Measurements of Slit-Width Effects in Young's Double-Slit Experiment for a Partially-Coherent Source

Brett J. Pearson  
Dickinson College

Natalie Ferris  
Dickinson College

Ruthie Strauss  
Dickinson College

Hongyi Li  
Dickinson College

David P. Jackson  
Dickinson College

Follow this and additional works at: https://scholar.dickinson.edu/faculty_publications

Part of the Physics Commons

Recommended Citation
Measurements of slit-width effects in Young’s double-slit experiment for a partially-coherent source

BRETT J. PEARSON,* NATALIE FERRIS, RUTHIE STRAUSS, HONGYI LI, AND DAVID P. JACKSON

Department of Physics and Astronomy, Dickinson College, Carlisle, Pennsylvania 17013, USA
*pearsonb@dickinson.edu

Abstract: We provide measurements of Young’s double-slit experiment using a partially-coherent light source consisting of a helium-neon laser incident on a rotating piece of white paper. The data allow a quantitative comparison with both the standard theory that does not account for the width of the slits, and a full, analytic model that does. The data agree much more favorably with the full calculation.

© 2018 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

1. Introduction

Young’s double-slit experiment is the standard method for demonstrating interference/diffraction in the wave model of light. With monochromatic, plane-wave illumination, a Fraunhofer diffraction analysis produces the well-known sinc² cos² pattern on the viewing screen, colloquially described as a “double-slit interference pattern modified by the single-slit diffraction pattern due to the finite width of the slits.” With a partially-coherent source of light, the analysis becomes more complicated. In particular, the detailed structure of the fringe visibility and spacing depends on the complex degree of spatial coherence of the source, and the double-slit arrangement or its extensions can be used to measure coherence aspects of the source. As one example of the usefulness of this phenomenon, the angular size of stars can be determined using the fringe visibility in a Michelson stellar interferometer [1].

The original experiment of Thompson and Wolf using partially-coherent light presented a theoretical model, along with experimental data recorded on photographic plates [2–4]. This model, which we refer to as the standard model, elegantly connects the resulting interference pattern to the complex degree of coherence of the source through the van Cittert-Zernike theorem [5]. However, the standard model considers the degree of coherence between two points in the plane of the mask, and thus does not fully account for the finite size of the two apertures. Since then, discrepancies between the standard model and experimental results have been noted, and different approaches have generalized the theory to incorporate finite-sized apertures [6,7]. In particular, the interference fringes are never observed to disappear completely in the experiments, even though this is predicted by the standard model.

Despite these discrepancies, we were extremely surprised to find no quantitative comparisons between experiment and theory for this situation published in the literature. While there have been a limited number of quantitative studies comparing data with the standard model [8–13], none address the issue of disagreement between the standard theory and experimental data at low visibilities. Here we present a detailed experimental investigation, along with a theoretical model, for the case of a rectangular source and rectangular slits. Our model extends that found in Ref. [6], and leads to a prediction in terms of well-known functions. A complete set of experimental data is presented and compared to both the standard model and the extended model. The data is shown to agree extremely well with the full theory for both rectangular and circular sources.
2. Theoretical models

2.1. Plane-wave Illumination

When illuminated by quasi-monochromatic plane waves (as from a well-collimated laser or a narrow-band point source located at infinity), the double-slit experiment is modeled using the Huygens-Fresnel principle by coherently adding point-source contributions from different locations on the slits by integrating over the aperture dimensions. In the Fraunhofer limit, the viewing screen is sufficiently far away such that the distance from any point over the aperture to a point on the screen can be written as a linear function of source extent. This regime is typically quantified by the condition $s \gg W^2/\lambda$, where $s$ is the distance to the viewing screen, $W$ the aperture size, and $\lambda$ the wavelength of the light.

In practice, the usual assumption for the double-slit experiment is that $s \gg a, d, y$, where $s$ is again the distance to the viewing screen, $a$ is the width of each slit (assumed to be the same), $d$ is the center-to-center separation of the slits, and $y$ is the perpendicular distance away from the optical axis on the viewing screen (see Fig. 1) [14]. This assumption is equivalent to the “small-angle approximation” for both the off-axis distance on the viewing screen as seen from the apertures, as well as the transverse size of the apertures as seen from the viewing screen. In these limits, the intensity on the viewing screen as a function of the off-axis distance for plane-wave illumination is given by

$$I(y) = \frac{1}{2} \sin^2 \left( \frac{ka}{2s} y \right) \cos^2 \left( \frac{kd}{2s} y \right) \left[ 1 + \cos \left( \frac{kd}{s} y \right) \right],$$

where $k = 2\pi/\lambda$ is the wave vector magnitude, $\sin(x) \equiv \sin(x)/x$, and we have normalized the intensity to have a (dimensionless) value of 1 at $y = 0$.

2.2. Extended source

If the source has finite extent, it can be written as a summation of point sources over the spatial extent of the source. We are interested in calculating the intensity pattern in the case that different points on the source emit independently of one another, i.e. the extended source is a collection of incoherent point sources. In this situation, the composite intensity pattern on the viewing screen is determined by adding the intensities of the patterns from each portion of the source (for coherent source one adds the fields; for incoherent sources one adds the intensities). We assume
a rectangular-sized source of width $b$ transverse to the optical axis (see Fig. 1) and effectively infinite in the direction perpendicular to plane of Fig. 1. Thus, both the source and the slit are quasi one-dimensional, which leads to a one-dimensional intensity pattern in the far field that depends only on $y$ in Fig. 1.

2.2.1. Standard model

Derivations of the standard model are available in the literature (e.g., see Refs. [2,3,15]) and will not be repeated here; we simply quote the result. Assuming the intensity across the two slits is the same, the double-slit intensity distribution is given by

$$I(y) = 2I^0(y) \left[ 1 + |\gamma_{12}| \cos \left( \frac{kd}{s} y + \beta_{12} \right) \right], \quad (2)$$

where $I^0(y)$ is the diffracted intensity distribution due to one of the slits by itself (either one, since they are assumed equal), $\gamma_{12}$ is the “complex degree of spatial coherence” between the two points 1 and 2 (the locations of the slits), and $\beta_{12}$ is the phase that arises due to the sign of $\gamma_{12}$ (not due to any path-length difference).

Equation (2) is a very general result that can be applied in many different contexts. For rectangular slits, the function $I^0(y)$ is the single-slit (of width $a$) intensity distribution $\text{sinc}^2(ka/d)$. To calculate $\gamma_{12}$ one typically makes use of the van Cittert-Zernike theorem, which states that when $d, b \ll \ell$, the complex degree of coherence is equal to the normalized Fourier transform of the irradiance distribution across the source [5]. For a rectangular source of width $b$, we find

$$\gamma_{12} = \text{sinc}(k\phi d/2), \quad \text{where } \phi = 2 \arctan(b/2\ell) \approx b/\ell \text{ is the angular source size as seen from the double-slit mask.}$$

Putting this all together, we find (using the standard normalization)

$$I(y) = \frac{1}{2} \text{sinc}^2 \left( \frac{ka}{2s} \right) \left[ 1 + \text{sinc} \left( \frac{k\phi d}{2} \right) \cos \left( \frac{kd}{s} y \right) \right]. \quad (3)$$

Equation (3) is the standard model for rectangular slits and a rectangular source, and predicts a fringe visibility of

$$V \equiv \frac{I_{\text{max}} - I_{\text{min}}}{I_{\text{max}} + I_{\text{min}}} = \left| \text{sinc} \left( \frac{k\phi d}{2} \right) \right|. \quad (4)$$

where $I_{\text{max}}$ and $I_{\text{min}}$ are the intensities of adjacent maxima and minima. Note that this visibility function periodically goes to zero whenever $\phi = n\lambda/d$, for integer $n$. As mentioned, the complex degree of spatial coherence $\gamma_{12}$ is a function of two points—the “locations” of the slits—and therefore ends up being a function of the distance $d$ between the slits. Importantly, this function is independent of the width $a$ of the slits. In this sense, the standard approach neglects the width of the slits, at least as far as the visibility function is concerned. Stated another way, Eq. (2) is only technically true when the aperture sizes are infinitesimally small (e.g., $a \to 0$). This limitation has been recognized, and more general models have been proposed [6,7]. However, precisely how these models compare to experimental data remains unexplored.

2.2.2. Full model

When the double-slits are not infinitely narrow (as is always the case experimentally), the analysis described above is no longer valid. The fundamental modification is that any theory must account for light coming from different spatial portions of the slits. Perhaps the most straightforward approach is to begin with the point-source intensity distribution for a double-slit mask that accounts for the finite width of the slits [given by Eq. (1)], and then add up the various contributions from the spatial extent of the source. As written, Eq. (1) is valid for a point source that resides on axis ($w = 0$) whose intensity distribution is centered on the point $y = 0$. For a point source that resides off-axis ($w \neq 0$), the intensity distribution will be centered at the point...
\( y = -(s/\ell)w \) (in the small-angle approximation). Therefore, to find the intensity distribution due to a finite incoherent source of width \( b \), we need to integrate the intensities for a series of point sources from \( w = -b/2 \) to \( w = +b/2 \). Now, the differential intensity \( dI \) will be proportional to a differential element \( dw \) of the source, so that

\[
I(y) = A \int_{-b/2}^{b/2} dw \frac{s^2}{k a^2} \left[ I_e(k y_+) + I_e(k y_-) - I_e(k y_+) - I_e(k y_-) \right],
\]

where \( A \) is an appropriate constant (if the height of the source were not constant, as in the case of a circular source, one would need to include a \( w \)-dependent factor inside the integral).

Equation (5) is equivalent to Eq. (7) in Ref. [6] (and essentially Eq. (5) in Ref. [7]); while not trivial (and with a little help from \textsc{Mathematica}), this integral can be computed in terms of well-known functions [16]. The (unnormalized) result is

\[
I(y) = \frac{s^2}{k a^2} \left[ I_e(k y_+) + I_e(k y_-) - I_e(k y_+) - I_e(k y_-) \right],
\]

where \( y = y \pm b s/2 \ell \approx y \pm \phi s/2 \) are the locations of the images of the source edges (in the viewing plane), and the functions \( I_e \) and \( I_s \) are given by

\[
I_e(x) \equiv \frac{1}{x} \left[ 1 + \cos(\bar{d}x) - \cos(\bar{a}x) - \frac{1}{2} \cos \left[ (\bar{d} - \bar{a})x \right] - \frac{1}{2} \cos \left[ (\bar{d} + \bar{a})x \right] \right]
\]

and

\[
I_s(x) \equiv \bar{d} \text{Si}(\bar{d}x) - \bar{a} \text{Si}(\bar{a}x) - (\bar{d} - \bar{a}) \text{Si} \left[ (\bar{d} - \bar{a})x \right] - (\bar{d} + \bar{a}) \text{Si} \left[ (\bar{d} + \bar{a})x \right],
\]

with \( \bar{d} = d/s, \bar{a} = a/s \), and

\[
\text{Si}(x) \equiv \int_0^x \frac{\sin t}{t} dt
\]

the sine integral function.

While Eq. (6) represents an exact prediction for the intensity distribution in terms of well-known functions, this result is rather complicated and is not terribly illuminating from a physics perspective. For example, there is no obvious way of determining the fringe spacing or visibility from Eq. (6). On the other hand, the standard result in Eq. (2) is much more useful in the sense that each portion of this equation conveys useful physical information. However, as we discuss in the next section, Eq. (6) agrees much more closely with experimental data.

3. **Experimental results**

The experimental arrangement is shown in Fig. 1. The partially-coherent source is created by expanding a helium-neon laser beam (\( \lambda = 632.8 \text{ nm} \)) and directing it onto a rotating white index card. An adjustable slit is placed immediately after the card and controls the transverse width of the extended source. The double-slit aperture is located a distance \( \ell \) away, implying a source of angular width \( \phi \approx b/\ell \). The distance \( \ell \) was typically 0.80 m, and \( b \) ranged from 0.15–2.00 mm.

The aperture mask has two slits of width \( a \), with a center-to-center separation of \( d \) (for the experimental results presented, \( a = 150 \mu\text{m} \) and \( d = 600 \mu\text{m} \)). A lens (\( f = 1 \text{ m} \)) is placed immediately after the mask to bring the Fraunhofer pattern to an accessible distance and the intensity pattern is recorded a distance \( s \) (nominally \( f \)) away from the lens. We use a single-photon, avalanche photodiode (APD) to count photons as the detector is scanned in the \( y \)-direction across the interference pattern. The count rates depend on the width of the adjustable slit (and any filtering), so the integration time is adjusted between scans to maintain a similar count rate each time. Typically, count rates in the hundreds of thousands per second are recorded at the maximum of an interference pattern.
Figure 2 shows the experimental data, along with predictions using the standard model of Eq. (3). Before plotting the data, we subtracted the background, defined \( y = 0 \) to be at the central fringe (maximum or minimum), and normalized the data. Using this approach, there are no adjustable parameters in the model; the theoretical prediction is a “zero-parameter fit.” As seen in the figure, the standard model agrees fairly well with the data when the visibility is large. However, as the visibility decreases, discrepancies begin to appear in both the fringe depth [panels (c) and (f)] and spacing [panels (b) and (e)]. The disagreement is particularly pronounced at the lowest visibilities, where the standard model predicts little to no modulation. As clearly seen here (and previously noted), fringes in the experimental data are always observed.

We can gain some insight into the lack of fringe disappearance by noting that the fringes are predicted to disappear, according to Eq. (4), for angular source sizes of \( \phi = n\lambda/d \). Taking \( n = 1 \), we see that this condition is such that the fringe pattern due to a point source at the edge of the source (\( w = -b/2 \)) is shifted so that it is centered at a position \( y \approx s\phi/2 = s\lambda/2d \). But this is precisely the location of the first minimum for the fringe pattern due to a point source at \( w = 0 \), using the condition \( d\sin\theta = \lambda/2 \) along with the small-angle approximation \( \theta \approx y/s \). In other words, the disappearance of fringes, according to the standard model, is a result of the fringe patterns from source points at \( w = 0 \) and \( w = -b/2 \) being exactly out of phase. Similarly, the fringe pattern from a source point just above \( w = 0 \) will be out of phase with that due to a source point just above \( w = -b/2 \), and so on, so that all across the source intensity pairs effectively
cancel out. Of course intensities are non-negative and do not sum to zero; instead, it is just
the fringes that cancel so that we are left with something that looks like a single-slit diffraction
pattern.

The above argument relies on the fact that there is a single value \( d \) for the slit spacing. In
reality, there are locations within the slits that are separated by distances ranging from \( d - a \) to
\( d + a \), and this range of widths gives rise to a range of coherence values and thus a range of
visibilities. Therefore, the true situation is one in which this range of visibilities plays a role in
determining the final interference pattern, with the end result being that the fringes never fully
disappear.

In Fig. 3 we compare the same experimental data shown in Fig. 2 to the (normalized) full-model
prediction of Eq. (6). The agreement here is excellent, all the way down to the lowest visibilities.
It is worth emphasizing that these predictions are not fits to the data. After subtracting the
background, defining \( y = 0 \) at the central fringe, and normalizing, there are no adjustable
parameters in Eq. (6).

4. Discussion

In addition to the discrepancies at low fringe visibility, it is also clear that the fringe locations
predicted by the standard model are not fully accurate. This deviation is most obvious in panels
(b) and (e) of Fig. 2. When viewed from the perspective that the slit spacing is what determines
the fringe positions, this inconsistency is not surprising. Again, there are a range of “spacings” between points within the slits, so presumably there will be some complicated averaging of different fringe positions, similar to what occurs for fringe visibilities. To examine this behavior in more detail, we simultaneously plotted predictions from the standard model and the full model for a range of source sizes and compiled these images into a video. Figure 4 shows an image from this video in which the fringe positions are clearly different between the two models. For this source size ($\phi = 0.97$ mrad), both models predict a maximum at $y = 0$, but as you move out from this point the fringe positions begin to differ, becoming almost completely out of phase by the third fringe. Interestingly, the fringe positions predicted by the two models are exactly the same whenever the visibility is at a local maximum. In addition, it is clear from the video that the positions of the outer fringes as predicted by the full model move steadily outward as the source size increases.

As previously mentioned, Eq. (2) can be applied to many different situations. In fact, one of the most useful features of the standard model is how simple it is to update for a new arrangement. In particular, we investigated the case of a circular source, swapping out the variable slit for a pinhole in our experiment. Making use of the van Cittert-Zernike theorem again, the complex degree of spatial coherence is simply the Fourier transform of a circle, leading to an intensity distribution in the standard model of

$$I(y) = \frac{1}{2} \sin^2 \left( \frac{ky}{2s} \right) \left[ 1 + \operatorname{Bessinc} \left( \frac{k\phi d}{2} \right) \cos \left( \frac{kd}{s} y \right) \right],$$

where, once again, $C$ is an appropriate constant. Unfortunately, we were unable to reduce this integral to well-known functions, so it was necessary to evaluate it numerically.

As previously mentioned, Eq. (2) can be applied to many different situations. In fact, one of the most useful features of the standard model is how simple it is to update for a new arrangement. In particular, we investigated the case of a circular source, swapping out the variable slit for a pinhole in our experiment. Making use of the van Cittert-Zernike theorem again, the complex degree of spatial coherence is simply the Fourier transform of a circle, leading to an intensity distribution in the standard model of

$$I(y) = \frac{1}{2} \sin^2 \left( \frac{ky}{2s} \right) \left[ 1 + \operatorname{Bessinc} \left( \frac{k\phi d}{2} \right) \cos \left( \frac{kd}{s} y \right) \right],$$

where, once again, $C$ is an appropriate constant. Unfortunately, we were unable to reduce this integral to well-known functions, so it was necessary to evaluate it numerically.

As previously mentioned, Eq. (2) can be applied to many different situations. In fact, one of the most useful features of the standard model is how simple it is to update for a new arrangement. In particular, we investigated the case of a circular source, swapping out the variable slit for a pinhole in our experiment. Making use of the van Cittert-Zernike theorem again, the complex degree of spatial coherence is simply the Fourier transform of a circle, leading to an intensity distribution in the standard model of

$$I(y) = \frac{1}{2} \sin^2 \left( \frac{ky}{2s} \right) \left[ 1 + \operatorname{Bessinc} \left( \frac{k\phi d}{2} \right) \cos \left( \frac{kd}{s} y \right) \right],$$

where, once again, $C$ is an appropriate constant. Unfortunately, we were unable to reduce this integral to well-known functions, so it was necessary to evaluate it numerically.

As previously mentioned, Eq. (2) can be applied to many different situations. In fact, one of the most useful features of the standard model is how simple it is to update for a new arrangement. In particular, we investigated the case of a circular source, swapping out the variable slit for a pinhole in our experiment. Making use of the van Cittert-Zernike theorem again, the complex degree of spatial coherence is simply the Fourier transform of a circle, leading to an intensity distribution in the standard model of

$$I(y) = \frac{1}{2} \sin^2 \left( \frac{ky}{2s} \right) \left[ 1 + \operatorname{Bessinc} \left( \frac{k\phi d}{2} \right) \cos \left( \frac{kd}{s} y \right) \right],$$

where, once again, $C$ is an appropriate constant. Unfortunately, we were unable to reduce this integral to well-known functions, so it was necessary to evaluate it numerically.
The results for a circular source follow almost exactly the results for a rectangular source. The standard model does a reasonably good job of predicting the experimental data when the visibility is high, with clear discrepancies when the visibility is low. Once again, the full model does an excellent job of predicting the data for all visibilities. A sample of our results is shown in Fig. 5 for a source diameter of 1 mm ($\phi = 1.25$ mrad). Similar to what is observed in the rectangular slit situation, the standard model here predicts essentially zero fringe visibility for this source size, while the experimental data (and the full-model prediction) still shows fringes.

![Fig. 5. Sample data and predictions for a circular source of angular size $\phi = 1.25$ mrad. The full model (red) matches the data much more closely than the standard model (blue).](image)

5. Conclusions

We present quantitative measurements demonstrating slit-width effects in a double-slit experiment with a partially-coherent source. As expected, the experimental results do not fully agree with the standard model, particularly at low visibilities. A theoretical prediction was derived that agrees much more favorably with the data. The data presented here grew out of a senior-level undergraduate project, and we believe such an experiment would work well as either a senior project or as part of an undergraduate advanced lab or upper-level optics course [17]. Finally, we note that our approach using a rotating card source and a scanning APD detector is only one of many possible implementations (see, e.g., Refs. [8–13] for a variety of different approaches).

References

14. Often, the viewing screen is placed in the focal plane of a lens located immediately after the double-slit to ensure the viewing screen is in the far-field, as seen in Fig. 1.
16. We note that the integral can, of course, also be done numerically, and this may be a more instructive approach for students.