Functional Analysis

Notes by Albert Zhang

August 1, 2020

Contents

1 Normed Vector Spaces

We will almost always work with the fields $\mathbb{F} = \mathbb{R}, \mathbb{C}$, and use X to denote a vector space over $\mathbb{F}$. A *seminorm* on X is a function $x \mapsto \|x\|$ from X to $[0, \infty)$ such that

- $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.
- $\|\lambda x\| = |\lambda| \|x\|$ for all $x \in X$ and $\lambda \in \mathbb{F}$.

A seminorm such that $\|x\| = 0$ only when $x = 0$ is called a *norm*. Note that a norm also induces a metric, given by

$$d(x, y) = \|x - y\|.$$

We say two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on X are *equivalent* if there exist $C_1, C_2 > 0$ such that

$$C_1 \|x\|_1 \leq \|x\|_2 \leq C_2 \|x\|_1 \quad x \in X.$$

A normed vector space that is complete with respect to the norm metric is called a *Banach space*. Some examples are $B(X)$, $BC(X)$, with the uniform norm, and L^p with the p-norm.

Theorem 1.1 (Criterion for Completeness)

A normed vector space X is complete iff every absolutely convergent series in X converges.

Proof. Suppose X is complete, and let $\sum \|x_n\| < \infty$ be an absolutely convergent series. Since Cauchy sequences converge, it suffices to show that

$$\sum_{j=1}^{n} x_j - \sum_{j=1}^{m} x_j = \sum_{j=m}^{n} x_j$$

converges to 0 as $m, n > N \to \infty$. Indeed, we have

$$\left\| \sum_{j=m}^{n} x_j \right\| \leq \sum_{j=m}^{n} \|x_j\| \to 0.$$

Conversely, suppose every absolutely convergent series converges in X. Let $(x_n) \subset X$ be a Cauchy sequence. Pick a subsequence (x_{n_j}) such that for $m, n > n_j$,

$$\|x_m - x_n\| < 2^{-j}.$$

Then the series

$$\sum_{j=1}^{\infty} [x_{n_{j+1}} - x_{n_j}] = \lim_{j \to \infty} x_{n_j} - x_{n_1}$$

converges absolutely. In particular, the series converges, and so

$$\lim_{n \to \infty} x_n = \lim_{j \to \infty} x_{n_j}$$

exists. Thus X is complete. $\square$

We say a linear map $T : X \to Y$ between two normed vector spaces is *bounded* if there is a uniform constant $C \geq 0$ such that

$$\|Tx\| \leq C\|x\|, \quad x \in X.$$

It can then be verified that $L(X, Y)$, the space of bounded linear maps from X to Y is a vector space. In addition, we may equip it with the *operator norm*, defined by

$$\|T\| = \sup\{\|Tx\| : \|x\| = 1\} = \sup\left\{\frac{\|Tx\|}{\|x\|} : x \neq 0\right\}.$$

For linear maps on finite dimensional spaces, continuity holds. In general, however, there exist discontinuous linear maps on infinite dimensional spaces. The following result gives us some necessary and sufficient conditions for continuity.

Lemma 1.2

If X and Y are normed vector spaces and $T : X \to Y$ a linear map, the following are equivalent:

(i) T is continuous.
(ii) T is continuous at 0.
(iii) T is bounded.

Proof. (i) $\implies$ (ii): Trivial.

(ii) $\implies$ (iii): Let $x \in X$. For a sufficiently small ball $B = B_\delta(0)$, the image $T(B)$ must be bounded, say by M. Let λ be such that $x = \lambda x_0$, where $\|x_0\| = \delta/2$. Then we have

$$\frac{\|Tx\|}{\|x\|} = \frac{\|Tx_0\|}{\|x_0\|} \leq \frac{2M}{\delta}.$$

Hence T is bounded.

(iii) $\implies$ (i): Suppose $\|Tx\| \leq C\|x\|$ for all $x \in X$. Then note that T is Lipschitz with constant C. □

Theorem 1.3

If X is a normed vector space and Y is a Banach space, then $L(X, Y)$ is a Banach space.

Proof. Let $(T_n) \subset L(X, Y)$ be a Cauchy sequence. Given $x \in X$, note that $(T_n x)$ is Cauchy, since

$$\|T_n x - T_m x\| \leq \|T_n - T_m\|\|x\|,$$

which vanishes as $m, n \to \infty$. Since Y is complete, we can then define a linear map $T : X \to Y$ given by

$$Tx = \lim_n T_n x.$$

It remains to show that T is bounded, and $\|T_n - T\| \to 0$. Note that $\|T_n x\| \to \|Tx\|$, and that there exists a uniform bound $\|T_n\| < M$, so we have

$$\|Tx\| = \lim_n \|T_n x\| \leq \lim_n M\|x\| = M\|x\|,$$

so $T \in L(X, Y)$. To show convergence, let $\epsilon > 0$ and pick N large so that $n, m > N$ implies $\|T_n - T_m\| < \epsilon$. Then, since $\|\cdot\|$ is continuous, we can take $m \to \infty$ to get

$$\|T_n - T\| = \lim_{m\to\infty} \|T_n - T_m\| < \epsilon.$$

It follows that $\lim_{n\to\infty} \|T_n - T\| = 0$, as desired. $\square$

1.1 Linear Functionals

We turn out attention to the *dual space* $X^* = L(X, \mathbb{F})$ of bounded linear functionals from X to the underlying field $\mathbb{F}$. Note that it is a Banach space when equipped with the operator norm.

The goal of this section is to show that there exist nonzero bounded linear functionals for any normed vector space. The tool for doing this is the Hahn-Banach theorem. We define a *sublinear functional* on X as a map $p : X \to \mathbb{R}$ such that

- $p(x + y) \leq p(x) + p(y)$ for all $x, y \in X$.
- $p(\lambda x) = \lambda p(x)$ for all $x \in X$, $\lambda \geq 0$.

Theorem 1.4 (Hahn-Banach, Real Version)

Let X be a real vector space, p a sublinear functional on X, M a subspace of X, and f a linear functional on M. Suppose $f \leq p$ on M. Then there exists a linear functional F on X such that $F \leq p$ on X and $F|_M = f$.

Proof. First, we show that given any $x \in X \setminus M$, we can extend f to a linear functional g on $M + \mathbb{R}x$ satisfying $g \leq p$. Once this is shown, we can use a Zorn's lemma argument to obtain a maximal extension.

Define $g : M + \mathbb{R}x \to \mathbb{R}$ by

$$g(y + \lambda x) = f(y) + \lambda\alpha,$$

where $y \in M$, $\lambda \in \mathbb{R}$, and $\alpha \in \mathbb{R}$ is some constant to be determined later. Clearly g is linear, and $g|_M = f$. Now, we want for $\lambda \neq 0$ to have $g \leq p$. We have two cases:

- Case 1: $\lambda > 0$. Then we'd require

  $$\lambda[f(y/\lambda) + \alpha] = g(y + \lambda x) \leq p(y + \lambda x) = \lambda[p(x + y/\lambda)].$$

  This condition becomes

  $$\alpha \leq p(x + y/\lambda) - f(y/\lambda). \tag{1}$$

- Case 2: $\lambda = -\mu < 0$. Then we'd require

  $$\mu[f(y/\mu) - \alpha] = g(y + \lambda x) \le p(y + \lambda x) = \mu[p(y/\mu - x)].$$

  This condition becomes

  $$\alpha \ge f(y/\mu) - p(y/\mu - x). \tag{2}$$

So combining (1) and (2), we must find an α satisfying

$$f(y/\mu) - p(y/\mu - x) \le \alpha \le p(x + y/\lambda) - f(y/\lambda).$$

Since y is more or less arbitrary, it suffices to show

$$\sup_{y \in M}\{f(y) - p(y - x)\} \le \inf_{y \in M}\{p(x + y) - f(y)\}. \tag{3}$$

Let $y_1, y_2 \in M$. Then we have

$$f(y_1) + f(y_2) = f(y_1 + y_2) \le p(y_1 + y_2) \le p(y_1 - x) + p(x + y_2),$$

which upon rearranging becomes

$$f(y_1) - p(y_1 - x) \le p(x + y_2) - f(y_2).$$

This verifies (3), and thus we can pick α to be any real that fits in the gap of (3).

Now we complete the proof with a Zorn's lemma argument. Note that our above argument can be applied to any extension g satisfying $g \le p$ on some extended domain $M' \supset M$. Therefore, the domain of a maximal linear extension satisfying $F \le p$ must be the whole space X. To find this maximal element, first regard the maps from subspaces of X to $\mathbb{R}$ as subsets of $X \times \mathbb{R}$, and note that the family $\mathcal{F}$ of linear extensions F of f satisfying $F \le p$ is partially ordered by inclusion. Since the union of a nested chain of subspaces of X is still a subspace, we see that every chain has an upper bound in $\mathcal{F}$. By Zorn's lemma, there exists a maximal element, concluding the proof. □

Note that all seminorms (and hence norms as well) are sublinear functionals. Then it is not too hard to extend Hahn-Banach to the following more general version.

Theorem 1.5 (Hahn-Banach, Complex Version)

Let X be a complex vector space, p a seminorm on X, M a subspace of X, and f a complex linear functional on M such that $|f| \le p$ on M. Then there exists a complex linear functional $F : X \to \mathbb{C}$ such that $|F| \le p$ on X and $F|_M = f$.

Here are a few applications of Hahn-Banach.

Corollary 1.6

Let X be a normed vector space.

(a) If M is a closed subspace of X and $x_0 \in X \setminus M$, there exists $f \in X^*$ such that $f \not\equiv 0$ and $f|_M \equiv 0$. Furthermore, if we let $\delta = \inf_{x \in M} \|x_0 - x\|$, then f can be taken to satisfy $\|f\| = 1$ and $f(x_0) = \delta$.

(b) If $0 \neq x_0 \in X$, there exists $f \in X^*$ such that $\|f\| = 1$ and $f(x_0) = \|x_0\|$.

(c) The family X^* separates points in X.

(d) If $x \in X$, define $\hat{x} : X^* \to \mathbb{C}$ by $\hat{x}(f) = f(x)$. Then the map $x \mapsto \hat{x}$ is an isometry from X into X^{**}.

Proof. (a): First consider the linear functional f on $M + \mathbb{C}x_0$ given by $f(x + \lambda x_0) = \lambda\delta$, where $x \in M$, $\lambda \in \mathbb{C}$. Then $f(x_0) = \delta$, $f|_M = 0$. Now, for $\lambda \neq 0$, we have

$$|f(x + \lambda x_0)| = |\lambda|\delta \leq |\lambda|\|\lambda^{-1}x + x_0\| = \|x + \lambda x_0\|.$$

Thus we may apply Hahn-Banach with $p(x) = \|x\|$ to the subspace $M + \mathbb{C}x_0$. This gives an extension with the desired properties.

(b): This is simply the case where we take $M = \{0\}$ in (a).

(c): If $x \neq y \in X$, then there exists an $f \in X^*$ such that $f(y - x) = \|y - x\| \neq 0$, which implies $f(x) \neq f(y)$.

(d) Let $x, y \in X$ and $\lambda \in \mathbb{C}$. Then for $f \in X^*$,

$$\widehat{x + \lambda y}(f) = f(x + \lambda y) = f(x) + \lambda f(y) = \hat{x}(f) + \lambda \hat{y}(f),$$

so the map $x \mapsto \hat{x}$ is linear. Suppose $\hat{x} = \hat{y}$. Then for every $f \in X^*$, $f(x) = f(y)$. But since X^* separates points, this implies $x = y$. Thus $x \mapsto \hat{x}$ is injective. Next, we have $|\hat{x}(f)| = |f(x)| \leq \|f\|\|x\|$, so $\|\hat{x}\| \leq \|x\|$. Conversely, (b) implies $\|\hat{x}\| \geq \|x\|$. Thus $\|\hat{x}\| = \|x\|$, so the map $x \mapsto \hat{x}$ is an isometry. $\square$

Exercise 1.7. If X is a normed vector space, show that any finite dimensional subspace is closed.

1.2 Baire Category Theorem & Corollaries

If $E \subset X$, we say E is of the *first category*, or *meager*, if it is a countable union of nowhere dense sets, i.e. sets whose closure has empty interior. Otherwise we say E is of the *second category*.

Theorem 1.8 (The Baire Category Theorem)

Let X be a complete metric space.

(a) If (U_n) is a sequence of open dense subsets of X, then $\cap_n U_n$ is dense in X.

(b) X is not a countable union of nowhere dense sets. That is, X is of the second category.

Proof. For part (a), we need to show that for any nonempty open subset $W \subset X$, the intersection of W with $\cap_n U_n$ is nonempty. Note that $W \cap U_1$ is nonempty, and in particular contains a ball $B_{r_1}(x_1)$. We then construct nested balls $B_{r_n}(x_n)$ contained in $B_{r_{n-1}}(x_{n-1}) \cap U_n$. Note that (x_n) is a Cauchy sequence, and since X is complete, it converges to a limit x. We have for every N,

$$x \in \overline{B_{r_N}(x_N)} \subset U_N \cap B_{r_1}(x_1) \subset U_N \cap W,$$

so x lies in $W \cap (\cap_n U_n)$.

For part (b), let (E_n) be a sequence of nowhere dense sets. Then $(\overline{E}_n^c)$ is a sequence of open dense sets, and by part (a),

$$\cap_n \overline{E}_n^c \neq \varnothing.$$

Thus we have

$$\cup_n E_n \subset \cup \overline{E}_n \neq X,$$

completing the proof. □

Theorem 1.9 (The Open Mapping Theorem)

If X and Y are Banach spaces and $T \in L(X, Y)$ is surjective, then T is an open mapping.

Proof. Being open is equivalent to requiring that if B is a ball centered at any $x \in X$, then $f(B)$ contains a ball centered at $f(x)$. For linear maps, this further reduces to the case where $B = B_1(0)$ and requiring $f(B)$ to contain a ball centered at 0, since

$$T(B_r(x)) = T(x) + rT(B_1(0)).$$

Since T is surjective, we have

$$Y = \cup_{n=1}^{\infty} T(B_n).$$

By Baire's theorem, there exists an n such that $T(B_n)$ is nowhere dense, which implies $T(B_1)$ is nowhere dense, since $y \mapsto ny$ is a homeomorphism of Y. So there exists an $x \in B_1$ such that $Tx \in \operatorname{int}\overline{T(B_1)}$, which implies

$$0 \in \operatorname{int}\overline{T(B_1 - x)} \subset \operatorname{int}\overline{T(B_2)}.$$

Then there is an $r > 0$ such that $\overline{B_0(2r)} \subset \overline{2T(B)}$, or simply $\overline{B_0(r)} \subset \overline{T(B)}$. More generally, this gives us the statement

$$\|y\| < \frac{r}{2^n} \Rightarrow y \in \overline{T(B_{2^{-n}})}.$$

Now, suppose $\|y\| < r/2$. We can find an $x_1 \in B_{2^{-1}}$ such that

$$\|y - Tx_1\| < r/4.$$

Proceeding inductively, we can find $x_n \in B_{2^{-n}}$ such that

$$\|y - \sum_{j=1}^{n} Tx_j\| < r2^{-(n+1)}.$$

Since X is complete, every absolutely convergent series converges, and thus $\sum x_n$ converges, say to x. And in particular,

$$\|x\| < \sum_{n=1}^{\infty} 2^{-n} = 1.$$

It follows that $T(B_1)$ contains all y with $\|y\| < r/2$, so T is open. $\square$

Corollary 1.10 (Bounded Inverses)

If X and Y are Banach spaces and $T \in L(X, Y)$ is bijective, then T is an isomorphism, i.e. $T^{-1} \in L(Y, X)$.

For our next application, we define the *graph* of $T \in L(X, Y)$ to be

$$\Gamma(T) = \{(x, y) \in X \times Y : y = Tx\}.$$

We say the graph is *closed* if $\Gamma(T)$ is a closed subspace of $X \times Y$. If T is continuous, then T is closed. For Banach spaces we have the converse statement:

Theorem 1.11 (The Closed Graph Theorem)

If X and Y are Banach spaces and $T : X \to Y$ is a closed linear map, then T is bounded.

Proof. Let π_1 and π_2 be the projections of $\Gamma(T)$ onto X and Y, respectively. Since $\pi_1 \in L(\gamma(T), X)$ is a bijection, the previous corollary 1.10 tells us that π_1^{-1} is bounded. It follows that $T = \pi_2 \circ \pi_1^{-1}$ is bounded. $\square$

Note the subtle difference between being closed and being continuous. Continuity of a linear map $T : X \to Y$ means that if $x_n \to x$ then $Tx_n \to Tx$, whereas closedness means that if $x_n \to x$ and $Tx_n \to y$ then $y = Tx$. So the power of the Closed Graph theorem is that if we want to show boundedness, or equivalently continuity, we may assume that Tx_n at least converges to something, say y, and all we have to do is show that $Tx = y$. So we can skip the step where we need to show that Tx_n even converges.

We now come to the third application of Baire's theorem, which roughly says that we can derive uniform estimates from pointwise estimates.

Theorem 1.12 (The Uniform Boundedness Principle)
Let X be a Banach space and Y a normed vector space. Suppose $\mathcal{A}$ is a subset of $L(X, Y)$ such that
$$\sup_{T \in \mathcal{A}} \|Tx\| < \infty.$$
Then
$$\sup_{T \in \mathcal{A}} \|T\| < \infty.$$

Proof. Define
$$E_n = \{x \in X : \sup_{T \in \mathcal{A}} \|Tx\| \leq n\}.$$
Then by our hypothesis we have
$$X = \cup_{n=1}^{\infty} E_n.$$
By Baire's theorem, one of the E_n must be nonmeager, and thus contian some ball $B_r(x_0)$. But this implies $B_r(0) \subset E_{2n}$, since for $x \in B_r(0)$, we have
$$\|Tx\| \leq \|T(x_0 - x)\| + \|Tx_0\| \leq 2n.$$
It follows that
$$\sup_{T \in \mathcal{A}} \|T\| < 2n/r < \infty,$$
since $\|Tx\| \leq 2n$ whenever $\|x\| < r$ for all $T \in \mathcal{A}$. $\square$

2 Hilbert Spaces

Much of the content on Hilbert spaces is just linear algebra extended to the infinite dimensional case. As such, we start with a familiar definition. Let X be a complex vector space. An *inner product* is a map $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{C}$ such that:

(i) (bilinearity in first slot) $\langle ax + by, z\rangle = a\langle x, z\rangle + b\langle y, z\rangle$ for all $x, y, z \in X$ and $a, b \in \mathbb{C}$.

(ii) (conjugate symmetry) $\langle x, y\rangle = \overline{\langle y, x\rangle}$ for all $x, y \in X$.

(iii) (positive definiteness) $\langle x, x\rangle \in (0, \infty)$ for all nonzero $x \in X$.

The first order of business is to verify that $\|x\| := \sqrt{\langle x, x\rangle}$ is a norm on X. To do this, we will need a familiar inequality whose proof is different from the one in the Euclidean setting.

Lemma 2.1 (Schwarz Inequality)
For all $x, y \in X$,

$$|\langle x, y\rangle| \leq \|x\|\|y\|,$$

with equality occurring if and only $x = \lambda y$ for some $\lambda \in \mathbb{C}$.

Proof. If $\langle x, y\rangle = 0$, the result is trivial. So assume $\langle x, y\rangle \neq 0$, and that w.l.o.g. $y \neq 0$. Let $\alpha = \arg\langle x, y\rangle$ and $z = \alpha y$, so that $\langle x, z\rangle = \langle z, x\rangle = |\langle x, y\rangle|$ and $\|y\| = \|z\|$. Then for $t \in \mathbb{R}$, we have

$$0 \leq \langle x - tz, x - tz\rangle = \|x\|^2 - 2t|\langle x, y\rangle| + t^2\|y\|^2.$$

Optimizing over t, we see that the quadratic achieves its minimum at $t = \|y\|^{-2}|\langle x, y\rangle|$. Plugging this in, we get

$$0 \leq \|x - tz\|^2 = \|x\|^2 - \|y\|^{-2}|\langle x, y\rangle|^2,$$

with equality iff $x - tz = x - \alpha ty = 0$, completing the proof. ☐

Proposition 2.2
The map $x \mapsto \|x\|$ is a norm on X.

Proof. It's easy to see that $\|x\| = \sqrt{\langle x, x\rangle} = 0$ iff $x = 0$. Furthermore, we have

$$\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x\rangle} = \sqrt{|\lambda|^2\langle x, x\rangle} = |\lambda|\|x\|.$$

Finally, we have by the Schwarz inequality that

$$\begin{aligned}
\|x + y\|^2 &= \langle x + y, x + y\rangle \\
&= \|x\|^2 + 2\mathrm{Re}\langle x, y\rangle + \|y\|^2 \\
&\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\
&= (\|x\| + \|y\|)^2,
\end{aligned}$$

proving the triangle inequality. ☐

An inner product space that is complete with respect to the norm $\|x\| = \sqrt{\langle x, x \rangle}$ is called a *Hilbert space*. We will use $\mathcal{H}$ to denote a Hilbert space.

Proposition 2.3

Inner products are continuous. That is, if $x_n \to x$ and $y_n \to y$, then $\langle x_n, y_n \rangle \to \langle x, y \rangle$.

Proof. Note that by the Schwarz inequality, we have

$$\begin{aligned} |\langle x_n, y_n \rangle - \langle x, y \rangle| &= |\langle x_n, y_n \rangle - \langle x, y_n \rangle + \langle x, y_n \rangle - \langle x, y \rangle| \\ &= |\langle x_n - x, y_n \rangle + \langle x, y_n - y \rangle| \\ &\leq \|x_n - x\| \|y_n\| + \|x\| \|y_n - y\| \to 0, \end{aligned}$$

as desired. □

Proposition 2.4 (Parallelogram law)

For any $x, y \in \mathcal{H}$,

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

Proof. Sum the formulas $\|x \pm y\|^2 = \|x\|^2 \pm 2\mathrm{Re}\langle x, y \rangle + \|y\|^2$. □

2.1 Orthogonality

For $x, y \in \mathcal{H}$, we say x is *orthogonal* to y and write $x \perp y$ if $\langle x, y \rangle = 0$. For any subset $E \subset \mathcal{H}$, we define the orthogonal complement

$$E^\perp = \{x \in \mathcal{H} : \langle x, y \rangle = 0 \text{ for all } y \in E\}.$$

By linearity and continuity of the inner product (2.3), we see that $E^\perp$ is a closed subspace of $\mathcal{H}$.

Theorem 2.5 (Pythagoras)

If $x_1, \ldots, x_n \in \mathcal{H}$ and $x_i \perp x_j$ for $i \neq j$, then

$$\left\| \sum_{i=1}^n x_i \right\|^2 = \sum_{i=1}^n \|x_i\|^2.$$

Proof. Expand the LHS as an inner product. □

Theorem 2.6 (Unique Decomposition)
Let $\mathcal{G}$ be a closed subspace of $\mathcal{H}$. Then

$$\mathcal{H} = \mathcal{G} \oplus \mathcal{G}^{\perp}.$$

In particular, every $x \in \mathcal{H}$ can be decomposed uniquely as

$$x = y + z, \quad y \in \mathcal{G}, z \in \mathcal{G}^{\perp}.$$

Moreover, y and z are the unique elements of $\mathcal{G}$ and $\mathcal{G}^{\perp}$ minimizing their distance to x.

Proof. Fix $x \in \mathcal{H}$. Let $\delta = \inf_{y\in\mathcal{G}} \|x - y\|$. Then there is a sequence $(y_n) \subset \mathcal{G}$ such that $\|x - y_n\| \to \delta$. We wish to show y_n converges to some $y \in \mathcal{G}$. Note that by Parallelogram law,

$$2(\|y_n - x\|^2 + \|y_m - x\|^2) = \|y_n - y_m\|^2 + \|y_n + y_m - 2x\|^2.$$

So since $\frac{y_n+y_m}{2} \in \mathcal{G}$, we have

$$\|y_n - y_m\|^2 \leq 2\|y_n - x\|^2 + 2\|y_m - x\|^2 + 4\delta^2 \to 0,$$

which implies that (y_n) is Cauchy, and thus converges. Since $\mathcal{G}$ is closed, say $y_n \to y \in \mathcal{G}$.

Now, we wish to show $z = x - y \in \mathcal{G}^{\perp}$. Let $u \in \mathcal{G}$. W.l.o.g. we can let $\langle z, u\rangle$ be real, so that

$$f(t) = \|z - tu\|^2 = \|z\|^2 - 2t\langle z, u\rangle + t^2\|u\|^2$$

is a real-valued function for $t \in \mathbb{R}$. Since $z - tu = x - (y + tu)$, we see that f achieves its minimum value of δ^2 at $t = 0$. So in particular

$$-2\langle z, u\rangle = f'(0) = 0,$$

which implies $z \in \mathcal{G}^{\perp}$.

Suppose there were a $y' \in \mathcal{G}$ also minimizing its distance to x. Then by Pythagoras, we have

$$\|x - y'\|^2 = \|x - y\|^2 + \|y - y'\|^2 \geq \|x - y\|^2,$$

with equality iff $y = y'$. The same argument can be applied to $z \in \mathcal{G}^{\perp}$, so these are unique.

Finally suppose there were another decomposition $x = y' + z'$. Then we'd have

$$y' - y = z - z' \in \mathcal{G} \cap \mathcal{G} \wedge \perp,$$

which forces $y' = y$ and $z' = z$, so the decomposition is unique as well. □

Exercise 2.7. Adapt the proof above to show that a close convex subset of a Hilbert space has a unique element with minimal norm.

We now talk about the Reisz Representation theorem. For each $y \in \mathcal{H}$, consider the function

$$f_y(x) = \langle x, y \rangle.$$

Note that this defines a bounded linear functional on $\mathcal{H}$ with $\|f_y\| = \|y\|$.

Theorem 2.8 (Riesz Representation Theorem)
If $f \in \mathcal{H}^*$, there is a unique $y \in \mathcal{H}$ such that

$$f(x) = \langle x, y \rangle$$

for all $x \in \mathcal{H}$.

Proof. For uniqueness, suppose there exist y, y' such that $f(x) = \langle x, y \rangle = \langle x, y' \rangle$ for all x. Then by setting $x = y - y'$, we get $\|y - y'\|^2 = 0$, which implies $y = y'$.

Existence is trickier. Note that if $f \equiv 0$, then $y = 0$ necessarily, so assume otherwise. Let

$$\mathcal{G} = \{x \in \mathcal{H} : f(x) = 0\}$$

be the kernel of f. Then $\mathcal{G}$ is a proper closed subspace of $\mathcal{H}$, so $\mathcal{G}^\perp \neq \{0\}$. Pick $z \in \mathcal{G}^\perp$ with $\|z\| = 1$, and take $y = \overline{f(z)}z$. Note that

$$f(x)z - f(z)x \in \mathcal{H}$$

for all x, so we have

$$0 = \langle f(x)z - f(z)x, z \rangle = f(x)\|z\|^2 - f(z)\langle x, z \rangle = f(x) - \langle x, y \rangle.$$

It follows that $f(x) = \langle x, y \rangle$. $\square$

One of the implications of this result is that the map

$$y \mapsto f_y$$

is a conjugate-linear isometry between $\mathcal{H}$ and its dual $\mathcal{H}^*$. In particular, Hilbert spaces are reflexive in that $\mathcal{H}$, $\mathcal{H}^*$, and $\mathcal{H}^{**}$ are all naturally isomorphic.

2.2 Orthonormal Bases

Theorem 2.9 (Bessel's Inequality)
If $\{u_\alpha\}_{\alpha \in A}$ is an orthonormal set in $\mathcal{H}$, then for any $x \in \mathcal{H}$,

$$\sum_{\alpha \in A} |\langle x, u_\alpha \rangle \wedge 2 \leq \|x\|^2.$$

Note that this implies that the set $\{\alpha : \langle x, u_\alpha \rangle \neq 0\}$ is countable.

Proof. Let $F \subset A$ be any finite subset. Note that

$$x - \sum_{\alpha \in F} \langle x, u_\alpha \rangle u_\alpha, \quad \sum_{\alpha \in F} \langle x, u_\alpha \rangle u_\alpha$$

are orthogonal, so by Pythgoras, we get

$$\left\| x - \sum_{\alpha \in F} \langle x, u_\alpha \rangle u_\alpha \right\|^2 + \sum_{\alpha \in F} |\langle x, u_\alpha \rangle|^2 = \|x\|^2.$$

Rearranging, this tells us

$$\|x\|^2 - \sum_{\alpha \in F} |\langle x, u_\alpha \rangle|^2 \geq 0,$$

as desired. □

For infinite dimensional spaces, we say an orthonormal set $\{u_\alpha\}_{\alpha \in A}$ is a *basis* of $\mathcal{H}$ if finite linear combinations are dense in $\mathcal{H}$. The following theorem provides some useful equivalent conditions for being a basis.

Theorem 2.10

Let $\{u_\alpha\}_{\alpha \in A}$ be an orthonormal set in $\mathcal{H}$. Then the following are equivalent conditions for $\{u_\alpha\}_{\alpha \in A}$ to be an orthonormal basis:

(a) The only element of $\mathcal{H}$ orthogonal to the whole family $\{u_\alpha\}_{\alpha \in A}$ is 0.

(b) Parseval's Identity: For every $x \in \mathcal{H}$,

$$\|x\|^2 = \sum_{\alpha \in A} |\langle x, u_\alpha \rangle|^2.$$

(c) For every $x \in \mathcal{H}$, we can write

$$x = \sum_{\alpha \in A} \langle x, u_\alpha \rangle u_\alpha.$$

Proof. Let α_j be an enumeration of the α's with $\langle x, u_{\alpha_j} \rangle \neq 0$.

(a) $\implies$ (c): First we show that the series converges to something. We can abuse completeness of $\mathcal{H}$. By Bessel's inequality, the series $\sum_{j=1}^n |\langle x, u_{\alpha_j} \rangle|^2$ converges. So by Pythagoras we have

$$\left\| \sum_{j=n}^m \langle x, u_{\alpha_j} \rangle u_{\alpha_j} \right\|^2 = \sum_{j=n}^m |\langle x, u_{\alpha_j} \rangle|^2 \to 0.$$

Due to completeness, the series $\sum \langle x, u_{\alpha_j} \rangle u_{\alpha_j}$ converges. Note that

$$\left\langle x - \sum_{j=1}^\infty \langle x, u_{\alpha_j} \rangle u_{\alpha_j}, u_{\alpha_k} \right\rangle = 0$$

for every k, from which we deduce that $x = \sum \langle x, u_\alpha \rangle u_\alpha$.

(c) $\implies$ (b): By Pythagoras, we have

$$\|x\|^2 - \sum_{j=1}^{n} |\langle x, u_\alpha \rangle|^2 = \left\| x - \sum_{j=1}^{n} \langle x, u_{\alpha_j} u_{\alpha_j} \right\|^2 \to 0$$

as $n \to \infty$.

(b) $\implies$ (a): Immediate. □

Theorem 2.11 (Existence)

Every Hilbert space has an orthonormal basis.

Proof. Consider the collection of orthonormal sets with the partial ordering of set inclusion. Any chain has an obvious upper bound, i.e. the union of its elements is another orthonormal set. Therefore by Zorn's lemma, there exists a maximal element. Proof by contradiction using (a) from 2.10 tells us that this orthonormal set is a basis. □

Theorem 2.12 (Countability)

A Hilbert space $\mathcal{H}$ has a countable orthonormal basis iff it is separable.

Proof. Let $\{x_n\}$ be a countable dense subbset of $\mathcal{H}$. Discard linearly dependent vectors and apply Gram-Schmidt to obtain a countable orthonormal basis.

Conversely, let $\{u_n\}$ be a countable orthonormal basis. Then finite linear combinations of the u_n's with coefficients in $\mathbb{Q} + i\mathbb{Q}$ is a countable dense subset in $\mathcal{H}$. □

Another important fact is that every Hilbert space looks like an ℓ^2 space. In particular, let $\{u_\alpha\}_{\alpha \in A}$ be an orthonormal basis of $\mathcal{H}$. Then the Fourier transform $x \mapsto \hat{x}$ given by

$$\hat{x}(\alpha) = \langle x, u_\alpha \rangle$$

is a unitary map from $\mathcal{H}$ to $\ell^2(A)$. This can be shown with Parseval's identity and the fact that a linear map is unitary iff it is isometric and surjective.

It also turns out that $\ell^2(A)$ is complete, i.e. it is a Hilbert space. So the transform above can be seen as an isomorphism in the category of Hilbert spaces.

Theorem 2.13

For any nonempty A, the space $\ell^2(A)$ is complete.

Proof. Let $(f_n) \subset \ell^2(A)$ be a Cauchy sequence. Then for any $\epsilon > 0$, and m, n sufficiently large,

$$\sum_{\alpha \in A} |f_n(\alpha) - f_m(\alpha)|^2 < \epsilon.$$

So in particular for each $\alpha \in A$, the sequence $\{f_n(\alpha)\}_{n=1}^{\infty}$ is Cauchy, and thus convergent to some $f(\alpha)$. Now we just need to show that $f_n \to f$ in norm. To see that $f \in \ell^2(A)$, note that $f_n \to f$ uniformly, so we can write

$$\sum_{\alpha \in A} |f(\alpha)|^2 = \sum_{\alpha \in A} \lim_n |f_n(\alpha)|^2 = \lim_{n\to\infty} \sum_{\alpha\in A} |f_n(\alpha)|^2 = \lim_{n\to\infty} \|x_n\|^2 < \infty.$$

Next, with

$$|f(\alpha) - f_n(\alpha)|^2 \leq |f(\alpha)|^2 + |f_n(\alpha)|^2,$$

we may use Dominated Convergence to conclude that $f_n \to f$ in norm. □

2.3 Adjoints

Let V and W be Hilbert spaces and $T \in L(V, W)$. Then the *adjoint* of T is the map $T^* : W \to V$ such that

$$\langle Tf, g\rangle = \langle f, T^*g\rangle$$

for all $f \in V$ and $g \in W$.

Exercise 2.14. Show that the adjoint exists and is unique, using Theorem 2.8. Then check that T^* is a bounded linear map.

Exercise 2.15. All maps are linear and bounded. Verify the following properties:

(i) $(T^*)^* = T$.
(ii) $\|T^*\| = \|T\|$.
(iii) $(S + T)^* = S^* + T^*$.
(iv) $(\lambda T)^* = \overline{\lambda} T^*$ for all $\lambda \in \mathbb{F}$.
(v) $I^* = I$.
(vi) $(S \circ T)^* = T^* \circ S^*$.

Example 2.16 (Multiplication Operator)**.** Let $h \in L^\infty(\mu)$ and define the multiplication operator $M_h : L^2(\mu) \to L^2(\mu)$ by

$$M_h f = fg.$$

Then M_h is a bounded linear map with adjoint

$$M_h^* = M_{\overline{h}}.$$

Example 2.17 (Integral Operator)**.** Let $(X, \mathcal{F}, \mu)$ and $(Y, \mathcal{G}, \nu)$ be σ-finite measure spaces and $K \in L^2(\mu \times \nu)$. Define the integral operator $\mathcal{I}_K : L^2(\nu) \to L^2(\mu)$ by

$$(\mathcal{I}_K f)(x) = \int_Y K(x, y) f(y) d\nu,$$

where $f \in L^2(\nu)$ and $x \in X$. We need to check that $\mathcal{I}_K$ is bounded and then compute its adjoint. By the Schwarz inequality, we have

$$\int_Y |K(x, y)||f(y)| d\nu \leq \left(\int_Y |K(x, y)|^2 d\nu\right)^{1/2} \|f\|_{L^2(\nu)}. \tag{4}$$

Then we have

$$\begin{aligned}\|\mathcal{I}_K f\|_{L^2(\mu)}^2 &= \int_X |(\mathcal{I}_K f)(x)|^2 d\mu(x)\\ &\le \int_X \left(\int_Y |K(x,y)||f(y)|d\nu\right)^2 d\mu\\ &\le \left(\int_X \int_Y |K(x,y)|^2 d\nu d\mu\right) \|f\|_{L^2(\nu)}^2\\ &= \|K\|_{L^2(\mu\times\nu)}^2 \|f\|_{L^2(\nu)}^2,\end{aligned}$$

where the third line is a result of (4) and the final line follows from Fubini's theorem. Thus

$$\|\mathcal{I}_K\| \le \|K\|_{L^2(\mu\times\nu)}$$

so $\mathcal{I}_K$ is bounded.

Now, consider the map $K^* : Y \times X \to \mathbb{F}$ given by

$$K^*(y,x) = \overline{K(x,y)}.$$

It's not too hard to verify that this determines the adjoint of $\mathcal{I}_K$. That is,

$$(\mathcal{I}_K)^* = \mathcal{I}_{K^*}.$$

Example 2.18 (Matrix Operator)**.** As a special case of the previous example, let's consider a finite setting where μ is the counting measure on $X = [m]$ and ν is the counting measure on $Y = [n]$. Then K is an $m \times n$ matrix and the operator $\mathcal{I}_K : \mathbb{F}^n \to \mathbb{F}^m$ is just the operation of multiplication by matrix K. In particular,

$$(\mathcal{I}_K f)(i) = \sum_{j=1}^{n} K(i,j) f(j)$$

is just the formula for matrix multiplication. Finally, note that K^* is just the conjugate transpose of K.

We have a result that is essentially the fundamental theorem of linear algebra, but for Hilbert spaces.

Theorem 2.19

Suppose V and W are Hilbert spaces and $T \in L(V,W)$. Then

(a) $\ker T^* = (\operatorname{ran} T)^\perp$.
(b) $\overline{\operatorname{ran} T^*} = (\ker T)^\perp$.
(c) $\ker T = (\operatorname{ran} T^*)^\perp$.
(d) $\overline{\operatorname{ran} T} = (\ker T^*)^\perp$.

Proof. It suffices to show (a). Suppose $g \in W$. Then

$$\begin{aligned} g \in \ker T^* &\iff T^*g = 0 \\ &\iff \langle f, T^*g \rangle = 0 \text{ for all } f \in V \\ &\iff \langle Tf, g \rangle = 0 \text{ for all } f \in V \\ &\iff g \in (\operatorname{ran} T)^\perp, \end{aligned}$$

so $\ker T^* = (\operatorname{ran} T)^\perp$. The rest follows by taking adjoints and orthogonal complements. □

Corollary 2.20

Let T be as above. Then T has dense range iff T^* is injective.

Proof. From 2.19(d) we have that T has dense range iff $(\ker T^*)^\perp = W$ iff $\ker T^* = \{0\}$ iff T^* is injective. □

Example 2.21 (Volterra Operator)**.** Consider the linear map $\mathcal{V} : L^2[0,1] \to L^2[0,1]$ given by

$$(\mathcal{V}f)(x) = \int_0^x f(y)dy$$

for $f \in L^2[0,1]$ and $x \in [0,1]$. It can be shown that the adjoint is given by

$$(\mathcal{V}^*f)(x) = \int_x^1 f(y)dy = \int_0^1 f(y)dy - \int_0^x f(y)dy.$$

We will now show that $\mathcal{V}$ has dense range, so let $f \in L^2[0,1]$ such that $\mathcal{V}^*f = 0$. Differentiating both sides of the above, we deduce that $f = 0$, and so $\mathcal{V}^*$ is injective. Thus the Volterra operator $\mathcal{V}$ has dense range. In particular, note that the range gives us an example of an infinite dimensional subspace which is not closed, since every element of $\operatorname{ran}\mathcal{V}$ is a continuous function that vanishes at 0.

An *operator* is a linear map from a vector space to itself. For a vector space V, we denote $L(V) := L(V,V)$ as the normed vector space of bounded operators on V. With operators it is possible to talk about inverses.

Theorem 2.22

An operator $T \in L(V)$ is invertible iff T^* is invertible. If this occurs, then

$$(T^*)^{-1} = (T^{-1})^*.$$

Proof. Suppose T is invertible. Then

$$T^*(T^{-1})^* = (T^{-1}T)^* = I^* = I,$$

and likewise $(T^{-1})^*T^* = I$, so T^* is invertible with inverse $(T^{-1})^*$.

Conversely, suppose T^* is invertible. Since $(T^*)^* = T$, the argument above works, implying that T is invertible and that $(T^*)^{-1} = (T^{-1})^*$. □

Next we have a useful result that says the formula for a geometric series works for operators on Banach spaces.

Lemma 2.23

If $T \in L(V)$ and V is a Banach space, with $\|T\| < 1$, then $I - T$ is invertible and

$$(I - T)^{-1} = \sum_{k=0}^{\infty} T^k.$$

Proof. Since V is Banach, so is $L(V)$. Then to show that the series converges, it suffices to show that it converges absolutely. Indeed, we have

$$\sum_{k=0}^{\infty} \|T^k\| \leq \sum_{k=0}^{\infty} \|T\|^k = \frac{1}{1 - \|T\|} < \infty.$$

Therefore we may write

$$(I - T) \sum_{k=0}^{\infty} = \lim_{n\to\infty} (I - T) \sum_{k=0}^{n} T^k = \lim_{n\to\infty} (I - T^{n+1}) = I.$$

The last equality holds because $\|T\| < 1$ and so $\|T^{n+1}\| \leq \|T\|^{n+1} \to 0$. Similarly, we have

$$\left(\sum_{k=0}^{\infty} T^k \right) (I - T) = I.$$

Thus $I - T$ is invertible with the desired inverse. □

Theorem 2.24

Suppose V is a Banach space. Then $\{T \in L(V) : T \text{ is invertible}\}$ is an open subset of $L(V)$.

Proof. Suppose $T \in L(V)$ is invertible. We wish to find a sufficiently small ball about T that contains only invertible maps. For $S \in L(V)$ to be invertible, it suffices to make $T^{-1}S$ invertible. From lemma 2.23 it suffices to show that $\|I - T^{-1}S\| < 1$. Since

$$\|I - T^{-1}S\| \leq \|T^{-1}\|\|T - S\|,$$

and we know $\|T^{-1}\|$ exists by the open mapping theorem, we can take $S \in L(V)$ such that

$$\|T - S\| < \frac{1}{\|T^{-1}\|}.$$

As all such S are invertible, it follows that the subset of invertible maps is open in $L(V)$. □

2.4 Spectrum

Let $T \in L(V)$ where V is a Banach space. We say $\alpha \in \mathbb{F}$ is an *eigenvalue* of T if $T - \lambda I$ is not injective, and a nonzero vector $v \in V$ is an *eigenvector* of T corresponding to the eigenvalue λ if $Tv = \lambda v$. The *spectrum* of T, denoted $\sigma(T)$, is defined as

$$\sigma(T) = \{\lambda \in \mathbb{F} : T - \lambda I \text{ is not invertible}\}.$$

Note that the set of eigenvalues is contained in the spectrum of T. For finite dimensional Banach spaces V, the spectrum of T equals the set of eigenvalues of T, since injectivity is equivalent to invertibility.

Exercise 2.25. Verify the set of eigenvalues and spectrum for each of the following examples:

- Let (b_n) be a bounded sequence in $\mathbb{F}$. Consider the bounded linear map $T : \ell^2 \to \ell^2$ given by
$$T(a_1, a_2, \ldots) = (a_1 b_1, a_2, b_2, \ldots).$$
Then the set of eigenvalues equals $\{b_k : k \in \mathbb{Z}^+\}$ and the spectrum equals the closure $\overline{\{b_k : k \in \mathbb{Z}^+\}}$.
- Let $h \in L^\infty(\mathbb{R})$. Consider the multiplication operator $M_h : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ given by
$$M_h f = hf.$$
Then $\lambda \in \mathbb{F}$ is an eigenvalue of M_h iff $\mu\{t \in \mathbb{R} : h(t) = \lambda\} > 0$. Also, $\lambda \in \sigma(M_h)$ iff $\mu\{t \in \mathbb{R} : |h(t) - \lambda| < \epsilon\} > 0$ for all $\epsilon > 0$. (*hint:* use corollary 1.10)

Note that for an eigenvalue λ and corresponding eigenvector f of T,

$$\|Tf\| = \|\lambda f\| = |\lambda|\|f\|,$$

which implies that $|\lambda| \leq \|T\|$. It turns out that the same is true for elements of the spectrum of T. In particular, suppose $|\lambda| > \|T\|$. Then by lemma 2.23 the operator

$$T - \lambda I = -\lambda(I - T/\lambda)$$

is invertible since $\|T/\lambda\| < 1$. Furthermore, for $|\lambda| > \|T\|$, we have

$$\begin{aligned}
\|(T - \lambda I)^{-1}\| &\leq \frac{1}{|\lambda|} \sum_{k=0}^{\infty} \frac{\|T\|^k}{|\lambda|^k} \\
&= \frac{1}{|\lambda|} \frac{1}{1 - \frac{\|T\|}{|\lambda|}} \\
&= \frac{1}{|\lambda| - \|T\|}.
\end{aligned}$$

Therefore,

$$\lim_{|\lambda| \to \infty} \|(T - \lambda I)^{-1}\| = 0. \tag{5}$$

Theorem 2.26

The spectrum of a bounded operator on a Banach space is a closed subset of $\mathbb{F}$.

Proof. Suppose $\lambda_n \to \lambda \in \mathbb{F}$, where each $\lambda_n \in \sigma(T)$. Then

$$\lim_{n\to\infty}(T - \lambda_n I) = T - \lambda I,$$

where convergence is with respect to the operator norm. Since the set of invertible elements of $L(V)$ is open, the complement is closed. So, with all the $T - \lambda_n I$ being non invertible, this implies that $T - \lambda I$ is non invertible as well. Then $\lambda \in \sigma(T)$ and so the spectrum is closed in $L(V)$. □

Recall that in finite dimensional linear algebra, every operator on a complex vector space has an eigenvalue. This result does not extend to general Hilbert spaces though, for consider the right shift $R : \ell^2 \to \ell^2$ given by

$$R(a_1, a_2, \ldots) = (0, a_1, a_2, \ldots).$$

It's evident that R has no eigenvalues.

However, the result does extend when we use instead the spectrum as a substitute. The proof of this result will use a bit of complex analysis. First we prove the result for complex Hilbert spaces, then leave the extension to complex Banach spaces as an exercise.

Lemma 2.27

Let T be a bounded operator on a complex Hilbert space V. Then the map

$$\varphi : \lambda \mapsto \langle (T - \lambda I)^{-1} f, g\rangle$$

is holomorphic on $\mathbb{C} \setminus \sigma(T)$ for every $f, g \in V$.

Proof. We will show that φ has a power series expansion about $\alpha \in \mathbb{C} \setminus \sigma(T)$ for a sufficiently small radius. In particular, let $\beta \in \mathbb{C}$ be such that

$$|\alpha - \beta| < \frac{1}{\|(T - \alpha I)^{-1}\|}.$$

By lemma 2.23, the operator $I - (\alpha - \beta)(T - \alpha I)^{-1}$ is invertible, with

$$(I - (\alpha - \beta)(T - \alpha I)^{-1})^{-1} = \sum_{k=0}^{\infty} (\alpha - \beta)^k ((T - \alpha I)^{-1})^k.$$

Multiplying by $(T - \alpha I)^{-1}$ on both sides, we get

$$(T - \beta I)^{-1} = \sum_{k=0}^{\infty} (\alpha - \beta)^k ((T - \alpha I)^{-1})^{k+1}.$$

Since inner products commute with summations, for all $f, g \in V$ we have

$$\left\langle (T-\beta I)^{-1} f, g \right\rangle = \sum_{k=0}^{\infty} \left\langle ((T-\alpha I)^{-1})^{k+1} f, g \right\rangle (\alpha-\beta)^k.$$

Thus φ has a power series expansion about α. ☐

Theorem 2.28

The spectrum of $T \in L(V)$, where V is a complex Hilbert space, is a nonempty subset of $\mathbb{C}$.

Proof. Suppose $\sigma(T) = \varnothing$. Let $0 \neq f \in V$. Then the function

$$\phi : \lambda \mapsto \langle (T-\lambda I)^{-1} f, T^{-1} f \rangle$$

is holomorphic on $\mathbb{C}$. By (5), we can deduce that this map converges uniformly to 0 as $|\lambda| \to \infty$. Recall the mean value property, which follows from Cauchy's integral formula,

$$\begin{aligned} f(z_0) &= \frac{1}{2\pi i} \int_{C_r} \frac{f(z)}{z-z_0} dz \\ &= \frac{1}{2\pi i} \int_0^{2\pi} \frac{z_0 + re^{it}}{re^{it}} ire^{it} dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + re^{it}) dt. \end{aligned}$$

We can exchange the limit with integral due to uniform convergence to deduce that $\varphi(0) = 0$. But then

$$\langle T^{-1} f, T^{-1} f \rangle = 0,$$

which implies $f = 0$, a contradiction. Thus $\sigma(T) \neq \varnothing$. ☐

Exercise 2.29. Let T be a bounded operator on a complex Banach space V.

(a) Prove that the map

$$\lambda \mapsto \phi((T-\lambda I)^{-1} f)$$

is analytic on $\mathbb{C} \setminus \sigma(T)$ for every $f \in V$ and $\phi \in V^*$.

(b) Prove that $\sigma(T) \neq \varnothing$.

Exercise 2.30 (Spectral Mapping Theorem)**.** Let T be a bounded operator on a complex Banach space, and $p \in \mathbb{C}[z]$. Show that

$$\sigma(p(T)) = p(\sigma(T)).$$

2.5 Special Classes of Bounded Operators

In this section, we cover self-adjoint operators, normal operators, isometries, unitary operators, and compact operators in $L(V)$.

2.5.1 Self-Adjoint Operators

We say a bounded operator T on a Hilbert space is *self-adjoint* if $T^* = T$. That is, we require that

$$\langle Tf, g\rangle = \langle f, Tg\rangle$$

for all $f, g \in V$.

Exercise 2.31. Verify each of the following examples.

- Recall the bounded operator given by

  $$T(a_1, a_2, \ldots) = (a_1 b_1, a_2 b_2, \ldots),$$

  where (b_n) is bounded. Then the adjoint $T^* : \ell^2 \to \ell^2$ is given by

  $$T^*(a_1, a_2, \ldots) = (a_1 \overline{b_1}, a_2 \overline{b_2}, \ldots),$$

  and T is self-adjoint iff $b_n \in \mathbb{R}$ for all n.
- As a generalization of the previous example, recall the multiplication operator M_h, with adjoint $M_h^* = M_{\overline{h}}$. Then M_h is self-adjoint iff h is real almost everywhere.
- Recall the operator of matrix multiplication in example 2.18. It is self-adjoint iff $K = \overline{K}^\top$.
- As a generalization of the previous example, recall the integral operator $\mathcal{I}_K$ in example 2.17. It is self-adjoint iff $K(x,y) = \overline{K(y,x)}$ for every (x,y).
- Let P_U denote the orthogonal projection of a Hilbert space V onto a closed subspace U.

  $$\begin{aligned}\langle P_U f, g\rangle &= \langle P_U f, P_U g + (I - P_U)g\rangle \\ &= \langle P_U f, P_U g\rangle \\ &= \langle f - (I - P_U)f, P_U g\rangle \\ &= \langle f, P_U g\rangle.\end{aligned}$$

  Hence P_U is self-adjoint.

Theorem 2.32

If $T \in L(V)$ where V is a Hilbert space, and $\langle Tf, f\rangle = 0$ for all $f \in V$, then

(i) If $\mathbb{F} = \mathbb{C}$, then $T = 0$.

(ii) If $\mathbb{F} = \mathbb{R}$ and T is self-adjoint, then $T = 0$.

Proof. Write

$$\begin{aligned}\langle Tg, h\rangle &= \frac{\langle T(g+h), g+h\rangle - \langle T(g-h), g-h\rangle}{4} \\ &\quad + \frac{\langle T(g+ih), g+ih\rangle - \langle T(g-ih), g-ih\rangle}{4} i \\ &= 0,\end{aligned}$$

and take $h = Tg$ to prove (a).

For (b), we instead have

$$\langle Tg, h \rangle = \frac{\langle T(g+h), g+h \rangle - \langle T(g-h), g-h \rangle}{4} = 0.$$

Taking $h = Tg$ implies $T = 0$. □

For intuition, we can think of the adjoint operation $T \mapsto T^*$ on $L(V)$ as akin to the conjugation operation $z \mapsto \overline{z}$ on $\mathbb{C}$. Then self-adjoint operators are akin to real numbers. For example, recall the multiplication operator, which is self-adjoint iff it is real almost everywhere. The next two results further reflect this philosophy.

Theorem 2.33

Let $T \in L(V)$. Then T is self-adjoint iff $\langle Tf, f \rangle \in \mathbb{R}$ for every $f \in V$.

Proof. For $f \in V$, we have

$$\langle Tf, f \rangle - \overline{\langle Tf, f \rangle} = \langle Tf, f \rangle - \langle f, Tf \rangle = \langle (T - T^*)f, f \rangle.$$

Thus $\langle Tf, f \rangle$ iff $\langle (T - T^*)f, f \rangle = 0$, which happens iff $T - T^* = 0$, i.e. T is self-adjoint. □

Theorem 2.34

Let $T \in L(V)$. Then $\sigma(T) \subset \mathbb{R}$.

Proof. See Axler. □

2.5.2 Normal Operators

We say a bounded operator T on a Hilbert space is *normal* if

$$T^*T = TT^*.$$

Every self-adjoint operator is normal, but there exist normal operators which are not self-adjoint.

Exercise 2.35. Verify each of the following examples:

- The multiplication operator M_h is normal for every h, and is only self-adjoint when h is $\mathbb{R}$-valued.
- The matrix
$$\begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}$$
is not self-adjoint, but is normal.
- The right-shift on ℓ^2 is not normal.

Theorem 2.36
Let $T \in L(V)$, where V is a Hilbert space. Then T is normal iff

$$\|Tf\| = \|T^* f\|$$

for all $f \in V$.

Proof. For $f \in V$ we have

$$\|Tf\|^2 - \|T^* f\|^2 = \langle Tf, Tf\rangle - \langle T^* f, T^* f\rangle = \langle (T^*T - TT^*)f, f\rangle.$$

The result follows. $\square$

Theorem 2.37
Let $T \in L(V)$, where V is a complex Hilbert space.

(a) There exist unique real and imaginary "parts", i.e. self adjoint operators A and B, such that $T = A + iB$.
(b) T is normal iff $AB = BA$.

Proof. For uniqueness, let $T = A + iB$. Then $T^* = A - iB$. This gives us formulas for A and B:

$$A = \frac{T + T^*}{2}, \quad B = \frac{T - T^*}{2}.$$

For existence, take the formulas above and verify that they are self-adjoint.

As for part (b), take the formulas above and note that

$$AB - BA = \frac{T^*T - TT^*}{2i},$$

which implies $AB = BA$ iff T is normal. $\square$

Theorem 2.38
Let T be as before. Then $\alpha \in \mathbb{F}$ is an eigenvalue of T with eigenvector f iff $\overline{\alpha}$ is an eigenvalue of T^* with eigenvector f.

Proof. Note that $T - \alpha I$ is normal since T is normal. Thus

$$\|(T - \alpha I)f\| = \|(T^* - \overline{\alpha} I)f\|$$

for every f. This implies that $T - \alpha I$ is injective exactly when $T^* - \overline{\alpha} I$ is injective. $\square$

Theorem 2.39

Eigenvectors of a normal operator corresponding to distinct eigenvalues are orthogonal.

Proof. Let α, β be distinct eigenvalues of T, with respective eigenvectors f, g. By the previous result, we have

$$(\beta - \alpha)\langle f, g\rangle = \langle f, \overline{\beta} g\rangle - \langle \alpha f, g\rangle = \langle f, T^* g\rangle - \langle Tf, g\rangle = 0.$$

Since $\alpha \neq \beta$, this implies that $\langle f, g\rangle = 0$. □

2.5.3 Isometries & Unitary Operators

Let $T \in L(V)$ for a Hilbert space V. Then we say T is an *isometry* if

$$\|Tf\| = \|f\|$$

for every $f \in V$. We say T is *unitary* if

$$T^*T = TT^* = I.$$

Exercise 2.40. Verify each of the following examples:

- The right shift on $\ell^2(\mathbb{N})$ is an isometry but not unitary.
- The right shift on $\ell^2(\mathbb{Z})$ is an isometry and is unitary.
- The multiplication operator M_h is an isometry iff it is unitary iff $h = 1$ almost everywhere.

Theorem 2.41

The following are equivalent:

(a) T is an isometry.
(b) $\langle Tf, Tg\rangle = \langle f, g\rangle$ for every $f, g \in V$.
(c) $T^*T = I$.
(d) $\{Te_k\}$ is an orthonormal family for every orthonormal family $\{e_k\}$.
(e) $\{Te_k\}$ is an orthonormal family for some orthonormal basis $\{e_k\}$.

Proof. Clearly (b) implies (a), if we take $f = g$. To show (a) implies (c), write

$$0 = \|Tf\|^2 - \|f\|^2 = \langle Tf, Tf\rangle - \langle f, f\rangle = \langle (T^*T - I)f, f\rangle.$$

Since $T^*T - I$ is self-adjoint, this happens only when $T^*T - I = 0$. Thus we've verified the equivalence of (a), (b), and (c).

Clearly (b) imlies (d), if we just write

$$\langle Te_k, Te_j\rangle = \langle e_k, e_j\rangle = 0$$

for each k, j. Also, (d) implies (e) is trivial.

We now show (e) implies (b). Let $\{e_k\}$ be an orthonormal basis of V. Then we may write

$$f = \sum a_k e_k, \quad g = \sum b_k e_k.$$

By continuity of inner products and of T, we have

$$\begin{aligned}\langle Tf, Tg\rangle &= \langle \sum T(a_k e_k), \sum T(b_k e_k)\rangle \\ &= \sum a_k \overline{b_k} \|Te_k\|^2 \\ &= \sum a_k \overline{b_k}.\end{aligned}$$

Similarly, we have

$$\langle f, g\rangle = \sum a_k \overline{b_k}.$$

Thus inner products are preserved for all $f, g \in V$, proving the five-way equivalence. □

We now provide the analogous equivalences for unitary operators.

Theorem 2.42

The following are equivalent:

(a) T is unitary.
(b) T is a surjective isometry.
(c) T and T^* are both isometries.
(d) T^* is unitary.
(e) T is invertible with $T^{-1} = T^*$.
(f) $\{Te_k\}$ is an orthonormal basis for every orthonormal basis $\{e_k\}$.
(g) $\{Te_k\}$ is an orthonormal basis for some orthonormal basis $\{e_k\}$.

Proof. First note that (a), (d), and (e) are clearly equivalent by definition. Part (c) of the previous theorem gives us equivalence of (c).

Note that (a) implies (b), since $TT^* = I$ forces T to be surjective, and we know unitary operators are isometries. Next, suppose (b) holds. Then T is surjective and injective, and $T^*T = I$ implies $T^{-1} = T^*$. We now have equivalence of (b) as well.

Again, suppose (b) holds. Then for every orthonormal basis $\{e_k\}$, we have by part (d) of previous theorem that $\{Te_k\}$ is an orthonormal family. But since T is surjective, this implies $\{Te_k\}$ must be a basis. So (b) implies (f). Clearly (f) implies (g).

It suffices to show (g) implies (b). If (g) holds, then the previous theorem says T is an isometry. But furthermore, since $\{Te_k\}$ is an orthonormal basis, for any $g \in V$, we can write

$$g = \sum a_k T e_k = T\left(\sum a_k e_k\right),$$

so T is surjective as well. This completes the seven-way equivalence. □

We can think of the condition $T^*T = TT^* = I$ as akin to $|z|^2 = 1$ for $z \in \mathbb{C}$. In fact, we now extend this idea to the spectrum of unitary operators.

Theorem 2.43

If T is a unitary operator on a Hilbert space, then

$$\sigma(T) \subset \{\lambda \in \mathbb{F} : |\lambda| = 1\}.$$

Proof. Let $|\lambda| \neq 1$. We aim to show that $T - \lambda I$ is invertible. Note that $T - \lambda I$ is normal, so it suffices to show that $(T - \lambda I)^*(T - \lambda I) = (T - \lambda I)(T - \lambda I)^*$ is invertible. We write

$$\begin{aligned}
(T - \lambda I)^*(T - \lambda I) &= (T^* - \overline{\lambda} I)(T - \lambda I) \\
&= (1 + |\lambda|^2)I - (\lambda T^* + \overline{\lambda} T) \\
&= (1 + |\lambda|^2)\left(I - \frac{\lambda T^* + \overline{\lambda} T}{1 + |\lambda|^2}\right).
\end{aligned}$$

It suffices to show the second term is invertible. Notice

$$\left\| \frac{\lambda T^* + \overline{\lambda} T}{1 + |\lambda|^2} \right\| \leq \frac{2|\lambda|}{1 + |\lambda|^2} < 1.$$

By lemma 2.23, this tells us that $(T - \lambda I)^*(T - \lambda I)$ is invertible. Thus $\alpha \notin \sigma(T)$. □

Next, we have the analogous result for isometries.

Theorem 2.44

If T is an isometry on a Hilbert space, and T is not unitary, then

$$\sigma(T) = \{\lambda \in \mathbb{F} : |\lambda| \leq 1\}.$$

Proof. First recall that elements of the spectrum are bounded by $\|T\|$, which equals 1 in this case. So we have

$$\sigma(T) \subset \{\lambda \in \mathbb{F} : |\lambda| \leq 1\}. \tag{6}$$

Next, suppose $|\lambda| < 1$. We wish to show $T - \lambda I$ is not invertible. We know T cannot be surjective, so in particular T is not invertible, and neither is T^*. But write

$$T^*(T - \lambda I) = I - \lambda T^*,$$

which is invertible by lemma 2.23. If $T - \lambda I$ were invertible, then we could write

$$T^* = (I - \lambda T^*)(T - \lambda I)^{-1},$$

which is invertible, contradiction. Therefore

$$\{\lambda \in \mathbb{F} : |\lambda| < 1\} \subset \sigma(T) \tag{7}$$

Together, (6), (7), and the fact that the spectrum is a closed set, implies that

$$\sigma(T) = \{\lambda \in \mathbb{F} : |\lambda| \leq 1\},$$

as desired. $\square$

2.5.4 Compact Operators