Math 255A' Lecture Notes

Professor: Tim Austin
Scribe: Daniel Raban

Contents

1 Hilbert Space Review

1.1 Inner products

In functional analysis, we need to use a field with a topological structure. In this course, we will use the fields $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$.

Definition 1.1. Let H be a vector space over $\mathbb{F}$. A **semi-inner product** $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{F}$ is a function such that

1. $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$
2. $\langle x, y \rangle = \overline{\langle y, x \rangle}$
3. $\langle x, x \rangle \geq 0$.

This is an **inner product** if $\langle x, x \rangle = 0 \implies x = 0$.[1]

Example 1.1. $\mathbb{F}^n$ has the inner product $\langle x, y \rangle = \sum_{i=1}^n x_i \overline{y}_i$.

Example 1.2. $\mathbb{F}^\infty = \{(x_i)_{i=1}^\infty \in \mathbb{F}^\mathbb{N} : x_i = 0 \text{ for all sufficiently large } i\}$ has the inner product $\langle x, y \rangle = \sum_{i=1}^\infty x_i \overline{y}_i$.

Example 1.3. $L^2_\mathbb{F}(\mu) = \{f : X \to \mathbb{F} : f \text{ measurable}, \int |f|^2 \, d\mu < \infty\}$ has the inner product $\langle f, g \rangle = \int f \overline{g} \, d\mu$.

1.2 Norm and metric structure

Theorem 1.1 (Cauchy-Bunyakowski-Schwarz inequality)**.** *Any semi-inner product satisfies*

$$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle} \sqrt{\langle y, y \rangle}.$$

Corollary 1.1. *If we set $\|x\| := \sqrt{\langle x, x \rangle}$, then*

- $\|x + y\| \leq \|x\| + \|y\|$
- $\|\lambda x\| = |\lambda| \cdot \|x\| \qquad \forall \lambda \in \mathbb{F}, x \in H.$

Definition 1.2. $\| \cdot \|$ is called the **(semi-) norm** associated to the (semi-) inner product.

Proposition 1.1 (Polar identity)**.**

$$\|x + y\|^2 = \|x\|^2 + 2 \operatorname{Re}(\langle x, y \rangle) + \|y\|^2$$

Remark 1.1. We get the imaginary part, too, because

$$\operatorname{Re} \langle -ix, y \rangle = \operatorname{Re}(-i \langle x, y \rangle) = \operatorname{Im} \langle x, y \rangle .$$

[1] This is sometimes referred to as the inequality being “coercive.”

Definition 1.3. The **associated metric** to an inner product is $d(x,y) := \|x-y\|$.

Definition 1.4. A **Hilbert space** is an inner product space which is complete with respect to this metric.

Example 1.4. $\mathbb{F}^n$ is a Hilbert space.

Example 1.5. $\mathbb{F}^\infty$ is not complete, so it is not a Hilbert space.

Example 1.6. $L^2(\mu)$ is a Hilbert space.

Proposition 1.2. *If $(H, \langle \cdot,\cdot \rangle)$ is an inner product space, then there is a Hilbert space $(H', \langle \cdot,\cdot \rangle')$ such that*

- *$H \subseteq H'$, and H is dense,*
- *$\langle \cdot,\cdot \rangle' |_{H\times H} = \langle \cdot,\cdot \rangle$.*

The space H' is called the **completion** of H.

Example 1.7. The completion of $\mathbb{F}^\infty$ is $\ell^2 = \{(x_i)_{i=1}^\infty \in \mathbb{F}^{\mathbb{N}} : \sum_{i=1}^\infty |x_i|^2 < \infty\}$ with the inner product $\langle x,y \rangle = \sum_{i=1}^\infty x_i \overline{y}_i$. This is also $L^2(m)$, where m is counting measure on $\mathbb{N}$.

Example 1.8. Let $G \subseteq \mathbb{C}$ be open. Then the **Bergman space** $L_a^2(G)$, the set of L^2 functions that are analytic in G, is a Hilbert space.

1.3 Orthogonality

Definition 1.5. Elements $x, y \in H$ are **orthogonal** (denoted $x \perp y$) if $\langle x,y \rangle = 0$. If $A, B \subseteq H$, we say $A \perp B$ if $x \perp y$ for all $(x,y) \in A \times B$.

Theorem 1.2 (Pythagorean identity)**.** *Let H be a semi-inner product space, and let $x_n \in H$ be such that $x_i \perp x_j$ for all $i \neq j$. Then*

$$\|x_1 + \cdots + x_n\|^2 = \|x_1\|^2 + \cdots + \|x_n\|^2.$$

Corollary 1.2 (Parallelogram law)**.** *For any $x, y \in H$,*

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

Definition 1.6. $A \subseteq H$ is **convex** if whenever $x, y \in A$, $tx + (1-t)y \in A$ for all $t \in [0,1]$.

Proposition 1.3. *Let H be a Hilbert space, let $h \in H$, and let $K \subseteq H$ be nonempty, closed, and convex. Then there is a unique $k \in K$ such that $\|h-k\| \le \|h-k'\|$ for all $k' \in K$.*

Corollary 1.3. *This holds if K is a closed subspace of H.*

Theorem 1.3. *If M is a closed subspace of a Hilbert space and $h \in H$, then $f \in M$ is the closest point to h iff $f \in M$ and $h - f \perp M$.*

Definition 1.7. If $A \subseteq H$, the **orthogonal complement** of A is $A^{\perp} = \{h \in H : h \perp A\}$.

Remark 1.2. For any A, $A^{\perp}$ is a closed, linear subspace.[2]

Theorem 1.4. *Let $M \subseteq H$, $h \in H$, and let Ph be the closest point in M to h. Then*

1. $P(ah + h') = aPh + Ph'$
2. $\|Ph\| \leq \|h\|$
3. $P^2 h = Ph$
4. $\ker P = M^{\perp}$, *and* $\operatorname{im} P = M$.

Definition 1.8. $P = P_M$ is called the **orthogonal projection** onto M.

Corollary 1.4. $(A^{\perp})^{\perp} = \overline{\operatorname{span}} A$.

Corollary 1.5. *If Y is a linear subspace of H, then Y is dense in H if and only if $Y^{\perp} = \{0\}$.*

[2] You could put in a picture of a rabbit, and $A^{\perp}$ would be a closed subspace.

2 More Hilbert Space Review

2.1 Linear functionals

Let H be a Hilbert space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. We want to consider linear functionals $L : H \to \mathbb{F}$.

Proposition 2.1. *Let $L : H \to \mathbb{F}$ be linear. The following are equivalent.*

1. *L is continuous.*
2. *L is continuous at 0.*
3. *L is continuous at one point.*
4. *L is bounded ($\exists c < \infty$ such that $|L(h)| \leq c\|h\|$ for all $h \in H$).*

Definition 2.1. For a bounded linear functional L its **norm** is

$$\begin{aligned}
\|L\| &= \inf\{c > 0 : |L(h)| \leq c\|h\|\} \\
&= \sup\left\{\frac{|L(h)|}{\|h\|} : h \in H \setminus \{0\}\right\} \\
&= \sup\{|L(h)| : \|h\| = 1\}.
\end{aligned}$$

Theorem 2.1 (Riesz representation)**.** *If $L : H \to \mathbb{F}$ is a bounded linear functional, then there is a unique $h_0 \in H$ such that $L(h) = \langle h, h_0 \rangle$ for all $h \in H$. Moreover, $\|L\| = \|h_0\|$.*

Corollary 2.1. *If $L : L^2_{\mathbb{R}}(\mu) \to \mathbb{R}$ is a bounded linear functional, then there exists a unique $h_0 \in L^2_{\mathbb{R}}(\mu)$ such that $L(h) = \int h\overline{h}_0 \, d\mu$ for all $h \in L^2_{\mathbb{R}}(\mu)$.*

2.2 Orthonormal sets and bases

Definition 2.2. A subset $\mathcal{E} \subseteq H$ is **orthonormal** if $\langle e, e' \rangle = \delta_{e,e'}$ for all $e, e' \in \mathcal{E}$. $\mathcal{E}$ is a **basis** if it is maximal under inclusion.

Proposition 2.2. *Any orthonormal set is contained in a basis.*

The proof uses Zorn's lemma.[3]

Example 2.1. In $L^2_{\mathbb{C}}([0, 2\pi])$, let $e_n(t) = \frac{1/\sqrt{2\pi}}{e}^{int}$. The set $\{e_n : n \in \mathbb{Z}\}$ is an orthonormal set (it is actually a basis, too).

Example 2.2. In $\mathbb{F}^n$, let e_k be the vector with all 0s except a 1 in the k-th coordinate. Then $\{e_1, \dots, e_n\}$ is an orthonormal basis.

[3] You cannot do this without waving the magic set theory wand.

Example 2.3. In $\ell^2 = \{(x_i)_{i=1}^\infty : \sum_i |x_i|^2 < \infty\}$, let e_n be the vector with all 0s except a 1 in the n-th coordinate. Then $\{e_n : n \in \mathbb{N}^+\}$ is an orthonormal basis.

Theorem 2.2 (Gram-Schmidt procedure)**.** *If $(h_n)_{n\geq 1}$ is linearly independent, then there is an orthonormal sequence $(e_n)_{n\geq 1}$ such that for all $N \in \mathbb{N}$, we have s*$\operatorname{span}\{h_1, \ldots, h_N\} = \operatorname{span}\{e_1, \ldots, e_N\}$*.*

Proposition 2.3. *Let $\{e_1, \ldots, e_n\}$ be an orthonormal set in H, and let their span be $M = \operatorname{span}\{e_1, \ldots, e_n\}$. Then $P_M h = \sum_{i=1}^n \langle h, e_i \rangle e_i$.*

Proof. Recall that $P_M h$ is the unique vector in M such that $h - P_M h \perp M$. Check this property. □

Theorem 2.3 (Bessel's inequality)**.** *If $(e_n)_{n\geq 1}$ is an orhonormal sequence in H and $h \in H$, then $\sum_{i\geq 1} |\langle h, e_n \rangle|^2 \leq \|h\|^2$.*

Proof. Fix $n \in \mathbb{N}$. Then consider $\langle h, e_1 \rangle, \cdots, \langle h, e_n \rangle e_n, h - P_n h$. The Pythagorean identity gives $\sum_{i=1}^n |\langle h, e_i \rangle|^2 + \|h - P_n h\|^2 = \|h\|^2$. Removing the term $\|h - P_n h\|^2$ gives the inequality for n. □

Corollary 2.2. *If $\mathcal{E}$ is an orthonormal set in H and $h \in H$, then $\mathcal{E}_0 = \{e \in \mathcal{E} : \langle h, e \rangle \neq 0\}$ is countable.*

Proof. We have $\mathcal{E}_0 = \bigcup_{n\geq 1} \mathcal{E}_n$, where $\mathcal{E}_n = \{e \in \mathcal{E} : |\langle h, e \rangle| \geq 1/n\}$. So Bessel's inequality implies $|\mathcal{E}_n| \leq n^2 \|h\|^2$. In particular, each $\mathcal{E}_n$ is finite. □

Corollary 2.3. *If $\mathcal{E}$ is orthonormal in H and $h \in H$, then*

$$\sum_{e\in\mathcal{E}} |\langle h, e \rangle|^2 \leq \|h\|^2.$$

Remark 2.1. By the sum over all $e \in \mathcal{E}$, we mean that it is a countable sum, since all but countably many terms in the sum are 0 for each $h \in H$.

What if we want to talk about uncountable sums in general?

Definition 2.3. Let $(h_i)_{i\in I}$ be an indexed family in H. Then

$$\sum_{i\in I} h_i = k$$

means that for every $\varepsilon > 0$, there is a finite $F \subseteq I$ such that whenever $F \subseteq F \subseteq I$ and $|G| < \infty$, we have $\|k - \sum_{i\in G} h_i\| < \varepsilon$.[4]

[4] This can be rephrased in terms of nets. Let's not do that.

Lemma 2.1. *If $\mathcal{E}$ is an orthonormal set in H, $M = \overline{\operatorname{span}}\, \mathcal{E}$, and $\mathbb{P} = P_M$, then*

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

Theorem 2.4. *Let $\mathcal{E}$ be an orthonormal set in M. The following are equivalent:*

1. *$\mathcal{E}$ is a basis*
2. *If $h \perp \mathcal{E}$, then $h = 0$.*
3. *$\overline{\operatorname{span}}\, \mathcal{E} = H$*
4. *For all $h \in H$, $h = \sum_e \langle h, e \rangle\, e$.*
5. *For all $g, h \in H$, $\langle g, h \rangle = \sum_e \langle g, e \rangle \langle e, h \rangle$.*
6. *For all $h \in H$, $\|h\|^2 \sum_e |\langle h, e \rangle|^2$.*

Corollary 2.4. *Any two bases of H have the same cardinality.*

Definition 2.4. The **dimension** $\dim H$ is the cardinality of a basis of H.

Proposition 2.4. *An infinite-dimensional Hilbert space is separable if and only if its dimension is* $\dim H = \aleph_0$.

2.3 Isomorphisms and isometries

Definition 2.5. An **isomorphism** $A : H \to K$ is a surjective linear operator such that

1. $\langle Ax, Ay \rangle = \langle x, y \rangle$ for all $x, y \in H$
2. A is surjective.

If A only satisfies 1, it is called an **isometry**.

Example 2.4. $A : \ell^2 \to \ell^2$ sending $A(x_1, x_2, \ldots) = (0, x_1, x_2, \ldots)$ is an isometry but not an isomorphism.

Proposition 2.5. *A is an isometry if and only if $\|Ax\|_K = \|x\|_H$ or all $x \in H$*

Proof. ($\implies$) follows from the definition. To get ($\impliedby$), use the Polar identity. □

Theorem 2.5. $\dim H = \dim K$ *if and only if H is isomorphic to K.*

Proof. ($\implies$) Let $\mathcal{E}$ be a basis for H. Then define $A : H \to \ell^2(\mathcal{E})$ as $h \mapsto (\langle h, e \rangle)_{e \in \mathcal{E}}$. We get

$$\begin{tikzcd} H \arrow[d, leftrightarrow] & K \arrow[d, leftrightarrow] \\ \ell^2(\mathcal{E}) \arrow[r, leftrightarrow] & \ell^2(\mathcal{F}) \end{tikzcd}$$

□

Corollary 2.5. *An infinite-dimensional Hilbert space is separable if and only if it is isomorphic to* $\ell^2(\mathbb{N})$.

Example 2.5. The Fourier transform is an isomorphism $L^2_{\mathbb{C}}[0, 2\pi) \to \ell^2_{\mathbb{C}}(\mathbb{Z})$ sending $f \mapsto \int_0^{2\pi} f e_n \, dt$.

2.4 Direct sums

Definition 2.6. Let H, K be inner product spaces. The **direct sum** $H \times K$ is an innerproduct space with coordinatewise addition and the inner product $\langle h \oplus k, h' \oplus k' \rangle = \langle h, h' \rangle + \langle k, k' \rangle$. For an arbitrary family $(H_i)_{i \in I}$, we define

$$\bigoplus_{i \in I} H_i = \left\{ (h_i)_{i \in I} \in \prod_{i \in I} h_i : \sum_{i \in I} \|h_i\|^2 < \infty \right\}, \qquad \langle (h_i)_i, (k_i)_i \rangle = \sum_i \langle h_i, k_i \rangle \, .$$

3 Brief Introduction to Banach Spaces

3.1 Seminorms and norms

We will denote X as a vector space over $\mathbb{F}$.

Definition 3.1. A **seminorm** on X is a function $p : X \to [0, \infty)$ such that

1. $p(x + y) \leq p(x) + p(y)$
2. $p(\lambda x) = |\lambda| p(x)$ for all $x \in X, \lambda \in F$.

We call p a **norm** if $p(x) = 0 \implies x = 0$ (coercivity of p).

Remark 3.1. The second property implies $p(0) = 0$.

A norm has an associated metric $d(x, y) = p(x - y)$.

Definition 3.2. If p is a norm, the pair (X, p) is called a **normed space**. If X is complete with respect to this metric, we call it a **Banach space**.

Proposition 3.1. *In a normed space, addition and scalar multiplication are continuous.*

Lemma 3.1. *Let p, q be seminorms on X. The following are equivalent:*

1. $p \leq q$
2. $\{x \in X : q(x) \leq 1\} \subseteq \{x \in X : p(x) \leq 1\}$
3. $\{x \in X : q(x) < 1\} \subseteq \{x \in X : p(x) < 1\}$
4. $\{x \in X : q(x) < 1\} \subseteq \{x \in X : p(x) \leq 1\}$

Proof. (4) $\implies$ (1): :et $x \in X$ be such that $q(x) \leq a$. Let $\varepsilon > 0$ be arbitrary. Then

$$q\left(\frac{x}{a + \varepsilon}\right) \leq \frac{a}{a + \varepsilon} < 1,$$

so $p(x/(a + \varepsilon)) \leq 1$. This implies $p(x) \leq a + \varepsilon$. □

Proposition 3.2. *For all $x, y \in X$, $|p(x) - p(y)| \leq p(x - y)$.*

Proof. The triangle inequality gives $p(x) \leq p(y) + p(x - y)$, so $p(x) - p(y) \leq p(x - y)$. Flip x and y to get the negative version. □

Remark 3.2. This tells us that the norm in a normed space is Lipschitz.

Definition 3.3. Two norms are **equivalent** if they generate the same topology.

Proposition 3.3. *$\|\cdot\|_1$ and $\|\cdot\|_2$ on X are equivalent if and only if there are constants $c, C > 0$ such that*

$$c\|\cdot\|_1 \leq \|\cdot\|_2 \leq C\|\cdot\|_1.$$

Proof. ($\Longleftarrow$): Given these inequalities, consider $B^2(x,\varepsilon) = \{y \in X : \|y-x\|_2 < \varepsilon\}$. This contains $B^1(x, \varepsilon/c)$. So the topology $\mathcal{T}_{\|\cdot\|_2}$ contains $\mathcal{T}_{\|\cdot\|_1}$. The other inequality gives the reverse inclusion.

($\Longrightarrow$): Assume $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent. Consider $B^1(0,1)$. It must contain some $\|\cdot\|_2$ open neighborhood U of 0. So there is some $\varepsilon > 0$ such that $B^1(0,1) \supseteq B^2(0,\varepsilon)$. This tells you that $\|\cdot\|_1 \leq (1/\varepsilon)\|\cdot\|_2$ by the lemma. We can do the reverse to get another inequality. $\square$

Definition 3.4. $(X, \|\cdot\|)$ and $(X', \|\cdot\|')$ are **isometric**[5] if there is a linear bijection $A : X \to X'$ such that $\|Ax\|' = \|x\|$ for all $x \in X$. They are **isomorphic** if $\|\cdot\|$ and $\|A(\cdot)\|'$ are equivalent.

3.2 Examples of Banach spaces

Example 3.1. Let X be a Hausdorff[6] topological space. Then let the space $C_b(X) = \{\text{bounded continuous functions } X \to \mathbb{F}\}$ equipped with the **uniform/sup norm** $\|f\| := \sup_{x\in X}|f(x)|$. Then $(C_b(X).\|\cdot\|)$ is a Banach space.

Example 3.2. If I is any set with the discrete topology, the previous example gives $C_b(I) = \ell^\infty(I) = \{(x_i)_{i\in I} \in \mathbb{F}^I : \sup_i |x_i| < \infty\}$. If $I = \mathbb{N}$, we call $\ell^\infty(\mathbb{N}) = \ell^\infty$.

Example 3.3. If X is locally compact, $C_0(X) = \{f \in C_b(X) : \forall\varepsilon > 0, \{|f| \geq \varepsilon\} \text{ is compact}\}$ is a closed subspace of $C_b(X)$. If X is compact, $C_b(X) = C_0(X) =: C(X)$.

We call $c_0 = C_0(\mathbb{N}) = \{(x_i)_i \in \mathbb{F}^{\mathbb{N}} : x_i \xrightarrow{i\to\infty} 0\}$.

Example 3.4. Let $(X.\Sigma, \mu)$ be a measure space. Then $L^p(\mu)$ for $1 \leq p \leq \infty$ is a Banach space with the norm $\|f\|_p = (\int |f|^p)^{1/p}$ if $p < \infty$ and $\|f\|_\infty = \operatorname{ess\,sup}|f|$.

Example 3.5. Fix $n \geq 1$, and let $C^{(n)}([0,1]) = \{f : [0,1] \to \mathbb{F} \text{ with } n\text{-fold conts. derivs.}\}$. With the norm $\|f\| = \max_{-\leq k \leq n} \sup_x |f^{(k)}(x)|$, $C^{(n)}([0,1])$ is a Banach space.

Similar spaces called **Sobolev spaces**, where we do not require the last derivative to be continuous. These are useful for PDEs; people will define a Banach space of functions with the correct amount of regularity to find a solution to a PDE inside.

[5] This is generally a really rigid condition. Theorems about isometry are almost always easy or false.

[6] We don't actually need this, but analysts don't like thinking about non-Hausdorff psaces. If you ever wonder why the Hausdorff condition is there in a situation, it might be sociological prejudice.

3.3 Bounded linear operators

Definition 3.5. A **continuous linear operator** $X \to X'$ is a linear operator which is continuous according to the norm topologies.

Proposition 3.4. *Let $T : X \to X'$ be linear. The following are equivalent:*

1. *T is continuous.*
2. *T is continuous at 0 ($= 0_X$).*
3. *T is continuous at some point in X.*
4. *There exists some $c < \infty$ such that $\|Tx\|' \leq c\|x\|$ for all $x \in X$.*

The proof is similar to the proof of the lemma from before. Because of condition 4, continuous linear operators are often referred to as **bounded**.

Definition 3.6. $\mathcal{B}(X, X')$ denotes the vector space of bounded linear operators $X \to X'$. This has the **operator norm**

$$\begin{aligned}\|T\| &= \inf\{c > 0 : \|Tx\|' \leq c\|x\| \ \forall x \in X\} \\ &= \sup\left\{\frac{\|Tx\|'}{\|x\|} : x \in X \setminus \{0\}\right\} \\ &= \sup\{\|Tx\|' : \|x\| = 1\}.\end{aligned}$$

Example 3.6. Fix a measure space (X, Σ, μ) and $1 \leq p \leq \infty$, and let $\varphi \in C^\infty(\mu)$. Then the **multiplication operator** $M_\varphi : L^p(\mu) \to L^p(\mu)$ sending $f \mapsto \varphi f$ is bounded:

$$\|M_\varphi f\|_p = \left(\int |\varphi f|^p\right)^{1/p} \leq \|\varphi\|_\infty \|f\|_p.$$

We can choose a positive measure set where φ is close to its essential supremum and let f be the indicator of that set. This makes $\|M_\varphi f\|_p$ arbitrarily close to $\|\varphi\|_\infty \|f\|_p$, so we get $\|M_\varphi\| = \|\varphi\|_\infty$.

Example 3.7. Consider $L^p(\mu)$. Assume $K : X \times X \to \mathbb{F}$ is such that there exist constants $c_1, c_2 < \infty$ such that

$$\int |K(x,y)| \, d\mu(x) \leq C_1 \qquad \text{for } \mu\text{-a.e. } y,$$
$$\int |K(x,y)| \, d\mu(y) \leq C_2 \qquad \text{for } \mu\text{-a.e. } x.$$

Then the operator $M : L^p \to L^p$ defined by

$$Mf(x) := \int K(x,y) f(y) \, d\mu(y)$$

is well-defined, and $\|M\| \leq C_1^{1/q} C_2^{1/p}$, where $1/p + 1/q = 1$.

Example 3.8. Let X, Y be compact, Hausdorff spaces, and let $\tau : Y \to X$ be continuous. Then the **pullback/composition operator** $\tau^* : C(X) \to C(Y)$ given by $f \mapsto f \circ \tau$ is bounded with $\|\tau^*\| \leq 1$. If $Y \neq \varnothing$, then $\|\tau^*\| = 1$.

4 Finite Dimensional Normed Spaces, Quotients, Products, and Dual Spaces

4.1 Norms on finite dimensional space

Theorem 4.1. *Let X be a normed space. If $\dim X < \infty$, then any two norms on X are equivalent.*

Proof. We can assume $X = \mathbb{F}^n$. If $\|\cdot\|$ is a mystery norm, we show that $\|\cdot\|$ is equivalent to the ℓ^1 norm $|x| = \sum_{i=1}^n |x_i|$.

Step 1: Let $e_1, \ldots, e_n$ be the standard orthonormal basis of $\mathbb{F}^n$. Then let $M := \max_{1 \leq i \leq n} \|e_i\|$. Then

$$\|x\| = \left\| \sum_{i=1}^n x_i e_i \right\| \leq \sum_i |x_i| \|e_i\| \leq M|x|.$$

Step 1.5: This shows that $\mathrm{Id} : \mathbb{F}^n \to \mathbb{F}^n$ is continuous from $|\cdot|$ to $\|\cdot\|$. So $\{x : |x| = 1\}$ is compact according to $\|\cdot\|$.

Step 2: So we get $\varepsilon > 0$ such that any x with $|x| = 1$ has $\|x\| \geq \varepsilon$. So $\{\|\cdot\|/\varepsilon < 1\} \subseteq \{|\cdot| < 1\}$. That is, $|\cdot| \leq (1/\varepsilon)\|\cdot\|$. □

Remark 4.1. A result called John's theorem gives explicit constants dependent on n.[7]

Corollary 4.1. *Any finite dimensional subspace of a normed space is closed.*

Corollary 4.2. *Let X, Y be a normed spaces with $\dim X < \infty$. Then if $T : X \to Y$ is linear, it must be continuous.*

Proof. $\|x\|_X + \|Tx\|_Y$ is a norm for X, so there is a constant $M < \infty$ such that $\|x\|_X + \|Tx\|_Y \leq M\|x\|_X$. □

4.2 Quotients in normed spaces

Let X be a normed space over $\mathbb{F}$ with a subspace M. Linear algebra tells you that the quotient $X/M = \{x + M : x \in X\}$ is a vector space.

Definition 4.1. The quotient space X/M has the **quotient seminorm** $\|x + M\| := \inf\{\|x - y\| : y \in M\} = \mathrm{dist}(x, M)$.

Lemma 4.1. *The quotient seminorm is a norm if and only if M is closed.*

Definition 4.2. The **quotient map** is the map $Q : X \to X/M$ given by $x \mapsto x + M$.

Theorem 4.2. *The quotient has the following properties:*

[7] Check out the proof of this one!

1. $\|Qx\| \leq \|x\|$ *for all* $x \in X$.

2. *If* X *is a Banach space and* M *is closed, then* X/M *is a Banach space.*

3. $U \subseteq X/M$ *is open if and only if* $Q^{-1}(U)$ *is open in* X.

4. Q *is an open mapping.*

Proof. 1. Since $0 \in M$, $\|x + M\| \leq \|x + 0\| = \|x\|$.

2. Suppose $(x_n + M)_n$ is Cauchy in X/M. Then there is a subsequence $(x_{n_i} + M)_i$ such that $\|x_{n_i} - x_{n_{i+1}} + M\| < 2^{-i}$ for all i. Then there is a $y_i \in M + (x_{n_i} - x_{n_{i+1}})$ such that $\|y_i\| < 2^{-i}$. Now $x_{n_2} \in x_{n_1} + y_1 + M$, $x_{n_3} \in x_{n_1} + y_1 + y_2 + M$, and so on, giving us $x_{n_{i+1}} \in x_{n_1} + y_1 + \cdots y_i + M$, where $x_{n_i} + y_1 + \cdots y_i$ is a Cauchy sequence in X.

   Now suppose that $x_{n_i} + y_1 + \cdots y_i \to z$. Then $\|x_{n_{i+1}} - z + M\| \to 0$. Then $x_n + M \to z + M$ in X/M

3. This is a rephrasing of (4).

4. Continuity follows from part(1). If $U \subseteq X$ is open, $x \in U$, and $B(x, \varepsilon) \subseteq U$, then any $y + M \in B_{X/M}(x + M, \varepsilon) = Q(B(x, \varepsilon)) \subseteq Q(U)$.

□

Definition 4.3. We use $M \leq X$ to say that M is a *closed* subspace of X.

Theorem 4.3. *If* X *is a normed space,* $M \leq X$*, and* N *is any finite dimensional subspace, then* $M + N = \{x + y : x \in M, y \in N\}$ *is closed.*

Proof. Observe that $M + N = Q^{-1}(Q(N))$. $Q(N)$ is finite dimensional, so it is closed. Q is continuous, so $Q^{-1}(Q(N))$ is closed. □

Remark 4.2. This is surprisingly tricky to prove without using the quotient X/M.

4.3 Products of normed spaces

If we have a general family $(X_i)_{i \in I}$ of normed spaces, there is no canonical norm on the product. We may define notions of product by considering various subspaces of $\prod_{i \in I} X_i$.

Example 4.1. Fix $1 \leq p < \infty$. The ℓ^p**-direct sum** $\bigoplus_p X_i = \{(x_i)_{i \in I} \in \prod_i X_i : \sqrt{\sum_i \|x_i\|_i^p} < \infty\}$ is a normed space with the norm $\|(x_i)_i := \sqrt{\sum_i \|x_i\|_i^p}$.

Example 4.2. The ℓ^∞**-direct sum** $\bigoplus_\infty X_i = \{(x_i)_{i \in I} \in \prod_i X_i : \sup_i \|x_i\|_i < \infty\}$ is a normed space with the norm $\|(x_i)_i := \sup_i \|x_i\|_i < \infty$.

Example 4.3. If $I = \mathbb{N}$, we also have $\bigoplus_0 X_i = \{(x_i)_{i \in I} \in \prod_i X_i : \|x_i\|_i \to 0\}$.

Proposition 4.1. *1. For each of these notions of product X, X is complete if and only if X_i is complete for all i.*

2. $X \to X/\{(x_i)_{i\in I} : x_i = 0\}$ is an isometry to X_i.

3. Each coordinate projection $X \to X_i$ has norm 1 and is open.

4.4 Dual spaces

Definition 4.4. The **dual** of X is the space $X^* := \mathcal{B}(X, \mathbb{F})$ of bounded linear functionals. The **dual norm** is $\|L\|_* := \sup\{|L(x)| : \|x\| = 1\}$.

Proposition 4.2. *If Y is complete, $\mathcal{B}(X, Y)$ is complete.*

Corollary 4.3. *X^* is a Banach space.*

Here is a proof of this fact independent of the general fact about operators.

Proof. Let $L \in X^*$, and consider $L|_B$ restricted to the closed unit ball. Then $L|_B \in C_b(B)$. So the map ρ sending $L \mapsto L|_B$ gives us that $\rho(X^*)$ is a lienar subspace of $C_b(X)$. Moreover, $\rho(X^*)$ is closed. Since $C_b(B)$ is complete, so is $\square$

Example 4.4. Let $X = c_0 = \{(x_i)_{i\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} : x_i \to 0\}$. Then $L(x_1, x_2, \dots) = x_1$ is a linear functional.

Let e_i be the vector with all 0s but a 1 in the i-th coordinate. Then $\{e_1, e_2, \dots,\} \cup \{(1, 1/2, 1/3.1/4, \dots)\}$ is linearly independent. So there exists a linear functional $L : c_0 \to \mathbb{R}$ such that $L(e_i) = 0$ for all i and $L(1, 1/2, 1/3, \dots) = 1$. This L is not continuous.

5 The Hahn-Banach Theorem

5.1 Examples of Dual Spaces

Here are examples of concrete descriptions of some dual spaces.

Example 5.1. If (X, Σ, μ) is a measure space, $1 < p < \infty$, and $p^{-1} + q^{-1} = 1$, then the map $L^q \to (L^p(\mu))^*$ given by $g \mapsto L_g$ is a linear isometry, where $L_g(f) = \int_X fg \, d\mu$.

Example 5.2. If (X, Σ, μ) is σ-finite, then $L^\infty(\mu) \mapsto (L^1(\mu))^*$ given by $g \mapsto L_g$ is an isometric isomorphism, where $L_g(f) = \int fg \, d\mu$.

Example 5.3. Let X be a locally compact Hausdorff space, and let $M(X)$ be the set of $\mathbb{F}$-valued regular[8] Borel measures on X with $\|\mu\|$ equalling the total variation of μ. Then the map $M(X) \to C_0(X)^*$ given by $\mu \mapsto L_\mu$ is an isometric isomorphism, where $M_\mu(f) = \int_X f \, d\mu$.

5.2 The Hahn-Banach theorem

Let X be a vector space over $\mathbb{F}$.

Definition 5.1. A **sublinear functional** on X is a function $p : X \to \mathbb{R}$ such that

1. $p(x + y) \leq p(x) + p(y)$
2. $p(\alpha y) = \alpha p(y)$ for all $\alpha \in [0, \infty)$.

Example 5.4. Any seminorm is a sublinear functional.

Theorem 5.1 (Hahn-Banach). *Let $\mathbb{F} = \mathbb{R}$, let M be a linear subspace of X, and let p be a sublinear functional on X. If $f : M \to \mathbb{R}$ is linear and $f \leq p|_M$, then there is a linear $F : X \to \mathbb{R}$ such that $F|_M = f$ and $F \leq p$.*

Proof. Step 1: Assume $\dim(X/M) = 1$. Then there is some $x_0 \in X$ such that $M + \mathbb{R}x_0 = X$. We must find something of the form $F(y + tx_0) = f(y) + t\alpha_0$ for some $\alpha_0 \in \mathbb{R}$ such that $F \leq p$. What must α_0 satisfy? We need $f(y) + t\alpha_0 \leq p(y) + tx_0)$ for all $y \in M, t \in \mathbb{R}$.

- If $t > 0$, divide by t to get $f(y') + \alpha_0 \leq p(y' + x_0)$ for all $y' \in M$. That is, we need $\alpha_0 \leq \inf_{y' \in M} p(y' + x_0) - f(y')$.
- If $t < 0$, divide by $-t$ to get $f(y') - \alpha_0 \leq p(y' - x_0)$ for all $y' \in M$. That is, we need $\alpha_0 \geq \sup_{y' \in M} f(y') - p(y' - x_0)$.

[8] For general locally compact spaces, "regular" can have different meanings. Take it to have the meaning that makes this theorem work.

It remains to check that for any $y', y'' \in M$, $f(y'') = p(y'' - x_0) \leq p(y' - x_0) - f(y')$ (so such an α_0 exists). We can rearrange this to get $f(y' + y'') \leq p(t' + x_0) + p(y'' - x_0)$. But this is true because

$$f(y + y'') \leq p(y' + y'') \leq p(y' + x - 0) + p(y'' - x_0)$$

by the subadditivity of p.

Step 2: The idea is to "iterate" Step 1 to get the general case. Let $\mathcal{P}$ be the collection of pairs (N, g) where N is a linear subspace such that $M \subseteq N \subseteq X$, $g : N \to \mathbb{R}$ is linear, and $g|_M = f$. We have the partial ordering $(N, g) \leq (N'.g')$ if $N \subseteq N'$ and $g'|_N = g$. If $((N_i, g_i))_i$ is a chain in $\mathcal{P}$, then $(\bigcup_i N_i, \bigcup_i g_i) \in \mathcal{P}$ is an upper bound for the chain. By Zorn's lemma, there is a maximal element $(N, g) \in \mathcal{P}$. We now must have $N = X$; otherwise, apply Step 1 to $N \subseteq N + \mathbb{R}x_1$ for some $x_1 \in X \setminus N$ to contradict the maximality of N. $\square$

Theorem 5.2 (complex Hahn-Banach)**.** *Let $\mathbb{F} = \mathbb{C}$, let M be a linear subspace of X, and let p be a sublinear functional on X. If $f : M \to \mathbb{C}$ is such that $|f(x)| \leq p(x)$ for all $x \in M$, then there exists some linear $F : X \to \mathbb{C}$ such that $F|_M = f$ and $|F| \leq p$.*

Proof. Here is the sketch. Treat X as a real vector space. Then $g = \mathrm{Re}(f)$ is an $\mathbb{R}$-linear functional $M \to \mathbb{R}$. Extend g via the real Hahn-Banach theorem to get G on all of X. If $G : X \to \mathbb{R}$ is $\mathbb{R}$-linear, then $F(x) = F(x) - iG(ix)$ is $\mathbb{C}$-linear. Then $\|F\| = \|G\|$. $\square$

Here is the special case where p is a norm.

Corollary 5.1. *Let X be a normed space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. If M is a linear subspace and $f \in M^*$, then there is an $F \in X^*$ such that $F|_M = f$ and $\|F\| = \|f\|$.*

5.3 Corollaries of Hahn-Banach

Corollary 5.2. *Let X be a normed space over $\mathbb{F}$, let $x_1, \dots, x_n \in X$ be linearly independent, and let $\alpha_1, \dots, \alpha_n \in \mathbb{F}$. Then there is some $f \in X^*$ such that $f(x_i) = \alpha_i$ for all i.*

Proof. Define f by $f(x_i) = \alpha_i$ on $\mathrm{span}\{x_1, \dots, x_n\}$. This is automatically bounded since this is a finite dimensional subspace. Now apply Hahn-Banach. $\square$

Corollary 5.3. *Let X be a normed space over $\mathbb{F}$, and let $x \in X$. Then $\|x\| = \max\{|f(x)| : f \in X^*, \|f\| \leq 1\}$.*

Proof. ($\geq$): This follows from the definition of the dual norm.

($\leq$): Apply the previous corollary with $x_1 = x$ and $\alpha_1 = \|x\|$. $\square$

Corollary 5.4. *Let X be a normed space over $\mathbb{F}$, let M be a non-dense linear subspace, and let $x \in X$. Then there is an $f \in X^*$ such that $f|_M = 0$, $\|f\| = 1$ and $f(x) = \mathrm{dist}(x, M)$.*

Proof. Consider the quotient map $Q : X \to X/M$. By Hahn-Banach, there exists an $f_0 \in (X/M)^*$ such that $\|f_0\| = 1$ and $f_0(x + M) = \operatorname{dist}(x, M)$. Let $f := f_0 \circ Q$. Then $f(y) = f_0(y + M)$ for all $y \in X$. $\square$

Corollary 5.5. *If X is a normed space and M is a linear subpsace, then*

$$\overline{M} = \bigcap_{\substack{f \in X^* \\ f|_M = 0}} \ker f.$$

Proof. ($\subseteq$): $\ker f \supseteq M$ for each element in the intersection, and each $\ker f$ is closed.

($\supseteq$): If $x \in X \setminus \overline{M}$, then take f from the previous corollary. Then $f(x) > 0$, so $x \notin \bigcap_f \ker f$. $\square$

6 Applications of The Hahn-Banach Theorem

The last application we will go over is not in Conway's textbook.

6.1 Banach limits

Let $\mathbb{F} = \mathbb{R}$ and $\ell^\infty = \ell^\infty(\mathbb{N})$. Then $c = \{(x_n) \in \ell^\infty : \lim_n x_n \text{ exists}\}$ is a closed subspace. lim is a bounded linear functional on c, so we can extend it to all of ℓ^∞.

Theorem 6.1. *There exists an $L \in (\ell^\infty)^*$ such that*

1. *$L(x) = \lim_n x_n$ for all $x \in c$,*
2. *$L(x) \geq 0$ if $x_n \geq 0$ for all n,*
3. *$L(\sigma(x)) = L(x)$, where $\sigma(x) = (x_2, x_3, \dots)$.*

Proof. Let $M = \{x - \sigma(x) : x \in \ell^\infty\}$, which is a linear subspace of ℓ^∞. We will apply a corollary of Hahn-Banach to get an L that kills M and $L((1, 1, \dots)) = 1$.

Claim: $\operatorname{dist}(\mathbf{1} = (1, 1, \dots), M) = 1$. Let $x - \sigma(x) \in M$. Then

$$\operatorname{dist}(\mathbf{1}, x - \sigma(x)) = \sup_n |1 - (x_n - x_{n+1})|.$$

Since $(x_n) \in \ell^\infty$, the right hand side gets arbitrarily close to 1 when x_n is close to $\inf_m x_m$. So there exists an $L \in (\ell^\infty)^*$ such that $L(M) = 0$, $L(\mathbf{1}) = 1$ and $\|L\| = 1$. This covers property 3.

For property 2, use $\|L\| = 1$ and $L(\mathbf{1}) = 1$. It's similar to the fact that if μ is a signed measure, then $|\mu|(X) = |\mu(X)| \implies \mu = 0$.

For property 1, suppose $x \in c$ and let $\alpha := \lim_n x_n$. We claim that $\|\sigma^n(x)\| - \alpha\mathbf{1}\| \to 0$ as $n \to \infty$; this is a rewording of $\alpha := \lim_n x_n$. So $L(x) = L(\sigma^n(x)) \to \alpha L(\mathbf{1}) = \alpha$. □

Corollary 6.1. *$c_0 \subseteq \ker L$.*

6.2 Dual of quotients by subspaces

Let X be a normed space, and let $M \leq X$ (i.e. M is a closed subspace).

Definition 6.1. The **annihilator** of M is $M^\perp := \{L \in X^* : L|_M = 0\}$.

Theorem 6.2. *Let X be a normed space, and let $M \leq X$. Then the map $X^*/M^\perp \to M^*$ sending $f + M^\perp \mapsto f|_M$ is an isometric isomorphism.*

Proof. This map is linear. We need to show that it is surjective and that $\|f|_M\| = \|f + M^\perp\|$. We have the inequality $\leq$. The rest of the proof follows from the following claim: Given $g \in M^*$, there exists some $f \in X^*$ such that $f|_M = g$ and $\|f\| = \|g\|$. This is just Hahn-Banach. □

Theorem 6.3. *Let $O : X \to X/M$ be the quotient map. Then the map $(X/M)^* \to M^\perp \subseteq X^*$ given by $g \mapsto g \circ Q$ is an isometric isomorphism.*

Proof. Any $v\ inM^\perp$ defines a linear functional on X/M. We want it to be bounded with the same norm:

$$\|f\| = \sup\{|f(x)| : x \in X, \|x\| \leq 1\} \implies f(x) \leq \|f\| \cdot \|x + M^\perp\|$$

for all x. And Q is surjective. $\square$

6.3 The double dual

We can keep taking the dual spaces of dual spaces to get $X^*, X^{**}, X^{***}, \dots$.

Definition 6.2. The **natural map** $X \to X^{**}$ is given by $x \mapsto \hat{x}$, where $\hat{x}(f) = f(x)$.

Lemma 6.1. *The natural map is isometric.*

Proof. Since $|\hat{x}(f)| \leq \|f\| \cdot \|x\|$, we have $\|\hat{x}\| \leq \|x\|$. Equality is by Hahn-Banach. $\square$

Definition 6.3. X is **reflexive** if $\hat{X} = X^{**}$; i.e. the natural map is surjective.

Example 6.1. Let $1 < p < \infty$. Then $L^p(\mu)$ is reflexive by Riesz representation.

Example 6.2. Id $\dim X < \infty$, then X is reflexive.

Example 6.3. $c_0 = C_0(\mathbb{N})$ is not reflexive. c_0^* is the collection of signed finite measures on $\mathbb{N}$, which is $\ell^1(\mathbb{N})$ by Riesz-representation. So c_0^{**} is $\ell^\infty(\mathbb{N})$, which is bigger than c_0.

6.4 Optimal transport

Let (X, ρ) be a metric, and let $\mu, \nu \in \mathrm{Prob}(X)$. We want to move the mass according to the distribution μ to that of ν. This is called the **transport problem**. For infinitesimal regions dx, dy, think of $\lambda(dx, dy)$ as how much mass moves from dx to dy. We interpret $\lambda \in \mathrm{Prob}(X \times X)$.

λ has to satisfy

$$\sum_{dy} \lambda(dx, dy) = \mu(dx), \qquad \sum_{dx} \lambda(dx, dy) = \nu(dy)$$

so we want

$$\lambda(A \times X) = \mu(A), \qquad \lambda(X \times B) = \nu(B)$$

for all measurable $A, B \subseteq X$.

Definition 6.4. A measure λ with these properties is called a **coupling** of μ, ν.

Call the collection of all such couplings C.

Example 6.4. The product measure $\lambda = \mu \times \nu$ is a coupling.

Let's suppose it costs us to move mass from dx to dy. Then we want to find λ and estimate

$$\min_{\lambda \in C} \int \rho(x,y)\, d\lambda(x,y).$$

This is not the inf because it is weak*-continuous.

Obstructions: Suppose $f \in L$, the collection of 1-Lipschitz functions $X \to \mathbb{R}$. Then for $\lambda \in C$,

$$\int f\, d\mu - \int f\, d\nu = \int (f(x) - f(y))\, d\lambda(x,y) \le \int |f(x) - f(y)|\, d\lambda(x,y) \le \int \rho\, d\lambda.$$

Theorem 6.4. *Let $D := \sup_{f \in L} |\int f\, d\mu - \int f\, d\nu|$. THen there exists a $\lambda \in C$ such that $\int \rho\, d\lambda = D$.*

Remark 6.1. This theorem says that these obstructions are the only ones.

Proof. Equivalently, by Riesz representation, we want $\phi \in C(X \times X)^*$ (which corresponds to λ) such that

1. (coupling) $\phi(f(x) \cdot 1(y)) = \int f\, d\mu$ and $\phi(1(x) \cdot g(y)) = \int g\, d\nu$ for all $f, g \in C(X)$,
2. (minimizer) $\phi(\rho) = D$,
3. (probability measure) $\|\phi\| = \phi(\mathbb{1}_{X \times X}) = 1$.

So define $M := \{f(x) + g(y) + a\rho(x,y); f, g \in C(X), a \in \mathbb{R}\} \subseteq C(X \times Y)$. Define ψ on M by

$$\psi(f + g + a\rho) = \int f\, d\mu + \int g\, d\nu + aD.$$

We need to check that ψ is well-defined and $\|\psi\|_{M^*} \le 1$. Taking $b = 0, 1$ it follows from the claim: If $f + g + a\rho \le b$, then $\psi(f + g + a\rho) \le b$. It is equivalent to show that if $f + g \le b + a\rho$ for all x, y then $\int f\, d\mu + \int g\, d\nu \le b + aD$.

Case 1: $a \le 0$. This is straightforward and is in the online notes.

Case 2: $a > 0$. We may assume $a = 1$. Rewrite this as $f(x) \le \inf_y [b - g(y) + \rho(x,y)] =: h(x) \in L$. Then $f \le h \le b - g(x)$, so

$$\int f\, d\mu \le \int h\, d\mu \le \int d\, d\nu + D \le b - \int g\, d\nu + D.$$

- Start with $b - g(x)$
- Draw cones at each point on the graph.
- Take h to be the minimum of the cones.

Then applying Hahn-Banach to ψ gives ϕ. $\square$

7 The Open Mapping Theorem, Closed Graph Theorem, and Uniform Boundedness Principle

7.1 The open mapping theorem

Definition 7.1. Let X and Y be topological spaces. A function $f : X \to Y$ is **open** if $f[U]$ is open for all open $U \subseteq X$.

Remark 7.1. In metric spaces, this is equivalent to: For all $B(x,r) \subseteq X$, there is an $\varepsilon > 0$ such that $f[B(x,r)] \supseteq B(f(x), \varepsilon)$. In normed spaces, it is enough to check this at $x = 0_X$.

Theorem 7.1 (Open mapping theorem)**.** *Let X, Y be Banach spaces. If $A : X \to Y$ is a bounded linear surjection, then A is open.*

Proof. Step 1: Write $Y = \bigcup_{n=1}^\infty \overline{A(B_X(0,n))}$. By the Baire category theorem, these cannot all be nowhere dense. So there exist $n \in \mathbb{N}$, $y \in Y$, $t > 0$ such that $\overline{A(B_X(0,n))} \supseteq B_y(y,t)$. The left hand side is symmetric under $z \mapsto -z$, so $\overline{A(B_X(0,n))} \supseteq B_Y(-y,t)$, as well. By convexity,

$$\begin{aligned}\overline{A(B_X(0,n))} &\supseteq \left\{\frac{1}{2}(y+z) + \frac{1}{2}(-y+w) : \|z\|_Y, \|w\|_Y < t\right\} \\ &= \left\{\frac{1}{2}z + \frac{1}{2}w : \|z\|_Y, \|w\|_Y < t\right\} \\ &= B_Y(0,t).\end{aligned}$$

Step 2: For any $a > 0$,

$$\overline{A(B_X(0,an))} \supseteq B(0,at).$$

Step 3: We will show that $A(B_X(0,2)) \supseteq B(0,r)$. for $r = t/n$. Let $y \in B_Y(0,r)$. By step 2, there is an $x_1 \in B_X(0,1)$ such that $\|y - Ax_1\| < r/2$. Let $y_1 = y - Ax_1$, and choose $x_2 \in B_x(0,1/2)$ such that $\|y_1 - Ax_2\| < r/2$. In this way, pick y_n, x_{n+1} for each n. Let $x = \sum_{n=1}^\infty x_n$; this converges because the lengths are bounded by a convergent geometric series: $\|x\| \le \sum_n \|x_n\| < 2$. Then $Ax = \sum_{n=1}^\infty Ax_n$. For each $N \in \mathbb{N}$,

$$y - \sum_{n=1}^N Ax_n = y_1 - \sum_{n=2}^N Ax_n = y_2 - \sum_{n=3}^N Ax_n = \cdots = y_N,$$

and $\|y_N\| - r/2^{N-1} \to 0$. So $y = Ax$. $\square$

Corollary 7.1. *A bounded linear bijection between Banach spaces is an isomorphism.*

Proof. Since $A : X \to Y$ is a bijection, A^{-1} exists as a linear transformation $Y \to X$. Boundedness of A^{-1} is precisely the openness of A. $\square$

Definition 7.2. If $A : X \to Y$, then **graph** of A is $\mathrm{gra}(A) := \{(x, Ax) : x \in X\} \subseteq X \oplus Y$. It is a linear subspace of $X \oplus Y$ with the **graph norm** $\|(x,y)\| = \|x\|_X + \|y\|_Y$.

7.2 The closed graph theorem

Corollary 7.2 (Closed graph theorem)**.** *Let X, Y be Banach spaces, and let $A : X \to Y$ be a linear transformation. If* $\mathrm{gra}(A)$ *is closed, then A is continuous.*

Proof. $\mathrm{gra}(A)$ is a closed subspace of a Banach space, so it is complete. In the following diagram, $A = P_2 \circ \tilde{A}$, so it is enough to show that $\tilde{A}$ is continuous.

$$\begin{tikzcd} A \arrow[r, "\tilde{A}: x \mapsto (x,\, Ax)"] \arrow[dr, "A"'] & \mathrm{gra}(A) \arrow[d, "P_2 : (x,\, y) \mapsto y"] \\ & Y \end{tikzcd}$$

But $\tilde{A} = (P_1|_{\mathrm{gra}(A)})^{-1}$, so it is continuous by the previous corollary. $\square$

Example 7.1. Let $X = C^{(1)}[0,1]$ and $Y = C[0,1]$, both with the uniform norm. Then A sending $f \mapsto f'$ is not continuous. But its graph, $\mathrm{gra}(A) = \{(f, f') : f \in C^{(1)}\}$ is closed: Suppose $(f_n)_n$ is such that $f_n \to g$ uniformly, and $f_n' \to h$ uniformly. Then $f_n - f \to 0$, which means that $f_n' - g' \to h - g'$ uniformly; sp we may assume that $f_n \to 0$ and $f_n' \to h$. We must show that $h = 0$. We have that for all $t \in [0,1]$. so

$$\int_0^t h(s)\, ds = \lim_n \int_0^t f_n' = \lim_n [f_n(t) - f_n(0)] = 0.$$

So $h = 0$.

In general, $\mathrm{gra}(A)$ is closed if $x_n \to 0$ and $Ax_n \to y$ implies $y \to 0$. This is often easier to check than continuity.

7.3 The principle of uniform boundedness

Theorem 7.2 (Principle of uniform boundedness)**.** *Let X be a Banach space, let Y be a normed space, and let $\mathcal{A} \subseteq \mathcal{B}(X, Y)$. Assume that* $\sup\{\|Ax\| : A \in \mathcal{A}\} < \infty$ *for all $x \in X$. Then* $\sup\{\|A\| : A \in \mathcal{A}\} < \infty$.

Instead of citing Baire category, we will adapt the proof of that theorem to prove this.

Proof. Assume, towards a contradiction, that $M(x) := \sup\{\|Ax\| : A \in \mathcal{A}\} < \infty$ for all x, but $\sup_{A \in \mathcal{A}} \|A\| = \infty$. So for every $\varepsilon > 0$, there is an $x \in X$ and $A \in \mathcal{A}$ such that $\|x\| < \varepsilon$ and $\|Ax\| > 1/\varepsilon$.

Construct sequences (x_n) in x and (A_n) in $\mathcal{A}$ by recursion: Pick any $\|x_1\| = 1$ and any A_1. Now choose (x_2, A_2) such that $\|x_2\| \leq 1/2$, $\|A_1 x_2\| \leq 1/2$, and $\|A_2 x_2\| > 2 + M(x_1)$. Now choose (x_3, A_3) such that $\|x_3\|, \|A_1 x_3\|, \|A_2 x_3\| < 1/4$ but $\|A_3 x_3\| > 3 + M(x_1) +$

$M(x_2)$. At the n-th stage, choose (x_n, A_n) such that $\|x_n\|, \|A_1 x_n\|, \dots, \|A_{n-1}x_n\| < 1/2^n$ but $\|A_n x_n\| ? n + M(x_1) + M(x_2) + \cdots + M(x_{n-1})$.

Now let $x = \sum_{n=1}^{\infty} x_n$. Then

$$\begin{aligned} A_k x &= \sum_{n=1}^{\infty} A_k x_n \\ &= \underbrace{\sum_{n=1}^{k-1} A_k x_n}_{\|\cdot\| \le M(x_1)+\cdots+M(x_{k+1})} + \underbrace{A_k x_k}_{\|\cdot\| > k+M(x_1)+\cdots+M(x_{k-1})} + \underbrace{\sum_{k+1}^{\infty} A_k x_n}_{\|\cdot\| \le 2^{-k}} \end{aligned}$$

So $\|A_k x\| > k - 1$, which implies that $M(x) = \infty$. This is a contradiction. $\square$

Corollary 7.3. *Let X be a Banach space. If $A \subseteq X^*$ is such that* $\sup\{|L(x)| : L \in A\}$ *for all x, then* $\sup_{L \in A} \|L\| < \infty$.

Corollary 7.4. *Let Y be a normed space. If $A \subseteq Y$ and* $\sup\{|L(a)| : a \in A\} < \infty$ *for all $L \in Y^*$, then* $\sup_{a \in A} \|a\| < \infty$.

Proof. Consider the natural embedding of A into $\hat{A} \subseteq Y^{**}$. $\square$

Corollary 7.5. *Let X be a Banach space, let Y be a normed space, and let $A \subseteq \mathcal{B}(C, Y)$. If* $\sup\{|L(Ax)| : A \in \mathcal{A}\} < \infty$ *for all $x \in X$ and $L \in Y^*$, then A is uniformly bounded.*

Proof. This is a double application of the principle of uniform boundedness. $\square$

Theorem 7.3 (Banach-Steinhaus). *Let X, Y be Banach spaces. Let $(A_n)_n$ be a sequence in $\mathcal{B}(X, Y)$. If for every x, there is a y such that $A_n x \to y$, then*

1. $\sup_n \|A_n\| < \infty$,
2. *There exists some $A \in \mathcal{B}(X < Y)$ such that $A_n x \to Ax$.*

8 Locally Convex Topological Vector Spaces

8.1 Topologies generated by seminorms

Definition 8.1. A **topological vector space (TVS)** over $\mathbb{F}$ is a vector space $(X, \mathcal{T})$ with a topology such that

1. $X \times X \to X$ sending $(x, y) \mapsto x + y$ is continuous,
2. $\mathbb{F} \times X \to X$ sending $(\alpha, x) \mapsto \alpha x$ is continuous.

Let X be a vector space over $\mathbb{F}$, and let $\mathcal{P}$ be a family of seminorms on X. We can use $\mathcal{P}$ to generate a topology (like how we do with norms). We get a base for the topology given by

$$\left\{ \bigcap_{i=1}^{k} \{x : p(x - x_i) < \varepsilon_i\} : p_1, \dots, p_k \in \mathcal{P}, x_1, \dots, x_k \in X, \varepsilon_i, \dots, \varepsilon_k > 0 \right\}.$$

Definition 8.2. A TVS X is a **locally convex space (LCS)** if the topology is is generated by some family $\mathcal{P}$ of seminorms and $\bigcap_{p \in \mathcal{P}} \{x : p(x) = 0\} - \{0\}$ (the seminorms separate points).

Proposition 8.1. *Let X be a TVS, and let p be a seminorm on X. The following are equivelent:*

1. *p is continuous.*
2. *$\{x : p(x) < 1\}$ is open.*
3. *$0 \in \operatorname{int}\{x : p(x) < 1\}$.*
4. *$0 \in \operatorname{int}\{x : p(x) \leq 1\}$.*
5. *p is continuous at 0.*
6. *There is a continuous seminorm q such that $p \leq q$.*

Proof. The first four statements get weaker, so we have (1) $\implies$ (2) $\implies$ (3) $\implies$ (4).

(4) $\implies$ (5): Let $\varepsilon > 0$. Then

$$U = \operatorname{int}\{x : p(x) \leq \varepsilon/2\} = \varepsilon/2 \cdot (\operatorname{int}(\{x : p(x) \leq 1\})).$$

(5) $\implies$ (1): Compose with translations.

(6) $\implies$ (1): Suppose taht $p \leq q$. Then $0 \in \{q < 1\} \subseteq \operatorname{int}\{p < 1\}$. $\square$

Proposition 8.2. *Let $p_1, \dots, p_n$ be continuous seminorms. Then $p_1 + \cdots + p_n$ and $\max_i p_i$ are continuous seminorms.*

Proposition 8.3. *If $(p_i)_i$ is a family of continuous seminorms and $p_i \le q$ for all i, where q is a continuous seminorm, then $\sup_i p_i$ is continuous.*

Example 8.1. Let X be a (Tychonoff)[9] topological space, let $K \subseteq X$ be compact, and let $p_K(f) = \|f|_K\|_{\text{sup}}$. Then $\{p_K : K \subseteq X \text{ compact}\}$ generate a locally convex topology.

On $\mathbb{R}^n$, this topology is generated by $\{p_{\overline{B(0,n)}} : n \in \mathbb{N}\}$.

Example 8.2. Let X be a normed space. For any $f \in X^*$, let $p_f(x) = |f(x)|$. Then X with the resulting LCS structure is called X with the **weak topology**.

8.2 Convex sets

Definition 8.3. Let X be a vector space, and let $A \subseteq X$. The **convex hull** of A is

$$\operatorname{co} A := \bigcap \{C : C \supseteq A, C \text{ convex}\}.$$

The **closed convex hull** of A is

$$\overline{\operatorname{co}} A := \bigcap \{C : C \supseteq A, C \text{ convex and closed}\}.$$

Proposition 8.4. *$\overline{\operatorname{co}} A = \overline{\operatorname{co} A}$.*

Proof. ($\supseteq$): The left hand side, closed, convex and contains A.

($\subseteq$): It suffices to show that $\overline{\operatorname{co} A}$ is convex. Consider $c = ta + (1-t)b$ for a $a, b \in \overline{\operatorname{co} A}$ and $0 < t < 1$. Consider $F : X \times X \to X$ given by $(x, y) \mapsto tx + (1-t)y$; F is continuous. Then for any neighborhood $W \ni c$, there is a neighborhood $W' \ni (a, b)$ such that $F[W'] \subseteq W$. By the definition of the product topology, we can find a neighborhood $U \times V \subseteq W'$ with the same property. Now pick $a' \in U \cap \operatorname{co} A$ and $b' \in B \cap \operatorname{co} A$. Now $F(a', b') \in W \cap \operatorname{co} A$. So $c \in \overline{\operatorname{co} A}$, as desired. $\square$

8.3 Correspondence between nice convex sets and seminorms

Definition 8.4. Let X be a vector space over $\mathbb{F}$, and let $A \subseteq X$ be convex.

1. A is **balanced** if $\alpha A \subseteq A$ for all $\alpha \in \mathbb{F}$ and $|\alpha| \le 1$.

2. A is **absorbing** if for all $x \in X$, there is a $\beta \in (0, \infty)$ such that $x \in \beta A$.

3. A is **absorbing at** $a \in A$ if $A - a$ is absorbing.

Proposition 8.5. *Let X be a vector space over $\mathbb{F}$. If V is a nonempty, balanced, convex set which is absorbing at all its points, then there is a unique seminorm on X such that $V = \{x : p(x) < 1\}$.*

[9] This means that it is Hausdorff and whenever $x \imath n X$ and $A \subseteq X$ is closed, there is an $f \in C(X)$ such that $f(x) = 0$ and $f|_A = 1$. If X is not Tychonoff, this still works, but the space is actually very small.

Proof. Define $p(x) := \inf\{t \geq 0 : x \in tV\}$. This is called the **Minkowski functional** of V. Then p is a seminorm:

- (homogeneity): $p(\alpha x) = \inf\{t : \alpha x \in tV\} = |\alpha| \inf\{t : \frac{\alpha}{|\alpha|} x\} = |\alpha| p(x)$.
- (subadditivity): If x, y, suppose $x \in tV$ and $y \in sV$. Then $x + y \in tV + sV = (t+s)V$ (by convexity). So if $p(x) \leq t$ and $p(y) \leq s$, then $p(x+y) \leq t + s$.

If $p(x) < 1$, then $x \in tV$ for some $t < 1$. Because V is balanced, $V \supseteq tV$, so $x \in V$. This gives $\{p < 1\} \subseteq V$.

Conversely, suppose $x \in V$. Then $p(x) \leq 1$. Since V is absorbing at x, there exists some $\varepsilon > 0$ such that $x + \varepsilon x \in V$. So $p(x) \leq 1/(1+\varepsilon) < 1$. This gives $V \subseteq \{p < 1\}$.

Uniqueness: if seminorms satisfy $\{p < 1\} = \{q < 1\}$, then $p = q$ (from lecture 1). □

Corollary 8.1. *A TVS is a LCS if and only if the collection of convex, balanced sets absorbing all their own points is a neighborhood base at 0.*

Proposition 8.6. *A LCS is generated by a translation invariant metric if and only if it is generated by a countable family of seminorms.*

Proof. If $(p_n)_{n=1}^{\infty}$ is a sequence of seminorms, then we can define

$$d(x, y) := \sum_{n=1}^{\infty} 2^{-n} \frac{p_n(x-y)}{1 + p_n(x-y)}.$$

□

Definition 8.5. A convex set A in a TVS X is **bounded** if for any neighborhood U of 0, there is a $t < \infty$ such that $tU \supseteq A$.

Theorem 8.1. *A LCS is normable if and only if it has a bounded, open neighborhood of 0.*

9 Metrizibility and Normability of LCSs and The Geometric Hahn-Banach Theorem

9.1 Metrizable locally convex spaces

When is a LCS topology metrizable?

Theorem 9.1. *Let X be a LCS. Then X is metrizable (with a translation invariant metric) if and only if its topology can be generated by a countable family of seminorms.*

Proof. Suppose the topology is generated by $(p_n)_n$. Define

$$d(x,y)\sum_{n=1}^{\infty}2^{-n}\frac{p_n(x-y)}{1+p_n(x-y)}.$$

For every $\varepsilon>0$ and $N\in\mathbb{N}$, there is a $\delta>0$ such that

$$\{y:d(x,y)<\delta\}\subseteq\bigcap_{n=1}^{N}\{y:p_n(x-y)<\varepsilon\}.$$

Conversely, for any $\varepsilon>0$ and $N\in\mathbb{N}$ such that

$$\{y:d(x,y)<\delta\}\supseteq\bigcap_{n=1}^{N}\{y:p_n(x-y)<\varepsilon\}.$$

Now assume d is a translation invariant metric generating the topology of X. Then $\{x : d(0,x)<1/n\}$ for $n\in\mathbb{N}$ form a neighborhood base at 0. Let $\mathcal{P}$ be any family of seminorms generating the topology. Then for any n, there exist seminorms $p_{n,1},\ldots,p_{n,N_n}\in\mathcal{P}$ and $\varepsilon_n>0$ such that

$$\bigcap_{i=1}^{n}\{x:p_{n,i}(x)<\varepsilon_n\}\subseteq\{x:d(0,x)<1/n\}.$$

Now $\mathcal{P}_0=\bigcup_{n=1}^{\infty}\{p_{n,1},\ldots,p_{n,N_n}\}$ is countable and generates the same topology. $\square$

Example 9.1. $C(\mathbb{R}^n)$ has the metric

$$d(f,g):=\sum_{n=1}^{\infty}2^{-n}\frac{\|f|_{B_n}-g|_{B_n}\|_\infty}{1+\|f|_{B_n}-g|_{B_n}\|_\infty}.$$

Definition 9.1. A TVS is a **Fréchet space** if its topology can be generated by a complete, translation invariant metric.

9.2 Normable locally convex spaces

When does a LCS have a norm?

Definition 9.2. $A \subseteq X$ is **bounded** if for any neighborhood $U \ni 0$, there is an $\varepsilon > 0$ such that $U \supseteq \varepsilon A$.

Theorem 9.2. *A LCS X is normable if and only if it has a nonempty, open, bounded neighborhood of 0.*

Proof. Let B be a nonempty, open, bounded subset $B \ni 0$. By openness, there is a continuous seminorm p such that $B \supseteq \{p < \varepsilon\}$ for some ε. We can assume that $B \supseteq \{p < 1\}$. We must show that p generates the topology. Let q be another continuous seminorm on X, and consider $\{q < \delta\}$. By boundedness, there exists some $\varepsilon > 0$ such that $\varepsilon\{p < 1\} = \{p < \varepsilon\} \subseteq \{q < \delta\}$. So p generates the topology. Since an LCS must separate points, p must actually be a norm. $\square$

9.3 The geometric Hahn-Banach theorem

Since continuous linear functionals make sense for LCS spaces, we still denote the dual space as X^*. It will have a topology, but we will not discuss which topology yet.

Proposition 9.1. *Let $f : X \to \mathbb{F}$ be a linear functional. The following are equivalent:*

1. *f is continuous.*
2. *f is continuous at 0.*
3. *f is continuous at some point.*
4. *$\ker f$ is closed*
5. *$x \mapsto |f(x)|$ is a continuous seminorm.*

If X is an LCS generated by $\mathcal{P}$, then also iff

6. *There exist $p_1, \dots, p_n \in \mathcal{P}$ and $\alpha_1, \dots, \alpha_n \in [0, \infty)$ such that $|f| \leq \sum_{i=1}^n \alpha_i p_i$.*

Proof. (5) $\implies$ (2): f is continuous at 0 iff for every $\varepsilon > 0$, the set $\{x : |f(x)| < \varepsilon\}$ is a neighborhood of 0.

(5) $\implies$ (6): For any $\varepsilon > 0$, there exist $p_1, \dots, p_n \in \mathcal{P}$ and $\beta_1, \dots, \beta_n > 0$ such that $\{|f| < \varepsilon\} \supseteq \bigcap_{i=1}^n \{p_i < \beta_i\}$. So $|f| < \frac{\varepsilon}{\sum_i \beta_i} \sum_i p_i$. $\square$

Proposition 9.2. *Let X be a TVS, and let $G \subseteq X$ be an open, convex neighborhood of 0. Then $q(x) := \inf\{t \geq 0 : tG \ni x\}$ is a nonnegative continuous sublinear functional (and $G = \{q < 1\}$).*

Theorem 9.3 (Geometric Hahn-Banach theorem)**.** *Let X be a TVS, and let $G \subseteq X$ be a nonempty, open, convex set with $G \not\ni 0$. Then there is a closed hyperplane $M \subseteq X$ such that $M \cap G = \varnothing$.*

Proof. Suppose $\mathbb{F} = \mathbb{R}$. Let $x_0 \in G$, and let $H := G - x_0$ be an open, convex neighborhood of 0. Then $0 \in H$, but $-x_0 \notin H$; as H is convex, $tH \not\ni -x_0$ for any $0 \leq t < 1$. Let $q(x) := \inf\{t \geq 0 : tH \ni x\}$ as in the proposition. Then $q(-x_0) \geq 1$. Now let $Y = \operatorname{span}\{-x_0\}$. Then $g : Y \to \mathbb{R}$ with $g(-x_0) = 1$ is a continuous linear functional, and Hahn-Banach gives a linear $f : X \to \mathbb{R}$ such that $f(-x_0) = 1$, $|f| \leq q$; so f is continuous. Now $\{f = 1\} \cap H = \varnothing$, so $\ker(f) \cap G = \varnothing$. So pick $M = \ker(f)$.

In the case $\mathbb{F} = \mathbb{C}$, applied the theorem to X (viewed as a vector space over $\mathbb{R}$). We get a continuous $\mathbb{R}$-linear $f : X \to \mathbb{R}$ such that $\ker(f) \cap G = \varnothing$. Construct $g(x) := f(x) - if(ix)$, which is a complex linear functional. Then $\ker g = (\ker f) \cap i(\ker f)$. □

Corollary 9.1. *Let X is a TVS, $Y \subseteq X$ be a closed affine subspace, and $G \neq 0$ be an open convex subset with $Y \cap G \neq \varnothing$. Then there is a closed affine hyperplane $M \supseteq Y$ such that $M \cap G = \varnothing$.*

Proof. Suppose $0 \in Y$. Consider the quotient map $Q : X \to X/Y$. Then $Q(G)$ is an open, convex subset of X/Y with $Q(G) \not\ni 0$. Find a hyperplane $\overline{M} \subseteq M/Y$ such that $\overline{M} \cap Q(G) = \varnothing$, and let $M := Q^{-1}[M]$.

If $0 \notin Y$, do the same with a translation. □

9.4 Half-spaces and separated sets

Definition 9.3. In a real TVS an open **half-space** is a is a set of the form $\{f > \alpha\}$ for some $f \in X^*$ and $\alpha \in \mathbb{R}$. A closed **half-space** is a is a set of the form $\{f \geq \alpha\}$ for some $f \in X^*$ and $\alpha \in \mathbb{R}$.

Definition 9.4. $A, B \subseteq X$ are **separated** of there exist closed half-spaces H, K such that $A \subseteq H$, $B \subseteq K$, and $H \cap K$ is an affine hyperplane. A and B are **strictly separated** if there are open half-spaces $H \supseteq A$ and $K \supseteq B$ with $H \cap K = \varnothing$.

Theorem 9.4. *Half-spaces and separated sets have the following properties:*

1. *The closure of an open half-space is a closed half-space.*
2. *The interior of a closed half-space is an open half-space.*
3. *If A, B are separated, then there exists an $f \in X^*$ and $\alpha \in \mathbb{R}$ such that $f|_A \leq \alpha$ and $f|_B \geq \alpha$.*
4. *If A, B are strictly separated, then there exists an $f \in X^*$ and $\alpha \in \mathbb{R}$ such that $f|_A < \alpha$ and $f|_B > \alpha$.*

Theorem 9.5. *Let X be a real TVS, and let A, B be disjoint, convex sets with A open. Then there exist an $f \in X^*$ and $\alpha \in \mathbb{R}$ such that $f|_A < \alpha$, $f|_B \geq \alpha$. If B is also open, then A and B are strictly separated.*

We will get this as a consequence of geometric Hahn-Banach next time.

10 Separation Results and Weak Topologies

10.1 Separation results in topological vector spaces

Last time, we had the geometric Hahn-Banach theorem.

Theorem 10.1. *Let X be a real TVS, and let G be a nonempty, open, convex subset with $x \in X \setminus G$. Then there exists*

- *an $f \in X^*$ and $\alpha \in \mathbb{R}$ such that $f(x) = \alpha$ and $f(G) \subseteq (-\infty, \alpha)$,*
- *a closed affine hyperplane $M = \{f = \alpha\}$ such that $x \in M$ and $M \cap G = \varnothing$.*

This separates a point from a convex set. What about separating two convex sets?

Theorem 10.2. *Let X be a real TVS, and let A, B be disjoint convex sets with A open. Then there are an $f \in X^*$ and $\alpha \in \mathbb{R}$ such that $A \subseteq \{f < \alpha\}$ and $B \subseteq \{f \geq \alpha\}$. If B is also open, then $B \subseteq \{f > \alpha\}$ (strict separation).*

Remark 10.1. This proof is difficult to imagine algebraically, but the main idea is only 1 step on top of the previous theorem.

Proof. Let $G := A - B = \{a - b : a \in A, b \in B\}$. This is convex, and we can also write $G = \bigcup_{b \in B}(A - b)$, which shows that G is open. Since $A \cap B = \varnothing$, $0 \notin G$. By the previous theorem, we find $f \in X^*$ such that $f[G] \subseteq (-\infty, 0)$. This set is $f[G] = f(A) - f(B)$. So $\alpha := \sup f[A] \leq \inf[B]$. Then $A \subseteq \{f \leq \alpha\}$ and $B \subseteq \{f \geq \alpha\}$. Because A is open, we can get $A \subseteq \{f < \alpha\}$. If B is open, we can do the same. $\square$

10.2 Separation results in locally convex spaces

Theorem 10.3. *Let X be a real LCS, and let A, B be disjoint, closed, convex subsets. If B is compact, they are strictly separated.*

Lemma 10.1. *Let $K \subseteq X$ be compact, and let $V \supseteq K$ be open. Then there is an open neighborhood $U \ni 0$ such that $K + U \subseteq V$.*

Proof. For each $x \in K$, there is a neighborhood U_x of 0 such that $x + Ux \subseteq V$. Because addition is continuous in X, there is a smaller neighborhood $W_x \ni 0$ such that $W_x - W_x \subseteq U$. Let $x \in K$, and suppose $x \in x_i + W_{x_i}$ By compactness, there exist $x_1, \dots, x_n \in K$ such that $K \subseteq \bigcup_{i=1}^n (x_i + W_{x_i})$. Now take $W := \bigcap_{i=1}^n W_{x_i}$.

Let $x \in K$, and say $x \in x_i + W_{x_i}$. Then $x + W \subseteq x_i + W_{x_i} + W \subseteq x_i + W_{x_i} + W_{x_i} \subseteq V$. $\square$

Corollary 10.1. *If X is an LCS, we may take U to be convex.*

Proof. $B \subseteq X \setminus A$. The lemma gives a convex open $U \ni 0$ such taht $(B+U) \cap A = \varnothing$. The previous version of the theorem gives $f \in X^*$ and $\alpha \in \mathbb{R}$ such that $f[B] + f[U] \subseteq \{f < \alpha\}$ and $A \subseteq \{f \geq \alpha\}$. B is compact, so $f[B]$ is compact; so there exists some $\varepsilon > 0$ such that $f[B] \leq \alpha - \varepsilon$. Also, $f[A] \geq \alpha$. □

Corollary 10.2. *Let X be a real LCS, let A be closed and convex, and let $x \in X \setminus A$. Then x, A are strictly separated.*

Corollary 10.3. *Let X be a real LCS, and let $A \subseteq X$. Then $\overline{\text{co}}A$ is the intersection of all closed half-spaces containing A.*

Corollary 10.4. *Let X be a real LCS, and let $A \subseteq X$. Then $\overline{\text{span}}A$ is the intersection of all closed hyperplanes containing A.*

Remark 10.2. These theorems all hold for complex vector spaces, as well. Here's how we get the complex vector space cases: If $f : X \to \mathbb{C}$. then let $f = \operatorname{Re} f$. Then $f = g(x) - ig(ix)$. In this case, when we say that two sets are separated, we mean the real part of f separates them.

10.3 Weak topologies

Example 10.1. Let (X, Σ, μ) be a finite measure space with no atoms (like the unit interval). Let $L^0(\mu)$ be the space of measurable functions $X \to \mathbb{C}$ with the topology of convergence in measure. So the topology is generated by sets of the form $\{f : \mu\{|f-g| > \varepsilon\} < \varepsilon\}$ for each $g \in L^0$.

There are no opne, convex sets besides the whole space. Assume U is convex with $U \ni 0$. Then $U \supseteq \{f : \mu\{|f| > \varepsilon\} < \varepsilon\}$. Let $n > 1/\varepsilon$. If $1 \leq i \leq n$, define $g_i = n(g\mathbb{1}_{[\frac{i-1}{n}, \frac{i}{n}]})$. Then $g = \frac{1}{n}(g_1, \ldots, g_n) \in U$.

Example 10.2. Let $C(\mathbb{R}^n)$ with the seminorms $p_K(f) := \|f|_K\|$ for all compact $K \subseteq \mathbb{R}^n$. Then if $L \in C(\mathbb{R}^N)^*$, then $|L| \leq \alpha p_K$ for some K. There exists a finite signed (or complex-valued) Borel measure $\mu \in M(K)$ such that $L(f) = \int f \, d\mu$ for all $f \in C(\mathbb{R}^n)$.

Let X be an LCS over $\mathbb{F}$.

Definition 10.1. If $x \in X$ and $x^* \in X^*$, we can write $\langle x, x^* \rangle$ as $x^*(x)$; we may also write $\langle x^*, x \rangle$.[10]

Definition 10.2. The **weak topology** on X is the topology generated by the seminorms $\{|f| : f \in X^*\}$. The **weak* topology** on X^* is the toplogy generated by $\{|\hat{x}| : x \in X\}$, where $\hat{x}(f) := f(x)$.

[10] Conway's textbook says you can write it either way around because of some category theoretic duality. Professor Austin is pretty sure that it is because no one can remember which way it goes.

This is not stronger than the original topology.

Remark 10.3. Some authors refer to $\sigma(X, X^*)$ as the weak topology on X and $\sigma(X^*, X)$ as the weak* topology on X^*.

Theorem 10.4. *Let X be a locally convex space.*

1. $(X, \mathrm{wk})^* = X^*$.
2. $(X^*, \mathrm{wk}^*)^* = X$.

Lemma 10.2. *Let X be any vector space, and let $f, g_1, \dots, g_n$ be linear functionals such that* $\ker(f) \supseteq \bigcap_{i=1}^n \ker g_i$. *Then* $f \in \operatorname{span}\{g_1, \dots, g_n\}$.

Now let's prove the theorem.

Proof.
1. We need to check that $X^* \subseteq (X, \mathrm{wk})^*$. If $f \in X^*$, then $|f|$ is a generating seminorm for the weak topology on X, so $f \in (X, \mathrm{wk})^*$.
2. $(\subseteq)$: This is from the definition of wk*.

   $(\supseteq)$: Suppose $f : X^* \to \mathbb{F}$ is continuous for the wk* topology. Then there exist scalars $\alpha_1, \dots, \alpha_n > 0$ and $x_1, \dots, x_n$ such that $|f| \leq \sum_i \alpha_i |\hat{x}_i|$. Then $\ker(f) \supseteq \bigcap_{i=1}^n \ker \hat{x}_i$. The lemma then tells us that $f = \sum_{i=1}^n \beta_i \hat{x}_i = (\sum_{i=1}^n \beta_i x_i)^\wedge \in X$. $\square$

For clarity, we will use the terms open, closed, and continuous to refer to the original toplogy on a space. We will use the terms weak-open, weak-closed, and weak-continuous to refer to the weak/weak* topology on a space.

11 Weak Closure of Convex Sets, Polars, and Alaoglu's Theorem

11.1 Weak closure of convex sets

Last time, we proved the following theorem:

Theorem 11.1. *Let X be a locally convex space.*

1. $(X, \mathrm{wk})^* = X^*$.
2. $(X^*, \mathrm{wk}^*)^* = X$.

Unlike with normed spaces, we can't just keep constructing duals and duals of duals. We can only construct $(X*, \mathrm{wk}^*)$ and its dual, (X, wk).

Theorem 11.2. *Let $A \subseteq X$. Then $\overline{\mathrm{co}\, A} = \overline{\mathrm{co}\, A}^{\mathrm{wk}}$.*

Proof. $(\subseteq)$: The weak topology has fewer closed sets.

$(\supseteq)$: Suppose $x \notin \overline{\mathrm{co}\, A}$. Then there exist an $f \in X^*$ and $\alpha \in \mathbb{R}$ such that $\operatorname{Re} f[\overline{\mathrm{co}\, A}] \leq \alpha < \operatorname{Re} f(x)$. So $\overline{\mathrm{co}\, A} \subseteq \{\operatorname{Re} f \leq \alpha\}$. □

Corollary 11.1. *If A is convex, $\overline{A} = \overline{A}^{\mathrm{wk}}$.*

Remark 11.1. The weak topology is the weakest topology with all closed, convex sets (in the original topology) still closed.

11.2 Polars and quotients

Definition 11.1. Let $A \subseteq X$. Its **polar** is $A^o = \{f \in X^* : |f(x)| \leq 1 \, \forall x \in A\}$.

Definition 11.2. Let $B \subseteq X^*$. Its **pre-polar** is ${}^oB := \{x \in X : |f(x)| \leq 1 \, \forall f \in B\}$.

Definition 11.3. If $A \subseteq X$, its **bipolar** is ${}^o(A^o)$.

Proposition 11.1. *Let $A \subseteq X$.*

1. *A^o is convex and balanced.*
2. *If $A_1 \subseteq A$, then $A_1^o \supseteq A^o$.*
3. *If $\alpha \in \mathbb{F} \setminus \{0\}$, then $(\alpha A)^o = \alpha^{-1} A^o$.*
4. *$A \subseteq^o A^o$.*
5. *$A^o = ({}^oA^o)^o$.*

Remark 11.2. There is an analogous version of this proposition for pre-polars if we start from $B \subseteq X^*$.

Theorem 11.3. *Let $A \subseteq X$. Then ${}^oA^o$ is the closed, convex, balanced hull of A (i.e. the intersection of all closed, convex, balanced sets containing A).*

Proof. ($\supseteq$): From the proposition, ${}^oA^o$ is closed, convex, and balanced.

($\subseteq$): Suppose there exists some convex, balanced, closed $A_1 \supseteq A$ and $x \in X \setminus A_1$; we need to show that $x \notin {}^oA^o$. Then there exist some $f \in X^*$ and $\alpha \in \mathbb{F}$ such that $\operatorname{Re} f[A_1] \leq \alpha < \operatorname{Re} f(x)$. Since $\operatorname{Re} f[A_1] \ni 0$, $\alpha \geq 0$; we can assume $\alpha > 0$. Since we have the balanced assumption, we can assume $\alpha = 1$.

If $f(x) \in \mathbb{R}$, then we are done, since $x \notin {}^oA^o$. So our only worry is that $f(x) \notin \mathbb{R}$. Then choose $w := \overline{f(x)}/|f(x)|$. Now let $g := wf$. Then $g(x) = \operatorname{Re} f(x)$, and $g[A_1] = f[wA_1] = f[A_1]$. So we can use the argument for when $f(x) \in \mathbb{R}$. $\square$

Definition 11.4. Let X be a locally convex space, and let M be a linear subspace. The **annihilator** of M is $M^\perp := \{f \in X^* : f|_M = 0\}$.

Proposition 11.2. *Let X be a vector space over $\mathbb{F}$, and let M be a linear subspace. Let p be a seminorm on X, and define*

$$\overline{p}(x + M) := \inf\{p(x + y) : y \in M\}.$$

Then the function $\overline{p}$ is a seminorm on X/M. If X is an LCS and $\mathcal{P}$ is the collection of continuous seminorms on X, then $\{\overline{p} : p \in \mathcal{P}\}$ generates the quotient topology on X/M. This is an LCS if M is closed.

Remark 11.3. This doesn't work unless we take $\mathcal{P}$ to be the collection of all continuous seminorms on X. What we need is $\overline{p_1 + p_2} \geq \overline{p}_1 + \overline{p}_2$, so we want a generating family of seminorms that is closed under addition (max is okay, too).

Theorem 11.4. *Let $Q : X \to X/M$ be the quotient map. Define $(X/m)^* \to M^\perp$ sending $f \mapsto f \circ Q$. This is an isomorphism of LCSs.*

Proof. Onto: Let $g \in M^\perp$. Then $g = f \circ Q$ for some linear $f : X/M \to \mathbb{F}$; we need to show that f is continuous. We have that $|g|$ is a continuous seminorm on X. Then $\overline{|g|}(x + M) = |f|(x)|$, so $|f|$ is a continuous seminorm. So f is continuous.

To check that the topologies are the same, he have that $\{f \in (X/M)^* : |f(x+M)| < \varepsilon\}$ corresponds to $\{g \in M^\perp : |g(x)| < \varepsilon\}$. These generate the respective topologies for the domain and codomain. $\square$

Theorem 11.5. *The map $X^* \to M^*$ sending $f \mapsto f|_M$ quotients to $X^*/M^\perp \to M^*$. This is an isomorphism of LCSs.*

Remark 11.4. We want $X^*/M^\perp$ to be Hausdorff, so we want $M^\perp$ to be closed. But $M^\perp$ is always closed, so this is okay.

Proof. Onto: If $g \in M^*$, then there is a continuous seminorm p on X such that $g \leq p$. Now apply Hahn-Banach to extend g to a continuous seminorm bounded by p. $\square$

11.3 Alaoglu's theorem

Theorem 11.6. *Let X be a normed space, and let $B^* = \{f \in X^*; \|f\| \leq 1\}$ be the closed unit ball in the dual space of X. Then B^* is weak*-compact.*

Proof. Consider the map $\varphi : B^* \to \prod_{x \in X, \|x\| \leq 1} \overline{\mathbb{D}}$, where $\mathrm{d} = \{z \in \mathbb{C} : |z| \leq 1\}$, given by $f \mapsto \langle f(x) \rangle_{\|x\| \leq 1}$. We claim that φ is a homeomorphism to a closed subset of $\prod_{\|x\| \leq 1} \overline{\mathbb{D}}$.

We claim that we have

$$\varphi[B^*] = \{\langle \alpha(x) \rangle_{\|x\| \leq 1} \in \prod \overline{\mathbb{D}} : \alpha(x+cy) = \alpha(x) + c\alpha(y) \text{ if } \|x\|, \|y\| \leq 1, c \in \mathbb{F}, \|x + cy\| \leq 1\}.$$

For ($\supseteq$): If α is in the right hand side, define $f(x) := \varepsilon^{-1}\alpha(\varepsilon x)$ for all $x \in X$ with $\varepsilon < 1/\|x\|$. Then $f \in B^*$.

Closed: Suppose $\alpha \notin \varphi[B^*]$. Then there are x, y, c such that

$$|\alpha(x + cy) - \alpha(x) - c\alpha(y)| > \varepsilon > 0.$$

If $|\alpha'(x) - \alpha(x)|$, $|\alpha'(y) - \alpha(y)|$, $|\alpha'(x+cy) - \alpha(x+cy)| < \varepsilon/3$, then $\alpha'(x+cy) \neq \alpha'(x) + c\alpha'(y)$. So $\varphi[B^*]$ is closed.

Check that the topologies agree. $\square$

Theorem 11.7. *For any normed space $\mathcal{X}$, there exists a compact Hausdorff space Z such that $\mathcal{X}$ embeds isometrically as a subspace of $C(Z)$.*

Proof. Let $Z = B^*$. For the mapping, take $x \mapsto \hat{x}|_{B^*}$. $\square$

12 Reflexive Banach Spaces and Metrizability of the Unit Ball in the Weak* Topology

Today's lecture was given by a guest lecturer, Professor Dimitri Shlyakhtenko.

12.1 Reflexive spaces

Let X be a Banach space. For this lecture, we will denote the weak topology by $\sigma(X, X^*)$ and the weak*-topology by $\sigma(X^*, X)$. Lst time we proved the following theorem:

Theorem 12.1 (Alaoglu)**.** *$(X^*)_1$ is $\sigma(X^*, X)$ compact.*

Definition 12.1. X is **reflexive** if $X = X^{**}$.

Example 12.1. For $1 < p < \infty$, L^p and ℓ^p are reflexive.

Proposition 12.1. *$(X)_1 \subseteq (X^{**})_1$ is $\sigma(X^{**}, X^*)$-dense.*

Remark 12.1. $\sigma(X^{**}, X^*)|_X = \sigma(X, X^*)$.

Proof. Take B to be the closure of $(X)_1$ in the $\sigma(X^{**}, X)$ topology. Then $B \subsetneq (X^{**})_1$ as $(X^{**})_1$ is closed. If $x^{**} \in (X^{**})_1 \setminus B$, then by Hahn-Banach (on $(X^{**}, \sigma * X^{**}, X^*)$), there exist an $x^* \in X^*$ and $\alpha \in \mathbb{R}$ such that $\operatorname{Re}\langle x, x^*\rangle < \alpha < \alpha + \varepsilon < \operatorname{Re}\langle x^*, x_0^{**}\rangle$ for all $x \in (X)_1$. So there is an $x^* \in X^*$ such that $\operatorname{Re}\langle x, x^*\rangle < 1 < 1 + \varepsilon < \operatorname{Re}\langle x^*, x_0^{**}\rangle$ for all $x \in X_0$. Then $|\langle x, x^*\rangle| = 1$ if $\|x\| \leq 1$, so $x^* \in (X^*)_1$. We now get

$$1 + \varepsilon < \operatorname{Re}\langle x^*, x_0^{**}\rangle \leq |\langle x^{,} x_0^{**}\rangle| \leq \|x_0^{**}\| \leq 1.$$

This is a contradiction. $\square$

Theorem 12.2. *Let X be a Banach space. The following are equivalent:*

1. *X is reflexive.*
2. *X^* is reflexive.*
3. *$\sigma(X^*, X) = \sigma(X^*, X^{**})$.*
4. *$(X)_1$ is compact in $\sigma(X, X^*)$.*

Proof. (1) $\implies$ (3): This is because $X = X^{**}$.

(1) $\implies$ (4): If $X = X^{**}$, Then $(X)_1 = (X^{**})_1$. So $\sigma(X, X^*) = \sigma(X^{**}, X^*)$, the weak* topology on X^{**}. So $(X)_1$ is compact by Alaoglu's theorem.

(4) $\implies$ (1): Note that $\sigma(X^{**}, X^*)|_X = \sigma(X, X^*)$. Thus, if $(X)_1$ is $\sigma(X, X^*)$ compact, then $(X)_1$ is compact in $\sigma(X^{**}, X^*)$ as a subset of X^{**}. By the proposition, $(X)_1$ is $\sigma(X^*, X^{**})$-dense in X^{**}. And compact implies closed.

(3) $\implies$ (2): By Alaoglu's theorem, $(X^*)_1$ is $\sigma(X^*, X)$-compact. By assumption, $(X^*)_1$ is $\sigma(X^*, X^{**})$-compact. Now apply the argument of (4) $\implies$ (1) to X^*. So X^* is reflexive.

(2) $\implies$ (1): Observe that $(X)_1$ is norm-closed in X^{**} (because this is an isometric inclusion). Therefore, $(X)_1$ is $\sigma(X^{**}, X^{***})$-closed. Assuming (2), $(X)_1$ is $\sigma(X^{**}, X^*)$-closed. Bt the proposition, $(X)_1$ is $\sigma(X^{**}, X^*)$-dense in $(X^{**})_1$. So $(X)_1 = (X^{**})_1$. □

12.2 Additional properties of reflexive spaces

Corollary 12.1. *Let X be a Banach space. If $Y \subseteq X$ is a closed subspace, then Y is reflexive.*

Proof. We have $(Y)_1 = Y \cap (X)_1$. So if X is reflexive, then $(Y)_1$ is (X, X^*)-compact. But then $\sigma(X, X^*)|_Y = \sigma(Y, Y^*)$ (check this using Hahn-Banach). So Y is reflexive. □

Example 12.2. Can you embed $\ell^\infty \subseteq \ell^2$ isometrically? No. Since ℓ^∞ is not reflexive. What about embedding L^2 into L^∞?

Corollary 12.2. *Let X be a Banach space. If X is reflexive, X is **weakly sequentially complete**. That is, any $\sigma(X, X^*)$-Cauchy sequence has a $\sigma(X, X^*)$-limit.*

Proof. Suppose $\{x_n\}_n$ is weakly Cauchy: for all $x \in X^*$, $\{\langle x_n, x^* \rangle\}$ is Cauchy. Then it is bounded. The principle of uniform boundedness implies that there exists some M such that $|\langle x_n, x^* \rangle| \leq M$ for any $x^* \in (X^*)_1$. So $\|x_n\| \leq M$ for all n. Now by reflexivity, $M(X)_1$ is compact in $\sigma(X, X^*)$. So there exists a limit point $x \in M(X)_1$ such that $\langle x_n, x^* \rangle \to \langle x, x^* \rangle$ for all $x^* \in X^*$. Then $x_n \to x$ in the $\sigma(X, X^*)$-topology. □

Example 12.3. Let $X = C([0, 1])$, and let

$$f_n = \begin{cases} -nx + 1 & x \in [0, 1/n] \\ 0 & x \in (1/n, 1]. \end{cases}$$

Then for any signed measure $\mu \in C([0, 1])^*$, $\langle f_n, \mu \rangle = \int f_n(t)\, d\mu(t) \to \mu(\{0\})$, but $f_n \not\to f$ weakly. So be careful.

Definition 12.2. A subspace $Y \subseteq X$ is called **proximal** if for all $x_0 \in X$, there exists some $y_0 \in Y$ such that $\|x_0 - y_0\| = \operatorname{dist}(Y, x_0)$

Corollary 12.3. *Let X be a Banach space. If X is reflexive, then any subspace $Y \subseteq X$ is proximal.*

Proof. The map $x \mapsto \|x - x_0\|$ is $\sigma(X, X^*)$-semicontinous. Then $Y \cap \{x : \|x - x_0\| \leq 2 \operatorname{dist}(x_0, Y)\}$ is a $\sigma(X, X^*)$-compact set (by reflexivity). Semicontinuous functions achieve their minima. □

Proposition 12.2. *If $x^* \in X^*$, then $\ker x^*$ is proximal if and only if there is an $x \in (X)_1$ such that $\langle x, x^* \rangle = \|x^*\|$.*

Example 12.4. Let $L : [0,1] \to \mathbb{C}$ be $\int_0^{1/2} f\,dx - \int_{1/2}^1 f\,dx$. The norm of L is never achieved (you want a step function, but this is not continuous), so $\ker L$ is not proximal.

Theorem 12.3 (James, 60s)**.** *X is reflexive if and only if every closed hyperplane is proximal.*

12.3 Metrizability of the closed unit ball in the weak* topology

Theorem 12.4. *Let X be a Banach space. $(X^*)_1$ is $\sigma(X^*, X)$-metrizable if and only if X is separable.*

Proof. ($\Longleftarrow$): Assume X is separable. Let $\{x_n\}_n$ be a countable dense subset of X. Let $\mathbb{D} = \{z : |z| \leq 1\}$, and let $Y = \prod_{\mathbb{N}} \mathbb{D}$. Then the map $\tau_1^* \to X$ by $\tau(x^*) = \{\langle x^*, x_n \rangle\}_n$ gives a homeomorphism from $((X^*)_1, \sigma(X^*, X)) \to Y$.

($\Longrightarrow$): If $(X^*)_1$ is $\sigma(X^*, X)$-metrizable, then there are open sets $U_n \subseteq (X^*)_1$ with $0 \in U_n$ such that $\bigcap_n U_n = \{0\}$. By the definition of the weak* topology, there exist finites subsets F_n of X such that $\{x^* \in (X^*)_1 : |\langle x^*, x \rangle| < 1\, \forall x \in F_n\} \subseteq U_n$. Let $F = \bigcup_n F_n$.

We claim that F is dense. Then $\overline{F} =^{\perp} (F^{\perp})$. So it is enough to prove that $F^{\perp} = \{0\}$. If $x^{\in} F^{\perp} \setminus \{0\}$, then for all $x \in F_n$,

$$0 = \left| \left\langle \frac{x^*}{\|x^*\|}, x \right\rangle \right| < 1 \implies \frac{x^*}{\|x^*\|} \in U_n \implies \frac{x^*}{\|x^*\|} = 0.$$

So $F^{\perp} = \{0\}$. $\square$

13 The Krein-Milman Theorem and The Markov-Kakutani Fixed Point Theorem

Today's lecture was given by a guest lecturer, Professor Dimitri Shlyakhtenko.

13.1 The Krein-Milman theorem

Definition 13.1. Let $K \subseteq X$ be convex. Then a is an **extreme point** of K if $a \in K$ and if whenever $a = \alpha + (1-\alpha)y$ for some $\alpha \in [0,1]$ and $x, y \in K$, then $\alpha = 0$ or $\alpha = 1$.

So extreme points cannot be on the interior of a line segment in K. The set of extreme points is denoted as $\operatorname{ext}(K)$.

Example 13.1. Suppose $K = \{f \in L^1([0,1]) : \|f\|_1 \leq 1\}$. What are the extreme points of K? If $\|f\|_1 = 1$, then $\int_0^1 |f(t)|\, dt = 1$. The primitive $F(T) = \int_0^T |f(t)|\, dt$ is continuous, so there is a T such that $\int_0^T |f(T)|\, dt = 1/2$. Now define

$$h(t) = \begin{cases} 2f(T) & t \leq T \\ 0 & \text{otherwise,} \end{cases} \qquad g(t) = \begin{cases} 0 & t \leq T \\ 2f(t) & \text{otherwise.} \end{cases}$$

Then $\|h\|_2 = \|g\|_2 = 1$, and $f = \frac{1}{2}h + \frac{1}{2}g$. So there are no extreme points.

Theorem 13.1 (Krein-Milman)**.** *Let X be an LCS, and let K be a nonempty, compact, convex subset. Then $K = \overline{\operatorname{co}}(\operatorname{ext} K)$. In particular, $\operatorname{ext} K \neq \varnothing$.*

Corollary 13.1. *If $B \subseteq X$ is a nonempty, convex subset such that $\operatorname{ext}(B) \neq \varnothing$, then no LCS structure on X makes B compact.*

Corollary 13.2. *If X is a Banach space and $\operatorname{ext}(X)_1 = \varnothing$, then $X \neq Y*$ for any Y.*

Example 13.2. This shows that L^1 is not the dual of anything.

Proposition 13.1. *Let $K \subseteq X$ be convex. The following are equivalent:*

1. *$a \in \operatorname{ext} K$.*
2. *If $a = \frac{1}{2}(x_1 + x_2)$ with $x_1, x_2 \in K$, then $x_1 = x_2$.*
3. *If $x_1, \ldots, x_k \in K$ and $a \in \operatorname{co}\{x_1, \ldots, x_k\}$, then $a = x_j$ for some j.*
4. *$K \setminus \{a\}$ is convex.*

Here is the idea of the proof of the Krein-Milman theorem: Look for maximal (non-trivial) relatively open convex subsets (and hope that these are the same as $\{K \setminus \{a\} : a \in \operatorname{ext} K\}$).

Proof. We want to use Zorn's lemma. Let $\mathcal{U} = \{U \subseteq K : U \text{ rel. open, convex}, U \neq \varnothing, U \neq K\}$. This is nonempty and ordered by inclusion. Assume that $\mathcal{U}_0 \subseteq \mathcal{U}$ is a chain. Let $U_0 = \bigcup_{U \in \mathcal{U}_0} U$; this is open (as a union of open sets), and it is convex.[11] U_0 is nontrivial, as well: if $U_0 = K$, then $\mathcal{U}_0$ is an open cover for K, which means that $K \subseteq U$ for some $U \in \mathcal{U}_0$, This is a contradiction.

By Zorn's lemma there exists a maximal element $U \in \mathcal{U}$. Let $x \in L$ and $\lambda \in [0, 1]$. Define $T_{x,\lambda} : K \to K$ by $T_{x,\lambda}(y) = \lambda y + (1 - \lambda)x$. This is continuous and **affine** (i.e. $T_{x,\lambda}(\sum_j \alpha_j y_j) = \sum_j \alpha_j T(y_j)$ if $\alpha_j \geq 0$ and $\sum_j \alpha_j = 1$).

We claim that if $\lambda < 1$ and $x \in U$, then $T_{x,\lambda}(U) \subseteq U$. Thus, $U \subseteq T_{x,\lambda}^{-1}(U)$, which is an open, convex set. If $y \in \overline{U} \setminus U$, then $T_{x,\lambda}(y) \in [x, y) \subseteq U$. So if $\overline{U} \subseteq T_{x,\lambda}^{-1}(U)$, then $T_{x,\lambda}^{-1}(U) = K$. Thus, $T_{x,\lambda}(K) \subseteq U$ for all $x \in U$ and $\lambda \in [0, 1)$.

We claim that if $V \subseteq K$ is open and convex, then $V \cup U = U$ or $V \cup U = K$. This is because $V \cup U$ is open, and the conclusion above implies that $V \cup U$ is convex. If $V \cup U \neq K$, then $V \cup U \subseteq U$ by maximality.

We now claim that $K \setminus U$ is one point. If $a, b \in K \setminus U$ and $a \neq b$, then choose disjoint, open, convex subsets $V_a, V_b \subseteq K$ with $a \in V_a, b \in V_b$. Then $V_a \cup U \neq U$, so $V_a \cup U = K$. However, this implies $b \in V_a \cap V_b$, which gives a contradiction.

We now claim that if $V \subseteq X$ is open, convex, and $\operatorname{ext} K \subseteq V$, then $K \subseteq V$: Suppose not, so there exists an open, convex $V \subseteq X$ such that $\operatorname{ext} K \subseteq V$ by $V \cap K \neq K$. Then $V \cap K \subseteq \mathcal{U}$, so there is a maximal $U \in \mathcal{U}$ such that $V \cap K \subseteq U = K \setminus \{a\}$ and $a \in \operatorname{ext}(K)$. Then $a \notin V$, which is a contradiction.

To finish the proof: Let $E = \overline{\operatorname{co}}(\operatorname{ext} K)$. If $x^* \in X^*$, $\alpha \in \mathbb{R}$, and $E \subseteq \{x \in X : \operatorname{Re}\langle x, x^* \rangle < \alpha\} = V$, then $K \subseteq V$. Hahn-Banach says that E is the intersection of such sets V. So $E \supseteq K$. $\square$

Here is another theorem. This is

Theorem 13.2. *Let X be an LCS, and let $X \subseteq K$ be compact, and convex. Assume that $F \subseteq K$ is such that $K = \overline{\operatorname{co}}(F)$. Then $\operatorname{ext}(K) \subseteq \overline{F}$.*

13.2 The Markov-Kakutani fixed point theorem

Fixed point theorems allow us to show the existence of desired objects by expressing them as a fixed point of some map(s).

Theorem 13.3 (Markov-Kakutani fixed point theorem)**.** *Let $K \subseteq X$ be a nonempty, compact, convex set. Let $\mathcal{F}$ be a family of affine maps $K \to K$ which is **abelian** ($ST = TS$ for all $S, T \in \mathcal{F}$). Then there exists a fixed point $x_0 \in K$ such that $T(x_0) = x_0$ for all $T \in \mathcal{F}$.*

[11] It also has a hilarious notation.

Proof. Let $T \in \mathcal{F}$. Define $T^{(n)} = \frac{1}{n}\sum_{k=0}^{n-1} T^k$. Then $T^{(n)}$ is a again an affine map taking $K \to K$. If $S, T \in \mathcal{F}$, then $S^{(n)}, T^{(m)}$ commute for all n, m. Let $\mathcal{K} = \{T^{(n)}(K) : T \in \mathcal{F}, n \geq 1\}$, which is a collection of compact, convex sets. If $T_1, \ldots, T_p \in \mathcal{F}$ and $n_1, \ldots, n_p \geq 1$, then

$$T_1^{(n_1)} \circ \cdots \circ T_p^{(n_p)}(K) \subseteq \bigcap_{j=1}^{p} T_j^{(n_j)}(K).$$

These are arbitrary elements of $\mathcal{K}$, then $\mathcal{K}$ has the finite intersection property. So there exists an $x_0 \in \bigcap_{K' \in \mathcal{K}} K'$.

We claim that x_0 is the desired fixed point. Take $t \in \mathcal{F}$, and let $n \geq 1$. Then $x_0 \in T^{(n)}(K)$, so $x_0 = T^{(n)}(x)$ for some x. In particular,

$$x_0 = \frac{1}{n}x + T(x) + \cdots + T^{n-1}(x)).$$

Applying T, we get

$$T(x_0) = \frac{1}{n}(T(x) + \cdots + T^{n-1}(x) + T^n(x)).$$

Subtracting this, we get

$$T(x_0) - x_0 = \frac{1}{n}(T^n(x) - x) \in \frac{1}{n}(K - K),$$

where $K - K$ is compact. This is true for any n. If U is an open neighborhood of 0, then there exists some n such that $\frac{1}{n}(K - K) \subseteq U$. Then $T(x_0) - x_0 \in U$ for all open neighborhoods U of 0, so $T(x_0) = x_0$. $\square$

14 Adjoints

14.1 Adjoints of linear maps

If $T : X \to Y$ is a linear map, then $f \mapsto f \circ T$ is a lienar operator on linear functionals. If T is bounded, then $f \circ T$ is continuous, so this restricts to a linear map $T^* : Y^* \to X^*$.

Definition 14.1. T^* is called the **adjoint** of T.

Proposition 14.1. *If T is bounded, then $\|T^*\| \leq \|T\|$.*

Proof.

$$\begin{aligned}\|T^* f\| &= \sup\{|T^* f(x)| : \|x\|_X \leq 1\}\}\\ &= \sup\{|f(Tx)| : \|x\|_X \leq 1\}\}\\ &\leq \|f\|\|T\|.\end{aligned}$$

□

Proposition 14.2. *Let X, Y be normed spaces, and let $T : X \to Y$ be linear. The following are equivalent:*

1. *T is bounded.*
2. *$f \circ T \in X^*$ for all $f \in Y^*$.*
3. *T is continuous $(X, \mathrm{wk}) \to (Y, \mathrm{wk})$.*

Proof. (1) $\implies$ (2): This is because T is continuous.

(2) $\implies$ (3): Consider

$$\begin{aligned}T^{-1}\left[\bigcap_{i=1}^{m}\{y : |\langle f_i, y\rangle| < \varepsilon_i\}\right] &= \bigcap_{i=1}^{n}\{x : |\langle f_i, Tx\rangle| < \varepsilon_i\}\\ &= \bigcap_{i=1}^{n}\{x : |\langle f_i \circ T, x\rangle| < \varepsilon_i\}\end{aligned}$$

(3) $\implies$ (1): We must show that $T[B_X] \subseteq MB_Y$ for some $M < \infty$. Given $f \in X^*$, consider

$$f[T[B_X]] = \{f(Tx) : \|x\|_X \leq 1\}$$

We know that there is a weak neighborhood $U \ni 0_X$ such that $T[U] \subseteq \{y : |f(y)| < 1\}$. The weak topology is weaker than the norm topology, so there exists some $\varepsilon > 0$ such that $T[B_X] \subseteq \{y : |f(y)| < 1/\varepsilon\}$. So $|f[T[B_X]]| \leq 1/\varepsilon$. □

Proposition 14.3. *Adjoints have the following properties:*

1. $(\alpha A + \beta B)^* = \alpha A^* + \beta B^*$ *for all* $A, B \in \mathcal{B}(X, Y)$.

2. *If* $T \in \mathcal{B}(X, Y)$, *then* T^* *is continuous from* (Y^*, wk^*) *to* (X^*, wk^*).

Remark 14.1. Riesz representation gives $H^* \cong H$ via $L_h := \langle \cdot, h \rangle \mapsto h$, which is conjugate-linear in h. So $L_{\alpha h} = \overline{\alpha} \cdot L_h$.

If $H = \mathbb{C}^n$, then A is represntaed by $[a_{i,j}] \in M_{n,n}(\mathbb{C})$. Then $H^* \cong \mathbb{C}^n$, so A^* is represented by $[a_{j,i}]$. But A^* on H itself is represented by $[\overline{a}_{j,i}]$.

Proposition 14.4. *Let* X, Y *be Banach, and let* $A \in \mathcal{B}(X, Y)$.

1. $A^{**}|_X = A$.

2. $\|A^*\| = \|A\|$.

3. *If* A *is invertible, so is* A^*, *and* $(A^*)^{-1} = (A^{-1})^*$.

4. *If* $B \in \mathcal{B}(Y, Z)$, *then* $(BA)^* = A^* B^*$.

Proof. For (2), we need to show that $\|A^*\| \geq \|A\|$. We know that $\|A^{**}\| \leq \|A^*\| \leq \|A\|$. Since A^{**} is an extension of A to a larger space, $\|A^{**}\| \geq \|A\|$. So these are all equal. $\square$

Example 14.1. Let $1 < p, p' < \infty$. Consider an operator $L^p(\mu) \to \mu^{p'}(\nu)$ given by

$$Tf(y) = \int f(x) K(x, y) \, d\mu(x).$$

Then $T^* : L^{q'}(\nu) \to L^q(\mu)$. For $g \in L^{q'}(\nu)$ and $f \in L^p(\mu)$,

$$\langle T^* g, f \rangle = \langle g, Tf \rangle = \iint g(y) f(x) K(x, y) \, d\mu(x) \, d\nu(y).$$

So

$$T^* g(y) = \int (y) K(y, x) \, d\nu(y),$$

If we are in a Hilbert space, we may want to do $\langle f, g \rangle = \int f \overline{g} \, d\mu$ instead.

Proposition 14.5. *Let* $A \in \mathcal{B}(X, Y)$. *Then* $\ker A^* = (\operatorname{ran} A)^\perp$, *and* $\ker A = {}^\perp(\operatorname{ran} A^*)$.

Proof. We prove the second one; the first is similar. We have

$$\begin{aligned} x \in \ker A &\iff Ax = 0 \\ &\iff \langle Ax, y^* \rangle = 0 \qquad \forall y^* \in Y^* \\ &\iff \langle x, A^* y^* \rangle = 0 \qquad \forall y^* \in Y^* \\ &\iff x \in {}^\perp(\operatorname{ran} A^*). \end{aligned}$$

$\square$

Proposition 14.6. *Let $A \in \mathcal{B}(X,Y)$. Then A is invertible if and only if A^* is invertible.*

Proof. ($\Longleftarrow$): If $\ker A^* = 0$, then $\operatorname{ran} A$ is dense. If $\operatorname{ran} A^* = X^*$, then $\ker A = \{0\}$. To finish, we need to show that $\operatorname{ran} A$ is closed. This follows because if $y = Ax$, then

$$\begin{aligned} \|Ax\| &= \sup\{|f(Ax)| : \|f\|_{Y^*} \leq 1\} \\ &= \sup\{|A^*f(x)| : \|f\|_{Y^*} \leq 1\} \\ &= \sup\{|g(x)| : f \in A^*[B_{Y^*}]\} \end{aligned}$$

For some $c > 0$,

$$\begin{aligned} &\geq \sup\{|g(x)| : g \in cB_{X^*}\} \\ &= c\|x\|_X. \end{aligned}$$

So $\operatorname{ran} A$ is closed. □

14.2 The Banach-Stone theorem

Example 14.2. Let X, Y be compact, Hausdorff spaces, let $\tau : Y \to X$ be a homeomorphism, and let $\alpha : Y \to S^1$ be continuous. Define $T : C(X) \to C(Y)$ by $Tf(y) = \alpha(y) \cdot f(\tau(y))$. Then T is an isometric isomorphism.

Theorem 14.1 (Banach-Stone)**.** *Any isometric isomorphism $C(X) \to C(Y)$ is of this form.*

The key is to tell you what Banach space structure of $C(X)$ to look at to recover what X is.

We know that T^* is an isometric isomorphism from $M(Y) \to M(X)$.

Proposition 14.7. *Let X be a compact, Hausdorff space.*

1. *Let $X \times S^1 \to M(X)$ send $(x, \alpha) \mapsto x \cdot \delta_x$. This is a homeomorphism from $X \times S^1$ to $(\operatorname{ext}(B_{M(X)}), \mathrm{wk}^*)$.*
2. *Let $X \times \{1\} \to M(X)$ send $x \mapsto \delta_x$. This is a homeomorphism $X \to \operatorname{ext} P(X)$.*

Proof. We prove (1). We must show that $\mu \in B_{M(X)}$ is extreme if and only if $\mu = \alpha\delta_x$ for some α, x. □

15 The Banach-Stone Theorem and Compact Operators

15.1 The Banach-Stone theorem

Lemma 15.1. *Let X be a compact, Hausdorff space. Let $X \times S^1 \to M(X)$ send $(x, \alpha) \mapsto \alpha \cdot \delta_x$. This is a homeomorphism from $X \times S^1$ to* $(\mathrm{ext}(B_{M(X)}), \mathrm{wk}^*)$.

Proof. First, we show that $\alpha\delta_x$ is an extreme point. If $\alpha\delta_x = t\mu_(1-t)\nu$, then the total variation of μ or ν must be 1. So μ and ν are supported on $\{x\}$. Since $\alpha \in S^1$, we must have $\mu = \delta_x$ or $\nu = \delta_x$.

Let $\varphi(x, \alpha) = \alpha\delta_x$. Then

$$\{(x', \alpha') : |\alpha' f(x') - \alpha f(x)| < \varepsilon\} = \varphi^{-1}\left\{\mu : |\int f\, d\mu - \alpha f(x)| < \varepsilon\right\}.$$

So this is continuous.

Injectivity: If $|\alpha\delta_x| = |\alpha'\delta_{x'}|$, then $x = x'$ and $\alpha = \alpha'$.

Finally, assume that $\mu \in \mathrm{ext}(\overline{B_{M(X)}}$. Then $|\mu|$ is a regular positive Borel measure. The **support** K of $|\mu|$ is the set

$$K = \bigcap_{\substack{C \subseteq X \text{ closed} \\ |\mu|(X \setminus C) = 0}}.$$

Then $|\mu|(X \setminus K)$ (because the measure is regular).

We need to show that K is a singleton. Suppose not. Suppose that $\overline{U} \cap \overline{V} = \varnothing$, where μ has positive measure in each. Then there is an $f : X \to [0, 1]$ such that $f|_U = 0$ and $f|_V = 1$. If μ is positive, write

$$\mu = \int f\, d\mu \cdot \frac{f\mu}{\int f\, d\mu} + \int (1-f)\, d\mu \frac{(1-f)\mu}{\int (1-f)\, d\mu}.$$

These two measures are different, which contradicts the fact that μ is an extreme point. For general μ, use $\mu = \frac{d\mu}{d|\mu|}|\mu|$.

This argument shows that $K = \{x\}$. This implies that $\mu = \alpha\delta_x$ for some $\alpha \in S^1$. □

Theorem 15.1 (Banach-Stone)**.** *Any isometric isomorphism $C(X) \to C(Y)$ is of the form $Tf(y) = \alpha(y) f(\tau(y))$, where $\tau : Y \to X$ is a homeomorphism and $\alpha : Y \to S^1$.*

Proof. The adjoint $T^* : M(Y) \to M(X)$ resitrcts to a continuous map $(\mathrm{ext}\, \overline{B}_{M(Y)}, \mathrm{wk}^*) \to (\mathrm{ext}\, \overline{B}_{M(X)}, \mathrm{wk}^*)$. By the lemma, we have a continuous map $Y \times S^1 \to X \times S^1$. We can view $Y = Y \times \{1\} \subseteq Y \times S^1$ and same for X. Then $T^*(\delta_y) = \alpha(y) \cdot \delta_{\tau(y)}$ for some α, τ, both continuous. Moreover, τ must be invertible. Now we have

$$Tf(y) = \langle Tf, \delta_y \rangle = \langle f, T^* \delta_y \rangle = \alpha(y) f(\tau(y)),$$

as desired. □

15.2 Compact operators

Let X, Y be Banach spaces.

Definition 15.1. $A \in \mathcal{B}(X,Y)$ is **compact** if any of the following equivalent statements hold:

- $\overline{A(\overline{B}_X)}$ is norm compact.
- $\overline{A(\overline{B}_X)}$ is totally bounded.
- For any bounded sequence $(x_n)_n$ in X¡ $(Tx_n)_n$ has a norm-Cauchy subsequence.

Example 15.1. If $\dim \operatorname{ran}(A) < \infty$, then A is compact.

Proposition 15.1. *If* $\dim X = \infty$*, then* Id_X *is not compact.*

Proof. Assume (towards a contradiction that B_X is compact and hence totally bounded. So there exist$x_1, \dots, x_n \in X$ such that $\overline{B}_X \subseteq \bigcup_{i=1}^n B(X_i, 1/2)$. Then let $y \in \overline{B}_X$ and $z \in \operatorname{span}\{x_1, \dots, x_n\}$ be such that $\|y - z\| < (1 = \varepsilon \operatorname{dist}(y, M) > 0$. Then $\|y - z\| < (1+\varepsilon)\operatorname{dist}(y-z, M)$. So $1 < (1+\varepsilon)\operatorname{dist}(\frac{y-z}{\|y-z\|}M)$; i.e.$\operatorname{dist}(\frac{y-z}{\|y-z\|}M) > 1/(1+\varepsilon) > 1/2$. $\square$

Theorem 15.2. *Let X, Y be Banach spaces, and let $A \in \mathcal{B}(X,Y)$. Then A is compact if and only if A^* is compact.*

Proof. ($\implies$): Let $(f_n)_n \in \overline{B_{Y^*}}$. Observe that

$$\|A^* f\| = \sup\{|f(Ax) : x \in \overline{B_X}\} = \|f|_{\overline{A(\overline{B_X})}}\|_\infty.$$

If $f \in \overline{B_{Y^*}}$, then f is 1-Lipschitz and bounded by 1 on the compact space $\overline{A(\overline{B_X})}$. So $\{f|_{\overline{A(\overline{B_Y})}} : f \in \overline{B_{Y^*}}\}$ is norm-compact. So there is a Cauchy subsequence in $(A^* f_n)_n$ by Arzelà-Ascoli.

($\impliedby$): $A^{**}|_X = A$. $\square$

Definition 15.2. Denote $\mathcal{B}_0(X,Y)$ as the collection of compact operators $X \to Y$ and $\mathcal{B}_{00}(X,Y)$ as the collection of finite-rank operators $X \to Y$.

Proposition 15.2. *Let X, Y be Banach spaces.*

1. *$\mathcal{B}_0$ is a closed subspaces of $\mathcal{B}(X,Y)$.*
2. *Suppose $A : X \to Y$ and $B : Y \to Z$ are bounded. If either A or B is compact, then BA is compact.*

Proof. Suppose $A \in \overline{\mathcal{B}_0(X,Y)}$; we want to show that A is compact. Consider $A(\overline{B_X}) \subseteq B_\varepsilon(C(\overline{B_X}))$. For every $\varepsilon > 0$, there is a $C \in \mathcal{B}_0(X,Y)$ such that $\|A - C\|_{\text{op}} < \varepsilon$. So we can cover $A(\overline{B_X})$ with finitely many balls of radius 2ε.

Assume A is compact, then $A(\overline{B_X})$ is totally bounded, and $B(A(\overline{B_X})) \subseteq B(\overline{A(\overline{B_X})})$. □

Corollary 15.1. *$\mathcal{B}_0(X)$ is an ideal in $\mathcal{B}(X)$. So $\mathcal{B}_0(X)$ is an algebra.*

Example 15.2. Let (X, Σ, μ) be a measure space, an dlet $k \in L^2(X \times X, \mu \times \mu)$. Then define the kernel operator

$$Kf(y) = \int k(x,y)f(x)\,d\mu(x).$$

Then $K \in \mathcal{B}(L^2(\mu), L^2(\mu))$, and $\|K\|_{\text{op}} \leq \|k\|_{L^2(\mu\times\mu)}$.

K is comapct because for all $\varepsilon > 0$, there exist $\varphi_1, \ldots, \varphi_n \in L^2(\mu)$ and $\psi_1, \ldots, \psi_n \in L^2(\mu)$ such that

$$\left\| k(x,y) - \sum_{i=1}^n \varphi_i(x), \psi_i(y) \right\|_{L^2} < \varepsilon.$$

So a finite rank approximation gives us that K is compact.

It is not always true that we can approximate by finite rank operators, but the counterexamples tend to be complicated.

Theorem 15.3. *Let X be a compact Hausdorff space. Then the space $\mathcal{B}_{00}(C(X))$ is dense in $\mathcal{B}_0(C(X))$.*

Proof. Assume $A(\overline{B_{C(X)}})$ is totally bounded. Pick $\varepsilon > 0$, and let $U_1, \ldots, U_n$ be an cover X with $x_i \in U_i$. For any $f \in \overline{B_{C(X)}}$ and $x \in U_i$, we have $|Af(x) - Af(x_i)| < \varepsilon$. There exists a partition of unity: $\varphi_1, \ldots, \varphi_n$ with $0 \leq \varphi_o \leq 1$ such that $\varphi_i|_{U_i^c} = 0$ ad $\sum_{i=1}^n \varphi_i = 1$. Define

$$A_\varepsilon f(x) := \sum_{i=1}^n Af(x_i) \cdot \varphi_i(x).$$

This is finite rank because it takes values in the span of the φ_i. We then have

$$|Af(x) - A_\varepsilon f(x)| \leq \sum_{i=1}^n |Af(x) - Af(x_i)|\,\varphi_i(x) < \varepsilon.$$ □

16 Adjoints and Hermitian Operators on Hilbert Spaces

Today's lecture was given by a guest lecturer, Professor Sorin Popa.

16.1 Sesquilinear forms and adjoints

If $T \in \mathcal{B}(X, Y)$, we have the adjoint operator $T^* \in \mathcal{B}(Y^*, X^*)$. If H, K are Hilbert spaces, then $H^* \cong \overline{H}$, the conjugate of H (i.e. H itself). So if $T \in \mathcal{B}(H, K)$, we get $T^* \in \mathcal{B}(K, H)$.

Definition 16.1. A **sesquilinear form** is a function $u : H \times K \to \mathbb{C}$ which is linear in the first variable, antilinear in the second variable, and bounded (as a bilinear map): $|u(\xi, \eta)| \leq C\|\xi\|_H \|\eta\|_K$ for all $\xi \in H$ and $\eta \in K$.

Example 16.1. Let $A \in \mathcal{B}(H, K)$ and $B \in \mathcal{B}(K, H)$. Then $u_A(\xi, eta) = \langle A\xi, \eta \rangle_K$ and $u_B(\xi, \eta) = \langle \xi, B\eta \rangle_H$ are sesquilinear.

Theorem 16.1. *Let H, K be Hilbert spaces. If $u : H \times K \to \mathbb{C}$ is sesquilinear and bounded by C, then there exist unique $A \in \mathcal{B}(H, K)$ such that $u = u_A = u_B$ with $\|A\|, \|B\| \leq K$.*

Remark 16.1. In fact, $\|u\| = \|A\| = \|B\|$.

Proof. For each $\xi \in H$, let $L_\xi : K \to \mathbb{C}$ with $L_\xi(\eta) = \overline{u(\xi, \eta)}$. This is linear, and $|L_\xi(\eta)| \leq C\|\xi\|\|\eta\| =: C_\xi\|\eta\|$ for all η, so $L_\xi \in K^*$. By Riesz representation, there is an $f \in K$ with $\|f\| \leq C\|\xi\|$ such that $L_\xi(\eta) = \langle \eta, f \rangle$. Thus, $A : H \to K$ defined by $A(\xi) = f$ is linear: $A(\alpha_1\xi_1 + \alpha_2\xi_2) = \alpha_1 A(\xi_1) + \alpha_2 A(\xi_2)$ by the uniqueness in the Riesz representation theorem. We also have $\|A(\xi)\| \leq C\|\xi\|$, so A is bounded. $\square$

Definition 16.2. If $A \in \mathcal{B}(H, K)$, the unique $B \in \mathcal{B}(K, H)$ that satisfies $u_A(\xi, \eta) = \langle A\xi, \eta \rangle_K = u_B(\xi, \eta) = \langle \xi, B\eta \rangle_H$ is called the **adjoint** of A (denoted A^*).

Proposition 16.1. *$u \in \mathcal{B}(H, K)$ is an isomorphism of Hilbert spaces if and only if u is invertible and $u^{-1} = u^*$.*

Proof. We have that
$$\|u\xi\|^2 = \langle u\xi, u\xi \rangle = \langle u^*u\xi, \xi \rangle = \langle \xi, \xi \rangle$$
for all $\xi \in H$ if and only if $u^*u = 1$. Since u is invertible, $u^* = u^{-1}$. $\square$

Proposition 16.2. *Let $A, B \in \mathcal{B}(H, K)$, and let $C \in \mathcal{B}(K, K')$.*

1. *$(\alpha A + \beta B)^* = \overline{\alpha} A^* + \overline{\beta} B^*$.*

2. *$(CA)^* = A^*C^*$.*

3. *If $H = K$ (so $A \in \mathcal{B}(H)$), then $(A^*)^* = A$.*

4. *If A is invertible, then A^* is invertible and $(A^*)^{-1} = (A^{-1})^*$.*

Proposition 16.3. *If $A \in \mathcal{B}(H)$, $\|A^*\| = \|A\| = \|A^*A\|^{1/2}$.*

Remark 16.2. The second equality is something you don't get in Banach spaces.

Proof.

$$\begin{aligned}\|A\|^2 &= \sup_{\xi \in (H)_1} \langle A\xi, A\xi \rangle = \sup_{\xi \in (H)_1} \langle A^*A\xi, \xi \rangle \\ &\leq \sup_{\xi \in (H)_1} \langle A^*A\xi, \xi \rangle \leq \sup_{\xi \in (H)_1} \|A^*A\xi\| \|\xi\| \\ &= \|A^*A\| \leq \|A^*\| \|A\|.\end{aligned}$$

So we get that $\|A\| \leq \|A^*\|$. In particular, this holds for $\|A^*\|$, so we get $\|A^*\| \leq \|A\|$. Then all inequalities are equalities, so $\|A^*\| = \|A\| = \|A^*A\|^{1/2}$. □

Example 16.2. If $M_\varphi \in \mathcal{B}(L^2(X, \mu))$ with $\varphi \in L^\infty(X, \mu)$, is multiplication by φ, then $(M_\varphi)^* = M_{\overline{\varphi}}$.

Example 16.3. The right shift $S : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ given by $S(\alpha_1, \alpha_2, \dots) = (0, \alpha_1, \alpha_2, \dots)$ is isometric. Then $S^*(\alpha_1, \alpha_2, \dots) = (\alpha_2, \alpha_3, \dots)$.

16.2 Hermitian operators

Definition 16.3. $A \in \mathcal{B}(H)$ is **Hermitian** (or **self adjoint**) if $A = A^*$.

Proposition 16.4. *A is Hermitian if and only if $\langle A\xi, \xi \rangle \in \mathbb{R}$ for all $\xi \in H$.*

Proof. ($\implies$): We have

$$\langle A\xi, \xi \rangle = \langle \xi, A\xi \rangle = \overline{\langle A\xi, \xi \rangle},$$

so $\langle A\xi, \xi \rangle \in \mathbb{R}$.

($\impliedby$): We would like to prove that if $\langle A\xi, \xi \rangle = \langle \xi, A\xi \rangle$ for all $\xi \in H$, then $\langle A\xi, \eta \rangle = \langle \xi, A\eta \rangle$ for all $\xi, \eta \in H$. We use a **polarization** trick: check that

$$\langle A\xi, \eta \rangle = \frac{1}{4} \sum_{i=0}^{3} i^k \left\langle A(\xi + i^k \eta), \xi + i^k \eta \right\rangle,$$

$$\langle \xi, A, \eta \rangle = \frac{1}{4} \sum_{i=0}^{3} i^k \left\langle \xi + i^k \eta, A(\xi + i^k \eta) \right\rangle.$$

The right hand sides are equal, so the left hand sides are, as well. □

Proposition 16.5. *Let $A \in \mathcal{B}(H)$.*

1. $\|A\| = \sup_{\xi, \eta \in (H)_1} |\langle A\xi, \eta \rangle|$.

2. If $A = A^*$, then $\|A\| = \sup_{\xi \in (H)_1} |\langle A\xi, \xi \rangle|$.

Proof. For (1), we have $\geq$. For $\leq$, take $\eta = \frac{A\xi}{\|A\xi\|}$ for ξ with $A\xi \neq 0$.

For (2), we use

$$\begin{aligned} \langle A(\xi \pm \eta), \xi \pm \eta \rangle &= \langle A\xi, \xi \rangle \pm \langle A\xi, \eta \rangle \pm \langle A\eta, \xi \rangle + \langle A\eta, \eta \rangle \\ &= \langle A\xi, \xi \rangle \pm \langle A\xi, \eta \rangle \pm \overline{\langle A\xi, \eta \rangle} + \langle A\eta, \eta \rangle \\ &= \langle A\xi, \xi \rangle \pm 2 \operatorname{Re} \langle A\xi, \eta \rangle + \langle A\eta, \eta \rangle \end{aligned}$$

By subtracting one from the other, we get

$$\begin{aligned} 4 \operatorname{Re} \langle A\xi, \eta \rangle &= \langle A(\xi + \eta), \xi + \eta \rangle - \langle A(\xi - \eta), \xi - eta \rangle \\ &\leq \left(\sup_{\xi \in (H)_1} |\langle A\xi, \xi \rangle| \right) (\|\xi + \eta\|^2 + \|\eta - \eta\|^2) \\ &= 2 \left(\sup_{\xi \in (H)_1} |\langle A\xi, \xi \rangle| \right) (\|\xi\|^2 + \|\eta\|^2) \\ &\leq 4 \sup_{\xi \in (H)_1} |\langle A\xi, \xi \rangle|. \end{aligned}$$

By part 1, we get $\|A\| \leq \sup_{\xi \in (H)_1} |\langle A\xi, \xi \rangle|$. ☐

Corollary 16.1. *If $\langle A\xi, \xi \rangle = 0$ for all $\xi \in H$, then $A = 0$.*

Proof. For any $A \in \mathcal{B}(H)$, we can decompose A as two self-adjoint operators:

$$A = \frac{A + A^*}{2} + \frac{A - A^*}{2i}.$$

If $\langle A\xi, \xi \rangle = 0$, then this is true for each of these two parts. So each of these parts has norm equal to 0 by the previous proposition. ☐

17 Projections and Idempotents in Hilbert Spaces

17.1 Projections and idempotents

Let H be a Hilbert space over $\mathbb{F}$.

Definition 17.1. An operator $E \in \mathcal{B}(H)$ is **idempotent** if $E^2 = E$. E is a **projection** if $E^2 = E$ and $\ker E = (\operatorname{ran} E)^{\perp}$.

Proposition 17.1. *Let $E \in \mathcal{H}$.*

1. *E is idempotent if and only if $1 - E$ is idempotent.*
2. *$\operatorname{ran} E = \ker(1 - E)$, $\ker E = \operatorname{ran}(1 - E)$, and these are closed subspaces of H.*
3. *$\ker E \cap \operatorname{ran} E = \{0\}$, and $\ker E + \operatorname{ran} E = H$.*

Proof. 1. $(1 - E)^2 = 1 - 2E + E^2$.

2.

$$\begin{aligned} h \in \operatorname{ran} E &\iff hEk \text{ for some } k \\ &\iff Eh = E^2 k = Ek = h \\ &\iff (1 - E)h = 0. \end{aligned}$$

3. $h = Eh + (1 - E)h$. □

Remark 17.1. This also holds for Banach spaces, but we will not use it in that generality.

Proposition 17.2. *Let P be a nonzero idempotent in $\mathcal{B}(H)$. The following are equivalent:*

1. *P is a projection.*
2. *P is the projection onto $\operatorname{ran} P$.*
3. *$\|P\| = 1$.*
4. *$P = P^*$.*
5. *P is normal.*
6. *$\langle Ph, h \rangle \geq 0$ for all h (nonnegativity).*

Proof. (1) $\implies$ (2): Let $M = \operatorname{ran} P$, which is closed. Then the projection $P_M h$ is characterized by $P_M h - h \perp M$; we show that P has this property. For any $h \in H$, $h - Ph = (1-P)h \in \operatorname{ran}(1-P) = \ker P$, and $\operatorname{ran} P \subseteq (\ker P)^\perp$. So $h - Ph \perp M$.

(2) $\implies$ (3): Write $h_1 = Ph$, so $h = h_1 + (h - h_1)$. Then $\|h_1\| \le \|h\|$ if $h \in M$.

(3) $\implies$ (1): We want to show that $\ker P = (\operatorname{ran} P)^\perp$. We will show that $(\ker P)^\perp = \operatorname{ran} P$. Assume $h \perp \ker P$; we will deduce that $h \in \operatorname{ran} P$. We get

$$0 = \langle h, h - Ph\rangle \implies \|h\|^2 = \langle h, Ph\rangle \implies \|h\|^2 \le \|h\|\|Ph\| \le \|P\|\|h\|^2 = \|h\|^2,$$

so all these are equal. Then

$$\|h - Ph\|^2 = \|h\|^2 + \|Ph\|^2 - 2\operatorname{Re}\langle h, Ph\rangle = 0,$$

so $h \in \operatorname{ran} P$.

Suppose $h \in \operatorname{ran} P$. Then $h = h_1 = h_2$, where $h_1 \in (\ker P)^\perp$, and $h_2 \in \operatorname{ran} P$. and is orthogonal to ranP $\cap (\ker P)^\perp$. This means $h_2 \in \operatorname{ran} P \cap \ker P = \{0\}$. so $h = h_1$.

(2) $\implies$ (4): Suppose $P = P_M$. Then $h = h_1 + h_2$ and $k = k_1 + k_2$, where $h_1, k_1 \in M$ and $h_2, k_2 \perp M$. Then

$$\langle Ph, k\rangle = \langle h_1, k_1 + k_2\rangle = \langle h_1, k_1\rangle = \langle h, Pk\rangle .$$

(4) $\implies$ (5): if $P = P^*$, then P commutes with P^*.

(5) $\implies$ (1): If $PP^* = P^*P$, then $\ker PP^* = \ker P^*P$. If $PP^*h = 0$, then

$$\langle PP^*h, h\rangle = \langle P^*h, P^*h\rangle = \|P^*h\|^2,$$

so multiplying by an adjoint does not change the kernel. So $\ker(PP^*) = \ker(P^*) = (\operatorname{ran} P)^\perp$. On the other hand, the same argument gives $\ker P^*P = \ker P$.

(6) $\implies$ (1): Suppose (1) does not hold, so there are an $h = Ph$ and $k \in \ker P$ such that $\langle h, k\rangle \ne 0$. Then

$$\langle P(\alpha h + \beta k), \alpha h + \beta k\rangle = \langle \alpha h, \alpha h + \beta k\rangle = \|\alpha h\|^2 + \alpha\overline{\beta}\langle h, k\rangle ,$$

where $\langle h, k\rangle$ is not necessarily ≥ 0. $\square$

17.2 Invariant and reducible subspaces

If P is a projection, then the map $h \mapsto (Ph, h - P - h)$ is a Hilbert space isomorphism $H \to \operatorname{ran} P \oplus \ker P$. So if we have an operator on H, we can think of it as an operator acting on this direct sum. More generally, if we have a closed subspace M, then $H \cong M \oplus M^\perp$. If $A \in \mathcal{B}(H)$, we identify it with

$$\begin{bmatrix} X & Y \\ W & Z \end{bmatrix}, \qquad Ah = \begin{bmatrix} X & Y \\ W & Z \end{bmatrix}\begin{bmatrix} h_1 \\ h_2 \end{bmatrix},$$

where $X \in \mathcal{B}(M)$, $Z \in \mathcal{B}(M^\perp)$, $Y \in \mathcal{B}(M^\perp, M)$, and $W \in \mathcal{B}(M, M^\perp)$. This gets us partway to diagonalization if we can show that $W, Y = 0$.

Definition 17.2. A subspace $M \leq H$ is **invariant** for $A \in \mathcal{B}(H)$ if $AM \subseteq M$. $M \leq H$ is **reducing** for $A \in \mathcal{B}(H)$ if $AM \subseteq M$ and $AM^{\perp} \subseteq M^{\perp}$.

Here's how we find X, Y, Z, W:

$$A(h_1+h_2) = PA(h_1+h_2)+(1-P)A(h_1+h_2) = PAPh+PA(1-P)h+(1-P)APh+(1-P)A(1-P)h.$$

In other words,

$$\begin{bmatrix} X & Y \\ W & Z \end{bmatrix} = \begin{bmatrix} PA|_M & PA|_{M^{\perp}} \\ (1-P)A|_M & (1-P)A|_{M^{\perp}} \end{bmatrix}.$$

Proposition 17.3. *1. M is invariant for $A \iff PAP = AP \iff W = 0$.*

2. M is reducing for $A \iff PA = AP \iff W = 0, Y = 0$.

Proof. 1. If $PAP = AP$, then $W = (1 - PA|_M = (1 - P)AP|_M = 0$.

If M is invariant, then

$$\begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix} \begin{bmatrix} h_1 \\ 0 \end{bmatrix} = \begin{bmatrix} Xh_1 \\ 0 \end{bmatrix}.$$

2. If $PA = AP$, then $PAP = AP$, so M is invariant. On the other hand $PA = PAP$, so $PA(1 - P) = 0$. So $M^{\perp}$ is invariant, making M reducing. $\square$

Idea on the route to the spectral theorem for self-adjoint operators: Break A into

$$\begin{bmatrix} X & 0 \\ 0 & Y \end{bmatrix}$$

such that X, Y are both "simpler" than A was originally. Keep doing this to "diagonalize" A.

18 Approximation and Eigenvalues of Compact Operators

18.1 Approximation of compact operators by finite rank operators

Last time, we were talking about invariant and reducing subspaces of a Hilbert space M. Here, we have $H = M \oplus M^\perp$ and $A \in \mathcal{B}(H)$ is

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

We saw that M is reducing $\iff P_M A = A P_M \iff Y = 0, W = 0$.

Proposition 18.1. *M is reducing if and only if M is invariant under both A and A^*.*

Proof. This is because

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

Then M is invariant for A^* iff $Y = 0$ iff $M^\perp$ is invariant for A. $\square$

Recall that $\mathcal{B}_0(H)$ is the space of compact operators, and $\mathcal{B}_{00}$ is the space of finite rank operators.

Theorem 18.1. *$\mathcal{B}_{00}(H)$ is dense in $\mathcal{B}_0(H)$.*

Proof. If $T \in \mathcal{B}_0(H)$, then $\overline{T(B_H)}$ is a compact metric space. So it is countable. Then $\operatorname{ran} T \subseteq \operatorname{span} \overline{T(B_H)} \subseteq \overline{\operatorname{span}} D$, where D is any countable dense set in $\overline{T(B_H)}$. So there is an orthonormal $\langle e_n \rangle_n$ such that $\operatorname{ran} T \subseteq \overline{\operatorname{span}}\{e_n\}$. Let P_m be the projection onto $\operatorname{span}\{e_1, \dots, e_m\}$. We will show that $\|P_m T - T\|_{\mathrm{op}} \to 0$.

Observe that for any $h \in \overline{B_H}$, we have $Th = \sum_n \langle Th, e_j \rangle e_h$. Then $P_n Th \to T_h$ in the norm of H. Let $\varepsilon > 0$. We can choose $h_1, \dots, j_k \in \overline{B_H}$ such that for all $h \in \overline{B_H}$, there is some i such that $\|Th - Th_i\| < \varepsilon$. Choose m such that $\|P_m Th_i - Th_i\| < \varepsilon$ for all $i = 1, \dots, k$. Then

$$\|P_m Th - Th\| < \|P_m(Th - Th_i)\| + \|P_m Th_i - Th_i\| + \|Th - i - Th\| < 3\varepsilon.$$

So $\|P_m T - T\|_{\mathrm{op}} < 3\varepsilon$. $\square$

Remark 18.1. If you try to do this with general Banach spaces, it fails. The issue is that you cannot guarantee that $\|P_m\| = 1$ for all m. So you lose control of the bound at the end.

Suppose $\langle e_n \rangle_n$ is an orthonormal basis for H. Define an operator by $Te_n = \alpha_n e_n$ for $\alpha_n \in \mathbb{F}$.

Lemma 18.1. *$T \in \mathcal{B}_0(H)$ if and only if $|\alpha_n| \to 0$.*

Proof. ($\Longrightarrow$): Assume there exist some $\varepsilon > 0$ and $n_1 < n_2 < \cdots$ such that$|\alpha_{n_i} > \varepsilon$. Then $\{Te_{n_1}, Te_{n_2}, \cdots = \{\alpha_{n_1} e_n, \alpha_{n_2} e_{n_2}, \dots\} \subseteq T(\overline{B_H})$. These all are distance $\geq \varepsilon$ to each other and are orthonormal to each other.

($\Longleftarrow$): Let

$$T_m e - n = \begin{cases} \alpha_n e_n & n \leq m \\ 0 & n > m \end{cases} = P_m T.$$

Then we have the diagonal matrix:

$$T - P_m T = \begin{bmatrix} 0 & & & & & \\ & \ddots & & & & \\ & & 0 & & & \\ & & & \alpha_{m+1} & & \\ & & & & \alpha_{m+2} & \\ & & & & & \ddots \end{bmatrix}.$$

So we can see that $\|T - P_m T\|_{\text{op}} \leq \max_{n>m} |\alpha_n|$. $\square$

Example 18.1. Let $k \in L^2(\mu \times \mu)$ and let

$$Kf(x) = \int k(x,y) f(y) \, d\mu(y).$$

For example, if $h \in L^2(-\pi, \pi)$, we have

$$Kf(x) = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} h(x-y) f(y) \, dy.$$

Let the Fourier basis be $e_n(x) = \frac{1}{\sqrt{2\pi}} e^{-inx}$ for all $n \in \mathbb{Z}$. Then we can check

$$Ke_n(x) = \widehat{h}(x) \cdot e_n(x).$$

18.2 Eigenvalues of compact operators

Definition 18.1. If $A \in \mathcal{B}(H)$, an **eigenvalue** of A is a $\lambda \in \mathbb{F}$ such that $\ker(A - \lambda) \neq \{0\}$. The **$\lambda$-eigenspace** is the set of **eigenvectors** corresponding to the eigenvalue λ. We denote the **point spectrum** $\sigma_p(A)$ to be the set of eigenvalues of A.

Remark 18.2. This is a special subset of the **spectrum**, which is the set of $\lambda \in \mathbb{F}$ such that $A - \lambda 1$ is not invertible.

Example 18.2. In $\mathbb{C}^4$, the matrix

$$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

has no nonzero eigenvalues but is nonzero. This kind of phenomenon becomes much richer in infinite dimensions, and we can have compact operators with no nonzero eigenvalues but with interesting properties.

Example 18.3. On $L^2([0,1])$, the **Volterra operator** is

$$Vf(x) = \int_0^x f(y)\, dy = \int_0^1 \mathbb{1}_{\{y \leq x\}} f(y)\, dy.$$

Proposition 18.2. *The Volterra operator is compact but has no eigenvalues.*

Proof. Suppose $Vf = \lambda f$ with $f \neq 0$. If $\lambda = 0$, then the integral of f only any interval is 0, so $f = 0$. Suppose $\lambda \neq 0$. Then we get $f(x) = \lambda^{-1} \int_0^x f$, so f is absolutely continuous, f' exists, and $f'(x) = \lambda^{-1} f(x)$ a.e. Since we must have f is continuous, this gives $f'(x) = \lambda^{-1} f(x)$ everywhere. The solution to this differential equation is $f(x) = Ce^{cx}$. But we must have $C = 0$ because the original equation implies $f(0) = 0$. So $f = 0$. $\square$

Proposition 18.3. *Let $T \in \mathcal{B}_0(H)$ and $\lambda \in \sigma_p(T) \setminus \{0\}$. Then* $\dim \ker(T - \lambda 1) < \infty$.

Proof. Call $M = \ker(T - \lambda 1)$. Then $Tx = \lambda x$ for all $x \in M$. We have $T(\overline{B}_H) \supseteq T(\overline{B}_H \cap M) = \lambda \overline{B}_M$, which is not totally bounded unless $\dim M < \infty$. $\square$

Proposition 18.4. *Let $T \in \mathcal{B}_0(H)$, and let $\lambda \neq 0$. Assume that*

$$\inf\{\|(T - \lambda)h\| : \|h\| = 1\} = 0.$$

Then $\lambda \in \sigma_p(T)$.

Remark 18.3. This says that "approximate eigenvalues" are actually eigenvalues for compact operators.

Proof. Choose $h_1, h_2, \dots$ with $\|h_n\| = 1$ such that $Th_n - \lambda h_h = (T - \lambda)h_n \to 0$ in $\|\cdot\|$. Choose $n_1 < n_2 < \cdots$ susch that $Th_n \to g$. Then $\lambda h_n = Th_n - (Th_n - \lambda h_n) \to g$, so $h_n \to \lambda^{-1} g$. So $Th_n \to \lambda^{-1} Tg = g$. $\square$

Corollary 18.1. *Let $T \in \mathcal{B}_0(H)$, and suppose that $\lambda \notin \sigma_p(T) \cap \{0\}$ and $\overline{\lambda} \notin \sigma_p(T^*)$. Then $T - \lambda$ is invertible.*

Remark 18.4. In fact, we will see that $\overline{\lambda} \notin \sigma_p(T^*)$ is implied by $\lambda \notin \sigma_p(T) \cap \{0\}$.

Proof. We know that $\ker(T-\lambda)=\{0\}$. On the other hand,

$$(\operatorname{ran}(T-\lambda))^{\perp}=\ker(T^*-\overline{\lambda})=\{0\}.$$

To finish, we will show that $\operatorname{ran}(T-\lambda)$ is closed. $(T-\lambda)h=0$ has no nonzero solutions, so there is a $c>0$ such that $\|(T-\lambda)h\|\geq c\|h\|$ for all h. So $(T-\lambda)$ is an open mapping, which forces $\operatorname{ran}(T-\lambda)$ to be closed. $\square$

18.3 The spectral theorem for self-adjoint operators

We will prove the following theorem.

Theorem 18.2 (Spectral theorem for self-adjoint operators)**.** *Suppose T is comapct and self adjoint. Then*

1. *$\sigma_p(T)$ is countable.*
2. *If $\sigma_p(T)\setminus\{0\}=\{\lambda_1,\lambda_2,\dots\}$ and P_n is the projection onto $\ker(T-\lambda_n)$, then*
   - *$P_nP_m=P_mP_n=0$ for all $m\neq n$ (i.e. $\ker(T-\lambda_n)\perp\ker(T-\lambda_m)$).*
   - *$\lambda_n\in\mathbb{R}$ for all n.*
   - *$T=\sum_{n=1}^{\infty}\lambda_nP_n$ in $\|\cdot\|_{\mathrm{op}}$.*

This is an infinite-dimensional diagonalization of T.

19 Spectral Theorems for Compact Operators

19.1 Spectral theorem for compact, self-adjoint operators

Last time, we proved the following propositions about eigenvalues of compact operators.

Proposition 19.1. *Let $T \in \mathcal{B}_0(H)$ and $\lambda \in \sigma_p(T) \setminus \{0\}$. Then* $\dim \ker(T - \lambda 1) < \infty$.

Proposition 19.2. *Let $T \in \mathcal{B}_0(H)$, and let $\lambda \neq 0$. Assume that*

$$\inf\{\|(T-\lambda)h\| : \|h\| = 1\} = 0.$$

Then $\lambda \in \sigma_p(T)$.

Theorem 19.1 (Spectral theorem[12] for self-adjoint operators)**.** *Suppose T is comapct and self adjoint. Then*

1. *$\sigma_p(T)$ is countable.*
2. *If $\sigma_p(T) \setminus \{0\} = \{\lambda_1, \lambda_2, \dots\}$ and P_n is the projection onto* $\ker(T - \lambda_n)$*, then*
   - *$P_n P_m = P_m P_n = 0$ for all $m \neq n$ (i.e.* $\ker(T-\lambda_n) \perp \ker(T - \lambda_m)$*).*
   - *$\lambda_n \in \mathbb{R}$ for all n.*
   - *$T = \sum_{n=1}^{\infty} \lambda_n P_n$ in $\|\cdot\|_{\mathrm{op}}$.*

The last sum should be thought of as an infinite block diagonal matrix where the blocks are $\lambda_i I_{\operatorname{ran} P_i}$.

Lemma 19.1. *If T is normal,* $\ker(T - \lambda) = \ker(T^* - \lambda)$ *is a reducing subspace.*

Proof. If $x \in \ker(T-\lambda)$, then $(T-\lambda)Tx = T(T-\lambda x) = 0$. Then $Tx \in \ker(T-\lambda)$, and same for T^*. ☐

Lemma 19.2. *Let T be self-adjoint. If λ, μ are eigenvalues with $\lambda \neq \overline{\mu}$, then* $\ker(T-\lambda) \perp \ker(T-\mu)$.

Proof. Let $x \in \ker(T-\lambda)$ and $y \in \ker(T-\mu)$. Then

$$\lambda \langle x, \mu \rangle = \langle \lambda x, y \rangle = \langle Tx, y \rangle = \langle x, Ty \rangle = \langle x, \mu y \rangle = \overline{\mu} \langle x, y \rangle$$

So $\langle x, y \rangle = 0$. ☐

Lemma 19.3. *If T is self-adjoint, then $\sigma_p(T) \subseteq \mathbb{R}$.*

[12]Tim learned about the spectral theorem at the same time when he was preparing for his driving test. This was a dangerous idea.

Proof. If $x \in \ker(T-\lambda) \setminus \{0\}$, then

$$\lambda \langle x, x \rangle = \langle \lambda x, x \rangle = \langle Tx, x \rangle = \langle x, Tx \rangle = \overline{\lambda} \langle x, x \rangle ,$$

so $\lambda = \overline{\lambda}$. □

Lemma 19.4. *Let T be compact and self-adjoint. Then at least one of $\|T\|_{\text{op}}, -\|T\|_{\text{op}} \in \sigma_p(T)$.*

Proof. Recall that

$$\|T\| = \sup\{|\langle Tx, x \rangle| : \|x\| = 1\}.$$

Since $\langle Tx, x \rangle = \langle x, Tx \rangle = \overline{\langle Tx, x \rangle}$, we will assume that this equals $\sup\{\langle Tx, x \rangle : \|x\| = 1\}$ (the other case is the negative case). Suppose $\|x_n\| = 1$ and $\langle Tx_n, x_n \rangle \to 1$. Then

$$\|Tx_n - \lambda x_n\|^2 = \underbrace{\langle Tx_n, Tx_n \rangle}_{\leq \lambda^2} - \underbrace{2\lambda \langle Tx_n, x-N \rangle}_{\to -2\lambda^2} + \lambda^2 \|x_n\|^2.$$

By the lemma, $\lambda \in \sigma_p(X)$. □

Proof. If $\|T\| \in \sigma_p(T)$, let $\lambda_1 = \|T\|$, and let P_1 be the projection onto $\ker(T-\lambda_1)$ (this is reducing). Now consider $T_1 = T|_{\ker(T-\lambda_1)^\perp}$. This is compact, self-adjoint, and $\|T_1\| \leq \|T\|$. If $-\|T\| \in \sigma_p(T_1)$, let $\lambda_2 = -\|T\|$ and $P_2 = P_{\ker(T-\lambda_2)}$. Then let $T_2 := T(1-P_1)(1-P_2) = T|_{(\ker(T-\lambda_1)+\ker(T-\lambda_2))^{|perp}}$. Now $\|T_2\| < \|T\|$.

Continue to produce a sequence of eigenvalues $\|lambda_3, \lambda_4, \lambda_5, \dots$ such that $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq |\lambda_4| \geq \cdots$ and a sequence of projections P_i onto $\ker(T-\lambda_i)$. In this sequence of eigenvalues, there are no consecutive equalities. Also, we have $|\lambda_{i+1}| = \|T_{i+1}\|$ and $T_{i+1} := T(1-P_1)\cdots(1-P_i)$.

Next, we show that $|\lambda_i| \to 0$. If not, let $x_i \in \ker(T-\lambda_i)$ be such that $\|x_I\| = 1$. Then $Tx_i = \lambda_i x_i$ is a sequence of orthogonal vectors not going to 0, contradicting compactness.

Now consider $S = \sum_{i=1}^\infty \lambda_i P_i$. We want to show that $S = T$. Call $S_N = \sum_{i=1}^N \lambda_i P_i$. We have by Parseval's theorem that

$$\|(S-S_N)x\|^2 = \left\| \sum_{i=N+1}^\infty \lambda_i P_i x \right\|^2 = \sum_{i=N+1}^\infty |\lambda|^2 \|P_i x\|^2 \leq |\lambda_{N+1}| \sum_{i=N+1}^\infty \|P_i x\|^2 \to 0.$$

Now

$$(T-S_N)x = (T-S_N)x_1 + (T-S_N)x_2$$

where $x_1 = (P_1 + \cdots + P_N)x, x_2 = x - x_1 \perp \operatorname{span}(\ker(T-\lambda_1), \dots, \ker(T-\lambda_N)$. Now split $x_1 = P_1 x_1 + \cdots + P_N x_1 = x_{1,1} + \cdots + x_{1,N}$ to get

$$= \sum_{i=1}^N (T-S_N)x_{1,i} + (T-S_n)X_2$$

$$= Tx_2.$$

And we also have

$$\|Tx_2\| = \|T_{N+1}x_2\| \leq |\lambda_{N+1}|\|x_2\| \leq |\lambda_{N+1}|\|x\| \to 0.$$

Finally, we have enumerated all the eigenvalues, so there are only countably many. □

The proof gives us the following facts, as well.

Corollary 19.1. *Let T be compact and self-adjoint.*

1. *The P_n each have finite rank.*
2. $|\lambda_n| \to 0$.
3. $\ker T = (\sum_n \operatorname{ran} P_n)^\perp$

Here is a formulation which makes this look even more like diagonalization:

Corollary 19.2. *There exist an orthonormal basis $(e_n)_n$ for $(\ker T)^\perp$ and $(\mu_n)_n$ in $\mathbb{R}$ with $\mu_n \to 0$ such that*

$$Tx = \sum_n \mu_n \langle x, e_n \rangle e_n, \qquad \forall x \in H.$$

Proof. Let $T = \sum_m \lambda_m P_m$. Convert to the above form. Each λ_m appears $\dim P_m$-many times as a μ_m. □

19.2 Spectral theorem for compact, normal operators

If N is normal, then $N = S + iT$, where S, T are self-adjoint and $ST = TS$. T and S are linear combinations of N and N^*, so if N is comapct, so are S, T.

Proposition 19.3. *Suppose $S = \sum_{i=1}^\infty \alpha_i P_i$ with $\alpha_i \in \mathbb{F}$ distinct (and nonzero) and P_i orthogonal projections, If $ST = TS$, then $P_i T P_i = TP_i$ for all i. If S is self-adjoint, then $P_i T = TP_i$ for all i.*

Proof. Check that $\ker(S - \alpha_i) = \operatorname{ran} P_i$. If $Sx = \alpha_i x$, then $STx = TSx = T(\alpha_i x) = \alpha_i Tx$. This shows that $P_i T P_i = TP_i$.

If $S = S^*$, then P_i reduces T for all i:

$$ST^* = S^*T^* = (TS)^* = (ST)^* = T^*S^* = T^*S.$$

So $P_i T = TP_i$. □

Theorem 19.2 (Spectral theorem for compact, normal operators)**.** *Let N be comapct and normal. Then*

1. *$\sigma_p(T)$ is countable.*

2. *If $\sigma_p(T) \setminus \{0\} = \{\lambda_1, \lambda_2, \dots\}$ and P_n is the projection onto* $\ker(T - \lambda_n)$*, then*

   - $P_n P_m = P_m P_n = 0$ *for all $m \neq n$ (i.e.* $\ker(N - \lambda_n) \perp \ker(N - \lambda_m)$*).*
   - $N = \sum_{n=1}^{\infty} \lambda_n P_n$ *in* $\|\cdot\|_{\mathrm{op}}$.

Proof. Let $N = S + iT$ with S, T self-adjoint, and write $S = \sum_{k\geq 1} \lambda_k^S P_k^S$. Now P_k^S reduces T for all k. Now choose a further decomposition $P_k^S = Q_{k,1} + \cdots + Q_{k,m_k}$ such that $TP_k^S = TQ_{k,1} + \cdots + TQ_{k,m_k} = \beta_{k,1}^T Q_{k,1} + \cdots + \beta_{k,m_k}^T Q_{k,m_k}$. Now $S = \sum_k \sum_{i=1}^{m_k} \lambda_k^S Q_{k,i}$, and $T = \sum_k \sum_{i=1}^{m_k} \beta_{k,1} Q_{k,i}$. So

$$S + iT = \sum_k \sum_{i=1}^{m_k} (\lambda_k^S + i\beta_{k,i}) Q_{k,i}.$$

Check that $\beta_{k,i} \to 0$ and that $Q_{k,i} Q_{\ell,j} = Q_{\ell,j} Q_{k,i} = 0$. $\square$

For non-compact operators, we will have an analogous result that gives $T = \int_a^b \lambda \, dE(\lambda)$. We have to make sense of this integral.

20 Positive Operators and Spectral Families

20.1 Positive operators

We want to generalize the following theorem, without the assumption of compactness.

Theorem 20.1 (Spectral theorem in finite dimensions)**.** *Let* $\dim(H) < \infty$*, and let* $T : H \to H$ *be a self-adjoint operator with eigenvalues* $a \leq \lambda_1 < \lambda_2 < \cdots < \lambda_m = b$*. Then*

$$T = \sum_{i=1}^{n} \lambda_i P_{\lambda_i},$$

where P_{λ_i} *is the projection onto* $\ker(T - \lambda_i)$*.*

Example 20.1. On $L^2([0,1])$ we have $Tf(x) = xf(x)$, the multiplication operator. Then $\|T\|_{\text{op}} \leq 1$, and

$$\langle Tf, g \rangle = \int_0^1 \overline{x} f(x) \overline{g(x)} \, dx = \langle f, Tg \rangle .$$

However, T has no eigenvectors! If $Tf = \lambda f$, then $xf(x) = \lambda f(x)$ for a.e. x. So $f = 0$ a.e.

Observe that if $V = \ker(T - \lambda) \neq \{0\}$, then V is reducing and $T|_V = \lambda I_V$. We want to loosen this to $\mu I_V \leq T|_V \leq \lambda I_V$ for $\mu < \lambda$.

Definition 20.1. $T \in \mathcal{B}(H)$ is **positive** (written $T \geq 0$) if T is self-adjoint and $\langle Tx, x \rangle \geq 0$. If S, T are self-adjoint, we say $S \leq T$ if $T - S \geq 0$.

This defines a partial order on the set of self-adjoint operators. How does this relate to our previous examples?

Example 20.2. In the finite dimensional case, for $\lambda \in \mathbb{R}$, define

$$E(\lambda) := \sum_{i : \lambda_i \leq \lambda} P_{\lambda_i}, \qquad E(\mu, \lambda) := \sum_{\mu < \lambda_i \leq \lambda} P_{\lambda_i} = E(\lambda) - E(\mu).$$

These all reduce T, and

$$\mu E(\mu, \lambda) \leq TE(\mu, \lambda) \leq \lambda E(\mu, \lambda)$$

for all $\mu \leq \lambda$. If λ_i is the unique element of $\sigma_p(T) \cap (\mu, \lambda]$, $\lambda_i P_i \leq TP_{\lambda_i} \leq \lambda_i P_i$.

Example 20.3. With the multiplication operator T on L^2, let $V(\mu, \lambda) := \{f \in L^2([0,1]) : f = f \mathbb{1}_{(\mu, \lambda]}\}$ for any $\mu \leq \lambda$. Then let $E(\mu, \lambda) = P_{V(\mu, \lambda)}$. We can check that

$$TE(\mu, \lambda) f(x) = xf \mathbb{1}_{(\mu, \lambda]}(x).$$

Then $\mu E(\mu, \lambda) \leq TE(\mu, \lambda) \leq \lambda E(\mu, \lambda)$.

Lemma 20.1. *Let T be self-adjoint, and let $a = \inf_{\|x\|=1} \langle Tx, x \rangle$. and $b = \sup_{\|x\|=1} \langle Tx, x \rangle$. Then $a \leq T \leq b$ and $\|T\| = \max(|a|, |b|)$.*

Proof. If $\|x\| = 1$, then

$$\langle (T - a)x, x \rangle = \langle Tx, x \rangle = a \geq 0.$$

The upper bound is the same.

We have seen already that $\|T\| = \sup | \langle Tx, x \rangle |$. $\square$

Corollary 20.1. *If $S \leq T$ and $T \leq S$ then $S = T$.*

Proof. This implies that $\langle (S - T)x, x \rangle = 0$ for all x. So the norm is $\|S - T\| = 0$. $\square$

Lemma 20.2. *For projections P, Q, the following are equivalent:*

1. *$P \leq Q$.*
2. *$QP = PQ = P$.*
3. *$Q - P$ is a projection.*
4. *$\|Px\| \leq \|Qx\|$.*
5. *$\operatorname{ran} P \subseteq \operatorname{ran} Q$.*

Proof. (1) $\implies$ (5): If (5) is false, then there is some $x \neq 0$ such that $Px = x$ but $Qx \neq x$. Then $\|x\|^2 = \langle Px, x \rangle$, but $\langle Qx, x \rangle = \|Qx\|^2 < \|x\|^2$. This contradicts (1).

(5) $\implies$ (2): $QP = P$ by the condition of (5), and we get $(QP)^* = P^*Q^* = PQ$ by self-adjointness.

(2) $\implies$ (4): $\|Px\| = \|PQx\| \leq \|Qx\|$.

(2) $\implies$ (3): $\langle (Q - P)x, x \rangle = \langle Q(1 - P)x, x \rangle = \langle Q(1 - P)x, Qx \rangle \geq 0$.

(3) $\implies$ (1): $Q - P$ is a projection, so $Q - P \geq 0$. $\square$

20.2 Spectral families and the spectral theorem

Definition 20.2. A **spectral family** on H is a map $\lambda \mapsto E(\lambda)$ from $\mathbb{R} \to \{\text{proj. on } H\}$ such that

1. If $\lambda > \mu$, then $E(\lambda) \geq E(\mu)$
2. There exist $a, b \in \mathbb{R}$ such that $E(\lambda) = 0$ if $|lambda < a$ and $E(\lambda) = I$ if $\lambda \geq b$.
3. $E(\lambda)x \to E(\mu)x$ as $\lambda \downarrow \mu$ for all $x \in H$ (convergence in the strong operator topology).

Theorem 20.2. *Let T be a self-adjoint operator on H. Then there exists a spectral family $(E(\lambda))_{\lambda\in\mathbb{R}}$ such that*

$$a = \inf_{\|x\|=1} \langle Tx, x\rangle, \qquad b = \sup_{\|x\|=1} \langle Tx, x\rangle$$

$$T = \int_{\mathbb{R}} \lambda \, dE(\lambda).$$

This means $\langle Tx, y\rangle = \int_{[a,b]} \lambda \, d\mu_{x,y}$ for all $x, y \in H$, where $\mu_{x,y}$ is the Lebesgue-Stieltjes measure corresponding to $F_{x,y}$.

To interpret this integral, we need the following lemma.

Lemma 20.3. *If E is a spectral family, then for any $x, y \in H$, then function $F_{x,y} : \lambda \mapsto \langle E(\lambda)x, y\rangle$ is right-continuous and of bounded variation.*

Proof. Right continuity follows from property (3) of a spectral family. For bounded variation,

Step 1: If $y = x$, then $F_{x,x}(\lambda) = \|E(\lambda)x\|^2$, which is increasing with λ.

Step 2:

$$F_{x,y}(\lambda) = \langle E(\lambda)x, y\rangle = \frac{\langle E(\lambda)(x+y), x+y\rangle - \langle E(\lambda)x, x\rangle - \langle E(\lambda)y, y\rangle}{2}$$

is a difference of nondecreasing functions, so it is of bounded variation. $\square$

Example 20.4. In the finite dimensional case, $E(\lambda)$ is constant, except for finitely many jumps. So the integral becomes a finite sum.

Example 20.5. Returning to the multiplication operator on L^2, if $f, g \in L^2([0,1])$, then

$$\langle Tf, g\rangle = \int_0^1 x f(x)\overline{g(x)} \, dx, \qquad dx = d\mu_{f,g}.$$

Here, $E(\lambda)$ is the proejction onto $\{f = f\mathbb{1}_{[0,\lambda]}\}$,and $\langle E(\lambda)f, g\rangle = \int_0^\lambda f\overline{g} \, dx$.

20.3 Functional calculus

How do we find this map $\lambda \mapsto E(\lambda)$? In the finite dimensional case, we have a self-adjoint T with eigenvalues $a = \lambda_1 < \lambda_2 < \cdots < \lambda_m = b$ and $T = \sum_i \lambda_i P_{\lambda_i}$. If $p(t) = \sum_{j=1}^k x_j t^j \in \mathbb{R}[t]$ is a polynomial, we can write $p(T) = \sum_{j=1}^k c_j T^j$. Since $T^j = \sum_i \lambda_i^j P_i$, we have $p(T) = \sum_i p(\lambda_i) P_{\lambda_i}$.

Choose any $p_\lambda \in \mathbb{R}[t]$ such that

$$p_\lambda(t) = \begin{cases} 1 & t = \lambda_i \le \lambda \\ 0 & t = \lambda_i > \lambda. \end{cases}$$

Then

$$p_\lambda(T) = \sum_{\lambda_i \le \lambda} P_{\lambda_i} = E(\lambda).$$

We need to make this work in infinite dimensions. But $\mathbb{R}[t]$ is not rich enough. We must extend the map $\mathbb{R}[T] \to \mathcal{B}(H)$ taking $p \mapsto p(T)$ to a larger class of functions. After doing so, we get the **functional calculus** of T. In particular, we want to be able to get the function $p(T)$, where $p(t) = \mathbb{1}_{(-\infty,\lambda]}(t)$.

21 Continuous Functional Calculus for Self-Adjoint Operators

21.1 Idea for proving the general spectral theorem for self-adjoint operators

Theorem 21.1. *Let T be a self-adjoint operator on H with*

$$a = \inf_{\|x\|=1} \langle Tx, x \rangle, \qquad b = \sup_{\|x\|=1} \langle Tx, x \rangle .$$

Then there exists a spectral family $(E(\lambda))_{\lambda \in \mathbb{R}}$ such that

$$T = \int_{\mathbb{R}} \lambda \, dE(\lambda).$$

This means $\langle Tx, y \rangle = \int_{[a,b]} \lambda \, d\mu_{x,y}$ for all $x, y \in H$, where $\mu_{x,y}$ is the Lebesgue-Stieltjes measure corresponding to $F_{x,y}(\lambda) = \langle E(\lambda), x, y \rangle$.

Method: Consider the map $\mathbb{R}[x] \to \mathcal{B}(H)$ sending $p(t) = \sum_{i=1}^n c_i t^i \mapsto \sum_{i=1}^n c_i T^i$.

Remark 21.1. This is a homomorphism of algebras over $\mathbb{R}$.

The idea is to enrich the domain of this homomorphism to produce many more operators out of T. Why is this relevant? Suppose

$$T = \sum_{i=1}^N \lambda_i E(\lambda_{i-1}, \lambda_i)$$

like in the finite case. Then

$$T^2 = \sum_{i=1}^N \lambda_i^2 E(\lambda_{i-1}, \lambda_i).$$

This generalizes to any polynomial of T. Assume we can do this for the functions $p(t) = \mathbb{1}_{(-\infty,\mu]}(t)$. Then

$$p(T) = \int_a^\mu p(\lambda) \, dE(\lambda) = E(\mu).$$

So this should tell us what $E(\mu)$ is. The proof of the spectral theorem is basically this idea in reverse.

21.2 Continuous functional calculus

When we extend our functional calculus to non-polynomial functions, we only really care what the functions only do on the spectrum of T. In particular, the function $p(T)$ should only depend on the values of p in $[a,b]$, where a,b are as above.

Our next goal is to show that if $p \in R[t]$ and $c \le p(t) \le d$ for all $t \in [a,b]$, then $cI \le p(T) \le dI$. It is enough to show one side of the inequality, and so it is enough to show it when $c = 0$. So we will show that if $p|_{[a,b]} \ge 0$, then $p(T) \ge 0$.

Lemma 21.1 (sum of squares decomposition)**.** *If $p \in \mathbb{R}[t]$ and $p \ge 0$, then there exist $q_1, \dots, q_m \in \mathbb{R}[t]$ such that $p(t) = \sum_{i=1}^m q_i(t)^2$.*

Remark 21.2. This can actually be shown for $m = 2$.

Proof. Let $p(t) = \sum_{i=0}^n c_i t^i$. If $n = 0$, we are done. Suppose $n \ge 1$ and $p \ge 0$. Then n is even, and $c_n > 0$. Then there exists some $u \in \mathbb{R}$ such that $p(u) = \min p(\mathbb{R}) =: c$. Let $p_1(t) := p(t) - c$. Then $p_1 \ge 0$, and $p_1(0) = 0$. This implies that $(t-u)^2$ divides p in $\mathbb{R}[t]$. so $p_1(t) = (t-u)^2 p_2(t)$. Since p_2 is continuous, $p_2 \ge 0$ with $\deg p_2 = n-2$. By induction, $p_2(t) = \sum_j q_j(t)^2$, and we get

$$p(t) = \sum_j ((t-u)q_j(t))^2 + (\sqrt{c})^2. \qquad \square$$

Proposition 21.1. *Let T be self-adjoint. If $p|_{[a,b]} \ge 0$, then $p(T) \ge 0$.*

Proof. Step 1: Assume $p \ge 0$. Then $p(t) = \sum_j q_j(t)^2$, so $p(T) = \sum_j (q_j(T))^2$, which is a sum of positive operators: $\langle (q(T))^2 x, x \rangle = \|q(T)x\|^2$.

Step 2: Assume $a = -1$, $b = 1$, so $p|_{[-1,1]} \ge 0$. Let $\varepsilon > 0$, and choose $\delta > 0$ such that $(p+\varepsilon)|_{[-(1+\delta),1+\delta]} \ge 0$. Define $p_n(t) = p(t) = \varepsilon + (\frac{t}{1+\delta})^{2n}$. Then $p_n \ge 0$ for all sufficiently large n. So by case 1, $p_n(T) \ge 0$ for all sufficiently large n. But

$$\|(p+\varepsilon)(T) - p_n(T)\| = \left\| \left(\frac{T}{1+\delta} \right)^n \right\| \le \frac{\|T\|^n}{(1+\delta)^n} = \frac{1}{(1+\delta)^n} \to 0.$$

So $p_n(T) \xrightarrow{op} (p+\varepsilon)(T)$, which makes $(p+\varepsilon)(T) \ge 0$. Then $p(T) \ge 0$. $\square$

So $p \mapsto p(T)$ satisfies $\|p(T)\|_{\text{op}} \le \|p|_{[a,b]}\|_{\text{sup}}$. So if $f \in C([a,b])$, define $f(T) = \lim_n p_n(T)$. Then for all $p_n \in \mathbb{R}[t]$ such that $p_n \to f$ uniformly on $[a,b]$ this is well-defined. So $f \mapsto f(T)$ is still a homomorphism: If $p_n \to f$ and $q_n \to g$ in $C([a,b])$, then $p_n q_n \to fg$ in $C([a,b])$. Therefore,

$$(fg)(t) = \lim_n (p_n q_n)(T) = \lim_n p_n(T) q_n(T) = f(T)g(T).$$

This gives us a well-defined functional calculus for continuous functions.

21.3 Weak operator limits of positive functions

Definition 21.1. If $\langle T_n \rangle \in \mathcal{B}(H)$ and $T \in \mathcal{B}(H)$, $T_n \to T$ in the **weak operator topology** if $\langle T_n x, y \rangle \to \langle Tx, y \rangle$.

Remark 21.3. We could also define the **strong operator topology** by $T_n x \to Tx$ for all $x \in H$.

Proposition 21.2. *Suppose $\langle T_n \rangle_n \in \mathcal{B}(H)$ with $T_n \geq 0$ and $T_n \geq T_{n+1}$ for all n. Then there exists a positive $T \in \mathcal{B}(H)$ such that $T_n \xrightarrow{WOT} T$.*

Remark 21.4. We actually get SOT convergence here, but the proof is a bit harder.

Proof. For all $x \in H$, $\langle T_n x, x \rangle$ must converge to some $Q(x,x) \geq 0$. FOr all x, y we get

$$\begin{aligned} \langle T_n x, y \rangle &= \frac{1}{2}(\langle T_n(x+y), x+y \rangle - \langle T_n x, x \rangle - \langle T_n y, y \rangle) \\ &\to \frac{1}{2}(Q(x+y, x+y) - Q(x,x) - Q(y,y)) \\ &=: Q(x,y). \end{aligned}$$

Check that $(x,y) \mapsto Q(x,y)$ is symmetric, bilinear, positive, and $Q(x,x) \leq M\|x\|^2$ for some M.

So for each x, the map $Q(x,\cdot) \in H^* = H$ and is bounded: $\|Q(x,\cdot)\| \leq M\|x\|$. By Riesz representation, there exists some $Tx \in H$ such that $Q(x,y) = \langle Tx, y \rangle$ for all x, y. Check that $T \in \mathcal{B}(H)$ is self-adjoint with $T \geq 0$. $\square$

The idea for the next step is to let $\langle f_n \rangle_n \in C([a,b])$ be bounded below with $f_n \downarrow g$ (possible e.g. if $g = \mathbb{1}_{(-\infty,\mu]}$). Then if $f_n \geq f$ in $C([a,b]$, $f_n(T) \geq f_{n+1}(T)$ in $\mathcal{B}(H)$. We will define $g(T)$ as the WOT limit of the $f_n(T)$.

22 Extension of Functional Calculus and Proof of The Spectral Theorem

22.1 Proof of the spectral theorem for self-adjoint operators

So far, we've constructed continuous functional calculus: a map $C([a,b]) \to \mathcal{B}_{\mathrm{sa}}(H)$ sending $f \mapsto f(T)$ which is

- linear,
- $f(g)(T) = f(T)g(T)$,
- $f \geq g \implies f(T) \geq g(T)$,
- $1(T) = I$,
- $\|f(T)\| \leq \|f\|_{\sup}$.

If $(f_n)_n$ is a sequence in $C([a,b])$ with $f_n \geq 0$ and $f_n(x) \downarrow g(x)$ for all $x \in [a,b]$, then we want to define $g(T)$ by $\langle g(T)x, y\rangle = \lim_n \langle f_n(T)x, y\rangle$. Last time, we showed that this limit exists (as a weak operator topology limit).

Lemma 22.1. *Suppose $f_n, f_n' \downarrow g$. Then the limit, $g(T)$, is the same.*

Proof. Let $f_n, f_n' \downarrow g$. For every x, $\varepsilon > 0$, and $n \in \mathbb{N}$, there exists an $n'(x,\varepsilon)$ such that $f'_{n'}(x) < g(x) + \varepsilon \leq f_n(x) + \varepsilon$. Then for each $n \in \mathbb{N}$, $\varepsilon > 0$ and x, we get $n'(n,x,\varepsilon)$ and a neighborhood $U(n,x,\varepsilon)$ of x such that $f'_{u'}|_{U_{n,x,\varepsilon}} < (f_n + \varepsilon)|_{U(n,x,\varepsilon)}$. Choose $x_1, \dots, x_t$ such that $\bigcup_{i=1}^t U(n, x_i, \varepsilon) = [a,b]$. Let $n'' = \max(n'(n,x_1,\varepsilon), \dots, n'(n,x_t,\varepsilon))$. Now $f'_{n''} < f_n + \varepsilon$ on $[a,b]$. Then $f'_{n''}(T) \leq f_n(T) + \varepsilon I$, so $\lim_{n''} f'_{n''}(T) \leq f_n(T) + \varepsilon$ for all n, ε. Since ε is arbitrary, and by symmetry, we get that $\lim f_n'(T) \leq \lim_n f_n(T)$ and $\lim f_n'(T) \geq \lim_n f_n(T)$. So the limits are equal. $\square$

Now, if we have $f_n \downarrow g \geq 0$, we get $g(T) \geq 0$. This is

- still additive: If $f_n \downarrow g$, $f_n' \downarrow g$;, then $f_n + f_n' \downarrow g + g'$. We have

$$(g+g')(T) = \mathrm{WO}\lim_n (f_n(T) + f_n'(T)) = g(T) + g'(T).$$

Lemma 22.2. *If $f_n \downarrow g \geq 0$, and $f_n' \downarrow g_n' \geq 0$, then*

$$(gg')(T) = g(T)g'(T).$$

Proof. We have $f_n f_n' \downarrow gg'$, so $(gg')(T) = \mathrm{WO}\lim_n (f_n f_n')(T)$. We want to show that the is the product of the limits of $f_n(T)$ and $f_n'(T)$. By polarization, it is enough to show that $\lim_n \langle (f_n f_n')Tx, x\rangle = \lim_n \lim_m \langle f_n(T) f_m'(T)x, x\rangle$. The limit of the diagonal terms is the same as $\lim_n \lim_m$ because the array is decreasing in n, m (a basic real analysis fact). $\square$

Given $\lambda\ in [a,b]$ and $n \in \mathbb{N}$, define

$$\varphi_n^\lambda(t) = \begin{cases} 1 & t \leq \lambda \\ -n(x - (\lambda + 1/n)) & \lambda < t \leq \lambda + 1/n \\ 0 & t > \lambda + 1/n \end{cases}$$

Then $\varphi_n^\lambda \downarrow \mathbb{1}_{(-\infty,\lambda]}$ as $n \to \infty$. Define $E(\lambda) := \lim_n \varphi_n^\lambda(T)$.

Here are the properties of $E(\lambda)$:

1. $E(\lambda)$ is self adjoint (as a WO limit of self-adjoints).
2. $E(\lambda) = \mathbb{1}_{(-\infty,\lambda]}(T) = \mathbb{1}^2_{(-\infty,\lambda]}(T) = E(\lambda)^2$.
3. If $\lambda \geq \mu$, then
$$E(\mu)E(\lambda) = E(\lambda)E(\mu) = (\mathbb{1}_{(-\infty,\lambda]}\mathbb{1}_{(-\infty,\mu]})(T) = E(\mu).$$
4. Declare $E(\lambda) = 0$ if $\lambda < a$ and $E(b) = \lim_n 1(T) = I$.
5. Fix $\lambda \in [a,b]$. Then $E(\mu)x \to E(\lambda)x$ as $\mu \downarrow \lambda$ for all $x \in H$. Equivalently, $\langle (E(\mu) - E(\lambda))x, x\rangle \to 0$.

   To show this, we know $\langle E(\lambda)x, x\rangle = \lim_n \left\langle \varphi_n^\lambda(T)x, x\right\rangle$. Pick n large enough so that $\left\langle \varphi_n^\lambda(T)x, x\right\rangle < \langle E(\lambda)x, x\rangle + \varepsilon$. This is also $\lim_{\mu\downarrow\lambda} \langle \varphi_n^\mu(T)x, x\rangle$. So for μ close enough to λ, we get
$$\langle E(\mu)x, x\rangle \leq \langle \varphi_n^\mu(T)x, x\rangle < \langle \varphi_n^\lambda(T)x, x\rangle + \varepsilon < \langle E(\lambda)x, x\rangle + 2\varepsilon.$$

This gives us a spectral family for T. If $a \leq \mu \leq \lambda \leq b$, then

$$E(\mu, \lambda] := E(\lambda) - E(\mu) = \mathrm{WO}\lim_n [\varphi_n^\lambda(T) - \varphi_n^\mu(T)].$$

This gives us

$$TE(\mu,\lambda] = \mathrm{WO}\lim_n T[\varphi_n^\lambda(T) - \varphi_n^\mu(T)] = \mathrm{WO}\lim_n [(t \cdot (\varphi_n^\lambda(t) - \varphi_n^\mu(t)))(T)].$$

Now

$$\mu\mathbb{1}_{(\mu+1/n,\lambda]} \leq t(\varphi_n^\lambda(t) - \varphi_n^\mu(\lambda)) \leq \lambda\mathbb{1}_{(\mu,\lambda+1/n]}$$

Taking the weak operator limit, we get

$$\mu E(\mu,\lambda) \leq TE(\mu,\lambda) \leq \lambda E(\mu,\lambda).$$

Now let $a = \lambda_0 < \lambda_1 < \cdots < \lambda_m = b$. Then

$$I = E(B)$$

$$= (E(\lambda_n) - E(\lambda_{n-1})) + \cdots + (E(\lambda_1) - E(\lambda_0))$$
$$= E(a, \lambda_1] + E(\lambda_1, \lambda_2] + \cdots + E(\lambda_{n-1}, b].$$

Multiplying by T, we get

$$T = TE(a, \lambda_1] + TE(\lambda_1, \lambda_2] + \cdots + TE(\lambda_{n-1}, b].$$

So we get

$$\sum_{i=1}^{m} \lambda_{i-1} E(\lambda_{i-1}, \lambda_i] \leq T \leq \sum_{i=1}^{n} \lambda_i E(\lambda_{i-1}, \lambda_i].$$

This gives

$$\sum_{i=1}^{m} \lambda_{i-1} \langle E(\lambda_{i-1}, \lambda_i]x, x\rangle \leq \langle Tx, x\rangle \leq \sum_{i=1}^{n} \lambda_i \langle E(\lambda_{i-1}, \lambda_i]x, x\rangle .$$

These are partial sums in the definition of the Riemann-Stieltjes integral. So taking the limit as $\max_i |\lambda_i - \lambda_{i-1}| \to 0$, we get

$$\langle Tx, x\rangle = \int \lambda \, d \langle E(\lambda)x, x\rangle .$$

This completes the proof of the spectral theorem.

22.2 Borel functional calculus and spectral measure

How far can we take this functional calculus? Here is another method which allows us to extend to all Borel functions. Assume we have a continuous functional calculus: $f \mapsto f(T)$ for all $f \in C([a, b])$. Given $x, y \in H$, consider

$$f \mapsto \langle f(T)x, y\rangle .$$

This is bounded by $|\langle f(T), x, y\rangle| \leq \|f\|_{\sup}\|x\|\|y\|$. So there exists some $\mu_{x,y} \in M([a, b])$ such that $\|\mu_{x,y}\| \leq \|x\|\|y\|$ and $\langle f(T)x, y\rangle = \int f \, d\mu_{x,y}$. So given g bounded and Borel, define

$$Q_g(x, y) := \int g \, d\mu_{x,y}.$$

This is bilinear in x, y and bounded: $|Q_g(x, y)| \leq \|g\|_\infty \|x\|\|y\|$. At each step, our construction is symmetric in x, y, so $Q_g(x, y)$ is symmetric in x, y. Now define $g(T)$ by $\langle g(T)x, y\rangle = Q_g(x, y)$. We can now define, as before, $\mathbb{1}_{(-\infty,\lambda]}(T)$.

The advantage of this method is that we can also define $E(A) := \mathbb{1}_A(T)$ for all $A \in \mathcal{B}([a, b])$. We can now show that

- Every $E(A)$ is a projection.

- $E(A \cap B) = E(A)E(B)$.
- $E(\varnothing) = 0$, and $E([a, b]) = I$.
- $E(\bigcup_n A_n) = \sum_n E(A_n)$.

This gives a **spectral measure**, which has the properties of a measure but takes values in projections. More advanced versions of the spectral theorem use this approach.

23 Spectral Theorem for Normal Operators

23.1 Spectral theorem for normal operators

Let T be a self-adjoint, bounded operator on a Hilbert space H. We have shown that

$$T = \int_{[a,b]} \lambda \, dE(\lambda),$$

in the sense that

$$\langle Tx, y \rangle = \int_{[a,b]} \lambda \, d\langle E(\lambda), x, y \rangle,$$

where $d\,\langle E(\lambda), x, y\rangle$ is the Lebesgue-Stieltjes measure given by the map $\lambda \mapsto \langle E(\lambda), x, y \rangle$.

We can extend our functional calculus to Borel-measurable functions by defining $f(T)$ to satisfy

$$\langle f(T)x, y \rangle = \int_{[a,b]} f(\lambda) \, d\langle E(\lambda), x, y \rangle.$$

So we can construct a **spectral measure** $E : \mathcal{B}([a, b]) \to \mathrm{Proj}(H)$ such that

- $E(\varnothing) = 0$, $E([a, b]) = I$,
- If A_n are disjoint, $E(\bigcup_{n=1}^\infty A_n)x = \sum_{n=1}^\infty E(A_n)x$ for all x (this is a weak operator convergent sum).
- $E(A \cap B) = E(A)E(B)$ for all $A, B \in \mathcal{B}([a, b])$.

Diagonalization of an operator looks like

$$T = \sum_{\lambda \in \sigma_p(T)} \lambda P_\lambda.$$

In the self-adjoint case, all λs must be real. In the case of normal operators, λ may be complex.

Theorem 23.1 (Spectral theorem for bounded normal operators)**.** *Let H be a Hilbert space over $\mathbb{C}$, and let $N \in \mathcal{B}(H)$ be normal. Then there is a compact $D \subseteq \mathbb{C}$ and a spectral measure $E : \mathcal{B}(D) \to \mathrm{Proj}(H)$ such that*

$$N = \int_D z \, dE(z).$$

In other words,

$$\langle Nx, y \rangle = \int z \, d\langle E(z)x, y \rangle,$$

where $U \mapsto \langle E(U)x, y \rangle$ is a complex-valued measure for each $x, y \in H$.

Given N, we can write $S + iT$, where S, T are both self-adjoint and commute.

Lemma 23.1. *In the spectral representation of T, every $E(\lambda)$ commutes with every operator that commutes with T.*

Proof. If $p \in \mathbb{R}[t])$ and S commutes with T, then S commutes with $p(T)$. Commutativity survives for $f(T)$ with $f \in C([a, b])$ by convergence in operator norm. Finally, if $T_n \xrightarrow{WO} T$ and $ST_n = T_n S$ for all n, then

$$\langle STx, y \rangle = \langle Tx, S^* y \rangle = \lim_n \langle T_n x, S^* y \rangle = \langle TSx, y \rangle,$$

for all $x, y \in H$, so $ST = TS$. □

Corollary 23.1. $E^S(\mu)E^T(\lambda) = E^T(\lambda)E^S(\mu)$ *for all* λ, μ.

Proof. Apply the lemma twice. □

Now, given $(p, q] + i(r, s] \subseteq \mathbb{C}$, define

$$\begin{aligned} E((p, q] + i(r, s]) &:= E^S((p, q]) + iE^T((r, s]) \\ &= (E^S(q) - E^S(p)) + i(E^T(s) - E^T(r)). \end{aligned}$$

Warning: We may have $\left\langle E^S((p, q])E^T((r, s])x, u \right\rangle \neq \left\langle E^S((p, q])x, y \right\rangle \left\langle E^T((r, s]x, y \right\rangle$.

Let $a^S = \inf \langle Sx, x \rangle$ and $b^S = \sup \langle Sx, x \rangle$, and define a^T and b^T similarly. We can now define

$$N' = \int_D z \, dE(z), \qquad D = [a^S, b^S] + i[a^T, b^T].$$

We want to check that $N' = N$. Using the spectral theorem for self-adjoint operators,

$$N' = \int_D x \, dE(z) + i \int_D y \, dE(z) = S + iT = N.$$

The middle step is by a "Fubini"-type argument.

23.2 Approximate eigenvalues

In the compact case, we had $T = \sum_{i=1}^{\infty} \lambda_i P_{\lambda_i}$, where the λ_i were the eigenvalues of T.

Definition 23.1. Let X be a normed space. $\lambda \in \mathbb{C}$ is an **approximate eigenvalue** for $T \in \mathcal{B}(X)$ if

$$\inf\{\|(T - \lambda)x\| : \|x\| = 1\} = 0.$$

For compact operators, we saw that these were actually eigenvalues of the operator. In general, this isn't true. Here is an example:

Proposition 23.1. *Let $H = L^2([0,1])$, and let $Tf(x) = xf(x)$. Then λ is an approximate eigenvalue if and only if $\lambda \in [0,1]$.*

Proof. Let $0 \le \lambda \le 1$, so $(T-\lambda)f(x) = (x-\lambda)f(x)$. For any $\varepsilon > 0$, pick $f \in L^2([0,1])$ with $\|f\| = 1$ such that $f(x) = 0$ if $x \notin (\lambda - \varepsilon, \lambda + \varepsilon)$. Then $\|Tf\| \le \varepsilon \|f\|$. $\square$

How does this play into our spectral representation, $T = \int_{[a,b]} \lambda \, dE(\lambda)$?

Definition 23.2. The **support** of E is $\{\lambda : E(\lambda + \varepsilon) \neq E(\lambda - \varepsilon) \, \forall \varepsilon > 0\}$.

Proposition 23.2. *The support of E is the set of approximate eigenvalues for T.*

Proof. ($\supseteq$): Suppose that $E(c) = E(d)$ for some $a \le c < d \le b$, and let $c < \mu < d$; we will show that μ is not an approximate eigenvalue. Then

$$T = \int_{[a,b]} \lambda \, dE(\lambda) = \int_{[a,c]} \lambda \, dE(\lambda) + \int_{[d,b]} \lambda \, dE(\lambda),$$

so we get

$$T - \mu = \int_{[a,b]} \lambda \, dE(\lambda) = \underbrace{\int_{[a,c]} (\lambda - \mu) \, dE(\lambda)}_{T_1} + \underbrace{\int_{[d,b]} (\lambda - \mu) \, dE(\lambda)}_{T_2}.$$

Both T_1 and T_2 are reduced by $I = E(x) + E(d,b]$. If we write $x = x_1 + x_2$, then

$$\|T_1 x_1\| \ge |c - \mu| \|x_1\|, \qquad \|T_2 x_2\| \ge |d - \mu| \|x_2\|,$$

so we cannot make these arbitrarily small.

($\supseteq$): If $\mu \in \operatorname{spt}(E)$, let $\varepsilon > 0$. Then $E(\mu - \varepsilon, \mu + \varepsilon] \neq 0$. Pick x with $\|x\| = 1$ such that $E(\mu - \varepsilon, \mu + \varepsilon]x = x$. Then

$$\langle (T - \mu)x, y \rangle = \int_{[a,b]} (\lambda - \mu) \, dE(\lambda)x = \int_{[\mu - \varepsilon, \mu + \varepsilon} (\lambda - \mu) \, d\langle E(\lambda)x, y \rangle \le \varepsilon \|x\| \|y\|. \qquad \square$$

This will give us a better idea of what the set D is. It will be a set of eigenvalues.

23.3 Banach algebras

Definition 23.3. An **algebra** over $\mathbb{F}$ is a vector space A over $\mathbb{R}$ togetehr with a multiplication $A \times A \to A : (a,b) \mapsto ab$ which is associative and distributive with addition. An algebra has an **identity** if there is some $e \in A$ such that $ea = ae = a$ for all $a \in A$.

Definition 23.4. A **Banach algebra** is a Banach space A which is also an algebra such that

$$\|ab\| \le \|a\| \|b\|, \qquad \forall a, b \in A.$$

Example 23.1. Let X be compact and Hausdorff. Then $C(X)$ is a Banach algebra. If X is locally compact, $C_0(X)$ is a Banach algebra.

Example 23.2. $L^\infty(\mu)$ is a Banach algebra.

These are all commutative. Here are some noncommutative examples.

Example 23.3. $\mathcal{B}(X)$ is a Banach algebra if X is a Banach space.

Example 23.4. The collection of compact operators, $\mathcal{B}(X)$, is a Banach algebra (when X is a Banach space).

24 Banach Algebras

24.1 Convolution of measures

Understanding Banach algebras will help us obtain a better understanding of the spectral theorem.

Here is a motivating example.

Example 24.1. Let G be a locally compact Hausdorff group (so you can think of $G = \mathbb{R}^d$). Then there is a space $M(G)$ of finite, regular, $\mathbb{F}$-valued Borel measures. Given $\mu, \nu \in M(G)$, define the **convolution**

$$\mu * \nu(A) := \mu \times (\{(g,h) : gh \in A\}).$$

This is related to convolution of functions: $d\mu = f\,dm$ and $d\nu = f'\,dm$, then $f(\mu * \nu) = (f * f')\,dm$. We could alternatively define this by its action on $f \in C_0(G)$:

$$\int f\,d(\mu * \nu) = \iint f(gh)\,d\mu(g)\,d\nu(h).$$

This is distributive over addition, and associative:

$$\int f\,d((\mu * \nu) * \lambda) = \iint f(ghk)\,f\mu(g)\,d\nu(h)\,d\lambda(k).$$

Observe that

$$\left| \int f\,d(\mu * \nu) \right| = \left| \iint f(gh)\,d\mu(g)\,d\nu(h) \right| \leq \iint |f(gh)|\,d|\mu|\,d|\nu| \leq \|f\| \cdot \|\mu\| \cdot \|\nu\|,$$

so $\|\mu * \nu\| \leq \|\mu\|\|\nu\|$.

There is also an identity element with respect to convolution, δ_e. We have

$$(\delta_e * \mu)(A) = (\delta_e \times \mu)(\{(g,h) : gh \in A\}) = \mu(\{h : h \in A\}) = \mu(A),$$

and a similar property holds for right multiplication by δ_e. You can also check that $\delta_g * \delta_h = \delta_{gh}$. So $M(X)$ is a unital Banach algebra with convolution as the multiplication.

24.2 Invertibility and ideals

Definition 24.1. Let $\mathscr{A}$ be a Banch algebra. Then $x \in \mathscr{A}$ is **left-invertible** if there is some $y \in \mathscr{A}$ such that $yx = 1$, **right-invertible** if there is some $y \in \mathscr{A}$ such that $xy = 1$, and **invertible** if it is left and right invertible.

If x is left and right invertible, the inverses are the same: $z = yxz = y$. We write this as x^{-1}.

One important question is: Given an algebra, can we recover information about what generated it?

Definition 24.2. M is a **left ideal** in $\mathscr{A}$ if M is a vector subspace and $xy \in M$ for all $x \in A$ and $y \in M$. M is a **right ideal** in $\mathscr{A}$ if $\mathscr{M}$ is a vector subspace and $yx \in M$ for all $x \in A$ and $y \in M$. M is an **ideal** if it is a left and right ideal.

Example 24.2. The compact operators, $\mathcal{B}_0(X) \subseteq \mathcal{B}(X)$, form an ideal.

Example 24.3. Let $X \neq \varnothing$ be compact, and let $K \subsetneq X$ be closed and nonempty. Then $C(X) \supseteq \{f \in C(X) : f|_K = 0\} =: I(K)$. Then $K \subseteq L \iff I(L) \subseteq I(K)$.

These get bigger if K gets smaller In fact, there is a correspondence between maximal ideals of $C(X)$ and points of X. So we can recover X from $C(X)$.

Lemma 24.1. *Let $\mathscr{A}$ is a Banach algebra with identity, and let $x \in \mathscr{A}$ have $\|x-1\| < 1$. Then x is invertible.*

Proof. Let $y := \sum_{k=0}^{\infty}(1-x)^k$. The norm of the k-th term is $\leq 1\|1-x\|^k$. So this is an absolutely convergent series. So for any $z \in \mathcal{A}$, we have $zy = \sum_{k=0}^{\infty} z(1-x)^k$. This gives

$$(1-x)y = \sum_{k=0}^{\infty}(1-x)(1-x)^k = \sum_{k=0}^{\infty}(1-x)^{k+1} = y - (1-x)^0.$$

So we get $xy = (1-x)^0 = 1$. □

Corollary 24.1. *If $\|x-1\| < \varepsilon < 1$, then $\|x^{-1}-1\| < \frac{\varepsilon}{1-\varepsilon}$.*

Proof.

$$\|x^{-1}-1\| = \left\|\sum_{k=1}^{\infty}(1-x)^k\right\| \leq \sum_{k=1}^{\infty}\|1-x\|^k < \frac{\varepsilon}{1-\varepsilon}.$$ □

Corollary 24.2. *If $ba = 1$ and $\|c-a\| < 1/\|b\|$, then c is left-invertible.*

Proof. We have

$$\|bc-1\| = \|bc-ba\| \leq \|b\|\|c-a\| < 1.$$

so there is an $x = (bc)^{-1}$. So $(xb)c = 1$ means that xb is the inverse of c. □

Proposition 24.1. *Let $\mathscr{A}$ be a Banach algebra with identity, let G_ℓ be the left-invertible elements, let G_r be the right-invertible elements, and let $G = G_\ell \cap G_r$. Moreover, the map $G \to G : x \mapsto x^{-1}$ is continuous.*

Proof. Openness follows from the previous corollary. For continuity, if $x \in G$, suppose that $\|y-x\| < \varepsilon^{-1}$ for some small enough $\varepsilon > 0$. Then $\|x^{-1}y-1\| < \varepsilon\|x^{-1}\| < 1$. So

$$\|(x^{-1}y)^{-1}-1\| < \frac{\varepsilon\|x^{-1}\|}{1-\varepsilon\|x^{-1}\|}.$$

Then y^{-1} exists (because it is equal to $(x^{-1}y)^{-1}x^{-1}$, and

$$\|y^{-1}-x^{-1}\| < \frac{\varepsilon\|x^{-1}\|^2}{1-\varepsilon\|x^{-1}\|}.$$ □

24.3 Maximal ideals and quotients

Definition 24.3. A left/right/two-sided ideal M is **maximal** if it is

1. proper ($M \neq A$),

2. M is not properly contained in any other proper ideal.

Corollary 24.3. *If $\mathscr{A}$ has an identity, then*

1. *The closed of a left/right/two-sided ideal is an ideal of the same kind.*

2. *Maximal ideals are closed.*

Proof. Check the proof of (1).

If M is a maximal (e.g. two-sided) ideal, then $M \cap G_\ell = \varnothing$. This is because if $x \in M \cap G_\ell$, then there exists some y such that $yx = 1$. So $1 \in M$, but then $a = a1 \in M$ for all $a \in \mathscr{A}$. So $M = \mathscr{A}$. In fact, we have $\overline{M} \cap G_\ell = \varnothing$. Now $M = \overline{M}$ by maximality. $\square$

Example 24.4. The algebra $C_0(\mathbb{R}) \supseteq C_c(\mathbb{R}) = \{f : f|_{[-a,a]^c} = 0 \text{ for some } a\}$. This is a dense ideal. This tells us that this fact really relies on the existence of an identity.

Proposition 24.2. *Any proper (left/right/two-sided) ideal in any algebra is contained in a maximal (left/right/two-sided) ideal.*

Proof. Zorn's lemma. $\square$

Lemma 24.2. *Let $\mathscr{A}$ be a Banach algebra, and let M be a closed idea in $\mathscr{A}$. Then $\mathscr{A}/M$ is still a Banach algebra.*

Proof. Given $(x+M), (y+M) \in \mathscr{A}/M$, define $(x+M)(y+M) := xy + M$. To show that this is well-defined, we have that for any $m, n \in M$,

$$(x+m+M)(y+n+M) = xy + \underbrace{my + xn + mn}_{\in M} + M = xy + M.$$

To check that $\mathscr{A}/M$ is a Banach algebra, we have

$$\|(x+M)(y+M)\| = \|xy + M\| \leq \|xy\| \leq \|x\|\|y\|.$$

This is true for all x, y, so we can take the inf over x and y to get $\|(x+M)(y+M)\| \leq \|(x+M)\|\|(y+M)\|$. $\square$

24.4 The spectrum of an element

Definition 24.4. Let $\mathscr{A}$ have an identity, and let $x \in A$. The **spectrum** is $\sigma(x) = \{\lambda \in \mathbb{F} : x - \lambda$ not invertible, the **left-spectrum** is $\sigma_\ell(x) = \{\lambda \in \mathbb{F} : x - \lambda$ not left-invertible, and **right-spectrum** is $\sigma_r(x) = \{\lambda \in \mathbb{F} : x - \lambda$ not right-invertible. The **resolvent** is $\rho(x) = \mathbb{F} \setminus \sigma(x)$.

Example 24.5. Let X be a compact, Hausdorff space, and let $f \in C(X)$. Then $\sigma(f) = f(X)$ is the image of f. If $g(f - \lambda) = 1$, then $g(x) = \frac{1}{f(x)-\lambda}$ for all x.

25 Weakly Compact Operators

25.1 Weak compactness and reflexivity

In this lecture, X, Y, etc. will be real[13] Banach spaces. We will write B_X as the closed unit ball in $\|\cdot\|_X$.

Definition 25.1. $T \in \mathcal{B}(X,Y)$ is **weakly compact** if $\overline{T(B_X)}^{\mathrm{wk}(Y)}$ is is weakly compact in Y.

We will start with a bit of a digression. Suppose we have a Banach space X. We can embed it inside its dual X^** by $x \mapsto \widehat{x}$. The weak topology of X is the restriction of the weak* topology on X^{**} to X. We will denote by τ the weak* topology on X^{**}.

Proposition 25.1. *Let X be a Banach space, and let τ be the weak* topology on X^{**}. Then $\overline{B_X}^\tau = B_{X^{**}}$; i.e. B_X is τ-dense in $B_{X^{**}}$.*

Proof. Let $C := \overline{B_X}^\tau \subseteq B_{X^{**}}$. Suppose that $z \in B_{X^{**}} \setminus C$. Then, by Hahn-Banach, there exists a continuous linear functional f on (X^{**}, τ) and $\alpha \in \mathbb{R}$ such that $f(C) \leq \alpha < \alpha + \varepsilon \leq f(z)$. That is, there is a continuous linear functional on X such that

$$C(f) \leq \alpha < \alpha + \varepsilon \leq z(f)$$

Moreover, $C(f)$ contains a neighborhood of 0. By rescaling f, we can take $\alpha = 1$. Then $C(f) := \{y(f) : y \in C\} \supseteq \{f(x) : x \in B_X\}$. What this says is that $\|f\|_{X^*} \leq 1$. However, since $z(f)$ is the pairing of elements in the unit balls of their respective spaces, we should not have $z(f) > 1$, $\square$

Corollary 25.1. *X is dense in X^{**}.*

Corollary 25.2. *X is reflexive if and only if B_X is weakly compact.*

Proof. ($\Longrightarrow$): This is Banach-Alaoglu.

($\Longleftarrow$): If B_X is weakly compact, then $B_X \subseteq X^{**}$ is compact for τ. So B_X is closed in (X^{**}, τ). Then $B_X = \overline{B_X}^\tau = B_{X^{**}}$. $\square$

We can rephrase this corollary as the following:

Corollary 25.3. *X is reflexive if and only if I_X is weakly compact.*

Proposition 25.2. *If X or Y is reflexive, then every $T \in \mathcal{B}(X,Y)$ is weakly comapct.*

Proof. Consider $\overline{T(B_X)}^Y \subseteq Y$; we want to show that this is weakly compacts. If X is reflexive, then B_X is comapct, so $T(B_X)$ is weakly compact. On the other hand, if Y is reflelxvibe, $r(B_Y)$ is compact for all r. Now take r large enough so that $\overline{T(B_X)}^Y \subseteq rB_Y$. $\square$

[13] The story is not so different for the complex case.

Proposition 25.3. *If S or T is weakly compact, so is $S \circ T$.*

This is the same proof as before.

25.2 Characterization of weak compactness

Corollary 25.4. *$T \in \mathcal{B}(X, Y)$ is weakly compact if it has a factorization*

$$\begin{tikzcd} X \arrow[r, "T"] \arrow[d, "R"'] & Y \\ W \arrow[ur, "S"'] & \end{tikzcd}$$

where W is reflexive.

Theorem 25.1. *This is an exact characterization of weak compactness.*

Proof. Every T has the factorization

$$\begin{tikzcd} X \arrow[r, "T"] \arrow[d, "Q"'] & Y \\ X/\ker T \arrow[ur, "\overline{T}"'] & \end{tikzcd}$$

where $T(B_X) = \overline{T}(B_{X/\ker T})$. So it is enough to treat $\overline{T}$. So we may assume that $\ker T = \{0\}$.

Switch to regarding $X \subseteq Y$ with different norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, where $\|\cdot\|_Y|_X \lesssim \|\cdot\|_X$ (meaning there is an implicit constant in the inequality). We will find a W and $\|\cdot\|_W$ with $X \leq W \leq Y$ such that $(W, \|\cdot\|_W)$ is reflexive, $\|\cdot\|_Y|_W \lesssim \|\cdot\|_W$, and $\|\cdot\|_W|_X \leq \|\cdot\|_X$.

The idea here comes from the theory of **interpolated Banach spaces**. For $w \in Y$, let $p_n(w) := \inf\{2^{-n}\|x\|_X + 2^n\|y\|_W : x \in X, y \in Y, x + y = w\}$.[14] These are new norms on Y. Let

$$p(w) := \sqrt{\sum_n p_n(w)^2}, \qquad W := \{w : p(w) < \infty\}.$$

Check that

1. The p_n satisfy the triangle inequality, so p does, too. Then p is a norm on W, and (W, p) is a normed space. Moreover, W is a Banach space.

2. If $x \in X$, then $p_n(x) \leq 2^{-n}\|x\|_X$, so $p(x) \lesssim \|x\|_X$.

[14] Imagine you can pay for $x \in X$ with $\|\cdot\|_X$ and $y \in Y\|$ with $\|\cdot\|_Y$. Then this is the least you have to pay for w.

3. If $w \in W$, then $p_1(w) \leq p(w)$. So there exists a decompostion $w = x + y$ such that $\|x\|_X + \|y\|_W \leq p(w)$. So $\|w\|_Y = \|x + y\|_Y \lesssim p(w)$.

To finish, we will show that W is reflexive. What is the dual of (W, p)? We claim that $f \in W^*$ if and only if there is a sequence $(f_n)_n \in Y^*$ such that $f(w) = \sum_n f_n(w)$ for all w and $\sum_n p_n^*(f_n)^2 < \infty$, where p_n^* is the dual norm on Y^* induced by p_n.

Let $Y_n = (Y, p_n)$. Then W is isometrically isomorphic to a subspace $\{(y_n)_n \in \bigoplus_{L^2} Y_n : y_n = y_m \,\forall n, m\}$. Check that the dual of $\bigoplus_{L^2} Y_n$ is $\bigoplus_{L^2} Y_n^*$. So W^* is the quotient $(\bigoplus_{L^2} Y_n^*)/W^\perp$. This proves the claim.

To show that W is reflexive, we will show that I_W is weakly compact. Now suppose $(z_j)_j \subseteq B_W$; we want to find a weakly convergent subsequence in W. We may assume that $p(z_j) < 1$ for all j. Write $z_j = x_{j,n} + y_{j,n}$ such that $2^{-n}\|x_{j,n}\|_X + 2^n\|y_{j,n}\|_Y$ is very close to $p_n(z_j)$. In particular, $p_n(z_j) < 1$ for every n, so $\|x_{j,n}\|_X \leq 2^n$ and $\|y_{j,n}\|_Y \leq 2^{-n}$. Because B_X is weakly precompact in Y, for each n, there is a $y_n \in Y$ such that $x_{j,n} \xrightarrow{\text{wk}} y_n$ as $j \to \infty$. We also have that for each j, $\|x_{j,n} - x_{j,m}\|_Y = \|y_{j,n} - y_{j,m}\| \leq 2^{-n} + 2^{-m}$. Taking $j \to \infty$, we get $\|y_n - y_m\|_y \leq 2^{-n} + 2^{-m}$ (weak limits cannot increase norm). So $y_n \xrightarrow{\|\cdot\|_Y} y \in Y$ as $n \to \infty$. So we get that $z_j \to y$ (weakly in Y); check this from the definition.

To finish, we need $y \in W$, and we need to show that $z_j \to y$ weakly in W. The point is that for each n, $p_n(y) \leq \liminf_j p_n(z_j)$. Then Fatou's lemma gives $p(y) \leq \liminf_j p(z_j) < \infty$. So $y \in W$. Now for $f(y) = \sum_n f_n(y) = \sum_n \lim_j f_n(z_j)$. We can take out the limit outside the sum because $|f_n(y)| \leq p_n^*(f_n)p_n(y)$., which is a uniform bound. $\square$

26 The Spectrum and The Spectral Radius

26.1 The spectrum of an element

Let $\mathscr{A}$ be a Banach algebra with identity 1. Recall that if $\|a-1\|<1$, then a^{-1} exists and equals $\sum_{k\geq 0}(1-a)^k$. The **spectrum** of a is $\sigma(a)=\{z\in\mathbb{F}: z-a \text{ is not invertible in } \mathcal{A}\}$ (and similarly for right/left spectrum σ_r,σ_ℓ). The **resolvent** is $\rho(a)=\mathbb{F}\setminus\sigma(a)$.

Example 26.1. Let X be a compact, Hausdorff space. If $f\in C(X)$, then $\sigma(f)=f[X]$.

Example 26.2. Let X be a Banach space, and let $A\in\mathcal{B}(X)$. Then

$$\sigma(A)=\{\lambda\in\mathbb{F}: A-\lambda \textbf{ is not a bijection } X\to X\},$$

$$\sigma(A)=\{\lambda\in\mathbb{F}: \inf\{\|(A-\lambda)x : \|x\|=1\}=0\}.$$

Example 26.3. If $\mathbb{F}=\mathbb{R}$, we can have elements with empty spectrum. For example, take

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\in M_2(\mathbb{R}).$$

If we take this as an element in $M_2(\mathbb{C})$, the spectrum is nonempty. So the spectrum depends on the space the element is sitting in.

Theorem 26.1. *If $\mathbb{F}=\mathbb{C}$ and $a\in\mathscr{A}$, then $\sigma(a)$ is a nonempty, compact subset of $\{z\in\mathbb{C}:|z|\leq\|a\|\}$.*

Proof. Consider $z-a=z(1-a/z)$. If $|z|>\|a\|$, then $\|a/z\|<1$. Then $(z-a)^{-1}=\frac{1}{z}(1-\frac{a}{z})^{-1}$ exists. This tells us that $\sigma(a)\subseteq\{z\in\mathbb{C}:|z|\leq\|a\|\}$.

Also $\rho(a)=g^{-1}(\{\text{invertible elements}\})$, where $g(z)=z-a$ is continuous. Since the invertible elements form an open set, we have $\rho(a)$ is open. So $\sigma(a)$ is closed and bounded.

Consider the **resolvent function** $f:\rho(a)\to\mathscr{A}$ by $z\mapsto(z-a)^{-1}$. This is a continuous map from $\rho(a)\to\{\text{invertible elements in } \mathscr{A}\}$. If $|z|>\|a\|$, then

$$f(z)=\frac{1}{z}\left(1-\frac{a}{z}\right)^{-1}=\frac{1}{z}\sum_{k\geq 0}\frac{a^k}{z^k},$$

so we can get

$$\|f(z)\|\leq\frac{1}{|z|}\sum_{k\geq 0}\frac{\|a\|^k}{|z|^k}=O(1/|z|)\qquad\text{as } |z|\to\infty.$$

If $z\in\rho(a)$, then

$$\frac{f(z+h)+f(z)}{h}\xrightarrow{h\to 0}((z-a)^{-1})^2$$

(proven below). So we can say "f is holomorphic on $\rho(a)$."[15]

This shows that if $\sigma(a) = \varnothing$, then f is a holomorphic and bounded function. By a version of Liouville's theorem (proven below), f is constant. So $f = 0$. But this is a contradiction. $\square$

Lemma 26.1. *If $z \in \rho(a)$, then*

$$\frac{f(z+h)+f(z)}{h} \xrightarrow{h\to 0} ((z-a)^{-1})^2.$$

Proof. If $, y \in \mathscr{A}$ are invertible, then

$$\begin{aligned} x^{-1} - y^{-1} &= x^{-1}yy{-}1 - x^{-1}xy^{-1} \\ &= xx^{-1}(y-x)y^{-1}. \end{aligned}$$

This is called the **resolvent identity**. So

$$\begin{aligned} \frac{1}{h}[(z+h-a)^{-1} - (z-a)^{-1}] &= (z+h-a)^{-1}(z-a)^{-1} \\ &\xrightarrow{h\to 0} ((z-a)^{-1})^2. \end{aligned}$$ $\square$

Lemma 26.2 (Liouville's theorem for Banach-valued holomorphic functions)**.** *Let X be a Banach space. If $f : \mathbb{C} \to X$ is holomorphic and bounded, it is constant.*

Proof. For any $\varphi \in X^*$, $\varphi \circ f : \mathbb{C} \to \mathbb{C}$ is holomorphic and bounded, so it is constant. $\square$

This trick is a common way to transfer results from complex-valued holomorphic functions to Banach-valued ones.

26.2 Spectral radius

Definition 26.1. Let $a \in \mathscr{A}$. The **spectral radius** of a is $r(a) := \sup\{|z| : z \in \sigma(a)\}$.

Example 26.4. In $M_2(\mathbb{C})$, let

$$a = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Then $a^2 = 0$, so $\sigma(a) = \{0\}$. So the spectrum of a is nonempty, but it has zero spectral radius.

Theorem 26.2 (spectral radius formula)**.** *Let $a \in \mathscr{A}$. Then*

$$r(a) = \lim_{n\to\infty} \|a^n\|^{1/n}.$$

[15]This is a notion of holomorphic functions that take values in a Banach algebra.

Remark 26.1. Since the norm is submultiplicative, this is $\leq \|a^m\|^{1/m}$ for any m. So this equals $\inf_n \|a^n\|^{1/n}$.

Proof. ($\leq$): We know that $r(a) \leq \|a\|$. We claim that $\sigma(a^m) = \{z^m : z \in \sigma(a)\}$.[16] If $\lambda \in \mathbb{C}$, then $a^m - \lambda = \prod_{i=1}^m (a - \omega_i)$, where the ω_i are the m-th roots of λ. Since each $a - \omega_i$ is invertible, $a^m - \lambda$ is invertible. If $a^m - \lambda$ is invertible, then $(a - \omega_1)^{-1} = \prod_{i=2}^m (a - \omega_i)(a^m - \lambda)^{-1}$. This proves the claim and gives us $r(a^m) = r(a)^m$ for any m. So $r(a) = r(a^m)^{1/m} \leq \|a^m\|^{1/m}$ for all m.

($\geq$): Let $h(w) = (\frac{1}{w} - a)^{-1}$ for w such that $\frac{1}{w} \in \rho(a)$. Extend this so $h(0) = 0$. As before,

$$h(w) = w \sum_{k \geq 0} w^k a^k \qquad \forall |w| < \|a\|^{-1},$$

and h is holomorphic on $\{0\} \cup \{\frac{1}{z} : z \in \rho(a)\}$. Now we use a fact from complex analysis (which extends to this case): By Hadamard's formula for the radius of convergence of a series, the supremal R such that h has a holomorphic extension to the ball $B_{\mathbb{C}}(0, R)$ equals the radius of convergence of the series; this is $\lim_n 1/\|a^n\|^{1/n}$. So $\inf\{1/|z| : z \in \sigma(a)\} = \lim_n 1/\|a^n\|^{1/n}$. $\square$

Remark 26.2. Here is another way to show that the sequence converges. We have $\|a^{n+m}\| \leq \|a^n\| \cdot \|a^m\|$, so $\log \|a^{n+m}\| \leq \log \|a^n\| + \log \|a^m\|$. Now use Fekete's subadditive lemma.

Example 26.5. For $f \in L^2([0,1])$, the Volterra operator is

$$Vf(x) = \int_0^x f = \int_0^1 \mathbb{1}_{\{y \leq x\}} f(y)\, dy.$$

Then $\sigma(V) = \{0\}$, so $r(V) = 0$.

Proposition 26.1. *If* $z \in \rho(a)$, *then* $\|(z-a)^{-1}\| \geq \frac{1}{\operatorname{dist}(z, \sigma(a))}$. *In other words,* $\operatorname{dist}(z, \sigma(a)) \geq 1/\|(z-a)^{-1}\|$.

If z is in the spectrum, $(z-a)^{-1}$ doesn't exist. This says that if z is close to the spectrum, then this blows up.

Proof. If $h \in \mathbb{C}$ with $|h| < \frac{1}{\|(z-a)^{-1}\|}$, then

$$z + h - a = (z - a)(h(z-a)^{-1} + 1)$$

is invertible. So $B(z, 1/\|(z-a)^{-1}\|) \subseteq \rho(a)$. $\square$

[16] This is a special case of the spectral mapping theorem, which we will discuss later.

26.3 Riesz functional calculus

Here is a teaser for what we will discuss next time.

If $a \in \mathscr{A}$, then the resolvent map $f : \rho(a) \mapsto \mathscr{A}$ takes $z \mapsto (z-a)^{-1}$. Any holomorphic $f : G \to \mathscr{A}$ satisfies Cauchy's integral formula. As a result, if G is an open subset of $\mathbb{C}$ with $G \supseteq \sigma(a)$, then let $\Gamma = \gamma_1 \cup \cdots \cup \gamma_m$ wind once around any $z \in \sigma(a)$ and 0 times around any $z \in \mathbb{C} \setminus G$. Then if $f : G \to \mathbb{C}$ is holomorphic, define

$$f(a) = \frac{1}{2\pi i} \int_\Gamma f(z)(z-a)^{-1}\, dz \in \mathscr{A}.$$

This allows us to produce more elements of our Banach algebra.

27 Riesz Functional Calculus and The Gelfand Transform

27.1 Riesz functional calculus

Theorem 27.1 (Cauchy's integral formula)**.** *Let $G \subseteq \mathbb{C}$ be open, let $f : G \to \mathbb{C}$ be holomorphic, and let $\Gamma = \gamma_1 \cup \cdots \cup \gamma_n$ be a system of contours such that the total winding number around any point in $\mathbb{C} \setminus G$ is 0. Let $z \in G \setminus \Gamma$. Then*

$$(\textit{winding \# of } \Gamma \textit{ around } z) f^{(k)}(z) = \frac{k!}{2\pi i} \oint_\Gamma \frac{1}{(z-w)^{k+1}} f(w)\, dw \qquad \forall k \geq 0.$$

This is ok when the target space is a Banach space X. The idea is that if $\mathscr{A}$ is a Banach algebra over $\mathbb{C}$ with identity and $a \in \mathscr{A}$, we let G be an open neighborhood of $\sigma(a)$. Then there exists some $\Gamma = \gamma_1 \cup \cdots \ cup \gamma_m$ such that the winding number is 0 around any point in $\mathbb{C} \setminus G$ and 1 around any point in $\sigma(a)$.

Now define
$$f(a) := \frac{1}{2\pi} \oint_\Gamma f(w) \cdot (a-w)^{-1}\, dw.$$
This is well-defined because if we define this with Γ and Γ', the difference is a sum of zeros by the Cauchy integral formula. Here are the properties of the functional calculus we get from this:

Proposition 27.1. *Let $\mathscr{A}$ be a Banach algebra with identity, let $a \ in \mathscr{A}$, and* $\mathrm{Hol}(a)$ *be the functions holomorphic on the spectrum of a. Then*

1. $\mathrm{Hol}(a) \to \mathscr{A} : f \mapsto f(a)$ *is an algebra homomorphism.*
2. *If $f(z) = \sum_{k \geq 0} \alpha_k z^k$ has radius of convergence $> r(a)$, then*
$$f(a) = \sum_k \alpha_k a^k$$
3. *If $f_1, f_2, \ldots, f$ are all holomorphic on $G \supseteq \sigma(a)$ and $f_n \to f$ uniformly on compact subsets of G, then $f_n(a) \to f(a)$ in $\| \cdot \|_\infty$.*

Remark 27.1. These properties uniquely determine this algebra homomorphism. The proof uses Runge's theorem; first do this for rational functions, and then extend via density.

27.2 Abelian Banach algebras

To return to the spectral theorem, we first need some considerations about abelian Banach algebras.

Example 27.1. Let $C(X)$ be a compact Hausdorff space. Then $C(X)$ is an abelian, Banach algebra. The maximal ideals in $C(X)$ are the sets $\{f : f(x) = 0\}$. This tells you that you can recover X by looking at the maximal ideals of $C(X)$.

Theorem 27.2 (Gelfand-Mazur). *Let $\mathscr{A}$ be a Banach algebra with identity. Assume that $\mathscr{A}$ is a division ring (every nonzero element of $\mathscr{A}$ is invertible). Then $\mathscr{A} = \mathbb{C} \cdot 1$.*

Remark 27.2. We are not assuming $\mathscr{A}$ is abelian, but this follows from the proof.

Proof. Let $a \in \mathscr{A}$, and choose $\lambda \in \sigma(a)$. Then $a - \lambda$ is not invertible, so $a - \lambda = 0$. So $a = \lambda 1$. $\square$

Proposition 27.2. *Let $\mathscr{A}$ be a unital, abelian Banach algebra over $\mathbb{C}$ If $h : \mathscr{A} \to \mathbb{C}$ is a homomorphism (sending $1 \mapsto 1$), then* $\ker h$ *is a maximal ideal, and all maximals ideals arise this way uniquely.*

Proof. If $a \in \ker h$ and $b \in \mathscr{A}$, then $h(ab) = 0h(b) = 0$, so $ab \in \ker h$. So $\ker h$ is an ideal.

If $\ker H \subseteq M \subsetneq \mathscr{A}$, where M is an ideal, then $h(M)$ is a subspace of $\mathbb{C}$ (and actually an ideal). Then $h(M) = \mathbb{C}$ or $\{0\}$. Since M is proper, we get $h(M) = \{0\}$. So $M = \ker h$ is an ideal and is in fact maximal.

Now let M be a maximal ideal. There is the quotient map $Q : \mathscr{A} \to \mathscr{A}/M$. Since M is maximal, $\mathscr{A}$ has no nontrivial ideals. Then all nonzero elements are invertible, so by Gelfand-Mazur, we get that $\mathscr{A}/M = \mathbb{C}1_{\mathscr{A}/M}$. If we call the isomorphism $\pi : \mathscr{A}/M \to \mathbb{C}1_{\mathscr{A}/M}$, then $M = \ker(\pi \circ Q)$. $\square$

Lemma 27.1.

$$\{a \in \mathscr{A} : a \textit{ not invertible}\} = \bigcup_{\substack{M\, max. \\ ideal}} M.$$

Proof. If a is in the left hand side, then $\{ab : b \in \mathscr{A}\}$ is an ideal without 1. So it is contained in a maximal ideal.

On the other hand, if $ab = 1$ and $a \in M$ for some ideal, then $1 \in M$. So $b = 1b \in M$, making $M = \mathscr{A}$. $\square$

We take the convention that $\|1\|_{\mathscr{A}} = 1$.

Proposition 27.3. *Any homomorphism $h : \mathscr{A} \to \mathbb{C}$ is continuous with $\|h\|_{\mathscr{A}^*} = 1$.*

The idea is that homomorphisms are a special kind of linear functional, the ones that preserve multiplication. This says that they are all contained in the unit ball of $\mathscr{A}^*$.

Proof. h is continuous because $\ker h = M$ is closed. To show the norm estimate, we have $h(1) = 1$, which gives $\|h\|_{\mathscr{A}^*} \geq 1$. Now let $a \in \mathscr{A}$, and let $\lambda = h(a) \in \mathbb{C} \setminus \{0\}$. Then $h(1 - \frac{a}{\lambda}) = 0$, so $1 - \frac{a}{\lambda}$ is not invertible. Then $\|\frac{a}{\lambda}\| \geq 1$. This gives $|\lambda| \leq \|a\|$. $\square$

27.3 Maximal ideal spaces and the Gelfand transform

Definition 27.1. The **maximal ideal space** of $\mathscr{A}$ is

$$\Sigma = \{h \in \mathscr{A}^* : h \text{ unital homomorphism}\}.$$

Proposition 27.4. *Σ is compact for the weak* topology.*

Proof. $\Sigma \subseteq B_{\mathscr{A}^*}$, which is compact by Banach-Alaoglu. Also,

$$\Sigma = \{h \in B_{\mathscr{A}^*} : h(1) = 1\} \cap \bigcap_{a,b \in \mathscr{A}} \{h \in B_{\mathscr{A}^*} : h(ab) - h(a)h(b) = 0\},$$

which is an intersection of weak*-closed sets. So Σ is compact. $\square$

Theorem 27.3. *If X is a nonempty, compact, Hausdorff space, then $x \mapsto \delta_x$ is a homeomorphism $X \to \Sigma$, the maximal ideal space at $C(X)$.*

Proof. We only need to show that every maximal ideal M in $C(X)$ has the form $\{f : f(x) = 0\}$. By Riesz representation, $h(f) = \int_X f\, d\mu$. Since $\|h\| = 1$, $h(1) = 1$. So μ is a probability measure. Now $f \in M \iff \int f\, d\mu = 0$. And if $f \in M$, then $|f|^2 = f\overline{f} \in M$¡ so $\int |f|^2\, d\mu = 0$. Check that this implies that the support of μ is a singleton. $\square$

Proposition 27.5. *If $a \in \mathscr{A}$, then $\sigma(a) = \{h(a) : h \in \Sigma\}$.*

Proof.

$$\begin{aligned} \lambda \in \sigma(a) &\iff a - \lambda \text{ not invertible} \\ &\iff a - \lambda \text{ is contained in some maximal } M \\ &\iff h(a) - \lambda = 0 \text{ for some } h \in \Sigma. \end{aligned}$$

$\square$

Definition 27.2. The **Gelfand transform** of $a \in \mathscr{A}$ is the function $\widehat{a} : \Sigma \to \mathbb{C}$ with $\widehat{a}(h) = h(a)$.

Now we can basically write the functional calculus but in reverse:

Theorem 27.4. *$a \mapsto \widehat{a}$ is a continuous homomorphism $\mathscr{A} \to C(\Sigma)$. Its kernel is $\operatorname{rad}(A) = \bigcap_{h \in \Sigma} \ker h$, and*

$$\|\widehat{a}\|_{\sup} = \lim_n \|a^n\|^{1/n} \le \|a\|.$$

Proof. $\widehat{a} \in C(\Sigma)$ because $\widehat{a} = h(a)$ is the kind of functional which defines the weak* topology. The expression for the norms is the spectral radius formula. Lastly, $\widehat{a} = 0$ for all h if and only if $h \in \operatorname{rad} A$. This is an ideal (and intersection of ideals is an ideal). $\square$

The Gelfand transform is the canonical "best possible way to compare $\mathscr{A}$ to continuous functions on something." It's the best way because if we have another map $\mathscr{A} \to C(X)$, the radical will still get sent to 0. Next time, we will discuss conditions under which this map is surjective.

Example 27.2. Let $V \in \mathcal{B}(L^2([0,1]))$ be the Volterra operator, so $\sigma(V) = \{0\}$. Then $\|V^n\|^{1/n} \to 0$. Let $\mathscr{A} = \overline{\{p(V) : p \in \mathbb{C}[x]\}}$. This is an abelian Banach algebra with identity. The Gelfand transform sends $V \mapsto \widehat{V} = 0$. Then if $p(x) = \sum_{k=0}^n a_k x^k$, $\widehat{p(V)} = a_0 \cdot 1$. So the kernel of the Gelfand transform is the unique maximal ideal. You can check that this is $\overline{\{p(V) : p \in \mathbb{C}[X], p(0) = 0\}}$.

28 C^*-Algebras and Normal Functional Calculus

28.1 C^*-algebras

Definition 28.1. On an algebra $\mathscr{A}$ over $\mathbb{C}$, an **involution** is a map $\mathscr{A} \to \mathscr{A} : a \mapsto a^*$ such that

1. $(a^*)^* = a$,
2. $(ab)^* = b^* a^*$,
3. $(\lambda a + b))^* = \overline{\lambda} a^* + b^*$ for all $\lambda \in \mathbb{C}$, $a, b \in \mathscr{A}$.

This this $\mathscr{A}$ is a ***-algebra**.

Definition 28.2. A Banach algebra with an involution is a C^***-algebra** if

$$\|a\|^2 = \|a^* a\| \qquad \forall a \in \mathscr{A}.$$

Example 28.1. Operators on a Hilbert space form a C^*-algebra:

$$\|Tx\|^2 = \langle Tx, Tx \rangle = \langle T^*Tx, x \rangle \leq \|T^*T\| \|x\|^2.$$

Example 28.2. $\mathscr{B}_0(H)$ is a C^*-algebra (without identity, unless $\dim H < \infty$).

Example 28.3. If X is compact and Hausdorff, then $C_{\mathbb{C}}(X)$ is a C^*-algebra with $f^* := \overline{f}$.

Henceforth, we will only deal with unital C^*-algebras.

Proposition 28.1. *Let $\mathscr{A}$ be a C^*-algebra. Then for all $a \in \mathscr{A}$, $\|a^*\| = \|a\|$. If $\mathscr{A}$ is unital, then $1^* = 1$ and $\|1\| = 1$.*

Proof. We have $\|a\|^2 = \|a^*\| \|a\| \leq \|a^*\| \|a\|$, which gives $\|a\| \leq \|a^*\|$. Switching a and a^*, we get the other inequality.

Suppose $a \in \mathscr{A}$. Then $1^* a = (a^* 1)^* = (a^*)^* = a$ (and same for right multiplication), so $1^* = 1$. This gives $\|1\|^2 = \|1^* 1\| = \|1\|$, so $\|1\| = 0$ or 1. But this is a norm, so $\|1\| = 1$. □

28.2 Self-adjoint, normal, and unitary elements

Definition 28.3. $a \in \mathscr{A}$ is

- **self-adjoint** if $a = a^*$,
- **normal** if $aa^* = a^* a$
- **unitary** if $a^* = a^{-1}$.

Proposition 28.2. *Let $a \in \mathscr{A}$.*

1. *If a is invertible, then a^* is invertible, and $(a^*)^{-1} = (a^{-1})^*$.*
2. *$a = x + iy$, where x, y are self-adjoint.*
3. *If u is unitary, $\|u\| = 1$.*
4. *If a is normal, its spectral radius is $r(a) = \|a\|$.*

Proof. 1. We have $a^*(a^{-1})^* = (a^{-1}a)^* = 1^* = 1$.

2. Let $x = \frac{a+a^*}{2}$ and $y = \frac{a-a^*}{2i}$.
3. $\|u\|^2 = \|u^*u\| = 1$.
4. We know that $r(a) = \lim_n \|a^n\|^{1/n}$. In particular, we can take a subsequence with powers of 2. We have

$$\|a^{2^k}\|^{2^{-k}} = \|a^{2^{k-1}}a^{2^{k-1}}\|^{2^{-k}} = \|a^{2^{k-1}}\|^{2^{-(k-1)}} = \cdots = \|a\|.$$

So $\lim_k \|a^{2^k}\|^{2^{-k}} = \|a\|$. □

Proposition 28.3. *Let $h : \mathscr{A} \to \mathbb{C}$ be a nonzero homomorphism. Then*

1. *If $a = a^*$, then $h(a) \in \mathbb{R}$. In particular, if $\mathscr{A}$ is abelian, $\sigma(a) \subseteq \mathbb{R}$.*
2. *$h(a^*) = \overline{h(a)}$.*
3. *$h(a^*a) \geq 0$.*
4. *If u is unitary, then $|h(u)| = 1$.*

Proof. 1. We know $\|h\|_{A^*} \leq 1$. Let $t \in \mathbb{R}$, and consider $h(a + it)$. We have

$$\begin{aligned} |h(a) + it|^2 &= |h(a + it)|^2 \\ &\leq \|a + it\|^2 \\ &= (a + it)^*(a + it) \\ &= \|(a - it)(a + it) \\ &= \|a^2 + t^2\| \\ &\leq \|a^2\| + t^2. \end{aligned}$$

If $h(a) = x + iy$, then we get $x^2 + (y + t)^2 \leq \|a\|^2 + t^2$ for all t. This gives us $x^2 + y^2 + 2yt \leq \|a\|^2$ for all t. So we get $y = 0$.

2. If $a = a + iy$, where x, y are self-adjoint, then $a^* = x - iy$. Now apply h.
3. $h(a^*a) = h(a^*)h * (a) = |h(a)|^2$.
4. We have $1 = uu^*$. Now apply h. □

28.3 The Gelfand Transform and functional calculus for normal elements

The extra structure here makes it clear why the spectral theorem is true.

Theorem 28.1. *If $\mathscr{A}$ is an abelian C^*-algebra, then the Gelfand transform $\mathscr{A} \to C(\Sigma)$ is an isometric *-isomorphism,*

Proof. It preserves the involution because

$$\widehat{a^*}(h) = h(a^*) = \overline{h(a)} = \overline{\widehat{a}}(h).$$

If $a \in \mathscr{A}$, then a is normal, so $\|\widehat{a}\|_{\sup} = r(a) = \|a\|$; so the transform is isometric.

To check that this is surjective, by the Stone-Weierstrass theorem, we need only check that $\widehat{\mathscr{A}}$ separates points. If $h_1 \neq h_2$, then let $a \in \mathscr{A}$ be such that $h_1(a) \neq h_2(a)$. Then $\widehat{a}(h_1) \neq \widehat{a}(h_2)$. $\square$

This gives us a full functional calculus: if $\mathscr{A}$ is any abelian C^*-subalgebra of $\mathcal{B}(H)$, then there exists an isometric *-algebra isomorphism $C(\Sigma) \to \mathscr{A}$, namely the inverse of the Gelfand transform. If $N \in \mathcal{B}(H)$ is normal, then $C^*(N) := \overline{\{p(N, N^*) : p \in C[z, \overline{z}]\}}$ is an abelian C^*-algebra which contains N. So normal operators are precisely the ones that have a functional calculus like this.

Proposition 28.4. *In this example, $\Sigma_{C^*(N)}$ is homeomorphic to $\sigma(N) \subseteq \mathbb{C}$ under the homeomorphism $\widehat{N} : \Sigma_{C^*(N)} \to \mathbb{C}$.*

Proof. We know that $\widehat{N}(\Sigma) = \{h(N) : h \in \Sigma\} = \sigma(N)$. We need to check that if $\widehat{N}(h_1) = \widehat{N}(h_2)$, then $h_1 = h_2$. We have $h_1(N) = h_2(N)$, so

$$h_1(N^*) = \overline{h_1(N)} = \overline{h_2(N)} h_2(N^*).$$

So h_1, h_2 agree on any polynomial in N, N^*, which means $h_1 = h_2$. $\square$

Let $\Phi : C(\sigma(N)) \to C^*(N)$ be our functional calculus. For any $f \in C(\sigma(N))$ and $x, y \in H$, consider

$$\langle \Phi(f)x, y \rangle = \int_{\sigma(N)} f \, d\mu_{x,y}$$

for some complex-valued Borel measure $\mu_{x,y}$. The right hand side is defined for all bounded Borel functions f on $\sigma(N)$. Use this to define $\Phi(f)$ for some functions. This extends Φ to a functional calculus from all bounded, Borel functions on $\sigma(N)$ to $\mathcal{B}(H)$. To get a spectral measure of N, use $\Phi(\mathbb{1}_A)$ for all Borel $A \subseteq \sigma(N)$.

Remark 28.1. We can look at abelian algebras generated by multiple commuting operators. There is a form of the spectral theorem in that setting, too.