

[SQUEAKING]

[RUSTLING]

[CLICKING]

CASEY
RODRIGUEZ: OK, so let's continue with our discussion of Hilbert spaces. And so last time, we finished with Riesz representation theorem, the only representation theory I remember, Riesz representation theorem, which states that if you have a Hilbert space, then for all f in the dual space-- which remember, this is a set of bounded linear maps from H to complex numbers-- there exists a unique v in H such that $f(u)$ is equal to $\langle u, v \rangle$ for all u and v in H .

So every continuous linear functional on a Hilbert space can be realized as an inner product with a fixed vector. So using this, we can revisit the subject which you touched upon in the assignment of adjoints. So let me just state this as a theorem.

Let H be a Hilbert space and A going from H to H be a bounded linear operator. The conclusion is then, there exists a unique bounded linear operator A^* from H to H , which we call the adjoint of A such that A^* satisfies the following property-- for all u, v in H , $\langle Au, v \rangle$ is equal to $\langle u, A^*v \rangle$. And moreover, the operator norm of the adjoint equals the operator norm of A .

So again, the same way we proved uniqueness for the Riesz representation theorem, you can prove that such a linear operator satisfying this identity for all u and v has to be unique. So I'll just state here that uniqueness of such an operator A^* follows from the identity it satisfies, just like the vector v , which spits out $f(u)$ when you take the inner product with u with it. Just like that vector is unique regardless of how you chose to define it-- I mean, how you created it based on that identity-- the same thing holds for the adjoint. So it's just the same argument is what I mean.

So how do we define this operator? Well, we basically define it by this relation. So let v be in H . Define a map f_v from H to \mathbb{C} via $f_v(u)$ is equal to $\langle Au, v \rangle$.

Now, this is clearly a linear map. Maybe I shouldn't check this, maybe I should, but let's check it anyways. If I take f_v of $\lambda_1 u_1 + \lambda_2 u_2$, this is equal to $\langle A(\lambda_1 u_1 + \lambda_2 u_2), v \rangle$.

And now, we use the fact that A is a linear operator so that it acts on the linear combination by linearity. So then I get $\lambda_1 \langle Au_1, v \rangle + \lambda_2 \langle Au_2, v \rangle$. And because the inner product is linear in the first entry-- it's conjugate linear in the second entry, but it's linear in the first entry-- this is equal to $\lambda_1 \langle Au_1, v \rangle + \lambda_2 \langle Au_2, v \rangle$, and which is $\lambda_1 f_v(u_1) + \lambda_2 f_v(u_2)$.

So this thing is a linear map. I claim it's also continuous, i.e., it's an element of the dual. And we just check.

If the norm of u equals 1, then if I take the absolute value of $f_v(u)$, this is, by definition, $|\langle Au, v \rangle|$, which is less than or equal to by the Cauchy-Schwarz inequality, the norm of Au times the norm of v . And the norm of Au is less than or equal to the norm of A , since A is a bounded linear operator, times the norm of u .

So this holds for all u and $\|u\| = 1$. And therefore, norm of this operator is less than or equal to norm of A times norm of v . So that proves that it's an element. So thus, f_v is an element of the dual.

And now we use the Riesz representation theorem. So by the Riesz representation theorem, there exists a unique element, which we denote A^*v , in H such that for all u in H , $\langle u, v \rangle = \langle u, A^*v \rangle$. I.e., for all u in H $\langle u, v \rangle$ because that's the definition of this guy on the left-- is equal to $\langle u, A^*v \rangle$.

So this just defines, for each v , an element A^*v in H . So now let's prove that this map v to A^*v is a bounded linear operator. That then defines the adjoint.

So first, we claim v , which maps to A^*v -- the element that satisfies this identity-- is linear. So let λ_1, λ_2 be in H , λ_1, λ_2 be complex numbers. Then for all u in H , we have what?

If I take-- let's see, how did I spell this out-- $\langle u, \lambda_1 v_1 + \lambda_2 v_2 \rangle$ -- so remember, this is the element or this is the element such that this inner product is defined by-- this is equal to $\langle u, \lambda_1 v_1 + \lambda_2 v_2 \rangle$. And this is equal to-- so first off, the inner product is conjugate linear in the second entry. So this is equal to $\lambda_1 \overline{\langle u, v_1 \rangle} + \lambda_2 \overline{\langle u, v_2 \rangle}$.

Now, how A^* -- how this symbol, if you like-- is defined is that it's equal to this quantity here is equal to $\lambda_1 \langle u, v_1 \rangle + \lambda_2 \langle u, v_2 \rangle$. And pulling the complex numbers back in, this is equal to $\lambda_1 \langle u, A^*v_1 \rangle + \lambda_2 \langle u, A^*v_2 \rangle$.

Now, this holds for all u in H , i.e., the inner product of u with this minus this is 0 for all u in H . And the only thing orthogonal to everything in H is the zero vector. So we conclude that this element is equal to-- so this element corresponding to this linear combination is a linear combination of these elements.

Thus, the map v that's mapped to A^*v is linear. And so we denote this map now simply by A^* . It acts on a vector v by operating by A^*v , by multiplying by A^* .

So now, A^* is a linear operator. When I write A^* here, I mean the map that takes v to A^*v . So now I want to check that this linear operator is a bounded linear operator. I need to check that it's continuous. And in the process, we'll end up showing that the norm of A^* equals the norm of A .

Now, suppose norm of v equals 1. So I want to bound A^* applied to v . So if A^*v equals 0, then clearly, A^*v is less than or equal to the norm of A . This, in the end, is what I will show. But I'm kind of dealing with the stupid case, which is when this thing on the left-hand side is 0.

So suppose A^*v is non-zero. Then, if I take the norm of A^*v squared, this is equal to $\langle A^*v, A^*v \rangle$. Now, by the definition of the inner product, this is equal to $\langle A^*v, v \rangle$, taking this A^* and moving it over here.

And by Cauchy-Schwarz-- well, first off, all of these numbers are real so it makes sense to actually write less than or equal to. But by Cauchy-Schwarz, this is equal to, or less than or equal to, $\langle A^*v, v \rangle$ norm times norm of v , which equals 1. Now, this equals 1.

And the norm of A applied to A^*v , that's bounded by the norm of A times the norm of the input, which is A^*v . And therefore-- so we started off with $\langle A^*v, A^*v \rangle$ less than or equal to the norm of A times the norm of A^*v . This is non-zero.

So I can divide by it and conclude that $\|A^*v\|$ is less than or equal to $\|v\|$. Now remember, the norm of an operator or a linear operator is the sup of this for all v which have $\|v\| = 1$. And since this is bounded above by the norm of A , that implies that $\|A^*\|$ is less than or equal to $\|A\|$.

So we've shown that basically all of this theorem, with the exception of equality of the norms. So every bounded linear operator has an adjoint which is a bounded linear operator and satisfies-- and it's a unique linear operator satisfying that identity up there for A^* .

Now, note that for all u, v , and H , if I look at A^*u inner product v , this is equal to the complex conjugate of v inner product A^*u , which is equal to, by the definition of A^* , moves over to v inner product Au . And again, applying the complex conjugate switches the order of the entries. So this is equal to u inner product A^*v .

So what have we shown? We've shown that the adjoint of the adjoint-- that is equal to A . So we've shown that A^*u inner product v is equal to u inner product A^*v . And recall this is supposed to be equal to u inner product A^*v , so the adjoint of the adjoint.

And we've shown that this quantity is equal to this quantity for all u and v . And therefore, the adjoint of the adjoint is equal to the operator again. Thus, by what we've done previously-- the previous argument, if you like-- applied now where A is replaced by A^* , the norm of A^* , which is equal to the norm of A -- the adjoint of the adjoint because this operator is equal to A -- by what we've done, we've shown that the norm of the adjoint is always less than or equal to the norm of the thing you're taking the adjoint of.

So this is less than or equal to $\|A\|$. And so we have this. And we have this, which implies $\|A^*\| = \|A\|$.

So I mean, what is this creature in practice? So let's take maybe a second simplest example. The simplest example would be if you're on \mathbb{C}^n and \mathbb{R}^n -- well, so let's talk about that.

Suppose now we have a matrix. So u is equal to u_1, \dots, u_n . And this is in \mathbb{C}^n . And A applied to u , if I want the i -th coordinate of that, is given by $\sum_{j=1}^n a_{ij} u_j$, where these are just complex numbers, some fixed complex numbers.

So A is just a matrix. A linear transformation on \mathbb{C}^n is always represented by a matrix. And here, I'm writing u in terms of the standard basis vectors e_1, \dots, e_n , and so on.

Then, to determine what the adjoint is, we figure out what's the operator that satisfies the identity the adjoint has to satisfy. So if we take Au and take its inner product with v , we want to write this-- so let me write this in a box-- we want to be able to write this as u inner product with some operator B applied to v . And then this would be-- and therefore, that would be the adjoint.

So if we write Au inner product v , this is equal to $\sum_{i=1}^n v_i \overline{\sum_{j=1}^n a_{ij} u_j}$. And this is, by definition, equal to-- I'm just going to write $\sum_{i,j} v_i \overline{a_{ij} u_j}$. And if I now switch this over and sum first with respect to i and then with respect to j , I can write this as $\sum_{j=1}^n \overline{u_j} \sum_{i=1}^n v_i a_{ij}$ applied to v inner product v , the complex conjugate there.

And so this tells me that-- so this is equal to $\sum_{j=1}^n u_j$ and times what I'll call the adjoint applied to v_j where-- what is the adjoint? Or how does the adjoint operate on an element v ? This is equal to $\sum_{j=1}^n$ let's say I want the i -th member now-- this is $\sum_{j=1}^n A_{ji}$ complex conjugate applied to v_j .

So for matrices-- again, this should be review-- if A is represented by the matrix, then the adjoint, which is also representable by a matrix, is representable by the matrix. So I should say the matrix for the adjoint A_{ij} is equal to A_{ji} complex conjugate.

So we did this for C^n . There's no reason we can't do it for, let's say, little l_2 now and make it a little more interesting. So suppose now I have now infinitely many numbers A_{ij} , so a double sequence in C^n , such that $\sum_{i,j} A_{ij}^2$, which is linked as n goes to infinity, of $\sum_{i=1}^n \sum_{j=1}^n$, is finite.

Now, let's define A now from little l_2 to little l_2 via A applied to a sequence, which I put an underline underneath, or let's make it u . OK, we've been using A 's for elements in little l_p as $\sum_{j=1}^{\infty} A_{ij} A_j$. And here, this is for A_{ij} l_2 .

Now, by the Cauchy-Schwarz inequality, you can check that if this condition is satisfied, then this is a bounded linear operator. The absolute value of this thing will be less than or equal to the $\sum_{j=1}^{\infty} A_{ij}^2$ sum of j squared, square root times the l_2 norm of A . And therefore, I can sum that using-- that's l_2 summable using the fact that this is finite.

Don't worry about the fact that I'm summing this double sum in a certain way by j going from 1 to n , i going from 1 to n . In fact, if it's absolutely summable in this, which it is-- it's the sum of non-negative numbers-- this sum doesn't depend on how I'm summing it in this way.

So then this is a bounded linear operator. And what's its adjoint? Its adjoint is going to be of the same flavor as what we got for the finite dimensional case.

And for all A, B and little l_2 , if I take the inner product of A applied to u , A, B , and little l_2 , this is equal to $\sum_{i=1}^{\infty}$ I mean, the same proof applies, essentially-- $\sum_{i=1}^{\infty} A_{ij} B_{ij}$ complex conjugate, which is equal to $\sum_{j=1}^{\infty} A_{ij} B_{ij}$ complex conjugate, complex conjugate, which is equal to inner product with $A^* B$, where $A^* B$ is defined via $\sum_{i=1}^{\infty}$ so I should have said this is the i -th entry in this new sequence in l_2 .

And so here, the i -th entry in this new sequence in l_2 is defined to be $\sum_{j=1}^{\infty} A_{ji}$ complex conjugate B_j . So just like in the finite dimensional case, for this case, if I switch the variables of this A -- if you'd like a double infinite matrix, if you want to think of it that way-- it's the same as in the finite dimensional case.

Let's do one last example. So I'm not going to go through again the computations, which are very similar. Here, we're having kind of finite sums. If you can do something with sums, you should consider whether or not you can do it for integrals.

So let's suppose K is a continuous function on $[0, 1] \times [0, 1]$. And we define a map A which we can show goes from L^2 to L^2 via $A f$ of x is equal to the integral from 0 to 1 of $K(x, y) f(y) dy$. In fact, this is for each f in L^2 , $A f$ -- this thing is, in fact, a continuous function.

It's more than just in L^2 . It's in fact, a continuous function. But continuous functions on $[0, 1]$ are elements of L^2 on $[0, 1]$.

Then you can check, just as we've done in these two other examples, that the adjoint of f is equal to-- or let's make this different-- say adjoint applied to g is equal to the integral from 0 to 1 $\int_0^1 \overline{g(y)} f(x, y) dy$. So again, it's like flipping the indices and taking the complex conjugate.

I said at the end of last class that you can tell something about the solvability of equations based on how the adjoint-- on properties of the adjoint. We'll see another way this is manifested later. That's kind of the simplest way.

And so suppose H is a Hilbert space. And A from H to H is a bounded linear operator. Then-- I used Ran in the assignments, so I'll use it this way here-- the range of A , orthogonal complement-- so the range you can show is a subspace of H . In fact, I think that was in one of the earlier assignments.

If you take the orthogonal complement of that, that is equal to the null space of the adjoint. So here, let me recall that range of an operator B -- this is exactly what it is. This is a set of all vectors Bu where u is in H and null space of B -- I might have called it the kernel, might have labeled it as the kernel of B , but both mean the same set. This is a set of all u in H such that $Bu = 0$.

So in particular, if we know that we have an operator A such that the range is a closed subspace, then being able to solve the equation $Au = v$ is equivalent to showing that the null space of the adjoint is simply $\{0\}$. So let me make that as a remark.

Suppose that the range of A is closed. Then A from H to H is surjective if and only if the adjoint is injective, meaning the only thing that gets set to 0 is the zero vector. Because if the null space is equal to the $\{0\}$ -- so first off, it's easy to see if the range of A is equal to H , then the orthogonal complement of H is the zero vector. So null space of the adjoint is just a set containing the zero vector.

Now, if the null space of the adjoint is just a zero vector, then taking the orthogonal complements of both sides, and using the fact that the range is closed so that the orthogonal complement of the orthogonal complement gives me back the set, I conclude that the range of A equals the orthogonal complement of the zero vector, which is H , the entire space. So we know that we have an operator which has closed range. Then being surjective is equivalent to the adjoint being injective.

So the proof of this, though, is pretty easy. So B is in the null space of A^* if and only if $\langle A^*B, u \rangle = 0$ for all u in H . $A^*v = 0$ if and only if this holds for all u in H .

And this is equivalent to, by the property of the adjoint, $\langle A^*v, u \rangle = \langle v, Au \rangle$ for all u in H . So $\langle v, Au \rangle = 0$ for all u in H . So v is a fixed thing. So this says that v is orthogonal to all of the elements of H of the form Au . But that's just the range of A . So this is equivalent to B is in the orthogonal complement of the range of A .

So we're soon going to get into the realm of more refined things we can say about solving certain equations involving operators, meaning when can you solve $Au = v$ and so on. But now, the best theorem about that you should know or maybe heard at some point from linear algebra is the rank-nullity theorem, that says the dimension of the range plus the dimension of the null space equals the dimension of the-- let's say you're going from spaces with the same dimension, then the dimension of the range plus the dimension of the kernel-- I may be actually getting this wrong; it's the end of a day; it's the end of a long day-- equals the dimension of the whole space.

So what this says is that in essence, in order to be able to solve a given equation, your input has to satisfy finitely many linear relations. And the solution of that equation is unique up to finally many-- or a finite dimensional subspace. Now this is the best thing you can say. Input has to satisfy finitely many linear relations and unique. So that existence is predicated upon your data satisfying finitely many linear conditions, and you have uniqueness up to a finite dimensional subspace, meaning the null space.

Now, this is from finite dimensional linear algebra. It would be great if we could do the same thing for infinite dimensions, not just because it would be fun, but in the end because we need to do these things, or I should say need to, but not in the sense that an airplane wing is going to fall off if we don't. But in order to learn more about certain problems that arise, it would be great to be able to do these things. And we'll be able to do them for certain operators, which are in some way-- these aren't the only operators, but for certain operators that are close to being matrices in a certain sense. So we'll get to that in a minute.

Now, these operators that we will eventually study the solvability properties of have a very special nature or property on how it acts on bounded sequences. Now, something that you take for granted in \mathbb{R}^n is that if I have a linear operator and I have a bounded sequence of vectors, then A maps that to a bounded sequence of vectors because a matrix is a bounded linear operator. It takes bounded sets of bounded sets. And therefore, by the Heine-Borel theorem, there will be a subsequence of the images of these vectors, which converges.

So there's some compactness hidden in what you're doing on \mathbb{R}^n and \mathbb{C}^n . And why am I rambling right now? I'm getting to the final point that we need to study a little bit about compactness and Hilbert spaces to eventually get to the place where we can say more about being able to solve equations involving bounded linear operators, now in infinite dimensions.

So with that rambling build-up, hopefully, I convinced you that we should study a little bit about what it means for sets to be compact in a Hilbert space. So let's look at compactness in Hilbert space. So let me recall for you what it means for a set to be compact in a metric space.

If X is a metric space, We say a subset K in X is compact if every sequence of elements in K has a subsequence converging to an element in K . So a set is compact if every sequence of elements in the set has a subsequence which converges to an element in the set.

So the simplest on Earth types of compact sets are finite sets. This just follows from the pigeonhole principle, basically. So the simplest example are finite sets of finite subsets of a metric space, of any metric space.

Now, we have this very cool theorem from Intro Analysis, which goes by the name of Heine-Borel, that says a subset K in \mathbb{R} -- you can replace \mathbb{R} by \mathbb{R}^n or \mathbb{C}^n , but this follows from the one-dimensional case, essentially-- a subset K of \mathbb{R} is compact if and only if K is closed subset of \mathbb{R} and bounded. So for example, the closed and bounded intervals a, b -- those are all compact sets. a, b , the set consisting of $1/n$, n a natural number, union the element 0. That's also a compact set. It's closed and it's bounded.

Now, this is a theorem about elements of compact subsets of \mathbb{R} . One can build off of the proof of this, really, to get the result for \mathbb{R}^n and \mathbb{C}^n . The metric properties of \mathbb{C}^n are the same as \mathbb{R}^{2n} , so let's say just \mathbb{R}^n -- that a subset of \mathbb{R}^n is compact if and only if it's closed and bounded.

Now, does that hold for arbitrary metric spaces? No. Does it hold for Banach spaces? No. You did an example on one of the assignments that proved that little l_p is not compact.

What about if we specialize the Hilbert spaces? The answer is still no. So let's make this a non example, since we did examples a minute ago. So all of these, as subsets of \mathbb{R} , these are all compact.

So let's suppose H is a infinite-dimensional Hilbert space. Then the closed ball B equals the set of u and H -- let's make it F -- such that length of u is less than or equal to 1. This is a closed and bounded set. You can check that. This is not compact.

Let's get back to compactness. So let's suppose H is a infinite-dimensional-- let's even make it separable, as all good Hilbert spaces are. Then the closed unit ball is not compact. So why is this?

Let e_n equals 1 to infinity b_n -- or in fact, it doesn't have to be separable. In fact, if you just take any infinite-dimensional Hilbert space, you can find by the-- we don't need to make it separable. You can always find an orthonormal, countably infinite orthonormal, subset of H .

Now, in the separable case, we could choose it so that it's a orthonormal basis. But that really doesn't matter. Just the fact that H is an infinite-dimensional Hilbert space, we can find a countable subset.

So we can find a countably infinite subset of orthonormal vectors. Why is this? Because we can find a countable set of linearly independent vectors-- an infinite-dimensional Hilbert space or infinite-dimensional vector space.

And then we can make them orthonormal by applying the Gram-Schmidt process to that collection, and come up with countably infinite orthonormal subset. Then, for all n not equal to k , we get that $\|e_n - e_k\|^2$ - this is equal to, again, using the parallelogram law-- well, I don't need to apply the parallelogram law. I can just apply what this is.

This is equal to $\|e_n\|^2 + \|e_k\|^2 - 2 \operatorname{Re} \langle e_n, e_k \rangle$. And since they're orthogonal, this is 0. And since they're normalized to 1, this is equal to 2.

So for any two distinct elements in the sequence, the distance between the two of them is equal to $\sqrt{2}$. So there's no way the sequence of these orthonormal vectors can have a convergent subsequence. Because a convergent subsequence would mean that if I go along the subsequence, or this subsequence has to be Cauchy so that the difference between vectors has to get small. But it's never small. It's always at least bigger than or equal to $\sqrt{2}$, at least the length squared or distance squared.

So this is a closed and bounded subset which is not compact. There exist sequences which have no-- or this sequence has no convergent subsequence.

So the question is, what's the additional condition I need to conclude that a subset of a metric space is compact? From Real Analysis 18.100, every compact set has to be closed and bounded. One implication is that if something is compact, it's closed and it's bounded.

What additional condition on the set-- so being closed and bounded is a necessity of the set. What additional condition guarantees that it's compact? Maybe not if and only if, like we have for the Heine-Borel, but at least something simple to verify so that the subset of the Hilbert space is compact.

Now, if you didn't see this in 18.100, that's fine. But this question can or could have been already asked in a different context in 18.100. And at least when I taught it, 100B, it was that you can ask the same question about subsets of continuous functions.

So definition of a compact set is with respect to a metric space. Doesn't have to be an inner product space. So the space of continuous functions on, say, $[0, 1]$ -- that's a perfectly good metric space with the infinity norm.

You could ask, what are the conditions on a set to ensure that the set is compact? And the three conditions are closed, bounded-- because these are necessary if the set is to be compact-- and what's called "equicontinuous." So for subsets of the space of continuous functions, if the subset is closed, bounded, and what's called equicontinuous, then the subset is compact. This is the famous Arzela-Ascoli theorem.

And that extra condition, equicontinuity, is in some sense what allows you to reduce your problem of showing your space is compact to the finite-dimensional case. It takes care of the infinite dimensions or it takes care of the infinite part, in a sense. That's not very specific, but it's needed to control all but finitely many things, in some sense. Equicontinuity in Arzela-Ascoli is what helps you control all but finitely many things.

Now, there is a similar theorem that you can prove, now in Hilbert spaces. Let's just state the theorem. Part of the theorem will be a definition, unfortunately, but let H be a Hilbert space.

Well, we're not going to prove the theorem I want to prove just yet. I just skipped ahead by one. So we'll motivate a certain condition. So I got ahead of myself just a minute by just one theorem.

But let me introduce the definition. So let H be a Hilbert space. A subset K of H sets has equismall tails with respect to a countable orthonormal subset.

So the new bit of terminology is "equismall tails--" e_n if-- let's make this K -- in some sense-- so by Bessel's inequality, the sum of the, if you like, Fourier coefficients or the sum of the squares of the inner product e_n with a fixed vector-- this is always bounded by the norm of the vector squared. And therefore, it always converges. So equismall tails means that somehow, that series converges, or the tail end of that series converges, uniformly with respect to this orthonormal subset.

So in other words, for all ϵ positive, there exists a natural number N such that for all v in the set K , if I look at the tail, this is less than ϵ squared. So somehow, given ϵ , I can always choose an N which is independent of the element in K so that the tail end of, if you like, the Fourier series is small. And this N can be chosen independent of the element in K .

So I leave it to you to verify if K is simply a finite set, this implies K has equismall tails with respect to any orthonormal subset. So what's the motivation for this definition is that the next simplest type of set which is compact. So like I said, any finite subset of a metric space is compact.

Now, I'll give an example of another, more interesting compact set. But that compact set also satisfies this property, that it has equismall tails with respect to any countable orthonormal subset. And so then, we should hope that if a set is closed, bounded, and has the property that, with respect to an orthonormal basis it has equismall tails, then that set is compact. And that's eventually what will prove, but I don't think we'll get to it today.

But so one more bit of motivation on why this might be the right extra condition to add to closed and bounded to produce compact sets is provided by the following. So again, let H be a Hilbert space and be in a convergent sequence. Now let e_n be some arbitrary orthonormal, countable orthonormal subset.

So then two things hold-- one is that the set K equals a set of elements v_N union the limit. This is compact. The proof is given in [? Mel ?] [? Rose's ?] notes, and I'm going to leave it to you to look up or work your way through. It's not that difficult to work your way through.

And K has equismall tails with respect to ϵ . So this subset consisting of a convergent sequence along with this limit is compact. And it has equismall details with respect to any countable orthonormal subset.

So like I said, I'm going to leave one for you. 2 we will prove. How much time do I got? I don't have much time left, but we should be able to finish this.

We have to verify this definition. So let ϵ be positive. So why should we-- what's the thinking on why we expect this set to have equismall tails? So I have to check that I can find a large enough integer that all the squares of these coefficients is small.

Now, the point is that the v_N 's are essentially very close to being v , at least for N very large. And for a fixed v , I can always choose a capital N so that this holds. And so that N will take care of infinitely many, and then I just need to choose N big enough for v , and also the finitely many that are not very close-- finitely many elements of this sequence which are not close to v . That's essentially the argument.

So since v_N converges to v , there exists a natural number M such that all N bigger than or equal to M , I have $\|v_N - v\|$ is less than $\epsilon/2$. Now, choose natural number N so large so that $\sum_{k \text{ bigger than } N} v_k^2$ is small, so here, v is fixed. And we always have that this whole sum converges by Bessel's inequality.

And therefore, we can always choose an N so that the tail is small. And we can do the same for the first $M - 1$ elements of the sequence, so plus \max of 1, so that this is all less than $\epsilon^2/4$. So for each v_N , I can find a capital-- so for v_1 , I can find an N_1 so that this is less than $\epsilon^2/8$, say.

And then for N equals 2, N equals 3, all the way up to $N - 1$, I can find a capital $N_{sub\ n}$. I then take capital N to be the maximum of those finitely many capital $N_{sub\ n}$'s along with the N that I need for this. So that's the argument on why I can find such a capital N . I just have finitely many things I need to make small, which I individually can.

Now, I claim this capital N works. So I have to show that if I take any element in K , then it has a small tail. Then I have that, by how I've chosen N , $\sum_{k \text{ bigger than } N} v_k^2$ is less than $\epsilon^2/4$, which is less than ϵ^2 . And for all $1 \leq n \leq N - 1$, the same bound.

So I just need to check now that this is small when n is bigger than or equal to M . Now, if n is bigger than or equal to M , by Bessel's inequality, we have that $\sum_{k \text{ bigger than } n} v_k^2 \leq \epsilon^2/4$, if I can take the $1/2$ power and show that's less than ϵ . So this is equal to $\sum_{k \text{ bigger than } N} v_k^2 + \sum_{k \text{ bigger than } n} v_k^2$.

So I can-- $\epsilon^2/4 + \sum_{k \text{ bigger than } n} v_k^2$ so let's write it this way. Now, this is the little l_2 norm of a sum of two sequences in K , not N , but in K . So by the triangle inequality for sequences in little l_2 , this is less than or equal to $\sum_{k \text{ bigger than } N} v_k^2 + \sum_{k \text{ bigger than } n} v_k^2$ plus $\sum_{k \text{ bigger than } n} v_k^2$ to the $1/2$ power.

Now, this quantity I already have control over. That's less than $\epsilon/2$. I'm taking the $1/2$ power here. And I can bound this by Bessel's inequality. This is less than or equal to $\|v_N - v\|^2$ and then using the bound for this-- $\epsilon/2$.

And now remember, how is capital M chosen? Capital M is chosen to ensure that this is less than $\epsilon/2$, which we can do since the v_N 's are converging to v . So this is less than $\epsilon/2$ plus $\epsilon/2$ equals ϵ , proving that this set consisting of the elements of a convergent sequence along with the limit, has equismall tails with respect to an arbitrary orthonormal subset.

So next time, we will prove that if we have a subset of a separable Hilbert space which is closed, bounded, and has equismall details with respect to A orthonormal basis, which exists because it's a separable Hilbert space, then the set is compact. And then, we'll rephrase that in a way that doesn't involve Hilbert spaces, and go from there, and start looking at some of these operators I said, which are close to matrices. All right, we'll stop there.