Functional Analysis I

Course Notes by Stefan Richter
Transcribed and Annotated by Gregory $\mathbb{Z}$itelli

Bounded Linear Operators

Let $\varphi \in L^\infty(\mu)$ and $X = Y = L^p(\mu)$. Then we can define a linear operator $M_\varphi : L^p \to L^p$ by pointwise multiplication $M_\varphi(f)(x) \to \varphi(x)f(x)$. Furthermore, this operator is bounded, since

$$\|M_\varphi f\|^p_{L^p(\mu)} = \int_X |\varphi f|^p d\mu \leq \|\varphi\|^p_{L^\infty(\mu)} \int_X |f|^p d\mu = \|\varphi\|^p_{L^\infty(\mu)} \|f\|^p_{L^p(\mu)}$$

which imples that $\|M_\varphi\| \leq \|\varphi\|_{L^\infty(\mu)}$. In many situations, we actually have equality for operators defined in a similar way.

Now let $(X, \mathcal{M}, \mu)$ be a finite measure space. Then there is a natural inclusion $L^p(\mu) \subseteq L^q(\mu)$ for $p > q \geq 1$. This fact follows directly from Holder's inequality, by letting $r = \frac{p}{q} > 1$ and $\frac{1}{s} + \frac{1}{r} = 1$,

$$\int_X |f|^q d\mu \leq \left(\int_X 1^s d\mu\right)^{1/s} \left(\int_X |f|^{qr}\right)^{1/r} = \mu(X) \left(\|f\|_{L^p(\mu)}\right)^{1/r}$$

The inclusion $i : L^p(\mu) \to L^q(\mu)$ is a bounded operator, with $\|i\| \leq \mu(X)^{1/sq}$.

The differential properties of the Sobolev spaces are not preserved by a differential operator within an individual space itself, however we doe have that $\frac{d}{dx} : W^{k,p}[a, b] \to W^{k-1,p}[a, b]$ is a bounded linear operator, since

$$\|f\|^p_{W^{k,p}} = \sum_{j=0}^{k} \int |f^{(j)}|^p d\mu$$

$$\left\|\frac{d}{dx} f\right\|^p_{W^{k-1,p}} = \sum_{j=1}^{k} \int |f^{(j)}|^p d\mu \leq \|f\|^p_{W^{k,p}}$$

Taking the derivative of functions in $C^1[a,b]$ with the supremum norm $\|f\|_{C^1} = |f(0)| + \|f\|_\infty$ is still a linear operator, but is certainly unbounded since $\|x^n\|_{C^1} = 1$ but $\left\|\frac{d}{dx}x^n\right\|_{C^1} = n$. Note that $C^1[a,b]$ is not a Banach space since it is not complete.

Hamel Bases

Recall that if X is a vector space, then a Hamel basis of X is a maximal linearly independent set. A Hamel basis B for a vector space always exists, such that for any $y \in X$ there exists unique elements $b_1, \dots b_n \in B$ and scalars $\alpha_1, \dots, \alpha_n \in \mathbb{F}$ where $y = \sum_{i=1}^n \alpha_i b_i$. If $E \subseteq X$ is a linearly independent set, then there always exists a Hamel basis B for X such that $E \subseteq B$, by Zorn's lemma. This allows us to make a precise definition of finite and infinite-dimension vector spaces, where depending on the cardinality of the space's Hamel basis.

Hamel basis can also be used to construct examples of unbounded functions on infinite dimensional Banach spaces. The typical construction follows. If B is a Hamel basis for X, and if we are given a specific scalar $a_b \in \mathbb{F}$ for each $b \in B$, then we can define a linear transformation on X given by $f : X \to \mathbb{F}$ such that $f(b) = a_b$. Such a transformation can be extended linearly so that for $y \in X$, $y = \sum_{i=1}^n \alpha_i b_i$, we have $f(y) = \sum_{i=1}^n \alpha_i a^{b_i}$.

Suppose that we have $C[0,1]$ under the supremum norm, and we let $\{f_n\}_{n=1}^\infty \subseteq C[0,1]$ be a countably infinite, linearly independent set of elements. By the preceding remarks there is a Hamel basis $B = \{f_n\} \cup E$ for $C[0,1]$. Then we can define a function $\varphi : B \to \mathbb{F}$ by $\varphi(f_n) = 2^n \|f_n\|$ and $\varphi(b) = 0$ for $b \in E$, which by the previous paragraph can be extended linearly to all of X. Now if we set $g_N = \sum_{i=1}^N \frac{1}{2^j \|f_j\|} f_j$, where $\|g_N\| \leq 1$, then we have the interesting situation

$$|\varphi(g_N)| = \sum_{j=1}^N \frac{1}{2^j \|f_j\|} \varphi(f_j) = N$$

which demonstrates that φ is unbounded. Such a construction relies heavily on the axiom of choice due to the very existence of the Hamel basis. In the finite dimensional case (where the axiom of choice cannot be evoked), all linear operators are bounded.

Proposition. *Let X be a normed space, $x_0 \in X$, $\varphi : X \to \mathbb{F}$ a linear functional with $\varphi(x_0) \neq 0$. Then*

$$X = \ker \varphi + \{\alpha x_0 : \alpha \in \mathbb{F}\}$$

and φ is bounded $\iff$ $\ker \varphi$ is closed $\iff$ $\ker \varphi$ is not dense in X.

Proof. Let $x \in X$, then

$$x = x - \frac{\varphi(x)}{\varphi(x_0)}x_0 + \frac{\varphi(x)}{\varphi(x_0)}x_0$$

where $x - \frac{\varphi(x)}{\varphi(x_0)}x_0 \in \ker\varphi$ and $\frac{\varphi(x)}{\varphi(x_0)}x_0 = \alpha x_0$ for some appropriate α.

Now suppose that φ is bounded. Let $x_n \in \ker\varphi$ with $x_n \to x$, then $\varphi(x_n) = 0$ for all n which implies that $\varphi(x) = 0$ and so $x \in \ker\varphi$.

Next, suppose that $\ker\varphi$ is closed. Since $\varphi(x_0) \neq 0$ we know that $x_0 \notin \ker\varphi$, and so $\ker\varphi$ is not dense in X.

Lastly, suppose that φ is not bounded, and we want to show that $\ker\varphi$ is dense in X. It is enough to show that there exists $z_n \in \ker\varphi$ such that $z_n \to x_0$, since then we can use the decomposition of X into $\ker\varphi + \{\alpha x_0 : \alpha \in \mathbb{F}\}$ to prove denseness.

Using the assumption that φ is unbounded, take x_n with $\|x_n\| = 1$ where $|\varphi(x_n)| \to \infty$. Now write $x_n = y_n + \alpha_n x_0$ with $y_n \in \ker\varphi$, so that $|\varphi(x_n)| = |\alpha_n||\varphi(x_0)|$, which implies that $|\alpha_n| \to \infty$. Then the elements $z_n = -\frac{1}{\alpha_n}y_n$ approximate x_0. ■

Finite-Dimensional Spaces

Definition. Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms over a vector space X. We say that the norms are equivalent if there exist constants $c, C > 0$ such that $c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1$ for all $x \in X$.

Two norms are equivalent whenever their convergence sequences are equivalent, or similar when their Cauchy sequences are equivalent. This is also the same as saying that the two norms generate the same topologies on X.

Theorem. *If X is a finite dimensional normed space, then any two norms are equivalent.*

Proof. Let $\|\cdot\|$ be some norm on X, and $\{e_j\}_{j=1}^N$ a Hamel basis. Then for any $x \in X$, there exist unique scalars $\varphi_j(x) \in \mathbb{F}$ (which can be thought of as linear functionals on X) such that $x = \sum_{j=1}^N \varphi_j(x)e_j$. Let $\|x\|_\infty = \max_{1\leq j\leq n}|\varphi_j(x)|$, which is easy to verify as a norm on X. Now we can simply use the triangle inequality on the original norm so that

$$\|x\| = \left\| \sum_{j=1}^{N} \varphi_j(x) e_j \right\| \leq \sum_{j=1}^{N} |\varphi_j(x)| \|e_j\| \leq \|x\|_\infty \sum_{j=1}^{N} \|e_j\|$$

so that we can take $C = \sum_{j=1}^{N} \|e_j\|$.

Now suppose that the lower inequality is not true, so that there is a sequence of $x_n \in X$ such that $\|x_n\|_\infty = 1$ but $\|x_n\| \to 0$. However, $|\varphi_j(x_n)| \leq \|x_n\|_\infty \leq 1$. Using a truncated diagonalization argument, there is a subsequence $x_{n_k} = \sum_{j=1}^{N} \varphi_j(x_{n_k}) e_j$ which converges to $\sum_{j=1}^{N} \alpha_j e_j$ in the original norm. Since $\|x_n\| \to 0$, we have that $\alpha_j \to 0$ for each j, and so $\varphi_j(x_{n_k}) \to 0$ which contradicts the fact that $\|x_{n_k}\| = 1$. Therefore, the lower inequality must hold, and so the original norm $\|\cdot\|$ is equivalent to $\|\cdot\|_\infty$. Since this can be done for any $\|\cdot\|$, we can show the equivalence of any two norms on the space. ■

Since all norms on a finite dimensional space are equivalent, there is a natural isomorphism between them and C^n under the Euclidean norm.

Corollary. *If X and Y are normed spaces, with X finite dimensional, then every linear operator $T : X \to Y$ is bounded.*

Proof. Let $T : (X, \|\cdot\|_\infty) \to Y$, then

$$\|Tx\|_Y = \left\| T\left(\sum_{j=1}^{N} \varphi_j(x) e_j \right) \right\| \leq \sum_{j=1}^{N} |\varphi_j(x)| \|Te_j\| \leq \|x\|_\infty \sum_{j=1}^{N} \|Te_j\|$$

■

Riesz Representation Theorems

Recall that the dual space X^* to a vector space X is the collection of all bounded linear functionals from X into a field, typically either $\mathbb{R}$ or $\mathbb{C}$. For Hilbert spaces, the two spaces X and X^* are naturally isomorphic.

Theorem (Riesz Representation)**.** *Let $\mathcal{H}$ be a Hilbert space. Then for every $\varphi \in \mathcal{H}^*$ there exists a unique element $x_0 \in \mathcal{H}$ such that $\varphi(x) = \langle x, x_0 \rangle$. Similarly, every function $\varphi(x) = \langle x, x_0 \rangle$ for $x_0 \in \mathcal{H}$ defines an element of $\mathcal{H}^*$ with $\|\varphi\|_* = \|x_0\|$. Consequently, there is a natural isomorphism between $\mathcal{H}$ and $\mathcal{H}^*$.*

Proof. For the second part, clearly $\varphi(x) = \langle x, x_0 \rangle$ is linear and continuous since $|\varphi(x)| = |\langle x, x_0 \rangle| \leq \|x\| \|x_0\|$.

For the first part, φ is continuous means that $\ker \varphi = \mathcal{M}$ is a closed subspace and $\mathcal{M} \neq \mathcal{H}$. Then there exists an element $y_0 \in \mathcal{H}$ such that $y_0 \neq 0$ and $y_0 \perp \mathcal{M}$. Then for $x \in \mathcal{H}$ we can write

$$x = x - \frac{\varphi(x)}{\varphi(y_0)} y_0 + \frac{\varphi(x)}{\varphi(y_0)} y_0$$

where the first difference is in $\mathcal{M}$ and the last term is orthogonal to $\mathcal{M}$. Then

$$\langle x, y_0 \rangle = \left\langle x - \frac{\varphi(x)}{\varphi(y_0)} y_0, y_0 \right\rangle + \frac{\varphi(x)}{\varphi(y_0)} \|y_0\|^2 = \frac{\varphi(x)}{\varphi(y_0)} \|y_0\|^2$$

Therefore, $\varphi(x) = \frac{\varphi(y_0)}{\|y_0\|^2} \langle x, y_0 \rangle$. ■

There are a variety of other kinds of representation theorems as well. If $(X, \mathcal{M}, \mu)$ is a measure space, then the spaces $L^p(\mu)$ and $L^q(\mu)$ are isomorphically dual to one another when $1 < p, q < \infty$ where $\frac{1}{p} + \frac{1}{q} = 1$. The proof follows by naturally pairing elements of $g \in L^q(\mu)$ with the dual elements $G \in (L^p(\mu))^*$ by $G(f) = \int f\overline{g}\, d\mu$. The correspondence is one-to-one, and the norms $\|G\|_{(L^p(\mu))^*} = \|g\|_{L^q(\mu)}$ correspond as in the Hilbert space setting.

In the case where $p = 1$ and the measure μ is σ-finite, the dual of $(L^1(\mu))^*$ is isomorphic to $L^\infty(\mu)$. On the other hand, the dual space to $L^\infty(\mu)$ is larger than $L^1(\mu)$ if the axiom of choice is assumed.

If X be a locally compact Hausdorff space, we define $C_0(X)$ to be the set of functions $f : X \to \mathbb{F}$ which are continuous, and such that for any $\epsilon > 0$ there is a compact set in X outside of which $|f| < \epsilon$. We can actually view $C_0(X)$ is the closure of $C_c(X)$ (continuous functions with compact support) under the supremum norm on functions on X. Additionally, we define $M_0(X)$ as the collection of regular, compactly supported Borel measures on X under the total variation norm.

Theorem. *The dual space of $C_0(X)$ is isomorphic to $M_0(X)$.*

Orthonormal Sets in Hilbert Spaces

Definition. Let $\mathcal{H}$ be a Hilbert space. A set $\mathcal{E} \subseteq \mathcal{H}$ is called orthonormal if for any distinct $e, f \in \mathcal{E}$ we have $\langle e, f\rangle = 0$ and $\|e\| = 1$. We say that $\mathcal{E} \subseteq \mathcal{H}$ is an orthonormal basis if it is orthonormal and maximal with respect to other orthonormal sets.

Our goal now is to be able to represent elements $h \in \mathcal{H}$ as a countable sum $\sum_{e\in\mathcal{E}}\langle h, e\rangle e$ of elements in an orthonormal basis. The sets $\mathcal{E}$ we are dealing with, however, may not be countable. We need a method of defining such a sum for a potentially uncountable collection, in such a way that the sums will converge unconditionally with these orthonormal bases.

If $\mathcal{E} = \{e_1, \ldots, e_n\}$ is an orthonormal set, we write $\mathcal{M}_n$ for the closed linear span of $\mathcal{E}$, the intersection of all subspaces containing $\mathcal{E}$. We can also express $\mathcal{M}_n$ as the collection $\{\sum_{i=1}^n \alpha_i e_i : \alpha_i \in \mathbb{F}\}$. It can be shown rather painlessly that $P_{\mathcal{M}_n} h = \sum_{k=1}^n \langle h, e_k\rangle e_k$ and $\|P_{\mathcal{M}_n} h\|^2 = \sum_{k=1}^n |\langle h, e_k\rangle|^2 \leq \|h\|$.

Next we claim that if $\mathcal{E}$ is an orthonormal set, then $\langle h, e\rangle$ is nonzero for at most a countable number of elements $e \in \mathcal{E}$. This can be accomplished by letting $\mathcal{E}_n = \{e \in \mathcal{E} : |\langle h, e\rangle| \geq 1/n\}$ so that $\bigcup_{n=1}^\infty \mathcal{E}_n$, and then showing that each $\mathcal{E}_n$ is finite. If we let $\{e_1, e_2, \ldots, e_j\} \subseteq \mathcal{E}_n$ be a distinct collection, then

$$\sum_{j=1}^k \frac{1}{n^2} \leq \sum_{j=1}^k |\langle h, e_j\rangle|^2 \leq \|h\|^2$$

However, this implies that $k \leq n^2\|h\|^2$, and so $\mathcal{E}_n$ must be finite.

Definition. We say that $(\Gamma, \leq)$ is a directed set if $\leq$ is a partial ordering on the set Γ with the properties that $\gamma \leq \gamma$ for all $\gamma \in \Gamma$ and if $\gamma, \mu \in \Gamma$, then there always exists some element $\lambda \in \Gamma$ such that $\gamma \leq \lambda$ and $\mu \leq \lambda$.

Let S be any kind of set. If we let $\mathcal{F}$ be the collection of all finite subsets of S, then there is a partial ordering on $\mathcal{F}$ given by inclusion. Here, if $\gamma, \mu \in \mathcal{F}$ then we can take $\lambda = \gamma \cup \mu$ to be the element that dominates both of them.

Definition. Let (X, τ) be a topological space. A net in X is a collection of elements indexed by a directed set, $\{x_\gamma\}_{\gamma\in\Gamma}$. Given a net, we say that $\{x_\gamma\}_{\gamma\in\Gamma} \to x \in X$ or that $\lim_{\gamma\in\Gamma} x_\gamma = x$ if for all open sets $U \in \tau$ such that $x \in U$ there exists some index γ_0 such that for all elements $\gamma \geq \gamma_0$ we have $x_\gamma \in U$.

Nets naturally generalize sequences, since the natural numbers form a directed set with the normal $\leq$ operation.

Definition. Let $A \subseteq \mathcal{H}$, and let $\mathcal{F}$ be the collection of finite subsets of A. For $F \in \mathcal{F}$ we define $S_F = \sum_{a \in F} a$, such that $\{S_F\}_{F \in \mathcal{F}}$ is a net in $\mathcal{H}$. Then we say that the sum $\sum_{a \in A} a = \lim_{F \in \mathcal{F}} S_F$, if such a limit exists.

Our next claim will be that if $\mathcal{E} \subseteq \mathcal{H}$ is an orthonormal subset, then $\sum_{e \in \mathcal{E}} |\langle h, e \rangle|^2$ exists and is no bigger than $\|h\|^2$, which will further show that $\sum_{e \in \mathcal{E}} \langle h, e \rangle e$ exists.

Dimensions of Hilbert Spaces

Definition. Let (X, τ), (Y, σ) be two topological spaces, $f : X \to Y$, and $x_0 \in X$. Then we say that f is continuous at x_0 if for all $U \in \sigma$ with $f(x_0) \in U$, there exists some $V \in \tau$ containing x_0 such that $f(V) \subseteq U$.

Theorem. *A function $f : X \to Y$ is continuous at a point x_0 if and only if for all nets $\{x_\gamma\}_\gamma \to x_0$ in X, we have that $\{f(x_\gamma)\}_\gamma \to f(x_0)$.*

Proof. Suppose that f is continuous at x_0, and let $\{x_\gamma\}_\gamma \to x_0$ be a convergent net. Let $U \in \sigma$ such that $f(x_0) \in U$, so that there exists $V \in \tau$ containing x_0 where $f(V) \subseteq U$. Then since $\{x_\gamma\}_\gamma \to x_0$ there must exist some γ_0 such that $\gamma \geq \gamma_0$ implies that $x_\gamma \in V$, and hence $f(x_\gamma) \in U$.

Next, let $\Gamma = \{V \in \tau : x_0 \in V\}$, with the partial ordering $V_1 \leq V_2$ if $V_2 \subseteq V_1$. Although this is a little counter intuitive, the idea is that larger elements in the ordering are more refined. It is not difficult to show that this turns Γ into a directed set. Now suppose that f is not continuous at x_0. Then there exists $U \in \sigma$ with $f(x_0) \in U$ such that for all $V \in \tau$ with $x_0 \in V$, $f(V)$ is not contained in U. This further implies that there exists $x_V \in V$ with $f(x_V) \notin U$. However, this lets us use $\{x_V\}_{V \in \Gamma}$ as a net converging to x_0. This implies that $\{f(x_V)\}_{V \in \Gamma}$ does not converge to $f(x_0)$, since $f(x_V) \notin U$ for all V. ∎

Returning to our original application of nets, we were considering an orthonormal set $\mathcal{E} \subseteq \mathcal{H}$. Recall that we said that $\mathcal{E}$ is an orthonormal basis if it was orthonormal and maximal. Furthermore, we had the following theorem.

Theorem. *If $\mathcal{E}$ is orthonormal in $\mathcal{H}$, then there exists an orthonormal basis $\mathcal{B}$ with $\mathcal{E} \subseteq \mathcal{B}$.*

We also showed that if $F \subseteq \mathcal{E}$ is a finite subset of $\mathcal{E}$, then for any $h \in \mathcal{H}$ we have that

$$\sum_{e \in F} |\langle h, e\rangle|^2 \leq \|h\|^2$$

Additionally, $\mathcal{F}_h = \{e \in \mathcal{E} : \langle h, e\rangle \neq 0\}$ is either fininte or countably infinite.

Definition. Let $S_f = \sum_{e \in F} \langle h, e\rangle e$. Then if $\{S_F\}_{F \subseteq \mathcal{E}}$ (with F finite) converges to some S as a net, then we write

$$S = \sum_{e \in \mathcal{E}} \langle h, e\rangle e$$

Theorem. *If $h \in \mathcal{H}$, $\mathcal{E}$ is orthonormal, then $\sum_{e \in \mathcal{E}} \langle h, e\rangle e$ exists.*

Proof. Let $\{e \in \mathcal{E} : \langle h, e\rangle \neq 0\} = \{e_n\}$ be a countable (possibly finite) collection of elements, as we've shown before. Then write $S_n = \sum_{k=1}^{n} \langle h, e_k\rangle e_k$. We've also seen before that $\|S_n\|^2 \leq \|h\|^2$, which implies that $\sum_{k=1}^{\infty} |\langle h, e_k\rangle|^2 < \infty$. Now let $\epsilon > 0$, then there exists some N such that $\sum_{N+1}^{\infty} |\langle h, e_k\rangle|^2 \leq \epsilon^2$, which implies that for all $n, m > N$ we have $\|s_n - s_m\|^2 \leq \epsilon^2$. However, this implies that $S_n \to S \in \mathcal{H}$.

We aren't done yet, though. We still need to show that $\{S_F\}_{F \subseteq \mathcal{E}}$ converges to the element S. Set $F_0 = \{e_1, e_2, \ldots, e_N\}$. Then $F \subseteq \mathcal{E}$, F is finite, and if $F_0 \subseteq F$ then

$$\|S - S_F\| \leq \|S - S_{F_0}\| + \|S_{F_0} - S_F\| \leq 2\epsilon$$

■

Theorem. *If $\mathcal{E} \subseteq \mathcal{H}$ is an orthonormal set, then the following are equivalent.*

(i) $\mathcal{E}$ is an orthonormal basis for $\mathcal{H}$.
(ii) If $h \in \mathcal{H}$, and $h \perp \mathcal{E}$, then $h = 0$.
(iii) The closed linear span of $\mathcal{E}$ is $\mathcal{H}$.
(iv) If $h \in \mathcal{H}$ then $h = \sum_{e \in \mathcal{E}} \langle h, e\rangle e$.
(v) If $g, h \in \mathcal{H}$, then $\langle g, h\rangle = \sum_{e \in \mathcal{E}} \langle g, e\rangle \langle e, h\rangle$.
(vi) If $h \in \mathcal{H}$ then $\|h\|^2 = \sum_{e \in \mathcal{E}} |\langle h, e\rangle|^2$

Proof. If $h \perp \mathcal{E}$ and $h \neq 0$, then we can adjoint the unit of h to $\mathcal{E}$ and maintain independent, contradicting the fact that $\mathcal{E}$ is maximal. Therefore, (i) implies (ii).

If the closed linear span of a $\mathcal{E}$ is all of $\mathcal{H}$ then its orthogonal complement is simply $\{0\}$, and so (ii) and (iii) are equivalent.

Let $e_0 \in \mathcal{E}$, then

$$\langle h - y, e_0 \rangle = \langle h, e_0 \rangle - \langle y, e_0 \rangle = \langle h, e_0 \rangle - \sum_{e \in \mathcal{E}} \langle h, e \rangle \langle e, e_0 \rangle = 0$$

So if we assume (ii), then (iv) holds.

Now if (iv) holds, then $h = \sum_{e \in \mathcal{E}} \langle h, e \rangle e$, so that $\langle h, g \rangle = \left\langle \sum_{e \in \mathcal{E}} \langle h, e \rangle e, g \right\rangle = \sum_{e \in \mathcal{E}} \langle h, e \rangle \langle e, g \rangle$. Therefore, (iv) implies (v).

To see that (v) implies (vi), take $g = h$.

Finallly, suppose that (i) does not hold. Then there exists some $\mathcal{E} \subset \mathcal{F}$, where the inclusion is strict and $\mathcal{F}$ is orthonormal. Then there exists $h \in \mathcal{F} \backslash \mathcal{E}$, such that $h \perp \mathcal{E}$. However then $\sum_{e \in \mathcal{E}} |\langle h, e \rangle|^2 = 0$ while $\|h\| = 1$, contradicting (vi). Thus, (vi) implies (i). ■

Corollary. *If $\mathcal{M} \subseteq \mathcal{H}$ is a subspace, and if $\{e_\alpha\}_\alpha$ is an orthonormal basis for $\mathcal{M}$, then for all $x \in H$, $P_{\mathcal{M}} x = \sum_\alpha \langle x, e_\alpha \rangle_\alpha$.*

Theorem. *If $\mathcal{E}$ and $\mathcal{F}$ are two orthonormal bases for $\mathcal{H}$, then the cardinality of $\mathcal{E}$ and $\mathcal{F}$ are the same. We call this cardinality the dimension of $\mathcal{H}$.*

Proposition. *If the dimension of a Hilbert space $\mathcal{H}$ is infinite, then $\mathcal{H}$ is separable if and only if its dimension is countable.*

Proof. Let $\mathcal{E}$ be an orthonormal basis for $\mathcal{H}$, and let $e, f \in \mathcal{F}$ such that $\|e\| = \|f\| = 1$ and $\|e - f\| = \sqrt{2}$. Then $B(e, 1/\sqrt{2}) \cap B(f, 1/\sqrt{2}) = \varnothing$. Then if the dimension of $\mathcal{H}$ is uncountable then there cannot be a countable dense subset, since there are an uncountable number of mutually disjoint balls.

On the other hand, if $\mathcal{E}$ is countable then the collection of elements $\sum_{i=1}^n r_i e_i$ where $r_i \in \mathbb{Q}$ (or $\mathbb{Q}[i]$) is certainly dense in $\mathcal{H}$. ■

Definition. If X and Y are Banach spaces, then we say that a linear map $T : X \to Y$ is isometric if $\|Tx\|_Y = \|x\|_X$ for all $x \in X$. Furthermore, we say that T is an isomorphism if it is isometric and onto. In this case, we say that X and Y are isomorphic.

Note that if $T : X \to Y$ is isometric then it must be injective and have closed range in Y. Therefore, if T is onto then it is clearly bijective, and its inverse is continuous. In fact, T is onto if and only if its range is dense in Y.

Definition. If $\mathcal{H}$ and $\mathcal{K}$ are Hilbert spaces, then $T : \mathcal{H} \to \mathcal{K}$ is called a Hilbert space isomorphism if it is linear, onto, and $\langle Tx, Ty\rangle_{\mathcal{K}} = \langle x, y\rangle_{\mathcal{H}}$ for all $x, y \in \mathcal{H}$.

Although the definition for a Hilbert space isomorphism seems stronger than the Banach space isomorphism because of the inclusion of the inner product, they are in fact equivalent. If there exists a Banach space isomorphism between two Hilbert spaces, then it is actually a Hilbert space isomorphism. This is fairly easy to show using the polarization identity, which equates the inner product with a continuous function of the norm.

If A is any set whatsoever, we can define the space $\ell^2(A)$ to be the collection of functions $x : A \to \mathbb{F}$ such that $\sum_{\alpha\in A} |x(\alpha)|^2 < \infty$. We have already shown how to define these sums in the case where A is uncountable, so this collection is well defined. We can equip this space with an inner product given by

$$\langle x, y\rangle_{\ell^2(A)} = \sum_{\alpha\in A} x(\alpha)\overline{y(\alpha)}$$

which transformed it into a Hilbert space. Furthermore, the elements $e_\alpha = \chi_{\{\alpha\}} \in \ell^2(A)$ form an orthonormal basis. Consequently, there exist Hilbert spaces of arbitrary cardinality.

Lemma. *If $\mathcal{H}$ and $\mathcal{K}$ are Hilbert spaces, and if $\{e_\alpha\}_{\alpha\in A}$ and $\{f_\beta\}_{\beta\in B}$ are orthonormal bases for $\mathcal{H}$ and $\mathcal{K}$ respectively, then if a bijection $\varphi : A \to B$ exists, we can define $V : \mathcal{H} \to \mathcal{K}$ given by $Vh = \sum_{\alpha\in A}\langle h, e_\alpha\rangle f_{\varphi(\alpha)}$ for $h = \sum_{\alpha\in A}\langle h, e_\alpha\rangle e_\alpha$ is a Hilbert space isomorphism.*

Theorem. *Two Hilbert space $\mathcal{H}$ and $\mathcal{K}$ are isomorphic if and only if they have the same dimension.*

Proof. The proof that two spaces having the same cardinality implies they are isomorphic is provided by the preceding lemma. For the other direction, let $V : \mathcal{H} \to \mathcal{K}$ be an isomorphism, with orthonormal bases $\{e_\alpha\}_{\alpha \in A}$ for $\mathcal{H}$. Set $f_\alpha = Ve_\alpha$, such that $\langle f_\alpha, f_\beta \rangle = \langle e_\alpha, e_\beta \rangle = \delta_{\alpha\beta}$. This demonstrates that $\{f_\alpha\}$ is orthonormal.

Now suppose that $g \in \mathcal{K}$ with $g \perp f_\alpha$ for all $\alpha \in A$. Then $g = Vh$ for some $h \in \mathcal{H}$ since V is onto, but $0 = \langle g, f_\alpha \rangle = \langle Vh, f_\alpha \rangle = \langle h, e_\alpha \rangle$, which implies that $h = 0$ and so $g = 0$ as well. ■

The traditional example for a separable Hilbert space is the sequence space $\ell^2(\mathbb{N})$. Being separable, there is an orthonormal basis given by elements with a single entry of 1 somewhere in the sequence, which is a countable collection. However, the Hamel basis for the space is gigantic, which can be demonstrated from the fact that the elements $(1, t, t^2, t^3, \ldots)$ are all linearly independent for varying $t \in (0, 1)$.

Hilbert Space Operators

Just a operators on the finite Hilbert spaces $\mathbb{C}^2$ can be represented by complex matrices, operators on arbitrary Hilbert spaces $T : \ell^2 \to \ell^2$ can also be represented by their action of the basis elements. When ℓ^2 is separable, we can write $T = (a_{ij})_{1 \leq i,j < \infty}$ such that $a_{ij} = \langle Te_j, e_i \rangle$, where $\{e_i\}$ is an orthonormal basis for ℓ^2. We are interested in what the entries a_{ij} can say about T, most importantly when T is bounded.

Lemma. *If $\sum_{j=0}^{\infty} |a_{ij}| \leq M_1$ for all i and $\sum_{i=0}^{\infty} |a_{ij}| \leq M_2$ for all j, then T is a bounded operator with norm $\|T\| \leq \sqrt{M_1 M_2}$.*

Proof. Let $x \in \ell^2$ such that $x = (x_0, x_1, \ldots)$, then

$$|(Tx)_i| = \left| \sum_{j=0}^{\infty} a_{ij} x_j \right| \leq \sqrt{\sum_{j=0}^{\infty} |a_{ij}|} \sqrt{\sum_{j=0}^{\infty} |a_{ij}||x_j|^2} \leq \sqrt{M_1} \sqrt{\sum_{j=0}^{\infty} |a_{ij}||x_j|^2}$$

$$\|Tx\|^2 = \sum_{i=0}^{\infty} |(Tx)_i|^2 \leq M_1 \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} |a_{ij}||x_j|^2 \leq M_1 M_2 \|x\|^2$$

■

The Hilbert matrices provide a counterexample to the converse of the last statement. They are the infinite matrices acting on ℓ^2 given by

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots \\ \frac{1}{2} & \frac{1}{3} & \cdots\cdots & \\ \frac{1}{3} & \cdots\cdots\cdots & & \\ \vdots & & & \end{pmatrix} \qquad \begin{pmatrix} 1 & -\frac{1}{2} & \frac{1}{3} & \cdots \\ -\frac{1}{2} & \frac{1}{3} & \cdots\cdots & \\ \frac{1}{3} & \cdots\cdots\cdots & & \\ \vdots & & & \end{pmatrix}$$

Although the entries are not summable, the operators are indeed bounded with norm $\leq \pi$. This kind of matrix with constant skew diagonal is called a Hankel matrix, and the associated operator a Hankel operator.

Hardy Hilbert Space

The fact that the trigonometric polynomials are dense in $L^2(\mathbb{T})$ (which we will prove later) provides a natural isomorphism between $L^2(\mathbb{T})$ to $\ell^2(\mathbb{Z})$. If $f \in L^(\mathbb{T})$, then we write $\hat{f}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{it})e^{-int}dt$ for its Fourier coefficients. The Fourier transform is the isomorphism from $L^2(\mathbb{T})$ to $\ell^2(\mathbb{Z})$ given by $f \mapsto \{\hat{f}(n)\}_{n\in\mathbb{Z}}$.

Starting with a function $f \in L^2(\mathbb{T})$, it makes sense to think of $f = \sum_{n=-\infty}^{\infty} \langle f, e_n\rangle e_n = \sum_{n=-\infty}^{\infty} \hat{f}(n)e_n$ in the sense of norm convergence of the sum. We would then assume that $f(e^{it}) = \sum_{n=-\infty}^{\infty} \hat{f}(n)e^{int}$. Since the convergence holds on $L^2(\mathbb{T})$, this convergence only holds almost everywhere with respect to Lebesgue measure on $\mathbb{T}$, but that still represents most points in a satisfying sense. Such conditions do not exist for functions in $L^1(\mathbb{T})$, however. In fact, Kolmogorov demonstrated a function in $L^1(\mathbb{T})$ such that its Fourier series $\sum_{n=-\infty}^{\infty} \hat{f}(n)e^{int}$ converges for no t whatsoever.

Theorem. *For $p > 1$, if $f \in L^p(\mathbb{T})$, then $\sum_{n=-\infty}^{\infty} \hat{f}(n)e^{int}$ converges almost everywhere.*

The preceding theorem is incredibly difficult. It was originally proved by Carleson for the case $p = 2$, and then by Hunt for $p > 1$.

Definition. The Hardy space of the circle $H^2(\mathbb{T})$ is the collection of elements in $L^2(\mathbb{T})$ such that the negative Fourier coefficients are zero. $H^2(\mathbb{T})$ is a closed subspace of $L^2(\mathbb{T})$.

Function in the Hardy space only have nonnegative Fourier coefficients, and the Hardy space itself is naturally isomorphic to $\ell^2(\mathbb{N})$ in the same way that $L^2(\mathbb{T})$ is isomorphic to $\ell^2(\mathbb{Z})$. If we take some $\varphi \in L^\infty(\mathbb{T})$, then the operator $M_\varphi : L^2(\mathbb{T}) \to L^2(\mathbb{T})$ given by $f \mapsto \varphi f$. It is clear that $\|M_\varphi\| \leq \|\varphi\|_\infty$. Then we can define $H_\varphi : H^2(\mathbb{T}) \to H^2(\mathbb{T})^\perp$ by $H_\varphi = P\ M_\varphi|_{H^2(\mathbb{T})}$, where P is the orthogonal projection from $L^2(\mathbb{T})$ to $H^2(\mathbb{T})^\perp$. Then $\|H_\varphi\| \leq \|\varphi\|_\infty$ as well. We will see that these operators give rise to Hankel operators.

Hankel Operators

Given an operator A on a separable Hilbert space, we can represent A by an infinite matrix with entries $\langle Ae_j, e_i\rangle = a_{ij}$ denoting its action on the underlying orthonormal basis. By representing elements of $L^2(\mathbb{T})$ by their Fourier series, and noting that $H^2(\mathbb{T}) \subseteq L^2(\mathbb{T})$ is the closed subspace of elements in $L^2(\mathbb{T})$ whose negative Fourier coefficients are all zero, we defined a class of operators H_φ where $\varphi \in L^\infty(\mathbb{T})$ as

$$H_\varphi = P_{H^{2\perp}} M_\varphi|_{H^2}$$

Then the entries of such an operator are then given by

$$\begin{aligned}
a_{mn} &= \langle H_\varphi e_n, f_m\rangle = \langle P_{H^{2\perp}} M_\varphi e_n, f_m\rangle = \langle M_\varphi e_n, f_m\rangle \\
&= \int_{-\pi}^{\pi} \varphi(e^{it}) e_n(e^{it}) \overline{f_m(e^{it})} \frac{dt}{2\pi} = \int_{-\pi}^{\pi} \varphi(e^{it}) e^{i(n+m+1)t} \frac{dt}{2\pi} \\
&= \hat{\varphi}(-(n+m+1))
\end{aligned}$$

Such an operator has constant skew diagonals, making it a Hankel operator. For example, when $\varphi(e^{it}) = t$ we have that $\|\varphi\|_\infty \leq \pi$ and so $\|H_\varphi\| \leq \|\varphi\|_\infty \leq \pi$ and the Hankel operator H_φ has a matrix representation

$$\begin{pmatrix}
1 & -\frac{1}{2} & \frac{1}{3} & \cdots \\
-\frac{1}{2} & \frac{1}{3} & \cdots\cdots & \\
\frac{1}{3} & \cdots\cdots\cdots\cdots & & \\
\vdots & & &
\end{pmatrix}$$

Hilbert Space Adjoints

Definition. Let $\mathcal{H}$ and $\mathcal{J}$ be Hilbert spaces. A sesquilinear form is a function $u : \mathcal{H} \times \mathcal{K} \to \mathbb{F}$ such that

(i) $u(\alpha x + \beta y, z) = \alpha u(x, z) + \beta u(y, z)$ for all $x, y \in \mathcal{H}$, $z \in \mathcal{K}$, and $\alpha, \beta \in \mathbb{F}$.
(ii) $u(x, \alpha y + \beta z) = \overline{\alpha} u(x, y) + \overline{\beta} u(x, z)$ for all $x \in \mathcal{H}$, $y, z \in \mathcal{K}$, and $\alpha, \beta \in \mathbb{F}$.
(iii) We say a sesquilinear form is bounded if there exists some M such that $|u(x, y)| \leq M\|x\|\|y\|$ for all $x \in \mathcal{H}$ and $y \in \mathcal{K}$.

Clearly the inner product is a sesquilinear form. If A is a bounded operator from $\mathcal{H}$ to $\mathcal{K}$ then $u(x, y) = \langle Ax, y\rangle_{\mathcal{K}}$ is a bounded sesquilinear form, and similarly for bounded operators from $\mathcal{K}$ to $\mathcal{H}$. As the next theorem shows, these are the only sesquilinear forms.

Theorem. *If $u : \mathcal{H} \times \mathcal{K} \to \mathbb{F}$ is a bounded sesquilinear form, then there exists unique elements $A \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ and $B \in \mathcal{B}(\mathcal{K}, \mathcal{H})$ such that*

$$u(x, y) = \langle Ax, y\rangle_{\mathcal{K}} = \langle x, By\rangle_{\mathcal{H}}$$

Proof. Let $x \in \mathcal{H}$, and define $L_x : \mathcal{K} \to \mathbb{F}$ be the bounded linear functional given by $\overline{u(x, y)}$. By the Riesz representation theorem, there exists $z \in \mathcal{K}$ such that $L_x(y) = \langle y, z\rangle_{\mathcal{K}}$. Then define $Ax = z$, which applies to every $x \in \mathcal{H}$. It is not hard to show that A is linear and bounded by the same constant that bounds u. A similar construction holds for B. ■

The theorem demonstrates that for any bounded operator U, a unique adjoint operator always exists. This operator is denoted by U^*, so that $\langle Ux, y\rangle = \langle x, U^*y\rangle$, where the inner products may be in different spaces. We can now say that an operator U is a Hilbert space isomorphism if and only if $U^* = U^{-1}$.

Proposition. *Let $\alpha, \beta \in \mathbb{C}$, $A, B \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ and $C \in \mathcal{B}(\mathcal{K}, \mathcal{H})$. Then $(\alpha A + \beta B)^* = \overline{\alpha}A^* + \overline{\beta}B^*$. Furthermore, $(CA)^* = A^*C^*$.*

Proposition. *If $A \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ then $\|A\| = \|A^*\| = \sqrt{\|AA^*\|} = \sqrt{\|A^*A\|}$.*

Self adjoint operators where $A^* = A$ are called hermitian, and operators which commute with their own adjoints are called normal. If $\varphi \in L^\infty(\mu)$, then we have a natural operator $M_\varphi : L^2(\mu) \to L^2(\mu)$ given by $M\varphi f = \varphi f$. These operators M_φ are always normal, and will be self adjoint when φ is strictly real valued. We will later see that all normal operators can be represented as some M_φ for an appropriate measure space $L^2(\mu)$.

Theorem. *If $\mathcal{H}$ is a complex Hilbert space, and if $T \in \mathcal{B}(\mathcal{H})$, then there exists some unique self adjoint operators A and B such that $T = A + iB$.*

Proof. The operators $A = \frac{T+T^*}{2}$ and $B = \frac{T-T^*}{2i}$ can be shown to be self adjoint, and $A + iB = T$. ■

Proposition. *If A is self adjoint, then $\|A\| = \sup\{|\langle Ax, x\rangle| : \|x\| \leq 1\}$.*

Corollary. *Let T be a bounded linear operator on a complex Hilbert space, and suppose $\langle Tx, x\rangle = 0$ for all x, then T is the zero operator.*

Note that the preceding corollary is false for real Hilbert space. For instance, if the Hilbert space is simply $\mathbb{R}^2$ and the operator is

$$T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Corollary. *If $\mathcal{H}$ is a complex Hilbert space, then A is self adjoint if and only if $\langle Ax, x\rangle \in \mathbb{R}$ for every $x \in \mathcal{H}$.*

Proof. If A is self adjoin the $\langle Ax, x\rangle = \langle x, Ax, \rangle = \overline{\langle Ax, x\rangle}$. On the other hand, if $\langle Tx, x\rangle \in \mathbb{R}$ for all $x \in \mathcal{H}$, then write $T = A + iB$ with A, B self adjoint. Then $\langle Tx, x\rangle = \langle Ax, x\rangle + i\langle Bx, x\rangle$, where the two terms are real and so $\langle Bx, x\rangle = 0$ for all $x \in \mathcal{H}$. Then $T = A$, and so T is self adjoint. ■

If A is a self adjoint operator on $\mathbb{C}^n$, then A is diagonalizable with $A = UDU^{-1}$ with U unitary. This is an instance of the spectral theorem for self adjoint operators.

Theorem. *Let $T \in \mathcal{B}(\mathcal{H})$. Then the following are equivalent.*

(a) T is normal.
(b) $\|Tx\| = \|T^*x\|$ for all $x \in \mathcal{H}$.

If $\mathcal{H}$ is a complex Hilbert space, then these are also equivalent to $AB = BA$ for $T = A + iB$ where A and B are self adjoint.

Proof. Suppose $T^*T = TT^*$, then $\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^*x, T^*x\rangle = \|T^*x\|^2$, so (a) implies (b). Next, suppose that $\|Tx\|^2 = \|T^*x\|^2$, then

$$\langle Tx, Tx\rangle - \langle T^*x, T^*x\rangle = 0$$

which leads to $\langle (T^*T - TT^*)x, x\rangle = 0$ for every $x \in \mathcal{H}$. However, $T^*T - TT^*$ is self adjoint, and so $TT^* = T^*T$ and (b) implies (a).

Now suppose that (a) and (b) hold, then if $A = \frac{T+T^*}{2}$ and $B = \frac{T-T^*}{2i}$ then $AB = BA$ by an easy calculation. On the other hand, if (c) holds then $T = A + ib$ and $T^* = A - iB$. Then $TT^* = (A + iB)(A - iB) = (A - iB)(A + iB) = T^*T$, which is precisely (a). ■

Definition. An operator $U \in \mathcal{B}(\mathcal{H})$ is called unitary if it is an $\mathcal{H}$ isomorphism with $\langle Ux, Uy\rangle = \langle x, y\rangle$ for all $x, y \in \mathcal{H}$, or equivalently that $U^* = U^{-1}$.

Theorem. *Let $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$. Then*

(i) $\ker T = (\operatorname{ran} T^*)^\perp$
(ii) $(\ker T)^\perp = \overline{\operatorname{ran} T^*}$
(iii) $\ker T^* = (\operatorname{ran} T)^\perp$
(iv) $(\ker T^*)^\perp = \overline{\operatorname{ran} T}$

Projections and Idempotents

Definition. An operator $E \in \mathcal{B}(\mathcal{H})$ is said to be idempotent if $E^2 = E$.

If E is idempotent, then $I - E$ is also idempotent. Furthermore, the range of E is always closed, and $\operatorname{ran} I - E = \ker E$. Common examples of idempotents are the rank one operators $Az = \langle z, y\rangle x$, usually written $A = x \otimes y$, where $\langle x, y\rangle = 1$. In this case, $A^2z = A(Az) = A(\langle z, y\rangle x) = \langle z, y\rangle Ax = \langle z, y\rangle\langle x, y\rangle x = Az$.

One important property of idempotents is that $\operatorname{ran} E \oplus \ker E = \mathcal{H}$. The kernel and the range may not be perpendicular, however. In such a case, we say that the idempotent is a projection.

Definition. An elements $P \in \mathcal{B}(\mathcal{H})$ is called a projection if it is idempotent and $\operatorname{ran} E = (\ker E)^{\perp}$.

Proposition. *Let E be a nonzero idempotent. Then the following are equivalent.*

(i) E is a projection.
(ii) E is the orthogonal projection onto the range of E.
(iii) $\|E\| = 1$
(iv) E is self adjoint.
(v) E is normal.
(vi) E is positive, such that $\langle Ex, x\rangle \geq 0$ for all $x \in \mathcal{H}$.

Direct Sums

Recall that if $\{e_\alpha\}_\alpha$ is an orthonormal basis for a Hilbert space, then we had an idea of what it meant to have $\sum_\alpha |\langle x, e_\alpha\rangle|^2$. If $\mathcal{H}$ and $\mathcal{K}$ are two Hilbert spaces, then we would like to be able to define their direct sum $\mathcal{H} \oplus \mathcal{K} = \{(x, y) : x \in \mathcal{H}, y \in \mathcal{K}\}$. An appropriate inner product on such a space would be $\langle (x_1, y_1), (x_2, y_2)\rangle = \langle x_1, x_2\rangle_{\mathcal{H}} + \langle y_1, y_2, \rangle_{\mathcal{K}}$, since we get $\|(x, y)\|^2 = \|x\|^2_{\mathcal{H}} + \|y\|^2_{\mathcal{K}}$.

For an arbitrarily indexed collection of Hilbert spaces $\{\mathcal{H}_\alpha\}_\alpha$, we can similarly define $\sum_\alpha \oplus \mathcal{H}_\alpha$, called the external direct sum. Similarly, if $\mathcal{M}, \mathcal{N} \subseteq \mathcal{H}$ are subspaces such that $\mathcal{M} \perp \mathcal{N}$ and $\mathcal{M} \oplus \mathcal{N} = \mathcal{H}$, then we say that $\mathcal{M} \oplus \mathcal{N}$ is an internal direct sum.

Consequently, whenever $\mathcal{M} \subseteq \mathcal{H}$ is a closed subspace, we have the decomposition $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^{\perp}$. If A is a bounded operator on $\mathcal{H}$, it is often useful to express $\mathcal{H}$ in this way so we can write

$$A = \begin{pmatrix} X & Y \\ Z & W \end{pmatrix}$$

where the decomposition means that $A = P_{\mathcal{M}} A P_{\mathcal{M}} + P_{\mathcal{M}^\perp} A P_{\mathcal{M}} + P_{\mathcal{M}} A P_{\mathcal{M}^\perp} + P_{\mathcal{M}^\perp} A P_{\mathcal{M}^\perp} = X + Z + Y + W$. In fact, we also have that

$$A^* = \begin{pmatrix} X^* & Z^* \\ Y^* & W^* \end{pmatrix}$$

Definition. Let $\mathcal{M} \subseteq \mathcal{H}$ be a closed subspace, and let T be a bounded linear operator on $\mathcal{H}$. Then $\mathcal{M}$ is an invariant subspace for T if $Tx \in \mathcal{M}$ for all $x \in \mathcal{M}$. We say that $\mathcal{M}$ is in the lattice of T, written $\text{Lat}(T)$. Furthermore, we say that $\mathcal{M}$ is a reducing subspace for T if both $\mathcal{M}$ and $\mathcal{M}^\perp$ are invariant.

One of the largest open problems in functional analysis is the invariant subspace problem: If $\mathcal{H}$ is a complex Hilbert space with dimension at least two, and if T is some bounded operator on $\mathcal{H}$, is there a nontrivial (not $\{0\}$ or $\mathcal{H}$) invariant subspace for T. If $\mathcal{H}$ is finite dimensional, then every nontrivial operator does indeed have nontrivial invariant subspaces since eigenvectors are unavoidable.

The question also exists for Banach spaces, but has shown to be false for some very strangely constructed counterexamples.

Theorem. *Let $\mathcal{M} \subseteq \mathcal{H}$ be a closed subspace, T a bounded operator on $\mathcal{H}$, and write $P = P_{\mathcal{M}}$. Then the following are equivalent:*

(i) $\mathcal{M} \in \mathit{Lat}(T)$
(ii) $PTP = TP$
(iii) We have the matrix form

$$T = \begin{pmatrix} X & Y \\ 0 & W \end{pmatrix}$$

Similarly, the following are also equivalent.

(i) $\mathcal{M}$ *is reducing for* T.
(ii) $TP = PT$
(iii) We have the matrix form

$$T = \begin{pmatrix} X & 0 \\ 0 & W \end{pmatrix}$$

(iv) $\mathcal{M} \in \mathit{Lat}(T) \cap \mathit{Lat}(T^*)$

For example, consider the bilateral shift $M_{e^{it}} : L^2(\mathbb{T}) \to L^2(\mathbb{T})$ given by $(M_{e^{it}} f)(e^{it}) = e^{it} f(e^{it})$. The operator is in fact unitary, since $(M_{e^{it}})^* = M_{e^{-it}} = (M_{e^{it}})^{-1}$. Then we have the following result.

Theorem (Wiener)**.** *A subspace $\mathcal{M}$ is a reducing subspace for $M_{e^{it}}$ if and only if $\mathcal{M} = \chi_E L^2(\mathbb{T})$ for some measurable $E \subseteq \mathbb{T}$.*

Lemma. *Let $\varphi : \mathbb{T} \to \mathbb{C}$ be measurable. Then M_φ is bounded on $L^2(\mathbb{T})$ if and only if $\varphi \in L^\infty(\mathbb{T})$, in which case $\|M_\varphi\| = \|\varphi\|_\infty$.*

Compact Operators

Definition. Let X and Y be Banach spaces with $A : X \to Y$ linear. We say that A is compact if, setting $X_1 = \{x \in X : \|x\| \le 1\}$, we have that $\overline{AX_1}$ is compact in Y. Not that if A is compact, it must also be bounded.

It should be fairly apparent that if A and B are compact operators, then $\alpha A + \beta B$ is also compact. Consequently, the compact operators form a vector space, called $\mathcal{B}_0(X, Y)$. It is also true that $\mathcal{B}_0(X, Y)$ is norm closed inside of $\mathcal{B}(X, Y)$.

Definition. A set $S \subseteq X$ in a metric space is called totally bounded if for any $\epsilon > 0$ there exists $x_1, \dots, x_n \in X$ such that $S \subseteq \bigcup_{i=1}^n B(x_i, \epsilon)$.

Note that the property of being compact is equivalent to being complete and totally bounded. Additionally, $\overline{S}$ is compact if and only if S is totally bounded.

Theorem. *If S is totally bounded, then for all $\epsilon > 0$ there exists $y_1, \dots, y_n \in S$ such that $S \subseteq \bigcup_{i=1}^n B(y_i, \epsilon)$.*

Lemma. *Let $T_n \in \mathcal{B}_0(\mathcal{E}, \mathcal{F})$ for some Banach spaces $\mathcal{E}$ and $\mathcal{F}$. Then if $T_n \to T$ converges in the norm, then $T \in \mathcal{B}_0(\mathcal{E}, \mathcal{F})$.*

Proof. Let $\epsilon > 0$. Choose n such that $\|T - T_n\| < \frac{\epsilon}{3}$. Since T_n is compact and, $T_n\mathcal{E}_1$ is totally bounded, hence there exists $x_1, \dots, x_m \in \mathcal{E}_1$ such that $T_n\mathcal{E}_1 \subseteq \bigcup_{i=1}^m B\left(T_n x_i, \frac{\epsilon}{3}\right)$. We would like to show that $T\mathcal{E}_1$ is also contained in this union.

Let $x \in \mathcal{E}$, then there exists x_i such that $T_n x \in B\left(T_n x_i, \frac{\epsilon}{3}\right)$. Then $\|Tx - T_n x_i\| \le \|Tx - T_n x\| + \|T_n x - T_n x_i\| \le \|T - T_n\|\|x\| + \frac{\epsilon}{3} < \frac{\epsilon}{3} + \frac{\epsilon}{3} < \epsilon$. ■

Definition. We say that a bounded linear operator $T : \mathcal{E} \to \mathcal{F}$ is a finite rank operator if the rank of T is finite dimensional. We denote these operators by $\mathcal{B}_{00}(\mathcal{E}, \mathcal{F}) \subseteq \mathcal{B}_0(\mathcal{E}, \mathcal{F})$.

Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces, with $T \in \mathcal{B}_{00}(\mathcal{H}, \mathcal{K})$. Write $\mathcal{H} = \ker T \oplus (\ker T)^\perp$, so that $T : (\ker T)^\perp \to \operatorname{ran} T$ is injective. Now let $\{e_k\}_{k=1}^n$ be a finite orthonormal basis for $(\ker T)^\perp$. Let $x \in \mathcal{H}$, then we write $x \sum_{k=1}^n \langle x, e_k \rangle e_k + y$ where $y \in \ker T$, and so

$$Tx\sum_{k=1}^{n}\langle x, e_k\rangle Te_k = \sum_{k=1}^{n}(Te_k \otimes e_k)(x)$$

Consequently, we have the expression $T = \sum_{k=1}^{n} x_k \otimes y_k$, and it van be verified that $T^* = \sum_{k=1}^{n} y_k \otimes x_k$.

Lemma. *Let $T \in \mathcal{B}_0(\mathcal{H}, \mathcal{K})$, then $\overline{\operatorname{ran} T}$ is separable.*

Corollary. *Any polynomial combination of an element $T \in \mathcal{B}_{00}(\mathcal{H}, \mathcal{K})$ between Hilbert spaces is also in $\mathcal{B}_{00}(\mathcal{H}, \mathcal{K})$, and $\mathcal{B}_{00}(\mathcal{H}, \mathcal{K})$ is dense in $\mathcal{B}_0(\mathcal{H}, \mathcal{K})$.*

Corollary. *Let $T \in \mathcal{B}(\mathcal{H}, \mathcal{K})$ be a bounded linear operator between Hilbert spaces. Then the following are equivalent:*

(i) T is compact.
(ii) T^* is compact.
(iii) There exists a sequence of finite rank operators T_n that converge to T in the norm.

Note that the property that compact operators can be approximated by finite rank operators may not be true in a general Banach space setting.

Let T be a bounded linear operator on a Banach space. If $\lambda \in \mathbb{C}$ and x is a nonzero element, then if $Tx = \lambda x$ we say that λ is an eigenvalue of T. This is equivalent to saying that $\ker(T - \lambda) \neq 0$ (here we use $T - \lambda$ to indicate the operator $T - \lambda I$). We define $\sigma_p(T)$, the pointwise spectrum of T, to be the subset of $\mathbb{C}$ of eigenvalues of T. The more general spectrum $\sigma(T)$ of T is the subset of $\mathbb{C}$ such that $(T - \lambda)^{-1}$ does not exist as a bounded operator.

Proposition. *If $T \in \mathcal{B}_0(\mathcal{H})$, and $\lambda \in \mathbb{C}\backslash\{0\}$, then the dimension of $\ker(T - \lambda)$ is finite.*

Proof. Let $\{e_n\}$ be an orthonormal basis for $\ker(T - \lambda)$, so that $Te_n = \lambda e_n$ for all n. Then there exists a subsequence e_{n_k} such that $Te_{n_k} \to z \in \mathcal{H}$, and so $\lambda e_{n_k} \to z$. Then Te_{n_k} is Cauchy, but $\|Te_{n_k} - Te_{n_j}\|^2 = |\lambda|^2 \|e_{n_k} - e_{n_j}\|^2 = 2|\lambda|^2$ which is not going to zero. Therefore, the set e_n must be finite. ■

Corollary. *If T is a nonzero compact operator, then either $\lambda \in \sigma_p(T)$, or $\overline{\lambda} \in \sigma_p(T^*)$, or else $(T-\lambda)^{-1} \in \mathcal{B}(\mathcal{H})$.*

Definition. Let $\mathcal{E}$ be a Banach space, with $T \in \mathcal{B}(\mathcal{E})$. We define the resolvent set of T, denoted $\rho(T)$, as the subset of $\mathbb{C}$ such that $(T-\lambda)^{-1} \in \mathcal{B}(\mathcal{E})$.

The spectrum of an operator can now be viewed as the compliment of the resolvent, so that $\sigma(T) = \mathbb{C}\backslash\rho(T)$. The point spectrum $\sigma_p(T)$ is defined as $\{\lambda \in \mathbb{C} : \ker(T_\lambda) \neq (0)\}$, and the approximate point spectrum $\sigma_{ap}(T)$ as $\{\lambda \in \mathbb{C} : \exists \|x_n\| = 1, \|(T-\lambda)x_n\| \to 0\}$. Clearly,

$$\sigma_p(T) \subseteq \sigma_{ap}(T) \subseteq \sigma(T) = \mathbb{C}\backslash\rho(T)$$

Lemma. *If $A \in \mathcal{B}(\mathcal{E})$, and if there exists $c > 0$ such that $\|Ax\| \geq c\|x\|$ for all $x \in \mathcal{E}$, then $\operatorname{ran} A$ is closed.*

Corollary. *Let $T \in \mathcal{B}(\mathcal{H})$, with $\mathcal{H}$ a Hilbert space. Then $\sigma_{ap}(T) \cup \overline{\sigma_p(T^*)} = \sigma(T)$.*

Proposition. *If $T \in \mathcal{B}_0(\mathcal{H})$, $\lambda \in \sigma_p(T)$ with $\lambda \neq 0$, then the dimension of the kernel of $T - \lambda$ is finite.*

Lemma. *Let $T \in \mathcal{B}_0(\mathcal{H})$, $\lambda \in \mathbb{C}\backslash\{0\}$, then $\lambda\sigma_p(T)$ if and only if $\lambda \in \sigma_{ap}(T)$.*

Proof. If $\lambda \in \sigma_p(T)$ then $\lambda \in \sigma_{ap}(T)$. Alternatively, suppose that $\lambda \in \sigma_{ap}(T)$. Then there exists a sequence of elements $x_n \in \mathcal{H}$ such that $\|x_n\| = 1$ and $\|(T-\lambda)x_n\| \to 0$. Since T is compact, there is a subsequence Tx_{n_k} which converges to some $z \in \mathcal{H}$. However, this means that $\lambda x_{n_k} \to z$ as well, and consequently $|\lambda| = \|z\| \neq 0$. Then $(T-\lambda)z = Tz - \lambda z = \lim_k T(\lambda x_{n_k}) - \lambda z = 0$. Therefore, $\ker(T-\lambda) \neq (0)$, so $\lambda \in \sigma_p(T)$. ■

Theorem. *Let $T \in \mathcal{B}_0(\mathcal{H})$ with $\lambda \neq 0$. Then either $\lambda \in \sigma_p(T)$ or $\overline{\lambda} \in \sigma_p(T^*)$, or else $(T-\lambda)^{-1} \in \mathcal{B}(\mathcal{H})$.*

Proof. Suppose that $\lambda \notin \sigma_p(T)$ and $\overline{\lambda} \notin \sigma_p(T^*)$. Then $\lambda \notin \sigma_{ap}(T)$ by the previous theorem, and since $\sigma(T) = \sigma_{ap}(T) \cup \overline{\sigma_p(T^*)}$ it follows that $\lambda \notin \sigma(T)$. Therefore, $\lambda \in \rho(T)$, so that $(T-\lambda)^{-1} \in \mathcal{B}(\mathcal{H})$. ■

Lemma. *Let $T \in \mathcal{B}(\mathcal{H})$, and $\{\lambda_\alpha\}_{\alpha\in A} \subseteq \sigma_p(T)$. Assume that $\lambda_\alpha \neq \lambda_\beta$ for all $\alpha \neq \beta$, and that if $x_\alpha \in \ker(T-\lambda_\alpha)$ with $x_\alpha \neq 0$ for all α, then $\{x_\alpha\}_{\alpha\in A}$ is linearly independent.*

Proof. We want to show that if $\alpha_1,\ldots,\alpha_n \in A$ and $\gamma_1,\ldots,\gamma_n \in \mathbb{F}$, then $\sum_{j=1}^n \gamma_j x_{\alpha_j} = 0$ implies that $\gamma_j = 0$. Fix some j_0, and let

$$p(z) = \prod_{\substack{j=1\\ j\neq j_0}}^{n} (z-\alpha_j) \qquad\qquad p(T) = \prod_{\substack{j=1\\ j\neq j_0}}^{n} (T-\alpha_j)$$

$$0 = p(T)\sum_{k=1}^n \gamma_k x_{\alpha_k} = \sum_{k=1}^n \gamma_k \prod_{\substack{j=1\\ j\neq j_0}}^{n} (\lambda_{\alpha_k}-\lambda_{\alpha_j})x_{\alpha_k} = \gamma_{j_0}\prod_{\substack{j=1\\ j\neq j_0}}^{n}\left(\lambda_{\alpha_{j_0}}-\lambda_{\alpha_j}\right)x_{\alpha_{j_0}}$$

which implies that $\gamma_{j_0} = 0$. ■

Theorem. *Let $T \in \mathcal{B}_0(\mathcal{H})$. If $\lambda_n \in \sigma_p(T)$, $\lambda_n \neq \lambda_m$ for $n \neq m$, then $\lambda_n \to 0$.*

Theorem (Spectral Theorem for Compact, Self Adjoint Operators)**.** *Let $T \in \mathcal{B}_0(\mathcal{H})$ be self adjoint, and let $\{\lambda_1,\lambda_2,\ldots\}$ be distinct eigenvalues of T. Let P_n be the projection of $\mathcal{H}$ onto the kernel of $T-\lambda_n$. Then*

(i) $P_nP_m = 0$ *for all* $n \neq m$.
(ii) $\lambda_n \in \mathbb{R}$ *for all* n.
(iii) $T = \sum_{k\geq 1}\lambda_k P_k$

Corollary. *If T is compact and self adjoint, then there exists an orthonormal basis $\{e_n\}$ for $\overline{\operatorname{ran} T}$ and $\mu_n \in \mathbb{R}$ such that $T = \sum_{n\geq 1}\mu_n e_n \otimes e_n$.*

Lemma. *Let N be normal, $\lambda,\mu \in \sigma_p(N)$ with $\lambda \neq \mu$. Then if $Ne = \lambda e$ and $Nf = \mu f$, then $e \perp f$.*

Proof. $\|(N^*-\overline{\mu})f\| = \|(N-\mu)f\|$ which implies that $N^*f = \overline{\mu}f$. Then $(\lambda-\mu)\langle e,f\rangle = \langle \lambda e,f\rangle - \langle e,\overline{\mu}f\rangle = \langle Ne,f\rangle - \langle e,N^*f\rangle = 0$. ■

Lemma. *If T is self adjoint, then $\lambda \in \sigma_p(T)$ implies that $\lambda \in \mathbb{R}$. Furthermore, if $\langle Tx, x\rangle \geq 0$ for all $x \in \mathcal{H}$, then $\lambda \geq 0$.*

Lemma. *Let T be self adjoint and compact, then either $\pm\|T\| \in \sigma_p(T)$.*

Proof of Spectral Theorem. Let $\lambda_1 = \pm\|T\|$, depending on what is in $\sigma_p(T)$. If $T = 0$ then $\lambda_1 = 0$ which is trivial, so assume not. Otherwise, assume $\lambda_1 \neq 0$. Then $\ker(T-\lambda_1) \in \text{Lat}T = \text{Lat}T^*$. If $x \in \ker(T-\lambda_1)$ then $(T-\lambda_1)Tx = T(T-\lambda_1)x = 0$, so that $\ker(T-\lambda_1)$ is a reducing subspace for T. Then $TP_1 = P_1T$, and $T = (P_1 + P_1^\perp)T(P_1 + P_1^\perp) = P_1TP_1 + P_1^\perp TP_1^\perp = \lambda_1 P_1 + P_a^\perp TP_1^\perp$. This decomposition can be continued, so that $T = \lambda_1 P_1 + \lambda_2 P_2 + \ldots$, and must either terminate or continue in a countable fashion, where $\lambda_n \to 0$. If the steps are countable, then $\|T - \sum_{k=1}^n \lambda_k P_k\| = |\lambda_{n+1}| \to 0$, so that the sum converges in norm. ■

Lemma. *If T is a bounded operator, then $\sigma(T) \subseteq \{z \in \mathbb{C} : |z| \leq \|T\|\}$.*

Proof. Let $\lambda \in \mathbb{C}$ such that $|\lambda| > \|T\|$. Because of this, $\frac{\|T^n\|}{|\lambda|^{n+1}} \leq \frac{\|T\|^n}{|\lambda|^{n+1}} \to 0$. Note that $\frac{1}{z-\lambda} = -\sum_{n=0}^{\infty} \frac{z^n}{\lambda^{n+1}}$ for $|\lambda| > |z|$.

Now define $T_n = -\sum_{n=0}^{N} \frac{1}{\lambda^{n+1}} T^n$, with $T^0 = I$. Then $\|T_N - T_M\| \leq \sum_{n=N+1}^{M} \frac{\|T\|^n}{|\lambda|^{n+1}}$, so that T_n is a Cauchy sequence. Then $T_N(T-\lambda) = -\frac{1}{\lambda^{N+1}} T^{N+1} + I$, so $\|I - T_n(T-\lambda)\| = \left\|\frac{1}{\lambda^{N+1}} T^{N+1}\right\| \to 0$, so that $T - \lambda$ has a bounded inverse. Therefore, it follows that $\lambda \notin \sigma(T)$. ■

Lemma. *Let T be self adjoint, then $\sigma(T) \subseteq \mathbb{R}$.*

Proof. Recall that $\sigma(T) = \sigma_{ap}(T) \cup \overline{\sigma_p(T^*)} = \sigma_{ap}(T) \cup \overline{\sigma_p(T)}$. We already know that $\sigma_p(T) \subseteq \mathbb{R}$, so it follows that $\sigma(T) = \sigma_{ap}(T)$. ■

Definition. For a bounded operator T, we define the spectral radius as $r_\sigma(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\}$.

Theorem. *If T is self adjoint, $\|T\| = r_\sigma(T)$.*

Theorem (Spectral Mapping Theorem)**.** *Let $p(t)$ be some polynomial in a variable t, and let $p(T)$ be the polynomial applied to a bounded operator T. Then $\mu \in \sigma(p(T))$ if and only if $\mu \in p(\sigma(T))$, which happens if and only if there exists some $\lambda \in \sigma(T)$ such that $\mu = p(\lambda)$.*

Operator Algebras

Definition. A Banach algebra $\mathcal{A}$ is a Banach space with a multiplication operator which transforms it into a ring, such that the multiplication of two elements commutes with multiplication by scalars, and $\|xy\| \leq \|x\|\|y\|$. We say that a map $\phi : \mathcal{A} \to \mathcal{B}$ between two Banach algebras is an algebra homomorphism if it is linear and $\phi(xy) = \phi(x)\phi(y)$.

Theorem. *Let $T \in \mathcal{B}(\mathcal{H})$ be self adjoint. Then there is a unique algebra homomorphism $phi_T : C_{\mathbb{R}}(\sigma(T)) \to \mathcal{B}(\mathcal{H})$ such that $\phi_T(1) = I$ and $\phi_T(g) = T$ for the identity function $g(t) = t$. Furthermore, $\phi_T(f)$ is self adjoint for every $f \in C_{\mathbb{R}}(\sigma(T))$, and commutes with every bounded operator that commutes with T.*

Definition. Let T be self adjoint, with $C_{\mathbb{R}}(\sigma(T))$, then we write $f(T) = \phi_T(f)$.

Lemma. *Let X, Y be Banach spaces, $\mathcal{D} \subseteq X$ dense, and $A : \mathcal{D} \to Y$ linear with $\|Ax\| \leq C\|x\|$ for all $x \in \mathcal{D}$. Then A has a unique extension to all of X.*

As an aside from advanced calculus, if we have a function $f : A \to Y$ where Y is a complete metric space, and f is continuous on a dense set A in a metric space X, f may not have a continuous extension onto all of X. Take, for example, $A = (0, 1] \subseteq [0, 1]$ and $f(x) = \sin \frac{1}{x}$. The necessary condition that we are missing is that f is uniformly continuous. Although the above lemma looks like it may be lacking this, the condition $\|Ax\| \leq C\|x\|$ actually implies both continuity and uniform continuity.

Proof of Theorem. Let $\mathcal{P}$ be the real valued polynomials, which is dense in $C_{\mathbb{R}}(\sigma(T))$ by Weierstrass. We define $\phi_T(p) = p(T)$ for $p \in \mathcal{P}$, making ϕ_T a linear map on the dense set. Furthermore, $\|\phi_T(p)\| = \|p(T)\| = \|p\|_{\infty,\sigma(T)}$. Then by the lemma, ϕ_T extends to $C_{\mathbb{R}}(\sigma(T)) \to \mathcal{B}(\mathcal{H})$ uniquely, with $\|\phi_T(f)\| \leq \|f\|_{\infty,\sigma(T)}$. In fact, the previous inequality is actually equality, by the continuity of ϕ. The various other properties (commuting with other operators, being a homomorphism) follow from the approximation of elements in $C_{\mathbb{R}}(\sigma(T))$ by polynomials. ∎

Corollary. *If $S \in \mathcal{B}(\mathcal{H})$, $S \geq 0$, then S has a unique positive square root $T \in \mathcal{B}(\mathcal{H})$ with $T \geq 0$ and $T^2 = S$.*

Proof. It isn't too difficult to show that $S \geq 0$ implies that $\sigma(S) \subseteq [0, \|S\|]$. Now define $f(t) = \sqrt{t}$ with $f \in C_{\mathbb{R}}(\sigma(S))$. Set $T = \phi_S(f) = \sqrt{S}$. Similarly, let $g(t) = \sqrt[4]{t}$, so that $g^2 = f$ and then

$$\langle Th, h \rangle = \langle f(S)h, h \rangle = \langle g^2(S)h, h \rangle = \langle g(S)h, g(S)h \rangle = \|g(S)h\|^2 \geq 0$$

since the operator $g(S)$ will be self adjoint. Furthermore, $T^* = T$ and $T^2 = f^2(S) = S$. ■