадку $d\Sigma$ входного зрачка оптической системы в пределах телесного угла $d\omega$, как показано на рис. 2.6, определяется выражением (2.1) в виде:

$$d^2\Phi = L_\sigma d\omega dS \cos \sigma, \quad (2.22)$$

где L_σ – яркость излучающего элемента в направлении, образующем угол σ с оптической осью; $d\omega$ – элементарный телесный угол с вершиной в осевой точке площадки dS ; σ – угол между осью телесного угла $d\omega$ и оптической осью.

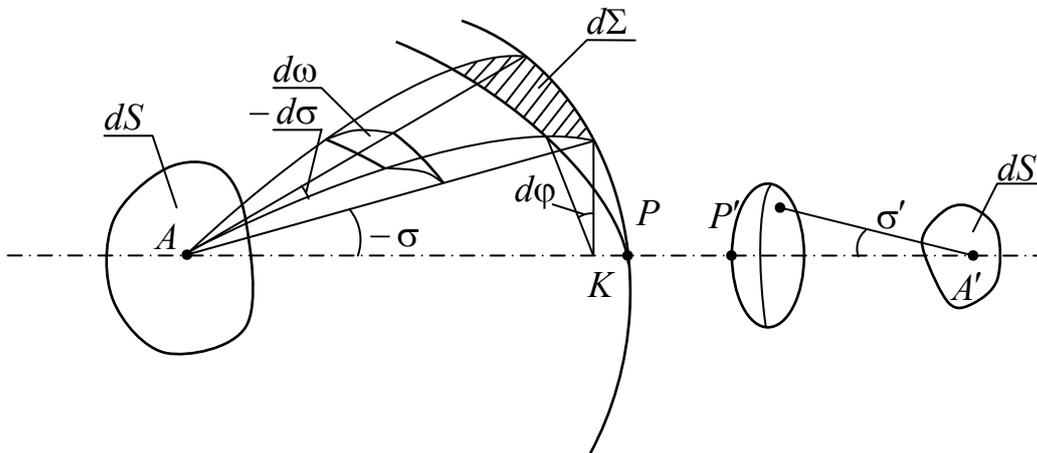


Рис. 2.6. Элементарный световой поток, излучаемый осевым элементом поверхности предмета

Обозначив отрезок $AP = -R$, в соответствии с рисунком получаем

$$d\omega = \frac{d\Sigma}{R^2} = \frac{R \sin \sigma d\varphi R d\sigma}{R^2} = \sin \sigma d\sigma d\varphi. \quad (2.23)$$

При этом

$$d^2\Phi = L_\sigma dS \sin \sigma \cos \sigma d\sigma d\varphi, \quad (2.24)$$

где $d\varphi$ – элементарный двугранный угол между двумя меридиональными (проходящими через оптическую ось) плоскостями, как показано на рис. 2.6, и составляющими боковые стенки телесного угла $d\omega$.

Для определения светового потока $d\Phi$, излучаемого элементарной площадкой dS и заполняющего весь входной зрачок оптической системы, проинтегрируем выражение (2.24) по всей площади входного зрачка, т.е. в пределах изменения переменных $0 \leq \varphi \leq 2\pi$ и при круглой форме зрачка $0 \leq \sigma \leq \sigma_K$ где σ_K – значение угла σ , соответствующего краю входного

зрачка, т.е. апертурный угол осевого пучка лучей в пространстве предметов.

Таким образом,

$$d\Phi = \int_0^{\sigma_K} \int_0^{2\pi} L_\sigma dS \sin \sigma \cos \sigma d\sigma d\varphi. \quad (2.25)$$

Задача определения потока $d\Phi$ существенно упрощается, если предположить, что яркость $L_\sigma = L = const$, т.е. является величиной постоянной во всех точках излучающего элемента dS и не зависит от угла излучения σ (при этом говорят, что элемент dS излучает по закону Ламберта). При этом условии выражение (2.25) можно переписать в виде:

$$d\Phi = L dS \int_0^{\sigma_K} \int_0^{2\pi} \sin \sigma \cos \sigma d\sigma d\varphi. \quad (2.26)$$

В результате интегрирования выражение (2.26) принимает вид:

$$d\Phi = \pi L dS \sin^2 \sigma. \quad (2.27)$$

Этот световой поток проходит сквозь оптическую систему и падает на элементарную площадку dS' , которая становится изображением площадки dS , как показано на рис. 2.6. Вполне очевидно, что световой поток $d\Phi'$, падающий на площадку dS' , определится аналогичным выражением вида:

$$d\Phi' = \pi L' dS' \sin^2 \sigma', \quad (2.28)$$

где σ' – апертурный угол осевого пучка лучей в пространстве изображений (или задний апертурный угол оптической системы).

Как было показано, $L = n^2 L_0$; $L' = n'^2 L_0$, где L_0 – редуцированная яркость; n и n' – показатели преломления сред в пространстве предметов и изображений соответственно.

Если принять во внимание тот факт, что в реальной оптической системе неизбежны потери светового потока (поглощение, френелево отражение на поверхностях), учитываемые коэффициентом пропускания τ ($\tau < 1$), то вместо светового потока $d\Phi$ из системы выходит поток $d\Phi' = \tau d\Phi$, а, следовательно, меньший потока $d\Phi$. При этом

$$d\Phi' = \tau \pi L' dS' \sin^2 \sigma', \quad (2.29)$$

где $L' = \left(\frac{n'}{n}\right)^2 L$. При $n' = n$: $L' = L$.

Условие синусов Аббе

Учитывая выражение (2.27), получаем

$$d\Phi' = \tau \pi L dS \sin^2 \sigma. \quad (2.30)$$

Приравняв правые части выражений (2.29) и (2.30), находим, что

$$n^2 dS \sin^2 \sigma = n'^2 dS' \sin^2 \sigma'. \quad (2.31)$$

Учитывая геометрическое подобие оптически сопряженных плоскостей фигур, одна из которых является предметом, а другая – его изображением, можно написать

$$\frac{dS'}{dS} = \left(\frac{dy'}{dy} \right)^2,$$

где dy и dy' – линейные отрезки оптически сопряженных элементов площадок dS и dS' . При этом выражение (2.31) можно переписать в виде:

$$n^2 dy^2 \sin^2 \sigma = n'^2 dy'^2 \sin^2 \sigma'.$$

Извлекая, квадратный корень из обеих частей этого равенства, получаем инвариант вида:

$$n dy \sin \sigma = n' dy' \sin \sigma'. \quad (2.32)$$

Отсюда находим, что

$$\frac{dy'}{dy} = V = \frac{n \sin \sigma}{n' \sin \sigma'}. \quad (2.33)$$

Это и есть известная формулировка закона синусов. Закон синусов впервые был получен немецким физиком, одним из основателей термодинамики, Рудольфом Юлиусом Эмануэлем Клаузиусом (1822–1888) в 1874 году и немецким физиком, математиком, физиологом и психологом Германом Людвигом Фердинандом Гельмгольцем (1821–1894) в 1874 году из термодинамических соображений. Однако, практическое значение этого условия было осознано лишь после того, как его сформулировал в 1879 году выдающийся немецкий физик – оптотехник Эрнст Аббе (1840–1905).

Из условия $V = V_0 = const$ следует так называемое условие синусов Аббе:

$$\frac{n \sin \sigma}{n' \sin \sigma'} = const. \quad (2.34)$$

Если в изображении A' осевой точки предмета A полностью устранена сферическая aberrация, то соблюдение условия (2.34) является необходимым и достаточным условием совершенного изображения элементарного отрезка dy , перпендикулярного оптической оси, а, следовательно, и всей элементарной площадки dS , проходящей через точку A и содержащей отрезок dy . При этом изображение формируют сколь угодно широкие пучки лучей, образующих углы σ и σ' с оптической осью любой величины.

Пусть расстояние от передней главной плоскости оптической системы до осевой точки предмета равно $-a$, а от задней главной плоскости до параксиального изображения этой точки равно a' . При этом поперечное увеличение изображения, образованного оптической системой пучками

лучей в параксиальной области, равно $V_0 = \frac{na'}{n'a}$. При соблюдении условия

синусов справедливо равенство

$$\frac{na'}{n'a} = \frac{n \sin \sigma}{n' \sin \sigma'}$$

Отсюда следует, что

$$a \sin \sigma = a' \sin \sigma'. \quad (2.35)$$

Но $a \sin \sigma = m$, а $a' \sin \sigma' = m'$, где m и m' – расстояния от оптической оси до точек пересечения крайних лучей, образующих углы σ и σ' с оптической осью, со сферами радиусов a и a' соответственно. Таким образом, применительно к реальным оптическим системам представления о главных плоскостях, справедливые в параксиальной области, для осевого пучка лучей в пределах конечного телесного угла должны быть заменены представлениями о главных сферах, понимая под последними сферы, которые проходят через главные точки оптической системы, а радиусы сфер равны a (передняя главная сфера) и a' (задняя главная сфера).

Заметим, что при $a \rightarrow \infty$ при $m = a \sin \sigma \neq 0$ угол $\sigma \rightarrow 0$. При этом расстояние a' стремится к величине f' , называемой задним фокусным расстоянием оптической системы. Как следует из выражения (2.35), при $a \rightarrow \infty$, а $\sigma \rightarrow 0$ имеем

$$f' = \frac{m}{\sin \sigma'}. \quad (2.36)$$

Условие синусов Аббе в этом случае определяется равенством $\frac{m}{\sin \sigma'} = \text{const}$. Следуя Аббе, назовем пару оптически сопряженных точек A и A' на оси оптической системы, в изображении которых отсутствует сферическая aberrация лучей широкого пучка и, кроме того, выполнено условие синусов, парой апланатических точек.

2.5. Освещенность в осевой точке изображения

При анализе свойств световых трубок по умолчанию предполагалось стигматичное отображение каждой точки элементарных площадок, при этом явления дифракции не учитывались. Результаты такого анализа дают практически важное представление об идеальном распределении светового потока в плоскости изображения. При этом освещенность в каждой точке изображения естественным образом определяется отношением элементарного светового потока к элементарной площадке, на которую он падает, т.е.

$$E_0 = \frac{d\Phi'}{dS'}$$

Используя формулу (2.24), находим, что освещенность E_0 в осевой точке изображения определяется выражением

$$E_0 = \tau \pi L' \sin^2 \sigma'. \quad (2.37)$$

Это выражение можно переписать в виде:

$$E_0 = \tau \pi L_0 A'^2. \quad (2.38)$$

Величину $A' = n' \sin \sigma'$ в оптике называют числовой апертурой со стороны изображения или задней числовой апертурой, при этом редуцированная яркость излучения (пучка лучей) $L_0 = \frac{L'}{n'^2} = \frac{L}{n^2}$. Пусть

числовая апертура светового пучка лучей в пространстве предметов (передняя числовая апертура) $A = n \sin \sigma$. Тогда, учитывая формулу (2.33), при соблюдении условия (2.34) получаем

$$E_0 = \frac{\tau \pi}{V_0^2} L_0 A^2, \quad (2.39)$$

где V_0 – поперечное увеличение изображения, образованного оптической системой.

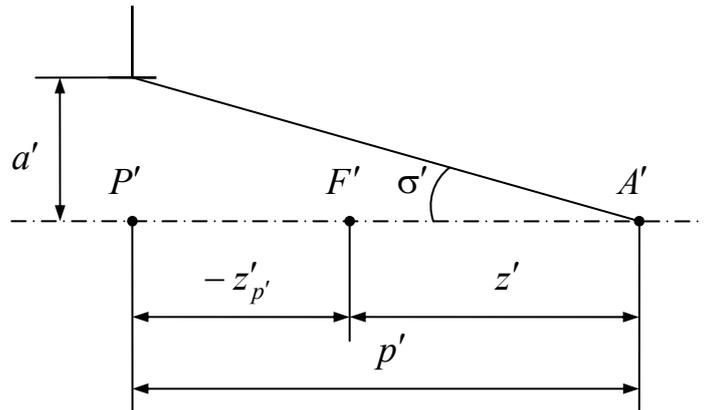


Рис. 2.7. Формирование задней угловой апертуры осевого пучка лучей

Пусть радиус выходного зрачка оптической системы равен a' , а расстояние от центра выходного зрачка до осевой точки изображения $P'A' = p'$, как показано на рис. 2.7. При этом в соответствии с рисунком имеем

$$\sin^2 \sigma' = \frac{a'^2}{a'^2 + p'^2} = \frac{a'^2}{p'^2} \frac{1}{1 + \frac{a'^2}{p'^2}}.$$

Подставив это соотношение в выражение (2.37), получаем

$$E_0 = \tau L' \frac{\pi a'^2}{p'^2} \frac{1}{1 + \frac{a'^2}{p'^2}}. \quad (2.40)$$

У большинства светосильных зеркальных и зеркально-линзовых систем выходной зрачок имеет форму кольца. Пусть a'_η – радиус экранируемой части зрачка. При этом освещенность в осевой точке изображения определится очевидным выражением вида:

$$E_0 = \tau\pi L'(\sin^2 \sigma'_H - \sin^2 \sigma'_{BH}), \quad (2.41)$$

где $\sin^2 \sigma'_H = \frac{a'^2}{a'^2 + p'^2}$; $\sin^2 \sigma'_{BH} = \frac{a'^2_\eta}{a'^2_\eta + p'^2}$.

Это выражение можно записать в виде

$$E_0 \approx \tau\pi L'(1 - \eta^2)\sin^2 \sigma', \quad (2.42)$$

где $\sigma' = \sigma'_H$; $\eta = \frac{a'_\eta}{a'}$ – коэффициент экранирования зрачка по диаметру (или коэффициент линейного экранирования).

Синус апертурного угла эквивалентной оптической системы, имеющей сплошной выходной зрачок круглой формы и создающей такую же освещенность в осевой точке изображения, равен

$$\sin \sigma'_{ЭКВ} = \sqrt{\sin^2 \sigma'_H - \sin^2 \sigma'_{BH}}.$$

В соответствии с рис. 2.7 имеем

$$p' = z' - z'_{p'} = f' \left(\frac{z'}{f'} - \frac{z'_{p'}}{f'} \right),$$

где f' – фокусное расстояние оптической системы.

При этом

$$p' = f'(-V_0 + V_{03P}), \quad (2.43)$$

где $V_0 = -\frac{z'}{f'}$ – поперечное увеличение изображения предмета,

образованного оптической системой, а $V_{03P} = -\frac{z'_{p'}}{f'}$ – поперечное

увеличение изображения входного зрачка.

Пусть $D_{3P} = 2a$ – диаметр входного зрачка, а $D'_{3P} = 2a'$ – диаметр выходного зрачка. При этом

$$|V_{03P}| = \frac{a'}{a}. \quad (2.44)$$

Заметим, что при $p' \geq 10a'$ с погрешностью, не превышающей 1%, освещенность в изображении осевой точки определяется выражением

$$E_0 = \tau L' \frac{\pi a'^2}{p'^2}. \quad (2.45)$$

Подставив в это выражение соотношения (2.43) и (2.44), получаем

$$E_0 = \frac{1}{4} \tau \pi L' \left(\frac{2a}{f'} \right)^2 \frac{V_{03P}^2}{(V_{03P} - V_0)^2}, \quad (2.46)$$

где $\frac{2a}{f'} = \frac{D_{3P}}{f'} = 2A$ – относительное отверстие оптической системы.

Величину $(2A)^2$ – называют геометрической светосилой оптической системы, а величину $\tau(2A)^2$ называют физической светосилой системы.

Если предмет расположен на весьма большом расстоянии от оптической системы, то величину z' можно считать достаточно малой, полагая $V_0 = -\frac{z'}{f'} \approx 0$. При этом формула (2.46) принимает вид:

$$E_0 = \frac{1}{4} \tau \pi L' \left(\frac{2a}{f'} \right)^2. \quad (2.47)$$

При увеличении изображения предмета $V_0 = -1^X$, а изображения входного зрачка $V_{03P} = 1^X$ формула (2.46) принимает вид:

$$E_0 = \frac{1}{16} \tau \pi L' \left(\frac{2a}{f'} \right)^2.$$

Формула Манжена-Чиколева

Рассмотрим важный в технике дальнего освещения (случай освещения прожектором) вопрос об освещенности бесконечно удаленного изображения.

Пусть источник света (предмет) расположен вблизи переднего фокуса F оптической системы. В этом случае его изображение окажется расположенным на весьма большом расстоянии p' от выходного зрачка оптической системы, как показано на рис. 5.8, где оптическая система представлена бесконечно тонким компонентом φ , а источник света представлен элементом $dS = \pi y^2$, все точки которого обладают одинаковой яркостью и излучают равномерно во всех направлениях.

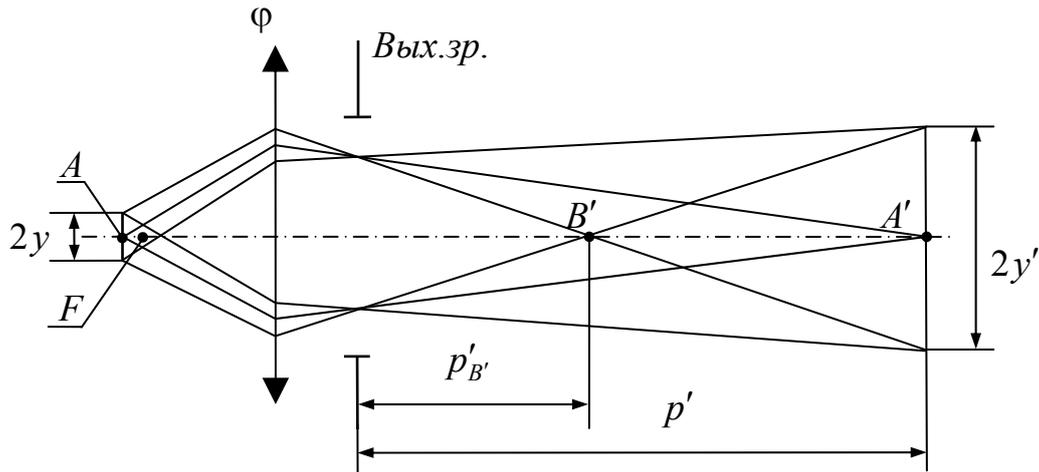


Рис. 2.8. Освещенность удаленного изображения

Освещенность E_0 в точке A' , расположенной на оптической оси системы, определится формулой (2.37):

$$E_0 = \tau \pi L' \sin^2 \sigma'.$$

В рассматриваемом случае с достаточной степенью приближения можно считать, что $\sin \sigma' = \frac{a'}{p'}$, где радиус выходного зрачка системы

$$a' = \frac{1}{2} D'_{зр}.$$

Пусть $S'_{зр}$ – площадь выходного зрачка оптической системы. При этом имеем

$$E_0 = \frac{\tau L' S'_{зр}}{p'^2}, \quad (2.48)$$

где $S'_{зр} = \pi a'^2$.

Сопоставив полученное соотношение с формулой (I.5.17): $dE_n = \frac{dI}{r^2}$, можно сделать вывод о том, что рассматриваемая оптическая система с источником излучения dS в отношении светового действия эквивалентна излучателю с осевой силой света

$$I_0 = \tilde{L}' S'_{зр},$$

где яркость $\tilde{L}' = \tau L' = \tau \left(\frac{n'}{n} \right)^2 L$; L – яркость источника света. При этом выражение (2.48) можно представить в виде:

$$E_0 = \frac{I_0}{p'^2}. \quad (2.49)$$

При $n' = n = 1$:

$$I_0 = \tau L S'_{зр}. \quad (2.50)$$

Полученную формулу, т.е. выражение (2.49), называют формулой Манжена-Чиколева.

Как следует из рис. 2.8, эта формула справедлива для всех точек на оптической оси в пространстве изображений, расположенных не ближе точки B' , так как для более близких точек действующей будет не вся световая поверхность выходного зрачка, а только часть ее.

В соответствии с рис. 2.8 имеем

$$\frac{2y'}{p' - p'_{B'}} = \frac{D'_{3P}}{p'_{B'}}$$

где D'_{3P} – диаметр выходного зрачка.

При этом

$$p'_{B'} \geq p' \frac{D'_{3P}}{D'_{3P} + 2y'}$$

Обозначим $2y = d$. При этом $2y' = \frac{z'}{f'} d$.

При $z' \approx p' - f'$ получаем

$$p'_{B'} \geq p' \frac{f' D'_{3P}}{f' D'_{3P} + (p' - f')d} \quad (2.51)$$

При $p' = \infty$:

$$p'_{B'(\infty)} \geq f' \frac{D'_{3P}}{d}$$

Формула Манжена-Чиколева может быть применена и в тех случаях, когда выходной зрачок имеет не только форму круга, но и более сложную форму, например, в проекционных системах, когда источник света в виде нитей или спирали лампы накаливания изображается конденсором в плоскость входного зрачка проекционного объектива.

Полученные соотношения позволяют весьма просто оценить эффективность применения оптической системы в составе осветительного устройства. Пусть источник излучения имеет площадь $dS_{ИСТ}$ и яркость L .

Его сила света $I_{ИСТ}$ в направлении нормали к поверхности будет равна

$$I_{ИСТ} = LdS_{ИСТ}$$

В соответствии с формулой (2.50) осевая сила света источника в сочетании с оптической системой равна

$$I_0 = \tau L S'_{3P}$$

Отношение рассматриваемых величин силы света называется коэффициентом усиления оптической системы:

$$m = \frac{I_0}{I_{ИСТ}} = \tau \frac{S'_{3P}}{dS_{ИСТ}} \quad (2.52)$$

площади $d\Sigma'$ одного из колец, ограниченный окружностями, радиусы которых равны r' и $r' + dr'$, и углом $d\varphi$ между двумя меридиональными плоскостями, одна из которых образует некоторый угол φ с осью Y . Площадь этого элемента зрачка равна $r'dr'd\varphi$. При этом световой поток, проходящий через элементарную световую трубку, образованную двумя площадками (сечениями) $d\Sigma' = r'dr'd\varphi$ и dS' , равен

$$d^2\Phi' = \frac{L'dS'r'dr'd\varphi \cos^2(q', Z)}{q'^2}, \quad (2.54)$$

где q' – расстояние от элемента $d\Sigma'$ кольца с координатами $(r' \sin \varphi; r' \cos \varphi; 0)$ до элемента dS' с координатами $(0; y'; p')$. В соответствии с рис. 2.9 имеем

$$\begin{aligned} \cos(q', Z) &= \frac{p'}{q'}; \\ q'^2 &= p'^2 + (y' - r' \cos \varphi)^2 + r'^2 \sin^2 \varphi = \\ &= p'^2 + y'^2 + r'^2 - 2r'y' \cos \varphi. \end{aligned} \quad (2.55)$$

Поток $d\Phi'$, проходящий через элемент изображения dS' , находим, выполняя интегрирование по всей площади выходного зрачка круглой формы радиуса a' :

$$d\Phi' = L'p'^2 dS' \int_0^{a'} r' dr' \int_0^{2\pi} \frac{d\varphi}{(p'^2 + y'^2 + r'^2 - 2r'y' \cos \varphi)^2}. \quad (2.56)$$

Представим это выражение в виде:

$$d\Phi' = L'p'^2 dS' \int_0^{a'} r' dr' \int_0^{2\pi} \frac{d\varphi}{(a + b \cos \varphi)^2}.$$

Первое интегрирование выполняем по формуле

$$\int \frac{d\varphi}{(a + b \cos \varphi)^2} = \frac{A \sin \varphi}{a + b \cos \varphi} + B \int \frac{d\varphi}{a + b \cos \varphi},$$

где $A = \frac{b}{b^2 - a^2}$ и $B = \frac{a}{a^2 - b^2}$. При этом

$$\int \frac{d\varphi}{a + b \cos \varphi} = \frac{2}{\sqrt{a^2 - b^2}} \operatorname{arctg} \left(\sqrt{\frac{a-b}{a+b}} \operatorname{tg} \frac{1}{2} \varphi \right), \text{ при условии, что } a > b.$$

Заметим, что

$$\begin{aligned} \int_0^{2\pi} \frac{d\varphi}{a + b \cos \varphi} &= \int_0^{\pi} \frac{d\varphi}{a + b \cos \varphi} + \int_{\pi}^{2\pi} \frac{d\varphi}{a + b \cos \varphi} = \\ &= \frac{2}{\sqrt{a^2 - b^2}} \left[\operatorname{arctg} \left(\sqrt{\frac{a-b}{a+b}} \operatorname{tg} \frac{1}{2} \varphi \right) \Big|_0^{\pi} + \operatorname{arctg} \left(\sqrt{\frac{a-b}{a+b}} \operatorname{tg} \frac{1}{2} \varphi \right) \Big|_{\pi}^{2\pi} \right] = \end{aligned}$$

$$= \frac{2}{\sqrt{a^2 - b^2}} [\operatorname{arctg} \infty - \operatorname{arctg} 0 + \operatorname{arctg} 0 - \operatorname{arctg}(-\infty)] =$$

$$= \frac{2}{\sqrt{a^2 - b^2}} \left[\frac{\pi}{2} - \left(-\frac{\pi}{2} \right) \right] = \frac{2\pi}{\sqrt{a^2 - b^2}}.$$

Кроме того, $A \frac{\sin \varphi}{a + b \cos \varphi} \Big|_0^{2\pi} = 0$. Тогда

$$\int_0^{2\pi} \frac{d\varphi}{(a + b \cos \varphi)^2} = \frac{a}{a^2 - b^2} \frac{2\pi}{\sqrt{a^2 - b^2}} = \frac{2\pi a}{(a^2 - b^2)^{3/2}}.$$

При этом выражение (2.56) принимает вид

$$d\Phi = 2\pi L' p'^2 dS' \int_0^{a'} \frac{(p'^2 + r'^2 + y'^2) r' dr'}{\left[(p'^2 + r'^2 + y'^2)^2 - 4y'^2 r'^2 \right]^{3/2}}. \quad (2.57)$$

Для выполнения интегрирования в выражении (2.57) введем новую переменную α' – угол между лучами $B_1 A'$ и $B_2 A'$, проведенными из точки A' к концам диаметра $B_1 B_2$ элементарного кольца в меридиональной плоскости. Как следует из рис. 2.9,

$$\alpha' = \beta_1 - \beta_2,$$

где β_1 и β_2 – углы с осью Y лучей $B_1 A'$ и $B_2 A'$.

Так как $\operatorname{tg} \beta_1 = \frac{p'}{y' - r'}$, а $\operatorname{tg} \beta_2 = \frac{p'}{y' + r'}$, то

$$\operatorname{tg} \alpha' = \frac{\operatorname{tg} \beta_1 - \operatorname{tg} \beta_2}{1 + \operatorname{tg} \beta_1 \operatorname{tg} \beta_2} = p' \frac{y' + r' - y' + r'}{y'^2 - r'^2} \left(1 + \frac{p'^2}{y'^2 - r'^2} \right)^{-1} = \frac{2p'r'}{p'^2 + y'^2 - r'^2};$$

$$\cos \alpha' = (1 + \operatorname{tg}^2 \alpha')^{-1/2} = \frac{p'^2 + y'^2 - r'^2}{\sqrt{(p'^2 + y'^2 - r'^2)^2 + 4p'^2 r'^2}} =$$

$$= \frac{p'^2 + y'^2 - r'^2}{\sqrt{(p'^2 + y'^2 + r'^2)^2 - 4r'^2(p'^2 + y'^2 + r'^2) + 4r'^4 + 4p'^2 r'^2}} =$$

$$= \frac{p'^2 + y'^2 - r'^2}{\sqrt{(p'^2 + y'^2 + r'^2)^2 - 4r'^2(r'^2 + p'^2) + 4r'^2(r'^2 + p'^2) - 4y'^2 r'^2}} =$$

$$= \frac{p'^2 + y'^2 - r'^2}{\sqrt{(p'^2 + y'^2 + r'^2)^2 - 4y'^2 r'^2}}.$$

Продифференцируем последнее выражение:

$$\begin{aligned}
-\sin \alpha d\alpha &= -\frac{2r'dr'}{\sqrt{(p'^2 + y'^2 + r'^2)^2 - 4y'^2 r'^2}} - \\
&= -\frac{(p'^2 + y'^2 + r'^2) \cdot 2(p'^2 + y'^2 + r'^2)2r' - 8y'^2 r'}{2[(p'^2 + y'^2 + r'^2)^2 - 4y'^2 r'^2]^{\frac{3}{2}}} dr' = \\
&= -\frac{2r'dr' \cdot (p'^2 + y'^2 + r'^2)^2 - 4y'^2 r'^2 + (p'^2 + y'^2 - r'^2)(p'^2 - y'^2 + r'^2)}{[(p'^2 + y'^2 + r'^2)^2 - 4y'^2 r'^2]^{\frac{3}{2}}} = \\
&= -\frac{2r'dr' \cdot (p'^2 + y'^2 + r'^2)(p'^2 + y'^2 + r'^2 + p'^2 + y'^2 - r'^2) - 4y'^2 r'^2}{[(p'^2 + y'^2 + r'^2)^2 - 4y'^2 r'^2]^{\frac{3}{2}}} - \\
&= -\frac{2y'^2(p'^2 + y'^2 - r'^2)}{[(p'^2 + y'^2 + r'^2)^2 - 4y'^2 r'^2]^{\frac{3}{2}}} = \\
&= -\frac{2(p'^2 + y'^2 + r'^2)(p'^2 + y'^2) - 2y'^2(p'^2 + y'^2 + r'^2)}{[(p'^2 + y'^2 + r'^2)^2 - 4y'^2 r'^2]^{\frac{3}{2}}} = \\
&= -4p'^2 \frac{p'^2 + y'^2 + r'^2}{[(p'^2 + y'^2 + r'^2)^2 - 4y'^2 r'^2]^{\frac{3}{2}}} r'dr'.
\end{aligned}$$

Итак,

$$\sin \alpha' d\alpha' = 4p'^2 \frac{p'^2 + y'^2 + r'^2}{[(p'^2 + y'^2 + r'^2)^2 - 4y'^2 r'^2]^{\frac{3}{2}}} r'dr'. \quad (2.58)$$

Сопоставив выражения (2.57) и (2.58), получаем

$$d\Phi' = \frac{1}{2} \pi L' dS' \int_0^{\alpha'_0} \sin \alpha' d\alpha', \quad (2.59)$$

где α_0 – угол между лучами, выходящими из крайних точек выходного зрачка, лежащих в меридиональной плоскости, и пересекающихся в точке A' в плоскости изображения. В результате интегрирования в выражении (2.59) получаем

$$d\Phi' = \frac{1}{2} \pi L' dS' (1 - \cos \alpha'_0) = \pi L' dS' \sin^2 \frac{1}{2} \alpha_0, \quad (2.60)$$

$$\text{где } \cos \alpha'_0 = \frac{p'^2 + y'^2 - a'^2}{\sqrt{(p'^2 + y'^2 + a'^2)^2 - 4y'^2 a'^2}}.$$

Заметим, что для точки A'_0 , расположенной на оптической оси системы, $y' = 0$, а, следовательно,

$\cos \alpha_0 = \frac{p'^2 - a'^2}{p'^2 + a'^2}$. При этом

$$d\Phi' = \frac{1}{2} \pi L' dS' \frac{p'^2 + a'^2 - p'^2 + a'^2}{p'^2 + a'^2} = \pi L' dS' \frac{a'^2}{p'^2 + a'^2} = \pi L' dS' \sin^2 \sigma',$$

где σ' – апертурный угол оптической системы в пространстве изображений. В этом случае формула (2.60) естественным образом переходит в формулу (2.28).

В реальной оптической системе выходящий световой поток $d\Phi'$ меньше входящего потока $d\Phi$ и равен $d\Phi' = \tau d\Phi$, где τ – коэффициент пропускания оптической системы.

2.7. Освещенность элемента изображения, расположенного вне оптической оси системы

Освещенность в некоторой точке, расположенной вне оптической оси в плоскости изображения, определяется отношением потока $d\Phi'$ к площади элемента изображения dS' , содержащего эту точку. В соответствии с формулой (2.60) это отношение равно:

$$E = \pi \tau L' \sin^2 \frac{1}{2} \alpha'_0. \quad (2.61)$$

Полагая заднюю апертуру наклонного пучка лучей сравнительно малой, что справедливо для подавляющего большинства применяемых систем, формулу (2.61) можно преобразовать к более удобному виду.

Действительно, если величина $\sin \frac{1}{2} \alpha'_0$ сравнительно невелика, то можно

принять, что

$$\sin \frac{1}{2} \alpha'_0 \approx \frac{a' \cos w'}{P'A'},$$

где в соответствии с рис. 2.9 $P'A'$ – отрезок, соединяющий центр выходного зрачка P' с рассматриваемой точкой A' ; a' – радиус кругового сечения наклонного пучка лучей, сходящихся в точке A' , плоскостью выходного зрачка.

В соответствии с рисунком $P'A' = \frac{p'}{\cos w'}$. При этом

$$\sin \frac{1}{2} \alpha'_0 \approx \frac{a'}{p'} \cos^2 w'.$$

Последнее соотношение позволяет выражение (2.61) представить в виде:

$$E = \pi \tau L' \frac{a'^2}{p'^2} \cos^4 w'. \quad (2.62)$$

Положив в формуле (6.37) $\sin \sigma' = \frac{a'}{p'}$, получаем

$$E_0 = \pi \tau L' \frac{a'^2}{p'^2}. \quad (2.63)$$

Из сопоставления соотношений (2.62) и (2.63) следует, что

$$E = E_0 \cos^4 w'. \quad (2.64)$$

При выводе формулы (2.64) предполагалось, что оптическая система образует безабберационное изображение предмета и апертурной диафрагмы. В случае реальной оптической системы формулу (2.63) удобно переписать в виде:

$$E_0 = \frac{d\Phi'}{dS'_0} = \tau \pi L' \frac{a_0'^2}{p'^2}, \quad (2.65)$$

а формулу (2.62) удобно представить в таком виде:

$$E = \frac{d\Phi'}{dS'_0} \frac{dS'_0}{dS'} = \frac{dS'_0}{dS'} \tau \pi L' \frac{a_0'^2}{p'^2} \cos^4 w'. \quad (2.66)$$

Из сопоставления формул (2.65) и (2.66) следует, что

$$E = \frac{a'^2}{a_0'^2} \frac{dS'_0}{dS'} E_0 \cos^4 w'. \quad (2.67)$$

Вполне очевидно, что aberrации в изображении апертурной диафрагмы, образованном следующей за ней частью оптической системы, определяют неравенство $\frac{a'^2}{a_0'^2} \neq 1$, а неравенство $\frac{dS'_0}{dS'} \neq 1$ определяется дисторсией в изображении предмета.

Назовем произведение рассматриваемых отношений aberrационным коэффициентом Русина (коэффициентом aberrационного виньетирования) и обозначим его через $P_{VC}(w')$:

$$P_{VC}(w') = \frac{a'^2}{a_0'^2} \frac{dS'_0}{dS'}. \quad (2.68)$$

При этом формулу (6.62) можно представить в виде:

$$E = P_{VC}(w') E_0 \cos^4 w'. \quad (2.69)$$

3. Структура изображения, образованного оптической системой, и оценка его качества

Распределение освещенности в плоскости изображения определяется выражением

$$E_0(x', y') = \tau \pi L'(x', y') \sin^2 \sigma' \cos^4 w', \quad (3.1)$$

где $L'(x', y')$ – распределение яркости излучения в направлении главных лучей, осредненной в пределах светового пучка лучей; τ – коэффициент пропускания света оптической системой и средами, разделяющими предмет, оптическую систему и плоскость изображений; σ' – угол, образованный крайним лучом осевого пучка с оптической осью в пространстве изображений; w' – угол (полевой угол) между главным лучом и оптической осью в пространстве изображений. Распределение яркости

$L'(x', y') = \frac{n'^2}{n^2} L(x, y)$, где $L(x, y)$ – распределение яркости излучения поверхностью предмета; n и n' – показатели преломления сред в пространстве предметов и изображений соответственно. Следовательно,

$$E_0(x', y') = \tau \pi L(x, y) \frac{n'^2}{n^2} \sin^2 \sigma' \cos^4 w'. \quad (3.2)$$

Поперечное увеличение изображения, образованного оптической системой, определяется соотношением вида:

$$V = \frac{n \sin \sigma}{n' \sin \sigma'},$$

где σ – угол, образованный крайним лучом, исходящим из осевой точки предмета, с оптической осью. При этом выражение (3.2) можно представить в виде:

$$E_0(x', y') = \tau \pi L(x, y) \frac{\sin^2 \sigma}{V^2} \cos^4 w'.$$

Пусть, например, $\sin w' = 0,1$, что соответствует углу $w' \approx 6^\circ$. Тогда $\cos^4 w' = (1 - \sin^2 w')^2 = 1 - 2 \sin^2 w' + \sin^4 w' = 0,9801$. Таким образом, при достаточно большом, например, для объективов зрительных труб, угловом поле изображения приближенно можно считать, что

$$E_0(x', y') = \tau \pi \frac{\sin^2 \sigma}{V^2} L(x, y) \quad (3.3)$$

или

$$E_0(x', y') = KL(x, y), \quad (3.4)$$

где $K = \tau \pi \frac{\sin^2 \sigma}{V^2} = const.$

Итак, при сравнительно малом, но достаточно большом, например, для объективов зрительной трубы, угловом поле распределение

освещенности на плоскости изображения идеально отображает распределение яркости излучения поверхностью предмета.

Важно отметить, что формула (3.1) получена для идеальной оптической системы на основе представлений геометрической оптики. Оптическую систему будем считать геометрически идеальной, если она удовлетворяет следующим условиям, сформулированным Д.К. Максвеллом в 1858 году:

– все лучи, вышедшие из точки предмета $A(x, y)$ и прошедшие через оптическую систему, должны сойтись (пересечься) в точке изображения $A'(x', y')$;

– каждый элемент плоскости предмета, нормальной к оптической оси и содержащей точку $A(x, y)$, должен быть изображен элементом плоскости, нормальной к оптической оси и содержащей точку $A'(x', y')$;

– высота изображения y' должна быть пропорциональна высоте предмета y , причем коэффициент пропорциональности должен быть постоянным независимо от местоположения точки $A(x, y)$ в плоскости предмета.

Отклонения от первого условия определяют поперечные aberrации широких и узких пучков лучей, образующих изображение; отклонения от второго условия определяют кривизну поверхности изображения, а отклонения от третьего условия нарушают подобие изображения предмету и называются дисторсией. Однако, в изображении, образованном геометрически идеальной оптической системой, распределение освещенности не является идеальным отображением яркости излучения поверхностью предмета, поскольку в этом случае каждая точка предмета изображается в виде дифракционного пятна, радиус первого темного кольца в котором (радиус кружка или пятна Эри) при круглой форме зрачков определяется формулой:

$$r'_д = \frac{1,22\lambda}{2n' \sin \sigma'}. \quad (3.5)$$

Вполне очевидно, что при прочих равных условиях размер дифракционного пятна в изображении точки тем меньше, чем меньше длина волны излучения λ . Этим, например, объясняется тот факт, что для повышения эффективности применения планарной технологии производства современной полупроводниковой электроники наблюдается постоянное стремление использовать излучение более коротких длин волн.

Из соотношения (3.5) следует, что только при $\lambda = 0$ безабберационная оптическая система формирует геометрически идеальное изображение. Если предположить, что величина светового потока, излучаемого точечным источником (точечным предметом), мала, но конечна, то вполне очевидно, что яркость излучения такого источника должна быть бесконечно большой, при этом бесконечно велика должна быть и

освещенность геометрически идеального изображения точки. Освещенность идеализированного изображения точки можно описать с помощью так называемой δ -функции Дирака, которая обладает следующими свойствами:

$$\delta(x' - x'_0, y' - y'_0) = \begin{cases} \infty, & \text{при } x' = x'_0; y' = y'_0; \\ 0, & \text{при } x' \neq x'_0 \text{ и (или) } y' \neq y'_0; \end{cases}$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x' - x'_0, y' - y'_0) dx' dy' = 1.$$

Отсюда следует, что если δ -функция описывает распределение освещенности в идеализированном изображении точки, то световой поток, формирующий это изображение, равен единице. С другой стороны, если яркость излучения точечного источника света равна нулю, то освещенность в изображении точки, а, следовательно, и поток, его формирующий, тоже будут равны нулю. Поэтому освещенность в идеализированном изображении точки определим произведением вида:

$$E_0(x', y') \delta(x' - x'_0, y' - y'_0).$$

Пусть $f(x', y')$ – любая функция переменных x' и y' . Если $f(x', y')$ непрерывна при $x' = x'_0$ и $y' = y'_0$, то справедливо следующее важное соотношение

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', y') \delta(x' - x'_0, y' - y'_0) dx' dy' = f(x'_0, y'_0),$$

т.е. справедливо соотношение

$$f(x', y') \delta(x' - x'_0, y' - y'_0) = f(x'_0, y'_0) \delta(x' - x'_0, y' - y'_0). \quad (3.7)$$

Свойство δ -функции, определяемое соотношением (3.7), называется фильтрующим свойством δ -функции.

Таким образом, справедливо выражение

$$E_0(x', y') \delta(x' - x'_0, y' - y'_0) = E_0(x'_0, y'_0) \delta(x' - x'_0, y' - y'_0).$$

Вполне очевидно, что световой поток, формирующий как реальное изображение точки (с учетом дифракции и аберраций), так и идеализированное, должен быть одним и тем же, равным входящему в оптическую систему с учетом коэффициента пропускания сред оптической системы.

Следовательно,

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(x' - x'_0, y' - y'_0) dx' dy' &= \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E_0(x', y') \delta(x' - x'_0, y' - y'_0) dx' dy' = E_0(x'_0, y'_0), \end{aligned} \quad (3.8)$$

где $E_0(x'_0, y'_0)$ – величина освещенности в изображении точки, определяемая формулой (3.1) или формулой (3.3).

Функцию $D(x' - x'_0, y' - y'_0)$, определяемую отношением

$$D(x' - x'_0, y' - y'_0) = \frac{E(x' - x'_0, y' - y'_0)}{E_0(x'_0, y'_0)},$$

называют функцией рассеяния точки (ФРТ). При этом, учитывая выражение (3.8), находим, что

$$\iint_{-\infty}^{\infty} D(x' - x'_0, y' - y'_0) dx' dy' = 1. \quad (3.9)$$

Таким образом, ФРТ можно интерпретировать как изображение точечного источника излучения, в котором распределен поток излучения, равный единице.

Заметим, что

$$E(x' - x'_0, y' - y'_0) = D(x' - x'_0, y' - y'_0) E_0(x'_0, y'_0), \quad (3.10)$$

где $E(x' - x'_0, y' - y'_0)$ – распределение освещенности в реальном изображении точки, образованном оптической системой.

Если ФРТ постоянна для всех точек плоскости изображения (абберации практически не изменяются в пределе всей плоскости изображения), то освещенность в рассматриваемой точке с координатами x', y' зависит только от расстояния, на котором она находится от точки идеализированного изображения с координатами x'_0, y'_0 . Это означает, что оптическая система удовлетворяет условию пространственной инвариантности, т.е. в соответствии с терминологией, принятой в оптике, она обладает изопланатической коррекцией абберацией или, просто, она изопланатична. В общем случае оптические системы этому условию не удовлетворяют. Однако, плоскость изображения всегда можно разделить на кольцевые зоны, концентричные осевой точке изображения, в пределах которых условие изопланатичности практически соблюдается, т.е. в пределах которых ФРТ изменяется в допустимых пределах.

Поскольку в изображении точки освещенность распределена на малой, но конечной площади, то вполне очевидно, что распределение освещенности в изображении других точек будет влиять на общее распределение освещенности в изображении этих точек. Будем считать излучаемый каждой точкой предмета свет некогерентным, при этом освещенность в каждой точке плоскости изображения равна сумме освещенностей от каждой изображаемой точки. В общем случае изображающая система линейна, если $E(x', y') = \sum_i a_i E_i(x', y')$.

3.1. Оптическая система как фильтр пространственных частот

В общем случае влияние оптической системы на структуру образованного ею изображения можно определить с помощью математического оператора, который показывает, как нужно преобразовать функцию $E_0(x'_0, y'_0)$ идеализированного изображения, чтобы получить функцию распределения освещенности $E(x', y')$ в действительном изображении. Такое воздействие оптической системы, удовлетворяющей требованиям линейности и изопланатичности, можно описать с помощью передаточной функции. Передаточная функция должна быть достаточно универсальна, чтобы ее можно было применить к системам различного назначения. Другое важное требование, предъявляемое к передаточной функции, состоит в том, чтобы ее можно было измерить.

Рассмотрим распределение освещенности на плоскости изображения вдоль линии, параллельной, например, оси x' . В этом случае $E_0(x'_0, y'_0) = E_0(x'_0, y'_0 = const)$. Если эту линию совместить с осью x' , то $y'_0 = 0$. При этом $E_0(x'_0, y'_0) = E_0(x'_0)$ и, соответственно,

$$E(x' - x'_0) = E(x'_0) D_{\text{л}}(x' - x'_0), \quad (3.11)$$

где $D_{\text{л}}(x' - x'_0)$ – функция рассеяния линии (ФРЛ), причем

$$D_{\text{л}}(x' - x'_0) = \int_{-\infty}^{\infty} D(x' - x'_0, y' - y'_0) dy'.$$

Условие нормировки ФРЛ таково, что

$$\int_{-\infty}^{\infty} D_{\text{л}}(x' - x'_0) dx' = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} D(x' - x'_0, y' - y'_0) dx' dy' = 1.$$

В случае изопланатической системы

$$D_{\text{л}}(x' - x'_{0i}) = D_{\text{л}}(x' - x'_{0j}).$$

Удобно обозначить $x' - x'_0 = \xi'$. При этом ФРЛ можно представить в виде, показанном на рис. 3.1.

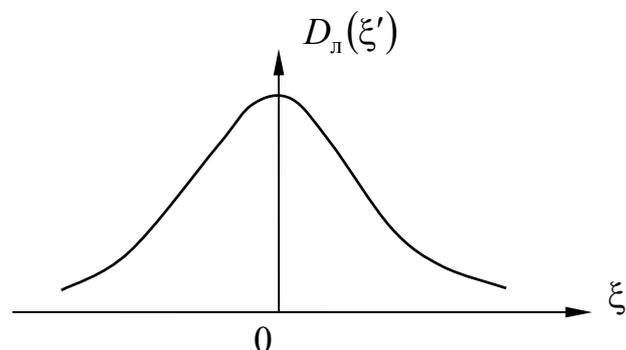


Рис. 3.1. Функция рассеяния линии

Обратимся к рис. 3.2, на котором распределение освещенности в идеализированном изображении в направлении оси x' представлено

кривой $E_0 = E_0(x')$. Определим освещенность изображения $E(x')$ в произвольной точке с координатой x' .

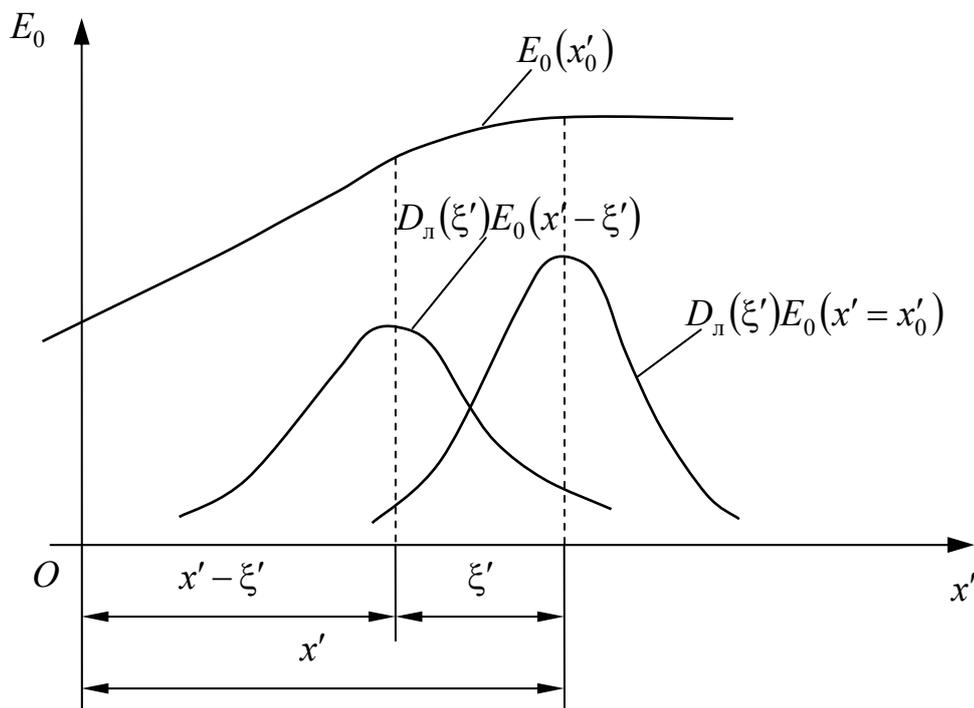


Рис. 3.2. К выводу свертки функций $D_{\text{л}}(\xi')$ и $E_0(x' - \xi')$

Для этого надо найти сумму ординат, равных $D_{\text{л}}(\xi')E_0(x' - \xi')$ при изменении координаты ξ' (отсчитываемой от точки с координатой x') теоретически от $-\infty$ до ∞ . При каждом значении координаты ξ' соответствующую ординату можно считать принадлежащей ФРЛ в изображении предмета с абсциссой $x'_0 = x' - \xi'$ и отстоящей от центра этой функции (т.е. от абсциссы x'_0) на расстоянии, равном ξ' , как показано на рис. 3.2. Таким образом, распределение освещенности в изображении предмета определяется выражением

$$E(x') = \int_{-\infty}^{\infty} D_{\text{л}}(\xi')E_0(x' - \xi')d\xi'. \quad (3.12)$$

Этот интеграл представляет собой так называемую свертку функции $E_0(x'_0)$ распределения освещенности вдоль оси x' в идеализированном изображении предмета, образованного оптической системой, с функцией рассеяния линии $D_{\text{л}}(\xi')$. Свертку функций часто обозначают значком \otimes . При этом $E = D_{\text{л}} \otimes E_0$.

В общем случае двумерного распределения освещенности $E_0(x'_0, y'_0)$ в идеализированном изображении предмета формула (3.12) принимает вид:

$$E(x', y') = \iint_{-\infty}^{\infty} D(\xi', \eta')E_0(x' - \xi', y' - \eta')d\xi'd\eta', \quad (3.13)$$

где ξ', η' – координаты ФРТ.

Можно показать, что линейный процесс, определяемый интегралом (3.13), на языке преобразования Фурье записывается в чрезвычайно простой форме, что приводит к интересным и полезным соображениям о действии оптической системы как фильтра пространственных частот.

Уместно напомнить, что формула (3.12), а, соответственно, и формула (3.13), получены в предположении, что вид функции $D(\xi', \eta')$ в пределах рассматриваемых участков изображения не изменяется, т.е. в пределах рассматриваемых участков изображения оптическая система обладает изопланатической коррекцией аберраций, при этом каждая точка предмета излучает некогерентный свет.

Пусть некоторая функция $f(x)$ задана на всей вещественной оси. При этом будем считать, что $f(x)$ в любом конечном интервале подчиняется условиям Дирихле и является абсолютно интегрируемой, т.е. $\int_{-\infty}^{\infty} |f(x)| dx < \infty$.

Такую функцию можно представить интегралом

$$f(x) = \int_{-\infty}^{\infty} g(N) \exp(i2\pi Nx) dN, \quad (3.14)$$

где N – пространственная частота. При этом функция $g(N)$ определяется через $f(x)$ формулой

$$g(N) = \int_{-\infty}^{\infty} f(x) \exp(-i2\pi Nx) dx. \quad (3.15)$$

Выражение (3.15) определяет прямое преобразование Фурье функции $f(x)$ и записывается в виде:

$$F[f(x)] = g(N) = \int_{-\infty}^{\infty} f(x) \exp(-i2\pi Nx) dx, \quad (3.16)$$

а обратное преобразование Фурье записывается в таком виде:

$$F^{-1}[g(N)] = f(x) = \int_{-\infty}^{\infty} g(N) \exp(i2\pi Nx) dN. \quad (3.17)$$

Это выражение можно рассматривать как представление функции $f(x)$ в виде линейной комбинации (т.е. интеграла) элементарных функций характерного вида $\exp(i2\pi Nx)$. Вполне очевидно, что комплексное число $g(N)$ представляет собой просто весовой коэффициент, на который следует умножить элементарную функцию частоты N при синтезе искомой функции $f(x)$.

Преобразование Фурье функции двух переменных $f(x, y)$ определяется формулой вида:

$$\begin{aligned}
 F[f(x, y)] &= g(N_x, N_y) = \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \exp[-i2\pi(N_x x + N_y y)] dx dy,
 \end{aligned}
 \tag{3.18}$$

при этом обратное преобразование Фурье записывается в виде:

$$\begin{aligned}
 F^{-1}[g(N_x, N_y)] &= f(x, y) = \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(N_x, N_y) \exp[i2\pi(N_x x + N_y y)] dN_x dN_y.
 \end{aligned}
 \tag{3.19}$$

Таким образом, точно так же, как одномерное, двумерное преобразование Фурье можно рассматривать как представление функции $f(x, y)$ в виде линейной комбинации элементарных функций вида $\exp[i2\pi(N_x x + N_y y)]$. Из определения преобразования Фурье следует, что преобразование Фурье свертки двух функций полностью эквивалентно более простой операции умножения преобразований Фурье свертываемых функций, т.е. их образов.

В соответствии с теоремой свертки преобразование Фурье распределения освещенности $E(x', y')$, определяемого формулой (3.13), равно произведению преобразований Фурье ФРТ $D(\xi', \eta')$ и распределения освещенности $E_0(x'_0, y'_0)$ в идеализированном изображении предмета:

$$e(N_x, N_y) = d(N_x, N_y) e_0(N_x, N_y).
 \tag{3.20}$$

Формула (3.20) определяет механизм образования изображения: каждой составляющей пространственного спектра $e_0(N_x, N_y)$ распределения освещенности при идеализированном изображении предмета соответствует коэффициент $d(N_x, N_y)$, который можно назвать «множителем передачи» или «множителем контраста» оптической системы. Таким образом, функция $d(N_x, N_y)$, характеризующая распределение освещенности в изображении точки, образованном оптической системой, по сути дела выполняет роль фильтра пространственных частот.

Функция $d(N_x, N_y)$, определяющая то, каким образом каждая частотная составляющая передается оптической системой с учетом дифракции, остаточных aberrаций и ошибок изготовления оптической системы, называется оптической передаточной функцией (ОПФ) системы.

ОПФ, являющаяся преобразованием Фурье ФРТ, представляет собой в общем случае комплексную функцию, которую можно записать в показательной форме через модуль комплексной функции $T(N)$ и аргумент $\varphi(N)$:

$$d(N_x, N_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} D(x', y') \exp[-i2\pi(N_x x' + N_y y')] dx' dy' =$$

$$= T(N_x, N_y) \exp[-i\varphi(N_x, N_y)]. \quad (3.21)$$

В случае одномерного распределения освещенности в идеализированном изображении предмета (этот случай чаще всего встречается при измерении ОПФ, когда предмет представляет собой решетку из параллельных штрихов, узкую щель и т.п.) ОПФ становится функцией одного аргумента – пространственной частоты N , например, вдоль оси x'

$$d(N) = \int_{-\infty}^{\infty} D_{\perp}(x') \exp[-i2\pi N x'] dx'.$$

Используя формулу Эйлера, получаем

$$d(N) = \int_{-\infty}^{\infty} D_{\perp}(x') \cos 2\pi N x' dx' - i \int_{-\infty}^{\infty} D_{\perp}(x') \sin 2\pi N x' dx' =$$

$$= C(N) - iS(N), \quad (3.22)$$

где интегралы, обозначенные через C и S , называются косинус-преобразованием и синус-преобразованием функции рассеяния линии. Модуль комплексной функции по определению равен

$$T(N) = \sqrt{[C(N)]^2 + [S(N)]^2} =$$

$$= \sqrt{\left[\int_{-\infty}^{\infty} D_{\perp}(x') \cos 2\pi N x' dx' \right]^2 + \left[\int_{-\infty}^{\infty} D_{\perp}(x') \sin 2\pi N x' dx' \right]^2}, \quad (3.23)$$

а аргумент $\varphi(N)$ может быть определен из соотношений:

$$\sin \varphi(N) = \frac{S(N)}{T(N)}, \quad (3.24)$$

$$\cos \varphi(N) = \frac{C(N)}{T(N)}. \quad (3.25)$$

Для лучшего понимания смысла модуля $T(N)$ и аргумента $\varphi(N)$ комплексной ОПФ рассмотрим более подробно, как изображается отдельная частотная составляющая спектра идеализированного изображения предмета оптической системой, имеющей известную функцию рассеяния.

Пусть распределение освещенности в идеализированном изображении предмета (или в его частотной составляющей) определяется косинусоидой с постоянной составляющей E_0 , как показано на рис. 3.3а, при этом

$$E_0(x') = E_0 + E_1 \cos 2\pi N x', \quad (3.26)$$

где N – пространственная частота; заметим, что период $\tilde{p} = \frac{1}{N}$.

Пусть $D_{\pi}(\xi')$ – функция рассеяния линии.

Тогда

$$E(x') = \int_{-\infty}^{\infty} D_{\pi}(\xi') E_0(x' - \xi') d\xi' = E_0 + E_1 \left[\cos 2\pi N x' \int_{-\infty}^{\infty} D_{\pi}(\xi') \cos 2\pi N \xi' d\xi' + \right. \\ \left. + \sin 2\pi N x' \int_{-\infty}^{\infty} D_{\pi}(\xi') \sin 2\pi N \xi' d\xi' \right] = E_0 + E_1 [C(N) \cos 2\pi N x' + \\ + S(N) \sin 2\pi N x'] = E_0 + E_1 T(N) [\cos \varphi(N) \cos 2\pi N x' + \sin \varphi(N) \sin 2\pi N x'].$$

В результате получаем, что

$$E(x') = E_0 + E_1 T(N) \cos[2\pi N x' - \varphi(N)]. \quad (3.27)$$

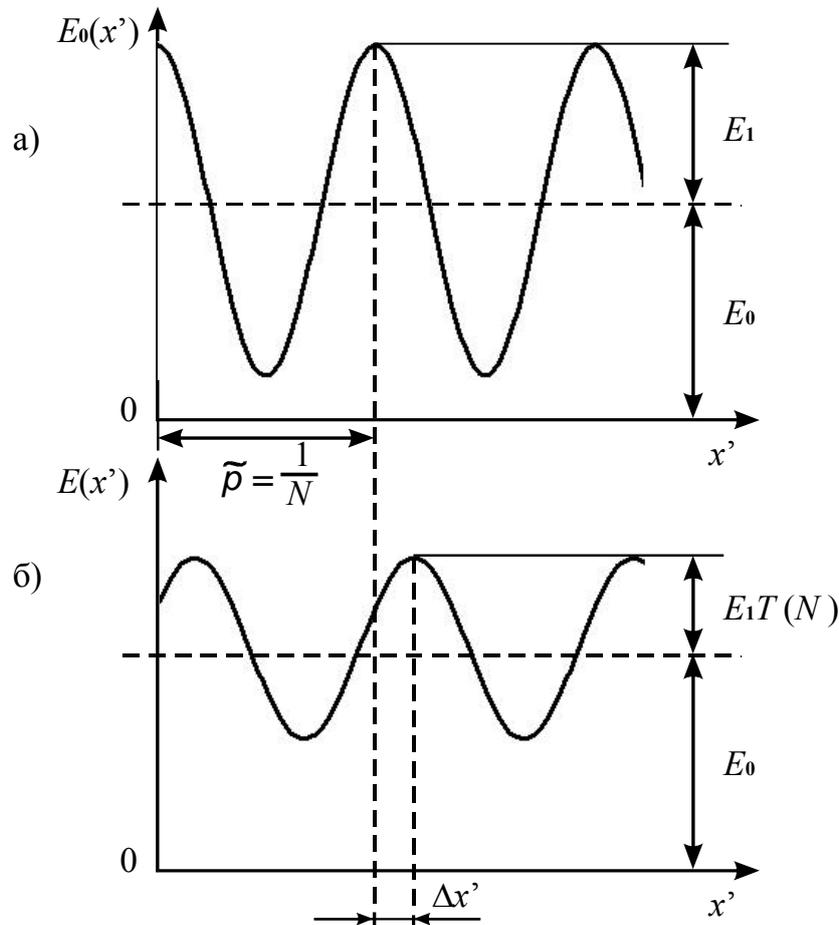


Рис. 3.3. Распределение освещенности в идеализированном (а) и действительном (б) изображениях

Отсюда следует, что полученное распределение освещенности в изображении предмета при косинусоидальном распределении освещенности в его идеализированном изображении остается косинусоидальным и той же пространственной частоты. Однако, модуляция, равная отношению амплитуды переменной составляющей распределения освещенности к среднему ее значению (т.е. к постоянной составляющей) для полученного распределения освещенности оказывается

меньше, чем для исходного, как показано на рис. 3.3б. Действительно, для идеализированного изображения имеем

$$m = \frac{E_1}{E_0}, \quad (3.28)$$

а при действительном распределении освещенности величина

$$m' = \frac{E_1}{E_0} T(N), \quad (3.29)$$

$$\text{т.е. } T(N) = \frac{m'}{m}. \quad (3.30)$$

Таким образом, значение модуля ОПФ, т.е. значение $T(N)$, для каждой пространственной частоты равно отношению модуляции распределения освещенности в действительном изображении к модуляции распределения освещенности в идеализированном изображении гармонической составляющей соответствующей частоты и называется коэффициентом передачи модуляции (КПМ) системы. Совокупность значений КПМ для различных пространственных частот составляет функцию передачи модуляции (ФПМ) или частотно-контрастную характеристику оптической системы.

Распределение освещенности в действительном изображении отличается от распределения освещенности в идеализированном изображении еще и сдвигом косинусоиды на угол $\varphi(N)$, определяющим линейное смещение косинусоиды на расстояние $\Delta x' = \frac{\varphi(N)}{2\pi N}$. Совокупность значений сдвига (смещения) фазы $\varphi(N)$ для различных пространственных частот составляет функцию передачи фазы (ФПФ) системы. Как следует из выражений (3.22)–(3.25), если $N = 0$, то $T(0) = 1$, а $\varphi(0) = 0$.

Определим контраст идеализированного изображения выражением

$$k = \frac{E_0(x')_{\max} - E_0(x')_{\min}}{E_0(x')_{\max} + E_0(x')_{\min}}. \quad (3.31)$$

Но $E_0(x')_{\max} = E_0 + E_1$, а $E_0(x')_{\min} = E_0 - E_1$. Тогда

$$k = \frac{E_0 + E_1 - E_0 + E_1}{E_0 + E_1 + E_0 - E_1} = \frac{E_1}{E_0} = m.$$

При этом контраст действительного изображения равен

$$\begin{aligned} k' &= \frac{E(x')_{\max} - E(x')_{\min}}{E(x')_{\max} + E(x')_{\min}} = \\ &= \frac{E_0 + E_1 T(N) - E_0 + E_1 T(N)}{E_0 + E_1 T(N) + E_0 - E_1 T(N)} = \frac{E_1}{E_0} T(N) = m T(N) = m'. \end{aligned}$$

Отсюда следует, что

$$T(N) = \frac{k'}{k} = \frac{m'}{m}.$$

При $k = 1$, т.е. при $E_1 = E_0$, $T(N) = k' = m'$.

Итак, нормированное распределение освещенности в изображении точечного предмета, образованном оптической системой, называется функцией рассеяния точки (ФРТ). Другими названиями этой же характеристики являются следующие: выходной сигнал импульсного типа, функция Грина, дифракционная картина Фраунгофера. ФРТ является одной из двух наиболее полных характеристик, описывающих качество изображения, образованного оптической системой. Распределение освещенности $E(x', y')$ в изображении любого предмета в принципе можно рассчитать как свертку распределения освещенности в идеализированном изображении предмета, т.е. $E_0(x', y')$, и ФРТ. Практически же для любого реального предмета, распределение освещенности в идеализированном изображении которого непрерывно, подобный расчет является довольно сложной процедурой. В то же время одиночные светящиеся точки или совокупность таких дискретных точек часто являются предметами для оптических систем, применяемых в астрономии, что придает ФРТ не только теоретическое значение, но и определенный практический смысл.

Второй из двух наиболее полных характеристик качества изображения является комплексная оптическая передаточная функция (ОПФ), удобная для характеристики качества изображения при непрерывном распределении освещенности в нем. Математически ФРТ и ОПФ взаимозаменяемы, так как каждая из них может быть получена путем Фурье-преобразования другой.

Применение ОПФ удобно потому, что, во-первых, процессы Фурье-преобразования функции $E_0(x', y')$, умножения ее на ОПФ и обратного Фурье-преобразования результата для получения функции $E(x', y')$ выполнить проще, чем эквивалентную математическую процедуру свертки; во-вторых, ОПФ для ряда независимых факторов, ухудшающих качество изображения, могут быть объединены простым их перемножением, в результате чего получается полная ОПФ изображающей системы.

3.2. Структура изображения точечного предмета, образованного оптической системой

3.2.1. Введение в дифракционную теорию образования изображения

Первое упоминание о дифракционных явлениях света появилось в работе Леонардо да Винчи (1452–1519). Однако, впервые они были описаны в книге монаха иезуита, физика и профессора математики в иезуитской коллегии в Болонье, Франческо Марии Гримальди (1618–1663), опубликованной в 1665 году спустя два года после его смерти. Пропустив

сквозь маленькое отверстие солнечный свет в темную комнату и поместив в конический пучок света на довольно большом расстоянии от отверстия обычный стержень, направляя тень от него на белую поверхность экрана, Гримальди наблюдал по обе стороны тени одну, две или три полосы (в зависимости от силы света), которые с края, обращенного к тени, были голубыми, а с противоположного – красными, при этом при ярком солнечном свете замечались цветные полосы и в самой тени. Такое свойство света Гримальди назвал дифракцией. Затем он проделал в ставне комнаты два отверстия и получил на экране изображения на таком расстоянии, чтобы они отчасти покрывали друг друга. В результате получились два темных пересекающихся кольца вокруг каждого из светлых кругов. Поле пересечения колец оказалось гораздо светлее полей каждого кольца в отдельности, при этом край каждого круга казался темным на освещенном поле другого. Из всех этих явлений, вместе взятых, Гримальди вывел следующее заключение: освещенное тело может сделаться темнее, если к полученному им свету прибавится новое количество света. Таким образом, он точно описывает результат интерференции света, хотя и не объясняет его.

Ф. Гримальди, без сомнения, думал о волнообразном движении света. Зачатки волновой теории света ясно выражены в его работе, где он писал: «Подобно тому, как вокруг камня, брошенного в воду (как вокруг центра), образуются круговые возвышения воды, точно так же вокруг тени непрозрачного предмета возникают блестящие полосы, которые соответственно форме последнего либо распространяются в длину, или же изгибаются дугообразно. Далее, подобно тому, как те круговые волны представляют простое скопление воды, вокруг которых с обеих сторон тянется углубление, так и блестящие полосы суть не что иное, как свет, распределенный неравномерно вследствие сильного рассеяния и прорезанный теньвыми промежутками».

Рассуждая о цветах, он, например, высказывает мнение, что цвета должны, по всей вероятности, обуславливаться различной скоростью колебаний светового вещества, подобно тому, как различные тоны зависят от неодинаковой скорости колебаний воздуха. Он считает цвета составными частями света, поясняя, что цвета могут существовать только в самом свете и что так называемые истинные, или постоянные, цвета тел зависят от способности последних отражать падающий на них свет с особыми видоизменениями. Ф. Гримальди заметил, что свет, пройдя через призму, рассеивается и, следовательно, одна часть светового луча преломляется сильнее другой. Отсюда вывод: там, где свет наименее преломлен, он всего гуще и ярче, т.е. окрашен в красный цвет; в месте же наибольшего преломления он более разряжен и получает голубую или фиолетовую окраски. Христиан Гюйгенс (1629–1695), впервые обосновавший волновую теорию, очевидно, не знал об открытии

Ф. Гримальди, иначе он, несомненно, сослался бы на него для подтверждения своей точки зрения. О возможности объяснить явление дифракции в рамках волновой теории нигде не упоминается вплоть до 1818 года, когда появился прекрасный мемуар Огюстена Жака Френеля (1788–1827), где было показано, что явление дифракции можно объяснить с помощью построения Гюйгенса и применения принципа интерференции. Позднее Густав Роберт Кирхгоф (1824–1887) придал исследованиям Френеля строго математическое обоснование и с этого времени началось широкое изучение дифракции.

Проблемы, возникающие при изучении дифракционных явлений, относятся к наиболее трудным в оптике и их редко удается довести до строгого решения. В большинстве же случаев, представляющих практический интерес, из-за математических трудностей приходится прибегать к приближенным методам и тут теория Гюйгенса и Френеля служит чрезвычайно мощным инструментом, позволяющим решить большинство вопросов, встречающихся в инструментальной оптике.

Согласно построению Гюйгенса каждую точку волнового фронта можно считать центром вторичного возмущения, которое вызывает элементарную сферическую волну, а волновой фронт в любой более поздний момент времени – огибающей этих волн. О. Френель смог объяснить явление дифракции, дополнив построение Гюйгенса утверждением, что вторичные волны интерферируют между собой. Это сочетание построения Гюйгенса с принципом интерференции называется принципом Гюйгенса-Френеля.

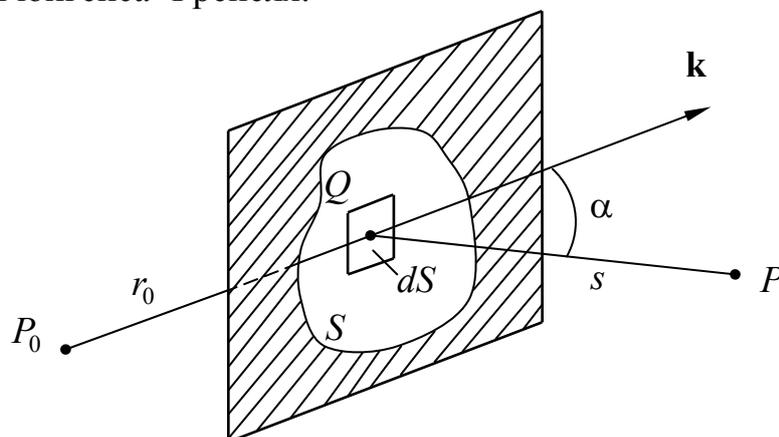


Рис. 3.4. Дифракция на отверстии в экране

Рассмотрим какой-либо экран с отверстием, через которое проходит свет от точечного монохроматического источника P_0 , как показано на рис. 3.4. Размеры отверстия много больше длины волны света. Будем под A понимать любой компонент векторов \mathbf{E} и \mathbf{H} электромагнитного поля световой волны, опуская при этом множитель $\exp(-i\omega t)$, определяющий зависимость от времени. Пусть требуется определить световое возмущение в точке P . Возмущение в точке Q волнового фронта можно представить в

виде $A \frac{\exp(ikr_0)}{r_0}$, где A – амплитуда светового возмущения на расстоянии

единицы длины от источника; $k = \frac{2\pi}{\lambda}$ – волновое число; λ – длина волны

излучения. Разделим мысленно поверхность отверстия на элементарные участки площадью dS , малые по сравнению с размерами отверстия, но большие по сравнению с длиной волны. Можно считать, что каждый из этих участков сам становится источником световой волны, распространяющейся во все стороны. Световое возмущение $du(P)$, создаваемое элементарным участком dS в точке P , пропорционально амплитуде светового возмущения в самом участке dS (в точке Q) и проекции dS_n площадки dS этого участка на плоскость, перпендикулярную вектору \mathbf{k} луча, пришедшего из источника света в dS , поскольку именно проекция площадки dS_n определяет световой пучок лучей, проходящих из источника через площадку dS .

При вычислении вклада участка dS отверстия в световое возмущение $du(P)$ необходимо учесть изменения амплитуды и фазы вторичной волны при ее распространении от dS к P . Это приводит к появлению в выражении для $du(P)$ множителя $\frac{\exp(iks)}{s}$, где s – расстояние от dS до P .

Таким образом,

$$du(P) = K(\alpha) A \frac{\exp(ikr_0)}{r_0} \frac{\exp(iks)}{s} dS_n,$$

где $K(\alpha)$ – некоторый коэффициент наклона, учитывающий зависимость амплитуды вторичных волн от угла α (часто называемого углом дифракции) между вектором \mathbf{k} и направлением на точку наблюдения. Следуя Френелю, предположим, что коэффициент $K(\alpha)$ максимален при $\alpha = 0$, плавно убывает с увеличением угла α и равен нулю при $\alpha = \frac{\pi}{2}$.

Многие практически важные задачи дифракции можно решить при этих весьма общих предположениях относительно $K(\alpha)$, не уточняя конкретного вида его зависимости от α . И, наконец, предположим, что в точке P сказывается влияние только той части первоначальной волны, которая не загороживается каким-либо препятствием, находящемся между P_0 и P . Поэтому полное возмущение в точке P равно

$$u(P) = A \int \int \frac{\exp(ikr_0)}{r_0} \frac{\exp(iks)}{s} K(\alpha) dS_n. \quad (3.32)$$

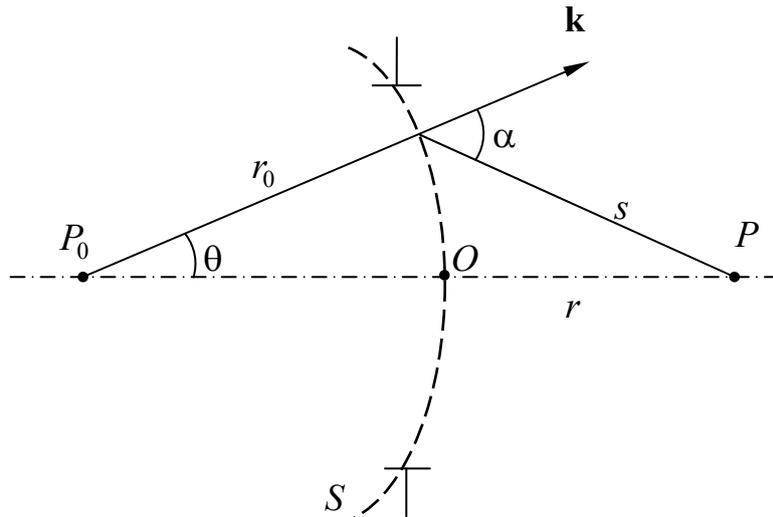


Рис. 3.5. Принцип Гюйгенса-Френеля

Формула (3.32) дает математическое выражение принципа Гюйгенса-Френеля. Применим ее для определения светового возмущения в точке P за круглым отверстием в экране.

Будем считать, что точечный источник P_0 и точка P лежат на прямой, проходящей через центр отверстия и перпендикулярной его плоскости, как показано на рис. 3.5. В качестве вспомогательной поверхности S выберем часть сферы, проходящей через края отверстия, радиус кривизны которой равен r_0 , а центр ее расположен в точке источника. Световое возмущение в каждой точке поверхности S равно

$$u(0) = u_0 \exp(ikr_0) = \frac{A}{r_0} \exp(ikr_0). \quad (3.33)$$

При этом выражение (3.32) принимает вид:

$$u(P) = \frac{A \exp(ikr_0)}{r_0} \iint_S \frac{\exp(iks)}{s} K(\alpha) dS. \quad (3.34)$$

Для того, чтобы найти величину, определяемую полученным выражением, воспользуемся так называемыми зонами Френеля. Для этого построим вокруг точки P сферы с радиусами, равными $s = r, r + \frac{\lambda}{2}, r + 2\frac{\lambda}{2}, r + 3\frac{\lambda}{2}, \dots, r + j\frac{\lambda}{2}, \dots$, где $r = OP$, а O – точка пересечения отрезка P_0P со сферой (волновым фронтом) S , как показано на рис. 3.6.

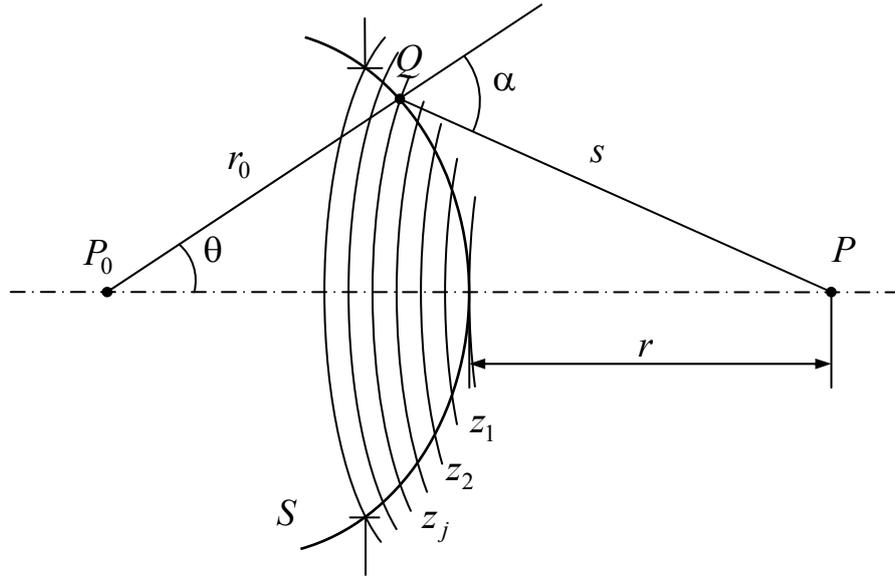


Рис. 3.6. Зоны Френеля

Сферы делят поверхность S на целый ряд зон $z_1, z_2, \dots, z_j, \dots$. Пусть r_0 и r велики по сравнению с длиной волны. Тогда можно предположить, что в любой зоне величина $K(\alpha)$ постоянна и в зоне j равна K_j . Из рисунка следует, что

$$s^2 = r_0^2 + (r_0 + r)^2 - 2r_0(r_0 + r)\cos\theta. \quad (3.35)$$

Дифференцируя это уравнение, получаем $sds = r_0(r_0 + r)\sin\theta d\theta$.

При этом площадь поверхности j -ой зоны равна

$$dS_j = 2\pi r_0^2 \sin\theta d\theta = 2\pi \frac{r_0}{r_0 + r} sds.$$

Следовательно, вклад j -ой зоны в световое возмущение в точке P равен

$$\begin{aligned} u_j(P) &= 2\pi \frac{A \exp(ikr_0)}{r_0 + r} K_j \int_{r+(j-1)\frac{\lambda}{2}}^{r+j\frac{\lambda}{2}} \exp(iks) ds = \\ &= -\frac{2\pi i}{k} \frac{A \exp[ik(r_0 + r)]}{r_0 + r} K_j \exp\left(ikj\frac{\lambda}{2}\right) \left[1 - \exp\left(-ik\frac{\lambda}{2}\right)\right]. \end{aligned}$$

Так как $k\lambda = 2\pi$, последние два множителя этого выражения легко привести к виду

$$\exp\left(ikj\frac{\lambda}{2}\right) \left[1 - \exp\left(-ik\frac{\lambda}{2}\right)\right] = \exp(i\pi j) [1 - \exp(-i\pi)] = 2(-1)^j.$$

В результате получаем, что

$$u_j(P) = 2i\lambda(-1)^{j+1} K_j \frac{A \exp[ik(r_0 + r)]}{r_0 + r}. \quad (3.36)$$

Отсюда следует, что вклад j -ой зоны отличается от вклада любой другой зоны лишь знаком и величиной коэффициента $K(\alpha)$, т.е. площади поверхностей зон Френеля равны между собой.

Заметим, что вклады следующих друг за другом зон имеют разные знаки. Результирующий эффект в точке P получается суммированием вкладов от всех зон, т.е.

$$u(P) = 2i\lambda \frac{A \exp[ik(r_0 + r)]}{r_0 + r} \sum_{j=1}^n (-1)^{j+1} K_j. \quad (3.37)$$

Обозначим

$$\sum_{j=1}^n (-1)^{j+1} K_j = K_1 - K_2 + K_3 - \dots + (-1)^{n+1} K_n = \Sigma.$$

Перепишем этот ряд в виде

$$\Sigma = \frac{1}{2} K_1 + \left(\frac{1}{2} K_1 - K_2 + \frac{1}{2} K_3 \right) + \left(\frac{1}{2} K_3 - K_4 + \frac{1}{2} K_5 \right) + \dots$$

Предположим, что изменение коэффициента K по зонам таково, что

$$K_{2j} > \frac{1}{2} K_{2j-1} + \frac{1}{2} K_{2j+1}.$$

Тогда

$$\Sigma < \frac{1}{2} K_1 + \frac{1}{2} K_n, \text{ если } n \text{ нечетно};$$

$$\Sigma < \frac{1}{2} K_1 + \frac{1}{2} K_{n-1} - K_n, \text{ если } n \text{ четно}.$$

Представим тот же ряд в виде

$$\Sigma = K_1 - \frac{1}{2} K_2 - \left(\frac{1}{2} K_2 - K_3 + \frac{1}{2} K_4 \right) - \left(\frac{1}{2} K_4 - K_5 + \frac{1}{2} K_6 \right) - \dots$$

В этом случае, поскольку

$$K_{2j+1} > \frac{1}{2} K_{2j} + \frac{1}{2} K_{2j+2}, \text{ имеем}$$

$$\Sigma > K_1 - \frac{1}{2} K_2 - \frac{1}{2} K_{n-1} + K_n, \text{ если } n \text{ нечетно};$$

$$\Sigma > K_1 - \frac{1}{2} K_2 - \frac{1}{2} K_n, \text{ если } n \text{ четно}.$$

Из полученных соотношений следует, что

$$K_1 - \frac{1}{2} K_2 - \frac{1}{2} K_{n-1} + K_n < \Sigma < \frac{1}{2} K_1 + \frac{1}{2} K_n, \text{ если } n \text{ нечетно};$$

$$K_1 - \frac{1}{2} K_2 - \frac{1}{2} K_n < \Sigma < \frac{1}{2} K_1 + \frac{1}{2} K_{n-1} - K_n, \text{ если } n \text{ четно}.$$

Представим эти выражения в виде

$$\frac{1}{2}K_1 + \frac{1}{2}K_n + \frac{1}{2}(K_1 - K_2) - \frac{1}{2}(K_{n-1} - K_n) < \Sigma < \frac{1}{2}K_1 + \frac{1}{2}K_n;$$

$$\frac{1}{2}K_1 - \frac{1}{2}K_n + \frac{1}{2}(K_1 - K_2) - \frac{1}{2}(K_{n-1} - K_n) < \Sigma < \frac{1}{2}K_1 - \frac{1}{2}K_n.$$

Поскольку величина K_j мало отличается от величин соседних коэффициентов K_{j-1} и K_{j+1} , приближенно можно считать, что

$$\Sigma = \frac{1}{2}K_1 + \frac{1}{2}K_n, \text{ если } n \text{ нечетно};$$

$$\Sigma = \frac{1}{2}K_1 - \frac{1}{2}K_n, \text{ если } n \text{ четно}.$$

Легко проверить, что полученные соотношения остаются справедливыми и в том случае, если предположить, что K_j меньше среднего арифметического соседних членов. А тогда

$$u(P) = i\lambda(K_1 \pm K_n) \frac{A \exp[ik(r_0 + r)]}{r_0 + r} \quad (3.38)$$

или

$$u(P) = \frac{1}{2}[u_1(P) + u_n(P)]. \quad (3.39)$$

Для последней зоны z_n , видимой из P , линия QP становится касательной к волновому фронту, при этом угол $\alpha = \frac{1}{2}\pi$. По

предположению Френеля $K_n = K\left(\frac{1}{2}\pi\right) = 0$. При этом выражение (3.38)

принимает вид

$$u(P) = i\lambda K_1 \frac{A \exp[ik(r_0 + r)]}{r_0 + r} = \frac{1}{2}u_1(P), \quad (3.40)$$

т.е. полное световое возмущение в точке P равняется половине возмущения, обусловленного действием первой зоны. Однако, в отсутствие экрана с отверстием аналогично выражению (3.33) имеем

$$u(P) = \frac{A}{r_0 + r} \exp[ik(r_0 + r)]. \quad (3.41)$$

Сопоставляя правые части выражений (3.40) и (3.41) получаем $i\lambda K_1 = 1$, т.е.

$$K_1 = -\frac{i}{\lambda} = \frac{\exp\left(-i\frac{\pi}{2}\right)}{\lambda}.$$

Множитель $\exp\left(-i\frac{\pi}{2}\right)$ можно объяснить, если предположить, что вторичные волны отстают по фазе на четверть периода от первичной волны.

Хорошей иллюстрацией, подтверждающей приведенный метод рассуждения Френеля, может служить опыт с зонной пластинкой. Обозначим $r_0 \sin \theta = R_n$, а $s_n = r + n\frac{\lambda}{2}$. Тогда

$$s_n^2 = \left(r + n\frac{\lambda}{2}\right)^2 = r^2 + nr\lambda + \left(n\frac{\lambda}{2}\right)^2 = r^2 \left[1 + \frac{n}{r}\lambda + \left(\frac{n\lambda}{r}\right)^2\right] \approx r^2 + nr\lambda;$$

$$\cos \theta = \sqrt{1 - \sin^2 \theta} = \sqrt{1 - \frac{R_n^2}{r_0^2}} \approx 1 - \frac{1}{2} \frac{R_n^2}{r_0^2}.$$

Подставив эти соотношения в уравнение (3.35), получаем

$$R_n \approx \sqrt{n\lambda \frac{r_0 r}{r_0 + r}}. \quad (3.42)$$

Рассмотрим пластинку, состоящую из последовательно чередующихся прозрачных и непрозрачных колец, радиусы которых удовлетворяют соотношению (3.42). Если пластинка прикрывает все четные зоны и оставит свободными все нечетные, начиная с центральной, то амплитуда светового возмущения в точке P будет пропорциональна

$$\Sigma = K_1 + K_3 + K_5 + \dots + K_{2n+1} + \dots$$

В противном случае амплитуда светового возмущения в точке P будет пропорциональна

$$\Sigma = K_2 + K_4 + \dots + K_{2n} + \dots$$

Опыт полностью подтверждает тот факт, что зонная пластинка Френеля увеличивает освещенность в точке P , действуя подобно собирающей линзе.

Положив в выражении (3.42) отношение $\frac{R_n^2}{n\lambda} = f'$, получаем

$$\frac{1}{f'} = \frac{1}{r_0} + \frac{1}{r},$$

т.е. r_0 и r взаимосвязаны соотношением, подобным формуле отрезков (без учета принятого правила знаков), а величина f' играет роль фокусного расстояния. При $r_0 = \infty$, т.е. если на пластинку падает плоская волна, то $r = f'$.

В отличие от линзы зонная пластинка образует не одно, а целую последовательность в направлении оси изображений источника. Действительно, если сместить точку наблюдения таким образом, чтобы в пределах каждого прозрачного кольца зонной пластинки укладывалась не

одна, а три зоны Френеля, действие двух из них будет взаимно скомпенсировано, а амплитуда колебаний в точке наблюдения будет определяться лишь третьей зоной, при этом волны, приходящие в точку наблюдения от нескомпенсированных зон всех колец пластинки, остаются синфазными, т.е. амплитуда колебаний в смещенной точке наблюдения имеет повышенное значение. Разность фаз между волнами от нескомпенсированных зон соседних колец при смещенной точке наблюдения увеличивается в три раза, при этом положение смещенной точки определится соотношением $\frac{1}{r_0} + \frac{1}{r_1} = \frac{3}{f'}$. В общем случае, если в

пределах каждого кольца пластинки укладывается любое нечетное число $2n+1$ зон Френеля, то в соответствующих точках будет наблюдаться повышенное значение амплитуды светового возмущения. При этом положение точек наблюдения определяется соотношением

$$\frac{1}{r_0} + \frac{1}{r_n} = \frac{2n+1}{f'} = \frac{1}{f'_n},$$

т.е. $f'_n = \frac{f'}{2n+1}$, где $n = 0, 1, 2, \dots$,

которое можно истолковать, как наличие у зонной пластинки последовательности (многих) фокусных расстояний f'_n .

Итак, за зонной пластинкой создается сложное волновое поле с множеством точек повышенной освещенности на осевой линии. Возникновение многих изображений обусловлено дифракцией падающей волны на сложном экране, который представляет собой зонная пластинка.

При этом следует иметь в виду, что фокусное расстояние зонной пластинки определяется величиной, обратно пропорциональной длине волны, а, следовательно, положение точек повышенной освещенности зависит от длины волны излучения, т.е. зонная пластинка обладает значительной хроматической аберрацией в изображении точки.

Можно достичь еще большей освещенности в изображении точки, если не перекрывать световую волну в пределах четных зон, а сообщить ей в пределах этих зон изменение фаз на π . Такую фазовую зонную пластинку впервые изготовил Р. Вуд, покрыв стекло тонким слоем лака и выгравировав на нем зонную пластинку так, что оптическая толщина нечетных зон отличалась от толщины четных на величину $\frac{1}{2}\lambda$.

3.2.2. Интегральная теорема Гельмгольца-Кирхгофа

В 1882 году идее Гюйгенса-Френеля Г. Кирхгоф придал строгий математический вид и показал, что принцип Гюйгенса-Френеля можно считать приближенной формой определенной интегральной теоремы. В этой теореме решение однородного волнового уравнения в произвольной

точке поля выражается через значения искомой величины и ее первой производной во всех точках произвольной замкнутой поверхности, окружающей эту точку.

Максвелл показал, что в однородной среде в областях, где отсутствуют заряды и токи, любой из декартовых компонентов $u(\mathbf{r}, t)$ векторов электрической \mathbf{E} или магнитной \mathbf{H} напряженностей поля удовлетворяют однородному волновому уравнению

$$\nabla^2 u - \frac{1}{V^2} \frac{\partial^2 u}{\partial t^2} = 0, \quad (3.43)$$

где V – скорость распространения волны, $V = \frac{c}{n}$, c – скорость света в вакууме; n – показатель преломления среды; ∇^2 – дифференциальный оператор Лапласа:

$$\nabla^2 = \Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

Пусть $\mathbf{r}(x, y, z)$ – радиус-вектор некоторой точки P , а $\mathbf{s}(s_x, s_y, s_z)$ – единичный вектор с фиксированным направлением. При этом любое решение уравнения (3.43) в виде $u = u(\mathbf{r} \cdot \mathbf{s}, t)$ определяет плоскую волну, так как в каждый момент времени величина u постоянна в плоскостях $\mathbf{r} \cdot \mathbf{s} = \text{const}$, которые перпендикулярны к единичному вектору \mathbf{s} . Общее решение уравнения (3.43) можно записать в виде:

$$u = u_1(\mathbf{r} \cdot \mathbf{s} - Vt) + u_2(\mathbf{r} \cdot \mathbf{s} + Vt), \quad (3.44)$$

где u_1 и u_2 – произвольные функции.

Если дан точечный источник излучения, излучающий равномерно во всех направлениях (т.е. изотропный), то амплитуду волны этого источника в свободном пространстве на расстоянии r от источника можно определить выражением

$$u = u(r, t), \quad (3.45)$$

где $r = |\mathbf{r}| = \sqrt{x^2 + y^2 + z^2}$.

Используя соотношения $\frac{\partial}{\partial x} = \frac{\partial r}{\partial x} \frac{\partial}{\partial r} = \frac{x}{r} \frac{\partial}{\partial r}$:

$$\begin{aligned} \frac{\partial^2}{\partial x^2} &= \frac{\partial}{\partial x} \left(\frac{x}{r} \frac{\partial}{\partial r} \right) = \frac{1}{r} \frac{\partial}{\partial r} + x \frac{\partial r}{\partial x} \frac{\partial}{\partial r} \left(\frac{1}{r} \frac{\partial}{\partial r} \right) = \\ &= \frac{1}{r} \frac{\partial}{\partial r} + \frac{x^2}{r^2} \frac{\partial^2}{\partial r^2} - \frac{x^2}{r^3} \frac{\partial}{\partial r} = \frac{x^2}{r^2} \frac{\partial^2}{\partial r^2} + \frac{1}{r} \left(1 - \frac{x^2}{r^2} \right) \frac{\partial}{\partial r} \text{ и т.д.,} \end{aligned}$$

получаем

$$\nabla^2 u = \frac{\partial^2 u}{\partial r^2} + \frac{2}{r} \frac{\partial u}{\partial r}.$$

С другой стороны

$$\frac{\partial^2}{\partial r^2}(ur) = \frac{\partial}{\partial r} \left(u + r \frac{\partial u}{\partial r} \right) = \frac{\partial u}{\partial r} + \frac{\partial u}{\partial r} + r \frac{\partial^2 u}{\partial r^2} = r \frac{\partial^2 u}{\partial r^2} + 2 \frac{\partial u}{\partial r}.$$

Сопоставляя правые части полученных выражений, имеем

$$\nabla^2 u = \frac{1}{r} \frac{\partial^2(ru)}{\partial r^2}.$$

При этом волновое уравнение (3.43) принимает вид:

$$\frac{\partial^2(ru)}{\partial r^2} - \frac{1}{V^2} \frac{\partial^2(ru)}{\partial t^2} = 0. \quad (3.46)$$

Заметим, что уравнение (3.46) совпадает с волновым уравнением (3.43), если одну из координатных осей совместить с направлением r , а величину u заменить на ru . Следовательно, решение уравнения (3.46) можно сразу же написать в виде выражения (3.44), т.е.

$$u = \frac{u_1(r-Vt)}{r} + \frac{u_2(r+Vt)}{r}, \quad (3.47)$$

где u_1 и u_2 – по-прежнему произвольные функции.

В правой части равенства (3.47) первый член представляет сферическую волну, расходящуюся от начала координат, а второй – сферическую волну, сходящуюся к началу координат, причем скорость распространения обеих волн равна V .

Вид функции u зависит от природы источника излучения. В случае монохроматической волны функция u является синусоидальной или косинусоидальной функцией, например, в виде:

$$u(r, t) = \frac{a}{r} \cos[k(r-Vt) + \theta], \quad (3.48)$$

где a – положительная постоянная; $k = \frac{2\pi}{\lambda}$ – волновое число; θ – фазовый угол.

Функцию u в случае монохроматической волны можно записать в комплексной форме:

$$u(r, t) = \frac{A}{r} \exp i(kr - \omega t), \quad (3.49)$$

где $A = a \exp i\theta$; $\omega = kV$ – угловая частота волны.

Следует отметить, что если две различные функции u являются решением уравнения (3.43), то в соответствии с принципом суперпозиции решением является и их сумма.

Пусть световое возмущение в точке P в момент времени t описывается скалярной функцией $u(P, t)$. В случае линейно поляризованных волн можно считать, что эта функция определяет

напряженность электрического или магнитного полей. Развернутое выражение для поля монохроматической волны можно представить в виде:

$$u(P, t) = U(P) \exp(-i\omega t), \quad (3.50)$$

где $U(P)$ – комплексная функция координат (иногда называемая фазором); $U(P) = U_0(P) \exp[-i\varphi(P)]$, где $U_0(P)$ и $\varphi(P)$ – амплитуда и фаза волны в точке P соответственно.

Если функция светового возмущения $u(P, t)$ описывает световую волну, то в каждой точке, где нет источников излучения, она должна удовлетворять скалярному волновому уравнению (3.43), при этом для описания светового возмущения в точке P достаточно комплексной функции $U(P)$, так как зависимость u от времени известна заранее.

Из выражения (3.50) находим, что

$$\nabla^2 u(P, t) = \exp(-i\omega t) \nabla^2 U(P).$$

Дифференцируя выражение (3.50), получаем

$$\frac{\partial u(P, t)}{\partial t} = -i\omega U(P) \exp(-i\omega t);$$

$$\frac{\partial^2 u(P, t)}{\partial t^2} = i^2 \omega^2 U(P) \exp(-i\omega t) = -\omega^2 U(P) \exp(-i\omega t).$$

Подставив эти соотношения в уравнение (3.43), получаем

$$\exp(-i\omega t) \nabla^2 U(P) + \frac{\omega^2}{V^2} U(P) \exp(-i\omega t) = 0.$$

Отсюда следует, что

$$\nabla^2 U(P) + k^2 U(P) = 0,$$

где $k = \frac{2\pi}{\lambda} = 2\pi \frac{\nu}{V} = \frac{\omega}{V}$; ν – частота колебаний.

Это выражение можно записать в виде:

$$(\nabla^2 + k^2)U = 0. \quad (3.51)$$

Комплексная амплитуда любого светового монохроматического возмущения, распространяющегося в свободном пространстве, должна удовлетворять уравнению (3.51). Комплексную функцию светового возмущения U в выбранной точке пространства определим при помощи соотношения, называемого теоремой Грина.

В теории поля справедливы следующие формулы, заменяющие тройной интеграл двойным:

– формула для градиента

$$\int_V \text{grad} f dV = \int_S f d\mathbf{S}; \quad (3.52)$$

– формула для дивергенции (теорема Остроградского-Гаусса)

$$\int_V \text{div} \mathbf{A} dV = \int_S \mathbf{A} \cdot d\mathbf{S}, \quad (3.53)$$

где $d\mathbf{S} = \mathbf{n}dS$; \mathbf{n} – единичный вектор внешней нормали к площадке dS . В соответствии с определением дивергенции находим

$$\begin{aligned} \operatorname{div}(f\mathbf{A}) &= \frac{\partial(fA_x)}{\partial x} + \frac{\partial(fA_y)}{\partial y} + \frac{\partial(fA_z)}{\partial z} = \\ &= f\left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}\right) + A_x \frac{\partial f}{\partial x} + A_y \frac{\partial f}{\partial y} + A_z \frac{\partial f}{\partial z} = \\ &= f\operatorname{div}\mathbf{A} + \mathbf{A}\operatorname{grad}f. \end{aligned} \quad (3.54)$$

Пусть S – замкнутая поверхность, ограничивающая объем V ; p и q – две скалярные функции точки M . Заменяв в выражении (3.54) f на p , а \mathbf{A} на $\operatorname{grad}q$, получаем

$$\operatorname{div}(p\operatorname{grad}q) = p\Delta q + \operatorname{grad}q \operatorname{grad}p. \quad (3.55)$$

Подставив в выражение (3.53) вместо вектора \mathbf{A} вектор $p\operatorname{grad}q$, с учетом выражения (3.55) получаем

$$\int_V (p\Delta q + \operatorname{grad}q \operatorname{grad}p) dV = \int_S p\operatorname{grad}q d\mathbf{S}.$$

Меняя роли функций p и q , имеем

$$\int_V (q\Delta p + \operatorname{grad}q \operatorname{grad}p) dV = \int_S q\operatorname{grad}p d\mathbf{S}.$$

Вычитая последнее выражение из предыдущего, получаем формулу Грина в виде:

$$\int_V (p\Delta q - q\Delta p) dV = \int_S (p\operatorname{grad}q - q\operatorname{grad}p) d\mathbf{S}. \quad (3.56)$$

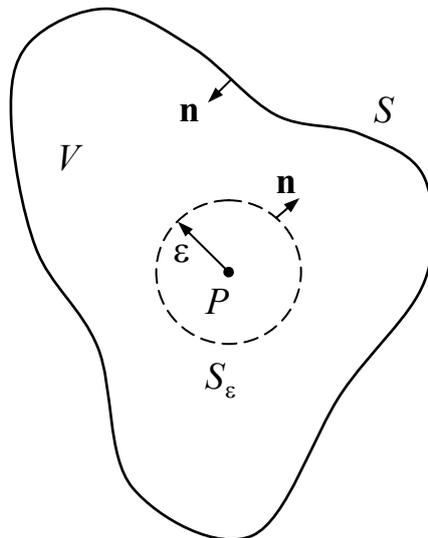


Рис. 3.7. К выводу теоремы Грина

Пусть V – объем, ограниченный произвольной замкнутой поверхностью S , а P – какая-нибудь точка внутри него, как показано на рис. 3.7, при этом в пространстве, ограниченном поверхностью S , нет источников света. Предположим, что комплексная функция светового

монохроматического возмущения U имеет непрерывные частные производные первого и второго порядков внутри этого объема и на поверхности S . Если G – любая другая функция, удовлетворяющая таким же требованиям непрерывности, как и U , то в соответствии с теоремой Грина имеем

$$\iiint_V (U\nabla^2 G - G\nabla^2 U) dV = -\iint_S \left(U \frac{\partial G}{\partial n} - G \frac{\partial U}{\partial n} \right) dS, \quad (3.57)$$

где $\frac{\partial}{\partial n}$ означает частную производную в каждой точке поверхности S , взятую по нормали, идущей внутрь пространства, ограниченного поверхностью S . Однако, только при осторожном выборе так называемой функции Грина G и замкнутой поверхности S теореме можно непосредственно применить к решению дифракционных задач.

В рассматриваемом случае задача состоит в том, чтобы выразить световое возмущение в точке P через его значения на поверхности S . Для решения этой задачи, следуя Кирхгофу, используем теорему Грина и выберем в качестве функции G сферическую волну единичной амплитуды, т.е. так называемую функцию Грина свободного пространства. При этом для функции G в произвольной точке Q поверхности S имеем

$$G(Q) = \frac{\exp(iks)}{s}, \quad (3.58)$$

где через s обозначено переменное расстояние от точки P до точки Q . Однако, в точке P функция (3.58) обращается в бесконечность, а, следовательно, не удовлетворяет требованию непрерывности и дифференцируемости. Для того, чтобы исключить точку P из области интегрирования, окружим ее небольшой сферой S_ε радиуса ε и применим теорему Грина, причем интегрирование будем вести по объему V' , заключенному между поверхностями S и S_ε , так что поверхностью интегрирования будет составная поверхность $S' = S + S_\varepsilon$, как показано на рис. 3.7. Заметим, что «внутренняя» нормаль к составной поверхности направлена внутрь (в общепринятом смысле) на поверхности S и наружу на поверхности S_ε (по направлению от точки P). Внутри объема V' световое возмущение G , представляющее собой сходящуюся в точке P сферическую волну, удовлетворяет уравнению Гельмгольца в виде

$$(\nabla^2 + k^2)G = 0. \quad (3.59)$$

Используя обе записи уравнений Гельмгольца, т.е. уравнения (3.51) и (3.59), преобразуем левую часть формулы Грина к виду:

$$\iiint_{V'} (U\nabla^2 G - G\nabla^2 U) dV = -\iiint_{V'} (UGk^2 - UGk^2) dV \equiv 0.$$

При этом выражение (3.57) принимает вид:

$$\iiint_{S'} \left(U \frac{\partial G}{\partial n} - G \frac{\partial U}{\partial n} \right) dS = 0$$

или

$$\iint_S \left(U \frac{\partial G}{\partial n} - G \frac{\partial U}{\partial n} \right) dS = - \iint_{S_\varepsilon} \left(U \frac{\partial G}{\partial n} - G \frac{\partial U}{\partial n} \right) dS. \quad (3.60)$$

Применив выражение (3.58) для произвольной точки на поверхности S' , получаем

$$\begin{aligned} \frac{\partial G(Q)}{\partial n} &= \frac{\partial G(Q)}{\partial s} \frac{\partial s}{\partial n} = \cos(\mathbf{n}, \mathbf{s}) \frac{\partial G(Q)}{\partial s} = \\ &= \cos(\mathbf{n}, \mathbf{s}) \left[ik \frac{\exp(iks)}{s} - \frac{\exp(iks)}{s^2} \right] = \\ &= \cos(\mathbf{n}, \mathbf{s}) \left(ik - \frac{1}{s} \right) \frac{\exp(iks)}{s}, \end{aligned} \quad (3.61)$$

где $\cos(\mathbf{n}, \mathbf{s})$ – косинус угла между направлением нормали \mathbf{n} и вектором \mathbf{s} , соединяющим точки P и Q . Для точки на поверхности S_ε имеем $\cos(\mathbf{n}, \mathbf{s}) = 1$, при этом

$$G(Q) = \frac{\exp(ik\varepsilon)}{\varepsilon}, \text{ а } \frac{\partial G(Q)}{\partial n} = \left(ik - \frac{1}{\varepsilon} \right) \frac{\exp(ik\varepsilon)}{\varepsilon}.$$

Если величина ε стремится к нулю, то в силу непрерывности функции U и ее производных можно записать, что

$$\begin{aligned} \iint_{S_\varepsilon} \left(U \frac{\partial G}{\partial n} - G \frac{\partial U}{\partial n} \right) dS = \\ = 4\pi\varepsilon^2 \left[U(P) \left(ik - \frac{1}{\varepsilon} \right) \frac{\exp(ik\varepsilon)}{\varepsilon} - \frac{\exp(ik\varepsilon)}{\varepsilon} \frac{\partial U(P)}{\partial n} \right]. \end{aligned}$$

При этом

$$\lim_{\varepsilon \rightarrow 0} \iint_{S_\varepsilon} \left(U \frac{\partial G}{\partial n} - G \frac{\partial U}{\partial n} \right) dS = -4\pi U(P).$$

Полученный результат позволяет представить выражение (3.60)

$$U(P) = \frac{1}{4\pi} \iint_S \left\{ U \frac{\partial}{\partial n} \left[\frac{\exp(iks)}{s} \right] - \left[\frac{\exp(iks)}{s} \right] \frac{\partial U}{\partial n} \right\} dS. \quad (3.62)$$

Это выражение, известное под названием интегральной теоремы Гельмгольца-Кирхгофа, играет важную роль в скалярной теории дифракции, поскольку позволяет выразить поле в любой точке P через «граничные значения» волны на любой замкнутой поверхности, окружающей эту точку.

3.2.3. Теория дифракции Кирхгофа

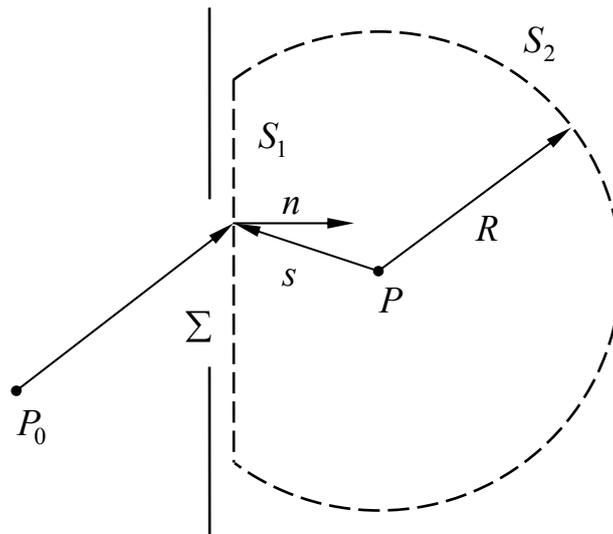


Рис. 3.8. Дифракция света на отверстии в непрозрачном экране

Рассмотрим задачу о дифракции на отверстии в бесконечном непрозрачном экране. Пусть световая волна падает слева на экран с отверстием, как показано на рис. 3.8. Необходимо рассчитать поле в точке P в пространстве справа от отверстия. Следуя Кирхгофу, возьмем замкнутую поверхность S таким образом, чтобы она состояла из двух частей: S_1 и S_2 , как показано на рисунке. Пусть плоская поверхность S_1 , лежащая сразу же за дифрагирующим экраном, замыкается большим сферическим колпаком S_2 радиуса R с центром в рассматриваемой точке P . Полная замкнутая поверхность S представляет собой просто сумму S_1 и S_2 . Применив интегральную теорему Гельмгольца-Кирхгофа в форме (3.62), получаем

$$U(P) = \frac{1}{4\pi} \int \int_{S_1+S_2} \left(U \frac{\partial G}{\partial n} - G \frac{\partial U}{\partial n} \right) dS, \quad (3.63)$$

где, как и прежде, $G = \frac{\exp(iks)}{s}$.

Функция G в точках на поверхности S_2 определяется выражением

$$G = \frac{\exp(ikR)}{R}. \quad (3.64)$$

Заметим, что с увеличением расстояния R значение функции G на поверхности S_2 уменьшается. Однако, следует иметь в виду, что площадь интегрирования при этом возрастает пропорционально R^2 . Исследуем вклад интегрирования по поверхности S_2 более обстоятельно.

В рассматриваемом случае в соответствии с выражением (3.61) имеем

$$\frac{\partial G}{\partial n} = \cos(\mathbf{n}, \mathbf{R}) \left(ik - \frac{1}{R} \right) \frac{\exp(ikR)}{R} = \left(\frac{1}{R} - ik \right) \frac{\exp(ikR)}{R} = \left(\frac{1}{R} - ik \right) G.$$

При этом

$$\iint_{S_2} \left(U \frac{\partial G}{\partial n} - G \frac{\partial U}{\partial n} \right) dS = \iint_{\Omega} \left[\left(\frac{1}{R} - ik \right) U - \frac{\partial U}{\partial n} \right] GR^2 d\Omega,$$

где Ω – телесный угол с вершиной в точке P , стягиваемый поверхностью S_2 . Величина $RG = \exp(ikR)$ равномерно ограничена на поверхности S_2 . Поэтому полный интеграл по S_2 будет стремиться к нулю по мере стремления R к бесконечности при условии, что

$$\lim_{R \rightarrow \infty} R \left[\left(\frac{1}{R} - ik \right) U - \frac{\partial U}{\partial n} \right] = 0$$

равномерно во всем телесном угле. Это условие называется условием Зоммерфельда для излучения.

Покажем, что сходящаяся сферическая волна (или линейная комбинация сферических волн), определяющая световое возмущение в точке P , удовлетворяет этому условию. Для этого предположим, что все источники света $S_{i1}, S_{i2}, S_{i3}, \dots$ лежат в конечной области пространства. Окружим эту область замкнутой поверхностью $S_{i\Sigma}$, как показано на рис. 3.9.

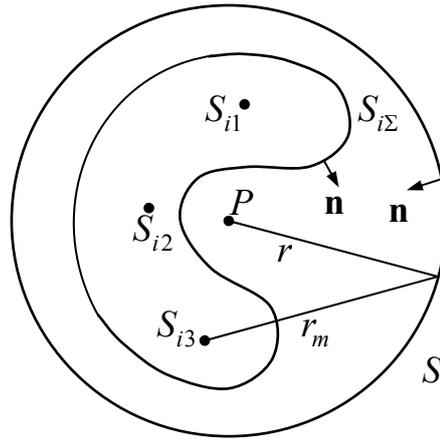


Рис. 3.9. К доказательству условия Зоммерфельда

Пусть точка P находится в пространстве вне поверхности $S_{i\Sigma}$. Опишем из точки P как из центра сферу настолько большего радиуса, чтобы она целиком окружала поверхность $S_{i\Sigma}$. Тогда в пространстве между поверхностями S и $S_{i\Sigma}$ не будет источников света, а поэтому можно для вычисления $U(P)$ применить формулу (3.62) в виде:

$$U(P) = \frac{1}{4\pi} \iint_{S_{i\Sigma} + S} \left(U \frac{\partial G}{\partial n} - G \frac{\partial U}{\partial n} \right) dS.$$

Докажем, что интеграл по сфере S стремится к нулю, когда ее радиус стремится к бесконечности. Для этого необходимо доказать, что соблюдается условие Зоммерфельда.

Предположим, что в пространстве, ограниченном поверхностью $S_{i\Sigma}$, находится один или несколько точечных источников света: S_1, S_2, S_3, \dots . Тогда поле этих источников на поверхности S можно представить в виде

$$U = \sum C_m \frac{\exp(ikr_m)}{r_m},$$

где C_m – постоянные коэффициенты.

Если при этом r стремится к ∞ , то и r_m стремится к ∞ , однако, разность $r_m - r$ будет оставаться конечной.

Представим световое возмущение U в виде:

$$U = \sum C_m \frac{\exp(ik(r + a_m))}{r + a_m} = \frac{\exp(ikr)}{r} \sum \frac{A_m}{1 + \frac{a_m}{r}},$$

где $A_m = C_m \exp(ika_m)$ – новые постоянные.

При этом

$$\begin{aligned} \frac{\partial U}{\partial n} &= \frac{\partial U}{\partial r} \frac{\partial r}{\partial n} = \frac{\partial U}{\partial r} \text{grad} r \cdot \mathbf{n} = \frac{\partial U}{\partial r} \frac{\mathbf{r}}{r} \cdot \mathbf{n} = \frac{\partial U}{\partial r} \cos(\mathbf{n} \cdot \mathbf{r}) = -\frac{\partial U}{\partial r} = \\ &= -\left(ik \frac{\exp(ikr)}{r} - \frac{\exp(ikr)}{r^2} \right) \sum \frac{A_m}{1 + \frac{a_m}{r}} - \frac{\exp(ikr)}{r^3} \sum \frac{A_m a_m}{\left(1 + \frac{a_m}{r}\right)^2} = \\ &= -\left(ik - \frac{1}{r} \right) U - \frac{\exp(ikr)}{r^3} \sum \frac{A_m a_m}{\left(1 + \frac{a_m}{r}\right)^2}. \end{aligned}$$

При $r = R$ имеем

$$\begin{aligned} \lim_{R \rightarrow \infty} R \left[\left(\frac{1}{R} - ik \right) U - \frac{\partial U}{\partial n} \right] &= \\ &= \lim_{R \rightarrow \infty} R \left[\left(\frac{1}{R} - ik \right) U + \left(ik - \frac{1}{R} \right) U + \frac{\exp(ikR)}{R^3} \sum \frac{A_m a_m}{\left(1 + \frac{a_m}{R}\right)^2} \right] = \\ &= \lim_{R \rightarrow \infty} \frac{\exp(ikR)}{R^2} \sum \frac{A_m a_m}{\left(1 + \frac{a_m}{R}\right)^2} = 0. \end{aligned}$$

Отсюда следует, что если некоторый участок поверхности S удаляется в бесконечность, то интеграл по этому участку поверхности стремится к нулю.

Устранив, таким образом, в выражении (3.63) интегрирование по поверхности S_2 , возмущение в точке P можно выразить через функцию светового возмущения и ее производную по нормали, взятые на бесконечной плоскости S_1 , расположенной непосредственно за экраном, т.е.

$$U(P) = \frac{1}{4\pi} \iint_{S_1} \left(U \frac{\partial G}{\partial n} - G \frac{\partial U}{\partial n} \right) dS. \quad (3.65)$$

Экран не прозрачен везде, кроме открытого отверстия, которое на рис. 3.8 обозначено через Σ . Поэтому Кирхгоф принял следующие интуитивные предположения:

1. На отверстии Σ распределение поля U и его производная $\frac{\partial U}{\partial n}$ 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 U и его производная $\frac{\partial U}{\partial n}$ тождественно равны нулю.

Эти условия известны под названием граничных условий Кирхгофа. Приняв эти условия, выражение (3.65) можно записать в виде:

$$U(P) = \frac{1}{4\pi} \iint_{\Sigma} \left(U \frac{\partial G}{\partial n} - G \frac{\partial U}{\partial n} \right) dS. \quad (3.66)$$

Граничные условия Кирхгофа значительно упрощают решение задачи. Тем не менее, следует помнить, что ни одно из них не может быть абсолютно справедливым. Однако, если размеры отверстия велики по сравнению с длиной волны, то получаемые результаты хорошо согласуются с результатами эксперимента.

Следует обратить внимание на то, что расстояние s от отверстия до точки наблюдения обычно во много раз больше длины волны, а, следовательно, $k \gg \frac{1}{s}$. При этом выражение (3.61) можно преобразовать к виду:

$$\frac{\partial G(Q)}{\partial n} = \cos(\mathbf{n}, \mathbf{s}) \left(ik - \frac{1}{s} \right) \frac{\exp(iks)}{s} \approx ik \cos(\mathbf{n}, \mathbf{s}) \frac{\exp(iks)}{s}.$$

Подставив это приближенное выражение и выражение (3.58) в выражение (3.66), получаем

$$U(P) = \frac{1}{4\pi} \iint_{\Sigma} \frac{\exp(iks)}{s} \left[ikU \cos(\mathbf{n}, \mathbf{s}) - \frac{\partial U}{\partial n} \right] dS. \quad (3.67)$$

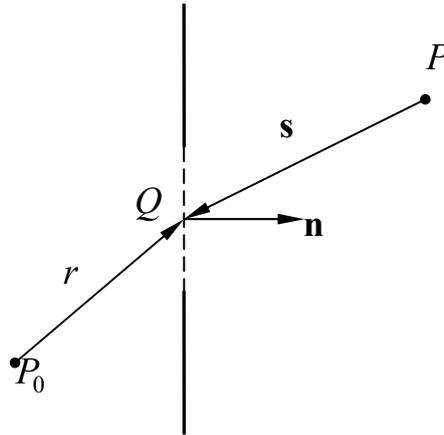


Рис. 3.10. К выводу дифракционной формулы Френеля-Кирхгофа

Теперь предположим, что на отверстие падает сферическая волна

$$U(Q) = \frac{A \exp(ikr)}{r},$$

исходящая из одиночного точечного источника, расположенного в точке P_0 на расстоянии r от точки Q , как показано на рис. 3.10. Если расстояние r во много раз больше длины волны, то

$$\frac{\partial U}{\partial n} \approx ik \cos(\mathbf{n}, \mathbf{r}) \frac{A \exp(ikr)}{r}.$$

При этом выражение (3.67) принимает вид:

$$U(P) = \frac{A_{ik}}{4\pi} \iint_{\Sigma} \frac{\exp(iks) \exp(ikr)}{sr} [\cos(\mathbf{n}, \mathbf{s}) - \cos(\mathbf{n}, \mathbf{r})] dS$$

или

$$U(P) = \frac{A}{2i\lambda} \iint_{\Sigma} \frac{\exp[ik(s+r)]}{sr} [\cos(\mathbf{n}, \mathbf{r}) - \cos(\mathbf{n}, \mathbf{s})] dS. \quad (3.68)$$

Это выражение называется дифракционной формулой Френеля-Кирхгофа.

Вполне очевидно, что вместо плоской поверхности Σ можно выбрать любую другую незамкнутую поверхность, границы которой совпадают с краем отверстия. В частности, при круглой апертуре и в случае, когда точка P расположена на оси апертуры, в качестве поверхности Σ удобно взять часть сферы L , концентричной точке P_0 , определяющей падающий волновой фронт. В этом случае расстояние r будет постоянной величиной, а угол (\mathbf{n}, \mathbf{r}) всегда равен нулю. Кроме того, угол (\mathbf{n}, \mathbf{s}) равен углу $\pi - \alpha$, как показано на рис. 3.11.

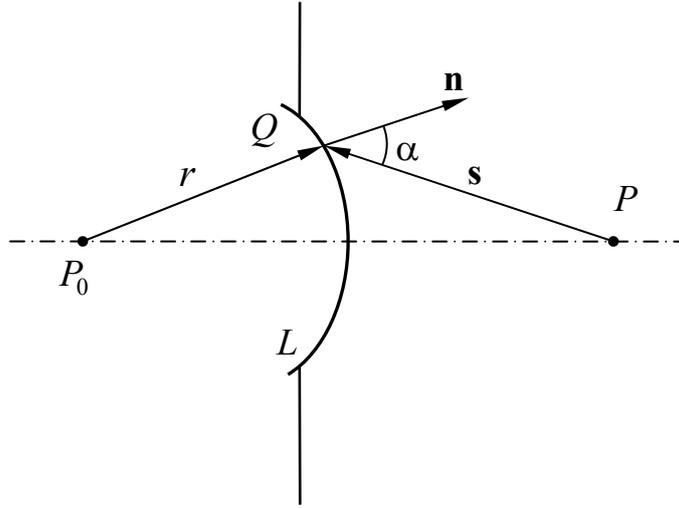


Рис. 3.11. Сферический волновой фронт на отверстии в плоском экране

При этом выражение (3.68) упрощается и принимает вид:

$$U(P) = \frac{1}{2i\lambda} \frac{A \exp(ikr)}{r} \iint_L \frac{\exp(iks)}{s} (1 + \cos \alpha) dS.$$

Этот результат находится в согласии с формулировкой принципа Гюйгенса-Френеля, если вкладом от элемента dS волнового фронта L считать световое возмущение

$$dU(P) = -\frac{i}{2\lambda} \frac{A \exp(ikr)}{r} \frac{\exp(iks)}{s} (1 + \cos \alpha) dS.$$

Из сопоставления с выражением (3.34) находим, что коэффициент наклона, фигурирующий в теории Френеля, равен

$$K(\alpha) = -\frac{i}{2\lambda} (1 + \cos \alpha).$$

Для центральной зоны $\alpha = 0$, при этом $K_1 = K(0) = -\frac{i}{\lambda}$, что согласуется с величиной, полученной Френелем. Однако, при $\alpha = \frac{\pi}{2}$ получаем $K\left(\frac{\pi}{2}\right) = -\frac{i}{2\lambda} = \frac{1}{2} K_1$, т.е. предположение о том, что $K\left(\frac{\pi}{2}\right) = 0$, неверно.

Возвращаясь снова к дифракционной формуле (3.68) Френеля-Кирхгофа, отметим, что она симметрична относительно источника и точки наблюдения. Это означает, что точечный источник, находящийся в P , производит в P_0 такое же действие, какое производил бы в точке P точечный источник равной интенсивности, помещенный в точку P_0 . Этот вывод иногда называют теоремой взаимности (или теоремой обратимости) Гельмгольца.

3.2.4. Дифракция Фраунгофера и Френеля

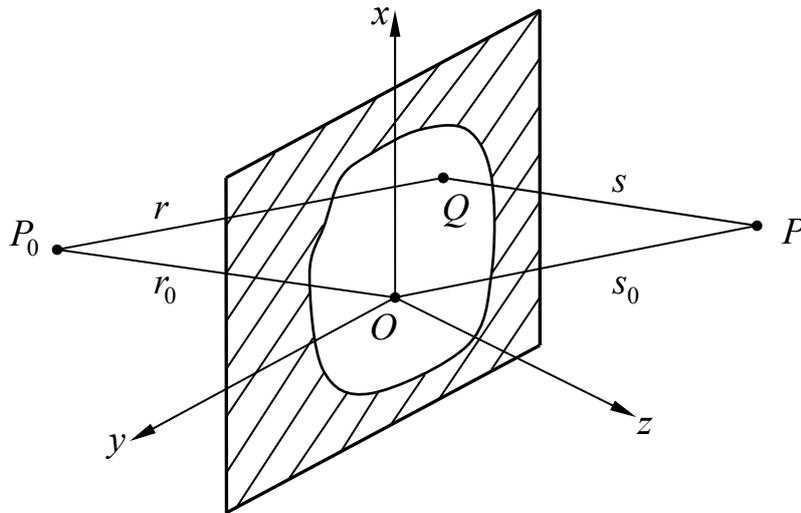


Рис. 3.12. Дифракция на отверстии в плоском экране

Пусть монохроматическая волна, излучаемая точечным источником P_0 , проходит сквозь отверстие в плоском непрозрачном экране, как показано на рис. 3.12. Будем считать, что линейные размеры отверстия велики по сравнению с длиной волны света, но малы по сравнению с расстояниями от точек P_0 и P до плоскости экрана. В соответствии с рисунком расстояние $P_0Q = r$, а расстояние $QP = s$. Определим световое возмущение в некоторой точке P справа от экрана. Для этого обратимся вновь к дифракционному интегралу Френеля-Кирхгофа

$$U(P) = -\frac{iA}{2\lambda} \iint_{\Sigma} \frac{\exp[ik(r+s)]}{rs} [\cos(\mathbf{n}, \mathbf{r}) - \cos(\mathbf{n}, \mathbf{s})] dS. \quad (3.69)$$

Величина $r+s$ в процессе интегрирования будет весьма заметно изменяться (на много длин волн) и поэтому величина $\exp[ik(r+s)]$ будет быстро осциллировать. Кроме того, если расстояния от точек P_0 и P до экрана велики по сравнению с линейными размерами отверстия, то множитель $[\cos(\mathbf{n}, \mathbf{r}) - \cos(\mathbf{n}, \mathbf{s})]$ изменяется по отверстию незначительно.

Возьмем за начало декартовой системы координат точку O отверстия, а оси Ox и Oy выберем в плоскости отверстия. Будем считать, что ось Oz направлена в сторону полупространства, в котором находится рассматриваемая точка P . Далее, предполагая, что углы, образованные линиями P_0O и OP с линией P_0P , не слишком велики, можно заменить этот множитель на $2\cos\delta$, где δ – угол между линией P_0P и нормалью к экрану. И, наконец, множитель $\frac{1}{rs}$ можно заменить множителем $\frac{1}{r_0s_0}$, где

r_0 и s_0 – расстояния от точек P_0 и P до начала координат. При этом уравнение (3.69) можно представить в виде:

$$U(P) \approx -\frac{iA \cos \delta}{\lambda r_0 s_0} \iint_{\Sigma} \exp[ik(r+s)] dS. \quad (3.70)$$

Пусть (x_0, y_0, z_0) и (x, y, z) – координаты точек P_0 и P соответственно, а (ξ, η) – координаты точки Q отверстия.

Тогда

$$\left. \begin{aligned} r^2 &= (x_0 - \xi)^2 + (y_0 - \eta)^2 + z_0^2, \\ s^2 &= (x - \xi)^2 + (y - \eta)^2 + z^2, \end{aligned} \right\} \quad (3.71)$$

$$\left. \begin{aligned} r_0^2 &= x_0^2 + y_0^2 + z_0^2, \\ s_0^2 &= x^2 + y^2 + z^2. \end{aligned} \right\} \quad (3.72)$$

Используя соотношения (3.72), уравнения (3.71) можно представить в виде:

$$\left. \begin{aligned} r^2 &= r_0^2 - 2(x_0 \xi + y_0 \eta) + \xi^2 + \eta^2, \\ s^2 &= s_0^2 - 2(x \xi + y \eta) + \xi^2 + \eta^2. \end{aligned} \right\} \quad (3.73)$$

Отсюда находим, что

$$\left. \begin{aligned} r &= r_0 \sqrt{1 - 2 \frac{x_0 \xi + y_0 \eta}{r_0^2} + \frac{\xi^2 + \eta^2}{r_0^2}}, \\ s &= s_0 \sqrt{1 - 2 \frac{x \xi + y \eta}{s_0^2} + \frac{\xi^2 + \eta^2}{s_0^2}}. \end{aligned} \right\} \quad (3.74)$$

Если линейные размеры отверстия в экране малы по сравнению с длиной отрезков r_0 и s_0 ; то правые части выражений (3.74) можно

разложить в степенные ряды по $\frac{\xi}{r_0}$, $\frac{\eta}{r_0}$, $\frac{\xi}{s_0}$ и $\frac{\eta}{s_0}$. При этом получим

$$\left. \begin{aligned} r &\approx r_0 - \frac{x_0 \xi + y_0 \eta}{r_0} + \frac{\xi^2 + \eta^2}{2r_0} - \frac{(x_0 \xi + y_0 \eta)^2}{2r_0^3} + \dots, \\ s &\approx s_0 - \frac{x \xi + y \eta}{s_0} + \frac{\xi^2 + \eta^2}{2s_0} - \frac{(x \xi + y \eta)^2}{2s_0^3} + \dots \end{aligned} \right\} \quad (3.75)$$

Полученные соотношения позволяют выражение (3.70) представить в виде:

$$U(P) = -\frac{i \cos \delta}{\lambda} \frac{A \exp[ik(r_0 + s_0)]}{r_0 s_0} \iint_{\Sigma} \exp[ikf(\xi, \eta)] d\xi d\eta, \quad (3.76)$$

$$\text{где } f(\xi, \eta) = -\frac{x_0 \xi + y_0 \eta}{r_0} - \frac{x \xi + y \eta}{s_0} + \frac{\xi^2 + \eta^2}{2r_0} + \frac{\xi^2 + \eta^2}{2s_0} - \frac{(x_0 \xi + y_0 \eta)^2}{2r_0^3} - \frac{(x \xi + y \eta)^2}{2s_0^3} + \dots \quad (3.77)$$

Обозначим первые два направляющих косинуса через (l_0, m_0) и (l, m) , т.е.

$$\left. \begin{aligned} l_0 &= \frac{0 - x_0}{r_0} = -\frac{x_0}{r_0}, & l &= \frac{x - 0}{s_0} = \frac{x}{s_0}, \\ m_0 &= \frac{0 - y_0}{r_0} = -\frac{y_0}{r_0}, & m &= \frac{y - 0}{s_0} = \frac{y}{s_0}. \end{aligned} \right\} \quad (3.78)$$

При этом выражение (3.77) можно переписать в виде:

$$f(\xi, \eta) = (l_0 - l)\xi + (m_0 - m)\eta + \frac{1}{2} \left[\left(\frac{1}{r_0} + \frac{1}{s_0} \right) (\xi^2 + \eta^2) - \frac{(l_0 \xi + m_0 \eta)^2}{r_0} - \frac{(l \xi + m \eta)^2}{s_0} \right] + \dots \quad (3.79)$$

Итак, задача определения светового возмущения в точке P свелась к вычислению интеграла (3.76). Распределение интенсивности (освещенности) в дифракционной картине определяется квадратом модуля светового возмущения $U(P)$. Следовательно, экспоненциальный множитель перед интегралом никакого влияния на распределение интенсивности в дифракционной картине не оказывает, поскольку по модулю равен единице.

Вычисление интеграла (3.76) существенно упростится, если в выражении (3.79) пренебречь квадратичными членами и членами более высоких порядков относительно ξ и η . В этом случае дифракцию, описываемую интегралом (3.76), называют дифракцией Фраунгофера. Если же квадратичными членами пренебречь нельзя, — то дифракцией Френеля. Строго говоря, члены второго и более высоких порядков обращаются в ноль только в том предельном случае, когда $r_0 \rightarrow \infty$ и $s_0 \rightarrow \infty$, т.е. когда и источник, и точка наблюдения расположены в

бесконечности. При этом надо допустить, что отношение $\frac{A}{r_0 s_0}$ в рассматриваемом случае равно конечной величине.

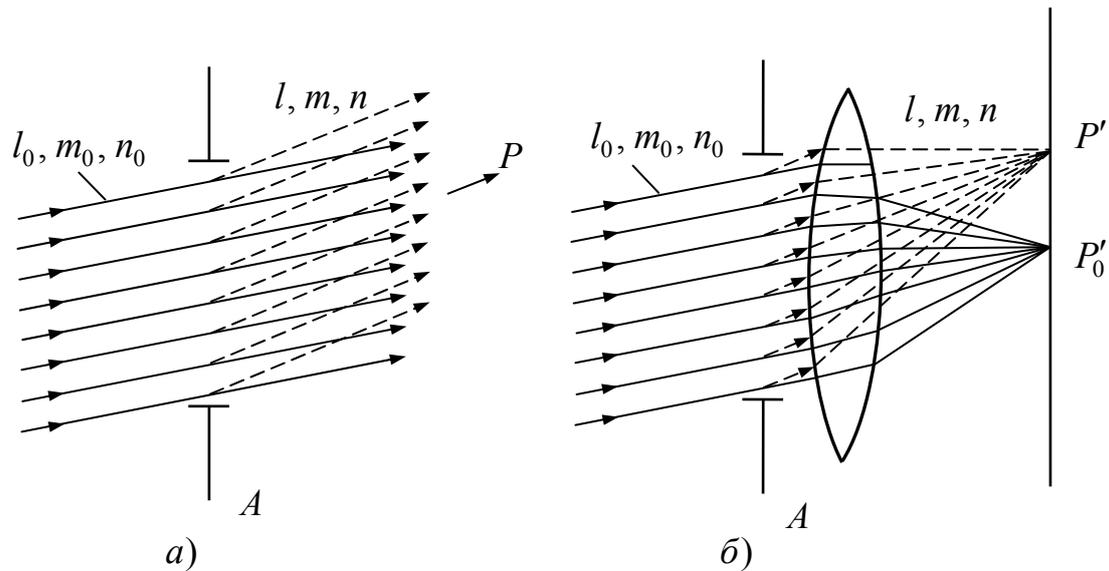


Рис. 3.13. Сравнение двух случаев дифракции Фраунгофера

На рис. 3.13а пучок лучей от бесконечно удаленной точки падает на отверстие в направлении, определяемом направляющими косинусами l_0, m_0, n_0 . Можно считать, что дифракция, наблюдаемая в направлении l, m, n в очень удаленной точке P , возникла в результате суперпозиции плоских волн, исходящих из каждой точки отверстия в этом направлении. Такие волны (не существующие в рамках геометрической оптики) можно назвать дифрагированными волнами, а соответствующие волновые нормали – дифрагированными лучами. Если теперь за экраном поместить хорошо коррегированный объектив, как показано на рис. 3.13б, то весь свет, дифрагированный в направлении (l, m, n) , соберется в точке P' в фокальной плоскости объектива. Так как длины оптических путей всех лучей, приходящих в точку P' от волнового фронта дифрагированного пучка лучей, равны, то, по сути дела, интерференционные эффекты остаются такими же, как и для исходного направления пучка лучей при условии, что объектив достаточно велик, чтобы не вносить дополнительной дифракции.

В выражении (3.79) имеем

$$\begin{aligned} (l_0\xi + m_0\eta)^2 &= l_0^2\xi^2 + 2l_0m_0\xi\eta + m_0^2\eta^2 = \\ &= (l_0^2 + m_0^2)(\xi^2 + \eta^2) - (l_0\eta - m_0\xi)^2 \leq (l_0^2 + m_0^2)(\xi^2 + \eta^2) \end{aligned}$$

и, соответственно, $(l\xi + m\eta)^2 \leq (l^2 + m^2)(\xi^2 + \eta^2)$.

В большинстве практически важных случаев $l_0^2 + m_0^2 \ll 1$ и $l^2 + m^2 \ll 1$, а поэтому для описания дифракции достаточно пользоваться приближением:

$$f(\xi, \eta) = (l_0 - l)\xi + (m_0 - m)\eta + \frac{1}{2} \left(\frac{1}{r_0} + \frac{1}{s_0} \right) (\xi^2 + \eta^2). \quad (3.80)$$

Такое приближение называют приближением Френеля.

Заметим, что в выражение (3.80) четыре величины l_0, m_0, l и m входят только в комбинации

$$p = l - l_0, \quad q = m - m_0.$$

Следовательно, в той области, где справедливо приближение Френеля, дифракционная картина не изменится, если отверстие сместится в своей собственной плоскости. При этом дифракционный интеграл Френеля-Кирхгофа можно записать в виде:

$$U(P) = C \iint_{\Sigma} \psi(\xi, \eta) \exp[-ik(p\xi + q\eta)] d\xi d\eta, \quad (3.81)$$

где C – комплексная постоянная величина

$$C = -\frac{ik \cos \delta \exp[ik(r_0 + s_0)]}{2\pi r_0 s_0}; \quad \psi(\xi, \eta) = A \exp \frac{ik}{2} \left(\frac{1}{r_0} + \frac{1}{s_0} \right) (\xi^2 + \eta^2).$$

В полученном выражении интегрирование предполагается в пределах площади Σ отверстия. Однако, учитывая, что в точках непрозрачных частей экрана $\psi(\xi, \eta) = 0$, интегрирование по ξ и η можно производить в пределах от $-\infty$ до ∞ . При этом дифракционная картина Френеля с точностью до комплексной постоянной величины определяется преобразованием Фурье функции $\psi(\xi, \eta)$:

$$U(P) = C \iint_{-\infty}^{\infty} \psi(\xi, \eta) \exp[-ik(p\xi + q\eta)] d\xi d\eta.$$

В приближении Френеля квадратичные члены относительно ξ и η отсутствуют при $\frac{1}{r_0} + \frac{1}{s_0} = 0$. (3.82)

Условие (3.82) означает, что дифракция Фраунгофера имеет место и тогда, когда точка наблюдения находится в плоскости, параллельной плоскости отверстия при условии, что точка наблюдения источник света достаточно близки к оси z . При этом возможны два случая. Если r_0 отрицательно, то падающие на отверстие волновые фронты имеют вогнутость в направлении распространения и точка P_0 является центром схождения, а не расхождения падающей волны. Этот случай очень важен для практики, так как осуществляется в пространстве изображений хорошо скорректированной центрированной системы, изображающей точечный источник, расположенный вблизи оси. Дифракционная картина Фраунгофера образуется в параксиальной плоскости изображений и может рассматриваться как результат дифракции волнового фронта на выходном зрачке.

Пусть $s_0 = -r_0 = R'$. Рассматривая изображение осевой точки предмета, можно принять $x_0 = 0$ и $y_0 = 0$, при этом $\cos \delta \approx 1$. Обозначив $x = x'$, $y = y'$, получаем $p = \frac{x}{s_0} = \frac{x'}{R'}$, $q = \frac{y}{s_0} = \frac{y'}{R'}$. При этом дифракционный интеграл, описывающий дифракцию Фраунгофера в рассматриваемом случае, принимает вид:

$$U(x', y') = -\frac{i}{\lambda R'^2} \iint_{\Sigma} \psi(\xi, \eta) \exp\left[-ik \frac{x'\xi + y'\eta}{R'}\right] d\xi d\eta, \quad (3.83)$$

где $\psi(\xi, \eta) = A$.

Если r_0 положительно, то волновые фронты имеют выпуклости в направлении распространения. При этом дифракционные картины мнимых изображений кажутся образованными на экране, проходящем через источник P_0 . Этот случай имеет место, например, тогда когда отверстие в экране находится непосредственно перед глазом наблюдателя.

Распределение светового возмущения в дифракционной картине мало отличается от распределения, определяемого дифракцией Фраунгофера, если в приближении Френеля выражение $\frac{1}{2} \left(\frac{1}{r_0} + \frac{1}{s_0} \right) (\xi^2 + \eta^2) \leq \frac{1}{4} \lambda$. При

этом

$$\frac{1}{r_0} + \frac{1}{s_0} \leq \frac{1}{2} \frac{\lambda}{\xi^2 + \eta^2}, \text{ где } r_0 + s_0 \neq 0.$$

$$\text{При } r_0 = s_0 = R \geq 4 \frac{\rho^2}{\lambda}, \text{ где } \rho^2 = \xi^2 + \eta^2.$$

$$\text{При этом при } \lambda = 5 \cdot 10^{-4} \text{ мм: } R_{\min} = 8 \cdot 10^3 \rho^2 \text{ мм.}$$

$$\text{При } \rho = 10 \text{ мм: } R_{\min} = 8 \cdot 10^5 \text{ мм} = 800 \text{ м.}$$

$$\text{При } \rho = 1 \text{ мм: } R_{\min} = 8 \cdot 10^3 \text{ мм} = 8 \text{ м.}$$

$$\text{При } \rho = 0,1 \text{ мм: } R_{\min} = 80 \text{ мм} = 0,08 \text{ м.}$$

$$\text{При } r_0 \rightarrow \infty: s_0 \geq 2 \frac{\rho^2}{\lambda}.$$

$$\text{При этом при } \lambda = 5 \cdot 10^{-4} \text{ мм: } s_{0\min} = 4 \cdot 10^3 \rho^2 \text{ мм.}$$

3.2.5. Дифракция Фраунгофера на отверстиях разной формы

Прямоугольное отверстие

Рассмотрим прямоугольное отверстие со сторонами $2a$ и $2b$. Если начало координат O находится в центре прямоугольника, а оси $O\xi$ и $O\eta$ параллельны его сторонам (рис. 3.14), то дифракционный интеграл Фраунгофера принимает вид

$$u(P) = C \int_{-a}^a \int_{-b}^b \exp[-ik(p\xi + q\eta)] d\xi d\eta = C \int_{-a}^a \exp(-ikp\xi) d\xi \int_{-b}^b \exp(-ikq\eta) d\eta, \text{ где}$$

$$C = \frac{ikA}{2\pi R'^2}.$$

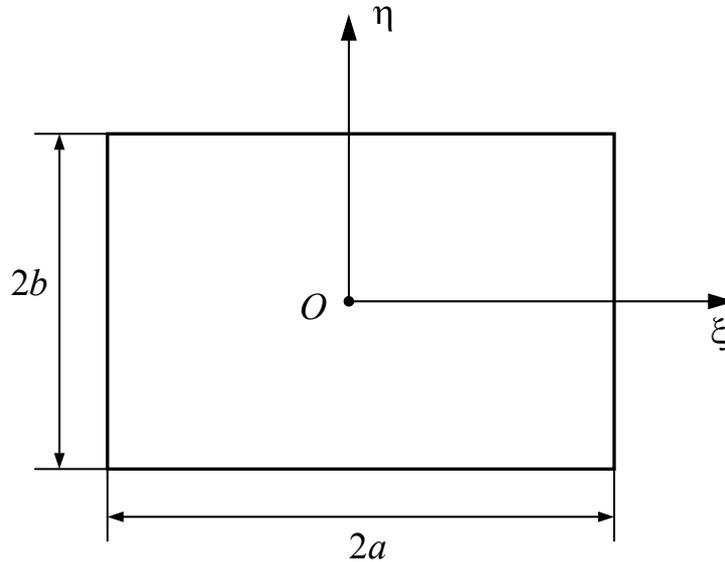


Рис. 3.14. Прямоугольное отверстие в плоском экране

Но

$$\int_{-a}^a \exp(-ikp\xi) d\xi = -\frac{1}{ikp} [\exp(-ikpa) - \exp(ikpa)] = 2 \frac{\sin kpa}{kp}.$$

Аналогично

$$\int_{-b}^b \exp(-ikq\eta) d\eta = 2 \frac{\sin kqb}{kq}.$$

При этом

$$u(P) = 4C \frac{\sin kpa}{kp} \frac{\sin kqb}{kq}.$$

При $p = 0, q = 0$: $u(0, 0) = 4Cab$. Тогда относительное распределение светового возмущения будет равно

$$\tilde{u}(P) = \frac{u(P)}{u(0, 0)} = \frac{\sin kpa}{kpa} \frac{\sin kqb}{kqb}.$$

При этом относительное распределение освещенности в дифракционном пятне изображения осевой точки определится выражением:

$$\tilde{E}(P) = \tilde{u}(P) \cdot \tilde{u}^*(P) = \left(\frac{\sin kpa}{kpa} \right)^2 \left(\frac{\sin kqb}{kqb} \right)^2.$$

$$\text{При } q = 0: \tilde{E}(P) = \tilde{E}(p) = \left(\frac{\sin kpa}{kpa} \right)^2.$$

$$\text{При } p = 0: \tilde{E}(q) = \left(\frac{\sin kqb}{kqb} \right)^2. \quad (3.84)$$

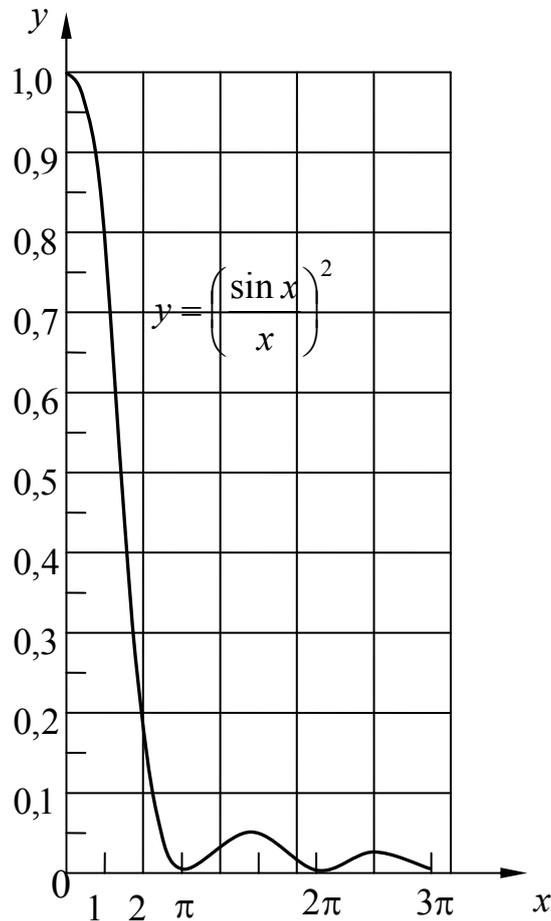


Рис. 3.15. Дифракционное изображение осевой точки при прямоугольной форме отверстия

График функции $y = \left(\frac{\sin x}{x} \right)^2$ приведен на рис. 3.15. Она имеет главный максимум $y=1$ при $x=0$ и минимумы, равные нулю, при $x = \pm\pi, \pm 2\pi, \pm 3\pi, \dots$. Эти минимумы разделяют вторичные максимумы, положение которых определяется из условия $\frac{dy}{dx} = 0$, т.е. из условия $\frac{dy}{dx} = \frac{\sin 2x}{x^3}(x - \operatorname{tg}x) = 0$. Отсюда следует, что положение вторичных максимумов определяется корнями уравнения $\operatorname{tg}x - x = 0$. Пять первых максимумов функции $y = \left(\frac{\sin x}{x} \right)^2$ приведены в таблице:

x	y
0	1
$1,430\pi = 4,493$	0,04718
$2,459\pi = 7,725$	0,01694
$3,470\pi = 10,90$	0,00834
$4,479\pi = 14,07$	0,00503

Из таблицы следует, что интенсивности в последовательности максимумов убывают очень быстро. Итак, интенсивность $\tilde{E}(P)$ равна нулю вдоль двух рядов линий, параллельных сторонам прямоугольника. Положение этих линий определяется из соотношений

$$kra = \pm \mu\pi, \quad kqb = \pm \nu\pi \quad (\mu, \nu = 1, 2, 3, \dots),$$

или, так как $p = l - l_0$, $q = m - m_0$, $k = \frac{2\pi}{\lambda}$,

$$l - l_0 = \pm \frac{\mu\lambda}{2a}, \quad m - m_0 = \pm \frac{\nu\lambda}{2b}.$$

Внутри каждого прямоугольника, образованного последовательными парами темных полос, интенсивности достигают максимумов, которые, однако, составляют лишь малую часть интенсивности центрального максимума и быстро уменьшаются по мере удаления от центра. Следует отметить, что большему размеру отверстия соответствуют меньшие эффективные размеры дифракционной картины.

Круглое отверстие

В этом случае удобно использовать полярные координаты вместо прямоугольных. Пусть (ρ, θ) – полярные координаты произвольной точки отверстия, т.е.

$$\rho \cos \theta = \xi, \quad \rho \sin \theta = \eta,$$

и пусть (ν, φ) – координаты точки P в дифракционной картине, относящейся к геометрическому изображению осевой точки источника, т.е. $\nu \cos \varphi = p$, $\nu \sin \varphi = q$.

Из определения p и q следует, что $\nu = \sqrt{p^2 + q^2}$ равно синусу угла между направлением (p, q) и центральным направлением $p = q = 0$. В таком случае, если a – радиус круглого отверстия, то дифракционный интеграл принимает вид

$$u(P) = C \int_0^a \int_0^{2\pi} \exp[-ik\rho\nu \cos(\theta - \varphi)] \rho d\rho d\theta.$$

Используя тождество

$$J_0(a) = \frac{1}{2\pi} \int_0^{2\pi} \exp[-ia \cos(\theta - \varphi)] d\theta,$$

получаем

$$u(P) = 2\pi C \int_0^a J_0(k\rho v) \rho d\rho.$$

Но $\frac{d}{dz} [z^{n+1} J_{n+1}(z)] = z^{n+1} J_n(z)$.

При $n = 0$: $d[zJ_1(z)] = zJ_0(z)dz$. А тогда

$$u(P) = \frac{2\pi C}{k^2 v^2} \int_0^a k v \rho J_0(k\rho v) d(k\rho v) = \frac{2\pi C}{k^2 v^2} k v a J_1(k a v) =$$

$$= 2\pi C a^2 \frac{J_1(k a v)}{k a v}.$$

При $v = 0$: $u(0) = \pi C a^2$. Тогда

$$\tilde{u}(P) = \frac{2J_1(k a v)}{k a v}. \text{ При этом}$$

$$\tilde{E}(P) = \left[\frac{2J_1(k a v)}{k a v} \right]^2. \tag{3.85}$$

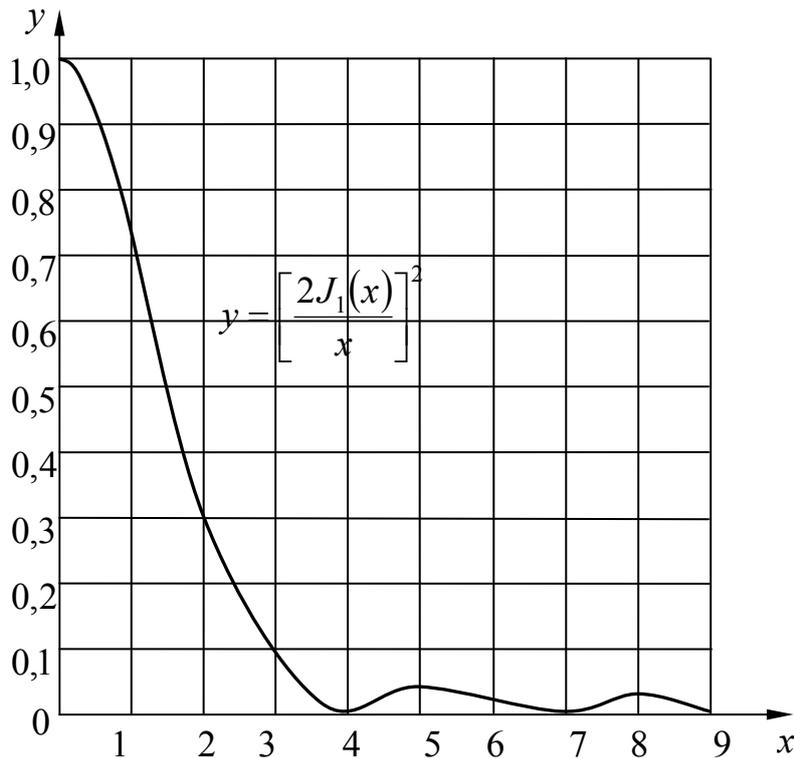


Рис. 3.16. Дифракционное изображение осевой точки при круглом отверстии в плоском экране

Распределение интенсивности в окрестности геометрического изображения точки описывается функцией $y = \left(\frac{2J_1(x)}{x} \right)^2$, график которой

приведен на рис. 3.16. Она имеет главный максимум $y=1$ при $x=0$ и с увеличением x осциллирует с постепенным уменьшением амплитуды подобно функции $\left(\frac{\sin x}{x}\right)^2$. Интенсивность равна нулю (минимум) при значениях x , определяемых корнями уравнения $J_1(x)=0$. Минимумы здесь уже не строго эквидистантны, что и следует из таблицы:

x	$\left(\frac{2J_1(x)}{x}\right)^2$	
0	1	максимум
$1,220\pi = 3,832$	0	минимум
$1,635\pi = 5,136$	0,0175	максимум
$2,233\pi = 7,016$	0	минимум
$2,679\pi = 8,417$	0,0042	максимум
$3,238\pi = 10,174$	0	минимум
$3,699\pi = 11,620$	0,0016	максимум

Положение вторичных максимумов определяется значениями x , удовлетворяющими уравнению

$$\frac{d}{dx} \left[\frac{J_1(x)}{x} \right] = 0,$$

или в соответствии с формулой $\frac{d}{dx} [x^{-n} J_n(x)] = -x^{-n} J_{n+1}(x)$ корнями уравнения $J_2(x)=0$. При увеличении x расстояния между последовательными минимумами или последовательными максимумами приближаются к π , как и в предыдущем случае.

Из полученных результатов следует, что наблюдаемая картина имеет вид светлого диска с центром в геометрическом изображении источника $p=q=0$, окруженного светлыми и темными кольцами. Интенсивность светлых колец быстро уменьшается с увеличением радиуса и обычно освещенность только одного или двух первых колец достаточна, чтобы их можно было наблюдать невооруженным глазом.

В рассматриваемом случае $x = kav$. При этом радиусы темных колец равны

$$v = \sqrt{p^2 + q^2} = \frac{x \lambda}{2\pi a} = 0,610 \frac{\lambda}{a}, 1,116 \frac{\lambda}{a}, 1,619 \frac{\lambda}{a}, \dots$$

Расстояние между двумя соседними кольцами асимптотически приближается к величине $\frac{\lambda}{2a}$. Здесь мы снова видим, что эффективные размеры дифракционной картины обратно пропорциональны линейным размерам отверстия.

Дифракционный интеграл Фраунгофера для круглого отверстия запишем в виде:

$$u(v, \varphi) = \frac{ik}{2\pi R'^2} \int_0^a \int_0^{2\pi} A \exp[-ik\rho v \cos(\theta - \varphi)] \rho d\rho d\theta.$$

Пусть $A = A(\rho, \theta)$. Из условия круговой симметрии рассматриваемой задачи следует, что

$$A(\rho, \theta) = A(\rho).$$

Пусть $D' = 2a$ – диаметр отверстия в экране. Тогда $\rho = a\rho'$, где $0 \leq \rho' \leq 1$. Обозначим

$$\tilde{\rho} = ka\rho' = k\rho. \text{ Тогда}$$

$$\begin{aligned} u(v, \varphi) &= \frac{i\lambda}{(2\pi R')^2} \int_0^{\tilde{\rho}_m} \int_0^{2\pi} A(\tilde{\rho}) \exp[-i\tilde{\rho}v \cos(\theta - \varphi)] \tilde{\rho} d\tilde{\rho} d\theta = \\ &= \frac{i}{kR'^2} \int_0^{\tilde{\rho}_m} A(\tilde{\rho}) J_0(\tilde{\rho}v) \tilde{\rho} d\tilde{\rho}. \end{aligned}$$

Здесь в весьма общем виде представлена зрачковая функция $G(\rho, \theta) = A(\tilde{\rho})$. Для ее аппроксимации удобно применить определенный интеграл Сонина:

$$J_{\mu+\nu+1}(x) = \frac{x^{\nu+1}}{2^\nu \Gamma(\nu+1)} \int_0^{\pi/2} J_\mu(x \sin \gamma) \sin^{\mu+1} \gamma \cos^{2\nu+1} \gamma d\gamma.$$

Эта формула справедлива, когда действительные числа $\mu, \nu > -1$. В дифракционный интеграл входит функция Бесселя первого рода нулевого порядка. Поэтому в интеграле Сонина в интересах рассматриваемого вопроса естественно принять $\mu = 0$.

Тогда

$$\int_0^{\pi/2} J_0(x \sin \gamma) \sin \gamma \cos^{2\nu+1} \gamma d\gamma = 2^\nu \Gamma(\nu+1) \frac{J_{\nu+1}(x)}{x^{\nu+1}}.$$

Это выражение можно переписать в виде

$$\int_0^1 J_0(x \sin \gamma) \cos^{2\nu} \gamma \sin \gamma d(\sin \gamma) = 2^\nu \Gamma(\nu+1) \frac{J_{\nu+1}(x)}{x^{\nu+1}}.$$

Введем обозначения:

$$\tilde{\rho} = \tilde{\rho}_m \sin \gamma, \quad x = \nu \tilde{\rho}_m.$$

Тогда

$$\int_0^{\tilde{\rho}_m} J_0(\nu \tilde{\rho}) \left[1 - \left(\frac{\tilde{\rho}}{\tilde{\rho}_m} \right)^2 \right]^\nu \tilde{\rho} d\tilde{\rho} = \tilde{\rho}_m^2 2^\nu \Gamma(\nu+1) \frac{J_{\nu+1}(\nu \tilde{\rho}_m)}{(\nu \tilde{\rho}_m)^{\nu+1}}.$$

Аппроксимируем зрачковую функцию выражением

$$G(\tilde{\rho}) = A_0 \sum_v q_v \left[1 - \left(\frac{\tilde{\rho}}{\tilde{\rho}_m} \right)^2 \right]^v.$$

Подставив это выражение в дифракционный интеграл и сопоставив полученный результат с преобразованным интегралом Сонина получаем

$$u(v) = \tilde{\rho}_m^2 \sum_v 2^v \Gamma(v+1) q_v \frac{J_{v+1}(v\tilde{\rho}_m)}{(v\tilde{\rho}_m)^{v+1}}.$$

Для кольцевой формы отверстия диаметр внутреннего экрана равен $\eta D'$, где η – коэффициент линейного экранирования.

В соответствии с принципом Бабинне

$$\begin{aligned} u(v) &= \frac{i}{kR'^2} \int_0^{\tilde{\rho}_m} G(\tilde{\rho}) J_0(\tilde{\rho}v) \tilde{\rho} d\tilde{\rho} - \frac{i}{kR'^2} \int_0^{\eta\tilde{\rho}_m} G(\tilde{\rho}) J_0(\tilde{\rho}v) \tilde{\rho} d\tilde{\rho} = \\ &= \frac{i}{kR'^2} \int_{\eta\tilde{\rho}_m}^{\tilde{\rho}_m} G(\tilde{\rho}) J_0(\tilde{\rho}v) \tilde{\rho} d\tilde{\rho}. \end{aligned}$$

$$\text{Пусть } G(\tilde{\rho}) = A_0 \left\{ q_0 + q_1 \left[1 - \left(\frac{\tilde{\rho}}{\tilde{\rho}_m} \right)^2 \right] \right\}.$$

При этом

$$\begin{aligned} u(v) &= \frac{iA_0}{kR'^2} \int_{\eta\tilde{\rho}_m}^{\tilde{\rho}_m} \left\{ q_0 + q_1 \left[1 - \left(\frac{\tilde{\rho}}{\tilde{\rho}_m} \right)^2 \right] \right\} J_0(\tilde{\rho}v) \tilde{\rho} d\tilde{\rho} = \\ &= \frac{iA_0(q_0 + q_1)}{kR'^2} \int_{\eta\tilde{\rho}_m}^{\tilde{\rho}_m} J_0(\tilde{\rho}v) \tilde{\rho} d\tilde{\rho} - \frac{iA_0q_1}{kR'^2\tilde{\rho}_m^2} \int_{\eta\tilde{\rho}_m}^{\tilde{\rho}_m} J_0(\tilde{\rho}v) \tilde{\rho}^3 d\tilde{\rho}. \end{aligned}$$

Поскольку $d[zJ_1(z)] = zJ_0(z)dz$, то

$$\begin{aligned} \frac{1}{v^2} \int_{\eta\tilde{\rho}_m}^{\tilde{\rho}_m} J_0(\tilde{\rho}v) (\tilde{\rho}v) d(\tilde{\rho}v) &= \frac{1}{v^2} (\tilde{\rho}v) J_1(\tilde{\rho}v) \Big|_{\eta\tilde{\rho}_m}^{\tilde{\rho}_m} = \\ &= \tilde{\rho}^2 \frac{J_1(\tilde{\rho}v)}{\tilde{\rho}v} \Big|_{\eta\tilde{\rho}_m}^{\tilde{\rho}_m} = \frac{1}{2} \tilde{\rho}_m^2 \left[\frac{2J_1(\tilde{\rho}_m v)}{\tilde{\rho}_m v} - \eta^2 \frac{2J_1(\eta\tilde{\rho}_m v)}{\eta\tilde{\rho}_m v} \right]. \end{aligned}$$

Применяя так называемое интегрирование по частям, получаем

$$\begin{aligned} \frac{1}{v^4} \int_{\eta\tilde{\rho}_m}^{\tilde{\rho}_m} J_0(\tilde{\rho}v) (\tilde{\rho}v)^3 d(\tilde{\rho}v) &= \frac{1}{v^4} (\tilde{\rho}v)^3 J_1(\tilde{\rho}v) \Big|_{\eta\tilde{\rho}_m}^{\tilde{\rho}_m} - \\ &- \frac{2}{v^4} \int_{\eta\tilde{\rho}_m}^{\tilde{\rho}_m} (\tilde{\rho}v)^2 J_1(\tilde{\rho}v) d(\tilde{\rho}v) = \tilde{\rho}_m^4 \left[\frac{J_1(\tilde{\rho}_m v)}{\tilde{\rho}_m v} - \eta^4 \frac{J_1(\eta\tilde{\rho}_m v)}{\eta\tilde{\rho}_m v} \right] - \end{aligned}$$

$$-\frac{2(\tilde{\rho}v)^2}{v^4} J_2(\tilde{\rho}v) \Big|_{\eta\tilde{\rho}_m}^{\tilde{\rho}_m} = \tilde{\rho}_m^4 \left[\frac{J_1(\tilde{\rho}_m v)}{\tilde{\rho}_m v} - \eta^4 \frac{J_1(\eta\tilde{\rho}_m v)}{\eta\tilde{\rho}_m v} \right] -$$

$$- 2\tilde{\rho}_m^4 \left[\frac{J_2(\tilde{\rho}_m v)}{\tilde{\rho}_m^2 v^2} - \eta^4 \frac{J_2(\eta\tilde{\rho}_m v)}{\eta^2 \tilde{\rho}_m^2 v^2} \right].$$

Подставив в дифракционный интеграл, получаем

$$u(v) = \frac{iA_0}{kR'^2} \tilde{\rho}_m^2 \left\{ \frac{q_0 + q_1}{2} \left[\frac{2J_1(v\tilde{\rho}_m)}{v\tilde{\rho}_m} - \eta^2 \frac{2J_1(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} \right] - \right.$$

$$\left. - \frac{q_1}{2} \left[\frac{2J_1(v\tilde{\rho}_m)}{v\tilde{\rho}_m} - \eta^4 \frac{2J_1(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} - \frac{4J_2(v\tilde{\rho}_m)}{v^2 \tilde{\rho}_m^2} + \eta^4 \frac{4J_2(v\eta\tilde{\rho}_m)}{v^2 \eta^2 \tilde{\rho}_m^2} \right] \right\}.$$

Но $2 \frac{v}{z} J_\nu(z) = J_{\nu+1}(z) + J_{\nu-1}(z)$, т.е. в нашем случае

$\frac{4}{z} J_2(z) = J_3(z) + J_1(z)$. А тогда

$$u(v) = \frac{iA_0}{kR'^2} \tilde{\rho}_m^2 \left\{ \frac{q_0 + q_1}{2} \left[\frac{2J_1(v\tilde{\rho}_m)}{v\tilde{\rho}_m} - \eta^2 \frac{2J_1(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} \right] - \right.$$

$$\left. - \frac{q_1}{2} \left[\frac{2J_1(v\tilde{\rho}_m)}{v\tilde{\rho}_m} - \eta^4 \frac{2J_1(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} - \frac{J_3(v\tilde{\rho}_m)}{v\tilde{\rho}_m} - \frac{J_1(v\tilde{\rho}_m)}{v\tilde{\rho}_m} + \right. \right.$$

$$\left. \left. + \eta^4 \frac{J_3(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} + \eta^4 \frac{J_1(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} \right] \right\} =$$

$$= \frac{iA_0}{2kR'^2} \tilde{\rho}_m^2 \left\{ (q_0 + q_1) \left[\frac{2J_1(v\tilde{\rho}_m)}{v\tilde{\rho}_m} - \eta^2 \frac{2J_1(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} \right] - \right.$$

$$\left. - \frac{q_1}{2} \left[\frac{2J_1(v\tilde{\rho}_m)}{v\tilde{\rho}_m} - \frac{2J_3(v\tilde{\rho}_m)}{v\tilde{\rho}_m} \right] + \frac{q_1}{2} \eta^4 \left[\frac{2J_1(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} - \frac{2J_3(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} \right] \right\}.$$

При $v = 0$:

$$u(0) = \frac{iA_0}{2kR'^2} \tilde{\rho}_m^2 \left[(q_0 + q_1)(1 - \eta^2) - \frac{q_1}{2}(1 - \eta^4) \right] =$$

$$= \frac{iA_0}{2kR'^2} \tilde{\rho}_m^2 (1 - \eta^2) \left[q_0 + q_1 - \frac{q_1}{2}(1 + \eta^2) \right] =$$

$$= \frac{iA_0}{2kR'^2} \tilde{\rho}_m^2 (1 - \eta^2) \left[q_0 + \frac{q_1}{2}(1 - \eta^2) \right].$$

При этом относительное распределение освещенности в изображении точки определится выражением:

$$\begin{aligned} \tilde{E}(v) &= \left[\frac{u(v)}{u(0)} \right]^2 = \frac{1}{(1-\eta^2)^2 \left[q_0 + \frac{q_1}{2}(1-\eta^2) \right]^2} \times \\ &\times \left\{ (q_0 + q_1) \left[\frac{2J_1(v\tilde{\rho}_m)}{v\tilde{\rho}_m} - \eta^2 \frac{2J_1(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} \right] - \right. \\ &\left. - \frac{q_1}{2} \left[\frac{2J_1(v\tilde{\rho}_m)}{v\tilde{\rho}_m} - \frac{2J_3(v\tilde{\rho}_m)}{v\tilde{\rho}_m} \right] + \frac{q_1}{2} \eta^4 \left[\frac{2J_1(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} - \frac{2J_3(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} \right] \right\}^2. \end{aligned}$$

Легко убедиться, что при $\tilde{\rho} = 0$: $G(0) = A_0(q_0 + q_1)$, а при $\tilde{\rho} = \tilde{\rho}_m$: $G(\tilde{\rho}_m) = A_0q_0$. Естественно положить $G(0) = A_0(q_0 + q_1) = A_0$. Отсюда $q_0 + q_1 = 1$. Тогда $q_1 = 1 - q_0$. При этом

$$\begin{aligned} \tilde{E}(v) &= \frac{1}{(1-\eta^2)^2 \left[q_0 + \frac{1-q_0}{2}(1-\eta^2) \right]^2} \left\{ \frac{2J_1(v\tilde{\rho}_m)}{v\tilde{\rho}_m} - \eta^2 \frac{2J_1(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} - \right. \\ &\left. - \frac{1-q_0}{2} \left[\frac{2J_1(v\tilde{\rho}_m)}{v\tilde{\rho}_m} - \frac{2J_3(v\tilde{\rho}_m)}{v\tilde{\rho}_m} \right] + \frac{1-q_0}{2} \eta^4 \left[\frac{2J_1(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} - \frac{2J_3(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} \right] \right\}^2. \end{aligned}$$

Пусть на краю отверстия $G(\tilde{\rho}_m) = A_0q_0 = 0$, т.е. $q_0 = 0$. При этом

$$\begin{aligned} \tilde{E}(v) &= \frac{4}{(1-\eta^2)^4} \left\{ \frac{2J_1(v\tilde{\rho}_m)}{v\tilde{\rho}_m} - \eta^2 \frac{2J_1(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} - \right. \\ &\left. - \frac{1}{2} \left[\frac{2J_1(v\tilde{\rho}_m)}{v\tilde{\rho}_m} - \frac{2J_3(v\tilde{\rho}_m)}{v\tilde{\rho}_m} - \eta^4 \left(\frac{2J_1(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} - \frac{2J_3(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} \right) \right] \right\}^2. \end{aligned} \quad (3.86)$$

Пусть $G(\tilde{\rho}) \equiv A_0$, т.е. $q_0 = 1$, а $q_1 = 0$.

При этом

$$\tilde{E}(v) = \frac{1}{(1-\eta^2)^2} \left[\frac{2J_1(v\tilde{\rho}_m)}{v\tilde{\rho}_m} - \eta^2 \frac{2J_1(v\eta\tilde{\rho}_m)}{v\eta\tilde{\rho}_m} \right]^2. \quad (3.87)$$

При $\eta = 0$:

$$\tilde{E}(v) = \left[\frac{2J_1(v\tilde{\rho}_m)}{v\tilde{\rho}_m} \right]^2 \quad \text{или} \quad \tilde{E}(v) = \left[\frac{2J_1(kav)}{kav} \right]^2. \quad (3.88)$$

Эллиптическое отверстие

Запишем уравнение окружности в виде

$$x^2 + y^2 = r^2.$$

Пусть $x = \frac{x'}{\mu}$; $y = y'$. Подставив в уравнение окружности, получаем $\frac{x'^2}{\mu^2 r^2} + \frac{y'^2}{r^2} = 1$, т.е. получили уравнение эллипса $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, где $a = \mu r$, а $b = r$. Таким образом, эллипс представляет собой окружность, растянутую или сжатую в одном из направлений.

Рассмотрим решение задачи в общем случае. Пусть Σ_1 и Σ_2 – два отверстия, причем в некотором направлении (0ξ) размеры Σ_2 в μ раз больше, чем Σ_1 . Для дифракции Фраунгофера на Σ_1 имеем

$$u_1(p, q) = C \iint_{\Sigma_1} \exp[-ik(p\xi + q\eta)] d\xi d\eta.$$

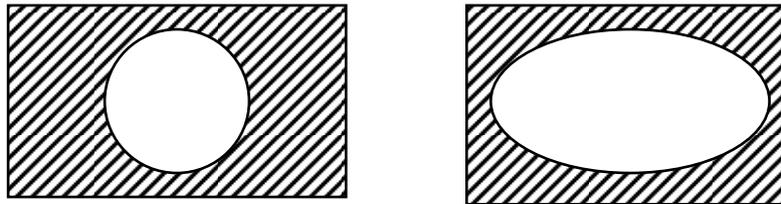
Аналогично для дифракции Фраунгофера на Σ_2 находим $u_2(p, q) = C \iint_{\Sigma_2} \exp[-ik(p\xi' + q\eta')] d\xi' d\eta'$. Заменяя здесь переменные интегрирования (ξ', η') на (ξ, η) , где $\xi' = \mu\xi$, $\eta' = \eta$, получим

$$\begin{aligned} u_2(p, q) &= \mu C \iint_{\Sigma_2} \exp[-ik(\mu p\xi + q\eta)] d\xi d\eta = \\ &= \mu C \iint_{\Sigma_1} \exp[-ik(\tilde{p}\xi + q\eta)] d\xi d\eta = \mu u_1(\tilde{p}, q) \end{aligned}$$

или

$$u_2(p, q) = \mu u_1(\mu p, q).$$

Отверстие



Дифракционная картина

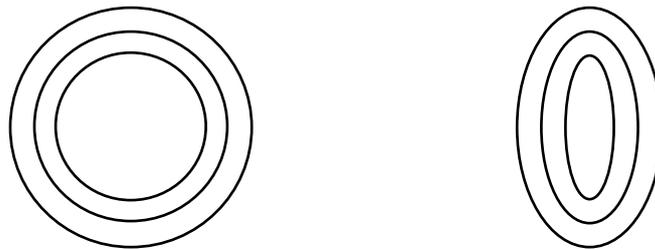


Рис. 3.17. Зависимость вида дифракционного изображения точки от формы отверстия

Отсюда следует, что если отверстие симметрично расширяется в каком-либо направлении в μ раз, то дифракционная картина Фраунгофера сжимается в том же направлении в μ раз, а интенсивность в некоторой точке новой картины становится в μ^2 раз больше интенсивности в соответствующей точке первоначальной картины. Используя этот результат, можно, например, сразу найти картину дифракции Фраунгофера от эллиптического отверстия из картины дифракции от круга, как показано на рис. 3.17.

3.2.6. Дифракция Фраунгофера на системе одинаковых и одинаково ориентированных отверстий

Зрачковую функцию экрана, состоящего из N отверстий можно представить в виде

$$G(\xi, \eta) = \sum_{j=1}^N G_j(\tilde{\xi}, \tilde{\eta}). \quad (3.89)$$

Начала координатных осей (ξ, η) и (x', y') расположены на общей нормали к плоскостям экрана и рассматриваемой дифракционной картины. Пусть система координат $(\tilde{\xi}, \tilde{\eta})$, начало которой совпадает с центром j -го отверстия, смещена относительно системы осей (ξ, η) в точку с координатами (ξ_{0j}, η_{0j}) . Тогда

$$\begin{aligned} u_j(p, q) &= \frac{ik}{2\pi R'^2} \iint_{\Sigma_j} G_j(\xi, \eta) \exp[-ik(p\xi + q\eta)] d\xi d\eta = \\ &= \frac{ik}{2\pi R'^2} \exp[-ik(p\xi_{0j} + q\eta_{0j})] \iint_{\Sigma_j} G_j(\tilde{\xi}, \tilde{\eta}) \exp[-ik(p\tilde{\xi} + q\tilde{\eta})] d\tilde{\xi} d\tilde{\eta}. \end{aligned}$$

Введем обозначения

$$\frac{\xi_{0j}}{R'} = \alpha_{0j}; \quad \frac{\tilde{\xi}}{R'} = \tilde{\alpha}; \quad \frac{\eta_{0j}}{R'} = \beta_{0j}; \quad \frac{\tilde{\eta}}{R'} = \tilde{\beta}; \quad kx' = \tilde{x}'; \quad ky' = \tilde{y}'; \quad \left(k = \frac{2\pi}{\lambda} \right).$$

Предположим, что область j -го отверстия ограничена прямоугольником, при этом

$$G_j(\tilde{\xi}, \tilde{\eta}) = A_j(\tilde{\xi}, \tilde{\eta}) = A_j(0, 0) = A_{0j}.$$

Тогда при $x = y = 0$ имеем

$$\begin{aligned}
u_j(\tilde{x}', \tilde{y}') &= \frac{i}{\lambda} A_{0j} \exp[-i(\alpha_{0j}\tilde{x}' + \beta_{0j}\tilde{y}')] \int_{-\alpha_j}^{\alpha_j} \exp(-i\tilde{\alpha}\tilde{x}') d\tilde{\alpha} \times \\
&\times \int_{-\beta_j}^{\beta_j} \exp(-i\tilde{\beta}\tilde{y}') d\tilde{\beta} = \frac{i}{\lambda} 4\alpha_j\beta_j A_{0j} \exp[-i(\alpha_{0j}\tilde{x}' + \beta_{0j}\tilde{y}')] \times \\
&\times \frac{\sin \alpha_j \tilde{x}' \sin \beta_j \tilde{y}'}{\alpha_j \tilde{x}' \beta_j \tilde{y}'}.
\end{aligned} \tag{3.90}$$

Для всей системы отверстий в целом

$$\begin{aligned}
u(\tilde{x}', \tilde{y}') &= \frac{4i}{\lambda} \sum_{j=1}^N A_{0j} \alpha_j \beta_j \frac{\sin \alpha_j \tilde{x}' \sin \beta_j \tilde{y}'}{\alpha_j \tilde{x}' \beta_j \tilde{y}'} \times \\
&\times \exp[-i(\alpha_{0j}\tilde{x}' + \beta_{0j}\tilde{y}')].
\end{aligned} \tag{3.91}$$

Пусть $\alpha_1 = \alpha_2 = \dots = \alpha_j = \dots = \alpha_N = \alpha$;

$\beta_1 = \beta_2 = \dots = \beta_j = \dots = \beta_N = \beta$; $A_{01} = A_{02} = \dots = A_{0j} = \dots = A_{0N} = A_0$.

Тогда

$$\begin{aligned}
u(\tilde{x}', \tilde{y}') &= \frac{4}{\lambda} A_0 \alpha \beta \frac{\sin \alpha \tilde{x}' \sin \beta \tilde{y}'}{\alpha \tilde{x}' \beta \tilde{y}'} \times \\
&\times \sum_{j=1}^N [\sin(\alpha_{0j}\tilde{x}' + \beta_{0j}\tilde{y}') + i \cos(\alpha_{0j}\tilde{x}' + \beta_{0j}\tilde{y}')].
\end{aligned}$$

При этом распределение освещенности в дифракционном пятне определится выражением вида

$$\begin{aligned}
E(\tilde{x}', \tilde{y}') &= u(\tilde{x}', \tilde{y}') u^*(\tilde{x}', \tilde{y}') = \left(\frac{4}{\lambda} A_0 \alpha \beta \frac{\sin \alpha \tilde{x}' \sin \beta \tilde{y}'}{\alpha \tilde{x}' \beta \tilde{y}'} \right)^2 \times \\
&\times \left\{ \left[\sum_{j=1}^N \sin(\alpha_{0j}\tilde{x}' + \beta_{0j}\tilde{y}') \right]^2 + \left[\sum_{j=1}^N \cos(\alpha_{0j}\tilde{x}' + \beta_{0j}\tilde{y}') \right]^2 \right\}.
\end{aligned} \tag{3.92}$$

При $N = 1$

$$E(\tilde{x}', \tilde{y}') = E_0(\tilde{x}', \tilde{y}') = \left(\frac{4}{\lambda} A_0 \alpha \beta \frac{\sin \alpha \tilde{x}' \sin \beta \tilde{y}'}{\alpha \tilde{x}' \beta \tilde{y}'} \right)^2.$$

$$\text{При } \tilde{x}' = 0; \tilde{y}' = 0: E(0, 0) = N^2 \left(\frac{4}{\lambda} A_0 \alpha \beta \right)^2 = N^2 E_0(0, 0).$$

Следует обратить внимание, что в результате перераспределения энергии освещенность в центре дифракционной картины увеличивается в N^2 раз.

При этом

$$\tilde{E}(\tilde{x}', \tilde{y}') = \Phi_M(\alpha_{0j}\tilde{x}', \beta_{0j}\tilde{y}') \tilde{E}_0(\tilde{x}', \tilde{y}'),$$

где $\tilde{E}_0(\tilde{x}', \tilde{y}') = \text{sinc}^2(\alpha\tilde{x}')\text{sinc}^2(\beta\tilde{y}')$; $\text{sinc}(z) = \frac{\sin z}{z}$; $\Phi_M(\alpha_{0j}\tilde{x}', \beta_{0j}\tilde{y}')$ – функция амплитудной модуляции пятна рассеяния в дифракционной картине изображения осевой точки, образованной отдельным отверстием экрана, при этом

$$\begin{aligned} N^2\Phi_M(\alpha_{0j}\tilde{x}', \beta_{0j}\tilde{y}') &= \left[\sum_{j=1}^N \sin(\alpha_{0j}\tilde{x}' + \beta_{0j}\tilde{y}') \right]^2 + \\ &+ \left[\sum_{j=1}^N \cos(\alpha_{0j}\tilde{x}' + \beta_{0j}\tilde{y}') \right]^2; \\ 0 &\leq \Phi_M(\alpha_{0j}\tilde{x}', \beta_{0j}\tilde{y}') \leq 1. \end{aligned} \quad (3.93)$$

Легко видеть, что $\tilde{E}_0(\tilde{x}', \tilde{y}') = 0$ при $\alpha\tilde{x}'_m = \pm\mu\pi$; $\beta\tilde{y}'_m = \pm\nu\pi$, где $\mu, \nu = 1, 2, \dots$. Отсюда $\tilde{x}'_m = \pm\frac{\mu}{\alpha}\pi$; $\tilde{y}'_m = \pm\frac{\nu}{\beta}\pi$. При этом будем считать

$$\tilde{x}' = \pm\frac{\mu}{\alpha}\pi\tilde{x}''; \quad \tilde{y}' = \pm\frac{\nu}{\beta}\pi\tilde{y}'', \quad \text{где } 0 \leq \tilde{x}'' \leq 1; \quad 0 \leq \tilde{y}'' \leq 1.$$

Тогда

$$\begin{aligned} \tilde{E}_0(\tilde{x}'', \tilde{y}'') &= \text{sinc}^2(\mu\pi\tilde{x}'')\text{sinc}^2(\nu\pi\tilde{y}''); \\ N^2\Phi_M(\tilde{\alpha}_{0j}\tilde{x}'', \tilde{\beta}_{0j}\tilde{y}'') &= \left[\sum_{j=1}^N \sin\left(\frac{\alpha_{0j}}{\alpha}\mu\pi\tilde{x}'' + \frac{\beta_{0j}}{\beta}\nu\pi\tilde{y}''\right) \right]^2 + \\ &+ \left[\sum_{j=1}^N \cos\left(\frac{\alpha_{0j}}{\alpha}\mu\pi\tilde{x}'' + \frac{\beta_{0j}}{\beta}\nu\pi\tilde{y}''\right) \right]^2. \end{aligned} \quad (3.94)$$

Пусть $\mu = \nu = 1$. При этом

$$\begin{aligned} \tilde{E}_0(\tilde{x}'', \tilde{y}'') &= \text{sinc}^2(\pi\tilde{x}'')\text{sinc}^2(\pi\tilde{y}''); \\ N^2\Phi_M(\tilde{\alpha}_{0j}\tilde{x}'', \tilde{\beta}_{0j}\tilde{y}'') &= \left[\sum_{j=1}^N \sin\left(\frac{\alpha_{0j}}{\alpha}\pi\tilde{x}'' + \frac{\beta_{0j}}{\beta}\pi\tilde{y}''\right) \right]^2 + \\ &+ \left[\sum_{j=1}^N \cos\left(\frac{\alpha_{0j}}{\alpha}\pi\tilde{x}'' + \frac{\beta_{0j}}{\beta}\pi\tilde{y}''\right) \right]^2. \end{aligned}$$

Пусть при этом $N = 2$. Полагая $\alpha_{01} = -\alpha_{02} = \alpha_0$; $\beta_{01} = -\beta_{02} = \beta_0$, получаем

$$\Phi_M(\tilde{\alpha}_0\tilde{x}'', \tilde{\beta}_0\tilde{y}'') = \frac{1 + \cos\left[2\pi\left(\frac{\alpha_0}{\alpha}\tilde{x}'' + \frac{\beta_0}{\beta}\tilde{y}''\right)\right]}{2}.$$

Пусть $y'' = 0$. Тогда

$$\Phi_M(\tilde{\alpha}_0 \tilde{x}'') = \frac{1 + \cos\left(2\pi \frac{\alpha_0}{\alpha} \tilde{x}''\right)}{2}.$$

Пусть $\Phi_M(\tilde{\alpha}_0 \tilde{x}'') = 0$. При этом $\cos\left(2\pi \frac{\alpha_0}{\alpha} \tilde{x}''\right) = -1$. Отсюда

$$2\pi \frac{\alpha_0}{\alpha} \tilde{x}'' = (2m+1)\pi.$$

Пусть $\tilde{x}'' = 1$. Тогда $2\pi \frac{\alpha_0}{\alpha} = (2m+1)\pi$.

Отсюда $\frac{\alpha_0}{\alpha} = m + \frac{1}{2}$. При $m = 0$: $\frac{\alpha_0}{\alpha} = \frac{1}{2}$, т.е. это решение физически не реализуемо.

При $m = 1$: $\frac{\alpha_0}{\alpha} = \frac{1}{2}$ и т.д.

Пусть $\alpha_j \rightarrow 0$, $\beta_j \rightarrow 0$. При этом

$$\frac{\sin \alpha_j \tilde{x}'}{\alpha_j \tilde{x}'} \rightarrow 1, \quad \frac{\sin \beta_j \tilde{y}'}{\beta_j \tilde{y}'} \rightarrow 1.$$

Будем считать, что

$$S_j = 4\alpha_j \beta_j = \lim_{\substack{\Delta\alpha_{0j} \rightarrow 0 \\ \Delta\beta_{0j} \rightarrow 0}} \Delta\alpha_{0j} \Delta\beta_{0j} = dS.$$

А тогда

$$u(\tilde{x}', \tilde{y}') = \frac{i}{\lambda} \lim_{\substack{\Delta\alpha_{0j} \rightarrow 0 \\ \Delta\beta_{0j} \rightarrow 0 \\ N \rightarrow \infty}} \sum_{j=1}^N A_{0j} \exp[-i(\alpha_{0j} \tilde{x}' + \beta_{0j} \tilde{y}')] \Delta\alpha_{0j} \Delta\beta_{0j}.$$

Положив $A_{0j} = A_0$, находим

$$\begin{aligned} u(\tilde{x}', \tilde{y}') &= \frac{i}{\lambda} A_{0j} \int_{-\alpha}^{\alpha} \exp(-i\alpha_{0j} \tilde{x}') d\alpha_{0j} \int_{-\beta}^{\beta} \exp(-i\beta_{0j} \tilde{y}') d\beta_{0j} = \\ &= \frac{4i}{\lambda} A_0 \alpha \beta \frac{\sin \alpha \tilde{x}'}{\alpha \tilde{x}'} \frac{\sin \beta \tilde{y}'}{\beta \tilde{y}'}, \end{aligned}$$

т.е. получили выражение, определяющее распределение светового возмущения в дифракционной картине при сплошном отверстии прямоугольной формы в экране.

Предположим, что область j -го отверстия ограничена окружностью, при этом удобно ввести полярную систему координат

$$\begin{aligned} \xi &= a\rho' \cos \theta, \\ \eta &= a\rho' \sin \theta, \quad \rho' = \sqrt{\frac{\xi^2 + \eta^2}{a^2}}, \quad 0 \leq \rho' \leq 1, \quad a = \frac{D'}{2}; \end{aligned}$$

$$x' = r' \cos \varphi, \quad y' = r' \sin \varphi, \quad r' = \sqrt{x'^2 + y'^2} = R' \sqrt{p^2 + q^2}.$$

Здесь a – радиус круглого экрана с отверстиями; ρ', θ – полярные координаты в плоскости экрана; r', φ – полярные координаты в плоскости наблюдения дифракционной картины.

При этом выражение, определяющее световое возмущение в дифракционной картине изображения осевой точки, примет вид

$$u(r', \varphi) = \frac{ika^2}{2\pi R'^2} \iint_S G(\rho', \theta) \times \exp\left[-ik \frac{a}{R'} r' \rho' \cos(\theta - \varphi)\right] \rho d\rho' d\theta. \quad (3.95)$$

Обратимся к рис. 3.18, на котором показана точка Q_j j -го отверстия в экране в системе полярных координат (ρ', θ) и (ρ'_j, θ_j) . Полюс O системы координат (ρ', θ) и полюс системы координат (r', φ) лежат на одной прямой, принятой в качестве оси, а полюс C системы координат (ρ'_j, θ_j) совмещен с центром j -го отверстия, положение которого в системе координат (ρ', θ) определено координатами $(\rho'_{0j}, \theta_{0j})$, при этом полярные оси обеих систем координат параллельны.

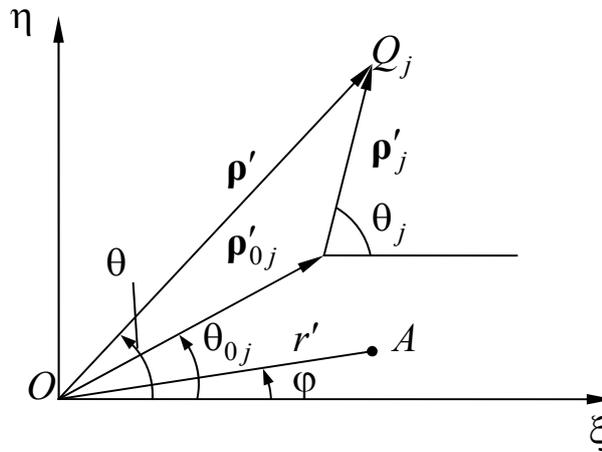


Рис. 3.18. Координаты элемента системы отверстий круглой формы

На этом же рисунке показана точка A плоскости изображения, положение которой определяется координатами (r', φ) . Из рисунка следует, что

$$\rho' \cos(\theta - \varphi) = \rho'_{0j} \cos(\theta_{0j} - \varphi) + \rho'_j \cos(\theta_j - \varphi).$$

$$\text{Обозначим } \theta_{0j} - \varphi = \psi_{0j}, \quad \theta_j - \varphi = \psi_j.$$

Выполнив соответствующую подстановку с учетом принятых обозначений, получаем

$$u_j(r') = \frac{ika_j^2}{2\pi R'^2} \exp\left(-ik \frac{a}{R'} r' \rho'_{0j} \cos \psi_{0j}\right) \times \int \int_{S_j} G(\rho'_j, \psi_j) \exp\left(-ik \frac{a_j}{R'} r' \rho'_j \cos \psi_j\right) \rho'_j d\rho'_j d\psi_j, \quad (3.96)$$

где a – радиус круглого экрана с отверстиями; при этом $0 \leq \rho'_{0j} \leq 1$; a_j – радиус j -го отверстия экрана, при этом $0 \leq \rho'_j \leq 1$.

Обозначим $\frac{a_j}{R'} = A_j$; $kA_j r' = \tilde{r}'_j$; $\tilde{\rho}'_{0j} = \frac{a}{a_j} \rho'_{0j}$, при этом $0 \leq \tilde{\rho}'_{0j} \leq \frac{a}{a_j}$.

Положив $G(\rho'_j, \psi_j) = \frac{u_j(\rho'_j, \psi_j)}{R'} = \frac{u_j(0, 0)}{R'} = \frac{u_{0j}}{R'}$,

где u_{0j} – амплитуда светового возмущения на расстоянии единицы длины от источника, получаем

$$u_j(\tilde{r}'_j) = i \frac{R'}{\lambda} A_j^2 u_{0j} \exp(-i\tilde{r}'_j \tilde{\rho}'_{0j} \cos \psi_{0j}) \times \int_0^1 \int_0^{2\pi} \exp(-i\tilde{r}'_j \rho'_j \cos \psi_j) \rho'_j d\rho'_j d\psi_j, \\ u_j(\tilde{r}'_j, \psi_j) = i u_{0j} \frac{R'}{\lambda} A_j^2 \exp(-i\tilde{r}'_j \tilde{\rho}'_{0j} \cos \psi_{0j}) \times \int_0^1 \rho'_j d\rho'_j \int_0^{2\pi} \exp(-i\tilde{r}'_j \rho'_j \cos \psi_j) d\psi_j. \quad (3.97)$$

Напомним, что

$$\frac{1}{2\pi} \int_0^{2\pi} \exp(-iz \cos \psi) d\psi = J_0(z), \text{ а } \frac{d}{dz} [zJ_1(z)] = zJ_0(z).$$

Применив эти соотношения, получаем

$$u_j(\tilde{r}'_j) = iku_{0j} R' A_j^2 \exp(-i\tilde{r}'_j \tilde{\rho}'_{0j} \cos \psi_{0j}) \int_0^1 J_0(\tilde{r}'_j \rho'_j) \rho'_j d\rho'_j = \\ = \frac{i}{2} ku_{0j} R' A_j^2 \frac{2J_1(\tilde{r}'_j)}{\tilde{r}'_j} \exp(-i\tilde{r}'_j \tilde{\rho}'_{0j} \cos \psi_{0j}).$$

При заполнении отверстий в экране когерентным излучением распределение светового возмущения в дифракционной картине определится выражением вида

$$u(r') = \frac{1}{2} kR' \left[\sum_{j=1}^N u_{0j} A_j^2 \frac{2J_1(\tilde{r}'_j)}{\tilde{r}'_j} \sin(\tilde{r}'_j \tilde{\rho}'_{0j} \cos \psi_{0j}) + \right. \\ \left. + i \sum_{j=1}^N u_{0j} A_j^2 \frac{2J_1(\tilde{r}'_j)}{\tilde{r}'_j} \cos(\tilde{r}'_j \tilde{\rho}'_{0j} \cos \psi_{0j}) \right]. \quad (3.98)$$

При этом соответствующее распределение освещенности в дифракционной картине равно произведению взаимно сопряженных величин $u(r')u^*(r')$, т.е.

$$E(r') = \frac{1}{4} k^2 R'^2 \left\{ \left[\sum_{j=1}^N u_{0j} A_j^2 \frac{2J_1(\tilde{r}'_j)}{\tilde{r}'_j} \sin(\tilde{r}'_j \tilde{\rho}'_{0j} \cos \psi_{0j}) \right]^2 + \right. \\ \left. + \left[\sum_{j=1}^N u_{0j} A_j^2 \frac{2J_1(\tilde{r}'_j)}{\tilde{r}'_j} \cos(\tilde{r}'_j \tilde{\rho}'_{0j} \cos \psi_{0j}) \right]^2 \right\}. \quad (3.99)$$

При $N = 1$:

$$E(r') = E_0(r) = \left[\frac{1}{2} kR' u_0 A^2 \frac{2J_1(\tilde{r}')}{\tilde{r}'} \right]^2.$$

Пусть $A_1 = A_2 = \dots = A_j = \dots = A$, $u_{01} = u_{02} = \dots = u_{0j} = \dots = u_0$.

Тогда

$$E(r') = \left[\frac{1}{2} kR' u_0 A^2 \frac{2J_1(\tilde{r}')}{\tilde{r}'} \right]^2 \left\{ \left[\sum_{j=1}^N \sin(\tilde{r}' \tilde{\rho}'_{0j} \cos \psi_{0j}) \right]^2 + \right. \\ \left. + \left[\sum_{j=1}^N \cos(\tilde{r}' \tilde{\rho}'_{0j} \cos \psi_{0j}) \right]^2 \right\}. \quad (3.100)$$

При $r' = 0$:

$$E(0) = N^2 \left(\frac{1}{2} kR' u_0 A^2 \right)^2 = N^2 E_0(0).$$

При этом

$$\tilde{E}(r') = \Phi_M(\tilde{r}', \tilde{\rho}'_{0j}, \psi_{0j}) \tilde{E}_0(r'),$$

где $\tilde{E}_0(r') = \left[\frac{2J_1(\tilde{r}')}{\tilde{r}'} \right]^2$; $\Phi_M(\tilde{r}', \tilde{\rho}'_{0j}, \psi_{0j})$ – функция амплитудной модуляции

пятна рассеяния в дифракционной картине изображения осевой точки, образованной отдельным отверстием экрана, при этом

$$N^2 \Phi_M(\tilde{r}', \tilde{\rho}'_{0j}, \psi_{0j}) = \left[\sum_{j=1}^N \sin(\tilde{r}' \tilde{\rho}'_{0j} \cos \psi_{0j}) \right]^2 + \left[\sum_{j=1}^N \cos(\tilde{r}' \tilde{\rho}'_{0j} \cos \psi_{0j}) \right]^2; \quad (3.101)$$

$$0 \leq \Phi_M(\tilde{r}', \tilde{\rho}'_{0j}, \psi_{0j}) \leq 1.$$

Пусть $N = 2$. Полагая $\tilde{\rho}'_{01} = \tilde{\rho}'_{02} = \tilde{\rho}'_0$, получаем

$$\tilde{E}(r') = \Phi_M(\tilde{r}', \tilde{\rho}'_0, \psi_{01}, \psi_{02}) \tilde{E}_0(r'),$$

где $\Phi_M(\tilde{r}', \tilde{\rho}'_0, \psi_{01}, \psi_{02}) = \frac{1 + \cos[\tilde{r}' \tilde{\rho}'_0 (\cos \psi_{02} - \cos \psi_{01})]}{2}$.

Легко видеть, что из условия $\cos \psi_{02} - \cos \psi_{01} = 0$ следует $\psi_{01} = \pm \psi_{02} = \psi_0$. При этом $\Phi_M(\tilde{r}', \tilde{\rho}'_0, \psi_0) = 1$, а $\tilde{E}(r') = \tilde{E}_0(r')$. Практический интерес представляет условие $\psi_{01} = \psi_0$; $\psi_{02} = \psi_0 + \pi$. При этом

$$\Phi_M(\tilde{r}', \tilde{\rho}'_0, \psi_0) = \frac{1 + \cos(2\tilde{r}' \tilde{\rho}'_0 \cos \psi_0)}{2}.$$

Отсюда следует, что при $\psi_0 = \frac{\pi}{2}$ $\tilde{E}(r') = \tilde{E}_0(r') = \left[\frac{2J_1(\tilde{r}')}{\tilde{r}'} \right]^2$.

При $\psi_0 = 0$: $\Phi_M(\tilde{r}', \tilde{\rho}'_0) = \frac{1 + \cos(2\tilde{r}' \tilde{\rho}'_0)}{2}$,

где $0 \leq \Phi_M(\tilde{r}', \tilde{\rho}'_0) \leq 1$.

Наглядное представление влияния функции $\Phi_M(\tilde{r}', \tilde{\rho}'_0)$ на распределение освещенности в изображении точки дает следующий пример.

Пусть $\tilde{E}(r') = 0$. Это возможно при следующих условиях:

- $\tilde{E}_0(r') = 0$;
- $\Phi_M(\tilde{r}', \tilde{\rho}'_0) = 0$;
- $\Phi_M(\tilde{r}', \tilde{\rho}'_0) = \tilde{E}_0(r') = 0$.

Пусть при $\tilde{r}' = \tilde{r}'_{ок}$ выполняется условие $\tilde{E}_0(r'_{ок}) = \left[\frac{2J_1(\tilde{r}'_{ок})}{\tilde{r}'_{ок}} \right]^2 = 0$.

Пусть при этом и $\Phi_M(\tilde{r}'_{ок}, \tilde{\rho}'_0) = 0$. Тогда $2\tilde{r}'_{ок} \tilde{\rho}'_0 = (1 + 2q)\pi$, где $q = 0, 1, 2, \dots$. Отсюда следует, что $\tilde{\rho}'_0 = \tilde{\rho}'_{0q} = \frac{1 + 2q}{2\tilde{r}'_{ок}} \pi$.

При $\tilde{\rho}'_0 = \tilde{\rho}'_{0q}$ из условия $\Phi_M(\tilde{r}'_{0t}, \tilde{\rho}'_{0q}) = 0$ находим

$$\tilde{r}'_{0t} = \frac{1 + 2t}{2\tilde{\rho}'_{0q}} \pi,$$

где $t = 0, 1, 2, \dots, q$. В результате соответствующей подстановки получаем

$$\tilde{r}'_{0t} = \frac{1+2t}{1+2q} \tilde{r}'_{ок}.$$

При $t=0$ $\tilde{r}'_{0t} = \frac{\tilde{r}'_{ок}}{1+2q}$, а при $t=q$ $\tilde{r}'_{0t} = \tilde{r}'_{ок}$.

При $t=q=0$: $\tilde{r}'_{0t} = \tilde{r}'_0 = 3,832$.

В общем случае распределение освещенности в дифракционной картине осевой точки определяется выражением (3.98) при условии одновременного изменения углов φ_{0j} на одну и ту же величину φ , т.е.

$$E(r', \varphi) = \frac{1}{4} k^2 R'^2 \left\{ \left[\sum_{j=1}^N u_{0j} A_j^2 \frac{2J_1(\tilde{r}'_j)}{\tilde{r}'_j} \sin(\tilde{r}'_j \tilde{\rho}'_{0j} \cos(\psi_{0j} + \varphi)) \right]^2 + \right. \\ \left. + \left[\sum_{j=1}^N u_{0j} A_j^2 \frac{2J_1(\tilde{r}'_j)}{\tilde{r}'_j} \cos(\tilde{r}'_j \tilde{\rho}'_{0j} \cos(\psi_{0j} + \varphi)) \right]^2 \right\}, \quad (3.102)$$

где $0 \leq \varphi \leq 2\pi$.

При $u_{01} = u_{02} = \dots = u_{0j} = \dots = u_0$; $A_1 = A_2 = \dots = A_j = \dots = A$:

$\tilde{E}(r', \varphi) = \Phi_M(\tilde{r}', \tilde{\rho}'_0, \psi_0, \varphi) E_0(r')$, где

$$\Phi_M(\tilde{r}', \tilde{\rho}'_0, \psi_0, \varphi) = \left\{ \sum_{j=1}^N \sin[\tilde{r}'_j \tilde{\rho}'_{0j} \cos(\psi_{0j} + \varphi)] \right\}^2 + \\ + \left\{ \sum_{j=1}^N \cos[\tilde{r}'_j \tilde{\rho}'_{0j} \cos(\psi_{0j} + \varphi)] \right\}^2. \quad (3.103)$$

3.2.7. Дифракционный интеграл при наличии aberrаций

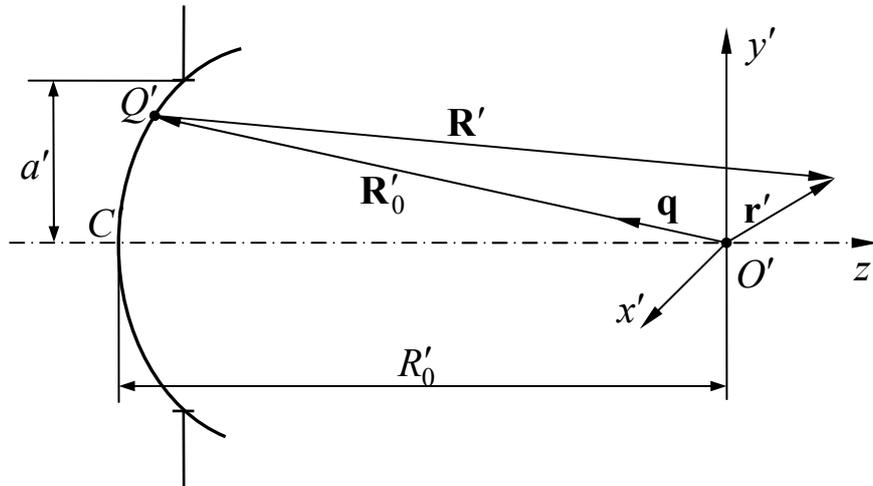


Рис. 3.19. Световое возмущение в дифракционном изображении точки

Пусть сферический монохроматический волновой фронт падает на круглое отверстие и сходится в осевой точке O' . Рассмотрим световое возмущение $u(P')$ в произвольной точке P' вблизи точки O' . Положение точки P' относительно точки O' определим вектором \mathbf{r}' . Будем считать, что расстояние $r' = O'P'$ и радиус a' ($\gg \lambda$) отверстия малы по сравнению с радиусом $R'_0 = O'C'$ волнового фронта L' , который заполняет это отверстие, как показано на рис. 3.19. Обозначим через R' расстояние от точки Q' волнового фронта L' до точки наблюдения P' , а через $\frac{A}{R'_0}$ – световое возмущение (амплитуду) падающего волнового фронта в точке Q' . Тогда, применяя принцип Гюйгенса-Френеля, получаем

$$u(P') = \frac{i}{\lambda} \frac{A \exp(-ikR'_0)}{R'_0} \iint_S \frac{\exp(ikR')}{R'} dS. \quad (3.104)$$

Поскольку рассматриваемые расстояния r' предполагаются малыми, вариации коэффициента наклона по волновому фронту пренебрежимо малы. В соответствии с рис. 3.19 имеем

$$\mathbf{R}'_0 + \mathbf{R}' = \mathbf{r}'.$$

Обозначая через \mathbf{q} единичный вектор в направлении $O'Q'$, получаем

$$\mathbf{R}'_0 + \mathbf{q} \cdot \mathbf{R}' = \mathbf{q} \cdot \mathbf{r}'.$$

Но $\mathbf{q} \cdot \mathbf{R}' = R' \cos(180^\circ - \angle \mathbf{q}, \mathbf{R}') = -R' \cos(\mathbf{q}, \mathbf{R}')$.

С хорошим приближением можно принять $\cos(\mathbf{q}, \mathbf{R}') = 1$. Тогда $\mathbf{q} \cdot \mathbf{R}' = -R'$. При этом

$$\mathbf{R}' - \mathbf{R}'_0 = -\mathbf{q} \cdot \mathbf{r}'. \quad (3.105)$$

Элемент dS волнового фронта равен

$$dS = R'_0{}^2 d\Omega,$$

где $d\Omega$ – элемент телесного угла, под которым элемент dS виден из точки O' . Без заметной ошибки можно заменить в знаменателе подынтегрального выражения R' на R'_0 . При этом

$$u(P') = \frac{i}{\lambda} A \int_{\Omega} \exp(-ik\mathbf{q} \cdot \mathbf{r}') d\Omega. \quad (3.106)$$

Теперь интегрирование производится по всему телесному углу Ω , под которым отверстие видно из точки O' . Полученное выражение представляет собой интеграл Дебая и выражает поле как результат суперпозиции плоских волн, распространяющихся во всех направлениях (определяемых векторами \mathbf{q}), попадающих в Ω .

Для вычисления интеграла представим подынтегральное выражение в более явном виде. Примем за начало декартовых координат точку O' , при этом $P' = P'(x', y', z')$, а $Q' = Q'(\xi', \eta', \zeta')$. Будем считать, что

$$\xi' = a' \rho \sin \theta, \quad x' = r' \sin \varphi,$$

$$\eta' = a' \rho \cos \theta; \quad y' = r' \cos \varphi.$$

Так как точка Q' лежит на сферическом волновом фронте L , имеем

$$\zeta' = -\sqrt{R_0'^2 - a'^2 \rho^2} = -R_0' \left(1 - \frac{1}{2} \frac{a'^2 \rho^2}{R_0'^2} - \dots \right).$$

Следовательно,

$$\mathbf{q} \cdot \mathbf{r}' = \frac{x' \xi' + y' \eta' + z' \zeta'}{R_0'} = \frac{a' \rho r' \cos(\theta - \varphi)}{R_0'} - z' \left(1 - \frac{1}{2} \frac{a'^2 \rho^2}{R_0'^2} - \dots \right).$$

Введем безразмерные переменные u и v , которые в сочетании с углом φ определяют положение точки P' , а именно

$$u' = \frac{2\pi}{\lambda} \left(\frac{a'}{R_0'} \right)^2 z', \quad v' = \frac{2\pi}{\lambda} \left(\frac{a'}{R_0'} \right) r' = \frac{2\pi}{\lambda} \frac{a'}{R_0'} \sqrt{x'^2 + y'^2}.$$

Если члены, содержащие отношение $\frac{\rho}{R_0'}$ в степени, превышающей вторую, пренебрежимо малы по сравнению с единицей, то

$$k\mathbf{q} \cdot \mathbf{r}' = v' \rho \cos(\theta - \varphi) - \left(\frac{R_0'}{a'} \right)^2 u' + \frac{1}{2} u' \rho^2.$$

Кроме того, элемент телесного угла $d\Omega$ равен

$$d\Omega = \frac{dS}{R_0'^2} = \frac{a'^2 \rho d\rho d\theta}{R_0'^2}.$$

С учетом этих соотношений получаем

$$u(P') = \frac{i}{\lambda} \frac{a'^2 A}{R_0'^2} \exp \left[i \left(\frac{R_0'}{a'} \right)^2 u' \right] \int_0^1 \int_0^{2\pi} \exp \left\{ -i \left[v' \rho \cos(\theta - \varphi) + \frac{1}{2} u' \rho^2 \right] \right\} \rho d\rho d\theta. \quad (3.107)$$

Рассмотрим центрированную оптическую систему, образующую изображение точечного источника монохроматического излучения P_0 , как показано на рис. 3.20. Выберем начало декартовой системы координат в месте паракиального изображения точки P_0 , т.е. в точке P'_0 , при этом ось z' направим вдоль луча $C'P'_0$, где C' – центр выходного зрачка. Ось y' расположим в меридиональной плоскости, т.е. в плоскости, в которой расположена точка P_0 и оптическая ось системы. Обозначим через y_0 и y'_0 расстояния от оси до точек P_0 и P'_0 соответственно. Деформацию волнового фронта (отклонение от сферы) в области выходного зрачка опишем функцией aberrаций или aberrационной функцией W . Пусть \tilde{Q}' и

Q' – точки пересечения луча в пространстве изображений с волновым фронтом, проходящим через точку C' , и с опорной сферой Гаусса соответственно. Если предположить, что показатель преломления среды в пространстве изображений равен единице, то волновая aberrация W определяется расстоянием $\tilde{Q}'Q'$, измеренным вдоль луча (на рисунке $W > 0$).

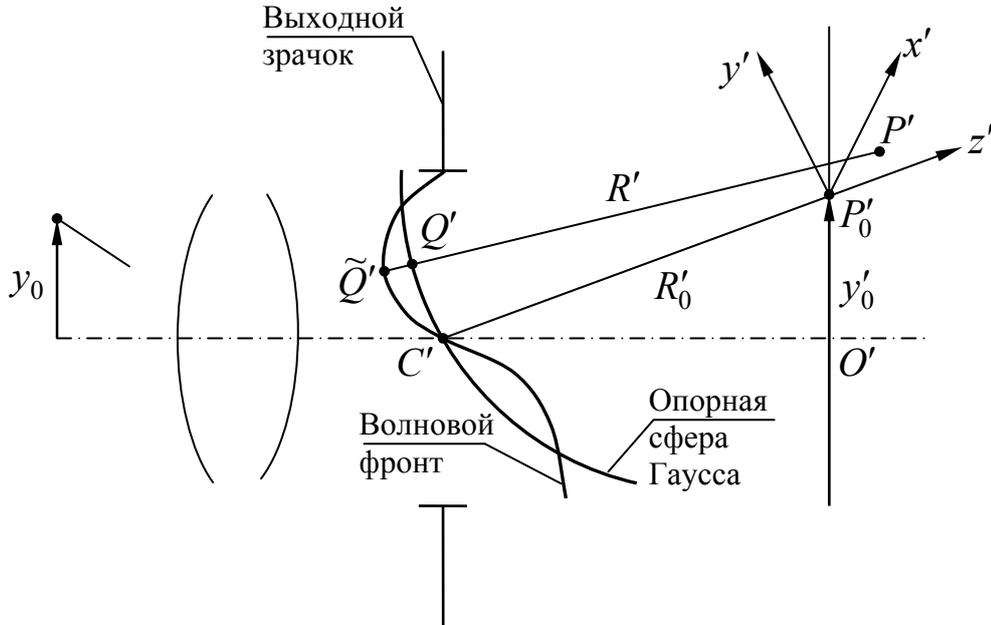


Рис. 3.20. Аберрации в изображении точки

Пусть R'_0 – радиус опорной сферы Гаусса, равный отрезку CP'_0 , R' – расстояние между точкой Q' и произвольной точкой P' , расположенной вблизи точки P'_0 . Возмущение в точке Q' описывается величиной

$$\frac{A \exp[ik(W - R'_0)]}{R'_0},$$

где $\frac{A}{R'_0}$ – амплитуда светового возмущения в точке Q' . Согласно принципу

Гюйгенса-Френеля возмущение в точке P' равно

$$u(P') = -\frac{i}{\lambda} \frac{A \exp(-ikR'_0)}{R'_0} \iint_S \frac{\exp[ik(W + R')]}{R'} dS, \quad (3.108)$$

где интегрирование проводится по той части опорной сферы Гаусса, которая открыта выходным зрачком, при этом предполагается, что смещение точки P' относительно точки P'_0 невелико, а поэтому вариациями коэффициента наклона по волновому фронту можно пренебречь. Кроме того, предполагаем, что амплитуда световых колебаний на поверхности волнового фронта практически постоянна, т.е. коэффициент A можно вынести из под знака интеграла.

Пусть a' – радиус выходного зрачка, а ξ', η', ζ' и x', y', z' – координаты точек Q' и P' соответственно, при этом

$$\xi' = a' \rho \sin \theta, \quad x' = r' \sin \varphi,$$

$$\eta' = a' \rho \cos \theta; \quad y' = r' \cos \varphi.$$

В этом случае, как было показано,

$$k(R' - R'_0) = -v' \rho \cos(\theta - \varphi) - \frac{1}{2} u' \rho^2 + \left(\frac{R'_0}{a'} \right) u',$$

где u' и v' – «оптические координаты» точки P' , т.е.

$$u' = \frac{2\pi}{\lambda} \left(\frac{a'}{R'_0} \right)^2 z', \quad v' = \frac{2\pi}{\lambda} \left(\frac{a'}{R'_0} \right) \sqrt{x'^2 + y'^2}.$$

При $z' = 0$: $k(R' - R'_0) = -v' \rho \cos(\theta - \varphi) = -k \frac{\xi' x' + \eta' y'}{R'_0}$. При этом

$$u(x', y') = -\frac{i}{\lambda} \frac{A}{R'_0} \iint_S \frac{\exp(ikW)}{R'} \exp\left(-ik \frac{\xi' x' + \eta' y'}{R'_0}\right) d\xi' d\eta'. \quad (3.109)$$

Рассмотрим величину W как функцию y'_0, ρ и θ : $W = W(y'_0, \rho, \theta)$.

Элемент поверхности опорной сферы Гаусса $dS = a^2 \rho d\rho d\theta$. Если угол, образованный лучом $C'P'_0$ с оптической осью системы, мал, то пределами интегрирования могут служить $0 \leq \rho \leq 1$, $0 \leq \theta \leq 2\pi$. Кроме того, для точек P' , расположенных вблизи точки P'_0 , величину R' , стоящую в знаменателе подынтегрального выражения, можно заменить на R'_0 . В результате получаем

$$\begin{aligned} u(P') &= u(u', v', \varphi) = \\ &= -\frac{i}{\lambda} \frac{A a'^2}{R_0'^2} \exp\left[i \left(\frac{R'_0}{a'} \right)^2 u' \int_0^1 \int_0^{2\pi} \exp\left\{ i \left[kW(y'_0, \rho, \theta) - \right. \right. \right. \\ &\quad \left. \left. \left. - v' \rho \cos(\theta - \varphi) - \frac{1}{2} u' \rho^2 \right] \right\} \rho d\rho d\theta, \end{aligned} \quad (3.110)$$

при этом интенсивность (освещенность) света в точке P' равна

$$\begin{aligned} E(P') &= u(P') u^*(P') = \\ &= \left(\frac{A a'^2}{\lambda R_0'^2} \right)^2 \left| \int_0^1 \int_0^{2\pi} \exp\left\{ i \left[kW(y'_0, \rho, \theta) - v' \rho \cos(\theta - \varphi) - \frac{1}{2} u' \rho^2 \right] \right\} \rho d\rho d\theta \right|^2. \end{aligned}$$

В точке параксиального изображения ($v' = 0, u' = 0$) при $W(y'_0, \rho, \theta) = 0$ имеем

$$E_0(P') = \left(\frac{A a'^2}{\lambda R_0'^2} \right)^2 \left| \int_0^1 \int_0^{2\pi} \rho d\rho d\theta \right|^2 = \pi^2 \left(\frac{A a'^2}{\lambda R_0'^2} \right)^2.$$

Тогда нормированная интенсивность определится выражением вида:

$$i(P') = \frac{E(P')}{E_0(P')} = \frac{1}{\pi^2} \left| \int_0^1 \int_0^{2\pi} \exp \left\{ i \left[kW(y'_0, \rho, \theta) - v' \rho \cos(\theta - \varphi) - \frac{1}{2} u' \rho^2 \right] \right\} \rho d\rho d\theta \right|^2.$$

В отсутствие aberrаций интенсивность света максимальна в точке параксиального изображения. При наличии же aberrаций максимум интенсивности в общем случае может сместиться из точки параксиального изображения в другую точку. Практический интерес представляет максимальная величина нормированной интенсивности и ее положение в определенной плоскости наблюдения. Эту величину называют числом Штреля или определительной яркостью.

Теорема смещения

Пусть W и W' – две функции aberrаций, причем

$$W' = W + H\rho^2 + K\rho \sin \theta + L\rho \cos \theta + M,$$

где H, K, L, M – постоянные порядка λ .

Выражение для нормированной интенсивности удобно записать в виде:

$$i(u', v', \varphi) = \frac{1}{\pi^2} \left| \int_0^1 \int_0^{2\pi} \exp [if(u', v', \varphi; \rho, \theta)] \rho d\rho d\theta \right|^2,$$

где $f(u', v', \varphi; \rho, \theta) = kW - v' \rho \cos(\theta - \varphi) - \frac{1}{2} u' \rho^2$. Последнее выражение

можно представить в виде:

$$\begin{aligned} f(u', v', \varphi; \rho, \theta) &= kW' - k[H\rho^2 + K\rho \sin \theta + L\rho \cos \theta + M] - \\ &- v' \rho \cos(\theta - \varphi) - \frac{1}{2} u' \rho^2 = \\ &= kW' - \tilde{v}' \rho \cos(\theta - \tilde{\varphi}) - \frac{1}{2} \tilde{u}' \rho^2 - kM = f'(\tilde{u}', \tilde{v}', \tilde{\varphi}; \rho, \theta) - kM, \end{aligned}$$

где $\tilde{u}' = u' + 2kH$, $\tilde{v}' \sin \tilde{\varphi} = v' \sin \varphi + kK$, $\tilde{v}' \cos \tilde{\varphi} = v' \cos \varphi + kL$.

$$\text{Но } \tilde{u}' = k \left(\frac{a'}{R'_0} \right)^2 \tilde{z}' = u' + 2kH = k \left(\frac{a'}{R'_0} \right)^2 z' + 2kH, \text{ т.е.}$$

$$\tilde{z}' = z' + 2 \left(\frac{R'_0}{a'} \right)^2 H; \quad \tilde{v}' = k \left(\frac{a'}{R'_0} \right) \sqrt{\tilde{x}'^2 + \tilde{y}'^2}, \text{ где}$$

$$\tilde{x}' = \tilde{r}' \sin \tilde{\varphi} = \frac{\tilde{v}'}{k} \left(\frac{R'_0}{a'} \right) \sin \tilde{\varphi} = \frac{1}{k} \left(\frac{R'_0}{a'} \right) v' \sin \varphi + \left(\frac{R'_0}{a'} \right) K = x' + \left(\frac{R'_0}{a'} \right) K;$$

$$\tilde{y}' = \tilde{r}' \cos \tilde{\varphi} = \frac{\tilde{v}'}{k} \left(\frac{R'_0}{a'} \right) \cos \tilde{\varphi} = \frac{1}{k} \left(\frac{R'_0}{a'} \right) v' \cos \varphi + \left(\frac{R'_0}{a'} \right) L = y' + \left(\frac{R'_0}{a'} \right) L.$$

Итак,

$$\begin{aligned} i(u', v', \varphi) &= \frac{1}{\pi^2} \left| \int_0^1 \int_0^{2\pi} \exp[if(u', v', \varphi; \rho, \theta)] \rho d\rho d\theta \right|^2 = \\ &= \frac{1}{\pi^2} \left| \exp(-ikM) \int_0^1 \int_0^{2\pi} \exp[if'(\tilde{u}', \tilde{v}', \tilde{\varphi}; \rho, \theta)] \rho d\rho d\theta \right|^2 = i(\tilde{u}', \tilde{v}', \tilde{\varphi}). \end{aligned}$$

Таким образом, доказана следующая теорема смещения: добавление к функции аберраций члена $H\rho^2 + K\rho \sin \theta + L\rho \cos \theta + M$, где H, K, L, M – постоянные порядка λ , не изменяет трехмерного распределения интенсивности света в параксиальном изображении точки, а только смещает его как целое в соответствии с преобразованиями:

$$\tilde{x}' = x' + \left(\frac{R'_0}{a'} \right) K, \quad \tilde{y}' = y' + \left(\frac{R'_0}{a'} \right) L, \quad \tilde{z}' = z' + 2 \left(\frac{R'_0}{a'} \right)^2 H. \text{ Иными словами,}$$

происходит смещение изображения точки на величину $2 \left(\frac{R'_0}{a'} \right)^2 H$ вдоль луча $C'P'_0$ (вдоль оси z') и на величины $\left(\frac{R'_0}{a'} \right) K$ и $\left(\frac{R'_0}{a'} \right) L$ вдоль направлений осей x' и y' соответственно.

Изменение опорной сферы

Аддитивные члены в правой части выражения

$$W' = W + H\rho^2 + K\rho \sin \theta + L\rho \cos \theta + M$$

можно рассматривать как величины, характеризующие изменение опорной сферы Гаусса. Предположим, что выбрана новая опорная сфера с центром в точке $\tilde{P}'_0(\tilde{x}', \tilde{y}', \tilde{z}')$ вблизи параксиального изображения точки P'_0 , при этом радиус кривизны ее равен $\tilde{R}' = N\tilde{P}'_0$, как показано на рис. 3.21, а расстояние от новой опорной сферы до опорной сферы Гаусса не превышает нескольких длин волн.

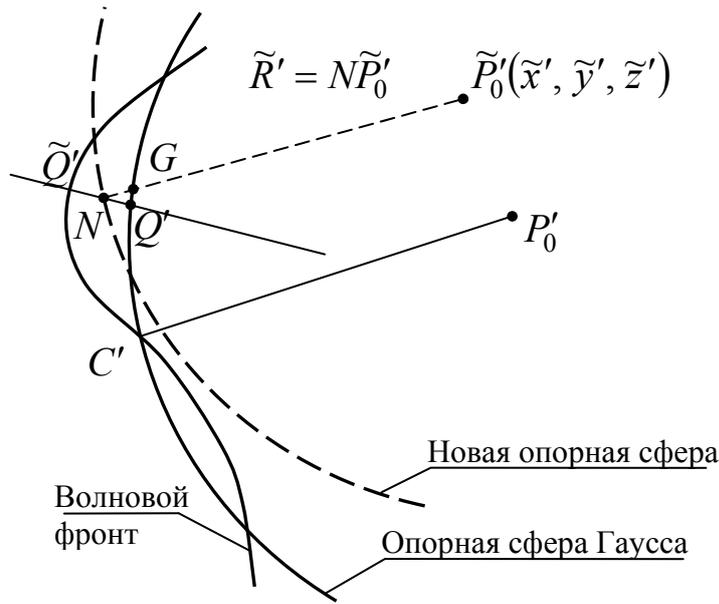


Рис. 3.21. Изменение опорной сферы

Пусть N – точка пересечения луча $\tilde{Q}'Q'$ с этой новой опорной сферой. Тогда функция aberrаций W' , отнесенная к новой сфере, в соответствии с рис. 3.21 равна $W' = \tilde{Q}'N = \tilde{Q}'Q' - NQ' \approx \tilde{Q}'Q' - NG$, где G – точка пересечения линии $N\tilde{P}'_0$ с опорной сферой Гаусса, при этом показатель преломления среды в пространстве изображений предполагается равным единице. Величина $\tilde{Q}'Q' = W$ равна волновой aberrации, отнесенной к опорной сфере Гаусса, а расстояние $NG = N\tilde{P}'_0 - G\tilde{P}'_0 = \tilde{R}' - R'$, где R' – расстояние между точками G и \tilde{P}'_0 . Тогда

$$\begin{aligned} W' &\approx W + R' - \tilde{R}' = W + R' - R'_0 + R'_0 - \tilde{R}' = \\ &= W + \frac{\lambda}{2\pi} \left[-\tilde{v}'\rho \cos(\theta - \tilde{\varphi}) - \frac{1}{2}\tilde{u}'\rho^2 + \left(\frac{R'_0}{a'}\right)^2 \tilde{u}' \right] + R'_0 - \tilde{R}'. \end{aligned}$$

$$\text{Напомним, что } \tilde{u}' = k \left(\frac{a'}{R'_0}\right)^2 \tilde{z}', \quad \tilde{v}' = k \left(\frac{a'}{R'_0}\right) \sqrt{\tilde{x}'^2 + \tilde{y}'^2};$$

$$\tilde{x}' = \tilde{r}' \sin \tilde{\varphi}, \quad \tilde{y}' = \tilde{r}' \cos \tilde{\varphi}.$$

Раскрыв последнее выражение, получаем

$$W' = W - \left(\frac{a'}{R'_0}\right) \sqrt{\tilde{x}'^2 + \tilde{y}'^2} \rho \cos \theta \cos \tilde{\varphi} - \left(\frac{a'}{R'_0}\right) \sqrt{\tilde{x}'^2 + \tilde{y}'^2} \rho \sin \theta \sin \tilde{\varphi} -$$

$$-\frac{1}{2}\left(\frac{a'}{R'_0}\right)^2 \tilde{z}' \rho^2 + \left(\frac{R'_0}{a'}\right)^2 \left(\frac{a'}{R'_0}\right)^2 \tilde{z}' + R'_0 - \tilde{R}' = W - \left(\frac{a'}{R'_0}\right) \tilde{r}' \cos \tilde{\varphi} \rho \cos \theta -$$

$$-\left(\frac{a'}{R'_0}\right) \tilde{r}' \sin \tilde{\varphi} \rho \sin \theta - \frac{1}{2}\left(\frac{a'}{R'_0}\right)^2 \tilde{z}' \rho^2 + \tilde{z}' + R'_0 - \tilde{R}'.$$

Итак,

$$W' = W + H\rho^2 + K\rho \sin \theta + L\rho \cos \theta + M,$$

где $H = -\frac{1}{2}\left(\frac{a'}{R'_0}\right)^2 \tilde{z}'$, $K = -\left(\frac{a'}{R'_0}\right) \tilde{x}'$, $L = -\left(\frac{a'}{R'_0}\right) \tilde{y}'$, $M = \tilde{z}' + R'_0 - \tilde{R}'$.

Формула Марешала

Световое возмущение в параксиальном изображении точки, т.е. при $x' = 0$ и $y' = 0$, в соответствии с формулой (3.110) определяется выражением

$$u(0, 0) = -\frac{i}{\lambda} \frac{Aa'^2}{R_0'^2} \exp\left[i\left(\frac{R'_0}{a'}\right)^2 u'\right] \int_0^1 \int_0^{2\pi} \exp(ikW) \rho d\rho d\theta, \quad (3.111)$$

где при $y'_0 = const$, $W(y'_0, \rho, \theta) = W(\rho, \theta)$.

Как известно, показательную функцию можно представить степенным рядом вида: $\exp x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots$

При этом выражение (3.111) можно представить в виде:

$$u(0, 0) = -\frac{i}{\lambda} \frac{Aa'^2}{R_0'^2} \exp\left[i\left(\frac{R'_0}{a'}\right)^2 u'\right] \int_0^1 \int_0^{2\pi} \left[1 + ikW + \frac{1}{2}(ikW)^2 + \dots\right] \rho d\rho d\theta.$$

Пусть \overline{W}^n – среднее значение n -й степени W , т.е.

$$\overline{W}^n = \frac{\int_0^1 \int_0^{2\pi} W^n \rho d\rho d\theta}{\int_0^1 \int_0^{2\pi} \rho d\rho d\theta} = \frac{1}{\pi} \int_0^1 \int_0^{2\pi} W^n \rho d\rho d\theta.$$

Тогда

$$u(0, 0) = -\frac{i}{\lambda} \frac{Aa'^2}{\pi R_0'^2} \exp\left[i\left(\frac{R'_0}{a'}\right)^2 u'\right] \left(1 + ik\overline{W} + \frac{1}{2}k^2\overline{W}^2 + \dots\right).$$

Освещенность в параксиальном изображении точки равна

$$E(0, 0) = u(0, 0) u^*(0, 0) =$$

$$= \left(\frac{Aa'^2}{\lambda \pi R_0'^2}\right)^2 \left(1 + ik\overline{W} - \frac{1}{2}k^2\overline{W}^2 + \dots\right) \left(1 - ik\overline{W} - \frac{1}{2}k^2\overline{W}^2 - \dots\right).$$

В том случае, когда волновые aberrации настолько малы, чтобы можно было пренебречь величинами $k\bar{W}$ в степени выше второй, получаем

$$E(0, 0) \approx \left(\frac{Aa'^2}{\lambda\pi R_0'^2} \right)^2 \left[1 + ik\bar{W} - \frac{1}{2}k^2\bar{W}^2 - ik\bar{W} + (k\bar{W})^2 - \frac{1}{2}k^2\bar{W}^2 \right] =$$

$$= \left(\frac{Aa'^2}{\lambda\pi R_0'^2} \right)^2 \left[1 - k^2\bar{W}^2 + (k\bar{W})^2 \right].$$

При $W \equiv 0$: $E_0(0, 0) = \left(\frac{Aa'^2}{\lambda\pi R_0'^2} \right)^2$. При этом нормированная освещенность в параксиальном изображении точки равна

$$i(0, 0) = \frac{E(0, 0)}{E_0(0, 0)} = 1 - k^2 \left[\bar{W}^2 - (\bar{W})^2 \right]. \quad (3.112)$$

Определим среднеквадратическую деформацию волнового фронта

$$W_{\text{скв}} = \left\{ \frac{\int_0^1 \int_0^{2\pi} [W(\rho, \theta) - \bar{W}]^2 \rho d\rho d\theta}{\int_0^1 \int_0^{2\pi} \rho d\rho d\theta} \right\}^{\frac{1}{2}} =$$

$$= \frac{1}{\sqrt{\pi}} \left[\int_0^1 \int_0^{2\pi} W^2(\rho, \theta) \rho d\rho d\theta - 2\bar{W} \int_0^1 \int_0^{2\pi} W(\rho, \theta) \rho d\rho d\theta + (\bar{W})^2 \int_0^1 \int_0^{2\pi} \rho d\rho d\theta \right]^{\frac{1}{2}} =$$

$$= \left[\bar{W}^2 - 2(\bar{W})^2 + (\bar{W})^2 \right]^{\frac{1}{2}} = \left[\bar{W}^2 - (\bar{W})^2 \right]^{\frac{1}{2}}.$$

Полученный результат позволяет придать формуле (3.112) вид известной формулы Марешаля

$$i(0, 0) = \tilde{E}(0, 0) \approx 1 - \left(\frac{2\pi}{\lambda} \right)^2 W_{\text{скв}}^2. \quad (3.113)$$

В случае малой величины aberrаций число Штреля равно нормированной освещенности, определяемой формулой Марешаля.

Согласно правилу четверти волны Рэлея качество изображения можно считать достаточно совершенным, если максимальная деформация волнового фронта $|W|_{\text{max}} \leq \frac{\lambda}{4}$. Следуя Марешалю, принято считать качество

изображения достаточно совершенным, если нормированная освещенность (число Штреля) не менее 0,8. Из формулы (3.113) следует, что в этом

случае $W_{\text{скв}} \leq \frac{\lambda}{14}$.

4. Геометрическая теория aberrаций

4.1. Уравнение луча. Волновая и поперечная aberrации

Обратимся к рис. 4.1, на котором показаны входной и выходной зрачки, а также плоскость предмета x, y и плоскость изображения x', y' .

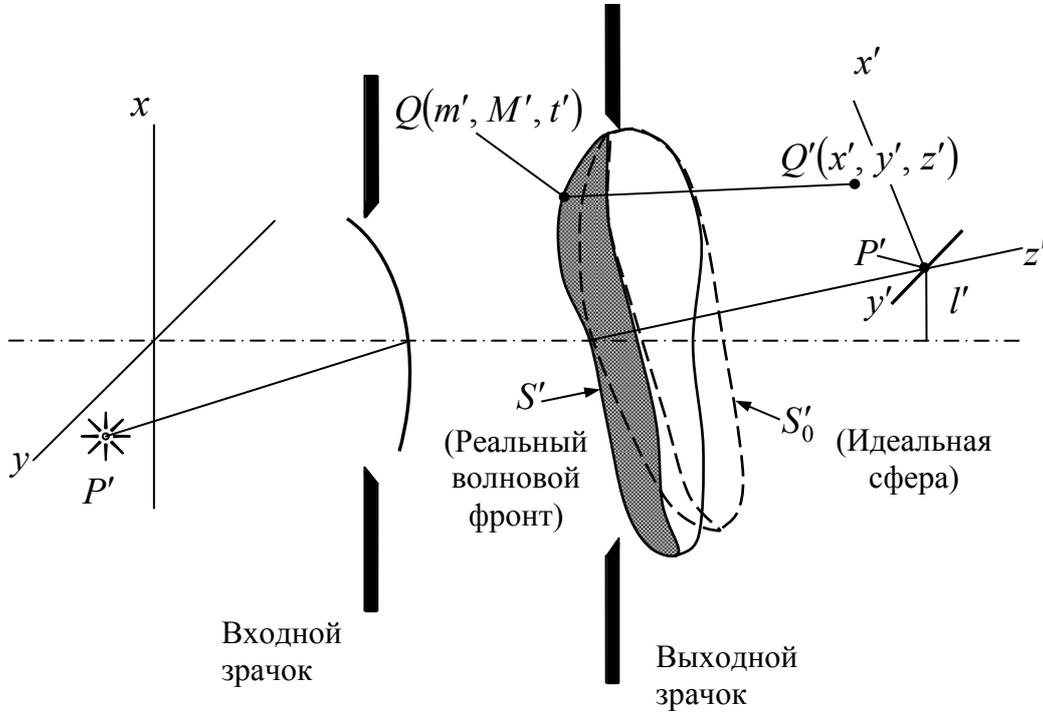


Рис. 4.1. К выводу уравнения идеальной волновой поверхности

Построим сферу с центром в точке P' , которая является гауссовым изображением точки P в плоскости предмета. Сфера радиуса R пересекает оптическую ось в центре выходного зрачка. Если бы оптический путь был одинаков для каждого луча, выходящего из точки P и проходящего сквозь оптическую систему и ее выходной зрачок, то сфера радиуса R совпала бы с поверхностью постоянного оптического пути (равного эйконала) и все лучи собирались бы в точке P' . Однако, в общем случае эта поверхность не является сферой и функцию, характеризующую отклонение формы этой поверхности от сферы, будем называть волновой aberrацией $W(m', M', l')$. Определим уравнение прямой линии, проходящей через точки $Q(m', M', t')$ и $Q'(x', y', z')$ с дополнительным условием, чтобы эта линия была перпендикулярна к S' в точке Q . Уравнение реальной волновой поверхности S' можно записать в виде

$$m'^2 + M'^2 + t'^2 = (R + W)^2.$$

Это уравнение приводится к уравнению идеальной сферы при $W \equiv 0$. Поскольку нас интересуют деформации волнового фронта, не

превышающие единиц длины волны света, величиной $\frac{W^2}{R^2}$ можно

пренебречь. При этом получаем

$$m'^2 + M'^2 + t'^2 - 2RW - R^2 = 0. \quad (4.1)$$

Направляющие косинусы единичного вектора

$$\mathbf{n} = \mathbf{i} \cos \alpha + \mathbf{j} \cos \beta + \mathbf{k} \cos \gamma,$$

нормального к поверхности S' и определяющего направление линии QQ' , найдем из следующих соображений.

Отношение $\frac{df}{|d\mathbf{r}|}$ называется производной скалярной функции f в

точке M по направлению $d\mathbf{r}$, при этом

$$\text{grad} f d\mathbf{r} = df.$$

Отсюда

$$\frac{df}{|d\mathbf{r}|} = \frac{d\mathbf{r}}{|d\mathbf{r}|} \text{grad} f = \mathbf{n} \text{grad} f' = |\text{grad} f'| \cos(\mathbf{n}, \text{grad} f'),$$

т.е. производная равна проекции градиента на направление $d\mathbf{r}$. В рассматриваемом случае, используя (4.1), получаем

$$\frac{dS'}{dm'} = \frac{dS'}{|dm'\mathbf{i}|} = \frac{dm'\mathbf{i}}{|dm'\mathbf{i}|} \text{grad} S' = \mathbf{i} \text{grad} S' = |\text{grad} S'| \cos \alpha.$$

Отсюда

$$\cos \alpha = \frac{1}{|\text{grad} S'|} \frac{dS'}{dm'} = \frac{2}{|\text{grad} S'|} \left(m' - R \frac{\partial W}{\partial m'} \right).$$

С другой стороны, из рисунка следует, что

$$\cos \alpha = \frac{x - m'}{QQ'}.$$

Таким образом

$$\cos \alpha = 2 \frac{m' - R \frac{\partial W}{\partial m'}}{|\text{grad} S'|} = \frac{x' - m'}{QQ'}.$$

Аналогично получаем

$$\cos \beta = \frac{M' - R \frac{\partial W}{\partial M'}}{|\text{grad} S'|} = \frac{y' - M'}{QQ'};$$

$$\cos \gamma = \frac{t'}{|\text{grad} S'|} = \frac{z' - t'}{QQ'}.$$

Полученные равенства позволяют определить уравнение луча, выходящего из точки Q выходного зрачка по нормали к фронту волны и пересекающего плоскость изображения в точке Q' , в виде

$$\frac{x' - m'}{m' - R \frac{\partial W}{\partial m'}} = \frac{y' - M'}{M' - R \frac{\partial W}{\partial M'}} = \frac{z' - t'}{t'}. \quad (4.2)$$

Поскольку в большинстве практических применений $t' \approx -R$, уравнение (4.2) можно переписать в таком виде:

$$\frac{x' - m'}{m' - R \frac{\partial W}{\partial m'}} = \frac{y' - M'}{M' - R \frac{\partial W}{\partial M'}} = -\frac{z'}{R} - 1.$$

Отсюда следует, что

$$x' \approx -\frac{m'z'}{R} + R \frac{\partial W}{\partial m'} \left(1 + \frac{z'}{R}\right), \quad (4.3)$$

$$y' \approx -\frac{M'z'}{R} + R \frac{\partial W}{\partial M'} \left(1 + \frac{z'}{R}\right).$$

В гауссовой плоскости изображения ($z' = 0$) выражения (4.3) принимают вид:

$$x'_0 \approx R \frac{\partial W}{\partial m'}; \quad y'_0 \approx R \frac{\partial W}{\partial M'}. \quad (4.4)$$

Итак, мы рассмотрели общий случай геометрического изображения внеосевой точки, при этом величины x' и y' определяют отклонения точки пересечения реального луча с гауссовой плоскостью изображения (точка Q') от гауссова изображения предметной точки P , т.е. от точки P' .

В общем случае $\frac{z'}{R} \ll 1$. При этом продольный сдвиг изображения (величину z') можно определить из выражений (4.3), положив при этом $x' = 0$ или $y' = 0$:

$$z'_x \approx \frac{R^2}{m'} \frac{\partial W}{\partial m'}; \quad z'_y \approx \frac{R^2}{M'} \frac{\partial W}{\partial M'}.$$

В случае круговой симметрии функции волновой аберрации должно выполняться очевидное равенство

$$z'_x = z'_y = z'_0.$$

В полярной системе координат имеем

$$m' = a' \rho \cos \varphi; \quad (4.5)$$

$$M' = a' \rho \sin \varphi,$$

при этом $\rho = \sqrt{\frac{m'^2 + M'^2}{a'^2}}$; $0 \leq \rho \leq 1$; $\varphi = \arctg \frac{M'}{m'}$, a' – радиус выходного зрачка. В этом случае для выбранной точки предмета $W = W(\rho, \varphi)$. При этом

$$\frac{\partial W}{\partial m'} = \frac{\partial W}{\partial \rho} \frac{\partial \rho}{\partial m'} + \frac{\partial W}{\partial \varphi} \frac{\partial \varphi}{\partial m'};$$

$$\frac{\partial W}{\partial M'} = \frac{\partial W}{\partial \rho} \frac{\partial \rho}{\partial M'} + \frac{\partial W}{\partial \varphi} \frac{\partial \varphi}{\partial M'}.$$
(4.6)

Из равенства $m'^2 + M'^2 = a'^2 \rho^2$ следует

$$a'^2 \rho d\rho = m' dm' + M' dM'.$$
(4.7)

Дифференцируя равенство $\operatorname{tg} \varphi = \frac{M'}{m'}$, получаем

$$\frac{d\varphi}{\cos^2 \varphi} = \frac{m' dM' - M' dm'}{m'^2}.$$
(4.8)

Подставим (4.7) и (4.8) в (4.6):

$$\frac{\partial W}{\partial m'} = \frac{\partial W}{\partial \rho} \frac{m'}{a'^2 \rho} - \frac{\partial W}{\partial \varphi} \frac{M' \cos^2 \varphi}{m'^2},$$

$$\frac{\partial W}{\partial M'} = \frac{\partial W}{\partial \rho} \frac{M'}{a'^2 \rho} + \frac{\partial W}{\partial \varphi} \frac{\cos^2 \varphi}{m'}.$$
(4.9)

Подставив (4.5) в (4.9), получаем

$$\frac{\partial W}{\partial m'} = \frac{\partial W}{\partial \rho} \frac{a' \rho \cos \varphi}{a'^2 \rho} - \frac{\partial W}{\partial \varphi} \frac{a' \rho \sin \varphi \cos^2 \varphi}{a'^2 \rho^2 \cos^2 \varphi} = \frac{\partial W}{\partial \rho} \frac{\cos \varphi}{a'} - \frac{\partial W}{\partial \varphi} \frac{\sin \varphi}{a' \rho},$$

$$\frac{\partial W}{\partial M'} = \frac{\partial W}{\partial \rho} \frac{a' \rho \sin \varphi}{a'^2 \rho} + \frac{\partial W}{\partial \varphi} \frac{\cos^2 \varphi}{a' \rho \cos \varphi} = \frac{\partial W}{\partial \rho} \frac{\sin \varphi}{a'} + \frac{\partial W}{\partial \varphi} \frac{\cos \varphi}{a' \rho}.$$

Итак,

$$\delta g' \approx \frac{R}{a'} \left(\cos \varphi \frac{\partial W}{\partial \rho} - \frac{\sin \varphi}{\rho} \frac{\partial W}{\partial \varphi} \right);$$

$$\delta G' \approx \frac{R}{a'} \left(\sin \varphi \frac{\partial W}{\partial \rho} + \frac{\cos \varphi}{\rho} \frac{\partial W}{\partial \varphi} \right).$$
(4.10)

Формулы (4.4) и (4.10) определяют взаимосвязь волновой и поперечной аберраций.

4.2. Функция волновой аберрации. Первичные аберрации

Для аберрационной функции возможны различные представления. При осесимметричной оптической системе, если использовать для точек выходного зрачка полярные координаты ρ и φ , величина W зависит только от трех инвариантов вращения: l'^2 , ρ^2 и $l' \rho \cos \varphi$. В общем случае можно написать

$$\begin{aligned}
W(\rho, \varphi, l') = & a'_0 + b_0 l'^2 + b_1 \rho^2 + b_2 l' \rho \cos \varphi + c_0 l'^4 + c_1 \rho^4 + \\
& + c_2 l'^2 \rho^2 \cos^2 \varphi + c_3 l'^2 \rho^2 + c_4 l'^3 \rho \cos \varphi + c_5 l' \rho^3 \cos \varphi + \\
& + d_0 l'^6 + d_1 \rho^6 + \dots,
\end{aligned} \tag{4.11}$$

где $m' = a' \rho \cos \varphi$, $\rho = \sqrt{\frac{m'^2 + M'^2}{a'^2}}$; $0 \leq \rho \leq 1$; $M' = a' \rho \sin \varphi$, $\varphi = \arctg \frac{M'}{m'}$,
 a' – радиус выходного зрачка.

Функция волновой аберрации в соответствии с выражениями (4.10) определяет поперечную аберрацию

$$\delta r' = \sqrt{\delta g'^2 + \delta G'^2},$$

равную отклонению точки пересечения реального луча с гауссовой плоскостью изображения в общем случае от внеосевой точки $P'(0, 0)$ в той

же плоскости. Если $W(\rho, \varphi, l') = W(\rho, l')$ при $\frac{\partial W}{\partial \varphi} = 0$, то семейство лучей,

для которых $\rho = const = \rho_1$, будет пересекаться с гауссовой плоскостью изображения в кольце радиуса $\delta r'_0$, определяемом выражением

$$\delta r'_0 = \sqrt{\delta g_0'^2 + \delta G_0'^2} = \frac{R}{a'} \frac{\partial W}{\partial \rho} \Big|_{\rho=\rho_1}. \tag{4.12}$$

Продольный разброс (вдоль оси) для этого частного случая можно определить в полярных координатах, применив формулы (4.3), полагая в них $x' = y' = 0$.

Из формулы (4.12) следует, что для рассматриваемого кольца в выходном зрачке, т.е. при $\rho = \rho_1$, продольный разброс равен

$$z' \approx \frac{R^2}{a'^2 \rho} \frac{\partial W}{\partial \rho} \Big|_{\rho=\rho_1}, \tag{4.13}$$

а общий эффект можно определить путем сопоставления составляющих от каждого кольца при $0 \leq \rho_1 \leq 1$. Проведем подробное исследование каждого члена степенного ряда (4.11).

Постоянные члены ($a_0, c_0 l'^4, d_0 l'^6, \dots$).

Как следует из (4.10), величина постоянных членов ряда (4.11) не влияет на величину аберраций и, по сути дела, определяет изменение величины радиуса идеальной сферы (сферы сравнения) без изменения положения ее центра (см. рис. 4.2а).

Погрешность фокусировки изображения

а) Продольная расфокусировка изображения

Если все аберрации, кроме продольной расфокусировки, отсутствуют, то выражение (4.11) принимает вид

$$W = b_1 \rho^2.$$

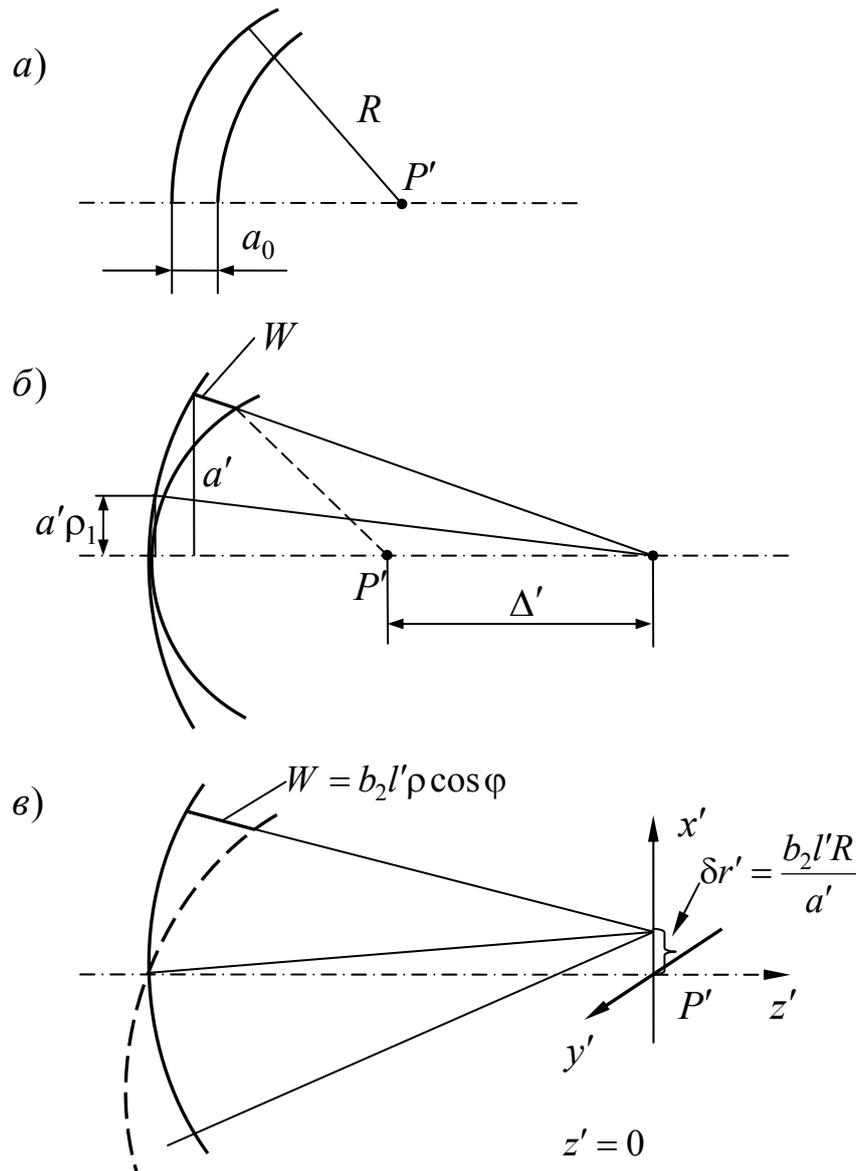


Рис. 4.2. Смещение изображения точки

При этом согласно (4.12) и (4.13) получаем

$$\delta r'_0 = \frac{2b_1 R}{a'} \rho_1, \quad 0 \leq \rho_1 \leq 1, \quad z' \approx \frac{2b_1 R^2}{a'^2} = \Delta', \quad \text{т.е. все лучи собираются в}$$

точке $\left(0, 0, \frac{2b_1 R^2}{a'^2}\right)$. Следовательно, это не абберация, а погрешность

продольной фокусировки изображения, которую можно устранить продольным смещением плоскости наблюдения изображения из точки P' на величину Δ' , как показано на рис. 4.2б. Из полученных выражений

следует соотношение $\frac{\Delta'}{R} = \frac{2b_1 R}{a'^2} = \frac{\delta r'_0}{a'}$, очевидное и из геометрических

соображений. Отсюда следует, что $b_1 = \frac{a'^2 \Delta'}{2R^2} = \frac{a'}{2R} \delta r'_0$.

б) *Поперечная погрешность фокусировки изображения*

В этом случае выражение (4.11) принимает вид

$$W = b_2 l' \rho \cos \varphi = b_2 l' \frac{m'}{a'}. \text{ При этом из выражений (4.4) следует, что}$$

$$\delta g' = b_2 l' \frac{R}{a'}; \quad \delta G' = 0, \text{ т.е. все лучи собираются в точке } \left(\frac{b_2 l' R}{a'}, 0, 0 \right).$$

Следовательно, и эта погрешность фокусировки изображения не является aberrацией, а определяется поперечным сдвигом изображения на величину

$$\delta r' = b_2 l' \frac{R}{a'}, \text{ как показано на рис. 4.2в. Отсюда } b_2 l' = \frac{a'}{R} \delta r'.$$

Первичные aberrации (aberrации Зейделя)

Следующие пять членов степенного ряда (4.11) представляют собой первичные aberrации или, иначе, классические aberrации Зейделя третьего порядка. Третий порядок aberrаций определяется суммарным показателем степени l' и ρ членов ряда, определяющего поперечную aberrацию оптической системы.

а) *Сферическая aberrация*

Эта aberrация определяется коэффициентом C_1 выражения (4.11), т.е.

$$W_c = c_1 \rho^4.$$

При этом согласно (4.12) и (4.13) получаем

$$\delta r'_0 = \frac{4R}{a'} c_1 \rho^3; \quad \Delta s' = \frac{4R^2}{a'^2} c_1 \rho^2,$$

т.е. каждому кольцу при $\rho = \rho_1$ соответствует кольцо в гауссовой плоскости изображения. При $\rho = 1$ эти величины приобретают максимальное значение, равное

$$\delta r'_0 = \frac{4R}{a'} c_1; \quad \Delta s' = \frac{4R^2}{a'^2} c_1. \text{ При этом } c_1 = \frac{a'}{4R} \delta r'_0 = \frac{a'^2}{4R^2} \Delta s'.$$

Напомним, что $b_1 = \frac{a'^2}{2R^2} \Delta'$ при $W_b = b_1 \rho^2$. Отсюда следует возможность уменьшения эффекта влияния сферической aberrации путем введения продольной расфокусировки изображения. В этом случае

$$\Delta W_{\text{ост}} = W_c - W_b = c_1 \rho^4 - b_1 \rho^2 = c_1 \left(\rho^4 - \frac{b_1}{c_1} \rho^2 \right) = c_1 (\rho^4 - 2\eta \rho^2), \text{ где } \eta = \frac{\Delta'}{\Delta s'}.$$

б) *Кома*

Пусть в отличие от других коэффициент $c_5 \neq 0$. При этом

$$W = c_5 l' \rho^3 \cos \varphi.$$

Подставив это выражение в соотношение (4.10), получаем

$$\delta g' = \frac{R}{a'} \left(3c_5 l' \rho^2 \cos^2 \varphi + \frac{\sin^2 \varphi}{\rho} c_5 l' \rho^3 \right);$$

$$\delta G' = \frac{R}{a'} \left(3c_5 l' \rho^2 \sin \varphi \cos \varphi - \frac{1}{\rho} \sin \varphi \cos \varphi c_5 l' \rho^3 \right).$$

Эти выражения легко преобразовать к виду

$$\delta g' = \psi \rho^2 (2 + \cos 2\varphi); \quad (4.14)$$

$$\delta G' = \psi \rho^2 \sin 2\varphi, \quad (4.15)$$

где $\psi = \frac{Rl'}{a'} c_5$.

Выражение (4.14) удобно представить в виде:

$$(\delta g' - 2\psi \rho^2) = \psi \rho^2 \cos 2\varphi. \quad (4.16)$$

Возведя выражения (4.15) и (4.16) в квадрат и сложив, приходим к уравнению вида:

$$(\delta g' - 2\psi \rho^2)^2 + (\delta G')^2 = \psi^2 \rho^4, \quad (4.17)$$

из которого следует, что каждое кольцо выходного зрачка ($\rho = const$) создает кружок в гауссовой плоскости изображения, радиус которого увеличивается при возрастании ρ , при этом центр кружка расположен на оси x' и смещается при возрастании ρ на величину $2\psi \rho^2$.

Кроме того, вследствие удвоения угла 2φ в аргументе одному повороту луча по кольцу в выходном зрачке соответствуют два поворота в плоскости изображения. В результате действия всех зон ($0 \leq \rho \leq 1$) образуется картина, напоминающая комету, как показано на рис. 4.3.

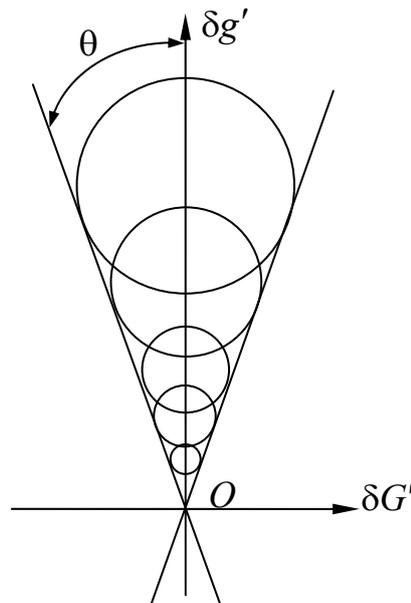


Рис. 4.3. Вид абберационного пятна в изображении точки при первичной коме

в) *Астигматизм и кривизна поверхности изображения*

При наличии только этих aberrаций выражение (4.11) принимает вид:

$$W = c_2 l'^2 \rho^2 \cos^2 \varphi + c_3 l'^2 \rho^2 \quad (4.18)$$

При $\varphi = 0$ (в меридиональной плоскости): $W_t = (c_2 + c_3) l'^2 \rho^2$.

При $\varphi = \frac{\pi}{2}$ (в сагиттальной плоскости): $W_s = c_3 l'^2 \rho^2$.

Таким образом, в меридиональной и сагиттальной плоскостях эта aberrация имеет вид продольной расфокусировки. В общем случае $W_s \neq W_t$.

Равенство (4.18) удобно переписать в виде

$$W = l'^2 \left[(c_2 + c_3) \rho^2 \cos^2 \varphi + c_3 \rho^2 \sin^2 \varphi \right] = \frac{l'^2}{a'^2} \left[(c_2 + c_3) m'^2 + c_3 M'^2 \right].$$

Из соотношений (4.2) следует

$$x' \approx -\frac{m'}{R} z' + R \frac{\partial W}{\partial m'}; \quad y' \approx -\frac{M'}{R} z' + \frac{\partial W}{\partial M'}. \quad \text{При этом}$$

$$x' = \left[-\frac{z'}{R} + \frac{2Rl'^2}{a'^2} (c_2 + c_3) \right] m', \quad y' = \left[-\frac{z'}{R} + \frac{2Rl'^2}{a'^2} c_3 \right] M'.$$

Возведя эти равенства в квадрат и сложив, получаем

$$\frac{x'^2}{\left[-\frac{z'}{R} + \frac{2Rl'^2}{a'^2} (c_2 + c_3) \right]^2} + \frac{y'^2}{\left[-\frac{z'}{R} + \frac{2Rl'^2}{a'^2} c_3 \right]^2} = a'^2 \rho^2,$$

откуда следует, что каждое кольцо выходного зрачка изображается в виде эллипса в плоскости z' (плоскости изображения). Максимальные размеры эллипса получаем при $\rho = 1$. При этом

$\frac{x'^2}{A^2} + \frac{y'^2}{B^2} = 1$, где большая и малая полуоси эллипса определяются

$$\text{выражениями } A = \left[z' - \frac{2R^2 l'^2}{a'^2} (c_2 + c_3) \right] \frac{a'}{R}; \quad B = \left[z' - \frac{2R^2 l'^2}{a'^2} c_3 \right] \frac{a'}{R}.$$

Введем обозначение $\frac{2R^2 l'^2}{a'^2} = \sigma$. При этом $x' = -A \frac{m'}{R}$; $y' = -B \frac{M'}{R}$,

где $A = \left[z' - \sigma (c_2 + c_3) \right] \frac{a'}{R}$; $B = \left[z' - \sigma c_3 \right] \frac{a'}{R}$.

При $A = 0$: $x' = 0$; $y' = -B \frac{M'}{R}$, т.е. лучи фокусируются в меридиональной плоскости, при этом изображение точки имеет вид линии. Поверхность, на которой это происходит, описывается уравнением $z'_t = \sigma (c_2 + c_3)$.

Аналогично получаем, что лучи фокусируются в сагиттальной плоскости ($x' = -A \frac{m'}{R}; y' = 0$ при $B = 0$) на поверхности $z'_s = \sigma c_3$. Расстояние между этими линиями равно $z'_t - z'_s = \sigma c_2$.

Таким образом, коэффициент c_2 определяет астигматизм в изображении точки. Если $c_2 = 0$, то $A = B$, а $z'_t = z'_s$. При этом лучи собираются в точке на расстоянии $z'_t = z'_s = z'_p$ от плоскости изображения Гаусса, причем $z'_p = \frac{2R^2}{a'^2} c_3 l'^2 = \tilde{c}_3 l'^2$.

Это уравнение определяет отклонение поверхности точечного изображения от плоскости, а поэтому и называется кривизной поверхности изображения или кривизной поля изображения.

В общем случае c_2 и c_3 не равны нулю. При этом имеется поверхность z'_c , на которой изображение точки превращается в кружок.

Положив $A^2 = B^2$, находим, что $z' = z'_c$ при $z'_c = \frac{\sigma}{2}(c_2 + 2c_3)$ или при

$z'_c = \frac{z'_t + z'_s}{2}$, т.е. поверхность, на которой геометрическое изображение точки имеет вид так называемого «кружка наименьшего рассеяния», лежит посередине между поверхностями меридионального и сагиттального изображения.

Все возможные состояния изображения описываются соотношениями

$$\frac{x'^2}{(z' - z'_t)^2} + \frac{y'^2}{(z' - z'_s)^2} = \frac{a'^2}{R^2} \rho^2 \quad \text{при} \quad c_3 = \frac{z'_s}{\sigma}; \quad c_2 = \frac{z'_t - z'_s}{\sigma}; \quad \sigma = \frac{2R^2 l'^2}{a'^2}.$$

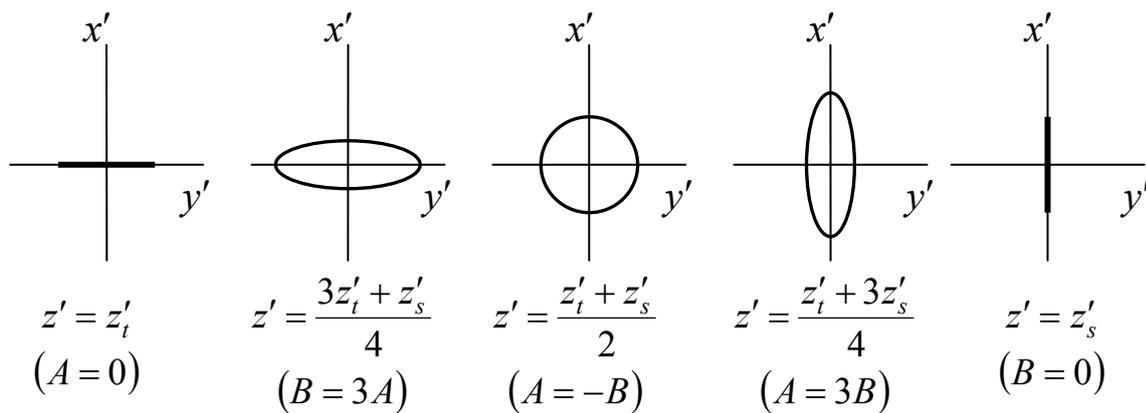


Рис. 4.4 Вид абберационного пятна в изображении точки при первичном астигматизме и первичной кривизне поверхности изображения

г) Дисторсия

В этом случае имеем $W = c_4 l'^3 \rho \cos \varphi = \frac{l'^3}{a'} c_4 m'$.

При этом $\delta g' = \frac{R}{a'} c_4 l'^3$; $\delta G' = 0$.

Таким образом, для точки поля, удаленной от оси на расстояние l' , все лучи фокусируются в точке $\left(\frac{R}{a'} c_4 l'^3, 0, 0\right)$. Но при этом поперечный сдвиг изображения изменяется пропорционально кубу линейного поля и в зависимости от знака c_4 приводит к хорошо известной либо «бочкообразной» ($c_4 < 0$), либо «подушкообразной» ($c_4 > 0$) дисторсии изображения предмета, если последний имеет вид сетки с квадратными ячейками, как показано на рис. 4.5.

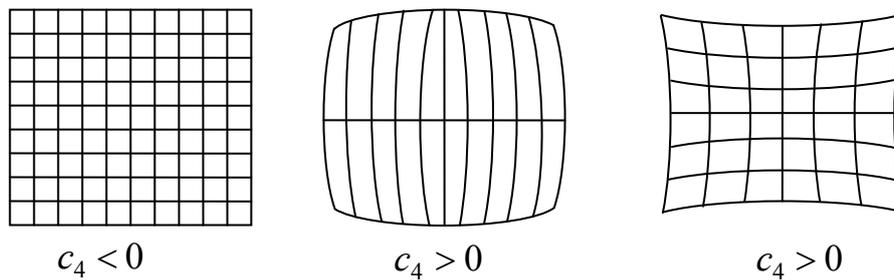


Рис. 4.5. Дисторсия изображения

4.3. Аберрационные свойства сферической поверхности

К числу простейших элементов, образующих оптическую систему, относится, прежде всего, сферическая поверхность, разделяющая среды с различными показателями преломления. Изучение аберрационных свойств сферической поверхности определяет основу грамотного построения исходной конструкции оптической системы для последующей оптимизации ее параметров.

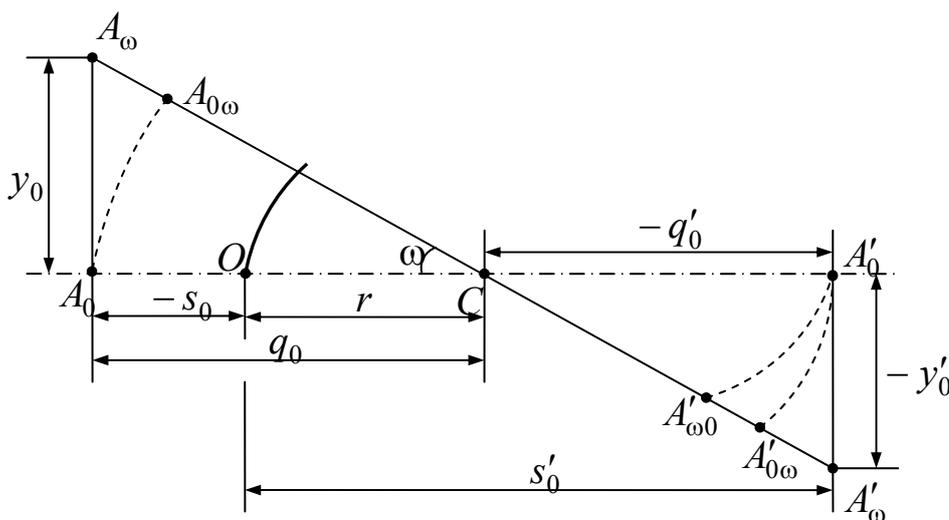


Рис. 4.6. Пецвалева кривизна поверхности изображения, образованного сферической поверхностью

Обратимся к рис. 4.6, на котором показана сферическая поверхность, разделяющая среды с показателями преломления n и n' . Положение идеального изображения A'_0 осевой точки A_0 предмета определяется формулой Аббе:

$$n' \left(\frac{1}{s'_0} - \frac{1}{r} \right) = n \left(\frac{1}{s_0} - \frac{1}{r} \right). \quad (4.19)$$

Вполне очевидно, что при смещении точки A_0 в плоскости рисунка по окружности, концентричной центру кривизны C сферической поверхности, в положение $A_{0\omega}$ ее изображение $A'_{0\omega}$ будет расположено в точке пересечения продолжения линии $A_{0\omega}C$ с окружностью, концентричной точке C и проходящей через точку A'_0 . Линия, проходящая через точки $A'_{0\omega}$ и C , пересекает плоскость, перпендикулярную оптической оси и проходящую через точку A_0 , в точке A_ω на расстоянии, равном $-y_0$ от оси, а плоскость, перпендикулярную оптической оси в точке A'_0 , пересекает в точке A'_ω на расстоянии, равном y'_0 от оси. Из рисунка следует, что

$$\frac{y'_0}{y_0} = \frac{q'_0}{q_0}, \text{ где } q_0 = r - s_0, \text{ а } q'_0 = r - s'_0. \text{ Из формулы Аббе получаем}$$

$$\frac{q'_0}{q_0} = \frac{ns'_0}{n's_0} = V_0, \text{ т.е. точка } A'_\omega \text{ определяет идеальное положение изображения}$$

точки A_ω . Однако, при смещении точки $A_{0\omega}$ пространства предметов в положение точки A_ω параксиальное изображение смещается из точки $A'_{0\omega}$ в точку $A'_{\omega 0}$, положение которой определяется отрезком $q'_{\omega 0}$. Определим проекцию отрезка $A'_\omega A'_{\omega 0}$ на оптическую ось, равную $-\Delta s'_{\omega 0}$, при этом, как следует, из рисунка,

$$\Delta s'_{\omega 0} = q'_0 - q'_{\omega 0} \cos \omega. \quad (4.20)$$

Из формулы Аббе следует, что

$$q_0 = r - s_0 = \frac{n'r q'_0}{nr + (n' - n)q'_0}. \quad (4.21)$$

Тогда

$$q_{\omega 0} = \frac{q_0}{\cos \omega}.$$

В этом случае

$$q'_{\omega 0} = \frac{nr q_{\omega 0}}{n'r - (n' - n)q_{\omega 0}} = \frac{nr q_0}{n'r \cos \omega - (n' - n)q_0}. \quad (4.22)$$

Заменив отрезок q_0 в этом выражении соотношением (4.21), находим, что

$$q'_{\omega 0} = \frac{1}{\cos \omega} \frac{q'_0}{1 - \frac{n' - n}{n} \frac{1 - \cos \omega}{r \cos \omega} q'_0}. \quad (4.23)$$

Подставив соотношение (4.23) в выражение (4.20), в результате преобразований получаем

$$\Delta s'_{\omega 0} = -\frac{n' - n}{2nr} \frac{1 - \operatorname{tg}^2 \frac{\omega}{2}}{1 - \frac{n' - n}{nr} y'_0 \operatorname{tg} \frac{\omega}{2}} y_0'^2. \quad (4.24)$$

При малой величине угла ω можно принять

$$\Delta s'_{\omega 0} \approx -\frac{n' - n}{2nr} y_0'^2. \quad (4.25)$$

Пусть рассматриваемая поверхность – одна из k поверхностей, образующих оптическую систему. Кривизна поверхности изображения $\Delta s'_{\omega 0v}$ после v -ой сферической поверхности соответствует кривизне поверхности изображения $\Delta s'_{0v}$ после всей системы поверхностей, равная

$$\Delta s'_{0v} = \Delta s'_{\omega 0v} \frac{n'}{n'_v} \prod_{v+1}^k V_v^2,$$

где V_v – поперечное увеличение изображения, образованного v -ой поверхностью. Положив в формуле (4.25) $\Delta s'_{\omega 0} = \Delta s'_{\omega 0v}$, получаем

$$\Delta s'_{0v} = -\frac{n'_v - n_v}{2n'_v r_v} y_0'^2 \frac{n'}{n'_v} \prod_{v+1}^k V_v^2 = -\frac{n' n'_v - n_v}{2 n'_v n_v r_v} y_0'^2.$$

При этом кривизна поверхности изображения, образованного оптической системой в целом, равна

$$\begin{aligned} \Delta s'_0 &= \sum_{v=1}^k \Delta s'_{0v} = -\frac{n'}{2} y_0'^2 \sum_{v=1}^k \frac{n'_v - n_v}{n'_v n_v r_v} = -\frac{n'}{2} y_0'^2 \sum_{v=1}^k \frac{1}{r_v} \left(\frac{1}{n_v} - \frac{1}{n'_v} \right) = \\ &= \frac{1}{2} n' y_0'^2 \sum_{v=1}^k \frac{1}{r_v} \Delta \frac{1}{n_v} = -\frac{1}{2} n' y_0'^2 S_{IV}, \end{aligned}$$

где S_{IV} – коэффициент пецвалеовой кривизны (или четвертая сумма Зейделя), равный

$$S_{IV} = -\sum_{v=1}^k \frac{1}{r_v} \Delta \frac{1}{n_v}.$$

Итак, параксиальное изображение точки A_ω расположено в точке $A'_{\omega 0}$. Однако, в общем случае пучок лучей, выходящих из точки A_ω плоскости предметов, не сходится в точке $A'_{\omega 0}$, что обусловлено сферической aberrацией, вносимой в структуру пучка лучей рассматриваемой преломляющей поверхностью.

На рис. 4.7 показан ход луча из осевой точки A в пространстве предметов через сферическую поверхность радиуса r , разделяющую среды с показателями преломления n и n' , в осевую точку A' в пространстве изображений. Смысл остальных обозначений вполне понятен из рисунка. В соответствии с рисунком имеем

$$r \sin \varepsilon = q \sin \sigma, \quad (4.26)$$

$$r \sin \varepsilon' = q' \sin \sigma'. \quad (4.27)$$

Но $n \sin \varepsilon = n' \sin \varepsilon'$. Тогда

$$nq \sin \sigma = n'q' \sin \sigma'. \quad (4.28)$$

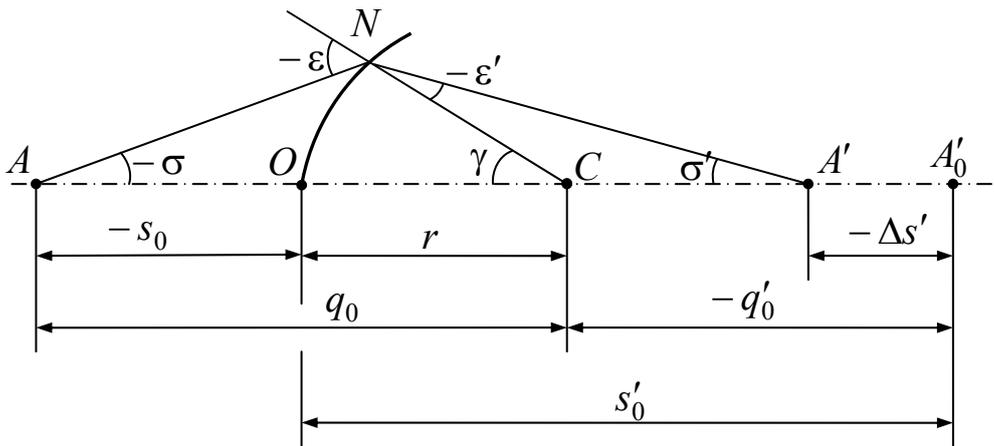


Рис. 4.7. Сферическая aberrация в изображении точки, образованной сферической поверхностью

Отсюда следует, что

$$q' = \frac{n \sin \sigma}{n' \sin \sigma'} q = Vq,$$

где V – поперечное увеличение изображения. При этом величина продольной сферической aberrации в изображении точки A равна

$$\Delta s' = s' - s'_0 = r - s'_0 - r + s' = q'_0 - q' = (V_0 - V)q_0. \quad (4.29)$$

Выразив угол γ через углы σ и ε , а затем через углы σ' и ε' , получаем $\sigma' = \sigma + \varepsilon' - \varepsilon$. При этом

$$\begin{aligned} \sin \sigma' &= \sin \sigma \cos \varepsilon \cos \varepsilon' + \sin \sigma \sin \varepsilon \sin \varepsilon' - \cos \sigma \sin \varepsilon \cos \varepsilon' + \\ &+ \cos \sigma \cos \varepsilon \sin \varepsilon'. \end{aligned}$$

Соотношение (4.26) с учетом закона преломления позволяет преобразовать полученное выражение к виду

$$V = \tilde{n} \psi^{-1}, \quad (4.30)$$

где $\psi = \tilde{n} \tilde{q}_0^2 \sin^2 \sigma + \sqrt{1 - \tilde{q}_0^2 \sin^2 \sigma} \sqrt{1 - \tilde{n}^2 \tilde{q}_0^2 \sin^2 \sigma} +$

$$+ \tilde{q}_0 \sqrt{1 - \sin^2 \sigma} \left(\tilde{n} \sqrt{1 - \tilde{q}_0^2 \sin^2 \sigma} - \sqrt{1 - \tilde{n}^2 \tilde{q}_0^2 \sin^2 \sigma} \right), \quad \tilde{n} = \frac{n}{n'}, \quad \tilde{q}_0 = \frac{q_0}{r}.$$

Напомним, что некоторую функцию $y = f(x)$, непрерывную и имеющую все производные при $x=0$, можно представить степенным рядом Маклорена в виде:

$$f(x) = f(0) + \frac{x}{1!} f'(0) + \frac{x^2}{2!} f''(0) + \dots + \frac{x^n}{n!} f^{(n)}(0) + \dots$$

Разложив выражение (4.30) в ряд Маклорена и подставив его в формулу (8.31), получаем

$$\Delta s' = q_0 (a_3 \sin^2 \sigma + a_5 \sin^4 \sigma + a_7 \sin^6 \sigma + a_9 \sin^8 \sigma + \dots), \quad (4.31)$$

где $a_3 = \frac{1}{2} V_0 \frac{1-\tilde{n}}{\tilde{n}} \tilde{q}_0 (1-\tilde{q}_0) (1+\tilde{n}\tilde{q}_0)$,

$$a_5 = \frac{1}{4} \alpha a_3 - a_3^2,$$

$$a_7 = \frac{1}{8} \beta a_3 - \frac{1}{2} \alpha a_3^2 + a_3^3,$$

$$a_9 = \frac{1}{64} \gamma a_3 - \frac{1}{16} (4\beta + \alpha^2) a_3^2 + \frac{3}{4} \alpha a_3^3 - a_3^4,$$

$$a_{11} = \dots;$$

$$\alpha = 1 + (1-\tilde{n}) \tilde{q}_0 + \frac{1-\tilde{n}^3}{1-\tilde{n}} \tilde{q}_0^2,$$

$$\beta = \alpha - \tilde{n} \tilde{q}_0^2 + (1-\tilde{n}^3) \tilde{q}_0^3 + \frac{1-\tilde{n}^5}{1-\tilde{n}} \tilde{q}_0^4,$$

$$\gamma = 5 \left[\beta + (1-\tilde{n}^5) \tilde{q}_0^5 + \frac{1-\tilde{n}^7}{1-\tilde{n}} \tilde{q}_0^6 \right] - \tilde{n} \tilde{q}_0^2 \alpha + \tilde{n} \left(2\tilde{n} - 3 \frac{1-\tilde{n}^3}{1-\tilde{n}} \right) \tilde{q}_0^4 +$$

$$+ \tilde{n} (1-\tilde{n}) \left(\tilde{n} + \frac{1-\tilde{n}^3}{1-\tilde{n}} \right) \tilde{q}_0^5;$$

$$V_0 = \frac{\tilde{n}}{1 + (\tilde{n}-1) \tilde{q}_0}.$$

Легко убедиться, что при $r = \infty$ или при $s_0 = \infty$ выражение (4.29) приводит к неопределенности вида $0 \cdot \infty$. Однако, и в этих частных случаях можно получить выражения, аналогичные выражению (4.31). Заметим, что коэффициент $a_3 = 0$ при $\tilde{q}_0 = 0$, $\tilde{q}_0 = 1$ и при $\tilde{n}\tilde{q}_0 = -1$. При $\tilde{q}_0 = 0$ имеем $s_0 = r_0$, при $\tilde{q}_0 = 1$ отрезок $s_0 = 0$, а при $\tilde{n}\tilde{q}_0 = -1$ радиус кривизны сферической поверхности определяется соотношением

$$r = \frac{ns_0}{n' + n},$$

откуда следует, что во всех этих случаях осевые точки предмета совпадают с апланатическими точками сферической поверхности. Важно отметить,

что при $a_3 = 0$ обращаются в ноль и все остальные коэффициенты выражения (4.31).

Обратимся к рис. 4.8, на котором показана сферическая поверхность, разделяющая среды с показателями преломления n и n' . Будем считать, что точка предмета A_0 и ее параксиальное изображение A'_0 , образованное сферической преломляющей поверхностью, определяют положение оптической оси. В плоскости рисунка на следе предметной плоскости, перпендикулярной линии $A_0A'_0$, принятой за оптическую ось, на произвольном расстоянии y_0 от точки A_0 выбираем предметную точку A . Из точки A через точку C проводим прямую линию до пересечения в точке A'_0 со следом плоскости изображения, параллельной предметной плоскости и проходящей через точку A'_0 . Назовем луч AA'_0 , проходящий под углом $-\omega$ к оптической оси через центр сферической поверхности C , центральным главным лучом (ЦГЛ).

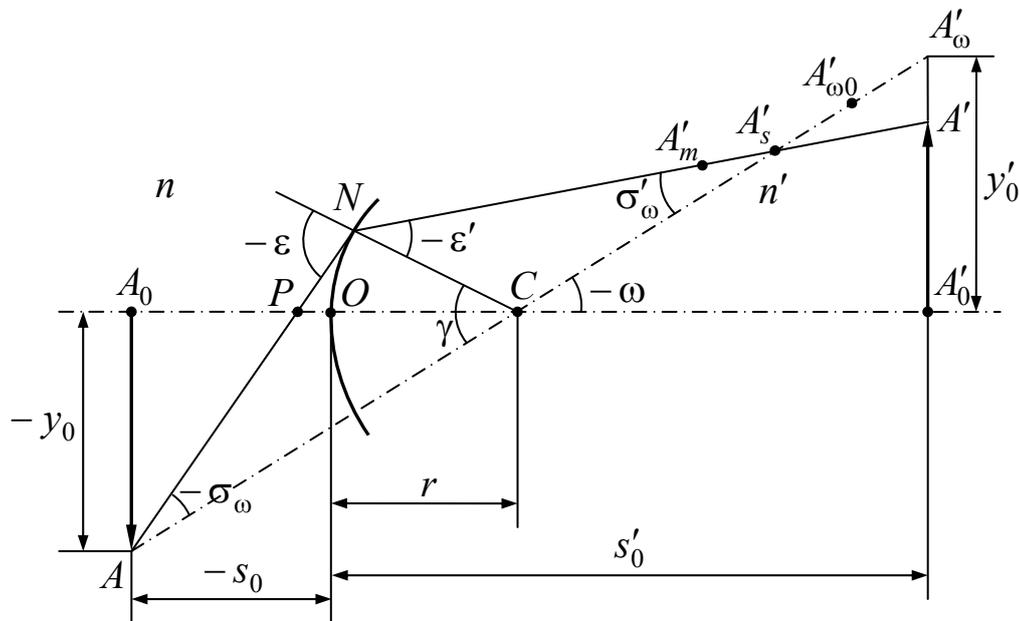


Рис. 4.8. Схема хода лучей через сферическую поверхность

Определим центр входного зрачка преломляющей поверхности точкой P на оси $A_0A'_0$. При этом главный луч из точки A под углом $-\sigma'_\omega$ к ЦГЛ проходит через точку P в точку N сферической преломляющей поверхности, после преломления на которой направляется в точку A' плоскости изображения, пересекает ЦГЛ в точке A'_s , образуя с ним угол σ'_ω в пространстве изображений. Качнув главный луч ANA' вокруг ЦГЛ ACA'_0 на малый угол в противоположные стороны, образуем узкий пучок лучей в сагиттальной плоскости, исходящей из точки A и собирающийся после преломления на сферической поверхности в точке пересечения главного луча с ЦГЛ, т.е. в точке A'_s . При этом сагиттальная составляющая

искривления поверхности изображения в направлении главного луча равно отрезку $A'A'_s = -\Delta s'_s$. На рис. 4.8 отрезок $A'_\omega A'_{\omega 0} = -\Delta s'_{\omega 0}$ определяет составляющую пецвалеовой кривизны в направлении ЦГЛ, а отрезок $A'_{\omega 0} A'_s = -\Delta s'_{\omega \sigma}$ определяет продольную сферическую абберацию преломляющей поверхности для луча ANA' на линии ЦГЛ. Используя теорему синусов, из треугольника $A'_\omega A'_s A'$ находим

$$\frac{\Delta \tilde{s}'_{\omega 0} + \Delta s'_{\omega \sigma'}}{\sin(90^\circ - \omega - \sigma'_\omega)} = \frac{\Delta s'_s}{\sin(90^\circ + \omega)} = \frac{\Delta y'}{\sin \sigma'_\omega}. \quad (4.32)$$

Здесь $\Delta y'$ – линейная величина дисторсии, равная $\Delta y' = A'_0 A' - A'_0 A'_\omega = y' - y'_0$. Из первого равенства (4.32) находим, что

$$\Delta s'_s = \frac{\Delta \tilde{s}'_{\omega 0} + \Delta s'_{\omega \sigma'}}{\cos(\omega + \sigma'_\omega)} \cos \omega, \quad (4.33)$$

а из второго –

$$\Delta y' = \frac{\Delta \tilde{s}'_{\omega 0} + \Delta s'_{\omega \sigma'}}{\cos(\omega + \sigma'_\omega)} \sin \sigma'_\omega, \quad (4.34)$$

где $\Delta \tilde{s}'_{\omega 0} = \frac{\Delta s'_{\omega 0}}{\cos \omega}$. Обозначим отрезок $PC = r - s_p = q_p$. Из треугольника

ACP находим, что $\frac{\sin \sigma_\omega}{q_p} = \frac{\sin(\omega + \sigma_\omega)}{q_{0\omega}}$. В результате последующих

преобразований этого соотношения получаем

$\sin^2 \sigma_\omega = \frac{q_p^2 \sin^2 \omega}{q_0^2 \operatorname{tg}^2 \omega + (q_0 - q_p)^2}$, где $q_0 = q_{0\omega} \cos \omega$. Но $\operatorname{tg} \omega = \frac{y_0}{q_0}$. Тогда

$$\sin^2 \sigma_\omega = \frac{q_p^2}{(q_0 - q_p)^2 + y_0^2} \frac{y_0^2}{q_0^2 + y_0^2}. \quad (4.35)$$

Применим выражение (4.31) для определения величины $\Delta s'_{\omega \sigma'}$. Заменив при этом отрезок q_0 отрезком $q_{0\omega}$, получаем

$$\Delta s'_{\omega \sigma'} = q_{0\omega} (a_{3\omega} \sin^2 \sigma_\omega + a_{5\omega} \sin^4 \sigma_\omega + \dots). \quad (4.36)$$

В результате последовательной подстановки выражения (4.35) в (4.36), а выражений (4.36) и (4.24) в (4.33) и (4.34) находим величину сагиттальной составляющей искривления поверхности изображения в направлении главного луча, а также величину линейной дисторсии соответственно. При $V_\omega = \tilde{n} \psi_\omega^{-1}$ величину угла σ'_ω находим из равенства $\sin \sigma'_\omega = \psi_m \sin \sigma_\omega$.

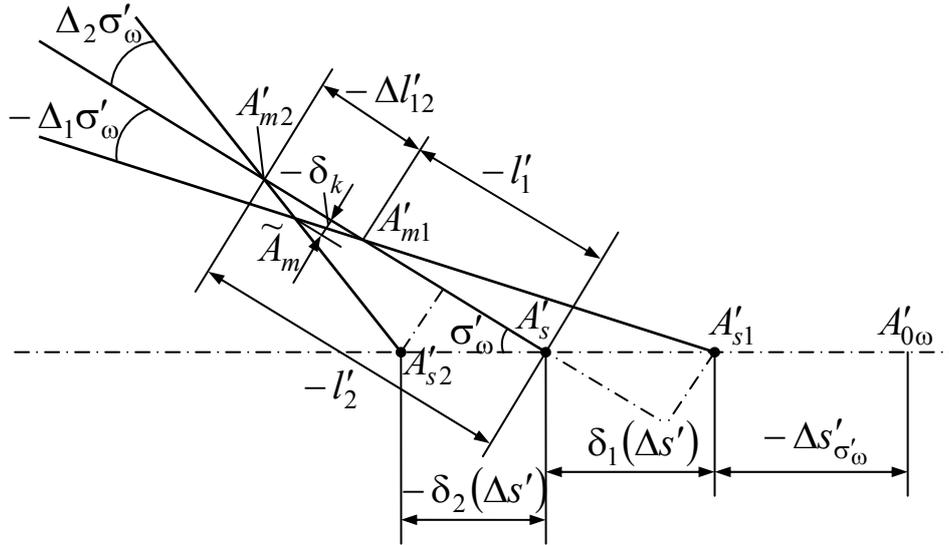


Рис. 4.9. Структура узкого полевого пучка лучей

Проведем из точки A под малыми углами к главному лучу два луча, образующие в пространстве изображений с главным лучом углы $-\Delta_1\sigma'_\omega$ и $\Delta_2\sigma'_\omega$ и пересекающие ЦГЛ в точках A'_{s1} и A'_{s2} , как показано на рис. 4.9. Из рисунка следует, что $\Delta l'_{12} = l'_2 - l'_1$. Полагая углы $\Delta_1\sigma'_\omega$ и $\Delta_2\sigma'_\omega$ малыми, находим, что

$$l'_1 = \frac{\delta_1(\Delta s') \sin \sigma'_\omega}{\Delta_1 \sigma'_\omega} + \delta_1(\Delta s') \cos \sigma'_\omega,$$

$$l'_2 = \frac{\delta_2(\Delta s') \sin \sigma'_\omega}{\Delta_2 \sigma'_\omega} + \delta_2(\Delta s') \cos \sigma'_\omega.$$

При этом

$$\Delta l'_{12} = \left[\frac{\delta_2(\Delta s')}{\Delta_2 \sigma'_\omega} - \frac{\delta_1(\Delta s')}{\Delta_1 \sigma'_\omega} \right] \sin \sigma'_\omega + [\delta_2(\Delta s') - \delta_1(\Delta s')] \cos \sigma'_\omega. \quad (4.37)$$

Вполне очевидно, что величина сферической аберрации определяется величиной угла σ'_ω , а, соответственно, и угла σ_ω . Пусть $\Delta s' = f(\sin \sigma'_\omega)$. Дифференцируя это выражение и заменяя дифференциалы конечными разностями, получаем

$$\delta(\Delta s') = f'(\sin \sigma'_\omega) \cos \sigma'_\omega \Delta \sigma'_\omega.$$

В пределах такого приближения имеем

$$\frac{\delta_2(\Delta s')}{\Delta_2 \sigma'_\omega} = \frac{\delta_1(\Delta s')}{\Delta_1 \sigma'_\omega} = f'(\sin \sigma'_\omega) \cos \sigma'_\omega.$$

При этом

$$\Delta l'_{12} = f'(\sin \sigma'_\omega) \cos^2 \sigma'_\omega (\Delta_2 \sigma'_\omega - \Delta_1 \sigma'_\omega).$$

В области дифференциальных соотношений (при малых отклонениях рассматриваемых лучей от главного) справедливо равенство $\Delta_2 \sigma'_\omega = -\Delta_1 \sigma'_\omega = \Delta \sigma'_\omega$. Тогда

$$\Delta l'_{12} = 2f'(\sin \sigma'_\omega) \cos^2 \sigma'_\omega \Delta \sigma'_\omega.$$

Смещение точки пересечения лучей относительно главного луча определяет абберацию, называемую комой. В соответствии с рис. 4.9 из треугольника $A'_{m1} \tilde{A}'_m A'_{m2}$ находим, что величина комы δ_k равна

$$\delta_k = \frac{1}{2} \Delta l'_{12} \Delta \sigma'_\omega = f'(\sin \sigma'_\omega) \cos^2 \sigma'_\omega (\Delta \sigma'_\omega)^2.$$

В этом выражении при фиксированном положении входного зрачка величина $f'(\sin \sigma'_\omega) \cos^2 \sigma'_\omega$ постоянна. Обозначим $\Delta \sigma'_\omega = \sigma'$. Тогда выражение, определяющее меридиональную кому третьего порядка, приобретает привычную форму

$$\delta_k = b\sigma'^2, \text{ где } b = f'(\sin \sigma'_\omega) \cos^2 \sigma'_\omega.$$

Пусть $\Delta s' = \tilde{a}_{3\omega} \sin^2 \sigma'_\omega$. Тогда $b = 2\tilde{a}_{3\omega} \sin \sigma'_\omega \cos^2 \sigma'_\omega$. При малой величине угла σ'_ω имеем

$$\delta_k = 2\tilde{a}_{3\omega} \sigma'_\omega \sigma'^2. \quad (4.38)$$

Вполне очевидно, что при $\Delta s' \neq 0$ при $\sigma' \rightarrow 0$ точки A'_{m1}, A'_{m2} и \tilde{A}'_m сливаются в одну точку A'_m , расположенную на главном луче, как показано на рис. 4.8. При этом отрезок $A'A'_m = -\Delta s'_m$ определяет меридиональную составляющую искривления поверхности изображения в направлении главного луча. Астигматическая разность меридиональной и сагиттальной составляющих искривления поверхности изображения равна отрезку $A'_s A'_m = -\Delta s'_a$. Как следует из рис. 4.9, величина $\Delta s'_a$ равна

$$\Delta s'_a = \frac{d(\Delta s')}{d\sigma'_\omega} = \sin \sigma'_\omega.$$

При $\Delta s' = \tilde{a}_{3\omega} \sin^2 \sigma'_\omega$ имеем

$$\Delta s'_a = 2\tilde{a}_{3\omega} \sin^2 \sigma'_\omega \cos \sigma'_\omega.$$

Полагая величину угла σ'_ω малой, получаем

$$\Delta s'_a = 2\tilde{a}_{3\omega} \sigma_\omega'^2. \quad (4.39)$$

Рассматриваемый подход к анализу абберационных свойств сферической поверхности дает наглядное представление о взаимосвязи и взаимной обусловленности аббераций различного вида в изображении, образованном сферической поверхностью, позволяет оценить сложность построения исходной конструкции оптической системы, обеспечивающей требуемое качество изображения каждой точки предмета.

Школа геометрической оптики в Санкт-Петербургском государственном университете информационных технологий, механики и оптики была создана на основе классической немецкой оптики трудами выдающегося физика члена-корреспондента АН СССР профессора Игнатовского Владимира Сергеевича и его ученика заслуженного деятеля

науки и техники России, доктора технических наук, профессора Чуриловского Владимира Николаевича, первого декана оптического факультета, организатора и заведующего первой в университете кафедрой теории оптических приборов. Систематическое изучение абберационных свойств отдельных элементов оптической системы (поверхностей, одиночных линз, поверхностей склейки линз, воздушной прослойки и т.д.) в области реальных полей и числовых апертур впервые предпринял выдающийся ученый-изобретатель, заслуженный деятель науки и техники России, доктор технических наук, профессор Русинов Михаил Михайлович. Развивая идеи синтеза как принципа создания рациональной конструкции оптической системы, он ввел понятие базового элемента оптической системы и принцип построения ее путем последовательного развития, основанного на постепенном введении необходимых коррекционных элементов. «Следует заметить, – писал М.М. Русинов в книге «Композиция оптических систем» – что удовлетворение требований, предъявляемых к разрабатываемой оптической системе, во многих случаях может обеспечиваться различными принципиальными схемами, что свидетельствует о существовании нескольких возможных решений. Следовательно, создание той или иной оптической системы нельзя сводить лишь к синтезу ее из ряда выбранных конструктивных элементов, т.е. необходим более широкий подход, который может быть назван композицией оптических систем».

Результаты исследований, выполнявшихся М.М. Русиновым на протяжении многих десятилетий (начиная с 30-х годов прошлого века), по сути дела, определили создание русской оптической школы композиции оптических систем. Эта школа трудами ее создателя М.М. Русинова и его учеников прошла стадию становления и продолжает успешно развиваться.

5. Оптотехника приборов визуального наблюдения

Первое достоверное описание способности линз создавать увеличенное изображение предмета мы находим в трудах монаха францисканского ордена Роджера Бэкона (1214–1294), выпускника Оксфордского университета, одного из замечательных ученых и мыслителей XIII века. В одном из своих трудов Р. Бэкон писал: «Прозрачные тела могут быть так обработаны, что отдаленные предметы покажутся приближенными и наоборот, так что на невероятном расстоянии будем читать малейшие буквы и различать мельчайшие вещи, а также будем в состоянии усматривать звезды, как пожелаем». Из этих строк со всей очевидностью следует, что автор прекрасно понимал свойства линз, выполненных в виде обратного телеобъектива, и понимал, что если фокальный отрезок $S'_{F'} \rightarrow \infty$, то оптическая сила $\varphi \rightarrow 0$, при этом линза превращается в однокомпонентную зрительную трубу, о чем, собственно, и пишет Р. Бэкон в конце фразы. Спустя более трех столетий подобную зрительную трубу, но из двух линзовых компонентов, изготовил выдающийся итальянский ученый Галилео Галилей (1564–1642). Практически одновременно с изобретением зрительной трубы был изобретен и сложный микроскоп. Один из претендентов на изобретение сложного микроскопа – Г. Галилей обнаружил, что его зрительная труба в сильно раздвинутом состоянии позволяет сильно увеличивать мелкие предметы. Факт изобретения микроскопа Галилеем подтверждает его биограф Д. Вивиани: «Изобретение телескопа привело великого мужа к изобретению микроскопа ... В 1612 году он послал один свой микроскоп польскому королю Казимиру». Термин «микроскоп», насколько известно, впервые был применен в 1625 году членом римской Академии Г. Фабером именно к построенному Галилеем прибору.

Известно, что микроскопы, изготовленные Корнелием Дребелем из Алькмара в Голландии, показывали предмет в перевернутом виде. Отсюда следует, что микроскопы Дребеля состояли из двух положительных линз, дающих перевернутое изображение наблюдаемого предмета. В своем труде «Диоптрика» Христиан Гюйгенс писал: «В 1621 году Дребель, голландец, живший в Лондоне, был известен как обладатель таких сложных микроскопов и считался всеми их изобретателем». Сложный микроскоп, состоящий из двояковыпуклого объектива и двояко- или плосковыпуклого окуляра, явился прототипом современных сложных микроскопов. Обладая, однако, большими и неустранимыми хроматическими aberrациями сложный микроскоп такой конструкции не был в состоянии конкурировать с простой лупой.

Около 1660 года итальянец делла Торре изготавливал шарообразные лупы (застывшие капли стекла), диаметр которых доходил до 0,17 мм, дававшие 1500-кратное увеличение изображения. Наиболее знаменитой

личностью в истории микроскопии является Антони ван Левенгук (1632–1723 гг.), живший в голландском городе Дельфте, строивший и применявший для исследований простые микроскопы (по сути дела, короткофокусные лупы) с увеличением образованного изображения до трехсот крат. Левенгук впервые наблюдал и описал мир микроскопических организмов, в том числе и одноклеточных (бактерий). Для наблюдения движения кровяных телец в капиллярных сосудах хвоста молодого угря он даже построил специальный микроскоп. В 1698 году Левенгука посетил русский царь Петр I, находившийся в то время в Голландии. Несомненно, что сам Петр и его соратники закупили и привезли из своих заграничных поездок ряд простых и сложных микроскопов для организованной в Петербурге Кунсткамеры. После создания Академии наук микроскопы перешли в ее ведение. Здесь они были под рукой у молодых русских ученых, возглавляемых М.В. Ломоносовым, которые не только по достоинству оценили значение этого оптического инструмента для развития биологии, медицины и других естественных наук, но и сами приняли участие в его усовершенствовании.

В 1747 году действительный член Петербургской академии наук Леонард Эйлер (1707–1783) предложил идею создания ахроматического объектива микроскопа. Фундаментальная работа Л. Эйлера в области геометрической оптики нашла отражение в его трехтомной «Диоптрике», вышедшей в 1769–1771 годах. О вопросах, рассмотренных в «Диоптрике», дает достаточно конкретное представление опубликованная в 1774 году Николаем Фуссом – учеником Л. Эйлера – работа: «Подробное наставление по приведению телескопов самых разнообразных видов к наивысшей возможной степени совершенства, извлеченное из диоптрической теории Г. Эйлера старшего и доступно изложенное для всех мастеров этого дела. С описанием микроскопа, который можно считать наиболее совершенным в своем роде и который может давать любые желательные увеличения». В 1784 году была опубликована работа действительного члена Петербургской академии наук Франца Ульриха Теодора Эпинуса (1724–1802) «Ахроматический микроскоп новой конструкции, пригодный для рассматривания объективов в свете, отраженном их поверхностью». В этом же году микроскоп Ф. Эпинуса был изготовлен. Воспитатель императора Александра I Ф.Ц. Лагарк видел этот инструмент и в письме к Ж.М. Фавру от 8 августа 1785 года об Эпинусе и его инструменте писал: «Он только что усовершенствовал микроскоп до такой степени, что больше сделать невозможно. Каждый предмет виден с необыкновенной ясностью и сохраняет свою природную окраску. Это еще не единственное преимущество. Микроскоп приделан к подзорной трубе в три фута длиною, им можно управлять с необыкновенной легкостью, и, укорачивая или удлиняя трубу, можно видеть весь предмет или только его часть, достигая произвольного увеличения, если желают с большим или

меньшим вниманием рассмотреть какую-либо часть его. Наконец, так как объектив находится на расстоянии 10 дюймов от предмета, а наблюдатель на три фута позади, то вы понимаете, сударь, что можно наблюдать за всеми движениями насекомых, не пугая их. Это изобретение представляет великую возможность для отечественной истории и наделает много шума, когда с ним более познакомятся». Это, весьма лаконичное, но предельно четкое, описание микроскопа Ф. Эпинуса дает ясное представление о принципиальной схеме, положенной в основу его построения. Отсюда следует, что в укомплектованном шестью сменными объективами ахроматическом микроскопе Ф. Эпинуса предусмотрена возможность плавного изменения увеличения за счет изменения расстояния между предметом и изображением (за счет изменения длины тубуса). Таким образом, благодаря трудам Л. Эйлера, Н. Фусса, Ф. Эпинуса и других именно в России была разработана и осуществлена первая в мире конструкция ахроматического микроскопа переменного увеличения. Идея изменения увеличения микроскопа путем изменения длины его тубуса не получила продолжения в последующих схемах. Тем не менее, сам факт ее осуществления имеет огромное значение для понимания развития прикладной оптики. Один из ахроматических микроскопов Эпинуса сохранился до наших дней и находится в Политехническом музее в Москве (принадлежит Институту истории, естествознания и техники РАН).

В 1807 году голландский оптик ван Дейль опубликовал описание сконструированного им ахроматического микроскопа, который западноевропейские историки обычно признают первым удовлетворительным микроскопом-ахроматом. Однако, он во всех отношениях уступает микроскопу Эпинуса. Еще менее совершенными были ахроматические микроскопы, которые в 1811 году начал выпускать известный оптик Йозеф Фраунгофер (1787–1826 гг.).

Развитию микроскопии много содействовал итальянский оптик, ботаник и астроном Джамбаттиста Амичи (1786–1863 гг.). В 1827 году он разработал и изготовил объектив-ахромат с числовой апертурой 0,60 при хорошей коррекции aberrаций. В 1844 году Амичи начал опыты применения водной и масляной иммерсии, приведшие в 1850 году к созданию объектива с водной иммерсией при апертуре 1,30. Однако, современные объективы с масляной иммерсией с числовой апертурой 1,50 стали возможными только после выдающихся работ немецкого оптика Эрнста Аббе (1840–1905 гг.), установившего закон синусов, позволяющий устранять кому в пределах малой величины предмета (в пределах малого линейного поля), что очень важно именно при больших апертурах. Кроме того, на основании развитой им теории образования изображения в микроскопе он внес ясность в вопрос о разрешающей способности микроскопа. Характеризуя роль творчества Э. Аббе, академик Д.С. Рождественский писал: «Аббе впервые ясно показал, что каждой

остроте инструмента соответствует свой предел возможности. Нельзя грубыми пальцами обрабатывать даже мягкий материал с точностью до сотой доли миллиметра, для этого нужны тонкие инструменты. Тончайший же из всех инструментов – это длина волны. Нельзя видеть объекта меньше полудлины волны – утверждает дифракционная теория Аббе, и нельзя получить изображение меньше полудлины волны, т.е. меньше 1/4 микрона ... Таким образом, гением Аббе установлено сознательное творчество в микроскопии и достигнуты пределы возможного». Теория образования изображения Э. Аббе получила развитие в трудах отечественных ученых Л.И. Мандельштама и академика Д.С. Рождественского.

5.1. Опотехника микроскопа

5.1.1. Лупа, простой микроскоп

Оптическая система простого микроскопа представляет собой короткофокусную лупу (микроскопы Левенгука). Пусть положение предмета относительно передней главной плоскости лупы определяется отрезком a , а положение его изображения, образованного лупой, относительно задней главной плоскости лупы определяется отрезком a' .

Из формулы отрезков $\frac{1}{a'} - \frac{1}{a} = \frac{1}{f'}$ следует, что

$$a' = \frac{a}{1 + a\varphi}, \quad (5.1)$$

где φ – оптическая сила лупы. Пусть L' – расстояние от изображения предмета до зрачка глаза наблюдателя ($L' > 0$), как показано на рис. 5.1. В соответствии с рисунком расстояние

$$L' = p - a', \quad (5.2)$$

где p – расстояние от задней главной плоскости лупы до зрачка глаза.

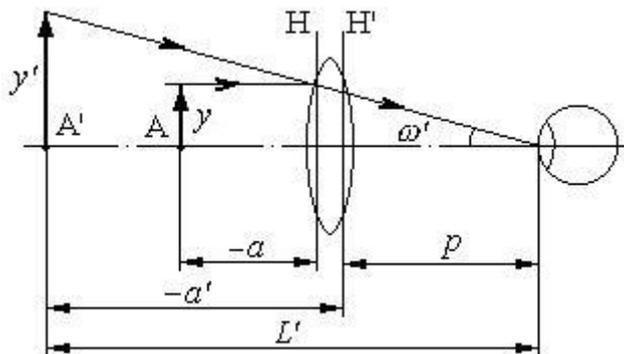


Рис. 5.1. Увеличительное действие лупы

Поперечное увеличение изображения, образованного лупой, равно

$$V = \frac{y'}{y} = \frac{a'}{a} = 1 - a'\varphi = 1 + (L' - p)\varphi. \quad (5.3)$$

Из формулы (5.3) следует, что при совмещении лупы с плоскостью предмета, т.е. при $L' = p$, величина $V=1^{\times}$. При перемещении лупы в направлении от предмета величина образованного ею изображения растёт, достигая максимальной величины при $p=0$, т.е. при совмещении задней главной плоскости лупы со зрачком глаза. При этом $V = 1 + L'\varphi$. В общем случае угловая величина наблюдаемого глазом изображения определяется соотношением:

$$tgw' = \frac{y'}{L'} . \quad (5.4)$$

Угловая величина предмета, наблюдаемого глазом с расстояния L , в соответствии с рис. 5.2 равна

$$tgw = \frac{y}{L} . \quad (5.5)$$

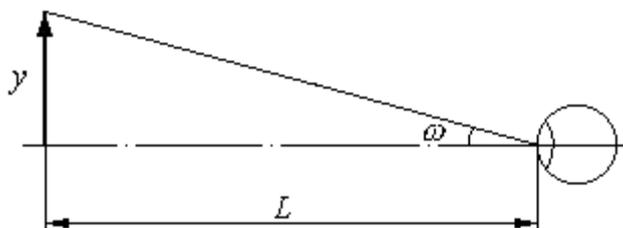


Рис. 5.2. Угловая величина наблюдаемого предмета

При этом видимое увеличение наблюдаемого изображения определяется величиной углового увеличения, равного

$$W = \frac{tgw'}{tgw} = \frac{y' L}{y L'} = V \frac{L}{L'} . \quad (5.6)$$

Используя выражение (5.3), получаем

$$W = [1 + (L' - p)\varphi] \frac{L}{L'} . \quad (5.7)$$

Положив $p=0$, при одинаковой аккомодации глаза в обоих случаях, т.е. при $L=L'=L_0$, имеем

$$W = 1 + L_0\varphi . \quad (5.8)$$

При $L_0=250$ мм: $W = 1 + \frac{250}{f'}$.

Изображение, образованное лупой, должно находиться на таком расстоянии от глаза, чтобы он мог на него аккомодировать без утомления даже при длительном наблюдении. Для нормального глаза это соответствует случаю, когда аккомодационный аппарат находится в состоянии покоя и глаз установлен на бесконечность. При этом предмет должен быть расположен в передней фокальной плоскости лупы. Угловая величина наблюдаемого в этом случае изображения определяется соотношением

$$\operatorname{tg} w' = \frac{y}{f'}. \quad (5.9)$$

Видимое увеличение наблюдаемого изображения и в этом случае определится угловым увеличением, равным

$$W = L_0 \varphi. \quad (5.10)$$

Утрата единицы в выражении (5.10) по сравнению с выражением (5.8) при $W \gg 1$ несущественна и вполне окупается комфортом наблюдения.

5.1.2. Сложный микроскоп

Схема оптической системы сложного микроскопа, состоящая из объектива $\varphi_{об}$ и окуляра $\varphi_{ок}$, с указанием расположения главных плоскостей, фокусов и фокусных расстояний, а также расстояний между ее компонентами представлена на рис. 5.3. С целью упрощения рисунка объектив и окуляр представлены простыми линзами.

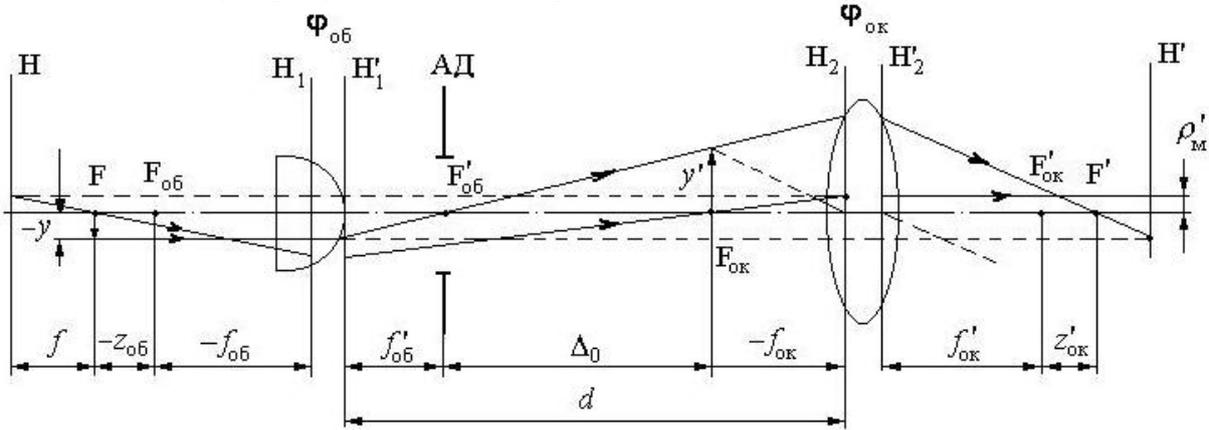


Рис. 5.3. Схема оптической системы сложного микроскопа

Объектив образует изображение y' предмета y (действительное, перевернутое и увеличенное) на расстоянии оптической длины тубуса Δ_0 от своего заднего фокуса $F'_{об}$. При этом полезное увеличение изображения

$$V_{об} = -\frac{\Delta_0}{f'_{об}}.$$

Окуляр, по сути дела, выполняет роль лупы, с помощью которой рассматривается изображение, образованное объективом. Угловое увеличение изображения, образованного объективом, определим выражением (5.10) в виде

$$W_{ок} = L_0 \varphi_{ок}.$$

Видимая угловая величина изображения предмета, образованного оптикой микроскопа, равна

$$2\operatorname{tg} w' = \frac{2y'}{f'_{ок}} = \frac{2y}{f'_{ок}} V_{об} = -\frac{2\Delta_0 y}{f'_{ок} f'_{об}}.$$

Видимая угловая величина предмета определяется формулой (2.5):

$$2tgw = \frac{2y}{L_0}.$$

При этом угловое (видимое) увеличение изображения, образованного оптикой микроскопа, равно

$$W_M = \frac{tgw'}{tgw} = -\frac{L_0\Delta_0}{f'_{об}f'_{ок}}. \quad (5.11)$$

Оптическая сила оптики микроскопа равна

$$\varphi_M = \varphi_{об} + \varphi_{ок} - \varphi_{об}\varphi_{ок}d = \varphi_{об} + \varphi_{ок} - \varphi_{об}\varphi_{ок}(f'_{об} + \Delta_0 + f'_{ок}) = -\varphi_{об}\varphi_{ок}\Delta_0,$$

т.е. угловое увеличение изображения, образованного оптикой микроскопа, определяется формулой:

$$W_M = L_0\varphi_M. \quad (5.12)$$

Действительное изображение, образованное объективом, ограничивается диафрагмой, размещаемой в плоскости предмета окуляра. Она служит полевой диафрагмой. Изображение полевой диафрагмы, образованное объективом в пространстве предметов, называется входным люком (окном) микроскопа, а изображение полевой диафрагмы, образованное окуляром в пространстве изображений, называется выходным люком (окном).

В случае наблюдения самосветящихся (или квазисамосветящихся) объектов апертурная диафрагма размещается в объективе между линзами или позади линз. Входным зрачком микроскопа является изображение апертурной диафрагмы, образованное объективом в пространстве предметов. Положение выходного зрачка объектива определяется положением изображения апертурной диафрагмы, образованного оставшейся частью объектива в пространстве между объективом и окуляром, а выходным зрачком всего микроскопа является изображение выходного зрачка объектива, образованное окуляром.

Большинство объектов, изучаемых под микроскопом, самостоятельно не излучают света. Для того чтобы наблюдать такие объекты, их необходимо освещать, т.е. необходимо, чтобы на них падал или сквозь них проходил свет от какого-либо источника излучения. В зависимости от того, с какой стороны на объект направляются освещающие его пучки лучей, различают два способа освещения:

1. Освещение падающим на объект светом.

В объектив и обратно через него в наблюдательную ветвь проходит свет, отраженный (диффузно или направленно) от объекта.

2. Освещение проходящим через объект светом.

Если разность между углом падения, т.е. углом между главным лучом освещающего пучка лучей и оптической осью системы, и половиной апертурного угла выходящего из осветительного устройства пучка лучей больше, чем половина апертурного угла объектива, то свет, идущий прямо от источника (наблюдение в проходящем свете) или после направленного

отражения от объекта (наблюдение в отраженном свете), не может попасть в объектив. Попасть в объектив могут лишь те лучи, которые отклонились от своего первоначального пути благодаря диффузному рассеянию, преломлению или дифракции в объекте. В этом случае светлыми на темном фоне выглядят лишь те участки объекта, которые как-либо отклоняют лучи света. Такое освещение называется освещением по методу темного поля. При этом объекты носят характер самосветящихся и называются квазисамосветящимися.

Если свет (по меньшей мере, часть его), идущий из осветительного устройства, попадает непосредственно в объектив, то фон выглядит светлым, а поэтому такое освещение называется освещением по методу светлого поля.

Если размер источника света для освещения объекта достаточно велик, то в простейшем случае его располагают на определенном расстоянии a от плоскости предмета.

Если расстояние a определяет положение плоскости, оптически сопряженной с апертурной диафрагмой объектива, как показано на рис. 5.4, то минимальный размер l^* источника света, заполняющего светом числовую апертуру A объектива определяется соотношением:

$$A = n \sin \sigma = \frac{l^*}{\sqrt{a^2 + l^{*2}}} \approx \frac{l^*}{a}.$$

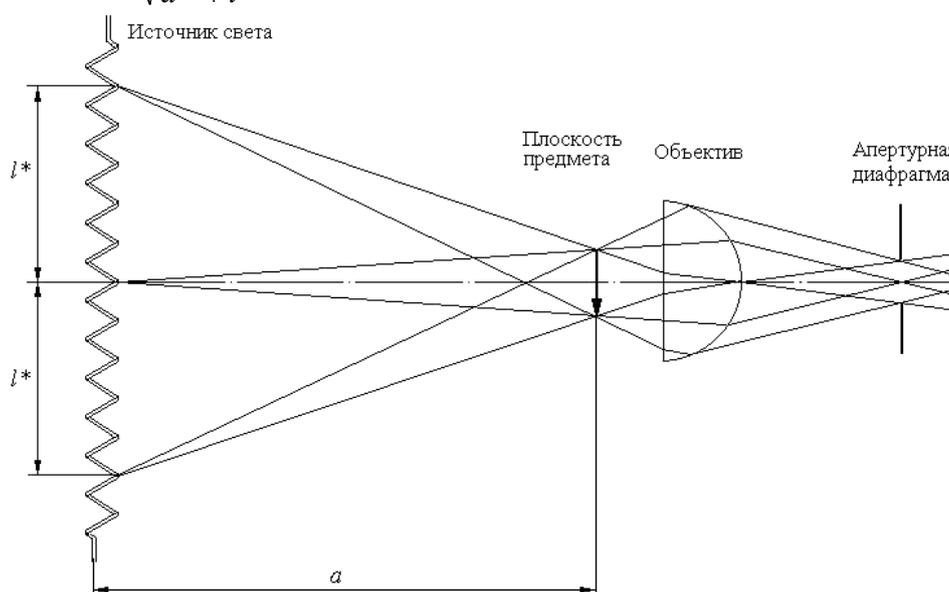


Рис. 5.4. Освещение предмета непосредственно источником света

Вполне очевидно, что источник света можно расположить ближе или дальше этой плоскости при условии его изображения в апертурной диафрагме объектива с допустимой величиной расфокусировки, т.е. в этом случае все лучи, проходящие через каждую точку плоскости предмета в обратном направлении, должны попадать на поверхность источника света.

Из конструктивных соображений между плоскостью предмета и источником света удобно расположить плоское зеркало, как показано на рис. 5.5. Если источник света достаточно мал, то плоское зеркало можно заменить вогнутым, как показано на рис. 5.6, которое формирует увеличенное изображение маленького источника света.

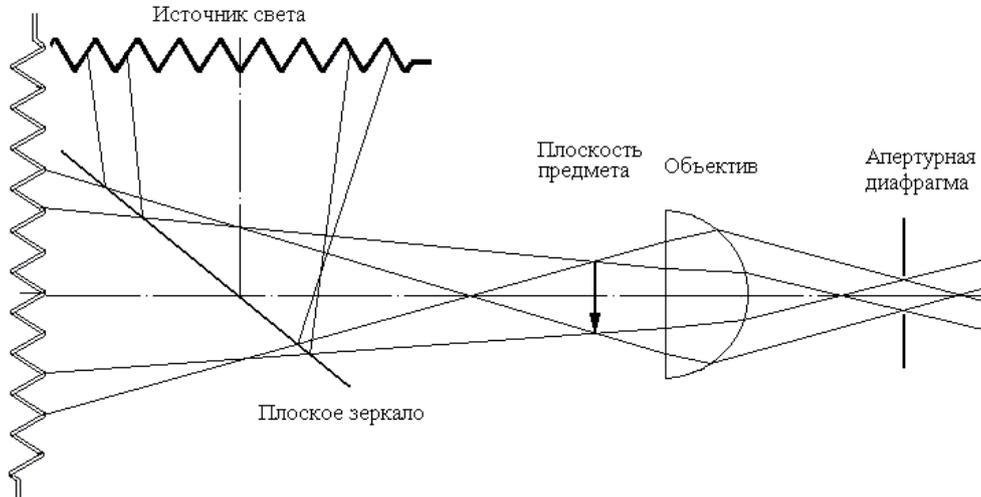


Рис. 5.5. Освещение плоским зеркалом

Однако, при средних и больших числовых апертурах микрообъектива протяженности источника света или увеличения ее с помощью вогнутого зеркала, как правило, недостаточно. Для достижения требуемых размеров изображения источника света применяют специальные оптические системы, называемые конденсорами.

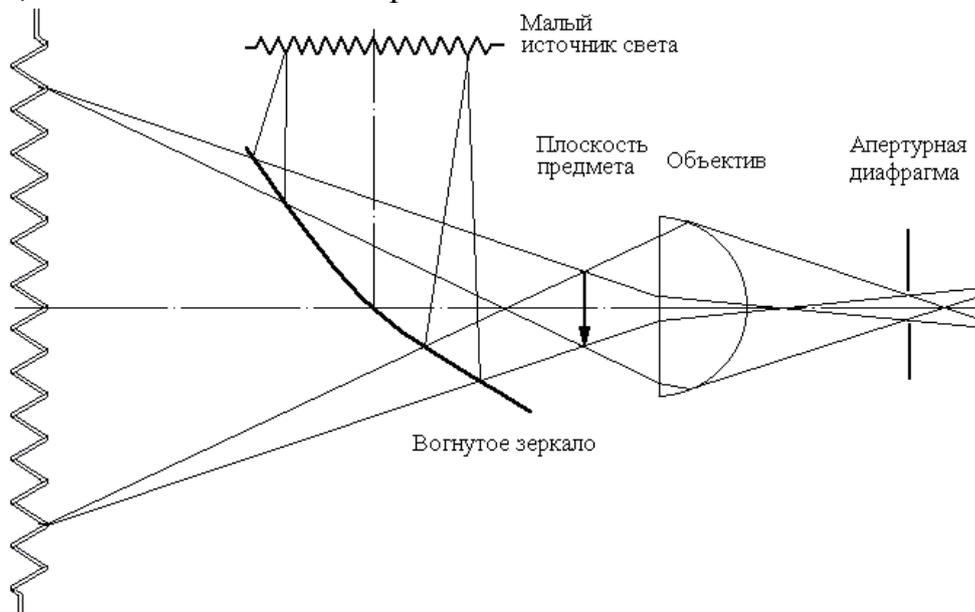


Рис. 5.6. Освещение с вогнутым зеркалом

Способ освещения, когда источник света конденсором изображается непосредственно в плоскости предмета, раньше применялся очень широко. В Англии он был назван критическим способом освещения. Для того

чтобы плоскость предмета была освещена равномерно, применяемый источник света должен излучать энергию равномерно по всей своей поверхности. На практике такие источники света встречаются редко. Поэтому в настоящее время используются обычные источники света в соединении с коллектором.

Итак, входной зрачок конденсора оптически сопряжен с апертурной диафрагмой объектива; в плоскости входного зрачка конденсора располагается ирисовая диафрагма, выполняющая роль апертурной диафрагмы оптической системы микроскопа. Если наблюдать отверстие коллектора из точек образованного им изображения источника света, то оно будет выглядеть одинаково ярким по всей поверхности. Конденсор, в свою очередь, изображает ирисовую диафрагму, расположенную в непосредственной близости к коллектору, в плоскости предмета. Таким образом, ирисовая диафрагма во входном зрачке конденсора служит для регулировки заполнения светом выходного зрачка объектива, а ирисовая диафрагма вблизи коллектора – для регулировки величины освещенного поля. Такой способ освещения впервые был предложен Кёлером и называется освещением по методу Кёлера.

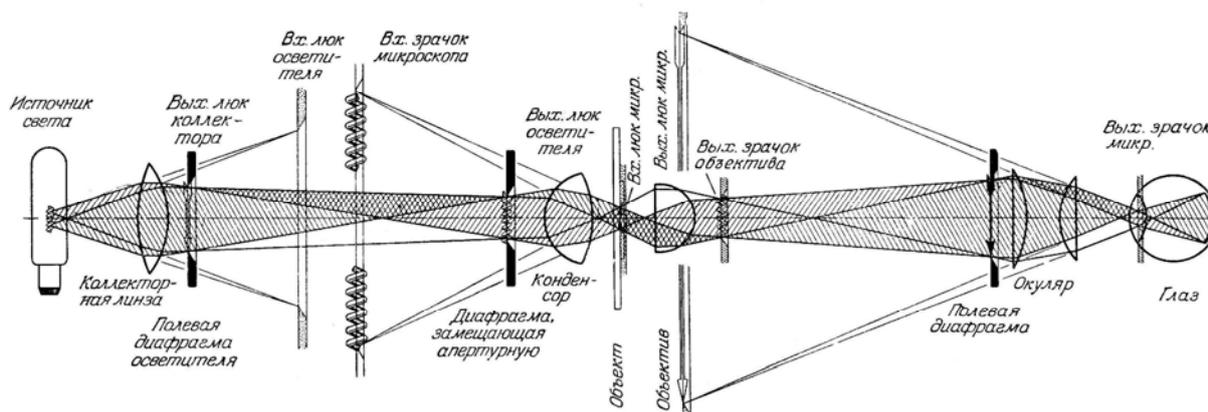


Рис. 5.7. Схема хода лучей в микроскопе с усовершенствованным устройством освещения по способу Кёлера

Оптическая система микроскопа с осветительным устройством, формирующим освещение по методу Кёлера, представлена на рис. 5.7. Коллекторная линза изображает источник света в апертурную ирисовую диафрагму конденсора, который проецирует полевую диафрагму коллектора в плоскость предмета. Выходящие из конденсора параллельные пучки лучей равномерно освещают прозрачный предмет в пределах изображения полевой диафрагмы коллектора и дальше падают на объектив микроскопа. Угол, образуемый с оптической осью лучом, выходящим из осевой точки предмета и направляющимся в край зрачка, является передним апертурным углом микрообъектива. Микрообъектив образует действительное, перевернутое и увеличенное изображение предмета в передней фокальной плоскости окуляра (т.е. в плоскости полевой

диафрагмы микроскопа). В задней фокальной плоскости объектива, где образуется изображение источника света, располагается апертурная диафрагма микроскопа. Выходной зрачок микроскопа, где располагается глаз наблюдателя, является изображением апертурной диафрагмы.

Наличие в осветительной системе ирисовых апертурной и полевой диафрагм позволяет отдельно регулировать величину числовой апертуры и диаметр освещаемого поля.

5.1.3. Разрешающая способность оптической системы микроскопа и полезное увеличение образованного ею изображения

Известно, что радиус дифракционного пятна (радиус кольца первого минимума) в изображении точки определяется формулой:

$$r = \frac{1,22\lambda}{2n \sin \sigma}, \quad (5.13)$$

где λ – длина волны света, $2n \sin \sigma = 2A$ – удвоенная числовая апертура оптической системы микрообъектива. Вполне естественно разрешающую способность оптической системы микроскопа в линейной мере определить соотношением:

$$\delta = \frac{\psi \lambda}{2n \sin \sigma}, \quad (5.14)$$

где ψ – постоянный коэффициент.

Разрешающая способность микроскопа в угловой мере равна

$$\gamma = \frac{\delta}{f'_m} = \frac{\psi \lambda}{d'}, \quad (5.15)$$

где d' – диаметр выходного зрачка. При визуальном наблюдении величина γ не может быть меньше величины γ' , где γ' – угловая разрешающая способность глаза.

Исследования, выполненные А. Арнюльфом в Оптическом институте в Париже, показали, что визуальная разрешающая способность безаберрационной оптической системы микроскопа при различных диаметрах выходного зрачка различна и для тест-объекта Фуко с контрастом, равным единице, для ряда значений диаметра выходного зрачка определяется соотношениями:

$$\begin{aligned} \text{при } d'=0,7 \text{ мм } \delta &= \frac{1,03\lambda}{2n \sin \sigma}, & \text{при } d'=2 \text{ мм } \delta &= \frac{2,5\lambda}{2n \sin \sigma}, \\ \text{при } d'=3 \text{ мм } \delta &= \frac{3,45\lambda}{2n \sin \sigma}, & \text{при } d'=4 \text{ мм } \delta &= \frac{5\lambda}{2n \sin \sigma}. \end{aligned}$$

Заметим, что диаметр зрачка глаза непостоянен и зависит в основном от яркости картины, воспринимаемой глазом, и, прежде всего, от яркости фона. По мере увеличения яркости фона происходит уменьшение диаметра зрачка глаза. В общем случае эта зависимость определяется формулой:

$$d' = 5 - 3th(0,4 \cdot \lg L), \quad (5.16)$$

где d' – диаметр зрачка глаза, мм; L – яркость фона, кд/м²; th – гиперболический тангенс: $th x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$.

Легко убедиться, что в предельных случаях:

- при $L=0$: $d' = 5 - 3 \cdot (-1) = 8 \text{ мм}$,
- при $L = \infty$: $d' = 5 - 3 \cdot 1 = 2 \text{ мм}$.

Как показал Н.И. Пинегин, диаметр зрачка глаза определяется не только яркостью фона, но и его угловой величиной: уменьшение углового размера наблюдаемого поля при постоянной яркости фона приводит к увеличению зрачка глаза. Однако, при постоянной яркости фона увеличение поля свыше 5°–10° практически не оказывает влияния на размер зрачка глаза.

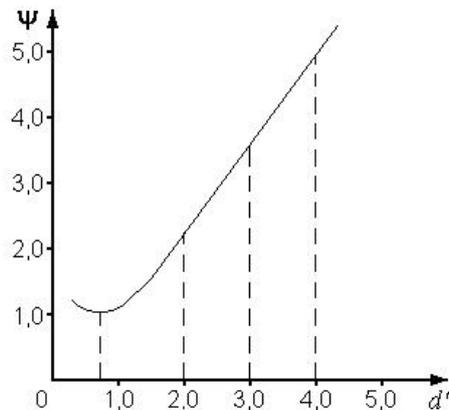


Рис. 5.8. Кривая зависимости $\psi = \psi(d')$

На рис. 5.8 показана кривая, полученная А. Арнюльфом, характеризующая визуальную разрешающую способность оптического прибора при контрасте мира Фучо, равном единице. По оси ординат отложены значения величины ψ , а по оси абсцисс – диаметр выходного зрачка d' . Падение разрешающей способности глаза при $d' < 0,75$ мм определяется, в частности, так называемыми энтоптическими [энто... (от греческого *entós* – внутри) – часть сложных слов, означающая «внутренний»] явлениями, которые сильно ухудшают видение и действительно наблюдаются при малых величинах выходных зрачков (при больших увеличениях изображения предмета) оптических приборов.

Итак, вполне естественное стремление сделать видимое увеличение изображения сколь угодно большим ограничивается визуальной разрешающей способностью, определяемой размером выходного зрачка,

равного $d' = 2 f'_m n \sin \sigma = 2 \frac{L}{W_m} n \sin \sigma$. При этом принято считать, что

диаметр выходного зрачка d' не должен быть меньше 0,5 мм, а, с другой стороны, нет необходимости в его величине, превышающей 1 мм. Отсюда

следует, что при $L=250$ мм предельные значения полезного углового (видимого) увеличения изображения, образованного оптической системой микроскопа, определяются диапазоном

$$500A \leq W_m \leq 1000A, \text{ где } A = n \sin \sigma.$$

5.1.4. Анализ влияния освещения предмета в микроскопе на его изображение

«Существует мнение, – писал академик Д.С. Рождественский, – что изображение в микроскопе мало зависит от метода освещения объекта ... Однако, с одной стороны, известно, что разрешающая сила микроскопа согласно теореме Гельмгольца для самосветящегося объекта вдвое больше, чем при когерентном (не косом) освещении. С другой стороны, обратно, Ван-Циттерт вычисляет, что изображение яркой щели, когда ширина не слишком велика и не слишком мала, а порядка разрешаемой величины, почти вдвое уже при освещении когерентным светом. Точно также Рэлей вычислял, что непрозрачная нить порядка десятых долей так называемой разрешаемой величины вдвое лучше видна при когерентном освещении, чем при некогерентном... Поэтому знать роль освещения в видении микроскопических объектов является существенно важным. Только тот владеет микроскопом до конца, кто знает также, как нужно манипулировать ирисовой диафрагмой под конденсором и чего можно и нужно добиваться». Для этого важно обстоятельно рассмотреть влияние освещения на структуру изображения, образованного оптической системой микроскопа.

Образование изображения в частично когерентном свете

Прежде чем анализировать образование изображения в частично когерентном свете, определим степень когерентности в изображении протяженного некогерентного источника света, полученного с помощью центрированной оптической системы. Свет от каждой точки источника вследствие дифракции (а в общем случае – и аберраций) не собирается в одну точку, а распределяется в пределах пятна рассеяния, при этом пятна от соседних точек перекрываются, а поэтому в интенсивности в достаточно близких друг к другу точках в плоскости изображения вносится как когерентный, так и некогерентный вклад.

Предположим, что S – однородный квазимонохроматический некогерентный источник излучения в виде круга радиуса ρ , излучающий свет со средней (в вакууме) длиной волны $\bar{\lambda}_0$ и расположенный в пространстве предметов в однородной среде с показателем преломления n . Обозначим расстояние между плоскостью предмета и плоскостью входного зрачка через t . Соответствующие величины в пространстве

изображений обозначим теми же символами, но со штрихом, как показано на рис. 5.9.

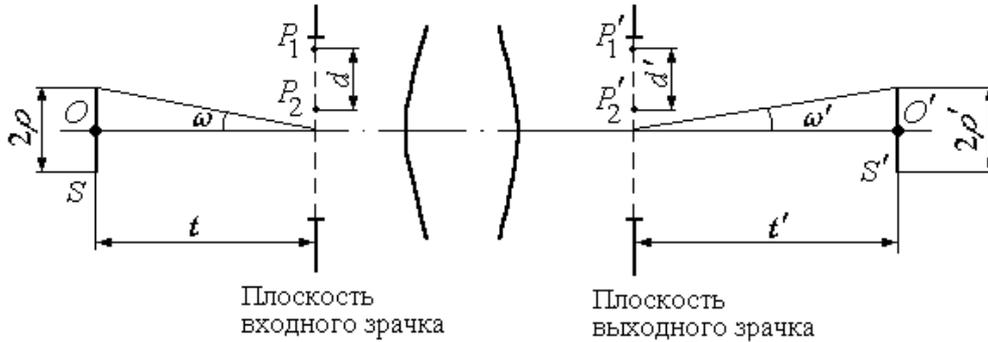


Рис. 5.9. К вычислению степени когерентности в изображении некогерентного источника света

Обозначим через d расстояние между двумя точками P_1 и P_2 входного зрачка. Предположим, что $\frac{\rho}{t} \ll 1$, $\frac{d}{t} \ll 1$ и $OP_1 - OP_2 \ll \bar{\lambda}_0$, где O – центральная (осевая) точка источника. В этом случае комплексная степень когерентности $\mu(P_1, P_2)$ в соответствии с выражением (I.4.64) равна:

$$\mu(P_1, P_2) = \frac{2J_1(v)}{v}, \quad (5.17)$$

$$v = \frac{2\pi}{\lambda} d \sin \omega = \frac{2\pi}{\lambda_0} nd \sin \omega, \quad (5.18)$$

где $2\omega \approx 2\sin \omega \approx \frac{2\rho}{t}$ – угол, под которым источник излучения виден из центра входного зрачка, как показано на рис. 5.9. Заметим, что если условие $OP_1 - OP_2 \ll \bar{\lambda}_0$ не выполняется, то величина $|\mu(P_1, P_2)|$ не изменится, но фаза $\mu(P_1, P_2)$ согласно выражению (I.4.64) увеличится на величину $\psi = 2\pi \frac{OP_1 - OP_2}{\lambda_0}$.

Пусть $U(G, P_1)$ и $U(G, P_2)$ – комплексные возмущения в точках P_1 и P_2 , обусловленные возмущением в точке G ассоциированного монохроматического источника. Тогда возмущения, обусловленные возмущением в точке G , в точках выходного зрачка, сопряженных точкам P_1 и P_2 , определяются формулами:

$$U(G, P_1') = K_{11}U(G, P_1), \quad U(G, P_2') = K_{22}U(G, P_2). \quad (5.19)$$

Здесь $K_{11} = K(P_1, P_1')$ – соответствующая функция пропускания для распространения световых возмущений между сопряженными точками P_1 и P_1' в плоскостях зрачков. При этом в соответствии с формулой Гопкинса (I.4.70) имеем:

$$\mu(P'_1, P'_2) = \frac{1}{\sqrt{I(P'_1)}\sqrt{I(P'_2)}} \int_S U(G, P'_1)U^*(G, P'_2)dS. \quad (5.20)$$

Интенсивности $I(P'_1)$ и $I(P_1)$ в плоскостях входного и выходного зрачков связаны соотношением:

$$I(P'_1) = \int_S |U(G, P'_1)|^2 dS = |K_{11}|^2 \int_S |U(G, P_1)|^2 dS = |K_{11}|^2 I(P_1). \quad (5.21)$$

Аналогичным соотношением связаны $I(P'_2)$ и $I(P_2)$. Используя выражения (5.19), (5.20) и (5.21), получаем:

$$\begin{aligned} \mu(P'_1, P'_2) &= \frac{K_{11} \cdot K_{22}^*}{|K_{11}| \cdot |K_{22}|} \cdot \frac{1}{\sqrt{I(P_1)}\sqrt{I(P_2)}} \int_S U(G, P_1)U^*(G, P_2)dS = \\ &= \exp[i(\Phi_{11} - \Phi_{22})]\mu(P_1, P_2), \end{aligned} \quad (5.22)$$

где Φ_{11} и Φ_{22} – фазы функций K_{11} и K_{22} соответственно. Это соотношение означает, что степень когерентности $|\mu|$ для любых двух точек выходного зрачка равна степени когерентности для сопряженных точек входного зрачка, а фазы соответствующих значений комплексных степеней когерентности для соответствующих пар точек отличаются на величину $\Phi_{11} - \Phi_{22}$, т.е. на геометрическую разность фаз $2\pi \frac{[P_1 P'_1] - [P_2 P'_2]}{\lambda_0}$.

Пусть $v' = \frac{2\pi \cdot n'}{\lambda_0} d' \sin \omega'$. Так как точка P'_1 оптически сопряжена с точкой P_1 , а точка P'_2 – с точкой P_2 , то в рамках параксиальной оптики в соответствии с инвариантом Лагранжа-Гельмгольца имеем: $v' = v$. Следовательно, комплексную степень когерентности для пары точек в выходном зрачке можно в соответствии с выражениями (5.17) и (5.22) записать в виде:

$$\mu(P'_1, P'_2) = \left(\frac{2J_1(v')}{v'} \right) \exp[i(\Phi_{11} - \Phi_{22})]. \quad (5.23)$$

Заметим, что при $v \leq 1$ величина $\left| \frac{2J_1(v')}{v'} \right| \geq 0,88$. При этом с достаточно хорошим приближением можно считать, что некогерентный квазимонохроматический однородный круглый источник будет когерентно освещать в выходном зрачке площадки диаметром

$$d'_{\text{коз}} \approx \frac{0,16\lambda_0}{n' \sin \omega'}, \quad (5.24)$$

где $2 \sin \omega' \approx 2\omega' \approx 2 \frac{\rho'}{t'}$ – угол, под которым диаметр изображения источника виден из центра выходного зрачка, а $\frac{\bar{\lambda}_0}{n'} = \lambda$ – средняя длина волны света в пространстве изображений.

Пусть r'_g обозначает радиус первого темного кольца в дифракционном изображении точки (в пятне Эри), образованном оптической системой, т.е.

$$r'_g = \frac{0,61 \bar{\lambda}_0}{n' \sin \theta'}, \quad (5.25)$$

где $n' \sin \theta' \approx n' \frac{a'}{t'}$ – числовая апертура со стороны изображения (a' – радиус выходного зрачка). При этом из выражений (5.24) и (5.25) находим, что

$$\frac{d'_{\text{коз}}}{r'_g} = \frac{0,16 \sin \theta'}{0,61 \sin \omega'} \text{ и, следовательно,}$$

$$d'_{\text{коз}} \approx 0,26 a' \frac{r'_g}{\rho'}. \quad (5.26)$$

В этой формуле размеры когерентно освещаемых областей выходного зрачка выражены через «физические параметры», а именно, через радиус r'_g первого темного кольца в дифракционной картине Эри, радиус ρ' геометрического изображения источника излучения и радиус a' выходного зрачка.

Выходной зрачок, а, следовательно, и плоскость изображения освещаются почти когерентно, если $d'_{\text{коз}} \geq a'$, т.е. если

$$\rho' \leq 0,26 r'_g. \quad (5.27)$$

Когда $d'_{\text{коз}} \ll a'$, т.е. когда

$$\rho' \gg 0,26 r'_g, \quad (5.28)$$

когерентно освещаемые площадки выходного зрачка будут малы по сравнению с самим выходным зрачком, так что в этом случае выходной зрачок фактически освещается некогерентно. При этом комплексная степень когерентности для пар точек Q'_1, Q'_2 в плоскости изображения будет совпадать со степенью когерентности, обусловленной некогерентным источником, расположенным там же, где и выходной зрачок, и обладающим теми же размерами и формой, причем распределение интенсивности вдоль источника и выходного зрачка одинаково. Следовательно, согласно теореме Ван-Циттерта-Цернике (формула (I.4.57)), имеем:

$$\mu(Q'_1, Q'_2) = \frac{1}{\sqrt{I(Q'_1)} \sqrt{I(Q'_2)}} \int_{\Sigma'} I(P') \frac{\exp[ik(s_1 - s_2)]}{s_1 s_2} dP', \quad (5.29)$$

$$I(Q'_1) = \int_{\Sigma'} \frac{I(P')}{s_1^2} dP', \quad I(Q'_2) = \int_{\Sigma'} \frac{I(P')}{s_2^2} dP'. \quad (5.30)$$

Интегрирование проводится по выходному зрачку Σ' , а через s_1 и s_2 обозначены расстояния от произвольной точки P' в выходном зрачке Σ' до точек Q'_1 и Q'_2 соответственно, как показано на рис. 5.10.

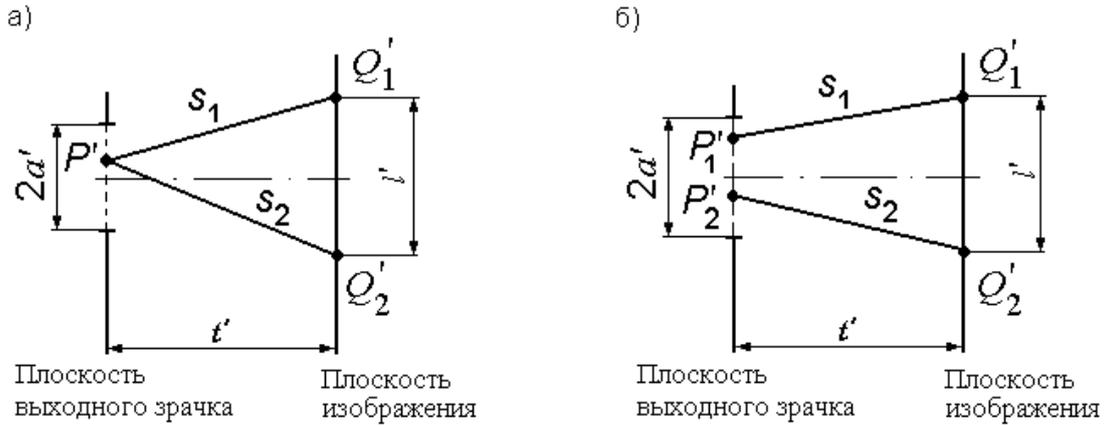


Рис. 5.10. К вычислению комплексной степени когерентности в плоскости изображения: некогерентное освещение выходного зрачка (а); частично когерентное освещение выходного зрачка (б)

Интенсивность $I(P')$ можно рассчитать, зная интенсивность $I(P)$ в сопряженной точке входного зрачка, с помощью соотношения (5.21). Поскольку в это соотношение не входит фаза функции пропускания, величина $\mu(Q'_1, Q'_2)$ не зависит от aberrаций системы. Как правило, интенсивность $I(P')$ практически постоянна. Кроме того, если точки Q'_1 и Q'_2 достаточно близки друг к другу, выражение (5.29) принимает вид:

$$\mu(Q'_1, Q'_2) = \frac{2J_1(u')}{u'}, \quad u' = \frac{2\pi n' a'}{\lambda_0} \frac{a'}{t'} l', \quad (5.31)$$

где l' – расстояние между точками Q'_1 и Q'_2 .

В общем случае, когда не выполняются ни условие (5.27), ни условие (5.28), выходной зрачок освещается частично когерентным светом, который характеризуется комплексной степенью когерентности (5.23). Тогда значение комплексной степени когерентности для пары точек в плоскости изображения следует рассчитывать с помощью закона распространения (I.4.74), что приводит к выражению:

$$\mu(Q'_1, Q'_2) = \frac{1}{\sqrt{I(Q'_1)}\sqrt{I(Q'_2)}} \iint_{\Sigma' \Sigma'} \sqrt{I(P'_1)}\sqrt{I(P'_2)} \left(\frac{2J_1(v')}{v'} \right) \times \exp\left\{i[\Phi_{11} - \Phi_{22} + \bar{k}(s_1 - s_2)]\right\} x_1 x_2^* dP'_1 dP'_2. \quad (5.32)$$

Пользуясь этим выражением, можно также рассчитать интенсивности $I(Q'_1)$ и $I(Q'_2)$, если учесть, что $\mu(Q'_1, Q'_1) = \mu(Q'_2, Q'_2) = 1$. Отметим, что в

этом случае комплексная степень когерентности зависит от aberrаций оптической системы, поскольку подынтегральное выражение содержит фазы Φ_{11} и Φ_{22} функции пропускания.

Влияние освещения предмета на разрешающую способность оптической системы микроскопа

Для наблюдения с помощью микроскопа небольшого несветящегося предмета его необходимо осветить. Если предмет почти прозрачен, как это обычно и бывает, его освещают снизу, т.е. предмет освещается проходящим светом, который проходит сквозь предмет и фокусируется в плоскости изображения, образованного объективом микроскопа. Для получения достаточной освещенности в изображении предмета требуемой площади применяют систему линз, образующих осветительную систему микроскопа. Существуют различные методы освещения предмета. Принципиально важными и широко распространенными методами являются следующие два: критическое освещение и освещение по методу Кёлера. Рассмотрим разрешающую способность оптической системы микроскопа, которую можно достичь при их применении.

Критическое освещение

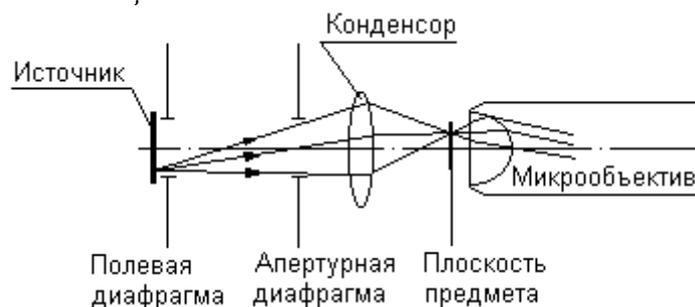


Рис. 5.11. Схема критического освещения

При этом методе освещения равномерно яркий источник света располагается непосредственно перед полевой диафрагмой и с помощью конденсора изображается в плоскости предмета объектива микроскопа, как показано на рис. 5.11, при этом величина изображения полевой диафрагмы, образованного конденсором, должна быть равна величине наблюдаемого предмета. Размер освещенной площади предмета значительно больше эффективного пятна Эри в изображении точки источника света, образованном конденсором в плоскости предмета, т.е. в этом случае $\rho' \gg r'_g$. Как было показано, при таких условиях комплексная степень когерентности для любой пары точек в плоскости предмета объектива совпадает со степенью когерентности, обусловленной некогерентным источником излучения, заполняющим светом конденсор. Кроме того, степень когерентности не зависит от aberrаций конденсора. Поскольку разрешающая способность оптики микроскопа зависит только

от степени когерентности света, падающего на предмет, и от свойств и параметров объектива, aberrации конденсора совершенно не влияют на разрешающую способность микроскопа.

Этим определяется возможность применения одного и того же конденсора при изменении величины линейного поля и задней числовой апертуры в достаточно широких пределах.

Для оценки влияния осветительной системы микроскопа на его разрешающую способность рассмотрим два небольших отверстия $P_1(X_1, Y_1)$ и $P_2(X_2, Y_2)$ в плоскости предмета. Комплексная степень когерентности света, освещающего эти отверстия, определяется формулой, подобной формуле (3.91), а именно:

$$\mu(P_1, P_2) = \frac{2J_1(u_{12})}{u_{12}}, \quad u_{12} = \frac{2\pi}{\lambda_0} \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2} n'_s \sin \sigma'_s, \quad (5.33)$$

где $n'_s \sin \sigma'_s$ – задняя (со стороны объектива микроскопа) числовая апертура конденсора.

Пусть $P(X, Y)$ – любая другая точка в плоскости предмета, а P' – ее изображение, образованное объективом. Будем считать, что в изображении предмета, образованном объективом микроскопа, aberrации отсутствуют. Тогда распределение интенсивности в плоскости изображения предмета, обусловленное светом, приходящим только из точки P_1 , представляет собой пятно Эри, центром которого является изображение P'_1 точки P_1 . Следовательно, интенсивность $I^{(1)}(P')$, обусловленная светом, который приходит в точку P' только из точки P_1 , с точностью до постоянного множителя равна:

$$I^{(1)}(P') = \left(\frac{2J_1(v'_1)}{v'_1} \right)^2, \quad v'_1 = \frac{2\pi}{\lambda_0} \sqrt{(X' - X_1)^2 + (Y' - Y_1)^2} n'_0 \sin \sigma'_0, \quad (3.34)$$

где $n'_0 \sin \sigma'_0$ – задняя числовая апертура микрообъектива (в пространстве изображений).

Однако для безaberrационной системы в соответствии с инвариантом Лагранжа-Гельмгольца имеем:

$$v'_1 = v_1 = \frac{2\pi}{\lambda_0} \sqrt{(X - X_1)^2 + (Y - Y_1)^2} n_0 \sin \sigma_0,$$

где $n_0 \sin \sigma_0$ – передняя числовая апертура микрообъектива (в пространстве предметов).

$$\text{При этом } I^{(1)}(P') = \left(\frac{2J_1(v_1)}{v_1} \right)^2. \quad (5.35)$$

Относительное распределение интенсивности света в плоскости изображения предмета, обусловленное светом, приходящим только из точки P_2 , определяется аналогичным выражением вида:

$$I^{(2)}(P') = \left(\frac{2J_1(v_2)}{v_2} \right)^2, \quad v_2 = \frac{2\pi}{\lambda_0} \sqrt{(X - X_2)^2 + (Y - Y_2)^2} n_0 \sin \sigma_0. \quad (5.36)$$

Таким образом, если два отверстия освещаются с помощью конденсора, то интенсивность $I(P')$ в плоскости изображения, образованного объективом микроскопа, определяется суперпозицией двух частично когерентных пучков световых лучей. Интенсивность каждого из пучков определяется выражениями (5.35) и (5.36), а комплексная степень когерентности – соотношением (5.33). Подставив указанные соотношения в формулу (I.4.47), получим выражение, определяющее $I(P')$. Если предположить, что точка P' достаточно близка к геометрическим изображениям точек P_1 и P_2 (точнее, что $\delta = [P_1P'] - [P_2P'] \ll \bar{\lambda}$), то формула, определяющая $I(P')$, примет вид:

$$I(P') = \left(\frac{2J_1(v_1)}{v_1} \right)^2 + \left(\frac{2J_1(v_2)}{v_2} \right)^2 + 2 \left(\frac{2J_1(mv_{12})}{mv_{12}} \right) \left(\frac{2J_1(v_1)}{v_1} \right) \left(\frac{2J_1(v_2)}{v_2} \right), \quad (5.37)$$

$$\text{где } m = \frac{n'_s \sin \sigma'_s}{n_0 \sin \sigma_0}, \quad v_{12} = \frac{u_{12}}{m} = \frac{2\pi}{\lambda_0} \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2} n_0 \sin \sigma_0. \quad (5.38)$$

Из соотношения (5.37) следует ряд важных выводов. Если mv_{12} является отличным от нуля корнем уравнения $J_1(mv_{12})=0$, то соотношение (5.37) принимает вид:

$$I(P') = \left(\frac{2J_1(v_1)}{v_1} \right)^2 + \left(\frac{2J_1(v_2)}{v_2} \right)^2. \quad (5.39)$$

Таким образом, когда задняя (выходная) числовая апертура конденсора равна передней числовой апертуре объектива микроскопа (т.е., когда $m=1$), а расстояние между геометрическими изображениями отверстий равно радиусу одного из темных колец в безабберационном изображении точки (в пятне Эри), образованном объективом, распределение интенсивности в плоскости изображения такое же, как и при некогерентном освещении отверстий P_1 и P_2 .

Если числовая апертура конденсора очень мала ($m \rightarrow 0$), то $\frac{2J_1(mv_{12})}{mv_{12}} \approx 1$ и соотношение (5.37) принимает вид:

$$I(P') = \left[\frac{2J_1(v_1)}{v_1} + \frac{2J_1(v_2)}{v_2} \right]^2. \quad (5.40)$$

При этом для любого расстояния между отверстиями распределение интенсивности остается таким же, как и в случае полностью когерентного освещения.

Формула (5.37) позволяет исследовать зависимость распределения интенсивности в плоскости изображения, образованного объективом микроскопа, от отношения m числовых апертур. В частности, определим интенсивность в точке, расположенной посередине между точками P'_1 и P'_2 .

Пусть расстояние от точки P' до точки P'_1 равно q . Тогда расстояние от точки P'_2 до точки P' равно разности расстояний $P'_1P'_2 - q$. При этом по смыслу решаемой задачи имеем:

$$\begin{aligned}(X - X_1)^2 + (Y - Y_1)^2 &= q^2, \\(X_2 - X_1)^2 + (Y_2 - Y_1)^2 &= (P'_1P'_2)^2, \\(X_2 - X)^2 + (Y_2 - Y)^2 &= (P'_1P'_2 - q)^2.\end{aligned}$$

Введем обозначения:

$$q \frac{\bar{\lambda}_0}{2n_0 \sin \sigma_0} = \delta, \quad P'_1P'_2 = \frac{\psi \bar{\lambda}_0}{2n_0 \sin \sigma_0}.$$

Тогда $v_1 = \pi\delta$, $v_2 = \pi(\psi - \delta)$, $v_{12} = \pi\psi$. Введенные обозначения позволяют формулу (5.37) переписать в виде:

$$\begin{aligned}I(P') &= \left(\frac{2J_1(\pi\delta)}{\pi\delta} \right)^2 + \left(\frac{2J_1[\pi(\psi - \delta)]}{\pi(\psi - \delta)} \right)^2 + \\ &+ 2 \frac{2J_1(\pi m\psi)}{\pi m\psi} \cdot \frac{2J_1(\pi\delta)}{\pi\delta} \cdot \frac{2J_1[\pi(\psi - \delta)]}{\pi(\psi - \delta)}.\end{aligned} \quad (5.41)$$

При $\delta=0$:

$$I(P'_1) = 1 + \left(\frac{2J_1(\pi\psi)}{\pi\psi} \right)^2 + 2 \frac{2J_1(\pi m\psi)}{\pi m\psi} \cdot \frac{2J_1(\pi\psi)}{\pi\psi}. \quad (5.42)$$

При $\delta=\psi$:

$$I(P'_2) = 1 + \left(\frac{2J_1(\pi\psi)}{\pi\psi} \right)^2 + 2 \frac{2J_1(\pi m\psi)}{\pi m\psi} \cdot \frac{2J_1(\pi\psi)}{\pi\psi}. \quad (5.43)$$

Отсюда следует, что $I(P'_1) = I(P'_2)$.

Будем считать, что точки P_1 и P_2 изображаются отдельно, если интенсивность в точке, расположенной посередине между точками P'_1 и P'_2 , меньше интенсивности в изображении каждой из этих точек на величину, принятую в качестве предельной. Положив $\delta = \frac{1}{2}\psi$, получаем:

$$I(P') = 2 \left[\frac{2J_1\left(\frac{1}{2}\pi\psi\right)}{\frac{1}{2}\pi\psi} \right]^2 \left[1 + \frac{2J_1(\pi m\psi)}{\pi m\psi} \right]. \quad (5.44)$$

Относительную величину разности интенсивностей в точках P' и P'_1 или P'_2 определим соотношением вида:

$$i(P') = \frac{I(P'_j) - I(P')}{I(P'_j)}, \quad (5.45)$$

где $j=1, 2$. Подставив в это соотношение выражения (5.42) или (5.43) и (5.44), получаем:

$$i(P') = 1 - \frac{2 \left[\frac{2J_1\left(\frac{1}{2}\pi\psi\right)}{\frac{1}{2}\pi\psi} \right]^2 \left[1 + \frac{2J_1(\pi m\psi)}{\pi m\psi} \right]}{1 + 2 \frac{2J_1(\pi m\psi)}{\pi m\psi} \cdot \frac{2J_1(\pi\psi)}{\pi\psi} + \left[\frac{2J_1(\pi\psi)}{\pi\psi} \right]^2}. \quad (5.46)$$

Отсюда следует, что при $J_1\left(\frac{1}{2}\pi\psi\right) = 0$ $i(P') = 1$. Заметим, что $J_1\left(\frac{1}{2}\pi\psi\right) = 0$ при $\frac{1}{2}\pi\psi = 3,832$, т.е. при $\psi = 2,440$.

При $m=0$ выражение (5.46) можно представить в виде:

$$i(P') = 1 - \frac{4 \left[\frac{2J_1\left(\frac{1}{2}\pi\psi\right)}{\frac{1}{2}\pi\psi} \right]^2}{\left[1 + \frac{2J_1(\pi\psi)}{\pi\psi} \right]^2}.$$

В этом случае $i(P') = 0$ при $\pi\psi = 4,77$, т.е. при $\psi = 1,5183$.

Положив в выражении (5.46) величину $m=1$, получаем:

$$i(P') = 1 - \frac{2 \left[\frac{2J_1\left(\frac{1}{2}\pi\psi\right)}{\frac{1}{2}\pi\psi} \right]^2 \left[1 + \frac{2J_1(\pi\psi)}{\pi\psi} \right]}{1 + 3 \left[\frac{2J_1(\pi\psi)}{\pi\psi} \right]^2}.$$

При $\pi\psi = 3,378$ величина $i(P') = 0$.

При $\pi\psi = 3,832$ ($\psi=1,22$) $J_1(\pi\psi)=0$. При этом

$$i(P') = 1 - 2 \left[\frac{2J_1\left(\frac{1}{2}\pi\psi\right)}{\frac{1}{2}\pi\psi} \right]^2 = 0,268.$$

Итак, при $\psi=2,440$ независимо от величины m контраст изображения, определяемый отношением $\tau = \frac{I(P'_j) - I(P')}{I(P'_j) + I(P')}$, равен 1. При $m=0$ (при когерентном освещении предмета) и при $\psi=1,5183$ контраст изображения $\tau=0$. Положив $m=1$ (полагая освещение предмета некогерентным), находим, что контраст изображения $\tau=0$ при $\psi=1,075$.

Изменяя величину m в диапазоне от 0 до 2 с интервалом, равным 0,25, и используя формулу (5.46), вычисляем зависимость $i(P') = i(P')(\psi)$.

На рис. 5.12 представлены кривые зависимости контраста $i(P')$ изображения двух одинаковых точек P'_j ($j=1, 2$) от величины коэффициента ψ для трех значений $m = 0; 1; 1,5$. Вычисления выполнены с помощью программы MathCad

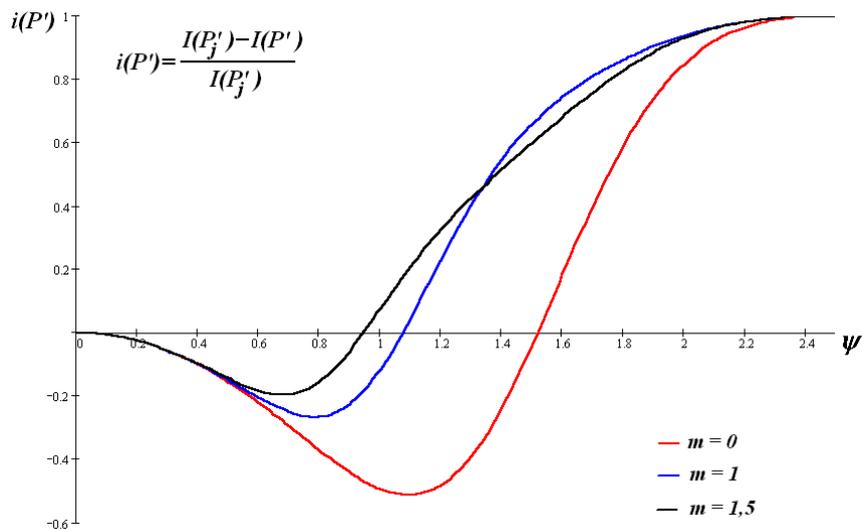


Рис. 5.12. Зависимость контраста $i(P')$ изображения двух одинаковых точек P'_j ($j=1, 2$) от величины коэффициента ψ для трех значений $m=0; 1; 1,5$

Используя ту же формулу, вычислены координаты зависимости $\psi=\psi(m)$ для четырех значений контраста $i(P')=0; 0,150; 0,268; 0,350$. Результаты вычислений представлены в виде кривых на рис. 5.13.

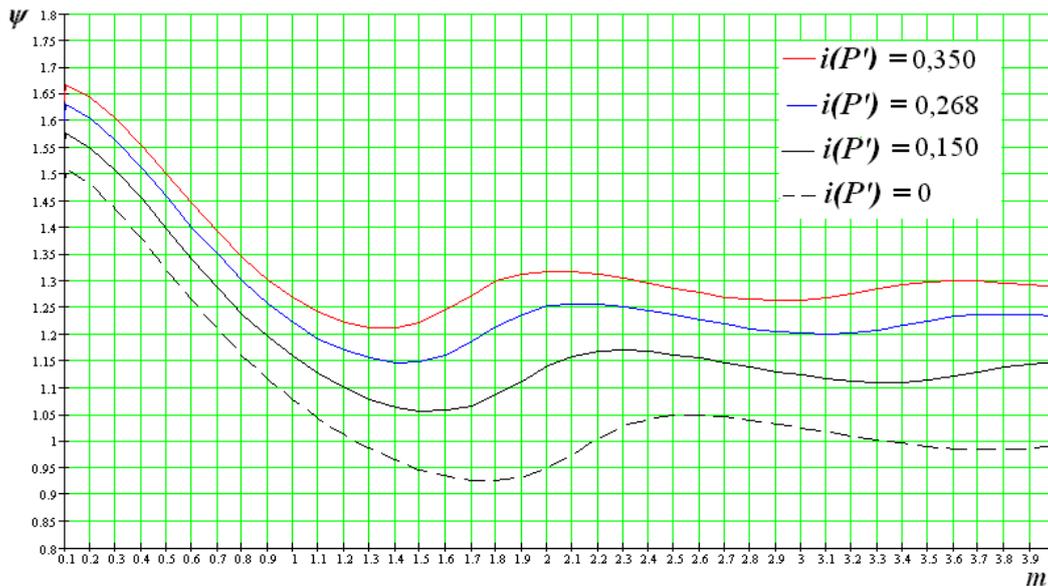


Рис. 5.13. Зависимость коэффициента ψ от величины отношения числовых апертур m для четырех значений контраста $i(P') = 0; 0,150; 0,268; 0,350$

Отсюда следует, что максимальная разрешающая способность достигается при $m \approx 1,5$.

Освещение по Кёлеру

Оптическая схема освещения по Кёлеру представлена на рис. 5.14.

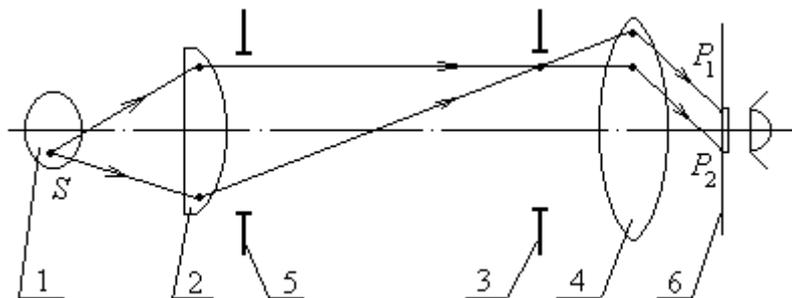


Рис. 5.14. Схема освещения по методу Кёлера

Источник света 1 коллектором 2 изображается в плоскости апертурной диафрагмы 3, расположенной в переднем фокусе конденсора 4. Полевая диафрагма 5, расположенная вблизи коллектора 2, изображается конденсором 4 в плоскости предмета 6. Два луча, исходящие из некоторой точки источника света, квазипараллельны в пространстве изображений конденсора, образуют углы α , β и γ с осями координат и пересекают плоскость предмета в точках P_1 и P_2 , как показано на рис. 5.14. Чтобы оценить предельное разрешение, достигаемое при освещении предмета по методу Кёлера, определим, прежде всего, комплексную степень когерентности μ для этой пары точек в предметной плоскости объектива микроскопа.

Пусть комплексные возмущения в точках $P_1(X_1, Y_1)$ и $P_2(X_2, Y_2)$ плоскости предмета объектива микроскопа, обусловленные возмущением в

точке S монохроматического ассоциированного источника, определяются соответственно выражениями вида:

$$U(S, P_1) = A_1 \exp(i\varphi_1), \quad U(S, P_2) = A_2 \exp(i\varphi_2). \quad (5.47)$$

Разность фаз:

$$\varphi_1 - \varphi_2 = \frac{2\pi}{\lambda_0} [p(X_1 - X_2) + q(Y_1 - Y_2)], \quad (5.48)$$

где p и q – первые два лучевые компоненты двух квазипараллельных лучей, испускаемых точкой S источника и проходящих через точки P_1 и P_2 , причем $p = n \cos \alpha$, $q = n \cos \beta$, где n – показатель преломления среды. Заметим, что малые отклонения от параллельности лучей обусловлены остаточными аберрациями оптической системы конденсора, влиянием которых вполне можно пренебречь. Подставляя выражения (5.47) и (5.48) в формулу Гопкинса (I.4.70), получаем:

$$\begin{aligned} \mu(P_1, P_2) &= \frac{1}{\sqrt{I(P_1)}\sqrt{I(P_2)}} \times \\ &\times \int_S A_1 A_2 \exp\{i k_0 [p(X_1 - X_2) + q(Y_1 - Y_2)]\} dS, \end{aligned} \quad (5.49)$$

где $I(P_1) = \int_S A_1^2 dS$, $I(P_2) = \int_S A_2^2 dS$.

Так как каждой точке $S(\xi, \eta)$ источника соответствует пара лучевых компонентов (p, q) , можно от интегрирования по S перейти к интегрированию по телесному углу

$$p^2 + q^2 \leq n_s'^2 \sin^2 \sigma_s', \quad (5.50)$$

образованному лучами, падающими на предмет.

В приближении параксиальной оптики имеем:

$$\xi = f' \cdot p, \quad \eta = f' \cdot q,$$

где f' – фокусное расстояние конденсора.

Следовательно, в принятом приближении якобиан равен постоянной величине. При этом воспользуемся формулой вида:

$$\int_S f(p, q) dS = \int_{\Omega} f(p, q) |D| dp dq,$$

где $|D|$ – абсолютная величина якобиана:

$$|D| = \frac{D(\xi, \eta)}{D(p, q)} = \begin{vmatrix} \frac{\partial \xi}{\partial p} & \frac{\partial \xi}{\partial q} \\ \frac{\partial \eta}{\partial p} & \frac{\partial \eta}{\partial q} \end{vmatrix}.$$

В общем случае якобиан изменяется в пределах области интегрирования. Однако, этим изменением по сравнению с изменением экспоненциального члена можно пренебречь. Если пренебречь также

изменением амплитуд A_1 и A_2 (если считать их постоянными величинами), то формулу (5.49) можно представить в виде:

$$\mu(P_1, P_2) = \frac{\iint_{\Omega} \exp\{ik_0[p(X_1 - X_2) + q(Y_1 - Y_2)]\} dpdq}{\iint_{\Omega} dpdq}, \quad (5.51)$$

где Ω – область интегрирования, определяемая соотношением (5.50). Вычисляя интеграл (5.51), получаем:

$$\mu(P_1, P_2) = \frac{2J_1(u_{12})}{u_{12}}, \quad u_{12} = \frac{2\pi}{\lambda_0} \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2} n'_s \sin \sigma'_s. \quad (5.52)$$

Эта формула идентична формуле (5.43) в случае критического освещения. Следовательно, комплексная степень когерентности света, падающего на плоскость предмета микроскопа, одинакова как при критическом освещении предмета, так и при освещении по методу Кёлера.

5.2. Опотехника телескопических приборов

5.2.1. Увеличительное действие телескопических систем

Предположим, что главный луч, проведенный из крайней точки наблюдаемого предмета в центр входного зрачка глаза наблюдателя, образует с оптической осью глаза некоторый угол ω . Вполне очевидно, что при пропорциональном изменении расстояния от глаза до предмета и самого угла ω не изменяет своего значения. Опыт показывает, что не изменится при этом и впечатление, производимое на наблюдателя воспринимаемым предметом, т.е. предмет глазу будет казаться неизменным.

Пусть в одной и той же плоскости на одном и том же расстоянии от наблюдателя расположены предметы l_1 и l_2 . Главные лучи, проведенные из края предметов в центр входного зрачка глаза, образуют с его оптической осью углы ω_1 и ω_2 . При этом видимая относительная величина наблюдаемых предметов определится отношением $\frac{l_1}{l_2} = \frac{tg\omega_1}{tg\omega_2}$.

Телескопическая система обладает свойством изменять видимую величину изображаемого предмета и служит для наблюдения удаленных предметов. Так как расстояние до наблюдаемых предметов значительно больше диаметра входного зрачка системы, то всякий пучок лучей, идущий от удаленной точки предмета и ограниченный входным зрачком системы можно считать параллельным. Для удобства наблюдения (без напряжения аккомодации) выходящие из системы пучки лучей тоже должны быть параллельными. При соблюдении этих условий фокусное расстояние телескопической системы будет равно бесконечности. Оптическая сила одиночной линзы в воздухе определяется формулой:

$$\varphi = (n-1) \left(\frac{1}{r_1} - \frac{1}{r_2} \right) + \frac{(n-1)^2 d}{nr_1 r_2}, \text{ где } r_1 \text{ и } r_2 - \text{ радиусы кривизны поверхностей}$$

линз; d – толщина линзы. При $\varphi = 0$ эта формула определяет взаимосвязь конструктивных параметров однолинзовой телескопической системы, свойства которой вполне соответствуют той, которую описал Роджер Бэкон. Практического применения однолинзовая телескопическая система не получила.

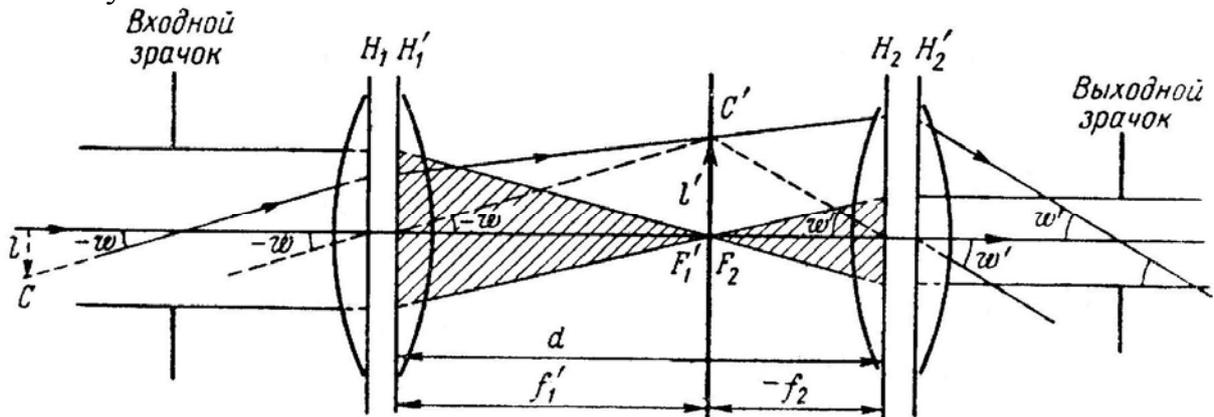


Рис. 5.15. Оптическая схема телескопической системы из двух компонентов

В простейшем случае телескопическая система состоит из двух компонентов: объектива и окуляра, как показано на рис. 5.15. Величина изображения, образованного объективом в его задней фокальной плоскости, равна $l' = -f'_{об} \text{tg } \omega$. Образованное объективом изображение наблюдатель рассматривает с помощью лупы, роль которой выполняет окуляр. Изображение, расположенное в передней фокальной плоскости окуляра, наблюдатель видит под углом, равным $2\omega'$, где $\text{tg } \omega' = \frac{l'}{f'_{ок}}$. При

этом видимое увеличение изображаемого предмета определится отношением

$$\Gamma = \frac{\text{tg } \omega'}{\text{tg } \omega} = -\frac{f'_{об}}{f'_{ок}}. \quad (5.53)$$

Пусть $D_{зр.вх.}$ – диаметр зрачка входа; $D_{зр.вых.}$ – диаметр зрачка выхода.

Тогда из подобия заштрихованных треугольников на рис. 5.15 следует, что

$$\frac{D_{зр.вх.}}{D_{зр.вых.}} = -\frac{f'_{об}}{f'_{ок}}. \text{ В результате получаем, что видимое увеличение}$$

изображения, образованного телескопической системой, равно

$$\Gamma = \frac{\text{tg } \omega'}{\text{tg } \omega} = \frac{D_{зр.вх.}}{D_{зр.вых.}} = -\frac{f'_{об}}{f'_{ок}}. \quad (5.54)$$

5.2.2. Светотехника телескопических систем

Итак, оптическая система глаза строит на сетчатке изображение наблюдаемого предмета. При этом предмету, видимому под большим углом, соответствует большее изображение на сетчатке. Увеличительное действие визуального оптического прибора в конечном счете определяется отношением линейных размеров изображения предмета на сетчатке глаза при наблюдении с помощью оптического прибора к линейным размерам изображения того же предмета при непосредственном наблюдении невооруженным глазом.

Предположим, что угловая величина наблюдаемого предмета, имеющего форму круга, равна 2ω . Световой поток, входящий в зрачок глаза, в соответствии с формулой (2.27) равен

$$\Phi_G = \frac{1}{4}(L\pi^2 d^2 \sin^2 \omega), \text{ где } d \text{ – диаметр входного зрачка глаза.}$$

Величину изображения на сетчатке глаза определим выражением $l' = f_G \operatorname{tg} \omega$, где f_G – переднее фокусное расстояние оптической системы глаза. При этом освещенность изображения на сетчатке глаза равна

$$E = \frac{\tau_G \Phi_G}{\pi l'^2} = \frac{1}{4} \tau_G L \frac{\pi^2 d^2 \sin^2 \omega}{\pi f_G^2 \operatorname{tg}^2 \omega} = \frac{\tau_G \pi L d^2}{4 f_G^2} \cos^2 \omega, \quad (5.55)$$

где τ_G – коэффициент пропускания оптических сред глаза. При наблюдении того же предмета с помощью телескопической системы (зрительной трубы) световой поток, выходящий из выходного зрачка, равен

$$\Phi_T = \frac{1}{4} L' \pi^2 D'^2 \sin^2 \omega',$$

где D' – диаметр выходного зрачка зрительной трубы; $L' = \tau_T \frac{n'^2}{n^2} L$, где τ_T – коэффициент пропускания оптической системы зрительной трубы; для зрительной трубы, как правило, $n = n'$. Величину изображения на сетчатке глаза при наблюдении предмета с помощью зрительной трубы определим выражением $l'_T = f_G \operatorname{tg} \omega'$. При этом освещенность изображения на сетчатке глаза равна

$$E_T = \frac{\tau_G \Phi_T}{\pi l'_T{}^2} = \frac{1}{4} \tau_G \tau_T L \frac{\pi^2 D'^2 \sin^2 \omega'}{\pi f_G^2 \operatorname{tg}^2 \omega'} = \frac{\tau_G \tau_T \pi L D'^2}{4 f_G^2} \cos^2 \omega' \quad (5.56)$$

$$\text{или } E_T = \frac{\tau_G \tau_T \pi L k^2 d^2}{4 f_G^2} \cos^2 \omega', \quad (5.57)$$

где $k = \frac{D'}{d}$, если $D' \leq d$ (увеличение больше нормального или равно ему);
 $k = 1$, если $D' > d$ (увеличение меньше нормального).

Субъективную яркость невозможно выразить числом, потому что она зависит не только от освещенности на сетчатке, но и от адаптации глаза. И, тем не менее, можно сравнивать две субъективные яркости в одинаковых условиях наблюдения, устанавливая их равенство или неравенство. Найдем отношение освещенностей на сетчатке глаза, определяемых выражениями (5.57) и (5.55): $i = \frac{E_T}{E} = k^2 \tau_T \frac{\cos^2 \omega'}{\cos^2 \omega}$. Величина углов ω и ω' в рассматриваемом случае не превышает единиц градусов. Поэтому можно считать, что

$$i = k^2 \tau_T. \quad (5.58)$$

Отсюда следует, что яркости L протяженного объекта при наблюдении через телескопическую систему соответствует субъективная яркость $k^2 \tau_T L$. Это означает, что применение зрительной трубы (телескопа) для наблюдения протяженных объектов не приводит к повышению их видимой яркости и в то же время может как угодно снизить ее.

Если размер светового пятна в изображении точки не превышает величины чувствительного элемента сетчатки глаза, то субъективная яркость наблюдаемого точечного источника света определится величиной излучаемого им светового потока, прошедшего через входной зрачок глаза. При наблюдении невооруженным глазом световой поток, падающий на чувствительный элемент сетчатки глаза от точечного источника света, равен освещенности на входном зрачке глаза, умноженной на его площадь:

$$\Phi_G = \frac{\pi}{4} \tau_G d^2 E. \quad (5.59)$$

При наблюдении с помощью зрительной трубы световой поток, выходящий из выходного зрачка зрительной трубы, с учетом потерь света на поверхностях и в средах оптической системы трубы равен

$$\Phi_T = \frac{\pi}{4} \tau_T D^2 E = \frac{\pi}{4} \tau_T \Gamma^2 D'^2 E. \quad (5.60)$$

При этом световой поток, падающий на чувствительный элемент сетчатки глаза от точечного источника света, определится выражением

$$\Phi_T = \frac{\pi}{4} \tau_T \tau_G \Gamma^2 k^2 d^2 E. \quad (5.61)$$

Таким образом, световой поток на сетчатке глаза при наблюдении с помощью телескопа больше светового потока при непосредственном наблюдении точечного источника света в i раз, равное отношению:

$$i = \frac{\Phi_T}{\Phi_G} = \tau_T \Gamma^2 k^2. \quad (5.61)$$

Отсюда следует, что освещенности E на входном зрачке невооруженного глаза, создаваемой точечным источником излучения, соответствует субъективная освещенность, равная при наблюдении его

через телескопическую систему. В общем случае видимость точечного источника света определяется его блеском (освещенностью, создаваемой в точке наблюдения при нормальном падении света) и яркостью фона, на котором он наблюдается. Эффективность телескопической системы для наблюдения точечных источников света можно определить величиной q , равной отношению

$$q = \frac{E_0}{E}, \quad (5.62)$$

где E_0 и E – пороговые значения блеска при наблюдении невооруженным глазом и с помощью оптического прибора. При одном и том же блеске звезд видимость их определяется яркостью неба (фона) L :

$$E_0 = f(L). \quad (5.63)$$

При наблюдении через телескопическую систему яркости фона L , как протяженного объекта, соответствует субъективная яркость $\tau_T k^2 L$, а освещенности E от точечного источника света – субъективный блеск $\tau_T \Gamma^2 k^2 E$. При этом пороговые условия видимости могут быть определены соотношением $\tau_T \Gamma^2 k^2 E = f(\tau_T k^2 L)$.

Пользуясь формулами (5.62), (5.63) и (5.64), получаем

$$q = \frac{E_0}{E} = \tau_T \Gamma^2 k^2 \frac{f(L)}{f(\tau_T k^2 L)}. \quad (5.65)$$

Если при яркости фона 10^{-6} кд/м² пороговый блеск для центрального зрения постоянен и равен $2 \cdot 10^{-8}$ лк, а для периферического зрения – $2 \cdot 10^{-9}$ лк, то при увеличении яркости фона пороговый блеск увеличивается. Зависимость $E_0 = f(L)$ для центрального зрения в пределах изменения яркости $L = 10^{-3} \dots 30$ кд/м² аппроксимируется функцией: $E_0 = 1,6 \cdot 10^{-7} L^n$, где $n = 0,6 + 0,1 \lg L$. При $L > 30$ кд/м² величина $n = 0,75$, а при $L < 10^{-3}$ кд/м² пороговый блеск постоянен и равен $2 \cdot 10^{-8}$ лк. Пороговый блеск для периферии в диапазоне изменения яркости $L = 10^{-6} \dots 10^{-3}$ кд/м² рассчитывается по формуле $E_0 = 1,6 \cdot 10^{-7} \cdot L^{0,3}$.

5.2.3. Проницающая сила телескопа

В 1846 году Вебер установил, что чувствительность глаза к изменению освещенности пропорциональна относительному изменению освещенности, т.е.

$$\Delta m = c \frac{\Delta E}{E}. \quad (5.66)$$

Действительно, освещенность поверхности шестью свечами заметно отличается от освещенности той же поверхности пятью свечами, тогда как отличие освещенности 150 свечами от освещенности 151 свечей

практически незаметно. Фехнер в 1860 году сделал допущение, что едва заметные изменения в ощущении глазом света можно считать дифференциалами. Это позволило ему записать выражение (5.66) в виде дифференциального уравнения $dm = c \frac{dE}{E}$. Интегрирование этого уравнения привело к установлению зависимости интенсивности раздражения глаза светом от освещенности в виде

$$m = c \ln E + k, \quad (5.67)$$

где c и k – некоторые постоянные. Эту зависимость принято называть основным психофизическим законом Вебера-Фехнера. Выражение (5.67) можно переписать в виде

$$m = c \frac{\lg E}{\lg e} + k = \tilde{c} \lg E + k. \quad (5.68)$$

Первой фотометрической работой в истории науки было разделение звезд по их блеску на шесть классов – звездных величин, выполненное Гиппархом во II веке до Р.Х. В звездном каталоге Птолемея, приложенном к его «Альмагесту», содержатся оценки звездных величин свыше 1000 звезд. Сопоставляя их с современными объективными измерениями освещенности от этих звезд, установили, что когда одна звезда имеет звездную величину на единицу больше, чем другая (это записывается так: $m_1 - m_2 = 1$), то освещенность от первой в 2,5 раза меньше, чем от второй. Из выражения (5.68) следует, что

$$m_1 - m_2 = \tilde{c} \lg \frac{E_1}{E_2}. \quad (5.69)$$

Пусть величина m в этом выражении представлена в звездных величинах. Тогда при $m_1 - m_2 = 1$ отношение $\frac{E_1}{E_2} = \frac{1}{2,5} = 0,4$. Отсюда

следует, что $\tilde{c} = -\frac{1}{\lg 2,5} = -\frac{1}{0,398}$. В 1850 году по предложению Погсона

отношение блеска звезд при разности в пять звездных величин было принято равным 100. Это означает, что при $m_1 - m_2 = 5$ отношение $\frac{E_1}{E_2} = \frac{1}{100}$, и, следовательно, $\tilde{c} = -2,5$. Тогда выражение (5.69) можно

записать в виде

$$m_1 - m_2 = -2,5 \lg \frac{E_1}{E_2}. \quad (5.70)$$

Отсюда $\lg \frac{E_1}{E_2} = 0,4(m_2 - m_1) = (m_2 - m_1) \lg 2,512$

$$\text{или } \frac{E_1}{E_2} = 2,512^{m_2 - m_1}. \quad (5.71)$$

Это выражение, определяющее зависимость между блеском двух звезд E_1 и E_2 и их видимыми звездными величинами m_1 и m_2 , называется формулой Погсона.

Для расчета энергетических и световых характеристик приемников различной спектральной чувствительности необходимо иметь эталонную звезду нулевой звездной величины. В видимой области спектра за эталонную звезду нулевой звездной величины принимается та, которая создает на границе атмосферы освещенность, равную

$$E_0 = 2,78 \cdot 10^{-6} \text{ лк} = 2,78 \cdot 10^{-10} \text{ лм/см}^2.$$

Величина эталонной энергетической освещенности при $m_0 = 0$ принимается равной $E_0 = 3,92 \cdot 10^{-13} \text{ Вт/см}^2$.

Интегральную освещенность, создаваемую звездой звездной величины m , можно вычислить, используя формулу Погсона:

$$E, \text{ лк} = E_0 \cdot 2,512^{-m} = 2,78 \cdot 10^{-6} \cdot 2,512^{-m};$$

$$E, \text{ Вт/см}^2 = E_0 \cdot 2,512^{-m} = 3,92 \cdot 10^{-13} \cdot 2,512^{-m}.$$

Одним из важнейших параметров, определяющих эффективность применения большого телескопа, является его проникающая сила. Проникающей силой телескопа называют звездную величину наиболее слабой звезды, различимой с его помощью в ясную безлунную ночь близ зенита. Рассмотрим соотношения, определяющие проникающую силу телескопа при произвольной яркости неба. Этот вопрос представляет практический интерес для астрономических наблюдений и оптической сигнализации ночью и днем.

Пусть в формуле (5.70) $m_2 = 0$ и, следовательно, $E_2 = 2,78 \cdot 10^{-6} \text{ лк}$. При этом шкала звездных величин определяется формулой

$$m = -2,5 \lg E_m - 13,89.$$

При пороговом блеске E_0 наблюдаемого точечного источника света невооруженным глазом соответствующая звездная величина его равна

$$m_0 = -2,5 \lg E_0 + \tilde{k}, \quad (5.72)$$

где $\tilde{k} = -13,89$. При этом зависимость порогового блеска E_0 точечного источника света от яркости фона L определяется соотношением вида (5.63). Зависимость величины m_0 от $\lg L$, где L – яркость фона в кд/м^2 , изображена в виде кривой на рис. 5.16. Для наглядности у соответствующих участков горизонтальной оси указаны условия, при которых небо имеет соответствующую яркость.

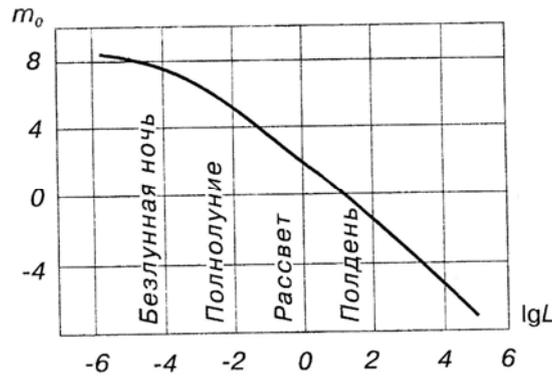


Рис. 5.16. Зависимость величины m_0 от $\lg L$

Обозначим через E пороговое значение блеска точечного источника света при наблюдении в телескоп, а через m – проникающую силу телескопа. При этом

$$m = -2,5 \lg E + \tilde{k}. \quad (5.73)$$

Яркости фона L при наблюдении с помощью телескопа будет соответствовать субъективная яркость $\tau_T k^2 L$, а освещенности E от наблюдаемого источника – субъективный блеск $\tau_T \Gamma^2 k^2 E$. Будем считать, что пороговые условия видимости определены соотношением (5.64). В результате получаем

$$m - m_0 = 2,5 \lg \frac{E_0}{E} = 2,5 \lg \frac{\tau_T \Gamma^2 k^2 f(L)}{f(\tau_T k^2 L)}.$$

Считая, что в рассматриваемом интервале значений яркости $E_0 = f(L) = cL^n$, а $E = f(\tau_T k^2 L) = c(\tau_T k^2 L)^n$, получаем

$$m - m_0 = 5 \lg \tau_T^{\frac{1-n}{2}} \Gamma k^{1-n}. \quad (5.74)$$

Величины n и d зависят от яркости L . Эта зависимость с достаточной для практики точностью характеризуется величинами, представленными в табл. 5.1.

Таблица 5.1. Зависимость коэффициента n и диаметра зрачка глаза d от яркости L фона

Яркость фона, кд/м ²	n	d , см
10^{-6}	0	0,6
10^{-3}	1/4	0,5
0,3	1/2	0,4
30	3/4	0,3
3000	1	0,25

При исчезающе малой яркости фона $n = 0$. При этом

$$m - m_0 = 5 \lg \tau_T^{\frac{1}{2}} \Gamma k. \text{ При } D' > d \text{ отношение } k = 1. \text{ В этом случае}$$

$m - m_0 = 5 \lg \tau^2 \frac{1}{d} D$. При наблюдении звезд в дневное время (при больших величинах яркости фона) коэффициент $n = 1$. В этом случае, как следует из формулы (5.74), имеем $m - m_0 = 5 \lg \Gamma$. В диапазоне изменения яркости ночного неба можно считать, что коэффициент $n = \frac{1}{4}$. Тогда при $k = 1$ получаем $m - m_0 = 5 \lg \tau^{\frac{3}{8}} \Gamma$.

Итак, если используется вся площадь входного зрачка прибора, то при черном фоне неба проникающая сила телескопа полностью определяется диаметром объектива. На фоне ночного неба проникающая сила телескопа в основном определяется диаметром объектива. Этим определяется потребность в создании телескопов с предельно достижимым при существующих технологических возможностях диаметром входного зрачка. При наблюдении звезд в условиях светлого неба проникающая сила телескопа полностью определяется угловым (видимым) увеличением образованного изображения.

Приложение

Математический инструментарий в оптотехнике

А. Векторное исчисление

1. Векторная алгебра

Величины, значения которых могут быть выражены действительными числами, называются *скалярами* (например, масса, заряд, температура, работа и т.п.). Величины же, значения которых определяются как числовым значением, так и направлением в пространстве, называются *векторами* (например, скорость, ускорение, сила, напряженность электрического и магнитного полей и т.п.).

Геометрический вектор – это направленный отрезок в пространстве. *Длина* вектора **a** называется его *модулем* и обозначается $|\mathbf{a}|$. *Единичные векторы* – это векторы, длина которых равна единице. *Нулевой вектор* **0** – это вектор, начало и конец которого совпадают; его модуль равен нулю при неопределенном направлении. Два вектора считаются *равными*, если равны их модули и совпадают их направления. Векторы, которые получаются из данного вектора **a** путем параллельного переноса (все такие векторы *равны a*), называют *свободными векторами*, порождаемыми вектором **a**. Два одинаково направленных и равных по длине вектора называются *эквивалентными*. Векторы, которые получаются из **a** путем параллельного переноса вдоль **a**, называются *скользящими векторами*. Если вектор нельзя переносить по физическим причинам (постоянная точка приложения), то говорят о *связанном векторе*.

1.1. Умножение на скаляр и сложение

Если α – действительное число и **a** – вектор, то произведение $\alpha \cdot \mathbf{a}$

Также есть вектор с длиной $|\alpha||\mathbf{a}|$ и направлением, совпадающим с направлением вектора **a** при $\alpha > 0$, и с направлением, противоположным направлению вектора **a** при $\alpha < 0$.

Сумма **a** + **b** двух векторов **a** и **b** есть вектор, который имеет начало, совпадающее с началом **a**, и конец, совпадающий с концом **b** (правило треугольника). Подобным образом сумма **a** + **b** + ... + **e** нескольких векторов **a**, **b**, ..., **e** определяется как вектор **f**, который замыкает ломаную, составленную из **a**, **b**, ..., **e**. *Разность* **a** – **b** векторов **a**, **b** рассматривается как сумма векторов **a** и $-\mathbf{b}$.

Правила действий с векторами

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}, \quad \mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c},$$

$$\alpha(\beta\mathbf{a}) = (\alpha\beta)\mathbf{a}, \quad (\alpha + \beta)\mathbf{a} = \alpha\mathbf{a} + \beta\mathbf{a},$$

$$\alpha(\mathbf{a} + \mathbf{b}) = \alpha\mathbf{a} + \alpha\mathbf{b}, \quad \alpha\mathbf{a} = |\alpha| \cdot |\mathbf{a}|,$$

$$\left| |\mathbf{a}| - |\mathbf{b}| \right| \leq |\mathbf{a} \pm \mathbf{b}| \leq |\mathbf{a}| + |\mathbf{b}|.$$

Под *линейной комбинацией векторов* \mathbf{a} , \mathbf{b} , ..., \mathbf{d} с действительными коэффициентами α , β , ..., δ понимают вектор, имеющий вид

$$\mathbf{e} = \alpha\mathbf{a} + \beta\mathbf{b} + \dots + \delta\mathbf{d}.$$

Два вектора \mathbf{a} , \mathbf{b} называются *коллинеарными*, если имеются такие действительные числа α , β , что $\alpha\mathbf{a} + \beta\mathbf{b} = \mathbf{0}$, причем α , β не равны одновременно нулю (геометрический смысл: прямые, проходящие в направлениях \mathbf{a} и \mathbf{b} , параллельны). Три вектора \mathbf{a} , \mathbf{b} , \mathbf{c} называются *компланарными*, если существуют такие действительные числа α , β , γ , что

$\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} = \mathbf{0}$ и α , β , γ не являются одновременно нулями (геометрический смысл: \mathbf{a} , \mathbf{b} , \mathbf{c} параллельны одной плоскости). Если \mathbf{a} , \mathbf{b} не коллинеарны или \mathbf{a} , \mathbf{b} , \mathbf{c} не компланарны, то их называют *линейно независимыми* на плоскости или в пространстве. Два ненулевых вектора \mathbf{a} , \mathbf{b} *ортогональны* (обозначение: $\mathbf{a} \perp \mathbf{b}$), если они взаимно перпендикулярны. Такие векторы всегда линейно независимы. Три попарно ортогональных ненулевых вектора \mathbf{a} , \mathbf{b} , \mathbf{c} также образуют тройку линейно независимых векторов.

Координаты вектора

Если заданы три линейно независимых вектора \mathbf{e}_1 , \mathbf{e}_2 , \mathbf{e}_3 , то каждый вектор \mathbf{a} можно однозначно представить в виде

$$\mathbf{a} = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3.$$

Величины a_i называются *аффинными (или контравариантными) координатами* \mathbf{a} относительно $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. Это кратко записывается так: $\mathbf{a} = (a_1, a_2, a_3)$. Равные векторы обладают одинаковыми аффинными координатами. Если $\mathbf{a} = (a_1, a_2, a_3)$, $\mathbf{b} = (b_1, b_2, b_3)$, то справедливы следующие соотношения: $\alpha\mathbf{a} = (\alpha a_1, \alpha a_2, \alpha a_3)$,

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2, a_3 + b_3).$$

Векторы \mathbf{e}_1 , \mathbf{e}_2 , \mathbf{e}_3 образуют *правую систему координат*, если они имеют такую же ориентацию, как соответственно большой, указательный и средний пальцы правой руки; в противном случае говорят о *левой системе координат*. Если, в частности, в качестве \mathbf{e}_1 , \mathbf{e}_2 , \mathbf{e}_3 выбирают правую систему из трех единичных векторов \mathbf{i} , \mathbf{j} , \mathbf{k} , попарно перпендикулярных друг другу, то $\mathbf{a} = a_x\mathbf{i} + a_y\mathbf{j} + a_z\mathbf{k}$, где a_x, a_y, a_z называются *прямоугольными декартовыми координатами* вектора \mathbf{a} .

1.2. Умножение векторов

Скалярное произведение векторов \mathbf{a} , \mathbf{b} , обозначаемое \mathbf{ab} , (\mathbf{a}, \mathbf{b}) , $\mathbf{a} \cdot \mathbf{b}$ или $(\mathbf{a} \cdot \mathbf{b})$ есть число $\mathbf{ab} = |\mathbf{a}| \cdot |\mathbf{b}| \cos\varphi$, где φ – угол между векторами \mathbf{a} и \mathbf{b} .

Под *векторным произведением* векторов \mathbf{a} , \mathbf{b} , обозначаемым $\mathbf{a} \times \mathbf{b}$, $[\mathbf{a} \cdot \mathbf{b}]$, $[\mathbf{a}, \mathbf{b}]$, $[\mathbf{ab}]$, или $[\mathbf{a} \times \mathbf{b}]$, понимают вектор \mathbf{c} , имеющий длину $|\mathbf{c}| = |\mathbf{a} \times$

$|\mathbf{b}| = |\mathbf{a}| \cdot |\mathbf{b}| \sin \varphi$ (площадь параллелограмма, построенного на \mathbf{a} и \mathbf{b} как на сторонах) и направленный перпендикулярно к \mathbf{a} и \mathbf{b} , причем так, что векторы \mathbf{a} , \mathbf{b} , $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ образуют правую тройку векторов.

Свойства произведений.

- 1) $\mathbf{ab} = \mathbf{ba}$ (коммутативность), но $\mathbf{a} \times \mathbf{b} = -(\mathbf{b} \times \mathbf{a})$ (антикоммутативность);
- 2) $(\alpha\mathbf{a})\mathbf{b} = \alpha(\mathbf{ab})$ и $(\alpha\mathbf{a}) \times \mathbf{b} = \alpha(\mathbf{a} \times \mathbf{b})$ (ассоциативность) при умножении на действительное число;
- 3) $\mathbf{a}(\mathbf{b} + \mathbf{c}) = \mathbf{ab} + \mathbf{ac}$ и $\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$ (дистрибутивность);
- 4) $\mathbf{ab} = 0$, если \mathbf{a} и \mathbf{b} взаимно перпендикулярны; $\mathbf{a} \times \mathbf{b} = 0$, если \mathbf{a} и \mathbf{b} коллинеарны;
- 5) $\mathbf{aa} = \mathbf{a}^2 = |\mathbf{a}|^2$, $\mathbf{a} \times \mathbf{a} = 0$;
- 6) в общем случае $\mathbf{a}(\mathbf{bc}) \neq (\mathbf{ab})\mathbf{c}$;
- 7) линейные комбинации векторов можно перемножать, как скалярные многочлены, однако для векторного произведения важна последовательность сомножителей.

Двойное векторное произведение $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$ – вектор, компланарный векторам \mathbf{b} и \mathbf{c} , вычисляемый по формуле $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{ac}) - \mathbf{c}(\mathbf{ab})$.

В общем случае $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \neq (\mathbf{a} \times \mathbf{b}) \times \mathbf{c}$.

Смешанное произведение $(\mathbf{a} \times \mathbf{b})\mathbf{c}$ есть скаляр, абсолютная величина которого равна объему параллелепипеда, построенного на векторах \mathbf{a} , \mathbf{b} , \mathbf{c} как на ребрах. Смешанное произведение положительно, если $\mathbf{a}, \mathbf{b}, \mathbf{c}$ образуют правую систему; в противном случае оно отрицательно. Вместо $(\mathbf{a} \times \mathbf{b})\mathbf{c}$ пишут также (\mathbf{abc}) или \mathbf{abc} . Перестановка двух сомножителей в смешанном произведении \mathbf{abc} меняет знак: $\mathbf{abc} = -\mathbf{acb} = -\mathbf{bac} = -\mathbf{cba}$, циклическая перестановка не меняет знака: $\mathbf{abc} = \mathbf{bca} = \mathbf{cab}$. Три вектора тогда и только тогда линейно независимы, когда $\mathbf{abc} \neq 0$. Кроме того, имеют место формулы

$$(\mathbf{a} \times \mathbf{b})(\mathbf{c} \times \mathbf{d}) = (\mathbf{ac})(\mathbf{bd}) - (\mathbf{bc})(\mathbf{ad}) \quad (\text{тождество Лагранжа})$$

$$(\mathbf{abc})(\mathbf{efg}) = \begin{vmatrix} \mathbf{ae} & \mathbf{af} & \mathbf{ag} \\ \mathbf{be} & \mathbf{bf} & \mathbf{bg} \\ \mathbf{ce} & \mathbf{cf} & \mathbf{cg} \end{vmatrix}.$$

Если векторы \mathbf{a} , \mathbf{b} , \mathbf{c} заданы в прямоугольных декартовых координатах:

$$\mathbf{a} = a_x \mathbf{i} + a_y \mathbf{j} + a_z \mathbf{k}, \quad \mathbf{b} = b_x \mathbf{i} + b_y \mathbf{j} + b_z \mathbf{k}, \quad \mathbf{c} = c_x \mathbf{i} + c_y \mathbf{j} + c_z \mathbf{k},$$

то произведения вычисляются следующим образом.

$$\text{Скалярное произведение: } \mathbf{ab} = a_x b_x + a_y b_y + a_z b_z.$$

Векторное произведение:

$$\mathbf{a} \times \mathbf{b} = (a_y b_z - a_z b_y) \mathbf{i} + (a_z b_x - a_x b_z) \mathbf{j} +$$

$$+(a_x b_y - a_y b_x) \mathbf{k} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix}.$$

$$\text{Смешанное произведение: } \mathbf{abc} = \begin{vmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{vmatrix}.$$

2. Векторный анализ

2.1. Векторные функции скалярного аргумента

Если каждому значению скалярного аргумента t поставить в соответствие вектор $\mathbf{r}(t)$, то $\mathbf{r}(t)$ называется *векторной функцией* (*вектор-функцией*) скалярного аргумента t . Если начало вектора $\mathbf{r}(t)$ (радиус-вектора) поместить в постоянную точку O , то конец радиус-вектора $\mathbf{r}(t)$ опишет пространственную кривую, которую называют *годографом* векторной функции. Если t означает время, то $\mathbf{r}(t)$ описывает *траекторию* движения материальной точки. Если $\mathbf{r}(t)$ разложить по базисным векторам $\mathbf{i}, \mathbf{j}, \mathbf{k}$ прямоугольной декартовой системы координат, то

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k},$$

причем компоненты $x(t), y(t), z(t)$ являются функциями от t . Параметрическое представление пространственной кривой (годографа) или траектории движения имеет вид $x = x(t), y = y(t), z = z(t)$.

Дифференцирование векторной функции

Если существует предел $\frac{d\mathbf{r}}{dt} = \lim_{\Delta t \rightarrow 0} \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{\Delta \mathbf{r}}{\Delta t}$, то $\frac{d\mathbf{r}}{dt}$

называется производной от $\mathbf{r}(t)$ в точке t . В другой записи $\mathbf{r}'(t)$ или $\dot{\mathbf{r}}(t)$. В декартовой системе координат: $\mathbf{r}'(t) = x'(t)\mathbf{i} + y'(t)\mathbf{j} + z'(t)\mathbf{k}$.

Вектор $\mathbf{r}'(t)$ имеет направление касательной к годографу в точке t и направлен в сторону, отвечающую возрастанию параметра t . Длина $\mathbf{r}'(t)$ зависит от выбора параметра t . Если t есть длина дуги, то $\left| \frac{d\mathbf{r}}{dt} \right| = 1$. Если t

означает время, а $\mathbf{r}(t)$ – траекторию движения материальной точки, то $\mathbf{r}'(t)$ – *вектор скорости*, $|\mathbf{r}'(t)|$ – величина скорости.

Правила дифференцирования

$$\frac{d}{dt}(r_1 + r_2) = \frac{dr_1}{dt} + \frac{dr_2}{dt},$$

$$\frac{d}{dt}(\varphi r) = \frac{d\varphi}{dt} r + \varphi \frac{dr}{dt} \quad (\varphi(t) \text{ – скалярная функция от } t),$$

$$\frac{d}{dt}(r_1 r_2) = \frac{dr_1}{dt} r_2 + r_1 \frac{dr_2}{dt},$$

$$\frac{d}{dt}(r_1 \times r_2) = \frac{dr_1}{dt} \times r_2 + r_1 \times \frac{dr_2}{dt} \quad (\text{множители нельзя менять местами}),$$

$$\frac{d}{dt}r(\varphi(t)) = \frac{dr}{d\varphi} \frac{d\varphi}{dt}.$$

Если $\mathbf{r}(t)$ – единичный вектор, то годограф лежит на единичной сфере и касательная всегда перпендикулярна радиус-вектору, т.е. $\mathbf{r} \frac{d\mathbf{r}}{dt} = 0$.

2.2. Введение в теорию поля

Говорят, что в пространстве задано *поле некоторой величины*, если в каждой точке пространства (или в некоторой его части) *определено значение этой величины*. Поле может быть *скалярным* или *векторным* в зависимости от характера исследуемой величины. Например, плотность заряда в различных точках изолированного наэлектризованного тела представляет собой скалярную функцию точки. Электрические заряды создают скалярное поле плотности и векторное поле электрических сил.

В математической физике часто рассматриваются величины, которые зависят не только от положения точки, но и еще от какой-либо другой переменной, в большинстве случаев от времени. Отсюда следует, что поле может быть *стационарным* (установившимся), если оно не меняется с течением времени в каждой точке пространства, или *нестационарным* (неустановившимся), если такое изменение имеет место.

Итак, поле скалярной величины f можно, например, записать в виде

$$f = f(x, y, z, t),$$

где t – время.

Однако, время, как правило, рассматривают не как равноправную четвертую координату, а скорее как некоторый дополнительный параметр в трехмерном пространстве, при этом поле рассматривается в любой, но фиксированный момент времени. Таким образом, с формальной точки зрения стационарное поле – это просто функция трех переменных x, y, z . Однако, надо иметь в виду, что координаты в пространстве можно ввести по-разному. При этом выражение $f(x, y, z)$ будет изменяться, но в любой данной точке M значение f , конечно, не зависит от выбора системы координат. Поэтому часто говорят, что f представляет собой «функцию точки» $f = f(M)$, так как задание точки M полностью определяет соответствующее значение f , т.е. значение величины f в точке M . Функция точки при рассмотрении поля является первичной по отношению к функции координат, так как поле имеет смысл и может быть исследовано без всяких систем координат.

Пусть S – замкнутая поверхность, ограничивающая объем V ; M – переменная точка в объеме V (или на поверхности S); dS – вектор, длина которого равна площади элемента поверхности dS , окружающей точку M , а направление совпадает с внешней нормалью к поверхности S . Пусть, далее, f – скалярная, а \mathbf{A} – векторная функция точки M . Предполагается, что они непрерывны вместе со своими первыми производными в любой точке объема V и его ограничивающей поверхности S .

Наиболее важные поля характеризуются следующими тремя функциями:

1. *Градиент – векторная функция*, аргументом которой является скалярная функция f точки:

$$\text{grad } f = \bar{i} \frac{\partial f}{\partial x} + \bar{j} \frac{\partial f}{\partial y} + \bar{k} \frac{\partial f}{\partial z}. \quad (1)$$

Выберем в пространстве некоторую точку M . Рассмотрим изменение *стационарного скалярного* поля $f = f(x, y, z)$ при перемещении из точки M в направлении, определяемом вектором \mathbf{r} на величину $d\mathbf{r}$. Используя правило скалярного умножения двух векторов, получаем

$$\text{grad } f \cdot d\mathbf{r} = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz = df.$$

Поверхности уровня определяются равенством $f(x, y, z) = \text{const}$. Уравнение поверхности уровня, проходящей через точку (x_0, y_0, z_0) , имеет вид

$$f(x, y, z) = f(x_0, y_0, z_0).$$

Оно геометрически определяет одну поверхность, если функция $f(x, y, z)$ однозначна, что почти всегда и встречается в физике. Так как при любом перемещении $d\mathbf{r}$ точки M по этой поверхности $df = 0$, то $d\mathbf{r}$ и $\text{grad } f$ перпендикулярны друг другу, а, следовательно, вектор $\text{grad } f$ в каждой точке нормален к рассматриваемой поверхности уровня, проходящей через эту точку. Вектор $\text{grad } f$ полностью описывает поведение функции f в окрестности рассматриваемой точки M . В частности, самое быстрое изменение f происходит при перемещении по нормали к поверхности уровня, величина и направление которого определяются вектором $\text{grad } f$.

2. *Дивергенция – скалярная функция*, аргументом которой является векторная функция \mathbf{A} точки:

$$\text{div } \mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}. \quad (2)$$

Поток вектора \mathbf{A} через площадку dS равен

$$dQ = \mathbf{A} \cdot \mathbf{n} \cdot dS,$$

где \mathbf{n} – единичный вектор внешней нормали к площадке dS .

При этом поток Q вектора \mathbf{A} через конечную поверхность S равен $Q = \int_S \mathbf{A} \cdot \mathbf{n} \cdot dS$.

Величину Q можно истолковать как количество векторных линий, начинающихся внутри объема V , ограниченного поверхностью S . Если $Q > 0$, то говорят, что в V имеется источник векторных линий, а Q называется *обильностью* этого источника.

Пусть векторные линии возникают по всему пространству. При этом объемная плотность источника в любой точке M пространства равна

$$\lim_{\Delta V \rightarrow M} \frac{\Delta Q}{\Delta V} = \lim_{\Delta V \leftarrow M} \frac{\int_{\Delta S} \mathbf{A} \cdot d\mathbf{S}}{\Delta V}. \quad (3)$$

Объемная плотность источника называется также *дивергенцией* (по-русски, расходимостью) векторного поля \mathbf{A} и обозначается через $\text{div} \mathbf{A}$. Таким образом, можно сказать, что дивергенция векторного поля – это количество векторных линий, начинающихся в бесконечно малом объеме (или, что то же самое, поток поля \mathbf{A} через поверхность этого объема), в расчете на единицу этого объема. Предел, определяемый формулой (3), представляет собой объемную производную векторного поля. Поэтому формулу (3) можно переписать в виде

$$\text{div} \mathbf{A} = \frac{dQ}{dV},$$

т.е. $dQ = \text{div} \mathbf{A} \cdot dV$. (4)

Но $dQ = \mathbf{A} \cdot d\mathbf{S}$. Производя суммирование, получаем формулу Остроградского-Гаусса

$$\oint_S \mathbf{A} \cdot d\mathbf{S} = \int_V \text{div} \mathbf{A} \cdot dV,$$

где V – конечный объем; S – его поверхность.

Определение дивергенции (3) было дано в инвариантной форме, не зависящей от выбора системы координат. Представляет интерес также вывести формулу для вычисления дивергенции с помощью декартовой системы координат x, y, z . Для этого воспользуемся тем, что в формуле (4) форма элементарного объема dV несущественна, и выберем в качестве этого объема бесконечно малый прямоугольный параллелепипед с ребрами, параллельными осям координат, как показано на рис. 1.

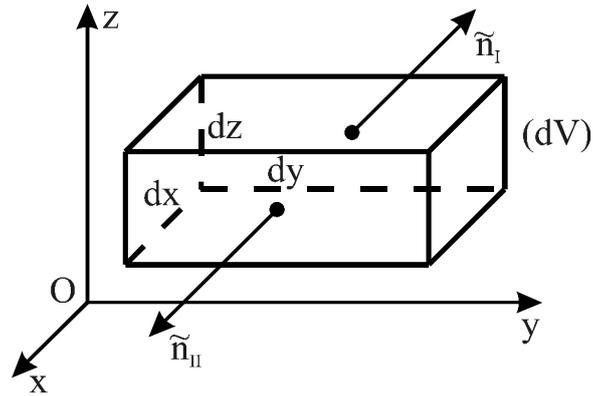


Рис. 1. Определение дивергенции поля в декартовой системе координат

Тогда поток dQ вектора \mathbf{A} через поверхность параллелепипеда (т.е. числитель дроби, стоящей в первой формуле (4)) можно представить в виде суммы шести слагаемых, соответствующих шести граням параллелепипеда. Рассмотрим сумму двух из этих слагаемых, отвечающих задней и передней граням, которые мы обозначим соответственно цифрами I и II.

$$\text{Тогда } (A_n)_{II} = (A_x)_{II}, (A_n)_I = -(A_x)_I,$$

а указанную сумму можно записать в виде

$$(\int A_n d\sigma)_I + (\int A_n d\sigma)_{II} = -\int (A_x)_I d\sigma + \int (A_x)_{II} d\sigma = \int [(A_x)_{II} - (A_x)_I] d\sigma.$$

Подынтегральная функция с точностью до малых высшего порядка равна $\partial_x A_x = \frac{\partial A_x}{\partial x} dx$; это «частный дифференциал» от A_x по x ,

получающийся из-за того, что точки передней грани отличаются от соответствующих точек задней грани значением координаты x . Поэтому весь интеграл, с точностью до членов высшего порядка малости, равен

$$\frac{\partial A_x}{\partial x} dx \int d\sigma = \frac{\partial A_x}{\partial x} dx dy dz .$$

Проводя аналогичные вычисления для двух других пар граней, получим весь элементарный поток

$$d\theta = \left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right) dx dy dz .$$

Но так как $dV = dx dy dz$, то в силу первой формулы (4) окончательно получаем формулу (2):

$$\text{div } \mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} .$$

3. Вихрь – векторная функция, аргументом которой является векторная функция \mathbf{A} точки

$$\text{rot } \mathbf{A} = \bar{i} \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) + \bar{j} \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) + \bar{k} \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right). \quad (5)$$

Пусть в пространстве, в котором задано поле вектора \mathbf{A} , выбрана ориентированная линия (L) , т.е. линия, для которой указано направление ее обхода. Криволинейным называется интеграл векторной функции \mathbf{A} по линии (L) : $\int_{(L)} \mathbf{A} \cdot d\mathbf{r}$, где вектор $d\mathbf{r}$ совпадает с касательной в

рассматриваемой точке в направлении обхода. Если линия (L) замкнутая, то линейный интеграл

$$\Gamma = \oint_{(L)} \mathbf{A} \cdot d\mathbf{r} = \oint_{(L)} (A_x dx + A_y dy + A_z dz) \quad (6)$$

называется циркуляцией вектора \mathbf{A} по линии (L) . Циркуляция обладает следующим важным свойством аддитивности (сложения). Допустим, что некоторая ориентированная незамкнутая поверхность (S) разбита на несколько частей, например, на три части: (S_1) , (S_2) и (S_3) , как показано на рис. 2.

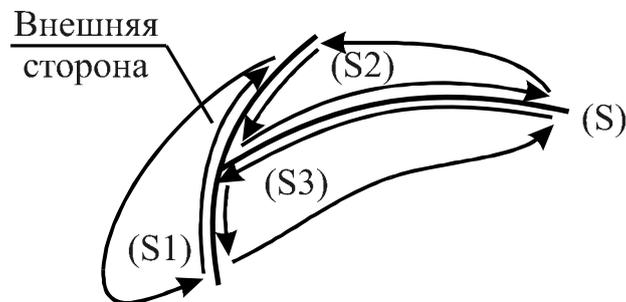


Рис. 2. Свойство аддитивности циркуляции вектора

Обозначив контуры (S) и этих частей в соответствии с ориентацией (S) (по правилу правого винта) через (L) , (L_1) , (L_2) и (L_3) , а соответствующие циркуляции – через Γ , Γ_1 , Γ_2 и Γ_3 , получаем $\Gamma = \Gamma_1 + \Gamma_2 + \Gamma_3$. Свойство аддитивности дает возможность говорить о том, что циркуляция (6) «порождается» на поверхности (S) , а тем самым и о «плотности порождения циркуляции», т.е. о циркуляции, порождаемой бесконечно малым куском поверхности в расчете на единицу площади этого куска. При этом общая циркуляция равна

$$\Gamma = \int_{(S)} d\Gamma.$$

Опишем из некоторой точки векторного поля \mathbf{A} , как из центра, окружность C малого радиуса r , лежащую в плоскости, перпендикулярной любому фиксированному направлению n . При этом поверхностная плотность порождения циркуляции равна

$$\frac{d\Gamma}{dS} = \lim_{r \rightarrow 0} \frac{\int_{(C)} \mathbf{A} \cdot d\mathbf{r}}{\pi r^2} = n \cdot \text{rot } \mathbf{A}. \quad (7)$$

Отсюда следует, что направление вектора $\text{rot } \mathbf{A}$ в пространстве определяется как направление нормали к такой площадке, для которой поверхностная плотность порождения циркуляции максимальна. Чтобы

подсчитать указанную плотность порождения циркуляции, вычислим циркуляцию вектора \mathbf{A} по бесконечно малому контуру. Сначала допустим, что этот контур лежит в плоскости $z = const$.

Кроме того, так как при подсчете плотности форма контура несущественна, то возьмем за этот контур прямоугольник со сторонами, параллельными осям координат, как показано на рис. 3, где размер этого прямоугольника преувеличен.

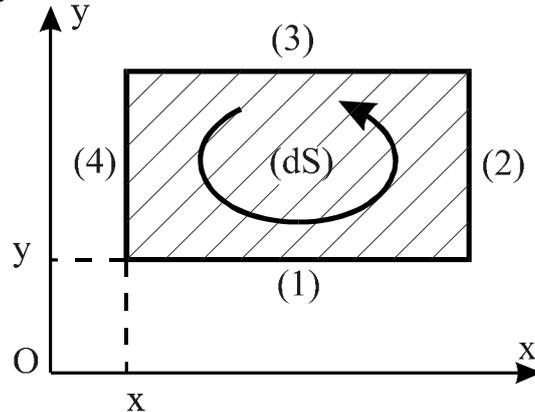


Рис. 3. Определение плотности порождения циркуляции

По формуле (6) соответствующая циркуляция равна

$$d\Gamma = \int_{(1)} A_x dx + \int_{(2)} A_y dy + \int_{(3)} A_x dx + \int_{(4)} A_y dy \quad (8)$$

(цифрами обозначены последовательные стороны прямоугольника, как показано на рис. 3), так как на каждой стороне изменяется только одна переменная, а остальные дифференциалы равны нулю. Учитывая направление обхода указанных сторон, из (8) получаем

$$\begin{aligned} d\Gamma &= (A_x)_1 dx + (A_y)_2 dy - (A_x)_3 dx - (A_y)_4 dy = \\ &= [(A_y)_2 - (A_y)_4] dy - [(A_x)_3 - (A_x)_1] dx, \end{aligned} \quad (9)$$

где цифровой индекс указывает, на какой стороне берется соответствующая проекция. Однако, с точностью до малых высшего порядка

$$(A_y)_2 - (A_y)_4 = \frac{\partial A_y}{\partial x} dx, \quad (A_x)_3 - (A_x)_1 = \frac{\partial A_x}{\partial y} dy,$$

а потому формула (9) дает

$$d\Gamma = \frac{\partial A_y}{\partial x} dx dy - \frac{\partial A_x}{\partial y} dy dx = \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) dS.$$

Итак, для бесконечно малого замкнутого контура

$$\frac{d\Gamma_{xy}}{dS_{xy}} = \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \quad (10)$$

(индексы в левой части указывают на то, что контур параллелен плоскости xy).

При этом имеется в виду, что контур обходится в положительном направлении; в противном случае надо поменять знак или, что то же, считать $dS_{XY} < 0$.

Декартовы координаты в пространстве полностью равноправны, а поэтому из любой формулы, содержащей эти координаты, можно получить другую верную формулу, заменив x, y, z соответственно на y, z, x или на z, x, y (такая перестановка называется «циклической» или «круговой»; при ней правая система координат остается правой). Поэтому из (10) вытекает

$$\frac{d\Gamma_{YZ}}{dS_{YZ}} = \frac{\partial A_Z}{\partial y} - \frac{\partial A_Y}{\partial z}; \quad \frac{d\Gamma_{ZX}}{dS_{ZX}} = \frac{\partial A_X}{\partial z} - \frac{\partial A_Z}{\partial x}. \quad (11)$$

Рассмотрим теперь бесконечно малую ориентированную площадку (dS), произвольно наклоненную относительно оси координат. Поверхность, замкнутая или незамкнутая, называется ориентированной, если указано, какая ее сторона считается наружной, а какая – внутренней. Для подсчета циркуляции эту площадку удобнее всего взять в форме треугольника, как показано на рис. 4.

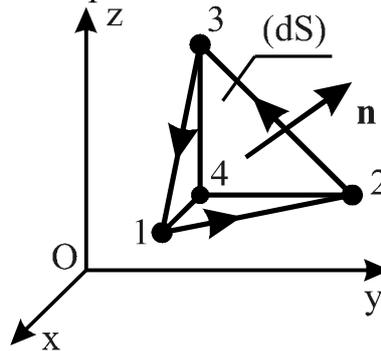


Рис. 4. Определение ротора поля

Построим на этом треугольнике тетраэдр с гранями, параллельными координатным плоскостям, и обозначим вершины этого тетраэдра цифрами. Легко убедиться, что

$$d\Gamma = d\Gamma_{123} = d\Gamma_{124} + d\Gamma_{234} + d\Gamma_{431},$$

так как в правой части интегралы по отрезкам 41, 42 и 43 взаимно уничтожаются. Но правую часть можно подсчитать по формулам (10) и (11):

$$d\Gamma = \left(\frac{\partial A_Y}{\partial x} - \frac{\partial A_X}{\partial y} \right) dS_{124} + \left(\frac{\partial A_Z}{\partial y} - \frac{\partial A_Y}{\partial z} \right) dS_{234} + \left(\frac{\partial A_X}{\partial z} - \frac{\partial A_Z}{\partial x} \right) dS_{431}, \quad (12)$$

где цифровые индексы показывают, о каких площадях идет речь.

Пусть $\mathbf{n}_{123} = \bar{n}$ – орт нормали к поверхности $dS_{123} = dS$;

$$\begin{aligned}\mathbf{n}_{234} &= \bar{i} - \text{орт нормали к грани } dS_{234}; \\ \mathbf{n}_{431} &= \bar{j} - \text{орт нормали к грани } dS_{431}, \\ \mathbf{n}_{124} &= \bar{k} - \text{орт нормали к грани } dS_{124}.\end{aligned}$$

Из рис. 4 следует, что сумма проекций граней на плоскость, перпендикулярную нормали к поверхности (dS), равна поверхности (dS), т.е.

$$dS = \mathbf{n} \mathbf{n}_{123} dS_{123} = \mathbf{n} (\bar{i} dS_{234} + \bar{j} dS_{431} + \bar{k} dS_{124}).$$

Отсюда

$$dS = \mathbf{n} dS = dS_{234} + dS_{431} + dS_{124}.$$

Удобно ввести вектор, называемый ротором (вихрем или вихревым вектором) поля \mathbf{A} и обозначаемый $rot \mathbf{A}$, определяемый формулой

$$rot \mathbf{A} = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \bar{i} + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \bar{j} + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \bar{k}. \quad (13)$$

При этом, в соответствии с (12),

$$d\Gamma = (rot \mathbf{A}) \mathbf{n} \cdot dS.$$

Тогда

$$\frac{d\Gamma}{dS} = (rot \mathbf{A}) \cdot \mathbf{n} = rot_n \mathbf{A}. \quad (14)$$

В этом выражении индекс n указывает на то, что берется проекция ротора на нормаль \mathbf{n} . Эта формула дает циркуляцию по бесконечно малому контуру, отнесенную к единице площади, охватываемой этим контуром. Отсюда следует, что направление вектора $rot \mathbf{A}$ в пространстве определяется как направление нормали к такой площадке, для которой $\frac{d\Gamma}{dS}$ максимально. Это определение сходно с определением направления градиента скаляра f как направления линии (l), в котором достигает максимума $\frac{df}{dl}$.

Из выражения (7) следует

$$d\Gamma = \mathbf{n} rot \mathbf{A} dS = rot \mathbf{A} \cdot d\mathbf{S}.$$

Учитывая (6), получаем

$$\oint_{(L)} \mathbf{A} \cdot d\mathbf{r} = \int_{(S)} rot \mathbf{A} \cdot d\mathbf{S}. \quad (15)$$

Эта важная формула называется *формулой Стокса*.

Пусть внешнее воздействие на какой-либо объект описывается функцией f , а результат этого воздействия – функцией F , т.е. получается, что каждая функция f по какому-то определенному закону преобразуется в новую функцию F . Такой закон преобразования функций – прообразов в функции – образы в математике называется *оператором*. Обозначим оператор перехода от функции внешнего воздействия f к функции –

«отклику» F через L , так что $F = L \cdot f$. Предположим, что действует закон линейности или, как говорят, принцип *суперпозиции*: при сложении внешних воздействий их результаты также складываются. Этот закон часто применяется, когда внешние воздействия не слишком велики. Его можно записать в форме

$$F = L(f_1 + f_2) = Lf_1 + Lf_2 = F_1 + F_2.$$

Оператор, обладающий таким свойством, называется *линейным*. Из свойства линейности следует

$$L(cf) = cLf,$$

где $c = \text{const}$.

Английский математик У. Гамильтон заметил, что рассмотренные нами три операции можно более просто записать, если ввести символ

$$\nabla = \bar{i} \frac{\partial}{\partial x} + \bar{j} \frac{\partial}{\partial y} + \bar{k} \frac{\partial}{\partial z},$$

называемой «набла» (это слово по-гречески означает «арфа», т.е. музыкальный инструмент, вид которого напоминает значок ∇). Отдельно взятый, этот символ представляет собой знак действия, т.е. «оператор», причем векторно-дифференциальный, сохраняющий при своем применении, как черты вектора, так и черты оператора дифференцирования. «Умножение» (т.е. действие) оператора Гамильтона на скаляр (точнее на скалярное поле) f и на вектор \mathbf{A} производится по следующим естественным правилам

$$\nabla f = (\bar{i} \frac{\partial}{\partial x} + \bar{j} \frac{\partial}{\partial y} + \bar{k} \frac{\partial}{\partial z})f = \bar{i} \frac{\partial f}{\partial x} + \bar{j} \frac{\partial f}{\partial y} + \bar{k} \frac{\partial f}{\partial z} = \text{grad } f,$$

$$\begin{aligned} \nabla \mathbf{A} &= (\bar{i} \frac{\partial}{\partial x} + \bar{j} \frac{\partial}{\partial y} + \bar{k} \frac{\partial}{\partial z})(\bar{i} A_x + \bar{j} A_y + \bar{k} A_z) = \\ &= \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} = \text{div } \mathbf{A}, \end{aligned}$$

$$\begin{aligned} \nabla \times \mathbf{A} &= \begin{vmatrix} \bar{i} & \bar{j} & \bar{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix} = \bar{i} \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) + \bar{j} \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) + \\ &+ \bar{k} \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) = \text{rot } \mathbf{A}. \end{aligned}$$

После применения оператора набла к полю получается новое поле, к которому можно вновь применить этот оператор. Так, например,

$$\text{grad div } \mathbf{A} = \nabla \cdot (\nabla \mathbf{A}) = (\bar{i} \frac{\partial}{\partial x} + \bar{j} \frac{\partial}{\partial y} + \bar{k} \frac{\partial}{\partial z}) (\bar{i} \frac{\partial A}{\partial x} + \bar{j} \frac{\partial A}{\partial y} + \bar{k} \frac{\partial A}{\partial z}) =$$

$$= (\bar{i} \frac{\partial}{\partial x} + \bar{j} \frac{\partial}{\partial y} + \bar{k} \frac{\partial}{\partial z}) (\frac{\partial A_X}{\partial x} + \frac{\partial A_Y}{\partial y} + \frac{\partial A_Z}{\partial z}) = \bar{i} (\frac{\partial^2 A_X}{\partial x^2} + \frac{\partial^2 A_Y}{\partial x \partial y} + \frac{\partial^2 A_Z}{\partial x \partial z}) +$$

$$+ \bar{j} (\frac{\partial^2 A_X}{\partial x \partial y} + \frac{\partial^2 A_Y}{\partial y^2} + \frac{\partial^2 A_Z}{\partial y \partial z}) + \bar{k} (\frac{\partial^2 A_X}{\partial x \partial z} + \frac{\partial^2 A_Y}{\partial y \partial z} + \frac{\partial^2 A_Z}{\partial z^2}); \quad (16)$$

$$\operatorname{div} \operatorname{grad} f = \nabla \cdot \nabla f = (\bar{i} \frac{\partial}{\partial x} + \bar{j} \frac{\partial}{\partial y} + \bar{k} \frac{\partial}{\partial z}) (\bar{i} \frac{\partial f}{\partial x} + \bar{j} \frac{\partial f}{\partial y} + \bar{k} \frac{\partial f}{\partial z}) =$$

$$= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = \Delta f, \quad (17)$$

где Δ – оператор Лапласа, равный

$$\Delta = \nabla \cdot \nabla = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

$$\operatorname{rot} \operatorname{grad} f = \nabla \times (\nabla f) = \begin{vmatrix} \bar{i} & \bar{j} & \bar{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{vmatrix} =$$

$$= \bar{i} (\frac{\partial^2 f}{\partial y \partial z} - \frac{\partial^2 f}{\partial y \partial z}) - \bar{j} (\frac{\partial^2 f}{\partial x \partial z} - \frac{\partial^2 f}{\partial x \partial z}) + \bar{k} (\frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial x \partial y}) = 0 \quad (18)$$

$$\operatorname{div} \operatorname{rot} \mathbf{A} = \frac{\partial}{\partial x} (\frac{\partial A_Z}{\partial y} - \frac{\partial A_Y}{\partial z}) + \frac{\partial}{\partial y} (\frac{\partial A_X}{\partial z} - \frac{\partial A_Z}{\partial x}) +$$

$$+ \frac{\partial}{\partial z} (\frac{\partial A_Y}{\partial x} - \frac{\partial A_X}{\partial y}) = 0, \quad (19)$$

$$\operatorname{rot} \operatorname{rot} \mathbf{A} = \nabla \times (\nabla \times \mathbf{A}) = \nabla (\nabla \cdot \mathbf{A}) - (\nabla \cdot \nabla) \mathbf{A} = \operatorname{grad} \operatorname{div} \mathbf{A} - \Delta \mathbf{A} =$$

$$= \bar{i} (\frac{\partial^2 A_X}{\partial x^2} + \frac{\partial^2 A_Y}{\partial x \partial y} + \frac{\partial^2 A_Z}{\partial x \partial z}) + \bar{j} (\frac{\partial^2 A_X}{\partial x \partial y} + \frac{\partial^2 A_Y}{\partial y^2} + \frac{\partial^2 A_Z}{\partial y \partial z}) +$$

$$+ \bar{k} (\frac{\partial^2 A_X}{\partial x \partial z} + \frac{\partial^2 A_Y}{\partial y \partial z} + \frac{\partial^2 A_Z}{\partial z^2}) - \frac{\partial^2 \bar{A}}{\partial x^2} - \frac{\partial^2 \bar{A}}{\partial y^2} - \frac{\partial^2 \bar{A}}{\partial z^2}. \quad (20)$$

Справедливы следующие три формулы, заменяющие тройной интеграл двойным:

- 1) формула для градиента
$$\int_V \operatorname{grad} f \cdot dV = \int_S f d\mathbf{S}; \quad (21)$$

- 2) формула для дивергенции (теорема Остроградского-Гаусса)
$$\int_V \operatorname{div} \mathbf{A} dV = \int_S \mathbf{A} d\mathbf{S}; \quad (22)$$

3) формула для вихря

$$\int_V \text{rot} \mathbf{A} dV = \int_S d\mathbf{S} \times \mathbf{A}. \quad (23)$$

В соответствии с определением дивергенции находим

$$\begin{aligned} \text{div}(f \mathbf{A}) &= \frac{\partial(fA_x)}{\partial x} + \frac{\partial(fA_y)}{\partial y} + \frac{\partial(fA_z)}{\partial z} = \\ &= f \left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right) + A_x \frac{\partial f}{\partial x} + A_y \frac{\partial f}{\partial y} + A_z \frac{\partial f}{\partial z} = \\ &= f \cdot \text{div} \mathbf{A} + \mathbf{A} \cdot \text{grad} f. \end{aligned} \quad (24)$$

Пусть S – замкнутая поверхность, ограничивающая объем V ; p и q – две скалярные функции точки M . Заменив в (24) f на p , а \mathbf{A} на $\text{grad} q$, получаем

$$\text{div}(p \text{grad} q) = p \Delta q + \text{grad} q \cdot \text{grad} p. \quad (25)$$

Подставив в формулу (22) вместо вектора \mathbf{A} вектор $p \text{grad} q$, с учетом (25) получаем

$$\int_V (p \Delta q + \text{grad} q \cdot \text{grad} p) dV = \int_S p \text{grad} q \cdot d\mathbf{S}.$$

Меняя роли функций p и q , имеем

$$\int_V (q \Delta p + \text{grad} p \cdot \text{grad} q) dV = \int_S q \text{grad} p \cdot d\mathbf{S}.$$

Вычитая последнее выражение из предыдущего, получаем формулу Грина

$$\int_V (p \Delta q - q \Delta p) dV = \int_S (p \text{grad} q - q \text{grad} p) \cdot d\mathbf{S}. \quad (26)$$

Б. Разложение в ряд произвольных функций

В оптике нередко возникает необходимость решения задачи разложения в ряд произвольно взятой функции по заданной системе функций. Подобные задачи возникают, например, при изучении основ геометрической оптики, при изучении дифракционно-абберационной структуры изображения, образованного оптической системой, при проектировании оптических систем, при обработке поверхностей оптических деталей и оценке их формы и т.д.

Напомним, что световое возмущение в изображении точки, образованном оптической системой, определяется преобразованием Фурье зрачковой функции, в которую входит функция волновой абберации. Для абберационной функции возможны различные представления. В оптике достаточно широко применяют разложение аббераций в степенной ряд. Если использовать для точек выходного зрачка полярные координаты ρ и

φ , то для осесимметричной оптической системы разложение функции волновой aberrации в степенной ряд можно представить в виде:

$$W(\rho, \varphi) = \sum_i \sum_j W_{ij} \rho^i \cos^j \varphi = W_{00} + W_{20} \rho^2 + W_{11} \rho \cos \varphi + \\ + W_{40} \rho^4 + W_{31} \rho^3 \cos \varphi + W_{22} \rho^2 \cos^2 \varphi + W_{60} \rho^6 + \dots$$

Здесь W_{ij} – коэффициент волновой aberrации, равный волновой aberrации на краю зрачка при $\rho = 1$, $\varphi = 0$; $i \geq j$, $i - j$ – четное число; W_{00} – постоянная составляющая, которая не влияет на качество изображения и обычно в разложение не включается.

Поскольку при круглой форме апертурной диафрагмы световое возмущение в изображении точки определяется интегрированием по единичному кругу, функцию волновой aberrации удобнее представлять в виде разложения в ряд по полной системе полиномов, ортогональных внутри единичного круга.

Разложения в ряд произвольных функций по полной ортогональной системе функций имеют замечательное геометрическое истолкование.

Относительно рассматриваемых функций будем предполагать, что они кусочно-гладкие, т.е. что они кусочно-непрерывны и имеют кусочно-непрерывные первые производные.

1. Ортогональные системы функций

Известно, что любую совокупность объектов, в которой можно производить линейные действия (сложение этих объектов и умножение их на скаляры, т.е. на числа), можно представить в виде многомерного векторного пространства. Такими объектами могут быть, в частности, некоторые функции $f(x)$, рассматриваемые на некотором интервале оси x . Последний может быть конечным или бесконечным, может совпадать со всей осью, но он должен быть одним и тем же для всех рассматриваемых функций, поскольку эти функции можно как складывать, так и умножать на скаляр. Отсюда следует, что функции $f(x)$ можно рассматривать как векторы, при этом совокупность таких функций образует соответствующее векторное пространство.

Пусть рассматриваемые функции принимают только вещественные значения. Выберем n значений независимой переменной $x = x_1, x_2, \dots, x_n$ и будем рассматривать значения любой функции f только в этих точках: $f(x_1), f(x_2), \dots, f(x_n)$. Вполне очевидно, что при этом имеется n степеней свободы, т.е. такие функции образуют n -мерное векторное пространство, в котором скалярное произведение можно ввести по формуле

$$(f, \varphi) = \sum_{k=1}^n f(x_k) \varphi(x_k). \quad (1)$$

При «сгущении» узловых значений переменной x , т.е. значений x_k , n -мерное пространство переходит в бесконечномерное, а сумма (1), определяющая скалярное произведение функций, естественным образом заменяется интегралом. Таким образом, в пространстве функций скалярное произведение можно ввести по формуле

$$(f, \varphi) = \int_a^b f(x) \varphi(x) dx, \quad (2)$$

где a и b начало, и конец интервала оси x , на котором рассматриваются функции. Согласно этой формуле и правилам векторной алгебры вводится «длина вектора f », называемая нормой функции f и обозначаемая $\|f\|$. Норма функции f вычисляется по формуле

$$\|f\|^2 = (f, f) = \int_a^b [f(x)]^2 dx. \quad (3)$$

Пространство функций со скалярным произведением, определяемым формулой (2), и с нормой, определяемой формулой (3), называется гильбертовым пространством или, полнее, вещественным гильбертовым пространством функций на заданном интервале с концами a, b . В гильбертовом пространстве естественным образом вводится понятие ортогональности. Принимается, что две функции $g_1(x)$ и $g_2(x)$ ортогональны друг другу (на интервале с концами a, b), если их скалярное произведение равно нулю, т.е. если

$$\int_a^b g_1(x) g_2(x) dx = 0.$$

Две ортогональные функции, очевидно, подобны двум перпендикулярным друг другу векторам.

Если имеется ортогональная система функций, т.е. совокупность попарно ортогональных функций

$$g_1(x), g_2(x), \dots, g_n(x), \dots, \quad (4)$$

то часто возникает задача о разложении любой заданной функции $f(x)$ по этим функциям, т.е. о разложении $f(x)$ в ряд вида:

$$f(x) = \sum_{k=1}^{\infty} a_k g_k(x). \quad (5)$$

В обычном (трехмерном) пространстве любой вектор можно разложить по трем ортогональным, т.е. перпендикулярным друг другу, векторам. Такая тройка векторов называется полной, при этом, как

говорят, ее можно принять за базис в пространстве. Если же ортогональных векторов два, то по ним можно разложить только такие векторы, которые лежат в плоскости этих двух векторов. Такая пара векторов в трехмерном пространстве не является полной. Она становится полной только после присоединения к ней третьего ортогонального вектора. В конечномерном векторном пространстве система ортогональных векторов полная, если число векторов в системе равно размерности пространства. Подобно этому ортогональная система функций (4) называется полной, если по ней можно разложить любую функцию $f(x)$ в ряд вида (5). Любая такая система функций образует базис в пространстве Гильберта.

Обычно установить полноту системы ортогональных функций (4) бывает далеко не просто. Если все же полнота системы функций (4) так или иначе, установлена, то найти коэффициенты разложения заданной функции $f(x)$ в ряд (5) очень легко. Для этого обе части равенства

$$f(x) = a_1 g_1(x) + a_2 g_2(x) + a_3 g_3(x) + \dots \quad (6)$$

надо скалярно умножить на одну из функций $g_k(x)$. При этом в правой части равенства (6) в соответствии с соотношением ортогональности все члены обратятся в нуль, кроме одного, в котором функция умножается сама на себя. В результате получаем равенство $(f, g_k) = a_k (g_k, g_k)$, отсюда

$$a_k = \frac{(f, g_k)}{(g_k, g_k)} = \frac{\int_a^b f(x) g_k(x) dx}{\int_a^b [g_k(x)]^2 dx}. \quad (7)$$

Если левую и правую части равенства (6) скалярно умножить самих на себя, то получается интересное обобщение теоремы Пифагора на пространство Гильберта. При этом в правой части все попарные скалярные произведения в соответствии с соотношением ортогональности оказываются равными нулю, а остаются только скалярные квадраты всех слагаемых. В результате имеем

$$(f, f) = a_1^2 (g_1, g_1) + a_2^2 (g_2, g_2) + \dots,$$

т.е.

$$\|f\|^2 = a_1^2 \|g_1\|^2 + a_2^2 \|g_2\|^2 + a_3^2 \|g_3\|^2 + \dots$$

Таким образом, получаем, что квадрат длины вектора f равен сумме квадратов его проекций на базисные векторы g_1, g_2, \dots . Полученное выражение называется равенством Парсеваля.

Функцию, норма которой равна единице, называют нормированной функцией. Систему нормированных функций $\varphi_1(x), \varphi_2(x), \dots$, в которой каждые две различные функции взаимно ортогональны, называют ортогональной нормированной системой функций, а характеризующие ее соотношения

$$(\varphi_\nu, \varphi_\mu) = \exp(\nu\mu) \quad (e_{\nu\nu} = 1, e_{\nu\mu} = 0 \text{ при } \mu \neq \nu)$$

называют соотношениями ортогональности.

Пример ортогональной нормированной системы функций в интервале $0 \leq x \leq 2\pi$ или вообще в любом интервале длиной 2π представляют функцию:

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \dots, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \dots$$

Если рассматривается комплексное гильбертово пространство, т.е. если рассматриваемые функции от вещественного аргумента могут принимать комплексные значения, то формула для скалярного произведения будет иметь вид:

$$(f, g) = \int_a^b f(x)g^*(x)dx,$$

где звездочкой обозначена комплексно сопряженная величина. Соответственно изменится и формула, определяющая квадрат нормы функции f :

$$\|f\|^2 = (f, f) = \int_a^b f(x)f^*(x)dx = \int_a^b |f(x)|^2 dx.$$

Две комплексозначные функции $f(x)$ и $g(x)$ называются ортогональными, если соблюдаются соотношения:

$$(f, g^*) = (f^*, g) = 0.$$

Функция $f(x)$ называется нормированной, если

$$\|f\| = \int_a^b f(x)f^*(x)dx = \int_a^b |f(x)|^2 dx = 1.$$

Простейший пример комплексной ортогональной нормированной системы представляют в интервале $-\pi \leq x \leq \pi$ показательные функции:

$$\frac{1}{\sqrt{2\pi}}, \frac{\exp(ix)}{\sqrt{2\pi}}, \frac{\exp(2ix)}{\sqrt{2\pi}}, \dots,$$

что непосредственно следует из соотношений ортогональности:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \exp[i(\mu - \nu)x]dx = e_{\mu\nu} \quad (e_{\nu\nu} = 1, e_{\nu\mu} = 0 \text{ при } \mu \neq \nu).$$

Функции f_1, \dots, f_r называются линейно зависимыми, если они удовлетворяют тождественно относительно x однородному линейному соотношению

$$\sum_{i=1}^r c_i f_i = 0$$

с постоянными коэффициентами c_i ($i = 1, 2, \dots, r$), которые не все равны нулю. В противном случае эти n функции называются линейно независимыми. Важно заметить, что функции ортогональной системы всегда линейно независимы. Действительно, из соотношения

$$c_1 f_1 + c_2 f_2 + \dots + c_n f_n = 0$$

следовало бы, если его умножить на f_v и проинтегрировать, что $c_v = 0$.

2. Ортогонализация функций

Пусть задана бесконечная система функций q_1, q_2, \dots , обладающая тем свойством, что при любом r каждые r произвольно выбранных функций линейно независимы. Из этой системы функций можно при помощи процесса ортогонализации получить ортогональную нормированную систему функций $\varphi_1, \varphi_2, \dots$, выбирая φ_n как соответствующую линейную комбинацию функций q_1, \dots, q_n .

Сначала полагаем $\varphi_1 = \frac{q_1}{\|q_1\|}$. Затем выбираем такое число α , чтобы

функция $\varphi'_2 = \alpha\varphi_1 + q_2$ была ортогональной к φ_1 , т.е. чтобы соблюдалось равенство:

$$\alpha + (\varphi_1, q_2) = 0.$$

Отсюда следует, что $\alpha = -(\varphi_1, q_2)$. Тогда

$$\varphi_2 = q_2 - \varphi_1(\varphi_1, q_2).$$

Функция φ'_2 в силу линейной независимости q_1 и q_2 , а, следовательно, и функций φ_1 и q_2 , не может тождественно равняться нулю. Таким образом,

$$\varphi_2 = \frac{\varphi'_2}{\|\varphi'_2\|}.$$

Теперь образуем функцию $\varphi'_3 = \beta\varphi_1 + \gamma\varphi_2 + q_3$, выбирая два одновременно не равных нулю числа β и γ , удовлетворяющих двум линейным однородным уравнениям:

$$(\varphi'_3, \varphi_1) = \beta + (\varphi_1, q_3) = 0,$$

$$(\varphi'_3, \varphi_2) = \gamma + (\varphi_2, q_3) = 0.$$

Из этих равенств, следует, что $\beta = -(\varphi_1, q_3)$, а $\gamma = -(\varphi_2, q_3)$. При этом $\varphi'_3 = q_3 - \varphi_1(\varphi_1, q_3) - \varphi_2(\varphi_2, q_3)$.

В силу линейной независимости функций q_1, q_2 и q_3 , а вместе с тем и φ_1, φ_2 и q_3 , функция φ'_3 не может тождественно равняться нулю, а поэтому соотношение

$$\varphi_3 = \frac{\varphi'_3}{\|\varphi'_3\|}$$

определяет нормированную функцию, ортогональную к φ_1 и φ_2 .

Продолжая неограниченно этот процесс, получим искомую ортогональную нормированную систему функций с помощью рекуррентной формулы:

$$\varphi_{n+1} = \frac{\varphi'_{n+1}}{\|\varphi'_{n+1}\|}, \quad \varphi'_{n+1} = q_{n+1} - \sum_{k=1}^n \varphi_k(\varphi_k, q_{n+1}). \quad (8)$$

Заметим, что рассмотренная процедура, известна под названием процедуры ортогонализации Грама-Шмидта, при этом одновременно с ортогонализацией осуществляет и нормирование функций.

3. Круговые полиномы Цернике

Существует бесчисленное множество полных систем полиномов от двух вещественных переменных x и y , ортогональных внутри единичного круга, т.е. удовлетворяющих условию

$$\iint_{x^2+y^2 \leq 1} V_{(\alpha)}^*(x, y) V_{(\beta)}(x, y) dx dy = A_{\alpha\beta} \delta_{\alpha\beta}. \quad (9)$$

Здесь $V_{(\alpha)}$ и $V_{(\beta)}$ – два произвольных полинома системы (звездочка обозначает комплексное сопряжение); $A_{\alpha\beta}$ – нормировочная постоянная, которая будет определена позднее; $\delta_{\alpha\beta}$ – символ Кронекера, определяемый условиями

$$\delta_{\alpha\beta} = \begin{cases} 1, & \text{если } \alpha = \beta, \\ 0, & \text{если } \alpha \neq \beta. \end{cases}$$

Однако, в оптике наиболее широкое применение нашла система круговых полиномов. Впервые круговые полиномы были введены и исследованы Ф. Цернике в его важной работе, посвященной исследованию методов темного поля и фазового контраста. Круговые полиномы Цернике отличаются от полиномов других систем некоторыми простыми свойствами инвариантности, которые проще всего объяснить в рамках теории групп. Однако, с помощью своего рода нормировки можно избежать введения абстрактного формализма теории групп.

Рассмотрим сначала такие системы полиномов, которые «инвариантны по форме» относительно поворота координатных осей вокруг начала координат. Такая инвариантность означает, что при любом повороте

$$\left. \begin{aligned} x' &= x \cos \varphi + y \sin \varphi \\ y' &= -x \sin \varphi + y \cos \varphi \end{aligned} \right\} \quad (10)$$

каждый полином $V(x, y)$ переходит в полином такого же вида, т.е. при использовании преобразования (10) полином V удовлетворяет следующему соотношению:

$$V(x, y) = \Psi(\varphi)V(x', y'), \quad (11)$$

где $\Psi(\varphi)$ – непрерывная периодическая функция угла поворота φ с периодом 2π при $\Psi(0) = 1$. Осуществление двух последовательных поворотов на углы φ_1 и φ_2 эквивалентно одному повороту на угол $\varphi_1 + \varphi_2$. При этом, из соотношения (11) следует, что величина Ψ должна удовлетворять функциональному уравнению

$$\Psi(\varphi_1)\Psi(\varphi_2) = \Psi(\varphi_1 + \varphi_2). \quad (12)$$

Общее решение этого уравнения, имеющее период 2π , хорошо известно:

$$\Psi(\varphi) = \exp(il\varphi). \quad (13)$$

Здесь l – любое целое число, положительное, отрицательное или нуль. Положив $x = \rho \cos \varphi$, $y = \rho \sin \varphi$, $x' = \rho$, а $y' = 0$ и подставив при этом выражение (13) в соотношение (11), получаем

$$V(\rho \cos \varphi, \rho \sin \varphi) = R(\rho)\exp(il\varphi), \quad (14)$$

где функция $R(\rho) = V(\rho, 0)$ зависит только от ρ .

Разложим теперь $\exp(il\varphi)$ в ряд по степеням $\cos \varphi$ и $\sin \varphi$. Предположим, что V – полином степени n от переменных $x = \rho \cos \varphi$, $y = \rho \sin \varphi$. Тогда из выражения (14) следует, что функция $R(\rho)$ представляет собой полином по ρ степени n и не содержит степеней ρ , меньших $|l|$. Это означает, что радиальные функции $R(\rho)$ являются полиномами по ρ , содержащими степени $\rho: \rho^n, \rho^{n-2}, \dots, \rho^{|l|}$. Более того, полином $R(\rho)$ является, очевидно, четным или нечетным полиномом в зависимости от четности числа l . Система круговых полиномов Цернике отличается от всех других подобных систем тем, что она содержит полином для каждой пары возможных значений n (степень) и l (угловая зависимость), т.е. для всех целых значений n и l , для которых $n \geq 0$, $l \geq 0$, $n \geq |l|$, а $n - |l|$ – четное число. Обозначим произвольный полином этой системы следующим образом:

$$V_n^l(\rho \cos \varphi, \rho \sin \varphi) = R_n^l(\rho) \exp(il\varphi). \quad (15)$$

Из выражений (9) и (15) следует, что радиальные полиномы $R_n^l(\rho)$ удовлетворяют соотношению:

$$\int_0^1 R_n^l(\rho) R_{n'}^l(\rho) \rho d\rho = a_n^l \delta_{nn'}, \quad (16)$$

где $a_n^l = \frac{A_n^l}{2\pi}$.

Для любого заданного значения l нижний индекс n может принимать только значения $|l|, |l|+2, |l|+4, \dots$. Соответствующая последовательность $R_{|l|}^l(\rho), R_{|l|+2}^l(\rho), R_{|l|+4}^l(\rho), \dots$ получается путем ортогонализации степеней ρ :

$$\rho^{|l|}, \rho^{|l|+2}, \rho^{|l|+4}, \dots \quad (17)$$

с весовой функцией ρ на интервале $0 \leq \rho \leq 1$. Поскольку в степени ρ (17) входит только модуль l , то

$$R_n^{-l}(\rho) = \beta_n^l R_n^l(\rho), \quad (18)$$

где β_n^l – постоянная, зависящая только от нормировки полиномов R_n^{-l} и R_n^l . В частности, она может быть выбрана так, чтобы $\beta_n^l = 1$ для всех значений l и n . При этом

$$V_n^{\pm m}(\rho \cos \varphi, \rho \sin \varphi) = R_n^m(\rho) \exp(\pm im\varphi), \quad (19)$$

где $m = |l|$ – неотрицательное целое число.

Система круговых полиномов содержит $\frac{1}{2}(n+1)(n+2)$ линейно независимых полиномов степени $\leq n$. Следовательно, любой одночлен $x^i y^j$ ($i \geq 0, j \geq 0$ – целые числа) и любой полином по x и y можно представить в виде линейной комбинации конечного числа круговых полиномов V_n^l . Тогда в соответствии с теоремой Вейерштрасса о приближении функций: любую функцию, непрерывную в интервале $a \leq x \leq b$, можно равномерно аппроксимировать в этом интервале с помощью полиномов, рассматриваемая система круговых полиномов будет полной.

Поскольку $R_n^{\pm m}(\rho)$ есть полином по ρ степени, n и он не содержит степени ρ , меньшей m , и является четным или нечетным в зависимости от того, четно или нечетно n , то его можно представить в виде:

$$R_n^{\pm m}(\rho) = t^{\frac{m}{2}} Q_{\frac{n-m}{2}}(t), \quad (20)$$

где $t = \rho^2$ и $Q_{\frac{n-m}{2}}(t)$ – полином по t степени $\frac{n-m}{2}$.

Согласно выражению (16) полиномы Q должны удовлетворять соотношению:

$$\frac{1}{2} \int_0^1 t^m Q_k(t) Q_{k'}(t) dt = a_n^{\pm m} \delta_{kk'}, \quad (21)$$

где $k = \frac{n-m}{2}$, $k' = \frac{n'-m}{2}$. Отсюда следует, что полиномы $Q_0(t), Q_1(t), \dots, Q_k(t), \dots$ можно получить ортогонализацией последовательности натуральных степеней

$$1, t, t^2, \dots, t^k, \dots \quad (22)$$

с весовой функцией $\omega(t) = t^m$ в области $0 \leq t \leq 1$.

Обратимся к выражению, определяющему полиномы Якоби (или гипергеометрические функции), которое в рассматриваемом случае можно представить в виде:

$$G_k(p, q, t) = \frac{(q-1)!}{(q+k-1)!} t^{1-q} (1-t)^{q-p} \frac{d^k}{dt^k} \left[t^{q-1+k} (1-t)^{p-q+k} \right] \quad (23)$$

или

$$G_k(p, q, t) = \frac{k!(q-1)!}{(p+k-1)!} \sum_{s=0}^k (-1)^s \frac{(p+k+s-1)!}{(k-s)!s!(q+s-1)!} t^s. \quad (24)$$

Здесь $k \geq 0, q > 0, p - q > -1$. Эти полиномы определяются как функции, получающиеся при ортогонализации степеней t (22) с весовой функцией более общего вида

$$\omega(t) = t^{q-1} (1-t)^{p-q}$$

в области $0 \leq t \leq 1$. Свойства ортонормируемости этих полиномов определим выражением

$$\int_0^1 t^{q-1} (1-t)^{p-q} G_k(p, q, t) G_{k'}(p, q, t) dt = b_k(p, q) \delta_{kk'}, \quad (25)$$

$$\text{где } b_k(p, q) = \frac{k![(q-1)!]^2 (p-q+k)!}{(q-1+k)!(p-1+k)!(p+2k)}. \quad (26)$$

При таком выборе b_k для всех k имеем $G_k(p, q, 0) = 1$. Сравнивая выражения (25) и (21), получаем

$$Q_k(t) = \sqrt{\frac{2a_n^{\pm m}}{b_k(m+1, m+1)}} G_k(m+1, m+1, t). \quad (27)$$

Подставив выражение (27) в выражение (20), получаем следующее соотношение между радиальными полиномами и полиномами Якоби:

$$R_n^{\pm m}(\rho) = \sqrt{\frac{2a_n^{\pm m}}{b_k(m+1, m+1)}} \rho^m G_k(m+1, m+1, \rho^2), \quad (28)$$

где $k = \frac{n-m}{2}$.

Следуя Цернике, выберем нормировку таким образом, чтобы при всех n и m осталось справедливое соотношение

$$R_n^{\pm m}(1) = 1. \quad (29)$$

Положив в выражении (28) $\rho = 1$ и учитывая соотношение (29), получаем

$$\sqrt{\frac{b_k(m+1, m+1)}{2a_n^{\pm m}}} = G_k(m+1, m+1, 1). \quad (30)$$

Значение $G_k(m+1, m+1, 1)$ можно получить с помощью производящей функции полиномов Якоби. Положив $\frac{1-x}{2} = \rho^2$, а $t = z$, получаем

$$\begin{aligned} & \frac{\left[z-1 + \sqrt{1-2z(1-2\rho^2)} + z^2 \right]^m}{(2z\rho^2)^m \sqrt{1-2z(1-2\rho^2)} + z^2} = \\ & = \sum_{s=0}^{\infty} \binom{m+s}{s} G_s(m+1, m+1, \rho^2) z^s. \end{aligned} \quad (31)$$

При $\rho = 1$ левая часть последнего соотношения равна $(1+z)^{-1}$. Разлагая при этом левую часть выражения (31) в степенной ряд и сравнивая полученный результат с правой частью, получаем

$$G_s(m+1, m+1, 1) = \frac{(-1)^s}{\binom{m+s}{s}}. \quad (32)$$

Из выражений (3.45) и (3.43) следует, что

$$\sqrt{\frac{2a_n^{\pm m}}{b_k(m+1, m+1)}} = (-1)^{\frac{n-m}{2}} \begin{pmatrix} \frac{1}{2}(n+m) \\ \frac{1}{2}(n-m) \end{pmatrix}. \quad (33)$$

Используя соотношения (23), (24) и (33), преобразуем выражение (28) к виду, определяющему окончательные выражения для радиальных полиномов:

$$R_n^{\pm m}(\rho) = \frac{1}{\left(\frac{n-m}{2}\right)! \rho^m} \left\{ \frac{d}{d(\rho^2)} \right\}^{\frac{n-m}{2}} \left\{ (\rho^2)^{\frac{m+n}{2}} (\rho^2 - 1)^{\frac{n-m}{2}} \right\} \quad (34)$$

или

$$R_n^{\pm m}(\rho) = \sum_{s=0}^{\frac{1}{2}(n-m)} (-1)^s \frac{(n-s)!}{s! \left(\frac{n+m}{2} - s\right)! \left(\frac{n-m}{2} - s\right)!} \rho^{n-2s}. \quad (35)$$

В табл. 1 приведены в явном виде выражения для ряда первых полиномов Цернике.

Соотношения (33) и (26) позволяют найти выражение, определяющее нормировочную постоянную $a_n^{\pm m}$, в виде:

$$a_n^{\pm m} = \frac{1}{2n+2}. \quad (36)$$

Для нахождения производящей функции радиальных полиномов в выражения (28) и (33) напомним s вместо $k = \frac{n-m}{2}$ и $m+2s$ вместо n и подставим выражения в выражение (31). В результате получаем

$$\frac{\left[z - 1 + \sqrt{1 - 2z(1 - 2\rho^2) + z^2} \right]^m}{(2z\rho)^m \sqrt{1 - 2z(1 - 2\rho^2) + z^2}} = \sum_{s=0}^{\infty} z^s R_{m+2s}^{\pm m}(\rho). \quad (37)$$

Вместо комплексных полиномов V можно использовать вещественные полиномы

$$U_n^m = \frac{1}{2} (V_n^m + V_n^{-m}) = R_n^m(\rho) \cos m\varphi,$$

$$U_n^{-m} = \frac{1}{2i} (V_n^m - V_n^{-m}) = R_n^m(\rho) \sin m\varphi.$$

Следуя Нижберу, разложим функцию волновой aberrации W по круговым полиномам Цернике. Так как волновые aberrации симметричны относительно меридиональной плоскости ($\varphi = 0$), то функция aberrаций является четной функцией угла φ . Поэтому в разложении функции волновой aberrации по круговым полиномам Цернике применяют только полиномы вида $U_n^m = R_n^m(\rho) \cos m\varphi$.

Таблица 1. Радиальные полиномы $R_n^m(\rho)$ для $m \leq 8, n \leq 8$

n	0	1	2	3	4	5	6	7	8
0	1	$2\rho^2 - 1$			$6\rho^4 - 6\rho^2 + 1$		$20\rho^6 - 30\rho^4 + 12\rho^2 - 1$		$70\rho^8 - 140\rho^6 + 90\rho^4 - 20\rho^2 + 1$
1		ρ	$3\rho^3 - 2\rho$			$10\rho^5 - 12\rho^3 + 3\rho$		$35\rho^7 - 60\rho^5 + 30\rho^3 - 4\rho$	
2			ρ^2		$4\rho^4 - 3\rho^2$		$15\rho^6 - 20\rho^4 + 6\rho^2$		$56\rho^8 - 105\rho^6 + 60\rho^4 - 10\rho^2$
3				ρ^3		$5\rho^5 - 4\rho^3$		$21\rho^7 - 30\rho^5 + 10\rho^3$	
4					ρ^4		$6\rho^6 - 5\rho^4$		$28\rho^8 - 42\rho^6 + 15\rho^4$
5						ρ^5		$7\rho^7 - 6\rho^5$	
6							ρ^6		$8\rho^8 - 7\rho^6$
7								ρ^7	
8									ρ^8

Из условия симметрии следует, что разложение функции волновой аберрации в ряд должно содержать лишь комбинации переменных y'^2 , ρ^2 и $y'\rho \cos \varphi$, где y' – координата некоторой точки P' изображения в меридиональной плоскости. Это означает, что функцию волновой аберрации с помощью круговых полиномов Цернике можно представить в виде:

$$W(y', \rho, \varphi) = \sum_l \sum_n \sum_m a_{lnm} y'^{2l+m} R_n^m(\rho) \cos m\varphi,$$

где l, n, m – неотрицательные целые числа, при этом $n \geq m$; $n - m$ – целое число; a_{lnm} – постоянная величина. Поскольку обычно определяют изображение фиксированной точки предмета ($y' = \text{const}$), удобно не выделять явную зависимость W от y' и представлять функцию волновой аберрации в виде:

$$W(\rho, \varphi) = A_{00} + \frac{1}{\sqrt{2}} \sum_{n=2}^{\infty} A_{n0} R_n^0(\rho) + \sum_{n=1}^{\infty} \sum_{m=1}^n A_{nm} R_n^m(\rho) \cos m\varphi. \quad (38)$$

Здесь коэффициенты A_{nm} зависят от величины y' , а множитель $\frac{1}{\sqrt{2}}$ перед вторым членом введен для упрощения окончательных формул.

Некоторые свойства круговых полиномов Цернике определяют их широкое применение в оптике.

Пусть, например, требуется оценить качество изображения, образованного оптической системой. Для оценки качества изображения удобно пользоваться распределением так называемой нормированной освещенности в изображении рассматриваемой точки предмета, определяемой отношением

$$I(\xi, \eta) = \frac{E(\xi, \eta)}{E_0(0, 0)},$$

где $E(\xi, \eta)$ – распределение освещенности в изображении точки, определяемое в системе координат ξ, η , начало которой совмещено с точкой параксиального изображения; $E_0(0, 0)$ – значение освещенности в изображении точечного предмета, образованном безабберационной системой, в точке параксиального изображения.

Наибольшее значение нормированной освещенности называют определяющей яркостью или числом Штреля. Обозначая эту величину через St , имеем

$$St = \frac{E_{\max}(\xi, \eta)}{E_0(0, 0)} = I_{\max}(\xi, \eta).$$

Пусть $\bar{W}_{p'}^n$ – среднее значение n -й степени $W_{p'}$, т.е.

$$\overline{W}_{p'}^n = \frac{\int_0^1 \int_0^{2\pi} W_{p'}^n \rho d\rho d\varphi}{\int_0^1 \int_0^{2\pi} \rho d\rho d\varphi} = \frac{1}{\pi} \int_0^1 \int_0^{2\pi} W_{p'}^n \rho d\rho d\varphi. \quad (39)$$

Если aberrации достаточно малы (строго говоря, при $\left(|W_{p'}|_{\max} < \frac{\lambda}{4} \right)$,

то выражение, определяющее число Штреля, приобретает вид:

$$St \approx 1 - \left(\frac{2\pi}{\lambda} \right)^2 (\Delta W_{p'})^2, \quad (40)$$

где $(\Delta W_{p'})^2 = \overline{W}_{p'}^2 - (\overline{W}_{p'})^2$.

Из условия ортогональности полиномов Цернике следует, что

$$\int_0^1 \int_0^{2\pi} A_{00} A_{nm} R_n^m(\rho) \cos m\rho d\rho d\varphi = 0,$$

откуда для любых значений величины m имеем

$$\int_0^1 \int_0^{2\pi} A_{nm} R_n^m(\rho) \cos m\rho d\rho d\varphi = 0.$$

Подставив выражение (38) в (39) и учитывая условие ортогональности полиномов, получаем

$$\left. \begin{aligned} \overline{W} &= A_{00}, \\ \overline{W}^2 &= A_{00}^2 + \frac{1}{2} \sum_{n=1}^{\infty} \sum_{m=0}^n \frac{A_{nm}^2}{n+1}. \end{aligned} \right\} \quad (41)$$

Из первого равенства (41) следует, что коэффициент A_{00} характеризует среднее запаздывание волнового фронта относительно опорной сферы Гаусса. Второе соотношение представляет собой равенство Парсеваля для ортогональной системы функций $R_n^m(\rho) \cos m\varphi$.

Подставив соотношения (41) в выражение (40), получаем достаточно простое выражение для нормированной интенсивности в параксиальном изображении точки:

$$St \approx 1 - 2 \frac{\pi^2}{\lambda^2} \sum_{n=1}^{\infty} \sum_{m=0}^n \frac{A_{nm}^2}{n+1}.$$

Обработанная отражающая поверхность крупногабаритных зеркал имеет форму, мало отличающуюся от номинальной, однако, достаточно сложную и, как правило, ассиметричную. Поэтому отклонение формы поверхности (отклонение формы отраженного ею волнового фронта) от

номинальной моделируют с помощью системы круговых ортогональных полиномов Цернике разложением вида:

$$W(\rho, \varphi) = A_{00} + \frac{1}{\sqrt{2}} \sum_{n=2}^{\infty} A_{n0} R_n^0(\rho) + \sum_{n=1}^{\infty} \sum_{m=1}^n R_n^m(\rho) (A_{nm} \cos m\varphi + B_{nm} \sin m\varphi).$$

В. Преобразование Фурье

1. Ряды Фурье

При решении многих задач физики и математики необходимо осуществить разложение периодической функции с периодом, равным p (мм), в ряд по тригонометрическим функциям:

$$f(x) = a_0 + a_1 \sin \frac{2\pi}{p} x + b_1 \cos \frac{2\pi}{p} x + \dots + a_n \sin n \frac{2\pi}{p} x + b_n \cos n \frac{2\pi}{p} x + \dots = a_0 + \sum_{n=1}^{\infty} \left(a_n \sin 2\pi \frac{nx}{p} + b_n \cos 2\pi \frac{nx}{p} \right). \quad (1)$$

Ряд вида (1) называется рядом Фурье, а разложение функции $f(x)$ в ряд Фурье составляет задачу гармонического анализа. В приложениях зачастую ограничиваются конечным числом членов и получают при этом приближение функции тригонометрическим многочленом.

Легко убедиться, что коэффициенты такого ряда Фурье определяются формулами:

$$a_0 = \frac{1}{p} \int_{-p/2}^{p/2} f(x) dx,$$

$$a_n = \frac{2}{p} \int_{-p/2}^{p/2} f(x) \sin 2\pi \frac{nx}{p} dx \quad (n = 1, 2, \dots),$$

$$b_n = \frac{2}{p} \int_{-p/2}^{p/2} f(x) \cos 2\pi \frac{nx}{p} dx \quad (n = 1, 2, \dots).$$

Во многих случаях, когда ряд Фурье сходится в обычном смысле, т.е. поточечно, представляет интерес ответ на вопрос о том, каким образом он описывает функцию $f(x)$. На этот вопрос в значительной степени дает ответ теорема Дирихле. Пусть функция $f(x)$ удовлетворяет в интервале

$\left(-\frac{1}{2}p, \frac{1}{2}p\right)$ так называемым условиям Дирихле:

– интервал $\left(-\frac{1}{2}p, \frac{1}{2}p\right)$ можно разбить на конечное число интервалов, в которых $f(x)$ непрерывна и монотонна;

– если x_0 является точкой разрыва функции $f(x)$, то существуют $f(x_0+0)$ и $f(x_0-0)$. Тогда ряд Фурье функции $f(x)$ сходится и имеет место

$$\lim_{n \rightarrow \infty} \left[a_0 + \sum_{n=1}^n \left(a_n \sin 2\pi \frac{nx}{p} + b_n \cos 2\pi \frac{nx}{p} \right) \right] =$$

$$= \begin{cases} f(x), & \text{если } f(x) \text{ непрерывна в } x; \\ \frac{f(x+0) + f(x-0)}{2} & \text{в противном случае.} \end{cases}$$

Запишем ряд Фурье в виде:

$$f(x) = a_0 + \sum_{n=1}^{n=\infty} a_n \sin n\omega x + \sum_{n=1}^{n=\infty} b_n \cos n\omega x,$$

где $\omega = \frac{2\pi}{p}$ – круговая частота.

Представим функцию $\exp(ix)$ в виде степенного ряда

$$\exp(ix) = 1 + ix + \frac{1}{2!} i^2 x^2 + \frac{1}{3!} i^3 x^3 + \frac{1}{4!} i^4 x^4 + \frac{1}{5!} i^5 x^5 + \dots + \frac{1}{n!} i^n x^n + \dots =$$

$$= \left(1 - \frac{1}{2!} x^2 + \frac{1}{4!} x^4 - \dots \right) + i \left(x - \frac{1}{3!} x^3 + \frac{1}{5!} x^5 - \dots \right) = \cos x + i \sin x.$$

Таким образом, $\exp(ix) = \cos x + i \sin x$, т.е. получаем известную формулу Эйлера.

При этом

$$f(x) = a_0 + \sum_{n=1}^{n=\infty} \frac{1}{2i} a_n [\exp(in\omega x) - \exp(-in\omega x)] +$$

$$+ \sum_{n=1}^{n=\infty} \frac{1}{2} b_n [\exp(in\omega x) + \exp(-in\omega x)] =$$

$$= a_0 + \sum_{n=1}^{n=\infty} \left[\frac{b_n - ia_n}{2} \exp(in\omega x) + \frac{b_n + ia_n}{2} \exp(-in\omega x) \right],$$

$$\begin{aligned} \text{где } \frac{b_n - ia_n}{2} &= \frac{1}{p} \int_{-p/2}^{p/2} f(x) (\cos n\omega x - i \sin n\omega x) dx = \\ &= \frac{1}{p} \int_{-p/2}^{p/2} f(x) \exp(-in\omega x) dx; \end{aligned}$$

$$\frac{b_n + ia_n}{2} = \frac{1}{p} \int_{-p/2}^{p/2} f(x) (\cos n\omega x + i \sin n\omega x) dx = \frac{1}{p} \int_{-p/2}^{p/2} f(x) \exp(in\omega x) dx.$$

$$\text{При } n = 0: \frac{1}{p} \int_{-p/2}^{p/2} f(x) e^0 dx = \frac{1}{p} \int_{-p/2}^{p/2} f(x) dx = a_0.$$

Заменяя знак величины n в формуле $\frac{b_n + ia_n}{2}$, получаем формулу для

$$\frac{b_n - ia_n}{2}. \text{ При этом } f(x) = \sum_{n=-\infty}^{n=\infty} C_n \exp(in\omega x), \quad (2)$$

$$\text{где } C_n = \frac{1}{p} \int_{-p/2}^{p/2} f(x) \exp(-in\omega x) dx. \quad (3)$$

2. Интегралы Фурье

Пусть N – пространственная частота функций (гармоник), составляющих ряд Фурье, при этом $N = \frac{n}{p}$. Единичное изменение частоты

N определим разностью

$$\Delta N = \frac{n_{i+1}}{p} - \frac{n_i}{p} = \frac{1}{p}.$$

При этом

$$f(x) = \sum_{n=-\infty}^{n=\infty} g(N) \exp(i2\pi N x) \Delta N, \text{ где } g(N) = \int_{-p/2}^{p/2} f(x) \exp(-i2\pi N x) dx.$$

Пусть некоторая функция $f(x)$ задана на всей вещественной оси. Предположим, что нас интересуют ее значения в интервале $[x_0, x_0 + p]$. При этом функцию $f(x)$ на всей вещественной оси можно заменить последовательностью ее значений в интервале $[x_0, x_0 + p]$ с периодом,

равным p , т.е. можно заменить функцию $f(x)$ соответствующей периодической функцией. Тогда функцию $f(x)$ в интервале $[x_0, x_0 + p]$ можно представить рядом Фурье. Вполне очевидно, что в этом случае функция, представленная рядом Фурье, вне интервала $[x_0, x_0 + p]$ не будет совпадать с функцией $f(x)$.

Для целого ряда задач было бы полезно иметь выражение, подобное ряду Фурье, представляющее функцию $f(x)$, заданную в промежутке от $-\infty$ до ∞ . При этом будем считать, что $f(x)$ в любом конечном интервале подчиняется условиям Дирихле и является абсолютно интегрируемой, т.е. $\int_{-\infty}^{\infty} |f(x)| dx < \infty$.

Чтобы распространить полученное выражение на случай непериодических функций, можно устремить период p к бесконечности. Основным интервал тогда становится бесконечно большим, при этом величину $\frac{1}{p}$ можно считать бесконечно малым изменением

пространственной частоты N , т.е. $\frac{1}{p} = dN$. Тогда суммирование бесконечно большого числа колебаний всех частот от $-\infty$ до ∞ , определяемое полученным бесконечным рядом, можно представить интегралом

$$f(x) = \int_{-\infty}^{\infty} g(N) \exp(i2\pi Nx) dN, \quad (4)$$

при этом функция $g(N)$ определяется через $f(x)$ формулой

$$g(N) = \int_{-\infty}^{\infty} f(x) \exp(-i2\pi Nx) dx. \quad (5)$$

Введем обозначение: $2\pi N = 2\pi \frac{n}{p} = \omega$; при $p \rightarrow \infty$ соотношение

$2\pi \frac{1}{p} = d\omega$. Тогда выражения (4) и (5) можно представить в виде:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} g(\omega) \exp(i\omega x) d\omega \quad (6)$$

и, соответственно,

$$g(\omega) = \int_{-\infty}^{\infty} f(x) \exp(-i\omega x) dx. \quad (7)$$

Выражение (6) можно представить в виде:

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) \exp(-i\omega x) dx \right] \exp(i\omega x) d\omega =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{g}(\omega) \exp(i\omega x) d\omega. \quad (8)$$

При этом

$$\tilde{g}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) \exp(-i\omega x) dx. \quad (9)$$

Выражение (4) определяет прямое преобразование Фурье функции $f(x)$ и записывается в виде:

$$F [f(x)] = g(N) = \int_{-\infty}^{\infty} f(x) \exp(-i2\pi Nx) dx, \quad (10)$$

а выражение, определяющее обратное преобразование Фурье функции $g(N)$, записывается в виде:

$$F^{-1}[g(N)] = f(x) = \int_{-\infty}^{\infty} g(N) \exp(i2\pi Nx) dN. \quad (11)$$

Это выражение можно рассматривать как представление функции $f(x)$ в виде линейной комбинации (т.е. интеграла) элементарных функций характерного вида $\exp(i2\pi Nx)$. Вполне очевидно, что комплексное число $g(N)$ представляет собой просто весовой коэффициент, на который следует умножить элементарную функцию частоты N при синтезе искомой функции $f(x)$.

Преобразование Фурье функции двух переменных $f(x, y)$ определяется формулой

$$F [f(x, y)] = g(N_x, N_y) =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \exp[-i2\pi(N_x x + N_y y)] dx dy, \quad (12)$$

при этом обратное преобразование Фурье записывается в виде:

$$F^{-1}[g(N_x, N_y)] = f(x, y) =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(N_x, N_y) \exp[i2\pi(N_x x + N_y y)] dN_x dN_y. \quad (13)$$

Заметим, что для любой пары частот N_x, N_y уравнение $\exp[i2\pi(N_x x + N_y y)] = 1$ справедливо при $i2\pi(N_x x + N_y y) = i2\pi n$, где n –

целое число. При этом $y = -\frac{N_x}{N_y}x + \frac{n}{N_y} = x \operatorname{tg} \alpha + y_0$, где $\operatorname{tg} \alpha = -\frac{N_x}{N_y}$.

Направление, перпендикулярное к этой линии, определится очевидным уравнением вида:

$$y = x \operatorname{tg} \left(\alpha + \frac{\pi}{2} \right) + y_0 = x \operatorname{tg} \theta + y_0, \text{ где } \operatorname{tg} \theta = -\frac{1}{\operatorname{tg} \alpha} = \frac{N_y}{N_x}.$$

Отсюда $N_y = N_x \operatorname{tg} \theta$. При этом

$$N_x x + N_y y = N_x x + N_x y \operatorname{tg} \theta = \frac{N_x}{\cos \theta} (x \cos \theta + y \sin \theta).$$

Повернем систему координат xOy вокруг начала координат в положение $x'Oy'$ в соответствии с уравнением

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

При этом

$$\begin{aligned} \exp[i2\pi(N_x x + N_y y)] &= \exp\left[i2\pi \frac{N_x}{\cos \theta} (x \cos \theta + y \sin \theta)\right] = \\ &= \exp\left[i2\pi \frac{N_x}{\cos \theta} (x' \cos^2 \theta - y' \sin \theta \cos \theta + x' \sin^2 \theta + y' \sin \theta \cos \theta)\right] = \\ &= \exp(i2\pi N x'), \end{aligned}$$

где $N = \frac{N_x}{\cos \theta}$, т.е. $N_x = N \cos \theta$, а $N_y = N_x \operatorname{tg} \theta = N \sin \theta$.

Таким образом, точно так же, как одномерное, двумерное преобразование Фурье можно рассматривать как представление функции $f(x, y)$ в виде линейной комбинации элементарных функций вида $\exp[i2\pi(N_x x + N_y y)]$, при этом каждая элементарная функция в плоскости xOy «направлена» по отношению к оси x под углом θ , как показано на

рис. 1. Здесь $N_x = \frac{n}{p_x}$; $N_y = \frac{n}{p_y}$, при этом $N = \sqrt{N_x^2 + N_y^2}$; величина угла

θ равна: $\theta = \operatorname{arctg} \frac{N_y}{N_x}$. Пространственный период (т.е. расстояние между

линиями равной фазы) можно определить очевидным выражением вида:

$$\tilde{p} = \frac{p}{n} = \frac{1}{N} = \frac{1}{\sqrt{N_x^2 + N_y^2}}.$$

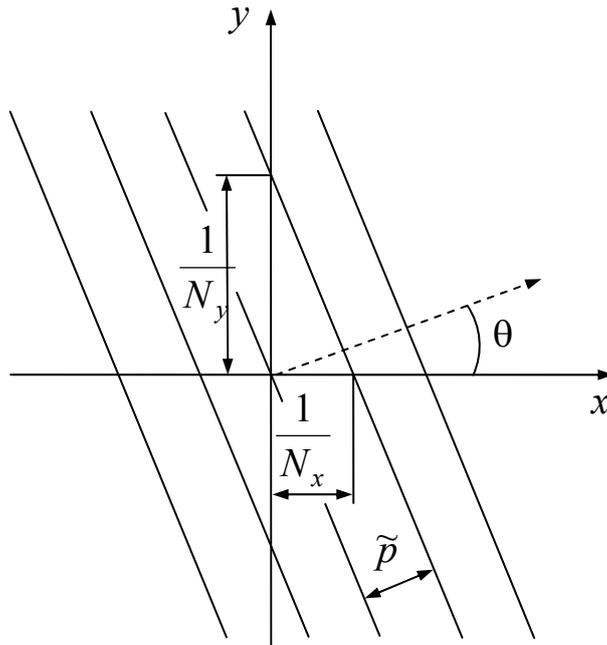


Рис. 1. Линии равной фазы

Из определения преобразования Фурье вытекает ряд математических следствий, облегчающих нахождение фурье-образа (иначе спектра Фурье или частотного спектра). Эти следствия сформулируем в виде математических теорем и приведем краткое их доказательство.

Теорема линейности

$$F [\alpha f(x, y) + \beta F(x, y)] = \alpha F [f(x, y)] + \beta F [F(x, y)].$$

Доказательство

Эта теорема непосредственно следует из свойств интеграла.

Теорема подобия

$$\text{Если } F [f(x, y)] = g(N_x, N_y), \text{ то } F [f(ax, by)] = \frac{1}{|ab|} g\left(\frac{N_x}{a}, \frac{N_y}{b}\right).$$

Доказательство

$$\begin{aligned} F [f(ax, by)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(ax, by) \exp[-i2\pi(N_x x + N_y y)] dx dy = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(ax, by) \exp\left[-i2\pi\left(\frac{N_x}{a} ax + \frac{N_y}{b} by\right)\right] d \frac{ax}{|a|} d \frac{by}{|b|} = \frac{1}{|ab|} g\left(\frac{N_x}{a}, \frac{N_y}{b}\right), \end{aligned}$$

т.е. «растяжение» координат в пространственной области (x, y) приводит к «сжатию» координат в области частот (N_x, N_y) и к изменению общей амплитуды спектра.

Теорема смещения

Если $F [f(x, y)] = g(N_x, N_y)$, то

$$F [f(x - a, y - b)] = g(N_x, N_y) \exp[-i2\pi(N_x a + N_y b)].$$

Доказательство

$$\begin{aligned} F [f(x - a, y - b)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x - a, y - b) \exp[-i2\pi(N_x x + N_y y)] dx dy = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', y') \exp\{-i2\pi[N_x(x' + a) + N_y(y' + b)]\} dx' dy' = \\ &= g(N_x, N_y) \exp[-i2\pi(N_x a + N_y b)], \end{aligned}$$

т.е. смещение функции в пространственной области вызывает линейный фазовый сдвиг в области частот.

Теорема Парсеваля

Если $F [f(x, y)] = g(N_x, N_y)$, то

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(x, y)|^2 dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |g(N_x, N_y)|^2 dN_x dN_y.$$

Доказательство

Если $F [f(x, y)] = g(N_x, N_y)$, а $F [F(x, y)] = G(N_x, N_y)$, то

$$\begin{aligned} F [f(x, y)F(x, y)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)F(x, y) \exp[-i2\pi(N_x x + N_y y)] dx dy = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \exp[-i2\pi(N_x x + N_y y)] \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(\tilde{N}_x, \tilde{N}_y) \times \right. \\ &\times \exp[i2\pi(\tilde{N}_x x + \tilde{N}_y y)] d\tilde{N}_x d\tilde{N}_y \left. \right\} dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(\tilde{N}_x, \tilde{N}_y) \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \times \right. \\ &\times \exp\langle -i2\pi[(N_x - \tilde{N}_x)x + (N_y - \tilde{N}_y)y] \rangle dx dy \left. \right\} d\tilde{N}_x d\tilde{N}_y = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} G(\tilde{N}_x, \tilde{N}_y) g(N_x - \tilde{N}_x, N_y - \tilde{N}_y) d\tilde{N}_x d\tilde{N}_y, \end{aligned}$$

т.е. преобразование Фурье произведения двух функций выражается просто через преобразования Фурье умножаемых функций g и G с помощью особой операции интегрирования, называемой сверткой функций g и G .

Свертку функций часто обозначают значком \otimes . При этом

$$F [f(x, y)F(x, y)] = g(N_x, N_y) \otimes G(N_x, N_y).$$

Преобразование Фурье функции $f(x, y)$ можно записать в виде:

$$g(N_x, N_y) = \iint_{-\infty}^{\infty} f(x, y) \cos[2\pi(N_x x + N_y y)] dx dy - \\ - i \iint_{-\infty}^{\infty} f(x, y) \sin[2\pi(N_x x + N_y y)] dx dy.$$

Отсюда следует, что

$$g(-N_x, -N_y) = \iint_{-\infty}^{\infty} f(x, y) \cos[2\pi(N_x x + N_y y)] dx dy + \\ + i \iint_{-\infty}^{\infty} f(x, y) \sin[2\pi(N_x x + N_y y)] dx dy = g^*(N_x, N_y).$$

При этом $F(x, y) = f^*(x, y)$ имеем $G(N_x, N_y) = g^*(N_x, N_y)$,

где $g^*(N_x, N_y) = \iint_{-\infty}^{\infty} f^*(x, y) \exp[i2\pi(N_x x + N_y y)] dx dy.$

Тогда

$$F[f(x, y)f^*(x, y)] = F|f(x, y)|^2 = \iint_{-\infty}^{\infty} |f(x, y)|^2 \exp[-i2\pi(N_x x + \\ + N_y y)] dx dy = \iint_{-\infty}^{\infty} f^*(x, y) \exp[-i2\pi(N_x x + N_y y)] \left\{ \iint_{-\infty}^{\infty} g(\tilde{N}_x, \tilde{N}_y) \times \right. \\ \times \exp[i2\pi(\tilde{N}_x x + \tilde{N}_y y)] d\tilde{N}_x d\tilde{N}_y \left. \right\} dx dy = \iint_{-\infty}^{\infty} g(\tilde{N}_x, \tilde{N}_y) \left\{ \iint_{-\infty}^{\infty} f^*(x, y) \times \right. \\ \times \exp[i2\pi(\tilde{N}_x - N_x)x + (\tilde{N}_y - N_y)y] dx dy \left. \right\} d\tilde{N}_x d\tilde{N}_y = \\ = \iint_{-\infty}^{\infty} g(\tilde{N}_x, \tilde{N}_y) g^*(\tilde{N}_x - N_x, \tilde{N}_y - N_y) d\tilde{N}_x d\tilde{N}_y.$$

Положив $N_x = N_y = 0$, получаем

$$\iint_{-\infty}^{\infty} |f(x, y)|^2 dx dy = \iint_{-\infty}^{\infty} g(\tilde{N}_x, \tilde{N}_y) g^*(\tilde{N}_x, \tilde{N}_y) d\tilde{N}_x d\tilde{N}_y = \\ = \iint_{-\infty}^{\infty} |g(\tilde{N}_x, \tilde{N}_y)|^2 d\tilde{N}_x d\tilde{N}_y.$$

Последнее выражение определяет теорему Парсеваля, которая обычно интерпретируется как закон сохранения.

Теорема свертки

Если $F[f(x, y)] = g(N_x, N_y)$, а $F[F(x, y)] = G(N_x, N_y)$, то

$$F \left[\iint_{-\infty}^{\infty} f(\xi, \eta) F(x - \xi, y - \eta) d\xi d\eta \right] = g(N_x, N_y) G(N_x, N_y).$$

Доказательство

$$\begin{aligned} F \left[\iint_{-\infty}^{\infty} f(\xi, \eta) F(x - \xi, y - \eta) d\xi d\eta \right] &= \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\iint_{-\infty}^{\infty} f(\xi, \eta) F(x - \xi, y - \eta) d\xi d\eta \right] \exp[-i2\pi(N_x x + N_y y)] dx dy = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\xi, \eta) \left\{ \iint_{-\infty}^{\infty} F(x - \xi, y - \eta) \exp[-i2\pi(N_x x + N_y y)] dx dy \right\} d\xi d\eta. \end{aligned}$$

Введем новые переменные $q = x - \xi$ и $p = y - \eta$.

При этом

$$\begin{aligned} \exp[-i2\pi(N_x x + N_y y)] &= \\ &= \exp[-i2\pi(N_x \xi + N_y \eta)] \exp[-i2\pi(N_x q + N_y p)]. \end{aligned}$$

Тогда

$$\begin{aligned} F [f(x, y) \otimes F(x, y)] &= \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\xi, \eta) \exp[-i2\pi(N_x \xi + N_y \eta)] d\xi d\eta \times \\ &\times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(q, p) \exp[-i2\pi(N_x q + N_y p)] dq dp = g(N_x, N_y) G(N_x, N_y), \end{aligned}$$

т.е. преобразование Фурье свертки двух функций полностью эквивалентно более простой операции умножения их образов.

Теорема свертки известна как теорема Бореля.

Теорема автокорреляции

Если $F [f(x, y)] = g(N_x, N_y)$, то

$$F \left[\iint_{-\infty}^{\infty} f(\xi, \eta) f^*(x - \xi, y - \eta) d\xi d\eta \right] = |g(N_x, N_y)|^2.$$

Доказательство

$$\begin{aligned} F \left[\iint_{-\infty}^{\infty} f(\xi, \eta) f^*(x - \xi, y - \eta) d\xi d\eta \right] &= \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\iint_{-\infty}^{\infty} f(\xi, \eta) f^*(x - \xi, y - \eta) d\xi d\eta \right] \exp[-i2\pi(N_x x + N_y y)] dx dy. \end{aligned}$$

Введем переменные $q = x - \xi$ и $p = y - \eta$.

Тогда

$$\begin{aligned} F[f(x, y) \otimes f^*(x, y)] &= \iint_{-\infty}^{\infty} f(\xi, \eta) \exp[-i2\pi(N_x \xi + N_y \eta)] d\xi d\eta \times \\ &\times \iint_{-\infty}^{\infty} f^*(q, p) \exp[-i2\pi(N_x q + N_y p)] dq dp = \\ &= g(N_x, N_y) g^*(N_x, N_y) = |g(N_x, N_y)|^2. \end{aligned}$$

Эту теорему можно рассматривать как частный случай теоремы свертки.

Интегральная теорема Фурье

Во всех точках, где функция $f(x, y)$ непрерывна,

$$FF^{-1}[f(x, y)] = F^{-1}F[f(x, y)] = f(x, y),$$

т.е. производя последовательно прямое и обратное преобразования Фурье функции, мы снова получаем эту функцию за исключением точек разрыва. В каждой точке разрыва функции $f(x, y)$ два последовательных преобразования Фурье ее дают среднее значение величины $f(x, y)$ в окрестности этой точки.

Рассмотренные теоремы преобразования Фурье определяют основные правила проведения преобразований и могут существенно упростить решение задач, связанных с использованием анализа Фурье.

Световое возмущение в плоскости изображения, возникающее из-за наличия точечного источника излучения в плоскости предмета, определяется двумерным преобразованием Фурье светового возмущения на сфере в пределах выходного зрачка (комплексного пропускания оптической системы или зрачковой функции) и может быть записано в виде:

$$u(x', y') = -\frac{iR'}{\lambda} \iint_{\Sigma'} F(\mu', \nu') \exp[-ik(\mu'x' + \nu'y')] d\mu' d\nu',$$

где x', y' – координаты в плоскости изображения; R' – радиус сферы сравнения; λ – длина волны; m', M' – координаты элемента поверхности волнового фронта, при этом $\mu' = \frac{m'}{R'}$, $\nu' = \frac{M'}{R'}$; Σ' – площадь выходного зрачка.

Начало координатных осей m', M' совпадает с осевой точкой выходного зрачка, а координатных осей x', y' – с осевой точкой изображения.

В этом выражении $F(\mu', \nu')$ – так называемая зрачковая функция, при этом

$$F(\mu', \nu') = P(\mu', \nu')U(\mu', \nu')\exp[-ikW(\mu', \nu')],$$

где $P(\mu', \nu')$ – единично-нулевая функция, определяемая областью выходного зрачка:

$$P(\mu', \nu') = \begin{cases} 1, & (m', M') \in \Sigma', \\ 0, & (m', M') \notin \Sigma'; \end{cases}$$

$U(\mu', \nu')$ – амплитуда светового возмущения на выходной сфере;
 $W(\mu', \nu')$ – функция волновой аберрации.

Распределение освещенности в изображении точки определяется соотношением

$$E_p(x', y') = U(x', y')U^*(x', y'),$$

при этом относительное распределение освещенности в изображении точечного объекта, равное

$$I(x', y') = \frac{E_p(x', y')}{\int\int_{-\infty}^{\infty} E_p(x', y')dx'dy'}, \quad (14)$$

принято называть функцией рассеяния точки (ФРТ). Отсюда следует, что

$$\int\int_{-\infty}^{\infty} I(x', y')dx'dy' = 1.$$

Распределение освещенности в изображении предмета, образованном оптической системой, определяется выражением

$$E_0(x', y') = \pi L(x', y')\sin^2 \sigma' \cos^4 W', \quad (15)$$

где $L(x', y')$ – приведенное к плоскости изображения распределение яркости излучения изображаемого предмета; σ' – задняя угловая амплитуда оптической системы (угол между крайним лучом осевого пучка и оптической осью в пространстве изображений); W' – полевой угол (угол между главным лучом и оптической осью) в пространстве изображений.

Выражение (15) определяет идеальное отображение распределения яркости на предмете в распределении освещенности в плоскости изображения, т.е. это выражение не содержит ограничений на разрешающую способность оптической системы. Если предположить, что в рассматриваемой зоне изображения $\cos^4 W' \approx 1$, то при изопланатической коррекции аберраций оптической системы реальное распределение освещенности в изображении предмета определяется в результате свертки функции распределения освещенности, определяемой выражением (15) при $W' \approx 0$, и функции рассеяния точки, определяемой выражением (14):

$$E(x', y') = E_0(x', y') \otimes I(x', y').$$

Процесс вычисления интеграла свертки функций в рассматриваемом случае весьма трудоемкий. Однако, в соответствии с пятой теоремой (теоремой Бореля) преобразование Фурье свертки двух функций равно произведению преобразований Фурье свертываемых функций, т.е.

$$e(N'_x, N'_y) = e_0(N'_x, N'_y) i(N'_x, N'_y),$$

где e , e_0 и i – фурье-образы соответствующих функций. Это равенство определяет оптическую систему как фильтр пространственных частот. Функция $i(N'_x, N'_y)$, определяющая каким образом, каждая частотная составляющая передается оптической системой с учетом дифракции, aberrаций, ошибок изготовления и т.д., называется оптической передаточной функцией (ОПФ) оптической системы. Значение модуля ОПФ для каждой пространственной частоты называется коэффициентом передачи модуляции (КПМ). Совокупность значений КПМ для различных пространственных частот составляет функцию передачи модуляции (ФПМ) или частотно-контрастную характеристику оптической системы. Аргумент ОПФ определяет функцию передачи фазы (ФПФ) оптической системы.

ЛИТЕРАТУРА

1. Борн М., Вольф Э. Основы оптики. – М.: Наука, 1970. – 856 с.
2. Волосов Д.С., Цивкин М.В. Теория и расчет светооптических систем. – М.: Искусство, 1960. – 526 с.
3. Гершун А.А. Избранные труды по фотометрии и светотехнике. – М.: Физматгиз, 1958. – 548 с.
4. Гуревич М.М. Фотометрия (теория, методы и приборы). – Л.: Энергоатомиздат, 1983. – 272 с.
5. Зверев В.А., Точилина Т.В. Основы оптоотехники. Учебное пособие. – СПб: СПбГУ ИТМО, 2005. – 293 с.
6. Максудов Д.Д. Астрономическая оптика. – М.-Л.: Гостехиздат, 1946. – 368 с.
7. Марешаль А., Франсон М. Структура оптического изображения. – М.: Мир, 1964. – 295 с.
8. Михель Курт. Основы теории микроскопа. Пер. с нем. под ред. Г.Г. Слюсарева. – М.: ГИТТЛ, 1955. – 276 с.
9. Михельсон Н.Н. Оптические телескопы. Теория и конструкция. – М.: Наука, 1976. – 512 с.
10. Панов В.А., Андреев Л.Н. Оптика микроскопов. Расчет и проектирование. – Л.: Машиностроение, 1976. – 432 с.
11. Русинов М.М. Габаритные расчеты оптических систем. – М.: Геодезиздат, 1959. – 258 с.
12. Тамм И.Е. Основы теории электричества. – М.: Наука, 1976. – 616 с.
13. Чуриловский В.Н. Теория оптических приборов. – М.-Л.: Машиностроение, 1966. – 564 с.
14. Эргономика зрительной деятельности человека / В.В. Волков, А.В. Луизов, Б.В. Овчинников, Н.П. Травникова. – Л.: Машиностроение, 1989. – 112 с.
15. Андреев Л.Н. Прикладная теория аберраций. Учебное пособие. – СПб: СПбГИТМО (ТУ), 2002. – 98 с.
16. Анго А. Математика для электро- и радиоинженеров. – М.: Наука, 1967. – 780 с.
17. Бронштейн И.Н., Семендяев К.А. Справочник по математике для инженеров и учащихся ВТУЗов. – М.: Наука, 1986. – 544 с.
18. Грамматин А.П. Методы синтеза оптических систем. Учебное пособие. – СПб: СПбГИТМО (ТУ), 2002. – 65 с.
19. Забелина И.А. Расчет видимости звезд и далеких огней. – Л.: Машиностроение, 1978. – 184 с.
20. Запрыгаева Л.А., Свешникова И.С. Расчет и проектирование оптических систем: Учебник для вузов. – М.: Логос, 2000. – 584 с.

21. Зверев В.А. Введение в оплотехнику проектирования оптических приборов. Учебное пособие. – СПб: СПбГИТМО (ТУ), 1995. – 104 с.
22. Зверев В.А. Основы геометрической оптики. – СПб: СПбГИТМО (ТУ), 2002. – 218 с.
23. Зверев В.А., Точилина Т.В. Специальные разделы прикладной математики для оплотехников. Учебное пособие. – СПб: СПбГУ ИТМО, 2005. – 235 с.
24. Зверев В.А. Эквивалентные преобразования световой трубки и их приложения // ОМП, 1990. – №9. – С. 43–46.
25. Зверев В.А., Кривоустова Е.В. Оплотехника несферических поверхностей. Учебное пособие. – СПб: СПбГУ ИТМО, 2006. – 203 с.
26. Зверев В.А., Кривоустова Е.В., Точилина Т.В. Оптические материалы. Часть первая. Учебное пособие. – СПб: СПбГУ ИТМО, 2009. – 241 с.
27. Игнатовский В.С. Элементарные основы теории оптических приборов. – Л.-М.: ГТТИ, 1933. – 184 с.
28. Ишанин Г.Г., Панков Э.Д., Андреев А.Л., Польщиков Г.В. Источники и приемники излучения: Учебное пособие для студентов оптических специальностей вузов. – СПб: Политехника, 1991. – 240 с.
29. Ишанин Г.Г., Козлов В.В. Источники излучения. Учебное пособие. – СПб: СПбГУ ИТМО, 2004. – 395 с.
30. Ишанин Г.Г., Козлов М.Г., Томский К.А. Основы светотехники. Учебное пособие. – СПб: Береста, 2004. – 292 с.
31. О'Нейл Э. Введение в статистическую оптику. – М.: Мир, 1966. – 254 с.
32. Проектирование оптических систем: Пер. с англ. / Под ред. Р. Шеннона, Дж. Вайанта. – М.: Мир, 1983. – 432 с.
33. Родионов С.А. Основы оптики. Конспект лекций. – СПб: СПбГИТМО (ТУ), 2000. – 167 с.
34. Русинов М.М. Техническая оптика: Учебное пособие для вузов. – Л.: Машиностроение, 1979. – 488 с.
35. Русинов М.М. Композиция оптических систем. – Л.: Машиностроение, 1989. – 383 с.
36. Слюсарев Г.Г. Геометрическая оптика. – М.-Л. АН СССР, 1946. – 332 с.
37. Чуриловский В.Н. Теория хроматизма и аберраций третьего порядка. – Л.: Машиностроение, 1968. – 312 с.



СПбГУ ИТМО стал победителем конкурса инновационных образовательных программ вузов России на 2007–2008 годы и успешно реализовал инновационную образовательную программу «Инновационная система подготовки специалистов нового поколения в области информационных и оптических технологий», что позволило выйти на качественно новый уровень подготовки выпускников и удовлетворять возрастающий спрос на специалистов в информационной, оптической и других высокотехнологичных отраслях науки. Реализация этой программы создала основу формирования программы дальнейшего развития вуза до 2015 года, включая внедрение современной модели образования.

КАФЕДРА ПРИКЛАДНОЙ И КОМПЬЮТЕРНОЙ ОПТИКИ

Кафедра прикладной и компьютерной оптики была образована в 1996 году путем объединения двух старейших кафедр оптического факультета: кафедры теории оптических приборов и кафедры оптических приборов и компьютерной оптики, в результате чего кафедра стала крупнейшей и ведущей в России в области прикладной оптики, оптических приборов и компьютерной оптики.

Кафедра имеет восемь учебно-исследовательских лабораторий: оптических измерений, контрольно-измерительных оптических приборов, микроскопов и медицинских оптических приборов, геодезических приборов, кинофотоприборов, компьютерных средств контроля оптики, исследования качества оптического изображения, а также лабораторию компьютерной оптики. На кафедре работают два сетевых компьютерных класса.

В учебном процессе используются научный потенциал и лабораторная база крупнейшего в России научного центра в области оптики – ВНЦ ГОИ им. С.И. Вавилова, ведущего оптического предприятия – ОАО «ЛОМО».

Достижения кафедры отмечены двумя Ленинскими премиями, пятью Государственными премиями, премией Совета Министров, премией французской Академии Наук. Кроме того, работы, выполненные на кафедре, отмечались многочисленными медалями и дипломами международных и российских выставок, медалями С.П. Королева, Ю.А. Гагарина, премиями Минвуза.

В настоящее время кафедра ведет подготовку специалистов и научную работу в следующих взаимосвязанных направлениях:

- теория и проектирование оптических систем, традиционно называемые вычислительной оптикой;
- оптические измерения и контроль оптических элементов и систем;

- оптические приборы, включая общую теорию, микроскопы и контрольно-измерительные приборы, геодезические приборы, офтальмологическую оптику;

- математические методы, алгоритмы, программы и компьютерные системы решения оптических задач, что можно коротко назвать компьютерной оптикой.

С 1965 года на кафедре оптико-механических приборов благодаря новаторским идеям и методам С.А. Родионова получает развитие новое направление в прикладной оптике – разработка математических методов, алгоритмов и программного обеспечения для решения оптических задач. Результаты фундаментальных исследований С.А. Родионова в области формирования изображения, теории дифракции и математического моделирования были положены в основу разработки пакета программ для автоматизированного проектирования оптических систем ОПАЛ, широко известного в оптической промышленности.

В настоящее время объектами исследований в области компьютерной оптики, проводимых на кафедре являются:

- методы и программы проектирования оптических систем (методы синтеза, анализа aberrаций и качества изображения, методы оптимизации, расчет допусков);

- обработка данных контроля оптических элементов и систем (интерферограммы, гартманогаммы и т.п.);

- компьютерное моделирование изображения с учетом различных факторов (частичная когерентность, поляризация, нелинейности и неизопланатизм и т.д.);

- теория и методы оптимальной компьютерной юстировки;

- моделирование процессов ближнепольной оптики.

Общепрофессиональная и специальная подготовка специалистов базируется на следующих составляющих:

- фундаментальных и широких знаниях в области предмета разработки и исследования (прикладной оптики, оптотехники, оптических приборов и систем);

- глубоких знаниях прикладной математики, включая математическое моделирование, численные методы, методы оптимизации и т.п.;

- владении современными компьютерными технологиями, особенно применяемыми в инженерной и научной деятельности.

В настоящее время кафедра прикладной и компьютерной оптики факультета оптико-информационных систем и технологий является одним из крупнейших подразделений Санкт-Петербургского государственного университета информационных технологий, механики и оптики, ориентированным на выпуск высококвалифицированных специалистов.

С информацией о кафедре можно ознакомиться на сервере: aco.ifmo.ru

Виктор Алексеевич Зверев
Методология и инструментарий
в современной оптотехнике
Учебное пособие

В авторской редакции

Дизайн

Верстка

Т.В. Точилина

Редакционно-издательский отдел Санкт-Петербургского государственного
университета информационных технологий, механики и оптики

Зав. РИО

Н.Ф. Гусарова

Лицензия ИД № 00408 от 05.11.09

Подписано к печати

Заказ № _____

Тираж _____ экз.

Отпечатано на ризографе