

[SQUEAKING]

[RUSTLING]

[CLICKING]

CASEY OK, so let's continue our discussion about spectral theory for self adjoint compact operators. So let me just briefly recall the spectrum of a bounded operator, which was supposed to be a generalization of the eigenvalues of a matrix.

So we defined the resolvent set of A to be those complex numbers such that A minus λ times the identity, which I just write is $A - \lambda I$, is an invertible, bounded linear operator, meaning it is bijective and, which by the open mapping theorem, tells you that the inverse is also continuous.

And the spectrum of A is simply those λ so that $A - \lambda I$ is not invertible, so the complement of resolvent set of A . So from linear algebra, you have the following characterization of what the spectrum is, that if H is $\mathbb{C}^n$, and A is just therefore matrix on $\mathbb{C}^n$, or $\mathbb{R}^n$ if you like, then the spectrum is just simply the set of eigenvalues of A .

And if we restrict our attention to Hermitian matrices, which are referred to as-- which are also referred to as self adjoint also in linear algebra, or symmetric if we're just looking at real vector spaces $\mathbb{R}^n$ then the eigenvalues are real.

And you can find an orthonormal basis of the space $\mathbb{C}^n$ or $\mathbb{R}^n$, depending on what you're looking at for this symmetric matrix so that in that basis, the matrix is completely diagonalized. and. What are the diagonal elements there? The eigenvalues of the matrix.

And what we're going to end up proving is that that picture, this picture where for a self adjoint matrix on $\mathbb{C}^n$, that the spectrum is given by the eigenvalues. And you can diagonalize, essentially diagonalize this operator, meaning you can find an orthonormal basis consisting entirely of eigenvectors of the operator is also true for self-adjoint compact operators.

This shouldn't come too much of a surprise, as too much of a surprise, since compact operators are limits of finite rank, i.e. matrices in the space of bounded linear operators. So that's where we're headed.

Now, so just to follow up on this and for a little bit of review, so in the finite dimensional case, the spectrum can be-- or the spectrum is always just the eigenvalues of the matrix A -- not so in the infinite dimensional setting, for example, if we're looking at l_2 . And then A times a is equal, to let's say, a_1 over 1, a_2 over 2, and so on, for a , a sequence in l_2 .

Then what you can prove is that 0 is in the spectrum of A . One way to see that is that each of the basis vectors of H given where you just have 1 in the n -th slot and 0 otherwise, each of those is an eigenvector of this operator A with eigenvalue $1/n$, where the n tells you where the 1 is and 0 otherwise.

So $1/n$ is an eigenvalue of this operator for each n . And $1/n$ convergence to 0. And since the spectrum of a bounded linear operator is a compact set, in particular closed, 0, which is the limit of that sequence, has to also be in the spectrum.

And, in fact, what we're going to show is that in the nondegenerate case, what we see here is what in general happens for compact self-adjoint operators, that if it's not a finite rank operator, then it has countably many-- or countably infinite many-- that's not a very good string of words-- countably infinite eigenvalues which converge to 0. And 0 may be an eigenvalue, may not be.

And what's more, that's going to be the only that completely characterizes the spectrum. So let me just write here for this example that, in fact, the spectrum of A is equal to the point 0 union $1/n$ in a natural number.

And these are-- again for this operator here, these are the eigenvalues. And this is just-- well, 0 is 0. But 0 is not an eigenvalue, OK. Now, this is not the general picture, meaning that for a compact self-adjoint operator, you'll have infinitely many eigenvalues. And then 0 will not be an eigenvalue. You could have 0 in eigenvalue as well.

But in general what the picture is is that anything in the spectrum that's not 0 has to be an eigenvalue. And that's essentially what will prove our first result of this lecture is the following. So this is the Fredholm alternative. So let A be a self-adjoint compact operator and λ be a nonzero real number.

Then the range of $A - \lambda$ is closed. So this is the conclusion. And thus the range of $A - \lambda$ is equal to the orthogonal complement of the orthogonal complement of itself, which-- the orthogonal complement of the range of $A - \lambda$ is equal to the null space of the adjoint. And since A is self-adjoint and λ is a real number, the adjoint is just $A - \lambda$.

So this right here is the main conclusion from which you get this. And therefore just to spell out the conclusion of this equality, therefore either one of two alternatives happens. So that's the name, alternative-- either $A - \lambda$ is bijective, or the null space of $A - \lambda$, which is just the eigenspace of $A - \lambda$, or I should say the eigenspace corresponding to λ , is non-trivial and finite dimensional.

Moreover, this equality here tells you when you can solve the equation $(A - \lambda)u = f$. You can solve $(A - \lambda)u = f$ if and only if f is in the range of $A - \lambda$, which is if and only if f is orthogonal to the null space of $A - \lambda$. So let me make a little remark here.

So first off, the fact that this space has to be finite dimensional we proved last time. We proved at the end of lecture, the last lecture, that for a compact self-adjoint operator, the null space, the eigenspace corresponding to a given nonzero eigenvalue, is finite dimensional. And then we also prove the eigenspaces corresponding to two different eigenvalues are orthogonal. And we also prove that the eigenvalues that are nonzero, or any eigenvalue, has to be real.

So now let me make a couple of remarks. The first is just a rephrasing of what's in the theorem. Therefore f is in the range of $A - \lambda$, meaning you can solve the equation $(A - \lambda)u = f$ if and only if f is in the null space of-- or the orthogonal complement of $A - \lambda$ -- or the null space of $A - \lambda$.

So what this says is that you can solve for f , given that f satisfies finitely many linear conditions because the null space of $A - \lambda$ is finite dimensional. So f being orthogonal to that means pick a finite base, a finite orthonormal basis of $A - \lambda$ null space of $A - \lambda$. Then f in our product with those finitely many vectors has to be 0.

So you have finitely many conditions on f to be able to solve for f . And not only that, this solution that you compute is also unique up to finitely many conditions because, again, the null space of $A - \lambda$ is-- or unique up to a finite dimensional subspace, again, because the null space of $A - \lambda$ is finite dimensional.

And the second is that for a self-adjoint operator, we have that the spectrum is a subset of the real numbers. For a self-adjoint operator, this is something we proved last lecture. It doesn't have to be compact, just a self-adjoint operator.

So since the spectrum is a subset of the reals, this proves that for a compact self-adjoint operator, the spectrum of A is equal to the set of eigenvalues of A , or I should say nonzero eigenvalues of A . Let's write it this way.

If I look at what's in the spectrum other than 0 possibly, then the nonzero numbers that are in the spectrum have to be eigenvalues. So for a compact self-adjoint operator, and just in the case of matrices, the spectrum, the nonzero spectrum has to be-- are nothing but eigenvalues.

And last time, remember, we proved that the eigenvalues are countably infinite, or countable. They're either finite or countably infinite. And if they're countably infinite, they converge to 0. So, again, by last lecture, we conclude that the spectrum of A take away 0 equals either finitely many eigenvalues or countably infinite eigenvalues that are converging to 0.

So from the Fredholm alternative we get a lot of information about when we can solve equations. But it also tells us-- I mean from that ability to say when we can solve the equations, we can also characterize the nonzero spectrum of a self-adjoint compact operator.

So we need to prove that-- so remember all of this just followed from stuff we had proven and the main conclusion of the theorem, which is that the range of $A - \lambda$ is closed. So we need to prove that the range of $A - \lambda$ is closed when λ is a nonzero real number.

So suppose that you have a sequence in the range, which I'll write as $A - \lambda$ times u_n converging to some element f in H . So what we'd like to be able to show is that f is in the range. So we want to show f is in the range of H -- or not range of H , range of $A - \lambda$, sorry. Hopefully I didn't make that mistake elsewhere-- no, just-- OK.

Now, we're only assuming that $A - \lambda$ when it hits u_n converges to f . We're not a priori assuming the use of u_n 's converge. In fact, we can't. But in the end we would like to come up with maybe a subsequence or a part of the u_n 's up to a subsequence which does converge, and then we conclude that f is in the range.

So first I want to get rid of the useless part of the u_n 's. So let W be-- well, I don't need to give it a name. Really, where is my eraser? So let v_n be the projection onto orthogonal complement of the null space of $A - \lambda$ of u_n .

So now this is just part of-- so every u_n is written-- you can write as something in the null space of $A - \lambda$. So since null space of $A - \lambda$ is a closed subspace of H , it has an orthogonal complement so that it and its orthogonal complement gives the direct product-- or when you take their direct product, gives you H .

So why am I saying that? Because then if I take $A - \lambda u_n$, this is equal to $A - \lambda$ applied to u_n , so the projection onto the null space of $A - \lambda$ plus the projection onto the orthogonal complement of the null space, which I defined as v_n .

Now, this element here is in the null space of $A - \lambda$. So when $A - \lambda$ hits it, I get 0. So I get $A - \lambda$ applied to v_n , all right. Then $A - \lambda v_n = A - \lambda u_n$, which converges to f .

So, basically, I've taken away some noise, all right, the part that when it hits $A - \lambda$, I get 0. So now I just have these v_n 's which lie in the orthogonal complement of the null space of $A - \lambda$.

So my claim is, first, is that the sequence v_n is bounded. Basically, once I can show this, then I'm done because if I can show v_n is bounded, then since A is a compact operator, when A hits v_n up to a subsequence, that converges.

Now, this whole expression converges and therefore λv_n converges. λ is nonzero, so then v_n converges up to a subsequence to something. And therefore $A - \lambda v_n$ then converges to $A - \lambda v$ for some v , which shows that f is in the range of $A - \lambda$. So this is really the whole ball game.

And we'll use here crucially as well that we threw away useless parts of u_n , useless at least to this argument. So I claim this is bounded, so suppose not. Then there exists subsequence v_{n_j} such that $\|v_{n_j}\|$ goes to infinity as j goes to infinity.

All right, now if I look at $A - \lambda$ applied to v_{n_j} over norm of v_{n_j} , this converges-- so first off, since A is a linear operator, this is equal to $\frac{1}{\|v_{n_j}\|} (A - \lambda) v_{n_j}$. And so this scalar, $\frac{1}{\|v_{n_j}\|}$ converges to 0. This converges to f . So I get the 0 vector in H .

So this thing converges to 0 in the Hilbert space H . Now, why is that bad? Because essentially what this is going to say is that there exists some element, or that this sequence converges at least up to a subsequence to an element v with norm 1 because they all have norm 1 so that $(A - \lambda)v = 0$. But all of these are in the null space of $A - \lambda$ and we get a contradiction.

So we have that $A - \lambda$ -- so we have that part. Since A is a compact operator, there exists a subsequence, so a further subsequence-- I'm just going to call it v_{n_k} instead of v_{n_j} -- v_{n_k} such that the sequence v_{n_k} converges.

But then I get that v_{n_k} -- or I should say $\frac{v_{n_k}}{\|v_{n_k}\|}$. So these all have norm 1. And A applied to something that has unit length-- or the image by A of the closed unit ball is a precompact-- or the closure of it is compact. And therefore every sequence has a convergent subsequence.

Then $\frac{v_{n_k}}{\|v_{n_k}\|} \cdot \frac{1}{\lambda} (A - \lambda I)v_{n_k}$, this is equal to $\frac{1}{\lambda} (A - \lambda I)v_{n_k}$ over norm of v_{n_k} minus $A - \lambda I$ applied to v_{n_k} . Now this sequence of elements converges to 0. That's, in fact, what we just proved. And this converges by how we've taken this subsequence because A is a compact operator.

So I have this sequence of vectors is equal to-- and here we can divide by λ because λ is nonzero. It's equal to a linear combination of two sequences which converge. And therefore we get that $\frac{v_{n_k}}{\|v_{n_k}\|}$ converges to an element v .

And now the null space or the orthogonal complement of the null space of $A - \lambda I$, each of these is in-- remember the orthogonal complement of the null space. And since it's converging to an element and this is closed, this element has to be in the same set. So, again, this follows from-- the fact that v is in here is because this set is closed. The orthogonal complement of any subset of a Hilbert space is closed.

Then by continuity of the norm, basically, the norm of v has to be equal to limit as k goes to infinity of the norm of the elements converging to it, which all equal 1. And if I compute $(A - \lambda I)v$, this is equal to-- since the $\frac{v_{n_k}}{\|v_{n_k}\|}$'s are converging to v , $(A - \lambda I)v_{n_k}$ over norm v_{n_k} is converging to $(A - \lambda I)v$ over norm v .

And, remember, this is a subsequence of v_{n_j} 's. And when $A - \lambda I$ hits that, they're converging to 0, so-- all of that predicated upon the assumption that the norm of v_{n_k} 's converges to infinity or that the sequence is unbounded.

So we have this element in the null space that has-- or the orthogonal complement of the null space that has norm 1 but also gives you 0. And therefore we get that v is in the null space of $A - \lambda I$ from this computation and it's orthogonal complement.

But the only possible vector that's in a space and it's orthogonal complement is the 0 vector, or the only vector in a subspace and its orthogonal complement is the zero vector. And therefore v equals 0, which this is a contradiction to the fact that the norm of v equals 1.

So we started off with the sequence v_n 's. Assuming that they are unbounded, $\frac{v_n}{\|v_n\|}$ is a sequence of essentially-- yeah. So maybe you got lost in the subsequences. But let's just assume I'm talking about the entire sequence.

Then $\frac{v_n}{\|v_n\|}$, when $A - \lambda I$ hits it, converges to 0. Since A is a compact operator, we can show essentially that A applied to $\frac{v_n}{\|v_n\|}$ converges to something, because those things have unit length.

And since λ is nonzero, we can then conclude that those vectors converge, in fact, to something, not just their images by A or their images by $A - \lambda I$, again, because λ is nonzero.

And since they all have unit length, their limit must have unit length. And since when $A - \lambda I$ hits these guys they go to 0, the limit must also, when $A - \lambda I$ hit it, equals 0. And that gives us our contradiction because this limit v had to be in the orthogonal complement of the null space, but then also in the null space, and have unit length. Those three things can all happen at once.

So thus the sequence v_n is bounded. So, remember, what were the v_n 's to start with? They were so that $A - \lambda I$ applied to v_n converged to this element f . And we wanted to show that f is in the range.

So we coming back over here. We had these v_n 's so that $A - \lambda I$ applied to v_n converges to f . And we want to show f is in the range to conclude that the range of $A - \lambda I$ is closed. But now since it's bounded and we've done this argument already, this is essentially the whole ball game.

Since the subsequence v_{n_j} is bounded, and A is a compact operator, we conclude that there exists a subsequence $v_{n_{j_k}}$ such that-- so this has nothing to do with the previous argument now, but I just don't feel like using different letters-- such that A applied to $v_{n_{j_k}}$ converges.

So, remember, A is a compact operator, which we stated it in terms of the closure of the image of the closed unit ball being compact. Equivalently by scaling the unit ball, it means that A takes any bounded sequence to a sequence that has a convergent subsequence.

So we showed v_n is bounded. And therefore since A is a compact operator, we can find a subsequence so that when A hits it, we have a convergence sequence. And by that same trick we used a minute ago, we conclude that $v_{n_{j_k}}$, which is-- and we can do this because λ is nonzero. We can divide by it.

So $v_{n_{j_k}} - \lambda^{-1}(A - \lambda I)v_{n_{j_k}}$ now, again, this here is converging to something. This here is converging to f . So this linear combination of convergent sequences is convergent-- converges to an element v .

And therefore I get that f which is the limit as n goes to infinity of the v_n 's. But convergence still holds if I look at a subsequence, $(A - \lambda I)v_{n_{j_k}}$ equals-- and since A -- so A is a bounded linear operator. λI times the identity is a bounded linear operator. This is equal to $(A - \lambda I)v$. And therefore f is in the range of $A - \lambda I$.

OK, so Fredholm alternative tells you that the range of $A - \lambda I$ for a compact self-adjoint operator is closed. So where did we really use the fact that it was self adjoint? Nowhere in this argument. So this fact that the range of $A - \lambda I$ is closed is still true if A is just a compact operator and λ is just some nonzero complex number.

But where we use that it's self adjoint is I guess in the rest of the conclusion, that the range of $A - \lambda I$ is therefore equal to the null space of $A - \lambda I$. And therefore either $A - \lambda I$ is bijective, or the null space of $A - \lambda I$ is non-trivial and finite dimensional by what we did in the previous lecture.

OK, now, again, this is a very powerful theorem. And, again, what this says is that if I look at the nonzero spectrum of a compact self-adjoint operator, then that consists entirely of eigenvalues of A .

Now earlier we proved that plus or minus the norm of A has to be in the spectrum of a self-adjoint operator. And therefore what we can conclude is that if we have a non-trivial self-adjoint compact operator, then it has at least one eigenvalue.

And we can characterize that eigenvalue. So this has a following theorem. Let A be a non-trivial compact self-adjoint operator, $A = A^*$. Then A has a non-trivial eigenvalue λ .

And we can characterize λ_1 or at least the absolute value of λ_1 as the supremum over norm u equals 1 Au , u equals-- and this supremum is actually achieved where u_1 is a normalized eigenvector corresponding to λ_1 .

So why do we have-- or I should say, let me make sure I have everything here. OK, so why is this? So first off, we've shown that plus or minus the norm of A is, in fact, in the spectrum of A . for any self-adjoint bounded linear operator, not necessarily compact, that plus or minus-- not necessarily both of them, but plus or minus, one of these, at least one of plus or minus, since A is self adjoint, meaning $A^* = A$.

Then λ_1 -- then I'm going to say plus or minus, meaning not both, actually at least one of these, is an eigenvalue of A by the Fredholm alternative. The Fredholm alternative and the fact that-- so λ_1 is going to be either plus norm of A or minus norm of A , depending on whether which one is in the spectrum. Let's say plus if it's in the spectrum, minus if plus is not.

And the fact that we can identify it as this quantity here is because we have that for self-adjoint operators, norm of A , which is the absolute value of one of those plus or minuses that are in the spectrum is equal to-- so this is where-- so this is equal to $\sup_{u \text{ equals } 1} |Au|$.

So that's the end of the proof. And now what we're going to do is we are going to keep going. So the end result will be that we can determine all of the eigenvalues via a certain maximum principle and build up a sequence of eigenvectors, normalized eigenvectors, which are pairwise orthogonal simply because they come from how they're built up. But we'll see.

And what? And then we'll show that essentially that set of eigenvectors that we get, along with a set of-- or an orthonormal basis chosen for the null space of A , form an orthonormal basis for a separable Hilbert space H .

So that's where we're going. But we can basically take this theorem and keep applying it. So we have the following maximum principle. If you like, this is the first step in a maximum principle. This says if you want to find the largest eigenvalue, λ_1 -- I should say-- I can even say largest eigenvalue, largest in the sense of absolute value.

Why? Because remember the spectrum is contained in the interval minus norm A plus norm of A . So anything in the spectrum has to have absolute value less than or equal to the norm of A . And if it's eigenvalue, or if it's a-- anything other than 0 has to be an eigenvalue so we get this.

So what's this maximum principle? And why was I saying all that? Oh, so this gives you a way to-- if you'd like try and find or at least approximate the first eigenvalue of a bounded linear operator, this is a maximization problem with a constraint. So you could use the method of Lagrange multipliers.

That doesn't maybe make sense to you to be able to do on an infinite dimensional Hilbert space. But let's say you choose a big basis of your Hilbert space. And you just restrict to looking at that big but finite dimensional, or finite basis, or span of that finite basis, and try to solve the approximate problem, then you should get close to the eigenvalue and get an approximate eigenvector because as I've said, the eigenvector will be achieved-- or I should say that the eigenvector achieves this maximum here, or the supremum.

So the maximum principle is the following. So let A -- again, we're only looking at compact self-adjoint operators, compact operator-- then the nonzero eigenvalues of A can be ordered.

Now, this part we already know. They can be ordered $\lambda_1 \leq \lambda_2 \leq \lambda_3$, counted with multiplicity meaning if λ_1 has a two-dimensional eigenspace then λ_2 or the absolute value of λ_2 will be λ_1 . So we'll repeat it according to multiplicity.

So we know that there's finitely many distinct eigenvalues. That was the part that I was saying. We know we can order them, but maybe it's not clear that you can order them with multiplicity with corresponding orthonormal eigenbasis functions u_k . So u_1 is a normalized eigenfunction for λ_1 . u_2 will be a normalized function for λ_2 , which is orthogonal to u_1 . So these are pairwise orthonormal.

And how do we obtain the eigenvalues in this order and these eigenfunctions via the following process so that λ_j is equal to the supremum over all unit vectors that are orthogonal to the first $j-1$? And this is equal to-- this will be achieved on the u_j .

So we have the first one, if you like. We built up λ_1 , u_1 . and. Now what this maximum principle theorem is saying is that we can repeat this, is that if we now look at this quantity here, the supremum over all norm 1 vectors that are orthogonal to u_1 , then we're going to pick up the next largest eigenvalue counted with multiplicity, meaning it'll be the absolute value of λ_2 , which will be less than λ_1 if λ_1 only has a one dimensional eigenspace.

Or it'll be λ_1 or the absolute value of λ_2 will be the absolute value of λ_1 again if λ_1 has a, let's say, two-dimensional eigenspace. And if it had a three-dimensional eigenspace, we would get the absolute value of λ_1 for-- it would equal this number and this number as well. So, again, these nonzero eigenvalues can be ordered in this way. Oh, and I left off one. And the λ_j is going to 0.

OK, so we have the first one, the absolute value of λ_1 , and the first eigenvector u_1 . and. Now we're just going to basically repeat or apply the previous theorem to a modification of A . Now I should say this is not entirely true because there may only be finitely many.

So really this should be a remark, if the decreasing sequence does not terminate, then the absolute value of the λ_j 's goes to 0. Now, so what's the new bit of information here? We do know that we can-- for each, let's say, capital N , the number of eigenvalues outside of-- or with absolute value bigger than $1/N$ has to be finite.

So the fact that we can order them is not really that much new information. And the fact that they go to 0 if this sequence is infinite, that's also not new information. I mean, we did prove that in a previous lecture that if you have-- so there we proved it for a sequence of distinct eigenvalues.

But because we now know that each of these has a finite dimensional eigenspace, if you just look back at the same proof or that proof, you can make a small adjustment to be able to say that if you count the eigenvalues with multiplicity, then also the sequence has to go to 0, not just necessarily the sequence consisting of the distinct eigenvalues. So let me make that small remark.

What the new piece is that we can compute the eigenvalues in this way and choose the eigenvectors in this way. That's a. New piece and that's what, in the end, we'll use to be able to show that H can be-- or that A can basically be diagonalized, or that you can find an orthonormal basis of a Hilbert space, separable Hilbert space consisting entirely of eigenvectors of A .

So the construction proceeds inductively. So at one point I said, if you can do it for one then you can do it for the next. And then I won't be so formal every time I do an inductive construction. But in this case it pays to be careful. So we'll construct the sequence of λ_j 's and u_k 's in this way via an inductive argument.

So coming up with k equals 1, that was the previous theorem. So previous-- or I should say j equals 1, that's the previous theorem. We found the largest eigenvalue, or the eigenvalue with-- an eigenvalue with the largest absolute value via this theorem. And then we obtained an eigenvector this way, just basically as a consequence of the Fredholm alternative. And then that we know that plus or minus the norm of A has to be in the spectrum.

[the fact that we can find the first eigenvalue of λ_1 , that follows from the previous theorem. So now we want to do the inductive step, i.e. we want to suppose we have found λ_1 up to λ_n -- let's say, what did I use here?

λ_n , so j equals n and with orthonormal eigenvectors u_1 up to u_n , satisfying this maximum principal. So j starting from-- starting at 1 up to n , we've constructed or found a λ_2 , λ_3 , up to λ_n , and the u went up to u_n , satisfying that maximum property.

So then there's two cases. So in case 1 is the fact that A minus-- A is equal to the sum from k equals 1 to n of $\lambda_k u$ inner product u_k, u_k . And therefore what this shows is that we found all the eigenvalues and the process terminates.

So this is kind of the degenerate case that A is a finite rank operator. So I could have started this whole theorem off with let's assume A is not a finite rank operator, and therefore we wouldn't have to deal with this case. But I just stated it for an arbitrary A .

So there is the possibility that that sequence stopped. We found all of the eigenvalues with multiplicity in this process. And then the theorem-- or construction is done in that case. And the case 2 is that it is not a finite rank operator, so it's not equal to k equals 1 to n .

Now we have to show how can we find λ_{n+1} and the eigenvector u_{n+1} . Let A_{n+1} be the operator A minus k equals 1 to n λ_k inner product u_k . So should say Au minus u sub k , all right.

Now. Note since we are in case 2, this operator is nonzero. So we're basically going to apply the previous theorem now to this operator. But let me first make a few remarks. Then A_{n+1} is A sub n is a-- this is something you can check. It's a self-adjoint compact operator. So why is it self adjoint? Basically because A is self adjoint and these λ_k 's are real numbers because eigenvalues have to be real, because A is self adjoint and these are orthonormal.

So that is why they're self adjoint. Why is it a compact operator? Well, it's the sum of a compact operator A and a finite rank operator here. So it's also a compact operator. And it's a nontrivial one because it's not identically equal to 0. So I shouldn't say not equal to 0, but which is not identically the zero operator.

So here's a couple of facts that if u is in the span of u_1 up to u_n , then I get that A_{n+1} applied to u is the zero vector. Why is that? So it suffices to check that this gives me-- this formula gives me 0 when u is one of the u_k 's.

Now if u is one of the u_k 's, then this will be 1 only when $j = k$ or let's say it's one of the u_j 's. I need a different letter. So if u is one of these u_j 's, then this will give me 1 only when j equals k and I pick up λ_j times u_j . And then I have A hitting u_j . And that spits out λ_j times u_j because u_j is an eigenvector of A . So then I get the same thing here and there and subtracting them gives me 0.

Now, if u is orthogonal to the span of the u_1 up to u_n , then I get $A_n u = Au$. So you just see this. If u is orthogonal to these, then this whole term is 0. Can't I just pick up $A_n u = Au$.

For all u in H , in the span of u_1 up to u_n , if I compute A and u inner product v , this is equal to u and v since A is self adjoint. And since v is in the span of the u_1 's up to u_n 's, by the first property that's going to be 0.

So this equals 0 and therefore, what have I shown? I've shown that A and u inner product v is 0 no matter what u is in H . And therefore, that means that v has to be in the orthogonal complement of the range. So I said that backwards.

Anyways, here u is fixed. v is a thing that's changing. So what does this say? u as fixed. A and u inner product v when v is in the span has to be 0. This holds for all v in the span. And therefore $A_n u$ has to be in the orthogonal complement of the span of these guys. So this proves that the range of A_n is a subset of your orthogonal complement of u_1 up to u_n .

And so from this previous property we get one last property, that if $A_n u = \lambda u$ is nonzero, meaning we have a nonzero eigenvalue of A_n , then that implies that u is equal to $1/\lambda A_n u$ applied to u .

In other words, u is equal to $A_n u$ applied to u over λ . So that implies u is in the range of A_n , which again is contained in the orthogonal complement of u_1 up to u_n . And since it's in the orthogonal complement, that means $A_n u = 0$. So any nonzero eigenvalue of A_n has to be a nonzero eigenvalue of A .

Now we just apply the previous theorem to A_n to the next eigenvalue λ_{n+1} and eigenvector u_{n+1} . By the previous theorem, A_n has a nonzero eigenvalue, which I will call λ_{n+1} , with unit eigenvector u_{n+1} so that λ_{n+1} is equal to the sup over all norm-- ah, OK, sup over all norm or unit length vectors of absolute value of $A_n u$ applied to u , inner product u .

Now, this sup is the same sup over all u so that $A_n u$ is nonzero. So that in particular includes-- OK. So if u is in the span u_1 up to u_n , then $A_n u = 0$.

So the supremum over all unit length is the same as the supremum over everything in the orthogonal complement of these guys because these guys, when you stick them into A , and gives me a 0. So that's the same supremum.

But now also when u is in the orthogonal complement of these u_1 's up to u_n , this is equal to sup of u on the orthogonal complement of the u_1 up to u_n 's. $A_n u$, remember, equals Au . So that gives us the fact that I can find this from this supremum.

And so why does this eigenvector has to be in the orthogonal complement of u_1 up to u_n ? it's. Because I can choose it that way because when A sub n hits anything in here, I get 0. So I should say this is also equal to A n apply to u_n plus 1. But that's the same as A applied to u_n plus 1-- u_n plus 1.

And this is less than or equal to the sup over norm u equals 1 u in span u_1 up to u_n minus 1, orthogonal complement Au applied to u , which equals the absolute value of the n -th eigenvalue counted with multiplicity. So we found the next eigenvalue in this sequence of eigenvalues countable multiplicity.

So now, we can conclude the following spectral theorem for compact self-adjoint operators. Let A be a self-adjoint compact operator on a separable Hilbert space H .

Let λ_1 bigger than or equal to λ_2 be corresponding eigenvalues, be the-- eigenvalues-- or I should say nonzero eigenvalues-- of A counted again with multiplicity, so counted with multiplicity, as we've constructed in this theorem which I called the maximum principle, with corresponding orthonormal eigenvectors u_k .

So we have these eigenvectors coming from this process. And the conclusion is this subset of eigenvectors, or the orthonormal eigenvectors is an orthonormal basis for the closure-- or for the range of A . In fact, we can upgrade that the u_k is, in fact, an orthonormal basis for the range of A closure.

And there exists an orthonormal basis, call it f_j , of the null space of A if it's nonzero, such that $u_k \perp f_j$. So, first off, the union of these two sets of orthonormal vectors is then going to be again a subset of orthonormal vectors because all the f_j 's would correspond to the eigenvalue 0. And these u_k 's would correspond to eigenvalues that are nonzero.

And by what we proved last time, any eigenvector for two distinct eigenvalues would have to be orthonormal. So this subset is orthonormal from this subset. But, moreover, is orthonormal like I said, but also an orthonormal basis for H .

So, in other words, I can find an orthonormal basis for H consisting entirely of eigenvectors of this self-adjoint compact operator, OK. So I will have one piece of this basis coming from the null space and the other piece corresponding to nonzero eigenvalues.

So, really, 2 follows from 1. Not sure why I decided to state them separately, but here we are. So proof of 2 will show that u_k is in orthonormal basis for the range of A .

So first off note as we did in the previous proof that the process of obtaining the λ_1 for the eigenvalues and eigenvectors, or orthonormal eigenvectors terminates if and only if A was finite rank.

In other words, there exists an n so that Au is this finite rank operator u in a product u_k , u_k . In this case, if it's a finite rank operator, then the range is contained in u_k , u_k , which is what we wanted to prove for this case, that A is this finite rank operator.

So suppose otherwise, in other words the process does not terminate, so that we have countably infinite many nonzero eigenvalues counted with multiplicity and corresponding orthonormal eigenvectors u_k . So this is a more interesting-- so the eigenvalues are countably infinite.

OK, so now we want to show-- so as in the remark that I made afterwards, you know that these λ_k 's have to go to 0. Now, we want to prove that the u_k 's are an orthonormal basis for the range of A . What does that mean?

By definition, that means that they're in maximal orthonormal subset of the range of A . So we have to show that if something's in the range and it's orthogonal to every one of these eigenvectors, then that thing has to be 0. So the claim is if f is in the range of A , and for all k $\langle f, u_k \rangle = 0$, then f is a 0 vector.

So this is the claim we want to prove. So suppose we have something in the range. That means we can write it as Au and $\langle Au, u_k \rangle = 0$ for all k . Then for all k , if I look at $\lambda_k \langle u, u_k \rangle$, this is equal to-- λ_k is a real number so I can bring it all the way in, get $\lambda_k \langle u, u_k \rangle$. And this is equal to $\langle u, A u_k \rangle$ apply to u_k .

Now, A is self adjoint so I can move this A over here to u . And this is equal to $\langle u, A u_k \rangle = 0$ for all k . So by this maximum principle which we proved just a minute ago we conclude that the norm of f , which is equal to the norm of Au , which is equal to the norm of A minus sum from $k=1$ to n of $\lambda_k \langle u, u_k \rangle^2$, $\langle u, u_k \rangle^2$ applied to u because every one of these numbers is 0, so I haven't subtracted off anything.

So I can write this in terms of this A applied to u , where A was this thing-- or A is this thing, which by the proof of the maximization-- or the maximum principle is less than or equal to $\lambda_n + 1$ times the norm of u because, again, remember this thing here is the supremum over all u 's of unit length of A applied to u .

So $\lambda_n + 1$ is less than-- or this quantity here divided by the norm of u so that you as a unit length is always less than or equal to this quantity here. But now λ_n -- so this-- I had a fixed thing here, norm of f . And I've shown it's less than or equal to $\lambda_n + 1$ times the norm of u . This is a fixed thing. And the λ_n 's are converging to 0.

And thus I started off with something non-negative, less than or equal to something converging to 0. And therefore that thing had to be 0. Therefore f is 0. So this proves, 1, that these eigenvectors are a maximal orthonormal subset of the range of A .

And for 2 we simply note that by 1, we have that the range of A closure, this is since the eigenvectors are in orthonormal basis for the range of A , the closure is contained in-- closure of the spans of the u_k 's, here this is a finite span, which, remember, this is by an assignment exercise, this is equal to $\sum c_k u_k$ such that-- and therefore, this implies that is an orthonormal basis for the range of A closure.

And this is equal to the range of A orthogonal complement, orthogonal complement, which is equal to the null space of-- or the orthogonal complement of the null space of A^* . A^* equals A , so we get that.

So the u_k 's-- the eigenvectors form an orthonormal basis for the orthogonal complement of the null space of A . So once we've chosen orthonormal basis for null space of A , that's it. Since H is separable and the null space of A is a closed subspace of H , null space of A is separable.

And we've proven that every separable Hilbert space, or even just $\mathbb{R}^n$ -- so this is a close one so it's a Hilbert space anyways. But every separable Hilbert space has an orthonormal basis of-- and therefore, again, so since these two-- call this F -- this is an orthonormal basis for the null space of A direct product null space of A orthogonal complement. That equals H .

And just in the nick of time I'm finished. So next time where do we go from here? We'll see some of this applied in a concrete setting of differentiable equations and also discuss the functional calculus.