PROFESSOR: So last time we introduced pre-Hilbert spaces as vector spaces that come equipped with a Hermitian inner product. Our Hermitian inner product is linear. So it's a pairing between two elements that gives you a complex number that's linear in the first entry.

And if you switch the entries, that's equal to the pairing of the original two entries, taking the complex conjugate of that. And it's positive definite, meaning if I take the inner product of a vector with itself, it's non-negative and 0 if and only if the element is zero. And we proved the Cauchy-Schwarz inequality.

So let me actually just recall that we defined for if h is a pre-Hilbert space. We defined what we called, or using this norm notation, although I haven't proved it's a norm yet to be the inner product of v with itself raised to the 1/2 power. And this is a non-negative real number.

This is a non-negative number. So taking the 1/2 power is meaningful. And we proved at the end of last time the Cauchy-Schwarz inequality that for all u, v in H if I take the inner product of u and v, take its absolute value, this is less than or equal to norm of u times the norm of v. OK, so now let's use this to actually prove or to prove that this thing that I have been denoting with this norm notation is, in fact, a norm on a pre-Hilbert space.

So theorem H is a pre-Hilbert space, then this thing here defined in this way is norm on h, OK. So remember, we have to prove two-- or we have to prove three things for something to be a norm. We have to prove it's positive definite. And then we also have to prove homogeneity and the triangle inequality.

So note that this quantity here equals 0 if and only if the inner product of v with itself is zero, which by the positive definite quantity of a Hermitian inner product implies that v equals 0. So that proves that this function here on h is in fact positive definite.

Now if lambda is a complex number and v is in H, if I take-- so we kind of saw this at the proof of the Cauchy-Schwarz inequality. But if I take lambda times v and inner product lambda times v, this is equal to lambda times lambda bar. A scalar pulls out of the first entry unfazed.

If a scalar is in the second entry, it comes out with a complex conjugate. And then of course for a complex number of times this complex conjugate, I get the norm of that or the length of that-- the absolute value of that complex number squared. So taking the square root of both sides of this, which is equal to this, I get that this quantity here is equal to the absolute value of lambda times this norm of v, which proves homogeneity of this function on h.

OK, so all that remains to prove that this is an actual norm-- so I don't have to keep being or at least trying to be careful with what words I'm using-- will prove that this is a norm. OK, so now we need to prove the triangle inequality. You let uv be in H, then I compute u plus v squared.

This is equal to u plus v plus v, which is equal to just how we computed it for when we had a t here and a t here, when we did the proof of the Cauchy-Schwarz inequality. And we'll use this identity quite often that the norm of u plus v squared is norm u squared plus norm b squared plus 2 times the real part of u and v, the real part of the inner product of u and v.

OK, now this is less than or equal to norm squared plus v squared plus the absolute value of the real part. Now the absolute value of the real part of a complex number is less than or equal to the absolute value of that complex number. So this is less than or equal to the absolute value of the inner product of u and v.

And by Cauchy-Schwarz that is less than or equal to still what I have before plus 2 times the inner product of u and v. And I guess let me go to the next board. All the proof is essentially done.

What I had before is equal to norm of u plus normal v squared, i.e. So I had started off with the norm of u plus v squared, and I had proved that's less than or equal to this thing here squared. So taking the square root of both sides, I get the norm of u plus v is less than or equal to the norm u plus the norm of v.

OK, so this thing that I defined before is, in fact, the norm. I'm not just using that notation to denote an impostor. OK, now using the Cauchy-Schwarz inequality, we can also prove that taking the inner product is-- it's a function on h cross h that's continuous, OK. So let me label this as continuity of the inner product.

So let me state it as the following. If we're in a Hilbert space and un converges to u and vm converges to v in a pre-Hilbert space with norm as defined as before-- so now we have a norm on a pre-Hilbert space, so we can define convergence since it's a norm space. Pre-Hilbert space H with norm this, then un, inner product vn converges to u inner product v.

OK. All right, so why is this? So last time in the previous proof, I mean, we really didn't use the full strength of the Cauchy-Schwarz inequality. We could have just gotten by with the real part, having just proven that the real part of the inner product is an absolute value less than or equal to the inner product of-- or the product of the norms of u and v. Here we'll actually use the full Cauchy-Schwarz inequality.

So the proof is quite simple. If un converges to u and vn converges to v-- i.e, let me just spell this out for you-- un minus u converges to 0. And norm of vn minus v converges to 0.

As n converges to-- as n goes to infinity, then we'll use the squeeze theorem to show that this quantity converges to u in a product with uv. So I have to show that this quantity here-- hold on. Let me not put that.

We're not in 18100. Typically in 18100 when I first started teaching the squeeze theorem, I always included the lower bound. But it should be clear that this is always bigger than or equal to zero since it's absolute value of a complex number.

OK, so this is equal to un minus u, vn plus-- let's see. U, inner product, vn minus v. By the triangle inequality now, just for the modulus or absolute value of complex numbers, that's less than or equal to un minus u, vn, plus-- that's the value of u, vn minus v.

And this is less than or equal to by the Cauchy-Schwarz inequality norm of un minus u times the norm of vn plus the norm of u times the norm of vn minus v. And now the vns are. So the vns are converging to v. And therefore, the norms of the vns are converging to the norm of v.

So recall that if the simple fact that vn convergence to v, then norm of vn converges to norm of v, OK. All right. And, again, so let me add-- since one can prove as you do in real analysis the following reverse triangle inequality since if I have two vectors v and h.

So this just works in any Banach space, not necessarily a pre-Hilbert space. The absolute value of the difference in norms is less than or equal to the norm of the difference, OK. So this is a convergent sequence of real numbers. So it must be bounded. So I can say that's less than or equal to times sup of n.

Or let's use a different letter, k, times plus norm of vn minus v times norm of u. Now, this is something converging to 0 times a fixed number. This is something converging to 0 times a fixed number.

And therefore, this goes to 0 as n goes to infinity. And therefore, the thing which we started with, so this quantity here, an absolute value, is less than or equal to something converging to zero. And it's also non-negative. So therefore I get that. So this is by the squeeze theorem, i.e.

OK, so on pre-Hilbert space, we can define a norm using the inner product. And this inner product is continuous with respect to this norm, OK. Because once we have a norm, we have convergence. So we can talk about-- once we have a norm, we have a notion of distance, so we can talk about convergence of sequences and things like that.

So now we have pre-Hilbert spaces. What is a Hilbert space? This is simply a pre-Hilbert space which is complete with respect to this norm. So a Hilbert space, H, is a pre-Hilbert space, which is complete with respect to the norm, with respect to this norm, which again remember was defined as-- I took a vector.

The norm of a vector here is equal to the inner product of the vector with itself raised to the 1/2 power. All right, so this is the new terminology. And as we'll see, or at least we'll explicitly spell out in the form of a theorem-- but we will see probably by the end of this lecture that really there's-- for a reasonable Hilbert space, you're either one of two things.

So let me first give the examples, basic examples of Hilbert space. So for example Cn which is a set of N-tuples. How do I denote this? Complex numbers where the inner product of two vectors, z and w, is equal to sum j equals 1 to the n, zj, wj. OK.

So this is an example of a Hilbert space, finite dimensional one, meaning it's a vector space of finite dimension. The other example we have is this space, little l, to which this was the space of sequences such that each of these is a complex number. And this series is convergent, is finite.

Sum of the absolute value of the ak squares, this is a Hilbert space as well with what's the inner product of two elements of little l2. This is the sum from a plus 1 to infinity of ak, bk. OK. All right. So this is pretty clear. But note that the norm I get from this inner product, this is simply-- which was little l2 norm, right?

OK. So these are two basic examples of a Hilbert space. We will in fact show that every separable Hilbert space can be in a inner product. And therefore length preserving way-- so that's usually called an isomorphism form-- can be mapped isometrically to either Cn or little l2.

So these two kind kof-- if one wants to go about categorizing all of the possible Hilbert spaces as far as separable Hilbert spaces, which are the only reasonable ones, of course, one can come up with wild examples of Hilbert spaces which are not. Then these are the only two-- I shouldn't say two because this one is indexed by its dimension. But these are the only two types that you come up with, either a finite dimensional Cn, or it's isometric to little l2, OK.

So but I will still write down another example for-- since we went to all that work, if E is a measurable subset of r and l2 of e, the big L2 of e, which, remember, this was the space of measurable functions f from e to c such that f squared e is finite. This is a Hilbert space. What's the inner product?

The inner product of two elements in L2 defines v, just the analog of little l2 where we replace the sum with an integral, f times g, OK, the Lesbesgue integral of f times the complex conjugate of g. OK, now, I wrote down capital L2 I wrote down little l2.

But what about the other little lp and big LP spaces or any of those Hilbert spaces? So, of course, if I define the inner product in this way, as I did before, then that only induces the little l2 and the big L2 norm. But is there perhaps some other kind of magic inner product out there that I can put on little lp or big LP so that I would get out the little lp or big LP norm, when I define the norm according to how I've been doing it.

So the question on hand is-- the other little lp or big LP spaces, also Hilbert spaces. All right, so again it's clear that if I were to define the inner product in the way I did in the two examples, then that only is going to give the l2 norm. I'm asking now is there some magical inner product I can define on these spaces that spits out the little lp or big LP norm?

So the answer is no. And there's in fact, a way to determine whether or not a space is a Hilbert space because if you think about it, the way we've come about introducing what a pre-Hilbert space is and a Hilbert space is, we first had a norm in our hand. And then we'd define an inner product. Well, if that's the way you're building your space, then you know automatically that it's a Hilbert space because you had an inner product first and then you define the norm second.

Let's suppose the data you're given is some norm on a space. When can you determine if that norm comes from an inner product? That's the question that is underlying this question I wrote on the board. If I have a norm space-- all right, so my initial data is the norm. When can I tell that that norm comes from an inner product?

And so you'll prove this basically by direct calculation. And this is the following parallelogram law, which is the following. If H is a pre-Hilbert space, then for all uv in H, if I take the norm of u plus v squared, and I add the norm of u minus v squared, this is equal to twice the norm of u squared plus the norm of v squared.

So this is a condition that is stated purely in terms of the norm. Moreover, if h is a norm space satisfying star, then h is a pre-Hilbert space. OK, in other words, although it seems we defined Hilbert spaces or pre-Hilbert spaces initially in terms of an inner product, in fact, you can say a norm space is a pre-Hilbert space if and only if it satisfies this parallelogram law.

So if you have on your hands, just the norm, then as long as that norm satisfies this identity, that norm can be derived from an inner product. So using this theorem, you can check that the answer is only for people to-- so in other words, if you can come with u and v so that this inequality is not satisfied when p is not equal to 2. OK.

OK, so now we have the notion of a Hilbert space where this norm given in terms of an inner product, this space is complete. Now, we have an inner product, so we can start talking about vectors being orthogonal to other vectors or orthonormal sets, which when I first started lecturing about this stuff, I was already using the terminology, this thing being orthogonal to this thing or something like that. But of course, if you don't remember what those words mean from linear algebra, I'll quickly remind you.

So suppose we're in a pre-Hilbert space. We say that two elements, u and v, are orthogonal, if their inner product is zero. If instead of saying orthogonal I want to write this in words, I'll write u per v, OK.

So throughout H is going to be-- well, a pre-Hilbert space if I don't actually write it. If h is a pre-Hilbert space, a subset, which I'll denote by a lambda, lambda, and capital lambda, so just some subset indexed by some indexing set capital lambda. We say this set is orthonormal if for all lambda each one of these vectors in the subset has unit length. And I need two different indexed elements, implies that they are orthogonal.

OK. So maybe this notation scares you because what's this indexing set? Typically we just use the natural number. So let me just make that remark although, I'll make a few remarks in general about orthonormal sets that are not necessarily indexed by the natural numbers will mainly be interested in a finite set or a countably infinite set.

OK, so although an orthonormal subset of Hs could be a very crazy type of subset, mainly we're going to be interested only in finite or countable, countably infinite orthonormal sets. So what are some examples? OK, so simplest examples if-- so this is an example of a set of orthonormal-- or an orthonormal subset of c2. Or this is a 1.

This is also an example of an orthonormal subset of c3, let's say, using notation from before. If I denote by e sub n, this is the sequence consisting of zeros up until I hit the end spot and then 0 afterwards and entry, which is an element of little l2, then this is an orthonormal subset of l2, OK. One other example is--

Let's look at the functions 1 over square root of 2pi times e to the i and x. And let's think of these as elements of l2 minus pi to pi. Then this is an orthonormal subset of-- I might write O-N instead of writing out orthonormal, but this is an orthonormal subset of L2.

If I take the inner product of two of these guys, I get-- let's say I just look at two different ones. let's say m does not equal n, so then I take the inner product of imx with inner product of inx, complex conjugate. So that's the inner product on big L2, and this is equal to 1 over 2 pi times minus pi to pi.

Now if m equals n, I just get the length of e to the i, nmx squared. The length of e to the i, mx is 1. And therefore, I would just get 1 here, dx over 2 pi, which gives me 1. But now they're not equal. So I get e to the i, m minus nx, dx. And now the integral of this creature-- so here e to the i, let's say I had to come up with some e to the i y when y is a real number, this is by definition cosine of y plus i sine y.

Now you can check that fundamental theorem of calculus still holds for e to the i, m minus n over x And this equals e to the im minus nx over i times m minus n. And e to the i, m minus n times x is 2 pi periodic. So when I evaluate it at minus pi, I get the same value, if I evaluated at pi. So this equals 0.

OK. So that's an orthonormal subset of big L2. So this collection of vectors-- and I'm calling them vectors, but this collection of elements in big L2 is still countable, even though I'm indexing it by the integers rather than the natural numbers. OK, so we have the following.

Now most of what I'm going to say with regards to countable subsets of-- so countable orthonormal sets still carries through to possibly uncountable orthonormal sets where now a infinite sum over an uncountable number of elements has to be defined in a precise way. But I will really just stick mostly to the countable case, and if you're interested, you can always look that stuff up.

So I should say-- what I haven't said is that whether or not the lp spaces, let's say over an interval ab or over r, are separable or not or even over a-- or even over a-- so let's stick to either a closed and bounded interval or r. Now, why are these spaces separable, meaning they have a countably dense subset, so including big L2?

The reason is the following, so what you proved in the assignment is that the continuous functions are dense and little and big LP for p between p equals 1 to infinity, strictly less than infinity. OK, they're not dense in l infinity. So as long as you stay away from that, they're dense, OK.

So continuous functions are dense in lp. Now, what's one way to approximate a continuous function? Going back to your introductory analysis class, hopefully you covered what's called the Weierstrauss approximation theorem, which says that for every continuous function, you can approximate it uniformly on the interval by a polynomial, OK. So that shows that polynomials are dense and all the lp spaces, of course, not l infinity.

Now how do you go from the set of polynomials, which is uncountable to a count, which is dense in LP, but is uncountable to a countable dense subset, which is dense in LP. Make everything rational. OK, the set of polynomials with rational coefficients is, in fact, countable. And you can approximate every polynomial with real coefficients on a closed and bounded interval by a polynomial with rational coefficients. It's not too difficult to believe.

And therefore, the polynomials with rational coefficients are dense in the lp spaces as long as I'm not in l infinity. And therefore all the lp spaces are separable. All right, now the little lp spaces are also separable because except for L infinity also, which is not separable-- because, first off, a dense subset of the, let's say, little l2-- let's make things definite-- is the subspace consisting of all sequences, which terminate after some entry. In other words, it's 0 after that, OK.

Convince yourself that this subspace of all finitely terminating sequences, this is dense in little lp for every-- for p between one an infinity, not equal to infinity. So finitely terminating sequences are dense and little lp. Unfortunately, that's, again, an uncountable-- I mean, it's a subspace. So it's going to be uncountable.

So how do I get now a countable thing again? I replace everything by rational numbers. So if I can approximate every sequence which terminates after a certain point consisting of real numbers by a sequence full of just rationals terminating after a certain point by just choosing the rationals very close to those real numbers. And I didn't say this explicitly a minute ago that we still have the density of the rationals and the reals.

For every real, I can find a rational very close to it. So this is thinking that goes on. But now the set of all finitely terminating sequences with rational coefficients. This is a countable set. And that countable set is dense in little lp. And therefore little lp is separable as long as p is between 1 and infinity, excluding infinity.

OK, so I said at the beginning that we're going to be mainly interested in separable spaces without actually saying why little lp or little l2 and big L2 are separable. But I just gave you the argument by word of mouth now instead of actually writing it down. OK, so we have the following Bessel's inequality equality for countable orthonormal subsets.

So if the n is-- well, let me just put n here. If this is a countable, meaning it's either finite or it's countably infinite, when it's a countable orthonormal subset of a pre-Hilbert space, h, and for all u [? en ?] h, if I look at the sum of squares u in a product e sub n, this is less than or equal to the norm of u squared, OK.

So our discussion here of orthonormal subsets is taking place within a pre-Hilbert space. We don't need Hilbert spaces to talk about these concepts. Yeah, but when we're in a Hilbert space and we have a certain or the normal subset, that would be important that we're in a Hilbert space. OK, so the proof is-- let's do the finite case first.

So suppose I have a finite collection of orthonormal or a finite orthonormal subset of h or a finite-- or I'll often say finite collection of all normal vectors in h. And subset of h in on, standing in for orthonormal. Then let me just record a few identities, which are pretty easy to verify sum from n equals 1 to n of u inner product en, en.

If I take the norm of this thing squared-- let's compute this out. This is the inner product of with the understanding that n is going from 1 to capital N. Let's use a different index here. And this is equal to nmuen. Now I have this sum here of-- oh, I'm leaving out em of u, vm, the complex conjugate times the inner product of en with em.

That's the two vectors that are taking the complex conjugate of. This is a number here. It just comes out. This number here gets hit with a complex conjugate when it comes out. Now the inner product of en with em is 0. When n does not equal m, it is equal to 1 because it's equal to the norm of en squared when n equals m.

So all I pick up from this double sum, which is just a finite double sum, is when n equals m and n going from 1 to n. And therefore, this is equal to uen squared, OK. And that's one formula I want to have. Another one is that if I take the inner product of u with m equals 1 to the n of the sum.

And maybe you recognize what this sum here actually is. I'll say so in a minute. So this is equal to sum from n equals 1 to n of-- so u inner product e sub n times this number. This number comes out and gives me a complex conjugate. OK.

And therefore, zero, which is bigger than or equal to the norm of u minus n, equals 1 to n of uen, en. Now if you remember back from linear algebra or from calculus that if I have orthonormal vectors, and I have a vector u, this quantity is nothing but the projection of u onto the span of those orthonormal vectors. So what I'm looking at here is, if you like, the norm part of u that's orthogonal to these finitely many vectors.

OK, so this thing is bigger than or equal to zero. We use that formula of how to compute the norm of something plus something. And this is equal to norm of u squared plus norm n equals 1 to nuen, en squared minus 2 times the real part of u inner product sum from n equals 1 to n of uen, en.

OK, and now we know what all of these things are. This is equal to-- I'm not even going to go over-- this is equal to this quantity here as is this inner product. It's also equal to this thing here. So the real part is equal to this thing here because this is a real number.

Since this comes with a 2, I cancel one of those. So I get norm of u squared minus sum from n equals 1 to n of u, en squared, which is exactly what I wanted to prove for the finit case. But the infinite case then follows from the finite case. And by letting capital N go to infinity-- so infinite case, suppose en equals 1 to infinity is an orthonormal normal subset of H, then we know that for all N, capital N, we have that sum over n equals 1 to n of uen.

This is less than or equal to the norm of u squared. So I can just send capital N to infinity to get that the sum equals 1 to infinity of norm of un squared is less than or equal to the norm of u squared. OK. OK. So orthonormal subsets we can define as a collection of vectors that have unit length and are mutually orthogonal to each other.

Now just orthonormal subsets-- just any old orthonormal subset is not really the most useful thing, if we're trying to study the entire Hilbert's, or pre-Hilbert space h, because we may miss something if we leave out certain orthonromal vectors. But a more useful type orthonormal set is a maximal orthonormal set, which is defined as follows, an orthonormal subset e lambda, lambda and capital lambda of a pre-Hilbert space H is maximal.

So again, if having a possibly uncountable collection of orthonormal subsets indexed by some indexing set makes you uncomfortable, replace this with n, where in is going from 1 to capital N or n is going from 1 to infinity, so a countable collection, if you like. But I'm stating this so that you know that something more general is true. So this is maximal if what? The following holds.

If u is an h and u is orthogonal to everything in this orthonormal subset, this implies that u is 0. OK. So an example, of course, is you can check that this collection here is maximal orthonormal subset of c2, a non-example. You can maybe see this coming, the one we had a minute ago.

This is not maximal since there's a vector that has inner product zero with both of these but is not zero, since this should be c3. You just write it loosely this way since this vector 010 is orthogonal to both of these, but it's not 0. Maximal means if you're orthogonal to everything in your collection, then it has to be zero.

But this is non-zero, and it's orthogonal to everything to these two vectors there. Another example is, again, with the notation from before. This is where this is the sequence that is 0, except for the end spot where it is 1. This is maximal orthonormal subset of little l2. OK.

Now, what we're going to see very shortly is that if we have a countably infinite maximal subset of a Hilbert space, then that set serves kind of the same purpose as a basis as an orthonormal basis does in linear algebra. I mean, if you look at here, already this orthonormal set is maximal. And it also forms the basis for c2.

This was not maximal. And you see it doesn't form a basis for c3. So maximal is going to give us a condition, which is equally useful as being a basis. But it won't be a Hamel base. These subsets won't be a Hamel basis in the sense that every vector can be written as a finite linear combination of the elements of a maximal orthonormal subset. But what is true is that we can write it as possibly an infinite sum involving the maximal orthonormal normal subset, which is, in most cases, just as good if you want to use that, all right.

OK, so first off, when-- does every pre-Hilbert space have a maximal, a subset? So let me state this as a theorem. In fact, I'll state two theorems. The first is every-- I should say non-trivial because we could have the Hilbert space-- this be the 0 vector, pre-Hilbert space as a maximal orthonormal subset. Whether it's separable or not, it has a maximal orthonormal subset.

And the way you prove this is using-- so I'm not going to give a proof of this, I'm going to give you a proof of something a little less strong but about as useful as we'll need. One proves this by using Zorn's Lemma, by taking a set, your set that you're going to put a partial order on to be the collections of orthonormal basis, or not orthonormal basis, max of orthonormal subsets and then ordered by inclusion. And then one can do a Zorn's Lemma argument and apply Zorn's Lemma to obtain a maximal orthonormal normal subset.

But that's kind of hands off. And maybe that scares you a little bit because Zorn's Lemma is equivalent to the axiom of choice. So if you don't like using the axiom of choice, maybe you have a problem with using it to construct a maximal orthonormal subset of a pre-Hilbert. But we can actually do this by hand, if the Hilbert space is separable.

So this is a theorem we'll actually prove every non-trivial separable, meaning the pre-Hilbert space has a countably dense subset. Every non-trivial separable pre-Hilbert space as a countable maximal orthonormal subset, OK, which as I said, this is the main types of orthonormal subsets we'll be interested in just because defining infinite sums is easier to do over a countable index than it is an uncountable index.

But anyways, so we're going to prove and actually construct this essentially by hand, using the process that's the name of this little section, the Gram-Schmidt process. So if you remember from linear algebra, if you have a collection of vectors, you can always find an orthonormal collection of vectors that span the vector space that is spanned by the original set of vectors. OK. And that's what we'll do.

So since H is separable, let be a countable dense subset of H. And this is a non-trivial pre-Hilbert space. So we can always make sure that the first one is a non-zero vector. OK. So you know what countable means? Dense, remember, means that for any element in h there exists an element from this sequence, or from this collection, that's within epsilon of that vector from H.

So now I'm going to make the following claim. This is essentially the Gram-Schmidt process, which we'll prove by induction. For all n, natural number, there exists another integer natural number m of n and an orthonormal subset e1 to em such that following is true. First is the span of e1 up to em of n equals the span of v1 up to vn.

And so you see n is changing. So maybe for each time I change n, I get a different orthonormal subset or a wildly different orthonormal subset from the integer before it. So the property of these subsets are that I'm just simply adding a vector or not. And e1 up to em of n, so this collection is equal to the previous collection, union either the empty set if vn is in the span of the 1 up to vn minus 1 and some new vector e of m sub n.

And I'll tell you what em of m sub n is otherwise. OK, so what I'm saying here is that I have this countable infinite list of vs. And for each n, I can come up with a finite orthonormal subset that spans the same span as v1 up to vn. And at each stage, all I do is add a vector or not depending on if the next v is in the span of the previous or not. OK, I hope that's clear.

OK, so I proved this by induction. So proof of claim, this is by induction. So let's do the base case n equals 1. We take e1 to be e1 over length of v. All right. So now we're started. Now we've got our first vector in this list that we're building up, inductive step.

So let's call this, what I want to prove, star. So suppose stars hold for a-- so this whole claim here, I shouldn't say just a and b, but the whole claim holds for n equals k. And now I want to prove it holds for n equals k plus 1. So I want to-- what kind of vector do I need to add to the previous collection of vectors to now span v1 up to vk plus 1? OK.

So if vk plus 1 is in the span of v1 up to vk, then e1 up to em of k-- I should say the span of these vectors-- equals span of v1 up to vk equals because vk plus 1 is in this span. This span is also equal to vk plus 1. OK, and therefore, this case is handled. And we're in this spot where we don't add anything to the previous collection, OK.

This proves what we wanted for n equals k plus 1 by not adding anything. So this was in the case that vk plus 1 is in the span of v1 up to vk. So now let's do the more interesting case of v1 not equal or not in the span of v1 up to vk. So now suppose vk plus 1 is not in the span of v1 to vk. I define wk plus 1 to be vk plus 1 minus its projection onto the previous list of orthonormal vectors. So sum from j equals 1 m of k will be k plus 1 ej, ej.

Then first note that this vector cannot be-- so wk plus 1 this vector cannot be zero. Otherwise, this would imply that vk plus 1 is equal to this quantity here. And therefore, vk plus 1 is in the span of the ejs for j between one of between 1 and imps of k, which is equal to the span of the vjs for 1 up to k. But we're assuming this is not in the span.

I should say this is equal to the span of e1 up to em sub k. That's what we're assuming. That's the inductive hypothesis, right? So then this vector does not equal 0. And define e n sub k plus 1 or m to be the vector wk over its length for plus 1. Now, this is a unit vector.

And what? And if I take-- so now I claim this is orthogonal to e1 up to e of m sub k. If I take emk plus 1 inner product e sub n sub j-- this is equal to, simply by definition, 1 over length of mk plus 1 times the inner product of vk plus 1, minus the sum from j equals 1 to m sub k. So remember, this is supposed to be the projection onto the e1 up to em sub k. And vk plus 1 minus that is supposed to be orthogonal to those guys. So we'll see that this ends up being 0.

And let me change this j to an l since I'm-- or no. And I get 1 over time the sum, times vk plus 1, e sub l, minus now e sub l inner product with the sum. Now this is the sum from j equals 1 to m sub k of this number here times the inner product of e sub j inner product e sub l.

Now, the inner product of e sub j and e sub l is zero unless j equals and 1 if j equals l. So I only pick up the j equals l part of this sum when this inner product hits e sub l. In particular, I only pick up the coefficient in front since these have unit length. So I get v sub k plus 1, v sub l equals 0.

Then this guy now does. the job. So I would write a little bit more, but I'm running short on time. OK. So we proved the claim. And now let's use this to conclude that this collection of all of these els are maximal.

So we let s to be v sub n. And s is orthonormal subset of H. So I may not be adding any more vectors after a certain point and therefore just have a finite collection. That's also possible. But I may also have a countably infinite collection.

So this is an orthonormal set by v. And so we now show x is maximal, OK. All right. So we haven't used anything about the nature of this subset of H. All we view is that it's countable. So we could do this step by step, where we construct these orthonormal vectors that span each finite-- equal the span of each finite collection of vectors.

Now, to show it's maximal, this is where we use the fact that this collection of vs were dense or are dense in h. So suppose u is in H and for all l, u of e sub l equals 0. So this is either a finite set. So l is between 1 and n. Or it's a countable set, and l is going from 1 to infinity.

Either this is equal to e1 up to e m sub n for some n. Or e1 is countably infinite? OK. So since vj-- this is a dense subset of h-- we can find a sequence of elements from this collection. So v of j sub k, k equals 1 to infinity such that vj of k converges to u as k goes to infinity.

OK, now by property a, the fact that this span is equal to that span, v of j sub k is in the span of e1 up to em of j sub k. And therefore, by Bessel's inequality, and the fact that u is orthogonal to all of these orthonormal vectors, I get that v of j sub k squared-- now v sub j of k, this is in this span.

So you can show that, in fact, this norm is in fact equal to the sum of the coefficients that I get from taking the inner product of-- OK, if I have a vector in the span of a collection of orthonormal vectors, a finite collection of all the normal vectors, then the norm squared is equal to the sum of the coefficients, just like an rn. And since u is orthogonal to each of these, this I can write as--

And by Bessel's inequality, this is less than or equal to-- the sum of these things squared is less than or equal to squared, and this goes to 0 as k goes to infinity since the v of j sub k is converted to u. But we started off with the norm of e of j sub k, which implies zero. And therefore u, which is the limit of these, must be zero, proving that this is a maximal subset.

OK. So next time we'll prove that these maximal orthonormal, these maximal countable orthonormal subsets, in fact, form a pretty good analog of basis that you find in finite dimensional linear algebra. And we'll stop there.