Applied Optics Professor Akhilesh Kumar Mishra Department of Physics Indian Institute of Technology, Roorkee Lecture: 33 Single Slit Diffraction

Hello everyone, welcome to the class. Today we will proceed ahead of what we started in our last class. In the last class we talked about Fraunhofer diffraction and therein we calculated the total field at a point of observation due to a line charge distribution, therein we considered very large number of point oscillators, and due to those oscillators we calculated the resultant field at a point of observation.

(Refer Slide Time: 1:05)

Today we will start single slit diffraction. And here we will calculate the irradiance pattern on the screen obtained due to single slit.

Now to start with we will consider the same figure, which we talked about in the last class. And we again have this line charge distribution or line oscillator distribution. And the length of this line oscillator is D. And in this length, we pick a particular section of length Δy_i , this array of oscillator is centered at the origin of the coordinate system where x axis is pointing in this direction, y is pointing vertically up, while the z axis is coming out of the plane of the paper.

The point of observation is P which is at a distance R from the origin and from this section y_i , this length element y_i , the distance of point of observation P is r_i . Now since, we are in the Fraunhofer regime. Let me make it very clear that the derivation, which we did in the last class

is a generalized derivation which is valid for both Fraunhofer and Fresnel because the r term the small r which is a function of y, there no approximation or no restrictions has been imposed on that.

Now, we will impose the restriction of Fraunhofer diffraction and what is this restriction this restriction is that R is much-much larger than D. It means the point of observation is very far away and the length of this oscillator, the array of oscillator is very small as compared to R, the distance between the line oscillator array and the point of observation P.

Under these circumstances r (y) which is appearing in the last expression which we derived in the last class, in the denominator of the amplitude as well as in the phase part, it never deviates appreciably from its midpoint value R. Because R which is a function of y is the distance of this oscillator from the point of observation P.

But if the length of this point oscillator array is very small as compared to the distance from the point of observation then r is almost fixed, it will not vary. And let us assume this fixed value of r is equal to R which is the distance of point of observation from the midpoint of the line oscillator array.

Now, therefore the amplitude which is ϵ_L/R will be fixed. Initially in the expression which we derived in the last class it was this, but now today under Fraunhofer approximation we will replace it with ϵ_L/R . This is the transition which we are making in the Fraunhofer domain.

And therefore, this amplitude of the field would be constant and this field is due to the all oscillators in the length element dy. Now the expression of this field at point P due to differential segment of source dy is therefore dE is equal to ϵ_L/R which is the amplitude and this is the phase part $sin(\omega t - kr)$ and dy also contribute to the amplitude. These two terms is the amplitude and which is written here, $(\epsilon_L/R) dy$ is the amplitude of the wave.

Now, you notice that the r which earlier was appearing in the denominator of the amplitude is replaced by R but the r which is appearing in the phase term here is untouched. We are not touching r from the phase. Why? Because phase is more sensitive quantity, with r in the phase we are multiplying k which is $2\pi/\lambda$, λ is a very small quantity which is sitting in the denominator.

Therefore, this phase become very sensitive and this is what is written here, notice that the phase is much more sensitive to variations in r (y) than is the amplitude so that we will have to be more careful about introducing this approximation into it. And therefore, we did not replace r in phase with R. This replacement is done only in the amplitude part. The r is sitting as it is in the phase part.

Since, r is function of y we can expand $r(y)$ to make it explicit function of y and this expansion would be in this form. How? How can we get this expression? Now let us go back. Now you see that we have a length element and r_i is the distance of this length element from the point of observation P or point of observation P is r_i distance away from this length element Δy_i .

(Refer Slide Time: 7:10)

But let us go to a generalized case, suppose we have two point sources, one is S_1 , other is S_2 means which are separated by a distance a and the point of observation P is sitting here and the distance between S_2 and the point of observation is r_2 while the distance between S_1 and point of observation P is r_1 . Now in this case, if we drop a perpendicular from S_2 to $S_1 P$ then this would be the your optical path length difference OPD.

Now this angle, if we draw a horizontal line then this angle will also be equal to angle θ , which is angle here. Now in triangle S_2PS_1 , we can do the following we will use the very basic trigonometry and we can write that $r_2^2 = r_1^2 + a^2 - 2ar_1\sin\theta$, very basic trigonometry. And this relation can be rewritten as follows, we take r_1 common from the right hand side and then divide it to the left hand side we are taking r_1 common from the right hand side and then the bringing it to the denominator of the left hand side.

And therefore, the relation on the right hand side reduces to this expression, a/r_1 whole square to the power half. Now what we will do is that, we will expand this in McLaurin series. Therefore, McLaurin series expansion yields $r_2 = r_1 - a \sin\theta + (a^2/2r_1)\cos^2\theta$, this is what we get. Now if you consider the Fraunhofer diffraction, in Fraunhofer domain, in Fraunhofer domain what will happen we know that this conditions prevail, under this condition $r_1 - r_2 =$ $asin\theta$ this we know.

But let us go back to our previous relation this equation number 22. Now just compare this equation number 22 with this equation derived, they are the same relation, the 22 is nothing but this relation which we derived. Now we have R, instead of R we are having r_1 , $asin\theta$, now instead of asin θ we are having $y\sin\theta$, a is replaced with y, similarly $y^2/2R$ is here in the third term and same thing is here too. And some additional terms are also there but they would be very small therefore we have neglected them.

(Refer Slide Time: 11:06)

Now in this expression 22 where θ is measured from the xz plane and xz plane is this plane. z is coming out of the paper and x is in the plane of the paper and therefore xz plane is a plane which is coming out of the plane of the paper and is perpendicular to the plane of the paper. And therefore, this is the array of line of this point oscillators, and y is also pointing in this direction, and xz plane is a plane which is perpendicular to the length of this array of point oscillator, and all the angles are measured with respect to this horizontal plane.

It means angles are measured like this up and for downward angle it is measured like this. Every angle is measured with this horizontal plane, which is coming out of the plane of the paper in this figure 11. Therefore, make it a point that θ is measured from the xz plane and the third term can be ignored as long as it contributes to its contribution to the phase is insignificant.

This term can be neglected. Why? Because we know y which is nothing but it represents the extension of this line oscillator and we know that the D, the total length of this line oscillator, is very small as compared to R. Therefore, this term can safely be neglected, the third term. And if you neglect the third term because the maximum value of y is $\pm D/2$ and which is this. And D is much much smaller than r therefore D^2 would be even more smaller than R.

And therefore, this third term can be neglected. Here in this case, I have multiplied this third term with a phase and therefore it gives, with a k, therefore this term gives phase. So, we can see here that this term is very small therefore we neglected and this will be true for all values of θ when R is adequately large. This sentence says that when we are in the Fraunhofer domain and when the observation screen is very far then irrespective of the value of the θ the third term wouldn't contribute and it will not contribute anything to the resultant field and therefore it can safely be neglected.

And therefore, the resultant expression of r which we get is $R - y \sin\theta$ then let us substitute this expression of r into equation 21 here. After the substitution, we get this relation. But apart from substituting the expression of r into equation 21 here, we have also integrated it within the limit -D/2 to +D/2 to have the total field distribution at the point of observation P due to all the small length segments.

And it is integrated over dy, it means that we are taking into account all the small length segment along the length of this point oscillator, this array of point oscillator. Now we will have to just solve this integration to have the final value of the resultant disturbance at the point of observation P, a resultant field at the point of observation P.

(Refer Slide Time: 14:44)

Now we solve this integration this is very easy integration and this will give this expression of the field. Now you see that equation number 24 looks very cumbersome but this term is common both in the numerator and the denominator. Therefore, just to simplify the appearance let us introduce a new parameter which we name as β and $\beta = (kD/2) \sin\theta$. Once we replace (kD/2) sin θ by β / in equation number 24, we get equation number 26. Now you its equation number 26 looks a bit simpler.

(Refer Slide Time: 15:24)

Now once the field is known we can also calculate the irradiance. And we know how to calculate it, then this is the expression of the irradiance, here we see that this term is gone because it is a phase part and this will contribute to half only, and this contribution is looking here, and this is the amplitude part which is squared and this is also the amplitude part which is $(\sin\beta/\beta)^2$.

Now, what would be the intensity when θ is equal to 0. Now when θ is equal to 0 this term which is also called sinc function, $sin\theta/\theta$ is also called sinc θ and this term which we call sinc function its value is equal to 1 when θ is equal to 0. Because we know $\beta = (kD/2) \sin \theta$ and when θ is equal to 0 we will have 0 by 0 and once you evaluate it then you will get 1. In this

situation I (θ) would be equal to I (0) which correspond to principle maximum, at the axis the principle maximum irradiance would be equal to I (0).

(Refer Slide Time: 16:46)

Now the irradiance resulting from an idealized coherent line source in the Fraunhofer approximation is therefore given by this expression. Where I (0) is this, this is very much clear from equation number 27. Because when you substitute $(\sin\beta/\beta)^2$ by 1, then what is left is this term, and this is your I (0), the irradiance of principle maxima.

And therefore, the resultant irradiance is given by equation number 29 which is I (θ) is equal to I (0) $\sin^2 \beta / \beta^2$ where $\beta = (\pi D/\lambda) \sin \theta$ which initially was k D/2 $\sin \theta$ and k is $2\pi/\lambda$. Now when D is much-much larger than λ , it means we have a array of point sources and whose length is D and if this point sources emit light whose wavelength is much-much smaller than D then you see that the irradiance drops extremely rapidly as θ deviate from 0.

When D is much-much larger than λ a then this term would be very large, D/ λ would be very large. And the $(sin\beta/\beta)^2$ would behave in such a way that as soon as the θ deviates from 0 it drops off rapidly. And now from equation 26, you see, this is your equation number 26 which is the expression for the field the phase part of the line source is equivalent to that of a point source located at the center of the array at a distance R from P.

Now you see here in this phase part we have $\omega t - kr$, this is a constant R. And what is R? And this R is this distance and this is $+D/2$, this is $-D/2$, origin is situated here, it means R is the distance of center point oscillator of this array. It means in phase term only this center of the array is appearing.

(Refer Slide Time: 19:18)

And therefore, when D is much larger than λ , then radiation will predominantly be in θ is equal to 0 direction. And emission resembles a circular wave in xz plane. Let me explain it in with detail. Now this is our x direction and this is our z direction, y is pointing up. And this is our array of point oscillator and θ is being measured from xz plane, xz plane is this plane.

There are two things, which we must keep into account, the first thing that the case which we are considering here is D is much larger than λ , wavelength is very small as compared to the length extent of the array. In this case, since wavelength is very small D/λ would be very large. Now as long as θ is equal to 0, we are having 0 by 0, and we have some appreciable value of intensity but as soon as we deviate from θ , the sinc function will reduce down to 0, almost 0.

It means, if you plot the intensity here and θ in this direction, in this horizontal direction then as long as θ is 0, you have appreciable intensity and then it is rapidly decaying down its going down to 0. And also make it a point θ is being measured from xz plane. This is our xz plane and this is how we are measuring the θ , the only dominant intensity is only appearing when θ is equal to 0, this angle is 0.

It means we have a line charge distribution which is dominantly emitting in this horizontal plane only. And as soon as we either go up or down, the intensity quickly goes down to 0. It means we will see a circular wave front coming out of this array of point oscillators, closely spaced array of point oscillators.

Therefore, what we will see is that around this array of point oscillator, we will see a circular wave front. And also, if you go to equation number 26 then you see that the phase part, from the phase part we can just predict that it resembles to a phase which have its origin to a point oscillator which is sitting at the origin at the center of the array of point oscillator.

It means that the whole array will emit in a circle and the phase behaves in a way that gives us a feeling that all the emissions are coming from a point source which is at the center of this array. And all the emissions are in a form that generates a circular wave front or circular wave.

It means all this point array will only emit in one plane in R direction the emission would be confined in one plane which is our zx plane. I repeat the emission is confined in a in zx plane only, it is a line charger the line charge is supposed to emit circular wave front but in through our calculation, we came to know that in case when D is much-much larger than λ whole line charge will emit in a way that gives us the feeling that whole line charge had has reduced down to a point charge and this special point charge is emitting circular wave.

But conventionally we know that points charge emit a spherical waveform but no, it is not a spherical wavefront it is a circular wave is being emitted, this specialized point charge. In case when D is much larger than λ . Now let us go to the other extremity where λ is larger than D, where wavelength is very large as compared to the length extent of the point oscillator array.

In this case, what will happen? In this case, let us go back to the expression of β , β = $(\pi D/\lambda)$ sin θ . Now since λ is very large, D/λ would be very small. And this will give very small β and if β is very small then $sin \beta = \beta$. And in this case I(θ) = I(0).

It means the irradiance is then constant for all θ irrespective of the value of θ the irradiance would be equal to $I(0)$, which is a constant. Now in this particular case, our line charge will again resemble to a point source which is again sitting at the center and which is now emitting a spherical wavefront.

It is now whole line source will now resemble to a point source which is emitting a spherical wavefront. Now in all directions because $I(\theta)$ is now a constant the irrespective value of the θ , irrespective of θ, θ is this angle, irrespective of θ, it will emit the same irradiance. And therefore, spherical waves would be emitted.

Therefore, we can safely say that in both the cases when λ is very large or when D is very large, we can anyway replace the line source to a point source. In one extremity when D is muchmuch larger than λ , a specialized kind of spherical sorry circular wave will be emitted. While when λ is much-much larger than D a spherical wave front would be emitted, spherical wave would be limit emitted which is very much true also.

Because whenever we say there is a point charge which is emitting spherical wave. And what is a point charge? How the point is defined? A point is defined as 0 dimensional entity it has neither length not width nor height it is a 0 dimensional entity which is much-much smaller than any wavelength which we consider because wavelength has a certain nonzero extent, its wavelength is always finite but a point has no dimension, it is a zero dimensional entity. Therefore, the second case correspond to the usual point sources, which usually emit a spherical wavefront which is also coming out to be true through this analysis.

While in the other case, when wavelength is very-very small as compared to the length extent of the line source, then in that case we get some special kind of wave emission which we name as at circular wave which remains confined in this horizontal plane xz plane. And the lines source is in this direction so vertically up and down, here it is in vertical direction. Now we can move to the realistic case of single slit.

Till now, we were considering a line source which have some finite length but width is 0 but in real situation, in realistic scenario, all single slits have certain width. The realistic single slit has length as well as some nonzero width. Now how to deal with this type of slits? How to mathematically calculate or mathematically model such a system?

(Refer Slide Time: 28:27)

Now the realistic single slit, we name it as elongated narrow rectangular hole. We name it as elongated narrow rectangular hole. Now we will follow the same analysis, what we did in the in the previous slides. Now how to use that previous analysis, now suppose we have this slit, now this slit has now certain width and let us assume it is b and it has certain length and let us assume that it is l.

Now what we will do is that we will split this slit into a very thin smaller slits. Now we are splitting it into thinner slits and say the thickness of this individual slit is dz, the large number of smaller slits will appear out of this broad slit. Now if we do this then for each smaller width slit, we can use our previous analysis, previous formulation. Now let us see, how it is done.

The usual procedure to follow in the analysis is to divide the slit into a series of long differential strips here. How to divide its dz/l, dz is the width and l is the length. We are dividing all this, the usual slit, into a smaller width slit with the same length. Now we can take out all these smaller width slit out of this and then analyze all these slits independently. Now each strip is a long coherent line source since the thickness is very small then we can use our previous analysis.

Because then thickness can now be neglected, it is now a line source. It corresponds to a line source, so therefore each strip is a long coherent line source and can therefore be replaced by point emitter on the z axis here. Now axis's are shown in this figure, this is our x axis, the vertical axis is y axis and z axis is this, this come again coming out of the plane of the paper,

the slit is in y z plane and the wave is propagating in x direction, x direction is perpendicular to the plane of the slit.

Now you see that you have a source here, it is emitting spherical wave front then we use a lens to make these beams parallel and then they fall on the slit and then they go in different direction due to diffraction and then another lens is used to focus them on a screen. And this is the fringe pattern which we usually see. Why do we see this type of fringe pattern? We will see in next few slides.

(Refer Slide Time: 31:47)

Each such emitter, which emitter? This line strips, this small strips which we have taken out of this very wide single slit is that each such emitter radiates a circular wave in xz plane. Why would it radiate circular wave? Because if you remember when D was much-much larger than λ then this slit behaves, this line source, behave as a point which is centered at the center and is emitting a circular wave.

And this is the case wavelength is much-much smaller than the length of the slit. And here in this particular case in the present example D is l, l is the length of the slit. Then we are taking a slit of width, infinite decimal width dz and length l. Therefore, in this case l is much-much larger than $λ$. $λ$ is very small quantity, the wavelength is very small. Therefore, each such emitter or each such strip will radiate a circular wave in xz plane.

Similar, to what we observed, in the previous analysis. Now there are an array of such strips, array of such emitter then each emitter will emit its own circular wave. Therefore, there will thus be a very little diffraction parallel to the edge of the slit. And since we know that the circular wave means every emission is appearing in θ is equal to 0 direction, as you move up or down, as you increase θ, the radiation irradiance drops down very rapidly.

Therefore, a very little diffraction would appear in a direction which is parallel to the edge of the slit. No intensity, almost 0 intensity, you will observe if you go up or down, everything will be confined in the xz plane, which is perpendicular to the length of the slit. The third point, the problem has been reduced to that of finding the field in xz plane due to infinite number of point sources extending across the width of the slit along z axis.

Now, let me explain it further, now this was our initial slit which had width b and length l. Now we created very huge number of small line sources and each such strip due to point number 1 and point number 2 here can we reduce to a point source. Each such strip can now be reduced to a point source and the width of this array of point source would be b but now it would be length here. Because each line source can be reduced down to a point source now, which is emitting a circular wave.

Therefore, each strip I am replacing with a point source and each point source is emitting a circular wave. Now the problem is reduced down to a line source which is of length b now and which is in this direction with here, which is now in z direction, this is the z direction. And this is now our array of point sources, this is our line source now which is now pointing along z axis. Thus, we created an array of point sources or line source which is now along z axis and whose length is now b.

Now, we then need only to evaluate the integral of the contribution dE from each element dz in the Fraunhofer approximation. In the Fraunhofer approximation, now we will evaluate the field due to this horizontal line source.

(Refer Slide Time: 36:22)

And what is the formula for irradiance of a line source, we have already derived I_{θ} = $I_0(\sin\beta/\beta)^2$. But the β here is different from what we did in our earlier analysis, in earlier analysis in β we were having kd/2, d was the length of the line source but here the length is now b , therefore I put it in red. And this equation number 30 is the same equation which we derived earlier here, equation number 29 is the same expression.

But the definition of β is now changed and parameter b which is the width of the single slit is now length of this line array, line oscillator array. Now this θ is also different. What is the difference? θ is now measured from the xy plane earlier the line source was like this and θ was measured from the horizontal plane. But now since this line source is now in this direction. Therefore, θ would be measured from this plane, this vertical plane, this is how the θ would be now measured.

And this is why it is written, θ is measured now from xy plane and if you go to the figure here, then you see that now since the line charge is now along z axis and the perpendicular plane here is now xy , xy is the perpendicular plane, then the angle would be measured from this perpendicular plane, which is your xy plane. Therefore, θ is now measured from xy plane.

Make it a point, now note that here the line source is short. Why short? Because this is the single slit we started with, its length was l and width was b and width at the beginning, at the very beginning, was chosen such that it is very much small as compared to l. In our previous analysis the line source was of length D which was a bit big. Because length is always big as compared to width, this is the convention.

But the length of now the new line source is b only which is very small. Therefore, D is in the definition of β is replaced by b. Now β is not large because b is very small here and although the irradiance falls off rapidly, higher order subsidiary maxima will be observable. As we said the intensity goes rapidly reduces to 0 as b deviates from θ is equal to 0 but now the length of the array is very small that the intensity will or irradiance will go down to 0 but it will again build up, and then again go down to 0, again build up, again go down to 0.

Then how to calculate all these maxima and minima? We have the expression of irradiance here in equation number 30, a very quick answer is to utilize the knowledge of differential calculus, calculate the first order derivative of irradiance with respect to β and equate it to 0. From there we will get the maxima minima. Let us do this, the extrema of $I(\theta)$ occur at values of β that causes $dI/d\beta$ to 0, well known.

Let us take the derivative of equation number 30 with respect to β and this derivation gives this expression. Now if you equate to 0 then you see two things, in the numerator we have multiplication of $\sin \beta$ with $\beta cos \beta - sin \beta$. It means either this term would be 0 or this term is 0.

(Refer Slide Time: 40:49)

Now the irradiance has minima equal to 0, when $\sin \beta = 0$. Let us assume that $\sin \beta = 0$. When $sin \beta = 0$ then β will have these values for $sin \beta = 0$, $\beta = \pm \pi$, $\pm 2\pi$, $\pm 3\pi$ and so on values. It means, we will have minimum for these values of β . And we know the expression of β from there we can calculate the values of θ and once the θ is known we can easily predict the positions of minima in single slit diffraction pattern.

Now, there is second possibility to the second term, second multiplicative term, it may also assume 0 value. And if let us equate this term is 0 then from we get $tan\beta = \beta$. Now this is a transcendental equation and which is very difficult to solve, people usually solve it numerically using bisection method. And therefore, the solution of this transcendental equation can be determined graphically, the solution to this transcendental equation represents the extremum.

Now, how to solve this equation number 35? On the left hand side, we have $tan\beta$, on the right hand side we have β , to solve it, we plot $tan\beta$ and β separately and the point of intersection would be the solution which will satisfy both left and right hand side.

(Refer Slide Time: 42:26)

Now the point of intersection of curve f_1 is equal to $tan\beta$ with a straight line f_2 is equal to β or common to both and only one such extremum exist between adjacent minima. Why? Now see in this figure, figure on the right hand side the f (β) , the β is plotted on the horizontal axis and functions of β are plotted on the vertical axis. This $f_2(\beta) = \beta$ is the straight line which is plotted here, while the f_1 which is equal to tan β is plotted here, these are the tan β .

Now you see that the straight line is intersecting with this tan β curve at several points. The first is appearing here, the second is appearing here, the third is appearing here and we also know that we have minima at β is equal to $\pm \pi$, $\pm 2\pi$ which are here, this is the minima, this is the minima, this is the minima and between two adjacent minima we have only one maxima. See this is the maxima, this is the maxima, this is the maxima. And this is what is said, only one

such extremum exist between adjacent minima, maximum is extremum we cannot comment right now what it is.

But the zeros of irradiance occur when $\beta = m\pi$, this we have calculated here, $\beta = m\pi$. And from here because $\beta = b \sin \theta_m$ which is m λ now, this is the same expression which is written in an expanded form where m is this integer.

(Refer Slide Time: 44:11)

And with this, if we plot the irradiance here on the vertical axis the relative irradiance is plotted on the horizontal axis both β and sin θ is here shown. Now you see that along θ is equal to 0 we have very huge irradiance and its drops off very rapidly, like its falls very rapidly as you increase θ.

Now it goes down to 0 at π , 2π , 3π and between adjacent minima, we have one maxima, between adjacent minima we have one maxima. The β values are mentioned here and sin θ values are mentioned here. And this is how a single slit diffraction pattern looks like, slit which have a certain width. Now suppose you increase or reduce the width of the slit, suppose this is for this slit. What will happen if you reduce down the width of the slit. In this case what will happen is that your pattern would be broadened, you will get like this. While for this, you will get this.

Smaller is the width of the single slit wider would be the pattern and this is also clear from this figure. This smaller is the b, larger would be the θ. Now this is all for single slit, I finish my lecture with this, thank you for listening me.