Applied Optics

Professor Akhilesh Kumar Mishra
Department of Physics
Indian Institute of Technology, Roorkee

Lecture 37

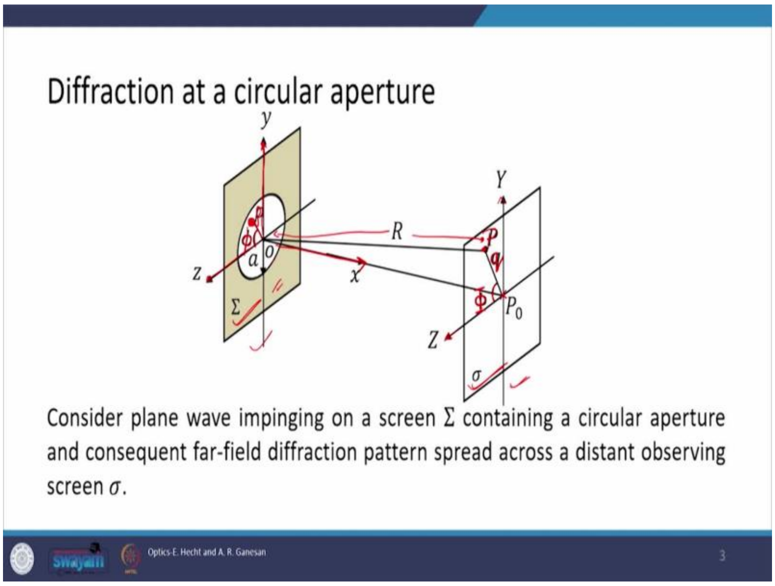
Diffraction at a Circular Aperture

Hello everyone, welcome back to class today, we will start our new topic in module 8, which is circular aperture. In the last class, we talked about a diffraction pattern which is formed using rectangular aperture. And in case of rectangular aperture, we observe that the fringes are spread both horizontally and vertically. And if the vertical extent of the aperture, rectangular aperture is larger as compared to the horizontal extent, then the fringes in the horizontal directions are spread more, their extent is larger, the extent of a spread of these fringes are inversely proportional to the extent of real aperture.

If the aperture width is smaller, the fringes will spread more, but it would be in opposite direction while in case of circular aperture, since there is a circular symmetry or rotational symmetry we assume or we may predict that the fringes would be circular, the symmetry predicts that the fringes would be circularly symmetric, we will repeat the same mathematics which we did in case of rectangular aperture. And the expression which we derived in the rectangular aperture, we will make use of that.

Here in circular aperture, we will introduce some spatial functions which are called Bessel functions, we will make use of Bessel function and with this we will see the kind of fringes we observe and the prospective applications of the observations which we will make in a while.

(Refer Slide Time: 2:27)



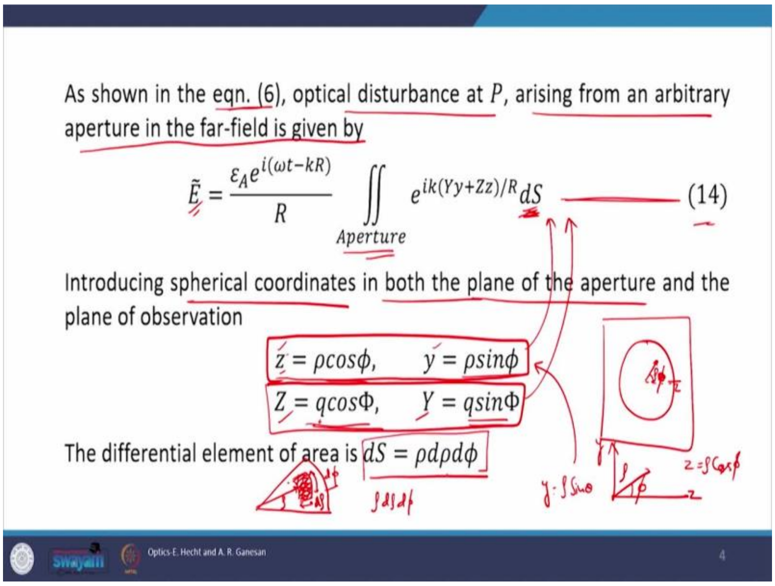
Now, in this figure the Σ plane represents the aperture plane while σ plane represents the screen plane or plane of observation. The circular aperture which are considered here in the Σ plane, it is of radius small a and the origin is at the center of the circular aperture, the x axis is coming out of the plane of the circular aperture while y and z axes are extending in this vertical and horizontal direction as is shown in this figure.

Now, similar to aperture plane, in the screen of observation plane, we have another axis system which are designated by Y and Z, x axis in both the systems they coincide. Now, we launch again plane wave on the aperture and then we observed diffraction pattern on the screen. Now, here again we assume small differential area which is at a distance a ρ from the center in aperture plane.

Now, this is at a distance ρ and ρ which is joining the origin and the differential area, it is inclined at angle φ(small phi) from the z axis as is shown here in this figure. Similarly, in the screen plane or plane of observation, the point of observation P is at a distance q from P_0 which is origin in screen plane and the corresponding angle here is Φ (capital phi).

Observe the difference between the two phis. Here, in the aperture plane, the angle is designated by φ while in the observation plane the angle is designated by Φ. Now the point of observation P is at a distance R from the origin in aperture plane and since we are in the Fraunhofer regime R is much much larger than any other relevant distance in the problem at hand. It is larger than the radius of circular aperture.

(Refer Slide Time: 5:17)



Now, we will borrow an expression from our previous class which we derived for random aperture and this expression is equation number 6 which is given here and the same equation is now renamed by equation number 14. Now, that optical disturbance P arising from an arbitrary aperture in the far field is given by this expression this we have derived in the last class therefore, I will not devote more time on this.

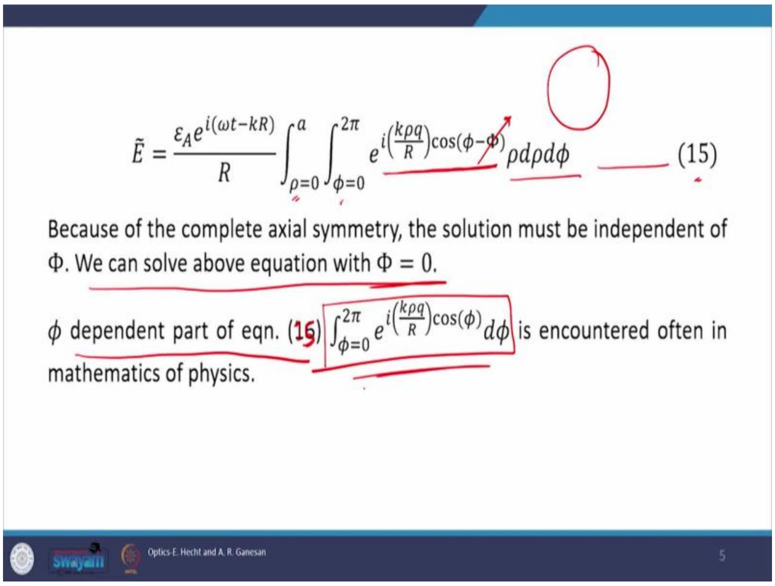
The integration is done over the aperture and dS is the area of the differential element, differential area element. Now, since the aperture at hand is circular aperture, which holds a circular symmetry therefore, to enjoy the symmetry of the system, we will move into spherical coordinate. Now, in the spherical coordinates in both of the planes, that is aperture plane and the plane of observation, we can replace the coordinate system y, z and Y, Z with these expressions because we know, like suppose this is your aperture plane and this is your circular aperture and any elemental area which is at a distance ρ and which is making an angle θ or angle φ with the Z axis then this ρ and φ, these are the two variables in the polar or spherical coordinate system. In Euclidean geometry or in x y coordinate system, x y are the independent variables or variables while in polar geometry or in a spherical coordinate, it is ρ and φ which are independent variables or variables only.

Now, we can relate these two in aperture plane, we can represent z= ρ cosφ, which is very much obvious in this figure because this is ρ, this is your z axis and this is your φ therefore, z= ρ cosφ. Similarly, y which is pointing up, y= ρ sinφ this is what is done here. In the aperture plane z is replaced by ρ cosφ and y= ρ sinφ.

Similarly, in screen plane, screen of observation the point of observation is to distance q and this point the line joining the origin at the screen plane and the point of observation is making an angle Φ with a Z axis therefore, Z, Y here would be replaced by q cosΦ and q sinΦ, this is very much obvious.

Now, we will substitute these new expressions of z, y, Z, Y in equation number 14. Now, one more thing which we must not forget is that dS which is the area of the element in cartesian coordinate system dS was express by dx dy while in spherical coordinates dS would be represented by ρ dρ dφ why? Because suppose we have arc and in this arc this is the area element which we are picking up suppose, this distance is ρ and this extension in radial direction is dρ and the change in the angle like this differential angle is dφ then the area of this element would be ρ dρ dφ. This is what is written here dS= ρ dρ dφ.

(Refer Slide Time: 9:38)



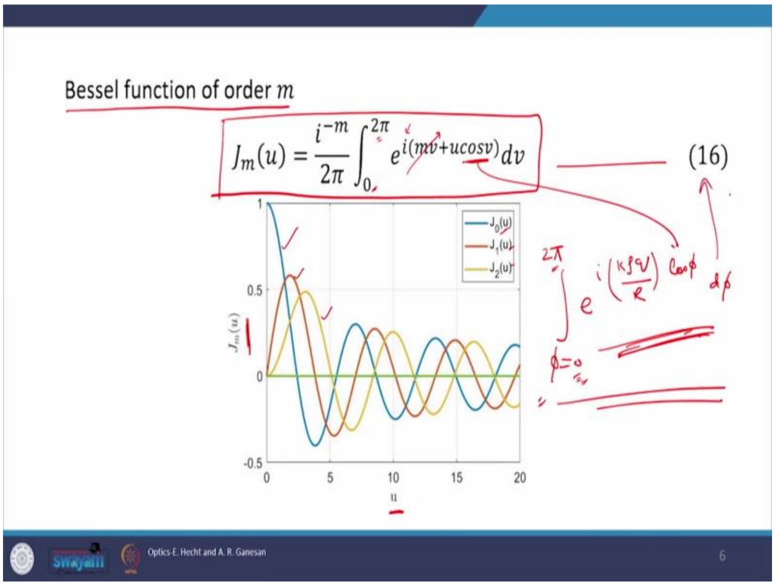
Now, with all this substitution and after putting the appropriate limit, the expression of total field at point of observation P would be given by equation number 15 where dS is replaced by ρ dρ dφ, y, z, Y, Z, they are also replaced with the appropriate expression. Of course, the limit of integration for φ would vary from 0 to 2 π because it is two full circle therefore, φ will go from 0 to 2 π and ρ which is the radial distance of area element this will vary from 0 to a, a is the radius of the circle or radius of the aperture.

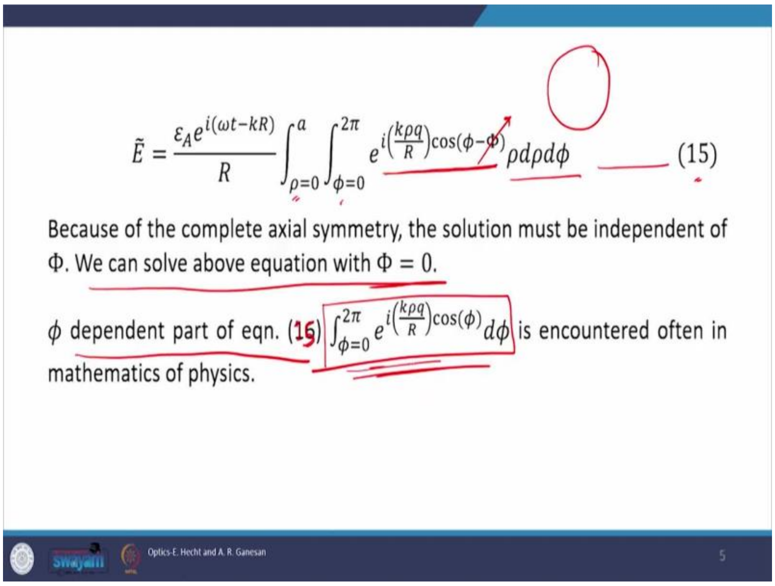
Now, because of the complete axial symmetry the solution must be independent of φ, because irrespective at what angle you are looking at, the system is symmetric. Therefore, we can just by looking at equation number 15, we can guess or just by looking at the aperture shape we can guess that we can solve above equation with Φ=0 and this will give us the same result because the solution must be independent of Φ.

Therefore, in equation 15, let us look for φ dependent part and what are the φ dependent part here it should be equation number 15, I mistaken here. A φ dependent part in equation number 15 is this because we replaced this Φ=0 now. Now, this particular integration is very tough to deal with. And it is encountered often in mathematical physics problem so how to deal with this type of complex integrals.

Now, there are some special ways to solve such an integration and one of them is using Bessel functions, which is a special kind of functions, how the Bessel functions are defined?

(Refer Slide Time: 11:51)

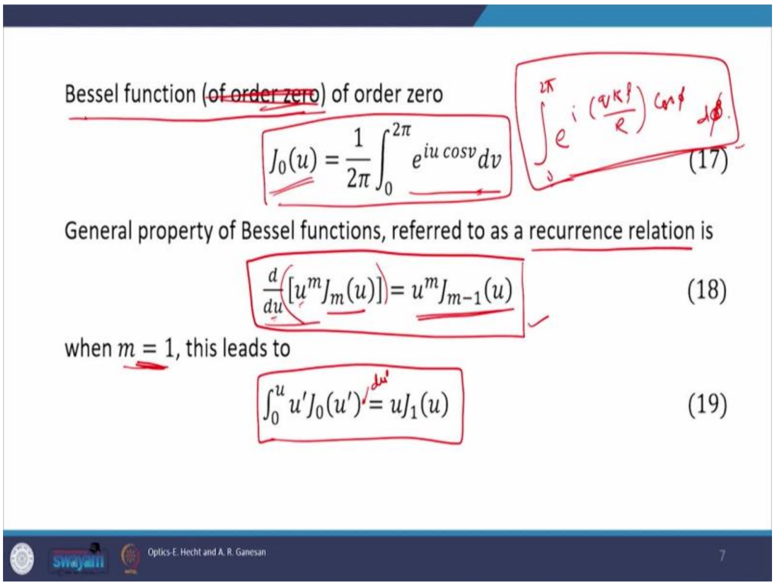


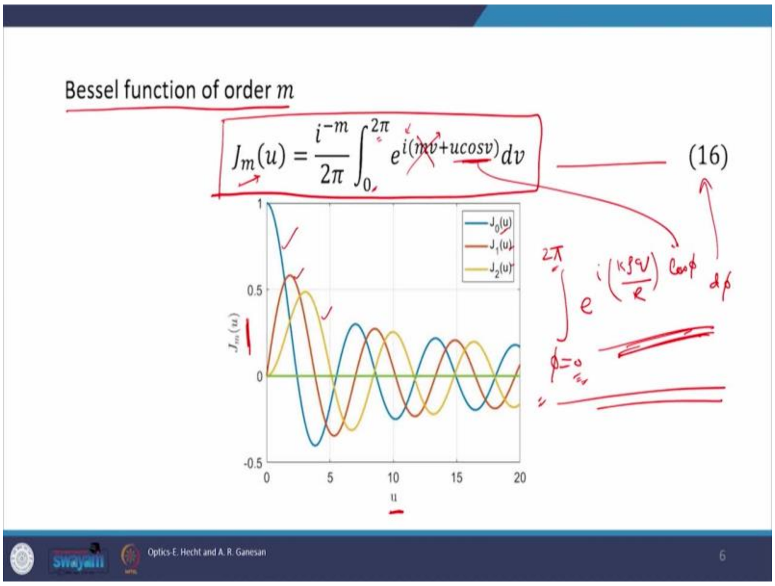


Bessel functions of order m is given by this expression, expression number 16 and if you plot this Bessel function with respect to the independent variable u then what you see is that for different orders of this function, m is the order of this Bessel function, you see different kinds of variations here yeah, for zeroth order you get this type of variation, for first order this type of variation, for second order this type of variation, these are the functional forms of Bessel function of order m.

Now, you see that this function which we want to solve, let us write it in the next slide, it is φ is equal to 0 to 2π and then $e^{ik\rho q/R}$ and then in the exponent again $cos\varphi\ d\varphi$. Now, let us compare this expression with equation number 16. And then you see that cos function is here. And Iota (i) is also here, it is also similar the limit of integration is again 0 to 2π here, it is matching, if we somehow neglect this, then we this integration now, then resembles with equation number 16.

(Refer Slide Time: 13:28)





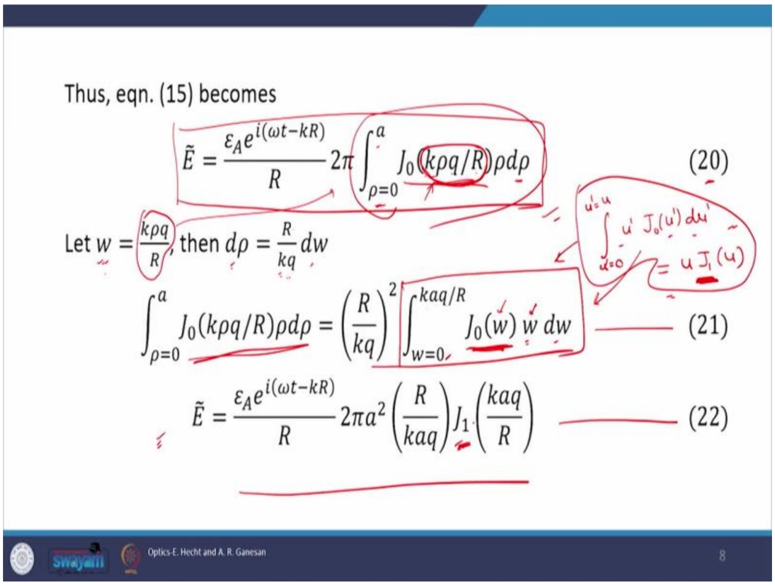
Now, let us see more into it. Let us see how does a Bessel function of order zero looks like. Now Bessel function of order zero looks like this. Here what I did is that in this expression, I just substituted 0 for m, if you substitute m is equal to 0 then this term will go away you would be left with u cos v dv.

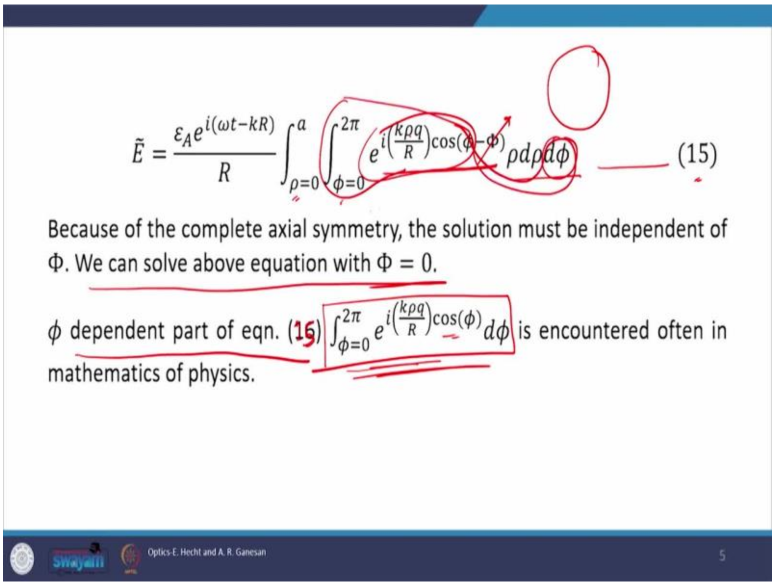
Now, this is more close to our expression which is $\int_0^{2\pi} e^{ik\rho q/R} cos\varphi d\varphi$. Now, you see that perfectly resembles with equation number 17 therefore, once we know the value of $J_0(u)$, we know the solution of this integration. Now, there is one more very important property of Bessel function which we will use and this property of Bessel function is called recurrence relation.

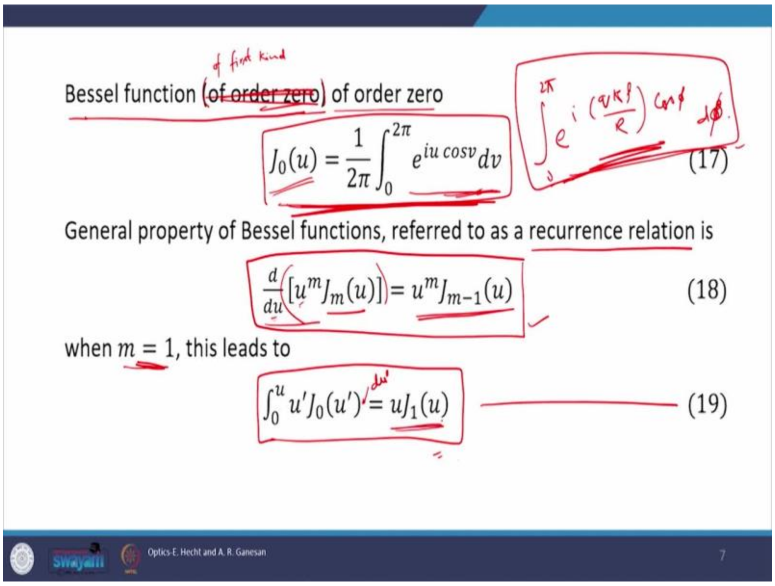
Now, what is the recurrence relation? This recurrence relation says that, in Bessel functions if you multiply u^m, m is nothing but a number and u is independent variable here and if you differentiate this with respect to the independent variable that is u, then you get this function $u^m J_{m-1}(u)$, now this is called a recurrence relation.

Now, when m=1, if you substitute m=1 and equation number 18 then we get this expression here, du' is missing here, the integration is with respect to u', u' is another variable. Now, with this in hand we will try to evaluate the integration.

(Refer Slide Time: 15:29)







This is the expression which we get using Bessel function in equation number 15. Now, let us go back to our equation number 15, this is our equation number 15. Now, in this equation number 15, we are replacing this part with this taken into account, this whole part is now being replaced with J_0 , we know because there is a one to one correspondence of Bessel function of 0 order with our integration.

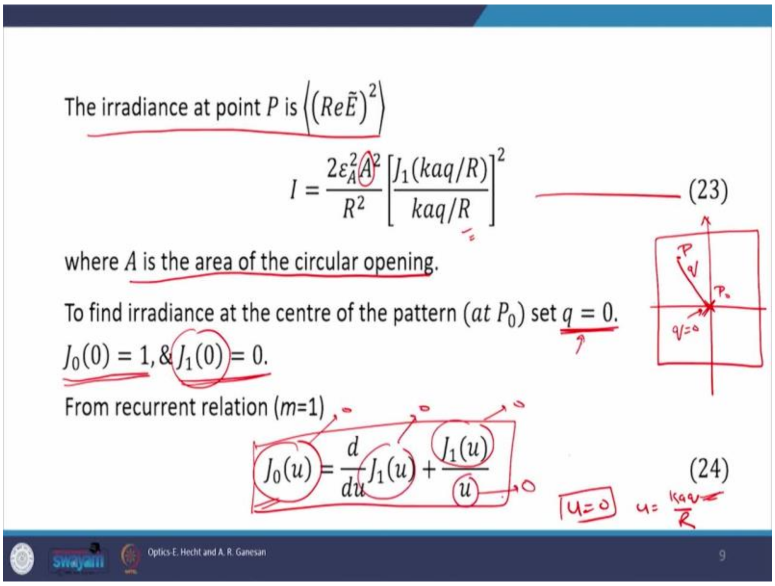
Therefore, we are only left with d ρ integrand, the integrand which is now depend upon ρ and ρ is varying from 0 to a, the equation number 15 now get a bit simplified. Now we will have to integrate the Bessel function of order 0. This Bessel function J is called order 0 and it is Bessel function of kind one. Let us also write it here instead of writing it of order 0 let us also say that in bracket it is of first kind, there are different kind of Bessel function, but the first kind Bessel function of order 0 looks like this.

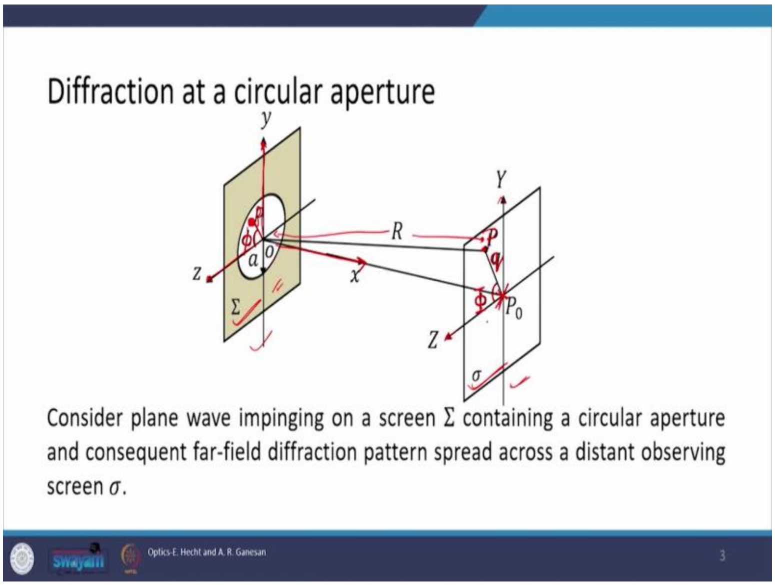
Now, those who study mathematical physics they must be knowing about these kinds of special functions Bessel function, Hermite Gauss functions. Now, once we have equation number 20 then we see that here it is a very big like so, many parameters are involved here kρq/R. So, let us replace them with w. Let us introduce a new parameter w which takes care of all these k, ρ, q and R.

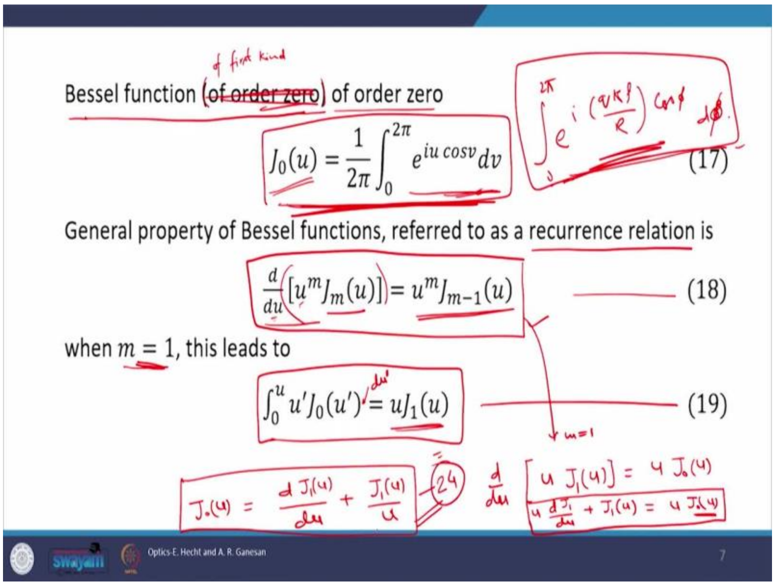
With this dρ would be given by (R/k q) dw and then let us substitute them back into question number 20. And after a bit of mathematics, we get equation number 21, this integration now looks like this. Now, you seen in equation number 21, we have an integral, this integral which is done on J_0 (w) and w varies from 0 to kaq/R. So, let us go back to the recurrence relation which we just looked at a minute before let us go particularly to equation number 19 and equation number 19 says that if you have, let us write equation number 19 in the next slide. Equation number 19 says that if you have a function which is varying from 0 to, sorry u'=u and here it is $\int_{u'=0}^{u'=u} u' J_0(u')\, du'$ then this is equal to let us go again back to question number 1 this must be equal to $uJ_1(u)$.

Let us apply this in this, here you see that we have J_0 and instead of u' we have w again, w this is one to one correspondence now. Therefore, we can use this property here in equation number 21 so after using it now equation number 20 reduces to equation number 22, a bit simplified but now here we have Bessel function of first order and instead of zero order Bessel function. Now we have first order Bessel function, but this integral is now gone. It is now in terms of J_1 only. How to evaluate it, we will see.

(Refer Slide Time: 19:21)







But once the field is known, we can easily calculate the total irradiance at the point of observation P which would be given by equation number 23. Here in this expression, you see that a new parameter which is A which is the area of the circular aperture, which is nothing but πa^2. Now to find the irradiance at the center of the pattern that is P_0 here, let us go back to the first slide, where we have all the arrangement here.

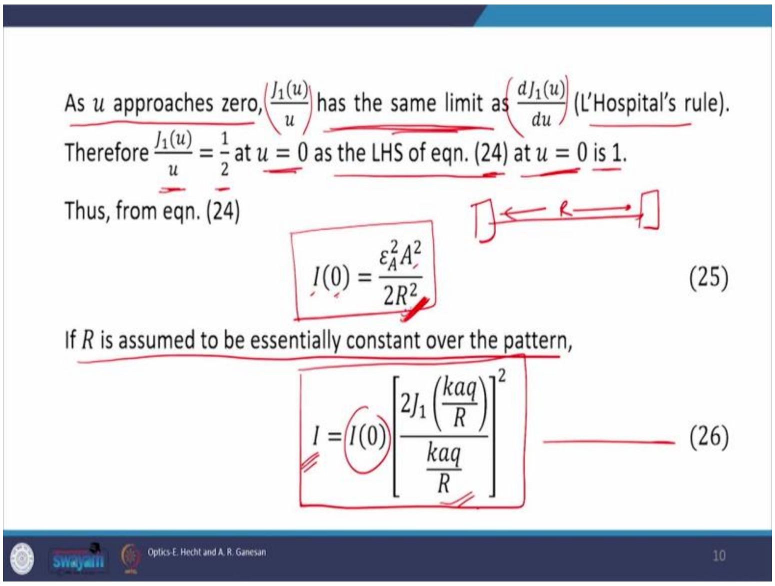
This is the center of the pattern on the screen. Now to calculate the irradiance at that point what we will have to do is that we will set q is equal to 0 why? Because if this is the screen plane and these are your axes and this is your point P and this is your point P_0, then P_0 to P, this separation was q if you want to calculate the irradiance at point P then you should put q is equal to 0, this represent point P_0.

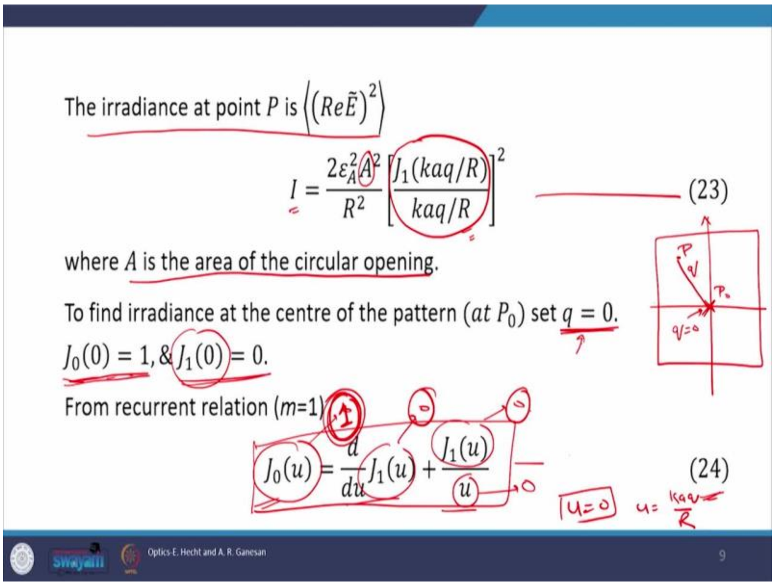
Now, if you set q=0 in equation number 23 and also use J_0 (0)=1 and J_1(0)=0. Then from the recurrence relation, where is the recurrence relation, this is our recurrence relation, question number 18. Now, let us rewrite the recurrence relation, for m =1 then this recurrence relation for m=1 will be like this$d/du[uJ_1(u)] = uJ_0(u)$.

Now, let us differentiate it with respect to u then you will get $u\, dJ_1/du + J_1 = uJ_0(u)$. Now, this is the expression which we got from the recurrence relation. Now, with this recurrence relation we can calculate J_0 (u), which is given here J_0 (u), J_0(u) would be given by from there $J_0(u) = dJ_1(u)/du\ + J_1(u)/u$ this is what we get and this is what exactly is written in that slide. In this equation number 24 is nothing but this expression, this is your equation number 24 in the slide ahead.

Now, with equation number 24 we will now try to evaluate the irradiance at q is equal to 0 but when q=0, u is 0 because u is nothing but k a q/R, once you substitute q=0, u=0, when u is equal to 0, then $J_1(0)$ as we know J_1(0)= 0, J_0(0)=1. Therefore, this quantity would be 0, this quantity would be 0, this quantity would be 0, this quantity would be 0. Is not?

(Refer Slide Time: 22:48)





Now if this is so then we can write that as u approaches 0, $J_1(u)/u$ has the same limit as $dJ_1(u)/du$ here you can use L'Hospital's rule because the numerator and denominator they both are simultaneously 0. Therefore, let us go back, now in this equation 24. Now $J_0(0) = 1$, I mistakenly wrote it out as 0 but it is 1, you see that on the left hand side we have 1 while on the right hand side we have 0 by 0, here we will use L'Hospital rule and with this since the both quantities they has the same limit with L'Hospital rule, we can clearly see that as u approaches 0 the terms in equation number 24, the terms on right hand side of equation number 24 they both has the same limit while on the left hand side we have 1 unity here.

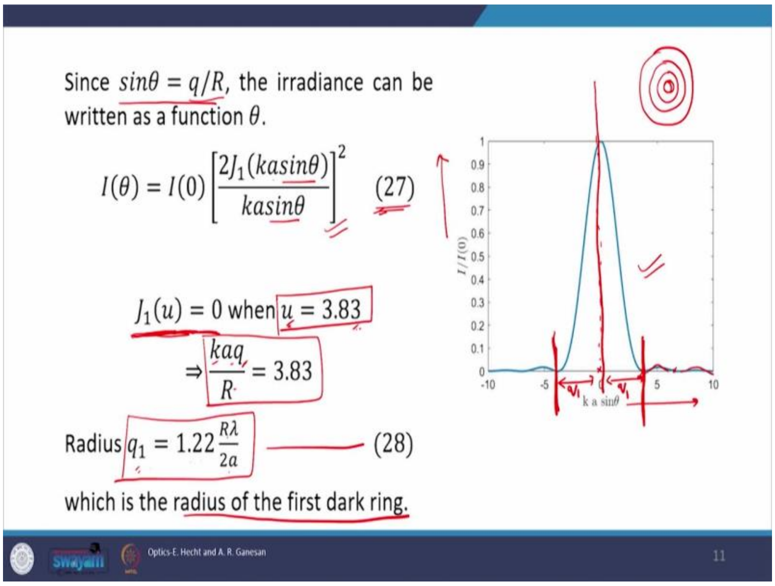
The left hand side term has value which is equal to 1 while the right hand side of the equation 24 has two terms which are having the same limit. It means these two terms would be equal and would be equal to half. Therefore, $J_1(u)/u = 1/2$ at u=0 as the L.H.S of equation 24 at u=0 is 1.

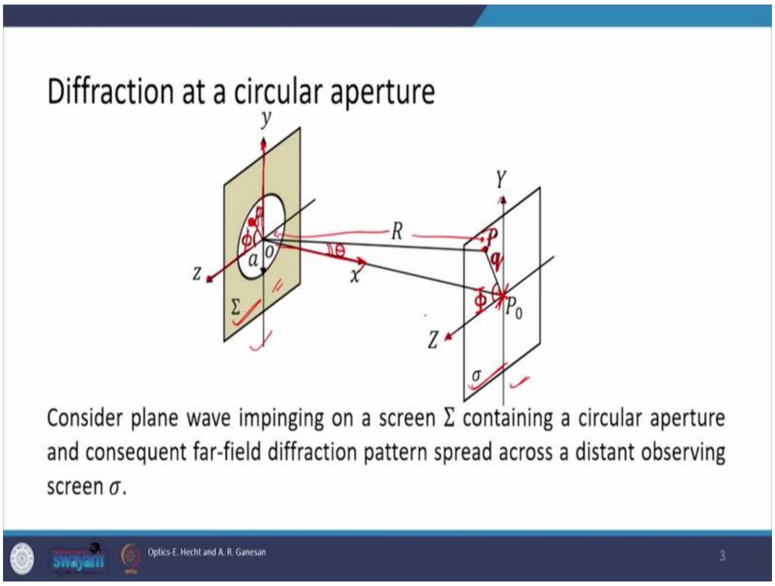
Now if $J_1(u)/u = 1/2$ then again let us go back to question number 23 and see here it is $J_1(u)/u$, and this value would be half and if we square it then this would be 1 by 4. Then after this substitution, the irradiance at center. These all calculations are done for u=0, it means at the center of the screen. I at the center of the screen would be equal to $\epsilon_A^2 A^2/2R^2$, which is irradiance at the center at θ= 0 direction.

Now, if R is assumed to be essentially constant over the pattern because it is everything is symmetric because there is a circular symmetry. This is the screen and this is your aperture plane and R is very huge. Now, if R is assumed to be sincerely constant over the pattern, then the final irradiance would be given by equation number 26.

This is the irradiance at the center which is I (0) and the rest of the term is here which we borrowed from equation number 23, this is the expression of irradiance, final expression of an irradiance due to the circular aperture.

(Refer Slide Time: 26:04)





Now, since $sin\theta = q/R$, let us go back to the first figure, you see here that this is q and this is R. And if you say that it is θ then sinθ would be q/R, therefore, this expression of irradiance here q/R can be replaced by sinθ and sinθ will appear here both in the numerator and denominator.

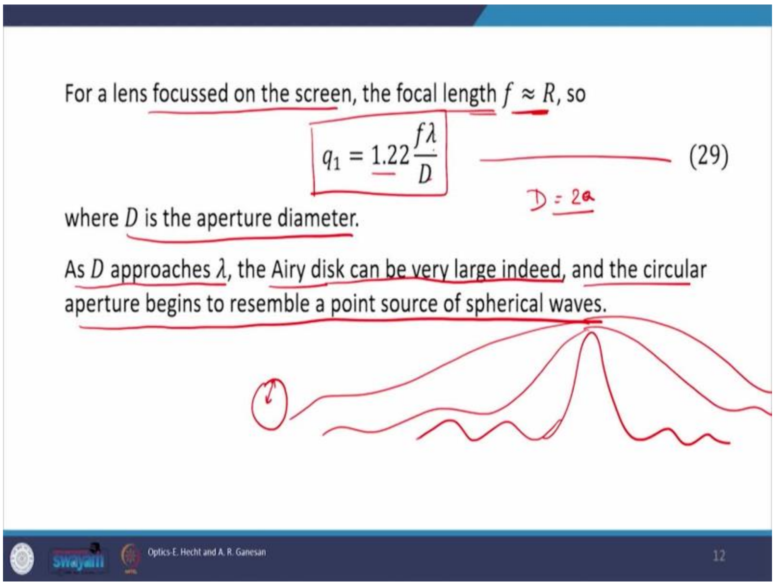
Now, if you plot this equation then you see this type of curve, here on the vertical axis relative irradiance has been plotted while on the horizontal axis k$a$sinθ is plotted. Now, this has a maxima here which is at k$a$sinθ is equal to 0 or θ =0 and then a minima and then secondary maxima again minima and then again maxima and then minima, this type of pattern is called Bessel pattern.

And if you rotate it around this axis, then you will find a concentric circular ring pattern. Therefore, circular aperture gives rise to a Bessel pattern and this would look like a concentric circular ring and this is airy function. Now, if we want to calculate the position of first minima, then what we will have to do is that we will somehow make right hand side of equation number 27 to 0 .

For this, let us see when J_1 goes to 0 and then let us calculate that value of u which makes $J_1(u) = 0$ for this we substituted $J_1 = 0$ and from there we what we found is that at u is equal to 3.83, J_1 goes to 0, it means when u is equal to 3.83, we will have first minima and when you substitute for the expression of u which is ka sinθ or k a q/R therefore, from there you can calculate the radius of the first dark ring which would be given by q value.

Let us designate the radius of first dark ring by q_1 and q_1 would be equal to 1.22 R λ/2 a and q_1 would be the radius of the first dark ring, this radius would be given by q_1, the radius of the first dark ring in circular aperture diffraction pattern. It is a very important relation equation number 28, $q_1 = 1.22R\lambda/2a$, where 2 a is the diameter of the circular aperture.

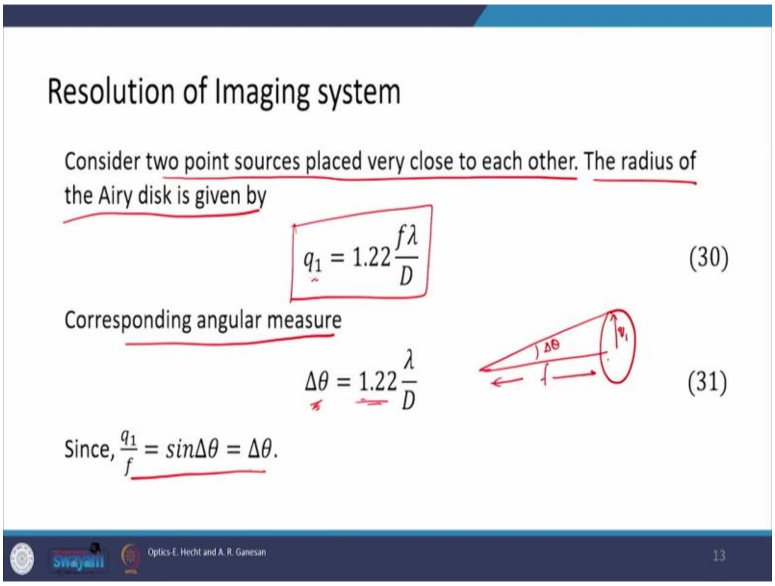
(Refer Slide Time: 29:21)



Now, once we have this, now this is a very broad implication because circular aperture can easily be replaced with a lens also, if you have a lens of diameter D, then this lens will also give similar type of diffraction pattern, for a lens which is focused on the screen, the focal length would be given by R. Now in this particular case, the radius of the first dark ring would be given by 1.22 f λ/D where D is the aperture diameter which is equal to twice of a.

Now from equation number 29 what we can see is that as D approaches λ, the airy disk, this pattern which we saw in the last slide, this airy disk can be very large indeed and the circular aperture begins to resemble to a point source of a spherical wave. Because if you reduce the radius then what will happen is that, the pattern will broaden, you will have this type of pattern then for larger D or for larger a and with a further reduction there would be more broadening and therefore, when the size of the circle is reduced to a point then you will see that the spherical waves now get generated, you will see spherical waves coming out of the this point source.

(Refer Slide Time: 31:13)



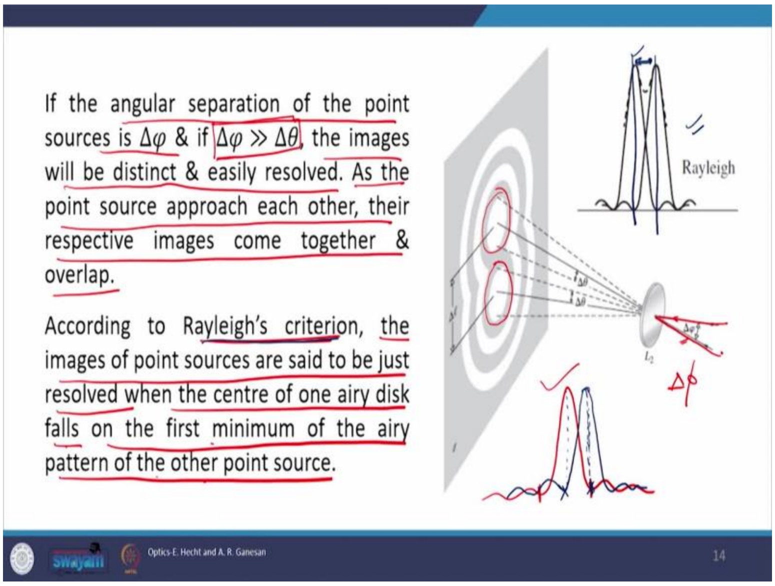
Now, coming back to the applications of the circular aperture diffraction study, this circular aperture as you saw it can be replaced by a lens or it can be replaced by point source. Therefore, the same study can be used to evaluate or decide the resolution of some imaging system how it is done, we will see here.

Now consider two points sources placed very close to each other. Now, they are point sources which is a limiting case of circular aperture, they will form their own diffraction pattern which would be nothing but airy disc. Now, the radius of the airy disc as derived in the last slide would be given by q_1 which is the radius of the first minima or first dark ring.

Now, once the radius is known, we can also calculate the corresponding angular width of the first dark ring. Then corresponding angular measure would therefore be given by $\Delta\theta$ where $\Delta\theta$

is nothing but q_1/f which is given here, the q_1/f would give $\Delta\theta$ which would be equal to 1.22 λ/D, these are the angular width of the ring. Suppose this is your first dark ring then the radius of the dark ring is given by q_1 and the corresponding width is given by $\Delta\theta$ where this is f.

(Refer Slide Time: 33:01)



Now, once it is known, now suppose we as stated in the previous slide suppose we have two points sources and these two points sources has certain angular separation which is given by $\Delta\varphi$. There are two sources and the rays or the light is coming from these two sources at an angle $\Delta\varphi$.

Now, if the angular separation of the point sources is larger than $\Delta\theta$, if this condition holds, the images will be distinct and easily resolved, why? because, if the angular separation between the sources is larger than $\Delta\theta$ then the first minima of two sources would be separated and you can clearly resolve them as is shown here in this figure. The first minima of the first sources being found here and the second source first minima is being found here which are well separated.

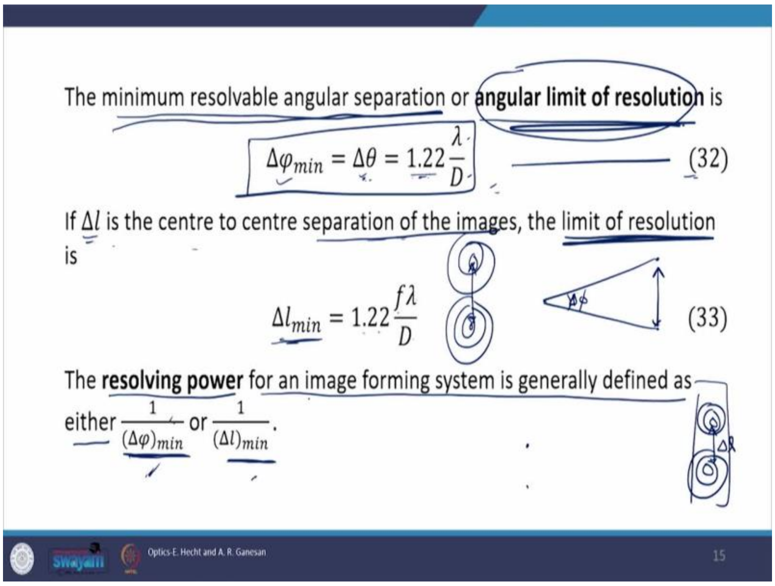
As the points source approach each other, their respective images come together and overlap. And once they are overlapped, we will not be able to clearly resolve them or distinct them. Now, here again Rayleigh's gave a criterion which is called Rayleigh's criterion. And, according to this criterion, the images of the point sources are set to be just resolved when the center of one airy disk falls on the first minimum of the airy pattern of the other point source.

What it says is that suppose this is the airy pattern of the first point source and this is the first minima of this first point source. Now, as per this statement when the center of one airy disk falls on the first minimum of the airy pattern of the other point source means the second source center must fall on this point in this particular case and if they are closer than this criterion, then they would not be resolved. This is the minimum possible separation where in the two point sources can be resolved.

Now, as per the Rayleigh's criterion, the maxima of the other should fall on the minima of the first, this how, you see that on the first minima the center of the first, second falls. Similarly, on this first minima of the other the center of the first falls and this is what is shown here, you see the center is falling on first minima and this center is falling on the first minima of the other source in this case only and if the sources are closer than this separation, then these peaks or the center of the two pattern would now be closer than the separation and we would not be able to resolve them.

The least separation between the two point sources is given by this Rayleigh's criterion and under this least separation, the center of one must fall on the first minima of the other. If they are closer than this, they would be said to be unresolved, if they are farther than this, they are obviously resolved. The minimum separation is given by Rayleigh's criterion. Now, suppose that $\Delta\theta$ and $\Delta\varphi$ is equal that is they are at the limit this is the minimum separation between the two sources where they are said to be resolved.

(Refer Slide Time: 37:25)



In this case, the minimum resolvable angular separation, the minimum resolvable angular separation or angular limit of resolution is given by equation number 32 which say $\Delta\varphi_{min} = \Delta\theta$a and $\Delta\theta$ as we derived in the previous slides it is equal to 1.22λ/D.

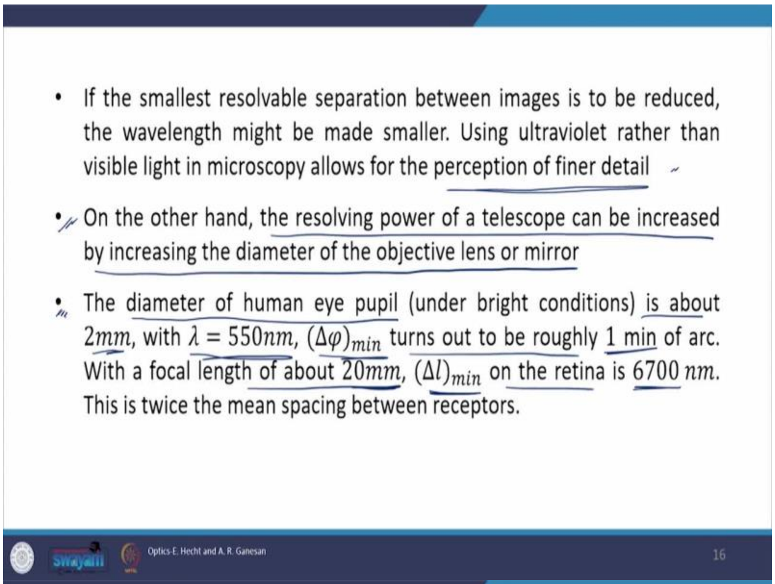
Now, we were having two sources and their angular separation is $\Delta\varphi$ and the minimum value of $\Delta\varphi$ is given by equation number 32. Now, suppose that the two sources, the center to center separation between the images of these two sources is Δl, there are two sources which are the angular separation between the two sources $\Delta\varphi$ and they are forming their own airy disk.

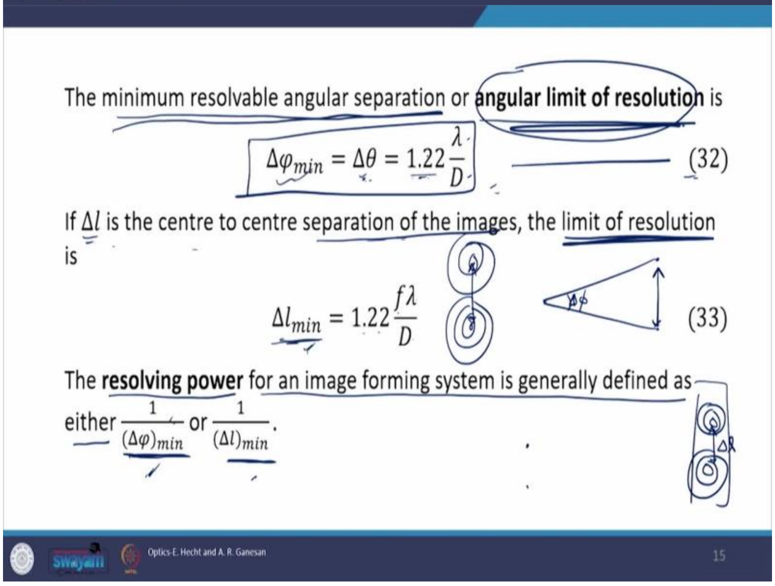
Now, if the center to center separation of the images of this airy disk is Δl then we define limit of resolution which is Δl_{min} as 1.22 fλ/D. Another very important terminology is defined as resolving power. The resolving power of an image forming system is generally defined as either $1/\Delta\varphi_{min}$ or $1/\Delta l_{min}$. If we are in the angular domain then we use $\Delta\varphi_{min}$ and if we are in the linear domain, length domain then we use Δl_{min}, where Δl is the center to center separation of the images.

We have point sources here, we have a screen here, these two points sources form their own images or their own disk and center to center separation of this disk is Δl. The minimum value of this distance is Δl_{min} which is called limit of resolution and inverse of limit of resolution is resolving power or inverse of angular limit of resolution is also resolving power.

Now, depending whether we are talking in terms of angle or length, the two expression varies they both are equally valid $1/\Delta\varphi_{min}$ and $1/\Delta l_{min}$.

(Refer Slide Time: 39:56)





Now, there are few takeaways from this lecture. Now, if the smallest resolvable separation between images is to be reduced, the wavelength might be made smaller, because from this expression equation 32 and 33 you see, this is the minimum resolvable angular limit and Δl_{min} is the limit of resolution in linear domain. In both the equations λ is there in the numerator and D is in the denominator.

Now, if you want to increase the angular limit of resolution or limit of resolution, then you will have to either increase λ or decreased D, this is what is being said here in this point exactly.

If the smallest resolvable separation between images is to be reduced, the wavelength might be made smaller, if we use a smaller wavelength, then we can separate out the closer sources. And therefore, using ultraviolet rather than visible light in microscopy, allows further perception of finer detail. If two images are very close to each other, go for the lower wavelength, if the wavelength lower, limit of resolution is also lower, and then we can look into the finer details, which is not visible in the larger wavelength limit.

Now, on the other hand, the resolving power of a telescope can be increased by increasing the diameter of the objective lens or mirror, because diameter is sitting in the denominator. Now, as an example, to have a perception of these parameters, let us talk about our eyes here, the diameter of human eye people is about 2 millimeter under bright conditions.

And suppose, we are using green light which is a λ is equal to 550 nanometer, $\Delta\varphi_{min}$ turns out to be roughly 1 minute of an arc. And with the focal length of about 20 millimeter, here we are assuming that this is the focal length of our eye lenses, Δl_{min} on the retina is 6700 nanometer and this is twice the mean spacing between the receptor, this is how the eye images, this is the resolving power of the eye is how it is calculated.

Now with this I end this lecture, and thank you for your patience, see you all in the next lecture.