Single-slit diffraction and Babinet's principle

Phys 294/5 Experiment 3

Student's manual

June 24, 2020

University of Alberta, Department of Physics

1 Theory

We will be studying far field (Fraunhofer) diffraction of light as light encounters a thin barrier, for example, a thin wire-like structure. To start, we first consider the diffraction of light through a slit, which is exactly the opposite of light encountering a wire. As we will see later, these two situations are basically the same. The schematic of single-slit diffraction is shown in Figure 1. When light passes through the (small enough) slit on the left side, an intensity pattern of light is projected on a screen placed at the right of the slit. This intensity pattern arises from the bending of light as it passes through the slit, which makes light to interfere with itself at the projection screen. Think of the incoming light as made of different light rays, all exactly the same and parallel to each other. Then theses light rays are bent as they pass through the slit. The intensity pattern shows points of minima (where there is no light reflected off the projection screen) and points of maxima (where the light is brightest compared to the minima immediately located around them). Points of minima occur when light interferes with itself *destructively*, and points of maxima occur when light interferes with itself *constructively*.
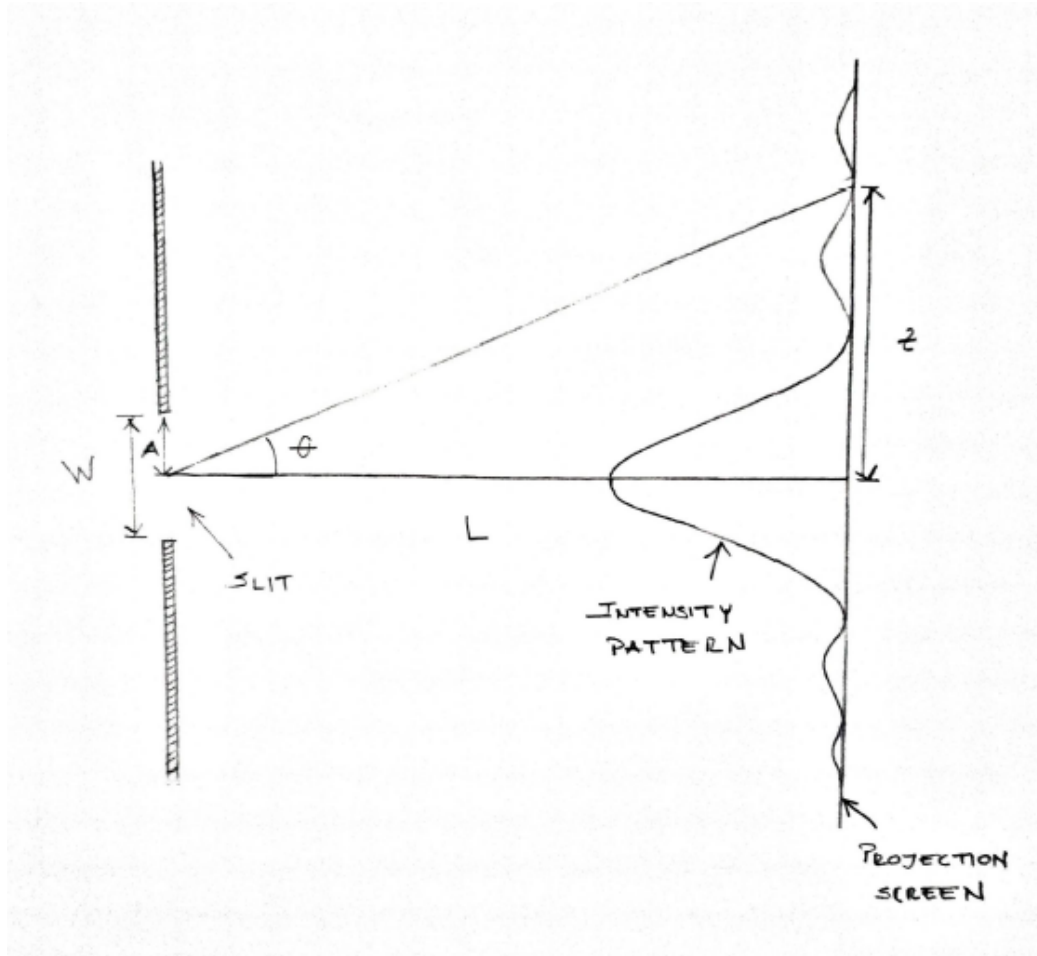
Figure 1: Far field single-slit diffraction. Not to scale for illustrative purposes. Based on Figure 1, Section 4 in Meldrum [3].

In the wave model of light, light is described by its intensity (related to its electric field amplitude), frequency (related to how fast it oscillates, which determines its color) and phase (which is the information on the general displacement of the wave with respect to a reference point in space, so the phase is defined up to a constant). Differences in phases, arising from the bending of light going through the slit, are the source of interference, which can be constructive or destructive, as mentioned above. The type of interference and conditions under which it happens depend on several factors, as we will see later.

Far field diffraction occurs when the distance L in Figure 1 between the slit and the projection screen is big enough that the light rays (although bent at the slit) are considered to arrive approximately parallel to each other at the

screen. This allows us to worry only about the phase differences between the light rays and to consider that all bent light rays travelled the same distance to a point on the screen; this will become clearer shortly. For far field single-slit diffraction, the condition for a minimum to occur in the projection screen is the following:

$$W sin(\theta) = m\lambda, \tag{1}$$

where W is the width of the slit, θ is the angle between the line perpendicular to the projection screen and the light ray, m is an integer number larger than zero labelling the points of minima (i.e., 1 for the first minima, 2 for the second minima, etc., starting the counting from the center of the intensity pattern), λ is the wavelength of light, and L is the distance between the slit and the screen. We will derive this equation shortly, but first we will look at the near field (Fresnel) diffraction situation; this is when the aperture and the projection screen are close together, so that the light rays arriving at the projection screen can no longer be considered parallel to each other.

Huygens (1629-1695). Huygens' principle was built with the assumption that light was a wave and, as a wave, that it needed a medium to propagate, like any other wave does. The medium was thought to be the ether, although this was later ruled out by the Michelson-Morley experiment, which demonstrated that no ether existed and that light can propagate in vacuum. Nevertheless, Huygens' model with the ether in it is still very useful. His idea was that light was a series of wave-fronts originated at each point of a luminous body, and that each point on the wave-fronts was a disturbance in the ether, capable of generating new light wave-fronts. [4] See Figure 2. Think of the light rays mentioned before as lines bisecting the wave-fronts (if the ether existed, these would be lines cutting through the ether, just like boats cut through the water).
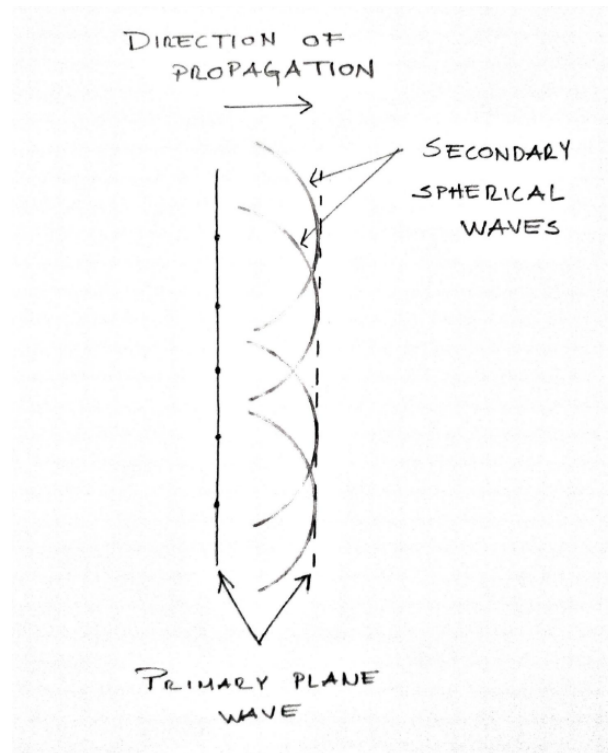
(a)
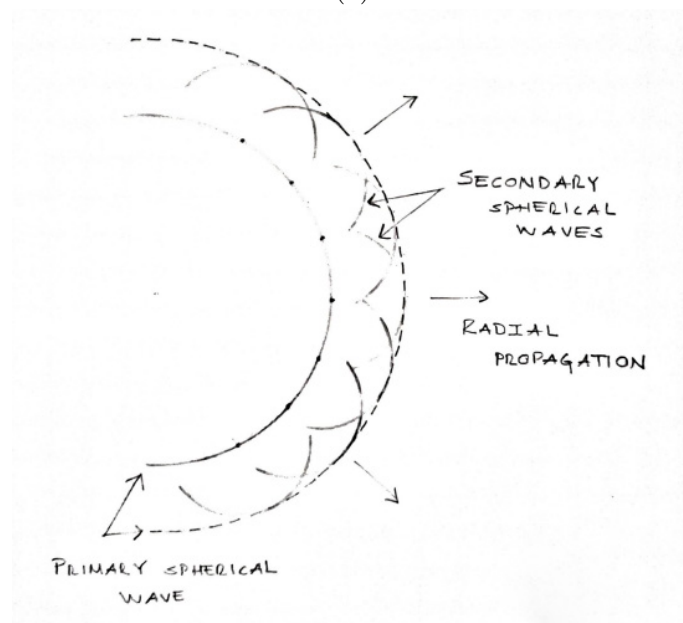
(b)

Figure 2: Huygens' principle for (a) plane wave and (b) spherical wave. Based on Figure 2-2 in Pedrotti[3] [4].

Now, Huygens' principle assumes that the amplitudes of the secondary wave-fronts arising at the surface of the primary wave-front are the same in all directions. This is not the case, however; there is a directional dependence, called the obliquity factor, of secondary wave-front amplitudes. See Figure 3. This obliquity factor can be derived from the Kirchhoff diffraction integral.[1] [3] Read the following first, before attempting to understand the derivation in Hecht's optics. The obliquity factor is: [4]
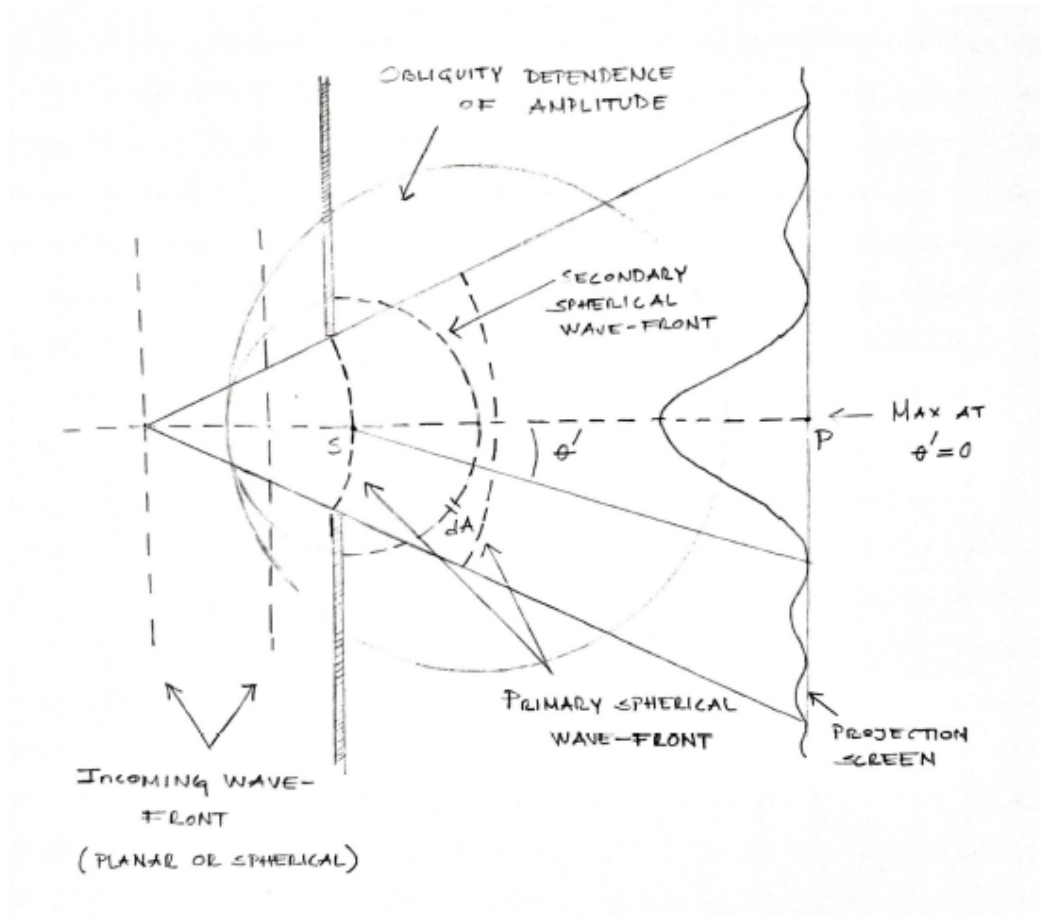

Figure 3: Spherical wave-front at slit. Not to scale for illustrative purposes; specially, the secondary wave-front was made exagerately large. Based on Figure 2-3 in Pedrotti[3] [4], Figure 3, Section 5 in Meldrum [3], and Figure 10.50 in Hecht [2].

$$F(\theta) = \frac{1}{2}(1 + cos(\theta')), \tag{2}$$

---

[1]For a nice derivation, I recommend reading section 10.4 in Hecht's optics.

where θ' is the angle between the plane bisecting a secondary spherical wave-front of light and some point in that plane P on the projection screen, as shown in Figure 3. The wave-fronts, as they pass through the slit, are spherical, and so they can be described with the mathematical expression of a harmonic spherical oscillation. For the following, refer to Figure 4. Consider a point-like light source of spherical wave-fronts located at a point S. The expression describing the light's electric field at a point at the surface of the wave-front after a time $t = t'$, from $t = 0$, is the following: [4]

$$E = \frac{E_0}{\rho} cos(\omega t' - k\rho), \tag{3}$$

where E_0 is the strength per unit area of electric field from the source S, ρ is the radius of this sphere, ω is the angular frequency, and k the wave number. E_0/ρ is the amplitude of the electric field at a distance ρ from the point S. Equation 3 is the field at any point on dA in Figure 4. Equation 3 assumes that all point sources within a dA element are coherent (i.e., all wavefronts from each point on dA, according to Huygens' principle, are exactly the same [same frequency and phase]); this means that any point P outside the sphere placed on a line passing through S, at $\theta' = 0$, will see the same properties on the light rays coming from all the points in a ring dA. Equation 3 also assumes that each of the points in dA radiates *secondary* light waves that are in phase with the *primary* light wave coming from S. All secondary wavelets generated at any of the rings dA travel a distance r to reach the point P at a time t, with the same phase $\phi = \omega t - k(\rho + r)$. Assuming that the field amplitude from the primary source S at dA, E_0/ρ, is proportional to the field strength per unit area of the secondary source dA, $E_S \propto E_0/\rho$, and taking the obliquity factor into account, we conclude that the field amplitude at the point P from the wave-front coming from dA is given by the modified expression for a harmonic spherical wave: [4]

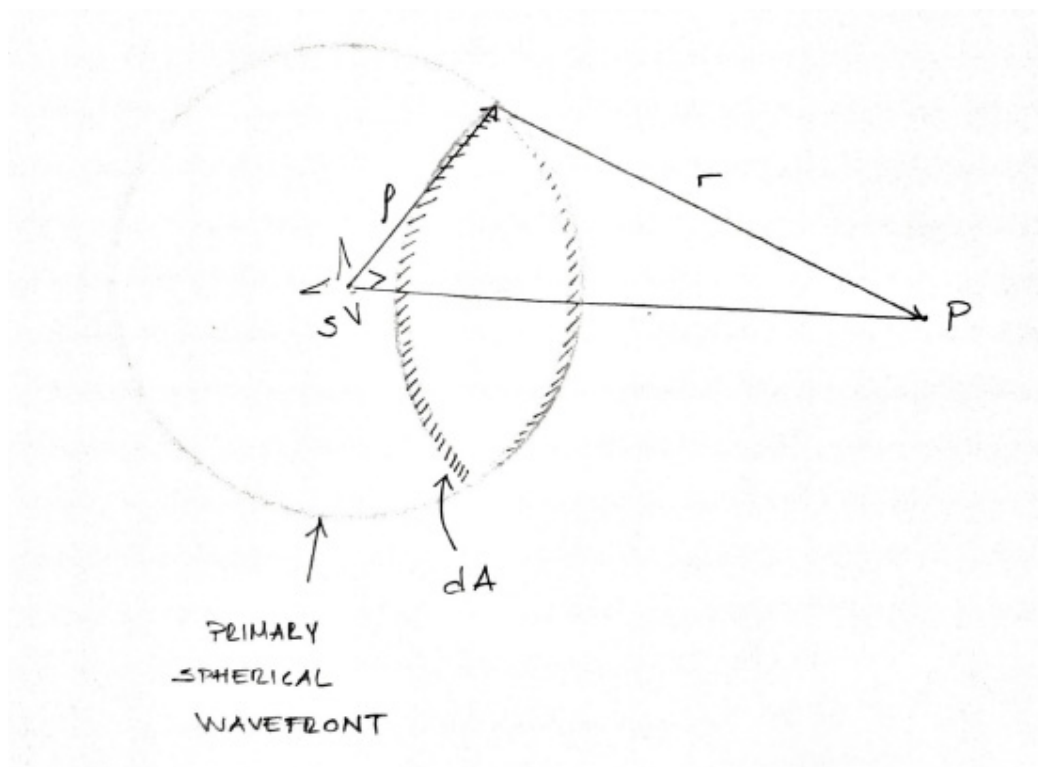$$dE = K(\theta)\frac{E_S}{r}cos(\omega t - k(\rho + r))dA. \tag{4}$$

Figure 4: Spherical wave-front, close look. Based on Figure 10.51 in Hecht [2].
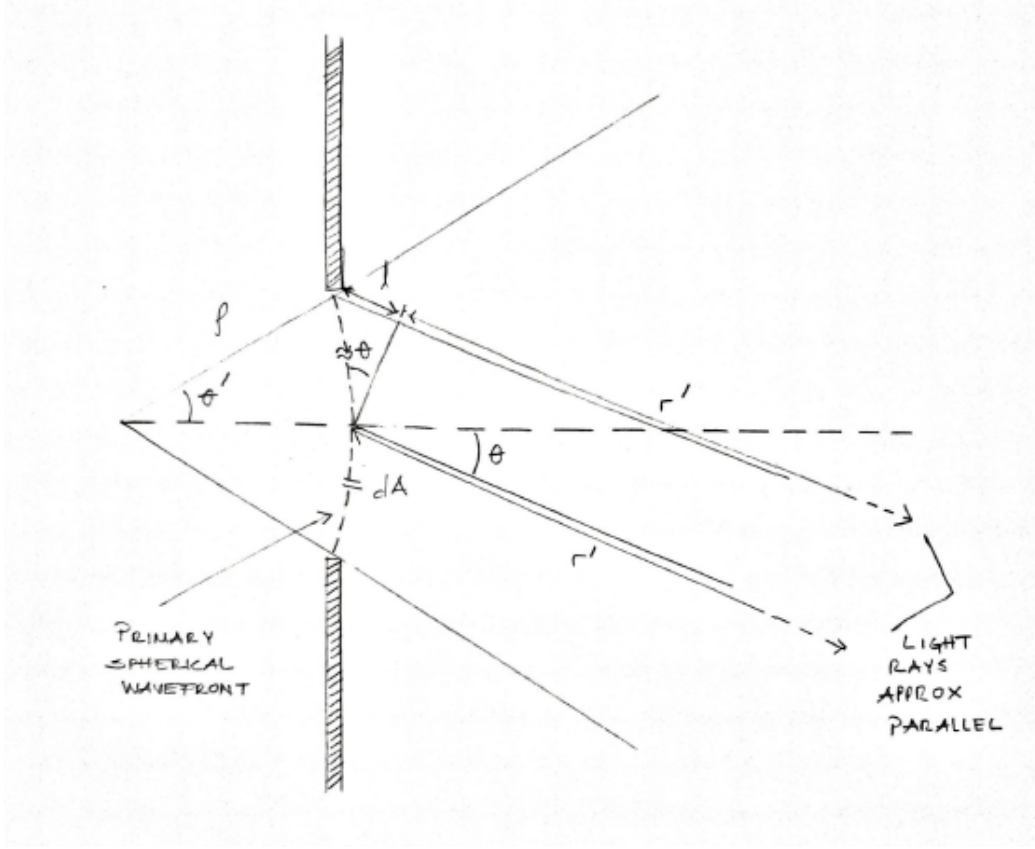
Figure 5: Far field single-slit diffraction approximations. Not to scale for illustrative purposes. Based on Figure 11-1 in Pedrotti[3] [4].

Now that we have this information, we can refer to Figure 5. The field at the point P can be described as

$$dE = K(\theta)\frac{E_S}{r}e^{i[\omega t-k(r+\rho)]}dA, r = r' + l, \tag{5}$$

where ρ is as before, the distance between the point of origin of the primary wave-front and point of origin of the secondary wave-front. In the far field limit, we can approximate $\rho \approx 0$, and so the elements dA lie approximately at the slit. Also, in the far field limit, the light rays at the slit to a point P in the projection screen are approximately parallel, so θ' and θ are very small: $F(\theta') \approx 1$, $l = Asin(\theta)$, $r = r' + l \approx r'$. See Figure 5. Then Equation 5 becomes:

$$dE \approx \frac{E_S}{r}e^{i[\omega t-k(r+Asin(\theta))]}dA. \tag{6}$$

8

We can ignore the curvature of the primary spherical wave-front at the aperture, since we are talking of a very small rectangular slit. Integrating both sides of Equation 6, we get:

$E = \frac{E_S}{r}e^{i(\omega t - kr)} \int_{-W/2}^{W/2} e^{ikA sin(\theta)} dA$

$E = \frac{E_S}{r}e^{i(\omega t - kr)} \frac{e^{ikW sin\theta/2} - e^{-ikW sin(\theta)/2}}{ik sin(\theta)}$

Define $\beta \equiv \frac{1}{2}kW sin(\theta) = kl$ (it is the phase difference between the light coming from a point dA in the slit and the center of the slit):

$E = \frac{E_S}{r}e^{i(\omega t - kr)} \frac{W(e^{i\beta} - e^{-i\beta})}{2i\beta}$

$E = \frac{E_S}{r}e^{i(\omega t - kr)} \frac{W}{\beta} sin(\beta).$

We take the real part of this expression, the one with the cosine term. Then the formula for the irradiance (which is proportional to the time average of the square of the electric field) at P is (recall that the average of the square of a cosine is 1/2 and that $sinc(\beta) \equiv sin(\beta)/\beta$):

$$I = \epsilon_0 c \langle E \rangle^2 = \epsilon_0 \frac{c}{2} \left(\frac{E_S W}{r} \right)^2 sinc^2(\beta), \qquad (7)$$

where ϵ_0 is the vacuum permittivity and c is the speed of light in vacuum. Here we have further assumed that r is constant and equal to the distance between the center of the slit and the point P; this is not too far away from reality in the far field limit, where the light rays arrive approximately parallel to each other at some point P in the projection screen. Defining $I_0 \equiv \epsilon_0 c \frac{1}{2}(\frac{E_S W}{r})^2$, we get:

$$I = I_0 sinc^2(\beta). \qquad (8)$$

Then the minima occur wherever $sinc^2(\beta) = 0$, that is, where $sin(\beta) = 0$, or $\beta = m\pi$, $m = 1, 2, 3, ...$ So $\beta = \frac{1}{2}kW sin(\theta) = m\pi$ for minima. Noticing that $2\pi/k = \lambda$, we obtain Equation 1:

$$W sin(\theta) = m\lambda. \qquad (9)$$

But in the far field limit, where θ is very small, $sin(\theta) \approx tan(\theta) \approx z/L$, where z is the distance from the center of the diffraction pattern to the mth minimum. See Figure 1. Then we can calculate the width, W, of the slit, with the condition for minima:

$$W = \frac{Lm\lambda}{z}. \tag{10}$$

This is for the far field diffraction of light passing through a single slit.

Babinet (1794-1872). From Babinet's principle. it can be concluded that the diffraction patterns from a single slit and its opposite, a thin wire of the same width, are exactly the same except at their center. These two situations, single slit and thin wire are called *complementary* apertures. Generalizing this, two diffraction screens are complementary when the transparent regions in one exactly correspond to the opaque regions in the other. [2] Babinet's principle says that the sum of the amplitudes at the projection screen from the two different diffraction screens alone equals the amplitude of the unobstructed light: [4]

$$E_1 + E_2 = E. \tag{11}$$

An important conclusion can be drawn from Equation 11: [1] if $E_1 = 0$, then $E_2 = E$. This means that at the points where the intensity on the projection screen is zero in the presence of the diffraction screen number one, in the presence of the diffraction screen number two the intensity is the same as if there was no diffraction screen present at all.[2]. Then where you would expect a minimum in the projection screen when using a single-slit diffraction screen, you will obtain a maximum in the projection screen when using a wire with the exact same width of the slit. So Equation 10 can be used to determine the with of a wire blocking a coherent source of light, **except that this time z will be the distance from the central maximum to the mth maximum in the diffraction pattern at the projection screen**.[3]

2   Experimental details

Use Babinet's principle to find the thickness of your hair, or the thin wire provided. The wavelength of the laser given to you is: $\lambda = 532 \pm 0.1$ nm (?). Since we are studying far field diffraction, make sure that the distance between the hair/wire and the wall onto which you are projecting the diffracted light is very large compared to the thickness of the hair/wire.

Perform the calculations for the different maxima that you see in the diffraction pattern on the wall. See if they all are consistent with each other.

---

[2]This is not true for the central maximum

[3]For a more rigorous derivation of this conclusion, see Pedrotti, Section 13-10, and Born & Wolf, Sections 8.3.2 and 11.3.

Then take the average of the calculations, that will give you the thickness of your hair/wire. Does the result seem to be reasonable? Show pictures of your work, and include uncertainties in your results.

References

[1] Born, M. & Wolf, E. (1980). Principles of optics: electromagnetic theory of propagation, interference and diffraction of light. 6th ed. with corrections. Pergamon Press.

[2] Hecht, E. (2017). Optics. 5th ed. Pearson: USA.

[3] Meldrum, A. (2018). Phys 362: Optics and Lasers. Fall term 2018 lecture notes, Sections 4 - 5. University of Alberta.

[4] Pedrotti, F. L.; Pedrotti, L. M.; Pedrotti, L. S. (2018). Introduction to optics. 3rd ed. Cambridge University Press: USA.