Real Analysis I: Final Review

Sam Auyeung

January 13, 2018

Notes from *Real Analysis: Modern Techniques and Their Applications*, 2nd ed. by Gerald B. Folland.

1 Measures

The concept of measures has roots in geometry where area and volume are discussed. We would like to find a function $\mu : X \to \mathbb{R}$ such that:

1. If $E_1, E_2, ...$ is a finite or infinite sequence of disjoint sets, then

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

2. If E is congruent to F (E can be transformed to F via translations, rotations, and reflections), then $\mu(E) = \mu(F)$.

3. $\mu(Q) = 1$ where Q is the unit cube in $\mathbb{R}^n$.

1.1 Non-Measurable Sets

However, there is no such measure μ. To see this, let $n = 1$ and define the equivalence relation on $[0, 1)$: $x \sim y$ iff $x - y \in \mathbb{Q}$. Let $N \subset [0, 1)$ that contains precisely one member of each equivalence class (invoke Axiom of Choice). Now let $R = \mathbb{Q} \cap [0, 1)$ and for each $r \in R$, let

$$N_r = \{x + r : x \in N \cap [0, 1 - r)\} \cup \{x + r - 1 : x \in N \cap [1 - r, 1)\}.$$

In English, we create N_r by shifting N to the right by r units and then the part that sticks out beyond $[0, 1)$, we shift left by one unit. Then $N_r \subset [0, 1)$ and every $x \in [0, 1)$ belongs to precisely one N_r. If $y \in N$ and $y \sim x$, then $x \in N_r$ where $r = x - y$ if $x \geq y$ or $r = x - y + 1$ if $x < y$. Furthermore, $N_r \cap N_s = \varnothing$ because if there were an element x in the intersetion, this would mean that we have two distinct elements of N belonging to the same equivalence class which is a contradiction by how we constructed N.

Suppose then that $\mu : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ satisfies (1), (2), and (3). By (1) and (2),

$$\mu(N) = \mu(N \cap [0, 1 - r)) + \mu(N \cap [1 - r, 1) = \mu(N_r)$$

for any $r \in R$. Since R is countable and $[0, 1)$ is the disjoint union of the N_r's,

$$\mu([0, 1)) = \sum_{r \in R} \mu(N_r) = \sum_{n=1}^{\infty} \mu(N)$$

by (1). But (3) requires $\mu([0, 1)) = 1$. If $\mu(N) > 0$, then the RHS is ∞. If $\mu(N) = 0$, then the RHS is 0. In either case, we have $1 = \infty$ or $1 = 0$, both leading to contradictions.

1.2 σ-Algebras

For the following definitions, let X be a nonempty set.

Definition 1.1. *An* ***algebra*** *of sets on X is a nonempty collection $\mathcal{A}$ of subsets of X that is closed under finite unions and complements. Observe that*

$$\bigcap_{j=1}^{n} E_j = (\bigcup_{j=1}^{n} E_j^c)^c$$

which means that it is also closed under finite intersections.

Note that if $\mathcal{A}$ is an algebra, it contains $\varnothing, X$ because if $E \in \mathcal{A}$, then $X = E \cup E^c, \varnothing = E \cap E^c$.

Definition 1.2. *A* σ***-algebra*** *is an algebra closed under countable unions. By a similar observation from above, a σ-algebra is closed under countable intersections.*

Lemma 1.3. *An algebra $\mathcal{A}$ is a σ-algebra if it is closed under countable disjoint unions.*

Proof. Suppose $\{E_j\}^\infty \subset \mathcal{A}$. Set

$$F_k = E_k \setminus [\bigcup^{k-1} E_j] = E_k \cap [\bigcup^{k-1} E_j]^c.$$

Then the F_k's belong to $\mathcal{A}$ and are disjoint while $\bigcup^\infty E_j = \bigcup^\infty F_k$. □

This trick of creating a disjoint sequence of sets is good to remember.

The intersection of a family of σ-algebras is a σ-algebra. Furthermore, for $\mathcal{E} \subset \mathcal{P}(X)$, there is a unique smallest σ-algebra $\mathcal{M}(\mathcal{E})$ containing $\mathcal{E}$; $\mathcal{M}(\mathcal{E})$ is the intersection of all σ-algebras containing $\mathcal{E}$. We say that $\mathcal{M}(\mathcal{E})$ is **generated** by $\mathcal{E}$.

Definition 1.4. *Let X be a topological space and $\mathcal{T} \subset \mathcal{P}(X)$ the topology of X. Then $\mathcal{M}(\mathcal{T})$ is the σ-algebra generated by open (or equivalently, closed) sets of X. This σ-algebra is called the* ***Borel*** σ***-algebra*** *on X and is denoted $\mathcal{B}_X$.*

Proposition 1.5. *$\mathcal{B}_{\mathbb{R}}$ is generated by any one of the following:*

1. *the open intervals of the form (a, b),*
2. *the closed intervals of the form $[a, b]$,*
3. *the half-open intervals of the form $(a, b]$ or $[a, b)$,*
4. *the open rays of the form (a, ∞) or $(-\infty, a)$,*
5. *the closed rays of the form $[a, \infty)$ or $(-\infty, a]$.*

Definition 1.6. *Let $\{X_\alpha\}_{\alpha\in A}$ be an indexed collection of nonempty sets, $X = \prod_{\alpha\in A} X_\alpha$, and $\pi_\alpha : X \to X_\alpha$ the coordinate maps. If $\mathcal{M}_\alpha$ is a σ-algebraon X_α for each α, the* ***product*** σ***-algebra*** *on X is the σ-algebra generated by*

$$\{\pi_\alpha^{-1}(E_\alpha : E_\alpha \in \mathcal{M}_\alpha, \alpha \in A)\}.$$

We denote it $\bigotimes_{\alpha\in A} \mathcal{M}_\alpha$.

Look at p. 23 for some propositions about $\bigotimes_{\alpha\in A} \mathcal{M}_\alpha$.

Definition 1.7. *An* ***elementary family*** *is a collection $\mathcal{E}$ of subsets of X such that*

- $\varnothing \in \mathcal{E}$,
- *if* $E, F \in \mathcal{E}$, *then* $E \cap F \in \mathcal{E}$,
- *if* $E \in \mathcal{E}$ *then* E^c *is a finite disjoint union of members of* $\mathcal{E}$.

Proposition 1.8. *If* $\mathcal{E}$ *is an elementary family, the collection* $\mathcal{A}$ *of finite disjoint unions of members of* $\mathcal{E}$ *is an algebra.*

1.3 Measures

Definition 1.9. *A* ***measure*** *on* $\mathcal{M}$ *is a function* $\mu : \mathcal{M} \to [0, \infty]$ *such that:*

1. $\mu(\varnothing) = 0$,
2. *(countable additivity) if* $\{E_j\}^\infty$ *is a sequence of disjoint sets in* $\mathcal{M}$, *then* $\mu(\bigcup^\infty E_j) = \sum^\infty \mu(E_j)$.

Here are some definitions about the "size" of μ.

Definition 1.10.

1. *If* $\mu(X) < \infty$, μ *is called* ***finite***. *Note that this implies* $\mu(E) < \infty$ *for all* $E \in \mathcal{M}$.
2. *If* $X = \bigcup^\infty E_j$ *where* $E_j \in \mathcal{M}$ *and* $\mu(E_j) < \infty$ *for all* j, *then* μ *is said to be* ***σ-finite***.
3. *If for each* $E \in \mathcal{M}$ *with* $\mu(E) = \infty$, *there exists* $F \in \mathcal{M}$ *with* $F \subset E$ *and* $0 < \mu(F) < \infty$, μ *is called* ***semifinite***.

The counting measure is not σ-finite. See p. 25 for other examples of measures.

Theorem 1.11. *Let* $(X, \mathcal{M}, \mu)$ *be a measure space.*

1. *(Monotonicity) If* $E, F \in \mathcal{M}$ *and* $E \subset F$, *then* $\mu(E) \leq \mu(F)$.
2. *(Subadditivity) If* $\{E_j\}^\infty \subset \mathcal{M}$, *then* $\mu(\bigcup^\infty E_j) \leq \sum^\infty \mu(E_j)$.
3. *(Continuity from Below) If* $\{E_j\}^\infty \subset \mathcal{M}$ *and* $E_1 \subset E_2 \subset ...$, *then* $\mu(\bigcup^\infty E_j) = \lim_{j\to\infty} \mu(E_j)$.
4. *(Continuity from Above) If* $\{E_j\}^\infty \subset \mathcal{M}$ *and* $E_1 \supset E_2 \supset ...$, *and* $\mu(E_1) < \infty$, *then* $\mu(\bigcap^\infty E_j) = \lim_{j\to\infty} \mu(E_j)$.

Definition 1.12. *A set* E *with measure zero is called a* ***null set***. *If a statement about points* $x \in X$ *is true for all* x *except for* x *in some null set, we say it is true* ***almost everywhere*** *(abbreviate a.e.). If* $F \subset E$ *and* $\mu(E) = 0$, *we say that* $\mathcal{M}$ *is complete if and only if* $F \in \mathcal{M}$.

The next theorem shows that we needn't worry too much about incomplete spaces.

Theorem 1.13. *Suppose that* $(X, \mathcal{M}, \mu)$ *is a measure space. Let* $\mathcal{N} = \{N \in \mathcal{M} : \mu(N) = 0\}$ *and* $\overline{\mathcal{M}} = \{E \cap F : E \in \mathcal{M}; F \subset N \text{ where } N \in \mathcal{N}\}$. *Then* $\overline{\mathcal{M}}$ *is a* σ*-algebra and there is a unique extension* $\overline{\mu}$ *of* μ *to a complete measure on* $\overline{\mathcal{M}}$.

1.4 Outer Measures

The concept of outer measure deals with approximating something from above; e.g. estimating the area of a shape by calculating the area of rectangles covering the shape. The area of the rectangles are easy to calculate and on a first pass, we have a rough approximation. By letting the grid of rectangles get finer, the approximation improves.

Definition 1.14. *An **outer measure** on a nonempty set X is a function $\mu^* : \mathcal{P}(X) \to [0, \infty]$ that satisfies*

- $\mu^*(\varnothing) = 0$,
- $\mu^*(A) \le \mu^*(B)$ *if* $A \subset B$,
- $\mu^*(\bigcup^\infty A_j) \le \sum^\infty \mu^*(A_j)$.

Proposition 1.15. *Let $\mathcal{E} \subset \mathcal{P}(X)$ and $\rho : \mathcal{E} \to [0, \infty]$ be such that $\varnothing \in \mathcal{E}$ and $\rho(\varnothing) = 0$. For any $A \subset X$, define*

$$\mu^*(A) = \inf \left\{ \sum^\infty \rho(E_j) : E_j \in \mathcal{E} \text{ and } A \subset \bigcup^\infty E_j \right\}.$$

Then μ^* is an outer measure.

Definition 1.16. *If μ^* is an outer measure on X, a set $A \subset X$ is called μ^***-measurable** if*

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c) \text{ for all } E \subset X.$$

If $A \subset E$, and E is "well-behaved, then we may think of $\mu^*(A) = \mu^*(E \cap A)$ as the outer measure of A as it gives a larger value for what $E \cap A$ is. The inner measure is $\mu^*(E) - \mu^*(E \cap A^c)$ because $\mu^*(E \cap A^c)$ overshoots on giving an estimate for the measure of what is between E and A^c. Then, subtracting an overshoot gives a smaller-than-actual value. To say something is μ^*-measurable is to say that $\mu^*(A) = \mu^*(E \cap A) = \mu^*(E) - \mu^*(E \cap A^c)$.

Theorem 1.17 (**Carathéodory's Theorem**, p. 29)**.** *If μ^* is an outer measure on X, the collection $\mathcal{M}$ of μ^*-measurable sets is a σ-algebra, and the restriction of μ^* to $\mathcal{M}$ is a complete measure.*

One application of Carath éodory's Theorem is extending measures from algebras to σ-algebras.

Definition 1.18. *If $\mathcal{A} \subset \mathcal{P}(X)$ is an algebra, a function $\mu_0 : \mathcal{A} \to [0, \infty]$ is a **premeasure** if*

1. $\mu_0(\varnothing) = 0$,
2. *if $\{A_j\}^\infty$ is a sequence of disjoint sets in $\mathcal{A}$ such that $\bigcup^\infty A_j \in \mathcal{A}$, then $\mu_0(\bigcup^\infty A_j) = \sum^\infty \mu_0(A_j)$.*

Proposition 1.19. *If μ_0 is a premeasure on $\mathcal{A}$ and μ^* is defined as*

$$\mu^*(E) = \inf \left\{ \sum^\infty \mu_0(A_j) : A_j \in \mathcal{A}, E \subset \bigcup^\infty A_j \right\}, \tag{1.1}$$

then

1. $\mu^*|_\mathcal{A} = \mu_0$,
2. *every set in $\mathcal{A}$ is μ^*-measurable.*

Theorem 1.20. *Let $\mathcal{A} \subset \mathcal{P}(X)$ be an algebra, μ_0 a premeasure on $\mathcal{A}$, and $\mathcal{M}$ the σ-algebra generated by $\mathcal{A}$. There exists a measure μ on $\mathcal{M}$ such that $\mu|_{\mathcal{A}} = \mu_0$ — namely, $\mu = \mu^*|_{\mathcal{M}}$ where μ^* is given by (1.1), If ν is another measure on $\mathcal{M}$ that extends μ_0, then $\nu(E) \leq \mu(E)$ for all $E \in \mathcal{M}$, with equality when $\mu(E) < \infty$. If μ_0 is σ-finite, then μ is the unique extension of μ_0 to a measure on $\mathcal{M}$.*

The proof of this theorem uses Carathéodory's Theorem.

1.5 Borel Measures on the Real Line

We want to build a measure on $\mathcal{B}_{\mathbb{R}}$ from an increasing, right continuous function F. Such a measure is called a **Borel measure** on $\mathbb{R}$. When $F(x) = x$, we'll have the usual "length" measure.

Proposition 1.21. *Let $F : \mathbb{R} \to \mathbb{R}$ be increasing and right continuous; i.e. $\lim_{t\to x^+} f(t) = f(x)$. If $(a_j, b_j]$ $(j = 1, ..., n)$ are disjoint half-open intervals (h-intervals), let*

$$\mu_0\left(\bigcup^n (a_j, b_j]\right) = \sum^n [F(b_j) - F(a_j)],$$

and let $\mu_0(\varnothing) = 0$. Then μ_0 is a premeasure on the algebra $\mathcal{A}$.

Theorem 1.22. *If $F : \mathbb{R} \to \mathbb{R}$ is any increasing, right continuous function, there is a unique Borel measure μ_F on $\mathbb{R}$ such that $\mu_F((a, b]) = F(b) - F(a)$ for all a, b. If G is another such function, we have $\mu_F = \mu_G$ if and only if $F - G$ is constant. Conversely, if μ is a Borel measure on $\mathbb{R}$ that is finite on all bounded Borel sets and we define*

$$F(x) = \begin{cases} \mu((0, x]) & x > 0, \\ 0 & x = 0 \\ -\mu((x, 0]) & x < 0, \end{cases}$$

then F is increasing and right continuous, and $\mu = \mu_F$.

We may obtain a complete measure which we usually also call μ_F from F. μ_F has a strictly larger domain than $\mathcal{B}_{\mathbb{R}}$ and is called the **Lebesgue-Stieltjes measure** associated to F.

Theorem 1.23. *If $E \in \mathcal{M}$, then*

$$\begin{aligned} \mu(E) &= \inf\{\mu(U) : E \subset U \text{ and } U \text{ is open}\} \\ &= \sup\{\mu(K) : K \subset E \text{ and } K \text{ is compact}\} \end{aligned}$$

See pp. 36-37 for some more theorems and propositions. The next theorem says that the **Lebesgue measure** (μ_F where $F(x) = x$), denoted m, is invariant under translations and predictable under dilations.

Theorem 1.24. *If $E \in \mathcal{L}$, then $E + s := \{x + s : x \in E\} \in \mathcal{L}$ and $rE := \{rx : x \in E\} \in \mathcal{L}$ for all $s, r \in \mathbb{R}$. Moreover, $m(E + s) = m(E)$ and $m(rE) = |r|m(E)$.*

Example 1.25. Let $\{r_j^\infty\}$ be an enumeration of the rationals in $[0, 1]$ and let $\epsilon > 0$. Let I_j be the open interval centered at r_j of length $e2^{-j}$. Then the set $U = (0, 1) \cap \bigcup^\infty I_j$ is open and dense in $[0, 1]$ but $m(U) \leq \sum^\infty \epsilon 2^{-j} = \epsilon$. Its complement $K = [0, 1] \setminus U$ is closed and nowhere dense but $m(K) \geq 1 - \epsilon$. Hence, sets which are topologically "large" can be measure-theoretically small, and vice versa. However, nonempty open sets cannot have Lebesgue measure zero.

Example 1.26. The Cantor set C is a standard example of an uncountable set with measure zero. it is also compact, nowhere dense (the closure has empty interior), and totally disconnected.

Example 1.27. Suppose $x \in C$ so that $x = \sum^{\infty} a_j 3^{-j}$ where $a_j = 0$ or 2 for all j. Let $g(x) = \sum^{\infty} b_j 2^{-j}$ where $b_j = a_j/2$. Then g maps C to $[0,1]$. We may extend the g to a function f so that it is constant on each interval missing from C. f is continuous and called the **Cantor function**.

Not every Lebesgue measurable set is Borel. See exercise 9 in Chapter 2 (homework).

2 Integration

2.1 Measurable Functions

We consider a definition similar to that of continuous functions.

Definition 2.1. *If $(X, \mathcal{M}), (Y, \mathcal{N})$ are measurable spaces, a mapping $f : X \to Y$ is called $(\mathcal{M}, \mathcal{N})$-**measurable** if $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{N}$.*

Since Borel sets are generated from the topology of a space, continuous functions $f : X \to Y$ on generic topological spaces X, Y are automatically $(\mathcal{B}_X, \mathcal{B}_Y)$ measurable. For the rest of the chapter, Folland, when referring to measurable functions, always means in the Borel sense; particularl with $\mathcal{B}_{\mathbb{R}}$ or $\mathcal{B}_{\mathbb{C}}$ as the σ-algebras of the range space. So we'll use **this** definition for measurable.

Warning: Something can be Lebesgue measurable but not Borel measurable. If $f, g : \mathbb{R} \to \mathbb{R}$ are Lebesgue measurable, it does not follow that $f \circ g$ is Lebesgue measurable, even if g is continuous. If $E \subset \mathcal{B}_{\mathbb{R}}$, we have $f^{-1}(E) \in \mathcal{L}$ but unless $f^{-1}(E) \in \mathcal{B}_{\mathbb{R}}$, there is no guarantee $g^{-1}(f^{-1}(E))$ will be in $\mathcal{L}$. If f is Borel measurable, then $f \circ g$ is Lebesgue or Borel measurable whenever g is.

Proposition 2.2. *If $\mathcal{N}$ is generated by $\mathcal{E}$, then $f : X \to Y$ is $(\mathcal{M}, \mathcal{N})$-measurable if and only if $f^{-1}(E) \in \mathcal{M}$ for all $E \in \mathcal{E}$.*

Proposition 1.5 gives us some ways of generating $\mathcal{B}_{\mathbb{R}}$. The next proposition shows some equivalent statements about real-valued measurable functions.

Proposition 2.3. *If $(x, \mathcal{M})$ is a measurable space and $f : X \to \mathbb{R}$, the following are equivalent:*

1. *f is $\mathcal{M}$-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 $a \in \mathbb{R}$.*

Proposition 2.4. *If $\{f_j\}$ is a sequence of $\overline{\mathbb{R}}$-valued measurable functions on $(X, \mathcal{M})$, then the functions*

$$g_1(x) = \sup_j f_j(x), \qquad g_3(x) = \lim_{j\to\infty} \sup f_j(x),$$
$$g_2(x) = \inf_j f_j(x), \qquad g_4(x) = \lim_{j\to\infty} \inf f_j(x)$$

are all measurable. If $f(x) = \lim_{j\to\infty} f_j(x)$ exists for every $x \in X$, then f is measurable.

Definition 2.5. *The **characteristic function** χ_E is defined by*

$$\chi_E(x) = \begin{cases} 1 & x \in E, \\ 0 & x \notin E. \end{cases}$$

χ_E is measurable if and only if $E \in \mathcal{M}$.

Definition 2.6. *A **simple function** on X is a finite linear combination, with complex coefficients, of characteristic functions on sets in $\mathcal{M}$. Equivalently, $f : X \to \mathbb{C}$ is simple if and only if f is measurable and the range of f is finite.*

$$f = \sum^n z_j \chi_{E_j}.$$

Theorem 2.7. *Let $(X, \mathcal{M})$ be a measurable space.*

1. *If $f : X \to [0, \infty]$ is measurable, there is a sequence $\{\phi_n\}$ of simple functions such that $0 \leq \phi_1 \leq \phi_2 \leq ... \leq f$, $\phi_n \to f$ pointwise, and $\phi_n \to f$ uniformly on any set on which f is bounded.*

2. *If $f : X \to \mathbb{C}$ is measurable, there is a sequence $\{\phi_n\}$ of simple functions such that $0 \leq |\phi_1| \leq |\phi_2| \leq ... \leq |f|$, $\phi_n \to f$ pointwise, and $\phi_n \to f$ uniformly on any set on which f is bounded.*

2.2 Integration of Nonnegative Functions

Definition 2.8. $L^+ = \{f : X \to [0, \infty] : f \text{ is measurable}\}$

Definition 2.9. *If $\phi \in L^+$ and has representation $\phi = \sum^n a_j \chi_{E_j}$, then we define the **integral** of ϕ with respect to μ by*

$$\int \phi \, d\mu = \sum^n a_j \mu(E_j).$$

We let $0 \cdot \infty = 0$ and allow $\int \phi = \infty$.

Definition 2.10. *Extending to all functions $f \in L^+$,*

$$\int f \, d\mu = \sup \left\{ \int \phi \, d\mu : 0 \leq \phi \leq f, \phi \text{ simple} \right\}.$$

Theorem 2.11 (The Monotone Convergence Theorem). *If $\{f_n\}$ is a sequence in L^+ such that $f_j \leq f_{j+1}$ for all j, and $f = \lim_{n\to\infty} f_n (= \sup_n f_n)$, then $\int f = \lim_{n\to\infty} \int f_n$.*

In general, if the condition $f_j \leq f_{j+1}$ isn't met, then the theorem may fail. The next example demonstrates this:

Example 2.12. Let $f_n(x) = 1$ for $x \geq n$ and 0 elsewhere. It is true that $\{f_n\} \subset L^+$ but $f_n \geq f_{n+1}$ for all n. $f_n \to f = 0$ but $\int f_n = \infty$ for all n and so $\lim_{n\to\infty} \int f_n = \infty$ while $\int f = 0$.

Next, we give an application of the MCT to a theorem which gives us conditions for when we may interchange $\int$ and $\sum$ while mantaining equality. See proof on p. 51.

Theorem 2.13. *If $\{f_n\}$ is a finite or infinite sequence in L^+ and $f = \sum_n f_n$, then $\int f = \sum_n \int f_n$.*

Proposition 2.14. *If $f \in L^+$, then $\int f = 0$ if and only if $f = 0$ a.e.*

Corollary 2.15. *If $\{f_n\} \subset L^+$, $f \in L^+$, $f_n \leq f_{n+1}$, and $f_n \to f$ a.e., then $\int f = \lim \int f_n$.*

The next lemma is used often:

Lemma 2.16 (**Fatou's Lemma**). *If $\{f_n\}$ is any sequence in L^+, then*

$$\int \liminf f_n \leq \liminf \int f_n.$$

To remember which way the inequality goes, remember that IL (integrate then limit) comes before LI (limit then integrate), alphabetically. Also, remember this example:

Example 2.17. Let $f_{2k} = \chi_{[-1,0]}$ and $f_{2k+1} = \chi_{[0,1]}$. Then $\liminf f_n = 0$ but $\int f_n = 1$ for all n. Thus, $0 = \int \liminf f_n < \liminf \int f_n = 1$. Similarly, if we let $f_n = \chi_{[n,n+1]}$, we have a similar strict inequality.

2.3 Integration of Complex Functions

We may extend the definition of integral from the previous section to all real-valued measurable functions f. Let f^+ be the positive part of f:

$$f^+(x) = \begin{cases} f(x), & f(x) \geq 0; \\ 0, & f(x) < 0. \end{cases}$$

Let f^- be the negative part:

$$f^-(x) = \begin{cases} -f(x), & f(x) \leq 0; \\ 0, & f(x) > 0. \end{cases}$$

Note that f^- is a positive function. Then $f = f^+ - f^-$. We say that f is **integrable** if and only if $\int f^+ < \infty$ and $\int f^- < \infty$ if and only if $\int |f| < \infty$. Integrable functions live in the vector space L^1; integration is a linear functional on L^1.

Example 2.18. $f(x) = \frac{1}{x} \sin x$ is not Lebesgue integrable because $\int_{[0,\infty)} |f| = \infty$ (behaves roughly like $\sum^\infty 1/n$). However, f is improperly Riemann integrable; $\int_0^\infty f\, dx = \pi/2$.

Proposition 2.19. *If $f \in L^1$, then $|\int f| \leq \int |f|$.*

For the purposes of having a metric which makes sense on L^1, we take equivalence classes of functions where $f \sim g \Leftrightarrow f = g$ a.e. The reason for this is because the integral of f and g are the same if they equal a.e. Then, the distance between them is well-defined.

Theorem 2.20 (**The Dominated Convergence Theorem**). *Let $\{f_n\}$ be a sequence in L^1 such that*

1. *$f_n \to f$ a.e.*

2. *There exists a nonnegative $g \in L^1$ such that $|f_n| \leq g$ a.e. for all n.*

Then $f \in L^1$ and $\int f = \lim_{n\to\infty} \int f_n$.

Thus, g and $-g$ bound all the f_n.

Example 2.21. Consider the sliding block scenario: $f_n(x)$ equals 1 on $[n, n+1]$ and 0 elsewhere. Then $f_n \to f = 0$ pointwise. It is bounded above by nonnegative g which equals 1 on $[0, \infty)$ and 0 elsewhere; so $|f_n| \le g$ for all n. However, $g \notin L^1$. Thus, $0 = \int f \neq \lim \int f_n = 1$.

The next theorem gives us conditions for interchanging $\int$ and $\sum$, similar to Theorem 2.13. However, note that we gain some more information when we assume the sequence is in L^1.

Theorem 2.22. *Suppose that $\{f_j\}$ is a sequence in L^1 such that $\sum^\infty \int |f_j| < \infty$. Then $\sum^\infty f_j$ converges a.e. to a function in L^1 and $\int \sum^\infty f_j = \sum^\infty \int f_j$.*

Theorem 2.23. *If $f \in L^1$ and $\epsilon > 0$, there is an integrable simple function $\phi = \sum a_j \chi_{E_j}$ such that $\int |f - \phi| d\mu < \epsilon$; i.e. the integrable simple functions are dense in L^1 in the L^1 metric.*

If μ is a Lebesgue-Stieljes measure on $\mathbb{R}$ (see Theorem 1.22), the sets E_j in the definition of ϕ can be taken to be finite unions of open intervals; moreover, there is a continuous function g that vanishes outside a bounded interval such that $\int |f - g| d\mu < \epsilon$.

Theorem 2.24. *Suppose that $f : X \times [a, b] \to \mathbb{C}$ and that $f_t : X \to \mathbb{C}$ is integrable for each $t \in [a, b]$. Let $F(t) = \int_X f(x, t)\, d\mu(x)$.*

1. *Suppose that there exists $g \in L^1$ such that $|f(x, t)| \le g(x)$ for all x, t. If $\lim_{t \to t_0} f(x, t) = f(x, t_0)$ for every x, then $\lim_{t \to t_0} F(t) = F(t_0)$. In particular, if $f(x, \cdot)$ is continuous for each x, then F is continuous.*

2. *Suppose that $\partial f / \partial t$ exists and there is a $g \in L^1$ such that $|(\partial f/\partial t)(x, t)| \le g(x)$ for all x, t. Then F is differentiable and $F'(t) = \int (\partial f/\partial t)(x, t)\, d\mu(x)$.*

The next theorem compares Riemann integrals to Lebesgue integrals.

Theorem 2.25. *Let f be a bounded real-valued function on $[a, b]$.*

1. *If f is Riemann integrable, then f is Lebesgue measurable (and hence integrable on [a,b] since it is bounded), and $\int_a^b f(x)\, dx = \int_{[a,b]} f\, dm$.*

2. *f is Riemann integrable if and only if the set of points at which f is discontinuous has Lebesgue measure zero.*

2.4 Modes of Convergence

Definition 2.26. *$f_n \to f$ in L^1 means $\int |f_n - f| \to 0$ pointwise.*

Definition 2.27. *We say $\{f_n\}$ is **Cauchy in measure** if for every $\epsilon > 0$,*

$$\mu(\{x : |f_n(x) - f_m(x)| \ge \epsilon\}) \to 0 \text{ as } m, n \to \infty.$$

*$\{f_n\}$ **converges in measure** to f if for every $\epsilon > 0$,*

$$\mu(\{x : |f_n(x) - f(x)| \ge \epsilon\}) \to 0 \text{ as } n \to \infty.$$

Definition 2.28. *$f_n \to f$ almost uniformly means that for every $\epsilon > 0$ there exists a set E such that $\mu(E^c) < \epsilon$ and $f_n \to f$ uniformly on E.*

The following functions will be useful as counterexamples:

Example 2.29. These examples are crucial.

1. $f_n = n^{-1}\chi_{(0,n)}$.

2. $f_n = \chi_{(n,n+1)}$.

3. $f_n = n\chi_{[0,1/n]}$.

4. $f_1 = \chi_{[0,1]}, f_2 = \chi_{[0,1/2]}, f_3 = \chi_{[1/2,1]}, f_4 = \chi_{[0,1/4]}, f_5 = \chi_{[1/4,1/2]}, f_6 = \chi_{[1/2,3/4]}, f_7 = \chi_{[3/4,1]}$, and in general, $f_n = \chi_{[j/2^k,(j+1)/2^k]}$ where $n = 2^k + j$ with $0 \leq j < 2^k$.

5. $f_{2k+1} = \chi_{[0,1]}, f_{2k} = \chi_{[-1,0]}$.

In (1),(2), (3), $f_n \to 0$ uniformly, pointwise, and a.e., respectively, but none converges to 0 in L^1 because for each of them, $\int |f_n| = \int f_n = 1$.

In (4), $f_n \to 0$ in L^1 since $\int |f_n| = 2^{-k}$ for $2^k \leq n < 2^{k+1}$, but $f_n(x)$ does not converge for any $x \in [0,1]$ because there are, for infinitey many n, $f_n(x) = 0$ and infinitely many n for which $f_n(x) = 1$.

In (3), the sequence converges to $f = 0$ in measure since the intervals on which f_n is nonzero gets smaller. However, $\lim_{n\to\infty} \int |f_n - f| = \lim_{n\to\infty} \int f_n = 1 \neq 0$. So (3) is also an example of convergence in measure but not in L^1.

In (5), there is a subsequence on which it converges everywhere. However, the sequence itself does not converge in any mode.

Lemma 2.30. *If $f_n \to f$ a.e. and $|f_n| \leq g \in L^1$, for all n, then $f_n \to f$ in L^1 (by the dominated convergence theorem).*

Proposition 2.31. *If $f_n \to f$ in L^1, then $f_n \to f$ in measure.*

Theorem 2.32. *Suppose that $\{f_n\}$ is Cauchy in measure. Then there is a measurable function f such that $f_n \to f$ in measure, and there is a subsequence $\{f_{n_j}\}$ that converges to f a.e. Moreover, if also $f_n \to g$ in measure, then $g = f$ a.e.*

Proposition 2.33. *If $f_n \to f$ in L^1, there is a subsequence $\{f_{n_j}\}$ such that $f_{n_j} \to f$.*

Theorem 2.34 (**Egoroff's Theorem**). *Suppose that $\mu(X) < \infty$, and $f_1, f_2, ...$ and f are measurable complex-valued functions on X such that $f_n \to f$ a.e. Then for every $\epsilon > 0$ there exists $E \subset X$ such that $\mu(E) < \epsilon$ and $f_n \to f$ uniformly on E^c.*

Proposition 2.35. *Almost uniform convergence implies convergence a.e. and convergence in measure.*

Here is a useful theorem (proved in homework 5):

Theorem 2.36 (**Lusin's Theorem**). *If $f : [a,b] \to \mathbb{C}$ is Lebesgue measurable and $\epsilon > 0$, then there is a compact set $E \subset [a,b]$ such that $\mu(E^c) < \epsilon$ and $f|_E$ is continuous.*

Below is a summarizing diagram which shows how the diffrent modes imply each other. The dashed lines show that there exists a subsequence which has the desired convergence. Furthermore, we have convergence in measure if and only if we have Cauchy in measure (Theorem 2.32).

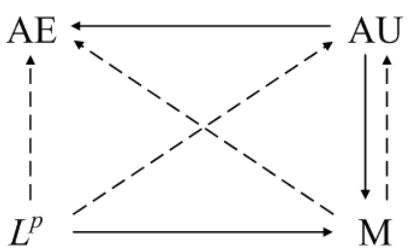

Fig. 1: Modes of Convergence in General Measure Spaces

For finite measure spaces X (i.e. $\mu(X) < \infty$), we have a few more relations; note that AE $\Rightarrow$ AU by Egoroff's Theorem.

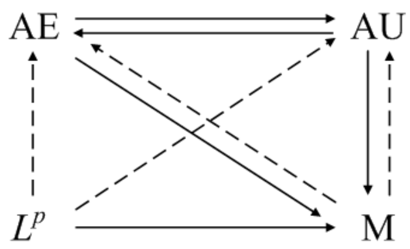

Fig. 2: Modes of Convergence in Finite Measure Spaces

2.5 Product Measures

Definition 2.37. *A measurable* ***rectangle*** *is a set of the form $A \times B$ where $A \in \mathcal{M}, B \in \mathcal{N}$. It is the case that $(A \times B) \cap (E \times F) = (A \cap E) \times (B \cap F)$ and $(A \times B)^c = (X \times B^c) \cup (A^c \times B)$. Thus, the collection of finite disjoint unions of rectangles $\mathcal{A}$ is an algebra which generates the σ-algebra $\mathcal{M} \otimes \mathcal{N}$.*

If $E \in \mathcal{A}$ is a disjoint union of rectangles $A_1 \times B_1, ..., A_n \times B_n$, we'll let π be a map such that

$$\pi(E) = \sum^{n} \mu(A_j)\nu(B_j).$$

By Theorem 1.20, π is a premeasure and generates an outer measure on $X \times Y$ whose restriction to $\mathcal{M} \times \mathcal{N}$ is a measure that extends π. We call this measure $\mu \times \nu$. If μ, ν are both σ-finite, then so is $\mu \times \nu$ and it is the unique measure such that $\mu \times \nu(A \times B) = \mu(A)\nu(B)$. We can do this for any finite number of factors and if all the measures in the factors are σ-finite, then associativity holds; i.e. $(\mathcal{M}_1 \otimes \mathcal{M}_2) \otimes \mathcal{M}_3 = \mathcal{M}_1 \otimes (\mathcal{M}_2 \otimes \mathcal{M}_3)$.

Definition 2.38. *If $E \subset X \times Y$, for $x \in X, y \in T$, we define x-**section** E_x and y-**section** E^y of E by*

$$E_x = \{y \in Y : (x, y) \in E\}, \qquad E^y = \{x \in X : (x, y) \in E\}.$$

Definition 2.39. *If f is a function on $X \times Y$, we define the x-**section** f_x and y-**section** f^y of f by*

$$f_x(y) = f^y(x) = f(x, y).$$

Definition 2.40. *A* ***monotone class*** *on a space X is a subset $\mathcal{C} \subset \mathcal{P}(X)$ that is closed under countable increasing unions and countable decreasing intersections; i.e. if $E_j \in \mathcal{C}$ and $E_1 \subset E_2 \subset ...$ then $\cup E_j \in \mathcal{C}$ and likewise for intersections.*

Every σ-algebra is a monotone class and the intersection of monotone classes is a monotone class. Then, there is a unique smallest monotone class containing $\mathcal{E} \subset \mathcal{P}(X)$.

Lemma 2.41 (The Monotone Class Lemma). *If $\mathcal{A}$ is an algebra of subsets of X, then the monotone class $\mathcal{C}$ generated by $\mathcal{A}$ coincides with the σ-algebra $\mathcal{M}$ generated by $\mathcal{A}$.*

Theorem 2.42. *Suppose $(X, \mathcal{M}, \mu), (Y, \mathcal{N}, \nu)$ are σ-finite measure spaces. If $E \in \mathcal{M} \otimes \mathcal{N}$, then the functions $x \mapsto \nu(E_x)$ and $y \mapsto \mu(E^y)$ are measurable on X and Y, respectively, and*

$$\mu \times \nu(E) = \int \nu(E_x)\, d\mu(x) = \int \mu(E^y)\, d\nu(y).$$

Theorem 2.43 (The Fubini-Tonelli Theorem). *Suppose that $(X, \mathcal{M}, \mu), (Y, \mathcal{N}, \nu)$ are σ-finite measure spaces.*

1. *(Tonelli) If $f \in L^+(X \times Y)$, then the functions $g(x) = \int f_x \, d\nu$ and $h(y) = \int f^y \, d\mu$ are in $L^+(X)$ and $L^+(Y)$, respectively and*

$$\int f \, d(\mu \times \nu) = \int \left[\int f(x,y) \, d\nu(y) \right] d\mu(x)$$
$$= \int \left[\int f(x,y) \, d\mu(x) \right] d\nu(y).$$

2. *(Fubini) If $f \in L^+(\mu \times \nu)$, then $f_x \in L^1(\nu)$ for almost every $x \in X$ and $f^y \in L^1(\mu)$ for almost every $y \in Y$. Moreover, the almost everywhere defined functions $g(x) = \int f_x \, d\nu$ and $h(y) = \int f^y \, d\mu$ are in $L^1(\mu)$ and $L^1(\nu)$, respectively and the integrals in (a) hold.*

2.6 Further Examples for Chapters 1 and 2

Example 2.44. The counting measure is a translation invariant measure on $\mathbb{R}$ that is not a multiple of Lebesgue measure.

Example 2.45. A set $E \subset \mathbb{R}$ of finite Lebesgue measure, $m(E) < \infty$ so that $m(E \cap I) > 0$ for every non-empty open interval $I \subset \mathbb{R}$.

Enumerate the rationals $\{r_n\}$ and take $E = \bigcup_{n=1}^{\infty} (r_n - 1/2^{n+1}, r_n + 1/2^{n+1})$. Then $m(E) \leq \sum_{n=1}^{\infty} 1/2^n = 1 < \infty$. But every open interval I contains a rational since $\mathbb{Q}$ is dense and thus, $I \cap E$ contains an interval so $m(I \cap E) > 0$.

Example 2.46. A Lebesgue integrable function f on $\mathbb{R}$ such that f^2 is not integrable.

$$f(x) = \begin{cases} x^{-1/2}, & 0 \leq x \leq 1 \\ 0, & \text{else} \end{cases}.$$

Then $\int f = 2$. However, $\lim_{a \to 0} \int_a^1 f^2 = \infty$.

Example 2.47. A non-measurable function f such that f^2 is measurable.

Let N be a non-measurable set. Let $f = \chi_{\mathbb{R} \setminus N} - \chi_N$. Thus, f is 1 or -1 everywhere. Since N is non-measurable, χ_N is non-measurable, forcing f to be non-measurable. However, $f^2 = 1$ everywhere; this is a continuous and thus measurable function.

Example 2.48. A Lebesgue integrable function on $\mathbb{R}$ that is unbounded on every open interval.

Enumerate the rationals as $\{r_n\}$. Let $f(x) = x^{-1/2}$ for $x \geq 0$ and 0 elsewhere. Let $g = \sum^{\infty} f(x - r_n)/2^n$. Then by Exercise 2.25 (see homework 5), $g \in L^1(m)$ but is discontinuous everywhere and unbounded on every set, even after modification on a Lebesgue null set. Furthermore, $g < \infty$ and $g^2 < \infty$ but g^2 is not integrable on any interval.

Example 2.49. Consider $\mathcal{E} = \{(-n, n) : n \in \mathbb{N}\}$. This is its own Monotone class but is not closed under complements. Thus, it is an example of a family not generated by an algebra and thus, not the same as a σ-algebra.

3 Signed Measures and Differentiation

The principal theme of this chapter is the concept of differentiating a measure ν with respect to another measure μ on the same σ-algebra.

3.1 Signed Measures

Definition 3.1. *Let $(X, \mathcal{M})$ be a measurable space. A* ***signed measure*** *on $(X, \mathcal{M})$ is a function $\nu : \mathcal{M} \to [-\infty, \infty]$ such that*

- *$\nu(\varnothing) = 0$;*
- *ν assumes at most one of the values $\pm\infty$.*
- *If $\{E_j\}$ is a sequence of disjoint sets in $\mathcal{M}$, then $\nu(\bigcup_1^\infty E_j) = \sum_1^\infty \nu(E_j)$, where the sum converges absolutely if $\nu(\bigcup_1^\infty E_j)$ is finite.*

If ν is a signed measure on $(X, \mathcal{M})$, a set $E \in \mathcal{M}$ is called **positive** (resp. **negative, null**) for ν if $\nu(F) \geq 0$ (resp. $\nu(F) \leq 0, \nu(F) = 0$) for all $F \in \mathcal{M}$ such that $F \subset E$.

Example 3.2. We will see later that the two examples below are in fact, exhaustive. Every signed measure can be represented in one of two ways.

- If μ_1, μ_2 are positive measures on $\mathcal{M}$ and at least one of them is finite, then $\nu = \mu_1 - \mu_2$ is a signed measure.
- If μ is a measure on $\mathcal{M}$ and $f : X \to [-\infty, \infty]$ is a measurable function such that a tleast one of $\int f^+ \, d\mu$ and $\int f^- \, d\mu$ is finite, then the set function $\nu(E) := \int_E f \, d\mu$ is a signed measure.

Proposition 3.3 (3.1)**.** *Let ν be a signed measure on $(X, \mathcal{M})$. If $\{E_j\}$ is an increasing sequence in $\mathcal{M}$, then $\nu(\bigcup_1^\infty E_j) = \lim_{j\to\infty} \nu(E_j)$. If $\{E_j\}$ is a decreasing sequence in $\mathcal{M}$ and $\nu(E_1)$ is finite, then $\nu(\bigcap_1^\infty E_j) = \lim_{j\to\infty} \nu(E_j)$.*

Lemma 3.4 (3.2)**.** *Any measurable subset of a positive set is positive and the union of any countable family of positive sets is positive.*

Theorem 3.5 (3.3: The Hahn Decomposition Theorem)**.** *If ν is a signed measure on $(X, \mathcal{M})$, there exist a positive set P and a negative set N for ν such that $P \cup N = X$ and $P \cap N = \varnothing$. If P', N' is another such pair, then $P \Delta P'$ (or $N \Delta N'$) is null for ν.*

The decomposition of $X = P \cup N$ as the disjoint union of a positive and negative set is called a **Hahn decomposition** for ν. It's not unique but it ldeas to a canonical representation of ν as the difference of two positive measures. To state this result, we first need a definition.

Definition 3.6. *Let μ and ν be signed measures on $(X, \mathcal{M})$. We say they are* ***mutually singular*** *or that ν is* ***singular with respect to*** *μ (or vice versa) if there exist $E, F \in \mathcal{M}$ such that $E \cap F = \varnothing$, $E \cup F = X$, E is null for μ, F is null for ν. We denote this as $\mu \perp \nu$.*

Informally, we might think of mutual singularity as μ and ν "living on disjoints sets."

Example 3.7. Consider the Dirac measure μ (point mass with respect to x_0).

$$\mu(E) = \begin{cases} 1, & x_0 \in E \\ 0, & x_0 \notin E. \end{cases}$$

μ is singular to the Lebesgue measure m if we let $E = \{x_0\}$ and $F = \mathbb{R} \setminus \{x_0\}$..

Theorem 3.8 (3.4: The Jordan Decomposition Theorem)**.** *If ν is a signed measure, there exist unique positive measures ν^+ and ν^- such that $\nu = \nu^+ - \nu^-$ and $\nu^+ \perp \nu^-$.*

The measures ν^+ and ν^- are called the **positive** and **negative variations** of ν. $\nu = \nu^+ - \nu^-$ is called the **Jordan decomposition** of ν.

Definition 3.9. *The **total variation** of ν is the measure $|\nu| := \nu^+ + \nu^-$.*

It is easy to see that $\nu(E) = 0 \Leftrightarrow |\nu|(E) = 0$ and $\nu \perp \mu \Leftrightarrow |\nu| \perp \mu \Leftrightarrow \nu^+ \perp \mu$ and $\nu^- \perp \mu$. Further observe that if ν omits the value ∞ then $\nu^+(X) = \nu(P) < \infty$ so that ν^+ is a finite measure and $|nu$ is bounded above by $\nu^+(X)$. A similar result holds when ν omits $-\infty$ as a value. If the range of ν is in $\mathbb{R}$, then ν is bounded.

Also observe that for any signed ν, we may write it as follows: $\nu(E) = \int_E f \, d\mu$ where $\mu = |\nu|$ and $f = \chi_P - \chi_N$ where $X = P \cup N$ is a Hahn decomposition for ν.

Integration with respect to a signed measure ν is defined in the following way: $L^1(\nu) = L^1(\nu^+) \cap L^1(\nu^-)$,

$$\int f \, d\nu = \int f \, d\nu^+ - \int f \, d\nu^- \quad (f \in L^1(\nu)).$$

A signed measure ν is called **finite** (resp. **σ-finite**) if $|\nu|$ is finite (resp. σ-finite).

3.2 The Lebesgue-Radon-Nikodym Theorem

Definition 3.10. *Suppose that ν is a signed measure and μ is a positive measure on $(X, \mathcal{M})$. ν is **absolutely continuous** with respect to μ if $\nu(E) = 0$ whenever $\mu(E) = 0$ ($E \in \mathcal{M}$). We denote this as $\nu \ll \mu$.*

It's easy to check that $\nu \ll \mu \Leftrightarrow |\nu| \ll \mu \Leftrightarrow \nu^+ \ll \mu$ and $\nu^- \ll \mu$. In a way, absolute continuity is the antithesis of mutual singularity. More precisely, if $\nu \perp \mu$ and $\nu \ll \mu$, then $\nu = 0$. This is because if E and F are disjoint such that $X = E \cup F$ and $\mu(E) = |\nu|(F) = 0$, then $\nu \ll \mu$ implies that $|\nu|(E) = 0$ and thus, $|\nu| = 0 \Rightarrow \nu = 0$.

Theorem 3.11 (3.5)**.** *Let ν be a finite signed measure and μ a positive measure on $(X, \mathcal{M})$. Then $\nu \ll \mu$ iff for every $\epsilon > 0$, there exists $\delta > 0$ such that $|\nu(E)| < \epsilon$ whenever $\mu(E) < \delta$.*

If μ is a measure and f is an extended μ-integrable function, the signed measure ν defined by $\nu(E) = \int_E f \, d\mu$ is absolutely continuous wrt μ. It is finite iff $f \in L^1(\mu)$. For any complex-valued $f \in L^1(\mu)$, the preceding theorem can be applied to Re f and Im f. We then get the following result:

Corollary 3.12 (3.6)**.** *If $f \in L^1(\mu)$, for every $\epsilon > 0$, ther exists $\delta > 0$ such that $|\int_E f \, d\mu| < \epsilon$ whenever $\mu(E) < \delta$.*

We use this notation to express $\nu(E) = \int_E f \, d\mu$: $d\nu = f \, d\mu$.

Lemma 3.13 (3.7)**.** *Suppose that ν and μ are finite measures on $(X, \mathcal{M})$. Either $\nu \perp \mu$, or there exist $\epsilon > 0$ and $E \in \mathcal{M}$ such that $0 < \mu(E)$ and $\nu \geq \epsilon\mu$ on E. That is, E is a positive set for $\nu - \epsilon\mu$.*

The next theorem gives a complete picture of the structure of a signed measure relative to a given positive measure.

Theorem 3.14 (3.8: The Lebesgue-Radon-Nikodym Theorem)**.** *Let ν be a σ-finite signed measure and μ a σ-finite positive measure on $(X, \mathcal{M})$. There exist unique σ-finite signed measures λ, ρ on $(X, \mathcal{M})$ such that*

$$\lambda \perp \mu, \quad \rho \ll \mu, \quad \nu = \lambda + \rho.$$

Moreover, there is an extended μ-integrable function $f : X \to \mathbb{R}$ such that $d\rho = f\,d\mu$; i.e. $\rho(E) = \int_E f\,d\mu$. Any two such functions are equal μ-almost everywhere.

The decomposition $\nu = \lambda + \rho$ where $\lambda \perp \mu$ and $\rho \ll \mu$ is called the **Lebesgue decomposition** of ν wrt μ. In the case where $\nu \ll \mu$, the LRN Theorem says that $d\nu = f\,d\mu$ for some f. f is called the **Radon-Nikodym derivative** of ν wrt μ. We denote it $d\nu/d\mu$:

$$d\nu = \frac{d\nu}{d\mu} d\mu.$$

We have a "chain rule" for these derivatives.

Proposition 3.15 (3.9)**.** *Suppose that ν is a σ-finite signed measure and μ, λ are σ-finite measures on $(X, \mathcal{M})$ such that $\nu \ll \mu$ and $\mu \ll \lambda$.*

1. *If $g \in L^1(\nu)$, then $g(d\nu/d\mu) \in L^1(\mu)$ and*

$$\int g\,d\nu = \int g \frac{d\nu}{d\mu} d\mu.$$

2. *We have $\nu \ll \lambda$ and*

$$\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu}\frac{d\mu}{d\lambda} \quad \lambda\text{-almost everywhere.}$$

Corollary 3.16 (3.10)**.** *If $\mu \ll \lambda$ and $\lambda \ll \mu$, then $(d\lambda/d\mu)(d\mu/d\lambda) = 1$ a.e. wrt either λ or μ.*

Example 3.17. This is a **nonexample**. Let m be Lebesgue measure and ν the point mass measure at 0 on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. It is clear that $\nu \perp m$. The nonexistent Radon-Nikodym derivative $d\nu/d\mu$ is popularly known as the Dirac δ-function.

Proposition 3.18. *If $\mu_1, ..., \mu_n$ are measures on $(X, \mathcal{M})$, there is a measure μ such that $\mu_j \ll \mu$ for all j—namely, $\mu = \sum_1^n \mu_j$.*

3.3 Complex Measures

Definition 3.19. *A **complex measure** on a measurable space $(X, \mathcal{M})$ is a map $\nu : \mathcal{M} \to \mathbb{C}$ such that*

- *$\nu(\varnothing) = 0$;*
- *If $\{E_j\}$ is a sequence of disjoint sets in $\mathcal{M}$, then $\nu(\bigcup^\infty E_j) = \sum^\infty \nu(E_j)$, where the series converges absolutely.*

Note that complex measures cannot take on infinite values by this definition. So positive measures are complex only if they are finite.

Example 3.20. If μ is a positive measure and $f \in L^1(\mu)$, then $\nu(E) = \int_E f\,d\mu$ is a complex measure.

If ν is a complex measure, we write ν_r and ν_i for the real and imaginary parts of ν; these are both signed measures and must be finite as ν is complex. Thus, the range of ν is a bounded subset of $\mathbb{C}$.

We define for complex measure ν, $L^1(\nu) = L^1(\nu_r) \cap L^1(\nu_i)$ and for $f \in L^1(\nu)$, $\int f\, d\nu = \int f\, d\nu_r + i \int f\, d\nu_i$. If ν, μ are complex measures, we say that $\nu \perp \mu$ if $\nu_a \perp \mu_b$ for $a, b = r, i$ and if λ is a positive measure, we say that $\nu \ll \lambda$ if $\nu_r \ll \lambda$ and $\nu_i \ll \lambda$. The other theorems in §3.2 also generalize by applying to to the real an dimaginary parts. In particular,

Theorem 3.21 (3.12: Complex Lebesgue-Radon-Nikodym Theorem)**.** *If ν is a complex measure and μ is σ-finite, positive on $(X, \mathcal{M})$, there exist a complex measure λ and an $f \in L^1(\mu)$ such that $\lambda \perp \mu$ and $d\nu = d\lambda + f\, d\mu$. If also $\tilde{\lambda} \perp \mu$ and $d\nu = d\tilde{\lambda} + \tilde{f}\, d\mu$, then $\lambda = \tilde{\lambda}$ and $f = \tilde{f}$ μ-a.e.*

As before, if $\nu \ll \mu$, we denote f by $d\nu/d\mu$. The **total variation** of a complex measure ν is the positive measure $|\nu|$ determined by the property that if $d\nu = f\, d\mu$ where μ is a positive measure, then $d|\nu| = |f|\, d\mu$. This is well defined independent of the choice of μ and f.

Proposition 3.22 (3.13)**.** *Let ν be a complex measure on $(X, \mathcal{M})$.*

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

Proposition 3.23 (3.14)**.** *If ν_1, ν_2 are complex measures on $(X, \mathcal{M})$, then $|\nu_1 + \nu_2| \leq |\nu_1| + |\nu_2|$.*

3.4 Differentiation on Euclidean Space

The Radon-Nikodym theorem provides an abstract notion of the "derivative" of a signed/complex measure ν wrt a measure μ. In this section, we consider the Lebesgue measure on Euclidean space. The **pointwise** derivative of ν wrt m can be defined in the following way:

Let $B(r, x)$ be the open ball of radius r about $x \in \mathbb{R}^n$. Consider the limit:

$$F(x) = \lim_{r \to 0} \frac{\nu(B(r, x))}{m(B(r, x))}$$

when it exists. If $\nu \ll m$, so that $d\nu = f\, dm$, then $\nu(B(r, x))/m(B(r, x))$ is simply the average value of f on $B(r, x)$ so one would hope that $F = f$ almost everywhere wrt m. This is the case provided that $\nu(B(r, x))$ is finite for all r, x. This is regarded as a generalization of the Fundamental Theorem of Calculus: the derivative of the indefinite integral of f (namely, $\nu = \int f\, dm$) is f.

Lemma 3.24 (3.15)**.** *Let $\mathcal{C}$ be a collection of open balls in $\mathbb{R}^n$ and let $U = \bigcup_{B \in \mathcal{C}} B$. If $c < m(U)$, there exist disjoint $B_1, ..., B_k \in \mathcal{C}$ such that $3^{-n} c < \sum^k m(B_j)$.*

Basically this lemma is saying that for any (possibly uncountable) collection of open balls, if the measure of their union is larger than a constant c, a finite number of disjoint balls in the collection make up a percentage of the measure on par with c, scaled by a factor dependent on the dimension of $\mathbb{R}^n$.

Definition 3.25. *A measurable function $f : \mathbb{R}^n \to \mathbb{C}$ is called **locally integrable** wrt m if $\int_K |f(x)|\, dx < \infty$ for every bounded measurable set $K \subset \mathbb{R}^n$. The space of locally integrable functions is denoted L^1_{loc}.*

Definition 3.26. *If $f \in L^1_{loc}, x \in \mathbb{R}^n$, and $r > 0$, we define the average value of f on $B(r, x)$ by*

$$A_r f(x) = \frac{1}{m(B(r,x))} \int_{B(r,x)} f(y)\, dy.$$

Lemma 3.27. *If $f \in L^1_{loc}$, $A_r f(x)$ is jointly continuous in r and x.*

Definition 3.28. *If $f \in L^1_{loc}$, the* ***Hardy-Littlewood maximal function*** *is defined as*

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

Hf is measurable because $(Hf)^{-1}((a,\infty)) = \bigcup_{r>0}(A_r|f|)^{-1}((a,\infty))$ is open for any $a \in \mathbb{R}$ since $A_r|f|(x)$ is continuous in r and x.

Theorem 3.29 (3.17: The Maximal Theorem)**.** *There is a constant $C > 0$ such that for all $f \in L^1$ and all $\alpha > 0$,*

$$m(\{x : Hf(x) > \alpha\}) \leq \frac{C}{\alpha} \int |f(x)|\, dx.$$

Theorem 3.30 (3.18)**.** *If $f \in L^1_{loc}$, then $\lim_{r\to\infty} A_r f(x) = f(x)$ for a.e. $x \in \mathbb{R}^n$.*

We might think of this in a different way: If $f \in L^1_{\text{loc}}$,

$$\lim_{r\to 0} \frac{1}{m(B(r,x))} \int_{B(r,x)} (f(y) - f(x))\, dy = 0 \text{ for a.e. } x. \tag{3.1}$$

We can say something stronger; we can have as integrand $|f(y) - f(x)|$ in Equation 3.1 and it will remain valid. That is, if we define

$$L_f = \{x : \lim_{r\to 0} \frac{1}{m(B(r,x))} \int_{B(r,x)} (f(y) - f(x))\, dy = 0\}$$

Theorem 3.31 (3.20)**.** *If $f \in L^1_{loc}$, then $m((L_f)^c) = 0$.*

We can also make these results more general by defining the limits on sets more general than balls. Take the next definition:

Definition 3.32. *A family $\{E_r\}_{r>0}$ of* ***Borel*** *subsets of $\mathbb{R}^n$ is said to* ***shrink nicely*** *to $x \in \mathbb{R}^n$ if*

- *$E_r \subset B(r, x)$ for each r;*
- *there is a constant $\alpha > 0$, independent of r, such that $m(E_r) > \alpha m(B(r, x))$.*

Note that the sets E_r need not contain x. For instance, if $U \subset B(1, 0)$ is Borel and $m(U) > 0$ and $E_r = \{x + ry : y \in U\}$, then $\{E_r\}$ shrinks nicely to x.

Theorem 3.33 (3.21: The Lebesgue Differentiation Theorem)**.** *Suppose that $f \in L^1_{loc}$. For every x in the Lebesgue set of f —in particular, for almost every x — we have*

$$\lim_{r\to 0} \frac{1}{m(E_r)} \int_{E_r} |f(y) - f(x)|\, dy = 0 \text{ and } \lim_{r\to 0} \frac{1}{m(E_r)} \int_{E_r} f(y)\, dy = f(x)$$

for every family $\{E_r\}_{r>0}$ that shrinks nicely to x.

Definition 3.34. *A Borel measure ν on $\mathbb{R}^n$ will be called* ***regular*** *if*

1. $\nu(K) < \infty$ *for every compact K;*

2. $\nu(E) = \inf\{\nu(U) : U \text{ open}, E \subset U\}$ *for every $E \in \mathcal{B}_{\mathbb{R}^n}$.*

Condition 2 is actually implied by condition 1 but we include that here. Signed or complex Borel measures are regular if their total variation are regular.

Example 3.35. If $f \in L^+(\mathbb{R}^n)$, the measure $f\,dm$ is regular iff $f \in L^1_{\text{loc}}$.

Theorem 3.36 (3.22)**.** *Let ν be a regular signed or complex Borel measure on $\mathbb{R}^n$ and let $d\nu = d\lambda + f\,dm$ be its Lebesgue-Radon-Nikodym representation. Then for m-almost every $x \in \mathbb{R}^n$,*

$$\lim_{r \to 0} \frac{\nu(E_r)}{m(E_r)} = f(x)$$

for every family $\{E_r\}_{r>0}$ that shrinks nicely to x.

3.5 Functions of Bounded Variation

Theorem 3.37 (3.23)**.** *Let $F : \mathbb{R} \to \mathbb{R}$ be increasing and let $G(x) = F(x+)$.*

1. *The set of points at which F is discontinuous is countable.*

2. *F and G are differentiable a.e. and $F' = G'$ a.e.*

Definition 3.38. *If $F : \mathbb{R} \to \mathbb{C}$ and $x \in \mathbb{R}$, define*

$$T_F(x) = \sup\left\{\sum^n |F(x_j) - F(x_{j-1})| : n \in \mathbb{N}, -\infty < x_0 < ... < x_n = x\right\}.$$

T_F is called the ***total variation function*** *of F.*

The idea behind this definition is that of a particle traveling along an interval. The total variation is described by a function F as the total distance travelled. If F is smooth, we can simply integrate F. If not, we need to approximate variation on small subintervals and pass to a limit. It follows that

$$T_F(b) - T_F(a) = \sup\left\{\sum^n |F(x_j) - F(x_{j-1})| : n \in \mathbb{N}, a < x_0 < ... < x_n = b\right\}.$$

T_F is an increasing function with values in $[0, \infty]$. If $T_F(\infty) = \lim_{x\to\infty} T_F(x)$ is finite, we say that F is of **bounded variation** on $\mathbb{R}$ and denote the space of all such F by BV.

The supremum on the RHS of the equation with $T_F(b) - T_F(a)$ is called the **total variation** of F on $[a, b]$. It depends only on the values of F on $[a, b]$ so we may define $BV[a, b]$ to be the set of all functions on $[a, b]$ whose total variation on $[a, b]$ is finite. If $F \in BV$, then its restriction to $[a, b]$ is in $BV[a, b]$ for all a, b. The total variation of the restriction is simply $T_F(b) - T_F(a)$. Conversely, if $F \in BV[a, b]$, then we may set $F(x) = F(a)$ for $x < a$ and $F(x) = F(b)$ for $x > b$. Then $F \in BV$. Thus, results for BV can be applied to $BV[a, b]$.

Example 3.39.

1. If $F : \mathbb{R} \to \mathbb{R}$ is bounded and increasing, then $F \in BV$ and $T_F(x) = F(x) - F(-\infty)$.

2. If $F, G \in BV$ and $a, b \in \mathbb{C}$, then $aF + bG \in BV$.

3. If F is differentiable on $\mathbb{R}$ and F' is bounded, then $F \in BV[a,b]$ for $-\infty < a < b < \infty$ (by the mean value theorem).

4. If $F(x) = \sin x$ then $F \in BV[a,b]$ but $F \notin BV$.

5. If $F(x) = x\sin(1/x)$ for $x \neq 0$ and $F(0) = 0$, then $F \notin BV[a,b]$ if $0 \in [a,b]$.

Lemma 3.40. *If $F \in BV$ is real-valued, then $T_F + F$ and $T_F - F$ are increasing.*

Theorem 3.41.

1. *$F \in BV$ iff Re $F \in BV$ and Im $F \in BV$.*

2. *If $F : \mathbb{R} \to \mathbb{R}$, then $F \in BV$ iff F is the difference of two bounded increasing functions; for $F \in BV$ these functions may be taken to be $\frac{1}{2}(T_F + F)$ and $\frac{1}{2}(T_F - F)$.*

3. *If $F \in BV$, 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 $F(\pm\infty) = \lim_{y \to \pm} F(y)$.*

4. *If $F \in BV$ the set of points at which F is discontinuous is countable.*

5. *If $F \in BV$ and $G(x) = F(x+)$, then F' and G' exist and are equal a.e.*

The representation $F = \frac{1}{2}(T_F + F) - \frac{1}{2}(T_F - F)$ of a real-valued $F \in BV$ is called the **Jordan decomposition** of F. Since $x^+ = \max(x,0) = \frac{1}{2}(|x| + x)$ and $x^- = \max(-x,0) = \frac{1}{2}(|x| - x)$ for $x \in \mathbb{R}$, we have

$$\frac{1}{2}(T_F \pm F)(x) = \sup\left\{\sum^n (F(x_j) - F(x_{j-1}))^{\pm} : x_0 < ... < x_n = x\right\} \pm \frac{1}{2}F(-\infty).$$

Definition 3.42. *The space of **normalized bounded variation** is defined as*

$$NBV = \{F \in BV : F \text{ is right continuous and } F(-\infty) = 0\}.$$

We observe that if $F \in BV$, then the functions G defined by $G(x) = F(x+) - F(-\infty)$ is in NBV and $G' = F'$ a.e.

Lemma 3.43 (3.28)**.** *If $F \in BV$, then $T_f(-\infty) = 0$. If F is also right continuous, then so is T_F.*

Theorem 3.44 (3.29)**.** *If μ is a complex Borel measure on $\mathbb{R}$ and $F(x) = \mu((-\infty, x])$, then $F \in NBV$. Conversely, if $F \in NBV$, there is a unique complex Borel measure μ_F such that $F(x) = \mu_F((-\infty, x])$; moreover, $\mu_F| = \mu_{T_F}$.*

A question to ask is what functions in NBV correspond to measures μ such that $\mu \perp m$ or $\mu \ll m$?

Proposition 3.45 (3.30)**.** *If $F \in NBV$, then $F' \in L^1(m)$. Moreover, $\mu_F \perp m$ iff $F' = 0$ a.e. and $\mu_F \ll m$ iff $F(x) = \int_{-\infty}^x F'(t)\,dt$.*

The condition $\mu_F \ll m$ can also be expressed directly in terms of F, as follows. A function $F : \mathbb{R} \to \mathbb{C}$ is called **absolutely continuous** if for every $\epsilon > 0$ there exists $\delta > 0$ such that for *any* finite set of disjoint intervals $(a_1, b_1), ..., (a_N, b_N)$,

$$\sum_1^n (b_j - a_j) < \delta \Longrightarrow \sum_1^N |F(b_j) - F(a_j)| < \epsilon.$$

F is **absolutely continuous** on $[a,b]$ if this condition is satisfied whenever the intervals (a_j, b_j) are all in $[a,b]$. Taking $N = 1$ shows that F is *uniformly continuous*. On the other hand, if F is everywhere differentiable and F' is bounded, then F is absolutely continuous, for $|F(b_j) - F(a_j)| \leq (\max |F'|)(b_j - a_j)$ by the Mean Value Theorem.

Proposition 3.46 (3.32)**.** *If $F \in NBV$, then F is absolutely continuous iff $\mu_F \ll m$.*

Corollary 3.47 (3.33)**.** *If $f \in L^1(m)$, then the function $F(x) = \int_{-\infty}^x f(t)\,dt$ is in NBV and is absolutely continuous, and $f = F'$ a.e. Conversely, if $F \in NBV$ is absolutely continuous, then $F' \in L^1(m)$ and $F(x) = \int_{-\infty}^x F'(t)\,dt$.*

If we consider functions on bounded intervals, this result can be refined.

Lemma 3.48 (3.34)**.** *If F is absolutely continuous on $[a,b]$, then $F \in BV[a,b]$.*

Theorem 3.49 (3.35: The Fundamental Theorem of Calculus for Lebesgue Integrals)**.** *If $-\infty < a < b < \infty$ and $F : [a,b] \to \mathbb{C}$, the following are equivalent:*

1. *F is absolutely continuous on $[a,b]$.*
2. *$F(x) - F(a) = \int_a^x f(t)\,dt$ for some $f \in L^1([a,b], m)$.*
3. *F is differentiable a.e. on $[a,b]$, $F' \in L^1([a,b], m)$, and $F(x) - F(a) = \int_a^x F'(t)\,dt$.*

The following decomposition of Borel measures on $\mathbb{R}^n$ is sometimes important.

Definition 3.50. *A complex Borel measure μ on $\mathbb{R}^n$ is called **discrete** if there is a countable set $\{x_j\} \subset \mathbb{R}^n$ and complex numbers c_j such that $\sum |c_j| < \infty$ and $\mu = \sum c_j \delta_{x_j}$, where δ_x is the point mass at x.*

Definition 3.51. *μ is called **continuous** if $\mu(\{x\}) = 0$ for all $x \in \mathbb{R}^n$.*

Any complex measure μ can be written uniquely as $\mu = \mu_d + \mu_c$ where μ_d is discrete and μ_c is continuous. Indeed, let $E = \{x : \mu(\{x\}) \neq 0\}$. Then $\mu_d(A) = \mu(A \cap E)$ is discrete and $\mu_c(A) = \mu(A \setminus E)$ is continuous.

If μ is discrete, then $\mu \perp m$ and if $\mu \ll m$ then μ is continuous. Then, any regular complex Borel measure μ on $\mathbb{R}^n$ can be written *uniquely* as

$$\mu = \mu_d + \mu_{ac} + \mu_{sc}$$

where μ_d is discrete, μ_{ac} is absolutely continuous wrt to m and μ_{sc} is continuous and singular to m; i.e. $\mu_{sc} \perp m$.

The existence of nonzero singular continuous measures in $\mathbb{R}^n$ is not too tricky when $n > 2$. When $n = 1$, they correspond to nonconstant functions $F \in NBV$ such that F is continuous but $F' = 0$ a.e. An example of such a function is the Cantor function. There even exist strictly increasing functions F such that $F' = 0$ a.e.

Example 3.52. [3.5.40] Let F denote the cantor function on $[0,1]$ and set $F(x) = 0$ for $x < 0$ and $F(x) = 1$ for $x > 1$. Let $\{[a_n, b_n]\}$ be an enumeration of the closed subintervals of $[0,1]$ with rational endpoints and let $F_n(x) = F((x - a_n)/(b_n - a_n))$. Then $G = \sum_1^\infty 2^{-n} F_n$ is continuous and strictly increasing on $[0,1]$, and $G' = 0$ a.e.

If $F \in NBV$, it is customary to denote the integral of a function g wrt the measure μ_F by $\int g\, dF$ or $\int g(x)\, dF(x)$. Such integrals are called **Lebesgue-Stieltjes integrals**. The next theorem is an integration-by-parts formula for such integrals.

Theorem 3.53 (3.36)**.** *If F and G are in NBV and at least one of them is continuous, then for $-\infty < a < b < \infty$,*

$$\int_{(a,b]} F\, dG + \int_{(a,b]} G\, dF = F(b)G(b) - F(a)G(a).$$

4 Point Set Topology

4.1 Basics about Sets

Definition 4.1. *The **interior** of a set A is the union of all open sets contained in A. The **closure** of A is the intersection of all closed sets containing A.*

Definition 4.2. *A set $A \subset X$ is called **dense** if its closure is X. **nowhere dense** if the interior of the closure is empty.*

Proposition 4.3 (4.5)**.** *Every second countable space is separable; i.e. if X has a countable basis, then there exists a countable dense subset in X.*

Proposition 4.4 (4.6)**.** *If X is first countable (has local countable basis) and $A \subset X$, then $x \in \overline{A}$ iff there is a sequence $\{x_j\}$ in A that converges to x.*

4.2 Continuous Maps

Proposition 4.5 (4.9)**.** *If the topology on Y is generated by a family of sets $\mathcal{E}$, then $f : X \to Y$ is continuous iff $f^{-1}(V)$ is open in X for every $V \in \mathcal{E}$.*

Definition 4.6. *The space of all bounded real-(resp. complex-) valued functions on X is denoted $B(X, \mathbb{R})$ (resp. $B(X, \mathbb{C})$). We will usually not include the target space in notation. The space of all bounded continuous functions is $BC(X) = B(X) \cap C(X)$. The uniform norm of $f \in B(X)$ is $\|f\|_u = \sup\{|f(x)| : x \in X\}$.*

Proposition 4.7 (4.13)**.** *If X is a topological space, $BC(X)$ is a closed subspace of $B(X)$ in the uniform metric; in particular, $BC(X)$ is complete.*

Theorem 4.8 (4.15: Urysohn's Lemma)**.** *Let X be a normal space. If A and B are disjoint closed sets in X, there exists $f \in C(X, [0, 1])$ such that $f = 0$ on A and $f = 1$ on B.*

Theorem 4.9 (4.16: The Tietze Extension Theorem)**.** *Let X be a normal space. If A is a closed subset of X and $f \in C(A, [a, b])$, there exists $F \in C(X, [a, b])$ such that $F|_A = f$.*

Theorem 4.10 (4.58: The Urysohn Metrization Theorem)**.** *Every second countable regular space is metrizable.*

4.3 Nets

Definition 4.11. *A **directed set** is a set A with a binary relation $\preceq$ such that*

- *$a \preceq a$ for all $a \in A$;*
- *if $a \preceq b$ and $b \preceq c$, then $a \preceq c$;*

- *for any $a, b \in A$, there exists $c \in A$ such that $a \preceq c$ and $b \preceq c$.*

Note that this is different from partially ordered sets.

Definition 4.12. *A **net** in a set X is a mapping $a \mapsto x_a$ from a directed set A into X. Nets are denoted by $\langle x \rangle_{a\in A}$. These generalize sequences in which $A = \mathbb{N}$.*

Definition 4.13. *A net $\langle x_a \rangle_{a\in A}$ is **eventually** in a set E if there exists $a_0 \in A$ such that for every $a \succeq a_0$, $x_a \in E$. $\langle x_a \rangle_{a\in A}$ **converges** to x if for every neighborhood U of x, $\langle x_a \rangle_{a\in A}$ is eventually in U.*

Definition 4.14. *A net $\langle x_a \rangle_{a\in A}$ is **frequently** in a set E if for every $a \in A$, there exists $b \succeq a$ such that $x_b \in E$. x is a **cluster point** of $\langle x_a \rangle_{a\in A}$ if for every neighborhood U of x, $\langle x_a \rangle_{a\in A}$ is frequently in U.*

4.4 Compact Spaces

Definition 4.15. *X is **sequentially compact** if every sequence in X has a convergent subsequence.*

Definition 4.16. *A family $\{F_a\}_{a\in A}$ of subsets of X is said to have the **finite intersection property** if $\bigcap_{a\in B} F_a \neq \varnothing$ for all finite $B \subset A$.*

Proposition 4.17 (4.21)**.** *A topological space X is compact iff for every family $\{F_a\}_{a\in A}$ of closed sets with the finite intersection property, $\bigcap_{a\in A} F_a$.*

Proposition 4.18 (4.22)**.** *A closed subset of a compact space is compact.*

Proposition 4.19 (4.24)**.** *Every compact subset of a Hausdorff space is closed.*

Proposition 4.20 (4.25)**.** *Every compact Hausdorff space is normal.*

Proposition 4.21 (4.28)**.** *If X is compact and Y is Hausdorff, then any continuous bijection $f : X \to Y$ is a homeomorphism.*

Theorem 4.22 (4.29)**.** *If X is a topological space, the following are equivalent:*

1. *X is compact.*
2. *Every net in X has a cluster point.*
3. *Every net in X has a convergent subnet.*

4.5 Locally Compact Hausdorff

Definition 4.23. *X is **σ-compact** if it is a countable union of compact sets.*

4.6 Two Compactness Theorems

Theorem 4.24 (4.42: Tychonoff's Theorem)**.** *If $\{X_a\}_{a\in A}$ is any family of compact topological spaces, then $X = \prod_{a\in A} X_a$ (with the product topology) is compact.*

Definition 4.25. *If $\mathcal{F} \subset C(X)$, $\mathcal{F}$ is called **equicontinuous at** $x \in X$ if for every $\epsilon > 0$, there is a neighborhood U of x such that $|f(y) - f(x)| < \epsilon$ for all $y \in U$ and all $f \in \mathcal{F}$.*

Definition 4.26. *$\mathcal{F} \subset C(X)$ is called **pointwise bounded** if $\{f(x) : f \in \mathcal{F}\}$ is a bounded subset of $\mathbb{C}$ for each $x \in X$.*

Theorem 4.27 (4.43: Arzelà-Ascoli Theorem I)**.** *Let X be a compact Hausdorrf space. If $\mathcal{F}$ is an equicontinuous, pointwise bounded subset of $C(X)$, then $\mathcal{F}$ is totally bounded in the uniform metric; i.e. for every $\epsilon > 0$, $\mathcal{F}$ can be covered by a finite number of ϵ-balls. Moreover, the closure of $\mathcal{F}$ is compact.*

Theorem 4.28 (4.44: Arzelà-Ascoli Theorem II)**.** *Let X be σ-compact and locally compact Hausdorff. If $\{f_n\}$ is an equicontinuous, pointwise bounded sequence in $C(X)$, there exist $f \in C(X)$ and a subsequence of $\{f_n\}$ that converges to f uniformly on compact sets.*

4.7 The Stone-Weierstrass Theorem

Definition 4.29. *A subset $\mathcal{A} \subset C(X)$ is said to **separate points** if for every distinct $x, y \in X$, there exists $f \in \mathcal{A}$ such that $f(x) \neq f(y)$.*

Definition 4.30. *A subset $\mathcal{A} \subset C(X)$ is called an **algebra** if it is a real (resp. complex) vector space of $C(X, \mathbb{R})$ (resp. $C(X, \mathbb{C})$) such that $fg \in \mathcal{A}$ whenever $f, g \in \mathcal{A}$.*

Definition 4.31. *A subset $\mathcal{A} \subset C(X)$ is called an **lattice** if $\max(f, g)$ and $\min(f, g)$ are in $\mathcal{A}$ whenever $f, g \in \mathcal{A}$.*

Theorem 4.32 (4.45: The Stone-Weierstrass Theorem)**.** *Let X be a compact Hausdorff space. If $\mathcal{A}$ is a closed subalgebra of $C(X, \mathbb{R})$ that separates points, then either $\mathcal{A} = C(X, \mathbb{R})$ or $\mathcal{A} = \{f \in C(X, \mathbb{R}) : f(x_0) = 0\}$ for some $x_0 \in X$. The first alternative holds iff $\mathcal{A}$ contains the constant functions.*

Theorem 4.33 (4.51: The Complex Stone-Weierstrass Theorem)**.** *Let X be a compact Hausdorff space. If $\mathcal{A}$ is a closed complex subalgebra of $C(X)$ that separates points and is closed under complex conjugation, then either $\mathcal{A} = C(X)$ or $\mathcal{A} = \{f \in C(X) : f(x_0) = 0\}$ for some $x_0 \in X$.*

5 Elements of Functional Analysis

Functional analysis is the study of infinite-dimensional vector spaces over $\mathbb{R}$ or $\mathbb{C}$ and the linear maps between them. Topology is what distinguishes the finite and infinite case. In finite vector spaces, there is usually only one reasonable topology and linear maps are automatically continuous. The infinite case is trickier.

5.1 Normed Vector Spaces

Definition 5.1. *A **seminorm** on X is a function $x \mapsto \|x\|$ from X to $[0, \infty)$ such that*

- *$\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$ (the **triangle inequality**),*
- *$\|\lambda x\| = |\lambda| \|x\|$ for all $x \in X$ and $\lambda \in K$.*

*If $\|x\| = 0$ only when $x = 0$, then $\| \cdot \|$ is a **norm**.*

Definition 5.2. *A normed vector space has a norm topology. Two norms $\| \cdot \|_1, \| \cdot \|_2$ are called **equivalent** if there exists $C_1, C_2 > 0$ such that*

$$C_1 \|x\|_1 \leq \|x\|_2 \leq C_2 \|x\|_1, \quad \forall x \in X.$$

If every Cauchy sequence converges in a normed space, the space is called a complete normed vector space or **Banach**.

Theorem 5.3 (5.1)**.** *A normed vector space X is complete iff every absolutely convergent series in X converges; i.e. given that $\sum^{\infty} \|x_n\|$ converges, $\sum^{\infty} x_n$ converges.*

If X, Y are normed space, $X \times Y$ may be given a **product norm** $\|(x,y)\| = \max(\|x\|, \|y\|)$. Other equivalent norms include $\|x\| + \|y\|$ or $(\|x\|^2 + \|y\|^2)^{1/2}$.

If M is a subspace of X, the **quotient space** is defined by modding out by the relation $x \sim y$ iff $x - y \in M$. X/M inherits the **quotient norm** $\|x + M\| = \inf_{y \in M} \|x + y\|$.

Definition 5.4. *A linear map $T : X \to Y$ is called **bounded** if there exists $C \geq 0$ such that $\|Tx\| \leq C\|x\|$ for all $x \in X$.*

Proposition 5.5 (5.2)**.** *If X and Y are normed vector spaces and $T : X \to Y$ is linear, the following are equivalent:*

1. *T is continuous.*
2. *T is continuous at 0.*
3. *T is bounded.*

Definition 5.6. *If X, Y are normed, the space of all **bounded linear maps** from X to Y is denoted $L(X, Y)$ and is given the norm:*

$$\begin{aligned} \|T\| &= \sup\{\|Tx\| : \|x\| = 1\} \\ &= \sup\left\{\frac{\|Tx\|}{\|x\|} : x \neq 0\right\} \\ &= \inf\{C : \|Tx\| \leq C\|x\|, x \in X\} \end{aligned}$$

Proposition 5.7 (5.4)**.** *If Y is complete, so is $L(X, Y)$.*

If $T \in L(X, Y), S \in L(Y, Z)$, then $\|STx\| \leq \|S\|\|Tx\| \leq \|S\|\|T\|\|x\|$. Thus, $ST \in L(X, Z)$.

Definition 5.8. *If $T \in L(X, Y)$ T is said to be **invertible** or an **isomorphism** if T is bijective and T^{-1} is bounded (i.e. $\|Tx\| \geq C\|x\|$ for some $x \in X$). T is an **isometry** of $\|Tx\| = \|x\|$ for all $x \in X$. An isometry is injective but not necessarily surjective; it is an isomorphism onto its image.*

5.2 Linear Functionals

Let X be a vector space over $K = \mathbb{R}$ or $\mathbb{C}$. Then if X is normed, $L(X, K)$ is called the space of **bounded linear functionals** on X or also called the **dual space** of X and is usually denoted X^*. X^* is complete since K is complete.

Proposition 5.9 (5.5)**.** *Let X be a vector space over $\mathbb{C}$. If f is a complex linear functional on X and $u = Re\ f$, then u is a real linear functional and $f(x) = u(x) - iu(ix)$ for all $x \in X$. COnversely, if u is a real linear functional on X and $f : X \to \mathbb{C}$ is defined by $f(x) = u(x) - iu(ix)$, then f is complex linear. In this case, if X is normed, we have $\|u\| = \|f\|$.*

Definition 5.10. *If X is a real vector space, a **sublinear functional** on X is a map $p : X \to \mathbb{R}$ such that*

$$p(x + y) \leq p(x) + p(y), p(\lambda x) = \lambda p(x), \quad \forall x, y \in X, \lambda \geq 0.$$

Theorem 5.11 (5.6: The Hahn-Banach Theorem)**.** *Let X be a real vector space, p a sublinear functional on X, M a subspace of X, and f a linear functional on M such that $f(x) \leq p(x)$ for all $x \in M$. Then there exists a linear functional F on X such that $F(x) \leq p(x)$ for all $x \in X$ and $F|_M = f$.*

Theorem 5.12 (5.7: The Complex Hahn-Banach Theorem)**.** *Let X be a complex vector space, p a seminorm on X, M a subspace of X, and f a complex linear functional on M such that $|f(x)| \leq p(x)$ for all $x \in M$. Then there exists a complex linear functional F on X such that $|F(x)| \leq p(x)$ for all $x \in X$ and $F|_M = f$.*

Theorem 5.13 (5.8)**.** *Let X be a normed vector space.*

1. *If M is a closed subspace of X and $x \in X \setminus M$, there exists $f \in X^*$ such that $f(x) \neq 0$ and $f|_M = 0$. In fact, if $\delta = \inf_{y \in M} \|x - y\|$.*
2. *If $x \neq 0$, there exists $f \in X^*$ such that $\|f\| = 1$ and $f(x) = \|x\|$.*
3. *The bounded linear functionals on X separate points.*
4. *If $x \in X$, define $\hat{x} : X^* \to \mathbb{C}$ by $\hat{x}(f) = f(x)$. Then the map $x \mapsto \hat{x}$ is a linear isometry from X into X^{**}.*

Let $\widehat{X} = \{\hat{x} : x \in X\}$. Since X^{**} is always complete, the closure $\overline{\widehat{X}}$ of $\widehat{X}$ is Banach and the map $x \mapsto \hat{x}$ embeds X into $\overline{\widehat{X}}$ as a dense subspace. $\overline{\widehat{X}}$ is called the **completion** of X. In particular, if X is Banach, then $\overline{\widehat{X}} = \widehat{X}$.

If X is finite dimensional, $\widehat{X} = X^{**}$ since they have the same dimension. In infinite dimensions, this might not hold. If $\widehat{X} = X^{**}$, X is called **reflexive**. Since we may identify $\hat{x}$ with x, reflexivity means $X^{**} = X$.

5.3 The Baire Category Theorem and its Consequences

Theorem 5.14 (5.9: The Baire Category Theorem)**.** *Let C be a complete* ***metric*** *space.*

1. *If $\{U_n\}^\infty$ is a sequence of open dense subsets of X, then $\bigcap^\infty U_n$ is dense in X.*
2. *X is not a countable union of nowhere dense sets.*

Since the Baire Category Theorem is a toplogical statement, we may apply it to spaces homeomorphic to complete metric spaces. For example, $(0, 1)$ is not complete but is homeomorphic to $\mathbb{R}$ which is complete.

A set which is a countable union of nowhere dense sets is called **meager** or of the **first category**. The complement of a meager set is called residual. Nonmeager sets are of the **second category**.

Theorem 5.15 (5.10: The Open Mapping Theorem)**.** *Let X, Y be Banach spaces. If $T \in L(X, Y)$ is surjective, then T is open.*

Corollary 5.16 (5.11)**.** *If X, Y are Banach spaces and $T \in L(X, Y)$ is bijective, then T is an isomorphism; that is $T^{-1} \in L(Y, X)$.*

Definition 5.17. *The* ***graph*** *of T is defined as*

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

$\Gamma(T)$ is a subspace of $X \times Y$. We say that T ***closed*** *if $\Gamma(T)$ is closed in $X \times Y$.*

If T is continuous, then T is closed. The next theorem gives conditions for the converse to hold.

Theorem 5.18 (5.12: The Closed Graph Theorem)**.** *If X, Y are Banach spaces and $T : X \to Y$ is a closed linear map, then T is bounded.*

Continuity of a linear map $T : X \to Y$ means that if $x_n \to x$, then $Tx_n \to Tx$ wherease closedness means that if $x_n \to x$ and $Tx_n \to y$, then $y = Tx$. Thus, the Closed Graph Theorem says that Tx_n does converge to something and we only need to check that it converges to Tx.

Theorem 5.19 (5.13: The Uniform Boundedness Principle)**.** *Suppose that X, Y are normed vector spaces and $\mathcal{A} \subset L(X, Y)$.*

1. *If $\sup_{T\in\mathcal{A}} \|Tx\| < \infty$ for all x in some nonmeager subset of X, then $\sup_{T\in\mathcal{A}} \|T\| < \infty$.*
2. *If X is a Banach space and $\sup_{T\in\mathcal{A}} \|Tx\| < \infty$ for all $x \in X$, then $\sup_{T\in\mathcal{A}} \|T\| < \infty$.*

5.4 Topological Vector Spaces

Definition 5.20. *A* ***topological vector space*** *is a vector space endowed with a topology such that $(x, y) \mapsto x + y$ and $(\lambda, x) \mapsto \lambda x$ are continuous maps.*

Definition 5.21. *A topological vector space is called* ***locally convex*** *if there is a basis for the topology consisting of convex sets; i.e. sets A such that if $x, y \in A$, then $tx + (1 - t)y \in A$ for $t \in (0, 1)$.*

Another way to think of convex sets is that for any point not in the set, one can separate the point and the set with a hyperplane. On $\mathbb{R}^2$, every norm gives a convex, balanced (contains x and $-x$), unit ball. Thus, there is a 1-1 correspondence between convex, balanced sets containing 0 and norms on $\mathbb{R}^2$. However, inner products correspond to ellipses.

The next theorem shows how we can generate "balls" for a topology without a norm.

Theorem 5.22 (5.14)**.** *Let $\{p_a\}_{a\in A}$ be a family of seminorms on the vector space X. If $x \in X$, $a \in A$, and $\epsilon > 0$, let*

$$U_{xa\epsilon} = \{y \in X : p_a(y - x) < \epsilon\},$$

and let $\mathcal{T}$ be the topology generated by the sets $U_{xa\epsilon}$.

1. *For each $x \in X$, the finite intersections of the sets $U_{xa\epsilon}$ form a neighborhood basis at x.*
2. *If $\langle x_i \rangle_{i\in I}$ is a net in X, then $x_i \to x$ iff $p_a(x_i - x) \to 0$ for all $a \in A$.*
3. *$(X, \mathcal{T})$ is a locally convex topological vector space.*

Proposition 5.23 (5.15)**.** *Suppose X, Y are vector spaces with topologies defined, respectively, by the families $\{p_a\}_{a\in A}$ and $\{q_b\}_{b\in B}$ of seminorms, and $T : X \to Y$ is a linear map. Then T is continuous iff for each $b \in B$, there exist $a_1, ..., a_l \in A$ and $C > 0$ such that $q_b(Tx) \le C \sum^k p_{a_j}(x)$.*

Proposition 5.24 (5.16)**.** *Let X be a vector space equipped with the topology defined by a family $\{p_a\}_{a\in A}$ of seminorms.*

1. *X is Hausdorff iff for each $x \neq 0$ there exists $a \in A$ such that $p_a(x) \neq 0$.*
2. *If X is Hausdorff and A is countable, then X is metrizable with a translation-invariant metric; i.e. $\rho(x, y) = \rho(x + z, y + z)$ for al $x, y, z \in X$.*

If X has the topology defined by a family of seminorms, the Hahn-Banach theorem guarantees the existence of lots of continuous linear functionals on X — enough to separate points if X is Hausdorff.

Definition 5.25. *A complete Hausdorff topolgical vector space whose topology is defined by a countable family of seminorms is called a **Fréchet space**.*

It is a result that no norm on $C^\infty[0,1]$ makes $\frac{d}{dx}$ a bounded operator. Simply consider $f_n(x) = x^n$. However, the differential operator can be bounded if we define a topology with seminorms.

Definition 5.26. *The **weak topology** on X is the coarsest topology on X such that every $f \in X^*$ is continuous. A net $\langle x_a \rangle$ in X converges weakly to x iff $f(x_a) \to f(x)$ for all $f \in X^*$.*

Definition 5.27. *The **weak*-topology** is the topology generated by $X \hookrightarrow X^{**}$. It is weaker than the weak topology on X^* and is the topology of pointwise convergence: $f_n \to f$ iff $\hat{x}(f_n) = f_n(x) \to f(x) = \hat{x}(f)$.*

Definition 5.28. *Let X, Y be Banach. The topology on $L(X,Y)$ generated by the evaluation maps $T \mapsto Tx$ ($x \in X$) is called the **strong operator topology** on $L(X,Y)$ and the topology generated by the linear functionals $T \mapsto f(Tx)$ ($x \in X, f \in Y^*$) is called the **weak operator topology** on $L(X,Y)$.*

We may understand these topologies in terms of convergence: $T_\alpha \to T$ strongly iff $T_\alpha x \mapsto Tx$ in the norm topology of Y for each $x \in X$ whereas $T_\alpha \to T$ weakly iff $T_\alpha x \to Tx$ in the weak topology of Y for each $x \in X$. Thus, the strong operator topology is finer than the weak operator topology but coarser than the norm topology on $L(X,Y)$.

Proposition 5.29 (5.17)**.** *Suppose $\{T_n\}^\infty \subset L(X,Y)$, $\sup_n \|T_n\| < \infty$ and $T \in L(X,Y)$. If $\|T_n x - Tx\| \to 0$ for all x in a dense subset D of X, then $T_n \to T$ strongly.*

Theorem 5.30 (5.18: Alaoglu's Theorem)**.** *If X is a normed vector space, the closed unit ball $B^* = \{f \in X^* : \|f\| \leq 1\}$ in X^* is compact in the weak* topology.*

5.5 Hilbert Spaces

Definition 5.31. *An **inner product** on a complex vector space H is a map $(x,y) \mapsto \langle x,y \rangle$ from $H \times H \to \mathbb{C}$ satisfying:*

1. *$\langle ax + by, z \rangle = a\langle x,z \rangle + b\langle y,z \rangle$ for all $x,y,z \in H$, $a,b \in \mathbb{C}$.*
2. *$\langle y,x \rangle = \overline{\langle x,y \rangle}$.*
3. *$\langle x,x \rangle \in (0,\infty)$ for all nonzero $x \in H$.*

It follows that $\langle x, ay + bz \rangle = \bar{a}\langle x,y \rangle + \bar{b}\langle x,z \rangle$ for all $x,y,z \in H$ and $a,b \in \mathbb{C}$.

Definition 5.32. *A complex vector space with an inner product is called a **pre-Hilbert space**. We define $\|x\| = \sqrt{\langle x,x \rangle}$ to be the norm of x. If H is complete wrt $\|\cdot\|$, then H is called a **Hilbert space**.*

Theorem 5.33 (5.19: The Cauchy-Schwarz Inequality)**.** *$|\langle x,y \rangle| \leq \|x\|\|y\|$ for all $x,y \in H$ with equality iff x,y are linearly dependent.*

From here, we assume H is a Hilbert space.

Proposition 5.34 (5.21)**.** *If $x_n \to x$ and $y_n \to y$, then $\langle x_n, y_n \rangle \to \langle x,y \rangle$.*

Theorem 5.35 (5.22: The Parallelogram Law)**.** *For all $x,y \in H$, $\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$.*

If a norm does not satisfy the parallelogram law, then it is not induced by an inner product.

Definition 5.36. *If $x, y \in H$ and $\langle x, y\rangle = 0$, we write $x \perp y$ and say x is **orthogonal** to y. If $E \subset H$, we define $E^\perp = \{x \in H : \langle x, y\rangle y = 0, \quad \forall y \in E\}$.*

$E^\perp$ is a closed subspace of H.

Theorem 5.37 (5.23: The Pythagorean Theorem)**.** *If $x_1, ..., x_n \in H$ and $x_j \perp x_k$ for $j \neq k$,*

$$\left\| \sum_1^n x_j \right\|^2 = \sum_1^n \|x_j\|^2$$

Theorem 5.38 (5.24)**.** *If M is a closed subspace of H, then $H = M \oplus M^\perp$; i.e. every $x \in H$ can be uniquely expressed as $x = y + z$ where $y \in M, z \in M^\perp$. Moreover, y and z are the unique elements of M and $M^\perp$ whose distance to x is minimal.*

Theorem 5.39 (5.25: The Riesz Representation Theorem)**.** *If H is a Hilbert space and $f \in H^*$, there is a unique $y \in H$ such that $f(x) = \langle x, y\rangle$ for all $x \in H$.*

If H is not complete, then this doesn't hold. The Riesz Representation Theorem shows that Hilbert spaces are reflexive in a strong sense: not only is H naturally isomorphic to H^{**}, it is naturally isomorphic to H^*.

The **Gram-Schmidt process** converts a linearly independent sequence $\{x_n\}^\infty$ into an orthonormal sequence.

Theorem 5.40 (5.26: Bessel's Inequality)**.** *If $\{u_a\}_{a\in A}$ is an orthonormal set in H, then for any $x \in H$,*

$$\sum_{a\in A} |\langle x, u_a\rangle|^2 \leq \|x\|^2.$$

In particular, $\{a : \langle x, u_a\rangle \neq 0\}$ is countable.

Theorem 5.41 (5.27)**.** *If $\{u_a\}_{a\in A}$ is an orthonormal set in H, the following are equivalent:*

1. *(Completeness) If $\langle x, u_a\rangle = 0$ for all a, then $x = 0$.*
2. *(Parseval's Identity) $\|x\|^2 = \sum_{a\in A} |\langle x, u_a\rangle|^2$ for all $x \in H$.*
3. *For each $x \in H$, $x = \sum_{a\in A}\langle x, u_a\rangle u_a$ where the sum on the right has only countable many nonzero terms and converges in the norm topology no matter how these terms are ordered.*

An orthonormal set having the properties of Theorem 5.41 is called an **orthonormal basis** for H.

Proposition 5.42 (5.28)**.** *Every Hilbert space has an orthonormal basis.*

The proof uses Zorn's Lemma.

Proposition 5.43 (5.29)**.** *A Hilbert space H is separable iff it has a countable orthonormal basis, in which case, every orthonormal basis for H is countable.*

Example 5.44. $\mathbb{R}$ considered as a vector space over $\mathbb{Q}$ has an inner product and is complete with respect to the induced norm. Also, $\mathbb{Q}$ is a countable dense subset of $\mathbb{R}$; i.e $\mathbb{R}$ is separable. However, $\mathbb{R}$ does not have a countable orthonormal basis. For it to have an orthonormal basis, it can only have one basis element; namely ± 1. But this is not enough to generate all of $\mathbb{R}$ and in fact, $\mathbb{R}$ does not even have a countable basis let alone an orthonormal one. This is because $\mathbb{R}$ over $\mathbb{Q}$ is not a Hilbert space; Hilbert spaces are over $\mathbb{R}$ or $\mathbb{C}$.

Definition 5.45. *Let H_1, H_2 be Hilbert spaces with inner products $\langle \cdot, \cdot \rangle_1, \langle \cdot, \cdot \rangle_2$, respectively. A **unitary map** $U : H_1 \to H_2$ is an invertible linear map that preserves inner products: $\langle Ux, Uy \rangle_1 = \langle x, y \rangle_1$ for all $x, y \in H_1$.*

Every unitary map is an isometry; i.e. $\|U_x\| = \|x\|$. Conversely, every surjective isometry is unitary. Thus, unitary maps are the true "isomorphisms" in the category of Hilbert spaces. They preserve linear structure, topology, norms, and inner products. From the point of view of abstract structure, every Hilbert space looks like an ℓ^2 space and every finite subspace of a Hilbert space looks Euclidean.

Definition 5.46. *Let A be any nonempty set. $\ell^2(A)$ is the set of functions $f : A \to \mathbb{C}$ such that the sum $\sum_{a \in A} |f(a)|^2$ is finite. This is defined to be the supremum of its finite partial sums.*

Proposition 5.47 (5.30)**.** *Let $\{e_a\}_{a \in A}$ be an orthonormal basis for X. Then the correspondence $x \mapsto \hat{x}$ defined by $\hat{x}(a) = \langle x, u_a \rangle$ is a unitary map from H to $\ell^2(A)$.*

6 Tricks and Methods

In this section, we mention some useful methods for solving problems in analysis.

- Instead of showing $x \leq y$ directly, it is usually a lot easier to show that, for all $\epsilon > 0$, $x \leq y + \epsilon$. To show $x = y$, perhaps show that $x \leq y \leq x + \epsilon$ for all ϵ.
- Work with "discretized" $\epsilon = 1/n$ and let $n \to \infty$.
- Enumerate the rationals $\{r_n\}$ and build functions f_n. One may want to consider their sum: $f = \sum^{\infty} f_n$. Another thing to remember is that every interval in $\mathbb{R}$, no matter the size, contains a rational.
- Splitting a set into disjoint pieces can be useful such as $E = (E \setminus F) \cup (F \cap E)$. Also, taking a sequence of sets $\{E_j\}^{\infty}$ and creating a disjoint sequence whose union is still the same:

$$F_k = E_k \setminus \left[\bigcup^{k-1} E_j \right] = E_k \cap \left[\bigcup^{k-1} E_j \right]^c .$$

  $\bigcup^{\infty} F_k = \bigcup^{\infty} E_j$.

- Buldings sets $\{A_n\}$ and considering $A = \bigcup^{\infty} A_n$ or $\bigcap^{\infty} A_n$.
- Let $\epsilon > 0$ and consider $\epsilon/2^n$. At some point, take $\sum^{\infty} \epsilon/2^n = \epsilon$.
- Cantor Set, Cantor function.