MITOCW | ocw-18-085-f08-lec31_300k

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.

OK, so. Pretty full day again. I had last time introduced the Fourier matrix, the discrete Fourier transform. Well, more strictly, the discrete Fourier transform is usually this one. It takes the function values and produces the coefficients. And then I started with the coefficients, added back, added up the series to get the function values. So F or F inverse. So we didn't do examples yet. And one natural example is the discrete delta function that has a one in the zero position. That's easy to do. And then we should do also a shift, to see what's the effect, if you shift the function, what happens to the transform. That's an important rule, important for Fourier series and Fourier integrals, too. Because often you do that, and it's going to be a simple rule. If you shift the function, the transform does something nice. OK, and then I want to describe a little about the FFT and then start on the next section, convolutions. So that's fun and that's a big deal.

OK, about reviews, I'll be here as usual today. I think maybe the 26th, just hours before Thanksgiving, we can give ourselves a holiday. So not next Wednesday but then certainly the Wednesday of the following week would be a sort of major guiz review in the review session. And in class. OK, ready to go? On this example gives us a chance just to remember what the matrix looks like, because we're just going to, if I multiply this inverse matrix by that vector it's just going to pick off the first column and it'll be totally easy so let's just do it. So if y is this one, I want to know about f. What are the Fourier coefficients of the delta function? Discrete delta function? OK, before I even do it, we got a pretty good idea what to expect. Because we remember what happened to the ordinary delta function, in continuous time. Or rather, I guess it was the periodic delta function. Do you remember the coefficients, what were the coefficients for the periodic delta function? You remember those? We took the integral from minus pi to pi of our function, which was delta(x). And then we had to remember to divide by 2pi, and you remember the coefficients are e^(-ikx). This is c_k in the periodic case, the 2pi periodic case. The function is the delta function, and you remember that if we want coefficient k we multiply by e[^](-ikx). That's the thing that will pick out the e^{+ikx} term. And of course everybody knows what we get here, this delta, this spike at x=0, means we take the value of that function at zero, which is one, so we just get 1/(2pi). For all the Fourier coefficients of the delta function. The point being that they're all the same. That all frequencies are in the delta function to the same amount. I mean that's kind of nice. We created the delta function for other reasons, but then here in Fourier space it's just clean as could be. And we'll expect something here, too.

You remember what F inverse is? F inverse is 1/N, instead of 1/2pi, and then the entries of F inverse come from F bar, the conjugate. So it's just one, one, one, minus-- Well, I've made it four by four here. This is minus omega--

No, it isn't minus omega. It's it's omega bar, which is minus i, in this case. Omega bar, so it's minus i, that's omega bar. And the next one would be omega bar squared, and cubed, and so on. All the way up to the ninth power. But we're multiplying by one, zero, zero, so none of that matters. What's the answer? I'm doing this discrete Fourier transform, so I'm multiplying by the matrix with the complex conjugate guys. But I'm multiplying by that simple thing so it's just going to pick out the zeroth column. In other words, constant. All the discrete Fourier coefficients of the discrete delta are the same. Just again. And what are they? So it picks out this column, but of course it divides by N, so the answer was 1/N [1, 1, 1, 1]. It's just constant with the 1/N, where in the continuous case we had 1/(2pi). No problem. OK. And, of course, everybody knows, suppose that I now start with these coefficients and add back to get the function. What would I get? Because, just to be sure that we believe that F and F inverse really are what they are supposed to be. If I start with these coefficients, add back to put those in here and reconstruct, so 1/N [1, 1, 1], what will I get? Well, what am I supposed to get? The delta, right? I'm supposed to get back to y. If I started with that, did F inverse to get the coefficients, that was the discrete Fourier transform, now I add back to get, add the Fourier series up again to come back here, well I'll certainly get [1, 0, 0, 0], and you see why? If I multiply F, that zeroth row of F is [1, 1, 1, 1], times [1, 1, 1, 1] will give me N. The N will cancel. I get the one. And all the other guys add to zeroes. So, sure enough, it works. We're really just seeing an example, an important example of the DFT. And the homework, then, would have some other examples. But I've forgotten whether the homework has this example.

But let's think about it now. Suppose that's my function value instead. OK, so now I'm starting with the delta. Again it's a delta, but it's moved over. And I could ask, I really should ask first, in the continuous case, suppose I, I can do it with just a little erasing here. Let me do the continuous case for the delta that we met first in this course. I'll shift it to a. If I shift the delta function to a point a, well, I said we'd met the delta function first in this course. At a, good. But when we did it wasn't 2pi periodic. So we still have, in fact, the Fourier integrals next week. Will have a similar formula. The integral will go from minus infinity to infinity and then we'll have the real delta, not periodic. Here, we have, and people call it a train of deltas. A train of spikes. Sort of you have one every 2pi. Anyway, that's what we've got. Now, you can see the answer. This is like perfect practice in doing an integral with a delta. What's the integral equal? Well, the spike is at x=a. So it picks out this function at x=a, which is e^{-ia} . So not constant any more. They depend on k. The 1/(2pi) is still there. So it's-- But still, the delta function shifted over. I mean, it didn't change energy. It didn't change, it just changed phase, so to speak. And we see that-- I would call this like a modulation. So it's staying of absolute value one, still. But it's not the number one, it's going around the circle. Going around the unit circle. So it's a phase factor, right. And that's what I'm going to expect to see here in the discrete case too. If I do this multiplication by one there, it picks out this column, right? That one will pick out this column, so you see it's-- Maybe I come up here now. Shall I just-- When I pick out that column, the answer then, I guess I've got the column circle, there it is. Minus i, minus i squared, minus i cubed. You see it's like k equals zero, one, two, three. Just the way here, we had k, well we had all integers k in that function case. Here we've got four

integers, k equals zero, one, two, and three, but again it's the minus i, it's the e to the, it's the w bar. In other words, the answer was one, w bar, w bar squared, w bar cubed. Just the powers of w with this factor 1/N. Here we had a modulation. It's the same picture. Absolute value's one.

And and what about energy? Having mentioned energy, so that's another key rule. The key rules for the Fourier series, just let's think back. What were the key rules? First, the rule to find the coefficients. Good. Then the rule for the derivatives. This is so important. These are rules. Let's say, for Fourier series. For Fourier series. Let's just make this a quick review. What were the important rules? The important rules were, if I had the Fourier series of f. Start with the Fourier series of f. Then the question was, what's the Fourier series of df/dx. And now I'm saying the next important rule is the Fourier series of f shifted. And then the last important rule is the energy. OK, and let's just, maybe this is a bad idea to, since we're kind of doing all of Fourier in, it's coming in three parts. Functions, discrete, integrals, but they all match. So this is what happens to the function. What happens to the coefficient? So this starts with coefficient c_k, for f, what are the coefficients for the derivative, just remind me? If f(x), so I'm starting with f(x) equals sum of c_k*e^(ikx). Start with that. And now take the derivative. When I take the derivative, down comes ik. So you remember that rule. Those are the Fourier coefficients of the derivative. Now what's the Fourier coefficients of the shift? If I've just shifted, translated the function, if my original x was this, now let me look at f(x-a). You'll see it. It'll jump out at us, it'll be a sum of the same c k's $e^{(ik(x-a))}$. So what are the Fourier coefficients of that? Well there is the e^(ikx). Whatever's multiplying it has got to be the Fourier coefficient, and we see it is c k times e^(ik(-a)). e^(-ika) times c k. Right? And, of course, that's just what we discovered here. That's just what we found there, that when we shifted the delta, we've multiplied by this modulation, this phase factor came into the Fourier coefficients.

And now finally, the energy stuff. You remember the energy was, what's the energy? The integral from minus pi to pi, of f(x) squared dx is the same as the sum from minus infinity to infinity of the coefficient squared. And somebody correctly sent me an email to say energy and length squared are you really, is there much difference? No. No. You could say length squared here, I'm just using the word energy. Now, I left a space because I know that there's a stupid 2pi somewhere. Where does it come? You remember how to get this? You put that whole series in there, multiply by its complex conjugate to get squared. And integrate. Right? That was the idea. Isn't that how we figured out, we got to this? We started with this, length squared. We plugged in the Fourier series. This is f times f bar, so that's this times its conjugate. And we integrated, and all the cross terms vanished. And only the ones where e^(ikx) multiplied e^(-ikx) came. And those had c_k squared. And when we integrated that one, we probably got a 2pi. So that's the energy inequality, right. For functions. And now what I was going to say is, you shouldn't miss the fact that in the discrete case, there'll be a similar energy inequality. So we had y was the Fourier matrix times c. Now, if I take the length squared of both, y, so I'm going to-- Right? That's the same as that. Now I'm going to do the same as this. I'm going to find the length squared, which will be y transpose y. No, it

won't be y transpose y. What is length squared? y bar transpose y. I have to do that right. That will be, substituting, that's c bar F bar transpose times y is Fc. Just plugged it in, and now what do I use? The key fact, the fact that the columns are orthogonal. That's what made all these integrals simple, right? When I put that into there, a whole lot of integrals had to be zero. When I put this in, a whole lot of dot products have to be zero. Rows of F bar times columns of F, all zero, except when I'm hitting the same row. And when I'm hitting that same row I get an N. So I get, this is N c bar transpose c. And that's c squared. That's the energy inequality, it's just orthogonality once again. Everything in these weeks is coming out of orthogonality.

Orthogonality is the fact that this is N times the identity. Right? Well, OK that's a quick recall of a bunch of stuff for functions. And just seeing maybe for the first time the discrete analogs. I guess I don't have a brilliant idea for the discrete analog of the derivative. Well, guess there's a natural idea, it would be a finite difference, but somehow that isn't a rule that gets like, high marks. But we saw the discrete analog of the shift and now we see the energy inequality is just that the length of the function squared is equal to N times the length of the coefficients squared. OK with that? Lots of formulas here. Let's see. Do some examples. I mean, these were simple examples. And I think the homework gives you some more. You should be able to take the Fourier transform and go backwards. And when we do convolution in a few minutes, we're certainly going to be taking the Fourier, we're going to be going both ways. And use all these facts. OK, I'll pause a moment. That's topic one.

Topic two, fast Fourier transform. Wow. What's the good way to-- I mean, any decent machine comes with the FFT hardwired in. So you might say, OK I'll just use it. And that seems totally reasonable to me. But you might like to see just on one board, what's the key idea? What's the little bit of algebra that makes it work? So I'll just have one board here for the FFT and a little bit of algebra. Simple, simple but once it hit the world, well, computer scientists just love the recursion that comes in there. So they look for that in every possible other algorithm now. OK, let me see that point. So here's the main point. That if I want to take the-- multiply by F of size 1,024, the fast Fourier transform connects that full matrix to a half-full matrix. It connects that to the half-full matrix that takes the half-size transforms separately. So it's half-full because of these zeroes. That's the point. That the 1,024 matrix is connected to the 512 matrix. And what's underlying that? The 1,024 matrix is full of e to the 2pi*i, the w, over 1,024, right? That's the w for this guy. And then the w for this guy, for both of these, is e^(2pi*i/512). So if there's a connection between that matrix and this matrix, there'd better be a connection between that number and that number. Because this is the number that fills this one, and this is the number that fills these. So what's the connection? It's just perfect, right? If I take this number, which is one part of the whole circle, 1/1,024, a fraction of the whole circle, what do I do to get this guy? To get one 512th of the way around the circle? I square it. The square of this w is this w. Let me call this w N, and this one w M. Maybe I'll use caps, yeah I'm using cap N, this is my w. This is the one I want. The w N, the N by N one, N is 1,024, and the point is, everybody saw that, when I squared it I doubled the angle? When I doubled that angle the two over that gave me 1/512. Fantastic. Of course,

that doesn't make this equal to that, but it suggests that there is a close connection.

So let me finish here the key idea of the Fourier transform in block matrix form. What's the key idea? So instead of doing the big transform, the full size, I'm going to do two half-sizes. But what am I going to apply those to? Here's the trick. These two separate guys apply to the odd-numbered coefficients. The odd-numbered component. And the even. So I have to first do a little permutation, and even comes first, always. Even means zero, two, four, up to a 1,022. So this is zero, two; zero up to 1,022. And then come all the odd guys. One up to 1,023. So this is a permutation, a simple permutation. Just take your 1,024 numbers, pick out the even ones, put them on top. Right? In other words, put them on top where 512 is going to act on it. Put the odd ones at the bottom, the last 512, that guy will act on that. So there's 512 numbers there, with the even coefficients, this acts. OK, now we've got two half-size transforms. Because we're applying this to the y, to a to a typical y. OK. But I've just written F without a y so I don't really need a y here. This is a matrix identity. It's a matrix identity that's got a bunch of zeroes there. Of course, that matrix is full of zeroes. I mean, this is instant speed to do that permutation. Grab the evens, put them in front of the odds.

OK, so now I've got two half-sizes, but then I have to put them back together to get the full-size matrix. And that is also a matrix. Turns out to be a diagonal there, and a minus the diagonal goes there. So actually, that looks great too, right? Full of zeroes, the identity. No multiplications whatever. Well, these are the only multiplications, sometimes they're called twiddle factors, give it a fancy name. Official sounding name, twiddle factors. OK, so that diagonal matrix D, what's that? That diagonal matrix D happens to be just the powers of w, sit along D. 1, w, up to w to the, this is w N, we're talking, the big w, and it's only half-size so it goes up to M-1. Half, M is half of N. Everybody's got that, right? M is half of N here. I'll get that written on the board. M is 512, N is 1,024, here we have the powers of this guy up to 511. The total size being 512 because that's a 512 by 512 matrix. I guess I can remember somewhere, being at a conference, this was probably soon after the FFT became sort of famous. And then somebody who was just presenting the idea and as soon as it was presented that way, I was happy. I guess. I thought OK, there you see the idea. Permutation, reorder the even-odd. Two half-size transforms, put them back together. And what's happened here? The work of this matrix, multiplying by this matrix, which would be 1,024 squared is now practically cut in half. Because this is nothing. And we have just this diagonal multiplication to do. and of course this is the same as that, just with minus signs. So the total multiplications we have to do, the total number of twiddle factors, is just 512, and then we're golden. So we've got half the work plus 512, 512 operations. That's pretty good. And of course it gets better.

How? Once you have the idea of getting down to 512, what are you going to do now? This is the computer scientist's favorite idea. Do it again. That's what it comes to. Whatever worked for 1,024 to get to 512 is going to work. So now I'll split this up into, well, each 512, so now I have to do, yeah. Let me write F_512 will be, it's now

smaller. An I, an I, and a D and a minus D for the 512 size, of F_256, 256 and then the permutation, the odd-even permutation P. So we're doing this idea in there, and in there. So it's just recursive. Recursive. And now, if we go all the way, so you see why I keep taking powers of two. It's natural to have powers of two. Two or three. Three is also good. I mean, all this gets so optimized that powers of two or three are pretty good. And you just use the same idea. There'd be a similar idea here for, if I was doing instead of odd-even, even-odd I was doing maybe three groups. But stick with two, that's fine. Then you might ask, if you were a worrier I guess you might ask what if it's not a power of two. I think you just add in zeroes. Just pad it out to be the next power of two. Nothing difficult there. I think that's right. Hope that's right. And once this idea came out, of course, people started looking. What if the number here was prime? And found another neat bit of algebra that worked OK for prime numbers. Using a little bit of number theory. But the ultimate winner was this one. So maybe I'll just refer you to those pages in the book, where you'll spot this matrix equality. And then next to it is the algebra that you have to do to check it. I can just say, because I want you to look to the right spot there, maybe I'll take out the great-- This is sometimes called after Parseval, or some other person. Yeah, the algebra--

Let me start the algebra that made this thing work. We want the sum, when we multiply Fourier, F times something, we want the sum of w^(ik), right? That's the coefficient, that's the entry of F. Times c k. Summed from k=0 to N-1. To 1,023. That's the y i that we're trying to compute. We're computing 1,024 y's from 1,024 c's by adding up the Fourier series when we multiply by F. This is F, this is the equation y=Fc written with subscripts. So this is what the matrices are doing, and now where do I find that 512 thing? How do I get M into the picture, remembering that the w_N squared was w_M, right? That's what we saw. This is the big number, this is half of it, so this is a little bit of the part around the circle. When I go twice as far I get to the other one. So that's the thing that we've got to use. Everything is going to depend on that. So this was w N, of course. This is the N by N transform. So now comes what, what's the key idea? The key idea is split into even and odd. And then use this. So split into the even ones and the odd ones. So I write this as two separate sums, a sum for the even ones. w N, now, so now the even c k, so I'm going to multiply by c (2k). These are the ones with even, and the sum is only going to go from zero to M-1, right? This is only half of the terms. And then look on the other half, plus the same sum of-- But I didn't finish here. Let me finish. So I'm just picking out, I'm taking, instead of k I'm doing 2k. So I have j times 2k here. And now these will be the odd ones. c_(2k+1), and omega N to the j(2k+1). It's a lot to ask you. To focus on this bit of algebra. I hope you're going to go away feeling, well it's pretty darn simple. I mean there'll be a lot of indices, and if I push it all the way through there'll be a few more.

But the point is, it's pretty darn simple. For example, this term. What have I got there? Look at that term, that's beautiful. That's w_N squared. What is w_N squared? It's w_M. So instead of w_N squared, I'm going to replace that w_N squared by w_M. And the sum goes from zero to M-1, and what does that represent? That represents the F_512 multiplication, the half-size transform. It goes halfway, it operates on the even ones, and it uses the M.

It's just perfectly, so this is nothing but the Fourier matrix, the M by M Fourier matrix, acting on the even c's. That's what that is. And that's why we get these identities. That's why we get these identities, because they're acting on the even-- The top half is the even guy. OK, this is almost as good. This is the odd c's. Here I have, now do I have w_M here? I've got to have w_M. Well, you can see. It's not quite coming out right, and that's the twiddle factor. I have to take out a w_N to the power j to make things good. And when I take out that w_N to the power j, that's what goes in the D, in the diagonal matrix. Yeah. I won't go more than that. You couldn't, there's no reason to do everything here on the board when the main point is there. And then the point was recursion and then, oh, let me complete the recursion. So I recurse down to 256, then 128, then 64. And what do I get in the end? What do I get altogether, once I've got all the way down to size two? I have a whole lot of these factors down in the middle, the part that used to be hard is now down to F_2 or F_1 or something. So let's say F_1, one by one, just the identity. So I go all the way. Ten steps from 1024, 512, 256, every time I get twiddle factors. Every time I get P's. A lot of P's, but the F_512, what used to be the hard part, has gone to the easy part. And then what do I have? I've got just a permutation there. And this is the actual work.

This is the only work left, is the matrices like this, for different sizes. And I have to do those. I have to do all those twiddle factors. So how many matrices are there, there? It's the log, right? Every time I divided by two. So if I started at 1,024, I do ten times, I have ten of those matrices and have me down to N=1. So I've got ten of these, and each one takes 1,024 or maybe only half of that. Actually, only half because this is a copy of that. I think the final count, and can I just put it here, this is the great number, is each of these took N multiplications. But there were only log to the base two-- Oh, no. Each of them took half of N multiplications, because the D and the minus D are just, I don't have to repeat. So half N for each factor and the number of factors is log to the base two of N. Ten, for 1,024. So that's the magic of the FFT. OK. It's almost all on one board, one and a half boards. To tell you the key point, odds and evens, recursion, twiddle factors, getting down to the point where you only have twiddle factors left and then those multiplications are only N log N. Good? Yes. Right, OK.

Now, that's discrete transform. The theory behind it and the fantastic algorithm that executes it. Are you ready for a convolution? Can we start on a topic that's really quite nice, and then Friday will be the focus on convolution. Friday will certainly be all convolution day. But maybe it's not a bad idea to see now, what's the question? Let me ask that question. OK, convolution. So we're into the next section of the book, Section 4.4, it must be. And let me do it first for a Fourier series. I have convolution of series, convolution of discrete. Convolution of integrals, but we haven't got there yet. So I'll do this one, this series. So let me start with a couple of series. f(x) is the sum of c_k*e^(ikx), say. g(x) is the sum of some other coefficients. And I'm going to ask you a simple question. What are the Fourier coefficients of f times g? If I multiply those functions, equals something. And let me call those coefficients, h maybe. h_k*e^(ikx), and my question is what are the coefficients h_k of f times g? That's the question that convolution answers. Actually, both this series and the discrete series are highly interesting. Highly

interesting. So here I wrote it for this series. If I write it for the discrete ones, you'll see, so let me do it over here for the discrete one. Because I can write it out for the discrete ones. My y's are c_0 plus c_1 e t-- No, w. I have this nice notation, w. plus c_(N-1)*w^(N-1), right? That's the-- Ooh. What's that? I haven't got that right. Yes, what do I want now? I need, yep. Sorry, I'm looking, really I'm looking at y_j, the j-th component of y. So I need a w^j, w^(j(N-1).) Yeah, OK, let me-- OK. Alright. And so that's my f. I'll come back to that, let me stay with this.

I'll stay with this to make the main point, and then Friday we'll see it in a neat way for the discrete one. So I'm coming back to this. f has its Fourier series. g has its Fourier series. I multiply. What happens when I multiply this times this? I'm not going to integrate. I mean, when I do this multiplication, I'm going to get a mass of terms. A real lot of terms. And I'm not going to integrate them away. So they're all there. So what am I asking? I'm asking to pick out all the terms that have the same exponential with them. Like, what's h 0? Yes, tell me what h 0 is? If you can pick out h_0 here, you'll get the idea of convolution. What's the constant term if I multiply this mess times this mess, and I look for the constant term, h 0, where do I get the constant terms when I multiply that by that? Just think about that. Where do I get a constant, without any k, without an e^(ikx)? If I multiply that by that. Well tell me one place I get something. c 0 times? d 0. Good. Is that the end of the story? No. If you thought that multiplying the functions, I just multiplied the Fourier coefficients, the first point is no. There's more stuff. Where else do I get a constant out of this? Just look at it, do that multiplication and ask yourself where's the constant. Another one, yep. You were going to say it is? c 1 times d (-1). Right. Right. c 1 times d (-1). And tell me all of them, now. c 2 times d (-2). And what about c (-1)? There's a c (-1). It multiplies d 1. And onwards. So the coefficient comes from, now how could you describe that? I guess I'll describe it as, I'll need a sum to multiply. This will be the sum of c k times d what? Minus k, right? That's what you told me, the piece at the start, and that's the pattern that keeps going. OK, that's h 0, the sum of c k times d (-k). Now, we have just time to do the next one. We've got time but not space, where the heck am I going to put it? I want to do h k, I guess. Or I better use a different letter h I, let me use the letter h I, and God there's no space. Alright, so can I-- Yes. h I. OK. So this was h 0, let me keep things sort of looking right for the moment. OK, now you're going to fix h I. So what does c 0 multiply, if I'm looking for h_I, I'm looking for the coefficient of e^(ilx). So ask yourself how do I get e^(ilx) when that multiplies that? When that multiplies that, and I'm looking for an $e^{(ix)}$, I get one when c 0 all multiplies what? This is it. d I, right. And what about for c 1? Think of here, I have a c 1*e^(i1x). What does it multiply down here to get the exponential to be I? ilx? It doesn't multiply d (-1). It multiplies, c 1 multiplies? d (I-1). Good, good. I-1, right. I-1, and what are you noticing here? c minus, I'll have to fill that in. But you're seeing the pattern here? And what was the pattern here? Those numbers added to this number. And now these numbers add to I. Those numbers add to I, whatever it is, the two indices have to add to I, so that when I multiply the exponential they'll add to e^(ilx). They'll multiply to e^(ilx). So what goes there? It's probably I+1, right? So that I+1 combined with minus one gives me the I. If you tell me what goes there, I'm a happy person. Let's make it h I. We're ready for the final formula for convolutions. Big star. To find h I, the coefficient of e^(ilx), when you multiply that by that, you look at c k, and

which d is going to show up in the e^(ilx) term? I-k, is that what you said? I hope, yeah. I-k. Right, that's it. That's it. So we've got a lot of computation here. But we've got the idea of what, we've got a formula. And most of all we have the magic rule. In convolutions, convolutions are, things multiply but indices add. Things multiply, numbers multiply, while their indices add. That's the key idea of convolution that we'll see clearly and completely on Friday. OK.