1-1-1994

A simple description of coherence

William Moebs
Loyola Marymount University

Jeff Sanny
Loyola Marymount University, jeff.sanny@lmu.edu

Recommended Citation

W. Moebs and J. Sanny, The Physics Teacher 32, 54 (1994); doi: 10.1119/1.2343900

A Simple Description of Coherence

William Moebs and Jeff Sanny

Department of Physics, Loyola Marymount University, Los Angeles, CA 90045-2699

In a typical general-physics textbook, the discussion of interference and diffraction is preceded by a description of coherent sources. These sources are usually defined in terms of two point sources that are radiating harmonic waves of the same frequency. They are called *coherent* if the phase difference between the waves at the two sources remains fixed. The sources are said to be *incoherent* if they are turned on and off such that the phase difference between the waves they produce varies randomly with time. Neglected in this model of coherence (rightly so for this level) is the fact that sources cannot really be turned on and off if they are to produce monochromatic radiation.

The description of coherent sources is usually followed by an analysis of Young's double-slit experiment. The light illuminating the slits is monochromatic, and the two coherent sources responsible for the interference pattern are the secondary Huygens oscillators at each slit. However, the interference patterns for nonmonochromatic light are barely considered. For example, the fact that white light produces fringes over only a few orders may be mentioned, but it is never explained. Students are left to ponder how monochromatic red light can produce many orders of fringes, yet somehow have its higher order fringes destroyed by the presence of the other colors of a white light source.

The purpose of this paper is to present a reasonably simple description of coherence that does include the effects of nonmonochromatic light. It is assumed that the students are already aware of Young's double-slit experiment for monochromatic light, and that they have been presented with a derivation of the intensity pattern

$$I = I_0 \cos^2\left(\frac{\pi d \sin\theta}{\lambda}\right) \qquad (1)$$

where (see Fig. 1) θ is the angular position on the screen, d is the distance between the slits, λ is the wavelength of the light, and I_0 is the maximum intensity of the interference pattern. The interference maxima then occur at those angular positions θ_m such that

$$\sin\theta_m = \frac{m\lambda}{d} \qquad (2)$$

The effects of a nonmonochromatic light source are determined by calculating the double-slit interference pattern formed when the monochromatic point source is replaced by a point source that emits light waves over a frequency range from $(f_0 - \Delta f/2)$ to $(f_0 + \Delta f/2)$. Assuming the students have been told about time-averaging of signals, they should know that interference occurs only for light waves from the two slits that have the same frequency. For example, if the source emits white light, the "blue waves" of frequency 6.5×10^{14} Hz from the two slits interfere, as do the "red waves" of frequency 4.7×10^{14} Hz, etc. However, a blue wave does not interfere with a red wave, nor does a green wave interfere with a yellow wave, etc. So, what is observed on the screen is a sum of interference patterns with intensity distributions given by Eq. (1) for every frequency. Each frequency component produces an interference

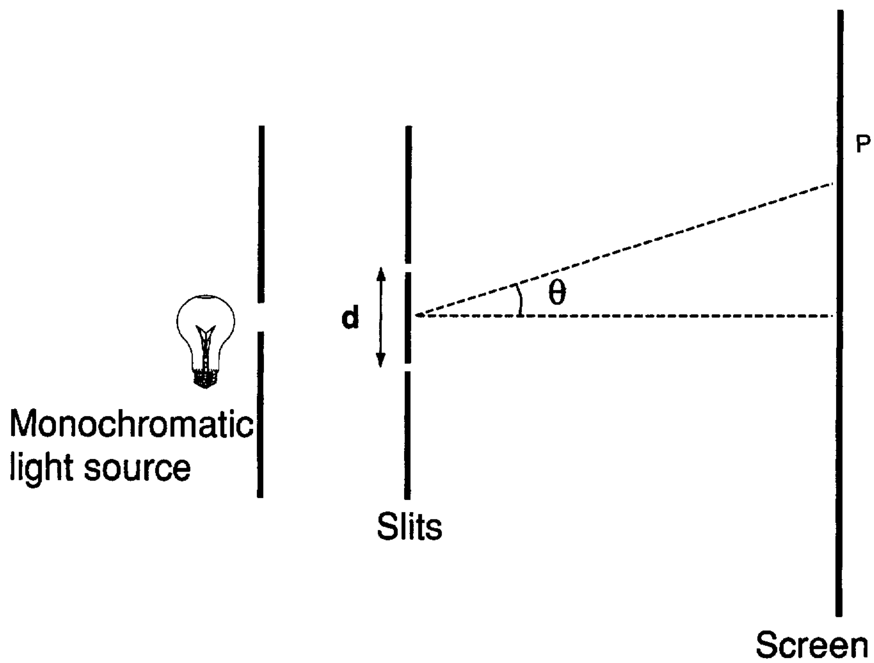

Fig. 1. A simple representation of Young's double-slit interference experiment.

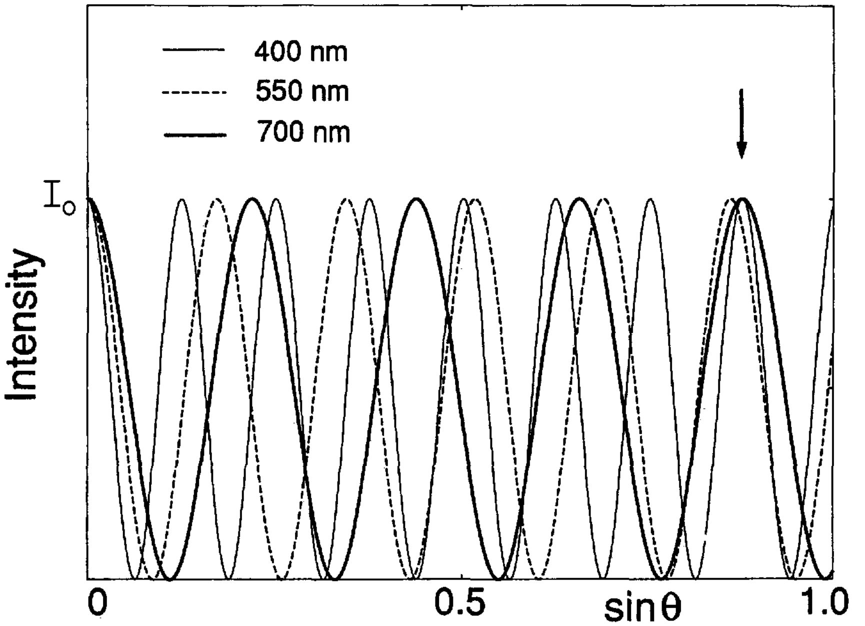

Fig. 2. Double-slit interference patterns for wavelengths 400 nm, 550 nm, and 700 nm. The arrow denotes the strong overlap of the $m = 7$ peak of the 400-nm light, the $m = 5$ peak of the 550-nm light, and the $m = 4$ peak of the 700-nm light.

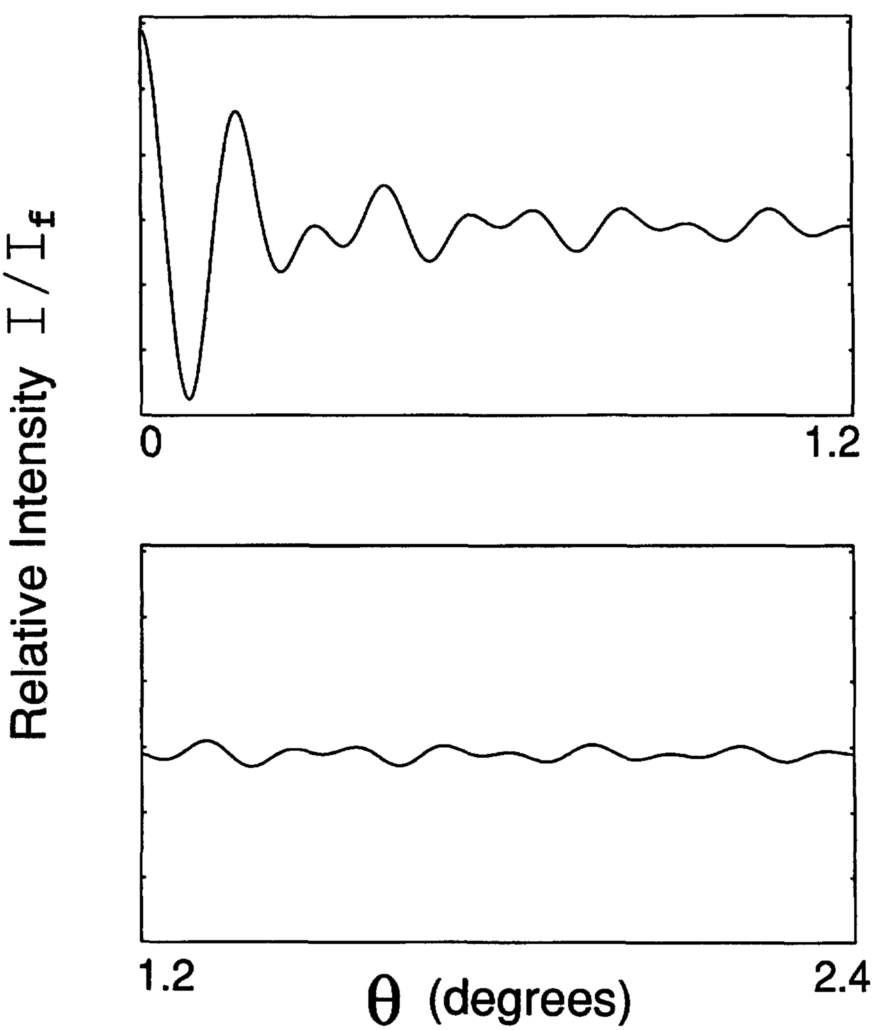

Fig. 3. Intensity pattern due to a source emanating over a continuous frequency band.

pattern, which is calculated by adding sinusoidal electric fields and then squaring to get the intensity. The net intensity pattern is then determined by summing the intensity patterns over all frequencies.

Figure 2 shows the double-slit interference patterns for three different frequencies. These frequencies correspond to the visible wavelengths 400 nm, 550 nm, and 700 nm. The zeroth-order fringe for each frequency component is centered at $\theta = 0$, a fact that is apparent from Eq. (2). Consequently, the color of the intensity pattern around $\theta = 0$ is the same as that of the source (the intensity pattern for a white-light source is white near $\theta = 0$). Away from the center of the pattern, the intensities of the different frequency components are maximized at different spots. This results in a separation of colors for small m in the interference pattern. For larger values of θ, there is strong overlap among different-order maxima for the various colors and the fringes are difficult to separate once again. As an example, notice in Fig. 2 the strong overlap of the $m = 7$ peak of the 400-nm light, the $m = 5$ peak of the 550-nm light, and the $m = 4$ peak of the 700-nm light.

When the source emits over a continuous frequency band, the net intensity is determined by integrating intensities for incremental frequencies over the specified frequency band. To keep the calculations relatively simple, we assume that all frequency components are present with the same intensity. Then for a particular θ, the net intensity contribution dI from a frequency interval df is, from Eq. (1),

$$dI = (I_f df)\cos^2\left(\frac{\pi d \sin\theta}{\lambda}\right) =$$

$$I_f \cos^2\left(\frac{f\pi d \sin\theta}{c}\right) df = I_f \cos^2(\gamma f)\, df$$

where $\gamma = \pi d \sin\theta / c$, and I_f is a constant representing the intensity per unit frequency. Integrating dI over the assumed frequency band, we have

$$I = I_f \int_{f_0 - \frac{\Delta f}{2}}^{f_0 + \frac{\Delta f}{2}} \cos^2 (\gamma f)\, df$$

This integrates to

$$I = I_f \left(\frac{f}{2} + \frac{\sin 2\gamma f}{4\gamma} \right) \Big|_{f_0 - \frac{\Delta f}{2}}^{f_0 + \frac{\Delta f}{2}} =$$

$$I_f \left\{ \frac{\Delta f}{2} + \frac{1}{4\gamma} \left[\sin 2\gamma \left(f_0 + \frac{\Delta f}{2}\right) - \sin 2\gamma \left(f_0 - \frac{\Delta f}{2}\right) \right] \right\}$$

which, with the help of

$$\sin (A \pm B) = \sin A \cos B \pm \cos A \sin B$$

reduces to

$$I = I_f \left[\frac{\Delta f}{2} + \frac{1}{2\gamma} (\cos 2\gamma f_0 \sin \gamma \Delta f) \right]$$

This intensity pattern is illustrated in Fig. 3 by plotting I/I_f versus θ. The values used are $d = 2.0 \times 10^{-4}$ m, $f_0 = 5.5 \times 10^{14}$ Hz, and $\Delta f = 3.0 \times 10^{14}$ Hz. For white light, the constant term $I_f\,(\Delta f/2)$ represents a uniform white background, which is visible everywhere along the interference pattern. A fluctuating intensity given by $(I_f \cos 2\gamma f_0 \sin \gamma \Delta f)/2\gamma$ is superimposed on this background. At some points this term is dominated by red, at other points by blue, and so forth. It gives us the fringes we see in the white-light interference pattern. Since the size of this term decreases as γ (which is proportional to sin θ) increases, the fluctuations become less and less prominent as θ gets larger. Consequently, the fringes become increasingly obscured by the uniform white background the more distant the observation point is from the center of the interference pattern.

Do the higher-order fringes become so faint relative to the background that they are no longer visible? To answer this question, we have to quantify "no longer visible." The criterion we use is that the fluctuating term has to be at least $1/\pi$ times* the constant term in order for a fringe to be observable; that is,

$$\frac{\cos 2\gamma f_0 \sin \gamma \Delta f}{\gamma \Delta f} \geq \frac{1}{\pi}$$

Since the cosine and sine functions are bounded by one, this condition can only be satisfied if

$$\frac{1}{\gamma \Delta f} \geq \frac{1}{\pi} \qquad (3)$$

Suppose we are observing a point P on the screen which is at an angle θ as shown in Fig. 1. The difference in optical path from the two slits to this location is $d \sin \theta$. If the light travels this distance in a time Δt, then $c\Delta t = d \sin \theta$, and

$$\gamma = \frac{\pi d \sin \theta}{c} = \frac{\pi c \Delta t}{c} = \pi \Delta t$$

Substituting this expression for γ into Eq. (3), we obtain

$$\Delta f \Delta t \leq 1 \qquad (4)$$

Thus fringes are visible only when the product of the frequency band width Δf of the source and the difference in transit time Δt between the slits is less than or equal to one.

Whether or not fringes will be visible at P can now be determined if the frequency band width Δf of the illumination is known. First, Eq. (4) with the equality is used to calculate Δt. This time interval is called the coherence time. Next, Δt is multiplied by c to get the coherence length $c\Delta t$ of the signal. Finally, if L is the difference in path lengths from the two slits to P, fringes are seen if $L \leq c\Delta t$, and they are not visible if $L > c\Delta t$.

For example, suppose the slits of Fig. 1 are $d = 2.0 \times 10^{-4}$ m apart and are illuminated by white light with $\Delta f = 3.8 \times 10^{14}$ Hz. To determine how many orders of yellow (5.1×10^{14} Hz) fringes will be visible, the coherence time must first be calculated with Eq. (4):

$$\Delta t = \frac{1}{3.8 \times 10^{14}\ \text{Hz}} = 2.6 \times 10^{-15}\ \text{s}$$

The coherence length is then $(3.0 \times 10^8\ \text{m/s})(2.6 \times 10^{-15}\ \text{s}) = 7.9 \times 10^{-7}$ m. Since the path difference between the two slits is $d \sin \theta = (2.0 \times 10^{-4}\ \text{m}) \sin \theta$, there are visible fringes if

$$(2.0 \times 10^{-4}\ \text{m}) \sin \theta \leq 7.9 \times 10^{-7}\ \text{m}$$

or

$$\sin \theta \leq 3.9 \times 10^{-3} \qquad (5)$$

From Eq. (2),

$$m = \frac{d \sin \theta}{\lambda} = \frac{df \sin \theta}{c} =$$

$$\frac{(2.0 \times 10^{-4}\ \text{m})(5.1 \times 10^{14}\ \text{Hz})}{(3.0 \times 10^8\ \text{m/s})} \sin \theta$$

$$= 340 \sin \theta$$

so with this and Eq. (5) combined,

$$\frac{m}{340} \leq 3.9 \times 10^{-3}$$

and $m \leq 1.3$. Thus only the $m = 0$ and the $m = 1$ yellow fringes can be seen.

On the other hand, if the slits are illuminated with light from a low-pressure Hg^{198} (λ = 546 nm) lamp with $\Delta f = 1.0 \times 10^9$ Hz, the coherence time is

$$\Delta t = \frac{1}{1.0 \times 10^9\ \text{Hz}} = 1.0 \times 10^{-9}\ \text{s}$$

and

$$c\Delta t =$$

$$(3.0 \times 10^8\ \text{m/s})(1.0 \times 10^{-9}\ \text{s}) = 0.30\ \text{m}$$

which is considerably greater than the maximum path difference

$$L = (2.0 \times 10^{-4}\ \text{m}) \sin 90° =$$

$$2.0 \times 10^{-4}\ \text{m}$$

Fringes are therefore visible for all θ from – 90° to +90°.

*We use $1/\pi$ to make Eq. (4) agree with the relationship normally used to estimate coherence time in terms of frequency bandwidth. Actually, this criterion somewhat underestimates our ability to distinguish fringes.