[SQUEAKING]

[RUSTLING]

[CLICKING]

CASEY RODRIGUEZ: OK, so let's continue with our discussion of Fourier series from last time. For a given function in L2, we define the Fourier coefficient f hat of n to be the integral of 1 over 2 pi minus pi to pi f of t e to the minus int dt, which up to a factor of 1 over root t, is equal to the inner product of f with e to the int over square root of 2 pi in the Hilbert space L2, OK? And the question that we had-- so let me-- and we also had that the n-th partial sum for the Fourier series associated to f was given by sum from n equals minus n to n f hat of n e to the inx.

And the question which we're trying to resolve is do we have for all f and L2 limit as n goes to infinity of f minus SN of f 2 equals 0, OK? All right? In other words, is f equal to its Fourier series, at least when we interpret equals as in this sense here? Now, based on what we've done for Hilbert spaces, this question is equivalent to the following statement.

If f is an L2, and the Fourier coefficients are all 0, does this imply f equals 0, right? So this question is-- by what we've done for Hilbert spaces, big L2 is a Hilbert space. This question here is equivalent to this statement, which is that the collection of orthonormal vectors and big L2 consisting of the exponentials divided by square root of 2-- is this a maximal orthonormal subset?

Or as we were using the terminology we had from last time, does that form an orthonormal basis? So this is the statement that we're going to prove this class. And we're going to proceed via Fejer's method, if you like to give it a name, where what we did last time was-- when we recall, we had the Cesaro Fourier mean we defined to be the average of the first n partial sums with the hope that somehow this behaves a little bit better than the partial sums, because that's the thing we're trying to study.

And that's a hard question. And typically, means of sequences might behave better than the sequences themselves. But if the original sequence converges, then the means converge. So we should expect this to converge to f but hopefully faster or have better, more recognizable properties than just studying the partial sums directly.

And we'll get to-- in the next statement, it's a little bit clearer why the Cesaro Fourier means converge to f. And so what our goal-- what we're going to show is that if f is an L2, then the Cesaro Fourier means converts to 0, I mean converts to f as n goes to infinity. OK?

And so once we've proven that, then that gives us what we want in the yellow box, right? That proves what's in the yellow box, because let's take f and L2 with 48 coefficients all 0. Then, all of the partial sums will be 0.

Then, all of the means will be 0. And since the means converge to f, that proves f is 0. And we get what's in the yellow box. And therefore, the partial the Fourier sums converts to f as capital N goes to infinity in L2, all right? And then, once we prove that, I'll make a couple of comments about other types of questions one can ask and what you can do, or a brief comment.

So this is our goal for this lecture. And we should be able to get through it. So let me first rewrite the Cesaro Fourier means slightly differently. How we did in the previous lecture for the partial Fourier sums-- we wrote them as what's called a convolution. I haven't defined convolution-- but an integral of a function depending on x and t times f of t dt.

And we're going to do the same now for the Cesaro means. And we'll see here in what-- it's a little bit more clear, although I didn't talk so much about the Dirichlet kernel that appears for these guys-- but why the Cesaro means converge to f. OK, all right, so the statement is for all f in L2 minus pi to pi, we have that the n-th Cesaro mean of f, which is our Fourier mean, I can write as the integral from minus pi to pi of a function kn of x minus t times f of t dt.

So remember, for the partial sums s sub n, we could write it as d sub n, where d denoted a Dirichlet kernel, where kn of x-- this is equal to n plus 1 over 2 pi, and then 1 over 2 pi times n plus 1 times sine n plus 1 over 2x over sine x over 2 squared. And this holds that x equals 0. This one holds at x0 equal to 0. OK?

And this thing, we call Fejer's or the Fejer kernel. So all right? And now, let me just list off a few properties that we'll get from this.

Moreover, we have the following properties. 1, kn is non-negative. kn of x is equal to kn of minus x. It's even. And kn is 2 pi periodic.

The second is that the integral of k sub n of x from minus pi to pi-- dx, or let's make this t-- dt equals 1. And the third is if delta is a positive number less than pi, less than or equal to pi, then for all x with absolute value bigger than or equal to delta and less than or equal to pi, we have that kn of x, which is equal-- I don't need the absolute values, because it's non-negative-- is less than or equal to 1 over 2 pi over n plus 1 times sine squared delta over 2, OK?

So OK. So let's prove this theorem. And then, I'm going to say a few comments about-- well, since I have these properties right here, let me go ahead and make a few comments before we prove it. What does that mean kn looks like?

Let me draw 0 pi minus pi. So kn is non-negative. It's even. And away from a small neighborhood, it's quite small if capital N is very big. So what it's looking like is maybe the first one-- and it's large at the origin.

OK, so maybe that's n equals let's say 1. And then, let's say this is delta, and then minus delta. If I were to now look at, let's say, n equals-- I don't know-- a billion, it looks more like something that's very concentrated at the origin, but in such a way that the area underneath the graph-- so the integral-- the area equals 1, OK? And the same with what I drew in white, because white was supposed to be n equals 1.

Yellow was supposed to be n equals-- I don't know-- 1,000. The area is always 1, OK? So this is telling you that if I look at sigma n of f-- so this is just some remarks. This is not to be taken completely literally.

This is just the intuition on why we believe that the Cesaro means converge to f. And I'll say how this picture differs from if we looked at just SN. So this means that sigma n of f is, in fact-- so remember, we're going to get, in the end, that this is equal to kn of x minus t f of t dt. Now, kn is very concentrated near where t equals x, OK?

So based on the picture, as n gets very large, this thing is getting more and more concentrated near where n equals x, OK? Now, and therefore, at least for let's say very nice f, if this thing is concentrated near where t equals x, then f of t will be approximately f of x. So f of x comes out of the integral because this is an integral dt.

So since this thing is concentrated at-- and because the area underneath the curve is always 1, this integral is always equal to the integral of kn over any. So kn is 2 pi periodic. This integral is equal to the same integral over any 2 pi periodic interval, which means I could put here-- I could add an x to both top and bottom, and therefore change variables to get this is kn of t dt, which equals-- because the integral is 1, I would get something like f of x, OK?

So this is a heuristic reason on why one should expect the Cesaro means to converge to f. OK? If you look back at the kernel that we had for the partial sums, it had some of the same-ish properties. It was 2 pi periodic, and also even.

The integral was 1. And it did decay away from 0. However, it's non-negative. I'm talking about the Dirichlet kernel dn, which if you look back in your notes, was sine of plus n plus 1/2 times x over sine x over 2 with a constant out in front. And that little difference, the fact that this kernel is non-negative-- and the Dirichlet kernel is not-- makes a big difference.

So although this heuristic argument-- maybe you don't see it there-- in the actual proof itself, that oscillation-- and what I mean by isolation is the fact that dn actually does oscillate between negative and positive values-- this bit of oscillation is actually what you can use to build up a continuous function whose partial sums do not converge to that continuous function at a point, OK? But as we'll see for the Cesaro means, the Cesaro Fourier means, basically, pick a space. And the Cesaro sums or Cesaro means converge to the function in whatever space you're talking about.

And I'll say a little bit more about that in a minute. But OK. So let's prove the theorem that the Cesaro means are written in this way, and the kernel has these three properties. So let me recall that we have SN of x-- or let's put a k there-- this is equal to, as we wrote last time, minus pi to pi DK of x minus t f of t dt, where DK of, say, t was, from last time, equal to 2n plus 1 over 2 pi at t equals 0 and sine n plus 1/2 t over sine t over 2, and then with 1 over 2 pi out in front, I believe. Let me make sure I got the right exponent. Right.

For t not equal to 0. OK? Oh, and this should be k. OK, so using this, we have that the Cesaro sum of x-- this is equal to 1 over n plus 1, sum from k equals 0 to n, the mean of the first and partial sums. And this is equal to-- now, sk f of x is equal to this. So I can write this as integral from minus pi to pi of 1 over n plus 1 sum from k equals 0 to n of DK x minus t f of t dt.

And so this here is kn of x minus t. All right? So now, I'm just going to verify that kn of x takes that form that we had before. And kn of x-- this is equal to 1 over n plus 1 sum from k equals 0 to n of DK of x.

And let's go to the next half board. So I can write this as 1 over 2 pi n plus 1 and times-- so I will look at the case that x is non-0. x equals 0 is, you'll get what you get. But let's look at x not equal to 0.

So then, I plug in this formula here and pull out a sine t over 2-- or sine x over 2 squared on the bottom. And then, I get k equals 0 to n of sine x over 2 times sine n plus 1/2 x, OK? And because I feel like it, let me put a 2 here and a 2 here.

Why do I feel like it? Well, it's because if I have 2 times sine of a sine b, I can write that as using my angle sum formulas from trigonometry. You wondered why those would be useful. Well, here they are appearing in the advanced MIT class.

You can write this as sum from k equals 0 to n of cosine n x minus cosine n plus 1x. Let me make sure I got that right. Or this should be k. I'm sorry. That should have been k.

k, k, all right? Now this, is a telescoping sum, right? I have a sum of cosine kx. I have a cosine k plus 1x.

So this is equal to-- so let's just write this out. And let me just indicate why this is a telescoping sum. We get cosine 0x minus cosine 1x plus cosine 1x minus cosine 2x dot dot dot plus the last one, which is cosine nx minus cosine n plus 1x. And OK, so this telescopes.

That cancels with this. That will cancel with so on. And that last one will cancel. So all that we're left with is this one minus this one divided by this 2 that I have right there.

And I get 1 over 2 pi n plus 1 times 1 over sine squared x over 2 times 1 minus cosine n plus 1 x over 2. And again, using a trig formula-- 1 minus cosine 2a equals sine squared-- divided by 2 is equal to sine squared a. So I get this is equal to 1 over 2 pi n plus 1 times sine squared n plus 1 over 2x divided by sine squared x over 2, OK?

So that verifies the formula for the Fejer kernel. What about the properties that we have there? These properties-- at least the first two-- follow directly from this formula and the definition. So 1, follows immediately.

This is clearly non-negative. It's even, taking x to minus x does not change this, because we have squares. And also because of the squares, it's 2 pi periodic rather than 4 pi periodic, OK? OK, so that's 1.

For 2, we note that if we take the integral for minus pi to pi of the Dirichlet kernel, this is-- OK, we had a formula for the Dirichlet kernel, but remember, this is nothing but-- this was defined to be the sum from n equals minus k to k of e to the int dt, OK? Now, e to the int when n is not equal to 0 is 2 pi periodic. And when I integrate it from minus pi to pi, the integral from minus pi to pi of this 2 pi periodic thing-- you can just check.

It's the integral of sine, nt, and cosine nt over its period. That's going to give me 0. So all I pick up is when n equals 0, right?

And so that's equal to just the n equals 0 term. So that gives me 1, OK? So since the integral of each kernel is 1, then the integral of the Fejer kernel-- which remember, this is equal to the average of the Dirichlet kernels. And each of these is 1 sum from k equals 0 to n 1. I get n plus 1 divided by n plus 1.

I get 1, OK? So that gives me 2. And for the third property, we have-- what do we have? Then, the function sine squared x over 2-- what does it look like? This is increasing. Or I should say it's even and increasing on 0 to pi.

So what it looks like is sine squared x over 2. So there's pi minus pi sine squared. Looks like it goes up to 1. So if I'm looking at all x outside of-- so in that shaded region-- then, if x is outside of this delta region, then I get that sine squared x over 2 is going to be bigger than or equal to whatever, so it sits above the value that I get here, which is sine squared delta over 2.

And therefore, I get that kn of x, which is equal to its absolute value, is less than or equal to 1 over 2 pi n plus 1 sine squared n plus 1 over 2x over-- I had sine squared x over 2, but since sine squared x over 2 is bigger than or equal to sine squared delta over 2, taking 1 over reverses the inequalities. And I get sine squared delta over 2 here. Sine of anything is always bounded above by 1. So I get this is less than or equal to 1 over 2 pi n plus 1 sine squared delta over 2, OK?

So for the moment, let me put this absolute value there. I'm not doing it because I think it looks better. It's because I'm going to make a comment in a minute. OK, let me just make a small comment.

Well, let me prove the next theorem. And then, I'll make the comment. OK, so we have these properties of the Fejer kernel. And now, what we're going to do is on our way to proving that we have convergence of the Cesaro means to a function in L2, we're first going to do it for continuous function.

So you proved in the assignments that in L2 minus pi to pi, the continuous functions vanishing at the two endpoints are dense in the space big L2, OK? Now, if a function's continuous and equals 0 at both of the endpoints, it's 2 pi periodic in the sense that it has the same value at both endpoints. And therefore, the subspace of continuous functions that are 2 pi periodic is dense in L2.

So if we're going to be able to show that the Cesaro means converge to a function in L2 for arbitrary L2 function, maybe it makes sense to try and do it first for continuous functions. And it's there that this argument that I just-- this heuristic argument I gave here will be more math-like. OK, so we have a following theorem due to Fejer, which is the following.

If f is continuous and 2 pi periodic, meaning f of pi is equal to f minus pi, then not only do we have the Cesaro means converging to f in L2, we actually have it in the best sense that you could for a continuous function. Then, sigma n of f converges to f uniformly in minus pi to pi, all right? So before, we were looking at Fourier series in L2.

So convergence in L2 was the way one makes sense of infinite series or something converging to something else, all right? If we're looking at continuous functions, then we already a different norm there if we want to just consider a complete space containing continuous functions. We have the uniform norm, or the infinity norm.

And so what this says is that even in this smaller space and in this stronger norm, we have convergence of the Cesaro means to the function f. But again, this doesn't imply that the Fourier series converges to f uniformly. Like I said, one can, in fact, use this oscillatory behavior of the Dirichlet kernel to prove there exist continuous functions whose Fourier series diverges at a point.

And therefore, it doesn't converge uniformly to the function. But this is true for the Cesaro means because of these properties of the Fejer kernel, because it has this shape where it's non-negative. It's peaking near the origin. And it has total mass 1, and total integral 1.

In some sense, you should think of, as n goes to infinity, sigma n is looking more and more like the Dirac delta function at 0, which maybe you encountered in physics. If that doesn't mean anything, don't worry about it. Just skip to the next part of the talk, which is supposed to have this magical property that it's 0 away from 0, which these are looking like, as integral 1.

And when you integrate it against a function, you get f evaluated at the origin, which is like what we're saying here, OK? So again, that's some more heuristics. But linear operators depending on a parameter that appear like this, where it's a function of this form times f of t integrated dt, pop up all the time in harmonic analysis, OK? And having these properties, in fact, pops up also in harmonic analysis, OK?

Harmonic analysis being a fancy name for Fourier analysis and other stuff. So let's prove this. So the first thing that I want to do is-- so f is a continuous function on minus pi to pi. That's 2 pi periodic.

So I can extend f to all of R by periodicity. In other words, so we extend to all of R, meaning I have-- so there's pi minus pi pi. Here's a 3 pi.

Here's minus 3 pi. So supposedly, I have this continuous function, which is 2 pi periodic. Now, I take that continuous function and just extend it by how it is here and so on, OK? I'm not saying I extend it by 0 outside.

I'm saying I extend it periodically, OK? OK, now, I can write down a formula for exactly how you do that. But just trust me. You can do that.

And also the following simple properties, then-- f, now referring to it as a function defined on all of R that's 2 pi periodic, this is also continuous, is 2 pi periodic, which implies that f is uniformly continuous and bounded, i.e. If I look at the infinity norm of f first off, because by periodicity, this is just equal to sup xn minus pi to pi, and because f is continuous, this thing is finite, OK? All right.

Now, it's not difficult to believe that f is-- if I extend it by periodicity, it's going to be continuous. But using that and the fact that it's 2 pi periodic, you can then also conclude that it's uniformly continuous, meaning-- let's just quickly review what uniformly continuous means. This means for all epsilon positive, there exists a delta positive such that if y minus z is less than delta, then f of y minus f of z is less than epsilon, meaning I can choose a delta independent of y and of the point, right?

Continuous at a point means I fix x. Then, for all epsilon, there exists a delta. Uniformly continuous means the delta doesn't depend on x, the point that I'm looking at. All right, so we have basic observation that we're going to make there.

And maybe I'll just leave this up for now. So we want to prove the sigma n's converge to f uniformly on minus pi to pi. So that means we should be able to find, for every epsilon, a capital M such that for all n bigger than or equal to M sigma int and for all x in minus pi to pi sigma n of f minus f is less than epsilon in absolute value.

All right, so let epsilon be positive. Since f is uniformly continuous, as I stated-- recalling the definition-- that implies that there exists a delta positive such that if y minus z is less than delta, then f of y minus f of z is less than-- and let me get this right so it comes out pretty in the end-- is less than epsilon over 2, OK? So now, what we're going to go through is make that argument which I just erased actually precise, all right?

So here, we're saying if f is very close to-- if any two points are sufficiently close, f is going to be close in value. OK, now choose M natural number so that for all n bigger than or equal to M, the quantity twice times the L infinity norm over n plus 1 times sine squared delta over 2 is less than epsilon over 2, OK? So n plus 1-- that's the thing that's changing. So I have these fixed numbers here now.

I've fixed delta. I have the L infinity norm of f. So I have this number here. And I'm just saying, choose a capital M so that for all n bigger than or equal to M, this number of times-- and I'll even put it here-- times 1 over n plus 1 is small, is less than epsilon over 2. And I can do that because this, as capital N goes to infinity, converges to 0, right?

OK. Now, since f and k sub n, the Dirichlet kernel, are 2 pi periodic, I can write the Cesaro mean, which is given by minus pi to pi kn of x minus t f of t dt. I can make a change of variables, set tau equal to x minus t. And then, this will be equal to-- what is it going to be equal to?

x minus pi x plus pi kn of tau f of x minus tau d tau, OK? All of this change of variable stuff is fine, because I'm dealing with continuous functions. I'm integrating continuous functions. So that's the Riemann integral.

We have a change of variables for the Riemann integral. So that's completely fine. OK, now this is the product of 2 pi periodic functions. And if I take the integral of that quantity of a 2 pi periodic function, the integral of that is equal to the integral over any interval of length 2 pi, all right?

So we're integrating over an interval of length 2 pi, right? We're going from x minus pi to x plus pi. That is equal to the integral of the same quantity over any interval of length 2 pi. So it's also equal to the integral over minus pi to pi.

OK? So all I'm saying is I can change variables and move the x minus t. And let me even go back to t instead of tau here. Because of periodicity, I can switch this x minus t over here to f. All right, now we're going to start seeing some magic happen.

And this is where that heuristic argument that I gave earlier actually starts to make sense. So then, I have that for all n bigger than or equal to M-- so I have that condition that quantity was less than epsilon over 2. And for all x and minus pi to pi, I have that sigma n f of x minus f of x-- so this is equal to minus pi to pi.

And again, I'm going to write this now as kn of t f of x minus t dt minus-- now, here's the trick. The Fejer kernel has integral 1. So I can actually write f of x as minus pi to pi integral kn of t times f of x dt. I'm integrating dt, right?

Then this just pops out. I get f of x times the integral of the Fejer kernel, which is 1, OK? And this equals minus pi pi. So just combining things-- kn of t f of x minus t minus f of x, which is good, because we have a continuous function. And now, we have something inside that looks like I'm subtracting f of some argument minus f of the argument minus something, OK?

Now, I'm going to split this integral into two parts, and then use the triangle inequality and bring the triangle inequality inside. In fact, I'm going to go ahead and do that here. This is less than or equal to if I combine terms like I did and then bring the absolute value inside. OK? And now, I'm going to split this integral up into two parts.

This is equal to the integral over t less than delta. And because kn is non-negative, this is just kn of t f of x minus t minus f of x dt plus the other term. OK? kn of t f of x minus t minus f of t dt, all right?

Now, what do we know? If the absolute value of t is less than delta, then x minus t minus x is equal to minus t, which is an absolute value less than delta. So note that x minus t minus x equals t is less than delta here, right? And therefore, this quantity here is less than epsilon over 2 by how we chose delta.

So this is less than epsilon over 2 times the integral over this region of kn of t. OK? But I can make this region larger and just go back to-- so let me just leave it here as it is. Plus now what do I do with this piece? I have this.

I bound by twice the L infinity norm of f. The absolute value of this is less than or equal to by the triangle inequality, the sum of the absolute values, which is less than or equal to the sup of this plus the sup of that and x, which is equal to twice the infinity norm. So I get 2 times the infinity norm of f popping out from this term, and kn of t-- oh, I'm away from t less than delta.

And this is where I use that third property that I have from before, that it's less than 1 over 2 pi. So let me leave this here. Sum 1 over 2 pi n plus 1, sine squared delta over 2 dt, OK? And now, this, I can say, is less than or equal to the whole integral over minus pi to pi, which is equal to 1. Plus again, making this an integral over the entire region, I get 2 pi times-- or divided by 2 pi gives me 1.

So I get twice infinity over n plus 1 sine squared delta over 2. And we chose n plus 1 so that this second quantity here is less than epsilon over 2. OK? And therefore, uniformly, we prove that for all capital N bigger than or equal to M for all x in minus pi to pi, the difference in sigma nf in f is less than epsilon, proving uniform convergence, OK? So here's the remark I was going to make, is that the same proof can be modified if instead of kn of x being bigger than or equal to 0-- let me make sure I'm saying the right thing.

So if instead of this property, which we had for the Fejer kernel, we have that sup over n of the integral from minus pi to pi of kn of x is finite, meaning if I have a function or if I have a sequence of functions, kn's, and I have the corresponding operators that look like that-- maybe they're not associated to any questions about Fourier analysis, but I'm just saying-- and it satisfies the three properties I had before with the exception of being non-negative, but instead of that, it satisfies this property, then I can do redo the same proof and show that those things converge to f uniformly, OK?

Why am I saying that? Because maybe you would like to then try your hand at replacing kn with dn, the Dirichlet kernel, OK? The Dirichlet kernel satisfies all of the other properties we had up there. The integral is 1.

In absolute value, it decays away from x is less than delta. And it's even in 2 pi periodic, OK? But it doesn't satisfy this. And if I look at minus pi to pi of the Dirichlet kernel, what one can prove is that this is something like log n for large enough n, OK?

All right? So that was just a tiny remark I wanted to say on why, if you thought about maybe redoing this proof using the Dirichlet kernel, which satisfies almost all the same properties with the exception of being non-negative, you could, if the Dirichlet kernel had satisfied this bound. But it doesn't. It satisfies this bound.

It's like log n. And therefore, if I take the sup, I don't get something finite, OK? All right, so we've proven that the Cesaro means of a continuous function convert uniformly to a continuous function.

So we're almost to the point where we can say that the Cesaro means of an L2 to function converge to an L2 function and conclude that the subset of exponentials divided by square root of 2 pi form a maximal orthonormal subset of L2, and therefore is in orthonormal basis so that the partial Fourier sums converge back to the function in L2. We just need one more bit of information. So we have the following theorem.

For all f in L2 of minus pi to pi, if I look at sigma n of f-- so first off, this is just a finite linear combination of exponentials, right? So this is clearly an L2. It's a continuous function. But if I take the L2 norm of that, it could depend on n.

But in fact, it's less than or equal to the L2 norm of f. OK? So how we'll prove this is we'll first-- and this bound is what allows us to go from the 2 pi periodic continuous functions to general L2 functions by a density argument, OK? So first, we'll do this for 2 pi periodic continuous functions, and then by density, conclude it for L2 functions.

So suppose first that f is 2 pi periodic. And then of course, extend it to R by periodicity like we did before. Then, as before, we had that the Cesaro mean of f is equal to the integral from minus pi to pi of f of x minus t kn of t dt. And so if I compute the integral sigma n f of x squared dx, this is equal to-- so each one of these is equal to an integral over minus pi to pi.

So I'm going to have three integrals. And f of x minus s times the complex conjugate fx minus t times kn of s and kn of t, And Then ds dt dx, OK? Now, all these functions are continuous. So we have a Fubini's theorem, which says we can reverse the order of integration however we please. So I can write this as the integral for minus pi to pi, minus pi to pi, of now integrating first with respect to x-- kn t. And now, integrating first with respect to x. dx, let's say ds, dt, OK? Now I do Cauchy-Schwarz on this. And so this is less than or equal to minus pi to pi minus pi to pi pi kn of s kn of t times-- I'm using Cauchy-Schwarz in x now-- so times the L2 norm of the function minus s. So I'm taking the L2 norm in this variable-- 2 times 2 ds dt, OK?

What I mean by this is I'm taking the L2 norm of this function depending on s, But In the first variable-- in this x variable, OK? So just write it out to see what I mean. Now, this is the integral of a function over an interval of length 2 pi. That's 2 pi periodic. That's equal to the integral of that function over any 2 pi periodic interval or any interval of length 2 pi. So I can, in fact, remove this s and remove this t, and just pick up the L2 norm of f in both places.

So this is, in fact, equal to-- and because these two things no longer depend on s and t, they come all the way out of the integral. And I get L2 norm squared times minus pi to pi kn of s ds times the integral from minus pi to p kn of t dt. Both of these integrals equal 1.

So I get norm squared. And I started off with the L2 norm squared, or the L2 norm squared of the Cesaro mean of f. So I get this for all 2 pi periodic continuous functions, OK? Now, how do we then get the bound for general f?

We use the density argument. So by what you've done in the assignments, there exists a sequence of 2 pi-- so let me start over real quick. Now, let's take a general element in L2.

OK, now we start. By assignments, you know that there exists a sequence of 2 pi periodic continuous functions converging to f in L2-- fn a of 2 pi periodic continuous functions such that the fn's converge to f in L2. And one can verify simply from the definition of each of the Cesaro means that then-- so this is as little n goes to infinity-- that then the Cesaro means also converge as little n goes to infinity. So capital N here is fixed, OK?

Just using the definition of what the Cesaro mean is and Cauchy-Schwarz, basically, OK? And the fact that fn's converge to f in L2. Thus we get that the L2 norm of the Cesaro mean is equal to the limit, as n goes to infinity, of-- so this is little n-- of the L2 norm of the Cesaro means of these continuous 2 pi periodic functions, which as we've proven already-- these are all less than or equal to the L2 norm of fn, because they're 2 pi periodic. And again, because f is converging to fn, the norms converge.

And I get the result I wanted for general L2 functions. OK? So now, we're almost there. What we have is this bound. And we have that the Cesaro remains converge to-- so if I take the Cesaro means of a continuous function, those converge to the continuous function uniformly on the interval.

We're going to use that, this bound, and the density, again, of the 2 pi periodic continuous functions in L2 to conclude the following theorem. For all f in L2, the Cesaro means converge to f as capital N goes to infinity. In particular, we get, as an immediate corollary, if all of the Fourier coefficients are 0, then f is 0, right? Because if I've proven this and all the Fourier coefficients are 0, then the Cesaro means are all 0.

And therefore, since this is 0 converging to f, f must be 0. OK? And therefore, the set of exponentials-- normalized, of course-- form a maximal orthonormal subset of L2, i.e. that they're an orthonormal basis for big L2, which answers the question we had about Fourier series converging to a function in L2, OK?

All right. So we'll do this just as a standard epsilon n argument. Let epsilon-- so let f be in L2. Let epsilon be positive.

So we know that the continuous 2 pi periodic functions are dense in L2, because we did this in the assignment that for any f in L2 over an interval, I can find a continuous function that vanishes at the endpoints and therefore is periodic, which is close to f in L2. So there exists a g that's continuous 2 pi periodic such that f minus g in L2 norm is less than or epsilon over 3.

So since sigma N g converges to g uniformly on minus pi to pi, there exists a natural number M such that for all N bigger than or equal to M, for all x minus pi to pi, I have that sigma N g of x minus g of x is less than epsilon over 3 square root of 2 pi. OK? Now, we go about the part where we replace f with g, OK?

Then, for all N bigger than or equal to M, if I look at the L2 norm of sigma N of f minus f in L2, and I apply the triangle-- I add and subtract terms and apply the triangle inequality-- I get that this is less than or equal to sigma N of f minus g 2 plus sigma N of g minus g in L2 plus g minus f in L2, OK? So sigma N of f minus g is equal to sigma N of f minus sigma N of g.

So I use that there without explicitly stating that. So let me say sigma N of f minus sigma N of g. Just from the definition, you can check this is equal to sigma N of f minus g, OK?

Now, by the bound I just proved, the L2 norm of the Cesaro mean is less than or equal to the L2 norm of the function here. So this is less than or equal to f minus g2. And then, I also have this L2 norm of f minus G there.

So I'll put a 2 there plus-- and I'll actually write out what this is-- sigma N g of x minus g of x squared dx 1/2, OK? Now, f minus g is less than epsilon over 3 in L2 norm. So this is less than twice epsilon over 3.

Sigma N g minus g is less than epsilon over 3 square root of 2 pi here. So I get epsilon over 3-- that pulls all the way out-- minus pi pi 1 over 2 pi dx. And I just get epsilon in the end, OK?

OK. So that concludes what I wanted to do for Fourier series, at least for now, which applies what we've done for Lebesgue integration, these big LP spaces, and also some of this general machinery we've built up for Hilbert spaces to actually answer a more concrete question rather than just trying to prove general statements. General statements are very, very useful. I'm not saying they're not.

But I'm just saying so that you can see a concrete problem why one would want and use functional analysis in the first place. Now, coming back to what we've done so far, so let me just make a couple of remarks about what we haven't shown. It's a very deep theorem due to Carleson.

So what we've shown is that the partial sums-- so we showed the set of exponentials normalized, or a maximal orthonormal set-- I mean that they're orthonormal basis. So the partial sums converge to f in L2. So this is what we've shown.

For all f in L2, the partial sums converge to f in L2, all right? But this does not translate into a point-wise statement. This does not say that the partial sums converge to f almost everywhere. OK?

There is a general theorem one can say that is covered in more advanced measure theory classes where one can say that there exists a subsequence converging to f almost everywhere. But that's not very good, or at least very clean. Now, for a long time, it was not necessarily believed that the partial sums converged to f almost everywhere.

But a theorem due to Carleson shows that for all f in L2, partial sums do converge to the function almost everywhere, OK? This is, in fact-- maybe this is true. Maybe this is not. I heard this from my advisor.

Carleson spent a few decades trying to prove the negation of the statement, trying to come up with an example of a function whose partial sums converge don't converge almost everywhere back to the function. And then, he came up with the bright idea that, well, maybe that's not true.

Let me spend some time trying to put myself in the other shoes. And within a year or a couple of years, he was able to prove this theorem, OK? So this is Carleson's theorem that we do have convergence almost everywhere.

Now, you can also ask, what about convergence? So this convergence in L2 of the partial sums. We have other LP spaces, right?

What about in those LP spaces? Can I replace this 2 with p? The Fourier coefficients and partial sums-- these all make sense for any big LP space. So what is known is that also-- and now, the name is escaping me, but I'll just state it.

For all p between 1 and infinity, the partial sums converge to the function in lp. When p equals 1, this is false, OK? And when p equals infinity, this is also false, because the partial sums-- these are a finite linear combination of exponentials, and therefore continuous function, OK? So you can't have, for an arbitrary function in L infinity-- which can be discontinuous, just has to be bounded-- these converging to L in such a function.

Because then, the limit would have to be continuous, OK? The uniform limit of continuous functions, which L infinity kind of is, has to be continuous, OK? So that's why you wouldn't expect it for L infinity. And for what one would call duality, because infinity is the dual of L1, you also don't get p equals 1.

But in fact, things are worse there. You can come up with an L1 function. So that the Fourier series-- I don't think I'm lying when I say this, but-- diverges almost everywhere, I want to say, OK? I don't think I'm lying. But if p equals 1, one can come up with an example where the partial sums diverge point-wise almost everywhere, OK?

OK. But to prove this flavor of statements requires deeper harmonic analysis, harmonic analysis being the umbrella that Fourier analysis sits in, and requires a knowledge of, or at least working with certain operators, which are called singular integral operators, which were developed back in the last century, middle of the last century at the University of Chicago by my mathematical grandfather and great grandfather, which gives you some beautiful results about, again, convergence of Fourier series, but also some applications to PDEs, which were why they were originally created in the first place and so on.

But perhaps you'll encounter that if you take a class in harmonic analysis or Fourier series. I haven't taught the Fourier series class, so I don't know what it's about. But that kind of material will not be covered in this class. And this will be as far as we go as far as these types of questions, all right?

So next time, we'll move on to minimizers over closed convex sets and consequences of that, one being that we can identify-- which is the most important application-- we can identify the dual of a Hilbert space with the Hilbert space in a canonical way. You can already prove that if you wish using the fact that every Hilbert space is asymmetrically isomorphic to little l2. You know that the dual of little l2 is 1 over q, is lq, where 1 over 2 plus 1 over q equals 1. And therefore, q equals 2.

So little l2 is a dual of itself. But we'll prove it for general Hilbert spaces, which has some very important and interesting consequences when it comes to now studying, solving equations in Hilbert spaces, meaning you have linear operators. When can you solve equations involving these linear operators, and so on? All right, so we'll stop there.