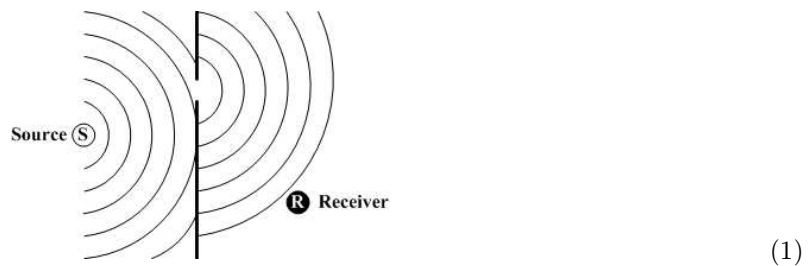


Lecture 19: Diffraction and resolution

1 Huygens' principle

Diffraction refers to what happens to a wave when it hits an obstacle. The key to understanding diffraction is a very simple observation first due to Huygens in 1678. Say a wave arrives at an opaque screen with a little hole in it. On the other side of the screen, the wave equation must still be satisfied with boundary conditions given by the motion of the wave in the hole. That is, the solution is identical to a situation where there was a source in the hole. The wavefront diagram looks like this:



The key point is that the receiver does not care about whatever happened on the other side of the screen. All that matters is what the wave was like as it passed through the hole.

Here is a picture of real wavefronts near a beach in Italy, taken from a satellite, showing this phenomenon:

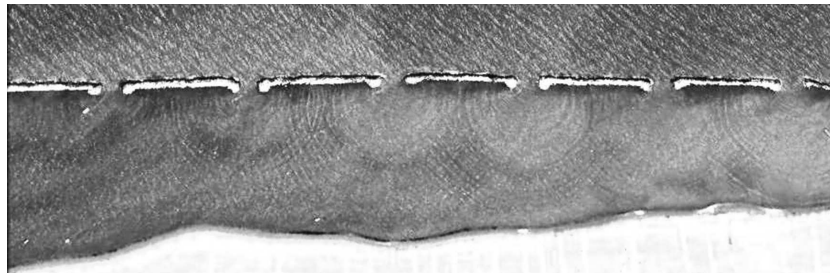
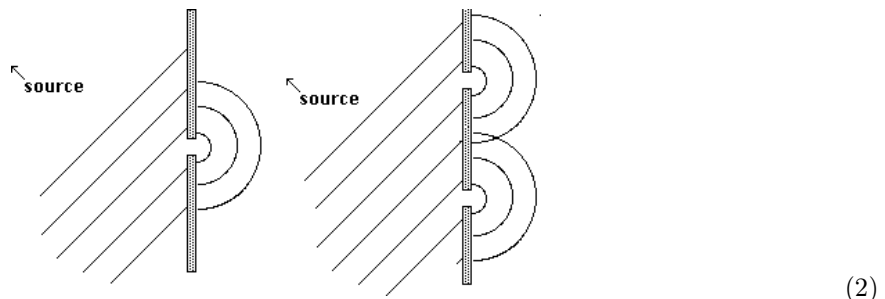


Figure 1. Waves near a shore somewhere in Italy. This image was taken from Google Earth.

The great thing about this way of thinking about diffraction is that, since the wave equation is linear, you can use this trick for any number of holes. You simply add the amplitude for the waves produced from a “source” at each hole:



Computing the amplitude by adding point sources in this way is known as **Huygens' principle**.

Huygen's principle works even if the holes are very close together. In fact, it works if they are connected, so instead of a hole, it's a slit. One just has to integrate over the holes rather than sum over them. We'll begin with a discrete set of holes as in a diffraction grating. Then

we'll take the continuum limit and talk about slits.

2 Multiple hole diffraction

Using Huygens' principle, we can easily compute the diffraction pattern from a plane wave passing through any number of holes. Say there are N holes in a row separated by a distance d . The solution will be as if there are N sources separated by a distance d . Thus the pattern will be the same as the antenna pattern for this source configuration we computed in Lecture 18!

Say one source has field $E = E_0 e^{-i\omega t} e^{ikr}$. Recall the diagram for two sources separated by a distance d :

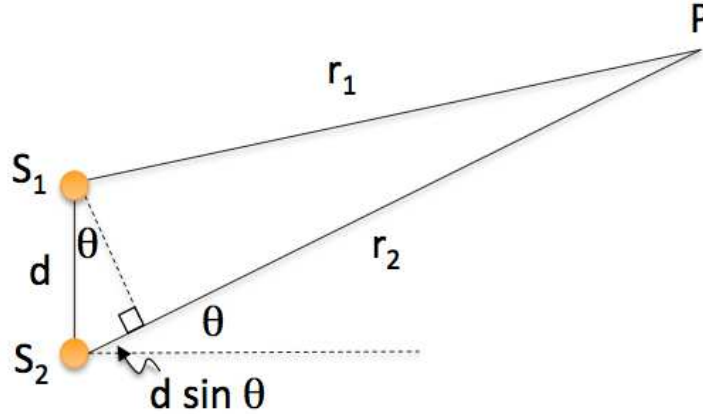


Figure 2. Phase difference from two holes is the same as from two sources.

We found that at a distant point P at an angle θ to the sources the field is

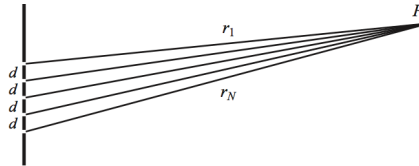
$$E_P = 2E_0 e^{-i\omega t} e^{ikr} \cos\left(\frac{\Delta}{2}\right) \quad (3)$$

where

$$\Delta = 2\pi \frac{d}{\lambda} \sin\theta + \delta \quad (4)$$

with δ the difference in phase between the two sources. This phase difference will be $\delta = 0$ if the sources are illuminated by a plane coming perpendicular to the separation between the sources. If the wave comes on an angle, $\delta \neq 0$.

For N slits, we get N sources, and the picture looks like this



This is called a **diffraction grating**. In this case, the amplitude at P is (again, see Lecture 18)

$$E_P = E_0 \left[\frac{e^{iN\frac{\Delta}{2}}}{e^{i\frac{\Delta}{2}}} \right] \left[\frac{\sin\left(N\frac{\Delta}{2}\right)}{\sin\left(\frac{\Delta}{2}\right)} \right] \quad (5)$$

So the intensity is

$$I = I_0 \frac{\sin^2\left(N\frac{\Delta}{2}\right)}{\sin^2\left(\frac{\Delta}{2}\right)} \quad (6)$$

where I_0 is the intensity from a single source.

Since the slits are spaced d apart, the total size of the diffraction grating is $a = Nd$. Now say we place a screen at a distance L from the slits with $L \gg a$. Then the height y up this screen is $y = L \tan\theta \approx L\theta \approx L \sin\theta$, and so

$$\Delta = 2\pi \frac{d y}{\lambda L} + \delta \tag{7}$$

If the plane wave hits the grating straight on, so $\delta = 0$ then

$$I(y) = I_0 \frac{\sin^2\left(\pi N \frac{dy}{\lambda L}\right)}{\sin^2\left(\pi \frac{dy}{\lambda L}\right)} \tag{8}$$

This is the intensity pattern from a diffraction grating. Note that for $N = 2$ this simplifies to $I(y) = 4I_0 \cos^2\left(\pi \frac{dy}{\lambda L}\right)$, which we computed in the previous lecture.

Here are some plots of $I(y)$ for $N = 2, N = 5$ and $N = 10$ slits:

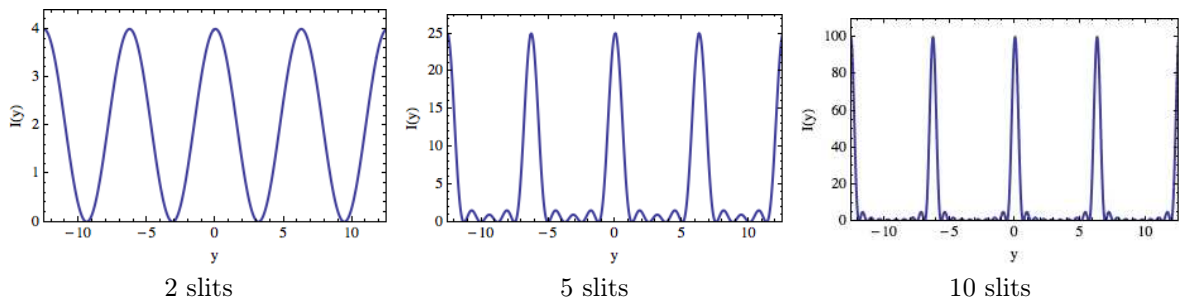


Figure 3. Patterns from diffraction gratings with $N = 2, 5$ and 10 slits. These plots show the intensity at a distance y up a screen placed at a distance L from the wall.

Here’s what this actual looks like with a laser passing through a diffraction grating

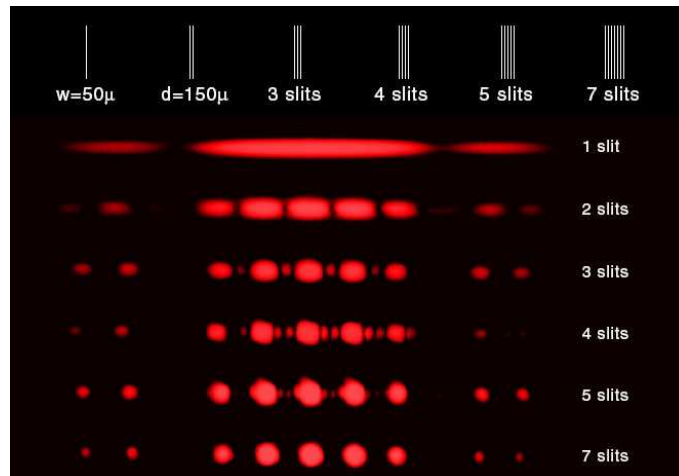


Figure 4. Pattern from a diffraction grating. As the number of slits is increased, the peaks become clearly separated. The extra dots on the right and left side after the gaps are due to the finite width of the slits, as discussed in Section 4.

You can see from the plots that each intensity pattern has a maximum at $y = 0$. Indeed, we find

$$I(0) = \lim_{y \rightarrow 0} I(y) = I_0 N^2 \tag{9}$$

which agrees with the vertical axes on the plots. So the more slits in the grating, the larger the central peak will be. The small bumps are due to the oscillation of the $\sin^2\left(\pi N \frac{yd}{\lambda L}\right)$ factor in the numerator. Indeed, the factor of N makes these oscillations have very high frequency. The overall periodicity is due to the denominator factor $\sin^2\left(\pi \frac{dy}{\lambda L}\right)$. This vanishes when

$$y_{\max} = j \frac{L\lambda}{d} = j \frac{L}{d} \frac{c}{\nu}, \quad j = 0, 1, 2, \dots \quad (10)$$

with ν the frequency.

Since the peaks are spaced by $\frac{L}{d}\lambda$, different wavelengths will have peaks in different places. Thus, diffraction gratings can be used to characterize the spectra of various things. For example, gases have interesting spectra which can be resolved with diffraction gratings.

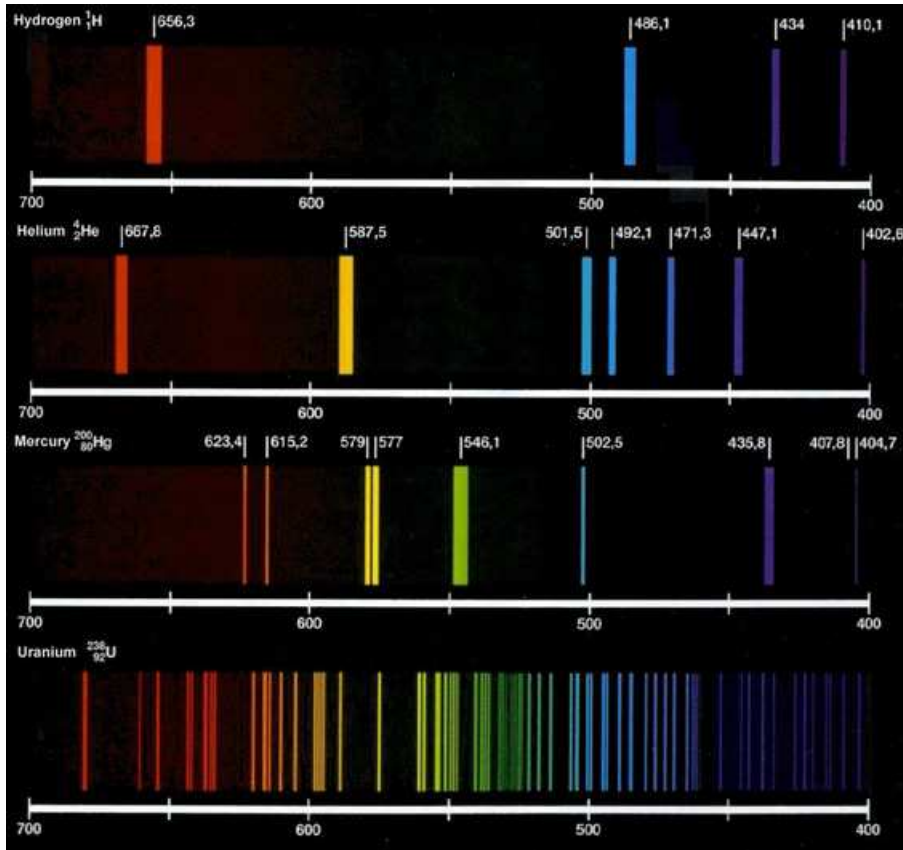


Figure 5. Spectra of hydrogen, helium, mercury and uranium as viewed through a diffraction grating.

The wavelengths in the hydrogen spectrum have the interesting pattern $\frac{1}{\lambda} = \frac{1}{91\text{nm}}\left(\frac{1}{4} - \frac{1}{j^2}\right)$ for $j = 3, 4, 5$ and 6 . This phenomenological observation was known in the 19th century. We will explain where it comes from in Lecture 21, on quantum mechanics.

3 Resolving power

An important use of diffraction gratings is to separate light into different wavelengths, as in Figure 5. Say a gas emits two colors, with wavelengths λ and λ' . How do we know if a given diffraction grating can separate the colors? Obviously, if the two wavelengths are infinitesimally close together, we will not be able to separate them. A reasonable criteria for whether frequencies can be separated is if the maximum of the intensity pattern from one wavelength lies on top of the minimum of the intensity pattern of the other:

$$\max \lambda' = \min \lambda \quad (11)$$

This is known as the **Rayleigh Criterion**.

All frequencies have maxima at $y = 0$, so we must look to the second peak to apply the criterion. The second maximum for λ' is at

$$y_{\max} = \frac{L}{d} \lambda' \quad (12)$$

The minima for λ happen only when the $\sin^2\left(\pi N \frac{yd}{\lambda L}\right)$ numerator in Eq. (8) vanishes, which is when

$$y_{\min} = m \frac{L \lambda}{d N}, \quad m = 1, 2, \dots, N - 1, N + 1, \dots \quad (13)$$

Values $m = 0$ and $m = N$ are excluded from this list since for these values the dominator in Eq. (8) vanishes as well and they are associated with maxima not minima.

For the second maximum of λ' to overlap with a minimum of λ we then want

$$\frac{L}{d} \lambda' = m \frac{L \lambda}{d N} \quad (14)$$

Or

$$\frac{\lambda'}{\lambda} = \frac{m}{N} \quad (15)$$

Note that this is independent of L and d . We want the value of m where this ratio is as close to 1 as possible (so that λ' is as close to λ as possible). So we take $m = N - 1$, referring to the minimum right next to the first maximum. Then $\frac{\lambda'}{\lambda} = \frac{N-1}{N} = 1 - \frac{1}{N}$. Thus $\frac{1}{N} = 1 - \frac{\lambda'}{\lambda} = \frac{\Delta\lambda}{\lambda}$ or

$$\boxed{\frac{\lambda}{\Delta\lambda} = N} \quad (16)$$

where $\Delta\lambda = \lambda' - \lambda$. In general $\frac{\lambda}{\Delta\lambda}$ with $\Delta\lambda$ is the smallest shift in λ which can be resolved is known as the **resolving power** of an optical system. Eq. (16) gives the resolving power of a diffraction grating.

For example, Hydrogen has wavelengths $\lambda_{\text{red}} = 656.2$ nm, $\lambda_{\text{blue}} = 486.1$ nm, $\lambda_{\text{violet-1}} = 434.0$ nm and $\lambda_{\text{violet-2}} = 410.1$ nm. As you can see from Figure 5 the closest are 434.0 nm and 410 nm with

$$\frac{410}{434 - 410} = 18 \quad (17)$$

So we would need 18 slits to see this separation. With 18 slits the lines would be barely separable. Typical diffraction gratings have 100s slits, so separations of $\frac{\Delta\lambda}{\lambda} \sim 1\%$ can be resolved. High quality diffraction gratings can have $\sim 10,000$ slits.

4 Wide slits

What happens if instead of having N holes separated d apart, we have one big hole? An easy way to derive the result is to take the limit of our previous result, Eq. (8) $N \rightarrow \infty$ holding the overall size $a = Nd$ fixed. Writing Eq. (8) in terms of N and I we have

$$I(y) = I_0 \frac{\sin^2\left(\pi \frac{Ndy}{\lambda L}\right)}{\sin^2\left(\pi \frac{dy}{\lambda L}\right)} \quad (18)$$

Note that $I(0) = \lim_{y \rightarrow 0} (y) = I_0 N^2$ which seems to blow up as $N \rightarrow \infty$. However, the relative intensity at a distance y compared to the intensity at $y = 0$ is finite

$$\frac{I(y)}{I(0)} = \lim_{N \rightarrow \infty} \frac{1}{N^2} \frac{\sin^2\left(\pi \frac{Ndy}{\lambda L}\right)}{\sin^2\left(\pi \frac{dy}{\lambda L}\right)} = \lim_{N \rightarrow \infty} \frac{1}{N^2} \frac{\sin^2\left(\pi \frac{ay}{\lambda L}\right)}{\sin^2\left(\pi \frac{ay}{N\lambda L}\right)} = \left(\frac{\sin\left(\pi \frac{ay}{\lambda L}\right)}{\pi \frac{ay}{\lambda L}}\right)^2 \quad (19)$$

This function is known as the **sinc** function

$$\text{sinc}(\beta) \equiv \frac{\sin\beta}{\beta} \quad (20)$$

So

$$\frac{I(y)}{I(0)} = \text{sinc}^2\left(\pi \frac{ay}{\lambda L}\right) \quad (21)$$

where a is the size of the slit, λ the wavelength, L the distance to the screen and y the distance up the screen.

This pattern looks like

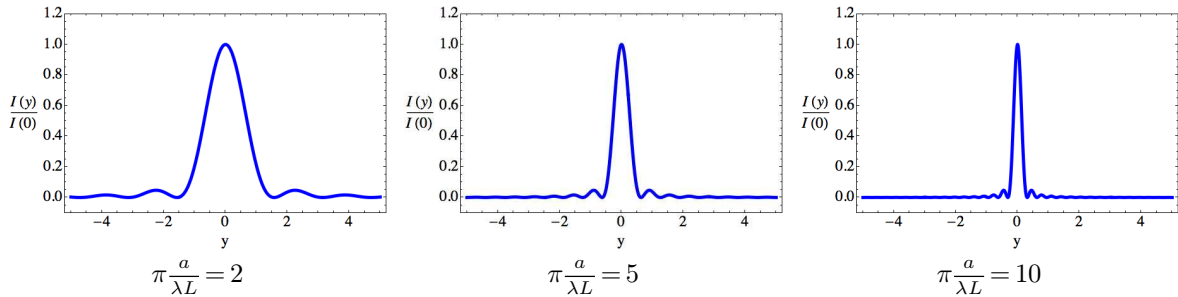


Figure 6. The $\text{sinc}^2\left(\frac{\pi a}{\lambda L}y\right)$ function gives the intensity from a wide slit. The wider the slit, the narrower the peak at fixed λ .

Or we can draw the picture like this

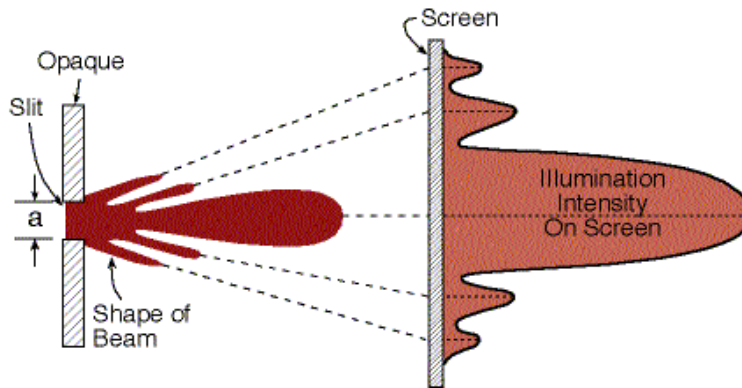


Figure 7. Diagram of wide-slit diffraction pattern.

The location of the first minimum in $\text{sinc}^2(\beta)$ is at $\beta = \pi$ or $\pi \frac{ay}{\lambda L} = \pi$, thus the width of the first peak is

$$\Delta y = \frac{2\lambda L}{a} \quad (22)$$

Note that as a gets smaller, y gets bigger. In the limit that a goes to a point, the outgoing wave is in phase everywhere and $\Delta y \rightarrow \infty$. Here's a photo of what this actually looks like

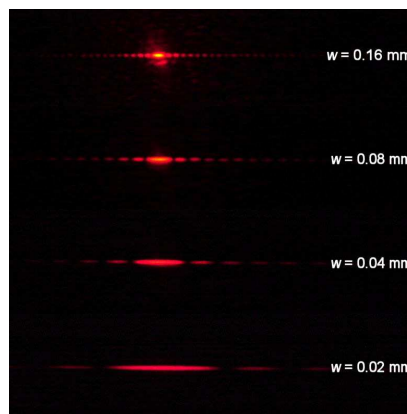


Figure 8. Wide slit interference patterns with variable width slits.

Now say we have two far away sources (like stars) and the light passes through a slit, making an image on a screen. What angular separation of the stars will we be able to distinguish on the screen? To find out, we apply the Rayleigh criterion: the maximum of one should hit the minimum of the other. For light coming in at an angle $\Delta\theta$, its plane-wave light will have slightly different phases at different parts of the slit. The net effect is to shift the pattern so that the maximum is not at $y=0$ but at $y=L\sin\theta$. So

$$\frac{I(y)}{I(0)} = \text{sinc}^2\left[\pi\frac{a}{\lambda L}(y - L\sin\theta)\right] \quad (23)$$

Say one star's light comes in at $\theta=0$ and the other at $\Delta\theta$. The one at $\theta=0$ is directly normal to the screen, so the *minimum* of its intensity pattern is at $y = \frac{\lambda L}{a}$. If the other one comes in at an angle $\Delta\theta$, so its *maximum* is at $y = L\sin\Delta\theta \approx L\Delta\theta$ for small angles. Setting these equal gives

$$\boxed{\Delta\theta_R = \frac{\lambda}{a}} \quad (24)$$

This is the resolving power of a slit. For a circular aperture, the formula is

$$\Delta\theta_R = 1.22\frac{\lambda}{a} \quad (25)$$

where the 1.22 comes from the circular geometry. If you have 20/20 vision, you can only resolve objects at around 12 times farther apart than the Rayleigh criterion. That is, ordinary vision is not diffraction limited.

If we have two wide slits, the interference pattern is a combination of the $\text{sinc}^2\left(\pi\frac{ay}{\lambda L}\right)$ behavior of a wide slit and the $\cos^2\left(\pi\frac{dy}{\lambda L}\right)$ behavior of the double slit. This is shown in Figure 9.

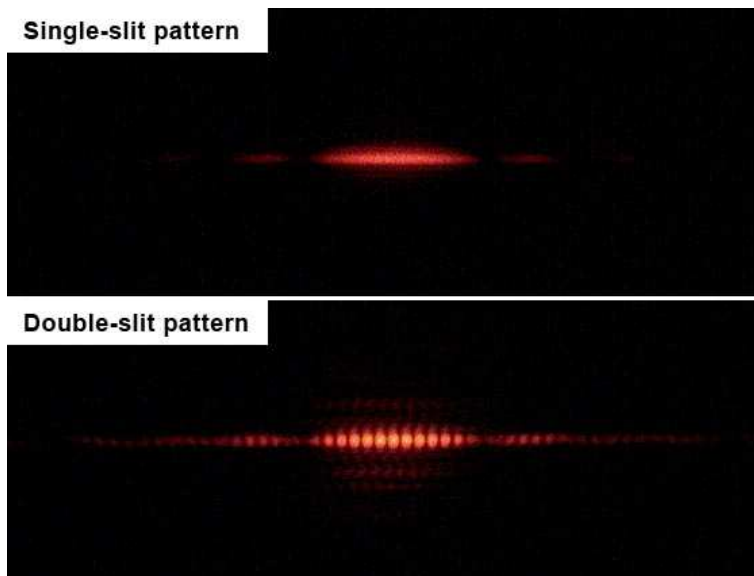


Figure 9. Comparison of the interference pattern for a single wide slit and two wide slits.

5 Fourier optics

What would happen if instead of a slit, we put a film in the slit which had a transparency $T(x)$ which varies along the length of the slit. Let's define $T=0$ as opaque and $T=1$ as totally transparent. Thus, the wide slit is $T(x)=1$. A half transparent slit would have $T(x)=\frac{1}{2}$. The diffraction grating is a bunch of small regions with $T=1$ surrounded by regions with $T=0$. We can model this with a set of δ -functions:

$$T(x) = \sum_{j=1}^N \delta(x - jd) \quad (26)$$

What is the intensity pattern on the screen from a general $T(x)$? We can compute the answer using Huygens' principle!

If a plane wave comes in with a field E_0 at the slit, then each point will come with a field given by $E_0T(x)$. At a distance y on a screen, we can then use our formula that the phase shift is

$$\Delta = 2\pi \frac{x y}{\lambda L} \quad (27)$$

So the sum of the phases gives

$$E(y) \propto \int dx T(x) e^{2\pi i x \frac{y}{L\lambda}} = \tilde{T}\left(2\pi \frac{y}{L\lambda}\right) \quad (28)$$

where

$$\tilde{T}(k) = \int dx T(x) e^{i k x} \quad (29)$$

is the Fourier transform of k . That is, the image on the screen is the **Fourier transform of the transparency**. So finally, we understand Fourier optics. You have been using this result in lab all semester, and we have finally derived it.

To make sure we're not nuts, let's check on our wide slit. There we expect the field E to be the square root of the intensity in Eq. (21):

$$E(y) \propto \text{sinc}\left(\pi \frac{a y}{\lambda L}\right) \quad (30)$$

The relevant $T(x)$ is given in Eq. (23). Its Fourier transform is

$$\tilde{T}(k) = \int_{-\infty}^{\infty} dx T(x) e^{i k x} = \int_{-\frac{a}{2}}^{\frac{a}{2}} dx e^{i k x} = \frac{1}{i k} \left[e^{i \frac{k a}{2}} - e^{-i \frac{k a}{2}} \right] = \frac{2}{k} \sin\left(\frac{k a}{2}\right) = a \text{sinc}\left(\frac{k a}{2}\right) \quad (31)$$

So

$$E(y) \propto \tilde{T}\left(\pi \frac{a y}{L\lambda}\right) \quad (32)$$

exactly as expected. We can also check the diffraction grating using Eq. (26),

$$\tilde{T}(k) = \int_{-\infty}^{\infty} dx T(x) e^{i k x} = \sum_{j=1}^N \int_{-\infty}^{\infty} dx \delta(x - j d) e^{i k x} = \sum_{j=1}^N e^{i k j d} = e^{i(N+1)\Delta} \frac{\sin\left(N \frac{\Delta}{2}\right)}{\sin\left(\frac{\Delta}{2}\right)} \quad (33)$$

where $\Delta = k d = 2\pi \frac{d}{\lambda}$. This agrees with our result for the diffraction grating in Eq. (5).

In summary, if we put a screen over the slit with transparency $T(x)$ then the image on the screen will be proportional to $\tilde{T}\left(\pi \frac{a y}{L\lambda}\right)$. Note that this holds even for monochromatic (fixed frequency) incoming light, as in a laser. Say the wavelength λ of light and the distance L to the screen are fixed. Then the intensity at a distance y up the screen tells us the strength of the Fourier component of the slit corresponding to a wavenumber of

$$k = 2\pi \frac{y}{L\lambda} \quad (34)$$

To be absolutely clear, this k is **not** $\frac{2\pi}{\lambda}$. The Fourier transform on the screen has nothing to do with the wavelength of the light entering the slit. The farther up the slit you look, the larger k , and hence smaller wavelength modes you are seeing. This is why the spot gets smaller as the slit gets bigger as in Figure 6.