Study Material - Sem. 2 - C4T

- Wave Optics & Interference - Dr. T. Kar

What is optics?

Optics is a branch of Physics in which the nature and properties of light are studied. The term light is commonly used to mean radiant energy which can produce the sensation of vision. However, at present the term light is used to mean all kinds of radiations, both visible and invisible. The subject of optics is conveniently divided into three distinct branches: (a) Geometrical Optics, (b) Physical Optics and (c) Quantum Optics. Each of these branches requires a distinct method of theoretical treatment. In geometrical optics, many fundamental principles concerning light are studied by purely geometrical means without assuming anything regarding the nature of light. It assumes rectilinear propagation of light. In physical optics, many experimental results are explained by considering primarily the wave nature of light. Quantum optics deals with the interaction of light with atomic entities of matter. Here one considers the corpuscular nature of light and for exact analysis it requires the method of quantum mechanics.

Nature of Light

In order to understand different optical phenomena, we are to consider the nature of light. Historically different theories have been proposed regarding the nature of light. In 1675, Newton first proposed the corpuscular theory of light. According to it, light consists of a stream of tiny particles moving with a great speed. Different colours are due to different size of the corpuscles. Though this model could explain many facts in geometrical optics, it failed to explain phenomena like interference, diffraction, polarisation etc. Next came the mechanical wave model of light as proposed by Huygens. Introduction of this model became very successful in many cases. However, Huygens' theory of longitudinal waves could not explain the phenomenon of polarisation of light and the rectilinear propagation of light. To overcome these difficulties, Fresnel introduced the concept of transverse waves and his idea of half period zones.

To explain the propagation of light through vacuum, the supporters of this wave theory of light had to imagine the presence of all-pervading ether medium. But now, from the theory of relativity, we know that ether does not exist.

A completely different conception regarding the nature of light wave was put forwarded by Maxwell. According to him, light is a transverse electromagnetic wave. A changing magnetic field produces a changing electric field and a changing electric field in turn produces a changing magnetic field. This makes the propagation of electromagnetic wave possible even through vacuum. The existence of such electromagnetic waves was verified experimentally by Hertz. This model was very fine upto the end of the nineteenth century. Trouble began with the discovery of photoelectric effect. This effect could not be explained on the basis of wave theory of light. It was explained by Einstein with the help of photon theory of light. According to this theory, light consist of quanta or packets of energy called photons. It may be called energy particles.

Thus we find that there are some phenomena which can be well explained by wave theory of light and there are some phenomena which can only be explained on the basis of particle (photon) nature of light. So we may say that light has dual character — sometimes as waves and sometimes as particle. A satisfactory theory must unify these two ideas into a general coordinating principle. However, a true unification of these two ideas can be obtained through modern quantum theory. Here the distinction between particles and waves vanishes because particles can show wave-like properties and vice-versa. Thus with the present day knowledge, we can say that the wave and particle ideas of light are not rival conceptions but are in reality complementary to each other.

Huygens' Principle and Propagation of Wavefront

A Wavefront is a surface upon which the phase of disturbance is same at any given instant of time. If we have a point source O in a homogeneous isotropic medium, this disturbance will travel with equal velocity in every direction and will, therefore, reach simultaneously to all points lying on the surface of a sphere drawn with the O as centre (Fig. 1). Thus the wavefront is spherical in this case. At long distances it may be

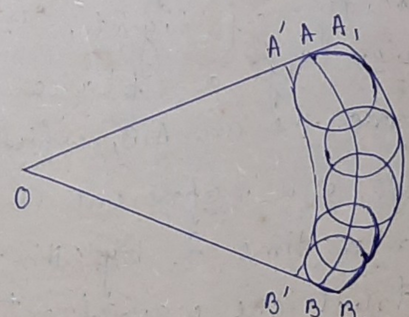


Fig. 1.

considered as a plane wavefront. As the wave advances, the wavefront moves parallel to itself in a direction normal to the wavefront.

If we have a line source we get cylindrical wavefronts. Now there arises one important question. Suppose we know the wavefront AB at any instant of time t. Then how to construct the wavefront at some later time instant $(t + \delta)$. The answer is obtained from Huygens' principle. According to this principle, in a homogeneous and isotropic medium every point of a primary wavefront serves as a source of spherical secondary wavelets that travel with a speed and frequency equal to those of the primary wave. The wavefront at some later instant of time is the envelope of these wavelets.

Thus in fig. 1., each point of the wavefront AB will be the sources of secondary disturbances and secondary wavelets will begin to form with those points as centres. The size of each secondary wavelet will grow with time and after a time δ, it will be a sphere of radius $c\delta$, where c is the velocity of light. The sections of these secondary waves by a plane will be circles as shown in the figure. The arc A_1B_1 which is the front envelope of these circles, will represent the section of the next wavefront after a time δ. By this type of construction, we can construct one wavefront from another.

If the medium is not homogeneous and isotropic, the velocity of propagation will be different at different points and in different directions. In such cases, appropriate velocity must be used at different points and different directions to construct the new wavefront. Also, for correct construction, we must take S very small.

In Huygens' construction as explained above there is one basic difficulty. As the secondary wavelets spread out in all directions, we can construct another wavefronts (A'B') on the back indicating backwave moving towards the source O. But this is never observed experimentally. The problem was taken care of theoretically by Fresnel and Kirchhoff. They showed that the intensity of the secondary waves in the rear is zero and hence the rear envelope of the circles cannot represent the section of the wavefront.

Huygens' principle in its original form had no firm theoretical background. Fresnel successfully modified the principle somewhat and later Kirchhoff showed that Huygen-Fresnel principle was a direct consequence of the differential wave equation. From the standpoint of wave theory, one can define a ray as a straight line perpendicular to the wavefront.

Interference of Light

Let monochromatic waves of light (having wavelength λ) from two sources A and B proceed in the same direction and superpose on a particle P of the medium (Fig. 1). The displacement of the particle and hence the intensity of light there, will be maximum or minimum according as the waves meet the particle in the same or in the opposite phase respectively. This phenomenon is known as interference of light. The waves meet in the same phase, when the crest or trough of one wave respectively superpose on the crest or trough of another wave. The waves meet in the opposite phases when the crest of one wave superposes on the trough of another. The conditions, which are essential to have the intensity at a point either permanently maximum or minimum, (i.e, to have a permanent interference phenomenon at the point) can be established analytically as follows—

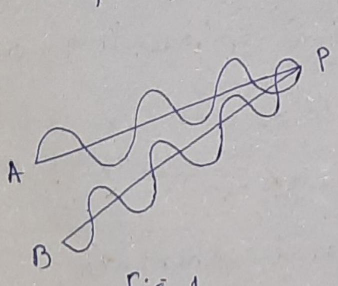


Fig. 1.

Let the two wave trains from two point sources A and B are given by—

$$y_1 = a\sin\frac{2\pi}{\lambda}(vt - x) = a\sin\theta$$

$$y_2 = b\sin\left[\frac{2\pi}{\lambda}(vt-x) + \delta\right] = b\sin(\theta+\delta)$$

On superposition at P, the resultant displacement

$$y = y_1 + y_2 = a\sin\theta + b\sin\theta\cos\delta + b\sin\delta\cos\theta$$

$$= (a + b\cos\delta)\sin\theta + b\sin\delta\cos\theta$$

$$= A\sin(\theta + \phi)$$

where,

$$A\cos\phi = a + b\cos\delta$$

$$A\sin\phi = b\sin\delta$$

$$\therefore \tan\phi = \frac{b\sin\delta}{a + b\cos\delta}$$

The amplitude after superposition is given by $A = (a^2 + b^2 + 2ab\cos\delta)^{1/2}$ and depends on the phase difference δ of the two waves at the point of superposition.

Let us define the term 'coherence'. Coherence is related to the definite phase relationship at different points of time and space. For a source to be coherent it must emit radiation of single frequency or the frequency spread must be very small and also the wavefront must have shape which remains constant in time. Thus we speak of two types of coherence — temporal coherence and spatial coherence. By the term, temporal coherence we mean a correlation between the electromagnetic disturbance produced by a wave field at a given point in space at a time t and that produced at earlier or later times. By the term, spatial coherence we mean a correlation between the phases of electromagnetic radiation emanating at two different points.

Therefore, δ depends on two factors —

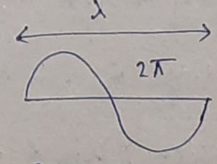
i) δ_1, due to temporal coherence

ii) δ_2, due to spatial coherence

$\therefore \delta = \delta_1 + \delta_2$

if $\delta_1 = 0$.

Then $\delta = \delta_2 = \frac{2\pi}{\lambda} x$



$\delta_2 = \frac{2\pi}{\lambda} \times$ path difference between two waves

$= \frac{2\pi}{\lambda} x$

The amplitude is maximum, when,

$\cos\delta = 1 = \cos 2n\pi$

or, $\delta = 2n\pi$

or, $\frac{2\pi}{\lambda} x = 2n\pi$ or, path difference $= x = n\lambda$ for bright fringes.

$\therefore A_{max} = (a+b)$

$\therefore$ Intensity $I_{max} = A_{max}^2 = (a+b)^2$

For minimum amplitude,

$\cos\delta = -1 = \cos(2n+1)\pi$

or, $\delta = (2n+1)\pi$

or, $\frac{2\pi}{\lambda} x = (2n+1)\pi$ or, path difference $= x = (2n+1)\frac{\lambda}{2}$ for dark fringes.

$\therefore A_{min} = a \sim b$

$\therefore$ Intensity $I_{min} = (a \sim b)^2$

If $a=b$ then, $I_{max} = 4a^2$ and $I_{min} = 0$.

When $n=0$; we get the central bright band for which the path difference is zero. On the both sides of the central bright band, there ~~and~~ are bands which are alternately of minimum and maximum intensity. This phenomenon is known as interference of light.

Why light from two different candles are not seen to interfere?

When the light waves from two separate candles meet at a point, the point will be bright or dark according as the waves meet in the same or opposite phases. This phase difference of the waves meeting at a point will be dependent on: (a) the path difference of the point from two sources and (b) phase relationship of the waves when they are emitted from the two sources. The path difference of the point in question is the same for all times but the phase relation of the two waves when they are emitted by the two different candles cannot be the same for all times. In fact, this phase difference of the waves as they start from the two sources is changing very rapidly with time, and the result is that at every point there is a rapid alteration of brightness and darkness and hence we get general illumination.

Explain – Light Energy is not destroyed by interference

By interference of light we get alternately brightness and darkness. The light energy at the dark region is not destroyed but it is simply redistributed in the neighbouring bright regions. Thus in interference, there is not the destruction of light energy but it is a case of redistribution of energy. The energy in the bright region is greater when interference takes place, than when there was general illumination. Thus the phenomenon of interference is consistent with the principle of conservation of energy.

Width and Shape of Interference Fringes

Fringe Width

Let us consider two coherent sources S_1 and S_2, which are sending monochromatic light of wavelength λ. The point C on the screen AB is equidistant from S_1 and S_2 and hence the waves from S_1 and S_2 arrive at

C at the same time. If the phase difference between

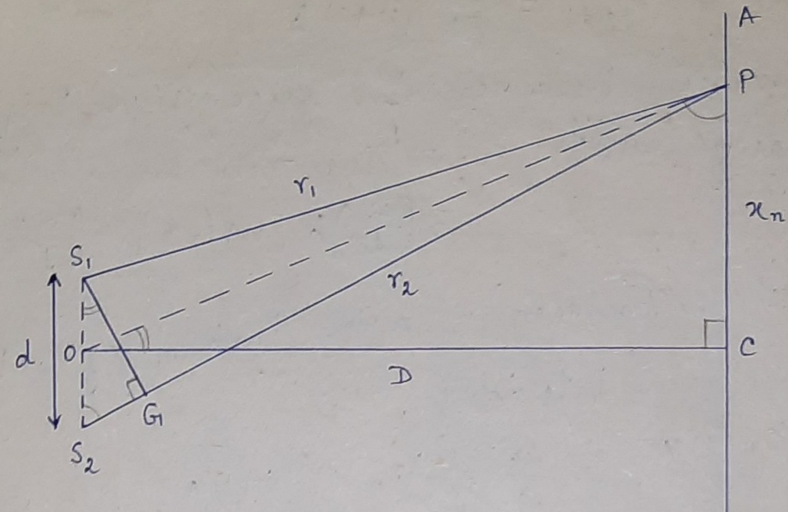


Fig. 2.

The two coherent sources is nil, the waves from S_1 and S_2 will meet at C in same phase and will produce a bright band, known as central bright band. Let the n^{th} bright band be formed at P whose distance from C is x_n (say). Now, in Fig. 2, D is the separation between the sources and screen; d is the separation between coherent sources; usually $D \gg d$. Path difference between the waves at P is $r_2 - r_1 = S_2G$.

From similar triangles, S_1S_2G and OPC,

$$\frac{S_2G}{PC} = \frac{S_1G}{OC} \quad \text{or,} \quad S_2G = \left(\frac{S_1G}{OC}\right) PC = \frac{d}{D} x_n \quad [\because S_1G \simeq S_1S_2 \simeq d]$$

If the initial phase difference is zero, then total phase difference is due to path difference only.

$$\therefore \delta = \frac{2\pi}{\lambda}\left(\frac{d}{D}\right) x_n$$

For maximum intensity, $\delta = 2n\pi$

$$\text{or,} \quad \frac{2\pi}{\lambda}\left(\frac{d}{D}\right) x_n = 2n\pi$$

$\therefore x_n = \left(\frac{D}{d}\right) n\lambda$, distance of the n^{th} bright fringe from C.

Similarly, the distance of $(n+1)$ th bright fringe from C is $x_{n+1} = \left(\frac{D}{d}\right)(n+1)\lambda$

$\therefore$ Separation between two successive bright fringes i.e, fringe width $= \beta = x_{n+1} - x_n$

$$= \left(\frac{D}{d}\right)\lambda\,[n+1-n] = \left(\frac{D}{d}\right)\lambda \rightarrow (1)$$

If P be the position of n th dark fringe, then, for minimum intensity,

$$\delta = (2n+1)\pi \quad \text{or,} \quad \frac{2\pi}{\lambda}\left(\frac{d}{D}\right)x_n = (2n+1)\pi$$

$$\text{or, } x_n = (2n+1)\left(\frac{D}{d}\right)\left(\frac{\lambda}{2}\right)$$

Similarly, for $(n+1)$ th dark fringe,

$$x_{n+1} = [2(n+1)+1]\left(\frac{D}{d}\right)\left(\frac{\lambda}{2}\right) = (2n+3)\left(\frac{D}{d}\right)\left(\frac{\lambda}{2}\right)$$

$\therefore$ Fringe width $= \beta = x_{n+1} - x_n = (2n+3-2n-1)\left(\frac{D}{d}\right)\left(\frac{\lambda}{2}\right)$

$$= \frac{D\lambda}{d} \rightarrow (2)$$

(1) & (2) show that distance between two successive bright or dark fringes are equal. Measuring β, d and D, one can find out λ.

Shape

An idea regarding the shape of interference fringes can be obtained by finding out the locus of points having constant path difference from the sources S_1 and S_2. Refering to Fig. 3., we choose the midpoint O between the sources as the origin of a coordinate system.

Let the x-axis be along OX and y-axis be perpendicular to the plane containing the sources. For any

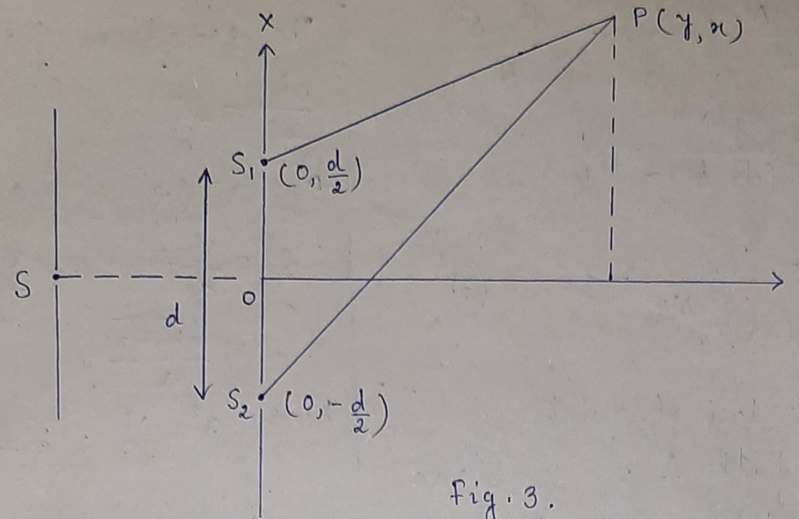


Fig. 3.

point $P(y,x)$, we can write,

$S_1P^2 = y^2 + \left(x-\frac{d}{2}\right)^2$ and $S_2P^2 = y^2 + \left(x+\frac{d}{2}\right)^2$ where d is the separation between the sources. The path difference is $\Delta = S_2P - S_1P$

or, $\Delta + S_1P = S_2P$

or, $\Delta + \left[y^2 + \left(x-\frac{d}{2}\right)^2\right]^{1/2} = \left[y^2 + \left(x+\frac{d}{2}\right)^2\right]^{1/2}$

Squaring both sides, we get,

$\Delta^2 + y^2 + \left(x-\frac{d}{2}\right)^2 + 2\Delta\left[y^2 + \left(x-\frac{d}{2}\right)^2\right]^{1/2} = y^2 + \left(x+\frac{d}{2}\right)^2$

or, $\Delta^2 + y^2 + x^2 - xd + \frac{d^2}{4} + 2\Delta\left[y^2 + \left(x-\frac{d}{2}\right)^2\right]^{1/2} = y^2 + x^2 + xd + \frac{d^2}{4}$

or, $2\Delta\left[y^2 + \left(x-\frac{d}{2}\right)^2\right]^{1/2} = 2xd - \Delta^2$

Again, squaring both sides, we get,

$4\Delta^2\left[y^2 + \left(x-\frac{d}{2}\right)^2\right] = 4x^2d^2 - 4xd\Delta^2 + \Delta^4$

or, $4\Delta^2y^2 + 4\Delta^2x^2 - 4xd\Delta^2 + \Delta^2d^2 = 4x^2d^2 - 4xd\Delta^2 + \Delta^4$

or, $\Delta^2y^2 + \Delta^2x^2 + \frac{\Delta^2d^2}{4} = x^2d^2 + \frac{\Delta^4}{4}$

or, $y^2 + x^2 + \frac{d^2}{4} = \frac{x^2d^2}{\Delta^2} + \frac{\Delta^2}{4}$

or, $x^2\left(\frac{d^2}{\Delta^2} - 1\right) - y^2 = \frac{d^2 - \Delta^2}{4}$

or, $x^2\left(\frac{d^2-\Delta^2}{\Delta^2}\right) - y^2 = \frac{d^2-\Delta^2}{4}$

or, $\frac{x^2}{\Delta^2} - \frac{y^2}{(d^2-\Delta^2)} = \frac{1}{4}$

or, $\frac{x^2}{(\Delta^2/4)} - \frac{y^2}{\left(\frac{d^2-\Delta^2}{4}\right)} = 1 \longrightarrow (3)$

Thus the loci of points of constant path Δ in xy plane are hyperbolae with S_1 and S_2 as foci on x-axis.

[The equation of a hyperbola, is

$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$

$b^2 = a^2(e^2-1)$, $e > 1$

$\therefore e^2 = \frac{b^2}{a^2} + 1 = \frac{b^2+a^2}{a^2}$

$\therefore e = \frac{\sqrt{b^2+a^2}}{a}$

here, $a^2 = \frac{\Delta^2}{4}$; $b^2 = \frac{d^2-\Delta^2}{4}$]

The eccentricities of the hyperbolae are given by,

$$e = \frac{[\Delta^2/4 + (d^2-\Delta^2)/4]^{1/2}}{\Delta/2} = \frac{d}{2}\frac{2}{\Delta} = \frac{d}{\Delta}$$

In optical experiments, the path difference $\Delta \sim 10^{-8}$ cm and $d \sim 10^{-2}$ cm. Therefore, e is very high, as a consequence of which eq. (3) becomes —

$$\frac{x^2}{(\Delta^2/4)} - \frac{y^2}{\left(\frac{d^2-\Delta^2}{4}\right)} = 1 \quad \text{or,} \quad \frac{x^2}{(\Delta^2/4)} - \frac{y^2}{\left[\Delta^2/4\left(\frac{d^2}{\Delta^2}-1\right)\right]} = 1$$

or, $x^2 - \frac{y^2}{\left(\frac{d^2}{\Delta^2}-1\right)} = \frac{\Delta^2}{4} \simeq 0$ as Δ is very small.

or, $y^2 = x^2\left(\frac{d^2}{\Delta^2}-1\right)$ or, $y = \pm\left(\frac{d^2-\Delta^2}{\Delta^2}\right)^{1/2} x$ — (4)

Equation (4) represents straight ~~line~~ line. Therefore, the hyperbolae become practically straight lines.

If instead of slits we use two coherent point sources S_1 and S_2 then in three dimensional space the loci of maxima i.e, bright fringes of different order numbers will form a system of confocal hyperboloids with S_1 and S_2 as foci (fig. 4). If a screen is placed parallel to the line joining S_1 and S_2 then short straight line fringes (parallel to the length of slits placed at S_1 and S_2) will be obtained. If the screen is placed perpendicular to the line joining S_1 and S_2, we shall get a number of alternately bright and dark concentric circles with their common centre lying on the point of intersection of the line S_1S_2 with screen. These fringes are called non-localized fringes for they can be obtained on a screen wherever it is placed.

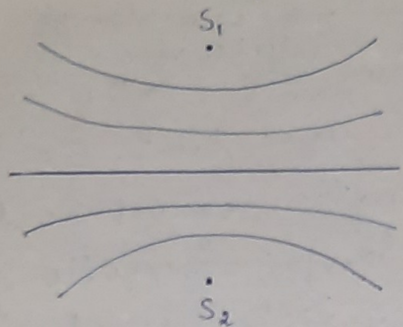


fig. 4.

Interference with white light and colour effect

The distance of the n^{th} bright fringe from the central one is given by $x_n = n\left(\frac{D}{d}\right)\lambda$, where d is the separation between two coherent sources and D is the distance of the screen from the source. x_n is a function of λ. So the fringes of different colours will be in step only for the central band ($n=0$). In this case, $x_n = 0$ and bright fringe of all wavelengths will coincide and the central band will be white with white light. For higher order bright fringes ($n>0$) x_n will be greater for a light of longer wavelength and less for a light of shorter wavelength. As the wavelength λ_r for red light is longer than the wavelength λ_v for violet we infer that all bright bands, excepting the central bright band, will be coloured in which red will be in

The outermost position while violet will be in the inner-most position. When the path difference between the interfering waves is large, the condition for constructive interference for one wavelength and the condition for destructive interference for another wavelength may be satisfied at the same point. In that case, the resultant illumination cannot be distinguished from white light. So for observable fringes with white light the path difference should be kept very small. In this case, we get a few coloured fringes on either side of central white fringe.