Math 245B Lecture Notes

Professor: Tim Austin
Scribe: Daniel Raban

Contents

1 Forms of The Axiom of Choice

1.1 The axiom of choice

We will need the axiom of choice later, so we will begin the course by introducing it now.

Definition 1.1. Let A be a nonempty set, and let X_α be a set for each $\alpha \in A$. The **Cartesian product** $\prod_{\alpha \in A} X_\alpha = \{\langle x_\alpha \rangle_{\alpha \in A} : a_\alpha \in X_\alpha \forall \alpha \in A\}$ is a function $A \to \bigcup_{\alpha \in A} X_\alpha$ such that $\alpha \mapsto x_\alpha$..

Definition 1.2. The Axiom of choice says that if $X_\alpha \neq \varnothing$ then $\prod_{\alpha \in A} X_\alpha \neq \varnothing$.

Theorem 1.1 (Cohen)**.** *The axiom of choice is not implied by the other standard axioms of set theory.*

This is difficult to apply, but we will use provably equivalent statements.

1.2 Posets and Zorn's lemma

Let X be a set.

Definition 1.3. A **partial order** on X is a relation "$\leq$" on X that is

1. Transitive: if $a \leq y$ and $y \leq z$, then $x \leq z$
2. Reflexive $x \leq x$ for all $x \in X$
3. Anti-symmetric: if $x \leq y$ and $y \leq x$, then $x = y$

Definition 1.4. A **total order** is a partial order where for all $x, y \in X$, either $x \leq y$ or $y \leq x$.

Example 1.1. Let S be a set and let $\mathscr{P}(S)$ be the set of subsets of S¿ Then $\subseteq$ is a partial order on $\mathscr{P}(S)$.

Example 1.2. On $\mathbb{R}$, $\leq$ is a partial order (and in fact a total order).

Example 1.3. Let $U \subseteq \mathbb{R}^2$ be a domain. Say $(x_1, y_1) \leq (x_2, x_2)$ if $y_2 \geq y_1$ and $|y_2 - y_1| \leq |x_2 - x_1|$. This is a partial order but not a total order.

Definition 1.5. Let $(X, \leq)$ be a poset with $U \subseteq X$. An element $x \in U$ is **maximal** if when $y \in U$ and $y \geq x$, we must have $y = x$. An element $x \in X$ is an **upper bound** for U if $x \geq u$ for all $u \in U$.

The definitions of minimal elements and lower bounds are analogous.

Definition 1.6. A **chain** in a partially ordered set $(X, \leq)$ is a subset $Y \subseteq X$ such that for all $y, z \in Y$, either $y \leq z$ or $z \leq y$.

Theorem 1.2 (Hausdorff Maximal Principal)**.** *Any nonempty poset $(X, \leq)$ has a maximal chain $Y \subseteq X$.*

Lemma 1.1 (Zorn)**.** *Let $(X, \leq)$ be a nonempty poset. If every chain in X has an upper bound,, then X has a maimal element.*

1.3 Proof sketch of Zorn's lemma and the Hausdorff maximality principle

Here is another incarnation of the axiom of choice.

Theorem 1.3. *Let $S \neq \varnothing$, and let $\mathcal{F} \subseteq \mathscr{P}(S)$ with $\mathcal{F} \neq \varnothing$. Assume $\mathcal{F}$ is*

1. *down-closed: If $A \subseteq B \in \mathcal{F}$, then $A \in \mathcal{F}$*
2. *chain-closed: If C is a chain with $C \subseteq \mathcal{F}$, then $\bigcup C \in \mathcal{F}$.*

Then $\mathcal{F}$ contains a maximal element.

Here is a sketch of the proof.

Proof. First, $\varnothing \in \mathcal{F}$, so $\mathcal{F} \neq \varnothing$. Assume the result is false. THen for all $A \in \mathcal{F}$, there exists a nonempty $B \in S \setminus A$ such that $A \cup B \in \mathcal{F}$. By property 1, we may assume $|B| = 1$. By the axiom of choice, there exists $f : \mathcal{F} \to S$ such that $f(A) \in S \setminus A$ and $A \cup \{f(A)\} \in \mathcal{F}$.

At this point, the idea is to start at the empty set and keep constructing chains, then taking the union of the chain, and then continuing. This requires a notion of the well-ordering principle, so we will choose a different explanation for our sketch.

Call a subfamily $T \subseteq \mathcal{F}$ a **tower** if

1. $\varnothing \in T$
2. $A \in T \implies A \cup \{f(A)\} \in T$
3. T is chain-closed.

Towers exist (e.g. $\mathcal{F}$). Any intersections of towers is a tower. So there exists a minimal tower $T_{\min}$.

Call $A \in T_{\min}$ a **bottleneck**[1] if $\forall B \in T_{\min}$, either $A \subseteq B$ or $B \subseteq A$. The idea is that the set of bottlenecks is a tower. So $T_{\min}$ is a chain. By property 3, $\bigcup T_{\min} \in T_{\min}$. So by property 2, $\bigcup T \cup \{f(\bigcup T_{\min})\} \in T_{\min}$. This is impossible. $\square$

Here is how we prove the Hausdorff maximal principle:

Proof. Let $(X, \leq)$ be nonempty. Let $\mathcal{F}$ be the set of chains in X. This satisfies the conditions of the theorem, which implies that there exists a maximal chain. $\square$

[1] This is not standard notation.

We can prove Zorn's lemma from this, as well.

Proof. Take an upper bound for a maximal chain Such an element is maximal. □

2 Point Set Topology

2.1 Topological spaces

A metric space defines a collection of open sets. To consider spaces without a metric, we define a collection of open sets with the same properties. This yields a more general theory than the theory of metric spaces.

Definition 2.1. Let X be a set. A **topology** on X is a a collection $\mathcal{T}$ of **open** subsets of X such that

1. $\varnothing, X \in \mathcal{T}$
2. If $\mathscr{A} \subseteq T$, then $\bigcup_{U \in \mathscr{A}} U \in \mathcal{T}$
3. If $U_1, \ldots, U_m \in \mathcal{T}$¡ then $\bigcap_{i=1}^{m} U_i \in \mathcal{T}$.

The pair $(X, \mathcal{T})$ is called a **topological space**.[2]

Definition 2.2. A subset $C \subseteq X$ is **closed** if $X \setminus C$ (denoted C^c) is open.

Example 2.1. Every metric space is a topological space.

Example 2.2. For every set X, $\mathcal{T} = \mathscr{P}(X)$ is called the **discrete topology**.

Example 2.3. For every set X, $\mathcal{T} = \{\varnothing, X\}$ is called the **trivial topology**.

Example 2.4. For every set X, $\mathcal{T} = \{U \subseteq X : U = \varnothing \text{ or } U^c \textbf{ is finite}\}$ is called the **cofinite topology**,

Example 2.5. If $(X, \mathcal{T})$ is a topological space, and $Y \subseteq X$¡ then $\mathcal{T}_Y = \{U \cap Y : U \in \mathcal{T}\}$ is called the **relative topology** of $\mathcal{T}$ on Y.

2.2 Closure and convergence

Let $(X, \mathcal{T})$ be a topological space.

Definition 2.3. If $Y \subseteq X$, then $V \subseteq Y$ is **relatively open (resp. closed)** in Y if $V = U \cap Y$, where U is open (resp. closed).

Definition 2.4. If $A \subseteq X$, $A^o = \bigcup\{U : U \subseteq A, U \text{ open}\}$ is the **interior** of A.

This is the largest open set contained in A.

Definition 2.5. If $A \subseteq X$, $\overline{A} = \bigcap\{C : C \supseteq A, C \text{ closed}\}$ is the **closure** of A.

[2] People usually just refer to X as the topological space when $\mathcal{T}$ is understood.

This is the smallest closed set contained in A.

Definition 2.6. $A \subseteq X$ is **dense** if $\overline{A} = X$.

Definition 2.7. The **boundary** of $A \subseteq X$ is $\partial A; = \overline{A} \setminus A^o$.

Definition 2.8. $A \subseteq X$ is **nowhere dense** if $(\overline{A})^o = \varnothing$.

Definition 2.9. A **neighborhood** of $x \in X$ is any $U \in \mathcal{T}$ such that $x \in U$. A **neighborhood** of $A \subseteq X$ is any $U \in \mathcal{T}$ such that $A \subseteq U$.

Definition 2.10. A **point of closure** of A is a point $x \in X$ such that $U \cap A \neq \varnothing$ for all neighborhoods U of x.

Proposition 2.1. *$\overline{A}$ is the set of points of closure of A.*

Proof. ($\supseteq$): Let x be a point of closure and let $C \supseteq A$. We want to show $x \in C$. If instead $x \in C^c$, then C^c is a neighborhood of x disjoint from A.So x is not a point of closure, which is a contradiction.

($\subseteq$): Let x be a non-point of closure. Then there exists a neighborhood $U \ni x$ such that $U \cap A = \varnothing$. So U^c is closed, $x \notin U^c$, and $U^c \supseteq A$. Then $x \notin \overline{A}$. □

Definition 2.11. Let $(x_n)_{n=1}^{\infty}$ be a sequence in X. Then x_n **converges** to x in $\mathcal{T}$ (written $x_n \to x$) if for every neighborhood U of x, there exists an n_0 such that $x_n \in U$ for all $n \geq n_0$.

Remark 2.1. Here are a few caveats. Convergence does not characterize points of closure like it does for metric spaces. Also, limits of sequences are not necessarily unique in topological spaces.

2.3 Generating topologies and bases

Definition 2.12. If $\mathcal{T}_1, \mathcal{T}_2$ are two topological spaces on X, then $\mathcal{T}_2$ is **stronger** (resp. **weaker**) than $\mathcal{T}_1$ if $\mathcal{T}_2 \supseteq \mathcal{T}_1$ (resp. $\mathcal{T}_2 \subseteq T_1$).

Lemma 2.1. *Any intersection of topologies is a topology.*

Corollary 2.1. *Any $\mathcal{E} \subseteq \mathscr{P}(X)$ generates a topology $\mathcal{T}(\mathcal{E})$.*

In this case, $\mathcal{E}$ is called a **sub-base** for the topology $\mathcal{T}(\mathcal{E})$.

Remark 2.2. Any family generates a unique topology, but a topology may be generated by many different families.

Definition 2.13. A **neighborhood base** at $x \in X$ is a collection $\mathcal{N}$ of neighborhoods of x such that for all neighborhoods $U \ni x$, ther exists a $V \in \mathcal{N}$ such that $x \in V \subseteq U$. A **base** for $\mathcal{T}$ is a family which includes a neighborhood base around every point.

Proposition 2.2. *Let $\mathcal{E} \subseteq \mathcal{T}$. Then $\mathcal{E}$ is a base for $\mathcal{T}$ if and only if every nonempty $U \in \mathcal{T}$ is a union of members of $\mathcal{E}$.*

Proof. ($\implies$): Assume $\mathcal{E}$ is a base, and let $\varnothing \neq U \in \mathcal{T}$. THen for all $x \in U$, there exists a $V_x \in \mathcal{E}$ such that $x \in V_x \subseteq U$. So $U = \bigcup_{x \in U} V_x$.

($\impliedby$): Let $x \in U \in \mathcal{T}$. Then $U = \bigcup_{V \in \mathcal{E}'} V$ for some $\mathcal{E}' \subseteq \mathcal{E}$. So $x \in$ for some $V \in \mathcal{E}'$. Now $x \in V \subseteq U$. □

These two characterizations generalize the notion of open balls in a metric space.

Proposition 2.3. *If $\mathcal{E} \subseteq \mathscr{P}(X)$, then $\mathcal{E}$ is a base for some $\mathcal{T}$ if and only if*

1. *$\bigcup \mathcal{E} = X$,*
2. *For all $U, V \in \mathcal{E}$ and for all $x \in U \cap V$, there exists a $W \in \mathcal{E}$ such that $x \in E \subseteq U \cap V$.*

Proof. ($\implies$): Try doing this direction yourself.

($\impliedby$): Let $\mathcal{T} = \{V \subseteq X : \forall x \in V, \exists U \in \mathcal{E} \text{ s.t. } x \in U \subseteq V\}$. Check that $\mathcal{T}$ is a topology, and then check that $\mathcal{E}$ is a base for $\mathcal{T}$: If $V_1, V_2 \in \mathcal{T}$ and $x \in V_1 \cap V_2$, then there exist $U_1, U_2 \in \mathcal{E}$ such that $x \in U_i \subseteq V_i$ for $i = 1, 2$. By the second property, there exists a $W \in \mathcal{E}$ such that $x \in W \subseteq U_1 \cap U_2 \subseteq V_1 \cap V_2$. So $V_1 \cap V_2 \in \mathcal{T}$. Finally, $\mathcal{E} \subseteq \mathcal{T}$, and the definition of $\mathcal{T}$ means that $\mathcal{E}$ concludes a neighborhood base at every point. □

Unlike with σ-algebras, this means that it is easy to see how we generate a topology.

Corollary 2.2. *If $\mathcal{E} \subseteq \mathscr{P}(X)$, then $\mathcal{T}(\mathcal{E}) = \{\varnothing, X\} \cup \{$unions of finite intersections from $\mathcal{E}\}$.*

Proof. Just show that $\mathcal{T}(\mathcal{E})$ is a topology. □

Example 2.6. $F \subseteq \mathbb{R}^{\mathbb{R}}$. For every $m \in \mathbb{N}$, $t_1, \dots, t_m \in \mathbb{R}$, $x_1, \dots, x_m \in \mathbb{R}$, and $\varepsilon > 0$, define $U(t_1, \dots, t_m, x_1, \dots, x_m, \varepsilon) := \{f \in F : |x_i - f(t_i)| < \varepsilon \, \forall i \leq m\}$. Let E be the set of all such $U(t_1, \dots, t_m, x_1, \dots, x_m, \varepsilon)$. As an exercise, show that this is a base for $\mathcal{T}(\mathcal{E})$. We claim that if $(f_n)_{n \in \mathbb{N}}$ is a sequence in F, then $f_n \to f$ in $\mathcal{T}(\mathcal{E})$ iff $f_n \to f$ pointwise. Next time, we will show that this topology is not defined by a metric.

3 Pointwise Convergence, Countability Axioms, Continuity, and Weak Topologies

3.1 The topology of pointwise convergence

Last time we had the following example of a topology.

Example 3.1. Let $F \subseteq \mathbb{R}^{\mathbb{R}}$, for example, $f = C(\mathbb{R})$. For every $m \in \mathbb{N}$, $t_1, \dots, t_m \in \mathbb{R}$, $x_1, \dots, x_m \in \mathbb{R}$, and $\varepsilon > 0$, define $U(t_1, \dots, t_m, x_1, \dots, x_m, \varepsilon) := \{f \in F : |x_i - f(t_i)| < \varepsilon \forall i \leq m\}$. Let $\mathcal{E}$ be the set of all such $U(t_1, \dots, t_m, x_1, \dots, x_m, \varepsilon)$. We claim that if $(f_n)_{n \in \mathbb{N}}$ is a sequence in F, then $f_n \to f$ in $\mathcal{T}(\mathcal{E})$ iff $f_n \to f$ pointwise. Next time, we will show that this topology is not defined by a metric.

Proposition 3.1. *$\mathcal{E}$ is a base for $\mathcal{T}$.*

Proof. We need to check two properties:

1. First, we need $\bigcup \mathcal{E} = F$. Given $f \in F$ and $t_1, \dots, t_m \in \mathbb{R}$, let $x_i = f(t_i)$. Then $f \in U(t_1, \dots, t_m, x_1, \dots, x_m, \varepsilon)$ for all $\varepsilon > 0$.

2. Let $U(t_1, \dots, t_m, x_1, \dots, x_m, \varepsilon), U(s_1, \dots, s_n, y_1, \dots, y_n, \delta) \in \mathcal{E}$. Consider f in their intersection. Choose η so smal that $(f(t_i) - \eta, f(t_i) + \eta) \subseteq (x_i - \varepsilon, x_i + \varepsilon)$ for all i and same for δ. Now $f \in U(t-1, \dots, t_m, s_1, \dots, s_n, f(t_1), \dots, f(t_m), f(s_1), \dots, f(s_n), \eta)$, which is contained in the intersection of the first two sets. $\square$

Proposition 3.2. *Let $(f_n)_{n \in \mathbb{N}}$ be a sequence in F. Then $f_n \to f$ in $\mathcal{T}$ iff $f_n \to f$ pointwise.*

Proof. ($\Longrightarrow$): Pick $t \in \mathbb{R}$ and $\varepsilon > 0$. Consider $U(t, f(t), \varepsilon)$. There exists n_0 such that $f_n \in U(t, f(t), \varepsilon)$ for all $n \geq n_0$; i.e. $|f(t) - f_n(t)| < \varepsilon$ for all $n \geq n_0$.

($\Longleftarrow$): Let $f \in F$, and let U be a neighborhood of f. Because $\mathcal{E}$ is a base, there exists $U(t_1, \dots, t_m, x_1, \dots, x_m, \varepsilon) \subseteq U$ containing f. By shrinking ε if necessary, we may assume that $x_i = f(t_i)$ for every i. We know that $f_n(t_i) \to f(t_i)$. There exists n_0 such that for all $n \geq n_0$, $|f_n(t_i) - f(t_i)| < \varepsilon$ for all $i \leq m$; i.e. $f_n \in U(t_1, \dots, t_m, x_1, \dots, x_m, \varepsilon)$. $\square$

3.2 Countability axioms and metrizability

Definition 3.1. A topology $\mathcal{T}$ on X is **metrizable** if it is generated by a metric on X.

There are natural and important topologies that are not metrizable. This is why we care about point set topology.

Definition 3.2. A topological space $(X, \mathcal{T})$ is **first countable at** x if it has a countable base at x. The space is **first countable** if it is first countable at every x.

Definition 3.3. A topological space $(X, \mathcal{T})$ is **second countable** if it has a countable base.

Definition 3.4. A topological space $(X, \mathcal{T})$ is **separable** if it has a countable dense subset.

Lemma 3.1. *A metrizable space is first countable.*

Proof. Let ρ generate $\mathcal{T}$. Fix $x \in X$. The collection $\{B(x,r) : r > 0, r \in \mathbb{Q}\}$ is a neighborhood base at x. $\square$

Lemma 3.2. *The topology of pointwise convergence on $\mathbb{R}^{\mathbb{R}}$ is not first countable.*

Proof. Suppose $U_1, U_2, \dots$ contain $f \in \mathbb{R}^{\mathbb{R}}$. We may replace if necessary so that $U_j = U(t_1^{(j)}, \dots, t_m^{(j)}, x_1^{(j)}, \dots, x_m^{(j)}, \varepsilon_j)$. Pick $\varepsilon \neq \infty$, and pick $t \in \mathbb{R} \setminus \{t_i^{(j)} : g \geq 1, i = 1, \dots, m_j\}$. Then $U(t, f(t), \varepsilon)$ is not contained in $U(t_1^{(j)}, \dots, t_m^{(j)}, x_1^{(j)}, \dots, x_m^{(j)}, \varepsilon_j)$ for all j. $\square$

Corollary 3.1. *The topology of pointwise convergence is not metrizable.*

3.3 Continuous functions

Definition 3.5. Let $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ be topological spaces, and let $f : X \to Y$. The function f is **continuous at** x if for every neighborhoddo f V of $f(x)$, there exists a neighborhood U of x such that $f[U] \subseteq V$. f is **continuous** if it is continuous at every point.

Proposition 3.3. *$f : X \to Y$ is continuous if and only if $f^{-1}[U] \in \mathcal{T}_X$ for every $U \in \mathcal{T}_Y$.*

Proof. The same proof from metric spaces works here. $\square$

Proposition 3.4. *If $\mathcal{T}_Y = \mathcal{T}(\mathcal{E})$, then $f : X \to Y$ is continuous if and only if $f^{-1}[U] \in \mathcal{T}_X$ for all $U \in \mathcal{E}$.*

Proof. The proof is the same as for the analogous statement for σ-algebras and measurable functions. $\square$

Definition 3.6. Let $K = \mathbb{R}$ or $\mathbb{C}$, and let $(X, \mathcal{T})$ be a topological space. Then $B(X, K)$ is the set of all **bounded functions** $f : X \to K$. $C(X, K)$ is the set of all **continuous functions** $f : X \to K$. $BC(X, K) = B(X, K) \cap C(X, K)$ is the set of **bounded continuous** functions.

Definition 3.7. On $B(X, K)$ or $BC(X, K)$, the **uniform norm** is $\|f\|_u := \sup_{x \in X} |f(x)|$, and the **uniform metric** is $\rho_u(f, g) := \|f - g\|_u$.

Proposition 3.5. *$BC(X, K)$ is complete with the metric ρ_u.*

3.4 The weak and product topologies

Definition 3.8. Let X be a set, let $((Y_\alpha, \mathcal{T}_\alpha))_{\alpha \in A}$ be topological spaces, and let $f_\alpha : X \to Y_\alpha$ for all $\alpha \in A$. The **weak topology** generated by the f_α is $\mathcal{T}\left(\bigcup_{\alpha \in A}\{f_\alpha^{-1}[U] : U \in \mathcal{T}_\alpha\}\right)$.

Definition 3.9. Let $((Y_\alpha, \mathcal{T}_\alpha))_{\alpha \in A}$ be topological spaces, let $X := \prod_{\alpha \in A} Y_\alpha$, and let $\pi_\alpha : X \to Y_\alpha$ send $(x_\beta)_{\beta \in A} \mapsto x_\alpha$ for all $\alpha \in A$. The **product topology** on X is the weak topology generated by $(\pi_\alpha)_{\alpha \in A}$.

The collection $\{\pi_\alpha^{-1}[U] : \alpha \in A, U \in \mathcal{T}_\alpha\}$ is a subbase for this topology. The collection $\{\bigcap_{j=1}^n \pi_{\alpha_j}^{-1}[U_{\alpha_j}] : \alpha_1, \cdots \alpha_n \in A, U_{\alpha_j} \in \mathcal{T}_{\alpha_j}\}$ is a base for this topology. Our previous topology on $\mathbb{R}^\mathbb{R}$ was actually the product topology.

4 Separation Axioms and Urysohn's Lemma

4.1 Second countability and separability

Proposition 4.1. *Every second countable topological space is separable. In metric spaces, the converse is true.*

Proof. Let E be a countable base for $\mathcal{T}$. Pick one point $x_U \in U$ for all $U \in \mathcal{E}$. Now $\{x_U : U \in E\}$ is dense. Let $y \in X$, let V be a neighborhood of y. Now $V = \bigcup_{U \in \mathcal{E}'} U$ for some $\mathcal{E}' \subseteq \mathcal{E}$. If $U \in \mathcal{E}'$, then $u_U \in V$, so $V \cap \{x_U\} \neq \varnothing$.

Let (X, ρ) be a separable metric space, and let $A \subseteq X$ be countable and dense. Let $\mathcal{E} = \{B_r(x) : x \in A, r > 0, r \in \mathbb{Q}\}$. Check that this is a base:

1. $\bigcup_{x \in A} B_1(x) = X$ by the density of A.

2. Let $x, y \in A$, $r, s \in \mathbb{Q} \cap (0, \infty)$. Let $z \in B_r(x) \cap B_s(y)$. Pick $\delta > 0$ with $\delta \in \mathbb{Q}$ such that $B_{2\delta}(z) \subseteq B_r(x) \cap B_s(y)$. Let $w \in A$ be such that $\rho(z, w) < \delta$. Now $B_\delta(w) \ni z$, and $B_\delta(w) \subseteq B_{2\delta}(z) \subseteq B_r(x) \cap B_s(y)$. $\square$

The reverse implication is not true in general, but you have to deal with a complicated set theory construction.

4.2 Separation axioms

Definition 4.1. A topological space $(X, \mathcal{T})$ has property

1. T_0: For all $x, y \in X$ with $x \neq y$, there exists $U \in \mathcal{T}$ such that $|U \cap \{x, y\}| = 1$.

2. T_1: For all $x, y \in X$ with $x \neq y$, there exists $U \in \mathcal{T}$ such that $U \cap \{x, y\} = \{x\}$.

3. T_2 (**Hausdorff property**): If $x \neq y \in X$, there exist $U \neq V \in \mathcal{T}$ such that $U \cap T = \varnothing$, $x \in U$, and $y \in V$.

4. T_3 (**regular**): T_1 and whenever $x \in X$ and $A \subseteq X$ is closed, there exist $U, V \in \mathcal{T}$ such that $U \cap V = \varnothing$, $x \in U$, and $A \subseteq V$.

5. T_4 (**normal**): T_1 and whenever $A, B \subseteq X$ are closed and disjoint, there exist $U, V \in \mathcal{T}$ open such that $U \cap V = \varnothing$, $A \subseteq U$, and $B \subseteq V$.

Proposition 4.2. $(X, \mathcal{T})$ *is* T_1 *if and only if singletons are closed sets.*

Proof. ($\implies$): Let $\{x\} \in X$. If $y \in X \setminus \{x\}$, then by T_1, there exists $U)y \in \mathcal{T}$ such that $y \, inU_y$, $x \notin U_y$. Now $X \setminus \{x\} = \bigcup_y U_y \in \mathcal{T}$.

($\impliedby$): If $x \neq y \in X$, then $X \setminus \{x\}$ is open and contains $y$but not x. $\square$

Corollary 4.1. $T_4 \implies T_3 \implies T_2 \implies T_1 \implies T_0$.

Lemma 4.1. *Any metric space $(X, \mathcal{T})$ is T_4.*

Proof. Assume $A, B \neq \varnothing$. $A \subseteq \{x : \rho(x, A) < \rho(x, B)\}$, and $B \subseteq \{x : \rho(x, A > \rho(x, B)\}$, where $\rho(x, A) = \inf_{y \in A} \rho(x, y)$. The function $x \mapsto \rho(x, A)$ is continuous, so these are open sets. $\square$

4.3 Urysohn's lemma

Lemma 4.2 (Urysohn)**.** *Let $(X, \mathcal{T})$ be T_4 and let $A, B \subseteq X$ be disjoint and closed. Then there exists $f \in C(X, [0, 1])$ such that $f|_A = 0$ and $f|_B = 1$.*

Remark 4.1. In metric spaces, we can just use the function

$$f(x) = \frac{\rho(x, A)}{\rho(x, A) + \rho(x, B)}.$$

The converse is true, as well; it is much easier to prove.

The idea is that if you had f, you could construct level sets (like when $f = 1/3$). We try to reconstruct f using its level sets.

Lemma 4.3. *Let $\Delta = \{k/2^n : n \geq 1, 0 < k < 2^n\}$. Then there exists $\{U_r : r \in \Delta\} \subseteq \mathcal{T}$ such that*

1. *$A \subseteq U_r \subseteq B^c$ for all r.*
2. *If $r < s \in \Delta$, then $\overline{U}_r \subseteq U_s$.*

Proof. We want to find $U_{1/2}$ such that $A \subseteq U_{1/2}$ and $\overline{U}_{1/2} \subseteq B^c$. By T_4, there exist $U \supseteq A$ and $V \supseteq B$ such that $U \cap V = \varnothing$; i.e. $U \subseteq V^c \subseteq B^c$, so $\overline{U} \subseteq B^c$. Now let $U_{1/2} := U$.

Suppose we have U_r for $r = k/2^n$ $n = 1, \ldots, m-1$. Consider U_s, where $s = d/2^n$. Let $r_1 = (\ell - 1)/2^n$,and $r_2 = (\ell + 1)/2^n$. Repeat the previous construction with the closed sets $\overline{U}_r$ and $U_{r_2}^c$. This gives us U_s. $\square$

We can now prove Urysohn's lemma.

Proof. Let $\{U_r : r \in \Delta\}$ be given by the lemma. Define

$$f(x) := \begin{cases} \inf\{r \in \Delta : x \in U_r\} & \exists r \in \Delta \text{ s.t. } x \in U_n \\ 1 & x \notin \bigcup_{r \in \Delta} U_r \end{cases}$$

Suppose $x \in X$, $0 < f(x) < 1$. Let $\varepsilon > 0$. Choose $r_1 < r_2 \in \Delta \cap (f(x) - \varepsilon, f(x))$, $s \in \Delta \cap (f(x), f(x) + \varepsilon)$. Now $x \in U_{r_2}^c \supseteq (\overline{U}_{r_1})^c$, but $x \in U_s$. Now $U_s \cap (\overline{U}_{r_1})^c$ us a neighborhood V of x and $f[V] \subseteq (f(x) - \varepsilon, f(x) + \varepsilon)$. $\square$

Here is the final (but not very useful) separation axiom.

Definition 4.2. A topological space $(X, \mathcal{T})$ is $T_{3\,1/2}$ if for all $x \in X$ and $A \subseteq X$ closed with $x \notin A$, there exists $f \in C(X, [0,1])$ such that $f(x) = 0$ and $f|_A = 1$.

This is a weakened version of the condition Urysohn's lemma which is weaker than T_4 and stronger than T_3.

Next time, we will use Urysohn's lemma to prove the following result.

Theorem 4.1 (Tietze's extension theorem)**.** *Let $(X, \mathcal{T})$ be T_4, let $A \subseteq X$ be closed, and let $f \in C(A, [a,b])$. Then there exists $F \in C(X, [a,b])$ such that $F|_A = f$. The same holds if $C(X, [a,b])$ is replaced with $C(X, K)$, where $K = \mathbb{R}$ or $\mathbb{C}$.*

5 Tietze's Extension Theorem and Compactness

5.1 Tietze's extension theorem

Let X be a normal topological space.

Theorem 5.1 (Tietze's extension theorem)**.** *Let $(X, \mathcal{T})$ be T_4, let $A \subseteq X$ be closed, and let $f \in C(A, [a, b])$. Then there exists $F \in C(X, [a, b])$ such that $F|_A = f$. The same holds if $C(X, [a, b])$ is replaced with $C(X, K)$, where $K = \mathbb{R}$ or $\mathbb{C}$.*

Proof. Without loss of generality, translate so that $a = 0$. We claim that if $f \in C(A, [0, b])$, then there exists $g \in C(X, [0, b/3])$ such that $0 \leq f - g \leq 2b/3$. Let $B = \{x \in A : f(x) \leq b/3\}$, and let $C = \{x \in A : f(x) \geq 2b/3\}$. These are relatively closed in A, and since A is closed, they are closed in X. By Urysohn's lemma, there exists $g \in C(X, [0, b/3])$ such that $g|_B = 0$ and $g|_C = b/3$. Now check that

1. $g|_A \leq f$,
2. $f \leq g|_A + 2b/3$.

Let g_1 be given by the claim, and let $f_1 = f - g_1|A$. Apply the claim again. There exists $g_2 \in C(X, [0, 2/3 \cdot b/3])$ such that $0 \leq f_1 - g_2|_A \leq (2/3)^2 b$. By recursion, we find $g_n \in C(X, [0, (2/3)^{n-1} \cdot b/3])$, and $f - (\sum_{i=1}^n g_i)|_A \leq (2/3)^n b$. Now, for any $m \geq n \geq n$,

$$\left\| \sum_{i=1}^m g_i - \sum_{i=1}^m g_i \right\|_u = \left\| \sum_{i=n+1}^m g_i \right\|_u \leq \sum_{n+1}^m \|g_i\|_u \leq \sum_{i=n+1}^m (2/3)^{i-1} \frac{b}{3} \leq C(2/3)^n b.$$

So $F := \sum_{i=1}^\infty g_i \in C(X, [0, b])$, and if $x \in A$,

$$|f(x) - F(x)| = \lim_{n \to \infty} |f(x) - \sum_{i=1}^n g_i(x)| = 0.$$

Now suppose $f \in C(X, \mathbb{R})$. Consider $f' = f/(1 + |f|) \in C(X, (-1, 1))$. This has an extension $F' \in C(X, [-1, 1])$. Let $H = \{x : F'(x) = \pm 1\}$. This is closed and disjoint from A. So by Urysohn's lemma, there exists $h \in C(X, [0, 1])$ such that $h|_A = 1$ and $h|_H = 0$. Let $G = F' \cdot h$. Now $G \in C(X, (-1, 1))$, and $G|_A = f'$. Now define $F := G/(1 - |G|)$. Then $F \in C(X, \mathbb{R})$ such that $F|_A = f$.

For $X = \mathbb{C}$, split into the real and imaginary parts of f. $\square$

5.2 Compact spaces

Definition 5.1. A topological space X is **compact** if every open cover has a finite subcover. The same is true for a subset of X. A subset $A \subseteq X$ is **precompact** if $\overline{A}$ is compact.

Remark 5.1. The characterization of compactness in metric spaces using sequences turns out to be not as useful in analysis, even though it can be defined in point set topology in general.

Definition 5.2. We say a family $\mathcal{F} \subseteq \mathscr{P}(X)$ has the **finite intersection property (FIP)** if $F_1 \cap \cdots \cap F_m \neq \varnothing$ whenever $m \in \mathbb{N}$ and $F_1, \dots, F_m \in \mathcal{F}$.

Lemma 5.1. *A topological space X his compact if and only if every FIP family of closed sets $\mathcal{F}$ has $\bigcap \mathcal{F} \neq \varnothing$.*

Proof. ($\implies$): Let $\mathcal{F}$ be an FIP family of closed sets. Let $\mathcal{U} = \{X \setminus F : F \in \mathcal{F}\}$. For any $X \setminus F_1, \dots, X \setminus F_m \in U$, we know that there exists $x \in F_1 \cap \cdots \cap F_m$, so $x \notin \bigcup_{i=1}^m (X \setminus F_i)$. So by compactness $\bigcup \mathcal{U} \neq X$. So $\bigcap \mathcal{F} \neq \varnothing$.

($\impliedby$): The reverse implication is just the same steps, but in reverse order. □

Proposition 5.1. *If X is compact and $A \subseteq X$ is closed, then A is compact.*

Proof. Suppose $\mathcal{U}$ is a family of open sets in X such that $A \subseteq \bigcup \mathcal{U}$. Define $\mathcal{V} = \mathcal{U} \cup \{X \setminus A\}$. This is an open cover of X, so it has a finite subcover $U_1 \dots, U_m \in \mathcal{U}$ such that $X = (X \setminus A) \cup \bigcup_{1=1}^m U_i$. So $U_1, \dots, U_m$ form a finite subcover of A. □

5.3 Compact Hausdorff spaces

Some topologies, like the trivial topology, give us undesirable compact spaces. We add the condition of Hausdorff to get spaces we do want.

Proposition 5.2. *Let X be Hausdorff, let $F \subseteq X$ be compact, and let $x \in X \setminus F$. Then there exist disjoint neighborhoods $U \ni x$ and $V \subseteq F$.*

Proof. For all $y\ inF$, we have $y \neq x$, so there exist disjoint open sets $U_y \ni x$ and $V_y \ni y$. Now $F \subseteq \bigcup_y V_y$, so there exist $y_1, \dots, y_m \in F$ such that $F \subseteq V_{y_1} \cup \cdots V_{y_m}$. Now $F \subseteq V : V_{y_1} \cup \cdots V_{y_m}$ is disjoint from $U := U_{y_1} \cap \cdots \cap U_{y_m}$. U is an open neighborhood of x. □

Proposition 5.3. *A compact subset of a Hausdorff space is closed.*

Proof. If F is a compact subset of X, then every $x \in X\ setminusF$ admits an open $U \ni x$ such that $F \cap U = \varnothing$. So $X \setminus F = \bigcup u_x$, so F is closed. □

Proposition 5.4. *A compact Hausdorff space is normal.*

Proof. Let $A, B \subseteq X$ be disjoint and closed. A and B are compact. So for all $x \in A$, there exist disjoint neighborhoods $U_x \ni x$ and $V_x \supseteq B$. Now $\{U_x : x \in A\}$ is an open cover. so there exist $U_{x_1} \cup \cdots, \cup U_{x_n} \supseteq A$ disjoint from $V_{x_1} \cap \cdots \cap V_{x_n} \supseteq B$. □

5.4 Continuous functions on compact spaces

Proposition 5.5. *If X is compact and $f : X \to Y$ is continuous, then $f(X)$ is compact.*

Proof. Let $\mathcal{U}$ be an open cover of $f(X)$. Then $f^{-1}[\mathcal{U}] = \{f^{-1}(U) : U \in \mathcal{U}\}$ is an open cover of X. By compactness, there exists a finite subcover $f^{-1}[U_1], \ldots, f^{-1}[U_m]$. Then $U_1 \cup \cdots \cup U_m \supseteq f(X)$ is an open cover. $\square$

Corollary 5.1. *If $Y = \mathbb{R}$, then extreme values are obtained. So $C(X) = BC(X)$.*

Proposition 5.6. *Let X be compact, Y be Hausdorff, and let $f : X \to Y$ be a continuous bijection. Then f is a homeomorphism; i.e. f^{-1} is also continuous.*

Proof. Let $C \subseteq X$ be closed. Then C is compact, so $f(C)$ is compact. Then $f(C)$ is closed, as Y is Hausdorff. So f sends closed sets to closed sets; i.e. f^{-1} is continuous. $\square$

6 Tychonoff's Theorem

6.1 Locally compact spaces

Sometimes, we want to generalize results for compact spaces to spaces that are not quite compact but can be broken up into compact pieces.

Definition 6.1. A topological space X is **locally compact** if for every $x \in X$ there exists a neighborhood U of x such that $\overline{U}$ is compact.

Example 6.1. $\mathbb{R}^n$ is not compact, but it is locally compact.

6.2 FIP closed families

Theorem 6.1 (Tychonoff). *Suppose $\langle x_\alpha \rangle_{\alpha \in A}$ is a collection of compact sets. Then $\prod_{\alpha \in A} X_\alpha$ is also compact.*

We will prove a special case of this theorem.[3]

Theorem 6.2. *Suppose $\langle X_\alpha \rangle_{\alpha \in A}$ is a collection of compact, Hausdorff sets. Then $\prod_{\alpha \in A} X_\alpha$ is also compact and Hausdorff.*

Recall that we showed last time that X is compact if any FIP family $\mathcal{F}$ of closed sets has $\bigcap \mathcal{F} \neq \varnothing$.

Lemma 6.1. *Let X be a topological space, and let $\mathcal{F}$ be an FIP closed family. Then there exists a maximal FIP closed family $\mathcal{G} \supseteq \mathcal{F}$.*

Proof. Let Γ be the collection of families $\mathcal{G} \subseteq \mathscr{P}(X)$ such that $\mathcal{G}$ consists of closed sets, is FIP, and $\mathcal{G} \subseteq \mathcal{F}$. We will use Zorn's lemma. Let's verify the conditions:

1. $\mathcal{F} \in \Gamma$, so $\Gamma \neq \varnothing$.

2. Every chain $\Lambda \subseteq \Gamma$ has an upper bound. Check that $\bigcup \Lambda \in \Gamma$; the crucial property is that $\bigcup \Lambda$ is FIP. Let $C_1, \dots, C_m \in \bigcup \Lambda$ Then $C_i \in \mathcal{G}_i \in \Lambda$ for all $i = 1, \dots, m$. Then there exists $i_i \leq m$ such that $C_1, \dots, C_m \in \mathcal{G}_{i_0}$. There $C_1 \cap \cdots \cap C_m \neq \varnothing$. □

Remark 6.1. This theorem actually needs Zorn's lemma. It is a theorem from the 1970s that it is possible to choose topological spaces such that Tychonoff's theorem implies Zorn's lemma.

Corollary 6.1. *X is compact if and only if every maximal FIP closed family $\mathcal{F}$ has $\bigcap \mathcal{F} \neq \varnothing$.*

[3] Professor Austin has never seen an application of Tychonoff's theorem where the spaces were not Hausdorff.

Proof. ($\Longleftarrow$): Let $\mathcal{F}$ be an arbitray FIP closed family. If $\mathcal{F} \subseteq \mathcal{G}$, then $\bigcap \mathcal{F} \supseteq \mathcal{G} \neq \varnothing$. $\square$

Lemma 6.2. *Let X be any topological space, and let $\mathcal{F}$ be a maximal FIP closed family.*

1. *If $C_1, \ldots, C_m \in \mathcal{F}$, then $C_1 \cap \cdots \cap C_m F$.*
2. *If $C \subseteq X$ is closed and $C \cap F \neq \varnothing$ for all $F \in \mathcal{F}$, then $C \in \mathcal{F}$.*

Proof. For the first statement, let $\mathcal{F}' = \{C_1 \cap \cdots \cap C_m : m \in \mathbb{N}, C_1, \ldots, C_m, \in \mathcal{F}\}$. Now $\mathcal{F} \subseteq \mathcal{F}'$, so $\mathcal{F} = \mathcal{F}'$.

For the second statement, let $\mathcal{F}'' = \mathcal{F} \cup \{C\}$. This is still FIP: if $D_1, \ldots, D_m \in \mathcal{F}$, then $C \cap (D_1 \cap \cdots \cap D_m) \neq \varnothing$, as $D_1 \cap \cdots \cap D_m$ is in $\mathcal{F}$ by property 1. $\square$

Remark 6.2. If X is compact, then every maximal FIP closed family equals $\{C : C \text{ closed }, C \ni x\}$ for some $x \in X$.

Lemma 6.3. *Let X be a topological space. The following are equivalent:*

1. *X is T_3.*
2. *If $U \subseteq X$ is an open neighborhood of x, then there exists an open $V \ni x$ such that $\overline{V} \subseteq U$.*

Proof. $X \setminus U$ is a closed set not containing x. T_3 is the statement that there is a closed set containing x and an open set in U not intersecting the closed set $X \setminus U$. $\square$

6.3 Proof of Tychonoff's theorem

We are now ready to prove Tychonoff's theorem.

Proof. Suppose X_α is compact, Hausdorff for all $\alpha \in A$. To show that X is Hausdorff, let $x = \langle x_\alpha \rangle, y = \langle y_\alpha \rangle \in C$, with $x \neq y$. So there exists an α such that $x_\alpha \neq y_\alpha$. Then there exist disjoint $U_\alpha \ni x_\alpha$ and $v_\alpha \ni y_\alpha$ (because X_α is Hausdorff). Now $U = \pi_\alpha^{-1}[Y_\alpha] \ni x$, and $V = \pi_\alpha^{-1}[V_\alpha] \ni y$.

To show that X is compact, let $\mathcal{F}$ be a maximal FIP closed family in X.

- Step 1: Find a good point candidate to be in the intersection of elements of $\mathcal{F}$: For all $\alpha \in A$, define $\mathcal{F}_\alpha = \{\pi_\alpha(F) : F \in \mathcal{F}\}$. This collection is FIP;

  $$\pi_\alpha(F_1) \cap \pi_\alpha(F_m) \supseteq \pi_\alpha(F_1 \cap \cdots \cap F_m) \neq \varnothing.$$

  Let $G_\alpha := \{\overline{\pi_\alpha(F)} : F \in \mathcal{F}\}$. This is an FIP closed family. By the compactness of X_α, there exists $x_\alpha \in \bigcap_{\mathcal{G}_\alpha}$. Let $x = \langle x_\alpha \rangle_{\alpha \in A}$.

- Step 2: Let V_α be any open set containing x_α.

$$\begin{aligned} x_\alpha \in \bigcup \mathcal{G}_\alpha &\implies V_\alpha \cap \pi_\alpha(F) \neq \varnothing \quad \forall F \in \mathcal{F} \\ &\implies \overline{V_\alpha} \cap \pi_\alpha(F) \neq \varnothing \quad \forall F \in \mathcal{F} \\ &\implies \pi_\alpha^{-1}[\overline{V_\alpha}] \cap F \neq \varnothing \quad \forall F \in \mathcal{F}. \end{aligned}$$

  So $\pi_\alpha^{-1}[\overline{V_\alpha}] \in \mathcal{F}$ for all α and for all open $V_\alpha \ni x_\alpha$.

- Step 3: Show that $x \in F$ for all $F \in \mathcal{F}$. It is enough to check that $U \cap F \neq \varnothing$ for all $F \in \mathcal{F}$ when U is an open neighborhood of x. It is enough to check for open sets in a neighborhood base. That is, we only need to check $U = \bigcap_{j=1}^m \pi_{\alpha_j}^{-1}[U_{\alpha_j}]$ for all open $U_\alpha \ni x_\alpha$. Because every x_{α_j} is Hausdorff, there exists an open set V_{α_j} such that $x_{\alpha_j} \in V_{\alpha_j} \subseteq \overline{V}_{\alpha_j} \subseteq U_{\alpha_j}$. Now, by step 2, $\pi_{\alpha_j}^{-1}[\overline{V}_{\alpha_j}] \in \mathcal{F}$ for all j. So

$$U \cap F = \left[\bigcap_{j=1}^n \pi_{\alpha_j}^{-1}[U_{\alpha_j}] \right] \cap F \supseteq \left[\bigcap_{j=1}^n \pi_{\alpha_j}^{-1}[\overline{V}_{\alpha_j}] \right] \cap F \neq \varnothing.$$

  So $x \in \bigcap \mathcal{F}$. $\square$

7 The Arzelà-Ascoli Theorem

7.1 Compactness of subsets of $C(X)$

Last time, we proved Tychonoff's theorem, which says that a product of compact spaces is compact. Really, we want to think about $\prod_{\alpha \in A} X_\alpha$ as the set of functions $f : A \to \bigcup_\alpha X_\alpha$ sending $\alpha \mapsto x_\alpha$, where $x_\alpha \in X_\alpha$ for all α. in analysis, finding compactness of spaces and subspaces of functions is very useful and important.

For this lecture, we will assume X is a compact, Hausdorff space. We will let $C(X) = C(X, \mathbb{C})$ (although everything here is true for $\mathbb{R}$ instead of $\mathbb{C}$. We also denote

$$\rho_u(f, g) = \|f - g\|_u = \sup_{x \in X} |f(x) - g(x)|.$$

We know that $(C(X), \rho_u)$ is a complete metric space. This is a big space. We will identify its compact subspaces. Let $\mathcal{F} \subseteq C(X)$. When is it closed and totally bounded? The point of this theorem is to give conditions for totally boundedness.

Definition 7.1. A family $\mathcal{F}$ is **equicontinuous at** $x \in X$ if for any $\varepsilon > 0$, there exists a neighborhood $U \ni x$ such that $|f(x) - f(y)| < \varepsilon$ for all $y \in U$ and for all $f \in \mathcal{F}$. $\mathcal{F}$ is **equicontinuous** if it is equicontinuous at every point.

This is the same neighborhood U for all $f \in \mathcal{F}$.

Definition 7.2. A family $\mathcal{F} \subseteq C(X)$ is **pointwise bounded** if the set $\{f(x) : f \in \mathcal{F}\}$ is bounded for all $x \in X$.

7.2 Statement and proof of the theorem

Theorem 7.1 (Arzelà-Ascoli)**.** *A subspace $\mathcal{F}$ is totally bounded if and only if it is equicontinuous and pointwise bounded. Moreover, $\overline{\mathcal{F}}$ is compact in $(C(X), \rho_u)$.*

Example 7.1. If X has a compact metric ρ, then $\mathcal{F}$ is equicontinuous if for all $x \in$ and $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ for all y such that $\rho(x, y) < \delta$ (for all $f \in \mathcal{F}$).

Example 7.2. Any finite subset of $C(X)$ is equicontinuous.

Example 7.3. Let $X = [-1, 1]$. Let $\mathcal{F}$ be the sequence of functions which are 0 on $[-1, 0]$ and increase continuously to 1 (with steeper and steeper slope). This sequence converges to $\mathbb{1}_{(0,1]}$, which is not in $\mathcal{F}$. This family is not totally bounded, and the ever-increasing steepness at 0 makes this family not equicontinuous.

We will only prove one direction of the equivalence.[4]

[4] Professor Austin says that the other direction is not very useful, in his experience.

Proof. ($\Longleftarrow$): Let $\varepsilon > 0$. We will cover $\mathcal{F}$ with finitely many subsets of ρ_u-diameter $< 4\varepsilon$. For every $x \in X$, there exists a neighborhood $U \ni x$ such that $|f(y) - f(x)| < \varepsilon$ for all $y \in U_x$ and $f \in \mathcal{F}$. By compactness, there exists a finite subcover $X = U_{x_1} \cup \cdots \cup U_{x_m}$. For each $i = 1, \dots, m$, the set $\{f(x_i) : f \in \mathcal{F}\}$ is bounded. So $\bigcup_{i=1}^m \{f(x_i) : f \in \mathcal{F}\}$ is bounded. That is, there exists a finite $B \subseteq \mathbb{C}$ such that for all i and $f \in \mathcal{F}$, there exists a $z \in B$ such that $|f(x_i) - z| < \varepsilon$.

For each $\varphi \in B^m$, define $\mathcal{F}_\varphi = \{f \in \mathcal{F} : |f(x_i) - \varphi(x_i)| < \varepsilon \, \forall i = 1, \dots, m\}$. To finish, condier $f, g \in \mathcal{F}_\varphi$. We know that $|f(x_i) - g(x_i)| < 2\varepsilon$ fpr a;; $i = 1, \dots, n$. For any other $y \in X$, we have $y \in U_{x_i}$ for some i, so

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f(x_i)| + |f(x_i) - g(x_i)| + |g(x_i) - g(y)| \\ &< \varepsilon + 2\varepsilon + \varepsilon \\ &= 4\varepsilon. \end{aligned}$$

So $\|f - g\|_u \leq 4\varepsilon$ for all $f, g \in \mathcal{F}_\varphi$. This gives total boundedness. □

Remark 7.1. $\overline{\mathcal{F}}$ is compact because it is still totally bounded. In general, closure preseves total boundedness.

7.3 Alternate proof of Arzelà-Ascoli

Here is an alternative proof.

Proof. ($\Longleftarrow$): Let $\mathcal{F}$ be equicontinuous and pointwise bounded. Let D_x be a closed, bounded disc in the complex plane containing $\{f(x) : f \in \mathcal{F}\}$.We can think of $\mathcal{F} \subseteq \prod_{x \in X} D_x$, which is a compact product by Tychonoff's theorem. The following lemma (left as an exercise) completes the proof. □

Lemma 7.1. *Let $\mathcal{F}$ be closed and equicontinuous in $C(X)$. Then*

1. *Restricted to $\mathcal{F}$, the uniform topology on $C(X)$ and the product topology on $\prod_{x \in X} D_x$ are the same.*

2. *$\mathcal{F}$ is also closed as a subset of $\prod_{x \in X} D_x$.*

7.4 Arzelà-Ascoli in $\mathbb{R}^n$

In general, we can extend Arzelà-Ascoli to spaces that are not compact but made of countably many compact pieces. Let's see how this works in $\mathbb{R}^n$. Let $(f_n)_{n \in \mathbb{N}} \subseteq C(\mathbb{R}^n)$ and $f \in C(\mathbb{R}^n)$. These are not even necessarily bounded. What is the appropriate notion of convergence for them?

Definition 7.3. The sequence $f_n \to f$ **locally uniformly** if $f_n|_K \to f|_K$ uniformly for all bounded $K \subseteq \mathbb{R}^n$.

Theorem 7.2. *If $(f_n)_{n\in\mathbb{N}} \subseteq C(\mathbb{R}^n)$ is equicontinuous and pointwise bounded, then there exist $f \in C(\mathbb{R}^n)$ and a subsequence $f_{n(k)} \to f$ locally uniformly as $k \to \infty$.*

Proof. For any $r \in \mathbb{N}$, Arzelà-Ascoli gives that $\{f_n : \overline{B_r(0)}\}$ is totally bounded in $C(\overline{B_r(0)})$. By a diagonal argument, there exists a subsequence $(f_{n(k)})_{k=1}^{\infty}$ such that $f_{n(k)}|_{\overline{B_r(0)}}$ converges uniformly to some $f^{(r)} \in C(\overline{B_r(0)})$. By the uniqueness of limits, we have $f^{(r)}_{\overline{B_r(0)}} = f^{(r)}$ for all $r < s$. So there exists a function $f : \mathbb{R}^n \to \mathbb{C}$ such that $f^{(r)} = f|_{\overline{B_r(0)}}$ for all r and $f \in C(\mathbb{R}^n)$. This is the same thing as locally uniform convergence $f_{n(k)} \to f$. $\square$

Remark 7.2. We can think of this proof as Arzelà-Ascoli applied to the image of $\mathcal{F}$ in $\prod_{r=1}^{\infty} C(\overline{B_r(0)})$.

Example 7.4. In $BC(\mathbb{R})$, consider $\{f_\alpha(x) = e^{i\alpha} : 1 \leq \alpha \leq 2\}$. This is pointwise bounded, and it is uniformly equicontinuous by the mean value theorem. Let $1 \leq \alpha < \beta \leq 2$. Then $\rho_u(f_\alpha, f_\beta) = 2$.

8 The Stone-Weierstrass Theorem

8.1 Algebras of functions

Last time, we characterized compact (in some sense 'small' subsets of $C(X)$). This time, we will characterize larger subsets, in the sense that $\mathcal{A} \subseteq C(X)$ is dense. This also generalizes the classic Weierstrass approximation theorem.

Let X be a compact Hausdorff space. In this lecture, we will denote $C(X) = C(X, \mathbb{R})$.

Definition 8.1. A subset $\mathcal{A} \subseteq C(X)$ **separates points** if for all distinct $x, y \in X$, there exists a function $f \in \mathcal{A}$ such that $f(x) \neq f(y)$.

Definition 8.2. An **algebra** of functions is a linear subspace $\mathcal{A} \subseteq C(X)$ such that if $f, g \in \mathcal{A}$, then $fg \in \mathcal{A}$.

Definition 8.3. A **lattice** of functions is a linear subspace $\mathcal{A} \subseteq C(X)$ such that if $f, g \in \mathcal{A}$, then $\max(f, g), \min(f, g) \in \mathcal{A}$.

Definition 8.4. $\mathcal{A}$ **vanishes** at $x \in X$ if $f(x) = 0$ for all $f \in \mathcal{A}$. $\mathcal{A}$ is **nowhere vanishing** if it does not vanish at any $x \in X$.

This means that for every $x \in X$, there is some $f \in \mathcal{A}$ such that $f(x) \neq 0$.

Theorem 8.1 (Stone-Weierstrass)**.** *Let $\mathcal{A}$ be an algebra, closed under ρ_u, and separate points.*

1. *If $\mathcal{A}$ is nowhere vanishing, then $\mathcal{A} = C(X)$.*
2. *Otherwise, there exists some $x_0 \in X$ such that $\mathcal{A} = \{f \in C(X) : f(x_0) = 0\}$.*

$\mathbb{R}^2$ is an algebra over $\mathbb{R}$ with the multiplication $(x, y) \cdot (u, v) := (xu, yv)$.

Lemma 8.1. *As an algebra over $\mathbb{R}$, the only subalgebras of $\mathbb{R}^2$ are $\{(0,0)\}$, $\{0\} \times \mathbb{R}$, $\mathbb{R} \times \{0\}$, $\{(t,t) : t \in \mathbb{R}\}$, and $\mathbb{R}^2$.*

Proof. Let A be a subalgebra of $\mathbb{R}^2$. We may assume that $\dim(A) = 1$. Let $(x, y) \in A$. Then $x^2, y^2 \in A$. These two ordered pairs must satisfy a linear relation, so $x = 0$, $y = 0$, or $x = y$. $\square$

Remark 8.1. This is a special case of Stone-Weierstrass. If $X = \{1, 2\}$, then $C(X) = \mathbb{R}^2$.

Lemma 8.2. *There exists a sequence $(p_n)_n$ of real polynomial with $p_n(0) = 0$ such that $p_n(t) \to |t|$ uniformly for $t \in [-1, 1]$.*

Proof. Consider the Maclaurin expoansion of $\sqrt{1-s}$, where $0 \leq s < 1$. Apply this, using the fact that $|t| = +\sqrt{1-(1-t^2)}$.[5] $\square$

[5] This is the way Folland proves this lemma. There are lots of equally good ways to prove this.

Lemma 8.3. *Let $\mathcal{A}$ be a closed subalgebra of $C(X)$, Then $f \in \mathcal{A} \implies |f| \in \mathcal{A}$, and $\mathcal{A}$ is a lattice.*

Proof. Let $f \in \mathcal{A}$ with $m := \|f\|_u > 0$. By considering $|f/m| = (1/m)|f|$, we may assume that $m \leq 1$. Let $(p_n)_n$ be given by the previous lemma. Then $p_n \circ f$ converges uniformly to $|f|$ as $n \to \infty$. All the $p_n \circ f$ lie in $\mathcal{A}$, as $\mathcal{A}$ is an algebra. Since $\mathcal{A}$ is closed, $|f| \in \mathcal{A}$.

If $f, g \in \mathcal{A}$, $\max(f, g) = (1/2)|f + g| + (1/2)|f - g|$, and $\min(f, g) = -\max(-f, -g)$. So these are still in $\mathcal{A}$. $\square$

8.2 Proof of the theorem

Now we can prove the theorem.

Proof. Suppose $\mathcal{A} \subseteq C(X)$ is a closed lattice that separates points. Also, assume $\mathcal{A}$ is nowhere vanishing.

Step 1: For all $x \neq y \in X$ consider $\mathcal{A}_{x,y} = \{(f(x), f(y)) : f \in \mathcal{A}\}$. Then $\mathcal{A}_{x,y}$ is a algebra of $\mathbb{R}^2$, separating points and nowhere vanishing. So $\mathcal{A}_{x,y} = \mathbb{R}^2$ for all x, y. Thus, for any $f \in C(X)$ and $x, y \in X$, there exists a function $g_{x,y} \in A$ such that $g - x, y(x) = f(x)$ and $g_{x,y}(y) = f(y)$.[6]

Step 2: First, here is the idea: Pin down a point x, and vary y. Each $g_{x,y}$ agrees with f at at least 2 points. Moreover, $g_x := \max_y g_{x,y}$ must satisfy $g_x(x) = f(x)$ and $g_x \geq f$ everywhere. Then we use compactness to only talk about finitely many points.

Fix $x \in X$. For all $y \in X$, we have $g_{x,y}$, as above. Fix $\varepsilon > 0$. Now there exists an open set $U_y \ni y$ such that $g_{x,y}|_{U_y} > f|_{U_y} - \varepsilon$. By compactness, there exists $X = U_{y_1} \cup \cdots \cup U_{y_m}$. Now let $g_x := ax(g - x, y_1, \ldots, g - x, y_m) \in \mathcal{A}$. We still have $g(x) = f(x)$, and for all $z \in X$¡ there exists an i such that $z \in U_{x,y_i}$. So $g_x(z) \geq g_{x,y_i}(z) > f(z) - \varepsilon$; i.e. $g_x(x) = f(x)$, and $f_x > f - \varepsilon$ everywhere.

Step 3: For every $x \in X$, there exists a neighborhood $V_x \ni x$ such that $g_x|_{V_x} < f|_{V_x} + \varepsilon$. By compactness there exists a finite subcover $X = V_{x_1} \cup \cdots \cup V_{x_m}$. Let $g = \min(g_{x_1}, \ldots, g_{x_m})$. Now $g < f + \varepsilon$ everywhere, and $g > f - \varepsilon$ from step 2.

If $\mathcal{A}$ vanishes at $x_0 \in X$, then it can't vanish anywhere else because it separates points. Rerun the previous proof, just altering Step 1. We know that $\mathcal{A} \subseteq \{f \in C(X) : f(x_0) = 0\}$. If $x \neq y \in X$, and $(x_0) = 0$, then we just need to show that there exists a $g_{x,y} \in \mathcal{A}$ such that $g - x, y = f(x)$ and $g_{x,y} = f(y)$. The proof is the same, except the subalgebra we get is $\mathcal{A}_{x_0,y} = \{0\} \times \mathbb{R}$. $\square$

Theorem 8.2. *Let $\mathcal{B} \subseteq C(X)$ be an algebra that separates points. Then*

1. *If $\mathcal{B}$ is nowhere vanishing, then $\mathcal{B}$ is dense in $C(X)$.*

2. *Otherwise, there exists $x_0 \in X$ such that $\mathcal{B}$ is dense in $\{f \in C(X) : f(x_0) = 0\}$.*

[6] It looks like we are using the axiom of choice here. You don't actually need it for this.

Proof. Let $\mathcal{A} = \overline{\mathcal{B}}$. This is still an algebra, and we can use the other version of the theorem. $\square$

8.3 The complex Stone-Weierstrass theorem

What about $C(X, \mathbb{C})$?

Definition 8.5. A ***-algebra** over $\mathbb{C}$ is an algebra such that $f \in \mathcal{A} \implies \overline{f} \in \mathcal{A}$.

Theorem 8.3 (complex Stone-Weierstrass)**.** *Let $\mathcal{A} \subseteq C(X, \mathbb{C})$ be a closed *-algebra that separates points. Then*

1. *If $\mathcal{B}$ is nowhere vanishing, then $\mathcal{B}$ is dense in $C(X)$.*
2. *Otherwise, there exists $x_0 \in X$ such that $\mathcal{B}$ is dense in $\{f \in C(X) : f(x_0) = 0\}$.*

Example 8.1. What if the algebra is not a *-algebra? Here is a counterexample in this case. Let $X = \{z \in \mathbb{C} : |z| = 1\}$, and let $\mathcal{A}$ be the set of complex polynomials in $C(X, \mathbb{C})$. We cannot approximate $z \mapsto \overline{z}$ by members of $\mathcal{A}$.

9 The Baire Category Theorem

9.1 Statement and proof of Baire's theorem

On $\mathbb{R}$, is $\mathbb{1}_{\mathbb{Q}}$ a pointwise limit of continuous functions? We will be able to answer this question and many more.

Let (X, ρ) be a complete metric space. Recall that $A \subseteq X$ is **nowhere dense** if $(\overline{A})^o = \varnothing$.

Theorem 9.1 (Baire category theorem)**.** *Let $(U_n)_{n=1}^{\infty}$ be a sequence of dense, open subsets of X. Then $\bigcap_{n=1}^{\infty} U_n$ is still dense in X. Equivalently, X is not a countable union of nowhere dense sets.*

Remark 9.1. The 2nd statement is the same statement but taking complements of all the sets involved.

Remark 9.2. This does not hold for uncountable intersections. For example, if $X = [0, 1]$, we can define $U_x = [0, 1] \setminus \{x\}$ for each $x \in X$. Then $\bigcup_x U_x = \varnothing$.

Proof. Let $x \in X$, $\delta > 0$. We must show that $B_\delta(x) \cap \bigcap_{n=1}^{\infty} U_n \neq \varnothing$. First, $B_\delta(x) \cap U_1 \neq \varnothing$; this intersection is open. Pick $B_{2\delta_1}(x_1) \subseteq B_\delta(x) \cap U_1$, where $2\delta_1 \leq \delta$. Then $B_{\delta_1}(x_1) \cap U_2 \neq \varnothing$; this intersection is also open. Pick $B_{2\delta_2}(x_2) \subseteq B_{\delta_1}(x_1) \cap U_2$ such that $2\delta_2 \leq \delta_1$. Continue this recursively. The end result is we get balls $B_{\delta_i}(x_i)$ with $i \geq 1$ such that $B_{2\delta_{i+1}}(x_{i+1}) \subseteq B_{\delta_i}(x_i) \cap U_{i+1}$ and $\delta_{u+1} \leq \delta_i/2$.

This tells us that $\delta_i \leq C/2^i$ for some constant C. We also get that $B_{\delta_j}(x_j) \subseteq B_{\delta_i}(x_i)$ for all $j \geq i$. So $\rho(x_j, x_i) < \delta_i$ for all $j \geq i \geq 1$, which means that the sequence $(x_i)_{i=1}^{\infty}$ is Cauchy. By completeness there exists a limit $y = \lim_i x_i$. Moreover, $y \in \overline{B_{\delta_i}(x_i)} \subseteq B_{2\delta_i}(x_i) \subseteq U_i$ for all i. Similarly, $y \in B_\delta(x)$. $\square$

The same proof gives a slightly more general statement.

Theorem 9.2. *If $V \subseteq X$ is open with $V \neq \varnothing$ and $(U_n)_{n=1}^{\infty}$ is open such that $\overline{U_n \cap V} \supseteq V$ for all n. Then $\overline{(\bigcap_{n=1}^{\infty} U_n) \cap V} \supseteq V$.*

9.2 Meager and residual sets

Definition 9.1. Let X be a topological space. Then $A \subseteq X$ is of **first category**[7] (or **meager**) if it is a bountable union of nowhere dense sets. $A \subseteq X$ is of **second category** (or **non-meager**) otherwise. $A \subseteq X$ is **co-meager** (or **residual**) if A^c is a meager set.

Example 9.1. The ambient space is important. $\mathbb{N}$ is meager inside $\mathbb{R}$ but residual in $\mathbb{N}$.

Corollary 9.1. *Let (X, ρ) be a complete metric space.*

[7] This was Baire's original terminology. I think meager is a much more intuitive term, though.

1. *If $X = \bigcup_{n=1}^\infty F_n$, where each F_n is closed, and $U \subseteq X$ is a nonempty open set, then there exists a nonempty open $V \subseteq U$ and $n \in \mathbb{N}$ such that $V \subseteq F_n^o$.*

2. *If $X = \bigcup_{n=1}^\infty A_n$, and $U \subseteq X$ is a nonempty open set, then there exists a nonempty open $V \subseteq U$ and $n \in \mathbb{N}$ such that $V \subseteq (\overline{A_n})^o$.*

3. *The collection of residual sets is closed under countable intersections.*

4. *The collection of meager sets is closed under countable unions.*

Corollary 9.2. *If $\mathbb{R} \setminus \mathbb{Q} = \bigcup_{n=1}^\infty E_n$, then there exists an n such that $\overline{E_n} \supseteq (a,b)$ for some $a < b$.*

Proof. The given condition implies that $\mathbb{R} = \bigcup_m \{q_m\} \cup \bigcup_{n=1}^\infty E_n$, where $(q_m)_m$ is an enumeration of $\mathbb{Q}$. Now apply Baire category. Either some $\{q_m\}$ or some E_n is dense in some (a,b). This cannot be any of the $\{q_m\}$. $\square$

Corollary 9.3. *$\mathbb{Q}$ is not the countable intersection of dense open sets.*

9.3 The Baire-Osgood theorem

Theorem 9.3 (Baire-Osgood)**.** *Let (X,ρ) be a complete metric space, and let $(f_n)_n \subseteq C(X,\mathbb{R})$ be such that $f_n \to f$ pointwise. Let A be the set of continuity points of f. Then A is residual.*

Proof. Call the **oscillation** of f at $x \in X$

$$\omega_f(x) := \inf_{\delta>0} \sup_{y,z \in B_\delta(x)} |f(y) - f(z)|.$$

f is continuous at A if and only if $\omega_f = 0$. So $X \setminus A = \{x : \omega_f(x) \neq 0\} = \bigcup_{k=1}^\infty \{x : \omega_f \geq 1/k\}$. By the Baire category theorem, it is enough to show that $\{x : \omega_f(x) > \varepsilon\}$ is closed and nowhere dense for all $\varepsilon > 0$.

To show that this is closed, let $x \notin \{x : \omega_f(x) > \varepsilon\}$. That is, let $\omega_f(x) < \varepsilon$. Then there exist $\delta > 0$ and $\varepsilon < \varepsilon$ such that $|f(y) - f(z)| \leq \varepsilon'$ for all $y, z \in B_\delta(x)$. If $x' \in B_{\delta/2}(x)$, then $B_{\delta/2}(x') \subseteq B_\delta(x)$. Then for all $y, z \in B_{\delta/2}(x')$, we have $y, z \in B_\delta(x)$. So $|f(y) - f(z)| \subseteq \varepsilon$; i.e. $\omega_f(x') \leq \varepsilon' < \varepsilon$. So $\{x : \omega_f(x) < \varepsilon\}$ is open.

Now let's show that $\{x : \omega_f(x) > \varepsilon\}$ is nowhere dense. Suppose U is a nonempty open set. Let $E_n := \bigcap_{i,j \geq n} \{x : |f_i(x) - f_j(x)| \leq \varepsilon\}$. These are closed, and $X = \bigcup_n E_n$. Then there exists an n and a nonempty open set $V \subseteq U \cap E_n^o$ containing $B_\delta(x) \ni y, z$. Then $|f_i(y) - f(y)| \leq \varepsilon$ for all $i \geq n$. Also, $\omega_f(x) = 0$, so tjere exists $|delta' < \delta$ such that $|f_i(y) - f_i(z)| \leq \varepsilon$ for all $y, z \in B_{\delta'}(x)$. By the triangle inequality, $|f(y) - f(z)| \leq 3\varepsilon$. This is a contradiction because $|f(y) - f(z)| \geq 4\varepsilon$. $\square$

Corollary 9.4. *$\mathbb{1}_{\mathbb{Q}}$ is not the pointwise limit of continuous funcitons.*

Proof. $\mathbb{1}_{\mathbb{Q}}$ is not continuous anywhere. $\square$

10 Universal Spaces

10.1 Embeddings into generalized cubes

In this lecture, $I = [0,1]$.

Definition 10.1. A **generalized cube** is I^A for some $A \neq \varnothing$, with the product topology.

Definition 10.2. Let X be a topological space. The family $\mathcal{F} \subseteq C(X, I)$ **separates points and closed sets** if for all closed $E \subseteq E$ and $x \in E^c$, there is some $f \in \mathcal{F}$ such that $f(x) \notin \overline{f(E)}$.

The existence of such functions in a T_4 space is given by Tietze's extension theorem.

Definition 10.3. If $\mathcal{F} \subseteq C(X, I)$ separates points and closed sets, then tere exists $\mathcal{G} \subseteq C(X, I)$ such that for all closed $E \subseteq X$ and $x \in E^c$, there exists some $g \in \mathcal{G}$ such that $g(x) = 1$ and $g|_E = 0$.

Proof. For all x, E as above, choose f which separates them; that is, $f(x) \notin \overline{f(E)}$. Then x is contained in an interval disjoint from E, so there exists some piecewise linear bump function φ such that $\varphi(x) = 1$ and $\varphi = 0$ outside of this interval. Then define $f_{x,E,f} = \varphi \circ f$. Let $\mathcal{G} = \{g_{x,E,f} : x, E, f \text{ as above}\}$. □

Definition 10.4. X is **completely regular** if it is T_1 and if for all closed $E \subseteq X$ and $x \in E^c$, there exists some $f \in C(X, I)$ such that $f(x) = 1$ and $f|_E = 0$.

This is sometimes called $T_{3\,1/2}$. So a T_1 space is completely regular if and only if $C(X, I)$ separates points and closed sets.

Definition 10.5. For $\mathcal{F} \subseteq C(X, I)$, the map associated to $\mathcal{F}$ is $e : X \to I^{\mathcal{F}} : x \mapsto (f(x))_{f \in \mathcal{F}}$.

We want to study when this is a homeomorphism.

Proposition 10.1. *Let $X, \mathcal{F}, e$ be as above.*

1. *e is continuous.*

2. *If $\mathcal{F}$ separates points, then e is injective.*

3. *If X is T_1 and $\mathcal{F}$ separates points and closed sets, then e is a homeomorphism $X \to e(X) \subseteq I^{\mathcal{F}}$.*

Proof. The first 2 mostly follow from the construction.

1. A canonical sub-base on $I^{\mathcal{F}}$ is sets of the form $\pi_f^{-1}[U] = \{(x_f)_{f \in \mathcal{F}} L x_f \in U\}$, where $U \subseteq [0,1]$ is open. Now $e^{-1}[\pi_f^{-1}[U]] = f^{-1}[U]$.

2. Let $x \neq y \in X$. Then there exists $f \in \mathcal{F}$ such that $(e_x)_f = f(x) \neq f(y) = (e(y))_f$. So $e(x) \neq e(y)$.

3. We must show that if U is open in X, then $e(U)$ is relatively open in $e(X)$. Pick $x \in U$. We will find an open subset V of $I^{\mathcal{F}}$ such that $e(x) \in V \cap e(X) \subseteq e(U)$; this implies that e^{-1} is continuous for the relative topology. Apply the assumption to x and $E = U^c$. Then there exists $f \in \mathcal{F}$ separating them, so $(e(x))_f \notin \overline{\pi_f(e[E])} = \overline{F(E)}$. Define $V = \{(y_g)_{g \in \mathcal{F}} : y_g \in I \setminus \overline{\pi_f(e[E])}\}$. This is open in $I^{\mathcal{F}}$. Then $e(x) \in V \cap e(X)$ by construction, and $V \cap e[E] = \varnothing$. So $V \cap e[X] \subseteq e[U]$. □

Corollary 10.1. *The following are equivalent:*

1. *X is completely regular.*

2. *X embeds into a cube.*

3. *X embeds into some compact Hausdorff space.*

Proof. (1) $\implies$ (2): Apply the proposition with $\mathcal{F} = C(X, I)$.

(2) $\implies$ (3): Cubes are compact Hausdorff spaces.

(3) $\implies$ (1): We just need that subsets of completely regular spaces are completely regular. Do this as an exercise. □

Corollary 10.2. *Any compact Hausdorff space is homeomorphic to a closed subset of a cube.*

Proof. X embeds into $e[X] \subseteq I^A$ for some A. Since X is compact, $e[X]$ is compact. I^A is Hausdorff, so $e[X]$ is closed. □

10.2 Compactification

In general, we can embed a completely regular space into a cube. Taking its closure, we get a compact, Hausdorff space.

Definition 10.6. A **compactification** of X is a pair (Y, φ), where Y is compact Hausdorff and φ is an embedding $X \to Y$ with $\varphi[X] = Y$.

Example 10.1. $\mathbb{R} \to S^1$ is an embedding. If we add in the extra point, we get a **one-point compactification**.

Example 10.2. $\mathbb{R} \to [-1, 1]$ is an embedding. If we add the endpoints, we can get a two-point compactification.

In general, the compactification $X \to \overline{e[X]} \subseteq I^{(C(X,I)}$ is called the **Stone-Čech compactification**.

$$\begin{tikzcd} X \arrow[r, "e"] \arrow[dr, "\varphi"'] & e[X] \arrow[d, dashed] \\ & Y \end{tikzcd}$$

10.3 Embeddings of compact spaces

Now let (X, ρ) be a compact metric space.

Lemma 10.1. *Compact metric spaces are separable.*

Proof. For all $n \in \mathbb{N}$, there exists a finite $S_n \subseteq X$ such that $\bigcup_{x \in S_n} B_{1/n}(x) = X$. Now $\bigcup_n S_n$ is countable and dense. $\square$

Corollary 10.3. *$C(X)$ is separable.*

Proof. Let $S \subseteq X$ be a countable dense subset. For $y \in S$, let $f_y(x) := \rho(y, x)$. Let $\mathcal{A}_{\mathbb{R}} := \{a_0 + \sum_{i=1}^m a_i f_{y_i} \cdots f_{y_i,m_i} : a_i \in \mathbb{R}, y_{i,j} \in S\}$. This is an algebra, it is nowhere vanishing, and it separates points : if $x \neq z$ in X, there exists $(y_n)_n \in S$ such that $y_n \to x$. So $f_{y_n}(x) \to 0$, and $f_{y-n}(z) \to \rho(x, z) \neq 0$. So $\overline{\mathcal{A}_{\mathbb{R}}}$ by the Stone-Weierstrass theorem, which means that $\overline{\mathcal{A}_{\mathbb{Q}}} = C(X)$. $\square$

Proposition 10.2. *Compact metric spaces embed into $[0,1]^{\mathbb{N}}$.*

Proof. Let $\mathcal{A}$ be some countable dense subset of $C(X, I)$. Then A separates points and closed sets. So $[0,1]^A \cong [0,1]^{\mathbb{N}}$. $\square$

Remark 10.1. We can do this explicitly whenever X is separable. Let $(x_n)_n$ be dense in X. Let $e(x) := (\min\{\rho(x, x_n), 1\})_n \in [0,1]^{\mathbb{N}}$. This is the embedding.

Theorem 10.1 (Urysohn's metrization theorem)**.** *Let X be 2nd countable. Then X is metrizable if and only if it is normal. Equivalently, X embeds into $[0,1]^{\mathbb{N}}$.*

Proof. Here is the idea for showing that normality implies that X embeds into $[0,1]^{\mathbb{N}}$. Let $\mathcal{E}$ be a countable base. Define the countabl collection $\mathcal{F}$ which separates U^c and V^c whenever $U, V \in \mathcal{E}$ and $U^c \cap V^c = \varnothing$. Now apply the embedding construction. $\square$

11 Introduction to Norms and Normed Vector Spaces

11.1 Normed vector spaces

First, here is our notation. We will denote $K = \mathbb{R}$ or $\mathbb{C}$ and $\mathcal{X}$ to be a vector space over K. We will denote $Kx = \{\lambda x : \lambda \in K\}$ and 0 as the origin in K or $\mathcal{X}$. If $\mathcal{M}, \mathcal{N}$ are vect spaces in $\mathcal{X}$, then we denote $\mathcal{M} + \mathcal{N} = \{x + y : x \in \mathcal{M}, y \in \mathcal{N}\}$.

Definition 11.1. A **seminorm** on $\mathcal{X}$ is a function $\|\cdot\| : \mathcal{X} \to [0, \infty)$ such that

1. $\|x + y| \leq \|x\| + \|y\|$ for all $x, y \in \mathcal{X}$
2. $\|\lambda x\| = |\lambda| \|x\|$ for all $x \in \mathcal{X}$ (homogeneous of order 1).

A **norm** is a seminorm such that $\|x\| = 0 \implies x = 0$. A pair $(\mathcal{X}, \|\cdot\|)$ is a **normed vector space**.

The second property of seminorms implies that $\|0\| = 0$.

Definition 11.2. The **norm metric** on $(\mathcal{X}, \|\cdot\|)$ is $\rho(x, y) = \|x - y\|$. This generates the **norm topology**.

This is the kind of definition

Example 11.1. $\mathbb{R}^n$ or $\mathbb{C}^n$ with the Euclidean norm are normed vector spaces.

Example 11.2. The space $BC(X, K)$ with $\|f\|_u := \sup_{x \in X} |f(x)|$.

Example 11.3. The space $\ell^\infty_K = \{(x_n)_{n=1}^\infty \in K^{\mathbb{N}} : \sup_n |x_n| < \infty\}$ is a normed vector space with the norm $\|x\|_\infty = \sup_n |x_n|$. This is actually $BC(\mathbb{N}, K)$.

Example 11.4. Let $(X, \mathcal{M}, \mu)$ be a measure space. Then $L^1_K(\mu)$, the set of measurable functions $f : X \to K$ such that $\|f\|_1 = \int |f| \, d\mu < \infty$, is not a normed vector space. In fact, $\|\cdot\|_1$ is a seminorm, so to get a normed vector space, we need to look at equivalence classes of functions that agree μ-a.e.

Example 11.5. The space $\ell^1_K = \{(x_n)_n \in K^{\mathbb{N}} : \|x_1\| = \sum_n |x_n| < \infty\}$ is a normed vector space.

Example 11.6. $\ell^2_K = \{(x_n)_n \in K^{\mathbb{N}} : \|x\|_2^2 = \sum_n |x_n|^2 < \infty\}$ is a normed vector space. In fact, if we replace 2 by p for $1 \leq p < \infty$, we also get a normed vector space.

11.2 Completeness and convergence

Definition 11.3. A **Banach space** over K is a normed vector space over K which is complete in the norm metric.

All the above examples are Banach spaces.

Example 11.7. Here is an incomplete Banach space.[8] Let $Y = \{x \in \ell^1_K : \exists n_0 \in \mathbb{N} \text{ s.t. } x_n = 0 \, \forall n \geq n_0\}$.

Definition 11.4. A series $\sum_{n=1}^\infty x_n$ in $(\mathcal{X}, \|\cdot\|)$ is **convergent** if there exists some $x \in \mathcal{X}$ such that $\|x - \sum_{n=1}^N x_n\| \to 0$ as $N \to \infty$. It is **absolutely convergent** if $\sum_{n=1}^\infty \|x_n\| < \infty$.

Proposition 11.1. *A normed space $(\mathcal{X}, \|\cdot\|)$ is complete if and only if every absolutely convergent sequence is convergent.*

Proof. ($\implies$): Assume $\mathcal{X}$ is complete. Let $S_N = \sum_{n=1}^n x_n$. Then for $M > N$,

$$\|S_N - S_m\| = \left\| \sum_{n=N+1}^M x_n \right\| \leq \sum_{n=N+1}^M \|x_n\| \xrightarrow{N,M\to\infty} 0.$$

Then S_N is Cauchy, which means it has a limit.

($\impliedby$): Suppose $(x_n)_n$ is Cauchy. Then $\|x_n - x_m\| \to 0$ as $n, m \to \infty$. Pick $n_1 < n_2 < \cdots$ such taht $\|x_n - x_m\| < 2^{i-1}$ for all $n, m \geq n_j$. Define $y_1 = x_{n_1}$ and $y_j = x_{n_j} - x_{n_{j-1}}$ for $j \geq 2$. Note that $\sum_{j=1}^k y_j = x_{n_k}$. Also,

$$\sum_{j=1}^k \|y_j\| = \|x_{n_1}\| + \sum_{j=2}^k \|x_{n_j} - x_{n_{j-1}}\| \leq \|x_{n_1}\| + \sum_{j=2}^\infty 2^{-(j-1)} < \infty.$$

So there exists some $x = \lim_{k\to\infty} \sum_{j=1}^k y_j = \lim_k x_{n_k}$. Then $x_n \to x$. $\square$

Remark 11.1. In this proof, we used a very useful technique: pass to a subsequence to upgrade the convergence to a much faster convergence.

Proposition 11.2. *$L^1_K(\mu)$ is complete.*

Proof. Assume $(f_j)_j \in L^1_{\mathbb{R}}(\mu)$ such that $\sum_j \int |f_j| \, d\mu < \infty$. Let $g_N = \sum_{j=1}^N |f_j|$ be non-negative and increasing in N. By the monotone convergence theorem, there exists some g such that $g = \lim_N g_N$, $g \geq 0$, and $\int g = \lim \int g_N < \infty$. Now if $F_N = \sum_{j=1}^N f_j$, then $|F_N| \leq g$. Moreover, $\sum_{j=N}^M |f_j| \leq g - g_N \to 0$ whenever $g < \infty$ (which holds a.e.). So $F(x) := \lim_N F_N(x)$ exists for a.e. x.By the dominated convergence theorem, we conclude that $\int F_N \, d\mu \to \int F \, d\mu$. Similarly, $|F_N - F| \leq 2g$ be the triangle inequality, and $|F_N - F| \to 0$ pointwise. So by the dominated convergence theorem again, $\int |F_N - F| \to \int 0 = 0$.

The case $K = \mathbb{C}$ is similar. $\square$

[8] If you ever wonder whether a property is using the completeness of a Banach space, try seeing if it still holds in this space.

11.3 Norms over finite dimensional vector spaces

Definition 11.5. Two norms $\|\cdot\|$ and $\|\cdot\|'$ are **equivalent** if there exists some $C \in (0, \infty)$ such that $(1/C)\|x\| \le \|x\|' \le C\|x\|$ for all $x \in \mathcal{X}$.

Theorem 11.1. *If* $\dim(\mathcal{X}) < \infty$*, all norms are equivalent.*

Proof. We will treat the $K = \mathbb{R}$ case; the $K = \mathbb{C}$ case is similar. It is enough to show this when $\mathcal{X} = \mathbb{R}^n$. Let $|\cdot|$ be the Euclidean norm and $\|\cdot\|$ be another norm. We will show that $|\cdot\|$ and $\|\cdot\|$ are equivalent.

Let $e_1, \dots, e_n$ be the standard basis. Then

$$\|x\| = \left\|\sum_{i=1}^n x_i e_i\right\| \le \sum_{i=1}^n |x_i|\|e_i\| \le \left(\sum_i |x_i|^2\right)^{1/2} \left(\sum_i \|e_i\|^2\right)^{1/2}$$

by Cauchy-Schwarz. In fact, this shows that $\|\cdot\|$ is continuous.

To finish, it is enough to show that $\inf\{\|x\| : |x| = 1\} > 0$. But this infimum is achieved at some x such that $|x| = 1$. We must still have $\|x\| > 0$ at this x. $\square$

The proof also showed us the following.

Corollary 11.1. $\|\cdot\|$ *is continuous for the usual topology on* $\mathbb{R}^n$*.*

12 Banach Space Constructions

12.1 Product spaces

Definition 12.1. Let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ and $(\mathcal{Y}, \|\cdot\|_{\mathcal{Y}})$ be normed vector spaces over K. The **Cartesian product** $\mathcal{X} \times \mathcal{Y}$ is a normed space with one of many possible norms:

1. $\|(x,y)\| := \max(|x\|_{\mathcal{X}}, \|y\|_{\mathcal{Y}})$
2. $\|(x,y)\| := |x\|_{\mathcal{X}} + \|y\|_{\mathcal{Y}}$
3. $\|(x,y)\| := \sqrt{\|x\|_{\mathcal{X}}^2 + \|y\|_{\mathcal{Y}}^2}$.

Remark 12.1. There are many natural options for what norm to use; not all of them are listed here. However, from a category theory perspective, none of these are "natural."

Proposition 12.1. *With any of these norms, $\mathcal{X} \times \mathcal{Y}$ is complete if and only if both $\mathcal{X}$ and $\mathcal{Y}$ are complete.*

12.2 Quotient spaces

Definition 12.2. Let $(\mathcal{X}, \|\cdot\|)$ be a normed space over K, and let $\mathcal{M} \subseteq \mathcal{X}$ be a vector subspace. The **quotient space** is $\mathcal{X}/\mathcal{M} = \{x + \mathcal{M} : x \in \mathcal{X}\}$ with the **quotient norm** $\|x + \mathcal{M}\| := \inf\{\|y\| : y \in x + \mathcal{M}\}$.

Lemma 12.1. *If $\mathcal{X}$ is complete and $\mathcal{M} \subseteq \mathcal{X}$ is a closed subspace, then $\mathcal{X}/\mathcal{M}$ is complete.*

Proof. Suppose $(x_n + \mathcal{M})_{n=1}^\infty \in \mathcal{X}/\ cM$ is a sequence such that $\sum_{n=1}^\infty \|x_n + \mathcal{M}\|$. For each n, pick $y_n \in x_n + \mathcal{M}$ such that $\|y_n\| < \|x_n + \mathcal{M}\| + 2^{-n}$. Then $\sum_{n=1}^\infty \|y_n\| < \infty$, so there exists some $y = \sum_{n=1}^\infty y_n \in \mathcal{X}$. So $\|y - \sum_{n=1}^N \| \to 0$ as $N \to \infty$. This is an element of $(y + \mathcal{M}) - \sum_{n=1}^N (y_n + M) = (y + \mathcal{M}) - \sum_{n=1}^N (x_n + M)$. So

$$\left\| (y + \mathcal{M}) - \sum_{n=1}^N (x_n + M) \right\| \le \left\| y - \sum_{n=1}^N \right\| \to 0. \qquad \square$$

12.3 Bounded linear maps

Definition 12.3. A linear map $T : \mathcal{X} \to \mathcal{Y}$ is called **bounded** if there exists some $C < \infty$ such that $\|T_x\|_{\mathcal{Y}} \le C\|x\|_{\mathcal{X}}$ for all $x \in \mathcal{X}$. The **vector space of bounded linear maps** is called $\mathcal{L}(\mathcal{X}, \mathcal{Y})$.

Proposition 12.2. *Let $T : \mathcal{X} \to \mathcal{Y}$ be linear. The following are equivalent:*

1. *T is continuous.*

2. T is continuous at 0.

3. T is bounded.

Proof. (1) $\implies$ (2): This is a special case.

(3) $\implies$ (1): For all $x, x' \in \mathcal{X}$, we have

$$\|Tx - Tx'\|_{\mathcal{Y}} = \|T(x - x')\|_{\mathcal{Y}} \leq C\|x - x'\|_{\mathcal{X}}.$$

(2) $\implies$ (3): For every $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$\|x\|_{\mathcal{X}} < \delta \implies \|T_x\|_{\mathcal{Y}} < \varepsilon.$$

So for all $x \in \mathcal{X} \setminus \{0\}$, let $x' = \frac{\delta}{2\|x\|_{\mathcal{X}}}$. Then $\|x'\|_{\mathcal{X}} < \delta$. Then

$$\|Tx'\|_{\mathcal{Y}} = \frac{\delta}{2\|x\|_{\mathcal{X}}}\|Tx\|_{\mathcal{Y}} < \varepsilon \implies \|T_x\|_{\mathcal{Y}} < \left(\frac{2\varepsilon}{\delta}\right)\|x\|_{\mathcal{X}}.$$

□

Lemma 12.2. *If $S, T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, say with constants C_S, C_T, then $S + T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ with constant at most $C_S + C_T$, and $\lambda S \mathcal{L}(\mathcal{X}, \mathcal{Y})$ with constant $\leq |\lambda| C_S$*

Proof.

$$\|(S + T)x\| \leq \|Sx\| + \|Tx\| \leq (C_S + C_T)\|x\|.$$

□

Definition 12.4. $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ is a normed space with the **operator norm**

$$\|T\|_{\text{op}} = \inf\{C : \|Tx\| \leq C\|x\| \; \forall x \in X\}.$$

Remark 12.2. Equivalently, we can define the operator norm as

$$\begin{aligned}\|T\|_{\text{op}} &= \sup\{C : \|Tx\|_{\mathcal{Y}} : x \in \mathcal{X}, \|x\|_{\mathcal{X}} = 1\}. \\ &= \sup\left\{C : \frac{\|Tx\|_{\mathcal{Y}}}{\|x\|_{\mathcal{X}}} : x \in \mathcal{X} \setminus \{0\}\right\}.\end{aligned}$$

Proposition 12.3. *If Y is complete, so is $\mathcal{L}(\mathcal{X}, \mathcal{Y})$.*

Proof. Let $(T_n)_n$ be Cauchy in $\mathcal{L}(\mathcal{X}, \mathcal{Y})$. Then for all $x \in \mathcal{X}$, we have

$$\|T_n x - T_m x\|_{\mathcal{Y}} \leq \|T_n - T_m\|_{\text{op}}\|x\|_{\mathcal{X}} \xrightarrow{n,m\to\infty} 0,$$

so there exists a $\lim_n T_n x =: Tx$. Now show that $T \in L(\mathcal{X}, \mathcal{Y})$,and $\|T_n - T\|_{\text{op}} \to 0$. □

Remark 12.3. If $S \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ and $T \in \mathcal{L}(\mathcal{Y}, \mathcal{Z})$, then for all $x \in \mathcal{X}$,

$$\|TSx\|_{\mathcal{Z}} = \|T\|\|Sx\|_{\mathcal{Y}} \leq \|T\|\|S\|\|x\|_{\mathcal{X}},$$

so $T \circ S \in \mathcal{L}(\mathcal{X}, \mathcal{Z})$, and $\|T \circ S\| \leq \|S\|\|T\|$. So $L(\mathcal{X}, \mathcal{X})$ is an algebra over $\mathcal{K}$, and it is a Banach algebra if $\mathcal{X}$ is complete.

Definition 12.5. A linear operator $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ is **invertible** (or an **isomorphism**) if T^{-1} exists an and is an element of $\mathcal{L}(\mathcal{Y}, \mathcal{X})$.

12.4 Dual spaces and the Hahn-Banach theorem

Definition 12.6. The space $\mathcal{X}^* := \mathcal{L}(\mathcal{X}, K)$ is the **dual space**. Its norm is called the **dual norm**, and its elements are **bounded linear functionals**.

Theorem 12.1 (Hahn-Banach)**.** *Let $(\mathcal{X}, \|\cdot\|)$ be a normed space, let $\mathcal{M}$ be a linear subspace, and let $f \in \mathcal{M}^*$. Then there exists $F \in \mathcal{X}^*$ such that $F|_{\mathcal{M}} = f$ and $\|F\|_{\mathcal{X}^*} = \|f\|_{\mathcal{M}^*}$.*

We will prove this theorem next time. Instead, let's look at a consequence.

Theorem 12.2. *If $\mathcal{M} \subseteq \mathcal{X}$ is a closed linear subspace and $x \in \mathcal{X} \setminus \mathcal{M}$, then there exists $f \in \mathcal{X}^*$ such that $f|_{\mathcal{M}} = 0$ but $f(x) \neq 0$. Moreover, we can take $\|f\| = 1$ and $f(x) = \inf_{y \in \mathcal{M}} \|x - y\|$.*

Proof. Let $\mathcal{N} = \mathcal{M} + Kx$. Let $\delta = \inf_{y \in M} \|x - y\| = \delta$. Define the function $g : \mathcal{N} \to K$ as $g(y + \lambda x) := 0 + \lambda \delta$. To show that g is well-defined and linear, note that

$$g((y + \lambda x) + (y' + \lambda' x)) = g((y + y') + (\lambda + \lambda')x) = (\lambda + \lambda')\delta.$$

For find the norm of g, we want $|g(y + \lambda x)| \leq \|y + x\|$ for all y, λ. Scaling by a constant, we can assume $\lambda = 1$. Then we want $\delta = |g(y + x)| \leq \|y + x\|$ for all $y \in \mathcal{M}$, which is true by definition. By the Hahn-Banach theorem, g has an extension $f \in \mathcal{X}^*$ with $\|f\| = \|g\| = 1$. $\square$

13 The Hahn-Banach theorem and Dual spaces

13.1 Reflexive spaces and dual spaces

Last time, we showed a consequence of the Hahn-Banach theorem. Here is a special case, where we separate $x \in \mathcal{X}$ from the closed subspace $\mathcal{M} = \{0\}$.

Proposition 13.1. *Let $x \in \mathcal{X} \setminus \{0\}$. Then there exists some $f \in \mathcal{X}^*$ such that $\|f\| = 1$,and $|f(x)| = \|x\|$. Moreover, $\mathcal{X}^*$ separates points of $\mathcal{X}$.*

Proposition 13.2. *If $x \in \mathcal{X}$, define $\hat{x} : \mathcal{X}^* :\to K$ by $\hat{x}(f) := f(x)$. Then $x \mapsto \hat{x}$ is a linear isometry $\mathcal{X} \to \mathcal{X}^{**}$.*

This is called the **canonical embedding**.

Proof. For linearity,

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

Multiplication by constants is the same. To show that it is an isometry,

$$\begin{aligned} \|\hat{x}\|_{**} &= \sup\{\hat{x}(f) : \|f\|_* \leq 1\} \\ &= \sup\{f(x) : \|f\|_* \leq 1\} \\ &\leq \|x\|. \end{aligned}$$

We can achieve equality by the above corollary of Hahn-Banach. □

Remark 13.1. Recall that if $\mathcal{Y}$ is complete, then $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ is complete. Then, since $\mathcal{X}^{**} = \mathcal{L}(\mathcal{X}^*, K)$, and $K = \mathbb{R}$ or $\mathbb{C}$ is complete, $\mathcal{X}^{**}$ is complete. So $\hat{\mathcal{X}} = \{\hat{x} : x \in \mathcal{X}\}$ is a canonical way of completing $\mathcal{X}$.

Definition 13.1. $\mathcal{X}$ is **reflexive** if $\hat{\mathcal{X}} = \mathcal{X}^{**}$.

Finite dimensional vector spaces are always reflexive. This is not always the case for infinite dimensional spaces.

Example 13.1. $C[0,1]$ is not reflexive. We will see this later, but its dual is not separable, and neither is its double dual.

Definition 13.2. The **adjoint** (or **transpose**) of $T : \mathcal{X} \to \mathcal{Y}$ is $T^* : \mathcal{Y}^* \to \mathcal{X}^*$ defined by $T^*(f) = f \circ T$.

$\mathcal{T}^*$ is linear, and it satisfies $\|T^*\| \leq \|T\|$.

Proposition 13.3. *Let $T : \mathcal{X} \to \mathcal{Y}$ be bounded and linear.*

1. $\|T^*\| = \|T\|$.

2. *Let $T^{**} : \mathcal{X}^{**} \to \mathcal{Y}^{**}$ be $T^{**} = (T^*)^*$. Then $T^{**}|_{\hat{\mathcal{X}}} = T$.*

3. *T^* is injective if and only if $T[\mathcal{X}]$ is dense in $\mathcal{Y}$.*

4. *If $T^*[\mathcal{Y}^*]$ is dense in $\mathcal{X}^*$, then T is injective.*

Proof. The verification of these either follows quickly from the definitions or is an application of Hahn-Banach. □

13.2 The Hahn-Banach theorem

For now, $\mathcal{X}$ will be a real vector space.

Definition 13.3. A **sublinear functional** is a function $p : \mathcal{X} \to \mathbb{R}$ such that

1. $p(x + y) \leq p(x) + p(y)$ for all $x, y \in \mathcal{X}$

2. $p(\lambda x) = \lambda p(x)$ for all $\lambda \geq 0$ and $x \in \mathcal{X}$.

Example 13.2. Any seminorm is a sublinear functional.

Theorem 13.1 (Hahn-Banach, general form)**.** *Let $\mathcal{X}$ be a real normed space, let p be a sublinear functional, let $\mathcal{M} \subseteq \mathcal{X}$ be a subspace, and let $f : \mathcal{M} \to \mathbb{R}$ be linear and such that $f(x) \leq p(x)$ for all $x \in \mathcal{M}$. Then there exists a linear functional $F : \mathcal{X} \to \mathbb{R}$ such taht $F|_{\mathcal{M}} = f$ and $F \leq p$.*

Proof. Step 1: Suppose $\mathcal{X} = \mathcal{M} + \mathbb{R}x$, where $x \notin \mathcal{M}$. Then any element of $\mathcal{X}$ is $m + \lambda x$ for some unique $m \in \mathcal{M}$ and $\lambda \in \mathbb{R}$. We want to find $\alpha \in \mathbb{R}$ such that if we set $F(m + \lambda x) := f(m) + \lambda\alpha$, then $F \leq p$; i.e. $f(m) + \lambda\alpha \leq p(m + \lambda x)$ for all $m \in \mathcal{M}$ and $\lambda \in \mathbb{R}$. Equivalently, we want $\lambda\alpha \leq p(m + \lambda x) - f(m)$. We have two cases:

1. If $\lambda > 0$, then this is equivalent to $\alpha \leq p(m + \lambda x)/\lambda - f(m)/\lambda$ for all $m \in M$. That is, $\alpha \leq p(m/\lambda + x) - f(m/\lambda)$. This is equivalent to $\alpha \leq p(m + x) - f(m)$ for all $m \in M$.

2. If $\lambda < 0$, divide by $-\lambda$ and rearrange similarly. We want $-\alpha \leq p(m - x) - f(m)$ for all $m \in \mathcal{M}$. This is $\alpha \geq f(m') - p(m' - x)$ for all $m' \in \mathcal{M}$.

So it remains to show that $f(m') - p(m' - x) \leq p(m + x) - f(m)$ for all $m, m' \in \mathcal{M}$; if we have this, then we can pick any α between these upper and lower bounds. We can rearrange to get $f(m) + f(m') \leq p(m' - x) + p(m + x)$ for all $m, m' \in \mathcal{M}$. But this is

$$\begin{aligned} f(m) + f(m') &= f(m + m') \\ &\leq p(m + m') \\ &= p((m + x) + (m' - x)) \end{aligned}$$

$$\leq p(m' - x) + p(m + x),$$

so we can have the desired α.

Step 2: Here is the general case. Let $\mathcal{E}$ be the collection of pairs $(\mathcal{N}, g)$ such that $\mathcal{N}$ is a subspace of $\mathcal{X}$ containing $\mathcal{M}$,and $g : \mathcal{N} \to \mathbb{R}$ is a linear functional such that $g|_{\mathcal{M}} = f$ and $g \leq p$. Define the partial order $(N, g) \leq (\mathcal{N}', g')$ if $N \subseteq \mathcal{N}'$, and $g'|_{\mathcal{N}} = g$. We will use Zorn's lemma. We wnt to show that every chain $((N_\alpha, g_\alpha))_\alpha$ has an upper bound. Let $\mathcal{N} = \bigcup_\alpha N_\alpha$, and let $g(x) = g_\alpha(x)$ for all $x \in \mathcal{N}_\alpha$. Then $g \leq p$, and $g|_{\mathcal{M}} = g_\alpha|_{\mathcal{M}} = f$. We must show that $\mathcal{N}$ and g are linear. If $x, y \in \mathcal{N}$, then $x \in N_\alpha$ and $y \in \mathcal{N}_\beta$, so $x, y \in N_\alpha$, where $N_\alpha \supseteq N_\beta$. So $x + y \in \mathcal{N}_\alpha < N$. So by Zorn's lemma there exists a maximal element $(\mathcal{N}, g) \in \mathcal{E}$. We must have $\mathcal{N} = \mathcal{X}$ otherwise step 1 contradicts the maximality of $\mathcal{N}$. $\square$

14 Applications of the Baire Category Theorem in Banach Spaces

14.1 The complex Hahn-Banach theorem

Here is a loose end from last time.

Theorem 14.1 (Hahn-Banach, complex version). *Let $(\mathcal{X}, \|\cdot\|)$ be a normed vector space over $\mathbb{C}$, let $\mathcal{M} \subseteq \mathcal{X}$ be a subspace, and let $f \in \mathcal{M}^*$. Then there exists an $F \in \mathcal{X}^*$ such that $F|_{\mathcal{M}} = f$ and $|F| = f$.*

Proof. Define $u = \operatorname{Re}(f)$. Observe that $f(ix) = if(x) = -\operatorname{Im}(f(x)) + i\operatorname{Re}(f(x))$. So $\operatorname{Im}(f) = -\operatorname{Re}(f(i\cdot)) = -u(i\cdot)$. By the real Hahn-Banach theorem, u extends to U, and let $F(x) = U(x) - iU(ix)$. Check that $|F| = |f|$. □

14.2 The open mapping theorem

In finite dimensional vector spaces, linear bijections have linear inverses. Does this still work for normed spaces and bounded linear functions? The answer is no, unless we are dealing with Banach spaces.

Definition 14.1. A function $f : X \to Y$ is called **open** for all open $U \subseteq X$, $f[U]$ is open in Y.

Lemma 14.1. *Let $T : \mathcal{X} \to \mathcal{Y}$ be a linear map between normed spaces. Then T is open if and only if $T[B_{\mathcal{X}}(0,1)] \supseteq B_y(0,r)$ for some $r > 0$.*

Proof. ($\Longrightarrow$): This follows from the definition.

($\Longleftarrow$): Assume the condition holds. Let $U \subseteq \mathcal{X}$ be open, and let $x \in U$. Since U is open, there exists some $s > 0$ such that $B_{\mathcal{X}}(x,s) \subseteq U$. Then

$$\begin{aligned}
T[U] &\supseteq T[B_{\mathcal{X}}(x,s)] \\
&= \{T(x+su) : u \in B_{\mathcal{X}}(0,1)\} \\
&= Tx + sT[B_{\mathcal{X}}(0,1)] \\
&\supseteq Tx + sB_{\mathcal{Y}}(0,r) \\
&= B_y(Tx, sr).
\end{aligned}$$

□

Theorem 14.2 (Open mapping theorem). *Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces, and let $T : \mathcal{X} \to \mathcal{Y}$ be surjective. Then T is open.*

Proof. Step 1: Write $\mathcal{X} = \bigcup_n nB_\mathcal{X}(0,1)$. So $\mathcal{Y} = T[\mathcal{X}] = \bigcup_n nT[B_\mathcal{X}(0,1)]$. By the Baire category theorem, there is some $n \in \mathbb{N}$ such that $\overline{nT[B_\mathcal{X}(0,1)]}$ contains some open ball $B_\mathcal{Y}(y,r)$. Then $\overline{T[B_\mathcal{X}(0,1)]} \supseteq B_y(y/n, r_n)$. Pick x_1 such that $\|Tx_1 - y/n\| < r/(4n)$. Then $\overline{T[-x_1 + B_\mathcal{X}(0,1)]} = -Tx_1 + \overline{T[B_\mathcal{X}(0,1)]} \supseteq B_\mathcal{Y}(y/n - Tx_1, r/n) \supseteq B_\mathcal{Y}(0, r/(2n))$. So we get $\overline{T[B_\mathcal{X}(0, 1 + \|x_1\|)]} \supseteq B_\mathcal{Y}(0, r/(2n))$. This gives us that $\overline{T[B_\mathcal{X}(0,1)]} \supseteq B_y(0,s)$ for some $s > 0$. By dilating by a constant (which is a homeomorphism from a Banach space to itself), we get $\overline{T[B_\mathcal{X}(0,r)]} \supseteq B_y(0,s)$ for all $r > 0$.

Step 2: Pick $y \in B_y(0,s)$, and pick $x_1 \in \mathcal{B}_\mathcal{X}(0,1)$ such that $\|y - Tx_1\|_\mathcal{Y} < s/2$. Call $y_1 = y - Tx_1$. Now pick $x_2 \in B_\mathcal{X}(0,1/2)$ such that $\|y_1 - T_{x_2}\|_\mathcal{Y} < s/4$, calling $y_2 = y_1 - T_{x_2}$. Continuing like this, we get a sequence $x_n \in B_\mathcal{X}(0, 1/2^{n-1})$ such that if $y_n = y_{n-1} - Tx_n$, then $\|y_n\| < s/2^n$. In the end, $x := \sum_n x_n \in \mathcal{X}$ as $\|x\| \leq \sum_n \|x_n\| < 2$, and $Tx = \sum_n Tx_n = y$. So $T[B_\mathcal{X}(0,2)] \supseteq B_\mathcal{Y}(0, s/2)$. □

Corollary 14.1. *If $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ is a bijection between Banach spaces, then $T^{-1} \in \mathcal{L}(\mathcal{Y}, \mathcal{X})$.*

Proof. T is open iff T^{-1} is continuous. □

Corollary 14.2. *If $\|\cdot\|_1 \leq \|\cdot\|_2$ are 2 norms on $\mathcal{X}$ that are both complete, then $\|\cdot\|_1 \geq C\|\cdot\|_2$ for some C.*

Proof. Apply the previous corollary to $\mathrm{id} : (\mathcal{X}, \|\cdot\|_2) \to (\mathcal{X}, \|\cdot\|_1)$. □

14.3 The closed graph theorem

Definition 14.2. The **graph** of $T : \mathcal{X} \to \mathcal{Y}$ is $\Gamma(T) = \{(x, Tx) : x \in \mathcal{X}\} \subseteq \mathcal{X} \times \mathcal{Y}$.

If $\mathcal{T}$ is linear, then $\Gamma(T)$ is a subspace of $\mathcal{X} \times \mathcal{Y}$.

Theorem 14.3. *If $T : \mathcal{X} \to \mathcal{Y}$ linear between Banach spaces and $\Gamma(T)$ is a closed subspace of $\mathcal{X} \times Y$, then T is continuous.*

Remark 14.1. In general, if T is continuous, its graph is closed. If $\Gamma(T)$ is closed, T is called a **closed operator**.

Proof. Factorize T into $S(x) = (x, Tx)$ and $R(y,z) = z$.

$$\begin{tikzcd} & \Gamma(T) \arrow[dr, "R"] & \\ \mathcal{X} \arrow[ur, "S"] \arrow[rr, "T"] & & \mathcal{Y} \end{tikzcd}$$

$\Gamma(T)$ is closed, so it is a Banach space. R is continuous, so it suffices to show that S is continuous. But S is a bijection, and $S^{-1} : (y,z) \to y$ is continuous, so the open mapping theorem implies that S is continuous. □

Why do we care? Continuous means that if $x_n \to x$, then $Tx_n \to Tx$. To show that something has a closed graph, we only need to show that if $(x_n, Tx_n) \to (x, y)$, then $y = Tx$. So we don't need to show that such an x exists; we only need to show that if it does, then x_n converges to the right thing.

14.4 The uniform boundedness principle

Theorem 14.4 (uniform boundedness principle)**.** *Let $\mathcal{X}, \mathcal{Y}$ be normed spaces, $\mathcal{X}$ be Banach, and let $\mathcal{A} \subseteq L(\mathcal{X}, \mathcal{Y})$. If $\sup_{T \in \mathcal{A}} \|Tx\| < \infty$ for all $x]in\mathcal{X}$, then $\sup_{T \in \mathcal{A}} \|T\| < \infty$.*

Proof. Let $E_n = \{x \in \mathcal{X} : \sup_{T \in \mathcal{A}} \|Tx\| \leq n\} = \bigcap_{T \in \mathcal{A}} \{x : \|Tx\| \leq n\}$. Then E_n is closed, and $\mathcal{X} = \bigcup_n E_n$, so by Biare category, $E_{n/r} \supseteq B_{\mathcal{X}}(x, 1)$ for some n, x, r. Then $B_{\mathcal{X}}(x, 1) - B_{\mathcal{X}}(x, 1) \subseteq E_{2n/r}$. But the left hand side contains $B_{\mathcal{X}}(0, 2)$. So $\|x\| < 2 \implies \|Tx\| \leq 2n/r$ for all $T \in \mathcal{A}$. This is independent of x, so $\|T\| \leq n/r$ for all $T \in \mathcal{A}$ $\square$

15 Locally Convex Topological Vector Spaces

15.1 A note on the uniform boundedness principle

Here is another perspective on the uniform boundedness principle.

Theorem 15.1 (uniform boundedness principle, weaker version)**.** *Suppose that (X, ρ) is a complete metric space, $(\mathcal{Y}, \|\cdot\|)$ is a normed space, and $\mathcal{F} \subseteq C(X, \mathcal{Y})$ is such that for all $x \in X$, $\sup_{f \in \mathcal{F}} \|f(x)\| < \infty$. Then there exists a nonempty, open $U \subseteq X$ such that $\sup\{\|f(x)\| : x \in U, f \in \mathcal{F}\} < \infty$.*

Proof. Let

$$\begin{aligned} E_n &= \{x \in X : \|f(x)\| \leq n \,\forall f \in \mathcal{F}\} \\ &= \bigcap_{f \in \mathcal{F}} \{\|f(\cdot)\| \leq n\}. \end{aligned}$$

Then each E_n is closed, and $X = \bigcup_n E_n$, so by the Baire category theorem. There exists an n such that $E_n^o \neq \varnothing$. $\square$

If X is Banach and $\mathcal{F} \subseteq L(\mathcal{X}, \mathcal{Y})$, then we actually get $\sup_{T \in \mathcal{F}} \|T\|_{\text{op}} < \infty$.

15.2 Topological vector spaces and convexity

Proposition 15.1. *Let $(\mathcal{X}, \|\cdot\|)$ be a normed space. Then*

1. *The addition map $\mathcal{X} \times \mathcal{X} \to \mathcal{X}$ sending $(x, y) \mapsto x + y$ is continuous.*
2. *The scalar multiplication map $K \times \mathcal{X} \to \mathcal{X}$ given by $(\lambda, x) \mapsto \lambda x$ is continuous.*

Proof. Use the fact that these maps are continuous over the scalar field. $\square$

Definition 15.1. A **topological vector space** is a pair $(\mathcal{X}, \mathcal{T})$ such that $\mathcal{X}$ is a vector space over $K = \mathbb{R}$ or $\mathbb{C}$, $\mathcal{T}$ is a topology on $\mathcal{X}$, and addition and scalar multiplication are continuous.

Definition 15.2. Let $\mathcal{X}$ be a vector space over K. A subset $A \subseteq \mathcal{X}$ is **convex** if $x, y \in A \implies tx + (1-t)y \in A$ for all $t \in [0, 1]$.

Definition 15.3. A topological vector space is **locally convex** if the origin in $\mathcal{X}$ has a neighborhood base consisting of convex open sets.

15.3 Topologies induced by seminorms

Theorem 15.2. *Let $(p_\alpha)_\alpha$ be a family of seminorms on $\mathcal{X}$. If $x \in \mathcal{X}$, $\alpha \in A$, and $\varnothing > 0$, define $U_{x,\alpha,\varepsilon} = \{y : p_\alpha(x-y) < \varepsilon\}$. Let $\mathcal{T}$ be the topology generated by the $U_{x,\alpha,\varepsilon}$.*

1. *For $x \in \mathcal{X}$, the set $\{\bigcap_{i=1}^n U_{x,\alpha_i,\varepsilon} : \alpha_i \in A, \varepsilon > 0\}$ is a neighborhood base at x.*
2. *If (x_n) is a sequence in $\mathcal{X}$, then $x_n \to x$ in $\mathcal{T}$ iff $p_\alpha(x_n - x) \to 0$ for all α.*
3. *$(\mathcal{X}, \mathcal{T})$ is a locally convex topological vector space.*

Proof. Here are the idea.

1. Suppose $x \in \bigcap_{i=1}^n U_{x_i,\alpha_i,\delta}$. Then $p_{\alpha_i}(x-x_i) < \delta_i$ for each i. Pick $\varepsilon_i < \delta_i - p_{\alpha_i}(x-x_i)$. Now $x \in U_{x_i,\alpha_i,\delta} \subseteq U_{x_i,\alpha_i,\delta_i}$. Let $\varepsilon = \min(\varepsilon_1, \ldots, \varepsilon_n)$.
2. Try it yourself!
3. We must show that addition and multiplication are continuous. Pick $\bigcap_{i=1}^n U_{x+y,\alpha_i,\varepsilon} \ni x+y$. Let $x' \in \bigcap_i U_{x,\alpha,\varepsilon/2}$ and same for y. Multiplication is the same.

   To get local convexity, if $y, z \in U_{x,\alpha,\varepsilon}$ and $t \in [0,1]$, then $p_\alpha(x - ty - (1-tz) \leq p_\alpha(tz - ty) = p_\alpha((1-t)x - (1-tz)) = tp_\alpha(x-y) + (1-t)p_\alpha(x-z) < \varepsilon$. Any intersection of convex sets is convex. □

Example 15.1. Let $\mathbb{R}^{\mathbb{N}}$ have the product topology. Let $p_i(x) = |x_i|$ for each i. These generate the product topology. Alternatively, we could define $\tilde{p}_m(x) = \max_{i \leq m} |x_i|$. Actually, we could also take $r_u(x) = |x_1| + \cdots + |x_i|$. This is a locally convex vector space. However, there is no norm that gives the product topology on $\mathbb{R}^{\mathbb{N}}$.

Example 15.2. There is a locally convex topology on $C(\mathbb{R}^n)$ that captures the notion of locally uniform convergence. Define the seminorms $p_m(f) = \|f|_{\overline{B_m(0)}}\|_u$ for each $m \in \mathbb{N}^+$. Now $f_n \to f$ in $\mathcal{T}$ iff $f_n \to f$ locally uniformly.

Example 15.3. Look at $L^1_{\text{loc}}(\mathbb{R}^n)$, Define the seminorms $P_m(f) = \int_{[-m,m]^n} |f| \, dx$. Then $f_n \to f$ in this topology iff $f_n \mathbb{1}_B \to f \mathbb{1}_B$ in L^1 for all bounded, measurable $B \; subseteq \mathbb{R}^n$.

Here is a non-example.

Example 15.4. Let $(X, \mathcal{M}, \mu)$ be a measure space, and define $L^0(\mu)$ to be the set of equivalence classes of measurable functions $X \to \mathbb{R}$ that agree μ-a.e. Let $\mathcal{T}$ be the topology generated by all sets of the form $V(f, \varepsilon) := \{g \in L^0(\mu) : \mu(\{|f-g| > \varepsilon\}) < \varepsilon\}$, where $f \in L^0(\mu)$ and $\varepsilon > 0$. Then $f_n \to f$ iff $f_n \to f$ in measure, but $\mathcal{T}$ is not locally convex.

In normed spaces, we saw that continuity was equivalent to boundedness. How does this play out in locally convex spaces?

15.4 Continuity in locally convex spaces

Proposition 15.2. *Let $\mathcal{X}, \mathcal{Y}$ be locally convex spaces generated by $(p_\alpha)_\alpha$ and $(q_\beta)_\beta$, respectively. Let $T : \mathcal{X} \to \mathcal{Y}$ be linear. The following are equivalent:*

1. *T is continuous.*
2. *For all $\beta \in B$, there exist $\{\alpha_1, \ldots, \alpha_n\} \subseteq A$ and $C > 0$ such that $q_\beta(Tx) \leq C \sum_{i=1}^n p_{\alpha_i}(x)$.*

Proof. (1) $\implies$ (2): Pick $\beta \in B$. If T is continuous, then $\{x : q_\beta(Tx) < 1\}$ is open in $\mathcal{X}$ and contains 0. So there exist $\alpha_1, \ldots, \alpha_n \in A$ and $\varepsilon > 0$ such that $\bigcap_{i=1}^n U_{0,\alpha_i,\varepsilon} \subseteq \{q_\beta \circ T < 1\}$. In particular, if $x \in \mathcal{X}$ and $\sum_{i=1}^n p_{\alpha_i}(x) < \varepsilon$, then $x \in U$, so $q_\beta(Tx) < 1$. That is, if $(1/\varepsilon) \sum_{i=1}^n p_{\alpha_i}(x) < 1$, then $q_\beta(Tx) < 1$. By homogeneity of order 1, we get $q_B \circ T \leq (1/\varepsilon) \sum_{i=1}^n p_{\alpha_i}$. □

Example 15.5. Take $\mathbb{R}^{\mathbb{N}}$ with the 3 families of seminorms $p_i(x) = |x_i|$, $q_i(x) = \max_{j \leq i} |x_j|$, and $r_i(x) = |x_1 + \cdots + |x_i|$. If we had that $\mathbb{R}^{\mathbb{N}}$ had a topology given by a norm, then $\|x\| \leq C \sum_{i=1}^n |x_i|$ for some C and n. But then, if we pick x to be nonzero but 0 in the first n coordinates, it has to have norm 0. This is impossible.

16 Fréchet Spaces, Weak Topologies, and The Weak* Topology

16.1 Fréchet spaces

Proposition 16.1. *Let $(\mathcal{X}, (p_\alpha)_\alpha, \mathcal{T})$ be a locally convex topological vector space generated by the seminorms p_α.*

1. *$\mathcal{T}$ is Hausdorff iff for all $x \in \mathcal{X} \setminus \{0\}$, there exists some α such that $p_\alpha(x) \neq 0$.*
2. *If $\mathcal{T}$ is Hausdorff and A is countable, then $(\mathcal{X}, \mathcal{T})$ is metrizable with a translation invariant metric: $\rho(x+z, y+z) = \rho(x,y)$ for all z.*

Proof. Proving the first statement is easiest done with the left implication and the contrapositive of the right implication.

1. ($\Longleftarrow$): Let $x, y \in \mathcal{X}$ such that $x \neq y$. Then there exists α such that $p_\alpha(y-x) > 0$. Consider $U_{x,\alpha,\varepsilon}, U_{y,\alpha,\varepsilon}$ for $\varepsilon < p_\alpha(y-x)/2$.

   ($\Longrightarrow$): Otherwise, there exists $x \neq 0$ such that $p_\alpha(x) = 0$ for all α. Then $x \in U_{0,\alpha,\varepsilon}$ for all α, ε. So x lies in any neighborhood of 0.

2. Given $(p_n)_{n \in \mathbb{N}}$, define

   $$\rho(x,y) = \max\{2^{-n} \min\{\rho(x-y), 1\} : n \in \mathbb{N}\}.$$

   The min inside satisfies the triangle inequality, and taking maxes preserves the triangle inequality. So this is a pseudometric. Since ρ is a function of $x-y$, it is translation invariant. Lastly, if $x \neq y$, then $p_x(x-y) \neq 0$ for some n, so $\rho(x,y) > 0$. $\square$

Definition 16.1. A **Fréchet space** a locally convex topological vector space with the above metric such that ρ can be chosen to be complete.

Example 16.1. $\mathbb{R}^{\mathbb{N}}$ with the product topology, $C(\mathbb{R}^n)$ with the topology of local uniform convergence, and L^1_{loc} are all Fréchet spaces.

16.2 Weak topologies

Definition 16.2. Let $T_\alpha : \mathcal{X} \to (\mathcal{Y}_\alpha, \|\cdot\|_\alpha)$ be a collection of linear maps with the resulting family of seminorms $p_\alpha(x) = \|T_\alpha x\|_\alpha$. These generate the **weak topology generated by** $(T_\alpha)_\alpha$.

Example 16.2. Let $T_m : C(\mathbb{R}^n) \to C([m,m]^d)$ send $f \mapsto f|_{[-m,m]^d}$. Then the topology of local uniform convergence is the weak topology generated by these maps.

Example 16.3. On $C^\infty([0,1])$, for each k, consider $(d/dx)^k : C^\infty([0,1]) \to C([0,1])$. Now take the weak topology generated by these.

Usually in the setting of normed spaces, we refer to a very specific weak topology.

Definition 16.3. The **weak topology** on $(\mathcal{X}, \|\cdot\|)$ is the topology generated by $\mathcal{X}^*$, the set of continuous linear functionals.

Remark 16.1. In general, $\mathcal{T}_{\text{weak}} \subseteq \mathcal{T}_{\text{norm}}$. These are equal iff $\dim(\mathcal{X}) < \infty$. If $f \in \mathcal{X}^*$, show that $U_{x,f,\varepsilon} = \{y : |f(y-x)| < \varepsilon\}$ is contained in a ball around x.

Remark 16.2. Convergence in the weak topology means the following:

$$x_n \to x \iff f(x_n) \to f(x) \ \forall f \in \mathcal{X}^*.$$

In norm topologies, we have $|f(x_n) - f(x)| \leq \|f\|\|x_n - x\|$. So the weak topology is weaker particularly because it does not give this uniformity of convergence.

16.3 The weak* topology

If $(\mathcal{X}^*, \|\cdot\|)$ is a Banach space, then we have the dual space $(\mathcal{X}^*, \|\cdot\|_*)$, This has its own dual $(\mathcal{X}^{**}, \|\cdot\|_{**})$. We have 2 choices for the weak topology on $\mathcal{X}^*$: we can take the usual weak topology, or we can restrict to the even weaker topology generated by $\mathcal{X}$ embedded into $\mathcal{X}^{**}$.

Definition 16.4. The **weak* topology** on $\mathcal{X}^*$ is generated by the family of maps $\hat{x} : f \mapsto f(x) \in \mathcal{K}$ where $x \in \mathcal{X}$ and $f \in \mathcal{X}^*$.

Theorem 16.1 (Alaoglu)**.** *$B^* = \{f \in \mathcal{X}^* : \|f\|_* \leq 1\}$ is compact for the weak* topology.*

Proof. Say $K = \mathbb{C}$.

$$\begin{aligned} B^* &= \{f : \mathcal{X} \to \mathbb{C} \mid f(x + \lambda y) = f(x) + \lambda f(y), |f(x)| \leq \|x\|\} \\ &= \{f : \mathcal{X} \to \mathbb{C} : f \text{ is linear}, f(x) \in \overline{B_{\mathbb{C}}(0, \|x\|)} \ \forall x\}. \end{aligned}$$

That is, $f \in \prod_{x \in \mathcal{X}} \overline{B_{\mathbb{C}}(0, \|x\|)}$.

$$\begin{aligned} &= \{f \in \prod_{x \in \mathcal{X}} \overline{B_{\mathbb{C}}(0, \|x\|)} : f(x+y) - g(x) = \lambda f(y) = 0 \ \forall x, y, \lambda\} \\ &= \bigcap_{x,y,\lambda} \{f \in \prod_{x \in \mathcal{X}} \overline{B_{\mathbb{C}}(0, \|x\|)} : f(x+y) - g(x) = \lambda f(y) = 0\}. \end{aligned}$$

By Tychonoff's theorem, we need only show that $\mathcal{T}_{\text{weak}^*}|_{B^*} = \mathcal{T}_{\text{prod}}|_{B^*}$. These are weak topologies generated by the same family of maps. $\square$

Proposition 16.2. *Let $(\mathcal{X}, \|\cdot\|)$ be separable. Then $\mathcal{T}_{\text{weak}^*}|_{B^*}$ is metrizable.*

Proof. Let $(x_n)_n$ be a dense sequence in $\mathcal{X}$. Then define ρ on B^* by

$$\rho(f, g) = \max\{2^{-n}/\|x_n\||f(x_n) - g(x_n)| : n \in \mathbb{N}\}.$$

This generates $\mathcal{T}_{\text{weak}^*}|_{B^*}$. For all $x \in \mathcal{X}$, there exists $x_{n_i} \to x$, and therefore $\hat{x}_{n_i}|_{B^*} \to \hat{x}|_{B^*}$ uniformly. $\square$

17 Weak* Metrizability, Operator Topologies, and Complex Measures

Many thanks to Anthony Graves-Mccleary, who provided me with notes when I missed this lecture.

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

Let's be a bit more thorough with a point we went over last time.

Proposition 17.1. *Let $(\mathcal{X}, \|\cdot\|)$ be separable. Then $\mathcal{T}_{\text{weak}^*}|_{B^*}$ is metrizable.*

Proof. Let $(x_n)_n$ be a dense sequence in $\mathcal{X}$. In $\mathcal{T}_{\text{weak}^*}|_{B^*}$, a neighborhood base of $f \in B^*$ is sets of the form

$$\bigcap_{i=1}^{m} \{g \in B^* : |g(x^{(i)}) - f(x^{(i)})| < \varepsilon\}$$

for some $x^{(1)}, \dots, x^{(m)} \in \mathcal{X}$ and $\varepsilon > 0$. Consider $\mathcal{T}'$ generated by such neighborhoods except only using $x^{(i)}$ from $\{x_1, x_2, \dots\}$. Then $\mathcal{T}' \subseteq \mathcal{T}_{\text{weak}^*}|_{B^*}$.

Step 1: $\mathcal{T}$ is metrizable: Let

$$\rho(f, g) = \max_{n \geq 1}(2^{-n} \min(|f(x_n) - g(x_n)|, 1)).$$

This is analogous to the construction of a metric on a weak topology.

Step 2: We know that $\mathcal{T}_{\text{weak}^*}|_{B^*}$ is the weakest topology on B^* that makes $\hat{x} = (f \mapsto f(x))$ continuous for each $x \in \mathcal{X}$. To finish, show that $\mathcal{T}'$ has this property; i.e. $\hat{x}$ is $\mathcal{T}'$-continuous. Suppose $x \in \mathcal{X}$. There exists a sequence (x_{n_i}) in the countable dense set such that $x_{n_i} \to x$ in norm. As a result, if $f \in B^*$, then

$$|\hat{x}(f) - \hat{x_{n_i}}(f)| = |f(x) - f(x_{n_i})| \leq \|f\| \cdot \|x - x_{n_i}\| \leq \|x - x_{n_i}\|,$$

which goes to 0 independently of f. So $\hat{x}_{n_i} \to \hat{x}$ uniformly on B^*. Thus, $\hat{x}$ is a uniform limit of $\mathcal{T}'$-continuous functions, so $\hat{x}$ is $\mathcal{T}'$-continuous. $\square$

Remark 17.1. The weak* topology is almost never metrizable for all of $\mathcal{X}^*$.

17.2 The strong and weak operator topologies

Let $\mathcal{X}, \mathcal{Y}$ be Banach spaces.

Definition 17.1. The **strong operator topology** on $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ is the topology generated by the linear operators $T \mapsto Tx$ for $x \in \mathcal{X}$; i.e. this is the weak generated by the seminorms $T \mapsto \|Tx\|$.

$T_n \to T$ in the strong operator topology if and only if $T_n x \to Tx$ in norm for all $x \in \mathcal{X}$.

Definition 17.2. The **weak operator topology** on $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ is the topology generated by the linear operators $T \mapsto \varphi(Tx)$ for $x \in \mathcal{X}$ and $\varphi \in \mathcal{Y}^*$; i.e. this is the weak topology generated by the seminorms $T \mapsto \|\varphi(Tx)\|$.

$T_n \to T$ in the weak operator topology if and only if $T_n x \to Tx$ weakly in $\mathcal{Y}$ for all $x \in \mathcal{X}$.

17.3 Signed measures, complex measures and the Lebesgue-Radon-Nikodym theorem

Recall the concept of signed measures. A signed measure ν cannot hit both $+\infty, -\infty$, and signed measures are related to two decompositions:

1. Hahn decomposition: $X = P \cup N$, where $\nu(A) \geq 0$ for all measurable $A \subseteq P$, and $\nu(B) \leq 0$ for all measurable $B \subseteq N$.
2. Jordan decomposition: $\nu = \nu^+ - \nu^-$, where ν^+ and ν^- are positive measures.

We write $|\nu| = \nu^+ + \nu^-$, and integration with respect to ν is $\int f\, d\nu = \int f d\, mu^+ - \int f\, d\nu^-$ for $f \in L^1(|\nu|)$.

Theorem 17.1 (Lebesgue-Radon-Nikodym)**.** *Let μ, ν be σ-finite positive and signed measures, respectively. Then there exists a unique decomposition $\nu = \lambda + \rho$ such that $\lambda \perp \mu$ and $\rho \ll \mu$. The Radon-Nikodym derivative, the function f such that $d\rho = f\, d\mu$, is unique μ-a.e.*

Definition 17.3. A **complex measure** on $(X, \mathcal{M})$ is a function $\nu : \mathcal{M} \to \mathbb{C}$ such that

1. $\nu(\varnothing) = 0$,
2. For (E_n) disjoint in $\mathcal{M}$, $\nu(\bigcup_{n=1}^\infty E_n) = \sum_{n=1}^\infty \nu(E_n)$, where the sum converges absolutely.

Here, we can write $\nu = \mathrm{Re}(\nu) + i\,\mathrm{Im}(\nu) = \nu_r + i\nu_i$, where ν_r, ν_i must be finite signed measures.

Definition 17.4. **Integration** with respect to a complex measure ν is given by

$$\int f\, d\nu = \int f\, d\nu_r + i \int f\, d\nu_i$$

for $f \in L^1(|\nu_r| + |\nu_i|)$.

Theorem 17.2 (Lebesgue-Radon-Nikodym for complex measures)**.** *Let μ, ν be σ-finite positive and signed measures, respectively. Then there exists a unique decomposition $\nu = \lambda + \rho$ such that $\lambda \perp \mu$ (i.e. $\lambda_r^\pm, \lambda_i^\pm$ all $\perp \mu$), $\rho \ll \mu$ (i.e. $\rho_r^\pm, \rho_i^\pm$ all $\ll \mu$), and the Radon-Nikodym derivative, $d\rho = f\, d\mu$ for some $f \in L^1_{\mathbb{C}}(\mu)$.*

17.4 Total variation of complex measures

If ν is a complex measure, then $\nu \ll |\nu_r| + |\nu_i|$. Now suppose $\nu \ll \mu$, where μ is σ-finite and positive. By Radon-Nikodym, $d\nu = f\,d\mu$ for some $f \in L^1_{\mathbb{C}}(\mu)$. We want to define $d|\nu| = |f|\,d\mu$.

Lemma 17.1. *If $f_1\,d\mu_1 = f_2\,d\mu_2$, then $|f_1|\,d\mu_1 = |f_2|\,d\mu_2$ (so $d|\nu|$ is well defined).*

Proof. For $i = 1, 2$, $\mu_i \ll \mu = \mu_1 + \mu_2$, so $d\mu_i = g_i d\mu$, where $g_i \geq 0$. Then $f_1 g_1\,d\mu = f_2 g_2\,d\mu$. So $f_1 g_2 = f_2 g_2$ μ-a.e., which gives $|f_1| g_1 = |f_1 g_1| = |f_2 g_2| = |f_2| g_2$ μ-a.e. So

$$|f_1|\,d\mu_1 = |f_1| g_1\,d\mu = |f_2| g_2\,d\mu = |f_2|\,d\mu_2. \qquad \square$$

Proposition 17.2. *Let ν be a complex measure. The total variation, $|\nu|$ has the following properties:*

1. $|\nu(E)| \leq |\nu|(E)$ *for all* $E \in \mathcal{M}$.
2. $\nu \ll |\nu|$, *and* $|\frac{d\nu}{d|\nu|}| = 1$ $|\nu|$*-a.e.*
3. $L^1(\nu) = L^1(|\nu|)$, *and if* $f \in L^1_{\mathbb{C}}(\nu)$, *then* $|\int f\,d\nu| \leq \int |f|\,d|\nu|$.

Proposition 17.3. *If ν_1, ν_2 are complex measures, then $|\nu_1 + \nu_2| \leq |\nu_1| + |\nu_2|$.*

18 The Riesz Representation Theorem

18.1 Triangle inequality for complex measures

Lemma 18.1. *Let $(X, \mathcal{M})$ be measurable with complex measures ν_1, ν_2. Then $|\nu_1 + \nu_2| \leq |\nu_1| + |\nu_2|$.*

Proof. Given ν, find a positive measure $\mu \gg \nu$. Then we get $d\nu = f\,d\mu$. Now $d|\nu| = |f|\,d\mu$. Similarly, let $d\nu_i = f_i\,d\mu$ for $\mu = |\nu_1| + |\nu_2|$. Then $d(\nu_1 + \nu_2) = (f_1 + f_2)\,d\mu$, so $d|\nu_1 + \nu_2| = |f_1 + f_2|\,d\mu \leq |f_1|\,d\mu + |f_2|\,d\mu$. $\square$

18.2 Positive linear functionals and the Riesz-Markov theorem

Let (X, ρ) be a compact metric space. The goal is to describe the dual of $C(X, \mathbb{R})$ or $C(X, \mathbb{C}) = C(X)$ with the uniform norm. Recall the Riesz-Markov theorem:

Definition 18.1. A linear functional $\ell : C(X, \mathbb{R}) \to \mathbb{R}$ is **positive** if $\ell(f) \geq 0$ whenever $f \geq 0$.

So if $f \geq g$ then $\ell(f) = \ell(g) + \ell(f - g) \geq \ell(g)$.

Theorem 18.1 (Riesz-Markov)**.** *For any positive linear functional $\ell : C(X, \mathbb{R}) \to \mathbb{R}$, there exists a unique finite positive Borel measure μ on X such that $\ell(f) = \int f\,d\mu$.*

Remark 18.1. If ℓ is a positive linear functional on $C(X, \mathbb{R})$ and $f \in C(X, \mathbb{R})$, then $-\|f\|_u \leq f \leq \|f\|_u$. Then $-\|f\|_u c_x \leq f \leq \|f\|_u c_x$, where c_x is the constant x function. So $-\|f\|_u \ell(c_x) \leq \ell(f) \leq \|f\|_u \ell(c_x)$, which gives $|\ell(f)| \leq \|f\|_u \ell(c_x)$ with equality if $f = c_x$. So $\|\ell\|_{C(X,\mathbb{R})^*} = \ell(c_x) = \mu(X)$.

18.3 The Riesz representation theorem

Let $M(M, K)$ be the space of all finite signed or complex measures on $(X, \mathcal{B}_X)$, where $K = \mathbb{R}$ or $\mathbb{C}$. This is a vector space over K.

Lemma 18.2. *$M(X, K)$ is a normed space over K with norm $\|\mu\| = |\mu|(X)$.*

Proof. If $\lambda \in K$ and $\mu \in M(X, K)$, then $d\mu = f\,d|\mu|$. So $d(\lambda\mu) = (\lambda f)\,d|\mu|$, and we get $d|\lambda\mu| = |\lambda f| d|\mu| = |\lambda|\,d|\mu|$. So $\|\lambda\mu\| = |\lambda|\|\mu\|$.

If $\nu_1, \nu_2 \in M(X, \mathbb{C})$, then by the lemma, we have $\|\nu_1 + \nu_2\| = |\nu_1 + \nu_2|(X) \leq |\nu_1|(X) + |\nu_2|(X) = \|\nu_1\| + \|\nu_2\|$.

If $\|\nu\| = 0$, then $|\nu|(X) = 0$, so $|\nu| = 0$ by monotonicity. Then $\nu = 0$ because $\nu \ll |\nu|$. $\square$

Theorem 18.2 (Riesz representation)**.** *For $\mu \in M(X, \mathbb{R})$, define $\ell_\mu \in C(X, \mathbb{R})^*$ by $\ell_\mu(f) = \int f\,d\mu$. Then $\mu \mapsto \ell_\mu$ is an isometric isomorphism $M(X, \mathbb{R}) \to C(X, \mathbb{R})^*$. The same holds if we replace $\mathbb{R}$ by $\mathbb{C}$.*

Here is a lemma we will need.

Lemma 18.3. *If $\ell \in C(X, \mathbb{R})^*$, then $\ell = \varphi - \psi$ for some positive linear functionals φ, ψ on $C(X, \mathbb{R})$.*

Proof. For $f \in C(X, \mathbb{R})$ with $f \geq 0$, define $\varphi(f) = \sup\{\ell(g) : 0 \leq g \leq f\}$. For general f define $\varphi(f) := \varphi(f^+) - \varphi(f^-)$. We need to show that φ is a positive linear functional such that $\varphi \geq \ell$. Then we can just define $\psi := \varphi - \ell$. By definition, we have $\varphi(f) \geq \ell(f)$ if $f \geq 0$, which gives us the inequality.

To show that φ is a linear functional, we take a few steps:

1. Suppose $f, h \geq 0$. Then for all $0 \leq g_1 \leq f$ and $0 \leq g_2 \leq h$, we have $0 \leq g_1 + g_2 \leq f + h$. So $\varphi(f+h) \geq \ell(g_1) + \ell(g_2)$ for all such g_1, g_2. Taking the sup over such g_1, g_2, we get $\varphi(f+h) \geq \varphi(f) + \varphi(h)$.

   Conversely, if $0 \leq g \leq f + h$, define $g_1 := \min\{g, f\}$. If $\min = f$ at some x, then $g(x) - g_1(x) = g(x) - f(x) \leq f(x) + h(x) - f(x) \leq h(x)$. So $g_2 := g - g_1 \leq h$. Take the sup over g to get $\varphi(f+h) \leq \varphi(f) + \varphi(h)$. So we get equality.

2. If $f = f^+ - f^- = g - h$ for $g, h \geq 0$, then $f^+ + h = g + f^-$, so step 1 gives $\varphi(f^+) + \varphi(h) = \varphi(g) + \varphi(f^-)$, so $\varphi(f) = \varphi(f^+) - \varphi(f^-) = \varphi(g) - \varphi(h)$.

3. For all f, h, we have $f + h = (f^+ + h^+) - (f^- + h^-)$. So $\varphi(f+h) = \varphi(f^+ + h^+) - \varphi(f^- + h^-) = (\varphi(f^+) - \varphi(f^-)) + (\varphi(h^+) - \varphi(h^-)) = \varphi(f) + \varphi(h)$. So φ is additive.

Similarly, $\varphi(\lambda f) = \lambda \varphi(f)$ for all $\lambda \in \mathbb{R}$. □

Proof. Definitely, $\ell_\mu \in C(X, K)^*$ for all $\mu \in M(X, K)$. Next, suppose $\ell \in C(X, \mathbb{R})^*$. Then $\ell = \varphi - \psi$, where $\varphi, \psi \geq 0$. By Riesz-Markov, we get $\ell = \ell_{\mu_1} - \ell_{\mu_2}$ for some $\mu_1, \mu_2 \geq 0$. So $\ell = \ell_{\mu_1 - \mu_2}$. If $\ell \in C(X, \mathbb{C})$, we can represent this as

$$\ell(f) = \ell_{\mu_1}(\mathrm{Re}(f)) - i\ell_{\mu_2}(i \,\mathrm{Im}(f)) = \ell_{\mu_1 - i\mu_2}(f).$$

It remains to show that $\|\ell_\mu\|$. Let's just prove this for the complex case; the real case is the same argument. We have $\ell_\mu(f) = \int f \;, d\mu$, so we get

$$|\ell_\mu(f)| = |\int f \, d\mu \leq \int |f| d|\mu| \leq \|f\|_u \int 1 \, d|\mu| = \|f\|_u \cdot \|u\|.$$

Let $d\mu = k \, d|\mu|$, where $k = d\mu/d|\mu|$ is measurable from $X \to S^1$. Now use the fact that for any ε_0, there exists $f \in C(X, \mathbb{C})$ such that $\|f - k\|_{L^1(|\mu|)} < \varepsilon$. We may assume $|f| \leq 1$. Now

$$\ell_\mu(\overline{f}) = \int \overline{f} \, d\mu = \int \overline{f} k \, d|\mu| \approx_\varepsilon \int \overline{k} k \, d|\mu| = \int 1 \, d|\mu| = \|\mu\|.$$

So $\|\ell_\mu\| \geq \|\mu\| - \varepsilon$ for all $\varepsilon > 0$. □

19 Applications of Riesz-Representation to Probability Measures

19.1 Probability measures on compact metric spaces

Let (X, ρ) be a compact metric space. Since $M(X, \mathbb{R}) = C(X, \mathbb{R})^*$, we may write $\mu(f) := \int f \, d\mu$.

Definition 19.1. A **probability measure** on a measurable space $(X, \mathcal{M})$ us a positive measure μ such that $\mu(X) = 1$.

On (X, ρ), let $P(x)$ be the collection of probability measures. Then $P(X) \subseteq M(X, \mathbb{R})$.

Lemma 19.1. *$P(x) = \{\mu \in C(X, \mathbb{R})^* : \|\mu\| \leq 1, \mu(\mathbb{1}_X) = 1\}$, where $\mathbb{1}_X$ is the constant 1 function.*

Proof. $\mu \in P(X)$ iff $\mu \geq 0$ and $\mu(X) = 1$. All the work is in showing $\supseteq$. We just need to show that if $\|\mu\| \leq 1$ and $\mu(X) = 1$, then $\mu \geq 1$. Let $\mu = \mu^+ - \mu^-$ be the Jordan decomposition of μ. Then $\mu(X) = \mu^+(X) - \mu^{-1}(X) = 1$, and $\|\mu\| = |\mu|(X) = \mu^+(X) + \mu^-(X) \leq 1$. So $\mu(X) = \mu^+(X) = 1$, and $\mu^- = 0$. □

Corollary 19.1. *$P(X)$ us compact and metrizable in the weak* topology.*

Proof. $P(X) = B^* \cap \{\mu \in C(X, \mathbb{R})^* : \mu(\mathbb{1}_X) = 1\}$. The latter set is closed, so $P(X)$ is a weak* closed subset of B^*. So Alaoglu's theorem gives us that $P(X)$ is compact. □

Remark 19.1. $(B^*, \mathcal{T}_{\text{weak}^*})$ is metrizable, so $C(X, \mathbb{R})$ is separable (this was a previous application of Stone-Weierstrass).

Remark 19.2. Here are explicit examples of suitable metrics. Le (f_n) be dense in the unit ball of $C(X, \mathbb{R})$. Then

$$\tilde{\rho}(\mu, \nu) = \max_n \left\{ 2^{-n} \left| \int f_n \, d\mu - \int f_n \, d\nu \right| \right\}$$

is a metric. The same works if we replace the max by a sum.

Remark 19.3. Embed $X \to P(X)$ by sending $x \mapsto \delta_x$. This is a homeomorphic embedding with this topology. The key point is that $x_n \to x$ iff $f(x_n) \to f(x)$ for all $f \in C(x)$; that is, $\int f \, d\delta_{x_n} \to \int f \, d\delta_x$.

Theorem 19.1 (Krylov-Bogoliubov)**.** *Let $X \neq \varnothing$ be a compact metric space, and let $T : X \to X$ be continuous. Then there exists $\mu \in P(X)$ such that $\mu(T^{-1}[A]) = \mu(A)$ for all $A \in \mathcal{B}_X$.*

Proof. Pick $x \in X$. For $n \in \mathbb{N}$, define $\mu_n := n^{-1} \sum_{i=0}^{n-1} \delta_{T^i(x)}$. Now let μ be $\lim_{k\to\infty} \mu_{n_k}$ for some subsequence that converges. Then for all $f \in C(X)$, we have

$$\int f \, d\mu = \lim_k \int f \, d\mu_{n_k} = \lim_k \frac{1}{n_k} \sum_{i=0}^{n_k-1} f(T^i(x)).$$

Similarly,

$$\int f \circ T \, d\mu = \lim_k \frac{1}{n_k} \sum_{i=0}^{n_k-1} f(T^{i+1}(x)) = \lim_k \frac{1}{n_k} \sum_{i=1}^{n_k} f(T^i(x)).$$

So

$$\left| \int f \, d\mu - \int f \circ T \, d\mu \right| = \lim_k \frac{1}{n_k} |f(x) - f(T^{n_k}(x))| \leq \frac{2\|f\|_u}{n_k} \to 0.$$

So we get $\int f \, d\mu = \int f \circ T \, d\mu$ for all $f \in C(X)$. This implies that μ is T-invariant (exercise in regularity). $\square$

Remark 19.4. We could write the last step as $\int f \, d\mu = \int f \, d(T_* \mu)$ for all $f \in C(X)$, where $T_* \mu$ is the push-forward measure of μ by T. This gives, $\mu = T_* \mu$.

19.2 Probability measures on non-compact metric spaces

What if our metric space is not compact? One nice way to do things is to work in locally compact spaces. Another important case is to look at complete and separable metric spaces. In either case, it is no longer true that $M(X, \mathbb{R}) = C(X, \mathbb{R})^*$.

Definition 19.2. Let (X, ρ) be a locally compact metric space with $(\mu_n) \subseteq p(X)$ and $\mu \in P(X)$. Then $\mu_n \to \mu$ **vaguely** if $\int f \, d\mu_n \to \int f \, d\mu$ for all $f \in C_0(X, R)$ (functions which tend to 0 at ∞). The **vague topology** is the corresponding topology.

Remark 19.5. The vague topology has a nice Banach space interpretation, but $P(X)$ is usually no longer compact. We can see this by looking at the embedding of $X \to P(X)$.

Remark 19.6. See Folland Proposition 7.19 for an interpretation of the vague topology of $P(\mathbb{R})$ in terms of $F(x) = \mu((-\infty, x])$.

Now suppose X is complete and separable.

Lemma 19.2. *If $\mu \in P(X)$, then for all $\varepsilon > 0$, ther eis a compact $K \subseteq X$ such that $\mu(K^c) < \varepsilon$.*

This motivates the following definition.

Definition 19.3. $\mathcal{A} \subseteq P(X)$ is **tight** if for all $\varepsilon > 0$, there exists a compact $K \subseteq X$ such that $\mu(K^c) < \infty$ for all $\mu \in \mathcal{A}$.

Theorem 19.2 (Prohorov)**.** *Let* (μ_n) *be a sequence in* $P(X)$*, and assume that* $\{\mu_n : n \in \mathbb{N}\}$ *is tight. Then there is a subsequence* $(\mu_{n_k})_k$ *and a measure* $\mu \in P(X)$ *such that*

$$\int f \, d\mu_{n_k} \xrightarrow{k\to\infty} \int f \, d\mu$$

for all $f \in BC(X)$*.*

Remark 19.7. Probabilists usually call the topology related to this type of convergence the **weak topology**.[9] You can instead call this the **BC-topology**.

[9] Please try not to do this.

20 Introduction to L^p Spaces

20.1 L^p spaces and norms

Fix a measure space $(X\mathcal{M}, \mu)$. We will deal with complex functions, but the real case is the same.

Definition 20.1. Let $0 < p < \infty$, and let $f : X \to \mathbb{C}$ be measurable. The L^p **norm**[10] is

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

If f doesn't have a lot of spiky parts in its graph, then the L^p norm of f is about the value of f. When the graph has huge peaks, as p gets bigger, the spikes are amplified. Likewise, as p gets bigger, tails of functions are suppressed.

Definition 20.2. The L^p **space** $L^p(X, \mathcal{M}, \mu) = L^p(\mu) = L^p$ is the space of measurable functions $f : X \to \mathbb{C}$ such that $\|f\|_p < \infty$.

Example 20.1. Let X be a countable set with the measure μ, counting measure on $(X, \mathscr{P}(X))$. Then $\ell^p(X) := L^p(\mu)$. As an example,

$$\ell^p(\mathbb{N}) = \ell^p = \left\{ (x_n)_n \in \mathbb{C}^{\mathbb{N}} : \sum_n |x_n|^p < \infty \right\}.$$

Lemma 20.1. *For all $p \in (0, \infty)$, $L^p(\mu)$ is a vector space over $\mathbb{C}$.*

Proof. If $\|f\|_p, \|g\|_p < \infty$,

$$|f + g|^p \le (2 \max(|f|, |g|))^p = 2^p \max(|f|^p, |g|^p) \le 2^p(|f|^p + |g|^p).$$

So

$$\int |f + g|^p \, d\mu \le 2^p \int |f|^p + 2^p \int |g|^p < \infty.$$

□

20.2 L^p norm inequalities

Now assume $p \ge 1$. We want to show that L^p is a normed space. These inequalities will help us, but they are very important to know on their own.

Lemma 20.2. *If $a, b \ge 0$ and $0 < \lambda < 1$, then*

$$a^\lambda b^{1-\lambda} \le \lambda a + (1 - \lambda) b.$$

[10] This is only really a norm when $p \ge 1$.

Proof. Assume $a, b, > 0$ and take logs:

$$\lambda \log(a) + (1_\lambda) \log(b) \leq \log(\lambda a + (1-\lambda)b)$$

by the convexity of log. $\square$

Lemma 20.3 (Hölder's inequality). *Let $1 < p < \infty$, and define $q \in (1, \infty)$ by $p^{-1} + q^{-1} = 1$. If $f, g : X \to \mathbb{C}$ are measurable, then*

$$\|fg\| \leq \|f\|_p \|g\|_q.$$

In particular, if $f \in L^p$ and $g \in L^q$, then $f, g \in L^1$. Equality holds if and only if $\alpha|f|^p = \beta|g|^q$ for some $\alpha, \beta \in \mathbb{C}$ not both zero.

Remark 20.1. In the statement of this lemma, q is called the **conjugate exponent** of p.

Proof. We may assume $9 < \|f\|_p, \|g\|_q < \infty$. The inequality holds for γf and λg for constants γ, λ iff it holds for f, g, so we may replace f, g by $f/\|f\|_p$ and $g/\|g\|_q$.[11] Let $\lambda = 1/p$, $1 - \lambda = 1/q$, and apply the previous inequality:

$$|f(x)g(x)| = (f(x)^p)^\lambda (|g(x)|)^{1-\lambda} \leq \lambda |f(x)|^p + (1-\lambda)|g(x)|^q.$$

Now integrate with respect to μ on both sides.

The equality case, after we do the reduction, is the case where $f^p = g^q$. $\square$

Lemma 20.4 (Minkowki's inequality). *If $1 \leq p < \infty$, then $\|\cdot\|_p$ satisfies the triangle inequality.*

Proof. Assume $f > 1$, and let $r, g \in L^-$. Then

$$|f+g|^p \leq (|f| + |g|)|f+g|^{p-1} = \int |f||f+g|^{p-1}\, d\mu + \int |g||f+g|^{p-1}\, d\mu$$

Apply Hölder's inequality again,

$$\leq \|f_p\| \| |f+g|^{p-1} \|_q + \|g\|_p \| |f+g|^{p-1} \|_q.$$

We can now check, using $q = p/(p-1)$, that

$$\| |f+g|^p \|_q = \left(\int |f+g|^{p(q-1)}\, d\mu \right)^{(p-1)/p} = \left(\int |f+g|^p\, d\mu \right)^{(p-1)/p} = \|f+g\|_p^{p-1}. \quad \square$$

Corollary 20.1. *Let $1 \leq p < \infty$.$(L^p, \|\cdot\|_p)$ is a normed space.*

Proof. We have shown that L^p is a vector space, and $\|\cdot\|_p$ satisfies the triangle inequality. The L^p norm is homogeneous of order 1, and if $\|f_p\| = 0$, then $\int |f|^p = 0$, which makes $f = 0$ μ-a.e. $\square$

[11] Terence Tao says that in situations like this, we have just "spent a symmetry." In this case, it is a symmetry under scalar multiplication.

20.3 Convergence in L^p spaces

Theorem 20.1. *Let $1 \leq p < \infty$. Then L^p is a Banach space.*

Proof. Assume $\sum_n f_n$ is absolutely convergent in L^p; i.e. $\sum_n \|f_n\|_p < \infty$. Let $G_n = \sum_{i=1}^n |f_i| \in L^p$. It satisfies $\|G_n\|_p \leq \sum_{i=1}^n \|f_i\|_p$ and $G_n(x) \uparrow G(x)$, where G is measurable and $[0,\infty]$-valued. By the monotone convergence theorem, $\|G_n\|_p \uparrow \|G\|_p$. Since $\|G_n\|_p \leq \sum_n \|f_n\|_p$, $\|G\|_p \leq \sum_n \|f_n\|_p$. So G is finite a,e,, and $G \in L^p$. So $\sum_n f_n(x)$ is absolutely convergent whenever $G(x) < \infty$ (i.e. a.e.). Let's call this pointwise limit f. $|f|^p \leq |g|^p$ a.e. so $|f^p| \in L^1$; that is, $|f| \in L^p$. Finally,

$$\left| f - \sum_{i=1}^n f_i \right|^p \leq 2^p |G|^p \in L^1.$$

By the dominated convergence theorem,

$$\int |f - \sum_{i=1}^n f_i|^p \, d\mu \xrightarrow{n \to \infty} 0,$$

so

$$\left(\int |f - \sum_{i=1}^n f_i|^p \, d\mu \right)^{1/p} \xrightarrow{n \to \infty} 0.$$

□

Proposition 20.1. *For $1 \leq p < \infty$, the set of integrable simple functions is dense in L^p.*

Proof. Let $f \in L^p$. There exist complex-valeued simple functions $(\psi_n)_n$ such that $\psi_n \to f$ a.e. and $|\psi_1| \leq |\psi_2| \leq \cdots \leq |f|$. Then $|f - \psi_n|^p \leq 2|f|^p \in L^1$, so $\|f - \psi_n\| \to 0$ by the dominated convergence theorem. □

Corollary 20.2. *Let m be Lebesgue measure on $\mathbb{R}^d$. Then the collection of functions $f \in C(\mathbb{R}^d, \mathbb{C})$ with bounded support is dense in $L^p(m)$.*

20.4 L^∞ spaces

Definition 20.3. Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $f : X \to \mathbb{C}$ be measurable. The L^∞ **norm** or **essential supremum** is

$$\|f\|_\infty = \operatorname*{ess\,sup}_x |f(x)| = \inf\{a \geq 0 : \mu(\{|f| > a\}) = 0\}.$$

Definition 20.4. $L^\infty(\mu)$ is the set of equivalence classes of functions f with $\|f\|_\infty < \infty$, under the equivalence realtion of a.e. equality.

Theorem 20.2. *L^∞ has the following properties:*

1. *For all* f, g, $\|fg\|_q \leq \|f\|_1 \|g\|_\infty$.
2. $\|\cdot\|_\infty$ *is a norm.*
3. L^∞ *is complete.*
4. $f_n \to f$ *in* L^∞ *iff there exists* $E \in \mathcal{M}$ *with* $\mu(E^c) = 0$ *such that* $f_n|_E \to f|_E$ *uniformly.*
5. *The set of simple functions (not necessarily integrable) is dense in* L^∞.

21 L^∞ Spaces and Duality of L^p Spaces

21.1 Properties of L^∞ spaces

Theorem 21.1. *L^∞ has the following properties:*

1. *For all measurable f, g, $\|fg\|_1 \le \|f\|_1 \|g\|_\infty$.*
2. *$\|\cdot\|_\infty$ is a norm.*
3. *L^∞ is complete.*
4. *$f_n \to f$ in L^∞ iff there exists $E \in \mathcal{M}$ with $\mu(E^c) = 0$ such that $f_n|_E \to f|_E$ uniformly.*
5. *The set of simple functions (not necessarily integrable) is dense in L^∞.*

Remark 21.1. If $\mu \not\equiv 0$, we can write

$$\|f\|_\infty := \inf\{a \ge 0 : \mu(\{|f| > a\}) = 0\} = \sup\{b : \mu(\{|f| > b\})\}.$$

The infimum in the definition is attained, but the supremum may not be. Let $a = \|f\|_\infty$. Let $a_n \downarrow 0$ and $\mu\{|f| > a_n\} = 0$. Now

$$\mu(\{|f| > a\}) = \mu\left(\bigcup_n \{|f| > a_n\}\right) = 0.$$

If $a = \|f\|_\infty$, then $\mu(\{|f| > a\}) = 0$. Define

$$g = \begin{cases} f & |f| \le a \\ 0 & |f| > a. \end{cases}$$

Now $g = f$ a.e., and $\|g\|_u = \|f\|_\infty$.

Remark 21.2. If $\mu \ll \nu$ and $\nu \ll \mu$, then $L^\infty(\mu) = L^\infty(\nu)$.

Remark 21.3. On $\mathbb{R}^n$, the set of continuous functions with bounded support is not dense in L^∞. Indeed, $C[0,1] \subseteq L^\infty([0,1], m)$. Then if f is continuous, $\|f\|_\infty = \|f\|_u$.

21.2 Duality of L^p spaces

Let $(X, \mathcal{M}, \mu)$ be a measure space, and let $1 \le p < \infty$. Let q be such that $p^{-1} + q^{-1} = 1$. So $1 < q \le \infty$.

Let $g \in L^q$. Then for all $f \in L^p$,

$$\left| \int fg \, d\mu \right| \le \int |fg| \, d\mu \le \|f\|_p \|g\|_q$$

by Hölder's inequality. So if we define $\varphi_g : L^p \to \mathbb{C}$ sendsing $f \mapsto f \mapsto \int fg \, d\mu$, then $\varphi_g \in (L^p)^*$, and $\|\varphi g\|_{(L^p)^*} \le \|g\|_{L^q}$.

Theorem 21.2 (Riesz representation[12]). *If $1 < p < \infty$, then $L^q \to (L^p)^*$ sending $g \mapsto \varphi_g$ is an isometric isomorphism. The same is true if $p = 1$, provided μ is σ-finite.*

Remark 21.4. When $p = \infty$, $q = 1$. For basically any nontrivial measure, $(L^\infty)^*$ is actually much bigger than L^1.

In this lecture, henceforth, μ is a finite measure. The extension to σ-finite measures is obtained by splitting up the space into countably many pieces and applying these results to each piece.

Proposition 21.1. *If $g \in L^q$, then $\|\varphi_g\| = \|g\|_q$.*

Proof. We have already shown one of the inequalities. If $q < \infty$, (i.e. $p > 1$), then let

$$f := \frac{|g|^{q-1}\overline{\operatorname{sgn}(g)}}{\|g\|_q^{q-1}}.$$

Then, because $p(q-1) = q$, we have

$$|f|^p = \frac{|g|^q}{\|g\|_q^q},$$

so $\int |f|^p = 1$. But now

$$fg = \frac{|g|^q}{\|g\|_q^{q-1}}|g|\operatorname{sgn}(g) = \frac{|g|^q}{\|g\|_q^q},$$

so $\int fg = \|g\|$. That is, $\|\varphi_g\| = \|\varphi_g\|\|f\|_p \geq \int fg = \|g\|_q$.

If $q = \infty$, i.e. g is essentially bounded, let $\varepsilon > 0$. Then $0 < \mu(\{|g| \geq \|g\|_\infty - \varepsilon\}) < \infty$. Now let

$$f = \mathbb{1}_{\{|g|\geq\|g\|_\infty-\varepsilon\}}\overline{\operatorname{sgn}(g)}.$$

Then $f \in L^1$, and $\|f\| = \mu(\{|g| \geq \|g\|_\infty - \varepsilon\})$. Also,

$$\int fg\, d\mu = \int_{\{|g|\geq\|g\|_\infty-\varepsilon\}} |g|\overline{\operatorname{sgn}(g)}\, d\mu \geq (\|g\| - \varepsilon)\|f\|,$$

so $\|\varphi_g\| \geq \|g\|_\infty - \varepsilon$ for all $\varepsilon > 0$. □

Remark 21.5. If μ is finite, $L^q \subseteq L^1$ for all $q \geq 1$.

Proposition 21.2. *Let $g \in L^1$, and let Σ be the set of simple functions on X. Then*

$$\|g\|_q = \sup\left\{\left|\int fg\right| : f \in \Sigma, \|f\|_p \leq 1\right\}.$$

[12] There are many theorems called the Riesz representation theorem, all by the same person. Riesz was a busy guy.

Proof. We already have that $\|g\|_q$ is at least as much as the right hand side, so it is enough to show the reverse.

Step 1: $|fg| \leq$ RHS for bounded, measurable functions: For all such f, $\|f\|_p \leq 1$. Given this f, there exist simple functions $f_n \to f$ pointwise such that $|f_n| \uparrow |f|$. In particular, $f_n \in \Sigma$, and $\|f_n\|_p \leq 1$. Also, $f_n g \to fg$ a.e., and $|f_n g| \leq |fg|$ for all n. Then $fg \in L^1$ because $\|fg\|_1 \leq \|f\|_\infty \|g\|_1$. So, by the DCT, $|\int f_n g| \to |\int fg|$; since the sequence terms are all bounded by the RHS of the inequality we want to show, so is the limit.

Step 2: $\|g\|_q \leq$ RHS. Assume $q < \infty$. There exist simple functions $\varphi_n \to g$ pointwise such that $|\varphi_n| \uparrow |g|$. By the previous proposition, there exist simple functions f_n such that $\|\varphi_n\|_q = |\int f_n \varphi_n|$. Then, by the monotone convergence theorem,

$$\|g\|_q = \lim_n \|\varphi_n\|_q = \lim_n |\int f_n \varphi_n| \leq \lim_n \int |f_n||\varphi_n| \leq \lim_n \int |f_n||g| \leq \text{RHS}\,. \qquad \square$$

We have shown so far that

$$\|g\|_q = \sup\left\{\left|\int fg\right| : f \in L^p, \|f\|_p = 1\right\} = \sup\left\{\left|\int fg\right| : f \in \Sigma, \|f\|_p = 1\right\}.$$

We will finish up the rest of the proof next time.

22 More L^p Duality and Existence of Kernel Operators

22.1 L^p duality, continued

Let's finish up our proof of L^p duality.

Theorem 22.1. *If $1 < p < \infty$, then the map $L^q \to (L^p)^*$ sending $g \mapsto \varphi_g$ is an isometric isomorphism. If μ is σ-finite, the same holds for $p = 1$.*

We are covering the case for when μ is finite. Here is a useful lemma.

Lemma 22.1. *If $\mu(X) < \infty$, then $L^p \subseteq L^1$ for all $p \geq 1$.*

Proof. By Hölder's inequality, if $f \in L^p$, then

$$\int |f| \, d\mu \int |f \mathbb{1}_X| \, d\mu \leq \|f\|_p \|\mathbb{1}_X\|_q \leq \|f\|_p (\mu(X))^{1/q}.$$ □

Last time, we showed the following proposiitions.

Proposition 22.1. $\|\varphi_g\|_{(L^p)^*} = \|g\|_{L^q}$

Proposition 22.2. *If $g \in L^1$ and Σ is the set of simple functions, then*

$$\|g\|_q = \sup\left\{\left|\int fg \, d\mu\right| : f \in \Sigma, \|f\|_p \leq 1\right\}.$$

In particular, the left hand side is ∞ if and only if the right hand side is, as well.

Now we can complete our proof of the main theorem.

Proof. Let $\varphi \in (L^p)^*$. We proceed in steps:

Step 1: For $E \subseteq \mathcal{M}$, define $\nu(E) := \varphi(\mathbb{1}_E)$. This uses the assumption that $\mu(X) < \infty$. We claim that ν is a complex measure on $(X, \mathcal{M})$. We have $\nu(\varnothing) = \varphi(0) = 0$, and finite additivity is not too hard to check. Let's check countable additivity. Let $(E_n)_n \subseteq \mathcal{M}$ be disjoint. Then $\mathbb{1}_{\bigcup_n E_n} = \sum_n \mathbb{1}_{E_n}$. To control the tail of this series, we have

$$\left\|\sum_{n=k}^{\infty} \mathbb{1}_{E_n}\right\| = \left\|\mathbb{1}_{\bigcup_{n=k}^{\infty} E_n}\right\| = \mu\left(\bigcup_{n=k}^{\infty} E_n\right)^{1/p},$$

which goes to 0 since $\mu(X) < \infty$ and $p < \infty$. So by continuity of φ on L^p, we have

$$\nu\left(\bigcup_n E_n\right) = \varphi\left(\mathbb{1}_{\bigcup_n E_n}\right) = \varphi\left(\sum_n \mathbb{1}_{E_n}\right) = \sum_n \nu(E_n).$$

Step 2: Also, $\nu \ll \mu$. Indeed, if $\mu(E) = 0$, then $\mathbb{1}_E = 0$ μ-a.e. So $\nu(E) = 0$. By the Radon-Nikodym theorem, there exists $g \in L^1_{\mathbb{C}}(\mu)$ such that $d\nu = g\,d\mu$.

Step 3: If $f \in \Sigma$, then $\int fg\,d\mu = \int f\,d\nu = \varphi(f)$ by linearity. We know this is bounded in absolute value by $\|\varphi\|_{(L^p)^*}\|f\|_p$. Our propositions give us that $g \in L^q$ and $\|g\|_q \leq \|\varphi\|_{(L^p)^*}$. We know that that $\varphi_g|_\Sigma = \varphi|_\Sigma$. So $\varphi_g = \varphi$ on a dense subspace of L^p, so continuity gives that $\varphi_g = \varphi$. □

Corollary 22.1. *If $1 < p < \infty$, then L^p is reflexive.*

Proof. We know $1 < q < \infty$, so $(L^p)^{**} = (L^q)^* = L^p$. □

Remark 22.1. For interesting μ, L^1 and L^∞ are not reflexive.

22.2 Existence of kernel operators in L^p

Theorem 22.2. *Let $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ be σ-finite measure spaces. Suppose*

1. *$K : X \times Y \to \mathbb{C}$ is measurable,*
2. *there exists $C > 0$ such that $\|(x, \cdot)\|_{L^1(\nu)} \leq C$ for μ-a.e. x,*
3. *there exists $C > 0$ such that $\|(\cdot, y)\|_{L^1(\mu)} \leq C$ for ν-a.e. y.*

Then there for all $p \in [1, \infty]$ and $f \in L^p(\nu)$, the integral

$$Tf(x) = \int_Y K(x, y)f(y)\,d\nu(y)$$

exists μ-a.e., $Tf \in L^p(\mu)$, and $\|Tf\|_{L^p(\mu)} \leq C\|f\|_{L^p(\nu)}$.

We will check the cases where $p \neq 1, \infty$.

Proof. The conjugate exponent $q \in (1, \infty)$. Let $x \in X$. Here is the key idea:

$$|K(x, y)f(y)| = |K(x, y)|^{1/q}\left(|K(x, y)|^{1/p}|f(y)|\right).$$

Apply Hölder's inequality to get

$$\begin{aligned}\int |K(x, y)f(y)|\,dy &\leq \left(\int |K(x, y)|\,d\nu(y)\right)^{1/q}\left(\int \left(|K(x, y)|^{1/p}|f(y)|\right)^p\right) d\nu(y)^{1/p} \\ &\leq C^{1/q}\left(\int |K(x, y)||f(g)|^p\,d\nu(y)\right)^{1/p}.\end{aligned}$$

By Tonelli's theorem,

$$\begin{aligned}\int \left[\int |K(x,y)||f(g)|^p \, d\nu(y)\right] \, d\mu(x) &= \int \left[\int |K(x,y)| \, d\mu(x)\right] |f(y)|^p |f(y)|^p \, d\nu(y) \\ &\leq C \int |f(y)|^p \, d\nu(y) = C\|f\|_{L^p(\nu)}.\end{aligned}$$

Overall, we get

$$\begin{aligned}\int \left[\int |K(x,y)||f(y)| \, d\nu(y)\right]^p \, d\mu(x) &\leq C^{p/q} \int \int |K(x,y)||f(y)|^p \, d\nu(y) \, d\mu(x) \\ &\leq C^{p/q} C \|f\|_{L^p(\nu)}^p \\ &= C^p \|f\|_{L^p(\nu)}^p.\end{aligned}$$

So $Tf(x)$ is well-defined μ-a.e., and $\|Tf\|_p^p \leq \text{LHS} \leq C^p \|f\|_p^p$, so $\|T\|_{\mathcal{L}(L^p, L^p)} \leq C$. $\square$

23 Translation Operators and Relationships Between L^p Spaces

23.1 Translation operators on L^p spaces

Let m be Lebesgue measure on $\mathbb{R}^d$, and let $t \in \mathbb{R}^d$. Let $\tau_t : \mathbb{R}^d \to \mathbb{R}^d$ send $v \mapsto v - \tau$. This is translation by t, and Lebesgue measure is translation invariant.

Lemma 23.1. *The map T_t sending $f \mapsto f \circ \tau_t$ is an isometry $L^p(m) \to L^p(m)$ for all p.*

However, the functions T_t are not kernel operators.

Lemma 23.2. *Let $p < \infty$. Let $(t_n)_n$ in $\mathbb{R}^d$ be such that $t_n \to 0$. Then $T_{t_n} \to \mathrm{Id}$ on $L^p(m)$ in the strong operator topology but not in $\| \cdot \|_{\mathrm{op}}$.*

Proof. $C_c(\mathbb{R}^d)$ is dense in $L^p(m)$. Let $f \in L^p$. Suppose first that $f \in C_c(\mathbb{R}^d)$. Pick R such that $f|_{B_R^c} = 0$. Given $\varepsilon > 0$, there exists $\delta > 0$ such that if $|z - y| < \delta \implies |f(x) - f(y)| < \varepsilon$. If $|t_n| < \delta$, then

$$\begin{aligned}
\|T_{t_n} f - f\|_p^p &= \int_{\mathbb{R}^d} |f(x - t_n) - f(x)|^p \, dm(x) \\
&= \int_{B_{R+1}} |f(x - t_n) - f(x)|^p \, dm(x) \\
&\leq \varepsilon^p m(B_{R+1}) \\
&\xrightarrow{t_n \to 0} 0.
\end{aligned}$$

Similarly, the map $\mathbb{R}^d \to \mathcal{L}(L^p(m), L^p(m))$ sending $t \mapsto T_t$ is continuous from $\mathbb{R}^d$ to the strong operator topology. For general $f \in L^p(m)$, let $\varepsilon > 0$. Choose $g\ inC_c(\mathbb{R}^d)$ such that $\|f - g\|_p < \varepsilon/3$. Choose n large enough such that $\|T_{t_n} g - g\|_p < \varepsilon/3$. Put together,

$$\begin{aligned}
\|T_{t_n} f - f\|_p &\leq \|T_{t_n}(f - g)\| + \|T_{t_n} g - g\|_p + \|f - g\|_p \\
&\leq \|f - g\| + \|T_{t_n} g - g\|_p + \|f - g\|_p \\
&< \varepsilon.
\end{aligned}$$

Now let's show that this convergence does not occur in the norm topology. For any $t \neq 0$, there exist $f \in C_c(\mathbb{R}^d$ such that $\|f\|_p = 1$ and $f|_{B_{t/2}^c} = 0$. Then

$$\|T_t f - f\|_p = 2^{1/p} \|f\|_p.$$

□

23.2 Relationships between L^p spaces

What is the relationship between L^p spaces for different p?

Example 23.1. Look at $((0, \infty)), \mathcal{B}_{(0,\infty)}, m)$. Let $1 \leq p < q < \infty$. Let $f_a(x) = x^{-a}$ for some choice of $a > 0$. Observe:

1. The function $f_a \mathbb{1}_{(0,1)} \in L^p$ iff $p < 1/a$.
2. The function $f_a \mathbb{1}_{(1,\infty)} \in L^p$ iff $p > 1/a$.

So $L^p \setminus L^q \neq \varnothing$, and $L^q \setminus L^p \neq \varnothing$.

Proposition 23.1. *If $0 < p < q < r \leq \infty$, then $L^q \subseteq L^p + L^r$.*

Proof. Let $f \in L^q$. Write $f = f\mathbb{1}_{\{|f|>1\}} + f\mathbb{1}_{\{|f|\leq 1\}}$. Then

$$\|f\mathbb{1}_{\{|f|>1\}}\|_p^p = \int_{\||f|>1\}} |f|^p \, d\mu \leq \int_{\||f|>1\}} |f|^q \, d\mu \leq \int |f|^q \, d\mu = \|f_q\|^q < \infty.$$

The same holds for $f\mathbb{1}_{\{|f|\leq 1\}}$. □

Proposition 23.2. *If $0 < p < q < r \leq \infty$, then $L^p \cap L^r \subseteq L^q$, and $\|f\|_q \leq \|f\|_p^\lambda \|f\|_r^{1-\lambda}$, where $1/q = \lambda(1/p) + (1-\lambda)(1/r)$.*

Proof. It suffices to prove the inequality.

$$\int |f|^q \, d\mu = \int |f|^{\lambda q} |f|^{(1-\lambda)q} \, d\mu$$

Use Hölder's inequality, where $1/s + 1/t = 1$. We will pick the values of s, t later to make sure they work out.

$$\leq \left(\int |f|^{\lambda q s} \, d\mu \right)^{1/s} \left(\int |f|^{(1-\lambda)qt} \, d\mu \right)^{1/t}$$

Pick $s = p/(\lambda q)$ to make things work out as stated in the theorem.

$$\leq \|f\|_p^{\lambda q} \|f\|_r^{(1-\lambda)q}.$$ □

There is, however, a case where the tails of functions do not count.

Lemma 23.3. *If $\mu(X) < \infty$ and $0 < p < q \leq \infty$, then $L^p \supseteq L^q$. In particular $\|f\|_p \leq \|f\|_q \mu(X)^{1/p-1/q}$.*

Proof. Let $f \in L^q$. Then, by Hölder's inequality,

$$\|f\|_p^p = \int |f|^p \mathbb{1}_X \, d\mu \leq \|f\|_q (\mu(X))^{1/p-1/q}.$$ □

Lemma 23.4. *Let A be any set. Let $\ell^p(A) = L^p(A, \mathscr{P}(A), \#)$. Then $\ell^p \subseteq \ell^q$.*

Proof. If $q = \infty$, then

$$\sup_\alpha |f(\alpha)| = (\sup_\alpha |f(\alpha)|^p)^{1/p} \leq \left(\sum_\alpha |f(\alpha)|^p \right)^{1/p} = \|f\|_p.$$

If $p < q < \infty$, then by the previous lemma,

$$\|f\|_q \leq \|f\|_p^\lambda \|f\|_\infty^{1-\lambda} \leq \|f\|_p^{\lambda+1-\lambda} = \|f\|_p.$$ □

23.3 Distribution functions

Fix $(X, \mathcal{M}, \mu)$, and let $f : X \to \mathbb{C}$ be measurable.

Definition 23.1. The **distribution function** of f, $\lambda_f : (0, \infty) \to [0, \infty]$, is

$$\lambda_f(\alpha) = \mu(\{|f| > \alpha\}).$$

Proposition 23.3. *Let λ_f be the distribution function of f.*

1. *λ_f is non-increasing and right-continuous.*
2. *If $|f| \leq |g|$, then $\lambda_f \leq \lambda_g$.*
3. *If $|f_n| \uparrow |f|$, then $\lambda_{f_n}(\alpha) \uparrow \lambda_f(\alpha)$.*
4. *If $f = g + h$, then $\lambda_f(\alpha) \leq \lambda_g(\alpha/2) + \lambda_h(\alpha/2)$.*

24 Distributions, Weak L^p, Strong Type, and Weak Type

24.1 Distributions

Last time, we introduced the notion of a distribution function $\lambda_f(\alpha) = \mu(\{|f| > \alpha\})$.

Definition 24.1. Let $f \geq 0$ and $\mu(X) < \infty$. Then the **distribution** of f is the measure $\nu(E) = \mu(\{x \in X : f(x) \in E\})$.

Observe that

$$\nu(a, b) = \mu(\{a < f \leq b\}) = \lambda_f(a) - \lambda_f(b) = [-\lambda_f(b)] - [-\lambda_f(a)].$$

So λ_f determines the measure ν and basically contains all the information about how much measure the range of f has in given sets.

Proposition 24.1 (Chebyshev's inequality)**.** *Let $0 < p < \infty$, and let $f \in L^p$. Then $\lambda_f(\alpha) \leq \|f\|_p^p / \alpha^p$.*

Remark 24.1. When $p = 1$, this is called Markov's inequality.[13]

Proof. $\lambda_f(\alpha) = \mu(\{f > \alpha\}) =: \mu(E_\alpha)$. By definition, $\mathbb{1}_{E_\alpha} \alpha^p \leq |f|^p$. Then

$$\mu(E_\alpha) = \alpha^p \int \mathbb{1}_{E_\alpha} \, d\mu \leq \int |f|^p \, d\mu. \qquad \square$$

24.2 Weak L^p

Definition 24.2. If $f : X \to \mathbb{C}$, the **"weak L^p norm** of f is

$$[f]_p = (\sup_{\alpha > 0} \alpha^p \lambda_f(\alpha))^{1/p}.$$

Remark 24.2. This is generally not actually a norm; the triangle inequality fails. Chebyshev's inequality says that

$$[f]_p \leq \|f\|_p.$$

Definition 24.3. The **weak L^p space** is

$$\operatorname{wk} L^p(\mu) = \{f : X \to \mathbb{C} \mid [f]_p < \infty\} / \sim,$$

under the equivalence relation of μ-a.e. equality.

By Chebyshev's inequality, $\operatorname{wk} L^p \supseteq L^p$.

[13] Markov was Chebyshev's advisor. Chebyshev is responsible for noticing that the inequality holds in general.

Example 24.1. Let m be Lebesgue measure on $(0, \infty)$. Consider $f(x) = x^{-1/p}$. Then $f \notin L^p(m)$. But

$$[f_p] = \sup_\alpha m(\{f > \alpha\}) = \sup_\alpha \alpha^p = \sup_\alpha m([0, 1/\alpha^p))\alpha^p = 1$$

Proposition 24.2. *Let $0 < p < \infty$, and let $f : X \to \mathbb{C}$. Then*

$$\|f\|_p^p = \int |f|^p \, d\mu = p \int_0^\infty \alpha^{p-1} \lambda_f(\alpha) \, d\alpha.$$

Proof. If there exists α such that $\lambda_f(\alpha) = \infty$, the the right hand side is infinite. By Chebyshev's inequality, so is the left hand side. So we may assume that $\lambda_f(\alpha) < \infty$ for all α. Then $\{f \neq 0\}$ is σ-finite. So we may assume μ is σ-finite.

Now consider $E = \{(x, y) \in X \times [0, \infty) : y < |f(x)|^p\}$. Now, by Tonelli's theorem,

$$\begin{aligned}
\int_X |f|^p \, d\mu &= \int_X \int_0^{|f(x)|^p} dy \, d\mu(x) \\
&= (\mu \otimes m)(E) \\
&= \int_0^\infty \mu(\{|f|^p > y\}) = p \int \alpha^{p-1} \lambda_f(\alpha) \, d\alpha,
\end{aligned}$$

where we have used the substitution $y = \alpha^p$. □

24.3 Strong type and weak type

Definition 24.4. Let $\mathcal{D}$ be some vector space of measurable $\mathbb{C}$-valued functions on $(X, \mathcal{M}, \mu)$, and let $T : \mathcal{D} \to L^0(Y, \mathcal{N}, \nu)$ (the space of measurable functions. T is **sublinear** if

1. $c > 0 \implies |T(cf)| = c|Tf|$ for all $f \in \mathcal{D}$
2. $|T(f_1 + f_2)(x)| \leq |Tf_1(x)| + |Tf_2(x)|$.

Example 24.2. Let $\mathcal{D} = L^1_{\text{loc}}$. The **Hardy-Littlewood maximal operator** is

$$Hf(x) = \sup_{r>0} \frac{1}{m(B(x, r))} \int_{B(x,r)} |f(y)| \, dy.$$

Then $H(f_1 + f_2) \leq Hf_1 + Hf_2$, so H is sublinear.

Remark 24.3. Often, sublinear functions arise from taking the pointwise supremum of a collection of linear functions.

Definition 24.5. T is **strong type** (p, q) for $1 \leq p, q \leq \infty$ if

1. $L^p(\mu) \subseteq \mathcal{D}$

2. $T[L^p(\mu)] \subseteq L^q(\nu)$, and $\|Tf\|_q \leq C\|f\|_p$ for some fixed $C > 0$.

Definition 24.6. T is **weak type** (p, q) for $1 \leq p, q \leq \infty$ if

1. $T[L^p(\mu)] \subseteq \operatorname{wk} L^q(\nu)$

2. $[Tf]_q \leq C[f]_p$ for some fixed $C > 0$.

Example 24.3. H is strong type $(, \infty, \infty)$. The Hardy-Littlewood maximal inequality says that H is weak type $(1, 1)$.

Theorem 24.1 (special case of Marciukiewicz' theorem)**.** *Suppose T is sublinear on $\mathcal{D} = L^1(\mu) + L^\infty(\mu)$. Suppose T is weak type $(1, 1)$ and strong type (∞, ∞). Then T is strong type (p, p) for all $p \in (0, \infty]$.*

Remark 24.4. $L^1(\mu) + L^\infty(\mu) \supseteq L^p(\mu)$ for all p.

Example 24.4. The Hardy-Littlewood maximal operator is strong type (p, p) for all $p \in [0, \infty]$. This is very difficult to prove by hand.

Proof. Pick a C such that $\|Tf_{\infty} \leq C\|f_{\infty}$ and $[Tf]_1 \leq C\|f\|_1$ for all $f \in L^\infty$ or L^1. Let $f \in L^p(\mu$, and let $A > 0$. Write $f = f_1 + f_2$, where $f_1 = f\mathbb{1}_{\{|f|>A\}}$ and $f_2 = f\mathbb{1}_{\{|f|\leq A\}}$. We will optimize over the value of A. We have

$$\|f_1\| = \int_{\{|f|>A\}} |f| \leq \int_{\{|f|>A\}} \frac{|f|^p|}{A^{p-1}} = \frac{A^{p-1}}{\int}{}_{\{|f|>A\}} |f|^p < \infty.$$

By subliniearity,

$$|Tf(x)| \leq |Tf_1(x)| + \underbrace{|Tf_2(x)|}_{\leq CA}.$$

So

$$\mu(\{|Tf| > 2CA\}) \leq \mu(\{|Tf_1 > CA\}) \leq \frac{C\|f_1\|}{CA} = \frac{1}{A}\int_{\{|f|>A\}} |f|.$$

So we have improved the weak type $(1, 1)$ inequality to get

$$\lambda_{Tf}(2CA) \leq \frac{1}{A}\int_{\{|f|>A\}} |f|.$$

Substitute this expression into the following:

$$\|Tf\|_p^p = p\int_0^\infty \alpha^{p-1}\lambda_{Tf}(\alpha)\, d\alpha \leq (2C)^p \frac{p}{p-1}\|f\|_p^p. \qquad \square$$

25 Introduction to Hilbert Spaces

25.1 Motivation

Consider $(X, \mathcal{M}, \mu) = (\{1, \dots, n\}, \mathscr{P}(X), \#)$. Then $L^p_{\mathbb{C}}(\mu) = \ell^p(n) = \mathbb{C}^n$. In this case, we are specifying a specific norm:

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}.$$

These give different shapes for the unit ball; try drawing the unit ball for different values of p when $n = 2$.

A linear functional φ on $\mathbb{C}^n$ has the form

$$\varphi(x) = \sum_{i=1}^n x_i \overline{y}_i = \langle x, y \rangle$$

for some $y = (y_1, \dots, y_n) \in \mathbb{C}^n$. So $\varphi \in (\ell^p(n))^*$; that is, every linear functional is continuous. The Riesz representation theorem says that

$$\|\varphi_y\|_{(\ell^p(n))^*} = \sup\{|\varphi_y(x)| : \|x\|_p \leq 1\} = \|y\|_{\ell^q},$$

where $1/p + 1/q = 1$.

There is a special case, when $p = 2$. We get that the dual norm is the original norm. So we can think of $\ell^2(n)$ as its own dual.

Definition 25.1. Let H be a vector space over $\mathbb{C}$. An **inner product** on H is a map $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{C}$ sending $(x, y) \mapsto \langle x, y \rangle$ such that

1. (bilinearity) $\langle ax + by, z \rangle = a \langle x, y \rangle + b \langle x, z \rangle$ for all $a, b \in \mathbb{C}$, $x, y, z \in H$,
2. (conjugate symmetry) $\langle x, y \rangle = \overline{\langle y, x \rangle}$,
3. $\langle x, x \rangle \in [0, \infty)$ with $\langle x, x \rangle = 0$ iff $x = 0$.

Example 25.1. $\mathbb{C}^n$ is a vector space with the usual inner product.

Example 25.2. $L^2_{\mathbb{C}}(\mu)$ is a vector space with the inner product $\langle f, g \rangle = \int_X f\overline{g}\, d\mu$.

Example 25.3. Let $X = \mathbb{N}$ with counting measure. Then

$$\ell^2 = \ell^2(\mathbb{N}) = \{(x_n)_n : \sum_n |x_n|^2 < \infty\}$$

has the inner product $\langle x, y \rangle = \sum_n x_n \overline{y}_n$.

Definition 25.2. A vector space $(H, \langle \cdot, \cdot \rangle)$ is a **pre-Hilbert space** (or **inner product space**).

25.2 Norms induced by inner products

An inner product space has the associated norm $\|x\| := \sqrt{\langle x, x\rangle}$. First, we have to show that this is actually a norm.

Lemma 25.1 (Cauchy-Bunyakowski-Schwarz inequality[14]). *For all $x, y \in H$,*

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

Proof. Consider $\langle x - ty, x - ty\rangle$. We get

$$\begin{aligned} 0 &\leq \langle x - ty, x - ty\rangle \\ &= \langle x, x\rangle - t\langle y, x\rangle - t\langle x, y\rangle + t^2\langle x, y\rangle \\ &= \|x\|^2 - 2t\operatorname{Re}(\langle x, y\rangle) + t^2\|y\|^2. \end{aligned}$$

This achieves its minimum at $t = \operatorname{Re}(\langle x, y\rangle)/\|y\|^2$. So we get

$$0 \leq \|x\|^2 - \frac{(\operatorname{Re}(\langle x, y\rangle))^2}{\|y\|^2},$$

which gives

$$|\operatorname{Re}(\langle x, y\rangle)| \leq \|x\|\|y\|.$$

Similarly, let $\alpha = \operatorname{sgn}(\langle x, y\rangle)$, an apply this to x and $y' = \alpha y$. Then

$$|\langle x, y\rangle| = |\operatorname{Re}(x, y')| \leq \|x\|\|y'\| = \|x\|\|y\|. \qquad \square$$

Corollary 25.1. $\|\cdot\|$ *is a norm.*

Proof.

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

Definition 25.3. A **Hilbert space** is a complete pre-Hilbert space.

Example 25.4. All the previous examples are complete.

Proposition 25.1. $\langle\cdot,\cdot\rangle : H \times H \to \mathbb{C}$ *is continuous for the norm topology on H.*

[14] Bunyakowski and Schwarz both knew the general form of this inequality, but, due to geopolitics, there was no way they could have ever met.

Proof. Suppose that $x_n \to x$ in norm and $y_n \to y$ in norm. Then

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

□

Proposition 25.2 (Parallelogram law)**.** *For all $x, y \in H$,*

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

Proof. Expand out $\langle x+y, x+y\rangle$ and cancel terms. □

25.3 Orthogonality

Definition 25.4. Elements $x, y \in H$ are **orthogonal** if $\langle x, y\rangle = 0$.

Definition 25.5. If $E \subseteq H$, its **orthogonal complement** is

$$E^\perp = \{x \in : \langle x, y\rangle = 0 \; \forall y \in E\}.$$

Theorem 25.1 (Pythagorean theorem[15])**.** *If $x_1, \ldots, x_n \in H$ are pairwise orthogonal, then*

$$\left\|\sum_i x_i\right\| = \sum_i \|x_i\|^2.$$

Proof. Expand $\left\langle \sum_i x_i, \sum_j x_j \right\rangle$ and cancel terms. □

Theorem 25.2. *Let H be a Hilbert space, and let M be a closed subspace. Then any $x \in H$ can be written uniquely as $x = y + z$, where $y \in M$ and$z \in M^\perp$. We write $H = M \oplus M^\perp$.*

Proof. Let $\delta = \inf\{\|x - y\| : y \in M\}$. Pick $(y_n)_n$ in M such that $\|x - y_n\| \to \delta$. We claim that (y_n) is Cauchy. We have

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

Rewrite this as

$$\|y_n - y_m\|^2 + 4\underbrace{\left\|\frac{y_n + y_m}{2} - x\right\|^2}_{\to\delta^2} = 2(\underbrace{\|y_n - x\|^2}_{\to\delta^2} + \underbrace{\|y_m - x\|^2}_{\to\delta^2}).$$

This is only possible if $\|y_n - y_m\| \to 0$.

So the limit $y = \lim_n y_n$ exists. This is the unique closest point in M to x. □

[15]A person named Pythagoras probably didn't exist. Nevertheless, the Pytagoreans almost surely did not know what a Hilbert space is.

25.4 Isomorphisms of Hilbert spaces

Definition 25.6. A **unitary operator** $U : H_1 \to H_2$ is linear operator such that $U \in \mathcal{L}(H_1, H_2)$ is an isomorphism and $\langle Ux, Uy \rangle_2 = \langle x, y \rangle_1$.

This is the true notion of isomorphism for inner product spaces. Next time, we will prove the following theorem:

Theorem 25.3. *Let H be a Hilbert space over $\mathbb{C}$.*

1. *If $\dim(H) = n < \infty$, then $H \cong \mathbb{C}^n$.*
2. *If $\dim(H) = \infty$ and H is separable, then $H \cong \ell^2(\mathbb{N})$.*

Example 25.5. $L^2(\mathbb{R})$ is separable, so $L^2(\mathbb{R}) \cong \ell^2(\mathbb{N})$.

Example 25.6. The **Fourier transform** is the unitary equivalence $L^2([0,1]) \cong \ell^2(\mathbb{Z})$.

26 Riesz Represetntation for Hilbert Spaces and Orthonormality

26.1 Riesz representation for Hilbert spaces

Let's finish up a proof from last time.

Theorem 26.1. *Let H be a Hilbert space, and let M be a closed subspace. Then any $x \in H$ can be written uniquely as $x = y + z$, where $y \in M$ and $z \in M^\perp$. We write $H = M \oplus M^\perp$.*

Proof. Let $\delta = \inf\{\|x - y\| : y \in M\}$. Pick $(y_n)_n$ in M such that $\|x - y_n\| \to \delta$. We claim that (y_n) is Cauchy. We have

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

Rewrite this as

$$\|y_n - y_m\|^2 + 4\underbrace{\left\|\frac{y_n + y_m}{2} - x\right\|^2}_{\to \delta^2} = 2(\underbrace{\|y_n - x\|^2}_{\to \delta^2} + \underbrace{\|y_m - x\|^2}_{\to \delta^2}).$$

This is only possible if $\|y_n - y_m\| \to 0$. So the limit $y = \lim_n y_n$ exists. Moreover, $\|y - x\| = \delta$. This point is unique; if we had y, y' with the same property, the same identity above gives $\|y - y'\| = 0$.

It now remains to show that $z = x - y \in M^\perp$. Suppose not, and choose $v \in M$ such that $|\langle z, v\rangle| \in (0, \infty)$. Now consider

$$\|x - (y + tv)\|^2 = \underbrace{\|z\|^2}_{=\delta^2} + t^2\|v\|^2 - 2t\,\mathrm{Re}(\{z, v\}).$$

This can be made $< \delta^2$ unless $\langle z, v\rangle = 0$. $\square$

The first part of this proof is appealing to a particular property which does not only hold just in Hilbert spaces.

Definition 26.1. A Banach space $(\mathcal{X}, \|\cdot\|)$ is **uniformly convex** if for all $\varepsilon > 0$, there exists a $\delta > 0$ such that whenever $x, y \in \mathcal{X}$ with $\|x\| = \|y\| = 1$, if $\|(x + y)/2\| > 1 - \delta$, then $\|x - y\| < \varepsilon$.

Example 26.1. For $1 < p < \infty$, $L^p(\mathbb{R})$ is uniformly convex.

Theorem 26.2 (Riesz[16])**.** *For any $f \in H^*$, there exists $y \in H$ such that $f(x) = \langle x, y\rangle$ and $\|f\| = \|y\|$.*

[16] This is yet another theorem called the Riesz representation theorem.

Proof. Assume $f \neq 0$, and let $M = \{x : f(x) = 0\}$. This is a closed, proper subspace of H. By the previous theorem, there must exist a $z \in H$ such that z is orthogonal to M. So pick $z \in M^\perp$ with $\|z\| = 1$. For any $x \in H$, consider $u = f(x)z - f(z)x$, which lies in M. So

$$0 = \langle u, z \rangle = f(x) \cdot 1 - f(z) \langle x, z \rangle .$$

That is, $f(x) = f(z) \langle x, z \rangle = \langle x, \overline{f(z)}z \rangle = \langle x, y \rangle$. ☐

Corollary 26.1. *Hilbert spaces are reflexive.*

26.2 Orthonormality

Definition 26.2. Let $(u_\alpha)_{\alpha \in H}$ be a collection of vectors in H. The collection is **orthonormal** if $\|u_\alpha\| = 1$ for all α, and when $\alpha \neq \beta$, $\langle u_\alpha, u_\beta \rangle = 0$.

Proposition 26.1 (Bessel's inequality)**.** *If* $(u_\alpha)_\alpha$ *is orthonormal in* H*, then*

$$\sum_\alpha | \langle x, u_\alpha \rangle |^2 \leq \|x\|.$$

Remark 26.1. When we are dealing with an uncountable set of vectors, we mean that all but countably many of them are orthonormal to x, so the sum makes sense.

Proof. Suppose $F \subseteq A$ and $|F| < \infty$. Then

$$\begin{aligned} 0 \leq \left\| x - \sum_{\alpha \in F} \langle x, u_\alpha \rangle u_\alpha \right\| &= \|x\|^2 - 2 \operatorname{Re} \left\langle x, \sum_{\alpha \in F} \langle x, u_\alpha \rangle u_\alpha \right\rangle + \sum_{\alpha \in F} | \langle x, u_\alpha \rangle |^2 \\ &= \|x\|^2 - 2 \sum_{\alpha \in F} | \langle x, u_\alpha \rangle |^2 + \sum_{\alpha \in F} | \langle x, u_\alpha \rangle |^2 \\ &= \|x\|^2 - \sum_{\alpha \in F} | \langle x, u_\alpha \rangle |^2. \end{aligned}$$

☐

Theorem 26.3. *Let* $(u_\alpha)_\alpha$ *be an orthonormal set in* H*. The following are equivalent:*

1. *(completeness[17]) If* $\langle x, u_\alpha \rangle = 0$ *for all* α*, then* $x = 0$*.*
2. *(Parseval's identity) Besel's inequality is an equality for all* x*.*
3. *For all* $x \in H$*, we have* $x = \sum_{\alpha \in A} \langle x, u_\alpha \rangle u_\alpha$*.*

Remark 26.2. It is possible for the sum of the lengths in (3) to be infinite, so this sum is not absolutely convergent. But the sum of the squares must be finite, as shown by part (b).

[17] Add this to the list of the most overused words in mathematics.

Proof. (1) $\implies$ (3). Pick x. Bessel's inequality gives $\|x\|^2 \geq \sum_\alpha |\langle x, u_\alpha\rangle|^2$. So there are only countably many nonzero terms. Enumerate them as $\alpha_1, \alpha_2, \dots$. Consider $\sum_{i=1}^n \langle x, u_{\alpha_i}\rangle u_{\alpha_i}$. If $m > n$,

$$\left\| \sum_{i=n+1}^{m} \langle x, u_{\alpha_i}\rangle u_{\alpha_i} \right\| = \sum_{i=n+1}^{m} |\langle x, u_\alpha\rangle|^2 \xrightarrow{n,m\to\infty} 0.$$

So $\sum_{i=1}^n \langle x, u_{\alpha_i}\rangle u_{\alpha_i}$ is a Cauchy sequence, so it converges to some y. Now $y = x$ because for all α,

$$\langle y, \alpha\rangle = \begin{cases} \langle x, u_\alpha\rangle & \alpha = \alpha_i \text{ for some } i \\ 0 = \langle x, u_\alpha\rangle & \alpha \notin \{\alpha_1, \alpha_2, \dots\}. \end{cases}$$

This implies $y = x$ by (a).

(3) $\implies$ (2): Look that

$$\left\| x - \sum_{i=1}^{n} \langle x, u_{\alpha_i}\rangle u_{\alpha_i} \right\|.$$

This is the gap we found in Bessel's inequality. So we get Parseval in the limit as $n \to \infty$.

(2) $\implies$ (1): If $\|x\|^2 = \sum_\alpha |\langle x, u_\alpha\rangle|^2$, and the left hand side is nonzero, then there exists α such that $\langle x, u_\alpha\rangle \neq 0$. □

Definition 26.3. Any orthonormal set satisfying the previous theorem is a **basis**.

Theorem 26.4. *Any Hilbert space H has an orthonormal basis.*

Proof. Use Zorn's lemma. □

Remark 26.3. A basis is also a maximal orthonormal set.

Proposition 26.2. *H is separable if and only if it has a countable basis.*

Theorem 26.5. *Let H be a Hilbert space over $\mathbb{C}$.*

1. *If $\dim(H) = n < \infty$, then $H \cong \mathbb{C}^n$.*
2. *If $\dim(H) = \infty$ and H is separable, then $H \cong \ell^2(\mathbb{N})$.*

Proof. Suppose $\dim(H) = \infty$. Pick a basis $\{u_1, u_2, u_3, \dots\}$. For each $x \in H$, map $x \mapsto (\langle x, u_1\rangle, \langle x, u_2\rangle, \dots) \in \ell^2$. Parseval's identity says exactly that this is a unitary equivalence. □