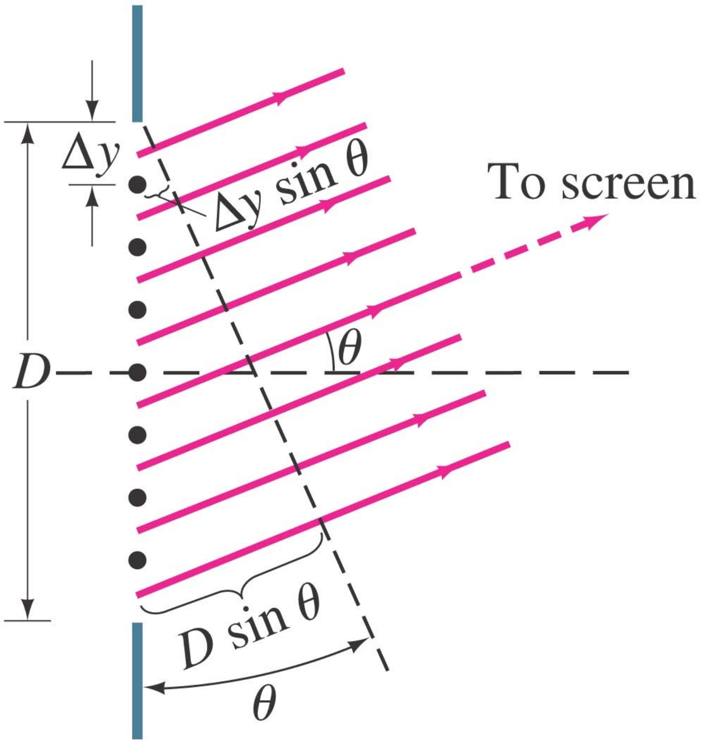
Intensity of single slit diffraction

1. General considerations

Following Giancoli, section 35-2 (and quoting some of the text), we consider the slit divided up into N very thin strips of width Δy as indicated in the figure below. Note that the width of the slit is $D = N\Delta y$.



Each strip sends light in all directions to a screen on the right. We take the rays heading for any particular point on the distant screen to be parallel, all making an angle θ with the horizontal as shown above. We choose the strip width $\Delta y \ll \lambda$ so that all the light from a given strip is in phase. The strips are of equal size, and if the whole slit is uniformly illuminated, we can take the electric field wave amplitudes from each thin strip to be equal. However, the separate amplitudes from the different strips will differ in phase. The phase difference in the light coming from adjacent strips is given by:

$$\Delta\beta = \frac{2\pi}{\lambda}\Delta y \sin\theta\,,$$

since the difference in path length is $\Delta y \sin\theta$. The total amplitude on the screen at any angle θ (denoted by E_θ) will be the sum of the separate wave amplitudes due to each strip. Finally, we take the limit of $N \to \infty$ and $\Delta y \to 0$, where the limits are taken such that the product, $D = N\Delta y$, is held fixed. Explicitly,[1]

$$E_\theta = \lim_{\substack{N\to\infty \\ \Delta y\to 0}} \frac{E_0}{N}\left[\sin\omega t + \sin(\omega t+\Delta\beta) + \sin(\omega t + 2\Delta\beta) + \ldots + \sin\bigl(\omega t + (N-1)\Delta\beta\bigr)\right],$$

where $\Delta\beta = 2\pi\Delta y\sin\theta/\lambda$. As in the case of the two-slit interference experiment, if the distance from the slit to the screen, L, is much larger than D, then the electric field vectors from the light rays originating from each of the strips are essentially parallel.

2. Calculation of the intensity

The (time-averaged) intensity of the resulting wave at the screen, located at an angle θ with respect to the symmetry axis of the slit (as shown in the figure), is proportional to E_θ^2 averaged over one full cycle of the wave. Thus, we can write:

$$I(\theta) = K\langle E_\theta^2\rangle\,,$$

where K is a constant (to be determined below) and the brackets $\langle\cdots\rangle$ indicate a time-average over one cycle of the wave. Our first task is to obtain a closed-form expression for E_θ, which we rewrite below using the summation notation:

$$E_\theta = \lim_{\substack{N\to\infty \\ \Delta y\to 0}} \frac{E_0}{N}\sum_{n=0}^{N-1}\sin\left(\omega t + n\Delta\beta\right)\,.$$

Since $D = N\Delta y$ is the width of the slit (which is fixed), it is convenient to substitute $1/N = \Delta y/D$ in the expression above, which yields

$$E_\theta = \lim_{\Delta y\to 0}\frac{E_0}{D}\sum_{n=0}^{N-1}\sin\left(\omega t + \frac{2\pi n\Delta y\sin\theta}{\lambda}\right)\Delta y = \frac{E_0}{D}\int_0^D \sin\left(\omega t + \frac{2\pi y\sin\theta}{\lambda}\right)dy\,.$$

The last step above is a consequence of the definition of the definite integral, which can be computed by approximating the area under the curve $\sin(\omega t + 2\pi y\sin\theta/\lambda)$ between $y=0$ and $y=D$ with N rectangular slices, each of width Δy. The exact result for the area then follows by taking $N\to\infty$ and $\Delta y\to 0$, keeping $D = N\Delta y$ fixed, as indicated by the equation above.

To compute the above integral, we introduce a change of variables:

$$z = \omega t + \frac{2\pi y\sin\theta}{\lambda}\,.$$

Then, $dz = (2\pi\sin\theta/\lambda)dy$. Writing dy in terms of dz, and expressing the integrand in terms of z then yields:

$$E_\theta = \frac{\lambda E_0}{2\pi D\sin\theta}\int_{\omega t}^{\omega t + 2\pi D\sin\theta/\lambda}\sin z\,dz\,,$$

after noticing that if $y=0$ then $z=\omega t$, and if $y=D$ then $z = \omega t + 2\pi D\sin\theta/\lambda$ (which determine the limits of integration over z). This last integral is elementary, and we obtain:

$$E_\theta = \frac{-\lambda E_0}{2\pi D\sin\theta}\left[\cos\left(\omega t + \frac{2\pi D\sin\theta}{\lambda}\right) - \cos\omega t\right]\,.$$

It is convenient to employ the trigonometric identity:

$$\cos A - \cos B = -2\sin\left(\frac{A+B}{2}\right)\sin\left(\frac{A-B}{2}\right)\,,$$

[1]Note that for $\theta = 0$, we obtain $E_{\theta=0} = E_0\sin\omega t$, which defines the amplitude E_0.

with $A \equiv \omega t + 2\pi D \sin\theta/\lambda$ and $B \equiv \omega t$ to obtain:

$$E_\theta = \frac{\lambda E_0}{\pi D \sin\theta} \sin\left(\omega t + \frac{\pi D \sin\theta}{\lambda}\right) \sin\left(\frac{\pi D \sin\theta}{\lambda}\right) .$$

Following Eq. 35-6 of Giancoli, we introduce the notation:

$$\beta \equiv \frac{2\pi D \sin\theta}{\lambda} .$$

Then, we can write:

$$E_\theta = E_0 \sin\left(\omega t + \tfrac{1}{2}\beta\right) \frac{\sin(\beta/2)}{\beta/2} .$$

As noted at the beginning of this section, the (time-averaged) intensity $I(\theta)$ is given by

$$I(\theta) = K\langle E_\theta^2 \rangle ,$$

where K is a constant (to be determined below) and the brackets $\langle \cdots \rangle$ indicate a time-average over one cycle of the wave. Note that the only time-dependence is in the factor $\sin\left(\omega t + \frac{1}{2}\beta\right)$, which is squared when one computes E_θ^2. Moreover,

$$\left\langle \sin^2\left(\omega t + \tfrac{1}{2}\beta\right)\right\rangle = \left\langle \cos^2\left(\omega t + \tfrac{1}{2}\beta\right)\right\rangle ,$$

since the functions $\sin^2\left(\omega t + \frac{1}{2}\beta\right)$ and $\cos^2\left(\omega t + \frac{1}{2}\beta\right)$ differ only in phase by 90°, and thus must average to the same result when averaged over a full cycle. Using $\sin^2\left(\omega t + \frac{1}{2}\beta\right) + \cos^2\left(\omega t + \frac{1}{2}\beta\right) = 1$, we conclude that:

$$\left\langle \sin^2\left(\omega t + \tfrac{1}{2}\beta\right)\right\rangle = \tfrac{1}{2} .$$

Hence

$$I(\theta) = \tfrac{1}{2} K E_0 \left(\frac{\sin(\beta/2)}{\beta/2}\right)^2 , \qquad \text{where} \quad \beta \equiv \frac{2\pi D \sin\theta}{\lambda} .$$

Finally, we define $I_0 \equiv I(\theta = 0)$. When $\theta = 0$, we see that $\beta = 0$. Noting that

$$\lim_{\beta \to 0} \frac{\sin(\beta/2)}{\beta/2} = 1 ,$$

it then follows that

$$I_0 \equiv I(\theta = 0) = \tfrac{1}{2} K E_0^2 .$$

Hence, we arrive at our final result:

$$\boxed{I(\theta) = I_0 \left(\frac{\sin(\beta/2)}{\beta/2}\right)^2 , \qquad \text{where} \quad \beta \equiv \frac{2\pi D \sin\theta}{\lambda} ,}$$

which coincides with the results of Eqs. 35-6 and 35-7 of Giancoli.

3. An alternative derivation

Let us return to the expression:

$$E_\theta = \lim_{\substack{N\to\infty \\ \Delta y\to 0}} \frac{E_0}{N}\left[\sin\omega t + \sin(\omega t+\Delta\beta) + \sin(\omega t + 2\Delta\beta) + \ldots + \sin\bigl(\omega t + (N-1)\Delta\beta\bigr)\right],$$

Another strategy for evaluating this limit is to first compute the sum of the N terms shown above in closed form. Then take the limit of $N\to\infty$ and $\Delta y\to 0$, with $D = N\Delta y$ held fixed. In fact, the technique of phasor diagrams employed by Giancoli in Section 35-2 effectively performs the above sum and then takes the limit using a geometrical representation of the sum. This method is quite powerful, since it allows one to employ simple geometric and trigonometric reasoning to explicitly evaluate the expression above.

However, one can also perform the sum directly using algebraic techniques. Although these methods require slightly more sophisticated manipulations (as compared to the derivation of section 2 above), this approach may hold some interest. Here, I will simply quote the result for the sum of the N terms above (with a proof relegated to an appendix):

$$\frac{1}{N}\left[\sin\omega t + \sin(\omega t+\Delta\beta) + \ldots + \sin\bigl(\omega t + (N-1)\Delta\beta\bigr)\right] = \sin\bigl(\omega t + \tfrac{1}{2}(N-1)\Delta\beta\bigr)\,\frac{\sin(N\Delta\beta/2)}{N\sin(\Delta\beta/2)}\,.$$

We can check that this formula produces a known result for $N=2$. If we write $\delta\equiv\Delta\beta$, then

$$\sin\omega t + \sin(\omega t+\delta) = \sin(\omega t+\delta/2)\,\frac{\sin\delta}{\sin(\delta/2)} = 2\sin(\omega t+\delta/2)\cos(\delta/2)\,,$$

after using $\sin\delta = 2\sin(\delta/2)\cos(\delta/2)$. The $N=2$ result was used in class to derive the intensity of the two-slit experiment. It is easily established using the trigonometric identity

$$\sin A + \sin B = 2\,\sin\left(\frac{A+B}{2}\right)\cos\left(\frac{A-B}{2}\right)\,.$$

Returning to the formula for the sum of the N sine terms above, we can now compute the intensity that results from the superposition of these N terms. As before, the (time-averaged) intensity is given by $I(\theta) = K\langle E_\theta^2\rangle$, where K is a constant to be determined. Using

$$\left\langle \sin^2\bigl(\omega t + \tfrac{1}{2}(N-1)\Delta\beta\bigr)\right\rangle = \tfrac{1}{2}\,,$$

it follows that

$$I_N(\theta) = \tfrac{1}{2}KE_0^2\left(\frac{\sin(N\Delta\beta/2)}{N\sin(\Delta\beta/2)}\right)^2\,,\qquad\qquad \text{where}\quad \Delta\beta\equiv 2\pi\Delta y\sin\theta/\lambda\,,$$

and $I_N(\theta)$ is the total intensity due to the superposition of N sources. Since $\Delta\beta=0$ when $\theta=0$, it follows that $I_{N0}\equiv I_N(\theta=0) = \frac{1}{2}KE_0^2$, where we have used the fact that:

$$\lim_{\Delta\beta\to 0}\frac{\sin(N\Delta\beta/2)}{N\sin(\Delta\beta/2)} = 1\,.$$

Hence,

$$I_N(\theta) = I_{N0}\left(\frac{\sin(N\Delta\beta/2)}{N\sin(\Delta\beta/2)}\right)^2 = I_{N0}\left(\frac{\sin(\pi D\sin\theta/\lambda)}{N\sin(\pi\Delta y\sin\theta/\lambda)}\right)^2\,,$$

after using $D = N\Delta y$.

Finally, we are ready to take the limit of $\Delta y \to 0$ [in which case $N = D/\Delta y \to \infty$]. Note that we can make use of the small angle approximation to obtain:

$$\lim_{\Delta y \to 0} N \sin(\pi \Delta y \sin\theta/\lambda) = \pi N \Delta y \sin\theta/\lambda = \pi D \sin\theta/\lambda \,.$$

Thus, if we define $I(\theta) \equiv \lim_{N\to\infty} I_N(\theta)$ and $I_0 \equiv \lim_{N\to\infty} I_{N0}$, it follows that

$$\boxed{I(\theta) = I_0 \left(\frac{\sin(\pi D \sin\theta/\lambda)}{\pi D \sin\theta/\lambda}\right)^2 .}$$

The above result can be rewritten as:

$$I(\theta) = I_0 \left(\frac{\sin(\beta/2)}{\beta/2}\right)^2 , \qquad\qquad \text{where} \quad \beta \equiv \frac{2\pi D \sin\theta}{\lambda} \,,$$

which is precisely the formula for the intensity as a function of the angle θ obtained in section 2.

Appendix: derivation of the sum of N sine functions

In this appendix, I provide a derivation of the formula:

$$\frac{1}{N}\Big[\sin\omega t+\sin(\omega t+\Delta\beta)+\ldots+\sin\big(\omega t+(N-1)\Delta\beta\big)\Big]=\sin\big(\omega t+\tfrac{1}{2}(N-1)\Delta\beta\big)\,\frac{\sin(N\Delta\beta/2)}{N\sin(\Delta\beta/2)}\,.$$

The derivation of the formula makes use of complex numbers, so it is beyond the scope of this class. But, you can return to this proof later if you take Physics 116A (or an equivalent course). To follow all the steps, you will need to be familiar with Euler's formula:

$$e^{i\theta}=\cos\theta+i\sin\theta\,.$$

It follows that $\sin\theta=\text{Im }e^{i\theta}$, where Im instructs you to take *the imaginary part* of the corresponding expression. Euler's formula also implies that:

$$e^{i\theta}+e^{-i\theta}=2\cos\theta\,,\qquad\text{and}\qquad e^{i\theta}-e^{-i\theta}=2i\sin\theta\,.$$

All these results are used at some point in the analysis below.

Using the summation notation, and denoting $\delta\equiv\Delta\beta$,

$$\begin{aligned}
\sum_{n=0}^{N-1}\sin(\omega t+n\delta)&=\text{Im}\sum_{n=0}^{N-1}e^{i(\omega t+n\delta)}=\text{Im }e^{i\omega t}\sum_{n=0}^{N-1}e^{in\delta}\\
&=\text{Im }e^{i\omega t}\left(\frac{1-e^{iN\delta}}{1-e^{i\delta}}\right)\\
&=\text{Im }e^{i\omega t}\left(\frac{1-e^{iN\delta}}{1-e^{i\delta}}\right)\left(\frac{1-e^{-i\delta}}{1-e^{-i\delta}}\right)\\
&=\text{Im }e^{i\omega t}\,\frac{(1-e^{iN\delta})(1-e^{-i\delta})}{2-e^{i\delta}-e^{-i\delta}}\\
&=\frac{-1}{2(1-\cos\delta)}\,\text{Im}\,\Big[e^{i\omega t}e^{-i\delta/2}e^{iN\delta/2}(e^{iN\delta/2}-e^{-iN\delta/2})(e^{i\delta/2}-e^{-i\delta/2})\Big]\\
&=\frac{2}{1-\cos\delta}\,\sin(N\delta/2)\sin(\delta/2)\,\text{Im}\,\Big[e^{i\omega t}e^{-i\delta/2}e^{iN\delta/2}\Big]\\
&=\frac{2}{1-\cos\delta}\,\sin(N\delta/2)\sin(\delta/2)\,\sin\big(\omega t+\tfrac{1}{2}(N-1)\delta\big)\,.
\end{aligned}$$

At line two of the above equation, we performed the sum of a geometric series according to the well known formula[2]

$$\sum_{n=0}^{N-1}x^n=\frac{1-x^N}{1-x}\,,$$

where $x\equiv e^{i\delta}$.

Finally, we note the trigonometric identity $\frac{1}{2}(1-\cos\delta)=\sin^2(\delta/2)$. It then follows that:

$$\sum_{n=0}^{N-1}\sin(\omega t+n\delta)=\sin\big(\omega t+\tfrac{1}{2}(N-1)\delta\big)\,\frac{\sin(N\delta/2)}{\sin(\delta/2)}\,,$$

which is the desired result. Dividing both sides of this equation by N ant putting $\delta\equiv\Delta\beta$ yields the summation formula quoted at the beginning of this appendix.

[2] To derive the sum of a finite geometric series, define $S_N\equiv 1+x+x^2+\ldots+x^{N-1}$. Then $xS_N=x+x^2+\ldots+x^N$. It follows immediately that $S_N-1=xS_N-x^N$. Solving this simple equation for S_N yields $S_N=(1-x^N)/(1-x)$.