The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high-quality educational resources for free. To make a donation, or to view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at ocw.mit.edu.

## PROFESSOR STRANG:

OK, thank you for coming today. The day before Thanksgiving. Day before my birthday, actually. So it's a special day. Everybody gets an A for showing up. Even you. So, let's see. Last time, I wrote down these formulas for the Fourier integral transform. And I thought I'd just write them again so you kind of photograph them and remember them. They're easy to remember. As always, you take the function, you multiply by e^(-ikx), and you integrate. To get the amount-- So k is my frequency variable. It could well have been omega or some other variable. I stayed with k because it was k in the Fourier series. So that's the calculation which as always, I mean, these are integrals that we may be able to do if the function is especially nice, or we may not. But that's the formula. And then to reconstruct the function, we combine all they e^(ikx)'s in that amount to get f(x) back. OK, nice formula.

So I did one example last time, and now could I just double it up? This is also in the textbook and so this is now going to be an even function. Last time the example I did was zero up to x=0. This time I'll make it symmetric, make the function even. And then I have two pieces in the integral. And if you remember what it was, you remember that this, I'll just remind you what we did. We wrote that it's  $e^{(-(a+ik)x)}$ . That was clear. And then when we integrated we got that same function divided by -(a+ik). And then we put in the limits. And the answer was, let me maybe write the answer down here. Was just at x equal infinity the limit was zero because this thing is tailing off. At x=0 this is one. It comes in with a minus because that's the lower limit. So it was 1/(a+ik) for the first half. And we were not surprised to see this 1/k because the first half all by itself has that jump, from zero to one. So we see that jump reflected in slow decay.

Alright, but now I'm making it even. What are you going to guess for the rate of decay of f hat of k for this function? This function no longer has a jump. But it does have - I don't know were we saying ramp, or corner? This is not a smooth point here, because the derivative going up is plus a, so I'll just put a circle right, the derivative is  $a^*e^*(ax)$ , and at x=0 that would be plus a going up. And here the derivative is  $-a^*e^*(-ax)$ . Put in x=0, and the derivative coming down is minus x. So there's a jump in the derivative. So what, just before we see it, what will you

expect for the rate of decay of the transform? One over k to what power, now? So it didn't have a jump, a jump was 1/k. This has a jump in the derivative, so we're expecting one over k squared. k squared, it'll be one order smoother. OK, you can easily see that happen. Because this part, well this is just  $e^{(a-ik)x}$ , which I'm going to integrate to get this thing over a-ik. And I'm going to plug in the limits minus infinity and zero. And at minus infinity I'll get nothing, this  $e^{(ax)}$  at minus infinity will be zero. Where the function starts. Way down at zero. So and at x=0, this is a one. So I just get 1/(a+ik). Over a minus, thank you. Right, over a minus.

And somehow it can't be an accident that this is the complex conjugate of that somehow. That's not a surprise. OK, so let's put those together into a single fraction and see what we have. So the denominator of that fraction will be this times this. And that's the most basic multiplication of complex numbers. That one times its conjugate gives me what? It gives me an a squared. And what else? A plus k squared because i times minus i is plus k squared, and no imaginary part. There's a plus i k a and a minus i k a, all we're seeing here is the sum of squares. The usual z times z bar. And in the numerator, let's see. When I put it over this-- So this was putting it over this common denominator, so I should have an a-ik going up on top there. And an a+ik going up on top here. Right? Those are my two fractions. That over this, and that over this. And now that numerator simplifies, oh look it's great. I'm getting a real answer. And because the minus ik and the plus ik cancel, and it's just the 2a. And probably no surprise that somehow that that's the jump in slope. That must have something to do with that 2a.

So we got a real even, Fourier-- f hat from my real even f. And it decays like k squared. OK, so that's another good example. A very useful example. Right. I could add other examples. One quite, before I use that in application, let's do just a few more examples. Suppose f(x) is the delta function. What's f hat of k? Can you just plug in f(x)=delta(x) here? Do that integration, and what does f hat of k come out to be? One. Because if it's a delta function there, at x, at x=0 the spike is at x=0, so I plug in x=0, I get one. So we're kind of, we'd be surprised if it wasn't a constant, right? A delta function in physical space goes in-- has all frequencies in equal amounts. And it's a constant in frequency space.

Then there's one more that takes a little trick to do, but it's a very neat one. f(x) is e to the minus x squared over two. Do you recognize that as an important function, e to the minus x squared or it usually has that e to the minus x squared over two or sometimes an e to the minus x squared over two sigma squared, a rescaling? But this would be the bell-shaped

curve. The bell-shaped curve. It decays very quickly. The variance, because that's a two and not a two sigma squared, the standard deviation is one here. The variance is one. So it's a bell shaped curve that has about 2/3 of its area between minus one and one. This is the all important function for probability. The normal distribution, the Gaussian, both of those words are used, it's the most important probability distribution. I need a one over square root of 2pi to make the total probability be one. But let me just leave it there. That's a very, very important function. It's also going to be important in the heat equation. In math finance, shows up all over the place. And its integral would not be easy to do from zero to one. The integral of that function, from zero to one, we have tables of it. To the nth place. But so there's no simple, elementary function whose derivative is this. That x squared is what's making the integral tricky.

So from zero to one, we just have to give it a name. So, error function. This would be E-R-F, error function, the integral of that thing correctly normalized. I'm just saying, important, important function. And it turns out that integrals from minus infinity to infinity can be done. So, beautifully, by some trickery, we can find the transform of this. We can find the transform of this, we can do this integral from minus infinity to infinity, where we could not do it from zero to one. So I'll just write down the answer for this guy. Only because it's such a key example. It's some constant that involves 2pi times e to the minus k squared over two. Boy, that's pretty amazing. Right, the Fourier integral transform, f hat of k, has the same form as the function. And of course this function is infinitely smooth. So its transform decays infinitely fast. Yeah, there's no problems like one over k squared here, there are no bumps in the bell shaped curve.

So I won't push that example except I'll use it. What else should I say just to, like, emphasize that this is such an important distribution in probability? Why is it important in probability? That's the question. Why does everybody assume, if you can get away with it and don't have any natural alternative, everybody assumes that noise, whatever, is coming with a normal distribution. So, in other words, with a sigma squared in there. So a normal distribution, that has mean zero because it's absolutely centered at the origin and it has variance one, but I could change the variance and that would just spread out or tighten the bell-shaped curve. Why is the bell-shaped curve so important? That's certainly, we're not going to launch into theory of probability but it's the central limit theorem. So let me just use those words. The central limit theorem that says that if I start with other probability distributions, like I'm flipping a coin. I flip a coin a million times. Then the mean, and let's say zero for tails, one for heads. OK,

so I flip, flip, flip. Well, the mean of that, the expected mean is what, half a million, right? Half tails, half heads. So if I give zero for tails, one for heads and flip a million times the mean would be about half a million. And then, so let me center the mean. I could have centered it by taking minus one and one. That would have been smarter. Minus one and one. Minus one for tails, one for heads would have had a mean of zero. And then it would be natural if I have a million of these to divide by a thousand, I think. Of course, the answer won't be zero, right? If I do a million flips it's not going to come out exactly half a million and half a million.

I'm remembering. I used to have a long discussion with a nice guy in college. He ran for Mayor of Boston, actually. But he had the idea that after a million flips, suppose there had been more heads than tails. Then the next flip, he figured, was more likely to be tails. I couldn't convince them that this was not mathematically the right thing to think about. And all I did was say don't go to Las Vegas. I mean, if you're thinking that way, save your money. So, anyway. But this is much studied. The variation, what that curve looks like, that's quite interesting. But my point is, that as the number gets bigger and bigger and we scale it properly, the distribution will approach the normal. All sorts of distributions. If I just repeat and repeat experiments and scale it, the central limit theorem says you're always going to the normal distribution. So that's highly important. OK, and it comes up different places. And it's quite a neat function. OK, so that's some examples.

Now, let me use, like every topics that I introduce, I want to find a use for. So now, can I start on this one? Constant coefficient differential equations. I'm going to write down a differential equation, which will look pretty much like the ones we started this course with. And I could, well, let me write it down. Minus d second u dx squared, we're used to that. Now let me put in an a squared u. Which is a lower lower order term, we could deal with that. Equals some f(x). And now, because I want to do Fourier integrals, I'm thinking all x. We're on the whole line. Instead of the interval (0,1) where I might use Fourier series and have sine series or cosine series, depending on the boundary conditions. Here, the boundary condition is just everything drops off at infinity. And minus infinity. So all these functions, we can do these integrals. OK, so there's a good question. What's the solution? We could tackle it, but I want to suggest to use Fourier. So it's not the only way, but it's one way to see it. So now if I use, what do I mean by using Fourier? It means I'm going to take the Fourier integral transform of every term. So when I take the Fourier transform of the right-hand side, I'm going to get f hat of k, whatever. This is known, of course. This guy is given. That's the source term. And u is the unknown. OK, so I'm going to take the Fourier transform of every term, well this is, a is a constant. a had to

be a constant, or I couldn't do, you know if a depended on x this would be some multiplication and the transform would be a mess. Fourier applies when you've got constant coefficients and nice boundary conditions. And here our boundary conditions are nice, they just go to zero fast.

OK, so the transform of this is, that's a constant. u hat of k. And what's the transform of that? So this is our chance to use probably the most important rule for Fourier integrals. Maybe you'll tell me what it is. You should think what it is. If I take a derivative, that's the rule. If I take a derivative of the function, what's happening in frequencies? I could make that happen here. If I took the derivative, yeah. So maybe if I take the derivative here, so here it's just remembering the rule. Suppose I take the derivative of this equation. I get this integral, and what would f' be? What would f hat, sorry. If I take the x derivative of this, if I take the x derivative of this equation, what happens on the right-hand side when I take the x derivative? Down comes ik. Down comes ik. So when ik is coming down, I won't even finish that equation. And ik is coming down, when I take the derivative. So the derivative, the transform, is multiplied by ik, higher frequencies are emphasized now because of that k factor. And now if I take two derivatives, I bring ik twice. So that's i squared k squared, the i squared and the minus give me a plus. So that's just k squared. u hat, of k. OK with that?

And now, we get an immediate formula for u hat of k, the solution. Well, it's the solution but it's in frequency space. If we wanted to know it in x space, as we do, we've got to transform back. But what do we get here? It's just f hat of k. Divided by, this is just multiplied by a squared plus k squared. OK, so that's the answer. In frequency space. That was simple. And then if I wanted it in x space, I take the reverse transform. Notice that this is, hidden here are the same three steps that I emphasize all the time about using eigenvectors and eigenvalues. Do you remember those three steps for solving differential equations? Difference equations, linear equations, whatever? The three steps were, find the coefficients, expand everything in eigenfunctions. I won't write, I'll talk. The three steps were expand in eigenfunctions, follow each eigenfunction function separately, that was the trivial step with just a division, like this division. And then use those coefficients of the eigenfunctions, combine them all back to get the answer. Right?

Step one: write it in the right basis, step two: easy in that basis. Step three: go back to your physical space. We're doing exactly the same thing here. These e^(ikx)'s are the eigenfunctions of this thing. They're the eigenfunctions. And here the eigenvalue of this stuff is k squared plus a squared. And that's what we divided by. And then the final job of going back

from u hat to u, by-- So write u there. Can I do, this f, now it's really u that I'm wanting to bring back to physical space. So just for the sake of your eyes seeing it, let me put a u in. So that's the answer in a way. It's the answer, it's a formula for the answer. It did depend on our being able to do two integrals. The integral from f to f hat may not have been easy, and then the integral from u hat back to u, this integral, might not have been easy. So it's a formula. OK, now I want to go with it a little longer. Because I want to show you how the delta function pays off. So let me do the example where f(x) is the delta function. So now we're really close to where this course began. Differential equation with a delta function. The only new thing is, we're not on an interval, we're on the whole line.

So I take transforms, so what's the transform now of this specific f(x) is? One. We just saw. OK, so now we get a one there. Now we just divide by here. And we've got a one here. So we were able to go, this was an integral we could easily do, to get from delta to delta hat, which was just one. And fantastically, this is an integral to go back to u(x), to go back to u(x), that's an integral we can do. Well, you may ask how can we do it. How do I find the u(x) that has this transform? Well, I either use complex variables to do integrals like this, residue methods that are in Chapter 5, or I look in a table. Or I look at the blackboard over there. I think that's the best way. Look at this blackboard. Right? Because this is the answer we got. We got that same answer apart from a constant factor 2a. So this is our function. This is our solution, u(x)is this. What am I going to call that? Two-sided pulse? I'll call that the two-sided pulse? Maybe I should give it a name but I'll just write out those words two-sided pulse. That's it, divided by 2a. So we've got the answer. Let me just make a little more space here. This was one over a squared plus k squared, and now having seen that already, I just say yep, that must be it. It's the two-sided pulse, and I have to divide by 2a. Do you see that that's the correct answer? We can substitute that in the equation and see that it works. I mean, so we have solved the problem. We have solved the problem when the right side was delta.

Let's put it into the equation. So this is just because we did this, it's nice to do it again after all this time. So I put it in the equation. This two-sided pulse over 2a. So what's my equation? Well, this is zero most of the time. So I believe that if I plug in this function, it will give me the zero. Do you want to just plug it in? I believe that if I plug in a^(-x), or a^x, either one, can I just check that u-- Try u=e^(-ax). Put it in and just see that I get zero. Because yeah, two derivatives bring down a squared with a minus. And there's a squared with a plus. It works, right? Two derivatives of this function bring down minus a twice, so that's a squared. So it's minus a squared, plus a squared. Works. And then, of course, the important point is x=0

where the spike is. What happens at the spike, going back to the beginning of the course? This term is going to be unimportant compared to this term. What do I see? With -u" equal a spike, what was the solution to that? u had a corner, right? The slope of u, what did the slope of u do? It dropped by one, was that right? The slope of u dropped by one. We used to have corners going up and down and the difference between the slopes was one. And here, the difference between the slopes, ah, look. The slope has dropped by 2a, and when we divided by the 2a, it was just right. And now, when I divide by the 2a, this has a slope of 1/2. This has a slope of minus 1/2, the drop is one. And we're right. It solves the equation. Nobody doubted that. OK, so that's great. We have found the solution to this equation, when the right side is delta. Good.

Now, can I ask you do you remember the name? There's a special name for the solution when the right side is a delta function. Whose name is associated with that? So that this particular u, I'm going to give it another letter. It's the particular u, the special u when the right side is delta, and whose name is associated with that solution? Green. It's the Green's function. Green's function. The famous Green's function. Green's function is just like an inverse to the problem. This is like having an identity on the right-hand side. It's like there it is. So let me just use G for Green's function. So that's the Fourier transform of the Green's function, and this is the Green's function. This is the Green's function. Now I can give it its name, Green's function, when I divide by the 2a. And now the slope is 1/2 going up, and minus 1/2 coming down. And it's all right. OK, so we found the Green's function. We found the fundamental solution to the equation, and this is it. It was straight lines, right? It was straight lines in the first weeks of the course. But now there's an exponential drop-off caused by this additional term. OK, good. So that's straightforward. Depending on our being able to recognize or do the transform back to the x space. Now comes the question, what about the original f(x)? How can the Green's function be used? So you're seeing now, what use is this Green's function? With that righthand side, when the right-hand side is something different? When the right-hand side is some different f(x)? So let me go back to an f(x) on the right. And then there's an f hat of k, after the transform. How can I use the Green's function for a general source? The general source term, a general load? This is a fundamental idea. I would say fundamental. How do you use the Green's function? And remember, the Green's function is like telling you the inverse matrix. So it can't be too hard. It's like solving a linear system when you know the inverse matrix. So that's the analogy, but let's just focus on the particular question. I think the intuition, you should have an intuition for how the Green's function works. So the Green's function was the solution when the source term was a delta. And here's the intuition. It's rough, but it works. Any source

term, f(x), is in some way a combination of delta. If f(x) is a combination of deltas, then our answer u(x) is the same combination of the Green's function, right? If the right-hand side is some combination of this special delta, then the solution will be the same. This is just linearity. Superposition, whatever short or long word you like to use. So if I can make sense of that statement, that f(x) is a combination of deltas, then I'm in.

Now, what do I mean by a combination of deltas? I mean, well, those deltas are going to be shifted deltas. Obviously the single delta, delta(x), is a spike at the origin. That's only one point. I want to combine delta of x and its shifts. So I'm going to have to expect to be using G(x) and its shifts, right? OK so now I'll just say this again. I'm thinking of f(x) as a combination of delta and its shifts, and then the solution u will be the same combination of G(x) and its shifts. So now you just have to tell me what combination. What combination of delta and its shifts? Maybe you'll allow me. Let me just do this maybe on that board. I just can't help writing down the discrete case. So, in the discrete case, the delta vector corresponds to something like [1, 0, 0, 0]. Right? That was a typical delta vector with a one in the zeroth position. Then its shifts would be [0, 1, 0, 0], that'd be a shift. And another shift would be [0, 0, 1, 0]. And another shift would be [0, 0, 0, 1]. So now there is the delta vector and its shifts. These four guys.

OK, now suppose my f, my right-hand side, is [1, 2, 3, 7]. I want to write that as a combination of those deltas. This would be in the case when if I know the solution for each of these guys, the Green's function, the inverse matrix. Everybody sees if I know the solution to those four, I know the inverse matrix. Right? Because if I can solve with those four right-hand sides, those four solutions are the columns of the inverse matrix. Right? You remember that if I had a matrix A and I was looking for its inverse, I solve A A inverse equal I. And I is just these four guys. So A inverse is the solution from these four guys. OK, now everybody's going to tell me what's the solution for this right-hand side [1, 2, 3, 7]? Suppose this guy has solution, so A inverse, the columns of A inverse are this Green's function. This Green's function with a shift. Maybe SG, Green's function with a shift. Yeah, S squared G, Green's function with a double shift. S cubed G. I'm just cooking up, I never used these letters before. But what's the answer? Then u is what? It's one times the Green's function. And it's two times the guy, right? This f is just, this f is just, is one times the Green's function, two times the shifted. Three times the double shifted, and seven times the triple shift. Right? Just taking four minutes to do something simple because over there, when I use the continuous case it'll look a little strange, but here it's so easy. That'll involve integrals, this involves a sum. So what is it? I have G, twice the shift of G. Three times the double shift of G, and seven times the triple shift of G. Right? By linearity, by superposition, if this is my f, this is my u. Everybody's with me here, right? That if I take a combination of f's, the answer is the corresponding combination of G's. OK, good.

I didn't mean that. That wasn't so great. That shouldn't have been called, I didn't mean to call those the Green's functions. I meant to call those the deltas, right? And the Green's functions were the answers. So this a shifted delta, this is a doubly shifted delta, this is a triply shifted delta, and now this is one of the delta, two of the shifted delta, three of the doubly shifted delta. The f is a combination of deltas and its shifts. The u is a combination of Green's function, and its shifts. Apologies for making that mistake, but maybe it's brought us back to the point. So the point is, express your function f, your source, as a combination of deltas, just exactly our plan. Then u is the same combination of the G's with the same shift. Alright, now back here. How do I express f(x), in what way is f(x) a combination of deltas? I mean, slow down just to see that point. How much of delta, of the spike at x=3, how much of delta(x-3), so that's a spike at three, right? How much of that do you think I need in f(x)? f of? f(3). Whatever f is at that point, x=3, that's the amount I need there. So this would be f(3). So that that part would sort of have the right pep, the right punch, at the point three. Now, three could be any point, that's any point along the line.

Let me, I can't use x for other points on the line because I've got an x in the formula. Let me use t. So this was t equal to three. Now I want to do it at all points. I want to take f(2), and f(pi), and f at everything and put them together. And of course putting them together in the continuous case means not sum, as I did there, but integrate. So I have to, now I'm going to change three to a t, because this is the amount of f of, so that's the amount, that's the delta functions which spike at t, multiplied by f(t), and now how do I get f(x) out of this? I put them together. dt. I add up, this is the combination I've been talking about. So this is like any f. This is like a crazy delta function identity. Actually, it's not crazy, it's the identity we've used our whole lives. Or at least our whole 18.085 lives, which is all that matters, right? OK, so I'm integrating a delta function. And I want to see, do I get that answer? And you're going to say absolutely, clearly you get that answer. No. Everybody knows that if I integrate something times a delta function that I plug in at the point t=x, where that spike happens, I put t=x, I get f(x), correct. So that's like any, that's an identity or whatever you would like to call it. And now just tell me the final answer.

Let me put it on the board above. Now, so this was, this expressed my f(x) as a combination of deltas. Just the way over here I expressed [1, 2, 3, 7] as a combination of deltas. Now, what's

that u? What's the function u? The solution that comes from f(x). Just erase here so that I can put all of them. The point of the whole example now is for you to tell me what's u(x)? Can you do it? You see what u(x) is going to come out? It's going to come out nicely. I've written the right-hand side f as a combination of deltas. I know the answer, for delta. It's G. What's the answer when the delta is shifted along? What's the Green's function when the spike is moved along to a point t? Just, because this constant-coefficient, translation-invariant, LTI problem is shift invariant, the answer, when the spike is moved along, is just the answer G moved along. So here's my answer. All this, this is my input and my output is the same integral, this from minus infinity to infinity, of -- where it was f, now it's G. Well no, what do I want? Help me out here. No, f(t) is just the amount of the delta, but now what's the solution? What do I write now? G of? x-t. That's it. You see why that works? Because this was the input, G's the output. The problem was shift invariant, so if I shifted the input I shifted the output. It was linear, so if I add up a bunch of deltas, the solution is add up a bunch of G. That's the answer. Oh, I could just say one thing more. But you've got it if you see math. So that's the point, that if you know the Green's function, well yeah.

Maybe from a practical point of view, what have we done? The original way we did it involved our computing two integrals. We had to, if we were given an f(x), we had to find its transform f hat of k, that we weren't sure we could do with pencil and paper. And then we got an answer with an f hat of k up here and then we had to transform back. Step one and step three, we had to do. Now this is better. This is like better, because we were able to get an explicit answer, G, when this was a delta. You could say I've got it down to one integral. Well, for whatever that's worth. I was going to say, probably can't do that one either depending what f(x) is. But that's a nice way to see the answer, you have to admit. And it's because the problem is shift-invariant and I can write the answer that way. OK, and now one more thing about this. And I've written the word, the key word, down here. Convolution. Do you see the nice way to write that answer? It's the convolution of f with G. We didn't do integral convolutions, we just did the discrete sums, but I mentioned that in the integral case, you have this same thing, t and x-t adding up to x, just the way we had k and n-k adding up to n. In other words, this is just notation. But I'm just going to write the answer in a nice way. So after all that lecture, the answer to the differential equations is in three symbols f star G. Where this just simply means this. This is the continuous convolution, not cyclic, and there's the answer.

I'll allow myself one more thing. Here we had a convolution for the right-hand side. We started with this, this is a convolution. Now, what are the three symbols that I write down? For the

shorthand for this equation? So this was to be true for any f. And now can I write down, how do I write? f(x) is equal to, what is this right-hand side? It's f convolved with? So any f is the same f convolved with delta. In convolution, delta is one, Because when I go to the other space and I get a multiplication, it really is one, right? In the other space. So in the other space, so this convolution in x space turns into multiplication in frequency space. And it just tells me that f hat is f hat times one. So that's the way to look at it in physical space. And this is the way to look at the solution. So, one more thought and I'll come back to that. So this G, the Green's function, this is what a CAT scan does, what an X-ray telescope does. What all sorts of physical things do. Provided we can assume this translation invariance, which is never perfectly true because the telescope is finite. But a telescope takes the star, takes the light signal, convolves it with the telescope's own little Green's function. It blurs it, it's the point spread function. It's the blurring function, G. The Green's function on a telescope is somehow, that's what's you're convolving with. And if you want to find that star as a bright, single point. You've got to do deconvolution. You've got to do a division to get the G out. May I just say those words and then it's Thanksgiving? That a machine, a sensor which is translation invariant, or you could say close enough to pretend it is, because nothing is going to be perfectly translation invariant all the way out to infinity. But if it's near the star. So I look at this point star in the telescope I see a blur. That's because the telescope has convolved the correct thing, that I should have seen, with G. It's blurred it by its point spread function. So what if the person, the factory where the telescope was built can test this whole thing on points. And it can find the point spread function. And if I knew G, then I could undo it. And get a clear picture. So it's that step that won the Nobel Prize for the CAT scan, and I'm sure is winning Nobel Prizes for astronomers. OK, have a great Thanksgiving, I'll see you Monday. Good.