CHAPTER 3

Hilbert spaces

There are really three 'types' of Hilbert spaces (over $\mathbb{C}$). The finite dimensional ones, essentially just $\mathbb{C}^n$, with which you are pretty familiar and two infinite dimensional cases corresponding to being separable (having a countable dense subset) or not. As we shall see, there is really only one separable infinite-dimensional Hilbert space and that is what we are mostly interested in. Nevertheless some proofs (usually the nicest ones) work in the non-separable case too.

I will first discuss the definition of pre-Hilbert and Hilbert spaces and prove Cauchy's inequality and the parallelogram law. This can be found in all the lecture notes listed earlier and many other places so the discussion here will be kept succinct. Another nice source is the book of G.F. Simmons, "Introduction to topology and modern analysis". I like it – but I think it is out of print.

1. pre-Hilbert spaces

A pre-Hilbert space, H, is a vector space (usually over the complex numbers but there is a real version as well) with a Hermitian inner product

$$\begin{gathered}(,):H\times H\longrightarrow\mathbb{C},\\ (\lambda_1v_1+\lambda_2v_2,w)=\lambda_1(v_1,w)+\lambda_2(v_2,w),\\ (w,v)=\overline{(v,w)}\end{gathered}\tag{3.1}$$

for any v_1, v_2, v and $w\in H$ and λ_1, $\lambda_2\in\mathbb{C}$ which is positive-definite

$$(v,v)\ge 0,\ (v,v)=0\Longrightarrow v=0.\tag{3.2}$$

Note that the reality of (v,v) follows from the second condition in (3.1), the positivity is an additional assumption as is the positive-definiteness.

The combination of the two conditions in (3.1) implies 'anti-linearity' in the second variable

$$(v,\lambda_1w_1+\lambda_2w_2)=\overline{\lambda_1}(v,w_1)+\overline{\lambda_2}(v,w_2)\tag{3.3}$$

which is used without comment below.

The notion of 'definiteness' for such an Hermitian inner product exists without the need for positivity – it just means

$$(u,v)=0\ \forall\ v\in H\Longrightarrow u=0.\tag{3.4}$$

Lemma 21. *If H is a pre-Hilbert space with Hermitian inner product $(,)$ then*

$$\|u\|=(u,u)^{\frac{1}{2}}\tag{3.5}$$

is a norm on H.

PROOF. The first condition on a norm follows from (3.2). Absolute homogeneity follows from (3.1) since

$$\|\lambda u\|^2 = (\lambda u, \lambda u) = |\lambda|^2 \|u\|^2. \tag{3.6}$$

So, it is only the triangle inequality we need. This follows from the next lemma, which is the Cauchy-Schwarz inequality in this setting – (3.8). Indeed, using the 'sesqui-linearity' to expand out the norm

$$\begin{aligned} \|u+v\|^2 &= (u+v, u+v) \\ &= \|u\|^2 + (u,v) + (v,u) + \|v\|^2 \le \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\ &= (\|u\| + \|v\|)^2. \end{aligned} \tag{3.7}$$

□

LEMMA 22. *The Cauchy-Schwarz inequality,*

$$|(u,v)| \le \|u\|\|v\| \ \forall\ u, v \in H \tag{3.8}$$

holds in any pre-Hilbert space.

PROOF. For any non-zero u, $v \in H$ and $s \in \mathbb{R}$ positivity of the norm shows that

$$0 \le \|u + sv\|^2 = \|u\|^2 + 2s \operatorname{Re}(u,v) + s^2 \|v\|^2. \tag{3.9}$$

This quadratic polynomial is non-zero for s large so can have only a single minimum at which point the derivative vanishes, i.e. it is where

$$2s\|v\|^2 + 2\operatorname{Re}(u,v) = 0. \tag{3.10}$$

Substituting this into (3.9) gives

$$\|u\|^2 - (\operatorname{Re}(u,v))^2/\|v\|^2 \ge 0 \Longrightarrow |\operatorname{Re}(u,v)| \le \|u\|\|v\| \tag{3.11}$$

which is what we want except that it is only the real part. However, we know that, for some $z \in \mathbb{C}$ with $|z| = 1$, $\operatorname{Re}(zu, v) = \operatorname{Re} z(u,v) = |(u,v)|$ and applying (3.11) with u replaced by zu gives (3.8). □

2. Hilbert spaces

DEFINITION 15. A Hilbert space H is a pre-Hilbert space which is complete with respect to the norm induced by the inner product.

As examples we know that $\mathbb{C}^n$ with the usual inner product

$$(z, z') = \sum_{j=1}^{n} z_j \overline{z'_j} \tag{3.12}$$

is a Hilbert space – since any finite dimensional normed space is complete. The example we had from the beginning of the course is l^2 with the extension of (3.12)

$$(a, b) = \sum_{j=1}^{\infty} a_j \overline{b_j},\ a, b \in l^2. \tag{3.13}$$

Completeness was shown earlier.

The whole outing into Lebesgue integration was so that we could have the 'standard example' at our disposal, namely

$$L^2(\mathbb{R}) = \{u \in \mathcal{L}^1_{\text{loc}}(\mathbb{R}); |u|^2 \in \mathcal{L}^1(\mathbb{R})\}/\mathcal{N} \tag{3.14}$$

where $\mathcal{N}$ is the space of null functions. and the inner product is

$$(u, v) = \int u\overline{v}. \tag{3.15}$$

Note that we showed that if u, $v \in \mathcal{L}^2(\mathbb{R})$ then $uv \in \mathcal{L}^1(\mathbb{R})$.

3. Orthonormal sets

Two elements of a pre-Hilbert space H are said to be orthogonal if

$$(u, v) = 0 \Longleftrightarrow u \perp v. \tag{3.16}$$

A sequence of elements $e_i \in H$, (finite or infinite) is said to be *orthonormal* if $\|e_i\| = 1$ for all i and $(e_i, e_j) = 0$ for all $i \neq j$.

Proposition 20 (Bessel's inequality). *If* e_i, $i \in \mathbb{N}$, *is an orthonormal sequence in a pre-Hilbert space* H, *then*

$$\sum_i |(u, e_i)|^2 \leq \|u\|^2 \ \forall \ u \in H. \tag{3.17}$$

Proof. Start with the finite case, $i = 1, \ldots, N$. Then, for any $u \in H$ set

$$v = \sum_{i=1}^{N} (u, e_i) e_i. \tag{3.18}$$

This is supposed to be 'the projection of u onto the span of the e_i'. Anyway, computing away we see that

$$(v, e_j) = \sum_{i=1}^{N} (u, e_i)(e_i, e_j) = (u, e_j) \tag{3.19}$$

using orthonormality. Thus, $u - v \perp e_j$ for all j so $u - v \perp v$ and hence

$$0 = (u - v, v) = (u, v) - \|v\|^2. \tag{3.20}$$

Thus $\|v\|^2 = |(u, v)|$ and applying the Cauchy-Schwarz inequality we conclude that $\|v\|^2 \leq \|v\|\|u\|$ so either $v = 0$ or $\|v\| \leq \|u\|$. Expanding out the norm (and observing that all cross-terms vanish)

$$\|v\|^2 = \sum_{i=1}^{N} |(u, e_i)|^2 \leq \|u\|^2$$

which is (3.17).

In case the sequence is infinite this argument applies to any finite subsequence, e_i, $i = 1, \ldots, N$ since it just uses orthonormality, so (3.17) follows by taking the supremum over N. □

4. Gram-Schmidt procedure

Definition 16. An orthonormal sequence, $\{e_i\}$, (finite or infinite) in a pre-Hilbert space is said to be *maximal* if

$$u \in H, \ (u, e_i) = 0 \ \forall \ i \Longrightarrow u = 0. \tag{3.21}$$

Theorem 12. *Every separable pre-Hilbert space contains a maximal orthonormal set.*

Proof. Take a countable dense subset – which can be arranged as a sequence $\{v_j\}$ and the existence of which is the definition of separability – and orthonormalize it. Thus if $v_1 \neq 0$ set $e_i = v_1/\|v_1\|$. Proceeding by induction we can suppose to have found for a given integer n elements e_i, $i = 1, \dots, m$, where $m \leq n$, which are orthonormal and such that the linear span

$$\operatorname{sp}(e_1, \dots, e_m) = \operatorname{sp}(v_1, \dots, v_n). \tag{3.22}$$

To show the inductive step observe that if v_{n+1} is in the span(s) in (3.22) then the same e_i's work for $n+1$. So we may as well assume that the next element, v_{n+1} is not in the span in (3.22). It follows that

$$w = v_{n+1} - \sum_{j=1}^{n} (v_{n+1}, e_j)e_j \neq 0 \text{ so } e_{m+1} = \frac{w}{\|w\|} \tag{3.23}$$

makes sense. By construction it is orthogonal to all the earlier e_i's so adding e_{m+1} gives the equality of the spans for $n+1$.

Thus we may continue indefinitely, since in fact the only way the dense set could be finite is if we were dealing with the space with one element, 0, in the first place. There are only two possibilities, either we get a finite set of e_i's or an infinite sequence. In either case this must be a maximal orthonormal sequence. That is, we claim

$$H \ni u \perp e_j \ \forall\ j \Longrightarrow u = 0. \tag{3.24}$$

This uses the density of the v_n's. There must exist a sequence w_j where each w_j is a v_n, such that $w_j \to u$ in H, assumed to satisfy (3.24). Now, each v_n, and hence each w_j, is a finite linear combination of e_k's so, by Bessel's inequality

$$\|w_j\|^2 = \sum_k |(w_j, e_k)|^2 = \sum_k |(u - w_j, e_k)|^2 \leq \|u - w_j\|^2 \tag{3.25}$$

where $(u, e_j) = 0$ for all j has been used. Thus $\|w_j\| \to 0$ and $u = 0$. □

Now, although a non-complete but separable pre-Hilbert space has maximal orthonormal sets, these are not much use without completeness.

5. Complete orthonormal bases

Definition 17. A maximal orthonormal sequence in a separable Hilbert space is called a complete orthonormal basis.

This notion of basis is not quite the same as in the finite dimensional case (although it is a legitimate extension of it).

Theorem 13. *If $\{e_i\}$ is a complete orthonormal basis in a Hilbert space then for any element $u \in H$ the 'Fourier-Bessel series' converges to u :*

$$u = \sum_{i=1}^{\infty} (u, e_i)e_i. \tag{3.26}$$

Proof. The sequence of partial sums of the Fourier-Bessel series

$$u_N = \sum_{i=1}^{N} (u, e_i)e_i \tag{3.27}$$

is Cauchy. Indeed, if $m < m'$ then

$$\|u_{m'} - u_m\|^2 = \sum_{i=m+1}^{m'} |(u, e_i)|^2 \le \sum_{i>m} |(u, e_i)|^2 \tag{3.28}$$

which is small for large m by Bessel's inequality. Since we are now assuming completeness, $u_m \to w$ in H. However, $(u_m, e_i) = (u, e_i)$ as soon as $m > i$ and $|(w - u_n, e_i)| \le \|w - u_n\|$ so in fact

$$(w, e_i) = \lim_{m\to\infty} (u_m, e_i) = (u, e_i) \tag{3.29}$$

for each i. Thus in fact $u - w$ is orthogonal to all the e_i so by the assumed completeness of the orthonormal basis must vanish. Thus indeed (3.26) holds. □

6. Isomorphism to l^2

A finite dimensional Hilbert space is isomorphic to $\mathbb{C}^n$ with its standard inner product. Similarly from the result above

Proposition 21. *Any infinite-dimensional separable Hilbert space (over the complex numbers) is isomorphic to* l^2, *that is there exists a linear map*

$$T : H \longrightarrow l^2 \tag{3.30}$$

which is 1-1, onto and satisfies $(Tu, Tv)_{l^2} = (u, v)_H$ *and* $\|Tu\|_{l^2} = \|u\|_H$ *for all* u, $v \in H$.

Proof. Choose an orthonormal basis – which exists by the discussion above and set

$$Tu = \{(u, e_j)\}_{j=1}^{\infty}. \tag{3.31}$$

This maps H into l^2 by Bessel's inequality. Moreover, it is linear since the entries in the sequence are linear in u. It is 1-1 since $Tu = 0$ implies $(u, e_j) = 0$ for all j implies $u = 0$ by the assumed completeness of the orthonormal basis. It is surjective since if $\{c_j\}_{j=1}^{\infty} \in l^2$ then

$$u = \sum_{j=1}^{\infty} c_j e_j \tag{3.32}$$

converges in H. This is the same argument as above – the sequence of partial sums is Cauchy since if $n > m$,

$$\| \sum_{j=m+1}^{n} c_j e_j \|_H^2 = \sum_{j=m+1}^{n} |c_|^2. \tag{3.33}$$

Again by continuity of the inner product, $Tu = \{c_j\}$ so T is surjective.

The equality of the norms follows from equality of the inner products and the latter follows by computation for finite linear combinations of the e_j and then in general by continuity. □

7. Parallelogram law

What exactly is the difference between a general Banach space and a Hilbert space? It is of course the existence of the inner product defining the norm. In fact it is possible to formulate this condition intrinsically in terms of the norm itself.

Proposition 22. *In any pre-Hilbert space the parallelogram law holds –*

$$\|v+w\|^2 + \|v-w\|^2 = 2\|v\|^2 + 2\|w\|^2,\ \forall\ v, w \in H. \tag{3.34}$$

Proof. Just expand out using the inner product

$$\|v+w\|^2 = \|v\|^2 + (v,w) + (w,v) + \|w\|^2 \tag{3.35}$$

and the same for $\|v-w\|^2$ and see the cancellation. □

Proposition 23. *Any normed space where the norm satisfies the parallelogram law,* (3.34)*, is a pre-Hilbert space in the sense that*

$$(v,w) = \frac{1}{4}\left(\|v+w\|^2 - \|v-w\|^2 + i\|v+iw\|^2 - i\|v-iw\|^2\right) \tag{3.36}$$

is a positive-definite Hermitian inner product which reproduces the norm.

Proof. A problem below. □

So, when we use the parallelogram law and completeness we are using the essence of the Hilbert space.

8. Convex sets and length minimizer

The following result does not need the hypothesis of separability of the Hilbert space and allows us to prove the subsequent results – especially Riesz' theorem – in full generality.

Proposition 24. *If $C \subset H$ is a subset of a Hilbert space which is*

1. *Non-empty*
2. *Closed*
3. *Convex, in the sense that $v_1, v_1 \in C$ implies $\frac{1}{2}(v_1+v_2) \in C$*

then there exists a unique element $v \in C$ closest to the origin, i.e. such that

$$\|v\|_H = \inf_{u\in C}\|u\|_H. \tag{3.37}$$

Proof. By definition of inf there must exist a sequence $\{v_n\}$ in C such that $\|v_n\| \to d = \inf_{u\in C}\|u\|_H$. We show that v_n converges and that the limit is the point we want. The parallelogram law can be written

$$\|v_n - v_m\|^2 = 2\|v_n\|^2 + 2\|v_m\|^2 - 4\|(v_n+v_m)/2\|^2. \tag{3.38}$$

Since $\|v_n\| \to d$, given $\epsilon > 0$ if N is large enough then $n > N$ implies $2\|v_n\|^2 < 2d^2 + \epsilon^2/2$. By convexity, $(v_n+v_m)/2 \in C$ so $\|(v_n+v_m)/2\|^2 \ge d^2$. Combining these estimates gives

$$n, m > N \Longrightarrow \|v_n - v_m\|^2 \le 4d^2 + \epsilon^2 - 4d^2 = \epsilon^2 \tag{3.39}$$

so $\{v_n\}$ is Cauchy. Since H is complete, $v_n \to v \in C$, since C is closed. Moreover, the distance is continuous so $\|v\|_H = \lim_{n\to\infty}\|v_n\| = d$.

Thus v exists and uniqueness follows again from the parallelogram law. If v and v' are two points in C with $\|v\| = \|v'\| = d$ then $(v + v')/2 \in C$ so

$$\|v - v'\|^2 = 2\|v\|^2 + 2\|v'\|^2 - 4\|(v + v')/2\|^2 \leq 0 \Longrightarrow v = v'. \tag{3.40}$$

□

9. Orthocomplements and projections

Proposition 25. *If $W \subset H$ is a linear subspace of a Hilbert space then*

$$W^\perp = \{u \in H; (u, w) = 0 \ \forall \ w \in W\} \tag{3.41}$$

is a closed linear subspace and $W \cap W^\perp = \{0\}$. If W is also closed then

$$H = W \oplus W^\perp \tag{3.42}$$

meaning that any $u \in H$ has a unique decomposition $u = w + w^\perp$ where $w \in W$ and $w^\perp \in W^\perp$.

Proof. That $W^\perp$ defined by (3.41) is a linear subspace follows from the linearity of the condition defining it. If $u \in W^\perp$ and $u \in W$ then $u \perp u$ by the definition so $(u, u) = \|u\|^2 = 0$ and $u = 0$. Since the map $H \ni u \longrightarrow (u, w) \in \mathbb{C}$ is continuous for each $w \in H$ its null space, the inverse image of 0, is closed. Thus

$$W^\perp = \bigcap_{w \in W} \{(u, w) = 0\} \tag{3.43}$$

is closed.

Now, suppose W is closed. If $W = H$ then $W^\perp = \{0\}$ and there is nothing to show. So consider $u \in H$, $u \notin W$ and set

$$C = u + W = \{u' \in H; u' = u + w, \ w \in W\}. \tag{3.44}$$

Then C is closed, since a sequence in it is of the form $u'_n = u + w_n$ where w_n is a sequence in W and u'_n converges if and only if w_n converges. Also, C is non-empty, since $u \in C$ and it is convex since $u' = u + w'$ and $u'' = u + w''$ in C implies $(u' + u'')/2 = u + (w' + w'')/2 \in C$.

Thus the length minimization result above applies and there exists a unique $v \in C$ such that $\|v\| = \inf_{u' \in C} \|u'\|$. The claim is that this v is perpendicular to W – draw a picture in two real dimensions! To see this consider an aritrary point $w \in W$ and $\lambda \in \mathbb{C}$ then $v + \lambda w \in C$ and

$$\|v + \lambda w\|^2 = \|v\|^2 + 2\operatorname{Re}(\lambda(v, w)) + |\lambda|^2\|w\|^2. \tag{3.45}$$

Choose $\lambda = te^{i\theta}$ where t is real and the phase is chosen so that $e^{i\theta}(v, w) = |(v, w)| \geq 0$. Then the fact that $\|v\|$ is minimal means that

$$\begin{gathered}\|v\|^2 + 2t|(v, w))| + t^2\|w\|^2 \geq \|v\|^2 \Longrightarrow \\ t(2|(v, w)| + t\|w\|^2) \geq 0 \ \forall \ t \in \mathbb{R} \Longrightarrow |(v, w)| = 0\end{gathered} \tag{3.46}$$

which is what we wanted to show.

Thus indeed, given $u \in H \setminus W$ we have constructed $v \in W^\perp$ such that $u = v + w$, $w \in W$. This is (3.42) with the uniqueness of the decomposition already shown since it reduces to 0 having only the decomposition $0 + 0$ and this in turn is $W \cap W^\perp = \{0\}$. □

Since the construction in the preceding proof associates a unique element in W, a closed linear subspace, to each $u \in H$, it defines a map

$$\Pi_W : H \longrightarrow W. \tag{3.47}$$

This map is linear, by the uniqueness since if $u_i = v_i + w_i$, $w_i \in W$, $(v_i, w_i) = 0$ are the decompositions of two elements then

$$\lambda_1 u_1 + \lambda_2 u_2 = (\lambda_1 v_1 + \lambda_2 v_2) + (\lambda_1 w_1 + \lambda_2 w_2) \tag{3.48}$$

must be the corresponding decomposition. Moreover $\Pi_W w = w$ for any $w \in W$ and $\|u\|^2 = \|v\|^2 + \|w\|^2$, Pythagoras' Theorem, shows that

$$\Pi_W^2 = \Pi_W,\ \|\Pi_W u\| \le \|u\| \Longrightarrow \|\Pi_W\| \le 1. \tag{3.49}$$

Thus, projection onto W is an operator of norm 1 (unless $W = \{0\}$) equal to its own square. Such an operator is called a projection or sometimes an idempotent (which sounds fancier).

Lemma 23. *If $\{e_j\}$ is any finite or countable orthonormal set in a Hilbert space then the orthogonal projection onto the closure of the span of these elements is*

$$Pu = \sum (u, e_k) e_k. \tag{3.50}$$

Proof. We know that the series in (3.50) converges and defines a bounded linear operator of norm at most one by Bessel's inequality. Clearly $P^2 = P$ by the same argument. If W is the closure of the span then $(u - Pu) \perp W$ since $(u - Pu) \perp e_k$ for each k and the inner product is continuous. Thus $u = (u - Pu) + Pu$ is the orthogonal decomposition with respect to W. □

10. Riesz' theorem

The most important application of these results is to prove Riesz' representation theorem (for Hilbert space, there is another one to do with measures).

Theorem 14. *If H is a Hilbert space then for any continuous linear functional $T : H \longrightarrow \mathbb{C}$ there exists a unique element $\phi \in H$ such that*

$$T(u) = (u, \phi)\ \forall\ u \in H. \tag{3.51}$$

Proof. If T is the zero functional then $\phi = 0$ gives (3.51). Otherwise there exists some $u' \in H$ such that $T(u') \ne 0$ and then there is some $u \in H$, namely $u = u'/T(u')$ will work, such that $T(u) = 1$. Thus

$$C = \{u \in H; T(u) = 1\} = T^{-1}(\{1\}) \ne \emptyset. \tag{3.52}$$

The continuity of T and the second form shows that C is closed, as the inverse image of a closed set under a continuous map. Moreover C is convex since

$$T((u + u')/2) = (T(u) + T(u'))/2. \tag{3.53}$$

Thus, by Proposition 24, there exists an element $v \in C$ of minimal length.

Notice that $C = \{v + w; w \in N\}$ where $N = T^{-1}(\{0\})$ is the null space of T. Thus, as in Proposition 25 above, v is orthogonal to N. In this case it is the unique element orthogonal to N with $T(v) = 1$.

Now, for any $u \in H$,

$$u - T(u)v \text{ satisfies } T(u - T(u)v) = T(u) - T(u)T(v) = 0 \Longrightarrow u = w + T(u)v,\ w \in N. \tag{3.54}$$

Then, $(u,v) = T(u)\|v\|^2$ since $(w,v) = 0$. Thus if $\phi = v/\|v\|^2$ then

$$u = w + (u,\phi)v \Longrightarrow T(u) = (u,\phi)T(v) = (u,\phi). \tag{3.55}$$

□

11. Adjoints of bounded operators

As an application of Riesz' we can see that to any bounded linear operator on a Hilbert space

$$A : H \longrightarrow H,\ \|Au\|_H \le C\|u\|_H\ \forall\ u \in H \tag{3.56}$$

there corresponds a unique adjoint operator.

Proposition 26. *For any bounded linear operator $A : H \longrightarrow H$ on a Hilbert space there is a unique bounded linear operator $A^* : H \longrightarrow H$ such that*

$$(Au,v)_H = (u,A^*v)_H\ \forall\ u,v \in H \text{ and } \|A\| = \|A^*\|. \tag{3.57}$$

Proof. To see the existence of A^*v we need to work out what $A^*v \in H$ should be for each fixed $v \in H$. So, fix v in the desired identity (3.57), which is to say consider

$$H \ni u \longrightarrow (Au,v) \in \mathbb{C}. \tag{3.58}$$

This is a linear map and it is clearly bounded, since

$$|(Au,v)| \le \|Au\|_H\|v\|_H \le (\|A\|\|v\|_H)\|u\|_H. \tag{3.59}$$

Thus it is a continuous linear functional on H which depends on v. In fact it is just the composite of two continuous linear maps

$$H \overset{u\longmapsto Au}{\longrightarrow} H \overset{w\longmapsto (w,v)}{\longrightarrow} \mathbb{C}. \tag{3.60}$$

By Riesz' theorem there is a unique element in H, which we can denote A^*v (since it only depends on v) such that

$$(Au,v) = (u,A^*v)\ \forall\ u \in H. \tag{3.61}$$

Now this defines the map $A^* : H \longrightarrow H$ but we need to check that it is linear and continuous. Linearity follows from the uniqueness part of Riesz' theorem. Thus if $v_1,\ v_2 \in H$ and $c_1,\ c_2 \in \mathbb{C}$ then

$$\begin{aligned}(Au, c_1v_1 + c_2v_2) &= \overline{c_1}(Au,v_1) + \overline{c_2}(Au,v_2)\\ &= \overline{c_1}(u,A^*v_1) + \overline{c_2}(u,A^*v_2) = (u, c_1A^*v_2 + c_2A^*v_2)\end{aligned} \tag{3.62}$$

where we have used the definitions of A^*v_1 and A^*v_2 – by uniqueness we must have $A^*(c_1v_1 + c_2v_2) = c_1A^*v_1 + c_2A^*v_2$.

Since we know the optimality of Cauchy's inequality

$$\|v\|_H = \sup_{\|u\|=1} |(u,v)| \tag{3.63}$$

it follows that

$$\|A^*v\| = \sup_{\|u\|=1} |(u,A^*v)| = \sup_{\|u\|=1} |(Au,v)| \le \|A\|\|v\|. \tag{3.64}$$

So in fact

$$\|A^*\| \le \|A\| \tag{3.65}$$

which shows that A^* is bounded.

The defining identity (3.57) also shows that $(A^*)^* = A$ so the reverse equality in (3.65) also holds and so

$$\|A^*\| = \|A\|. \tag{3.66}$$

□

12. Compactness and equi-small tails

A compact subset in a general metric space is one with the property that any sequence in it has a convergent subsequence, with its limit in the set. You will recall, with pleasure no doubt, the equivalence of this condition to the (more general since it makes good sense in an arbitrary topological space) covering condition, that *any* open cover of the set has a finite subcover. So, in a separable Hilbert space the notion of a compact set is already fixed. We want to characterize it, actually in several ways.

A general result in a metric space is that any compact set is both closed and bounded, so this must be true in a Hilbert space. The Heine-Borel theorem gives a converse to this, for $\mathbb{R}^n$ or $\mathbb{C}^n$ (and hence in any finite dimensional normed space) in which any closed and bounded set is compact. Also recall that the convergence of a sequence in $\mathbb{C}^n$ is equivalent to the convergence of the n sequences given by its components and this is what is used to pass first from $\mathbb{R}$ to $\mathbb{C}$ and then to $\mathbb{C}^n$. All of this fails in infinite dimensions and we need some condition in addition to being bounded and closed for a set to be compact.

To see where this might come from, observe that

LEMMA 24. *In any metric space a set, S, consisting of the points of a convergent sequence, $s : \mathbb{N} \longrightarrow M$, together with its limit, s, is compact.*

PROOF. The set here is the image of the sequence, thought of as a map from the integers into the metric space, together with the limit (which might or might not already be in the image of the sequence). Certainly this set is bounded, since the distance from the intial point is bounded. Moreover it is closed. Indeed, the complement $M \setminus S$ is open – if $p \in M \setminus S$ then it is not the limit of the sequence, so for some $\epsilon > 0$, and some N, if $n > N$ then $s(n) \notin B(p, \epsilon)$. Shrinking ϵ further if necessary, we can make sure that all the $s(k)$ for $k \le N$ are not in the ball either – since they are each at a positive distance from p. Thus $B(p, \epsilon) \subset M \setminus S$.

Finally, S is compact since any sequence in S has a convergent subsequence. To see this, observe that a sequence $\{t_j\}$ in S either has a subsequence converging to the limit s of the original sequence or it does not. So we only need consider the latter case, but this means that, for some $\epsilon > 0$, $d(t_j, s) > \epsilon$; but then t_j takes values in a finite set, since $S \setminus B(s, \epsilon)$ is finite – hence some value is repeated infinitely often and there is a convergent subsequence. □

LEMMA 25. *The image of a convergent sequence in a Hilbert space is a set with equi-small tails with respect to any orthonormal sequence, i.e. if e_k is an othonormal sequence and $u_n \to u$ is a convergent sequence then given $\epsilon > 0$ there exists N such that*

$$\sum_{k>N} |(u_n, e_k)|^2 < \epsilon^2 \ \forall \ n. \tag{3.67}$$

PROOF. Bessel's inequality shows that for any $u \in \mathcal{H}$,

$$\sum_k |(u, e_k)|^2 \le \|u\|^2. \tag{3.68}$$

The convergence of this series means that (3.67) can be arranged for any single element u_n or the limit u by choosing N large enough, thus given $\epsilon > 0$ we can choose N' so that

$$\sum_{k>N'} |(u, e_k)|^2 < \epsilon^2/2. \tag{3.69}$$

Consider the closure of the subspace spanned by the e_k with $k > N$. The orthogonal projection onto this space (see Lemma 23) is

$$P_N u = \sum_{k>N} (u, e_k) e_k. \tag{3.70}$$

Then the convergence $u_n \to u$ implies the convergence in norm $\|P_N u_n\| \to \|P_N u\|$, so

$$\|P_N u_n\|^2 = \sum_{k>N} |(u_n, e_k)|^2 < \epsilon^2, \ n > n'. \tag{3.71}$$

So, we have arranged (3.67) for $n > n'$ for some N. This estimate remains valid if N is increased – since the tails get smaller – and we may arrange it for $n \le n'$ by chossing N large enough. Thus indeed (3.67) holds for all n if N is chosen large enough. □

This suggests one useful characterization of compact sets in a separable Hilbert space.

PROPOSITION 27. *A set $K \subset \mathcal{H}$ in a separable Hilbert space is compact if and only if it is bounded, closed and the Fourier-Bessel sequence with respect to any (one) complete orthonormal basis converges* uniformly *on it.*

PROOF. We already know that a compact set in a metric space is closed and bounded. Suppose the equi-smallness of tails condition fails with respect to some orthonormal basis e_k. This means that for some $\epsilon > 0$ and all p there is an element $u_p \in K$, such that

$$\sum_{k>p} |(u_p, e_k)|^2 \ge \epsilon^2. \tag{3.72}$$

Consider the subsequence $\{u_p\}$ generated this way. No subsequence of it can have equi-small tails (recalling that the tail decreases with p). Thus, by Lemma 25, it cannot have a convergent subsequence, so K cannot be compact if the equi-smallness condition fails.

Thus we have proved the equi-smallness of tails condition to be necessary for the compactness of a closed, bounded set. It remains to show that it is sufficient.

So, suppose K is closed, bounded and satisfies the equi-small tails condition with respect to an orthonormal basis e_k and $\{u_n\}$ is a sequence in K. We only need show that $\{u_n\}$ has a Cauchy subsequence, since this will converge ($\mathcal{H}$ being complete) and the limit will be in K (since it is closed). Consider each of the sequences of coefficients (u_n, e_k) in $\mathbb{C}$. Here k is fixed. This sequence is bounded:

$$|(u_n, e_k)| \le \|u_n\| \le C \tag{3.73}$$

by the boundedness of K. So, by the Heine-Borel theorem, there is a subsequence u_{n_l} such that (u_{n_l}, e_k) converges as $l \to \infty$.

We can apply this argument for each $k = 1, 2, \dots$. First extract a subsequence $\{u_{n,1}\}$ of $\{u_n\}$ so that the sequence $(u_{n,1}, e_1)$ converges. Then extract a subsequence $u_{n,2}$ of $u_{n,1}$ so that $(u_{n,2}, e_2)$ *also* converges. Then continue inductively. Now pass to the 'diagonal' subsequence v_n of $\{u_n\}$ which has kth entry the kth term, $u_{k,k}$ in the kth subsequence. It is 'eventually' a subsequence of each of the subsequences previously constructed – meaning it coincides with a subsequence from some point onward (namely the kth term onward for the kth subsquence). Thus, for this subsequence *each* of the (v_n, e_k) converges.

Consider the identity (the orthonormal set e_k is complete by assumption) for the difference

$$\begin{aligned}\|v_n - v_{n+l}\|^2 &= \sum_{k\le N} |(v_n - v_{n+l}, e_k)|^2 + \sum_{k>N} |(v_n - v_{n+l}, e_k)|^2 \\ &\le \sum_{k\le N} |(v_n - v_{n+l}, e_k)|^2 + 2\sum_{k>N} |(v_n, e_k)|^2 + 2\sum_{k>N} |(v_{n+l}, e_k)|^2\end{aligned} \tag{3.74}$$

where the parallelogram law on $\mathbb{C}$ has been used. To make this sum less than ϵ^2 we may choose N so large that the last two terms are less than $\epsilon^2/2$ and this may be done for all n and l by the equi-smallness of the tails. Now, choose n so large that each of the terms in the first sum is less than $\epsilon^2/2N$, for all $l > 0$ using the Cauchy condition on each of the finite number of sequence (v_n, e_k). Thus, $\{v_n\}$ is a Cauchy subsequence of $\{u_n\}$ and hence as already noted convergent in K. Thus K is indeed compact. □

13. Finite rank operators

Now, we need to starting thinking a little more seriously about operators on a Hilbert space, remember that an operator is just a continuous linear map $T : \mathcal{H} \longrightarrow \mathcal{H}$ and the space of them (a Banach space) is denoted $\mathcal{B}(\mathcal{H})$ (rather than the more cumbersome $\mathcal{B}(\mathcal{H}, \mathcal{H})$ which is needed when the domain and target spaces are different).

Definition 18. An operator $T \in \mathcal{B}(\mathcal{H})$ is of *finite rank* if its range has finite dimension (and that dimension is called the rank of T); the set of finite rank operators will be denoted $\mathcal{R}(\mathcal{H})$.

Why not $\mathcal{F}(\mathcal{H})$? Because we want to use this for the *Fredholm operators*.

Clearly the sum of two operators of finite rank has finite rank, since the range is contained in the sum of the ranges (but is often smaller):

$$(T_1 + T_2)u \in \operatorname{Ran}(T_1) + \operatorname{Ran}(T_2) \ \forall \ u \in \mathcal{H}. \tag{3.75}$$

Since the range of a constant multiple of T is contained in the range of T it follows that the finite rank operators form a linear subspace of $\mathcal{B}(\mathcal{H})$.

What does a finite rank operator look like? It really looks like a matrix.

Lemma 26. *If $T : H \longrightarrow H$ has finite rank then there is a finite orthonormal set $\{e_k\}_{k=1}^{L}$ in H such that*

$$Tu = \sum_{i,j=1}^{L} c_{ij}(u, e_j)e_i. \tag{3.76}$$

Proof. By definition, the range of T, $R = T(H)$ is a finite dimensional subspace. So, it has a basis which we can diagonalize in H to get an orthonormal basis, e_i, $i = 1, \dots, p$. Now, since this is a basis of the range, Tu can be expanded relative to it for any $u \in H$:

$$Tu = \sum_{i=1}^{p} (Tu, e_i)e_i. \tag{3.77}$$

On the other hand, the map $u \longrightarrow (Tu, e_i)$ is a continuous linear functional on H, so $(Tu, e_i) = (u, v_i)$ for some $v_i \in H$; notice in fact that $v_i = T^* e_i$. This means the formula (3.77) becomes

$$Tu = \sum_{i=1}^{p} (u, v_i)e_i. \tag{3.78}$$

Now, the Gram-Schmidt procedure can be applied to orthonormalize the sequence $e_1, \dots, e_p, v_1 \dots, v_p$ resulting in $e_1, \dots, e_L$. This means that each v_i is a linear combination which we can write as

$$v_i = \sum_{j=1}^{L} \overline{c_{ij}} e_j. \tag{3.79}$$

Inserting this into (3.78) gives (3.76) (where the constants for $i > p$ are zero). $\square$

It is clear that

$$B \in \mathcal{B}(\mathcal{H}) \text{ and } T \in \mathcal{R}(\mathcal{H}) \text{ then } BT \in \mathcal{R}(\mathcal{H}). \tag{3.80}$$

Indeed, the range of BT is the range of B restricted to the range of T and this is certainly finite dimensional since it is spanned by the image of a basis of $\operatorname{Ran}(T)$. Similalry $TB \in \mathcal{R}(\mathcal{H})$ since the range of TB is contained in the range of T. Thus we have in fact proved most of

Proposition 28. *The finite rank operators form a $*$-closed two-sided ideal in $\mathcal{B}(\mathcal{H})$, which is to say a linear subspace such that*

$$B_1,\ B_2 \in \mathcal{B}(\mathcal{H}),\ T \in \mathcal{R}(\mathcal{H}) \Longrightarrow B_1 T B_2,\ T^* \in \mathcal{R}(\mathcal{H}). \tag{3.81}$$

Proof. It is only left to show that T^* is of finite rank if T is, but this is an immediate consequence of Lemma 26 since if T is given by (3.76) then

$$T^* u = \sum_{i,j=1}^{N} \overline{c_{ij}} (u, e_i) e_j \tag{3.82}$$

is also of finite rank. $\square$

Lemma 27 (Row rank=Colum rank). *For any finite rank operator on a Hilbert space, the dimension of the range of T is equal to the dimension of the range of T^*.*

Proof. From the formula (3.78) for a finite rank operator, it follows that the v_i, $i = 1, \dots, p$ must be linearly independent – since the e_i form a basis for the range and a linear relation between the v_i would show the range had dimension less than p. Thus in fact the null space of T is precisely the orthocomplement of the span of the v_i – the space of vectors orthogonal to each v_i. Since

$$\begin{gathered}(Tu, w) = \sum_{i=1}^{p} (u, v_i)(e_i, w) \Longrightarrow \\ (w, Tu) = \sum_{i=1}^{p} (v_i, u)(w, e_i) \Longrightarrow \\ T^*w = \sum_{i=1}^{p} (w, e_i) v_i\end{gathered} \tag{3.83}$$

the range of T^* is the span of the v_i, so is also of dimension p. □

14. Compact operators

Definition 19. An element $K \in \mathcal{B}(\mathcal{H})$, the bounded operators on a separable Hilbert space, is said to be *compact* (the old terminology was 'totally bounded' or 'completely continuous') if the image of the unit ball is precompact, i.e. has compact closure – that is if the closure of $K\{u \in \mathcal{H}; \|u\|_{\mathcal{H}} \le 1\}$ is compact in $\mathcal{H}$.

Notice that in a metric space, to say that a set has compact closure is the same as saying it is contained in a compact set.

Proposition 29. *An operator $K \in \mathcal{B}(\mathcal{H})$, bounded on a separable Hilbert space, is compact if and only if it is the limit of a norm-convergent sequence of finite rank operators.*

Proof. So, we need to show that a compact operator is the limit of a convergent sequence of finite rank operators. To do this we use the characterizations of compact subsets of a separable Hilbert space discussed earlier. Namely, if $\{e_i\}$ is an orthonormal basis of $\mathcal{H}$ then a subset $I \subset \mathcal{H}$ is compact if and only if it is closed and bounded and has equi-small tails with respect to $\{e_i\}$, meaning given $\epsilon > 0$ there exits N such that

$$\sum_{i>N} |(v, e_i)|^2 < \epsilon^2 \ \forall \ v \in I. \tag{3.84}$$

Now we shall apply this to the set $K(B(0, 1))$ where we assume that K is compact (as an operator, don't be confused by the double usage, in the end it turns out to be constructive) – so this set is *contained* in a compact set. Hence (3.84) applies to it. Namely this means that for any $\epsilon > 0$ there exists n such that

$$\sum_{i>n} |(Ku, e_i)|^2 < \epsilon^2 \ \forall \ u \in \mathcal{H}, \ \|u\|_{\mathcal{H}} \le 1. \tag{3.85}$$

For each n consider the first part of these sequences and define

$$K_n u = \sum_{k \le n} (Ku, e_i) e_i. \tag{3.86}$$

This is clearly a linear operator and has finite rank – since its range is contained in the span of the first n elements of $\{e_i\}$. Since this is an orthonormal basis,

$$\|Ku - K_n u\|_{\mathcal{H}}^2 = \sum_{i>n} |(Ku, e_i)|^2 \tag{3.87}$$

Thus (3.85) shows that $\|Ku - K_nu\|_{\mathcal{H}} \le \epsilon$. Now, increasing n makes $\|Ku - K_nu\|$ smaller, so given $\epsilon > 0$ there exists n such that for all $N \ge n$,

$$\|K - K_N\|_{\mathcal{B}} = \sup_{\|u\|\le 1} \|Ku - K_nu\|_{\mathcal{H}} \le \epsilon. \tag{3.88}$$

Thus indeed, $K_n \to K$ in norm and we have shown that the compact operators are contained in the norm closure of the finite rank operators.

For the converse we assume that $T_n \to K$ is a norm convergent sequence in $\mathcal{B}(\mathcal{H})$ where each of the T_n is of finite rank – of course we know nothing about the rank except that it is finite. We want to conclude that K is compact, so we need to show that $K(B(0,1))$ is precompact. It is certainly bounded, by the norm of K. By a result above on compactness of sets in a separable Hilbert space we know that it suffices to prove that the closure of the image of the unit ball has uniformly small tails. Let Π_N be the orthogonal projection *off* the first N elements of a complete orthonormal basis $\{e_k\}$ – so

$$u = \sum_{k\le N} (u, e_k)e_k + \Pi_N u. \tag{3.89}$$

Then we know that $\|\Pi_N\| = 1$ (assuming the Hilbert space is infinite dimensional) and $\|\Pi_N u\|$ is the 'tail'. So what we need to show is that given $\epsilon > 0$ there exists n such that

$$\|u\| \le 1 \Longrightarrow \|\Pi_N Ku\| < \epsilon. \tag{3.90}$$

Now,

$$\|\Pi_N Ku\| \le \|\Pi_N(K - T_n)u\| + \|\Pi_N T_n u\| \tag{3.91}$$

so choosing n large enough that $\|K - T_n\| < \epsilon/2$ and then using the compactness of T_n (which is finite rank) to choose N so large that

$$\|u\| \le 1 \Longrightarrow \|\Pi_N T_n u\| \le \epsilon/2 \tag{3.92}$$

shows that (3.90) holds and hence K is compact. □

Proposition 30. *For any separable Hilbert space, the compact operators form a closed and $*$-closed two-sided ideal in $\mathcal{B}(H)$.*

Proof. In any metric space (applied to $\mathcal{B}(H)$) the closure of a set is closed, so the compact operators are closed being the closure of the finite rank operators. Similarly the fact that it is closed under passage to adjoints follows from the same fact for finite rank operators. The ideal properties also follow from the corresponding properties for the finite rank operators, or we can prove them directly anyway. Namely if B is bounded and T is compact then for some $c > 0$ (namely $1/\|B\|$ unless it is zero) cB maps $B(0,1)$ into itself. Thus $cTB = TcB$ is compact since the image of the unit ball under it is contained in the image of the unit ball under T; hence TB is also compact. Similarly BT is compact since B is continuous and then

$$BT(B(0,1)) \subset B(\overline{T(B(0,1))}) \text{ is compact} \tag{3.93}$$

since it is the image under a continuous map of a compact set. □

15. Weak convergence

It is convenient to formalize the idea that a sequence be bounded and that each of the (u_n, e_k), the sequence of coefficients of some particular Fourier-Bessel series, should converge.

Definition 20. A sequence, $\{u_n\}$, in a Hilbert space, $\mathcal{H}$, is said to *converge weakly* to an element $u \in \mathcal{H}$ if it is bounded in norm and $(u_j, v) \to (u, v)$ converges in $\mathbb{C}$ for each $v \in \mathcal{H}$. This relationship is written

$$u_n \rightharpoonup u. \tag{3.94}$$

In fact as we shall see below, the assumption that $\|u_n\|$ is bounded and that u exists are both unnecessary. That is, a sequence converges weakly if and only if (u_n, v) converges in $\mathbb{C}$ for each $v \in \mathcal{H}$. Conversely, there is no harm in assuming it is bounded and that the 'weak limit' $u \in \mathcal{H}$ exists. Note that the weak limit is unique since if u and u' both have this property then $(u - u', v) = \lim_{n\to\infty}(u_n, v) - \lim_{n\to\infty}(u_n, v) = 0$ for all $v \in \mathcal{H}$ and setting $v = u - u'$ it follows that $u = u'$.

Lemma 28. *A (strongly) convergent sequence is weakly convergent with the same limit.*

Proof. This is the continuity of the inner product. If $u_n \to u$ then

$$|(u_n, v) - (u, v)| \le \|u_n - u\|\|v\| \to 0 \tag{3.95}$$

for each $v \in \mathcal{H}$ shows weak convergence. □

Lemma 29. *For a bounded sequence in a separable Hilbert space, weak convergence is equivalent to component convergence with respect to an orthonormal basis.*

Proof. Let e_k be an orthonormal basis. Then if u_n is weakly convergent it follows immediately that $(u_n, e_k) \to (u, e_k)$ converges for each k. Conversely, suppose this is true for a bounded sequence, just that $(u_n, e_k) \to c_k$ in $\mathbb{C}$ for each k. The norm boundedness and Bessel's inequality show that

$$\sum_{k \le p} |c_k|^2 = \lim_{n\to\infty} \sum_{k \le p} |(u_n, e_k)|^2 \le C^2 \sup_n \|u_n\|^2 \tag{3.96}$$

for all p. Thus in fact $\{c_k\} \in l^2$ and hence

$$u = \sum_k c_k e_k \in \mathcal{H} \tag{3.97}$$

by the completeness of $\mathcal{H}$. Clearly $(u_n, e_k) \to (u, e_k)$ for each k. It remains to show that $(u_n, v) \to (u, v)$ for all $v \in \mathcal{H}$. This is certainly true for any finite linear combination of the e_k and for a general v we can write

$$\begin{aligned}(u_n, v) - (u, v) = (u_n, v_p) - (u, v_p) + (u_n, v - v_p) - (u, v - v_p) \Longrightarrow \\ |(u_n, v) - (u, v)| \le |(u_n, v_p) - (u, v_p)| + 2C\|v - v_p\|\end{aligned} \tag{3.98}$$

where $v_p = \sum\limits_{k \le p} (v, e_k) e_k$ is a finite part of the Fourier-Bessel series for v and C is a bound for $\|u_n\|$. Now the convergence $v_p \to v$ implies that the last term in (3.98) can be made small by choosing p large, independent of n. Then the second last term can be made small by choosing n large since v_p is a finite linear combination of the

e_k. Thus indeed, $(u_n, v) \to (u, v)$ for all $v \in \mathcal{H}$ and it follows that u_n converges weakly to u. □

PROPOSITION 31. *Any bounded sequence $\{u_n\}$ in a separable Hilbert space has a weakly convergent subsequence.*

This can be thought of as an analogue in infinite dimensions of the Heine-Borel theorem if you say 'a bounded closed subset of a separable Hilbert space is *weakly* compact'.

PROOF. Choose an orthonormal basis $\{e_k\}$ and apply the procedure in the proof of Proposition 27 to extract a subsequence of the given bounded sequence such that (u_{n_p}, e_k) converges for each k. Now apply the preceeding Lemma to conclude that this subsequence converges weakly. □

LEMMA 30. *For a weakly convergent sequence $u_n \rightharpoonup u$*

$$\|u\| \le \liminf \|u_n\|. \tag{3.99}$$

PROOF. Choose an orthonormal basis e_k and observe that

$$\sum_{k \le p} |(u, e_k)|^2 = \lim_{n \to \infty} \sum_{k \le p} |(u_n, e_k)|^2. \tag{3.100}$$

The sum on the right is bounded by $\|u_n\|^2$ independently of p so

$$\sum_{k \le p} \|u, e_k\|^2 \le \liminf_n \|u_n\|^2 \tag{3.101}$$

by the definition of $\liminf$. Then let $p \to \infty$ to conclude that

$$\|u\|^2 \le \liminf_n \|u_n\|^2 \tag{3.102}$$

from which (3.99) follows. □

LEMMA 31. *An operator $K \in \mathcal{B}(\mathcal{H})$ is compact if and only if the image Ku_n of any weakly convergent sequence $\{u_n\}$ in $\mathcal{H}$ is strongly, i.e. norm, convergent.*

This is the origin of the old name 'completely continuous' for compact operators, since they turn even weakly convergent into strongly convergent sequences.

PROOF. First suppose that $u_n \rightharpoonup u$ is a weakly convergent sequence in $\mathcal{H}$ and that K is compact. We know that $\|u_n\| < C$ is bounded so the sequence Ku_n is contained in $CK(B(0, 1))$ and hence in a compact set (clearly if D is compact then so is cD for any constant c.) Thus, any subsequence of Ku_n has a convergent subseqeunce and the limit is necessarily Ku since $Ku_n \rightharpoonup Ku$ (true for any bounded operator by computing

$$(Ku_n, v) = (u_n, K^*v) \to (u, K^*v) = (Ku, v).) \tag{3.103}$$

But the condition on a sequence in a metric space that every subsequence of it has a subsequence which converges to a fixed limit implies convergence. (If you don't remember this, reconstruct the proof: To say a sequence v_n *does not* converge to v is to say that for some $\epsilon > 0$ there is a subsequence along which $d(v_{n_k}, v) \ge \epsilon$. This is impossible given the subsequence of subsequence condition (converging to the fixed limit v.))

Conversely, suppose that K has this property of turning weakly convergent into strongly convergent sequences. We want to show that $K(B(0, 1))$ has compact closure. This just means that any sequence in $K(B(0,1))$ has a (strongly) convergent subsequence – where we do not have to worry about whether the limit is in the set or not. Such a sequence is of the form Ku_n where u_n is a sequence in $B(0,1)$. However we know that the ball is weakly compact, that is we can pass to a subsequence which converges weakly, $u_{n_j} \rightharpoonup u$. Then, by the assumption of the Lemma, $Ku_{n_j} \to Ku$ converges strongly. Thus u_n does indeed have a convergent subsequence and hence $K(B(0,1))$ must have compact closure. □

As noted above, it is not really necessary to assume that a sequence in a Hilbert space is bounded, provided one has the Uniform Boundedness Principle, Theorem 3, at the ready.

Proposition 32. *If $u_n \in H$ is a sequence in a Hilbert space and for all $v \in H$*

$$(u_n, v) \to F(v) \text{ converges in } \mathbb{C} \tag{3.104}$$

then $\|u_n\|_H$ is bounded and there exists $w \in H$ such that $u_n \rightharpoonup w$ (converges weakly).

Proof. Apply the Uniform Boundedness Theorem to the continuous functionals

$$T_n(u) = (u, u_n),\ T_n : H \longrightarrow \mathbb{C} \tag{3.105}$$

where we reverse the order to make them linear rather than anti-linear. Thus, each set $|T_n(u)|$ is bounded in $\mathbb{C}$ since it is convergent. It follows from the Uniform Boundedness Principle that there is a bound

$$\|T_n\| \le C. \tag{3.106}$$

However, this norm as a functional is just $\|T_n\| = \|u_n\|_H$ so the original sequence must be bounded in H. Define $T : H \longrightarrow \mathbb{C}$ as the limit for each u :

$$T(u) = \lim_{n\to\infty} T_n(u) = \lim_{n\to\infty} (u, u_n). \tag{3.107}$$

This exists for each u by hypothesis. It is a linear map and from (3.106) it is bounded, $\|T\| \le C$. Thus by the Riesz Representation theorem, there exists $w \in H$ such that

$$T(u) = (u, w)\ \forall\ u \in H. \tag{3.108}$$

Thus $(u_n, u) \to (w, u)$ for all $u \in H$ so $u_n \rightharpoonup w$ as claimed. □

16. The algebra $\mathcal{B}(H)$

Recall the basic properties of the Banach space, and algebra, of bounded operators $\mathcal{B}(\mathcal{H})$ on a separable Hilbert space $\mathcal{H}$. In particular that it is a Banach space with respect to the norm

$$\|A\| = \sup_{\|u\|_{\mathcal{H}}=1} \|Au\|_{\mathcal{H}} \tag{3.109}$$

and that the norm satisfies

$$\|AB\| \le \|A\|\|B\| \tag{3.110}$$

as follows from the fact that

$$\|ABu\| \le \|A\|\|Bu\| \le \|A\|\|B\|\|u\|.$$

Consider the set of invertible elements:

$$\mathrm{GL}(\mathcal{H}) = \{A \in \mathcal{B}(\mathcal{H}); \exists\ B \in \mathcal{B}(\mathcal{H}),\ BA = AB = \mathrm{Id}\}. \tag{3.111}$$

Note that this is equivalent to saying A is 1-1 and onto in view of the Open Mapping Theorem, Theorem 4.

This set is open, to see this consider a neighbourhood of the identity.

Lemma 32. *If* $A \in \mathcal{B}(\mathcal{H})$ *and* $\|A\| < 1$ *then*

$$\mathrm{Id} - A \in \mathrm{GL}(\mathcal{H}). \tag{3.112}$$

Proof. This follows from the convergence of the Neumann series. If $\|A\| < 1$ then $\|A^j\| \le \|A\|^j$, from (3.110), and it follows that

$$B = \sum_{j=0}^{\infty} A^j \tag{3.113}$$

(where $A^0 = \mathrm{Id}$ by definition) is absolutely summable in $\mathcal{B}(\mathcal{H})$ since $\sum\limits_{j=0}^{\infty} \|A^j\|$ converges. Since $\mathcal{B}(H)$ is a Banach space, the sum converges. Moreover by the continuity of the product with respect to the norm

$$AB = A \lim_{n\to\infty} \sum_{j=0}^{n} A^j = \lim_{n\to\infty} \sum_{j=1}^{n+1} A^j = B - \mathrm{Id} \tag{3.114}$$

and similarly $BA = B - \mathrm{Id}$. Thus $(\mathrm{Id} - A)B = B(\mathrm{Id} - A) = \mathrm{Id}$ shows that B is a (and hence *the*) 2-sided inverse of $\mathrm{Id} - A$. □

Proposition 33. *The invertible elements form an open subset* $\mathrm{GL}(\mathcal{H}) \subset \mathcal{B}(\mathcal{H})$.

Proof. Suppose $G \in \mathrm{GL}(\mathcal{H})$, meaning it has a two-sided (and unique) inverse $G^{-1} \in \mathcal{B}(\mathcal{H})$:

$$G^{-1}G = GG^{-1} = \mathrm{Id}\,. \tag{3.115}$$

Then we wish to show that $B(G; \epsilon) \subset \mathrm{GL}(\mathcal{H})$ for some $\epsilon > 0$. In fact we shall see that we can take $\epsilon = \|G^{-1}\|^{-1}$. To show that $G + B$ is invertible set

$$E = -G^{-1}B \Longrightarrow G + B = G(\mathrm{Id} + G^{-1}B) = G(\mathrm{Id} - E) \tag{3.116}$$

From Lemma 32 we know that

$$\|B\| < 1/\|G^{-1}\| \Longrightarrow \|G^{-1}B\| < 1 \Longrightarrow \mathrm{Id} - E \text{ is invertible.} \tag{3.117}$$

Then $(\mathrm{Id} - E)^{-1}G^{-1}$ satisfies

$$(\mathrm{Id} - E)^{-1}G^{-1}(G + B) = (\mathrm{Id} - E)^{-1}(\mathrm{Id} - E) = \mathrm{Id}\,. \tag{3.118}$$

Moreover $E' = -BG^{-1}$ also satisfies $\|E'\| \le \|B\|\|G^{-1}\| < 1$ and

$$(G + B)G^{-1}(\mathrm{Id} - E')^{-1} = (\mathrm{Id} - E')(\mathrm{Id} - E')^{-1} = \mathrm{Id}\,. \tag{3.119}$$

Thus $G + B$ has both a 'left' and a 'right' inverse. The associtivity of the operator product (that $A(BC) = (AB)C$) then shows that

$$G^{-1}(\mathrm{Id} - E')^{-1} = (\mathrm{Id} - E)^{-1}G^{-1}(G+B)G^{-1}(\mathrm{Id} - E')^{-1} = (\mathrm{Id} - E)^{-1}G^{-1} \tag{3.120}$$

so the left and right inverses are equal and hence $G + B$ is invertible. □

Thus $\mathrm{GL}(\mathcal{H}) \subset \mathcal{B}(\mathcal{H})$, the set of invertible elements, is open. It is also a group – since the inverse of G_1G_2 if G_1, $G_2 \in \mathrm{GL}(\mathcal{H})$ is $G_2^{-1}G_1^{-1}$.

This group of invertible elements has a smaller subgroup, $\mathrm{U}(\mathcal{H})$, the unitary group, defined by

$$\mathrm{U}(\mathcal{H}) = \{U \in \mathrm{GL}(\mathcal{H}); U^{-1} = U^*\}. \tag{3.121}$$

The unitary group consists of the linear isometric isomorphisms of $\mathcal{H}$ onto itself – thus

$$(Uu, Uv) = (u, v),\ \|Uu\| = \|u\|\ \forall\ u, v \in \mathcal{H},\ U \in \mathrm{U}(\mathcal{H}). \tag{3.122}$$

This is an important object and we will use it a little bit later on.

The groups $\mathrm{GL}(H)$ and $\mathrm{U}(H)$ for a separable Hilbert space may seem very similar to the familiar groups of invertible and unitary $n \times n$ matrices, $\mathrm{GL}(n)$ and $\mathrm{U}(n)$, but this is somewhat deceptive. For one thing they are much bigger. In fact there are other important qualitative differences – you can find some of this in the problems. One important fact that you should know, even though we will not try prove it here, is that both $\mathrm{GL}(H)$ and $\mathrm{U}(\mathcal{H})$ are contractible as a metric spaces – they have no significant topology. This is to be constrasted with the $\mathrm{GL}(n)$ and $\mathrm{U}(n)$ which have a lot of topology, and are not at all simple spaces – especially for large n. One upshot of this is that $\mathrm{U}(\mathcal{H})$ does not look much like the limit of the $\mathrm{U}(n)$ as $n \to \infty$. Another important fact that we will show is that $\mathrm{GL}(H)$ is *not* dense in $\mathcal{B}(H)$, in contrast to the finite dimensional case.

17. Spectrum of an operator

Another direct application of Lemma 32, the convergence of the Neumann series, is that if $A \in \mathcal{B}(H)$ and $\lambda \in \mathbb{C}$ has $|\lambda| > \|A\|$ then $\|\lambda^{-1}A\| < 1$ so $(\mathrm{Id} - \lambda^{-1}A)^{-1}$ exists and satisfies

$$(\lambda \,\mathrm{Id} - A)\lambda^{-1}(\mathrm{Id} - \lambda^{-1}A)^{-1} = \mathrm{Id} = \lambda^{-1}(\mathrm{Id} - \lambda^{-1}A)^{-1}(\lambda - A). \tag{3.123}$$

Thus, $\lambda - A \in \mathrm{GL}(H)$ has inverse $(\lambda - A)^{-1} = \lambda^{-1}(\mathrm{Id} - \lambda^{-1}A)^{-1}$. The set of λ for which this operator is invertible,

$$\{\lambda \in \mathbb{C}; (\lambda \,\mathrm{Id} - A) \in \mathrm{GL}(H)\} \subset \mathbb{C} \tag{3.124}$$

is an open, and non-empty, set called the *resolvent set* (usually $(A - \lambda)^{-1}$ is called the resolvent). The complement of the resolvent set is called the spectrum of A

$$\mathrm{Spec}(A) = \{\lambda \in \mathbb{C}; \lambda \,\mathrm{Id} - A \notin \mathrm{GL}(H)\}. \tag{3.125}$$

As follows from the discussion above it is a compact set – it cannot be empty. You should resist the temptation to think that this is the set of eigenvalues of A, that is not really true.

For a bounded self-adjoint operator we can say more quite a bit more.

Proposition 34. *If $A : H \longrightarrow H$ is a bounded operator on a Hilbert space and $A^* = A$ then $A - \lambda \,\mathrm{Id}$ is invertible for all $\lambda \in \mathbb{C} \setminus \mathbb{R}$ and at least one of $A - \|A\| \,\mathrm{Id}$ and $A + \|A\| \,\mathrm{Id}$ is not invertible.*

The proof of the last part depends on a different characterization of the norm in the self-adjoint case.

Lemma 33. *If $A^* = A$ then*

$$\|A\| = \sup_{\|u\|=1} |\langle Au, u\rangle|. \tag{3.126}$$

Proof. Certainly, $|\langle Au, u\rangle| \le \|A\|\|u\|^2$ so the right side can only be smaller than or equal to the left. Suppose that

$$\sup_{\|u\|=1} |\langle Au, u\rangle| = a.$$

Then for any $u, v \in H$, $|\langle Au, v\rangle| = \langle Ae^{i\theta}u, v\rangle$ for some $\theta \in [0, 2\pi)$, so we can arrange that $\langle Au, v\rangle = |\langle Au', v\rangle|$ is non-negative and $\|u'\| = 1 = \|u\| = \|v\|$. Dropping the primes and computing using the polarization identity (really just the parallelogram law)

$$4\langle Au, v\rangle = \langle A(u+v), u+v\rangle - \langle A(u-v), u-v\rangle + i\langle A(u+iv), u+iv\rangle - i\langle A(u-iv), u-iv\rangle. \tag{3.127}$$

By the reality of the left side we can drop the last two terms and use the bound to see that

$$4\langle Au, v\rangle \le a(\|u+v\|^2 + \|u-v\|^2) = 2a(\|u\|^2 + \|v\|^2) = 4a \tag{3.128}$$

Thus, $\|A\| = \sup_{\|u\|=\|v\|=1} |\langle Au, v\rangle| \le a$ and hence $\|A\| = a$. □

Proof of Proposition 34. If $\lambda = s + it$ where $t \ne 0$ then $A - \lambda = (A - s) - it$ and $A - s$ is bounded and selfadjoint, so it is enough to consider the special case that $\lambda = it$. Then for any $u \in H$,

$$\operatorname{Im}\langle (A - it)u, u\rangle = -t\|u\|^2. \tag{3.129}$$

So, certainly $A - it$ is injective, since $(A - it)u = 0$ implies $u = 0$ if $t \ne 0$. The adjoint of $A - it$ is $A + it$ so the adjoint is injective too. It follows that the range of $A - it$ is dense in H. Indeed, if $v \in H$ and $v \perp (A - it)u$ for all $u \in H$, so v is orthogonal to the range, then

$$0 = \operatorname{Im}\langle (A - it)v, v\rangle = -t\|v\|^2. \tag{3.130}$$

By this density of the range, if $w \in H$ there exists a sequence u_n in H with $(A - it)u_n \to w$. But this implies that $\|u_n\|$ is bounded, since $t\|u_n\|^2 = -\operatorname{Im}\langle (A - it)u_n, u_n\rangle$ and hence we can pass to a weakly convergent subsequence, $u_n \rightharpoonup u$. Then $(A - it)u_n \rightharpoonup (A - it)u = w$ so $A - it$ is 1-1 and onto. From the Open Mapping Theorem, $(A - it)$ is invertible.

Finally then we need to show that one of $A \pm \|A\| \operatorname{Id}$ is NOT invertible. This follows from (3.126). Indeed, by the definition of sup there is a sequence $u_n \in H$ with $\|u_n\| = 1$ such that either $\langle Au_n, u_n\rangle \to \|A\|$ or $\langle Au_n, u_n\rangle \to -\|A\|$. We may pass to a weakly convergent subsequence and so assume $u_n \rightharpoonup u$. Assume we are in the first case, so this means $\langle (A - \|A\|)u_n, u_n\rangle \to 0$. Then

$$\begin{gathered} \|(A - \|A\|)u_n\|^2 = \|Au_n\|^2 - 2\|A\|\rangle Au_n, u_n\rangle + \|A\|^2\|u_n\|^2 \\ \|Au_n\|^2 - 2\|A\|\rangle(A - \|A\|)u_n, u_n\rangle - \|A\|^2\|u_n\|^2. \end{gathered} \tag{3.131}$$

The second two terms here have limit $-\|A\|^2$ by assumption and the first term is less than or equal to $\|A\|^2$. Since the sequence is positive it follows that $\|(A - \|A\|)^2 u_n\| \to 0$. This means that $A - \|A\| \operatorname{Id}$ is not invertible, since if it had a bounded inverse B then $1 = \|u_n\| \le \|B\|\|(A - \|A\|)^2 u_n\|$ which is impossible.

The other case is similar (or you can replace A by $-A$) so one of $A \pm \|A\|$ is not invertible. □

18. Spectral theorem for compact self-adjoint operators

One of the important differences between a general bounded self-adjoint operator and a compact self-adjoint operator is that the latter has eigenvalues and eigenvectors – lots of them.

Theorem 15. *If $A \in \mathcal{K}(\mathcal{H})$ is a self-adjoint, compact operator on a separable Hilbert space, so $A^* = A$, then H has an orthonormal basis consisting of eigenvectors of A, u_j such that*

$$Au_j = \lambda_j u_j,\ \ \lambda_j \in \mathbb{R} \setminus \{0\}, \tag{3.132}$$

consisting of an orthonormal basis for the possibly infinite-dimensional (closed) null space and eigenvectors with non-zero eigenvalues which can be arranged into a sequence such that $|\lambda_j|$ is a non-increasing and $\lambda_j \to 0$ as $j \to \infty$ (in case $\operatorname{Nul}(A)^\perp$ *is finite dimensional, this sequence is finite).*

The operator A maps $\operatorname{Nul}(A)^\perp$ into itself so it may be clearer to first split off the null space and then look at the operator acting on $\operatorname{Nul}(A)^\perp$ which has an orthonormal basis of eigenvectors with non-vanishing eigenvalues.

Before going to the proof, let's notice some useful conclusions. One is that we have 'Fredholm's alternative' in this case.

Corollary 4. *If $A \in \mathcal{K}(\mathcal{H})$ is a compact self-adjoint operator on a separable Hilbert space then the equation*

$$u - Au = f \tag{3.133}$$

either has a unique solution for each $f \in \mathcal{H}$ or else there is a non-trivial finite dimensional space of solutions to

$$u - Au = 0 \tag{3.134}$$

and then (3.133) *has a solution if and only if f is orthogonal to all these solutions.*

Proof. This is just saying that the null space of $\operatorname{Id} - A$ is a complement to the range – which is closed. So, either $\operatorname{Id} - A$ is invertible or if not then the range is precisely the orthocomplement of $\operatorname{Nul}(\operatorname{Id} - A)$. You might say there is not much alternative from this point of view, since it just says the range is *always* the orthocomplement of the null space. □

Let me separate off the heart of the argument from the bookkeeping.

Lemma 34. *If $A \in \mathcal{K}(\mathcal{H})$ is a self-adjoint compact operator on a separable (possibly finite-dimensional) Hilbert space then*

$$F(u) = (Au, u),\ \ F : \{u \in \mathcal{H}; \|u\| = 1\} \longrightarrow \mathbb{R} \tag{3.135}$$

is a continuous function on the unit sphere which attains its supremum and infimum where

$$\sup_{\|u\|=1} |F(u)| = \|A\|. \tag{3.136}$$

Furthermore, if the maximum or minimum of $F(u)$ is non-zero it is attained at an eivenvector of A with this extremal value as eigenvalue.

Proof. Since $|F(u)|$ is the function considered in (3.126), (3.136) is a direct consequence of Lemma 33. Moreover, continuity of F follows from continuity of A and of the inner product so

$$|F(u)-F(u')| \le |(Au,u)-(Au,u')|+|(Au,u')-(Au',u')| \le 2\|A\|\|u-u'\| \tag{3.137}$$

since both u and u' have norm one.

If we were in finite dimensions this almost finishes the proof, since the sphere is then compact and a continuous function on a compact set attains its sup and inf. In the general case we need to use the compactness of A. Certainly F is bounded,

$$|F(u)| \le \sup_{\|u\|=1} |(Au,u)| \le \|A\|. \tag{3.138}$$

Thus, there is a sequence u_n^+ such that $F(u_n^+) \to \sup F$ and another u_n^- such that $F(u_n^-) \to \inf F$. The *weak* compactness of the unit sphere means that we can pass to a weakly convergent subsequence in each case, and so assume that $u_n^\pm \rightharpoonup u^\pm$ converges weakly. Then, by the compactness of A, $Au_n^\pm \to Au^\pm$ converges strongly, i.e. in norm. But then we can write

$$\begin{aligned}|F(u_n^\pm) - F(u^\pm)| &\le |(A(u_n^\pm - u^\pm), u_n^\pm)| + |(Au^\pm, u_n^\pm - u^\pm)| \\ &= |(A(u_n^\pm - u^\pm), u_n^\pm)| + |(u^\pm, A(u_n^\pm - u^\pm))| \le 2\|Au_n^\pm - Au^\pm\|\end{aligned} \tag{3.139}$$

to deduce that $F(u^\pm) = \lim F(u_n^\pm)$ are respectively the sup and inf of F. Thus indeed, as in the finite dimensional case, the sup and inf are attained, and hence are the max and min. Note that this is NOT typically true if A is not compact as well as self-adjoint.

Now, suppose that $\Lambda^+ = \sup F > 0$. Then for any $v \in \mathcal{H}$ with $v \perp u^+$ and $\|v\| = 1$, the curve

$$L_v : (-\pi,\pi) \ni \theta \longmapsto \cos\theta u^+ + \sin\theta v \tag{3.140}$$

lies in the unit sphere. Expanding out

$$\begin{aligned}&F(L_v(\theta)) = \\ &\quad (AL_v(\theta), L_v(\theta)) = \cos^2\theta F(u^+) + 2\sin(2\theta)\operatorname{Re}(Au^+, v) + \sin^2(\theta)F(v)\end{aligned} \tag{3.141}$$

we know that this function must take its maximum at $\theta = 0$. The derivative there (it is certainly continuously differentiable on $(-\pi,\pi)$) is $\operatorname{Re}(Au^+, v)$ which must therefore vanish. The same is true for iv in place of v so in fact

$$(Au^+, v) = 0 \ \forall\ v \perp u^+,\ \|v\| = 1. \tag{3.142}$$

Taking the span of these v's it follows that $(Au^+, v) = 0$ for all $v \perp u^+$ so A^+u must be a multiple of u^+ itself. Inserting this into the definition of F it follows that $Au^+ = \Lambda^+ u^+$ is an eigenvector with eigenvalue $\Lambda^+ = \sup F$.

The same argument applies to $\inf F$ if it is negative, for instance by replacing A by $-A$. This completes the proof of the Lemma. □

Proof of Theorem 15. First consider the Hilbert space $\mathcal{H}_0 = \operatorname{Nul}(A)^\perp \subset \mathcal{H}$. Then, as noted above, A maps $\mathcal{H}_0$ into itself, since

$$(Au, v) = (u, Av) = 0 \ \forall\ u \in \mathcal{H}_0,\ v \in \operatorname{Nul}(A) \Longrightarrow Au \in \mathcal{H}_0. \tag{3.143}$$

Moreover, A_0, which is A restricted to $\mathcal{H}_0$, is again a compact self-adjoint operator – where the compactness follows from the fact that $A(B(0,1))$ for $B(0,1) \subset \mathcal{H}_0$ is smaller than (actually of course equal to) the whole image of the unit ball.

Thus we can apply the Lemma above to A_0, with quadratic form F_0, and find an eigenvector. Let's agree to take the one associated to $\sup F_0$ unless $\sup F_0 < -\inf F_0$ in which case we take one associated to the inf . Now, what can go wrong here? Nothing except if $F_0 \equiv 0$. However in that case we know from Lemma 33 that $\|A\| = 0$ so $A = 0$.

So, we now know that we can find an eigenvector with non-zero eigenvalue unless $A \equiv 0$ which would implies $\operatorname{Nul}(A) = \mathcal{H}$. Now we proceed by induction. Suppose we have found N mutually orthogonal eigenvectors e_j for A all with norm 1 and eigenvectors λ_j – an orthonormal set of eigenvectors and all in $\mathcal{H}_0$. Then we consider

$$\mathcal{H}_N = \{u \in \mathcal{H}_0 = \operatorname{Nul}(A)^\perp; (u, e_j) = 0,\ j = 1, \dots, N\}. \tag{3.144}$$

From the argument above, A maps $\mathcal{H}_N$ into itself, since

$$(Au, e_j) = (u, Ae_j) = \lambda_j(u, e_j) = 0 \text{ if } u \in \mathcal{H}_N \Longrightarrow Au \in \mathcal{H}_N. \tag{3.145}$$

Moreover this restricted operator is self-adjoint and compact on $\mathcal{H}_N$ as before so we can again find an eigenvector, with eigenvalue either the max of min of the new F for $\mathcal{H}_N$. This process will not stop uness $F \equiv 0$ at some stage, but then $A \equiv 0$ on $\mathcal{H}_N$ and since $\mathcal{H}_N \perp \operatorname{Nul}(A)$ which implies $\mathcal{H}_N = \{0\}$ so $\mathcal{H}_0$ must have been finite dimensional.

Thus, either $\mathcal{H}_0$ is finite dimensional or we can grind out an infinite orthonormal sequence e_i of eigenvectors of A in $\mathcal{H}_0$ with the corresponding sequence of eigenvalues such that $|\lambda_i|$ is non-increasing – since the successive F_N's are restrictions of the previous ones the max and min are getting closer to (or at least no further from) 0.

So we need to rule out the possibility that there is an infinite orthonormal sequence of eigenfunctions e_j with corresponding eigenvalues λ_j where $\inf_j |\lambda_j| = a > 0$. Such a sequence cannot exist since $e_j \rightharpoonup 0$ so by the compactness of A, $Ae_j \to 0$ (in norm) but $|Ae_j| \ge a$ which is a contradiction. Thus if $\operatorname{null}(A)^\perp$ is not finite dimensional then the sequence of eigenvalues constructed above must converge to 0.

Finally then, we need to check that this orthonormal sequence of eigenvectors constitutes an orthonormal basis of $\mathcal{H}_0$. If not, then we can form the closure of the span of the e_i we have constructed, $\mathcal{H}'$, and its orthocomplement in $\mathcal{H}_0$ – which would have to be non-trivial. However, as before F restricts to this space to be F' for the restriction of A' to it, which is again a compact self-adjoint operator. So, if F' is not identically zero we can again construct an eigenfunction, with non-zero eigenvalue, which contracdicts the fact the we are always choosing a largest eigenvalue, in absolute value at least. Thus in fact $F' \equiv 0$ so $A' \equiv 0$ and the eigenvectors form and orthonormal basis of $\operatorname{Nul}(A)^\perp$. This completes the proof of the theorem. □

19. Functional Calculus

So the non-zero eigenvalues of a compact self-adjoint operator form the image of a sequence in $[-\|A\|, \|A\|]$ either converging to zero or finite. If $f \in \mathcal{C}^0([-\|A\|, \|A\|)$ then one can define an operator

$$f(A) \in \mathcal{B}(H),\ f(A)u = \sum_i f(\lambda_u)(u, e_i)e_i \tag{3.146}$$

where $\{e_i\}$ is a complete orthonormal basis of eigenfunctions. Provided $f(0) = 0$ this is compact and if f is real it is self-adjoint. This formula actually defines a linear map

$$\mathcal{C}^0([-\|A\|, \|A\|]) \longrightarrow \mathcal{B}(H) \text{ with } f(A)g(A) = (fg)(A). \tag{3.147}$$

Such a map exists for any bounded self-adjoint operator. Even though it may not have eigenfunctions – or not a complete orthonormal basis of them anyway, it is still possible to define $f(A)$ for a continous function defined on $[-\|A\|, \|A\|]$ (in fact it only has to be defined on $\operatorname{Spec}(A) \subset [-\|A\|, \|A\|]$ which might be quite a lot smaller). This is an effective replacement for the spectral theorem in the compact case.

How does one define $f(A)$? Well, it is easy enough in case f is a polynomial, since then we can factorize it and set

$$f(z) = c(z - z_1)(z - z_2) \dots (z - z_N) \Longrightarrow f(A) = c(A - z_1)(A - z_2) \dots (A - z_N). \tag{3.148}$$

Notice that the result does not depend on the order of the factors or anything like that. To pass to the case of a general continuous function on $[-\|A\|, \|A\|]$ one can use the norm estimate in the polynomial case, that

$$\|f(A)\| \le \sup_{z \in [-\|A\|, \|A\|} |f(z)|. \tag{3.149}$$

This allows one to pass f in the uniform closure of the polynomials, which by the Stone-Weierstrass theorem is the whole of $\mathcal{C}^0([-\|A\|, \|A\|])$. The proof of (3.149) is outlined in Problem 5.33 below.

20. Compact perturbations of the identity

I have generally not had a chance to discuss most of the material in this section, or the next, in the lectures.

Compact operators are, as we know, 'small' in the sense that the are norm limits of finite rank operators. If you accept this, then you will want to say that an operator such as

$$\operatorname{Id} - K,\ K \in \mathcal{K}(\mathcal{H}) \tag{3.150}$$

is 'big'. We are quite interested in this operator because of spectral theory. To say that $\lambda \in \mathbb{C}$ is an eigenvalue of K is to say that there is a non-trivial solution of

$$Ku - \lambda u = 0 \tag{3.151}$$

where non-trivial means other than than the solution $u = 0$ which always exists. If λ is an eigenvalue of K then certainly $\lambda \in \operatorname{Spec}(K)$, since $\lambda - K$ cannot be invertible. For general operators the converse is not correct, but for compact operators it is.

Lemma 35. *If $K \in \mathcal{B}(H)$ is a compact operator then $\lambda \in \mathbb{C}\backslash\{0\}$ is an eigenvalue of K if and only if $\lambda \in \operatorname{Spec}(K)$.*

Proof. Since we can divide by λ we may replace K by $\lambda^{-1}K$ and consider the special case $\lambda = 1$. Now, if K is actually finite rank the result is straightforward. By Lemma 26 we can choose a basis so that (3.76) holds. Let the span of the e_i be W – since it is finite dimensional it is closed. Then $\operatorname{Id} - K$ acts rather simply – decomposing $H = W \oplus W^\perp$, $u = w + w'$

$$(\operatorname{Id} - K)(w + w') = w + (\operatorname{Id}_W - K')w',\ K' : W \longrightarrow W \tag{3.152}$$

being a matrix with respect to the basis. Now, 1 is an eigenvalue of K if and only if 1 is an eigenvalue of K' as an operator on the finite-dimensional space W. Now, a matrix, such as $\operatorname{Id}_W - K'$, is invertible if and only if it is injective, or equivalently surjective. So, the same is true for $\operatorname{Id} - K$.

In the general case we use the approximability of K by finite rank operators. Thus, we can choose a finite rank operator F such that $\|K - F\| < 1/2$. Thus, $(\operatorname{Id} - K + F)^{-1} = \operatorname{Id} - B$ is invertible. Then we can write

$$\operatorname{Id} - K = \operatorname{Id} - (K - F) - F = (\operatorname{Id} - (K - F))(\operatorname{Id} - L),\ L = (\operatorname{Id} - B)F. \tag{3.153}$$

Thus, $\operatorname{Id} - K$ is invertible if and only if $\operatorname{Id} - L$ is invertible. Thus, if $\operatorname{Id} - K$ is *not* invertible then $\operatorname{Id} - L$ is not invertible and hence has null space and from (3.153) it follows that $\operatorname{Id} - K$ has non-trivial null space, i.e. K has 1 as an eigenvalue. □

A little more generally:-

Proposition 35. *If $K \in \mathcal{K}(\mathcal{H})$ is a compact operator on a separable Hilbert space then*

$$\begin{gathered} \operatorname{null}(\operatorname{Id} - K) = \{u \in \mathcal{H}; (\operatorname{Id}_K)u = 0\} \text{ is finite dimensional} \\ \operatorname{Ran}(\operatorname{Id} - K) = \{v \in \mathcal{H}; \exists u \in \mathcal{H},\ v = (\operatorname{Id} - K)u\} \text{ is closed and} \\ \operatorname{Ran}(\operatorname{Id} - K)^{\perp} = \{w \in \mathcal{H}; (w, Ku) = 0\ \forall\ u \in \mathcal{H}\} \text{ is finite dimensional} \end{gathered} \tag{3.154}$$

and moreover

$$\dim\left(\operatorname{null}(\operatorname{Id} - K)\right) = \dim\left(\operatorname{Ran}(\operatorname{Id} - K)^{\perp}\right). \tag{3.155}$$

Proof of Proposition 35. First let's check this in the case of a finite rank operator $K = T$. Then

$$\operatorname{Nul}(\operatorname{Id} - T) = \{u \in \mathcal{H}; u = Tu\} \subset \operatorname{Ran}(T). \tag{3.156}$$

A subspace of a finite dimensional space is certainly finite dimensional, so this proves the first condition in the finite rank case.

Similarly, still assuming that T is finite rank consider the range

$$\operatorname{Ran}(\operatorname{Id} - T) = \{v \in \mathcal{H}; v = (\operatorname{Id} - T)u \text{ for some } u \in \mathcal{H}\}. \tag{3.157}$$

Consider the subspace $\{u \in \mathcal{H}; Tu = 0\}$. We know that this this is closed, since T is certainly continuous. On the other hand from (3.157),

$$\operatorname{Ran}(\operatorname{Id} - T) \supset \operatorname{Nul}(T). \tag{3.158}$$

Remember that a finite rank operator can be written out as a finite sum

$$Tu = \sum_{i=1}^{N} (u, e_i) f_i \tag{3.159}$$

where we can take the f_i to be a basis of the range of T. We also know in this case that the e_i must be linearly independent – if they weren't then we could write one of them, say the last since we can renumber, out as a sum, $e_N = \sum\limits_{j<N} c_i e_j$, of multiples of the others and then find

$$Tu = \sum_{i=1}^{N-1} (u, e_i)(f_i + \overline{c_j} f_N) \tag{3.160}$$

showing that the range of T has dimension at most $N - 1$, contradicting the fact that the f_i span it.

So, going back to (3.159) we know that $\mathrm{Nul}(T)$ has finite *codimension* – every element of $\mathcal{H}$ is of the form

$$u = u' + \sum_{i=1}^{N} d_i e_i,\ u' \in \mathrm{Nul}(T). \tag{3.161}$$

So, going back to (3.158), if $\mathrm{Ran}(\mathrm{Id} - T) \neq \mathrm{Nul}(T)$, and it need not be equal, we can choose – using the fact that $\mathrm{Nul}(T)$ *is* closed – an element $g \in \mathrm{Ran}(\mathrm{Id} - T) \setminus \mathrm{Nul}(T)$ which is orthogonal to $\mathrm{Nul}(T)$. To do this, start with any a vector g' in $\mathrm{Ran}(\mathrm{Id} - T)$ which is not in $\mathrm{Nul}(T)$. It can be split as $g' = u'' + g$ where $g \perp \mathrm{Nul}(T)$ (being a closed subspace) and $u'' \in \mathrm{Nul}(T)$, then $g \neq 0$ is in $\mathrm{Ran}(\mathrm{Id} - T)$ and orthongonal to $\mathrm{Nul}(T)$. Now, the new space $\mathrm{Nul}(T) \oplus \mathbb{C}g$ is again closed and contained in $\mathrm{Ran}(\mathrm{Id} - T)$. But we can continue this process replacing $\mathrm{Nul}(T)$ by this larger closed subspace. After a a finite number of steps we conclude that $\mathrm{Ran}(\mathrm{Id} - T)$ itself is closed.

What we have just proved is:

Lemma 36. *If $V \subset \mathcal{H}$ is a subspace of a Hilbert space which contains a closed subspace of finite codimension in $\mathcal{H}$ – meaning $V \supset W$ where W is closed and there are finitely many elements $e_i \in \mathcal{H}$, $i = 1, \dots, N$ such that every element $u \in \mathcal{H}$ is of the form*

$$u = u' + \sum_{i=1}^{N} c_i e_i,\ c_i \in \mathbb{C}, \tag{3.162}$$

then V itself is closed.

So, this takes care of the case that $K = T$ has finite rank! What about the general case where K is compact? Here we just use a consequence of the approximation of compact operators by finite rank operators proved last time. Namely, if K is compact then there exists $B \in \mathcal{B}(\mathcal{H})$ and T of finite rank such that

$$K = B + T,\ \|B\| < \frac{1}{2}. \tag{3.163}$$

Now, consider the null space of $\mathrm{Id} - K$ and use (3.163) to write

$$\mathrm{Id} - K = (\mathrm{Id} - B) - T = (\mathrm{Id} - B)(\mathrm{Id} - T'),\ T' = (\mathrm{Id} - B)^{-1}T. \tag{3.164}$$

Here we have used the convergence of the Neumann series, so $(\mathrm{Id} - B)^{-1}$ does exist. Now, T' is of finite rank, by the ideal property, so

$$\mathrm{Nul}(\mathrm{Id} - K) = \mathrm{Nul}(\mathrm{Id} - T') \text{ is finite dimensional.} \tag{3.165}$$

Here of course we use the fact that $(\mathrm{Id} - K)u = 0$ is equivalent to $(\mathrm{Id} - T')u = 0$ since $\mathrm{Id} - B$ is invertible. So, this is the first condition in (3.154).

Similarly, to examine the second we do the same thing but the other way around and write

$$\mathrm{Id} - K = (\mathrm{Id} - B) - T = (\mathrm{Id} - T'')(\mathrm{Id} - B),\ T'' = T(\mathrm{Id} - B)^{-1}. \tag{3.166}$$

Now, T'' is again of finite rank and

$$\mathrm{Ran}(\mathrm{Id} - K) = \mathrm{Ran}(\mathrm{Id} - T'') \text{ is closed} \tag{3.167}$$

again using the fact that $\mathrm{Id} - B$ is invertible – so every element of the form $(\mathrm{Id} - K)u$ is of the form $(\mathrm{Id} - T'')u'$ where $u' = (\mathrm{Id} - B)u$ and conversely.

So, now we have proved all of (3.154) – the third part following from the first as discussed before.

What about (3.155)? This time let's first check that it is enough to consider the finite rank case. For a compact operator we have written

$$(\operatorname{Id}-K) = G(\operatorname{Id}-T) \tag{3.168}$$

where $G = \operatorname{Id}-B$ with $\|B\| < \frac{1}{2}$ is invertible and T is of finite rank. So what we want to see is that

$$\dim \operatorname{Nul}(\operatorname{Id}-K) = \dim \operatorname{Nul}(\operatorname{Id}-T) = \dim \operatorname{Nul}(\operatorname{Id}-K^*). \tag{3.169}$$

However, $\operatorname{Id}-K^* = (\operatorname{Id}-T^*)G^*$ and G^* is also invertible, so

$$\dim \operatorname{Nul}(\operatorname{Id}-K^*) = \dim \operatorname{Nul}(\operatorname{Id}-T^*) \tag{3.170}$$

and hence it is enough to check that $\dim \operatorname{Nul}(\operatorname{Id}-T) = \dim \operatorname{Nul}(\operatorname{Id}-T^*)$ – which is to say the same thing for finite rank operators.

Now, for a finite rank operator, written out as (3.159), we can look at the vector space W spanned by all the f_i's and all the e_i's together – note that there is nothing to stop there being dependence relations among the combination although separately they are independent. Now, $T : W \longrightarrow W$ as is immediately clear and

$$T^*v = \sum_{i=1}^{N} (v, f_i)e_i \tag{3.171}$$

so $T : W \longrightarrow W$ too. In fact $Tw' = 0$ and $T^*w' = 0$ if $w' \in W^\perp$ since then $(w', e_i) = 0$ and $(w', f_i) = 0$ for all i. It follows that if we write $R : W \longleftrightarrow W$ for the linear map on this finite dimensional space which is equal to $\operatorname{Id}-T$ acting on it, then R^* is given by $\operatorname{Id}-T^*$ acting on W and we use the Hilbert space structure on W induced as a subspace of $\mathcal{H}$. So, what we have just shown is that

$$(\operatorname{Id}-T)u = 0 \Longleftrightarrow u \in W \text{ and } Ru = 0,\ (\operatorname{Id}-T^*)u = 0 \Longleftrightarrow u \in W \text{ and } R^*u = 0. \tag{3.172}$$

Thus we really are reduced to the finite-dimensional theorem

$$\dim \operatorname{Nul}(R) = \dim \operatorname{Nul}(R^*) \text{ on } W. \tag{3.173}$$

You no doubt know this result. It follows by observing that in this case, everything now on W, $\operatorname{Ran}(W) = \operatorname{Nul}(R^*)^\perp$ and finite dimensions

$$\dim \operatorname{Nul}(R) + \dim \operatorname{Ran}(R) = \dim W = \dim \operatorname{Ran}(W) + \dim \operatorname{Nul}(R^*). \tag{3.174}$$

□

21. Fredholm operators

Definition 21. A bounded operator $F \in \mathcal{B}(\mathcal{H})$ on a Hilbert space is said to be *Fredholm*, written $F \in \mathcal{F}(H)$, if it has the three properties in (3.154) – its null space is finite dimensional, its range is closed and the orthocomplement of its range is finite dimensional.

For general Fredholm operators the row-rank=colum-rank result (3.155) does not hold. Indeed the difference of these two integers, called the index of the operator,

$$\operatorname{ind}(F) = \dim\left(\operatorname{null}(\operatorname{Id}-K)\right) - \dim\left(\operatorname{Ran}(\operatorname{Id}-K)^\perp\right) \tag{3.175}$$

is a very important number with lots of interesting properties and uses.

Notice that the last two conditions in (3.154) are really independent since the orthocomplement of a subspace is the same as the orthocomplement of its closure. There is for instance a bounded operator on a separable Hilbert space with trivial null space and dense range which is not closed. How could this be? Think for instance of the operator on $L^2(0,1)$ which is multiplication by the function x. This is assuredly bounded and an element of the null space would have to satisfy $xu(x) = 0$ almost everywhere, and hence vanish almost everywhere. Moreover the density of the L^2 functions vanishing in $x < \epsilon$ for some (non-fixed) $\epsilon > 0$ shows that the range is dense. However it is clearly not invertible.

Before proving this result let's check that, in the case of operators of the form $\operatorname{Id} - K$, with K compact the third conclusion in (3.154) really follows from the first. This is a general fact which I mentioned, at least, earlier but let me pause to prove it.

Proposition 36. *If $B \in \mathcal{B}(\mathcal{H})$ is a bounded operator on a Hilbert space and B^* is its adjoint then*

$$\operatorname{Ran}(B)^\perp = (\overline{\operatorname{Ran}}(B))^\perp = \{v \in \mathcal{H}; (v,w) = 0 \ \forall \ w \in \operatorname{Ran}(B)\} = \operatorname{Nul}(B^*). \tag{3.176}$$

Proof. The definition of the orthocomplement of $\operatorname{Ran}(B)$ shows immediately that

$$\begin{aligned} v \in (\operatorname{Ran}(B))^\perp &\Longleftrightarrow (v,w) = 0 \ \forall \ w \in \operatorname{Ran}(B) \longleftrightarrow (v, Bu) = 0 \ \forall \ u \in \mathcal{H} \\ &\Longleftrightarrow (B^*v, u) = 0 \ \forall \ u \in \mathcal{H} \Longleftrightarrow B^*v = 0 \Longleftrightarrow v \in \operatorname{Nul}(B^*). \end{aligned} \tag{3.177}$$

On the other hand we have already observed that $V^\perp = (\overline{V})^\perp$ for any subspace – since the right side is certainly contained in the left and $(u,v) = 0$ for all $v \in V$ implies that $(u,w) = 0$ for all $w \in \overline{V}$ by using the continuity of the inner product to pass to the limit of a sequence $v_n \to w$. □

Thus as a corrollary we see that if $\operatorname{Nul}(\operatorname{Id} - K)$ is always finite dimensional for K compact (i. e. we check it for all compact operators) then $\operatorname{Nul}(\operatorname{Id} - K^*)$ is finite dimensional and hence so is $\operatorname{Ran}(\operatorname{Id} - K)^\perp$.

There is a more 'analytic' way of characterizing Fredholm operators, rather than Definition 21.

Lemma 37. *An operator $F \in \mathcal{B}(H)$ is Fredholm, $F \in \mathcal{F}(H)$, if and only if it has a generalized inverse P satisfying*

$$\begin{aligned} PF &= \operatorname{Id} - \Pi_{(F)} \\ FP &= \operatorname{Id} - \Pi_{(F)^\perp} \end{aligned} \tag{3.178}$$

with the two projections of finite rank.

Proof. If (3.178) holds then F must be Fredholm, since its null space is finite dimensional, from the second identity the range of F must contain the range of $\operatorname{Id} - Pi_{(F)^\perp}$ and hence it must be closed and of finite codimension (and in fact be equal to this closed subspace.

Conversely, suppose that $F \in \mathcal{F}(H)$. We can divide H into two pieces in two ways as $H = (F) \oplus (F)^\perp$ and $H = \operatorname{Ran}(F)^\perp \oplus \operatorname{Ran}(F)$ where in each case the first summand is finite-dimensional. Then F defines four maps, from each of the two first summands to each of the two second ones but all but one of these is zero and so F corresponds to a bounded linear map $\tilde{F} : (F)^\perp \longrightarrow \operatorname{Ran}(F)$. These are two Hilbert spaces with bounded linear bijection between them, so the inverse map, $\tilde{P} : \mathrm{Ran}(F) \longrightarrow (F)^\perp$ is bounded by the Open Mapping Theorem and we can define

$$P = \tilde{P} \circ \Pi(F)^\perp v). \tag{3.179}$$

Then (3.178) follows directly. □

What we want to show is that the Fredholm operators form an open set in $\mathcal{B}(H)$ and that the index is locally constant. To do this we show that a weaker version of (3.178) also implies that F is Fredholm.

Lemma 38. *An operator $F \in \mathcal{F}(H)$ is Fredholm if and only if it has a* parametrix $Q \in \mathcal{B}(H)$ *in the sense that*

$$\begin{aligned} QF &= \mathrm{Id} - E_R \\ FQ &= \mathrm{Id} - E_L \end{aligned} \tag{3.180}$$

with E_R and E_L of finite rank. Moreover any two such parametrices differ by a finite rank operator.

Proof. If F is Fredholm then $Q = P$ certainly is a parameterix in this sense. Conversely suppose that Q as in (**??**) exists. Then $(\mathrm{Id} - E_R)$ is finite dimensional – from (3.154) for instance. However, from the first identity $(F) \subset (QF) = (\mathrm{Id} - E_R)$ so (F) is finite dimensional too. Similarly, the second identity shows that $\mathrm{Ran}(F) \supset \mathrm{Ran}(FQ) = \mathrm{Ran}(\mathrm{Id} - E_L)$ and the last space is closed and of finite codimension, hence so is the first.

Now if Q and Q' both satisfy (3.180) with finite ranke error terms E'_R and E'_L for Q' then

$$(Q' - Q)F = E_R - E'_R \tag{3.181}$$

is of finite rank. Applying the generalized inverse, P of F on the right shows that the difference

$$(Q' - Q) = (E_R - E'_R)P + (Q' - Q)\Pi_{(F)} \tag{3.182}$$

is indeed of finite rank. □

Now recall (in 2014 from Problems7) that finite-rank operators are of trace class, that the trace is well-defined and that the trace of a commutator where one factor is bounded and the other trace class vanishes. Using this we show

Lemma 39. *If Q and F satisfy* (3.180) *then*

$$\mathrm{ind}(F) = \mathrm{Tr}(E_L) - \mathrm{Tr}(E_R). \tag{3.183}$$

Proof. We certainly know that (3.183) holds in the special case that $Q = P$ is the generalized inverse of F, since then $E_L = \Pi_{(F)}$ and $E_R = \Pi_{\mathrm{Ran}(F)^\perp}$ and the traces are the dimensions of these spaces.

Now, if Q is a parameterix as in (3.180) consider the straight line of operators $Q_t = (1-t)P + tQ$. Using the two sets of identities for the generalized inverse and paramaterix

$$\begin{aligned} Q_tF &= (1-t)PF + tQF = \mathrm{Id} - (1-t)\Pi_{(F)} - tE_L, \\ FQ_t &= (1-t)FP + tFQ = \mathrm{Id} - (1-t)\Pi_{\mathrm{Ran}(F)^\perp} - tE_R. \end{aligned} \tag{3.184}$$

Thus Q_t is a curve of parameterices and what we need to show is that

$$J(t) = \mathrm{Tr}((1-t)\Pi_{(F)} + tE_L) - \mathrm{Tr}((1-t)\Pi_{\mathrm{Ran}(F)^\perp} + tE_R) \tag{3.185}$$

is constant. This is a linear function of t as is Q_t. We can differentiate (3.184) with respect to t and see that

$$\frac{d}{dt}((1-t)\Pi_{(F)}+tE_L)-\frac{d}{dt}((1-t)\Pi_{\mathrm{Ran}(F)^\perp}+tE_R)=[Q-P,F] \\ \Longrightarrow J'(t)=0 \tag{3.186}$$

since it is the trace of the commutator of a bounded and a finite rank operator (using the last part of Lemma 38. □

Proposition 37. *The Fredholm operators form an open set in $\mathcal{B}(H)$ on which the index is locally constant.*

Proof. We need to show that if F is Fredholm then there exists $\epsilon>0$ such that $F+B$ is Fredholm if $\|B\|<\epsilon$. Set $B'=\Pi_{\mathrm{Ran}(F)}B\Pi_{(F)^\perp}$ then $\|B'\|\le\|B\|$ and $B-B'$ is finite rank. If $\tilde F$ is the operator constructed in the proof of Lemma 37 then $\tilde F+B'$ is invertible as an operator from $(F)^\perp$ to $\mathrm{Ran}(F)$ if $\epsilon>0$ is small. The inverse, P'_B, extended as 0 to (F) as P is defined in that proof, satisfies

$$\begin{gathered} P'_B(F+B)=\mathrm{Id}-\Pi_{(F)}+P'_B(B-B'),\\ (F+B)P'_B=\mathrm{Id}-\Pi)\mathrm{Ran}(F)^\perp+(B-B])P'_B \end{gathered} \tag{3.187}$$

and so is a parametrix for $F+B$. Thus the set of Fredholm operators is open.

The index of $F+B$ is given by the difference of the trace of the finite rank error terms in the second and first lines here. It depends continuously on B in $\|B\|<\epsilon$ so, being integer valued, is constant. □

This shows in particular that there is an open subset of $\mathcal{B}(H)$ which contains no invertible operators, in strong contrast to the finite dimensional case. Still even the Fredholm operators do no form a dense subset of $\mathcal{B}(H)$. One such open subset consists of the *sem-Fredholm* operators, those with closed range and with *either* null space of complement of range finite-dimensional.

22. Kuiper's theorem

For finite dimensional spaces, such as $\mathbb{C}^N$, the group of invertible operators, denoted typically $\mathrm{GL}(N)$, is a particularly important example of a Lie group. One reason it is important is that it carries a good deal of 'topological' structure. In particular – I'm assuming you have done a little topology – its fundamental group is not trivial, in fact it is isomorphic to $\mathbb{Z}$. This corresponds to the fact that a continuous closed curve $c:\mathbb{S}\longrightarrow\mathrm{GL}(N)$ is *contractible* if and only if its winding number is zero – the effective number of times that the determinant goes around the origin in $\mathbb{C}$. There is a lot more topology than this and it is actually quite complicated.

Perhaps surprisingly, the corresponding group of the bounded operators on a separable (complex) infinite-dimensional Hilbert space which have bounded inverses (or equivalently those which are bijections in view of the open mapping theorem) is contractible. This is Kuiper's theorem, and means that this group, $\mathrm{GL}(H)$, has no 'topology' at all, no holes in any dimension and for topological purposes it is like a big open ball. The proof is not really hard, but it is not exactly obvious either. It depends on an earlier idea, 'Eilenberg's swindle', which shows how the infinite-dimensionality is exploited. As you can guess, this is sort of amusing (if you have the right attitude . . .).

Let's denote by $\mathrm{GL}(H)$ this group, as remarked above in view of the open mapping theorem we know that

$$\mathrm{GL}(H) = \{A \in \mathcal{B}(H); A \text{ is injective and surjective.}\}. \tag{3.188}$$

Contractibility is the topological notion of 'topologically trivial'. It means precisely that there is a continuous map

$$\begin{gathered}\gamma : [0, 1] \times \mathrm{GL}(H) \longrightarrow \mathrm{GL}(H) \text{ s.t.}\\ \gamma(0, A) = A,\ \gamma(1, A) = \mathrm{Id},\ \forall\ A \in \mathrm{GL}(H).\end{gathered} \tag{3.189}$$

Continuity here means for the metric space $[0, 1] \times \mathrm{GL}(H)$ where the metric comes from the norms on $\mathbb{R}$ and $\mathcal{B}(H)$.

As a warm-up exercise, let us show that the group $\mathrm{GL}(H)$ is contractible to the unitary subgroup

$$\mathrm{U}(H) = \{U \in \mathrm{GL}(H); U^{-1} = U^*\}. \tag{3.190}$$

These are the isometric isomorphisms.

Proposition 38. *There is a continuous map*

$$\Gamma : [0, 1] \times \mathrm{GL}(H) \longrightarrow \mathrm{GL}(H) \textit{ s.t. } \Gamma(0, A) = A,\ \Gamma(1, A) \in \mathrm{U}(H)\ \forall\ A \in \mathrm{GL}(H). \tag{3.191}$$

Proof. This is a consequence of the functional calculus, giving the 'polar decomposition' of invertible (and more generally bounded) operators. Namely, if $A\,\mathrm{GL}(H)$ then $AA^* \in \mathrm{GL}(H)$ is self-adjoint. Its spectrum is then contained in an interval $[a, b]$, where $0 < a \le b = \|A\|^2$. It follows from what we showed earlier that $R = (AA^*)^{\frac{1}{2}}$ is a well-defined bounded self-adjoint operator and $R^2 = AA^*$. Moreover, R is invertible and the operator $U_A = R^{-1}A \in \mathrm{U}(H)$. Certainly it is bounded and $U_A^* = A^*R^{-1}$ so $U_A^*U_A = A^*R^{-2}A = \mathrm{Id}$ since $R^{-2} = (AA^*)^{-1} = (A^*)^{-1}A^{-1}$. Thus U_A^* is a right inverse of U_A, and (since U_A is a bijection) is the unique inverse so $U_A \in \mathrm{U}(H)$. So we have shown $A = RU_A$ (this is the polar decomposition) and then

$$\Gamma(s, A) = (s\,\mathrm{Id} + (1 - s)R)U_A,\ s \in [0, 1] \tag{3.192}$$

satisfies (3.191). □

Initially we will consider only the notion of 'weak contractibility'. This has nothing to do with weak convergence, rather just means that we only look for an homotopy over compact sets. So, for any compact subset $X \subset \mathrm{GL}(H)$ we seek a continuous map

$$\begin{gathered}\gamma : [0, 1] \times X \longrightarrow \mathrm{GL}(H) \text{ s.t.}\\ \gamma(0, A) = A,\ \gamma(1, A) = \mathrm{Id},\ \forall\ A \in X,\end{gathered} \tag{3.193}$$

note that this is not contractibility *of* X, but *of* X in $\mathrm{GL}(H)$.

In fact, to carry out the construction without having to worry about too many things at one, just consider (path) connectedness of $\mathrm{GL}(H)$ meaning that there is a continuous map as in (3.193) where $X = \{A\}$ just consists of one point – so the map is just $\gamma : [0, 1] \longrightarrow \mathrm{GL}(H)$ such that $\gamma(0) = A$, $\gamma(1) = \mathrm{Id}$.

The construction of γ is in three stages

1. Creating a gap
2. Rotating to a trivial factor
3. Eilenberg's swindle.

This approach follows ideas of B. Mityagin, [**2**].

Lemma 40 (Creating a gap). *If* $A \in \mathcal{B}(H)$ *and* $\epsilon > 0$ *is given there is a decomposition* $H = H_K \oplus H_L \oplus H_O$ *into three closed mutually orthogonal infinite-dimensional subspaces such that if* Q_I *is the orthogonal projections onto* H_I *for* $I = K,\ L,\ O$ *then*

$$\|Q_L B Q_K\| < \epsilon. \tag{3.194}$$

Proof. Choose an orthonormal basis e_j, $j \in \mathbb{N}$, of H. The subspaces H_i will be determined by a corresponding decomposition

$$\mathbb{N} = K \cup L \cup O,\ K \cap L = K \cap O = L \cap O = \emptyset. \tag{3.195}$$

Thus H_I has orthonormal basis e_k, $k \in I$, $I = K,\ L,\ O$. To ensure (3.194) we choose the decomposition (3.195) so that all three sets are infinite and so that

$$|(e_l, Be_k)| < 2^{-l-1}\epsilon \ \forall\ l \in L,\ k \in K. \tag{3.196}$$

Once we have this, then for $u \in H$, $Q_K u \in H_K$ can be expanded to $\sum\limits_{k\in K}(Q_k u, e_k)e_k$ and expanding in H_L similalry,

$$\begin{aligned} Q_L B Q_K u &= \sum_{l\in L}(BQ_K u, e_l)e_l = \sum_{k\in L}\sum_{k\in K}(Be_k, e_l)(Q_K u, e_k)e_l \\ \Longrightarrow \|Q_L B Q_K u\|^2 &\le \sum_{k\in K}\left(|(Q_k u, e_k)|^2 \sum_{l\in L}|(Be_k, e_l)|^2\right) \\ &\le \frac{1}{2}\epsilon^2 \sum_{k\in K}|(Q_k u, e_k)|^2 \le \frac{1}{2}\epsilon^2\|u\|^2 \end{aligned} \tag{3.197}$$

giving (3.194). The absolute convergence of the series following from (3.196).

Thus, it remains to find a decomposition (3.195) for which (3.196) holds. This follows from Bessel's inequality. First choose $1 \in K$ then $(Be_1, e_l) \to 0$ as $l \to \infty$ so $|(Be_1, e_{l_1})| < \epsilon/4$ for l_1 large enough and we will take $l_1 > 2k_1$. Then we use induction on N, choosing $K(N)$, $L(N)$ and $O(N)$ with

$$\begin{gathered} K(N) = \{k_1 = 1 < k_2 < \dots, k_N\}, \\ L(N) = \{l_1 < l_2 < \cdots < l_N\},\ l_r > 2k_r,\ k_r > l_{r-1} \text{ for } 1 < r \le N \text{ and} \\ O(N) = \{1, \dots, l_N\} \setminus (K(N) \cup L(N)). \end{gathered}$$

Now, choose $k_{N+1} > l_N$ by such that $|(e_l, Be_{k_{N+1}})| < 2^{-l-N}\epsilon$, for all $l \in L(N)$, and then $l_{N+1} > 2k_{N+1}$ such that $|(e_{l_{N+1}}, B_k)| < e^{-N-1-k}\epsilon$ for $k \in K(N+1) = K(N) \cup \{k_{N+1}\}$ and the inductive hypothesis follows with $L(N+1) = N(N) \cup \{l_{N+1}\}$.

Given a fixed operator $A \in \mathrm{GL}(H)$ Lemma 40 can be applied with $\epsilon = \|A^{-1}\|^{-1}$. It then follows that the curve

$$A(s) = A - sQ_L A Q_K,\ s \in [0,1] \tag{3.198}$$

lies in $\mathrm{GL}(H)$ and has endpoint satisfying

$$Q_L B Q_K = 0,\ B = A(1),\ Q_L Q_K = 0 = Q_K Q_L,\ Q_K = Q_K^2,\ Q_L = Q_L^2 \tag{3.199}$$

where all three projections, Q_L, Q_K and $\mathrm{Id} - Q_K - Q_L$ have infinite rank.

These three projections given an identification of $H = H \oplus H \oplus H$ and so replace the bounded operators by 3×3 matrices with entries which are bounded operators on H. The condition (3.199) means that

$$B = \begin{pmatrix} B_{11} & B_{12} & B_{13} \\ 0 & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{pmatrix},\ Q_K = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ Q_L = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.200}$$

So, now we have a 'little hole'. Under the conditions (3.199) consider

$$P = BQ_KB^{-1}(\mathrm{Id} - Q_L). \tag{3.201}$$

The condition $Q_LBQ_K = 0$ and the definition show that $Q_LP = 0 = PQ_L$. Moreover,

$$P^2 = BQ_KB^{-1}(\mathrm{Id} - Q_L)BQ_KB^{-1}(\mathrm{Id} - Q_L) = BQ_KB^{-1}BQ_KB^{-1}(\mathrm{Id} - Q_L) = P.$$

So, P is a projection which acts on the range of $\mathrm{Id} - Q_L$; from its definition, the range of P is contained in the range of BQ_K. Since

$$PBQ_K = BQ_KB^{-1}(\mathrm{Id} - Q_L)BQ_K = BQ_K$$

it follows that P is a projection *onto* the range of BQ_K.

If $A = Q_LAP$ is an isomorphism between the ranges of P and Q_L and $A' = PA'Q_L$ is its inverse, it is possible to rotate the range of P to that of Q_L

$$R(\theta) = \cos\theta P + \sin\theta A - \sin\theta A' + \cos\theta Q_L + (\mathrm{Id} - P - Q_L). \tag{3.202}$$

That this is a rotation can be seen directly

$$R(\theta)R(-\theta) = \mathrm{Id}\,. \tag{3.203}$$

Thus the homotopy $R(\theta)B$, $\theta \in [0, \pi/2]$, connects B to

$$B' = (\mathrm{Id} - P - Q_L)B + AB \tag{3.204}$$

since $A'B = 0$ and $(\mathrm{Id} - Q_L)B'Q_K = (\mathrm{Id} - P - Q_L)BQ_K + (\mathrm{Id} - Q_L)ABQ_k = 0$. Thus B' maps the range of Q_K to the range of Q_L and as such is an isomorphism,

$$Q_LB'Q_K = Q_LABQ_K = Q_LAPQ_K = (Q_LAP)(PBQ_K) = APQ_K. \tag{3.205}$$

Now, a similar, simpler, rotation can be made from the range of Q_L to the range of Q_K using any isomorphism, which can be chosen to be $G = (APQ_K)^{-1}$,

$$R'(\theta) = \cos\theta Q_L + \sin\theta G - \sin\theta APQ_K + \cos\theta Q_K + Q_O,\ R'(\theta'_R(-\theta) = \mathrm{Id}\,. \tag{3.206}$$

The homotopy $R'(\theta)B'$ connects B' to B'' which has $Q_KB''Q_K = Q_K$ so with respect to the 2×2 decomposition given by Q_K and $\mathrm{Id} - Q_K$,

$$B'' = \begin{pmatrix} \mathrm{Id} & E \\ 0 & F \end{pmatrix}. \tag{3.207}$$

The invertibility of this is equivalent to the invertibility of F and the homotopy

$$B''(s) = \begin{pmatrix} \mathrm{Id} & (1-s)E \\ 0 & F \end{pmatrix} \tag{3.208}$$

connects it to

$$L = \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & F \end{pmatrix},\ (B''(s))^{-1} = \begin{pmatrix} \mathrm{Id} & -(1-s)EF^{-1} \\ 0 & F^{-1} \end{pmatrix} \tag{3.209}$$

through invertibles.

The final step is 'Eilenberg's swindle'. Start from the form of L in (3.209), choose an isomorphism $\operatorname{Ran}(Q_K) = l^2(H) \oplus l^2(H)$ and then consider the successive rotations in terms of this 2×2 decomposition

$$L(\theta) = \begin{pmatrix} \cos\theta & \sin\theta F^{-1} \\ -\sin\theta F & \cos\theta \end{pmatrix},\ \theta \in [0, \pi/2], \tag{3.210}$$
$$L(\theta) = \begin{pmatrix} \cos\theta F^{-1} & \sin\theta F^{-1} \\ -\sin\theta F & \cos\theta F \end{pmatrix},\ \theta \in [\pi/2, \pi]$$

extended to be the constant isomorphism F on the extra factor. Then take the isomorphism

$$l^2(H) \oplus l^2(H) \oplus H \longrightarrow L^2(H) \oplus l^2(H),\ (\{u_i\}, \{w_i\}, v) \longmapsto (\{u_i\}, \{v, w_i\}) \tag{3.211}$$

in which the last element of H is place at the beginning of the second sequence. Now the rotations in (3.210) act on this space and $L(\pi - \theta)$ gives a homotopy connecting $\tilde{B}$ to the identity.

□