Physics 228, Lecture 5
Thursday, Feb. 3, 2005

Diffraction. Ch 36:1-4, 7;

1 Diffraction in general

Huygens' principle has told us that waves do not strictly travel in straight lines as particles without forces on them do, and as we assumed when we were using geometrical (that is ray) optics. Parts of the waves will spread out from straight paths, in a process called diffraction. In this lecture we will investigate diffraction

First of all, a comment about the edges of shadows. If we have a point source of light and an opaque object with a straight edge, geometrical optics would say the shadow would have a straight line dividing a black region from a lit region. If we actually look closely at the dividing line, however, we find that there are fringes which fade into an even brightness. This is a diffraction pattern, and can be understood only in terms of the propagation of waves. Once your eyes are sufficiently dark adjusted, and now that you know what to look for, look at the shadow of a razor blade on the right side of the screen.

S&BV5 Fig. 38.2
Edge Diffraction

A really spectacular consequence is the shadow of a disk. Look at the shadow of a disk on the left of the screen. You see the geometrical circular dark area, but you also see fringes near the border. More surprizingly, if you look very closely you should be able to see a bright spot at the center! This is known as Poisson's spot, and there is a lovely story about that[1].

[1] In 1818 the French Academy sponsored a competition, to which Fresnel submitted an erudite paper on calculational methods for the wave theory of light. Poisson, who was one of the judges and who hated the wave theory, was dead against Fresnel getting the prize, and to shoot down this possibility, he used Fresnel's theory to show that a circular object should cast a shadow with a bright spot in the middle. This, he said, is absurd, so we can't possibily award the prize to Fresnel. But another judge, Arago, decided to test the prediction of the spot, and found that indeed it was there! So Fresnel got the prize, but curiously the spot got named for Poisson, "like a curse".

We are going to analyze a few cases involving slits, and then we will describe the results for circular apertures, which are important but more difficult to analyze. Previously in analyzing slits we treated them as infinitesimally wide, but now we will consider their width. In all cases, it makes the analysis easier if we can treat the interfering rays as coming in parallel through the slits or aperatures. There are two ways this can be appropriate. Either the screen can be very far away compared to the extent of the aperature, or we can use a converging lens with the screen at its focal point. Then the rays of light which converge at a point on the screen are those which came through the slit in parallel. When the situation satisfies this requirement, the result is called **Fraunhofer diffraction**.

Demo adj slit, start wide, get narror

S&BV5 Fig 38.4. Single slit diffraction
3"× 4 3/8", plot, no photo

We have a slit a fraction of a millimeter wide in the back of the room, through which a laser shines light at the screen. Clearly the rays hitting any point of the screen are very nearly parallel, even though they do intersect at that point of the screen. So we have Fraunhofer diffraction here. Notice the pattern, with a bright fringe in the middle, but also more fringes, much fainter, on either side.

2 Diffraction from a single slit

We will first analyze the diffraction from a single slit of width a. To find the intensity at a point P at an angle θ, we consider the extra path length that each ray must travel relative to the ray at the top which travels the shortest distance.

S&BV5 Fig 38.5. single slit split in two.
(3 1/2" × 4")

Before we consider the situation in general, let's consider the particular angle θ for which the bottom ray is exactly one wavelength further from the screen than the top ray. As that extra distance is $a \sin\theta$, we are talking about $\sin\theta = \lambda/a$. The ray which comes from the middle of the slit will then have travelled exactly one half wavelength more than the ray from the top, so it will be exactly 180° out of phase, and will interfere destructively with that ray. That is only one pair of rays which cancel, but in fact the same is true for every ray which passes through the upper half of the slit — it is exactly

180° out of phase with one ray which passed $a/2$ below it, so this pair cancels as well. And when we have counted all the rays in the upper half and their partners $a/2$ below, we have counted them all, so must have a dark spot at this angle. At this angle the top half of the slit cancels the bottom half.

In fact, if we divide the slit into any even number of equal slices and arrange the angle so that successive slices give phase shifts differing by 180°, they will cancel. So the condition for a dark fringe is

$$\text{Dark:} \qquad \sin\theta = m\frac{\lambda}{a} \quad \text{for } m \text{ an integer other than 0.}$$

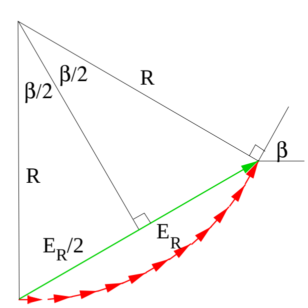
How can we analyze quantitatively the intensity as a function of angle? Let β be the net phase shift between the lowest and highest rays, so

$$\beta = 2\pi\frac{a\sin\theta}{\lambda}.$$

We can imagine the full slit as composed of a large number N of evenly spaced slits, a situation we met before. Each little slit will contribute a field E_0 proportional to its width, so inversely proportional to N,

$$E_0 = \frac{A}{N},$$

and the phase difference between successive little slits is $\phi = \beta/N$. The real problem, that is a single slit of width a, is the limit that we have an infinite number of infinitesimally thin slits, so we want to take the limit as $N \to \infty$. We recall that for small angles ϕ the phasors lined up for addition along a circle, so they add up to an arc. The length of the arc is just $NE_0 = A$, a constant independent of the angle θ (or β) and of how many imaginary slits we used to subdivide our problem. The last phasor is at an angle β, so the arc subtends an angle β as well. Thus the length of the arc is also given by the radius times β, $R\beta = A$. The bisector of the isosceles triange shows us that $R\sin(\beta/2) = E_R/2$, so



$$E_R = 2R\sin(\beta/2) = A\,\frac{\sin(\beta/2)}{\beta/2},$$

and the intensity, which is proportional to E_R^2, is given by

$$I = I_{\max}\left(\frac{\sin(\beta/2)}{\beta/2}\right)^2 = I_{\max}\left(\frac{\sin(\pi a\sin\theta/\lambda)}{\pi a\sin\theta/\lambda}\right)^2.$$

We can check that we get zero at $\sin\theta = m\lambda/a$, as our qualitative argument predicted.

Note that the calculation we did here is very similar to what we did for N slits last time, except that we took a limit $N\to\infty$ with $a = Nd$ fixed. The slide show from last time just needs to be changed to display the total phase difference from one side of the slit to the other, rather than between neighboring slits (which loose their distinction as $N\to\infty$) and to concentrate on a smaller total phase difference. Here it is.

Show slide show

stop at 1500°

Recall that our qualitative argument said we have minima for all integer m **except zero**. What does our formula show? If we evaluate

$$\frac{\sin(\beta/2)}{\beta/2}$$

at $\beta = 0$, we get 0 over 0, undetermined. We need to approach $\beta = 0$ as a limit. Using L'Hospital's rule

$$\lim_{x\to 0}\frac{\sin x}{x} = \lim_{x\to 0}\left(\frac{d\sin x}{dx}\Big/\frac{dx}{dx}\right) = \lim_{x\to 0}\frac{\cos x}{1} = 1,$$

so in fact the intensity at $\theta = 0$ is not zero but the maximum.

3 Two slit diffraction

When we considered multiple slit diffraction last time, we assumed the intensity from each slit would not be a function of angle. The interference patterns usually have a fairly small angular range, and if the slit is thin enough the light passing through it emerges with a broad angular distribution. But we now see that if each slit has a width a, each slit will contribute with an amplitude proportional to

$$\frac{\sin(\pi a\sin\theta/\lambda)}{\pi a\sin\theta/\lambda},$$

so we should correct the two slit formula for narrow slits

$$I = I_{\max}^{(a\to 0)} \cos^2\left(\frac{\pi d \sin\theta}{\lambda}\right) \qquad \text{(two infinitesimal slits)}$$

by replacing $I_{\max}^{(a\to 0)}$ with

$$I_{\max}\left(\frac{\sin(\pi a \sin\theta/\lambda)}{\pi a \sin\theta/\lambda}\right)^2,$$

which gives all together

$$I_{\max}\cos^2\left(\frac{\pi d \sin\theta}{\lambda}\right)\left(\frac{\sin(\pi a \sin\theta/\lambda)}{\pi a \sin\theta/\lambda}\right)^2.$$

If the spacing of the slits, d, is many times the width of each slit, a, the cosine term will oscillate much more rapidly than the sine term. The intensity pattern will then be the rapid oscillation of the double slit pattern *modulated* by the single slit pattern. Here it is for $d = 6a$:

S&BV5 Fig 38.11 multiple finite width slits
7" × 6"

4 Resolution

As we saw last time, when a plane wave passes through a narrow slit, some of the light is diffracted so that it is no longer travelling in exactly the same direction that it had been. If we have a lens behind the slit with a screen at its focal length, the deviated light will be focussed into a position other than where it would have gone had there been no slit. If there were two sources of light, each of them has a central maximum in the position where the source would have been imaged had there been no slit, but each of them also has a diffraction pattern from the slit. [Note that the two independent sources do not interfere with each other, only each ones light interferes with itself.]

Now if the two sources are not very far apart in angle, the images formed, which are no longer points but diffraction patterns, may start to overlap. If they overlap enough, we will not be able to tell that this is the light from two sources rather than one. This means that imaging through a slit imposes a limit on the **resolution**.

Although it is somewhat arbitrary when to declare images no longer resolved, the commonly accepted standard is called **Rayleigh's criterion**.

Rayleigh announced that we should consider the images just barely resolvable if the central maximum of one image falls on the first minimum of the other. As we have seen, for a single slit, this means that the angle between the sources must be no less than

$$\sin\theta = \frac{\lambda}{a}.$$

You are probably more familiar with optical systems which involve circular aperatures rather than slits. Working out the diffraction of a circular aperature is considerably more mathematically sophisticated than for a slit, so we will only quote the result, which is that the first minimum occurs at an angle

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

where D is the diameter of the circular aperature. So every optical instrument automatically has a limit on its resolving power. If you go outside on a dark night, your pupils will dilate to, say, 8 mm in diameter. If you look up at the stars, the light appears white, meaning that many different wavelengths are involved, but we may use 500 nm (green) as a good central value. Then according to Rayleigh's criterion, no matter how well your eyes focus or how closely together your retinal receptors are, you will not be able to resolve two stars which are at an angle less than[2]

$$\begin{aligned} \theta_{\text{min}} = 1.22\frac{\lambda}{D} = 1.22 \times \frac{5.00 \times 10^{-7}}{8 \times 10^{-3}} &= 7.6 \times 10^{-5}\,\text{radians} \\ &= 0.004^\circ \approx \frac{1}{4}\ \text{minutes of arc.} \end{aligned}$$

Rutgers is currently involved in building a telescope in South Africa with an aperature of 11 m. If its resolutions were limited by diffraction, what would be the minimum angle it could resolve? How big an object on the moon could be resolved as not pointlike?

$$\begin{aligned} \text{Now } \theta_{\text{min}} = 1.22\frac{\lambda}{D} = 1.22 \times \frac{5.00 \times 10^{-7}}{11} &= 5.5 \times 10^{-8}\,\text{radians} \\ &= 11\ \text{milliseconds of arc.} \end{aligned}$$

[2]Actually you can do better, because we should have used the wavelength in your eyeball, not in vacuum. The index of refraction in your vitreous humor is roughly like water's, 1.33, so the wavelength there is about 3/4 of the wavelength in air, and you so you can perhaps do about 30% better than we calculated.

The moon is $R = 3.84 \times 10^8$ m from Earth, so the minimum resolvable distance would be $R\theta_{\text{min}} = 3.84 \times 10^8 \times 5.5 \times 10^{-8}\,\text{m} = 21$ m.

Unfortunately atmospheric disturbances cause more resolution problems than diffraction for such big telescopes, and the resolution on a really good night will never be better than about 0.8 seconds of arc. The size of large ground based telescopes is primarily to be able to gather more light from weak sources, rather than to improve the resolution from diffraction.

5 Summary

- Waves do not actually move in straight lines, but have some spreading out when they pass boundaries, called **diffraction**.
- Fraunhofer diffraction occurs when the interfering light can be considered parallel as it passes through the aperature.
- Each portion of the aperature can be considered a source interfering with all the other portions.
- For a single slit of width a, the intensity at angle θ is

$$I = I_{\text{max}} \left(\frac{\sin(\pi a \sin\theta/\lambda)}{\pi a \sin\theta/\lambda} \right)^2 .$$

- For multiple slits each with finite width a, the intensity is the product of the formula for a single slit of width a with that for multiple infinitesimal slits.
- Rayleigh's criterion for being able to resolve two images is that the central maximum of the diffraction of one lie at the minimum of the central maximum of the other. If they are closer than that, it is hard to distinguish the images.
- For a slit of width a, this gives $\theta_{\text{min}} = \frac{\lambda}{a}$. For a circular aperature $\theta_{\text{min}} = 1.22\frac{\lambda}{D}$, where D is the diameter.