1. Topology

Suppose $f : X \to Y$ is a function. Then f induces functions

$$f : P(X) \to P(Y) \qquad \text{by} \qquad A \mapsto f(A) := \{f(a) | a \in A\}$$
$$f^{-1} : P(Y) \to P(X) \qquad \text{by} \qquad B \mapsto f^{-1}(B) := \{x \in X | f(x) \in B\}$$

Exercise 1.0.1.

(1) Determine the relationship between $f^{-1}(f(A))$ and $A \subset X$. When are they equal?
(2) Determine the relationship between $f(f^{-1}(B))$ and $B \subset X$. When are they equal?
(3) Prove that $A \mapsto f(A)$ preserves unions, but not necessarily intersections or complements. Under what conditions on f does this preserve intersections? complements?
(4) Prove that $B \mapsto f^{-1}(B)$ preserves unions, intersections, and complements.

1.1. Topology basics.

Definition 1.1.1. A *topology* on a set X is a collection $\mathcal{T}$ of subsets of X such that:

- $\emptyset, X \in \mathcal{T}$,
- $\mathcal{T}$ is closed under arbitrary unions, and
- $\mathcal{T}$ is closed under finite intersections.

The elements of $\mathcal{T}$ are called *open sets*. An open set containing $x \in X$ is called a *neighborhood* of x. Complements of elements of $\mathcal{T}$ are called *closed sets*.

Definition 1.1.2. Observe that if $\mathcal{S}, \mathcal{T}$ are topologies on X, then so is $\mathcal{S} \cap \mathcal{T}$. This means if $\mathcal{E} \subset P(X)$, there is a *smallest* topology $\mathcal{T}(\mathcal{E})$ which contains $\mathcal{E}$ called the topology *generated by* $\mathcal{E}$.

Definition 1.1.3. Suppose $(X, \mathcal{T})$ is a topological space. A *neighborhood/local base* for $\mathcal{T}$ at $x \in X$ is a subset $\mathcal{B}(x) \subset \mathcal{T}$ consisting of neighborhoods of x such that

- for all $U \in \mathcal{T}$ such that $x \in U$, there is a $V \in \mathcal{B}(x)$ such that $V \subset U$.

A *base* for $\mathcal{T}$ is a subset $\mathcal{B} \subset \mathcal{T}$ which contains a neighborhood base for $\mathcal{T}$ at every point of X.

Example 1.1.4. Given a topological space $(X, \mathcal{T})$, the set $\mathcal{T}(x)$ of all open subsets which contain x is a neighborhood base at x.

Exercise 1.1.5. Show that $\mathcal{B} \subset \mathcal{T}$ is a base if and only if every $U \in \mathcal{T}$ is a union of members of $\mathcal{B}$.

Definition 1.1.6. Suppose $(X, \mathcal{T})$ is a topological space. We call $(X, \mathcal{T})$:

- *first countable* if there is a countable neighborhood base for $\mathcal{T}$ at every $x \in X$
- *second countable* if there is a countable base for $\mathcal{T}$.

Exercise 1.1.7. Show that second countable implies *separable*, i.e., there is a countable dense subset.

Exercise 1.1.8. Suppose X is first countable and $A \subset X$. Then $x \in \overline{A}$ (the smallest closed subset of X containing A) if and only if there is a sequence $(x_n) \subset A$ such that $x_n \to x$ (for every open subset U containing x, (x_n) is *eventually* in U).

Definition 1.1.9. Suppose X, Y are topological spaces. A function $f : X \to Y$ is called *continuous* at $x \in X$ if for every neighborhood V of $f(x)$, there is a neighborhood U of x such that $f(U) \subset V$. We call f *continuous* if f is continuous at x for all $x \in X$.

Exercise 1.1.10. Show that $f : X \to Y$ is continuous if and only if the preimage of every open set in Y is open in X, i.e., for every $V \in \mathcal{T}_Y$,

$$f^{-1}(V) := \{x \in X | f(x) \in V\} \in \mathcal{T}_X.$$

Exercise 1.1.11. Show that the composite of continuous functions is continuous.

Exercise 1.1.12. Prove the following assertions.

(1) Given $f : X \to Y$ and a topology $\mathcal{T}$ on Y, $\{f^{-1}(U) | U \in \mathcal{T}\}$ is a topology on X. Moreover it is the weakest topology on X such that f is continuous.

(2) Given $f : X \to Y$ and a topology $\mathcal{S}$ on X, $\{U \subset Y | f^{-1}(U) \in \mathcal{S}\}$ is a topology on Y. Moreover it is the strongest topology on Y such that f is continuous.

1.1.1. *Metric spaces.*

Definition 1.1.13. A *metric space* is a set X together with a *distance* function $d : X \times X \to [0, \infty)$ satisfying

- (definite) $d(x, y) = 0$ if and only if $x = y$,
- (symmetric) $d(x, y) = d(y, x)$ for all $x, y \in X$, and
- (triangle inequality) $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$.

The topology $\mathcal{T}_d$ *induced* by d is generated by the *open balls of radius* r

$$B_r(x) := \{y \in X | d(x, y) < r\} \qquad r > 0.$$

That is, U is open with respect to d if and only if for every $x \in U$, there is an $r > 0$ such that $B_r(x) \subset U$. Observe that every metric space is first countable.

Exercise 1.1.14. Let (X, d) be a metric space. Show tha $(X, \mathcal{T}_d)$ is second countable if and only if $(X, \mathcal{T}_d)$ is separable.

Exercise 1.1.15. Two metrics d_1, d_2 on X are called *equivalent* if there is a $C > 0$ such that

$$C^{-1} d_1(x, y) \leq d_2(x, y) \leq C d_1(x, y) \qquad \forall x, y \in X.$$

Show that equivalent metrics induce the same topology on X. That is, show that $U \subset X$ is open with respect to d_1 if and only if U is open with respect to d_2.

Exercise 1.1.16 (Sarason)**.** Let (X, d) be a metric space.

(1) Let $\alpha : [0, \infty) \to [0, \infty)$ be a continuous non-decreasing function satisfying
  - $\alpha(s) = 0$ if and only if $s = 0$, and
  - $\alpha(s + t) \leq \alpha(s) + \alpha(t)$ for all $s, t \geq 0$.

  Define $\sigma(x, y) := \alpha(d(x, y))$. Show that σ is a metric, and σ induces the same topology on X as d.

(2) Define $d_1, d_2 : X \times X \to [0, \infty)$ by

$$d_1(x, y) := \begin{cases} d(x, y) & \text{if } d(x, y) \leq 1 \\ 1 & \text{otherwise.} \end{cases}$$

$$d_2(x, y) := \frac{d(x, y)}{1 + d(x, y)}.$$

Use part (1) to show that d_1 and d_2 are metrics on X which induce the same topology on X as d.

Exercise 1.1.17. Suppose V is a $\mathbb{F}$-vector space for $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. A *norm* on V is a function $\|\cdot\| : V \to [0, \infty)$ such that

- (definite) $\|v\| = 0$ if and only if $v = 0$.
- (homogeneous) $\|\lambda \cdot v\| = |\lambda| \cdot \|v\|$ for all $\lambda \in \mathbb{F}$ and $v \in V$.
- (subadditive) $\|u + v\| \leq \|u\| + \|v\|$.

(1) Prove that $d(u, v) := \|u, v\|$ defines a metric on V.
(2) Prove that the following conditions are equivalent:
 (a) (V, d) is a complete metric space, i.e., every Cauchy sequence converges.
 (b) For every sequence $(v_n) \subset V$ with $\sum \|v_n\| < \infty$, the sequence $(\sum^k v_n)$ converges.

1.1.2. *Connectedness.*

Definition 1.1.18 (Relative topology)**.** Suppose X is a topological space and $A \subset X$ is a subset. The *relative topology* on A is given by $U \subset A$ is open if and only if there is an open set $V \subset X$ such that $U = V \cap A$.

Exercise 1.1.19. Suppose X is a topological space and $A \subset X$ is a subset. Show that $F \subset A$ is closed if and only if there is a closed set $G \subset X$ such that $F = G \cap A$.

Definition 1.1.20 ((Dis)connected set)**.** Let X be a topological space. We call a subset X *disconnected* if there exist non-empty, disjoint open sets U, V such that $X = U \amalg V$. A subset $A \subset X$ is disconnected if it is disconnected in its relative topology. If a subset is not disconnected, it is called *connected*. That is, $A \subset X$ is connected if and only if whenever $A \subset X$ can be written as the disjoint union $A = U \amalg V$ with U, V relatively open in A, then U or V is empty.

Exercise 1.1.21. Prove that the unit interval $[0, 1] \subset \mathbb{R}$ is connected.

Exercise 1.1.22.

(1) Suppose $f : X \to Y$ is continuous and $A \subset X$ is connected. Prove $f(A) \subset Y$ is connected.
(2) A subset $A \subset X$ is called *path connected* if for every $x, y \in A$, there is a continuous map $\gamma : [0, 1] \to A$ (called a *path*) such that $\gamma(0) = x$ and $\gamma(1) = y$. Prove that a path connected subset is connected.

Exercise 1.1.23. Recall that an *interval* $I \subset \mathbb{R}$ is a subset such that $a < b < c$ and $a, c \in I$ implies $b \in I$.

(1) Show that all intervals in $\mathbb{R}$ are connected.
(2) Prove that if $X \subset \mathbb{R}$ is not an interval, then X is not connected.

Exercise 1.1.24.

(1) Show that every open subset of $\mathbb{R}$ is a countable disjoint union of open intervals.
(2) Show that every open subset of $\mathbb{R}$ is a countable union of open intervals where both endpoints are rational.

1.1.3. *Separation axioms.*

Definition 1.1.25. We have the following separation properties for a topological space $(X, \mathcal{T})$.

(T_0) For every $x, y \in X$ distinct, there is an open set $U \in \mathcal{T}$ which contains exactly one of x, y.

(T_1) For every $x, y \in X$ distinct, there is an open set $U \in \mathcal{T}$ which only contains x. (Observe that by swapping x and y, there is also an open set $V \in \mathcal{T}$ which only contains y.)

(T_2) (a.k.a. Hausdorff) for every $x, y \in X$ distinct, there are disjoint open sets $U, V \in \mathcal{T}$ such that $x \in U$ and $y \in V$.

(T_3) (a.k.a. Regular) (T_1) and for every closed $F \subset X$ and $x \in F^c$, there are disjoint open sets $U, V \in \mathcal{T}$ such that $F \subset U$ and $x \in V$.

(T_4) (a.k.a. Normal) (T_1) and for every disjoint closed sets $F, G \subset X$, there are disjoint open sets $U, V \in \mathcal{T}$ such that $F \subset U$ and $G \subset V$.

Exercise 1.1.26. Let X be a set. The *finite complement topology* $\mathcal{T}$ has its opens those sets U such that U^c is finite and the empty set. Show $\mathcal{T}$ is (T_1). When is $\mathcal{T}$ Hausdorff?

Exercise 1.1.27. Suppose X is a normal topological space and $F \subset G \subset X$ with F closed and G open. Show there is an open U such that $F \subset U \subset \overline{U} \subset G$.

Lemma 1.1.28. *Suppose X is a normal topological space and $A, B \subset X$ are disjoint non-empty closed sets. Consider the dyadic rationals:*

$$D := \left\{ \frac{k}{2^n} \middle| n \in \mathbb{N},\ k = 1, \ldots, 2^n - 1 \right\} \subset (0, 1) \tag{1.1.29}$$

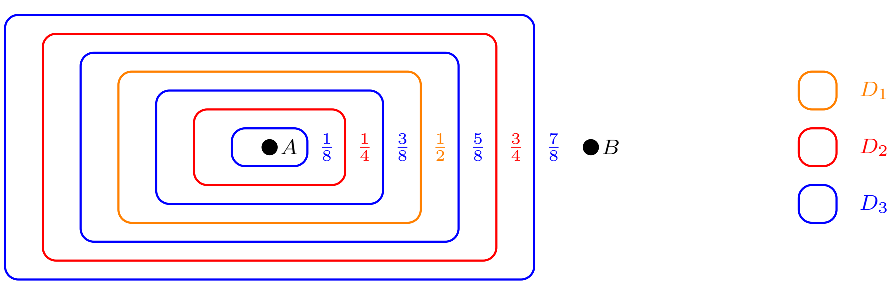
There are open sets $(U_d)_{d \in D}$ such that

- $A \subset U_d \subset \overline{U_d} \subset B^c$ *for all $d \in D$, and*
- $\overline{U_d} \subset U_{d'}$ *whenever $d < d'$.*

Proof. For $n \in \mathbb{N}$, set

$$D_n := \left\{ \frac{k}{2^n} \middle| k = 1, \ldots, 2^n - 1 \right\}.$$

We construct U_d for $d \in D_n$ inductively. Here is a cartoon of the main idea:



Base case: Let $U_{1/2}$ be any open set $A \subset U_{1/2} \subset \overline{U_{1/2}} \subset B^c$.

Inductive Step: Suppose that U_d have been defined for all $d \in D_1 \cup \ldots, \cup D_n$. Then, using the convention $U_0 := A$ and $U_1 := B^c$, we define $U_{\frac{2k+1}{2^{n+1}}}$ for $k = 0, 1, \ldots, 2^n - 1$ to be any open set such that

$$\overline{U_{k/2^n}} \subset U_{\frac{2k+1}{2^{n+1}}} \subset \overline{U_{\frac{2k+1}{2^{n+1}}}} \subset U_{\frac{k+1}{2^n}}. \qquad \square$$

Lemma 1.1.30 (Urysohn)**.** *Let X be a normal topological space. If $A, B \subset X$ are disjoint nonempty closed subsets, there is a continuous function $f : X \to [0,1]$ such that $f|_A = 0$ and $f|_B = 1$.*

Proof. For the dyadic rationals $D \subset (0,1)$ as in (1.1.29), we have open sets $(U_d)_{d\in D}$ satisfying the conditions in Lemma 1.1.28. Define $f : X \to [0,1]$ by $f(x) := \sup\{d | x \notin U_d\}$. It is clear by construction that $f|_A = 0$ and $f|_B = 1$. Also observe that

(D1) $f(x) > d$ implies that $x \notin \overline{U_d}$, and $f(x) < d'$, then $x \in U_{d'}$.
(D2) If $x \notin \overline{U_d}$, then $f(x) \geq d$, and if $x \in U_{d'}$, then $f(x) \leq d'$.

It remains to prove that f is continuous. Fix $x_0 \in X$ and $\varepsilon > 0$.

Case 1: Suppose $0 < f(x) < 1$. Choose $d, d' \in D$ such that $d < f(x_0) < d'$ and $d' - d < \varepsilon$. By (D1) above, $x_0 \in U_{d'} \setminus \overline{U_d}$. By (D2) above, $|f(x) - f(x_0)| < \varepsilon$ for all $x \in U_{d'} \setminus \overline{U_d}$.

Case 2: $f(x) = 0$ or 1. Similar to above and omitted. □

Theorem 1.1.31 (Tietze Extension)**.** *Suppose X is normal, $A \subset X$ is closed, and $f : A \to [a,b]$ is continuous. Then there is a continuous function $F : X \to [a,b]$ such that $F|_A = f$.*

Proof. Without loss of generality, $[a,b] = [0,1]$. (Otherwise, replace f with $(f-a)/(b-a)$.) We inductively construct a sequence of continuous functions (g_n) on X such that

- $0 \leq g_n \leq 2^{n-1}/3^n$ for all $n \in \mathbb{N}$, and
- $0 \leq f - \sum_{k=1}^n g_k \leq \left(\frac{2}{3}\right)^n$ on A for all $n \in \mathbb{N}$.

Then by (a), $\sum g_n$ converges uniformly to a continuous limit function F on X, and by (b), $F|_A = f$.

Base case: Set $B := f^{-1}([0,1/3]) \subset A$ and $C := f^{-1}([2/3,1]) \subset A$. Since f is continuous on A, $B, C \subset A \subset X$ are closed. By Urysohn's Lemma, there is a continuous function $g_1 : X \to [0,1/3]$ such that $g_1|_B = 0$ and $g_1|_C = 1/3$. Then

$$f - g_1 \leq \left\{\begin{array}{ll} \frac{1}{3} - 0 = \frac{1}{3} & \text{on } B \subset A \\ \frac{2}{3} - 0 = \frac{2}{3} & \text{on } A \setminus (B \cup C) \\ 1 - \frac{1}{3} = \frac{2}{3} & \text{on } C \subset A \end{array}\right\} \leq \frac{2}{3} \qquad \text{on } A.$$

Inductive Step: Suppose we have constructed $g_1, \ldots, g_{n-1}$. Then there is a continuous function $g_n : X \to [0, 2^{n-1}/3^n]$ such that $g_n = 0$ on $\{f - \sum_{k=1}^{n-1} g_k \leq 2^{n-1}/3^n\}$ and $g_n = 2^{n-1}/3^n$ on $\{f - \sum_{k=1}^{n-1} g_k \geq 2^n/3^n\}$. This implies that $f - \sum_{k=1}^n g_k \leq 2^n/3^n$ on A as in the base case. □

1.2. **Locally compact Hausdorff spaces.**

Definition 1.2.1. A topological space X is called *compact* if every open cover has a finite subcover.

Exercise 1.2.2. A collection of subsets of $(A_i)_{i\in I}$ of X has the *finite intersection property* if for any finite $J \subset I$, we have $\bigcap_{j\in J} A_j \neq \emptyset$. Prove that the following are equivalent.

(1) Every open cover of X has a finite subcover.
(2) For every collection of closed subsets $(F_i)_{i\in I}$ with the finite intersection property, $\bigcap_{i\in I} F_i \neq \emptyset$.

Fact 1.2.3. An interval in $\mathbb{R}$ is compact if and only if it is closed and bounded.

Exercise 1.2.4. In this exercise, you will prove that the half-open interval topology on $(0,1]$ is *Lindelöf*, i.e., every open cover has a countable sub-cover.

(1) Suppose $U \subset \mathbb{R}$ is open and suppose $((a_j, b_j))_{j\in J}$ is a collection of open intervals which cover U:
$$U \subset \bigcup_{j\in J}(a_j, b_j).$$
Show there is a countable sub-cover, i.e., show that there is a countable subset $I \subset J$ such that
$$U \subset \bigcup_{i\in I}(a_i, b_i).$$
Hint: Use Exercise 1.1.24.

(2) Suppose $((a_j, b_j])_{j\in J}$ is a collection of half-open intervals which cover $(0,1]$:
$$(0,1] \subset \bigcup_{j\in J}(a_j, b_j].$$
Show there is a countable sub-cover, i.e., show that there is a countable subset $I \subset J$ such that
$$(0,1] \subset \bigcup_{i\in I}(a_i, b_i].$$

Exercises 1.2.5. Suppose X is a topological space. Verify the following assertions.

(1) If X is compact and $F \subset X$ is closed, then F is compact.
(2) If X is Hausdorff, $K \subset X$ is compact, and $x \notin K$, then there are disjoint open U, V such that $x \in U$ and $K \subset V$. In particular, K is closed.
(3) If X is compact Hausdorff, then X is normal.
(4) If X is compact and $f : X \to Y$ is continuous, then $f(X)$ is compact.
(5) If X is compact and Y is Hausdorff, and continuous bijection $f : X \to Y$ is automatically a homeomorphism (i.e., f^{-1} is continuous).

Exercise 1.2.6 (Lebesgue Number Lemma)**.** Suppose (X, d) is a compact metric space. Prove that for every open cover $(U_i)_{i\in I}$, there is a $\delta > 0$ such that for every $x_0 \in X$, there is an $i_0 \in I$ such that $B_\delta(x_0) \subset U_{i_0}$.

Exercise 1.2.7. Consider the following conditions:

(1) For every $x \in X$, there is a neighborhood U of x such that $\overline{U}$ is compact.
(2) For every $x \in X$, there is a neighborhood base $\mathcal{B}(x)$ consisting of neighborhoods U of x such that $\overline{U}$ is compact.
(3) For every $x \in X$ and every neighborhood U of x, there is an open V with $x \in V \subset U$ with $\overline{V}$ compact.
(4) For every $x \in X$ and every neighborhood U of x, there is an open V with $x \in V \subset \overline{V} \subset U$ with $\overline{V}$ compact.

Determine which conditions imply which other conditions. Then show all the above conditions are equivalent when X is Hausdorff.

Definition 1.2.8. A Hausdorff space satisfying one (equivalently all) of the conditions in Exercise 1.2.7 is called a *locally compact Hausdorff* (LCH) space.

Exercise 1.2.9. Suppose X is a second countable LCH space. Prove the following assertions.

(1) X is *σ-compact*, i.e., there is a sequence (K_n) of compact subsets of X such that $X = \bigcup K_n$.
(2) Every compact $K \subset X$ is a *G_δ-set*, i.e., a countable intersection of open sets.

Exercise 1.2.10 (Baire Category)**.** Suppose X is either:

(1) a complete metric space, or
(2) an LCH space.

Suppose (U_n) is a sequence of open dense subsets of X. Prove that $\bigcap U_n$ is dense in X.
Hint: Let V_0 be an arbitrary non-empty open set. Inductively construct a decreasing sequence $(V_n)_{n\geq 1}$ of non-empty open subsets with $V_{n+1} \subset \overline{V_{n+1}} \subset U_{n+1} \cap V_n$ such that in the two cases above,

(1) V_n is a ball of radius $1/n$ for all $n \in \mathbb{N}$, or
(2) $\overline{V_n}$ is compact for all $n \in \mathbb{N}$.

Exercise 1.2.11. Suppose X is LCH. Verify the following assertions.

(1) If $K \subset U \subset X$ where K is compact and U is open, there is an open V with $K \subset V \subset \overline{V} \subset U$ with $\overline{V}$ compact.
Hint: Use Exercise 1.2.7(4).
(2) (Urysohn) If $K \subset U \subset X$ as above, there is a continuous $f : X \to [0,1]$ such that $f|_K = 1$ and $f = 0$ outside of a compact subset of U.
(3) (Tietze) Of $K \subset X$ is compact and $f \in C(K)$, there is an $F \in C_c(X)$ such that $F|_K = f$.

Definition 1.2.12. Let X be an LCH space. We define the following function algebras:

- $C(X)$ is the algebra of continuous ($\mathbb{C}$-valued) functions on X.
- $C_c(X)$ is the algebra of continuous functions of compact support, i.e., there is a compact set K such that $f|_{K^c} = 0$. We'll write $\operatorname{supp}(f) := \overline{\{x | f(x) \neq 0\}}$, so f has compact support if and only if $\operatorname{supp}(f)$ is compact.
- $C_0(X)$ is the algebra of continuous functions which *vanish at infinity*, i.e., for all $\varepsilon > 0$, $\{|f| \geq \varepsilon\}$ is compact.
- $C_b(X)$ is the algebra of continuous bounded functions.

We write $C(X, \mathbb{R}), C_c(X, \mathbb{R}), C_0(X, \mathbb{R}), C_b(X, \mathbb{R})$ for the real subalgebras of real-valued functions. Observe that

$$C_c(X) \subset C_0(X) \subset C_b(X) \subset C(X).$$

The *uniform/∞*-norm on $C_b(X)$ is given by

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

Exercise 1.2.13. Show that $C(X), C_c(X), C_0(X), C_b(X)$ are all complex algebras. Moreover, show $C_c(X), C_0(X)$ are unital if and only if X is compact.

Exercise 1.2.14 (Dini's Lemma)**.** Suppose X is a compact topological space and $(f_n) \subset C(X, [0,1])$. Show that if $f_n(x) \searrow 0$ pointwise, then $f_n \searrow 0$ uniformly.

Theorem 1.2.15. *Suppose X is LCH.*

(1) $\|\cdot\|_\infty$ is a norm on $C_b(X)$.

(2) $C_b(X)$ is complete with respect to $\|\cdot\|_\infty$.
(3) $C_0(X) \subset C_b(X)$ is closed (and thus complete).
(4) $\overline{C_c(X)}^{\|\cdot\|_\infty} = C_0(X)$.

Proof.
(1) Exercise.
(2) Suppose (f_n) is uniformly Cauchy. Then $(f_n(x))$ is Cauchy in $\mathbb{C}$ for every $x \in X$. Define $f(x) := \lim f_n(x)$, which is continuous (use $\varepsilon/3$ argument). Then one shows $\|f_n\|_\infty \subset [0, \infty)$ is bounded. Finally, you can show $f_n \to f$ uniformly, and $\sup|f(x)| \leq \sup \|f_n\| < \infty$.
(3) Suppose $(f_n) \subset C_0(X)$ such that $f_n \to f$ in $C_b(X)$. Let $\varepsilon > 0$. Pick $N \in \mathbb{N}$ such that $n \geq N$ implies $\|f - f_n\|_\infty < \varepsilon/2$. Since $f_N \in C_0(X)$, $\{|f_N| \geq \varepsilon/2\}$ is compact. Then $\{|f| \geq \varepsilon\} \subset \{|f_N| \geq \varepsilon/2\}$ is compact as a closed subset of a compact set.
(4) It suffices to prove that we can uniformly approximate any function in $C_0(X)$ by a function in $C_c(X)$. Let $f \in C_0(X)$ and $\varepsilon > 0$ so that $K := \{|f| \geq \varepsilon\}$ is compact. By the LCH Urysohn Lemma (Exercise 1.2.11(2)), there is a continuous function $g : X \to [0, 1]$ such that $g|_K = 1$ and g has compact support. Then $fg \in C_c(X)$, and $\|f - fg\|_\infty < \varepsilon$. □

Exercise 1.2.16. Suppose $(X, \mathcal{T})$ is a locally compact topological space and (f_n) is a sequence of continuous $\mathbb{C}$-valued functions on X. Show that the following are equivalent:

(1) There is a continuous function $f : X \to \mathbb{C}$ such that $f_n|_K \to f|_K$ uniformly on every compact $K \subset \mathbb{C}$.
(2) For every compact $K \subset X$, $(f_n|_K)$ is uniformly Cauchy.

Deduce that $C(X)$ is complete in the topology of local uniform convergence.

Exercise 1.2.17. Suppose X is a locally compact Hausdorff space, $K \subset X$ is compact, and $\{U_1, \ldots, U_n\}$ is an open cover of K. Prove that there are $g_i \in C_c(X, [0, 1])$ for $i = 1, \ldots, n$ such that $g_i = 0$ on U_i^c and $\sum_{i=1}^n g_i = 1$ everywhere on K.

1.3. **Convergence in topological spaces.** Let $(X, \mathcal{T})$ be a topological space. Recall that a sequence (x_n) converges to x, denoted $x_n \to x$ if for every open $U \in \mathcal{T}$ with $x \in U$, there is an $N \in \mathbb{N}$ such that $n > N$ implies $x_n \in U$ (x_n is *eventually* in U for every open neighborhood U of x). Not all spaces are first countable, so sequences do not suffice to describe the topology!

1.3.1. *Nets.*

Definition 1.3.1. A *directed set* is a set I equipped with a *preorder* (reflexive and transitive binary relation) $\leq$ satisfying

- for all $i, j \in I$, there is a $k \in I$ such that $i \leq k$ and $j \leq k$.

Examples 1.3.2.

(1) $\mathbb{N}, \mathbb{R}$, or any linearly ordered set.
(2) $\mathbb{R} \setminus \{a\}$ where $x \leq y$ if and only if $|x - a| \geq |y - a|$ (y is closer to a than x is).
(3) Any neighborhood base $\mathcal{T}(x)$ at $x \in X$, ordered by reverse inclusion ($U \leq V$ iff $V \subseteq U$).
(4) If X is any infinite set, $\{F \subset X | F \text{ is finite}\}$ ordered by inclusion.

Definition 1.3.3. Let X be a nonempty set and I a directed set. A *net in X based on I* (or an *I-net in X*) is a function $x : I \to X$, where write $x_i := x(i)$ and $x = (x_i)_{i \in I}$.

Given an I-net $(x_i)_{i \in I}$ and a subset $S \subset X$, we say

- (x_i) is *eventually* in S if there is some $j \in I$ such that for all $i \geq j$, $x_i \in S$.
- (x_i) is *frequently* in S if for every $j \in I$, there is an $i \geq j$ such that $x_i \in S$.

We say (x_i) *converges* to $x \in X$ if (x_i) is eventually in every neighborhood of x. We say x is a *cluster point* of (x_i) if (x_i) is frequently in every neighborhood of x.

Proposition 1.3.4. *Suppose X is a topological space and $A \subset X$. The following are equivalent for $x \in X$:*

(1) x is an accumulation/limit point of A (for all open U such that $x \in U$, $A \cap (U \setminus \{x\})$ is not empty), and
(2) there is a net in $A \setminus \{x\}$ that converges to x.

Proof.
(1) ⇒ (2): Let $\mathcal{B}(x)$ be any neighborhood base at x, ordered by reverse inclusion. (For example, we can take $\mathcal{T}(x)$, the set of all open sets which contain x.) For every $U \in \mathcal{B}(x)$, pick $x_U \in U \cap (A \setminus \{x\})$. (Observe this requires the Axiom of Choice!) Then observe that $(x_U)_{U \in \mathcal{B}(x)}$ converges to x.
(2) ⇒ (1): Exercise. □

Corollary 1.3.5. *A subset $A \subset X$ is closed if and only if every convergent net in A only converges to points in A.*

Proposition 1.3.6. *X is Hausdorff if and only if every convergent net has a unique limit.*

Proof.
⇒: If there is a net without a unique limit, any 2 distinct limit points of the same net cannot be separted by disjoint open sets.
⇐: We'll prove the contrapositive. Suppose X is not Hausdorff, so there are $x, y \in X$ such that for every neighborhoods U, V of x, y respectively, $U \cap V$ is nonempty. Let $\mathcal{B}(x), \mathcal{B}(y)$ be a neighborhood base for $\mathcal{T}$ at x, y respectively, both ordered by reverse inclusion. Direct $\mathcal{B}(x) \times \mathcal{B}(y)$ by $(U_1, V_1) \geq (U_2, V_2)$ if and only if $U_1 \subset U_2$ and $V_1 \subset V_2$. Then for all $(U, V) \in \mathcal{B}(x) \times \mathcal{B}(y)$, choose a point $x_{(U,V)} \in U \cap V$. (Again, this uses the Axiom of Choice!) This net converges to both x and y. □

Proposition 1.3.7. *A function $f : X \to Y$ is continuous if and only if for every convergent net $x_i \to x$ in X, $f(x_i) \to f(x)$ in Y.*

Proof.
⇒: Suppose $f : X \to Y$ is continuous. Let (x_i) be a convergent net with $x_i \to x$ in X. We need to show that $f(x_i) \to f(x)$ in Y. Let V be an open neighborhood of $f(x)$ in Y. Observe that $f^{-1}(V)$ is open in X, and $x \in f^{-1}(V)$. Since $x_i \to x$, (x_i) is eventually in $f^{-1}(V)$. Hence $f(x_i)$ is eventually in V.
⇐: We'll show that the preimage of every closed set is closed. Let $F \subset Y$ be closed. We may assume F is non-empty. By Corollary 1.3.5, it suffices to prove that every convergent net (x_i) in $f^{-1}(F)$ only converges to points of $f^{-1}(F)$. So suppose (x_i) is a convergent net in $f^{-1}(F)$, and say $x_i \to x$. Then $f(x_i) \in F$ for all i, and $f(x_i) \to f(x)$ by assumption. Since F is closed, by Corollary 1.3.5, $f(x) \in F$, and thus $x \in f^{-1}(F)$. □

Definition 1.3.8. A *subnet* of an I-net (x_i) consists of a J-net (y_j) together with a function $f : J \to I$ which need not be injective such that

- $y_j = x_{f(j)}$ for all $j \in J$, i.e., $y = x \circ f : J \to X$.
- for all $i \in I$, there is a $j_0 \in J$ such that $f(j) \geq i$ for all $j \geq j_0$, i.e., for every $i \in I$, $(f(j))$ is eventually greater than i.

Observe that if $x_i \to x$, then $y_j \to x$ for any subnet (y_j) of (x_i).

Proposition 1.3.9. *Suppose (x_i) is a net in X and $x \in X$. The following are equivalent:*

(1) x is a cluster point of (x_i).
(2) there is a subnet (y_j) of (x_i) such that $y_j \to x$.

Proof.
(1) ⇒ (2): Choose a neighborhood base $\mathcal{B}(x)$ at x. Define $J := I \times \mathcal{B}(x)$ where $(i_1, U_1) \leq (i_2, U_2)$ iff $i_1 \leq i_2$ and $U_1 \supset U_2$. For each $(i, U) \in J$, define $f(i, U) := i'$ to be *any* i' with $i' \geq i$ and $x_{i'} \in U$. Then if $(i_1, U_1) \leq (i_2, U_2)$, $i_1 \leq i_2 \leq f(i_2, U_2)$, and $x_{f(i_2,U_2)} \in U_2 \subset U_1$. This means $(x_{f(i,U)})$ is a subnet of (x_i) converging to x.
(2) ⇒ (1): Exercise. □

Exercise 1.3.10. When $(X, \mathcal{T})$ is first countable, then Propositions 1.3.4, 1.3.6, 1.3.7, and 1.3.9 and Corollary 1.3.5 all hold with sequences instead of nets.

Exercise 1.3.11. Suppose (X, d) is a metric space. Prove that the following are equivalent:

(1) X is compact.
(2) X is sequentially compact (every sequence has a convergent subsequence).
(3) X is complete and totally bounded.

Deduce that if in addition X is complete and $A \subset X$, then $\overline{A}$ is compact if and only if A is totally bounded.

Theorem 1.3.12. *Suppose X is a topological space. The following are equivalent:*

(1) X is compact.
(2) For every family of closed sets (F_i) with the finite intersection property, $\bigcap F_i$ is nonempty.
(3) Every net in X has a cluster point.
(4) Every net in X has a convergent subnet.

Proof.
(1) ⇔ (2): This is Exercise 1.2.2.
(3) ⇔ (4): This follows by Proposition 1.3.9.
(2) ⇒ (3): Let (x_i) be a net in X. For $i \in I$, define $A_i := \{x_j | j \geq i\}$. Observe $\bigcap \overline{A_i}$ is the set of cluster points of (x_i). Moreover, (A_i) has the finite intersection property, so $(\overline{A_i})$ also has the finite intersection property. We conclude by (2) that $\bigcap \overline{A_i}$ is nonempty, and thus (x_i) has a cluster point.
(3) ⇒ (2): We'll prove the contrapositive. If (2) fails, then there is a family of closed sets (F_i) with the finite intersection property such that $\bigcap F_i = \emptyset$. Define J to be the set of non-empty finite intersections of (F_i) ordered by reverse inclusion. Since (F_i) has the finite intersection property, for every $F \in J$, F is nonempty. Use the Axiom of Choice to pick $x_F \in F$ for every $F \in J$. Then any cluster point of (x_F) lies in $\bigcap_{F \in J} F = \bigcap F_i = \emptyset$. □

1.3.2. *Filters.*

Exercise 1.3.13 (Pedersen *Analysis Now*, E1.3.4 and E1.3.6)**.** A *filter* on a set X is a collection $\mathcal{F}$ of non-empty subsets of X satisfying

- $A, B \in \mathcal{F}$ implies $A \cap B \in \mathcal{F}$, and
- $A \in \mathcal{F}$ and $A \subset B$ implies $B \in \mathcal{F}$.

Suppose $\mathcal{T}$ is a topology on X. We say a filter $\mathcal{F}$ *converges* to $x \in X$ if every open neighborhood U of x lies in $\mathcal{F}$.

(1) Show that $A \subset X$ is open if and only if $A \in \mathcal{F}$ for every filter $\mathcal{F}$ that converges to a point in A.
(2) Show that if $\mathcal{F}$ and $\mathcal{G}$ are filters and $\mathcal{F} \subset \mathcal{G}$ ($\mathcal{G}$ is a *subfilter* of $\mathcal{F}$), then $\mathcal{G}$ converges to x whenever $\mathcal{F}$ converges to x.
(3) Suppose (x_λ) is a net in X. Let $\mathcal{F}$ be the collection of sets A such that (x_λ) is eventually in A. Show that $\mathcal{F}$ is a filter. Then show that $x_\lambda \to x$ if and only if $\mathcal{F}$ converges to x.
(4) Show that $(X, \mathcal{T})$ is Hausdorff if and only if every convergent filter has a unique limit.

Exercise 1.3.14 (Pedersen *Analysis Now*, E1.3.5)**.** A filter $\mathcal{F}$ on a set X is called an *ultrafilter* if it is not properly contained in any other filter.

(1) Show that a filter $\mathcal{F}$ is an ultrafilter if and only if for every $A \subset X$, we have either $A \in \mathcal{F}$ or $A^c \in \mathcal{F}$.
(2) Use Zorn's Lemma to prove that every filter is contained in an ultrafilter.

Exercise 1.3.15. Let X and Y be sets and $f : X \to Y$ a function. Let $\mathcal{F}$ be an ultrafilter on X. Prove that $f^*(\mathcal{F}) := \{A \subset Y | f^{-1}(A) \in \mathcal{F}\}$ is an ultrafilter on Y.

Exercise 1.3.16. Given a filter $\mathcal{F}$ on X, show that $\mathcal{F}$ is an ultrafilter if and only if $\bigcup_{i=1}^n A_i \in \mathcal{F}$ implies that $A_i \in \mathcal{F}$ for some $i \in \{1, \dots, n\}$.

Exercise 1.3.17. Let X be a nonempty set and let $\mathcal{U}$ be a collection of subsets of X.
Note: It is not assumed that $\mathcal{U}$ is a filter!
Show that the following two statements are equivalent.

(1) $\mathcal{U}$ is an ultrafilter on X.
(2) Whenever X can be partitioned into three disjoint sets $X = A_1 \amalg A_2 \amalg A_3$, there is a unique $i \in \{1, 2, 3\}$ such that $A_i \in \mathcal{U}$.
Hint: The A_i's need not be distinct nor non-empty.

Exercise 1.3.18. Let $(X, \mathcal{T})$ be a topological space. A net $(x_\lambda)_{\lambda \in \Lambda}$ is called *universal* if for every subset $Y \subset X$, (x_λ) is either eventually in Y or eventually in Y^c. Show that every net has a universal subnet.
Hint: Let (x_λ) be a net in X. We say a filter $\mathcal{F}$ on X is associated to *(x_λ) if (x_λ) is frequently in every $F \in \mathcal{F}$.*

(1) Show that the set of filters associated to (x_λ) is non-empty.
(2) Order the set of filters associated to (x_λ) by inclusion. Show that if $(\mathcal{F}_j)$ is a totally ordered set of filters for (x_λ), then $\cup \mathcal{F}_j$ is also a filter for (x_λ).
(3) Use Zorn's Lemma to assert there is a maximal filter $\mathcal{F}$ associated to (x_λ).
(4) Show that $\mathcal{F}$ is an ultrafilter.
(5) Find a subnet of (x_λ) that is universal.

Exercise 1.3.19. Let $(X, \mathcal{T})$ be a topological space. Prove that the following are equivalent:

(1) $(X, \mathcal{T})$ is compact
(2) every ultrafilter converges
(3) every universal net converges.

1.4. **Categories, universal properties, and product topology.**

Definition 1.4.1. A *category* $\mathcal{C}$ is a collection of *objects* together with a set of *morphisms* $\mathcal{C}(a \to b)$ for every ordered pair of objects $a, b \in \mathcal{C}$ and a composition operation $- \circ_{\mathcal{C}} - : \mathcal{C}(b \to c) \times \mathcal{C}(a \to b) \to \mathcal{C}(a \to c)$, i.e., $f : a \to b$ and $g : b \to c$, then $g \circ f : a \to c$ such that

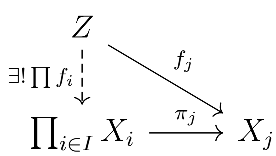
- composition is associative, i.e., $h \circ (g \circ f) = (h \circ g) \circ f$ for all $f : a \to b$, $g : b \to c$, and $h : c \to d$.
- every object has an identity morphism, i.e., for every $b \in \mathcal{C}$, there is a $\mathrm{id}_b : b \to b$ such that $\mathrm{id}_b \circ f = f$ for all $f : a \to b$ and $g \circ \mathrm{id}_b = g$ for all $g : b \to c$

Definition 1.4.2. Suppose $(X_i)_{i \in I}$ is a family of sets. The (categorical) *product* is the Cartesian product

$$\prod_{i \in I} X_i := \left\{ x : I \to \bigcup_{i \in I} X_i \,\middle|\, x_i := x(i) \in X_i \right\}$$

together with the canonical projection maps $\pi_j : \prod X_i \to X_j$ given by $\pi_j(x) = x_j$. It satisfies the following *universal property* :

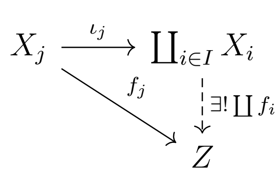
- (product) for any set Z and functions $f_i : Z \to X_i$ for $i \in I$, there is a unique function $\prod f_i : Z \to \prod X_i$ such that $\pi_j \circ \prod f_i = f_j$ for all $j \in I$.



Exercise 1.4.3. Suppose Y is another set together with functions $\theta_i : Y \to X_i$ for all $i \in I$ satisfying the universal property of the product. Show there is a unique bijection between Y and $\prod X_i$ which is compatible with the projection maps. In this sense, we say that the product is *unique up to unique isomorphism.*

Exercise 1.4.4. A set $\coprod X_i$ together with maps $\iota_j : X_j \to \coprod X_i$ for each $j \in I$ is called the *coproduct* of $(X_i)_{i \in I}$ if it satisfies the following universal property:

- (coproduct) for any set Z and functions $f_i : X_i \to Z$ for $i \in I$, there is a unique function $\coprod f_i : \coprod X_i \to Z$ such that $(\coprod f_i) \circ \iota_j = f_j$.



(1) Show that the coproduct, if it exists, is unique up to unique isomorphism.
(2) What is the coproduct in the category of sets?

Definition 1.4.5. Suppose $(X_i)_{i\in I}$ is a family of topological spaces. The (categorical) *product* is the Cartesian product $\prod_{i\in I} X_i$ equipped with the *weakest* topology such that the canonical projection maps $\pi_j : \prod X_i \to X_j$ are continuous for every $j \in I$. We call this topology the *product topology*.

Exercise 1.4.6. Prove that the open sets $\prod U_i$ with $U_i \subset X$ open where only finitely many of the U_i are not equal to X_i form a base for the product topology.

Exercise 1.4.7. Prove that $\prod X_i$ with the product topology together with the canonical projection maps $\pi_j : \prod X_i \to X_j$ is the categorical product in the category of topological spaces with continuous maps. That is, prove the product satisfies the universal property in Definition 1.4.2 subject to the additional condition that all functions are continuous.

Exercise 1.4.8. What is the categorical coproduct of topological spaces?

Theorem 1.4.9 (Tychonoff)**.** *Suppose $(X_i)_{i\in I}$ is a family of compact topological spaces. Then the product $\prod X_i$ is compact in the product topology.*

Proof. Discussion section. □

Definition 1.4.10. Suppose $\mathcal{C}, \mathcal{D}$ are categories. A (covariant) *functor* $F : \mathcal{C} \to \mathcal{D}$ assigns to each object $c \in \mathcal{C}$ an object $F(c) \in \mathcal{D}$ and to each morphism $f \in \mathcal{C}(a \to b)$ a morphism $F(f) \in \mathcal{D}(F(a) \to F(b))$ such that

- $F(\mathrm{id}_c) = \mathrm{id}_{F(c)}$ for all objects $c \in \mathcal{C}$, and
- $F(g \circ f) = F(g) \circ F(f)$ for all $f \in \mathcal{C}(a \to b)$ and $g \in \mathcal{C}(b \to c)$.

A *contravariant functor* $F : \mathcal{C} \to \mathcal{D}$ is similar to a functor, but instead of the second bullet point above, we have $F(g \circ f) = F(f) \circ F(g)$ for composable f, g.

Exercise 1.4.11. Let Set denote the category of sets and functions.

(1) For a function $f : X \to Y$, define $P(f) : P(X) \to P(Y)$ by $P(f)(A) = f(A) = \{f(a)|a \in A\}$. Show that $P\mathsf{Set} \to \mathsf{Set}$ is a functor.
(2) For a set X, define $P^{-1}(X) := P(X) = \{A \subset X\}$. For a function $f : X \to Y$ and $B \subset Y$, define $P^{-1}(f)(B) := f^{-1}(B) = \{x \in X | f(x) \in B\}$. Show that $P^{-1} : \mathsf{Set} \to \mathsf{Set}$ is a contravariant functor.

Exercise 1.4.12. Let Top denote the category topological spaces and continuous maps.

(1) There is a forgetful functor $\mathrm{Forget} : \mathsf{Top} \to \mathsf{Set}$ which forgets the topology.
(2) Given a set X, we can endow it with the discrete topology $\mathcal{T}_{\text{disc}} := P(X)$. This gives a functor $L : \mathsf{Set} \to \mathsf{Top}$. Show that if Y is any topological space, then every function $X \to Y$ is continuous with respect to the discrete topology on X. In other words,
$$\mathsf{Top}(L(X) \to Y) = \mathsf{Set}(X \to \mathrm{Forget}(Y)).$$
(3) Given a set Y, we can endow it with the trivial topology $\mathcal{T}_{\text{triv}} := \{\emptyset, Y\}$. This gives a functor $R : \mathsf{Set} \to \mathsf{Top}$. Show that if X is any topological space and Y is a set, then every function $X \to Y$ is continuous with respect to the trivial topology on Y. In other words,
$$\mathsf{Set}(\mathrm{Forget}(X) \to Y) = \mathsf{Top}(X \to R(Y)).$$

Exercise 1.4.13. Let CptHsd denote the category of compact Hausdorff topological spaces and continuous maps. Let $\mathsf{Alg_u}$ denote the category of unital complex algebras and unital algebra homomorphisms. Show that $X \mapsto C(X)$ and $f : X \to Y$ maps to $- \circ f : C(Y) \to C(X)$ gives a contravariant functor $\mathsf{CptHsd} \to \mathsf{Alg_u}$.

Exercise 1.4.14.

(1) Given LCH spaces X, Y and a continuous function $f : X \to Y$, when does the image of the map $- \circ f : C_0(Y) \to C(X)$ lie in $C_0(X)$?
(2) Show that on the correct category LCH of locally compact Hausdorff topological spaces, the assignments $X \mapsto C_0(X)$ and $f \mapsto - \circ f$ define a contravariant functor to Alg, the category of non-unital complex algebras and algebra homomorphisms

1.5. **The Stone-Weierstrass Theorem.** Weierstrass' original theorem from 1885:

(1) The polynomials are dense in $C[a, b]$ where $-\infty < a < b < \infty$.
(2) A continuous function on $\mathbb{R}$ with period 2π can be uniformly approximated by a finite linear combination of functions of the form $\sin(nx), \cos(nx)$ for $n \in \mathbb{N}$, i.e., a trigonometric polynomial.

Theorem 1.5.1 ($\mathbb{R}$-Stone-Weierstrass)**.** *Suppose X is compact Hausdorff and $A \subset C(X, \mathbb{R})$ is a closed $\mathbb{R}$-subalgebra which separates points (for all distinct $x, y \in X$, there is an $f \in A$ such that $f(x)$ and $f(y)$ are distinct).*

- *If A contains a non-vanishing function, then $A = C(X, \mathbb{R})$.*
- *If every $f \in A$ has a zero, then there exists a unique $x_0 \in X$ such that*

$$A = \{f \in C(X, \mathbb{R}) | f(x_0) = 0\}.$$

Exercise 1.5.2. Suppose X is compact Hausdorff and $A \subset C(X, \mathbb{F})$ is a subalgebra where $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$. Prove that $\overline{A}$ is also a subalgebra. Deduce that if A separates points, then so does $\overline{A}$.

Lemma 1.5.3. *On any compact $K \subset \mathbb{R}$, the function $x \mapsto |x|$ on $\mathbb{R}$ can be uniformly approximated on K by a polynomial which vanishes at zero.*

Proof. We give a proof of Sarason. We'll show for $R > 0$, there is a sequence of polynomials (p_n) which converges uniformly to $|\cdot|$ on $[-R, R]$ such that $p_n(0) = 0$ for all n. Without loss of generality, $R = 1$. It suffices to find a sequence (q_n) of polynomials converging to $q(t) := 1 - |t|$ on $[-1, 1]$ such that $q_n(0) = 1$ for all n. Observe that

$$q \text{ takes values in } [0, 1] \text{ and } (1 - q(t))^2 = t^2 \text{ for all } |t| \leq 1. \quad (*)$$

For a given $t \in [-1, 1]$, consider the equation $(1 - s)^2 = t^2$. It has 2 solutions, namely $s = 1 \pm |t|$, and exactly one of these values of s lies in $[0, 1]$. Hence $q(t)$ is unique function on $[-1, 1]$ satisfying $(*)$. We can rewrite $(*)$ as

$$q \text{ takes values in } [0, 1] \text{ and } q(t) = \frac{1}{2}(1 - t^2 + q(t)^2). \quad (**)$$

We define (q_n) inductively by

- $q_0(t) = 1$, and
- $q_{n+1}(t) = \frac{1}{2}(1 - t^2 + q_n(t)^2)$.

By induction, for all $n \geq 0$, we have q_n takes values in $[0,1]$, $q_n(0) = 1$, and

$$q_n - q_{n+1} = \frac{1}{2}(q_{n-1}^2 - q_n^2) = \frac{1}{2}(q_{n-1} - q_n)(q_{n-1} + q_n) \geq 0.$$

(Indeed, observe that $q_1(t) = 1 - \frac{1}{2}t^2$, so $q_0 - q_1 = \frac{1}{2}t^2 \geq 0$.) This means that (q_n) is monotone decreasing by construction. Let $\widetilde{q}$ be the pointwise limit. Observe that $\widetilde{q}$ satisfies ($**$) by construction, so $\widetilde{q} = q$ by uniqueness! Now as $q_n \searrow q$ on $[-1,1]$ pointwise, $q_n \to q$ uniformly by Dini's Lemma (Exercise 1.2.14). □

Lemma 1.5.4. *If $A \subset C(X, \mathbb{R})$ is a closed $\mathbb{R}$-subalgebra, then A is a lattice (for all $f, g \in A$, the functions $f \vee g := \max\{f, g\}$ and $f \vee g := \min\{f, g\}$ belong to A).*

Proof. Suppose $a \in A$ and $a \neq 0$. Then $\frac{a}{\|a\|_\infty} : X \to [-1,1]$. By Lemma 1.5.3, for all $\varepsilon > 0$, there is a polynomial p on $[-1,1]$ with $p(0) = 0$ and $|\,|t| - p(t)| < \varepsilon$ for all $t \in [-1,1]$. Hence

$$\left| \frac{|a(x)|}{\|a\|_\infty} - p\left(\frac{a(x)}{\|a\|_\infty}\right) \right| < \varepsilon \qquad \forall\, x \in X.$$

In other words,

$$\left\| \frac{|a|}{\|a\|_\infty} - \underbrace{p\left(\frac{a}{\|a\|_\infty}\right)}_{\in A} \right\|_\infty < \varepsilon.$$

Since $p(0) = 0$, $p(a/\|a\|_\infty) \in \operatorname{span}\{a^n | n \in \mathbb{N}\} \subset A$. Since the algebra A is closed and $\varepsilon > 0$ was arbitrary, $|a|/\|a\|_\infty \in A$, and thus $|a| \in A$. Hence for all $a, b \in A$,

$$\max\{a, b\} = \frac{1}{2}(a + b + |a - b|)$$
$$\min\{a, b\} = \frac{1}{2}(a + b - |a - b|)$$

are both elements of A. □

Lemma 1.5.5. *Suppose $A \subset C(X, \mathbb{R})$ is a $\mathbb{R}$-vector space which is also a lattice. Suppose $f \in C(X, \mathbb{R})$ satisfies*

- *for all $\varepsilon > 0$ and all distinct $x, y \in X$, there is an $a_{x,y} \in A$ such that*
$$|f(x) - a_{x,y}(x)| < \varepsilon \qquad \text{and} \qquad |f(y) - a_{x,y}(y)| < \varepsilon.$$

Then $f \in \overline{A}$.

Proof. For every $\varepsilon > 0$ and $x, y \in X$, pick $a_{x.y} \in A$ such that $|f(x) - a_{x,y}(x)| < \varepsilon$ and $|f(y) - a_{x,y}(y)| < \varepsilon$. Then x, y are both in:

$$U_{x,y} = \{z \in X | f(z) < a_{x,y}(z) + \varepsilon\}$$
$$V_{x,y} = \{z \in X | a_{x,y}(z) < f(z) + \varepsilon\}.$$

Fix $x \in X$. Then sets $(U_{x,y})_{y \in X}$ are an open cover of X. Since X is compact, $X \subset \bigcup_{i=1}^n U_{x,y_i}$ for some $y_1, \ldots, y_n \in X$. Then $a_x := \bigvee_{i=1}^n a_{x,y_i} \in A$, and $f(z) < a_x(z) + \varepsilon$ for all $z \in X$ in construction. Also, $a_x(z) < f(z) + \varepsilon$ for all $z \in W_x := \bigcap_{i=1}^n V_{x,y_i}$, which is some open neighborhood of x. Varying over $x \in X$, $(W_x)_{x \in X}$ are an open cover, so there are finitely many $x_1, \ldots, x_k \in X$ such that $X \subset \bigcup_{i=1}^k W_{x_i}$ by compactness. Setting $a_\varepsilon := \bigwedge_{i=1}^k a_{x_i}$ satisfies $\|f - a_\varepsilon\|_\infty < \varepsilon$. Since ε was arbitrary, we conclude that $f \in \overline{A}$. □

Proof of the $\mathbb{R}$*-Stone-Weierstrass Theorem 1.5.1.* Suppose $x \neq y$ in X. Since point evaluation is an $\mathbb{R}$-algebra homomorphism $A \to \mathbb{R}$, then

$$A_{x,y} := \{(f(x), f(y)) | f \in A\} \subset \mathbb{R}^2$$

is a $\mathbb{R}$-subalgebra. The only $\mathbb{R}$-subalgebras of $\mathbb{R}^2$ are:

$$(0,0) \qquad \mathbb{R} \times \{0\} \qquad \{0\} \times \mathbb{R} \qquad \Delta = \{(x,x) | x \in \mathbb{R}\} \qquad \mathbb{R}^2.$$

Since A separates points, $A_{x,y} \neq (0,0)$ or Δ for all $x \neq y$.

Claim. *$A_{x,y} = \mathbb{R}^2$ for all $x \neq y$ except for when x, y are equal to one possible $x_0 \in \mathbb{R}$.*

Proof. If there are $x \neq y$ such that $A_{x,y} \neq \mathbb{R}^2$, then without loss of generality, $A_{x,y} = \{0\} \times \mathbb{R}$. Thus $f(x) = 0$ for all $f \in A$. Since A separates points, $f(x') = 0$ for all $f \in A$ implies $x' = x$. So $A_{y,z} = \mathbb{R}^2$ for all $y \neq x \neq z$. □

Claim. *$A_{x,y} = \mathbb{R}^2$ for all $x \neq y$ if and only if A contains a non-vanishing function.*

Proof of Claim. If A contains a non-vanishing function, then $A_{x,y} = \mathbb{R}^2$ for all $x \neq y$. Conversely, suppose $A_{x,y} = \mathbb{R}^2$ for all $x \neq y$. Then for all $x \in X$, choose a continuous function $a_x \in A$ such that $a_x(x) \neq 0$. Observe that the sets $(U_x := \{a_x \neq 0\})_{x \in X}$ form an open cover of X, so by compactness, there are $x_1, \ldots, x_n$ such that $X \subset \bigcup_{i=1}^n U_{x_i}$. By Lemma 1.5.4, A is a lattice, so

$$a := \max\{a_{x_1}, \ldots, a_{x_n}, -a_{x_1}, \ldots, -a_{x_n}\} = \max\{|a_{x_1}|, \ldots, |a_{x_n}|\} \in A.$$

Since $|a_{x_i}| > 0$ on U_{x_i} for all $i = 1, \ldots, n$, we have $a(x) > 0$ for all $x \in X$ by construction. □

From these claims, we see that either A contains a non-vanishing function, in which case $A_{x,y} = \mathbb{R}^2$ for all $x \neq y$, or every function in A vanishes at some point of X, in which case there is a unique $x_0 \in X$ such that $a(x_0) = 0$ for all $a \in A$.

Case 1: For all $x \neq y$ in X and $f \in C(X, \mathbb{R})$, there is an $a_{x,y} \in A$ such that $f(x) = a_{x,y}(x)$ and $f(y) = a_{x,y}(y)$. By Lemma 1.5.4, A is a lattice, and by Lemma 1.5.5, $f \in A$.

Case 2: For all $x_0 \neq x \neq y \neq x_0$ and $f \in \{g \in C(X, \mathbb{R}) | g(x_0) = 0\}$ (which is a closed subalgebra/ideal of $C(X, \mathbb{R})$), there is an $a_{x,y} \in A$ such that $f(x) = a_{x,y}(x)$ and $f(y) = a_{x,y}(y)$. By Lemma 1.5.4, A is a lattice, and by Lemma 1.5.5, $f \in A$. □

Theorem 1.5.6 ($\mathbb{C}$-Stone-Weierstrass)**.** *Suppose X is a compact Hausdorff space. Let $A \subset C(X)$ be a closed subalgebra that separates points of X and is closed under complex conjugation.*

- *If A contains a non-vanishing function, then $A = C(X)$.*
- *If every $f \in A$ has a zero, then there exists a unique $x_0 \in X$ such that*

$$A = \{f \in C(X) | f(x_0) = 0\}.$$

Proof. Note that $A_{\text{sa}} := \left\{f \in A \middle| f = \overline{f}\right\}$ is an $\mathbb{R}$-subalgebra of A. (Here, 'sa' stands for *self-adjoint*.) Since A is closed under complex conjugation, for all $f \in A$, $\operatorname{Re}(f), \operatorname{Im}(f) \in A$, and thus $A = A_{\text{sa}} \oplus iA_{\text{sa}}$. Moreover, $C(X) = C(X, \mathbb{R}) \oplus iC(X, \mathbb{R})$ by similar reasoning. Hence the strategy is to apply the $\mathbb{R}$-Stone-Weierstrass Theorem 1.5.1 to $A_{\text{sa}} \subset C(X, \mathbb{R})$.

First, observe A_{sa} separates points, since if $f \in A$ separates x, y, then one of $\text{Re}(f), \text{Im}(f) \in A_{\text{sa}}$ separates x, y. Second, observe that A_{sa} is closed, since if $(f_n) \subset A_{\text{sa}}$ converges uniformly, then its limit lies in A as A is closed, and since (f_n) must converge pointwise, its limit only takes real values and thus lies in A_{sa}.

We now check the two cases in the statement of the theorem.

Case 1: If A contains a non-vanishing function f, then $|f|^2 = \overline{f}f \in A_{\text{sa}}$ does not vanish. By the $\mathbb{R}$-Stone-Weierstrass Theorem 1.5.1, $A_{\text{sa}} = C(X, \mathbb{R})$, and thus

$$A = A_{\text{sa}} \oplus iA_{\text{sa}} = C(X, \mathbb{R}) \oplus iC(X, \mathbb{R}) = C(X).$$

Case 2: If every element of A vanishes somewhere, then so does every element of $A_{\text{sa}} \subset A$. By the $\mathbb{R}$-Stone-Weierstrass Theorem 1.5.1, $A_{\text{sa}} = \{f \in C(X, \mathbb{R}) | f(x_0) = 0\}$, and thus

$$\begin{aligned} A &= A_{\text{sa}} \oplus iA_{\text{sa}} \\ &= \{f \in C(X, \mathbb{R}) | f(x_0) = 0\} \oplus i \{f \in C(X, \mathbb{R}) | f(x_0) = 0\} \\ &= \{f \in C(X) | f(x_0) = 0\}. \end{aligned}$$

□

Exercise 1.5.7. Suppose X is LCH and $A \subset C_0(X)$ is a closed subalgebra that separates points and is closed under complex conjugation. Then either $A = C_0(X)$ or $A = \{f \in C_0(X) | f(x_0) = 0\}$ for some $x_0 \in X$.
Hint: Use the one point (Alexandroff) compactification discussed in §1.6 below.

Exercise 1.5.8. Show the following collections of functions are uniformly dense in the appropriate algebras:

(1) For $a < b$ in $\mathbb{R}$, the polynomials $\mathbb{R}[t] \subset C([a, b], \mathbb{R})$.
(2) For $a < b$ in $\mathbb{R}$, the piece-wise linear functions $PWL \subset C([a, b], \mathbb{R})$.
(3) For $K \subset \mathbb{C}$ compact, the polynomials $\mathbb{C}[z, \overline{z}] \subset C(K)$.
(4) For $\mathbb{R}/\mathbb{Z}$, the trigonometric polynomials $\text{span}\{\sin(2\pi nx), \cos(2\pi nx) | n \in \mathbb{N} \cup \{0\}\} \subset C(\mathbb{R}/\mathbb{Z}, \mathbb{R})$.

Exercise 1.5.9.

(1) Use the difference quotient to show that complex cojugation $\overline{\cdot} : \mathbb{C} \to \mathbb{C}$ given by $z \mapsto \overline{z}$ is nowhere complex differentiable.
(2) Let $\mathbb{D} \subset \mathbb{C}$ be the open unit disk $\{|z| < 1\}$. Describe the uniform closure of $\mathbb{C}[z]$, the polynomials in z, in $C(\overline{\mathbb{D}})$.
Hint: You may use without proof Morera's Theorem from Complex Analysis which states on any open domain $U \subset \mathbb{C}$, the local uniform limit of complex differentiable functions is complex differentiable.
(3) Discuss your answer in the context of the Stone-Weierstrass Theorem.

Exercise 1.5.10. Let X, Y be compact Hausdorff spaces. For $f \in C(X)$ and $g \in C(Y)$, define $(f \otimes g)(x, y) := f(x)g(y)$. Prove that $\text{span}\{f \otimes g | f \in C(X) \text{ and } g \in C(Y)\}$ is uniformly dense in $C(X \times Y)$.

Exercise 1.5.11 (Sarason)**.** Suppose $f \in C([0, 1], \mathbb{R})$ such that $\int_0^1 x^n f(x)\, dx = 0$ for all $n \geq 2020$. Prove that $f = 0$.
Hint: Consider $A := \text{span}\{x^n | n \geq 2020\} \subset C([0, 1], \mathbb{R})$.

Exercise 1.5.12 (Sarason)**.** Find a sequence of polynomials in $\mathbb{R}[t] \subset C(\mathbb{R}, \mathbb{R})$ that simultaneously converges to 1 uniformly on every compact subinterval of $(0, \infty)$ and to -1 uniformly on every compact subinterval of $(-\infty, 0)$.

1.6. **One point (Alexandroff) and Stone-Čech compactification.**

Definition 1.6.1. Suppose X is a topological space. An *embedding* $\varphi : X \to Y$ is a continuous injection which is a homeomorphism onto its image, i.e., $\varphi^{-1} : \varphi(X) \to X$ is continuous with respect to the relative topology.

A *compactification* of a topological space X consists of a compact space K and an embedding $\varphi : X \to K$ such that $\varphi(X)$ is dense in K.

Example 1.6.2. Consider the map $[0, 1) \to S^1 := \{z \in \mathbb{C} | |z| = 1\}$ by $r \mapsto \exp(2\pi i r)$. This map is a continuous bijection, but not a homeomorphism onto its image.

Examples 1.6.3. Compactifications of $\mathbb{R}$ include:

1. the extended real numbers $\overline{\mathbb{R}} = [-\infty, \infty]$
2. the 'one point' compactification $\mathbb{R} \cup \{\infty\} \cong S^1$
3. You can add $(0, 0)$ and S^1 in $\mathbb{R}^2$ to an embedding $\mathbb{R} \hookrightarrow \mathbb{R}^2$ as a spiral.
4. You can add a circle S^1 embedded in a 2-toruse $\mathbb{T}^2 \subset \mathbb{R}^3$ to an embedding $\mathbb{R} \hookrightarrow \mathbb{T}^2$ which coils $\mathbb{R}$ around the torus from either side.

Definition 1.6.4. Suppose X is an LCH space, and choose any object $\infty \notin X$. Define $X^\bullet := X \amalg \{\infty\}$, where $\amalg$ denotes disjoint union (coproduct in Set). We say $U \subset X^\bullet$ is open if and only if either

- $U \subset X$ is open in X, or
- $\infty \in U$, and U^c is compact.

Due to the next theorem, we call $X^\bullet$ the (Alexandroff) *one point compactification* of X.

Theorem 1.6.5. *If X is LCH, then the space $X^\bullet$ is compact Hausdorff, and the inclusion $X \hookrightarrow X^\bullet$ is an embedding.*

Proof. The inclusion $X \hookrightarrow X^\bullet$ is obviously an embedding.

Compact: Suppose (U_i) is an open cover of $X^\bullet$. Then there is some U_0 such that $\infty \in U_0$ and U_0^c is compact. Then $(U_i \cap X)$ is an open cover of U_0^c, which is compact. So pick a finite subcover.

Hausdorff: Since X is Hausdorff, it suffices to separate $x \in X$ from $\infty \in X^\bullet$. Since X is LCH, there is an open neighborhood $U \subset X$ of x such that $\overline{U} \subset X$ is compact. Set $V := \overline{U}^c$ in $X^\bullet$, which is an open neighborhood of ∞ disjoint from U. ☐

Definition 1.6.6. A topological space X is *completely regular* if for every closed $F \subset X$ and $x \in F^c$, there is a continuous function $f : X \to [0, 1]$ such that $f(x) = 1$ and $f|_F = 0$.

We call X *Tychonoff* if X is completely regular and T_1.

Exercises 1.6.7.

1. X Tychonoff implies X is Hausdorff.
2. Every normal space is Tychonoff by Urysohn's Lemma.
3. Any subspace of a Tychonoff space is Tychonoff.

Lemma 1.6.8 (Embedding). *Suppose X is a topological space $\Phi \subset C(X,[0,1])$ is a family of continuous functions. Define $e : X \to [0,1]^\Phi := \{f : \Phi \to [0,1]\} = \prod_{f\in\Phi}[0,1]$ (which is compact in the product topology!) by $x \mapsto (f(x))_{f\in\Phi}$.*

(1) e is continuous.
(2) e is injective if and only if Φ separates points, i.e., for all $x \neq y$ in X, there is an $f \in \Phi$ such that $f(x) \neq f(y)$.
(3) If Φ separates points from closed sets (for all $F \subset X$ closed and $x \in F^c$, there is an $f \in \Phi$ such that $f(x) \notin \overline{f(F)}$), then e is an open map of X onto $e(X)$.
(4) If Φ separates points and Φ separates points from closed sets, then e is an embedding.

Proof.

(1) Observe that $\pi_f \circ e = f$ is continuous for all $f \in \Phi$. Thus e is continuous by the universal property defining the product in Top.
(2) $e(x) \neq e(y)$ if and only if there is an $f \in \Phi$ such that

$$f(x) = (\pi_f \circ e)(x) \neq (\pi_f \circ e)(y) = f(y).$$

(3) Suppose Φ separates points from closed sets. Let $U \subset X$ be open. Suppose $x \in U$. We want to find an open set $V \subset [0,1]^\Phi$ such that $e(x) \in V \cap e(X) \subset e(U)$. There is an $f \in \Phi$ such that $f(x) \notin \overline{f(U^c)}$. Then $W := [0,1] \setminus \overline{f(U^c)}$ is an open set containing $f(x)$, so $e(x) \in \pi_f^{-1}(W)$, which is open in $[0,1]^\Phi$. Observe that

$$e(y) \in \pi_f^{-1}(W) \cap e(X) \quad \Longleftrightarrow \quad f(y) \notin \overline{f(U^c)} \quad \Longrightarrow \quad y \in U.$$

Setting $V := \pi_f^{-1}(W)$, we have $e(x) \in V \cap e(X) \subset e(U)$ as desired.
(4) By (1) and (2), $e : X \to [0,1]^\Phi$ is a continuous injection. By (3), e^{-1} on $e(X)$ is continuous. So e is a homeomorphism onto its image. $\square$

Corollary 1.6.9. *X is Tychonoff if and only if there exists an embedding $X \hookrightarrow [0,1]^I$ for some set I.*

Definition 1.6.10. Suppose X is Tychonoff and set $\Phi := C(X,[0,1])$. Consider the embedding $e : X \hookrightarrow [0,1]^\Phi$ by $e(x)_f := f(x)$. The *Stone-Čech compactification* of X is $\beta X := \overline{e(X)}$, with $X \to \beta X$ being the corestriction of e, still denoted e.

Suppose $f : X \to Y$ is any continuous map between Tychonoff spaces. Define $F : [0,1]^{\Phi_X} \to [0,1]^{\Phi_Y}$ componentwise for $g \in \Phi_Y = C(Y,[0,1])$ by $\pi_g(F(p)) := \pi_{g\circ f}(p)$. Then F is continuous, since $\pi_g \circ F = \pi_{g\circ f} : [0,1]^{\Phi_X} \to [0,1]$ is continuous for all $g \in \Phi_Y$. Moreover, for all $x \in X$,

$$\pi_g(F(e_X(x))) = \pi_{g\circ f}(e_X(x)) = g(f(x)) = \pi_g(e_Y(f(x)).$$

This means that $F \circ e_X = e_Y \circ f : X \to [0,1]^{\Phi_Y}$. Hence $\operatorname{im}(F|_{\beta X}) \subset \overline{e_Y(Y)} = \beta Y$. Define $\beta f := F|_{\beta X} : \beta X \to \beta Y$. Observe we have the following commutative diagram:

$$\begin{array}{ccccc} X & \xhookrightarrow{e_X} & \beta X & \longhookrightarrow & [0,1]^{\Phi_X} \\ \downarrow f & & \downarrow \beta f & & \downarrow F \\ Y & \xhookrightarrow{e_Y} & \beta Y & \longhookrightarrow & [0,1]^{\Phi_Y}. \end{array} \tag{1.6.11}$$

Remark 1.6.12. Suppose X, Y are Tychonoff and $f : X \to Y$ is continuous. We note for future use that if every $h \in \Phi_X$ factorizes as $h = g \circ f$ for some $g \in \Phi_Y$, then F from Definition 1.6.10 is injective. Indeed, if $p, p' \in [0,1]^{\Phi_X}$, we have

$$\begin{aligned} F(p) = F(p') &\iff \pi_g(F(p)) = \pi_g(F(p')) && \forall\, g \in \Phi_Y \\ &\iff \pi_{g\circ f}(p) = \pi_{g\circ f}(p') && \forall\, g \in \Phi_Y \\ &\iff \pi_h(p) = \pi_h(p') && \forall\, h \in \Phi(X) \\ &\iff p = p'. \end{aligned}$$

Theorem 1.6.13. *The Stone-Čech compactification $(\beta X, e)$ satisfies the universal property*

- *For every compact Hausdorff space Z and continuous function $f : X \to Z$, there exists a unique continuous function $\beta f : \beta X \to Z$ such that $\widetilde{f} \circ e = f$.*

$$\begin{array}{ccc} \beta X & & \\ {\scriptstyle e}\uparrow & \searrow^{\exists!\widetilde{f}} & \\ X & \xrightarrow{f} & Z. \end{array}$$

Proof. First, given any compactification $\varphi : X \to K$, compact Hausdorff Z, and continuous map $f : X \to Z$, there exists *at most one* continuous function $g : K \to Z$ such that $g \circ \varphi = f$. So it suffices to prove existence of $\widetilde{f}$. Just observe that since Z is compact, $e_Z(Z) \subset \beta Z$ is dense and compact, so $e_Z(Z) = \beta Z$. Hence $e_Z : Z \to \beta Z$ is a continuous bijection from a compact space to a Hausdorff space, and is thus a homeomorphism. So the map $\widetilde{f} : \beta X \to Z$ given by

$$\beta X \xrightarrow{\beta f} \beta Z \xrightarrow{e_Z^{-1}} Z$$

satisfies $\widetilde{f} \circ e_X = f$ by the commutative diagram (1.6.11). □

Exercise 1.6.14. If $\varphi : X \hookrightarrow Y$ is any compactification of X satisfying the universal property in Theorem 1.6.13, then $\widetilde{\varphi} : \beta X \to Y$ is a homeomorphism.

Corollary 1.6.15. *Let X be Tychonoff and $\varphi : X \to K$ a compactification.*

(1) The unique lift $\widetilde{\varphi} : \beta X \to K$ is surjective.
(2) Suppose for all $f \in C_b(X)$ there is a $g \in C(K)$ such that $f = g \circ \varphi$. Then $\widetilde{\varphi} : \beta X \to K$ is a homeomorphism.

Proof.

(1) Since $\widetilde{\varphi} \circ e_X = \varphi$ and $\varphi(X)$ is dense in K, $\widetilde{\varphi}(\beta X)$ is dense in K. But βX is compact and $\widetilde{\varphi}$ is continuous, so $\widetilde{\varphi}(\beta X)$ is compact. Since K is compact Hausdorff, compact subsets are closed, and thus $\widetilde{\varphi}(\beta X) = K$.
(2) By (1), it suffices to prove that $\widetilde{\varphi} : \beta X \to K$ is injective. Then since βX is compact and K is Hausdorff, the continuous bijection $\widetilde{\varphi}$ is automatically a homeomorphism. Injectivity follows by Remark 1.6.12. Indeed, every $f \in \Phi_X \subset C_b(X)$ factorizes as $f = g \circ \varphi$ for some $g \in \Phi_K$. □

Proposition 1.6.16. *Stone-Čech compactification is a functor β : $\mathsf{Tych} \to \mathsf{CptHsd}$.*

Proof.

id: Since
$$\beta\,\mathrm{id}_X \circ e_X \underset{(1.6.11)}{=} e_X \circ \mathrm{id}_X = e_X = \mathrm{id}_{\beta X} \circ e_X$$
we must have $\beta\,\mathrm{id}_X = \mathrm{id}_{\beta X}$ as they agree on the dense subset $X \subset \beta X$.

$-\circ-$: Suppose $f : X \to Y$ and $g : Y \to Z$ are continuous with all spaces Tychonoff. Since
$$\beta(g \circ f) \circ e_X \underset{(1.6.11)}{=} e_Z \circ g \circ f \underset{(1.6.11)}{=} \beta g \circ e_Y \circ f \underset{(1.6.11)}{=} \beta g \circ \beta f \circ e_X,$$
$\beta(g \circ f) = \beta g \circ \beta f$ as the agree on the dense subset $X \subset \beta X$. □

Exercise 1.6.17 (Adapted from Folland §4.8, #74)**.** Consider $\mathbb{N}$(with the discrete topology) as a subset of its Stone-Čech compactification $\beta\mathbb{N}$.

(1) Prove that if A, B are non-empty disjoint subsets of $\mathbb{N}$, then their closures in $\beta\mathbb{N}$ are disjoint.
(2) Suppose $(x_n) \subset \mathbb{N}$ is a sequence which is not eventually constant. Show there exist non-empty disjoint subsets $A, B \subset \mathbb{N}$ such that (x_n) is frequently in A and frequently in B.
(3) Deduce that no sequence in $\mathbb{N}$ converges in $\beta\mathbb{N}$ unless it is eventually constant.

Exercise 1.6.18 (Adapted from `http://u.cs.biu.ac.il/~tsaban/RT/Book/Chapter3.pdf`)**.** Let $\mathcal{U}\mathbb{N}$ be the set of ultrafilters on $\mathbb{N}$. For a subset $S \subset \mathbb{N}$, define $[S] := \{\mathcal{F} \in \mathcal{U}\mathbb{N} | S \in \mathcal{F}\}$. Show that the function $S \mapsto [S]$ satisfies the following properties:

(1) $[\emptyset] = \emptyset$ and $[\mathbb{N}] = \mathcal{U}\mathbb{N}$.
(2) For all $S, T \subset \mathbb{N}$,
   (a) $[S] \subset [T]$ if and only if $S \subset T$.
   (b) $[S] = [T]$ if and only if $S = T$.
   (c) $[S] \cup [T] = [S \cup T]$.
   (d) $[S] \cap [T] = [S \cap T]$.
   (e) $[S^c] = [S]^c$.
(3) Find a sequence of subsets (S_n) of $\mathbb{N}$ such that $[\bigcup S_n] \neq \bigcup [S_n]$.
(4) Find a sequence of subsets (S_n) of $\mathbb{N}$ such that $[\bigcap S_n] \neq \bigcap [S_n]$.

Exercise 1.6.19 (Adapted from `http://u.cs.biu.ac.il/~tsaban/RT/Book/Chapter3.pdf`)**.** Assume the notation of Exercise 1.6.18.

(1) Show that $\{[S] | S \subset \mathbb{N}\}$ is a base for a topology on $\mathcal{U}\mathbb{N}$.
(2) Show that all the sets $[S]$ are both closed and open in $\mathcal{U}\mathbb{N}$.
(3) Show that $\mathcal{U}\mathbb{N}$ is compact.
(4) For $n \in \mathbb{N}$, let $\mathcal{F}_n = \{S \subset \mathbb{N} | n \in S\}$. Show $\mathcal{F}_n$ is an ultrafilter on $\mathbb{N}$.
   Note: Each $\mathcal{F}_n$ is called a principal *ultrafilter on* $\mathbb{N}$.
(5) Show that $\{\mathcal{F}_n | n \in \mathbb{N}\}$ is dense in $\mathcal{U}\mathbb{N}$.
(6) Show that for every compact Hausdorff space K and every function $f : \mathbb{N} \to K$, there is a continuous function $\widetilde{f} : \mathcal{U}\mathbb{N} \to K$ such that $\widetilde{f}(\mathcal{F}_n) = f(n)$ for every $n \in \mathbb{N}$. Deduce that $\mathcal{U}\mathbb{N}$ is homeomorphic to the Stone-Čech compactification $\beta\mathbb{N}$.
   Hint: Given $f : \mathbb{N} \to K$, use Exercise 1.3.15 to get an ultrafilter on K from an ultrafilter on $\mathbb{N}$. Then use Exercises 1.3.13(4) and 1.3.19(2) to define $\widetilde{f}(\mathcal{F})$ for $\mathcal{F} \in \mathcal{U}\mathbb{N}$.

2. Measures

We begin with an informal discussion.

Definition 2.0.1. Let X be a set. A *measure* on X is a function $\mu : \mathcal{M} \to [0, \infty]$ where $\mathcal{M} \subset P(X)$ is some collection of subsets (whose properties are to be determined) satisfying:

(1) $\mu(\emptyset) = 0$
(2) $\mu(\coprod E_n) = \sum \mu(E_n)$ when (E_n) is a collection of mutually disjoint subsets in $\mathcal{M}$, where $\coprod$ means *disjoint union*.

We now would like to discuss what kind of properties the subset $\mathcal{M} \subset P(X)$ should satisfy.

- $\emptyset, X \in \mathcal{M}$ ($\mathcal{M}$ is nonempty)
- closed under disjoint unions (finite? countable?)

Example 2.0.2 (Counting measure)**.** Let $\mathcal{M} = P(X)$ and $\mu(E) := |E|$.

Example 2.0.3 (Lebesgue measure)**.** There is a measure λ on some $\mathcal{M} \subset P(\mathbb{R})$ such that

- (normalized) $\lambda([0, 1)) = 1$, and
- (translation invariant) $\lambda(E + r) = \lambda(E)$ for all $E \in \mathcal{M}$ and $r \in \mathbb{R}$.

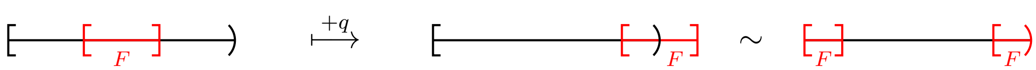
For this λ, we cannot have $\mathcal{M} = P(\mathbb{R})$! Indeed, define an equivalence relation on $[0, 1)$ by

$$x \sim y \qquad \Longleftrightarrow \qquad x - y \in \mathbb{Q}.$$

Using the Axiom of Choice, pick one representative from each equivalence class, and call this set E. For $q \in \mathbb{Q} \cap [0, 1)$, define

$$E_q := \{x + q | x \in E \cap [0, 1 - q)\} \cup \{x + q - 1 | x \in [1 - q, 1)\}.$$

Here is a cartoon of the basic idea:



Observe that there is some countable subset $Q \subset \mathbb{Q}$ such that $[0, 1) = \coprod_{q \in Q} E_q$.

Now if $\mathcal{M} = P(X)$, then we'd have

$$1 = \lambda([0, 1)) = \lambda\left(\coprod_{q \in Q} E_q\right) = \sum_{q \in Q} \lambda(E_q) = \sum_{q \in Q} \lambda(E) = \lambda(E) \sum 1 \in \{0, \infty\},$$

a contradiction.

Exercise 2.0.4. Let X be a nonempty set and $\mathcal{E} \subset P(X)$ any collection of subsets which is closed under finite unions and intersections. Suppose $\nu : P(X) \to [0, \infty]$ be a function which satisfies

- (finite additivity) for any disjoint sets $E_1, \ldots, E_n \in P(X)$, $\nu\left(\coprod_{i=1}^n E_i\right) = \sum_{i=1}^n \nu(E_i)$.

Prove that ν also has the following properties.

(1) (monotonicity) Show that if $A, B \in \mathcal{E}$ with $A \subset B$, then $\nu(A) \leq \nu(B)$.
(2) (finite subadditivity) Show that for any (not necessarily disjoint) sets $E_1, \ldots, E_n \in \mathcal{E}$, $\nu\left(\bigcup_{i=1}^n E_i\right) \leq \sum_{i=1}^n \nu(E_i)$.
(3) Show that for all $A, B \in \mathcal{E}$, $\nu(A) + \nu(B) = \nu(A \cup B) + \nu(A \cap B)$.

Exercise 2.0.5. Suppose $\mathcal{E} \subset P(\mathbb{R})$ is any collection of subsets which contains the bounded open intervals and is closed under countable unions. Let $\nu : \mathcal{E} \to [0, \infty]$ be a function which satisfies

- (monotonicity) If $E, F \in \mathcal{E}$ with $E \subset F$, then $\nu(E) \leq \nu(F)$.
- (subadditivty) for any sequence of sets $(E_n)_{n=1}^\infty \subset \mathcal{E}$, $\nu(\bigcup_{n=1}^\infty E_n) \leq \sum_{n=1}^\infty \nu(E_n)$.
- (extends length of open intervals) for all $a < b$ in $\mathbb{R}$, we have $\nu((a, b)) = b - a$.

Show that if $E \in \mathcal{E}$ is countable, then $\nu(E) = 0$.

2.1. σ-algebras.

Definition 2.1.1. A non-empty subset $\mathcal{M} \subset P(X)$ is called an *algebra* if

(1) $\mathcal{M}$ is closed under finite unions, and
(2) $\mathcal{M}$ is closed under complements.

Observe that every algebra

- contains $X = E \amalg E^c$ for some $E \in \mathcal{M}$, and thus $\emptyset = X^c$.
- is closed under finite intersections

$$\bigcap_1^k E_n = \left(\bigcap_1^k E_n\right)^{cc} = \left(\bigcup_1^k E_n^c\right)^c$$

If in addition an algebra $\mathcal{M}$ is closed under *countable* unions, then we call $\mathcal{M}$ a σ-algebra. Here, the 'σ' signifies 'countable'. We call the elements of a σ-algebra *measurable sets*.

Examples 2.1.2. Lex X be a set.

(1) $\{\emptyset, X\}$ is the *trivial* σ-algebra.
(2) $P(X)$ is the *discrete* σ-algebra.

Exercise 2.1.3. Define $\mathcal{M} := \{E \subset X | E \text{ or } E^c \text{ is countable}\}$. Show that $\mathcal{M}$ is a σ-algebra.

Exercise 2.1.4. Let X be a set. A *ring* $\mathcal{R} \subset P(X)$ is a collection of subsets of X which is closed under unions and set differences. That is, $E, F \in \mathcal{R}$ implies $E \cup F \in \mathcal{R}$ and $E \setminus F \in \mathcal{R}$.

(1) Let $\mathcal{R} \subset P(X)$ be a ring.
    (a) Prove that $\emptyset \in \mathcal{R}$.
    (b) Show that $E, F \in \mathcal{R}$ implies the symmetric difference $E \triangle F \in \mathcal{R}$.
    (c) Show that $E, F \in \mathcal{R}$ implies $E \cap F \in \mathcal{R}$.
(2) Show that any ring $\mathcal{R} \subset P(X)$ is an algebraic ring where the addition is symmetric difference and multiplication is intersection.
    (a) What is $0_\mathcal{R}$?
    (b) Show that this algebraic ring has *characteristic* 2, i.e., $E + E = 0_\mathcal{R}$ for all $E \in \mathcal{R}$.
    (c) When is the algebraic ring $\mathcal{R}$ unital? In this case, what is $1_\mathcal{R}$?
    (d) Determine the relationship (if any) between an algebra of sets in the sense of measure theory and an algebra in the algebraic sense.
    (e) Sometimes an algebra in measure theory is called a *field*. Why?

Trick. Suppose (E_n) is a sequence of subsets of X. Define

$$F_1 := E_1 \qquad F_k := E_k \setminus \bigcup_1^{k-1} E_n = E_k \cap \left(\bigcup_1^{k-1} E_n\right)^c. \tag{II}$$

Inductively, one proves (F_n) is a sequence of pairwise disjoint subsets of X such that $\bigcup E_n = \coprod F_n$. Moreover, observe that if $(E_n) \subset \mathcal{M}$ for some algebra $\mathcal{M}$, then $(F_n) \subset \mathcal{M}$.

Definition 2.1.5. Observe that if $\mathcal{M}, \mathcal{N}$ are σ-algebras, then so is $\mathcal{M} \cap \mathcal{N}$. This means if $\mathcal{E} \subset P(X)$, there is a *smallest* σ-algebra $\mathcal{M}(\mathcal{E})$ which contains $\mathcal{E}$ called the σ-algebra *generated by* $\mathcal{E}$.

Exercise 2.1.6. Let $\mathcal{A} \subset P(X)$ be an algebra. Show that the following are equivalent:

(1) $\mathcal{A}$ is a σ-algebra,
(2) $\mathcal{A}$ is closed under countable disjoint unions, and
(3) $\mathcal{A}$ is closed under countable increasing unions.

Fact 2.1.7. Suppose $\mathcal{E}, \mathcal{F} \subset P(X)$ with $\mathcal{E} \subset \mathcal{M}(\mathcal{F})$. Then $\mathcal{M}(\mathcal{E}) \subset \mathcal{M}(\mathcal{F})$.

Example 2.1.8. Suppose $(X, \mathcal{T})$ is a topological space. We call $\mathcal{B}_\mathcal{T} := \mathcal{M}(\mathcal{T})$ the *Borel* σ-algebra.

Remark 2.1.9.

- A countable intersection of open sets is called a G_δ set.
- A countable union of closed sets is called an F_σ set.
- A countable union of G_δ sets is called a $G_{\delta\sigma}$ set.
- A countable intersection of F_σ sets is called an $F_{\sigma\delta}$ set.

And so on and so forth. Observe that $\mathcal{B}_\mathcal{T}$ contains all these types of sets, so $\mathcal{B}_\mathcal{T}$ is much larger than $\mathcal{T}$.

Proposition 2.1.10. *The Borel σ-algebra $\mathcal{B}_\mathbb{R}$ on $\mathbb{R}$ generated by the usual topology (which is induced by the metric $\rho(x, y) = |x - y|$) is also generated by the following collections of sets:*

($\mathcal{B}_\mathbb{R}$1) open intervals (a, b)
($\mathcal{B}_\mathbb{R}$2) closed intervals $[a, b]$
($\mathcal{B}_\mathbb{R}$3) half-open intervals $(a, b]$
($\mathcal{B}_\mathbb{R}$4) half-open intervals $[a, b)$
($\mathcal{B}_\mathbb{R}$5) open rays (a, ∞) or $(-\infty, a)$
($\mathcal{B}_\mathbb{R}$6) closed rays $[a, \infty)$ or $(-\infty, a]$

Proof. First, observe that each of ($\mathcal{B}_\mathbb{R}$1), ($\mathcal{B}_\mathbb{R}$2), ($\mathcal{B}_\mathbb{R}$5), ($\mathcal{B}_\mathbb{R}$6) are all open or closed, so they lie in $\mathcal{B}_\mathbb{R}$. Also, $(a, b] = (a, \infty) \cap (b, \infty)^c$, so each of the sets ($\mathcal{B}_\mathbb{R}$3) are contained in $\mathcal{B}_\mathbb{R}$. Similarly for ($\mathcal{B}_\mathbb{R}$4). Hence each of ($\mathcal{B}_\mathbb{R}$1)–($\mathcal{B}_\mathbb{R}$6) lie in $\mathcal{B}_\mathbb{R}$, so their generated σ-algebras are contained in $\mathcal{B}_\mathbb{R}$ by Fact 2.1.7.

For the other directions, observe all open sets in $\mathbb{R}$ are countable unions of open intervals. (You proved this on HW1.) Hence $\mathcal{B}_\mathbb{R} \subset \mathcal{M}((\mathcal{B}_\mathbb{R}1))$ by Fact 2.1.7. For $(j) =$ ($\mathcal{B}_\mathbb{R}$2)–($\mathcal{B}_\mathbb{R}$6), one shows that ($\mathcal{B}_\mathbb{R}1$) is contained in $\mathcal{M}((j))$:

$$
\begin{aligned}
(a,b) &= \bigcup \left[a+\frac{1}{n}, b-\frac{1}{n}\right] && (\mathcal{B}_\mathbb{R}2)\\
&= \bigcup \left(a, b-\frac{1}{n}\right] && (\mathcal{B}_\mathbb{R}3)\\
&= \bigcup \left[a+\frac{1}{n}, b\right) && (\mathcal{B}_\mathbb{R}4)\\
&= (a,\infty)\cap(-\infty,b) && (\mathcal{B}_\mathbb{R}5)\\
&= ((-\infty,a]\cup[b,\infty))^c\,. && (\mathcal{B}_\mathbb{R}6)
\end{aligned}
$$

Again by Fact 2.1.7, we have $\mathcal{B}_\mathbb{R} \subset \mathcal{M}((\mathcal{B}_\mathbb{R}1)) \subset \mathcal{M}((j)) \subset \mathcal{B}_\mathbb{R}$. □

Exercise 2.1.11. Define the *h-intervals*

$$\mathcal{H} := \{\emptyset\} \cup \{(-a,b] \mid -\infty \le a < b < \infty\} \cup \{(a,\infty) \mid a \in \mathbb{R}\}\,.$$

Let $\mathcal{A}$ be the collection of finite disjoint unions of elements of $\mathcal{H}$. Show *directly from the definitions* that $\mathcal{A}$ is an algebra. Deduce that the σ-algebra $\mathcal{M}(\mathcal{A})$ generated by $\mathcal{A}$ is equal to the Borel σ-algebra $\mathcal{B}_\mathbb{R}$.

Exercise 2.1.12. Denote by $\overline{\mathbb{R}}$ the extended real numbers $[-\infty,\infty]$ with its usual topology. Prove the following assertions.

(1) The Borel σ-algebra on $\overline{\mathbb{R}}$ is generated by the open rays $(a,\infty]$ for $a \in \mathbb{R}$.
(2) If $\mathcal{E} \subset P(\mathbb{R})$ generates the Borel σ-algebra on $\mathbb{R}$, then $\mathcal{E} \cup \{\{\infty\}\}$ generates the Borel σ-algebra on $\overline{\mathbb{R}}$.

Exercise 2.1.13. Let X be a set. A π-*system* on X is a collection of subsets $\Pi \subset P(X)$ which is closed under finite intersections. A λ-*system* on X is a collection of subsets $\Lambda \subset P(X)$ such that

- $X \in \Lambda$
- Λ is closed under taking complements, and
- for every sequence of disjoint subsets (E_i) in Λ, $\bigcup E_i \in \Lambda$.

(1) Show that $\mathcal{M}$ is a σ-algebra if and only if $\mathcal{M}$ is both a π-system and a λ-system.
(2) Suppose Λ is a λ-system. Show that for every $E \in \Lambda$, the set

$$\Lambda(E) := \{F \subset X \mid F \cap E \in \Lambda\}$$

is also a λ-system.

Exercise 2.1.14 ($\pi - \lambda$ Theorem)**.** Let Π be a π-system, let Λ be the smallest λ-system containing Π, and let $\mathcal{M}$ be the smallest σ-algebra containing Π.

(1) Show that $\Lambda \subseteq \mathcal{M}$.
(2) Show that for every $E \in \Pi$, $\Pi \subset \Lambda(E)$ where $\Lambda(E)$ was defined in Exercise 2.1.13 above. Deduce that $\Lambda \subset \Lambda(E)$ for every $E \in \Pi$.
(3) Show that $\Pi \subset \Lambda(F)$ for every $F \in \Lambda$. Deduce that $\Lambda \subset \Lambda(F)$ for every $F \in \Lambda$.
(4) Deduce that Λ is a σ-algebra, and thus $\mathcal{M} = \Lambda$.

2.2. **Measures.**

Definition 2.2.1. A set X together with a σ-algebra $\mathcal{M}$ is called a *measurable space*. A *measure* on $(X, \mathcal{M})$ is a function $\mu : \mathcal{M} \to [0, \infty]$ such that

- (vacuum) $\mu(\emptyset) = 0$, and
- (countable additivity) for every sequence of disjoint sets $(E_n) \subset \mathcal{M}$, $\mu(\coprod E_n) = \sum \mu(E_n)$.

Observe that countable additivity implies finite additivity by taking all but finitely many of the E_n to be $\emptyset$.

We call the triple $(X, \mathcal{M}, \mu)$ a *measure space*. A measure space is called:

- *finite* if $\mu(X) < \infty$.
- *σ-finite* if $X = \bigcup E_n$ with $(E_n) \subset \mathcal{M}$ a sequence of measurable sets with $\mu(E_n) < \infty$. By disjointification (II), we may take such (E_n) to be disjoint.
- *semifinite* if for every $E \in \mathcal{M}$, $\mu(E) = \infty$, there is an $F \subset E$ with $F \in \mathcal{M}$ such that $0 < \mu(F) < \infty$.
- *complete* if $E \in \mathcal{M}$ with $\mu(E) = 0$ (E is *μ-null*) and $F \subset E$ implies $F \in \mathcal{M}$. *Note: We will see that $\mu(F) = 0$ by monotonicity below in (μ1) of Facts 2.2.4.*

Remark 2.2.2. In probability theory, a measure space is typically denoted $(\Omega, \mathcal{F}, P)$, and $P(\Omega) = 1$.

Examples 2.2.3.

(1) Counting measure on $P(X)$

(2) Pick $x_0 \in X$, and define μ_{x_0} on $P(X)$ by

$$\mu_{x_0}(E) = \delta_{x_0 \in E} := \begin{cases} 0 & \text{if } x_0 \notin E \\ 1 & \text{if } x_0 \in E. \end{cases}$$

We call μ_{x_0} the *point mass* or *Dirac measure* at x_0.

(3) Pick any $f : X \to [0, \infty]$. On $P(X)$, define

$$\mu_f(E) := \sum_{x \in E} f(x) := \sup \sum_{\substack{x \in F \\ F \text{ finite}}} f(x) = \lim_{\substack{\text{finite } F \\ \text{ordered by inclusion}}} \sum_{x \in F} f(x)$$

When $f = 1$, μ_f is counting measure. When $f = \delta_{x = x_0}$, we get the Dirac measure.

(4) On the σ-algebra of countable or co-countable sets, define

$$\mu(E) = \begin{cases} 0 & \text{if } E \text{ is countable} \\ 1 & \text{if } E \text{ is co-countable.} \end{cases}$$

Facts 2.2.4 (Basic properties of measures)**.** Let $(X, \mathcal{M}, \mu)$ be a measure space.

(μ1) (Monotonicity) If $E, F \in \mathcal{M}$, then $F \subset E$ implies $\mu(F) \leq \mu(E)$. In particular, if $\mu(E) = 0$, then $\mu(F) = 0$.

Proof. $\mu(E) = \mu(F \amalg (E \setminus F)) = \mu(F) + \mu(E \setminus F)$, and $\mu(E \setminus F) \geq 0$. □

($\mu 2$) (Subadditivity) If $(E_n) \subset \mathcal{M}$, then $\mu\left(\bigcup E_n\right) \leq \sum \mu(E_n)$.

Proof. Use disjointification (II). That is, setting $F_1 := E_1$ and $F_k := E_k \setminus \bigcup_1^{k-1} E_n$, we have $F_k \subset E_k$ for all k, and
$$\mu\left(\bigcup E_n\right) = \mu\left(\coprod F_n\right) = \sum \mu(F_n) \leq \sum \mu(E_n). \qquad \square$$

($\mu 3$) (Continuity from below) If $E_1 \subset E_2 \subset E_3 \subset \cdots$ is an increasing sequence of elements of $\mathcal{M}$, then
$$\mu\left(\bigcup E_n\right) = \lim_{n\to\infty} \mu(E_n).$$

Proof. Set $E_0 = \emptyset$. In this setting, disjointification (II) is easy; just set $F_n := E_n \setminus E_{n-1}$ for all $n \geq 1$. Then
$$\begin{aligned}\mu\left(\bigcup E_n\right) &= \mu\left(\coprod F_n\right) = \sum \mu(F_n) = \sum \mu(E_n \setminus E_{n-1})\\ &= \lim_{k\to\infty} \sum^k \mu(E_n \setminus E_{n-1}) = \lim_{k\to\infty} \mu(E_k). \qquad \square\end{aligned}$$

($\mu 4$) (Continuity from above) If $E_1 \supset E_2 \supset E_3 \supset \cdots$ is a decreasing sequence of elements of $\mathcal{M}$ with $\mu(E_k) < \infty$ for some $k \in \mathbb{N}$, then
$$\mu\left(\bigcap E_n\right) = \lim_{n\to\infty} \mu(E_n).$$

Proof. We may assume $\mu(E_1) < \infty$. Set $F_1 := E_1$ and $F_n := E_1 \setminus E_n$, so that $\mu(E_1) = \mu(E_n) + \mu(F_n)$ for all $n \geq 1$. Observe that
$$\bigcup F_n = \bigcup E_1 \cap E_n^c = E_1 \cap \left(\bigcup E_n^c\right) = E_1 \cap \left(\bigcap E_n\right)^c = E_1 \setminus \left(\bigcap E_n\right).$$
Hence
$$\begin{aligned}\mu\left(\bigcap E_n\right) &= \mu(E_1) - \mu\left(\bigcup F_n\right) \underset{(3)}{=} \mu(E_1) - \lim_{n\to\infty} \mu(F_n)\\ &= \mu(E_1) - \lim_{n\to\infty} (\mu(E_1) - \mu(E_n)) = \lim_{n\to\infty} \mu(E_n). \qquad \square\end{aligned}$$

Exercise 2.2.5. Suppose $(X, \mathcal{M}, \mu)$ is a measure space and $(E_n) \subset \mathcal{M}$. Recall that
$$\liminf E_n = \bigcup_k \bigcap_{n\geq k} E_n \qquad \text{and} \qquad \limsup E_n = \bigcap_k \bigcup_{n\geq k} E_n$$

(1) Prove that $\mu(\liminf E_n) \leq \liminf \mu(E_n)$.
(2) Suppose μ is finite. Prove that $\mu(\limsup E_n) \geq \limsup \mu(E_n)$.
(3) Does (2) above hold if μ is not finite? Give a proof or counterexample.

Theorem 2.2.6. *Suppose $(X, \mathcal{M}, \mu)$ is a measure space. Define*
$$\overline{\mathcal{M}} := \{E \cup F | E \in \mathcal{M} \text{ and } F \subset N \text{ for some } N \in \mathcal{M} \text{ with } \mu(N) = 0\}.$$

(1) $\overline{\mathcal{M}}$ is a σ-algebra containing $\mathcal{M}$.

(2) *There is a unique complete measure* $\overline{\mu}$ *on* $\overline{\mathcal{M}}$ *such that* $\overline{\mu}|_{\mathcal{M}} = \mu$. *We call* $\overline{\mu}$ *the* completion *of* μ.

Proof.
$\overline{\mathcal{M}}$ a σ-algebra:

(0) Observe that $\emptyset \in \mathcal{M} \subset \overline{\mathcal{M}}$, so $\overline{\mathcal{M}} \neq \emptyset$.

(1) If $(E_n \cup F_n) \subset \overline{\mathcal{M}}$, then
$$\bigcup E_n \cup F_n = \underbrace{\left(\bigcup E_n\right)}_{\in \mathcal{M}} \cup \underbrace{\left(\bigcup F_n\right)}_{\subset \bigcup N_n}.$$
Observe that each $F_n \subset N_n \in \mathcal{M}$ with $\mu(N_n) = 0$, so by countable subadditivity, we have $\mu\left(\bigcup N_n\right) \leq \sum \mu(N_n) = 0$. Hence $\overline{\mathcal{M}}$ is closed under countable unions.

(2) Suppose $E, N \in \mathcal{M}$ with $F \subset N$ μ-null. Observe that
$$\begin{aligned}(E \cup F)^c &= (E^c \cap F^c) = (E^c \cap F^c) \cap X = (E^c \cap F^c) \cap (N^c \amalg N)\\ &= (E^c \cap \underbrace{F^c \cap N^c}_{=N^c \in \mathcal{M}}) \amalg (E^c \cap F^c \cap N) = \underbrace{(E^c \cap N^c)}_{\in \mathcal{M}} \amalg \underbrace{(E^c \cap F^c \cap N)}_{\subset N}.\end{aligned}$$
Hence $\overline{\mathcal{M}}$ is closed under taking complements.

$\overline{\mu}$ unique: If $\overline{\mu}|_{\mathcal{M}} = \mu$, then for all $E \cup F \in \overline{\mathcal{M}}$ with $F \subset N$ μ-null, we have
$$\mu(E) = \overline{\mu}(E) \leq \overline{\mu}(E \cup F) \leq \overline{\mu}(E) + \overline{\mu}(F) \leq \overline{\mu}(E) + \overline{\mu}(N) = \mu(E) + \mu(N) = \mu(E).$$
Hence $\overline{\mu}(E \cup F) = \mu(E)$.

$\overline{\mu}$ exists: First, we show that $\overline{\mu}(E \cup F) := \mu(E)$ is a well-defined function on $\overline{\mathcal{M}}$. Suppose $E_1 \cup F_1 = E_2 \cup F_2$ with $F_i \subset N_i$ μ-null for $i = 1, 2$. Observe that
$$E_1 \subset E_1 \cup F_1 = E_2 \cup F_2 \subset E_2 \cup N_2 \qquad \Longrightarrow \qquad \mu(E_1) \leq \mu(E_2 \cup N_2) \leq \mu(E_2) + \mu(N_2) = \mu(E_2).$$
Swapping the roles of E_1, E_2, F_1, F_2, and N_1, N_2, we have $\mu(E_2) \leq \mu(E_1)$.

Next, we will show $\overline{\mu}$ is a measure on $\overline{\mathcal{M}}$:

(0) (Vacuum) Observe that $\overline{\mu}(\emptyset) = \mu(\emptyset) = 0$.

(1) (σ-additivity) Suppose $(E_n \cup F_n) \subset \overline{\mathcal{M}}$ is a sequence of disjoint sets with $F_n \subset N_n$ μ-null for each $n \in \mathbb{N}$. Then (E_n) and (F_n) are disjoint, and $\coprod F_n \subset \bigcup N_n$ is μ-null. Hence
$$\overline{\mu}\left(\coprod E_n \cup F_n\right) = \overline{\mu}\left(\coprod E_n \cup \coprod F_n\right) = \mu\left(\coprod E_n\right) = \sum \mu(E_n) = \sum \overline{\mu}(E_n \cup F_n).$$

$\overline{\mu}$ complete: First, note that if $F \subset N$ with N μ-null, then $F = \emptyset \cup F \in \overline{\mathcal{M}}$. Suppose $G \subset E \cup F$ where $F \subset N$ is μ-null, and $\mu(E) = 0$. Then observe $G \subset E \cup N \in \mathcal{M}$, and $\mu(E \cup N) \leq \mu(E) + \mu(N) = 0$. Hence $G \in \overline{\mathcal{M}}$. $\square$

Exercise 2.2.7. Let Π be a π-system, and let $\mathcal{M}$ be the smallest σ-algebra containing Π. Suppose μ, ν are two measures on $\mathcal{M}$ whose restrictions to Π agree.

(1) Suppose that μ, ν are finite and $\mu(X) = \nu(X)$. Show $\mu = \nu$.
Hint: Consider $\Lambda := \{E \in \mathcal{M} | \nu(E) = \mu(E)\}$.

(2) Suppose that $X = \coprod_{j=1}^{\infty} X_j$ with $(X_j) \subset \Pi$ and $\mu(X_j) = \nu(X_j) < \infty$ for all $j \in \mathbb{N}$. (Observe that μ and ν are σ-finite.) Show $\mu = \nu$.

Exercise 2.2.8 (Folland §1.3, #14 and #15)**.** Given a measure μ on $(X, \mathcal{M})$, define ν on $\mathcal{M}$ by

$$\nu(E) := \sup\{\mu(F) | F \subset E \text{ and } \mu(F) < \infty\}.$$

(1) Show that ν is a semifinite measure. We call it the *semifinite part* of μ.
(2) Suppose $E \in \mathcal{M}$ with $\nu(E) = \infty$. Show that for any $n > 0$, there is an $F \subset E$ such that $n < \nu(F) < \infty$.
This is exactly Folland §1.3, #14 applied to ν.
(3) Show that if μ is semifinite, then $\mu = \nu$.
(4) Show there is a measure ρ on $\mathcal{M}$ (which is generally not unique) which assumes only the values 0 and ∞ such that $\mu = \nu + \rho$.

Exercise 2.2.9. Suppose μ, ν are two measures on a measurable space $(X, \mathcal{M})$. We say μ is *absolutely continuous* with respect to ν if $\nu(E) = 0$ implies $\mu(E) = 0$. Prove that when μ is finite, the following are equivalent:

(1) μ is absolutely continuous with respect to ν.
(2) For every $\varepsilon > 0$, there is a $\delta > 0$ such that $E \in \mathcal{M}$ with $\nu(E) < \delta$ implies $\mu(E) < \varepsilon$.

Which direction(s) still hold if μ is infinite?

2.3. **Outer measures.**

Definition 2.3.1. Let X be a set. A function $\mu^* : P(X) \to [0, \infty]$ is called an *outer measure* if

(0) (vacuum) $\mu^*(\emptyset) = 0$.
(1) (monotonicity) $E \subset F$ implies $\mu^*(E) \le \mu^*(F)$.
(2) (countable subadditivity) $\mu^*\left(\bigcup E_n\right) \le \sum \mu^*(E_n)$ for every sequence (E_n).

Exercise 2.3.2. Suppose $(\mu_i^*)_{i \in I}$ is a family of outer measures on X. Show that

$$\mu^*(E) := \sup_{i \in I} \mu_i^*(E)$$

is an outer measure on X.

Proposition 2.3.3. *Let $\mathcal{E} \subset P(X)$ be any collection of subsets of X satisfying*

- *$\emptyset \in E$, and*
- *for all $E \subset X$, there is a sequence $(E_n) \subset \mathcal{E}$ such that $E \subset \bigcup E_n$. (Observe that if $X \in \mathcal{E}$, this condition is automatic.)*

Suppose $\rho : \mathcal{E} \to [0, \infty]$ is any function such that $\rho(\emptyset) = 0$. Then

$$\mu^*(E) := \inf\left\{\sum \rho(E_n) \middle| (E_n) \subset \mathcal{E} \text{ with } E \subset \bigcup E_n\right\} \tag{2.3.4}$$

is an outer measure, called the outer measure induced by $(\mathcal{E}, \rho)$.

Proof.

(0) Setting $E_n = \emptyset$ for all n gives $\mu^*(\emptyset) = 0$.
(1) Observe that whenever $F \subset \bigcup F_n$ with $F_n \in \mathcal{E}$ for all n, then $E \subset F \subset \bigcup F_n$. Hence the inf for E is less than or equal to the inf for F.

(2) We'll use the following two tricks:

> **Trick.** $\sum_1^\infty \frac{\varepsilon}{2^n} = \varepsilon$
>
> **Trick.** $r \leq s$ if and only if for all $\varepsilon > 0$, $r \leq s + \varepsilon$.

Suppose (E_n) is a sequence of sets and let $\varepsilon > 0$. For each n, there is a cover $(F_k^n)_k$ such that $E_n \subset \bigcup_k F_k^n$ such that

$$\sum_k \rho(F_k^n) \leq \mu^*(E_n) + \frac{\varepsilon}{2^n}.$$

Then $\bigcup E_n \subset \bigcup_n \bigcup_k F_k^n$, so

$$\mu^*\left(\bigcup E_n\right) \leq \sum_n \sum_k \rho(F_k^n) \leq \sum_n \mu^*(E_n) + \frac{\varepsilon}{2^n} = \sum \mu^*(E_n) + \sum \frac{\varepsilon}{2^n} = \sum \mu^*(E_n) + \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, $\mu^*\left(\bigcup E_n\right) \leq \sum \mu^*(E_n)$. $\square$

Exercise 2.3.5. Show that the second bullet point in Proposition 2.3.3 can be removed if we add the convention that $\inf \emptyset = \infty$.

Example 2.3.6. One can get an outer measure on $P(X)$ by taking *any* measure μ on a σ-algebra $\mathcal{M}$ and defining its induced outer measure μ^* as in (2.3.4).

We get a measure μ from an outer measure μ^* by restricting to the σ-algebra $\mathcal{M}^*$ of μ^*-measurable sets.

Definition 2.3.7. Given an outer measure μ^* on $P(X)$, we define the collection of *μ^*-measurable sets*

$$\mathcal{M}^* := \{E \subset X | \mu^*(E \cap F) + \mu^*(E^c \cap F) = \mu^*(F) \text{ for all } F \subset X\}.$$

That is, E is μ^*-measurable if it 'splits' every other set nicely with respect to μ^*.

Remarks 2.3.8.

(1) Clearly $\mu^*(F) \leq \mu^*(E \cap F) + \mu^*(E^c \cap F)$. So

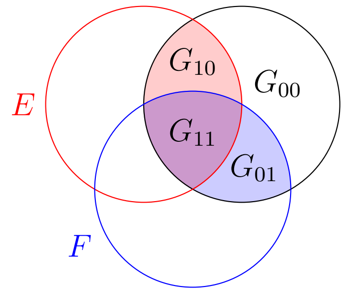
$$E \in \mathcal{M}^* \iff \mu^*(F) \geq \mu^*(E \cap F) + \mu^*(E^c \cap F) \qquad \forall F \subset X. \tag{2.3.9}$$

(2) All μ^*-null sets are in $\mathcal{M}^*$. That is, if $N \subset X$ with $\mu^*(N) = 0$, then for all $F \subset X$

$$\mu^*(\underbrace{F \cap N}_{\subset N}) + \mu^*(F \setminus N) = \mu^*(F \setminus N) \leq \mu^*(F).$$

Lemma 2.3.10. *For $G \subset X$ and $E, F \in \mathcal{M}^*$, define*

$$G_{00} := G \setminus (E \cup F)$$
$$G_{10} := G \cap (E \setminus F)$$
$$G_{01} := G \cap (F \setminus E)$$
$$G_{11} := G \cap E \cap F$$



.

Then we have

$$\mu^*(G) = \mu^*(G_{00}) + \mu^*(G_{01}) + \mu^*(G_{10}) + \mu^*(G_{11}). \tag{2.3.11}$$

Proof. Since $E \in \mathcal{M}^*$,

$$\mu^*(G) = \mu^*(G \cap E) + \mu^*(G \setminus E) = \mu^*(G_{11} \cup G_{10}) + \mu^*(G_{01} \cup G_{00}).$$

Since $F \in \mathcal{M}^*$,

$$\mu^*(G_{11} \cup G_{10}) = \mu^*(G_{11} \cup G_{10} \cap F) + \mu^*(G_{11} \cup G_{10} \setminus F) = \mu^*(G_{11}) + \mu^*(G_{10}).$$

Similarly, $\mu^*(G_{01} \cup G_{00}) = \mu^*(G_{01}) + \mu^*(G_{00})$. The result follows. □

Theorem 2.3.12 (Carathéodory)**.** *Let μ^* be an outer measure on X. The collection of μ^*-measurable sets $\mathcal{M}^*$ is a σ-algebra, and $\mu := \mu^*|_{\mathcal{M}^*}$ is a (complete) measure.*

Proof.

Step 1: $\mathcal{M}^*$ is an algebra.

(0) Clearly $\emptyset \in \mathcal{M}^*$ since it is μ^*-null by Remarks 2.3.8(2).

(1) If $E, F \in \mathcal{M}^*$, then for all $G \subset X$, (2.3.11) holds above. By applying (2.3.11) to $G_{10} \cup G_{11} \cup G_{01}$, we have

$$\mu^*((E \cup F) \cap G) = \mu^*(G_{10} \cup G_{11} \cup G_{01}) \underset{(2.3.11)}{=} \mu^*(G_{10}) + \mu^*(G_{11}) + \mu^*(G_{01}).$$

Moreover, $\mu^*((E \cup F)^c \cap G) = \mu^*(G_{00})$. Again by (2.3.11), we have

$$\mu^*((E \cup F) \cap G) + \mu^*((E \cup F)^c \cap G) = (\mu^*(G_{10}) + \mu^*(G_{11}) + \mu^*(G_{01})) + \mu^*(G_{00}) \underset{(2.3.11)}{=} \mu^*(G).$$

(2) Observe that the Carathéodory Criterion (2.3.9) is preserved under taking complements.

Step 2: $\mathcal{M}^*$ is a σ-algebra.

Suppose $(E_n) \subset \mathcal{M}^*$ is a sequence of disjoint sets, and set $E := \coprod E_n$. By Step 1, for all $N \in \mathbb{N}$, $\coprod^N E_n \in \mathcal{M}^*$. Let $F \subset X$, and define $G := F \cap \coprod^N E_n$. Then since $E_N \in \mathcal{M}^*$, we have

$$\mu^*\left(F \cap \coprod^N E_n\right) = \mu^*(G) = \mu^*(E_N^c \cap G) + \mu^*(E_N \cap G) = \mu^*\left(F \cap \coprod^{N-1} E_n\right) + \mu^*(F \cap E_N).$$

By iterating as $E_n \in \mathcal{M}^*$ for all $n \in \mathbb{N}$, we have

$$\mu^*\left(F \cap \coprod^N E_n\right) = \sum^N \mu^*(F \cap E_n) \qquad \forall N \in \mathbb{N}.$$

It follows that for all $N \in \mathbb{N}$,

$$\mu^*(F) = \mu^*\left(F \cap \coprod^N E_n\right) + \mu^*\left(\underbrace{F \setminus \coprod^N E_n}_{\supset F \setminus E}\right) \geq \sum^N \mu^*(F \cap E_n) + \mu^*(F \setminus E).$$

Taking limits in $[0,\infty]$ as $N \to \infty$, we have

$$\begin{aligned}\mu^*(F) &\geq \sum^{\infty} \mu^*(F \cap E_n) + \mu^*(F \setminus E)\\ &\geq \mu^*\left(\coprod F \cap E_n\right) + \mu^*(F \setminus E)\\ &= \mu^*(F \cap E) + \mu^*(F \setminus E).\end{aligned} \tag{2.3.13}$$

Thus $E = \coprod E_n \in \mathcal{M}^*$.

Step 3: $\mu = \mu^*|_{\mathcal{M}^*}$ is a measure.

It remains to show μ is σ-additive on $\mathcal{M}^*$. Suppose $(E_n) \subset \mathcal{M}^*$ is a sequence of disjoint sets as in Step 2. Taking $F = E$ in (2.3.13) above shows us

$$\mu^*(E) \geq \sum \mu^*(E_n) \geq \mu^*(E),$$

so equality holds. □

2.4. **Pre-measures.** In the last section, we gave a prescription for constructing a complete measure on X. Start with any collection of subsets $\mathcal{E} \subset P(X)$ with $\emptyset \in \mathcal{E}$ such that for every $E \subset X$, there is some sequence $(E_n) \subset \mathcal{E}$ with $E \subset \bigcup E_n$. Take any function $\rho : \mathcal{E} \to [0,\infty]$ such that $\rho(\emptyset) = 0$. We get an induced outer measure μ^* by (2.3.4). Taking the μ^*-measurable sets $\mathcal{M}^*$, we get a σ-algebra, and $\mu := \mu^*|_{\mathcal{M}^*}$ is a complete measure.

However, we get little control over $\mathcal{M}^*$ and μ. Consider the following two crucial questions:

(1) When is $\mathcal{E} \subset \mathcal{M}^*$?
(2) In this case, when does $\mu|_{\mathcal{E}} = \rho$?
Note: we always have $\mu^* \leq \rho$, since every $E \in \mathcal{E}$ is covered by itself. But there might be some cover $E \subset \bigcup E_n$ from $\mathcal{E}$ such that $\sum \rho(E_n) < \rho(E)$.

A sufficient condition to ensure a positive answer to both of these questions is that $\mathcal{E}$ is an algebra, and ρ is a *premeasure.*

Definition 2.4.1. Let $\mathcal{A} \subset P(X)$ be an algebra. A function $\mu_0 : \mathcal{A} \to [0,\infty]$ is called a *premeasure* if

(0) (vacuum) $\mu_0(\emptyset) = 0$, and
(1) (countable additivity) for every sequence $(E_n) \subset \mathcal{A}$ of disjoint sets such that $\coprod E_n \in \mathcal{A}$, we have $\mu_0\left(\coprod E_n\right) = \sum \mu_0(E_n)$.

The adjectives *finite*, *σ-finite*, and *semi-finite* for premeasures are defined analogously to those for measures.

Facts 2.4.2. The following are basic properties of a premeasure μ_0 on an algebra $\mathcal{A} \subset P(X)$.

(pre-μ1) (finite additivity) If $E_1, \ldots, E_n \in \mathcal{A}$ are disjoint, then $\mu_0\left(\coprod E_n\right) = \sum \mu_0(E_n)$.

Proof. If $E_1, \ldots, E_n \in \mathcal{A}$ are disjoint sets, then observe that $\coprod_{i=1}^n E_i \in \mathcal{A}$. So by setting $E_i = \emptyset$ for all $i > n$, we have

$$\mu_0\left(\coprod_{i=1}^n E_i\right) = \mu_0\left(\coprod E_i\right) = \sum \mu_0(E_i) = \sum_{i=1}^n \mu_0(E_i).$$ □

(pre-μ2) (monotonicity) If $E, F \in \mathcal{A}$ with $F \subset E$, then $\mu_0(F) \leq \mu_0(E)$.

Proof. Immediate by (pre-μ1) since $E = F \amalg (E \setminus F)$. □

(pre-μ3) (countable subadditivity) If $(E_n) \subset \mathcal{A}$ such that $\bigcup E_n \in \mathcal{A}$, then $\mu_0(\bigcup E_n) \leq \sum \mu_0(E_n)$.

Proof. We use disjointification (II). Set $F_1 := E_1$ and inductively define $F_n := E_n \setminus \bigcup_{i=1}^{n-1} E_i$. Then $F_n \in \mathcal{A}$ for all n, and $\coprod F_n = \bigcup E_n$. Thus

$$\mu_0\left(\bigcup E_n\right) = \mu_0\left(\coprod F_n\right) = \sum \mu_0(F_n) \underset{\text{(pre-}\mu 2)}{\leq} \sum \mu_0(E_n).$$ □

(pre-μ4) (monotone countable subadditivity) Suppose $E \in \mathcal{A}$ and $(E_n) \subset \mathcal{A}$ such that $E \subset \bigcup E_n$. Then $\mu_0(E) \leq \sum \mu_0(E_n)$.
Warning: This does not follow immediately by monotonicity and countable subadditivity, since we are not assured that $\bigcup E_n \in \mathcal{A}$!

Proof. Let $F_1 := E \cap E_1$ and inductively set $F_n := E \cap \left(E_n \setminus \bigcup_{i=1}^{n-1} E_i\right)$. Then $F_n \in \mathcal{A}$ for all n, and $\coprod F_n = E \in \mathcal{A}$. Hence

$$\mu_0(E) = \mu_0\left(\coprod F_n\right) = \sum \mu_0(F_n) \underset{\text{(pre-}\mu 2)}{\leq} \sum \mu_0(E_n).$$ □

Remark 2.4.3. Recall that if μ_0 is only known to be finitely additive and not necessarily countably additive, then μ_0 still satisfies monotonicity and finite subadditivity (cf. Exercise 2.0.4).

Lemma 2.4.4. *Suppose μ_0 is a premeasure on $\mathcal{A}$. Let μ^* be the induced outer measure given by (2.3.4).*

(1) $\mu^*|_{\mathcal{A}} = \mu_0$, and
(2) $\mathcal{A} \subset \mathcal{M}^*$.

Proof.
(1) Suppose $E \in \mathcal{A}$.

$\underline{\mu^* \leq \mu_0}$: Setting $E_1 := E$ and $E_n := \emptyset$ for all $n > 1$, $\mu^*(E) \leq \sum \mu_0(E_n) = \mu_0(E)$.

$\underline{\mu^* \geq \mu_0}$: Let $\varepsilon > 0$. By definition of μ^* as an infimum, there is a sequence $(E_n) \subset A$ such that $E \subset \bigcup E_n$ and $\sum \mu_0(E_n) \leq \mu^*(E) + \varepsilon$. But by monotone countable subadditivity, $\mu_0(E) \leq \sum \mu_0(E_n)$, and thus $\mu_0(E) \leq \mu^*(E) + \varepsilon$. Since $\varepsilon > 0$ was arbitrary, $\mu_0 \leq \mu^*$ on $\mathcal{A}$.

(2) Suppose $E \in \mathcal{A}$ and $F \subset X$ and $\varepsilon > 0$. Pick $(F_n) \subset \mathcal{A}$ such that $F \subset \bigcup F_n$ and $\sum \mu_0(F_n) \leq \mu^*(F) + \varepsilon$. Since μ_0 is σ-additive on $\mathcal{A}$,

$$\begin{aligned} \mu^*(F) + \varepsilon &\geq \sum \mu_0(F_n) = \sum \mu_0(F_n \cap E) + \mu_0(F_n \cap E^c) \\ &= \sum \mu_0(F_n \cap E) + \sum \mu_0(F_n \cap E^c) \\ &\geq \mu^*(F \cap E) + \mu^*(F \cap E^c). \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, we conclude that $\mu^*(F) \geq \mu^*(F \cap E) + \mu^*(F \cap E^c)$, and thus $E \in \mathcal{M}^*$. □

Construction 2.4.5. Starting with a premeasure μ_0 on an algebra $\mathcal{A}$, we get a σ-algebra $\mathcal{M}^*$ which contains $\mathcal{A}$, and a complete measure $\mu := \mu^*|_{\mathcal{M}^*}$ such that $\mu|_{\mathcal{A}} = \mu_0$.

Remark 2.4.6. Observe that by Fact 2.1.7, $\mathcal{M}^*$ contains $\mathcal{M} := \mathcal{M}(\mathcal{A})$, the σ-algebra generated by $\mathcal{A}$, and $\mu|_{\mathcal{M}}$ is a (possibly non-complete) measure.

Theorem 2.4.7. *Suppose μ_0 is a premeasure on an algebra $\mathcal{A}$, and μ is the measure on $\mathcal{M}^*$ from Construction 2.4.5. If ν is a measure on $\mathcal{M} = \mathcal{M}(\mathcal{A})$ such that $\nu|_{\mathcal{A}} = \mu_0$, then $\nu(E) \leq \mu(E)$ for all $E \in \mathcal{M}$, with equality when $\mu(E) < \infty$.*

Proof. Suppose $E \in \mathcal{M}$.
Step 1: $\nu(E) \leq \mu(E)$.

Since $E \in \mathcal{M}$, for all sequences $(E_n) \subset \mathcal{A}$ such that $E \subset \bigcup E_n$,

$$\nu(E) \leq \sum \nu(E_n) = \sum \mu_0(E_n).$$

Hence $\nu(E) \leq \inf\{\sum \mu_0(E_n) | E \subset \bigcup E_n\} = \mu^*(E) = \mu(E)$.

Step 2: When $\mu(E) < \infty$, we show $\mu(E) \leq \nu(E)$, and thus $\mu(E) = \nu(E)$.

Let $\varepsilon > 0$. Then there exists a sequence $(E_n) \subset \mathcal{A}$ such that $E \subset \bigcup E_n$ and

$$\mu\left(\bigcup E_n\right) \leq \sum \mu_0(E_n) \leq \mu(E) + \varepsilon < \infty.$$

Since $E \subset \bigcup E_n$ and $\mu(E) < \infty$, we have

$$\mu\left(\left(\bigcup E_n\right) \setminus E\right) = \mu\left(\bigcup E_n\right) - \mu(E) \leq \varepsilon. \tag{2.4.8}$$

Now by continuity from below (μ3) for both μ and ν, we have

$$\begin{aligned} \mu\left(\bigcup E_n\right) &= \lim_{N\to\infty} \mu\left(\bigcup^N E_n\right) = \lim_{N\to\infty} \mu_0\left(\bigcup^N E_n\right) \\ &= \lim_{N\to\infty} \nu\left(\bigcup^N E_n\right) = \nu\left(\bigcup E_n\right). \end{aligned} \tag{2.4.9}$$

Putting these two equations together, we have

$$\begin{aligned} \mu(E) \leq \mu\left(\bigcup E_n\right) &\underset{(2.4.9)}{=} \nu\left(\bigcup E_n\right) = \nu(E) + \nu\left(\left(\bigcup E_n\right) \setminus E\right) \\ &\leq \nu(E) + \mu\left(\left(\bigcup E_n\right) \setminus E\right) \underset{(2.4.8)}{\leq} \nu(E) + \varepsilon \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, $\mu(E) \leq \nu(E)$.

This concludes the proof. □

Corollary 2.4.10. *Suppose μ_0 is a premeasure on an algebra $\mathcal{A}$, and μ is the measure on $\mathcal{M}^*$ from Construction 2.4.5. If μ_0 is σ-finite, then μ is the unique extension of μ_0 to $\mathcal{M} = \mathcal{M}(\mathcal{A})$.*

Proof. Recall that μ_0 is σ-finite if there exists a sequence $(E_n) \subset \mathcal{A}$ such that $\bigcup E_n = X$ and $\mu_0(E_n) < \infty$ for all n. Observe that by disjointification (II), we may assume that the E_n are disjoint.

Now for any other ν extending μ_0 and $E \in \mathcal{M}$, we have

$$\mu(E) = \mu\left(\coprod E \cap E_n\right) = \sum \underbrace{\mu(E \cap E_n)}_{<\infty} = \sum \nu(E \cap E_n) = \nu\left(\coprod E \cap E_n\right) = \nu(E). \quad \square$$

Exercise 2.4.11. Suppose $\mathcal{A}$ is an algebra on X, μ_0 a premeasure on $\mathcal{A}$, and μ^* the induced outer measure on $P(X)$ given by (2.3.4). Show that for every $E \subset X$, there is a μ^*-measurable set $F \supset E$ such that $\mu^*(F) = \mu^*(E)$.

Exercise 2.4.12 (Adapted from Folland §1.4, #18 and #22)**.** Suppose $\mathcal{A}$ is an algebra, and let $\mathcal{M}$ be the σ-algebra generated by $\mathcal{A}$. Let μ_0 be a σ-finite premeasure on $\mathcal{A}$, μ^* the induced outer measure given by (2.3.4), and $\mathcal{M}^*$ the σ-algebra of μ^*-measurable sets. Show that the following are equivalent.

(1) $E \in \mathcal{M}^*$
(2) $E = F \setminus N$ where $F \in \mathcal{M}$ and $\mu^*(N) = 0$.
(3) $E = F \cup N$ where $F \in \mathcal{M}$ and $\mu^*(N) = 0$.

Deduce that if μ is a σ-finite measure on $\mathcal{M}$, then $\mu^*|_{\mathcal{M}^*}$ on $\mathcal{M}^*$ is the completion of μ on $\mathcal{M}$.

Exercise 2.4.13 (Folland §1.4, #20)**.** Let μ^* be an outer measure on $P(X)$, $\mathcal{M}^*$ the σ-algebra of μ^*-measurable sets, and $\mu := \mu^*|_{\mathcal{M}^*}$. Let μ^+ be the outer measure on $P(X)$ induced by the (pre)measure μ on the (σ-)algebra $\mathcal{M}^*$.

(1) Show that $\mu^*(E) \leq \mu^+(E)$ for all $E \subset X$ with equality if and only if there is an $F \in \mathcal{M}^*$ with $E \subset F$ and $\mu^*(E) = \mu^*(F)$.
(2) Show that if μ^* was induced from a premeasure μ_0 on an algebra $\mathcal{A}$, then $\mu^* = \mu^+$.
(3) Construct an outer measure μ^* on the two point set $X = \{0, 1\}$ such that $\mu^* \neq \mu^+$.

Exercise 2.4.14. Let X be a set, $\mathcal{A}$ an algebra on X, μ_0 a premeasure on $\mathcal{A}$, and μ^* the induced outer measure on $P(X)$ given by (2.3.4). Suppose that (E_n) is an *increasing* sequence of subsets of X, i.e., $E_1 \subset E_2 \subset E_3 \subset \cdots$. Prove that

$$\mu^*\left(\bigcup_{n=1}^{\infty} E_n\right) = \lim_{n\to\infty} \mu^*(E_n).$$

Exercise 2.4.15 (Sarason)**.** Suppose μ_0 is a finite premeasure on the algebra $\mathcal{A} \subset P(X)$, and let $\mu^* : P(X) \to [0, \infty]$ be the outer measure induced by μ_0. Prove that the following are equivalent for $E \subset X$.

(1) $E \in \mathcal{M}^*$, the μ^*-measurable sets.
(2) $\mu^*(E) + \mu^*(X \setminus E) = \mu(X)$.

Hint: Use Exercise 2.4.12.

2.5. **Lebesgue-Stieltjes measures on $\mathbb{R}$.**

2.5.1. *Construction of Lebesgue-Stieltjes measures.* Recall from Exercise 2.1.11 that we define the collection of *h-intervals* by

$$\mathcal{H} := \{\emptyset\} \cup \{(a,b] | -\infty \leq a < b < \infty\} \cup \{(a,\infty) | a \in \mathbb{R}\}.$$

Let $\mathcal{A} = \mathcal{A}(\mathcal{H})$ be the collection of finite disjoint unions of elements of $\mathcal{H}$. By Exercise 2.1.11, $\mathcal{A}$ is an algebra, and the σ-algebra generated by $\mathcal{A}$ is $\mathcal{M}(\mathcal{A}) = \mathcal{B}_{\mathbb{R}}$, the Borel σ-algebra. Our goal is to construct a nice class of premeasures on $\mathcal{A}$.

Construction 2.5.1. Let $F : \mathbb{R} \to \mathbb{R}$ be any function which is

- (non-decreasing) $r \leq s$ implies $F(r) \leq F(s)$, and
- (right continuous) if $r_n \searrow a$, then $F(r_n) \searrow F(a)$

Extend F to a function $\overline{\mathbb{R}} = [-\infty, \infty] \to \overline{\mathbb{R}}$ by

$$F(-\infty) := \lim_{a\to-\infty} F(a) \qquad \text{and} \qquad F(\infty) := \lim_{b\to\infty} F(b).$$

Define $\mu_0 : \mathcal{H} \to [0, \infty]$ by

- $\mu_0(\emptyset) := 0$,
- $\mu_0((a, b]) := F(b) - F(a)$ for all $a \geq -\infty$, and
- $\mu_0((a,\infty)) := F(\infty) - F(a)$ for all $a \geq -\infty$.

In (LS4) below, we extend $\mu_0 : \mathcal{H} \to [0, \infty]$ to a well-defined function $\mathcal{A} = \mathcal{A}(\mathcal{H}) \to [0, \infty]$. In Theorem 2.5.7 below, we prove this extension to $\mathcal{A}$ is a *premeasure*. By Carathéodory's outer measure construction, we get an outer measure μ^*_F on $(\mathbb{R}, P(\mathbb{R}))$ by (2.3.4). By taking the σ-algebra of μ^*_F-measurable sets $\mathcal{M}_F := \mathcal{M}^*$, we get a complete measure $\mu_F := \mu^*_F|_{\mathcal{M}_F}$.

Definition 2.5.2. We call μ_F the *Lebesgue-Stieltjes measure* associated to F.

Remark 2.5.3. Since μ_F is σ-finite by construction, it follows from Exercise 2.4.12 that $\mathcal{M}_F$ is the completion $\overline{\mathcal{B}_{\mathbb{R}}}$ of the Borel σ-algebra for $\mu_F|_{\mathcal{B}_{\mathbb{R}}}$. Thus, sets in $\mathcal{M}_F$ are unions of Borel sets and subsets of Borel sets which are μ_F-null.

In the remainder of this section, we prove that μ_0 extends to a premeasure on $\mathcal{A} = \mathcal{A}(\mathcal{H})$.

Facts 2.5.4. We have the following facts about the function μ_0.

(LS1) Splitting $(a,\infty) = (a, b] \amalg (b,\infty)$, we have $\mu_0((a,\infty)) = \mu_0((a, b]) + \mu_0((b,\infty))$.

(LS2) If $(a, b] = \coprod_{i=1}^n (a_i, b_i]$, then $\mu_0((a, b]) = \sum_{i=1}^n \mu_0((a_i, b_i])$.

> *Proof.* Re-indexing, we may assume $a = a_1 < b_1 = a_2 < b_2 = a_3 < \cdots < b_n$. Then
>
> $$\mu_0((a, b]) = F(b) - F(a) = \sum_{i=1}^n F(b_i) - F(a_i) = \sum_{i=1}^n \mu_0((a_i, b_i]).$$ □

(LS3) If $E_1, \ldots, E_n \in \mathcal{H}$ are disjoint and $F \in \mathcal{H}$ such that $F \subset \coprod_{i=1}^n E_i$, then $\mu_0(F) = \sum_{i=1}^n \mu_0(F \cap E_i)$.

> *Proof.* Removing elements of $(E_i)_{i=1}^n$ if necessary, we may assume that $F \cap E_i \neq \emptyset$ for all $i = 1, \ldots, n$. This means that $F \cap E_i \in \mathcal{H}$ for all i, and $F = \coprod_{i=1}^n F \cap E_i$. The result now follows by (LS1) and (LS2). □

(LS4) If $(E_1,\dots,E_m)\subset\mathcal{H}$ and $(F_1,\dots,F_n)\subset\mathcal{H}$ are two collections of disjoint h-intervals with $\coprod_{i=1}^m E_i=\coprod_{j=1}^n F_j$, then $\sum_{i=1}^m \mu_0(E_i)=\sum_{j=1}^n \mu_0(F_j)$.

Proof. By applying (LS3) twice, we have

$$\sum_{i=1}^m \mu_0(E_i) \underset{(3)}{=} \sum_{i=1}^m\sum_{j=1}^n \mu_0(E_i\cap F_j) = \sum_{j=1}^n\sum_{i=1}^m \mu_0(E_i\cap F_j) \underset{(3)}{=} \sum_{j=1}^n \mu_0(F_j). \qquad \square$$

Hence μ_0 extends to a well-defined function still denoted $\mu_0:\mathcal{A}=\mathcal{A}(\mathcal{H})\to[0,\infty]$ by

$$\mu_0\left(\coprod_{i=1}^n E_i\right) := \sum_{i=1}^n \mu_0(E_i) \qquad\qquad \forall \text{ disjoint } E_1,\dots,E_n\in\mathcal{H}.$$

Corollary 2.5.5. *The extension $\mu_0:\mathcal{A}\to[0,\infty]$ afforded by (LS4) is finitely additive and thus monotone and finitely subadditive by Exercise 2.0.4.*

Proof. Suppose $E=\coprod_{i=1}^n E_i$ with $E,E_1,\dots,E_n\in\mathcal{A}$. Then we may write each $E_i=\coprod_{j=1}^{m_i} E_j^i$ where $E_j^i\in\mathcal{H}$ for all $j=1,\dots,m_i$, and thus $E=\coprod_{i=1}^n\coprod_{j=1}^{m_i} E_j^i$. Then by countable additivity of μ_0 on $\mathcal{H}$ from (LS4), we have

$$\mu_0(E)=\sum_{i=1}^n\sum_{j=1}^{m_i}\mu_0(E_j^i)=\sum_{i=1}^n \mu_0(E_i). \qquad \square$$

Exercise 2.5.6. Describe to the best of your ability the set of accumulation points of right endpoints (b_j) for a disjoint collection of bounded h-intervals $((a_n,b_n])_{n=1}^\infty$ such that $\coprod(a_n,b_n]=(a,b]$ for some $a<b$ in $\mathbb{R}$.

Theorem 2.5.7. *The extension $\mu_0:\mathcal{A}\to[0,\infty]$ afforded by (LS4) is a premeasure on $\mathcal{A}$.*

Proof. It is clear that $\mu_0(\emptyset)=0$ by construction.

Suppose $(E_n)\subset\mathcal{A}$ is a disjoint sequence such that $\coprod E_n\in\mathcal{A}$. Then there are disjoint h-intervals $F_1,\dots,F_k\in\mathcal{H}$ such that $\coprod E_n=\coprod_{j=1}^k F_j$. We may assume that $E_n\cap F_j\neq\emptyset$ for at most one j. Thus we may partition the (E_n) into (E_n^j) such that $\coprod E_n^j=F_j$ for $j=1,\dots,k$. We make the following claim.

Claim. *Suppose $H\in\mathcal{H}$ is a single h-interval such that $H=\coprod H_n$ where $(H_n)\subset\mathcal{H}$ is a sequence of disjoint h-intervals. Then $\mu_0(H)=\sum\mu_0(H_n)$.*

Then by applying (LS4), we have

$$\mu_0\left(\coprod E_n\right)=\mu_0\left(\coprod_{j=1}^k F_j\right)=\sum_{j=1}^k\mu_0(F_j) \underset{\text{(Claim)}}{=} \sum_{j=1}^k\sum\mu_0(E_n^j)=\sum\mu_0(E_n).$$

Thus it remains to prove the claim.

Proof of claim for $H = (a, b]$, $a, b \in \mathbb{R}$. Suppose $(a, b] = \coprod (a_j, b_j]$. Then for all $n \in \mathbb{N}$, $\coprod_{j=1}^{n} (a_j, b_j] \subset (a, b]$. By (LS4) and monotonicity, we have

$$\sum_{j=1}^{n} \mu_0((a_j, b]) = \mu_0\left(\coprod_{j=1}^{n} (a_j, b_j]\right) \leq \mu_0((a, b]).$$

Taking $n \to \infty$, we have $\sum \mu_0((a_j, b_j]) \leq \mu_0((a, b])$.
To show the reverse inequality, let $\varepsilon > 0$. Since F is right continuous,

- there is $\delta > 0$ such that $F(a + \delta) - F(a) < \frac{\varepsilon}{2}$, and
- for all $j \geq 1$, there is $\delta_j > 0$ such that $F(b_j + \delta_j) - F(b_j) < \frac{\varepsilon}{2^{j+1}}$.

Observe now that $\{(a_j, b_j + \delta_j)\}_{j=1}^{\infty}$ is an open cover of the compact interval $[a + \delta, b]$. Hence there is a finite subcover, i.e., there is an $N \in \mathbb{N}$ such that $[a + \delta, b] \subset \bigcup_{j=1}^{N} (a_j, b_j + \delta_j)$. Then we calculate

$$\begin{aligned}
\mu_0((a, b]) &= F(b) - F(a) \\
&< F(b) - F(a + \delta) + \frac{\varepsilon}{2} \\
&= \mu_0((a + \delta, b]) + \frac{\varepsilon}{2} \\
&\leq \mu_0\left(\bigcup_{j=1}^{N} (a_j, b_j + \delta_j]\right) + \frac{\varepsilon}{2} \\
&\leq \sum_{j=1}^{N} \mu_0((a_j, b_j + \delta_j]) + \frac{\varepsilon}{2} \\
&= \sum_{j=1}^{N} (F(b_j + \delta_j) - F(a_j)) + \frac{\varepsilon}{2} \\
&\leq \sum_{j=1}^{N} \left(F(b_j) + \frac{\varepsilon}{2^{j+1}} - F(a_j)\right) + \frac{\varepsilon}{2} \\
&= \sum_{j=1}^{N} \mu_0((a_j, b_j]) + \sum_{j=1}^{N} \frac{\varepsilon}{2^{j+1}} + \frac{\varepsilon}{2} \\
&\leq \sum_{j=1}^{\infty} \mu_0((a_j, b_j]) + \sum_{j=1}^{\infty} \frac{\varepsilon}{2^{j+1}} + \frac{\varepsilon}{2} \\
&= \sum_{j=1}^{\infty} \mu_0((a_j, b_j]) + \varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, $\mu_0((a, b]) \leq \sum_{j=1}^{\infty} \mu_0((a_j, b_j])$. □

The cases $H = (-\infty, b]$ for some $b < \infty$ and $H = (a, \infty)$ for $-\infty \leq a$ are left as the following exercise. □

Exercise 2.5.8. Consider the extension $\mu_0 : \mathcal{A} \to [0, \infty]$ afforded by (LS4). Suppose H is $(-\infty, b]$ for some $b < \infty$ or (a, ∞) for $-\infty \le a$. If $H = \coprod H_n$ where $(H_n) \subset \mathcal{H}$ is a sequence of disjoint h-intervals, then $\mu_0(H) = \sum \mu_0(H_n)$.

Exercise 2.5.9 (Folland, §1.5, #28)**.** Let $F : \mathbb{R} \to \mathbb{R}$ be increasing and right continuous, and let μ_F be the associated Lebesgue-Stieltjes Borel measure on $\mathcal{B}_\mathbb{R}$. For $a \in \mathbb{R}$, define

$$F(a-) := \lim_{r \nearrow a} F(r).$$

Prove that:

(1) $\mu_F(\{a\}) = F(a) - F(a-)$,
(2) $\mu_F([a, b)) = F(b-) - F(a-)$,
(3) $\mu_F([a, b]) = F(b) - F(a-)$, and
(4) $\mu_F((a, b)) = F(b-) - F(a)$.

2.5.2. *Lebesgue measure.*

Definition 2.5.10. *Lebesgue measure* λ is the Lebesgue-Stieltjes measure μ_{id} where id : $\mathbb{R} \to \mathbb{R}$ is the identity function $\text{id}(r) = r$. The *Lebesgue σ-algebra* is $\mathcal{L} := \mathcal{M}^* = \overline{\mathcal{B}_\mathbb{R}}$ for $\lambda|_{\mathcal{B}_\mathbb{R}}$.

Definition 2.5.11. For $E \subset \mathbb{R}$ and $r, s \in \mathbb{R}$, define $rE := \{rx | x \in E\}$ and $s + E := \{s + x | x \in E\}$.

Theorem 2.5.12. *Suppose $E \in \mathcal{L}$.*

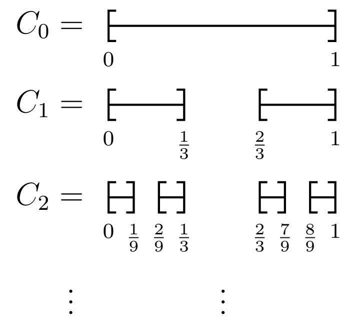
(1) (dilation homogeneity) If $r \in \mathbb{R}$, then $rE \in \mathcal{L}$ and $\lambda(rE) = |r| \cdot \lambda(E)$.
(2) (translation invariance) If $s \in \mathbb{R}$, then $s + E \in \mathcal{L}$ and $\lambda(s + E) = \lambda(E)$.

Proof. We will prove dilation homogeneity and leave translation invariance to the reader.
Step 1: $\mathcal{B}_\mathbb{R}$ is closed under $E \mapsto rE$. This is trivial if $r = 0$, so assume $r \neq 0$. Then multiplication by r is a bijection on $P(\mathbb{R})$ mapping open intervals to open intervals. Thus multiplication by r maps $\mathcal{B}_\mathbb{R}$ onto itself.
Step 2: It is a straightforward exercise to prove that $|r| \cdot \lambda$ is a measure on $\mathcal{L}$ and $\lambda^r(E) := \lambda(rE)$ is a measure on $\mathcal{B}_\mathbb{R}$.
Step 3: If $E \in \mathcal{H}$, then $\lambda^r(E) = |r| \cdot \lambda(E)$, so $\lambda^r = |r| \cdot \lambda$ on $\mathcal{A}(\mathcal{H})$ and thus all of $\mathcal{B}_\mathbb{R}$ by Corollary 2.4.10 (or Exercise 2.2.7) as λ^r and $|r| \cdot \lambda$ are both σ-finite.
Step 4: If $E \in \mathcal{L}$ is λ-null, then $rE \in \mathcal{L}$ is λ-null. Indeed, by Remark 2.5.3, $E \in \mathcal{L}$ is λ-null if and only if there is an $N \in \mathcal{B}_\mathbb{R}$ such that $E \subset N$ and $\lambda(N) = 0$. Now $rE \subset rN$, and $\lambda(rN) = |r| \cdot \lambda(N) = 0$ by Step 3.
Step 5: Finally, as $\mathcal{L} = \overline{\mathcal{B}_\mathbb{R}}$ for λ, we see λ^r and $|r| \cdot \lambda$ are both defined on $\mathcal{L}$ and agree. □

Exercise 2.5.13. Let $\mathcal{B}_\mathbb{R}$ be the Borel σ-algebra of $\mathbb{R}$. Suppose μ is a translation invariant measure on $\mathcal{B}_\mathbb{R}$ such that $\mu((0, 1]) = 1$. Prove that $\mu = \lambda|_{\mathcal{B}_\mathbb{R}}$, the restriction of Lebesgue measure on $\mathcal{L}$ to $\mathcal{B}_\mathbb{R}$.

Remark 2.5.14. By Exercise 2.5.9(1), $\lambda(\{r\}) = 0$ for all $r \in \mathbb{R}$, and thus $\lambda(E) = 0$ for all countable $E \subset \mathbb{R}$.

Example 2.5.15. The Cantor set C is defined as $\bigcap C_n$ where we define C_n inductively by 'removing middle thirds' of $[0, 1]$.



By continuity from above ($\mu 4$) for λ, we have $\lambda(C) = \lim_{n\to\infty} \lambda(C_n)$. By Exercise 2.5.9,

$$\begin{aligned}
&\lambda(C_0) = 1\\
&\lambda(C_1) = 1 - \frac{1}{3}\\
&\lambda(C_2) = 1 - \frac{1}{3} - \frac{2}{9}\\
&\lambda(C_3) = 1 - \frac{1}{3} - \frac{2}{9} - \frac{4}{27} \qquad \text{etc.}\\
\Longrightarrow \qquad &\lambda(C) = 1 - \sum_{n=1}^{\infty} \frac{2^{n-1}}{3^n} = 1 - \frac{1}{3}\sum_{n=0}^{\infty}\left(\frac{2}{3}\right)^n = 1 - \frac{1}{3}\left(\frac{1}{1-\frac{2}{3}}\right) = 0.
\end{aligned}$$

It is well known that C is uncountable; indeed it is in bijection with $\{0, 1\}^{\mathbb{N}}$ via base 3 decimal expansions where only the digits 0 and 2 occur. (Recall that decimal expansion is not unique; one must pick a particular convention here.)

Exercise 2.5.16. Show that the function $f : \{0, 1\}^{\mathbb{N}} \to C$ given by

$$f(x) := \sum_{n=1}^{\infty} \frac{2x_n}{3^n}$$

is a homeomorphism of $\{0, 1\}^{\mathbb{N}}$ onto the Cantor set.

Exercise 2.5.17. Suppose $E \in \mathcal{L}$ with $\lambda(E) > 0$. Show there is an $F \subset E$ such that $F \notin \mathcal{L}$. That is, show any Lebesgue measurable set with positive measure contains a non-measurable subset.

Exercise 2.5.18 (Sarason)**.** Suppose $E \in \mathcal{L}$ is Lebesgue null, and $\varphi : \mathbb{R} \to \mathbb{R}$ is a C^1 function (continuous with continuous derivative). Prove that $\varphi(E)$ is also Lebesgue null.

Exercise 2.5.19. Let (X, ρ) be a metric (or simply a topological) space. A subset $S \subset X$ is called *nowhere dense* if $\overline{S}$ does not contain any open set in X. A subset $T \subset X$ is called *meager* if it is a countable union of nowhere dense sets.

Construct a meager subset of $\mathbb{R}$ whose complement is Lebesgue null.

Exercise 2.5.20. Suppose $F : \mathbb{R} \to \mathbb{R}$ is a bounded, non-decreasing, right continuous function, and let μ_F be the corresponding Lebesgue-Stieltjes measure. (Observe μ_F is finite.) Prove the following are equivalent:

(1) μ_F is absolutely continuous (see Exercise 2.2.9) with respect to Lebesgue measure λ.
(2) F is *absolutely continuous*, i.e., for every $\varepsilon > 0$, there exists $\delta > 0$ such that for any disjoint set of open intervals $(a_1, b_1), \ldots, (a_N, b_N)$,

$$\sum_{i=1}^{n}(b_i - a_i) < \delta \qquad \Longrightarrow \qquad \sum_{i=1}^{N}(F(b_i) - F(a_i)) < \varepsilon.$$

2.5.3. *Regularity properties of Lebesgue-Stieltjes measures.*

Definition 2.5.21. Suppose $(X, \mathcal{T})$ is a Hausdorff topological space and $\mathcal{M} \subset P(X)$ is any σ-algebra containing the Borel σ-algebra $\mathcal{B}(\mathcal{T})$, i.e., $\mathcal{T} \subset \mathcal{M}$. A measure μ on $\mathcal{M}$ is called:

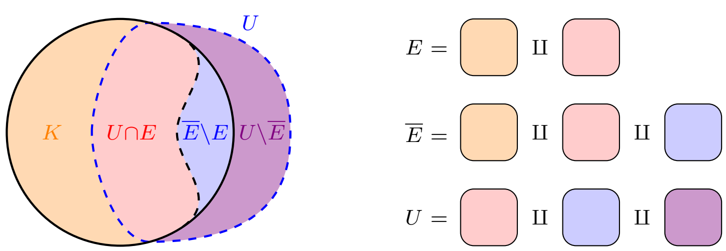
- *outer regular* if $\mu(E) = \inf\{\mu(U) | E \subset U \text{ open}\}$
- *inner regular* if $\mu(E) = \sup\{\mu(K) | \text{compact } K \subset E\}$
- *regular* if μ is both outer and inner regular.

Proposition 2.5.22. *Suppose $(X, \mathcal{T})$ is a Hausdorff topological space and μ is a measure on any σ-algebra $\mathcal{M}$ containing $\mathcal{B}_\mathcal{T}$. If $(X, \mathcal{T})$ is σ-compact and μ is outer regular and finite on compact sets, then μ is inner regular and thus regular (and thus Radon; see Exercise 2.5.24 below).*

Proof.
Step 1: Suppose X is compact and $E \in \mathcal{B}_\mathcal{T}$. Then $\overline{E}$ is compact. Let $\varepsilon > 0$. By outer regularity, there is an open $U \supset \overline{E} \setminus E$ such that $\mu(U) \leq \mu(\overline{E} \setminus E) + \varepsilon$. Observe that:

- $\overline{E} \setminus E \subset U \setminus E$,
- $K := \overline{E} \setminus U$ is compact and contained in E, and
- since $\overline{E} = K \amalg (U \cap \overline{E})$ and $E \subset \overline{E}$, $E = (K \cap E) \amalg (U \cap E)$, and thus $U \cap E = K^c \cap E$.

Here is a cartoon of $K, E, \overline{E}, U$:



We now calculate

$$\begin{aligned}
\mu(K) &= \mu(E) - \mu(K^c \cap E) && (E = K \amalg (K^c \cap E))\\
&= \mu(E) - \mu(U \cap E) && (E \cap U = E \cap K^c)\\
&= \mu(E) - (\mu(U) - \mu(U \setminus E)) && (U = (E \cap U) \amalg (U \setminus E))\\
&\geq \mu(E) \underbrace{-\mu(U) + \mu(\overline{E} \setminus E)}_{\geq -\varepsilon} && (\overline{E} \setminus E \subset U \setminus E)\\
&\geq \mu(E) - \varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, μ is inner regular.
Step 2: Since X is σ-compact, by disjointification, we may write $X = \coprod X_n$ where each X_n has compact closure in X. In particular, $\mu(X_n) < \infty$ for all n. Let $E \in \mathcal{B}_\mathcal{T}$, and write $E = \coprod E_n$

where $E_n := E \cap X_n$. By Step 1, for each n, there is a compact set $K_n \subset E_n \subset X_n \subset \overline{X_n}$ such that $\mu(K_n) \geq \mu(E_n) - \frac{\varepsilon}{2^{n+1}}$. Set $F_n := \coprod_{i=1}^n K_i$, which is still compact. Observe that

$$\mu(F_n) \geq \mu\left(\coprod_{i=1}^n E_i\right) - \frac{\varepsilon}{2}.$$

There are two cases to consider now.

If $\mu(E) = \infty$, since $\mu(\coprod_{i=1}^n E_i) \nearrow \infty$, eventually $\mu(F_n) > M$ for every $M > 0$. Hence $\sup\{\mu(F_n) | n \in \mathbb{N}\} = \infty = \mu(E)$. Otherwise, $\mu(E) < \infty$, and there is an $N \in \mathbb{N}$ such that

$$\mu(E) \leq \mu\left(\coprod_{i=1}^N E_i\right) + \frac{\varepsilon}{2} \leq \mu(F_N) + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \mu(F_N) + \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, we conclude μ is inner regular. □

Exercise 2.5.23. Suppose $(X, \mathcal{T})$ is a topological space, μ is a σ-finite regular Borel measure, and $E \in \mathcal{B}_\mathcal{T}$ is a Borel set. Prove the following assertions.

(1) For every $\varepsilon > 0$, there exist an open U and a closed F with $F \subset E \subset U$ and $\mu(U \setminus F) < \varepsilon$.
(2) There exist an F_σ-set A and a G_δ-set B such that $A \subset E \subset B$ and $\mu(B \setminus A) = 0$.

Exercise 2.5.24. Suppose $(X, \mathcal{T})$ is a topological space, μ is a Borel measure on $\mathcal{B}_\mathcal{T}$. We call μ a *Radon measure* if μ is outer regular, finite on compact sets, and inner regular on all open sets.

(1) Show that if μ is a σ-finite Radon measure, then μ is inner regular and thus regular. Deduce that the finite Radon measures are exactly the finite regular Borel measures.
(2) Suppose μ is a σ-finite regular Borel measure. Is μ Radon? That is, is μ finite on all compact sets? Give a proof or a counterexample.

Exercise 2.5.25 (Folland, §7.2, #7)**.** Suppose μ is a σ-finite Radon measure on $(X, \mathcal{T})$ and $E \in \mathcal{B}_\mathcal{T}$ is a Borel set. Show that $\mu_E(F) := \mu(E \cap F)$ defines another (σ-finite) Radon measure.

Remark 2.5.26. Once we have developed the theory of integration, we will be able to upgrade Proposition 2.5.22 considerably. In Corollary 5.6.10, we will show that if X is LCH such that every open set is σ-compact, then every Borel measure which is finite on compact sets is regular and thus Radon.

Exercise 2.5.27. Suppose X is a metric space (not necessarily locally compact) and let μ be a finite Borel measure. Show that the collection $\mathcal{M} \subset \mathcal{B}_X$ of sets such that

$$\begin{aligned}\mu(E) &= \inf\{\mu(U) | E \subseteq U \text{ open}\}\\ &= \sup\{\mu(F) | E \supseteq F \text{ closed}\}\end{aligned}$$

is a σ-algebra containing all closed (or open) sets and is thus equal to $\mathcal{B}_X$. Deduce that μ is outer regular.

Exercise 2.5.28. Suppose X is a compact Hausdorff topological space, $\mathcal{B}_X$ is the Borel σ-algebra, and μ is a regular measure on $\mathcal{B}_X$ such that $\mu(X) = 1$. Prove there is a compact $K \subset X$ such that $\mu(K) = 1$ and $\mu(F) < 1$ for every proper compact subset $F \subsetneq K$.
Remark: One strategy uses Zorn's Lemma, but it is not necessary.

We now analyze the regularity of the Lebesgue-Stieltjes measure μ_F on $\mathcal{M}_F$ where $F : \mathbb{R} \to \mathbb{R}$ is any non-decreasing right continuous function.

Exercise 2.5.29. For every $E \subset \mathbb{R}$, show that

$$\mu_F^*(E) = \inf\left\{\sum_{n=1}^{\infty} \mu_F((a_n, b_n]) \middle| E \subset \bigcup_{n=1}^{\infty} (a_n, b_n] \text{ with } a_n, b_n \in \mathbb{R}, \forall\, n \in \mathbb{N}\right\}.$$

Lemma 2.5.30. *For all $E \subset \mathbb{R}$, $\mu_F^*(E) = \inf\{\sum_{n=1}^{\infty} \mu_F((a_n, b_n)) | E \subset \bigcup_{n=1}^{\infty}(a_n, b_n)\}$.*

Proof. Denote the inf on the right hand side by $\nu(E)$.
Step 1: $\mu_F^*(E) \leq \nu(E)$.

Suppose $E \subset \bigcup(a_n, b_n)$. We can write each $(a_n, b_n) = \coprod_{i=1}^{\infty}(a_i^n, b_i^n]$. Then $E \subset \bigcup_{n=1}^{\infty} \coprod_{i=1}^{\infty}(a_i^n, b_i^n]$, and

$$\mu_F^*(E) \leq \sum_{n,i} \mu_F((a_i^n, b_i^n]) = \sum \mu_F((a_n, b_n)).$$

Step 2: $\mu_F(E) \geq \nu(E)$.

Let $\varepsilon > 0$. There exists $((a_n, b_n])$ such that $E \subset \bigcup(a_n, b_n]$ and $\sum \mu_F((a_n, b_n]) \leq \mu_F^*(E) + \frac{\varepsilon}{2}$. For each n, by right continuity of F, pick $\delta_n > 0$ such that $F(b_n + \delta_n) - F(b_n) < \frac{\varepsilon}{2^{n+1}}$. Then $E \subset \bigcup(a_n, b_n + \delta_n)$ and

$$\begin{aligned}
\sum \mu_F((a_n, b_n + \delta_n)) &\leq \sum F(b_n + \delta_n) - F(a_n) \\
&< \sum F(b_n) - F(a_n) + \frac{\varepsilon}{2^{n+1}} \\
&= \sum \mu_F((a_n, b_n]) + \sum \frac{\varepsilon}{2^{n+1}} \\
&\leq \mu_F^*(E) + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\
&= \mu_F^*(E) + \varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, the result follows.

This concludes the proof. □

Theorem 2.5.31. *The Lebesgue-Stieltjes measure μ_F on $\mathcal{M}_F$ is regular.*

Proof. Since $\mathbb{R}$ is σ-compact and μ_F is finite on all compact intervals by Exercise 2.5.9, by Proposition 2.5.22, it remains to show μ_F is outer regular. Let $E \in \mathcal{M}_F$. By Lemma 2.5.30, given $\varepsilon > 0$, there is a sequence $((a_n, b_n))$ of open intervals such that $E \subset \bigcup(a_n, b_n)$ and $\sum \mu_F((a_n, b_n)) \leq \mu(E) + \varepsilon$. Setting $U = \bigcup(a_n, b_n)$, we have $E \subset U$ and

$$\mu_F(E) \leq \mu_F(U) \leq \sum \mu_F((a_n, b_n)) \leq \mu(E) + \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, we have $\mu_F(E) = \inf\{\mu_F(U) | E \subset U \text{ open}\}$. □

Exercise 2.5.32. Show that $\mu_F^*(E) = \inf\{\mu_F(U) | E \subset U \text{ open}\}$ for every $E \subset \mathbb{R}$. Then find the error in the following 'proof' that $\mathcal{M}_F = P(\mathbb{R})$.

'Proof'. Suppose $E \subset \mathbb{R}$. By Lemma 2.5.30, for every $\varepsilon > 0$, there is an open subset U so that $\mu_F^*(E) \leq \mu_F(U) + \varepsilon$. Inductively construct a decreasing sequence of open sets $E \subset U_n \subset U_{n+1}$ such that $\mu_F^*(U_n \setminus E) \leq 1/n$. Then $F := \bigcap U_n \in \mathcal{B}_{\mathbb{R}} \subset \mathcal{M}_F$ and $\mu_F^*(F \setminus E) \leq \mu_F^*(U_n \setminus E) \leq 1/n$ for every $n \in \mathbb{N}$. Then $N := F \setminus E$ is μ^*-null, and $E = F \setminus N$, so by Exercise 2.4.12, $E \in \mathcal{M}_F$. □

Exercise 2.5.33 (Steinhaus Theorem, Folland §1.5, #30 and #31)**.** Suppose $E \in \mathcal{L}$ and $\lambda(E) > 0$.

(1) Show that for any $0 \leq \alpha < 1$, there is an open interval $I \subset \mathbb{R}$ such that $\lambda(E \cap I) > \alpha\lambda(I)$.
(2) Apply (1) with $\alpha = 3/4$ to show that the set

$$E - E := \{x - y | x, y \in E\}$$

contains the interval $(-\lambda(I)/2, \lambda(I)/2)$.

2.6. **Hausdorff measure.** Let (X, d) be a metric space. For $A, B \subset X$ nonempty, define

$$\begin{aligned} d(a, B) &:= \inf \{d(a, b) | b \in B\} & \forall a \in A \\ d(A, B) &:= \inf \{d(a, b) | a \in A,\, b \in B\}. \end{aligned}$$

For a set $Y \subset X$, define

$$\operatorname{diam}(Y) := \sup \{d(x, y) | x, y \in Y\}.$$

Definition 2.6.1. An outer measure μ^* on $P(X)$ is called a (Carathéodory) *metric outer measure* if

- (metric finite additivity) $d(A, B) > 0$ (which implies $A \cap B = \emptyset$) implies $\mu^*(A \coprod B) = \mu^*(A) + \mu^*(B)$.

Proposition 2.6.2. *If μ^* is a metric outer measure on $P(X)$, then the Borel σ-algebra $\mathcal{B}_d$ is contained in $\mathcal{M}^*$, the μ^*-measurable sets.*

Proof. Since $\mathcal{B}_d$ is generated by the open sets, it suffices to show all open sets are in $\mathcal{M}^*$. Let $U \subset X$ be open.
Step 1: We may assume $d(U, U^c) = 0$. Otherwise, for all $F \subset X$, $d(F \cap U, F \setminus U) > 0$, so $\mu^*(F) = \mu^*(F \cap U) + \mu^*(F \setminus U)$, and thus $U \in \mathcal{M}^*$.
Step 2: For $n \in \mathbb{N}$, define $A_n := \{x \in U | d(x, U^c) > 1/n\}$. Then (A_n) is increasing and $\bigcup A_n = \mathcal{U}$. Setting $A_0 = \emptyset$, define $B_n := A_n \setminus A_{n-1}$ for all $n \in \mathbb{N}$. Then $\coprod B_n = U$, and $B_n \neq \emptyset$ frequently. Indeed, observe $B_n = \emptyset$ for all $n > k$ if and only if $A_k = U$, which implies $d(U, U^c) \geq 1/k$.
Step 3: If $|m - n| > 1$ and $B_m \neq \emptyset \neq B_n$, then $d(B_m, B_n) > 0$.

Proof. Suppose $1 \leq m < n-1$. Let $x \in B_m$ and $y \in B_n$. Then $y \notin A_{n-1} \supset A_{m+1}$, so there is a $z \in U^c$ such that $d(y,z) \leq \frac{1}{m+1}$. But $x \in B_m$, so $d(x,z) > \frac{1}{m}$. By the triangle inequality,

$$d(x,y) \geq d(x,z) - d(y,z) > \frac{1}{m} - \frac{1}{m+1} = \frac{1}{m(m+1)}.$$

Taking sup over x, y, we have $d(B_m, B_n) \geq \frac{1}{m(m+1)} > 0$. $\square$

Step 4: Let $F \subset X$. If $\mu^*(F) = \infty$, then $\mu^*(F) \geq \mu^*(F \cap U) + \mu^*(F \setminus U)$. Assume $\mu^*(F) < \infty$. Then $\sum_{n=k}^{\infty} \mu^*(F \cap B_n) \to 0$ as $k \to \infty$.

Proof. By Step 3, for all $k \in \mathbb{N}$, we have

$$\sum^{k} \mu^*(F \cap B_{2n-1}) = \mu^* \left(\coprod^{k} F \cap B_{2n-1} \right) \leq \mu^*(F)$$

$$\sum^{k} \mu^*(F \cap B_{2n}) = \mu^* \left(\coprod^{k} F \cap B_{2n} \right) \leq \mu^*(F).$$

Taking $k \to \infty$, we have $\sum \mu^*(F \cap B_n) \leq 2\mu^*(F) < \infty$. Hence the tail of the sum must converge to zero. $\square$

Step 5: We now calculate for all $n \in \mathbb{N}$ and $F \subset X$:

$$\begin{aligned}
\mu^*(F \cap U) + \mu^*(F \setminus U) &\leq \mu^*(F \cap A_n) + \mu^*(F \cap (\underbrace{U \setminus A_n}_{\coprod_{k=n+1}^{\infty} B_k})) + \mu^*(F \setminus U) \\
&= \underbrace{\mu^*(F \cap A_n) + \mu^*(F \setminus U)}_{d(F \cap A_n, F \setminus U) \geq d(A_n, U^c) \geq \frac{1}{n}} + \mu^* \left(\coprod_{k=n+1}^{\infty} B_k \right) \\
&= \mu^*(F \cap (A_n \amalg F \setminus U)) + \mu^* \left(\coprod_{k=n+1}^{\infty} B_k \right) \\
&\leq \mu^*(F) + \underbrace{\sum_{k=n+1}^{\infty} \mu^*(F \cap B_k).}_{\to 0 \text{ as } n \to \infty \text{ by Step 4.}}
\end{aligned}$$

We conclude that $U \in \mathcal{M}^*$. $\square$

Definition 2.6.3. Suppose (X,d) is a metric space, $p \geq 0$, and $\varepsilon > 0$. For $E \subset X$, define

$$\eta^*_{p,\varepsilon}(E) := \inf \left\{ \sum_{1}^{\infty} (\operatorname{diam}(B_n))^p \,\middle|\, \begin{array}{l} (B_n) \text{ a } \leq \varepsilon\textit{-diameter cover}, \text{ i.e., a sequence of open} \\ \text{balls with } \operatorname{diam}(B_n) \leq \varepsilon \text{ for all } n \text{ and } E \subset \bigcup B_n \end{array} \right\},$$

where we use the convention that $\inf \emptyset = \infty$. By Exercise 2.3.5, $\eta^*_{p,\varepsilon}$ is the outer measure induced by

$$\begin{aligned} \rho_{p,\varepsilon} : \{\emptyset\} \cup \{B_r(x) | x \in X \text{ and } r \leq \varepsilon\} &\longrightarrow [0,\infty] \\ \emptyset &\longmapsto 0 \\ B_r(x) &\longmapsto (\operatorname{diam}(B_r(x)))^p. \end{aligned}$$

Moreover, if $\varepsilon < \varepsilon'$, then $\eta^*_{p,\varepsilon}(E) \geq \eta^*_{p,\varepsilon'}(E)$ as we are taking an infimum over a smaller set (every $\leq \varepsilon$-diameter cover is a $\leq \varepsilon'$-diameter cover). Hence

$$\eta^*_p(E) := \lim_{\varepsilon \to 0} \eta^*_{p,\varepsilon}(E) = \sup_{\varepsilon > 0} \eta^*_{p,\varepsilon}(E)$$

gives an outer measure by Exercise 2.3.2.

Proposition 2.6.4. *η^*_p is a metric outer measure.*

Proof. Suppose $d(E,F) > \varepsilon > 0$. If there is no ε-diameter cover of $E \amalg F$, then there is no ε-diameter cover of one of E, F, and thus

$$\eta^*_p(E) + \eta^*_p(F) = \infty = \eta^*_p(E \amalg F).$$

Now suppose there exists an ε-diameter cover (B_n) of $E \amalg F$. Then for all $n \in \mathbb{N}$, B_n intersects at most one of E, F. So we may partition (B_n) into (B^E_n) and (B^F_n) such that

- $E \subset \bigcup B^E_n$ and $B^E_n \cap E \neq \emptyset$, and
- $F \subset \bigcup B^F_n$ and $B^F_n \cap F \neq \emptyset$.

Thus

$$\eta^*_{p,\varepsilon}(E) + \eta^*_{p,\varepsilon}(F) \leq \sum \operatorname{diam}(B^E_n)^p + \operatorname{diam}(B^F_n)^p \leq \sum \operatorname{diam}(B_n)^p$$

for any ε-diameter cover. Hence for all $\varepsilon < d(E,F)$,

$$\eta^*_{p,\varepsilon}(E) + \eta^*_{p,\varepsilon}(F) \leq \eta^*_{p,\varepsilon}(E \amalg F).$$

Taking $\varepsilon \to 0$, we get

$$\eta^*_p(E \amalg F) \leq \eta^*_p(E) + \eta^*_p(F) \leq \eta^*_p(E \amalg F),$$

and thus equality holds. □

Definition 2.6.5. Since the Borel σ-algebra $\mathcal{B}_d$ for (X,d) is contained in the η^*_p-measurable sets $\mathcal{M}^*_p$ by Propositions 2.6.2 and 2.6.4, we get a Borel measure $\eta_p := \eta^*_p|_{\mathcal{B}_d}$ called *p-dimensional Hausdorff measure.*

Facts 2.6.6. Here are some elementary properties about Hausdorff measures.

(Hμ1) If $f : X \to X$ is an *isometry* ($d(f(x), f(y)) = d(x,y)$ for all $x, y \in X$), then for all $E \in \mathcal{B}_d$, $\eta_p(E) = \eta_p(f(E))$.

> *Proof.* For all $\varepsilon > 0$, $\eta^*_{p,\varepsilon}(E) = \eta^*_{p,\varepsilon}(f(E))$ since $E \subset \bigcup B_n$ if and only if $f(E) \subset \bigcup f(B_n)$ as isometries are injective. □

(Hμ2) $\eta_1 = \lambda|_{\mathcal{B}_\mathbb{R}}$ on $\mathbb{R}$ with the usual metric.

Proof. Since $\eta_1((0,1]) = 1$ (observe $\operatorname{diam}(B) = \lambda(B)$ for any open ball B and apply Lemma 2.5.30), this follows by uniqueness of the translation invariant Borel measure on $\mathbb{R}$ from Exercise 2.5.13. □

(Hμ3) If $\eta_p(E) < \infty$, then $\eta_q(E) = 0$ for all $q > p$.

Proof. Let $\varepsilon > 0$. Since $\eta_p(E) < \infty$, there is a sequence (B_n) of open balls with $\operatorname{diam}(B_n) \leq \varepsilon$ such that $\sum \operatorname{diam}(B_n)^p \leq \eta_p(E) + 1$. But if $q > p$, then

$$\begin{aligned}\eta^*_{q,\varepsilon}(E) &\leq \sum \operatorname{diam}(B_n)^q \\ &= \sum \underbrace{\operatorname{diam}(B_n)^{q-p}}_{\leq \varepsilon^{q-p}} \operatorname{diam}(B_n)^p \\ &\leq \varepsilon^{q-p} \sum \operatorname{diam}(B_n)^p \\ &\leq \varepsilon^{q-p}(\eta_p(E) + 1).\end{aligned}$$

Letting $\varepsilon \to 0$, we have

$$\eta_q(E) = \eta^*_q(E) = \lim_{\varepsilon \to 0} \eta^*_{q,\varepsilon}(E) \leq \lim_{\varepsilon \to 0} \varepsilon^{q-p}(\eta_p(E) + 1) = 0.$$ □

(Hμ4) If $\eta_p(E) > 0$, then $\eta_q(E) = \infty$ for all $q < p$.

Proof. This follows as the contrapositive of (Hμ3). □

Definition 2.6.7. The *Hausdorff dimension* of $E \in \mathcal{B}_d$ is

$$\inf \{p \geq 0 | \eta_p(E) = 0\} = \sup \{p \geq 0 | \eta_p(E) = \infty\}.$$

Remark 2.6.8. If $E \in \mathcal{B}_d$ and $p \geq 0$ such that $0 < \eta_p(E) < \infty$, then the Hausdorff dimension of E is necessarily p by Lemma 2.6.6(3,4).

Exercise 2.6.9. Prove that the Cantor set from Example 2.5.15 has Hausdorff dimension $\ln(2)/\ln(3)$.

Exercise 2.6.10. Find an uncountable subset of $\mathbb{R}$ with Hausdorff dimension zero.

3. Integration

3.1. Measurable functions.

Definition 3.1.1. If $(X,\mathcal{M})$ and $(Y,\mathcal{N})$ are measurable spaces, we say $f : X \to Y$ is $(\mathcal{M}-\mathcal{N})$ *measurable* if $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{N}$.

Exercise 3.1.2. Prove the following assertions.

(1) Given $f : X \to Y$ and a σ-algebra $\mathcal{N}$ on Y, $\{f^{-1}(E)|E \in \mathcal{N}\}$ is a σ-algebra on X. Moreover it is the smallest σ-algebra on X such that f is measurable.

(2) Given $f : X \to Y$ and a σ-algebra $\mathcal{M}$ on X, $\{E \subset Y|f^{-1}(E) \in \mathcal{M}\}$ is a σ-algebra on Y. Moreover it is the largest σ-algebra on Y such that f is measurable.

Exercise 3.1.3. Prove that the composite of two measurable functions is measurable. More precisely, if $f : (X,\mathcal{M}) \to (Y,\mathcal{N})$ is $\mathcal{M}-\mathcal{N}$ measurable and $g : (Y,\mathcal{N}) \to (Z,\mathcal{P})$ is $\mathcal{N}-\mathcal{P}$ measurable, then $g\circ f$ is $\mathcal{M}-\mathcal{P}$ measurable. Deduce that measurable spaces and measurable functions form a category.

Proposition 3.1.4. *Suppose $(X,\mathcal{M})$ and $(Y,\mathcal{N})$ are measurable spaces, $f : X \to Y$, and $\mathcal{N} = \langle\mathcal{E}\rangle$ for some $\mathcal{E} \subset P(Y)$. Then f is measurable if and only if $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{E}$.*

Proof. The forward direction is trivial. Suppose $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{E}$. Then $\mathcal{E}$ is contained in the σ-algebra $\mathcal{N}_f$ on Y co-induced by $\mathcal{M}, f$, i.e., the largest σ-algebra such that f is measurable. Since $\mathcal{N}_f$ is a σ-algebra containing $\mathcal{E}$, we see that $\mathcal{N}_f$ contains $\mathcal{N}$. Since f is $\mathcal{M}-\mathcal{N}_f$ measurable, f is $\mathcal{M}-\mathcal{N}$ measurable. □

Exercise 3.1.5. Show that every monotone increasing function $f : \mathbb{R} \to \mathbb{R}$ is Borel measurable.

Definition 3.1.6. Suppose X, Y are topological spaces. We call $f : X \to Y$ *Borel measurable* if it is $\mathcal{B}_X - \mathcal{B}_Y$ measurable.

Corollary 3.1.7. *Continuous functions are Borel measurable.*

Proof. Observe $f : X \to Y$ is continuous if and only if for all $U \in \mathcal{T}_Y$, $f^{-1}(U) \in \mathcal{T}_X \subset \mathcal{B}_X$. This implies f is Borel measurable by Proposition 3.1.4. □

Corollary 3.1.8. *Suppose $(X,\mathcal{M})$ is a measurable space and $f : X \to \mathbb{R}$ (where $\mathbb{R}$ is equipped with the Borel σ-algebra). The following are equivalent:*

(1) f is $\mathcal{M} - \mathcal{B}_\mathbb{R}$ measurable.
(2) $f^{-1}(a,\infty) \in \mathcal{M}$ for all $a \in \mathbb{R}$.
(3) $f^{-1}[a,\infty) \in \mathcal{M}$ for all $a \in \mathbb{R}$.
(4) $f^{-1}(-\infty, a) \in \mathcal{M}$ for all $a \in \mathbb{R}$.
(5) $f^{-1}(-\infty, a] \in \mathcal{M}$ for all $a \in \mathbb{R}$.

Observe that we can also use collections of intervals $(a,b), [a,b), (a,b], [a,b]$ for all $a,b \in \mathbb{R}$.

Corollary 3.1.9. *If $(X,\mathcal{M})$ is a measurable space and $f : X \to \overline{\mathbb{R}} = [-\infty,\infty]$, then Corollary 3.1.8 holds replacing $\mathbb{R}$ with $\overline{\mathbb{R}}$ and intervals excluding $\pm\infty$ with intervals including $\pm\infty$ respectively.*

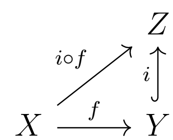
Proof. Use Exercise 2.1.12. □

Definition 3.1.10. Suppose $(X, \mathcal{M})$ is a measurable space. We say a function $f : X \to \mathbb{R}, \overline{\mathbb{R}}, \mathbb{C}$ is $\mathcal{M}$-*measurable* if f is $\mathcal{M} - \mathcal{B}_{\mathbb{R}}$, $\mathcal{M} - \mathcal{B}_{\overline{\mathbb{R}}}$, $\mathcal{M} - \mathcal{B}_{\mathbb{C}}$ measurable respectively.

Warning 3.1.11. If $f, g : \mathbb{R} \to \mathbb{R}$ are Lebesgue measurable (i.e., $\mathcal{L} - \mathcal{B}_{\mathbb{R}}$ measurable), then $f \circ g$ need not be Lebesgue measurable!

Exercise 3.1.12. Find examples of $f, g : \mathbb{R} \to \mathbb{R}$ are Lebesgue measurable with $f \circ g$ not Lebesgue measurable.
Note: First find an $E \in \mathcal{L} \setminus \mathcal{B}_{\mathbb{R}}$ and an $\mathcal{L}$-measurable function $f : \mathbb{R} \to \mathbb{R}$ such that $f^{-1}(E) \notin \mathcal{L}$. Then set $g := \chi_E$.

Exercise 3.1.13. Suppose $(X, \mathcal{M})$ is a measurable space and X, Y are topological spaces, $i : Y \to Z$ is a continuous injection which maps open sets to open sets, and $f : X \to Y$. (For example, $Y = \mathbb{R}$ and $Z = \overline{\mathbb{R}}$.)



Show that f is $\mathcal{M} - \mathcal{B}_Y$ measurable if and only if $i \circ f$ is $\mathcal{M} - \mathcal{B}_Z$ measurable. Deduce that if $f : (X, \mathcal{M}) \to \overline{\mathbb{R}}$ only takes values in $\mathbb{R}$, then f is $\mathcal{M} - \mathcal{B}_{\overline{\mathbb{R}}}$ measurable if and only if f is $\mathcal{M} - \mathcal{B}_{\mathbb{R}}$ measurable. Hence we can say f is $\mathcal{M}$-measurable without any confusion.

Exercise 3.1.14. Let $(X, \mathcal{M})$ be a measurable space.

(1) Prove that the Borel σ-algebra $\mathcal{B}_{\mathbb{C}}$ on $\mathbb{C}$ is generated by the 'open rectangles'
$$\{z \in \mathbb{C} | a < \operatorname{Re}(z) < b \text{ and } c < \operatorname{Im}(z) < d\}.$$
(2) Prove directly from the definitions that $f : X \to \mathbb{C}$ is $\mathcal{M} - \mathcal{B}_{\mathbb{C}}$ measurable if and only if $\operatorname{Re}(f)$ and $\operatorname{Im}(f)$ are $\mathcal{M} - \mathcal{B}_{\mathbb{R}}$ measurable.

Definition 3.1.15. Suppose $(X, \mathcal{M}, \mu)$ is a measure space. We say that a property P of a measurable function f from X into $\mathbb{R}$, $\overline{\mathbb{R}}$, or $\mathbb{C}$ holds *almost everywhere (a.e.)* if there is a μ-null set $E \in \mathcal{M}$ such that P holds on E^c. For example, $f \geq 0$ a.e. if there is a μ-null set $E \in \mathcal{M}$ such that $f|_{E^c} \geq 0$.

Exercise 3.1.16. Define a relation on the set of $\mathcal{M}$-measurable functions (into $\mathbb{R}$, $\overline{\mathbb{R}}$, or $\mathbb{C}$) by $f \sim g$ if and only if $f = g$ a.e. Prove $\sim$ is an equivalence relation.

Notation 3.1.17. Given $f : X \to \overline{\mathbb{R}}$, we write $\{a < f\} := f^{-1}(a, \infty]$. We define $\{a \leq f\}, \{f < b\}, \{f \leq b\}, \{a < f < b\}$, etc. similarly.

Facts 3.1.18. Suppose $(X, \mathcal{M})$ is a measurable space and $f, g : X \to \overline{\mathbb{R}}$ are $\mathcal{M}$-measurable. The following functions are all $\mathcal{M}$-measurable:

($\mathcal{M}$-meas1) $(f \vee g)(x) := \max\{f(x), g(x)\}$ and $(f \wedge g)(x) := \min\{f(x), g(x)\}$

Proof. If $a \in \mathbb{R}$, then
$$\{a < f \vee g\} = \{a < f\} \cup \{a < g\} \in \mathcal{M}$$
$$\{a < f \wedge g\} = \{a < f\} \cap \{a < g\} \in \mathcal{M}. \qquad \square$$

($\mathcal{M}$-meas2) any well-defined linear combination of f, g, where by convention, $0 \cdot \pm\infty = 0$ and $\pm\infty \pm \infty = \pm\infty$, but $\pm\infty \mp \infty$ is not defined.

Proof.
Step 1: For $a, c \in \mathbb{R}$,

$$\{cf > a\} = \begin{cases} \emptyset & \text{if } c = 0 \leq a \\ X & \text{if } c = 0 > a \\ \left\{\frac{a}{c} < f\right\} & \text{if } c > 0 \\ \left\{\frac{a}{c} > f\right\} & \text{if } c < 0 \end{cases} \qquad \text{which are all in } \mathcal{M}.$$

Step 2: If $f + g$ is well-defined, then for $a \in \mathbb{R}$,

$$\{a < f + g\} = \bigcup_{\substack{r,s \in \mathbb{Q} \\ a < r+s}} \{r < f\} \cap \{s < g\} \in \mathcal{M}.$$

□

($\mathcal{M}$-meas3) fg

Proof.
Step 1: Suppose f, g are non-negative. Then for all $a \geq 0$,

$$\{a < fg\} = \bigcup_{\substack{r,s \in \mathbb{Q}_{>0} \\ a < rs}} \{r < f\} \cap \{s < g\} \in \mathcal{M}.$$

Also, for all $a < 0$, $\{a < fg\} = X \in \mathcal{M}$.
Step 2: For f, g arbitrary, we use the following trick:

Trick. $f = f_+ - f_-$ where $f_+ := f \vee 0$ and $f_- := -(f \wedge 0)$. Observe that $f_\pm \cdot f_\mp = 0$.

Similarly, we can write $g = g_+ - g_-$. Then

$$fg = f_+g_+ - f_-g_+ - f_+g_- + f_-g_-,$$

all of which have disjoint support. Hence each of the summands of fg is measurable by Step 1, and the linear combination is measurable by (3) as it is well-defined. □

Exercise 3.1.19. Suppose $f : X \to \overline{\mathbb{R}}$. Show that $f = f_+ - f_-$ is the unique decomposition of f as $g - h$ such that $g, h \geq 0$ and $gh = 0$.

Exercise 3.1.20. Let $(X, \mathcal{M})$ be a measurable space.

(1) Prove that the $\mathcal{M} - \mathcal{B}_\mathbb{C}$ measurable functions form a $\mathbb{C}$-vector space.
(2) Show that if $f : X \to \mathbb{C}$ is $\mathcal{M} - \mathcal{B}_\mathbb{C}$ measurable, then $|f| : X \to [0, \infty)$ is $\mathcal{M} - \mathcal{B}_\mathbb{R}$ measurable.
(3) Show that if (f_n) is a sequence of $\mathcal{M} - \mathcal{B}_\mathbb{C}$ measurable functions $X \to \mathbb{C}$ and $f_n \to f$ pointwise, then f is $\mathcal{M} - \mathcal{B}_\mathbb{C}$ measurable.

Facts 3.1.21. Suppose (f_n) is a sequence of $\mathcal{M}$-measurable functions $X \to \overline{\mathbb{R}}$. The following functions are $\mathcal{M}$-measurable.

($\mathcal{M}$-meas4) $\sup f_n$ and $\inf f_n$

Proof. For all $a \in \mathbb{R}$,

$$\{a < \sup f_n\} = \bigcup_n \{a < f_n\} \in \mathcal{M}$$
$$\{a < \inf f_n\} = \bigcap_n \{a < f_n\} \in \mathcal{M}. \qquad \square$$

($\mathcal{M}$-meas5) $\limsup f_n$ and $\liminf f_n$

Proof. Observe that

$$\limsup f_n = \lim_{n\to\infty} \sup_{k>n} f_k = \inf_n \underbrace{\sup_{k>n} f_k}_{\text{measurable by } (\mathcal{M}\text{-meas4})}$$
$$\liminf f_n = \lim_{n\to\infty} \inf_{k>n} f_k = \sup_n \underbrace{\inf_{k>n} f_k}_{\text{measurable by } (\mathcal{M}\text{-meas4})}$$

Applying ($\mathcal{M}$-meas4) again, we see that $\limsup f_n$ and $\liminf f_n$ are $\mathcal{M}$-measurable. $\square$

3.2. **Measurable simple functions.** For this section, fix a measurable space $(X, \mathcal{M})$.

Definition 3.2.1. An $\mathcal{M}$-measurable function $\psi : X \to \mathbb{R}$ is *simple* if it takes finitely many values. Observe that if ψ is simple, we can write

$$\psi = \sum_{k=1}^{n} c_k \chi_{E_k} \qquad c_1, \ldots, c_n \in \mathbb{R} \qquad E_1, \ldots, E_n \in \mathcal{M}.$$

Here, we write χ_E for the *characteristic function* of E:

$$\chi_E(x) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \in E^c. \end{cases}$$

Observe that there is exactly one such expression of a simple function, called its *standard form*, such that

- $c_1, \ldots, c_n$ are distinct, and
- $E_1, \ldots, E_n$ are disjoint and non-empty such that $X = \coprod_{k=1}^{n} E_k$.

Denote by SF the collection of simple ($\mathcal{M}$-measurable) functions. Define $\mathsf{SF}^+ := \{\psi \in \mathsf{SF} | \psi \geq 0\}$.

Exercise 3.2.2. Verify the uniqueness of standard form of an simple function.

Exercise 3.2.3.

(1) Prove that SF is an $\mathbb{R}$-algebra and SF^+ is closed under addition, multiplication, and non-negative scalar multiplication.
(2) Prove SF is a lattice (closed under max and min) and $\mathsf{SF}^+ \subset \mathsf{SF}$ is a sublattice.

Proposition 3.2.4. *Suppose $f : X \to [0,\infty]$ is $\mathcal{M}$-measurable. There is a sequence $(\psi_n) \subset \mathsf{SF}^+$ such that*

- *$\psi_n(x) \nearrow f(x)$ for all $x \in X$, and*
- *for all $N \in \mathbb{N}$, $\psi_n \to f$ uniformly on $\{f \leq N\}$.*

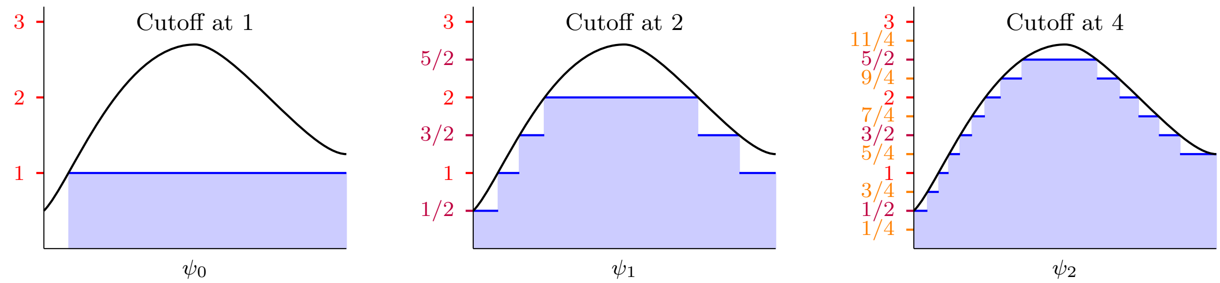
Proof. For $n \geq 0$ and $1 \leq k \leq 2^{2n}$, set

$$E_n^k := f^{-1}\left(\frac{k-1}{2^n}, \frac{k}{2^n}\right] \qquad \text{and} \qquad F_n := f^{-1}(2^n, \infty].$$

Observe that $X = f^{-1}(0) \amalg F_n \amalg \coprod_{k=1}^{2^{2n}} E_n^k$. Define

$$\psi_n := 2^n \chi_{F_n} + \sum_{k=1}^{2^{2n}} \frac{k-1}{2^n} \chi_{E_n^k}.$$

Here is a cartoon of ψ_0, ψ_1, ψ_2:



Observe that $\psi_n \leq \psi_{n+1}$ for all $n \geq 0$, and $0 \leq f - \psi_n \leq 2^{-n}$ on $\{f \leq 2^n\}$. The result follows. $\square$

Exercise 3.2.5. Let $(X, \overline{\mathcal{M}}, \overline{\mu})$ be the completion of the measure space $(X, \mathcal{M}, \mu)$.

(1) Show that if f is $\overline{\mathcal{M}}$-measurable and $g = f$ a.e., then g is $\overline{\mathcal{M}}$-measurable.
Optional: Does this hold with $\overline{\mathcal{M}}$ replaced by $\mathcal{M}$?

(2) Show that if f is $\overline{\mathcal{M}}$-measurable, there exists an $\mathcal{M}$-measurable g such that $f = g$ a.e.
Hint: First do the case f is $\mathbb{R}$-valued.

(3) Show that if (f_n) is a sequence of $\overline{\mathcal{M}}$-measurable functions and $f_n \to f$ a.e., then f is $\overline{\mathcal{M}}$-measurable.
Optional: Does this hold with $\overline{\mathcal{M}}$ replaced by $\mathcal{M}$?

(4) Show that if (f_n) is a sequence of $\mathcal{M}$-measurable functions and $f_n \to f$ a.e., then f is $\overline{\mathcal{M}}$-measurable. Deduce that there is an $\mathcal{M}$-measurable function g such that $f = g$ a.e., so $f_n \to g$ a.e.

For all parts, consider the cases of $\mathbb{R}$, $\overline{\mathbb{R}}$, and $\mathbb{C}$-valued functions.

3.3. **Integration of non-negative functions.** For this section, fix a measure space $(X, \mathcal{M}, \mu)$. Define

$$L^+ := L^+(X, \mathcal{M}, \mu) = \{\mathcal{M}\text{-measurable } f : X \to [0,\infty]\}.$$

Definition 3.3.1. For $\psi = \sum_{k=1}^n c_k \chi_{E_k} \in \mathsf{SF}^+ \subset L^+$ in standard form, define

$$\int \psi := \int_X \psi \, d\mu := \int_X \psi(x) \, d\mu(x) := \sum_{k=1}^n c_k \mu(E_k).$$

For $E \in \mathcal{M}$, we define $\int_E \psi := \int \psi \cdot \chi_E$. Observe that to calculate $\int_E \psi$, we must write the simple function $\psi \cdot \chi_E$ in standard form.

We say that $\psi \in \mathsf{SF}^+$ is *integrable* if $\int \psi < \infty$. We write $\mathsf{ISF}^+ := \left\{\psi \in \mathsf{SF}^+ \middle| \psi \text{ integrable}\right\}$.

Exercise 3.3.2. Suppose $f : (X, \mathcal{M}, \mu) \to [0, \infty]$ is $\mathcal{M}$-measurable and $\{f > 0\}$ is σ-finite. Show that there exists a sequence of $(\psi_n) \subset \mathsf{ISF}^+$ such that $\psi_n \nearrow f$ pointwise.
Optional: In what sense can you say $\psi_n \nearrow f$ uniformly?

Theorem 3.3.3. *The map $\int : \mathsf{SF}^+ \to [0, \infty]$ satisfies*

(1) (homogeneous) for all $r \geq 0$, $\int r\psi = r \int \psi$.
(2) (monotone) if $\phi \leq \psi$ everywhere, then $\int \phi \leq \int \psi$.
(3) (additive) $\int \phi + \psi = \int \phi + \int \psi$.

Hence $\int : \mathsf{SF}^+ \to [0, \infty]$ is an order-preserving $\mathbb{R}^+$-linear functional.

Proof.
(1) Observe if $r = 0$, then $\int r\psi = 0 = 0 \cdot \int \psi$. If $r > 0$ and $\psi = \sum^n c_k \chi_{E_k}$, then $r\psi = \sum^n rc_k \chi_{E_k}$ is in standard form, and

$$\int r\psi = \sum^n rc_k \mu(E_k) = r \sum^n c_k \mu(E_k) = r \int \psi.$$

(2) Suppose that $\phi = \sum^m a_j \chi_{E_j}$ and $\psi = \sum^n b_k \chi_{F_k}$ are in standard form. Here is the trick:

Trick. Since $X = \coprod^m E_j = \coprod^n F_k$, we have $E_j = \coprod_{k=1}^n E_j \cap F_k$ and $F_k = \coprod_{j=1}^m E_j \cap F_k$.

Since $\phi \leq \psi$ everywhere,

$$\phi = \sum_{j,k} a_j \chi_{E_j \cap F_k} \leq \sum_{j,k} b_k \chi_{E_j \cap F_k} = \psi,$$

and so $a_j \leq b_k$ whenever $E_j \cap F_k \neq \emptyset$. Thus

$$\int \phi = \sum_{j=1}^m a_j \mu(E_j) = \sum_{j=1}^m \sum_{k=1}^n a_j \mu(E_j \cap F_k) \leq \sum_{k=1}^n \sum_{j=1}^m b_k \mu(E_j \cap F_k) = \sum_{k=1}^n b_k \mu(F_k) = \int \psi.$$

(3) Suppose that $\phi = \sum^m a_j\chi_{E_j}$, $\psi = \sum^n b_k\chi_{F_k}$, and $\phi + \psi = \sum_{\ell=1}^p c_\ell\chi_{G_\ell}$ are in standard form. Similar to the argument in (2) above, $a_j + b_k = c_\ell$ whenever $E_j \cap F_k \cap G_\ell \neq \emptyset$. Then

$$\begin{aligned}\int \phi + \int \psi &= \sum_j a_j\mu(E_j) + \sum_k b_k\mu(F_k) \\ &= \sum_{j,k}(a_j + b_k)\mu(E_j \cap F_k) \\ &= \sum_{j,k,\ell}(a_j + b_k)\mu(E_j \cap F_k \cap G_\ell) \\ &= \sum_{j,k,\ell} c_\ell\mu(E_j \cap F_k \cap G_\ell) \\ &= \sum_\ell c_\ell\mu(G_\ell) \\ &= \int \phi + \psi. \end{aligned}$$

□

Remark 3.3.4. Observe that the map $\mathcal{M} \to [0,\infty]$ by $E \mapsto \int_E d\mu$ equals μ.

Lemma 3.3.5. *For $\psi \in \mathsf{SF}^+$, $\mu_\psi : \mathcal{M} \to [0,\infty]$ by $E \mapsto \int_E \psi$ is a measure.*

Proof.
(0) Observe that $\psi\chi_\emptyset = 0$, so

$$\mu_\psi(\emptyset) = \int_\emptyset \psi = \int \psi\chi_\emptyset = \int 0 = 0.$$

(1) Write $\psi = \sum_{j=1}^m a_j\chi_{E_j}$ in standard form. If $(F_n) \subset \mathcal{M}$ is a disjoint sequence, then observe $\psi\chi_{\coprod F_n} = \sum_{j=1}^m a_j\chi_{E_j \cap \coprod F_n}$ is also in standard form (up to a subset of $\{\psi\chi_{\coprod F_n} = 0\}$).

$$\begin{aligned}\mu_\psi\left(\coprod F_n\right) &= \int_{\coprod F_n} \psi \\ &= \int \psi\chi_{\coprod F_n} \\ &= \sum_j a_j\mu(E_j \cap \coprod F_n) \\ &= \sum_{j,n} a_j\mu(E_j \cap F_n) \\ &= \sum_n \int_{F_n} \psi. \end{aligned}$$

□

Definition 3.3.6. For $f \in L^+$, define

$$\int f := \int_X f\,d\mu := \int_X f(x)\,d\mu(x) := \sup\left\{\int \psi \,\middle|\, \psi \in \mathsf{SF}^+ \text{ such that } 0 \leq \psi \leq f\right\}.$$

Remarks 3.3.7.

(1) Observe that for $\psi \in \mathsf{SF}^+$, we have
$$\int \psi = \sup\left\{ \int \phi \middle| \phi \in \mathsf{SF}^+ \text{ such that } 0 \le \phi \le \psi \right\}.$$
Hence the above definition extends $\int \psi$ for $\psi \in \mathsf{SF}^+$ to $f \in L^+$.

(2) If $f, g \in L^+$ with $f \le g$, then $\int f \le \int g$ as we are taking sup over a larger set.

(3) If $f \in L^+$ and $r \in (0, \infty)$, then $\int rf = r \int f$, since if $S \subset [0, \infty]$, $\sup rS = r \cdot \sup S$. (Remember that $0 \cdot \infty = 0$.)

Proposition 3.3.8. *Suppose $f \in L^+$. The following are equivalent.*

(1) $\int f = 0$, and

(2) $f = 0$ a.e., i.e., there is a μ-null set $E \in \mathcal{M}$ such that $f|_{E^c} = 0$.

Proof.

(1) $\Rightarrow$ (2): We'll prove the contrapositive. If f is not zero a.e., there is an $n > 0$ such that $\mu(\{\frac{1}{n} < f\}) > 0$. Then $f > \frac{1}{n}\chi_{\{\frac{1}{n} < f\}}$, so
$$0 < \frac{1}{n} \cdot \mu\left(\left\{\frac{1}{n} < f\right\}\right) = \int \frac{1}{n}\chi_{\{\frac{1}{n} < f\}} \le \int f.$$
(2) $\Rightarrow$ (1): First, if $f = \sum_{k=1}^n c_k \chi_{E_k} \in \mathsf{SF}^+$ is in standard form, then $\int f = 0$ if and only if $\mu(E_k) = 0$ for all k such that $c_k \neq 0$ if and only if $f = 0$ a.e. Second, if $f \in L^+$ with $f = 0$ a.e., then for all $\psi \in \mathsf{SF}^+$ with $0 \le \psi \le f$, $\psi = 0$ a.e., so $\int f = \sup_{0 \le \psi \le f} \int \psi = 0$. □

Theorem 3.3.9 (Monotone Convergence, a.k.a MCT)**.** *Suppose $(f_n) \subset L^+$ is an increasing sequence and $f = \lim f_n = \sup f_n$. Then*
$$\int f = \lim_{n\to\infty} \int f_n.$$
Proof.

$\le$: Observe $(\int f_n)$ is increasing in $[0, \infty]$, and thus it converges. Moreover, $\int f_n \le \int f$ for all n, so $\lim_{n\to\infty} \int f_n \le \int f$.

$\ge$: Pick a $\psi \in \mathsf{SF}^+$ with $0 \le \psi \le f$ and $0 < \varepsilon < 1$. Set $E_n := \{\varepsilon\psi < f_n\}$. Then observe $(E_n) \subset \mathcal{M}$ is an increasing sequence such that $\bigcup E_n = X$, so by continuity from below (μ3), $\int_{E_n} \psi \nearrow \int \psi$. Thus
$$\int f_n \ge \int_{E_n} f_n \ge \varepsilon \int_{E_n} \psi \xrightarrow{n\to\infty} \varepsilon \int \psi.$$
Hence $\lim \int f_n \ge \varepsilon \int \psi$ for all $0 < \varepsilon < 1$. Since ε was arbitrary, letting $\varepsilon \to 1$, we have $\lim \int f_n \ge \int \psi$. Taking sup over all $0 \le \psi \le f$ gives $\lim \int f_n \ge \int f$. □

Facts 3.3.10 (Corollaries of the MCT)**.**

(MCT1) If $f \in L^+$, then $\int f = \lim \int \psi_n$ for all sequences $(\psi_n) \subset \mathsf{SF}^+$ such that $\psi_n \nearrow f$.

(MCT2) For all $f, g \in L^+$, $\int f + g = \int f + \int g$.

Proof. If $\phi_n \nearrow f$ and $\psi_n \nearrow g$, then $\phi_n + \psi_n \nearrow f + g$, so
$$\int f + g \underset{\text{(MCT)}}{=} \lim \int \phi_n + \psi_n = \lim \int \phi_n + \lim \int \psi_n = \int f + \int g.$$ □

(MCT3) For $f, g \in L^+$, if $f = g$ a.e., then $\int f = \int g$.

Proof. Let $E \in \mathcal{M}$ such that $f\chi_E = g\chi_E$ and E^c is μ-null. Then
$$\int f \underset{\text{(MCT2)}}{=} \int f\chi_E + \int f\chi_{E^c} = \int f\chi_E = \int g\chi_E = \int g\chi_E + \int g\chi_{E^c} \underset{\text{(MCT2)}}{=} \int g.$$
□

(MCT4) For all $(f_n) \subset L^+$, $\sum \int f_n = \int \sum f_n$, where $\sum f_n$ is the sup of the sequence of partial sums (which is a measurable function).

Proof. Observe
$$\int \sum f_n = \int \lim_{N\to\infty} \sum^N f_n \underset{\text{(MCT)}}{=} \lim_{N\to\infty} \int \sum^N f_n \underset{\text{(MCT2)}}{=} \lim_{N\to\infty} \sum^N \int f_n = \sum \int f_n.$$
□

(MCT5) If $(f_n) \subset L^+$, $f_n \nearrow f$ a.e., and $f \in L^+$ (which is automatic if μ is complete), then $\int f = \lim \int f_n$.

Proof. Suppose $f_n \nearrow f$ on $E \in \mathcal{M}$ and E^c is μ-null. Then
$$\int f \underset{\text{(MCT3)}}{=} \int f\chi_E \underset{\text{(MCT)}}{=} \lim \int f_n\chi_E \underset{\text{(MCT3)}}{=} \lim \int f_n.$$
□

(MCT6) (Fatou's Lemma) If $(f_n) \subset L^+$, then $\int \liminf f_n \leq \liminf \int f_n$.

Proof. For all $j \geq k \in \mathbb{N}$, $\inf_{n\geq k} f_n \leq f_j$, so
$$\int \inf_{n\geq k} f_n \leq \int f_j \qquad \text{for all} \qquad j \geq k.$$
Thus $\int \inf_{n\geq k} f_n \leq \inf_{j\geq k} \int f_j$. Letting $k \to \infty$, we have
$$\int \liminf f_n \underset{\text{(MCT)}}{=} \lim_{k\to\infty} \int \inf_{n\geq k} f_n \leq \lim_{k\to\infty} \inf_{j\geq k} \int f_j = \liminf \int f_n.$$
□

(MCT7) If $(f_n) \subset L^+$, $f_n \to f$ a.e., and $f \in L^+$ (which is automatic if μ is complete), then $\int f \leq \liminf \int f_n$.

Proof. Let $E \in \mathcal{M}$ such that $f_n \to f$ on E and E^c is μ-null. Then
$$\int f \underset{(3)}{=} \int f\chi_E \underset{\text{(MCT6)}}{\leq} \liminf \int f_n\chi_E \underset{\text{(MCT3)}}{=} \liminf \int f_n.$$
□

Exercise 3.3.11. Assume Fatou's Lemma (MCT6) and prove the Monotone Convergence Theorem from it.

Exercise 3.3.12. If $f \in L^+$ and $\int f < \infty$, then $\{f = \infty\}$ is μ-null and $\{0 < f\}$ is σ-finite.

Exercise 3.3.13. Suppose $(X, \mathcal{M}, \mu)$ is a finite measure space and $f : X \to \mathbb{C}$ is measurable. Prove that $\mu(\{n \leq |f|\}) \to 0$ as $n \to \infty$.

Exercise 3.3.14. Suppose $(X, \mathcal{M}, \mu)$ is a measure space and $f \in L^+$. For $E \in \mathcal{M}$, define

$$\nu(E) := \int_E f \, d\mu.$$

(1) Prove that ν is a measure on $\mathcal{M}$.
(2) Prove that $\int g \, d\nu = \int fg \, d\mu$ for all $g \in L^+$
Hint: First suppose g is simple.

3.4. **Integration of $\overline{\mathbb{R}}$-valued functions.** For this section, $(X, \mathcal{M}, \mu)$ is a fixed measure space.

Definition 3.4.1. An $\mathcal{M}$-measurable function $f : X \to \overline{\mathbb{R}}$ is called *integrable* if $\int f_\pm < \infty$ where $f = f_+ - f_-$ with $f_+ = 0 \vee f$ and $f_- = -(0 \wedge f)$. Since $|f| = f_+ + f_-$, observe that f is integrable if and only if $\int |f| < \infty$.

Define $L^1(\mu, \mathbb{R}) := \{\text{integrable } f : X \to \mathbb{R}\}$.

Exercise 3.4.2. Show that a simple function $\psi = \sum_{k=1}^n c_k \chi_{E_k} \in \mathsf{SF}$ with $c_1, \ldots, c_n$ distinct and $E_1, \ldots, E_n$ disjoint is integrable if and only if $\mu(E_k) < \infty$ for all k such that $c_k \neq 0$.

Proposition 3.4.3. *The set $L^1(\mu, \mathbb{R})$ is an $\mathbb{R}$-vector space. Moreover, $\int : L^1(\mu, \mathbb{R}) \to \mathbb{R}$ given by $\int f := \int f_+ - \int f_-$ is a linear functional.*

Proof. If $r \in \mathbb{R}$ and $f, g \in L^1(\mu, \mathbb{R})$, then $|rf + g| \leq |r| \cdot |f| + |g|$ which is integrable. Hence $L^1(\mu, \mathbb{R})$ is an $\mathbb{R}$-vector space.

If $r \in \mathbb{R}$ and $f \in L^1(\mu, \mathbb{R})$, then there are three cases:

$$(rf)_\pm = \begin{cases} rf_\pm & \text{if } r > 0 \\ 0 & \text{if } r = 0 \\ -rf_\mp & \text{if } r < 0. \end{cases}$$

In all three cases, by Remarks 3.3.7(3), we have

$$\int rf = \int (rf)_+ - \int (rf)_- = \left\{\begin{matrix} \displaystyle\int rf_+ - \int rf_- & \text{if } r > 0 \\ 0 & \text{if } r = 0 \\ \displaystyle\int (-r)f_- - \int (-r)f_+ & \text{if } r < 0 \end{matrix}\right\} = r\int f_+ - r\int f_-.$$

If $f, g \in L^1(\mu, \mathbb{R})$, observe

$$(f+g)_+ - (f+g)_- = f + g = f_+ + g_+ - f_- - g_-$$

which implies

$$(f+g)_+ + f_- + g_- = (f+g)_- + f_+ + g_+.$$

By (MCT2),

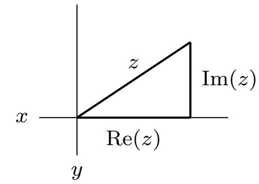
$$\int (f+g)_+ + \int f_- + \int g_- = \int (f+g)_- + \int f_+ + \int g_+,$$

and rearranging yields the result. □

3.5. **Integration of $\mathbb{C}$-valued functions.** For this section, fix a measure space $(X, \mathcal{M}, \mu)$. Recall from Exercise 3.1.14(2) that $f : X \to \mathbb{C}$ is measurable if and only if $\operatorname{Re}(f)$ and $\operatorname{Im}(f)$ are measurable. By Exercise 3.1.20(2), $|f|$ is measurable.

Definition 3.5.1. A measurable function $f : X \to \mathbb{C}$ is *integrable* if $\int |f| < \infty$, i.e., $|f| \in L^1(\mu, \mathbb{R})$. Since

$$|f| \le |\operatorname{Re}(f)| + |\operatorname{Im}(f)| \le 2|f|$$



,

f is integrable if and only if $\operatorname{Re}(f)$ and $\operatorname{Im}(f)$ are integrable. In this case, we define

$$\int f := \int \operatorname{Re}(f) + i \int \operatorname{Im}(f).$$

It follows from Proposition 3.4.3 that

$$L^1(\mu, \mathbb{C}) := \{\text{integrable } f : X \to \mathbb{C}\}$$

is a $\mathbb{C}$-vector space, and $\int : L^1(\mu, \mathbb{C}) \to \mathbb{C}$ is linear.

Proposition 3.5.2. *For all $f \in L^1(\mu, \mathbb{C})$, $\left|\int f\right| \le \int |f|$.*

Proof.
Step 1: If f is $\mathbb{R}$-valued, then $\left|\int f\right| = \left|\int f_+ - \int f_-\right| \le \int f_+ + \int f_- = \int |f|$.
Step 2: Suppose f is $\mathbb{C}$-valued. We may assume $\int f \ne 0$. We use the following trick:

Trick. Define $\operatorname{sgn}\left(\int f\right) := \frac{\int f}{\left|\int f\right|} \in \mathbb{T} := \{z \in \mathbb{C} \,|\, |z| = 1\}$. Then since $z^{-1} = \overline{z}$ for all $z \in \mathbb{T}$,

$$\left|\int f\right| = \overline{\operatorname{sgn}\left(\int f\right)} \int f = \underbrace{\int \overline{\operatorname{sgn}\left(\int f\right)} f}_{\in \mathbb{R}}.$$

We then calculate

$$\left|\int f\right| = \int \overline{\operatorname{sgn}\left(\int f\right)} f = \operatorname{Re} \int \overline{\operatorname{sgn}\left(\int f\right)} f = \int \operatorname{Re}\left(\overline{\operatorname{sgn}\left(\int f\right)} f\right)$$

$$\underset{\text{(Step 1)}}{\le} \int \left|\operatorname{Re}\left(\overline{\operatorname{sgn}\left(\int f\right)} f\right)\right| \le \int \left|\underbrace{\overline{\operatorname{sgn}\left(\int f\right)}}_{\in \mathbb{T}} f\right| = \int |f|.$$

□

Corollary 3.5.3. *For all $f, g \in L^1(\mu, \mathbb{C})$, the following are equivalent:*

(1) $f = g$ a.e.
(2) $\int |f - g| = 0$
(3) for all $E \in \mathcal{M}$, $\int_E f = \int_E g$.

Proof.
(1) ⇔ (2) Observe $f = g$ a.e. if and only if $|f - g| = 0$ a.e. if and only if $\int |f - g| = 0$ by Proposition 3.3.8.
(2) ⇒ (3) By Proposition 3.5.2, for all $E \in \mathcal{M}$,

$$\left|\int_E f - \int_E g\right| = \left|\int (f-g)\chi_E\right| \leq \int |f-g|\chi_E \leq \int |f-g| = 0.$$

(3) ⇒ (1) Recall that $\int_E f - g = \int_E \operatorname{Re}(f-g) + i\int_E \operatorname{Im}(f-g)$. So by assumption,

$$\int_E \operatorname{Re}(f-g) = 0 \qquad \text{and} \qquad \int_E \operatorname{Im}(f-g) = 0 \qquad\qquad \forall\, E \in \mathcal{M}.$$

We now look at the following particular $E \in \mathcal{M}$:

$$\begin{aligned} E &= \{0 \leq \operatorname{Re}(f-g)\} &&\Rightarrow && \operatorname{Re}(f-g)_+ = 0 \text{ a.e.}\\ E &= \{0 \geq \operatorname{Re}(f-g)\} &&\Rightarrow && \operatorname{Re}(f-g)_- = 0 \text{ a.e.}\\ E &= \{0 \leq \operatorname{Im}(f-g)\} &&\Rightarrow && \operatorname{Im}(f-g)_+ = 0 \text{ a.e.}\\ E &= \{0 \geq \operatorname{Im}(f-g)\} &&\Rightarrow && \operatorname{Im}(f-g)_- = 0 \text{ a.e.}\end{aligned}$$

Hence $\operatorname{Re}(f-g) = 0$ and $\operatorname{Im}(f-g) = 0$ a.e., which is equivalent to $f = g$ a.e. □

Exercise 3.5.4. Suppose $(X, \mathcal{M}, \mu)$ be a measure space and $f \in L^1(\mu, \mathbb{C})$. Prove that $\{f \neq 0\}$ is σ-finite.

Exercise 3.5.5. Suppose $(X, \mathcal{M}, \mu)$ is a measure space and $f \in L^1(\mu, \mathbb{C})$. Prove that for every $\varepsilon > 0$, there exists a $\delta > 0$ such that for every $E \in \mathcal{M}$ with $\mu(E) < \delta$, $\int_E |f| < \varepsilon$.

Definition 3.5.6. Define $\mathcal{L}^1(\mu, \mathbb{C}) := L^1(\mu, \mathbb{C})/\sim$ where $f \sim g$ if and only if $f = g$ a.e. We write $f \in \mathcal{L}^1(\mu, \mathbb{C})$ to mean $f \in L^1(\mu, \mathbb{C})$ representing its equivalence class in $\mathcal{L}^1(\mu, \mathbb{C})$.

Exercise 3.5.7. Let $(X, \mathcal{M}, \mu)$ be a measure space.

(1) Prove that $\|\cdot\|_1 : \mathcal{L}^1(\mu, \mathbb{C}) \to [0, \infty)$ given by $\|f\|_1 := \int |f|$ is a norm.
(2) Let $(X, \overline{\mathcal{M}}, \overline{\mu})$ be the completion of $(X, \mathcal{M}, \mu)$. Find a canonical $\mathbb{C}$-vector space isomorphism $\mathcal{L}^1(\mu, \mathbb{C}) \cong \mathcal{L}^1(\overline{\mu}, \mathbb{C})$ which preserves $\|\cdot\|_1$.
Hint: Use Exercise 3.2.5.

Theorem 3.5.8 (Dominated Convergence, a.k.a. DCT)**.** *Suppose $(f_n) \subset \mathcal{L}^1(\mu, \mathbb{C})$ such that $f_n \to f$ a.e. If there is a $g \in L^1(\mu, \mathbb{C}) \cap L^+$ such that eventually $|f_n| \leq g$ a.e., then $f \in \mathcal{L}^1(\mu, \mathbb{C})$ and $\int f = \lim \int f_n$.*

Proof. By redefining f on a μ-null set if necessary by Exercise 3.2.5, we may assume f is $\mathcal{M}$-measurable. Taking limits pointwise, $|f| \leq g$, so $f \in L^1(\mu, \mathbb{C})$. Taking real and imaginary parts of f, we may assume $(f_n), f$ are all $\mathbb{R}$-valued. Then $-g \leq f_n \leq g$ a.e., so

$$g + f_n \geq 0 \qquad \text{and} \qquad g - f_n \geq 0 \qquad\qquad \text{a.e.}$$

By Fatou's Lemma (MCT6),

$$\begin{aligned}\int g + \int f = \int g + f \leq \liminf \int g + f_n = \int g + \liminf \int f_n\\ \int g - \int f = \int g - f \leq \liminf \int g - f_n = \int g - \limsup \int f_n.\end{aligned}$$

Combining these inequalities,

$$\limsup \int f_n \leq \int f \leq \liminf \int f_n. \qquad \square$$

Corollary 3.5.9. *Suppose $(f_n) \subset \mathcal{L}^1(\mu, \mathbb{C})$ such that $\sum \int |f_n| < \infty$. Then $\sum f_n$ converges a.e. to a function in $\mathcal{L}^1(\mu, \mathbb{C})$, and $\int \sum f_n = \sum \int f_n$.*

Exercise 3.5.10. Suppose $(X, \mathcal{M}, \mu)$ is a measure space, and $(f_n) \subset L^+$ is a decreasing sequence of non-negative $\mathcal{M}$-measurable functions, i.e., $f_n \geq f_{n+1}$ for all $n \in \mathbb{N}$.

(1) Find an example of such a sequence such that $\int f_n$ does not converge to $\int f$.
(2) Suppose $\int f < \infty$. Find a necessary and sufficient condition so that $\int f_n \to \int f$ as $n \to \infty$.

Exercise 3.5.11. Suppose $(X, \mathcal{M}, \mu)$ is a finite measure space and $f : X \to [0, \infty]$ is measurable. Prove that $f \in \mathcal{L}^1(\mu)$ if and only if $\sum_{n=1}^{\infty} \mu(\{n < f\}) < \infty$.

Exercise 3.5.12. Prove that the metric d_1 on $\mathcal{L}^1(\mu, \mathbb{C})$ induced by $\|\cdot\|_1$ is complete. That is, prove every Cauchy sequence converges in $\mathcal{L}^1$.
Note: This follows immediately from Corollary 3.5.9 if one shows that completeness of a normed vector space V is equivalent to the property that every absolutely convergent series converges in V.

Exercise 3.5.13. Let μ be a Lebesgue-Stieltjes Borel measure on $\mathbb{R}$. Show that $C_c(\mathbb{R})$, the continuous functions of compact support ($\overline{\{f \neq 0\}}$ compact) is dense in $\mathcal{L}^1(\mu, \mathbb{R})$. Does the same hold for $\mathbb{C}$-valued functions?
Hint: You could proceed in this way:

(1) Reduce to the case $f \in L^1 \cap L^+$.
(2) Reduce to the case $f \in L^1 \cap \mathsf{SF}^+$.
(3) Reduce to the case $f = \chi_E$ with $E \in \mathcal{B}_{\mathbb{R}}$ and $\mu(E) < \infty$.
(4) Reduce to the case $f = \chi_U$ with $U \subset \mathbb{R}$ open and $\mu(U) < \infty$.
(5) Reduce to the case $f = \chi_{(a,b)}$ with $a < b$ in $\mathbb{R}$.

3.6. **Modes of convergence.** Let $(X, \mathcal{M}, \mu)$ be a measure space. For $(f_n), f$ all $\mathcal{M} - \mathcal{B}_{\mathbb{C}}$ measurable functions, $f_n \to f$ could mean many things:

- (pointwise) $f_n(x) \to f(x)$ for all $x \in X$.
- (a.e.) $f_n(x) \to f(x)$ for a.e. $x \in X$.
- (uniformly) for all $\varepsilon > 0$, there is an $N \in \mathbb{N}$ such that $n > N$ implies $|f_n(x) - f(x)| < \varepsilon$ for all $x \in X$.
- (almost uniformly, a.k.a. a.u.) for all $\varepsilon > 0$, there is an $E \in \mathcal{M}$ with $\mu(E) < \varepsilon$ such that $f_n \chi_{E^c} \to f \chi_{E^c}$ uniformly.
- (in $\mathcal{L}^1$) $\int |f_n - f| \to 0$ as $n \to \infty$.
- (in measure) for all $\varepsilon > 0$, $\mu(\{\varepsilon \leq |f - f_n|\}) \to 0$.

Observe that obviously uniform implies a.u., uniform implies pointwise, and pointwise implies a.e.

Proposition 3.6.1. *Almost uniform convergence implies almost everywhere convergence.*

Proof. Suppose $f_n \to f$ a.u. For $k \in \mathbb{N}$, let $E_k \in \mathcal{M}$ such that $\mu(E_k) < 1/k$ and $f_n \chi_{E_k^c} \to f \chi_{E_k^c}$ uniformly. Let $E := \bigcap E_k$. Then $\mu(E) = 0$ by continuity from above (μ4), and since $E^c = \bigcup E_k^c$, we have $f_n \chi_{E^c} \to f \chi_{E^c}$ pointwise. $\square$

Proposition 3.6.2. *Almost uniform convergence implies convergence in measure.*

Proof. Suppose $f_n \to f$ a.u. Let $\varepsilon > 0$. Show for all $\delta > 0$, there is an $N \in \mathbb{N}$ such that $n > N$ implies $\mu(\{\varepsilon \le |f - f_n|\}) < \delta$. Pick $E \in \mathcal{M}$ such that $\mu(E) < \delta$ and $f_n \chi_{E^c} \to f\chi_{E^c}$ uniformly. Then

$$\mu(\{\varepsilon \le |f - f_n|\}) = \underbrace{\mu(\{\varepsilon \le |f - f_n|\} \cap E)}_{\text{always } < \delta} + \underbrace{\mu(\{\varepsilon \le |f - f_n|\} \cap E^c)}_{= \emptyset \text{ for } n \text{ large}} < \delta$$

for n sufficiently large. □

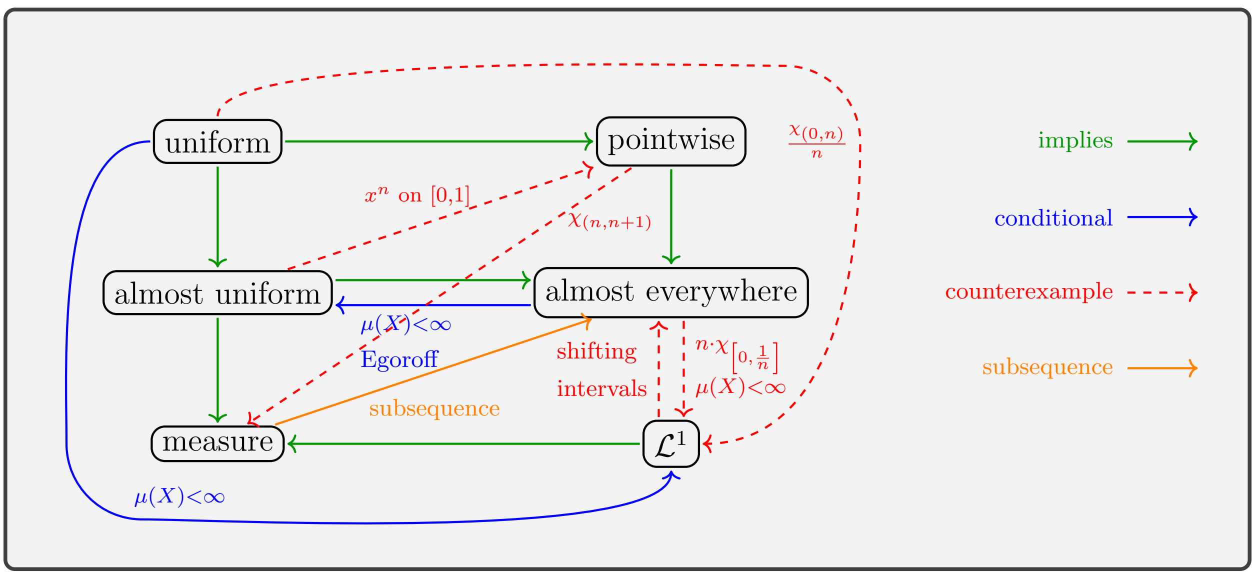
Proposition 3.6.3. *Convergence in $\mathcal{L}^1$ implies convergence in measure.*

Proof. Suppose $f_n \to f$ in $\mathcal{L}^1$. Let $\varepsilon > 0$, and set $E := \{\varepsilon \le |f - f_n|\}$. Then

$$\mu(E) = \int_E 1 = \frac{1}{\varepsilon}\int_E \varepsilon \le \frac{1}{\varepsilon}\int_E |f - f_n| \xrightarrow{n\to\infty} 0.$$ □

Facts 3.6.4 (Counterexamples)**.** We consider the following important counterexamples:

(1) $f_n = \frac{1}{n}\chi_{(0,n)}$ converges uniformly to zero, but not in $\mathcal{L}^1$.
(2) $f_n = \chi_{(n,n+1)}$ converges pointwise to zero, but not in measure.
(3) $f_n = n\chi_{[0,1/n]}$ converges a.e. to zero with $\mu(X) < \infty$, but not in $\mathcal{L}^1$.
(4) $f_n(x) := x^n$ on $[0,1]$ almost uniformly to zero, but not pointwise.
(5) (shifting intervals) $f_1 = \chi_{[0,1]}$, $f_2 = \chi_{[0,1/2]}$, $f_3 = \chi_{[1/2,1]}$, $f_4 = \chi_{[1,1/4]}$, $f_5 = \chi_{[1/4,1/2]}$, etc. converges in $\mathcal{L}^1$, but not a.e.



Lemma 3.6.5. *If $f_n \to f$ uniformly and $\mu(X) < \infty$, then $f_n \to f$ in $\mathcal{L}^1$.*

Proof. Observe that

$$\int |f_n - f| \le (\sup|f_n - f|) \cdot \int 1 = \underbrace{(\sup|f_n - f|)}_{\to 0 \text{ as } n\to\infty} \cdot \mu(X).$$ □

Theorem 3.6.6 (Egoroff)**.** *If $f_n \to f$ a.e. and $\mu(X) < \infty$, then $f_n \to f$ a.u.*

Proof. By replacing X with $X \setminus N$ for some μ-null set $N \in \mathcal{M}$, we may assume $f_n \to f$ pointwise. Now observe that for all $k \in \mathbb{N}$,

$$E_{n,k} := \bigcup_{j=n}^{\infty} \left\{ \frac{1}{k} \leq |f - f_j| \right\} \searrow \emptyset \qquad \text{as} \qquad n \to \infty.$$

Since $\mu(X) < \infty$, by continuity from above (μ4), $\mu(E_{n,k}) \to 0$ as $n \to \infty$. Let $\varepsilon > 0$. For all $k \in \mathbb{N}$, choose $n_k \in \mathbb{N}$ such that $\mu(E_{n_k,k}) < \varepsilon/2^k$. Setting $E := \bigcup_{k=1}^{\infty} E_{n_k,k}$, we have

$$\mu(E) \leq \sum_k \mu(E_{n_k,k}) < \varepsilon \sum 2^{-k} = \varepsilon.$$

Finally, observe that for all $n > n_k$, if $x \in E^c = \bigcap_{k=1}^{\infty} E_{n_k,k}^c$, then $|f(x) - f_n(x)| < 1/k$. Thus $f_n \to f$ uniformly on E^c. □

Definition 3.6.7. A sequence (f_n) of $\mathcal{M}$-measurable functions is *Cauchy in measure* if for all $\varepsilon > 0$,

$$\mu\left(\{\varepsilon \leq |f_m - f_n|\}\right) \xrightarrow{n,m\to\infty} 0.$$

Exercise 3.6.8. Prove that if $f_n \to f$ in measure, then (f_n) is Cauchy in measure.

Theorem 3.6.9. *If (f_n) is Cauchy in measure, then there exists a unique (up to μ-null set) $\mathcal{M}$-measurable function f such that $f_n \to f$ in measure. Moreover, there is a subsequence (f_{n_k}) such that $f_{n_k} \to f$ a.e.*

Proof.
Step 1: There is a subsequence (f_{n_k}) such that $\mu\left(\{2^{-k} \leq |f_{n_k} - f_{n_{k+1}}|\}\right) < 2^{-k}$.

> *Proof.* For all $k \in \mathbb{N}$, $\mu\left(\{2^{-k} \leq |f_n - f_m|\}\right) \to 0$ as $m, n \to \infty$. Pick n_k inductively so $n_{k+1} > n_k$ and $m, n \geq n_k$ implies $\mu\left(\{2^{-k} \leq |f_n - f_m|\}\right) < 2^{-k}$. □

Step 2: (f_{n_k}) is pointwise Cauchy off a μ-null set N.

> *Proof.* For $k \in \mathbb{N}$, set $E_k := \{2^{-k} \leq |f_{n_k} - f_{n_{k+1}}|\}$, and for $\ell \in \mathbb{N}$, set $N_\ell := \bigcup_{k=\ell} E_k$. Then $\mu(N_\ell) \leq \sum_{k=\ell} 2^{-k} = 2^{1-\ell}$. Setting $N = \bigcap N_\ell = \limsup E_k$, we have $\mu(N) = 0$ by continuity from above (μ4). If $x \in N^c$, then $x \notin N_\ell$ for some ℓ, and thus for all $\ell \leq i \leq j$,
>
> $$|f_{n_i}(x) - f_{n_j}(x)| \leq \sum_{k=i}^{j-1} |f_{n_k}(x) - f_{n_{k+1}}(x)| \leq \sum_{k=i}^{j-1} 2^{-k} \leq 2^{1-i}. \tag{3.6.10}$$
>
> We conclude that (f_{n_k}) is pointwise Cauchy on N^c. □

Step 3: Define

$$f(x) := \begin{cases} 0 & \text{if } x \in N \text{ (which is } \mu\text{-null)} \\ \lim_k f_{n_k}(x) & \text{if } x \in N^c. \end{cases}$$

Then f is $\mathcal{M}$-measurable and $f_{n_k} \to f$ a.e.

Proof. It remains to show f is measurable. Observe $f_{n_k} \cdot \chi_{N^c}$ is $\mathcal{M}$-measurable for all k, and thus so is $f = \lim f_{n_k} \cdot \chi_{N^c}$ by Exercise 3.2.5. □

Step 4: $f_{n_k} \to f$ in measure.

Proof. For all $x \in N_\ell^c$ and $k \geq \ell$, we have

$$|f_{n_k}(x) - f(x)| = \lim_{j\to\infty} |f_{n_k}(x) - f_{n_j}(x)| \underset{(3.6.10)}{\leq} 2^{1-k}.$$

Let $\varepsilon > 0$ and pick $\ell \in \mathbb{N}$ such that $0 < 2^{-\ell} < \varepsilon$. Then for all $k \geq \ell$,

$$\mu\left(\{\varepsilon \leq |f_{n_k} - f|\}\right) \leq \mu\left(\left\{\frac{1}{2^k} \leq |f_{n_k} - f|\right\}\right) < 2^{1-k} \xrightarrow{k\to\infty} 0.$$ □

Step 5: $f_n \to f$ in measure.

Proof. We use the following trick:

Trick. For non-negative $\mathcal{M}$-measurable f, g, $\{a+b \leq f+g\} \subset \{a \leq f\} \cup \{b \leq g\}$.

Now observe that

$$\{\varepsilon \leq |f_n - f|\} \subseteq \underbrace{\left\{\frac{\varepsilon}{2} \leq |f_n - f_{n_k}|\right\}}_{\substack{\mu \to 0 \text{ as } (f_n) \\ \text{Cauchy in measure}}} \cup \underbrace{\left\{\frac{\varepsilon}{2} \leq |f_{n_k} - f|\right\}}_{\mu \to 0 \text{ by Step 4}}.$$

Hence $\mu(\{\varepsilon \leq |f_n - f|\}) \to 0$ as $n \to \infty$. □

Step 6: f is unique (up to a μ-null set) such that $f_n \to f$ in measure.

Proof. Suppose g is another such candidate. Then using the same trick as in Step 5,

$$\{\varepsilon \leq |f - g|\} \subseteq \underbrace{\left\{\frac{\varepsilon}{2} \leq |f - f_n|\right\}}_{\mu \to 0 \text{ as } n \to \infty} \cup \underbrace{\left\{\frac{\varepsilon}{2} \leq |g - f_n|\right\}}_{\mu \to 0 \text{ as } n \to \infty}.$$

Hence $\mu(\{\varepsilon \leq |f - g|\}) = 0$ for all $\varepsilon > 0$, and thus $f = g$ a.e. □

This concludes the proof. □

Exercise 3.6.11 (Lusin's Theorem)**.** Suppose $f : [a, b] \to \mathbb{C}$ is Lebesgue measurable and $\varepsilon > 0$. There is a compact set $E \subset [a, b]$ such that $\lambda(E^c) < \varepsilon$ and $f|_E$ is continuous.
Hint: Use Exercise 3.3.13 and Egoroff's Theorem 3.6.6.

Exercise 3.6.12. Suppose $f \in \mathcal{L}^1([0, 1], \lambda)$ is an integrable non-negative function.

(1) Show that for every $n \in \mathbb{N}$, $\sqrt[n]{f} \in \mathcal{L}^1([0, 1], \lambda)$.
(2) Show that $(\sqrt[n]{f})$ converges in $\mathcal{L}^1$ and compute its limit.

Hint for both parts: Consider $\{f \geq 1\}$ and $\{f < 1\}$ separately.

Exercise 3.6.13. Suppose $(X, \mathcal{M}, \mu)$ is a measure space and $f_n \to f$ in measure and $g_n \to g$ in measure (these functions are assumed to be measurable). Show that

(1) $|f_n| \to |f|$ in measure.
(2) $f_n + g_n \to f + g$ in measure.
(3) $f_n g_n \to fg$ if $\mu(X) < \infty$, but not necessarily if $\mu(X) = \infty$.
Hint: First show $f_n g \to fg$ in measure. To do so, one could follow the following steps.
  (a) Show that for any $\varepsilon > 0$, by Exercise 3.3.13, $X = E \amalg E^c$ where $|g|_E| < M$ and $\mu(E^c) < \varepsilon/2$.
  (b) For $\delta > 0$ and carefully chosen $M > 0$ and E,
$$\begin{aligned}\{|f_n g - fg| > \delta\} &= (\{|f_n g - fg| > \delta\} \cap E) \amalg (\{|f_n g - fg| > \delta\} \cap E^c)\\ &\subseteq \left\{|f_n - f| > \frac{\delta}{M}\right\} \cup E^c.\end{aligned}$$

Exercise 3.6.14 (Folland §2.4, #33 and 34)**.** Suppose $(X, \mathcal{M}, \mu)$ is a measure space and $f_n \to f$ in measure (these functions are assumed to be measurable).

(1) Show that if $f_n \geq 0$ everywhere, then $\int f \leq \liminf \int f_n$.
(2) Suppose $|f_n| \leq g \in \mathcal{L}^1$. Prove that $\int f = \lim \int f_n$ and $f_n \to f$ in $\mathcal{L}^1$.

Exercise 3.6.15. Let $(X, \mathcal{M}, \mu)$ be a measure space. Suppose $(E_n) \subset \mathcal{M}$ is a sequence of measurable sets with $\mu(E_n) < \infty$ for all n. Show that if $\chi_{E_n} \to f$ in $\mathcal{L}^1$ (this assumes f is $\mathcal{M}$-measurable), then there is an $E \in \mathcal{M}$ such that $f = \chi_E$ a.e.

3.7. **Comparison of the Lebesgue and Riemann integrals.** We now review the Riemann integral for a Reimann integrable function $f : [a, b] \to \mathbb{R}$.

Definition 3.7.1. A *partition* of $[a, b]$ is a set of points $P = \{a = s_0 < s_1 < \cdots < s_m = b\}$. We say an interval $J \in P$ if $J = [s_{i-1}, s_i]$ for some $i = 1, \ldots, m$. We write
$$m_J := \inf \{f(x)|x \in J\} \qquad\qquad M_J := \sup \{f(x)|x \in J\}.$$
We define the:

- Lower sum: $L(f, P) := \sum_{J \in P} m_J \lambda(J)$
- Upper sum: $U(f, P) := \sum_{J \in P} M_J \lambda(J)$

Here, $\lambda(J)$ is the length (Lebesgue measure) of the interval. Observe $L(f, P) \leq U(f, P)$.

A *refinement* of P is a partition $Q = \{a = t_0 < t_1 < \cdots < t_n = b\} \supset P$. Observe that if Q refines P, then
$$L(f, P) \leq L(f, Q) \leq U(f, Q) \leq U(f, P).$$
Thus if P_1, P_2 are two partitions of $[a, b]$ and Q is a common refinement, then
$$\sup_{i=1,2} L(f, P_i) \leq L(f, Q) \leq U(f, Q) \leq \inf_{i=1,2} U(f, P_i).$$
We define the:

- Upper integral: $\overline{\int}_{[a,b]} f := \inf_P U(f, P)$
- Lower integral: $\underline{\int}_{[a,b]} f := \sup_P L(f, P)$

We say f is *Riemann integrable* on $[a, b]$ if $\overline{\int}_{[a,b]} f = \underline{\int}_{[a,b]} f$, and we denote this common value by $\int_a^b f(x)\, dx$.

Exercise 3.7.2. Suppose $f : [a, b] \to \mathbb{R}$. Prove the following are equivalent:

(1) f is Riemann integrable
(2) for all $\varepsilon > 0$, there is a partition P of $[a, b]$ such that $U(f, P) - L(f, P) < \varepsilon$.

Theorem 3.7.3. *If f is Riemann integrable on $[a, b]$, then f is Lebesgue integrable and $\int_{[a,b]} f\, d\lambda = \int_a^b f(x)\, dx$.*

Proof. Let (P_n) be a sequence of partitions of $[a, b]$ such that P_{n+1} refines P_n and $U(f, P_n) - L(f, P_n) < 1/n$ for all $n \in \mathbb{N}$. Here's the trick:

> **Trick.** Define the simple functions $\psi_n := \sum_{J\in P_n} m_J \chi_J$ and $\Psi_n := \sum_{J\in P_n} M_J \chi_J$.

Observe that $L(f, P_n) = \int \psi_n\, d\lambda$ and $U(f, P_n) = \int \Psi_n\, d\lambda$ and

$$\psi_n \le \psi_{n+1} \le f \le \Psi_{n+1} \le \Psi_n \qquad \forall\, n \in \mathbb{N}.$$

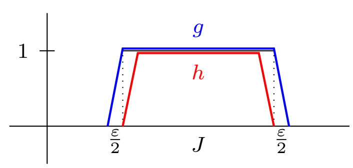
Define $\psi := \lim \psi_n$ and $\Psi := \lim \Psi_n$, which exists as (ψ_n) and (Ψ_n) are pointwise bounded and monotone. Then by (a slight modification of) the MCT 3.3.9, ψ, Ψ are integrable, and

$$\int \psi = \lim \int \psi_n = \int_a^b f(x)\, dx = \lim \int \Psi_n = \int \Psi.$$

But since $\Psi - \psi \ge 0$ everywhere, $\int \Psi - \psi = 0$ implies $\Psi = f = \psi$ a.e. So $f \in \mathcal{L}^1$ and $\int f = \int_a^b f(x)\, dx$. □

Lemma 3.7.4. *Suppose $f : [a, b] \to \mathbb{R}$ is Riemann integrable and bounded. Then for all $\varepsilon > 0$, there are continuous functions $g, h : [a, b] \to \mathbb{R}$ such that $h \le f \le g$ and $\int_{[a,b]}(g-h)\, d\lambda \le \varepsilon$.*

Proof.
Step 1: If $f = \chi_J$ for some interval J, then we can find piecewise linear functions g, h such that $h \le f \le g$ such as in the following cartoon:



Then $\int_{[a,b]} g = \lambda(J) + \varepsilon/2$ and $\int_{[a,b]} h = \lambda(J) - \varepsilon/2$, so $\int g - h = \varepsilon$.

Step 2: Without loss of generality, we may assume $f \ge 0$. (Otherwise, treat $f_\pm$ separately.) Take a partition P of $[a, b]$ such that $U(f, P) - L(f, P) < \varepsilon/2$. As in the trick in the previous theorem, define the simple functions

$$\psi_n := \sum_{J\in P} m_J \chi_J \le f \le \Psi_n := \sum_{J\in P} M_J \chi_J$$

so that $\int \psi = L(f, P)$ and $\int \Psi = U(f, P)$. Apply Step 1 to each χ_J to get continuous g_J, h_J with $h_J \le \chi_J \le g_J$ such that $\int g_J - h_J < \frac{\varepsilon}{2|P|M}$ where $|P|$ is the number of intervals of P and $M := \sup\{f(x) | a \le x \le b\}$. Setting $g := \sum_{J\in P} M_J g_J$ and $h := \sum_{J\in P} m_J h_J$, we have

$$h = \sum_{J\in P} m_J h_J \le \sum_{J\in P} m_J \chi_J = \psi \le f \le \psi = \sum_{J\in P} M_J \chi_J \le \sum_{J\in P} M_J g_J = g,$$

and thus

$$\begin{aligned}
\int g - h &= \sum_{J\in P} M_J \int g_J - m_J \int h_J \\
&= U(f,P) - L(f,P) + \sum_{J\in P} \underbrace{M_J}_{<M} \left(\int g_J - \lambda(J)\right) + \underbrace{m_J}_{<M}\left(\lambda(J) - \int h_J\right) \\
&< \underbrace{U(f,P) - L(f,P)}_{<\frac{\varepsilon}{2}} + M \sum_{J\in P} \underbrace{\int g_J - h_J}_{<\frac{\varepsilon}{2|P|M}} \\
&< \varepsilon. \qquad \square
\end{aligned}$$

Exercise 3.7.5. Let X be a topological space and let $g : X \to \mathbb{R}$. We say that g is *upper semicontinuous* at $x_0 \in X$ if for every $\varepsilon > 0$, there is an open neighborhood U of x_0 such that $x \in U$ implies $f(x) < f(x_0) + \varepsilon$. We say g is upper semicontinuous if g is upper semicontinuous at every $x \in X$.

(1) Show that g is upper semicontinuous if and only if $\{g < r\}$ is open in for all $r \in \mathbb{R}$.
(2) Define lower semicontinuity (both at $x_0 \in X$ and everywhere) and prove the analogous statement to (1).

Theorem 3.7.6 (Lebesgue)**.** *A bounded function $f : [a,b] \to \mathbb{R}$ is Riemann integrable if and only if it is continuous a.e.*

Proof.
$\Rightarrow$: Suppose f is Riemann integrable. By Lemma 3.7.4, there are sequence of continuous functions (h_n) and (g_n) on $[a,b]$ with $h_n \leq f \leq g_n$ such that $\int g_n - h_n < 1/n$ for all $n \in \mathbb{N}$. Since

$$g_{n+1} \wedge g_n - h_{n+1} \vee h_n \leq g_{n+1} - h_{n+1} \qquad \forall\, n \in \mathbb{N},$$

we may assume that

$$h_n \leq h_{n+1} \leq f \leq g_{n+1} \leq g_n \qquad \forall\, n \in \mathbb{N}.$$

Setting $h := \lim h_n$ and $g := \lim g_n$, we have $h \leq f \leq g$ and $\int h = \int f = \int g$ by MCT 3.3.9. Since $g - h \geq 0$, we know $g = f = h$ a.e. on $[a,b]$.

Claim. *Since $g_n \searrow g$, g is upper semicontinuous. Similarly, h is lower semicontinuous*

Proof. Let $x_0 \in [a,b]$ and $\varepsilon > 0$. Pick $N \in \mathbb{N}$ such that $n \geq N$ implies $g_n(x_0) - g(x_0) < \varepsilon/2$. Pick $\delta > 0$ such that $x \in (x_0 - \delta, x_0 + \delta) \cap [a,b]$ implies $|g_N(x) - g_N(x_0)| < \varepsilon/2$. Then for all $x \in (x_0 - \delta, x_0 + \delta) \cap [a,b]$,

$$g(x_0) > g_N(x_0) - \frac{\varepsilon}{2} > g_N(x) - \varepsilon \geq g(x) - \varepsilon. \qquad \square$$

Whenever $h(x_0) = f(x_0) = g(x_0)$, f is both upper semicontinuous and lower semicontinuous at x_0, i.e., f is continuous at x_0. This happens on $[a,b]$ a.e.

$\Leftarrow$: Suppose f is continuous on $[a,b]$ a.e. Let E be the λ-null set of discontinuities, and let $\varepsilon > 0$. We'll construct a partition P such that $U(f,P) - L(f,P) < \varepsilon$. By outer regularity of λ, there is an open $U \supset E$ such that $\lambda(U) < \varepsilon'$ to be determined later. Let

$K := [a,b] \setminus U$, which is compact, and observe that f is continuous at all points of K (not $f|_K$!). For each $x \in K$, pick $\delta_x > 0$ such that $y \in [a,b]$ (not K!) and $|x-y| < \delta_x$ implies $|f(x) - f(y)| < \varepsilon'$. Then $\{B_{\delta_x/2}(x)\}_{x\in K}$ is an open cover of K, so there are $x_1, \ldots, x_n \in K$ such that $K \subset \bigcup_{i=1}^n B_{\delta_{x_i}/2}(x_i)$. Set $\delta := \min\{\delta_{x_i}/2 | i = 1, \ldots, n\}$.

Claim. *If $x \in K$ and $y \in [a,b]$ and $|x-y| < \delta/2$, then $|f(x) - f(y)| < 2\varepsilon'$.*

Proof. Without loss of generality, $x \in B_{\delta_1/2}(x_1)$. Then $y \in B_{\delta_1}(x_1)$, and thus

$$|f(x) - f(y)| \le |f(x) - f(x_1)| + |f(x_1) - f(y)| < 2\varepsilon'.$$ □

Let P be any partition of $[a,b]$ whose intervals have length at most δ. Let P' consist of the intervals that intersect K and let P'' be the intervals that do not intersect K. By the claim, if $J \in P'$, then $M_J - m_j \le 4\varepsilon'$. Thus

$$\begin{aligned}
U(f,P) - L(f,P) &= \sum_{J\in P} (M_J - m_J)\lambda(J) \\
&= \sum_{J\in P'} (M_J - m_J)\lambda(J) + \sum_{J\in P''} (M_J - m_J)\lambda(J) \\
&\le \sum_{J\in P'} 4\varepsilon'\lambda(J) + \sum_{J\in P''} (M - m)\lambda(J) \\
&\le 4\varepsilon'(b-a) + (M-m)\lambda(U) && \left(\bigcup_{J\in P''} J \subseteq U\right) \\
&< \varepsilon'(4(b-a) + (M-m))
\end{aligned}$$

where $M = \sup_{x\in[a,b]} f(x)$ and $m := \inf_{x\in[a,b]} f(x)$. Taking $\varepsilon' = \varepsilon/(4(b-a) + (M-m))$ works. □

3.8. **Product measures.**

Definition 3.8.1. Given measurable spaces $(X, \mathcal{M})$ and $(Y, \mathcal{N})$, a *measurable rectangle* is a set of the form $E \times F \subset X \times Y$ where $E \in \mathcal{M}$ and $F \in \mathcal{N}$. The *product σ-algebra* $\mathcal{M} \times \mathcal{N} \subset P(X \times Y)$ is the σ-algebra generated by the measurable rectangles.

Exercise 3.8.2. Prove that $\mathcal{M} \times \mathcal{N}$ is the smallest σ-algebra such that the canonical projection maps $\pi_X : X \times Y \to Y$ and $\pi_Y : X \times Y \to X$ are measurable. Deduce that $\mathcal{M} \times \mathcal{N}$ is generated by $\pi_X^{-1}(\mathcal{E}_X) \cup \pi_Y^{-1}(\mathcal{E}_Y)$ for any generating sets $\mathcal{E}_X$ of $\mathcal{M}$ and $\mathcal{E}_Y$ of $\mathcal{N}$.

Warning 3.8.3. Recall that given topological spaces X, Y, the canonical projections $\pi_X : X \times Y \to X$ and $\pi_Y : X \times Y \to Y$ are open maps. When $(X, \mathcal{M}), (Y, \mathcal{N})$ are measurable, however, π_X, π_Y need not map measurable sets to measurable sets. (Unfortunately, actually constructing a set in $\mathcal{M} \times \mathcal{N}$ whose projection to X is not measurable is quite difficult.)

Exercise 3.8.4. Show that the subset of $P(X \times Y)$ consisting of finite disjoint unions of measurable rectangles is an algebra which generates $\mathcal{M} \times \mathcal{N}$.
Hint: For $E, E_1, E_2 \in \mathcal{M}$ and $F, F_1, F_2 \in \mathcal{N}$,

- $(E_1 \times F_1) \cap (E_2 \times F_2) = (E_1 \cap E_2) \times (F_1 \cap F_2)$, *and*
- $(E \times F)^c = (E \times F^c) \amalg (E^c \times F) \amalg (E^c \times F^c)$.

Proposition 3.8.5. *Suppose (X, d_X) and (Y, d_Y) are metric spaces.*

(1) *$\mathcal{B}_X \times \mathcal{B}_Y$ is generated by $(\mathcal{T}_X \times Y) \cup (X \times \mathcal{T}_Y)$.*
(2) *$\mathcal{B}_X \times \mathcal{B}_Y \subset \mathcal{B}_{X\times Y}$.*
(3) *If X, Y are separable, then $\mathcal{B}_X \times \mathcal{B}_Y = \mathcal{B}_{X\times Y}$.*

Proof.
(1) This is an immediate consequence of Exercise 3.8.2.
(2) Since $\mathcal{T}_X \times Y, X \times \mathcal{T}_Y \subset \mathcal{T}_X \times \mathcal{T}_Y$, we have $\mathcal{B}_X \times \mathcal{B}_Y \subset \mathcal{B}_{X\times Y}$.
(3) Suppose $C \subset X$ and $D \subset Y$ are countable dense subsets. Let $\mathcal{E}_X, \mathcal{E}_Y$ be the collections of open balls centered at C, D respectively with rational radii. Note that $C \times D$ is a countable dense subset of $X \times Y$, and thus $\mathcal{T}_X \times \mathcal{T}_Y$ is generated by $\mathcal{E}_X \times \mathcal{E}_Y \subset \mathcal{B}_X \times \mathcal{B}_Y$. Hence $\mathcal{B}_{X\times Y} \subset \mathcal{B}_X \times \mathcal{B}_Y$. □

Exercise 3.8.6.

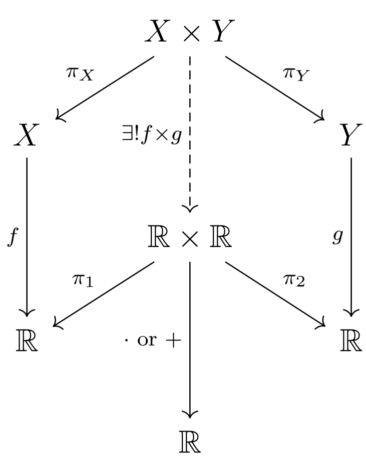
(1) Find an example of (non-separable) metric spaces X, Y such that $\mathcal{B}_X \times \mathcal{B}_Y \subsetneq \mathcal{B}_{X\times Y}$.
(2) If one of X or Y is separable, is $\mathcal{B}_X \times \mathcal{B}_Y = \mathcal{B}_{X\times Y}$? Find a proof or a counterexample.

Exercise 3.8.7. Suppose $(X, \mathcal{M}), (Y, \mathcal{N}), (Z, \mathcal{P})$ are measurable spaces and $f : Z \to X$ and $g : Z \to Y$. Show that $f \times g : Z \to X \times Y$ (the unique map from the universal property of the product) is measurable if and only if f and g are measurable. Deduce that the category of measurable spaces and measurable functions has finite categorical products.

Exercise 3.8.8. Prove that $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ are continuous and thus (Borel) measurable.

Corollary 3.8.9. *If $f : (X, \mathcal{M}) \to \mathbb{R}$ and $g : (Y, \mathcal{N}) \to \mathbb{R}$ are measurable, then so are $f + g$ and fg. (This also holds for other codomains such as $\mathbb{C}$ and $\overline{\mathbb{R}}$ if the sum is well-defined.)*

Proof. Observe that fg and $f + g$ are composites:



The composite of these measurable functions is $\mathcal{M} \times \mathcal{N}$-measurable. □

Exercise 3.8.10. Adapt the proof of Corollary 3.8.9 to give another proof that $f : (X, \mathcal{M}) \to \mathbb{C}$ is measurable if and only if $\mathrm{Re}(f), \mathrm{Im}(f)$ are measurable.

For the rest of this section, suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are measure spaces, and let $\mathcal{A}$ be the algebra of finite disjoint unions of measurable rectangles from Exercise 3.8.4.

Proposition 3.8.11. *For $G = \coprod_{k=1}^{n} E_k \times F_k \in \mathcal{A}$, define*

$$(\mu \times \nu)_0(G) := \sum_{k=1}^{n} \mu(E_k)\nu(F_k)$$

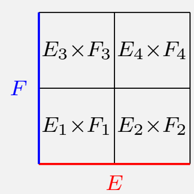
with the convention that $0 \cdot \infty = 0$. Then $(\mu \times \nu)_0$ is a premeasure on $\mathcal{A}$.

Proof. It suffices to show that if $E \in \mathcal{M}$ and $F \in \mathcal{N}$ such that $E \times F = \coprod E_n \times F_n$ for some (non-disjoint!) sequences $(E_n) \subset \mathcal{M}$ and $(F_n) \subset \mathcal{N}$, then $\mu(E)\nu(E) = \sum \mu(E_n)\nu(E_n)$.

Trick. For all $x \in E$ and $y \in F$, there is a unique k such that $(x, y) \in E_k \times F_k$. Hence, for any fixed $y \in F$, $(x, y) \in E \times F$ for all $x \in E$, and thus

$$E = \coprod_{k \text{ s.t. } y \in F_k} E_k.$$

This is a disjoint union, since if $x \in E_j \cap E_k$ and $y \in F_j \cap F_k$, then $(x, y) \in (E_j \times F_j) \cap (E_k \times F_k)$, so $j = k$. Here is a cartoon of this trick:



$E = E_1 \amalg E_2 = E_3 \amalg E_4$

$E_1 = E_3$ and $E_2 = E_4$

$F = F_1 \amalg F_3 = F_2 \amalg F_4$

$F_1 = F_2$ and $F_3 = F_4$

Hence for $y \in F$,

$$\mu(E) = \sum_{k \text{ s.t. } y \in F_k} \mu(E_k) = \sum \mu(E_k)\chi_{F_k}(y),$$

and thus $\mu(E)\chi_F(y) = \sum \mu(E_k)\chi_{F_k}(y)$. Integrating over y yields

$$\mu(E)\nu(F) = \int_Y \mu(E)\chi_F(y)\, d\nu(y) = \int_Y \sum \mu(E_k)\chi_{F_k}(y)\, d\nu(y)$$
$$\underset{\text{(MCT)}}{=} \sum \int_Y \mu(E_k)\chi_{F_k}(y)\, d\nu(y) = \sum \mu(E_k)\nu(F_k). \qquad \square$$

Now use Carathéodory's outer measure construction, we get an outer measure $(\mu \times \nu)^*$ on $P(X \times Y)$, which restricts to a measure $\mu \times \nu$ on the $(\mu \times \nu)^*$-measurable sets, which is a σ-algebra containing $\mathcal{M} \times \mathcal{N}$ (as sets in $\mathcal{A}$ are $(\mu \times \nu)^*$-measurable, and $\mathcal{A}$ generates $\mathcal{M} \times \mathcal{N}$).

Exercise 3.8.12. Suppose X, Y are topological spaces and μ, ν are σ-finite Borel measures on X, Y respectively.

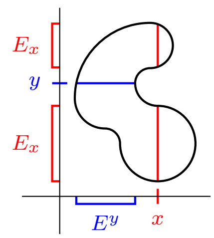
(1) Prove that $\mu \times \nu$ is σ-finite.
(2) Show that if μ, ν are both outer regular, then so is $\mu \times \nu$.
(3) Show that (2) fails when the σ-finite condition is dropped.
Hint: Consider a Dirac mass δ at x_0 such that $\delta(\{x_0\}) = \infty$.

3.9. **The Fubini and Tonelli Theorems.** For this section, fix measure spaces $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$.

Definition 3.9.1. For $E \subset X \times Y$, we define

- (x-section) $E_x := \{y \in Y | (x, y) \in E\} = \pi_Y(E \cap (\{x\} \times Y))$
- (y-section) $E^y := \{x \in X | (x, y) \in E\} = \pi_X(E \cap (X \times \{y\}))$

Here is a cartoon of x- and y-sections:



Exercise 3.9.2. Suppose $(E_n) \subset P(X \times Y)$. Prove the following assertions.

(1) $(\bigcup E_n)_x = \bigcup (E_n)_x$
(2) $(\bigcap E_n)_x = \bigcap (E_n)_x$
(3) $(E_n \setminus E_k)_x = (E_n)_x \setminus (E_k)_x$
(4) $\chi_{E_n}(x, y) = \chi_{(E_n)_x}(y)$.

Deduce similar statements also hold for y-sections.

Proposition 3.9.3. *Let $E \in \mathcal{M} \times \mathcal{N}$. For all $x \in X$, $E_x \in \mathcal{N}$ and for all $y \in Y$, $E^y \in \mathcal{M}$.*

Proof. We prove the first statement, and the second is similar.

> **Trick.** We'll show that the following set is a σ-algebra on $X \times Y$:
> $$\mathcal{S} := \{E \subset X \times Y | E_x \in \mathcal{N}\}.$$

This implies the result, since $\mathcal{S}$ contains the measurable rectangles in $\mathcal{M} \times \mathcal{N}$, which generates $\mathcal{M} \times \mathcal{N}$. Thus $\mathcal{M} \times \mathcal{N} \subset \mathcal{S}$.

(0) Observe $\emptyset \in \mathcal{N}$ implies $\emptyset \in \mathcal{S}$.
(1) If $(E_n) \subset \mathcal{S}$, then $(E_n)_x \in \mathcal{N}$ for all $n \in \mathbb{N}$. By Exercise 3.9.2, $(\bigcup E_n)_x = \bigcup (E_n)_x \in \mathcal{N}$. Thus $\bigcup E_n \in \mathcal{S}$.
(2) If $E \in \mathcal{S}$, then $E_x \in \mathcal{N}$. Observe $(E^c)_x = (E_x)^c \in \mathcal{N}$, and thus $E^c \in \mathcal{S}$. ☐

Exercise 3.9.4. Use Proposition 3.9.3 to show that $\mathcal{L} \times \mathcal{L}$ is not equal to $\mathcal{L}^2$, where $\mathcal{L}$ is the Lebesgue σ-algebra and $\mathcal{L}^2$ denotes the σ-algebra of $(\lambda \times \lambda)^*$-measurable sets in $\mathbb{R}^2$.

Definition 3.9.5. For $f : X \times Y \to \mathbb{R}$, $\overline{\mathbb{R}}$, or $\mathbb{C}$, we define

- (x-section) $f_x : Y \to \mathbb{R}$, $\overline{\mathbb{R}}$, or $\mathbb{C}$ by $f_x(y) := f(x, y)$, and
- (y-section) $f^y : X \to \mathbb{R}$, $\overline{\mathbb{R}}$, or $\mathbb{C}$ by $f^y(x) := f(x, y)$.

Corollary 3.9.6. *If $f : X \times Y \to \mathbb{R}$, $\overline{\mathbb{R}}$, or $\mathbb{C}$ is $\mathcal{M} \times \mathcal{N}$-measurable, then*

- *for all $x \in X$, f_x is $\mathcal{N}$-measurable, and*
- *for all $y \in Y$, f^y is $\mathcal{M}$-measurable.*

Proof. We'll prove the first statement, and the second is similar. Observe that for all $x \in X$ and measurable G contained in the codomain, $f_x^{-1}(G) = f^{-1}(G)_x \in \mathcal{N}$. ☐

Exercise 3.9.7. Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ is such that each x-section f_x is Borel measurable and each y-section f^y is continuous. Show f is Borel measurable.

Theorem 3.9.8 (Tonelli for characteristic functions)**.** *Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ-finite measure spaces. Then for all $E \in \mathcal{M} \times \mathcal{N}$,*

(1) The functions $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are measurable, and
(2) $(\mu \times \nu)(E) = \int \nu(E_x)\, d\mu(x) = \int \mu(E^y)\, d\nu(y)$.

Proof.
Step 1: First, we'll assume μ, ν are finite measures. Let $\Lambda \subset \mathcal{M} \times \mathcal{N}$ be the subset for which (1) and (2) above hold. Observe that $\Pi := \{\text{measurable rectangles in } \mathcal{M} \times \mathcal{N}\}$ is contained in Λ.
Step 1a: Π is a π-system.

Proof. The intersection of 2 measurable rectangles is a measurable rectangle. □

Step 1b: Λ is a λ-system. Thus by the $\pi - \lambda$ Theorem,

$$\mathcal{M} \times \mathcal{N} = \Lambda(\Pi) \subset \Lambda \subset \mathcal{M} \times \mathcal{N},$$

and thus equality holds.

Proof.
(0) First, note $X \times Y \in \Pi \subset \Lambda$.
(1) If $E \in \Lambda$ so that (1) and (2) hold for E, then as we assumed ν is finite,

$$x \longmapsto \nu((E^c)_x) = \nu((E_x)^c) = \nu(Y) - \nu(E_x)$$

is measurable (as a constant function minus a measurable function), as is $y \mapsto \mu((E^c)_y)$, so (1) holds for E^c. Moreover, $\mu \times \nu$ is finite, so

$$\begin{aligned}
(\mu \times \nu)(E^c) &= (\mu \times \nu)(X \times Y) - (\mu \times \nu)(E) \\
&= \int_X \nu(Y)\, d\mu(x) - \int \nu(E_x)\, d\mu(x) \\
&= \int_X (\nu(Y) - \nu(E_x))\, d\mu(x) \\
&= \int_X \nu((E_x)^c)\, d\mu(x) \\
&= \int_X \nu((E^c)_x)\, d\mu(x) && \text{proving part of (2) for } E^c \\
&= \int_Y \mu((E^c)^y)\, d\nu(y) && \text{similarly.}
\end{aligned}$$

Thus Λ is closed under taking complements.
(2) Suppose $(E_n) \subset \Lambda$ is a sequence of disjoint subsets. Observe for all $x \in X$, $((E_n)_x) \subset \mathcal{N}$ is disjoint. Then for all n, $x \mapsto \nu((E_n)_x)$ is measurable, and thus so is

$$x \longmapsto \sum \nu((E_n)_x) = \nu\left(\coprod (E_n)_x\right) = \nu\left(\left(\coprod E_n\right)_x\right).$$

Similarly, $y \mapsto \mu\left((\coprod E_n)^y\right)$ is measurable, proving (1) for $\coprod E_n$. We calculate

$$
\begin{aligned}
(\mu \times \nu)\left(\coprod E_n\right) &= \sum (\mu \times \nu)(E_n) \\
&= \sum \int_X \nu((E_n)_x)\, d\mu(x) \\
&= \int_X \sum \nu((E_n)_x)\, d\mu(x) && \text{(by the MCT 3.3.9)} \\
&= \int_X \nu\left(\coprod (E_n)_x\right) d\mu(x) \\
&= \int_X \nu\left(\left(\coprod E_n\right)_x\right) d\mu(x) && \text{proving part of (2) for } \coprod E_n \\
&= \int_Y \mu\left(\left(\coprod E_n\right)^y\right) d\nu(y) && \text{similarly.}
\end{aligned}
$$

Thus Λ is closed under taking countable disjoint unions. □

Step 2: When μ, ν are σ-finite, write $X \times Y$ as an *increasing* union $X \times Y = \bigcup X_n \times Y_n$ with $X_n \times Y_n$ measurable rectangles such that $\mu(X_n), \nu(Y_n) < \infty$ for all $n \in \mathbb{N}$. For $E \in \mathcal{M} \times \mathcal{N}$, write $E_n := E \cap (X_n \times Y_n)$, and observe $E_n \nearrow E$, so $(E_n)_x \nearrow E_x$. Thus the function

$$x \longmapsto \nu(E_x) = \lim \nu((E_n)_x)$$

is measurable (as a pointwise limit of measurable functions), as is $y \longmapsto \mu(E^y)$. We then calculate

$$
\begin{aligned}
(\mu \times \nu)(E) &= \lim (\mu \times \nu)(E_n) \\
&= \lim \int_X \nu((E_n)_x)\, d\mu(x) && \text{(by Step 1)} \\
&= \int_X \lim \nu((E_n)_x)\, d\mu(x) && \text{(by the MCT 3.3.9)} \\
&= \int_X \nu(E_x)\, d\mu(x) \\
&= \int_Y \mu(E^y)\, d\nu(y) && \text{similarly.}
\end{aligned}
$$

□

Theorem 3.9.9 (Tonelli). *Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ-finite measure spaces. For $f \in L^+(X \times Y, \mathcal{M} \times \mathcal{N})$,*

(1) $x \mapsto \int_Y f_x\, d\nu$ is $\mathcal{M}$-measurable (an element of $L^+(X, \mathcal{M})$),

(2) $y \mapsto \int_X f^y\, d\mu$ is $\mathcal{N}$-measurable (an element of $L^+(Y, \mathcal{N})$), and

(3) $\int_{X \times Y} f\, d(\mu \times \nu) = \int_X \left(\int_Y f_x\, d\nu\right) d\mu = \int_Y \left(\int_X f^y\, d\mu\right) d\nu.$

Proof. If $f = \chi_E$ for some $E \in \mathcal{M} \times \mathcal{N}$, this is exactly the previous theorem. Since $(cf + g)_x = c(f_x) + g_x$ (this is an exercise), we get the result for non-negative simple functions by linearity.

Suppose now $f \in L^+$ is arbitrary and $(\psi_n) \subset \mathsf{SF}^+$ such that $\psi_n \nearrow f$ everywhere. Then $(\psi_n)_x \nearrow f_x$ and $(\psi_n)^y \nearrow f^y$, so by the MCT 3.3.9,

$$\int_Y (\psi_n)_x \, d\nu \nearrow \int_Y f_x \, d\nu \qquad \text{and} \qquad \int_X (\psi_n)^y \, d\mu \nearrow \int_X f^y \, d\mu,$$

which implies (1) and (2) (countable supremums of measurable functions are measurable). Again by the MCT 3.3.9,

$$\begin{aligned}
\int_X \left(\int_Y f_x \, d\nu \right) d\mu &= \int_X \left(\lim \int_Y (\psi_n)_x \, d\nu \right) d\mu \\
&= \lim \int_X \left(\int_Y (\psi_n)_x \, d\nu \right) d\mu \\
&= \lim \int_{X \times Y} \psi_n \, d(\mu \times \nu) && \text{by previous theorem} \\
&= \int_{X \times Y} f \, d(\mu \times \nu) \\
&= \int_Y \left(\int_X f^y \, d\mu \right) d\nu && \text{similarly.}
\end{aligned}$$

□

Exercise 3.9.10 (Counterexample: Folland §2.5, #46)**.** Let $X = Y = [0, 1]$, $\mathcal{M} = \mathcal{N} = \mathcal{B}_{[0,1]}$, $\mu = \lambda$ Lebesgue measure, and ν counting measure. Let $\Delta = \{(x, x) | x \in [0, 1]\}$ be the diagonal. Prove that $\int \int \chi_\Delta \, d\mu \, d\nu$, $\int \int \chi_\Delta \, d\nu \, d\mu$, and $\int \chi_\Delta \, d(\mu \times \nu)$ are all distinct.

Exercise 3.9.11. Suppose $f : \mathbb{R} \to [0, \infty)$ is Borel measurable.

(1) Show that $E := \{(x, y) \in \mathbb{R}^2 | 0 \leq y \leq f(x)\}$ is Borel measurable.
(2) Show that $\int f(x) \, d\lambda(x) = (\lambda \times \lambda)(E)$.

Remark 3.9.12. Under the hypotheses of Tonelli's Theorem 3.9.9, if in addition $f \in L^+(X \times Y, \mathcal{M} \times \mathcal{N}) \cap L^1(\mu \times \nu)$, then

- $\displaystyle\int_Y f_x \, d\nu < \infty$ ($f_x \in L^1(\nu)$) a.e. $x \in X$, and
- $\displaystyle\int_X f^y \, d\mu < \infty$ ($f^y \in L^1(\mu)$) a.e. $y \in Y$.

Corollary 3.9.13 (Fubini)**.** *Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are σ-finite measure spaces. If $f \in L^1(\mu \times \nu)$, then*

(1) $f_x \in L^1(\nu)$ a.e. $x \in X$ and $f^y \in L^1(\mu)$ a.e. $y \in Y$,
(2) $\left(x \mapsto \displaystyle\int_Y f_x \, d\nu \right) \in L^1(\mu)$ and $\left(y \mapsto \displaystyle\int_X f^y \, d\mu \right) \in L^1(\nu)$, and
(3) $\displaystyle\int_{X \times Y} f \, d(\mu \times \nu) = \int_X \left(\int_Y f_x \, d\nu \right) d\mu = \int_Y \left(\int_X f^y \, d\mu \right) d\nu$.

Proof. Recall that

$$f = \mathrm{Re}(f)_+ - \mathrm{Re}(f)_- + i \, \mathrm{Im}(f)_+ - i \, \mathrm{Im}(f)_-,$$

where $\mathrm{Re}(f)_\pm, \mathrm{Im}(f)_\pm \in L^+(X \times Y, \mathcal{M} \times \mathcal{N}) \cap L^1(\mu \times \nu)$. Hence Tonelli's Theorem 3.9.9 applies to the 4 functions, as does Remark 3.9.12. The result follows. □

Exercise 3.9.14 (Counterexample: Folland §2.5, #48)**.** Let $X = Y = \mathbb{N}$, $\mathcal{M} = \mathcal{N} = P(\mathbb{N})$, and $\mu = \nu$ counting measure. Define

$$f(m,n) := \begin{cases} 1 & \text{if } m = n \\ -1 & \text{if } m = n+1 \\ 0 & \text{else.} \end{cases}$$

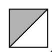
Prove that $\int |f| d(\mu \times \nu) = \infty$, and $\int\int f\, d\mu\, d\nu$ and $\int\int f\, d\nu\, d\mu$ both exist and are unequal.

Exercise 3.9.15. Let $f, g \in L^1([0,1], \lambda)$ where λ is Lebesgue measure. For $0 \leq x \leq 1$, define

$$F(x) := \int_{[x,1]} f\, d\lambda \qquad \text{and} \qquad G(x) := \int_{[x,1]} g\, d\lambda.$$

(1) Prove that F and G are continuous on $[0,1]$.
(2) Compute

$$\int_{\underbrace{[0,1]^2 = \square}_{\textit{Hint!}}} f(x)g(y)\, d(\lambda \times \lambda)$$



to prove the *integration by parts* formula:

$$\int_{[0,1]} Fg\, d\lambda = F(0)G(0) - \int_{[0,1]} Gf\, d\lambda.$$

Exercise 3.9.16. Prove the Fubini Theorem (Corollary 3.9.13) also holds replacing $(\mathcal{M} \times \mathcal{N}, \mu \times \nu)$ with its completion $(\overline{\mathcal{M} \times \mathcal{N}}, \overline{\mu \times \nu})$

Exercise 3.9.17. Show that the conclusions of the Fubini and Tonelli Theorems hold when $(X, \mathcal{M}, \mu)$ is an arbitrary measure space (not necessarily σ-finite) and Y is a countable set, $\mathcal{N} = P(Y)$, and ν is counting measure.

Exercise 3.9.18. Suppose $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are measure spaces *which are not assumed to be σ-finite*. Let $f \in L^1(\mu, \mathbb{R})$ and $g \in L^1(\nu, \mathbb{R})$, and define $h(x,y) := f(x)g(y)$.

(1) Prove that h is $\mathcal{M} \times \mathcal{N}$-measurable.
(2) Prove that $h \in L^1(\mu \times \nu)$.
(3) Prove that $\displaystyle\int_{X \times Y} h\, d(\mu \times \nu) = \int_X f\, d\mu \int_Y g\, d\nu$.

Remark: Since $(X, \mathcal{M}, \mu)$ and $(Y, \mathcal{N}, \nu)$ are not assumed to be σ-finite, you cannot directly apply the Fubini or Tonelli Theorems!

As an application, we give the following exercise on convolution multiplication on $\mathcal{L}^1(\mathbb{R}, \lambda)$.

Exercise 3.9.19. Suppose $f, g \in \mathcal{L}^1(\mathbb{R}, \lambda)$.

(1) Show that $y \mapsto f(x-y)g(y)$ is measurable for all $x \in \mathbb{R}$ and in $\mathcal{L}^1(\mathbb{R}, \lambda)$ for a.e. $x \in \mathbb{R}$.
(2) Define the *convolution* of f and g by

$$(f * g)(x) := \int_{\mathbb{R}} f(x-y)g(y)\, d\lambda(y).$$

Show that $f * g \in \mathcal{L}^1(\mathbb{R}, \lambda)$.
(3) Show that $\mathcal{L}^1(\mathbb{R}, \lambda)$ is a commutative $\mathbb{C}$-algebra under $\cdot, +, *$.

(4) Show that $\int_{\mathbb{R}} |f * g| \leq \int_{\mathbb{R}} |f| \int_{\mathbb{R}} |g|$, i.e., $\|\cdot\|_1$ is submultiplicative.

Since we already know that $\mathcal{L}^1(\mathbb{R}, \lambda)$ is complete, this shows that the $\mathbb{C}$-algebra $\mathcal{L}^1(\mathbb{R}, \lambda)$ is a *Banach algebra*.

3.10. **Then n-dimensional Lebesgue integral.** Recall that $\mathcal{L}$ is the Lebesgue σ-algebra on $\mathbb{R}$ and λ is Lebesgue measure.

Definition 3.10.1. We define $(\mathbb{R}^n, \mathcal{L}^n, \lambda^n)$ as the completion of $(\mathbb{R}^n, \underbrace{\mathcal{L} \times \cdots \times \mathcal{L}}_{n \text{ factors}}, \underbrace{\lambda \times \cdots \times \lambda}_{n \text{ factors}})$.

Facts 3.10.2. Here are some properties of Lebesgue measure. Verification is left as an exercise.

(1) λ^n is σ-finite.
(2) λ^n is regular.
(3) For all $E \in \mathcal{L}^n$ with $\lambda^n(E) < \infty$, for all $\varepsilon > 0$, there are disjoint rectangles $R_1, \ldots, R_n$ whose sides (projections) are intervals such that $\lambda^n(E \bigtriangleup \coprod^n R_k) < \varepsilon$, where $\bigtriangleup$ denotes symmetric difference.
(4) $\mathsf{ISF} = \mathsf{SF} \cap \mathcal{L}^1(\lambda^n)$ is dense in $\mathcal{L}^1(\lambda^n)$.
(5) $C_c(\mathbb{R}^n)$ is dense in $\mathcal{L}^1(\lambda^n)$.
(6) Suppose $E \in \mathcal{L}^n$.
   - For all $r \in \mathbb{R}^n$, $r + E \in \mathcal{L}^n$, and $\lambda^n(r + E) = \lambda^n(E)$.
   - For all $T \in GL(n, \mathbb{R})$, $TE \in \mathcal{L}^n$ and $\lambda^n(TE) = |\det(T)| \cdot \lambda^n(E)$.
(7) For all $\mathcal{L}^n$-measurable $f : \mathbb{R}^n \to \mathbb{C}$, the following functions are also $\mathcal{L}^n$-measurable:
$$x \longmapsto f(x + r) \text{ for } r \in \mathbb{R}^n, \text{ and}$$
$$x \longmapsto f(Tx) \text{ for } T \in GL(n, \mathbb{R}).$$
If moreover $f \in L^+$ or $\mathcal{L}^1(\lambda^n)$, then
$$\int f(x + r)\, d\lambda^n(x) = \int f(x)\, d\lambda^n(x) \qquad \text{and}$$
$$\int f(x)\, d\lambda^n(x) = |\det(T)| \cdot \int f(Tx)\, d\lambda^n(x).$$

Exercise 3.10.3. Suppose μ is a translation-invariant measure on $\mathcal{B}_{\mathbb{R}^n}$ such that $\mu([0, 1]^n) = 1$. Show that $\mu = \lambda^n|_{\mathcal{B}_{\mathbb{R}^n}}$.

Exercise 3.10.4. Prove some assertions from Facts 3.10.2.

Exercise 3.10.5. Suppose $T \in GL_n(\mathbb{C})$ and $f \in L^+$ or $\mathcal{L}^1(\lambda^n)$.

(1) Prove that $f \circ T \in L^+$ or $\mathcal{L}^1(\lambda^n)$ respectively.
(2) Show that
$$\int f(x)\, d\lambda^n(x) = |\det(T)| \cdot \int f(Tx)\, d\lambda^n(x).$$

4.1. **Signed measures.** For this section, let $(X, \mathcal{M})$ be a measurable space.

Definition 4.1.1. A function $\nu : \mathcal{M} \to \overline{\mathbb{R}}$ is called a *signed measure* if

- ν takes on *at most one* of the values $\pm\infty$,
- (vacuum) $\nu(\emptyset) = 0$, and
- (σ-additivity) for every disjoint sequence $(E_n) \subset \mathcal{M}$, $\nu\left(\coprod E_n\right) = \sum \nu(E_n)$.

We call ν *finite* if ν does not take on the values $\pm\infty$.

Remark 4.1.2. If ν is a signed measure and $(E_n) \subset \mathcal{M}$ are disjoint, then σ-additivity of ν *implies* that the sum $\sum \nu(E_n)$ must converge *absolutely* if $|\nu\left(\coprod E_n\right)| < \infty$. Indeed, reindexing the sets (E_n) does not change $\coprod E_n$, and thus it must not change the sum $\sum \nu(E_n)$.

Exercise 4.1.3.

(1) If μ_1, μ_2 are measures on $(X, \mathcal{M})$ with at least one of μ_1, μ_2 finite, then $\nu := \mu_1 - \mu_2$ is a signed measure.
(2) Suppose μ is a measure on $(X, \mathcal{M})$. If $f : X \to \overline{\mathbb{R}}$ is measurable and *extended μ-integrable*, i.e., at least one of $\int f_\pm < \infty$, then $\nu(E) := \int_E f\, d\mu$ is a signed measure.

It is now our goal to prove these are really the *only* ways to construct signed measures!

Definition 4.1.4. Suppose ν is a signed measure on $(X, \mathcal{M})$. We call $E \in \mathcal{M}$:

- *positive* if for all measurable $F \subseteq E$, $\mu(F) \geq 0$,
- *negative* if for all measurable $F \subseteq E$, $\mu(F) \leq 0$, and
- *null* if for all measurable $F \subseteq E$, $\mu(F) = 0$.

Observe that $N \in \mathcal{M}$ is null if and only if N is both positive and negative.

Facts 4.1.5. For ν a signed measure on $(X, \mathcal{M})$, we have the following facts about positive measurable sets. Similar statements hold for negative and null measurable sets.

(1) E positive implies $\nu(E) \geq 0$.
(2) E positive and $F \subseteq E$ measurable implies F is positive.
(3) $(E_n) \subset \mathcal{M}$ positive implies $\bigcup E_n$ positive.

Proof. Disjointify the E_n so that $\bigcup E_n = \coprod F_n$ where $F_1 := E_1$ and $F_n := E_n \setminus \bigcup^{n-1} E_k$ is positive for all $n \in \mathbb{N}$. If $G \subset \bigcup E_n = \coprod F_n$, then

$$\nu(G) = \nu\left(G \cap \coprod F_n\right) = \sum \nu(G \cap F_n) \geq 0. \qquad \square$$

(4) If $0 < \nu(E) < \infty$, there is a positive $F \subseteq E$ such that $\nu(F) > 0$.

Proof. If E is positive, we win. Otherwise, let $n_1 \in \mathbb{N}$ be minimal such that that there is a measurable $E_1 \subset E$ and $\nu(E_1) < -\frac{1}{n_1}$. Observe that $\nu(E \setminus E_1) > 0$, so if $E \setminus E_1$ is positive, we win. Otherwise, let $n_2 \in \mathbb{N}$ minimal such that there is a measurable $E_2 \subset E \setminus E_1$ with $\nu(E_2) < -\frac{1}{n_2}$. We can inductively iterate this procedure. Either $E \setminus \coprod^n E_k$ is positive for some n, or we have constructed a disjoint sequence (E_k) with $\nu(E_k) < -\frac{1}{n_k}$ for all k. Set $F := E \setminus \coprod E_k$. Since $\nu(E) < \infty$ and $E = F \amalg \coprod E_k$, by countable additivity, $\sum |\nu(E_k)| < \infty$, so $\sum_k -\frac{1}{n_k}$ converges. Hence $n_k \to \infty$ as $k \to \infty$. Since $\nu(E) > 0$ and $\nu(E_k) < 0$ for all k, $\nu(F) > 0$. Suppose $G \subset F$ is measurable. Then $\nu(G) \geq -\frac{1}{n_k - 1}$ for all k with $n_k > 1$, and thus $\nu(G) \geq 0$. So F is positive. □

Theorem 4.1.6 (Hahn Decomposition). *Let ν be a signed measure on $(X, \mathcal{M})$. There is a positive set $P \in \mathcal{M}$ such that P^c is negative. Moreover, if $Q \in \mathcal{M}$ is another positive set such that Q^c is negative, then $P \triangle Q$ and $P^c \triangle Q^c$ are null.*

A positive $P \in \mathcal{M}$ such that P^c is negative is called a Hahn decomposition *of X with respect to ν.*

Proof.

Existence: We may assume $\infty \notin \operatorname{im}(\nu) \subset \overline{\mathbb{R}}$ (otherwise, replace ν with $-\nu$). Define

$$r := \sup \{\nu(E) | E \text{ is positive}\}.$$

Then there is a sequence (E_n) of positive sets such that $\nu(E_n) \to r$. Take $P := \bigcup E_n$, which is positive. Since a signed measure restricted to a positive set is a positive measure, $\nu(P) = \lim \nu(E_n) = r$ by continuity from below ($\mu 3$). We claim that P^c is negative. If $F \subset P^c$ such that $\nu(F) > 0$, by Facts 4.1.5(4), there is a positive $G \subset F$ such that $\nu(G) > 0$. Then $P \amalg G$ is positive with $\nu(P \amalg G) = \nu(P) + \nu(G) > r$, a contradiction.

Uniqueness: Suppose $P, Q \subset X$ are positive such that P^c, Q^c are negative. Then

$$P \triangle Q = (P \setminus Q) \cup (Q \setminus P) = \underbrace{(P \cap Q^c)}_{\text{pos. and neg.}} \cup \underbrace{(Q \cap P^c)}_{\text{pos. and neg.}}$$

is ν-null. Similarly, $P^c \triangle Q^c$ is ν-null. □

Definition 4.1.7. We say positive measures μ_1, μ_2 on $(X, \mathcal{M})$ are *mutually singular*, denoted $\mu_1 \perp \mu_2$, if there exist disjoint $E, F \in \mathcal{M}$ such that $X = E \amalg F$ and $\mu_1(F) = 0 = \mu_2(E)$.

Theorem 4.1.8 (Jordan decomposition). *Let ν be a signed measure on $(X, \mathcal{M})$. There exist unique mutually singular measures $\nu_\pm$ on $(X, \mathcal{M})$ such that $\nu = \nu_+ - \nu_-$, which we call the* Jordan decomposition *of ν.*

Proof.

Existence: Given a Hahn decomposition $X = P \amalg P^c$, $\nu_+(E) := \nu(E \cap P)$ and $\nu_-(E) := -\nu(E \cap P^c)$ are positive measures on $\mathcal{M}$, such that $\nu_+(P^c) = 0 = \nu_-(P)$ and $\nu = \nu_+ - \nu_-$. (Observe $\nu_\pm$ are *independent* of the Hahn decomposition.)

Uniqueness: Suppose that $\nu = \mu_+ - \mu_- = \nu_+ - \nu_-$ where $\mu_\pm$ and $\nu_\pm$ are all positive measures with $\mu_+ \perp \mu_-$ and $\nu_+ \perp \nu_-$. Then by definition of mutual singularity, there exist two Hahn decompositions for ν: $X = P \amalg P^c$ such that $\mu_+(P^c) = 0 = \mu_-(P)$ and $X = Q \amalg Q^c$ such that $\nu_+(Q^c) = 0 = \nu_-(Q)$. Thus $P \triangle Q$ and $P^c \triangle Q^c$ are ν-null, and for all $E \in \mathcal{M}$,

$$\begin{aligned}\mu_+(E) = \mu_+(E \cap P) = \nu(E \cap P) &= \nu(E \cap P \cap Q) + \nu(E \cap \underbrace{P \cap Q^c}_{\nu\text{-null}})\\ &= \nu(E \cap P \cap Q) = \nu(E \cap P \cap Q) + \nu(E \cap \underbrace{P^c \cap Q}_{\nu\text{-null}}) = \nu(E \cap Q)\\ &= \nu_+(E \cap Q) = \nu_+(E).\end{aligned}$$

Hence $\mu_+ = \nu_+$, and similarly, $\mu_- = \nu_-$. □

Definition 4.1.9. For a signed measure ν on $(X, \mathcal{M})$, define $L^1(\nu) := L^1(\nu_+) \cap L^1(\nu_-)$. For $f \in L^1(\nu)$, define

$$\int f\, d\nu := \int f\, d\nu_+ - \int f\, d\nu_-.$$

Clearly $L^1(\nu)$ is a $\mathbb{C}$-vector space and $\int$ is a linear functional. We define $\mathcal{L}^1(\nu)$ to be the quotient of $L^1(\nu)$ by the equivalence relation $f = g$ ν_+-a.e. and ν_--a.e.

Exercise 4.1.10. Suppose ν is a signed measure on $(X, \mathcal{M})$. Prove that $E \in \mathcal{M}$ is ν-null if and only if E is ν_+-null and ν_--null. Deduce that $f = g$ ν_+-a.e. and ν_--a.e if and only if $f = g$ up to a ν-null set.

Definition 4.1.11. For a signed measure ν on $(X, \mathcal{M})$, define the *total variation* of $\nu = \nu_+ - \nu_-$ by $|\nu| := \nu_+ + \nu_-$, which is a positive measure. Observe that

$$|\nu(E)| = |\nu_+(E) - \nu_-(E)| \leq \nu_+(E) + \nu_-(E) = |\nu|(E) \qquad \forall\, E \in \mathcal{M}.$$

Hence ν is finite if and only if $|\nu|$ is finite.

Exercise 4.1.12. Suppose ν is a signed measure on $(X, \mathcal{M})$, let $\nu = \nu_+ - \nu_-$ be its Jordan decomposition, and let $|\nu|$ be its total variation.

(1) Prove that for $E \in \mathcal{M}$, $\nu_+(E) = \sup\{\nu(F) | F \in \mathcal{M} \text{ with } F \subset E\}$.
(2) Prove that for $E \in \mathcal{M}$, $\nu_-(E) = -\inf\{\nu(F) | F \in \mathcal{M} \text{ with } F \subset E\}$.
(3) Prove that for $E \in \mathcal{M}$,

$$|\nu|(E) = \sup\left\{\sum_{i=1}^n |\nu(E_i)| \,\middle|\, E_1, \ldots, E_n \in \mathcal{M} \text{ disjoint with } E = \coprod_{i=1}^n E_i\right\}.$$

Exercise 4.1.13. Suppose $(X, \mathcal{M})$ is a measurable space, ν is a signed measure on $(X, \mathcal{M})$, and λ, μ are positive measures on $(X, \mathcal{M})$ such that $\nu = \lambda - \mu$. Show that $\nu_+ \leq \lambda$ and $\nu_- \leq \mu$ where $\nu = \nu_+ - \nu_-$ is the Jordan decomposition of ν.

Lemma 4.1.14. *Suppose μ_1, μ_2 are measures on X with at least one of μ_1, μ_2 finite, and set $\nu = \mu_1 - \mu_2$. Then $|\nu|(X) \leq \mu_1(X) + \mu_2(X)$.*

Proof. Let $\nu = \nu_+ - \nu_-$ be the Jordan decomposition of ν, and let $X = P \amalg P^c$ be a Hahn decomposition such that $\nu_+(P^c) = 0 = \nu_-(P)$. Then

$$\begin{aligned}0 &\leq \nu_+(X) = \nu(X \cap P) = \nu(P) = \mu_1(P) - \mu_2(P) \leq \mu_1(P) \leq \mu_1(X)\\ 0 &\leq \nu_-(X) = -\nu(X \cap P^c) = -\nu(P^c) = \mu_2(P^c) - \mu_1(P^c) \leq \mu_2(P^c) \leq \mu_2(X)\end{aligned}$$

Hence $|\nu|(X) = \nu_+(X) + \nu_-(X) \leq \mu_1(X) + \mu_2(X)$. □

Exercise 4.1.15 (Folland §3.1, #3)**.** Suppose μ is a positive measure on $(X, \mathcal{M})$ and ν is a signed measure on $(X, \mathcal{M})$. Prove that the following are equivalent.

(1) $\nu \perp \mu$
(2) $|\nu| \perp \mu$
(3) $\nu_+ \perp \mu$ and $\nu_- \perp \mu$.

Exercise 4.1.16 (Folland §3.1, #3)**.** Let ν be a signed measure on $(X, \mathcal{M})$. Prove the following assertions:

(1) $\mathcal{L}^1(\nu) = \mathcal{L}^1(|\nu|)$.
(2) If $f \in \mathcal{L}^1(\nu)$, $\left|\int f\, d\nu\right| \leq \int |f| d|\nu|$.
(3) If $E \in \mathcal{M}$, $|\nu|(E) = \sup\left\{\left|\int_E f\, d\nu\right| \middle| -1 \leq f \leq 1\right\}$.

Exercise 4.1.17. Suppose μ, ν are finite signed measures on the measurable space $(X, \mathcal{M})$.

(1) Prove that the signed measure $\mu \wedge \nu := \frac{1}{2}(\mu + \nu - |\mu - \nu|)$ satisfies $(\mu \wedge \nu)(E) \leq \min\{\mu(E), \nu(E)\}$ for all $E \in \mathcal{M}$.
(2) Suppose in addition that μ, ν are positive. Prove that $\mu \perp \nu$ if and only if $\mu \wedge \nu = 0$.

Exercise 4.1.18 (Folland §3.1, #6)**.** Suppose

$$\nu(E) := \int_E f\, d\mu \qquad\qquad E \in \mathcal{M}$$

where μ is a positive measure on $(X, \mathcal{M})$ and and f is an extended μ-integrable function. Describe the Hahn decompositions of ν and the positive, negative, and total variations of ν in terms of f and μ.

Exercise 4.1.19. In this exercise, we will show that

$$M := M(X, \mathcal{M}, \mathbb{R}) := \{\text{finite signed measures on } (X, \mathcal{M})\}$$

is a Banach space with $\|\nu\| := |\nu|(X)$.

(1) Prove $\|\nu\| := |\nu|(X)$ is a norm on M.
(2) Show that $(\nu_n) \subset M$ Cauchy implies $(\nu_n(E)) \subset \mathbb{R}$ is uniformly Cauchy for all $E \in \mathcal{M}$. That is, show that for all $\varepsilon > 0$, there is an $N \in \mathbb{N}$ such that for all $n \geq N$ and $E \in \mathcal{M}$, $|\nu_m(E) - \nu_n(E)| < \varepsilon$.
(3) Use part (2) to define a candidate limit signed measure μ on $\mathcal{M}$. Prove that ν is σ-additive.
Hint: first prove ν is finitely additive.
(4) Prove that $\sum \nu(E_n)$ converges absolutely when $(E_n) \subset \mathcal{M}$ is disjoint, and thus ν is a finite signed measure.
(5) Show that $\nu_n \to \nu$ in M.

4.2. **Absolute continuity and the Lebesgue-Radon-Nikodym Theorem.** For this section, we fix a measurable space $(X, \mathcal{M})$.

Definition 4.2.1. Let ν be a signed measure and μ a positive measure on $(X, \mathcal{M})$. We say ν is *absolutely continuous with respect to* μ, denoted $\nu \ll \mu$, if $\mu(E) = 0$ implies $\nu(E) = 0$.

Example 4.2.2. Let $f \in L^1(\mu, \mathbb{R})$ and set $\nu(E) := \int_E f\, d\mu$. (This is sometimes denoted by $d\nu := f d\mu$.) Then $\nu \ll \mu$.

Exercise 4.2.3 (Folland §3.2, #8)**.** Suppose μ is a positive measure on $(X, \mathcal{M})$ and ν is a signed measure on $(X, \mathcal{M})$. Prove that the following are equivalent.

(1) $\nu \ll \mu$
(2) $|\nu| \ll \mu$
(3) $\nu_+ \ll \mu$ and $\nu_- \ll \mu$.

Exercise 4.2.4. Suppose $(X, \mathcal{M})$ is a measurable space and ν is a signed measure and λ, μ are positive measures on $(X, \mathcal{M})$ such that $\nu = \lambda - \mu$. Show that $\nu_+ \leq \lambda$ and $\nu_- \leq \mu$ where $\nu = \nu_+ - \nu_-$ is the Jordan decomposition of ν.

Exercise 4.2.5 (Adapted from Folland §3.2, #9)**.** Suppose $\{\nu_j\}$ is a sequence of positive measures on $(X, \mathcal{M})$ and μ is a positive measure on $(X, \mathcal{M})$. Prove the following assertions.

(1) If $\{\nu_j\}$ is a sequence of positive measures on $(X, \mathcal{M})$ with $\nu_j \perp \mu$ for all j, then $\sum_{j=1}^\infty \nu_j \perp \mu$.
(2) If ν_1, ν_2 are positive measures on $(X, \mathcal{M})$ with at least one of ν_1, ν_2 is finite and $\nu_j \perp \mu$ for $j = 1, 2$, then $(\nu_1 - \nu_2) \perp \mu$.
(3) If $\{\nu_j\}$ is a sequence of positive measures on $(X, \mathcal{M})$ with $\nu_j \ll \mu$ for all j, then $\sum_{j=1}^\infty \nu_j \ll \mu$.
(4) If ν_1, ν_2 are positive measures on $(X, \mathcal{M})$ with at least one of ν_1, ν_2 is finite and $\nu_j \ll \mu$ for $j = 1, 2$, then $(\nu_1 - \nu_2) \ll \mu$.

Proposition 4.2.6. *Suppose ν is a finite signed measure and μ is a positive measure on $(X, \mathcal{M})$. The following are equivalent:*

(1) $\nu \ll \mu$, and
(2) For all $\varepsilon > 0$, there is a $\delta > 0$ such that for all $E \in \mathcal{M}$, $|\nu(E)| < \varepsilon$ whenever $\mu(E) < \delta$.

Proof. Since $\nu \ll \mu$ if and only if $|\nu| \ll \mu$ and $|\nu(E)| \leq |\nu|(E)$, we may assume ν is positive. The result now follows from a previous exercise. For completeness, we'll provide the proof below.

First, it is clear that (2) implies (1). Suppose (2) fails. Then there is an $\varepsilon > 0$ such that for all $n \in \mathbb{N}$, there is an $E_n \in \mathcal{M}$ with $\mu(E_n) < 2^{-n}$, but $\nu(E_n) \geq \varepsilon$. Set $F := \bigcap_{k=1}^\infty \bigcup_{k=n}^\infty E_n$. Since

$$\mu\left(\bigcup_{k=n}^\infty E_n\right) < \sum_{n=k}^\infty 2^{-k} = 2^{1-k} \qquad \forall\, k \in \mathbb{N},$$

$\mu(F) = 0$. But since ν is finite, $\nu(F) = \lim_k \left(\bigcup_{k=n}^\infty E_n\right) \geq \varepsilon$. Hence (1) fails. □

Example 4.2.7. On $(\mathbb{N}, P(\mathbb{N}))$, define $\mu(E) := \sum_{n \in E} 2^{-n}$ and $\nu(E) := \sum_{n \in E} 2^n$. Then $\nu \ll \mu$ and $\mu \ll \nu$, but (2) above fails as ν is not finite.

Lemma 4.2.8. *Suppose μ, ν are finite measures on $(X, \mathcal{M})$. Either $\nu \perp \mu$ or there is an $\varepsilon > 0$ and $E \in \mathcal{M}$ such that $\mu(E) > 0$ and $\nu \geq \varepsilon\mu$ on E, i.e., E is positive for $\nu - \varepsilon\mu$.*

Proof. Let $X = P_n \amalg P_n^c$ be a Hahn decomposition for $\nu - n^{-1}\mu$ for all $n \in \mathbb{N}$. Set $P := \bigcup P_n$ so $P^c = \bigcap P_n^c$. Then P^c is negative for all $\nu - n^{-1}\mu$. Observe

$$0 \leq \nu(P^c) \leq \frac{1}{n} \underbrace{\mu(P^c)}_{<\infty} \qquad \forall\, n \in \mathbb{N},$$

so $\nu(P^c) = 0$. If $\mu(P) = 0$, then $\nu \perp \mu$. If $\mu(P) > 0$, then $\mu(P_n) > 0$ for some n, and P_n is positive for $\nu - n^{-1}\mu$. □

Theorem 4.2.9 (Lebesgue-Radon-Nikodym)**.** *Let ν be a σ-finite signed measure and μ a σ-finite positive measure on $(X, \mathcal{M})$. There are unique σ-finite signed measures λ, ρ on $(X, \mathcal{M})$ called the* Lebesgue decomposition of ν *such that*

$$\lambda \perp \mu, \qquad \rho \ll \mu, \qquad \text{and} \qquad \nu = \lambda + \rho.$$

Moreover, there exists a unique extended μ-integrable function f called the Radon-Nikodym derivative of ρ with respect to μ *such that $d\rho = f d\mu$. If ν is positive or finite, then so are λ and ρ respectively, and $f \in L^+$ or $L^1(\mu)$ respectively.*

Proof.
Case 1: Suppose μ, ν are finite positive measures.
Uniqueness: Suppose λ, λ' are finite signed measures such that $\lambda, \lambda' \perp \mu$ and $f, f' \in \mathcal{L}^1$ such that $d\nu = d\lambda + f d\mu = d\lambda' + f' d\mu$. Then as signed measures, $d(\lambda - \lambda') = (f' - f)d\mu$. But $(\lambda - \lambda') \perp \mu$ and $(f' - f)d\mu \ll d\mu$, so as signed measures by Exercise 4.2.5, $d(\lambda - \lambda') = 0 = (f' - f)d\mu$. We conclude that $\lambda = \lambda'$ and $f = f'$ in $\mathcal{L}^1$.
Existence: Set

$$A := \left\{ f \in L^1(X, \mu, [0, \infty]) \,\middle|\, \int_E f \, d\mu \leq \nu(E) \text{ for all } E \in \mathcal{M} \right\}.$$

Observe that $0 \in A$.

Claim. *$f, g \in A$ implies $f \vee g \in A$.*

Proof. For all $E \in \mathcal{M}$,

$$\int_E f \vee g \, d\mu = \int_{E \cap \{g < f\}} f \, d\mu + \int_{E \setminus \{g < f\}} g \, d\mu \leq \nu(E \cap \{g < f\}) + \nu(E \setminus \{g < f\}) = \nu(E).$$

□

Set $M := \sup \left\{ \int f \, d\mu \middle| f \in A \right\}$, and note that $M \leq \nu(X) < \infty$. Choose $(f_n) \subset A$ such that $\int f_n \, d\mu \nearrow M$. Set $g_n := \max\{f_1, \ldots, f_n\} \in A$ and $f := \sup g_n$. Then by the Squeeze Theorem,

$$\int f_n \, d\mu \leq \int g_n \, d\mu \nearrow M.$$

Since $g_n \nearrow f$ pointwise,

$$\int_E f \, d\mu \underset{\text{(MCT)}}{=} \lim_n \int_E g_n \, d\mu \leq \nu(E) \qquad \forall E \in \mathcal{M}.$$

So $f \in A$ and $\int f \, d\mu = M$.

Claim. $\lambda(E) := \nu(E) - \int_E f\,d\mu \geq 0$ *is mutually singular with respect to* μ*. So setting* $d\rho := f d\mu$*, we have* $\lambda \perp \mu$*,* $\rho \ll \mu$*,* $\nu = \lambda + \rho$*, and* $d\rho = f d\mu$*.*

Proof. Suppose λ is not mutually singular with respect to μ. Then by Lemma 4.2.8, there is a $E \in \mathcal{M}$ and $\varepsilon > 0$ such that $\mu(E) > 0$ and $\lambda \geq \varepsilon\mu$ on E. But then for all $F \in \mathcal{M}$,

$$\begin{aligned}
\int_F f + \varepsilon\chi_E\,d\mu &= \int_F f\,d\mu + \varepsilon\mu(E \cap F) \\
&\leq \int_F f\,d\mu + \lambda(E \cap F) \\
&= \int_F f\,d\mu + \nu(E \cap F) - \int_{E\cap F} f\,d\mu \\
&= \int_{F\setminus E} f\,d\mu + \nu(E \cap F) \\
&\leq \nu(F \setminus E) + \nu(E \cap F) \\
&= \nu(F).
\end{aligned}$$

Hence $f + \varepsilon\chi_E \in A$, but $\int f + \varepsilon\chi_E\,d\mu = M + \varepsilon\mu(E) > M$, a contradiction. □

Case 2: Suppose μ, ν are σ-finite positive measures.
Existence: Write $X = \coprod X_n$ such that $\mu(X_n) < \infty$ and $\nu(X_n) < \infty$ for all n. Set $\mu_n(E) := \mu(E \cap X_n)$ and $\nu_n(E) := \nu(E \cap X_n)$ for all n. By Case 1, there exist positive measures $\lambda_n \perp \mu_n$ and $f_n \in \mathcal{L}^1_+(X_n, \mu_n)$ such that $d\nu_n = d\lambda_n + f_n d\mu_n$. Since $\mu_n(X_n^c) = \nu_n(X_n^c) = 0$, we have

$$\lambda_n(X_n^c) = \nu_n(X_n^c) - \int_{X_n^c} f_n\,d\mu_n = 0.$$

Hence we may assume $f_n|_{X_n^c} = 0$. Set $\lambda := \sum \lambda_n$ and $f := \sum f_n \in L^+$. Then $\lambda \perp \mu$ by Exercise 4.2.5, λ and $f d\mu$ are σ-finite, and $d\nu = d\lambda + f d\mu$.
Uniqueness: If λ' is another positive measure such that $\lambda' \perp \mu$ and $f' \in L^+$ such that $d\nu = d\lambda' + f' d\mu$. Setting $\lambda'_n(E) := \lambda'(E \cap X_n)$ for $E \in \mathcal{M}$ and $f'_n := f'\chi_{X_n}$, by Uniqueness from Case 1, we have $\lambda'_n = \lambda_n$ and $f'_n = f_n$ in $\mathcal{L}^1(\mu_n)$. Then

$$\begin{aligned}
\lambda' &= \sum \lambda'_n = \sum \lambda_n = \lambda && \text{on } X\text{, and} \\
f' &= \sum f'_n = \sum f_n = f && \text{in } \mathcal{L}^1(\mu).
\end{aligned}$$

Case 3: Suppose μ is σ-finite positive and ν is σ-finite signed. In this case, we use the Jordan Decomposition Theorem 4.1.8 to get $\nu = \nu_+ - \nu_-$ with $\nu_+ \perp \nu_-$. We apply Case 2 to $\nu_\pm$ separately and subtract the results. This shows existence and uniqueness. □

Remark 4.2.10. If μ is σ-finite positive and ν is σ-finite signed with $\nu \ll \mu$, there is a unique extended μ-integrable function $\frac{d\nu}{d\mu}$ called the *Radon-Nikodym derivative of* ν *with respect to* μ such that $d\nu = \frac{d\nu}{d\mu} d\mu$.

Exercise 4.2.11. Suppose ν is a σ-finite signed measure.

(1) Show that $\left|\frac{d\nu}{d|\nu|}\right| = 1$, $|\nu|$-a.e.
(2) Suppose further that $\nu \ll \mu$ for some σ-finite positive measure μ on $(X, \mathcal{M})$. Show that for all $f \in \mathcal{L}^1(\nu)$, $f\frac{d\nu}{d\mu} \in \mathcal{L}^1(\mu)$ and $\int f\,d\nu = \int f\frac{d\nu}{d\mu}\,d\mu$.
(3) Suppose even further that $\mu \ll \lambda$ for some σ-finite positive measure λ. Show $\nu \ll \lambda$ and $\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu}\frac{d\mu}{d\lambda}$.

Definition 4.2.12. A signed measure ν on a topological space $(X, \mathcal{T})$ is called *regular* if $|\nu|$ is regular.

Exercise 4.2.13. Suppose ν is a finite signed Borel measure on the LCH space X. Determine which of the conditions below are equivalent.
(1) ν is regular.
(2) $\nu_\pm$ is regular.
(3) For every $E \in \mathcal{B}_X$ and $\varepsilon > 0$, there is an open $U \subset X$ with $E \subset U$ such that $|\nu(U) - \nu(E)| < \varepsilon$.

Which of the above conditions are equivalent if
- X is σ-compact?
- ν is not finite?

4.3. **Complex measures.** For this section, fix a measurable space $(X, \mathcal{M})$.

Definition 4.3.1. A function $\nu : \mathcal{M} \to \mathbb{C}$ is called a *complex measure* if
- (vacuum) $\nu(\emptyset) = 0$, and
- (σ-additivity) For every disjoint sequence $(E_n) \subset \mathcal{M}$, $\nu\left(\coprod E_n\right) = \sum \nu(E_n)$.

Observe that if ν is a complex measure on $(X, \mathcal{M})$, then $\operatorname{Re}(\nu)$ and $\operatorname{Im}(\nu)$ are finite signed measures on $(X, \mathcal{M})$.

Remark 4.3.2. As in Remark 4.1.2, if ν is a complex measure and $(E_n) \subset \mathcal{M}$ are disjoint, then σ-additivity of ν *implies* that the sum $\sum \nu(E_n)$ converges *absolutely*.

Exercise 4.3.3. Prove the following assertions.
(1) If $\mu_0, \mu_1, \mu_2, \mu_3$ are finite measures on $(X, \mathcal{M})$, then $\sum_{k=0}^3 i^k \mu_k$ is a complex measure.
(2) For μ a measure on $(X, \mathcal{M})$ and $f \in L^1(\mu)$, $\nu(E) := \int_E f\,d\mu$ is a complex measure on $(X, \mathcal{M})$.

By the Jordan Decomposition Theorem 4.1.8, we get the following corollary:

Corollary 4.3.4. *If ν is a complex measure on $(X, \mathcal{M})$, there exist unique pairs of mutually singular finite measures $\operatorname{Re}(\nu)_\pm$ and $\operatorname{Im}(\nu)_\pm$ such that*

$$\nu = \underbrace{\operatorname{Re}(\nu)_+}_{=:\nu_0} - \underbrace{\operatorname{Re}(\nu)_-}_{=:\nu_2} + i\big(\underbrace{\operatorname{Im}(\nu)_+}_{=:\nu_1} - \underbrace{\operatorname{Im}(\nu)_-}_{=:\nu_3}\big) =: \sum_{k=0}^{3} i^k \nu_k.$$

Definition 4.3.5. For a complex measure ν on $(X, \mathcal{M})$, we define $L^1(\nu) := \bigcap_{k=0}^3 L^1(\nu_k)$. We define $\mathcal{L}^1(\nu)$ to be the quotient under the equivalence relation $f = g$ ν_k-a.e. for $k = 0, 1, 2, 3$. For $f \in L^1(\nu_k)$, we define

$$\int f\,d\nu := \sum_{k=0}^{3} i^k \int f\,d\nu_k.$$

Warning 4.3.6. The total variation of a complex measure $\nu = \sum_{k=0}^{3} i^k \nu_k$ is *not* $\sum_{k=0}^{3} \nu_k$. We must use the complex Radon-Nikodym Theorem 4.3.9 below.

Definition 4.3.7. Suppose ν is a complex measure and μ is a positive measure on $(X, \mathcal{M})$. We say:

- $\nu \perp \mu$ if $\operatorname{Re}(\nu) \perp \mu$ and $\operatorname{Im}(\nu) \perp \mu$, and
- $\nu \ll \mu$ if $\operatorname{Re}(\nu) \ll \mu$ and $\operatorname{Im}(\nu) \ll \mu$.

Exercise 4.3.8. Suppose ν is a complex measure and μ is a positive measure on $(X, \mathcal{M})$. Show that $\nu \ll \mu$ if and only if for all $E \in \mathcal{M}$, $\mu(E) = 0$ implies $|\nu(E)| = 0$.

Theorem 4.3.9 (Complex Lebesgue-Radon-Nikodym)**.** *If ν is a complex measure on $(X, \mathcal{M})$ and μ is a σ-finite positive measure on $(X, \mathcal{M})$, there exists a unique complex measure λ on $(X, \mathcal{M})$ and a unique $f \in \mathcal{L}^1(\mu)$ such that $\lambda \perp \mu$ and $d\nu = d\lambda + f d\mu$.*

Proof. Apply the Lebesgue-Radon-Nikodym Theorem 4.2.9 to $\operatorname{Re}(\nu)$ and $\operatorname{Im}(\nu)$ separately and then recombine. □

Lemma 4.3.10. *Suppose ν is a complex measure on $(X, \mathcal{M})$. There is a unique positive measure $|\nu|$ on $(X, \mathcal{M})$ satisfying the following property:*

- *For all σ-finite positive measures μ on $(X, \mathcal{M})$ and $f \in \mathcal{L}^1(\mu)$ such that $d\nu = f d\mu$, $d|\nu| = |f| d\mu$.*

We call $|\nu|$ the total variation *of ν.*

Proof. First consider $\mu := |\operatorname{Re}(\nu)| + |\operatorname{Im}(\nu)|$. Since $|\operatorname{Re}(\nu)| \ll \mu$ and $|\operatorname{Im}(\nu)| \ll \mu$, we have $\operatorname{Re}(\nu) \ll \mu$ and $\operatorname{Im}(\nu) \ll \mu$, and thus $\nu \ll \mu$. By the complex Lebesgue-Radon-Nikodym Theorem 4.3.9, there is an $f \in \mathcal{L}^1(\mu)$ such that $d\nu = f d\mu$. Define $d|\nu| := |f| d\mu$. Observe this uniquely determines $|\nu|$ if it satisfies the uniqueness property in the bullet point above. So suppose that $d\nu = g d\rho$ for another σ-finite positive measure ρ on $(X, \mathcal{M})$ and $g \in \mathcal{L}^1(\rho)$. Consider $\mu + \rho$ on $(X, \mathcal{M})$ and observe that $\nu \ll \mu$, $\mu \ll \mu + \rho$, and $\rho \ll \mu + \rho$. Hence

$$d\mu = \frac{d\mu}{d(\mu+\rho)} d(\mu+\rho) \qquad \text{and} \qquad d\rho = \frac{d\rho}{d(\mu+\rho)} d(\mu+\rho).$$

Since

$$f \frac{d\mu}{d(\mu+\rho)} d(\mu+\rho) = f d\mu = d\nu = g d\rho = g \frac{d\rho}{d(\mu+\rho)} d(\mu+\rho),$$

by Exercise 4.2.11(2) we have

$$f \frac{d\mu}{d(\mu+\rho)} = \frac{d\nu}{d(\mu+\rho)} = g \frac{d\rho}{d(\mu+\rho)} \qquad (\mu+\rho)\text{-a.e.}$$

This implies

$$|f| \frac{d\mu}{d(\mu+\rho)} = \left| f \frac{d\mu}{d(\mu+\rho)} \right| = \left| g \frac{d\rho}{d(\mu+\rho)} \right| = |g| \frac{d\rho}{d(\mu+\rho)} \qquad (\mu+\rho)\text{-a.e.}$$

Again by Exercise 4.2.11(2), $|f| d\mu = d|\nu| = |g| d\rho$, and thus ν satisfies the uniqueness condition in the bullet point. □

Exercise 4.3.11. Repeat Exercise 4.2.11 for ν a complex measure on $(X, \mathcal{M})$.

Facts 4.3.12. Suppose ν is a complex measure on $(X, \mathcal{M})$.

(1) $\nu \ll |\nu|$, as

$$|\nu(E)| = \left|\int_E f\, d\mu\right| \leq \int_E |f|\, d\mu = |\nu|(E) \qquad\qquad \forall, E \in \mathcal{M}.$$

(2) If ν is a finite signed measure ($\operatorname{Im}(\nu) = 0$), then $d\nu = (\chi_P - \chi_{P^c})d|\nu|$, and so $d|\nu|' = (\chi_P + \chi_{P^c})d|\nu| = d|\nu|$ for any Hahn decomposition $X = P \amalg P^c$ for ν. This means this new definition $|\nu|'$ for a complex measure agrees with the old definition $|\nu|$ for a finite signed measure.

(3) Observe that if $d\nu = f d\mu$, then

$$\begin{matrix} d\operatorname{Re}(\nu) = \operatorname{Re}(f)d\mu \\ d\operatorname{Im}(\nu) = \operatorname{Im}(f)d\mu \end{matrix} \qquad \Longrightarrow \qquad \begin{matrix} d|\operatorname{Re}(\nu)| = |\operatorname{Re}(f)|d\mu \\ d|\operatorname{Im}(\nu)| = |\operatorname{Im}(f)|d\mu \end{matrix}$$

Since $|f|^2 = |\operatorname{Re}(f)|^2 + |\operatorname{Im}(f)|^2$, we have

$$\frac{d|\nu|}{d\mu} = |f| = \left(|\operatorname{Re}(f)|^2 + |\operatorname{Im}(f)|^2\right)^{1/2} = \left(\left(\frac{d|\operatorname{Re}(\nu)|}{d\mu}\right)^2 + \left(\frac{d|\operatorname{Im}(\nu)|}{d\mu}\right)^2\right)^{1/2}.$$

Exercise 4.3.13. Suppose ν is a complex measure on $(X, \mathcal{M})$. Prove that $|\operatorname{Re}(\nu)| \leq |\nu|$, $|\operatorname{Im}(\nu)| \leq |\nu|$, and $|\nu| \leq |\operatorname{Re}(\nu)| + |\operatorname{Im}(\nu)|$ as $[0, \infty)$-valued functions on $\mathcal{M}$.

Exercise 4.3.14. Suppose ν is a complex measure on $(X, \mathcal{M})$.

(1) Prove that $L^1(\nu) = L^1(|\nu|)$.

(2) Show that for $f \in L^1(\nu)$,

$$\left|\int f d\nu\right| \leq \int |f| d|\nu|.$$

Exercise 4.3.15. In this exercise, we will show that

$$M := M(X, \mathcal{M}, \mathbb{C}) := \{\text{complex measures on } (X, \mathcal{M})\}$$

is a Banach space with $\|\nu\| := |\nu|(X)$.

(1) Prove that $\max\{\|\operatorname{Re}(\nu)\|, \|\operatorname{Im}(\nu)\|\} \leq \|\nu\| \leq 2\max\{\|\operatorname{Re}(\nu)\|, \|\operatorname{Im}(\nu)\|\}$.

(2) Show that if $(V, \|\cdot\|_V), (W, \|\cdot\|_W)$ are normed vector spaces, then $\|(v, w)\|_\infty := \max\{\|v\|, \|w\|\}$ is a norm on $V \oplus W$. Moreover, show that if $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ are complete, then so is $(V \oplus W, \|\cdot\|_\infty)$.

(3) Show that $M(X, \mathcal{M}, \mathbb{C}) = M(X, \mathcal{M}, \mathbb{R}) \oplus iM(X, \mathcal{M}, \mathbb{R})$, where $M(X, \mathcal{M}, \mathbb{R})$ was defined in Exercise 4.1.19.

(4) Show that $\|\cdot\|$ on $M(X, \mathcal{M}, \mathbb{C})$ is equivalent to $\|\cdot\|_\infty$ on $M(X, \mathcal{M}, \mathbb{R}) \oplus iM(X, \mathcal{M}, \mathbb{R})$. Deduce that $M(X, \mathcal{M}, \mathbb{C})$ is complete.

Definition 4.3.16. A complex Borel measure ν on a topological space $(X, \mathcal{T})$ is called *regular* if $|\nu|$ is regular.

Exercise 4.3.17. Repeat Exercise 4.2.13 for a complex Borel measure ν, where (2) is replaced by

(2') $\operatorname{Re}(\nu)$ and $\operatorname{Im}(\nu)$ are regular signed measures.

4.4. **Lebesgue differentiation.** Here, I will be following notes from a graduate course I took in Fall 2005 at UC Berkeley from Sarason. We will treat differentiation of $f \in L^1(\lambda^n)$, and we'll then explain how to extend these results to

$$L^1_{\text{loc}} := L^1_{\text{loc}}(\lambda^n) := \{f : \mathbb{R}^n \to \mathbb{C} | f \text{ is integrable on bounded measurable sets}\}.$$

Definition 4.4.1. A *cube* in $\mathbb{R}^n$ is a set $Q \subset \mathbb{R}^n$ of the form $Q = \prod_{k=1}^n I_k$ where each I_k is a closed interval of the same length, which we denote by $\ell(Q)$.

- For $x \in \mathbb{R}^n$, define $\mathcal{C}(x) := \{\text{cubes } Q | x \in Q \text{ and } 0 < \ell(Q) < \infty\}$.
- For Q a cube and $r > 0$, rQ is the cube with the same center as Q, but with $\ell(rQ) = r\ell(Q)$.

Our goal is to prove the following theorem.

Theorem 4.4.2 (Lebesgue Differentiation)**.** *For all $f \in L^1_{\text{loc}}$,*

$$\lim_{\substack{\ell(Q)\to 0\\ x\in Q}} \frac{1}{\lambda^n(Q)} \int_Q f\, d\lambda^n = f(x) \qquad a.e. \qquad \text{(LDT)}$$

As a direct corollary, we get (for $n = 1$):

Theorem 4.4.3 (Fundamental Theorem of Calculus)**.** *Suppose $f \in \mathcal{L}^1(\lambda)$. Define $F(x) := \int_{(-\infty,x)} f\, d\lambda$. Then $F'(x) = f(x)$ a.e.*

Proof. Observe

$$\lim_{h\to 0} \frac{F(x+h) - F(x)}{h} = \lim_{\substack{h\to 0\\ x\in Q_h:=[x,x+h]}} \frac{1}{\lambda(Q_h)} \int_{Q_h} f\, d\lambda \underset{\text{(LDT)}}{=} f(x) \qquad a.e. \qquad \square$$

Definition 4.4.4 (Hardy-Littlewood Maximal Function)**.** For $f \in L^1_{\text{loc}}$, define $Mf := \mathbb{R}^n \to [0,\infty]$ by

$$(Mf)(x) := \sup\left\{ \frac{1}{\lambda^n(Q)} \int_Q |f|\, d\lambda^n \middle| Q \in \mathcal{C}(x)\right\}.$$

The function $M : L^1_{\text{loc}} \to \{f : \mathbb{R}^n \to [0,\infty]\}$ is called the *Hardy-Littlewood maximal function.*

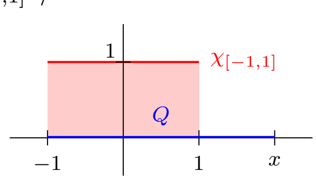
Facts 4.4.5. The Hardy-Littlewood maximal function satisfies the following properties:

(1) $M(rf) = |r| \cdot Mf$ for all $r \in \mathbb{R}$.
(2) $M(f+g) \le Mf + Mg$ for all $f, g \in L^1_{\text{loc}}$.
(3) $Mf > 0$ everywhere unless $f = 0$ a.e.
(4) Mf is lower semicontinuous ($\{Mf > r\}$ is open for all $r \in \mathbb{R}$), and thus measurable.

Example 4.4.6. For $\chi_{[-1,1]} : \mathbb{R} \to \mathbb{C}$,

$$M\chi_{[-1,1]}(x) = \begin{cases} 1 & x \in [-1,1] \\ \dfrac{2}{1+|x|} & x \notin [-1,1] \end{cases}$$

and thus $M\chi_{[-1,1]} \notin L^1$. Here is a cartoon:



$$\frac{1}{\lambda(Q)} \int_Q \chi_{[-1,1]}\, d\lambda = \frac{2}{1+x}.$$

Exercise 4.4.7 (Sarason)**.** Prove that for the f defined below, $f \in L^1(\lambda)$, but $Mf \notin L^1_{\text{loc}}$:

$$f(x) := \begin{cases} \dfrac{1}{|x|(\ln|x|)^2} & \text{if } |x| \le \frac{1}{2} \\ 0 & \text{if } |x| > \frac{1}{2}. \end{cases}$$

Theorem 4.4.8 (Hardy-Littlewood Maximal, a.k.a. HLMT)**.** *There is a $c > 0$, only depending on n, such that for all $f \in L^1(\lambda^n)$ and $a > 0$,*

$$\lambda^n(\{Mf > a\}) \le c \cdot \frac{\|f\|_1}{a}.$$

Remark 4.4.9. The HLMT 4.4.8 is a generalization of Chebyshev's Inequality for a measure space $(X, \mathcal{M}, \mu)$: for all $a \ge 0$, $\int_{\{a \le |f|\}} |f|\, d\mu \ge a\mu(\{a \le |f|\})$. Hence for all $f \in L^1(\mu)$ and $a \ge 0$,

$$\mu(\{a \le |f|\}) \le \frac{\|f\|_1}{a}. \tag{4.4.10}$$

To prove the HLMT 4.4.8, we'll use a variation of the Vitali Covering Lemma. We'll prove the more general Vitali Covering Lemma, and I'll leave the exact variation that we'll use to prove the HLMT as an exercise.

Lemma 4.4.11 (Vitali Covering)**.** *Let $\mathcal{B}$ be some collection of open balls in $\mathbb{R}^n$, and let $U = \bigcup_{B \in \mathcal{B}} B$. If $c < \lambda^n(U)$, then there exist disjoint $B_1, \dots, B_k \in \mathcal{B}$ such that $\sum_{j=1}^k \lambda^n(B_j) > 3^{-n}c$.*

Proof. Since λ^n is regular, there is a compact $K \subset U$ such that $c < \lambda^n(K)$. Then there exist finitely many balls in $\mathcal{B}$ which cover K, say $A_1, \dots, A_m$. Define B_1 to be the largest (in terms of radius) of the A_i, and inductively for $j \ge 2$, define B_j to be the larges of the the A_i *disjoint* from $B_1, \dots, B_{j-1}$. Since there are finitely many A_i, this process terminates, giving $B_1, \dots, B_k$.

> **Trick.** If A_i is *not* one of $B_1, \dots, B_k$, there is a smallest $1 \le j \le k$ such that $A_i \cap B_j \neq \emptyset$. Then $\text{rad}(A_i) \le \text{rad}(B_j)$, so $A_i \subset 3B_j$, where $3B_j$ has the same center as B_j, but three times the radius.

Then $K \subset \bigcup^k 3B_j$, so

$$c < \lambda^n(K) \le \sum^k \lambda^n(3B_j) = 3^n \sum^k \lambda^n(B_j). \qquad \square$$

Exercise 4.4.12 (Sarason, variation of Vitali Covering Lemma 4.4.11)**.** Suppose $E \subset \mathbb{R}^n$ (not assumed to be Borel measurable) and let $\mathcal{C}$ be a family of cubes covering E such that

$$\sup\{\ell(Q) | Q \in \mathcal{C}\} < \infty.$$

Show there exists a sequence $(Q_k) \subset \mathcal{C}$ of disjoint cubes such that

$$\sum_{k=1}^\infty \lambda^n(Q_k) \ge 5^{-n}(\lambda^n)^*(E).$$

Hint: Inductively choose Q_k such that $2\ell(Q_k)$ is larger than the sup of the lengths of all cubes which do not intersect $Q_1, \dots, Q_{k-1}$, with $Q_0 = \emptyset$ by convention.

Proof of HLMT 4.4.8. Suppose $f \in L^1(\lambda^n)$ and $a > 0$. Let $E = \{a < Mf\}$ and

$$\mathcal{C} = \left\{ \text{cubes } Q \middle| a < \frac{1}{\lambda^n(Q)} \int_Q |f| \, d\lambda^n \right\}.$$

By definition, the cubes in $\mathcal{C}$ cover E. Observe that $a < \ell(Q)^{-n}\|f\|_1$ implies $\ell(Q) < \left(\frac{\|f\|_1}{a}\right)^{1/n}$. By Exercise 4.4.12, there is a sequence $(Q_i) \subset \mathcal{C}$ of disjoint cubes such that $\sum \lambda^n(Q_i) \geq 5^{-n}\lambda^n(E)$. Then

$$\lambda^n(E) \leq 5^n \sum \lambda^n(Q_i) \leq 5^n \sum \frac{1}{a} \int_{Q_i} |f| \, d\lambda^n \leq 5^n \cdot \frac{\|f\|_1}{a}. \qquad \square$$

Proof of the Lebesgue Differentiation Theorem 4.4.2.
Step 1: (LDT) for all $f \in L^1$ implies (LDT) for all $f \in L^1_{\text{loc}}$.

Proof. Suppose $f \in L^1_{\text{loc}}$. It suffices to show that for all $R > 0$, (LDT) holds a.e. $x \in Q_R(0) := \prod^n [-R, R]$. For $x \in Q_R(0)$ and $Q \in \mathcal{C}(x)$ with $\ell(Q) \leq 1$, the value of $\frac{1}{\ell(Q)^n} \int_Q f \, d\lambda^n$ only depends on $f(y)$ for $y \in Q_{R+1}(0)$. So we can replace f with $f\chi_{Q_{R+1}(0)} \in L^1$. $\square$

Step 2: (LDT) for all $f \in C_c(\mathbb{R}^n)$ implies (LDT) for all $f \in L^1$.

Proof. For $Q \in \mathcal{C}(0)$ and $f \in L^1$, define $(I_Q f)(x) := \frac{1}{\lambda^n(Q)} \int_{Q+x} f \, d\lambda^n$. Observe I_Q is linear, and $|I_Q f| \leq Mf$ everywhere. Now fix $f \in L^1$ and $\varepsilon > 0$. Let

$$E_\varepsilon := \left\{ x \in \mathbb{R}^n \middle| \limsup_{\substack{\ell(Q) \to 0 \\ Q \in \mathcal{C}(0)}} |I_Q f(x) - f(x)| > \varepsilon \right\}.$$

We'll show $(\lambda^n)^*(E_\varepsilon) = 0$, which implies $E_\varepsilon \in \mathcal{L}^n$ and $\lambda^n(E_\varepsilon) = 0$. If $\varepsilon' < \varepsilon$, then $E_\varepsilon \subset E_{\varepsilon'}$. Hence $\bigcup E_{1/n}$ has measure zero, which implies the result.
In order to show $(\lambda^n)^*(E_\varepsilon) = 0$, let $\delta > 0$. Since $C_c(\mathbb{R}^n) \subset L^1$ is dense, there is a continuous $g \in C_c(\mathbb{R}^n)$ such that $\|f - g\|_1 < \delta$. Then

$$\begin{aligned} |I_Q f - f| &= |I_Q(f - g) + (I_Q g - g) + (g - f)| \\ &\leq |I_Q(f - g)| + |(I_Q g - g)| + |(g - f)| \\ &\leq M(f - g) + \underbrace{|(I_Q g - g)|}_{\to 0} + |g - f| \end{aligned}$$

By assumption, as $\ell(Q) \to 0$ for $Q \in \mathcal{C}(0)$, $|(I_Q g - g)| \to 0$. Hence

$$E_\varepsilon \subset \left\{ \frac{\varepsilon}{2} < M(f - g) \right\} \cup \left\{ \frac{\varepsilon}{2} < |f - g| \right\}.$$

By the HLMT 4.4.8 and Chebyshev's Inequality (4.4.10),

$$\begin{aligned}
(\lambda^n)^*(E_\varepsilon) &\leq \lambda^n\left(\left\{\frac{\varepsilon}{2} < M(f-g)\right\}\right) + \lambda^n\left(\left\{\frac{\varepsilon}{2} < |f-g|\right\}\right) \\
&\leq \frac{c\|f-g\|_1}{\varepsilon/2} + \frac{\|f-g\|_1}{\varepsilon/2} \\
&= \frac{2(c+1)}{\varepsilon} \cdot \|f-g\|_1 \\
&< \frac{2(c+1)}{\varepsilon} \cdot \delta.
\end{aligned}$$

But $\delta > 0$ was arbitrary, so $(\lambda^n)^*(E_\varepsilon) = 0$. □

Step 3: (LDT) holds for all $g \in C_c(\mathbb{R}^n)$.

Proof. Observe that g is uniformly continuous. Let $\varepsilon > 0$, and pick $\delta > 0$ such that $x, y \in Q$ with $\ell(Q) < \delta$ implies $|g(x) - g(y)| < \varepsilon$. Then for all such Q,

$$\left| g(x) - \frac{1}{\lambda^n(Q)} \int_Q g(y)\, d\lambda^n(y) \right| \leq \frac{1}{\lambda^n(Q)} \int_Q |g(x) - g(y)|\, d\lambda^n(y) < \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, the result follows. □

Combining Steps 1-3 yields the result. □

Definition 4.4.13. Suppose $E \in \mathcal{L}^n$. A point $x \in E$ is called a *Lebesgue point of density of* E if

$$\lim_{\substack{\ell(Q)\to 0 \\ Q \in \mathcal{C}(x)}} \frac{\lambda^n(Q \cap E)}{\lambda^n(Q)} = 1.$$

Corollary 4.4.14. *For $E \in \mathcal{L}^n$, almost all points of E are Lebesgue points of density.*

Proof. Apply the Lebesgue Differentiation Theorem 4.4.2 to χ_E. □

Exercise 4.4.15 (Steinhaus Theorem, version 2)**.** Suppose that $A, B \subset \mathbb{R}$ are sets with positive Lebesgue measure. Prove that there is an interval I with $\lambda(I) > 0$ such that

$$I \subseteq A + B = \{a + b | a \in A \text{ and } b \in B\}.$$

Definition 4.4.16. For $f \in L^1(\lambda^n)$, $x \in \mathbb{R}^n$ is called a *Lebesgue point of* f if

$$\lim_{\substack{\ell(Q)\to 0 \\ Q \in \mathcal{C}(x)}} \frac{1}{\lambda^n(Q)} \int_Q |f - f(x)|\, d\lambda^n = 0.$$

Corollary 4.4.17. *For $f \in L^1_{\text{loc}}$, almost all points of $\mathbb{R}^n$ are Lebesgue points of f.*

Proof. As in the proof of the Lebesgue Differentiation Theorem 4.4.2, we may assume $f \in L^1$. Let $D \subset \mathbb{C}$ be a countable dense subset ($\mathbb{Q} + i\mathbb{Q}$ will suffice). For $d \in D$, set

$$E_d := \left\{ x \in \mathbb{R}^n \,\middle|\, \lim_{\substack{\ell(Q)\to 0 \\ Q \in \mathcal{C}(x)}} \frac{1}{\lambda^n(Q)} \int_Q |f - d| - |f(x) - d|\, d\lambda^n = 0 \right\}.$$

By the Lebesgue Differentiation Theorem 4.4.2, E_d^c is λ^n-null, which implies $E_d \in \mathcal{L}^n$. Set $E := \bigcap_{d\in D} E_d \in \mathcal{L}^n$, and observe $E^c = \bigcup_{d\in D} E_d^c$ is still λ^n-null. We claim that every $x \in E$ is a Lebesgue point of f. Indeed, if $x \in E$, then for all $d \in D$,

$$|f - f(x)| \leq |f - d| + |f(x) - d| = (|f - d| - |f(x) - d|) + 2|f(x) - d|.$$

This implies for all $d \in D$,

$$\begin{aligned}\limsup_{\substack{\ell(Q)\to 0\\ Q\in\mathcal{C}(x)}} \frac{1}{\lambda^n(Q)} \int_Q |f - f(x)|\, d\lambda^n &\leq 2|f(x) - d| + \underbrace{\limsup_{\substack{\ell(Q)\to 0\\ Q\in\mathcal{C}(x)}} \frac{1}{\lambda^n(Q)} \int_Q |f - d| - |f(x) - d|\, d\lambda^n}_{=0} \\ &= 2|f(x) - d|.\end{aligned}$$

But since D is dense in $\mathbb{C}$, we can approximate $f(x)$ by $d \in D$ up to any $\varepsilon > 0$. We conclude that x is a Lebesgue point of f. □

4.5. **Functions of bounded variation.** Recall that the Lebesgue-Stieltjes measures on $\mathbb{R}$ were constructed from non-decreasing right continuous functions $F : \mathbb{R} \to \mathbb{R}$. They enjoyed the properties of being a complete measure which is equal to the completion of the restriction to $\mathcal{B}_{\mathbb{R}}$, which is a regular Borel measure.

We can adapt this construction to get a complex measure from a function $F : \mathbb{R} \to \mathbb{C}$ with *bounded variation*.

Definition 4.5.1. For a function $F : \mathbb{R} \to \mathbb{C}$, define its *total variation* $T_F : \mathbb{R} \to [0, \infty]$ by

$$T_F(x) := \sup\left\{\sum_{i=1}^n |F(x_i) - F(x_{i-1})| \,\middle|\, n \in \mathbb{N} \text{ and } -\infty < x_0 < x_1 < \cdots < x_n = x\right\}.$$

Observe that T_F is a non-decreasing function. We say F has *bounded variation* if T_F is bounded, which is equivalent to $\lim_{x\to\infty} T_F(x) < \infty$. We define

$$\mathsf{BV} := \{F : \mathbb{R} \to \mathbb{C} | F \text{ has bounded variation}\}.$$

Exercise 4.5.2. Prove that for all $a, b \in \mathbb{R}$ with $a < b$ and $F : \mathbb{R} \to \mathbb{C}$,

$$T_F(b) = T_F(a) + \sup\left\{\sum_{i=1}^n |F(x_i) - F(x_{i-1})| \,\middle|\, n \in \mathbb{N} \text{ and } a = x_0 < x_1 < \cdots < x_n = b\right\}.$$

The sup on the right hand side is called the *total variation of F on $[a, b]$*. We say F has bounded variation on $[a, b]$ if this number is bounded.

Exercise 4.5.3. Show that if F is differentiable and F' is bounded, then $F \in \mathsf{BV}[a, b]$ for all $a < b$ in $\mathbb{R}$.

Facts 4.5.4. Here are some facts about functions with bounded variation.

(BV1) If $F : \mathbb{R} \to \mathbb{R}$ is increasing, then $F \in \mathsf{BV}$ if and only if F is bounded.

Proof. For any $-\infty < x_0 < x_1 < \cdots < x_n = x$,

$$\sum_{i=1}^{n} |F(x_i) - F(x_{i-1})| = F(x) - F(x_0).$$

Hence T_F is bounded if and only if F is bounded. □

(BV2) $F \in \mathsf{BV}$ if and only if $T_F \in \mathsf{BV}$.

Proof. If $F \in \mathsf{BV}$, then $T_F : \mathbb{R} \to [0, \infty]$ is increasing and bounded, and thus in BV by (BV1). Conversely, if $T_F \in \mathsf{BV}$, then T_F is bounded by (BV1), and thus $F \in \mathsf{BV}$. □

(BV3) BV is a complex vector space which is closed under complex conjugation.

Proof. The triangle inequality implies $T_{F+G} \leq T_F + T_G$, homogeneity ($|wz| = |w| \cdot |z|$) implies $T_{zF} \leq |z| \cdot T_F$, and $|\overline{z}| = |z|$ for $z \in \mathbb{C}$ implies $T_{\overline{F}} = T_F$. □

(BV4) $F \in \mathsf{BV}$ if and only if $\mathrm{Re}(F), \mathrm{Im}(F) \in \mathsf{BV}$.

Proof. Just observe that $\mathrm{Re}(F) = \frac{1}{2}(F + \overline{F})$ and $\mathrm{Im}(F) = \frac{1}{2i}(F - \overline{F})$, so the result follows from (BV3). □

(BV5) If $F : \mathbb{R} \to \mathbb{R}$ and $F \in \mathsf{BV}$, then $T_F \pm F$ are increasing (and in BV).

Proof. Suppose $a < b$ in $\mathbb{R}$. Let $\varepsilon > 0$, and choose $x_0 < x_1 < \cdots < x_n = a$ such that

$$\sum_{i=1}^{n} |F(x_i) - F(x_{i-1})| \geq T_F(a) - \varepsilon.$$

Then since $F(b) = (F(b) - F(a)) + F(a)$,

$$\begin{aligned} T_F(b) \pm F(b) &\geq \underbrace{\sum_{i=1}^{n} |F(x_i) - F(x_{i-1})| + |F(b) - F(a)|}_{\leq T_F(b)} \pm F(b) \\ &= \sum_{i=1}^{n} |F(x_i) - F(x_{i-1})| + \underbrace{|F(b) - F(a)| \pm (F(b) - F(a))}_{\geq 0} \pm F(a) \\ &\geq T_F(a) - \varepsilon \pm F(a) \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, we have $T_F \pm F$ is increasing. (The parenthetical follows from (BV3).) □

(BV6) If $F : \mathbb{R} \to \mathbb{C}$, then $F \in \mathsf{BV}$ if and only if $F = \sum_{k=0}^{3} i^k F_k$ where $F_k : \mathbb{R} \to \mathbb{R}$ is bounded and increasing for $k = 0, 1, 2, 3$.

Proof. By (BV4), $F \in \mathsf{BV}$ if and only if $\mathrm{Re}(F), \mathrm{Im}(F) \in \mathsf{BV}$, so we may assume $F : \mathbb{R} \to \mathbb{R}$. If $F \in \mathsf{BV}$, just observe

$$F = \frac{1}{2}(T_F + F) - \frac{1}{2}(T_F - F).$$

The converse follows from (BV1) and (BV3). □

(BV7) If $F \in \mathsf{BV}$, then $F(x+) := \lim_{y \searrow x} F(y)$ and $F(x-) := \lim_{y \nearrow x} F(y)$ exist for all $x \in \mathbb{R}$, as do $F(\pm\infty) := \lim_{y \to \pm\infty} F(y)$.

Proof. This follows from (BV6). □

Remark 4.5.5. For an $\mathbb{R}$-valued $F \in \mathsf{BV}$, we call

$$F = \frac{1}{2}(T_F + F) - \frac{1}{2}(T_F - F)$$

the Jordan decomposition of F. We call $\frac{1}{2}(T_F \pm F)$ the positive/negative variations of F respectively.

Definition 4.5.6. The space of *normalized* functions of bounded variation is

$$\mathsf{NBV} := \{F \in \mathsf{BV} | F \text{ is right continuous and } F(-\infty) = 0\}.$$

Observe that NBV is a complex vector subspace of BV closed under complex conjugation.

Exercise 4.5.7. Suppose $f \in L^1(\lambda)$ where λ is Lebesgue measure on $\mathbb{R}$. Consider the function $F : \mathbb{R} \to \mathbb{C}$ by $F(x) = \int_{-\infty}^{x} f(t)\, dt$.

(1) Prove directly from the definitions that $F \in \mathsf{NBV}$.
(2) Describe T_F to the best of your ability. Justify your answer.

Lemma 4.5.8. *Suppose $F : \mathbb{R} \to \mathbb{C}$.*

(1) If $F \in \mathsf{BV}$, then $T_F(-\infty) = 0$.
(2) If moreover F is right-continuous, then so is T_F.

Hence $F \in \mathsf{NBV}$ implies $T_F \in \mathsf{NBV}$.

Proof.
(1) Let $\varepsilon > 0$. For $x \in \mathbb{R}$, choose $x_0 < x_1 < \cdots < x_n = x$ such that

$$\sum_{i=1}^{n} |F(x_i) - F(x_{i-1})| \geq T_F(x) - \varepsilon.$$

By Exercise 4.5.2

$$T_F(x) - T_F(x_0) \geq T_F(x) - \varepsilon,$$

and thus $T_F(y) \leq \varepsilon$ for all $y \leq x_0$. Since $\varepsilon > 0$ was arbitrary, $T_F(-\infty) = 0$.
(2) Now suppose F is right continuous. Fix $x \in \mathbb{R}$, and define

$$\alpha := \lim_{y \searrow x} T_F(y) - T_F(x).$$

To show $\alpha = 0$, fix $\varepsilon > 0$, and let $\delta > 0$ such that $0 < h < \delta$ implies both $|F(x+h)-F(x)| < \varepsilon$ and

$$T_F(x+h) - T_F(x) - \alpha = T_F(x+h) - \lim_{y \searrow x} T_F(y) < \varepsilon. \tag{4.5.9}$$

Now fixing $0 < h < \delta$, by Exercise 4.5.2, there are $x = x_0 < x_1 < \cdots < x_n = x+h$ such that

$$\frac{3}{4}\alpha \leq \frac{3}{4}(T_F(x+h) - T_F(x)) \leq \sum_{j=1}^{n} |F(x_j) - F(x_{j-1})|$$

which by right continuity of F and the choice of δ implies

$$\frac{3}{4}\alpha - \varepsilon \leq \frac{3}{4}(T_F(x+h) - T_F(x)) - |F(x_1) - F(x_0)| \leq \sum_{j=2}^{n} |F(x_j) - F(x_{j-1})|. \tag{4.5.10}$$

Again using Exercise 4.5.2,

$$\frac{3}{4}\alpha \leq \frac{3}{4}(T_F(x_1) - T_F(x)) \leq \sum_{i=1}^{k} |F(t_i) - F(t_{i-1})|. \tag{4.5.11}$$

Combining these inequalities, we have

$$\begin{aligned}
\alpha + \varepsilon &> T_F(x+h) - T_F(x) && \text{by (4.5.9)}\\
&\geq \sum_{i=1}^{k} |F(t_i) - F(t_{i-1})| + \sum_{j=2}^{n} |F(x_j) - F(x_{j-1})| && \text{by Exercise 4.5.2}\\
&\geq \frac{3}{4}\alpha + \frac{3}{4}\alpha - \varepsilon && \text{by (4.5.10) and (4.5.11)}\\
&= \frac{3}{2}\alpha - \varepsilon.
\end{aligned}$$

This implies $\alpha \leq 4\varepsilon$, but since $\varepsilon > 0$ was arbitrary, $\alpha = 0$. □

Theorem 4.5.12.

(1) *If ν is a complex Borel measure on $\mathbb{R}$, then $F_\nu(x) := \nu((-\infty, x])$ defines a function in* NBV.

(2) *If $F \in$ NBV, there is a unique complex Borel measure ν_F such that $F(x) = \nu_F((-\infty, x])$.*

Proof. For a complex Borel measure ν, we have $\nu = \sum_{k=0}^{3} i^k \nu_k$ where each ν_k is a finite positive measure. If we set $F_k := \nu_k((-\infty, x])$, then F_k is increasing and right continuous, $F_k(-\infty) = 0$, and $F_k(\infty) = \nu_k(\mathbb{R}) < \infty$. Thus each $F_k \in$ NBV, and thus $F_\nu := \sum_{k=0}^{3} i^k F_k$ is in NBV.

Conversely, by (BV6) and Lemma 4.5.8, any $F \in$ NBV can be written as $F = \sum_{k=0}^{3} i^k F_k$ where each $F_k : \mathbb{R} \to [0, \infty)$ is increasing and in NBV. By the Lebesgue-Stieltjes construction, for each F_k, there is a finite regular Borel measure ν_k on $\mathbb{R}$ with $\nu_k((-\infty, x]) = F_k(x)$. Setting $\nu := \sum_{k=0}^{3} i^k \nu_k$ gives a complex Borel measure such that $F(x) = \nu((-\infty, x])$. Uniqueness follows by being determined on h-intervals together with the $\pi - \lambda$ Theorem. □

Exercise 4.5.13. Suppose $F \in$ NBV, and let ν_F be the corresponding complex Borel measure from Theorem 4.5.12.

(1) Prove that ν_F is regular.

(2) Prove that $|\nu_F| = \nu_{T_F}$.
One could proceed as follows.
(a) *Define* $G(x) := |\nu_F|((-\infty, x])$. *Show that* $|\nu_F| = \nu_{T_F}$ *if and only if* $G = T_F$.
(b) *Show* $T_F \leq G$.
(c) *Show that* $|\nu_F(E)| \leq \nu_{T_F}(E)$ *whenever* E *is an interval.*
(d) *Show that* $|\nu_F| \leq \nu_{T_F}$.

Exercise 4.5.14. Show that if $F \in \mathsf{NBV}$, then $(\nu_F)_\pm = \nu_{\frac{1}{2}(T_F \pm F)}$, i.e., the positive/negative variations of F exactly correspond to the positive/negative parts of the Jordan decomposition of ν_F.
Hint: Use Exercise 4.5.13.

4.6. **Bounded variation, differentiation, and absolute continuity.** We now want to connect functions of bounded variation and ordinary differentiation on $\mathbb{R}$.

Definition 4.6.1. Recall that $F : \mathbb{R} \to \mathbb{C}$ is called *absolutely continuous* if for every $\varepsilon > 0$, there exists $\delta > 0$ such that for any finite set of disjoint open intervals $(a_1, b_1), \ldots, (a_n, b_n)$,

$$\sum_{i=1}^{n}(b_i - a_i) < \delta \qquad \Longrightarrow \qquad \sum_{i=1}^{n}|F(b_i) - F(a_i)| < \varepsilon.$$

Exercise 4.6.2. Suppose $F \in \mathsf{NBV}$. Show F is absolutely continuous if and only if T_F is absolutely continuous.
Hint: Use Exercise 4.5.2.

Proposition 4.6.3. *If* $F \in \mathsf{NBV}$, *then* F *is absolutely continuous if and only if* $\nu_F \ll \lambda$.

Proof.

Claim. *We may assume* F *is* $[0, \infty)$*-valued and increasing. Thus* $\nu_F = \mu_F$ *is an honest Lebesgue-Stieltjes measure.*

Proof. By Exercises 4.3.13 and 4.5.13(2), $\nu_F \ll \lambda$ if and only if $|\nu_F| = \nu_{T_F} \ll \lambda$. By Exercise 4.6.2, F is absolutely continuous if and only if T_F is absolutely continuous. Hence we may replace F with $T_F \in \mathsf{NBV}$ which is $[0, \infty)$-valued and increasing. $\square$

That $\mu_F \ll \lambda$ for a Lebesgue-Stieltjes measure is equivalent to absolute continuity of a bounded, right-continuous $F : \mathbb{R} \to [0, \infty)$ with $F(-\infty) = 0$ now follows Exercise 2.5.20. We provide a proof here for completeness and convenience using Proposition 4.2.6 which states:

- $\mu_F \ll \lambda$ if and only if for all $\varepsilon > 0$, there is a $\delta > 0$ such that for all $E \in \mathcal{M}$, $\mu_F(E) < \varepsilon$ whenever $\lambda(E) < \delta$.

First, suppose $\mu_F \ll \lambda$. For any finite set of disjoint h-intervals $((a_i, b_i])_{i=1}^n$, we have

$$\sum_{i=1}(b_i - a_i) = \lambda\left(\coprod (a_i, b_i]\right) < \delta \qquad \Longrightarrow \qquad \mu_F\left(\coprod (a_i, b_i]\right) = \sum_{i=1}^{n} \mu_F((a_i, b_i]) < \varepsilon.$$

This immediately implies F is absolutely continuous.

Conversely, suppose F is absolutely continuous, and $\varepsilon > 0$. Pick $\delta > 0$ for F as in the definition of absolute continuity for any $0 < \varepsilon' < \varepsilon$. Suppose $E \in \mathcal{L}$ such that $\lambda(E) < \delta$. By outer regularity of λ and μ_F (by Exercise 4.5.13(1)), there is an open set U with $E \subset U$ such that $\lambda(U) < \delta$. Then U is a countable disjoint union of open intervals by Exercise 1.1.24, say $U = \coprod(a_i, b_i)$. For each $n \in \mathbb{N}$,

$$\sum_{i=1}^{n}(b_i - a_i) \leq \lambda(U) < \delta \qquad \Longrightarrow \qquad \sum_{i=1}^{n} \mu_F((a_i, b_i]) = \sum_{i=1}^{n} F(b_i) - F(a_i) < \varepsilon'.$$

Taking the limit as $n \to \infty$, we have

$$\sum_{i=1}^{\infty}(b_i - a_i) \leq \lambda(U) < \delta \qquad \Longrightarrow \qquad \sum_{i=1}^{\infty} \mu_F((a_i, b_i]) = \sum_{i=1}^{\infty} F(b_i) - F(a_i) \leq \varepsilon' < \varepsilon.$$

Hence $\mu_F \ll \lambda$. $\square$

Exercise 4.6.4. Prove that if $F : [a, b] \to \mathbb{C}$ with $a, b \in \mathbb{R}$ is absolutely continuous, then $F \in \mathsf{BV}[a, b]$.

Exercise 4.6.5 (cf. Folland Thm. 3.22)**.** Denote by λ^n Lebesgue measure on $\mathbb{R}^n$. Suppose ν is a regular signed or complex Borel measure on $\mathbb{R}^n$ which is finite on compact sets (and thus Radon and σ-finite). Let $d\nu = d\rho + f d\lambda^n$ be its Lebesgue-Radon-Nikodym representation from Theorem 4.3.9. Then for λ^n-a.e. $x \in \mathbb{R}^n$,

$$\lim_{\substack{\ell(Q) \to 0 \\ Q \in \mathcal{C}(x)}} \frac{\nu(Q)}{\lambda^n(Q)} = f(x).$$

Hint: One could proceed as follows.

(1) Show that $d|\nu| = d|\rho| + |f| d\lambda^n$. Deduce that ρ and $f d\lambda^n$ are regular, and $f \in L^1_{\text{loc}}$.
(2) Use the Lebesgue Differentiation Theorem to reduce the problem to showing

$$\lim_{\substack{\ell(Q) \to 0 \\ Q \in \mathcal{C}(x)}} \frac{|\rho|(Q)}{\lambda^n(Q)} = 0 \qquad \lambda^n\text{-a.e. } x \in \mathbb{R}^n.$$

Thus we may assume ρ is positive.
(3) Since $\rho \perp \lambda^n$, pick $P \subset \mathbb{R}^n$ Borel measurable such that $\rho(P) = \lambda^n(P^c) = 0$. For $a > 0$, define

$$E_a := \left\{ x \in P \,\middle|\, \lim_{\substack{\ell(Q) \to 0 \\ Q \in \mathcal{C}(x)}} \frac{|\rho|(Q)}{\lambda^n(Q)} > a \right\}.$$

Let $\varepsilon > 0$. Since ρ is regular, there is an open $U_\varepsilon \supset P$ such that $\rho(U_\varepsilon) < \varepsilon$. Adapt the proof of the HLMT to show there is a constant $c > 0$, depending only on n, such that for all $a > 0$,

$$\lambda^n(E_a) \leq c \cdot \frac{\rho(U_\varepsilon)}{a} = c \cdot \frac{\varepsilon}{a}$$

(Choose your family of cubes to be contained in U_ε.) Deduce that $\lambda^n(E_a) = 0$.

Lemma 4.6.6. *Suppose that $F : \mathbb{R} \to \mathbb{R}$ is increasing or $F \in \mathsf{BV}$.*

(1) The set of points at which F is discontinuous is countable.

(2) *Suppose in addition F is right continuous. Let μ_F be the corresponding (regular, σ-finite) Lebesgue-Stieltjes measure, and let $d\lambda = d\rho + f d\lambda$ be its Lebesgue-Radon-Nikodym representation from Theorem 4.3.9. Then F is differentiable λ-a.e. with $F'(x) = f(x)$ λ-a.e.*

(3) *Setting $G(x) := \lim_{y \searrow x} F(y)$, F and G are differentiable a.e., with $F' = G'$ a.e.*

Proof. Since every $F \in \mathsf{BV}$ is a linear combination of four increasing, bounded functions $\mathbb{R} \to \mathbb{R}$ by (BV6), we may assume $F : \mathbb{R} \to \mathbb{R}$ is an arbitrary increasing function.

(1) Observe that at every discontinuity $x \in \mathbb{R}$, the open interval

$$\left(\lim_{y \nearrow x} F(y), \lim_{y \searrow x} F(y)\right) \neq \emptyset$$

and thus contains a rational point. Since F is increasing, these open intervals at distinct discontinuities will be disjoint, and we can construct an injective mapping from the set of discontinuities to $\mathbb{Q}$.

(2) Suppose in addition that F is right-continuous. Let $D \subset \mathbb{R}$ be the countable set of discontinuities of F, and observe that $\lambda(D) = 0$. By Exercise 4.6.5,

$$\lim_{\substack{\ell(Q) \to 0 \\ Q \in \mathcal{C}(x)}} \frac{\mu_F(Q)}{\lambda(Q)} = f(x) \qquad \lambda\text{-a.e. } x \in \mathbb{R}.$$

Now observe that for $x \notin D$ and $h > 0$, by Exercise 2.5.9,

$$\mu_F([x, x+h]) = \lim_{y \nearrow x} \mu_F((y, x+h]) = \lim_{y \nearrow x} F(x+h) - F(y) = F(x+h) - F(x)$$

If in addition $x - h \notin D$, then we also have

$$\mu_F([x-h, x]) = F(x) - F(x-h).$$

Since D is countable and F is increasing, we may take the following limit for $x \in D^c$ along $h \to 0$ such that $x - |h| \notin D$ to conclude that

$$\begin{aligned} \lim_{h \to 0} \frac{F(x+h) - F(x)}{h} &= \lim_{\substack{h \to 0 \\ x - |h| \notin D}} \frac{\mu_F([\min\{x, x+h\}, \max\{x, x+h\}])}{\lambda([\min\{x, x+h\}, \max\{x, x+h\}])} \\ &= f(x) \qquad \lambda\text{-a.e. } x \in D^c \text{ by Exercise 4.6.5.} \end{aligned}$$

<u>(3), Step 1:</u> G is increasing and right-continuous, and thus G is differentiable a.e. by (2).

If $a < b$ in $\mathbb{R}$, then since F is increasing,

$$G(a) = \lim_{x \searrow a} F(x) = \lim_{\substack{x \searrow a \\ a < x < b}} F(x) \leq F(b) \leq G(b),$$

and thus G is increasing. To show G is right continuous at $x \in \mathbb{R}$, let $\varepsilon > 0$. Since $G(x) = \lim_{y \searrow x} F(y)$, we can pick $\delta' > 0$ such that $0 < h' < \delta'$ implies $F(x+h') - G(x) < \varepsilon$. Then for any $0 \leq h < \delta < h' < \delta'$,

$$G(x+h) - G(x) \leq F(x+h') - G(x) < \varepsilon.$$

(2), Step 2: Setting $H := G - F \geq 0$, H' exists and is zero a.e.

First, note $H(d) > 0$ for all $d \in D$, and

$$\sum_{\substack{d\in D\\ |d|<N}} H(d) = \sum_{\substack{d\in D\\ |d|<N}} G(d) - F(d) \leq G(N) - F(N) < \infty. \tag{4.6.7}$$

Claim. *Setting $\eta := \sum_{d\in D} H(d)\delta_d$ where δ_d is the Dirac point mass at d, η is a regular Borel measure such that $\eta \perp \lambda$.*

Proof. Observe η is finite on compact sets by (4.6.7). We define $h : \mathbb{R} \to \mathbb{R}$ by picking an arbitrary $r_0 \in D^c$, setting $h(r_0) = 0$, and setting

$$h(r) := \begin{cases} \sum_{\substack{d\in D\\ r_0<d\leq r}} H(d) & \text{if } r > r_0 \\ -\sum_{\substack{d\in D\\ r<d<r_0}} H(d) & \text{if } r < r_0. \end{cases}$$

Observe that h is increasing and right-continuous, and by construction, the Lebesgue-Stieltjes measure $\mu_h = \eta$, which is thus regular. Since η is supported on D and $\lambda(D) = 0$, we have $\eta \perp \lambda$. □

Now for $|h| \neq 0$, again by Exercise 4.6.5,

$$\left|\frac{H(x+h) - H(x)}{h}\right| \leq \frac{H(x+h) + H(x)}{|h|} \leq 2\frac{\eta([x-|h|, x+|h|])}{\lambda([x-|h|, x+|h|])} \xrightarrow{h\to 0} 0 \quad \text{a.e. } x \in \mathbb{R}.$$

We conclude that $H' = 0$ a.e.

This concludes the proof. □

Facts 4.6.8. Suppose $F \in \mathsf{NBV}$, and let $\nu_F = \rho_F + f d\lambda$ where $f \in L^1(\lambda)$ be the Lebesgue-Radon-Nikodym Representation of ν_F from Theorem 4.3.9.

($\mathsf{NBV}'1$) F' exists λ-a.e. with $F' = f \in L^1(\lambda)$.

Proof. By (BV6), $F = \sum_{k=0}^3 i^k F_k$ where each $F_k : \mathbb{R} \to \mathbb{R}$ is an increasing right-continuous function in NBV. Let $\mu_{F_k} = \rho_{F_k} + f_k d\lambda$ where $f_k \in L^1(\lambda)$ for $k = 0, 1, 2, 3$ be the Lebesgue-Radon-Nikodym representation of the Lebesgue-Stieltjes measure μ_{F_k} from Theorem 4.2.9. By Lemma 4.6.6(2), F_k' exists λ-a.e., and $F_k' = f_k$ λ-a.e. By the proof of the Complex Lebesgue-Radon-Nikodym Theorem 4.3.9, we have $f = \sum_{k=0}^3 i^k f_k$. Hence

$$F' = \sum_{k=0}^{3} i^k F_k' = \sum_{k=0}^{3} i^k f_k = f \qquad \lambda\text{-a.e.}$$ □

($\mathsf{NBV}'2$) $\nu_F \perp \lambda$ if and only if $F' = 0$ a.e.

Proof. This follows immediately from ($\mathsf{NBV}'1$) and the Lebesgue-Radon-Nikodym Representation of ν_F. □

(NBV′3) $\nu_F \ll \lambda$ if and only if $F(x) = \int_{-\infty}^x F'(t)\,dt$.

Proof. Observe $\nu_F \ll \lambda$ if and only if $\rho_F = 0$ if and only if $d\nu_F = F' d\lambda$ by (NBV′1). This last condition is equivalent to

$$F(x) = \nu_F((-\infty, x]) = \int_{-\infty}^x F'(t)\,dt. \qquad \square$$

Proposition 4.6.9. *The following are equivalent for $F : \mathbb{R} \to \mathbb{C}$.*

(1) $F \in \mathsf{NBV}$ is absolutely continuous.
(2) F is differentiable a.e., $F' \in L^1(\lambda)$, and $F(x) = \int_{-\infty}^x F'(t)\,dt$.
(3) There is an $f \in L^1(\lambda)$ such that $F(x) = \int_{-\infty}^x f(t)\,dt$.

Proof.
(1) ⇒ (2): If $F \in \mathsf{NBV}$ is absolutely continuous, then $\nu_F \ll \lambda$ by Proposition 4.6.3. By (NBV′1), F is differentiable a.e. with $F' \in L^1(\lambda)$, and by (NBV′3), $F(x) = \int_{-\infty}^x F(t)\,dt$.
(2) ⇒ (3): Trivial.
(3) ⇒ (1): Since $f \in L^1(\lambda)$, $d\nu := f d\lambda$ is a complex Borel measure. Thus

$$F(x) = \int_{-\infty}^x f(t)\,dt = \nu((-\infty, x])$$

defines a function in NBV by Theorem 4.5.12(1). Since $\nu \ll \lambda$ by construction, F is absolutely continuous by Proposition 4.6.3. $\square$

We leave the proof of the following corollary to the reader.

Corollary 4.6.10 (Fundamental Theorem of Calculus for Lebesgue Integrals)**.** *Let $a, b \in \mathbb{R}$ with $a < b$, and suppose $F : [a, b] \to \mathbb{C}$. The following are equivalent.*

(1) F is absolutely continuous on $[a, b]$.
(2) F is differentiable a.e. on $[a, b]$, $F' \in L^1([a, b], \lambda)$, and $F(x) - F(a) = \int_a^x F'(t)\,dt$.
(3) $F(x) - F(a) = \int_a^x f(t)\,dt$ for some $f \in L^1([a, b], \lambda)$.

Exercise 4.6.11 (Folland §3.5, #37)**.** Show that $F : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous (there is an $M > 0$ such that $|F(x) - F(y)| \le M|x - y|$ for all $x, y \in \mathbb{R}$) if and only if F is absolutely continuous and $|F'| \le M$ a.e.

5.1. **Normed spaces and linear maps.** For this section, X will denote a vector space over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. (We will assume $\mathbb{F} = \mathbb{C}$ unless stated otherwise.)

Definition 5.1.1. A *seminorm* on X is a function $\|\cdot\| : X \to [0, \infty)$ which is

- (homogeneous) $\|\lambda x\| = |\lambda| \cdot \|x\|$
- (subadditive) $\|x + y\| \leq \|x\| + \|y\|$

We call $\|\cdot\|$ a *norm* if in addition it is

- (definite) $\|x\| = 0$ implies $x = 0$.

Recall that given a norm $\|\cdot\|$ on a vector space X, $d(x, y) := \|x - y\|$ is a metric which induces the *norm topology* on X. Two norms $\|\cdot\|_1, \|\cdot\|_2$ are called *equivalent* if there is a $c > 0$ such that

$$c^{-1}\|x\|_1 \leq \|x\|_2 \leq c\|x\|_1 \qquad \forall x \in X.$$

Exercise 5.1.2. Show that all norms on $\mathbb{F}^n$ are equivalent. Deduce that a finite dimensional subspace of a normed space is closed.
Note: You may assume that the unit ball of $\mathbb{F}^n$ is compact in the Euclidean topology.

Exercise 5.1.3. Show that two norms $\|\cdot\|_1, \|\cdot\|_2$ on X are equivalent if and only if they induce the same topology.

Definition 5.1.4. A *Banach space* is a normed vector space which is complete in the induced metric topology.

Examples 5.1.5.

(1) If X is an LCH topological space, then $C_0(X)$ and $C_b(X)$ are Banach spaces.
(2) If $(X, \mathcal{M}, \mu)$ is a measure space, $\mathcal{L}^1(X, \mathcal{M}, \mu)$ is a Banach space.
(3) $\ell^1 := \{(x_n) \subset \mathbb{F}^\infty | \sum |x_n| < \infty\}$

Definition 5.1.6. Suppose $(X, \|\cdot\|)$ is a normed space and $(x_n) \subset (X, \|\cdot\|)$ is a sequence. We say $\sum x_n$ *converges to* $x \in X$ if $\sum^N x_n \to x$ as $N \to \infty$. We say $\sum x_n$ *converges absolutely* if $\sum \|x_n\| < \infty$.

Proposition 5.1.7. *The following are equivalent for a normed space $(X, \|\cdot\|)$.*

(1) X is Banach, and
(2) Every absolutely convergent sequence converges.

Proof.
(1) $\Rightarrow$ (2): Suppose X is Banach and $\sum \|x_n\| < \infty$. Let $\varepsilon > 0$, and pick $N > 0$ such that $\sum_{n>N} \|x_n\| < \varepsilon$. Then for all $m \geq n > N$,

$$\left\|\sum^m x_i - \sum^n x_i\right\| = \left\|\sum_{n+1}^m x_i\right\| \leq \sum_{n+1}^m \|x_i\| \leq \sum_{n>N} \|x_i\| < \varepsilon.$$

(2) $\Rightarrow$ (1): Suppose (x_n) is Cauchy, and choose $n_1 < n_2 < \cdots$ such that $\|x_m - x_n\| < 2^{-k}$ whenever $m, n > n_k$. Define $y_0 := 0$ (think of this as x_{n_0} by convention), and inductively define $y_k := x_{n_k} - x_{n_{k-1}}$ for all $k \in \mathbb{N}$. Then

$$\sum \|y_k\| \leq \|x_{n_1}\| + \sum_{k\geq 1} 2^{-k} = \|x_{n_1}\| + 1 < \infty.$$

Hence $x := \lim x_{n_k} = \sum y_k$ exists in X. Since (x_n) is Cauchy, $x_n \to x$. □

Proposition 5.1.8. *Suppose X, Y are normed spaces and $T : X \to Y$ is linear. The following are equivalent:*

(1) T is uniformly continuous (with respect to the norm topologies),
(2) T is continuous,
(3) T is continuous at 0_X, and
(4) T is bounded, *i.e., there exists a $c > 0$ such that $\|Tx\| \leq c\|x\|$ for all $x \in X$.*

Proof.
(1) ⇒ (2) ⇒ (3): Trivial.
(3) ⇒ (4): Suppose T is continuous at 0_X. Then there is a neighborhood U of 0_X such that $TU \subset \{y \in Y | \|y\| \leq 1\}$. Since U is open, there is a $\delta > 0$ such that $\{x \in X | \|x\| \leq \delta\} \subset U$. Thus $\|x\| \leq \delta$ implies $\|Tx\| \leq 1$. Then for all $x \neq 0$

$$\left\| \delta \cdot \frac{x}{\|x\|} \right\| \leq \delta \qquad \Longrightarrow \qquad \left\| \delta \cdot \frac{Tx}{\|x\|} \right\| \leq 1 \qquad \Longrightarrow \qquad \|Tx\| \leq \delta^{-1}\|x\|.$$

(4) ⇒ (1): Let $\varepsilon > 0$. If $\|x_1 - x_2\| < c^{-1}\varepsilon$, then

$$\|Tx_1 - Tx_2\| = \|T(x_1 - x_2)\| \leq c\|x_1 - x_2\| < \varepsilon.$$ □

Exercise 5.1.9. Suppose X is a normed space and $Y \subset X$ is a subspace. Define $Q : X \to X/Y$ by $Qx = x + Y$. Define

$$\|Qx\|_{X/Y} = \inf \{\|x - y\|_X | y \in Y\}.$$

(1) Prove that $\|\cdot\|_{X/Y}$ is a well-defined seminorm.
(2) Show that if Y is closed, then $\|\cdot\|_{X/Y}$ is a norm.
(3) Show that in the case of (2) above, $Q : X \to X/Y$ is continuous and open.
Optional: is Q continuous or open only in the case of (1)?
(4) Show that if X is Banach, so is X/Y.

Exercise 5.1.10.

(1) Show that for any two finite dimensional normed spaces F_1 and F_2, all linear maps $T : F_1 \to F_2$ are continuous.
Optional: Show that for any two finite dimensional vector spaces F_1 and F_2 endowed with their vector space topologies from Exercise 5.1.2, all linear maps $T : F_1 \to F_2$ are continuous.
(2) Let X, F be normed spaces with F finite dimensional, and let $T : X \to F$ be a linear map. Prove that the following are equivalent:
 (a) T is bounded (there is an $c > 0$ such that $T(B_1(0_X)) \subseteq B_c(0_F)$), and
 (b) $\ker(T)$ is closed.
Hint: One way to do (b) implies (a) uses Exercise 5.1.9 part (3) and part (1) of this problem.

Definition 5.1.11. Suppose X, Y are normed spaces. Let

$$\mathcal{L}(X \to Y) := \{\text{bounded linear } T : X \to Y\}.$$

Define the *operator norm* on $\mathcal{L}(X \to Y)$ by

$$\begin{aligned}\|T\| &:= \sup\{\|Tx\| | \|x\| \leq 1\}\\ &= \sup\{\|Tx\| | \|x\| = 1\}\\ &= \sup\left\{\frac{\|Tx\|}{\|x\|}\middle| \|x\| \neq 0\right\}\\ &= \inf\{c > 0 | \|Tx\| \leq c\|x\| \text{ for all } x \in X\},\end{aligned}$$

Observe that if $S \in \mathcal{L}(Y \to Z)$ and $T \in \mathcal{L}(X \to Y)$, then $ST \in \mathcal{L}(X \to Z)$ and

$$\|STx\| \leq \|S\| \cdot \|Tx\| \leq \|S\| \cdot \|T\| \cdot \|x\| \qquad \forall\, x \in X.$$

So $\|ST\| \leq \|S\| \cdot \|T\|$.

Proposition 5.1.12. *If Y is Banach, then so is $\mathcal{L}(X \to Y)$.*

Proof. If (T_n) is Cauchy, then so is $(T_n x)$ for all $x \in X$. Set $Tx := \lim T_n x$ for $x \in X$. One verifies that T is linear, T is bounded, and $T_n \to T$. □

Corollary 5.1.13. *If X is complete, then $\mathcal{L}(X) := \mathcal{L}(X \to X)$ is a* Banach algebra *(an algebra with a complete submultiplicative norm).*

Exercise 5.1.14 (Folland §5.1, #7)**.** Suppose X is a Banach space and $T \in \mathcal{L}(X)$. Let $I \in \mathcal{L}(X)$ be the identity map.

(1) Show that if $\|I - T\| < 1$, then T is invertible.
Hint: Show that $\sum_{n\geq 0}(I - T)^n$ converges in $\mathcal{L}(X)$ to T^{-1}.
(2) Show that if $T \in \mathcal{L}(X)$ is invertible and $\|S - T\| < \|T^{-1}\|^{-1}$, then S is invertible.
(3) Deduce that the set of invertible operators $GL(X) \subset \mathcal{L}(X)$ is open.

Exercise 5.1.15. Consider the measure space $(M_n(\mathbb{C}) \cong \mathbb{C}^{n^2}, \lambda^{n^2})$. Show that $GL_n(\mathbb{C})^c \subset M_n(\mathbb{C})$ is λ^{n^2}-null.

Exercise 5.1.16 (Folland §5.2, #19)**.** Let X be an infinite dimensional normed space.

(1) Construct a sequence (x_n) such that $\|x_n\| = 1$ for all n and $\|x_m - x_n\| \geq 1/2$ for all $m \neq n$.
(2) Deduce X is not locally compact.

5.2. **Dual spaces.**

Definition 5.2.1. Let X be a (normed) vector space. A linear map $X \to \mathbb{F}$ is called a (linear) functional. The *dual space* of X is $X^* := \mathrm{Hom}(X \to \mathbb{F})$. Here, Hom means:

- linear maps if X is a vector space, and
- bounded linear maps if X is a normed space.

Exercise 5.2.2. Suppose $\varphi, \varphi_1, \ldots, \varphi_n$ are linear functionals on a vector space X. Prove that the following are equivalent.

(1) $\varphi = \sum_{k=1}^n \alpha_k \varphi_k$ for some $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$.
(2) There is an $\alpha > 0$ such that for all $x \in X$, $|\varphi(x)| \leq \alpha \max_{k=1,\ldots,n} |\varphi_k(x)|$.
(3) $\bigcap_{k=1}^n \ker(\varphi_k) \subset \ker(\varphi)$.

Exercise 5.2.3. Let X be a locally compact Hausdorff space and suppose $\varphi : C_0(X) \to \mathbb{C}$ is a linear functional such that $\varphi(f) \geq 0$ whenever $f \geq 0$. Prove that φ is bounded.
Hint: Argue by contradiction that $\{\varphi(f) | 0 \leq f \leq 1\}$ is bounded using Proposition 5.1.7.

Proposition 5.2.4. *Suppose X is a complex vector space.*

(1) *If $\varphi : X \to \mathbb{C}$ is $\mathbb{C}$-linear, then $\operatorname{Re}(\varphi) : X \to \mathbb{R}$ is $\mathbb{R}$-linear, and for all $x \in X$,*
$$\varphi(x) = \operatorname{Re}(\varphi)(x) - i \operatorname{Re}(\varphi)(ix).$$
(2) *If $f : X \to \mathbb{R}$ is $\mathbb{R}$-linear, then*
$$\varphi(x) := f(x) - if(ix)$$
defines a $\mathbb{C}$-linear functional.

(3) *Suppose X is normed and $\varphi : X \to \mathbb{C}$ is $\mathbb{C}$-linear.*
- *In Case (1), $\|\varphi\| < \infty$ implies $\|\operatorname{Re}(\varphi)\| \leq \|\varphi\|$*
- *In Case (2), $\|\operatorname{Re}(\varphi)\| < \infty$ implies $\|\varphi\| \leq \|\operatorname{Re}(\varphi)\|$.*

Thus $\|\varphi\| = \|\operatorname{Re}(\varphi)\|$.

Proof.
(1) Just observe $\operatorname{Im}(\varphi(x)) = -\operatorname{Re}(i\varphi(x)) = -\operatorname{Re}(\varphi)(ix)$.
(2) It is clear that φ is $\mathbb{R}$-linear. We now check
$$\varphi(ix) = f(ix) - if(i^2x) = f(ix) - if(-x) = if(x) + f(ix) = i(f(x) - if(ix)) = i\varphi(x).$$
(3, Case 1) Since $|\operatorname{Re}(\varphi)(x)| \leq |\varphi(x)|$ for all $x \in X$, $\|\operatorname{Re}(\varphi)\| \leq \|\varphi\|$.
(3, Case 2) If $\varphi(x) \neq 0$, then
$$|\varphi(x)| = \overline{\operatorname{sgn}(\varphi(x))}\varphi(x) = \varphi(\overline{\operatorname{sgn}(\varphi(x))} \cdot x) = \operatorname{Re}(\varphi)(\overline{\operatorname{sgn}(\varphi(x))} \cdot x).$$
Hence $|\varphi(x)| \leq \|\operatorname{Re}(\varphi)\| \cdot \|x\|$, which implies $\|\varphi\| \leq \|\operatorname{Re}(\varphi)\|$. □

Exercise 5.2.5. Consider the following sequence spaces.
$$\ell^1 := \left\{(x_n) \subset \mathbb{C}^\infty \middle| \sum |x_n| < \infty\right\} \qquad \|x\|_1 := \sum |x_n|$$
$$c_0 := \{(x_n) \subset \mathbb{C}^\infty | x_n \to 0 \text{ as } n \to \infty\} \qquad \|x\|_\infty := \sup |x_n|$$
$$c := \left\{(x_n) \subset \mathbb{C}^\infty \middle| \lim_{n\to\infty} x_n \text{ exists}\right\} \qquad \|x\|_\infty := \sup |x_n|$$
$$\ell^\infty := \{(x_n) \subset \mathbb{C}^\infty | \sup |x_n| < \infty\} \qquad \|x\|_\infty := \sup |x_n|$$
(1) Show that every space above is a Banach space.
Hint: First show ℓ^1 and ℓ^∞ are Banach. Then show c_0, c are closed in ℓ^∞.
(2) Construct isometric isomorphisms $c_0^* \cong \ell^1 \cong c^*$ and $(\ell^1)^* \cong \ell^\infty$.
(3) Which of the above spaces are separable?

Warning 5.2.6. If X is a normed space, constructing a non-zero bounded linear functional takes a considerable amount of work. One cannot get by simply choosing a basis for X as an ordinary linear space and mapping the basis to arbitrarily chosen elements of $\mathbb{F}$.

Definition 5.2.7. Suppose X is an $\mathbb{R}$-vector space. A *sublinear (Minkowski) functional* on X is a function $p : X \to \mathbb{R}$ such that
- (positive homogeneous) for all $x \in X$ and $r \geq 0$, $p(rx) = rp(x)$, and
- (subadditive) for all $x, y \in X$, $p(x + y) \leq p(x) + p(y)$.

Theorem 5.2.8 (Real Hahn-Banach)**.** *Let X be an $\mathbb{R}$-vector space, $p : X \to \mathbb{R}$ a sublinear functional, $Y \subset X$ a subspace, and $f : Y \to \mathbb{R}$ a linear functional such that $f(y) \le p(y)$ for all $y \in Y$. Then there is an $\mathbb{R}$-linear functional $g : X \to \mathbb{R}$ such that $g|_Y = f$ and $g(x) \le p(x)$ for all $x \in X$.*

Proof.
Step 1: For all $x \in X \setminus Y$, there is a linear $g : Y \oplus \mathbb{R}x \to \mathbb{R}$ such that $g|_Y = f$ and $g(z) \le p(z)$ on $Y \oplus \mathbb{R}x$.

Proof. Any extension g of f to $Y \oplus \mathbb{R}x$ is determined by $g(y + rx) = f(y) + r\alpha$ for all $r \in \mathbb{R}$, where $\alpha = g(x)$. We want to choose $\alpha \in \mathbb{R}$ such that

$$f(y) + r\alpha \le p(y + rx) \qquad \forall\, y \in Y \text{ and } \forall\, r \in \mathbb{R}. \tag{5.2.9}$$

Since f is $\mathbb{R}$-linear and p is positive homogeneous, we need only consider the cases $r = \pm 1$. Restricting to these 2 cases, (5.2.9) becomes:

$$f(y) - p(y - x) \le \alpha \le p(z + x) - f(z) \qquad \forall\, y, z \in Y.$$

Now observe that

$$p(z+x)-f(z)-f(y)+p(y-x) = p(z+x)+p(y-x)-f(y+z) \ge p(y+z)-f(y+z) \ge 0.$$

Hence there exists an α which lies in the interval

$$[\sup\{f(y) - p(y - x) | y \in Y\}, \inf\{p(z + x) - f(z) | z \in Y\}].$$ □

Step 2: Observe that Step 1 applies to any extension g of f to $Y \subset Z \subset X$ such that $g|_Y = f$ and $g \le p$ on Z. Thus any maximal extension g of f satisfying $g|_Y = f$ and $g \le p$ on its domain must have domain X. Note that

$$\left\{ (Z, g) \middle| \begin{matrix} Y \subseteq Z \subseteq X \text{ is a subspace and } g : Z \to \mathbb{R} \\ \text{such that } g|_Y = f \text{ and } g \le p \text{ on } Z \end{matrix} \right\}$$

is partially ordered by $(Z_1, g_1) \le (Z_2, g_2)$ if $Z_1 \subseteq Z_2$ and $g_2|_{Z_1} = g_1$. Since every ascending chain has an upper bound, there is a maximal extension by Zorn's Lemma. □

Remark 5.2.10. Suppose p is a seminorm on X and $f : X \to \mathbb{R}$ is $\mathbb{R}$-linear. Then $f \le p$ if and only if $|f| \le p$. Indeed,

$$|f(x)| = \pm f(x) = f(\pm x) \le p(\pm x) = p(x).$$

Theorem 5.2.11 (Complex Hahn-Banach)**.** *Let X be an $\mathbb{C}$-vector space, $p : X \to [0, \infty)$ a seminorm, $Y \subset X$ a subspace, and $\varphi : Y \to \mathbb{R}$ a linear functional such that $|\varphi(y)| \le p(y)$ for all $y \in Y$. Then there is a $\mathbb{C}$-linear functional $\psi : X \to \mathbb{C}$ such that $\psi|_Y = \varphi$ and $|\psi(x)| \le p(x)$ for all $x \in X$.*

Proof. By the Real Hahn-Banach Theorem 5.2.8 applied to $\mathrm{Re}(\varphi)$ which is bounded above by p, there is an $\mathbb{R}$-linear extension $g : X \to \mathbb{R}$ such that $g|_Y = \mathrm{Re}(\varphi)$ and $|g| \le p$. Define $\psi(x) := g(x) - ig(ix)$. By Proposition 5.2.4, $\psi|_Y = \varphi$. Finally, for all $x \in X$,

$$|\psi(x)| = \overline{\operatorname{sgn} \psi(x)} \cdot \psi(x) = \psi(\overline{\operatorname{sgn} \psi(x)} \cdot x) = g(\overline{\operatorname{sgn} \psi(x)} \cdot x) \le p(\overline{\operatorname{sgn} \psi(x)} \cdot x) = p(x).$$ □

Facts 5.2.12. Here are some corollaries of the Hahn-Banach Theorems 5.2.8 and 5.2.11. Let X be an $\mathbb{F}$-linear normed space.

(HB1) If $x \neq 0$, there is a $\varphi \in X^*$ such that $\varphi(x) = \|x\|$ and $\|\varphi\| = 1$.

Proof. Define $f : \mathbb{F}x \to \mathbb{F}$ by $f(\lambda x) := \lambda\|x\|$, and observe that $|f| \leq \|\cdot\|$. Now apply Hahn-Banach. □

(HB2) If $Y \subset X$ is closed and $x \notin Y$, there is a $\varphi \in X^*$ such that $\|\varphi\| = 1$ and

$$\varphi(x) = \|x + Y\|_{X/Y} := \inf_{y \in Y} \|x - y\|.$$

Proof. Apply (HB1) to $x + Y \in X/Y$ to get $f \in (X/Y)^*$ such that $\|f\| = 1$ and

$$f(x + Y) = \|x + Y\| = \inf_{y \in Y} \|x - y\|.$$

By Exercise 5.1.9, the canonical quotient map $Q : X \to X/Y$ is continuous. Since

$$\|x + Y\| = \inf_{y \in Y} \|x - y\| \leq \|x\| \qquad \forall\, x \in X,$$

we have $\|Q\| \leq 1$. Thus $\varphi := f \circ Q$ works. □

(HB3) X^* separates points of X.

Proof. If $x \neq y$, then by (HB1), there is a $\varphi \in X^*$ such that $\varphi(x - y) = \|x - y\| \neq 0$. □

(HB4) For $x \in X$, define $\operatorname{ev}_x : X^* \to \mathbb{F}$ by $\operatorname{ev}_x(\varphi) := \varphi(x)$. Then $\operatorname{ev} : X \to X^{**}$ is a linear isometry.

Proof. It is easy to see that ev is linear. For all $\varphi \in X^*$,

$$\|\operatorname{ev}_x(\varphi)\| = |\varphi(x)| \leq \|\varphi\| \cdot \|x\| \qquad \Longrightarrow \qquad \|\operatorname{ev}_x\| \leq \|x\|.$$

Thus $\operatorname{ev}_x \in X^{**}$. If $x \neq 0$, by (HB1) there is a $\varphi \in X^*$ such that $\varphi(x) = \|x\|$ and $\|\varphi\| = 1$. Thus $\|\operatorname{ev}_x\| = \|x\|$. □

Exercise 5.2.13 (Banach Limits)**.** Let $\ell^\infty(\mathbb{N}, \mathbb{R})$ denote the Banach space of bounded functions $\mathbb{N} \to \mathbb{R}$. Show that there is a $\varphi \in \ell^\infty(\mathbb{N}, \mathbb{R})^*$ satisfying the following two conditions:

(1) Letting $S : \ell^\infty(\mathbb{N}, \mathbb{R}) \to \ell^\infty(\mathbb{N}, \mathbb{R})$ be the shift operator $(Sx)_n = x_{n+1}$ for $x = (x_n)_{n \in \mathbb{N}}$, $\varphi = \varphi \circ S$.

(2) For all $x \in \ell^\infty$, $\liminf x_n \leq \varphi(x) \leq \limsup x_n$.

Hint: One could proceed as follows.

(1) Consider the subspace $Y = \operatorname{im}(S - I) = \{Sx - x | x \in \ell^\infty\}$. *Prove that for all* $y \in Y$ *and* $r \in \mathbb{R}$, $\|y + r \cdot \mathbf{1}\| \geq |r|$, *where* $\mathbf{1} = (1)_{n \in \mathbb{N}} \in \ell^\infty$.

(2) Show that the linear map $f : Y \oplus \mathbb{R}\mathbf{1} \to \mathbb{R}$ *given by* $f(y + r \cdot \mathbf{1}) := r$ *is well-defined, and* $|f(z)| \leq \|z\|$ *for all* $z \in Y \oplus \mathbb{R}\mathbf{1}$.

(3) Use the Real Hahn-Banach Theorem 5.2.8 to extend f *to a* $g \in \ell^\infty(\mathbb{N}, \mathbb{R})^*$ *which satisfies (1) and (2).*

Definition 5.2.14. For a normed space X, its *completion* is $\overline{X} := \overline{\text{ev}(X)} \subset X^{**}$, which is always Banach. Observe that if X is Banach, then $\text{ev}(X) \subset X^{**}$ is closed. In this case, if $\text{ev}(X) = X^{**}$, we call X *reflexive.*

Exercise 5.2.15. Show that X is reflexive if and only if X^* is reflexive.
Hint: Instead of the converse, try proving the inverse, i.e., if X is not reflexive, then X^* is not reflexive.

Exercise 5.2.16.

(1) (Folland §5.2, #25) Prove that if X is a Banach space such that X^* is separable, then X is separable.
(2) Find a separable Banach space X such that X^* is not separable.

5.3. The Baire Category Theorem and its consequences.

Theorem 5.3.1 (Baire Category)**.** *Suppose X is either:*

(1) a complete metric space, or
(2) an LCH space.

Suppose (U_n) is a sequence of open dense subsets of X. Then $\bigcap U_n$ is dense in X.

Proof. Let $V_0 \subset X$ be non-empty and open. We will inductively construct for $n \in \mathbb{N}$ a non-empty open set $V_n \subset \overline{V_n} \subset U_n \cap V_{n-1}$.
Case 1: Take V_n to be a ball of radius $< 1/n$.
Case 2: Take V_n such that $\overline{V_n}$ is compact, so $(\overline{V_n})$ are non-empty nested compact sets.

> **Claim.** *$K := \bigcap V_n$ is not empty.*
>
> *Proof of Claim.*
> Case 1: Let x_n be the center of V_n for all n. Then (x_n) is Cauchy, so it converges. The limit lies in K by construction.
> Case 2: Observe $(\overline{V_n})$ is a family of closed sets with the finite intersection property. Since $\overline{V_1}$ is compact, we have $K \neq \emptyset$. □

Now observe $\emptyset \neq K \subset (\bigcap U_n) \cap V_0$. Thus $\bigcap U_n$ is dense in X. □

Corollary 5.3.2. *If X is as in the Baire Category Theorem 5.3.1, then X is not* meager*, i.e., a countable union of nowhere dense sets.*

Proof. If (Y_n) is a sequence of nowhere dense sets, then $(U_n := \overline{Y_n}^c)$ is a sequence of open dense sets. Then
$$\bigcap U_n = \bigcap \overline{Y_n}^c = \left(\bigcup \overline{Y_n}\right)^c \subseteq \left(\bigcup Y_n\right)^c$$
is dense in X, so $\bigcup Y_n \neq X$. □

Lemma 5.3.3. *Suppose X, Y are Banach spaces and $T \in \mathcal{L}(X \to Y)$. Let $U \subset X$ be an open ball centered at 0_X and $V \subset Y$ be an open ball centered at 0_Y. If $V \subset \overline{TU}$, then $V \subset TU$.*

Proof. Let $y \in V$. Take $r \in (0,1)$ such that $y \in rV$. Let $\varepsilon \in (0,1)$ to be decided later. Observe that
$$y \in \overline{rV} \subset \overline{rTU} = \overline{TrU},$$

so there is an $x_0 \in rU$ such that

$$y - Tx_0 \in \varepsilon rV \subset \overline{\varepsilon rTU} = \overline{T(\varepsilon rU)}.$$

Then there is an $x_1 \in \varepsilon rU$ such that

$$y - Tx_0 - Tx_1 \in \varepsilon^2 rV \subset \overline{T(\varepsilon^2 rU)}.$$

Hence by induction, we can construct a sequence (x_n) such that

$$x_n \in \varepsilon^n rU \qquad \text{and} \qquad y - \sum_{j=0}^{n} Tx_j \in \varepsilon^{n+1} rV.$$

Observe that $\sum x_j$ converges as $\|x_j\| < \varepsilon^j rR$ (which is summable!), where $R := \operatorname{radius}(U)$. Moreover,

$$T\sum x_j = \lim_{n\to\infty} T\sum^{n} x_j = \lim_{n\to\infty} \sum^{n} Tx_j = y.$$

Finally, we have

$$\left\|\sum x_j\right\| \leq \sum \|x_j\| < \sum_{j=0}^{\infty} \varepsilon^j rR = \frac{rR}{1-\varepsilon},$$

so $\sum x_j \in \frac{r}{1-\varepsilon}U$. Thus if $\varepsilon < 1 - r$, then $\sum x_n \in U$, so $y \in TU$. □

Theorem 5.3.4 (Open Mapping)**.** *Suppose X, Y are Banach spaces and $T \in \mathcal{L}(X \to Y)$ is surjective. Then T is an open map.*

Proof. It suffices to prove T maps an open neighborhood of 0_X to an open neighborhood of 0_Y. Note $Y = \bigcup_n \overline{TB_n(0_X)}$. By the Baire Category Theorem 5.3.1, there is an $n \in \mathbb{N}$ such that $\overline{TB_n(0)}$ contains a non-empty open set, say $Tx_0 + V$ where $x_0 \in TB_n(0_X)$ and V is an open ball in Y with center 0_Y. Then $V \subset \overline{TB_n(0)} - Tx_0 \subset \overline{TB_{2n}(0_X)}$. By Lemma 5.3.3, $V \subset TB_{2n}(0_X)$. □

Facts 5.3.5. Here are some corollaries of the Open Mapping Theorem 5.3.4.

(OMT1) Suppose X, Y are Banach spaces and $T \in \mathcal{L}(X \to Y)$ is bijective. Then $T^{-1} \in \mathcal{L}(Y \to X)$, and we call T an *isomorphism*.

> *Proof.* When T is bijective, T^{-1} is continuous if and only if T is open. □

(OMT2) Suppose X is Banach under $\|\cdot\|_1$ and $\|\cdot\|_2$. If there is a $c \geq 0$ such that $\|x\|_1 \leq c\|x\|_2$ for all $x \in X$, then $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent.

> *Proof.* Apply (OMT1) to the identity map $\operatorname{id} : (X, \|\cdot\|_2) \to (X, \|\cdot\|_1)$. □

Definition 5.3.6. Suppose X, Y are normed spaces and $T : X \to Y$ is linear. The *graph* of T is the subspace

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

Here, we endow $X \times Y$ with the norm

$$\|(x, y)\|_\infty := \max\{\|x\|_X, \|y\|_Y\}.$$

We say T is *closed* if $\Gamma(T) \subset X \times Y$ is a closed subspace.

Remark 5.3.7. If $T \in \mathcal{L}(X \to Y)$, then $\Gamma(T)$ is closed. Indeed, $(x_n, Tx_n) \to (x, y)$ if and only if $x_n \to x$ and $Tx_n \to y$. Since T is continuous, $Tx_n \to Tx$. Since Y is Hausdorff, $Tx = y$.

Theorem 5.3.8 (Closed Graph)**.** *Suppose X, Y are Banach. If $T : X \to Y$ is a closed linear map, then $T \in \mathcal{L}(X \to Y)$, i.e., T is bounded.*

Proof. Since X, Y are Banach, so is $X \times Y$. Consider the canonical projection maps $\pi_X : X \times Y \to X$ and $\pi_Y : X \times Y \to Y$, which are continuous. Since $\pi_X|_{\Gamma(T)} : \Gamma(T) \to X$ by $(x, Tx) \mapsto x$ is norm decreasing and bijective, $\pi_X|_{\Gamma(T)}^{-1}$ is bounded by (OMT1). Now observe

$$x \xmapsto{\pi_X|_{\Gamma(T)}^{-1}} (x, Tx) \xmapsto{\pi_Y|_{\Gamma(T)}} Tx \qquad \Longrightarrow \qquad T = \pi_Y|_{\Gamma(T)} \circ \pi_X|_{\Gamma(T)}^{-1}$$

which is bounded as the composite of two bounded linear maps. □

Exercise 5.3.9. Suppose X, Y are Banach spaces and $S : X \to Y$ and $T : Y^* \to X^*$ are linear maps such that

$$\varphi(Sx) = (T\varphi)(x) \qquad \forall\, x \in X, \quad \forall\, \varphi \in Y^*.$$

Prove that S, T are bounded.

Definition 5.3.10. A subset S of a topological space $(X, \mathcal{T})$ is called:

- *meager* if S is a countable union of nowhere dense sets, and
- *residual* if S^c is meager.

Exercise 5.3.11. Construct a (non-closed) infinite dimensional meager subspace of ℓ^∞.

Theorem 5.3.12 (Banach-Steinhaus/Uniform Boundedness Principle)**.** *Suppose X, Y are normed spaces and $\mathcal{S} \subset \mathcal{L}(X \to Y)$.*

(1) If $\sup_{T\in\mathcal{S}} \|Tx\| < \infty$ for all x in a non-meager subset of X, then $\sup_{T\in\mathcal{S}} \|T\| < \infty$.
(2) If X is Banach and $\sup_{T\in\mathcal{S}} \|Tx\| < \infty$ for all $x \in X$, then $\sup_{T\in\mathcal{S}} \|T\| < \infty$.

Proof.
(1) Define

$$\begin{aligned} E_n :&= \left\{ x \in X \middle| \sup_{T\in\mathcal{S}} \|Tx\| \le n \right\} = \bigcap_{T\in\mathcal{S}} \{x \in X | \|Tx\| \le n\} \\ &= \bigcap_{T\in\mathcal{S}} (\underbrace{\|\cdot\| \circ T}_{\text{cts}})^{-1}([0, n]), \end{aligned} \tag{5.3.13}$$

which is closed in X. Since $\bigcup E_n$ is a non-meager subset of X, some E_n is non-meager. Thus there is an $x_0 \in X$, $r > 0$, and $n > 0$ such that $\overline{B_r(x_0)} \subset E_n$. Then $\overline{B_r(0)} \subset E_{2n}$:

$$\|Tx\| \le \|T(\underbrace{x - x_0}_{\in \overline{B_r(x_0)} \subset E_n})\| + \|Tx_0\| \le 2n \qquad \text{when } \|x\| \le r.$$

Thus for all $T \in \mathcal{S}$ and $\|x\| \le r$, we have $\|Tx\| \le 2n$. This implies

$$\sup_{T\in\mathcal{S}} \|T\| \le \frac{2n}{r}.$$

(2) Define E_n as in (5.3.13) above. Since $X = \bigcup E_n$ is Banach, the sets cannot all be meager by Corollary 5.3.2 to the Baire Category Theorem 5.3.1. The result now follows from (1). □

Exercise 5.3.14. Provide examples of the following:

(1) Normed spaces X, Y and a discontinuous linear map $T : X \to Y$ with closed graph.
(2) Normed spaces X, Y and a family of linear operators $\{T_\lambda\}_{\lambda\in\Lambda}$ such that $(T_\lambda x)_{\lambda\in\Lambda}$ is bounded for every $x \in X$, but $(\|T_\lambda\|)_{\lambda\in\Lambda}$ is not bounded.

Exercise 5.3.15. Suppose X and Y are Banach spaces and $T : X \to Y$ is a continuous linear map. Show that the following are equivalent.

(1) There exists a constant $c > 0$ such that $\|Tx\|_Y \geq c\|x\|_X$ for all $x \in X$.
(2) T is injective and has closed range.

Exercise 5.3.16 (Folland §5.3, #42)**.** Let $E_n \subset C([0, 1])$ be the space of all functions f such that there is an $x_0 \in [0, 1]$ such that $|f(x) - f(x_0)| \leq n|x - x_0|$ for all $x \in [0, 1]$.

(1) Prove that E_n is nowhere dense in $C([0, 1])$.
(2) Show that the subset of nowhere differentiable functions is residual in $C([0, 1])$.

Exercise 5.3.17. Suppose X, Y are Banach spaces and $(T_n) \subset \mathcal{L}(X \to Y)$ is a sequence of bounded linear maps such that $(T_n x)$ converges for all $x \in X$.

(1) Show that $Tx := \lim T_n x$ defines a bounded linear map.
(2) Does $T_n \to T$ in norm? Give a proof or a counterexample.
Hint: Think about shift operators on a sequence space.

5.4. **Topological vector spaces.**

Definition 5.4.1. An $\mathbb{F}$-vector space X equipped with a topology $\mathcal{T}$ is called a *topological vector space* if

$$+ : X \times X \longrightarrow X$$
$$\cdot : \mathbb{F} \times X \longrightarrow X$$

are continuous.

A subset $C \subseteq X$ is called *convex* if if

$$x, y \in C \qquad \Longrightarrow \qquad tx + (1 - t)y \in C \quad \forall\, t \in [0, 1].$$

A topological vector space is called *locally convex* if for all $x \in X$ and open neighborhoods $U \subset X$ of x, there is a convex open neighborhood V of x such that $V \subseteq U$.

Facts 5.4.2. Suppose $\mathcal{P}$ is a family of seminorms on the $\mathbb{F}$-vector space X. For $x \in X$, $p \in \mathcal{P}$, and $\varepsilon > 0$, define

$$U_{x,p,\varepsilon} := \{y \in X | p(x - y) < \varepsilon\}.$$

Let $\mathcal{T}$ be the topology generated by the sets $U_{x,p,\varepsilon}$, i.e., arbitrary unions of finite intersections of sets of this form.

(LCnvx1) Suppose $x_1, \ldots, x_n \in X$, $p_1, \ldots, p_n \in \mathcal{P}$, and $\varepsilon_1, \ldots, \varepsilon_n > 0$ and $x \in \bigcap_{i=1}^n U_{x_i,p_i,\varepsilon_i}$. Then there is a $\varepsilon > 0$ such that

$$\bigcap_{i=1}^{n} U_{x,p_i,\varepsilon} = \{y \in X | p_i(x - y) < \varepsilon \ \forall\, p_1, \ldots, p_n \in \mathcal{P}\} \subset \bigcap_{i=1}^{n} U_{x_i,p_i,\varepsilon_i}.$$

Hence sets of the form $\bigcap_{i=1}^n U_{x,p_i,\varepsilon} = \{y \in X | p_i(x-y) < \varepsilon \ \forall p_1, \dots, p_n \in \mathcal{P}\}$ form a neighborhood base for $\mathcal{T}$ at x.

Proof. Define $\varepsilon := \min\{\varepsilon_i - p_i(x - x_i) | i = 1, \dots, n\}$. Then for all $y \in \bigcap_{i=1}^n U_{x,p_i,\varepsilon}$ and $j = 1, \dots, n$,

$$p_j(x_j - y) \le p_j(x_j - x) + p_j(x - y) \le (\varepsilon_j - \varepsilon) + \varepsilon = \varepsilon_j.$$

Thus $y \in \bigcap_{i=1}^n U_{x_i,p_i,\varepsilon_i}$, and thus $\bigcap_{i=1}^n U_{x,p_i,\varepsilon} \subseteq \bigcap_{i=1}^n U_{x_i,p_i,\varepsilon_i}$. □

(LCnvx2) If $(x_i) \subset X$ is a net, $x_i \to x$ if and only if $p(x - x_i) \to 0$ for all $p \in \mathcal{P}$.

Proof. By (LCnvx1) $x_i \to x$ if and only if (x_i) is eventually in $U_{x,p,\varepsilon}$ for all $\varepsilon > 0$ and $p \in \mathcal{P}$ if and only if $p(x - x_i) \to 0$ for all $p \in \mathcal{P}$. □

(LCnvx3) $\mathcal{T}$ is the weakest topology such that the $p \in \mathcal{P}$ are continuous.

Proof. Exercise. □

(LCnvx4) $(X, \mathcal{T})$ is a topological vector space.

Proof.
<u>+ cts:</u> Suppose $x_i \to x$ and $y_i \to y$. Then for all $p \in \mathcal{P}$,

$$p(x + y - (x_i + y_i)) \le p(x - x_i) + p(y - y_i) \to 0.$$

<u>· cts:</u> Suppose $x_i \to x$ and $\alpha_i \to \alpha$. Then for all $p \in \mathcal{P}$,

$$\begin{aligned} p(\alpha_i x_i - \alpha x) &\le p(\alpha_i x_i - \alpha x_i) + p(\alpha x_i - \alpha x) \\ &\le \underbrace{|\alpha_i - \alpha|}_{\to 0} \cdot \underbrace{p(x_i)}_{\to p(x)} + |\alpha| \cdot \underbrace{p(x_i - x)}_{\to 0}. \end{aligned}$$

□

(LCnvx5) $(X, \mathcal{T})$ is locally convex.

Proof. Observe that each $U_{x,p,\varepsilon}$ is convex. Indeed, if $y, z \in U_{x,p,\varepsilon}$, then for all $t \in [0, 1]$,

$$\begin{aligned} p(x - (ty + (1-t)z)) &= p((tx + (1-t)x) - (ty + (1-t)z)) \\ &= p((t(x - y) + (1-t)(x - z)) \\ &\le t \cdot p(x - y) + (1-t) \cdot p(x - z) \\ &< t\varepsilon + (1-t)\varepsilon \\ &= \varepsilon. \end{aligned}$$

The result now follows from (LCnvx1) as the intersection of convex sets is convex. □

(LCnvx6) $(X, \mathcal{T})$ is Hausdorff if and only if $\mathcal{P}$ separates points if and only if for all $x \in X \setminus \{0\}$, there is a $p \in \mathcal{P}$ such that $p(x) \neq 0$.

Proof. Exercise. ◻

(LCnvx7) If $(X, \mathcal{T})$ is Hausdorff and $\mathcal{P}$ is countable, then there exists a metric $d : X \times X \to [0, \infty)$ which is *translation invariant* ($d(x + z, y + z) = d(x, y)$ for all $x, y, z \in X$) which induces the same topology as $\mathcal{P}$.

Proof. Let $\mathcal{P} = (p_n)$ be an enumeration and set

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

We leave it to the reader to verify that d is a translation invariant metric which induces the topology $\mathcal{T}$. ◻

(LCnvx8) If $(X, \mathcal{T})$ is locally convex Hausdorff TVS, then $\mathcal{T}$ is given by a separating family of seminorms.

Proof. Beyond the scope of this course; take Functional Analysis 7211. ◻

Proposition 5.4.3. *Suppose $(X, \mathcal{P})$ and $(Y, \mathcal{Q})$ are seminormed locally convex topological vector spaces. The following are equivalent for a linear map $T : X \to Y$:*

(1) T is continuous.
(2) T is continuous at 0_X.
(3) For all $q \in \mathcal{Q}$, there are $p_1, \ldots, p_n \in \mathcal{P}$ and $c > 0$ such that $q(Tx) \leq c \sum_{j=1}^{n} p_j(x)$ for all $x \in X$.

Proof.
(1) ⇒ (2): Trivial.
(2) ⇒ (3): Suppose T is continuous at 0_X and $q \in \mathcal{Q}$. Then there are $p_1, \ldots, p_n \in \mathcal{P}$ and $\varepsilon > 0$ such that for all $x \in V := \bigcap_{i=1}^{n} U_{0,p_i,\varepsilon}$, we have $q(Tx) < 1$. Fix $x \in X$. If $p_i(x) = 0$ for all $i = 1, \ldots, n$, then $rx \in V$ for all $r > 0$, so

$$rq(Tx) = q(T\underbrace{rx}_{\in V}) < 1 \qquad \forall r > 0.$$

This implies $q(Tx) = 0 \leq c \sum_{i=1}^{n} p_i(x)$ for all $c > 0$, so we may assume $p_1(x) > 0$. Then

$$y := \left(\frac{\varepsilon}{2 \sum_{i=1}^{n} p_i(x)} \right) \cdot x \in V$$

as $p_i(y) \leq \varepsilon/2 < \varepsilon$ for all $i = 1, \ldots, n$. Thus

$$q(Tx) = \left(\frac{2}{\varepsilon} \sum_{i=1}^{n} p_i(x) \right) q(Ty) < \frac{2}{\varepsilon} \sum_{i=1}^{n} p_i(x)$$

as desired.

(3) ⇒ (1): We must show if $x_i \to x$ in X, then $q(Tx_i - Tx) \to 0$ for all $q \in \mathcal{Q}$. Since $x_i \to x$, $p(x_i - x) \to 0$ for all $p \in \mathcal{P}$. Fix $q \in \mathcal{Q}$. By (3), there are $p_1, \dots, p_n \in \mathcal{P}$ and $c > 0$ such that

$$q(T(x_i - x)) \leq c \sum_{j=1}^{n} p_j(x_i - x) \longrightarrow 0 \qquad \forall\, x \in X. \qquad \square$$

Definition 5.4.4. Let X be a normed space. Recall that X^* separates points of X by the Hahn-Banach Theorem 5.2.8 or 5.2.11. Consider the family of seminorms

$$\mathcal{P} := \{\, x \mapsto |\varphi(x)| \,|\, \varphi \in X^*\}$$

on X, which separates points. Hence $\mathcal{P}$ induces a locally convex Hausdorff vector space topology on X in which $x_i \to x$ if and only if $\varphi(x_i) \to \varphi(x)$ for all $\varphi \in X^*$ by (LCnvx2). We call this topology the *weak topology* on X.

Proposition 5.4.5. *If $U \subset X$ is weakly open then U is $\|\cdot\|$-open.*

Proof. Observe that every basic open set $U_{x,\varphi,\varepsilon} = \{y \in X |\, |\varphi(x-y)| < \varepsilon\}$ is norm open in X. Indeed, $y \mapsto |\varphi(x-y)|$ is norm continuous as $\varphi \in X^*$ is norm continuous, the vector space operations are norm-continuous, and $|\cdot| : \mathbb{C} \to [0,\infty)$ is continuous. $\square$

Exercise 5.4.6. Let X be a normed space. Prove that the weak and norm topologies agree if and only if X is finite dimensional.

Proposition 5.4.7. *A linear functional $\varphi : X \to \mathbb{F}$ is weakly continuous (continuous with respect to the weak topology) if and only if $\varphi \in X^*$ (continuous with respect to the norm topology).*

Proof. Suppose $\varphi \in X^*$. Then $\varphi^{-1}(B_\varepsilon(0_{\mathbb{C}})) = \{x \in X |\, |\varphi(x)| < \varepsilon\} = U_{0,\varepsilon,\varepsilon}$ is weakly open. Hence φ is continuous at 0_X and thus weakly continuous by Proposition 5.4.3.

Now suppose $\varphi : X \to \mathbb{C}$ is weakly continuous. Then for all $U \subset \mathbb{C}$ open, $\varphi^{-1}(U)$ is weakly open and thus norm open by Proposition 5.4.5. Thus φ is $\|\cdot\|$-continuous and thus in X^*. $\square$

Definition 5.4.8. The weak* topology on X^* is the locally convex Hausdroff vector space topology induced by the separating family of seminorms

$$\mathcal{P} = \{\varphi \mapsto |\operatorname{ev}_x(\varphi)| = |\varphi(x)| \,|x \in X\}.$$

Observe that $\varphi_i \to \varphi$ if and only if $\varphi_i(x) \to \varphi(x)$ for all $x \in X$.

Theorem 5.4.9 (Banach-Alaoglu)**.** *The norm-closed unit ball B^* of X^* is weak*-compact.*

Proof.

> **Trick.** For $x \in X$, let $D_x = \{z \in \mathbb{C} |\, |z| \leq \|x\|\}$. By Tychonoff's Theorem, $D := \prod_{x \in X} D_x$ is compact Hausdorff. The elements $(d_x) \in D$ are precisely functions $f : X \to \mathbb{C}$ (not necessarily linear) such that $|f(x)| \leq \|x\|$ for all $x \in X$.

Observe $B^* \subset D$ is the subset of linear functions. The relative product topology on B^* is the relative weak* topology, as both are pointwise convergence. It remains to prove $B^* \subset D$ is closed. If $(\varphi_i) \subset B^*$ is a net with $\varphi_i \to \varphi \in D$, then

$$\varphi(\alpha x + y) = \lim \varphi_i(\alpha x + y) = \lim \alpha\varphi_i(x) + \varphi_i(y) = \alpha\varphi(x) + \varphi(y). \qquad \square$$

Exercise 5.4.10. Let X be a normed space.

(1) Show that every weakly convergent sequence in X is norm bounded.
(2) Suppose in addition that X is Banach. Show that every weak* convergent sequence in X^* is norm bounded.
(3) Give a counterexample to (2) when X is not Banach.
Hint: Under $\|\cdot\|_\infty$, $c_c^* \cong \ell^1$, *where* c_c *is the space of sequences which are eventually zero.*

Exercise 5.4.11 (Goldstine's Theorem)**.** Let X be a normed vector space with closed unit ball B. Let B^{**} be the unit ball in X^{**}, and let $i : X \to X^{**}$ be the canonical inclusion. Recall that the weak* topology on X^{**} is the weak topology induced by X^*. In this exercise, we will prove that $i(B)$ is weak* dense in B^{**}.
Note: You may use a Hahn-Banach separation theorem that we did not discuss in class to prove the result directly if you do not choose to proceed along the following steps.

(1) Show that for every $x^{**} \in B^{**}$, $\varphi_1, \dots, \varphi_n \in X^*$, and $\delta > 0$, there is an $x \in (1+\delta)B$ such that $\varphi_i(x) = x^{**}(\varphi_i)$ for all $1 \leq i \leq n$.
Hint: Here is a walkthrough for this first part. Fix $\varphi_1, \dots, \varphi_n \in X^*$.
  (a) *Find* $x \in X$ *such that* $\varphi_i(x) = x^{**}(\varphi_i)$ *for all* $1 \leq i \leq n$.
  (b) *Set* $Y := \bigcap \ker(\varphi_i)$ *and let* $\delta > 0$. *Show by contradiction that* $(x+Y) \cap (1+\delta)B \neq \emptyset$. *(This part uses the Hahn-Banach Theorem.)*
(2) Suppose U is a basic open neighborhood of $x^{**} \in B^{**}$. Deduce that for every $\delta > 0$, $(1+\delta)i(B) \cap U \neq \emptyset$. That is, there is an $x_\delta \in (1+\delta)B$ such that $i(x_\delta) \in U$.
(3) By part (2), $(1+\delta)^{-1}x_\delta \in B$. Show that for δ sufficiently small (which can be expressed in terms of the basic open neighborhood U), $(1+\delta)^{-1}i(x_\delta) \in i(B) \cap U$.

Exercise 5.4.12. Suppose X is a Banach space. Prove that X is reflexive if and only if the unit ball of X is weakly compact.
Hint: Use the Banach-Alaoglu Theorem 5.4.9 and Exercise 5.4.11.

Exercise 5.4.13. Suppose X, Y are Banach spaces and $T : X \to Y$ is a linear transformation.

(1) Show that if $T \in \mathcal{L}(X,Y)$, then T is weak-weak continuous. That is, if $x_\lambda \to x$ in the weak topology on X induced by X^*, then $Tx_\lambda \to Tx$ in the weak topology on Y induced by Y^*.
(2) Show that if T is norm-weak continuous, then $T \in \mathcal{L}(X,Y)$.
(3) Show that if T is weak-norm continuous, then T has finite rank, i.e., TX is finite dimensional.

Hint: For part (3), one could proceed as follows.

(1) *First, reduce to the case that* T *is injective by replacing* X *with* $Z = X/\ker(T)$ *and* T *with* $S : Z \to Y$ *given by* $x + \ker(T) \mapsto Tx$. *(You must show* S *is weak-norm continuous on* Z.*)*
(2) *Take a basic open set* $\mathcal{U} = \{z \in Z \,|\, |\varphi_i(z)| < \varepsilon$ *for all* $i = 1, \dots, n\} \subset S^{-1}B_1(0_Y)$. *Use that* S *is injective to prove that* $\bigcap_{i=1}^n \ker(\varphi_i) = (0)$.
(3) *Use Exercise 5.2.2 to deduce that* Z^* *is finite dimensional, and thus that* Z *and* $TX = SZ$ *are finite dimensional.*

Exercise 5.4.14. Suppose X is a Banach space. Prove the following are equivalent:

(1) X is separable.
(2) The relative weak* topology on the closed unit ball of X^* is metrizable.

Deduce that if X is separable, the closed unit ball of X^* is weak* sequentially compact.
Hint: For $(1) \Rightarrow (2)$*, you could adapt either the proof of (LCnvx7) or the trick in the proof of the Banach-Alaoglu Theorem 5.4.9 using a countable dense subset. For* $(2) \Rightarrow (1)$*, there a countable neighborhood base* $(U_n) \subset B^*$ *at* 0_X *such that* $\bigcap U_n = \{0\}$*. For each* $n \in \mathbb{N}$*, there is a finite set* $D_n \subset X$ *and an* $\varepsilon_n > 0$ *such that*

$$U_n \supseteq \{\varphi \in X^* | \, |\varphi(x)| < \varepsilon_n \text{ for all } x \in D_n\}.$$

Setting $D = \bigcup D_n$*, show that* $\operatorname{span}(D)$ *is dense in* X*. Deduce that* X *is separable.*

Exercise 5.4.15. Suppose X is a Banach space. Prove the following are equivalent:

(1) X^* is separable.
(2) The relative weak topology on the closed unit ball of X is metrizable.

Exercise 5.4.16. How do you reconcile Exercises 5.4.12, 5.4.14, and 5.4.15? That is, how do you reconcile the fact that there exist separable Banach spaces which are not reflexive?

Exercise 5.4.17.

(1) Prove that the norm closed unit ball of ℓ^∞ is weak* sequentially compact.
(2) Prove that the norm closed unit ball of ℓ^∞ is not weakly sequentially compact.
Hint: One could proceed as follows.
   (a) Prove that the weak topology on* $\ell^\infty \cong (\ell^1)^*$ *is contained in the weak topology, i.e., if* $x_i \to x$ *weakly, then* $x_i \to x$ *weak*.*
   (b) Consider the sequence $(x_n) \subset c \subset \ell^\infty$ *given by*

$$(x_n)(m) = \begin{cases} 0 & \text{if } n < m \\ 1 & \text{if } n \geq m. \end{cases}$$

   Show that $x_n \to 0$ *weak* in* ℓ^∞*.*
   (c) Show that (x_n) *does not converge weakly in* ℓ^∞ *by extending* $\lim : c \to \mathbb{C}$ *to* ℓ^∞*.*
   (d) Deduce no subsequence of (x_n) *converges weakly in* ℓ^∞*.*

Remark 5.4.18. The Eberlein-Šmulian Theorem (which we will not prove here) states that if X is a Banach space and $S \subset X$, the following are equivalent.

(1) S is weakly *precompact*, i.e., the weak closure of S is weakly compact.
(2) Every sequence of S has a weakly convergent subsequence (whose weak limit need not be in S).
(3) Every sequence of S has a weak cluster point.

Exercise 5.4.19. Let X be a compact Hausdorff topological space. For $x \in X$, define $\operatorname{ev}_x : C(X) \to \mathbb{F}$ by $\operatorname{ev}_x(f) = f(x)$.

(1) Prove that $\operatorname{ev}_x \in C(X)^*$, and find $\|\operatorname{ev}_x\|$.
(2) Show that the map $\operatorname{ev} : X \to C(X)^*$ given by $x \mapsto \operatorname{ev}_x$ is a homeomorphism onto its image, where the image has the relative weak* topology.

5.5. **Hilbert spaces.**

Definition 5.5.1. A *sesquilinear form* on an $\mathbb{F}$-vector space H is a function $\langle \cdot , \cdot \rangle : H \times H \to \mathbb{F}$ which is

- linear in the first variable: $\langle \alpha x + y, z\rangle = \alpha\langle x, z\rangle + \langle y, z\rangle$ for all $\alpha \in \mathbb{F}$ and $x, y, z \in H$, and
- conjugate linear in the second variable: $\langle x, \alpha y + z\rangle = \overline{\alpha}\langle x, y\rangle + \langle x, z\rangle$ for all $\alpha \in \mathbb{F}$ and $x, y, z \in H$.

We call $\langle \cdot , \cdot \rangle$:

- *self-adjoint* if $\langle x, y\rangle = \overline{\langle y, x\rangle}$ for all $x, y \in H$,
- *non-degenerate* if $\langle x, y\rangle = 0$ for all $y \in H$ implies $x = 0$
- *positive* if $\langle x, x\rangle \geq 0$ for all $x \in H$. A positive sesquilinear form is called *definite* if moreover $\langle x, x\rangle = 0$ implies $x = 0$.

A self-adjoint positive definite sesquilinear form is called an *inner product*.

Exercise 5.5.2. Suppose $\langle \cdot , \cdot \rangle$ is a self-adjoint sesquilinear form on the $\mathbb{R}$-vector space H. Show that:

- ($\mathbb{R}$-polarization) $4\langle x, y\rangle = \langle x + y, x + y\rangle - \langle x - y, x - y\rangle$ for all $x, y \in H$.

Now suppose $\langle \cdot , \cdot \rangle$ is a sesquilinear form on the $\mathbb{C}$-vector space H. Prove the following.

(1) ($\mathbb{C}$-polarization) $4\langle x, y\rangle = \sum_{k=0}^{3} i^k\langle x + i^k y, x + i^k y\rangle$ for all $x, y \in H$.
(2) $\langle \cdot , \cdot \rangle$ is self-adjoint if and only if $\langle x, x\rangle \in \mathbb{R}$ for all $x \in H$.
(3) Positive implies self-adjoint.

Definition 5.5.3. Suppose that $\langle \cdot , \cdot \rangle$ is positive and self-adjoint (so $(H, \langle \cdot , \cdot \rangle)$ is a *pre-Hilbert space*). Define

$$\|x\| := \langle x, x\rangle^{1/2}.$$

Observe that $\|\cdot\|$ is *homogeneous*: $\|\alpha x\| = |\alpha| \cdot \|x\|$ for all $\alpha \in \mathbb{F}$ and $x \in H$.

We say that x and y are *orthogonal*, denoted $x \perp y$, if $\langle x, y\rangle = 0$.

Facts 5.5.4. We have the following facts about pre-Hilbert spaces:

(H1) (Pythagorean Theorem) $x \perp y$ implies $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.

Proof. $\|x + y\|^2 = \|x\|^2 + 2\operatorname{Re}\langle x, y\rangle + \|y\|^2 = \|x\|^2 + \|y\|^2$. □

(H2) $x \perp y$ if and only if $\|x\|^2 \leq \|x + \alpha y\|^2$ for all $\alpha \in \mathbb{F}$.

Proof.
$\Rightarrow$: $\|x + \alpha y\|^2 \underset{\text{(H1)}}{=} \|x\|^2 + |\alpha|^2\|y\|^2 \geq \|x\|^2$ for all $\alpha \in \mathbb{F}$.

$\Leftarrow$: Suppose

$$\|x\|^2 + 2\operatorname{Re}(\alpha\langle x, y\rangle) + |\alpha|^2\|y\|^2 = \|x + \alpha y\|^2 \geq \|x\|^2 \qquad \forall\, \alpha \in \mathbb{F}.$$

Then for all $\alpha \in \mathbb{F}$,

$$0 \leq 2\operatorname{Re}(\alpha\langle x, y\rangle) + |\alpha|^2\|y\|^2.$$

Taking $\alpha \in \mathbb{F}$ sufficiently close to $0_{\mathbb{F}}$, the term $2\operatorname{Re}(\alpha\langle x, y\rangle)$ dominates, and this can only be non-negative for all $\alpha \in \mathbb{F}$ if $\langle x, y\rangle = 0$. □

(H3) The properties of being definite and non-degenerate are equivalent.

Proof.
$\Rightarrow$: Trivial; just take $y = x$ in the definition of non-degeneracy.
$\Leftarrow$: If $\|x\|^2 = 0$, then for all $\alpha \in \mathbb{F}$ and $y \in H$, $\|x\|^2 = 0 \leq \|x + \alpha y\|^2$ by positivity. Hence $x \perp y$ for all $y \in H$ by (H2). Thus $x = 0$ by non-degeneracy. □

(H4) (Cauchy-Schwarz Inequality) For all $x, y \in H$, $|\langle x, y\rangle| \leq \|x\| \cdot \|y\|$.

Proof. For all $r \in \mathbb{R}$,

$$0 \leq \|x - ry\|^2 = \|x\|^2 - 2r \operatorname{Re}\langle x, y\rangle + r^2\|y\|^2,$$

which is a non-negative quadratic in r. Therefore its discriminant

$$4(\operatorname{Re}\langle x, y\rangle)^2 - 4 \cdot \|x\|^2 \cdot \|y\|^2 \leq 0,$$

which implies $|\operatorname{Re}\langle x, y\rangle| \leq \|x\| \cdot \|y\|$.

Trick. $|\langle x, y\rangle| = \alpha\langle x, y\rangle$ for some $\alpha \in U(1) = \{z \in \mathbb{C} \mid |z| = 1\}$.

Then

$$|\langle x, y\rangle| = \alpha\langle x, y\rangle = \langle \alpha x, y\rangle \leq \|\alpha x\| \cdot \|y\| = \|x\| \cdot \|y\|.$$ □

(H5) (Cauchy-Schwarz Definiteness) If $\langle \cdot\,, \cdot\rangle$ is definite, then $|\langle x, y\rangle| = \|x\| \cdot \|y\|$ implies $\{x, y\}$ is linearly dependent.

Proof. We may assume $y \neq 0$. Set

$$\alpha := \frac{|\langle x, y\rangle|}{\|y\|^2}\overline{\operatorname{sgn}(\langle x, y\rangle)}.$$

Then we calculate

$$\begin{aligned}
\|x - \alpha y\|^2 &= \|x\|^2 - 2\operatorname{Re}(\alpha\langle x, y\rangle) + |\alpha|^2 \cdot \|y\|^2 \\
&= \|x\|^2 - 2\frac{|\langle x, y\rangle|^2}{\|y\|^2} + \frac{|\langle x, y\rangle|^2}{\|y\|^4}\|y\|^2 \\
&= \|x\|^2 - \frac{|\langle x, y\rangle|^2}{\|y\|^2} \\
&= \|x\|^2 - \frac{\|x\|^2 \cdot \|y\|^2}{\|y\|^2} \\
&= 0.
\end{aligned}$$

This implies $x = \alpha y$ by definiteness.
(The essential idea here was to minimize a quadratic in α.) □

(H6) $\|\cdot\| : H \to [0, \infty)$ is a seminorm. It is a norm exactly when $\langle \cdot\,, \cdot\rangle$ is definite, i.e., an inner product.

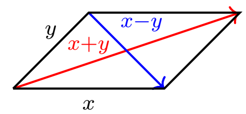
Proof. It remains to prove subadditivity of $\|\cdot\|$, which follows by the Cauchy-Schwarz Inequality (H4):

$$\begin{aligned}\|x+y\|^2 &= \langle x+y, x+y\rangle \\ &= \|x\|^2 + 2\operatorname{Re}\langle x,y\rangle + \|y\|^2 \\ &\le \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2 \\ &\le \|x\|^2 + 2\|x\|\cdot\|y\| + \|y\|^2 && \text{(H4)} \\ &= (\|x\|+\|y\|)^2.\end{aligned}$$

Now take square roots. The final claim follows immediately. □

Proposition 5.5.5. *A norm $\|\cdot\|$ on a $\mathbb{C}$-vector space comes from an inner product if and only if it satisfies the parallelogram identity:*

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



Proof.
⇒: If $\|\cdot\|$ comes from an inner product, then add together

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

⇐: If the parallelogram identity holds, just define

$$\langle x,y\rangle := \frac{1}{4}\sum_{k=0}^{3} i^k \|x + i^k y\|^2$$

by polarization. One checks this works. □

Definition 5.5.6. A *Hilbert space* is an inner product space whose induced norm is complete, i.e., Banach.

Exercise 5.5.7. Verify the follows spaces are Hilbert spaces.

(1) $\ell^2 := \{(x_n) \in \mathbb{C}^\infty | \sum |x_n|^2 < \infty\}$ with $\langle x,y\rangle := \sum x_n\overline{y_n}$.
(2) Suppose $(X, \mathcal{M}, \mu)$ is a measure space. Define

$$\mathcal{L}^2(X,\mu) := \frac{\left\{\text{measurable } f : X \to \mathbb{C} \middle| \int |f|^2\, d\mu < \infty\right\}}{\text{equality a.e.}}$$

with $\langle f,g\rangle := \int f\overline{g}\, d\mu$.

Exercise 5.5.8. Suppose H is a Hilbert space and $S, T : H \to H$ are linear operators such that for all $x, y \in H$, $\langle Sx, y\rangle = \langle x, Ty\rangle$. Prove that S and T are bounded.

From this point forward, H will denote a Hilbert space.

Theorem 5.5.9. *Suppose $C \subset H$ is a non-empty convex closed subset and $z \notin C$. There is a unique $x \in C$ such that*

$$\|x - z\| = \inf_{y\in C} \|y - z\|.$$

Proof. By translation, we may assume $z = 0 \notin C$. Suppose $(x_n) \subset C$ such that $\|x_n\| \to r := \inf_{y\in C} \|y\|$. Then by the parallelogram identity,

$$\left\|\frac{x_m - x_n}{2}\right\|^2 + \left\|\frac{x_m + x_n}{2}\right\|^2 = 2\left(\left\|\frac{x_m}{2}\right\|^2 + \left\|\frac{x_n}{2}\right\|^2\right)$$

Rearranging, we have

$$\|x_m - x_n\|^2 = 2\underbrace{\|x_m\|^2}_{\to r^2} + 2\underbrace{\|x_n\|^2}_{\to r^2} - 4\underbrace{\left\|\frac{x_m + x_n}{2}\right\|^2}_{\geq r^2}$$

where the last inequality follows since $(x_m + x_n)/2 \in C$ by convexity. This means that

$$\limsup_{m,n} \|x_m - x_n\|^2 \leq 2r^2 + 2r^2 - 4r^2 = 0,$$

and thus (x_n) is Cauchy. Since H is complete, there is an $x \in H$ such that $x_n \to x$, and $\|x\| = r$. Since C is closed, $x \in C$.

For uniqueness, observe that if $x' \in C$ satisfies $\|x'\| = r$, then $(x, x', x, x', \dots)$ is Cauchy by the above argument, and thus converges. We conclude that $x = x'$. ☐

Definition 5.5.10. For $S \subset H$, define the *orthogonal complement*

$$S^\perp := \{x \in H | \langle x, s\rangle = 0 \ , \forall\, s \in S\}.$$

Observe that $S^\perp$ is a closed subspace.

Facts 5.5.11. We have the following facts about orthogonal complements.

($\perp$1) If $S \subset T$, then $T^\perp \subset S^\perp$.

> *Proof.* Observe $x \in T^\perp$ if and only if $\langle x, t\rangle = 0$ for all $t \in T \supseteq S$. Hence $x \in S^\perp$. ☐

($\perp$2) $\overline{S} \subset S^{\perp\perp}$ and $S^\perp = S^{\perp\perp\perp}$.

> *Proof.* If $s \in S$, then $\langle s, x\rangle = \overline{\langle x, s\rangle} = 0$ for all $x \in S^\perp$. Thus $s \in S^{\perp\perp}$. Since $S^{\perp\perp}$ is closed, $\overline{S} \subset S^{\perp\perp}$.
> Now replacing S with $S^\perp$, we get $S^\perp \subset S^{\perp\perp\perp}$. But since $S \subseteq S^{\perp\perp}$, by ($\perp$1), we have $S^{\perp\perp\perp} \subseteq S^\perp$. ☐

($\perp$3) $S \cap S^\perp = \{0\}$.

> *Proof.* If $x \in S \cap S^\perp$, then $\langle x, x\rangle = 0$, so $x = 0$. ☐

($\perp$4) If $K \subset H$ is a subspace, then $H = \overline{K} \oplus K^\perp$.

Proof. By (⊥2) and (⊥3),

$$\{0\} \subseteq \overline{K} \cap K^\perp \subseteq K^{\perp\perp} \cap K^\perp = \{0\},$$

so equality holds everywhere.

Let $x \in H$. Since $\overline{K}$ is closed and convex, there is a unique $y \in \overline{K}$ minimizing the distance to x, i.e., $\|x - y\| \leq \inf_{k\in K} \|x - k\|$. We claim that $x - y \in K^\perp$, so that $x = y + (x - y)$, and $H = \overline{K} + K^\perp$. Indeed, for all $k \in K$ and $\alpha \in \mathbb{C}$,

$$\|x - y\|^2 \leq \|x - (y - \alpha k)\|^2 = \|(x - y) + \alpha k\|^2 \qquad \forall\, \alpha \in \mathbb{C}.$$

By (H2), we have $(x - y) \perp k$ for all $k \in K$, i.e., $x - y \in K^\perp$ as claimed. □

(⊥5) If $K \subset H$ is a subspace, then $\overline{K} = K^{\perp\perp}$.

Proof. Let $x \in K^{\perp\perp}$. By (⊥4), there are unique $y \in \overline{K}$ and $z \in K^\perp$ such that $x = y + z$. Then

$$0 = \langle x, z\rangle = \langle y + z, z\rangle = \underbrace{\langle y, z\rangle}_{=0 \text{ by } (\perp 2)} + \langle z, z\rangle.$$

Hence $z = 0$, and $x = y \in \overline{K}$. □

Notation 5.5.12 (Dirac bra-ket)**.** Let $(H, \langle\,\cdot\,,\,\cdot\,\rangle)$ be a Hilbert space, where $\langle\,\cdot\,,\,\cdot\,\rangle$ is linear on the left and conjugate linear on the right. Define $\langle\,\cdot\,|\,\cdot\,\rangle : H \times H \to \mathbb{F}$ by

$$\langle x|y\rangle := \langle y, x\rangle.$$

That is, $\langle\,\cdot\,|\,\cdot\,\rangle$ is the 'same' inner product, but linear on the right and conjugate linear on the left.

We may further denote a vector $x \in H$ by the *ket* $|x\rangle$. For $x \in H$, we denote the linear map $H \to \mathbb{F}$ by $y \mapsto \langle x|y\rangle$ by the *bra* $\langle x|$. Observe that the bra $\langle x|$ applied to the ket $|y\rangle$ gives the bracket $\langle x|y\rangle$.

Theorem 5.5.13 (Riesz Representation)**.** *Let H be a Hilbert space.*

(1) For all $y \in H$, $\langle y| \in H^*$ and $\|\langle y|\| = \|y\|$.
(2) For $\varphi \in H^*$, there is a unique $y \in H$ such that $\varphi = \langle y|$.
(3) The map $y \mapsto \langle y|$ is a conjugate-linear isometric isomorphism.

Proof.
(1) Clearly $\langle y|$ is linear. By Cauchy-Schwarz, $|\langle y|x\rangle| \leq \|x\| \cdot \|y\|$, so $\|\langle y|\| \leq \|y\|$. Taking $x = y$, we have $|\langle y|y\rangle| = \|y\|^2$, so $\|\langle y|\| = \|y\|$.
(2) If $\langle y| = \langle y'|$, then

$$\|y - y'\|^2 = \langle y - y'|y - y'\rangle = \langle y|y - y'\rangle - \langle y'|y - y'\rangle = 0,$$

and thus $y = y'$. Suppose now $\varphi \in H^*$. We may assume $\varphi \neq 0$. Then $\ker(\varphi) \subset H$ is a closed proper subspace. Pick $z \in \ker(\varphi)^\perp$ with $\varphi(z) = 1$. Now for all $x \in H$, $x - \varphi(x)z \in \ker(\varphi)$, so

$$\langle z|x\rangle = \langle z|x - \varphi(x)z + \varphi(x)z\rangle = \langle \underbrace{z}_{\in\ker(\varphi)^\perp} | \underbrace{x - \varphi(x)z}_{\in\ker(\varphi)}\rangle + \langle z|\varphi(x)z\rangle = \langle z|\varphi(x)z\rangle = \varphi(x)\|z\|^2.$$

We conclude that $\varphi = \left\langle \frac{z}{\|z\|^2} \right|$.

(3) $y \mapsto \langle y|$ is isometric by (1) and onto by (2). Conjugate linearity is straightforward. □

Exercise 5.5.14. Suppose H is a Hilbert space. Show that the dual space H^* with

$$\langle \langle x|, \langle y| \rangle_{H^*} := \langle y, x\rangle_H$$

is a Hilbert space whose induced norm is equal to the operator norm on H^*.

Definition 5.5.15. A subset $E \subset H$ is called *orthonormal* if $e, f \in E$ implies $\langle e, f\rangle = \delta_{e=f}$. Observe that $\|e - f\| = \sqrt{2}$ for all $e \neq f$ in E. Thus if H is separable, any orthonormal set is countable.

Exercise 5.5.16. Suppose H is a Hilbert space, $E \subset H$ is an orthonormal set, and $\{e_1, \ldots, e_n\} \subset E$. Prove the following assertions.

1. If $x = \sum_{i=1}^n c_i e_i$, then $c_j = \langle x, e_j\rangle$ for all $j = 1, \ldots, n$.
2. The set E is linearly independent.
3. For every $x \in H$, $\sum_{i=1}^n \langle x, e_i\rangle e_i$ is the unique element of $\operatorname{span}\{e_1, \ldots, e_n\}$ minimizing the distance to x.
4. (Bessel's Inequality) For every $x \in H$, $\|x\|^2 \geq \sum_{i=1}^n |\langle x, e_i\rangle|^2$.

Theorem 5.5.17. *For an orthonormal set $E \subset H$, the following are equivalent:*

(1) E is maximal,
(2) $\operatorname{span}(E)$*, the set of finite linear combinations of elements of E, is dense in H.*
(3) $\langle x, e\rangle = 0$ for all $e \in E$ implies $x = 0$.
(4) For all $x \in H$, $x = \sum_{e\in E} \langle x, e\rangle e$, where the sum on the right:
- *has at most countably many non-zero terms, and*
- *converges in the norm topology regardless of ordering.*

(5) For all $x \in H$, $\|x\|^2 = \sum_{e\in E} |\langle x, e\rangle|^2$.

If E satisfies the above properties, we call E an orthonormal basis *for H.*

Proof.

<u>(1) ⇒ (2):</u> If $\operatorname{span}(E)$ is not dense, there is an $e \in \operatorname{span}(E)^\perp$ with $\|e\| = 1$. Then $E \subsetneq E \cup \{e\}$, which is orthonormal.

<u>(2) ⇒ (3):</u> Suppose $\langle e, x\rangle = 0$ for all $e \in E$. Then $\langle x| = 0$ on $\operatorname{span}(E)$. Since $\operatorname{span}(E)$ is dense in H and $\langle x|$ is continuous, $\langle x| = 0$ on H, and thus $x = 0$ by the Riesz Representation Theorem 5.5.13.

<u>(3) ⇒ (1):</u> (3) is equivalent to $E^\perp = 0$. This means there is no strictly larger orthonormal set containing E.

<u>(3) ⇒ (4):</u> For all $e_1, \ldots, e_n \in E$, by Bessel's Inequality, $\|x\|^2 \geq \sum^n |\langle x, e_i\rangle|^2$. So for all countable subsets $F \subset E$, $\|x\|^2 \geq \sum_{f\in F} |\langle x, f\rangle|^2$. Hence $\{e \in E | \langle x, e\rangle \neq 0\}$ is countable. Let (e_i) be an enumeration of this set. Then

$$\left\| \sum_m^n \langle x, e_i\rangle e_i \right\|^2 = \sum_m^n |\langle x, e_i\rangle|^2 \xrightarrow{m,n\to\infty} 0.$$

So $\sum \langle x, e_i\rangle e_i$ converges as H is complete. Obsere that for all $e \in E$,

$$\left\langle x - \sum \langle x, e_i\rangle e_i, e \right\rangle = 0,$$

so $x = \sum \langle x, e_i \rangle e_i$ by (3).
(4) $\Rightarrow$ (5): Let $x \in H$ and let $\{e_i\}$ be an enumeration of $\{e \in E | \langle x, e \rangle \neq 0\}$. Then

$$\|x\|^2 - \sum^{n} |\langle x, e_i \rangle|^2 = \left\| x - \sum^{n} \langle x, e_i \rangle e_i \right\|^2 \xrightarrow{n \to \infty} 0.$$

(Indeed, expand the term on the right into 4 terms to see you get the term on the left.)
(5) $\Rightarrow$ (3): Immediate as $\| \cdot \|$ is definite. $\square$

Exercise 5.5.18. Suppose H is a Hilbert space. Prove the following assertions.

(1) Every orthonormal set E can be extended to an orthonormal basis.
(2) H is separable if and only if it has a countable orthonormal basis.
(3) Two Hilbert spaces are isomorphic (there is an invertible $U \in \mathcal{L}(H \to K)$ such that $\langle Ux, Uy \rangle_K = \langle x, y \rangle$ for all $x, y \in H$) if and only if H and K have orthonormal bases which are the same size.
(4) If E is an orthonormal basis, the map $H \to \ell^2(E)$ given by $x \mapsto (\langle x, \cdot \rangle : E \to \mathbb{C})$ is a unitary isomorphism of Hilbert spaces. Here, $\ell^2(E)$ denotes square integrable functions $E \to \mathbb{C}$ with respect to counting measure.

Exercise 5.5.19. Consider the space $L^2(\mathbb{T}) := L^2(\mathbb{R}/\mathbb{Z})$ of $\mathbb{Z}$-periodic functions $\mathbb{R} \to \mathbb{C}$ such that $\int_{[0,1]} |f|^2 < \infty$. Define

$$\langle f, g \rangle := \int_{[0,1]} f\overline{g}.$$

(1) Prove that $L^2(\mathbb{T})$ is a Hilbert space.
(2) Show that the subspace $C(\mathbb{T}) \subset L^2(\mathbb{T})$ of continuous $\mathbb{Z}$-periodic functions is dense.
(3) Prove that $\{e_n(x) := \exp(2\pi i n x) | n \in \mathbb{Z}\}$ is an orthonormal basis for $L^2(\mathbb{T})$.
Hint: Orthonormality is easy. Use (2) and the Stone-Weierstrass Theorem to show the linear span is dense.
(4) Define $\mathcal{F} : L^2(\mathbb{T}) \to \ell^2(\mathbb{Z})$ by $\mathcal{F}(f)_n := \langle f, e_n \rangle_{L^2(\mathbb{T})} = \int_0^1 f(x) \exp(-2\pi i n x)\, dx$. Show that if $f \in L^2(\mathbb{T})$ and $\mathcal{F}(f) \in \ell^1(\mathbb{Z})$, then $f \in C(\mathbb{T})$, i.e., f is a.e. equal to a continuous function.

5.6. **The dual of $C_0(X)$.** Let X be an LCH space. In this section, we prove the Reisz Representation Theorem which characterizes the dual of $C_0(X)$ in terms of Radon measures on X.

Definition 5.6.1. A *Radon measure* on X is a Borel measure which is

- finite on compact subsets of X,
- outer regular on all Borel subsets of X, and
- inner regular on all open subsets of X.

Facts 5.6.2. Recall the following facts about Radon measures on an LCH space X.

(R1) If μ is a Radon measure on X and $E \subset X$ is σ-finite, then μ is σ-finite on E by Exercise 2.5.24(1). Hence every σ-finite Radon measure is regular.
(R2) If X is σ-compact, every Radon measure is σ-finite and thus regular.
(R3) Finite Radon measures on X are exactly finite regular Borel measures on X.

Exercise 5.6.3. Suppose X is LCH and μ is a Radon measure on X. Prove $C_c(X)$ is dense in $\mathcal{L}^1(\mu)$.

Notation 5.6.4. Recall that the *support* of $f : X \to \mathbb{C}$ is $\operatorname{supp}(f) := \{f \neq 0\}$. We say f has *compact support* if $\operatorname{supp}(f) := \overline{\{f \neq 0\}}$ is compact, and we denote the (possibly non-unital) algebra of all continuous functions of compact support by $C_c(X)$. For an open set $U \subset X$, we write $f \prec U$ to denote $0 \leq f \leq 1$ and $\operatorname{supp}(f) \subset U$. Observe that if $f \prec U$, then $f \leq \chi_U$, but the converse need not be true.

Definition 5.6.5. A *Radon integral* on X is a *positive* linear functional $\varphi : C_c(X) \to \mathbb{C}$, i.e., $\varphi(f) \geq 0$ for all $f \in C_c(X)$ such that $f \geq 0$.

Lemma 5.6.6. *Radon integrals are bounded on compact subsets. That is, if $K \subset X$ is compact, there is a $c_K > 0$ such that for all $f \in C_c(X)$ with $\operatorname{supp}(f) \subset K$, $|\varphi(f)| \leq c_K \cdot \|f\|_\infty$.*

Proof. Let $K \subset X$ be compact. Choose $g \in C_c(X)$ such that $g = 1$ on K by the LCH Urysohn Lemma (Exercise 1.2.11(2)).
Step 1: If $f \in C_c(X, \mathbb{R})$ with $\operatorname{supp}(f) \subset K$, then $|f| \leq \|f\|_\infty \cdot g$ on X. So $\|f\|_\infty \cdot g - |f| \geq 0$, and $\|f\|_\infty \cdot g \pm f \geq 0$. Thus $\|f\|_\infty \cdot \varphi(g) \pm \varphi(f) \geq 0$. Hence

$$|\varphi(f)| \leq \varphi(g) \cdot \|f\|_\infty \qquad \forall f \in C_c(X, \mathbb{R}) \text{ with } \operatorname{supp}(f) \subset K.$$

Taking $c_K := \varphi(g)$ works for all $f \in C_c(X, \mathbb{R})$.
Step 2: Taking real and imaginary parts, we see $c_K := 2\varphi(g)$ works for all $f \in C_c(X)$. Indeed,

$$|\varphi(f)| \leq |\varphi(\operatorname{Re}(f))| + |\varphi(\operatorname{Im}(f))| \leq \varphi(g)\|\operatorname{Re}(f)\|_\infty + \varphi(g)\|\operatorname{Im}(f)\|_\infty \leq 2\varphi(g)\|f\|_\infty$$

for all $f \in C_c(X)$ with $\operatorname{supp}(f) \subset K$. □

Theorem 5.6.7 (Riesz Representation)**.** *If φ is a Radon integral on X, there is a unique Radon measure μ_φ on X such that*

$$\varphi(f) = \int f \, d\mu_\varphi \qquad \forall f \in C_c(X).$$

Moreover, μ_φ satisfies:

($\mu_\varphi 1$) For all open $U \subset X$, $\mu_\varphi(U) = \sup\{\varphi(f) | f \in C_c(X) \text{ with } f \prec U\}$, and
($\mu_\varphi 2$) For all compact $K \subset X$, $\mu_\varphi(K) = \inf\{\varphi(f) | f \in C_c(X) \text{ with } \chi_K \leq f\}$.

Proof.
Uniqueness: Suppose μ is a Radon measure such that $\varphi(f) = \int f \, d\mu$ for all $f \in C_c(X)$. If $U \subset X$ is open, then $\varphi(f) \leq \mu(U)$ for all $f \in C_c(X)$ with $f \prec U$. If $K \subset U$ is compact, then by the LCH Urysohn Lemma (Exercise 1.2.11(2)), there is an $f \in C_c(X)$ such that $f \prec U$ and $f|_K = 1$, and

$$\mu(K) \leq \int f \, d\mu = \varphi(f) \leq \mu(U).$$

But μ is inner regular on U as it is Radon, and thus

$$\mu(U) = \sup\{\mu(K) | U \supset K \text{ is compact}\} \leq \sup\{\varphi(f) | f \in C_c(X) \text{ with } f \prec U\} \leq \mu(U).$$

Hence μ satisfies ($\mu_\varphi 1$), so μ is determined on open sets. But since μ is outer regular, μ is determined on all Borel sets.

Existence: For $U \subset X$ open, define $\mu(U) := \sup\{\varphi(f) | f \in C_c(X) \text{ with } f \prec U\}$ and

$$\mu^*(E) := \inf\{\mu(U) | U \text{ is open and } E \subset U\} \qquad E \subset X.$$

Step 1: μ is monotone on inclusions of open sets, i.e., $U \subset V$ both open implies $\mu(U) \leq \mu(V)$. Hence $\mu^*(U) = \mu(U)$ for all open U.

> *Proof.* Just observe that if $U \subseteq V$ are open, then $f \in C_c(X)$ with $f \prec U$ implies $f \prec V$. Hence $\mu(U) \leq \mu(V)$ are we are taking sup over a super set. □

Step 2: μ^* is an outer measure on X.

> *Proof.* It suffices to prove that if (U_n) is a sequence of open sets, then $\mu\left(\bigcup U_n\right) \leq \sum \mu(U_n)$. This shows that
>
> $$\mu^*(E) = \inf\left\{\sum \mu(U_n) \middle| \text{the } U_n \text{ are open and } E \subset \bigcup U_n\right\},$$
>
> which we know is an outer measure by Proposition 2.3.3. Suppose $f \in C_c(X)$ with $f \prec \bigcup U_n$. Since $\operatorname{supp}(f)$ is compact, $\operatorname{supp}(f) \subset \bigcup_{n=1}^N U_n$ for some $N \in \mathbb{N}$.
>
> **Trick.** By Exercise 1.2.17, there are $g_1, \ldots, g_N \in C_c(X)$ such that $g_n \prec U_n$ and $\sum_{n=1}^N g_n = 1$ on $\operatorname{supp}(f)$.
>
> Then $f = f\sum_{n=1}^N g_n$ and $fg_n \prec U_n$ for each n, so
>
> $$\varphi(f) = \sum_{n=1}^N \varphi(fg_n) \leq \sum_{n=1}^N \varphi(\chi_{U_n}) = \sum_{n=1}^N \mu(U_n) \leq \sum \mu(U_n).$$
>
> Since $f \prec U$ was arbitrary,
>
> $$\mu\left(\bigcup U_n\right) = \sup\left\{\varphi(f) \middle| f \in C_c(X) \text{ with } f \prec \bigcup U_n\right\} \leq \sum \mu(U_n). \qquad \square$$

Step 3: Every open set is μ^*-measurable, and thus $\mathcal{B}_X \subset \mathcal{M}^*$, the μ^*-measurable sets. Hence $\mu_\varphi := \mu^*|_{\mathcal{B}_X}$ is a Borel measure which is by definition outer regular and satisfies ($\mu_\varphi 1$).

> *Proof.* Suppose $U \subset X$ is open. We must prove that for every $E \subset X$ such that $\mu^*(E) < \infty$, $\mu^*(E) \geq \mu^*(E \cap U) + \mu^*(E \setminus U)$.
>
> Case 1: If E is open, then $E \cap U$ is open. Given $\varepsilon > 0$, there is a $f \in C_c(X)$ with $f \prec E \cap U$ such that $\varphi(f) > \mu(E \cap U) - \varepsilon/2$. Since $E \setminus \operatorname{supp}(f)$ is open, there is a $g \prec E \setminus \operatorname{supp}(f)$ such that $\varphi(g) > \mu(E \setminus \operatorname{supp}(f)) - \varepsilon/2$. Then $f + g \prec E$, so
>
> $$\begin{aligned} \mu(E) &\geq \varphi(f) + \varphi(g) \\ &> \mu(E \cap U) + \mu(E \setminus \operatorname{supp}(f)) - \varepsilon \\ &\geq \mu^*(E \cap U) + \mu^*(E \setminus U) - \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, the result follows.

Case 2: For a general E, given $\varepsilon > 0$, there is an open $V \supseteq E$ such that $\mu(V) < \mu^*(E) + \varepsilon$. Thus

$$\begin{aligned}\mu^*(E) + \varepsilon &> \mu(V)\\ &\geq \mu^*(V \cap U) + \mu^*(V \setminus U)\\ &\geq \mu^*(E \cap U) + \mu^*(E \setminus U).\end{aligned}$$

Again, as $\varepsilon > 0$ was arbitrary, the result follows. □

Step 4: μ_φ satisfies $(\mu_\varphi 2)$ and is thus finite on compact sets.

Proof. Suppose $K \subset X$ is compact and $f \in C_c(X)$ with $\chi_K \leq f$. Let $\varepsilon > 0$, and set $U_\varepsilon := \{1 - \varepsilon < f\}$, which is open. If $g \in C_c(X)$ with $g \prec U_\varepsilon$, then $(1-\varepsilon)^{-1} f - g \geq 0$, so $\varphi(g) \leq (1-\varepsilon)^{-1}\varphi(f)$. Hence

$$\mu_\varphi(K) \leq \mu_\varphi(U_\varepsilon) = \sup\{\varphi(g) | g \prec U_\varepsilon\} \leq (1-\varepsilon)^{-1}\varphi(f).$$

As $\varepsilon > 0$ was arbitrary, we conclude that $\mu_\varphi(K) \leq \varphi(f)$.
Now, for all open $U \supset K$, by the LCH Urysohn Lemma (Exercise 1.2.11(2)), there is an $f \prec U$ such that $\chi_K \leq f$ ($f|_K = 1$), and by definition, $\varphi(f) \leq \mu_\varphi(U)$. Since μ_φ is outer regular on K by definition,

$$\mu_\varphi(K) = \inf\{\mu_\varphi(U) | K \subset U \text{ open}\} = \inf\{\varphi(f) | f \geq \chi_K\}.$$ □

Step 5: μ_φ is inner regular on open sets and thus Radon.

Proof. If $U \subset X$ is open and $0 \leq \alpha < \mu(U)$, choose $f \in C_c(X)$ such that $f \prec U$ and $\varphi(f) > \alpha$. For all $g \in C_c(X)$ with $\chi_{\operatorname{supp}(f)} \leq g$, we have $g - f \geq 0$, so $\alpha < \varphi(f) \leq \varphi(g)$. Since $(\mu_\varphi 2)$ holds, $\alpha < \mu(\operatorname{supp}(f)) \leq \mu(U)$. Hence μ is inner regular on U. □

Step 6: For all $f \in C_c(X)$, $\varphi(f) = \int f \, d\mu_\varphi$.

Proof. We may assume $f \in C_c(X, [0,1])$ as this set spans $C_c(X)$. Fix $N \in \mathbb{N}$, and set $K_j := \{f \geq j/N\}$ for $j = 1, \ldots, N+1$ and $K_0 := \operatorname{supp}(f)$ so that

$$\emptyset = K_{N+1} \subset K_N \subset \cdots \subset K_1 \subset K_0 = \operatorname{supp}(f).$$

for $j = 1, \ldots, N$, define

$$f_j := \left(\left(f - \frac{j-1}{N}\right) \vee 0\right) \wedge \frac{1}{N}$$

which is equivalent to

$$f_j(x) = \begin{cases} 0 & \text{if } x \notin K_{j-1} \\ f(x) - \frac{j-1}{N} & \text{if } x \in K_{j-1} \setminus K_j \\ N^{-1} & \text{if } x \in K_j. \end{cases}$$

Observe that this implies:

- $\frac{\chi_{K_j}}{N} \leq f_j \leq \frac{\chi_{K_{j-1}}}{N}$ for all $j = 1, \ldots, N$, and
- $\sum_{j=1}^N f_j = f$,

which gives us the inequalities

$$\frac{1}{N}\mu_\varphi(K_j) \leq \int f_j \, d\mu_\varphi \leq \frac{1}{N}\mu_\varphi(K_{j-1}). \tag{5.6.8}$$

Now for all open $U \supset K_{j-1}$, $Nf_j \prec U$, so $N\varphi(f_j) \leq \mu_\varphi(U)$. By ($\mu_\varphi 2$) and outer regularity of μ_φ, we have the inequalities

$$\frac{1}{N}\mu_\varphi(K_j) \leq \varphi(f_j) \leq \frac{1}{N}\mu_\varphi(K_{j-1}). \tag{5.6.9}$$

Now summing over $j = 1, \ldots N$ for both (5.6.8) and (5.6.9), we have the inequalities

$$\frac{1}{N}\sum_{j=1}^{N}\mu_\varphi(K_j) \leq \int f \, d\mu_\varphi \leq \frac{1}{N}\sum_{j=0}^{N-1}\mu_\varphi(K_j)$$
$$\frac{1}{N}\sum_{j=1}^{N}\mu_\varphi(K_j) \leq \varphi(f) \leq \frac{1}{N}\sum_{j=0}^{N-1}\mu_\varphi(K_j).$$

This implies that

$$\left|\varphi(f) - \int f \, d\mu_\varphi\right| \leq \frac{\mu_\varphi(K_0) - \mu_\varphi(K_N)}{N} \leq \frac{\mu_\varphi(\operatorname{supp}(f))}{N} \xrightarrow{N\to\infty} 0$$

as $\mu_\varphi(\operatorname{supp}(f)) < \infty$ and $N \in \mathbb{N}$ was arbitrary. □

This completes the proof. □

The following corollary is the upgrade of Proposition 2.5.22 promised in Remark 2.5.26.

Corollary 5.6.10. *Suppose X is LCH and every open subset of X is σ-compact (e.g., if X is second countable). Then every Borel measure on X which is finite on compact sets is Radon.*

Proof. Suppose μ is such a Borel measure. Since $C_c(X) \subset L^1(\mu)$, $\varphi(f) := \int f \, d\mu$ is a positive linear functional on $C_c(X)$. By the Riesz Representation Theorem 5.6.7, there is a unique Radon measure ν on C such that $\varphi(f) = \int f \, d\nu$ for all $C_c(X)$. It remains to prove $\mu = \nu$.

For an open $U \subset X$, write $U = \bigcup K_j$ with K_j compact for all j. We may inductively find $f_n \in C_c(X)$ such that $f_n \prec U$ and $f_n = 1$ on the compact set $\bigcup^n K_j \cup \bigcup^{n-1} \operatorname{supp}(f_j)$. Then $f_n \nearrow \chi_U$ pointwise, so by the MCT 3.3.9,

$$\mu(U) = \lim \int f_n \, d\mu = \lim \varphi(f_n) = \lim \int f_n \, d\nu = \nu(U).$$

Now suppose $E \in \mathcal{B}_X$ is arbitrary. By (R2), ν is a regular Borel measure, so by Exercise 2.5.23, given $\varepsilon > 0$, there are $F \subset E \subset U$ with F closed, U open, and $\nu(U \setminus F) < \varepsilon$. But since $U \setminus F$ is open,

$$\mu(U \setminus F) = \nu(U \setminus F) < \varepsilon,$$

and thus $\mu(U) - \varepsilon \leq \mu(E) \leq \mu(U)$. Hence μ is outer regular, and thus $\mu = \nu$. □

Lemma 5.6.11. *Suppose X is LCH and μ is a Radon measure on X. Define $\varphi(f) := \int f\, d\mu$ on $C_c(X)$. The following are equivalent:*

(1) φ extends continuously to $C_0(X)$.
(2) φ is bounded with respect to $\|\cdot\|_\infty$.
(3) $\mu(X)$ is finite.

Proof.
(1) $\Leftrightarrow$ (2): This follows as $C_c(X) \subset C_0(X)$ is dense with respect to $\|\cdot\|_\infty$ by the LCH Urysohn Lemma (Exercise 1.2.11(2)).
(2) $\Leftrightarrow$ (3): This follows as $\mu(X) = \sup\left\{\varphi(f) = \int f\, d\mu \middle| f \in C_c(X) \text{ with } 0 \leq f \leq 1\right\}$. □

Corollary 5.6.12. *A positive linear functional in $C_0(X)^*$ is of the form $\int \cdot\, d\mu$ for some finite Radon measure μ.*

Proposition 5.6.13. *If $\varphi \in C_0(X, \mathbb{R})^*$, there are positive $\varphi_\pm \in C_0(X, \mathbb{R})^*$ such that $\varphi = \varphi_+ - \varphi_-$. Hence there are finite Radon measures μ_1, μ_2 on X such that*

$$\varphi(f) = \int f\, d\mu_1 - \int f\, d\mu_2 = \int f\, d(\mu_1 - \mu_2) \qquad \forall f \in C_0(X, \mathbb{R}).$$

Proof. For $f \in C_0(X, [0\infty))$, define $\varphi_+(f) := \sup\{\varphi(g) | 0 \leq g \leq f\}$. For $f \in C_0(X, \mathbb{R})$, define $\varphi_+(f) := \varphi_+(f_+) - \varphi_+(f_-)$ as $f_\pm \in C_0(X, [0, \infty))$.
Step 1: For all $f_1, f_2 \in C_0(X, [0, \infty))$ and $c \geq 0$, $\varphi_+(cf_1 + f_2) = c\varphi_+(f_1) + \varphi_+(f_2)$.

> *Proof.* It suffices to show additivity. Whenever $0 \leq g_1 \leq f_1$ and $0 \leq g_2 \leq f_2$, $0 \leq g_1 + g_2 \leq f_1 + f_2$. This implies $\varphi_+(f_1 + f_2) \geq \varphi_+(f_1) + \varphi_+(f_2)$.
> Now if $0 \leq g \leq f_1 + f_2$, set $g_1 := g \wedge f_1$ and $g_2 := g - g_1$. Then $0 \leq g_1 \leq f_1$ and $0 \leq g_2 \leq f_2$, so
> $$\varphi(g) = \varphi(g_1) + \varphi(g_2) \leq \varphi_+(f_1) + \varphi_+(f_2).$$
> Taking sup over such g gives $\varphi_+(f_1 + f_2) \leq \varphi_+(f_1) + \varphi_+(f_2)$. □

Step 2: If $f \in C_0(X, \mathbb{R})$ with $f = g - h$ where $g, h \geq 0$, then $\varphi_+(f) = \varphi_+(g) - \varphi_+(h)$.

> *Proof.* Observe that $g + f_- = h + f_+ \geq 0$, so $\varphi_+(g) + \varphi(f_-) = \varphi_+(h) + \varphi_+(f_+)$ by Step 1. Rearranging gives the result. □

Step 3: φ_+ is linear on $C_0(X, \mathbb{R})$.

> *Proof.* Suppose $c \in \mathbb{R}$ and $f, g \in C_0(X, \mathbb{R})$. If $c \geq 0$, then $cf + g = cf_+ + g_+ - (cf_- + g_-)$ where $cf_\pm + g_\pm \geq 0$. Then
> $$\begin{aligned} \varphi_+(cf + g) &= \varphi_+(cf_+ + g_+) - \varphi_+(cf_- + g_-) && \text{(Step 2)} \\ &= c\varphi_+(f_+) + \varphi_+(g_+) - c\varphi_+(f_-) - \varphi_-(g_-) && \text{(Step 1)} \\ &= c(\varphi_+(f_+) - \varphi_+(f_-)) + (\varphi_+(g_+) - \varphi_+(g_-)) \\ &= c\varphi_+(f) + \varphi_+(g) && \text{(Step 2).} \end{aligned}$$ □

Step 4: $\varphi_+ \in C_0(X,\mathbb{R})^*$ is positive with $\|\varphi_+\| \leq \|\varphi\|$.

Proof. First suppose $f \in C_0(X,[0,\infty))$. Since

$$|\varphi(g)| \leq \|\varphi\| \cdot \|g\|_\infty \leq \|\varphi\| \cdot \|f\|_\infty \qquad \forall\, 0 \leq g \leq f,$$

we have that

$$0 = \varphi(0) \leq \varphi_+(f) \leq \|\varphi\| \cdot \|f\|_\infty \qquad \forall\, f \in C_0(X,[0,\infty)).$$

Now if $f \in C_0(X,\mathbb{R})$ is arbitrary,

$$|\varphi_+(f)| \leq \max\{\varphi_+(f_+), \varphi_+(f_-)\} \leq \|\varphi\| \cdot \max\{\|f_+\|_\infty, \|f_-\|_\infty\} \leq \|\varphi\| \cdot \|f\|_\infty.$$

Hence $\|\varphi_+\| \leq \|\varphi\|$. □

Step 5: Finally, the linear functional $\varphi_- := \varphi_+ - \varphi \in C_0(X,\mathbb{R})^*$ is also positive as $\varphi_+(f) \geq \varphi(f)$ for all $f \in C_0(X,[0,\infty))$ by definition of φ_+. □

Exercise 5.6.14. For $\varphi \in C_0(X)^*$, there are finite Radon measures $\mu_0, \mu_1, \mu_2, \mu_3$ on X such that

$$\varphi(f) = \sum_{k=0}^{3} i^k \int f\, d\mu_k = \int f\, d\left(\sum_{k=0}^{3} i^k \mu_k\right) \qquad \forall\, f \in C_0(X).$$

Definition 5.6.15. Let X be an LCH space.

- A signed Borel measure ν on X is called a *signed Radon measure* if $\nu_\pm$ are Radon, where $\nu = \nu_+ - \nu_-$ is the Jordan decomposition of ν. We denote by $\mathsf{RM}(X,\mathbb{R}) \subset M(X,\mathbb{R})$ the subset of finite signed Radon measures.
- A complex Borel measure ν on X is called a *complex Radon measure* if $\operatorname{Re}(\nu), \operatorname{Im}(\nu)$ are Radon. We denote by $\mathsf{RM}(X,\mathbb{C}) \subset M(X,\mathbb{C})$ the subset of complex Radon measures.

Exercise 5.6.16 (Lusin's Theorem)**.** Suppose X is LCH and μ is a Radon measure on X. If $f : X \to \mathbb{C}$ is measurable and vanishes outside a set of finite measure, then for all $\varepsilon > 0$, there is an $E \in \mathcal{B}_X$ with $\mu(E^c) < \varepsilon$ and a $g \in C_c(X)$ such that $g = f$ on E. Moreover:

- If $\|f\|_\infty < \infty$, we can arrange that $\|g\|_\infty \leq \|f\|_\infty$.
- If $\operatorname{im}(f) \subset \mathbb{R}$, we can arrange that $\operatorname{im}(g) \subset \mathbb{R}$.

Theorem 5.6.17 (Real Riesz Representation)**.** *Suppose X is LCH. Define $\Phi : \mathsf{RM}(X,\mathbb{R}) \to C_0(X,\mathbb{R})^*$ by $\nu \mapsto \varphi_\nu$ where $\varphi_\nu(f) := \int f\, d\nu$. Then Φ is an isometric linear isomorphism.*

Proof. Clearly Φ is linear. By Proposition 5.6.13, Φ is surjective. It remains to prove Φ is isometric, which also implies injectivity. Fix $\nu \in \mathsf{RM}$.

$\|\varphi_\nu\| \leq \|\nu\|$: For all $f \in C_0(X,\mathbb{R})$,

$$|\varphi_\nu(f)| = \left|\int f\, d\nu\right| = \left|\int f\, d\nu_+ - \int f\, d\nu_-\right| \leq \left|\int f\, d\nu_+\right| + \left|\int f\, d\nu_-\right|$$
$$\leq \int |f|\, d\nu_+ + \int |f|\, d\nu_- = \int |f|\, d|\nu| \leq \|f\|_\infty \cdot \|\nu\|_{\mathsf{RM}}.$$

Hence $\|\varphi_\nu\| \leq \|\nu\|$.

$\underline{\|\varphi_\nu\| \geq \|\nu\|}$: Since ν is finite, by Exercise 4.2.11, $\left|\frac{d\nu}{d|\nu|}\right| = 1$ on X $|\nu|$-a.e. Let $\varepsilon > 0$. Since $|\nu|$ is finite, by Lusin's Theorem (Exercise 5.6.16), there is an $f \in C_c(X, \mathbb{R})$ such that $\|f\|_\infty = 1$ and $f = \overline{\frac{d\nu}{d|\nu|}}$ on $E \in \mathcal{B}_X$ where $|\nu|(E^c) < \varepsilon/2$. Then

$$\begin{aligned}
\|\nu\| &= \int d|\nu| = \int \left|\frac{d\nu}{d|\nu|}\right|^2 d|\nu| = \int \overline{\frac{d\nu}{d|\nu|}} \cdot \frac{d\nu}{d|\nu|}\, d|\nu| \underset{(\text{Ex. } 4.2.11)}{=} \int \overline{\frac{d\nu}{d|\nu|}}\, d\nu \\
&\leq \left|\int f\, d\nu\right| + \left|\int f - \overline{\frac{d\nu}{d|\nu|}}\, d\nu\right| \leq \|\varphi_\nu\| \cdot \underbrace{\|f\|_\infty}_{=1} + \int \left|f - \overline{\frac{d\nu}{d|\nu|}}\right| d|\nu| \\
&\leq \|\varphi_\nu\| + 2|\nu|(E^c) \leq \|\varphi_\nu\| + \varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, $\|\nu\| \leq \|\varphi_\nu\|$. □

Exercise 5.6.18 (Complex Riesz Representation)**.** Suppose X is LCH. Define $\Phi : \mathsf{RM}(X, \mathbb{C}) \to C_0(X, \mathbb{C})^*$ by $\nu \mapsto \varphi_\nu$ where $\varphi_\nu(f) := \int f\, d\nu$. Show that Φ is an isometric linear isomorphism.