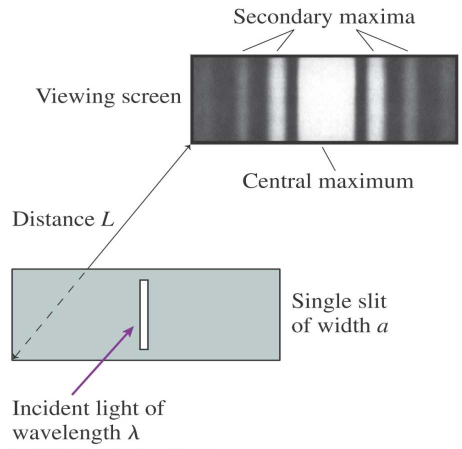
PHY 171
Lecture 9
(*February 6, 2012*)

Diffraction

Single Slit Diffraction

In class, we demonstrated that when monochromatic light from a laser is sent through a single slit, we get a pattern of bright and dark bands on a screen, as shown in the figure on the right. This phenomenon is known as diffraction.

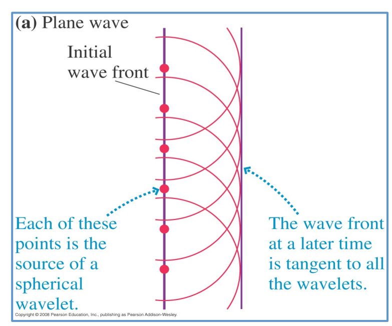


As we demonstrated in class, it is easy to do this experiment – place two fingers together and look toward a light source; if you vary the slit between the fingers just right, you will see alternate bright and dark fringes.

In this lecture, we will try to understand why diffraction occurs. We will assume that parallel rays of light fall on the slit, and pass through to fall on a viewing screen very far away (Fraunhofer diffraction) – if the screen is too near, use a lens to make the outgoing beam parallel.

The other case when the screen is close to the slit is called Fresnel diffraction, and its analysis is more involved (and will not be done here).

To understand why diffraction occurs, we will need to borrow the following idea from *Huygen's theory* of wave propagation. According to Huygen's theory, each point on a wavefront acts like a source of secondary waves. To find how far a wave moving with speed v has been displaced after time t, draw arcs of radius vt, then draw a larger arc that is the envelope of all these arcs (see figure on the right). This is the new wavefront. (Note that many authors have criticized this theory as being far too removed from reality, but it helps greatly in understanding diffraction conceptually, which is why we're using it).

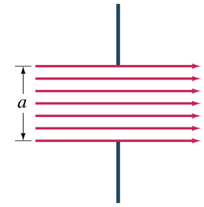


The key point to remember is that each point on the wave at the aperture (the single slit) is like a source of secondary waves. It is the interference between these waves that can be used to understand the cause of diffraction.

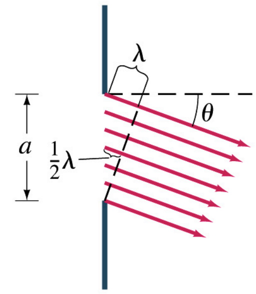
To avoid confusion with interference, we will designate the width of a slit by the letter "*a*."

Explanation of why we get bright and dark fringes in diffraction

See the figure on the right – rays passing straight through are all in phase, so they will form a *bright spot* on the screen. Therefore, we get a bright maximum at the center of the screen.

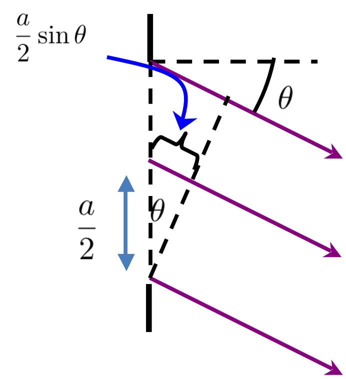


Next, see the figure on the right – rays going at an angle θ such that the ray from the top travels exactly one wavelength farther than the ray from the bottom edge.



In this case, a ray at the center of the slit will travel $\lambda/2$ farther than a ray at the bottom edge. Pairing off rays like this, we realize that all the rays in the top half travel $\lambda/2$ farther than the rays in the bottom half. Therefore, they interfere destructively, so no light will reach the viewing screen at this angle.

From the geometry of the figure, the path difference between a ray in the top half and a ray in the bottom half is then $(a/2)\sin\theta$; this is shown in the figure on the right, which is an enlarged version of the figure above. For destructive interference, this must be equal to $\lambda/2$. Therefore, setting $\lambda/2 = (a/2)\sin\theta$, then canceling 2 from each side gives

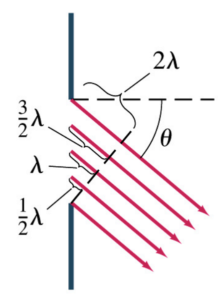


$$\lambda = a\sin\theta$$

This is then the condition for the first diffraction minimum.

This is then the condition for the first diffraction minimum.

To find the next minimum, let us move down the screen to a larger angle θ, where rays from the top travel 2λ farther than rays from the bottom of the slit, as shown in the figure on the right. Imagine dividing the slit into 4 parts. The rays in the bottom fourth will interfere destructively with the rays in the quarter just above them (third quarter), because there is a path difference of λ/2 between them. Likewise, rays in the second quarter (bottom quarter of the top half) will interfere destructively with the rays in the top quarter. Therefore, at this angle, we get again a minimum in the diffraction pattern.



From the geometry of the figure, the path difference between a ray in the bottom quarter and a ray in the third quarter is then $(a/4)\sin\theta$. For destructive interference, this must be equal to $\lambda/2$. Therefore, setting $\lambda/2 = (a/4)\sin\theta$, then canceling 2 from each side gives

$$a\sin\theta = 2\lambda$$

This is then the condition for the second diffraction minimum.

Continuing in this manner, we see that the diffraction minima will occur at

$$a \sin \theta = p\lambda$$

where $p = 1, 2, 3, \ldots,$

Note, however, that p is not zero; because the center always has a maximum.

Caution: To a casual observer, the *minima of a diffraction pattern* might appear to satisfy a condition that looks very similar to the *maxima of a 2-slit interference* pattern. Upon careful examination, you will realize this is not true, but that the two conditions are truly different. In the diffraction expression, we have "*a*" which is the *width of a slit*. In the interference expression, we have "*d*" which is the *distance between slits*. More important, the condition for destructive interference is still an odd integer multiple of a half-wavelength; it is only because of cancellation of terms on both sides that the expression for diffraction minima takes the final form written above.

Next, diffraction maxima other than the central maximum are located approximately midway between the diffraction minima. So, the first order maximum away from the center can be located by averaging the $p = 1$ and $p = 2$ expressions, which gives

$$(\lambda + 2\lambda)/2 = 3\lambda/2.$$

So, we get the condition for the first diffraction maximum away from the center to be

$$a \sin \theta \approx 3\lambda/2$$

Does this make sense? Consider an angle θ such that the top ray travels 3/2 λ farther than the bottom ray. Divide the slit into 3 parts. Rays from the bottom third will then cancel with rays from the middle third because they will be $\lambda/2$ out of phase. However, light from the top third will still reach the screen, so there will be a bright spot centered near

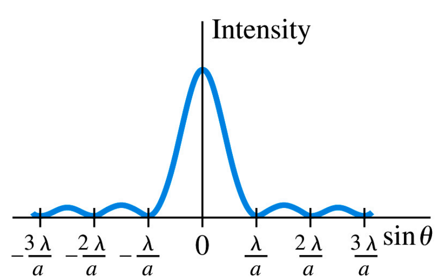
$$\sin \theta \approx 3\lambda/2a$$

So, we do get a maximum with the condition predicted above. Note also that since only 1/3 of rays coming through the slit are responsible for forming this maximum, it will not be as bright as the central spot, a fact which we have seen already in the demo.

Likewise, we can predict the positions of the other maxima by averaging the two nearest minima.

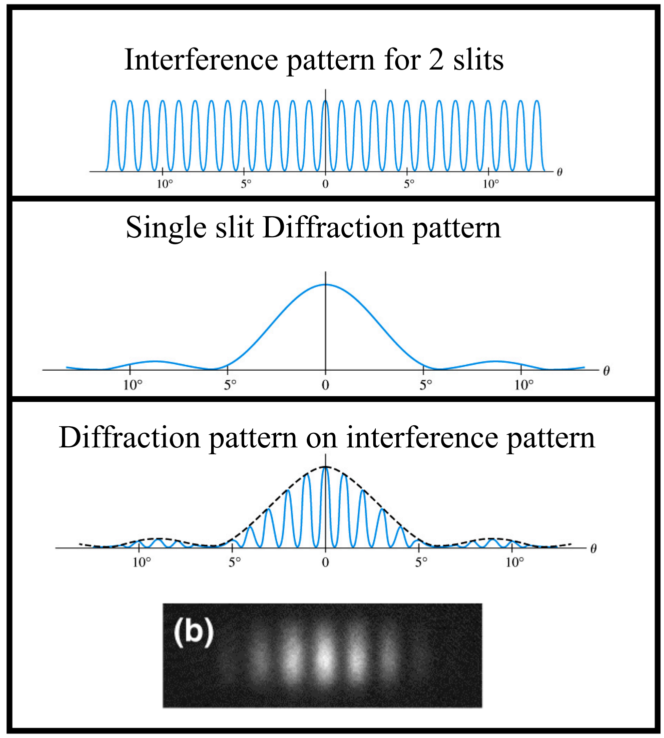
Thus we have shown that the diffraction pattern has alternate bright and dark bands.

On the right is a figure showing a graph of the intensity pattern for single slit diffraction plotted against sin θ.



Diffraction in the Double Slit Experiment

Recall the demo of the two-slit interference experiment, where you saw that the intensity falls off as you move away from the center of the pattern. We said that we would not worry about this decrease in intensity when we studied interference, for which the intensity expression predicts the same peak intensity as you move away from the central maximum. That is, if you plot intensity *vs.* θ in a double slit experiment, you should get a plot like that shown in the top panel below (labeled "Interference pattern for 2 slits").



In reality, though, the diffraction pattern of the single slit is superimposed on the interference pattern. The diffraction envelope is shown in the middle panel of the figure on the right (labeled "Single Slit Diffraction pattern").

Notice that the width of the central diffraction maximum (see middle panel) is much wider than the width of the central interference maximum (see top panel). This makes sense, since the first diffraction minimum is located at $\sin\theta_D = \pm\lambda/a$ (where a is the width of a single slit), whereas the first interference minimum is located at $\sin\theta_{INT} = \pm\lambda/2d$ (where d is the distance between the two slits). Since the distance between slits (d) is obviously much larger than the slit width (a), we will get θ_{INT} to be much smaller than θ_D, so the width seen in the graphs is as expected.

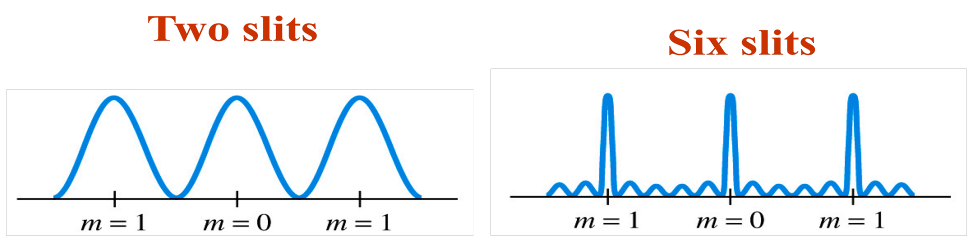
The diffraction pattern acts as an envelope that limits the interference peaks. When we combine the top panel of the figure above with the middle panel, the result will be the bottom panel. This is the two-slit interference pattern we see in reality, with a falloff in the intensity of the interference peaks as we move away from the central interference maximum. A photograph of the true observed 2-slit interference pattern is shown in the bottom of the third panel of the figure above.

Diffraction Grating

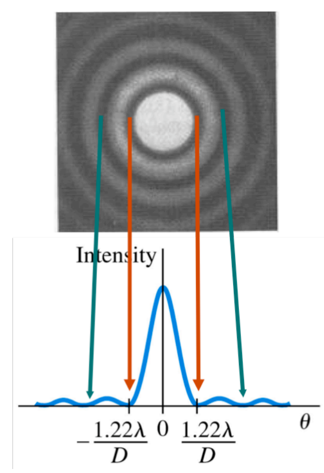
A diffraction grating contains a large number of equally spaced parallel slits, and should more properly be called an interference grating. The condition for maxima is the same as for 2-slit interference, namely $d\sin\theta = m\lambda$, where $m = 0, 1, 2, 3, \ldots$, but the interference peaks are now sharper and narrower.

> Why? In 2-slit interference, when you move away from the maxima point on the screen by a small angle θ, the two waves will be only slight out of phase, so they still interfere constructively. In a diffraction grating, even a slight movement away in angle on the screen means that destructive interference occurs for some pairs of slits, e.g., 1 & 1001, 2 & 1002, etc.

The figure below shows the interference pattern for 2 slits compared to the interference pattern for 6 slits.



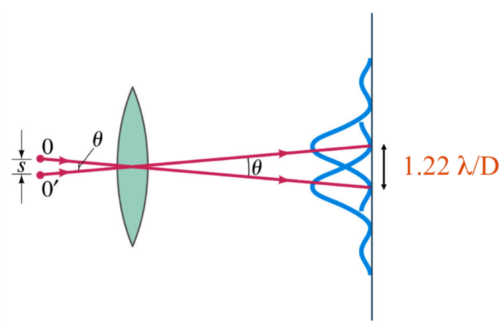
Diffraction through a Circular Aperture



Diffraction through a circular aperture is important because the eye and many optical instruments like telescopes are circular apertures. The diffraction pattern through a circular aperture is shown in the figure on the right. The intensity pattern corresponding to this diffraction pattern is shown below on the right. The first diffraction minima occur at 1.22 λ/D, where D is the diameter of the circular aperture.

This has important consequences. Angular or spatial resolution is the ability to distinguish detail (i.e., the ability to figure out that two nearby objects are really two and not perceived as being blended into one).

Angular (or spatial) resolution is usually characterized by the *Rayleigh criterion*, which states that two images are just resolvable when the central maximum of the diffraction pattern of one image is directly over the first minimum of the diffraction pattern of the other image – see the figure on the right. This gives us an expression for the angular resolution:



$$\theta = 1.22\frac{\lambda}{D}$$

If two points are closer than this, they will appear merged as one. Recall the picture of the field of stars we looked at in class, taken with a low and high angular resolution telescope respectively. Of course, other factors may dominate over this diffraction limit. For example, turbulence in the atmosphere is the limiting criterion for Earth-based telescopes. That's why the Hubble telescope with its ~2 m mirror can be better than Keck with its 10 m mirror.