PROFESSOR: Right, OK. So I guess we ought to start. We seem to be a bit depleted. Anyway, so today last time we spoke about 4f imaging systems and a bit about spatial filtering. And I guess we're going to do quite a lot more on that today, really. So I mean this is really the most basic part of Fourier optics, really, I think so you need to get lots of practice seeing how it all works until it comes second nature really.

So this is what we said last time. We went through this slide last time showing the significance of the point spread function of a low pass filter. The idea is that if you put in a point object-- i.e. You put in an opaque screen with a small hole in, illuminate that with light, this is going to then give a spherical wave. And then this lens is going to collimate that spherical wave to make a plane wave.

And in this back focal plane of that lens, you will get a nice smooth variation here-- no variation, uniform intensity, uniform amplitude, which is basically the Fourier transform of this delta function. Fourier transform of a delta function is a constant, isn't it? And then this light then goes through a pupil mask. The plane wave after the mask then gets to this second lens. And the plane wave is then focused down and produces a focused spot at some image plane, right?

So that's how it all works. And so now what do you actually see here? What's the image of a single point object? What you see is the Fourier transform of this pupil mask. Because again, this is also producing a Fourier transform. Right, so the point spread function-- what's called the amplitude point spread function is just the Fourier transform of this pupil mask. And we sometimes talk about an intensity point spread function, which is actually the model of the square of that, the intensity in that image of a point object.

OK, so and this is reminding you where you might have come across this sort of concept before. This is like-- we spoke in the last lecture about if in electronics, you might think of an electronic system like an amplifier or something as a sort of black box. You put in a signal, and you get out a signal. And you can characterize that black box by what's called the impulse response, which is the response that it would give to a sharp spike going in.

All right, so ideally you'd expect-- you'd want a sharp spike to come out as well. But in practice, you don't get that. You get some sort of broader function because of the imperfections of the transfer system. OK, so this is our point spread function. And it's denoted-- little h of x dash, y dash. This the x dash y dash plane. And I guess this terminology little h seems to be pretty universal. Goodman, I think, uses that in other places.

All right, so this is just now describing again how you how you calculate this 4 pi-- this 4f system. We notice everywhere that the input transparency of a delta function gives one here. You multiply by the pupil mask, and then you do another Fourier transform to find the image. All right, so that's-- I ought to say something bit more about the scaling of this.

Yeah, so you've got of course, this. The output point spread function is a function of x dashed. But, of course, when you do a Fourier transform of something, it's going to be a function of frequency, not of distance, isn't it? So really this h here, h of x dashed, we get by doing the Fourier transform of this, and then transforming the coordinates-- scaling the coordinates in order to get it into the coordinates x dash.

OK, so and then last time, we also-- I'll go through this very quickly, what we said last time, if you've got a low pass filter in 1D, like a slip pupil, then the pupil mask is a rect function. And so if we're trying to calculate the image of a point object again, the field when it gets to this pupil plane is just a constant. And so after it's gone through the aperture, it will have this sort of amplitude variation.

And then so this is what this is showing. On the positive side is what's on the negative side, which is just 1 times this, which is this. And then if you then Fourier transform that, you find that the point spread function is going to be the Fourier transform of this, which is just the sync. And then the point spread function is the Fourier transform of this.

And then this is doing the scaling of the coordinates and getting it into the right hand side. I'm not going to go through that. We went through it last time, but hopefully we're going to-- yeah, here we are. And there's the sync. Right, so this is that the amplitude point spread function or just normally called point spread function. Note that it goes negative. This is 0 here. So note that it goes negative, right. So you've got to convolve that with your objects to find the image of some arbitrary objects, which we'll carry on to talk about in a minute.

This was just another example of a phase filter and the Fourier transform of that. But I don't think there's point going through that again. We'll just carry on now where we've got up too, which was to do with shift invariance of the 4f system. Now what we find-- what we've looked at is if we've got a point object that we get from this hole in the screen, we get an image which is this amplitude point spread function. The amplitude here would be this Fourier transform of this pupil mask.

Then, yeah I'd just mention, there is this constant. This is the point spread function. It's the Fourier transform of this pupil mass. But there's this constant out at the front. And this constant out at the front is written like this. And you can see that it's made up of two parts. First of all, this phase part-- this just comes about because of the distance the wave is traveled through the optical system. Right, so if it starts off with zero phase here, it's going to come up with some particular finite phase when it travels all this distance through the system.

And then the other bet is this F2 over F1, which is a scaling factor, the ratio of the focal lengths. And basically you need that to satisfy conservation of energy. All right, so you can imagine that as the point spread function gets broader, then it has to also get lower because it has to contain the right amount of energy equal to the energy going into the system.

So when you work out the intensity of this image of a point, this would be F2 over F1 squared, which is the ratio of the heights-- the maximum intensities of the center of that image of a point. Right, now so what we'll be carrying on to look at now is what happens if we displace our point object. We move this delta function sideways. And then let's look, first of all, at just the physical situation of what's happening. This produces a spherical wave. And then when it goes through the lens, you can see that it's going to produce again-- it's going to be collimated into a plane wave.

But now this plane wave is not actually moving horizontally but is moving at an angle. And this comes about because when you Fourier transform this, you will get some exponential phase factor, which actually represents a wave which is moving at an angle. And then when this inclined plane wave goes through the pupil mask, of course, its angle doesn't change anymore. So it's still moving at the same angle here. And then again, it's collimated and produces an image.

And the image-- what we find-- is the same as it was before. The amplitude is exactly the same as it was before. But it's shifted. So this is what we found before when we looked at the image. You got this-- you remember-- this inverted magnified image. The inversion means that it's the opposite side of the axis. The scaling is this F2 over F1, which comes about from the fact that the focal lengths are not equal. But the important thing to notice is that apart from this shift, this function is the same as the function was before.

So what that means is that this is what's called shift invariance. It means that the image of a point is independent of where the point is. Right, so that's a very important property, which simplifies things tremendously. Yeah, and here's a note just to say that that is the lateral magnification of this system as we described before. OK, so now this is just showing this diagram again. These are not meant as sort of rays. These are just sort of showing regions where the light is as it goes through.

But we've got this amplitude here, which is Fourier transformed by a lens and then goes through a mask, Fourier transformed again [AUDIO OUT] field. So what this slide is trying to show is the significance of the point spread function. At some arbitrary objects-- we can think of an arbitrary object g of t, an arbitrary transparency. You can you can think of-- sorry. You're illuminating that with some field. So what actually you get after the transparency is the products of those two.

So is the g in that I really meant to talk about. The g in you can think of as being made up of lots of delta functions. Any arbitrary function, you can divide it up into lots of delta functions where the strength of each of them corresponds to the strength of that function at that point. Right, so this is what this expression here is writing down. So the g in is written as-- you can write it as an integral over delta functions with the strength given by the function.

But each of those delta functions is going to propagate through this system and produce a point spread function exactly like we described before. And then the image of the total object is going to be the sum of the images of all these delta functions. So all of these delta functions produce point spread functions. So you integrate over those point spread functions with the relative strengths that you need from here and you'll get the final image.

And yeah, you can see that this is a bit like Huygens principle, that you are breaking up this wavefront into lots of components that are all represented as delta functions. OK, so therefore what we get out is going to be this integral. Each of these delta functions is going to produce an image given by this point spread function. The point spread function is shifted a distance which depends on the original distance but is scaled according to this factor we just mentioned earlier.

And the strength of that component is given by this value g in at this point as we said there. So our final expression says that the output amplitude, but amplitude convolved with the point spread function. Right, so this expression-- you can see-- is exactly the same as an ordinary convolution expression, except [AUDIO OUT] But you can see it's like [AUDIO OUT]

OK, yeah so because of this scaling, we can-- and also the inversion as well, we sometimes express things in terms of a reduced optical output coordinates. So we think of the output coordinate as being [? measured ?] [AUDIO OUT] to get rid of the sine and also scaled it to get rid of the scaling factor. And if we do it in terms of these reduced quantities, then you can write this as are just a normal convolution expression. So all we've done really is scaled the output coordinates in order to make it so that you can use the normal convolution expression.

And then we think back to our case I was mentioning before of the electronic black box. So an electronic black box, you can either characterize in terms of an impulse function or in terms of a transfer function. So an impulse function-- you look at a delta function in time and look at what comes out. In terms of a transfer function, you put in a pure frequency and you look at the output that you get from that pure frequency.

Right, so we can do exactly the same for the optical case, instead of thinking in terms of a point spread function, which is the same as this impulse response, we can think in terms of a transfer function. And now one thing to be very clear here is of course, this whole thing, we're talking about coherent systems. So we're working with amplitudes all the time. And so it's the amplitudes that are going to be affected by the optical system. And that's why we have to think of what's called an amplitude transfer function.

All right, and we've written that as ATF, amplitude transfer function. So this-- if you do have just a Fourier transform of this convolution, you get a product. And so this is saying that the spectrum out-- the spatial frequency content of the image is equal to the spatial frequency content of the input multiplied by some transfer function, which is the Fourier transform of the point spread function.

Right, so this thing here is called our amplitude transfer function. So we have to multiply the spectral content of the object by the transfer function. And then we could do an inverse Fourier transfer and we'd end up now with the image, the amplitude in the image. That there's something. It wrote something, did it? Yeah, OK. Right, so we've already [? reached ?] that.

OK, so let's look at these examples we said again. So these are two pupil masks that we looked at before. The first one is just a low pass filter, lets through the low spatial frequencies but doesn't let through the high spatial frequencies. And this was another-- this was a phase filter we looked at before where we only let through the spatial frequencies within this region. But also within that region, we changed the phase-- the relative phase of the components which are in this region here.

Right, so you can think of this as being the modulus of the filter. This is the phase of the filter. And outside here, of course, it's zero anyway. So [AUDIO OUT] Right and then, this is how we calculate the transfer function from this. This is a distance, right? The pupil is actually a real physical thing. So it's got a physical size. So this is measured in centimeters here. It cuts off at 1.5 centimeters. And in order to convert that into spatial frequency, we have to divide that by lambda f.

And this is putting in some figures. That says that u is equal to x double dash in centimeters times 10 to the minus 3 centimeters to the minus [AUDIO OUT] [? Scaling ?] of our transfer function, and then finally this is what the ATF is. 5 is going to be 1.5 times 10 to the 3 cent-- [AUDIO OUT] which is what this shows here in this diagram. It's identical, just scaled. Of course, to the other example-- or any [AUDIO OUT] [INAUDIBLE] [AUDIO OUT]

AUDIENCE: Professor?

PROFESSOR: [INAUDIBLE] is going to have this modulus of phase components. Yeah?

AUDIENCE: Your sound is breaking up a little bit. And I wanted to ask how is the point spread function related to ATF? Is it just the-- ATF is the Fourier transform of the point spread function.

PROFESSOR: Yeah, they're a Fourier transform pair. Yeah, but you have to be careful with the scalings. OK, so I don't know why the sound. I moved that a bit closer. Is it?

TA: Maybe clip it on the other side.

PROFESSOR: I'll clip it on the other side, perhaps. Right, OK. Let's see if that's better. Maybe just walking around. OK, so that's that one finished. Where am I going to find something, George?

TA: [INAUDIBLE] That's proper.

PROFESSOR: OK, so now we're up to today. OK, so really today we're going to carry on talking about the same sort of stuff. I think this is the simplest way to describe it. And so what we've been looking at-- this is just revising again with the fact that we're thinking of the 4f system as being a linear shifting variance system. Right, so if you remember the shift invariance means that the point spread function is independent of position? Right, so there in that case, you can [? right ?] the image as a convolution. Yeah?

AUDIENCE: Mr. [INAUDIBLE] There's [? nothing ?] with [INAUDIBLE]

PROFESSOR: OK. So I hope that this is better. So we can take about a 4f system as being an linear shift invariant imaging system. And so this is going through then how you calculate that sort of system. So we start off with some illuminating field that comes in here. It illuminates some transparency. And so what you get on the far side of that is going to be the product of illumination and the transparency. All right, so this is basically the object of the imaging system,

And then it goes through an imaging system with a certain pupil mask, which is related to an amplitude transfer function, which is just a scaled version of that pupil mask. And by doing the Fourier transform of that, you will get the point spread function of the system. So the Fourier transform of the pupil mask gives this point spread function like this.

And then your output field is going to be your input field convolved with the point spread function of the system. So I noticed that-- so there are two versions of this that are described here, one in terms of x dashed, and one in terms of x naught dashed. This x naught dashed one is the rescaled one to make it look like a proper convolution. So we've got rid of the scaling factor F2 over F1. And you see what that does back here in the point spread function is it just changes an F2-- well, one over F2 to one over F1.

So you can work in either of those. There's no difference really. OK, so let's now think about how these concepts fit in with what we learn from geometrical optics. And first of all, then we look at natural magnification. What we find then is that if we have a point object here-- we just mentioned this before-- what it would give is an image which is going to be a blurred image of a point object. It's going to be this amplitude point spread function.

But let's for the moment-- we're trying to find out how this fits in with the geometric optics limit. Of course, in geometrical optics there is no diffraction. So you can look at the sort of limiting case as when that point spread function becomes a delta function. So in this case, this is not physical because you couldn't do this either very easily. But you could do it only by making the wavelength 10 to zero or you could make the aperture 10 to infinity, which of course, both things you can't really do.

But just think of it like this just to see what's going on. If you could think of the point spread function as being a delta function, the image of our point here would be a point here. And we've already found where that point is. We've found that it's a point which is on the opposite side of the axis. And the distance here is scaled relative to this one by the ratio of the focal lengths. And this is exactly what we found from geometrical optics.

So the lateral magnification is this n here. nt is the lateral magnification which is given by minus the ratio of the focal length in F2 over F1. OK, so if we assume that our amplitude point spread function is just a delta function, then the output amplitude is going to be equal to the input amplitude but with this scaling. OK, so that's the result we get. It's exactly the same as you get from geometrical optics. Analyzing the same system would also give you this same magnification.

Right, so this is another example. This is looking at the angular magnification of this system. So before we looked at what happens for a point in the input plane. Now we're going to look at what happens to a plane wave at a certain angle in the input plane. And you can see what that plane wave does as it goes through. Of course, the plane wave when it's Fourier transformed now becomes a converging wave, which converges on this point. It then spreads again and then is coming is collimated by this lens to produce an another parallel wave there.

But you'll notice that the angle of this wave is not the same as the angle of this wave. For a start, you'll notice they've got opposite signs. That one's going up. That one's coming down. But that's what you'd expect of course, because of this same inversion that we've been talking about. So this input field then is just an exponential face factor.

OK, so still assuming this ideal geometrical point spread function, the output field is going to be given by this. And you can then express that in terms of phases. And it comes out to be this. So this is the angle-- the angle in was 3 to 1 and the angle out is now minus F1 over F2 times 3 to 1. So again, it's scaled and got an inversion relative to the input angle.

So you can see the angle of magnification then is given by this expression, minus F1 over F2. And again, that is the same as you get from your geometrical optics. OK, so based on the interpretation of propagation of spatial frequency, the magnification results are also in agreement with the scaling similarity theorem of Fourier transforms.

Right, so you've got an input amplitude, an output amplitude related to an input amplitude, an output spatial frequency related to an input spatial frequency. And you notice this one's got an F1 over F2. And this one's got an F2 over F1. So this is back to where idea is that if you make the functions smaller, the Fourier transform is broader and vise versa.

OK, so now this geometry I guess you could see-- maybe I could just draw one. So this is our system. And what I was going to draw, this is F1. F1, F1, F2, F2. And you can see this wave is coming in like this. So we could have drawn a ray like this. And this would of course, have been refracted so that it's traveling parallel to the axis. And then this would come down like that.

So you can see here, this is our phase one. So f phase one is this distance. And that is also-- if you think of this as being theta out, we can see that this distance is the same as this distance. And so F2 times theta out is equal to F1 times theta in. OK, so everything agrees nicely with geometrical optics. So that's good.

So let's look at some examples now. Let's look at some examples of some apertures. So first of all, the case of let's put it as our pupil mask a rectangular aperture, which has got a distance a and b in the two directions. So we can think of this as being rect, x over a times directive rect of y over b. And we can then work out what the transfer function of that is. The transfer function-- we said-- is just equal to a scaled version of the pupil mask. So if this is a rectangle, this is also a rectangle.

But the scaling has got this lambda F1 coming in. Right, so this is now-- because it has to have something. This is a function of distance. Sorry, x is distance. a is distance. Right, so this is the dimension and its quantity. And in here, this is distance. This is distance. And this is distance. So this has to be, of course, inverse distance, doesn't it-- to make this whole thing dimensionless. So this is why you need to have something in there to get it the right dimensions.

And then the point spread function is given by the Fourier transform of that. And that's what this looks like. Fourier transform of rect is sinc. So this is separable. So you just do the two Fourier transforms separately. And you end up with this pattern that you've seen before. It looks like a central spot and then a lot of extra spots that decrease in amplitude or intensity as you get away from that. Yeah?

TA: Maybe we mentioned that in the notes that they posted the ATF and PSF were mislabeled.

PROFESSOR: Yeah, OK. In the notes that were on the web, this said PSF and this said ATF. So it's pretty obvious when you see them. It's obviously the wrong way around.

TA: And we posted corrections.

PROFESSOR: Yeah, OK. [? George ?] posted a correction. Now I was going to ask one question. This is presumably the intensity of the point spread function that your plotting here--

TA: Is the square.

PROFESSOR: Is the square. Yeah, the square of the amplitude point spread function. Right, So that's our first example, very important example. The second one is probably even more important. So this is a circular aperture. And of course, most optical systems do have circular apertures. So this is why this is of great importance. It seems to have turned out being bit like American football for some reason, but never mind. And practice because it's going to over to the US, I suppose.

TA: This would have been OK except the perception [INAUDIBLE] the other way around.

PROFESSOR: Oh yeah. That's true. Yes. It doesn't know how to do Fourier transforms obviously. Anyway, this is supposed to be a circle. And if this is the pupil mask, the transfer function-- the amplitude transfer function is just a scaled version of this. So it's again just a circular thing. And sorry, perhaps I just ought to stress this. It's obvious to you now, I'm sure, by this stage. So this is plotted here as black or white. Of course, white means one-- black means zero in this case.

So that means if spatial frequencies land on this, those inside this get through and those outside this don't get through, right? So this is what we have to multiply our frequency response of our object by in order to get the frequency response of the image. And that's just like before. We have to scale this with these lambda F1s. And then the Fourier transform of this will give the point spread function.

And this one has now gone into glorious technicolor. And so this is what we call the airy disc. The equation for best is given in terms of Bessel functions. You've defined this at some point, I presume. I'm not sure I was in the lecture where we defined this [INAUDIBLE] function, which is basically 2j 1 of something over something. And so this is our normal airy disc, the image of a point object, which has got this central bright spot, and then a dark ring, and then some bright rings outside a not so bright ring, and then another dark ring, and so on and so on.

All right, so what we find then is the size of this has got this magic factor 1.22 that comes up all the time in terms of when you look at Rayleigh resolution and things, you always get this magic factor, which is actually 3.83 which is the first zero of J1 divided by pi it turns out. So that's why you get this magic figure. OK, and this is just to remind you of course, when it comes to resolution, we've mentioned before how resolution is related to numerical aperture.

And so what we're doing here is producing an image of an object, aren't we? So if we've got an object, it's imaged with this lens, which has got a certain numerical aperture. Right, so this aperture here is what decides the resolution of the image that you will get in this system. This aperture here, you see is a ray that was drawn so it goes to the edge of this pupil mask.

And if you trace this strobe, you'll see that it comes back down here. Sorry, of course, the numerical aperture is actually-- the sine of this angle, if you remember. That's what this means. It's not the angle itself. NA is not the angle. NA is equal to N sine alpha, if you remember. The refractive index of the material times the sine of this angle. But what we find is that if F1 and F2 are not equal, this angle here, of course, is not going to be equal to this one.

You can see that if NA is small, if NA is small, you can see that these are going to be scaled in the same way as your F1 and your F2. It [INAUDIBLE] exactly be that. Of course, this NA would have to be the tan of that. But the sine-- if NA is small, then this distance is the same here and here. So this distance over this distance divided by this distance over this distance is pretty close to being the ratio of the focal length.

OK, so we've got a certain NA on the input side. And we've got a certain NA on the output side. And the ratio between those names is just equal to the magnification of the system. So actually, of course, we're well used to the fact that rays of light go equally well in both directions. So we could use this same ray diagram to look at what happens if this was the object and this was the image.

But in this case, of course, all these things are now inverted. So if it was producing-- as it stands here, it's producing a demagnified image in this way. So it would produce a magnified image going this way. And so you see that the object that you'd have to be imaging is scaled relative to this one. So the resolution would be the same basically in two directions. I mean, sorry-- that's not quite true. What am I trying to say?

I'm trying to say that the NA relative to the size of our object that we're imaging here will be the same when we think of it as going this way or that way, because in real terms, how small a thing you can actually see with an optical system depends on the numerical aperture of the lens that you're looking at with, not the lens at the far end.

OK, and then for our slit aperture, this is what the point spread function looks like. And then you can express that then very simply just replacing our a over F1 by numerical aperture, we've got now that the width at this spot is given by lambda over NA, which is exactly what we came about-- I guess you said that before somewhere.

TA: I can't remember.

PROFESSOR: But the resolution of a microscope or whatever and of an imaging system, you can write in this form. So of course, this distance will be smallest when this is biggest. So the biggest this could be is if NA was one and you made n, the refractive index nice and big. And this one, these distances are in terms-- this is distance in the image in terms of the NA in the output plane. So you could have converted this so this is distance in the objects. And then this would become NA in the input plane.

And right and then circular aperture, our point spread function looks like this. And so this is what we get. And this delta r dashed here given by this expression, this 0.61 is half of the 1.22 we had before. But notice this is delta r dashed is the radius of the PSF [INAUDIBLE] So it's this distance, not the whole distance. It's this distance, right? So this is why we got half-- why the 1.22 is divided by 2.

OK. Now I didn't mislead you also by saying from now to the end of the lecture's a breeze because it isn't, because the next bit is this thing about sampling. So that one more bit which is a bit difficult to take in. And this is really pointing out about effectively the number of degrees of freedom in an image. So imagine first of all, you've got an object that you're imaging. So this is this input field.

It's got a certain size which is called here delta x. And of course, in any real practical systems, you don't actually measure it as a continuous function. It's actually going to be measured with a certain pixel size. So here we're calling the pixel size delta x. Right, now so what we've got here then is this function is like an array of spikes in this direction. An array of spikes delta functions in this direction, isn't it-- so in the products of those.

And if you do the Fourier transform of that, you'll get something which is very similar again an array of spikes. If you look in the standard books on Fourier transforms like Bracewell or whatever, he calls this a cone function. And the Fourier transform of a cone function is another cone function. Right, so this is a cone array of delta functions multiplied by a rectangle. And the Fourier transform of a product is the convolution of the Fourier transforms.

Right, so the Fourier transform of the rect is actually-- so this is really-- I ought to get into that. So just think of this then as being in the spatial frequency domain. It's got a certain range of spatial frequencies, and it's made up of certain resolution of frequencies. So what we're going to say is something about Nyquist now. So we can say that-- we know that in order to image this input field, knowing what this bandwidth is here, we have to sample with this distance. Is that what I'm trying to say?

Yeah. The bandwidth is equal to 1 over twice the sampling distance. That's right. Yeah, and this is in the x direction. This is in the y direction. But we can also-- you can see by the similarity of these diagrams and the fact that Fourier transforms go either way, you can see that there will be a similar sort of Nyquist type criterion going from the other way around. So because this has got a finite field, this would mean that you've got a certain frequency resolution there. So there's a relationship between delta u and delta x.

The fact that this has got a 2 here, of course, you know from the normal definition of Nyquist. And the fact that this one doesn't have a 2 in, it just comes about from the fact that this is defined here as being two times this. And right, so just like before, the spatial frequency domain-- those spatial frequencies are then multiplied by the pupil mask to find the spatial frequencies in the image, aren't they?

Right, so this is looking now in the plane of the pupil mask. So this is actually the same as equal to this but just with a rescaling. And so these are the expressions for that. So this is just applying that rescaling, putting these lambda F1s in these expressions. So you'll get then what you need, how you need to sample in this plane here.

And what you can notice from that is if you you'll notice that these things are all equal and equal to some constant, therefore. And this thing is called the Space bandwidth product. So you look at what this means. Delta x over delta x is the field size divided by the pixel size. So it's basically the number of pixels in the field, isn't it? Right, so your original object here has got the number of pixels equal to n x times n y.

So that's the total amount of the number of pixels. And therefore if you think of it as being a binary thing, it's the amount of information that you've got in that original object. And of course, that information is going to be transferred through the system. So if you look at now the ratio in the next plane in terms of use, this distance over this distance is equal to again, the number of pixels. And as you know, when you do a fast Fourier transform of something which has got n elements, you get something which is n elements. So it all works right.

So effectively, you can think of this as being a bit like-- of course, you can't really-- if this has actually got some finite bandwidth, then of course you can't sample on points. So it's a bit of an idealization to thinking this for the terms really. But nevertheless, it gives you a good idea. So of course, for a repetitive type function like this, also you can think in terms of a Fourier series becomes a Fourier transfer.

Right, so all we're doing here is really resolving our function here into a Fourier series. And the Fourier series consists of a lot of sine terms, doesn't it? Sine of n, and sine of x, and sine of 2x, and so on and so on, and these correspond in the Fourier domain to these points on this grid.

Right, so these each of these points represents the particular Fourier series component in the original object if you think of it like that. So the idea of a space bandwidth product is quite a powerful way of thinking about the fundamental limits of an optical system. And people have gone on to study this sort of approach in much more detail, actually trying to get around the limitations that I just described.

The fact that, of course, that we've truncated this means that this is no longer a repetitive function. And similarly for here, right? So that actually you know this idea of thinking of truncated repetitive functions is a bit of a anomaly, really. But people have come up with mathematical ways of dealing with these sorts of systems which are probably way beyond what we are going to get to in this course.

OK, did something come up then? Oh, yeah. Next slide. So OK, this is just pointing out then we said that you put a mask in here and that mask produces a certain transfer function-- amplitude transfer function, and the Fourier transform of that gives the pupil function. And we've shown a couple of examples of what this pupil function looks like and what the image would look like.

But this is just pointing out that actually, of course, that means that we can choose whatever we want to put in there in order to produce the results we want. So this is what George has called here pupil engineering, which Ernst Stelzer doesn't believe in, I think, from what we were saying. But anyway, so the idea is that by changing this mask, you can improve the properties of that point spread function in some way, which would suit the that particular application you're looking at.

Actually, it turns out that you can't actually do a lot better normally than the very simple example of a circular aperture. And the circular aperture is actually very special because it does give according to some measures anyway, a very sharp point spread function. And you can come up with other pupil functions that make the central lobe narrower, but normally what happens is then the size of the side lobes get stronger.

So it's a sort of compromise. You can't have a very small central lobe together with very weak side lobes. OK, so some examples then. So let's imagine first of all, our object is our intensity at the input plane is this rectangle with the stripy pattern. And so we're going to then look at what happens if we image that with some mask that we put in here. And what should we put in there, then?

Well, the first one let's put in something like this. So now this is going to let through-- this is a low pass filter, isn't it? Right, so it's a very small low pass filter as well. So it only lets through the very slowly varying parts of that object. But it stops all the fast variations of the object. Right, so if you look at the object, you'll see that it's got this stripy pattern, which is actually a very fine structure. So these are actually really being calculated using Mat lab, not by me, by George.

And so this is what you'd see in the output plane of that. So there you are. You can see that this stripy pattern has disappeared virtually. We've lost some sharpness at the edge though because we've also lost the high spatial frequencies, of course, that defined the edge. And a sharp cut off at the edge will include quite a lot of high spatial frequencies. But anyway, so that's a good example of spatial filtering. And then the next ones we're going to look at something which has got a bit more structure.

And so this is MIT written as a binary mask. And so this is our original object. And we're going to image that in an optical system and look at what the image looks like in different cases. So the first one is quite straightforward. We've got now a circular aperture. So this is what the image looks like in this case. You see that there's already a bit of artifacts in that image. It's no longer-- you can still read it as MIT, but you see that inside the letters, you'll see there's some structure there, which is not really present.

So this is some sort of edge ringing, I guess from the fact that this aperture has got this sharp cut off at some particular spatial frequency. So if you made it so that it tailed off gradually, you could probably get rid of some of that-- those fringes there. But OK, so that's an aperture. So we've got rid of some high spatial frequencies. Let's look what happens if we reduce the spatial frequencies even more. So this is making the aperture smaller.

And now you can see that the resolution has degraded. It looks rather blurry now. You can you can't really see whether it says MIT of HIT or maybe HIY. It's all getting a bit vague, isn't it? And also you can see that it gets a bit sort of blurry around the edges. Yeah, sorry if we make that--

AUDIENCE: Question? One thing to point out here which is kind of obvious, but it's worth pointing is that if you do this also, the average brightness of the image would decrease as you stop down the aperture. But in this case, they cheated. They normalized the intensity to the full level so it is visible in the projector. But in a real system, it would also become less bright.

PROFESSOR: Right, OK so I didn't notice that because I didn't do the calculations. But yeah, OK. So of course, if this aperture gets smaller and smaller, the amount of light going through gets smaller or smaller. And so in the end, you'd end up with black. Right, so anyway we're now going to make this even smaller. So what would you expect?

Well, you'd expect that again the thing to become more blurry looking. And in fact, it really does become. Just before we carry on, just take this fully in. You notice that there are four legs here, four straight lines. And these two are closer together than these other spacings, right? So as we decrease the size of the pupil, now you see what's happened. This is quite amazing. I was quite surprised at this. The four legs have become three legs.

So I guess these two central legs are now not resolved. They're below the resolution limit. And you'd have no idea of course what that said now. It looks something really blurred. So we're not getting enough information obviously there to be able to see the image properly. OK, then let's try doing some other funny things. So this one is a high pass filter. Now, so what we're doing here is we're blocking out this central part and letting through all the other filters, all the other frequencies.

Now there's one other thing that we ought to mention here is that this, of course, is done-- these calculations are done using FFTs, right? So we're actually working on a field of a certain size when we calculate the FFTs. The fact that this-- as you know FFTs do like a sort of repeating, don't they? Right, so this is not really quite the same as this being going out to infinity in spatial frequency. So I think that the result in the image you might get could depend a little on the properties of the zero padding and things like that in this case.

But anyway, this is the image in this case. So you'll notice that we've got the high spatial frequencies. So we actually get some sort of enhancement of the edges. You see the edges come out really very crisply now. So you get this enhancement of the edges. It's like an edge enhancement filter. But on top of that, there's a rather strange sort of glowing patterns, which I'm not really quite sure what they refer to at all. But you could-- I guess-- think of them as being some sort of artifact. It looks a lot better on this screen than that one. Interesting.

But if we do another of these high pass filtering but with the central block even bigger, this effect of edge enhancement shows really nicely now. So all we're now seeing is the edges. And the bits inside the letters are not showing up or outside the letters, so neither. I can't remember what the object looked lie now. It's interesting. The object, whether it was black writing on a white background or white writing on the black background, probably this would look very similar. It wouldn't make a lot of difference-- I don't think.

OK, and so all these examples have been circularly symmetric type patterns. But you can also, of course, put in things which are not certainly symmetric. So this is throwing out these frequencies here and here. So we haven't done anything here. So we're going to have a high resolution in that direction. But in this direction here, we've reduced the resolution. So we'd expect it to be blurred in the vertical direction.

And of course, just to prove that we can do the opposite, and this one now is going to be blurred in the horizontal direction. So you see here again, you're almost losing the resolution of those four legs. OK, and then probably one of the most important applications of pupil filters is in phase imaging. And so this leads on to a Zernike face contrast, which Zernike got the Nobel Prize for.

It's quite interesting. It was a long time after he invented it though, I believe. He invented it sometime in the '30s. And it was commercialized-- actually he seems to have managed to get Zeiss and [? Leicher ?] and everyone to make it. So he must have sold his patent to lots of people.

But anyway, eventually it became very popular. But it didn't get it didn't get the Nobel Prize until the '50s some time. And very useful for a long time, I think eventually it got to a certain degree anyway, outclassed by some later techniques, but still used a fair bit in the lab. He has one advantage of being quite cheap as well.

Anyway, what you're trying to do with [? phase-phase ?] contrast is image of phase objects. And this is an example of what a phase object might look like. Imagine, for example, you've got some glass slide-- well, one thing which would approximate quite well to this is if you had a biological cell sitting on top of a glass slide. Right, so this shows that something like that. But actually what we're thinking-- well, if it was a cell, of course, the refractive index of this would be slightly different from this. But that's really not an important difference.

So this is looking on the top of it. So here we've got this object, which has got a certain thickness. And what we'd like to be able to do is to produce an image which actually shows up the shape of this object, right, which might be a cell or something. And so this phase object-- we can think of the light propagating through this and producing a phase shift then. So the phase shift is given by this expression. So n here is the refractive index of this cell.

We have to have n minus 1, of course, because the cell is displacing air. So the cell wasn't there, there would be air there. So we have to subtract away the air that would have been there. And so the phase is 2 pi times n minus 1 times the thickness, which is a function of x or of x of y in this case, divided by the wavelength. And so if we illuminate this with some certain illumination field this, then what we'll get on the output side is just this times this.

There's no-- each of the i phi, our object just behaves here as a pure phase object. There's no modulus change, only a phase change. Yeah, so useful for imaging biological objects such as cells, but not only that. It's got applications in material science, and so on. And yeah, so if we looked at the-- image this illumination is just unity. And we look at the modulus squared of this object function. You see, because it's just only a phase object, when you look at the modulus square, you'll see nothing at all.

So if you look at this in an ordinary microscope, you wouldn't see anything. You'd just see completely uniformly white. And so you wouldn't see the shape of this thing that we're trying to see. So the question is, how can we go about that? What can we do to solve that? So yeah, what we're now going to look at is back to this MIT pattern. And we're going to think of an MIT pattern, which is now not a binary object but a phase object. So imagine that this is our object.

So this represents phase. So red represents a phase change of 0.1 of a radian. And this one shows the modulus of this object, which we're assuming is completely uniform. And so it's a pure phase object. And if we looked at this in our ordinary optical imaging system, all we'd see is the modular square of this thing. So we'd see basically nothing. So the question is what can we do to try and improve that.

Well, Zernike came up with the answer. And actually, this is not quite the way he does it. We're not going to really do the way he does it. But the principle is here anyway. What you do is you introduce a pupil mask, which has also got a phase objects in it. It's got basically a small region at the center, i.e. where the low spatial frequencies are, which we change the relative phase of those spectral components relative to the other spatial frequencies by pi over 2.

And so we put that in our in our system. And what we find, as it says there, is we can actually get contrast of the phase. We actually see the phase now. And this will explain why that works. So this is what we're doing. We're putting in our phase MIT thing here. And we're putting in this pi over 2 phase mask here. So that's what it looks like. It's some bit of dielectric, which has got this central region which changes the phase by pi over 2.

Ideally, what we want to do, as it says here, is for this to be very small. What we're trying to do is to use it to change the phase of just the DC component, the unscattered light component relative to all the rest. But of course, you can never do that because it's got to have some definite real size. So whatever size it is, there might be some low spatial frequencies of the object, which are going to be-- the phase is going to be changing in the same way as the DC term.

So let's look at the mass of how that works. So what we do is we assume our phase object is what we call a weak phase object. It's got the form e to the i phi. But we think of that as being approximately one plus i phi. And we neglect all the other higher order terms in the Haus's series expansion for the exponential. And you can see then that in an ordinary system, what we'd see is you can think of it-- keeping all the terms, the modular square of this is obviously one. And this also, the modular square of this to the first order of small quantities is also going to be one.

Of course, if you square this, you'll get phi squared terms. But actually, there are also really some extra phi squared terms in this series, which actually, if you do it properly, all cancel out. So you still must get one.

OK, so what we're now going to say, this is our object. But now, this object is going to go through this mask here, which is going to change the relative phase of this relative to this. And so if we change the one to i, we're not going to get i plus i phi. And therefore now, the output that we get is going to be given by-- this should be into i, intensity actually.

Yeah, the intensity that you see is the modular square of this, which is 1 plus phi squared, which to small-- Because phi is small, we can expand that to first two terms of the series. And we get one plus two phi.

Right, so now we're getting an image where the phase is visible to the order of this-- order of the weak phase object phi. In this case here, we weren't seeing anything to the order of phi. The lowest things that were there would be of the order of phi squared, which is going to be negligible if phi is going to be small enough.

Right, so there we are. So by putting this in, we've managed to make this phase object visible. So this is the demonstration of that. So this is our looking out at our phase MIT. And this is the size of our phase mask. So it's quite big here actually. And so I guess, it's changing the phase of the central DC term. But it's also changing the phase of some low spatial frequency terms of the MIT, which I guess accounts for some sort of imperfections in this image.

And I think a lot of the ripples-- again, are coming from the fact that this has got a sharp edge to it. Probably if you smoothed this, you could make it so that you didn't get such a strong ringing in this thing here. But this is the next example. We've now made this a lot smaller. So it's now changing the phase of the DC term. But it seems that the object itself probably has got very, very weak components in that region. And you end up then with this really wonderful image.

So I guess that really shows the power of the Zernike phase contrast method. That's the end. How are we going? Good, we're up to time. I've been fast. Right, any questions anywhere? Anyone want to look at these nice pictures again?

AUDIENCE: We had a question actually.

PROFESSOR: Yes good.

AUDIENCE: It seemed like the last analysis you're doing, phi was much, much less than 1. But then you're saying that phi was pi over 2.

PROFESSOR: No, no. Two different phi's. How do I get back in this?

TA: You click the play button.

PROFESSOR: The play button. Ah, the play button. It's right there. Yeah, OK. So two different phi's. The object is the one that has to be weak. So here we said the object has got this phi. So we got e to the i phi as the object. And this phi has to be small. And somewhere it said-- maybe it was after. The phase MIT objects-- here we are. The phase MIT objects-- the phrase change is 0.1 of a radian.

So 0.1 of a radian we're taking as being small. But so red here represents 0.1 of a radian. Red there doesn't represent 0.1 of a radian. So this is our mask. So this has got a phase change of pi by 2. So this is much bigger than 0.1 of a radian. OK? Yeah, so that's-- it's the object that has to be a weak phase object, not the mask.

TA: I think the principle still works for strong phase objects. But the explanation is easier if you assume a weak phase object.

PROFESSOR: Yeah. I think I agree with that. Yeah, but actually for a strong phase object, you do also get an image in your ordinary imaging system. But it's very complicated to actually relate it to the object. It's a bit like frequency modulation.

TA: Or another example of a strong phrase object is the waves in a swimming pool. They create an image at the bottom which looks like very bright lines. In a way these lines are an image of the index of reflection of the water. And these lines are known as caustics. But it is very difficult to interpret them and relate them to the actual shape of the water.

PROFESSOR: Yeah. OK, any more questions? Not many from this end, today, either. I usually rely on [INAUDIBLE] to come up with some questions. You know, at FOM, he was in FOM, and he asked the very first question of the whole conference.

TA: Oh, that is quite the distinction. To the plenary speaker, I imagine?

PROFESSOR: Yeah, to the plenary speaker. All very clear? I think these demonstrations are really great because I think that you get a very nice feel for what happens.

TA: We have the quiz coming up. But we have enough time. Maybe we can introduce temporal coherence at the beginning of the next lecture or should we just stop here?

PROFESSOR: Up to you.

TA: If anybody has no questions at all, we have quiz coming up. This is your chance to ask questions, I suppose.

PROFESSOR: So when's the quiz?

TA: On Monday.

PROFESSOR: Oh, right. OK.

TA: That's why he could not answer the question. He was expecting a phone call.

AUDIENCE: When you have a phase grading, normally you try to solve for a system that has the 4f system, you normally play different tricks. One of them is you try to expand part of the grading in Fourier series. And after that, you somehow managed to decompose that into something easy, or you actually change the e to the exponent to something easier to handle in terms of you look at the phase-- and you look at it. And if it is like pi over two or it's pi, you change it to something different that it's easier to handle, any advice on how to solve those? I found those particularly tricky to actually get a grasp of how to solve things.

PROFESSOR: I'm not quite sure you're saying. I mean if you've got-- I mean, you start off by saying you've got some sort of gradient object. And then if you illuminate this, you'll get a series of diffraction orders. And of course, you know, the relative strengths of these might be complex to account for the phase changes. And so what you're saying is that by changing the phrase or whatever of these components, you can improve the image. But is that what you're saying? Something like that?

AUDIENCE: Yeah, my question was more in the order of like how to handle the math? Normally if you get something in terms of Fourier series, you get a bunch of terms. And it's kind of like it ends up being pretty messy.

PROFESSOR: Yeah. But you know, this is like going through the mask. So if we've got some mask which has got some phase variation. So the relative strength of these gradient orders, it will just be changed by the phase of the phase mask. So the forward problem-- I've got the feeling that maybe you were trying to solve the inverse problem, you know. If you know what the image looks like, how do you know how to improve the image or something like that?

But normally, of course, if you knew what the object was, then it probably wouldn't be a lot of point in imaging it. But anyway, to calculate this, this is all you have to do. Just multiply the strength of these gradient components by the strength of your mask, both in amplitude and phase. So it's very straightforward really.

TA: And I don't think there's a way out of the Fourier series expansion. Like if you really want to solve this kind of a problem with a phase gradient, you must compute the Fourier series expansion of the gradient. Unless, for example, if the phase is weak, then you can simplify it as we just saw in the example. But otherwise, yeah, you have to go through the pain of the Fourier series expansion.

OK, thanks.

PROFESSOR: Yeah. But you know, there are some times of course, for lots of other objects, you can just look them up in the tables of Fourier transforms. And even the Fourier series, you can sometimes get from that, of course. I mean, for example, you know, these square wave gradients like this. You can think of this as being this convolved with this.

So the Fourier transform of this is going to be the Fourier transform of this, which is a sinc times the Fourier transform of this, which is another one of these. Right, so you end up with basically an envelope, which is this and then some components within here. And the height of these gives you the strengths of these different gradient components. So I think there was a sketch like that in one of the earlier slides showing how you can think of that.

So these are the strengths of the Fourier series for this object. They're just given by the product of a sinc and an appropriately scaled sampling function.

TA: I think you need to push the paper a little bit up.

PROFESSOR: Sorry. Yeah. Can I do that? Oh, yeah. Sorry. There you can see what I'm saying. Yeah, so some sometimes you can find shortcuts to be able to deal with some of these various areas without having to go through that the hard work of doing what you did when you were a first year student.

TA: There was a homework like that actually where it was not quite phrased like this-- but yeah, I had you guys compute the Fourier series in this clever way.

PROFESSOR: OK no more questions. Yes, [INAUDIBLE] got one. Yeah?

AUDIENCE: How did the idea of numerical aperture come about? And why is it defined as a sine of an angle?

PROFESSOR: Rather the tangent of an angle. Yeah, good point. Well, I guess it's because George in one of the earlier lectures pointed out that according to Snell's law, you've got n sine alpha is the invariant. So as rays goes through a series of interfaces, then that will be an invariant. It won't change. So I think that was the justification for it. But yeah, you've hit something which is quite-- you know, the fact that sine theta isn't tan theta is basically why paraxial optics eventually breaks down.

And so a lot of the stuff that we've presented has been based on assuming that these angles are all small. And a lot of these things just fail if you can't assume that these angles are small.

TA: For example, aberrations are another type of not shift invariant behavior with the exception of spherical. But the aberrations are not shift invariant so our approximations fail. For example, astigmatism if you recall, astigmatism occurs if you have a plane wave entering the lens-- I'm sorry, I should've said coma. Coma occurs if you have a plane wave entering the lens at an angle. So of course, as you increase the angle, the coma becomes worse. So that is shift invariant

PROFESSOR: OK, do you want to stop or do you want to say something?

TA: We could talk about the Talbot effect that we skipped in the previous lecture. I can have to do to find it.

PROFESSOR: I think that might not be interested in that because I don't think my students have ever come across the Talbot effect. [INAUDIBLE] I was just going to say-- [INAUDIBLE] that's all right. I'll come and say it here. I was just going to say that this Talbot effect. He was actually a very long time ago. I can't quite remember the history, but it's a very old effect.

TA: And I think Talbot was also another professor of mechanics who ended up doing contributions in optics if I'm not mistaken, very similar to Maxwell and a bunch of other people. No, no. I went the other way around there. I shouldn't be expected to make the contributions in mechanics. [INAUDIBLE] OK, oh, I'm not showing anything. I'm sorry.

So the Talbot effect is a phenomenon that happens when we have a periodic pattern, for example a gradient and you illuminate it with a plane wave. And then what you do is you observe the intensity pattern forming after the field diffracts, after it has gone through the transparency. So I don't when the movie will show possibly because of something [INAUDIBLE] the projector, it might not show. But if you look carefully, you will see that as the field diffracts, the original grate and pattern becomes visible at certain distances. Then it disappears. Then it becomes visible again.

It will play it in the case of diagonal gradient and then I will play both of them again. So you can see, it diffracts, then the sinusoid reappears and so on. In this case, it is actually very clear when it will stop happening because as we said, in the case of a gradient, in the field after the gradient splits up into diffraction orders. Now this gradient is finite, so as you can see the diffraction orders separate. And by the time it reached here, you can see that you have basically three rectangles that have split up out of the gradient.

So the Talbot effect will stop happening when the rectangles completely separate. It hasn't quite stopped yet. But it can only happen that region of overlap between gradients. OK, so the next-- let me play this once again so you can see it. So again, look out for the repetition of the periodicity of this pattern now. And again for the next one.

I hope with this is showing up also in Boston. Here you can see it very clearly on the projector. OK, so yes? OK, so the next slide is cross sections?

AUDIENCE: George? Did you see that this effect occurs only when there is an overlap between the diffraction orders of the--

TA: Yes. So if the gradient is infinite, then, of course, it occurs forever. But if the gradient is finite, eventually the orders will separate and then it doesn't happen. So this is cross sections of the intensity pattern at different distances, which of course I chose strategically because I know the formula that this effect follows.

Of course, you don't have the formula yet. But anyway, this is what you see. Now what is really interesting and you may not have seen in the previous one, but there's also a plane where you see a periodic pattern at twice the frequency.

So this like the second harmonic of the original gradient. So the question is why does that happen. Well, there is the physical explanation, which is probably the most interesting. And that has to do with the following.

So this really goes to the heart of diffraction, why really diffraction happens and why we see all of this phenomenon. So imagine that you have a plane wave going into a sinusoid amplitude gradient. And of course, in this case you only get three diffraction orders, as we discussed, the plus 1, the minus 1, and the [INAUDIBLE]

And these diffraction orders, each one of them is really a plane wave. So you have now-- you started with one coming in, but after the gradient, you have three plane waves propagating out. And imagine that you pick any point of the gradient really, it doesn't matter. But let's pick this one. And we'll draw a sphere over radius z. we'll draw a sphere centered at this point.

Now if you compare the three plane waves and you take rays really, centered at the center of this sphere, and if you compare what happens to these plane waves by the time that each of the cell of this sphere, then you realize that they all have the same phase delay because they only propagated the same distance from the center to the edge of this sphere. They all propagated by the same amount. So they have the same phase delay.

However, when we observe the field in all of this cross-section that I showed before in my movie and in the calculation, I did not really observe the field on a sphere. I observed it on a plane. So if a center this plane at a distance z from the axis, then the central order that propagates a axis, it sustains itself in phase delay. But you can see that the plus one and minus one order, they propagate the longer distance from this sphere to that gradient.

And you can calculate this distance. It is the difference between the z itself and an expansion that is that you get from its Pythagoras's theorem. So this expression is basically the hypotenuse of this triangle all the way out to the plane. And if you do after that the paraxial approximation, then you find that is given by a quadratic. That is this is our familiar quadratic term that appears in the Fresnel kernel and in all of these expressions of Fresnel diffraction.

So however, now what am I really observing? Well, in this plane, I observe the interference pattern. You can think of it as a Fresnel diffraction pattern from the gradient or you can also think of it as an interference pattern between these three plane waves. It's like an interferometer, right? I mean, physically as I think Professor separately mentioned and I also mentioned at some point or another, diffraction is really no different than interference.

It is just that in the case of diffraction, you have many, many, many waves interfering produced by the Huygen's wavelengths originating at your original object. But in this case, instead of thinking of all these Huygens wavelengths, I'm better off just thinking about it as plane waves interfering because the physics of the gradient tell me that the diffracting field is really composed of three plane waves.

So now there's a phase delay between this guy and this guy. Actually, this guy and this guy, the plus 1 and minus 1 are always in phase. But they can have a difference in phase delay between the themselves and the 0-th order. And of course, if that phase delay happens to be 2 pi, then there's no phase delay. And when did these things start having no phase delay? Well, at the gradient, right?

So if I let this thing propagate enough, whatever parameter I can vary here is z. If I let z grow so that this phase delay becomes equal to 2 pi, then I reproduced my original gradient because of course, there was no phase delay here. So when the phase delay becomes 2 pi, then again I will see the same pattern. It is not as obvious why you get the subharmonic, why you get the second harmonic that I discussed before.

But there is a little bit fortunate that the lectures got shuffled in order because by now, you remember from one of the examples of Professor Sheppard showed, when you actually get rid of the DC component, you can get the second harmonic to pop out all of a sudden. A very similar thing happens if you get the pi phase shift. This pi phase shift can produce a very distinct sub harmonic. So this is that reason.

Now the rest, I will actually not go through. But there's a mathematical-- well, I can go through very quickly, I guess. There's a mathematical way to derive it. And it is also described in Goodman. And I also did it in the slides. So if you start with a grate with a field after the gradient-- of course, you can express it as a diffraction orders. We did this before. You can put it in a form the field after the gradient. We've done this before ask in the context of the 4f system.

But mathematically, it is really the same. Each one of these exponentials in the field will produce a delta function. Now, the question is what happens after the field propagates by a certain distance. So in order to do that, let me show the Fresnel diffraction again. These we're familiar with. We said many times that Fresnel diffraction is expressed as a convolution that has to do with the Huygen's wavelengths propagating and producing spherical waves that again interfere. So the interference is expressed as a convolution integral.

But because of the properties of the Fourier transform, if I take a Fourier transform of the output field, then it will be expressed as a product of the Fourier transform of the input field times the Fourier transform of this Fresnel kernel. And I really am not going to prove this. But you can look up in Goodman's book in the table of Fourier integrals. You can look up the Fourier transform of an expression like this one. And you'll find that it looks very similar. This is a quadratic phase delay

It's a Fourier transform. It also looks like a quadratic phase delay. But it has two key differences. One is that it has a minus sign. So it is actually quadratic phase advance if you wish. And the second is, of course, the [? scalene ?] theorem that will take up this factor of lambda z from the denominator and will pop it up in the numerator. And of course, the happy outcome of this scale operation is that also the units are preserved. This quantity has to be dimensionless.

So indeed, it is. I have distance square in the numerator as well as the denominator. In this case it uses spatial frequency which is inverse distance square. By multiplying by this distance squared, I get again a one dimensional quantity. So the [? scalene ?] theorem works quite nicely. And you can also think then of-- we talked about transfer functions in the context over the 4f system.

But really you can also think of a transfer function in the case of Fresnel propagation because it is shift invariant. Clearly if you have free space, that is it truly a shift invariant system. No matter where you start in free space, you would expect the field to shift by the same amount.

It would be really odd if anything else happened. Unless you believe in space and [INAUDIBLE] I guess. Only close to the Big Bang, this property was not observed. But since billions of years have lapsed since the Big Bang, space is isotropic. So we don't have to worry about it.

So of course, the Fourier integral approaches. And of course, then this quantity that we derive is simply the transfer function of free space. It is another quadratic phase delay. OK, so we can actually apply this formulation now to the mathematics of the Talbot effect.

All I have to do is multiply the Fourier transform of the gradient itself. All I have to do is multiply it with a Fourier transform of the Fresnel propagation kernel that is another quadratic phase delay. And this multiple multiplication is very easy to do because I'm multiplying the function that I know with a bunch of delta functions. So all that will remain really is the value of this quadratic phase delay computed at the spots where the delta functions occurred.

And this is what I'm doing the next slide. This looks a little bit nasty but really that's all you need to know. It looks like an expression very similar to the gradient. This is the DC component. This is the plus one diffracted order. This is a minus one diffracted order. But it picked up this additional quadratic phase delay, which is really it came out of the math now. I used Fourier transforms and convolutions and everything else.

But if you go back to my picture here with the sphere and the phase delays, you can really see that this is really the same expression. It came out of the math in one case. It really came out of a physical picture of phase delay of [? planes ?] as they propagate. But it's all consistent. It is nearly all the same answer. And after a little bit more pain, which involves really taking one Fourier transform, and then combining these things to form a sinusoid, you'll get this expression for the field.

And now you can see what I was saying before, if you-- this looks like a nasty phase delay that got appended to the diffracted field. This is really diffracted field now. So this is the Fresnel diffraction pattern from a gradient. But if this quantity and exponent happens to be 2 pi, then it is not there anymore. So you recover your original gradient.

OK, that's the intensity. And you can also see from this-- you can also get the condition for the second harmonic generation. Basically, if you kill this term, which means that this quantity becomes pi over 2, then this disappears. So all you're left is the cosine squared, which is of course the second harmonic. So this is really delineated in the next slide.

If the propagation distance-- if this quantity, this is also non-dimensional. It is distance squares over distance squared. If you propagate far enough so that this equals an even integer, then you get the replica of the original gradient. This is basically-- this quantity becomes a phase delay of 2 pi. If you get it to become equal to [? n, ?] you call this distance a Talbot play.

If you're get it to propagate so that it is an odd integer, then it is pi that shifts this term by pi. Basically, you flip the sign. Now the gradient has shifted. If you really do work it out, you still get the same sinusoid but shifted by half a period. And finally, if you completely eliminate the linear term by choosing-- this tends to be pi over 2 or 3 pi over 2 or 5 pi over 2, and so on and so forth, then you can actually get-- I forgot to put to close my bracket here, but anyway, then you get the period doubled. And you'll get the second harmonic.

OK, and this the schematic. These are basically now repeating periodically every time you add this propagation distance-- I'm sorry, these are flipped here. This should have been 2 lambda z over lambda squared, 2 lambda z over lambda squared. So basically every time you propagate by this quantity, you go from a Talbot sub plane where you have a second harmonic to a half period shift, to another second harmonic to a replica. And then it repeats. So these are the Talbot plains.

Now we're out of time so I will say something that's interesting for those of you who have time to stay. This-- as Professor Sheppard mentioned has been noticed since a very long time ago. But more recently, it became actually kind of a hot topic over [INAUDIBLE] because people noticed that it happens not only in sinusoidal gradients, but it can also happen in every periodic pattern. If, for example, you have a square gradient, it would also have its own Talbot planes.

And then, not only that, but people noticed that it can happen in some strange patterns that are called was quasi-crystals. So I don't know how many of you have taken solid state physics. But in solid state physics they teach you that there are certain patterns in nature that-- there's often symmetries in nature. For example, a very simple symmetries is rectangular symmetry.

If you take this as a unit cell and then you repeat it, you can keep filling out the space like this periodically. Another example is a hexagon. Let's see if I know how to do the hexagon. I ended up doing an octagon. Sorry about that. That's a pentagon. I need one more, OK. I got my hexagon, OK. That's nice because this does repeat. OK, never mind anyway. OK, that will repeat. Good.

OK, so fortunately I did a pentagon also. So the question is what will happen if you start sticking pentagons next to each other. Now this will be really impossible for me to do, but you can imagine if you tried to stick another pentagon, identical pentagon next to here, then you will get something like this, I guess. And then you can keep doing it. But you noticed when I did it, then I got to stick a pentagon but it is slightly rotated.

It is still is difficult for me to draw but if you google quasi-crystals. I don't know if it's the first website that comes up. There's a website at Caltech, in the physics department at Caltech. And you will see some very nice pictures of what happens if you stick pentagons next to each other. That's called a quasi-crystal.

OK, now why are people interested in this? Physicists are interested because in solid state they teach you that if you have atoms arranged in one of the patterns that produce symmetries, then if you pass X-rays through this symmetric lattice, you get a diffraction pattern, very similar to a gradient that we saw here, except that we're in 3D because the symmetries are very simple in 2D. I think in 2D you only get two-- you get the rectangular and the hexagonal symmetry.

In 3D, you get many more symmetries-- I forget, how many do you get? I think it is 7 groups or 12 or anyway. I should refresh my solid state physics. But anyway, each one of these gives rise to-- each symmetry gives rise to a distinct diffraction pattern. And then crystallographers basically bypass the next stage to a crystal and observe the diffraction pattern-- they can guess what was the original symmetry that produced it.

We also know that non symmetric patterns can produce diffraction. And that is what won Crick and Watson the Nobel Prize because they figured out that you have a helix, which not quite symmetric, it can still produce a diffraction pattern. And Crick was a very clever mathematician who solved it. And he won the Nobel Prize for it. Anyway, recently people have observed that quasi-crystals like this one-- if you take pentagons and you stick them next to each other, they also produce a diffraction pattern, which looks actually like a bunch of bright dots in a pentagonal kind of symmetry.

It looks very funny. If you go to this Caltech website, you will see it. And then people also noticed that this kind of thing also produces Talbot patterns. If you you diffract a field from a pentagonal quasi-crystal, it will produce Talbot patterns as you'll go behind it. And that caused quite a bit of stir. Actually there's a lot of papers analyzing this and looking into this interest.

Anyway, so I don't know if this has any use-- it's kind of interesting physics. But yeah, people spend some time on that. Oh, yeah. Actually, Talbot effect found a use recently in a very surprising way in lithography. It is a very easy way to reproduce a pattern lenslessly, as long as the pattern is periodic.

And so for planar lithography is not that interesting, but nowadays many people who want to make photonic crystals which are periodic. So in the photonic crystals, you want to start with the volume. So this is a cross-section, and you want to write patterns that look like this. So for example, what you do is you take a polymer-- OK, I take it back. You take a monomer and if you could expose some [INAUDIBLE] these regions and then you develop it, then you would end up with a structure that has low index in the unexposed regions and high index in exposed regions.

That is a photonic crystal. So the question is how could you do this. Well, one way is with a Talbot effect because a Talbot effect will actually focus the light very sharply in the Talbot planes and then you'll get your exposure. But everywhere else, the light is kind of diffuse. So it does not expose very much. So if you play with the threshold of the photoresist, you can actually get the pattern to work like this.

So there's a colleague of mine at MIT, Ned Thomas. I don't know if any of Ned's students are in the class. But he came up with a very clever system that he can expose and create photonic crystals this way. And then interestingly, a student of mine discovered that if you vary the duty cycle of this thing, I may be disclosing-- everybody here is MIT, right?

And those of you who are NUS are sort of bound by a vow of secrecy or whatever. But anyway, I may be disclosing something that we have a patent pending for. But anyway, so it is also well known-- people have analyzed patterns with a duty cycle that these different [INAUDIBLE] point have. So for example, if you have a gradient that is like this-- this also produces a Talbot effect.

So it will also produce replicas of that pattern. So but my student discovered that if you vary the duty cycle, that's really interesting and we don't fully understand why but it is true. I mean, it is physically correct-- we don't know how to explain it very intuitively. But mathematically, you find that you have a pattern with a fixed period. But the variable duty cycle, for example, like this, it also for this is a Talbot effect. Not forever, but over a finite region maybe about-- I don't know, it depends but maybe about 10 or so Talbot planes, it reproduces itself.

So it is also possible to make non-periodic patterns, which as [? Colin ?] knows, we're very interested in making non-periodic kind of-- they're not for [? tiny ?] crystals anymore, but they're not periodic strong index modulations. That is possibly one way to do it.

AUDIENCE: George, it seems like it seems like that was still a periodic pattern though.

TA: No, because the duty cycle varies. We call it periodic because you must have the same duty cycle in every period.

AUDIENCE: So even though it repeats every period with a different duty cycle, that's not periodic.

TA: I don't know, would you call this function periodic?

AUDIENCE: I understand what you mean about the duty cycle changing, but it's got a definite period to it where features are--

TA: Yeah. Yeah, but it's non-periodic. OK, what you're saying is a possible explanation for why it has a Talbot effect. It has a Talbot effect because it has this fundamental period underlined the variable duty cycle.

AUDIENCE: Right, I look at the duty cycle as changing the orders that are coming out for that particular region.

TA: Well, in a slowly varying approximation, perhaps. Yes. I mean, this thing-- I've not looking at the diffraction pattern of this thing in detail. But yes, it would produce diffracted orders. But it is not clear how you would compute their strength from a Fourier series. At least, I don't know how to do it. There may be a way with somebody adiabatic approximation. Because it is true that in our simulation, the duty cycle was very slowly varied.

And I believe it is also true. In fact, I think Will-- my student who did this-- he also did this simulation with a fast, varying duty cycle, and he found out that the Talbot effect actually does not happen if the duty cycle varies fast. So, yeah. There's certainly an adiabatic approximation happening here. But I guess we haven't really worked out all the details. But it is sort of interesting still.

Anyway, yeah, it's one of those things that-- the reason I'm bringing it up is because the Talbot effect doesn't look like something particularly useful. It looks more like a curiosity. But yeah, I guess sometimes even curiosities can find applications, right?