M. Sc. Semester II' 2020

Subject: Physics

Course No.: PHS 204 (CBCS)

Section: Important Developments of Physical Science before 20th century

Topic: Wave theory of light: Young

Lecture No. 3

by

Biswajit Das

Assistant Professor, Department of Physics, Vidyasagar University

1. Introduction:

In between the period of 17th and 18th centuries, many scientists proposed the wave theory of light based on experimental observations, including Robert Hooke, Christiaan Huygens and Leonhard Euler. They predicted that if light is a wave, we will see certain things. We will see that light can reflect off shiny surfaces, refract when it passes through the two media, and diffract around objects or when moving through slits. It is also possible to see interference, where crests or troughs of the propagating waves superpose to create brighter light, and crests and troughs cancel out to create darker areas. All of these things were seen in formal experiments by the 19th century.

In 1678, Dutch physicist, Christiaan Huygens, believed that light is made up of waves vibrating up and down perpendicular to the direction of the light travels, and therefore formulated a way of visualising wave propagation. He proposed that every point that a luminous disturbance meets turns into a source of the spherical wave itself. The sum of the secondary waves, which are the result of the disturbance, determines what form the new wave will take. This theory of light is known as 'Huygens' Huygens was successful in deriving the laws of reflection and refraction of light and in explaining the propagation of light using this theory. However, he was not able to explain the diffraction of light. Diffraction is the resultant of the wave-nature of light. When light waves pass through an aperture or across a sharp edge, the light waves spread out.

At the time, some of the experiments conducted on light theory, both the wave theory and particle theory, have some unexplained phenomenon; Newton could not explain the phenomenon of light interference. Interference is caused by waves interacting with each other as they intersect, causing a wave to either add together or cancel. In 1803, an English physician and physicist, Thomas Young, studied the interference of light waves by allowing light to pass through two closely set pinholes equally separated onto a screen; he found that the light emerging from the two pinholes is spread out according to Huygen's principle. The experiment conducted by Thomas Young on the interference of light proved Huygens' wave theory of light to be correct. Later in 1815, Augustin Fresnel supported Young's experiments with mathematical calculations. However, it is established that light can exhibit both wave nature and particle nature at the same time. Much of the time, light behaves like a wave.

In 1817, Young proposed that the light waves are transverse as they oscillate in the direction traverse to the direction of wave travel, and thus explained polarization, the alignment of light waves to vibrate in the same plane. The founder of physiological optics and father of the wave theory of light, Young was the first to give the word energy its scientific significance, and Young's modulus, a constant in the mathematical equation describing elasticity, was named in his honour. Young was the first to describe astigmatism in 1801 and in the same year he turned to the study of light.

2. Wave and its characteristics:

A wave is a disturbance that propagates through space. Waves transfer energy, momentum, and information, but not mass. From ripples on a pond to deep ocean swells, sound waves, and light, all waves share some basic characteristics. Most waves move through a supporting medium, with the disturbance being a physical displacement of the medium. The time dependence of the displacement at any single point in space is often an oscillation about some equilibrium position.

A simple and useful example of a periodic wave is a harmonic wave. A harmonic wave is a wave with a frequency that is a positive integer multiple of the frequency of the original wave, known as the fundamental frequency. The original wave is also called the 1^{st} harmonic, the following harmonics are known as higher harmonics such as 2^{nd}, 3^{rd}, 4^{th} harmonics and so on. As all harmonics are periodic at the fundamental frequency, the sum of harmonics is also periodic at that frequency. It is typically applied to repeating signals, such as sinusoidal waves. The sine wave is the fundamental waveform in nature. When dealing with light waves, we usually refer to the sine wave. At any frequency other than a harmonic frequency, the resulting disturbance of the medium is irregular and non-periodic.

The wavelength λ of the wave is the physical separation between successive crests. The maximum displacement of the wave, or amplitude, is denoted by a. The time between successive oscillations is called the period T of the wave. The number of oscillations per second is the wave frequency f, which is the reciprocal of the period, $1/T$. This is shown in Fig. 1.

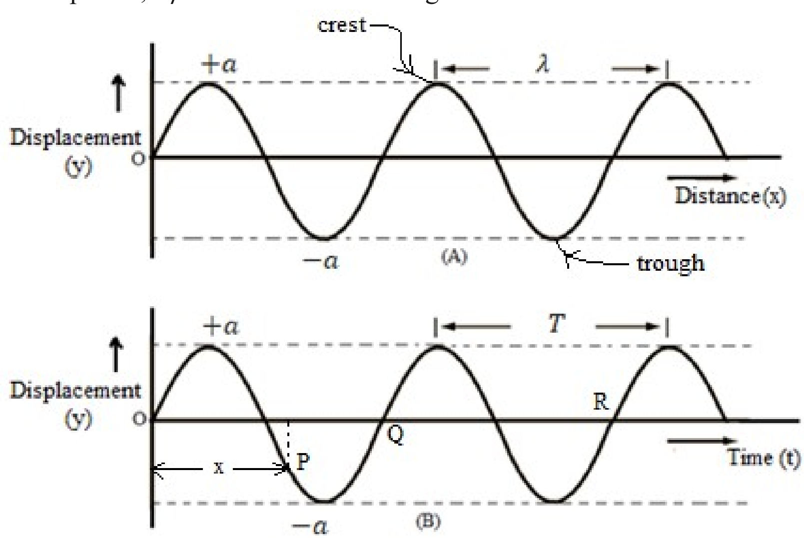

Fig. 1: Harmonic waves: (A) variation with position at one time (B) Variation with time at one place

Consider a single point is fixed in space, the number of wave crests that pass that point per second is the wave frequency f. The distance traveled past that point by any one crest in one second called the wave velocity v which is equal to the distance between crests λ multiplied by the frequency f, i.e., $v = \lambda f$.

Now, when the wave propagated from point O to point Q, its phase change is $= 2\pi$. So, when the wave propagated from point O to another point P at distance x from O (Fig. 1), we have the phase change ϕ as given by

$$\phi = \frac{2\pi}{\lambda}.x = kx$$

where, $k = \frac{2\pi}{\lambda}$ is the wave number.

The properties of harmonic waves are illustrated in the mathematical expression (1) for the displacement in both space and time. For a harmonic wave propagating in the x-direction, the spatial and time dependence of the displacement $y(x,t)$ is

$$y(x,t) = a\sin\frac{2\pi}{\lambda}(vt - x) \qquad \text{... (1)}$$
$$= a\sin\left(\frac{2\pi}{\lambda}vt - \frac{2\pi}{\lambda}x\right)$$
$$= a\sin\left(2\pi ft - \frac{2\pi}{\lambda}x\right)$$
$$= \sin(\omega t - kx)$$
$$= \sin(\omega t - \phi)$$

where, the term ω is the angular frequency of the wave, it is related with T and f as given by

$$\omega = 2\pi f = \frac{2\pi}{T}$$

3. Superposition of waves:

Superposition is a defining characteristic of all waves; it is used to describe the behaviour of overlapping waves. The principle of superposition of waves (or, simply, superposition principle) states that when two or more propagating waves overlap at a point in space, the resultant displacement at that point is equal to the vector sum of the displacement of the individual waves. When two waves from coherent sources superpose at a point, then the distribution of energy due to one wave is disturbed by the other.

4. Coherent sources:

Coherent sources of light which emit light waves continuously of same wavelength, and time period, frequency and amplitude and have zero phase difference or constant phase difference are coherent sources.

5. Graphical study of interference of light:

Interference is a phenomenon in which two waves superpose at a point on an intersecting path in space to form a resultant wave of greater, lower, or the same amplitude. The effect is that of the addition of the amplitudes of the individual waves the point affected by more than one wave.

Let S_1 and S_2 be two coherent sources of light. They are very close to each other from which monochromatic waves are proceeding towards a screen and to be superposed on it (Fig. 2). The amplitude and hence the intensity of the resultant wave will be maximum or minimum according as the waves reach the point of superposition in the same or in the opposite phases respectively. The waves will be in same phase when the crests or troughs of one wave become superposed respectively on the crests or troughs of the other. The phases will be opposite if the crests of one wave are superposed on the troughs of the other.

The condition for the two waves to meet in the same or in the opposite phase depends on two factors: (i) the phase-difference of the waves at the instant they leave the sources and (ii) the path-difference of the point of superposition of the waves from the two sources. If a source emits crests or troughs just when another source is emitting crests or troughs, then the phase-difference of the two waves at the instant of their start from the sources will be zero (0). If the waves are in opposite phases at the start, then it will be π. These conditions are fulfilled only if the two sources are coherent, i.e., if they are derived from a single source and are identical in all respect.

If the path-difference of two waves originating from coherent sources be even multiples of half of the wavelength, i.e., $2n\lambda/2$, then the waves reach the point of superposition in the same phase and the resultant intensity will be maximum. If this path-difference be odd multiples of half of the wavelength, i.e., $(2n+1)\lambda/2$, then the waves will reach there in opposite phase and the resultant intensity will be minimum. Where, $n = 0, 1, 2, 3, etc.$ and λ is the wavelength of light.

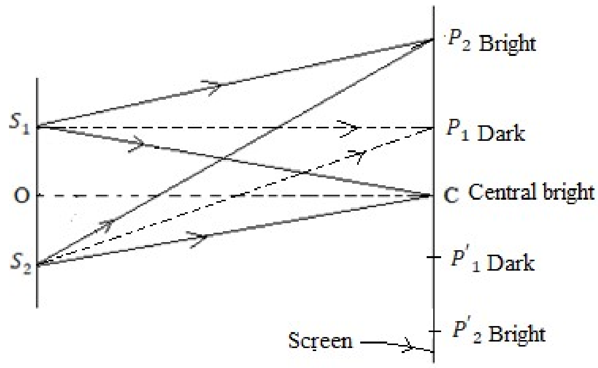

Fig. 2: Interference of two light waves

In Fig. 2, the central point C on the screen is equivalent from the two sources S_1 and S_2. Obviously, the path-difference of the waves at C is zero. The waves from S_1 and S_2 reach the point C in the same phase and a bright band is formed at C, this bright band is known as the **central** bright band. If the path-difference of the waves from S_1 and S_2 is $\lambda/2$, then the waves will reach at P_1 in the opposite phase and a dark band is formed at P_1. If the path-difference of the waves from S_1 and S_2 is $2\lambda/2$, then the waves will reach at P_2 in the same phases and a bright band will be produced at P_2, this bright band is referred to as **first order** bright band. In this way alternate bright and dark bands are produced on either side of the central band at C. thus the intensity of light will be alternately maximum and minimum. This gives an interpretation of the interference of light waves graphically.

Interference is of two types: constructive interference and destructive interference.

In constructive interference, the crests of two waves coincide at a point and the waves are said to be in phase with each other. Their superposition results in a reinforcement of the disturbance. Constructive interference results where the crest or trough of one coincides with the crest or trough of the other. If two of the components are of the same frequency and phase (*i.e.,* they vibrate at the same rate and are maximum at the same time), the amplitude of the resulting combined wave (or, resultant wave) is the sum of the individual amplitudes and its value will be double the amplitude of the individual waves if they are of equal amplitude. In case of destructive interference, the crest of one wave coincides with the trough of a second wave and they are said to be out of phase. That means if the two waves are out of phase by $1/2$ period (*i.e.,* one is minimum when the other is maximum), the result is destructive interference. The amplitude of the resultant wave equals the difference between the amplitudes of the individual waves. If the two individual amplitudes are equal, then the destructive interference is complete and the net disturbance to the medium is zero. Fig. 3 shows the interference of two waves. When in phase, the two lower waves create the constructive interference (Fig. 3A), resulting in a wave of greater amplitude. When $180°$ out of phase these two waves create the destructive interference (Fig. 3B).

Constructive interference occurs when the phase-difference between the waves is an even multiple of π (i.e., 2π, 4π, 6π, $etc.$) or the path-difference is even multiples of $\lambda/2$ (i.e., $0, \lambda, 2\lambda, 4\lambda, etc.$); whereas destructive interference occurs when the phase-difference is an odd multiple of π (i.e., π, 3π, $5\pi, etc.$) or the path-difference is odd multiples of $\lambda/2$, (i.e., $\lambda/2, 3\lambda/2$, $5\lambda/2, etc.$).

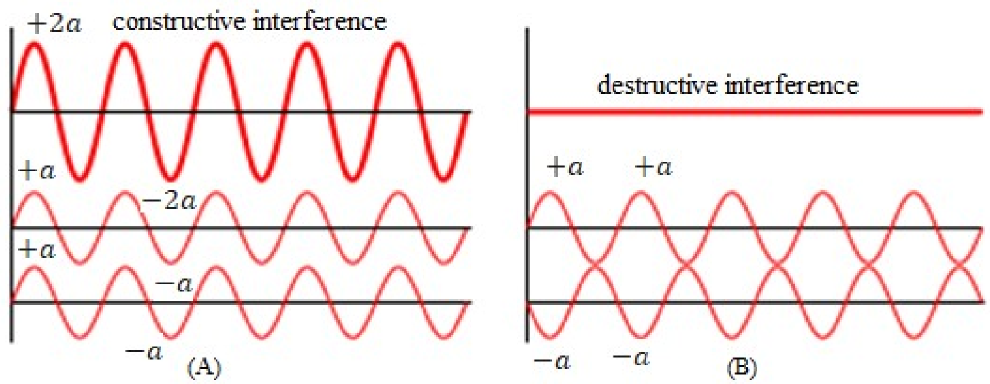

Fig. 3: Constructive and destructive interference

6. Analytical treatment of interference:

Let two waves of equal wavelength λ and amplitude a, differing in path δ are superpose on a point P. According to principle of superposition, the optical effect at the point P at any instant of time is the algebraic sum of the optical effects which would have been produced by each of the individual waves.

The displacements of the two individual waves at any instant t are given by

$$y_1 = a\sin\frac{2\pi}{\lambda}(vt - x) \qquad \text{... (1)}$$

$$y_2 = a\sin\frac{2\pi}{\lambda}[vt - (x+\delta)] \qquad \text{... (2)}$$

Here, v is the velocity of the two individual waves.

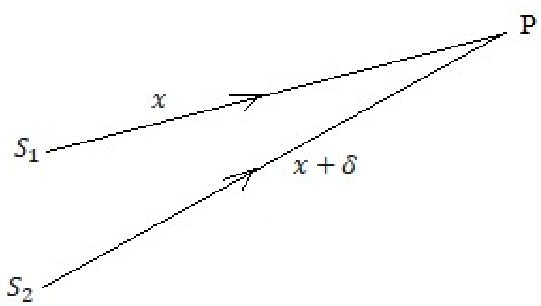

Fig. 4: Two waves superposed on a point P

The resultant displacement at the point P at the instant t is then given by

$$y = y_1 + y_2 = a\sin\frac{2\pi}{\lambda}(vt - x) + \sin\frac{2\pi}{\lambda}[vt - (x+\delta)]$$

$$= 2a\cos\frac{\pi\delta}{\lambda}.\sin\frac{2\pi}{\lambda}\left[vt - \left(x + \frac{\delta}{2}\right)\right]$$

$$= A\ \sin\frac{2\pi}{\lambda}\left[vt - \left(x + \frac{\delta}{2}\right)\right] \qquad \text{... (3)}$$

Equation (3) shows that the resultant wave is a sine wave of amplitude A which is given by

$$A = 2a\cos\frac{\pi\delta}{\lambda} \qquad \text{... (4)}$$

The amplitude A depends on δ and will change with δ.

When $\delta = \frac{\lambda}{2}, \frac{3\lambda}{2}, \frac{5\lambda}{2}, \dots \dots (2n+1)\frac{\lambda}{2}$, then $A = 0$, i.e., resultant amplitude is zero (0). As the intensity of light is proportional to the square of its amplitude, so the intensity of the resultant wave is zero at P. Thus a minimum intensity (i.e., a dark band) is obtained when the path-difference between the two waves is equal to odd multiple of $\frac{\lambda}{2}$. So, the condition for destructive interference the path-difference is odd multiples of $\lambda/2$, (i.e., $\lambda/2, 3\lambda/2,\ 5\lambda/2, etc.$) or the phase-difference is an odd multiples of π (i.e., $\pi,\ 3\pi,\ 5\pi, etc.$), where, $n = 0, 1, 2, 3, etc.$

When $\delta = 0, \lambda, 2\lambda, 4\lambda,\ \dots \dots 2n\frac{\lambda}{2}$, then $A = \pm 2a$, i.e., resultant amplitude is $|A|$. So, the intensity of the resultant wave is $4a^2$ at P. Thus for a maximum intensity, i.e., for a bright band, the path-difference between the two waves is equal to even multiple of λ. So, the condition for constructive interference the path-difference is even multiples of $\lambda/2$ (i.e., $0, \lambda, 2\lambda, 4\lambda, etc.$) or the phase-difference is even multiples of π (i.e., $0, 2\pi,\ 4\pi,\ 6\pi,\ etc.$).

6.1. Conditions for production of interference:

(i) The two sources of light should emit continuous waves of same wavelength and same time period i.e. the source should have phase coherence.
(ii) The two sources of light should be very close to each other.
(iii) The waves emitted by two sources should either have either zero phase difference or some constant difference in phase.
(iv) The two sources should be narrow.
(v) The amplitude of two waves should be preferably same.

7. Young's double slit experiment:

In Young's own judgment, of his many achievements the most important was to establish the wave theory of light. In the early 19th century Young put forth a number of theoretical reasons supporting the wave theory of light, and he developed two enduring demonstrations to support this viewpoint.

In his apparatus as shown in Fig. 5, Young allowed sunlight to pass through a pin-hole S and then at a short distance away this was passed through two pin-holes S_1 and S_2 subsequently. He replaced S_1 and S_2 by slits and S being illuminated by a monochromatic (single colour) source of light (Fig. 5).

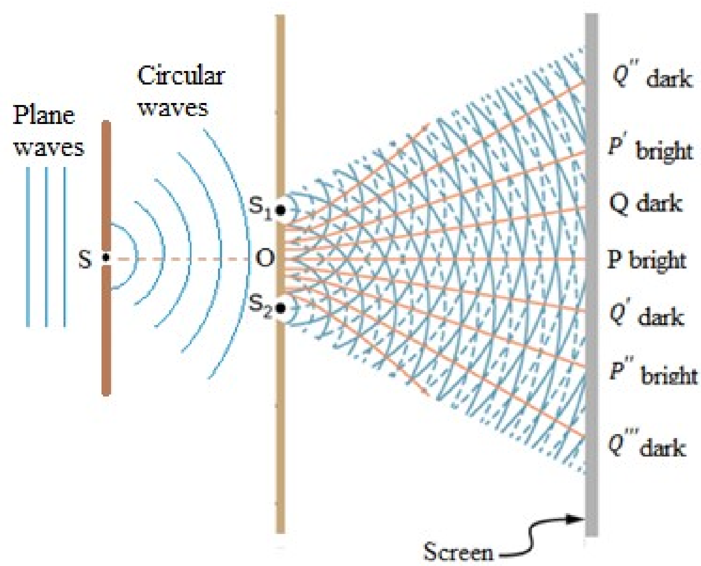

Fig. 5 Double slit produce interference pattern

Now, according to Huygens' principle, the slit S_1 and S_2 behave as new sources to produce circular waves. These waves are of transverse type. The crests of the waves are shown by solid circular lines and the troughs by broken circular lines. The loci of the points of intersection where the crests reinforce crests and troughs reinforce troughs are meeting on different positions on the screen. These positions are the position of maximum intensity and so produce constructive interference. The loci of the points of intersection where crests and troughs from S_1 and S_2 superpose are meeting on other different positions. These positions correspond to minimum intensity and produce destructive interference.

Two wave trains of light from a double slit produce interference, an effect that is visible on a screen as a pattern of alternating dark and light bands caused by intensification and extinction at points at which the waves are in phase and out of phase, respectively. This is shown in Fig. 5.

7.1. Theory of Young's experiment:

As shown in Fig. 6, let S_1 and S_2 be two narrow slits at a distance d wherefrom light waves of same wavelength and amplitude radiate. We assume that light waves (wavelength, λ) start from S_1 and S_2 in the same phase. So, these two point sources of light can behave as the two coherent sources. If a screen be kept at a distance D away in front of the sources, a number of alternate bright and dark bands will be seen on the screen if a monochromatic light is used.

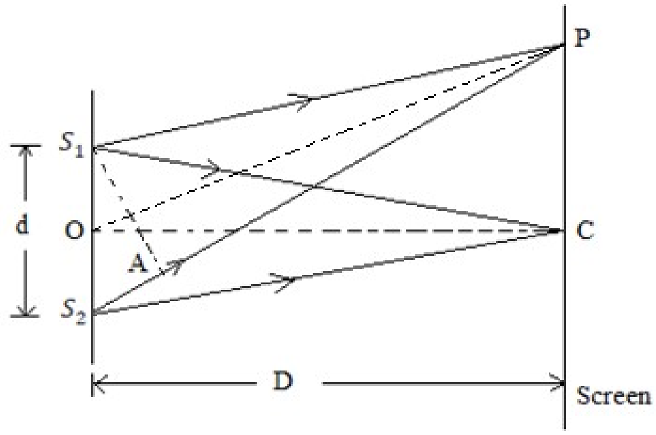

Fig. 6: Interference pattern

Now, we join S_1 and S_2 and let O be its mid-point. Through the point O, we draw a straight line OC perpendicular to S_1S_2 and the screen. As length of S_1C and S_2C are equal, the light waves starting from S_1 and S_2 in the same phase will reach the point C also in the same phase, and then reinforce one another on C. thus the point C will be a bright point.

We now determine the nature of illumination at any point, say P, situated at a distance CP on the screen. Let us join PS_1, PS_2 and PO. A normal S_1A is drawn on the line PS_2. If S_1S_2 is very small in comparison to PS_1, then S_1A will be perpendicular to both PS_2 and PO. Then, $S_2P - S_1P = S_2A$ is the path difference of the wave from S_2 with respect to that from S_1.

Now, for the point P to be bright, we must have $S_2A = 2n.\frac{\lambda}{2}$ and for the point P to be dark, we must have $S_2A = (2n+1).\frac{\lambda}{2}$.

Therefore, we have seen the alternate bright and dark bands on either side of C on the screen, it is due to the interference of light originating from two neighbouring sources of light. These bright and dark bands are called interference bands or fringes, they are equidistant from each other.

Next, we calculate the distance of any bright band from the central band. To calculate this, we note that the two triangles S_2S_1A and POC are similar, i.e.,

$$\Delta\, S_2S_1A \;\equiv \Delta\, POC$$

because, $S_1A \perp PO$ and $S_1S_2 \perp OC$ and also $\angle S_2S_1A = \angle POC$.

Hence, we have

$$\frac{PC}{OC} = \frac{S_2A}{S_1S_2} = \frac{S_2A}{S_1A}$$

$$\therefore \qquad S_2A = S_1A.\frac{PC}{OC}$$

Let the n^{th} bright band is produced at the point P at a distance x_n from the centre C, i.e., $PC = x_n$. For S_1S_2 is very small, we have $S_1A = S_1S_2 = d$. Also, $OC = D$. Therefore,

$$S_2A = d.\frac{x_n}{D}$$

$$or, \qquad x_n = \frac{D}{d}.S_2A \qquad ..(1)$$

For the n^{th} bright band, $S_2A = 2n.\frac{\lambda}{2}$. So, the relation (1) becomes

$$x_n = \frac{D}{d}.2n.\frac{\lambda}{2} = \frac{D}{d}.n\lambda$$

$$or, \qquad \lambda = \frac{x_n d}{nD} \qquad ...(2)$$

Relation (2) gives a method of determining the wavelength of a monochromatic light.

The distance between the centres of the n^{th} and $(n+1)^{th}$ bright band, i.e., the width of a bright band is then obtained from

$$\alpha = x_{n+1} - x_n = \frac{D}{d}.[(n+1)\lambda - n\lambda] = \frac{D}{d}.\lambda$$

$$or, \qquad \lambda = \frac{d}{D}.\alpha \qquad ...(3)$$

We find from equation (3) that the expression for the width of a band is independent of the order of the band. Thus with a monochromatic light the interference bands are equi-spaced. The relation (3) can also be used to find λ, the wavelength of light.

Some questions and answers:

Example 1: What will happen if the amplitudes of two interfering waves are not equal ?

Answer: If the amplitudes of the interfering waves is not equal, then the dark points will not be perfectly dark and the bright points lose the brightness to some extent. If the amplitudes be a and b $(a > b)$, the intensity of the bright point is $(a+b)^2$ which will be less than to $(2a)^2$ or $(2b)^2$ and the intensity of the dark pints is $(a-b)^2$ which will be greater than zero.

Example 2: Calculate the fringe-width of interference pattern produced in Young's double slit experiment with two slits 1 mm apart on a screen 1 m away. Wavelength of light is 5893 Å.

Answer: Here, $D = 1\,m$, $d = 1\,mm = 0.001\,m$, $\lambda = 5893\text{ Å} = 5893 \times 10^{-10}\,m$. Therefore, the fringe-width,

$$\alpha = \frac{D}{d}.\lambda = \frac{1}{0.001} \times 5893 \times 10^{-10}\,m = 5893 \times 10^{-7}\,m.$$

Questions:

1. Explain the principle on which interference of light is based ?

2. In a Young's double-slit experiment the slits are separated by $0.2\ cm$ and the screen is placed $1\ m$ away. The slits are illuminated by yellow light of wavelength 5893 Å. Calculate the fringe-width.
3. Deduce an expression for the intensity of light at a point due to superposition of waves coming from two light sources. Hence find the condition of constructive and destructive interference.
4. What is the highest-order maximum for $400\ nm$ light falling on double slits separated by $25.0\ \mu m$?

.

References:

1. Thomas Young: British Physician and Physicist. https://www.britannica.com/biography/Thomas-Young.
2. Born, M.; Wolf, E. (1999). *Principles of Optics*. Cambridge University Press. ISBN 978-0-521-64222-4.
3. A. B. Bhattacharyya and R. Bhattacharya (2008). Undergraduate Physics, Vol. II. New Central Book Agency (P) Ltd., Kolkata – 700009, India.