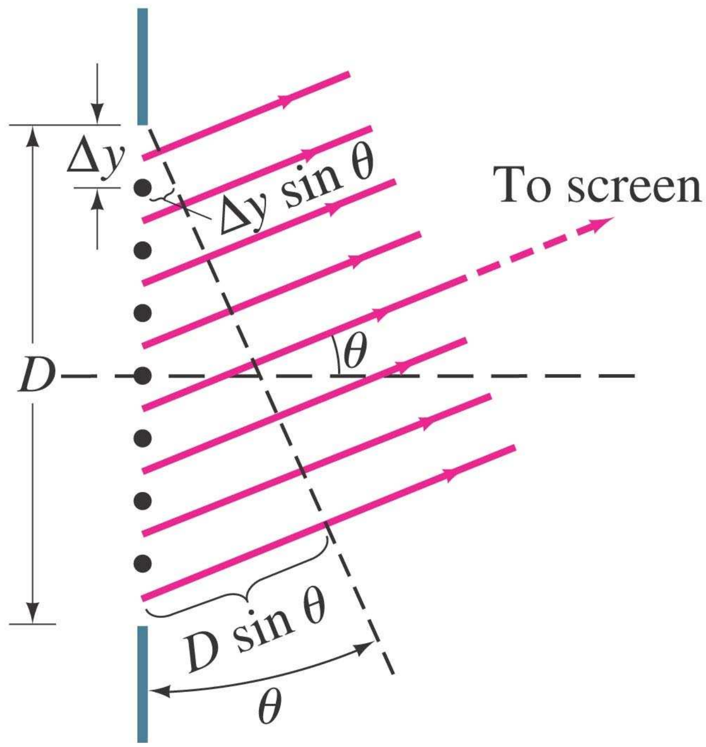
Intensity of single slit diffraction

1. General considerations

Following Giancoli, section 35-2 (and quoting some of the text), we consider the single 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 $D = N\Delta y$. This is an approximate description of an actual slit of width D. This approximation becomes exact in the limit of $N \to \infty$ and $\Delta y \to 0$, where these two limits are to be taken in such a way that the product $N\Delta y = D$ is fixed.



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, with all rays 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. The amplitude of the wave emanating from each thin strip is given by

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

where E_0 is the amplitude of the wave incident on the single slit. The outgoing waves from the different strips will differ in phase at the screen. The phase difference in the light coming from adjacent strips is given by:

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

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 coherent sum of the separate waves due to each strip:

$$E_\theta = E_{\theta 0} \sin(\omega t + \phi)\,.$$

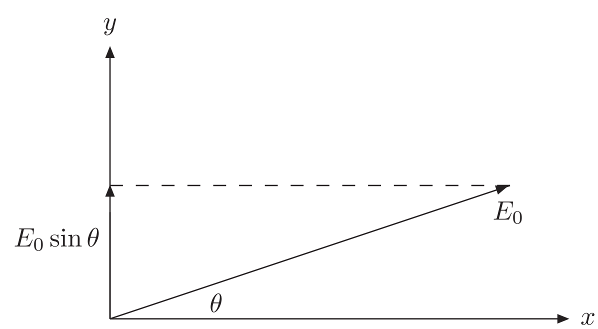
The object of the calculation presented in these notes is to compute the amplitude $E_{\theta 0}$ and the phase angle ϕ of the coherent sum of waves in the limit of $N \to \infty$ and $\Delta y \to 0$, where the limits are taken such that the product, $D = N\Delta y$, is held fixed. Since $\Delta E_0 = E_0/N$, the coherent sum of the N waves from the different strips at the screen is given by:

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

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, and we can ignore the vector nature of the electric field.

2. Calculation of the intensity by the method of phasors

To compute the sum of the sine functions in eq. (1), we use the method of phasors. To represent the quantity $E_0 \sin\theta$, we first consider a fictitious vector in a two-dimensional space, whose length is E_0, which makes an angle θ with respect to the x-axis as shown below:



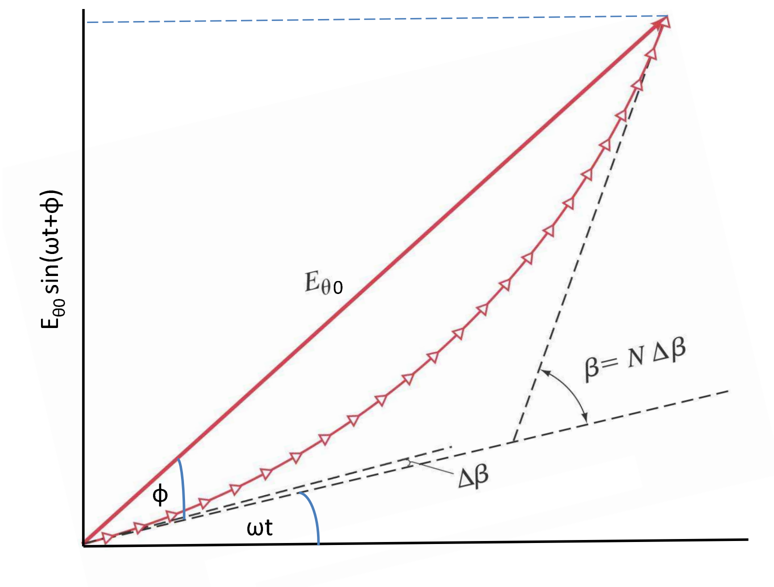
If we project this vector onto the y-axis, then its projected length is $E_0 \sin\theta$ as shown above. This vector is called a *phasor* and represents a quantity with an amplitude E_0 and an angle θ. The utility of such a representation is that we can perform the sum of eq. (1) by considering the phasors corresponding to each sine term in the sum, and then adding the phasors vectorially. The projection of the vector sum of the N phasors onto the y-axis is just the sum of the sine functions that we wish to compute. This vector sum can be carried out geometrically, and $E_{\theta 0}$ and ϕ can then be determined.

For example, in the case of $N = 2$, the phasor diagram for the sum of two sine functions is given by Figure 34-13 of Giancoli (also discussed in the class handout entitled "How to add sine functions of different amplitude and phase"). We now apply this technique to the sum of the N sine functions in eq. (1). Each phasor is oriented by an angle $\Delta\beta$ with respect to the adjacent phasor.

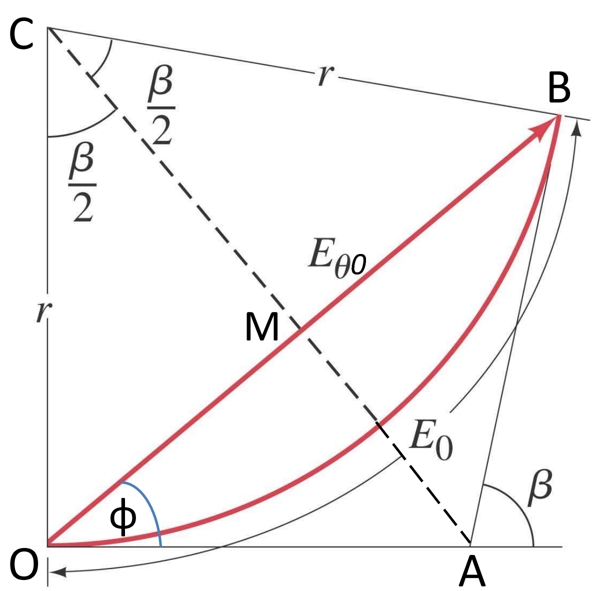
Consequently, the Nth phasor is oriented by an angle

$$\boxed{\beta \equiv N\Delta\beta = \frac{2\pi N\Delta y \sin\theta}{\lambda} = \frac{2\pi D \sin\theta}{\lambda}} \tag{2}$$

with respect to the first phasor as shown in the phasor diagram below.



Note that ΔE_0 is the length of each of the N phasors in the above figure. We can apply the limit of $N \to \infty$ and $\Delta y \to 0$ (with $N\Delta y = D$ held fixed) directly to the phasor diagram above. In this limit, the phasors follow the arc of a circle of radius r whose center is located at C. The arc length from the origin O to the point B shown in the figure below is then $N\Delta E_0 = E_0$.



We first prove that:

$$\boxed{\phi = \beta/2}$$

As a first step, draw the dashed line CM such that the lines CM and OB are perpendicular. Using the symmetry of the figure, we can extend the line CM to intersect the two tangent lines (OA and AB) at the point A. Since OCB is an isosceles triangle (with $OC = CB = r$), it follows that $OM = MB$, and we can conclude that OAB is also an isosceles triangle (with $OA = AB$). Using the fact that the sum of the angles of triangle OAB is 180^0, it then follows that $\phi = \beta/2$. (Likewise, we can also conclude that the angles OCM and MCB are both equal to $\beta/2$, as indicated on the figure above.)

Next, to determine $E_{\theta 0}$ (which is the length of the line OB), we note that from the triangle OCM it follows that

$$\frac{E_{\theta 0}}{2} = r \sin\left(\frac{\beta}{2}\right) ,$$

and the arc length $OB \equiv E_0 = r\beta$. Thus, $r = E_0/\beta$, which when substituted into the equation above yields:

$$E_{\theta 0} = E_0 \frac{\sin(\beta/2)}{\beta/2} .$$

We conclude that:

$$\boxed{E_\theta = E_{\theta 0}\sin(\omega t + \phi) = E_0 \sin(\omega t + \tfrac{1}{2}\beta)\,\frac{\sin(\beta/2)}{\beta/2}} \tag{3}$$

after using $\phi = \beta/2$, which was determined previously.

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. 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

$$I(\theta) = \tfrac{1}{2}KE_0\left(\frac{\sin(\beta/2)}{\beta/2}\right)^2 , \qquad \text{where} \quad \beta \equiv \frac{2\pi D\sin\theta}{\lambda} .$$

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

$$I_0 \equiv I(\theta = 0) = \tfrac{1}{2}KE_0^2 .$$

Hence, we arrive at our final result:

$$\boxed{I(\theta) = I_0\left(\frac{\sin(\beta/2)}{\beta/2}\right)^2 , \qquad \text{where} \quad \beta \equiv \frac{2\pi D\sin\theta}{\lambda}} \tag{4}$$

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

3. Algebraic method for computing E_θ

The phasor technique was a geometrical method for evaluating the sum given in eq. (1) in the limit of $N \to \infty$ and $\Delta y \to 0$ (with $D = N\Delta y$ held fixed). In this section, we will provide an algebraic technique for computing the same limiting sum. We first rewrite the expression for E_θ 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:

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

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

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

Introducing the angle β defined in eq. (2),

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

we end up with

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

which coincides with the result of eq. (3) obtained by the phasor technique.

4. An alternate algebraic 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 *before* taking the limit of $N \to \infty$ and $\Delta y \to 0$ (with $D = N\Delta y$ held fixed). The required limit is then taken at the end of the computation. 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 $\sin\delta = 2\sin(\delta/2)\cos(\delta/2)$. This $N = 2$ result was used in class to derive the intensity of the interference pattern arising in the two-slit experiment.

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

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

it follows that

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

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,

$$\boxed{I_N(\theta) = I_{N0}\left(\frac{\sin(N\Delta\beta/2)}{N\sin(\Delta\beta/2)}\right)^2 = I_{N0}\left(\frac{\sin(\pi D\sin\theta/\lambda)}{N\sin(\pi\Delta y\sin\theta/\lambda)}\right)^2} \tag{5}$$

after using $D = N\Delta y$.

Here, we shall apply the result of eq. (5) to the single slit diffraction problem by taking 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

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

which coincides with eq. (4) obtained by the phasor technique.

Appendix: derivation of the sum of N sine functions

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

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

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 $\sin\theta=\operatorname{Im}e^{i\theta}$, 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 $\delta\equiv\Delta\beta$,

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

At line two of the above computation, we performed the sum of a geometric series according to the well known formula[1]

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

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:

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

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

[1] 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 $S_N=(1-x^N)/(1-x)$.