

[SQUEAKING] [RUSTLING] [CLICKING]

PROFESSOR: OK, so ortho-- today, we're going to be discussing orthonormal bases of a Hilbert space. So let me just recall what we did at the end of last time. We introduced maximal orthonormal sets, so a collection of vectors in a-- we could say a pre-Hilbert space, but let's say in a Hilbert space.

So this is maximal if u in H and $\langle u, e_\lambda \rangle = 0$ for all λ in the index implies that $u = 0$. So a collection of-- so I should say that a collection of orthonormal-- orthonormal vectors is maximal if the only thing that's orthogonal to all of them is the zero vector.

And last time, we proved that-- let H be a separable Hilbert space, meaning it has a countable dense subset. Most-- I mean, I think all of the Hilbert spaces you've come into contact with are separable. C_n , R_n , also little ℓ_2 , big ℓ_2 -- this is what you're doing in the assignment. That's also separable-- then H has a countable maximal orthonormal subset, e_n .

So we proved this at the end of last time via the Gram-Schmidt process. We took a countable collection of-- a countable collection of elements in the Hilbert space, which were dense. And then we formed this countable orthonormal subset. And I shouldn't put N in, capital N because this could be a finite collection or it could be countably infinite collection. But then we applied the Gram-Schmidt process to that collection of dense elements and came up with this maximal orthonormal subset.

Now, we have a special name for countable maximal orthonormal subsets for a good reason as we'll see. So let H be a Hilbert space. An orthonormal basis of H is a countable maximal orthonormal subset-- I shouldn't say n and N again-- of H . So an orthonormal basis is just a special name that we give to those maximal orthonormal subsets which are countable, meaning they can be either finite or countably infinite.

So again, all of the Hilbert spaces we've encountered in practice-- for example, C_n , little ℓ_2 , big ℓ_2 , which are all separable, meaning they do have countable maximal orthonormal subset, meaning they have an orthonormal basis-- so they have these objects. Now, why do we call-- why do we call a countable orthonormal-- or a countable maximal orthonormal subset an orthonormal basis? So a basis is supposed to be something where you can write every vector in the space as some sort of linear combination of the elements of the basis, right?

Now, we dealt with Hamel basis-- or bases, which are you can write every element as a finite linear combination of the elements in this subset. Now, in this setting, in what sense is this a basis? Or what is-- I shouldn't say, is it a basis? But in what sense does this connect to what we've encountered, say, in finite dimensions?

An orthonormal basis-- so as stated in the next theorem, every element can be written now as an infinite linear combination of elements of an orthonormal basis, not necessarily finite. So the statement is the following-- if e_n is an orthonormal basis in the Hilbert space H , then for all u in H , if I look at the partial sum, $\langle \sum_{n=1}^N e_n, u \rangle$ inner product $\langle \sum_{n=1}^N e_n, u \rangle$, this converges as n goes to infinity to the element u in the space H .

Written in shorter form is that u is equal to sum from n equals 1 to infinity of u inner product $e_{\text{sub } n}$, $e_{\text{sub } n}$. Sometimes, this is referred to as a Fourier-Bessel series. So just like in finite dimensions, if you have an orthonormal basis consisting of only finitely many vectors, then you can expand every vector in that, say, C_n in this way. Now, the statement is, in a Hilbert space, if you have an orthonormal basis in this sense, meaning it's a countable maximal or orthonormal subset, then you can expand every element as an infinite linear combination of these elements.

OK, so to prove this, we will use Bessel's inequality and the completeness of H . This is where we're using the fact that H is not a pre-Hilbert space, but a Hilbert space, all right? What we're going to do is we're going to be able to show that the right-hand side of this equality-- so we're going to be able to show that this series here converges in H . That uses Bessel's inequality and the fact that H is a Hilbert space converges to some element. And then we're going to use the fact that this collection is maximal to show that the inner product of whatever this defines with each element $e_{\text{sub } n}$ is the same as the inner product of this thing with each element $e_{\text{sub } n}$, and then conclude that this thing on the right side has to equal this thing.

All right, so if we first prove that the sequence of partial sums-- so m is Cauchy. So let ϵ be positive. So by Bessel's inequality, we have that the sum of n equals 1 to infinity of u squared-- this thing converges. So the only thing we have to check is that, since this is a series involving non-negative terms, that this thing is bounded above. And we have by Bessel's inequality that this is always bounded above by the norm of u squared. So this sum, this series converges to a finite number.

In particular, the series-- the partial sums corresponding to this series is a Cauchy sequence of real numbers, of non-negative real numbers. Thus, there exists an N natural number such that for all-- let's see-- all N bigger than or equal to M , we have that the tail is small, meaning n equals N plus 1 infinity of less than ϵ . You can just look back into your 18.100 notes or whatever that if I have a series of non-negative terms which converges, then for every ϵ , I can find a capital N so that the tail, so starting the sum at a large enough entry, will be small.

Then for all m greater than l greater than this capital M , if I compute the norm of n equals 1 to m u inner product $e_{\text{sub } n}$, $e_{\text{sub } n}$ minus N equals 1 to l u $e_{\text{sub } n}$, $e_{\text{sub } n}$ squared. And let's make this a square. I can do that.

This you can just compute. Expanding this out as the inner product of this whole thing with itself, this gives you sum from n equals l plus 1 to m of u e_n squared. And since l is bigger than or equal to m , this is certainly less than or equal to-- well, I can put infinity here. n equals l plus 1 u inner product $e_{\text{sub } n}$ squared. And that's less than-- since l is bigger than or equal to capital M , we can use this inequality. And that's less than ϵ squared.

So we've shown that for n and l bigger than or equal to capital M , the norm squared of the difference is less than ϵ squared. And therefore, the norm of the difference is less than ϵ , proving that the sequence of partial sums is Cauchy.

OK, so we've proven that the sequence of partial sums is Cauchy. So since H is complete, there exists an element u in H such that u -bar is equal to limit as n goes to infinity equals 1 m of these partial sums, where this limit means that the norm of the difference in H converges to 0. And so just in shorthand, again, this means-- now, what I'd like to do is show that u -bar equals u . I don't mean complex conjugate. I just mean an element of H .

And now, how we're going to do that is show that the inner product of u minus $\bar{u}$ against every element of this maximal orthonormal subset is 0, and therefore conclude u equals $\bar{u}$. Now, we have this result, that the inner product is continuous with respect to the entry. So by continuity of the inner product, we have that the inner product of u -- so let me-- so then for all l natural number, the inner product of u minus $\bar{u}$ with e_l , this is equal to-- $\bar{u}$ is equal to the limit as m goes to infinity of this partial sum in H . So this inner product is equal to the limit as m goes to infinity of u minus sum from n equals 1 to m of u_n , e_n inner product e_l . And this is also-- I mean, I'm sure you've seen this at some point. But this is also why the coefficients appearing in front of the e_n 's are the way they are.

So this is equal to the limit as m goes to infinity of u inner product e_l minus the sum from n equals 1 to m of u_n , e_n inner product e_l . And when n does not equal l -- so remember, these are orthonormal. When n does not equal l , I get 0. And when n equals l , I get 1. So all I pick up is u -- this just reduces to u inner product e_l , which cancels with that one. And therefore, I get 0.

All right, so I've shown that for every l , u minus $\bar{u}$ inner product with e_l is 0. And since this collection is maximal, anything that inner product with-- so the only element that's orthogonal to every element in this collection is 0. That implies that u minus $\bar{u}$ equals 0. And therefore, u equals $\bar{u}$, which means u is equal to this series.

OK, so we've shown that if a Hilbert space has an orthonormal basis, then every element can be expanded in what's typically referred to as a Fourier-Bessel series in this way, in terms of the elements of the orthonormal basis. Now, let me just tie one thing up. So we know that every separable Hilbert space does have an orthonormal basis, a countable maximal orthonormal subset. And so if H is separable, that implies H has an orthonormal basis.

Now, what this theorem also proves is that if H has an orthonormal basis, then H is separable. So let me just state that as a simple theorem that follows from this. So we've shown that if H -- that was the first theorem my stated at the beginning, that if H is separable, then it has an orthonormal basis. If H has an orthonormal basis, H is a Hilbert space.

So I should have said that at the beginning. But H is a Hilbert space. Then H is separable, meaning it has a countably-- a countable dense subset. So what is this countable dense subset? I'm just going to give you the subset and then talk through why this subset works.

So suppose e_n is an orthonormal basis for H , again a Hilbert space. Then if I define S to be the union over all, let's say, m natural number of elements of the form sum from n equals 1 to m of $q_n e_n$ where q_1 -- these are just rational numbers. So first off, this is a countable subset of H .

Why is that? Well, each of these is in one to one correspondence with the m -fold Cartesian product of the rational numbers. And now, the rational numbers are countable. And you proved back in 18.100 that any-- in fact, well, any Cartesian product, finite Cartesian product of a countable set is, again, countable.

And then we're taking a countable union. Another thing you proved in 18.100 is that a countable union of countable subsets is countable. So that's why S is countable. And then so now, I'm just going to state by the previous theorem S is dense in H . And then I'm going to put a box here and explain why. So every element-- so if H has an orthonormal basis, by the theorem we proved, every element can be expanded in one of these Fourier-Bessel series as coefficients times-- as an infinite linear combination of orthonormal vectors, of these orthogonal vectors.

Now, this means the partial sums are converging to a given element u . So what we have to show for this thing to be dense-- we have to show for every ϵ , there exists something from this set within ϵ distance of that given vector. Now, give yourself a vector. Its Fourier-Bessel series converges to it. So we can cut off a-- we can cut the series off at a certain point and still be within ϵ over 2 to that element we're trying to get close to.

Then that finite-- that finite sum will be in one of these-- well, it won't be in one of these. But it will just be the sum from n equals 1 to m of some coefficients time the e sub n 's. Now, the rational numbers-- I should say-- there, now that's correct. Now, any complex number here can be approximated by a rational number here plus i times the rational number here. This is still in one to one correspondence with $\mathbb{Q}^2$, the Cartesian product of with itself. And therefore, this is still countable-- and so I should have said $2m$ here. I was thinking about real numbers-- so that you can get close to the actual Fourier-Bessel coefficients that appear in that sum by now just rational or complex numbers with rational real and imaginary part.

So I hope that explanation was clear. Maybe sit down, and think about it, and actually write down carefully the ϵ argument. But that's essentially why this is true.

So what we've shown is that a Hilbert space is separable if and only if it has an orthonormal basis. So let me make a remark that what we've shown up to this point is that if H is a Hilbert space, then H is separable if and only if H has an orthonormal basis.

All right, so we've proven that, in the Hilbert space, if a Hilbert space has an orthonormal basis, then every element can be expanded in this kind of infinite series involving or the orthonormal vectors. I say infinite. That's only if the orthonormal basis is countably infinite. If it was finite, then that's actually not an infinite sum. It's a finite sum.

Now, what follows from that is that-- so we have Bessel's inequality, which says that the sum of squares of the coefficients appearing in this Fourier-Bessel series is always less than or equal to the norm of u squared. That's no matter what you assume about the orthonormal subset. That's for any orthonormal subset.

But now, we have that if it's an orthonormal basis, then, in fact, we have equality. So if H is the Hilbert space and this is a countable orthonormal basis, then for all u in H , we have that the sum n equals 1-- I'll just put sum over n because this may be a finite sum-- $\sum u_n$, which by Bessel's inequality, we always had less than or equal to the norm of u squared, in fact, equals norm of u squared. And this is sometimes referred to as Parseval's identity.

All right, so what's the proof of that? We have that u equals the sum of $u_n e_n$. And therefore, immediately, by the continuity of the inner product, we can write norm of u squared as-- so this is for the case that we have countably infinite. If it's finite, if it's a finite orthonormal basis, this follows immediately. But sum from m equals-- or the sum as m goes to infinity of $\sum_{n=1}^m \langle u, e_n \rangle \langle u, e_n \rangle$ equals $\langle u, u \rangle$.

And this is equal to the limit as m goes to infinity of sum of n equals 1 to m of $\langle u, e_n \rangle \langle e_n, u \rangle$. And now, this constant here comes out with a complex conjugate-- $\langle u, e_n \rangle \langle e_n, u \rangle$. And now, we only pick up-- so n and l are going from 1 to m . For a fixed n , we only pick up the l where l equals n because these vectors are orthonormal. And when n equals l , this gives me 1.

So I only pick up-- when l equals n , I get $\langle u, e_n \rangle \langle e_n, u \rangle$ times the complex conjugate of $\langle u, e_n \rangle \langle e_n, u \rangle$, which is just the norm squared when I multiply by that. And I'm taking the limit as m goes to infinity. So I pick up the whole sum, the infinite sum, if it's infinite.

All right, so we've seen a couple of simple applications of the theorem. We proved that if I have a normal basis, then the Hilbert space must be separable. And also, we have equality in Bessel's inequality for the-- in terms of the orthonormal basis.

What's more is that the previous actually gives us a way to identify every separable Hilbert space with the one you were introduced to in the first week, although I didn't call it a Hilbert space. So this is the following theorem, that if H -- so if H is a finite dimensional Hilbert space, that means I can find finitely many orthonormal vectors that span the space. And it's quite easy to show that then that is isomorphic to $\mathbb{C}^n$ in an isometric way.

So I'm just going to state the infinite dimensional version. If H is an infinite dimensional separable Hilbert space, then H is isometrically isomorphic-- and I'll spell out what these words mean-- to ℓ^2 , little ℓ^2 , meaning what? Meaning there exists a bijective map T going from H to little ℓ^2 -- so it's one to one and on to-- so that for all u, v in H , we have that the norm of Tu -- I should say bijective linear map, bijective linear map. Or I should even say, I mean, bijective linear operator. So it's going to be a bounded linear map. And this follows immediately from what I'm about to write down-- such that if I take the norm of the image in little ℓ^2 , this is equal to the norm of the vector in H .

So that's the isometric part here, is that the map is not changing lengths. And basically, if I have two maps between-- or if I have a map between two Hilbert spaces that preserve the lengths, then it also preserves inner products by-- not the parallelogram law, it's the polarization identity, which you see in the assignment. But anyways-- and also, we have $\langle Tu, Tv \rangle$ -- taking this inner product and ℓ^2 because these are elements now of little ℓ^2 , this is the same as the inner product in H of u and v . And it really just follows kind of immediately from what we've done, once I just write down the map.

So what's the proof? So since H is a separable Hilbert space, it has an orthonormal basis, e_n , which is countably infinite, since we're in the infinite dimensional setting. And by the previous theorem, we have for all u in H , $\|u\|^2$ is equal to sum from n equals 1 to infinity $\langle u, e_n \rangle \langle e_n, u \rangle$. And with the norm of u squared equals-- if you'd like, let me remove the squares and I can write this as sum from n equals 1 to infinity of $\| \langle u, e_n \rangle \|^2$.

So now, this should just jump at you. How do I define my map from H to little ℓ^2 ? Then I define T of u to simply be the sequence of coefficients appearing here. And this is an element of little ℓ^2 by this identity here, Parseval's identity. Then T does the job.

So I didn't go through and check all the properties that I needed to check. I mean, it's clear that it's going to be linear in u because these coefficients are linear in u . And that's clear. It's one to one. It follows from the fact that every u is expanded this way. And therefore, if the coefficients are the same for two different u 's, then those two u 's have to be the same. So that makes it injective.

The fact that it's surjective-- again, it doesn't take too much to prove, although you just show that for every choice of sequence in little ℓ_2 , that forming such a sum-- so let's label these C_n . So now, you have just the element of little ℓ_2 . Now, if I put C_n here e_n , now, you can argue, as we did before, that this series is, in fact, Cauchy in H and therefore converges to some element in H . And then you can prove that T takes that element that you have to the sequence that you started with, proving that T is surjective. So I'm not going to go through all the details. But it should be kind of clear that this-- at least based on this identity here, what the map should be.

All right, so we've seen some applications of this general theory of orthonormal bases of countable maximal subsets in the Hilbert space applications, meaning if it has a basis, orthonormal basis, then the space has to be separable. Every separable infinite dimensional Hilbert space is basically the same. They're all isometrically isomorphic to little ℓ_2 .

But how can we use this in a more concrete setting, whatever concrete means? I mean, concrete kind of is by taste. So I thought we'd pause here on the general theory for Hilbert spaces that we've been doing and do something a little more specific and look at Fourier series, which will connect this general stuff that we've been doing with Hilbert spaces to more of the concrete bit of producing Lebesgue integration in these big L^p spaces, which we proved are complete spaces involving, in some sense, integrable functions.

So let's take a pause from general theory and talk a little bit about Fourier series. OK, Fourier series, which was the reason why a lot of especially the integration theory was created in the first place, was to understand a certain question, which we'll get to in a minute.

Let me just start off with a very simple theorem, that the subset of functions, e^{inx} over $-\pi$ to π in now an integer is an orthonormal-- orthonormal subset of L^2 minus π to π .

And here, let me-- just for a quick refresher, if t is a real number, e^{it} , this is simply defined to be the complex number of $\cos t$ plus $i \sin t$. And it satisfies all the things that you know and love about the exponential. If I multiply e^{it} times $e^{i\tau}$, that's equal to $e^{i(t+\tau)}$ just by referring to the definition and using angle sum formulas and so on.

All right, so this is not too difficult to prove. What's the proof of this? If I take e^{inx} and inner product it with e^{imx} , now this inner product is in big L^2 . This is, by definition, equal to the integral from minus π to π of e^{inx} times e^{-imx} complex conjugate.

And again, from the definition, when I take a complex conjugate, that flips this i to a minus i . And then I can take that minus that sits out here and put it in, and make this a minus t here and a minus t there, since $\cos t$ is even and $\sin t$ is odd. So then this becomes $\int_{-\pi}^{\pi} e^{i(n-m)x} dx$. And again, you can just go from the definition of what this is. You don't need any fancy complex analysis. This is $\int_{-\pi}^{\pi} \cos((n-m)x) dx$.

Now, this quantity here-- when n equals m , I just get 0 here. And therefore, this is just 1. And when I integrate that, I get 2π . So this equals 2π . And now, when n does not equal m , this has an antiderivative that you expect it to be. And therefore, this is equal to $\frac{e^{i(n-m)x}}{i(n-m)}$ evaluated at π minus $-\pi$. And this is if n equals m , n does not equal m .

Now, what is it about e to the it ? This thing is 2π periodic. And it's the same if I have an n here. Then that's going to be 2π over n periodic. So if I stick in π here, the value I get is going to be the same as when I stick in $-\pi$ here because the difference between these two values is 2π . This is 2π periodic. And I'm sticking in two numbers separated by 2π . So this will be 0 when n does not equal m .

And therefore, that proves the claim that this is an orthonormal subset. We divide by square root of 2π so that when I take the inner product of this with itself or the element with itself, I get 1. Here, I just did e to the inx , so I got 2π . So I divide by the square root of that to get the orthonormal thing.

So make a definition. So let f be in L^2 of $[-\pi, \pi]$. The n -th Fourier coefficient-- this is the new bit of terminology-- of f is the complex number $\hat{f}(n)$ equals $\frac{1}{2\pi}$ inner product-- I mean-- inner product-- integral of $[-\pi, \pi]$ of $f(t) e^{-int} dt$.

And the N -th partial Fourier sum of f is denoted by capital S_N of f of x . This is equal to the sum for n in absolute value less than or equal to capital N . So little n is an integer of $\hat{f}(n) e^{inx}$, which I can-- is the same as f inner product with e^{-inx} over square root of 2π e^{inx} .

All I did with my definition of the Fourier coefficient is I combined these two square root of 2π 's in my definition of $\hat{f}(n)$. But this partial sum is just equal to a partial sum in terms of this orthonormal subset.

OK, and we also associate to function f a formal object. The Fourier series of f is the formal series, because we are not making any claims about its convergence at all, summed from n in $\mathbb{Z}$ of $\hat{f}(n) e^{inx}$. So the question that we're going to answer-- and like I said, what motivated all of this to begin with is the following.

So back when Fourier was studying heat conduction, he made a claim that every function can be expanded in terms of-- essentially, he was doing cosines and sines. But in terms of this, he said every function can be expanded-- is equal to its Fourier series. Now, at the time, people said, no, that's not the case. Not every function is periodic. And these are 2π -- each of these is 2π periodic. So what are you talking about?

And then he said, maybe if we restrict to continuous functions, maybe it's equal to its Fourier series. That's not true. There's continuous functions that have-- where the Fourier series actually diverges at a point and doesn't converge back to the function.

But Fourier series, now if you're asking about pointwise convergence, that's a very tricky and delicate issue. But Fourier series has a very nice and beautiful answer when you look in terms of the space that you're supposed to. As I started off this discussion, I said these elements are orthonormal with respect to this inner product. And this inner product lives, as far as in a Hilbert space, in the space L^2 .

So the question that one should ask then is-- you have this or the normal subset of L^2 . So the question is, do we have for all f in L^2 $f(x)$ equals $\sum \hat{f}(n) e^{inx}$. So this is an infinite sum. So I should be talking about in what sense is this series converging. And I mean in the sense of in the space that we're asking this question, in L^2 , i.e. do I have that the partial sums converge to f in the L^2 norm?

So I'm not going to write the argument. OK, maybe I will. Does this equal-- does this-- OK, now, I'm butchering this. Instead of writing that limit, let's say, does this norm converge to 0 as capital N goes to infinity? So that's the question. Do we have convergence of this series, meaning the partial sums, do they converge to the function f in the L^2 norm? So this is something weaker than-- or maybe not even comparable to pointwise convergence. So sometimes, this is referred to as convergence in the mean.

And so what is the way to phrase this question based on what we've done? So this question is equivalent to the question, is this collection of orthonormal elements in big L^2 a maximal subset of big L^2 ? i.e. and so now, let's really put this question into equivalence with just a pure statement-- i.e. does the vanishing of all the Fourier coefficients imply that the function is 0?

So we proved that this statement here is equivalent to e to the ∞ being an orthonormal basis. So we proved one direction. We proved that if I have a collection of orthonormal vectors, then that's an orthonormal basis implies that every function can be expanded in this infinite series. But the converse clearly holds, too. If I can expand every function in an infinite series with those coefficients, then that implies that the subset has to be orthonormal-- has to be maximal because if something's orthogonal to everything in the collection, then all those coefficients appearing in the series are 0 and therefore the function is 0.

So this question here is equivalent to asking whether or not this collection is maximal, which means if I have an element in L^2 that is orthogonal to everything in here, then it has to be 0, which is this statement here. So let me put a box around it. Based on what we've done, of the question of convergence of Fourier series in big L^2 -- and what I'm using here really as well is the fact that big L^2 is complete.

This all only works in a Hilbert space. At certain points, we relied on the fact that Cauchy sequences converge to something in this space. So the fact that I can cleanly reduce this question down to what's in yellow relies on the fact that big L^2 is complete, which we did a lot of work to show and to construct.

So this is the question that we're going to try to answer. And let me go ahead and answer-- I asked a question, so I should give the answer. The answer is yes, but it's going to take some work. This is a non-trivial matter. So let's see, how much time do I have?

OK, so how we're going to proceed is via what may be referred to as Fejer's method. So again, our goal is to show this, to answer the question about convergence of Fourier series in L^2 . Why in L^2 ? Because that's the complete-- that's a complete Hilbert space that we're working in, that we're going to apply our general framework to.

So let me start off with a following simple calculation, that for all f in L^2 minus π to π and for all natural numbers n , including 0, if I want to look at the n -th partial sum of f , I can write this as the integral from minus π to π of a function evaluated at x minus t times f of t integrated dt . And again, this should be interpreted as a Lebesgue integral, even though I'm writing it using the notation that you use-- Lord, sorry, that e was awful. I just couldn't let it stay there.

Even though I was using this notation at one point to denote the Riemann integral, I'm now using this notation for the Lebesgue integral if the function's-- if I'm talking about Lebesgue integration, which is not such an abuse of notation because we found out that the Lebesgue integral of continuous functions is the Riemann integral.

So let me finish the statement. So we can write it as some function of x minus t times f of t integrate from minus π to π where dn of x , this is equal to the function, which is $2n$ plus 1 over 2π when x equals 0 and $\sin$ of n plus $1/2$ of x over 2π $\sin x$ over 2 when x does not equal 0 . And this function here-- so first off, note that this is a continuous function. As x converges to 0 , using L'Hopital's rule, if you like, this thing converges to $2n$ plus 1 over 2π . In fact, it's a smooth function. This function here is referred to as the Dirichlet kernel.

As a first step, I said, we're going to look at-- we're going to rewrite the partial sums in this way. I'll tell you why in a minute. But let's just take this as a warm-up calculation of some calculations to come.

OK, so what's the proof? We have that the n -th partial Fourier sum of f , this is equal to n less than or equal to N of-- let me just write out here what the Fourier coefficient is-- 1 over 2π minus π to π of e to the minus i of f of t , e to the minus $int dt$ e to the inx .

And this is just a finite sum. So I can bring this inside the integral and combine everything. And I get this is equal to minus π to π of f of t times dn of x minus t , where dn of-- so in fact, let me-- let me not jump ahead. I'll go ahead and write this out. This is equal to now sum times e to the inx minus $t dt$. So this is dn , all right?

So dn -- so call this thing, dn of x minus t , x minus t appearing there. And now, let's compute dn of x . So let me just rewrite what it is, sum in less than or equal to n e to the inx . Where'd the t go? Again, again this is dn of x minus t . So the argument's there. So dn of x is-- there's the argument. Maybe I should have put a y or something like that, but I didn't think that far ahead.

So this is the d capital N . And now, let's just massage this a little bit. We can write as e to the minus inx times now the sum from n equals 0 to $2n$ e to the inx . This is a sum from, if you'd like, n equals minus capital N to n . All right, so if I factor out an e to the minus inx , I can write this sum as this sum.

Now, this is a geometric sum. This e to the inx , again, this is something that you can just check from the definition. This is also equal to e to the ix raised to the n -th power. And I know how to sum things that involves something being raised to the n -th power-- finite sums, that is.

This is equal to 1 over 2π times e to the minus inx times 1 minus e to the $i 2n$ plus 1 x over 1 minus e to the ix , the thing that's appearing here. $2n$ plus 1 the thing that appears on top plus 1 .

Now, multiplying through by this e to the inx , pulling out a $1/2$ e to the ix over 2 and distributing that to the bottom, we get that the previous is equal to 1 over 2π . So first off, this is valid only when x -- what? When x is not equal to 0 .

When x equals 0 , I just get 1 here and I get 1 here. And then I get the sum from 0 to $2n$, which is equal to $2n$ plus 1 . That's why I get what I get when x equals 0 . So this is valid for x not equal to 0 . And so I get e to the inx plus 1 x . I should say, after I've taken away that $1/2$ x that appeared with this $1x$ and distributed it down on the bottom e to the ix over 2 minus e to the minus ix over 2 .

Now, if I have e to the i times something minus e to the minus i times that something and I subtract those two, I pick up $2i$ times $\sin$ of whatever's in the-- whatever this real number is. So I pick up 1 over 2π times $2i$ $\sin$ of n plus $1/2$ x over $2i$ $\sin$ of x over 2 . And these $2i$'s cancel. And that equals 1 over 2π $\sin$ of n plus $1/2$ x over $\sin x$ over 2 , which is what I wanted.

So what's the idea of trying to prove this? In the end-- I mean, in the beginning, I should say, we were asking about the convergence of the partial Fourier sums to f . And we don't know that going in. That's what we're trying to prove. So maybe working with the partial sums is not the best thing.

What am I getting at here? How about I introduce the next bit, and then I'll explain why we're interested in it. So if f is in L^2 from minus π to π , we define the N -th Cesaro-Fourier mean of f -- so this is the new bit of terminology-- to be the average of the partial sums of f . This we denote by σ_n of f of x . This is equal to the average of the first n partial sums of f .

In the end, what we'd like to do is we're trying to establish that this claim in the yellow box-- if the Fourier coefficients are all zero, then the function has to be zero. So you may think, well, let's prove that the partial sums converge to the function f . And that would give us immediately that the function is 0 because the partial sums would all be 0 because they involve the coefficients of f . But that's ridiculous because that's actually what we're trying to prove. That's equivalent to the question that we're trying to-- that's equivalent to what's in the yellow box. We're not making our life any easier by doing that.

So now, what one can do instead is instead of trying to prove what's in the yellow box, that if all the coefficients are zero the function is zero, what if we can prove that this object now converges to f ? This object here, which is this mean of the partial sums maybe has better properties than the partial sums originally that we were trying to study.

Remember if you look back to 18.100, if you have just say a sequence of real numbers which converges, you can define a Cesaro mean of that by averaging the first n terms, just like we did here. There's an $n+1$ here because we're starting at 0 and going up to n rather than going from 1 to n . But if you have a sequence of real numbers, you can look at its Cesaro sums or Cesaro means.

And what's great about the Cesaro means is that you don't-- it kind of behaves a little bit better than the sequence you start with. But you don't lose any information. So if the original sequence converges, then the Cesaro mean also converges. So if you're expecting the partial sums to converge to f , then the Cesaro means should converge to f .

But the Cesaro means have an even better quality. Let's go back to sequences of real numbers. You could have sequences of real numbers that don't converge whose Cesaro means do converge. So take the sequence 1, minus 1, 1 minus, 1, 1 minus 1, and so on, that sequence doesn't converge. But the Cesaro means do converge. The Cesaro mean is 1, 0, 1/3, 0, 1/5, 0, 1/7, 0. So the Cesaro means converge to 0.

So all of that is to say that this object here, which is the average of the partial sums, we expect to behave better than just how the partial sums converge. We expect it to have better convergence properties. And what does better convergence properties mean?

Quite honestly, it means we should be able to show it converges to f in hopefully some straightforward way-- or maybe not straightforward way, but in a-- yeah, I don't know how to describe that. And that's what we'll do. That's the plan, is what we're going to show is that for every f in L^2 , the Cesaro means converge to f in L^2 , meaning the limit as capital N -- so let me-- so what's the goal that we'll show is we'll show that the Cesaro means-- Cesaro-Fourier means of f converge to f in L^2 .

Now, if we can do this, then we will have answered and we will have shown what's in the yellow box because if you assume all of the Fourier coefficients are 0, then all of the partial sums are 0. And therefore, all the Cesaro means are 0. And by this thing that we are hoping to prove, this will prove that 0 converges to f . And therefore, f is 0, which is what we wanted to show.

So one more time-- we want to prove that if all the Fourier coefficients are 0, then the function has to be 0. Now, if all the Fourier coefficients are 0, then all the partial sums are 0. And therefore, all the 0 means are 0. And if we're able to prove this holds for all f in L^2 , then this would tell us that 0, which is what all of these Cesaro-Fourier means are, converges to f . And therefore, f is 0, our desired conclusion.

And from that, we conclude that the collection e^{inx} over square root of 2π is a maximal-- is a orthonormal basis for L^2 . And therefore, the Fourier series of a given L^2 function converges to the function in L^2 , meaning f minus the partial sum converges to 0 as n goes to infinity. All right, and so that'll be what we do next time, is prove this claim right here.