Math 202A

Topology and Analysis Notes

Instructor: Marc Rieffel

Bryan Li

Fall 2022

Contents

Part I

Topology

Administrative Matters

Midterm date:

1. Wed Oct 19

Weekly psets due Friday midnight.

1 Metric Spaces

Definition 1.1 (Metric Space). A metric on a set X is a function $d\colon X^2 \to \mathbb{R}_{\geq 0}$ such that

1. $d(x,y) = 0$ if and only if $x = y$.
2. $d(x,y) = d(y,x)$.
3. $d(x,z) \leq d(x,y) + d(y,z)$.

(X,d) is then called a metric space.

If condition 1 is omitted and replaced only with the condition that $d(x,x) = 0$, then d is called a semimetric (or quasimetric). If distances are allowed to be infinite, then d is called an extended metric.

If (X,d) is a semimetric space, then if we define a relation $\sim$ such that $x \sim y$ iff $d(x,y) = 0$, it is an equivalence relation. In this case, d "drops" to a metric on the equivalence classes $X/\sim$. To show this, we just need to prove that $\tilde{d}([x],[y]) = d(x,y)$ is a well-defined function, which utilizes the triangle inequality.

Example 1.1. If G is a connected weighted graph, then we can define a metric on G by defining the distance between two vertices to be the length of the minimal path connecting them.

If G is unconnected, by assigning the distance of infinity between points in separate components, we define a extended metric.

Definition 1.2 (Norm). If V is a vector space then $\|\cdot\|\colon V \to \mathbb{R}_{\geq 0}$ is a norm on V if

1. $\|v\| = 0$ iff $v = 0$
2. $\|\lambda v\| = |\lambda|\,\|v\|$
3. $\|v + w\| \leq \|v\| + \|w\|$

A norm on V clearly induces a metric on V by defining $d(v,w) = \|v - w\|$.

Example 1.2 (Euclidean metric). The real line $\mathbb{R}$ and the euclidean spaces $\mathbb{R}^n$ with the euclidean metric $d(x,y) = \|x - y\|_2$.

Example 1.3 (p-norms). The p-norms $\left\|(x_1,\ldots,x_n)\right\|_p = \left(\sum x_i^p\right)^{1/p}$ induce a metric on $\mathbb{R}^n$.

Definition 1.3 (Restricted metric). If (X,d) is a metric space and Y is a subset of X, then we can restrict d to Y and $d|_Y$ is a metric on Y.

It is common to consider the restriction of the Euclidean metric onto subsets of Euclidean space.

Example 1.4 (L^p norms). If $V = C^0([0,1])$, then examples of norms on V include

1. $\|f\|_\infty = \sup |f(t)|$.
2. $\|f\|_1 = \int_0^1 |f(t)|\,\mathrm{d}t$
3. $\|f\|_2 = \left(\int_0^1 |f(t)|^2\,\mathrm{d}t\right)^{1/2}$
4. $\|f\|_p = \left(\int_0^1 |f(t)|^p\,\mathrm{d}t\right)^{1/p}$ for $1 \le p < \infty$.[1]

This generalizes to bounded closed regions of $\mathbb{R}^n$, where all continuous functions are integrable (due to compactness and the extreme value theorem). The proof that these are actually norms that satisfy the triangle inequality will come later (see the Minkowski and Hölder inequalities).

1.1 Topology of Metric Spaces

Categories

Categories contain objects and morphisms, that satisfy the following properties:

1. There is a morphism from any object A to itself, called the identity 1_A.
2. Morphisms can be composed.

Metric spaces can form a category under a few choices of collections of morphisms:

- Isometries: $i\colon (X, d_X) \to (Y, d_Y)$ such that $d_X(x, y) = d_Y(i(x), i(y))$ for all x, y.

  Surjective isometries form a subcategory of the above category, ,where every map is an isomorphism.

- Lipschitz maps: $f\colon (X, d_X) \to (Y, d_Y)$, where there exists a constant C (which may depend on f), such that $d_Y(f(x), f(y)) \le C d_X(x, y)$.

  We can also form the subcategory of bi-Lipschitz isomorphisms.

- Uniformly continuous maps: $f\colon (X, d_Y) \to (Y, d_Y)$ such that for all $\varepsilon > 0$ there exists $\delta > 0$ such that when $d_X(x, y) \le \delta$, then $d_Y(f(x), f(y)) < \varepsilon$.

- Continuous maps: $f\colon (X, d_Y) \to (Y, d_Y)$ such that for all $x \in X$, for all $\varepsilon > 0$ there exists $\delta > 0$ such that when $d_X(x, y) \le \delta$, then $d_Y(f(x), f(y)) < \varepsilon$.

These conditions get weaker as we go down, and continuity can be generalized to non-metric spaces which gives rise to general topology. The study of uniformly continuous maps is much less common, but gives rise to uniformities.

Definition 1.4 (Convergence). If $\{x_n\}$ is a sequence in a metric space (X, d), then $x_n \to x$ if $\forall \varepsilon > 0$, $\exists N$ such that $n > N$ implies that $d(x_n, x)$.

Definition 1.5 (Cauchy sequence). If $\{x_n\}$ is a sequence in a metric space (X, d), then $x_n \to x$ if $\forall \varepsilon > 0$, $\exists N$ such that $n, m > N$ implies that $d(x_n, x_m)$.

[1] The ∞-norm can be viewed as the limit of the p-norms as $p \to \infty$.

Definition 1.6 (Completeness). (X,d) is complete if every cauchy sequence converges to a point.

For any metric space that isn't complete, we want to form a "completion" of it.

Definition 1.7 (Density). If $S \subseteq X$, S is dense if for all x in X and all $\varepsilon < 0$ there exists $s \in S$ where $d(s,x) < \varepsilon$.

Definition 1.8 (Completion). A completion of a metric space (X,d) is a metric space $(\overline{X},\overline{d})$ with an inclusion isometry $i\colon X \to \overline{X}$ such that the image of i is dense in $\overline{X}$. In addition, $\overline{X}$ is a complete metric space.

One way to define a completion for any metric space is by defining a semimetric $\tilde{d}$ on the set of Cauchy sequences $\mathrm{cs}(X,d)$, where $\tilde{d}\big(\{x_n\},\{y_n\}\big) = \lim_{n\to\infty} d(x_n,y_n)$, then by dropping that to a metric on the equivalence classes induced by the semimetric.

We can just define $f\colon X \to \overline{X}$ as $f(x) = [\{x,x,x,\dots\}]$. To show that $\overline{X}$ is complete, we consider Cauchy sequences of Cauchy sequences, approximate their elements by elements in X, and show that the sequence converges to the sequence generated by the approximation.

Example 1.5. If $f \in V = C^0([0,1])$, the norm $\|f\|_\infty = \sup(|f(x)|)$ induces a complete metric space (due to the limit of uniformly continous functions converging to a continuous function).

However, in the L^1 norm $\|f\|_1 = \int_0^1 |f(x)|\,\mathrm{d}x$, V is not complete (you can construct a sequence of functions that converges to a step function).

Now to answer the question "What is the completion of V under the L^p norms", we must look to measure theory (2nd half of this course).

If (X,d_X) and (Y,d_Y) are metric spaces, if $f\colon X \to Y$, recall that if $x \in X$, for all $\varepsilon > 0$ there exists $\delta > 0$ such that when $d_X(x,y) \leq \delta$, then $d_Y(f(x),f(y)) < \varepsilon$, then f is continuous.

Definition 1.9 (Open ball). The open ball of radius r, $B_r(x) = \big\{y \in X \mid d_X(x,y) < r\big\}$.

This allows us to refine our definition of continuity:

Definition 1.10 (Continuity). If $f\colon X \to Y$, then f is continuous if for all x, for all $\varepsilon > 0$, there exists $\delta > 0$ such that $f(B_\delta(x)) \subseteq B_\varepsilon(f(x))$.

Definition 1.11 (Open set). The set $A \subseteq X$ is open if for all $a \in A$, there exists an open ball around a that is contained within a.

Proposition 1.1. Open balls are open.

Proof. If $a \in B_r(x)$, then $B_{r-d_X(a,x)}(a)$ is contained within $B_r(x)$ due to the triangle inequality. $\square$

Proposition 1.2. If f is continuous, then if U is open in Y such that $x \in X$, then there is an open ball around x that is contained within the preimage of U.

A direct corollary of this proposition is

Corollary 1.3. $f\colon X \to Y$ is continuous is equivalent to the following condition: If U is open in Y, then $f^{-1}(U)$ is open in X.

Now, we consider properties of the set $\mathcal{T}$ consisting of open sets in X:

Theorem 1.4.

1. $\varnothing, X \in \mathcal{T}$.
2. Arbitrary unions of sets in $\mathcal{T}$ are also in $\mathcal{T}$.
3. Finite intersections of sets in $\mathcal{T}$ are in $\mathcal{T}$.

Proof. Trivial. □

These properties give rise to the definition of a topology.

Definition 1.12 (Topology)**.** A topology on a set X is a collection $\mathcal{T} \subseteq 2^X$ such that

1. $\varnothing, X \in \mathcal{T}$.
2. Arbitrary unions of sets in $\mathcal{T}$ are also in $\mathcal{T}$.
3. Finite intersections of sets in $\mathcal{T}$ are in $\mathcal{T}$.

The sets in $\mathcal{T}$ are called open sets.

2 Topological Spaces

We recall the definition of a topology:

Definition 2.1 (Topology). A topology on a set X is a collection $\mathcal{T} \subseteq 2^X$ such that

1. $\varnothing, X \in \mathcal{T}$.
2. Arbitrary unions of sets in $\mathcal{T}$ are also in $\mathcal{T}$.
3. Finite intersections of sets in $\mathcal{T}$ are in $\mathcal{T}$.

The sets in $\mathcal{T}$ are called open sets.

We then extend the definition of continuity to general topologies.

Definition 2.2 (Continuity). If $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ are topological spaces, then f is continuous if for all $U \in \mathcal{T}_Y$, $f^{-1}(U) \in \mathcal{T}_X$.

Topologies on a set X form a poset under the subset relation.

Definition 2.3 (Closed sets). The closed sets in $(X, \mathcal{T})$ are precisely the complement of open sets in X.

Proposition 2.1 (Properties of closed sets).

1. $\varnothing, X$ are closed.
2. Arbitrary intersections of closed sets are also closed.
3. Finite unions of closed sets are closed.

Definition 2.4 (Closure). The closure $\overline{A}$ of an arbitrary set A is the smallest closed set containing it.

Definition 2.5 (Density). Given $A \subseteq B$, we say A is dense in B if $B \subseteq \overline{A}$. Equivalently, $\overline{A} = \overline{B}$.

2.1 Bases and Subbases

Proposition 2.2. If X is a set, then the arbitrary intersection of topologies is a topology on X.

Proof. These details are routinely verified. □

Corollary 2.3. Given any collection $\mathcal{S}$ of subsets of X, there is a smallest topology containing $\mathcal{S}$.

Proof. This topology is the intersection of all the topologies that contain $\mathcal{S}$. We say that $\mathcal{S}$ generates this topology. □

Definition 2.6 (Subbase). Let $\mathcal{T}$ be a topology on X, and let $\mathcal{S} \subseteq 2^X$, with $X = \bigcup \mathcal{S}$. We say that $\mathcal{S}$ is a subbase of $\mathcal{T}$ if $\mathcal{T}$ is the topology generated by $\mathcal{S}$.

Subbases can be useful for studying topologies in the same way that a basis is useful for studying vector spaces. For example, see Proposition 2.7.

We can explicitly characterize the topology generated by a set $\mathcal{S}$. We first state a lemma:

Lemma 2.4. If $\mathcal{I} \subseteq 2^X$ be closed under finite intersections and $\bigcup \mathcal{I} = X$, then the topology generated by $\mathcal{I}$ consists of the arbitrary unions of sets in $\mathcal{I}$ (with union over 0 sets being $\varnothing$).

Proof. The only part that is not immediately obvious is the closure under finite intersections. If $\mathcal{O}_1 \bigcup_A U_\alpha$ and $\mathcal{O}_2 = \bigcup_B V_\beta$, then

$$\mathcal{O}_1 \cap \mathcal{O}_2 = \bigcup_{A,B} U_\alpha \cap V_\beta.$$

□

Proposition 2.5. If $\mathcal{S} \subseteq 2^X$ whose union is X, then the topology $\mathcal{T}$ generated by $\mathcal{S}$ consist of the arbitrary unions of $\mathcal{I}^{\mathcal{S}}$ which are defined as the set of finite intersections of sets in $\mathcal{S}$.

Definition 2.7 (Base). A collection $\mathcal{B} \subseteq 2^X$ is called a base for a topology $\mathcal{T}$ on X if $\mathcal{T}$ consists of the arbitrary unions of elements in $\mathcal{B}$.

Example 2.1.
1. The collection of finite intersections $\mathcal{I}^{\mathcal{S}}$ of a subbase $\mathcal{S}$.
2. The set of open balls in a metric space.
3. The set of intervals of the form $(-\infty, a)$ and (b, ∞) in $\mathbb{R}$ under the standard topology.

Notice how the second and third examples are not closed under finite intersections.

Proposition 2.6. For any base $\mathcal{B}$, any finite intersection of sets in $\mathcal{B}$ must be a union of elements in $\mathcal{B}$.

Proposition 2.7. If $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ are topological spaces, and let $f\colon X \to Y$. If $\mathcal{S}_Y$ is a subbase of $\mathcal{T}_Y$, then f is continuous iff for all $U \in \mathcal{S}_Y$, we have $f^{-1}(U)$ is open.

Proof. First show that the preimages of sets in $\mathcal{I}_{\mathcal{S}_Y}$ are open, and then show that the preimages of all sets in $\mathcal{T}_Y$ are open. □

2.2 New Topologies from Old

Definition 2.8 (Initial Topologies). If X is a set, $(Y_\alpha, \mathcal{T}_\alpha)$ are topological spaces, and for each α, $f_\alpha \colon X \to Y_\alpha$. Then, the initial topology is the smallest topology on X that makes all f_α continuous.

This initial topology is just the topology generated from the preimages of open sets in Y_α under f_α. This gives rise to a few important methods of constructing new topological spaces from old ones.

Definition 2.9 (Relative topology). If $X \subseteq Y$ and $(Y, \mathcal{T}_Y)$ is a topological space, and if j is the inclusion map from X to Y, then the initial topology generated by j consists of the intersection of open sets with X, which is also called the relative (or subspace) topology.

Definition 2.10 (Finite product topology). If $(X_1, \mathcal{T}_1)$ and $(X_2, \mathcal{T}_2)$ are topological spaces then if $X = X_1 \times X_2$. We have a couple of "natural" maps, $X \xrightarrow{p_1} X_1$, $X \xrightarrow{p_2} X_2$. We see that $p_1^{-1}(U) = U \times X_2$, and $p_2^{-1}(V) = X_1 \times V$, so it is easy to see that the product topology has a basis consisting of sets of the form $U \times V$, where U is open in X_1 and V is open in X_2.

This trivially generalizes to any **FINITE** number of topological spaces and their **FINITE** cartesian product. However, when we generalize to infinite products, the following case occurs:

Definition 2.11 (Product topology). Suppose we have topological spaces $(X_\alpha, \mathcal{T}_\alpha)$ over a (possibly infinite) indexing set A, where $X = \prod X_\alpha$. Then, p_α is the natural projection $X \to X_\alpha$, and $p_\alpha^{-1}(U)$ is $\prod U_\beta$, where $U_\beta = X_\beta$ when $\beta \neq \alpha$ and $U_\alpha = U$. These sets form a subbase of this topology. Now, since topologies only allow finite intersections, then the product topology is generated by products $\prod_\alpha U_\alpha$, where only a finite number of U_α are allowed to be not equal to X_α.

Example 2.2. If $X_i = \{0, 1\}$ with the discrete topology, then $\prod_0^\infty X_i$ is compact.

Definition 2.12 (Weak topology). For a normed vector space $(V, \|\cdot\|)$. If V^* is the vector space of all continuous linear functionals. The initial topology generated by V^* is the weak topology on V.

Example 2.3. For example, if we have $V = \mathcal{C}^0([0,1])$, with

$$\phi_g \colon V \to \mathbb{R}, f \mapsto \int_0^1 f(x) g(x) \, \mathrm{d}x.$$

Then, the initial topology generated by ϕ_g defines "a" weak topology on V.

As opposed to a initial topology, we can define final topologies:

Definition 2.13 (Final topology). If $(X_\alpha, \mathcal{T}_\alpha)$ are topological spaces and $f_\alpha \colon X_\alpha \to Y$, then the final topology $\mathcal{T}$ is the strongest topology making all f_α continuous.

To construct this topology, we can consider $\left\{A \in 2^Y \mid f_\alpha^{-1}(A) \in \mathcal{T}_\alpha\right\}$, which is a topology on Y for each α. We can just take the intersection of these topologies to obtain the strongest topology that makes all the f_α continuous.

Example 2.4 (Quotient topology). Any function $f \colon X \to Y$ induces an equivalence relation with the equivalence classes being the preimages $f^{-1}(\{y\})$. These equivalence classes partition X. It turns out any equivalence relation $\sim$ induces equivalence classes that partition X, which are denoted $X/\sim$. If X has a topology $\mathcal{T}_X$, then the final topology from the natural mapping $f \colon X \to X/\sim$ is called the quotient topology on $X/\sim$.

We can construct a bunch of new, interesting topologies using quotients. First we define what it means for topological spaces to be "the same." I will also start dropping the topology and use individual letters X to denote the topological spaces $(X, \mathcal{T}_X)$, depending on context.

Definition 2.14 (Homeomorphism). A function $f \colon X \to Y$ is a homeomorphism if it is an invertible continuous mapping, whose inverse is also continuous. Two spaces where there exists a homeomorphism between them are called "homeomorphic."

Example 2.5 (Circle). If $I = [0,1]$ with the equivalence relation such that only $0 \sim 1$, then $X/\sim$ is homeomorphic to the 1-sphere (circle), S^1. One homeomorphism from $I \to S^1$ is the mapping $t \mapsto e^{2\pi i t}$.

Example 2.6 (Cut-off cylinder). If $X = [0,2] \times [0,1]$ is endowed with the relation $(0,r) = (2,r)$ then the quotient $X/\sim$ is homeomorphic to a cylinder. If we instead identified $(0,r) \sim (2,1-r)$ then $X/\sim$ will be homeomorphic to a Möbius strip.

Example 2.7 (Projective space). If X is the unit sphere in $\mathbb{R}^3$, then if we define the equivalence $v \sim -v$, then $X/\sim$ is the real projective space $\mathbb{R}\mathrm{P}^2$.

Example 2.8. It can be seen that homeomorphisms from an object from itself can form (quite complicated) groups. These groups are the automorphism groups $\mathrm{Aut}(X)$ of topological spaces. Recall that a group action on X is a homomorphism $\alpha\colon G \to \mathrm{Aut}(X)$. Examples of group actions on spaces include $\mathbb{Z}$ acting on $\mathbb{R}$ by translation, or $\mathbb{Z}_2$ acting on S^2 by the antipodal map.

However, recall that group actions induce equivalence classes (orbits) on the topological space and thus induces a quotient topology A/α. The quotient topology of the examples of the actions above give rise to S^1 and $\mathbb{R}\mathrm{P}^2$.

3 Functions into the Reals

The following are properties of topological spaces that can kind of categorize how "nice" they are.

Definition 3.1 (Hausdorff (T_2)). A space X is Hausdorff if for all $x \neq y$, there exist disjoint open sets U, V such that $x \in U$ and $y \in V$.

Definition 3.2 (Normal (T_4)). A space X is Hausdorff if for all pairs of disjoint closed subsets C_1, C_2, there exist disjoint open sets O_1, O_1 such that $C_1 \subseteq O_1$ and $C_2 \subseteq C_2$.

3.1 Urysohn's Lemma

It can easily be seen that all metric spaces are Hausdorff, but it is also the case that metric spaces are normal. How do we prove this? We first consider the following important theorem:

Theorem 3.1 (Urysohn's Lemma). If X is a normal space, then for any disjoint closed subsets C_0, C_1 there exists a continuous function $f: X \to [0,1]$ where $f(C_0) = 0$ and $f(C_1) = 1$.

The converse of Urysohn's Lemma also holds, we simply take $O_0 = f^{-1}([0,1/3))$ and $O_1 = f^{-1}((2/3,1])$. We prove that metric spaces satisfy this property:

Proposition 3.2. Metric spaces are normal.

Definition 3.3 (Distance from a set). For a metric space (X,d), we define $d_A(x)$ or $d(x,A)$ where $A \subseteq X$, as $\inf_{a \in A} d(x,a)$.

It is easily seen that this function is Lipschitz with constant ≤ 1.

Now, we can proceed with a proof of Proposition 3.2.

Proof. Let C_0, C_1 be disjoint closed subsets in (X,d), then define

$$f: X \to [0,1], x \mapsto \frac{d(x,C_0)}{d(x,C_0) + d(x,C_1)}.$$

It is clear that $f(C_0) = 0$ and $f(C_1) = 1$, and that f is continuous. □

Now, we continue with a proof of Urysohn's Lemma.

Lemma 3.3. If X is normal, then for any closed set C and open set O containing C, there is an open set U "between" C and O, that is, $C \subseteq U \subseteq \overline{U} \subseteq O$.

Proof. If $D = X \setminus O$, then D is closed and disjoint from C and then we can find disjoint open sets U, V such that U contains C and V contains D. In addition, the closure of C must be contained within $X \setminus V$ which is contained by U. $\square$

Proof of Urysohn's Lemma. Given X normal and disjoint closed sets C_0, C_1, set $O_1 = C_1'$, then apply the preceding lemma to get a set $O_{1/2}$ "between" C_0 and O_1. Applying the lemma again, we then get a two open sets $O_{1/4}, O_{3/4}$ s.t. $C_0 \subseteq O_{1/4} \subseteq \overline{O_{1/4}} \subseteq O_{1/2} \subseteq \overline{O_{1/2}} \subseteq O_{3/4} \subseteq \overline{O_{3/4}} \subseteq O_1$. We can repeat this process to choose $O_{1/8}, O_{3/8}, O_{5/8}, O_{7/8}$, and then repeat the process over and over again.

Now, we are able to define O_d, for all $d \in \Delta$ (rational numbers of the form $k/2^n$ where $k \leq 2^n$), such that if $r < s$, $\overline{O_r} \subseteq O_s$. In addition, for all r, $C_0 \subseteq O_r$. Define f on X by $f(x) = 1$ if $x \in C_1$, otherwise, we set $f(x) = \inf\{r \in \Delta \colon x \in O_r\}$. We can see that $f(C_0) = 0$ in this case, since r is dense in $[0,1]$.

Now, to prove continuity, we choose a subbase of $[0,1]$ consisting of sets of the form $(0,a]$ and $(b,1]$. We then claim that

$$f^{-1}([0,a)) = \bigcup_{r<a} O_r, \text{ and } f^{-1}((b,1]) = \bigcup_{s>b} (\overline{O_s})'$$

This can be seen to be true due to density of Δ, as you can always find a dyadic rational between $f(x)$ and either a or b. $\square$

3.2 Tietze Extension Theorem

If X is a set and (Y,d) is a metric space, then let $B(X,Y)$ be the set of bounded functions from X to Y. Then, we can define a metric $\tilde{d}$ on $B(X,Y)$ as $\tilde{d}(f,g) = \sup_x |f(x) - g(x)|$.

Proposition 3.4. If Y is complete, then $B(X,Y)$ is complete.

Proof. If f_n is a Cauchy sequence then $f_n(x)$ is a Cauchy sequence for all x, and we can define $f(x) = \lim f_n(x)$. $\square$

Proposition 3.5. If we assume that X is instead of a topological space, then the set of continuous bounded functions $BC(X,Y)$ is complete if Y is complete.

The following is an important result using Urysohn's Lemma.

Theorem 3.6 (Tietze Extension Theorem)**.** If X is a normal space and A is a closed subset of X, and if $f\colon A \to \mathbb{R}$ is continuous, then there exists a continuous extension $g\colon X \to \mathbb{R}$ such that $g|_A = f$. Furthermore, if f maps A to $[a,b]$, then g maps X to that same interval.

Note that this theorem uses special properties of $\mathbb{R}$. For example, there is no continuous extension of the identity function from the unit circle to itself to the closed unit ball. To prove this theorem, we first prove a lemma.

Lemma 3.7. If X is normal and A is a closed subset of X, if there is an h mapping A to $[0,r]$, then there is a $g\colon X \to [0,r/3]$ such that $0 \leq h - g|_A \leq 2r/3$.

Proof. If $B = \{x \in A \mid h(x) \le r/3\}$, and $C = \{x \in A \mid h(x) \ge 2r/3\}$. Then, we can find a function $g\colon X \to [0, r/3]$ such that $g|_B = 0$ and $g|_C = r/3$ by Urysohn's Lemma. □

Proof of Tietze Extension Theorem. We prove it for the special case $[0,1]$, since all finite intervals are homeomorphic to the unit interval, this will hold for all finite initervals. To start, we can find a function $g^1\colon [0,1/3]$ such that $0 \le f - g|_A^1 \le 2/3$.

We then apply the lemma repeatedly: We apply it to $f - g|_A^1$ to get a function $g^2\colon X \to [0, 1/3(2/3)]$ such that $0 \le f - g|_A^1 - g|_A^2 \le 4/9$.

Repeat with $f - g|_A^1 - g|_A^2$ to get g^3, and in general we repeat this process with $f - g|_A^1 - \cdots - g|_A^n\colon A \to [0, (\frac{2}{3})^n]$ to get a function $g^{n+1}\colon [0, \frac{1}{3}(\frac{2}{3})^n]$ such that $f - g|_A^1 - \cdots - g|_A^{n+1}\colon A \to [0, (2/3)^{n+1}]$.

Now we have $\|g^n\|_\infty \le \frac{1}{3}(\frac{2}{3})^{n-1}$, so $\sum_{k=1}^n g^k$ is a Cauchy sequence and thus converges to $g = \sum g^n$, and we get that $\|g\|_\infty \le 1$ by summing up the geometric series. We see that $g = f$ because

$$\left\|f - g|_A\right\| = \lim \left\|f - \sum_{k=1}^{n} g|_A^k\right\|.$$

□

For infinite intervals, we use the homeomorphism h from $\mathbb{R}$ to $(0,1)$, and we set $g = h \circ f\colon A \to (0,1)$. Now if we apply Tietze's extension theorem we get $\tilde{g}\colon X \to [0,1]$, and let $D = \tilde{g}^{-1}(0)$, then D is disjoint from A and closed. We can take $U = \tilde{g}^{-1}([0,1/4)) \cap A'$. By Urysohn's Lemma we can find a function that takes $k\colon X \to [0,1]$ where $k(D) = 1/8$ and $k(U') = 0$, and we may add $\tilde{g} + k$ to obtain a continuous function whose output is not 0. Repeat this process for the point 1 and we get a final function that maps into $(0,1)$.

3.3 Completion of Metric Spaces

Recall that if X is a topological space then $\mathrm{BC}(X, \mathbb{R})$ with the supremum norm is complete. This is an example of the following.

Definition 3.4 (Banach space). A complete normed vector space is called a Banach space.

We can use this fact to define a completion of a metric space Y.

Lemma 3.8. If (Y, d_Y) and (M, d_M) are metric spaces and M is complete, then if $f\colon Y \to M$ is an isometry, it can be seen that $\overline{f(Y)}$ is a completion of M.

Theorem 3.9. All metric spaces have a completion.

Proof. Given (Y, d), we let $f_y(y') = d(y, y')$. This is unbounded, but if we choose $y_0 \in Y$ be any base point, we can define $h_y = f_y - g$, where $g = f_{y_0}$. By the triangle inequality, h_y is bounded.

We now show that $y \mapsto h_y$ is an isometry, as $(h_{y_1} - h_{y_2})(y) = (f_{y_1} - f_{y_2})(y)$ so the norm is at most $d(y_1, y_2)$. In addition, if we plug in y_1, we get $R(f_{y_1} - f_{y_2})(y_1) = d(y_1, y_2)$, so $\left\|h_{y_1} - h_{y_2}\right\| = d(y_1, y_2)$. According to the above lemma, we have a completion of Y. □

4 Compactness

Definition 4.1 (Open Cover, Subcover). If X is a topological space, by an open cover we mean a collection $\mathcal{C} = \{U_\alpha\}$ of open sets such that $\bigcup \mathcal{C} = X$.

By a subcover of $\mathcal{C}$, we mean a collection $\mathcal{D} \subseteq \mathcal{C}$ where $\bigcup \mathcal{D} = X$.

Definition 4.2 (Compactness). X is compact if and only if every open cover has a finite subcover. If $A \subseteq X$, then A is compact if A is compact in its relative topology.

Proposition 4.1. *A is compact iff for any collection $\mathcal{E}$ of open sets such that $\bigcup \mathcal{E} \supseteq A$, there is a finite subset $\mathcal{S} \subseteq \mathcal{E}$ that also covers A, that is, $\bigcup \mathcal{S} \supseteq A$.*

Proposition 4.2. *Any closed subset of a compact space is compact.*

Proof. If $\mathcal{E}$ is covers $A \subseteq X$, then $\mathcal{E} \bigcup \{A'\}$ covers X, so we find a finite subcover and take out A'. □

Note that the converse is not true: Any subset of a space with the indiscrete topology is compact. However, when X is Hausdorff, then this holds:

Proposition 4.3. *If X is Hausdorff, then if $A \subseteq X$ is compact, then for all $x \notin A$, then there exist disjoint open sets U and V such that $A \in U$ and $x \in V$.*

Proof. If x is in the complement of A, for all $y \in A$ we can find disjoint open sets U_y, V_y with $x \in U_y$ and $y \in V_y$. Then V_y covers A, and we can find a finite subcover V_{y_k}, $1 \le k \le n$. Then, we have $\bigcap U_{y_k}$ is an open set that contains X and does not intersect $\bigcup V_{y_k}$, which is an open set containing x. □

We get two immediate corollaries:

Corollary 4.4. *Any compact subset of a Hausdorff space is closed.*

Corollary 4.5. *Any compact Hausdorff space is regular (T_3), i.e. for any closed set $A \subset X$, and any $x \notin A$, we have disjoint open sets U, V such that $A \subset U$ and $x \in V$.*

We can prove a slightly stronger statement:

Corollary 4.6. *Any compact Hausdorff space is normal (T_4).*

Proof. If A and B are disjoint closed sets in a compact Hausdorff space X, then for any $y \in B$ there are disjoint open sets U_y, V_y such that $A \subseteq V_y$ and $y \in U_y$, and we can find a finite subcover U_{y_k} of the open cover U_y of B, Now, we let $U = \bigcup U_{y_k}$ and $V = \bigcap V_{y_k}$ and we have U disjoint from V, where $B \subseteq U$ and $A \subseteq V$. □

Now, we examine maps on compact spaces.

Proposition 4.7. If X is compact and $f: X \to Y$, then $f(X)$ is compact.

Proof. WLOG assume that $Y = f(X)$. If $\mathcal{C}$ is an open cover of $f(Y)$, then if $\mathcal{C}'$ is the set of preimages of sets in $\mathcal{C}$, then $\mathcal{C}'$ is an open cover of X, which we can take a finite subcover $\mathcal{S}' = \{f^{-1}(O_1), \dots, f^{-1}(O_n)\}$, then we have that $\mathcal{S} = \{O_1, \dots, O_n\}$ is a finite subcover of Y. □

Corollary 4.8 (Extreme Value Theorem). If $f: \mathbb{R} \to \mathbb{R}$ is continuous on an interval $[a, b]$, then f has a maximum on $[a, b]$ and attains it on $[a, b]$.

Proposition 4.9. If $f: X \to Y$ for X compact and Y Hausdorff, then if f is bijective, then it is a homeomorphism.

Proof. Images of closed sets are closed due to Proposition 4.7 and Corollary 4.4. □

We can give an alternative characterization of compactness using closed sets.

Proposition 4.10. X is compact iff for any collection of closed sets has the finite intersection property, the intersection is nonempty.

4.1 Tychonoff's Theorem

We first state Tychonoff's theorem:

Theorem 4.11 (Tychonoff's Theorem). If $\{X_\alpha\}$ are a collection of compact spaces, then the space $\prod_\alpha X_\alpha$ is compact.

It is clear that this requires the Axiom of Choice, as $\prod_\alpha X_\alpha$ may be empty without it. It is also interesting that Tychonoff's Theorem implies AC, although Tychonoff's Theorem for Hausdorff compact spaces does not imply AC.

The proof of this theorem also requires AC, in the form of

Theorem 4.12 (Zorn's Lemma). If P is a poset, then if any chain C (totally ordered subset) in P has an upper bound in P, then P has a maximal element. (We call P inductively ordered).

Zorn's Lemma has wide applications across many fields, such as in the proof that every vector space has a basis, or that every proper ideal is contained in a maximal ideal.

Definition 4.3. For any set X, a filter on X is a collection $\mathcal{F}$ of nonempty subsets that is closed under intersections, and where all supersets of a set in the filter is also in the filter.

An ultrafilter is a maximal filter.

Proof of Tychonoff's Theorem. If $\alpha \in \mathcal{A}$ indexes a family of compact spaces X_α, then if $X = \prod X_\alpha$, then let $\mathcal{C}$ be a collection of closed subsets of X with the finite intersection property. If $W_\mathcal{C}$ is the collection of families $\mathcal{D}$ of subsets of X such that $\mathcal{C} \subseteq \mathcal{D}$ and such that $\mathcal{D}$ has FIP.

Now, we show that $W_\mathcal{C}$ is inductively ordered: If Φ is a chain in W_C, then if $\mathcal{D}_\Phi = \bigcup \Phi$, then let $A_1, \dots, A_n \in \mathcal{D}_\phi$, then $A_i \in \mathcal{D}_{\alpha_i}$, which we can just find the maximal index out of the α_i, which we call α_0. Then, $A_i \in \alpha_0$ for all i, which proves that $\bigcap A_i$ has nonempty intersection.

We can then use Zorn's Lemma to obtain a maximal element $\mathcal{D}^*$ in $W_\mathcal{C}$. Now, we can see that $\mathcal{D}^*$ is closed under finite intersections because it is maximal. In addition, we also have that if

$B \cap A \neq \varnothing$ for all $A \in \mathcal{D}^*$, then $B \in \mathcal{D}^*$, again due to maximality (we can use the previous property to prove this statement).

Now, for any $\alpha \in \mathcal{A}$, we have the continuous projection $\pi_\alpha \colon X \to X_\alpha$, now we claim that for any $\mathcal{D} \in W_{\mathcal{C}}$, the set of $\{\pi_\alpha(A) \mid A \in \mathcal{D}\}$ has FIP since if $x \in \bigcap A_j$, then $\pi_\alpha(x) \in \bigcap \pi_\alpha(A_j)$. We also trivially have that the closures $\overline{\pi_\alpha(A)}$ has the FIP. Now, applying to $\mathcal{D}^*$, we have that $\bigcap_{A \in \mathcal{D}^*} \overline{\pi_\alpha(A)}$ is nonempty since X_α is compact.

Now, applying AC, we can pick $x \in X$ where $\pi_\alpha(x) \in \bigcap_{A \in \mathcal{D}^*} \overline{\pi_\alpha(A)}$. Now, we prove that $x \in \overline{A}$ for all $A \in \mathcal{D}^*$, using the subbase of the product topology. If we let $U = \pi_\alpha^{-1}(U_\alpha)$, then $x \in U$ implies that $U \cap A \neq 0$ for all $A \in \mathcal{D}^*$, since we have $x_\alpha \in \overline{\pi_\alpha(A)}$ and U_α, then $\pi_\alpha(A) \cap U_\alpha$ must be nonempty otherwise $\overline{\pi_\alpha(A)}$ would not contain x_α. This implies that $\pi^{-1}(U_\alpha)$ has nonempty intersection with any set in $\mathcal{D}^*$, and thus belongs to $\mathcal{D}^*$.

Now, by the first property of $\mathcal{D}^*$, then we have that finite intersections $\bigcap \pi_\alpha^{-1}(U_{\alpha_i})$ are in $\mathcal{D}^*$, thus $\mathcal{D}^*$ contains a neighborhood base of $\mathcal{D}^*$. Now, if $x \in U$ open in X, then there exists an open set V in $\mathcal{D}^*$. Thus, for all $A \in \mathcal{D}^*$, we have that $x \in \overline{A}$. Finally, this results in the fact that x is contained in $\bigcap \mathcal{C}$. □

This theorem has applications in 202B, such as the Aloglu theorem (unit ball in weak-* topology is compact), and when considering locally compact spaces and β-compactifications.

Now, we consider the following interesting result

Theorem 4.13 (Kelley). Tychonoff's Theorem is equivalent to AC.

Proof. The reverse implication was proven above, and now if we assume Tychonoff's theorem, we need to show that $\prod X_\alpha$ is nonempty. We can take $\omega = \bigcup X_\alpha$ which is not in $\bigcup X_\alpha$ due to regularity. Now, if $Y_\alpha = X_\alpha \cup \{\omega\}$, we can take $\mathcal{T}_\alpha = \{Y_\alpha, \varnothing, T_\alpha, \{\omega\}\}$, which is compact since there are finitely many open sets. The product $Y = \prod Y_\alpha$ is compact due to Tychonoff's theorem. If $F_\alpha = \pi_\alpha^{-1}(X_\alpha)$, then F_α is closed, and has the FIP, since for finite intersections $\bigcap F_{\alpha_i}$, we can consider $x_{\alpha_i} = F_{\alpha_i}$, and $x_\alpha = \omega$ otherwise. Thus due to compactness we have $\bigcap F_\alpha = \prod X_\alpha \neq \varnothing$, which proves the equivalence. □

4.2 Totally Bounded Sets

Definition 4.4. We say a metric space (X, d) is **totally bounded** if for all $\varepsilon > 0$, there exists a finite set $\{x_i\}$ such that

$$X = \bigcup_{i=1}^{n} B_\varepsilon(x_i).$$

It is trivial that compactness implies totally bounded. We now prove a partial converse:

Theorem 4.14. Complete and totally bounded metric spaces are compact.

Proof. If X is totally bounded, then if $\{x_n\}$ is a sequence in X, we have for any $\varepsilon > 0$ that $X = \bigcup B_\varepsilon(x_i)$. Due to finiteness, for at least one x_i, we have infinitely many points in the sequence in $B_\varepsilon(x_i)$. Taking $\varepsilon = 1$, we have $A_1 = B_1(y_1)$ which has infinitely many points of x_n inside, which means we can choose $A_{1/2} = B_{1/2}(y_2) \cap A_1$ with infinitely many points of x_n inside, and we can repeat this process to obtain a sequence $A_1 \supseteq A_{1/2} \supseteq A_{1/4} \supseteq \cdots$ with successively decreasing diameters. By completeness, there exists an (unique) point in the intersection, and thus we have found an accumulation point. $\square$

Corollary 4.15 (Heine-Borel Theorem). A subset of Euclidean space is compact iff it is closed and bounded.

We can now study compact sets in $\mathrm{BC}(X,M)$ with the d_∞ metric, where X is a topological space and M is a complete metric space. These correspond precisely to the totally bounded sets.

Definition 4.5. If X is a topological space and (M,d) is a metric space, then a family of functions $\mathcal{F} \subseteq \mathrm{C}(X,M)$ is **equicontinuous at a point** x if for any $\varepsilon > 0$, there exists an open neighborhood $\mathcal{O}$ such that $d(f(y),f(x)) < \varepsilon$ whenever $y \in \mathcal{O}$.

We say $\mathcal{F}$ is **equicontinuous** if it is equicontinuous at all points in X.

Theorem 4.16 (Arzela-Ascoli). Let X and M be as above, and let X be compact. If $\mathcal{F} \subseteq \mathrm{BC}(X,M)$ with $\mathcal{F}$ pointwise totally bounded, and equicontinuous, then $\mathcal{F}$ is totally bounded.

Proof. If $\varepsilon > 0$ is given, for each x there is an open set O_x such that $y \in O_x$, then $d(f(y),f(x)) < \varepsilon/4$. By compactness, there are $x_1,\ldots,x_n$ such that

$$X = \bigcup_{i=1}^{n} O_{x_i}.$$

Now, for each j, the set $I_j = \left\{f(x_j) \mid f \in \mathcal{F}\right\}$ is totally bounded, so if we choose S_j to be a finite subset of I_j that is $\varepsilon/4$-dense in I_j. We can then take $S = \bigcup S_j$. Now, if we let $\Psi = \left\{\psi\colon \{1,\ldots,n\} \to S\right\}$, which is finite, then we can define $\mathcal{F}_\psi = \left\{f \in \mathcal{F} \mid \forall j.\ d\left(f\left(x_j\right),\psi(j)\right) < \varepsilon/4\right\}$. We thus have $\mathcal{F} = \bigcup \mathcal{F}_\psi$.

Now, we prove that $\operatorname{diam}\left(\mathcal{F}_\psi\right) < \varepsilon$. For all y, we have that $y \in O_{x_i}$ for some i. Thus, if $f,g \in \mathcal{F}_\psi$, then

$$d(f(y),g(y)) \leq d(f(y),f(x_i)) + d(g(y),g(x_i)) + d(f(x_i),\psi(x_i)) + d(g(x_i),\psi(x_i)) < 4(\varepsilon/4) = \varepsilon.$$

Then, $\mathcal{F}_\psi$ is contained within $B_\varepsilon(f_\psi)$ for any $f_\psi \in \mathcal{F}_\psi$, and so $X = \bigcup B_\varepsilon(f_\psi)$. $\square$

4.3 Locally Compact Spaces

Definition 4.6. A topological space X is **locally compact** if for all $x \in X$, there exists a open set $x \in O$ such that $\overline{O}$ is compact.

For now, we assume that X is Hausdorff for simplicity.

Proposition 4.17. If X is locally compact, then for any point and open set $x \in O$, there exists an open set $x \in U \subseteq \overline{U} \subseteq O$.

Proof. Let $O' = O \cap O_1$, where $\overline{O_1}$ is compact. Now, due to regularity of $\overline{O'}$ which is also compact, and disjoint open sets U and V such that $x \in U$ and $\partial O \subseteq V$. Then $\overline{U}$ does not contain any points in ∂O and thus $\overline{U} \subseteq O' \subseteq O$. Furthermore, $\overline{U}$ is compact due to being a closed subset of $\overline{O'}$. □

A stronger proposition is the following:

Proposition 4.18. If X is locally compact, then for any compact subset of an open set $C \subseteq O$, there exists an open set $C \subseteq U \subseteq \overline{U} \subseteq O$.

Proof. This just follows by taking points in C and applying the previous proposition, then using compactness. □

Definition 4.7. For $f \in \mathrm{C}(X, V)$ where V is a vector space and X is locally compact, the **support** of f, $\operatorname{supp}(f)$ is defined as the closure of the set $\{x \colon f(x) \neq 0\}$.

Definition 4.8. We say that f has **compact support** if $\operatorname{supp}(f)$ is compact. We denote the set of continuous compactly-supported functions by $\mathrm{C}_{\mathrm{c}}(X, V)$.

Proposition 4.19. For X locally compact and C compact, O open, then there is a function $f \in \mathrm{C}_{\mathrm{c}}(X, \mathbb{R})$ where $f(C) = 1$ and $f(O') = 0$.

Proof. Find U such that $C \subseteq U \subseteq \overline{U} \subseteq O$ and apply Urysohn's Lemma. □

Definition 4.9. $\mathrm{C}_{\mathrm{b}}(X, \mathbb{R})$ denotes the set of all bounded functions $X \to \mathbb{R}$. Using the L^∞ norm, we have that the closure of $\mathrm{C}_{\mathrm{c}}(X, \mathbb{R})$ is $\mathrm{C}_\infty(X, \mathbb{R})$, which are the set of functions that vanish at infinity.

4.4 A Quick C*-algebra Interlude

We start with the following proposition:

Proposition 4.20. If X is a compact Hausdorff space, then the only unital homomorphisms between $\mathrm{C}(X)$ and $\mathbb{K}$ are the evaluation homomorphisms $\varphi_x(f) = f(x)$.

If A is any $\mathbb{K}$-algebra, then by $\hat{A}$ we mean the set of unital homomorphisms $A \to \mathbb{K}$. If $a \in A$, we define $\hat{a}$ to be the function $\hat{a} \colon \hat{A} \to \mathbb{K}$ to be $\hat{a}(\varphi) = \varphi(a)$. We give $\hat{A}$ the weakest topology that makes these functions continuous, that is, the initial topology generated by the $\hat{a}$.

Proposition 4.21. For $A = \mathrm{C}(X)$, we have that $x \mapsto \varphi_x$ is a homeomorphism from X to $\hat{A}$.

If $K = \mathbb{C}$, then $A = \mathrm{C}(X)$ has an involution, where $f^*(x) = \overline{f(x)}$. It is obvious that $\|f^* f\|_\infty = \|f\|_\infty^2$. This seemingly innocent property actually characterizes $C(X)$ with the ∞-norm:

Theorem 4.22 (Little Gelfand-Naimark theorem, 1943)**.** If A is a commutative Banach algebra with involution. If $\|a^* a\| \|a\|^2$ for all $a \in A$, then A is isomorphic (isometric *-isomorphism) to $C(\hat{A})$.

Now if we consider X locally compact, then we let $A = \mathrm{C}_\infty(X)$. Similarly to the cases above, $\mathrm{C}_\infty(X) \cong \mathrm{C}_\infty(\hat{A})$ and $X \cong \hat{A}$.

Definition 4.10. A **Hilbert space** is a inner product space that is complete under its induced norm.

We write $\mathrm{B}(\mathcal{H})$ to be the set of all continuous operators on $\mathcal{H}$.

$\mathrm{B}(\mathcal{H})$ has an involution, the adjoint, and $\|T^*T\| = \|T\|^2$. However, it isn't commutative, but this motivates the general form of the Gelfand-Naimark theorem, where any space satisfying this property is isomorphic in a sense to $\mathrm{B}(\mathcal{H})$ for some Hilbert space $\mathcal{H}$.

Algebras that satisfy this property are called C^*-algebras.

Part II

Measure and Integration

5 Building Measures

Previously, we've considered the L^p norms on $C([0,1])$ and shown that it is not complete under these norms. In order to complete $C([0,1])$, we need to define the integral on characteristic functions χ_E for many sets $E \subseteq [0,1]$, as they can be approximated by trapezoids.

To do this, we want to be able measure the sizes of sets "consistently." By "consistent", we mean that it satisfies a few properties that are intuitively true. For example, if μ is a function that "measures" the sizes of sets, then if E and F are two disjoint sets, then intuitively, $\mu(E \sqcup F)$ should be $\mu(E) + \mu(F)$.

5.1 Families and Measures

Definition 5.1. If X is a set, a **ring** of subsets is a collection $R \subseteq 2^X$ such that

1. If $E, F \in R$, then $E \cup F \in R$.
2. If $E, F \in R$, then $E \setminus X \in R$.

This additionally implies that $\varnothing$ and $E \cap F = E \cup F \setminus (E \triangle F)$ is in R.

A **finitely additive measure** on R is a function $\mu \colon X \to [0, \infty]$ where $\mu(E \sqcup F) = \mu(E) + \mu(F)$.

It is obvious that for a finitely-additive measure we have that $\mu(\varnothing) = 0$.

However, rings and finitely-additive measures aren't good enough, because of the key word *finite*. If E_n is a sequence of disjoint intervals $[a, b)$, we define $\mu(E_n) = [a, b)$. We consider that

$$\sum_{n=1}^{\infty} \mu(E_n) < \infty$$

and define

$$F_n = \bigsqcup_{k=1}^{n} E_k,$$

we see that χ_{F_n} is a Cauchy sequence, which "wants to" converge to χ_E, where $E = \bigcup E_k$. Thus it makes sense to define

$$\mu(E) = \sum_{n=1}^{\infty} \mu(E_n),$$

and by doing so, we introduce countable additivity.

Definition 5.2. A **σ-ring** is a ring that is closed under countable unions.

A **σ-algebra** (or σ-field) is a σ-ring with the whole X in it.

Similarly to topologies, we can order σ-rings (or σ-algebras) by inclusion, and we have the following proposition:

Proposition 5.1. If $\{\mathcal{S}_\alpha\}$ is any (arbitrarily large) collection of σ-rings (σ-algebras) on X, then the intersection $\bigcap \mathcal{S}_\alpha$ is a σ-ring (resp. σ-algebra) on X.

As a corollary we have that

Corollary 5.2. Any collection $\mathcal{C} \in 2^X$ has a smallest σ-ring (or σ-algebra) containing it.

We call this σ-ring $\mathcal{S}(\mathcal{C})$, and we say that it is generated by $\mathcal{C}$.

Example 5.1. If X is any set, the set of finite and countable subsets of X form a σ-ring on X. This is also the σ-ring generated by singletons.

Definition 5.3. A (countably additive) **measure** on a σ-ring $\mathcal{S}$ is a function $\mu\colon \mathcal{S} \to [0,\infty]$ such that for all disjoint families of sets $E_1, E_2, \dots \in \mathcal{S}$, we have

$$\mu\left(\bigsqcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mu(E_n).$$

One may ask why we only consider countable additivity and not uncountable additivity. Intuitively, on the real line, the length (size) of any singleton $\{x\}$ should be 0, as it is smaller than any interval of positive length. However, the interval $[0,1]$ is an uncountable disjoint union of these singletons, so uncountable additivity would imply that the length of $[0,1]$ would be 0.

Some examples of measures are below:

Example 5.2 (Counting Measure). For a space X, we can let $\mathcal{S}$ be the set of all countable subsets of X, and we can define a measure on $\mathcal{S}$ by $\mu(E) = \#(E)$, which is ∞ if E has infinite cardinality.

Example 5.3. The counting measure is a special case of the measure defined as

$$\mu_f(E) = \sum_{x \in E} f(x)$$

where $f\colon X \to [0,\infty]$.

The reason why this is countably additive is due to the ability to switch the order of summation when it is absolutely convergent.

Example 5.4. Another special case is when X itself is countable and

$$\sum_{x \in X} f(x) = 1.$$

This gives rise to discrete probability measures and distributions, and f is in this case known as the probability mass function.

As an example, taking $X = \{0, 1, \dots\}$ and $f(n) = p(1-p)^{n-1}$ gives us the geometric distribution with parameter p.

5.2 Prerings and Premeasures

However, directly defining measures for more complicated σ-rings is very difficult. For instance, it would be very difficult to explicitly define a measure on the σ-ring (which happens to be a σ-

algebra) generated by the intervals $[a,b)$ (we will see why we choose one end to be open later) in $\mathbb{R}$.

Thus, we try to define measures on smaller objects and then extend them to the σ-rings generated by those objects, analogouly to how we define a function on a basis of a vector space and extend it to the entire vector space. These smaller objects are called prerings.

Definition 5.4. A **prering** (or semiring) on a set X is a nonempty collection of subsets $\mathscr{P} \subseteq 2^X$ satisfying the following properties:

1. If $E, F \in \mathscr{S}$, then $E \cap F \in \mathscr{S}$.

2. If $E, F \in \mathscr{S}$, then there exists disjoint sets $E_1, \dots, E_n \in \mathscr{S}$ such that

$$E \setminus F = \bigsqcup_{i=1}^{n} E_i.$$

Definition 5.5. A countably additive set function $\mu\colon \mathscr{P} \to [0,\infty]$ is called a **premeasure**.

Example 5.5. As above, the sets $[a,b)$ form a prering on the reals. The proof of this is simple but tedious.

The function $\mu([a,b)) = b - a$ is a premeasure. The proof of this is less simple and even more tedious.

Example 5.6. If $\alpha\colon \mathbb{R} \to \mathbb{R}$ is a nondecreasing function, that is left-continuous, then $\mu_\alpha([a,b)) = \alpha(b) - \alpha(a)$ is a premeasure. We see that the preceding example is a special case of this.

Another special case is when $\alpha(x) \to 0$ as $x \to -\infty$, and $\alpha(x) \to 1$ as $x \to \infty$. This gives rise to continuous probability distributions on $\mathbb{R}$.

Proposition 5.3. μ_α is an example of a premeasure.

Proof. To show that this is indeed a premeasure, we consider $[a_0, b_0) = \bigsqcup [a_n', b_n')$ and for all ε we set $\varepsilon_j = 2^{-n-1}\varepsilon$. Then, we pick $b' < b$ such that $\alpha(b') + \varepsilon/2 > \alpha(b)$. Now, for each j, choose $a_j' < a_j$ such that $\alpha(a_j') + \varepsilon_j > \alpha(a_j)$. Then, (a_j', b_j) form an open covering of $[a_0, b_0']$. We can then find a finite subcover $\mathscr{C}$. We want to reorder and renumber inductively the sets in this interval, to make the sets disjoint members of the prering. As a first step, we choose $S = (a_1', b_1) \in \mathscr{C}$ such that $a_0 \in S$, and we call it S_1. Now, there's an interval containing b_1, and we do a similar process until we run out of sets. Now,

$$\sum_{j=1}^{n} \alpha(b_j) - \alpha(a_j') = -\alpha(a_1) + \sum_{j=1}^{n-1} \alpha(b_j) - \alpha(a_{j+1}') + \alpha(b_n) \ge \alpha(b_n) - \alpha(a_1') > \alpha(b_0) - \alpha(a_0) - \varepsilon/2.$$

We also have that

$$\sum_{j=1}^{n} \alpha(b_j) - \alpha(a_j') < \sum_{j=1}^{n} \alpha(b_j) - \alpha(a_j) + \varepsilon_j \le \varepsilon/2 + \sum_{j=1}^{\infty} \alpha(b_j) - \alpha(a_j).$$

Combining these we get that

$$\mu_\alpha([a_0,b_0)) - \varepsilon/2 < \alpha(b_0) - \alpha(a_0) - \varepsilon/2 < \sum_{j=1}^{n} \alpha(b_j) - \alpha(a_j') < \varepsilon/2 + \sum_{j=1}^{\infty} = \varepsilon/2 + \sum_{j=1}^{n} \mu_\alpha([a_j, b_j)$$

The fact that the sum of the measures is greater than or equal to the measure of the union follows from taking $\varepsilon \to 0$. The fact that the sum of the measures is at most the measure of the union follows from considering the partial sums. □

5.3 Caratheodory's Extension Theorem

The main reason why we were interested in prerings and premeasures is because we can extend premeasures on prerings to measures on σ-rings.

The main star of the show is

Theorem 5.4 (Caratheodory's Extension Theorem)**.** If $\mathcal{P}$ is a prering and μ is a premeasure, then it extends to a measure μ on $\mathcal{S} = S(\mathcal{P})$.

A few lemmas first, before we can build up to the theorem:

Lemma 5.5. If E, $F_1, \ldots, F_n$ are all in $\mathcal{P}$, then $E \setminus \bigcup F_k$ is the disjoint union of sets in $\mathcal{P}$.

Proof. Slightly messy, but pretty straightforward. □

This leads to a few important corollaries:

Corollary 5.6. If $E \subseteq F$, then $\mu(E) \leq\subseteq \mu(F)$, i.e. μ is monotone.

Corollary 5.7. If $E = \bigcup F_k$, then $\mu(E) \leq \sum \mu(F_k)$, i.e. μ is (countably) subadditive.

Proof. Given a sequence F_k then we can "disjointize" it by defining

$$G_k = F_k \setminus \bigcup_{j=1}^{k-1} F_k.$$

Now we apply the preceding results. □

Definition 5.6. Let $\mathcal{F}$ be any family of subsets of X, then $\mathcal{H}(\mathcal{F})$ is the **hereditary**[1] **σ-ring** of sets that are countably covered by $\mathcal{F}$.

Definition 5.7. If $\mathcal{F}$ is a family and $\mu\colon \mathcal{F} \to [0,\infty]$, then define $\mu^*\colon \mathcal{H}(\mathcal{F}) \to [0,\infty]$ by

$$\mu^*(A) = \inf\left\{ \sum_{j=1}^{\infty} \mu(E_j) : E_j \in \mathcal{F}, A \subseteq \bigcup_{j=1}^{\infty} E_j \right\}.$$

For $\mathcal{F}$ a prering and μ a premeasure, then μ^* is called the **outer measure** induced by μ.

Proposition 5.8. μ^* is countably subadditive and monotone, i.e., μ^* is an **outer measure**.

[1] Heriditary just means if E is in the family, then all subsets of E are also in the family.

Proof. Monotonicity is trivial, but for subadditivity, suppose $B_1, \dots, B_k$ are in $\mathcal{H}(\mathcal{F})$, then for any ε we let $\varepsilon_j = 2^{-j-1}\varepsilon$. Now, for each B_j we have sets E_k^j such that $B_j \subseteq \bigcup_k E_k^j$, and $\mu^*(B_j) + \varepsilon_j \geq \sum_k \mu(B_k^j)$. Now, we have

$$\mu^*\left(\bigcup_j B_j\right) \leq \sum_j \sum_k \mu(E_k^j) \leq \sum_j \mu^*(B_j) + \varepsilon_j \leq \varepsilon + \sum_j \mu^*(B_j).$$

Since ε can be made arbitrarily small, we prove the desired inequality. □

However, while μ^* is subadditive, it turns out that it isn't necessarily countably additive. If we want to make it countably additive, we have to restrict the domain, which leads to the following clever definition:

Definition 5.8. If $\mathcal{H}$ is a hereditary σ-ring on X and ν is an outer measure on $\mathcal{H}$, then we say a set $E \in \mathcal{H}$ is **ν-measurable** iff it "splits" every $A \in \mathcal{H}$, that is, $\nu(A) = \nu(A \setminus E) + \nu(A \cap E)$.

We denote the ν-measurable subsets by $M(\nu)$.

Proposition 5.9 (Caratheodory). $M(\nu)$ is a σ-ring and $\nu|_{M(\nu)}$ is a measure.

We first prove a lemma:

Lemma 5.10. $M(\nu)$ is a ring and for any $A \in \mathcal{H}$, $E, F \in \mathcal{H}$ disjoint, we have $\nu(A \cap (E \cup F)) = \nu(A \cap E) + \nu(A \cap F)$.

Proof. We first prove that $M(\nu)$ is closed under finite unions. If E, F are measurable, from subadditivity we have $\nu(A \cap (E \cup F)) + \nu(A \setminus (E \cup F)) \geq \nu(A)$. However, we also have that

$$\begin{aligned}
\nu(A \cap (E \cup F)) + \nu(A \setminus (E \cup F)) &= \nu((A \cap E) \cup (A \setminus E) \cap F) + \nu((A \setminus E) \setminus F) \\
&\leq \nu(A \cap E) + \nu((A \setminus E) \cap F) + \nu((A \setminus E) \setminus F) \\
&= \nu(A \cap E) + \nu(A \setminus E) \qquad \text{(because } F \text{ splits } A \setminus E\text{)} \\
&= \nu(A)
\end{aligned}$$

thus $\nu(A \cap (E \cup F)) + \nu(A \setminus (E \cup F)) = \nu(A)$.

Now, we prove that $M(\nu)$ is closed under set difference. If E, F are measurable, one direction of the inequality follows from subadditivity. Now, we have that

$$\begin{aligned}
\nu(A \cap (E \setminus F)) + \nu(A \setminus (E \setminus F)) &= \nu((A \cap E) \setminus F) + \nu((A \setminus E) \cup (A \cap (E \cap F))) \\
&\leq \nu((A \cap E) \setminus F) + \nu(A \setminus E) + \nu((A \cap E) \cap F) \\
&= \nu(A \cap E) + \nu(A \setminus E) \\
&= \nu(A).
\end{aligned}$$

Now, we prove the last property. Since E splits $A \cap (E \cup F)$, then we have $\nu(A \cap (E \cup F) \cap E) + \nu((A \cap (E \cup F)) \setminus E) = \nu(A \cap E) + \nu(A \cap F) = \nu(A \cap (E \cup F))$. □

We have a couple of corollaries:

Corollary 5.11. $\nu|_{M(\nu)}$ is finitely additive.

Corollary 5.12. If E_k are a disjoint finite sequence of measurable sets, then $\nu 1.(A \cap \bigsqcup E_k) = \sum \nu(\cap E_k)$

Now we prove that $M(\nu)$ is a σ-ring and that ν is countably additive on $M(\nu)$.

Proof. If E_j is a sequence of measurable sets, then we can disjointize the E_j by setting

$$F_k = E_k \setminus \bigcup_{j=1}^{k-1} E_j.$$

Now, each F_k is measurable due to $M(\nu)$ being a ring, and now we prove that $G = \bigsqcup F_j$ is measurable. Now, for all n,

$$\begin{aligned}\nu(A) &= \nu\left(A \cap \bigsqcup_{j=1}^{n} F_k\right) + \nu\left(A \setminus \bigsqcup_{j=1}^{n} F_k\right) \\ &\geq \nu(A \setminus G) + \sum_{j=1}^{n} \nu(A \cap E_j).\end{aligned}$$

and thus as a corollary,

$$\nu(A) \geq \nu(A \setminus G) + \sum_{j=1}^{\infty} \nu(A \cap E_j) \geq \nu(A \setminus G) + \nu(A \cap G).$$

Now, to prove countable additivity, just take $A = G$ in the previous proof. □

We have now proved Caratheodory's Extension Theorem in full, as if ν is the outer measure induced by the premeasure μ, it is easy to see that ν is an extension of μ to a σ-ring containing $\mathcal{S}(\mathcal{P})$.

An important case of this extension theorem is the following:

Example 5.7. Applying this extension to the premeasures μ_α as in Proposition 5.6, we get the **Lebesgue-Stiltjes measures**. When α is the identity, we get the **Lebesgue measure**.

6 Properties of Measures

6.1 Properties of Extensions

Now, we prove a few properties of this constructed measure.

Definition 6.1. If $\mathcal{F}$ is a heriditary family of subsets of X, and $\nu\colon \mathcal{F} \to [0,\infty]$, we say ν is **complete** if when $\nu(E) = 0$, for all $F \subseteq E$, we have $F \in \mathcal{F}$, and $\nu(F) = 0$.

Proposition 6.1. ν is complete.

Proof. If $\nu(E) = 0$ and $B \subseteq E$, then for any set A, $\nu(A \cap B) = 0$, and $\nu(A \setminus B)$ must be equal to $\nu(A)$ due to monotonicity and subadditivity. □

Note that a similar proof also implies that any set with outer measure 0 is measurable with measure 0.

We may assume extensions are unique, but this does not hold:

Proposition 6.2. Extensions are not necessarily unique.

Proof. Consider $\mathbb{R}$ with the prering $\mathcal{P}$ consisting of all countable sets, and define $\mu(P) = 0$ for all $P \in \mathcal{P}$. Then when extending this to $S(\mathcal{P})$, we can define the measure to be the zero measure or the measure that attains infinity on uncountable sets. □

Another example is given by:

Example 6.1. Consider $\mathcal{P}$ to be the prering of half open intervals $[a,b)$ and define μ to be infinite on all of these.

We notice that these all happen due to the extensions attaining infinite values, which motivates the following definition:

Definition 6.2. If $(X,\mathcal{P},\mu)$ is a premeasure space, then we say it is **σ-finite** if for all $E \in \mathcal{P}$ there exists a countable sequence of E_j in $\mathcal{P}$ such that $E = \bigcup E_j$ and $\mu(E_j) < \infty$.

Theorem 6.3. If $(X,\mathcal{P},\mu)$ is σ-finite then for all σ-rings $\mathcal{S} \subseteq M(\mu^*)$ containing $\mathcal{P}$, then $\mu^*|_{\mathcal{S}}$ is the unique measure on $\mathcal{S}$ extending μ.

Proof. If ν is any extension of μ to $\mathcal{S}$, if $G \in \mathcal{S}$ and $G \subseteq E \in \mathcal{P}$

If G is not contained in any E, then by the σ-finite condition, we can find a sequence of E_j covering G. □

However, there is still a characterization of the measure induced by the outer measure even in the non σ-finite case.

Proposition 6.4. If $(X, \mathscr{P}, \mu)$ is a premeasure, then for all σ-rings $\mathscr{S} \subseteq M(\mu^*)$ containing $\mathscr{P}$, then $\mu^*|_{\mathscr{S}}$ is the largest measure on $\mathscr{S}$ extending μ.

6.2 Continuity Properties of Measures

We have the following "continuity properties" of measures:

Proposition 6.5. If E_j is an increasing sequence of sets, then $\mu(E_j) \nearrow \mu(\bigcup E_j)$.

Proof. Disjointize the E_j by setting $F_j = E_j \setminus E_{j-1}$. Then we have $\mu(E_j) = \sum \mu(F_j)$, the latter of which is nondecreasing and converges upward to $\mu(\bigcup E_j)$. □

Proposition 6.6. Similarly, if E_j is a decreasing sequence of sets, then if the E_j have finite measure, then $\mu(E_j) \searrow \mu(\bigcap E_j)$.

Proof. Let $F_j = E_1 \setminus E_j$, then these sets are increasing and their union is $E_1 \setminus E := \bigcap E_j$. Thus we have that $\mu(E_1) - \mu(E_j) \to \mu(E_1) - \mu(E)$, implying that $\mu(E_j) \searrow \mu(E)$. □

6.3 Translation Invariance of Measures

The Lebesgue measure satisfies a property that we would intuitively expect.

Proposition 6.7. The Lebesgue measure is translation-invariant.

Proof. Problem Set. □

However, this property actually defines the Lebesgue measure:

Definition 6.3. We call the σ-ring generated by the sets of the form $[a, b)$ the **Borel σ-ring**.

Proposition 6.8. Every measure μ on the Borel σ-ring that is translation-invariant and finite on the sets $[a, b)$ is a positive scalar multiple of the Lebesgue measure.

This actually has a remarkable generalization.

Definition 6.4. If $(X, \mathscr{T})$ is a topological space, we call the σ-ring generated by the sets in $\mathscr{T}$ the **Borel σ-ring**.

However, when X is locally compact (Hasudorff), the **Borel σ-ring** is instead defined as the σ-ring generated by compact subsets of X.

Theorem 6.9 (Haar)**.** If G is a locally compact group, then there is a nonzero Borel measure that is finite on compact sets which is invariant under left-translation.

Moreover, any left-translation invariant Borel measure that is finite on compact sets is a non-negative scalar multiple of the preceding measure.

Theorem 6.10 (Weil)**.** If G is a group with σ-ring $\mathscr{S}$ and nonzero measure μ, under certain hypothesis, if $\mathscr{S}$ and μ are left-translation invariant, then we can construct a topology on G where it is a locally compact group and μ is its Haar measure.

On the problem set, we do a very special case where we construct the Haar measure on $\mathbb{R}/\mathbb{Z}$.

6.4 Nonmeasurable Sets

While the Borel σ-ring seems quite large, it is possible to find sets (using the Axiom of Choice) that are not Lebesgue-measurable.

Example 6.2 (Vitali Counterexample)**.** Let $\sim$ be an equivalence relation on $[0,1)$ with $x \sim y$ iff $y - x \in \mathbb{Q}$. Now, take A to be the set containing a representative of each equivalence class.

Proposition 6.11. A is not measurable.

Proof. Suppose A was measurable. Then, $\mu(A+r) = \mu(A)$ for all $r \in \mathbb{R}$, due to translation invariance of the Lebesgue measure. We clearly have that $[0,1)$ is the disjoint union of

$$((A \cap [0,r)) + r) \sqcup ((A \cap (r,1)) - 1 + r)$$

for $r \in \mathbb{Q} \cap [0,1)$.

Now, if $\mu(A) = 0$, that would imply that $\mu([0,1)) = 0$, and if $\mu(A) > 0$, that would imply that $\mu([0,1)) = \infty$. Thus A cannot be measurable. $\square$

This counterexample can also be constructed on the circle.

7 Integration of Functions

Now that we are able to measure the sizes of sets, we aim to define the integral of functions $f\colon (X,\mathcal{S},\mu) \to \mathcal{B}$, where $\mathcal{B}$ is a Banach space. The most important example will be when $\mathcal{B} = \mathbb{R}$. In this section, X will denote a measure space with σ-ring $\mathcal{S}$ and measure μ, and $\mathcal{B}$ will denote a Banach space over $\mathbb{K}$.

7.1 Integrating Simple Functions

From intuition, we want the integral of the indicator function of a set A to be the size (i.e. measure) of that set. However, when the sets are not measurable, this won't work, so we restrict ourselves to indicator functions of measurable sets.

Also, we want the integral to be linear. Through these two properties, we can define the integral of linear combinations of indicator functions of measurable sets.

Definition 7.1. We say $f\colon X \to \mathcal{B}$ is a **simple measurable function** if the image of f is finite and if $f^{-1}\big(\{y\}\big)$ is measurable for any $y \in \mathcal{B}$.

Lemma 7.1. Any simple measurable function can be written as

$$f = \sum_{y\in\operatorname{im} f} y\chi_{f^{-1}(\{y\})}.$$

Lemma 7.2. A function $f\colon X \to \mathcal{B}$ is simple measurable iff it can be written as

$$f = \sum_{i=1}^{n} y_i\chi_{E_i}$$

where $y_i \neq 0$ and are distinct, and the E_i are disjoint and nonempty.

Furthermore, if f can be written in such a form, then the y_i must range over $\operatorname{im}(f) \setminus \{0\}$, and the E_i must be the sets $f^{-1}(\{y_i\})$.

Proposition 7.3. The simple measurable functions from $X \to \mathcal{B}$ form a $\mathbb{K}$-vector space. In fact, if $\mathcal{B}$ is a Banach algebra, they form a $\mathbb{K}$-algebra.

Proof. Clearly kf is simple measurable since $\operatorname{im}(kf) = k\operatorname{im}(f)$, and $(kf)^{-1}(ky) = f^{-1}(y)$.

Thus, if f and g are simple measurable, then let

$$f = \sum_{i=1}^{m} a_i\chi_{A_i} \quad\text{and}\quad g = \sum_{i=1}^{n} b_i\chi_{B_i}.$$

as in the previous lemma.

For convinience, let $A_{m+1} = X \setminus \bigcup A_i$, $B_{n+1} = X \setminus \bigcup A_j$, and $a_{m+1} = 0 = b_{m+1}$. Then, let $C_{ij} = A_i \cap B_j$ and $c_{ij} = a_i + b_j$. We can write

$$f + g = \sum_{i=1}^{m+1} \sum_{j=1}^{n+1} c_{ij} \chi_{C_{ij}}.$$

However, the c_{ij} might be zero and the C_{ij} might be empty, so we remove those terms from the summation. In addition, if there are two c_{ij}s that are equal, we combine those terms in the summation. After doing these steps, we have a summation in the form required by the previous lemma, and thus we are done.

The proof for multiplication when $\mathscr{B}$ is an Banach algebra is similar. □

Corollary 7.4. The simple measurable functions are the span of the indicator functions of measurable sets.

Finally, we define the integral for simple measurable functions using the rules we stated at the beginning of the chapter.

Definition 7.2. For a simple measurable function $f\colon X \to \mathscr{B}$, we define the **integral** of f to be

$$\int f \, \mathrm{d}\mu = \sum_{y \in \operatorname{im}(f) \setminus \{0\}} \mu(f^{-1}(\{y\}))y.$$

Clearly $\int \chi_E \, \mathrm{d}\mu = \mu(E)$, and satisfies the two properties we proposed at the beginning of this section:

Proposition 7.5. $\int \chi_E \, \mathrm{d}\mu = \mu(E)$, and the integral is $\mathbb{K}$-linear.

7.2 Measurable Functions

Integrating simple functions isn't terribly interesting though, so we want to extend integration to functions that are non-simple. We can do this by considering the class of functions that are well approximable by SMFs, which turns out to be incredibly rich.

Definition 7.3. If $f\colon X \to \mathscr{B}$, (where $\mathscr{B}$ could be just a pointed metric space), we say f is $\mathcal{S}$-**measurable** if there exists a sequence $\{f_n\}$ of SMFs that converges pointwise to f.

If we have a measure μ on $\mathcal{S}$, the null sets form a σ-ring. In this situation, we say that f is μ-**measurable** if there exists a sequence $\{f_n\}$ that converges to f *μ-almost everywhere*, i.e., there exists a null set N such that $f_n \to f$ on $X \setminus N$.

To make the theory work smoothly, however, we need the following non-obvious fact:

Proposition 7.6. If $f_n \to f$, where the functions f_n are $\mathcal{S}$-measurable, then f is measurable.

We will prove this later, but first we try and isolate properties of measurable functions to make this easier to prove.

Proposition 7.7. If $\{f_n\}$ are SMFs, and $f_n \to f$ pointwise, then

$$\operatorname{im}(f) = \overline{\left(\bigcup_{i=1}^{\infty} \operatorname{im}(f_i)\right)},$$

and thus $\operatorname{im}(f)$ is separable, i.e. is contained in a closure of a countable set.

Proposition 7.8. The carrier of a measurable function f is contained in the union of the carriers of f_n, which are SMFs converging to f, and thus there is a set $E \in \mathcal{S}$ where carrier$(f) \subseteq E$.

Now we get the main property

Proposition 7.9. If f is measurable then if U is open in $\mathcal{B}$, we have $f^{-1}(U) \cap E \in \mathcal{S}$.

Proof. $x \in f^{-1}(U)$ iff $f(x) \in U$. Now, we define U_m to be the set $\{x \in U \mid d(x, U') > 1/m\}$. Then, we have $\overline{U_m} \subseteq U_{m-1}$. We can see that $f(x) \in U$ iff there exists n such that $f(x) \in U_n$. Now, if $f_k \to f$ is any sequence of functions converging to f pointwise, then using approximation and the triangle inequality, we see that $f(x) \in U$ if and only if there exists a n and K such for $k \geq K$ such that $f_k(x) \in U_n$.

Now we have

$$x \in f^{-1}(U) \cap E \iff x \in \bigcup_{n=1}^{\infty} \bigcup_{K=1}^{\infty} \bigcap_{k=K}^{\infty} f_k^{-1}(U_n) \cap E,$$

thus $f^{-1}(U) \cap E \in \mathcal{S}$ for f_k simple. □

Corollary 7.10. If f is measurable, for any Borel set $A \subseteq \mathcal{B}$, then $f^{-1}(A) \in \mathcal{S}$. **This is the traditional definition of measurable**.

Corollary 7.11. If f_n is a sequence of measurable functions converging to f then f satisfies

1. $\operatorname{im}(f)$ is separable.
2. There is a set $E \in \mathcal{S}$ where carrier$(f) \subseteq E$.
3. If f_n is a sequence of measurable functions and $f_n \to f$, then $f^{-1}(A) \cap E \in \mathcal{S}$ for all Borel sets $A \subseteq \mathcal{B}$.

Now, to prove that limits of measurable functions are measurable, we just need to prove that functions that satisfy 1, 2, and 3 are measurable.

Theorem 7.12. If $f : X \to \mathcal{B}$ has properties 1, 2, and 3, then it is measurable. That is, a function f is a $\mathcal{S}$-measurable function if and only if

1. $\operatorname{im}(f)$ is separable.
2. There is a set $E \in \mathcal{S}$ where carrier$(f) \subseteq E$.
3. If f_n is a sequence of measurable functions and $f_n \to f$, then $f^{-1}(A) \cap E \in \mathcal{S}$ for all Borel sets $A \subseteq \mathcal{B}$.

Proof. In this proof, we will work implicity with $X = E$ to avoid many intersections. Now, to construct a sequence of SMFs that converge to f, we enumerate a sequence $\{b_j\}$ that is dense in $\text{im}(f)$. Set $C_{ji} = B_{1/j}(b_i)$, whose preimage is in $\mathcal{S}$, and we order the pairs (j,i) lexicographically. Now, for fixed n, we construct f_n as follows: disjointize the C_{ji} for $j,i \le n$ in reverse of the order given. Thus, $E_{nn} = C_{nn}$, and $E_{ij} = C_{ji} \setminus \bigcup\{C_{\ell k} \mid ji < \ell k \le nn, i \le n\}$. Now we let $f_n = \sum_{ji \le nn, i \le n} b_i \chi_{E_{ji}}$.

Now we prove that $f_n \to f$ pointwise. If we fix $x \in X$ and $\varepsilon > 0$, there exists some point b_N where $\|f(x) - b_N\| < \varepsilon$, and thus x is contained within C_{NK} for $K = \lceil 1/\varepsilon \rceil$. Now, if we take $M = \max(N,K)$, then we can see for all $m \ge M$, $f_m(x)$ is either b_N, or some other b such that $\|f(x) - b\| < 1/M < \varepsilon$. □

Here are some standard facts about measurable functions:

Proposition 7.13.

1. The set of $\mathcal{S}$-measurable functions form a vector space.
2. If f is measurable, then $\|f\|$ is a measurable function from $X \to \mathbb{R}$.
3. If f is measurable and g is a real-valued measurable function, then gf is measurable.

7.3 Modes of Convergence

Now that we have developed the theory of measurable functions, we go back to integration.

We first define what it means for a simple function to be integrable:

Definition 7.4. A SMF f is **simple integrable** iff $\int \|f\| \,\mathrm{d}\mu < \infty$, which is equivalent to $\mu(E) < \infty$ for all $E = f^{-1}(\{y\})$ for $y \in \text{im}(f) \setminus \{0\}$.

Considering only SIFs, we see the following properties:

Proposition 7.14.

1. $f \mapsto \int f \,\mathrm{d}\mu$ is linear.
2. If f is a real-valued SIF where $f \ge 0$, then $\int f \,\mathrm{d}\mu \ge 0$.
3. The function
$$f \mapsto \int \|f\| \,\mathrm{d}\mu$$
is a seminorm.
4. The functions f such that $\int \|f\| \,\mathrm{d}\mu = 0$ are equal to 0 a.e.

We want to take the completion of this space and extend the integral operator.

However, we have bad news: A sequence $\{f_n\}$ of SIF that is Cauchy for $\|\cdot\|_1$ need not converge pointwise at any point.

Example 7.1. Let $X = [0,1]$ and μ be the Lebesgue measure, then we let $f_1 = \chi_{[0,1/2]}$, $f_2 = \chi_{[1/2,1]}$, and continue the sequence as so: $\chi_{[0,1/3]}, \chi_{[1/3,2/3]}, \chi_{[2/3,1]}, \chi_{[0,1/4]}$, and so on.

This sequence is Cauchy in mean (Cauchy in the L^1 norm), but does not converge pointwise, as it is not Cauchy anywhere.

Definition 7.5. If f_n is a sequence of measurable functions and f is measurable, then f_n **converges to f in measure** if for all $\varepsilon > 0$, we have $\mu\Big(\Big\{x \mid \left\|f_n(x) - f(x)\right\| \geq \varepsilon\Big\}\Big) \to 0$.

Definition 7.6. If f_n is a sequence of measurable functions then f_n is **Cauchy in measure** if for all $\varepsilon > 0$, we have $\sup_{n,m \geq N} \mu\Big(\Big\{x \mid \left\|f_n(x) - f_m(x)\right\| \geq \varepsilon\Big\}\Big)$ converges to 0 as $N \to \infty$.

Proposition 7.15. If a sequence if SIFs f_n is Cauchy in the L^1 (or L^p) norm, then it is Cauchy in measure.

Proof. For a given $\varepsilon > 0$, and $m, n \in \mathbb{N}$, then we let $E^{\varepsilon}_{mn} = \Big\{x \in X \mid \left\|f_m(x) - f_n(x)\right\| \geq \varepsilon\Big\}$. We note that $\varepsilon \chi_{E^{\varepsilon}_{mn}} \leq \left\|f_m(x) - f_n(x)\right\|$. Now we have $\int \left\|f_m(x) - f_n(x)\right\| \mathrm{d}\mu = \left\|f_n - f_m\right\| \geq \varepsilon \mu(E^{\varepsilon}_{mn})$. Thus, we have that $\mu(E^{\varepsilon}_{mn}) \to 0$ as $m, n \geq N$ and $N \to \infty$. □

Corollary 7.16 (Chebyshev's inequality)**.** If X is a random variable, then if $\mu := \mathbb{E}[X], \sigma^2 := \mathbb{E}[(X - \mu)^2]$, then $\mathbb{P}[|X - \mu| \geq \varepsilon] \leq \sigma^2/\varepsilon$.

We prove that convergence in measure is unique in a sense.

Proposition 7.17. If $f_n \to f, g$ where all functions are measurable, then $f = g$ almost everywhere.

Proof. $\|f(x) - g(x)\| \leq \left\|f(x) - f_n(x)\right\| + \left\|f_n(x) - g(x)\right\|$ by the triangle inequality. Then, we also have

$$\mu(\{x \mid \|f(x) - g(x)\| \geq \varepsilon\}) \leq \mu\Big(\Big\{x \mid \left\|f(x) - f_n(x)\right\| \geq \varepsilon/2\Big\}\Big) + \mu\Big(\Big\{x \mid \left\|f(x) - f_n(x)\right\| \geq \varepsilon/2\Big\}\Big)$$

where the RHS converges to 0 due to definition of measure. Thus we have that the LHS converges to 0 for all $\varepsilon > 0$. If we let $E^{\varepsilon} = \big\{x \mid \|f(x) - g(x)\|\big\} \geq \varepsilon$, then we see that the set $E = \big\{x \mid f(x) \neq g(x)\big\} = \bigcup_n E^{1/n}$. Due to countable additivity, $\mu(E) \leq \sum E^{1/n} = 0$. □

We also have some pretty standard properties:

Proposition 7.18. If $f_n \to f$ and $g_n \to g$ in measure, then $af_n + bg_n \to af + bg$ in measure. Furthermore, $x \mapsto \left\|f_n(x)\right\|$ converges to $x \mapsto \|f(x)\|$ in measure.

Definition 7.7. If f_n is any sequence of $\mathcal{B}$-valued functions, we say that $f_n \to f$ **almost uniformly** if for every $\varepsilon > 0$ there exists a set $E^{\varepsilon} \in \mathcal{S}$ such that $\mu(E^{\varepsilon}) < \varepsilon$, and $f_n \to f$ uniformly on $X \setminus E^{\varepsilon}$.

f_n is **almost uniformly Cauchy** if for all $\varepsilon > 0$ there exists a set $E^{\varepsilon} \in \mathcal{S}$ such that $\mu(E^{\varepsilon}) < \varepsilon$, and f_n is uniformly Cauchy on $X \setminus E^{\varepsilon}$.

Theorem 7.19 (Riesz-Weyl)**.** If f_n is a sequence of $\mathcal{S}$-measurable functions that is Cauchy in measure, then there is a subsequence of f_n that is almost uniformly Cauchy.

Proof. Consider the sequence n_k defined by $\mu(\Big\{x \mid \left|f_{n_k} - f\right| \geq 2^{-k}\Big\}) < 2^{-k}$. □

7.4 The General Integral

Proposition 7.20. If f_n is a sequence of measurable functions that converges to f almost uniformly, then it converges to f in measure.

Theorem 7.21. Let f be a measurable function, then the following are equivalent:

1. There is a sequence of SIFs $f_n \to f$ in measure, such that f_n is Cauchy in mean.
2. There is a sequence of SIFs $g_n \to f$ almost uniformly, such that g_n is Cauchy in mean.
3. There is a sequence of SIFs $h_n \to f$ almost everywhere, such that h_n is Cauchy in mean.

Proof. We prove this as a chain of implications:

(1) $\Rightarrow$ (2) By the Riesz-Weyl theorem, we have that some subsequence g_n of f_n that is almost uniformly Cauchy, which implies that it converges almost uniformly to some function g. Since by the previous proposition that implies it converges in measure to g, that must mean $g = f$ a.e. by Proposition 7.17, which implies that $g_n \to f$ almost uniformly.

(2) $\Rightarrow$ (3) Consider $A^{1/n}$ such that $f_n \to f$ uniformly on $A^{1/n}$ with $\mu(X \setminus A^{1/n}) < 1/n$. Then, $f_n \to f$ on $\bigcup A^{1/n}$.

(3) $\Rightarrow$ (1) Apply the Riesz-Weyl theorem to get a sequence g_n of h_n that is almost uniformly Cauchy, which implies that converges almost uniformly to f. Now, it is clear that for any δ, ε, we can choose A where $\mu(X \setminus A) < \varepsilon$, and $g_n \to f$ uniformly on A. In addition, we can choose N large enough that $\left\| g_n(x) - f(x) \right\| < \delta$ for all $x \in A$, which means that

$$\mu\Big(\Big\{x \mid \left\|g_n(x) - f(x)\right\| \geq \delta\Big\}\Big) < \varepsilon$$

for all $n \geq N$.

$\square$

Definition 7.8. We say f_n is **integrable** if one of these conditions holds. The set of integrable functions from $(X, \mathcal{S}, \mu) \to \mathcal{B}$ is denoted by $\mathcal{L}^1(X, \mathcal{S}, \mu, \mathcal{B})$.

We can note that $\mathcal{L}^1(X, \mathcal{S}, \mu, \mathcal{B})$ is a module over $\mathcal{L}^\infty(X, \mathcal{S}, \mu, \mathbb{K})$, which is the set of essentially bounded functions.

To properly complete the space of SIFs under the L^1 norm, we want some sort of uniqueness property of convergence.

Proposition 7.22. If the sequences f_n and g_n are Cauchy in mean and converge to f and g respectively, where $\left\|f_n - g_n\right\|_1 \to 0$, then $f = g$ a.e.

When we define the integral of a integrable function f, we want to approximate it with SIFs. The following proposition states that it does not matter which sequence of SIFs we wish to use to approximate f with.

Definition 7.9. If f is an integrable function, then we define the **integral** of f,

$$\int f \, \mathrm{d}\mu = \lim_{n \to \infty} \int f_n \, \mathrm{d}\mu,$$

where f_n is a mean-Cauchy sequence of SIFs converging in measure to f.

The following proposition implies that this value does not depend on the choice of sequence, and thus the definition above is a valid one.

Proposition 7.23. If f_n and g_n are mean Cauchy sequences of SIFs where both converge to f in measure. Then, they are equivalent, i.e. $\left\|f_n \to g_n\right\|_1 \to 0$.

Proof. We first prove that if $h_n \to 0$ in measure and is mean Cauchy, then it converges to 0 in mean. By the Riesz-Weyl theorem, there exists a subsequence j_n that is almost uniformly Cauchy. Let $E_n = \text{carrier}(j_n)$, and $E = \bigcup E_n$. Then, for N such that $\left\|j_n - j_m\right\| < \varepsilon/4$ when $n, m \geq N$, we have that

$$\int_{E \setminus E_N} \left\|j_n\right\| \mathrm{d}\mu = \left\|j_n - j_N\right\|_1 < \varepsilon/4$$

and we have that j_n almost uniformly converges to 0 on E_N, which has finite measure.

Now, we choose G such that $j_n \to 0$ uniformly on $E_N \setminus G$, where $\mu(G) < \varepsilon/(4 + 4\left\|g_N\right\|_\infty)$, and thus

$$\int_{E_N \setminus G} j_n \to 0.$$

Now, on G, we have

$$\int_G \left\|j_n\right\| \mathrm{d}\mu \leq \int_G \left\|j_n - j_N\right\| + \left\|j_N\right\| \mathrm{d}\mu \leq \varepsilon/4 + \mu(G)\left\|f_N\right\|_\infty < \varepsilon/2,$$

and thus picking $N' \geq N$ big enough where $\left\|j_n\right\| < \varepsilon/(4\mu(E_N \setminus G))$ for all $n \geq N'$, and we have that $\left\|j_n\right\|_1 < \varepsilon$ for all $n \geq N'$.

Then, it suffices to see that $f_n - g_n \to 0$ in measure. $\square$

8 Properties of the Integral

8.1 A Few Basic Facts

Clearly the integral is a linear function from $\mathcal{L}^1 \to \mathcal{B}$. If $\mathcal{B} = \mathbb{R}$, then the integral is additionally monotone. Thus, the function $\|f\|_1 = \int \|f\| \, d\mu$ is a seminorm on $\mathcal{L}^1$. If $f_n \to f$ (or is Cauchy) in this norm, we say $f_n \to f$ (or is Cauchy) in mean.

We now collect some basic properties about the Lebesgue integral here:

Proposition 8.1. The integral is linear.

Proposition 8.2. If $\mathcal{B} = \mathbb{R}$, then the integral is monotone.

Proof. Note that if $f_n \to f \geq 0$ converges to f in measure, then so does $\left|f_n\right|$. □

Proposition 8.3.

$$\left\| \int f \, d\mu \right\| \leq \int \|f\| \, d\mu$$

Proof. Clearly true for SIFs due to triangle inequality, extend by taking limit. □

8.2 Building Towards Completeness

Proposition 8.4. If $f \colon X \to \mathcal{B}$ is integrable, then f_n is a sequence of mean Cauchy SIFs such that $f_n \to f$ in measure, then $f_n \to f$ in mean.

Proof. Let N be such that if $n, m \geq N$ then $\left\|f_n - f_m\right\|_1 < \varepsilon/2$, then for all fixed $M \geq N$, we define $g_n^M = f_M - f_n$. Clearly, g_n^M is mean Cauchy and converges in measure to $f_M - f$.

Due to the reverse triangle inequality, the sequence $\left\|g_n^M\right\|$ is also mean Cauchy and converges to $\left\|f_M - f\right\|$ in measure, so

$$\left\|f_M - f\right\|_1 = \int \left\|f_M - f\right\| d\mu = \lim_{n\to\infty} \int \left\|g_n\right\| d\mu = \lim \left\|f_M - f_n\right\|_1.$$

Due to mean Cauchy-ness, the limit on the right hand side is at most $\varepsilon/2$ since for all $n \geq N$, $\left\|f_M - f_n\right\| < \varepsilon/2$, and thus for all $M \geq N$, we have $\left\|f_M - f\right\| < \varepsilon$, proving the proposition. □

Corollary 8.5. The space SIF is dense in the space $\mathcal{L}^1$.

Lemma 8.6. If $f \colon X \to \mathbb{R}$ is integrable, then for all ε there exists a set E such that $\mu(X \setminus E) < \infty$ and $\int_E f \, d\mu < \varepsilon$.

Proof. Take a SIF g such that $\|f - g\|_1 < \varepsilon$, then take E to be the complement of carrier(g). □

Lemma 8.7. If $f \geq \chi_E$ a.e. for some measurable E, then $\int f \,\mathrm{d}\mu \geq \mu(E)$.

Proof. We show that χ_E is integrable: Let $f_n \to f$ be a sequence of SIFs converging to f pointwise and f_n is Cauchy in mean. □

8.3 The Space L^1

One of our original goal for integration theory was to define the completion of $\mathrm{SIF}(X, \mathcal{S}, \mu, \mathcal{B})$. We say two functions $f, g \in \mathcal{L}^1(X, \mathcal{B})$ are equivalent iff $f = g$ a.e. Taking equivalence classes under this equivalence relation, we get the following definition:

Definition 8.1. $L^1(X, \mathcal{S}, \mu, \mathcal{B})$ denotes the set of integrable functions quotiented out by a.e.-equivalence.

From the previous section, one can deduce that there is a bijection between the Cauchy completion of $\mathrm{SIF}(X, \mathcal{B})$ and $L^1(X, \mathcal{B})$, and thus we have achieved our original goal.

Theorem 8.8. L^1 is isometric to the Cauchy completion of SIF quotiented by a.e.-equivalence.

Proof. To prove that the integral is well-defined, we proved that if f_n, g_n were equivalent mean Cauchy sequences (i.e. $\left\|f_n - g_n\right\|_1 \to 0$), then if $f_n \to f$ in measure and $g_n \to g$ in measure, then $f = g$ a.e., that is, they belong to the same one L^1 equivalence class. □

Corollary 8.9. L^1 is complete.

Proof. Completions of a metric space are complete themselves. □

8.4 The Indefinite Integral

Definition 8.2. If f is an integrable function then we denote by $\mu_f \colon \mathcal{S} \to \mathcal{B}$ the **indefinite integral** of f, defined by $\mu_f(E) = \int_E f \,\mathrm{d}\mu = \int f \chi_E \,\mathrm{d}\mu$.

Theorem 8.10 (Radon-Nikodym). Sufficiently nice $\mathcal{B}$-valued measures are the indefinite integral of some integrable f.

8.5 Interchange of Limit and Integral

We want to

Theorem 8.11 (Egorov's Theorem). If E is subset of X with finite measure $f_n \to f$ a.e. on E, then $f_n \to f$ almost uniformly on E.

Proof. I □

Theorem 8.12 (Dominated Convergence Theorem). If $f_n \to f$ a.e. and $\left\|f_n(x)\right\| < g(x)$ for some integrable $g \colon X \to \mathbb{R}$ and all x, then f_n is Cauchy in mean.

We specialize to the case where $\mathcal{B} = \mathbb{R}$.

Theorem 8.13 (Monotone Convergence Theorem).

8.6 L^p Spaces