Tomsk Polytechnic University

Antonov V. M., Pichugin V. F.

PHYSICS III

Optics. Quantum Physics. Atomic Physics

Textbook

Tomsk 2002

UDC 53 (076)

Antonov V. M., Pichugin V. F. PHYSICS III. Optics. Quantum Physics. Atomic Physics. Textbook. Tomsk: TPU Press, 2002, 193 pp.

This textbook is the third part of the general course in physics for technical universities. It includes Optics, Introduction to Quantum Physics, Quantum Mechanics, Atomic Physics, Solid State Physics, Nuclear Structure, Fission and Fusion, Particle Physics and Cosmology. The main theoretical concepts are formulated in logical fashion. There are many examples and practice problems to be solved. This textbook has been approved by the Department of Theoretical and Experimental Physics and the Department of General Physics of TPU.

Reviewed by: V.Ya. Epp, Professor of Department Physics and Mathematics, TSPU, D.Sc.

Contents

Preface

This is the third volume of a three-volume course in physics summing up the experience gained by the authors for more than ten years of teaching the subject in English at the Tomsk Polytechnic University.

The stuff of the University has a long history of teaching physics and all of the previous developed ideas have been used. But, by no means, this version is a direct translation from Russian into English. The manual has been written from the very beginning in English using all the advantages of the emotional and expressive language of William Shakespeare and Isaac Newton.

The accent is placed not only on imparting information, but also on the formation of physical thinking by the students and on their mastering of the ideas and methods of the science. Improved ways of treating a number of questions have been found. These made the exposition stricter, and at the same time simpler and easier to understand.

The main content of the present volume is the science of optics, which represents the last section of the classical physics; and the second part of our course, which we named the modern physics. The second part includes more then 70 items dedicated to the most important questions of the modern physics. In particular, the phenomenon of superconductivity and the high-temperature superconductors are discussed in detail.

The present course is intended for the foreign students who might have problems with Russian language, and above all for our Christian friends from Cyprus. Of course, this manual may be used by everybody who wants to improve his knowledge in Physics.

Vladimir M. Antonov
Vladimir F. Pichugin

Tomsk, March, 2002

Part I. Optics

Nature of Light. Laws of Geometrical Optics

1.1. Nature of light

The light is a complicated phenomenon. In some cases it behaves like an electromagnetic wave in others like a stream of special particles (photons).

What oscillates in an electromagnetic wave are the vectors **E** and **H**. Experiments show that physical actions of light are due to the electric vector which we shall call the **light vector**. The change in space and time of the projection of the light vector onto the direction along which it oscillates can be described by the equation:

$$\vec{E} = E_m \cos(\omega t - kr + \alpha) \tag{1.1}$$

Here k is wave number,

r is distance measured along the direction of the light wave.

The ratio of the speed of a light wave in a vacuum c to the phase velocity in a medium v is known as the **absolute refractive index** of medium and is designated by the letter n:

$$n = \frac{c}{v} \tag{1.2}$$

The values of the refractive index characterise the **optical density of** the medium. A medium with a greater n is called optically denser then that with a smaller n, and vice versa.

The wavelengths of visible light are within the following limits: λ = 0.40-0.70μm(4000-7600Å). These values relate to light waves in a vacuum. The velocity of light in a medium with index n is $v = c/n$, so $\lambda = V/\nu = C/\nu n = \lambda_0/n$.

The frequencies of visible light waves are within the limit:

$$\nu = (0.39 - 1.75)\cdot 10^{15}\,Hz \tag{1.3}$$

Neither our eye not any other receiver of luminous energy can track such frequent changes of the energy flux because they register the time-averaged flux .The magnitude of this quantity is called the light **intensity I** at the given point of space.

The density of the flux of electromagnetic energy equals the Pointing vector **S**. Hence,

$$I = \left|\langle \vec{S} \rangle\right| = \left|\langle [\vec{E}\vec{H}] \rangle\right| \tag{1.5}$$

We must note that when considering the propagation of light in a homogeneous medium

$$I \sim E^2_m \tag{1.6}$$

The lines along which light energy propagates are called **rays**. The direction of <**S**> in isotropic media coincides with a normal to the wave surface, i.e. with the direction of the wave vector **k**.The rays are perpendicular to the wave surface. All the light waves are transverse, they usually do not display asymmetry relative to a ray. The explanation is that in **natural light** (i.e. in light emitted by conventional sources) there are oscillations that occur in the most diverse directions perpendicular to a ray (Fig.1.1)

The radiation of a luminous body consists of the waves emitted by its atoms. The process of radiation in an individual atom continues about $10^{-8}\,s$. During this time, a sequence of crests and troughs (or, as it is said, a **wave train)** of about three meters in lengths is formed. The atom “dies out” and then “flares up” again after a certain time elapses. Many atoms “flares up” at the same time .The wave trains are superposed on one another and form the light wave emitted by the relevant body. The plane of

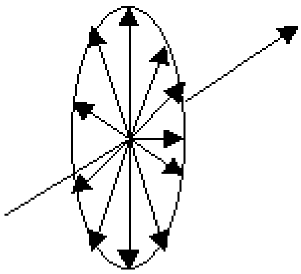


Fig.1

oscillation is oriented randomly for each wave train. Therefore, the resultant wave contains oscillations of different directions with an equal probability. In natural light, the oscillations in different directions follow one another rapidly and without any order. Light in which the direction of the oscillation has been brought into order in some way or other is called **polarized.**

We shall deal with natural light. The ways of obtaining polarized light and its properties are considered in section 3.

The great number of optical phenomena can be easily explained considering light to be an electromagnetic wave. But there are some phenomena (photoelectric effect, in particular) which can not be explained from the wave point of view. To explain the laws of photoeffect we are to use the quantum conceptions and assume light to be the stream of particles having an energy and momentum.

$$E = \hbar\omega \qquad (1.7)$$

$$\vec{p} = \hbar\vec{k} \qquad (1.8)$$

$\hbar$ is Plank constant.

The question ”what is the light: particles or waves” is the old one. In our opinion, the answer is such one. The light is an electromagnetic phenomenon. Sometimes it can be described by ways, sometimes – by particles.

1.2 Measurements of the Speed of Light

The speed of light is the very important world constant which characterise our universe. The most difficulty in measuring of this quantity is due to the very great its value. The first experiment concerning this problem and described in scientific literature was made by G.Galliley in 1601.

An observer *A* sends a light signal to observer *B*. The observer *B* after having received the signal sends it back to observer *A*. If the time needed for light to run the distance *AB* is measured, then it is very easy to calculate the speed of light: c = 2 *AB*/*t*. The distance in the Galliley’s experiment was several scores meters, so the light covered this distance during 10^{-9} s. It is quite obvious that with the aid of medieval technique it was impossible to measure such short time intervals. And instead of measuring the speed of light, the reaction of observers was measured. The speed of light was measured with the aid of astronomical methods by Romer (1676) and Bradly (1725-1728).

The magnitude of the speed of light appeared to be $3.0 \cdot 10^8$ m/s. But the errors were great (1-3%). These experiments were rather complicated and nowadays they are only of a historical interest. In 1849 Physo measured the light speed in laboratory conditions using method of interruption. The scheme of experiment is shown in Fig. 1.2.

The beam of light emitted by the source propagates between the cogs of the window to the mirror M, reflexes and goes back. If the wheel has angle velocity sufficient for next opening between the cogs to be inserted in the light beam, it can be seen throw the ocular E.
Let: the distance AM is D, the number of cogs is z, the number of revolutions is n. The first interruption of light beam occurs when during the time t=2D/c the wheel turns by the angleπ /z, i.e.

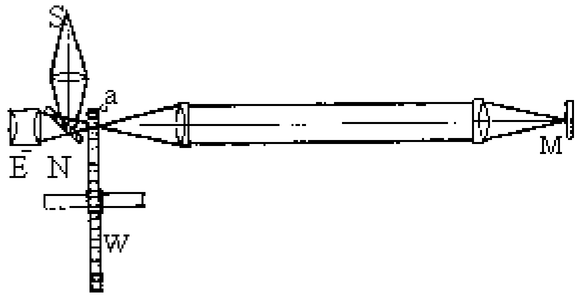


Fig. 1.2

$$\frac{2D}{c} = \frac{1}{2zn}; c = 4Dzn \quad (1.10)$$

Later various scientists made the analogue experiments. The accuracy of methods was improved, but the error was not less than 300 – 500 m/s. All the above-mentioned methods are the direct methods of measurement. In 1972 the speed of light was measured with the aid of an indirect method, namely measuring separately the frequency and wavelength of light ($c = nl$) The accuracy of the method was two orders greater (c = 299792456.2 ±1.1 m/s).

1.3 .**The Ray Approximation in Geometrical Optics**

The lengths of light waves perceived by the human eye are very small (of the order of 10^{-7} m). For this reason, the propagation of the visible light in first approximation can be considered without giving attention to its wave nature and assuming that light propagates along lines called rays. In the limiting case corresponding to $\lambda \to 0$ the laws of optics can be formulated using the language of geometry. Accordingly, the branch of optics in which the finiteness of the wavelength is disregarded is known as **geometrical optics.** Another name for it is **ray optics.**
Geometrical optics is based on four laws: (1) the law of propagation of light along a straight line; (2) the law of independence of light rays; (3) the law of reflection; (4) the law of refraction.
The **law of straight-line propagation** states that in homogeneous medium, light propagates in a straight line, (See Fig.1.3). This law is approximation – when light passes through very small (of the order of a wavelength) openings, deviations from a straight line are observed that increase with diminishing size of the opening.

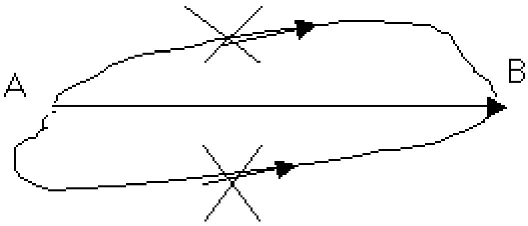


Fig.1.3

Light that travels directly from point A to point B follows the ray path of a straight line rather than some curve line.

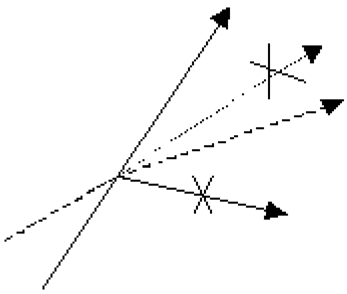


Fig.1.4

The **law of independence of light rays** states that rays do not disturb one another when they intersect, (See Fig.1.4). The intersection of rays does not hinder each of them from propagating independently of the others. This law holds only at not too great luminous intensities. At intensities reached with the aid of lasers, the independence of light rays stops being observed. The **law of reflection** states that reflected ray lies in one plane with the incident ray and the normal to the point of incidence (Fig.1.5). The angle of reflection equals the angle of incidence.

The **law of refraction of light** states that refracted ray lies in one plane with the incident ray and the normal to the point of incidence. The ratio of the sine of the angle of incidence ϑ to the sine of the angle of refraction ϑ'' is constant for the given substance

$$\frac{\sin\vartheta}{\sin\vartheta''} = n_{12} \tag{1.11}$$

The quantity n_{12} is known as the **relative retracted index** of the second substance with respect to the first one. We will discuss this law in detail in the following sections.

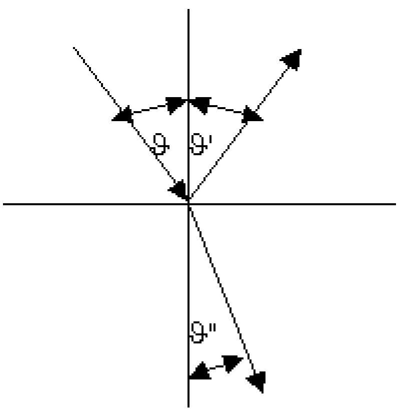


Fig.1.5

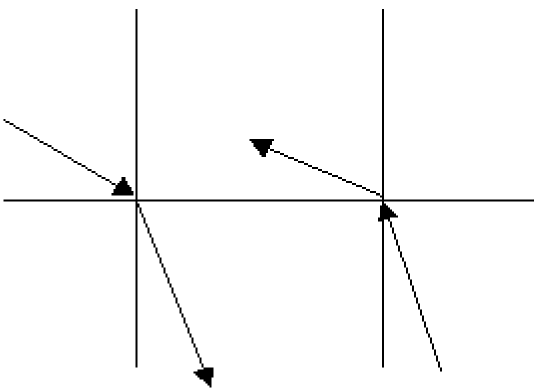


Fig.1.6

Inspection of the equation 1.11 shows that when light passes from an optically denser medium to an optically less dense one, the rays move away from a normal to the interface of the media. An increase in the angle of incidence ϑ is attended by more rapid increase in the angle of refraction ϑ''; when the angle ϑ reaches the value

$$\vartheta_{cr} = \arcsin n_{12} \tag{1.12}$$

the angle ϑ'' becomes equal to $\pi/2$. The angle determined by Eq.(1.12) is called the **critical angle**.

(5**) Principle of ray reversibility**: any actual ray of light in an optical system if reversed in direction will retrace the same path backward. Graphically the reversibility principle can be represented as follows: (see Fig.1.6).

In order to demonstrate the methods of geometrical optics let us discuss the case of refraction (and reflection) of light upon a spherical surface. Assume that the media with refractive indexes n_1 and n_2 are separated by a spherical surface Σ (Fig.1.7).

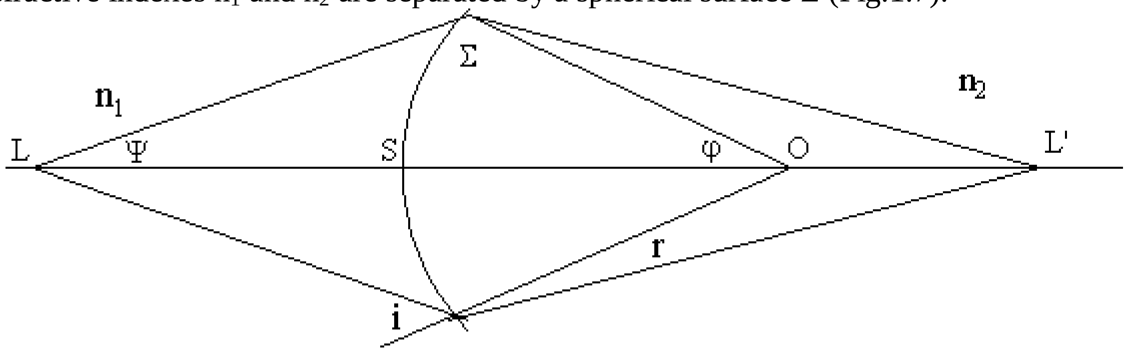


Fig.1.7

A point source L is located upon the line LL' passing through the center of the sphere. Assume that angle ψ is small and approximately the distances $LS = LA; L'S = L'A$ and so on. Such beam of light is called a paraxial one. From the triangle ALO, it follows LO/LA $LO/LA = \sin i / \sin\varphi$; from the triangle OAL' $AL'/OL' = \sin\varphi / \sin r$. Hence, $LO \cdot AL' / LA \cdot OL' = \sin i / \sin r = n_2 / n_1$. We consider the distances in the right direction from the point S to be positive and the distances in the left direction to be negative. Thus, $AL = SL = -a_1; AL' = SL' = a_2; AO = SO = R; LO = -a_1 + R; OL' = a_2 - R$ and accordingly with Eq.(1.13) we have:

$$n_1\left(\frac{1}{a_1} - \frac{1}{R}\right) = n_2\left(\frac{1}{a_2} - \frac{1}{R}\right) = const. \tag{1.14}$$

Eq.(1.14) shows that the quantity n(1/a - 1/R) in the process of refraction is an invariant which is known to be called the **Abbey invariant.** Sometimes it is useful to write the previous expression as follows:

$$\frac{n_1}{a_1} - \frac{n_2}{a_2} = \frac{(n_1 - n_2)}{R} \tag{1.15}$$

Using Eq.(1.15) it is possible to find the distance $a_2 = SL^1$ if the distance a_1 is given. It should be noted that a_2 for given n_1,n_2, and R depends only on a_1. Thus, all the paraxial rays intersect at the same point L^1.
Eq.(1.15) includes all possible variants (a_1, a_2 > or <; R > or < 0). We adviser the reader to investigate all these possibilities and represent them graphically. From the Eq.(1.15), it follows that when $a_{1} = -\infty$

$$a_2 = \frac{n_2 R}{n_2} - n_1 = f_2 \qquad (1.16)$$

And when $a_2 = \infty$

$$a_1 = -\frac{n_1 R}{n_2} - n_1 = f_1 \qquad (1.17)$$

The quantities f_1 and f_2 depend on radius of curvature and refractive indexes n_1, n_2. The quantities f_1 and f_2 are called the **focal distances:** f_1 = **forward focal distance**, the point F_1 is the **forward focus**, f_2 = **back focal distance**, the point F_2 is the **back focus** (Fig.1.8).

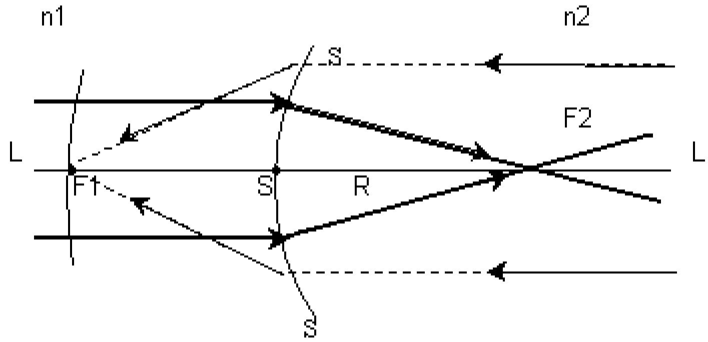


Fig. 1.8

The very important device in optics is a lens. The lens is a transparent material with spherical surface borders (See Fig.1.9). If the width of the lens $S_1S_2 \ll R_1$, R_2 (radii of curvature of the spherical surfaces), the lens is called a thin one. We assume that $S_1S_2 \Rightarrow 0$. I.e. both points may be considered as one point S. The point S is called an **optic centre of the lens**.

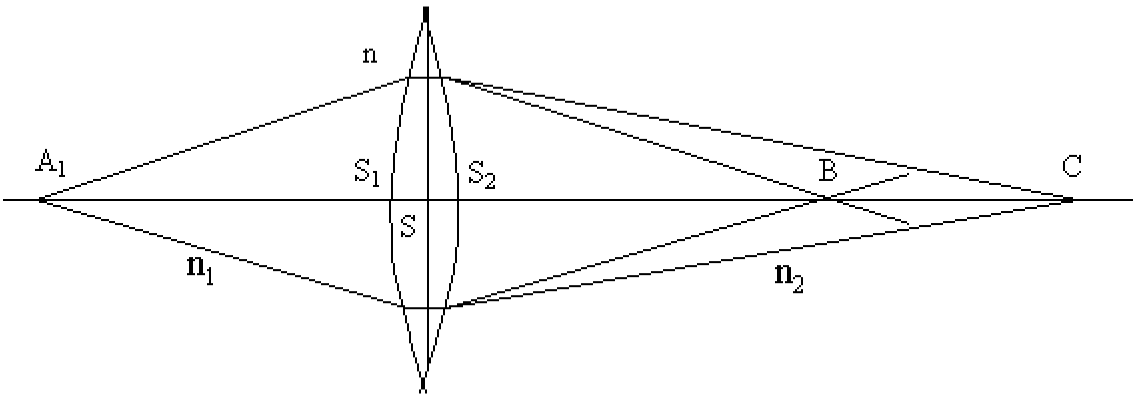


Fig.1.9

Refraction on the first spherical surface would produce in continuous glass with a refractive index n, an image C at a distance SC = a: $n_1 / a_1 - n / a = (n_1 - n) / R_1$ where a_1 =

SA, R = radius of curvature of the first spherical surface. Point C is like an unreal source of light with respect to the second surface. The image of this source is a point B at a distance a_2 = SB. We can write the following expression: $n/a - n_1/n_2 = (n - n_2)/R_2$, where R_2 is the radius of the second surface. There is an air from the both sides of the lens, so $n_1 = n_2$ and
$(n_1/a_1 - n/a) = (n_1 - n)/R_1; (n/a - n_1/a_2) = (n - n_1)/R_2$. By summation of these expressions we get:

$$n_1\left(\frac{1}{a_2} - \frac{1}{a_1}\right) = (n - n_1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) \tag{1.17}$$

Or using the relative refractive index $N = n/n_1$ we get

$$\frac{1}{a_2} - \frac{1}{a_1} = (N - 1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) \tag{1.18}$$

This expression is valid for convex and concave lenses, for any position of a source. It is necessary to have in mind the signs of a_1, a_2, R_1, and R_2 are considered to be positive if the mentioned distances are being measured in the same direction as that one of the previous light ray i.e. to the right, and to be negative if they are directed to the left. If the signs of a_1 and a_2 are like, then one of the focuses is imaginable. Using Eq.(1.18), we have for $a_1 = -\infty$

$$a_2 = f_2 = \frac{1}{(N-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)} \tag{1.19}$$

For $a_2 = -\infty$

$$a_1 = f_1 = -\frac{1}{(N-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)} \tag{1.20}$$

$$f_1 = -f_2 \tag{1.21}$$

Thus, the focal distances are of the same magnitude but of different signs. Introducing f_1 and f_2 in Eq.(1.18) we can write it as follows:

$$\frac{1}{a_2} - \frac{1}{a_1} = \frac{1}{f_1},$$

$$f = f_2 = -f_1 \tag{1.22}$$

It can be seen from Eq.(1.22) that an image moves in the same direction as an object. Only the point $a_1 = f_1$ is an exemption: $a_2 = +\infty$ changes to $a_2 = -\infty$.

1.4 **Reflection and Refraction**

Assume that a plane electromagnetic wave falls on plane interface between two homogeneous and isotropic dielectrics characterized by the permettivities ε_1 and ε_2 ($\mu_1 = \mu_2 = 1$). Experiments show that in this case, besides the plane-refracted wave propagating in the second dielectric, a plane reflected wave propagating in the first dielectric is present.
Let us determine the direction of propagation of the incident reflected and refracted waves with the aid of the wave vectors k, k', k'' correspondingly (Fig.1.10).
The following condition is to be observed:

$$E_{1\tau} = E_{2\tau} \tag{1.23}$$

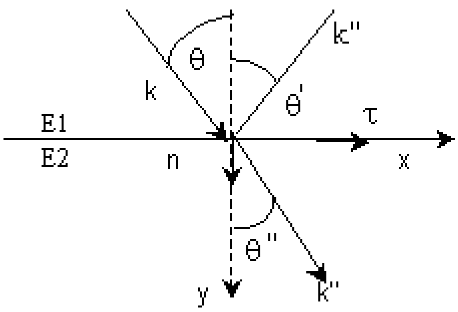


Fig.1.10

Here $E_{1\tau}$ and $E_{2\tau}$ are the tangential components of the electric field strengths in the first and the second media respectively. The plain in which the vectors **k** and **n** (normal to interface) are is called the **plain of incidence**. It is obvious from considerations of symmetry that the vectors $\vec{k}'$ and $\vec{k}''$ can only be in the plain of incidence (the media are homogenous and isotropic).

Let us separate from a naturally falling ray a plain-polarized component in which the direction of oscillation of the vector **E** makes an arbitrary angle with the plain of incidence. The oscillations of the vector **E** in the plain electromagnetic wave propagating in the direction of the vector **k** are described as follows: $\vec{E} = E_m \exp[i(\omega t - \vec{k}\vec{r})] = E_m \exp[i(\omega t - k_x x - k_y y)]$, (with our choice of the coordinate axes, the projection of the vector k onto the z-axis is zero, therefore the addend $k_z z$ is absent in the exponent).By correspondingly choosing the beginning of reading t, we have made the initial phase of the wave equal zero.
The field strengths in the reflected and refracted waves are determined by similar expressions

$$E' = E_m^{'} \exp\left[i\left(\omega' t - k_x' - k_y^{'} + \alpha'\right)\right] \tag{1.25}$$

$$E'' = E_m'' \exp\left[i\left(\omega'' t - k_x'' x - k_y'' + \alpha''\right)\right] \tag{1.26}$$

Where α' and α'' are the initial phases of the relevant waves.
The resultant field in the first medium is

$$E_1 = E + E' = E_m \exp[i(\omega t - k_x x - k_y y)] + E_m' \exp[i(\omega' t - k_x' x - k_y' y + \alpha')] \tag{1.27}$$

In the second medium

$$E_2 = E'' = E''_m \exp[i(\omega'' t - k''_x x - k''_y y + \alpha'')] \qquad (1.28)$$

Having in mind Eq.(1.23) we can write

$$E_{m,\tau} \exp[i(\omega t - k_x x)] + E'_{m,\tau} \exp[i\{\omega' t - k'_x x + \alpha')] = E''_{m,\tau} \exp[i(\omega'' t - k''_x x + \alpha'')] \qquad (1.29)$$

For condition (1.29) to be observed at any time, all the frequencies must be the same.

$$\omega = \omega' = \omega'' \qquad (1.30)$$

Condition (1.29) are to be observed at any place; so the projections of the wave vectors onto the x-axis must be equal, in other words

$$k_x = k_x{}' = k_x{}'' \qquad (1.31)$$

The angles θ, θ′, θ″ shown in Fig.1.10 are called the **angle of incidence**, the **angle of reflection** and the **angle of refraction**.
Obviously, in accordance with Eq.(1.31) we can write

$$k \sin \vartheta = k' \sin \vartheta' = k'' \sin \vartheta'' \qquad (1.32)$$

$$\frac{\omega}{v_1} \sin \vartheta = \frac{\omega}{v_1} \sin \vartheta' = \frac{\omega}{v_2} \sin \vartheta'' \qquad (1.33)$$

Then it follows that

$$\vartheta = \vartheta' \qquad (1.34)$$

$$\frac{\sin \vartheta}{\sin \vartheta''} = \frac{v_1}{v_2} = n_{12} \qquad (1.35)$$

Equation (1.34) is the law of reflection of light and equation (1.35) expresses the law of refraction of light. The quantity n_{12} (**relative refractive index**) can be written in the form:

$$n_{12} = \frac{v_1}{v_2} = \frac{cv_1}{cv_2} = \frac{n_2}{n_1} \qquad (1.36)$$

Thus, the relative refractive index of two substances equals the ratio of their absolute relative indices. Substituting the ratio n_2 / n_1 for n_{12} in equation (1.35), we can write the law of refraction in the form

$$n_1 \sin \vartheta = n_2 \sin \vartheta'' \qquad (1.37)$$

Inspection of this equation show that when light passes from an optically denser medium to an optically less dense medium, the ray moves away from a normal to the interface of the media. An increase in the angle of incidence ϑ is attended by a more rapid growth in the angle of refraction ϑ'', and when the angle ϑ reaches the value

$$\vartheta_{cr} = \arcsin n_{12} \quad (1.38),$$

the angle ϑ'' becomes equal to $\pi/2$.The angle determined by equation (1.38) is called the **critical angle**. The energy carried by an incident ray is distributed between the reflected and refracted rays. As the angle of incidence grows, the intensity of the reflected ray increases, while that of the refracted ray diminishes and vanishes at the critical angle. At angles of incidence within the limits from ϑ_{cr} to $\pi/2$, the light wave penetrates into the second medium to a distance of the order of a wavelength λ and then returns to the first medium. This phenomenon is called **total internal reflection**.

Let us find the relation between the amplitudes and phases of the incident, reflected, and refracted waves. For simplicity, we shall limit ourselves to the normal incidence of a wave onto the interface between dielectrics. Assume that the oscillations of the vector E in the falling wave occur along the direction which we shall take as x-axis. In this case, the condition of continuity of the tangential component of the electric field strength has the form

$$E_x + E'_x = E''_x \quad (1.40)$$

It is well known that the instantaneous values of the vectors **E** and **H** in an electromagnetic wave are related by the expression

$$E_m = \sqrt{\varepsilon\varepsilon_0} = H_m\sqrt{\mu\mu_0} = H_m\sqrt{\mu_0} \quad (1.41)$$

We assume that $\mu = 1$

$$H_m = \sqrt{\varepsilon}\sqrt{\frac{\varepsilon_0}{\mu_0}}E_m = n\sqrt{\frac{\varepsilon_0}{\mu_0}}E_m \quad (1.42)$$

Where n is the refractive index of the medium in which the wave propagates

$$n = \sqrt{\varepsilon\mu} = \sqrt{\varepsilon} \quad (\mu = 1) \quad (1.43)$$

Thus, H~E and the Pointing vector

$$\vec{S} = [\vec{E}\vec{H}] \approx nE^2 \quad (1.44)$$

It is easy to see that the energy conservation law leads to the equation

$$n_1 E_x^2 = n_1 E_x'^2 + n_2 E_x''^2 \quad (1.45)$$

Introducing E''_x - E_x into equation (1.45) instead of E'_x it is easy to see that

$$E''_x = \frac{2n_1}{n_1 + n_2} E_x \tag{1.46}$$

Using this value of E''_x in Eq.(1.40), we find that

$$E'_x = \frac{(n_1 - n_2)}{n_1 + n_2} E_x \tag{1.47}$$

Examination of Eq.(1.46) shows that the projection of the vectors **E** and E''_x have identical signs at each moment of time. Hence, we conclude that the oscillations in the incident wave and in the one passing into the second medium occur at the interface in the same phase – when a wave passes through the interface there is no jump in the phase.

It can be seen from Eq.(1.47) that when $n_2 < n_1$, the sign of E'_x coincides with that of E_x. This signifies that the oscillations in the incident and reflected waves occur at the interface in the same phase – the phase of a wave does not change upon reflection. If $n_2 > n_1$ then the sign of E'_x is opposite to that of E_x, the oscillations in the incident and reflected waves occur at the interface in counterphase – the phase of the wave changes in a jump by π upon reflection. The result obtained also holds upon the inclined falling of a wave at the interface between two transparent media.

Using equations (1.42) and (1.43) we can find the **reflection coefficient ρ** and **transmission coefficient** τ of a light wave (for normal incidence at the interface between two transparent media. Indeed, by definition

$$\rho = \frac{I'}{I} = \frac{n_1 {E'_m}^2}{n_1 E_m^2} \tag{1.48}$$

$$\tau = \frac{I''}{I} = \frac{n_2 {E''_m}^2}{n_1 E_m^2} \tag{1.49}$$

Here I, I' and I″ are intensities of incident, reflected and refracted waves. Using Eqs. (1.46) and (1.47) we get

$$\rho = \left(\frac{n_{12} - 1}{n_{12} + 1} \right)^2 \tag{1.50}$$

$$\tau = n_{12} \left(\frac{2}{n_{12} + 1} \right)^2 \tag{1.51}$$

where $n_{12} = n_1 / n_2$ is the refractive index of the second medium relative to the first one.

We must note that the substitution for n_{12} in Eq.(1.50) of its reciprocal $n_{21} = 1/n_{12}$ does not change the value of ρ. Hence, the coefficient of reflection of the interface between two given media has the same value for both directions of propagation of light.
The index of refraction for glass is close to 1.45. Introducing $n_{12} = 1.5$ into Eq. (1.50), we get ρ = 0.004. Thus, each surface of a glass plate reflects (with incidence close to normal) about four pert cent of the luminous energy falling on it.

1.5. **Fermat`s Principle**

Geometrical optics can be based on the principle established by the French mathematician Pierre de Fermat (1601- 1665). It underlies the laws of straight – line propagation, reflection, and refraction of light. As formulated by Fermat himself, this principle states that *any light ray will travel between two end [points along a line requiring the minimum transit time.*
Light needs the time $dt = ds/v$, where v is the speed of light at the given point of the medium, to cover distance ds (Fig.1.11). Replacing v with c/n, we find that $dt = (1/c)nds.$.
.

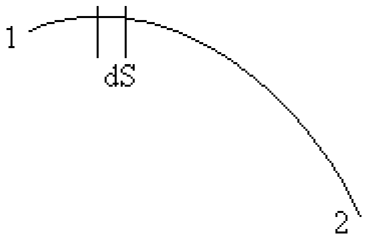


Fig.1.11

Time τ spent by light in covering the distance from point 1 to point 2 is

$$\tau = \frac{1}{c}\int_1^2 nds \qquad (1.52)$$

The quantity

$$L = \int_1^2 nds \text{ n} \qquad (1.53)$$

having the dimension of length is called the **optical path.** In a homogeneous medium, the optical path equals the product of the geometrical path s and the index of refraction n of the medium:

$$L = ns \qquad (1.54)$$

According to Eqs.(1.52), (1.53), we have

$$\tau = \frac{L}{c} \qquad (1.55)$$

The proportionality of the time τ of covering a path to the optical path L makes it possible to word Fermat’s principle as follows: *light travels along a path whose optical length is minimum*. More exactly, the optical path must be extremal, i.e. either minimum or maximum, or stationary - identical for all possible paths. In the last case, all the paths of

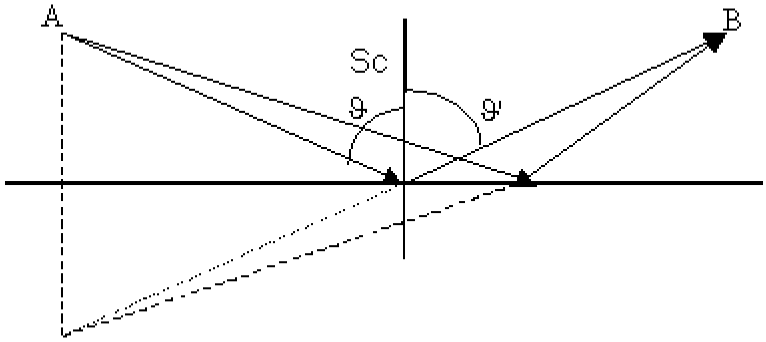


Fig.1.12

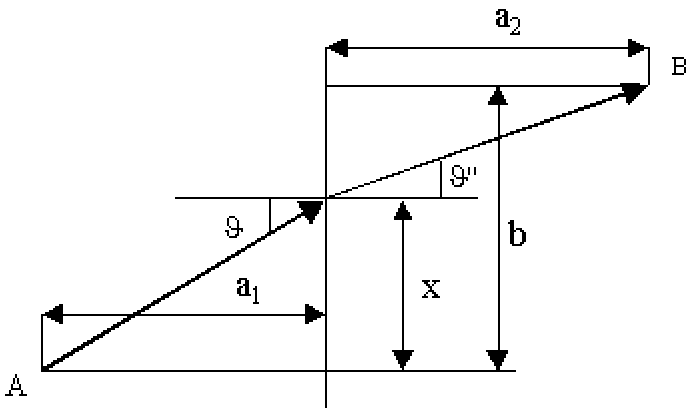


Fig.1.13

light between two points are **tautochronous** (requiring the same time for covering them). The reversibility of light rays ensues from Fermat's principle automatically.

Let us use Fermat,s principle to obtain the laws of reflection and refraction of light. Assume that a ray of light reaches point B from point A after being reflected from surface MN (Fig.1.12, the straight path from A to B is blocked by opaque screen Sc). The medium in which the ray travels is homogeneous. It is obvious that the shortest way is AOB and $\vartheta = \vartheta'$. Any other way (for example A0'B) is longer.

Now let us find the point at which a ray travelling from A to B must be refracted for the optical path to be extremal (Fig.1.13).

The optical path for an arbitrary ray is

$$L = n_1 s_1 + n_2 s_2 = n_1 \sqrt{a_1^2 + x^2} + n_2 \sqrt{a_2^2 + (b-x)^2}$$

To find the extreme value, we have to differentiate L with respect to x and equate the derivative to zero:

$$\frac{dL}{dx} = \frac{n_1 x}{\sqrt{a_1^2 + x^2}} - \frac{n_2 (b - x)}{\sqrt{a_2^2 + (b - x)^2}} = n_1 \frac{x}{s_1} - n_2 \frac{b - x}{s_2} = 0$$

The factors of n_1 and n_2 equal $\sin \vartheta$ and $\sin \vartheta''$, respectively. So, we get the relation

$$n_1 \sin \vartheta = n_2 \sin \vartheta''$$

expressing the law of refraction (Snellius law).
The optical paths are stationary when the rays passthrough a lens (Fig.1.14).

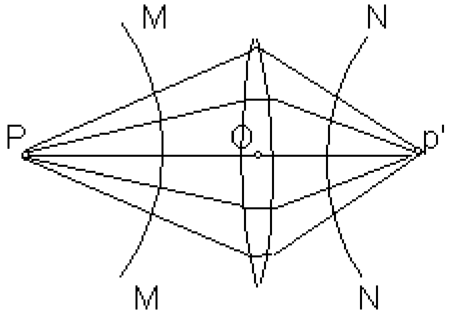


Fig.1.14

Ray POP' has the shortest path in the air (where the index of refraction n is virtually equal to unity) and the longest path in glass (n = 1.5). Ray PQQ'P' has the longest path in the air, but a shortest one in glass. As a result, the optical paths will be the same for all the rays. Hence, latter are tautochronous, avd the optical path is stationary.

1.6. Huygen's Principle

In the chapter 3 we shall have to do with processes taking place behind an opaque barrier with apertures when a light wave impinges on the barrier. In the approximation of geometrical optics, no light ought to penetrate beyond the barrier into the region of the geometrical shadow. Actually, however, a light wave in principle propagates throughout the entire space behind the barrier and penetrates into the region of the geometrical shadow, this penetration being more noticeable, the smaller are the dimensions of the apertures. With a diameter of the apertures or a width of slits comparable with the length of a light wave, the approximation of geometrical optics is absolutely illegitimate.

The behavior of light behind a barrier with an aperture can be explained qualitatively with the aid of **Huygens' principle,** named in honor of the Dutch physicist Christian Huygens (1629 – 1696) who discovered it. The principle establishes the way of constructing a wavefront at the moment of time $t + \Delta t$ accordingly to the known position of the wavefront at the moment t. According to Guygens' principle, every point on an advancing wavefront can be considered as a source of secondary wavelets, and the envelope of these wavelets, defines a new wavefront (Fig.1.15).

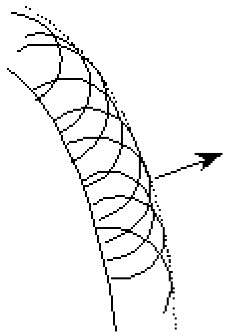


Fig. 1.15

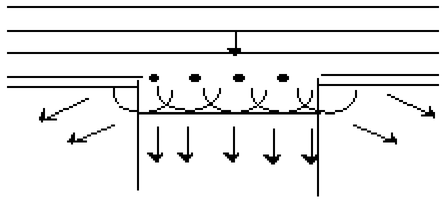


Fig. 1.16

Assume that a plane barrier with an aperture is struck by a wave front parallel to it (Fig.1.16)

According to Huygens, every point on the portion of the wave front bordering on the aperture is a centre of secondary waveles which will be spherical in a homogenous and isotropic medium. Constructing the envelope of these wavelets, we shall see that the wave penetrates beyond the aperture into the region of the geometrical shadow (these regions are shown by dash lines in the figure), bending around the edges of the barrier.

Huygens' principle gives no information on the intensity of waves propagating in various directions. The French physicist Augustin Fresnel (-1788 – 1827) eliminated this shortcoming. The improved Huygens-Fresnel principle is treated in Sec. 3.1, where a physical substantiation of the principle is also given.

2. Interference of Light Waves

2.1 Conditions for Interference

Let us assume that two waves of the same frequency being superposed on each other at a certain point in space produce oscillations of the same direction:

$$\Psi_1 = A_1 \cos(\omega t + \alpha_1); \Psi_2 = A_2 \cos(\omega t + \alpha_2) \tag{2.1}$$

The amplitude of the oscillation at the given point is described by the expression

$$A^2 = A_1^2 + 2A_1 A_2 \cos\delta + A_2^2 \tag{2.2}$$

$\delta = \alpha_2 - \alpha_1$ (Fig.2.1).

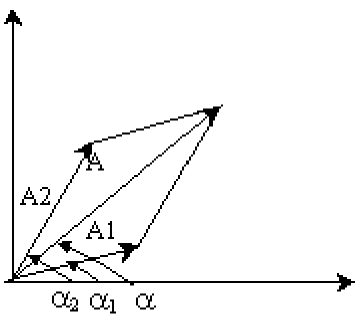


Fig. 2.1.

Expression (2.2) can be easily got with the aid of the thrigonometrical or graphical methods (see Fig.2.1). If the phase different δ of the oscillations set up by the waves remains constant in the time, then the waves are called **coherent.** The phase difference δ for incoherence waves varies continuously. Averaged value of $\cos\delta$ equals zero. Therefore:

$$\langle A \rangle = \langle A_1^2 \rangle + \langle A_2^2 \rangle \tag{2.3}$$

Having in mined that intensity of light I=A^2, we conclude that the intensity observed upon the superposition of incoherent waves equals the sum of the intensities produced by each of the waves individually:

$$I = I_1 + I_2 \tag{2.4}$$

For coherent waves, the value of $\cos\delta$ is a time constant (but a different one for each point of space), so that

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\delta \tag{2.5}$$

At the points of space for which $\cos\delta > 0$, the intensity I will exceed $(I_1 + I_2)$; at the points for which $\cos\delta < 0$, it will be smaller than $(I_1 + I_2)$. Thus, the superposition of coherent light waves is attended by redistribution of the light flux in space. As a result, maxima of the intensity will appear at some spots and minima at others. This phenomenon is called the **interference** of waves. Interference manifests itself especially clearly when the intensity of both interfering waves is the same: $I_1 = I_2$. Hence, according to Eq.(2.5), at the maxima $I = 4I_1$, while at the minima $I = 0$. For incoherent waves in the same condition, we get the same intensity everywhere [see Eq.(2.4)].

It follows from what has been said above that when a surface is illuminated by several sources of light (for example, by two lamps), an interference pattern ought to be observed with a characteristic alternation of maxima and minima of intensity. We know from our everyday experience, however, that in this case the illumination of the surface diminishes monotonously with an increasing distance from the light sources, and no interference pattern is observed. The explanation is that natural light sources are not coherent.

The incoherence of natural light sources is due to the fact that the radiation of a luminous body consists of the waves emitted by many atoms. The individual atoms emit the wave trains with a duration of about 10^{-8} s and a length of about 3 m. The phase of a new train is not related in any way to that of preceding one. In the light wave emitted by a body, the radiation of one group of atoms after about 10^{-8} s is replaced by the radiation of another group, and the phase of the resultant wave undergoes random changes.

Coherent light waves can be obtained by splitting (by means of reflection or refraction) the wave emitted by a single source into two parts. Assume that the splitting into two coherent waves occurs at point O (Fig. 2.2.).

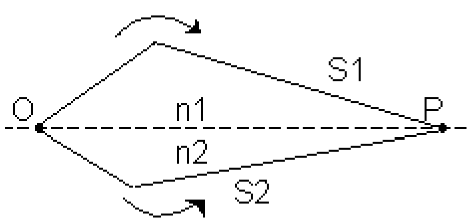


Fig.2.2

Up to point P, the first wave travels the path s_1 in a medium of refractive index n_1, and the second wave travels the path s_2 in a medium of refractive index n_2. It is obvious that the difference in optical path

$$\Delta = n_2 s_2 - n_1 s_1 = L_2 - L_1 \tag{2.6}$$

Hence, the phase difference

$$\delta = \omega\left(\frac{s_2}{v_2} - \frac{s_1}{v_1}\right) = \frac{\omega}{c}\left(n_2 s_2 - n_1 s_1\right) = \frac{2\pi\Delta}{\lambda_0} \tag{2.7}$$

π_0 is the wavelength in a vacuum.
A glance at Eq.(2.7) shows that if the difference in the optical path equals an integral number of wavelengths in a vacuum:

$$\Delta = \pm m\lambda_0, (m = 0,1,2,...) \qquad (2.8)$$

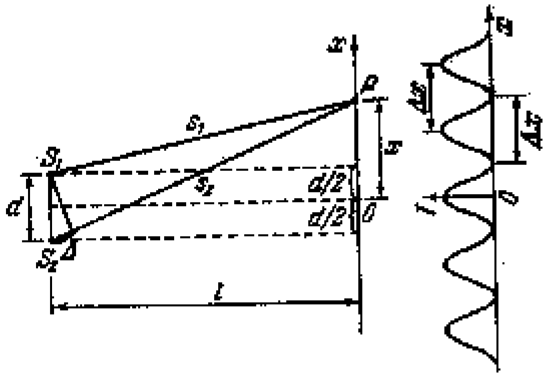


Fig2.3

then the phase difference δ is a multiple of 2π and the oscillations produced at point P by both waves will occur with the same phase. Thus, Eq. (2.8) is the condition for an interference maximum, i.e. for **constructive inteference.**
If the difference in the optical path equals a half-integral number of wavelengthsin a vacuum

$$\Delta = \pm\left(m + \frac{1}{2}\right)\lambda_0; (m = 0,1,2,...) \qquad (2.9),$$

$\delta = \pm(2m + 1)\pi$, so that the oscillations at point P are in counterphase. Thus, Eq. (2.9) is the condition for an interference minimum, i.e. for **destructive interference**.
Examination of Fig.2.3 shows that

$$s_1^2 = l^2 + \left(x - \frac{d}{2}\right)^2; s_2^2 = l^2 + \left(x + \frac{d}{2}\right)^2. \text{ Hence,}$$

$$s_2^2 - s_1^2 = (s_1 + s_2)(s_2 - s_1) = 2xd$$

The experiments show that to obtain a distinguishable interference pattern, the following conditions are to be hold: $d << 1; x << 1$. In these conditions we can assume that $2l = s_1 + s_2. Thus, s_2 - s_1 = \frac{xd}{l}$.Hence,

$$\Delta = n\frac{xd}{l} \qquad (2.10)$$

The introduction of this value of Δ into condition (2.8) shows that intensity maxima will be observed at values of x equal to

$$x_{\max} = \pm m\frac{l}{d}\lambda \quad (\text{m} = 0,1,2,\ldots) \qquad (2.11)$$

Here $\lambda = \lambda_0/n$ is the wavelength in the medium filling the space between the sources and the screen.
Using the value of ☟ given by Eq. (2.10) in condition (2.9), we get the co-ordinates of the intensity minima:

$$x_{\min} = \pm\left(m + \frac{1}{2}\right)\frac{l}{d}\lambda; (m = 0,1,2,...) \qquad (2.12)$$

Let us call the distance between two adjacent intensity maxima the **distance between interference fringes**, and the distance between adjacent minima the **width of an**

interference fringe. It can be seen from Eq. (2.11) and 92.12) that the distance between fringes and the width of a fringe have the same value equal to

$$\Delta x = \frac{l}{d}\lambda \qquad (2.13)$$

According to Eq. (2.13), the distance between the fringes grows with a decreasing distance d between the sources. If d were comparable with l, the distance between the fringes would be of the same order as λ, i.e. would be several scores of micrometers. In this case, the separate fringes would be absolutely indistinguishable. For an interference pattern to become distinct, the condition $d << l$ must be observed.

If the intensity of the interfering waves is the same ($I_1 = I_2 = I_0$), then according to Eq. (2.5) the resultant intensity at the points for which the phase difference is δ is determined by the expression

$$I = 2I_0(1 + \cos\delta) = 4I_0 \cos^2\frac{\delta}{2} \qquad (2.14)$$

The right – hand part of Fig.2.3 shows the dependence of I on x obtained in monochromatic light.

The width of interference fringes and their spacing depends on the wavelength λ The maxima of all wavelength will coincide only at the centre of a pattern when x = 0. With an increasing distance from the centre of the pattern, the maxima of different colours become displaced from one another more and more. The result is blurring of the interference pattern when it is observed in white light. The number of distinguishable interference fringes appreciably grows in monochromatic light. Having measured the distance between the fringes Δx and knowing l and d, we can use Eq.(2.13) to find the wavelength. It is exactly from experiment involving the interference of light that the wavelengths for light rays of various colours were determined for the first time.

2.2. Young's Double – Slit Experiment

The English scientists Thomas Young (1773 – 1829) in 1802 obtained interference from two slits. Young's experiment has tremendous historical significance since it instrumentally demonstrated that light indeed behaved like a wave. However its importance in optics is more than historical. This experiment is an ideal system for illustrating, understanding, and analysing numerous properties of the wave nature of light, including interference, diffraction and coherence.

Figure 2.4 shows the experimental arrangement. A point source of light illuminates two tiny apertures in a opaque screen; the resultant intensity pattern is observed on a viewing screen.

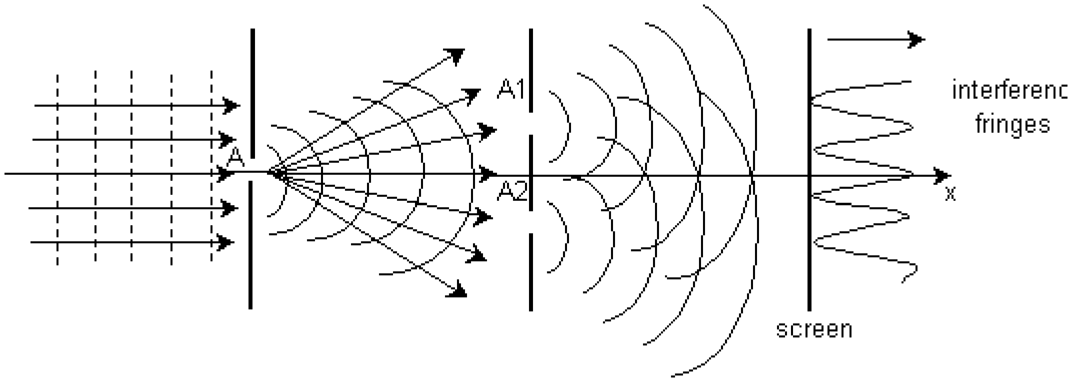


Fig. 2.4

The first aperture A ensures that the light waves that arrive at apertures at apertures A_1 and A_2 originate from a single point source and thus have equal (or at least correlated) phase. The degree to which the phase at two points in a light wave is correlated is called **coherence.** The fact **that** light spreads out as if from a point source when it is passed through a small aperture is a fundamental consequence of the wave nature of light that is called **diffraction**.

The second apertures A_1 and A_2 act like point sources of two spherically (in three dimensions) or cylindrically (in two dimensions) expanding waves which interfere and thus produce fringes on the viewing screen. Young's experiment corresponds to the classic interference experiment (see Fig.2.3). And all formulas of section 2.1 are valid. Thus, for the first time in history, Young observed the interference of light waves and determined the lengths of these waves.

It should be noted that that Young's layout is unique because no optical devices are used. Later the others classical interference layouts were used by other scientists using some optical devices (mirrors, lenses, prisms). For example, the Fresnel double mirror experiment uses reflection for splitting a light wave into two portions. The Fresnel biprism layout uses refraction of light. We shall not discuss these schemes.

2.3 Interference of Light Reflected from Thin Plates

Assume that a plane wave that can be considered as a parallel beam of rays falls on a transparent plane-parallel plate (Fig.2.5). The plate reflected upwards two parallel beams of light: one from the top and the other from the bottom surface. In Fig.2.5 each of these beams is represented by only one ray. In addition to these two beams, the plate throws upward beams produced as a result of three-, five-fold, etc. reflections from the plate surfaces. Owing to their small intensity, however, we shall take no account of these beams. Indeed, at n = 1.5, about 5% of the incident luminous flux is reflected from the surface of the plate. After two reflections, the intensity will be only 0.25%, after three reflections 0.0125% which is 1/ 400 of the intensity of the singly reflected beam. We shall also display no interest in the beams passing through the plate.

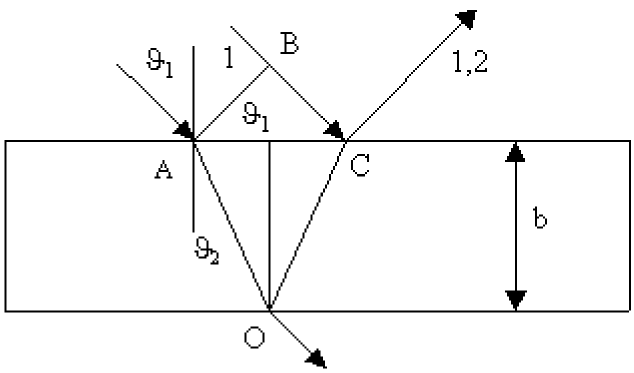


Fig.2.5

The path difference acquired by rays 1 and 2b before they meet at point C is

$$\Delta = ns_2 - s_1 \tag{2.15}$$

Here s_1 = length of segment BC

s_2 =total length of segments AO and OC

n =refractive index of the plate.

We assume that the refractive index of the medium surrounding the plate is unity. From the Fig. 2.5 it follows that $s_1 = 2b \tan \vartheta_2 . \sin \vartheta_1$, and $s_2 = 2b/\cos\vartheta_2$ (b is the thickness of the plate). Using these values in Eq.(2.15), we get

$$\Delta = \frac{2bn}{\cos \vartheta_2} - 2b \tan \vartheta_2 \sin \vartheta_1 = \frac{2b\left(n^2 - \sin \vartheta_2 \sin \vartheta_1\right)}{n \cos \vartheta_2}$$

Substituting $\sin \vartheta_1$ for $n \sin \vartheta_2$ and taking into account that

$$n \cos \vartheta_2 = \sqrt{n^2 - n^2 \sin^2 \vartheta_2} = \sqrt{n^2 - \sin^2 \vartheta_1}$$

we arrive to the equation

$$\Delta = 2b\sqrt{n^2 - \sin^2 \vartheta_1} \qquad (2.16)$$

We must take into account a phase jump at the point C where reflection occurs on the interface ($n_2 > n_1$). Hence, an additional phase difference equal π is produced between rays 1 and 2. It must be taken into account by adding to Δ (or subtracting from it) half a wavelength in a vacuum. The result is:

$$\Delta = 2b\sqrt{n^2 - \sin^2 \vartheta_1} - \frac{\lambda_0}{2} \qquad (2.17)$$

Interference from a plane-parallel plate is observed in practice by placing in the path of the reflected beams a lens that gathers rays at one of the points of the screen in the focal plane of the lens. The illumination at this point depends on the quality Δ. When $\Delta = m\lambda_0$, we get maxima, and when $\Delta = (m + 1/2)\lambda_0$ - minima of the intensity (m is an integer). The condition for the maximum intensity has the form

$$2b\sqrt{n^2 - \sin^2 \vartheta_1} = \left(m + \frac{1}{2}\right)\lambda_0 \qquad (2.18)$$

It is easy to understand that for given values of λ_0 and m there are a number of angles: $\vartheta', \vartheta'', \vartheta'''$ for which condition (2.18) holds. Thus, for a diffuse monochromatic light, the result will be the appearance of a system of alternating bright and dark fringes on the screen. Each fringe will be formed by the rays falling on the plate at the same angle ϑ. This is why the interference fringes produced under such conditions are known as **fringes of equal inclination**. The shape of the fringes depends on position of lens relative to the plate. For example, if the lens is parallel to the plate, the fringes are circular. It should be noted that in all cases, the plane of a screen must coincide with the focal plane of the lens.

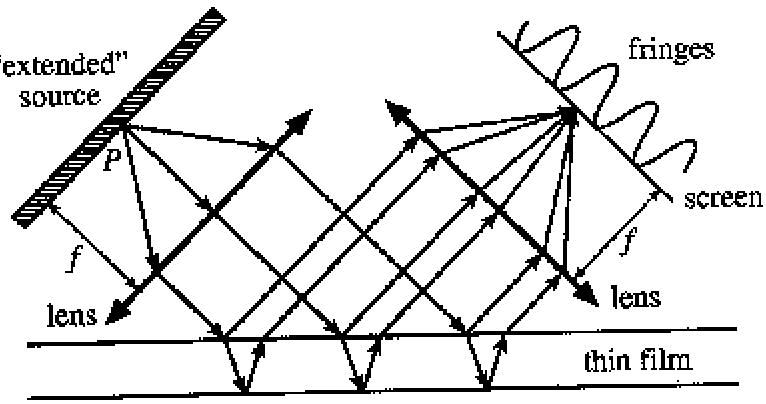


Fig.2.6

Fringes of equal inclination can be viewed as follows (Fig.2.6).
Each point P of the extended source produces rays of an unique angle of inclination (as a resultant action of the first lens) which converges on an unique point on the screen (as a resultant action of the second lens); thus, producing bright and dark fringes.
Suppose that a ray of monochromatic light falling upon a thin film is perfectly collimated, in such a way that there is a single, constant angle of incidence ϑ. Then, if the film thickness is not constant, we will see the alternating regions of constructive and destructive interference on the surface of the film (Fig.2.7).

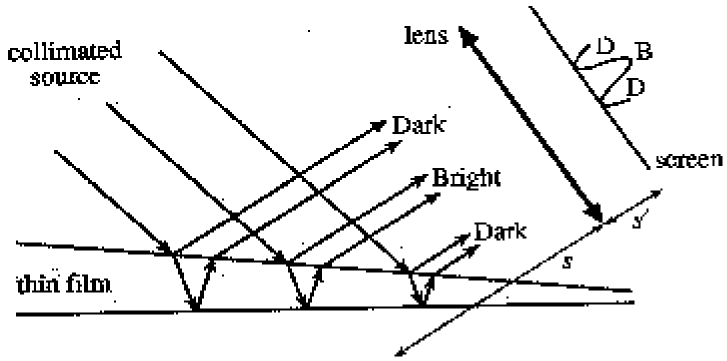


Fig.2.7

Since all locations of the equal thickness of the film produce similar interference, these fringes are called the **fringes of equal thickness**.
When observed in white light, the fringes will be coloured, so that the surface of the plate or film will have rainbow colouring. For example, the thin films of oil on the surface of water and soap films have such colouring. The temper colours appearing on the surface of steel articles when they are hardened are also due to interference from a film of transparent oxides.
A classical example of fringes of equal thickness is **Newton's rings**. They are observed when light is reflected from lens having large radius of curvature (Fig.2.8).

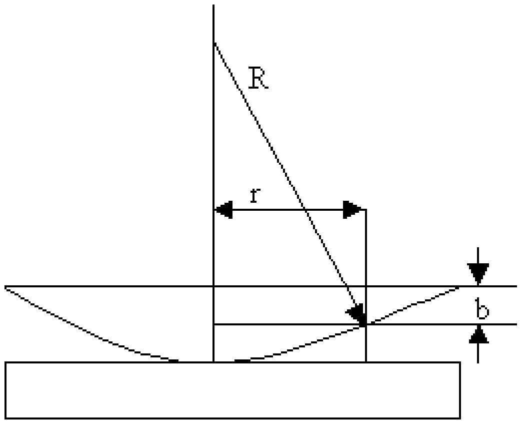


Fig. 2.8

The part of a thin film from whose surfaces coherent waves are reflected is played by the air gap between the plate and the lens (owing to the great thickness of the plate and the lens, no interference fringes appear as a result of reflection from other surfaces). With normal incidence of the light, fringes of equal thickness have the form of concentric rings, and with inclined incidence, of ellipses. Let us find the radii of Newton's rings produced when light falls along a normal to the plate. In this case, $\sin\vartheta_1 = 0$, and the optical path difference equals the double thickness of the gap (n = 1).

It follows from Fig.2.8 that

$$R^2 = (R-b)^2 + r^2 \cong R^2 - 2Rb + r^2 \qquad (2.19)$$

Here R = radius of curvature of the lens

R = radius of a circle with the identical gap b corresponding to all of its points.

To take into account, the change in the phase by π occurring upon reflecting from the plate, we must add λ_0 to $2b = r^2 / R$. Thus:

$$\Delta = \frac{r^2}{R} + \frac{\lambda_0}{2} \qquad (2.20)$$

At points for which $\Delta = m\lambda_0$, maxima appear and at points for which $\Delta = (m + 1/2)\ \lambda_0$, minima of the intensity appear. As a result find the radii of bright and dark Newton's rings:

$$r = \sqrt{\frac{R\lambda_0(m-1)}{2}}; (m = 1,2,3,...) \qquad (2.21)$$

Radii of bright rings correspond to even m's, and radii of dark rings to odd ones. The value
r = 0 corresponds to m = 1, i.e. to the point of the place of contact of the plate and the lens. A minimum of intensity is observed at this point. It is due to the change in the phase by π when a light wave is reflected from the plate.

The coating of lenses is based on the interference of light when reflected from thin films. The transmission of light through each refracting surface of a lens is attended by the reflection of about four per cent 0f the incident light. In multicomponent lenses, such reflections occur many times, and the total loss of the light flux reaches an appreciable value. In addition, the reflection from the lens surfaces result in the appearance of highlights. The reflection of light is eliminated by applying a thin film of a substance having a refractive index other than that of the lens to each free surface of the latter. The components obtained in this way are called **coated lenses**. The thickness of the coating is chosen so that the waves reflected from both its surfaces interfere destructively. An especially good result is obtained if the refractive index of the film equals the square root of the refractive index of the lens. When this condition is satisfied, the intensity of both waves reflected from the film surface is the same.

2.4 The Michelson Interferometr

Many varieties of interference instruments called **interferometers are in use.** Figure 2.9 is a schematic view of a **Michelson interferometer** (named after its inventor, the American physicist Albert Michelson (1852 – 1931)). Light beam from source S falls on semitransparent plate P_1 coated with a thin layer of silver (this layer is depicted by dots in the figure). Half of the incident light flux is reflected by plate P_1 in the direction of ray 1, and half passes through the plate and propagates in the direction of ray 2. Beam 1 is reflected from mirror M_1 and returns to $P_{1,}$ where it is split in two beams of equal intensity. One of them passes through the plate and forms beam 1', and the second one is reflected in the direction of S. The latter beam will no longer interest us. Beam 2 after being reflected by mirror M_2 returns to plate P_1 where it is divided into two parts: beam 2' reflected from the semitransparent layer, and the beam transmitted through the layer, which will also no longer interest us. Light beams 1' and 2' have the same intensity.

The result of the interference of beams 1' and 2' depends on the optical path difference from plate P_1 to mirrors M_1 and M_2 and back. Ray 2 passes through the plate three times, and ray 1 only once. To compensate the resulting change in the optical path difference (owing to dispersion) for waves of different length, plate P_2 is placed in the path of ray 1. Plates P_1 and P_2 except for the silver coating on the former. This arrangement makes the paths of rays 1 and 2 in glass equal. The interference pattern is observed with the aid of telescope T.

Michelson used the described instrument to carry out several experiments that entered annals of physics. The most famous of them, performed together with the American chemist Edward Morley (1838 – 1923) in 1887, had the aim of detecting the motion of the Earth relative to hypothetical ether. In 1890 – 1895, Michelson used the interferometer he had invented to make the first comparison of the wavelength of the red line of cadmium with the length of the standard metre.

In 1881, Michelson carried out his famous experiment by means of which he counted on detecting the motion of the Earth relative to ether (the ether wind). The experiment was based on the following reasoning. Let us assume that interferometer arm PM_2 (Fig.2.9) coincides with the direction of motion of the Earth relative to ether. Consequently, the time needed for ray 1 to cover the path to mirror M_1 and back will differ from the time needed for ray 2 to cover path PM_2P. As a result, even when the length of both arms are equal, rays 1 and 2 will acquire a certain path difference. If we turn the arrangement through 90 degrees, the arms will exchange places, and the path difference will change its sign. This should result in displacement of the interference pattern whose magnitude, as shown by calculations performed by Michelson, could be detected quite readily.

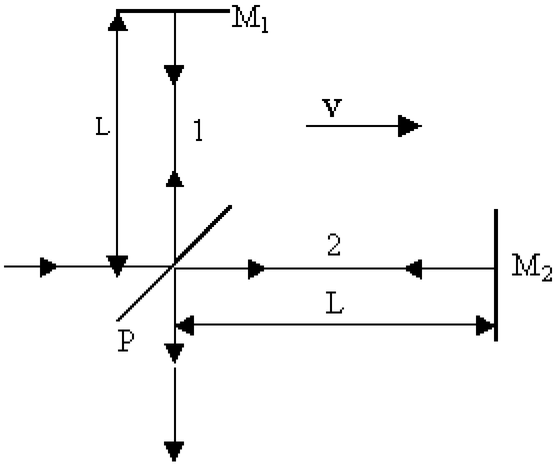


Fig.2.9

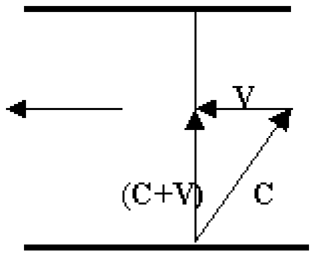


Fig.2.10

To calculate the expected displacement of the interference pattern, let us find the time spent by rays 1 and 2 to cover the relevant paths. Assume that the Earth velocity relative to ether is v. If the ether is not carried along by the Earth and the velocity of light relative to ether is c (refractive index of air is practically equal to unity). Then the velocity of light relative to the instrument will be c - v for direction PM_2 and c+v for direction M_2P. Hence, the time needed for ray 2 is determined by the expression

$$t_2 = \frac{1}{c-v} + \frac{1}{c+v} = \frac{2lc}{c^2 - v^2} = \frac{2l}{c} \cdot \frac{1}{1 - \frac{v^2}{c^2}} \approx \frac{2l}{c}\left(1 + \frac{v^2}{c^2}\right) \tag{2.22}$$

(the Earth velocity along its orbit is 30 km/s, therefore $v^2/c^2 = 10^{-8} \ll 1$).

Before commencing to calculate the time t_1, let us consider the following example from mechanics (Fig.2.10). Suppose that a launch developing the velocity c relative to the water has to cross a river with a current velocity of v in a direction strictly perpendicular to its banks

For the launch to travel in the required direction, its velocity c relative to the water must be directed as shown in the figure. Therefore, the velocity of the launch relative to the banks will be

$|\vec{c} + \vec{v}| = \sqrt{c^2 - v^2}$. The velocity of ray 1 relative to the arrangement (as assumed by Michelson) will be the same. Consequently, the time taken by ray 1 is

$$t_1 = \frac{2l}{\sqrt{c^2 - v^2}} = \frac{2l}{c} \cdot \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = \frac{2l}{c} \cdot \left(1 + \frac{1}{2} \cdot \frac{v^2}{c^2}\right) \tag{2.23}$$

Here we have used the formula $(1-x)^{-\frac{1}{2}} = 1 + \frac{1}{2}x$ holding for small values of x.

Substituting for t_2 and t_1 in the expression $\Delta = c(t_2 - t_1)$ their values from expressions (2.22) and (2.23), we get the path difference for rays 1 and 2:

$$\Delta = 2l\left[\left(1+\frac{v^2}{c^2}\right)-\left(1+\frac{1}{2}\cdot\frac{v^2}{c^2}\right)\right]= l\frac{v^2}{c^2} \tag{2.23}$$

When the arrangement is turned through 90 degrees, the path difference changes its sign. Consequently, the number of fringes by which the interference pattern will be displaced is

$$\Delta N = 2\frac{\Delta}{\lambda_0} = 2\frac{l}{\lambda_0}\cdot\frac{v^2}{c^2} \tag{2.24}$$

The arm length l (taking into account multifold reflections) was 11 m, $_0$ = 0.59µm. It gives
ΔN = 0.4 fringes
The arrangement made it possible to detect a displacement of the order of 0.012 fringes. But no displacement of the interference pattern was detected. The negative result of this classical experiment was explained in terms of Relativity.

2.5 Coherence

By **coherence** is meant the coordinated proceeding of several oscillatory or wave processes. We can accordingly introduce the concept of the **degree of coherence** of two waves.
Temporal and **spatial coherence** are distinguished. We shall begin with a discussion of temporal coherence.
Temporal Coherence. The process of interference described in the preceding section is idealized. This process is actually mach more complicated. The reason is that a monochromatic wave described by the expression $\Psi = A\cos(\omega t - kr + \alpha)$ (A, ω, and α are constants) is an abstraction. A real light wave is formed by superposition of oscillations of all possible frequencies (or wavelengths) confined within more or less narrow but finite range of frequencies Δω (or corresponding range of wavelengths Δλ). Even for light considered to be monochromatic the frequency interval Δω is finite (a "natural" width of spectral lines emitted by atoms have the value of the order of $\Delta\omega = 10^8$ rad/s ($\Delta x \sim 10^{-4}$Å). In addition, the amplitude of the wave A and the phase α undergo continuous random (chaotic) changes with time. Hence, the oscillations produced at a certain point of space by two superposed light waves have the form

$$\Psi_1 = A_1(t)\cos[\omega_1(t)t + \alpha_1(t)]; \Psi_2 = A_2(t)\cos[\omega_2(t) + \alpha_2(t)] \tag{2.24}$$

We shall assume for simplicity's sake that that the amplitudes A_1 and A_2 are constant. Changes in the frequency and phase can be reduced either to change only in the phase, or to a change only in the frequency. Let us write the function

$$f(t) = A\cos[\omega(t)t + \alpha(t)] \qquad (2.25)$$

in the form

$$f(t) = A\cos\{\omega_0 t + [\omega(t) - \omega_0]t + \alpha(t)\}$$

(ω_0 is a certain average value of the frequency), and introduce the notation $[\omega(t) - \omega_0] + \alpha(t) = \alpha'(t)$. Equation (2.25) will thus become

$$f(t) = A\cos[\omega_0 t + \alpha'(t)] \qquad (2.26)$$

We have obtained a function in which only the phase of the oscillation changes chaotically.

On the other hand, it is proved on mathematics that an inharmonic function, for example, function (2.37), can be represented in the form of sum of harmonic functions with frequencies confined within a certain interval $\Delta\omega$ [see Eq. (2.41)].

Thus when considering the matter of coherence, two approaches are possible: a "phase" one and a "frequency" one. Let us begin with the phase approach. Assume that the frequencies ω_1 and ω_2 are identical, i.e. $\omega_1 = \omega_2 = \omega_3 = const.$ According to Eq.(2.5), the intensity of light at a given point is determined by the expression

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\delta(t) \qquad (2.40)$$

$\delta(t) = \alpha_2(t) - \alpha_1(t)$. The last addend in this equation is called the **interference term**.

An instrument that can be used to observe an interference pattern has a certain inertia. Let t_{instr} be the time needed for operation of the instrument. The coherent properties of waves are characterized by introducing the **coherence time** t_{coh}. It is defined as the time during which a chance change in the wave phase $\alpha(t)$ reaches a value of the order of π. When $t_{instr.} << t_{coh.}$, the instrument will detect a sharp interference pattern.

The distance $l_{coh} = ct_{coh}$ is called the **coherence length** (or the **train length**). To obtain an interference pattern by splitting a natural wave into two parts, it is essential that that the optical path difference Δ be smaller than the coherence length.

Let us pass over to a consideration of the part of the non-monochromatic nature of light waves. Assume that consist of a sequence of identical trains of frequency of ω_0 and duration τ. When one train is replaced with another one, the phase experiences disordered changes. As a result, the trains are mutually incoherent. With these assumptions, the duration of a train τ virtually coincides with the coherence time t_{coh}.

In mathematics, the Fourier theorem is proved, according to which any finite and integrable function F(t) can be represented in the form of the sum of an infinite number of harmonic components with a continuously changing frequency:

$$F(t) = \int_{-\infty}^{\infty} A(t) e^{i\omega t} d\omega \qquad (2.41)$$

$$A(\omega) = 2\pi \int_{-\infty}^{\infty} F(t) e^{-i\omega t} dt \qquad (2.42)$$

Assume that the function F(t) describes a light disturbance at a certain point at the moment of time t due to a single wave train:

$$F(t) = A_0 e^{i\omega_0 t}; |t| < \frac{\tau}{2};$$

$$F(t) = 0; |t| > \frac{\tau}{2}.$$

A graph of the real part of this function is given in Fig.(2.11)
Outside the interval from -τ/2 to +τ /2, the function F(t) is zero. Therefore, expression (2.42) determining the amplitude of the harmonic component has the form

$$A(\omega) = 2\pi \int_{-\tau.2}^{+\tau.2} [A_0 \exp(i\omega_0 \xi)] \exp(-i\omega\xi) d\xi = 2\pi A_0 \int_{-\tau/2}^{+\tau/2} [\exp i(\omega_0 - \omega)\xi] d\xi =$$

$$= 2\pi A_0 \left. \frac{\exp[i(\omega_0 - \omega\xi]}{i(\omega_0 - \omega)} \right|_{-\tau/2}^{+\tau/2}$$

After introducing the integration limits and simple transformations, we arrive at the

equation

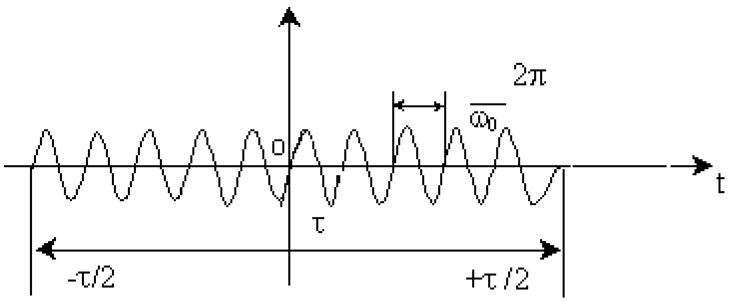


Fig.2.11

$$A(\omega) = \pi A_0 \tau \frac{\sin[(\omega - \omega^\circ)\tau/2]}{[(\omega - \omega^\circ)\tau/2}.$$

The intensity $I(\omega)$ of a harmonic wave component is proportional to the square of the amplitude, i.e. to the expression

$$f(\omega)=\frac{\sin^2\left[(\omega-\omega_0)\frac{\tau}{2}\right]}{\left[(\omega-\omega_0)\frac{\tau}{2}\right]^2} \tag{2.45}$$

A graph of function (2.45) is shown in Fig. 2.12.

A glance at the figure shows that the intensity of the component whose frequencies are within the interval of width $\Delta\omega = \pi 2/\tau$ considerably exceeds the intensity of the remaining components. This circumstance allows us to relate the duration of a train τ to the effective frequency range $\Delta\omega$ of a Fourier spectrum:

$$\tau=\frac{2\pi}{\Delta\omega}=\frac{1}{\Delta\nu} \tag{2.46}$$

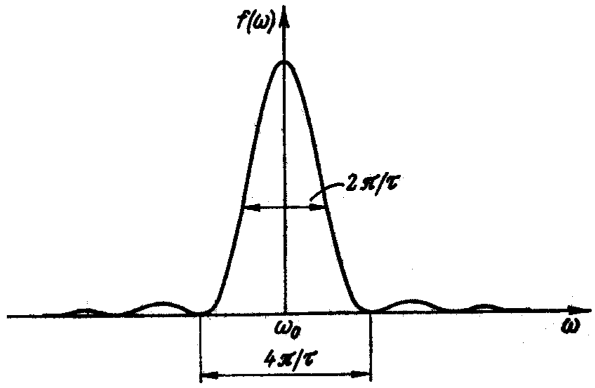


Fig.2.12

Identifying τ with the coherence time, we arrive at the relation

$$\tau_{coh}\approx\frac{1}{\Delta\nu} \tag{2.47}$$

(The sign "$\approx$"stands for “equal to in order of magnitude”).
It can be seen from expression (2.47) that the broader the interval of frequencies present in a given light wave, the smaller is the coherence time of this wave.
The frequency is related to the wavelength in a vacuum by the expression $\nu = c/\lambda_0$. Differentiation of this expression yields $\Delta\nu = c\Delta\lambda/\lambda_0^2 \approx c\Delta x/\lambda^2$. (We have omitted the minus sign obtained in differentiation and also assumed that $\lambda \approx \lambda_0$). Substituting for $\Delta\nu$ in Eq.(2.47) its expression through λ and $\Delta\lambda$, we obtain the following expression for the coherence time:

$$t_{coh}\approx\frac{\lambda}{c\Delta\lambda} \tag{2.48}$$

Hence, we get the following value for the coherence length:

$$L_{coh} = ct_{coh} \approx \frac{\lambda^2}{\Delta\lambda} \tag{2.49}$$

Examination of Eq.(2.35) shows that the path difference at which a maximum of the m-th order is obtained is determined by the relation

$$\Delta_m = \pm m\lambda_0 \cong \pm m\lambda$$

When this path difference reaches values of the order of the coherence length, the fringes become indistinguishable. Consequently, the extreme interference order observed is determined by the condition

$$m_{extr} \approx l_{coh} \approx \frac{\lambda^2}{\Delta\lambda} \tag{2.52}$$

It follows from Eq.(2.52) that the number of interference fringes observed according to the layout shown in Fig. 2.3 grows when the wavelength interval in the light used diminishes.

Spatial coherence. According to the equation $k = \omega/v = n\omega/c$k, scattering in the frequencies $\Delta\omega$ results in scattering of the values of k. We have established that the temporal coherence is determine by the value of $\Delta\omega$. Consequently, the temporal coherence is ass0siated with scattering of the values of the magnitude of the wave vector **k.** Spatial coherence is associated with scattering of the directions of vector **k** that is characterized by the quantity $\Delta\mathbf{e}_\mathbf{k}$.

The setting up at a certain point of space of oscillations produced by waves with different values of $\mathbf{e}_\mathbf{k}$ is possible if these waves are emitted by different sections of an extended (not a point) light source. Let us assume for simplicity's sake that the source has the form of a disc visible from a given point at the angle φ. It can be seen from Fig. 2.13 that the angle φ characterizes the interval confiding the unit vector $\mathbf{e}_\mathbf{k}$**.** We shall consider that this angle is small. This is needed for the degree of temporal coherencetobe sufficient for obtaining a sharp interference pattern

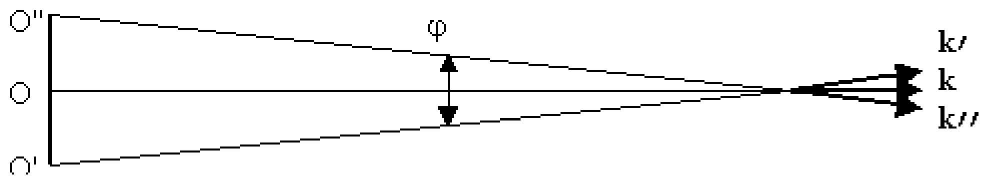


Fig.2.13

Assume that the light from the source falls on two narrow slits behind which there is a screen (Fig. 2.14). We shall consider that the interval of frequencies emitted by the source is very small. This is needed for the degree of temporal coherence to be sufficient for obtaining a sharp interference pattern.
The wave arriving from the section of the surface designated in Fig.2.14 by O produces a zero-order maximum M at the middle of the screen. The zero-order maximum M' produced by the wave arriving from section O' will be displaced from the middle of the screen by the distance x'. Owing to the smallness of the angle φ and the ratio d/l, we can consider that $x' = l\varphi/2$. The zero-order maximum M'' produced by the wave arriving from section O" is displaced in the opposite direction from the middle of the screen over the distance x'' equal x`. The zero-order maxima from the other sections of the source will be between the maxima M' and M".

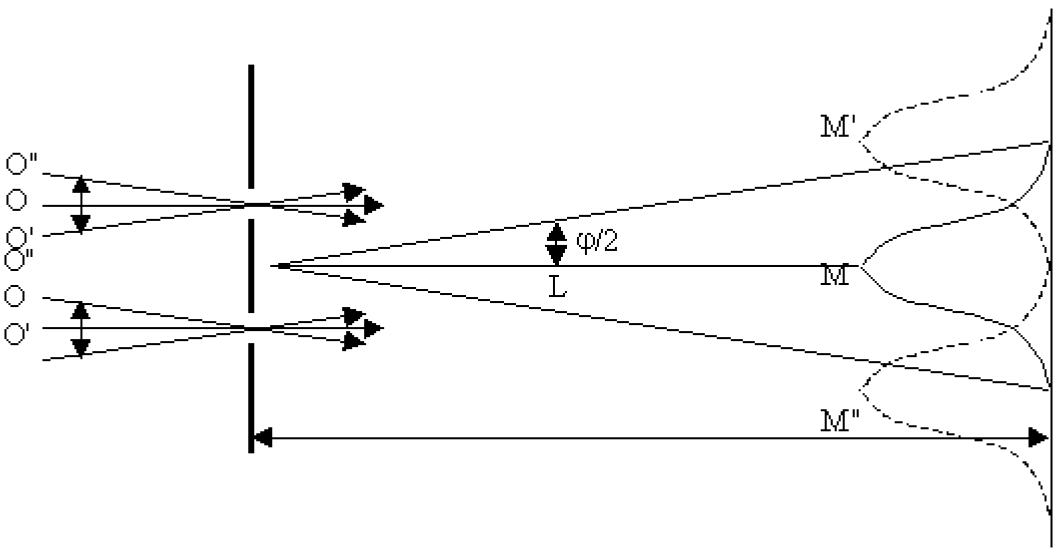


Fig. 2.14

The separate sections of the light source produce waves whose phases are in no way related to one another. For this reason, the interference pattern appearing on the screen will be a superposition of the patterns produced by each section separately. If the displacement x' is much smaller than the width of an interference fringe $\Delta x = l\lambda/d$ [see Eq. (adc)], then the maxima from different sections of the source will practically be superposed on one another, and the pattern will be like one produced by a point source. When x' ≈ Δx, the maxima from some sections will coincide with the minima from others, and no interference pattern will be observed. Thus, an interference pattern will be distinguishable provided that x' < Δx, i.e.

$$\frac{\lambda\varphi}{2} < \frac{l\lambda}{d} \qquad (2.52)$$

or

$$\varphi < \frac{\lambda}{d} \qquad (2.53)$$

We have omitted the factor 2 when passing over from expression (2.52) to (2.53).
Formula (2.53) determines the angular dimensions of a source at which interference is observed. We can also use this formula to find the greatest distance between the slits at which interference from a source with the angular dimension φ can still be observed. Multiplying inequality (2.53) by d/φ, we arrive at the condition

$$d < \frac{\lambda}{\varphi} \qquad (2.54)$$

A collection of waves with different values of $\mathbf{e_k}$ can be replaced with the resultant wave falling on a screen with slits. The absence of an interference pattern signifies that that the oscillations produced by this wave at the places where the first and the second slit s are situated are incoherent. Consequently, the oscillations in the wave itself at points at a distance d are incoherent too. If the source were ideally monochromatic (this means that $\Delta\nu = 0$ and $t_{coh}=\infty$), the surface passing through the slits would be a wave one, and the oscillations at all the points of this surface would occur in the same phase. We have established that when $\Delta\nu \neq 0$ and the source has finite dimensions ($\varphi \neq 0$, the oscillations at points of a surface at a distance of $d > \lambda/\varphi$ are incoherent.
We shall call a surface which would be a wave one if the source were monochromatic, a pseudowave surface for brevity. We could satisfy condition (2.53) by reducing the distance d between the slits, i.e. by taking closer points of the pseudowave surface. Consequently, oscillations produced by a wave at adequately close points of a pseudowave are coherent. Such coherence is called **spatial.**
Thus, the phase of oscillations changes chaotically when passing from one point of pseudowave surface to another. Let us introduce a distance ρ_{coh}, upon displacement by which along a pseudowave surface a random change in the phase reaches a value of about π. Oscillations at two points of a pseudowave surface spaced apart at a distance less than ρ_{coh} will be approximately coherent. The distance ρ_{coh} is called the **spatial coherence length** or the **coherence radius.** It can be seen from expression (2. 54) that

$$\rho_{coh} = \frac{\lambda}{\varphi} \qquad (2.54)$$

The angular dimension of the Sun is about 0.01 radian, and the length of its light waves is about 0.5μm. Hence, the coherence radius of he light waves arriving from the Sun has a value of the order of

$$\rho_{coh} = \frac{0.5}{0.01} = 50\mu m = 0.05mm. \qquad (2.55)$$

The entire space occupied by a wave can be divided into parts in each of whose the wave approximately retains coherence. The volume 0f such a part of space, called the **coherence volume,** in its order of magnitude equals the product of the temporal coherence length and the area of a circle of radius ρ_{coh}..
The spatial coherence of a light wave near the surface of the heated body emitting it is restricted by a value of ρ_{coh} of only a few wavelengths. With an increasing distance from the source, the degree of spatial coherence grows. The radiation of a laser has an enormous temporal and spatial coherence. At the outlet opening of a laser, spatial coherence is observed throughout the entire cross section of the light beam.
It would seem possible to observe interference by passing light propagating from an arbitrary source through two slits in opaque screen. With a small spatial coherence of

the wave falling on the slits, however, the beams of light passing through them will be incoherent, and an interference pattern will be absent. The English scientists Thomas Young (1773-1829) in 1802 obtained interference from two slits by increasing the spatial coherence of the light falling on the slits. Yong achieved such an increase by first passing the light through a small aperture in an opaque screen. This light was used to illuminate the slits in a second opaque screen. Thus, for the first time in history, Young observed the interference of light waves and determined the wavelengths of these waves.

2.6 Multibeam Interference

Up to now, we have dealt with two-beam interference. Now let us investigate the interference of many light rays.
Assume that N rays of the same intensity arrive at a given point of a screen, the phase of each following ray being shifted relative to that of the preceding one by the same value δ. Let us represent the oscillations set up by the rays in the form of exponents:

$$E_1 = ae^{i\omega t}; E_2 = ae^{i(\omega t+\delta)};...; E_m = ae^{i[\omega t+(m-1)\delta]};...; E_N = ae^{i[\omega t+(N-1)\delta]}.$$

Here, a is the amplitude of an oscillation. The resultant oscillation is determined by the formula

$$E = \sum_{m=1}^{N} Em = ae^{i\omega t} \sum_{m-1}^{N} e^{i(m-1)\delta} .$$

The expression obtained is the sum of N terms of a geometrical progression with its first term equal to unity and its common ratio equal to $e^{i\delta}$. Hence,

$$\mathrm{E} = ae^{i\omega t} \frac{1-e^{iN\delta}}{1-e^{i\delta}} = \mathrm{A}°e^{i\omega t}$$

$$\mathrm{A}° = a\frac{1-e^{iN\delta}}{1-e^{i\delta}} \tag{2.56}$$

is the complex amplitude that can be represented in the form

$$A° = Ae^{i\alpha} \tag{2.57}$$

(A is the usual amplitude of the resultant oscillation, α is the initial phase).
The product of quantity (2.57) and its complex conjugate gives the square of the amplitude of the resultant oscillation:

$$A°A°* = Ae^{i\alpha} A^{-i\alpha} = A^2 \tag{2.58}$$

Substituting for A in Eq.(2.58) its value from Eq.(2.56), we get the following expression for the square of the amplitude:

$$A^2 = A°A°* = a^2 \frac{1-e^{iN\delta})(1-e^{-iN\delta})}{(1-e^{i\delta})(1-e^{-i\delta})} = a^2 \frac{2-e^{iN\delta}-e^{-iN\delta}}{2-e^{i\delta}-e^{-i\delta}} =$$

$$= a^2\frac{1-\cos N\delta}{1-\cos\delta} = a^2 \frac{\sin^2(N\delta/2)}{\sin^2(\delta/2)} \tag{2,59}$$

The intensity is proportional to the square of the amplitude. Hence, the intensity produced upon the interference of the *N* rays being considered is determined by the expression

$$I(\delta) = Ka^2 \frac{\sin^2(N\delta/2)}{\sin^2(\delta/2)} = I^\circ \frac{\sin^2(N\delta/2)}{\sin^2(\delta/2)} \qquad (2.60)$$

(K is a constant of proportionality, $I_0 = Ka^2$ is the intensity produced by each of the rays separately).

At the values

$\delta = 2\pi m \quad (m = 0, \pm1, \pm2, \ldots)$

Eq.(2.60) becomes indeterminate. For this reason, we apply L'Hospital's rule. After rather simple calculations we get:

$$I = I_0 N^2 \qquad (2.61)$$

This result could be predicted. Indeed, all the oscillations arrive at points for which $\delta = \pi m$ in the same phase. Hence, the resultant amplitude is N times the amplitude of a separate oscillation, and the intensity is N^2 times that of a separate oscillation.

Let us call the spots where the intensity determined by condition (2.60) is observed the **principal maxima**. Their position is determined by condition (2.60). The number m is called the **order** of the principal maximum. It can be seen from Eq.(2.59) that the space between two adjacent principal maxima accommodates (N-1) minima of the intensity.

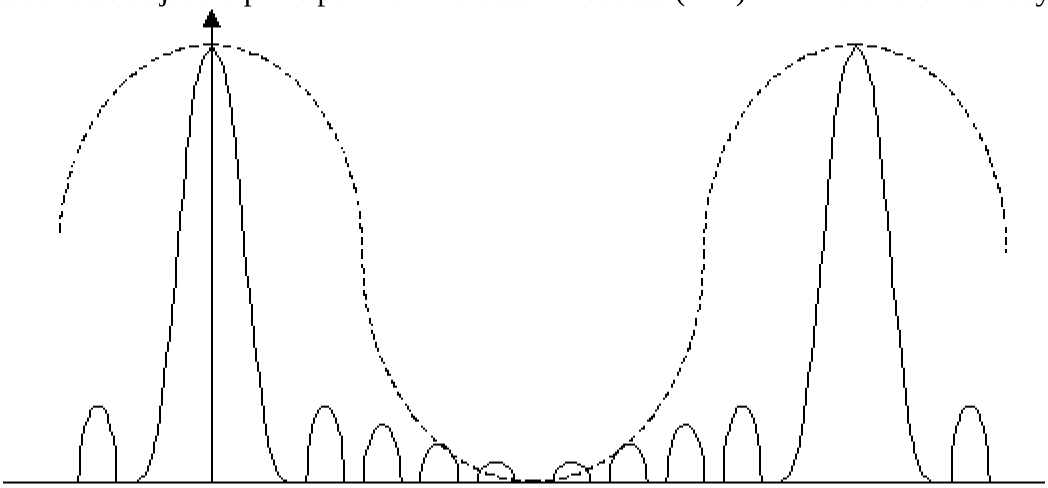


Fig.2.15

Fig.2.15 shows a plot of the function I(δ) for *N* = 10. For comparison, a plot of the intensity for *N*=2 is shown by a dash line. Inspection of the figure shows that the principal maxima become narrower and narrower with an increase in the number of interfering rays. The secondary maxima are so weak that the interference pattern practically has the form of narrow bright lines on a dark background.

3.Diffraction and Polarization

3.1 Introduction to Diffraction

By diffraction is meant the combination of phenomena observed when light propagates in a medium with sharp heterogeneity and associated with deviations from the laws of geometrical optics. Diffraction, in particular, leads to light waves bending around obstacles and to the penetration of light into the region of a geometrical shadow.

There is no appreciable physical difference between interference and diffraction. Both phenomena consist in redistribution of light flux as a result of the superposition of waves. For historical reasons, the redistribution of the intensity produced as a result of the superposition of waves emitted by a finite number of discrete coherent therefore speak about the interference pattern from two narrow slits and the diffraction pattern from one slit.

Two kinds of diffraction are distinguished. If the light source S and the point of observation P are so far from a barrier that the rays falling on the sources arranged continuously has been called the diffraction of waves. We barrier and those travelling to point P form virtually parallel beams, we have to do with **diffraction in parallel rays** or with **Fraunhofer diffraction** (Fig.3.1).

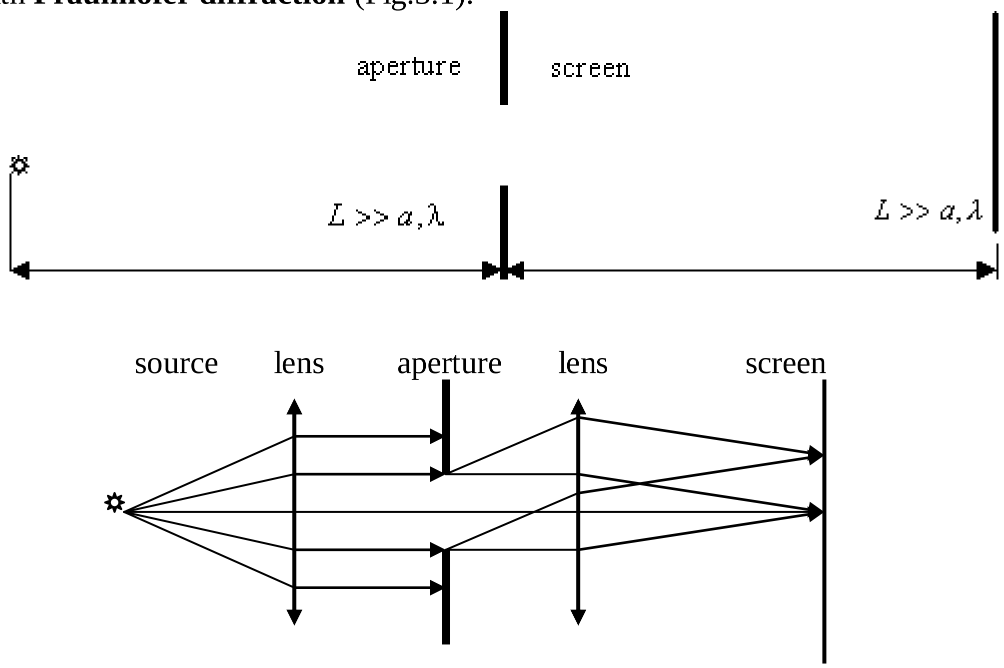


Fig.3.1

Otherwise, we have to do with **Fresnel diffraction** (Fig.3.2)

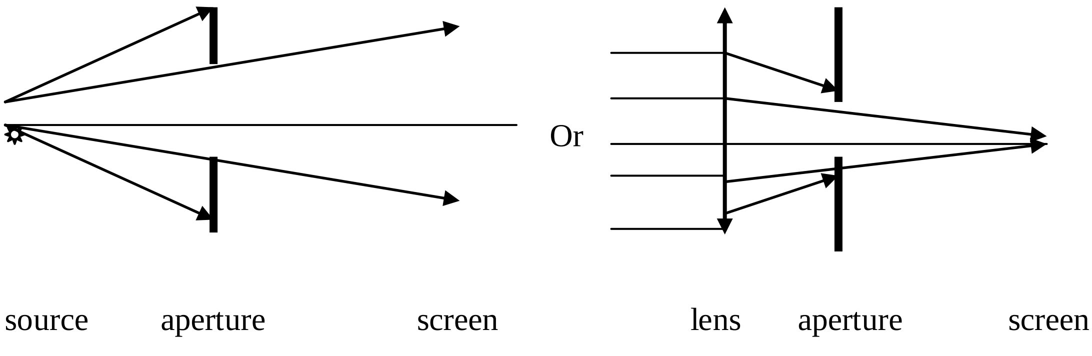


Fig3.2

Historically, a very important example of Fresnel diffraction is Poison's spot (Fig.3.3). The penetration of the light wave into the region of geometrical shadow can be explained with the aid of the Huygens principle.
This principle, however, gives no information on the amplitude, and consequently, on the intensity of waves propagating in different directions. The French physicist

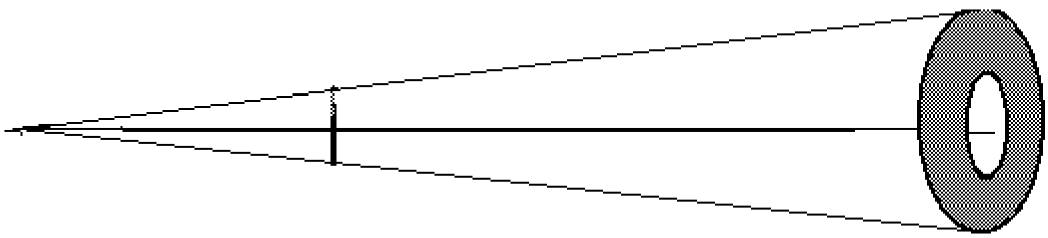


Fig.3.3

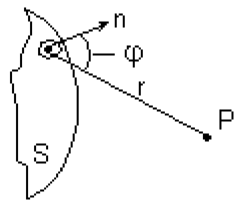


Fig.3.4

Augustin Fresnel (1788-1827) supplemented Huygens principle with the concept of interference of secondary waves. Taking into account the amplitudes and phases of the secondary waves makes it possible to find the amplitude of the resultant wave for any point of the space. Huygens's principle developed in this way was named the **Huygens-Fresnel principle.**

According to the Huygens-Fresnel principle, every element of wave surface (fig.3.4) is the source of a secondary spherical wave whose amplitude is proportional to the size of element dS.

The resultant oscillation at the point P is the superposition of oscillations produced by every element ds, i.e.

$$E = \int_S K(\varphi)\frac{a_0}{r}\cos(\omega t - kr + \alpha_0)dS \qquad (3.1)$$

The coefficient K(φ) depends on the angle φ between a normal **n** to the area dS and the direction from dS to point P. When $\varphi = 0$, this coefficient is maximum; when $\varphi = \pi/2$, it vanishes. The factor a_0 is determined by the amplitude of the light oscillation at the location of dS. This equation is an analytical expression of the Huygens-Fresnel principle.
What has been said above signifies that when calculating the amplitude of the oscillation set up at point P by a light wave propagating from a real source, we can replace this source with a collection of secondary sources arranged along the wave surface. This is exactly the essence of Huygens-Fresnel principle.
The performance of calculations by Eq.(3.1) is a very difficult task in general case. As Fresnel showed, however, the amplitude of resultant oscillation can be found by simple algebraic or geometrical summation in cases distinguished by symmetry (Fig.3.5)

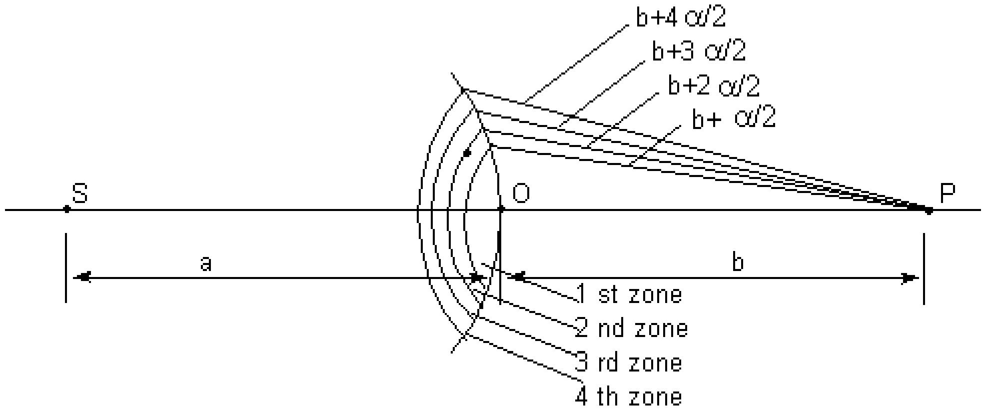


Fig.3.5

Rather simple calculations show that the area of the m zone (when m is not too great) is

$$S_m = \frac{\pi ab\lambda}{a+b} \tag{3.2}$$

and the radius of the outer bounding of the m zone is

$$r_m = \sqrt{\frac{ab}{a+b} m\lambda} \tag{3.3}$$

Zones having the properties shown in Fig.(3.4) are known as the **Fresnel zones.** It if clear that the oscillations arriving at point P from similar points of two adjacent zones (i.e. from points at the middle of the zones, or at the outer edges of the zones, etc.,) are in counter phase. Therefore, the resultant oscillations produced by each of the zones as a whole will differ in phase for adjacent zones by π, too.
If we assume that a = b = 1 m and λ = 0.5μm, then we get a value of r_1 = 0.5 mm for the radius of the first (central) zone. The radii of the following zones grow as $\sqrt{m}$.
Thus, the radii of Fresnel zones are aproximately identical. The distance b_m from a zone to point P slowly increases with the zone number m. The angle φ between a normal to the zone elements and the direction toward the point P also grows with m. All this leads to the fact that the amplitudes of the oscillations produced at point P by Fresnel zones form a monotonously diminishing sequence:

$$A_1 > A_1 > \quad A_2 > A_3 > \ldots A_{m-1} > A_m > \quad A_{m+1} \quad > \ldots \tag{3.4}$$

The phases of the oscillations produced by the adjacent zones differ by π. Therefore, the amplitude A of the resultant oscillation at point P can be represented in the form

$$A = A_1 - A_2 + A_3 - A_4 + \ldots \tag{3.5}$$

Let us write Eq.(3.5) in the form:

$$A = \frac{A_1}{2} + (\frac{A_1}{2} - A_2 + \frac{A_3}{2}) + (\frac{A_3}{2} - A_4 + \frac{A_5}{2}) + \cdots \quad (3.6)$$

Owing to the monotonous diminishing of A_m , we can approximately assume that

$$A_m = \frac{A_{m-1} + A_{m+1}}{2} \quad (3.7)$$

The expression in parentheses will therefore vanish, and Eq.(3.6) will be simplified as follows:

$$A = \frac{A_1}{2} \quad (3.8)$$

According to Eq.(3.8), the amplitude produced at point P by an entire spherical wave surface equals half of the amplitude produced by the central zone alone. If we put in the path of a wave an opaque screen having an aperture that leaves only the central Fresnel zone open, the amplitude at point P will equal A_1, i.e. it will be double the amplitude given by Eq.(3.8). Accordingly, the intensity of the light at point P will in this case be four times greater than where there are no barrier between points S and P.

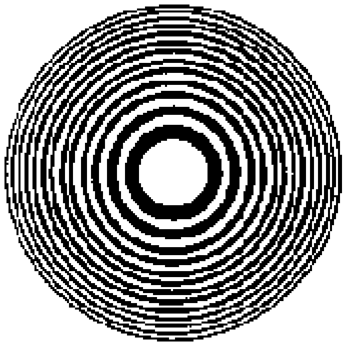


Fig.3.6

The oscillations from the even and odd Fresnel zones are in counterphase and, therefore mutually weaken one another. If we would place in the path of a light wave aplate that would cover all the even or odd zones, the intensity of the light at point P would sharply grow. Such a plate, known as a **zone** one functions like a converging lens. Fig.3.6 shows a plate covering the even zones

A still greater effect can be achieved by changing the phase of the even (or odd) zone oscillations by π instead of covering these zones. This can be done with the aid of a transparent plate whose thickness at the places corresponding to the even or odd zones differs by a proper selected value. Such a plate is called a **phase zone plate.** In comparison with the **amplitude zone plate** covering zones, a phase plate produces an additional two-fold increase in the amplitude, and a four-fold increase in the light intensity.

3.2 Fresnel Diffraction from a Simple Barriers

Diffraction from a Round Aperture. The scheme of experiment is shown in Fig.3.7.

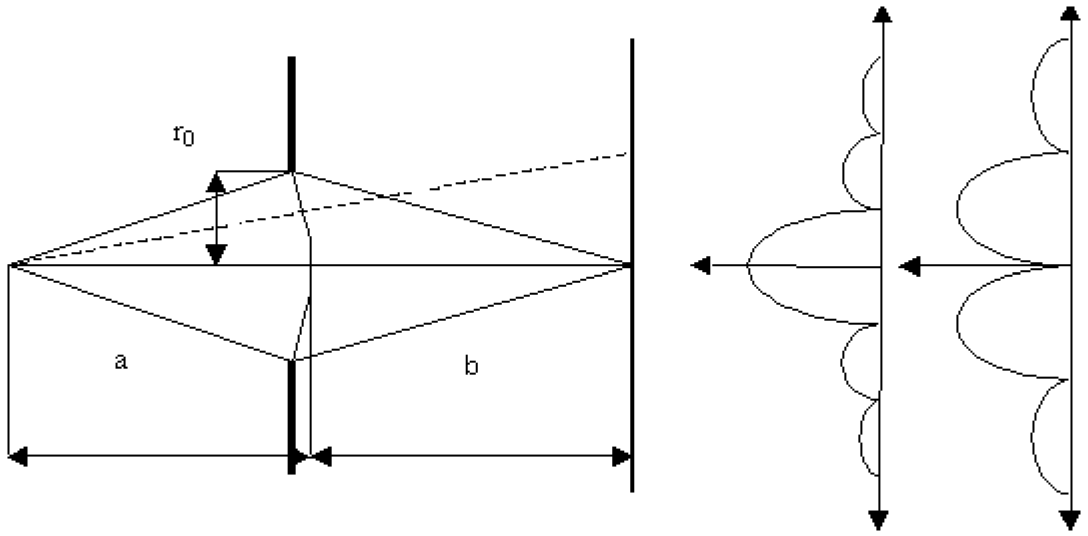


Fig.3.7 b) c)

If the distances a and b satisfy the relation

$$r_0 = \sqrt{\frac{ab}{a+b} m\lambda} \qquad (3.9)$$

where m is an integer, then the aperture will leave open exactly m first Fresnel zones constructed for [point P [see Eq.(3.3)]. Hence, the number of open Fresnel zones is determined by the expression

$$m = \frac{r_0^2}{\lambda}\left(\frac{1}{a} + \frac{1}{b}\right) \qquad (3.10)$$

According to Eq.(3.5), the amplitude at point P will be

$$A = A_1 - A_2 + A_3 - A_4 + \cdots \pm A_m \qquad (3.11)$$

The amplitude A_m is taken with a plus sign if m is odd and with a minus sign if m is even. Then in accordance with Eq.(3.6) we have

$$A = \frac{A_1}{2} + \frac{A_m}{2} \quad \text{(m is odd)} \qquad (3.12)$$

$$A = \frac{A_1}{2} + \frac{A_{m-1}}{2} - A_m \quad \text{(m is even)} \qquad (3.13)$$

The amplitudes from two adjacent zones are virtually the same. We e may therefore replace (A_{m-1} / 2) – A_m / 2. The result is

$$A = \frac{A_1}{2} \pm \frac{A_m}{2}. \qquad (3.14)$$

The plus sign is taken for odd and the minus sign for even m.

The amplitude A_m differs only slightly from A_1 for small m's. Hence, with odd m, the amplitude at point P will approximately equal A_1, and at even m , zero.

If we remove the barrier, the amplitude at point P will become equal to $A_1 / 2$ [see (Eq.3.8)]. Thus, a barrier with an aperture opening small odd number of zones not only does not weaken the illumination at point P, but, on the contrary, leads to an increase in the amplitude almost twice, and of the intensity, almost four times.

The diffraction pattern produced by a round aperture has the form of alternating bright and dark concentric rings. There will be either a bright (m is odd) or dark (m is even) spot at the centre of the pattern. The variation in the intensity I with the distance r from the centre of the pattern is shown in Fig.3.7b (for an odd m) and in Fig.3.7c (for an even m). When the screen is moved parallel to itself along straight line SP, the patterns shown in Fig.3.8 will replace one another [according to Eq.(3.10)], when b changes, the value of m becomes odd and even alternately.

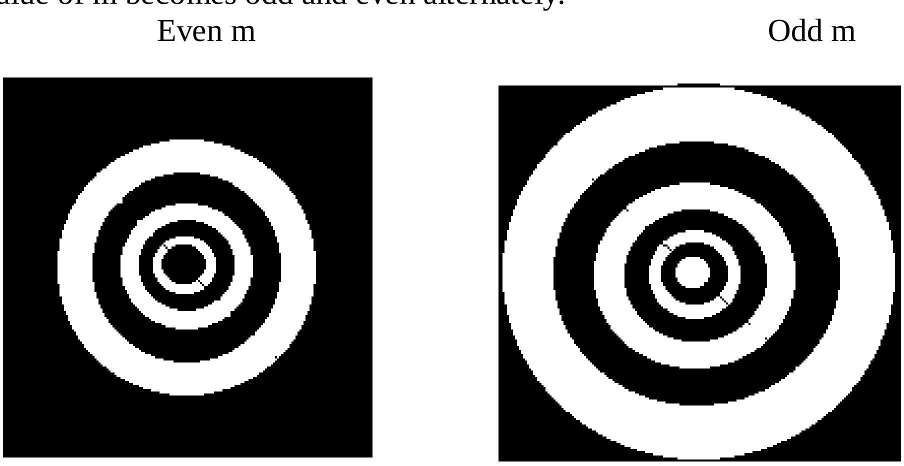


Fig.3.8

If the aperture opens only a part of the central Fresnel zone, a blurred bright spot is obtained on the screen; there is no alternation of dark and bright rings in this case. If thee aperture opens a great number of zones, the alternation of dark and bright rings is observed only in avery narrow region on the boundary of the geometrical shadow; inside this region the illumination is virtually constant.

Diffraction from a Disc. The scheme of experiment is shown in Fig.(3.9). If the dick covers m first Fresnel zones, the amplitude at point P will be

$$A = A_{m+1} - A_{m+2} + A_{m+3} - \cdots = \frac{A_{m+1}}{2} + \left(\frac{A_{m+1}}{2} - A_{m+2} + \frac{A_{m+2}}{2} \right) \qquad (3.15)$$

The expression in parentheses can be assumed to equal zero, consequently

$$A = \frac{A_{m+1}}{2} \tag{3.16}$$

Let us determinate the nature of the pattern obtained on the screen (see Fig.3.9). It is obvious that the illumination can depend only on the distance r from point P. With a small number of covered

Zones, the amplitude A_{m+1} differs slightly from A_1. The intensity at point P will be therefore almost the same as in absence of a barrier between source S and point P [see Eq.3.8)]. For point P' displaced relative to point P in any radial direction, the disc will cover a part of the (m+1) Fresnel zone, and a part of the m zone will be opened simultaneously . This will cause the intensity to diminish. At a certain position of point P', the intensity will reach its minimum. If the distance from the centre of the pattern is still greater, the disc will cover additionally a part of the (m+2) zone, and a part of the (m-1) zone will be opened simultaneously. As a result, the intensity grows and reaches a maximum at point P".

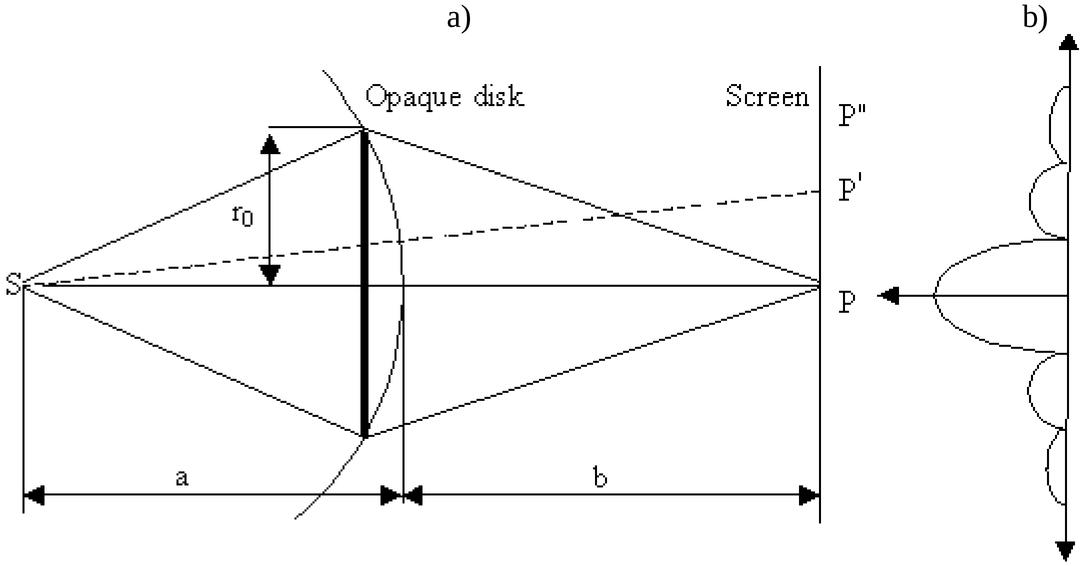


Fig.3.9

Thus, the diffraction pattern for an opaque disc has the form of alternating bright and dark concentric rings. The centre of the pattern contains a bright spot (see Fig.3.3). The light intensity I varieties with the distance r from point P as shown in Fig.3.8b.

If the disc covers only a small part of the central Fresnel zone, it does not form a shadow at all; the illumination of the screen everywhere is the same as in the absence of barriers. If the disc covers many Fresnel zones, alternation of the bright and dark rings is observed only in a narrow region on the boundary of the geometrical shadow. In this case, $A_{m+1} \ll A_1$, so that the bright spot at the centre is absent, and the illumination in the region of the geometrical shadow equals zero practically everywhere.

3.3 Fraunhofer Diffraction from a Slit

The scheme of experiment is shown in Fig.3.9. The wave surface of the incident wave, the plane of the slit, and the screen are parallel to one another.

Let us divide the open part of the wave surface into elementary zones of width dx parallel to the edges of the slit. The secondary waves emitted by the zones in the direction determined by the angle φ will gather at point P. The lens will gather plane (not spherical) waves in the focal plane. Therefore, the factor 1/r in Eq.(3.1) for dE will be absent for Fraunhofer diffraction. Limiting ourselves to a consideration of not too great angles φ, we can assume that the coefficient K in Eq.(3.1) is constant. Hence, the amplitude of the oscillation produced by a zone at any point of the screen will depend only on the area of the zone. The area is proportional to the width dx of the zone.

Consequently, the amplitude dA of the oscillation dE produced by a zone of width dx at any point of the screen will have the form

$$dA = Cdx$$

where C is a constant.

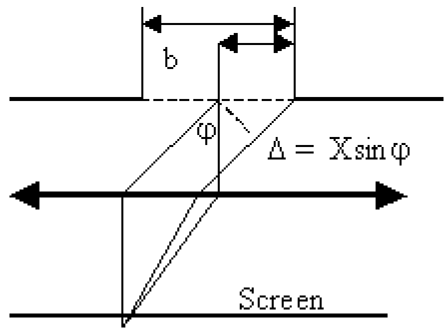


Fig.3.10

Let A_0 stand for the algebraic sum of the amplitudes of the oscillations produced by all the zones at a point of the screen. We can find A_0 by integrating dA over the entire width of the slit b:

$$A_0 = \int dA = \int_{-b/2}^{+b/2} Cdx = Cb \tag{3.18}$$

Hence, $C = A_0/b$ and, therefore,

$$dA = \frac{A_0}{b} dx \tag{3.19}$$

Now let us find the phase relation between the oscillations. We shall compare the phases of the oscillations produced at point P by the elementary zones having coordinates 0 and x. The optical paths OP and QP are tautochronous . Therefore, the path difference between the oscillations being considered is formed on the path Δ equal to x sin φ. If the initial phase of the oscillation produced at point P by the elementary zone at the middle of the slit (φ = 0) is assumed to equal zero, then the initial phase of the oscillation produced by the zone with the coordinate x will be

$$-2\pi\frac{\Delta}{\lambda} = -\frac{2\pi}{\lambda}x\sin\varphi \qquad (3.20)$$

λ is the wavelength in the given medium.
Thus, the oscillation produce by the elementary zone with the coordinate x at point P (whose position is determined by the angle φ) can be written in the form

$$dE_{\varphi} = \left(\frac{A_0}{b}dx\right)\exp\left[t\left(\omega t - \frac{2\pi}{\lambda}x\sin\varphi\right)\right] \qquad (3.21)$$

(we have in mind the real part of this expression).
Integrating Eq.(3.21) over the entire width of the slit, we shall find the resultant oscillation produced at point P by the part of the wave surface uncovered by the slit:

$$E_{\varphi} = \int_{-b/2}^{+b/2}\frac{A_0}{b}\exp\left[i\left(\omega t - \frac{2\pi}{\lambda}x\sin\varphi\right)\right]dx \qquad (3.22)$$

Let us put the multipliers not depending on x outside the integral. In addition, we shall introduce the symbol

$$\gamma = \frac{\pi}{\lambda}\sin\varphi \qquad (3.23)$$

As a result, we get

$$E_{\varphi} = \frac{A_0}{b}e^{i\omega t}\int_{-b/2}^{+b/2}e^{-2i\gamma x}dx = \frac{A_0}{b}e^{i\omega t}\frac{1}{(-2i\gamma)}e^{-2i\gamma x}\Big|_{-b/2}^{+b/2} =$$

$$= e^{i\omega t}\left\{\frac{A_0}{\gamma b(-2i)}\left[e^{-i\gamma b} - e^{i\gamma b}\right]\right\} = e^{i\omega t}\left\{\frac{A_0}{\gamma b}\frac{1}{2i}\left[e^{i\gamma b} - e^{-i\gamma b}\right]\right\} \qquad (3.24)$$

The expression in braces determines the complex amplitude A_{φ} of the resultant oscillation. Taking into account that the difference between the exponents divided by 2i is $\sin\gamma b$, we can write

$$A_{\varphi} = A_0\frac{\sin\gamma b}{\gamma b} = A_0\frac{\sin(\pi b\sin\varphi/\lambda)}{(\pi b\sin\varphi/\lambda)} \qquad (3.25)$$

[we have introduced the value of γ from Eq.(3.23)].
Equation (3.25) is a real one. Its magnitude is the usual amplitude of the resultant oscillation.

$$A_{\varphi} = \left| A_0 \frac{\sin(\pi b \sin\varphi / \lambda)}{(\pi b \sin\varphi / \lambda)} \right| \tag{3.26}$$

For a point opposite the centre of the lens, $\varphi = 0$. Introduction of this value into Eq.(3.26) gives the value A_0 for the amplitude
At values of φ satisfying the condition $\pi b \sin\varphi/\lambda \pm k\pi$, i.e. when

$$b \sin\varphi = \pm k\lambda; (k = 1,2,3,...) \tag{3.27}$$

the amplitude A_{φ} vanishes. Thus, condition (3.27) determines the positions of the minima of intensity. We must note that $b\sin\varphi$ is the path difference Δ of the rays travelling to point P from the edges of the slit (see Fig.3.10).
The intensity of light is proportional to the square of the amplitude. Hence, in accordance with Eq.(3.26),

$$I_{\varphi} = I_0 \frac{\sin^2(\pi b \sin\varphi / \lambda)}{(\pi b \sin\varphi / \lambda)^2} \tag{3.28}$$

Here I_0 is the intensity at the middle of the diffraction pattern (opposite the centre of the lens), and I_{φ} is the intensity at the point whose position is determined by the given value of φ.
We find from Eq.(3.28) that $I_{-\varphi} = I_{\varphi}$. This signifies that the diffraction pattern is symmetrical relative to the centre of the lens. We must note that when the slit is displaced parallel to the screen (along the x-axis in Fig.3.9), the diffraction pattern observed on the screen remains stationary (its middle is opposite the centre of the lens). Conversely, displacement of the lens with the slit stationary is attended by the same displacement of the pattern on the screen.
The relative intensity of maxima are arranged in the following proportion:

$$I_0 : I_1 : I_2 : I_3 ... = 1 : 0.045 : 0.016 : 0.008 : ... \tag{3.29}$$

Thus, the central maximum considerably exceeds the remaining maxima in intensity; the main fraction of the light flux passing through the slit is concentrated in it.
The number of intensity minima is determined by the ratio of the width of a slit b to the wavelength λ. It can be seen from condition (3.27) that $\sin\varphi = \pm k\lambda/b$. The magnitude of $\sin\varphi$ can not exceed unity. Hence, $k\lambda/b < 1$, whence

$$k \le \frac{b}{\lambda} \tag{3.30}$$

At a slit width less than a wavelength, minima do not appear at all. In this case, the intensity of the light monotonously diminishes from the middle of the pattern towards its edges.

The values of the angle φ obtained from the condition bsin $\varphi = \pm\lambda$ correspond to the edges of the central maximum. These values are ± arcsin (λ/b). Consequently, the angular width of the central maximum is

$$\delta\varphi = 2\arcsin\frac{\lambda}{b} \tag{3.31}$$

When b >> λ, the value of sin(λ/b) can be assumed equal to λ/b. The equation for the angular width of the central maximum is thus simplified as follows:

$$\delta\varphi = \frac{2\lambda}{b} \tag{3.32}$$

3.4 Diffraction Grating

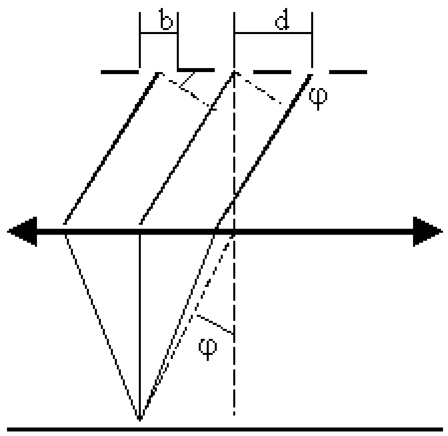


Fig.3.11

A diffraction grating is a collection of a large number of identical equispased slits (Fif.3.11). The distance d between the centres of adjacent slits is called the **period** of the grating. The resultant oscillation at point P whose position is determined by the angle φ is sum of N oscillations having the same amplitude A_φ shifted relative to one another in phase by the same amount δ. According to Eq.(2.60), the intensity in these conditions is

$$I_{gr} = I_\varphi \frac{\sin^2(N\delta/2)}{\sin^2(\delta/2)} \tag{3.33}$$

(here I_φ plays the part of I_0).

A glance at Fig.(3.11) shows that the path difference from adjacent slits is Δ d sin φ. Hence, the phase difference is

$$\delta = 2\pi\frac{\Delta}{\lambda} d\sin\varphi \tag{3.34}$$

Introducing into Eq.(3.33) Eqs.(3.28) and (3.34) for I_φ and δ, respectively, we get

$$I_{gr} = I_0 \frac{\sin^2(\pi b \sin\varphi/\lambda)}{(\pi b \sin\varphi/\lambda)^2} \frac{\sin^2(N\pi d \sin\varphi/\lambda)}{\sin^2(\pi d \sin\varphi/\lambda)} \tag{3.35}$$

(I_0 is the intensity produced by one slit opposite the centre of the lens).
The first multiplier of I_0 in Eq.(3.35) vanishes for points for which condition (3.27) is observed, i.e.,

b sin φ = $\pm k\lambda$ (k = 1, 2, 3, . . .)

At these points, the intensity produced by each slit individually equals zero.
The second multiplier of I_0 in Eq.(3.35) acquires the value N^2 for points satisfying the condition

$$d \sin\varphi = \pm m\lambda; (m = 0,1,2,...) \tag{3.36}$$

[see Eq.(2.61)]. For the directions determined by this condition, the oscillations from individual slits mutually amplify one another. As a result, the amplitude of the oscillations at the corresponding point of the screen is

$$A_{\max} = NA_{\varphi} \tag{3.37}$$

($A_{er\varphi}$ is the amplitude of the oscillation emitted by one of the slit at the angle φ).
Condition (3.36) determines the positions of the intensity maxima called the **principal** ones. The number m gives the **order** of the principal maximum. There is only one zero-order maximum, and there are two each of the maxima of the 1st, 2nd, etc. orders.

Squaring Eq.(3.37), we find that the intensity of the prinsipal maxima I_{max} is N^2 times greater than the intensity I_φ produced in the direction φ by a single slit:

$$I_{\max} = N^2 I_{\varphi} \tag{3.38}$$

Apart from the minima determined by condition (3.27), there are (N-1) additional minima in each interval between adjacent principal maxima . These minima appear in the directions for which the oscillations from individual slits mutually destroy one another. In accordance with Eq.(3.27), the directions of the additional minima are determined by the condition

$$d \sin\varphi = \pm \frac{k'}{N} \lambda \tag{3.39}$$

$(k' = 1,2,3,..., N-1, N+1,...,2N-1,2N+1,...)$

In Eq.(3.39), k' takes on all integral values except for 0, N, 2N, . . ., i.e. except for those at which Eq.(3.39) transforms into Eq.(3.36).
The number of principal maxima observed [see Eq.(3.36)

$$m \le \frac{d}{\lambda} \qquad (3.40)$$

The position of the principal maxima depends on the wavelength λ. Therefore, when white light is passed through a grating, all the maxima except for the central one will expand into a spectrum whose violet end faces the centre of the diffraction pattern, and whose red end faces outward. Thus, a diffraction grating is a spectral instrument. We must note that whereas a glass prism deflects violet rays the greatest, a diffraction grating, on the contrary, deflects red rays to a greater extent.
Fig.3.12 shows schematically the spectra of different orders produced by a grating when white light is passed through it.

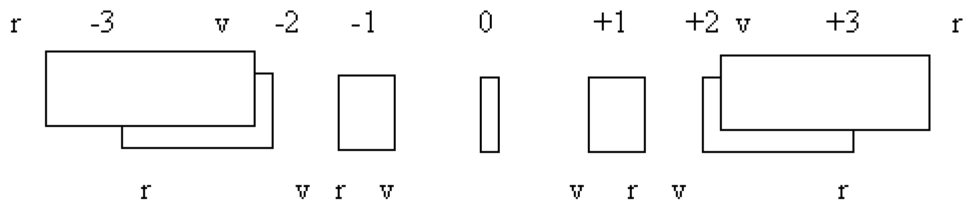


Fig.3.12

At the centre is a narrow zero-order maximum. At both sides of the central maximum are two first-order spectra, then two second-order spectra, etc. The positions of the red end of the m order spectrum and the violet end of the (m+1) order one are determined by the relations

$$\sin\varphi_r = m\frac{0.76}{d}, \sin\varphi_v = (m+1)\frac{0.40}{d}$$

Here d has been taken in micrometers. When the condition is observed that

$$0.76m > 0.40\,(m+1)$$

the spectra of the m and (m+1) orders partly overlap. The inequality gives m>10/9. Hence, partial overlapping begins from the spectra of the second and third orders (see Fig 3.12) in which for illustration the spectra of different orders are displaced relative to one another vertically.
The main characteristics of a spectral instrument are its **dispersion** and **resolving power.** The dispersion determines the angular or linear distance between two spectral lines differing in wavelength by one unit (for example by 1Å). The resolving power determines the minimum difference between wavelength δλ at which the two lines corresponding to them are perceived separately in the spectrum.
The **angular dispersion** is defined as the quantity

$$D = \frac{\delta\varphi}{\delta\lambda} \qquad (3.41)$$

$\delta\varphi$ is the angular distance between spectral lines differing in wavelength by $\delta\lambda$.

From eq.(3.36) we get (omitting the minus sign)

$$d \cos\varphi\delta\varphi = m\delta\lambda$$

$$D = \frac{\delta\varphi}{\delta\lambda} = \frac{m}{d\cos\varphi} \qquad (3.42)$$

Within the range of small angles $\cos\varphi \cong 1$. We can therefore assume that

$$D \cong \frac{m}{\lambda} \qquad (3.43)$$

Linear dispersion is defined as the quantity

$$D_{lin} = \frac{\delta l}{\delta\lambda} \qquad (3.44)$$

Here δl is the linear distance on a screen or photographic plate between spectral lines differing in wavelength by $\delta\lambda$. A glance at Fig.3.13 shows that for small values of the angle φ we can assume that $\delta l \approx f'\delta\varphi$, where f' is the focal length of the lens gathering the diffracted rays on a screen. Thus,

$$D_{lin} = f'D \qquad (3.45)$$

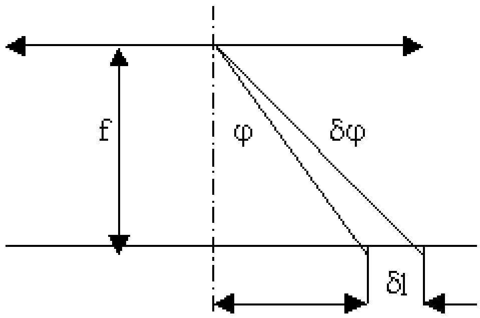


Fig.3.13

Taking into account Eq.(3.44), we get

$$D_{lin} = f'\frac{m}{\lambda} \; D_{lin} = f'\frac{m}{\lambda} D \qquad (3.46)$$

The resolving power of a spectral instrumrnt is defined as the dimensionless quantity

$$R = \frac{\lambda}{\Delta\lambda}$$

(3.47) where $\delta\lambda$ is the minimum difference between the wavelengths of two spectral lines at which these lines are perceived separately.

The possibility of resolving (i.e. perceiving separately) two close spectral lines depends not only on he distance between them (that is determined by the dispersion of the instrument), but also on the width of the spectral maximum. Fig.3.14 shows the resultant intensity (solid curves) observed in superposition of two close maxima (the dash curves). In case a, both maxima are perceived as a single one. In case b, there is a minimum between the maxima. Two close maxima are perceived by the eye separately if the intensity in the interval between them is not over 80 per cent of the intensity of a maximum. According to the criterion proposed by the British physicist John Rayleigh (1842-1919), such a ratio of the intensities occurs if the middle of one maximum coincides with the edge of another one (Fig.3.14b). Such a mutual arrangement of the maxima is obtained at a definite (for the given instrument) value of $\delta\lambda$

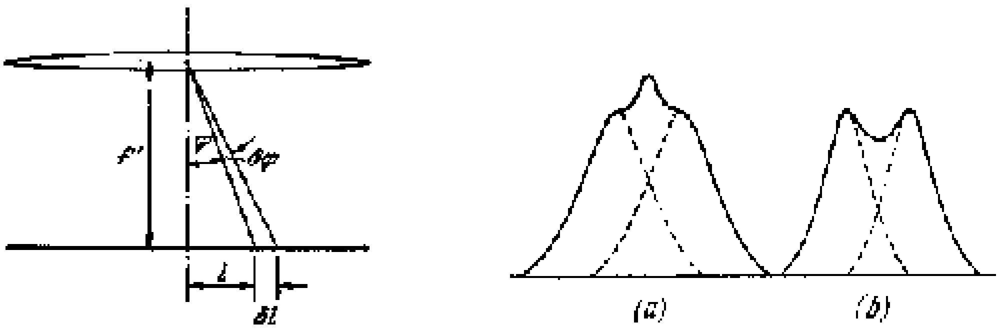


Fig.3.14

Let us find the resolving power of diffraction grating. The position of the middle of the m maximum for the wavelength $\lambda \pm \delta\lambda$ is determined by the condition

$$d \sin \varphi = m(\lambda \pm \delta\lambda)$$

The edges of the m maximum for the wavelength λ are at the angles complying with the condition

$$d \sin \varphi_{\min} = \left(m \pm \frac{1}{N} \right) \lambda$$

The middle of the maximum for the wavelength $\lambda \pm \delta\lambda$ coincides with the edge of the maximum for the wavelength λ if

$$m(\lambda + \delta\lambda) = \left(m + \frac{1}{N} \right) \lambda .$$

Whence

$$m\delta\lambda = \frac{\lambda}{N}$$

Solving this equation relative to λ/δλ, we get an expression for the resolving power:

$$R = mN \tag{3.48}$$

Thus, the resolving power of a diffraction grating is proportional to the order m of the spectrum and the number of slits N.

The best gratings have up to 1200 lines per mm ($d \cong 0.8 mkm$). It can be seen from Eq.(3.40) that no second-order spectra are observed in visible light with such a period. The total number of lines in such gratings reaches 200000 (they are about 200 mm long). With a focal length of the instrument f ' = 2 m, the length of the visible first-order spectrum in this case is over 700 mm.

3.5 Diffraction of X-Rays

Diffraction can be observed in three-dimensional structures, i.e., spatial formations displaying periodicity along three directions not in one plane. All crystalline bodies are such structures. Their period (~ 10^{-10} m), however , is too small for the observations of diffraction in visible light. The condition d > λ is observed for crystals only for X-rays. The diffraction of X-rays from crystals was first observed in 1913 in an experiment conducted by the German physicist Max fon Laue (1879-1959).

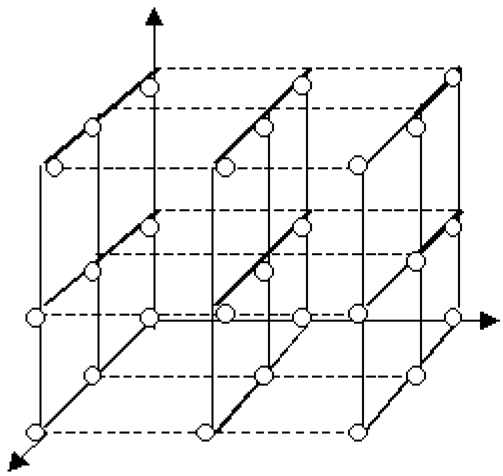


Fig.3.15

Let us find the conditions for the formation of diffraction maxima from a three-dimensional structure. We position the coordinate axes x, y, and z in the directions along which the properties of the structure display periodicity (Fig.3.15). The structure can be represented as a collection of equally spaced parallel trains of structural elements arranged along one of the coordinate axes. We shall consider the action of an individual linear train parallel, for instance, to the x-axis (Fig.3.16). Assume that a beam of parallel rays making the angle α_0 with the x-axis falls on the train. Every structural element is a source of secondary wavelets. An incident wave arrives at adjacent sources with a phase difference of $\delta_0 = 2\pi\Delta_0/\lambda$, where $\Delta_0 = d_1 \cos\alpha_0$ (here d_1 is the period of the structure along the x-axis). Apart from this, the additional path difference $\Delta = d_1 \cos\alpha$ is produced between the secondary wavelets propagating in directions that make the angle α with the x-axis (all such directions lie along the generatrices of a cone whose axis is the x-axis). The oscillations from different structural elements will be mutually amplified for the directions for which

$$d_1(\cos\alpha - \cos\alpha_0) = \pm m_1\lambda; (m_1 = 0,1,2,...) \quad (3.49)$$

There is a separate cone of directions for each value of m_1, and along these directions we get maxima of the intensity from one individually taken train parallel to the x-axis. The axis of this cone coincides with the x-axis.

The conditions of the maximum for a train parallel to the y-axis has the form

$$d_2(\cos\beta - \cos\beta_0) = \pm m_2\lambda; (m_2 = 0,1,2,...)$$
(3.50)

Here d_2 = period of the structure in the direction of y-axis,

β_0=angle between the incident beam and the y-axis,

β =angle between the y-axis and the directions along which diffraction maxima are obtained.

A cone of directions whose axis coincides with the y-axis corresponds to each value of m_2.

In directions satisfying condition (3.49) and (3.50) simultaneously, mutual amplification of the oscillations from sources in the same plane perpendicular to the z-axis occur. The directions of the intensity maxima produced lie along he lines of intersection of the direction cones, of which one is determined by condition (3.49), and the second one by condition (3.50).

Finally, for the train parallel to z-axis , the directions of the maxima are determined by the condition

$$d_3(\cos\gamma - \cos\gamma_0) = \pm m_3\lambda; (m_3 = 0,1,2,...) \quad (3.51)$$

Here d_3 = period of the structure in the direction of the z-axis,

γ_0=angle between the incident beam and the z-axis,

γ=angle between the z-axis and the directions along which diffraction maxima are obtained.

As in the preceding cases, a cone of directions whose axis coincides with the z-axis corresponds to each value of m_3.

In the directions satisfying conditions (3.49), (3.50), and (3.51) simultaneously, mutual amplification of the oscillations from all the elements forming the three-dimensional structure occurs. As a result, diffraction maxima are produced by the three-dimensional structure. The directions of these maxima are on the lines of intersection of threecones whose axes are parallel to the coordinate axes. The conditions (3.49), (3.50), and (3.51) are called **Laue's formulas.**

The angles α,β and γ are not independent:

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1 \qquad (3.52)$$

We have not treated the question of how rays travelling from different structural elements are made to converge to one point on a screen. A lens does this for visible light. A lens can not be used for X-rays because the refractive index of these rays in all substances is virtually equal to unity. For this reason, the interference of the secondary wavelets is achieved by using very narrow beams of rays producing spots of a very small size on a screen (or a photographic plate) even without a lens.

The Russian scientist Yuri Vulf (1863-1925) and the British physicists William Henry Bragg (1862-1942) and his son William Lawrence Bragg (1890-1971) showed independently of each other that the diffraction pattern from a crystal lattice can be calculated in the following simple way. Let us draw parallel equispaced planes through the points of a crystal lattice (Fig.3.16). We shall call these planes **atomic layers**. If the wave falling on the crystal is plane, the envelope of the secondary waves set up by the atoms in such a layer will also be a plane. Thus, the summary action of the atoms in one layer can be represented in the form of a plane wave reflected from an atom-covered surface according to the usual law of reflection.

The plane secondary wavelets reflected from different atomic layers are coherent and will interfere with one another like the waves emitted in the given direction by different slits of a diffraction grating. As in the case of a grating, the secondary wavelets will virtually destroy one another in all directions except those for which the path difference between adjacent wavelets is a multiply of λ. Inspection of Fig.3.16 shows that the difference between the path of two waves reflected from adjacent atomic layers is 2dsinθ, where d is the period of identity of the crystal in a direction at right angles to the layers being considered, and θ is angle supplementing the angle of incidence and called the **glancing angle** of the incident rays. Consequently the directions in which diffraction maxima are obtained are determined by the condition

$$2d\sin\vartheta = \pm m\lambda; (m = 1,2,3,...) \qquad (3.53)$$

This expression is known as the **Bragg-Wulf formula**

We should note that calculations by the Bragg-Wulf formula and by Laue's formulas lead to coinciding results.

The diffraction of X-rays from crystals has two principal applications.It is used to investigate the spectral composition of X-radiation (**X-ray spectroscopy)** and to study the structure of crystals **(X-ray structure analysis).**

3.6 Polarization of Light

We remind our reader that light is called polarized if the direction of oscillatons of the light vector in it are brought into order in some way or other. The simplest kind of polarization is known to be the **plane polarization** i.e. when the light vector oscillates in a fixed direction.

The plane in which the light vector oscillates in a plane polarized wave will be called the **plane f oscillations.** For historical reasons, the term **plane of polarization** was applied not to the plane in which the vector **E** oscillates , but to the plane perpendicular to it.

We can obtain more complicated types of polarization by summation of several plane-polarized waves. For example, two plane polarized waves whose planes of polarization are mutually perpendicular produce an elliptically polarized light (Fig.3.17).

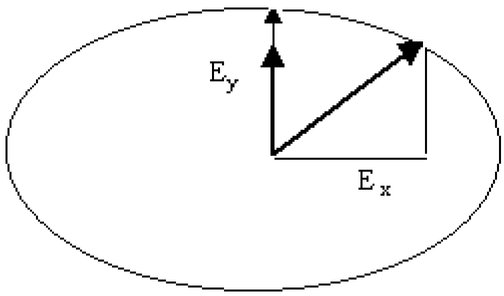


Fig.3.17

If the amplitudes of the waves are equal the ellipse transforms into a circle – circularly polarized light is produced.

Plane-polarized light can be obtained from natural light with the aid of devices called **polarizers.** These devices freely transmit oscillations parallel to the plane which we shall call the **polarizer plane** and completely or partly retain the oscillations perpendicular to this plane. We shall apply the adjective **imperfect** to a polarizer that only partly retain oscillations perpendicular to its plane. We shall apply the term "polarizer" for brevity to a perfect polarizer that completely retains the oscillations perpendicular to its plane and does not weaken the oscillations parallel to its plane.

Light produced at the outlet from an imperfect polarizer in which the oscillations in one direction predominate over the oscillations in other directions is called **partly polarized.** It can be considered as a mixture of natural and plane-polarized light. Partly polarized light, like natural light, can be represented in the form of a superposition of two incoherent plane-polarized waves with mutually perpendicular planes of oscillations. The difference is that for natural light, the intensity of these waves is the same, and for partly polarized light it is different.

If we pass partly polarized light through a polarizer, then when the device rotates about the direction of the ray , the intensity of the transmitted light will change within the limits from I_{max} to I_{min}. The transition from one of these values to the other one will occur upon rotation through an angle of $\pi/2$ (during one complete revolution both the maximum and the minimum intensity will be reached twice). The expression

$$P = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \qquad (3.54)$$

is known as the **degree of polarization.** For plane-polarized light, $I_{min} = 0$, and $P = 1$. For natural light, $I_{max} = I_{min}$, and $P = 0$.
The concept of the degree of polarization can not be applied to elliptically polarized light (in such light the oscillations are completely ordered, so that the degree of polarizaion always equals unity).
An oscillation of amplitude A occurring in a plane making the angle φ with the polarizer plane can be resolved into two oscillations having the amplitudes $A_{||} = A \cos \varphi$ and $A_{\perp} = A \sin \varphi$ (Fig.3.18; the ray is perpendicular to the plane of the drawing).

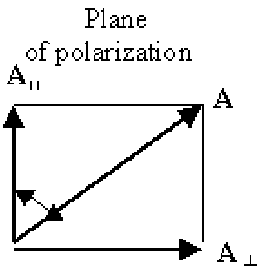


Fig.3.18

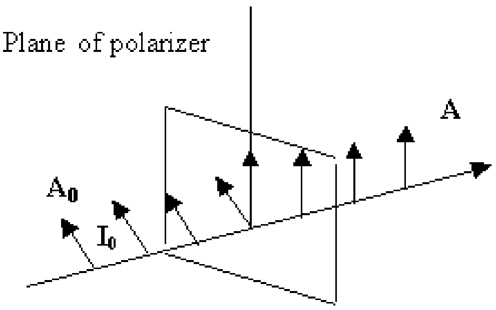


Fig.3.19

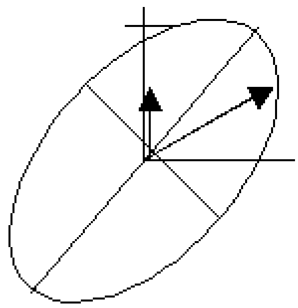
The first oscillation will pass through device, the second will be retained. The intensity of transmitted wave is proportional to $A_{||}^2 = A^2 \cos^2\varphi$, i.e. is $I \cos^2\varphi$ where I is the intensity of the oscillation of amplitude A. Consequently, an oscillation parallel to the plane of the polarizer carries along a fraction of the intensity equal to $\cos^2\varphi$. In natural light, all the values of φ are equally probable. Therefore, the fraction of the light transmitted through the polarizer will equal the average value of $\cos^2 \varphi$, i.e. one-half. When the polarizer is rotated about the direction of a natural ray, the intensity of the transmitted light remains the same. What changes is only the orientation of the plane of oscillations of the light leaving the device.
Assume that plane-polarized light of amplitude A_0 and intensity I_0 falls on a polarizer (Fig.3.20). The component of the oscillation having the amplitude $A = A_0 \cos \varphi$, where φ is the angle between the plane of oscillation of the incident light and the plane of polarizer, will pass through device. Hence, the intensity of the transmitted light I is determined by the expression

$$I = I_0 \cos^2 \varphi \qquad (3.55)$$

Relation (3.55) is known as **Malus's law.** It was first formulated by the French physicist Etienne Malus (1755-1812).
Let us put two polarizers whose planes make the angle φ in the path of a natural ray. Plane-polarized light whose intensity I_0 is half that of natural light I_{nat} will emerge from the first polarizer. According to Malus's law, light having an intensity of $I_0 \cos^2 \varphi$ will emerge from the second polarizer. The intensity of the light transmitted through both polarizers is

$$I = \frac{1}{2} I_{nat} \cos^2 \varphi \qquad (3.56)$$

The maximum intensity equal to (1/2)I_{nat} is obtained at φ = 0 (the polarizers are parallel). At φ = π/2, the intensity is zero – crossed polarizers transmit no light.
Assume that elliptically polarized light falls on a polarizer. The device transmits the component $E_{\Uparrow}$ of the vector **E** in the direction of the plane of polarizer (Fig.3.20).



The maximum value of this component is reached at points 1 and 2. Hence, the amplitude of the plane polarized light leaving the device equals the length of 01'. Rotating the polarizer arund the direction of the ray, we shall bserve changes in the intensity ranging from I_{max} (obtained when the plane of polarizer coincides with the semimajor axis of the ellipse) to I_{min} (obtained when the plane of the polarizer coincides with the semiminor axis of the ellipse). The intensity of light for partly polarized light will change in the same way upon rotation of the polarizer. For circularly polarized light, rotation of the polarizer is not attended (as for natural light) by a change in the intensity of the light transmitted through the device.

Fig. 3.20

Polarization in Reflection and Refraction. If the angle of incidence of light on interface between two dielectrics (for example, on the surface of a glass plate) differs from zero, the reflected and refracted rays will be partly polarized. Oscillations perpendicular to the plane of incidence predominate in the reflected ray (in Fig.3.22 these oscillations are denoted by points(, and oscillations parallel to the plane of incidence predominate in the refracted ray (the are depicted in the figure by double-headed arrows). The degree of polarization depends on the angle of incidence.

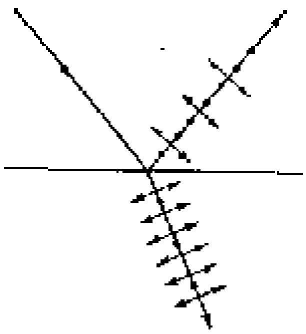


Fig.3.21

Let θ_{Br} stand for the angle satisfying the condition

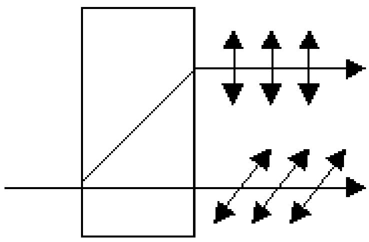
$$\tan \theta_{Br} = n_{12} \qquad (3.57)$$

(n_{12} is the refractive index of the second medium relative to the fist one). At a angle of incidence θ_1 equal to θ_{Br}, the reflected ray is completely polarized (it contains only oscillations perpendicular to the plane of incidence). The degree of polarization of the refracted ray at an angle of incidence equal to θ_{Br} reaches is maximum value, but this ray remains polarized only partly.

Relation (3.57) is known as **Brewster's law,** in honour of its discoverer, the British physisist David Brewster (1781-1868), and the angle θ_{Br} is called **Brewster's angle**. It is easy to see that when light falls at Brewster's angle, the reflected and refracted rays are mutually perpendicular.
The degree of polarization of the reflected and refracted rays for different angles of incidence can be obtained with the aid of **Fresnel formulas** which establish the relation between the amplitudes of the incident, reflected, and refracted waves. We shall not discuss this question.

37 **Polarization in Double Refraction.** When light passes through all transparent crystals except for those belonging to the cubic system, a phenomenon is observed called **double refraction.** It consists in that a ray falling on a crystal is split inside the latter into two rays propagating, generally speaking, with different velocities and in different directions.

38 Doubly refracting crystals are divided into **uniaxial** and **biaxial** ones. In uniaxial crystals, one of the refracted rays obeys the conventional law of refraction, in particular 9t is i the same plane as the incident ray and a normal to the refracting surface. This ray is called an **ordinary ray** and is designated by the symbol o. For the other ray, called **extraordinary ray** (designated by e), the ratio of the sinus of the angle of incidence and the angle of refraction does not remain constant when the angle of incidence varies. Even upon normal incidence of light on a crystal, an extraordinary ray, generally speaking, deviates from a normal (Fig.3.22). In biaxial crystals, both rays are extraordinary.



Uniaxial crystals have a direction along which ordinary and extraordinary rays propagate without separation and with the same velocity. This direction is known as the **optical axis** of the crystal. A plane passing through an optical axis is called a **principal section** or a **principal plane.**

39 Fig.3.22

40 Investigation of the ordinary and extraordinary rays shows that they are both completely polarized in mutually perpendicular directions (see Fig.3.22). The plane of oscillations of the ordinary ray is perpendicular to a principal section of the crystal. In the extraordinary ray, the oscillations of the light vector occur in a plane coinciding with the principal section. When they emerge from the crystal, the two rays differ from each other only in the direction of polarization so that the term “ordinary” and “extraordinary” have a meaning only inside the crystal. Double refraction is explained by the anisotropy of crystals.

Interference of Polarized Rays. When two coherent rays polarized in mutually perpendicular directions are superposed, no interference pattern with the characteristic alternation of maxima and minima of the intensity can be obtained. Interference occur only when the oscillations in the interacting rays occur along the same direction. The oscillations in two rays initially polarized in mutually perpendicular directions can be brought into the plane by passing these rays through a polarizer installed so that its plane does not coincide with the plane of oscillations of any of the rays.

Let us see what happens when an ordinary and extraordinary ray emerging ffrom a crystal plate are superposed. Assume that the plate has been out parallel to an optical axis (Fig.3.23). With normal incidence of the light on the plate, the ordinary and extraordinary rays will propagate without separating, but with different velocities. The following path difference appears between the rays while they pass through the plate:

$$\Delta = (n_0 - n_e)d \qquad (3.58)$$

or the following phase difference:

$$\delta = \frac{(n_0 - n_e)d}{\lambda_0} 2\pi \tag{3.59}$$

(d is the plate thickness, and λ_0 the wavelength in a vacuum).

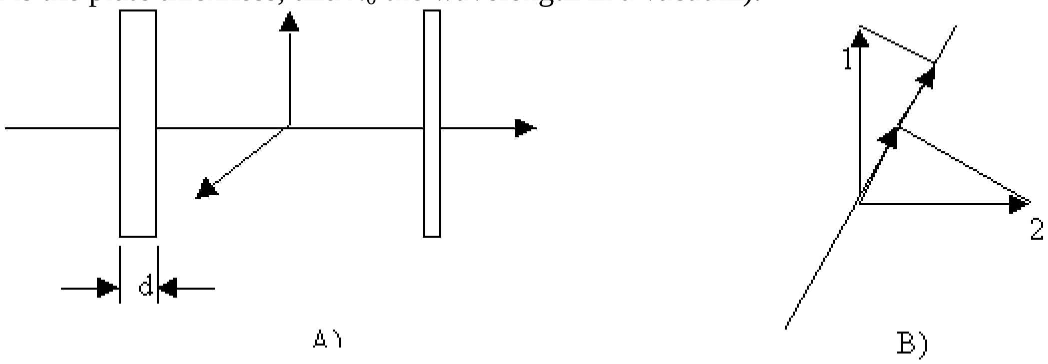


Fig.3.23

Thus, if we pass natural light through a crystal plate cut out parallel t to the optical axis (Fig.3.23a), two rays 1 and 2 that are polarized in mutually perpendicular planes will emerge from the plate, and between them there will be a phase difference determined by Eq. (3.59). Let us place a polarizer in the path of those rays. Both rays after passing through the polarizer will oscillate in one plane. Their amplitudes will equal the components of the amplitudes of rays 1 and 2 in the direction of the plane of the polarizer (Fig.3.23b). In the crystal, ray 1 was extraordinary and could be designated by the symbol (e), and ray 2 was ordinary (o). Upon emerging from the crystal, these rays lost their right to be called ordinary and extraordinary.

The rays emerging from the polarizer are produced as a result of division of the light obtained from a single source. Therefore, they ought to interfere. If rays 1 and 2 are produced as a result of natural light passing through the plate, however, the do not interfere. The explanation is very simple. Although the ordinary and extraordinary rays are produced by the same light source, they contain many oscillations belonging to different wave trains emitted by individual atoms. The oscillations in the ordinary ray are predominantly due to the trains whose oscillation planes are close to one direction in space, whereas those of the extraordinary ray are due to trains whose oscillation planes are close to another direction perpendicular to the first one. Since the individual trains are incoherent, the ordinary and extraordinary rays produced from natural light, and, consequently, rays 1 and 2 too, are also incoherent.

Matters are different if plane-polarized light falls on a crystal plate. In this case, the oscillations of each train are divided between the ordinary and extraordinary rays in some proportion. Consequently, rays (o) and (e), and therefore rays 1and 2 too, will be coherent and will interfere.

Rotation of Polarization Plane. Some substances known as **optically active** ones have the ability of causing rotation of the plane of polarization of plane-polarized light passing through them. Such substances include crystalline bodies (for example, quartz, cinnabar), pure liquids (turpentine, nicotine), and solutions of optically active substances in inactive solvents aqueous solutions of sugar, tartaric acid, etc.).

Crystalline substances rotate the plane of polarization to the greatest extent when the light propagates along the optical axis of the crystal. The angle of rotation φ is proportional to the path l travelled by a ray in the crystal:

$$\varphi = \alpha l \varphi \qquad (3.60)$$

The coefficient α is called the **rotational constant.** It depends on the wavelength.
In solutions, the angle of rotation of the plane of polarization is proportional to the path of the light in the solution and to the concentration of the active substance c:

$$\varphi = [\alpha] c l \qquad (3.61)$$

Here, [α] is a quantity called the **specific rotational constant.**
Depending on the direction of rotation of the polarization plane, optically active substances are divided into **right-hand** and **left-hand ones.** All optically active substances exists in two varieties – right-hand and left-hand. There exists right-hand and left-hand quartz, right-hand and left-hand sugar, etc.
Optically inactive substances acquire the ability of rotating the plane of polarization under action of magnetic field. This phenomenon was discovered by Michael Faraday and is therefore sometimes called the **Faraday effect**. It is observed only when light propagates along the direction of magnetization. The angle of rotation of the polarization plane φ is proportional to the distance l traveled by the light in the substance and to the magnetization of the latter. The magnetization, in turn, is proportional to the magnetic field strength H. We can therefore write that

$$\varphi = VlH \qquad (3.62)$$

The coefficient V is known as the **Verdet constant** or the **specific magnetic rotation.** The constant V, like the rotational constant α, depends on the wavelength.
Optically active substances when acted upon by a magnetic field acquire an additional ability of rotating the plane of polarization that is added to their natural ability.
Kerr effect. The appearance of double refraction in liquids and amorphous solids under action of an electric field was discovered by the Scotch physicist John Kerr (1824-1907) in 1875. This effect was named the **Kerr effect** after its discoverer. In 1930, it was also observed in gases.
An arrangement for studying Kerr effect consists of a **Kerr cell** placed between crossed polarizers P and P'. A Kerr cell is a sealed vessel containing a liquid into which capacitor plates have been introduced. When a voltage is applied across the plates, a virtually homogeneous electric field is set up between them. Under its action, the liquid acquires the properties of a uniaxial crystal with an optical axis oriented along the field.
The resulting difference between the refractive index n_o and n_e is proportional to the square of the field strength E:

$$n_o - n_e = kE^2 \qquad (3.63)$$

The path difference
$\Delta = (n_o - n_e)l = klF^2$
appears between the ordinary and extraordinary rays along the path l. The corresponding phase difference is

$$\delta = 2\pi \frac{\Delta}{\lambda_o} = 2\pi \frac{k}{\lambda_0} lE^2$$

The latter expression is conventionally written in the form

$$\delta = 2\pi B l E^2 \qquad (3.64)$$

where B is a quantity characteristic of a given substance and known as the **Kerr constant.**

The Kerr constant depends on he temperature of a substance and on the wavelengthof the light. Among known liquids, nitrobenzene ($C_6H_5NO_2$) has the highest Kerr constant. The Kerr effect is explained by the different polarzation of molecules in various directions. In the absence of a field, the molecules are oriented chaotically, therefore a liquid as a whole displays no anisotropy. Under the action of a field,, the molecules turn so that either their electric dipole moments (in polar molecules) or their directions of maximum polarization (in non-polar molecules) are oriented in the direction of the field. As a result, the liquid becomes optically anisotropic. The thermal motion of the molecules counteracts the orienting action of the field. This explains the reduction in the Kerr constant with elevation of the temperature.

The time during which the prevailing orientation of the molecules sets in (when the field is switched on) or vanishes (when the field is switched off) is about 10^{-10} s. Therefore, a Kerr cell placed between crossed polarizers can be used as a virtually inertialess light shutter. In the absence of a voltage across the capacitor plates, the shutter will be closed. When the voltage is switched on, the shutter transmits a considerable part of the light falling on the first polaizer.

3.7 **Holography**

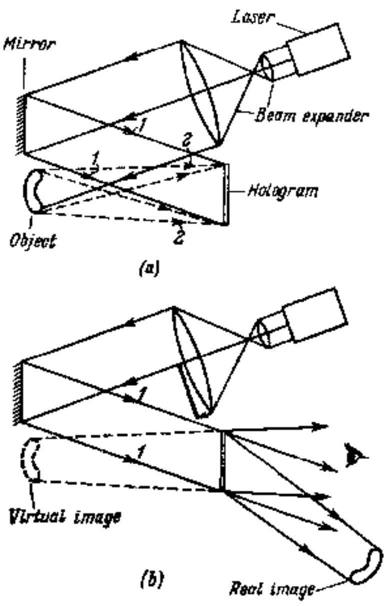


Fig.3.24

Holography (i.e. “complete recording”, from the Greek “holos” meaning “the whole” and “grapho” – “write”) is a special way of recording the structure of the light way reflected by an object on a photographic plate. When this plate (a hologram) is illuminated with a beam of light, the wave recorded on it is reconstructed in practically its original form, so that when the eye perceives the reconstructed wave, the visual sensation is virtually the same as it would be if the object itself were observed.

Holography was invented in 1947 by the British physicist Dennis Gabor. The complete embodiment of Gabor’s idea became possible, however only after the appearance in 1960 of light sources having a high degree of coherence – lasers.

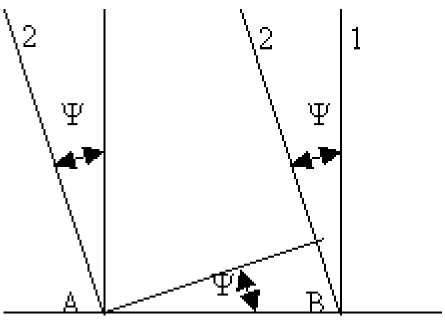
We shall limit ourselves to an elementary consideration of the method of recording holograms on a thin - layer emulsion. Figure 3.24a contains a schematic view of an arrangement for recording holograms, and Fig.3.24b – a schematic view of reconstruction of the image. The light beam emitted by the laser, expanded by the system of lenses, is split into two parts. One part forming the so-called reference wave 1is reflected by the mirror to the photographic plate. The second part reaches the plate after being reflected from the object; it forms object beam 2. Both beams must be coherent. This requirement is satisfied because laser radiation has a high degree of spatial coherence (the light oscillations are coherent over the entire cross section of a laser beam. The reference and object beams superpose and form an interference pattern that is recorded by the
photographic plate. A plate exposed in this way and developed is a **hologram.** Two beams of light participate in forming the hologram. In this connection, the arrangement described above is called two-beam or split-beam holography.

To reconstruct the image, the developed photographic plate is put in the same place where it was recording the hologram, and is illuminated with the reference beam of light (the part of the laser beam that illuminated the object in recording the hologram is now stopped). The reference beam diffracts on the hologram, and as a result a wave is produced having exactly the same structure as the one reflected y the object. This wave produces the virtual image of the object that is seen by the observer. In addition to the wave forming the virtual image, another wave is produced that gives the real image of the object. This image is pseudoscopic; this means that it has a relief which is the opposite of the relief of the object – the convex spots are replaced by concave ones, and vice versa.

Let us consider the nature of hologram and the process of image reconstruction. Assume that two parallel beams of light rays fall on the photographic plate, with the angle □ between the beams (Fig.3.25). Beam 1 is a reference one, and beam 2, the object one. (The object in the given case is an infinitely remote point). We shall assume for simplicity that beam 1 is normal to the plate. All he results obtained below also hold when the reference beam falls at an angle, but the formulas will be more cumbersome.

Owing to the interference of the reference and object beams, a system of alternating straight maxima and minima of the intensity is formed on the plate. Let points A and B correspond to the midle of adjacent interference maxima. Hence, the path difference Δ equals λ.Examination of Fig.3.25 shows that Δ = d sinφ;(($d = AB$).Hence,

$$d \sin \phi = \lambda \qquad (3.65)$$



Having recorded the interference pattern on the plate (by exposure and developing), we direct reference beam 1 at it. For this beam, the plate plays the part of a diffraction grating whose period grating is the circumstance that its transmittance changes in a direction perpendicular to the “lines” according to a cosine law. The result of this feature is that the intensity of all the diffraction maxima of orders higher than the first one virtually equals 0.

When the plate is illuminated with the reference beam (Fig.3.26), a diffraction pattern appears whose maxima form the angles φ with a normal to the plate. These angles are determined by the condition

$$d \sin\varphi = m\lambda \qquad (m=0, \pm 1)$$

The maximum corresponding to m = 0 is on the continuation of the reference beam. The maximum corresponding to m = +1 has the same direction as object beam 2 did during the exposure. In addition, a maximum corresponding to m = -1 appears.
It can be shown that the result we obtained also holds when object beam 2 consists of diverging rays instead of parallel ones. The maximum corresponding to m = +1 has the nature of diverging beam of rays 2'(it produces a virtual image of the point from which rays 2 emerged during the exposure); the maximum corresponding to m = -1, on the other hand, has the nature of a converging beam of rays 2" (it forms a real image of the point which rays 2 emerged from during the exposure).

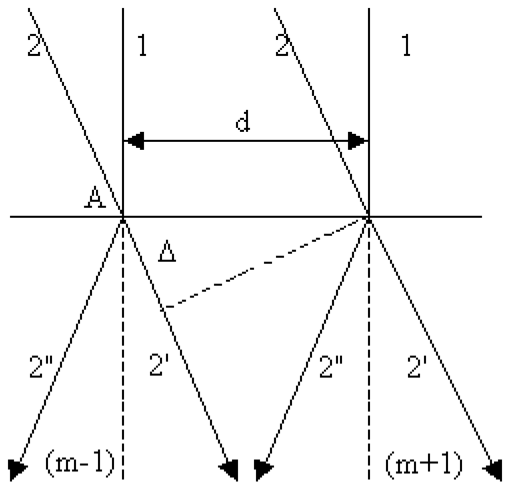


Fig.3.26

In recording the hologram, the plate is illuminated by reference beam 1and numerous diverging beams 2 reflected by different points of object. An intricate interference pattern is formed on the plate as a result of superposition of the patterns produced by each of the beams 2 separately. When the hologram is illuminated with reference beam 1, all beams 2 are reconstructed, i.e. the complete light wave reflected by the object (m= +1) corresponds to it. Two other waves appear in addition to it (corresponding to m=0 and m=-1). But these waves propagate in other directions and do not hinder the perceptions of the wave producing a virtual image of the object.

The image of an object produced by a hologram is three-dimensional. It can be viewed from different positions. If in recording a hologram closest objects concealed more remote ones, then by moving to a side we can look behind the closer object (more exactly, behind its image) and see the objects that had been concealed previously. The explanation is that when moving to a side, we see the image reconstructed from the peripheral part of the hologram onto which the rays reflected from the concealed objects also fell during the exposure. When looking at the images of close and far objects, we have to accommodate our eyes as when looking at the objects themselves.
If a hologram is broken in several pieces, then each of them when illuminated will produce the same picture as the original hologram. But the smaller the part of the hologram used to reconstruct the image, the lower is its sharpness. This is easy to understand by taking into account that when the number of lines of a diffraction grating

is reduced, its resolving power decreases. The application of holography are very diverse. A far from complete list of them includes holographic motion pictures and television, holographic microscopes, and control of the quality of processing articles. The statement can be encountered in publications on the subject that holography can be compared in publications on the subject that holography can be compared as regards its consequences with the setting up of radio communication.

Part II. Modern physics

4. Introduction to Quantum Physics

4.1 Blackbody Radiation and Planck's Hypothesis

In this section, we are going to discuss some phenomena, which can not be explained by the classical physics. Radiation of the black body is just one of these phenomena. Radiation of the black body is an electromagnetic radiation, which is in the state of equilibrium with the bodies, which emit and at the same time absorb it. A cavity (Fig.4.1) which walls are at the temperature T is a very good model of the black body. At any point in the cavity, the density of the energy flux within a spatial angle dΩ

$$dj = \frac{cu}{4\pi} d\Omega \qquad (4.1)$$

Here, u = energy density, c = the speed of light in a vacuum.
An elementary surface ΔS of the cavity (Fig.4.2) sends (within spatial angle $d\Omega = \sin\vartheta d\vartheta d\varphi$) an energy flux

$$d\Phi = dj\Delta S\cos\vartheta = \frac{cu}{4\pi} d\Omega\Delta S\cos\vartheta = \frac{cu}{4\pi}\Delta S\cos\vartheta\sin\vartheta d\vartheta d\varphi \qquad (4.2)$$

The total energy flux emitted by the surface element ΔS is

$$\Delta\varphi = \int d\Phi = \frac{cu}{4\pi}\int_{o}^{\pi/2}(\cos\vartheta\sin\vartheta)d\vartheta\int_{0}^{2\pi} d\varphi = \frac{c}{4}u\Delta S \qquad (4.3)$$

We can also represent 👎☞ by multiplication the magnitude of the **emittance** R by 👎ΔS, i.e. $\Delta\Phi$ = R ΔS

$R = \int_{o}^{\infty} f(\omega, T)d\omega$, Here f ($\omega$, T) is the **spectral emittance.** Thus,

$$R = \frac{c}{4}u \qquad (4.4)$$

Equation (4.4) holds for every frequency, so we get

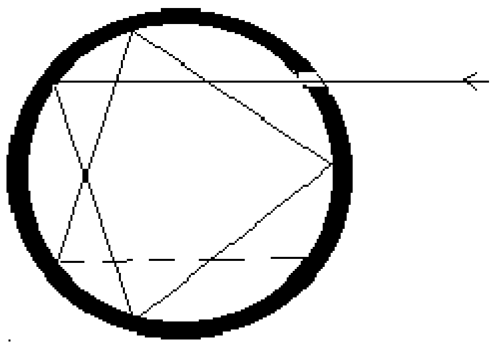


Fig.4.1

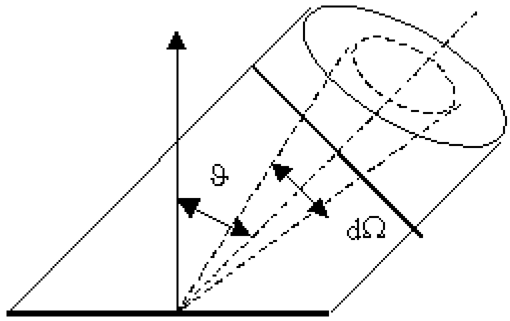


Fig.4.2

$$f(\omega,T) = \frac{c}{4}u(\omega,T) \qquad (4.5)$$

We remind our reader that $u = \int_0^{\infty} u(\omega,T)d\omega$

The expression (4.5) gives relation between the spectral emittance of the black body f(ω,T) and equilibrium density of energy of the heat radiation u(ω,T).
At the beginning of the twentieth century, the function u(ω,T) was found experimentally and its form is shown in Fig.4.3

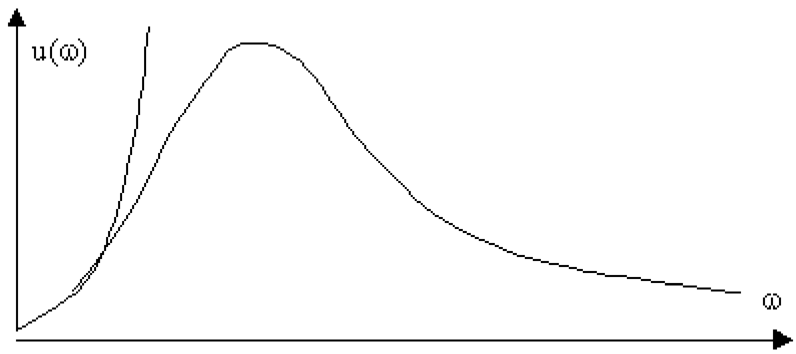


Fig.4.3

All the attempts to deduce the function u(ω,T) with the aid of the classical physics were in vain .The classical expression coincides with the experimental curve only for the small frequencies. It is so called the **Rayleigh-Jeans formula:**

$$u(\omega,T)) = \frac{\omega^2}{\pi^2 c^3} kT \tag{4.6}$$

Here k = Bolzmann's constant.
It is easy to see that integration of the Eq.(4.6) within the limits from 0 to ∞ gives for the magnitude of the energy density u unlimited value and is of no physical meaning. This result was historically called the **ultraviolet catastrophe.**
From the classical point of view, the deduction of the Rayleigh-Jeans formula was faultless. This manifested some deep phenomena, which contradicted with classical physics.
In 1900 Max Planck theoretically found the function u(ω,T) which coincided with the experimental data. In order to deduce it, he had to suggest that the electromagnetic radiation was emitted and absorbed by separate portions of energy (or **quanta**) the magnitude of which is proportional to the frequency of radiation

$$E = \hbar\omega \tag{4.7}$$

The constant ħ was called afterwards the **Planck constant** ($\hbar = 1.054 \cdot 10^{-34}$ Joule ·s). (In reality, Planck used the quantity h = 2πħ)
If the radiation us emitted by portions ·ω, the energy ε_n should be integer of this quantity i.e.

$$\varepsilon_n = n\hbar\omega \tag{4.8}$$

In the state of equilibrium, the distribution of oscillations must be in accordance with Bolzmann,s law. According to this law, the probability P_n for the energy of oscillation of the frequency ω to have the magnitude ε_n is given by the expression:

$$P_n = \frac{N_n}{N} = \frac{e^{-\varepsilon_n/kT}}{\sum_n e^{-\varepsilon_n/kT}} \tag{4.9}$$

The average energy is given by

$$\langle \varepsilon \rangle = \sum_n P_n \varepsilon_n \tag{4.10}$$

Using Eqs.(4.8) and (4.9) we can write:

$$\langle \varepsilon \rangle = \frac{\sum_{n=0}^{\infty} n\hbar\omega e^{-n\hbar\omega/kT}}{\sum_{n=0}^{\infty} e^{-n\hbar\omega/kT}} \tag{4.11}$$

Designating $\hbar\omega / kT = x$ we can express Eq.(4.11) in the form

$$\langle \varepsilon \rangle = \hbar\omega \frac{\sum_{n=0}^{\infty} n e^{-nx}}{\sum_{n=0}^{\infty} e^{-nx}} = -\hbar\omega \frac{d}{dx} \ln \sum_{n=0}^{\infty} e^{-nx} \tag{4.12}$$

Obviously,

$$\sum_{n=o}^{\infty} e^{-nx} = \frac{1}{1-e^{-x}} \tag{4.13}$$

Thus,we get:

$$\langle \varepsilon \rangle = -\hbar\omega \frac{d}{dx} \ln \frac{1}{1-e^{-x}} = \hbar\omega \frac{e^{x}}{1-e^{-x}} = \frac{\hbar\omega}{e^{x}-1} = \frac{\hbar\omega}{e^{\hbar\omega/kT}-1} \tag{4.14}$$

The average number of quanta having the frequency ω

$$\langle n \rangle = \frac{1}{e^{\hbar\omega/kT}-1} \tag{4.15}$$

The number of levels within the frequency interval from ω up to ω+dω is given by the well-known formula

$$dn_{\omega} = \frac{\omega^2 d\omega}{\pi^2 c^3} \tag{4.16}$$

(In appendix of this section, the derivation of the expression (4.16) is given.
Multiplying Eqs.(4.13) and (4.16) we get the expression for the density of energy:

$$u(\omega, T) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{e^{\hbar\omega/kT}-1} \tag{4.17}$$

. Using Eq.(4.5) we arrive at the expression:

$$f(\omega,T)=\frac{\hbar\omega^3}{4\pi^2c^3}\frac{1}{e^{\hbar\omega/kT}-1} \qquad (4.18)$$

Expressions (4.17) and (4.18) are called **Planck's formulas**. The calculations made in accordance with these formulas coincide with experimental data.
Sometimes it is convenient to use the expression for spectral emittance $\varphi(\lambda,T)$ given in terms of λ:

Obviously, $\varphi(\lambda,T)d\lambda = f(\omega,T)d\omega$ (4.19)

With the aid of Eq.(4.19) we easily get

$$\varphi(\lambda,T)=\frac{4\pi^2\hbar c^3}{\lambda^5}\frac{1}{e^{2\pi\hbar c/kT\lambda}-1} \qquad (4.20)$$

In Fig.4.4, the functions f(ω,T and φ(λ,T) for T=5000K are shown. The logarithmically scales are used. A glance on figure shows that the frequencies corresponding to the maxima of intensity do not coincide.

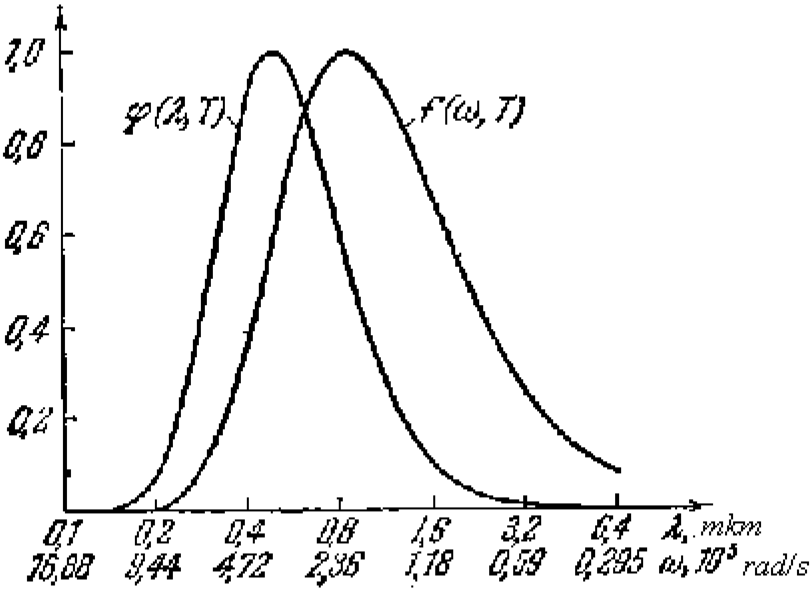


Fig.4.4

We can write the energy emittance of the black body in the following form

$$R=\int_0^\infty f(\omega,\mathrm{T})d\omega=\int_0^\infty \frac{\hbar\omega^3}{4\pi^2c^3}\frac{d\omega}{e^{\hbar\omega/kT}-1}$$

Introducing a new argument $x=\hbar\omega/kT$ we write this expression in the form

$$R=\frac{\hbar}{4\pi^2c^3}\left(\frac{kT}{\hbar}\right)^4\int\frac{x^3dx}{e^x-1}$$

The integral in this expression equals $\pi^4/15 \approx 6.5$. And we have the **Stefan-Bolzmann law:**

$$R = \frac{\pi^2 k^4}{60 c^2 \hbar^3} T^4 = \sigma T^4 \tag{4.21}$$

Using the numerical values of k, c, ħ we get for the **Stefan-Bolzmann constant** $\delta = 5.6696 \cdot 10^{-8}$ W/(m^2 · K^4) which coincides with the experimental value.
In conclusion, let us find the value of the constant in the **Wien displacement law.**

$$T\lambda_m = b \tag{4.22}$$

Here, λ_m is the wavelength corresponding to maximum of function $\varphi(\lambda, T)$. Taking derivative from Eq.(4.20) and assuming this derivative to be zero, i.e.

$$\frac{\partial \varphi}{\partial \lambda} = 0; (\lambda = \lambda_m) \tag{4.23}$$

we arrive at the equation $xe^x - 5(x-1) = 0$. Here $x = 2\pi\hbar c / kT\lambda_m$. Solution of this equation is x = 4.965.Hence,

$$E\lambda_m = \frac{2\pi\hbar c}{4.965k} = b = 2.90 \cdot 10^{-3} m \cdot K \tag{4.24}$$

Thus, Planck's formula describes all aspects associated with the radiation of the black body. Historically the world constant h appeared in the world scientific literature in 1900. This year can be considered as the beginning of new physics – the quantum physics.

41 Appendix I

An electromagnetic field in the cavity (see Fig.4.1) must satisfy the wave equation

$$\nabla^2 \psi - \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} = 0 \tag{4.25}$$

If to chose Ψ to be zero at the boundaries of a cube with the side L, then the solution of Eq.(4.25) can be written as follows:

$$\Psi = \Psi_0 \sin(\omega t - \varphi) \sin(k_x x) \sin(k_y y) \sin(k_z z) \tag{4.26}$$

$$\text{Here } k_x L = \pi l; k_y L = \pi m; k_z L = \pi n. \tag{4.27}$$

Here l, m, n are integer (1, 2, 3, . . .) and ω is connected with k_x, k_y, k_z by the condition

$$k_x^2 + k_y^2 + k_z^2 - \frac{\omega^2}{c^2} = 0 \qquad (4.28)$$

$$\omega = c\sqrt{k_x^2 + k_y^2 + k_z^2} = c\frac{\pi}{L}\sqrt{l^2 + m^2 + n^2} = \omega_{lmn} \qquad (4.29)$$

A general solution of Eq.(4.25) can be represented as the sum of equations (4.26) with all possible magnitudes of l, m, n.

In order to calculate the number of possible partial solutions (abstract oscillators) within the frequency interval dω, let us investigate the space of vector c**k** having components ck_x, ck_y, ck_z and the length ck=ω (Fig.4.5).

Obviously, vectors c**k** having its tails in nudes of a spatial cubical lattice build by elementary cubes with sides cπ/L represent all possible vectors c**k** and $\omega_{lmn.}$

In the space of vector c**k,** let us imagine a spherical surface of radius ω; and designate by N(ω) the number of nodes of a cubical lattice within the limits of the first octant of this space. Obviously, N(ω) is the total number of abstract oscillators having a frequency not greater than ω.

Let us assume that the radius f the sphere is much greater than the side of an elementary cube, than N(ω) is approximately equal to the number of elementary cubes which are inside the octant. Hence, N(ω) can be represented as follows:

$$N(\omega) = \frac{\frac{1}{8}\cdot\frac{4}{3}\pi\omega^3}{\left(c\frac{\pi}{L}\right)^3} \qquad (4.30)$$

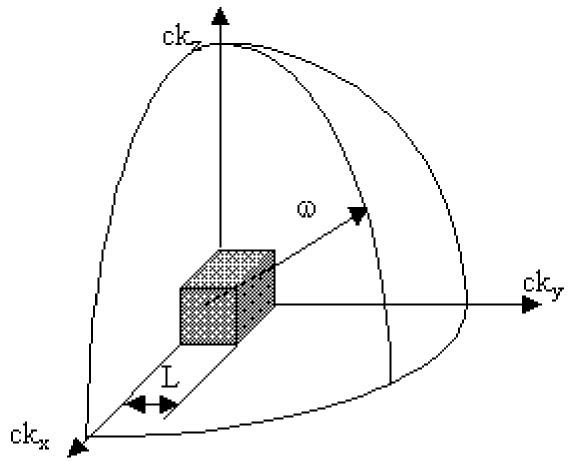


Fig.4.5

Or, having in mind that $L^3 = V$, where V is the volume of a cavity inside the walls at which Ψ=0, we get

$$N(\omega) = \frac{\omega^3 V}{6\pi^2 c^3} \tag{4.31}$$

Having in mind two ways of polarization

$$N(\omega) = \frac{\omega^3 V}{3\pi^2 c^3} \tag{4.32}$$

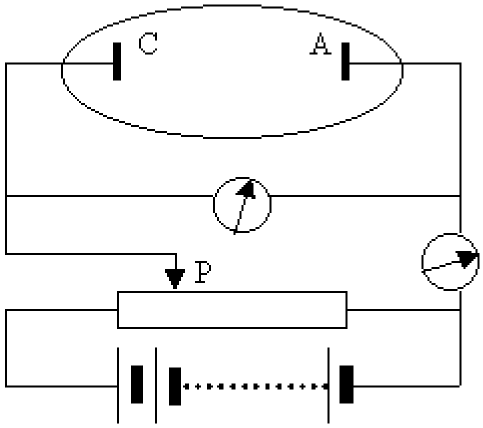
The number of abstract oscillators per unit volume which frequencies are within the frequency interval dω

$$dn_\omega = \frac{1}{V}\frac{\partial N(\omega)}{\partial \omega} d\omega = \frac{\omega^2}{\pi^2 c^3} d\omega$$

and we arrive to Eq.(4.16).

4.2 Photoelectric Effect

The phenomenon associated with emittance of electrons from the surface of metals when irradiated by light is called the **photoelectric effect** or simply **photoeffect.** It was discovered by Hertz in 1887, and investigated systematically the first time by Stoletov and later by Lennard, Thompson, and others.
The typical scheme of experiment is shown in Fig.4.6



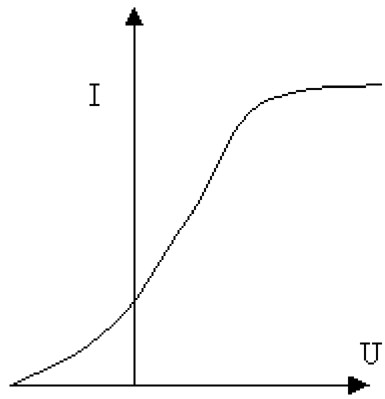


Fig.4.6 Fig.4.7

The light propagating through the quartz window illuminates a cathode. The electrons emitted from cathode C move in the direction of anode A. As a result, an electric current is produced. The voltage applied between cathode and anode can be changed with the aid of potentiometer P. A typical dependence between the electric current and voltage applied through the electrodes is shown in Fig.4.7.

Experiments show that the energy of photoelectrons does not depend on the intensity of light, but depends only on wavelength.

It was established that if the frequency of light is less than a certain quantity (called the **red border**), there is no photoeffect at all. This situation can not be explained by the classical theory of electromagnetic waves. Albert Einstein was the first who explained
Fig.4.6
the laws of pfotoeffect using the conception of quanta developed by Max Planck. According to Einstein, the energy $\hbar\omega$ is needed for an electron to leave the bound state A, (A is usually called the **work of exit)** and to acquire the kinetic energy $mv^2/2$:

$$\hbar\omega = \frac{1}{2}mv^2 + A \tag{4.33}$$

The magnitudes of the energy of exit for some metals are given in the following table.

Metal	Work of exit [eV]	Metal	Work of exit [eV]
Li	2.38	W	4.54
Na	2.35	Pd	4.80
K	2.22	Pt	5.32
Ca	1.81		

The red boundary frequency ω_0 of photoelectric effect can be easily calculated if the work of exit is known

$$A = \hbar\omega_0 = 2\pi\hbar\frac{c}{\lambda_0} \tag{4.34}$$

For K, λ_0 = 0.6μm; and for W, λ_0 = 0.27μm. It can be easily shown that the photoeffect is impossible if the electron is in a free state. Indeed, according to the energy conservation law, we can write

$$\hbar\omega = \frac{mv^2}{2}, \tag{4.35}$$

$$\frac{\hbar\omega}{c} = mv \tag{4.36}$$

Equations (4.35) and (4.36) lead to physically senseless result: v=2c !?
We can try to use the formulas of the special relativity theory:

$$\begin{aligned} \hbar\omega + mc^2 &= \frac{mc^2}{\sqrt{1-\frac{v^2}{c^2}}} \\ \frac{\hbar\omega}{c} &= \frac{mv}{\sqrt{1-\frac{v^2}{c^2}}} \end{aligned} \tag{4.37 a, b}$$

The Eqs.(4.37) again lead to physically senseless result: v=0 or v=c.
The photoelectric effect is widely used in techniques. We are not going to discuss this question. We just wanted to show to the reader that in order to explain the main features of photoelectric effect it is necessary to use the ideas quite different from the classical ones.

4.3 Bothe experiment

In order to explain the energy distribution in the spectrum of black-body radiation it is quite sufficient (it was shown by Planck) to admit that the light is emitted by discreet portions $\cdot\omega$. In order to explain the laws of photoelectric effect it is sufficient to assume that the light is absorbed by the same portions. But A.Einstein had made a more important step, postulating that the light itself propagates as the flux of discreet particles which were called the **light quanta.** Later (in 1926) these particles were called the **photons.**
This idea was confirmed by the experiment made by the German physicist Hans Bothe (Fig. 3.333). The thin metallic plate was put between two gas-discharged counters. The foil was being illuminated by the weak beam of X-rays under action of which it itself became the source of X-rays (this phenomenon is called the roentgen fluorescence). The initial flux of X-quanta was very small, so the number of the second quanta emitted by the foil was also small. When an X-quantum entered the counter, a special mechanism which made a mark upon the moving band was put in action. If the emitted radiation propagated uniformly in all directions (as it follows from the wave concepts), the both counters should act at the same time and the marks on the band should be one against the other. In reality, the chaotic recording of marks was registered. It could be explained only by the fact that the particles generated in the foil were moving in this or that direction.

Thus, the existence of light particles (**photons)** was established. The energy of a photon is given by its frequency

$$E = \hbar\omega \tag{4.38}$$

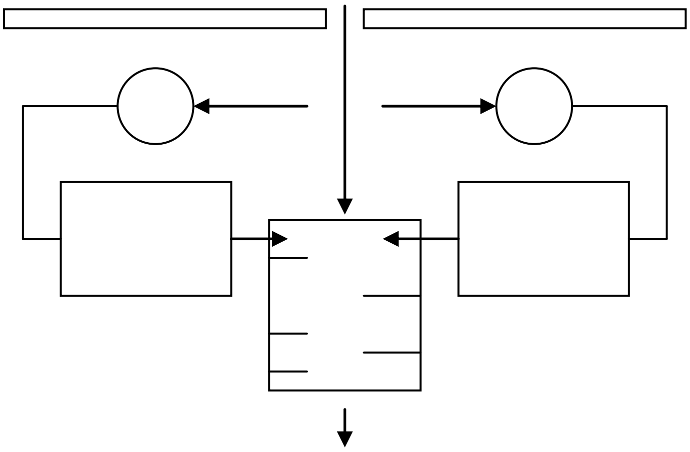


Fig.4.8

We recommend our reader to calculate that the wavelength λ=5000Å (the green region of spectrum) corresponds to the photon's energy $\hbar\omega = 2.5eV$; when λ = 1Å ħω = 12.5 keV.

In order to find the momentum of a photon let us use the relations of Relativity. Let two reference systems K and K' move with the relative velocity $\mathbf{v_0}$. Let a **photon** to move along $\mathbf{v_0}$. The direction of the coordinate axis X and X' coincide with that of velocity. The energy of photon in the system K and K' is correspondingly equal to ħω and ħω'. The frequencies ω and ω' are connected by the relation

$$\omega^1 = \omega \frac{1 - \frac{v_o}{c}}{\sqrt{1 - \frac{v_0^2}{c^2}}} \tag{4.39}$$

Hence,

$$E' = E \frac{1 - \frac{v_0}{c}}{\sqrt{1 - \frac{v_0^2}{c^2}}} \tag{4.40}$$

Let us designate the momentum of the photon in the reference frame K by the symbol **p,** and in the system K' – by the symbol **p'.** From the principle of symmetry, it follows that the momentum of the photon should be directed along the X-axis. Thus, we have: $p_x = p$, $p_x^1 = p'$ **.** The energy and momentum are transformed, in accordance with the formulas of relativity, as follows:

$$E' = E\frac{E - v_0 p_x}{\sqrt{1 - \frac{v_0^2}{c^2}}} \qquad (4.41)$$

Comparing the last two formulas we have:

$$E\left(1 - \frac{v_0}{c}\right) = E - v_0 p \qquad (4.42)$$

Hence

$$p = \frac{E}{c} = \frac{\hbar\omega}{c} \qquad (4.43)$$

Relation (4.43) is true for the particles which have the rest mass equal zero and move with the speed of light in a vacuum.
Taking into account that $\omega = \pi/\lambda$ we get

$$p = \frac{2\pi\hbar}{\lambda} = \hbar k \qquad (4.44)$$

(k is a wave number). A photon moves in the same direction as the electromagnetic wave does. Hence, we can write the previous formula as follows

$$\vec{p} = \hbar\vec{k} \qquad (4.45)$$

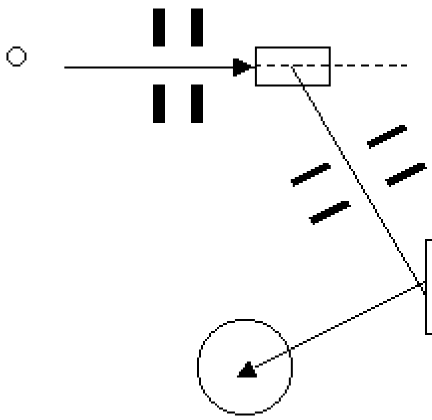


Fig.4.9

4.4 Compton Effect

In 1923 the British physicist A. Compton , studying the scattering of X-rays from some substances, had found that in the scattered rays besides the radiation with the previous wavelength λ, the rays having wavelengths $\lambda' > \lambda$ were present. The difference $\Delta\lambda = \lambda' - \lambda$ is dependent on the angle θ between the previous and secondary directions of propagation. It does not depend on the wavelength and the nature of substance.
The scheme of the experiment is shown in Fig.4.9. A monochromatic narrow beam of X-rays emitted by the source after having passed through collimator falls on the substance under investigation. The scattered beam is being studied with the aid of a X-ray spectrometer consisting of a crystal and ionization chamber.

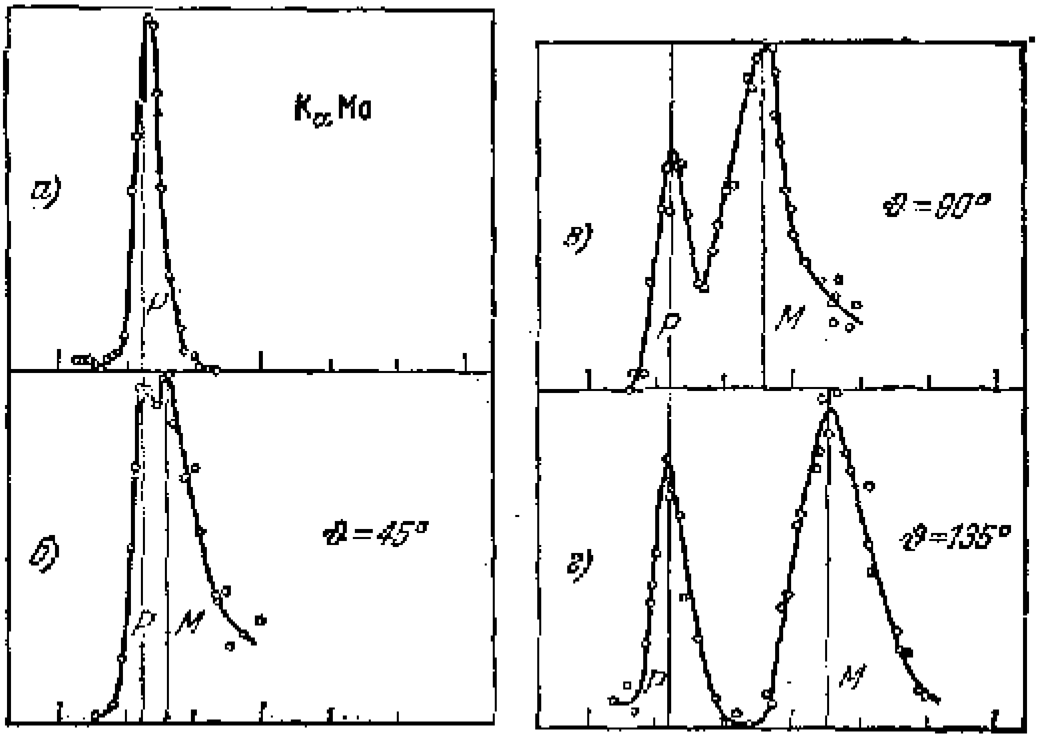


Fig.410

Fig.4.10 shows the results of scattering of monochromaic X-rays (K_α-line of Molibden) from graphite. The curve 'a' represents the initial spectrum. The other curves are the results of scattering at various angles ϑ. The axis X represents the wavelengths and the axis Y represents the intensity of radiation.

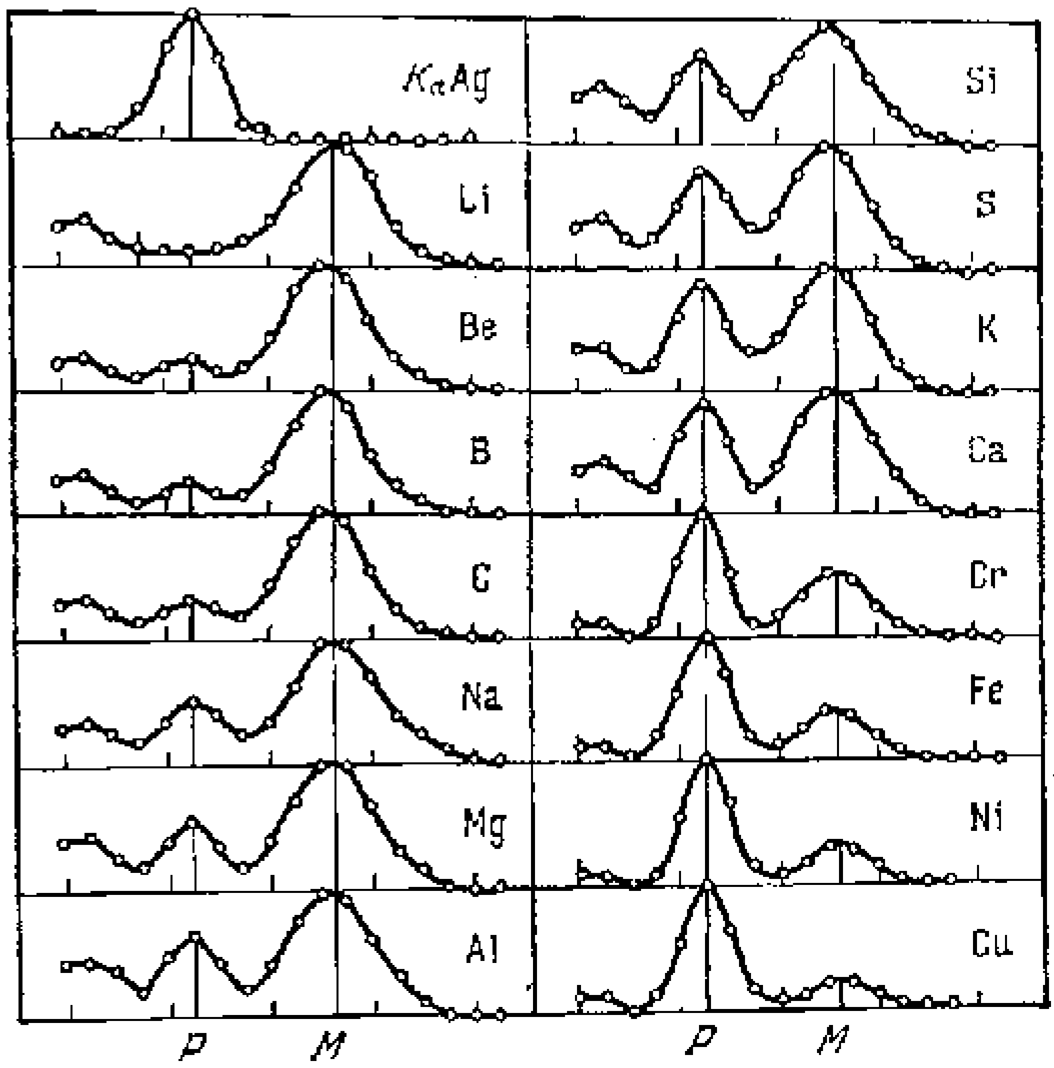


Fig.4.11

Fig.4.11 shows the ratio between the intensities of displaced and nondisplaced components of radiation as function of atomic number of scattering substance. The upper left-hand curve represents the initial spectrum (K-line of Silver). When the atomic number is small (Li, Be, B), the scattered radiation contains the great amount of radiation with displaced wavelength; when the atomic number increases, the quantity of displaced radiation decreases.

All particularities of the Compton effect can be easily explained as nonelastic scattering of X-photons from practically free electrons.

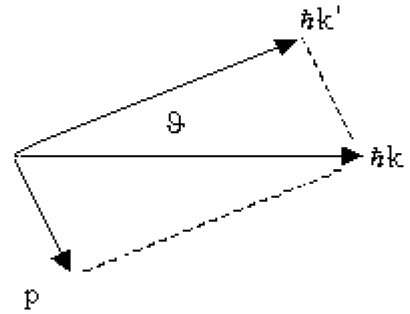


Fig.4.12

Let us assume that a photon with energy ħω and momentum ħω/c falls upon an free electron at rest (Fig.4.12). The initial energy of electron is mc^2 (m is the rest mass of an electron), its momentum is zero. After collision, the electron has the momentum **p** and energy $c\sqrt{p^2+m^2c^2}$. After the collision, the energy and momentum of photon will be ħω and ħk'. The energy and momentum conservation laws lead to equations

$$\hbar\omega + mc^2 = \hbar\omega' + c\sqrt{p^2+m^2c^2} \tag{4.46}$$

$$\hbar\vec{k} = \vec{p} + \hbar\vec{k}' \tag{4.47}$$

Let us divide the first equation by c and rewrite i n the form

$$\sqrt{p^2+m^2c^2} = \hbar(k-k') + mc$$

(◆/c = k). Rising this equation in the second power leads to

$$p^2 = \hbar^2(k^2+k'^2-2kk') + 2\hbar mc(k-k') \tag{4.48}$$

From Eq.(4.40) it follows that

$$p^2 = \hbar^2(k-k')^2 = \hbar^2(k^2+k'^2-2kk'\cos\vartheta) \tag{4.49}$$

(ϑ is the angle between the vectors **k** and **k**'; see Fig.4.12).

Comparing these last two equations we get

$$mc(k-k') = \hbar kk'(1-\cos\vartheta). \tag{4.50}$$

Multiplying this equation by 2π and dividing it by mckk' we arrive at the equation

$$\frac{2\pi}{k'} - \frac{2\pi}{k} = \frac{2\pi\hbar}{mc}(1-\cos\theta) \tag{4.51}$$

Having in mind that that $2\pi/k = \lambda$, we get to the equation

$$\Delta\lambda = \lambda' - \lambda = \lambda_C(1-\cos\vartheta). \qquad (4.52)$$

Here,

$$\lambda_C = \frac{2\pi\hbar}{mc} \qquad (4.53)$$

is called the **Compton wave length** .
When a photon is scattered by electrons which are strongly bound, then the energy and momentum exchange occurs with the atom as a whole. The atom's mass is much greater than the mass of an electron. Thus, the magnitudes of λ and λ' are practically the same ($\Delta\lambda \ll \lambda$). When the atom number grows, the relative quantity of strongly bound electrons is also grows. That leads to weakening of the intensity of the displaced line [see Fig.4.11].

4.5 Atomic Spectra

The electromagnetic radiation of separate atoms consists from separate spectral lines (line spectra). The lines can be united in spectral groups. The typical line spectrum is that of a hydrogen atom. The frequency hydrogen spectrum lines can be represented as follows:

Layman series $\omega = R\left(\frac{1}{1^2} - \frac{1}{n^2}\right)$ (n = 2, 3, 4, . . .), ultraviolet light (4.54)

Balmer series $\omega = R\left(\frac{1}{2^2} - \frac{1}{n^2}\right)$ (n = 3, 4, 5, . . .), visible light (4.55)

Pashen series $\omega = R\left(\frac{1}{3^2} - \frac{1}{n^2}\right)$ (n = 4, 5, 6, . . .), infrared light (4.56)

Bracket series $\omega = R\left(\frac{1}{4^2} - \frac{1}{n^2}\right)$ (n = 5, 6, 7, . . .), infrared light (4.57)

Pfund series $\omega = R\left(\frac{1}{5^2} - \frac{1}{n^2}\right)$ (n = 6, 7, 8, . . .), infrared light (4.58)

Here, R is the **Rydberg constant,** $R = 2.07 \cdot 10^{16}$ rad/s.
Sometimes, it is more convenient to characterize the spectral lines by the inverse wavelength

$$\nu' = \frac{1}{\lambda} = \frac{\omega}{2\pi c} \qquad (4.59)$$

Using this quantity, we can write, for example, the formula for the Balmer series in the form:

$$\nu' = R'\left(\frac{1}{2^2} - \frac{1}{n^2}\right) \tag{4.60}$$

Here R = 109737.309 ± 0.012 cm^{-1}.
Thus, the frequencies of spectral lines for the Hydrogen spectrum can be given as follows

$$\omega = R\left(\frac{1}{m^2} - \frac{1}{n^2}\right) \tag{4.61}$$

where m = 1for the Layman series, m=2 for the Balmer series, and so on. It should be noted that this expression is called the **general Balmer formula.**
When n is being increased ($n \to \infty$), the frequency of the spectral line has its maximal value R/m^2 which is called the **border of the series.**
The quantities

$$T(n) = \frac{R}{n^2}. \text{ I.e. } \frac{R}{1^2}, \frac{R}{2^2}, \frac{R}{3^2}, \ldots \tag{4.62}$$

are called the **spectral terms.** The frequency of any spectral line can be expressed as the difference of two terms. For example, the frequency of the first line of Balmer series is equal T(2) – T(3); the frequency of the second line of Pfund series is equal T(5)-T(7), and so on.
Studying the spectra of other atoms showed that the properties of lines can also be expressed as the difference of two terms. But the terms have more complicated form than those of a hydrogen atom.

4.6 Bohr's Model of Atom

In 1913 the Dutch physicist Niles Bohr postulated principles of quantum model of atom. These principles can be formulated as follows. *The energies of electron in atom are discreet and only those orbits are possible for which the angular momentum $L=n\hbar$ (n=1,2,3,. . .); the energy emitted by transition of an electron from E_n-level to E_m-level is $\hbar\omega_{mn}=E_n-E_m$.*
These two postulates explain properties of a hydrogen atom. Indeed, let us assume that an electron rotates about a proton. We can write the equations

$$mvr = n\hbar \tag{4.63}$$

$$\frac{mv^2}{r} = \frac{e^2}{r^2} \tag{4.64}$$

Here: m = the mass of the electron,
e = the electric charge, r = the radius of orbit, v =velocity.
(We use CGSE units).

From Eqs.(4.55) and (4.56), it follows that the radius of n-th orbit

$$r_n = \frac{\hbar^2}{ml^2} n^2 \tag{4.65}$$

r_1 is called the **Bohr radius,** and is designated by r_0.

$$r_0 = \frac{\hbar^2}{mc^2} = 0.529 A^\circ \tag{4.66}$$

The inner energy of atom is equal

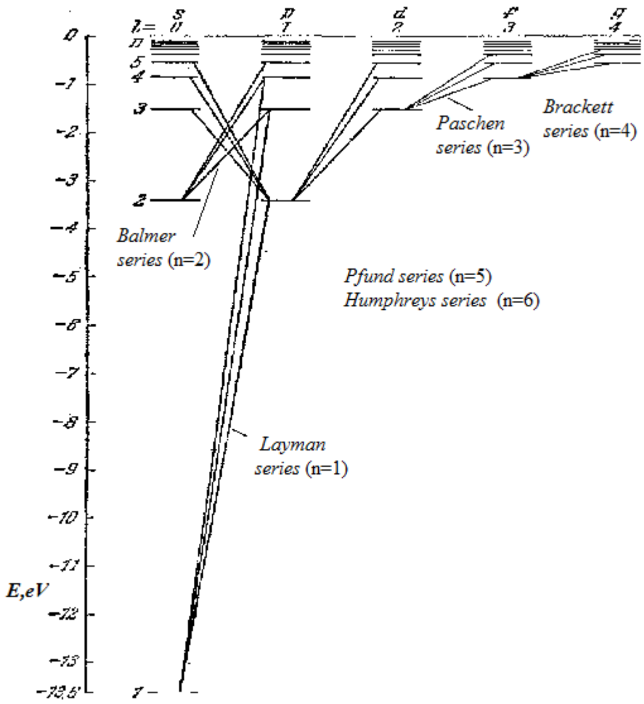
$$E = \frac{mv^2}{2} - \frac{e^2}{r} \tag{4.67}$$

Using Eqs.(4.55) and (4.56) we have

$$\frac{mv^2}{2} = \frac{e^2}{2r}$$
$$E = \frac{e^2}{2r} - \frac{e^2}{r} = -\frac{e^2}{2r} \tag{4.68a,b}$$

Or substituting for r expression (4.57), we can write:

$$E_n = -\frac{me^4}{2\hbar^2} \frac{1}{n^2} \tag{4.69}$$



The scheme of the energy levels is shown in Fig.(4.13).

In accordance with Bohr's postulate, the frequency of emitted photons is given by the expression

$$\omega_{mn} = \frac{me^4}{2\hbar^3}\left(\frac{1}{m^2} - \frac{1}{n^2}\right) \tag{4.70}$$

A glance at the expression (4.61) shows that the Rydberg constant

$$R = \frac{me^4}{2\hbar^3} \tag{4.71}$$

This quantity coincides with the experimental one very closely.

But all attempts to describe other atoms (even the Helium atom) were not successful. Nowadays, the Bohr theory is only of historical interest.