

Lecture Notes to Functional Analysis
Winter 2020, Technion

Lectures by Liran Rotem
Typed by Elad Tzorani

February 28, 2021

Contents

1	Preliminaries	5
1.1	Banach Spaces & Examples	5
1.1.1	Applications of Duality & Some Measure Theory	11
2	Linear Functionals	15
2.1	Bounded Linear Functionals	15
2.1.1	The Hahn-Banach Theorem	15
2.1.2	Application of Hahn-Banach - Banach Limits	18
2.1.3	Application of Hahn-Banach - Reflexive Spaces	20
2.1.4	Application of Hahn-Banach - Approximation Theory	21
2.1.5	Application of Hahn-Banach - Convex Separation	23
2.2	Banach Spaces	26
2.2.1	Baire Category Theorem & Uniform Boundedness	26
2.2.2	Applications to Harmonic Analysis	27
2.2.3	The Open Mapping Theorem	29
2.2.4	Application of the Open Mapping Theorem to Harmonic Analysis	31
2.2.5	The Closed Graph Theorem	32
2.2.6	Projections and Quotient Spaces	33
2.2.7	Quotient Spaces	34
2.2.8	Schauder Bases	36
2.3	Finite-Dimensional Spaces	38
2.3.1	Definitions	38
2.3.2	Geometric Interpretation	42
3	Weak Topologies	45
3.1	Weak Topologies	45
3.1.1	Definitions	45
3.1.2	Weak-* Topologies	49
3.1.3	Application: Haar Measures	54
3.2	Locally-Convex Spaces	55

4	Banach Algebras	63
4.1	Banach Algebras	63
4.1.1	Definition & Examples	63
4.1.2	Banach-Valued Analytic Functions	66
4.1.3	Ideals	70
4.1.4	Multiplicative Characters	72
4.1.5	The Gelfand Transform	73
4.1.6	C* Algebras	76
4.1.7	The Spectral Theorem	79
5	Unbounded Operators	89
5.0.1	The Spectral Theorem	95
5.0.2	Self-Adjoint Extensions	99

Chapter 1

Preliminaries

1.1 Banach Spaces & Examples

Definition 1.1.1 (Norm). Let X be a vector space over $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ (and assume from now until mentioned otherwise these are the only fields). A *norm* on X is a function

$$\|\cdot\| : X \rightarrow [0, \infty)$$

such that the following hold.

1. $\|x\| = 0$ iff $x = 0$.
2. $\|\lambda x\| = |\lambda| \|x\|$ for all $\lambda \in \mathbb{F}$.
3. $\|x + y\| \leq \|x\| + \|y\|$.

Remark 1.1.2. A norm $\|\cdot\|$ defines a metric by $d(x, y) = \|x - y\|$. This defines a topology generated by the open balls in the metric.

Definition 1.1.3 (Banach Space). A normed space $(X, \|\cdot\|)$ is a *Banach space* if it's complete (i.e. if every Cauchy sequence in it converges).

Definition 1.1.4 (Inner Product). Let X be a vector space over \mathbb{F} . An *inner product* on X is a function

$$\langle \cdot, \cdot \rangle : X \times X \rightarrow \mathbb{F}$$

such that the following hold.

1. $\langle \lambda x_1 + x_2, y \rangle = \lambda \langle x_1, y \rangle + \langle x_2, y \rangle$.
2. $\langle x, y \rangle = \overline{\langle y, x \rangle}$.
3. $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0$ iff $x = 0$.

Remark 1.1.5. An inner product $\langle \cdot, \cdot \rangle$ defines a norm $\|x\| = \sqrt{\langle x, x \rangle}$. This can be proven by Cauchy-Schwarz $|\langle x, y \rangle| \leq \sqrt{\|x\| \|y\|}$.

Definition 1.1.6 (Hilbert Space). A *Hilbert space* is a complete inner product space.

Example 1.1.7 (ℓ_p Spaces). Fix $p \in [1, \infty)$ and $(a_n)_{n=1}^\infty$ a sequence in \mathbb{F} . Define

$$\|(a_n)_{n=1}^\infty\|_p := \left(\sum_{n=1}^\infty |a_n|^p \right)^{\frac{1}{p}}$$

$$\|(a_n)_{n=1}^\infty\|_\infty := \sup_{n \in \mathbb{N}_+} |a_n|$$

and

$$\ell_p := \left\{ (a_n)_{n=1}^\infty \mid \|(a_n)_{n=1}^\infty\|_p < \infty \right\}$$

$$\ell_\infty := \left\{ (a_n)_{n=1}^\infty \mid \|(a_n)_{n=1}^\infty\|_\infty < \infty \right\}$$

Example 1.1.8. Define

$$c := \left\{ (a_n)_{n \in \mathbb{N}_+} \in \ell_\infty \mid \exists \lim_{n \rightarrow \infty} a_n \in \mathbb{C} \right\}.$$

Consider this with the ∞ -norm, as a closed subspace of ℓ_∞ .

Similarly, define

$$c_0 := \left\{ (a_n)_{n \in \mathbb{N}_+} \in \ell_\infty \mid \lim_{n \rightarrow \infty} a_n = 0 \right\}.$$

Example 1.1.9 (Continuous Functions over a Compact Space). Let X be a compact topological space. Define

$$\mathcal{C}(X) := \{f: X \rightarrow \mathbb{F} \mid f \text{ is continuous}\}.$$

This is a Banach space with $\|f\| = \max_{x \in X} |f(x)|$.

Definition 1.1.10 (Support of a Function). Let $f: X \rightarrow V$ be a function of sets from a topological space into a vector space. Define *the support of f* to be

$$\text{supp}(f) := \overline{\{x \mid f(x) \neq 0\}}.$$

Example 1.1.11 (Continuous Functions over Locally Compact Hausdorff Spaces). Let X be a locally compact Hausdorff topological space (e.g. $X = \mathbb{R}^n$). Define

$$\mathcal{C}_c(X) := \{f: X \rightarrow \mathbb{F} \mid f \text{ is compactly supported and continuous}\}.$$

This is a normed space which is *not* complete.

Define $f(x) = e^{-x^2}$. This can be approximated by continuous functions f_n which agree with it on $[-n, n]$ and go linearly to zero outside $[-n, n]$ until the boundary of $[-n-1, n+1]$. $(f_n)_{n=1}^\infty \subseteq \mathcal{C}_c(X)$ is Cauchy but not convergent because $f \notin \mathcal{C}_c(X)$.

Definition 1.1.12 (Completion of a Normed Space). Let Y be a normed space. There is a Banach space \hat{Y} such that $Y \leq \hat{Y}$ and $\bar{Y} = \hat{Y}$. This is called *the completion of Y* .

Example 1.1.13.

$$\widehat{\mathcal{C}_c(X)} = \mathcal{C}_0(X) := \left\{ f: X \rightarrow \mathbb{F} \mid \begin{array}{l} f \text{ is continuous} \\ \lim_{|x| \rightarrow \infty} f(x) = 0 \end{array} \right\}.$$

Example 1.1.14. Let (X, Σ, μ) be a measure space and let $p \in [1, \infty)$. We define

$$L_p(X, \Sigma, \mu) := \left\{ f: X \rightarrow \mathbb{F} \mid \begin{array}{l} f \text{ is Borel measurable} \\ \|f\|_p < \infty \end{array} \right\}$$

where

$$\|f\|_p := \left(\int_X |f|^p d\mu \right)^{\frac{1}{p}}$$

and where we take the Borel measure on \mathbb{F} .

Remark 1.1.15. If $X = \mathbb{N}_+$ and μ is the counting measure $\mu(a) = |A|$ then $L_p(X, \Sigma, \mu) = \ell_p$.

Remark 1.1.16. $\|\cdot\|_p$ which we defined isn't *exactly* a norm, since there are measurable functions $f: X \rightarrow \mathbb{F}$ which are 0 almost everywhere, but not everywhere.

We therefore look at L_p as the space of equivalence classes of functions up to equivalence almost-everywhere of functions.

On $L_p(X, \Sigma, \mu)$, $\|\cdot\|_p$ is a norm. This follows from the following inequality.

Theorem 1.1.17 (Hölder Inequality). If $p, q \geq 1$ and $\frac{1}{p} + \frac{1}{q} = 1$ then for all $f, g: X \rightarrow \mathbb{F}$ it holds that

$$\left| \int_X fg d\mu \right| \leq \left(\int_X |f|^p d\mu \right)^{\frac{1}{p}} \left(\int_X |g|^q d\mu \right)^{\frac{1}{q}}.$$

Corollary 1.1.18.

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Corollary 1.1.19. On $L_p(X, \Sigma, \mu)$, $\|\cdot\|_p$ is a norm.

Theorem 1.1.20. $L_p(X, \Sigma, \mu)$ is a Banach space.

Proof. It's enough to prove that if

$$\sum_{n=1}^{\infty} \|f_n\|_p < \infty$$

then $\sum_{n=1}^{\infty} f_n$ converges.

Assume first $f_n \geq 0$. Define $g := \sum_{n=1}^{\infty} f_n \in [0, \infty]$ and $g_N := \sum_{n=1}^N f_n$. We know by the triangle inequality that

$$\|g_N\|_p \leq \sum_{n=1}^N \|f_n\|_p \leq \sum_{n=1}^{\infty} \|f_n\|_p =: S.$$

Now

$$\|g\|_p^p = \int_X |g|^p d\mu = \int_X \lim_{N \rightarrow \infty} |g_N|^p d\mu \stackrel{\text{Monotone Convergence}}{=} \lim_{N \rightarrow \infty} \int_X |g_N|^p d\mu \leq S^p$$

hence $g \in L_p$. In particular, $g < \infty$ almost-everywhere.

Finally,

$$\|g - g_N\|_p = \left\| \sum_{n=N+1}^{\infty} f_n \right\|_p \leq \sum_{n=N+1}^{\infty} \|f_n\|_p \xrightarrow{N \rightarrow \infty} 0.$$

In general, we can write $f_n = f_n^+ - f_n^-$ where f_n^+, f_n^- are positive. We have $|f_n^\pm| \leq |f_n|$ so $\|f_n^\pm\|_p \leq \|f_n\|_p$. So $\sum_{n=1}^{\infty} \|f_n^\pm\|_p < \infty$ so $\sum_{n=1}^{\infty} f_n^\pm$ converge and so does $\sum_{n=1}^{\infty} (f_n^+ - f_n^-)$. \blacksquare

Definition 1.1.21. Let $f: X \rightarrow \mathbb{F}$. Define

$$\text{essup}(f) := \inf \{M > 0 \mid \mu(\{x \mid |f(x)| \geq M\}) = 0\}.$$

Definition 1.1.22 (L_∞). Define $\|f\|_\infty = \text{essup}(f)$ and

$$L_\infty(X, \Sigma, \mu) = \{f: X \rightarrow \mathbb{F} \mid \|f\|_\infty < \infty\}.$$

Exercise 1. Assume $\mu(X) < \infty$.

1. If $1 \leq p < q \leq \infty$ then $L_q \subseteq L_p$.
2. If $f \in L_\infty(X, \Sigma, \mu)$ then $\|f\|_\infty = \lim_{p \rightarrow \infty} \|f\|_p$.

Definition 1.1.23 (The Operator Norm). Let X, Y be normed spaces and let $T: X \rightarrow Y$ be linear. Define the *operator norm* as

$$\|T\| = \sup_{x \in X \setminus \{0\}} \frac{\|Tx\|_Y}{\|x\|_X} = \sup_{\substack{x \in X \\ 0 < \|x\| \leq 1}} \|Tx\|.$$

Definition 1.1.24 (Bounded Linear Map). A linear map $T: X \rightarrow Y$ between normed spaces is *bounded* if $\|T\| < \infty$.

Fact 1.1.25. Let $T: X \rightarrow Y$ be a linear map between normed spaces. The following are equivalent.

1. T is bounded.

2. T is continuous.

3. T is continuous at $x = 0$.

4. T is Lipschitz.

Notation 1.1.26. Let X, Y be normed spaces. We denote the class of bounded linear functions $X \rightarrow Y$ by $\mathcal{L}(X, Y)$.

Theorem 1.1.27. If Y is a Banach space, so is $\mathcal{L}(X, Y)$.

Definition 1.1.28 (The Dual Space). The *dual space* of a normed space X is $X^* := \mathcal{L}(X, \mathbb{F})$.

Remark 1.1.29. If X is a Banach space, so is X^* .

Example 1.1.30. Let H be a Hilbert space. By the Riesz representation theorem we know that every $f \in H^*$ is of the form

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

for some $y \in H$.

The maps $y \mapsto f_y$ is a bijection, it holds that $\|f_y\| = \|y\|$, and

$$f_{\alpha y}(x) = \langle x, \alpha y \rangle = \bar{\alpha} \langle x, y \rangle = \bar{\alpha} f_y(x).$$

Example 1.1.31. Let $1 \leq p < \infty$. It holds that $(\ell_p)^* = \ell_q$ where $\frac{1}{p} + \frac{1}{q} = 1$.

For $c := (c_n)_{n \in \mathbb{N}_+}$ let

$$f_c(a_n) = \sum_{n \in \mathbb{N}_+} a_n c_n.$$

Then $c \mapsto f_c$ is an isometry and $\ell_q \cong (\ell_p)^*$.

Example 1.1.32. The same construction gives an embedding $\ell_1 \hookrightarrow (\ell_\infty)^*$, which is linear, norm-preserving and 1-1, but *not* surjective.

For $p < \infty$ it holds that

$$f\left((a_n)_{n \in \mathbb{N}_+}\right) = f\left(\sum_{n \in \mathbb{N}_+} a_n e_n\right) = \sum_{n \in \mathbb{N}_+} a_n f(e_n)$$

where e_n is the n^{th} basis vector. But, in ℓ^∞ it doesn't hold that $(a_n)_{n \in \mathbb{N}_+} \neq \sum_{n \in \mathbb{N}_+} a_n e_n$. For example,

$$\left\| (1, 1, 1, \dots) - \sum_{n \in [N]} e_n \right\|_\infty = \|(0, 0, \dots, 0, 1, 1, 1, \dots)\| = 1 \not\rightarrow 0.$$

But, we do have $c_0^* \cong \ell_1$ with the same proof as for ℓ_p .

Example 1.1.33. Let X be a compact metric space. We want to understand $\mathcal{C}(X)^*$. If μ is a finite Borel measure on X , we can define

$$\begin{aligned}\Phi_\mu : \mathcal{C}(X) &\rightarrow \mathbb{C} \\ f &\mapsto \int_X f \, d\mu.\end{aligned}$$

This is a linear functional. It holds that

$$|\Phi_\mu(f)| = \left| \int_X f \, d\mu \right| \leq \int_X |f| \, d\mu \leq \|f\|_\infty \mu(X).$$

So, $\Phi_\mu \in \mathcal{C}(X)^*$ and $\|\Phi_\mu\|^* \leq \mu(X)$. Taking $f \equiv 1$ we see actually that $\|\Phi_\mu\|^* = \mu(X)$.

If μ, ν are two measure on X , let

$$\Phi(f) = \int_X f \, d\mu - i \int_X f \, d\nu \in \mathcal{C}(X)^* = \int_X f \, d(\mu - i\nu).$$

Then $\Phi \in \mathcal{C}(X)^*$.

Definition 1.1.34 (Complex Measure). Let (X, Σ) be a measurable space. A *complex measure* μ on (X, Σ) is a map $\mu : \Sigma \rightarrow \mathbb{C}$ such that if $(A_n)_{n \in \mathbb{N}_+} \subseteq \Sigma$ are pairwise disjoint it holds that

$$\mu \left(\bigsqcup_{n \in \mathbb{N}_+} A_n \right) = \sum_{n \in \mathbb{N}_+} \mu(A_n).$$

Fact 1.1.35. 1. If μ is a complex measure, $\Re\mu, \Im\mu$ are real signed measures.

2. For every signed measure on a space X , there is a decomposition $X = P \sqcup N$ such that $\mu(A) \geq 0$ for every $A \in \Sigma$ such that $A \subseteq P$, and $\mu(A) \leq 0$ for every $A \in \Sigma$ such that $A \subseteq N$.

This is called the *Hahn decomposition* of a signed measure.

3. If μ is a complex measure on a space X , we can write

$$\mu = \mu_1 - \mu_2 + i(\mu_3 - \mu_4)$$

where $(\mu_i)_{i \in [4]}$ are (real, non-negative) measures on X . Then

$$\int_X f \, d\mu = \int_X f \, d\mu_1 - \int_X f \, d\mu_2 + i \left(\int_X f \, d\mu_3 - \int_X f \, d\mu_4 \right).$$

Definition 1.1.36 (Norm of a Measure). Given a complex norm μ , we define

$$\|\mu\| := \sup \left\{ \sum_{i \in \mathbb{N}_+} |\mu(A_i)| \mid X = \bigsqcup_{i \in [n]} A_i \right\}.$$

Theorem 1.1.37. Every $\Phi \in \mathcal{C}(X)^*$ is of the form $\Phi(f) = \int_X f \, d\mu$ for some complex measure μ . Also $\|\Phi\| = \|\mu\|$.

1.1.1 Applications of Duality & Some Measure Theory

Theorem 1.1.38 (Radon Nikodym). *Let (X, Σ, μ) be a finite measure space. Let ν be another measure on (X, Σ) such that $\nu \ll \mu$. Then $\exists g \geq 0$ such that*

$$\nu(A) = \int_A g \, d\mu$$

for all $A \in \Sigma$.

Proof. Let $H := L_2(X, \mu + \nu)$ and let

$$\begin{aligned} \ell: H &\rightarrow \mathbb{R} \\ f &\mapsto \int_X f \, d\mu. \end{aligned}$$

Then

$$\begin{aligned} |\ell(f)| &= \left| \int_X f \, d\mu \right| \\ &\leq \left(\int_X f^2 \, d\mu \right)^{\frac{1}{2}} \cdot \left(\int_X 1^2 \, d\mu \right)^{\frac{1}{2}} \\ &\leq \mu(X)^{\frac{1}{2}} \cdot \left(\int_X f^2 \, d(\mu + \nu) \right)^{\frac{1}{2}} \\ &= \mu(X) \|f\|_{L_2(\mu + \nu)}. \end{aligned}$$

By Riesz, there's $h \in H$ such that

$$\int f \, d\mu = \ell(f) = \langle f, h \rangle = \int_X f h \, d(\mu + \nu).$$

Then

$$\int_X f(1-h) \, d\mu = \int f h \, d\nu.$$

One can check $0 \leq h \leq 1$. Define $g = \frac{1-h}{h}$, then

$$\int_X f g \, d\mu = \int_X \frac{f}{h} (1-h) \, d\mu = \int_X \frac{f}{h} \cdot h \, d\nu = \int_X f \, d\nu.$$

Take $f = \chi_A$, from which

$$\int_A g \, d\mu = \int_A 1 \, d\nu = \nu(A). \quad \blacksquare$$

Remark 1.1.39. The theorem is also true if ν is a complex measure, with $g: X \rightarrow \mathbb{C}$ in $L_1(\mu)$. We get this by writing $\nu = (\nu_1 - \nu_2) + i(\nu_3 - \nu_4)$.

Theorem 1.1.40. *Let (X, Σ, μ) a finite measure space. Fix $1 \leq p < \infty$. Then*

$$(L_p(X, \Sigma, \mu))^* = L_q(X, \Sigma, \mu)$$

where $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. We work over \mathbb{C} , and assume $p > 1$. For every $g \in L_q(X, \Sigma, \mu)$ define $\Phi_g: L_p \rightarrow \mathbb{C}$ by

$$\Phi_g(f) = \int_X fg \, d\mu.$$

Obviously, Φ_g is linear. By Hölder we know

$$|\Phi_g(f)| \leq \|f\|_p \|g\|_q,$$

so $\Phi_g \in L_p^*$ and $\|\Phi_g\| \leq \|g\|_q$. Notice also that the map $g \mapsto \Phi_g$ is linear.

Fix $\Phi \in L_p(\mu)^*$, we should prove $\Phi = \Phi_g$ for some $g \in L_q$. Define a new measure ν on X by

$$\nu(A) = \Phi(\chi_A).$$

Here, $\chi_A \in L_p(\mu)$ since μ is a finite measure. We prove ν is indeed a (complex) measure. Fix $(A_n)_{n=1}^\infty$ such that $A = \sqcup_{n=1}^\infty A_n$. Note that $\sum_{n=1}^\infty \chi_{A_n} = \chi_A$ in L_p . Indeed,

$$\begin{aligned} \left\| \chi_A - \sum_{n=1}^N \chi_{A_n} \right\|_p &= \left\| \chi_{\sqcup_{n=N+1}^\infty A_n} \right\|_p \\ &= \mu \left(\bigsqcup_{n=N+1}^\infty A_n \right)^{\frac{1}{p}} \\ &= \left(\sum_{n=N+1}^\infty \mu(A_n) \right)^{\frac{1}{p}} \\ &\xrightarrow{N \rightarrow \infty} 0. \end{aligned}$$

Now, ν is indeed a measure since

$$\begin{aligned} \nu(A) &= \Phi(\chi_A) \\ &= \Phi \left(\sum_{n=1}^\infty \chi_{A_n} \right) \\ &= \sum_{n=1}^\infty \Phi(\chi_{A_n}) \\ &= \sum_{n=1}^\infty \nu(A_n). \end{aligned}$$

Now, if $\mu(A) = 0$ then $\chi_A = 0$ in $L_p(\mu)$. So

$$\nu(A) < \Phi(\chi_A) = \Phi(0) = 0,$$

so by Radon-Nikodym there exists $g \in L_1(\mu)$ such that

$$\Phi(\chi_A) = \nu(A) = \int_A g \, d\mu = \int \chi_A g \, d\mu = \Phi_g(\chi_A).$$

We got Φ_g such that $\Phi = \Phi_g$ on indicators. If $f = \sum_{i=1}^n \alpha_i \chi_{A_i}$ is a simple function, we have

$$\begin{aligned} \Phi(f) &= \sum_{i=1}^n \alpha_i \Phi(\chi_{A_i}) \\ &= \sum_{i=1}^n \alpha_i \Phi_g(\chi_{A_i}) \\ &= \Phi_g(f). \end{aligned}$$

Assume now that $f \in L_p$ is bounded with $|f| \leq M$. Choose simple functions $(f_n)_{n \in \mathbb{N}}$ such that $f_n \rightarrow f$ almost-everywhere and that $|f_n| \leq M$. Now

$$\begin{aligned} \Phi(f) &\stackrel{\text{DCT}}{=} \lim_{n \rightarrow \infty} \Phi(f_n) \\ &= \lim_{n \rightarrow \infty} \int f_n g \, d\mu \\ &\stackrel{\text{DCT}}{=} \int \lim_{n \rightarrow \infty} f_n g \, d\mu \\ &= \int f g \, d\mu \\ &= \Phi_g(f). \end{aligned}$$

Define

$$A_n := \{x \mid |g(x)| \leq n\}$$

and set

$$\rho(z) := \begin{cases} \frac{\bar{z}}{|z|} & z \neq 0 \\ 0 & z = 0 \end{cases}$$

so that $z\rho(z) = |z|$. Define $f_n := \chi_{A_n} \rho(g) |g|^{q-1}$ which is bounded.

For all $n \in \mathbb{N}$ we have

$$\begin{aligned}
 \|\Phi\| \left(\int_{A_n} |g|^q d\mu \right)^{\frac{1}{p}} &= \|\Phi\| \left(\int_{A_n} |g|^{p(q-1)} d\mu \right)^{\frac{1}{p}} \\
 &= \|\Phi\| \|f_n\|_p \\
 &\geq \Phi(f_n) \\
 &= \Phi_g(f_n) \\
 &= \int g f_n d\mu \\
 &= \int_{A_n} g \rho(g) |g|^{q-1} d\mu \\
 &= \int_{A_n} |g|^q d\mu.
 \end{aligned}$$

Hence

$$\left(\int_{A_n} |g|^q d\mu \right)^{\frac{1}{q}} \leq \|\Phi\|.$$

Then by MCT we have

$$\|g\|_q \leq \|\Phi\| < \infty.$$

Hence $g \in L_q(\mu)$ so $\|\Phi_g\| \leq \|g\|_q < \infty$ so Φ_g, Φ are continuous. Since they're equal on simple functions, they're equal. Also $\|\Phi_g\| = \|\Phi\| \geq \|g\|_q$ where we've just shown the inequality. Hence $\|\Phi_g\| = \|g\|_q$. \blacksquare

Exercise 2. Extend the above theorem to σ -finite spaces.

Chapter 2

Linear Functionals

2.1 Bounded Linear Functionals

2.1.1 The Hahn-Banach Theorem

We said that $(\ell_p)^* \cong \ell_q$ for $1 \leq p < \infty$, but *not* for ℓ_∞ . To show this, we need to find some linear functional $f \in (\ell_\infty)^*$ not of the form $f(a_n) = \sum_{n=1}^\infty a_n b_n$.

Definition 2.1.1 (Sublinear Functions). Let $X \in \mathbf{Vect}_\mathbb{R}$. A function $p: X \rightarrow \mathbb{R}$ is called *sublinear* if the following holds for all $x, y \in X$ and $\lambda \in \mathbb{R}_{\geq 0}$.

1. $p(x + y) \leq p(x) + p(y)$.
2. $p(\lambda x) = \lambda p(x)$.

Theorem 2.1.2 (Hahn-Banach, #1). Let $X \in \mathbf{Vect}_\mathbb{R}$ and $p: X \rightarrow \mathbb{R}$ sublinear. Assume $Y \subseteq X$ is a subspace, $\ell: Y \rightarrow \mathbb{R}$ is linear and $\ell \leq p$ on Y . Then $\exists \tilde{\ell}: X \rightarrow \mathbb{R}$ linear such that $\tilde{\ell} \leq p$ and $\tilde{\ell}|_Y = \ell$.

Lemma 2.1.3. Theorem 2.1.2 holds if $\text{codim } Y = 1$.

Proof. Every $x \in X$ is of the form $x = y + \lambda x_0$. We want to define

$$\tilde{\ell}(x) = \tilde{\ell}(y) + \lambda \tilde{\ell}(x_0) = \ell(y) + \lambda \tilde{\ell}(x_0).$$

We only get to choose $a = \tilde{\ell}(x_0)$. The goal is to choose a to have

$$\forall y \in Y \forall \lambda \in \mathbb{R}: \ell(y) + \lambda a = \tilde{\ell}(y + \lambda x_0) \leq p(y + \lambda x_0).$$

For $\lambda > 0$ we get the requirement

$$a \leq \frac{1}{\lambda} (p(y + \lambda x_0) - \ell(y)) = p\left(\frac{y}{\lambda} + x_0\right) - \ell\left(\frac{y}{\lambda}\right).$$

For $\lambda < 0$, write $\lambda = -\mu$ with $\mu > 0$. Denote also $z := y$. We get the requirement

$$\ell(z) - \mu a \leq p(z - \mu x_0)$$

so

$$a \geq \frac{1}{\mu} (\ell(z) - p(z - \mu x_0)) = \ell\left(\frac{z}{\mu}\right) - p\left(\frac{z}{\mu} - x_0\right).$$

Choosing such a is possible iff

$$\sup_{\substack{Z \in Y \\ \mu > 0}} \left[\ell\left(\frac{z}{\mu}\right) - p\left(\frac{z}{\mu} - x_0\right) \right] \leq \inf_{\substack{y \in Y \\ \lambda \geq 0}} \left[p\left(\frac{y}{\lambda} + x_0\right) - \ell\left(\frac{y}{\lambda}\right) \right].$$

We prove this is the case.

Let $y, z \in Y$ and $\lambda, \mu \in \mathbb{R}_+$. We indeed have

$$\begin{aligned} \ell\left(\frac{z}{\mu}\right) + \ell\left(\frac{y}{\lambda}\right) &= \ell\left(\frac{z}{\mu} + \frac{y}{\lambda}\right) \\ &\leq p\left(\frac{z}{\mu} - x_0 + \frac{y}{\lambda} + x_0\right) \\ &\leq p\left(\frac{z}{\mu} - x_0\right) + p\left(\frac{y}{\lambda} + x_0\right). \end{aligned} \quad \blacksquare$$

Proof (2.1.2). Let

$$\mathcal{F} := \left\{ (Z, f) \left| \begin{array}{l} Y \subseteq Z \subseteq X \\ f: Z \rightarrow \mathbb{R} \\ f|_Y = \ell \\ f \leq p \end{array} \right. \right\}.$$

We have $(Y, \ell) \in S$ so $S \neq \emptyset$. Say $(Z_1, f_1) \leq (Z_2, f_2)$ if $Z_1 \subseteq Z_2$ and $f_2|_{Z_1} = f_1$.

Take a chain $\{(Z_i, f_i)\}_{i \in I}$ in \mathcal{F} . We define $Z_\infty = \bigcup_{i \in I} Z_i$ and $f_\infty(x) = f_i(x)$ if $x \in Z_i$. Check that $(Z_\infty, f_\infty) \in S$ and this is an upper bound to the chain. By Zorn's lemma, there's a maximal element (Z_0, f_0) of \mathcal{F} . By 2.1.3 we get $Z_0 = X$, hence we're done. \blacksquare

Theorem 2.1.4 (Hahn Banach #2). *Let $(X, \|\cdot\|)$ be a normed space over \mathbb{F} , let $Y \subseteq X$ be a subspace and let $f \in Y^*$. Then $\exists \tilde{f} \in X^*$ such that $\tilde{f}|_Y = f$ and $\|\tilde{f}\|_{X^*} = \|f\|_{Y^*}$.*

Proof. • Assume first $\mathbb{F} = \mathbb{R}$. Just take $\ell = f$ and $p(x) = \|f\|_{Y^*} \|x\|$. Then for all $y \in Y$ we have

$$f(y) = \ell(y) \leq p(y) = \|f\|_{Y^*} \|y\|.$$

Hence by 2.1.2 there's \tilde{f} such that $\tilde{f}|_Y = f$ and

$$\forall x \in X: \tilde{f}(x) \leq \|f\|_{Y^*} \|x\|.$$

Taking $-x$ instead of x we get

$$-\tilde{f}(x) \leq \|f\|_{Y^*} \|x\|.$$

Hence $\|\tilde{f}\|_{X^*} \leq \|f\|_{Y^*}$.

- Assume now that $\mathbb{F} = \mathbb{C}$. We can think of $(X, \|\cdot\|)$ also as a space over \mathbb{R} . Write

$$f(y) = g(y) + ih(y)$$

where $g, h: Y \rightarrow \mathbb{R}$ are \mathbb{R} -linear. We have

$$g(iy) + ih(iy) = f(iy) = if(y) = -h(y) + ig(y)$$

so

$$h(y) = g(iy).$$

Hence

$$f(y) = g(y) - ig(iy).$$

Obviously,

$$|g(y)| \leq |f(y)|,$$

so

$$\|g\|_{Y^*} \leq \|f\|_{Y^*}.$$

By the real case, there's $\tilde{g}: X \rightarrow \mathbb{R}$ which is \mathbb{R} -linear such that $\tilde{g}|_Y = g$ and

$$\|\tilde{g}\|_{X^*} \leq \|g\|_{Y^*} \leq \|f\|_{Y^*}.$$

Define

$$\tilde{f}(x) = \tilde{g}(x) - i\tilde{g}(ix).$$

We get that $\tilde{f}(x)$ is \mathbb{R} -linear, but also

$$i\tilde{f}(x) = i\tilde{g}(x) + \tilde{g}(ix) = \tilde{f}(ix)$$

so \tilde{f} is \mathbb{C} -linear. Also $\tilde{f}|_Y = f$, so we're left to check the norm of \tilde{f} . Fix $x \in X$ and write $\tilde{f}(x) = re^{i\theta}$. Now

$$\begin{aligned} |\tilde{f}(x)| &= r \\ &= e^{-i\theta} \tilde{f}(x) \\ &= \tilde{f}(e^{-i\theta}x) \\ &= \tilde{g}(e^{-i\theta}x) \\ &\leq \|\tilde{g}\| \|e^{-i\theta}x\| \\ &= \|\tilde{g}\| \|x\|. \end{aligned}$$

So,

$$\|\tilde{f}\|_{X^*} \leq \|\tilde{g}\|_{X^*} \leq \|f\|_{Y^*}.$$

■

Corollary 2.1.5. *For every $x_0 \in X \setminus \{0\}$ there's $f \in X^*$ such that $\|f\| = 1$ and $f(x_0) = \|x_0\|$.*

Proof. Let $Y = \text{Span}\{x_0\}$ and

$$\begin{aligned} f: Y &\rightarrow \mathbb{F} \\ \lambda x_0 &\mapsto \lambda \|x_0\|. \end{aligned}$$

Then $\|f\|_{Y^*} = 1$ and $f(x_0) = \|x_0\|$. Now extend f to all of X . ■

Corollary 2.1.6. *If $f(x_1) = f(x_2)$ for all $f \in X^*$, then $x_1 = x_2$.*

Proof. If $x_1 \neq x_2$, by 2.1.5 there's $f \in X^*$ with $\|f\| = 1$ such that $f(x_1 - x_2) = \|x_1 - x_2\| \neq 0$. Then $f(x_1) \neq f(x_2)$. ■

Corollary 2.1.7. *Let $E \leq X$ and let x_0 such that*

$$d := d(x_0, E) > 0.$$

There's $f \in X^$ with $\|f\| = 1$ such that $f(E) = 0$ and $f(x_0) = d$.*

Proof. Take $Y = \text{Span}\{E, x_0\}$ and let

$$\begin{aligned} f: Y &\rightarrow \mathbb{F} \\ e + \lambda x_0 &\mapsto \lambda \cdot d. \end{aligned}$$

Then

$$\begin{aligned} \|e + \lambda x_0\| &= |\lambda| \left\| x_0 - \left(-\frac{1}{\lambda}\right) \right\| \\ &\geq |\lambda| d \\ &= |f(e + \lambda x_0)|. \end{aligned}$$

So $\|f\| \leq 1$, and one can check that actually $\|f\| = 1$.

By 2.1.4 we can extend f to X , which finishes the proof. ■

Corollary 2.1.8. *Let $A \subseteq X$ and $x \in X$. The following are equivalent.*

1. $x \in \overline{\text{Span}(A)}$.
2. For every $f \in X^*$ such that $f(A) = 0$ it holds that $f(x) = 0$.

Proof. 1 \implies 2 is obvious.

For the other direction, let $E := \overline{\text{Span}(A)}$. If $x \notin E$ then $d(x, E) = d > 0$. From 2.1.7 there's $f \in X^*$ with $\|f\| = 1$ such that $f(E) = 0$ and $f(x) = d \neq 0$. ■

2.1.2 Application of Hahn-Banach - Banach Limits

Definition 2.1.9 (Banach Limit). A *Banach limit* is a linear functional $f: \ell_\infty \rightarrow \mathbb{F}$ such that the following holds.

1. If $(a_n)_{n \in \mathbb{N}_+}$ is such that $a_n \geq 0$ for all $n \in \mathbb{N}_+$, then $f((a_n)_{n \in \mathbb{N}_+}) \geq 0$.

2. $f\left((a_{n+1})_{n \in \mathbb{N}_+}\right) = f\left((a_n)_{n \in \mathbb{N}_+}\right)$.
3. If $a_n \xrightarrow{n \rightarrow \infty} L$ then $f\left((a_n)_{n \in \mathbb{N}_+}\right) = L$.

Remark 2.1.10. It holds that

$$f\left((a_n)_{n \in \mathbb{N}_+}\right) \leq f\left(\left(\left\|(a_n)_{n \in \mathbb{N}_+}\right\|_\infty, \left\|(a_n)_{n \in \mathbb{N}_+}\right\|_\infty, \left\|(a_n)_{n \in \mathbb{N}_+}\right\|_\infty, \dots\right)\right) = \left\|(a_n)_{n \in \mathbb{N}_+}\right\|_\infty.$$

Similarly

$$-f\left((a_n)_{n \in \mathbb{N}_+}\right) \leq \left\|(a_n)_{n \in \mathbb{N}_+}\right\|_\infty.$$

From this it follows that $\|f\| \leq 1$, and by looking at a constant sequence we actually get $\|f\| = 1$.

Remark 2.1.11. A Banach limit f is in $(\ell_\infty)^*$ and this is *not* of the form

$$f\left((a_n)_{n \in \mathbb{N}_+}\right) = \sum_{n \in \mathbb{N}_+} a_n b_n.$$

We show the latter part. Assume $f\left((a_n)_{n \in \mathbb{N}_+}\right) = \sum_{n=1}^\infty a_n b_n$. Then

$$\begin{aligned} 0 &= f(e_k) \\ &= \sum_{n \in \mathbb{N}_+} (e_k)_n b_n \\ &= \sum_{n \in \mathbb{N}_+} \delta_{k,n} b_n \\ &= b_k \end{aligned}$$

so $f = 0$, which is a contradiction to $f(1, 1, 1, \dots) = 1$.

Theorem 2.1.12. *There exists a Banach limit.*

Remark 2.1.13. Without parts (1,2) of the definition, this is obvious from 2.1.4. One can extend $f\left((a_n)_{n \in \mathbb{N}_+}\right) = \lim_{n \rightarrow \infty} a_n$ from c to ℓ_∞ .

Proof. Let

$$E := \left\{ (a_{n+1} - a_n)_{n \in \mathbb{N}_+} \mid (a_n)_{n \in \mathbb{N}_+} \in \ell_\infty \right\}$$

and $x_0 = (1, 1, 1, \dots)$ the constant sequence 1. We claim $d(x_0, E) = 1$. Let $b := (b_n)_{n \in \mathbb{N}_+} \in E$ and let $d := \|x_0 - b\|$. Then

$$\begin{aligned} nd &\geq \sum_{k \in [n]} |(x_0)_k - b_k| \\ &\geq \sum_{k \in [n]} (1 - b_k) \\ &= n - \sum_{k \in [n]} b_k. \end{aligned}$$

We claim the last sum is bounded. Indeed, if $b_k = a_{k+1} - a_k$ then

$$\sum_{k \in [n]} b_k = a_{n+1} - a_1$$

which is bounded as $(a_n)_{n \in \mathbb{N}_+} \in \ell_\infty$. Hence, dividing by n we get

$$d \geq 1 - \frac{1}{n} \sum_{k \in [n]} b_k \xrightarrow{n \rightarrow \infty} 1.$$

Hence $d(x_0, E) \geq 1$, but obviously $d(x_0, E) \leq \|x_0\| = 1$.

By 2.1.7 there's $f \in (\ell_\infty)^*$ such that $f(E) = 0$, $f(x_0) = 1$ and $\|f\| = 1$. Hence we have property 2 of the Banach limit. To get property 1 fix a sequence $a = (a_n)_{n \in \mathbb{N}_+}$ such that $a_n \geq 0$ and assume without loss of generality that $\|a\|_\infty = 1$. Then

$$\begin{aligned} f(a) &= f(x_0 - (x_0 - a)) \\ &= f(x_0) - f(x_0 - a) \geq 1 - \|x_0 - a\| \\ &\geq 1 - 1 \\ &= 0 \end{aligned}$$

where in the last inequality we use $a_n \in [0, 1]$ for all $n \in \mathbb{N}_+$. This proves property 1.

Note that

$$\forall a \in \ell_\infty \forall k \in \mathbb{N} f((a_n)_{n \in \mathbb{N}_+}) = f((a_{n+k})_{n \in \mathbb{N}_+}) \leq \sup_{n \geq k} a_n.$$

Hence

$$f((a_n)_{n \in \mathbb{N}_+}) \leq \inf_{k \in \mathbb{N}} \sup_{n \geq k} a_n = \limsup_{n \rightarrow \infty} a_n.$$

Taking $-a$ instead of a we get

$$-f(a) \leq \limsup_{n \rightarrow \infty} (-a) = -\liminf_{n \rightarrow \infty} a_n$$

so

$$f(a) \geq \liminf_{n \rightarrow \infty} a_n.$$

Hence if $a_n \rightarrow L$ also $f(a) = L$. ■

2.1.3 Application of Hahn-Banach - Reflexive Spaces

Definition 2.1.14. For every $x \in X$ define $\text{ev}_x \in X^{**} = (X^*)^*$ by

$$\forall f \in X^*: \text{ev}_x(f) = f(x).$$

Proposition 2.1.15. The map $\iota: x \mapsto \text{ev}_x$ is a linear norm-preserving map $X \rightarrow X^{**}$.

Proof. Linearity is easy. Note that

$$|\mathrm{ev}_x(f)| = |f(x)| \leq \|f\| \|x\| = \|x\| \|f\|,$$

so by definition $\|\mathrm{ev}_x\| \leq \|x\|$. By Hahn-Banach there's f with $\|f\| = 1$ and $f(x) = \|x\|$. Then

$$\|x\| = |f(x)| = |\mathrm{ev}_x(f)| \leq \|\mathrm{ev}_x\| \|f\| = \|\mathrm{ev}_x\|.$$

Hence $x \mapsto \mathrm{ev}_x$ is norm-preserving. ■

Definition 2.1.16. A normed space X is called *reflexive* if $\iota: X \rightarrow X^{**}$ is surjective.

Example 2.1.17. If $p \in (1, \infty)$, the spaces ℓ_p, L_p are reflexive since for $\frac{1}{p} + \frac{1}{q} = 1$ we know $(L_p)^* \cong L_q$ and $(L_q)^* \cong L_p$, and similarly for ℓ_p .

However, ℓ_1 is not reflexive since $(\ell_1)^* \cong \ell_\infty$ and $(\ell_\infty)^* \not\cong \ell_1$. Similarly, c_0 is not reflexive since $c_0^* \cong \ell_1$ but $\ell_1^* \cong \ell_\infty \not\cong c_0$.

2.1.4 Application of Hahn-Banach - Approximation Theory

Example 2.1.18. The space $\mathrm{Span}\{x^k \mid k \in \mathbb{Z}_{\geq 0}\}$ is dense in $\mathcal{C}([0, 1])$ by the Weierstrass approximation theorem.

Example 2.1.19. Consider

$$\mathcal{C}(\mathbb{T}) := \{f: \mathbb{R} \rightarrow \mathbb{C} \mid f \text{ is continuous and } 2\pi\text{-periodic}\}.$$

Then $\mathrm{Span}\{e^{2\pi i n} \mid n \in \mathbb{Z}_{\geq 0}\}$ is dense in $\mathcal{C}(\mathbb{T})$ by the second Weierstrass approximation theorem / Fejer's theorem.

Theorem 2.1.20. Fix a weight function $w: \mathbb{R} \rightarrow \mathbb{R}$ such that $0 < w(t) < Ae^{-B|t|}$. Then

$$E := \mathrm{Span}\{t^n w \mid n \in \mathbb{Z}_{\geq 0}\}$$

is dense in $c_0(\mathbb{R})$.

Proof. Step 1: We show that $c^{i\alpha t} w \in \bar{E}$ for every $\alpha \in \mathbb{R}$. Let $f \in c_0(\mathbb{R})^*$ such that $f(E) = 0$. For $z \in \mathbb{C}$ write $\rho_z(t) := e^{izt} w$. It holds that

$$\begin{aligned} |\rho_z(t)| &= |e^{izt}| w \\ &= e^{\Re(izt)} w \\ &\leq e^{-(\Im z)t} A^{-B|t|} \\ &\leq Ae^{(|\Im z| - B)|t|}. \end{aligned}$$

If $z \in S := \{z \mid |\Im z| < B\}$ then $\rho_z \in c_0(\mathbb{R})$. Define $\varphi: S \rightarrow \mathbb{C}$ by $\varphi(t) = f(\rho_t)$.

Note that

$$\begin{aligned} \left(\frac{\rho_{z+h} - \rho_z}{h} \right) (t) &= \frac{e^{i(z+h)t} w(t) - e^{izt} w(t)}{h} \\ &= w(t) e^{izt} \frac{e^{iht} - 1}{h} \\ &= w(t) e^{izt} it \cdot \frac{e^{iht} - 1}{ith} \end{aligned}$$

where the last expression converges uniformly to e^{izt} as $h \rightarrow 0$ since $w(t) e^{izt} it$ is bounded and since $\frac{e^{iht} - 1}{ith}$ converges uniformly to 1.

Hence

$$\begin{aligned} \varphi'(z) &= \lim_{h \rightarrow 0} \frac{\varphi(z+h) - \varphi(z)}{h} \\ &= \lim_{h \rightarrow 0} f \left(\frac{\rho_{z+h} - \rho_z}{h} \right) \\ &= f \left(\lim_{h \rightarrow 0} \frac{\rho_{z+h} - \rho_h}{h} \right) \\ &= f(it e^{izt} w(t)). \end{aligned}$$

In particular, φ is analytic and $\varphi'(0) = f(f(itw(t))) = 0$ where the latter equality is by the choice of f .

Similarly

$$\varphi^{(n)}(t) = f((it)^n e^{izt} w(t))$$

so

$$\varphi^{(n)}(0) = f(i^n t^n w(t)) = 0.$$

Hence by analyticity and complex analysis, $\varphi \equiv 0$. Hence $f(\rho_t) = 0$ for every $z \in S$ and every $f \in c_0(\mathbb{R})^*$ for which $f(E) = 0$. Hence $\rho_z \in \bar{E}$ for every $z \in S$.

Step 2: We'll show that $\mathcal{C}_c(\mathbb{R}) \subseteq \bar{E}$ which finishes the proof because $\overline{\mathcal{C}_c(\mathbb{R})} = c_0(\mathbb{R})$.

Let $u \in \mathcal{C}_c(\mathbb{R})$ and let $(-M, M)$ contain $\text{supp}(u)$. Define

$$u(t) := \frac{u\left(\frac{M}{\pi}t\right)}{w\left(\frac{M}{\pi}t\right)}.$$

This is supported in $(-\pi, \pi)$. We can then extend v to $\mathcal{C}(\mathbb{T})$. By Weirstrass' approximation there are trigonometric polynomials $(p_n)_{n \in \mathbb{N}}$ such that $p_n \xrightarrow{n \rightarrow \infty} v$ uniformly. Then

$$p_n\left(\frac{\pi}{M}t\right) w(t) \xrightarrow{n \rightarrow \infty} v\left(\frac{\pi}{M}t\right) w(t) = u(t)$$

uniformly since $w(t)$ is bounded. By the first step, we have $p_n\left(\frac{\pi}{M}t\right) w(t) \in \bar{E}$ hence also $u \in \bar{E}$, as required. ■

2.1.5 Application of Hahn-Banach - Convex Separation

Definition 2.1.21 (The Minkowski Functional). Let $X \in \mathbf{Vect}_{\mathbb{R}}$ and let $K \subseteq X$ be convex such that $0 \in K$. Define the *Minkowski function* on X

$$p_K: X \rightarrow [0, \infty]$$

$$x \mapsto \inf \left\{ \lambda > 0 \mid \frac{x}{\lambda} \in K \right\}.$$

Definition 2.1.22 (Internal Point). Let $X \in \mathbf{Vect}_{\mathbb{R}}$ and $A \subseteq X$. A point $a \in A$ is *internal* if

$$\forall y \in X \exists t_0 \in \mathbb{R} \forall t \leq t_0: a + ty \in A.$$

Remark 2.1.23. If X is normed and $A \subseteq X$, every interior point of A is internal but the converse is generally false. E.g. take the union of two tangent circles and the line tangent to both of them. The tangential point is an internal point but not an interior point.

Fact 2.1.24. Let $X \in \mathbf{Vect}_{\mathbb{R}}$ and let $K \subseteq X$ be convex such that $0 \in K$ is internal. Then $p_K: X \rightarrow [0, \infty)$ is sub-linear.

Proof. One checks that $p_K(\lambda x) = \lambda p_K(x)$ for $\lambda > 0$.

We check the triangle inequality. Let $x, y \in X$. There are $(\lambda_n)_{n \in \mathbb{N}}$ and $(\mu_n)_{n \in \mathbb{N}}$ such that $\frac{x}{\lambda_n}, \frac{y}{\mu_n} \in K$ for all $n \in \mathbb{N}$ and that

$$\lambda_n \xrightarrow{n \rightarrow \infty} p_K(x)$$

$$\mu_n \xrightarrow{n \rightarrow \infty} p_K(y).$$

We have

$$\frac{x+y}{\lambda_n + \mu_n} = \frac{\lambda_n}{\lambda_n + \mu_n} \frac{x}{\lambda_n} + \frac{\mu_n}{\lambda_n + \mu_n} \frac{y}{\mu_n} \in K$$

by convexity of K . So $p_K(x+y)$. So

$$p_K(x+y) \leq \lambda_n + \mu_n \xrightarrow{n \rightarrow \infty} p_K(x) + p_K(y). \quad \blacksquare$$

Remark 2.1.25. We cannot in general reconstruct K from p_K . We know

$$p_K^{-1}([0, 1)) \subseteq K \subseteq p_K^{-1}([0, 1]),$$

but don't know if a point $p \in X$ for which $p_K(p) = 1$ are in K or not.

Definition 2.1.26 (Separating Function). Let X be a real Banach space. Let $A, B \subseteq X$ be disjoint. A linear functional $f: X \rightarrow \mathbb{R}$ separates A and B if $f \neq 0$ and

$$\sup_{a \in A} f(a) \leq \inf_{b \in B} f(b).$$

Theorem 2.1.27. *Let X be a real vector space and $A, B \subseteq X$ be convex and disjoint such that A has an internal point. Then there exists $f: X \rightarrow \mathbb{R}$ which separates A, B .*

Proof. Step 1: Assume first that $B = \{x_0\}$ and that $0 \in A$ is internal. Let $Y = \text{Span}\{x_0\}$ and define

$$\begin{aligned}\ell: Y &\rightarrow \mathbb{R} \\ \lambda x_0 &\mapsto \lambda p_A(x_0).\end{aligned}$$

Then

$$\begin{aligned}\forall \lambda > 0: \ell(\lambda x_0) &= \lambda p_A(x_0) = p_A(\lambda x_0) \\ \forall \lambda \leq 0: \ell(\lambda x_0) &= \lambda p_A(x_0) \leq 0 \leq p_A(\lambda x_0)\end{aligned}$$

so $\ell \leq p_A$.

We can extend ℓ to a function $f: X \rightarrow \mathbb{R}$ such that $f|_Y = \ell$ and $f \leq p_A$. But

$$\sup_{a \in A} f(a) \leq \sup_{a \in A} p_A(a) \leq 1 \leq p_A(x_0) = f(x_0).$$

Step 2: In general, let $a_0 \in A$ be internal and let $b_0 \in B$. Let

$$C := A - B + b_0 - a_0 = \{a - b + b_0 - a_0 \mid a \in A, b \in B\}.$$

This is convex by writing down the definition, and $0 = a_0 - b_0 + b_0 - a_0 \in C$ is an internal point because a_0 is internal in A .

But, $b_0 - a_0 \notin C$ for otherwise $a = b$ which implies $A \cap B \neq \emptyset$. Hence, there's $f: X \rightarrow \mathbb{R}$ such that

$$\sup_{\substack{a \in A \\ b \in B}} (f(a) - f(b) + f(b_0) - f(a_0)) = \sup_{c \in C} f(c) \leq f(b_0 - a_0) = f(b_0) - f(a_0).$$

So

$$\sup_{\substack{a \in A \\ b \in B}} (f(a) - f(b)) \leq 0$$

which implies

$$\sup_{a \in A} f(a) \leq \inf_{b \in B} f(b).$$

■

Example 2.1.28. Let $X := \mathbb{R}[x]$ and let A be the subspace of polynomials with positive leading coefficient. This is convex. Take $B = \{0\}$, so that $A \cap B = \emptyset$.

Assume there's a separating $f: X \rightarrow \mathbb{R}$ for A, B . I.e.

$$\sup_{a \in A} f(a) \leq f(a) = 0.$$

Let $p \in X$ and $M > \deg p$. Then

$$\forall \delta > 0: \delta x^m + p \in A.$$

Then

$$\forall \delta > 0: \delta f(x^m) + f(p) = f(\delta x^m + p) \leq 0.$$

By sending $\delta \rightarrow 0$ we get $f(p) \leq 0$. By considering $f(-p)$ we get $f(p) \geq 0$ so $f \equiv 0$, a contradiction.

Note that A doesn't have an internal point, for otherwise the theorem would imply there's a separation for A, B .

We've so far discussed separation when X isn't normed, but there might be a norm on it.

Proposition 2.1.29. *Let X be a real Banach space and let $f: X \rightarrow \mathbb{R}$ be linear. Let $\lambda \in \mathbb{R}$ and $H := f^{-1}(\lambda)$. Then*

1. *If f is bounded, H is closed.*
2. *If f is not bounded, H is dense.*

Proof. 1. This is immediate from topology.

2. It holds that

$$\sup_{x \in X} \frac{|f(x)|}{\|x\|} = \infty.$$

Take $(x_n)_{n \in \mathbb{N}}$ such that $\frac{|f(x_n)|}{\|x_n\|} \xrightarrow{n \rightarrow \infty} \infty$ and $y_n := \frac{x_n}{f(x_n)}$, so $f(y_n) = 1$ and $\|y_n\| \rightarrow 0$. In particular $y_n \rightarrow 0$. Then

$$\forall x \in X: x = \lim_{n \rightarrow \infty} (x - f(x)y_n + \lambda y_n).$$

But,

$$f(x - f(x)y_n + \lambda y_n) = f(x) - f(x) \cdot 1 + \lambda \cdot 1 = \lambda$$

so

$$x - f(x)y_n + \lambda y_n \in H,$$

so x is the limit of elements of H , which implies $x \in \bar{H}$. ■

Corollary 2.1.30. *Let X be a real Banach space and let $A, B \subseteq X$ be convex disjoint sets such that A has an interior point y_0 . There's $f \in X^*$ separating A, B .*

Proof. We know there's some linear functional $f: X \rightarrow \mathbb{R}$, which we don't know is continuous. We have

$$\sup_{a \in B(y_0, r)} f(a) \leq \sup_{a \in A} f(a) \leq \lambda = \inf_{b \in B} f(b).$$

Let

$$H := f^{-1}(\lambda + 1).$$

Then $H \cap B(y_0, r) = \emptyset$ so H is not dense, and by the proposition this implies f is bounded, i.e. continuous. ■

2.2 Banach Spaces

2.2.1 Baire Category Theorem & Uniform Boundedness

Definition 2.2.1. Let X be a topological space and let $A \subseteq X$. We say A is *nowhere dense* if $\bar{A} = \emptyset$. We say A is *meagre* (alternatively, a set of the set category) if $A = \bigcup_{n \in \mathbb{N}} A_n$ where each A_n is nowhere dense. We say A is *comeagre* if A^C is meagre.

Theorem 2.2.2 (Baire). *A complete metric space is not meagre.*

Theorem 2.2.3 (Uniform Boundedness Principle). *Let X be a Banach space and Y a normed space. Let $\mathcal{F} \subseteq \mathcal{L}(X, Y)$. Assume*

$$\forall x \in X: \sup_{T \in \mathcal{F}} \|Tx\| < \infty.$$

Then

$$\sup_{T \in \mathcal{F}} \|T\| < \infty.$$

This follows immediately from the following theorem and Baire's category theorem.

Theorem 2.2.4. *Let X be a Banach space, Y a normed space and $\mathcal{F} \subseteq \mathcal{L}(X, Y)$. Assume $\{x \mid \sup_{T \in \mathcal{F}} \|Tx\| < \infty\}$ is not meagre. Then $\sup_{T \in \mathcal{F}} \|T\| < \infty$.*

Proof. Let

$$A_n := \left\{ x \mid \sup_{T \in \mathcal{F}} \|Tx\| \leq n \right\}.$$

Then A_n are closed because if $\varphi_T(x) = \|Tx\|$ we get

$$A_n = \bigcap_{T \in \mathcal{F}} \varphi_T^{-1}([0, n]).$$

Moreover, $\bigcup_{n \in \mathbb{N}_+} A_n = A$ so there's $n \in \mathbb{N}$ such that $(\overline{A_n})^\circ \neq \emptyset$. Then there's $y_0 \in X$ and $r > 0$ such that $B(y_0, r) \subseteq A_n$. If $|z| < r$ we can write

$$\|Tz\| = \left\| T\left(\frac{y_0 + z}{2}\right) - T\left(\frac{y_0 - z}{2}\right) \right\| \leq \frac{1}{2} (\|T(y_0 + z)\| + \|T(y_0 - z)\|) \leq \frac{1}{2} (n + n) = n,$$

where in the last inequality we used $y_0 \pm z \in A_n$. Now

$$\forall T \in \mathcal{F} \forall x \in X: \|Tx\| = \left\| T\left(\frac{rx}{2\|x\|}\right) \right\| \cdot \frac{2\|x\|}{r} \leq \frac{2n}{r} \|x\|.$$

Hence

$$\sup_{T \in \mathcal{F}} \|T\| \leq \frac{2n}{r} < \infty. \quad \blacksquare$$

Corollary 2.2.5. *Let X be a Banach space and $A \subseteq X$. Then A is bounded iff $f(A)$ is bounded for all $f \in X^*$.*

Proof. • Assume A is bounded and $f \in X^*$. We get

$$|f(a)| \leq \|f\| \|a\| \leq \left(\sup_{a \in A} \|a\| \right) \|f\|$$

so $f(A)$ is bounded.

- Assume $f(A)$ is bounded for all $f \in X^*$. Define

$$\mathcal{F} := \{\text{ev}_x \mid x \in A\} \subseteq X^{**} \subseteq \mathcal{L}(X^*, \mathbb{F}).$$

Then

$$\forall f \in X^*: \sup_{T \in \mathcal{F}} \|Tf\| = \sup_{x \in A} |\text{ev}_x(f)| = \sup_{x \in A} |f(x)| < \infty.$$

By the uniform boundedness theorem this implies

$$\sup_{T \in \mathcal{F}} \|T\| = \sup_{x \in A} \|\text{ev}_x\| = \sup_{x \in A} \|x\| < \infty. \quad \blacksquare$$

Proposition 2.2.6. *Let H be a Hilbert space and $T: H \rightarrow H$ be linear and self-adjoint. Then T is continuous.*

Proof. Let

$$A := \{Tx \mid \|x\| \leq 1\}.$$

We know that every $f \in H^*$ is of the form $f(x) = \langle x, y \rangle$. Hence

$$\begin{aligned} f(A) &= \{f(a) \mid a \in A\} \\ &= \{f(Tx) \mid \|x\| \leq 1\} \\ &= \{\langle Tx, y \rangle \mid \|x\| \leq 1\} \\ &= \{\langle x, Ty \rangle \mid \|x\| \leq 1\}. \end{aligned}$$

By Cauchy-Schwarz,

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

so $f(A)$ is bounded.

By the uniform boundedness theorem it follows that A is bounded, hence T is continuous. \blacksquare

2.2.2 Applications to Harmonic Analysis

Definition 2.2.7. Given $f \in \mathcal{C}(\mathbb{T})$ and define

$$\hat{f}(n) = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx.$$

The Fourier series of f is

$$\sum_{n \in \mathbb{Z}} \hat{f}(n) e^{inx}.$$

Theorem 2.2.8 (Dirichlet). Let $f \in \mathcal{C}_{\text{pr}}\mathbb{T}$, it holds that

$$\sum_{n \in \mathbb{Z}} \hat{f}(n) e^{inx} = f(x)$$

and the partial sums $S_N f = \sum_{n=-N}^N \hat{f}(n) e^{inx}$ converge uniformly to f .

Theorem 2.2.9. Let

$$A := \{f \in \mathcal{C}(\mathbb{T}) \mid \forall q \in \mathbb{Q}: (S_N f(q)) \text{ isn't bounded}\}.$$

A is comeagre in $\mathcal{C}(\mathbb{T})$.

Proof. For $q \in \mathbb{Q}$ define

$$A_q := \{f \in \mathcal{C}(\mathbb{T}) \mid (S_N f(q))_{N \in \mathbb{N}} \text{ isn't bounded}\}.$$

Then $A = \bigcap_{q \in \mathbb{Q}} A_q$, so it's enough to show that each A_q is comeagre. WLOG assume $q = 0$. Define

$$\begin{aligned} T_n: \mathcal{C}(\mathbb{T}) &\rightarrow \mathbb{C} \\ f &\mapsto S_N f(0). \end{aligned}$$

We claim $T_n \in \mathcal{C}(\mathbb{T})^*$ but $\sup_{n \in \mathbb{N}} \|T_n\| = \infty$. By uniform boundedness this would imply

$$A_0^C = \{f \mid T_n f \text{ is bounded}\}$$

is meagre.

Indeed,

$$\begin{aligned} T_n f &= \sum_{n=-N}^N \hat{f}(n) \\ &= \sum_{n=-N}^N \left(\frac{1}{2\pi} \int_0^{2\pi} f(x) e^{\pm inx} dx \right) \\ &= \frac{1}{2\pi} \int_0^{2\pi} f(x) \cdot \underbrace{\left(\sum_{n=-N}^N e^{inx} \right)}_{D_N(x)} dx \\ &= \frac{1}{2\pi} \int_0^{2\pi} f(x) D_N(x) dx \end{aligned}$$

from which

$$\begin{aligned} |T_n f| &\leq \frac{1}{2\pi} \int_0^{2\pi} |f(x)| |D_N(x)| dx \\ &\leq \|f\| \cdot \underbrace{\frac{1}{2\pi} \int_0^{2\pi} |D_n(x)| dx}_{I_n} \end{aligned}$$

so $\|T_n\| \leq I_n < \infty$.

Actually $\|T_n\| = I_n$ by picking $f = \operatorname{sgn} D_N$ and approximating it by continuous functions. But

$$I_N := \frac{1}{2\pi} \int_0^{2\pi} \frac{\sin \left[\left(N + \frac{1}{2} \right) x \right]}{\sin \left(\frac{x}{2} \right)} dx \xrightarrow{n \rightarrow \infty} \infty,$$

so $\sup_{n \in \mathbb{N}} \|T_n\| = \infty$ as required by the above reduction. ■

2.2.3 The Open Mapping Theorem

Definition 2.2.10 (Open Map). Let X, Y be topological spaces. A map $\varphi: X \rightarrow Y$ is called *open* if $\varphi(U)$ is open for every open $U \subseteq X$.

Theorem 2.2.11 (The Open Mapping Theorem). Let X, Y be Banach spaces and let $T \in \mathcal{L}(X, Y)$. If T is onto, it's open.

Proof. We prove that for any open ball $B(x, r)$ the set $T(B(x, r))$ contains a ball around Tx . Note that

$$T(B(x, r)) = T(rB(0, 1) + x) = r \cdot T(B(0, 1)) + Tx.$$

Hence it's enough to show that $T(B(0, 1))$ contains a ball around 0. Note that

$$Y = T(X) = \bigcup_{n \in \mathbb{N}_+} T(B_X(0, 1)).$$

By Baire's theorem

$$\exists n \in \mathbb{N}_+ : \operatorname{int} \overline{T(B(0, 1))} \neq \emptyset$$

or in other words

$$\overline{T(B(0, 1))} \supseteq B(y, r).$$

We have to fix three things. We want a ball around zero, we want the image of a ball of radius 1, and we want the actual image to contain it, rather than the closure of the image.

Ball Around Zero & Image of 1-Ball: Write

$$\begin{aligned} B(0, r) &= B(y, r) - y \\ &\subseteq \overline{T(B(0, n)) - Tx} \\ &= \overline{T(B(0, n) - x)} \\ &\subseteq \overline{T(B(0, n + \|x\|))}. \end{aligned}$$

Dividing by $n + \|x\|$ we get

$$\overline{T(B(0, 1))} \supseteq B\left(0, \frac{r}{n + \|x\|}\right) =: B(0, \varepsilon).$$

Closure: Note that for every $a > 0$ it holds that

$$\overline{T(B(0, a))} \supseteq B(0, a\varepsilon). \quad (2.1)$$

We show that $T(B(0, 1)) \supseteq B(0, \frac{\varepsilon}{2})$. Let $y \in B(0, \frac{\varepsilon}{2})$. By (2.1) with $a = \frac{1}{2}$ there's $x_1 \in B(\frac{1}{2})$ such that

$$\|y - Tx_1\| < \frac{\varepsilon}{4}.$$

By (2.1) with $\frac{1}{2^n}$ there are $x_n \in B(0, \frac{1}{2^n})$ such that

$$\left\| y - T \left(\sum_{i \in [n]} x_i \right) \right\| < \frac{\varepsilon}{2^{n+1}}. \quad (2.2)$$

Since

$$\sum_{n \in \mathbb{N}_+} \|x_n\| < \sum_{n \in \mathbb{N}_+} \frac{1}{2^n} = 1 < \infty$$

and X is complete, we get that $\sum_{n \in \mathbb{N}_+} x_n$ converges to sum x . We have

$$\|x\| \leq \sum_{n \in \mathbb{N}_+} \|x_n\| < 1$$

so $x \in B(0, 1)$.

Let $n \rightarrow \infty$ in (2.2), we get $\|y - Tx\| \leq 0$ so $y = Tx$. ■

Corollary 2.2.12. *Let X, Y be Banach spaces and $T \in \mathcal{L}(X, Y)$. If T is a bijection, T^{-1} is bounded and there's $c > 0$ such that $\|Tx\|_Y \geq c\|x\|_X$.*

Proof. We know by the open mapping theorem that $T(B(0, 1)) \supseteq T(0, \varepsilon)$ for some $\varepsilon > 0$, or in other words $B(0, 1) \supseteq T^{-1}(B(0, \varepsilon))$. Then

$$\|T^{-1}x\| = \left\| T^{-1} \left(\frac{\varepsilon}{2} \cdot \frac{x}{\|x\|} \right) \right\| \cdot \frac{2\|x\|}{\varepsilon} \leq \frac{2}{\varepsilon} \|x\|$$

where in the last inequality we use $\frac{\varepsilon}{2} \frac{x}{\|x\|} \in B(0, \varepsilon)$. Hence $\|T^{-1}\| \leq \frac{2}{\varepsilon}$. In fact one can check $\|T^{-1}\| \leq \frac{1}{\varepsilon}$.

For every $x \in X$ we now have

$$\|x\| = \|T^{-1}Tx\| \leq \|T^{-1}\| \|Tx\| \leq \frac{2}{\varepsilon} \|Tx\|.$$

Hence $\|Tx\| \geq \frac{\varepsilon}{2} \|x\|$. ■

Remark 2.2.13. The fact that T^{-1} is bounded can be shown more directly. Since T is surjective, it's open so T^{-1} is continuous and therefore bounded.

Corollary 2.2.14. *Let X be a complete Banach space with respect to two norms $\|\cdot\|_1$ and $\|\cdot\|_2$. If*

$$\exists C > 0 \forall x \in X: \|x\|_1 \leq C \|x\|_2$$

then

$$\exists \tilde{C} > 0 \forall x \|x\|_2 \leq \tilde{C} \|x\|_1,$$

so the norms are equivalent.

Proof. Apply 2.2.12 to $i: (X, \|\cdot\|_2) \rightarrow (X, \|\cdot\|_1)$. ■

2.2.4 Application of the Open Mapping Theorem to Harmonic Analysis

Definition 2.2.15 (Fourier Coefficients for Functions on the Circle). Write

$$L_1(\mathbb{T}) := \left\{ f: \mathbb{R} \rightarrow \mathbb{C} \mid \begin{array}{l} f \text{ is } 2\pi\text{-periodic} \\ \int_0^{2\pi} |f| dx < \infty \end{array} \right\} \cong L_1([0, 2\pi])$$

with the norm

$$\|f\| = \frac{1}{2\pi} \int_0^{2\pi} |f(x)| dx.$$

This is a Banach space. For every $f \in L_1(\mathbb{T})$ and $z \in \mathbb{Z}$ define

$$\hat{f}(n) = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx.$$

Fact 2.2.16. 1. *If $\hat{f}(n) = 0$ for every $n \in \mathbb{Z}$ then $f = 0$ in L_1 .*

2. *Riemann-Lebesgue: $\hat{f}(n) \xrightarrow{n \rightarrow \pm\infty} 0$.*

Given $(a_n)_{n \in \mathbb{Z}} \subseteq \mathbb{C}$ such that $a_n \xrightarrow{n \rightarrow \pm\infty} 0$ we want to ask if there's $f \in L_1(\mathbb{T})$ such that $\hat{f}(n) = a_n$ for every $n \in \mathbb{Z}$. It turns out that the answer is no, which we prove using the open mapping theorem.

Definition 2.2.17. Define

$$c_0(\mathbb{Z}) := \left\{ (a_n)_{n \in \mathbb{Z}} \mid a_n \xrightarrow{n \rightarrow \pm\infty} 0 \right\}$$

with the supremum norm.

Definition 2.2.18. Define

$$\begin{aligned} \mathcal{F}: L_1(\mathbb{T}) &\rightarrow c_0(\mathbb{Z}) \\ f &\mapsto \hat{f}. \end{aligned}$$

Remark 2.2.19. \mathcal{F} is linear. It's bounded, because

$$\left| \hat{f}(n) \right| \leq \frac{1}{2\pi} \int_0^{2\pi} |f(x) e^{-inx}| dx = \|f\|.$$

It's injective by the first part of 2.2.16. If \mathcal{F} is onto, then by the corollary $\|\mathcal{F}(f)\|_\infty \geq c \|f\|_1$. But, take $f = D_N = \sum_{n \in [-N, N]} e^{inx}$. Then

$$\|\mathcal{F}(f)\| = \|(0, 0, \dots, 0, 1, 1, \dots, 1, 1, 0, 0, \dots)\| = 1.$$

But, $\|D_N\|_1 \xrightarrow{N \rightarrow \infty} \infty$. This is a contradiction, hence \mathcal{F} is not onto $c_0(\mathbb{Z})$.

2.2.5 The Closed Graph Theorem

Definition 2.2.20 (Graph of a Map). Let X, Y be Banach spaces and let $E \subseteq X$. Let

$$T: E \rightarrow Y$$

be linear. The *graph* of T is

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

Remark 2.2.21. We can define a norm on $X \times Y$ by

$$\|(x, y)\| = \|x\| + \|y\|.$$

$X \times Y$ with this norm is denote $X \oplus_1 Y$ or sometimes $X \oplus Y$. This is a Banach space.

Definition 2.2.22. T is called closed if $\Gamma(T) \subseteq X \oplus Y$ is a closed set.

Proposition 2.2.23. T is closed iff for every $(x_n)_{n \in \mathbb{N}_+} \subseteq E$ such that $x_n \rightarrow x$ implies $Tx_n \rightarrow y$, it holds that $x \in E$ and $y = Tx$.

Proof. • Assume T is closed. If $x_n \rightarrow x$ and $Tx_n \rightarrow y$ then $(x_n, Tx_n) \rightarrow (x, y)$ so $(x, y) \in \Gamma(T)$, so $x \in E$ and $y = Tx$.

• The other direction is left as an exercise. ■

Example 2.2.24. Let $X = Y = \mathcal{C}[0, 1]$ and let $E = \mathcal{C}^1[0, 1]$. Define

$$\begin{aligned} T: E &\rightarrow Y \\ f &\mapsto f'. \end{aligned}$$

Then T is closed. Assume $(f_n)_{n \in \mathbb{N}_+} \subseteq E$ is such that $f_n \rightarrow f$ and $f'_n \rightarrow g$. Then

$$f_n(x) - f_n(0) = \int_0^x f'_n(t) dt \rightarrow \int_0^x g(t) dt$$

where the first expression converges also to $f(x) - f(0)$. Hence

$$f(x) = f(0) + \int_0^x g(t) dt.$$

Hence $f \in E$ and $f' = g$.

Note that E is not closed (it's in fact dense) and that T is not bounded.

Theorem 2.2.25. *Let X, Y be Banach spaces. A closed map $T: X \rightarrow Y$ is continuous.*

Proof. T is closed, hence $\Gamma(T)$ is closed in $X \oplus Y$ and is therefore a Banach space. Let π_X, π_Y be the projections from Γ to X and to Y . π_X is a continuous bijection so π_X^{-1} is continuous by 2.2.12. Then $T = \pi_Y \circ \pi_X^{-1}$ is continuous as a composition of continuous maps. ■

2.2.6 Projections and Quotient Spaces

Definition 2.2.26 (Projection). A *projection* is a linear map $P: X \rightarrow X$ such that $P^2 = P$.

Proposition 2.2.27. *Given a projection $P: X \rightarrow X$ we have $X = \text{Im } P \oplus \ker P$.*

Proof. Let $x \in X$, we can write $x = (x - Px) + Px$ where $Px \in \text{Im}(P)$ and $x - Px \in \ker(P)$. If $x \in \text{Im}(P) \cap \ker(P)$ there's $y \in X$ such that $x = Py$ then

$$0 = Px = P^2y = Py = x$$

so $x = 0$, so the sum is direct. ■

Remark 2.2.28. If $x = e + f$ where $e \in \text{Im } P$ and $f \in \ker P$ we get

$$Px = Pe + Pf = Pe = e.$$

Definition 2.2.29 (Projection onto a Subspace). Let $P: X \rightarrow X$ be a projection, let $E = \text{Im } P$ and $F = \ker P$. We say P is *the projection onto E parallel to F* .

Definition 2.2.30 (Complemented Subspace). A closed subspace $E \leq X$ of a Banach space is called *complemented* if there exists $F \leq X$ closed such that $X = E \oplus F$.

Theorem 2.2.31. *Let X be a Banach space. For a closed $E \leq X$ the following are equivalent.*

1. E is complemented.
2. There is a continuous projection $P: X \rightarrow E$.

Proof. **2** \implies **1**: Take $F = \ker P$ which is closed, and $X = E \oplus F$ since $E = \operatorname{Im} P$ and $F = \ker P$.

1 \implies **2**: Assume $X = E \oplus F$ where E, F are closed subspaces. Take P to be the projection onto E parallel to F .

We show that P is closed. Assume $(x_n)_{n \in \mathbb{N}_+}$ is such that $x_n \rightarrow x$ and $Px_n \rightarrow y$. Since E is closed, $y \in E$. But,

$$x - y = \lim_{n \rightarrow \infty} (x_n - Px_n) \in F$$

since F is closed. Hence

$$x = y + (x - y)$$

where $y \in E$ and $(x - y) \in F$. So, by definition, $Px = y$. Hence P is closed.

By the closed graph theorem, P is then continuous. ■

Fact 2.2.32. 1. c_0 is not complemented in ℓ_∞ .

2. Lindenstaruss-Tzafriri: Every closed subspace of X is complemented if and only if X is isomorphic (in the sense that there's a bijection $T: X \rightarrow H$ such that $c\|x\| \leq \|Tx\| \leq C\|x\|$ for constants $c, C > 0$) to a Hilbert space.

2.2.7 Quotient Spaces

Definition 2.2.33 (Quotient Space). Let X be a Banach space and $E \leq X$ a closed subspace. We define the quotient space

$$X/E = \{x + E \mid x \in X\} = \{[x] \mid x \in X\}$$

where we identify $x \sim y$ iff $x - y \in E$. This is a Banach space with

$$\|x + E\| := \inf_{y \sim x} \|y\| = \inf_{e \in E} \|x - e\| = d(x, E).$$

Fact 2.2.34. If X is a Banach space and E is closed, then X/E is a Banach spaces.

Proof. One should check that X/E is a vector space and that $\|\cdot\|$ defines a norm on it. We prove completeness.

Let $(x_n)_{n \in \mathbb{N}_+} \subseteq X$ such that $\sum_{n \in \mathbb{N}_+} \|[x_n]\| < \infty$. It suffices to show that $(x_n)_{n \in \mathbb{N}_+}$ is convergent. Pick $y_n \sim x_n$ such that $\|y_n\| \leq \|[x_n]\| + \frac{1}{2^n}$. Then

$$\sum_{n \in \mathbb{N}_+} \|y_n\| \leq \sum_{n \in \mathbb{N}_+} \|[x_n]\| + \sum_{n \in \mathbb{N}_+} \frac{1}{2^n} < \infty.$$

Since X is complete, we get $\sum_{n \in \mathbb{N}_+} y_n = y$ for some $y \in X$. Then

$$\begin{aligned} \left\| \sum_{n \in [N]} [x_n] - [y] \right\| &= \left\| \left[\sum_{n \in [N]} x_n - y \right] \right\| \\ &= \left\| \left[\sum_{n \in [N]} y_n - y \right] \right\| \\ &\leq \left\| \sum_{n \in [N]} y_n - y \right\| \\ &\xrightarrow{N \rightarrow \infty} 0. \end{aligned}$$

Hence X/E is complete. ■

Theorem 2.2.35. *Let X, Y be Banach spaces and let $T \in \mathcal{L}(X, Y)$ be surjective. There is an isomorphism*

$$X/\ker T \cong Y.$$

Proof. Let $E := \ker T$ and define

$$\begin{aligned} S: X/E &\rightarrow Y \\ [x] &\mapsto Tx. \end{aligned}$$

Note that

$$[x] = [y] \iff x - y \in E = \ker T \iff T(x - y) = 0 \iff Tx = Ty \iff S([x]) = S([y])$$

so S is well-defined and injective. S is surjective since T is surjective

Let $[x] \in X/E$. For every $\varepsilon > 0$ there's $y \in X/E$ such that $[y] = [x]$ and $\|y\| \leq \|[x]\| + \varepsilon$. Then

$$\begin{aligned} \|S([x])\| &= \|S([y])\| \\ &= \|Ty\| \\ &\leq \|T\| \|y\| \\ &\leq \|T\| (\|[x]\| + \varepsilon). \end{aligned}$$

Letting $\varepsilon > 0$ we get $\|S\| \leq \|T\|$ (in fact there's equality $\|S\| = \|T\|$). Hence S is continuous.

By the open mapping theorem a continuous bijection is an isomorphism, hence so is S . ■

Corollary 2.2.36. *Let X be a Banach space and let $E, F \leq X$ be closed subspaces such that $X = E \oplus F$. It holds that*

$$X/E \cong F.$$

Proof. Let $P: X \rightarrow F$ be the projection onto F parallel to E . This is a surjective and continuous (which follows from the open graph theorem), hence 2.2.35 gives

$$X/E = X/\ker P \cong F. \quad \blacksquare$$

2.2.8 Schauder Bases

Definition 2.2.37 (Schauder Base). A sequence $(e_n)_{n \in \mathbb{N}_+}$ in a Banach space X is a (Schauder) base if for every $x \in X$ there are unique $(\alpha_n)_{n \in \mathbb{N}_+} \subseteq \mathbb{F}$ such that

$$x = \sum_{n \in \mathbb{N}_+} \alpha_n e_n.$$

Example 2.2.38. In a Hilbert space, an orthonormal base is a (Schauder) base, and $\alpha_n = \langle x, e_n \rangle$.

Example 2.2.39. In ℓ_p for $p \in [1, \infty)$, the vectors e_n defined by $(e_n)_i = \delta_{n,i}$, form a base. Indeed $a = (a_1, a_2, a_3, \dots)$ can be written uniquely as $\sum_{n \in \mathbb{N}_+} a_n e_n$.

Example 2.2.40. In $\mathcal{C}[0, 1]$, take $e_n(t) = t^n$ for every $n \in \mathbb{N}$. We know by the Weierstrass approximation theorem that $\overline{\text{Span}\{e_n \mid n \in \mathbb{N}\}} = \mathcal{C}[0, 1]$.

However, $(e_n)_{n \in \mathbb{N}}$ is not a basis. If $f = \sum_{n=0}^{\infty} a_n t^n$ then $f \in \mathcal{C}^\infty$. Hence we cannot write any $f \in \mathcal{C}[0, 1] \setminus \mathcal{C}^\infty[0, 1]$ as $\sum_{n \in \mathbb{N}} a_n t^n$.

Definition 2.2.41. Given a basis $(e_n)_{n \in \mathbb{N}_+}$, define $\alpha_n: X \rightarrow \mathbb{F}$ by

$$x = \sum_{n \in \mathbb{N}_+} \alpha_n(x) e_n$$

and

$$\begin{aligned} P_N: X &\rightarrow X \\ X &\mapsto \sum_{n \in [N]} \alpha_n(x) e_n. \end{aligned}$$

Remark 2.2.42. In Hilbert spaces with an orthonormal basis it holds that $\|\alpha_n\| = 1$ and $\|P_N\| = 1$.

Remark 2.2.43. If X has a countable base, it's separable. One might ask if any separable space has a base. Enlfó showed in 1973 that the answer is no.

Theorem 2.2.44. If $(e_n)_{n \in \mathbb{N}_+}$ is a base for a Banach space X . There exists $C > 0$ such that $\|P_k\| \leq C$, and $\alpha_n \in X^*$.

Proof. Define $\|\cdot\|_1$ on X by

$$\|x\|_1 := \sup_{k \in \mathbb{N}_+} \|P_k x\|.$$

This is finite since $P_k x \xrightarrow{n \rightarrow \infty} x$ and it is easily checked that this is a norm. We have

$$\forall x \in X: \|x\|_1 = \lim_{k \rightarrow \infty} \|P_k x\| = \|x\|.$$

- We show $(X, \|\cdot\|_1)$ is complete. Let $(x_n)_{n \in \mathbb{N}_+}$ be $\|\cdot\|_1$ -Cauchy. By our bound, it's also $\|\cdot\|$ -Cauchy. Hence there's

$$x := \|\cdot\| - \lim_{n \rightarrow \infty} x_n.$$

For every $k \in \mathbb{N}_+$, $(P_k x_n)_{n \in \mathbb{N}_+}$ is also $\|\cdot\|$ -Cauchy, so $P_k x_n \xrightarrow{n \rightarrow \infty} y_k$ for some $y_k \in X$. We want to show

$$\lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} P_k x_n = \lim_{k \rightarrow \infty} \lim_{n \rightarrow \infty} P_k x_n$$

which then implies

$$\begin{aligned} x &= \lim_{n \rightarrow \infty} P_k x \\ &= \lim_{n \rightarrow \infty} \lim_{k \rightarrow \infty} P_k x_n \\ &= \lim_{k \rightarrow \infty} \lim_{n \rightarrow \infty} P_k x_n \\ &= \lim_{k \rightarrow \infty} y_k \end{aligned}$$

Fix $\varepsilon > 0$. Because $(x_n)_{n \in \mathbb{N}_+}$ is Cauchy, there's $n_0 \in \mathbb{N}_+$ such that for all $m \geq n \geq n_0$ we have $\|x_m - x_n\|_1 < \varepsilon$. Hence

1. For every $m \geq n \geq n_0$ it holds that $\|x_n - x_m\| < \varepsilon$. Letting $m \rightarrow \infty$ we get $\|x_n - x\| < \varepsilon$.
2. For every $k \in \mathbb{N}_+$ and every $m \geq n \geq n_0$ we know $\|P_k x_m - P_k x_n\| < \varepsilon$ so by letting $m \rightarrow \infty$ we get $\|y_k - P_k x_n\| < \varepsilon$.

Choose $k_0 \in \mathbb{N}_+$ such that $\|x_{n_0} - P_{k_0} x_{n_1}\| < \varepsilon$ for all $k \geq k_0$. Then

$$\|y_k - x\| \leq \|y_k - P_k x_{n_0}\| + \|P_k x_{n_0} - x_{n_0}\| + \|x_{n_0} - x\| < 3\varepsilon.$$

Then indeed $y_k \xrightarrow{k \rightarrow \infty} x$ in $\|\cdot\|$.

Fix $1 \leq j \leq k$. Then α_j is bounded on $E_k := \text{Span}(e_i)_{i \in [k]}$. We have

$$\lim_{n \rightarrow \infty} = \lim_{n \rightarrow \infty} \alpha_j(P_k x_n) = \alpha_j(y_k).$$

Hence $\alpha_j(y_k)$ doesn't depend on k . Denote $c_j := \alpha_j(y_k)$. Hence

$$y_k = \sum_{j \in [k]} c_j e_j.$$

Hence

$$x = \sum_{j \in \mathbb{N}_+} c_j e_j.$$

We now show $x_n \xrightarrow{n \rightarrow \infty} x$ in $\|\cdot\|_1$. Fix $\varepsilon > 0$ as before. Then

$$\exists n_0 \in \mathbb{N}_+ \forall k \in \mathbb{N}_+ \forall n \geq n_0: \|y_k - P_k x_n\| < \varepsilon.$$

Hence

$$\begin{aligned} \|x_n - x\|_1 &= \sup_{k \in \mathbb{N}_+} \|P_k x_n - P_k x\| \\ &= \sup_{k \in \mathbb{N}_+} \|P_k x_n - y_k\| \\ &< \varepsilon. \end{aligned}$$

Hence $(X, \|\cdot\|_1)$ is complete.

- By the corollary we have $\|x\|_1 \leq C \|x\|$. This implies

$$\forall k \in \mathbb{N}_+: \|P_k x\| \leq C \|x\|$$

so $\|P_k\| \leq C$.

For the part $\alpha_n \in X^*$, we have

$$\begin{aligned} |\alpha_n(x)| \|e_n\| &= \|\alpha_n(x) e_n\| \\ &= \|P_n x - P_{n-1} x\| \\ &\leq 2C \|x\| \end{aligned}$$

$$\text{so } \|\alpha_n\| \leq \frac{2C}{\|e_n\|}. \quad \blacksquare$$

2.3 Finite-Dimensional Spaces

2.3.1 Definitions

Definition 2.3.1 (Isomorphic Normed Spaces). Let X, Y be normed spaces. We say X, Y are *isomorphic* if there's a bijection $T: X \rightarrow Y$ and there are $c, C \in \mathbb{R}_+$ such that

$$\forall x \in X: c \|x\|_X \leq \|Tx\|_Y \leq C \|x\|_X.$$

Theorem 2.3.2. *Let X, Y be normed spaces with $\dim X = \dim Y = n \in \mathbb{N}$. Then X, Y are isomorphic.*

Corollary 2.3.3. *Every finite-dimensional normed space is a Banach space.*

Corollary 2.3.4. *If X is a finite-dimensional normed space, any linear $T: X \rightarrow Y$ is bounded.*

Definition 2.3.5 (The Banach-Mazur distance). Let X, Y be isomorphic Banach spaces. The *Banach-Mazur distance* is

$$\begin{aligned} d_{\text{BM}}(X, Y) &:= \inf \left\{ \|T\| \|T^{-1}\| \mid T: X \rightarrow Y \text{ is a linear bijection} \right\} \\ &= \inf \left\{ \frac{b}{a} \mid \begin{array}{c} \text{There's a linear bijection } T: X \rightarrow Y \text{ such that} \\ \forall x \in X: a\|x\| \leq \|Tx\| \leq b\|x\| \end{array} \right\}. \end{aligned}$$

Proposition 2.3.6. 1. $d_{\text{BM}}(X, Y) \geq 1$ and $d_{\text{BM}}(X, Y) = 1$ if and only if X, Y are isometric.

2. $d_{\text{BM}}(X, Y) = d_{\text{BM}}(Y, X)$.

3. $d_{\text{BM}}(X, Z) \leq d_{\text{BM}}(X, Y) d_{\text{BM}}(Y, Z)$.

Corollary 2.3.7. $\log d_{\text{BM}}$ is a metric on the space of isomorphism classes of finite-dimensional normed spaces quotiented by the isometric relation.

Proposition 2.3.8. $d_{\text{BM}}(\ell_1^n, \ell_2^n) = \sqrt{n}$.

Proof. • We first show that $d_{\text{BM}}(\ell_1^n, \ell_2^n) \leq \sqrt{n}$.

Recall that

$$\|x\|_1 = \sum_{i \in [n]} |x_i| \cdot 1 \leq \left(\sum_{i \in [n]} |x_i|^2 \right)^{\frac{1}{2}} \cdot \left(\sum_{i \in [n]} 1^2 \right)^{\frac{1}{2}} = \sqrt{n} \|x\|_2.$$

For the other inequality, assume $\|x\|_1 = 1$. The $\forall i \in [n]: |x_i| \leq 1$. Then $\forall i \in [n]: |x_i|^2 \leq |x_i|$. Hence

$$\|x\|_2 = \left(\sum_{i \in [n]} |x_i|^2 \right)^{\frac{1}{2}} \leq \left(\sum_{i \in [n]} |x_i| \right)^{\frac{1}{2}} = 1.$$

Hence by homogeneity

$$\forall x \in X: \|x\|_2 = \left\| \frac{x}{\|x\|_1} \right\|_2 \cdot \|x\|_1 \leq 1 \cdot \|x\|_1 = \|x\|_1.$$

Hence

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n} \|x\|_2.$$

Hence

$$d_{\text{BM}}(\ell_1^n, \ell_2^n) \leq \frac{\sqrt{n}}{1} = \sqrt{n}.$$

- Let $T: \ell_1^n \rightarrow \ell_2^n$ be a bijection such that

$$\forall x \in X: a \|x\|_1 \leq \|Tx\|_2 \leq b \|x\|_1.$$

By the parallelogram law

$$\begin{aligned} \sum_{(\theta_i)_{i \in [n]} \subseteq \{\pm 1\}} \left\| \sum_{i \in [n]} \theta_i T e_i \right\|_2^2 &= \sum_{(\theta_i)_{i \in [n]} \subseteq \{\pm 1\}} \left\langle \sum_{i \in [n]} \theta_i T e_i, \sum_{j \in [n]} \theta_j T e_j \right\rangle \\ &= \sum_{(\theta_i)_{i \in [n]} \subseteq \{\pm 1\}} \sum_{i, j \in [n]} \theta_i \theta_j \langle T e_i, T e_j \rangle \\ &= \sum_{i, j \in [n]} \left(\sum_{(\theta_i)_{i \in [n]} \subseteq \{\pm 1\}} \theta_i \theta_j \right) \langle T e_i, T e_j \rangle \\ &= 2^n \sum_{i \in [n]} \langle T e_i, T e_i \rangle \\ &= 2^n \sum_{i \in [n]} \|T e_i\|_2^2 \\ &\leq 2^n \sum_{i \in [n]} (b \|e_i\|_1)^2 \\ &= 2^n \cdot n \cdot b^2. \end{aligned}$$

On the other hand,

$$\begin{aligned} \sum_{(\theta_i)_{i \in [n]} \subseteq \{\pm 1\}} \left\| \sum_{i \in [n]} \theta_i T e_i \right\|_2^2 &\geq \sum_{(\theta_i)_{i \in [n]} \subseteq \{\pm 1\}} \left(a \left\| \sum_{i \in [n]} \theta_i e_i \right\| \right)^2 \\ &= \sum_{(\theta_i)_{i \in [n]} \subseteq \{\pm 1\}} (an)^2 \\ &= 2^n n^2 a^2. \end{aligned}$$

Combining the inequalities we get $2^n n^2 a^2 \leq 2^n n b^2$, from which we get $\frac{b}{a} \geq \sqrt{n}$. ■

Theorem 2.3.9 (Auerbach). *Let X be a real normed space of dimension $n \in \mathbb{N}$. There exists a basis $(e_i)_{i \in [n]}$ such that for every $x = \sum_{i \in [n]} \alpha_i e_i$ in X we have $|\alpha_i| \leq \|x\|$ and $\|e_i\| = 1$.*

Equivalently, if we write $x = \sum_{i \in [n]} \alpha_i(x) e_i$, we have $\|\alpha_i\| = 1$ (where the norm on α_i is in X^) and $\|e_i\| = 1$.*

We call such a basis an Auerbach basis.

Proof. Without loss of generality, write $X = (\mathbb{R}^n, \|\cdot\|)$ for some norm $\|\cdot\|$. For $(y_i)_{i \in [n]} \subseteq \mathbb{R}^n$ define

$$D(y_1, \dots, y_n) := \det \begin{pmatrix} | & & | \\ y_1 & \cdots & y_n \\ | & & | \end{pmatrix}.$$

Choose $(e_i)_{i \in [n]}$ to maximise $D(e_1, \dots, e_n)$ such that $\|e_i\| = 1$ for all $i \in [n]$. This exists by compactness. Since $D(e_1, \dots, e_n) > 0$ we get that $(e_i)_{i \in [n]}$ is a base. Let

$$f_i(x) := \frac{D(e_1, \dots, e_{i-1}x, e_{i+1}, \dots, e_n)}{D(e_1, \dots, e_{i-1}, e_i, e_{i+1}, \dots, e_n)}.$$

Then f_i is linear, it holds that $f_i(e_j) = \delta_{i,j}$. Hence $f_i = \alpha_i$. Hence if $\|x\| = 1$ we have

$$|\alpha_i(x)| = |f_i(x)| \leq 1$$

where inequality is by maximality of $D(e_1, \dots, e_n)$. Hence $\|\alpha_i\| \leq 1$. ■

Corollary 2.3.10. *If $\dim X = n \in \mathbb{N}$ it holds that $d_{\text{BM}}(X, \ell^n) \leq n$.*

Proof. Let $(\tilde{e}_i)_{i \in [n]}$ be an Auerbach base of X and let

$$\begin{aligned} T: X &\rightarrow \ell_1^n \\ \tilde{e}_i &\mapsto e_i. \end{aligned}$$

Then

$$\begin{aligned} \|Tx\| &= \left\| T \left(\sum_{i \in [n]} \alpha_i(x) \tilde{e}_i \right) \right\| \\ &\leq \sum_{i \in [n]} \alpha_i(x) \|T\tilde{e}_i\| \\ &= \sum_{i \in [n]} \alpha_i(x) \\ &\leq n \|x\|. \end{aligned}$$

Conversely, for $a = (a_1, \dots, a_n) \in \ell_1$ we have

$$\begin{aligned} \|T^{-1}a\| &= \left\| T^{-1} \sum_{i \in [n]} a_i e_i \right\| \\ &= \left\| \sum_{i \in [n]} a_i \tilde{e}_i \right\| \\ &\leq \sum_{i \in [n]} |a_i| \underbrace{\|\tilde{e}_i\|}_1 = \|a\|_1. \end{aligned}$$

Hence

$$d_{\text{BM}}(X, \ell_1^n) \leq \|T\| \|T^{-1}\| \leq n \cdot 1 = n. \quad \blacksquare$$

Corollary 2.3.11. *If $\dim X = \dim Y = n$, then $d_{\text{BM}}(X, Y) \leq n^2$.*

Fact 2.3.12 (F. John). *If $\dim X = n$ it holds that $d_{\text{BM}}(X, \ell_2^n) \leq \sqrt{n}$. So, for every X, Y of dimension n it holds that $d_{\text{BM}}(X, Y) \leq n$.*

Fact 2.3.13 (Gluskin). *There are X, Y of dimension n such that $d_{\text{BM}}(X, Y) \geq \frac{n}{1000000}$.*

2.3.2 Geometric Interpretation

Let $X = (\mathbb{R}^n, \|\cdot\|)$ and let

$$K = \{x \mid \|x\| \leq 1\}$$

which is convex.

Aurbach's theorem says that we can apply a linear transformation to a convex body such that the image contains $\pm e_i$ for every base element e_i , and such that the 1-norm of points in the image is less than 1.

Theorem 2.3.14. *Let X be a Banach space. Then $\dim X < \infty$ if and only if $\bar{B}_X := \bar{B}(0, 1)$ is compact.*

Lemma 2.3.15 (Riesz Lemma). *Let X be a normed space and let $E \subsetneq X$ be a closed subspace. Then for every $\varepsilon > 0$ there's $x_\varepsilon \in X$ such that $\|x_\varepsilon\| = 1$ and $d(x_\varepsilon, E) \geq 1 - \varepsilon$.*

Proof. Fix $x \in X \setminus E$ and write $d := d(x, E) > 0$. Choose $e \in E$ such that $\|x - e\| \leq d(1 + \varepsilon)$ and define $x_\varepsilon = \frac{x - e}{\|x - e\|}$. Then $\|x_\varepsilon\| = 1$ and for all $y \in E$ it holds that

$$\|x_\varepsilon - y\| = \frac{\|x - e - y\|}{\|x - e\|} \geq \frac{d}{d(1 + \varepsilon)} = \frac{1}{1 + \varepsilon} \geq 1 - \varepsilon. \quad \blacksquare$$

Proof (2.3.14). If X is finite-dimensional, it's isomorphic to \mathbb{R}^n , so the closed unit ball is compact.

Assume $\dim X = \infty$, we want to show that $\bar{B}_X := \bar{B}(0, 1)$ is not compact. Let $x_1 \in X$ with $\|x_1\| = 1$, and let $E_1 = \text{Span}\{x_1\} \neq X$. This is closed and by Riesz's lemma there's $x_2 \in X \setminus E_1$ such that $\|x_2\| = 1$ and $d(x_2, E_1) \geq 0.9$. Set $E_2 = \text{Span}\{x_1, x_2\} \neq X$ which is closed. By Riesz lemma there's again $x_3 \in X \setminus E_2$ of norm 1 such that $d(x_3, E_2) \geq 0.9$, and so on. Then $(x_n)_{n \in \mathbb{N}_+} \subseteq \bar{B}_X$ but

$$\|x_n - x_m\| \geq d(x_n, E_n) \geq d(x_n, E_{n-1}) \geq 0.9$$

so there is no convergent subsequence, so \bar{B}_X is not compact. \blacksquare

Corollary 2.3.16. *If $\dim X = \infty$ and $A \subseteq X$ is compact, it holds that $\int A = \emptyset$.*

Proof. Otherwise, $\bar{B}(X_0, \frac{r}{2}) \subseteq B(x_0, r) \subseteq A$ so $\bar{B}(x_0, \frac{r}{2}) = \frac{r}{2}\bar{B}_X + x_0$ is compact, so \bar{B}_X is compact, a contradiction. ■

The above corollary raises a problem. Compact sets are comfortable to work with, and one wants to have compact sets other than those with empty interior. To fix this we later define new topologies on Banach spaces.

Chapter 3

Weak Topologies

3.1 Weak Topologies

3.1.1 Definitions

Definition 3.1.1 (Weak Convergence). Let X be a Banach space and let $(x_n)_{n \in \mathbb{N}} \subseteq X$ and $x \in X$. We say $x_n \xrightarrow{n \rightarrow \infty} x$ weakly if for every $f \in X^*$ it holds that $f(x_n) \xrightarrow{n \rightarrow \infty} f(x)$.

Notation 3.1.2. We denote weak convergence by $x_n \rightharpoonup x$, by $\text{w-lim}_{n \rightarrow \infty} x_n = x$ or by $x_n \xrightarrow{w} x$.

Example 3.1.3. Let H be a Hilbert space and let $(e_n)_{n \in \mathbb{N}} \subseteq H$ be orthonormal. Every $f \in H^*$ is of the form $f(x) = \langle x, y \rangle$ and by Bessel it holds that

$$\sum_{n \in \mathbb{N}} |\langle y, e_n \rangle|^2 \leq \|y\|^2 < \infty.$$

Then $\langle y, e_n \rangle \rightarrow 0$ so $f(e_n) = \langle e_n, y \rangle \xrightarrow{n \rightarrow \infty} 0 = f(0)$. Then $e_n \xrightarrow{n \rightarrow \infty} 0$. But, of course $\|e_n - e_m\| = \sqrt{2}$ for every n, m different, then $(e_n)_{n \in \mathbb{N}}$ doesn't converge (strongly).

Proposition 3.1.4. *A weak limit is unique.*

Proof. Assume that for every $f \in X^*$ it holds that $f(x_n) \xrightarrow{n \rightarrow \infty} f(x), f(y)$. Then $f(x) = f(y)$ for every $f \in X^*$, and by Hahn-Banach it follows that $x = y$. ■

Proposition 3.1.5. 1. *If $x_n \xrightarrow{n \rightarrow \infty} x$ then $(x_n)_{n \in \mathbb{N}}$ is bounded.*

2. *If $x_n \xrightarrow{n \rightarrow \infty} x$ then $\|x\| \leq \liminf_{n \rightarrow \infty} \|x_n\|$.*

Proof. 1. For every f we have $f(x_n) \xrightarrow{n \rightarrow \infty} f(x)$. Hence $(f(x_n))_{n \in \mathbb{N}}$ is bounded. By uniform boundedness, $(x_n)_{n \in \mathbb{N}}$ is bounded.

2. By Hahn-Banach there's $f \in X^*$ such that $\|f\| = 1$ and $f(x) = \|x\|$. Now

$$\begin{aligned} \|x\| &< |f(x)| \\ &= \lim_{n \rightarrow \infty} |f(x_n)| \\ &\leq \liminf_{n \rightarrow \infty} \|f\| \|x_n\| \\ &= \liminf_{n \rightarrow \infty} \|x_n\|. \end{aligned} \quad \blacksquare$$

Proposition 3.1.6. *Let S be a compact metric space and let $X = \mathcal{C}(S)$. Let $(f_n)_{n \in \mathbb{N}}, f \in X$. Then the following are equivalent.*

1. $f_n \xrightarrow{n \rightarrow \infty} f$.
2. $f_n(s) \xrightarrow{n \rightarrow \infty} f(s)$ for all s , and $\sup_{n \in \mathbb{N}} \sup_{s \in S} |f_n(s)| < \infty$.

Proof. 1. We prove 2 assuming 1. Pick $\delta_s \in \mathcal{C}(S)^*$ which is $\delta_s(f) = f(s)$. We know $\delta_s(f_n) \xrightarrow{n \rightarrow \infty} \delta_s(f)$. Also, $\sup_{n \in \mathbb{N}} \sup_{s \in S} |f_n(s)| = \sup_{n \in \mathbb{N}} \|\delta_s(f_n)\| < \infty$ by the proposition.

We prove the other direction. By measure theory, $\varphi \in \mathcal{C}(S)^*$ is of the form $\varphi(f) = \int f d\mu$. Now $\varphi(f_n) \xrightarrow{n \rightarrow \infty} \varphi(f)$ by dominated convergence. \blacksquare

Example 3.1.7 (Schur's Theorem). In ℓ_1 , weak convergence implies strong convergence.

Definition 3.1.8 (The Weak Topology). The weak topology on a Banach space X is that with the sub-base $\left\{ V_{f,a,\delta} \mid \begin{array}{l} f \in X^* \\ a \in \mathbb{F} \\ \delta > 0 \end{array} \right\}$ where

$$V_{f,a,\delta} = \{x \in X \mid |f(x) - a| < \delta\}.$$

Proposition 3.1.9. 1. *A local base for every $z \in X$ is given by sets of the form*

$$U_{f_i,z,\delta} := \{x \in X \mid \forall i \in [m]: |f_i(x) - f_i(z)| < \delta\}$$

for $\delta > 0$ and $f_1, \dots, f_m \in X^$.*

2. *The weak topology is the weakest topology on X such that every $f \in X^*$ is continuous.*
3. *$x_n \xrightarrow{n \rightarrow \infty} x$ in the weak topology if and only if $x_n \xrightarrow{n \rightarrow \infty} x$.*
4. *The weak topology is weaker than the norm topology. It is strictly weaker if and only if $\dim X = \infty$.*
5. *The weak topology is Hausdorff.*
6. *$+: X \times X \rightarrow X$ and $\cdot: \mathbb{F} \times X \rightarrow X$ are continuous in the weak topology.*

Proof. 1. Let U is weakly open and $z \in U$. Then

$$z \in \bigcap_{i \in [m]} V_{f_i, a_i, \delta_i} \subseteq U.$$

Then $f_i(z) \in B_{\mathbb{F}}(a_i, \delta_i)$ so we can choose $\delta > 0$ such that $B_{\mathbb{F}}(f_i(z), \delta) \subseteq B_{\mathbb{F}}(a_i, \delta_i)$ for all $i \in [m]$. Then $z \in U_{f_i, z, \delta} \subseteq U$.

2. Fix $f \in X^*$. For every $a \in \mathbb{F}$ and $\delta > 0$ we have

$$V_{f, a, \delta} = f^{-1}(B_{\mathbb{F}}(a, \delta))$$

is weakly-open. Hence f is weakly-continuous. If every $f \in X^*$ is continuous with respect to some topology τ , then $V_{f, a, \delta} \in \tau$ by the same argument, so the weak topology w is contained in τ .

3. Assume convergence in the weak topology. Since every $f \in X^*$ is continuous in the weak topology, we have weak convergence.

For the other direction, assume $x_n \xrightarrow{n \rightarrow \infty} x$. Fix $U_{f_i, x, \delta}$. Then

$$\forall i \in [m]: f_i(x_n) \xrightarrow{n \rightarrow \infty} f_i(x),$$

so

$$\exists n_0 \in \mathbb{N} \forall n \geq n_0: |f_i(x_n) - f_i(x)| < \delta.$$

Then

$$\forall n \geq n_0: x_n \in U_{f_i, x, \delta}.$$

4. We prove that if $\dim X = \infty$ then every weakly-open set U is unbounded. In particular $B(0, 1)$ is not weakly open.

It's enough to prove the statement in the case $U = U_{f, x, \delta}$. Note that by linear algebra there's $y \neq 0$ such that $f_i(y) = 0$ for all $i \in [m]$. Then $x + ty \in U_{f, x, \delta}$ for all $t \in \mathbb{F}$, so $U_{f, x, \delta}$ is unbounded.

The case where $\dim X < \infty$ is an exercise.

5. Fix $x, y \in X$ different. There's $f \in X^*$ such that $f(x) \neq f(y)$. Write $f = |f(x) - f(y)|$. Then

$$U_{f, x, \frac{\delta}{2}} \cap U_{f, y, \frac{\delta}{2}} = \emptyset$$

where x is in the first set and y in the second.

6. We prove that $+$ is continuous. It's enough to note

$$U_{f, x, \frac{\delta}{2}} + U_{f, y, \frac{\delta}{2}} \subseteq U_{f, x+y, \delta}$$

which follows from the triangle inequality in \mathbb{F} . ■

Proposition 3.1.10. *If $\dim X = \infty$ then $\exists A \subseteq X$ such that $0 \in \bar{A}^w$ but there is no $(x_n)_{n \in \mathbb{N}} \subseteq A$ such that $x_n \xrightarrow{n \rightarrow \infty} 0$.*

Corollary 3.1.11. *Let X be a Banach space such that $\dim X = \infty$ and let w be the weak topology on it. (X, w) is not metrisable.*

Proof. Choose subspaces $E_1 \subseteq E_2 \subseteq \dots$ such that $\dim E_n = n$. Let

$$A = \bigcup_{n \in \mathbb{N}_+} \{x \in E_n \mid \|x\| = n\}.$$

Fix an open set U around zero and assume WLOG that $U = U_{f_i, 0, \delta}$. There's $y \in E_{m+1}$ such that $f_i(y) = 0$ for every $i \in [m]$. Then

$$(m+1) \frac{y}{\|y\|} \in U_{f_i, 0, \delta} \cap A \neq \emptyset$$

so $0 \in \bar{A}^w$.

Assume $(x_n)_{n \in \mathbb{N}} \subseteq A$ converges weakly to zero. Then $\|x_n\| \leq m$ for some m , so $(x_n)_{n \in \mathbb{N}} \subseteq E_m$. Since $\dim E_m < \infty$ we have $x_n \xrightarrow{n \rightarrow \infty} 0$ in norm, which is impossible since $\|x\| \geq 1$. ■

Proposition 3.1.12. *Let $K \subseteq X$ be convex. Then K is closed iff it's weakly-closed.*

Proof. If K is weakly-closed it's closed because the weak topology is weaker.

Assume K is norm-closed. Fix $x_0 \notin K$. Choose $d > 0$ such that $B(x_0, d) \cap K = \emptyset$. By Convex separation there exists $f \in X^*$ such that $\sup_{a \in K} f(a) \leq \inf_{b \in B(x_0, d)} f(b)$. Pick any $y \in X$ such that $f(y) > 0$ and $\|y\| = 1$. Then

$$f(x_0) = f\left(x_0 - \frac{d}{2}y\right) + f\left(\frac{d}{2}y\right)$$

where $x_0 - \frac{d}{2}y \in B(x_0, d)$ so $f(x_0) > c + 0 = c$. But now

$$U = \{x \mid f(x) > c\}$$

is weakly-open, $x_0 \in U$ and $U \cap K = \emptyset$. Hence K is weakly-closed. ■

Corollary 3.1.13. 1. $\bar{B}(0, 1)$ is weakly-closed.

2. If $(x_n)_{n \in \mathbb{N}} \subseteq K$ where K is convex and closed, and $x_n \xrightarrow{n \rightarrow \infty} x$, then $x \in K$.

Corollary 3.1.14 (Mazur). *If $x_n \xrightarrow{n \rightarrow \infty} x$ then $\exists (y_n)_{n \in \mathbb{N}}$ such that*

1. $y_k \in \text{Conv}(x_n)_{n \in \mathbb{N}}$ for every $k \in \mathbb{N}$.

2. $y_n \xrightarrow{n \rightarrow \infty} x$.

Proof. Take $K = \overline{\text{Conv}(x_n)_{n \in \mathbb{N}}}$. This is convex and closed so it's w-closed. Then $K = \overline{\text{Conv}(x_n)_{n \in \mathbb{N}}}^w$. Therefore $x \in \overline{(x_n)_{n \in \mathbb{N}}}^w \subseteq K$. Hence there's $(y_n)_{n \in \mathbb{N}} \subseteq \text{Conv}(x_n)_{n \in \mathbb{N}}$ such that $y_n \xrightarrow{n \rightarrow \infty} x$. ■

Theorem 3.1.15 (Banach-Saks, 1). *Let H be a Hilbert space and let $(x_n)_{n \in \mathbb{N}}$ weakly convergent to x . Then there's a subsequence $(x_{n_k})_{k \in \mathbb{N}}$ such that*

$$\frac{1}{k} \sum_{i \in [k]} x_{n_i} \xrightarrow{k \rightarrow \infty} x.$$

Proof. Assume without loss of generality that $x = 0$. We have

$$\begin{aligned} I &:= \left\| \frac{1}{k} \sum_{i \in [k]} x_{n_i} \right\|^2 = \left\langle \frac{1}{k} \sum_{i \in [k]} x_{n_i}, \frac{1}{k} \sum_{j \in [k]} x_{n_j} \right\rangle \\ &\leq \frac{1}{k^2} \sum_{i \in [k]} \|x_{n_i}\|^2 + \frac{2}{k^2} \sum_{i, j \in [k]} |\langle x_{n_i}, x_{n_j} \rangle|. \end{aligned}$$

Choose $x_{n_1} = x_1$. Assume we already chose x_{n_i} for $i \in [j-1]$. Since $x_n \xrightarrow{n \rightarrow \infty} 0$ we have $\langle x_m, x_{n_i} \rangle \xrightarrow{m \rightarrow \infty} 0$ for all $i \in [k-1]$. Hence we can choose x_{n_k} such that $|\langle x_{n_k}, x_{n_i} \rangle| \leq \frac{1}{2^k}$ for all $i \in [k-1]$. Since $x_n \xrightarrow{n \rightarrow \infty} 0$, we have also $\|x_n\| \leq C$ for all $n \in \mathbb{N}$ and some $C > 0$. Then

$$\begin{aligned} I &\leq \frac{1}{k^2} Ck + \frac{2}{k^2} \sum_{i, j \in [k]} \frac{1}{2^j} \\ &= \frac{C}{K} + \frac{2}{k^2} \sum_{j \in [k]} \frac{j}{2^j} \\ &\xrightarrow{k \rightarrow \infty} 0. \end{aligned} \quad \blacksquare$$

3.1.2 Weak-* Topologies

Definition 3.1.16. Let X be a Banach space and let $F \leq X^*$. The $\sigma(X, F)$ topology on X is the weakest topology such that every $f \in F$ is continuous.

Proposition 3.1.17. 1. A local base at X for the $\sigma(X, F)$ topology is given by sets

$$U_{f_i, x, \delta} := \{y \in X \mid \forall i \in [m]: |f_i(y) - f_i(x)| < \delta\}$$

for $f_1, \dots, f_m \in F$ and $\delta > 0$.

2. $\sigma(X, F) \subseteq w$.

3. $+, \cdot$ are continuous with respect to the $\sigma(X, F)$ topology.

4. $(X, \sigma(X, F))$ is Hausdorff if F separates points.

Proposition 3.1.18. *If $f \in X^*$ then $f: (X, \sigma(X, F)) \rightarrow \mathbb{F}$ is continuous if and only if $f \in F$.*

Lemma 3.1.19. *Let V be a vector space over \mathbb{F} and let $f_1, \dots, f_m, g: V \rightarrow \mathbb{F}$ be linear functionals. Assume that for every x , $f_1(x) = \dots = f_m(x) = 0$ implies $g(x) = 0$. Then $g \in \text{Span}\{f_i \mid i \in [m]\}$.*

Proof. Define

$$\begin{aligned} T: V &\rightarrow \mathbb{F}^m \\ x &\mapsto (f_i(x))_{i \in [m]} \end{aligned}$$

and

$$\begin{aligned} \varphi: \text{Im } T &\rightarrow \mathbb{F} \\ Tx &\mapsto g(x). \end{aligned}$$

φ is well-defined because $Tx = Ty$ implies $\forall i \in [m]: f_i(x) = f_i(y)$ which implies $f_i(x - y) = 0$ so $g(x - y) = 0$, so $g(x) = g(y)$. We know $\varphi(a_1, \dots, a_m) = \sum_{i \in [m]} \lambda_i a_i$ so

$$g(x) = \varphi(Tx) = \varphi(f_1(x), \dots, f_m(x)) = \sum_{i \in [m]} \lambda_i f_i(x). \quad \blacksquare$$

Proposition 3.1.20. *If $f \in X^*$ then $f: (X, \sigma(X, F)) \rightarrow \mathbb{F}$ is continuous if and only if $f \in F$.*

Proof. Assume f is continuous. It holds that $f(0) = 0$ so continuity implies

$$f^{-1}(B(0, 1)) \supseteq U_{f_i, 0, \delta}$$

for $f_1, \dots, f_m \in F$. So, if $|f_i(z)| < \delta$ for all $i \in [m]$ then $\text{abs } f(z) < 1$. If $f_i(z) = 0$ for all $i \in [m]$ then for all $\lambda > 0$ it holds that $|f_i(\lambda z)| < \delta$, so $|\lambda| |f(z)| = |f(\lambda z)| < 1$, so $f(z) = 0$. By the lemma, $f \in \text{Span}\{f_1, \dots, f_m\} \subseteq F$, as required. \blacksquare

Example 3.1.21. $\sigma(X, X^*) = w$.

Example 3.1.22. Fix X a Banach space. The topology $\sigma(X^*, X)$ on X^* is called the weak-* topology on X^* . This is denoted w^* and is the weakest topology on X^* such that $f \mapsto f(x)$ are all continuous.

Example 3.1.23. Let $X = \mathcal{C}([-1, 1])$. Then $X^* = M([-1, 1])$ is the space of measures on $[-1, 1]$. Choose $\mu_n \in X^*$ by $\mu_n(f) = \int f d\mu_n = \frac{n}{2} \int_{[-\frac{1}{n}, \frac{1}{n}]} f dt$.

The for every $f \in X$ it holds that $\mu_n(f) \xrightarrow{n \rightarrow \infty} f(0) = \delta_0(f)$, so $\mu_n \xrightarrow{w^*} \delta_0$.

But, $\mu_n \not\xrightarrow{w} \delta_0$. Indeed, choose $\varphi \in X^{**}$ by $\varphi(\mu) = \mu(\{0\})$, it holds that

$$\varphi(\mu_n) = \mu_n(\{0\}) = 0 \not\xrightarrow{w} 1 = \delta_0(\{0\}) = \varphi(\delta_0).$$

Proposition 3.1.24. $\sigma(X^*, X^{**}) = \sigma(X^*, X)$ if and only if X is reflexive.

Proof. If X is reflexive, clearly $\sigma(X^*, X^{**}) = \sigma(X^*, X)$.

On the other hand, assume X isn't reflexive and fix $\varphi \in X^{**} \setminus X$. Then φ is continuous on (X^*, w) but not on (X^*, w^*) , so $w \neq w^*$. ■

Theorem 3.1.25 (Banach Alaoglu). For every Banach space X , (\bar{B}_{X^*}, w^*) is compact.

Proof. Let

$$S = \prod_{x \in X} \bar{B}_{\mathbb{F}}(0, \|x\|)$$

with the product topology. We identify this with

$$\{g: X \rightarrow \mathbb{F} \mid \forall x \in X: |g(x)| \leq \|x\|\}.$$

Then

$$\bar{B}_{X^*} \subseteq S$$

and in fact

$$\bar{B}_{X^*} = \{g \in S \mid g \text{ is linear}\}.$$

In fact,

$$(\bar{B}_{X^*}, w^*) \subseteq (S, \text{product})$$

by definition of the product topology. By Tychonoff, S is compact, so we just need to show that \bar{B}_{X^*} is closed in it. For every $x, y \in X$ and $\lambda \in \mathbb{F}$ define

$$\begin{aligned} \varphi_{x,y,\lambda}: S &\rightarrow \mathbb{F} \\ g &\mapsto g(\lambda x + y) - (\lambda g(x) + g(y)). \end{aligned}$$

Then $\varphi_{x,y,\lambda}$ is continuous by definition of the product topology. Then

$$\bar{B}_{X^*} = \{g \in S \mid g \text{ is linear}\} = \bigcap_{x,y,\lambda} \ker \varphi_{x,y,\lambda}$$

is closed, so by the above it's compact. ■

Corollary 3.1.26. $A \subseteq X^*$ is w^* compact if and only if it's w^* -closed and bounded.

Proof. Assume A is w^* -closed bounded, so there's $r > 0$ such that $A \subseteq r \cdot \bar{B}_X$. $r\bar{B}_X$ is compact and A is w^* , hence A is w^* -compact.

For the other direction, assume A is w^* -compact. For every $x \in X$ define

$$\text{ev}_x(A) = \{f(x) \mid f \in A\} \subseteq \mathbb{F},$$

which is compact as the continuous image of a compact set. Hence this is bounded in \mathbb{F} . By the uniform boundedness principle, A is then bounded. ■

Corollary 3.1.27. *If X is reflexive, $A \subseteq X$ is w -compact if and only if it's w -closed and bounded.*

Proof. This is true because when X is reflexive, $(X, w) = ((X^*)^*, w^*)$. ■

Example 3.1.28. Let $X = \ell_\infty$ and define $\varphi \in X^*$ by

$$f_n(a_i)_{i \in \mathbb{N}} = a_n.$$

Then $|f_n(a)| \leq \|a\|$ and specifically $\|f_n\| \leq 1$ so $(f_n)_{n \in \mathbb{N}} \subseteq \bar{B}_{X^*}$.

Assume $f_{n_k} \xrightarrow{w^*} f$ for some subsequence. Define $(a_m)_{m \in \mathbb{N}}$ by

$$a_m = \begin{cases} (-1)^k & m = n_k \\ 17 & \forall k \in \mathbb{N}: m \neq n_k \end{cases}.$$

Then $(a_m)_{m \in \mathbb{N}} \in \ell_\infty$. Note that

$$f_{n_k}(a) = (-1)^k \not\xrightarrow[k \rightarrow \infty]{} f(a),$$

which is a contradiction. Hence no subsequence of $(f_n)_{n \in \mathbb{N}}$ is convergent. This isn't a contradiction to Banach-Alaoglu since in (X^*, w^*) compactness doesn't imply sequential compactness.

In separable spaces however, compactness and sequential compactness are equivalent, which is more comfortable.

Theorem 3.1.29. *Assume X is separable. Then (\bar{B}_{X^*}, w^*) is metrisable.*

Proof. Let $\{x_n\}_{n \in \mathbb{N}}$ be a countable dense set in X . Define a metric on X^* by

$$d(f, g) = \sum_{n \in \mathbb{N}} \frac{\min(|f(x_n) - g(x_n)|, 1)}{2^n}.$$

We claim d is indeed a metric, which is a standard proof. We do show that d separates points. Assume $d(f, g) = 0$. Then $|f(x_n) - g(x_n)| = 0$ for all $n \in \mathbb{N}$, so f, g agree on a dense set and hence $f = g$.

We claim that $i: (\bar{B}_{X^*}, w^*) \rightarrow (\bar{B}_{X^*}, d)$ is continuous. Let $g_0 \in \bar{B}_{X^*}$ and $\varepsilon > 0$. Choose N such that $\frac{1}{2^N} < \frac{\varepsilon}{2}$. Define

$$U := \left\{ f \in \bar{B}_{X^*} \mid \forall i \in [N]: |f(x_i) - g_0(x_i)| < \frac{\varepsilon}{2} \right\} \in w^*.$$

We claim $U \subseteq B_d(g_0, \varepsilon)$, which implies continuity. Indeed, for every $f \in U$ we have

$$\forall f \in U: d(f, g) = \sum_{n \in \mathbb{N}} \frac{\min(|f(x_n) - g_0(x_n)|, 1)}{2^n} \leq \sum_{n \in [N]} \frac{|f(x_n) - g_0(x_n)|}{2^n} + \sum_{n=N+1}^{\infty} \frac{1}{2^n} \leq \sum_{n \in [N]} \frac{\varepsilon}{2^{n+1}} + \frac{1}{2^N}$$

Hence i is continuous. But by Banach-Alaoglu (\bar{B}_{X^*}, w^*) is compact, and (\bar{B}_{X^*}, d) is Hausdorff, hence i is in fact a homeomorphism. Hence (\bar{B}_{X^*}, w^*) is metrisable. ■

Corollary 3.1.30. *If X is separable, \bar{B}_{X^*} is sequentially compact, i.e. every bounded sequence $(f_n) \subseteq X^*$ has a w^* -convergent subsequence.*

Exercise 3. If $\dim X = \infty$, (X^*, w^*) is not metrisable.

Example 3.1.31. Let S be a compact metric space and let $X = \mathcal{C}(S)$. Then X is separable, which is clear e.g. for $S = [0, 1]$, but is true in general as we see later. (\bar{B}_{X^*}, w^*) is metrisable and compact. We identify this with the space of Borel measures on S . Let $\mathcal{P}(S)$ denote the space of probability measures on S . We have by definition of the norm that

$$\mathcal{P}(S) \subseteq \bar{B}_{M(S)}$$

where $M(S)$ is the space of measures on S . In fact,

$$\mathcal{P}(S) = \left\{ \mu \in \bar{B}_{M(S)} \mid \forall f \in \mathcal{S}: f \geq 0 \implies \int f \, d\mu \geq 0 \right\},$$

so $\mathcal{P}(S)$ is closed in $\bar{B}_{M(S)}$, so $(\mathcal{P}(S), w^*)$ is compact and metrisable.

Hence, every sequence $(\mu_n)_{n \in \mathbb{N}}$ of probability measures on a compact set S has a w^* -convergent subsequence.

Theorem 3.1.32 (Eberlein-Šmulian). *For every Banach space X , $A \subseteq X$ is w -compact if and only if A is w -sequentially-compact.*

Proof. If $A \subseteq X$ is w -compact, it follows similarly to the proof of the last theorem that A is w -sequentially-compact.

The other direction is much harder but also much less useful. ■

Corollary 3.1.33 (Banach-Saks, strong form). *Let H be a Hilbert space and let $(x_n)_{n \in \mathbb{N}} \subseteq H$ be bounded. There exists a subsequence $(x_{n_k})_{k \in \mathbb{N}}$ converge (in norm).*

Proof. Since $(x_n)_{n \in \mathbb{N}}$ is bounded, it's contained in some ball $r \cdot \bar{B}_H$ which is w^* -compact by Banach-Alaoglu and then w -compact since a Hilbert space is reflexive. Then $r\bar{B}_H$ is r -sequentially-compact, so there's a subsequence $(x_{n_k})_{k \in \mathbb{N}}$ that is convergent to some x . By the first form of Banach-Saks there's a subsequence $(x_{n_{k_\ell}})_{\ell \in \mathbb{N}}$ such that

$$\frac{1}{\ell} \sum_{i \in [\ell]} x_{n_{k_i}} \xrightarrow{\ell \rightarrow \infty} x$$

in norm. ■

3.1.3 Application: Haar Measures

Let (X, d) be a compact metric space and let G be a group that acts isometrically on X .

Theorem 3.1.34. *There exists a Borel probability measure μ on X such that $\mu(A) = \mu(gA)$ for all $g \in G$ and $A \subseteq X$ borel. μ is called the Haar measure.*

Example 3.1.35. 1. \mathbb{R}^n acts on $\mathbb{R}^n/\mathbb{Z}^n$ by translations. The Haar measure here is the Lebesgue measure.

2. $\mathcal{O}(n)$, the group of orthogonal $n \times n$ matrices, acts on $S^{n-1} \subseteq \mathbb{R}^n$. The Haar measure is the uniform measure on S^{n-1} .

Theorem 3.1.36 (Hall's Marriage Theorem). *Let H be a bipartite graph with vertex sets V and W . Assume that for every $A \subseteq V$ it holds that*

$$|n(A)| = |\{w \in W \mid \exists v \in A: w \sim v\}| \geq |A|.$$

Then there's $\varphi: V \rightarrow W$ injective such that $v \sim \varphi(v)$ for all $v \in V$.

Lemma 3.1.37. *Recall that $N_\varepsilon \subseteq X$ is called an ε -net if for every $x \in X$ there's $y \in N_\varepsilon$ such that $d(x, y) \leq \varepsilon$. Then for every $\varepsilon > 0$ there exists a finite ε -net. We say that N_ε is a minimal ε -net if $|N_\varepsilon|$ is minimal. We claim that if $N_\varepsilon, \tilde{N}_\varepsilon$ are minimal ε -nets, there's a bijection $\varphi: N_\varepsilon \rightarrow \tilde{N}_\varepsilon$ such that $d(y, \varphi(y)) \leq 2\varepsilon$.*

Proof. Take a graph H with vertices $N_\varepsilon \sqcup \tilde{N}_\varepsilon$ and edges $y \sim z \iff \bar{B}(y, \varepsilon) \cap \bar{B}(z, \varepsilon) \neq \emptyset$. For every $\Gamma \subseteq N_\varepsilon$ the set $(N_\varepsilon \setminus \Gamma) \cup n(\Gamma)$ is also an ε -net. Indeed, $y \notin \Gamma$ implies $y \in (N_\varepsilon \setminus \Gamma) \cup n(\Gamma)$, and $y \in \Gamma$ implies $z \in (N_\varepsilon \setminus \Gamma) \cup n(\Gamma)$ since $z \sim y$.

Then $|(N_\varepsilon \setminus \Gamma) \cup n(\Gamma)| \geq |N_\varepsilon|$ so $|n(\Gamma)| \geq |\Gamma|$. By Hall's theorem there's then $\varphi: N_\varepsilon \rightarrow \tilde{N}_\varepsilon$ such that

$$\forall y \in N_\varepsilon: y \sim \varphi(y) \implies d(y, \varphi(y)) \leq 2\varepsilon. \quad \blacksquare$$

Proof (3.1.34). For N_ε a minimal ε -net, and $\varepsilon > 0$, define

$$\mu_\varepsilon = \frac{1}{|N_\varepsilon|} \sum_{y \in N_\varepsilon} \delta_y$$

which is the uniform probability measure on N_ε . By Alaoglu / Prokhorov, there's a sequence $\varepsilon_i \rightarrow 0$ such that $\mu_{\varepsilon_i} \xrightarrow{w^*} \mu$. We claim μ is a Haar measure. We show that

$$\forall f \in \mathcal{C}(X) \forall g \in G: \int g(gx) d\mu(x) = \int f(x) d\mu(x),$$

which gives the result.

Note that weak convergence and the definition of μ_ε imply

$$\begin{aligned} \int f(gx) d\mu(x) &= \lim_{i \rightarrow \infty} \int f(gx) d\mu_{\varepsilon_i}(x) \\ &= \lim_{i \rightarrow \infty} \frac{1}{|N_{\varepsilon_i}|} \sum_{y \in N_{\varepsilon_i}} f(gy) \\ &= \lim_{i \rightarrow \infty} \frac{1}{|N_{\varepsilon_i}|} \sum_{y \in gN_{\varepsilon_i}} f(y). \end{aligned}$$

Similarly,

$$\int f(x) d\mu(x) = \lim_{i \rightarrow \infty} \frac{1}{|N_{\varepsilon_i}|} \sum_{y \in N_{\varepsilon_i}} f(y).$$

Because g acts by an isometry, gN_{ε_i} is a minimal ε -net.

f is continuous on a compact set X , so it's uniformly continuous on it. Fix $\eta > 0$, there's $\varepsilon > 0$ such that $d(x, y) < \varepsilon$ implies $|f(x) - f(y)| < \eta$. For large enough i we have $\varepsilon_i < \frac{\varepsilon}{2}$. Then, for every $y \in N_{\varepsilon_i}$ we have

$$d(y, \varphi(y)) \leq 2\varepsilon_i < \varepsilon$$

where $\varphi: N_{\varepsilon_i} \rightarrow gN_{\varepsilon_i}$ is the map from the lemma. Then

$$|f(y) - f(\varphi(y))| < \eta$$

so

$$\left| \frac{1}{|N_{\varepsilon_i}|} \sum_{y \in g(N_{\varepsilon_i})} f(y) - \frac{1}{|N_{\varepsilon_i}|} \sum_{y \in N_{\varepsilon_i}} f(y) \right| \leq \frac{1}{|N_{\varepsilon_i}|} \sum_{y \in N_{\varepsilon_i}} |f(\varphi(y)) - f(y)| < \eta.$$

Then

$$\forall \eta > 0: \left| \int f(gx) d\mu - \int f(x) d\mu \right| < \eta$$

so

$$\int f(gx) d\mu = \int f(x) d\mu$$

so μ is a Haar measure. ■

3.2 Locally-Convex Spaces

Definition 3.2.1 (Locally Convex Space (LCS)). A *locally-convex space (LCS)* X is a vector space over \mathbb{R} with a topology τ such that

1. (X, τ) is Hausdorff.
2. $+: X \times X \rightarrow X$ and $\cdot: \mathbb{R} \times X \rightarrow X$ are continuous.

3. 0 has a local base of convex sets.

Remark 3.2.2. For every $y \in X$, the map $x \mapsto x + y$ is a homeomorphism, so having a local base of convex sets at 0 implies there's a local base of convex sets at every point.

Example 3.2.3. Let $(X, \|\cdot\|)$ be a Banach space with the norm topology. It obviously satisfies the first two conditions to be an LCS. Balls $\{x \mid \|x\| < \varepsilon\}$ are a convex local base at 0, so X is an LCS.

Example 3.2.4. Let X be a Banach space, let $F \subseteq X^*$ which separates points and consider $(X, \sigma(X, F))$. This is an LCS. It's Hausdorff since F separates points. We've seen addition and multiplication by scalars are continuous. We've seen that a local base at 0 is given by sets

$$U_{f_i, 0, \varepsilon} = \{x \in X \mid \forall i \in [m]: |f_i(x)| < \varepsilon\}.$$

These are convex sets.

Example 3.2.5. Recall $X = \mathcal{C}^1([0, 1])$ with the norm

$$\|f\| = \max |f| + \max |f'|$$

is a Banach space, where $f_n \xrightarrow{\|\cdot\|} f$ iff $f_n \rightarrow f$ and $f'_n \rightarrow f'$ uniformly.

We can generalise this and look at $X = \mathcal{C}^k([0, 1])$ and $\|f\| = \sum_{i=0}^k \max |f^{(i)}|$. Here convergence in norm is equivalent to uniform convergence of all the derivatives.

Example 3.2.6. Let $X = \mathcal{C}^\infty([0, 1])$. Set

$$\|f\|_n = \max_{t \in [0, 1]} |f^{(n)}(t)|$$

and take the weakest topology on X such that all $\|\cdot\|_n$ are continuous. A local base at g is given by

$$U_{g, N, \varepsilon} = \{f \in X \mid \forall n \in \{0, \dots, N\}: \|f - g\|_n < \varepsilon\}.$$

X is metrisable with metric

$$d(f, g) = \sum_{n=0}^{\infty} \frac{\min(\|f - g\|_n, 1)}{2^n}.$$

This is the subspace topology from embedding $\mathcal{C}([0, 1])$ in $\mathcal{C}([0, 1])^{\mathbb{N}}$ by taking a function h to its derivatives.

We have $f_n \rightarrow f$ in X iff $f_n^{(i)} \rightarrow f^{(i)}$ uniformly for all i . This is an LCS because the $U_{g, N, \varepsilon}$ define a convex local base.

Remark 3.2.7. For any sequence $(\|\cdot\|_n)_{n \in \mathbb{N}}$ of semi-norms we can do exactly the same as in the above example, with the assumption that if $\|x\|_n = 0$ for all n then $x = 0$. Such an LCS is called a Fréchet space.

Example 3.2.8. Let $D := \{x \in \mathbb{R}^d \mid \|x\|_2 < 1\} \subseteq \mathbb{R}^d$. Take $X = \mathcal{C}(D)$. This is a Fréchet space with $\|f\|_n = \max_{\|t\|_2 < 1 - \frac{1}{n}} |f(t)|$. If $\|f\|_n = 0$ for all n indeed $f = 0$. We have $f_n \rightarrow f$ iff $f_n \rightarrow f$ uniformly on each $\{\|x\|_2 < 1 - \frac{1}{n}\}$ iff $f_n \rightarrow f$ uniformly on all compact subsets $K \subseteq D$.

Theorem 3.2.9. Let X be an LCS, $K \subseteq X$ convex and closed and $a \notin K$. There exists $f: X \rightarrow \mathbb{R}$ continuous and linear such that

$$\sup_{x \in K} f(x) < f(a).$$

Remark 3.2.10. We can also separate open sets, we can separate a close set from a compact set, etc. as in the case of Banach spaces.

Proof. Without loss of generality assume $a = 0$ and let W be an open neighbourhood of 0 disjoint from K . Without loss of generality, by local convexity, we may assume W is convex. Let $U = W \cap (-W)$ which is a symmetric open neighbourhood of 0 disjoint from K . U is also convex, as the intersection of two convex sets.

We show that $0 \in U$ is internal. For $y \in X$, the map $t \mapsto t \cdot y$ from \mathbb{R} to X is continuous. So, $V_y := \{t \in \mathbb{R} \mid ty \in U\}$ is open in \mathbb{R} . We have $0 \in V_y$ and V_y is open, so there's $\varepsilon > 0$ such that $(-\varepsilon, \varepsilon) \subseteq V_y$. Hence $ty \in U$ for all $|t| < \varepsilon$.

We now have two convex disjoint sets U, K where U has an internal point. Hence there's $f: X \rightarrow \mathbb{R}$ linear and non-zero such that

$$\sup_{x \in K} f(x) \leq \inf_{u \in U} f(u).$$

Fix $y \in X$ such that $f(y) < 0$ and $t > 0$ such that $ty \in U$. Then

$$\sup_{x \in K} f(x) \leq \inf_{u \in U} f(u) \leq f(ty) = tf(y) < 0 = f(0).$$

We're left to show f is continuous. Fix $(\alpha, \beta) \subseteq \mathbb{R}$ and $x \in f^{-1}((\alpha, \beta))$. Choose $0 < r < \min(\beta - f(x), f(x) - \alpha)$. Denote $-m = \sup_{x \in K} f(x)$, where $m > 0$. For every $y \in x + \frac{r}{m}U$ we have

$$\begin{aligned} f(y) &= f\left(x + \frac{r}{m}U\right) \\ &= f(x) + \frac{r}{m}f(u) \\ &\geq f(x) - \frac{r}{m}(-m) \\ &= f(x) - r \\ &> f(x) - (f(x) - \alpha) \\ &= \alpha. \end{aligned}$$

Similarly $f(y) < \beta$: Since U is symmetric we have $-u \in U$ so

$$\begin{aligned} f(y) &= f(x) - \frac{r}{m} f(-u) \\ &\leq f(x) - \frac{r}{m} (-m) \\ &= f(x) + r \\ &< f(x) + (\beta - f(x)) \\ &= \beta. \end{aligned}$$

Then

$$x + \frac{r}{m}U \subseteq f^{-1}((\alpha, \beta))$$

so

$$f^{-1}((\alpha, \beta))$$

is open, hence f is continuous. ■

Remark 3.2.11. We actually showed that $f: X \rightarrow \mathbb{R}$ is linear and bounded from below on an open set, it's continuous.

Example 3.2.12. Let $p \in (0, 1)$ and let $X = L^p([0, 1])$. Define $\|f\| = \|f\|_p^p = \int_0^1 |f|^p$. This is *not* a norm since $\|\lambda f\| \neq |\lambda| \|f\|$. But, $\|f + g\| \leq \|f\| + \|g\|$ since $(a + b)^p < a^p + b^p$ for $p \in (0, 1)$. Then $d(f, g) = \|f - g\|$. Then (X, d) is a complete metric space, but it's not locally convex since $X^* = \{0\}$ as we now show.

Assume towards a contradiction that $0 \neq \Phi \in X^*$. Choose $f \in X$ such that $a := \Phi(f) > 0$, and write $b := \|f\|$. By the mean value theorem, there's t_0 such that $\int_0^{t_0} |f|^p = \frac{b}{2}$.

Take $g = 2f\chi_{[0, t_0]}$ and $h = 2f\chi_{[t_0, 1]}$. Then

$$a = \Phi(f) = \Phi\left(\frac{g+h}{2}\right) = \frac{\Phi(g) + \Phi(h)}{2}.$$

Then $\Phi(g) \geq a$ or $\Phi(h) \geq a$. Also,

$$\|g\| = \int_0^1 |g|^p = \int_0^{t_0} |2f|^p = 2^p \cdot \frac{b}{2} = \frac{b}{2^{1-p}}$$

and similarly $\|h\| = \frac{b}{2^{1-p}} = \|g\|$. Hence we can choose f_1 to be either g or h (take g if $\Phi(g) \geq a$ and otherwise take h) such that $\Phi(f_1) \geq a$ and $\|f_1\| = \frac{b}{2^{1-p}}$.

Repeating this, we get a sequence $(f_n)_{n \in \mathbb{N}_+} \subseteq X$ such that $\Phi(f_n) \geq a$ and $\|f_n\| = \frac{b}{(2^{1-p})^n} \xrightarrow{n \rightarrow \infty} 0$. Since $\Phi(f_n) \not\rightarrow 0$ we get that Φ is not continuous.

Definition 3.2.13 (Extreme Point). Let X be a vector space and let $K \subseteq X$ be convex. A point $a \in K$ is an *extreme point* of K if

$$\forall \lambda \in (0, 1) \forall y, z \in K: a = (1 - \lambda)y + \lambda z \implies y = z = a.$$

Example 3.2.14. Let H be a Hilbert space and let $K = \bar{B}_H$. Then $\text{Ext}(K)$, the set of extreme points of K , is the set $\{x \in H \mid \|x\| = 1\}$.

Assume $\|x\| = 1$, and write $x = (1 - \lambda)y + \lambda z$ where $\lambda \in (0, 1)$ and $y, z \in K$. Then

$$1 = \|x\| = \|(1 - \lambda)y + \lambda z\| \leq (1 - \lambda)\|y\| + \lambda\|z\| \leq (1 - \lambda) + \lambda = 1$$

so there are equalities, so there's equality in the triangle inequality, which implies x, y, z are colinear, but $\|y\| = \|z\|$ since there's equality in the above inequality, hence $y = z = x$.

For the other direction, assume $\|x\| < 1$ and fix $y \perp x$ non-zero. Write

$$x = \frac{(x + ty) + (x - ty)}{2},$$

and $x \pm ty \in K$ for t small enough.

Example 3.2.15. Let $X = c_0$ and $K = \bar{B}_{c_0}$. Assume $a \in K$. Then $a_n \xrightarrow{n \rightarrow \infty} 0$ so there's $m \in \mathbb{N}$ such that $|a_m| < \frac{1}{2}$. Now write

$$a = \frac{(a + \frac{1}{2}e_m) + (a - \frac{1}{2}e_m)}{2}$$

where $a \pm \frac{1}{2}e_m \in K$ by assumption on a_m . Hence a is not an extreme point, so $\text{Ext}(K) = \emptyset$.

Corollary 3.2.16. c, c_0 are not isometric.

Proof. If $T: c \rightarrow c_0$ is an isometry then $T(\bar{B}_c) = \bar{B}_{c_0}$ and so $T(\text{Ext}(\bar{B}_c)) = \text{Ext}(\bar{B}_{c_0})$. But, $\text{Ext}(\bar{B}_{c_0}) = \emptyset$ and $(1)_{n \in \mathbb{N}} \in \text{Ext}(\bar{B}_c)$. ■

Remark 3.2.17. c, c_0 are isomorphic. Define

$$\begin{aligned} T: c_0 &\rightarrow c \\ (a_i)_{i \in \mathbb{N}} &\mapsto (a_0 + a_{i+1})_{i \in \mathbb{N}}. \end{aligned}$$

This is an isomorphism.

Theorem 3.2.18 (Krein-Milman). If X is an LCS and $K \subseteq X$ is compact and convex, then $K = \overline{\text{conv}(\text{Ext}(K))}$.

Corollary 3.2.19. There is no space X such that $X^* = c_0$ (isometrically).

Proof. If $X^* = c_0$, we can look at $\bar{B}_{c_0} \subseteq (c_0, w^*)$ which is convex and compact (by Banach-Alaoglu). By Krein-Milman, $\bar{B}_{c_0} = \overline{\text{conv}(\text{Ext}(\bar{B}_{c_0}))} = \emptyset$, a contradiction. ■

Definition 3.2.20 (Extremal Set). Let K be a convex subset of a vector space X . A subset $E \subseteq K$ is called an *extremal set of K* if it's convex and

$$\forall \lambda \in (0, 1) \forall y, z \in K: (1 - \lambda)y + \lambda z \in E \implies y, z \in E.$$

Lemma 3.2.21. *In an LCS X , if K is convex, \bar{K} is also convex.*

Lemma 3.2.22. *Let $K \subseteq X$ be a compact and convex subset of an LCS. Then for every extremal set $E \subseteq K$ it holds that $E \cap \text{Ext}(K) \neq \emptyset$.*

Proof. Define

$$Z := \{A \in \mathcal{P}(E) \setminus \{\emptyset\} \mid A \text{ is closed and extremal}\}.$$

Then $Z \neq \emptyset$ since $E \in Z$. If $(A_\alpha)_\alpha \subseteq Z$ is a chain in Z then $A = \bigcap_\alpha A_\alpha$ is closed and extremal. We show it's also non-empty. We have $A_{\alpha_1} \cap \dots \cap A_{\alpha_k}$ for every k values of α , since $(A_\alpha)_\alpha$ is a chain.

By Zorn's lemma there's $A \in Z$ minimal. Assume $x_0, y_0 \in A$ are different. There's $f \in X^*$ such that $f(x_0) \neq f(y_0)$. Write $m := \max_A f$ and $B := \{x \in A \mid f(x) = m\}$. Then B is closed and convex (by linearity), and extremal: Assume $y, z \in K$ and $\lambda \in (0, 1)$. Then $(1 - \lambda)y + \lambda z \in B \subseteq A$ and since A is extremal we have $y, z \in A$. But,

$$m = f((1 - \lambda)y + \lambda z) = (1 - \lambda)f(y) + \lambda f(z) \leq (1 - \lambda)m + \lambda m = m$$

so there's in fact equality so $f(y) = f(z) = m$, so $y, z \in B$. By minimality of A we have $B = A$. Then $f|_A \equiv m$, a contradiction to $f(x_0) \neq f(y_0)$. Then $A = \{x_0\}$ for some x_0 , which is then $x_0 \in E \cap \text{Ext}(K)$. ■

Proof (3.2.18). Let $\hat{K} = \overline{(\text{Ext}(K))}$. This is closed and is convex as the closure of a convex set. Also $\hat{K} \subseteq K$ because K itself is convex and closed.

Assume towards a contradiction that $a \in K \setminus \hat{K}$ and let $f \in X^*$ such that $\max_{x \in \hat{K}} f(x) < f(a)$. Let $m := \max_K f$ and $E := \{x \in K \mid f(x) = m\}$ where the latter is a closed extremal set as in the lemma. By the lemma, there's $p \in \text{Ext}(K) \cap E$, so

$$\begin{aligned} m &= f(p) \\ &\leq \sup_{x \in \text{Ext}(K)} f(x) \\ &\leq \sup_{x \in \hat{K}} f(x) \\ &< f(a) \\ &\leq m, \end{aligned}$$

from which $m < m$, a contradiction. Hence $\hat{K} = K$. ■

Example 3.2.23. Let S be a compact metric space and let $P(S)$ be a borel probability measure on S . Discrete measures (finitely-supported) are w^* -dense in $P(S)$.

Proof. $P(S)$ is convex, and $\text{Ext}(P(S)) = \{\delta_x \mid x \in S\}$ which we show soon. Then $P(S)$ is w^* -compact by the Banach-Alaoglu theorem. Then by Krein-Milman

$$P(S) = \overline{\text{Conv}\{\delta_x \mid x \in S\}}^{w^*}.$$

Assume $\mu \in \text{Ext}(P(S))$. If μ is not a δ -measure there's $A \subseteq S$ borel such that $\mu(A) \in (0, 1)$. Define a measure ν by

$$\nu(B) = \mu(B \cap A)(1-t) - \mu(B \cap A^C)t.$$

Note that $\nu(S) = t(1-t) - (1-t)t = 0$, but $\nu(A) = t(1-t) \neq 0$ so $\nu \neq 0$. We have

$$\mu = \frac{(\mu + \nu) + (\mu - \nu)}{2}.$$

We have $(\mu + \nu)(S) = (\mu - \nu)(S) = 1$. We want to show $\mu \pm \nu \geq 0$ from which we get that μ is not an extreme point. Indeed,

$$\begin{aligned} (\mu + \nu)(B) &= \mu(B \cap A) + \mu(B \cap A^C) + \mu(B \cap A)(1-t) - \mu(B \cap A^C)t \\ &= \mu(B \cap A)(2-t) + \mu(B \cap A^C)(1-t) \geq 0 \end{aligned}$$

and similarly for $\mu - \nu$. Then $\mu \pm \nu \in P(S)$ and we get the result. \blacksquare

Example 3.2.24. A theorem by Birkhoff says the following.

Let

$$D = \left\{ A \in \mathbb{C}^{n \times n} \mid \begin{array}{l} \forall i, j \in [n]: a_{i,j} \geq 0 \\ \sum_j a_{i,j} = 1 \\ \sum_i a_{i,j} = 1 \end{array} \right\}$$

which we call *doubly-stochastic matrices* and let \mathbb{T} be $n \times n$ permutation matrices. Then $D = \text{Conv}(\mathbb{T})$.

Proof. We should show that $\text{Ext}(D) = \mathbb{T}$, then by Krein-Milman

$$D = \overline{\text{Conv}(\text{Ext}(D))} = \overline{\text{Conv}(\mathbb{T})}.$$

Since \mathbb{T} is finite, $\text{Conv}(\mathbb{T})$ is compact and so is closed. Then the above equation gives $D = \text{Conv}(\mathbb{T})$.

We now show that indeed $\text{Ext}(D) = \mathbb{T}$. Define $I := \{(i, j) \in [n]^2 \mid a_{i,j} \in (0, 1)\}$. We want to show that $I \neq \emptyset$ implies $A \notin \text{Ext}(D)$. We can pick some (i, j) , then choose another (i, j') in I by definition of I . We can then pick another pair of indices with the same value of j' , and continue by alternately changing the column and row numbers. This eventually closes a cycle of even length. Take B by adding $(-1)^k$ in the k^{th} . Then $A = \frac{(A+\varepsilon B) + (A-\varepsilon B)}{2}$ is in D for small enough ε . \blacksquare

Chapter 4

Banach Algebras

4.1 Banach Algebras

4.1.1 Definition & Examples

Definition 4.1.1 (Banach Algebra). A *Banach Algebra* A is a Banach space with a bilinear map $\cdot : A \times A \rightarrow A$ such that the following hold.

1. \cdot is associative.
2. There's $e \in A$ such that $e \cdot x = x \cdot e = x$ for all $x \in A$.
3. $\|x \cdot y\| \leq \|x\| \|y\|$ and $\|e\| = 1$.

Example 4.1.2. Let $A = \mathcal{C}(S)$ where S is a compact and Hausdorff space and with the supremum norm. Take on it the product $(f \cdot g)(s) = f(s)g(s)$. Here $e(s) \equiv 1$ gives $\|e\| = 1$. We indeed have

$$\begin{aligned}\|fg\| &= \max_{s \in S} (|f(s)| |g(s)|) \\ &\leq \max_{s \in S} |f(s)| \cdot \max_{s \in S} |g(s)| \\ &= \|f\| \|g\| .\end{aligned}$$

Example 4.1.3. Let X be a Banach space and $A := \mathcal{L}(X, X) = \mathcal{L}(X)$ with the operator norm. $T \cdot S$ is the composition and $e = \text{id}$ is the unit. In particular if $A = M_{n \times n}(\mathbb{C})$ is such an example, where X is a finite-dimensional vector space.

Remark 4.1.4. A Banach algebra can be defined over \mathbb{R} , although we will soon deal with spectral theory where work over an algebraically-closed field is much easier. We therefore treat only Banach Algebras over \mathbb{C} .

Example 4.1.5. Let $A := \ell^1(\mathbb{Z})$. Take the product to be the convolution

$$(a * b)_n := \sum_{k \in \mathbb{Z}} a_k b_{n-k}.$$

With $e = e_0$ this is indeed a Banach algebra. Indeed

$$\begin{aligned} \|a * b\|_1 &= \sum_{n \in \mathbb{Z}} \left| \sum_{k \in \mathbb{Z}} a_k b_{n-k} \right| \\ &\leq \sum_{n \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} |a_k| |b_{n-k}| \\ &= \sum_{k \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} |a_k| |b_{n-k}| \\ &= \sum_{k \in \mathbb{Z}} |a_k| \sum_{n \in \mathbb{Z}} |b_{n-k}| \\ &= \sum_{k \in \mathbb{Z}} |a_k| \|b\| \\ &= \|a\| \|b\|. \end{aligned}$$

Example 4.1.6 (The Wiener Algebra). Let

$$W := \left\{ f \in \mathcal{C}(\mathbb{T}) \mid \begin{array}{l} f(t) = \sum_{n \in \mathbb{Z}} \hat{f}(n) e^{int} \\ \sum_{n \in \mathbb{Z}} |\hat{f}(n)| < \infty \end{array} \right\}$$

with norm

$$\|f\| = \sum_{n \in \mathbb{Z}} |\hat{f}(n)|,$$

the point-wise product and $e \equiv 1$ as the identity.

Note that

$$\begin{aligned} T: W &\rightarrow \ell^1(\mathbb{Z}) \\ f &\mapsto \left(\hat{f}(n) \right)_{n \in \mathbb{Z}} \end{aligned}$$

is an isometric isomorphism of Banach Algebras since

$$\left(\widehat{fg}(n) \right)_{n \in \mathbb{Z}} = \left(\hat{f}(n) \right)_{n \in \mathbb{Z}} * \left(\hat{g}(n) \right)_{n \in \mathbb{Z}}.$$

Definition 4.1.7 (Invertible Element). $x \in A$ is *invertible* if there's $x^{-1} \in A$ such that $x = x^{-1} = x^{-1}x = ex$.

Fact 4.1.8. 1. If z_1, z_2 are such that $z_1x = xz_2 = e$ then $z_1 = z_2 = x^{-1}$.

2. If x, y are invertible, so is xy and $(xy)^{-1} = y^{-1}x^{-1}$.

3. If xy, yx are invertible, so are x, y . E.g. write $xy(xy)^{-1} = e$ and $(yx)^{-1}yx = e$ so x has inverses from both sides, and by the first property has an inverse.

Proposition 4.1.9. 1. The set \mathcal{O} of invertible elements is open. In fact for every $x \in \mathcal{O}$ and every y such that $\|y\| < \frac{1}{\|x^{-1}\|}$ we have $x \cdot y \in \mathcal{O}$.

2. The map $()^{-1} : \mathcal{O} \rightarrow \mathcal{O}$ is continuous.

Proof. 1. Assume first that $x = e$. Note that

$$\sum_{n \in \mathbb{N}} \|y^n\| \leq \sum_{n \in \mathbb{N}} \|y\|^n < \infty$$

since $\|y\| < 1$. By compactness, $z = \sum_{n \in \mathbb{N}} y^n = \lim_{N \rightarrow \infty} \sum_{n=0}^N y^n$ exists. We have

$$\begin{aligned} z(e - y) &= \lim_{N \rightarrow \infty} \left(e \sum_{n=0}^N y^n - y \sum_{n=0}^N y^n \right) \\ &= \lim_{N \rightarrow \infty} \left(\sum_{n=0}^N y^n - \sum_{n=1}^{N+1} y^n \right) \\ &= \lim_{N \rightarrow \infty} (e - y^{N+1}) \\ &= e. \end{aligned}$$

Similarly $(e - y)z = e$. Generally, $x - y = x(e - x^{-1}y) \in \mathcal{O}$.

2. Fix $x \in \mathcal{O}$ and let y such that $\|y\| \leq \frac{1}{\|x^{-1}\|}$. We have

$$\begin{aligned} \|x^{-1} - (x - y)^{-1}\| &= \|x^{-1} - x^{-1}(e - yx^{-1})^{-1}\| \\ &\leq \|x^{-1}\| \|e - (e - yx^{-1})^{-1}\| \\ &= \|x^{-1}\| \left\| e - \sum_{n \in \mathbb{N}} (yx^{-1})^n \right\| \\ &= \|x^{-1}\| \left\| \sum_{n \in \mathbb{N}_+} (yx^{-1})^n \right\| \\ &\leq \|x^{-1}\| \sum_{n \in \mathbb{N}_+} (\|y\| \|x^{-1}\|)^n \\ &= \|x^{-1}\| \frac{\|y\| \|x^{-1}\|}{1 - \|y\| \|x^{-1}\|} \\ &\xrightarrow{y \rightarrow 0} 0. \end{aligned}$$

■

Definition 4.1.10 (Regular Value). $\lambda \in \mathbb{C}$ is a *regular value* for $x \in A$ if $x - \lambda e \in \mathcal{O}$.

Definition 4.1.11 (Resolvent Set). The set of regular values $\rho(x)$ is called the *resolvent set*.

Definition 4.1.12 (Spectrum). The set $\sigma(x) = \mathbb{C} \setminus \rho(x)$ is called the *spectrum* of X .

Example 4.1.13. If $A = \mathcal{L}(X)$ and $T \in A$, the set $\sigma(T)$ is the spectrum in the classical sense.

In the classical setting we defined the point spectrum

$$\sigma_p(T) := \{\lambda \mid T - \lambda I \text{ is not injective}\} = \{\text{eigenvalues}\}.$$

We defined also the continuous spectrum

$$\sigma_c := \left\{ \lambda \notin \sigma_p(T) \mid \overline{\text{Im}(T - \lambda I)} = X \right\}$$

and the residual spectrum

$$\sigma_r(T) := \left\{ \lambda \notin \sigma_p(T) \mid \overline{\text{Im}(T - \lambda I)} \neq X \right\}.$$

We have

$$\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$$

where the equality encapsulates the fact that if $S \in \mathcal{L}(X)$ is bijective then $S^{-1} \in \mathcal{L}(X)$.

Example 4.1.14. Let $A = \mathcal{C}(S)$. If $f(s) \neq 0$ for all s , we have $\frac{1}{f} \in A$. Here

$$\begin{aligned} \sigma(f) &= \{\lambda \mid (f - \lambda) \text{ is not invertible}\} \\ &= \{\lambda \mid \exists s: (f - \lambda)(s) = 0\} \\ &= \{\lambda \mid \exists s \in S: f(s) = \lambda\} \\ &= \text{Im } f. \end{aligned}$$

4.1.2 Banach-Valued Analytic Functions

Definition 4.1.15 (Analytic Function). Let $D \subseteq \mathbb{C}$ open and X a complex Banach space. $\varphi: D \rightarrow X$ is *analytic* if for every $\lambda \in D$ the limit

$$\lim_{h \rightarrow 0} \frac{\varphi(\lambda + h) - \varphi(\lambda)}{h}$$

exists.

Theorem 4.1.16. $\varphi: D \rightarrow X$ is analytic if and only if for every $f \in X^*$ the function $f \circ \varphi: D \rightarrow \mathbb{C}$ is analytic.

Proof. We show only the implication. Assume $\varphi: D \rightarrow X$ is analytic. We have

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{f(\varphi(\lambda + h)) - f(\varphi(\lambda))}{h} &= \lim_{h \rightarrow 0} f\left(\frac{\varphi(\lambda + h) - \varphi(\lambda)}{h}\right) \\ &= f(\varphi'(\lambda)) \end{aligned}$$

where the last equality is by continuity of f , hence the limit exists so $f \circ \varphi$ is analytic. ■

Remark 4.1.17. If $\psi: [a, b] \rightarrow X$ is continuous, we can define the

$$\int_a^b \psi(t) dt$$

using Riemann sums. In particular, if $\Gamma \in \mathbb{C}$ is a smooth curve with parametrisation $z: [a, b] \rightarrow \Gamma$, we can define

$$\int_{\Gamma} \varphi dz = \int_a^b \varphi(z(t)) z'(t) dt.$$

Proposition 4.1.18. Let $D \subseteq \mathbb{C}$ be simply connected and let $\Gamma = \partial D$. Let \tilde{D} be a domain containing \bar{D} and let $\varphi: \tilde{D} \rightarrow \mathbb{C}$ analytic in a neighbourhood of \bar{D} . Then for every $\lambda \in D$ it holds that

$$\varphi(\lambda) = \frac{1}{2\pi i} \int_{\Gamma} \frac{\varphi(\eta)}{\eta - \lambda} d\eta.$$

Proof. For every $f \in X^*$ we have that $f \circ \varphi$ is analytic (in the classical sense), so by Cauchy's formula

$$f(\varphi(\lambda)) = \frac{1}{2\pi i} \int_{\Gamma} \frac{f(\varphi(\eta))}{\eta - \lambda} d\eta \stackrel{\text{linearity}}{=} \frac{1}{2\pi i} \int_{\Gamma} f\left(\frac{\varphi(\eta)}{\eta - \lambda}\right) d\eta \stackrel{\text{linearity}}{=} f\left(\frac{1}{2\pi i} \int_{\Gamma} \frac{\varphi(\eta)}{\eta - \lambda} d\eta\right).$$

We are done since X^* separates points. ■

Corollary 4.1.19 (Generalised Cauchy Formula). Let $\varphi: D \rightarrow X$ be analytic. Then φ is infinitely-differentiable and

$$\varphi^{(n)}(\lambda) = \frac{n!}{2\pi i} \int_{\Gamma} \frac{\varphi(\eta)}{(\eta - \lambda)^{n+1}} d\eta.$$

Proof. Differentiate Cauchy's formula n times. ■

Corollary 4.1.20 (Taylor Series). Let $\varphi: D \rightarrow X$ be analytic in a neighbourhood of $\lambda \in D$. Let d be the distance from λ to the nearest singular point of φ . Then if $|\mu - \lambda| < d$ we have

$$\varphi(\mu) = \sum_{n \in \mathbb{N}} \frac{\varphi^{(n)}(\lambda)}{n!} (\mu - \lambda)^n.$$

In particular

$$d = R := \left(\limsup_{n \rightarrow \infty} \left\| \frac{\varphi^{(n)}(\lambda)}{n!} \right\|^{\frac{1}{n}} \right)^{-1}.$$

Proof. Expand the right-hand side of Cauchy's formula to a power series. ■

Definition 4.1.21 (Spectral Radius). If $x \in A$, the spectral radius of A is

$$r(x) := \lim_{n \rightarrow \infty} \|x^n\|^{\frac{1}{n}}.$$

Lemma 4.1.22 (Fekete's Lemma). If $(a_n)_{n \in \mathbb{N}_+} \subseteq \mathbb{R}$ and $a_{n+m} \leq a_n + a_m$, the limit

$$\lim_{n \rightarrow \infty} \frac{a_n}{n} = \inf_{n \in \mathbb{N}_+} \frac{a_n}{n}.$$

Proof. Write $L := \inf_{n \in \mathbb{N}_+} \frac{a_n}{n}$. Fix $\varepsilon > 0$ and choose m such that $\frac{a_m}{m} \leq L + \varepsilon$. For every $n \in \mathbb{N}_+$ we can write $n = qm + r$ where $r \in \{0, \dots, m-1\}$. Then

$$\frac{a_n}{n} = \frac{a_{m+m+\dots+m+r}}{n} \leq \frac{qa_m + a_r}{n} \leq \frac{qm(L + \varepsilon) + a_r}{n} \leq L + \varepsilon + \frac{a_r}{n} \leq L + \varepsilon + \frac{\max_{a_0, \dots, a_{m-1}}}{n} \xrightarrow{n \rightarrow \infty} L + \varepsilon.$$

Hence

$$\limsup_{n \rightarrow \infty} \frac{a_n}{n} \leq L$$

so $\frac{a_n}{n} \xrightarrow{n \rightarrow \infty} L$. ■

Remark 4.1.23. L could be $-\infty$.

Corollary 4.1.24. $r(x) = \lim_{n \rightarrow \infty} \|x^n\|^{\frac{1}{n}}$ exists and $r(x) \in (0, \|x\|)$.

Proof. Take $a_n = \log \|x^n\|$ in Fekete's lemma. ■

Theorem 4.1.25 (Gelfand). Let A be a Banach algebra and $x \in A$. Then

1. $\sigma(x) \subseteq \mathbb{C}$ is compact and non-empty.
2. (Gelfand's Formula:) $\max\{|\lambda| \mid \lambda \in \sigma(x)\} = r(x)$.

Proof. 1. We have

$$\sigma(x) = \{\lambda \mid x - \lambda e \neq \mathcal{O}\}$$

and \mathcal{O} is open, hence $\sigma(x)$ is closed. We want to show it's also non-empty. Define $\varphi: \rho(x) \rightarrow X$ by $\varphi(\lambda) = (x - \lambda e)^{-1}$. We claim φ is analytic: We have

$$\begin{aligned} \frac{\varphi(\lambda + h) - \varphi(\lambda)}{h} &= \frac{(x - \lambda e - he)^{-1} - (x - \lambda e)^{-1}}{h} \\ &= (x - \lambda e - he)^{-1} \frac{(x - \lambda e) - (x - \lambda e - he)}{h} (x - \lambda e)^{-1} \\ &= (x - \lambda e - he)^{-1} (x - \lambda e)^{-1} \xrightarrow{h \rightarrow 0} (x - \lambda e)^{-2}. \end{aligned}$$

Note that

$$\begin{aligned}\lim_{|\lambda| \rightarrow \infty} \varphi(\lambda) &= \lim_{|\lambda| \rightarrow \infty} (x - \lambda e)^{-1} \\ &= - \lim_{|\lambda| \rightarrow \infty} \left(\frac{1}{\lambda} \left(e - \frac{x}{\lambda} \right)^{-1} \right) \\ &= 0.\end{aligned}$$

If $\sigma(x) = \emptyset$ we have $\rho(x) = \mathbb{C}$ so $\varphi: \mathbb{C} \rightarrow X$ is a bounded entire function. Hence for every $f \in X^*$, $f \circ \varphi$ is bounded and entire, so $f \circ \varphi$ is constant for every $f \in X^*$. Hence φ is constant (because X^* separates points). However, φ isn't constant, hence a contradiction.

Boundedness follows from the second part.

2. Consider the power series $\sum_{n \in \mathbb{N}} x^n \mu^n$. We have

$$\begin{aligned}\sum_{n \in \mathbb{N}} x^n \mu^n &= (e - \mu x)^{-1} \\ &= -\frac{1}{\mu} \left(x - \frac{1}{\mu e} \right)^{-1}.\end{aligned}$$

We find the radius of convergence for this in two ways. On one hand this is

$$R = \left(\limsup_{n \rightarrow \infty} \|x^n\|^{\frac{1}{n}} \right)^{-1} = \frac{1}{r(x)}.$$

On the other hand,

$$\begin{aligned}R &= \min \left\{ |\mu - 0| \mid \mu \text{ is a singular point of } (e - \mu x)^{-1} \right\} \\ &= \min \left\{ |\mu| \mid \frac{1}{\mu} \notin \rho(x) \right\} \\ &= \frac{1}{\max \{ |\lambda| \mid \lambda \in \sigma(x) \}}.\end{aligned}$$

By comparing these values of R we get the result. ■

Corollary 4.1.26 (Banach-Mazur Theorem). *Let A be a Banach algebra where every $x \neq 0$ is invertible. Then $A \cong \mathbb{C}$.*

Proof. Fix $x \in A$ and $\lambda \in \sigma(x) \neq \emptyset$. Then $x - \lambda e$ is not invertible. Hence $x - \lambda e = 0$ so $x = \lambda e \cong \mathbb{C}$ by $\lambda e \leftrightarrow \lambda$. ■

Remark 4.1.27. In the real case, the dimension of a Banach algebra with inverses is either 1, 2 or 4, in which case it's either \mathbb{R}, \mathbb{C} or the quaternions.

Example 4.1.28 (The Volterra Operator). Let $A = \mathcal{L}(\mathcal{C}[0, 1])$. Take $V \in A$ to be

$$(Vf)(x) = \int_0^x f(t) dt.$$

We want to find $\sigma(V)$. For this we use Gelfand's formula.

Assume f is such that $\|f\| = 1$. We have

$$|(Vf)(x)| = \left| \int_0^x f(t) dt \right| \leq \int_0^x |f(t)| dt \leq \int_0^x 1 dt = x.$$

We then have

$$|(V^2 f)(x)| \leq \int_0^x |Vf(t)| dt \leq \int_0^x t dt = \frac{x^2}{2}$$

and similarly

$$|V^n(f)(x)| \leq \frac{x^n}{n!}.$$

Hence $\|V^n f\| \leq \frac{1}{n!}$. Hence $\|V^n\| \leq \frac{1}{n!}$ (in fact, there's equality as one can check by starting with the constant function 1). So,

$$r(V) = \lim_{n \rightarrow \infty} \|V^n\|^{\frac{1}{n}} = \lim_{n \rightarrow \infty} \left(\frac{1}{n!} \right)^{\frac{1}{n}} = 0.$$

So $\sigma(V) = 0 \{0\}$.

Definition 4.1.29 (Quasi-Nilpotent Element). An element $V \in A$ such that $\sigma(V) = \{0\}$ is called *quasi-nilpotent*.

Remark 4.1.30. The Volterra operator V is quasi-nilpotent but not nilpotent.

4.1.3 Ideals

Definition 4.1.31 (Ideal). $I \leq A$ is an ideal if it's a linear subspace and for every $x \in A$ we have $xI, Ix \subseteq I$.

From now on, assume A is a commutative Banach algebra.

Fact 4.1.32. $x \in A$ is not invertible if and only if there exists a proper ideal $I \leq A$ such that $x \in I$.

Proposition 4.1.33. If $I \leq A$ is a proper ideal, so is \bar{I} .

Proof. \bar{I} is a subspace. If $x \in \bar{I}$ and $y \in A$, there's $(x_n)_{n \in \mathbb{N}} \subseteq I$ such that $x_n \xrightarrow{n \rightarrow \infty} x$. Then $x_n y \xrightarrow{n \rightarrow \infty} xy$ but $x_n y \in I$ for all $n \in \mathbb{N}$, so $xy \in \bar{I}$.

Since I is proper, $I \cap \mathcal{O} = \emptyset$. Since \mathcal{O} is open we get $\bar{I} \cap \mathcal{O} = \emptyset$ so \bar{I} is proper. ■

Definition 4.1.34 (Maximal Ideal). $M \leq X$ is called a *maximal ideal* if M is a proper ideal which is maximal with respect to inclusion.

Proposition 4.1.35. *A maximal ideal is closed.*

Proof. Otherwise, take $I = \bar{M}$ which is a proper ideal. ■

Fact 4.1.36. *Every ideal is contained in a maximal ideal.*

Proof. Use Zorn's lemma. ■

Example 4.1.37. Let $A = \mathcal{C}(S)$. We want to find the maximal ideals in A .

For $s \in S$, let $M_s := \{f \in A \mid f(s) = 0\}$. Here M_s is a maximal ideal, e.g. since $\text{codim } M_s = 1$. We see that by defining $\text{ev}_s(f) = f(s)$ and since

$$A/M_s = A/\ker \text{ev}_s \cong \text{Im ev}_s \cong \mathbb{C}.$$

Assume M is a maximal ideal, we want to show it's of the form M_s for some $s \in S$. Assume towards a contradiction this isn't the case. Then $M \not\subseteq M_s$ for all $s \in S$ so there's $f_s \in M$ such that $f_s(s) \neq 0$. There's U_s an open neighbourhood of s such that $f_s \neq 0$ on U_s . By compactness, we have finitely-many $(U_{s_i})_{i \in [n]}$. Let

$$g = \sum_{i \in [n]} f_{s_i}^2 = \sum_{i \in [n]} f_{s_i} \overline{f_{s_i}},$$

which is in M since all $f_{s_i}, \overline{f_{s_i}}$ are. But, $g > 0$ so g is invertible, so $M = A$, a contradiction.

Fact 4.1.38. 1. *If A is an algebra and $I \leq A$, there's a quotient algebra A/I .*

2. *Every ideal $K \leq A/I$ is of the form $K = J/I$ where $J \leq A$ contains I .*

Fact 4.1.39. *Let A be a commutative Banach algebra and $I \leq A$ a proper closed ideal. We know A/I is a Banach space with $\|[x]\| = \inf_{z \in I} \|x + z\|$, and an algebra with the product $[x] \cdot [y]$.*

Proposition 4.1.40. *A/I is a Banach algebra.*

Proof. Fix $[x], [y] \in A/I$ and $\varepsilon > 0$. Choose $z_1, z_2 \in I$ such that $\|[x]\| > \|x + z_1\| - \varepsilon$ and $\|[y]\| \geq \|y + z_2\| - \varepsilon$. Then

$$\begin{aligned} \|[x][y]\| &= \|[x + z_1][y + z_2]\| \\ &= \|[(x + z_1)(y + z_2)]\| \\ &\leq \|(x + z_1)(y + z_2)\| \\ &\leq \|x + z_1\| \cdot \|y + z_2\| \\ &\leq (\|[x]\| + \varepsilon)(\|[y]\| + \varepsilon), \end{aligned}$$

and taking $\varepsilon \rightarrow 0$ we get $\|[x][y]\| \leq \|[x]\| \|[y]\|$.

We should now show that $\|[e]\| = 1$. Clearly $\|[e]\| \leq \|e\| = 1$. Assume towards a contradiction that also $\|[e]\| < 1$. Then there's $z \in I$ such that $\|e + z\| < 1$, but then $-z = e - (e + z) \in I$ is invertible, so $I = A$, a contradiction. ■

Proposition 4.1.41. *If $M \leq A$ is a maximal ideal, then $\text{codim } M = 1$, i.e. $A/M \cong \mathbb{C}$.*

Proof. Let $[x] \in A/M$ be non-invertible. Then $[x] \in K \subsetneq A/M$. By the correspondence theorem, $K = I/M$ where $M \leq I \subsetneq A$, but M is maximal so $I = M$ so $K = \{[0]\}$ so $[x] = [0]$.

By Gelfand-Mazur the only Banach algebra where every element except 0 is invertible is \mathbb{C} , hence $A/M \cong \mathbb{C}$. ■

4.1.4 Multiplicative Characters

Definition 4.1.42. A multiplicative functional on A is a linear map $\varphi: A \rightarrow \mathbb{C}$ nonzero such that $\varphi(xy) = \varphi(x)\varphi(y)$.

Proposition 4.1.43. *If φ is a multiplicative character, it holds that $\varphi(e) = 1$ and $\|\varphi\| = 1$.*

Proof. Fix $x \in A$ such that $\varphi(x) \neq 0$. Then $\varphi(x) = \varphi(xe) = \varphi(x)\varphi(e)$ so $\varphi(e) = 1$.

Now, $\|\varphi\| \geq 1$ since $\varphi(e) = 1$. Note that if x is invertible, $1 = \varphi(e) = \varphi(xx^{-1}) = \varphi(x)\varphi(x^{-1})$ so $\varphi(x) \neq 0$ (and $\varphi(x^{-1}) = \frac{1}{\varphi(x)}$). Fix $x \in A$ with $\|x\| \leq 1$. For every $\lambda \in \mathbb{C}$ with $|\lambda| > 1$ the element $e - \frac{x}{\lambda}$ is invertible, so

$$0 \neq \varphi\left(e - \frac{x}{\lambda}\right) = \varphi(e) - \frac{1}{\lambda}\varphi(x) = 1 - \frac{\varphi(x)}{\lambda}$$

so $\varphi(x) \neq \lambda$. Hence $|\varphi(x)| < 1$ so $\|\varphi\| \leq 1$. ■

Theorem 4.1.44. *The map $\ker: \varphi \mapsto \ker \varphi$ is a bijection between (multiplicative) characters and maximal ideals.*

Proof. • We know $\ker \varphi$ is an ideal as the kernel of a homomorphism. It's maximal since $A/\ker \varphi \cong \mathbb{C}$.

- If $\ker \varphi = \ker \psi$ then $\varphi = \lambda\psi$, but since $\varphi(e) = \psi(e) = 1$ we get $\lambda = 1$ and $\varphi = \psi$.
- Fix a maximal ideal M and consider the projection

$$\varphi: A \rightarrow A/M \cong \mathbb{C}.$$

φ is a character and $\ker \varphi = M$. ■

Corollary 4.1.45. *Let A be a Banach algebra. $x \in A$ is invertible if and only if $\varphi(x) \neq 0$ for every character φ .*

Proof. If x is invertible, we've seen previously that $\varphi(x) \neq 0$ (with inverse $\varphi(x^{-1})$).

If x isn't invertible, there's a proper ideal $I \subsetneq A$. I is contained in a maximal ideal M . M is a kernel of $\varphi: A \rightarrow A/M \cong \mathbb{C}$, and since $x \in M$ we have $\varphi(x) = 0$. ■

Recall

$$W = \left\{ f \in \mathcal{C}(\mathbb{T}) \mid f = \sum_{n \in \mathbb{Z}} \hat{f}(n) e^{inx}, \sum_{n \in \mathbb{Z}} |\hat{f}(n)| < \infty \right\}.$$

Theorem 4.1.46 (Wiener). *If $f \in W$ and $f(t) \neq 0$ for all $t \in \mathbb{T}$, then $\frac{1}{f} \in W$.*

Proof. Assume $\varphi: W \rightarrow \mathbb{C}$ is a character. Write $a := \varphi(e^{it})$. Then

$$|a| = |\varphi(e^{it})| \leq \|\varphi\| \|e^{it}\| = 1 \cdot 1 = 1,$$

but also

$$\left| \frac{1}{a} \right| = |\varphi(e^{-it})| \leq \|\varphi\| \|e^{-it}\| = 1,$$

so $|a| = 1$. Write $a = e^{i\theta}$. Now

$$\varphi(e^{int}) = \varphi(e^{it})^n = (e^{i\theta})^n = e^{in\theta}.$$

For any trigonometric polynomial we therefore have $\varphi(p) = p(\theta)$. Since trigonometric polynomials are dense and φ is continuous we get $\varphi(f) = f(\theta)$ for all $f \in W$. If $f(t) \neq 0$ for all $t \in \mathbb{T}$ then $\varphi(f) \neq 0$ for every character φ , so f is invertible in W . ■

4.1.5 The Gelfand Transform

Proposition 4.1.47. *Let X be any Banach space. There exists a compact Hausdorff space K such that there's an isometric embedding $X \hookrightarrow \mathcal{C}(K)$.*

Proof. Let $\iota: X \hookrightarrow X^{**}$ send x to ev_x . Take $K = (\bar{B}_{X^*}, w^*)$ which is compact by Banach-Alaoglu. Define $\Phi: X \rightarrow \mathcal{C}(K)$ by $\Phi(x) = \text{ev}_x|_{\bar{B}_{X^*}}$. This is linear and

$$\|\Phi(x)\| = \max_{\varphi \in \bar{B}_{X^*}} |\text{ev}_x(\varphi)| = \max_{\|\varphi\| \leq 1} |\text{ev}_x(\varphi)| = \|\text{ev}_x\|^{**} = \|x\|. \quad \blacksquare$$

Given a commutative Banach algebra A , we want to find an embedding (of Banach algebras) $\Gamma: A \rightarrow \mathcal{C}(\chi)$ where χ is a compact Hausdorff space. This isn't always possible, however under some conditions this exists, and under some conditions which we shall describe, it's an isomorphism.

Proposition 4.1.48. *Let $\chi = \chi(A)$ be the space of (multiplicative) characters on A , which is contained in A^* since we've seen that characters are continuous. Then (χ, w^*) is a compact Hausdorff space.*

Proof. We have

$$\chi = \left\{ \varphi \in \bar{B}_{A^*} \mid \forall x, y \in A: \begin{array}{l} \varphi(xy) = \varphi(x)\varphi(y) \\ \varphi(e) = 1 \end{array} \right\}$$

where \bar{B}_{A^*} is a compact Hausdorff space by Banach Alaoglu. The conditions are closed, so χ is closed within \bar{B}_{A^*} and is so a compact Hausdorff space. ■

Definition 4.1.49 (The Gelfand Transform). Let

$$\begin{aligned} \Gamma: A &\rightarrow \mathcal{C}(\chi) \\ x &\mapsto \text{ev}_x. \end{aligned}$$

We sometimes denote $\hat{x} := \Gamma(x)$.

Proposition 4.1.50. 1. Γ is a well-defined homomorphism which isn't identically 0.

2. x is invertible in A if and only if $\hat{x}(\varphi) \neq 0$ for all $\varphi \in \chi$.

3. $\sigma(x) = \text{Im } \hat{x} = \{\hat{x}(\varphi) \mid \varphi \in \chi\}$.

4. $\|\hat{x}\|_{\mathcal{C}(\chi)} = r(x) \leq \|x\|$ so $\|r\| \leq 1$.

Proof. 1. Γ is well-defined by definition of the w^* topology. We have

$$\widehat{xy}(\varphi) = \varphi(xy) = \varphi(x)\varphi(y) = \hat{x}(\varphi)\hat{y}(\varphi),$$

so

$$\Gamma(xy) = \overline{xy} = \hat{x}\hat{y} = \Gamma(x)\Gamma(y).$$

Similarly, Γ is linear.

We have $\hat{e}(\varphi) = \varphi(e) = 1$ so $\Gamma(e) = 1 \neq 0$ so Γ is not identically 0.

2. We've seen that above in different notation.

3. $\lambda \in \sigma(x)$ if and only if $x - \lambda e$ is not invertible, if and only if there's $\varphi \in \chi$ such that

$$\hat{x}(\varphi) - \lambda = \hat{x}(\varphi) - \lambda\hat{e}(\varphi) = \widehat{(x - \lambda e)}(\varphi) = 0,$$

if and only if there's $\varphi \in \chi$ such that $\hat{x}(\varphi) = \lambda$, if and only if $\lambda \in \text{Im } \hat{x}$.

4. By the previous part and by Gelfand's formula, we have

$$\|\hat{x}\| = \max_{\varphi \in \chi} |\hat{x}(\varphi)| = \max_{\lambda \in \sigma(x)} |\lambda| = r(x). \quad \blacksquare$$

Proposition 4.1.51. Assume τ is a topology on χ such that (χ, τ) is compact and every \hat{x} is continuous. Then $\tau = w^*$.

Proof. Since every $\hat{x}(\varphi) = \varphi(x)$ is continuous, we know $w^* \leq \tau$. But τ is compact and w^* is Hausdorff, so $\tau = w^*$. ■

Example 4.1.52. Let $A = \mathcal{C}(S)$ where S is compact and Hausdorff. We saw the every maximal ideal $M \trianglelefteq A$ is of the form $\mathfrak{m}_s = \{f \in A \mid f(s) = 0\}$ where $s \in S$. Hence $\chi(A) = \{\text{ev}_s \mid s \in S\}$. So

$$\begin{aligned} \Gamma: A &\rightarrow \mathcal{C}(\chi) \\ f &\mapsto \hat{f} \end{aligned}$$

and

$$\Gamma(f)(\text{ev}_s) = \hat{f}(\text{ev}_s) = \text{ev}_s(f) = f(s).$$

Note that we have a homeomorphism $\chi(A) \cong S$ by $\text{ev}_s \leftrightarrow s$ by the proposition. Under the identification $\chi(A)$, the map $\Gamma: A \rightarrow \mathcal{C}(S)$ is the identity.

Example 4.1.53. Let $A = W$ be the Wiener algebra. Then $\chi(W) = \{\text{ev}_t \mid t \in \mathbb{T}\}$. Then $\chi(W) \cong \mathbb{T}$ so we can think of $\Gamma: W \rightarrow \mathcal{C}(\mathbb{T})$ as the identity. Here Γ is injective but not surjective. But, $\overline{\Gamma(W)} = \mathcal{C}(\mathbb{T})$ for example since the image contains trigonometric polynomials.

Example 4.1.54 (The Disc Algebra). Let

$$\mathbb{D} = \{z \in \mathbb{C} \mid |z| < 1\}$$

and

$$A = A(\mathbb{D}) = \{f \in \mathcal{C}(\bar{\mathbb{D}}) \mid f \text{ is analytic on } \mathbb{D}\}$$

with $\|f\| = \max_{z \in \bar{\mathbb{D}}} |f(z)|$. Show as an exercise that

$$\chi(A) = \{\text{ev}_a \mid a \in \bar{\mathbb{D}}\}.$$

Again $\chi(A) \cong \bar{\mathbb{D}}$ and $\Gamma: A \rightarrow \mathcal{C}(\bar{\mathbb{D}})$ can be thought of as the identity. This time $\overline{\Gamma(A)} = \Gamma(A) \neq \mathcal{C}(\bar{\mathbb{D}})$.

Example 4.1.55. Define

$$\begin{aligned} T\mathbb{C}^n &\rightarrow \mathbb{C}^n \\ (x_1, \dots, x_n) &\mapsto (x_2, \dots, x_n, 0). \end{aligned}$$

Take

$$A = \left\{ \sum_{i=0}^{n-1} c_i T^i \mid c_i \in \mathbb{C} \right\} \subseteq \mathcal{L}(\mathbb{C}^n) = M_{n \times n}(\mathbb{C}).$$

This is closed under multiplication since $T^n = 0$ and is complete since $\dim A < \infty$.

Assume $\varphi \in \chi(A)$. Then $0 = \varphi(e) = \varphi(T^n) = \varphi(T)^n$ so $\varphi(T) = 0$. Then

$$\varphi\left(\sum_{i=0}^{n-1} c_i T^i\right) = \sum_{i=0}^{n-1} c_i \varphi(T)^i = c_0.$$

Hence $\chi(A) = \{\text{ev}_0\}$. Hence $\mathcal{C}(\chi(A)) \cong \mathbb{C}$. Hence $\Gamma: A \rightarrow \mathbb{C}$ is given by $\Gamma\left(\sum_{i=0}^{n-1} c_i T^i\right) = c_0$. Hence Γ is not injective in this case.

Definition 4.1.56 (The Radical of a Banach Algebra). The set

$$\ker \Gamma = \{x \in A \mid \hat{x} = 0\} = \{x \in A \mid r(x) = 0\} = \{x \in A \mid x \text{ is quasi-nilpotent}\} = \{x \in A \mid \sigma(x) = \{0\}\}$$

is called the radical of A .

Definition 4.1.57 (Semisimple Banach Algebra). If $\ker \Gamma = \{0\}$, A is called semi-simple.

4.1.6 C^* Algebras

We want to understand when $\Gamma: A \rightarrow \mathcal{C}(\chi)$ is an isometric isomorphism.

On $\mathcal{C}(\chi)$ there is complex conjugation. We want an analogue for abstract algebras.

Definition 4.1.58 ($*$ -Algebra). Let A be a Banach algebra. A is called a $*$ -algebra if there exists a map $*$: $A \rightarrow A$ such that

1. $x^{**} = x$.
2. $(x + \lambda y)^* = x^* + \bar{\lambda}y$.
3. $(xy)^* = y^*x^*$.
4. $\|x^*\| = \|x\|$.

Definition 4.1.59 (C^* -Algebra). A is called a C^* -algebra if it's a $*$ -algebra such that $\|x^*x\| = \|x\|^2$.

We want to show that if A is a commutative C^* -algebra then $\Gamma: A \rightarrow \mathcal{C}(\chi)$ is an isometric $*$ -isomorphism.

Remark 4.1.60. 1. $e^* = e^*e^{**} = (e^*e)^* = e^{**} = e$.

2. If $\|x\|^2 \leq \|x^*x\|$, the property $\|x^*\| = \|x\|$ follows since

$$\|x\|^2 \leq \|x^*x\| \leq \|x^*\| \|x\|$$

implies $\|x\| \leq \|x^*\|$ and by applying the same for x^* we get $\|x\| = \|x^*\|$.

Then $\|x^*x\| \leq \|x^*\| \|x\| = \|x\|^2$ so the C^* identity also follows.

Example 4.1.61. Let $A = \mathcal{C}(S)$. This is a C^* -algebra with $f^* = \bar{f}$. We have

$$\|f^*f\| = \max |\bar{f} \cdot f| = \max |f|^2 = (\max |f|)^2 = \|f\|^2.$$

Example 4.1.62. Let $A = W$ be the Wiener algebra with $f^* = \bar{f}$. This is a $*$ -algebra but not a C^* -algebra. Take $f = e^{it} + 2^{it} - e^{3it}$. Then $\bar{f}f = 3 - e^{2it} - e^{-2it}$. So $\|f\| = 3$ and $\|\bar{f}f\| = 5$ so $\|\bar{f}f\| < \|f\|^2$.

Example 4.1.63 (Important Example). Let H be a Hilbert space and let $A = \mathcal{L}(H)$. Let T^* be the adjoint of T satisfying $\langle Tx, y \rangle = \langle x, T^*y \rangle$. We've seen the first three conditions for A to be a $*$ -algebra. We have

$$\|Tx\|^2 = \langle Tx, Tx \rangle = \langle x, T^*Tx \rangle \leq \|x\| \|T^*Tx\| \leq \|T^*T\| \|x\|^2$$

so

$$\|Tx\| \leq \sqrt{\|T^*T\|} \cdot \|x\|$$

so $\|T\| \leq \sqrt{\|T^*T\|}$ so $\|T\|^2 \leq \|T^*T\|$. By the above remark it follows that A is a C^* -algebra.

Proposition 4.1.64. In a $*$ -algebra, $\lambda \in \sigma(x)$ if and only if $\bar{\lambda} \in \sigma(x^*)$.

Proof. $\lambda \notin \sigma(x)$ if and only if there's $y \in A$ such that $y(x - \lambda e) = (x - \lambda e)y = e$ if and only if there's $y \in A$ such that $(x - \lambda e)^* y^* = y^* (x - \lambda e)^* = e$ if and only if $\bar{\lambda} \notin \sigma(x^*)$. ■

Definition 4.1.65 (Self-Adjoint). $x \in A$ is *self-adjoint* if $x = x^*$.

Definition 4.1.66 ($*$ -Homomorphism). A homomorphism $F: A \rightarrow B$ is a $*$ -homomorphism if it's a homomorphism of Banach algebras such that $F(x^*) = F(x)^*$.

Proposition 4.1.67. Let A be a commutative $*$ -algebra and assume that $\sigma(x) \subseteq \mathbb{R}$ for every $x \in A$ which is self-adjoint. Then $\Gamma: A \rightarrow \mathcal{C}(\chi)$ is a $*$ -homomorphism with dense image.

Example 4.1.68. Take $A = W$. Self-adjointness of $f \in A$ means $f = \bar{f}$ so f is real-valued. Then

$$\sigma(f) = \text{Im } \Gamma(f) = \{\varphi(f) \mid \varphi \in \chi\} = \{f(t) \mid t \in \mathbb{T}\} = \text{Im } f,$$

so $f = f^*$ implies $\sigma(f) \subseteq \mathbb{R}$. Indeed, Γ has a dense image.

Proof. For every $x \in A$ write $y = \frac{x+x^*}{2}$ and $z = \frac{x-x^*}{2i}$. Then $y^* = y$ and $z^* = z$, and we can write $x = y + iz$. We have

$$\begin{aligned}
 \Gamma(x)^* &= \overline{\Gamma(x)} \\
 &= \overline{\Gamma(y + iz)} \\
 &= \overline{\Gamma(y)} + i\overline{\Gamma(z)} \\
 &= \overline{\Gamma(y)} - i\overline{\Gamma(z)} \\
 &\stackrel{\text{Im } \Gamma(y) = \sigma(y) \subseteq \mathbb{R}}{=} \Gamma(y) - i\Gamma(z) \\
 &= \Gamma(y - iz) \\
 &= \Gamma(x^*).
 \end{aligned}$$

Let $B := \text{Im } \Gamma \leq \mathcal{C}(\chi)$. B is an algebra, is closed under complex conjugation and separates points. If $\varphi \neq \psi$, there's x such that $\varphi(x) \neq \psi(x)$, so $\Gamma(x)(\varphi) \neq \Gamma(x)(\psi)$. By Stone-Weierstrass, it follows that $\bar{B} = \mathcal{C}(\chi)$. ■

Theorem 4.1.69 (Gelfand-Naimark). *If A is a commutative C^* -algebra then $\Gamma: A \rightarrow \mathcal{C}(\chi)$ is an isometric $*$ -isomorphism.*

Proof. Assume $x = x^*$ and $\alpha + i\beta \in \sigma(x)$. We want to prove $\beta = 0$. We know $\alpha - i\beta \in \sigma(x)$. Let $t \in \mathbb{R}$. We have

$$\alpha + i(\beta + T) \in \sigma(x + ite).$$

Then

$$|\alpha + i(\beta + t)| \leq r(x + ite) \leq \|x + ite\|.$$

Similarly

$$\alpha - i(\beta + T) \in \sigma(x - ite)$$

so

$$|\alpha - i(\beta + t)| \leq \|x - ite\| = \|(x + ite)^*\| = \|x + ite\|.$$

Multiplying the inequalities we have

$$\begin{aligned}
 \alpha^2 + (\beta + t)^2 &\leq \|x + ite\|^2 \\
 &= \|(x + ite)^*(x + ite)\| \\
 &= \|(x - ite)(x + ite)\| \\
 &= \|x^2 + t^2e\| \\
 &< \|x^2\| + t^2.
 \end{aligned}$$

Then

$$\alpha^2 + \beta^2 + 2\beta t + t^2 \leq \|x^2\| + t^2$$

for all $t \in \mathbb{R}$, so $\beta = 0$ (otherwise we can make the left-hand-side as big as we want) so $\sigma(t) \subseteq \mathbb{R}$.

We claim that $\|x^2\| = \|x\|^2$ for all $x \in A$. Indeed,

$$\begin{aligned} \|x^2\|^2 &= \|(x^2)^* x^2\| \\ &= \|(xx^*)^* (xx^*)\| \\ &= \|xx^*\|^2 \\ &= \|x^*x\|^2 \\ &= \|x\|^4. \end{aligned}$$

Then $\|x^{2^k}\| = \|x\|^{2^k}$ for all $k \in \mathbb{N}$ so

$$\|\Gamma(x)\| = r(x) = \lim_{n \rightarrow \infty} \|x^n\|^{\frac{1}{n}} = \lim_{k \rightarrow \infty} \|x^{2^k}\|^{\frac{1}{2^k}} = \|x\|$$

so Γ is an isometric $*$ -homomorphism. Since A is complete (as a Banach algebra), $\Gamma(A)$ is complete (since Γ is an isometry), hence $\Gamma(A) \subseteq \mathcal{C}(\chi)$ is closed. Since it's closed and dense, it's everything. ■

4.1.7 The Spectral Theorem

We remind the spectral theorem.

Theorem 4.1.70 (The Spectral Theorem for Normal Operators). *Assume $T \in M_{n \times n}(\mathbb{C})$ is normal (i.e. $TT^* = T^*T$). Then there's an orthonormal basis $(e_i)_{i \in [n]}$ of eigenvectors of T . Let $(\lambda_i)_{i \in [m]}$ be the eigenvalues of T and $(V_i)_{i \in [m]}$ the corresponding eigenspaces with orthogonal projections $(E_i)_{i \in [m]}$ respectively. Then*

1. $E_i^2 = E_i = E_i^*$.
2. $E_i E_j = 0$ for all $i \neq j$.
3. $\sum_{i \in [m]} E_i = \text{id}$.
4. $Tx = T\left(\sum_{i \in [m]} E_i x\right) = \sum_{i \in [m]} T E_i x = \sum_{i \in [m]} \lambda_i E_i x$ so $T = \sum_{i \in [m]} \lambda_i E_i$.

Example 4.1.71. Let $H = L_2([0, 1])$ and $T \in \mathcal{L}(H)$ by $(Tf)(t) = tf(t)$. First, if $\lambda \notin [0, 1]$, $T - \lambda I$ is invertible since

$$\left((T - \lambda I)^{-1} f\right)(t) = \frac{f(t)}{t - \lambda}$$

which is well-defined since $\frac{1}{t - \lambda}$ is bounded.

If $\lambda \in [0, 1]$, $T - \lambda I$ is not onto, e.g. since $1 \notin \text{Im}(T - \lambda I)$ since $\frac{1}{t - \lambda} \notin L_2([0, 1])$. So $\sigma(T) = [0, 1]$.

But, $\sigma_p(T) = \emptyset$: If $Tf = \lambda f$ then for almost every t it holds that $tf(t) = \lambda f(t)$ so $(t - \lambda)f(t) = 0$ so for almost every t , $f(t) = 0$.

Here there is no decomposition as in the finite case since there are no eigenspaces. But, for $A \subseteq [0, 1]$ define $E(A)f = f \cdot \chi_A$. Then $E(A)^2 = E(A) = E(A)^*$ and

$$\begin{aligned} \sum_{i \in [N]} E\left(\left[\frac{i-1}{N}, \frac{i}{N}\right]\right) &= \text{id} \\ \sum_{i \in [N]} \frac{i}{N} E\left(\left[\frac{i-1}{N}, \frac{i}{N}\right]\right) &\xrightarrow{N \rightarrow \infty} T. \end{aligned}$$

We say that $T = \int_0^1 \lambda dE$. Here dE is not a measure in the usual sense. For every $A \subseteq [0, 1]$ it gives an orthogonal projection. We formalise this later.

Theorem 4.1.72. *If $A \subseteq B$ are C^* -algebras and $a \in A$, then $\sigma_A(a) = \sigma_B(a)$.*

Proposition 4.1.73. *If A is a C^* algebra and $a \in A$ is invertible, then $a^{-1} \in C^*(a)$ where $C^*(a)$ is the smallest C^* algebra containing a , namely*

$$C^*(a) = \overline{\{p(a, a^*)\}}p \in \mathbb{C}[x, y].$$

Proof. Assume first that $a = a^*$ and consider

$$B = C^*(a, a^{-1}) := \overline{\{p(a, a^{-1}) \mid p \in \mathbb{C}[x, y]\}}.$$

B is commutative, so $\Gamma: B \rightarrow \mathcal{C}(\chi(B))$ is an isometric $*$ -isomorphism. Let $K = \sigma_B(a)$. $K \subseteq \mathbb{R}$ since $a = a^*$ and $0 \notin K$ since $a^{-1} \in B$. Hence $h(\lambda) = \frac{1}{\lambda}$ is continuous on K . Choose $(p_n)_{n \in \mathbb{N}}$ polynomials such that $p_n \xrightarrow{n \rightarrow \infty} h$ uniformly on K . Then

$$\begin{aligned} \|p_n(a) - a^{-1}\| &= \|\Gamma(p_n(a) - a^{-1})\| \\ &= \max_{\varphi \in \chi(B)} |\Gamma(p_n(a) - a^{-1})(\varphi)| \\ &= \max_{\varphi \in \chi(B)} |\varphi(p_n(a) - a^{-1})| \\ &= \max_{\varphi \in \chi(B)} \left| p_n(\varphi(a)) - \varphi(a)^{-1} \right| \max_{\lambda \in \sigma_B(a)} \left| p_n(\lambda) - \frac{1}{\lambda} \right| \\ &\xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

So $p_n(a) \rightarrow a^{-1}$ so $a^{-1} \in C^*(a)$.

In the general case, write $a^{-1} = (a^*a)^{-1}a^*$. a^*a is self-adjoint so $(a^*a)^{-1} \in C^*(a^*a) \subseteq C^*(a)$ and $a^* \in C^*(a)$, so we are done. \blacksquare

Theorem 4.1.74 (The Spectral Theorem, Form #1). *Let A be a C^* algebra and let $a \in A$ be normal. There's a unique isometric $*$ -isomorphism $\rho: \mathcal{C}(\sigma(a)) \rightarrow C^*(a)$ such that $h(\lambda) = \lambda$ implies $\rho(h) = a$. Here $\rho(f)$ is denoted $f(a)$.*

Proof. $C^*(a)$ is a commutative C^* -algebra (since a is normal). So $\Gamma: C^*(a) \rightarrow \mathcal{C}(\chi(C^*(a)))$ is an isometric $*$ -isomorphism.

We claim that $\chi(C^*(a))$ is homeomorphic to $\sigma(a)$. Consider $\Gamma(a) = \hat{a}: \chi(C^*(a)) \rightarrow \mathbb{C}$. We know that \hat{a} is continuous and we know that $\text{Im } \hat{a} = \sigma(a)$. We claim also that \hat{a} is injective. If $\hat{a}(\varphi) = \hat{a}(\psi)$ it holds that $\varphi(a) = \psi(a)$ so $\varphi(a^*) = \overline{\varphi(a)} = \overline{\psi(a)} = \psi(a^*)$, so $\varphi(p(aa^*)) = \psi(p(aa^*))$ for every polynomial $p \in \mathbb{C}[x]$. So $\varphi(x) = \psi(x)$ for all $x \in C^*(a)$ by continuity of φ, ψ . Hence \hat{a} is continuous and bijective. Since $\chi(C^*(a))$ is compact and $\sigma(a)$ is Hausdorff it follows that \hat{a} is a homeomorphism.

For every $f \in \mathcal{C}(\sigma(a))$ define $\rho(f) = \Gamma^{-1}(f \circ \hat{a}) \in C^*(a)$ which is obviously an isometric $*$ -isomorphism. Not that

$$\rho(h) = \Gamma^{-1}(h \circ \hat{a}) = \Gamma^{-1}(\hat{a}) = a.$$

If ρ_1, ρ_2 are two such maps, $\rho_1(\lambda) = \rho_2(\lambda)$ so $\rho_1(\bar{\lambda}) = \rho_2(\bar{\lambda})$ so $\rho_1(p(\lambda, \bar{\lambda})) = \rho_2(p(\lambda, \bar{\lambda}))$ for every polynomial $p \in \mathbb{C}[x, y]$, so by continuity $\rho_1(f) = \rho_2(f)$ for all $f \in \mathcal{C}(\sigma(a))$. ■

Corollary 4.1.75 (The Spectral Mapping Theorem). *For every $f \in \mathcal{C}(\sigma(a))$ it holds that $\sigma(f(a)) = f(\sigma(a))$.*

Proof. We have

$$\sigma(f(a)) = \sigma(f) = \text{Im } f - f(\sigma(a)). \quad \blacksquare$$

Example 4.1.76. Let $T \in M_{n \times n}(\mathbb{C})$ be normal and write $\sigma(T) = \{\lambda_i\}_{i \in [m]}$. Let $f_i := \chi_{\{\lambda_i\}} \in \mathcal{C}(\sigma(T))$ (this is continuous since the spectrum is discrete). Take $E_i = f_i(T)$. We have

1. $f_i = f_i^2 = \bar{f}_i$ (on $\sigma(T)$) so $E_i = E_i^2 = E_i^*$.
2. $f_i f_j = 0$ for every $i \neq j$ (on $\sigma(T)$) so $E_i E_j = 0$.
3. $\sum_{i \in [m]} f_i = 1$ (...) so $\sum_{i \in [m]} E_i = I$.
4. $\sum_{i \in [m]} \lambda_i f_i = \lambda$ so $\sum_{i \in [m]} \lambda_i E_i = T$.

Proposition 4.1.77. *Let $T \in \mathcal{L}(H)$ be normal. Then*

1. *T is self-adjoint if and only if $\sigma(T) \subseteq \mathbb{R}$.*
2. *T is unitary (i.e. $TT^* = T^*T = \text{id}$) if and only if $\sigma(T) \subseteq S^1 \subseteq \mathbb{C}$.*

Proof. Take $h(\lambda) = \lambda$.

1. $T = T^*$ if and only if $h(T) = h(T)^* = \bar{h}(T)$ if and only if $h = \bar{h}$ on $\sigma(T)$, if and only if $\sigma(T) \subseteq \mathbb{R}$.
2. $TT^* = I$ if and only if $(h\bar{h})(T) = h(T)\bar{h}(T) = \text{id} = 1(T)$ if and only if $h\bar{h} = 1$ on $\sigma(T)$ if and only if $|\lambda| = 1$ on $\sigma(T)$. ■

We now want to define “ $T = \int \lambda dE$ ”. For this we want $E(A) = \chi_A(T)$, but χ_A isn’t in general continuous.

Lemma 4.1.78. *Let H be a Hilbert space and $B: H \times H \rightarrow \mathbb{C}$ be sesquilinear and bounded. Then there’s a unique $S \in \mathcal{L}(H)$ such that $\langle Sx, y \rangle = B(x, y)$, and moreover $\|S\| = \|B\|$.*

Proof. Fix $x \in H$ and consider $\psi_x(y) := \overline{B(x, y)}$ which is linear. We have

$$|\psi_x(y)| = |B(x, y)| \leq \|B\| \|x\| \|y\|,$$

so $\|\psi_x\| \leq \|B\| \|x\|$. By Riesz there’s $x^* = Sx$ such that $\psi_x(y) = \langle y, Sx \rangle$, but then

$$B(x, y) = \overline{\psi_x(y)} = \langle y, Sx \rangle = \langle Sx, y \rangle.$$

We know

$$\|Sx\| = \|x^*\| = \|\psi_x\| \leq \|B\| \|x\|,$$

so $\|S\| \leq \|B\|$. Conversely,

$$|B(x, y)| = |\langle Sx, y \rangle| \leq \|Sx\| \|y\| \leq \|S\| \|x\| \|y\|$$

so $\|B\| \leq \|S\|$.

For uniqueness, if $\langle S_1x, y \rangle = \langle S_2x, y \rangle$ for all x, y then $S_1x - S_2x \perp H$ so $S_1x = S_2x$. ■

Every $\varphi \in \mathcal{C}(\sigma(a))^*$ is of the form $\varphi(f) = \int f d\mu$. But then, one get integrate all measurable and bounded functions, not necessarily continuous ones. ρ we defined above returns operators, so we have to go from operators through numbers and back to operators.

Formally, given $x, y \in H$ we consider $\ell_{x,y}(f) = \langle f(T)x, y \rangle$. Then

$$|\ell_{x,y}(f)| \leq \|f(T)\| \|x\| \|y\| = \|f\| \|x\| \|y\|$$

so $\|\ell_{x,y}\| \leq \|x\| \|y\|$ (and there’s in fact equality). So, $\ell_{x,y} \in \mathcal{C}(\sigma(a))^*$ so there’s a complex measure $\mu_{x,y}$ such that

$$\ell_{x,y}(f) = \langle f(T)x, y \rangle = \int_{\sigma(T)} f d\mu_{x,y}$$

for all $f \in \mathcal{C}(\sigma(T))$. Write $B(\sigma(T))$ for all bounded Borel measurable functions. $\ell_{x,y}$ is sesquilinear hence so is $\mu_{x,y}$. Hence, for every $g \in B(\sigma(T))$ one has that

$$\int_{\sigma(T)} g d\mu_{x,y}$$

is sesquilinear in x, y , and bounded since

$$\left| \int_{\sigma(T)} g d\mu_{x,y} \right| \leq \sup |g| \cdot \|\mu_{x,y}\| \leq \sup |g| \cdot \|x\| \|y\|.$$

By the lemma, there's $g(T)$ such that

$$\langle g(T)x, y \rangle = \int_{\sigma(T)} g \, d\mu_{x,y}.$$

Theorem 4.1.79. *The map*

$$\begin{aligned} B(\sigma(T)) &\rightarrow \mathcal{L}(H) \\ g &\mapsto g(T) \end{aligned}$$

*is a *-homomorphism and $\|g(T)\| \leq \sup_{\sigma(T)} |g|$.*

Proof. Linearity: For every $f, g \in B(\sigma(T))$ we have

$$\begin{aligned} \langle (\alpha f + g)(T)x, y \rangle &= \int (\alpha f + g) \, d\mu_{x,y} \\ &= \int \alpha f \, d\mu_{x,y} + \int g \, d\mu_{x,y} \\ &= \langle \alpha f(T)x, y \rangle + \langle g(T)x, y \rangle \\ &= \langle \alpha f(T) + g(T)x, y \rangle \end{aligned}$$

$$\text{so } (\alpha f + g)(T) = \alpha f(T) + g(T).$$

***-preserving:** For $f \in \mathcal{C}(\sigma(T))$ we know

$$\langle \bar{f}(T)x, y \rangle = \langle f(T)^*x, y \rangle.$$

The left-hand-side equals $\int \bar{f} \, d\mu_{x,y}$. The right hand side equals

$$\begin{aligned} \langle x, f(T)y \rangle &= \overline{\langle f(T)y, x \rangle} \\ &= \overline{\int f \, d\mu_{y,x}} \\ &= \int \bar{f} \, d\mu_{y,x}. \end{aligned}$$

So, $\mu_{x,y} = \overline{\mu_{y,x}}$.

But then, for every $f \in B(\sigma(T))$ we have

$$\begin{aligned} \langle \bar{f}(T)x, y \rangle &= \int \bar{f} \, d\mu_{x,y} \\ &= \overline{\int f \, d\mu_{x,y}} \\ &= \overline{\int f \, d\mu_{y,x}} \\ &= \overline{\langle f(T)y, x \rangle} \\ &= \overline{\langle y, f(T)^*x \rangle} \\ &= \langle f(T)^*x, y \rangle, \end{aligned}$$

so $\bar{f}(T) = f(T)^*$.

Multiplicativity: Fix $f, g \in \mathcal{C}(\sigma(T))$. We have

$$\begin{aligned}
 \int fg \, d\mu_{x,y} &= \langle (fg)(T)x, y \rangle \\
 &= \langle f(T)g(T)x, y \rangle \\
 &= \langle g(T)f(T)x, y \rangle \\
 &= \langle f(T)x, g(T)^*y \rangle \\
 &= \int f \, d\mu_{x, g(T)^*y}
 \end{aligned}$$

so $g \, d\mu_{x,y} = \mu_{x, g(T)^*y}$. Now, if $f \in B(\sigma(T))$ and $g \in \mathcal{C}(\sigma(T))$ we have

$$\begin{aligned}
 \int gf \, d\mu_{x,y} &= \langle (fg)(T)x, y \rangle \\
 &= \int fg \, d\mu_{x,y} \\
 &= \int f \, d\mu_{x, g(T)^*y} \\
 &= \langle f(T)x, g(T)^*y \rangle \\
 &= \langle g(T)f(T)x, y \rangle \\
 &= \int g \, d\mu_{f(T)x, y},
 \end{aligned}$$

so $f \, d\mu_{x,y} = \mu_{f(T)x, y}$.

Now if $f, g \in B(\sigma(T))$ we have

$$\begin{aligned}
 \langle (fg)(T)x, y \rangle &= \int fg \, d\mu_{x,y} \\
 &= \int g \, d\mu_{f(T)x, y} \\
 &= \langle g(T)f(T)x, y \rangle,
 \end{aligned}$$

so $(fg)(T) = g(T)f(T)$.

Norm Bound: We have

$$\begin{aligned}
 |\langle g(T)x, y \rangle| &= \left| \int g \, d\mu_{x,y} \right| \\
 &\leq \sup |g| \cdot \|\mu_{x,y}\| \\
 &\leq \sup |g| \cdot \|x\| \|y\|
 \end{aligned}$$

so $\|g(T)\| \leq \sup |g|$. ■

Definition 4.1.80 (The Spectral Resolution of an Operator). The spectral resolution of a normal operator $T \in \mathcal{L}(H)$ is

$$\begin{aligned} E: \text{Borel}(\sigma(T)) &\rightarrow \mathcal{L}(H) \\ A &\mapsto \chi_A(T). \end{aligned}$$

Proposition 4.1.81. E is resolution of the identity in the following sense:

1. $E(A)^2 = E(A) = E(A)^*$.
2. If $(A_n)_{n \in \mathbb{N}_+}$ are pairwise disjoint and $A = \bigcup_{n \in \mathbb{N}} A_n$ then $E(A)x = \sum_{n \in \mathbb{N}_+} E(A_n)x$ for all $x \in H$.
3. $E(\emptyset) = 0$ and $E(\sigma(T)) = I$.
4. $E(A \cap B) = E(A)E(B)$ (and specifically if $A \cap B = \emptyset$ by the above $E(A)E(B) = 0$).

Also, the following holds.

5. $E(A)$ commutes with $f(T)$ for all $f \in B(\sigma(T))$.
6. If $A \subseteq \sigma(T)$ is relatively open then $E(A) \neq 0$.

Proof. 1. This is true since $\chi_A^2 = \chi_A = \overline{\chi_A}$ and this is a $*$ -homomorphism.

2. Note that

$$\begin{aligned} \langle E(A)x, y \rangle &= \langle \chi_A(T)x, y \rangle \\ &= \int \chi_A d\mu_{x,y} \\ &= \mu_{x,y}(A) \\ &= \sum_{n \in \mathbb{N}_+} \mu_{x,y}(A_n) \\ &= \sum_{n \in \mathbb{N}_+} \langle E(A_n)x, y \rangle. \end{aligned}$$

This shows weak convergence, but note

$$\sum_{n \in [N]} \|E(A_n)x\|^2 = \left\| \sum_{n \in [N]} E(A_n)x \right\|^2 = \left\| E\left(\bigcup_{n \in [N]} A_n\right)x \right\|^2 \leq \|x\|^2,$$

so $\sum_{n \in \mathbb{N}_+} \|E(A_n)x\|^2 < \infty$. Hence

$$\left\| \sum_{n=M}^N E(A_n)x \right\|^2 = \sum_{n=M}^N \|E(A_n)x\|^2 \xrightarrow{N, M \rightarrow \infty} 0,$$

so the sequence is Cauchy and so the weak convergence is a strong convergence.

3. This is true since $\chi_\emptyset = 0$ and $\chi_{\sigma(T)} = 1$ and this is a $*$ -homomorphism.
4. This is true since $\chi_{A \cap B} = \chi_A \chi_B$ and this is a $*$ -homomorphism.
5. This is true since $\chi_A \cdot f = f \cdot \chi_A$ and this is a $*$ -homomorphism.
6. Pick $f \in \mathcal{C}(\sigma(T))$ non-zero with $\text{supp}(f) \subseteq A$ (e.g. $f(\lambda) = d(\lambda, \sigma(T) \setminus A)$). Then $f \cdot \chi_A = f$ so $f(T)E(A) = f(T) \neq 0$. Hence $E(A) \neq 0$. ■

Proposition 4.1.82. *Let (Ω, \mathcal{E}) be a measurable space and let $E: \mathcal{E} \rightarrow \mathcal{L}(H)$ be a resolution of the identity. Then $\mu_{x,y}(A) = \langle E(A)x, y \rangle$ are complex measure satisfying $\|\mu_{x,y}\| \leq \|x\| \|y\|$ and $\mu_{x,x} \geq 0$.*

Proposition 4.1.83. *Let $E: \mathcal{E} \rightarrow \mathcal{L}(H)$ be a resolution of the identity on (Ω, \mathcal{E}) . Fix $f \in B(\Omega)$. There exists a unique $S \in \mathcal{L}(H)$ such that the following holds.*

- If $\Omega = \bigsqcup_{n \in [N]} A_n$ and $|f(\lambda_1) - f(\lambda_2)| < \varepsilon$ for all $\lambda_1, \lambda_2 \in A_i$, then for every choice $t_n \in A_n$ we have

$$\left\| S - \sum_{n \in [N]} f(t_n) E(A_n) \right\| < \varepsilon.$$

We write $S = \int_\Omega f dE$.

Proof. Take as usual $\mu_{x,y}(A) = \langle E(A)x, y \rangle$. The map

$$(x, y) \mapsto \int f d\mu_{x,y}$$

is sesquilinear and bounded, so there's S such that $\langle Sx, y \rangle = \int f d\mu_{x,y}$ for all x, y . Now,

$$\begin{aligned} \left| \left\langle \left(S - \sum_{n \in [N]} f(t_n) E(A_n) \right) x, y \right\rangle \right| &= \left| \int f d\mu_{x,y} - \sum_{n \in [N]} f(t_n) \int_{A_n} d\mu_{x,y} \right| \\ &= \left| \sum_{n \in [N]} \int_{A_n} (f - f(t_n)) d\mu_{x,y} \right| \\ &\leq \varepsilon \|\mu_{x,y}\| \\ &= \varepsilon \|x\| \|y\|. \end{aligned}$$

From this we get

$$\left\| S - \sum_{n \in [N]} f(t_n) E(A_n) \right\| \leq \varepsilon. \quad \blacksquare$$

Proposition 4.1.84. *Let T be normal, E is its spectral resolution and $f \in B(\sigma(T))$. Then*

$$f(T) = \int_{\sigma(T)} f \, dE.$$

Proof.

$$\langle f(T)x, y \rangle = \int_{\sigma(T)} f \, d\mu_{x,y} = \left\langle \left(\int_{\sigma(T)} f \, dE \right) x, y \right\rangle. \quad \blacksquare$$

Corollary 4.1.85 (The Spectral Theorem). $I = \int_{\sigma(T)} dE$ and $T = \int_{\sigma(T)} \lambda \, dE$.

Proposition 4.1.86. *Let $T \in \mathcal{L}(H)$ be normal and let $\lambda_0 \in \sigma(T)$. Then*

$$V_{\lambda_0} = \ker(T - \lambda_0 I) = \operatorname{Im} E(\{\lambda_0\}).$$

In particular $E(\{\lambda_0\}) \neq 0$ if and only if $\lambda_0 = \sigma_p(T)$.

Proof. $(\lambda - \lambda_0) \chi_{\{\lambda_0\}} = 0$ so $(\lambda - \lambda_0) \chi_{\{\lambda_0\}}(T) = 0$. But this equals to $(T - \lambda_0 I) E(\{\lambda_0\})$, so $\operatorname{Im} E(\{\lambda_0\}) \subseteq \ker(T - \lambda_0 I)$.

Take $A_n = \{\lambda \in \sigma(T) \mid |\lambda - \lambda_0| > \frac{1}{n}\}$ and $f_n = \frac{1}{\lambda - \lambda_0} \chi_{A_n} \in B(\sigma(T))$. Then

$$E(A_n) = \chi_{A_n}(T) = (f_n \cdot (\lambda - \lambda_0))(T) = f_n(T)(T - \lambda_0 I),$$

so if $x \in V_{\lambda_0}$ then $E(A_n)x = 0$. Hence

$$E\left(\bigcup_{n \in \mathbb{N}_+} A_n\right)x = 0$$

by continuity of the measure. But, since $\bigcup_{n \in \mathbb{N}_+} A_n = \sigma(T) \setminus \{\lambda_0\}$ and $E(\{\lambda_0\})x + E(\sigma(T) \setminus \{\lambda_0\})x = x$ so $x \in \operatorname{Im} E(\{\lambda_0\})$. \blacksquare

Chapter 5

Unbounded Operators

Definition 5.0.1. An operator on a Hilbert space H over \mathbb{C} is a subspace $D(T) \subseteq H$ and a linear map $T: D(T) \rightarrow H$, not necessarily bounded. We assume that $\overline{D(T)} = H$.

Definition 5.0.2. For an operator T on H define

$$D(T^*) := \{y \in H \mid \psi_y(x) = \langle Tx, y \rangle \text{ is bounded}\}.$$

Proposition 5.0.3. *There exists a unique operator T^* on $D(T^*)$ such that $\langle Tx, y \rangle = \langle x, T^*y \rangle$ for all $x \in D(T)$ and $y \in D(T^*)$.*

Proof. For every $y \in D(T^*)$, the functional $\psi_y: D(T) \rightarrow H$ is bounded so it can be uniquely extended to H . By Riesz there's $y^* := T^*y$ such that $\psi_y(x) = \langle Tx, y \rangle = \langle x, T^*y \rangle$.

Note that if

$$\forall x \in D(T) : \langle x, y_2 \rangle = \langle Tx, y \rangle = \langle x, y_1 \rangle$$

then $y_1 - y_2 = 0$ on $D(T)$ so $y_1 = y_2$. ■

Definition 5.0.4. 1. $T \subseteq S$ if $D(T) \subseteq D(S)$ and $S|_{D(T)} = T$.

2. T is called symmetric if $T \subseteq T^*$, in which case for all $x, y \in D(T)$ it holds that $\langle Tx, y \rangle = \langle x, Ty \rangle$.

3. T is called self-adjoint if $T = T^*$, then in particular $D(T) = D(T^*)$.

Definition 5.0.5 (Absolutely Continuous Function). A function $f: [a, b] \rightarrow \mathbb{C}$ is called absolutely continuous if

$$f(x) = f(a) + \int_0^x g(t) dt$$

for some $g \in L^1([a, b])$.

Remark 5.0.6. If f is absolutely continuous, it's differentiable almost-everywhere and $f' = g$.

Example 5.0.7. Define T_1, T_2, T_3 on $L^2([0, 1])$ by $T_j f = if'$ with the domains

$$\begin{aligned} D(T_1) &= \{f \in L^2 \mid f \text{ is absolutely continuous and } f' \in L^2\} \\ D(T_2) &= \{f \in L^2 \mid f(0) = f(1)\} \cap D(T_1) \\ D(T_3) &= \{f \in L^2 \mid f(0) = f(1) = 0\} \cap D(T_1). \end{aligned}$$

Clearly $T_3 \subseteq T_2 \subseteq T_1$. Note that $\overline{D(T_3)} = L^2([0, 1])$.

We claim $T_1^* = T_3, T_2^* = T_2, T_3^* = T_1$. This implies in particular T_2 is self-adjoint, T_3 is symmetric and T_1 is neither. So, these properties are very sensitive to a change of domain.

Note that for every $f, g \in D(T_j)$ we have

$$\begin{aligned} \langle T_j f, g \rangle &= \langle if', g \rangle \\ &= \int_0^1 if' \bar{g} \\ &= if\bar{g}'|_0^1 - i \int_0^1 f\bar{g}' \\ &= if\bar{g}'|_0^1 + \int_0^1 f \overline{(ig)'} \\ &= if\bar{g}'|_0^1 + \langle f, T_\ell g \rangle \end{aligned}$$

for all $\ell \in \{1, 2, 3\}$. In particular, we see $T_1^* \supseteq T_3, T_2^* \supseteq T_2$ and $T_3^* \supseteq T_1$.

We check that $T_2^* = T_2$. We claim that $\text{Im } T_2 = \text{Span}\{1\}^\perp$. Indeed,

$$\langle T_2 f, 1 \rangle = \int_0^1 if' \cdot 1 = i(f(1) - f(0)) = 0$$

and if $\langle g, 1 \rangle = 0$ then $\int_0^1 g = 0$ so

$$g = T_2 \left(-i \int_0^x g(t) dt \right) \in \text{Im}(T_2).$$

Now, fix $g \in (T_2^*)$. Write $h = T_2^* g$ and $H \int_0^x h(t) dt$. We know that for all

$f \in D(T_2)$ we have

$$\begin{aligned}
 \langle T_2 f, g \rangle &= \langle f, T_2^* g \rangle \\
 &= \langle f, h \rangle \\
 &= \int_0^1 f \bar{h} \\
 &= \int_0^1 f \bar{H}' \\
 &= f \bar{H}|_0^1 - \int_0^1 f' \bar{H} \\
 &= f(1) \bar{H}(1) - \int_0^1 f' \bar{H}.
 \end{aligned}$$

Taking $f = 1$ we get $0 = \bar{H}(1) - 0 = \bar{H}(1)$ so $H(1) = 0$, so

$$\langle i f', g \rangle = -\langle f', H \rangle.$$

This implies $\langle f', H - ig \rangle = 0$ which implies $H - ig \perp \text{Im } T_2$. By the above claim this means $H - ig$ is a constant c . Then $g = i(c - H)$. Hence g is absolutely-continuous (because H is), $g' = -ih \in L^2$ and $g(0) = g(1) = ic$, so $g \in D(T_2)$. Hence $T_2^* = T_2$.

Proposition 5.0.8. *For every T it holds that $(\text{Im } T)^\perp = \ker T^*$.*

Proof. Let $z \in (\text{Im } T)^\perp$. Then $\langle Tx, z \rangle = 0$ for every $x \in D(T)$ where $\langle Tx, z \rangle$ is continuous in x . Then $z \in D(T^*)$ and $0 = \langle Tx, z \rangle = \langle x, T^* z \rangle$ so $T^* z \perp D(T)$ so $z \in \ker(T^*)$.

The other inclusion is the same argument in reverse. ■

Definition 5.0.9 (Graph of an Operator). Let T be an operator. We define

$$\Gamma(T) = \{(x, Tx) \mid x \in D(T)\} \subseteq H \oplus H.$$

T is closed if $\Gamma(T)$ is closed.

Proposition 5.0.10. *Define $V: H \oplus H \rightarrow H \oplus H$ by $V(x, y) = (-y, x)$. Then*

1. V is a unitary map.

2. $\Gamma(T^*) = (V\Gamma(T))^\perp$.

Proof. 1. We have

$$\begin{aligned}
 \langle V(x_1, y_1), V(x_2, y_2) \rangle &= \langle (-y_1, x_1), (-y_2, x_2) \rangle \\
 &= \langle y_1, y_2 \rangle + \langle x_1, x_2 \rangle \\
 &= \langle (x_1, y_1), (x_2, y_2) \rangle.
 \end{aligned}$$

2. Fix $(x, Tx) \in \Gamma(T)$ and $(y, T^*y) \in \Gamma(T^*)$. We have

$$\begin{aligned} \langle (y, T^*y), V(x, Tx) \rangle &= \langle (y, T^*y), (-Tx, x) \rangle \\ &= -\langle y, Tx \rangle + \langle T^*y, x \rangle \\ &= -\overline{\langle Tx, y \rangle} + \langle x, T^*y \rangle \\ &= 0. \end{aligned}$$

Hence $\Gamma(T^*) \subseteq V(\Gamma(T))^\perp$.

Fix $(a, b) \in V(\Gamma(T))^\perp$. For every $x \in D(T)$ we have

$$\begin{aligned} 0 &= \langle (a, b), V(x, Tx) \rangle \\ &= \langle (a, b), (-Tx, x) \rangle \\ &= -\langle a, Tx \rangle + \langle b, x \rangle \end{aligned}$$

so $\langle Tx, a \rangle = \langle x, b \rangle$, where the latter is continuous in x . Hence $a \in D(T^*)$, and

$$\langle x, T^*a \rangle = \langle Tx, a \rangle = \langle x, b \rangle$$

so $b - T^*a \perp D(T)$, so $b = T^*a$. Hence $(a, b) = (a, T^*a) \in \Gamma(T^*)$. ■

Corollary 5.0.11. *T^* is always closed. In particular, if T is self-adjoint, T is closed.*

Proposition 5.0.12. *If T is closed, $\overline{D(T^*)} = H$ and $T^{**} = T$.*

Proof. We need to prove the first part in order for T^{**} to be defined. Let $z \perp D(T^*)$. Then $(z, 0) \perp \Gamma(T^*)$ since $\langle (z, 0), (y, T^*y) \rangle = \langle z, y \rangle$. So

$$(z, 0) \in \Gamma(T^*)^\perp = V(\Gamma(T))^{\perp\perp} = \overline{V(\Gamma(T))} = V(\Gamma(T))$$

where the last equality is true since T is closed and V is an isometry. Hence $(z, 0) = V(x, Tx) = (-Tx, x)$ for some x and by comparison $x = 0$ so $z = 0$. Hence $\overline{D(T^*)} = H$.

Now,

$$\begin{aligned} \Gamma(T^{**}) &= (V(\Gamma(T^*)))^\perp \\ &= V\left(V(\Gamma(T))^\perp\right)^\perp \\ &= V^2(\Gamma(T))^{\perp\perp} \\ &= (-\Gamma(T))^{\perp\perp} \\ &= \Gamma(T)^{\perp\perp} \\ &= \Gamma(T). \end{aligned}$$

Hence $T^{**} = T$. ■

Definition 5.0.13 (Resolvent Set). Say λ is in the resolvent set of T if $T - \lambda I: D(T) \rightarrow H$ has a bounded inverse.

Denote the resolvent set of T by $\rho(T)$.

Definition 5.0.14 (Spectrum). Let $\sigma(T) := \mathbb{C} \setminus \rho(T)$ be the spectrum of T .

Proposition 5.0.15. 1. If T is closed, $\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$.

2. If T isn't closed, $\sigma(T) = \mathbb{C}$.

Proof. 1. We should prove that if $T - \lambda I: D(T) \rightarrow \mathbb{H}$ is a bijection then $(T - \lambda I)^{-1}$ is bounded. But,

$$\begin{aligned} \Gamma((T - \lambda I)^{-1}) &= \left\{ (y, (T - \lambda I)^{-1}y) \mid y \in \mathbb{H} \right\} \\ &= \{ ((T - \lambda I)x, x) \mid x \in D(T) \} \end{aligned}$$

where the latter is closed since $T - \lambda I$ is. Hence $(T - \lambda I)^{-1}$ is closed by the closed graph theorem.

2. If $\lambda \notin \sigma(T)$, $(T - \lambda I)$ has a bounded inverse. Then $(T - \lambda I)^{-1}$ is closed. Hence $T - \lambda I$ is also closed. ■

Proposition 5.0.16. $\sigma(T)$ is closed.

Proof. We saw that if A has a bounded inverse and $\|B - A\| < \frac{1}{\|A^{-1}\|}$ then B has a bounded inverse. Take $A = T - \lambda I$ and $B = T - \lambda I - \varepsilon I$. Then repeat the same proof as for bounded operators. ■

Example 5.0.17. Let $T = T_2$ on $L^2([0, 1])$ as in example 5.0.7. Then T is self-adjoint, and we compute $\sigma(T)$. Assume $T(f) = \lambda f$. Then $if' - \lambda f = 0$ so $f(t) = ce^{-i\lambda t}$. We want $f \in D(T)$ so we need $c = f(0) = f(1) = ce^{-\lambda}$. If $f \neq 0$ we have $c \neq 0$ so $e^{-i\lambda} = 1$ so $\lambda \in 2\pi\mathbb{Z}$. Hence $\sigma_p(T) = 2\pi\mathbb{Z}$.

To check if $\lambda \in \sigma_c(T) \cup \sigma_r(T)$ we need to solve

$$if' - \lambda f = (T - \lambda)f = g \in L^2([0, 1])$$

for f . This is a first-order linear ODE with solution

$$f(t) = Ce^{-i\lambda t} - ie^{-i\lambda t} \int_0^t e^{i\lambda s} g(s) ds.$$

We require $f(0) = f(1)$. If $e^{-\lambda} \neq 1$ we can choose c to have $f(0) = f(1)$. Hence $\sigma(T) = \sigma_p(T) = 2\pi\mathbb{Z}$.

Note that here $\sigma(\mathbb{Z})$ is unbounded.

Example 5.0.18. Let $T = T_3$ from example 5.0.7. Then T is symmetric and $T = T_1^*$ so T is closed.

Note that $\sigma_p(T^*) = \mathbb{C}$ (since $D(T^*) = \{f \in L^2[0, 1] \mid f' \in L^2\}$ and we can always take $f(t) = e^{-i\lambda t} \in D(T^*)$ eigenfunction with eigenvalue λ).

For T we have

$$\operatorname{Im}(T - \lambda I)^\perp = \ker(T - \lambda I)^\perp = \ker(T^* - \bar{\lambda}I) \neq \{0\}$$

so $\overline{\operatorname{Im}(T - \lambda I)} \neq H$, so $\sigma(T) = \sigma_r(T) = \mathbb{C}$.

This is terrible and is an example for why we later treat only self-adjoint operators.

Proposition 5.0.19. Let T be symmetric. Fix $\lambda = \alpha + i\beta \in \mathbb{C}$.

$$1. \|(T - \lambda I)x\|^2 = \|(T - \alpha I)x\|^2 + |\beta|^2 \|x\|^2.$$

2. If $\beta \neq 0$, $T - \lambda I$ is injective.

3. If $\beta \neq 0$ and T is closed, $\operatorname{Im}(T - \lambda I)$ is closed.

Proof. 1. By replacing T with $T - \alpha I$ we may assume $\alpha = 0$. We have

$$\begin{aligned} \|(T - \lambda I)x\|^2 &= \|Tx - i\beta x\|^2 \\ &= \|Tx\|^2 - 2\Re\langle Tx, i\beta x \rangle + \|i\beta x\|^2. \end{aligned}$$

Note that $\langle Tx, x \rangle = \langle x, Tx \rangle = \overline{\langle Tx, x \rangle}$ so $\langle Tx, x \rangle \in \mathbb{R}$. So

$$\|(T - \lambda I)x\|^2 = \|Tx\|^2 - 2\Re(i\beta \langle Tx, x \rangle) + |\beta|^2 \|x\|^2.$$

2. We have $\|(T - \lambda I)x\| \geq |\beta| \|x\|$. So if $(T - \lambda I)x = 0$ we get $x = 0$.

3. Assume $(y_n)_{n \in \mathbb{N}}$ and $y_n \xrightarrow{n \rightarrow \infty} y$. We can write $y_n = (T - \lambda I)x_n$ for some sequence $(x_n)_{n \in \mathbb{N}}$. But $\|x_n - x_m\| \leq \frac{1}{|\beta|} \|y_n - y_m\| \xrightarrow{n, m \rightarrow \infty} 0$ so $(x_n)_{n \in \mathbb{N}}$ is Cauchy and hence convergent to some x . Since $(T - \lambda I)x_n = y_n \xrightarrow{n \rightarrow \infty} y$ and since $T - \lambda I$ is closed, we get $y = (T - \lambda I)x$. ■

Proposition 5.0.20. If T is self-adjoint, $\sigma(T) \subseteq \mathbb{R}$ and $\sigma_r(T) = \emptyset$.

Proof. Fix $\lambda \notin \mathbb{C} \setminus \mathbb{R}$. Then

$$\begin{aligned} \operatorname{Im}(T - \lambda I)^\perp &= \ker(T - \lambda I)^* \\ &= \ker(T - \bar{\lambda}I) \\ &= \{0\} \end{aligned}$$

where the last equality is $\bar{\lambda} \notin \sigma_p(T)$. Hence

$$\operatorname{Im}(T - \lambda I) = \overline{\operatorname{Im}(T - \lambda I)} = \{0\}^\perp = H$$

so $\lambda \in \rho(T)$. Also, if $\lambda \in \mathbb{R}$ we get

$$\operatorname{Im}(T - \lambda I)^\perp = \ker(T - \lambda I).$$

If $\lambda \in \sigma_r(T)$ we get the the left-hand side is nontrivial but the right-hand side is, a contradiction. ■

5.0.1 The Spectral Theorem

We want to write $T = \int_{\sigma(T)} \lambda dE$ for some resolution E . We cannot prove this because we don't yet know what this means. Since $\sigma(T)$ isn't bounded, λ isn't bounded on it. But, we haven't treated integration of unbounded functions.

Definition 5.0.21. Let S be a normal bounded operator on H with spectral resolution E . Fix $f: \sigma(S) \rightarrow \mathbb{C}$ measurable. Take

$$A_n := \{\lambda \in \sigma(S) \mid |f(\lambda)| < n\}$$

and define

$$\left(\int_{\sigma(T)} f dE \right) x = f(S)x := \lim_{n \rightarrow \infty} [(f\chi_{A_n})(S)x]$$

for all $x \in H$ such that the limit exists.

Proposition 5.0.22. For $T = f(S)$ we have the following:

1. $D(T) = \left\{ x \in H \mid \int |f|^2 dm_{x,x} < \infty \right\}$ where $m_{x,x}(A) = \langle E(A)x, x \rangle$.
2. $\overline{D(T)} = H$.
3. $T^* = \int \bar{f} dE = \bar{f}(S)$.
4. $f(S)g(S) = (fg)(S)$. In fact, $D(f(S)g(S)) = D((fg)(S)) \cap D(g(S))$.
5. If f is real, T is self-adjoint.

We prove the proposition later and go back to the main theorem. The idea for the proof is as follows. A self-adjoint operator has a spectrum which is unbounded in \mathbb{R} . We apply to this a function $f(\lambda) = \frac{\lambda-i}{\lambda+i}$ which maps \mathbb{R} to S^1 . The operator T is sent to $f(T)$ which is unitary and for which we know the spectral theorem. We then apply the inverse map $g(\lambda) = i\frac{1+\lambda}{1-\lambda}$ and get a spectral resolution for T .

Definition 5.0.23. Let T be self-adjoint and define

$$U: H \rightarrow H$$

by $(T - iI)(T + iI)^{-1}$ which we call the Cayley transform of T .

Proposition 5.0.24. 1. U is well-defined and unitary (and in particular bounded).

2. $I - U$ is 1 to 1, $\text{Im}(I - U) = D(T)$ and we have $T = i(I + U)(I - U)^{-1}$.
3. $\lambda \in \sigma(T)$ if and only if $f(\lambda) = \frac{\lambda-i}{\lambda+i} \in \sigma(U)$.

Proof. 1. $\sigma(T) \subseteq \mathbb{R}$ so $T \pm iI: D(T) \rightarrow H$ are bijective, hence U is well-defined. We also proved $\|(T + iI)x\|^2 = \|Tx\|^2 + \|x\|^2$. Fix $x \in H$ and write $y = (T + iI)^{-1}x$. We have

$$\begin{aligned}\|x\|^2 &= \|(T + iI)y\|^2 = \|Ty\|^2 + \|y\|^2 \\ \|Ux\|^2 &= \|(T - iI)y\|^2 = \|Ty\|^2 + \|y\|^2\end{aligned}$$

so $\|x\| = \|Ux\|$. Since U is a bijection, this implies it's unitary.

2. Fix $x \in H$ and write $y = (T + iI)^{-1}x$. Then $x = (T + iI)y$ and $Ux = (T - iI)y$. Add and subtract these equations to get $(I + U)(x) = 2Ty$ and $(I - U)x = 2iy$. So

- (a) $(I - U)x = 0$ implies $2iy = 0$ so $y = 0$ so $x = 0$.
- (b) $\text{Im}(I - U) = \{(I - U)x \mid x \in H\} = \{2iy \mid y \in D(T)\} = D(T)$.

Finally,

$$i(I + U)(I - U)^{-1}y = i(I + U)\left(\frac{x}{2i}\right) = \frac{1}{2}(I + U)(x) = Ty.$$

3. We have

$$\begin{aligned}T - \lambda I &= i(I + U)(I - U)^{-1} - \lambda I \\ &= i((I + U) + i\lambda(I - U))(I - U)^{-1} \\ &= ((i - \lambda)I + (i\lambda)U)(I - U)^{-1} \\ &= (i + \lambda)\left(U - \frac{\lambda - i}{\lambda + i}I\right)(I - U)^{-1}.\end{aligned}$$

We know $(I - U)^{-1}$ is bijective, so $T - \lambda I$ is bijective iff $U - f(\lambda)I$ is bijective, hence the result. ■

Theorem 5.0.25 (The Spectral Theorem). *Let T be self-adjoint. There's a resolution of the identity E on $\sigma(T) \subseteq \mathbb{R}$ such that $T = \int_{\sigma(T)} \lambda dE$.*

Proof. Let U be the Cayley transform of T and let

$$\tilde{E}: \text{Borel}(\sigma(U)) \rightarrow \mathcal{L}(H)$$

be the spectral resolution of U . We have $\tilde{E}(\{1\}) = 0$ since $1 \notin \sigma_p(U)$ (so that $U - I$ is injective). Take $X = S^1 \setminus \{1\}$, we can think of \tilde{E} as a resolution on X .

Define $g(U) := S = \int_X g d\tilde{E}$ where $g(\lambda) = i\frac{1+\lambda}{1-\lambda}$. We claim $S = T$. We have $g(\lambda)(1 - \lambda) = i(1 + \lambda)$. Applying this to U we get

$$i(I + U) = (g \cdot (1 - \lambda))(U) \supseteq g(U)(1 - \lambda)(U) = S(I - U).$$

This is in fact an equality because

$$D(g(U)(1-\lambda)(U)) = D((g \cdot (1-\lambda))(U)) \cap \subseteq = HD(I-U).$$

So $i(I+U) = S(I-U)$ so for every $x \in D(T)$ we get

$$Sx = i(I+U)(I-U)x = Tx.$$

Hence $S \supseteq T$. But then $S = S^* \subseteq T^* = T$ (where $S = S^*$ since g is real on S^1), so $S = T$.

We therefore have

$$T = \int_{\sigma(S)} g(\lambda) d\tilde{E}.$$

We define

$$\begin{aligned} E: \text{Borel}(\sigma(T)) &\rightarrow \mathcal{L} \\ A &\mapsto \tilde{E}(g^{-1}(A)). \end{aligned}$$

This is the pushforward measure of E by g and one gets the appropriate change of variables

$$T = \int_{\sigma(S)} g d\tilde{E} = \int_{\sigma(T)} \lambda dE. \quad \blacksquare$$

Proof (5.0.22). 1. For $B \in \Sigma$ measurable write $T_B = \int f \cdot \chi_B dE$. We claim

$$\|T_B x\|^2 = \int_B |f|^2 dm_{x,x}.$$

Indeed,

$$\begin{aligned} \|T_B x\|^2 &= \langle T_B x, T_B x \rangle \\ &= \langle T_B^* T_B x, x \rangle \\ &= \left\langle \left(\int \overline{f \chi_B} \cdot f \chi_B dE \right) x, x \right\rangle \\ &= \left\langle \left(\int_B |f|^2 dE \right) x, x \right\rangle \\ &= \int_B |f|^2 dm_{x,x}. \end{aligned}$$

So, if $x \in D(T)$ we have

$$\begin{aligned} \int_{\Omega} |f|^2 &= \lim_{n \rightarrow \infty} \int_{A_n} |f|^2 dm_{x,x} \\ &= \lim_{n \rightarrow \infty} \|T_{A_n} x\|^2 \\ &= \|Tx\|^2 \\ &< \infty. \end{aligned}$$

Conversely, if $\int |f|^2 d\mu_{x,x} < \infty$ then for $n > m$ we have

$$\begin{aligned} \|T_{A_n} - T_{A_m}x\|^2 &= \|T_{A_n \setminus A_m}x\|^2 \\ &= \int_{A_n \setminus A_m} |f|^2 dm_{x,x} \\ &\xrightarrow{n,m \rightarrow \infty} 0. \end{aligned}$$

Hence $(T_{A_n})_{n \in \mathbb{N}}$ is Cauchy, so it converges.

2. We claim something even stronger, that $\text{Im } E(A_k) \subseteq D(T)$. This is enough since for every $x \in D(T)$ we have

$$x = E(\Omega)x = \lim_{k \rightarrow \infty} E(A_k)x.$$

Indeed, if $x \in \text{Im } E(A_k)$ we have

$$\begin{aligned} m_{x,y}(B) &= \langle E(B)x, y \rangle \\ &= \langle E(B)E(A_k)x, y \rangle \\ &= \langle E(B \cap A_k)x, y \rangle \\ &= m_{x,y}(B \cap A_k) \end{aligned}$$

and then for $n > k$ we have

$$\langle T_{A_n}x, y \rangle = \int_{A_n} f dm_{x,y} = \int_{A_k} f dm_{x,y} = \langle T_{A_k}x, y \rangle$$

which implies $T_{A_n}x = T_{A_k}x$, so the limit exists.

3. Fix $y \in D(T^*)$. For every $x \in H$ we want to understand T^*y . It's easier to study its projections, which we study via the inner product. We have

$$\begin{aligned} \langle x, E(A_n)T^*y \rangle &= \langle TE(A_n)x, y \rangle \\ &= \langle T_{A_n}E(A_n)x, y \rangle \\ &= \left\langle \left(\int f \chi_{A_n} dE \right) \left(\int \chi_{A_n} dE \right) x, y \right\rangle \\ &= \langle T_{A_n}x, y \rangle \\ &= \langle x, T_{A_n}^*y \rangle \end{aligned}$$

for all $x \in H$. Hence $E(A_n)T^*y = (T_{A_n})^*y$. Taking $n \rightarrow \infty$ we get

$$\begin{aligned} T^*y &= \lim_{n \rightarrow \infty} (T_{A_n})^*y \\ &= \lim_{n \rightarrow \infty} \left(\int \bar{f} \chi_{A_n} dE \right) y \\ &= \left(\int \bar{f} dE \right) y. \end{aligned}$$

4. (Sketch): Write $S := \int g \, dE$ and $T := \int f \, dE$. Then

$$\begin{aligned} D(\text{LHS}) &= \{x \in H \mid x \in D(S), Sx \in D(T)\} \\ &= D(S) \cap \left\{ x \in H \mid \int |f|^2 \, dm_{Sx, Sx} < \infty \right\} \end{aligned}$$

and when g is bounded we saw that $dm_{Sx, Sx} = g \, dm_{x, x}$ and that $m_{x, y} = \overline{m_{y, x}}$. Hence $dm_{x, Sx} = \bar{g} \, dm_{x, x}$, so $dm_{Sx, Sx} = \bar{g}^2 \, dm_{x, x}$. Hence

$$D(\text{LHS}) = D(S) \cap \left\{ x \mid \int |f|^2 |g|^2 \, dm_{x, x} < \infty \right\} = D(S) \cap D(\text{RHS})$$

. For unbounded g one checks the same via approximation by bounded functions.

Now let

$$A_n = \{x \mid |f|(x) < n\}, B_n = \{x \mid |g|(x) < n\}, C_n = \{x \mid |fg|(x) < n\}.$$

Then

$$\begin{aligned} \left(\int f \, dE \cdot \int g \, dE \right) x &= \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \left(\int_{A_n} f \, dE \cdot \int_{B_m} g \, dE \right) x = \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \left(\int_{A_n \cap B_m} fg \, dE \right) x, \\ \left(\int fg \, dE \right) x &= \lim_{n \rightarrow \infty} \left(\int_{C_n} fg \, dE \right) x. \end{aligned}$$

We now have

$$\begin{aligned} \left\| \left(\int f \, dE \cdot \int g \, dE \right) x - \left(\int fg \, dE \right) x \right\|^2 &= \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \left\| \left(\int_{A_n \cap B_m} fg \, dE - \int_{C_n} fg \, dE \right) x \right\|^2 \\ &= \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \int_{C_n \Delta (A_n \cap B_m)} |fg|^2 \, dm_{x, x} = 0 \end{aligned}$$

where the last equation is by dominant-convergence. \blacksquare

5.0.2 Self-Adjoint Extensions

We have a spectral theorem for self-adjoint operators, but it's generally difficult to construct self-adjoint operators or find out if an operator is self-adjoint. We therefore want a way to extend symmetric operators to self-adjoint ones.

Proposition 5.0.26. *Let T be a symmetric such that $\pm i \notin \sigma(T)$. Then T is self-adjoint.*

Proof. We first show that

$$\left((T - iI)^{-1} \right)^* = (T + iI)^{-1}.$$

Fix $v, w \in H$ and write $x = (T - iI)^{-1}v$ and $y = (T + iI)^{-1}w$. Then

$$\begin{aligned} \langle (T - iI)^{-1}v, w \rangle &= \langle x, (T + iI)y \rangle \\ &= \langle x, Ty \rangle - i \langle x, y \rangle \\ &= \langle Tx, y \rangle - y \langle x, y \rangle \\ &= \langle (T - iI)x, y \rangle \\ &= \langle v, (T + iI)^{-1}w \rangle. \end{aligned}$$

Now, fix $y \in D(T^*)$ and $z = T^*y$. For every $x \in D(T)$ we have

$$\begin{aligned} \langle (T - iI)x, y \rangle &= \langle Tx, y \rangle - i \langle x, y \rangle \\ &= \langle x, z \rangle - i \langle x, y \rangle \\ &= \langle x, z + iy \rangle \\ &= \langle (T - iI)^{-1}(T - iI)x, z + iy \rangle \\ &= \langle (T - iI)x, (T + iI)^{-1}(z + iy) \rangle. \end{aligned}$$

Since $T - iI$ is a bijection, $(T - iI)D(T) = H$ so the above implies $y = (T + iI)^{-1}(z + iy) \in D(T)$. Hence $T = T^*$. \blacksquare

Proposition 5.0.27. *Let T be a symmetric operator on H .*

1. *There exists a closed symmetric operator $\bar{T} \supseteq T$ such that $\Gamma(\bar{T}) = \overline{\Gamma(T)}$. We call \bar{T} the closure of T .*
2. *For every $\lambda \in \mathbb{C} \setminus \mathbb{R}$, $\text{Im}(\bar{T} - \lambda I) = \overline{\text{Im}(T - \lambda I)}$.*

Proof. 1. We should show that $\overline{\Gamma(T)}$ is a graph. Fix $(x, a), (x, b) \in \overline{\Gamma(T)}$, we should show that $a = b$. Fix two sequences $(x_n)_{n \in \mathbb{N}}, (x'_n)_{n \in \mathbb{N}} \subseteq D(T)$ such that $(x_n, Tx_n) \rightarrow (x, a)$ and $(x'_n, Tx'_n) \rightarrow (x, b)$. For $y \in D(T)$ we have

$$\langle T(x_n - x'_n), y \rangle = \langle x_n - x'_n, Ty \rangle.$$

Taking $n \rightarrow \infty$ we get

$$a - b = \langle x - x, Ty \rangle = 0.$$

Hence $a - b \perp y$ for all $y \in D(T)$ so $a - b = 0$, so $a = b$.

Hence \bar{T} is well-defined. We show symmetry, as linearity is easier. Fix $x, y \in D(\bar{T})$. Choose $(x_n)_{n \in \mathbb{N}}, (y_n)_{n \in \mathbb{N}} \subseteq D(T)$ such that $(x_n, Tx_n) \rightarrow (x, \bar{T}x)$ and $(y_n, Ty_n) \rightarrow (y, \bar{T}y)$. Then

$$\langle Tx_n, y_n \rangle = \langle x_n, Ty_n \rangle.$$

Taking $n \rightarrow \infty$ we get $\langle \bar{T}x, y \rangle = \langle x, \bar{T}y \rangle$, as required.

2. Fix $y \in \operatorname{Im}(\bar{T} - \lambda I)$ and write $y = \bar{T}x - \lambda x$. Choose $(x_n)_{n \in \mathbb{N}} \subseteq D(T)$ such that $(x_n, Tx_n) \rightarrow (x, \bar{T}x)$. Then $Tx_n - \lambda x_n \rightarrow y \in \overline{\operatorname{Im}(T - \lambda I)}$.

For the other inclusion, we saw $\operatorname{Im}(\bar{T} - \lambda I)$ is closed. Then $\operatorname{Im}(\bar{T} - \lambda I) \subseteq \overline{\operatorname{Im}(T - \lambda I)}$. ■

Remark 5.0.28. There are non-closable (non-symmetric) operators.

Theorem 5.0.29. Let T be a symmetric operator. Define $n_+ := \operatorname{codim} \overline{\operatorname{Im}(T - iI)}$ and $n_- = \operatorname{codim} \overline{\operatorname{Im}(T + iI)}$. Then T has a self-adjoint extension if and only if $n_+ = n_-$.

Remark 5.0.30. 1. $\operatorname{codim}(\overline{\operatorname{Im}(T - \lambda I)})$ is constant on the upper/lower half-planes. So, it doesn't matter what two numbers we take in the upper/lower half-planes.

2. We can have n_+, n_- infinite, so we treat n_+, n_- as cardinal dimensions.

Proof. Assume $n_+ = n_-$. Without loss of generality assume T is closed. Define $H_1 = \operatorname{Im}(T + iI)$ and $H_2 = \operatorname{Im}(T - iI)$, which are both closed subspaces of H . We define $U: H_1 \rightarrow H_2$ by $U = (T - iI)(T + iI)^{-1}$, which we call the Cayley transform of T as before. This is a norm-preserving bijection. Note $H_1^\perp \cong H/H_1 \cong H/H_2 \cong H_2^\perp$ since Hilbert spaces of the same dimension are unique up to an isometric isomorphism. Hence there's an isometric isomorphism $\tilde{U}: H_1^\perp \rightarrow H_2^\perp$. Define $V: H \rightarrow H$ by $V(x + y) := Ux + \tilde{U}(y)$ where $x \in H_1$ and $y \in H_1^\perp$, which is a bijection since U, \tilde{U} are, and is an isometry since

$$\|V(x + y)\|^2 = \|Ux\|^2 + \|\tilde{U}y\|^2 = \|x\|^2 + \|y\|^2 = \|x + y\|^2$$

for all $x, y \in H$. Hence V is unitary.

Define $S = i(I + V)(I - V)^{-1}$ which we show is a self-adjoint extension of T . Note that

$$\ker(I - V) = \ker(VV^* - V) = \ker(V(V^* - I)) = \ker(V^* - I)$$

where the last equality is since V is injective. Then

$$\ker(I - V) = \ker(V - I)^* = \operatorname{Im}(V - I)^\perp \subseteq \operatorname{Im}(U - I)^\perp = D(T)^\perp = \{0\}$$

so $I - V$ is injective and S is well-defined.

We first show S is symmetric.

$$\begin{aligned}
\langle Sx, y \rangle &= i \left\langle (I + V) (I - V)^{-1} x, y \right\rangle \\
&= i \left\langle x, \left((I - V)^{-1} \right)^* (I + V)^* y \right\rangle \\
&= i \left\langle x, (I - V^{-1})^{-1} (I + V^{-1}) y \right\rangle \\
&= i \left\langle x, (I - V^{-1})^{-1} V^{-1} V (I + V^{-1}) y \right\rangle \\
&= i \left\langle x, (V - I)^{-1} (V + I) y \right\rangle \\
&= \langle x, i (I + V) (I - V) y \rangle \\
&= \langle x, Sy \rangle.
\end{aligned}$$

We now want to show S is self-adjoint. We proved that $\sigma(S) = g(\sigma(V))$ where $g(\lambda) = i \frac{1+\lambda}{1-\lambda}$. Since V is unitary, $\sigma(V) \subseteq S^1$ so $\sigma(S) \subseteq g(S^1) = \mathbb{R}$. Hence S is self-adjoint.

For the other direction, assume S is self-adjoint and $S \supseteq T$. Let V be the Cayley transform of S . We know $V(H_1) = H_2$ so $V(H_1^\perp) = H_2^\perp$, so $\dim H_1^* = \dim H_2^\perp$. ■

Fact 5.0.31. Consider $H = L^2(I)$ for $I \subseteq \mathbb{R}$ an interval. Let $g: I \rightarrow \mathbb{C}$ be continuous, and consider

$$E = \left\{ f \in H \mid f \text{ is compactly-supported and } \int_I f \bar{g} = 0 \right\}.$$

If $g \in H$ then $\bar{E} = \text{Span}\{g\}^\perp$, and if $g \notin H$ then $\bar{E} = H$.

Example 5.0.32. Let $H = L^2(\mathbb{R})$ and let $Tf = if'$ on

$$D(T) = \{f \in H \mid f \text{ is smooth and compact-supported}\}.$$

We know T is symmetric (by integration by parts).

We compute $\overline{\text{Im}(T - iI)}$. Given some $g \in H$, the solution to $if' - if = g$ is

$$f(t) = Ce^t - ie^t \int_{-\infty}^t e^{-s} g(s) ds.$$

For f to be in $D(T)$ we need g to be compactly-supported and smooth, and we need $C = 0$ and that $\int_{-\infty}^{\infty} e^{-s} g(s) ds = 0$. Since $e^{-s} \notin L^2(\mathbb{R})$ and by the fact, $\overline{\text{Im}(T - iI)} = H$, so $n_+ = 0$.

Similarly, if we solve $if' + if = g$ we get $f(t) = Ce^{-t} - ie^{-t} \int_{-\infty}^t e^s g(s) ds$ and $e^s \notin L^2(\mathbb{R})$ so $\overline{\text{Im}(T + iI)} = H$ in the same when. Then $n_- = 0$.

Then \bar{T} is self-adjoint (in which case we call T essentially self-adjoint).

Example 5.0.33. Let $H = L^2((0, \infty))$ and $Tf = if'$ with

$$D(T) = \{f \in H \mid f \text{ is smooth and supported on } [a, b] \subseteq (0, \infty)\}.$$

In contrast with the above example, here $e^{-s} \in L^2(0, \infty)$ so $\overline{\text{Im}(T - iI)} = \text{Span}\{e^{-s}\}^\perp$ so $n_+ = 1$. But, $e^s \notin L^2(0, \infty)$ so $\overline{\text{Im}(T + iI)} = H$, so $n_- = 0$, so there's no self-adjoint extension to T .

Example 5.0.34. Let $H = L^2(0, 1)$ and $Tf = if'$ with

$$D(T) = \{f \in H \mid f \text{ is supported in } [a, b] \subseteq (0, 1)\}.$$

We have $e^s, e^{-s} \in L^2(0, 1)$ so $n_+ = n_- = 1$. So \bar{T} is not self-adjoint, but there is a self-adjoint extension.

Proposition 5.0.35. Let $H = L^2(\Omega)$ and assume T is symmetric and $T\bar{f} = \overline{Tf}$ (with \bar{f} the complex conjugate of f , and so that $f \in D(T) \iff \bar{f} \in D(T)$). Then T has a self-adjoint extension.

Proof. Consider $C: H \rightarrow H$ by $Cf = \bar{f}$ which is a skew-linear bijective isometry mapping subspaces to subspaces. We have

$$C(\text{Im}(T - iI)) = \{\overline{Tf - if} \mid f \in D(T)\} = \{T\bar{f} + i\bar{f} \mid f \in D(T)\} = \{Tf + if \mid f \in D(T)\} = \text{Im}(T + iI).$$

This implies $C(\overline{\text{Im}(T - iI)}) = \overline{\text{Im}(T + iI)}$, so $n_+ = n_-$. ■

Example 5.0.36. Let $H = L^2(0, \infty)$ and $Tf = -f''$ which is symmetric, with domain

$$D(T) = \{f \in H \mid f \text{ is smooth and compactly supported on } (0, \infty)\}.$$

(This is the square of a previous example). Then T has a self-adjoint extension.

Theorem 5.0.37 (Friedrich's Extension Theorem). Assume T is symmetric and $\langle Tx, x \rangle \geq C\|x\|^2$ for some $c \in \mathbb{R}$ and all $x \in H$. Then T has a self-adjoint extension.

Remark 5.0.38. This is a useful theorem e.g. in the theory of elliptic PDE's.

Proof (Sketch). Without loss of generality assume $c = 1$. Define on $D(T)$ the product $\langle x, y \rangle_T = \langle Tx, y \rangle = \langle x, Ty \rangle$ which is an inner product with $\|x\|_T \geq \|x\|$. $(D(T), \|\cdot\|_T)$ is not complete. Find $D(T) \subseteq H_T \subseteq H$ which is the completion of $(D(T), \|\cdot\|_T)$. Note that $\langle \cdot, \cdot \rangle_T$ is defined on H_T even though T isn't.

Fix $y \in H$ and consider $\psi_y(z) = \langle z, y \rangle$ which is a continuous linear functional on H in the usual norm. If $z \in H_T$ we have

$$|\psi_y(z)| = |\langle z, y \rangle| \leq \|z\| \|y\| \leq \|z\|_T \|y\|,$$

so $\psi_y \in H_T^*$. There exists $x := x(y) \in H_T$ such that

$$\psi_y(z) = \langle z, y \rangle = \langle z, x \rangle_T.$$

The map $y \mapsto x(y)$ is well-defined and injective. Define

$$D(S) := \{x(y) \mid y \in H\}, \quad S(x(y)) = y.$$

If $x \in D(T)$ we have

$$\langle z, Sx \rangle = \langle z, x \rangle_T = \langle z, Tx \rangle$$

so $Tx = Sx$. S is symmetric since $\langle \cdot, \cdot \rangle_T$ is symmetric. Then $0 \notin \sigma(S)$ so S is self-adjoint (if it isn't there's a point in the spectrum in the upper/lower half-plane, then all the half plane is in the spectrum, but since the spectrum is closed we have $\sigma(S) \subseteq \mathbb{R}$). ■