CHAPTER 3

Hilbert spaces

There are really three 'types' of Hilbert spaces (over $\mathbb{C}$). The finite dimensional ones, essentially just $\mathbb{C}^n$, for different integer values of n, with which you are pretty familiar, and two infinite dimensional types 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 (no doubt you realize that the $\mathbb{C}^n$ are separable) and that is what we are mostly interested in. Nevertheless we try to state results in general and then give proofs (usually they are the nicest ones) which work in the non-separable cases too.

I will first discuss the definition of pre-Hilbert and Hilbert spaces and prove Cauchy's inequality and the parallelogram law. This material can be found in many other places, so the discussion here will be kept succinct. One nice source is the book of G.F. Simmons, "Introduction to topology and modern analysis" [**5**]. I like it – but I think it is long out of print.

RBM:Add description of contents when complete and mention problems

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

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

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

$$\langle v, v\rangle \geq 0,\ \langle v, v\rangle = 0 \Longrightarrow v = 0. \tag{3.2}$$

Note that the reality of $\langle v, v\rangle$ follows from the 'Hermitian symmetry' 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

$$\langle v, \lambda_1 w_1 + \lambda_2 w_2\rangle = \overline{\lambda_1}\langle v, w_1\rangle + \overline{\lambda_2}\langle v, w_2\rangle \tag{3.3}$$

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

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

Lemma 3.1. *If H is a pre-Hilbert space with Hermitian inner product $\langle,\rangle$ then*

$$\|u\| = \langle u, u\rangle^{\frac{1}{2}} \tag{3.5}$$

is a norm on H.

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

$$\|\lambda u\|^2 = \langle \lambda u, \lambda u\rangle = |\lambda|^2\|u\|^2. \tag{3.6}$$

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

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

□

LEMMA 3.2. *The Cauchy-Schwarz inequality,*

$$|\langle u, v\rangle| \le \|u\|\|v\| \ \forall\ u, v \in H \tag{3.8}$$

holds in any pre-Hilbert space.

PROOF. This inequality is trivial if either u or v vanishes. For any non-zero u, $v \in H$ and $s \in \mathbb{R}$ positivity of the norm shows that

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

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

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

Substituting this into (3.9) gives

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

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

COROLLARY 3.1. *The inner product is continuous on the metric space (i.e. with respect to the norm) $H \times H$.*

PROOF. Corollaries really aren't supposed to require proof! If $(u_j, v_j) \to (u, v)$ then, by definition $\|u - u_j\| \to 0$ and $\|v - v_j\| \to 0$ so from

$$\begin{aligned}|\langle u, v\rangle - \langle u_j, v_j\rangle| &\le |\langle u, v\rangle - \langle u, v_j\rangle| + |\langle u, v_j\rangle - \langle u_j, v_j\rangle| \\ &\le \|u\|\|v - v_j\| + \|u - u_j\|\|v_j\|\end{aligned} \tag{3.12}$$

continuity follows. □

COROLLARY 3.2. *The Cauchy-Scwharz inequality is optimal in the sense that*

$$\|u\| = \sup_{v \in H; \|v\| \le 1} |\langle u, v\rangle|. \tag{3.13}$$

I really will leave this one to you.

2. Hilbert spaces

Definition 3.1. 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

$$\langle z, z' \rangle = \sum_{j=1}^{n} z_j \overline{z'_j} \tag{3.14}$$

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 (3.14)

$$\langle a, b \rangle = \sum_{j=1}^{\infty} a_j \overline{b_j},\ a, b \in l^2. \tag{3.15}$$

Completeness was shown earlier.

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

$$L^2(\mathbb{R}) = \{u \in \mathcal{L}^1_{\text{loc}}(\mathbb{R}); |u|^2 \in \mathcal{L}^1(\mathbb{R})\}/\mathcal{N} \tag{3.16}$$

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

$$\langle u, v \rangle = \int u\overline{v}. \tag{3.17}$$

Note that we showed that if $u,\ v \in \mathcal{L}^2(\mathbb{R})$ then $uv \in \mathcal{L}^1(\mathbb{R})$. We also showed that if $\int |u|^2 = 0$ then $u = 0$ almost everywhere, i.e. $u \in \mathcal{N}$, which is the definiteness of the inner product (3.17). It is fair to say that we went to some trouble to prove the completeness of this norm, so $L^2(\mathbb{R})$ is indeed a Hilbert space.

3. Orthonormal sequences

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

$$\langle u, v \rangle = 0 \text{ which can be written } u \perp v. \tag{3.18}$$

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

Proposition 3.1 (Bessel's inequality). *If e_i, $i \in \mathbb{N}$, is an orthonormal sequence in a pre-Hilbert space H, then*

$$\sum_i |\langle u, e_i \rangle|^2 \leq \|u\|^2\ \forall\ u \in H. \tag{3.19}$$

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

$$v = \sum_{i=1}^{N} \langle u, e_i \rangle e_i. \tag{3.20}$$

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

$$\langle v, e_j \rangle = \sum_{i=1}^{N} \langle u, e_i \rangle \langle e_i, e_j \rangle = \langle u, e_j \rangle \tag{3.21}$$

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

$$0 = \langle u - v, v\rangle = \langle u, v\rangle - \|v\|^2. \tag{3.22}$$

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

$$\|v\|^2 = \sum_{i=1}^{N} |\langle u, e_i\rangle|^2 \le \|u\|^2$$

which is (3.19).

In case the sequence is infinite this argument applies to any finite subsequence, e_i, $i = 1, \dots, N$ since it just uses orthonormality, so (3.19) follows by taking the supremum over N. $\square$

4. Gram-Schmidt procedure

DEFINITION 3.2. An orthonormal sequence, $\{e_i\}$, (finite or infinite) in a pre-Hilbert space is said to be *maximal* if

$$u \in H,\ \langle u, e_i\rangle = 0\ \forall\ i \Longrightarrow u = 0. \tag{3.23}$$

THEOREM 3.1. *Every separable pre-Hilbert space contains a maximal orthonormal sequence.*

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. First if $v_1 \neq 0$ set $e_i = v_1/\|v_1\|$. Proceeding by induction we can suppose we have found, for a given integer n, elements e_i, $i = 1, \dots, m$, where $m \le n$, which are orthonormal and such that the linear span

$$\operatorname{sp}(e_1, \dots, e_m) = \operatorname{sp}(v_1, \dots, v_n). \tag{3.24}$$

We certainly have this for $n = 1$. To show the inductive step observe that if v_{n+1} is in the span(s) in (3.24) 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 (3.24). It follows that

$$w = v_{n+1} - \sum_{j=1}^{n} \langle v_{n+1}, e_j\rangle e_j \neq 0 \text{ so } e_{m+1} = \frac{w}{\|w\|} \tag{3.25}$$

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

$$H \ni u \perp e_j\ \forall\ j \Longrightarrow u = 0. \tag{3.26}$$

This uses the density of the v_j's. There must exist a sequence w_k where each w_k is a v_j, such that $w_k \to u$ in H, assumed to satisfy (3.26). Now, each v_j, and hence each w_k, is a finite linear combination of e_l's so, by Bessel's inequality

$$\|w_k\|^2 = \sum_l |\langle w_k, e_l\rangle|^2 = \sum_l |\langle u - w_k, e_l\rangle|^2 \le \|u - w_k\|^2 \tag{3.27}$$

where $\langle u, e_l\rangle = 0$ for all l has been used. Thus $\|w_k\| \to 0$ and hence $u = 0$. $\square$

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

5. Orthonormal bases

Definition 3.3. In view of the following result, a maximal orthonormal sequence in a separable Hilbert space will be called an orthonormal basis; it is often called a 'complete orthonormal basis' but the 'complete' is really redundant.

This notion of basis is not quite the same as in the finite dimensional case (although it is a legitimate extension of it). There are other, quite different, notions of a basis in infinite dimensions. See for instance 'Hamel basis' which arises in some settings – it is discussed briefly in §1.12 and can be used to show the existence of a non-continuous functional on a Banach space.

Theorem 3.2. *If $\{e_i\}$ is an orthonormal basis (a maximal orthonormal sequence) in a Hilbert space then for any element $u \in H$ the 'Fourier-Bessel series' converges to u :*

$$u = \sum_{i=1}^{\infty} \langle u, e_i\rangle e_i. \tag{3.28}$$

In particular a Hilbert space with an orthonormal basis is separable!

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

$$u_N = \sum_{i=1}^{N} \langle u, e_i\rangle e_i \tag{3.29}$$

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

$$\|u_{m'} - u_m\|^2 = \sum_{i=m+1}^{m'} |\langle u, e_i\rangle|^2 \le \sum_{i>m} |\langle u, e_i\rangle|^2 \tag{3.30}$$

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

$$\langle w, e_i\rangle = \lim_{m\to\infty} \langle u_m, e_i\rangle = \langle u, e_i\rangle \tag{3.31}$$

for each i. Thus $u - w$ is orthogonal to all the e_i so by the assumed completeness of the orthonormal basis must vanish. Thus indeed (3.28) holds. $\square$

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

$$T : H \longrightarrow l^2 \tag{3.32}$$

which is 1-1, onto and satisfies $\langle Tu, Tv\rangle_{l^2} = \langle u, v\rangle_H$ 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

$$Tu = \{\langle u, e_j \rangle\}_{j=1}^{\infty}. \tag{3.33}$$

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 $\langle u, e_j \rangle = 0$ 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

$$u = \sum_{j=1}^{\infty} c_j e_j \tag{3.34}$$

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

$$\|\sum_{j=m+1}^{n} c_j e_j\|_H^2 = \sum_{j=m+1}^{n} |c_j|^2. \tag{3.35}$$

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

$$\|v + w\|^2 + \|v - w\|^2 = 2\|v\|^2 + 2\|w\|^2, \ \forall \ v, w \in H. \tag{3.36}$$

PROOF. Just expand out using the inner product

$$\|v + w\|^2 = \|v\|^2 + \langle v, w \rangle + \langle w, v \rangle + \|w\|^2 \tag{3.37}$$

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

PROPOSITION 3.4. *Any normed space where the norm satisfies the parallelogram law,* (3.36)*, is a pre-Hilbert space in the sense that*

$$\langle v, w \rangle = \frac{1}{4}\left(\|v + w\|^2 - \|v - w\|^2 + i\|v + iw\|^2 - i\|v - iw\|^2\right) \tag{3.38}$$

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 3.5. *If $C \subset H$ is a subset of a Hilbert space which is*
(1) *Non-empty*
(2) *Closed*
(3) *Convex, in the sense that $v_1,\ v_2 \in C$ 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

$$\|v\| = \inf_{u \in C} \|u\|. \tag{3.39}$$

PROOF. By definition of the infimum of a non-empty set of real numbers which is bounded below (in this case by 0) there must exist a sequence $\{v_n\}$ in C such that $\|v_n\| \to d = \inf_{u \in C} \|u\|$. We show that v_n converges and that the limit is the point we want. The parallelogram law can be written

$$\|v_n - v_m\|^2 = 2\|v_n\|^2 + 2\|v_m\|^2 - 4\|(v_n + v_m)/2\|^2. \tag{3.40}$$

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 $\|(v_n + v_m)/2\|^2 \geq d^2$. Combining these estimates gives

$$n, m > N \Longrightarrow \|v_n - v_m\|^2 \leq 4d^2 + \epsilon^2 - 4d^2 = \epsilon^2 \tag{3.41}$$

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 $\|v\| = \lim_{n \to \infty} \|v_n\| = d$.

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

$$\|v - v'\|^2 = 2\|v\|^2 + 2\|v'\|^2 - 4\|(v + v')/2\|^2 \leq 0 \Longrightarrow v = v'. \tag{3.42}$$

Alternatively you can just observe that we have actually shown above that *any* sequence in C such that $\|v_n\| \to d$ converges, so this is true for the alternating sequence of v and v'. □

9. Orthocomplements and projections

PROPOSITION 3.6. *If $W \subset H$ is a linear subspace of a Hilbert space then*

$$W^\perp = \{u \in H; \langle u, w \rangle = 0\ \forall\ w \in W\} \tag{3.43}$$

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

$$H = W \oplus W^\perp \tag{3.44}$$

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 (3.43) 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 $\langle u, u \rangle = \|u\|^2 = 0$ and $u = 0$; thus $W \cap W^\perp = \{0\}$. Since the map $H \ni u \longrightarrow \langle u, w \rangle \in \mathbb{C}$ is continuous for each $w \in H$ its null space, the inverse image of 0, is closed. Thus

$$W^\perp = \bigcap_{w \in W} \{u \in H; \langle u, w \rangle = 0\} \tag{3.45}$$

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

$$C = u + W = \{u' \in H; u' = u + w,\ w \in W\}. \tag{3.46}$$

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 orthogonal 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

$$\|v + \lambda w\|^2 = \|v\|^2 + 2\operatorname{Re}(\lambda\langle v, w\rangle) + |\lambda|^2\|w\|^2. \tag{3.47}$$

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

$$\begin{gathered}\|v\|^2 + 2t|\langle v, w\rangle)| + t^2\|w\|^2 \geq \|v\|^2 \Longrightarrow \\ t(2|\langle v, w\rangle| + t\|w\|^2) \geq 0\ \forall\ t \in \mathbb{R} \Longrightarrow |\langle v, w\rangle| = 0\end{gathered} \tag{3.48}$$

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 (3.44) 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

$$\Pi_W : H \longrightarrow W. \tag{3.49}$$

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

$$\lambda_1 u_1 + \lambda_2 u_2 = (\lambda_1 v_1 + \lambda_2 v_2) + (\lambda_1 w_1 + \lambda_2 w_2) \tag{3.50}$$

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

$$\Pi_W^2 = \Pi_W,\ \|\Pi_W u\| \leq \|u\| \Longrightarrow \|\Pi_W\| \leq 1. \tag{3.51}$$

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).

Finite-dimensional subspaces are always closed by the Heine-Borel theorem.

Lemma 3.3. *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*

$$Pu = \sum \langle u, e_k\rangle e_k. \tag{3.52}$$

Proof. We know that the series in (3.52) 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. $\square$

LEMMA 3.4. *If $W \subset H$ is a linear subspace of a Hilbert space which contains the orthocomplement of a finite dimensional space then W is closed and $W^\perp$ is finite-dimensional.*

PROOF. If $U \subset W$ is a closed subspace with finite-dimensional orthocomplement then each of the N elements, v_i, of a basis of $(\mathrm{Id} - \Pi_U)W$ is the image of some $w_i \in W$. Since U is the null space of $\mathrm{Id} - \Pi_U$ it follows that any element of W can be written uniquely in the form

$$w = u + \sum_{i=1}^{N} c_i v_i,\ u = \Pi_U w \in U,\ c_i = \langle w, v_i \rangle. \tag{3.53}$$

Then if ϕ_n is a sequence in W which converges in H it follows that $\Pi_U \phi_n$ converges in U and $\langle \phi_n, v_i \rangle$ converges and hence the limit is in W. $\square$

Note that the existence of a non-continuous linear functional $H \longrightarrow \mathbb{C}$ is equivalent to the existence of a non-closed subspace of H with a one-dimensional complement. Namely the null space of a non-continuous linear functional cannot be closed, since from this continuity follows, but it does have a one-dimensional complement (not orthocomplement!)

QUESTION 1. Does there exist a non-continuous linear functional on an infinite-dimensional Hilbert space? [1]

10. Riesz' theorem

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

THEOREM 3.3. *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*

$$T(u) = \langle u, \phi \rangle\ \forall\ u \in H. \tag{3.54}$$

PROOF. If T is the zero functional then $\phi = 0$ gives (3.54). Otherwise there exists some $u' \in H$ such that $T(u') \neq 0$ and then there is some $u \in H$, namely $u = u'/T(u')$ will work, such that $T(u) = 1$. Thus

$$C = \{u \in H; T(u) = 1\} = T^{-1}(\{1\}) \neq \emptyset. \tag{3.55}$$

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

$$T((u + u')/2) = (T(u) + T(u'))/2. \tag{3.56}$$

Thus, by Proposition 3.5, 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 3.6 above, v is orthogonal to N. In this case it is the unique element orthogonal to N with $T(v) = 1$.

---

[1] The existence of such a functional requires some form of the Axiom of Choice (maybe a little weaker in the separable case). You are free to believe that all linear functionals are continuous but you will make your life difficult this way.

Now, for any $u \in H$,

$$u - T(u)v \text{ satisfies } T(u - T(u)v) = T(u) - T(u)T(v) = 0 \Longrightarrow \\ u = w + T(u)v,\ w \in N. \tag{3.57}$$

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

$$u = w + \langle u, \phi\rangle v \Longrightarrow T(u) = \langle u, \phi\rangle T(v) = \langle u, \phi\rangle. \tag{3.58}$$

□

11. Adjoints of bounded operators

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

$$A : H \longrightarrow H,\ \|Au\| \le C\|u\|\ \forall\ u \in H \tag{3.59}$$

there corresponds a unique adjoint operator. This has profound consequences for the theory of operators on a Hilbert space, as we shall see.

Proposition 3.7. *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*

$$\langle Au, v\rangle_H = \langle u, A^*v\rangle_H\ \forall\ u, v \in H \text{ and } \|A\| = \|A^*\|. \tag{3.60}$$

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 (3.60), which is to say consider

$$H \ni u \longrightarrow \langle Au, v\rangle \in \mathbb{C}. \tag{3.61}$$

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

$$|\langle Au, v\rangle| \le \|Au\|\|v\| \le (\|A\|\|v\|)\|u\|. \tag{3.62}$$

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

$$H \xrightarrow{u \longmapsto Au} H \xrightarrow{w \longmapsto \langle w, v\rangle} \mathbb{C}. \tag{3.63}$$

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

$$\langle Au, v\rangle = \langle u, A^*v\rangle\ \forall\ u \in H. \tag{3.64}$$

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

$$\langle Au, c_1v_1 + c_2v_2\rangle = \overline{c_1}\langle Au, v_1\rangle + \overline{c_2}\langle Au, v_2\rangle \\ = \overline{c_1}\langle u, A^*v_1\rangle + \overline{c_2}\langle u, A^*v_2\rangle = \langle u, c_1A^*v_2 + c_2A^*v_2\rangle \tag{3.65}$$

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

Using the optimality of Cauchy's inequality

$$\|A^*v\| = \sup_{\|u\|=1} |\langle u, A^*v\rangle| = \sup_{\|u\|=1} |\langle Au, v\rangle| \le \|A\|\|v\|. \tag{3.66}$$

This shows that A^* is bounded and that

$$\|A^*\| \le \|A\|. \tag{3.67}$$

The defining identity (3.60) also shows that $(A^*)^* = A$ so the reverse equality in (3.67) also holds and therefore

$$\|A^*\| = \|A\|. \tag{3.68}$$

□

One useful property of the adjoint operator is that

$$\mathrm{Nul}(A^*) = (\mathrm{Ran}(A))^\perp. \tag{3.69}$$

Indeed $w \in (\mathrm{Ran}(A))^\perp$ means precisely that $\langle w, Av\rangle = 0$ for all $v \in \mathcal{H}$ which translates to

$$w \in (\mathrm{Ran}(A))^\perp \Longleftrightarrow \langle A^*w, v\rangle = 0 \Longleftrightarrow A^*w = 0. \tag{3.70}$$

Note that in the finite dimensional case (3.69) is equivalent to $\mathrm{Ran}(A) = (\mathrm{Nul}(A^*))^\perp$ but in the infinite dimensional case $\mathrm{Ran}(A)$ is often not closed in which case this cannot be true and you can only be sure that

$$\overline{\mathrm{Ran}(A)} = (\mathrm{Nul}(A^*))^\perp. \tag{3.71}$$

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 Hilbert space the notion of a compact set is already fixed. We want to characterize it, actually in two closely related ways.

In any metric space a 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) 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 3.5. *In any metric space the set, S, consisting of the points of a convergent sequence, together with its limit, is compact.*

Proof. We show that S is compact by checking that 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 3.6. *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*

$$\sum_{k>N} |\langle u_n, e_k\rangle|^2 < \epsilon^2 \ \forall\ n. \tag{3.72}$$

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

$$\sum_{k} |\langle u, e_k\rangle|^2 \le \|u\|^2. \tag{3.73}$$

The convergence of this series means that (3.72) 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

$$\sum_{k>N'} |\langle u, e_k\rangle|^2 < \epsilon^2/2. \tag{3.74}$$

Consider the closure of the subspace spanned by the e_k with $k > N$. The orthogonal projection onto this space (see Lemma 3.3) is

$$P_N u = \sum_{k>N} \langle u, e_k\rangle e_k. \tag{3.75}$$

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

$$\|P_N u_n\|^2 = \sum_{k>N} |\langle u_n, e_k\rangle|^2 < \epsilon^2,\ n > n'. \tag{3.76}$$

So, we have arranged (3.72) 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 choosing N large enough. Thus indeed (3.72) holds for all n if N is chosen large enough. □

This suggests one useful characterization of compact sets in a separable Hilbert space since the equi-smallness of the tails, as in (3.72), for all $u \in K$ just means that the Fourier-Bessel series converges uniformly.

Proposition 3.8. *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

$$\sum_{k>p} |\langle u_p, e_k\rangle|^2 \ge \epsilon^2. \tag{3.77}$$

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 3.6, 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 $\langle u_n, e_k\rangle$ in $\mathbb{C}$. Here k is fixed. This sequence is bounded:

$$|\langle u_n, e_k\rangle| \le \|u_n\| \le C \tag{3.78}$$

by the boundedness of K. So, by the Heine-Borel theorem, there is a subsequence u_{n_l} such that $\langle u_{n_l}, e_k\rangle$ 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 $\langle u_{n,1}, e_1\rangle$ converges. Then extract a subsequence $u_{n,2}$ of $u_{n,1}$ so that $\langle u_{n,2}, e_2\rangle$ *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 $\langle v_n, e_k\rangle$ converges.

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

$$\begin{aligned}\|v_n - v_{n+l}\|^2 &= \sum_{k\le N} |\langle v_n - v_{n+l}, e_k\rangle|^2 + \sum_{k>N} |\langle v_n - v_{n+l}, e_k\rangle|^2 \\ &\le \sum_{k\le N} |\langle v_n - v_{n+l}, e_k\rangle|^2 + 2\sum_{k>N} |\langle v_n, e_k\rangle|^2 + 2\sum_{k>N} |\langle v_{n+l}, e_k\rangle|^2\end{aligned} \tag{3.79}$$

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 $\langle v_n, e_k\rangle$. Thus, $\{v_n\}$ is a Cauchy subsequence of $\{u_n\}$ and hence as already noted convergent in K. Thus K is indeed compact. □

This criterion for compactness is useful but is too closely tied to the existence of an orthonormal basis to be easily applicable. However the condition can be restated in a way that holds even in the non-separable case (and of course in the finite-dimensional case, where it is trivial).

PROPOSITION 3.9. *A subset $K \subset H$ of a Hilbert space is compact if and only if it is closed and bounded and for every $\epsilon > 0$ there is a finite-dimensional subspace $W \subset H$ such that*

$$\sup_{u\in K} \inf_{w\in W} \|u - w\| < \epsilon. \tag{3.80}$$

So we see that the extra condition needed is 'finite-dimensional approximability'.

PROOF. Before proceeding to the proof consider (3.80). Since W is finite-dimensional we know it is closed and hence the discussion in §9 above applies. In particular $u = w + w^\perp$ with $w \in W$ and $w^\perp \perp W$ where

$$\inf_{w \in W} \|u - w\| = \|w^\perp\|. \tag{3.81}$$

This can be restated in the form

$$\sup_{u \in K} \|(\mathrm{Id} - \Pi_W)u\| < \epsilon \tag{3.82}$$

where Π_W is the orthogonal projection onto W (so $\mathrm{Id} - \Pi_W$ is the orthogonal projection onto $W^\perp$).

Now, let us first assume that H is separable, so we already have a condition for compactness in Proposition 3.8. Then if K is compact we can consider an orthonormal basis of H and the finite-dimensional spaces W_N spanned by the first N elements in the basis with Π_N the orthogonal projection onto it. Then $\|(\mathrm{Id} - \Pi_N)u\|$ is precisely the length of the 'tail' of u with respect to the basis. So indeed, by Proposition 3.8, given $\epsilon > 0$ there is an N such that $\|(\mathrm{Id} - \Pi_N)u\| < \epsilon/2$ for all $u \in K$ and hence (3.82) holds for $W = W_N$.

Now suppose that $K \subset H$ and for each $\epsilon > 0$ we can find a finite dimensional subspace W such that (3.82) holds. Take a sequence $\{u_n\}$ in K. The sequence $\Pi_W u_n \in W$ is bounded in a finite-dimensional space so has a convergent subsequence. Now, for each $j \in \mathbb{N}$ there is a finite-dimensional subspace W_j (not necessarily corresponding to an orthonormal basis) so that (3.82) holds for $\epsilon = 1/j$. Proceeding as above, we can find successive subsequence of u_n such that the image under Π_j in W_j converges for each j. Passing to the diagonal subsequence u_{n_l} it follows that $\Pi_j u_{n_i}$ converges for each j since it is eventually a subsequence of the jth choice of subsequence above. Now, the triangle inequality shows that

$$\|u_{n_i} - u_{n_k}\| \le \|\Pi_j(u_{n_i} - u_{n_k})\|_{W_j} + \|(\mathrm{Id} - \Pi_j)u_{n_i}\| + \|(\mathrm{Id} - \Pi_j)u_{n_k}\|. \tag{3.83}$$

Given $\epsilon > 0$ first choose j so large that the last two terms are each less than $1/j < \epsilon/3$ using the choice of W_j. Then if $i, k > N$ is large enough the first term on the right in (3.83) is also less than $\epsilon/3$ by the convergence of $\Pi_j u_{n_i}$. Thus u_{n_i} is Cauchy in H and hence converges and it follows that K is compact.

This converse argument does not require the separability of H so to complete the proof we only need to show the necessity of (3.81) in the non-separable case. Thus suppose K is compact. Then K itself is separable – has a countable dense subset – using the finite covering property (for each $p > 0$ there are finitely many balls of radius $1/p$ which cover K so take the set consisting of all the centers for all p). It follows that the closure of the span of K, the finite linear combinations of elements of K, is a separable Hilbert subspace of H which contains K. Thus any compact subset of a non-separable Hilbert space is contained in a separable Hilbert subspace and hence (3.80) holds. □

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 3.4. 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):

$$(T_1 + T_2)u \in \mathrm{Ran}(T_1) + \mathrm{Ran}(T_2) \ \forall \ u \in \mathcal{H}. \tag{3.84}$$

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 3.7. *If $T : H \longrightarrow H$ has finite rank then there is a finite orthonormal set $\{e_k\}_{k=1}^L$ in H and constants $c_{ij} \in \mathbb{C}$ such that*

$$Tu = \sum_{i,j=1}^{L} c_{ij} \langle u, e_j \rangle e_i. \tag{3.85}$$

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 $u \in H$:

$$Tu = \sum_{i=1}^{p} \langle Tu, e_i \rangle e_i. \tag{3.86}$$

On the other hand, the map $u \longrightarrow \langle Tu, e_i \rangle$ is a continuous linear functional on H, so $\langle Tu, e_i \rangle = \langle u, v_i \rangle$ for some $v_i \in H$; notice in fact that $v_i = T^* e_i$. This means the formula (3.86) becomes

$$Tu = \sum_{i=1}^{p} \langle u, v_i \rangle e_i. \tag{3.87}$$

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

$$v_i = \sum_{j=1}^{L} \overline{c_{ij}} e_j. \tag{3.88}$$

Inserting this into (3.87) gives (3.85) (where the constants for $i > p$ are zero). □

It is clear that

$$B \in \mathcal{B}(\mathcal{H}) \text{ and } T \in \mathcal{R}(\mathcal{H}) \text{ then } BT \in \mathcal{R}(\mathcal{H}). \tag{3.89}$$

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 $\mathrm{Ran}(T)$. 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 3.10. *The finite rank operators form a $*$-closed two-sided ideal in $\mathcal{B}(\mathcal{H})$, which is to say a linear subspace such that*

$$B_1,\ B_2 \in \mathcal{B}(\mathcal{H}),\ T \in \mathcal{R}(\mathcal{H}) \Longrightarrow B_1 T B_2,\ T^* \in \mathcal{R}(\mathcal{H}). \tag{3.90}$$

PROOF. It is only left to show that T^* is of finite rank if T is, but this is an immediate consequence of Lemma 3.7 since if T is given by (3.85) then

$$T^*u = \sum_{i,j=1}^{N} \overline{c_{ij}} \langle u, e_i \rangle e_j \tag{3.91}$$

is also of finite rank. □

LEMMA 3.8 (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 (3.87) 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

$$\begin{gathered} \langle Tu, w \rangle = \sum_{i=1}^{p} \langle u, v_i \rangle \langle e_i, w \rangle \Longrightarrow \\ \langle w, Tu \rangle = \sum_{i=1}^{p} \langle v_i, u \rangle \langle w, e_i \rangle \Longrightarrow \\ T^*w = \sum_{i=1}^{p} \langle w, e_i \rangle v_i \end{gathered} \tag{3.92}$$

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

14. Compact operators

DEFINITION 3.5. 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; such a set is said to be *precompact*.

PROPOSITION 3.11. *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

$$\sum_{i>N} |\langle v, e_i \rangle|^2 < \epsilon^2 \ \forall \ v \in I. \tag{3.93}$$

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 (3.93) applies to it. Namely this means that for any $\epsilon > 0$ there exists n such that

$$\sum_{i>n} |\langle Ku, e_i\rangle|^2 < \epsilon^2 \ \forall \ u \in \mathcal{H}, \ \|u\|_{\mathcal{H}} \le 1. \tag{3.94}$$

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

$$K_n u = \sum_{k\le n} \langle Ku, e_i\rangle e_i. \tag{3.95}$$

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,

$$\|Ku - K_n u\|^2_{\mathcal{H}} = \sum_{i>n} |\langle Ku, e_i\rangle|^2 \tag{3.96}$$

Thus (3.94) shows that $\|Ku - K_n u\|_{\mathcal{H}} \le \epsilon$. Now, increasing n makes $\|Ku - K_n u\|$ smaller, so given $\epsilon > 0$ there exists n such that for all $N \ge n$,

$$\|K - K_N\|_{\mathcal{B}} = \sup_{\|u\|\le 1} \|Ku - K_n u\|_{\mathcal{H}} \le \epsilon. \tag{3.97}$$

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. Let $W_n = T_n\mathcal{H}$ be the range of T_n. By definition it is a finite dimensional subspace and hence closed. Let Π_n be the orthogonal projection onto W_n, so $\operatorname{Id} - \Pi_n$ is projection onto $W_n^\perp$. Thus the composite $(\operatorname{Id} - \Pi_n)T_n = 0$ and hence

$$(\operatorname{Id} - \Pi_n)K = (\operatorname{Id} - \Pi_n)(K - T_n) \Longrightarrow \|(\operatorname{Id} - \Pi_n)K\| \to 0 \text{ as } n \to \infty. \tag{3.98}$$

So, for any $\epsilon > 0$ there exists n such that

$$\sup_{u\in B(0,1)} \inf_{w\in W_n} \|Ku - w\| \le \sup_{\|u\|\le 1} \|(\operatorname{Id} - \Pi_n)Ku\| < \epsilon \tag{3.99}$$

and it follows from Proposition 3.9 that $K(B(0, 1))$ is precompact and hence K is compact. $\square$

Proposition 3.12. *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

$$BT(B(0, 1)) \subset B(\overline{T(B(0, 1))}) \text{ is compact} \tag{3.100}$$

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 $\langle u_n, e_k\rangle$, the sequence of coefficients of some particular Fourier-Bessel series, should converge.

DEFINITION 3.6. 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 $\langle u_j, v\rangle \to \langle u, v\rangle$ converges in $\mathbb{C}$ for each $v \in \mathcal{H}$. This relationship is written

$$u_n \rightharpoonup u. \tag{3.101}$$

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 $\langle u_n, v\rangle$ 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 $\langle u - u', v\rangle = \lim_{n\to\infty}\langle u_n, v\rangle - \lim_{n\to\infty}\langle u_n, v\rangle = 0$ for all $v \in \mathcal{H}$ and setting $v = u - u'$ it follows that $u = u'$.

LEMMA 3.9. *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

$$|\langle u_n, v\rangle - \langle u, v\rangle| \le \|u_n - u\|\|v\| \to 0 \tag{3.102}$$

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

LEMMA 3.10. *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 $\langle u_n, e_k\rangle \to \langle u, e_k\rangle$ converges for each k. Conversely, suppose this is true for a bounded sequence, just that $\langle u_n, e_k\rangle \to c_k$ in $\mathbb{C}$ for each k. The norm boundedness and Bessel's inequality show that

$$\sum_{k\le p} |c_k|^2 = \lim_{n\to\infty} \sum_{k\le p} |\langle u_n, e_k\rangle|^2 \le \sup_n \|u_n\|^2 \tag{3.103}$$

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

$$u = \sum_k c_k e_k \in \mathcal{H} \tag{3.104}$$

by the completeness of $\mathcal{H}$. Clearly $\langle u_n, e_k\rangle \to \langle u, e_k\rangle$ for each k. It remains to show that $\langle u_n, v\rangle \to \langle u, v\rangle$ 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

$$\begin{aligned}\langle u_n, v\rangle - \langle u, v\rangle = \langle u_n, v_p\rangle - \langle u, v_p\rangle + \langle u_n, v - v_p\rangle - \langle u, v - v_p\rangle \Longrightarrow \\ |\langle u_n, v\rangle - \langle u, v\rangle| \le |\langle u_n, v_p\rangle - \langle u, v_p\rangle| + 2C\|v - v_p\|\end{aligned} \tag{3.105}$$

where $v_p = \sum\limits_{k\le p} \langle v, e_k\rangle e_k$ 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 (3.105) 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, $\langle u_n, v\rangle \to \langle u, v\rangle$ for all $v \in \mathcal{H}$ and it follows that u_n converges weakly to u. □

PROPOSITION 3.13. *Any bounded sequence $\{u_n\}$ in a separable Hilbert space has a weakly convergent subsequence.*

This can be thought of as different extension to infinite dimensions of the Heine-Borel theorem. As opposed to the characterization of compact sets above, which involves adding the extra condition of finite-dimensional approximability, here we weaken the notion of convergence.

PROOF. Choose an orthonormal basis $\{e_k\}$ and apply the procedure in the proof of Proposition 3.8 to it. Thus, we may extract successive subsequence along the kth of which $\langle u_{n_p}, e_k\rangle \to c_k \in \mathbb{C}$. Passing to the diagonal subsequence, v_n, which is *eventually* a subsequence of each of these ensures that $\langle v_n, e_k\rangle \to c_k$ for each k. Now apply the preceeding Lemma to conclude that this subsequence converges weakly. □

LEMMA 3.11. *For a weakly convergent sequence $u_n \rightharpoonup u$*

$$\|u\| \le \liminf \|u_n\| \tag{3.106}$$

and a weakly convergent sequence converges strongly if and only if the weak limit satisfies $\|u\| = \lim_{n\to\infty} \|u_n\|$.

PROOF. Choose an orthonormal basis e_k and observe that

$$\sum_{k\le p} |\langle u, e_k\rangle|^2 = \lim_{n\to\infty} \sum_{k\le p} |\langle u_n, e_k\rangle|^2. \tag{3.107}$$

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

$$\sum_{k\le p} |\langle u, e_k\rangle|^2 \le \liminf_n \|u_n\|^2 \tag{3.108}$$

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

$$\|u\|^2 \le \liminf_n \|u_n\|^2 \tag{3.109}$$

from which (3.106) follows.

Now, suppose $u_n \rightharpoonup u$ then

$$\|u - u_n\|^2 = \|u\|^2 - 2\operatorname{Re}\langle u, u_n\rangle + \|u_n\|^2. \tag{3.110}$$

Weak convergence implies that the middle term converges to $-2\|u\|^2$ so if the last term converges to $\|u\|^2$ then $u \to u_n$. □

Observe that for any $A \in \mathcal{B}(H)$, if $u_n \rightharpoonup u$ then $Au_n \rightharpoonup Au$ using the existence of the adjoint:-

$$\langle Au_n, v\rangle = \langle u_n, A^*v\rangle \to \langle u, A^*v\rangle = \langle Au, v\rangle \ \forall\ v \in \mathcal{H}. \tag{3.111}$$

LEMMA 3.12. *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 subsequence and the limit is necessarily Ku since $Ku_n \rightharpoonup Ku$. 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 $d(v_{n_k}, v) \geq \epsilon$. 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 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 weakly convergent sequence in a Hilbert space is bounded, provided one has the Uniform Boundedness Principle, Theorem 1.3, at the ready.

Proposition 3.14. *If $u_n \in H$ is a sequence in a Hilbert space and for all $v \in H$*

$$\langle u_n, v\rangle \to F(v) \text{ converges in } \mathbb{C} \tag{3.112}$$

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

Proof. Apply the Uniform Boundedness Theorem to the continuous functionals

$$T_n(u) = \langle u, u_n\rangle,\ T_n : H \longrightarrow \mathbb{C} \tag{3.113}$$

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

$$\|T_n\| \leq C. \tag{3.114}$$

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 :

$$T(u) = \lim_{n\to\infty} T_n(u) = \lim_{n\to\infty} \langle u, u_n\rangle. \tag{3.115}$$

This exists for each u by hypothesis. It is a linear map and from (3.114) it is bounded, $\|T\| \leq C$. Thus by the Riesz Representation theorem, there exists $w \in H$ such that

$$T(u) = \langle u, w\rangle\ \forall\ u \in H. \tag{3.116}$$

Thus $\langle u_n, u\rangle \to \langle w, u\rangle$ 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

$$\|A\| = \sup_{\|u\|_{\mathcal{H}}=1} \|Au\|_{\mathcal{H}} \tag{3.117}$$

and that the norm satisfies

$$\|AB\| \le \|A\|\|B\| \tag{3.118}$$

as follows from the fact that

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

Consider the set of invertible elements:

$$\mathrm{GL}(\mathcal{H}) = \{A \in \mathcal{B}(\mathcal{H}); \exists\ B \in \mathcal{B}(\mathcal{H}),\ BA = AB = \mathrm{Id}\}. \tag{3.119}$$

Note that this is equivalent to saying A is 1-1 and onto in view of the Open Mapping Theorem, Theorem 1.4.

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

Lemma 3.13. *If $A \in \mathcal{B}(\mathcal{H})$ and $\|A\| < 1$ then*

$$\mathrm{Id} - A \in \mathrm{GL}(\mathcal{H}). \tag{3.120}$$

Proof. This follows from the convergence of the Neumann series. If $\|A\| < 1$ then $\|A^j\| \le \|A\|^j$, from (3.118), and it follows that

$$B = \sum_{j=0}^{\infty} A^j \tag{3.121}$$

(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

$$AB = A \lim_{n\to\infty} \sum_{j=0}^{n} A^j = \lim_{n\to\infty} \sum_{j=1}^{n+1} A^j = B - \mathrm{Id} \tag{3.122}$$

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

Proposition 3.15. *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 $G^{-1} \in \mathcal{B}(\mathcal{H})$:

$$G^{-1}G = GG^{-1} = \mathrm{Id}. \tag{3.123}$$

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

$$E = -G^{-1}B \Longrightarrow G + B = G(\mathrm{Id} + G^{-1}B) = G(\mathrm{Id} - E) \tag{3.124}$$

From Lemma 3.13 we know that

$$\|B\| < 1/\|G^{-1}\| \Longrightarrow \|G^{-1}B\| < 1 \Longrightarrow \mathrm{Id} - E \text{ is invertible.} \tag{3.125}$$

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

$$(\mathrm{Id}-E)^{-1}G^{-1}(G+B) = (\mathrm{Id}-E)^{-1}(\mathrm{Id}-E) = \mathrm{Id}\,. \tag{3.126}$$

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

$$(G+B)G^{-1}(\mathrm{Id}-E')^{-1} = (\mathrm{Id}-E')(\mathrm{Id}-E')^{-1} = \mathrm{Id}\,. \tag{3.127}$$

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

$$G^{-1}(\mathrm{Id}-E')^{-1} = (\mathrm{Id}-E)^{-1}G^{-1}(G+B)G^{-1}(\mathrm{Id}-E')^{-1} = (\mathrm{Id}-E)^{-1}G^{-1} \tag{3.128}$$

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

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

$$\mathrm{U}(\mathcal{H}) = \{U \in \mathrm{GL}(\mathcal{H}); U^{-1} = U^*\}. \tag{3.129}$$

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

$$\langle Uu, Uv\rangle = \langle u, v\rangle,\ \|Uu\| = \|u\|\ \forall\ u, v \in \mathcal{H},\ U \in \mathrm{U}(\mathcal{H}). \tag{3.130}$$

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. One important fact that you should know, and there is a proof towards the end of this chapter, is that both $\mathrm{GL}(H)$ and $\mathrm{U}(\mathcal{H})$ are contractible as 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 $n \to \infty$. In fact there is another group which *is* essentially the large n limit of the $\mathrm{U}(n)$, namely

$$\mathrm{U}^{-\infty}(H) = \{\mathrm{Id} + K \in \mathrm{U}(H); K \in \mathcal{K}(H)\}. \tag{3.131}$$

It does have lots of interesting (and useful) topology.

Another important fact that we will discuss below is that $\mathrm{GL}(H)$ is *not* dense in $\mathcal{B}(H)$, in contrast to the finite dimensional case. In other words there are operators which are not invertible and cannot be made invertible by small perturbations.

17. Spectrum of an operator

Another direct application of Lemma 3.13, 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

$$(\lambda\,\mathrm{Id}-A)\lambda^{-1}(\mathrm{Id}-\lambda^{-1}A)^{-1} = \mathrm{Id} = \lambda^{-1}(\mathrm{Id}-\lambda^{-1}A)^{-1}(\lambda - A). \tag{3.132}$$

Thus, $\lambda\,\mathrm{Id}-A \in \mathrm{GL}(H)$, which we usually abbreviate to $\lambda - A$, has inverse $(\lambda - A)^{-1} = \lambda^{-1}(\mathrm{Id}-\lambda^{-1}A)^{-1}$. The set of λ for which this operator is invertible is called the *resolvent set* and we have shown

$$\begin{aligned}\operatorname{Res}(A) = \{\lambda \in \mathbb{C}; (\lambda \operatorname{Id} - A) \in \operatorname{GL}(H)\} \subset \mathbb{C}\\ \{|\lambda| > \|A\|\} \subset \operatorname{Res}(A).\end{aligned} \tag{3.133}$$

From the discussion above, it is an open, and non-empty, set on which $(A-\lambda)^{-1}$, called the resolvent of A, is defined. The complement of the resolvent set is called the spectrum of A

$$\operatorname{Spec}(A) = \{\lambda \in \mathbb{C}; \lambda \operatorname{Id} - A \notin \operatorname{GL}(H)\} \subset \{\lambda \in \mathbb{C}; |\lambda| \le \|A\|\}. \tag{3.134}$$

As follows from the discussion above it is a compact set – in fact it cannot be empty. This is quite easy to see if you know a little complex analysis since it follows from Liouville's Theorem. One way to show that $\lambda \in \operatorname{Spec}(A)$ is to check that $\lambda - A$ is not injective, since then it cannot be invertible. This means precisely that λ is an *eigenvalue* of A :

$$\exists\ 0 \ne u \in \mathcal{H} \text{ s.t. } Au = \lambda u. \tag{3.135}$$

However, you should strongly resist the temptation to think that the spectrum is the set of eigenvalues of A, this is sometimes but by no means always true. The other way to show that $\lambda \in \operatorname{Spec}(A)$ is to prove that $\lambda - A$ is not *surjective*. Note that by the Open Mapping Theorem if $\lambda - A$ is both surjective and injective then $\lambda \in \operatorname{Res}(A)$.

For a finite rank operator the spectrum does consist of the set of eigenvalues.

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

Proposition 3.16. *If $A : H \longrightarrow H$ is a bounded operator on a Hilbert space and $A^* = A$ then $A - \lambda \operatorname{Id}$ is invertible for all $\lambda \in \mathbb{C} \setminus [-\|A\|, \|A\|]$ and conversely at least one of $A - \|A\| \operatorname{Id}$ and $A + \|A\| \operatorname{Id}$ is not invertible.*

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

Lemma 3.14. *If $A^* = A \in \mathcal{B}(H)$ then*

$$\|A\| = \sup_{\|u\|=1} |\langle Au, u\rangle|. \tag{3.136}$$

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

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

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

$$\begin{aligned}&4\langle Au, v\rangle\\ &= \langle A(u+v), u+v\rangle - \langle A(u-v), u-v\rangle + i\langle A(u+iv), u+iv\rangle - i\langle A(u-iv), u-iv\rangle.\end{aligned} \tag{3.137}$$

By the reality of the left side we can drop the last two terms and use the bound $|\langle Aw, w\rangle| \le a\|w\|^w$ on the first two to see that

$$4\langle Au, v\rangle \le a(\|u+v\|^2 + \|u-v\|^2) = 2a(\|u\|^2 + \|v\|^2) = 4a \tag{3.138}$$

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

This suggests an improvement on the last part of the statement of Proposition 3.16, namely

$$\begin{aligned}&\text{If } A^* = A \in \mathcal{B}(H) \text{ then } a_-, a_+ \in \operatorname{Spec}(A) \text{ and } \operatorname{Spec}(A) \subset [a_-, a_+]\\ &\qquad \text{where } a_- = \inf_{\|u\|=1} \langle Au, u\rangle,\ a_+ = \sup_{\|u\|=1} \langle Au, u\rangle.\end{aligned} \tag{3.139}$$

Observe that Lemma 3.14 shows that $\|A\| = \max(a_+, -a_-)$.

Proof of Proposition 3.16. First we show that if $A^* = A$ then $\operatorname{Spec}(A) \subset \mathbb{R}$. Thus we need to show that if $\lambda = s + it$ where $t \neq 0$ then $A - \lambda$ is invertible. Now $A - \lambda = (A - s) - it$ and $A - s$ is bounded and selfadjoint, so it is enough to consider the special case that $\lambda = it$. Then for any $u \in \mathcal{H}$,

$$\operatorname{Im}\langle (A - it)u, u\rangle = -t\|u\|^2. \tag{3.140}$$

So, certainly $A - it$ is injective, since $(A - it)u = 0$ implies $u = 0$ if $t \neq 0$. The adjoint of $A - it$ is $A + it$ so the adjoint is injective too. It follows from (3.71) that the range of $A - it$ is dense in $\mathcal{H}$. By this density of the range, if $w \in H$ there exists a sequence $u_n \in \mathcal{H}$ with $w_n = (A - it)u_n \to w$. So again we find that

$$\begin{aligned}&|\operatorname{Im}\langle (A - it)\langle u_n - u_m), (u_n - u_m)\rangle| = |t|\|u_n - u_m\|^2\\ &\qquad = |\operatorname{Im}\langle (w_n - w_m), (u_n - u_m)\rangle| \le \|w_n - w_m\|\|u_n - u_m\|\\ &\qquad\qquad \Longrightarrow \|u_n - u_m\| \le \frac{1}{|t|}\|w_n - w_m\|.\end{aligned} \tag{3.141}$$

Since $w_n \to w$ it is a Cauchy sequence and hence u_n is Cauchy so by completeness, $u_n \to u$ and hence $(A - it)u = w$. Thus $A - it$ is 1-1 and onto and $\|A^{-1}\| \le 1/|t|$. So we have shown that $\operatorname{Spec}(A) \subset \mathbb{R}$.

We already know that $\operatorname{Spec}(A) \subset \{z \in \mathbb{C}; |z| \le \|A\|\}$ so finally then we need to show that one of $A \pm \|A\| \operatorname{Id}$ is NOT invertible. This follows from (3.136). 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 $\langle Au_n, u_n\rangle \to -\|A\|$. Assume we are in the first case, so $\langle Au_n, u_n\rangle \to \|A\|$. Then

$$\begin{aligned}\|(A - \|A\|)u_n\|^2 &= \|Au_n\|^2 - 2\|A\|\langle Au_n, u_n\rangle + \|A\|^2\|u_n\|^2\\ &\le 2\|A\|^2 - 2\|A\|\langle Au_n, u_n\rangle \to 0.\end{aligned} \tag{3.142}$$

Since the sequence is positive it follows that $\|(A - \|A\|)u_n\| \to 0$. This means that $A - \|A\| \operatorname{Id}$ is not invertible, since if it had a bounded inverse B then $1 = \|u_n\| \le \|B\|\|(A - \|A\|)u_n\|$ which is impossible. In the other case it follows similarly that $A + \|A\|$ is not invertible, or one can replace A by $-A$ and use the same argument. So one of $A \pm \|A\|$ is not invertible. □

Only slight modifications of this proof are needed to give (3.139) which we restate in a slightly different form.

Lemma 3.15. *If $A = A^* \in \mathcal{B}(H)$ then*

$$\operatorname{Spec}(A) \subset [\alpha-, \alpha_+] \Longleftrightarrow \alpha_- \le \langle Au, u\rangle \le \alpha_+ \ \forall\ u \in H,\ \|u\| = 1. \tag{3.143}$$

Proof. Take $a_\pm$ to be defined as in (3.139) then set $b = (a_+ - a_-)/2$ and consider $B = A - b \operatorname{Id}$ which is self-adjoint and clearly satisfies

$$\sup_{\|u\|=1} |\langle Bu, u\rangle| = b \tag{3.144}$$

Thus $\|B\| = b$ and $\operatorname{Spec}(B) \subset [-b, b]$ and the argument in the proof above shows that both end-points are in the spectrum. It follows that

$$\{a_-\} \cup \{a_+\} \subset \operatorname{Spec}(A) \subset [a_-, a_+] \tag{3.145}$$

from which the statement follows. □

In particular if $A = A^*$ then

$$\operatorname{Spec}(A) \subset [0, \infty) \Longleftrightarrow \langle Au, u\rangle \geq 0. \tag{3.146}$$

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 3.4. *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*

$$Au_j = \lambda_j u_j,\ \ \lambda_j \in \mathbb{R} \setminus \{0\}, \tag{3.147}$$

combining 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 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 3.3. *If $A \in \mathcal{K}(\mathcal{H})$ is a compact self-adjoint operator on a separable Hilbert space then the equation*

$$u - Au = f \tag{3.148}$$

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

$$u - Au = 0 \tag{3.149}$$

and then (3.148) *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 3.16. *If $A \in \mathcal{K}(\mathcal{H})$ is a self-adjoint compact operator on a separable (possibly finite-dimensional) Hilbert space then*

$$F(u) = \langle Au, u\rangle,\ \ F : \{u \in \mathcal{H}; \|u\| = 1\} \longrightarrow \mathbb{R} \tag{3.150}$$

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

$$\sup_{\|u\|=1} |F(u)| = \|A\|. \tag{3.151}$$

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 (3.136), (3.151) is a direct consequence of Lemma 3.14. Moreover, continuity of F follows from continuity of A and of the inner product so

$$|F(u)-F(u')| \le |\langle Au, u\rangle - \langle Au, u'\rangle| + |\langle Au, u'\rangle - \langle Au', u'\rangle| \le 2\|A\|\|u-u'\| \tag{3.152}$$

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 supremum and infimum. In the general case we need to use the compactness of A. Certainly F is bounded,

$$|F(u)| \le \sup_{\|u\|=1} |\langle Au, u\rangle| \le \|A\|. \tag{3.153}$$

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 properties of weak convergence mean 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 $\|u^\pm\| \le 1$ by the properties of weak convergence. The compactness of A means that $Au_n^\pm \to Au^\pm$ converges strongly, i.e. in norm. But then we can write

$$\begin{aligned} |F(u_n^\pm) - F(u^\pm)| &\le |\langle A(u_n^\pm - u^\pm), u_n^\pm\rangle| + |\langle Au^\pm, u_n^\pm - u^\pm\rangle| \\ &= |\langle A(u_n^\pm - u^\pm), u_n^\pm\rangle| + |\langle u^\pm, A(u_n^\pm - u^\pm)\rangle| \le 2\|Au_n^\pm - Au^\pm\| \end{aligned} \tag{3.154}$$

to deduce that $F(u^\pm) = \lim F(u_n^\pm)$ are respectively the supremum and infimum of F. Thus indeed, as in the finite dimensional case, the supremum and infimum 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

$$L_v : (-\pi, \pi) \ni \theta \longmapsto \cos\theta u^+ + \sin\theta v \tag{3.155}$$

lies in the unit sphere. Expanding out

$$\begin{aligned} &F(L_v(\theta)) = \\ &\qquad \langle AL_v(\theta), L_v(\theta)\rangle = \cos^2\theta F(u^+) + \sin(2\theta)\operatorname{Re}(Au^+, v) + \sin^2(\theta)F(v) \end{aligned} \tag{3.156}$$

we know that this function must take its maximum at $\theta = 0$. The derivative there (it is certainly continuously differentiable on $(-\pi, \pi)$) is $2\operatorname{Re}\langle Au^+, v\rangle$ which must therefore vanish. The same is true for iv in place of v so in fact

$$\langle Au^+, v\rangle = 0 \ \forall\ v \perp u^+,\ \|v\| = 1. \tag{3.157}$$

Taking the span of these v's it follows that $\langle Au^+, v\rangle = 0$ for all $v \perp u^+$ so Au^+ 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. $\square$

PROOF OF THEOREM 3.4. 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

$$\langle Au, v\rangle = \langle u, Av\rangle = 0 \ \forall\ u \in \mathcal{H}_0,\ v \in \operatorname{Nul}(A) \Longrightarrow Au \in \mathcal{H}_0. \tag{3.158}$$

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 3.14 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

$$\mathcal{H}_N = \{u \in \mathcal{H}_0 = \operatorname{Nul}(A)^\perp; \langle u, e_j\rangle = 0,\ j = 1, \dots, N\}. \tag{3.159}$$

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

$$\langle Au, e_j\rangle = \langle u, Ae_j\rangle = \lambda_j\langle u, e_j\rangle = 0 \text{ if } u \in \mathcal{H}_N \Longrightarrow Au \in \mathcal{H}_N. \tag{3.160}$$

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 $\|Ae_j\| \geq a$ 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 contradicts 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 $\mathrm{Nul}(A)^\perp$. This completes the proof of the theorem. $\square$

19. Functional Calculus

As we have seen, the non-zero eigenvalues of a compact self-adjoint operator A form the image of a sequence in $[-\|A\|, \|A\|]$ either converging to zero or finite. If e_j is an orthonormal sequence of eigenfunctions which spans $\mathrm{Nul}(A)^\perp$ with associated eigenvalues λ_i then

$$A = \sum_i \lambda_i P_i,\ \ P_i u = \langle u, e_i\rangle e_i \tag{3.161}$$

being the projection onto the span $\mathbb{C}e_i$. Since $P_iP_j = 0$ if $i \neq j$ and $P_i^2 = P_i$ it follows inductively that the positive powers of A are given by similar sums converging in $\mathcal{B}(H)$:

$$A^k = \sum_i \lambda_i^k P_i,\ \ P_i u = \langle u, e_i\rangle e_i,\ \ k \in \mathbb{N}. \tag{3.162}$$

There is a similar formula for the identity of course, except we need to remember that the null space of A then appears (and the series does *not* usually converge in the norm topology on $\mathcal{B}(H)$) :

$$\mathrm{Id} = \sum_i P_i + P_N,\ \ N = \mathrm{Nul}(A). \tag{3.163}$$

The sum (3.163) can be interpreted in terms of a *strong limit* of operators, meaning that the result converges when applied term by term to an element of H, so

$$u = \sum_i P_i u + P_N u,\ \ \forall\ u \in H \tag{3.164}$$

which is a form of the Fourier-Bessel series. Combining these formulæ we see that for any polynomial $p(z)$

$$p(A) = \sum_i p(\lambda_i) P_i + p(0) P_N \tag{3.165}$$

converges strongly, and in norm provided $p(0) = 0$.

In fact we can do this more generally, by choosing $f \in \mathcal{C}([-\|A\|, \|A\|)$ and defining an operator by

$$f(A) \in \mathcal{B}(H),\ \ f(A)u = \sum_i f(\lambda_i)(u, e_i)e_i \tag{3.166}$$

This series converges in the norm topology provided $f(0) = 0$ so to a compact operator and if f is real it is self-adjoint. You can easily check that, always for $A = A^*$ compact here, this formula defines a bounded linear map

$$\mathcal{C}([-\|A\|, \|A\|] \longrightarrow \mathcal{B}(H) \tag{3.167}$$

which has nice properties. Most importantly

$$(fg)(A) = f(A)g(A),\ \ (f(A))^* = \bar{f}(A) \tag{3.168}$$

so it takes the product of two continuous functions to the product of the operators.

We will proceed to show that such a map exists for any bounded self-adjoint operator. Even though it may not have eigenfunctions – or even if it does, it might not have an orthonormal basis of eigenvectors. Even so, it is still possible to define $f(A)$ for a continous function defined on $[a_-,a_+]$ if $\operatorname{Spec}(A)\subset[a_-,a_+]$. (In fact it only has to be defined on the compact set $\operatorname{Spec}(A)$ which might be quite a lot smaller). This is an effective extension of the spectral theorem to the case of non-compact self-adjoint operators.

How does one define $f(A)$? Well, it is easy enough in case f is a polynomial, since then we can simply substitute A^n in place of z^n. If we factorize the polynomial this is the same as setting

$$f(z)=c(z-z_1)(z-z_2)\dots(z-z_N)\Longrightarrow f(A)=c(A-z_1)(A-z_2)\dots(A-z_N) \tag{3.169}$$

and this is equivalent to (3.166) in case A is also compact.

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 we need to estimate the norm in the polynomial case.

Proposition 3.17. *If $A=A^*\in\mathcal{B}(H)$ is a bounded self-adjoint operator on a Hilbert space then for any polynomial with real coefficients*

$$\|f(A)\|\le\sup_{z\in[a_-,a_+]}|f(z)|,\ \operatorname{Spec}(A)\subset[a_-,a_+]. \tag{3.170}$$

Proof. For a polynomial we have defined $f(A)$ by (3.169). We can drop the constant c since it will just contribute a factor of $|c|$ to both sides of (3.170). Now, recall from Lemma 3.14 that for a self-adjoint operator the norm can be realized as

$$\|f(A)\|=\sup\{|t|;t\in\operatorname{Spec}(f(A))\}. \tag{3.171}$$

That is, we need to think about when $f(A)-t$ is invertible. However, $f(z)-t$ is another polynomial (with leading term z^N because we normalized the leading coefficient to be 1). Thus it can also be factorized:

$$\begin{gathered}f(z)-t=\prod_{j=1}^N(z-\zeta_j(t)),\\ f(A)-t=\prod_{j=1}^N(A-\zeta_j(t))\end{gathered} \tag{3.172}$$

where the $\zeta_j\in\mathbb{C}$ are the roots (which might be complex even though the polynomial is real). Written in this way we can see that

$$(f(A)-t)^{-1}=\prod_{j=1}^N(A-\zeta_j(t))^{-1}\text{ if }\zeta_j(t)\notin\operatorname{Spec}(A)\ \forall\ j. \tag{3.173}$$

Indeed the converse is also true, i.e. the inverse exists if and only if all the $A-\zeta_j(t)$ are invertible, but in any case we see that

$$\operatorname{Spec}(f(A))\subset\{t\in\mathbb{C};\zeta_j(t)\in\operatorname{Spec}(A),\text{ for some }j=1,\dots,N\} \tag{3.174}$$

since if t is not in the right side then $f(A)-t$ is invertible.

Now this can be restated as

$$\operatorname{Spec}(f(A))\subset f(\operatorname{Spec}(A)) \tag{3.175}$$

since $t\notin f(\operatorname{Spec}(A))$ means $f(z)\neq t$ for $z\in\operatorname{Spec}(A)$ which means that there is no root of $f(z)=t$ in $\operatorname{Spec}(A)$ and hence (3.174) shows that $t\notin\operatorname{Spec}(f(A))$. In fact it is easy to see that there is equality in (3.175).

Then (3.170) follows from (3.171), the norm is the sup of $|z|$, for $z \in \operatorname{Spec}(f(A))$ so

$$\|f(A)\| \le \sup_{t \in \operatorname{Spec}(A)} |f(t)|.$$

□

This allows one to pass by continuity to f in the uniform closure of the polynomials, which by the Stone-Weierstrass theorem is the whole of $\mathcal{C}([a_-, a_+])$.

Theorem 3.5. *If $A = A^* \in \mathcal{B}(H)$ for a Hilbert space H then the map defined on polynomials, through* (3.169) *extends by continuity to a bounded linear map*

$$\mathcal{C}([a_-, a_+]) \longrightarrow \mathcal{B}(H) \text{ if } \operatorname{Spec}(A) \subset [a_-, a_+],\ \operatorname{Spec}(f(A)) \subset f([a_-, a_+]). \tag{3.176}$$

Proof. By the Stone-Weierstrass theorem polynomials are dense in continous functions on any compact interval, in the supremum norm. □

Remark 3.1. You should check the properties of this map, which also follow by continuity, especially that (3.168) holds in this more general context. In particular, $f(A)$ is self-adjoint if $f \in \mathcal{C}([a_-, a_+])$ is real-valued and is non-negative if $f \ge 0$ on $\operatorname{Spec}(A)$.

20. Spectral projection

I have not discussed this in lectures but it is natural at this point to push a little further towards the full spectral theorem for bounded self-adjoint operators. If $A \in \mathcal{B}(H)$ is self-adjoint, and $[a_-, a_+] \supset \operatorname{Spec}(A)$, we have defined $f(A) \in \mathcal{B}(H)$ for $A \in \mathcal{C}([a_-, a_+])$ real-valued and hence, for each $u \in H$,

$$\mathcal{C}([a_-, a_+]) \ni f \longmapsto \langle f(A)u, u\rangle \in \mathbb{R}. \tag{3.177}$$

Thinking back to the treatment of the Lebesgue integral, you can think of this as a replacement for the Riemann integral and ask whether it can be extended further, to functions which are not necessarily continuous.

In fact (3.177) is essentially given by a Riemann-Stieltjes integral and this suggests finding the increasing function which defines it. Of course we have the rather large issue that this depends on a vector in Hilbert space as well – clearly we want to allow this to vary too.

One direct approach is to try to define the 'integral' of the characteristic function $(-\infty, a]$ for fixed $a \in \mathbb{R}$. To do this is consider

$$Q_a(u) = \inf\{\langle f(A)u, u\rangle; f \in \mathcal{C}([a_-, a_+]),\ f(t) \ge 0,\ f(t) \ge 1 \text{ on } [a_-, a]\}. \tag{3.178}$$

Since $f \ge 0$ we know that $\langle f(A)u, u\rangle \ge 0$ so the infimum exists and is non-negative. In fact there must exist a sequence f_n such that

$$Q_a(u) = \lim\langle f_n(A)u, u\rangle,\ f_n \in \mathcal{C}([a_-, a_+]),\ f_n \ge 0,\ f_n(t) \ge 1,\ a_- \le t \le a, \tag{3.179}$$

where the sequence f_n could depend on u. Consider an obvious choice for f_n given what we did earlier, namely

$$f_n(t) = \begin{cases} 1 & a_- \le t \le a \\ 1 - (t-a)/n & a \le t \le a + 1/n \\ 0 & t > a + 1/n. \end{cases} \tag{3.180}$$

Certainly

$$Q_a(u) \le \lim\langle f_n(A)u, u\rangle \tag{3.181}$$

where the limit exists since the sequence is decreasing.

LEMMA 3.17. *For any* $a \in [a_-, a_+]$,

$$Q_a(u) = \lim \langle f_n(A)u, u\rangle. \tag{3.182}$$

PROOF. For any given f as in (3.178), and any $\epsilon > 0$ there exists n such that $f(t) \geq 1/(1+\epsilon)$ in $a \leq t \leq a + 1/n$, by continuity. This means that $(1+\epsilon)f \geq g_n$ and hence $\langle f(A)u, u\rangle \geq (1+\epsilon)^{-1}\langle f_n(A)u, u\rangle$ from which (3.182) follows, given (3.181). □

Thus in fact one sequence gives the infimum for all u. Now, use the polarization identity to define

$$Q_a(u, v) = \frac{1}{4}\left(Q_a(u+v) - Q_a(u-v) + iQ_a(u+iv) - iQ_a(u-iv)\right). \tag{3.183}$$

The corresponding identity holds for $\langle f_n(A)u, v\rangle$ so in fact

$$Q_a(u, v) = \lim_{n\to\infty} \langle f_n(A)u, v\rangle. \tag{3.184}$$

It follows that $Q_a(u, v)$ is a sesquilinear form, linear in the first variable and antilinear in the second. Moreover the $f_n(A)$ are uniformly bounded in $\mathcal{B}(H)$ (with norm 1 in fact) so

$$|Q_a(u, v)| \leq C\|u\|\|v\|. \tag{3.185}$$

Now, using the linearity in v of $\overline{Q_a(u, v)}$ and the Riesz Representation theorem it follows that for each $u \in H$ there exists a unique $Q_a u \in H$ such that

$$Q_a(u, v) = \langle Q_a u, v\rangle,\ \forall\ v \in H,\ \|Q_a u\| \leq \|u\|. \tag{3.186}$$

From the uniqueness, $H \ni u \longmapsto Q_a u$ is linear so (3.186) shows that it is a bounded linear operator. Thus we have proved most of

PROPOSITION 3.18. *For each* $a \in [a_-, a_+] \supset \operatorname{Spec}(A)$ *there is a uniquely defined operator* $Q_a \in \mathcal{B}(H)$ *such that*

$$Q_a(u) = \langle Q_a u, u\rangle \tag{3.187}$$

recovers (3.182) *and* $Q_a^* = Q_a = Q_a^2$ *is a projection satisfying*

$$Q_a Q_b = Q_b Q_a = Q_b \text{ if } b \leq a,\ [Q_a, f(A)] = 0\ \forall\ f \in \mathcal{C}([a_-, a_+]). \tag{3.188}$$

This operator, or really the whole family Q_a, is called the spectral projection of A.

PROOF. We have already shown the existence of $Q_a \in \mathcal{B}(H)$ with the property (3.187) and since we defined it directly from $Q_a(u)$ it is unique. Self-adjointness follows from the reality of $Q_a(u) \geq 0$ since $\overline{\langle Q_a u, v\rangle} = \langle u, Q_a v\rangle$ then follows from (3.186).

From (3.184) it follows that

$$\begin{gathered}\langle Q_a u, v\rangle = \lim_{n\to\infty} \langle f_n(A)u, v\rangle \Longrightarrow \\ \langle Q_a u, f(A)v\rangle = \lim_{n\to\infty} \langle f_n(A)u, f(A)v\rangle = \langle Q_a f(A)u, v\rangle\end{gathered} \tag{3.189}$$

since $f(A)$ commutes with $f_n(A)$ for any continuous f. This proves the statement in (3.188). Since $f_n f_m \leq f_n$ is admissible in the definition of Q_a in (3.178)

$$\langle Q_a u, v\rangle = \lim_{n\to\infty} \langle (f_n f_m)(A)u, v\rangle = \lim_{n\to\infty} \langle f_n(A)u, f_m(A)v\rangle = \langle Q_a(A)u, f_m(A)v\rangle \tag{3.190}$$

and now letting $m \to \infty$ shows that $Q_a^2 = Q_a$. A similar argument shows the first identity in (3.188). $\square$

Returning to the original thought that (3.177) represents a Riemann-Stieltjes integral for each u we see that collectively what we have is a map

$$[a_-, a_+] \ni a \longmapsto Q_a \in \mathcal{B}(H) \tag{3.191}$$

taking values in the self-adjoint projections and *increasing* in the sense of (3.188). A little more application allows one to recover the functional calculus as an integral which can be written

$$f(A) = \int_{[a_-, a_+]} f(t) dQ_t \tag{3.192}$$

which does indeed reduce to a Riemann-Stieltjes integral for each u :

$$\langle f(A)u, u\rangle = \int_{[a_-, a_+]} f(t) d\langle Q_t u, u\rangle. \tag{3.193}$$

This, meaning (3.192), is the spectral resolution of the self-adjoint operator A, replacing (and reducing to) the decomposition as a sum in the compact case

$$f(A) = \sum_n f(\lambda_j) P_j \tag{3.194}$$

where the P_j are the orthogonal projections onto the eigenspaces for λ_j.

21. Polar Decomposition

One nice application of the functional calculus for self-adjoint operators is to get the polar decomposition of a general bounded operator.

Lemma 3.18. *If $A \in \mathcal{B}(H)$ then $E = (A^*A)^{\frac{1}{2}}$, defined by the functional calculus, is a non-negative self-adjoint operator.*

Proof. That E exists as a self-adjoint operator satisfying $E^2 = A^*A$ follows directly from Theorem 3.5 and positivity follows as in Remark 3.1. $\square$

Proposition 3.19. *Any bounded operator A can be written as a product*

$$A = U(A^*A)^{\frac{1}{2}},\ U \in \mathcal{B}(H),\ U^*U = \operatorname{Id} - \Pi_{\operatorname{Nul}(A)},\ UU^* = \Pi_{\overline{\operatorname{Ran}(A)}}. \tag{3.195}$$

Proof. Set $E = (A^*A)^{\frac{1}{2}}$. We want to define U and we can see from the first condition, $A = UE$, that

$$U(w) = Av, \quad \text{if } w = Ev. \tag{3.196}$$

This makes sense since $Ev = 0$ implies $\langle Ev, Ev\rangle = 0$ and hence $\langle A^*Av, v\rangle = 0$ so $\|Av\| = 0$ and $Av = 0$. So let us define

$$U(w) = \begin{cases} Av & \text{if } w \in \operatorname{Ran}(E),\ w = Ev \\ 0 & \text{if } w \in (\operatorname{Ran}(E))^{\perp}. \end{cases} \tag{3.197}$$

So U is defined on a dense subspace of H, $\mathrm{Ran}(E) \oplus (\mathrm{Ran}(E))^\perp$ which may not be closed if $\mathrm{Ran}(E)$ is not closed. It follows that

$$\begin{gathered}U(w_1+w_2)=U(w_1)=Av_1 \Longrightarrow \\ \|U(w_1+w_2)\|^2 = |\langle Av_1, Av_1\rangle|^2 = \langle E^2v_1, v_1\rangle = \|Ev_1\|^2 = \|w_1\|^2 \le \|w_1+w_2\|^2 \\ \text{if } w_1 = Ev,\ w_2 \in (\mathrm{Ran}\,E)^\perp. \end{gathered} \tag{3.198}$$

Thus U is bounded on the dense subspace on which it is defined, so has a unique continuous extension to a bounded operator $U \in \mathcal{B}(H)$. From the definition of U the first, factorization, condition in (3.195) holds.

From the definition U vanishes on $\mathrm{Ran}(E)^\perp$. We can now check that the continuous extension is a bijection

$$U : \overline{\mathrm{Ran}(E)} \longrightarrow \overline{\mathrm{Ran}(A)}. \tag{3.199}$$

Indeed, if $w \in \overline{\mathrm{Ran}(E)}$ then $\|w\| = \|Uw\|$ from (3.198) so (3.199) is injective. The same identity shows that the range of U in (3.199) is closed since if Uw_n converges, $\|w_n - w_m\| = \|U(w_n - w_m)\|$ shows that the sequence w_n is Cauchy and hence converges; the range is therefore $\overline{\mathrm{Ran}(A)}$. This same identity, $\|Uw\| = \|w\|$, for $w \in \mathrm{Ran}(E)$, implies that

$$\langle Uw, Uw'\rangle = \langle w, w'\rangle,\ \ w, w' \in Ran(E). \tag{3.200}$$

This follows from the polarization identity

$$\begin{aligned}4\langle Uw, Uw'\rangle &= \|U(w+w')\|^2 - \|U(w-w')\|^2 + i\|U(w+iw')\|^2 - i\|U(w-iw')\|^2 \\ &= \|w+w'\|^2 - \|w-w'\|^2 + i\|w+iw'\|^2 - i\|w-iw'\|^2 = 4\langle w, w'\rangle\end{aligned} \tag{3.201}$$

The adjoint U^* of U has range contained in the orthocomplement of the null space of U, so in $\overline{\mathrm{Ran}(E)}$, and null space precisely $\mathrm{Ran}(A)^\perp$ so defines a linear map from $\mathrm{Ran}(A)$ to $\overline{\mathrm{Ran}(E)}$. As such it follows from (3.201) that

$$U^*U = \mathrm{Id} \ \text{ on } \ \mathrm{Ran}(E) \Longrightarrow U^* = U^{-1} \ \text{ on } \ \mathrm{Ran}(A) \tag{3.202}$$

since U is a bijection it follows that U^* is the two-sided inverse of U as a map in (3.199). The remainder of (3.195) follows from this, so completing the proof of the Proposition. □

A bounded linear operator with the properties of U above, that there are two decompositions of $H = H_1 \oplus H_2 = H_3 \oplus H_4$ into orthogonal closed subspaces, such that $U = 0$ on H_2 and $U : H_1 \longrightarrow H_3$ is a bijection with $\|Uw\| = \|w\|$ for all $w \in H_1$ is called a *partial isometry.* So the polar decomposition writes a general bounded operator as product $A = UE$ where U is a partial isometry from $\overline{\mathrm{Ran}(E)}$ onto $\overline{\mathrm{Ran}(A)}$ and $E = (A^*A)^{\frac{1}{2}}$. If A is injective then U is actually unitary.

Exercise 1. Show that in the same sense, $A = FV$ where $F = (AA^*)^{\frac{1}{2}}$ and V is a partial isometry from $\overline{\mathrm{Ran}(A^*)}$ to $\overline{\mathrm{Ran}\,F}$.

22. 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 they are norm limits of finite rank operators. Accepting this, then you will want to say that an operator such as

$$\mathrm{Id} - K,\ K \in \mathcal{K}(\mathcal{H}) \tag{3.203}$$

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

$$Ku - \lambda u = 0 \tag{3.204}$$

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

Lemma 3.19. *If $K \in \mathcal{B}(H)$ is a compact operator then $\lambda \in \mathbb{C} \setminus \{0\}$ is an eigenvalue of K if and only if $\lambda \in \mathrm{Spec}(K)$.*

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 3.7 we can choose a basis so that (3.85) holds. Let the span of the e_i be W – since it is finite dimensional it is closed. Then $\mathrm{Id} - K$ acts rather simply – decomposing $H = W \oplus W^\perp$, $u = w + w'$

$$(\mathrm{Id} - K)(w + w') = w + (\mathrm{Id}_W - K')w',\ K' : W \longrightarrow W \tag{3.205}$$

being a matrix with respect to the basis. It follows that 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. A matrix, such as $\mathrm{Id}_W - K'$, is invertible if and only if it is injective, or equivalently surjective. So, the same is true for $\mathrm{Id} - K$.

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, $(\mathrm{Id} - K + F)^{-1} = \mathrm{Id} - B$ is invertible. Then we can write

$$\mathrm{Id} - K = \mathrm{Id} - (K - F) - F = (\mathrm{Id} - (K - F))(\mathrm{Id} - L),\ L = (\mathrm{Id} - B)F. \tag{3.206}$$

Thus, $\mathrm{Id} - K$ is invertible if and only if $\mathrm{Id} - L$ is invertible. Thus, if $\mathrm{Id} - K$ is *not* invertible then $\mathrm{Id} - L$ is not invertible and hence has null space and from (3.206) it follows that $\mathrm{Id} - K$ has non-trivial null space, i.e. K has 1 as an eigenvalue. □

A little more generally:-

Proposition 3.20. *If $K \in \mathcal{K}(\mathcal{H})$ is a compact operator on a separable Hilbert space then*

$$\begin{gathered}
\mathrm{null}(\mathrm{Id} - K) = \{u \in \mathcal{H}; (\mathrm{Id}_K)u = 0\} \text{ is finite dimensional} \\
\mathrm{Ran}(\mathrm{Id} - K) = \{v \in \mathcal{H}; \exists u \in \mathcal{H},\ v = (\mathrm{Id} - K)u\} \text{ is closed and} \\
\mathrm{Ran}(\mathrm{Id} - K)^\perp = \{w \in \mathcal{H}; (w, Ku) = 0\ \forall\ u \in \mathcal{H}\} \text{ is finite dimensional}
\end{gathered} \tag{3.207}$$

and moreover

$$\dim\left(\mathrm{null}(\mathrm{Id} - K)\right) = \dim\left(\mathrm{Ran}(\mathrm{Id} - K)^\perp\right). \tag{3.208}$$

PROOF OF PROPOSITION 3.20. First let's check this in the case of a finite rank operator $K = T$. Then

$$\operatorname{Nul}(\operatorname{Id} - T) = \{u \in \mathcal{H}; u = Tu\} \subset \operatorname{Ran}(T). \tag{3.209}$$

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

$$\operatorname{Ran}(\operatorname{Id} - T) = \{v \in \mathcal{H}; v = (\operatorname{Id} - T)u \text{ for some } u \in \mathcal{H}\}. \tag{3.210}$$

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 (3.210),

$$\operatorname{Ran}(\operatorname{Id} - T) \supset \operatorname{Nul}(T). \tag{3.211}$$

Now, $\operatorname{Nul}(T)$ is closed and has finite *codimension* – it's orthocomplement is spanned by a finite set which maps to span the image. As shown in Lemma 3.4 it follows from this that $\operatorname{Ran}(\operatorname{Id} - T)$ itself is closed with finite dimensional complement.

This takes care of the case that $K = T$ has finite rank! What about the general case where K is compact? If K is compact then there exists $B \in \mathcal{B}(\mathcal{H})$ and T of finite rank such that

$$K = B + T, \ \|B\| < \frac{1}{2}. \tag{3.212}$$

Now, consider the null space of $\operatorname{Id} - K$ and use (3.212) to write

$$\operatorname{Id} - K = (\operatorname{Id} - B) - T = (\operatorname{Id} - B)(\operatorname{Id} - T'), \ T' = (\operatorname{Id} - B)^{-1}T. \tag{3.213}$$

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

$$\operatorname{Nul}(\operatorname{Id} - K) = \operatorname{Nul}(\operatorname{Id} - T') \text{ is finite dimensional.} \tag{3.214}$$

Here of course we use the fact that $(\operatorname{Id} - K)u = 0$ is equivalent to $(\operatorname{Id} - T')u = 0$ since $\operatorname{Id} - B$ is invertible. So, this is the first condition in (3.207).

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

$$\operatorname{Id} - K = (\operatorname{Id} - B) - T = (\operatorname{Id} - T'')(\operatorname{Id} - B), \ T'' = T(\operatorname{Id} - B)^{-1}. \tag{3.215}$$

Now, T'' is again of finite rank and

$$\operatorname{Ran}(\operatorname{Id} - K) = \operatorname{Ran}(\operatorname{Id} - T'') \text{ is closed and of finite codimension.} \tag{3.216}$$

What about (3.208)? This time let's first check first that it is enough to consider the finite rank case. For a compact operator we have written

$$(\operatorname{Id} - K) = G(\operatorname{Id} - T) \tag{3.217}$$

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

$$\dim \operatorname{Nul}(\operatorname{Id} - K) = \dim \operatorname{Nul}(\operatorname{Id} - T) = \dim \operatorname{Nul}(\operatorname{Id} - K^*). \tag{3.218}$$

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

$$\dim \operatorname{Nul}(\operatorname{Id} - K^*) = \dim \operatorname{Nul}(\operatorname{Id} - T^*) \tag{3.219}$$

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 (3.85), 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

$$T^*v = \sum_{i=1}^{N}(v, f_i)e_i \tag{3.220}$$

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

$$(\operatorname{Id} -T)u = 0 \Longleftrightarrow u \in W \text{ and } Ru = 0,\ (\operatorname{Id} -T^*)u = 0 \Longleftrightarrow u \in W \text{ and } R^*u = 0. \tag{3.221}$$

Thus we really are reduced to the finite-dimensional theorem

$$\dim \operatorname{Nul}(R) = \dim \operatorname{Nul}(R^*) \text{ on } W. \tag{3.222}$$

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

$$\dim \operatorname{Nul}(R) + \dim \operatorname{Ran}(R) = \dim W = \dim \operatorname{Ran}(W) + \dim \operatorname{Nul}(R^*). \tag{3.223}$$

□

23. Hilbert-Schmidt, Trace and Schatten ideals

As well as the finite rank and compact operators there are other important ideals. Since these results are not exploited in the subsequence sections, the many proofs are relegated to exercises.

First consider the Hilbert-Schmidt operators. The definition is based on

LEMMA 3.20. *For a separable Hilbert space, H, if $A \in \mathcal{B}(H)$ then once the sum for any one orthonormal basis $\{e_i\}$*

$$\|A\|^2_{\mathrm{HS}} = \sum_i \|Ae_i\|^2 \tag{3.224}$$

is finite it is finite for any other orthonormal basis and is independent of the choice of basis.

It is straightforward to show that the operators of finite rank satisfy (3.224); this is basically Bessel's inequality.

PROOF. This is Problem XX. Starting from (3.224) for some orthonormal basis e_i, consider any other orthonormal basis f_j. Using the completeness, expand using Bessel's identity

$$\|Ae_i\|^2 = \sum_j |\langle Ae_i, f_j\rangle|^2 = \sum_j |\langle e_i, A^*f_j\rangle|^2. \tag{3.225}$$

This converges absolutely, so the convergence of (3.224) implies the convergence of the double sum, which can then be rearranged to give

$$\sum_i \|Ae_i\|^2 = \sum_i\sum_j |\langle e_i, A^*f_j\rangle|^2 = \sum_j\sum_i |\langle e_i, A^*f_j\rangle|^2 = \sum_j \|A^*f_j\|^2 \tag{3.226}$$

where Bessel's identity is used again. Thus the sum for A^* with respect to the new basis is finite. Applying this argument again shows that the sum is independent of the basis, and the same for the adjoint. □

PROPOSITION 3.21. *The operators for which* (3.224) *is finite form a 2-sided $*$ -ideal* $\mathrm{HS}(H) \subset \mathcal{B}(H)$, *contained in the ideal of compact operators, it is a Hilbert space and the norm satisfies*

$$\|A\|_{\mathcal{B}} \le \|A\|_{\mathrm{HS}} = \left(\sum_i \|Ae_i\|^2\right)^{\frac{1}{2}}, \tag{3.227}$$
$$\|AD\|_{\mathrm{HS}} \le \|A\|_{\mathrm{HS}}\|D\|_{\mathcal{B}},\ A \in \mathrm{HS}(H),\ D \in \mathcal{B}(H).$$

The inner product is

$$\langle A, B\rangle_{\mathrm{HS}} = \sum_i \langle Ae_i, Be_i\rangle_H,\ A,\ B \in \mathrm{HS}(H).$$

For a compact operator the polar decomposition can be given a more explicit form and we can use this to give another characterization of the Hilbert-Schmidt operators.

PROPOSITION 3.22. *If* $A \in \mathcal{K}(H)$ *then there exist orthonormal bases* e_i *of* $\mathrm{Nul}(A)^\perp$ *and* f_j *of* $\mathrm{Nul}(A^*)^\perp$ *such that*

$$Au = \sum_i s_i \langle u, e_i\rangle f_i$$

where the s_i *are the non-zero eigenvalues of* $(A^*A)^{\frac{1}{2}}$ *repeated with multiplicity.*

The s_i are called the *characteristic values* of A.

PROOF. First take a basis e_i of eigenvectors of A^*A restricted to $\mathrm{Nul}(A)^\perp = \mathrm{Nul}(A^*A)^\perp$ with eigenvalues $s_i^2 > 0$, so the s_i are the non-zero eigenvalues of $|A| = (A^*A)^{\frac{1}{2}}$. Then $A = U|A|$ with U a unitary operator from $\overline{\mathrm{Ran}(|A|)} = \mathrm{Nul}(A)^\perp$ to $\overline{\mathrm{Ran}(A)}$ so (3.22) follows by taking $f_i = Ue_i$. □

Extending the e_i to an orthonormal basis of H it follows that

$$\|A\|_{\mathrm{HS}} = \left(\sum_i s_i^2\right)^{\frac{1}{2}} = \|s_*\|_{l^2}.$$

So to say that A is Hilbert-Schmidt is to say that the sequence of its characteristic values is in l^2 (with the caveat that the sequence might be finite).

One reason that the Hilbert-Schmidt operators are of interest is their relation to the ideal of operators 'of trace class', $\mathcal{T}(H)$.

DEFINITION 3.7. The space $\mathcal{T}(H) \subset \mathcal{B}(H)$ for a separable Hilbert space consists of those operators A for which

$$\|A\|_{\mathrm{Tr}} = \sup \sum_i |\langle Ae_i, f_i\rangle| < \infty \tag{3.228}$$

where the supremum is over pairs of orthonormal sequences $\{e_i\}$ and $\{f_i\}$.

PROPOSITION 3.23. *The trace class operators form an ideal, $\mathcal{T}(H) \subset \mathrm{HS}(H)$, which is a Banach space with respect to the norm (3.228) which satisfies*

$$\|A\|_{\mathcal{B}} \le \|A\|_{\mathrm{Tr}},\ \|A\|_{\mathrm{HS}} \le \|A\|_{\mathcal{B}}^{\frac{1}{2}}\|A\|_{\mathrm{Tr}}^{\frac{1}{2}}; \tag{3.229}$$

the following two conditions are equivalent to $A \in \mathcal{T}(H)$:

(1) *The operator defined by the functional calculus,*

$$|A|^{\frac{1}{2}} = (A^*A)^{\frac{1}{4}} \in \mathrm{HS}(H). \tag{3.230}$$

(2) *There are operators $B_i,\ B_i' \in \mathrm{HS}(H)$ such that*

$$A = \sum_{i=1}^{N} B_i' B_i. \tag{3.231}$$

PROOF. Note first that $\mathcal{T}(H)$ is a linear space and that $\|\cdot\|_{\mathrm{Tr}}$ is a norm on it. Now suppose $A \in \mathcal{T}(H)$ and consider its polar decomposition $A = U(A^*A)^{\frac{1}{2}}$. Here U is a partial isometry mapping $\overline{\mathrm{Ran}(A^*A)^{\frac{1}{2}}}$ to $\overline{\mathrm{Ran}(A)}$ and vanishing on $\overline{\mathrm{Ran}(A^*A)^{\frac{1}{2}}}^{\perp}$. Consider an orthonormal basis $\{e_i\}$ of $\overline{\mathrm{Ran}(A^*A)^{\frac{1}{2}}}$. This is an orthonormal sequence in H as is $f_i = Ue_i$. Inserting these into (3.228) shows that

$$\sum_i |\langle U(A^*A)^{\frac{1}{2}} e_i, f_i\rangle| = \sum_i |\langle (A^*A)^{\frac{1}{4}} e_i, (A^*A)^{\frac{1}{4}} e_i\rangle| < \infty \tag{3.232}$$

where we use the fact that $U^* f_i = U^* U e_i = e_i$. Since the closure of the range of $(A^*A)^{\frac{1}{4}}$ is the same as the closure of the range of $(A^*A)^{\frac{1}{2}}$ it follows from (3.232) that (3.230) holds (since adding an orthonormal basis of $\mathrm{Ran}((A^*A)^{\frac{1}{4}})^{\perp}$ does not increase the sum).

Next assume that (3.230) holds for $A \in \mathcal{B}(H)$. Then the polar decomposition can be written $A = (U(A^*A)^{\frac{1}{4}})(A^*A)^{\frac{1}{4}}$ showing that A is the product of two Hilbert-Schmidt operators, so in particular of the form (3.231).

Now assume that A is of the form (3.231), so is a sum of products of Hilbert-Schmidt operators. The linearity of $\mathcal{T}(H)$ means it suffices to assume that $A = BB'$ where $B,\ B' \in \mathrm{HS}(H)$. Then,

$$|\langle Ae, f_i\rangle| = |\langle B' e_i, B^* f_i\rangle| \le \|B' e_i\|_H \|B^* f_i\|_H. \tag{3.233}$$

Taking a finite sum and applying Cauchy-Schwartz inequality

$$\sum_{i=1}^{N} |\langle Ae, f_i\rangle| \le (\sum_{i=1}^{N} \|B' e_i\|^2)^{\frac{1}{2}} (\sum_{i=1}^{N} \|B^* f_i\|^2)^{\frac{1}{2}}. \tag{3.234}$$

If the sequences are orthonormal the right side is bounded by the product of the Hilbert-Schmidt norms so

$$\|BB'\|_{\mathrm{Tr}} \le \|B\|_{\mathrm{HS}} \|B'\|_{\mathrm{HS}} \tag{3.235}$$

and $A = BB' \in \mathcal{T}(H)$.

The first inequality in (3.229) follows the choice of single unit vectors u and v as orthonormal sequences, so

$$|\langle Au, v\rangle| \le \|A\|_{\mathrm{Tr}} \Longrightarrow \|A\| \le \|A\|_{\mathrm{Tr}}. \tag{3.236}$$

The completeness of $\mathcal{T}(H)$ with respect to the trace norm follows standard arguments which can be summarized as follows

(1) If A_n is Cauchy in $\mathcal{T}(H)$ then by the equality just established, it is Cauchy in $\mathcal{B}(H)$ and so converges in norm to $A \in \mathcal{B}(H)$.
(2) A Cauchy sequence is bounded, so there is a constant $C = \sup_n \|A_m\|_{\mathrm{Tr}}$ such that for any N, any orthonormal sequences e_i, f_i,

$$\sum_{i=1}^{N} |\langle A_n e_i, f_i\rangle| \le C. \tag{3.237}$$

Passing to the limit $A_n \to A$ in the finite sum gives the same bound with A_n replaced by A and then allowing $N \to \infty$ shows that $A \in \mathcal{T}(H)$. Similarly the Cauchy condition means that for $\epsilon > 0$ there exists M such that for all N, and any orthonormal sequences e_i, f_i

$$m.n > M \Longrightarrow \sum_{i=1}^{N} |\langle (A_n - A_m)e_i, f_i\rangle| \le \epsilon. \tag{3.238}$$

Passing first to the limit $m \to \infty$ in the finite sum and then $N \to \infty$ shows that

$$n > M \Longrightarrow \|A_n - A\|_{\mathrm{Tr}} \le \epsilon$$

and so $A_n \to A$ in the trace norm.

□

Proposition 3.24. *The trace functional*

$$\mathcal{T}(H) \ni A \longmapsto \mathrm{Tr}(A) = \sum_i \langle Ae_i, e_i\rangle \tag{3.239}$$

is a continuous linear functional (with respect to the trace norm) which is independent of the choice of orthonormal basis $\{e_i\}$ and which satsifies

$$\mathrm{Tr}(AB - BA) = 0 \text{ if } A \in \mathcal{T}(H),\ B \in \mathcal{B}(H) \text{ or } A,\ B \in \mathrm{HS}(H). \tag{3.240}$$

Proof. It suffices to assume that A and B are self-adjoint and $A \in \mathcal{T}(H)$. Then we can choose an orthonormal basis of eigenvectors for A with eigenvalues λ_i which satisfy $\sum\limits_i |\lambda_i| < \infty$ and see that

$$\begin{aligned}\mathrm{Tr}(AB - BA) &= \sum_i (\langle Be_i, Ae_i\rangle - \langle Ae_i, Be_i\rangle) \\ &= \sum_i (\lambda_i \langle Be_i, e_i\rangle - \lambda_i\langle e_i, Be_i\rangle) = 0.\end{aligned} \tag{3.241}$$

□

This is the fundamental property of the trace functional, that it vanishes on commutators where one of the elements is of trace class and the other is bounded. Two other important properties are that

Lemma 3.21. (1) *If A, $B \in \mathrm{HS}(H)$ then*

$$\langle A, B\rangle_{\mathrm{HS}} = \mathrm{Tr}(A^*B) \tag{3.242}$$

(2) *If $T = T^* \in \mathcal{K}(H)$ then $T \in \mathcal{T}(H)$ if and only if the sequence of non-zero eigenvalues λ_j of T (repeated with multiplicity) is in l^1 and*

$$\mathrm{Tr}(T) = \sum_j \lambda_j,\ \ \|T\|_{\mathrm{Tr}} = \sum_j |\lambda_j|. \tag{3.243}$$

Next we turn to the more general Schatten classes.

DEFINITION 3.8. An operator $A \in \mathcal{K}(H)$ is 'of Schatten class,' $A \in \mathrm{Sc}_p(H)$, $p \in [1, \infty)$ if and only if $|A|^p \in \mathcal{T}(H)$, i.e.

$$\|T\|_{\mathrm{Sc}_p} = \left(\sum_i s_i^p\right)^{\frac{1}{p}} < \infty \tag{3.244}$$

where s_i are the non-zero characteristic values of A repeated with multiplicity.

So $\mathcal{T}(H) = \mathrm{Sc}_1(H)$, $\mathrm{HS}(H) = \mathrm{Sc}_2(H)$.

Of course the notation is suggestive, but we need to be a bit careful in proving the results which are implied by the notation!

PROPOSITION 3.25. *Each of the Schatten classes is a two-sided $*$-ideal in $\mathcal{B}(H)$ which is a Banach space with respect to the norm* (3.244)*; the norm is also given by*

$$\|T\|_{\mathrm{Sc}_p}^p = \sup \sum_i |\langle Te_i, f_i\rangle|^p \tag{3.245}$$

with the supremum over orthonormal sequences, with finiteness implying that $T \in \mathrm{Sc}_p(H)$. If q is the conjugate index to $p \in (0, \infty)$ then

$$A \in \mathrm{Sc}_q(H),\ B \in \mathrm{Sc}_p(H) \Longrightarrow AB \in \mathcal{T}(H),\ \|AB\|_{\mathrm{Tr}} \le \|A\|_{\mathrm{Sc}_q}\|B\|_{\mathrm{Sc}_p} \tag{3.246}$$

and conversely, if $A \in \mathcal{B}(H)$ then $A \in \mathrm{Sc}_p(H)$ if and only if $AB \in \mathcal{T}(H)$ for all $B \in \mathrm{Sc}_q(H)$ and

$$\|A\|_{\mathrm{Sc}_p} = \sup_{\|B\|_{\mathrm{Sc}_q}=1} \|AB\|_{\mathrm{Tr}}.$$

PROOF. The alternate realization of the Schatten norm in (3.245) is particularly useful since whilst it is clear from the definition that $cT \in \mathrm{Sc}_p(H)$ if $T \in \mathrm{Sc}_p(H)$ and $c \in \mathbb{C}$, it is not otherwise immediately clear that the space is linear (or that the triangle inequality holds).

From the definition (3.244), that if T is self-adjoint then $T \in \mathrm{Sc}_p(H)$ if and only if

$$\sup \sum_i |\langle Tf_i, f_i\rangle|^p = \|T\|_{\mathrm{Sc}_p}^p < \infty \tag{3.247}$$

with the supremum over orthonormal sequences. To see this let e_j be an orthonomal basis of eigenvectors for T. Then expanding in the Fourier-Bessel series

$$\begin{aligned}\langle Tf_i, f_i\rangle &= \sum_j \lambda_j |\langle f_i, e_j\rangle|^2 \le \sum_j |\lambda_j| |\langle f_i, e_j\rangle|^{2/p} |\langle f_i, e_j\rangle|^{2/q} \\ &\le (\sum_j |\lambda_j|^p \langle f_i, e_j\rangle|^2)^{\frac{1}{p}} (\sum_j |\langle f_i, e_j\rangle|^2)^{\frac{1}{q}} = (\sum_j |\lambda_j|^p \langle f_i, e_j\rangle|^2)^{\frac{1}{p}}\end{aligned} \tag{3.248}$$

by Hölder's inequality, so

$$\sum_i |\langle Tf_i, f_i\rangle|^p \le \sum_j |\lambda_j|^p \sum_i \langle f_i, e_j\rangle|^2 = \sum_j |\lambda_j|^p = \|T\|_{\mathrm{Sc}_p}^p. \tag{3.249}$$

This proves (3.247) when $T = T^* \in \mathcal{K}(H)$.

Now consider (3.245). Let P_N be the orthogonal projection onto the span of the eigenspaces corresponding to the the N largest eigenvalues of $|T|$. Then we replace T by $T_N = TP_N$; certainly $TP_N \to T$ in norm. Since T_N has finite rank, both $\mathrm{Nul}(T_N)$ and $\mathrm{Nul}(T_N^*)$ are infinite dimensional so we can write the polar decomposition as

$$T_N = U_N A_N,\ A_N = P_N|T|P_N$$

and take U_N to be unitary (rather than a partial isometry) by extending it by an isometric isomorphism $\mathrm{Nul}(T_N)^\perp \longrightarrow \mathrm{Nul}(T_N^*)^\perp$. Then using Cauchy-Schwartz inequality and then (3.247) for A_N,

$$\begin{aligned}(3.250)\quad \sum_i |\langle T_N e_i, f_i\rangle|^p &= \sum_i |\langle A_N^{\frac12} e_i, A_N^{\frac12} U_N^* f_i\rangle|^p \\ &\le (\sum_i \|A_N^{\frac12} e_i\|^{2p})^{\frac12} (\sum_i \|A_N^{\frac12} U_N^* f_i\|^{2p})^{\frac12} \\ &\le (\sum_i |\langle A_N e_i, e_i\rangle|^p)^{\frac12} |\sum_i |\langle A_N U_N^* f_i, U^* f_i\rangle|^p)^{\frac12} \\ &\le \|A_N\|^p_{\mathrm{Sc}_p} = \|T_N\|^p_{\mathrm{Sc}_p} \le \|T\|^p_{\mathrm{Sc}_p}.\end{aligned}$$

As usual dropping to a finite sum on the left we can pass to the limit as $N \to \infty$ and obtain a uniform bound on any finite sum for T from which (3.245) follows.

At this point we know that if $A \in \mathrm{Sc}_p(H)$ and U_1, U_2 are unitary then

$$U_1 A U_2 \in \mathrm{Sc}_p(H) \text{ and } \|U_1 A U_2\|_{\mathrm{Sc}_p} = \|A\|_{\mathrm{Sc}_p}.$$

From (3.245) it follows directly that $\mathrm{Sc}_p(H)$ is linear, that the triangle inequality holds, so that $\|\cdot\|_{\mathrm{Sc}_p}$ is a norm, and $\mathrm{Sc}_p(H)$ is complete and that it is $*$-closed.

Now, if $A \in \mathrm{Sc}_q(H)$ and $B \in \mathrm{Sc}_p(H)$ for conjugate indices p, $q \in (1, \infty)$ choose a finite rank orthogonal projection P and consider ABP which is of finite rank, and hence of trace class. We can compute its trace with respect to any orthonormal basis. Choose an orthonormal basis e_i of the range of PAP and f_i so that the polar decomposition of PAP becomes

$$PAPf_i = s_i e_i \Longrightarrow PA^*Pe_i = s_i f_i$$

where the s_i are the characteristic values of PA. With finite sums

$$\begin{aligned}(3.251)\quad |\operatorname{Tr}(PAPBP)| &= |\sum \langle PAP^2BPe_i, e_i\rangle| \\ &= |\sum \langle PBPe_i, PA^*Pe_i\rangle| \le \sum_{i=1}^N s_i |\langle PBPe_i, f_i\rangle| \\ &\le (\sum s_j^q)^{\frac1q} (\sum |\langle PBPe_i, f_i\rangle|^p)^{\frac1p} \le \|PAP\|_{\mathrm{Sc}_q} \|PBP\|_{\mathrm{Sc}_p}\end{aligned}$$

by Hölder's inequality. Now $|PBP| = P|B|P$ (and similarly for A) and from minimax arguments discussed earlier it follows that, $s_j(|PBP|) \le s_j(|B|)$ for all j. So we see that

$$(3.252)\qquad |\operatorname{Tr}(PAPBP)| \le \|A\|_{\mathrm{Sc}_q} \|B\|_{\mathrm{Sc}_p}.$$

Fixing B this is true for any A, so A can be replaced by UA with U unitary, in such a way that $APB = |APB|$. We also know that $\|UA\|_{\mathrm{Sc}_q} = \|A\|_{\mathrm{Sc}_q}$ and since $P|APB|P$ is positive and

$$(3.253)\qquad \operatorname{Tr}(PAPBP) = \operatorname{Tr}(P|APB|P) = \|P|APB|P\|_{\mathrm{Tr}} \le \|A\|_{\mathrm{Sc}_q} \|B\|_{\mathrm{Sc}_p}.$$

Taking an increasing sequence of projections P_N, it follows that $P_N|AP_NB|P_N \to |AB|$ in trace norm and that (3.246) holds.

The proof of optimality in this 'non-commutative Hölder inequality' is left as an exercise. That $\mathrm{Sc}_p(H)$ is an ideal then follows from the fact that $\mathcal{T}(H)$ is an ideal. $\square$

24. Fredholm operators

Definition 3.9. 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 (3.207) – its null space is finite dimensional, its range is closed and the orthocomplement of its range is finite dimensional.

In view of Proposition 3.20, if $K \in \mathcal{K}(H)$ then $\mathrm{Id} - K \in \mathcal{F}(H)$. For general Fredholm operators the row-rank=colum-rank result (3.208) does not hold. Indeed the difference of these two integers, called the index of the operator,

$$\operatorname{ind}(F) = \dim\left(\operatorname{null}(F)\right) - \dim\left(\operatorname{Ran}(F)^{\perp}\right) \tag{3.254}$$

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

Notice that the last two conditions in (3.207) 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 this operator is not invertible and not Fredholm.

On the other hand we do know that a subspace with finite codimension is closed, so we can replace the last two conditions in Definition 3.9 by saying that the range of the operator has finite codimension. I have not done this directly since it is a little too easy to fall into the trap of thinking that it is enough to check that the *closure* of the range has finite codimension; it isn't!

Before looking at general Fredholm operators let's check that, in the case of operators of the form $\mathrm{Id} - K$, with K compact the third conclusion in (3.207) really follows from the first. This is a general fact which I mentioned, at least, earlier but let me pause to prove it.

Proposition 3.26. *If $B \in \mathcal{B}(\mathcal{H})$ is a bounded operator on a Hilbert space and B^* is its adjoint then*

$$\operatorname{Ran}(B)^{\perp} = (\overline{\operatorname{Ran}}(B))^{\perp} = \{v \in \mathcal{H}; (v,w) = 0 \ \forall\ w \in \operatorname{Ran}(B)\} = \operatorname{Nul}(B^*). \tag{3.255}$$

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

$$\begin{aligned} v \in (\operatorname{Ran}(B))^{\perp} &\Longleftrightarrow (v,w) = 0 \ \forall\ w \in \operatorname{Ran}(B) \Longleftrightarrow (v, Bu) = 0 \ \forall\ u \in \mathcal{H} \\ &\Longleftrightarrow (B^*v, u) = 0 \ \forall\ u \in \mathcal{H} \Longleftrightarrow B^*v = 0 \Longleftrightarrow v \in \operatorname{Nul}(B^*). \end{aligned} \tag{3.256}$$

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 $v_n \to w$. $\square$

There is a more 'analytic' way of characterizing Fredholm operators, rather than Definition 3.9.

LEMMA 3.22. *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*

$$\begin{aligned} PF &= \mathrm{Id} - \Pi_{\mathrm{Nul}(F)} \\ FP &= \mathrm{Id} - \Pi_{\mathrm{Ran}(F)^{\perp}} \end{aligned} \tag{3.257}$$

with the two projections of finite rank.

PROOF. If (3.257) 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 $\mathrm{Id} - \Pi_{\mathrm{Nul}(F)^{\perp}}$ and hence it must be closed and of finite codimension.

Conversely, suppose that $F \in \mathcal{F}(H)$. We can divide H into two pieces in two ways as $H = \mathrm{Nul}(F) \oplus \mathrm{Nul}(F)^{\perp}$ and $H = \mathrm{Ran}(F)^{\perp} \oplus \mathrm{Ran}(F)$ 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 only one of these is non-zero and so F corresponds to a bounded linear map $\tilde{F} : \mathrm{Nul}(F)^{\perp} \longrightarrow \mathrm{Ran}(F)$. These are two Hilbert spaces with a bounded linear bijection between them, so the inverse map, $\tilde{P} : \mathrm{Ran}(F) \longrightarrow \mathrm{Nul}(F)^{\perp}$ is bounded by the Open Mapping Theorem and we can define

$$P = \tilde{P} \circ \Pi_{\mathrm{Nul}(F)^{\perp}}. \tag{3.258}$$

Then (3.257) 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 (3.257) also implies that F is Fredholm.

LEMMA 3.23. *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*

$$\begin{aligned} QF &= \mathrm{Id} - E_L \\ FQ &= \mathrm{Id} - E_R \end{aligned} \tag{3.259}$$

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

The term 'parametrix' refers to an inverse modulo an ideal. Here we are looking at the ideal of finite rank operators. In fact this is equivalent to the existence of an inverse modulo compact operators. One direction is obvious – since finite rank operators are compact – the other is covered by one of the problems. Notice that the parametrix Q is itself Fredholm, since reversing the two equations shows that F is a parameterix for Q. Similarly it follows that if F is Fredholm then so is F^* and that the product of two Fredholm operators is Fredholm.

PROOF. If F is Fredholm then $Q = P$ certainly is a parameterix in this sense. Conversely suppose that Q as in (3.259) exists. Then $\mathrm{Nul}(\mathrm{Id} - E_L)$ is finite dimensional – from (3.207) for instance. However, from the first identity $\mathrm{Nul}(F) \subset \mathrm{Nul}(QF) = \mathrm{Nul}(\mathrm{Id} - E_L)$ so $\mathrm{Nul}(F)$ is finite dimensional too. Similarly, the second identity shows that $\mathrm{Ran}(F) \supset \mathrm{Ran}(FQ) = \mathrm{Ran}(\mathrm{Id} - E_R)$ and the last space is closed and of finite codimension, hence so is the first. Thus the existence of such a parameterix Q implies that F is Fredholm.

Now if Q and Q' both satisfy (3.259) with finite rank error terms E'_R and E'_L for Q' then

$$(Q' - Q)F = E_L - E'_L \tag{3.260}$$

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

$$(Q' - Q) = (E_L - E'_L)P + (Q' - Q)\Pi_{\mathrm{Nul}(F)} \tag{3.261}$$

is indeed of finite rank. □

Observe that (3.260) can be reversed. If F is Fredholm, so has a parametrix Q then all the operators $Q + E$ where E is of finite rank are also parametrices. It is also the case that if F is Fredholm and K is compact then $F + K$ is Fredholm. Indeed, if you go through the proof above replacing 'finite rank' by 'compact' you can check this. Thus an operator is Fredholm if and only if it has invertible image in the Calkin algebra, $\mathcal{B}(H)/\mathcal{K}(H)$.

Now recall 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 3.24. *If Q and F satisfy* (3.259) *then*

$$\mathrm{ind}(F) = \mathrm{Tr}(E_L) - \mathrm{Tr}(E_R). \tag{3.262}$$

Proof. We certainly know that (3.262) holds in the special case that $Q = P$ is the generalized inverse of F, since then $E_L = \Pi_{\mathrm{Nul}(F)}$ 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 (3.259) consider the straight line of operators $Q_t = (1-t)P + tQ$. Using the two sets of identities for the generalized inverse and paramaterix

$$\begin{aligned} Q_tF &= (1-t)PF + tQF = \mathrm{Id} - (1-t)\Pi_{\mathrm{Nul}(F)} - tE_L,\\ FQ_t &= (1-t)FP + tFQ = \mathrm{Id} - (1-t)\Pi_{\mathrm{Ran}(F)^\perp} - tE_R. \end{aligned} \tag{3.263}$$

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

$$J(t) = \mathrm{Tr}((1-t)\Pi_{\mathrm{Nul}(F)} + tE_L) - \mathrm{Tr}((1-t)\Pi_{\mathrm{Ran}(F)^\perp} + tE_R) \tag{3.264}$$

is constant. This is a linear function of t as is Q_t. We can differentiate (3.263) with respect to t and see that

$$\begin{aligned} \frac{d}{dt}((1-t)\Pi_{\mathrm{Nul}(F)} + tE_L) - \frac{d}{dt}((1-t)\Pi_{\mathrm{Ran}(F)^\perp} + tE_R) = [Q - P, F]\\ \Longrightarrow J'(t) = 0 \end{aligned} \tag{3.265}$$

since it is the trace of the commutator of a bounded and a finite rank operator (using the last part of Lemma 3.23). □

Proposition 3.27. *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 $B' = \Pi_{\mathrm{Ran}(F)} B \Pi_{\mathrm{Nul}(F)^\perp}$ then $\|B'\| \le \|B\|$ and $B - B'$ is finite rank. If $\tilde{F}$ is the operator constructed in the proof of Lemma 3.22

then $\tilde{F} + B'$ is invertible as an operator from $\operatorname{Nul}(F)^\perp$ to $\operatorname{Ran}(F)$ if $\epsilon > 0$ is small. The inverse, P'_B, extended as 0 to $\operatorname{Nul}(F)$ as P is defined in that proof, satisfies

$$\begin{aligned} P'_B(F+B) &= \operatorname{Id} - \Pi_{\operatorname{Nul}(F)} + P'_B(B - B'), \\ (F+B)P'_B &= \operatorname{Id} - \Pi_{\operatorname{Ran}(F)^\perp} + (B - B')P'_B \end{aligned} \tag{3.266}$$

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 there is an open subset of $\mathcal{B}(H)$ which contains no invertible operators, in strong contrast to the finite dimensional case. In fact even the Fredholm operators do not form a dense subset of $\mathcal{B}(H)$. One such open subset consists of the *semi-Fredholm* operators – those with closed range and with *either* null space or complement of range finite-dimensional.

Why is the index important? For one thing it actually labels the components of the Fredholm operators – two Fredholm operators can be connected by a curve of Fredholms if and only if they have the same index. One of the main applications of the index is quite trivial to see – if the index of a Fredholm operator is positive then the operator must have non-trivial null space. This is a remarkably powerful method for showing that certain sorts ('elliptic' for one) of equations have non-trivial solutions.

25. Kuiper's theorem

For finite dimensional spaces, such as $\mathbb{C}^N$, the group of invertible operators – in this case matrices and 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 – if 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 invertible bounded operators on a separable (complex) infinite-dimensional Hilbert space 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' - it is an unusual name for a theorem - which shows how the infinite-dimensionality is exploited. As you can guess, this is sort of amusing (if you have the right attitude ...). The proof I give here is due to B. Mityagin, [**3**].

Let's denote by $\mathrm{GL}(H)$ this group. In view of the open mapping theorem we know that

$$\mathrm{GL}(H) = \{A \in \mathcal{B}(H); A \text{ is injective and surjective}\}. \tag{3.267}$$

Contractibility means precisely that there is a continuous map

$$\begin{aligned}\gamma : [0,1] \times \mathrm{GL}(H) \longrightarrow \mathrm{GL}(H) \text{ s.t.}\\ \gamma(0,A) = A,\ \gamma(1,A) = \mathrm{Id},\ \forall\ A \in \mathrm{GL}(H).\end{aligned} \tag{3.268}$$

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

I will only show 'weak contractibility' of $\mathrm{GL}(H)$. This has nothing to do with weak convergence, rather just means that we only look for an homotopy over compact sets.

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

$$\mathrm{U}(H) = \{U \in \mathrm{GL}(H); U^{-1} = U^*\}. \tag{3.269}$$

These are the isometric isomorphisms.

Proposition 3.28. *There is a continuous map*

$$\Gamma : [0,1] \times \mathrm{GL}(H) \longrightarrow \mathrm{GL}(H) \textit{ s.t. } \Gamma(0,A) = A,\ \Gamma(1,A) \in \mathrm{U}(H)\ \forall\ A \in \mathrm{GL}(H). \tag{3.270}$$

Proof. This is a consequence of the functional calculus, giving the 'polar decomposition' of invertible (and more generally bounded) operators. Namely, if $A \in \mathrm{GL}(H)$ 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 $A = RU_A$ then

$$\Gamma(s,A) = (s\,\mathrm{Id} + (1-s)R)U_A,\ \ s \in [0,1] \tag{3.271}$$

satisfies (3.270).

There is however the issue of continuity of this map. Continuity in s is clear enough but we also need to show that the map

$$\mathrm{GL}(H) \ni A \longmapsto (A^*A)^{\frac{1}{2}} \in \mathrm{GL}(H), \tag{3.272}$$

defining R, is continuous.

Certainly the map $A \longmapsto A^*A$ is (norm) continuous. Suppose $A_n \to A$ in $\mathrm{GL}(H)$ then given ϵ there is a polynomial p such that

$$\|(B^*B)^{\frac{1}{2}} - p(B^*B)\| \le \epsilon/3 \text{ for } B = A,\ A_n\ \forall\ n. \tag{3.273}$$

On the other hand $(A_n^*A_n)^k \to (A^*A)^k$ for any k so

$$\begin{aligned}&\|(A^*A)^{\frac{1}{2}} - (A_n^*A_n)^{\frac{1}{2}}\| \le\\ &\quad \|(A^*A)^{\frac{1}{2}} - p(A^*A)\| + \|p(A^*A) - p(A_n^*A_n)\| + \|(A_n^*A_n)^{\frac{1}{2}} - p(A_n^*A_n)\| \to 0.\end{aligned} \tag{3.274}$$

$\square$

So, for any compact subset $X \subset \mathrm{GL}(H)$ we seek a continuous map

$$\begin{aligned}\gamma : [0,1] \times X \longrightarrow \mathrm{GL}(H) \text{ s.t.}\\ \gamma(0,A) = A,\ \gamma(1,A) = \mathrm{Id},\ \forall\ A \in X,\end{aligned} \tag{3.275}$$

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 once, just consider (path) connectedness of $\mathrm{GL}(H)$ meaning that there is a continuous map as in (3.275) 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.

LEMMA 3.25 (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*

$$\|Q_L B Q_K\| < \epsilon. \tag{3.276}$$

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

$$\mathbb{N} = K \cup L \cup O,\ K \cap L = K \cap O = L \cap O = \emptyset. \tag{3.277}$$

Thus H_I has orthonormal basis e_k, $k \in I$, $I = K, L, O$. To ensure (3.276) we choose the decomposition (3.277) so that all three sets are infinite and so that

$$|(e_l, Be_k)| < 2^{-l-1}\epsilon \ \forall\ l \in L,\ k \in K. \tag{3.278}$$

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 similarly,

$$\begin{aligned} Q_L B Q_K u &= \sum_{l \in L} (BQ_K u, e_l) e_l = \sum_{k \in L} \sum_{k \in K} (Be_k, e_l)(Q_K u, e_k) e_l \\ \Longrightarrow \|Q_L B Q_K u\|^2 &\le \sum_{k \in K} \left(|(Q_k u, e_k)|^2 \sum_{l \in L} |(Be_k, e_l)|^2 \right) \\ &\le \frac{1}{2}\epsilon^2 \sum_{k \in K} |(Q_k u, e_k)|^2 \le \frac{1}{2}\epsilon^2 \|u\|^2 \end{aligned} \tag{3.279}$$

giving (3.276). The absolute convergence of the series following from (3.278).

Thus, it remains to find a decomposition (3.277) for which (3.278) 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

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

Now, choose $k_{N+1} > l_N$ 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 $L(N+1) = N(N) \cup \{l_{N+1}\}$. $\square$

Given a fixed operator $A \in \mathrm{GL}(H)$ Lemma 3.25 can be applied with $\epsilon = \|A^{-1}\|^{-1}$. It then follows, from the convergence of the Neumann series, that the curve

$$A(s) = A - sQ_L A Q_K,\ s \in [0,1] \tag{3.280}$$

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

$$Q_L B Q_K = 0,\ B = A(1),\ Q_L Q_K = 0 = Q_K Q_L,\ Q_K = Q_K^2,\ Q_L = Q_L^2 \tag{3.281}$$

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 (3.281) means that

$$B = \begin{pmatrix} B_{11} & B_{12} & B_{13} \\ 0 & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{pmatrix},\ Q_K = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ Q_L = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.282}$$

So, now we have a 'little hole'. Under the conditions (3.281) consider

$$P = BQ_K B^{-1}(\mathrm{Id} - Q_L). \tag{3.283}$$

The condition $Q_L B Q_K = 0$ and the definition show that $Q_L P = 0 = PQ_L$. Moreover,

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

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

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

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

The next part of the proof can be thought of as a result on 3×3 matrices but applied to a decomposition of Hilbert space. First, observe a little result on rotations.

Lemma 3.26. *If P and Q are projections on a Hilbert space with $PQ = QP = 0$ and $M = MP = QM$ restricts to an isomorphism from the range of P to the range of Q with 'inverse' $M' = M'Q = PM'$ (so $M'M = P$ and $MM' = Q$)*

$$[-\pi/2, \pi/2] \ni \theta \longmapsto R(\theta) = \cos\theta P + \sin\theta M - \sin\theta M' + \cos\theta Q + (\mathrm{Id} - P - Q) \tag{3.284}$$

is a path in the space of invertible operators such that

$$R(0)P = P,\ R(\pi/2)P = M'P. \tag{3.285}$$

Proof. Computing directly, $R(\theta)R(-\theta) = \mathrm{Id}$ from which the invertibility follows as does (3.285). □

We have shown above that the projection P has range equal to the range of BQ_K; apply Lemma 3.26 with $M = S(BQ_K)^{-1}P$ where S is a fixed isomorphism of the range of Q_K to the range of Q_L. Then

$$L_1(\theta) = R(\theta)B \text{ has } L_1(0) = B,\ L(\pi/2) = B' \text{ with } B'Q_K = Q_L S Q_K \tag{3.286}$$

an isomorphism onto the range of Q.

Next apply Lemma 3.26 again but for the projections Q_K and Q_L with the isomorphism S, giving

$$R'(\theta) = \cos\theta Q_K + \sin\theta S - \sin\theta S' + \cos\theta Q_L + Q_O. \tag{3.287}$$

Then the curve of invertibles

$$L_2(\theta) = R'(\theta - \theta')B' \text{ has } L(0) = B',\ L(\pi/2) = B'',\ B''Q_K = Q_K.$$

So, we have succeed by succesive homotopies through invertible elements in arriving at an operator

$$B'' = \begin{pmatrix} \mathrm{Id} & E \\ 0 & F \end{pmatrix} \tag{3.288}$$

where we are looking at the decomposition of $H = H \oplus H$ according to the projections Q_K and $\mathrm{Id} - Q_K$. The invertibility of this is equivalent to the invertibility of F and the homotopy

$$B''(s) = \begin{pmatrix} \mathrm{Id} & (1-s)E \\ 0 & F \end{pmatrix} \tag{3.289}$$

connects it to

$$L = \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & F \end{pmatrix},\ (B''(s))^{-1} = \begin{pmatrix} \mathrm{Id} & -(1-s)EF^{-1} \\ 0 & F^{-1} \end{pmatrix} \tag{3.290}$$

through invertibles.

The final step is 'Eilenberg's swindle'. In (3.290) we arrived at a family of operators on $H \oplus H$. Reversing the factors we can consider

$$\begin{pmatrix} F & 0 \\ 0 & \mathrm{Id} \end{pmatrix}. \tag{3.291}$$

Eilenberg's idea is to connect this to the identity by an explicit curve which 'only uses F' and so is uniform in parameters. So for the moment just take F to be a fixed unitary operator on H.

We use several isomorphism involving H and $l^2(H)$ which are isomorphic of course, as separable Hilbert spaces. First consider the two simple 'rotations' on $H \oplus H$

$$\begin{pmatrix} \mathrm{Id}\cos t & F\sin t \\ -F^{-1}\sin t & \mathrm{Id}\cos t \end{pmatrix},\ \begin{pmatrix} F\cos t & F\sin t \\ -F^{-1}\sin t & \cos tF^{-1} \end{pmatrix}. \tag{3.292}$$

These are both unitary norm-continuous curves, the first starts at the identity and is off-diagonal at $t = \pi/2$ and equal to the second at that point. So by reversing the second and going back we connect

$$\mathrm{Id} = \begin{pmatrix} \mathrm{Id} & 0 \\ 0 & \mathrm{Id} \end{pmatrix} \text{ to } \begin{pmatrix} F & 0 \\ 0 & F^{-1} \end{pmatrix}. \tag{3.293}$$

Now, we can also identify $H \oplus H$ with $l^2(H \oplus H)$. So an element of this space is an l^2 sequence with values in $H \oplus H$ and the identity just acts as the identity on each 2×2 block. We can perform the to-and-fro rotation in (3.292) in each block. That this is actually a norm-continuous curve acting on $l^2(H \oplus H)$ is a consequence of the fact that it is 'the same' in each block and so it is actually a sequence of operators, each on $H \oplus H$ which is uniformly continuous in the parameter i corresponding to a sequence in l^2, so is collectively continuous.

This connects Id to the second matrix (3.293) acting in each block of $l^2(H\oplus H)$. So, here is one part of the swindle, we can reorder the space so it becomes $l^2(H)$ where now the operator is diagonal but with alternating entries

$$\mathrm{Diag}(F^{-1},F,F^{-1},F,\dots)\text{ on }l^2(H). \tag{3.294}$$

This is just a unitary isomorphism corresponding to relabelling the basis elements.

Now, go back to the operator (3.291) and look at the lower left identity element acting on H. We can identify the H in this spot with $l^2(H)$ and then we have a curve linking this entry to (3.294). For the whole operator this gives a norm-continuous curve connecting

$$\begin{pmatrix}F & 0\\ 0 & \mathrm{Id}\end{pmatrix}\text{ to }\mathrm{Diag}(F,F^{-1},F,F^{-1},F,\dots)\text{ on }H\oplus l^2(H) \tag{3.295}$$

just adding the first entry. But now we reverse the procedure using F^{-1} in place of F so the end-point in (3.295) is connected to the identity on $l^2(H)=H$!

The fact that this construction only uses F itself and 2×2 matrices means that it works uniformly when F depends continously on parameters in a compact set. So we have constructed a curve as desired in (3.275) and hence we have proved:-

Theorem 3.6 (Kuiper). *For any compact subset $X\subset\mathrm{GL}(H)$ there is a retraction γ as in* (3.275).

Note that it follows from a result of Milnor (on CW complexes) that in this case contractibility follows from weak contractibility. If you are topologically inclined you might like to look up some applications of Kuiper's Theorem - for instance that the projective unitary group is a classifying space for two dimensional integral cohomology, an Eilenberg-MacLane space.