Notes on Real Analysis

Class notes for Math 6320 - 6321
Theory of Functions of a Real Variable
Fall 2010 - Spring 2011
University of Houston

Ricky Ng

Contents

Chapter 1

Abstract Integration

1.1 σ-algebras

The classical Riemann integral relies on continuity, which is a topological property. One of our main objectives is to generalize Riemann integral. Hence, it is essential to enlarge the class of integrable functions from merely continuous functions. In analogy to a topology, we want to construct an underlying space of subsets on which more functions can be integrated similarly. We will study this space and the corresponding functions from section 1.1 to 1.4.

Throughout this chapter, let X be a nonempty set and $\mathcal{P}(X) = \{S \colon S \subseteq X\}$ denote its power set.

Definition 1.1.1 (Topology). Let $I \neq \varnothing$. A **topology** $\tau \subseteq \mathcal{P}(X)$ of X satifies:

i. The empty set $\varnothing$ and $X \in \tau$.

ii. If $E_i \in \tau$, for all $i \in I$, then $\bigcup_{i\in I} E_i \in \tau$.

iii. If $E_i \in \tau$, for $1 \leq i \leq n$, then $\bigcap_{i=1}^{n} E_i \in \tau$.

Definition 1.1.2 (Basis for a topology)**.** A collection $B \subseteq \mathcal{P}(X)$ is called a **basis** for X satisfying:

i. For all $x \in X$, there is $E \in B$ such that $x \in E$.

ii. For all $x \in X$ with $x \in E_1 \cap E_2$, there is $E_3 \in B$, such that $x \in E_3 \subseteq E_1 \cap E_2$.

Proposition 1.1.3. *Let B be a basis for X. Then, $\tau := \{U \colon U = \bigcup_{E\in B} E\}$ is a topology. We call such τ the topology generated by B.*

Proof. Verify the axioms. ■

Definition 1.1.4 ($\boldsymbol{\sigma}$-algebra). A collection $M \subseteq \mathcal{P}(X)$ is called a **$\boldsymbol{\sigma}$-algebra** if it satisfies the following:

i. The set $X \in M$.

ii. If $A \in M$, then $X \setminus A \in M$.

iii. If $A_n \in M$, for each $n \in \mathbb{N}$, then $\bigcup_{n=1}^{\infty} A_n \in M$.

The pair (X, M) is called a **measurable space**. Any set $E \in M$ is called a **measurable set**.

REMARK. Note that M is "closed" under any finite and countable union, intersections, and set difference.

EXAMPLE 1.1.5. Let $X \neq \varnothing$.

- $\mathcal{P}(X)$, $\{\varnothing, X\}$ are σ-algebras.
- If X is uncountable, then $M := \{E \in \mathcal{P}(X) \colon \text{either } E \text{ or } E^c \text{ is countable}\}$ is a σ-algebra.

1.2 Measurable Maps and Borel Algebras

In this section, we want to study the relations between topological spaces and measurable spaces. To be precise, we will write the set X with its topology τ and σ-algebra M when necessary.

PROPOSITION 1.2.1 (σ-algebra generated by a collection)**.** *Suppose $F \subseteq \mathcal{P}(X)$. There exists a unique σ-algebra $M(F) \subseteq \mathcal{P}(X)$, such that*

i. $F \subseteq M(F)$

ii. If $N \supseteq F$ is a σ-algebra, then $N \supseteq M(F)$.

We call $M(F)$ the σ-algebra generated by F.

Proof. Define $\Omega := \{N \subseteq \mathcal{P}(X) \colon N \supset F,\ N \text{ is a } \sigma\text{-algebra}\}$. Define $M(F) := \bigcap_{N \in \Omega} N$. Verify the definition. For the uniqueness, show $M(F) \subseteq M'(F)$ and $M(F) \supseteq M'(F)$. ■

DEFINITION 1.2.2 (Borel σ-algebra). Let (X, τ) be a topological space. The σ-algebra generated by τ, $M(\tau)$, is called the **Borel σ-algebra**. Any set $B \in M(F)$ is an **Borel set**.

DEFINITION 1.2.3 (Measurable maps). Let (X, M) be a measurable space, (Y, τ) be a topological space. A map $f \colon (X, M) \to (Y, \tau)$ is **measurable** if for each open set U, $f^{-1}(U)$ is measurable in X.

DEFINITION 1.2.4 (Borel measurable maps). Let $(X, \tau_X), (Y, \tau_Y)$ be topological spaces. Then $f \colon (X, M(\tau_X)) \to (Y, \tau_Y)$ is **Borel measurable** if for each open set U in Y, $f^{-1}(U)$ is a Borel set in X. Hence, all continuous maps are Borel measurable.

Theorem 1.2.5. *The map $f\colon (X, M) \to (Y, \tau)$ is measurable if and only if for each Borel set B in Y, $f^{-1}(B)$ is measurable in X.*

Proof. ($\Leftarrow$) is trivial. ($\Rightarrow$) Define $N := \{U \subset Y : f^{-1}(U) \in M\}$. Show that N is a σ-algebra containing τ. By definition of Borel algebra, $N \supset M(\tau)$. ∎

Remark 1.2.6. From the definitions and Theorem 1.2.5., we immediately see that if $f\colon (X, M) \to (Y, \tau_Y)$ is measurable and $g\colon (Y, \tau_Y) \to (Z, \tau_Z)$ is Borel measurable, then $g \circ f\colon X \to Z$ is measurable.

Proposition 1.2.7. *Let $f\colon (X, M) \to (Y, \tau_Y)$ be measurable, $g\colon (Y, \tau_Y) \to (Z, \tau_Z)$ be continuous. Then $g \circ f\colon X \to Z$ is measurable.*

Proof. For $U \in M(\tau_Z)$, $g^{-1}(U) \in M(\tau_Y)$, so $f^{-1}(g^{-1}(U)) = (g \circ f)^{-1}(U) \in M$. ∎

Lemma 1.2.8. *Let $V \subseteq \mathbb{R}^2$ be open. There exists sequence of open rectangles $\{R_i\}_{i=1}^{\infty}$, $R_i = (a_i, b_i) \times (c_i, d_i)$, such that $V = \bigcup_{i\in\mathbb{N}} R_i$.*

Proof. $B := \{(a, b) \times (c, d) \colon a, b, c, d \in \mathbb{Q}\}$ is a countable basis for $\mathbb{R}^2$. ∎

Proposition 1.2.9. *Let $u, v\colon X \to \mathbb{R}$ be measurable. Then the map $f\colon X \to \mathbb{R}^2$, given by $f(x) := (u(x), v(x))$ is measurable.*

Proof. Let $R := (a, b) \times (c, d)$ be an open rectangle. Then,

$$f^{-1}(R) = \{x\colon u(x) \in (a, b)\} \cap \{x\colon v(x) \in (c, d)\} \in M.$$

Let $V \in \mathbb{R}^2$ be open. By Lemma (1.2.8), $V = \bigcup_{i=1}^{\infty} R_i$, where R_i is an open rectangle. Hence,

$$f^{-1}(V) = f^{-1}\left(\bigcup_{i=1}^{\infty}(R_i)\right) = \bigcup_{i=1}^{\infty} f^{-1}(R_i) \in M.$$ ∎

Corollary 1.2.10.

1. *If $f = u + iv$, where u, v are real measurable maps on X, then f is complex measurable on X.*
2. *If $f = u + iv$ is complex measurable on X, then u, v, and $|f|$ are real measurable on X.*
3. *If f, g are complex measurable on X, so are $f + g$ and fg.*

Definition 1.2.11 (Extended real line). In Analysis, we frequently deal with ∞, sequences and compactness. To generalize our results, we will work with the **extended real line**. From Topology, it is a compactification of $\mathbb{R}$. Roughly speaking, we add the symbols "$-\infty$" and "∞" to $\mathbb{R}$ and enlarge the standard topology on $\mathbb{R}$ by allowing sets in the form of $[-\infty, b), (a, \infty]$ to be open. Hence, a standard basis for this topology is

$$B = \{[-\infty, b)\} \cup \{(a, b)\} \cup \{(a, \infty]\}. \tag{1.2.1}$$

Under this topology, $\mathbb{R}$ is compact. We sometime write $[-\infty, \infty]$ for $\mathbb{R}$.

Theorem 1.2.12 (Test for measurability)**.** *Let $f\colon (X, M) \to [-\infty, \infty]$. Then, f is measurable if and only for all $a \in \mathbb{R}$, $f^{-1}((a, \infty]) \in M$.*

Proof. ($\Leftarrow$) is trivial by the definition of meaurable function. ($\Rightarrow$) First show that any basis set is measurable using $[-\infty, b) = \bigcup_{n=1}^{\infty}[-\infty, b - \frac{1}{n}]$ and intersection. Then, for every open $V \in [-\infty, \infty]$, V is either in $(-\infty, \infty)$ or not. If former, done previously. Otherwise, V is the union of a set containining $[-\infty, b)$ and/or $(a, \infty]$ with some $V_0 \subset (-\infty, \infty)$, hence measurable. ■

1.3 Sequences

Definition 1.3.1. Given $\{a_n\}_{n=1}^{\infty}$ in $[-\infty, \infty]$. Define

$$\limsup a_n := \inf_{k\in\mathbb{N}}\{\sup_{k\geq n} a_n\}, \qquad \text{and} \qquad \liminf a_n := \sup_{k\in\mathbb{N}}\{\inf_{k\geq n} a_n\}. \tag{1.3.1}$$

We call $\limsup a_n$ the **upper limit** and $\liminf a_n$ the **lower limit** of $\{a_n\}$, respectively.

Remark 1.3.2. Since $\{\sup_{k\geq n} a_n\}_{n=1}^{\infty}$ is a decreasing sequence in a compact set, $\limsup a_n$ always exists. Similarly, $\liminf a_n$ always exists. Also,

$$\liminf_{n\to\infty}(a_n) = -\limsup_{n\to\infty}(-a_n).$$

If $\{a_n\}$ converges in $[-\infty, \infty]$, then

$$\liminf a_n = \lim a_n = \limsup a_n.$$

Definition 1.3.3. Let $f_n\colon X \to [-\infty, \infty]$, for all $n \in \mathbb{N}$. Define the following functions pointwise:

- $(\sup f_n)(x) := \sup_{n\in\mathbb{N}}\{f_n(x)\}$.
- $(\limsup f_n)(x) := \limsup_{n\to\infty}\{f_n(x)\}$.

We define $\inf f_n$, $\liminf f_n$ similarly. Also, if for all $x \in X$, $\{f_n(x)\}$ converges, we call $f(x) := \lim_{n\to\infty} f_n(x)$ the pointwise limit of the sequence $\{f_n\}$.

Proposition 1.3.4. *Let (X, M) be measurable space, $f_n\colon X \to [-\infty, \infty]$ be measurable for each $n \in \mathbb{N}$. Then,*

$$g := \sup f_n \qquad \textit{and} \qquad h := \limsup f_n$$

are measurable.

Proof. Show that $g^{-1}((a, \infty]) = \bigcup_{n=1}^{\infty} f_n^{-1}((a, \infty]) \in M$. Same for $\inf f_n$, then apply on h. ■

Corollary 1.3.5. *There following corollaries are useful in the later chapters.*

1. *If $f_n \colon X \to \mathbb{C}$ and f is the pointwise convergent limit of f_n for all $x \in X$, then f is complex measurable.*
2. *If $f, g \colon X \to [-\infty, \infty]$ are measurable, then so are $\max\{f, g\}$ and $\min\{f, g\}$. In parictular, it is true for*

$$f^+ := \max\{f, 0\} \qquad \textit{and} \qquad f^- := -\min\{f, 0\}.$$

1.4 Simple Functions

Definition 1.4.1. A function $s \colon X \to \mathbb{C}$ is called a **simple function** if $s(X)$ is a finite set.

Definition 1.4.2. Let $A \in M$, define the **characteristic function** $\chi_A \colon X \to \mathbb{C}$ given by

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{else.} \end{cases} \tag{1.4.1}$$

Remark. Note that χ_A is measurable because $\chi_A^{-1}(U) = X, A^c, A$, or $\varnothing$.

Proposition 1.4.3. *A function $s \colon X \to \mathbb{C}$ is simple if and only if there are disjoint measurable sets $A_1, \dots, A_n$, and $\alpha_1, \dots, \alpha_n \in \mathbb{C}$, such that $s = \sum_{i=1}^n \alpha_i \chi_{A_i}$.*

Proof. ($\Rightarrow$) follows from definition. ($\Leftarrow$). Let $\{\alpha_1, \dots, \alpha_n\} = s(X)$. Define $A_i := s^{-1}(\alpha_i)$, for each i. Verify the claim. ∎

Theorem 1.4.4. *Let $f \colon X \to [0, \infty]$ be measurable, then there exists a sequence of simple functions $\{s_n\}_{n=1}^\infty$, $s_n(x) \geq 0$ such that*

i. *For each $n \in \mathbb{N}$, s_n is measurable.*

ii. *The sequence $\{s_n(x)\}$ is non-decreasing, for all $x \in X$.*

iii. *For all $x \in X$, $s_n(x) \to f(x)$.*

Proof. Fix $t \in [0, \infty)$. For each n, there is $k_n(t) \in \mathbb{N}_0$, such that $k_n(t) 2^{-n} \leq t < (k_n(t) + 1) 2^{-n}$.[1] Define a **staircase function** $\varphi_n \colon [0, \infty) \to [0, \infty)$ given by

$$\varphi_n(t) := \begin{cases} \frac{k_n(t)}{2^n}, & \text{if } 0 \leq t < n, \\ n, & \text{if } t \geq n. \end{cases} \tag{1.4.2}$$

Note that $\varphi_n(t) \nearrow t$, and $0 \leq \varphi_1 \leq \cdots \leq t$. Also,

$$\varphi_n^{-1}((a, \infty]) = \begin{cases} \varnothing, & \text{if } n < a, \\ [\frac{m+1}{2^n}, \infty], & \text{if } \frac{m}{2^n} \leq a < \frac{m+1}{2^n}, \text{ for some } m \in \mathbb{N}. \end{cases}$$

[1] The set $\{\frac{m}{2^n} : m, n \in \mathbb{N}\}$ is dense in $\mathbb{R}$.

Hence, φ_n is Borel measurable. Define $s_n := \varphi_n \circ f$, then $s_n \leq f$. Note that s_n's are simple, measurable, and $s_n \nearrow f$. ■

1.5 Measures

Definition 1.5.1. Let (X, M) be a measurable space. A set function $\mu \colon M \to [0, \infty]$ is called **countably additive**, or **σ-addivity**, if whenever $\{A_n\}_{n=1}^{\infty} \subset M$, with $A_i \cap A_j = \varnothing$, for all $i \neq j$, we have

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

Moreover, if $\mu(A) < \infty$ for some $A \in M$, then μ is called a **positive measure**. The triple (X, M, μ) is called a **measure space**.

Example 1.5.2. Let $X = \mathbb{N}$, $M := \mathcal{P}(X)$. Define $\mu(S) = |S|$. Such μ is known as the **counting measure**.

Theorem 1.5.3 (Elementary Properties of Positive Measures)**.** *Let (X, M, μ) be a positive measure space. Then*

(a) $\mu(\varnothing) = 0$.

(b) For finite disjoint collection $\{A_i\}_{i=1}^{n}$, $\mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i)$.

(c) If $A, B \in M$, *and* $A \subseteq B$, *then* $\mu(A) \leq \mu(B)$.

(d) If $A_1 \subseteq A_2 \subset \dots$ *in* M, *then* $\mu(\bigcup_{i=1}^{\infty} A_i) = \lim_{n\to\infty} \mu(A_n)$.

(e) If $A_1 \supseteq A_2 \supseteq \dots$ *in* M, *and there is* $k \in \mathbb{N}$, *so that* $\mu(A_k) < \infty$, *then* $\mu(\bigcap_{i=1}^{\infty} A_i) = \lim_{n\to\infty} \mu(A_n)$.

We call properties (b) ***finite addivitiy*** *and (c)* ***monotonicty***.

Proof. Follow from definition. For (d), take $B_n = A_n \setminus A_{n-1}$. For (e), similar. ■

1.6 Integration of Positive Functions

From now on, (X, M, μ) denotes a positive measure space.

Definition 1.6.1. Let $s \colon X \to [0, \infty]$ be a simple measurable function, of the form

$$s = \sum_{i=1}^{n} \alpha_i \chi_{A_i},$$

where $s(X) = \{\alpha_1, \ldots, \alpha_n\}$ and $A_i \cap A_j = \varnothing$, for all $i \neq j$. If $E \in M$, we define the integral of s over E by

$$\int_E s \,\mathrm{d}\mu := \sum_{i=1}^{n} \alpha_i \mu(A_i \cap E),$$

with the convention $0 \cdot \infty = 0$.

Remark. By definition, if $0 \leq t \leq s$ is a simple measurable function over a measurable set $E \in M$, we see that

$$\int_E t \,\mathrm{d}\mu \leq \int_E s \,\mathrm{d}\mu.$$

Definition 1.6.2 (Lebesgue integral). Let $f\colon X \to [0, \infty]$ be a measurable function and $E \in M$, we define the **Lebesgue integral** of f over E

$$\int_E f \,\mathrm{d}\mu := \sup \left\{ \int_E s \,\mathrm{d}\mu : s \text{ is simple, measurable with } 0 \leq s \leq f \right\}$$

Note that if f is simple, the definitions of $\int_E f \,\mathrm{d}\mu$ agree.

Theorem 1.6.3 (Properties of Lebesgue Integral). *Let $f, g\colon X \to [0, \infty]$ be measurable.*

(a) If $0 \leq f \leq g$, then $\int_E f \,\mathrm{d}\mu \leq \int_E g \,\mathrm{d}\mu$.

(b) If $A \subseteq B$ and $A, B \in M$, then $\int_A f \,\mathrm{d}\mu \leq \int_B g \,\mathrm{d}\mu$.

(c) Given $c \in [0, \infty), E \in M$, $\int_E cf \,\mathrm{d}\mu = c \int_E f \,\mathrm{d}\mu$.

(d) If $f(E) = 0$, or $\mu(E) = 0$, then $\int_E f \,\mathrm{d}\mu = 0$.

(e) For all $E \in M$, $\int_E f \,\mathrm{d}\mu = \int_X f \chi_E \,\mathrm{d}\mu$.

Proof. One should be able to prove these properties by the definitions. In general, when we need to prove the result for measurable f, we first prove that for simple functions. It is a lot easier because simple functions are just finite linear combinations. Then generalize the case for measurable function f. ■

We now come to some remarkable results on limits and sequences.

Lemma 1.6.4. *Let $s\colon X \to [0, \infty]$ be a simple measurable function. The function $\varphi\colon M \to [0, \infty]$ given by*

$$\varphi(E) := \int_E s \,\mathrm{d}\mu,$$

is a positive measure.

Proof. It is easy to verify the definition of positive measure since s only takes finitely many values on X. ■

Theorem 1.6.5 (**Lebesgue's Monotone Convergence Theorem**)**.** *Let $\{f_n\}$ be a sequence of measurable functions on $X \to [0, \infty]$, with*

(a) $0 \le f_1 \le f_2 \le \cdots \le \infty$, *and*

(b) $f_n(x) \to f(x)$, *for all* $x \in X$.

Then f is measurable, and

$$\lim_{n\to\infty} \int_X f_n \,\mathrm{d}\mu = \int_X f \,\mathrm{d}\mu. \tag{1.6.1}$$

Proof. Obviously, $\lim_{n\to\infty} \int_X f_n \,\mathrm{d}\mu \le \int_X f \,\mathrm{d}\mu$. For another inequality, for every $c \in (0, 1)$, fix a simple measurable function s, with $0 \le s \le f$. Let $E_n := \{x : f_n(x) \ge cs(x)\}$. Then, $E_1 \subseteq E_2 \subseteq \ldots$ and $\bigcup_{n=1}^\infty E_n = X$. By Lemma (1.6.4),

$$\begin{aligned}\int_X s \,\mathrm{d}\mu = \varphi(X) = \lim_{n\to\infty} \varphi(E_n) &= \lim_{n\to\infty} \int_{E_n} s \,\mathrm{d}\mu \\ &\le \lim_{n\to\infty} \int_{E_n} \frac{1}{c} f_n \,\mathrm{d}\mu \\ &\le \lim_{n\to\infty} \int_X \frac{1}{c} f_n \,\mathrm{d}\mu. \end{aligned} \tag{1.6.2}$$

Bying taking the supremum over all such s on the LHS of inequality (1.6.2),

$$\lim_{n\to\infty} \int_X f_n \,\mathrm{d}\mu \ge c \int_X f \,\mathrm{d}\mu, \quad \forall c \in (0, 1). \tag{1.6.3}$$

Thus, (1.6.3) also holds true when $c = 1$ and it completes the proof. ■

Corollary 1.6.6. *Let $f, g\colon X \to [0, \infty]$ be measurable, then for all $E \in M$,*

$$\int_E (f + g) \,\mathrm{d}\mu = \int_E f \,\mathrm{d}\mu + \int_E g \,\mathrm{d}\mu.$$

Proof. First show for simple functions using the lemma. Pick increasing sequences of simple functions, $s_n \nearrow f$, $t_n \nearrow g$ and apply Monotone Convergence. ■

Corollary 1.6.7 (**Monotone Convergence for Series**)**.** *If $f_n\colon X \to [0, \infty]$ is measurable for each $n \in \mathbb{N}$, then for all $E \in M$,*

$$\int_E \sum_{n=1}^\infty f_n \,\mathrm{d}\mu = \sum_{n=1}^\infty \int_E f_n \,\mathrm{d}\mu. \tag{1.6.4}$$

Proof. Define $g_n := f_1 + \cdots + f_n$ and apply Monotone Convergence on $\{g_n\}$. ■

Theorem 1.6.8 (Integral to measures)**.** *Let $f\colon X \to [0, \infty]$ be measurable. Then the function $\varphi\colon M \to [0, \infty]$ given by*

$$\varphi(E) := \int_E f \,\mathrm{d}\mu,$$

is a positive measure. Moreover, if $g\colon X \to [0,\infty]$ is measurable, then

$$\int_X g\,\mathrm{d}\varphi = \int_X gf\,\mathrm{d}\mu.$$

Proof. This is a general version of Lemma (1.6.4). Show that φ defines a measure using Monotone Convergence. For the second part, start with simple function, then use Monotone Convergence on $s_n \nearrow g$. ■

Theorem 1.6.9 (**Fatou's Lemma**)**.** *Let $f_n : X \to [0,\infty]$ be measurable for every $n \in \mathbb{N}$, then*

$$\int_X \left(\liminf_{n\to\infty} f_n\right)\mathrm{d}\mu \le \liminf_{n\to\infty} \int_X f_n\,\mathrm{d}\mu. \tag{1.6.5}$$

Proof. Let $g_n(x) := \inf\{f_k(x) : n \ge k\}$. So, $g_n \le f_m$ for all $m \ge n$, and $\{g_n\}$ is a non-decreasing sequence. By Monotone Convergence,

$$\begin{aligned}\int_X \liminf_{n\to\infty} f_n\,\mathrm{d}\mu &= \int_X \lim_{n\to\infty} g_n\,\mathrm{d}\mu = \lim_{n\to\infty}\int_X g_n\,\mathrm{d}\mu \\ &\le \lim_{n\to\infty}\left(\inf_{m\ge n}\left\{\int_X f_m\,\mathrm{d}\mu\right\}\right) = \liminf_{n\to\infty}\int_X f_n\,\mathrm{d}\mu.\end{aligned}$$ ■

1.7 Integration of Complex Functions

Definition 1.7.1 (L^1 **Space**)**.** Define $L^1(\mu)$ to be the set of all complex measurable functions $f\colon X \to \mathbb{C}$ such that

$$\int_X |f|\,\mathrm{d}\mu < \infty. \tag{1.7.1}$$

Recall that the measurability of f implies that of $|f|$. Any $f \in L^1(\mu)$ is called **Lebesgue integrable**, **absolutely integrable**, or ***L*¹-integrable**.

Definition 1.7.2 (Integral of Complex Functions)**.** Let $f \in L^1(\mu)$ and $f = u + iv$. For all $E \in M$, we define the integral of f to be

$$\int_E f\,\mathrm{d}\mu := \left(\int_E u^+\,\mathrm{d}\mu - \int_E u^-\,\mathrm{d}\mu\right) + i\left(\int_E v^+\,\mathrm{d}\mu - \int_E v^-\,\mathrm{d}\mu\right). \tag{1.7.2}$$

Note that $u^+, u^- \le |u| \le |f|$ and $v^+, v^- \le |v| \le |f|$. Thus, each of the integals above is finite. Also, if $f \in L^1(\mu)$ and $f\colon X \to [-\infty,\infty]$, $E \in M$, we write

$$\int_E f\,\mathrm{d}\mu = \int_E f^+\,\mathrm{d}\mu - \int_E f^-\,\mathrm{d}\mu. \tag{1.7.3}$$

Theorem 1.7.3. *For all $f, g \in L^1(\mu)$, for all $\alpha, \beta \in \mathbb{C}$, $\alpha f + \beta g \in L^1(\mu)$, and*

$$\int_X (\alpha f + \beta g)\,\mathrm{d}\mu = \alpha\int_X f\,\mathrm{d}\mu + \beta\int_X g\,\mathrm{d}\mu.$$

Proof. First, $\alpha f + \beta g$ is measurable. Also,

$$\int_X |\alpha f + \beta g| \,\mathrm{d}\mu \leq \int_X |\alpha||f| + |\beta||g| \,\mathrm{d}\mu = |\alpha| \int_X |f| \,\mathrm{d}\mu + |\beta| \int_X |g| \,\mathrm{d}\mu < \infty.$$

So $\alpha f + \beta g \in L^1(\mu)$. Now consider $h := f + g$. We see that

$$h^+ + f^- + g^- = h^- + f^+ + g^+,$$

which implies $\int_X f + g \,\mathrm{d}\mu = \int_X f \,\mathrm{d}\mu + \int_X g \,\mathrm{d}\mu$. Finally, if $\alpha = -1$, then use $(-u)^+ = u^-$. If $\alpha = i$, we have

$$\int_X if \,\mathrm{d}\mu = \int_X (-v) \,\mathrm{d}\mu + i \int_X u = i\Big(\int_X u \,\mathrm{d}\mu + i \int_X v \,\mathrm{d}\mu \Big) = i \int_X f \,\mathrm{d}\mu.$$

Therefore, $L^1(\mu)$ is a **complex vector space**. ■

Proposition 1.7.4. *If $f \in L^1(\mu)$, then $|\int_X f \,\mathrm{d}\mu| \leq \int_X |f| \,\mathrm{d}\mu$.*

Proof. First show that for real-valued f,

$$\begin{aligned} \left| \int_X f \,\mathrm{d}\mu \right| &= \left| \int_X f^+ \,\mathrm{d}\mu - \int_X f^- \,\mathrm{d}\mu \right| \leq \left| \int_X f^+ \,\mathrm{d}\mu \right| + \left| \int_X f^- \,\mathrm{d}\mu \right| \\ &= \int_X (f^+ + f^-) \,\mathrm{d}\mu = \int_X |f| \,\mathrm{d}\mu. \end{aligned}$$

For complex-valued f, suppose $z := \int_X f \,\mathrm{d}\mu \neq 0$. Then, there is $\alpha \in \mathbb{C}$ such that $\alpha z = |z|$. Hence,

$$\begin{aligned} \left| \int_X f \,\mathrm{d}\mu \right| &= \alpha \int_X f \,\mathrm{d}\mu = \int_X \alpha f \,\mathrm{d}\mu = \int_X \mathrm{Re}(\alpha f) \,\mathrm{d}\mu \\ &\leq \int_X |\alpha f| \,\mathrm{d}\mu = \int_X |f| \,\mathrm{d}\mu, \end{aligned}$$

where the third equality comes from the fact that $|\int_X f \,\mathrm{d}\mu|$ is real-valued. ■

Theorem 1.7.5 (Dominated Convergence Theorem). *For all $n \in \mathbb{N}$, let $f_n : X \to \mathbb{C}$ be measurable. Suppose*

$$f(x) := \lim_{n\to\infty} f_n(x)$$

exists for all $x \in X$ and there exists $g \in L^1(\mu)$ such that $|f_n(x)| \leq g(x)$, for all $x \in X$ and $n \in \mathbb{N}$. Then $f \in L^1(\mu)$,

$$\lim_{n\to\infty} \int_X |f_n - f| \,\mathrm{d}\mu = 0, \qquad \textit{and} \qquad \lim_{n\to\infty} \int_X f_n \,\mathrm{d}\mu = \int_X f \,\mathrm{d}\mu. \tag{1.7.4}$$

Proof. First, since $|f_n| \le g$, $|f| \le g$ and $f \in L^1(\mu)$. Applying Fatou's Lemma on $2g - |f_n - f|$, we see that

$$\begin{aligned}\int_X 2g\,\mathrm{d}\mu &\le \liminf_{n\to\infty} \int_X 2g - |f_n - f|\,\mathrm{d}\mu \\ &= \int_X 2g\,\mathrm{d}\mu + \liminf_{n\to\infty}\Big(-\int_X |f_n - f|\,\mathrm{d}\mu\Big) \\ &= \int_X 2g\,\mathrm{d}\mu - \limsup_{n\to\infty}\int_X |f_n - f|\,\mathrm{d}\mu. \qquad (1.7.5)\end{aligned}$$

Because $\int_X 2g\,\mathrm{d}\mu < \infty$, inequality (1.7.5) gives

$$\limsup_{n\to\infty}\int_X |f_n - f|\,\mathrm{d}\mu \le 0. \qquad (1.7.6)$$

However, $\int_X |f_n - f|\,\mathrm{d}\mu \ge 0$ forces the upper limit to be 0. On other hand, the lower limit is at least 0. Therefore,

$$0 \le \lim_{n\to\infty}\left|\int_X f_n\,\mathrm{d}\mu - \int_X f\,\mathrm{d}\mu\right| \le \limsup_{n\to\infty}\int_X |f_n - f|\,\mathrm{d}\mu = 0. \qquad \blacksquare$$

1.8 Sets of Measure Zero

Definition 1.8.1. Let (X, M, μ) be a measure space, and $E \in M$. We say a property P holds **almost everywhere** on E with respect to μ if there exists $N \subset E$, $N \in M$ such that P holds on $E \setminus N$ and $\mu(N) = 0$. We denote that by "a.e.", "μ-a.e.", or "a.e. $[\mu]$".

Example 1.8.2. If f, g are measurable functions on X, and if

$$\mu(\{x : f(x) \neq g(x)\}) = 0,$$

we say $f = g$ μ-a.e. on X, and denote that by $f \sim g$ because it is an *equivalence relation*. Moreover, for all $E \in M$,

$$\int_E f\,\mathrm{d}\mu = \int_E g\,\mathrm{d}\mu.$$

This is an important concept in measure theory. We will investigate further in the later chapters.

Definition 1.8.3. If for every $E \in M$ with $\mu(E) = 0$, $F \subseteq E$ implies $F \in M$, then μ is called a **complete measure** on M.

Theorem 1.8.4 (Completion of measure space). *Let (X, M, μ) be a measure space. Define $M^* := \{E \subseteq X : \exists A, B \in M \text{ s.t. } A \subset E \subset B, \mu(B \setminus A) = 0\}$. Define $\mu^*(E) = \mu(A)$. Then μ^* is complete on M^*.*

Proof. **Step 1:** We need to show that μ^* is well-defined, i.e. μ^* depends on the choice of E, not A. Suppose $A \subset E \subset B$, $A' \subset E \subset B'$. Then,

$$\begin{aligned} A \subset E \subset B' &\Rightarrow (A \setminus A') \subset (E \setminus A') \subset (B \setminus A') \\ &\Rightarrow \mu(A \setminus A') = 0. \end{aligned}$$

Similary, $\mu(A' \setminus A) = 0$, and

$$\begin{aligned} \mu(A) &= \mu(A \setminus A') + \mu(A \cap A') \\ &= \mu(A' \setminus A) + \mu(A' \cap A) \\ &= \mu(A') = \mu^*(E). \end{aligned}$$

Step 2: Verify that M^* is a σ-algebra. We see that $X \in M$. Let $E \in M$, if $A \subset E \subset B$ and $\mu(B \setminus A) = 0$, then $A^c \supset E^c \supset B^c$, and

$$\mu(A^c \setminus B^c) = \mu(A^c \cap B) = \mu(B \cap A^c) = \mu(B \setminus A) = 0.$$

So, $E^c \in M^*$. Let $E_n \in M^*$ and $A_n \subset E_n \subset B_n$. Define $A := \bigcup_{n=1}^{\infty} A_n$ and $B := \bigcup_{n=1}^{\infty} B_n$. Then, $A \subset \bigcup_{n=1}^{\infty} E_n \subset B$ and

$$\mu(B \setminus A) = \mu\left(\bigcup_{n=1}^{\infty} B_n \setminus A_n\right) = \sum_{n=1}^{\infty} \mu(B_n \setminus A_n) = 0.$$

Thus, $\bigcup_{n=1}^{\infty} E_n \in M^*$, and M^* is a σ-algebra.

Step 3: Finally, we need to show that μ^* is a measure on M^*, i.e. show countable additivity. Let $\{E_n\} \subset M^*$ be pairwise disjoint. Then there exists pairwise disjoint sequence $\{A_n\}$ such that

$$\mu^*(E) = \mu(A) = \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n) = \sum_{n=1}^{\infty} \mu^*(E_n).$$

Obviously, if $E \in M \subset M^*$, $\mu^*(E) < \infty$. Therefore, (X, M^*, μ^*) is a completion of (X, M, μ). ■

Theorem 1.8.5. *Let $f\colon X \to [0, \infty]$ be measurable and $\int_X f \, \mathrm{d}\mu = 0$. Then, $f = 0$ a.e.*

Proof. Let $E_n := \{x \in X : f(x) > \frac{1}{n}\}$. Then $E_n \in M$, for all $n \in \mathbb{N}$ and $E := \bigcup_{n \in \mathbb{N}} E_n = \{x : f(x) > 0\} \in M$. Consider

$$\begin{aligned} \int_X f \, \mathrm{d}\mu = 0 &\geq \int_X f \chi_{E_n} \, \mathrm{d}\mu \geq \frac{1}{n} \int_X \chi_{E_n} \, \mathrm{d}\mu \\ &= \frac{1}{n} \mu(E_n) \geq 0. \end{aligned}$$

Then, by σ-additivity, $\mu(E_n) = 0$ and $\mu(E) = 0$, which concludes $f = 0$ a.e. ■

Proposition 1.8.6. *If $f = 0$ a.e., then $\int_X f \, \mathrm{d}\mu = 0$.*

Proof. Let $E := \{x : f(x) \neq 0\}$. Then, $\mu(E) = 0$ and

$$\int_X f \,\mathrm{d}\mu = \int_X f(\chi_E + \chi_{E^c}) \,\mathrm{d}\mu = \int_E f \,\mathrm{d}\mu + \int_{E^c} f \,\mathrm{d}\mu = 0 + 0 = 0. \qquad \blacksquare$$

Theorem 1.8.7. *For all $n \in \mathbb{N}$, let f_n be a complex measurable function* ***defined a.e.*** *on X such that $\sum_{n=1}^{\infty} \int_X |f_n| \,\mathrm{d}\mu < \infty$. Then there exists $f \in L^1(\mu)$ such that*

$$f(x) = \sum_{n=1}^{\infty} f_n(x),\ \mu\text{-a.e.} \qquad \text{and} \qquad \int_X f \,\mathrm{d}\mu = \sum_{n=1}^{\infty} \int_X f_n \,\mathrm{d}\mu. \tag{1.8.1}$$

Proof. Define $S_n := \{x : f_n(x) \text{ is not defined.}\}$, so $\mu(S_n) = 0$. We want to show that there is S with $\mu(S) = 0$, and for all $x \notin S$, $\sum_{n=1}^{\infty} f_n(x)$ **converges absolutely**. Define $\varphi(x) := \sum_{n=1}^{\infty} |f_n(x)|$. By Monotone Convergence,

$$\int_X \varphi \,\mathrm{d}\mu = \sum_{n=1}^{\infty} \int_X |f_n| \,\mathrm{d}\mu < \infty.$$

Thus, $\{x \in X : \varphi(x) = \infty\}$ has measure zero. Hence, if $x \notin S_n, \forall n \in \mathbb{N}$ and $\varphi(x) \neq \infty$, $\sum_{n=1}^{\infty} f_n(x)$ must converge absolutely. Define $S := \bigcup_{n=1}^{\infty} S_n \cup \{x \in X : \varphi(x) = \infty\}$. Then $\mu(S) = 0$ and $\forall x \in S^c$, by absolute convergence we have

$$\sum_{n=1}^{\infty} f_n(x) = f(x).$$

Finally, define $g_n := f_1 + \cdots + f_n, \forall n \in \mathbb{N}$. Then, $|g_n| \leq \varphi$ and $g_n(x) \to f(x), \forall x \in S^c$. By Dominated Convergence on S^c,

$$\int_{S^c} f \,\mathrm{d}\mu = \lim_{n \to \infty} \int_{S^c} g_n \,\mathrm{d}\mu = \sum_{n=1}^{\infty} \int_{S^c} f_n \,\mathrm{d}\mu.$$

Since S has measure zero, we obtain the second equality. $\blacksquare$

Proposition 1.8.8 (An Average Argument)**.** *Let (X, M, μ) be a* ***finite*** *positive measure space. Let $f \in L^1(\mu)$ and S be a* ***closed set*** *in $\mathbb{C}$. If for every $E \in M$, the averages*

$$A_E(f) := \frac{1}{\mu(E)} \int_E f \,\mathrm{d}\mu \in S,$$

then $f(x) \in S$ for almost all $x \in X$.

Proof. Since S^c is open, $S^c = \bigcup_{n=1}^{\infty} E_n$ for open balls E_n of the form $B(\alpha, r) \subset S^c$. Let $E := f^{-1}(B(\alpha, r))$. Thus, it suffcies to show $\mu(E) = 0$. By contradiction, suppose $\mu(E) > 0$, then

$$\begin{aligned} |A_E(f) - \alpha| &= \frac{1}{\mu(E)} \left| \int_E (f - \alpha) \,\mathrm{d}\mu \right| \leq \frac{1}{\mu(E)} \int_E |f - \alpha| \,\mathrm{d}\mu \\ &\leq \frac{1}{\mu(E)} \int_E r \,\mathrm{d}\mu = r. \end{aligned}$$

However, it contradicts the hypothesis $A_E(f) \in S$. $\blacksquare$

Theorem 1.8.9. *Let $\{E_n\}$ be a sequence of measurable sets in X, such that $\sum_{n=1}^{\infty} \mu(E_n) < \infty$. Then almost all $x \in X$ lie in at most* ***finitely many*** *E_k's.*

Proof. Define $A := \{x : x \in \text{infinitely many } E_k\}$. Define $g(x) := \sum_{n=1}^{\infty} \chi_{E_n}(x)$. Then, $x \in A$ if and only if $g(x) = \infty$. By Monotone Convergence for series, $\int_X g \, \mathrm{d}\mu = \sum_{n=1}^{\infty} \mu(E_n) < \infty$. Hence, $g \in L^1(\mu)$ and $g(x) < \infty$ a.e. ∎

Chapter 2

Positive Borel Measures

2.1 Integration as a Linear Functional

Definition 2.1.1. Let V be a complex vector space. A **linear functional** is a linear transformation $\Lambda\colon V \to \mathbb{C}$.

Example 2.1.2. Let (X, M, μ) be a positive measure space. Recall that $L^1(\mu)$ is a complex vector space. The map $f \mapsto \int_X f \, \mathrm{d}\mu$ is a linear functional on $L^1(\mu)$. Likewise, for any bounded measurable function g, $f \mapsto \int_X fg \, \mathrm{d}\mu$ is also a linear functional.

Definition 2.1.3. Let V be a complex vector space of functions on X. We say $\Lambda\colon V \to \mathbb{C}$ is a **positive linear functional** if $f \geq 0$ implies $\Lambda(f) \geq 0$.

Example 2.1.4. Let $C([0,1])$ be the set of complex continuous functions on $[0,1]$. The map $f \mapsto \int_0^1 f \, \mathrm{d}x$ is a positive linear functional.

2.2 Topological Preliminaries

Let (X, τ) be a topological space.

Definition 2.2.1.

(a) A set $E \subseteq X$ is **closed** if E^C is open.

(b) The **closure** of E is defined as $\overline{E} := \bigcap_{F \supseteq E} F$, where F is any closed set.

(c) A set $K \subseteq X$ is **compact** if for any open cover $\{U_\alpha\}_{\alpha \in A}$ of K, there exists finite set $B \subset A$ such that $K \subseteq \bigcup_{\alpha \in B} U_\alpha$.

(d) A **neighborhood** of a point $p \in X$ is an open set $U \in \tau$ containing p.

(e) We say (X, τ) is **Hausdorff** if for all $p \neq q \in X$, there are open sets $U, V \in \tau$ such that $p \in U, q \in V$, and $U \cap V = \varnothing$.

(f) We say (X, τ) is **locally compact** if for each $p \in X$, there exists neighborhood U of p such that $\overline{U}$ is compact.

Theorem (Heine-Borel)**.** *Recall that a subset K of the Euclidean space $\mathbb{R}^n$ is compact if and only if K is closed and bounded.*

Proposition 2.2.2 (Finite Intersection Property)**.** *Let (X, τ) be a Hausdorff space, and $\{K_\alpha\}_{\alpha \in A}$ be a family of compact sets. If for every finite $B \subset A$, $\bigcap_{\alpha \in B} K_\alpha \neq \varnothing$, then $\bigcap_{\alpha \in A} K_\alpha \neq \varnothing$.*

Proof. By contrapositive, suppose $\bigcap_{\alpha \in A} K_\alpha = \varnothing$. Then $\bigcup_{\alpha \in A} K_\alpha^C = X$. There exists $\beta \in A$ such that $K_\beta \subseteq X = \bigcup_{\alpha \in A} K_\alpha^C$. By compactness, $K_\beta \subseteq \bigcup_{\alpha \in J} K_\alpha^c$ for some finite $J \subset A$. Hence, $K_\beta \cap \bigcap_{\alpha \in J} K_\alpha = \varnothing$. ■

Proposition 2.2.3. *Let (X, τ) be a topological space. If F is closed subset of a compact set K, then F is compact.*

Proof. Let $\{U_\alpha\}_{\alpha \in A}$ be an open cover on F. Then, $F^c \cup \bigcup_{\alpha \in A} U_\alpha \supseteq K$. So, $F^c \cup \bigcup_{\alpha \in B} U_\alpha \subseteq K$, for some finite $B \subset A$. Hence, F is covered by $\{U_{\alpha \in B}\}$. ■

Corollary 2.2.4.

1. *Compact subsets of Hausdorff spaces are closed.*
2. *If F is closed and K is compact in a Hausdorff space, then $F \cap K$ is compact.*

Proposition 2.2.5. *Let (X, τ) be a Hausdorff space, and K be compact. If $p \notin K$, then there are $U, V \in \tau, U \cap V = \varnothing$, such that $K \subseteq V, p \in U$.*

Proof. Fix $p \notin K$. For each $q \in K$, there are $U_q, V_q \in \tau, U_q \cap V_q = \varnothing$ such that $p \in U_q, q \in V_q$. Then by compactness,

$$K \subseteq \bigcup_{i=1}^{n} V_{q_i} = V \qquad \text{and} \qquad p \in \bigcap_{i=1}^{n} U_{q_i} = U.$$ ■

Theorem 2.2.6. *Let (X, τ) be a locally compact Hausdorff space. If K is compact, $K \subset U \in \tau$, then there exists $V \in \tau$ such that $\overline{V}$ is compact, and*

$$K \subset V \subset \overline{V} \subset U.$$

Proof. For every $x \in K$, there exists open neighborhood of x with compact closusre. By compactness, K is covered by a finite subcover of these neighborhoods. If $U = X$, let $V := G$ be the union of such subcover.

Otherwise, for each $p \in U^c$, there is open $W_p \in \tau$ such that $K \subset W_p$ and $p \notin \overline{W_p}$. Define $A := \{U^c \cap \overline{G} \cap \overline{W_p} : p \in U^c\}$. Observe that A is a collection of compact sets, and $\cap_{F \in A} F = \varnothing$. Hence, by FIP, there exists $p_1, \ldots, p_m \in U^c$ such that $\bigcap_{i=1}^m F_i = \varnothing$. Hence, $U^c \cap \overline{G} \cap (\bigcap_{i=1}^m \overline{W_{p_i}}) = \varnothing$. Define $V := G \cap (\bigcap_{i=1}^m W_{p_i})$ and verify the claim. ■

Definition 2.2.7. Let f be a real or extended-real valued function on (X, τ). If $\{x : f(x) > \alpha\} \in \tau$, $\forall \alpha \in \mathbb{R}$, f is said to be **lower semicontinuous**. Likewise, if $\{x : f(x) < \alpha\} \in \tau$, $\forall \alpha \in \mathbb{R}$, then f is **upper semicontinuous**.

Remark. A real valued function is continous if and only if it is both upper and lower semicontinuous.

Proposition 2.2.8. *If f, g are lower (upper) semicontinouos, so is $f + g$. If $u_1 \leq u_2 \leq \dots$ are lower semicontinuous, then so is $u := \lim_{n\to\infty} u_n$.*

Proof. Take $\{x : f(x) + g(x) > \alpha\} = \bigcup_{r\in\mathbb{Q}}(\{x : f(x) > r\} \cap \{x : g(x) > \alpha - r\})$, which is open. Take $\{x : u(x) > \alpha\} = \bigcup_{n=1}^{\infty}\{x : u_n(x) > \alpha\}$, which is open. ∎

Definition 2.2.9. Let (X, τ) be a topological space, and $f : X \to \mathbb{C}$. The **support** of f is defined as

$$\operatorname{supp}(f) := \overline{\{x \in X : f(x) \neq 0\}}. \tag{2.2.1}$$

We also define

$$C_c(X) := \{f : X \to \mathbb{C} : f \text{ continuous with compact support}\}.$$

Remark. Note that $\operatorname{supp}(f + g) \subseteq \operatorname{supp}(f) \cup \operatorname{supp}(g)$, which is compact. Thus, $C_c(X)$ is a complex vector subspace of $C(X)$.

Proposition 2.2.10. *Let X, Y be topological spaces and $f : X \to Y$ be continuous. If $K \subset X$ is compact, then $f(K)$ is compact in Y.*

Proof. Let $\{V_\alpha\}_{\alpha\in A}$ be an open cover of $f(K)$. So, $\{f^{-1}(V_\alpha)\}_{\alpha\in A}$ is an open cover of K, and there is a finite $B \subset A$ such that $K \subseteq \bigcup_{\alpha\in B} f^{-1}(V_\alpha)$, which concludes $f(K) \subseteq \bigcup_{\alpha\in B} V_\alpha$. ∎

Theorem 2.2.11 (Urysohn's Lemma). *Let (X, τ) be a locally compact Hausdorff space. Suppose K is a compact subset and $K \subseteq V \in \tau$. Then there exists $f \in C_c(X)$ such that $\chi_K \leq f \leq \chi_V$.*

Proof. First, define $P := [0, 1]_{\mathbb{Q}} = \{0, 1, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \dots\}$. By Theorem (2.2.6), there are $V_0, V_1 \in \tau$ such that

$$K \subset V_1 \subset \overline{V_1} \subset V_0 \subset \overline{V_0} \subset V.$$

For all $r, q \in P$, with $0 < r < q < 1$, choose open sets V_r, V_q such that

$$K \subset V_1 \subset \overline{V_1} \subset \dots V_q \subset \overline{V_q} \subset \dots \subset V_r \subset \overline{V_r} \subset \dots \subset V_0 \subset \overline{V_0} \subset V.$$

Define $f : X \to [0, 1]$ by

$$f(x) = \begin{cases} \sup\{r \in P : x \in V_r\}, & \text{if } x \in V, \\ 0, & \text{if } x \notin V. \end{cases}$$

Clearly, $f = 1$ on K and $0 \leq f \leq 1$ with $\operatorname{supp}(f) \subseteq \overline{V_0}$, which is compact. For continuity of f, suppose $x \in X$ with $f(x) = 0$, or $f(x) = 1$. Then f is continuous at x by sequential continuity.

Now suppose $f(x) \in (0,1)$. For all $(a,b) \subset (0,1)$ with $f(x) \in (a,b)$, choose $r, q \in P$ such that $a < r < f(x) < q < b$. Let $U := V_r \setminus \overline{V_q}$. We see that $U \in \tau$ with $f(x) \in f(U) \subset (a,b)$. Therefore, $f \in C_c(X)$. ■

Remark. The Urysohn's Lemma is an important tool in building more complicated functions.

Corollary 2.2.12. *Let (X, M, μ) be a positive measure space and a locally compact Hausdorff space. Then, every compact K has $\mu(K) < \infty$ if and only if $C_c(X) \subset L^1(\mu)$ and $\Lambda(f) := \int_X f \,\mathrm{d}\mu$ defines a positive linear functional on $C_c(X)$.*

Proof. ($\Rightarrow$) Suppose every compact subset of X has finite measure. For each $f \in C_c(X)$, $|f| \neq 0$ on some compact K. Thus, $|f| < \infty$ on K and $\mu(K) < \infty$. Hence,

$$\int_X |f| \,\mathrm{d}\mu = \int_K |f| \,\mathrm{d}\mu + \int_{K^c} |f| \,\mathrm{d}\mu < \infty + 0.$$

The second assertion is trivial.

($\Leftarrow$) For every compact K, by compactness we may choose open precompact V such that $K \subset V$. By Urysohn's Lemma, there exists $f \in C_c(X)$ such that $\chi_K \leq f$. By hypothesis,

$$\mu(K) = \int_X \chi_K \,\mathrm{d}\mu \leq \int_X f \,\mathrm{d}\mu < \infty.$$ ■

2.3 Riesz Representation Theorem and Borel Measures

Theorem 2.3.1 (Riesz Representation Theorem). *Let X be a locally compact Hausdorff space, and Λ be a **positive linear functional** on $C_C(X)$. Then exists a σ-algebra M which **contains all Borel sets** in X, and a **unique positive measure** μ such that:*

(a) Functional to Integral: $\Lambda(f) = \displaystyle\int_X f \,\mathrm{d}\mu$, *for all* $f \in C_C(X)$.

(b) Finite Measure on Compact Set: $\mu(K) < \infty$, *for all compact* $K \subset X$.

(c) Outer Regularity: If $E \in M$, $\mu(E) = \inf\{\mu(V) : E \subset V,\ V \text{ open}\}$.

(d) Inner Regularity: If E is open, or $E \in M$ with $\mu(E) < \infty$, then $\mu(E) = \sup\{\mu(K) : K \subset E,\ K \text{ compact}\}$.

(e) *Complete Measure: If $E \in M, A \subseteq E$, then $\mu(E) = 0$ implies $A \in M$.*

Proof. Read Rudin page 41. ■

Definition 2.3.2. Let (X, τ) be a locally compact Hausdorff space. A measure μ defined on the **Borel σ-algebra** M is called a **Borel measure**.

A Borel set E is **outer regular**, or **inner regular** if it satisfies property (c), or (d), respectively in the Riesz Representation Theorem. If every Borel set is both outer and inner regular, then μ called a **regular measure**.

Remark. Note that in Theorem (2.3.1), we only have inner regularity for open sets and Borel sets with finite measure. In general, it is the best we can have. However, with extra assumptions, we can obtain regularity.

Definition 2.3.3. Let X be a topological space. A set E is called **σ-compact** if $E = \bigcup_{n=1}^{\infty} K_n$, where K_n is compact for all n. A set is called F_σ if it is a **countable union of open sets**; likewise it is called G_δ if it is a **countable intersection of closed sets**.

Definition 2.3.4. A meaurable set E in a measure space is said to have **σ-finite measure** if $E = \bigcup_{n=1}^{\infty} E_n$, where E_n is measurable and $\mu(E_n) < \infty$ for all n. In particular, if X has σ-finite measure, then μ is called σ-finite.

Theorem 2.3.5 (Riesz Representation Theorem). *In addition to the hypotheses of Theorem (2.3.1), if X is σ-compact, then the followings hold:*

(a) *For all $E \in M$, and $\varepsilon > 0$, there are closed F and open V, so that $F \subset E \subset V$ and $\mu(V \setminus F) < \varepsilon$.*

(b) *μ is a regular Borel measure on X.*

(c) *For all $E \in M$, there are F_σ-set A, G_δ-set B, such that $A \subset E \subset B$ and $\mu(B \setminus A) = 0$.*

From (c), every $E \in M$ is the union of an F_σ and a set of measure zero.

Proof. (a). By hypothesis, $X = \bigcup_{n=1}^{\infty} K_n$, where K_n is compact and $\mu(K_n) < \infty$ for every n. Given any $E \in M, \varepsilon > 0$, for every n we have $\mu(K_n \cap E) < \infty$. By outer regularity, there exists open $V_n \supseteq (K_n \cap E)$, such that

$$\mu(V_n \setminus (K_n \cap E)) = \mu(V_n) - \mu(K_n \cap E) < \frac{\varepsilon}{2^{n+1}}.$$

Define $V := \bigcup_{n=1}^{\infty} V_n$. Then $V \setminus E \subseteq \bigcup_{n=1}^{\infty}(V_n \setminus (K_n \cap E))$, and

$$\mu(V \setminus E) \leq \sum_{n=1}^{\infty} \mu(V_n \setminus (K_n \cap E)) = \frac{\varepsilon}{2}.$$

Next, consider E^c. Similarly, outer regularity gives open $W \supseteq E^c$ with $\mu(W \setminus E^c) < \frac{\varepsilon}{2}$. Let $F := W^c$. Then we have $F \subseteq E$ with $\mu(E \setminus F) = \mu(W \setminus E^c) < \varepsilon/2$, and $\mu(V \setminus F) < \varepsilon$.

(b). For all $E \in M$, there exists F with $\mu(E \setminus F) < \varepsilon$. Let $F_n := F \cap (\bigcup_{i=1}^n K_i)$. Then F_n is compact in E, with $\mu(F_n) \nearrow \mu(F)$. Thus, μ is inner regular.

(c). Let $\varepsilon_n = \frac{1}{n}$. By (a), there exists open V_n and closed F_n such that $F_n \subseteq E \subseteq V_n$, and $\mu(V_n \setminus F_n) < \frac{1}{n}$. Then $\bigcup_{n=1}^\infty F_n$ and $\bigcap_{n=1}^\infty V_n$ are the corresponding F_σ and G_δ sets. ■

Theorem 2.3.6. *Let (X, τ) be a locally compact Housdorff space in which every open set is σ-compact. Let λ be a positive Borel measure on X such that for all compact set K, $\lambda(K) < \infty$. Then λ is regular.*

Proof. Define $\Lambda : C_c(X) \to \mathbb{C}$ by $\Lambda(f) := \int_X f \, \mathrm{d}\lambda$. Since $\lambda(K) < \infty$ for all compact set K, Λ is a positive linear functional on $C_c(X)$. By the **Riesz Representation Theorem** (2.3.1), there exists regular Borel measure μ such that

$$\int_X f \, \mathrm{d}\lambda = \int_X f \, \mathrm{d}\mu.$$

We will show that $\lambda = \mu$. First observe that they agree on open sets. Let V be open. By hypothesis, V is σ-compact. So, $V = \bigcup_{i=1}^\infty K_i$, where K_i is compact. By the Urysohn's Lemma (2.2.11), for each i, there is $f_i \in C_c(X)$ such that $\chi_{K_i} \leq f_i \leq \chi_V$.

Define $g_n := \max\{f_i \mid 1 \leq i \leq n\}$. Then, the g_n's are continuous with compact supports, and $g_n(x) \nearrow \chi_V(x)$. Hence, by Monotone Convergence,

$$\Lambda(V) = \lim_{n\to\infty} \int_X g_n \, \mathrm{d}\lambda = \lim_{n\to\infty} \int_X g_n \, \mathrm{d}\mu = \mu(V).$$

Now suppose E is a Borel set. Given any $\varepsilon > 0$, by Theorem (2.3.5), there are closed F and open V with $F \subseteq E \subseteq V$ such that $\mu(V \setminus E) < \varepsilon$. Hence, $\mu(V) \leq \mu(E) + \varepsilon$. Note that $V \setminus F$ is open. So, by the preceding step, $\lambda(V \setminus E) \leq \varepsilon$ and $\lambda(V) \leq \lambda(E) + \varepsilon$. Thus,

$$\begin{aligned} \lambda(E) \leq \lambda(V) = \mu(V) \leq \mu(E) + \varepsilon, \\ \mu(E) \leq \mu(V) = \lambda(V) \leq \lambda(E) + \varepsilon, \end{aligned}$$

and we have $|\lambda(E) - \mu(E)| < \varepsilon$. Therefore, $\lambda(E) = \mu(E)$. ■

2.4 Lebesgue Measure

In this section, we work with the familiar Euclidean space $\mathbb{R}^k$ with its Borel algebra. Recall that the Riemann integral is only defined for continuous functions over compact sets. We want to extend this integration from continuous functions to Borel measurable functions. Therefore, we must first find a suitable Borel measure on $\mathbb{R}^k$, which presevers some useful properties of the Riemann integrals.

Definition 2.4.1. Let $E \subseteq \mathbb{R}^k$ and $x \in \mathbb{R}^k$. The **translate** of E by x is the set

$$E + x := \{y + x : y \in E\}.$$

Definition 2.4.2. A k-**cell** in $\mathbb{R}^k$ is a set of the form $V = \prod_{i=1}^k I_i$, where I_i is a bounded interval with endpoints α_i and β_i. We also define the **volume** of V by

$$vol(V) := \prod_{i=1}^{k} (\beta_i - \alpha_i).$$

Definition 2.4.3. Let $\alpha = (\alpha_1, \ldots, \alpha_k) \in \mathbb{R}^k$ and $\delta > 0$. A δ-box with corner at α is a set $Q(\alpha, \delta) = \prod_{i=1}^k [\alpha_i, \alpha_i + \delta)$.

Theorem 2.4.4 (The Lebesgue Measure)**.** *There exists a σ-algebra $M \supsetneq \mathcal{B}(\mathbb{R}^k)$, the Borel algebra of $\mathbb{R}^k$, and a positive complete measure m on M satisfying the followings:*

(a) $m(V) = vol(V)$ for every k-cell V.

(b) $E \in M$ if and only if there are F_σ-set A and G_δ-set B such that $A \subset E \subset B$ and $m(A \setminus B) = 0$. Moreover, m is a regular Borel measure on $\mathcal{B}(\mathbb{R}^k)$.

(c) m is translation invariant, i.e. $m(E + x) = m(E)$, for all $E \in M$, $x \in \mathbb{R}^k$.

(d) If μ is a positive, translation-invariant, Borel measure on $\mathbb{R}^k$ such that $\mu(K) < \infty$ for all compact K, then there is $c \in \mathbb{R}$ such that $\mu(E) = cm(E)$, for all $E \in \mathcal{B}(\mathbb{R}^k)$.

(e) Let $T\colon \mathbb{R}^k \to \mathbb{R}^k$ be a linear map. Then there is $\triangle(T) \in [0, \infty)$ such that

$$m(T(E)) = \triangle(T)m(E), \qquad \forall E \in \mathcal{B}(\mathbb{R}^k).$$

In particular, $m(T(E)) = m(E)$ when T is a rotation.

Definition 2.4.5. The sets $E \in M$ are called **Lebesgue measurable** in $\mathbb{R}^k$; m is called the **Lebesgue measure** on $\mathbb{R}^k$.

Proof. **Step 1:** Define $\Lambda\colon C_c(\mathbb{R}^k) \to \mathbb{C}$ by

$$\Lambda(f) := \int_{\mathbb{R}^k} f \,\mathrm{d}V, \tag{2.4.1}$$

where $\mathrm{d}V$ stands for the Riemann Integral. Note that this definition is well-defined because $\operatorname{supp}(f)$ is a compact set in $\mathbb{R}^k$, and f is bounded. Also, Λ is a positive linear functional on $C_c(\mathbb{R})$. By the **Riesz Representation Theorem** (2.3.1), there is a σ-algebra $M \supset \mathcal{B}(\mathbb{R}^k)$, and a complete positive measure m, such that

$$\Lambda(f) = \int_{\mathbb{R}^k} f \,\mathrm{d}m, \tag{2.4.2}$$

for all $f \in C_c(\mathbb{R}^k)$. Since $\mathbb{R}^k$ is σ-compact, property (b) follows from the other Riesz Representation Theorem (2.3.5).

Step 2: For property (a), first suppose V is an open k-cell. There is a sequence of compact sets $K_1 \subset K_2 \subset \cdots \subset V$, such that $\bigcup_{n=1}^{\infty} K_n = V$. By the Uryshon's Lemma (2.2.11), there is $f_n \in C_c(\mathbb{R}^k)$ with $\chi_{K_n} \le f_n \le \chi_V$, for each n. Hence,

$$m(K_n) = \Lambda(\chi_{K_n}) = \int_{\mathbb{R}^k} \chi_{K_n}\, \mathrm{d}m \le \int_{\mathbb{R}^k} f_n\, \mathrm{d}m \le \int_{\mathbb{R}^k} \chi_V\, \mathrm{d}m = m(V).$$

On the other hand, from elementary Calculus, $m(K_n) = vol(K_n)$, and $vol(K_n) \to vol(V)$. By Monotone Convergence,

$$vol(V) \le \int_{\mathbb{R}^k} \lim_{n\to\infty} f_n\, \mathrm{d}m \le vol(V). \tag{2.4.3}$$

Since $f_n(x) \to \chi_V(x)$, for all $x \in \mathbb{R}^k$, we conclude

$$m(V) = \int_{\mathbb{R}^k} \chi_V\, \mathrm{d}m = vol(V). \tag{2.4.4}$$

Finally, suppose V is any k-cell. Pick a decreasing sequence of open sets $V_1 \supset V_2 \supset \cdots \supset V$, with $m(V_1) < \infty$. Then by monotonicity, $m(V_n) \to m(V) = vol(V)$.

Step 3:[1] For property (c), fix $x \in \mathbb{R}^k$ and define $\lambda_x(E) := m(E + x)$, for all $E \in M$. It is easy to see λ_x satisfies σ-addivity, hence is a measure. Also, for all k-cell V, $\lambda_x(V) = vol(V) = m(V)$. Since every open set is a countable union of k-cells, and λ_x is a measure, $\lambda_x(E) = m(E)$ on all open E. Then, $\lambda_x(K) < \infty$, for all compact K. By Theorem (2.3.6), λ_x is regular. Therefore, by regularity of λ_x and m, $m(E + x) = \lambda_x(E) = m(E)$, for every $E \in \mathcal{B}(\mathbb{R}^k)$. Finally, recall that F_σ and G_δ-sets are also Borel sets. Thus, the equality also holds by property (b).

Step 4: For property (d), let Q_0 be a 1-box and define $c := \mu(Q_0)$. For each $n \in \mathbb{N}$, suppose Q_n is a 2^{-n}-box. Note that Q_0 is a disjoint union of 2^{nk} many such Q_n boxes. By translation invariance,

$$2^{nk}\mu(Q_n) = \mu(Q_0) = c \cdot 1 = cm(Q_0) = 2^{nk} cm(Q_n).$$

We conclude that $\mu(Q_n) = cm(Q_n)$, for all $n \in \mathbb{N}$. Finally, since every k-cell V is a countable disjoint union these Q_n boxes, $\mu(V) = cm(V)$. Hence, similar to Step 3, $\mu(E) = cm(E)$, for all Borel set E.

Step 5: For property (e), first suppose $\operatorname{rank}(T) < k$. Then, $m(T(E)) = 0$, for all k-cells E, hence Borel sets; and we let $\triangle(T) := 0$. Now suppose $\operatorname{rank}(T) = k$. From linear algebra, T is bijective and linear, hence a **homeomorphism** from $\mathbb{R}^k$ to $\mathbb{R}^k$. More importantly, $T(E)$ is a Borel set for all Borel set E.

Define a positive Borel measure μ by $\mu(E) := m(T(E))$, for all $E \in \mathcal{B}(\mathbb{R}^k)$. For all $x \in \mathbb{R}^k$, by linearity of T and translation invariance of m,

$$\mu(E + x) = m(T(E + x)) = m(T(E) + T(x)) = m(T(E)) = \mu(E).$$

[1] This is a standard approach in proving two Borel measures are equal. First show that on open sets, then Borel sets by regularity. If necessary, all measurable sets with F_σ and G_δ. The usual tools are Monotone/Dominated Convergence, σ-addivity, monotonicity, and Uryshon's Lemma.

Hence, μ is a positive, translation invariant, Borel measure on $\mathbb{R}^k$, and $\mu(K) < \infty$ for all compact K. By property (d), there is $\triangle(T) \in \mathbb{R}$ such that

$$\mu(E) = m(T(E)) = \triangle(T)m(E), \qquad \forall E \in \mathcal{B}(\mathbb{R}^k).$$

In other words, one can compute $\triangle(T)$ easily by $m(T(E))/m(E)$, for some $m(E) \in (0, \infty)$. Finally, if T is a rotation, then in particular $T(B) = B$, where B is the unit open ball. Therefore, $\triangle(T) = 1$ and $m(T(E)) = m(E)$. ■

Proposition 2.4.6. *Let $T\colon \mathbb{R}^k \to \mathbb{R}^k$ be a linear map. Then,*

$$\triangle(T) = |det(T)|, \tag{2.4.5}$$

where $\triangle(T)$ is given in Theorem (2.4.4).

Proof. **Step 1:** From linear algebra, let $\{e_1, \dots, e_k\}$ be the standard basis for $\mathbb{R}^k$. For each j, $1 \le j \le k$,

$$T(e_j) = \sum_{i=1}^{n} \alpha_{ij} e_i,$$

for some $\alpha_{ij} \in \mathbb{R}$. Hence, the matrix representation of T is $[T] := (\alpha_{ij})$.

Step 2: Note that if $T = T_1 \circ T_2$, then $\triangle(T) = \triangle(T_1)\triangle(T_2)$, and $det(T) = det(T_1)det(T_2)$. Recall that every linear operator T on a finite dimensional vector space is a finite product of the following three types, each corresponds to one elementary row operation on $[T]$:

(I) Switching: $T(e_1) = e_2, T(e_2) = e_1$, and $T(e_i) = e_i$, for $3 \le i \le k$.

(II) Scaling: $T(e_1) = \alpha e_1$, and $T(e_i) = e_i$, for $2 \le i \le k$.

(III) Addition: $T(e_1) = e_1 + e_2$, and $T(e_i) = e_i$, for $2 \le i \le k$.

Thus, it suffices to show that each of these types satisfies equation (2.4.5). To determine $\triangle(T)$, let E be a 1-box cornered at 0.

Step 3: If T is of type (I), then $T(E) = E$, and $\triangle(T) = 1$. Also, $[T]$ has exactly one 1 on each coloumn and row. Hence, $det(T) = \pm 1$, and $\triangle(T) = |det(T)|$. If T is of type (II), then $m(T(E)) = |\alpha| m(E)$, and $\triangle(T) = |\alpha| = |det(T)|$.

Step 4: If T is of type (III), then $det(T) = 1$. For $\triangle(T)$, write $x \in \mathbb{R}^k$ as $x = (x_1, \dots, x_k)$. Then, $(x_1, x_2, \dots, x_k) \mapsto (x_1 + x_2, x_2, \dots, x_k)$, and

$$T(E) = \{y \in \mathbb{R}^k : y_2 \le y_1 < y_2 + 1,\ 0 \le y_i < 1,\ i \ne 1\}.$$

Let $S_1 := \{y \in T(E) : y_1 < 1\}$, and $S_2 := T(E) \setminus S_1$. Then, $S_1 \cap (S_2 - e_1) = \varnothing$, and $S_1 \cup (S_2 - e_1) = E$.[2] Hence,

$$\triangle(T) = m(T(E))/m(E) = m(S_1 \cup S_2)$$

[2]To see that, consider I^2 be the unit square in $\mathbb{R}^2$. Then $T(I^2)$ is the parallelogram with vertices $(0,0), (1,0), (1,1)$, and $(2,1)$. S_1 is the lower triangle, and S_2 is the upper.

$$
\begin{aligned}
&= m(S_1) + m(S_2) = m(S_1) + m(S_2 - e_1) \\
&= m(E) = 1.
\end{aligned}
$$

Therefore, $\triangle(T) = 1 = |det(T)|$. ■

REMARK 2.4.7. If m is the Lebesgue measure on $\mathbb{R}^k$, we usually write $L^1(\mathbb{R}^k)$ instead of $L^1(m)$. From equation (2.4.1), we see that for every complex continuous function f supported in a compact set K, the Riemann integral of f agrees with the Lebesgue integral of f.

REMARK 2.4.8. Recall that we are working in $M \supset \mathcal{B}(\mathbb{R}^k)$. It is important to know which Lebesgue measurable set $E \in M$ is a Borel set in $\mathcal{B}(\mathbb{R}^k)$. By a cardinality argument (Rudin, p.53), in fact most $E \in M$ are not Borel sets. Finally, we conclude the dicussion with the following observation: Every $E \in M$ with $m(E) > 0$ has non-measurable subsets. The proof is given as follows.

PROPOSITION 2.4.9. *Let M be a σ-algebra on $\mathbb{R}$ and $\lambda\colon M \to [0, \infty]$ be a translation invariant measure with $0 < \lambda([0, 1)) < \infty$. Then there exists $E \subset [0, 1)$ such that $E \notin M$.*

Proof. Note that $(\mathbb{R}, +)$ is a group and $(\mathbb{Q}, +)$ is its subgroup. Define an equivalence relation on $[0, 1)$ by $x \sim y \iff |x - y| \in \mathbb{Q}$. This gives an partition of $[0, 1)$ by the equivalence classes. By the Axiom of Choice, we can pick one representative from each equivlence class, and denote E the set of such representatives. Then E has the following property:

$$(E + r) \cap (E + s) = \varnothing, \qquad \forall r, s \in \mathbb{Q}, r \neq s.$$

To see it, suppose $x \in (E + r) \cap (E + s)$. Then, $y + r = x = z + s$, for some $y, z \in E, y \neq z$. Then, $y - z = s - r \in \mathbb{Q}$, which is a contradiction because $[y]$ and $[z]$ are different equivalence classes. Note that

$$E \subset [0, 1) \subseteq \bigcup_{r \in \mathbb{Q} \cap [-1,1]} (E + r) \subset [-1, 2).$$

Now, by way of contradiction suppose $E \in M$, then we have

$$
\begin{aligned}
\lambda([0, 1)) &\leq \sum_{r \in \mathbb{Q} \cap [-1,1]} \lambda(E + r) \leq \lambda([-1, 2)) \\
&= \sum_{r \in \mathbb{Q} \cap [-1,1]} \lambda(E) \leq 3\lambda([0, 1)) < \infty
\end{aligned}
$$

Since there are infinitely many such r's, $\lambda(E) = 0$. However, then $\lambda([0, 1])) = 0$, which is a contradiction. ■

2.5 Continuity Properties of Measurable Functions

The following two theorems establish important relations between continuous and measurable functions, i.e. approximation using continuous functions. In this section, let μ be a positive measure on a locally compact Hausdroff space X, which has the **five properties stated in Riesz Representation Theorem**. In particular, μ could be the Lebesgue measure on $\mathbb{R}^k$.

Theorem 2.5.1 (**Lusin's Theorem**). *Let $f : X \to \mathbb{C}$ be measurable. Suppose $A := \{x : f(x) \neq 0\}$ and $\mu(A) < \infty$. Then, given any $\varepsilon > 0$, there exists $g \in C_C(X)$ such that*

$$\mu(\{x : f(x) \neq g(x)\} < \varepsilon, \tag{2.5.1}$$

and

$$\sup_{x\in X} |g(x)| \leq \sup_{x\in X} |f(x)|. \tag{2.5.2}$$

Proof. **Step 1:** First suppose A is compact. By local compactness, there exists open V such that $A \subseteq V$ and $\overline{V}$ is compact. Let us begin with simple functions. Suppose f is simple,

$$f := \sum_{j=1}^{n} \alpha_j \chi_{A_j}, \qquad A_j \text{ disjoint}, \ \bigcup_{j=1}^{n} A_j = A, \ \alpha_j > 0.$$

By regularity, with $\mu(A_j) \leq \mu(A) < \infty$, there are compact sets $K_j \subseteq A_j$, such that $\mu(A_j \setminus K_j) < \frac{\epsilon}{2^j}$. Note that there are finitely many disjoint K_j's. Hence, we can find disjoint open V_j such that $K_j \subseteq V_j \subseteq V$, with $\overline{V_j}$ compact. Moreover, by outer regularity of μ, we may assume $\mu(V_j \setminus K_j) < \frac{\epsilon}{2^j}$.

By Urysohn's Lemma, there are $g_j \in C_c(X)$ such that $\chi_{K_j} \leq g_j \leq \chi_{V_j}$. Define

$$g(x) := \sum_{j=1}^{n} \alpha_j g_j, \qquad \forall x \in X.$$

Then, $g \in C_c(X)$ with $\operatorname{supp}(g) \subseteq \overline{V}$, and $|g(x)| \leq \max\{|\alpha_j|\} = \max_{x\in X} |f(x)|$. Also, if $x \in (\bigcup_{j=1}^{n} K_j) \cup (A^c \cap (\bigcap_{j=1}^{n} V_j^c))$, then $f(x) = g(x)$. Hence,

$$\begin{aligned}
\{x \in X : g(x) \neq f(x)\} &\subseteq \left(\bigcap_{j=1}^{n} K_j^c\right) \cap \left(A \cup (\bigcup_{j=1}^{n} V_j)\right) \\
&= \left(\bigcap_{j=1}^{n} K_j^c \cap A\right) \cup \left((\bigcap_{j=1}^{n} K_j^c) \cap (\bigcup_{j=1}^{n} V_j)\right) \\
&= \left(\bigcup_{j=1}^{n} (A_j \setminus K_j)\right) \cup \left(\bigcup_{j=1}^{n} (V_j \setminus K_j)\right).
\end{aligned}$$

Thus, $\mu(\{f \neq g\}) < 2\varepsilon$ and it concludes the case for simple functions.

Step 2: Next, suppose $0 \leq f \leq 1$. By Theorem (1.4.4), using the staircase functions φ_n, there exists a sequence of simple functions $\{s_n\}$ such that $0 \leq s_n \leq s_{n+1}$ and $s_n(x) \nearrow f(x)$, for all $x \in X$.

Define $t_n := s_n - s_{n-1}$, with $s_0 \equiv 0$. Then t_n's are simple, $t_n = 0$ on A^c, and $t_n \geq 0$. Moreover, from the definition of φ_n, $|t_n| \leq \frac{1}{2^{n-1}}$. By Step 1, there are $g_n \in C_c(X)$ such that (1): $\mu(\{g_n \neq t_n\}) < \frac{\varepsilon}{2^n}$; (2): $|g_n| \leq \max t_n \leq \frac{1}{2^{n-1}}$; and (3): $\operatorname{supp}(g_n) \subseteq \overline{V}$.

Define $g := \sum_{n=1}^{\infty} g_n$. Then $\operatorname{supp}(g) = \bigcup_{n=1}^{\infty} \operatorname{supp}(g_n) \subseteq \overline{V}$ and $\sum_{n=1}^{\infty} g_n$ converges uniformly by the M-test. Hence, $g \in C_c(X)$. If for some $x \in X, g_n(x) = t_n(x)$, for all $n \in \mathbb{N}$, then $g(x) = f(x)$. Thus,

$$\mu(\{f \neq g\}) \leq \mu\left(\bigcup_{n=1}^{\infty}\{t_n \neq g_n\}\right) \leq \sum_{n=1}^{\infty} \mu(\{t_n \neq g_n\}) = \varepsilon.$$

Consequently, if f is bounded, the results follow by scaling f.

Step 3: If $f\colon X \to [0, \infty)$ is measurable, then $\bigcap_{n=1}^{\infty}\{f \geq n\} = \varnothing$. Since $\mu(\{f \geq 1\}) \leq \mu(A) < \infty$ by hypothesis, the monotonicity of μ implies that

$$\mu(\{f \geq n\}) \to \mu(\varnothing) = 0.$$

Hence, given $\varepsilon > 0$, there is n such that $\mu(\{f \geq n\}) < \frac{\varepsilon}{2}$, and $f' := f\chi_{\{f<n\}}$ is bounded. By Step 2, we obtain $g \in C_c(X)$, such that $g = f'$, except on a set of measure $< \frac{\varepsilon}{2}$. Therefore, $g = f$, except on set of measure $< \varepsilon$, and satisfies both inequalities.

Step 4: Now if $f\colon X \to \mathbb{C}$ is measurable, write

$$f = (u_+ - u_-) + i(v_+ - v_-)$$

and perform the approximation separately on each term. Then we obtain g that satisifies inequality (2.5.1). To obtain inequality (2.5.2), if $|f|$ is bounded, define $M := \sup\{|f(x)| : x \in X\}$, and $\varphi\colon \mathbb{C} \to \mathbb{C}$ by

$$\varphi(z) := \begin{cases} z, & \text{if } |z| \leq M, \\ M\frac{z}{|z|}, & \text{if } |z| > M. \end{cases}$$

Note that $\varphi \in C_c(\mathbb{C})$. Therefore, $\varphi \circ g$ satisfies both inequalities. If $|f|$ is unbounded, then $f\chi_{\{|f|<n\}}$ is bounded, and $\mu(\{|f| \geq n\}) < \frac{\varepsilon}{2}$. We can then proceed as in Step 3.

Step 5: Finally, suppose A is not compact. Since $\mu(A) < \infty$, by inner regularity, there is a compact $K \subset A$ such that $\mu(A \setminus K) < \varepsilon/2$. Now apply the preceding steps on K, with $\varepsilon/2$ to obtain $g \in C_c(X)$. ■

Corollary 2.5.2. *Assume that the hypotheses of Lusin's Theorem are satisfied and $|f| \leq 1$. Then there is a sequence $\{g_n \in C_c(X)\}$ such that $|g_n| \leq 1$, and*

$$f(x) = \lim_{n\to\infty} g_n(x) \quad a.e.$$

Proof. By Lusin's Theorem, $\forall n \in \mathbb{N}$, there is a $g_n \in C_c(X)$ such that $|g_n| \leq 1$ and $\mu(E_n) \leq 2^{-n}$, where $E_n = \{x : f(x) \neq g(x)\}$. Since $\sum_{n=1}^{\infty} \mu(E_n) = 1 < \infty$, by Theorem 1.8.9., almost every $x \in X$ lies in finitely many of the E_n's. Hence, for every such fixed x, there is a large enough n such that $f(x) = g_n(x)$. ■

Theorem 2.5.3 (Vitali-Caratheodory Theorem). *Let $f : X \to \mathbb{R}$, $f \in L^1(\mu)$. Then, given $\varepsilon > 0$, there are u, v, such that $u \leq f \leq v$, u is upper semicontinuous and bounded above, v is lower semicontinuous and bounded below, and*

$$\int_X (v-u)\,\mathrm{d}\mu < \varepsilon.$$

Proof. First suppose $f \geq 0$. Using the staircase functions, we obtain a sequence of increasing functions $s_n \nearrow f$ pointwise. Take $s_0 \equiv 0$ and define $t_n := s_n - s_{n-1}$. Then t_n is simple and $f = \sum_{n=1}^{\infty} t_n$.

We can define constants $c_j > 0$ and measurable sets E_j, not necessarily disjoint, such that $f = \sum_{j=1}^{\infty} c_j \chi_{E_j}$. By Monotone Convergence,

$$\sum_{j=1}^{\infty} c_j \int_X \chi_{E_j}\,\mathrm{d}\mu = \int_X f\,\mathrm{d}\mu < \infty.$$

Hence, $\mu(E_j) < \infty$ for all j. By regularity, there are compact K_j and open V_j such that $K_j \subseteq E_j \subseteq V_j$ and $\mu(V_j \setminus K_j) < \frac{\varepsilon}{2^j c_j}$. Also, since the series converges, there exists N, such that $\sum_{j=N+1}^{\infty} c_j \mu(E_j) < \varepsilon$. Define

$$u := \sum_{j=1}^{N} c_j \chi_{K_j} \qquad \text{and} \qquad v := \sum_{j=1}^{\infty} c_j \chi_{V_j}.$$

Then, u is upper semicontinuous and v is lower semicontinuous with $u \leq f \leq v$. Finally,

$$\begin{aligned}
\int_X (v-u)\,\mathrm{d}\mu &= \int_X \left(\sum_{j=1}^{N} c_j \chi_{X_{V_j \setminus K_j}} \right) \mathrm{d}\mu + \int_X \left(\sum_{j=N+1}^{\infty} c_j \chi_{V_j} \right) \mathrm{d}\mu \\
&= \sum_{j=1}^{N} c_j \mu(V_j \setminus K_j) + \sum_{j=N+1}^{\infty} c_j \mu(V_j) \\
&= \sum_{j=1}^{N} c_j \mu(V_j \setminus K_j) + \sum_{j=N+1}^{\infty} c_j (\mu(V_j \setminus E_j) + \mu(E_j))
\end{aligned}$$

$$\leq \sum_{j=1}^{\infty} c_j \mu(V_j \setminus K_j) + \sum_{j=1}^{\infty} c_j \mu(E_j) < 2\varepsilon$$

Thus, the results hold true for $f \geq 0$. In general case, write $f = f^+ - f^-$. Find u_+, v_+ for f^+ and u_-, v_- for f^-. Let $u := u_+ - v_-$ and $v := v_+ - u_-$. Then,

$$u = u_+ - v_- \ \leq \ f = f^+ - f^- \ \leq \ v_+ - u_- = v.$$

Consequently, u is upper semicontinous, v is lower semicontinouous, and

$$\int_X (v - u) \, \mathrm{d}\mu < 4\varepsilon. \qquad \blacksquare$$

Chapter 3

L^p-Spaces

3.1 Convex Functions and Inequalities

Definition 3.1.1. A function $\varphi\colon (a,b) \to \mathbb{R}$ is **convex** if for all $x, y \in (a,b)$, and given $\lambda \in [0,1]$,

$$\varphi((1-\lambda)x + \lambda y) \leq (1-\lambda)\varphi(x) + \lambda\varphi(y). \tag{3.1.1}$$

Let $a < x = s < t < y = u < b$, $\lambda = \frac{t-s}{u-s}$. It is equivalent to

$$\begin{aligned} \varphi(t) &\leq \varphi(s)\frac{u-t}{u-s} + \varphi(u)\frac{t-s}{u-s}, \\ (u-s)\varphi(t) &\leq (u-t)\varphi(s) + (t-s)\varphi(u). \end{aligned}$$

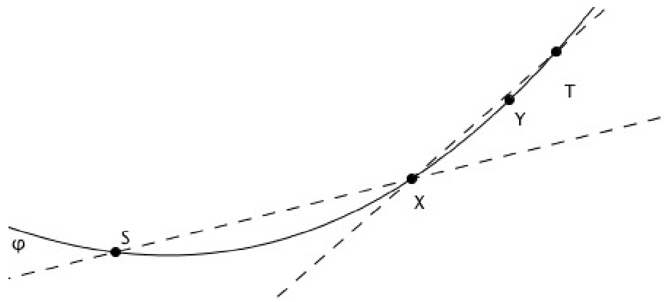
Subtract $(t-s)\varphi(t)$ from both sides and simplify:

$$\begin{aligned} (u-t)(\varphi(t)-\varphi(s)) &\leq (t-s)(\varphi(u)-\varphi(t)) \\ \frac{\varphi(t)-\varphi(s)}{t-s} &\leq \frac{\varphi(u)-\varphi(t)}{u-t}. \end{aligned} \tag{3.1.2}$$

By the Mean Value Theorem for Differentiation, a differentiable function φ is convex in (a,b) iff $\varphi'(s) \leq \varphi'(t), \forall s < t$ if and only if φ' is monotonically increasing.

Proposition 3.1.2. *If φ is convex on (a,b), then φ is continuous on (a,b).*

Proof. Suppose $a < s < x < y < t < b$. Let $S := (s, \varphi(s))$ and similar for x, y, t. Then Y is below the line XT and above SX. Check the picture, as $y \to x$, $Y \to X$, and vice-versa.



■

THEOREM 3.1.3 (Jensen's Inequality). *Let μ be a positive measure on (X, M) with $\mu(X) = 1$. Let $f : X \to [-\infty, \infty], f \in L^1(\mu)$, $a < f(x) < b$, for all $x \in X$, and φ be convex on (a, b). Then,*

$$\varphi\left(\int_X f \,\mathrm{d}\mu\right) \leq \int_X (\varphi \circ f) \,\mathrm{d}\mu. \tag{3.1.3}$$

Proof. For all $t \in (a, b)$ with $a < s \leq t \leq u < b$, inequality (3.1.2) gives

$$\frac{\varphi(t) - \varphi(s)}{t - s} \leq \frac{\varphi(u) - \varphi(t)}{u - t}.$$

Define $\beta := \sup\{\frac{\varphi(t)-\varphi(s)}{t-s} : s \in (a, t)\}$. Then,

$$\frac{\varphi(t) - \varphi(s)}{t - s} \leq \beta \leq \frac{\varphi(u) - \varphi(t)}{u - t}.$$

Hence,

$$\varphi(s) \geq \varphi(t) + \beta(s - t). \tag{3.1.4}$$

Let $s = f(x)$, then inequality (3.1.4) gives

$$\varphi(f(x)) - \varphi(t) + \beta(f(x) - t) \geq 0. \tag{3.1.5}$$

Let $t = \int_X f \,\mathrm{d}\mu$. Then $a < t < b$ because $\mu(X) = 1$. Now, integrating (3.1.5) gives

$$\begin{aligned}
0 &\leq \int_X \varphi \circ f \,\mathrm{d}\mu - \varphi(t) \int_X \mathrm{d}\mu + \beta\left(\int_X f \,\mathrm{d}\mu - t \int_X \mathrm{d}\mu\right) \\
&\leq \int_X \varphi \circ f \,\mathrm{d}\mu - \varphi(t)\mu(X) + \beta(t - t\mu(X)).
\end{aligned}$$

Since $\mu(X) = 1$, it follows that

$$\varphi \circ \int_X f \,\mathrm{d}\mu \leq \int_X \varphi \circ f \,\mathrm{d}\mu.$$

■

Example 3.1.4. Let μ be the probability measure on $S := \{1, \dots, n\}$, $\mu(\{j\}) = \alpha_j > 0$, and $\sum_{j=1}^{n} \alpha_j = 1$. Let $\beta_j = f(j)$, $\varphi(s) = e^s$, which is convex. Then,

$$\varphi\biggl(\int_S f \, \mathrm{d}\mu\biggr) = e^{\sum_{j=1}^{n} \alpha_j \beta_j} = \prod_{j=1}^{n} e^{\alpha_j \beta_j} \leq \int_S \varphi \circ f \, \mathrm{d}\mu = \sum_{j=1}^{n} \alpha_j e^{\beta_j}.$$

Now, let $\gamma_j = e^{\beta_j}$, then $\beta_j = \ln(\gamma_j)$ and we have

$$\prod_{j=1}^{n} \gamma_j^{\alpha_j} \leq \sum_{j=1}^{n} \alpha_j \gamma_j, \tag{3.1.6}$$

whenever $\sum_{j=1}^{n} \alpha_j = 1$. The left and right sides are often called the **geometric mean** and **arithmetic mean**, respectively.

Definition 3.1.5. If $p, q > 0$ such that $p + q = pq$ or $\frac{1}{p} + \frac{1}{q} = 1$, then we call p and q a pair of **conjugate exponents**.

Theorem 3.1.6 (Hölder's and Minkowski's Inequalities). *Let p, q be conjugate exponents, $1 < p < \infty$. Let X be a measure space, with measure μ. Let $f, g \colon X \to [0, \infty]$ be measurable. Then,*

$$\int_X fg \, \mathrm{d}\mu \leq \biggl(\int_X f^p \, \mathrm{d}\mu\biggr)^{1/p} \biggl(\int_X g^q \, \mathrm{d}\mu\biggr)^{1/q} = \|f\|_p \|g\|_q, \tag{3.1.7}$$

and

$$\biggl(\int_X (f+g)^p \, \mathrm{d}\mu\biggr)^{1/p} \leq \biggl(\int_X f^p \, \mathrm{d}\mu\biggr)^{1/p} + \biggl(\int_X g^p \, \mathrm{d}\mu\biggr)^{1/p}. \tag{3.1.8}$$

The inequality (3.1.7) is **Hölder's**; (3.1.8) is **Minkowski's**. [1]

Proof. For the Hölder's inequality, let $A := \|f\|_p, B := \|g\|_q$. If $A = 0$ or ∞, it is trivial. Suppose $0 < A, B < \infty$. Let $F(x) := \frac{1}{A}|f(x)|$, $G(x) := \frac{1}{B}|g(x)|$. Let $\varphi(z) := e^z$. Since e^z ranges over $(0, \infty)$, for every $x \in X$, $F(x) = e^{s/p}$ and $G(x) = e^{t/q}$ for some $s, t \in \mathbb{R}$. By convexity of φ, we have

$$\begin{aligned} e^{s/p + t/q} &\leq \frac{1}{p} e^s + \frac{1}{q} e^t \\ F(x)G(x) &\leq \frac{1}{p}(F(x))^p + \frac{1}{q}(G(x))^q. \end{aligned}$$

Integrating both sides gives

$$\begin{aligned} \int_X FG \, \mathrm{d}\mu &\leq \frac{1}{p} \int_X F^p \, \mathrm{d}\mu + \frac{1}{q} \int_X G^q \, \mathrm{d}\mu \\ \frac{1}{AB} \int_X |fg| \, \mathrm{d}\mu &\leq \frac{1}{p}\biggl(\frac{1}{A^p} \int_X |f|^p \, \mathrm{d}\mu\biggr) + \frac{1}{q}\biggl(\frac{1}{B^q} \int_X |g|^q \, \mathrm{d}\mu\biggr) \end{aligned}$$

[1] Here we use the shorthand notation $\|f\|_p$ for $(\int_X f^p \, \mathrm{d}\mu)^{1/p}$, although we have not yet defined $\|f\|_p$ for the suitable f.

$$\leq \frac{1}{p}\left(\frac{1}{A^p}\right)(A^p) + \frac{1}{q}\left(\frac{1}{B^q}\right)(B^q) = 1$$

$$\int_X |fg| \, \mathrm{d}\mu \leq AB = \|f\|_p \|g\|_q.$$

For Minkowski's inequality, fix p and consider

$$(f+g)^p = (f+g)(f+g)^p = f(f+g)^{p-1} + g(f+g)^{p-1}.$$

By Hölder's inequality,

$$\begin{aligned} \|f+g\|_p^p &\leq \|f\|_p \|(f+g)^{p-1}\|_q + \|g\|_p \|(f+g)^{p-1}\|_q \\ &\leq (\|f\|_p + \|g\|_p) \|(f+g)^{p-1}\|_q. \end{aligned} \tag{3.1.9}$$

It is sufficent to prove the case in which $\|f+g\|_p > 0$ and $\|f\|_p + \|g\|_p < \infty$. Consider the following with $\frac{p}{q} = p-1$:

$$\begin{aligned} \|(f+g)^{p-1}\|_q^q &= \int_X ((f+g)^{p-1})^q \, \mathrm{d}\mu \\ &= \int_X (f+g)^{q(p-1)} \, \mathrm{d}\mu \\ &= \int_X (f+g)^p \, \mathrm{d}\mu = \|f+g\|_p^p. \end{aligned}$$

Hence,

$$\|(f+g)^{p-1}\|_q = \|f+g\|_p^{p/q} = \|f+g\|_p^{p-1}.$$

By convexity of $\varphi(t) := t^p$,

$$\begin{aligned} \varphi\left(\frac{1}{2}f + \frac{1}{2}\right) &\leq \frac{1}{2}\varphi \circ f + \frac{1}{2}\varphi \circ g \\ \left(\frac{f+g}{2}\right)^p &\leq \frac{1}{2}f^p + \frac{1}{2}g^p \\ (f+g)^p &\leq 2^{p-1}f^p + 2^{p-1}g^p \end{aligned}$$

Integrating both sides gives

$$0 < \|f+g\|_p^p \leq 2^{p-1}\left(\|f\|_p^p + \|g\|_p^p\right) < \infty.$$

Thus, we can divide inequality (3.1.9) by $\|f+g\|_p^{p-1}$,

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

3.2 The L^p-Spaces

In this section, (X, M, μ) is a positive measure space.

Definition 3.2.1. Let $0 < p < \infty$, $f\colon X \to \mathbb{C}$ be measurable. Define the ***L^p*-norm** of f by

$$\|f\|_p := \left(\int_X |f|^p \, \mathrm{d}\mu \right)^{1/p}. \tag{3.2.1}$$

Define the ***L^p*-space** of X by $L^p(\mu) := \{f\colon X \to \mathbb{C},\ \text{measurable with } \|f\|_p < \infty\}$.

Remark 3.2.2. If μ is the counting measure on a countable set A, we denote the corresponding L^p-space by $l^p(A)$, or l^p. An element in $l^p(A)$ may be regarded as a complex sequence $x = \{x_n\}$ and

$$\|x\|_p = \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p}.$$

Definition 3.2.3. Let $f\colon X \to [0, \infty]$ be measurable. The **essential supremum** is

$$\text{ess sup}(f) := \inf\{\alpha : \mu(\{x : f(x) > \alpha\}) = 0\}. \tag{3.2.2}$$

Remark. Note that

$$\begin{aligned} \mu(\{x : f(x) > \text{ess sup}(f)\}) &= \mu\Big(\bigcup_{n=1}^{\infty} \{x : f > \text{ess sup}(f) + \frac{1}{n}\} \Big) \\ &= \sum_{n=1}^{\infty} \mu(\{x : f > \text{ess sup}(f) + \frac{1}{n}\}) = 0 \end{aligned}$$

Definition 3.2.4. If $f\colon X \to \mathbb{C}$ is measurable, we define

$$\|f\|_\infty := \text{ess sup}(|f|). \tag{3.2.3}$$

Define $L^\infty(\mu) := \{f\colon X \to \mathbb{C},\ \text{measurable with } \|f\|_\infty < \infty\}$. Sometimes we call the members of L^∞ **essentially bounded measurable functions** on X.

Remark. Hence, for almost all $x \in X$, $|f(x)| \le M \iff M \ge \|f\|_\infty$.

Theorem 3.2.5 (Hölder's Inequality for *L^p*-Spaces). *Let p, q be conjugate exponents, $1 \le p \le \infty$, and $f \in L^p(\mu), g \in L^q(\mu)$. Then, $fg \in L^1(\mu)$ and*

$$\|fg\|_1 \le \|f\|_p \|g\|_q.$$

Proof. For $1 < p < \infty$, it is done in Theorem (3.1.6). Suppose $p = 1, q = \infty$. Let $E := \{x \in X : |g(x)| > \|g\|_\infty\}$, so $\mu(E) = 0$. Then,

$$\|fg\|_1 = \int_X |fg| \, \mathrm{d}\mu = \int_{X \setminus E} |fg| \, \mathrm{d}\mu + \int_E |fg| \, \mathrm{d}\mu$$

$$\begin{aligned} &\leq \|g\|_\infty \cdot \int_{X\setminus E} |f| \,\mathrm{d}\mu + 0 \\ &\leq \|g\|_\infty \cdot \int_X |f| \,\mathrm{d}\mu \\ &= \|f\|_1 \|g\|_\infty < \infty. \end{aligned}$$

Therefore, $fg \in L^1(\mu)$. ■

THEOREM 3.2.6 (**$\triangle$-Inequality for L^p-Spaces**). *Let $1 \leq p \leq \infty$, and $f, g \in L^p(\mu)$. Then,*

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

Proof. For $1 < p < \infty$, it is done in Theorem (3.1.6). The case where $p = 1$ is trivial from $|f+g| \leq |f| + |g|$. So, suppose $p = \infty, f, g \in L^\infty(\mu)$. Define the following sets:

$$\begin{aligned} A &:= \{x \in X : |f(x)| + |g(x)| > \|f\|_\infty + \|g\|_\infty\} \\ B &:= \{x \in X : |f(x) + g(x)| > \|f\|_\infty + \|g\|_\infty\} \end{aligned}$$

Then $|f+g| \leq |f| + |g|$ gives $B \subseteq A$.

By basic set operations,

$$\begin{aligned} x \notin A \quad &\Rightarrow \quad x \notin \{|f(y)| > \|f\|_\infty\} \cap \{|g(y)| > \|g\|_\infty\} \\ &\Rightarrow \quad A \subseteq \{|f(y)| > \|f\|_\infty\} \cup \{|g(y)| > \|g\|_\infty\}. \end{aligned}$$

Hence,

$$\mu(A) \leq \mu(\{|f| > \|f\|_\infty\}) + \mu(\{|g| > \|g\|_\infty\}) = 0 + 0$$

gives $\mu(A) = 0$ and $\mu(B) = 0$. Recall that $\|f+g\|_\infty = \inf\{\alpha : \mu(\{|f+g| > \alpha\}) = 0\}$. In particular, for $\alpha = \|f\|_\infty + \|g\|_\infty$, we have $\|f+g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$. ■

REMARK 3.2.7. For $1 \leq p \leq \infty$, $f \in L^p(\mu), \alpha \in \mathbb{C}$, it is clear that $\alpha f \in L^p(\mu)$. Therefore, the $\triangle$-inequality implies that $L^p(\mu)$ is a **vector space**. Even better, we may define a *distance function* $d(f,g) := \|f-g\|_p$. The only problem here is that when $f = g$ μ-a.e., but $f \not\equiv g$, we have $d(f,g) = 0$.

To make such d into a *metric* on $L^p(\mu)$, we simply partition $L^p(\mu)$ into equivalence classes given by $f \sim g \iff d(f,g) = 0$. In this case, we have a quotient space of $L^p(\mu)$ whose memebers are in the form $[f]$. However, for simplicity, we still view $L^p(\mu)$ as a space of functions, and identify each $[f]$ by its representative f. Therefore, $(L^p(\mu), d)$ is a **metric space**, hence a **normed vector space**.

The following results show that $L^p(\mu)$ is complete with the norm $\|\cdot\|_p$.

LEMMA 3.2.8. *Let $\{f_n\}$ be an Cauchy sequence with respect to $\|\cdot\|_p$. Then there exists a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ which converges pointwise μ-a.e.; that is, $f_{n_k}(x) \to f(x)$, for μ-almost every x.*

Proof. **Case 1:** $1 \leq p < \infty$. By hypothesis, for each $k \in \mathbb{N}$, select f_{n_k} such that

$$\|f_{n_{k+1}} - f_{n_k}\|_p \leq \frac{1}{2^k}.$$

Define $g_m := \sum_{k=1}^m |f_{n_{k+1}} - f_{n_k}|$, and $g := \lim_{m\to\infty} g_m$. Then, $g_m^p \leq g_{m+1}^p \leq \ldots$ and $g_m^p \to g^p$. By Monotone Convergence,

$$\int_X g_m^p \, \mathrm{d}\mu \to \int_X g^p \, \mathrm{d}\mu.$$

By $\triangle$-inequality,

$$\|g_m\|_p = \left\| \sum_{k=1}^m |f_{n_{k+1}} - f_{n_k}| \right\|_p \leq \sum_{k=1}^m \|f_{n_{k+1}} - f_{n_k}\|_p \leq \sum_{k=1}^m \frac{1}{2^k} \leq 1.$$

Therefore, $\|g\|_p \leq 1$ and g is finite μ-a.e. Hence, $\sum_{k=1}^\infty (f_{n_{k+1}} - f_{n_k})$ is **absolutely convergent** μ-a.e. Thus, for μ-almost all x, define

$$f(x) := \lim_{m\to\infty} \sum_{k=1}^m (f_{n_{k+1}}(x) - f_{n_k}(x)) = \lim_{m\to\infty} f_{n_m}(x) - f_{n_1}(x).$$

Assume $f_{n_1}(x) = 0$, μ-a.e., we have

$$\lim_{k\to\infty} f_{n_k} = f \in L^p(\mu).$$

Case 2: $p = \infty$. Define the following sets:

$$E_{k,n} := \{x \in X : |f_n(x) - f_k(x)| > \|f_n - f_k\|_\infty\}.$$

Then, $\mu(E_{n,k}) = 0$. Let $E := \bigcup_{n,k\in\mathbb{N}} E_{k,n}$, with $\mu(E) = 0$.

On E^c, $|f_n(x) - f_k(x)| \leq \|f_n - f_k\|_\infty$. By hypothesis, $\{f_n\}$ is L^∞-Cauchy. Therefore, the sequence **converges uniformly** μ-a.e. ■

Theorem 3.2.9. *$L^p(\mu)$ is a **complete** metric space with the p-norm. Hence, every L^p-Cauchy sequence $\{f_n\}$ converges to $f \in L^P(\mu)$.*

Proof. If $p = \infty$, it is done by the previous Lemma by uniform convergence. We may define $f(x) = 0$, for every $x \in E$.

For $1 \leq p < \infty$, supppose $\{f_n\}$ is L^p-Cauchy. By the previous Lemma, there exists subsequence $f_{n_k} \to f$ pointwise μ-a.e. We will show that $f \in L^p(\mu)$, and $f_n \xrightarrow{L^p} f$.

For every fixed $n \in \mathbb{N}$, define $g_k := |f_n - f_{n_k}|^p$. By Fatou's Lemma (1.6.9),

$$\begin{aligned}
\liminf_{k\to\infty} \int_X |f_n - f_{n_k}|^p \, \mathrm{d}\mu &\geq \int_X \liminf_{k\to\infty} |f_n - f_{n_k}|^p \, \mathrm{d}\mu \\
&= \int_X |f_n - f|^p \, \mathrm{d}\mu. \qquad (3.2.4)
\end{aligned}$$

On the other hand, since $\{f_n\}$ is L^p-Cauchy, given $\varepsilon > 0$, there is $N \in \mathbb{N}$ such that for all $n, m > N$,

$$\int_X |f_n - f_m|^p \,\mathrm{d}\mu = \|f_n - f_m\|_p^p < \varepsilon^p.$$

For $n_k > N$, with $m = n_k$, by inequality (3.2.4),

$$\begin{aligned}\int_X |f_n - f|^p \,\mathrm{d}\mu &\le \liminf_{k\to\infty} \|f_n - f_{n_k}\|_p^p \\ &\le \lim_{m\to\infty} \|f_n - f_m\|_p^p \\ &< \varepsilon^p.\end{aligned}$$

for all $n > N$. Hence this shows that

$$\left(\int_X |f_n - f|^p \,\mathrm{d}\mu\right)^{1/p} = \|f_n - f\|_p < \varepsilon,$$

for all $n > N$. Finally, to see $f \in L^p(\mu)$, note that for some $n > N$,

$$\begin{aligned}\|f\|_p &= \|f - f_n + f_n\|_p \\ &\le \|f - f_n\|_p + \|f_n\|_p \\ &\le \varepsilon + \|f_n\|_p < \infty.\end{aligned}$$

Therefore, $f \in L^p(\mu)$ and $f_n \xrightarrow{L^p} f$. ■

3.3 More on L^p-Spaces

In general, for all $p \neq q$, $L^p(\mu) \not\subset L^q(\mu)$. For example, consider the Lebesgue measure m on $(0, \infty)$. Given $q > 0$, define $f\colon (0,\infty) \to \mathbb{R}$, by $f_q(x) := x^{-q}$. Then, $f\chi_{(0,1)} \in L^p$ if and only if $p < q^{-1}$, and $f\chi_{(1,\infty)} \in L^p$ if and only if $p > q^{-1}$. However, under certain conditions, we do have inclusion. In this section, we write L^p for $L^p(\mu)$.

Proposition 3.3.1. *If $0 < p < q < r \le \infty$, then $L^q = L^p + L^r$. That is, for all $f \in L^q$, there is $g \in L^P$ and $h \in L^r$ such that $f = g + h$.*

Proof. Let $f \in L^q$, and define $E := \{x : |f(x)| \ge 1\}$. Define $g := f\chi_E$, and $h := f\chi_{E^c}$. Note that $|g|^p = |f|^p\chi_E \le |f|^q\chi_E$, thus $\|g\|_p \le \|f\|_q < \infty$, and $g \in L^p$. Similarly, $h \in L^r$. ■

Proposition 3.3.2. *If $0 < p < q < r \le \infty$, then $L^p \cap L^r \subset L^q$. Moreover, for all $f \in L^p \cap L^r$, we have*

$$\|f\|_q \le \|f\|_p^\lambda \|f\|_r^{1-\lambda}, \tag{3.3.1}$$

where $0 < \lambda < 1$ is given by $q^{-1} = \lambda p^{-1} + (1-\lambda) r^{-1}$.

Proof. First suppose $r = \infty$. Then $q^{-1} = \lambda p^{-1}$, and for almost every $x \in X$,

$$|f(x)|^q = |f(x)|^{q-p}|f(x)|^p \leq \|f\|_\infty^{q-p}|f(x)|^p.$$

Integrating both sides and take the q^{th}-root,

$$\begin{aligned} \|f\|_q &\leq \|f\|_\infty^{(1-p/q)}\left(\int_X |f|^p \,\mathrm{d}\mu\right)^{1/q} = \|f\|_\infty^{(1-p/q)}\|f\|_p^{p/q} \\ &= \|f\|_p^\lambda \|f\|_r^{1-\lambda}. \end{aligned}$$

Now suppose $r < \infty$. Note that

$$\frac{\lambda q}{p} + \frac{(1-\lambda)q}{r} = 1.$$

Hence, $p' := p/(\lambda q)$ and $q' := r/(q(1-\lambda))$ are conjugate expontents. By Hölder's inequality (3.1.6),

$$\begin{aligned} \int_X |f|^q \,\mathrm{d}\mu &= \int_X |f|^{\lambda q}|f|^{(1-\lambda)q} \,\mathrm{d}\mu \leq \||f|^{\lambda q}\|_{p'}\||f|^{(1-\lambda)q}\|_{q'} \\ &= \left(\int_X (|f|^{\lambda q})^{p'} \,\mathrm{d}\mu\right)^{1/p'}\left(\int_X (|f|^{(1-\lambda)q})^{q'} \,\mathrm{d}\mu\right)^{1/q'} \\ &= \|f\|_p^{\lambda q}\|f\|_r^{(1-\lambda)q}. \end{aligned}$$

Take the q^{th}-root on both sides, we obtain inequality (3.3.1), and $f \in L^q$. ■

Proposition 3.3.3. *Let A be a nonempty set, $0 < p < q \leq \infty$. Then, $l^p(A) \subset l^q(A)$ and for all $f \in l^p(A)$, $\|f\|_q \leq \|f\|_p$.*

Proof. If $q = \infty$, then

$$\|f\|_\infty^p = (\sup_{\alpha \in A}|f(\alpha)|)^p \leq \sum_{\alpha \in A}|f(\alpha)|^p = \|f\|_p^p.$$

Thus, $\|f\|_\infty \leq \|f\|_p$. If $q < r := \infty$, by Proposition (3.3.2), we have $\lambda = p/q$, and

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

Proposition 3.3.4. *Let $\mu(X) < \infty$, and $0 < p < q \leq \infty$. Then $L^q \subset L^p$, and for all $f \in L^q$,*

$$\|f\|_p \leq \|f\|_q \mu(X)^{(1/p)-(1/q)}. \tag{3.3.2}$$

Proof. If $q = \infty$, then $1/q = 0$ and

$$\|f\|_p = \left(\int_X |f|^p \,\mathrm{d}\mu\right)^{1/p} \leq \|f\|_\infty \mu(X)^{1/p}.$$

If $q < \infty$, observe that $p' := q/p$ and $q' := q/(q-p)$ are conjugate exponents. Hence, by Hölder's inequality,

$$\begin{aligned}\|f\|_p^p &= \int_X |f|^p \, d\mu \le \||f|^p\|_{p'} \|1\|_{q'} \\ &= \left(\int_X (|f|^p)^{p'} \, d\mu \right)^{1/p'} \mu(X)^{1/q'} \\ &= \|f\|_q^p \mu(X)^{1-p/q}.\end{aligned}$$

Taking the p^{th}-root on both sides, we obtain inequality (3.3.2). ■

3.4 Approximations in L^p-Spaces

Proposition 3.4.1. *Let*

$$S := \{s \colon X \to \mathbb{C} \mid s \textit{ simple, measurable }, \mu(\{s \neq 0\}) < \infty\}.$$

Then for $1 \le p < \infty$, $S \subset L^p(\mu)$ is L^p-dense in $L^p(\mu)$.

Proof. Obviously, for all $s \in S$, $\|s\|_p < \infty$, and $S \subset L^p(\mu)$. We will show that whenever $f \in L^p(\mu)$, there is $s_n \xrightarrow{L^p} f$.

First, suppose $f \colon X \to [0, \infty)$. Using the staircase functions, there is a sequence of simple functions $s_n \nearrow f$, pointwise. Since $0 \le s_n \le f$, we have $s_n \in L^p(\mu)$, hence $s_n \in S$. Note that $|f - s_n|^p \le |f|^p$ and by Dominated Convergence, we have

$$\lim_{n\to\infty} \int_X |f - s_n|^p \, d\mu = \int_X \lim_{n\to\infty} |f - s_n| \, d\mu = 0.$$

Hence, $\|f - s_n\|_p \to 0$, or equivalently, $s_n \xrightarrow{L^p} f$. In general, if $f \colon X \to \mathbb{C}$, write $f = (u_+ - u_-) + i(v_+ - v_-)$. From the preceding step we obtain corresponding sequences $\{s_n^+\}, \{s_n^-\}, \{t_n^+\}, \{t_n^-\}$. Then apply the $\triangle$-inequality. ■

Approximation by Continuous Functions

Now let μ be a measure on a locally compact Hausdroff space X, which has the **five properties stated in Riesz Representation Theorem**. In particular, μ could be the Lebesgue measure on $\mathbb{R}^k$.

Theorem 3.4.2. *For $1 \le p < \infty$, $C_c(X)$ is L^p-dense in $L^p(\mu)$.*

Proof. From Propoisition (3.4.1), it suffices to show that $C_c(X)$ is dense in S. For every $s \in S$, define $A := \{s \neq 0\}$. By definition of $S, \mu(A) < \infty$ and $s(x) = 0$, for all $x \in A^c$. Therefore, by Lusin's Theorem (2.5.1), given $\varepsilon > 0$, there is $g \in C_c(X)$ such that

$$|g| \le \sup\{s(x) : x \in X\} = \|s\|_\infty \qquad \text{and} \qquad \mu(\{g \neq s\}) < \varepsilon.$$

Then, using $|g - s| < 2\|s\|_\infty$, we have

$$\|g - s\|_p = \left(\int_X |g - s|^p \, \mathrm{d}\mu \right)^{1/p} = \left(\int_{\{g \neq s\}} |g - s|^p \, \mathrm{d}\mu \right)^{1/p}$$
$$\leq 2\|s\|_\infty \left(\int_X \chi_{\{g \neq s\}} \, \mathrm{d}\mu \right)^{1/p} = 2\|s\|_\infty \varepsilon^{1/p}$$

Therefore, $C_c(X)$ is dense in S, hence in $L^p(\mu)$. ■

Remark 3.4.3. Consider the Lebesgue measure m on $\mathbb{R}^k$. For $1 \leq p \leq \infty$, the metric $\|f - g\|_p$ on $C_c(\mathbb{R}^k)$ is a genuine metric, i.e. we do not have to pass to equivalence classes. It is because if $f \not\equiv g$, then they must differ on some open set U, and $m(U) > 0$. Hence, if $\|f - g\|_p = 0$, then $f = g, m$-a.e., and $f \equiv g$. Also, note that in $C_c(\mathbb{R}^k)$, $\|f\|_\infty = \sup\{|f(x)| : x \in \mathbb{R}^k\}$.

For $1 \leq p < \infty$, $L^p(\mathbb{R}^k)$ is the L^p-completion of $C_c(\mathbb{R}^k)$ by Theorem (3.4.2). In particular, when $p = 1$, $f \in C_c(\mathbb{R}^k)$, $\|f\|_1$ is precisely the Reimann integral $\int_{\mathbb{R}^k} |f(x)| \, \mathrm{d}x$; and $L^1(\mathbb{R}^k)$ is the L^1-completion of $C_c(\mathbb{R}^k)$.

However, when $p = \infty$, the L^∞-completion of $C_c(\mathbb{R}^k)$ is not $L^\infty(\mathbb{R}^k)$, but $C_0(\mathbb{R}^k)$, *the space of continuous functions on* $\mathbb{R}^k$ *which vanish at infinity.* We shall see that in Proposition (3.4.5).

Definition 3.4.4. A complex function f on a locally compact Hausdorff space X is said to **vanish at infinity** if given $\varepsilon > 0$, there exists a compact set $K \subseteq X$ such that $|f(x)| < \varepsilon$, for all $x \notin K$.

We denote $C_0 := \{f \colon X \to \mathbb{C} \mid \text{continous } f \text{ vanishes at infinity.}\}$. Obviously $C_c(X) \subseteq C_0(X)$, and they are equal if X is compact. In this case we simply denote it as $C(X)$.

Proposition 3.4.5. *If X is a locally compact Hausdorff space, then $C_0(X)$ is the completion of $C_c(X)$, relative to the metric defined by*

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

Proof. It is obvious that $(C_0(X), \|\cdot\|)$ is a metric space with $d(f, g) = \|f - g\|$. We will show that $C_c(X)$ is dense in $C_0(X)$, and $C_0(X)$ is complete.

For density, let $f \in C_0(X)$ and $\varepsilon > 0$. By definition of $C_0(X)$, there is a compact $K \subset X$ such that $|f(x)| < \varepsilon$ ouside K. By Urysohn's Lemma (2.2.11), there is $g \in C_c(X)$ such that $0 \leq g \leq 1$ and $g(x) = 1$ on K. Define $h := fg$, then $h \in C_c(X)$ and $\|f - h\| < \varepsilon$.

For completion, let $\{f_n\}$ be Cauchy in $C_0(X)$. Hence, $\{f_n\}$ convereges uniformly because $|f_n(x) - f_m(x)| < \|f_n - f_m\|$. Therefore, the pointwise limit f is continuous. To see $f \in C_0(X)$, given $\varepsilon > 0$, there is n such that $\|f_n - f\| < \varepsilon/2$. Also there is a compact $K \subset X$ so that $|f_n(x)| < \varepsilon/2, \forall x \notin K$. Hence, $\forall x \notin K$,

$$|f(x)| = |f(x) - f_n(x) + f_n(x)| \leq |f(x) - f_n(X)| + |f_n(x)| < \varepsilon.$$

Thus f vanishes at infinity and $C_0(X)$ is complete. ■

3.5 Additional : Egoroff's Theorem

We conclude this chapter with the Egoroff's Theorem concerning on uniform convergence of a sequence of measurable functions. This type of convergence is sometimes called **almost uniform convergence**. Moreover, one can prove Lusin's Theorem (2.5.1) easily, using Egoroff's Theorem and Tietze Extension Theorem.

Theorem 3.5.1 (Egoroff's Theorem). *Let $\mu(X) < \infty$, $f_n \colon X \to \mathbb{C}$ be measurable and $f_n(x) \to f(x)$, for μ-almost every x. Then given $\varepsilon > 0$, there is a measurable set $E \subset X$ with $\mu(E^c) < \varepsilon$ such that $\{f_n\}$ converges uniformly on E.*

Proof. Without loss of generality, we assume $f_n(x) \to f(x)$, for all $x \in X$. Define

$$S_{n,k} := \bigcap_{i,j \geq n} \left\{ x \in X : |f_i(x) - f_j(x)| < \frac{1}{k} \right\}.$$

For all $x \in X$, $\{f_n(x)\}$ is Cauchy. Thus, for all $k \in \mathbb{N}$, $S_{1,k} \subseteq S_{2,k} \subseteq \cdots \to X$. Hence, for each $k \in \mathbb{N}$, there is $n_k \in \mathbb{N}$ such that

$$\begin{aligned} \mu(S_{n_k,k}) > \mu(X) - \frac{\varepsilon}{2^k} \quad &\Rightarrow \quad \mu(S^c_{n_k,k}) < \frac{\varepsilon}{2^k} \\ &\Rightarrow \quad \mu\left(\bigcup_{k=1}^{\infty} S^c_{n_k,k} \right) < \varepsilon. \end{aligned}$$

Choose such pair (k, k_n) and define $E := \bigcap_{k=1}^{\infty} S_{n_k,k}$. We will show that $\{f_n\}$ converges uniformly on E. For every $\varepsilon > 0$, choose $k \in \mathbb{N}$ with $\frac{1}{k} < \varepsilon$, and let $N := n_k$. Then if $x \in E$, $x \in S_{N,k}$ gives

$$|f_i(x) - f_j(x)| < \frac{1}{k} < \varepsilon, \quad \text{for all } i, j > N.$$

Therefore, $\{f_n\}$ is uniformly Cauchy, hence uniformly convergent to f on E. ∎

Remark 3.5.2. Egoroff's Theorem does not hold in σ-finite space. For example, let $X := [0, \infty)$ with the Lebesgue measure m. Define $f_n = \chi_{[n,\infty)}$, then $f_n(x) \to 0$.

Let $\varepsilon = \frac{1}{2}$ and take E such that $\mu(E^c) < \frac{1}{2}$. Then for each $n > k$, $[k, n) \cap E \neq \varnothing$ because $\mu([k, n)) \geq 1$. However, then there is $x \in E$, such that $f_k(x) = 1$, and $f_n(x) = 0$, for all $n > k$. Hence,

$$|f_k(x) - f_n(x)| = 1$$

and $\{f_n\}$ does not converge at x. ∎

Theorem 3.5.3 (Tietze Extension). *Let (X, τ) be a locally compact Hausdorff space, $K \subset X$ be compact. Suppose $f \colon K \to \mathbb{C}$ is continuous, then there is $g \in C_c(X)$ such that $g|_K = f$. Moreover, supp(g) is a subset of some open U.*

Proof. It suffices to prove for real-valued f. Since f is continuous on K, f is bounded on K. Without loss of generality, we assume $f(K) \subseteq [-2, 2]$. We shall proceed by induction.

Step 1: Define $A_1 := f^{-1}([-2, -\frac{2}{3}])$ and $B_1 := f^{-1}([\frac{2}{3}, 2])$. Note that A_1, B_1 are closed in K, hence closed and compact in X, with $B_1 \subset X \setminus A_1$. By Urysohn's Lemma (2.2.11), there is $h_1 \in C_c(X)$, such that $h_1|_{B_1} = 1$, $h_1|_{A_1} = 0$, and $h_1(X) = [0, 1]$. Define $g_1 = \frac{1}{2}(\frac{2}{3})(h_1 - \frac{1}{2})$. Then $g_1 \in C_c(X)$, and $g_1|_{B_1} = \frac{2}{3}$, $g_1|_{A_1} = -\frac{2}{3}$, and $g_1(X) = [-\frac{2}{3}, \frac{2}{3}]$. Let $f_1 := f$, and $f_2 := f_1 - g_1$ on K. Then, $f_2 \in C_c(K)$ and $f_2(K) = [-\frac{4}{3}, \frac{4}{3}]$.

Step 2: Let $A_2 := f_2^{-1}([-\frac{4}{3}, -(\frac{2}{3})^2])$ and $B_2 := f_2^{-1}([(\frac{2}{3})^2, \frac{4}{3}])$. There exists $h_2 \in C_c(X)$, such that $h_2|_{B_2} = 1$, $h_2|_{A_2} = 0$, and $h_2(X) = [0, 1]$. Let $g_2 := \frac{1}{2}(\frac{2}{3})^2(h_2 - \frac{1}{2})$. Then, $g_2|_{B_2} = (\frac{2}{3})^2$, $g_2|_{A_2} = -(\frac{2}{3})^2$, and $g_2(X) = [-(\frac{2}{3})^2, (\frac{2}{3})^2]$. Define $f_3 := f_2 - g_2$ in $C_c(K)$ and $f_3(K) = [-3(\frac{2}{3})^3, 3(\frac{2}{3})^3]$.

Step 3: Proceed inductively. For each n, we obtain $g_n \in C_c(X)$ so that $g_n(X) = [-(\frac{2}{3})^n, (\frac{2}{3})^n]$, and $f_n \in C_c(K)$ with $f_n(K) = [-3(\frac{2}{3})^n, 3(\frac{2}{3})^n]$. Also, on K, $g_n = f_n - f_{n+1}$. Finally, define $g(x) := \sum_{n=1}^{\infty} g_n(x)$, for all $x \in X$. Since $|g_n| \leq (\frac{2}{3})^n$, $\sum_{n=1}^{\infty} g_n$ converges uniformly to g. Hence, $g \in C_c(X)$. Moreover on K,

$$g = \lim_{N\to\infty} \sum_{n=1}^{N} (f_n - f_{n+1}) = \lim_{N\to\infty} f_1 - f_{N+1} = f,$$

because $|f_N| \leq 3(\frac{2}{3})^N \to 0$. ■

Theorem (Lusin's Theorem). *Let (X, M, μ) be a locally compact Hausdorff space, and μ be a regular Borel measure. Suppose $f : X \to \mathbb{C}$ is measurable, $A := \{x : f(x) \neq 0\}$ and $\mu(A) < \infty$. Then, given any $\varepsilon > 0$, there exists $g \in C_C(X)$ such that*

$$\mu(\{x : f(x) \neq g(x)\} < \varepsilon, \tag{3.5.1}$$

and

$$\sup_{x\in X} |g(x)| \leq \sup_{x\in X} |f(x)|. \tag{3.5.2}$$

Proof. Pick a sequence of simple functions $s_n(x) \nearrow f(x)$, for all $x \in X$. By the Egoroff's Theorem (3.5.1) on A, there is $E \in M$ such that $\mu(A \setminus E) < \frac{\varepsilon}{3}$, and $s_n|_E \to f|_E$, uniformly.

By regularity of μ, there is compact K and open U, such that $K \subset E \subset A \subset U$, with $\mu(E \setminus K) < \frac{\varepsilon}{3}$, and $\mu(U \setminus A) < \frac{\varepsilon}{3}$. By the Tietze Extension (3.5.3), for each n, there is $g_n \in C_c(X)$ such that $g_n|_K = s_n|_K \to f|_K$, uniformly.

Hence, $f|_K \in C_c(K)$. Apply the Tietze Extension again on $f|_K$, we find $g \in C_c(X)$, such that $g|_K = f|_K$ and $\text{supp}(g) \subset U$. Moreover, $\{x \in A : g(x) \neq f(x)\} \subset U \setminus K$, which has measure $< \varepsilon$.

Finally, we obtain inequality (3.5.2) exactly as in the proof of Theorem (2.5.1). ■

Chapter 4

Elementary Hilbert Space Theory

4.1 Inner Products and Linear Functionals

Definition 4.1.1. Let H be a complex vector space. A **sesquilinear form** is a map $\langle \cdot, \cdot \rangle \colon H \times H \to \mathbb{C}$ satisfying the followings: For all $x, y, z \in H$, $\alpha \in \mathbb{C}$,

(a) $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

(b) $\langle x + \alpha z, y \rangle = \langle x, y \rangle + \alpha \langle z, y \rangle$.

(c) $\langle x, x \rangle \geq 0$. (**Positive Semidefinite**)

In addition, if $\langle x, x \rangle = 0$ if and only $x = 0$, then H is called an **inner product space**.

Remark. (a) implies $\langle x, x \rangle$ is positive real. (b) implies the map $x \mapsto \langle x, y \rangle$ is a linear functional on H. (a) and (b) imply that $\langle x, y + \alpha z \rangle = \langle x, y \rangle + \overline{\alpha} \langle x, z \rangle$.

Proposition 4.1.2. *If $\langle \cdot, \cdot \rangle$ is positive semidefinite, and $\langle x, x \rangle = 0$, then $\langle x, y \rangle = 0$, for each $y \in H$.*

Proof. For all $\alpha \in \mathbb{C} \setminus \{0\}$,

$$\begin{aligned}
\langle x + \alpha y, x + \alpha y \rangle &= \langle x, x + \alpha y \rangle + \alpha \langle y, x + \alpha y \rangle \\
&= \langle x, x \rangle + \overline{\alpha} \langle x, y \rangle + \alpha \langle y, x \rangle + |\alpha|^2 \langle y, y \rangle \\
&= 0 + 2 \operatorname{Re}(\alpha \langle y, x \rangle) + |\alpha|^2 \langle y, y \rangle \\
&= 2 \operatorname{Re}(\alpha \langle y, x \rangle) + |\alpha|^2 \langle y, y \rangle \geq 0
\end{aligned}$$

By way of contradiction, suppose $\langle x, y \rangle \neq 0$. Then either (1): $\langle y, y \rangle = 0$ and $2 \operatorname{Re}(\alpha \langle y, x \rangle) \geq 0$, or (2): $\langle y, y \rangle \neq 0$.

For (1), let $\alpha := \frac{-1}{\langle y, x \rangle}$, then $2 \operatorname{Re}(\alpha \langle y, x \rangle) = -2 \geq 0$. $\rightarrow\leftarrow$

For (2), let $\alpha := -\frac{|\langle x, y \rangle|^2}{\langle y, x \rangle \langle y, y \rangle}$, then $|\alpha| = \frac{|\langle x, y \rangle|}{\langle y, y \rangle}$, and

$$2 \operatorname{Re}(\alpha \langle y, x \rangle) + |\alpha|^2 \langle y, y \rangle = -\frac{|\langle x, y \rangle|^2}{\langle y, y \rangle} \geq 0. \quad \rightarrow\leftarrow$$

■

Definition 4.1.3. If $\langle \cdot, \cdot \rangle$ is a sesquilinear positive semidefinite form on H, then the **seminorm** of $x \in H$ is defined to be

$$\|x\| := \sqrt{\langle x, x \rangle}.$$

Proposition 4.1.4 (Cauchy-Schwarz Inequality). *If $\langle \cdot, \cdot \rangle$ is a sesquilinear positive semidefinite form on H, then for all $x, y \in H$,*

$$|\langle x, y \rangle| \leq \|x\| \, \|y\|.$$

Proof. Suppose $\|y\| \neq 0$. Let $\lambda := \frac{\langle x,y \rangle}{\langle y,y \rangle}$. Then,

$$\begin{aligned} \langle x - \lambda y, x - \lambda y \rangle &= \|x\|^2 - 2\operatorname{Re}(\lambda \langle y, x \rangle) + |\lambda|^2 \|y\|^2 \\ &= \|x\|^2 - 2\operatorname{Re}\left(\frac{|\langle x, y \rangle|^2}{\langle y, y \rangle}\right) + \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle} \quad \geq 0. \end{aligned}$$

Hence, $\|x\|^2 \geq \frac{|\langle x,y \rangle|^2}{\langle y,y \rangle}$ and $|\langle x, y \rangle| \leq \|x\| \, \|y\|$. ∎

Proposition 4.1.5 ($\triangle$-inequality). *For the seminorm, for all $x, y \in H$,*

$$\|x + y\| \leq \|x\| + \|y\|. \tag{4.1.1}$$

Proof. By the Cauchy-Schwarz Inequality,

$$\begin{aligned} \|x + y\|^2 = \langle x + y, x + y \rangle &= \|x\|^2 + 2\operatorname{Re}(\langle x, y \rangle) + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

Hence, $\|x + y\| \leq \|x\| + \|y\|$. ∎

Remark 4.1.6. If H is an inner product space, then $\|x\|$ is a norm. The metric $d(x, y) := \|x - y\|$ gives a metric topology on H.

Proposition 4.1.7 (Parallelogram Law). *For the seminorm, for all $x, y \in H$,*

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2. \tag{4.1.2}$$

(The sum of the squares of the diagonals of a parallelogram is equal to the sum of the squares of its sides.)

Proof. Sum the identities $\|x \pm y\|^2 = \|x\|^2 \pm 2\operatorname{Re}(\langle x, y \rangle) + \|y\|$. ∎

Proposition 4.1.8 (Polarization Identity)**.** *For all $x, y \in H$,*

$$4\langle x, y \rangle = \|x + y\|^2 - \|x - y\|^2 + i(\|x + iy\|^2 - \|x - iy\|^2). \tag{4.1.3}$$

Proof. First suppose H is a real vector space. Then $\langle x, y \rangle = \langle y, x \rangle$. Consider

$$\begin{aligned} \|x + y\|^2 - \|x - y\|^2 &= \langle x + y, x + y \rangle - \langle x - y, x - y \rangle \\ &= \|x\|^2 + \langle x, y \rangle + \langle y, x \rangle + \|y\|^2 - (\|x\|^2 - \langle x, y \rangle - \langle y, x \rangle + \|y\|^2) \end{aligned}$$

$$\begin{aligned} &= 2\langle x, y\rangle + 2\langle y, x\rangle \\ &= 4\langle x, y\rangle. \end{aligned}$$

If H is a complex vector space, then from the previous calculations,

$$\|x+y\|^2 - \|x-y\|^2 = 2\langle x, y\rangle + 2\langle y, x\rangle,$$

and

$$\|x+iy\|^2 - \|x-iy\|^2 = 2\langle x, iy\rangle + 2\langle iy, x\rangle.$$

Subsituting into formula (4.1.3), the RHS gives

$$\begin{aligned} RHS &= 2\langle x, y\rangle + 2\langle y, x\rangle - 2(i^2)\langle x, y\rangle + 2(i^2)\langle y, x\rangle \\ &= 2\langle x, y\rangle + 2\langle y, x\rangle + 2\langle x, y\rangle - 2\langle y, x\rangle \\ &= 4\langle x, y\rangle. \end{aligned}$$

■

Definition 4.1.9 (**Hilbert Space**). Let H be an inner product space. If H is complete with respect to $\|\cdot\|$, then H is an **Hilbert Space**.

Example 4.1.10. If μ is any positive measure, $L^2(\mu)$ is an inner product space with

$$\langle f, g\rangle := \int_X f\overline{g}\,\mathrm{d}\mu. \tag{4.1.4}$$

Note that

$$\|f\| = \langle f, f\rangle^{1/2} = \left(\int_X |f|^2\,\mathrm{d}\mu\right)^{1/2} = \|f\|_2.$$

Also, recall that $L^p(\mu)$ is complete for $1 \le p < \infty$. Hence, $L^2(\mu)$ is a Hilbert space.

Throughout this chapter, let H be a Hilbert space.

Proposition 4.1.11. *Let $g \in H$, then $\lambda_g \colon H \to \mathbb{C}$ given by $\lambda_g(f) := \langle f, g\rangle$ is a linear functional, and uniformly continuous. Consequently, the maps $f \mapsto \langle g, f\rangle$ and $f \mapsto \|f\|$ are also uniformly continuous.*

Proof. Linearity is done previously. For uniform continuity, $\forall \varepsilon > 0$, if $g = 0$, then $\lambda_g(f) = 0$. If $g \neq 0$, pick $\delta = \frac{\varepsilon}{\|g\|}$. Then, $\forall f, h \in H$, with $\|f - h\| < \frac{\varepsilon}{\|g\|}$, we have

$$|\lambda_g(f) - \lambda_g(h)| = |\langle f, g\rangle - \langle h, g\rangle| = |\langle f - h, g\rangle| \le \|f - h\|\|g\| < \varepsilon.$$

■

Definition 4.1.12. A **closed subspace** of H is a subspace that is a closed set under the metric topology of H.

Remark. If M is a closed space of H, so is its closure $\overline{M}$. To see it, pick convergent sequences $\{x_n\}, \{y_n\}$ in H, and $\alpha \in \mathbb{C}$. It is easy to see that $\alpha x_n + y_n \to \alpha x + y \in \overline{M}$.

Definition 4.1.13 (Convex Sets)**.** A set E in a complex vector space V is said to be **convex** if $\forall x, y \in E, \forall t \in (0,1)$,

$$z_t := (1-t)x + ty \quad \in E.$$

One may visualize z_t as a stright line segement from x to y, lying inside E. Obviously, every subspace of V is convex. Also, if E is convex, so is the translate $E + x := \{y + x : y \in E\}$.

4.2 Orthogonality

Definition 4.2.1 (**Orthogonality**)**.** We say $x, y \in H$ are **orthogonal** if $\langle x, y\rangle = 0$; we denote it as $x \perp y$. If $S \subset H$, we write $S^\perp := \{x \in H : \langle x, y\rangle = 0, \forall y \in S\}$.

Proposition 4.2.2. *Let $S \subset H$, then $S^\perp$ is a closed subspace of H.*

Proof. Let $z \in H$, and $\lambda_z(x) := \langle z, x\rangle$. Then $\{z\}^\perp = \lambda_z^{-1}\{0\}$. Observe that $\{z\}^\perp$ is closed by continuity, and is a subspace by linearity. Now, note that

$$S^\perp = \bigcap_{z \in S}\{z\},$$

which is a closed subspace. ■

Remark. There is a subspace that is not closed. For example, $C([0,1]) \subset L^2([0,1])$. There exists a sequence of continuous functions converges to a non-continuous function with respect to the L^2-norm.

Lemma 4.2.3. *Let M be a closed subspace in H. Then for all $h \in H$, there is $m \in M$ that is nearest to h.*

Proof. For every $h \in H$, define $\delta := \inf\{\|m - h\| : m \in M\}$. Let $\{m_i\}_{i=1}^\infty$ be a sequence such that $\|m_i - h\| \to \delta$. We will show that $\{m_i\}_{i=1}^\infty$ is Cauchy.

Recall the Parallelogram law, $\|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 - \|x+y\|^2$. Let $x = m_i - h$, $y = m_j - h$. Then,

$$\frac{x+y}{2} = \frac{m_i + m_j}{2} - h \qquad \text{and} \qquad \left\|\frac{x+y}{2}\right\| \geq \delta.$$

Substitute x, y, we have

$$\begin{aligned}\|m_i - m_j\|^2 &= 2(\|m_i - h\|^2 + \|m_j - h\|^2) - \|m_i + m_j - 2h\|^2 \\ &\leq 2\|m_i - h\|^2 + 2\|m_j - h\|^2 - 4\delta^2.\end{aligned}$$

Since $\|m_i - h\|^2, \|m_j - h\|^2 \searrow \delta^2$, given $\varepsilon > 0$, $\exists N \in \mathbb{N}$, such that for all $i, j > N$, we have

$$\|m_i - m_j\| < \varepsilon.$$

Therefore, $\{m_i\}_{i=1}^\infty$ is Cauchy and converges to $m \in M$ because M is closed. Hence, we have $\|m - h\| = \delta$, which is the minimum by definition. ■

Theorem 4.2.4 (**Orthogonal Projections**). *If M is a closed subspace of H, then $\forall h \in H$, there is a unique pair $m \in M$, $n \in M^\perp$ such that $h = m + n$ and $\|h\|^2 = \|m\|^2 + \|n\|^2$. Moreover, the maps $P\colon h \mapsto m$, and $Q\colon h \mapsto n$ are linear. We write $m = Ph$ and $n = Qh$.*

Proof. Fix $h \in H$, by the previous Lemma, pick $m \in M$ that is nearest to h, and let $n := h - m$. We will show that $n \in M^\perp$. For every $x \in M$, given $\alpha \in \mathbb{C}$,

$$\|n - \alpha x\|^2 = \|n\|^2 - 2\operatorname{Re}(\alpha\langle x, n\rangle) + |\alpha|^2\|x\|^2.$$

Suppose $\langle n, x\rangle \neq 0$. Let $\alpha = \alpha(t) = \frac{t}{\langle x,n\rangle}$, $t \in \mathbb{R}$. Then,

$$\|n - \alpha x\|^2 = \|n\|^2 - 2t + \frac{t^2\|x\|^2}{|\langle x, n\rangle|^2}. \tag{4.2.1}$$

However, for sufficiently small t, $2t > \frac{t^2\|x\|^2}{|\langle x,n\rangle|^2}$, and equation (4.2.1) gives

$$\|n - \alpha x\|^2 < \|n\|^2 \quad\Rightarrow\quad \|h - (m + \alpha x)\|^2 < \|h - m\|^2,$$

which contradicts to m being the nearest point in M. Hence, $\langle n, x\rangle = 0$ and $n \in M^\perp$. Moreover, $\|h\|^2 = \|m + n\|^2 = \|m\|^2 + \|n\|^2$.

For uniqueness, let $h = m' + n'$. Then, $\underbrace{m - m'}_{\in M} = \underbrace{n' - n}_{\in M^\perp}$. Thus, $m - m' = 0$ and $m = m'$; $n = n'$ likewise.

For linearity, let $h = h_1 + \alpha h_2$. Then,

$$h = (m_1 + n_1) + \alpha(m_2 + n_2) = \underbrace{(m_1 + \alpha m_2)}_{\in M} + \underbrace{(n_1 + \alpha n_2)}_{\in M^\perp}. \qquad \blacksquare$$

Definition 4.2.5. P and Q are called the **orthogonal projections** of H onto M and $M^\perp$.

Corollary 4.2.6. *If $M \subsetneq H$ is a closed subspace, then $M^\perp \neq \{0\}$.*

Proof. Let $h \in H \setminus M$, then $Qh \notin \{0\}$ because $\|h - Ph\|^2 = \|Qh\|^2 \neq 0$. $\blacksquare$

Theorem 4.2.7 (**Riesz Representation Theorem on Hilbert Space**). *Let $\Lambda\colon H \to \mathbb{C}$ be a continuous (hence bounded) linear functional. Then there is a unique $y \in H$, such that $\Lambda(x) = \langle x, y\rangle$, for every $x \in H$.*

Proof. Suppose $\Lambda \not\equiv 0$. Let $M := \ker(\Lambda) = \{x \in H : \Lambda(x) = 0\}$. Then M is a proper closed subspace in H. Hence, $M^\perp \neq \{0\}$.

If $v, w \in M^\perp$ and $v, w \neq 0$, then $\Lambda(v), \Lambda(w) \neq 0$. Then,

$$\Lambda\left(\frac{v}{\Lambda(v)} - \frac{w}{\Lambda(w)}\right) = 1 - 1 = 0 \quad\Rightarrow\quad \frac{v}{\Lambda(v)} - \frac{w}{\Lambda(w)} \in M.$$

However, linearity of $M^\perp$ gives $\frac{1}{\Lambda(v)}v - \frac{1}{\Lambda(w)}w \in M^\perp$. Hence,

$$\frac{v}{\Lambda(v)} - \frac{w}{\Lambda(w)} = 0 \quad \Rightarrow \quad w = \frac{\Lambda(w)}{\Lambda(v)}v,$$

and $M^\perp = \mathbb{C}v$. Using orthogonal projections on M and $M^\perp$, given any $x \in H$, $x = Px + Qx$, where $Qx = \frac{\alpha v}{\Lambda(v)}$, for some $\alpha \in \mathbb{C}$. Consequently,

$$\begin{aligned}
\Lambda(x) &= \Lambda(Px) + \Lambda\left(\frac{\alpha v}{\Lambda(v)}\right) = 0 + \alpha \\
&= \alpha\left\langle v, \frac{v}{\overline{\Lambda(v)}} \cdot \frac{\overline{\Lambda(v)}}{\|v\|^2}\right\rangle = \alpha\left\langle \frac{v}{\Lambda(v)}, \frac{\overline{\Lambda(v)}v}{\|v\|^2}\right\rangle \\
&= \left\langle \frac{\alpha v}{\Lambda(v)}, \frac{\overline{\Lambda(v)}v}{\|v\|^2}\right\rangle = \left\langle \underbrace{Px + \frac{\alpha v}{\Lambda(v)}}_{x}, \underbrace{\frac{\overline{\Lambda(v)}v}{\|v\|^2}}_{y}\right\rangle.
\end{aligned}$$

For uniqueness, suppose $\langle x, y\rangle = \langle x, y'\rangle$ for each $x \in H$. Let $z := y - y'$, then $\langle x, z\rangle = 0$. In particular, $\langle z, z\rangle = 0$ gives $z = 0$ and $y = y'$. ∎

4.3 Orthonormal Sets

Definition 4.3.1. A family $\{u_\alpha\}_{\alpha\in A} \subset H$ is called **orthonormal** if $\langle u_\alpha, u_\beta\rangle = 0$, $\forall \alpha \neq \beta$, and $\|u_\alpha\| = 1$, $\forall \alpha \in A$. If $x \in H$, the complex numbers $\langle x, u_\alpha\rangle$ are called the **Fourier coefficients** of x relative to the set $\{u_\alpha\}$, or **coordinate orthogonal projections** onto $\mathrm{Span}(u_\alpha : \alpha \in A)$.

We begin with *finite* othonormal sets.

Proposition 4.3.2. *Let $\{u_\alpha\}_{\alpha\in A}$ be an orthonormal set, and $F \subseteq A$ be finite. Let $M_F := \mathit{Span}(u_\alpha : \alpha \in F)$.*

(a) If $\varphi\colon A \to \mathbb{C}$ with $\varphi|_{A\setminus F} = 0$, then there exists $y \in M_F$, namely

$$y = \sum_{\alpha\in F} \varphi(\alpha)u_\alpha,$$

such that $\langle y, u_\alpha\rangle = \varphi(\alpha)$, $\forall \alpha \in A$. Also,

$$\|y\|^2 = \sum_{\alpha\in F} |\langle y, u_\alpha\rangle|^2.$$

(b) If $x \in H$, then

$$\left\|x - \sum_{\alpha\in F}\langle x, u_\alpha\rangle u_\alpha\right\| < \|x - s\|, \tag{4.3.1}$$

for all $s \in M_F$, except $s = \sum_{\alpha\in F}\langle x, u_\alpha\rangle u_\alpha$. Moreover,

$$\sum_{\alpha\in F} |\langle x, u_\alpha\rangle|^2 \leq \|x\|^2. \tag{4.3.2}$$

Proof. (a) is a direct calculation using orthonormality.

For (b), let $s(x) := \sum_{\alpha\in F}\langle x, u_\alpha\rangle u_\alpha$. Note that $\langle s(x), u_\alpha\rangle = \langle x, u_\alpha\rangle, \forall\alpha \in F$. Hence, $\langle x - s(x), u_\alpha\rangle = 0, \forall\alpha \in F$. Because M_F is spanned by u_α's, we have $(x - s(x)) \perp (s(x) - s), \forall s \in M_F$. Therefore,

$$\|x - s\|^2 = \|(x - s(x)) + (s(x) - s)\|^2 = \|x - s(x)\|^2 + \|s(x) - s\|^2. \tag{4.3.3}$$

Hence, it gives (4.3.1). For (4.3.2), let $s = 0$ and the result from (a). ■

Remark. Note that (4.3.1) states that the "Fourier series" $\sum_{\alpha\in F}\langle x, u_\alpha\rangle u_\alpha$ of x is the unique best approximation to x in M_F.

Remark 4.3.3. Note that $s(x) = (\sum_{\alpha\in F}\langle x, u_\alpha\rangle u_\alpha) \perp x$, and we have the Pathagorean Theorem:

$$\|x\|^2 = \|x - s(x)\|^2 + \|s(x)\|^2. \tag{4.3.4}$$

Remark 4.3.4. Now we want to extend the results to uncountable sets. Because of that, we need to clarify the meaning of $\sum_{\alpha\in A}\varphi(\alpha)$. Suppose $0 \le \varphi(\alpha) \le \infty$, then we define

$$\sum_{\alpha\in A}\varphi(\alpha) := \sup\left\{\sum_{\alpha\in F}\varphi(\alpha) : F \subset A, F \text{ finite.}\right\}.$$

In fact, if μ denotes the counting measure on A, then

$$\sum_{\alpha\in A}\varphi(\alpha) = \int_A \varphi\,\mathrm{d}\mu.$$

In this case, we write $l^p(A)$ instead of $L^p(\mu)$. Moreover, if $\varphi\colon A \to \mathbb{C}$, then

$$\varphi \in l^2(A) \iff \int_A |\varphi|^2\,\mathrm{d}\mu = \sum_{\alpha\in A}|\varphi(\alpha)|^2 < \infty.$$

Example (4.1.10) shows that $l^2(A)$ is a Hilbert space with inner product:

$$\langle\varphi, \psi\rangle = \int_A \varphi\overline{\psi}\,\mathrm{d}\mu = \sum_{\alpha\in A}\varphi(\alpha)\overline{\psi(\alpha)}. \tag{4.3.5}$$

Remark 4.3.5. If $\varphi \in l^2(A)$, then $S := \{\alpha \in A : \varphi(\alpha) \ne 0\}$ is **at most countable**. To see this, let $A_n := \{\alpha : |\varphi(\alpha)| > 1/n\}$. Hence,

$$|A_n| < \sum_{\alpha\in A_n}|n\varphi(\alpha)|^2 \le n^2\sum_{\alpha\in A_n}|\varphi(\alpha)|^2 < \infty.$$

Since every A_n is finite, $S = \bigcup_{n=1}^\infty A_n$ is countable.

Theorem 4.3.6 (Bessel's Inequality). *Let $\{u_\alpha\}$ be an orthonormal set in H. Then $\forall h \in H$,*

$$\|h\|^2 \ge \sum_{\alpha\in A}|\langle h, u_\alpha\rangle|^2. \tag{4.3.6}$$

Proof. For every finite $F \subset A$, by Proposition 4.3.2, we have

$$\|h\|^2 \geq \sum_{\alpha \in F} |\langle h, u_\alpha \rangle|^2.$$

Let $\varphi(\alpha) := \langle h, u_\alpha \rangle$. By Remark (4.3.5), $S = \{\alpha \in A : \varphi(\alpha) \neq 0\}$ is at most countable. Therefore, taking the supremum of all finite F, hence all countable S, we obtain

$$\begin{aligned} \|h\|^2 &\geq \sup_{F \subset A, \text{ finite}} \left\{ \sum_{\alpha \in F} |\langle h, u_\alpha \rangle|^2 \right\} \\ &= \sup_{S \subset A} \left\{ \sum_{\alpha \in S} |\langle h, u_\alpha \rangle|^2 \right\} \\ &= \sum_{\alpha \in A} |\langle h, u_\alpha \rangle|^2 \end{aligned}$$

■

4.4 Orthonormal Basis

Definition 4.4.1. An orthonormal set $\{u_\alpha\}_{\alpha \in A}$ is called complete or an **orthonormal basis** if for all $h \in H$,

$$\|h\|^2 = \sum_{\alpha \in A} |\langle h, u_\alpha \rangle|^2. \tag{4.4.1}$$

Remark. Note that it is not a basis in the sense of vector space.

Definition 4.4.2. The set $U = \{u_\alpha\}$ is called a **maximal orthonormal set** if V is an orthonormal set containing U, then $V = U$.

Theorem 4.4.3. *Let $\{u_\alpha\}_{\alpha \in A}$ be an orthonormal set in H. The following are equivanlent:*

(1) $\{u_\alpha\}$ is an orthonormal basis.

(2) The set of all finite linear combinations of $\{u_\alpha\}$, denoted P, is dense in H.

(3) $\{u_\alpha\}$ is a maximal orthonormal set.

Proof. (1) $\Rightarrow$ (2). Given $h \in H$, by (1) we have

$$\|h\|^2 = \sum_{\alpha \in A} |\langle h, u_\alpha \rangle|^2.$$

Let $B \subset A$ be finite, and define $g_B := \sum_{\alpha \in B} \langle h, u_\alpha \rangle u_\alpha$. By Remark (4.3.3),

$$\|h - g_B\|^2 = \|h\|^2 - \|g_B\|^2 = \|h\|^2 - \sum_{\alpha \in B} |\langle h, u_\alpha \rangle|^2.$$

Now, taking the supremum over all such B, we have $\sum_{\alpha\in B}|\langle h,u_\alpha\rangle|^2 \to \|h\|^2$. Therefore, $\forall\varepsilon>0$, we can find such B so that $\|h-g_B\|<\varepsilon$. Therefore,

$$h = \sup_{B\subset A,\ \text{finite}} \left\{\sum_{\alpha\in B}\langle h,u_\alpha\rangle u_\alpha\right\},$$

and P is dense in H.

(2) $\Rightarrow$ (3). By contrapositive, suppose $\exists u\in H, \|u\|=1$, and $\langle u,u_\alpha\rangle = 0, \forall\alpha\in A$. Then, for all finite $B\subset A$, $\forall c_\alpha\in\mathbb{C}$,

$$\left\|u-\sum_{\alpha\in B}c_\alpha u_\alpha\right\|^2 = \|u\|^2+\sum_{\alpha\in B}|c_\alpha|^2 \geq 1.$$

Hence, $u\notin\overline{P}$, and P is not dense in H.

(3) $\Rightarrow$ (1). **Step 1:** By contrapositive, suppose there exists a $h\in H$, such that $\|h\|^2>\sum_{\alpha\in A}|\langle h,u_\alpha\rangle|^2$. Let $A_n := \{\alpha : |\langle h,u_\alpha\rangle|\geq 1/n\}$, and $A_0 := \{\alpha : \langle h,u_\alpha\rangle \neq 0\}$. Recall that $A_0=\bigcup_{n=1}^\infty A_n$, $A_1\subseteq A_2\subseteq\dots$, and every A_n is finite.

Define $g_n := \sum_{\alpha\in A_n}\langle h,u_\alpha\rangle u_\alpha$. Then, given $\varepsilon>0$, there exists $N\in\mathbb{N}$, such that

$$\sum_{\alpha\in A}|\langle h,u_\alpha\rangle|^2 < \sum_{\alpha\in A_N}|\langle h,u_\alpha\rangle|^2+\varepsilon.$$

Step 2: Now, for all $m>n>N$, by orthogonality,

$$\begin{aligned}
\|g_m-g_n\|^2 = \left\|\sum_{\alpha\in A_m\setminus A_n}\langle h,u_\alpha\rangle u_\alpha\right\|^2 &= \sum_{\alpha\in A_m\setminus A_n}|\langle h,u_\alpha\rangle|^2\\
&= \sum_{\alpha\in A_m}|\langle h,u_\alpha\rangle|^2 - \sum_{\alpha\in A_n}|\langle h,u_\alpha\rangle|^2\\
&\leq \sum_{\alpha\in A_0}|\langle h,u_\alpha\rangle|^2 - \sum_{\alpha\in A_N}|\langle h,u_\alpha\rangle|^2\\
&= \sum_{\alpha\in A}|\langle h,u_\alpha\rangle|^2 - \sum_{\alpha\in A_N}|\langle h,u_\alpha\rangle|^2 < \varepsilon
\end{aligned}$$

Therefore, $\{g_n\}$ is Cauchy and $g_n\to g\in H$.

Step 3: By monotonicity of $\{A_n\}$, for all $\gamma\in A_0$, there is N such that $\gamma\in A_n$, whenever $n>N$. Hence, $\langle h,u_\gamma\rangle = \langle g_n,u_\gamma\rangle$, for all $n>N$. On the other hand,

$$\langle h-g,u_\gamma\rangle = \langle h,u_\gamma\rangle - \langle g,u_\gamma\rangle = \langle h,u_\gamma\rangle - \lim_{k\to\infty}\langle g_k,u_\gamma\rangle$$

gives $\langle h-g,u_\gamma\rangle = 0$. Thus, $h-g\perp u_\gamma$, for every $\gamma\in A_0$.

Step 4: Moreover, if $\gamma\notin A_0$, then $\langle u_\gamma,u_\alpha\rangle=0$, for all $\alpha\in A_0$, and

$$\langle h-g,u_\gamma\rangle = \langle h,u_\gamma\rangle - \langle g,u_\gamma\rangle = 0-0=0.$$

Hence, $h - g \perp u_\alpha$, for every $\alpha \in A$. Now, by assumption and Remark (4.3.3),

$$\|h - g\|^2 = \lim_{n\to\infty} \|h - g_n\|^2 = \|h\|^2 - \lim_{n\to\infty} \|g_n\|^2 > 0.$$

Finally, let $u := \frac{h-g}{\|h-g\|}$. Then $\{u\} \cup \{u_\alpha\}$ is a bigger orthonormal set, thus $\{u_\alpha\}$ is not maximal. ∎

Corollary 4.4.4 (Parseval's Identity). *Let $\{u_\alpha\}_{\alpha\in A}$ be an orthonormal basis of H. Then, for every $h \in H$,*

$$\|h\|^2 = \sum_{\alpha\in A} |\langle h, u_\alpha\rangle|^2 \iff h = \sum_{\alpha\in A} \langle h, u_\alpha\rangle u_\alpha. \tag{4.4.2}$$

Moreover, for all $x, y \in H$,

$$\sum_{\alpha\in A} \langle x, u_\alpha\rangle \overline{\langle y, u_\alpha\rangle} = \langle x, y\rangle. \tag{4.4.3}$$

Proof. ($\Rightarrow$) is done in the proof of Theorem (4.4.3): (1) $\Rightarrow$ (2). Conversely, Let $\{c_\alpha\} \subset \mathbb{C}$ be square summable, and $\{u_\alpha\}$ be an orthonormal basis. Define A_n as in Theorem (4.4.3), then $\{g_n := \sum_{\alpha\in A_n} c_\alpha u_\alpha\}$ is Cauchy. Let $c_\alpha := \langle h, u_\alpha\rangle$, and $g_n \to h$ completes the proof.

Finally, recall that from Remark (4.3.4), $l^2(A)$ is a Hilbert space. For each $x \in H$, we associate a function $x(\alpha)$ on A by $x(\alpha) := \langle x, u_\alpha\rangle$. Then, from equation (4.4.2), $\|x\|^2_{l^2(A)} = \|x\|^2_H$. By Polarization Identity, inner products in $l^2(A)$ can be expressed in terms of norm in $l^2(A)$, which is equivalent to the norm in H. Therefore,

$$\sum_{\alpha\in A} \langle x, u_\alpha\rangle \overline{\langle y, u_\alpha\rangle} = \sum_{\alpha\in A} x(\alpha)\overline{y(\alpha)} = \langle x, y\rangle_{l^2(A)} = \langle x, y\rangle_H. \quad ∎$$

Corollary 4.4.5. *Let $\{u_\alpha\}_{\alpha\in A}$ be an orthonormal basis of H. For every $h \in H$, let $A_0 := \{\alpha : \langle h, u_\alpha\rangle \neq 0\}$. Then, there exists nested finite sequence A_n's, such that $A_0 = \bigcup_{n=1}^\infty A_n$, and*

$$\lim_{n\to\infty} \sum_{\alpha\in A_k}^{n} \langle h, u_\alpha\rangle u_\alpha = h \tag{4.4.4}$$

Proof. Done in Theorem (4.4.3). A useful result. ∎

Remark 4.4.6. If $\{u_\alpha\}_{\alpha\in A}$ is an orthonormal basis of H. Then the map $\Lambda\colon H \to l^2(A)$, given by

$$(\Lambda(x))(\alpha) := \langle x, u_\alpha\rangle,$$

is a bijection. Moreover, it preserves the distances, hence an isomorphic isometry.

Example 4.4.7. Let $H = l^2(A)$. The set $\{\chi_{\{\alpha\}} : \alpha \in A\}$ is an orthonormal basis.

EXAMPLE 4.4.8. Let $H = L^2([-\pi,\pi])$ with

$$\langle f, g\rangle := \frac{1}{2\pi}\int_{[-\pi,\pi]} f\overline{g}\,\mathrm{d}m. \tag{4.4.5}$$

Let $e_n(\theta) := e^{in\theta}$. Then $\{e_n : n \in \mathbb{Z}\}$ is an orthonormal basis, by showing the set of finite linear combinations is dense.

Idea: For any $f \in L^2([-\pi,\pi])$, take $g \in C([-\pi,\pi])$ such that $\|f-g\|_2 < \varepsilon$. Note that g might not be 2π-periodic. Then pick $h \in C([-\pi,\pi])$ with $h(-\pi) = h(\pi)$ so that $\|h-g\|_2 < \varepsilon$. Use Stone-Weierstrass to pick trigonometric polynomial $p(\theta) = \sum_{k=-m}^{m} c_k e_k(\theta)$ with $\|h-p\|_\infty < \varepsilon$. Then Hölder's inequality gives $\|h-p\|_2 < \sqrt{2\pi}\varepsilon$, and $\|f-p\|_2 < 2\varepsilon + \sqrt{2\pi}\varepsilon$.

COROLLARY 4.4.9 (Riesz-Fischer). *If $f \in L^2([-\pi,\pi])$, then*

$$f = \sum_{n\in\mathbb{Z}} \langle f, e_n\rangle e_n, \qquad \text{and} \qquad \|f\|^2 = \sum_{n\in\mathbb{Z}} |\langle f, e_n\rangle|^2.$$

Moreover, if $\{c_m\} \in l^2(\mathbb{Z})$, i.e. $\sum_{m=1}^{\infty} |c_m|^2 < \infty$, then

$$g := \sum_{m\in\mathbb{Z}} c_m e_m \in L^2([-\pi,\pi]),$$

and the partial sum converges to g in L^2.

Proof. Since $\{e_n\}$ is an orthonormal basis, all follow from the results above. ■

REMARK 4.4.10. The complex numbers $\frac{1}{\sqrt{2\pi}}\langle f, e_n\rangle$ are precisely the **Fourier coefficients** of f.

THEOREM 4.4.11. *Every nontrivial Hilbert space H has an orthonormal basis.*

Proof. We will use the Zorn's Lemma. Let $S := \{U \subseteq H : U \text{ is orthonormal}\}$. Then $(S, \subseteq)$ is partially ordered. Suppose C is a totally ordered subset of S.

Define $V := \bigcup_{U\in C} U$. If $x_1, x_2 \in V$, then there are $U_1, U_2 \in C$ such that $x_1 \in U_1$ and $x_2 \in U_2$. Without loss of generality, suppose $U_1 \subset U_2$. Hence, $x_1, x_2 \in U_2$, and

$$\langle x_1, x_2\rangle = \begin{cases} 1, & \text{if } x_1 = x_2, \\ 0, & \text{else.} \end{cases}$$

Thus, V is orthonormal and V is an upper bounded of C. By Zorn's Lemma, S has a maximal element W, which is an maximal orthonormal set. Therefore, W is an orthonormal basis for H. ■

4.5 Isometries

Definition 4.5.1. Let $(H, \langle\cdot,\cdot\rangle_H), (K, \langle\cdot,\cdot\rangle_K)$ be Hilbert spaces. A linear map $\Lambda\colon H \to K$ is called an **isometry** if for all $h, g \in H$,

$$\langle \Lambda(h), \Lambda(g)\rangle_K = \langle h, g\rangle_H.$$

In addition, if Λ is surjective, we say Λ is an **unitary**. If such Λ exists, then H and K are isomorphic. Then, Λ is a **Hilbert space isometric isomorphism**.

Remark. Injectivity is automatic since $\Lambda(h) = 0 \iff \|\Lambda(h)\|^2 = 0 \iff h = 0$.

Theorem. *The Parseval's Identity together with Remark (4.4.6) give the following:*

If $\{u_\alpha\}_{\alpha\in A}$ is an orthonormal basis of a Hilbert space H, then the map $\Lambda\colon H \to l^2(A)$, by

$$(\Lambda(x))_{\alpha\in A} := (\langle x, u_\alpha\rangle)_{\alpha\in A},$$

is a Hilbert space isometric isomorphism.

Proposition 4.5.2. *A linear map $\Lambda\colon H \to K$ is an isometry if and only if $\|\Lambda(h)\| = \|h\|$, for all $h \in H$.*

Proof. $(\Rightarrow)$ is obvious with $g = h$. Conversely, first suppose H and K are over $\mathbb{R}$. By the Parallelogram law,

$$\begin{aligned}
\langle \Lambda(h), \Lambda(g)\rangle_K &= \frac{1}{4}(\|\Lambda(h) + \Lambda(g)\|_K^2 - \|\Lambda(h) - \Lambda(g)\|_K^2 \\
&= \frac{1}{4}(\langle \Lambda(h+g), \Lambda(h+g)\rangle_K - \langle \Lambda(h-g), \Lambda(h-g)\rangle_K) \\
&= \frac{1}{4}(\langle h+g, h+g\rangle_H - \langle h-g, h-g\rangle_H \\
&= \frac{1}{4}(\|h+g\|_H^2 - \|h-g\|_H^2 \\
&= \langle h, g\rangle_H
\end{aligned}$$

For complex case, split it into real and imaginary parts. ■

Example 4.5.3. Let $S\colon l^2(\mathbb{N}) \to l^2(\mathbb{N})$ given by $(x_1, x_2, x_3, \dots) \mapsto (0, x_1, x_2, \dots)$. Such S an **isometry but not unitary**; it is known as the *shift operation* or *unitary shift*.

Theorem 4.5.4. *Let $\{u_\alpha\}_{\alpha\in A}$ be an orthonormal basis of H. Then there is a unitary map $U\colon H \to l^2(A)$ such that $U(u_\alpha) = \chi_{\{\alpha\}}$.*

Proof. **Step 1:** Define $Q := \{\sum_{i=1}^n c_i u_{\alpha_i} : \alpha_i \in A, c_i \in \mathbb{C}, n \in \mathbb{N}\}$. Hence, Q is the set of all finite linear combinations, and $\overline{Q} = H$. Consider $p \in Q$ as linear combination of all u_α's with $c_\alpha = 0$ when necesary. We define $W\colon Q \to l^2(A)$, by

$$W(p) := (c_\alpha)_{\alpha\in A}. \tag{4.5.1}$$

Then, by Theorem (4.4.3),

$$\|W(p)\|_{l^2(A)} = \|p\|_H. \tag{4.5.2}$$

Step 2: Now, for every h, there is a sequence $p_n \to h$. Hence, $\{p_n\}$ is Cauchy in H and $\|W(p_n - p_m)\|_{l^2(A)} = \|p_n - p_m\|_H$ implies that $\{W(p_n)\}$ is Cauchy in $l^2(A)$. Recall that $l^2(A)$ is complete. Thus, there is $g \in l^2(A)$ so that $W(p_n) \to g$. Define $U(h) := g$. However, we need to check that U is well defined, i.e. $U(h)$ does not depends on the choice of $\{p_n\}$.

Step 3: Suppose $p'_n \to h$ and $W(p'_n) \to g$. We want show that $g' = g$. Given $\varepsilon > 0$, there is $N \in \mathbb{N}$, such that for each $n > N$,

$$\|p_n - h\|_H < \varepsilon \qquad \text{and} \qquad \|p'_n - h\|_H < \varepsilon.$$

Hence, by $\triangle$-inequality,

$$\|p_n - p'_n\|_H < \|p_n - h\|_H + \|p'_n - h\|_H < 2\varepsilon.$$

By isometry of W in equation (4.5.2),

$$\|W(p_n - p'_n)\|_{l^2(A)} = \|W(p_n) - W(p'_n)\|_{l^2(A)} < 2\varepsilon.$$

Therefore,

$$\begin{aligned}
\|g - g'\|_{l^2(A)} &= \|g - W(p_n) + W(p_n) - W(p'_n) + W(p'_n) - g'\|_{l^2(A)} \\
&\leq \|g - W(p_n)\|_{l^2(A)} + \|W(p_n) - W(p'_n)\|_{l^2(A)} + \|W(p'_n) - g'\|_{l^2(A)} \\
&= 4\varepsilon.
\end{aligned}$$

By convergence, $g = g'$, and U is well defined.

Step 4: Now we will show U is surjective and isometric. For isometry, note that

$$\|U(h)\| = \lim_{n\to\infty} \|W(p_n)\| = \lim_{n\to\infty} \|p_n\| = \|h\|.$$

For surjectivity, the set of finite linear combinations of characteristic functions of $\{\alpha\}$, $P := \{\sum_{i=1}^{n} c_i \chi_{\{\alpha_i\}}\}$ is dense in $l^2(A)$. Also, $U(H)$ is closed by the isometry above. Hence,

$$U(H) = \overline{U(H)} = \overline{U(Q)} = \overline{P} = l^2(A).$$

Therefore, U an unitary from H to $l^2(A)$. ■

Chapter 5

Examples of Banach Space Techniques

5.1 Banach Spaces

Definition 5.1.1. A complex vector space $(X, \|\cdot\|)$, with a **norm** $\|\cdot\|\colon X \to [0, \infty)$ is a **normed vector space** if it satisfies the followings: For all $x, y \in X$,

i. $\|x + y\| \le \|x\| + \|y\|$.

ii. $\|\alpha x\| = |\alpha|\|x\|$, for all $\alpha \in \mathbb{C}$.

iii. If $\|x\| = 0$, then $x = 0$.

Definition 5.1.2. If $\{x_n\}$ is a sequence in $(X, \|\cdot\|)$, the series $\sum_{n=1}^{\infty} x_n$ is said to converge to x if for some $x \in X$, $\sum_{n=1}^{N} x_n \to x$, as $N \to \infty$. The series is called **absolutely convergent** if $\sum_{n=1}^{\infty} \|x_n\| < \infty$.

Proposition 5.1.3. *A normed vector space X is complete if and only if every absolutely convergent series in X converges.*

Proof. ($\Rightarrow$). Suppose X is complete and $\sum_{n=1}^{\infty} \|x_n\| < \infty$. Define $s_k := \sum_{n=1}^{k} x_n$. Given $\varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $\forall k > m > N$, $\sum_{n=m}^{\infty} \|x_n\| < \varepsilon$. Therefore,

$$\|s_k - s_m\| = \sum_{n=m+1}^{k} \|x_n\| < \varepsilon.$$

Thus, $\{s_n\}$ is Cauchy and $s_n \to x = \sum_{n=1}^{\infty} x_n \in X$.

($\Leftarrow$). Let $\{x_n\}$ be Cauchy. For each $k \in \mathbb{N}$, $\exists N_k$ such that $\forall m, n > N_k > N_{k-1}$, $\|x_m - x_n\| < 2^{-k}$. Define $s_1 := x_{N_1}$ and $s_k := x_{N_k} - x_{N_{k-1}}, \forall k > 1$. Note that $\|s_k\| \le 2^{-k}$ and $s_1 + \cdots + s_k = x_{N_k}$. Hence,

$$\sum_{k=1}^{\infty} \|s_k\| \le \|s_1\| + \sum_{k=2}^{\infty} 2^{-k} < \infty.$$

By absolute convergence,

$$\lim_{k\to\infty} s_k = \lim_{k\to\infty} x_{N_k} = x \in X.$$

Finally, every Cauchy sequence with a convergent subsequence converges to the same limit. Therefore, $x_n \to x \in X$. ■

Definition 5.1.4. A **Banach space** is a normed vector space that is **complete** in the metric topology induced by the norm.

Example 5.1.5. Every Hilbert space is a Banach space, so is $L^p(\mu)$, where $1 \leq p \leq \infty$.

Definition 5.1.6. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed vector spaces. Let $\Lambda\colon X \to Y$ be a linear map. We define the **operator norm** of Λ by

$$\|\Lambda\| := \sup\{\|\Lambda(x)\|_Y : \|x\|_X \leq 1\}. \tag{5.1.1}$$

If $\|\Lambda\| < \infty$, then we say Λ is **bounded**.

Remark 5.1.7. Note that in (5.1.1), we may modify the definition to

$$\|\Lambda\| = \sup\{\|\Lambda(x)\|_Y : \|x\|_X = 1\}, \tag{5.1.2}$$

since if $x \in X$, $\|\Lambda(\alpha x)\|_Y = \|\alpha\Lambda(x)\|_Y = |\alpha|\|\Lambda(x)\|_Y$.

Remark 5.1.8. Note that $\|\Lambda\|$ is the smallest number such that $\forall x \in X$,

$$\|\Lambda(x)\|_Y \leq \|\Lambda\|\|x\|_X. \tag{5.1.3}$$

Remark. From formula (5.1.1), Λ maps the closed unit ball in X into a closed ball in Y with center 0 and radius $\|\Lambda\|$.

Theorem 5.1.9. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed vector spaces, and $\Lambda\colon X \to Y$ be linear. The following are equivalent:*

(1) Λ is bounded.

(2) Λ is uniformly continuous.

(3) Λ is continuous at some $x_0 \in X$.

Proof. (1) $\Rightarrow$ (2). For all $x, y \in X$,

$$\|\Lambda(x) - \Lambda(y)\|_Y = \|\Lambda(x-y)\|_Y \leq \|\Lambda\|\|x-y\|_X.$$

Hence, given any $\varepsilon > 0$, simply pick $\delta < \frac{\varepsilon}{\|\Lambda\|}$.

(2) $\Rightarrow$ (3). Obvious.

(3) $\Rightarrow$ (1). Given $\varepsilon > 0$, there is $\delta > 0$ such that $\forall x \in X$, with $\|x - x_0\|_X < \delta$, $\|\Lambda(x - x_0)\|_Y < \varepsilon$. Then, if $y \in X$ with $\|y\|_X < \delta$, we have

$$\|\Lambda(x_0 + y) - \Lambda(x_0)\|_X = \|\Lambda(y)\|_Y < \varepsilon.$$

By linearity, $\forall z \in X$ with $\|z\|_X < 1$, $\|\Lambda(z)\|_Y < \varepsilon/\delta$. By continuities of Λ and the norm function, $\|\Lambda\| = \sup\{\|\Lambda(z)\|_y : \|z\|_X \leq 1\} \leq \varepsilon/\delta < \infty$. ■

5.2 Consequences of Baire's Theorem

The completeness of Banach spaces is useful in application. In fact, two of the three fundamental theorems in Functional Analysis use completeness and Baire's Theorem.

Theorem 5.2.1 (**Baire's Category Theorem**). *Let (X, d) be a* ***complete metric space****, and U_n be an open dense subset of X, $\forall n \in \mathbb{N}$. Then, $\bigcap_{n=1}^{\infty} U_n$ is dense in X.*

Proof. We want to show that for all open W, $W \cap \bigcap_{n=1}^{\infty} U_n \neq \varnothing$. Denote $B(x,r)$ the open ball with radius r centered at x, and $\overline{B}(x,r)$ be its closure. (Note: $\overline{B}(x,r) \neq \{y : d(x,y) \leq r\}$ in general.)

Now, since U_1 is dense, $U_1 \cap W$ is open and nonempty. Then there is $x_1 \in U_1 \cap W$ and $r_1 \in (0,1)$ such that $\overline{B}(x_1, r_1) \subset V_1 \cap W$. Next, since V_2 is dense, $V_2 \cap B(x_1, r_1) \subset W$. So, there exists $\overline{B}(x_2, r_2) \subset V_2 \cap B(x_1, r_1)$ with $r_2 \in (0, 1/2)$. Proceed inductively, we obtain

$$\cdots \subset \overline{B}(x_n, r_n) \subset U_n \cap B(x_{n-1}, r_{n-1}) \subset \overline{B}(x_{n-1}, r_{n-1}) \subset \cdots \subset W,$$

with $0 < r_n < 1/n$.

Also, for all $i, j > n$, with $x_i, x_j \in B(x_n, r_n)$, we have $d(x_i, x_j) < \frac{2}{n}$. Hence, $\{x_n\}$ is Cauchy and by completeness, $x_n \to x \in X$. Moreover, $x \in \overline{B}(x_n, r_n) \subset U_n \cap W$, for every $n \in \mathbb{N}$. Therefore, $W \cap \bigcap_{n=1}^{\infty} U_n \neq \varnothing$, and $\bigcap_{n=1}^{\infty} U_n$ is dense. ■

Corollary 5.2.2. *If $\{U_n\}_{n=1}^{\infty}$ is a sequence of G_δ sets, and U_n is dense for all $n \in \mathbb{N}$, then $\bigcap_{n=1}^{\infty} U_n$ is dense and G_δ.*

Proof. For each $n \in \mathbb{N}$, $U_n = \bigcap_{k_n=1}^{\infty} V_{n,k_n}$, where V_{n,k_n} is open, and dense. Therefore,

$$\bigcap_{n=1}^{\infty} = \bigcap_{n=1}^{\infty} \bigcap_{k_n=1}^{\infty} V_{n,k_n}$$

is dense and a countable intersection of open sets, hence G_δ. ■

Definition 5.2.3. A set $E \subset X$ is called **nowhere dense** if $\overline{E}$ does not contain any open set in X. A countable union of such E is called a set of the **first category**. Otherwise, it is of the **second category**.

Corollary 5.2.4. *Let (X, d) be a complete metric space. Then X is of the second category.*

Proof. Let $\{E_n\}$ be a sequenece of nowhere dense sets in X. Then $\{\overline{E_n}^c\}$ is a sequence of open dense sets. By Baire's Theorem, $\bigcap_{n=1}^{\infty} \overline{E_n}^c \neq \varnothing$. Hence, $\bigcup_{n=1}^{\infty} E_n \subset \bigcup_{n=1}^{\infty} \overline{E_n} \neq X$. ■

Corollary 5.2.5. *Contraposition of Baire's Theorem: Let (X, d) be a complete metric space, and $\{U_n\}$ be a sequence of open sets. If $\bigcap_{n=1}^{\infty} U_n$ is not dense in X, then there exists $n \in \mathbb{N}$ such that U_n is not dense in X. Hence, there is a nonempty open set $V \subseteq X \setminus \overline{U}$.*

Corollary 5.2.6. *In a complete metric space (X, d) which has no isolated points, no dense G_δ-subset is countable.*

Proof. By way of contradiction, suppose $E = \{x_1, x_2, \dots\} \subset X$ is a countable dense G_δ. Then, $E = \bigcap_{k=1}^{\infty} V_k$, where V_k is dense and open. Then,

$$U_k := V_k \setminus \bigcup_{n=1}^{k} \{x_n\}$$

is a dense open set. However, $\bigcap_{k=1}^{\infty} U_k = \varnothing$, contradiction. ■

Theorem 5.2.7 (Banach-Steinhaus Theorem). *Let $(X, \|\cdot\|_X)$ be a Banach space and $(Y, \|\cdot\|_Y)$ be a normed vector space. Let $\{\Lambda_\alpha : \alpha \in A\}$ be a collection of bounded linear maps from X to Y. Then, either*

(a) Bounded uniformly: There is $M > 0$, such that for all $\alpha \in A, \|\Lambda_\alpha\| \leq M$, or

(b) All blows up: There is a dense G_δ-set $S \subset X$, such that for all $x \in S$, $\sup\limits_{\alpha \in A} \|\Lambda_\alpha(x)\|_Y = \infty$.

Remark. Geometrically, either there is a ball $B(0, M)$ in Y such that every Λ_α maps the unit ball in X into B; or there is a dense G_δ-set S such that for all $x \in S$ no ball in Y contains $\Lambda_\alpha(x)$, for all α.

Proof. For each $\alpha \in A$, let $\varphi_\alpha(x) := \|\Lambda_\alpha(x)\|_Y$. Since Λ_α and norm function are continuous, φ_α is continuous, hence lower semicontinuous. Define

$$\varphi(x) := \sup_{\alpha \in A} \varphi_\alpha(x), \qquad \text{and} \qquad V_n := \varphi^{-1}((n, \infty)), \quad \forall n \in \mathbb{N}.$$

Note that φ is lower semicontinous, so V_n are open sets.

Suppose each V_n is dense in X. Then by Baire's Theorem (5.2.1), the set $S := \bigcap_{n=1}^{\infty} V_n$ is G_δ and dense in X with $\varphi(S) = \{\infty\}$. Hence, it proves (b).

Otherwise, $\exists n \in \mathbb{N}$ with V_n is not dense. Thus, $\exists B(y, \delta) \subset V_n^c$, for some $y \in X$, $\delta > 0$. Therefore, for all $x \in B(y, \delta)$, $\|\Lambda_\alpha(x)\|_Y \leq n$, for all $\alpha \in A$. By linearity, if $z \in X$ with $\|z\|_X < \delta$,

$$\|\Lambda_\alpha(z)\|_Y = \|\Lambda_\alpha(z + y) - \Lambda_\alpha(y)\|_Y \leq 2n.$$

Hence, for every $\alpha \in A$ and $z \in X$ with $\|z\|_X \leq 1$, $\|\Lambda_\alpha(z)\|_Y < 2n/\delta$. Consequently,

$$\sup_{\alpha \in A} \|\Lambda_\alpha\| < \frac{2n}{\delta} = M < \infty,$$

and it proves (a). ■

COROLLARY 5.2.8. *Let X, Y, and $\{\Lambda_\alpha\}_{\alpha\in A}$ be as in Theorem* (5.2.7). *If for every $x \in X$, $\sup\limits_{\alpha\in A} \|\Lambda_\alpha(x)\|_Y < \infty$, then $\|\Lambda_\alpha\| \le M$, for some M.*

THEOREM 5.2.9 (Open Mapping Theorem). *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be Banach spaces. If $\Lambda : X \to Y$ is a surjective bounded linear map. then Λ is an open map.*

Proof. Let $B^X(x, r)$ denote the open ball in X with radius r about $x \in X$. We shall divide the proof into two parts. Let $(*)$ be the following statement:

$$\text{There exists } \delta > 0, \text{ such that } \Lambda(B^X(0,1)) \supset B^Y(0,\delta).$$

If $(*)$ is true, then we will be done by the argument below.

Let $U \subseteq X$ be open. We want to find an open neighborhood for each $y \in \Lambda(U)$ lying inside $\Lambda(U)$. For each $y \in \Lambda(U)$, $\exists u \in U$ so that $\Lambda(u) = y$, and $\exists \varepsilon > 0$ with $B^X(u,\varepsilon) \subseteq U$. Consider $B^Y(y, \varepsilon\delta)$. For each $w \in B^Y(y, \varepsilon\delta)$,

$$\frac{1}{\varepsilon}\|w - \Lambda(u)\|_Y < \delta \quad \Rightarrow \frac{1}{\varepsilon}(w - \Lambda(u)) \in B^Y(0,\delta).$$

Thus, by $(*)$, $\exists x \in B^X(0,1)$, so that $\Lambda(x) = \frac{1}{\varepsilon}(w - \Lambda(u))$, and $w = \Lambda(u + \varepsilon x)$. Define $v := u + \varepsilon x$. Then, $\Lambda(v) = w$ and

$$\|u - v\|_X = \varepsilon\|x\| < \varepsilon \quad \Rightarrow \quad v \in B^X(u,\varepsilon) \subseteq U.$$

Therefore, $B^Y(y, \varepsilon\delta) \subset \Lambda(U)$, is an open neighborhood of y. Hence, $\Lambda(U)$ is open and this completes the proof.

Proof of $(*)$

Step 1: By Corollary 5.2.4., Y is not nowhere dense. Since Λ is surjective, $Y = \bigcup_{n=1}^\infty \Lambda(B^X(0,n))$. By linearity, for each n, the map $y \mapsto ny$ is a homeomorphism from $\Lambda(B^X(0,1))$ to $\Lambda(B^X(0,n))$. Thus, $\Lambda(B^X(0,1))$ cannot be nowhere dense.

Step 2: Hence, there exists $B^Y(y_0, 4r) \subset \overline{\Lambda(B^X(0,1))}$, for some $y_0 \in Y$, $r > 0$. Pick $y' \in B^Y(y_0, 2r) \cap \overline{\Lambda(B^X(0,1))}$, with $y' = \Lambda(x')$, for some $x' \in B^X(0,1)$. Then, $B^Y(y', 2r) \subset B^Y(y_0, 4r) \subset \overline{\Lambda(B^X(0,1))}$. For all $y \in Y$, with $\|y\|_Y < 2r$,

$$\begin{aligned} y = -y' + (y + y') &\in -y' + \overline{\Lambda(B^X(0,1))} \\ &\subseteq \overline{\Lambda(-x' + B^X(0,1))} \subset \overline{\Lambda(B^X(0,2))}. \end{aligned}$$

Dividing both sides by 2, we have

$$B^Y(0,r) \subset \overline{\Lambda(B^X(0,1))}. \tag{5.2.1}$$

If we can pick sufficently small $\delta < r$, such that $B^Y(0,\delta) \subset \Lambda(B^X(0,1))$, then $(*)$ is proved.

Step 3: Claim: let $\delta = \frac{r}{2}$. By linearity, for each $n \in \mathbb{N}$, condition (5.2.1) gives

$$B^Y(0, r2^{-n}) \subseteq \overline{\Lambda(B^X(0,2^{-n}))}.$$

By definition of closure, given $y \in B^Y(0,\delta)$, there is $x_1 \in B^X(0,2^{-1})$, such that $\|y-\Lambda(x_1)\|_Y < r2^{-2}$. Let $y_1 := y-\Lambda(x_1)$, then $y_1 \in B^Y(0,r2^{-2}) \subseteq \overline{\Lambda(B^X(0,2^{-n}))}$.

Then, we can pick $x_2 \in B^X(0,2^{-2})$, such that

$$y_2 := y_1 - \Lambda(x_2) = y - \Lambda(x_1 + x_2) \quad \in B^Y(0, r2^{-3}).$$

Proceed inductively, for each $n \in \mathbb{N}$, we choose $x_n \in B^X(0,2^{-n})$, such that $y_n := y - \Lambda(\sum_{i=1}^n x_i) \in B^Y(0, r2^{-n-1})$, i.e.

$$\|y_n\|_Y = \left\| y - \Lambda\left(\sum_{i=1}^n x_i\right)\right\|_Y < \frac{r}{2^{n+1}}$$

Step 4: Note that $\sum_{n=1}^\infty \|x_n\|_X < \sum_{n=1}^\infty 2^{-n} = 1$. Hence, $\sum_{n=1}^\infty x_n = x$ for some $x \in X$, by absolute convergence. On the other hand, $y_n \to 0$ since $\|y_n\| \to 0$. Therefore, $\Lambda(x) = y$, and we conclude $B^Y(0,\delta) \subset \Lambda(B^X(0,1))$. ∎

Corollary 5.2.10 (Bounded Inverse Theorem). *Let X, Y be Banach spaces and $\Lambda\colon X \to Y$ be bijective. Then Λ^{-1} is bounded.*

Proof. Because Λ is open, Λ^{-1} is continuous, hence bounded. ∎

Theorem 5.2.11 (Closed Graph Theorem). *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be Banach spaces. Let $\Lambda : X \to Y$ be linear. Then Λ is bounded if and only if $\mathcal{G}(\Lambda) := \{(x,\Lambda(x)) : x \in X\}$ is closed in $X \times Y$.*

Proof. Define $\|(x,y)\| := \|x\|_X + \|y\|_Y$ on $X \times Y$. It is easy to check that $(X, \times Y, \|\cdot\|)$ is a Banach space. ($\Rightarrow$). Since Λ is continuous, the map $x \mapsto (x, \Lambda(x))$ is also continuous. Hence, by sequential continuity, $\mathcal{G}(\Lambda)$ is closed in $X \times Y$.

($\Leftarrow$). Let $\mathcal{G}(\Lambda)$ be closed and $\pi\colon \mathcal{G}(\Lambda) \to X$ be $\pi(x,\Lambda(x)) := x$. Note that π is linear, bijective, and bounded by $\|x\|_X \le \|(x,y)\|$. Also, $\mathcal{G}(\Lambda)$ is a Banach space with the induced norm because of closedness. The Bounded Inverse Theorem gives $\|\pi^{-1}(x)\| = \|x\|_X + \|\Lambda(x)\|_Y \le C\|x\|_X, \forall x \in X$ for some $C > 0$. Therefore, Λ is bounded. ∎

5.3 Fourier Series of Continuous Functions

Question (Pointwise Convergence of Fourier Series). Let $T := [-\pi,\pi]$, $C(T) := \{f\colon T \to \mathbb{C} \mid f \text{ is } 2\pi\text{-periodic and continuous}\}$. Since $C(T) \subset L^2(T)$, by Riesz-Fischer Theorem (4.4.9), if $f \in C(T)$, f has the Fourier series $\mathcal{F}(f)$ with coefficients

$$c_n := \langle f, e_n\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,\mathrm{d}t, \quad n \in \mathbb{Z},$$

and

$$s_n(x) := \sum_{k=-n}^{n} c_k e^{nkx}, \quad \lim_{n\to\infty} s_n = \mathcal{F}(f) = f, \ \mu\text{-a.e.}$$

Now here is a natural question to ask: Is it true that for all $f \in C(T)$, $\mathcal{F}(f)(x) = f(x)$, for every $x \in X$?

The Banach-Stienhaus Theorem answers it negatively as follows.

Solution. **Step 1:** Define the *Dirichlet kernel*,

$$D_n(t) := \sum_{k=-n}^{n} e^{ikt} = \frac{\sin(n + \frac{1}{2}t)}{\sin(\frac{t}{2})}. \tag{5.3.1}$$

Observe that

$$\begin{aligned} \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)D_n(x-t)\,\mathrm{d}t &= \sum_{k=-n}^{n}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-ikt}\,\mathrm{d}t\right)e^{ikx} \\ &= \sum_{k=-n}^{n} c_k e^{ikx} = s_n(x). \end{aligned}$$

For each $n \in \mathbb{N}$, let $\Lambda_n\colon C(T) \to \mathbb{C}$ by $\Lambda_n(f) := s_n(0)$. Then, $\{\Lambda_n\}_{n=1}^{\infty}$ is a sequence of linear functionals.

Equip $C(T)$ with the sup-norm $\|\cdot\|_\infty$, then $\forall f \in C(T)$, with $\|f\|_\infty \le 1$, by Hölder's inequality,

$$\begin{aligned} |\Lambda_n(f)| &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)D_n(t)\,\mathrm{d}t\right| \\ &\le \frac{1}{2\pi}\|f\|_\infty\|D_n\|_1 \\ &\le \frac{1}{2\pi}\|D_n\|_1 < \infty. \end{aligned}$$

Hence, for each n, Λ_n is a bounded linear functional.

Step 2: We will show that $\{\Lambda_n\}$ is not bounded uniformly. First, consider $\lim_{n\to\infty}\|D_n\|_1$. With $|\sin(x)| \le |x|$, we have

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |D_n(t)|\,\mathrm{d}t \ge \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{|\sin((n+\frac{1}{2})t)|}{|\frac{t}{2}|}\,\mathrm{d}t = \frac{2}{\pi}\int_{0}^{\pi}\frac{|\sin((n+\frac{1}{2})t)|}{t}\,\mathrm{d}t.$$

Integrating the right-hand side of the expression with $u = (n+1/2)t$, we have

$$\begin{aligned} \frac{2}{\pi}\int_{0}^{\pi}\frac{|\sin((n+\frac{1}{2})t)|}{t}\,\mathrm{d}t &\ge \frac{2}{\pi}\int_{0}^{(n+1/2)\pi}\frac{|\sin(u)|}{u}\,\mathrm{d}u \\ &\ge \frac{2}{\pi}\sum_{k=1}^{n}\int_{(k-1)\pi}^{k\pi}\frac{|\sin(u)|}{u}\,\mathrm{d}u \end{aligned}$$

$$
\begin{aligned}
&\geq \frac{2}{\pi}\sum_{k=1}^{n}\int_{(k-1)\pi}^{k\pi}\frac{|\sin(u)|}{k\pi}\,\mathrm{d}u\\
&= \frac{2}{\pi}\sum_{k=1}^{n}\frac{1}{k\pi}\int_{0}^{\pi}|\sin(u)|\,\mathrm{d}u,
\end{aligned}
$$

which $\to \infty$, as $n \to \infty$. Hence, $\|D_n\|_1 \to \infty$.

Next, we claim that $\|\Lambda_n\| = \|D_n\|_1$, for each $n \in \mathbb{N}$. Define

$$
g(x) = \begin{cases} 1, & D_n(t) > 0, \\ -1, & D_n(t) < 0, \\ 0, & \text{else.} \end{cases}
$$

Note that $g \notin C(T)$. Pick $\{f_k\} \subset C(T)$ such that $-1 \leq f_k \leq 1$, and $f_k(t) \to g(t), \forall t \in T$. Then by Dominated Convergence,

$$
\Lambda_n(f_k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_k(t)D_n(t)\,\mathrm{d}t \xrightarrow{k\to\infty} \frac{1}{2\pi}\int_{-\pi}^{\pi} g(t)D_n(t)\,\mathrm{d}t = \frac{1}{2\pi}\|D_n\|_1.
$$

Thus, $\|\Lambda_n\| = \|D_n\|_1 \to \infty$, and $\{\Lambda_n\}$ is not uniformly bounded.

Step 3: By the Banach-Steinhaus Theorem (5.2.7), there is a dense G_δ-subset G in $(C(T), \|\cdot\|_\infty)$, such that

$$
\mathcal{F}(f)(0) = \lim_{n\to\infty} |\Lambda_n(f)| = \infty, \quad \forall f \in G.
$$

Therefore, the Fourier series of $f \in G$ does not converge at $x = 0$. ■

Remark. In fact, if $x \in T$, we can find such a corresponding dense G_δ-subset G_x in $C(T)$ so that for all $f \in G_x$, $\mathcal{F}(f)(x) = \infty$. Let us take countably many such $x_n \in T$. Then by Baire's Theorem, $G := \bigcap_{n=1}^{\infty} G_n$ is again a dense G_δ in $C(T)$ so that

$$
\mathcal{F}(f)(x_i) = \infty, \quad \forall f \in G, \forall i \in \mathbb{N}.
$$

Also, if we choose the $\{x_n\}$ such that it is dense (e.g. the rationals) in T, then we can conclude that for all fixed $f \in G$,

$$
E_f := \{x : \mathcal{F}(f)(x) = \infty\}
$$

is a dense G_δ in $\mathbb{R}$ by periodicity. Moreover, by Corollary 5.2.6, each E_f and G are uncountable.

5.4 Fourier Coefficients of L^1-functions

Lemma 5.4.1 (Riemann-Lebesgue). *Let $T := [-\pi, \pi]$, and $f \in L^1(T)$, and define*

$$
\hat{f}(n) := \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,\mathrm{d}t. \tag{5.4.1}
$$

Then, $\lim_{n\to\pm\infty} \hat{f}(n) = 0$.

Proof. Recall that the set of all trigonometric polynomials is dense in $L^1(T)$. For $\varepsilon > 0$, there exists trigonometric polynomial $p = \sum_{k=-m}^{m} p_k e^{ikt}$ such that $\|f-p\|_1 < 2\pi\varepsilon$. Hence, for every n, by Hölder's inequality,

$$\begin{aligned}
|\hat{f}(n) - \hat{p}(n)| &= \frac{1}{2\pi}\left| \int_{-\pi}^{\pi} (f-p)e^{-int}\,\mathrm{d}t \right| \\
&\leq \frac{1}{2\pi}\|f-p\|_1 \|e^{-int}\|_\infty \\
&= \frac{1}{2\pi}\|f-p\|_1 < \varepsilon.
\end{aligned}$$

Since $\hat{p}(n) \to 0$, given any $\varepsilon > 0$,

$$\lim_{n\to\pm\infty} |\hat{f}(n)| < \varepsilon \quad \Rightarrow \quad \lim_{n\to\pm\infty} |\hat{f}(n)| = 0. \qquad \blacksquare$$

Question. Is the converse of the Riemann-Lebesgue Lemma true?

Solution. Let $C_0 := \{f \colon \mathbb{Z} \to \mathbb{C} \mid \lim_{n\to\pm\infty} f(n) = 0\}$. Then, $(C_0, \|\cdot\|_\infty)$ is a Banach space. Define $\Lambda \colon L^1(T) \to C_0$ by $(\Lambda f)(n) := \hat{f}(n)$.

Step 1: We first show that Λ is bounded. Note that

$$\begin{aligned}
\|\Lambda\| &= \sup\left\{ \|\Lambda(f)(n)\|_\infty : \|f\|_1 \leq 1 \right\} \\
&= \sup\left\{ \sup_{n\in\mathbb{Z}} |\Lambda(f)(n)| : \|f\|_1 \leq 1 \right\} \\
&= \sup_{\|f\|_1 \leq 1} \left(\sup_{n\in\mathbb{Z}} \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)e^{-int}\,\mathrm{d}t \right\} \right) \\
&\leq \sup_{\|f\|_1 \leq 1} \left(\sup_{n\in\mathbb{Z}} \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t)|\,\mathrm{d}t \right\} \right) \\
&= \frac{1}{2\pi}.
\end{aligned}$$

In fact, if $f = \frac{1}{2\pi}$, then we obtain $\|\Lambda\| = \frac{1}{2\pi}$. Thus, Λ is bounded.

Step 2: To see that Λ is injective, let $f \in L^1(T)$ with $\hat{f} \equiv 0$. If p is a trigonometric polynomial, $p = \sum_{k=-n}^{n} c_k e^{ikt}$, then by assumption

$$\int_{-\pi}^{\pi} fp\,\mathrm{d}t = 0.$$

Given $g \in C(T)$, there is a sequence of bounded trigonometric polynomials $\{p_n\}$, such that $f(x)p_n(x) \to f(x)g(x)$, for each $x \in T$. Thus, by Dominated Convergence (1.7.5),

$$\int_{-\pi}^{\pi} fg\,\mathrm{d}t = \lim_{n\to\infty} \int_{-\pi}^{\pi} fp_n\,\mathrm{d}t = 0.$$

By Corollary (2.5.2), there is $\{g_n\} \subset C(T)$ such that $|g_n| < |f|$ and $g_n(x) \to f(x)$ a.e. Therefore, by Dominated Convergence, we have

$$\int_{\pi}^{-\pi} f \,\mathrm{d}t = \lim_{n\to\infty} \int_{\{g_n = f\}} g_n \,\mathrm{d}t = 0.$$

Therefore, by Theorem (1.8.5), $f = 0$ a.e., and Λ is injective.

Step 3: Finally, if Λ is surjective, then the Bounded Inverse Theorem implies that $\|\Lambda^{-1}\| < \infty$. Hence, there is $M > 0$, such that for all $\hat{f} \in C_0$, with $\|\hat{f}\|_\infty \leq 1$, $\|\Lambda^{-1}(\hat{f})\|_1 \leq M$. Consider the sequence of Dirichlet's kernels D_n. By 2π-periodicity,

$$\begin{aligned}
\Lambda(D_n)(k) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \Bigg(\sum_{j=-n}^{n} e^{ijt} \Bigg) e^{-ikt} \,\mathrm{d}t \\
&= \frac{1}{2\pi} \sum_{j=-n}^{n} \int_{-\pi}^{\pi} e^{it(j-k)} \,\mathrm{d}t \\
&= \frac{1}{2\pi} \sum_{j=-n}^{n} \int_{-\pi}^{\pi} \cos((j-k)t) + \sin((j-k)t) \,\mathrm{d}t \\
&= \begin{cases} 1, & |k| \leq n, \\ 0, & \text{else.} \end{cases}
\end{aligned}$$

We see that $\|\hat{D_n}\|_\infty = 1$, but $\|D_n\|_1 \to \infty$, which is a contradiction. So, the inverse of the Riemann-Lebesgue Lemma does not hold. ■

5.5 The Hahn-Banach Theorem

Proposition 5.5.1. *Let V be a complex vector space.*

(a) Let $f\colon V \to \mathbb{C}$ be linear and $u := \mathrm{Re}(f)$. Then

$$f(x) = u(x) - iu(ix), \quad \textit{for all } x \in V. \tag{5.5.1}$$

(b) If $u\colon V \to \mathbb{R}$ is linear and $f\colon V \to \mathbb{C}$ defined by (5.5.1), then f is linear.

(c) If V is a a normed vector space and f and u are related as in (5.5.1), then $\|u\| = \|f\|$.

Proof. (a). Let $z = \alpha + i\beta$, $\alpha, \beta \in \mathbb{R}$. Then $\mathrm{Re}(iz) = -\beta$, and $z = \mathrm{Re}(z) - i\,\mathrm{Re}(iz)$. Also, $\mathrm{Re}(if(x)) = \mathrm{Re}(f(ix)) = u(ix)$.

(b). Obviously f is real linear. Moreover,

$$f(ix) = u(ix) - iu(-x) = u(ix) + iu(x) = i(u(x) - iu(ix)) = if(x).$$

Hence f is also complex linear.

(c). It is obvious that $|u(x)| \leq |f(x)|$ hence $\|u\| \leq \|f\|$. On the other hand, $\forall x \neq 0 \in V$, let $\alpha = \frac{|f(x)|}{f(x)}$. Then, $|f(x)| = \alpha f(x) = f(\alpha x) \in \mathbb{R}$, which equals $u(\alpha x) \leq \|u\|\|\alpha x\| = \|u\|$. ∎

Theorem 5.5.2 (**Hahn-Banach Theorem**). *Let $(X, \|\cdot\|)$ be a normed vector space. Let M be a subspace of X, and $\lambda : M \to \mathbb{C}$ be bounded linear. Then, λ can be extended to a bounded linear $\Lambda : X \to \mathbb{C}$, such that $\Lambda|_M = \lambda$ and $\|\Lambda\| = \|\lambda\|$.*

Proof. Without loss of generality, we assume that $|\lambda(x)| \leq \|x\|$, for all $x \in X$, and $M \subsetneq X$. **Step 1:** First suppose $\lambda\colon X \to \mathbb{R}$. We will extend λ from M to some subspace N. Pick $x_0 \in X \setminus M$, then for all $x, y \in M$, we have

$$\begin{aligned} \lambda(x) - \lambda(y) &= \lambda(x - y) \\ &\leq \|x - y\| \\ &\leq \|x - x_0\| + \|x_0 - y\| \\ \lambda(x) - \|x - x_0\| &\leq \lambda(y) + \|x_0 - y\|. \end{aligned} \tag{5.5.2}$$

Since inequality (5.5.2) holds for any $x, y \in M$,

$$\sup_{x\in M}\{\lambda(x) - \|x - x_0\|\} \leq \inf_{x\in M}\{\lambda(x) + \|x - x_0\|\},$$

and there is $\alpha \in \mathbb{R}$ such that for all $x \in M$,

$$\lambda(x) - \|x - x_0\| \leq \alpha \leq \lambda(x) + \|x - x_0\|.$$

Hence, $|\lambda(x) - \alpha| \leq \|x - x_0\|$. Now, if $c \neq 0 \in \mathbb{R}$, then $x/c \in M$ and thus

$$\begin{aligned} |\lambda(x/c) - \alpha| \leq \|x/c - x_0\| \quad &\Rightarrow \quad |c||\lambda(x/c) - \alpha| \leq |c|\|x/c - x_0\| \\ &\Rightarrow \quad |\lambda(x - c\alpha)| \leq \|x - cx_0\|. \end{aligned} \tag{5.5.3}$$

Define $N := \{m + cx_0 : c \in \mathbb{R}, m \in M\}$. Then, N is a linear subspace and $M \subsetneq N$. Define $f\colon N \to \mathbb{R}$ by

$$f(m + cx_0) := \lambda(m) + c\alpha.$$

We see that $f|_M = \lambda$. By (5.5.3), $|f(m + cx_0)| \leq \|m + cx_0\|$, hence $\|f\| = \|\lambda\|$.

Step 2: Use Zorn's Lemma. Define

$$\mathcal{S} := \{(N, f) : M \subset N \subset X,\ f|_M = \lambda,\ f \text{ linear },\|f\| = \|\lambda\|\}.$$

From Step 1, we know $\mathcal{S}$ is not empty. Define a partial order on $\mathcal{S}$ by $(N', g') < (N, g) \iff N' \subset N$ and g extends g'. Suppose $\{(N_i, f_i)\}_{i\in I}$ is a totally ordered chain in $\mathcal{S}$. Then, $\bigcup_{i\in I} N_i$ is a linear subspace. Define $h(x) = f_i(x), i \in I$. Note that h is well-defined because of extension. Moreover, h is linear, $h|_M = \lambda$ and $|h(x)| = |f_i(x)| \leq \|x\|$. Therefore, $(\bigcup_{i\in I} N_i, h) \in \mathcal{S}$ and it is a maximal element of such chain.

Zorn's Lemma gives $(Z, \Lambda) \in \mathcal{S}$ such that Λ on Z extends λ. Obviously, $Z = X$; otherwise, we can pick $x' \in X \setminus Z$ as in Step 1.

Step 3: Now, if $\lambda\colon M \to \mathbb{C}$, let $u := \mathrm{Re}(\lambda)$. Then $u\colon M \to \mathbb{R}$. From the previous steps, we can extend u to $U\colon X \to \mathbb{R}$. Define $F\colon X \to \mathbb{C}$ as in (5.5.1), $\Lambda(x) := U(x) - iU(ix)$. By Proposition (5.1.1), Λ is a complex linear functional with $\|\Lambda\| = \|U\| = \|u\|$, and for each $x \in M$,

$$\Lambda(x) = U(x) - iU(ix) = u(x) - iu(ix) = \lambda(x). \qquad \blacksquare$$

Corollary 5.5.3. *Let X, M be the same as above. A point $x_0 \in \overline{M} \iff \nexists$ bounded linear functional Λ on X s.t. $\Lambda|_M = 0$ and $\Lambda(x_0) \neq 0$.*

Equivalently, a point $x_0 \notin \overline{M} \iff$ there is a nonzero bounded linear functional Λ on X such that $\Lambda|_M = 0$.

Proof. ($\Rightarrow$). Let $x_0 \in \overline{M}$, and $\Lambda|_M = 0$. Then, by sequential continuity of Λ, $\Lambda(x_0) = 0$. ($\Leftarrow$). By contraposition, if $x_0 \notin \overline{M}$, then there is a $B(x_0, \delta) \subset \overline{M}^C$. Let $N := \mathrm{Span}(M, \{x_0\})$, and $\lambda\colon N \to \mathbb{C}$, by $\lambda(x + cx_0) := c$. Note that λ is a linear functional on N, and

$$\|x + cx_0\| = |c|\|x/c + x_0\| \leq |c|\delta.$$

By the Hahn-Banach Theorem, we can extend λ to Λ on X, then we see that $\Lambda(x_0) = \lambda(x_0) = 1$. $\blacksquare$

Definition 5.5.4 (Dual spaces). Let X be a nomred vector space. We define the **dual space** of X as

$$X^* := \{\Lambda\colon X \to \mathbb{C} \mid \Lambda \text{ is bounded linear}\}.$$

It is trivial to see that X^* is a normed vector space. In fact, if X is a Banach space, then so is X^*.

Corollary 5.5.5. *Let $(X, \|\cdot\|)$ be a normed vector space and $x_0 \neq 0 \in X$. Then $\exists \Lambda \in X^*$ such that $\|\Lambda\| = 1$ and $\Lambda(x_0) = \|x_0\|$.*

Proof. Let $M := \mathbb{C}x_0$ and $\lambda\colon M \to \mathbb{C}$ by $\lambda(cx_0) := c\|x_0\|$. Then, λ is a bounded linear functional with $\|\lambda\| = 1$. By Hahn-Banach Theorem, we obtain $\Lambda\colon X \to \mathbb{C}$ as desired. $\blacksquare$

Corollary 5.5.6. *Let X be a normed vector space.*

(i) If $x_1 \neq x_2 \in X$, then there is $\Lambda \in X^*$ such that $\Lambda(x_1) \neq \Lambda(x_2)$.

(ii) For each $x \in X$, define $\lambda_x\colon X^* \to \mathbb{C}$ by $\lambda_x(f) := f(x)$. Then the map $x \mapsto \lambda_x$ is an isometry from X to $(X^*)^*$.

Proof. To see (i), simply apply the previous corollary on $x_1 - x_2 \neq 0$. For (ii), it is not hard to check that λ_x is a linear functional on X^*. We call such map **point-evaluation functional**. Also, $|\lambda_x(f)| = |f(x)| \leq \|f\|\|x\|$, thus $\|\lambda_x\| \leq \|x\|$. On the other hand, the previous corollary gives $f \in X^*$ with $\|\lambda_x\| \geq |f(x)| = \|x\|$. $\blacksquare$

Remark. Note that in (ii), the map $x \mapsto \lambda_x$ is just an isometry, not necessarily bijective. In general, if X is infinite dimensional, then $(X^*)^*$ is a "superset" of X.

Question. Is the extension given by the Hahn-Banach Theorem unique?

Solution. No. Consider $X := L^\infty(T)$ with only real-valued functions and let $M := C(T) \subset X$. Define $\Lambda\colon M \to \mathbb{R}$ by $\Lambda(f) := f(0)$. Then Λ is linear and bounded since $|\Lambda(f)| = |f(0)| \leq \|f\|_\infty$. By the Hahn-Banach Theorem, there is an extension $\widetilde{\Lambda}$ on X. We will find another extension.

Consider $C' := \{f \in X \mid f \text{ is continuous at } x \neq 0, \text{ and } f(0^-), f(0^+) \text{ exist}\}$. We see that $C(T) \subsetneq C' \subsetneq X$. Now, for all $\alpha \in [0,1]$, define $\Lambda'_\alpha\colon C' \to \mathbb{R}$ by

$$\Lambda'_\alpha(f) := \alpha f(0^-) + (1-\alpha) f(0^+).$$

Note that if $f \in C(T)$, then $\Lambda'_\alpha(f) = \Lambda(f) = f(0)$. Hence, Λ'_α extends Λ. Also,

$$\begin{aligned}|\Lambda'_\alpha(f)| &= |\alpha f(0^-) + (1-\alpha) f(0^+)| \\ &\leq \alpha |f(0^-)| + (1-\alpha)|f(0^+)| \\ &\leq (\alpha + 1 - \alpha)\|f\|_\infty = \|f\|_\infty.\end{aligned}$$

So Λ'_α is a bounded linear functional on C'. Now, applying the Hahn-Banach Theorem on $\{\Lambda'_\alpha : \alpha \in [0,1]\}$, we obtain distinct extensions of Λ on X. ■

5.6 Uniqueness of Point Evaluation Funcionals and the Poisson Integral

In the previous section, we see that in general not all point-evaluation functionals can be extended uniquely. We shall see that there is a unique extension of such functional on certain spaces. We begin with the following theorem.

Theorem 5.6.1. *Let $D := \{z \in \mathbb{C} : |z| < 1\}$ be the open unit disk. Let $p(z) := \sum_{k=0}^n p_k z^k$ be a polynomial on $\overline{D}$. Then,* $\max\{|p(z)| : z \in \overline{D}\} = \max\{|p(z)| : z \in \partial D\}$. *Equivalently,* $\|p\|_{\infty,\overline{D}} = \|p\|_{\infty,\partial D}$.

Proof. First, since $\overline{D}$ is compact, we know $|p|$ attains its maxium at some $z \in \overline{D}$. Suppose there is $z_0 \in D$ such that $p(z_0) \geq p(z)$, for all $z \in \overline{D}$. We will show that p must be a constant function. Write

$$p(z) = \sum_{k=0}^n q_k (z - z_0)^k,$$

for some $q_k \in \mathbb{C}$. Since $z_0 \in D$, there is an open disk $B(z_0, r) \subset D$, and if $z \in B(z_0, r), z = z_\theta = z_0 + re^{i\theta}$, for some θ. Since $\int_0^{2\pi} (e^{i\theta})^m \,\mathrm{d}\theta = 0$, given $m \neq 0$,

$$\frac{1}{2\pi} \int_0^{2\pi} |p(z_\theta)|^2 \,\mathrm{d}\theta = \frac{1}{2\pi} \int_0^{2\pi} \left| \sum_{k=0}^n q_k (re^{i\theta})^k \right|^2 \mathrm{d}\theta$$

$$= \sum_{k=0}^{n} |q_k|^2.$$

On the other hand,

$$\frac{1}{2\pi}\int_0^{2\pi} |p(z_\theta)|^2 \, \mathrm{d}\theta \le \frac{1}{2\pi}\int_0^{2\pi} |p(z_0)|^2 \, \mathrm{d}\theta = |q_0|^2.$$

Thus, for every $k > 0$, $q_k = 0$, and $p(z) = q_0$, is constant on $\overline{D}$.

In conclusion, every polynomial attains its maxium modulus on ∂D. ∎

Remark. This is a special case of the **maxmium modulus theorem**.

Solution. Let D be the open unit disk in $\mathbb{C}$, $A(\overline{D})$ be a subspace of $C(\overline{D})$ such that if $f \in A(\overline{D})$, $\|f\|_{\infty,\overline{D}} = \|f\|_{\infty,\partial D}$. In this example, we choose $A(\overline{D})$ to be the closure of all the polynomials on $\overline{D}$, with the $\|\cdot\|_\infty$ norm, for the following reason:

Claim: For all $f \in A(\overline{D})$, $\|f\|_{\infty,\overline{D}} = \|f\|_{\infty,\partial D}$. In fact, if $f \in C(\overline{D})$, there exists polynomials $p_n \to f$ uniformly by the Stone-Weierstrass Theorem. Hence, $\|p_n - f\|_{\infty,\overline{D}} \to 0$. By the previous theorem, for all polynomial p on $\overline{D}$, $\|p\|_{\infty,\overline{D}} = \|p\|_{\infty,\partial D}$. Therefore, $\|f\|_{\infty,\overline{D}} = \|f\|_{\infty,\partial D}$.

Step 1: Let $A(\partial D) \subset A(\overline{D})$ be the subspace whose functions are restriced on ∂D. From the norm-preserving property, we see that the linear functional $f \mapsto f|_{\partial D}$ is an isomorphic isometry from $A(\overline{D})$ to $A(\partial D)$, with respect to $\|\cdot\|_\infty$. In other words, $A(\partial D) = A(\overline{D})$ as Banach spaces.

Step 2: For each $z \in D$, define $\lambda_z \colon A(\overline{D}) \to \mathbb{C}$ by $\lambda_z(f) := f(z)$. Thus,

$$|\lambda_z(f)| = |f(z)| \le \|f\|_{\infty,D}.$$

is bounded. In fact, $\|\lambda_z\| = 1$ since $\lambda_z(1) = 1$. From Step 1, $A(\overline{D}) = A(\partial D) \subset C(\partial D)$. By the Hahn-Banach Theorem, we extend λ_z to Λ_z on $C(\partial D)$. Thus, $\Lambda_z(f) = \lambda_z(f) = f(z)$, for all $f \in A(\overline{D})$, and $\|\Lambda_z\| = \|\lambda_z\| = 1$.

Step 3: Claim: Λ_z is a **positive linear functional**. To see this, without loss of generality, suppose $f \in C(\partial D)$, with $0 \le f \le 1$. Define $g := 2f - 1$, so $-1 \le g \le 1$ and $|\Lambda_z(g)| \le 1$. Then, given $r \in \mathbb{R}$,

$$|g + ir|^2 \le \|g\|_{\infty,\partial D}^2 + |r|^2 \le 1 + r^2.$$

Thus, viewing ir as a constant function, we see that

$$\begin{aligned}(\mathrm{Im}(\Lambda_z(g)) + r)^2 &\le |\operatorname{Re}(\Lambda_z(g)) + i(\mathrm{Im}(\Lambda_z(g)) + r)|^2 \\ &= |\Lambda_z(g) + ir|^2 = |\Lambda_z(g + ir)|^2 \\ &\le \|\Lambda_z\|^2 \cdot |g + ir|^2 \le 1 + r^2.\end{aligned}$$

It follows that

$$0 \le (1 + r)^2 - (\mathrm{Im}(\Lambda_z(g)) + r)^2$$

$$= -(\mathrm{Im}(\Lambda_z(g)))^2 - 2r\,\mathrm{Im}(\Lambda_z(g)) + 1, \quad \forall r \in \mathbb{R},$$

which is only possible if $\mathrm{Im}(\Lambda_z(g)) = 0$. Hence, $\Lambda_z(g) \in \mathbb{R}$. Moreover,

$$\Lambda_z(f) = \Lambda_z(\frac{1}{2} + \frac{g}{2}) = \frac{1}{2}\Lambda_z(1) + \frac{1}{2}\Lambda_z(g).$$

Recall that $\Lambda_z(g) \in [-1, 1]$. Then, $\Lambda_z(f) \geq 0$, for all $f \geq 0$, and Λ_z is a positive linear functional.

Step 4: By the **Riesz Representation Theorem** (2.3.1), for each $z \in \overline{D}$, there is a *unique* regular positive Borel measure μ_z on ∂D such that

$$\Lambda_z(f) = \int_{\partial D} f \,\mathrm{d}\mu_z, \quad \text{for all } f \in C_c(\partial D) = C(\partial D). \tag{5.6.1}$$

In particular, we have

$$f(z) = \Lambda_z(f) = \int_{\partial D} f \,\mathrm{d}\mu_z, \quad \text{for all } f \in A(\overline{D}) = A(\partial D). \tag{5.6.2}$$

Step 5: Finally, for each $f \in C(\partial D)$, the Stone-Weierstrass Theorem gives us a sequence of trigonometric polynomials $\{p_n\}_{n=1}^{\infty}$ in $A(\overline{D})$, such that $p_n \to f$ uniformly. So, we may assume $|p_n| \leq |f|$. By Dominated Convergence,

$$\int_{\partial D} |p_n - f| \,\mathrm{d}\mu_{z_0} \to 0 \quad \Rightarrow \quad \int_{\partial D} p_n \,\mathrm{d}\mu_{z_0} \to \int_{\partial D} f \,\mathrm{d}\mu_{z_0}.$$

Consequently,

$$\lim_{n\to\infty} p_n(z_0) = \lim_{n\to\infty} \Lambda_{z_0}(p_n) = \Lambda_{z_0}(f),$$

which must be unique. ■

Remark 5.6.2. Note that from Step 1 to Step 4, it can be done abstractly by replacing $\overline{D}$ by a compact Hausdorff space K, and ∂D by a compact subset H of K. All the results hold true up to this point. In equation (5.6.1), we see that μ_z is uniquely determined by Λ_z, but the extension itself might not be unique. In Step 5, the Stone-Weierstrass Theorem passes the functional to sequential limit, which is then unique.

Remark 5.6.3. When we identify $A(\overline{D})$ with $A(\partial D)$, it seems like we are losing information on D. However, we can determine the value of f on D by the representation in equation (5.6.2). This result is remarkable as it proves that in fact, we do not lose anything. However, in practice the problem in equation (5.6.2) arises when it comes to finding the measure μ_z. We want to find a more concrete formula to compute $f(z)$. Here we introduce the Poisson integral.

Poisson Integral.

Step 1: A consequence of equation (5.6.2): Fix $z_0 = re^{in\theta} \in \overline{D}$, for some $0 \leq r < 1$, $\theta \in \mathbb{R}$. For each $f \in C(\partial D)$,

$$\Lambda_{z_0}(\overline{f}) = \int_{\partial D} \overline{f} \,\mathrm{d}\mu_{z_0} = \overline{\int_{\partial D} f \,\mathrm{d}\mu_{z_)}} = \overline{\Lambda_{z_0}(f)}.$$

For $n \in \mathbb{N}$, define $u_n(x) := x^n$. Then $u_n \in C(\partial D)$, and $u_{-n}(x) = x^{-n} = \overline{x}^n = \overline{u_n}(x)$. We conclude that

$$\Lambda_{z_0}(u_{-n}) = \Lambda_{z_0}(\overline{u_n}) = \overline{\Lambda_{z_0}(u_n)} = \overline{r^n e^{in\theta}} = r^n e^{-in\theta}.$$

Therefore, for each $n \in \mathbb{Z}$,

$$\Lambda_{z_0}(u_n) = \int_{\partial D} u_n \, \mathrm{d}\mu_{z_0} = r^{|n|} e^{in\theta}. \tag{5.6.3}$$

Step 2: Consider the function $P_{r,\theta} \in C(\partial D)$, for $t \in [0, 2\pi]$,

$$P_{r,\theta}(t) := \sum_{n\in\mathbb{Z}} r^{|n|} e^{in(\theta - t)} = \sum_{n\in\mathbb{Z}} u_n(z_0) e^{-int}. \tag{5.6.4}$$

Since $|r| < 1$, $P_{r,\theta}(t)$ is absolutely convergent. For each $k \in \mathbb{Z}$, we can integrate the following term by term:

$$\begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} P_{r,\theta}(t) e^{ikt} \, \mathrm{d}t &= \frac{1}{2\pi} \int_0^{2\pi} \sum_{n\in\mathbb{Z}} u_n(z_0) e^{-int} \cdot e^{ikt} \, \mathrm{d}t \\ &= \frac{1}{2\pi} \sum_{n\in\mathbb{Z}} \int_0^{2\pi} u_n(z_0) e^{i(-n+k)t} \, \mathrm{d}t \\ &= u_k(z_0) = r^{|k|} e^{ik\theta}. \end{aligned} \tag{5.6.5}$$

Step 3: Therefore, if $f = u_n$, by equations (5.6.3) and (5.6.5), we see that

$$\int_{\partial D} f \, \mathrm{d}\mu_{z_0} = \frac{1}{2\pi} \int_0^{2\pi} P_{r,\theta}(t) f(e^{it}) \, \mathrm{d}t. \tag{5.6.6}$$

Since every trigonometric polynomial p is a finite linear combination of the u_n's, (5.6.6) also holds for p. By the Stone-Weierstrass and Dominated Convergence, it also holds for all $f \in C(\partial D)$. In particular, if $f \in A(\overline{D}) = A(\partial D)$, equation (5.6.2) gives

$$f(z_0) = \frac{1}{2\pi} \int_0^{2\pi} P_{r,\theta}(t) f(e^{it}) \, \mathrm{d}t. \tag{5.6.7}$$

Step 4: From equation (5.6.4), $P_{r,\theta} = \overline{P_{r,\theta}}$ shows that $P_{r,\theta}$ is a real-valued function. So,

$$\begin{aligned} P_{r,\theta}(t) &= \operatorname{Re}\left(\sum_{n\in\mathbb{Z}} r^{|n|} e^{in(\theta - t)} \right) = 1 + 2 \operatorname{Re}\left(\sum_{n=1}^{\infty} r^n e^{in(\theta - t)} \right) \\ &= 1 + 2 \operatorname{Re}\left(\sum_{n=1}^{\infty} (z_0 e^{-it})^n \right) = 1 + \operatorname{Re}\left(\frac{2 z_0 e^{-it}}{1 - z_0 e^{-it}} \right) \\ &= \operatorname{Re}\left(\frac{1 + z_0 e^{-it}}{1 - z_0 e^{-it}} \right) = \frac{\operatorname{Re}((1 + z_0 e^{-it})(1 - \overline{z_0 e^{-it}}))}{|1 - z_0 e^{-it}|^2} \end{aligned}$$

$$= \frac{\operatorname{Re}(1 + 2ir\sin(\theta - t) - r^2)}{|1 - z_0 e^{-it}|^2} = \frac{1 - r^2}{1 - 2r\cos(\theta - t) + r^2}.$$

We call

$$P_{r,\theta}(t) = \frac{1 - r^2}{1 - 2r\cos(\theta - t) + r^2} \tag{5.6.8}$$

the Poisson kernel. ■

Finally, we conclude this section by summarizing the result.

Theorem 5.6.4. *Let $A(\overline{D})$ be the space of continuous complex functions on $\overline{D}$. Suppose A contains all polynomials and for each $f \in A$,*

$$\sup\{|f(z)| : z \in D\} = \sup\{|f(z)| : z \in \partial D\}.$$

Then for all $f \in A(\overline{D})$, the Poisson integral representation

$$f(z) = \frac{1}{2\pi}\int_0^{2\pi} \frac{1 - r^2}{1 - 2r\cos(\theta - t) + r^2} f(e^{it})\,\mathrm{d}t \tag{5.6.9}$$

holds for every $z \in D$, where $z = re^{i\theta}$.

Chapter 6

Complex Measure

Let (X, M) be a measure space throughout the chapter.

6.1 Total Variation Measure

Definition 6.1.1. A **measurable partition** of $E \in M$ is a sequence $\{E_n\}_{n=1}^{\infty} \subset M$, such that $E_i \cap E_j = \varnothing$, for all $i \neq j$ and $\bigcup_{n=1}^{\infty} = E$.

Definition 6.1.2 (Complex measure)**.** Let M be a σ-algebra. A **complex measure** μ on M is a set function $\mu\colon M \to \mathbb{C}$ such that

$$\mu(E) = \sum_{n=1}^{\infty} \mu(E_n), \tag{6.1.1}$$

for each measurable partition $\{E_n\}_{n=1}^{\infty}$ of E.

Remark 6.1.3. Unlike positive measure, the *convergence of the series* in $\mathbb{C}$ in equality (6.1.1) is now required. Thus, a positive measure is not necessarily a complex measure!

Remark 6.1.4. Permutation of the $\mu(E_n)$'s does not change $\mu(E)$. Hence, $\sum_{n=1}^{\infty} \mu(E_n)$ is absolutely convergent by the Riemann Series Theorem.

Definition 6.1.5 (Total variation)**.** The **total variation** of μ is a set function $|\mu|\colon M \to \mathbb{R}$, defined as

$$|\mu|(E) := \sup\left\{ \sum_{n=1}^{\infty} |\mu(E_n)| : \{E_n\}_{n=1}^{\infty} \text{ is a partition of } E \right\}. \tag{6.1.2}$$

Remark. For all $E \in M$, $|\mu(E)| \leq |\mu|(E)$.

Proposition 6.1.6. *The total variation $|\mu|$ is a positive measure on X.*

Proof. We will prove (6.1.1). Let $E \in M$ and $\{E_n\}_{n=1}^{\infty}$ be a partition of E. For each $n \in \mathbb{N}$, $\forall t_n \in \mathbb{R}$, such that $t_n < |\mu|(E_n)$, there exists partition $\{A_{n,m}\}_{m=1}^{\infty}$ of E_n such that

$$t_n < \sum_{m=1}^{\infty} |\mu(A_{n,m})|,$$

by definition of $|\mu|(E_n)$. Note that $\{A_{n,m} : n, m \in \mathbb{N}\}$ is a partition of E. Hence,

$$\sum_{n=1}^{\infty} t_n \leq \sum_{n=1}^{\infty} \left(\sum_{m=1}^{\infty} |\mu(A_{n,m})| \right) \leq |\mu|(E).$$

Take the supremum over all $t_n < |\mu|(E_n)$, we have

$$\sum_{n=1}^{\infty} |\mu|(E_n) \leq |\mu|(E).$$

On the other hand, for each partition $\{A_m\}_{m=1}^{\infty}$ of E, we have

$$\begin{aligned}
\sum_{m=1}^{\infty} |\mu(A_m)| = \sum_{m=1}^{\infty} \left| \sum_{n=1}^{\infty} \mu(A_m \cap E_n) \right| &\leq \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |\mu(A_m \cap E_n)| \\
= \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |\mu(A_m \cap E_n)| &\leq \sum_{n=1}^{\infty} |\mu|(E_n).
\end{aligned}$$

Since it holds for any partition $\{A_m\}_{m=1}^{\infty}$, so does the supremum over all partitions,

$$|\mu|(E) \leq \sum_{n=1}^{\infty} |\mu|(E_n).$$

Hence, $|\mu|$ satisfies countable addivity. Also, $|\mu|(\varnothing) = 0$. We will see $|\mu|(X) < \infty$ in the next proposition. ■

Lemma 6.1.7. *If $\{z_1, \ldots, z_N\} \subset \mathbb{C}$, then there is $S \subseteq \{1, \ldots, N\}$ such that*

$$\left| \sum_{k \in S} z_k \right| \geq \frac{1}{\pi} \sum_{i=k}^{N} |z_i|.$$

Proof. Write $z_k = |z_k| e^{i\alpha_k}$. Fix $\theta \in [0, 2\pi]$, and define $S(\theta) := \{k : \cos(\theta - \alpha_k) > 0\}$. Then,

$$\begin{aligned}
\left| \sum_{k \in S(\theta)} z_k \right| &= \left| |e^{-i\theta}| \sum_{k \in S(\theta)} z_k \right| = \left| \sum_{k \in S(\theta)} e^{-i\theta} z_k \right| \\
&\geq \mathrm{Re} \left(\sum_{k \in S(\theta)} e^{-i\theta} z_k \right) = \sum_{k \in S(\theta)} \mathrm{Re}(|z_k| e^{i(\alpha_k - \theta)}) \\
&= \sum_{k \in S(\theta)} |z_k| \cos(\alpha_k - \theta) = \sum_{k=}^{N} |z_k| \cos^{+}(\alpha_k - \theta), \qquad (6.1.3)
\end{aligned}$$

where $\cos^+(x) := \max\{\cos(x), 0\}$. Integrating inequality (6.1.3) over θ gives

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi}\left|\sum_{k\in S(\theta)} z_k\right| \mathrm{d}\theta &\geq \frac{1}{2\pi}\int_0^{2\pi}\sum_{k=1}^{N}|z_k|\cos^+(\alpha_k-\theta)\,\mathrm{d}\theta \\ &= \frac{1}{2\pi}\sum_{k=1}^{N}|z_k|\int_0^{2\pi}\cos^+(\alpha_k-\theta)\,\mathrm{d}\theta \\ &= \frac{1}{\pi}\sum_{k=1}^{N}|z_k|.\end{aligned}$$

The Mean Value Theorem for Integration asserts there is a θ such that

$$\left|\sum_{k\in S(\theta)} z_k\right| \geq \frac{1}{\pi}\sum_{k=1}^{N}|z_k|.$$

■

Proposition 6.1.8. *Let μ be a complex measure on X. Then $|\mu|(X) < \infty$.*

Proof. Suppose there is $E \in M$ with $|\mu|(E) = \infty$. Define $t := \pi(1 + |\mu(E)|) < \infty$. The definition of $|\mu|$ asserts a finite partition $\{E_k\}_{k=1}^N$ of E such that

$$\sum_{k=1}^{N}|\mu(E_k)| > t.$$

Lemma (6.1.7) gives $S \subset \{1, \ldots, N\}$, with $A := \bigcup_{k\in S} E_k$ such that

$$|\mu(A)| = \left|\sum_{k\in S}\mu(E_k)\right| \geq \frac{1}{\pi}\sum_{k=1}^{N}|\mu(E_k)| > \frac{t}{\pi} > 1.$$

Let $B := E \setminus A$, then $A \cap B = \varnothing$ and

$$|\mu(B)| = |\mu(E) - \mu(A)| \geq |\mu(A)| - |\mu(E)| > \frac{t}{\pi} - |\mu(E)| > 1.$$

Since $|\mu|(E) = \infty$, without loss of generality, we assume $|\mu|(B) = \infty$. Now, let $E = X$, then $X = A_1 \cup B_1$, with $|\mu|(B_1) = \infty$. Then let $B_1 = A_2 \cup B_2$ with $|\mu(B_2) = \infty$ and proceed inductively. Observe that $A_i \cap A_j = \varnothing, \forall i \neq j$. Hence,

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty}\mu(A_i),$$

with $|\mu(A_i)| > 1$. However, then $\sum_{i=1}^{\infty}\mu(A_i)$ diverges, which is a contradiction to μ being a complex measure. ■

Remark. Hence the range of μ is a subset of a finite disk in $\mathbb{C}$. We sometime say μ is of **bounded variation**.

Definition 6.1.9. Define $\mathcal{M}(M)$ to be the set of all complex measures on (X, M). For all $\mu, \nu \in \mathcal{M}(M)$, $\alpha \in \mathbb{C}$, define $\mu + \alpha\nu \colon M \to \mathbb{C}$ by $(\mu + \alpha\nu)(E) := \mu(E) + \alpha\nu(E)$. Hence, $\mathcal{M}(M)$ is a complex vector space. Moreover, define $\|\mu\| := |\mu|(X)$. Then, $\mathcal{M}(M)$ is a normed vector space.

Definition 6.1.10 (Positive and Negative Variations)**.** Let μ be a real measure on X, define

$$\mu_+ := \frac{1}{2}(|\mu| + \mu) \qquad \text{and} \qquad \mu_- := \frac{1}{2}(|\mu| - \mu). \tag{6.1.4}$$

Both μ_+ and μ_- are positive real measures. They are called **positive** and **negative variations** of μ, respectively. The representation is also called the **Jordan decomposition**.

6.2 Absolute Continuity

Definition 6.2.1. Let μ be a positive measure, and λ be any measure (positive or complex) on (X, M). Let $A, B \in M$.

- λ is **absolutely continuous** with respect to μ if $\lambda(E) = 0 \Rightarrow \mu(E) = 0, \forall E \in M$. We write it as $\mu \ll \lambda$.
- λ is **concentrated on** A if $\lambda(E) = \lambda(E \cap A), \forall E \in M$.
- Suppose $A \cap B = \varnothing$ and λ_1, λ_2 are measures on M. If λ_1 is concentrated on A and λ_2 is concentrated on B, then we say λ_1 and λ_2 are **mutually disjoint**, and denote it $\lambda_1 \perp \lambda_2$.

Proposition 6.2.2. *Let $\mu, \lambda, \lambda_1, \lambda_2$ be measures on (X, M) and μ be positive.*

(a) If λ is concentrated on A, so is $|\lambda|$.

(b) $\lambda_1 \perp \lambda_2 \quad \Rightarrow \quad |\lambda_1| \perp |\lambda_2|$.

(c) $\lambda_1 \perp \mu$ *and* $\lambda_2 \perp \mu \quad \Rightarrow \quad (\lambda_1 + \lambda_2) \perp \mu$.

(d) $\lambda_1 \ll \mu$ *and* $\lambda_2 \ll \mu \quad \Rightarrow \quad (\lambda_1 + \lambda_2) \ll \mu$.

(e) $\lambda \ll \mu \quad \Rightarrow \quad |\lambda| \ll \mu$.

(f) $\lambda_1 \ll \mu$ *and* $\lambda_2 \perp \mu \quad \Rightarrow \quad \lambda_1 \perp \lambda_2$.

(g) $\lambda \ll \mu$ *and* $\lambda \perp \mu \quad \Rightarrow \quad \lambda \equiv 0$.

Proof. All are obvious from the definitions. ■

We now come the core of this chapter, the Lebesgue-Radon-Nikodym Theorem. It is one of the most important theorems in measure thoery. It uniquely decomposes any complex measure into its *absolute continuous* and *mutually singular* parts relative to a positive σ-finite measure. More importantly, it provides a conditional converse to Theorem (1.6.8), which passes integrals to measures. This result is remarkable as we have seen the deep connections among integrals, measures and linear functionals throughout the course.

Theorem 6.2.3 (**Lebesgue-Radon-Nikodym Theorem**). *Let μ be a positive σ-finite, λ be a complex measure on (X, M), λ. Then,*

(a) There is a unique pair of complex measures λ_a and λ_s such that

$$\lambda = \lambda_a + \lambda_s, \qquad \lambda_a \ll \mu, \qquad \lambda_s \perp \mu.$$

(b) There is a unique $h \in L^1(\mu)$ such that

$$\lambda_a(E) = \int_E h \,\mathrm{d}\mu, \quad \forall E \in M. \tag{6.2.1}$$

Remark 6.2.4. The pair (λ_a, λ_s) is called the **Lebesgue decomposition** of λ relative to μ. Note that if $\lambda \ll \mu$, then $\lambda_a = \lambda$, and we can pass measure to integral: $\lambda(E) = \int_E h \,\mathrm{d}\mu$. The function $h \in L^1(\mu)$ is called the **Radon-Nikodym derivative** of λ_a with respect to μ. We write $d\lambda_a = h \,\mathrm{d}\mu$ or $h = \frac{d\lambda_a}{d\mu}$.

Proof of Uniqueness. Due to the length of the proof, we shall show the uniqueness part here. If (λ_a, λ_s) and (λ'_a, λ'_s) both satisfies (a), then $\lambda'_a - \lambda_a = \lambda_s - \lambda'_s$. Since $\lambda'_a - \lambda_a \ll \mu$ and $\lambda_s - \lambda'_s \perp \mu$, the equality must be 0, and the uniqueness follows. For h, if $h' - h \neq 0$ on $E \in M$, then $\mu(E) = 0$, and so $h = h'$, μ-a.e. ∎

Proof. **Step 1:** First suppose λ, μ are positive finite measures. Define $\varphi := \mu + \lambda$. Let $f \geq 0$ be measurable, then by characteristic, simple functions and the definition of integral,

$$\int_X f \,\mathrm{d}\varphi = \int_X f \,\mathrm{d}\mu + \int_X f \,\mathrm{d}\lambda.$$

For all $f \in L^2(\varphi)$, the Cauchy-Schwarz inequality asserts that

$$\begin{aligned}
\left| \int_X f \,\mathrm{d}\lambda \right| &\leq \int_X |f| \,\mathrm{d}\lambda \leq \int_X |f| \,\mathrm{d}\varphi \\
&\leq \left(\int_X |f|^2 \,\mathrm{d}\varphi \right)^{1/2} + \left(\int_X 1^2 \,\mathrm{d}\varphi \right)^{1/2} \\
&= \|f\|_{L^2(\varphi)} \cdot (\varphi(X))^{1/2} < \infty.
\end{aligned}$$

Hence, $\Lambda \colon L^2(\varphi) \to \mathbb{C}$, by $\Lambda(f) := \int_X f \,\mathrm{d}\lambda$ is a **bouned linear functional**.

Step 2: By the **Riesz Representation Theorem on Hilbert space** (4.2.7), there is a unique $g \in L^2(\varphi)$ such that

$$\Lambda(f) = \int_X f \,\mathrm{d}\lambda = \langle f, g \rangle = \int_X f\overline{g} \,\mathrm{d}\varphi, \quad \forall f \in L^2(\varphi). \tag{6.2.2}$$

Note that g is unique as a point function on X up to φ-a.e.

Step 3: Let $f := \chi_E$ for $E \in M$, with $\varphi(E) > 0$. Then,

$$\lambda(E) = \int_E \overline{g}\,\mathrm{d}\varphi \geq 0.$$

Dividing both sides by $\varphi(E)$, we have

$$0 \leq \frac{1}{\varphi(E)} \int_E \overline{g}\,\mathrm{d}\varphi = \frac{\lambda(E)}{\varphi(E)} \leq 1.$$

Therefore, by the average argument Propoistion (1.8.8), $0 \leq \overline{g} \leq 1$, φ-a.e. We may assume $\overline{g}(x) \in [0,1]$, for all $x \in X$. Thus, $g = \overline{g}$ ranges in $[0,1]$, and we conclude that

$$\begin{aligned} \int_X f\,\mathrm{d}\lambda = \int_X fg\,\mathrm{d}\varphi &= \int_X f\,\mathrm{d}\lambda + \int_X f\,\mathrm{d}\mu \\ \int_X f(1-g)\,\mathrm{d}\lambda &= \int_X fg\,\mathrm{d}\mu \end{aligned} \tag{6.2.3}$$

Step 4: Define $A := \{x : g(x) < 1\}$, $B := \{x : g(x) = 1\}$. Define the measures

$$\lambda_a(E) := \lambda(E \cap A) \quad \text{and} \quad \lambda_s := \lambda(E \cap B).$$

Consider $f := \chi_B$, then equation (6.2.3) gives

$$\mu(B) = \int_B \mathrm{d}\mu = \int_X \chi_B g\,\mathrm{d}\mu = \int_X \chi_B(1-g)\,\mathrm{d}\lambda = 0.$$

Hence, $\mu(E) = \mu(E \cap A), \forall E \in M$, and $\lambda_s \perp \mu$ since A and B are disjoint.

Step 5: To see $\lambda_a \ll \mu$, define $f_n := \sum_{k=0}^{n} g^k$. By equation (6.2.3),

$$\int_X (1-g^{n+1})\,\mathrm{d}\lambda = \int_X \sum_{k=0}^{n} g^k(1-g)\,\mathrm{d}\lambda = \int_X \sum_{k=0}^{n} g^k \cdot g\,\mathrm{d}\mu. \tag{6.2.4}$$

For each $E \in M$, by definition of A and Dominated Convergence, the LHS of equation (6.2.4) yields

$$\begin{aligned} \lim_{n\to\infty} \int_E (1-g^{n+1})\,\mathrm{d}\lambda &= \lim_{n\to\infty} \int_{E\cap A} (1-g^{n+1})\,\mathrm{d}\lambda \\ &= \lambda(E \cap A) = \lambda_a(E). \end{aligned}$$

On the RHS of equation (6.2.4), by Monotone Convergence, we have

$$\begin{aligned} \lim_{n\to\infty} \int_E \sum_{k=0}^{n} g^k \cdot g\,\mathrm{d}\mu &= \int_E \sum_{k=0}^{\infty} g^k \cdot g\,\mathrm{d}\mu \\ &= \int_E \frac{1}{1-g} g\,\mathrm{d}\mu. \end{aligned}$$

Step 6: Hence,

$$\lambda_a(E) = \int_E \frac{g}{1-g}\,\mathrm{d}\mu, \quad \forall E \in M.$$

Now, define

$$h(x) := \begin{cases} \frac{g}{1-g}(x), & x \in A, \\ 0, & x \notin A. \end{cases}$$

Because $\lambda_a(A) = \int_A h\,\mathrm{d}\mu = \int_X |h|\,\mathrm{d}\mu < \infty$, (b) is proved. Also, $\mu(E) = 0$ gives $\int_E h\,\mathrm{d}\mu = \lambda_a(E) = 0$ which shows $\lambda_a \ll \mu$ and completes the proof of (a).

Step 7: Now we will generalize the proof. Let μ be σ-finite and λ be finite positive. By σ-finiteness, let $\{X_i\}_{i=1}^{\infty}$ be a paritition of X and $\mu(X_i) < \infty, \forall i$. Define $\lambda_i(E) := \lambda(E \cap X_i), \forall E \in M$, and $M_i := \{E \cap X_i : E \in M\}$.

Apply the previous result on each (X_i, M_i) with μ and λ_i. Then we obtain $(\lambda_{i_a}, \lambda_{i_s})$ on X_i with A_i, B_i, and $h_i \in L^1(\mu)$ such that $\forall E \in M_i$,

$$\lambda_{i_a}(E_i) = \int_{E_i} h_i\,\mathrm{d}\mu.$$

Define $\lambda_a(E) := \sum_{i=1}^{\infty} \lambda_{i_a}(E \cap X_i), \forall E \in M$, and likewise for λ_s with $A := \bigcup_{i=1}^{\infty} A_i$ and $B := \bigcup_{i=1}^{\infty} B_i$. Define $h(x) := h_i(x)$, where $x \in X_i$, which is well-defined because X_i's are disjoint. By Monotone Convergence, $\lambda_a(E) = \int_E h\,\mathrm{d}\mu$, and $\lambda(X) < \infty$ gives $h \in L^1(\mu)$.

Finally suppose μ is σ-finite and $\lambda = \lambda_1 + i\lambda_2$ is complex. For $k = 1, 2$, use Jordan decomposition on $\lambda_k = \lambda_k^+ - \lambda_k^-$ and apply the previous results. ■

Remark 6.2.5. If both μ and λ are σ-finite positive measure, using the techniques above, we can still obtain a function h which satisfies equation (6.2.1). However, in general $h \notin L^1(\mu)$, although $\int_{X_n} |h|\,\mathrm{d}\mu < \infty$ for each n. If we go beyond σ-finiteness, then both (a) and (b) fail. To see this, take μ to be the Lebesgue measure, and λ the counting measure on $(0, 1)$ and consider a singleton $\{x\}$.

Proposition 6.2.6 (Absolute Continuity). *Let μ be positive and λ be complex measures on (X, M). Then, the followings are equivalent:*

1. *$\lambda \ll \mu$.*
2. *Given $\varepsilon > 0$, there is $\delta > 0$, such that for every $E \in M$ with $\mu(E) < \delta$, $|\lambda(E)| < \varepsilon$.*

Proof. (2) $\Rightarrow$ (1). Given $E \in M$ with $\mu(E) = 0 < \delta, \forall \delta$. Hence, $|\lambda(E)| < \varepsilon, \forall \varepsilon > 0$ and $|\lambda(E)| = 0$, thus $\lambda \ll \mu$.

(1) $\Rightarrow$ (2). Proof by contrapositive. Suppose there is $\varepsilon > 0$ such that for each $\delta_n := 2^{-n} > 0$, there is $E_n \in M$ so that $\mu(E_n) < \delta_n$ but $|\lambda(E_n)| \geq \varepsilon$. Let $A_n := \bigcup_{i=n}^{\infty} E_i$, then

$$\mu(A_n) \leq \sum_{i=n}^{\infty} \mu(E_i) \leq \sum_{i=n}^{\infty} \frac{1}{2^i} = \frac{1}{2^{n-1}}.$$

Also, $\mu(A_1) < \infty$ and $A_1 \supseteq A_2 \supseteq \dots$, thus by monotonicity of μ,

$$\mu\bigg(\bigcap_{n=1}^{\infty} A_n\bigg) = \lim_{n\to\infty} \mu(A_n) = 0.$$

On the other hand, for each n, $|\lambda|(A_n) \geq |\lambda|(E_n) \geq |\lambda(E_n)| \geq \varepsilon$ gives

$$|\lambda|\bigg(\bigcap_{n=1}^{\infty} A_n\bigg) = \lim_{n\to\infty} |\lambda|(A_n) \geq \varepsilon > 0. \tag{6.2.5}$$

Then, $\lambda \ll |\lambda| \not\ll \mu$ concludes that $\lambda \not\ll \mu$. ■

Remark 6.2.7. Note that if λ is a positive unbounded measure, then (1) $\not\Rightarrow$ (2), as we use the boundedness in inequality (6.2.5).

6.3 Consequences of the Radon-Nikodym Theorem

Theorem 6.3.1 (Polar Decomposition/Representation Theorem). *Let μ be a complex measure on (X, M). Then is a measurable function h such that $|h(x)| = 1, \forall x \in X$, and $\mathrm{d}\mu = h\,\mathrm{d}|\mu|$.*

Proof. Observe that $|\mu|$ is finite and $\mu \ll |\mu|$. Hence, by Radon-Nikodym Theorem (6.2.3), there is $h \in L^1(\mu)$ such that $\mathrm{d}\mu = h\,\mathrm{d}|\mu|$. Consider the set $A_n := \{x : |h(x)| < 1 - \frac{1}{n}\}$, for each n. For every partition $\{E_k\}_{k=1}^{\infty}$ of A_n,

$$\begin{aligned}
\sum_{k=1}^{\infty} |\mu(E_k)| &= \sum_{k=1}^{\infty} \bigg|\int_{E_k} h\,\mathrm{d}|\mu|\bigg| \leq \sum_{k=1}^{\infty} \int_{E_k} (1 - \frac{1}{n})\,\mathrm{d}|\mu| \\
&= (1 - \frac{1}{n}) \sum_{k=1}^{\infty} \int_{E_k} \mathrm{d}|\mu| = (1 - \frac{1}{n}) \sum_{k=1}^{\infty} |\mu|(E_k) \\
&= (1 - \frac{1}{n})|\mu|(A_n).
\end{aligned}$$

By taking the supremum over all such partitions, we see that $|\mu|(A_n) \leq (1 - \frac{1}{n})|\mu|(A)$, which is only possible when $|\mu|(A_n) = 0$. Hence,

$$|\mu|(\{x : |h(x)| < 1\}) = |\mu|\bigg(\bigcup_{n=1}^{\infty} A_n\bigg) = \lim_{n\to\infty} |\mu|(A_n) = 0. \tag{6.3.1}$$

On the other hand, if $|\mu|(E) > 0$, then

$$1 \geq \frac{|\mu(E)|}{|\mu|(E)} = \frac{1}{|\mu|(E)} \int_E h\,\mathrm{d}|\mu| \geq 0.$$

By the average argument (1.8.8), $|h| \leq 1\,|\mu|$-a.e. Together with equation (6.3.1), we conclude that $|h| = 1$, $|\mu|$-a.e. Finally, we redefine $h(x) := 1$ on the set of $|\mu|$-measure zero and complete the proof. ■

Corollary 6.3.2. *Let μ be a positive measure on (X, M), $g \in L^1(\mu)$ and $\lambda(E) := \int_E g \, \mathrm{d}\mu$. Then,*

$$|\lambda|(E) = \int_E |g| \, \mathrm{d}\mu. \tag{6.3.2}$$

Proof. Since $g \in L^1(\mu)$, λ is a complex measure. By Polar decomposition (6.3.1), there is measurable h so that $|h| = 1$ and $\mathrm{d}\lambda = h \, \mathrm{d}|\lambda|$. By hypothesis,

$$\mathrm{d}\lambda = g \, \mathrm{d}\mu = h \, \mathrm{d}|\lambda|.$$

By viewing $h \, \mathrm{d}|\lambda|$ as a measure, we obtain $\mathrm{d}|\lambda| = \overline{h} g \, \mathrm{d}\mu$, hence $|\lambda| \ll \mu$. Then by Radon-Nikodym Theorem (6.2.3), there is $\varphi \in L^1(\mu)$ such that $\mathrm{d}|\lambda| = \varphi \, \mathrm{d}\mu$. Thus, $g \, \mathrm{d}\mu = h\varphi \, \mathrm{d}\mu$. By positivity of $|\lambda|$ and μ, $\varphi \geq 0$. By uniqueness,

$$\begin{aligned} g = h\varphi \quad &\Rightarrow \quad \overline{h} g = \varphi \geq 0, \quad \mu\text{-a.e.} \\ &\Rightarrow \quad \overline{h} g = |g|, \quad \mu\text{-a.e.} \end{aligned}$$

Therefore, $\mathrm{d}|\lambda| = \varphi \, \mathrm{d}\mu = |g| \, \mathrm{d}\mu$, and $|\lambda|(E) = \int_E |g| \, \mathrm{d}\mu$. ∎

Theorem 6.3.3 (Hahn Decomposition Theorem). *Let μ be a real measure on (X, M). Then there are $A, B \in M$, $A \cup B = X, A \cap B = \varnothing$, such that $\mu^+(E) = \mu(E \cap A)$ and $\mu^-(E) = -\mu(E \cap B), \forall E \in M$.*

Remark 6.3.4. Recall the Jordan decomposition: $\mu^+ = \frac{1}{2}(|\mu| + \mu)$, $\mu^- = \frac{1}{2}(|\mu| - \mu)$. The pair (A, B) is called a **Hahn decomposition** of X, induced by μ. Basically, X is split into two, where A contains the "positive mass" of μ, and B contains the "negative mass" of μ.

Proof. By Polar Decomposition (6.3.1), there is measurable h so that $|h| = 1$ and $\mathrm{d}\mu = h \, \mathrm{d}|\mu|$. Since μ is real, so is h and $h = \pm 1$ everywhere by redefining. Define $A := \{x : h(x) = 1\}$ and $B := \{x : h(x) = -1\}$. Note that

$$\frac{1}{2}(1 + h) = \begin{cases} h, & \text{on } A, \\ 0, & \text{on } B. \end{cases} \tag{6.3.3}$$

Hence, $\forall E \in M$,

$$\begin{aligned} \mu^+(E) &= \frac{1}{2}(|\mu|(E) + \mu(E)) = \frac{1}{2}\left(\int_E \mathrm{d}|\mu| + \int_E h \, \mathrm{d}|\mu| \right) \\ &= \frac{1}{2} \int_E (1 + h) \, \mathrm{d}|\mu| = \int_{E \cap A} h \, \mathrm{d}|\mu| \\ &= \mu(E \cap A). \end{aligned}$$

Since $\mu(E) = \mu(E \cap A) + \mu(E \cap A) = \mu^+(E) - \mu^-(E)$, $\mu^-(E) = -\mu(E \cap B)$. ∎

Corollary 6.3.5. *If $\mu = \lambda_1 - \lambda_2$, where λ_1 and λ_2 are positive measures, then $\mu^+ \leq \lambda_1$ and $\mu^- \leq \lambda_2$.*

Proof. If $\mu = \lambda_1 - \lambda_2$, then by positivity of λ_1 and λ_2, for all $E \in M$,

$$\begin{aligned}\mu^+(E) &= \mu(E \cap A) = \lambda_1(E \cap A) - \lambda_2(E \cap A)\\ &\leq \lambda_1(E \cap A) \leq \lambda_1(E).\end{aligned}$$

Similarly, we obtain $\mu^-(E) \leq \lambda_2(E)$, for all $E \in M$. ■

6.4 Bounded Linear Functionals on L^p

LEMMA 6.4.1. *If μ is a σ-finite positive measure on (X, M), then there exists $w \in L^1(\mu)$ such that $0 < w(x) < 1$, $\forall x \in X$.*

Proof. Let $\{X_n\}_{n=1}^{\infty}$ be a partition of X with $\mu(X_n) < \infty$, for each n. Define

$$w(x) := \sum_{n=1}^{\infty} \frac{1}{2^n} \cdot \frac{1}{1 + \mu(X_n)} \chi_{X_n}(x).$$

Then $0 < w(x) < 1$ as claimed. ■

REMARK 6.4.2. If μ is σ-finite, then $\tilde{\mu}$ given by $\mathrm{d}\tilde{\mu} := w \,\mathrm{d}\mu$ is finite. Moreover, because of the strictly positivity of w, $\tilde{\mu}$ has precisely the same sets of measure 0 as μ. Moreover, the map $f \mapsto \omega^{1/p} f$ is a linear isometry of $L^p(\hat{\mu})$ onto $L^p(\mu)$.

THEOREM 6.4.3 (L^p-Isometry). *Let $1 \leq p < \infty$, q be conjugate exponent, μ be σ-finite positive measure on (X, M). Then, for all bounded linear functional $\Lambda \in L^p(\mu)^*$, there is a unique $g \in L^q$, such that for each $f \in L^p(\mu)$,*

$$\Lambda(f) = \int_X f g \,\mathrm{d}\mu. \tag{6.4.1}$$

Moreover, $\|\Lambda\| = \|g\|_q$. Hence, $L^q(\mu)$ is isometrically isomorphic to $L^p(\mu)^*$.

Proof. **Step 1:** First suppose $\mu(X) < \infty$, $\Lambda \in L^p(\mu)^*$. Define $\lambda\colon M \to \mathbb{C}$ by $\lambda(E) := \Lambda(\chi_E)$. Let $\{E_n\}_{n=1}^{\infty}$ be a partition of $E \in M$. By linearity, for each $N \in \mathbb{N}$,

$$\lambda\left(\sum_{n=1}^{N} E_n\right) = \sum_{n=1}^{N} \lambda(E_n).$$

Also, $\|\chi_E - \chi_{\bigcup_{n=1}^N E_n}\|_p \to 0$, as $N \to \infty$. Hence, by continuity of Λ, we have $|\lambda(E) - \lambda(\bigcup_{n=1}^N E_n)| \to 0$, and λ is a complex measure. Moreover, if $\mu(E) = 0$, then $\chi_E = 0$ and $\lambda(E) = 0$. So, $\lambda \ll \mu$.

Step 2: By Radon-Nikodym Theorem (6.2.3), there is a unique $g \in L^1(\mu)$ such that $\mathrm{d}\lambda = g \,\mathrm{d}\mu$. Thererfore,

$$\Lambda(f) = \int_X f \,\mathrm{d}\lambda = \int_X f g \,\mathrm{d}\mu \tag{6.4.2}$$

holds for characteristic functions f, hence simple functions. Also, if $f \in L^\infty(\mu)$, there exists simple functions $s_n \to f$ uniformly. Hence, $\|s_n - f\|_p \to 0$ as $n \to \infty$, and equation (6.4.2) holds for all $f \in L^\infty(\mu)$. In order to complete the proof, we divide it into two cases.

Step 3: Case 1. $p = 1$. We will show every $f \in L^1(\mu)$ satisfies equation (6.4.2), and $\|g\|_\infty = \|\Lambda\|$. By assumption, $\mu(X) < \infty$, so $L^\infty(\mu)$ is dense in $L^1(\mu)$, by density of simple functions. Let $\{f_n\}$ be a sequence in L^∞ and $f_n \xrightarrow{L^1} f$. By continuity of Λ, the LHS of equation (6.4.2) gives $\Lambda(f_n) \to \Lambda(f)$.

Recall that Λ is bounded. For all $E \in M$ with $\mu(E) > 0$,

$$\Lambda(\chi_E) \leq \|\Lambda\| \cdot \|\chi_E\|_1 = \|\Lambda\| \cdot \mu(E).$$

Dividing both sides by $\mu(E)$, we see that

$$\left| \frac{1}{\mu(E)} \int_E g \,\mathrm{d}\mu \right| \leq \|\Lambda\|.$$

By the average argument (1.8.8), $|g(x)| \leq \|\Lambda\|$ μ-a.e., thus $\|g\|_\infty \leq \|\Lambda\|$. On the RHS of equation (6.4.2), by Hölder's inequality,

$$\left| \int_X |f_n - f| g \,\mathrm{d}\mu \right| \leq \|g\|_\infty \cdot \|f_n - f\|_1 \leq \|\Lambda\| \cdot \|f_n - f\|_1,$$

which $\to 0$, as $n \to \infty$. Hence, $\int_X f_n g \,\mathrm{d}\mu \to \int_X fg \,\mathrm{d}\mu$, and equation (6.4.2) holds for all $f \in L^1(\mu)$. Finally, by Hölder's inequality, for all $f \in L^1(\mu)$ with $\|f\|_1 \leq 1$,

$$\Lambda(f) = \int_X fg \,\mathrm{d}\mu \leq \|f\|_1 \cdot \|g\|_\infty \leq \|g\|_\infty.$$

Thus, $\|\Lambda\| \leq \|g\|_\infty$, and we conclude Case 1.

Step 4: Case 2. $1 < p < \infty$. Define $E_n := \{x : |g(x)| \leq n\}$. For each $n \in \mathbb{N}$, define

$$f_n := \frac{\overline{g}}{|g|} |g|^{q-1} \chi_{E_n}. \tag{6.4.3}$$

Note that $|f_n| = |g|^{q-1}$ and $f_n \in L^\infty(\mu)$. Moreover,

$$\begin{aligned} \int_X |f_n|^p \,\mathrm{d}\mu &= \int_{E_n} |g|^{p(q-1)} \,\mathrm{d}\mu \\ &= \int_{E_n} |g|^q \,\mathrm{d}\mu \leq n^q \mu(E_n) < \infty. \end{aligned}$$

So, $f_n \in L^\infty(\mu) \cap L^p(\mu)$. Restricting Λ on $L^\infty(\mu) \cap L^p(\mu)$, as a **Banach space with** $\|\cdot\|_p$, we see that

$$\left| \int_{E_n} |g|^q \,\mathrm{d}\mu \right| = |\Lambda(f_n)| \leq \|\Lambda\| \cdot \|f_n\|_p = \|\Lambda\| \left(\int_{E_n} |g|^q \,\mathrm{d}\mu \right)^{1/p}.$$

Therefore,

$$\|\Lambda\| \geq \left| \int_X |g\chi_{E_n}|^q \,\mathrm{d}\mu \right|^{1-1/p} = \left| \int_X |g\chi_{E_n}|^q \,\mathrm{d}\mu \right|^{1/q} = \|g\chi_{E_n}\|_q. \tag{6.4.4}$$

Note that $|g\chi_{E_1}| \leq |g\chi_{E_2}| < \dots$. By Monotone Convergence on equation (6.4.4),

$$\|g\|_q = \| \lim_{n\to\infty} |g\chi_{E_n}| \|_q = \lim_{n\to\infty} \| |g\chi_{E_n}| \|_q \leq \|\Lambda\|. \tag{6.4.5}$$

Hence, $g \in L^q(\mu)$ and by letting $f = 1$, μ-a.e., we obtain $\|g\|_q = \|\Lambda\|$.

Finally, by hypothesis Λ is continuous on $L^p(\mu)$. On the other hand, by Hölder's inequality, if $\|f_n - f\|_p \to 0$,

$$\int_X |f_n - f| g \,\mathrm{d}\mu \leq \|f_n - f\|_p \|g\|_q \to 0.$$

By sequential continuity, the map $f \mapsto \int_X fg \,\mathrm{d}\mu$ is continuous on $L^p(\mu)$. By equation (6.4.2), both continuous maps agree on the dense subset $L^\infty(\mu) \cap L^p(\mu)$ of $L^p(\mu)$. Therefore, they conincide everywhere on $L^p(\mu)$, and equation (6.4.2) holds for all $f \in L^p(\mu)$.

Step 5: Now suppose μ is σ-finite. By Lemma (6.4.1), define $\mathrm{d}\tilde{\mu} := w \,\mathrm{d}\mu$, and $\tilde{\mu}$ is a finite measure on X. Moreover, the map $\iota \colon L^p(\tilde{\mu}) \to L^p(\mu)$ given by

$$\iota(\tilde{f}) := w^{1/p} \tilde{f}, \tag{6.4.6}$$

is linearly isometric, and bijective since $w(x) \neq 0, \forall x \in X$. Consequently, the map $\Lambda \mapsto \Lambda \circ \iota$ defines an isomorphic isometry from $L^p(\mu)^*$ to $L^p(\tilde{\mu})^*$.

Step 6: Let $\Lambda \in L^p(\mu)^*$. Define $\tilde{\Lambda} := \Lambda \circ \iota \in L^p(\tilde{\mu})^*$. From the preceding steps, we obtain $\tilde{g} \in L^q(\tilde{\mu})$ so that $\|\tilde{g}\|_{q,\tilde{\mu}} = \|\tilde{\Lambda}\| = \|\Lambda\|$, and

$$\tilde{\Lambda}(\tilde{f}) = \int_X \tilde{f}\tilde{g} \,\mathrm{d}\tilde{\mu}, \quad \forall \tilde{f} \in L^p(\tilde{\mu}). \tag{6.4.7}$$

Define $g := \iota(\tilde{g}) = w^{1/q}\tilde{g}$. Then,

$$\begin{aligned} \int_X |g|^q \,\mathrm{d}\mu &= \int_X |w^{1/q}\tilde{g}|^q \,\mathrm{d}\mu = \int_X |\tilde{g}|^q w \,\mathrm{d}\mu \\ &= \int_X |\tilde{g}|^q \,\mathrm{d}\tilde{\mu} = \|\tilde{\Lambda}\|^q = \|\Lambda\|^q. \end{aligned}$$

Therefore, $\|g\|_q = \|\Lambda\|$, and for each $f \in L^p(\mu)$,

$$\begin{aligned} \Lambda(f) &= \tilde{\Lambda} \circ \iota^{-1}(f) = \tilde{\Lambda}(w^{-1/p} f) \\ &= \int_X (w^{-1/p} f)\tilde{g} \,\mathrm{d}\tilde{\mu} = \int_X (w^{-1/p} f)(w^{-1/q} g) \,\mathrm{d}\tilde{\mu} \\ &= \int_X fgw^{-1} \,\mathrm{d}\tilde{\mu} = \int_X fg \,\mathrm{d}\mu. \end{aligned}$$

■

6.5 The Riesz Representation Theorem

In this section, X denotes a locally compact Hausdorff space, $C_0(X)$ denotes the space of all complex continuous functions on X which vanish to infinity, and $C_c(X) \subset C_0(X)$ contains functions with compact support. By Riesz Representation Theorem (2.3.1), we have seen that every positive linear functional on $C_c(X)$ can be represented uniquely by a Borel measure on X. Now we will characterize all bounded linear functionals on $C_0(X)$ similarly.

Proposition 6.5.1. *The space $C_0(X)$ with the supremum norm $\|\cdot\|_\infty$ is a Banach space; and $C_c(X)$ is dense in $(C_0(X), \|\cdot\|_\infty)$.*

Proof. **Step 1:** Obviously, $C_0(X)$ is a normed vector space. For completeness, let $\{f_n\}_{n=1}^\infty$ be a Cauchy sequence in $C_0(X)$. Then, given $\varepsilon > 0$, there is $N \in \mathbb{N}$, such that for all $m, n > N$, for all $x \in X$,

$$|f_n(x) - f_m(x)| \le \|f_n - f_m\|_\infty < \varepsilon.$$

Thus, $\{f_n\}_{n=1}^\infty$ is **uniformly Cauchy** and converges to a continuous function f uniformly.[1]

Step 2: To see that f vanishes at infinity, let $\varepsilon > 0$, and choose N as above. Then, there is a compact K so that $|f_N| < \varepsilon$ on K^c. Then for all $x \in K^c$,

$$|f(x)| = |f(x) - f_N(x) + f_N(x)| \le |f(x) - f_N(x)| + |f_N(x)| \le 2\varepsilon.$$

Therefore, $f \in C_0(X)$ and $C_0(X)$ is Banach Space.

Step 3: Finally, suppose $f \in C_0(X)$. Then given $\varepsilon > 0$, there is a compact set K such that $|f| < \varepsilon$ on K^c. By the Urysohn's Lemman (2.2.11), choose $g \in C_c(X)$ such that $0 \le g \le 1$ and $g = 1$ on K. Then the function $fg \in C_c(X)$, and

$$|fg(x) - f(x)| \begin{cases} = 0, & x \in K, \\ \le \|f\|_\infty < \varepsilon, & x \in K^c. \end{cases}$$

Hence, $\|fg - f\|_\infty < \varepsilon$ and $C_c(X)$ is dense. ∎

[1] Since $\{f_n(x)\}_{n=1}^\infty$ is Cauchy at each $x \in X$, define f by $f_n(x) \to f(x)$, pointwise. Fix $x_0 \in X$, for every $\varepsilon > 0$, let $U := f_N^{-1}(B(f_N(x_0), \varepsilon))$. Then $x_0 \in U$, and for all $x \in U$,

$$\begin{aligned} |f(x) - f(x_0)| &= |f(x) - f_N(x) + f_N(x) - f_N(x_0) + f_N(x_0) - f(x_0)| \\ &\le |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| \\ &\le 3\varepsilon. \end{aligned}$$

So, $U \subset f^{-1}(B(f(x_0), 3\varepsilon))$. Now suppose V is open in $\mathbb{C}$. For each $x_0 \in f^{-1}(V)$, pick ε such that $B(f(x_0), 3\varepsilon) \subset V$. Then pick U as above, we obtain

$$x_0 \in U \subset f^{-1}(B(f(x_0), 3\varepsilon) \subset f^{-1}(V).$$

Thus, x_0 is an interior point of $f^{-1}(V)$. It follows that $f^{-1}(V)$ is open and f is continuous.

REMARK. Let S be a topological space, M be a complete metric space, and $f_n\colon S \to M$ be continuous. If $\{f_n\}$ is uniformly Cauchy, then $f_n \to f$ uniformly and f is continuous. The proof is the same.

DEFINITION 6.5.2. A complex Borel measure μ on X is regular if $|\mu|$ is **regular** on X. Denote $\mathcal{M}(X) := \{\mu : \text{regular complex Borel measure on } X\}$. Note that $\mathcal{M}(X)$ is a Banach space with the norm $\|\mu\| := |\mu|(X)$.

REMARK 6.5.3. Let μ be a complex Borel measure on X. By Polar decomposition (6.3.1), there is a complex Borel function h, with $|h| = 1$ so that $\mathrm{d}\mu = h\,\mathrm{d}|\mu|$. Thus, for all $f \in C_0(X)$,

$$\left|\int_X f\,\mathrm{d}\mu\right| = \left|\int_X fh\,\mathrm{d}|\mu|\right| \leq \|fh\|_\infty \int_X \mathrm{d}|\mu| = \|f\|_\infty |\mu|(X).$$

Therefore, $\Lambda_\mu(f) := \int_X f\,\mathrm{d}\mu$ defines a bounded linear functional on $C_0(X)$, and $\|\Lambda_\mu\| \leq |\mu|(X)$.

REMARK 6.5.4. Now suppose μ is complex *regular* Borel measure. Given a compact K, by the Urysohn's Lemma (2.2.11), there is $g_K \in C_c(X)$ such that $g_K = 1$ on K and $0 \leq g_K \leq 1$. Then,

$$\Lambda_\mu(\overline{h}g_K) = \int_X g_K\,\mathrm{d}|\mu| \geq \int_X \chi_K\,\mathrm{d}|\mu| = |\mu|(K).$$

By regularity and taking the supremum over all K, we see that

$$\|\Lambda_\mu\| \geq |\sup_{K\subset X} \Lambda_\mu(\overline{h}g_K)| \geq \sup_{K\subset X} |\mu|(K) = |\mu|(X).$$

Hence, $\|\Lambda_\mu\| = |\mu|(X)$. In other words, if we restrict to regular complex Borel measures, the map $\mu \mapsto \Lambda_\mu$ is an isometry.

QUESTION. Can every bounded linear functional on $C_0(X)$ be obtained this way, while preseving the norm? The answer is positive, and it is another version of the Riesz Representation Theorem. To prove it, we first introduce the following lemma which in fact is the technical part of the proof.

LEMMA 6.5.5. *Let $\lambda : C_c(X) \to \mathbb{R}$ be a **bounded linear functional**. Then there is a **positive linear functional** ρ on $C_c(X)$ such that $|\lambda(f)| \leq \rho(f) \leq \|f\|_\infty$.*

Proof. **Step 1:** First consider for $f \geq 0$, define $\rho(f) := \sup\{|\lambda(h)| : h \in C_c(X), |h| \leq f\}$. Observe that $\rho(f) \geq 0$, $|\Lambda(f)| \leq \rho(f) \leq \|f\|_\infty$, and ρ preserves scalar multiplication. Also, if $f_1 \geq f_2$, then $\rho(f_1) \geq \rho(f_2)$. We shall prove addivity.

Step 2: Suppose $f, g \in C_c(X), f, g \geq 0$. Given $\varepsilon > 0$, by definition of ρ, there is $h_1, h_2 \in C_c(X)$, $|h_1| \leq f, |h_2| \leq g$, such that

$$\rho(f) \leq |\lambda(h_1)| + \varepsilon \qquad \text{and} \qquad \rho(g) \leq |\lambda(h_2)| + \varepsilon.$$

Hence,

$$\rho(f) + \rho(g) \leq |\lambda(h_1)| + |\lambda(h_2)| + 2\varepsilon.$$

Moreover, for $i = 1, 2$ there is $\alpha_i \in \mathbb{C}$ with $|\alpha_i| = 1$ so that $\lambda(\alpha_i h_i) = |\lambda(h_i)|$. It follows that

$$\rho(f) + \rho(g) \leq \lambda(\alpha_1 h_1 + \alpha_2 h_2) + 2\varepsilon. \tag{6.5.1}$$

Since $|\alpha_1 h_1 + \alpha_2 h_2| \leq f + g$, inequality (6.5.1) gives

$$\begin{aligned} \rho(f) + \rho(g) &\leq |\lambda(\alpha_1 h_1 + \alpha_2 h_2)| + 2\varepsilon \\ &\leq \rho(|\alpha_1 h_1 + \alpha_2 h_2|) + 2\varepsilon \\ &\leq \rho(f + g) + 2\varepsilon, \end{aligned}$$

for all $\varepsilon > 0$. Therefore, $\rho(f) + \rho(g) \leq \rho(f + g)$.

Step 3: For the other inequality, pick $h \in C_c(X)$ such that $|h| \leq f + g$. Define the followings:

$$h_1(x) := \begin{cases} \frac{f(x)}{f(x)+g(x)} h(x), & f(x) + g(x) > 0, \\ 0, & \text{else.} \end{cases}$$

and

$$h_2(x) := h(x) - h_1(x).$$

Note that $|h_1| \leq f$ and $|h_2| \leq g$, and both are continuous everywhere. Consider

$$\begin{aligned} |\lambda(h)| &= |\lambda(h_1) + \lambda(h_2)| \\ &\leq |\lambda(h_1)| + |\lambda(h_2)| \\ &\leq \rho(f) + \rho(g). \end{aligned}$$

By taking the supremum over all such h, we obtain $\rho(f + g) \leq \rho(f) + \rho(g)$. Consequently, $\rho(f) + \rho(g) = \rho(f + g)$.

Step 4: Now suppose $f \in C_c(X)$, f is real-valued. Define $f^+ := \frac{1}{2}(|f| + f)$ and $f^- := \frac{1}{2}(|f| - f)$. Then $f = f^+ - f^-$ and $f^+, f^- \geq 0$. We define $\rho(f) := \rho(f^+) - \rho(f^-)$.

Given real-valued $f, g \in C_c(X)$, let $h := f + g$. Then,

$$\begin{gathered} f^+ + g^+ + h^- = f^- + g^- + h^+ \\ \rho(f^+) + \rho(g^+) + \rho(h^-) = \rho(f^-) + \rho(g^-) + \rho(h^+) \\ \rho(f) + \rho(g) = \rho(h) = \rho(f + g). \end{gathered}$$

Also, for $c \in \mathbb{R}$, it is easy to see that $\rho(cf) = c\rho(f)$. Finally, if f is complex-valued, write $f = u + iv$. Define $\rho(f) := \rho(u) + i\rho(v)$ and proceed similarly. ■

Theorem 6.5.6 (Riesz Representation Theorem of Bounded Linear Functionals). *Let X be a locally compact Hausdorff space. Then every* **bounded linear functional** *Φ on $(C_0(X), \|\cdot\|_\infty)$ is represented* ***uniquely*** *by a* ***regular complex Borel measure*** *μ such that*

$$\Phi(f) = \int_X f \,\mathrm{d}\mu, \quad \textit{for all } f \in C_0(X). \tag{6.5.2}$$

Moreover, $\|\Phi\| = |\mu|(X)$. In other words, $\mathcal{M}(X) \cong (C_0(X))^*$.

Proof of Uniqueness. Because both $\mathcal{M}(X)$ and $(C_0(X))^*$ are vector spaces and we proved surjectivity, it suffices to show that $\Phi \equiv 0$ implies $\mu \equiv 0$. Suppose $\Phi(f) = 0$, for all $f \in C_0(X)$, and μ is the corresponding measure.

By Polar decomposition (6.3.1), there is a complex Borel function h with $|h| = 1$, such that $\mathrm{d}\mu = h\,\mathrm{d}|\mu|$. Note that $h \in L^1(|\mu|)$. Since $C_c(X)$ is dense in $L^1(|\mu|)$, there is a sequence f_n in $C_c(X)$ such that $f_n \xrightarrow{L^1(|\mu|)} \overline{h}$. Therefore,

$$|\mu|(X) = \int_X 1\,\mathrm{d}|\mu| + 0 = \int_X (\overline{h} - f_n)h\,\mathrm{d}|\mu| \le \int_X |\overline{h} - f_n|\,\mathrm{d}|\mu|.$$

As $n \to \infty$, $|\mu|(X) = 0$. It follows that $|\mu| \equiv 0$ implies $\mu \equiv 0$. ■

Proof. **Step 1:** Let Φ be a bounded linear functional on $C_0(X)$. Without loss of generality, we assume $\|\Phi\| = 1$. By Lemma (6.6.5), there is a **positive linear functional** Λ on $C_c(X)$ such that

$$|\Phi(f)| \le \Lambda(|f|) \le \|f\|_\infty. \tag{6.5.3}$$

By **Riesz Representation Theorem** (2.3.1), there is a positive Borel measure λ such that

$$\Lambda(f) = \int_X f\,\mathrm{d}\lambda, \quad \text{for all } f \in C_c(X). \tag{6.5.4}$$

Step 2: Note that λ is outer regular. To show inner regularity, it suffices to show $\lambda(X) < \infty$. (Then every $E \in M$ has finite measure and automatically is inner regular by Riesz.) Since X is open,

$$\lambda(X) = \sup\{\lambda(K) : K \subset X, K \text{ is compact}\}.$$

By the Urysohn's Lemma (2.2.11), for every such K, there is $f \in C_c(X)$, with $0 \le f \le 1$ so that $\chi_K \le f \le \chi_X$. Conversely, every $f \in C_c(X)$ with $0 \le f \le 1$ is bounded above by χ_K, where $K = \operatorname{supp}(f)$. Thus,

$$\lambda(X) = \sup\{\Lambda(f) : 0 \le f \le 1, f \in C_c(X)\}.$$

Recall that Λ is bounded. Hence, for $\|f\|_\infty \le 1$, $|\Lambda(f)| \le \|\Lambda\| \cdot \|f\|_\infty \le 1$. Consequently, $\lambda(X) \le 1$ and λ is a positive regular Borel measure.

Step 3: From equation (6.5.4), note that

$$|\Phi(f)| = \Lambda(|f|) = \int_X |f|\,\mathrm{d}\lambda = \|f\|_1, \quad \text{for all } f \in C_c(X). \tag{6.5.5}$$

Now, consider $(C_c(X), \|\cdot\|_1)$ as a $L^1(\lambda)$-space. Then, Φ is a **bounded linear functional** on $L^1(\lambda)$. By $\boldsymbol{L^p}$**-Isometry** Theorem (6.4.3), there is a complex Borel measurable function $g \in L^\infty(\lambda)$, with $\|g\|_\infty = \|\Phi\| = 1$ such that

$$\Phi(f) = \int_X fg\,\mathrm{d}\lambda, \quad \text{for all } f \in C_c(X). \tag{6.5.6}$$

Step 4: Next, we will extend equation (6.5.6) to $C_0(X)$. Recall that Φ is continuous on $(C_0(X), \|\cdot\|_\infty)$. On the other hand, given $f \in C_c(X)$, and a sequence $f_n \xrightarrow{L^\infty} f$, by Hölder's inequality, we see that

$$\begin{aligned}\int_X |f_n - f| g \,\mathrm{d}\lambda &\leq \||f_n - f|g\|_\infty \int_X \mathrm{d}\lambda \\ &\leq \|f_n - f\|_\infty \|g\|_\infty \lambda(X) \\ &\leq \|f_n - f\| \to 0,\end{aligned}$$

as $n \to \infty$. Thus, by sequential continuity, the map $f \mapsto \int_X fg \,\mathrm{d}\lambda$ is also continuous. Since both continuous maps conincide on the dense subset $C_c(X)$ in $(C_0(X), \|\cdot\|_\infty)$, they agree everywhere. Define $\mathrm{d}\mu := g \,\mathrm{d}\lambda$. So, μ is a regular complex Borel measure, and we conclude that

$$\Phi(f) = \int_X fg \,\mathrm{d}\lambda = \int_X f \,\mathrm{d}\mu, \quad \text{for all } f \in C_0(X). \tag{6.5.7}$$

Step 5: Finally, we will show the isometry: $\|\Phi\| = |\mu|(X) = 1$. Since $\|\Phi\| = 1$, for all $f \in C_0(X)$ with $\|f\|_\infty \leq 1$, we see that

$$\begin{aligned}|\Phi(f)| = \left| \int_X fg \,\mathrm{d}\lambda \right| &= \int_X |fg| \,\mathrm{d}\lambda \\ &\leq \||f|\|_\infty \cdot \int_X |g| \,\mathrm{d}\lambda \leq \int_X |g| \,\mathrm{d}\lambda.\end{aligned}$$

Hence, $\int_X |g| \,\mathrm{d}\lambda \geq \sup\{|\Phi(f) : \|f\|_\infty \leq 1\} = \|\Phi\| = 1$. However, recall that $|g| \leq 1$ and $\lambda(X) \leq 1$. It is only possible when $\lambda(X) = 1$, and $g = 1$, λ-a.e. It follows that $g \in L^1(\lambda)$, and by Corollary (6.3.2), $\mathrm{d}|\mu| = |g| \,\mathrm{d}\lambda = \mathrm{d}\lambda$. Therefore,

$$|\mu|(X) = \lambda(X) = 1 = \|\Phi\|. \qquad \blacksquare$$

Chapter 7

Differentiation

In this chapter, $\mathbb{R}^k$ denotes the k-dimensional Euclidean space; m is the Lebesgue measure on $\mathbb{R}^k$; M denotes its complete Borel σ-algebra; μ is a complex Borel measure on $\mathbb{R}^k$. We write $B(x,r)$ for the open ball in $\mathbb{R}^k$ centered at $x \in \mathbb{R}^k$ with radius $r > 0$.

7.1 Derivatives of Measures

Definition 7.1.1. A function $f\colon \mathbb{R} \to \mathbb{C}$ is differentiable at $x_0 \in \mathbb{R}$ if there exists $A(x_0) \in \mathbb{C}$, such that given $\varepsilon > 0$, there is $\delta > 0$, so that

$$\left| \frac{f(b) - f(a)}{b - a} - A(x_0) \right| < \varepsilon, \tag{7.1.1}$$

whenever $|b - a| < \delta$, and for all $x \in (a, b)$. If such $A(x_0)$ exists, we denote it by $f'(x_0)$.

Remark 7.1.2. Note that $m((a, b)) = b - a$. Hence, if we define $f\colon \mathbb{R} \to \mathbb{C}$ by $f(x) := \mu((-\infty, x))$, then

$$\left| \frac{\mu((a, b))}{m((a, b))} - f'(x) \right| < \varepsilon, \tag{7.1.2}$$

provided that $f'(x)$ exists for $x \in \mathbb{R}$.

Remark 7.1.3. Observe that in $\mathbb{R}^k$, (a, b) can be replaced by $B(x, r)$, which is a Borel set. Thus, inequality (7.1.2) sugguests that we might want to define the "derivative of μ with respect to m" as the limit of the quotient $\frac{\mu(B(x,r))}{m(B(x,r))}$, as $r \to 0$. To do so, we now introduce some definitions.

Definition 7.1.4. The **symmetric derivative** of μ at x is defined to be

$$(D\mu)(x) := \lim_{r \to 0} \frac{\mu(B(x, r))}{m(B(x, r))}, \tag{7.1.3}$$

if it exists (in sense of $\mathbb{C}$). If it exists for every x, we simply denote it as $D\mu$.

REMARK. Just as we have a dominating positive measure $|\mu|$ on μ, we also want to introduce a dominating positive function on $D\mu$.

DEFINITION 7.1.5. The **maximal function** of μ is defined to be $M\mu\colon \mathbb{R}^k \to [0, \infty]$,

$$(M\mu)(x) := \sup_{r>0} \frac{|\mu|(B(x,r))}{m(B(x,r))}. \tag{7.1.4}$$

Note that $M\mu$ always exists since its range includes infinity.

PROPOSITION 7.1.6. *The maximal function $M\mu$ is* lower semicontinuous, *i.e. $(M\mu)^{-1}((\alpha,\infty))$ is open for all $\alpha > 0$. Hence, $M\mu$ is both μ and m-measurable.*

Proof. Given $\alpha > 0$ and $x \in E := (M\mu)^{-1}((\alpha,\infty))$, we will show that x is an interior point of E. By definition,

$$M\mu(x) = \sup_{r>0} \frac{|\mu|(B(x,r))}{m(B(x,r))} > \alpha.$$

Hence, there exists $r, t > 0$ such that

$$\frac{|\mu|(B(x,r))}{m(B(x,r))} = t > \alpha, \quad \text{and} \quad |\mu|(B(x,r)) = tm(B(x,r)).$$

Since $t > \alpha$, there is δ such that $(r+\delta)^k < r^k \frac{t}{\alpha}$. Consider the open ball $B(x,\delta)$. We will show that $B(x,\delta) \subset E$. In fact, for each $y \in B(x,r)$, $\triangle$-inequality gives $B(y, r+\delta) \supset B(x,r)$. Therefore, by the ratio of radii and translation invariance of the Lebesgue measure,

$$\begin{aligned} M\mu(y) &\geq \frac{|\mu|(B(y,r+\delta)}{m(B(y,r+\delta))} \geq \frac{|\mu|(B(x,r))}{m(B(y,r+\delta))} \\ &= \frac{tm(B(x,r))}{m(B(y,r+\delta))} = t \cdot \frac{r^k}{(r+\delta)^k} \\ &> \alpha. \end{aligned}$$

Hence, $M\mu(y) \in (\alpha, \infty)$, and $B(x,\delta) \subset E$. Consequently, E is open, and $M\mu$ is lower-semicontinuous. ■

LEMMA 7.1.7. *Let $W := \bigcup_{i=1}^N B(x_i, r_i) \subset \mathbb{R}^k$. Then there exists $S \subset \{1, \ldots, N\}$ such that*

(a) For $i \neq j \in S$, $B(x_i, r_i) \cap B(x_j, r_j) = \varnothing$.

(b) $W \subset \bigcup_{i \in S} B(x_i, 3r_i)$.

(c) $m(W) \leq 3^k m(\bigcup_{i\in S} B(x_i, r_i)) = 3^k \sum_{i\in S} m(B(x_i, r_i))$.

Proof. (a). Denote $B_i = B(x_i, r_i)$. Order $\{B_i\}$ such that $r_1 \geq r_2 \geq \cdots \geq r_N$. Choose $B_{i_1} := B_1$ with the largest radius $r_{i_1} = r_1$. Remove B_j from the collection $\{B_i\}$ if $B_j \cap B_{i_1} \neq \varnothing$. Add i_1 to S and reorder the remaining collection. Then, pick

B_{i_2} with the second largest radius r_{i_2} from the remaining collection and iterate. Since the collection is finite, after finitly many iterations, we will obtain (a).

(b). If B_j is removed in iteration n, then $B_{i_n} \cap B_j \neq \varnothing$, and $r_j < r_{i_n}$. Hence, for each $y \in B(x_j, r_j)$,

$$\begin{aligned} |y - x_{i_n}| &= |y - x_j + x_j - x_{i_n}| \\ &\leq |y - x_j| + |x_j - x_{i_n}| \\ &\leq r_{i_n} + 2r_{i_n} = 3r_{i_n}. \end{aligned}$$

So, $B(x_j, r_j) \subset B(x_{i_n}, 3r_{i_n})$ for $1 \leq j \leq N$. It follows that $W \subset \bigcup_{i \in S} B(x_i, 3r_i)$.

(c). By scaling (more precisely, by property (e) in Theorem (2.4.4) with the linear map $T(x) := 3x$, and $\triangle(T) = \det(T) = 3^k$), from (b) we see that

$$m(W) \leq m\left(\bigcup_{i \in S} B(x_i, 3r_i)\right) = \sum_{i \in S} m(B(x_i, 3r_i)) \leq 3^k \sum_{i \in S} m(B(x_i, r_i)). \qquad \blacksquare$$

Proposition 7.1.8. *For all $\lambda > 0$, $m(\{x : M\mu(x) > \lambda\}) \leq 3^k \lambda^{-1} |\mu|(\mathbb{R}^k)$.*

Remark. Recall that $|\mu|(X) < \infty$. Thus, as $\lambda \to 0$, $m(\{x : M\mu > \lambda\}) \to 0$. In other words, the maximal function cannot be large on a large set, in sense of the Lebesgue measure.

Proof. Given $\lambda > 0$, define $E := \{x : M\mu(x) > \lambda\}$. Since $M\mu$ is lower semi-continuous, E is open. Suppose $K \subset E$ is compact. For each $x \in K$, by definition of $M\mu$, there is $r_x > 0$ such that

$$|\mu|(B(x, r_x)) > \lambda m(B(x, r_x)).$$

Hence, $\{B(x, r_x) : x \in K\}$ is an open cover of K. By compactness, $K \subset \bigcup_{i=1}^N B(x_i, r_i)$, for some $N \in \mathbb{N}$. By Lemma (7.1.7), there is a finite $S \subset \{1, \ldots, N\}$ such that

$$\begin{aligned} m(K) &\leq 3^k \sum_{i \in S} m(B(x_i, r_i)) \\ &\leq 3^k \frac{1}{\lambda} \sum_{i \in S} |\mu|(B(x_i, r_i)) \\ &\leq 3^k \frac{1}{\lambda} |\mu|(\mathbb{R}^k). \end{aligned} \tag{7.1.5}$$

Since inequality (7.1.5) holds for all compact $K \subset E$, by inner regularity of m, $m(E) \leq 3^k \lambda^{-1} |\mu|(\mathbb{R}^k)$. $\blacksquare$

Definition 7.1.9 (Weak L^1)**.** Let $f\colon \mathbb{R}^k \to \mathbb{C}$ be m-measurable. We say $f \in$ **weak $\boldsymbol{L^1}$** if there is $M > 0$, such that for all $\lambda \in (0, \infty)$,

$$\lambda \cdot m(\{|f| > \lambda\}) \leq M. \tag{7.1.6}$$

Remark 7.1.10. Every $f \in L^1(\mathbb{R}^k)$ is in weak L^1 because

$$\|f\|_1 \geq \int_{\{|f|>\lambda\}} |f| \, \mathrm{d}m > \lambda \cdot m(\{|f| > \lambda\}).$$

Certainly weak L^1 is strictly large than L^1. Consider the function $f(x) := 1/x$ on $(0,1)$. Then $f \notin L^1$, but

$$\lambda \cdot m(\{1/x > \lambda\}) = \lambda \cdot m((0, 1/\lambda)) = 1, \quad \forall \lambda > 0.$$

Definition 7.1.11. Let $f : \mathbb{R}^k \to \mathbb{C}$. Define the **maximal function** of f to be

$$(Mf)(x) := \sup_{r>0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |f| \, \mathrm{d}m. \tag{7.1.7}$$

Remark 7.1.12. Let $f \in L^1(\mathbb{R}^k)$ and define the a complex Borel measure μ by $\mathrm{d}\mu := f \, \mathrm{d}m$. Then, Mf is exactly $M\mu$, and by Proposition (7.1.8)

$$\lambda m(\{Mf > \lambda\}) \leq 3^k |\mu|(\mathbb{R}^k) = 3^k \|f\|_1, \tag{7.1.8}$$

for all $\lambda > 0$. This is a special case of **Hardy-Littlewood maximal inequality**. Moreover, the operator $\boldsymbol{M}$ **sends** $\boldsymbol{L^1}$ **to weak** $\boldsymbol{L^1}$ with a bound 3^k.

7.2 Lebesgue Points

Definition 7.2.1 (Lebesgue points). If $f \in L^1(\mathbb{R}^k)$, we say $x_0 \in \mathbb{R}^k$ is a **Lebesgue point** of f if

$$\lim_{r\to 0} \frac{1}{m(B(x_0,r))} \int_{B(x_0,r)} |f(x) - f(x_0)| \, \mathrm{d}m(x) = 0. \tag{7.2.1}$$

Remark 7.2.2. If x_0 is a Lebesgue point of f, then

$$\begin{aligned}
\lim_{r\to 0} \frac{1}{m(B(x_0,r))} \left| \int_{B(x_0,r)} f(x) - f(x_0) \, \mathrm{d}m(x) \right| &= 0 \\
\left| \lim_{r\to 0} \frac{1}{m(B(x_0,r))} \int_{B(x_0,r)} f(x) \, \mathrm{d}m(x) - f(x_0) \right| &= 0 \\
\lim_{r\to 0} \frac{1}{m(B(x_0,r))} \int_{B(x_0,r)} f(x) \, \mathrm{d}m(x) &= f(x_0).
\end{aligned} \tag{7.2.2}$$

In general, equation (7.2.1) asserts that the averages of $|f - f(x)|$ are small on small open balls at x. Thus, the Lebesgue points of f are the points where f does not oscillate too much in average. Also, if f is continuous at x, then x is a Lebesgue point of f because $|f(y) - f(x)| < \varepsilon \to 0$, as $|y - x| < r \to 0$.

Theorem 7.2.3 (Lebesgue Differentiation Theorem). *If $f \in L^1(\mathbb{R}^k)$, then m-almost every $x \in \mathbb{R}^k$ is a Lebesgue point of f.*

Proof. **Step 1:** For each $r > 0$, define $T_r f : \mathbb{R}^k \to [0, \infty]$, by

$$(T_r f)(x) := \frac{1}{m(B(x,r))} \int_{B(x,r)} |f - f(x)| \, \mathrm{d}m. \tag{7.2.3}$$

Also, define $Tf : \mathbb{R}^k \to [0, \infty]$, by

$$(Tf)(x) := \limsup_{r \to 0} (T_r f)(x). \tag{7.2.4}$$

We want to show $(Tf)(x) = 0$, for m-almost every $x \in \mathbb{R}^k$.

Step 2: Fix an $n \in \mathbb{N}$. Since $C_c(\mathbb{R}^k)$ is dense in $L^1(\mathbb{R}^k)$, we can pick $g \in C_c(\mathbb{R}^k)$ such that $\|f - g\|_1 < 1/n$. Moreover, by continuity of g, $(Tg)(x) = 0$, for all $x \in \mathbb{R}^k$.

Step 3: Let $h := f - g \in L^1(\mathbb{R}^k)$. Then for all $x \in \mathbb{R}^k$,

$$\begin{aligned}
(T_r h)(x) &= \frac{1}{m(B(x,r))} \int_{B(x,r)} |(f-g) - (f(x) - g(x))| \, \mathrm{d}m \\
&\leq \left(\frac{1}{m(B(x,r))} \int_{B(x,r)} |(f-g) \, \mathrm{d}m \right) + |f(x) - g(x)| \\
&\leq \left(\frac{1}{m(B(x,r))} \int_{B(x,r)} |h| \, \mathrm{d}m \right) + |h(x)|.
\end{aligned} \tag{7.2.5}$$

Step 4: Also, for all $x \in \mathbb{R}^k$,

$$\begin{aligned}
(T_r f)(x) = (T_r(g+h)(x) &= \frac{1}{m(B(x,r))} \int_{B(x,r)} |g + h - (g(x) + h(x))| \, \mathrm{d}m \\
&\leq \frac{1}{m(B(x,r))} \left(\int_{B(x,r)} |g - g(x)| \, \mathrm{d}m + \int_{B(x,r)} |h - h(x)| \, \mathrm{d}m \right) \\
&\leq (T_r g)(x) + (T_r h)(x).
\end{aligned} \tag{7.2.6}$$

Step 5: Take the $\limsup_{r \to 0}$ on inequality (7.2.6) and apply inequality (7.2.5):

$$\begin{aligned}
(Tf)(x) &\leq (Tg)(x) + \limsup_{r \to 0} (T_r h)(x) \\
&\leq 0 + \sup_{r > 0} (T_r h)(x) \\
&\leq \sup_{r > 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} |h| \, \mathrm{d}m + |h(x)| \\
&= (Mh)(x) + |h(x)|.
\end{aligned} \tag{7.2.7}$$

Step 6: For all $y > 0$ such that $(Tf)(x) > 2y$, we have $(Mh)(x) > y$ or $|h(x)| > y$. Thus,

$$\{x : (Tf)(x) > 2y\} \subset \{x : (Mh)(x) > y\} \cup \{x : |h(x)| > y\}. \tag{7.2.8}$$

By inequality (7.1.8), and $ym(\{|h| > y\}) \leq \|h\|_1$, we see that

$$\begin{aligned} m(\{Tf > 2y\}) &\leq m(\{Mh > y\}) + m(\{|h| > y\}) \\ &\leq 3^k \frac{1}{y}\|h\|_1 + \frac{1}{y}\|h\|_1 \\ &\leq (3^k + 1)\frac{1}{ny}, \end{aligned} \tag{7.2.9}$$

where the last part is given by $\|h\|_1 \leq 1/n$.

Step 7: Note that $m(\{Tf > 2y\})$ is independent of n. (Although we have done a lot of approximations starting with n, the Lebesgue measure is fixed once we pick y.) Since inequality (7.2.9) holds for all $n \in \mathbb{N}$, as $n \to \infty$, we obtain $m(\{Tf > 2y\}) = 0$. By completeness of m, $\{Tf > 2y\}$ is m-measurable, for all $y > 0$. By monotonicity,

$$m(\{Tf > 0\}) = \lim_{N\to\infty} m(\{Tf > 1/N\}) = 0.$$

Therefore, the set for which $Tf \neq 0$ has Lebesgue measure 0; equivalently, m-almost every $x \in \mathbb{R}^k$ is a Lebesgue point of f. ■

Corollary 7.2.4. *If $\mu \ll m$ and $f := \frac{\mathrm{d}\mu}{\mathrm{d}m}$ is the Radon-Nikodym derivative of μ with respect to m, then $f = D\mu$, m-a.e.*

Proof. If x is a Lebesgue point of f,

$$\begin{aligned} f(x) &= \lim_{r\to 0} \frac{1}{m(B(x,r))} \int_{B(x,r)} f \,\mathrm{d}m \\ &= \lim_{r\to 0} \frac{\mu(B(x,r))}{m(B(x,r))} = D\mu(x). \end{aligned}$$

By Lebesgue Differentiation Theorem (7.2.3), $D\mu(x)$ exists and $D\mu(x) = f(x)$ for m-a.e. all $x \in \mathbb{R}^k$. ■

Corollary 7.2.5. *If $f \in L^1(\mathbb{R})$, then for m-a.e. $x \in \mathbb{R}$,*

$$f(x) = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \int_x^{x+\varepsilon} f \,\mathrm{d}m. \tag{7.2.10}$$

Proof. We will show that

$$\frac{1}{\varepsilon} \int_x^{x+\varepsilon} |f - f(x)| \,\mathrm{d}m \to 0, \quad \text{as } \varepsilon \to 0.$$

In fact, by positivity of m and $|f - f(x)|$,

$$\begin{aligned} \frac{1}{\varepsilon} \int_x^{x+\varepsilon} |f - f(x)| \,\mathrm{d}m &\leq \frac{2}{2\varepsilon} \int_{x-\varepsilon}^{x+\varepsilon} |f - f(x)| \,\mathrm{d}m \\ &\leq \frac{2}{m(B(x,\varepsilon))} \int_{B(x,\varepsilon)} |f - f(x)| \,\mathrm{d}m. \end{aligned}$$

By Lebesgue Differentiation Theorem (7.2.3), almost every x is a Lebesgue point. Thus, as $\varepsilon \to 0$, RHS $\to 0$, and we obtain

$$\frac{1}{\varepsilon}\int_{x}^{x+\varepsilon} |f - f(x)| \,\mathrm{d}m \to 0. \qquad \blacksquare$$

Definition 7.2.6 (Nicely shrinking sets)**.** Let $x \in \mathbb{R}^k$ and $\{E_n\}$ be a sequence of Borel sets in $\mathbb{R}^k$. We say $\{E_n\}$ **shrinks to x nicely** if there is $\alpha > 0$ and a sequence of positive numbers $r_n \to 0$, such that for all $n \in \mathbb{N}$, $E_n \subseteq B(x, r_n)$, and $m(E_n) \geq \alpha m(B(x, r_n))$.

Remark 7.2.7. Note that E_n need not contain x itself. For example, $E_n := (0, \frac{1}{n})$ shrinks to 0 nicely in $\mathbb{R}$. The condition of α requires each E_n to occupy certain portion of the ball $B(x, r_n)$. To illustrate this, $E_n := (0, \frac{1}{n}) \times (0, \frac{1}{n^2})$ does not shrink nicely to $(0, 0)$ in $\mathbb{R}^2$.

Proposition 7.2.8. *Suppose $f \in L^1(\mathbb{R}^k)$ and $x \in \mathbb{R}^k$ is a Lebesgue point of f. If $\{E_n\}$ shrinks to x nicely, then*

$$f(x) = \lim_{n\to\infty} \frac{1}{m(E_n)} \int_{E_n} f \,\mathrm{d}m. \tag{7.2.11}$$

Hence, it holds for almost every $x \in \mathbb{R}^k$.

Proof. Let α and $\{r_n\}$ be the positive number and sequence that are associated to $\{E_n\}$. Hence, $E_n \subset B(x, r_n)$ and

$$\begin{aligned}\frac{1}{m(E_n)} \int_{E_n} |f - f(x)| \,\mathrm{d}m &\leq \frac{1}{m(E_n)} \int_{B(x,r_n)} |f - f(x)| \,\mathrm{d}m \\ &\leq \frac{1}{\alpha m(B(x, r_n)} \int_{B(x,r_n)} |f - f(x)| \,\mathrm{d}m.\end{aligned}$$

Since $r_n \to 0$, the RHS $\to 0$ by definition of Lebesgue point, and we obtain equation (7.2.11). $\blacksquare$

Theorem 7.2.9. *Let $f \in L^1(\mathbb{R})$ and for all $x \in \mathbb{R}$,*

$$F(x) := \int_{\infty}^{x} f \,\mathrm{d}m.$$

Then, $F'(x) = f(x)$, at every Lebesgue point x of f, hence $F' = f$, m-a.e.

Proof. Let x be a Lebesgue point of f. Suppose $r_n > 0$, for all $n \in \mathbb{N}$, and $r_n \to 0$. Then $E_n := [x, x + r_n]$ shrinks to x nicely. By Proposition (7.2.8),

$$\begin{aligned}F'_+(x) &= \lim_{n\to\infty} \frac{1}{r_n}(F(x + r_n) - F(r_n)) \\ &= \lim_{n\to\infty} \frac{1}{r_n} \int_{x}^{x+r_n} f \,\mathrm{d}m = f(x).\end{aligned}$$

Likewise, $S_n := [x, x - r_n]$ also shrinks to x nicely, and $F'_-(x) = f(x)$. Hence, $F' = f$, m-a.e. $\blacksquare$

7.3 The Fundamental Theorem of Calculus

QUESTION. Let $f\colon [a,b] \to \mathbb{C}$. Recall that if f is continuous on $[a,b]$ and differentiable everywhere on (a,b), then

$$f(x) - f(a) = \int_a^x f' \,\mathrm{d}m, \quad x \in [a,b]. \tag{7.3.1}$$

However, if f' no longer exists everywhere on (a,b), what other assumptions are necessary?

REMARK 7.3.1. It turns out it is not enough even with f continuous on $[a,b]$, f' defined m-a.e. on $[a,b]$, and $f' \in L^1([a,b])$. We shall see that in the following example.

EXAMPLE 7.3.2 (**The Cantor Function**). **Step 1:** We proceed the standard Cantor set construction on $[0,1]$. For $n = 0$, remove $E_{0,1} := (\frac{1}{3}, \frac{2}{3})$ from $[0,1]$. Let $f := \frac{1}{2}$ on $E_{0,1}$. We denote that $f(E_{0,1}) = \frac{1}{2}$. For $n = 1$, remove the middle third intervals of the remaing intervals $[0,1] \setminus E_{0,1}$. We have $E_{1,1} := (\frac{1}{9}, \frac{2}{9})$, $E_{1,2} := (\frac{7}{9}, \frac{8}{9})$. Let $f(E_{1,0}) := \frac{1}{2^2}, f(E_{1,2}) := \frac{3}{2^2}$.

Step 2: In general, for each n, there are 2^n many disjoint E_{n,k_n}'s, each with measure 3^{-n-1}, from the remaining disjoint intervals. Hence, $E_n := \bigcup_{k=1}^{2^n} E_{n,k_n}$ has measure $\frac{1}{3} \cdot (\frac{2}{3})^n$. Let $E := \bigcup_{n=0}^{\infty} E_n$. We see that

$$m(E) = \sum_{n=0}^{\infty} \frac{1}{3} \left(\frac{2}{3}\right)^n = 1.$$

Define $f\colon E \to [0,1]$, given by

$$f(x) := \frac{2k_n - 1}{2^{n+1}}, \quad x \in E_{n,k_n}. \tag{7.3.2}$$

Note that $f(E) := \{\frac{1}{2}, \frac{1}{4}, \frac{3}{4}, \frac{1}{8}, \dots\}$, which is dense in $[0,1]$. Moreover, since f is constant on each open set E_{n,k_n}, $f'|_{E_{n,k_n}} = 0$. Thus, $f' = 0$ on E.

Step 3: Now we want to extend f continuously from E to $[0,1]$. Let $f(0) := 0$, for all $x \in [0,1] \setminus E$, define $f(x) := \sup\{f(t) : t < x\}$. Note that f is increasing. To see that f is continuous, we first show that f is surjective.

For each $y \in [0,1]$, let $S := \{x \in E : f(x) \leq y\}$, and $t := \sup(S)$. By monotonicty, $f(t) \leq y$. Suppose $f(t) < y$, then by density of $f(E)$, there is E_{n,k_n} such that $f(t) < f(E_{n,k_n}) < f(y)$. Hence, $t \neq \sup(S)$, which is a contradiction; and f is surjective.

Step 4: For all $x \in [0,1]$, $\varepsilon > 0$, there are $E_{n,k_n} = (a,b)$ and $E_{n',k'_n} = (a',b')$, such that $b < x < a'$, and

$$f(x) - \varepsilon < f(E_{n,k_n}) < f(x) < f(E_{n',k'_n}) < f(x) + \varepsilon.$$

By monotonicity and surjectivity, $f((b,a')) \subset (f(x)-\varepsilon, f(x)+\varepsilon)$. Therefore, f is continuous from $[0,1]$ to $[0,1]$.

Step 5: Finally, recall that $m(E)=1$, and the Cantor set $C=[0,1]\setminus E$ has measure zero. Thus, $f'=0$, m-a.e. and $f' \in L^1([a,b])$. However,

$$f(1)-f(0)=1 \neq \int_0^1 f' \,\mathrm{d}m.$$

Answer. In order to obtain equation (7.3.1), we need to introduce a stronger condition than merely continuity on f: absolute continuity.

Definition 7.3.3 (Absolute continuity). A function $f\colon I := [a,b] \to \mathbb{C}$ is called **absolutely continuous** if for any $\varepsilon>0$, there is $\delta>0$, such that whenever $\{(\alpha_i,\beta_i)\}_{i=1}^n$ is a finite collection of disjoint intervals in I, with $\sum_{i=1}^n(\beta_i-\alpha_i)<\delta$, we have

$$\sum_{i=1}^n |f(\beta_i)-f(\alpha_i)| < \varepsilon.$$

Remark 7.3.4. Obviously, absolute continuity implies uniform continuity, hence continuity. Also, the space of absolutely continuous functions is a vector space.

Theorem 7.3.5. *Let $I=[a,b]$, $f: I \to \mathbb{R}$ be continuous and non-decreasing. Then, the following are equivalent:*

(i) f is absolutely continuous.

(ii) f maps sets of measure zero to sets of measure zero.

(iii) f is differentiable m-a.e. on I, $f' \in L^1(I)$, and $f(x)-f(a)=\int_a^x f'\,\mathrm{d}m$.

Proof. $(i) \Rightarrow (ii)$. Let M be the σ-algebra of Lebesgue measurable sets. Let $E \subset I$, $E \in M$ and $m(E)=0$. We will show that $m(f(E))=0$.

Step 1: Without loss of generality, supose $E \subset (a,b)$. Given $\varepsilon>0$, by absolute continuity there is $\delta>0$ such that there exists an open set $V \supset E$ with $m(V)<\delta$, by outer regularity.

Step 2: Since V is open, we may write $V=\bigcup_{i=1}^\infty(\alpha_i,\beta_i)$, where (α_i,β_i)'s are disjoint. (To see this, let $\{q_1,q_2,\dots\}$ be rationals in V, let $B(q_1,r_1)\subset V$ with r_1 maximum. Remove $B(q_1,r_1)$ and iterate.) Thus,

$$\sum_{i=1}^\infty(\beta_i-\alpha_i)=m(V)<\delta \quad\Rightarrow\quad \sum_{i=1}^\infty |f(\beta_i)-f(\alpha_i)| \le \varepsilon.$$

Step 3: Since m is a positive measure and $E\subset V$, we have

$$m(f(E)) \le m(f(V)) = \sum_{i=1}^\infty |f(\beta_i)-f(\alpha_i)| \le \varepsilon.$$

Let $\varepsilon \to 0$, then $m(f(E))=0$, and we conclude (ii).

$(ii) \Rightarrow (iii)$. Suppose f maps sets of measure zero to sets of measure zero.

Step 1: Define $g(x) := x + f(x), x \in I$. For all segment $(a, b) \subset I$,

$$m(g((a,b))) = m((a,b)) + m(f((a,b))) = (b-a) + m(f((a,b))).$$

If $m(E) = 0$, then $m(f(E)) = 0$ and E does not contain any segments. Hence, g also satisfies (ii).

Step 2: Since g is continuous, and strictly increasing on $[g(a), g(b)]$, g^{-1} is also continuous. Thus, $g\colon I \to [g(a), g(b)]$ is a homeomorphism. Consequently, g preserves all the topological properties. Then, E is a Borel set in $[g(a), g(b)]$ if and only if $g^{-1}(E)$ is a Borel set in I.

Step 3: Moreover, for all $E \subset I$, $E \in M$, by regularity $E = K \cup F$, where K is an F_δ-set and $m(F) = 0$ by Theorem (2.3.1)(c). Thus, $K = \bigcup_{i=1}^\infty C_i$, where C_i is closed, hence compact in I. Since g is a homeomorphism, $g(C_i)$ is compact and closed in $[g(a), g(b)]$. Also, recall that g satisfies (ii), so $m(g(F)) = 0$. Hence, $g(E)$ is also a union of an F_δ-set, which is measurable, and a set of measure zero. We conclude that $g(E) \in M$.

Step 4: Define a measure $\mu\colon M \to \mathbb{R}$ by $\mu(E) := m(g(E))$. Note that μ is well-defined because $g(E) \in M$. Since g is injective, disjoint sets in I are mapped to disjoint images. By σ-addivity of m, we see that μ is indeed a positive and bounded measure. Moreover, $\mu \ll m$ because satisfies (ii).

Step 5: By Radon-Nikodym Theorem (6.2.3), $\mathrm{d}\mu = h\,\mathrm{d}m$, for some $h \in L^1(m)$. If follows that

$$\begin{aligned} g(x) - g(a) &= m([g(a), g(x)]) = m(g([a,x])) \\ &= \mu([a,x]) = \int_a^x h\,\mathrm{d}m. \end{aligned}$$

Hence,

$$\begin{aligned} (x + f(x)) - (a + f(a)) &= \int_a^x h\,\mathrm{d}m \\ f(x) - f(a) &= \int_a^x (h-1)\,\mathrm{d}m. \end{aligned}$$

By Theorem (7.2.9), $f'(x) = h(x) - 1$ for m-a.e. x, and we conclude (iii).

$(iii) \Rightarrow (i)$. Let f be differentiable m-a.e. on I, $f' \in L^1(m)$ with $f(x) - f(a) = \int_a^x f'\,\mathrm{d}m$. Define a measure $\mu\colon M(I) \to \mathbb{R}$, by

$$\mu(E) := \int_E f'\,\mathrm{d}m. \tag{7.3.3}$$

It follows that $\mu \ll m$. By absolute continuity of **measures**, Proposition (6.2.6), given $\varepsilon > 0$, there is δ such that whenever $m(E) < \delta$, then $|\mu(E)| < \varepsilon$. Moreover, since it holds for all $E \in M$, we conclude that $|\mu|(E) \leq \varepsilon$, and thus $|\mu| \ll m$. (Because $|\mu|$ takes supremum of μ over all finite partitions, and m is positive.)

Finally, suppose $E := \bigcup_{i=1}^{n}(\alpha_i, \beta_i)$ is a finite union of disjoint intervals with $m(E) = \sum_{i=1}^{n}(\beta_i - \alpha_i) < \delta$. Then,

$$\begin{aligned}\sum_{i=1}^{n} |f(\beta_i) - f(\alpha_i)| &= \sum_{i=1}^{n} |\mu((\alpha_i, \beta_i))| \\ &\leq \sum_{i=1}^{n} |\mu|((\alpha_i, \beta_i)) \\ &\leq |\mu|(E) \leq \varepsilon.\end{aligned}$$

Therefore, f is absolutely continuous and we conclude (i). ■

Definition 7.3.6 (Total variation function)**.** Let $f : [a, b] \to \mathbb{R}$ be absolutely continuous. The **total variation function** $F\colon [a, b] \to [0, \infty)$ is defined by

$$F(x) := \sup_{\{t_i\}_{i=0}^{N}} \sum_{i=1}^{N} |f(t_i) - f(t_{i-1})|, \tag{7.3.4}$$

where $\{t_i\}_{i=0}^{N}$ is any finite partition of $[a, b]$ with $a = t_0 < t_1 < \cdots < t_N = b$.

Lemma 7.3.7. *If $f\colon [a, b] \to \mathbb{R}$ is absolutely continuous, then the total variation function F, $F + f$, and $F - f$ are absolutely continuous and non-decreasing.*

Proof. First, we will show they are non-decreasing. Suppose $x, y \in I$, with $x < y$, and $\{t_i\}_{i=1}^{n}$ is a partition of $[a, x]$. Then, the set $\{t_i\}_{i=1}^{n} \cup \{t_{n+1} := y\}$ is a partition of $[a, y]$. By definition of F, we have

$$F(y) \geq |f(y) - f(x)| + \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})|. \tag{7.3.5}$$

Since it holds for any partition, by taking the supremum on the RHS of inequality (7.3.5), we see that

$$F(y) \geq |f(y) - f(x)| + F(x). \tag{7.3.6}$$

By simple arithmetics,

$$\begin{aligned}F(y) &\geq F(x), \\ F(y) - f(y) &\geq F(x) - f(x), \\ F(y) + f(y) &\geq F(x) + f(x).\end{aligned}$$

Therefore, F, $F - f$, and $F + f$ are all non-decreasing.

Step 1: Now to show that F is abolutely continuous, first observe that by definition of F, for all $(\alpha, \beta) \subset I$,

$$F(\beta) - F(\alpha) = \sup_{\{t_i\}_{i=0}^{N}} \sum_{i=1}^{N} |f(t_i) - f(t_{i-1})|, \tag{7.3.7}$$

where $\{t_i\}_{i=0}^N$ is a finite partition of $[\alpha, \beta]$.

Step 2: Since f is absolutely continuous, for all $\varepsilon > 0$, there is $\delta > 0$ such that if $\{(\alpha_i, \beta_i)\}_{i=1}^n$ is a collection of disjoint intervals with $\sum_{i=1}^n (\beta_i - \alpha_i) < \delta$, we have $\sum_{i=1}^n |f(\beta_i - f(\alpha_i)| < \varepsilon$. For each i, let $\{t_{i,j}\}_{j=0}^{N_j}$ be an arbitrary finite partition of (α_i, β_i). Then,

$$\sum_{i=1}^N \left(\sum_{j=1}^{N_j} (t_{i,j} - t_{i,j-1}) \right) = \sum_{i=1}^n (\beta_i - \alpha_i) < \delta.$$

Note that $\{(t_{i,j}, t_{i,j+1}) : 0 \le i \le n, 0 \le j \le N_j\}$ is also a finite partition of disjoint intervals. Thus,

$$\sum_{i=1}^N \sum_{j=1}^{N_j} |f(t_{i,j}) - f(t_{i,j-1})| < \varepsilon. \tag{7.3.8}$$

Step 3: Again, since equation (7.3.8) holds for any finite partition $\{t_{i,j}\}$, taking the supremum over all $\{t_{i,j} : 0 \le j \le N_j\}$ gives

$$\sum_{i=1}^N F(\beta_i) - F(\alpha_i) \le \varepsilon. \tag{7.3.9}$$

Since F is non-decreasing, $F(\beta_i) - F(\alpha_i) = |F(\beta_i) - F(\alpha_i)|$, and we obtain absolute continuity for F simply by using $\frac{\varepsilon}{2}$. Therefore, $F + f$ and $F - f$ are also absolutely continuous. ■

Theorem 7.3.8 (The Fundamental Theorem of Calculus). *If $f : I \to \mathbb{C}$ is absolutely continuous, then f' exists m-a.e., $f' \in L^1(I)$ and for all $x \in I$,*

$$f(x) - f(a) = \int_{[a,x]} f' \,\mathrm{d}m. \tag{7.3.10}$$

Remark. Note that f need not be non-decreasing.

Proof. It suffices to prove for $f : I \to \mathbb{R}$, by taking $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$ as usual. Consider the total variation F of f. Define $f_1 := \frac{1}{2}(F + f)$ and $f_2 := \frac{1}{2}(F - f)$. By the previous Lemma, f_1 and f_2 are absolutley continuous and non-decreasing. Applying the previous results, we see that

$$f_1(x) - f_1(a) = \int_{[a,x]} f_1' \,\mathrm{d}m \qquad \text{and} \qquad f_2(x) - f_2(a) = \int_{[a,x]} f_2' \,\mathrm{d}m.$$

Hence,

$$f(x) - f(a) = \int_{[a,x]} f' \,\mathrm{d}m,$$

where $f' = f_1' - f_2'$. ■

Chapter 8

Product Spaces

8.1 Measurability on Cartesian Products

In this section, let (X, M) and (Y, N) be measurable spaces. We want to extend our results on integrability, measurability, as well as measures, from measure spaces to thier Cartesian products. Hence, it is essential to construct a σ-algebra on $X \times Y$ from a set-theoretic approach that is relevant to both (X, M) and (Y, N).

Definition 8.1.1 (Algebra)**.** An **algebra** $\mathcal{A} \subset \mathcal{P}(X)$ is a nonempty collection of subsets in X such that $\mathcal{A}$ is closed under finite unions and complement. Note that by the De'Morgan's Laws, $\mathcal{A}$ is closed under finite intersections and set differences. Obviously, if in addition $\mathcal{A}$ is closed under infinite unions, then $\mathcal{A}$ is a σ-algebra.

Definition 8.1.2 (Monotone class)**.** A **monotone class** $\mathcal{M} \subset \mathcal{P}(X)$ is a nonempty collection of subsets in X such that $\mathcal{M}$ is closed under countable increasing unions and countable decreasing intersections. That is, if for each $i \in \mathbb{N}$, $A_i \subset A_{i+1}, B_i \supset B_{i+1}$, and $A_i, B_i \in \mathcal{M}$, then $\bigcup_{i=1}^{\infty} A_i$ and $\bigcap_{i=1}^{\infty} B_i \in \mathcal{M}$.

Let $S \subset \mathcal{P}(X)$. The monotone class generated by S is the intersection of all monotone classes that contains S. It is defined similarly as in Proposition (1.2.1). Thus, it is the smallest monotone class that contains S.

Remark 8.1.3. Hence every σ-algebra is a monotone class.

Theorem 8.1.4 (Monotone Class Theorem)**.** *If $\mathcal{A}$ is an algebra of a set X, then the monotone class $\mathcal{M}$ generated by $\mathcal{A}$ is precisely the σ-algebra Σ generated by $\mathcal{A}$.*

Proof. Obviously, $\mathcal{M} \subseteq \Sigma$, we will show $\Sigma \subseteq \mathcal{M}$. By assumption, $\mathcal{A} \subset \mathcal{M}$. Observe that given disjoint sequence $\{E_i\}_{i=1}^{\infty}$ in $\mathcal{M}$, if we let $E_k' := \bigcup_{i=1}^{k} E_i$, then $\{E_k'\}_{k=1}^{\infty}$ is an increasing sequence, and let

$$E := \bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{\infty} E_i'.$$

If $\mathcal{M}$ is an algebra, then $E'_k \in M$ and by definition of monotone class, $E \in \mathcal{M}$. Hence, $\mathcal{M}$ is a σ-algebra of $\mathcal{A}$, and we obtain $\Sigma \subseteq M$. Claim: $\mathcal{M}$ **is an algebra.**

Step 1: For all $E \in \mathcal{M}$, define

$$\mathcal{M}(E) := \{F \in \mathcal{M} : E \setminus F, F \setminus E, E \cap F \in \mathcal{M}\}. \tag{8.1.1}$$

Note that $\varnothing, E \in \mathcal{M}$. Also, the properties listed above are symmetric. That is, $F \in \mathcal{M}(F)$ if and only if $E \in M(E)$.

Step 2: We will show $\mathcal{M}(E)$ is a monotone class. Suppose $F_1 \subseteq F_2 \subseteq \dots$ is an increasing sequence in $\mathcal{M}(E)$. Let $F := \bigcup_{i=1}^{\infty} F_i$. We need to show $E \setminus F$, $F \setminus E$, and $E \cap F \in \mathcal{M}(E)$. Observe that $(E \setminus F_i) \supseteq (E \setminus F_{i+1})$ are in $\mathcal{M}$, thus

$$E \setminus F = \bigcap_{i=1}^{\infty}(E \setminus F_i) \in \mathcal{M}.$$

Similarly, $F \setminus E$ and $E \cap F \in \mathcal{M}$ and it follows that $F \in \mathcal{M}(E)$. Likewise, for a decreasing sequence $F_1 \supseteq F_2 \supseteq \dots$ in $\mathcal{M}(E)$, $\bigcap_{i=1}^{\infty} F_i \in \mathcal{M}(E)$. Consequently, $\mathcal{M}(E)$ is a monotone class.

Step 3: Fix an $E \in \mathcal{A}$. For all $F \in A$, by definition of algebra, $F \setminus E$, $E \setminus F$, $F \cap E \in \mathcal{A} \subset \mathcal{M}$. Then, $F \in \mathcal{M}(E)$ and thus $\mathcal{A} \subset \mathcal{M}(E)$. Recall that $\mathcal{M}(E)$ is a monotone class. Therefore, $\mathcal{M} \subseteq \mathcal{M}(E)$, for every $E \in \mathcal{A}$.

Moreover, if $F \in \mathcal{M}$, then $F \in \mathcal{M}(E)$, for all $E \in \mathcal{A}$. From Step 1, we see that $E \in \mathcal{M}(F)$. Thus $\mathcal{A} \subseteq \mathcal{M}(F)$ and $\mathcal{M} \subseteq \mathcal{M}(F)$. Conclusion[1]: For all $E \in \mathcal{M}$, $\mathcal{M} \subseteq \mathcal{M}(E)$.

Step 4: Finally, given any $E, F \in \mathcal{M} \subseteq \mathcal{M}(E)$, by definition of $\mathcal{M}(E)$, $E \setminus F, F \setminus F, E \cap F \in \mathcal{M}$. Furthermore, since $X \in \mathcal{A} \subseteq \mathcal{M}$, $E^c \in \mathcal{M}$. Together with the finite unions property of monotone class, $\mathcal{M}$ is an algebra. ■

Definition 8.1.5. A **measurable rectangle** $E \in X \times Y$ is in the form $A \times B$, where $A \in M$, $B \in N$. An **elementary set** is a finite union of disjoint measurable rectangles. Denote $\mathcal{E}$ the collection of all elementary sets, and $M \otimes N$ the σ-algebra generated by $\mathcal{E}$. Note that $\mathcal{E}$ is an algebra.

Corollary 8.1.6. *The σ-algebra $M \otimes N$ is the smallest monotone class containing $\mathcal{E}$.*

Proof. Since $\mathcal{E}$ is an algebra, $M \otimes N$ is the monotone class generated by $\mathcal{E}$, by the Monotone Class Theorem. ■

Definition 8.1.7 (Cross section)**.** Let $E \subset X \times Y$. Define the x-section and y-section respectively by

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

[1] This step is purely set-theoretic and definition-based. Read and proceed carefully. The main idea here is to extend the fact that $\mathcal{M}(E)$ is a monotone class that contains $\mathcal{M}$ for all $E \in \mathcal{A}$ to that for all $E \in \mathcal{M}$, which leads to the argument in Step 4.

Proposition 8.1.8. *If $E \in M \otimes N$, then $E^y \in M$ and $E_x \in N$, for all $x \in X$ and $y \in Y$.*

Proof. Let $\Omega := \{E : E^y \in M, \forall y \in Y\}$. We will show that Ω is a σ-algebra that contains $M \otimes N$. First, we begin with rectangle. Suppose $E = A \times B$, where $A \in M$ and $B \in N$. Then,

$$E^y = \begin{cases} A & \text{if } y \in B, \\ \varnothing & y \notin B, \end{cases}$$

and $E^y \in M$. Next, we show that Ω is a σ-algebra. In fact,

(i) $X \times Y \in \Omega$ because it is a rectangle.

(ii) If $E_n \in \Omega$ for each n, then $(\bigcup_{n=1}^{\infty} E_n)^y = \bigcup_{n=1}^{\infty} E_n^y \in M$.

(iii) Given $E \in \Omega$, $(E^c)^y = \{x : (x, y) \notin E\} = \{x : (x, y) \in E\}^c = (E^y)^c \in M$.

Thus, Ω is a σ-algebra that contains all measurable recetangles. Hence, $\Omega \supset \mathcal{E}$, and $\Omega \supset M \otimes N$. The proof for E_x is the same. ∎

Proposition 8.1.9. *If $f : X \times Y \to \mathbb{C}$ is $M \otimes N$-measurable, then $f^y(x) := f(x, y)$ is M-measurable. Likewise, $f_x(y) := f(x, y)$ is N-measurable.*

Proof. Let V be open in $\mathbb{C}$, by measurability of f, $f^{-1}(V) \in M \otimes N$. Hence,

$$\begin{aligned}(f^{-1}(V))^y &= \{x : (x, y) \in f^{-1}(V)\} = \{x : f(x, y) \in V\} \\ &= \{x : (f^y)(x) \in V\} = (f^y)^{-1}(V) \in M,\end{aligned}$$

by Proposition (8.1.8). Thus, f^y is M-measurable; likewise for f_x. ∎

8.2 Product Measures

In this section, let (X, M, μ) and (Y, N, ν) be positive σ-finite measure spaces. After the construction of the σ-algebra $M \otimes N$ on $X \times Y$, we want to construct a natural measure on $M \otimes N$, again relevant to both μ and ν.

Theorem 8.2.1. *Suppose $Q \in M \otimes N$. Define $\varphi(x) := \nu(Q_x)$ and $\psi(y) := \mu(Q^y)$. Then, φ is M-measurable and ψ is N-measurable. Also,*

$$\int_X \nu(Q_x)\, \mathrm{d}\mu(x) = \int_Y \mu(Q^y)\, \mathrm{d}\nu(y). \tag{8.2.1}$$

Proof. First suppose Q is a measurable rectangle, $Q = A \times B$, for some $A \in M$, $B \in N$. Then,

$$\varphi(x) = \nu(Q_x) = \nu(B)\chi_A(x), \qquad \text{and} \qquad \psi(y) = \mu(Q^y) = \mu(A)\chi_B(y).$$

Both functions are respectively M and N-measurable. Also,

$$\int_X \nu(Q_x)\,\mathrm{d}\mu(x) = \nu(B)\mu(A) = \int_Y \mu(Q^y)\,\mathrm{d}\nu(y).$$

Let Ω be the collection of $Q \in M \otimes N$ for which the conclusions hold. Recall that $M \otimes N$ is the monotone class generated by $\mathcal{E}$. Thus, if we can show Ω is a monotone class that contains $\mathcal{E}$, then $M \otimes N = \Omega$.

We will first show $\mathcal{E} \subset \Omega$. Let $\{Q_i\}_{i=1}^N$ be a finite collection of disjoint measurable rectangles. Then, by addivity of ν,

$$\nu\bigg((\bigcup_{i=1}^N Q_i)_x\bigg) = \nu\bigg(\bigcup_{i=1}^N (Q_i)_x\bigg) = \sum_{i=1}^N \nu((Q_i)_x),$$

which is M-measurable. Similarly, we have N-measurability by

$$\mu\bigg((\bigcup_{i=1}^N Q_i)^y\bigg) = \sum_{i=1}^n \mu((Q_i)^y).$$

To check equation (8.2.1),

$$\begin{aligned}
\int_X \nu\bigg((\bigcup_{i=1}^N Q_i)_x\bigg)\,\mathrm{d}\mu(x) &= \sum_{i=1}^N \int_X \nu((Q_i)_x)\,\mathrm{d}\mu(x) \\
&= \sum_{i=1}^N \int_Y \mu((Q_i)^y)\,\mathrm{d}\nu(y) \\
&= \int_X \mu\bigg((\bigcup_{i=1}^N Q_i)^y\bigg)\,\mathrm{d}\nu(y).
\end{aligned}$$

Therefore, $\mathcal{E} \subset \Omega$. To see that Ω is a monotone class, let $\{Q_n\}_{n=1}^\infty$ be a sequence in Ω. We will show countable union and intersection properties as follows.

Step 1: Suppose $Q_n \subseteq Q_{n+1}$. We will show $Q := \bigcup_{n=1}^\infty Q_n \in \Omega$. For each $x \in X$, by countable addivitiy and positivity of ν, $\nu((Q_n)_x) \leq \nu((Q_{n+1})_x)$, and $\nu((Q_n)_x) \to \nu(Q_x)$ by monotonicty. By Monotone Convergence, $\varphi(x) = \nu((Q_x))$ is M-measurable. Likewise, $\psi(y) = \mu(Q^y)$ is N-measurable. Moreover,

$$\begin{aligned}
\lim_{n\to\infty} \int_X \nu((Q_n)_x)\,\mathrm{d}\mu(x) &= \int_X \nu(Q_x)\,\mathrm{d}\mu(x), \\
\lim_{n\to\infty} \int_Y \mu((Q_n)^y)\,\mathrm{d}\nu(y) &= \int_Y \mu(Q^y)\,\mathrm{d}\nu(y).
\end{aligned}$$

Recall that

$$\int_X \nu((Q_n)_x)\,\mathrm{d}\mu(x) = \int_Y \mu((Q_n)^y)\,\mathrm{d}\nu(y),$$

for all n. So, the limits are equal and $Q \in \Omega$.

Step 2: Suppose the Q_n's are disjoint. We will show $\bigcup_{n=1}^{\infty} Q_n \in \Omega$. From Step 1, it suffices to show that if $Q, Q' \in \Omega$ and $Q \cap Q' = \varnothing$, then $Q \cup Q' \in \Omega$. In fact, $\nu((Q \cup Q')_x) = \nu(Q_x) + \nu(Q'_x)$. So, $\varphi(x)$ is M-measurable; and $\psi(y)$ is N-measurable likewise. Moreover, the equality of integral follows simply by linearity. We conclude that $\bigcup_{n=1}^{\infty} Q_n \in \Omega$.

Step 3: Now suppose $Q_n \supseteq Q_{n+1}$. We will show $Q := \bigcap_{n=1}^{\infty} \in \Omega$. We first assume $Q_1 \subseteq A \times B$, for some measurable rectangle with $\mu(A), \nu(B) < \infty$. Similar to Step 1, $\nu((Q_n)_x) \to \nu(Q_x)$ by monotonicity of ν, and hence $\varphi(x)$ is M-measurable by Corollary 1.3.5. Likewise, $\psi(y)$ is N-measurable. Also, $\nu((Q_n)_x), \mu((Q_n)^y) \leq \mu(A)\nu(B) < \infty$. By Dominated Convergence,

$$\lim_{n\to\infty} \int_X \nu((Q_n)_x) \,\mathrm{d}\mu(x) = \int_X \nu(Q_x) \,\mathrm{d}\mu(x),$$
$$\lim_{n\to\infty} \int_Y \mu((Q_n)^y) \,\mathrm{d}\nu(y) = \int_Y \mu(Q^y) \,\mathrm{d}\nu(y).$$

Thus, the limits are equal.

Step 4: In general, by σ-finiteness, there are disjoint partitions $\{X_i\}_{i=1}^{\infty}$ and $\{Y_j\}_{j=1}^{\infty}$ of X and Y with $\mu(X_i), \nu(Y_j) < \infty$. For each Q_n, let $Q_{n,(i,j)} := Q_n \cap (X_i \times Y_j)$. By Step 3, $\bigcap_{n=1}^{\infty} Q_{n,(i,j)} \in \Omega$. Since $\{\bigcap_{n=1}^{\infty} Q_{n,(i,j)} : i, j \in \mathbb{N}\}$ is a disjoint sequence, from Step 2,

$$\begin{aligned}\bigcup_{i,j\in\mathbb{N}} \left(\bigcap_{n=1}^{\infty} Q_{n,(i,j)} \right) &= \bigcap_{n=1}^{\infty} \left(\bigcup_{i,j\in\mathbb{N}} Q_{n,(i,j)} \right) \\ &= \bigcap_{n=1}^{\infty} Q_n = Q \in \Omega.\end{aligned}$$

Therefore, $\Omega = M \otimes N$. ■

Definition 8.2.2. The **product measure** $\mu \times \nu \colon M \otimes N \to [0, \infty]$ is defined

$$(\mu \times \nu)(Q) := \int_X \nu(Q_x) \,\mathrm{d}\mu(x) = \int_Y \mu(Q^y) \,\mathrm{d}\nu(y). \tag{8.2.2}$$

Remark 8.2.3. Note that $\mu \times \nu$ is indeed a measure because σ-addivity follows from Step 2 in the proof of Theorem (8.2.1), using Monotone Convergence for series.

Remark 8.2.4. Note that $M \otimes N$ is not necessarily the completion of $\mu \times \nu$ even if M, N are the completions of μ, ν. To see that, pick $A := \varnothing$, $B \notin N$. Then $(\mu \times \nu)(A \times Y) = 0$, but $A \times B \notin M \otimes N$.

Theorem 8.2.5 (Fubini's Theorem). *Let* (X, M, μ) *and* (Y, N, ν) *be* σ**-finite** *measure spaces,* f *be complex* $M \otimes N$*-measurable.*

(a) If $f \geq 0$*, then*

$$\int_X \underbrace{\int_Y f_x \,\mathrm{d}\nu}_{M\text{-measurable}} \,\mathrm{d}\mu = \int_{X\times Y} f \,\mathrm{d}(\mu \times \nu) = \int_Y \underbrace{\int_X f^y \,\mathrm{d}\mu}_{N\text{-measurable}} \,\mathrm{d}\nu. \tag{8.2.3}$$

(b) If $f \in L^1(\mu \times \nu)$, then for μ-a.e. $x \in X$, $f_x \in L^1(\nu)$,

$$\int_Y |f_x| \, \mathrm{d}\nu < \infty,$$

and $\psi(x) := \int_Y f_x \, \mathrm{d}\nu \in L^1(\mu)$. Likewise for f^y and $\varphi(y) := \int_X f^y \, \mathrm{d}\mu$.

(c) If f is complex and

$$\int_X \left(\int_Y |f_x| \, \mathrm{d}\nu \right) \mathrm{d}\mu < \infty,$$

then $f \in L^1(\mu \times \nu)$ and (b) holds.

(d) For all $M \otimes N$-measurable $f \in L^1(\mu \times \nu)$,

$$\int_X \left(\int_Y f_x \, \mathrm{d}\nu \right) \mathrm{d}\mu = \int_{X \times Y} f \, \mathrm{d}(\mu \times \nu) = \int_Y \left(\int_X f^y \, \mathrm{d}\mu \right) \mathrm{d}\nu.$$

Remark 8.2.6. From (b) and (c), if f is complex $M \otimes N$-measurable and

$$\int_X \mathrm{d}\mu \int_Y |f(x,y)| \, \mathrm{d}\nu < \infty,$$

then

$$\int_X \mathrm{d}\mu \int_Y f(x,y) \, \mathrm{d}\nu = \int_Y \mathrm{d}\nu \int_X f(x,y) \, \mathrm{d}\mu < \infty.$$

Proof. (a). From Theorem (8.2.1), we see that equation (8.2.3) holds for characteristic functions, hence simple functions. Now choose a sequence of simple measurable simple functions $s_n \geq 0$, such that $s_n(x,y) \nearrow f(x,y)$ pointwise. Thus, $(s_n)_x$, and s_n are increasing sequences to f_x and f, respectively. By Monotone Convergence on (Y, N, ν) and $(X \times Y, M \otimes N, \mu \times \nu)$,

$$\lim_{n\to\infty} \int_Y (s_n)_x \, \mathrm{d}\nu = \int_Y f_x \, \mathrm{d}\nu,$$
$$\lim_{n\to\infty} \int_{X \times Y} s_n \, \mathrm{d}(\mu \times \nu) = \int_{X \times Y} f \, \mathrm{d}(\mu \times \nu).$$

Also, $\psi_n(x) = \int_Y (s_n)_x \, \mathrm{d}\nu$ is M-measurable. By Monotone Convergence on (X, M, μ), $\int_Y f_x \, \mathrm{d}\nu$ is M-measurable, and

$$\int_X \lim_{n\to\infty} \psi_n \, \mathrm{d}\mu = \lim_{n\to\infty} \int_X \psi_n \, \mathrm{d}\mu.$$

Hence,

$$\int_X \left(\int_Y f_x \, \mathrm{d}\nu \right) \mathrm{d}\mu = \int_X \left(\lim_{n\to\infty} \int_Y (s_n)_x \, \mathrm{d}\nu \right) \mathrm{d}\mu$$

$$\begin{aligned}
&= \lim_{n\to\infty} \int_X \Big(\int_Y (s_n)_x \,\mathrm{d}\nu\Big)\,\mathrm{d}\mu \\
&= \lim_{n\to\infty} \int_{X\times Y} s_n \,\mathrm{d}(\mu\times\nu) \\
&= \int_{X\times Y} f \,\mathrm{d}(\mu\times\nu).
\end{aligned}$$

Likewise, the proof for f^y is the same. ■

(b). Since $|f| \in L^1(\mu\times\nu)$, from (a) on $|f|$,

$$\begin{aligned}
\int_X \left| \int_Y f_x \,\mathrm{d}\nu \right| \mathrm{d}\mu &\le \int_X \int_Y |f_x| \,\mathrm{d}\nu \,\mathrm{d}\mu \\
&= \int_{X\times Y} |f| \,\mathrm{d}(\mu\times\nu) < \infty.
\end{aligned}$$

Therefore, $\psi \in L^1(\mu)$ and likewise for f^y and φ. ■

(c). Again from (a),

$$\infty > \int_X \left(\int_Y |f_x| \,\mathrm{d}\nu \right) \mathrm{d}\mu \ge \int_{X\times Y} |f| \,\mathrm{d}(\mu\times\nu).$$

Hence, $f \in L^1(\mu\times\nu)$. ■

(d). Write $f = (u^+ - u^-) + i(v^+ - v^-)$, then apply (a) on each function as usual. ■

We will see that the hypotheses in Fubini's Theorem are necessary with the following counterexamples.

Example 8.2.7 (L^1). Suppose $X, Y := \mathbb{N}$, and μ, ν are the counting measures. Consider the function $a\colon X\times Y \to \mathbb{R}$, by

$$a = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ -\frac{1}{2} & 1 & 0 & 0 & \cdots \\ -\frac{1}{4} & -\frac{1}{2} & 1 & 0 & \cdots \\ -\frac{1}{8} & -\frac{1}{4} & -\frac{1}{2} & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Note that $a \notin L^1(\mu\times\nu)$. Then, the row sum is given by

$$\int_X \int_Y a \,\mathrm{d}\nu \,\mathrm{d}\mu = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} a_{n,m} = 2.$$

On the other hand, the coloum sum is

$$\int_Y \int_X a \,\mathrm{d}\mu \,\mathrm{d}\nu = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{n,m} = 0.$$

Hence, the integrals are not equal.

EXAMPLE 8.2.8 (σ-finiteness). Let $X, Y := [0,1]$, μ be the Lebesgue measure, and ν be the counting measure. Note that ν is not σ-finite on Y. Define

$$f(x,y) := \begin{cases} 1, & x = y, \\ 0, & x \neq y. \end{cases} \tag{8.2.4}$$

Then, $f^y = 0$, μ-a.e., so

$$\int_Y \int_X f^y \, \mathrm{d}\mu \, \mathrm{d}\nu = \int_Y 0 \, \mathrm{d}\nu = 0.$$

On the other hand,

$$\int_X \int_Y f_x \, \mathrm{d}\nu \, \mathrm{d}\mu = \int_X 1 \, \mathrm{d}\mu = 1.$$

Thus, again the integrals are not equal.

EXAMPLE 8.2.9 ($M \otimes N$-measurability). Assume the continuum hypothesis. By Zermelo-Frankel, there is bijective $j\colon [0,1] \to W$, where W is well-ordered. Moreover, for all $x \in W$, $\{y : y < x\}$ is at most countable. Define $Q := \{(x,y) \in [0,1]^2 : j(x) < j(y)\}$. Define $f := \chi_Q$ and Consider

$$f_x(y) = \begin{cases} 1, & (x,y) \in Q, j(x) < j(y), \\ 0, & \text{else.} \end{cases}$$

Thus, f_x is N-measurable. Fix x, then $\{y : j(y) < j(x)\} = \{y : f_x(y) = 0\}$ is at most countable. Therefore, integrating with the Lebesgue measure gives

$$\int_{[0,1]} \int_{[0,1]} f_x \, \mathrm{d}m \, \mathrm{d}m = 1.$$

On the other hand, f^y is M-measurable, and $\{x : f^y(x) = 1\} = \{x : f(x,y) = 1\} = \{x : j(x) < j(y)\}$ is also at most countable. Hence,

$$\int_{[0,1]} \int_{[0,1]} f^y \, \mathrm{d}m \, \mathrm{d}m = 0.$$

Since m is finite and $f \geq 0$, f must not be $M \otimes N$-measurable.

8.3 Completion of Product Measures

As we have seen, in general $M \otimes N$ is not the completion of the σ-algebra underlying $\mu \times \nu$ even though M and N belong to, respectively, the completions of μ and ν. In this section, we will see that the completion of $m^r \times m^s$ is in fact m^k on $\mathbb{R}^r \times \mathbb{R}^s = \mathbb{R}^k$. Also, we will present an alternative statement of the Fubini's Theorem with complete measure spaces.

Proposition 8.3.1. *Let m^k be the Lebesgue measure on $\mathbb{R}^k = \mathbb{R}^r \times \mathbb{R}^s$. Then, m^k is the completion of $m^r \times m^s$.*

Proof. **Step 1:** Denote $\mathcal{B}^k$ the Borel algebra of $\mathbb{R}^k$, and M^k the completion of $\mathcal{B}^k$. First, we want to show $\mathcal{B}^k \subset M^r \otimes M^s \subset M^k$. Every k-cell E can be written as

$$E = \prod_{i=1}^{r} I_i \times \prod_{i=1}^{s} I_i,$$

where I_i is an finite interval, so $E \in M^r \otimes M^s$. Also, every open set in $\mathbb{R}^k$ is a countable union of the k-cells. Since $\mathcal{B}^k$ is generated by these open sets, $\mathcal{B}^k \subset M^r \otimes M^s$.

Step 2: Suppose $E \in M^r, F \in M^s$. By Theorem (2.4.4), there is F_σ-set A and G_δ-set B in M^r such that $A \subset E \subset B$ and $m^r(B \setminus A) = 0$. Note that $A \times \mathbb{R}^s$ and $B \times \mathbb{R}^s$ are F_σ and G_δ in M^k, respectively. Thus, $A \times \mathbb{R}^s \subset E \times \mathbb{R}^s \subset B \times \mathbb{R}^s$, and

$$m^k\left((B \times \mathbb{R}^s) \setminus (A \times \mathbb{R}^s)\right) = m^k((B \setminus A) \times \mathbb{R}^s) = 0.$$

We conclude that $E \times \mathbb{R}^s \in M^k$. Likewise, $\mathbb{R}^r \times F \in M^k$. Therefore, $E \times F = E \times \mathbb{R}^s \cap \mathbb{R}^r \times F \in M^k$; and $M^r \otimes M^s \subset M^k$.

Step 3: Let $Q \in M^r \otimes M^s$, we will show $m^k(Q) = (m^r \times m^s)(Q)$. Since $Q \in M^k$, there are Borel sets $P_1, P_2 \in \mathcal{B}^k$, such that $P_1 \subset Q \subset P_2$, and $m^k(P_2 \setminus P_1) = 0$. However, recall from Step 1 that m^k and $m^r \times m^s$ agree on k-cells, hence open sets, hence Borel sets. Thus,

$$(m^r \times m^s)(Q \setminus P_1) \leq (m^r \times m^s)(P_2 \setminus P_1) = m^k(P_1 \setminus P_2) = 0,$$

we conclude that $(m^r \times m^s)(Q) = (m^r \times m^s)(P_1) = m^k(P_1) = m^k(Q)$.[2] Therefore, M^k is the completion of $M^r \otimes M^s$ because m^k is a complete measure on M^k. ∎

Lemma 8.3.2. *Let (X, M^*, μ) be the completion of (X, M, μ). If f is M^*-measurable, then there is M-measurable g such that $f = g$ μ-a.e.*

Proof. Suppose f is M^*-measurable and $f \geq 0$. Let $\{s_n\}$ be a sequence of M^*-measurable non-negative simple functions such that $s_n \nearrow f$. Note that

$$f = \sum_{n=1}^{\infty}(s_{n+1} - s_n),$$

and $s_n \to f$. Since each $s_{n+1} - s_n$ is a finite linear combination of characteristics functions,

$$f = \sum_{i=1}^{\infty} c_i \chi_{E_i}, \tag{8.3.1}$$

[2] This is not hard to see using regularity.

for some $c_i > 0$, and $E_i \in M^*$. By definition of M^*, for each E_i, there are $A_i, B_i \in M$ such that $A_i \subseteq E_i \subseteq B_i$, and $\mu(B_i \setminus A_i) = 0$. Define

$$g := \sum_{i=1}^{\infty} c_i \chi_{A_i},$$

which is M-measurable. Also, $g \neq f$ on $(E_i \setminus A_i) \subset (B_i \setminus A_i)$. Hence, $g = f$, μ-a.e. The general case follows as usual. ■

Lemma 8.3.3. *Let (X, M, μ) and (Y, N, ν) be complete measure spaces. If h is $(M \times N)^*$-measurable and $h = 0$ $(\mu \times \nu)$-a.e., then for ν-almost all $y \in Y$, h^y is M-measurable and $h^y(x) = 0$ for μ-a.e. $x \in X$. Analogous statement can be made for h_x.*

Proof. Let $P := \{(x, y) : h(x, y) \neq 0\}$. Then $P \in (M \otimes N)^*$ and $(\mu \times \nu)(P) = 0$. By completeness, there is $Q \in M \otimes N$ so that $P \subset Q$ and $(\mu \times \nu)(Q) = 0$. Then by Theorem (8.2.1),

$$\int_X \mu(Q^y)\,\mathrm{d}\nu = 0,$$

and $\mu(Q^y) > 0$ on a set T with $\nu(T) = 0$. Hence, if $y \notin T$, then $\mu(Q^y) = 0$, and $P^y \in M$ by completeness. Therefore, for ν-almost all $y \in Y$, h^y is M-measurable and $h^y = 0$, μ-a.e.; likewise for h_x. ■

Proof. Let $P := \{(x, y) : h(x, y) \neq 0\}$. Then $P \in (M \times N)^*$ and $(\mu \times \nu)(P) = 0$. Let $Q \supset P$ with $Q \in M \times N$ and $(\mu \times \nu)(Q) = 0$. Use Fubini on Q^y, use completeness of μ to show $P^y \in M$. Definition of P^y and all subset of $P^y \in M$. ■

Theorem 8.3.4 (Fubini's Theorem with Completion)**.** *Let (X, M, μ) and (Y, N, ν) be complete σ-finite measure spaces. Let $(M \otimes N)^*$ be the completion of $M \otimes N$ with respect to $\mu \times \nu$. Suppose f is a $(M \otimes N)^*$-measurable function on $X \times Y$. Then, the conclusions of Fubini's Theorem* (8.2.5) *still hold,* **except** *that f_x is N-measurable for μ-a.e. $x \in X$ and f^y is M-measurable for ν-a.e. $y \in Y$.*

Proof. By Lemma (8.3.1), we can replace f by an $M \otimes N$-measurable function g such that $g = f$, $(\mu \times \nu)$-a.e. Applying Fubini's Theorem on g, together with Lemma (8.3.2), we see that for μ-almost all $x \in X$, and ν-almost all $y \in Y$,

$$f_x = g_x,\ \nu\text{-a.e.} \qquad \text{and} \qquad f^y = g^y,\ \mu\text{-a.e.}$$

Therefore,

$$\int_{X\times Y} f\,\mathrm{d}(\mu \times \nu) = \int_{X\times Y} g\,\mathrm{d}(\mu \times \nu)$$
$$\int_Y \int_X f^y\,\mathrm{d}\mu\,\mathrm{d}\nu = \int_Y \int_X g^y\,\mathrm{d}\mu\,\mathrm{d}\nu$$

$$\int_X \int_Y f_x \, \mathrm{d}\nu \, \mathrm{d}\mu = \int_X \int_Y g_x \, \mathrm{d}\nu \, \mathrm{d}\mu$$

Since the right-hand sides are all equal by Fubini's Theorem, so are the left-hand sides. ■

8.4 Convolutions

Theorem 8.4.1 (Convolution)**.** *Suppose $f, g \in L^1(\mathbb{R})$. Define F on $\mathbb{R}^2$, given by $F(x,y) := f(x-y)g(y)$. Then, for m-almost all $x \in \mathbb{R}$, $F_x \in L^1(\mathbb{R})$. Moreover, for these x's, the function*

$$h(x) := \int_{\mathbb{R}} f(x-y)g(y) \, \mathrm{d}y \tag{8.4.1}$$

is in $L^1(\mathbb{R})$, with $\|h\|_1 \leq \|f\|_1 \|g\|_1$.

Proof. Without loss of generality, we assume both f and g are Borel measurable functions. Thus, F is also m^2-measurable since $(x,y) \mapsto x-y$ and $y \mapsto y$ are both measurable maps. By Fubini's Theorem (8.2.5)(a),

$$\begin{aligned}
\int_{\mathbb{R}^2} |F| \, \mathrm{d}x \, \mathrm{d}y &= \int_{\mathbb{R}} \int_{\mathbb{R}} |F_x| \, \mathrm{d}x \, \mathrm{d}y = \int_{\mathbb{R}} \int_{\mathbb{R}} |f(x-y)| \, \mathrm{d}x |g(y)| \, \mathrm{d}y \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} |f(x)| \, \mathrm{d}x |g(y)| \, \mathrm{d}y = \int_{\mathbb{R}} \|f\|_1 |g(y)| \, \mathrm{d}y \\
&= \|f\|_1 \|g\|_1,
\end{aligned}$$

where the third equality is given by translation invariance of the Lebesgue measure. Hence, $F \in L^1(\mathbb{R})$. By Fubini's Theorem (8.2.5)(b), $F_x \in L^1(\mathbb{R})$ for m-almost all $x \in \mathbb{R}$. Moreover,

$$h(x) = \int_{\mathbb{R}} F_x \, \mathrm{d}y$$

is in $L^1(\mathbb{R})$ and

$$\int_{\mathbb{R}} |h(x)| \, \mathrm{d}x = \int_{\mathbb{R}} \left| \int_{\mathbb{R}} F_x \, \mathrm{d}y \right| \mathrm{d}x \leq \int_{\mathbb{R}} \int_{\mathbb{R}} |F_x| \, \mathrm{d}y \, \mathrm{d}x = \|f\|_1 \|g\|_1.$$ ■

Definition 8.4.2. We call h the **convolution** of f and g, and denote it by $f * g$.

Convolution of Measures

Let μ and λ be complex Borel measures on $\mathbb{R}$. Let $F \colon \mathbb{R}^2 \to \mathbb{R}$, by $F(x,y) = x + y$. Define the **convolution** of μ and λ by

$$(\mu * \lambda)(E) := (\mu \times \lambda)(F^{-1}(E)), \tag{8.4.2}$$

for all Borel set E in $\mathbb{R}$.

Theorem 8.4.3. *The convolution $\mu * \lambda$ defines a complex Borel measure on $\mathbb{R}$.*

Proof. Since F is continuous, $F^{-1}(E)$ is a Borel set in $\mathbb{R}^2$. Recall that Fubini's Theorem applies on positive σ-finite measures. We will first show that $\mu \times \lambda$ defines a complex Borel measure. It will follow that $\mu * \lambda$ is well-defined.

By polar decomposition (6.3.1), there are Borel measurable $h_1 \in L^1(\mu)$, $h_2 \in L^1(\lambda)$, such that

$$\mathrm{d}\mu = h_1 \,\mathrm{d}|\mu| \qquad \text{and} \qquad \mathrm{d}\lambda = h_2 \,\mathrm{d}|\lambda|,$$

with $|h_1| = 1$, μ-a.e., and $|h_2| = 1$, λ-a.e. Thus, for all Borel set B in $\mathbb{R}^2$,

$$\begin{aligned}(\mu \times \lambda)(B) = \int_{\mathbb{R}} \lambda(B_x) \,\mathrm{d}\mu &= \int_{\mathbb{R}} \left(\int_{B_x} h_2(y) \,\mathrm{d}|\lambda|(y) \right) h_1(x) \,\mathrm{d}|\mu|(x) \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \chi_{B_x} h_1(x) h_2(y) \,\mathrm{d}|\lambda|(y) \right) \mathrm{d}|\mu|(x).\end{aligned}$$

Also, because $|\chi_{B_x} h_1 h_2| \leq 1$, for μ-a.e. x and ν-a.e. y,

$$\begin{gathered}\int_{\mathbb{R}} |\chi_{B_x} h_1 h_2| \,\mathrm{d}|\lambda| \leq \|\lambda\|, \quad \text{and} \\ \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |\chi_{B_x} h_1 h_2| \,\mathrm{d}|\lambda| \right) \mathrm{d}|\mu| \leq \|\mu\| \|\lambda\| < \infty.\end{gathered}$$

By Fubini's Theorem (8.2.5)(c), $\chi_{B_x} h_1 h_2 \in L^1(\mu \times \lambda)$. Define $h(x, y) := h_1(x) h_2(y)$. Then, $h \in L^1(\mathbb{R}^2)$ is Borel measurable and $|h(x, y)| = 1$, $(|\mu| \times |\lambda|)$-a.e. More importantly,

$$(\mu \times \lambda)(B) = \int_{\mathbb{R}^2} \chi_B h \,\mathrm{d}(|\mu| \times |\lambda|) \tag{8.4.3}$$

shows that $\mu \times \lambda$ is indeed a complex Borel measure. [3]

Finally, if $\{E_n\}$ are disjoint Borel sets in $\mathbb{R}$, then $\{B_n := F^{-1}(E_n)\}$ are disjoint Borel sets in $\mathbb{R}^2$. Consequently, $\mu * \lambda$ satisfies σ-addivity and defines a complex Borel measure on $\mathbb{R}$. ■

Remark 8.4.4. Observe that $|\mu| * |\lambda| = |\mu * \lambda|$ by Polar decomposition. Also, all open sets are σ-compact because the collection $\{(a, b) : a, b \in \mathbb{Q}\}$ is a basis for $\mathbb{R}$. Since $|\mu|$, $|\lambda|$, and $|\mu * \lambda|$ are finite on all Borel sets, μ, λ, and $\mu * \lambda$ are regular complex Borel measures.

Theorem 8.4.5. *The convolution $\mu * \lambda$ is a unique complex Borel measure ν on $\mathbb{R}$ such that*

$$\int_{\mathbb{R}} f \,\mathrm{d}\nu = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x + y) \,\mathrm{d}\mu(x) \,\mathrm{d}\lambda(y), \tag{8.4.4}$$

for all $f \in C_0(\mathbb{R})$.

[3] By now, we should be able to prove that easily by: Theorem (1.6.8), Monotone Convergence with splitting, or Dominated Convergence.

Proof. Define $L\colon C_0(\mathbb{R}) \to \mathbb{C}$ by

$$L(f) := \int_{\mathbb{R}} \int_{\mathbb{R}} f(x+y)\,\mathrm{d}\mu(x)\,\mathrm{d}\lambda(y). \tag{8.4.5}$$

We will show that L is a bouned linear functional on $C_0(\mathbb{R})$. From equation (8.4.3), for all $f \in C_0(\mathbb{R})$, with $\|f\|_\infty \le 1$, by Hölder's inequality,

$$\begin{aligned}\left| \int_{\mathbb{R}} \int_{\mathbb{R}} f(x+y)h(x,y)\,\mathrm{d}|\mu|(x)\,\mathrm{d}|\lambda|(y) \right| &\le \int_{\mathbb{R}} \int_{\mathbb{R}} |f(x+y)|\,\mathrm{d}|\mu|(x)\,\mathrm{d}|\lambda|(y) \\ &\le \|\mu\|\|\lambda\| < \infty.\end{aligned}$$

Therefore, $\|L\| \le \|\mu\|\|\lambda\|$. By **Riesz Representation Theorem** of bounded linear functionals over C_0, Theorem (6.5.6), there is a unique complex Borel measure ν on $\mathbb{R}$ such that

$$L(f) = \int_{\mathbb{R}} f\,\mathrm{d}\nu. \tag{8.4.6}$$

We will show $\nu = \mu \times \lambda$. Suppose E is an open set in $\mathbb{R}$. By inner regularity, there is an increasing sequence of compact sets $\{K_n\}_{n=1}^\infty$ in E such that $\bigcup_{n=1}^\infty K_n = E$. Moreover, by the Urysohn's Lemma (2.2.11), there is $f_n \in C_c(\mathbb{R}) \subset C_0(\mathbb{R})$ such that $\chi_{K_n} \le f_n \le \chi_E$, $f_n \nearrow \chi_E$. Hence, equations (8.4.5) and (8.4.6) give

$$\begin{aligned}\int_{\mathbb{R}} f_n\,\mathrm{d}\nu &= \int_{\mathbb{R}} \int_{\mathbb{R}} f_n(x+y)h(x,y)\,\mathrm{d}|\mu|(x)\,\mathrm{d}|\lambda|(y) \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} (f_n \circ F)\,\mathrm{d}\mu(x)\,\mathrm{d}\lambda(y).\end{aligned} \tag{8.4.7}$$

By Monotone Convergence on both sides of equation (8.4.7),

$$\begin{aligned}\lim_{n\to\infty} \int_{\mathbb{R}} f_n\,\mathrm{d}\nu &= \lim_{n\to\infty} \int_{\mathbb{R}} \int_{\mathbb{R}} (f_n \circ F)\,\mathrm{d}\mu(x)\,\mathrm{d}\lambda(y) \\ \nu(E) &= \int_{\mathbb{R}} \int_{\mathbb{R}} (\chi_E \circ F)\,\mathrm{d}\mu(x)\,\mathrm{d}\lambda(y) \\ &= \int_{\mathbb{R}^2} \chi_{F^{-1}(E)}\,\mathrm{d}(\mu \times \lambda) \\ &= (\mu * \lambda)(E).\end{aligned}$$

Hence, $\nu = \mu * \lambda$ on all open sets. Finally, since $|\nu|$ and $|\mu * \lambda|$ are finite measures, by outer regularity, $\nu(E) = (\mu * \lambda)(E)$ on all measurable sets E. ■

Remark 8.4.6. The motivation is to see that $\|f * g\| \le \|f\|\|g\|$ and $\|\mu * \lambda\| \le \|\mu\|\|\lambda\|$ in $L^1(\mathbb{R})$ and $M(\mathbb{R})$. If we view $*$ as a multiplication on these Banach spaces, then they become algebras. In general, a Banach algebra is a Banach space B with $\|\cdot\|$ and $*$ such that $(B, *)$ is an algebra (not necessarily commutative), and $\|x * y\| \le \|x\|\|y\|$ for all $x, y \in B$.

Note that $(M(\mathbb{R}), *)$ is commutative. Moreover, there is a unit in $M(\mathbb{R})$, by

$$\delta_o(E) := \begin{cases} 1, & o \in E, \\ 0, & \text{else}. \end{cases}$$

It is not hard to see $\delta_o * \mu = \mu$.

Chapter 9

The Fourier Transform

9.1 Formal Properties

Definition 9.1.1. Let $f, g \in L^1$. We define the **convolution** of f and g by

$$(f * g)(x) := \int_{\mathbb{R}} f(x-y)g(y)\,\mathrm{d}y, \quad x \in \mathbb{R}, \tag{9.1.1}$$

and the **Fourier transform** of f as

$$\hat{f}(t) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x)e^{-ixt}\,\mathrm{d}x, \quad t \in \mathbb{R}. \tag{9.1.2}$$

Sometimes we also call the map $\mathcal{F}\colon f \mapsto \hat{f}$ the Fourier transform.

Remark 9.1.2. Observe that $|\hat{f}(t)| \leq \frac{1}{\sqrt{2\pi}}\|f\|_1$, for all $t \in \mathbb{R}$. Thus, $\mathcal{F}\colon L^1 \to L^\infty$ is a bounded linear translation. Hence, we can show that it is uniformly continuous.

Definition 9.1.3. A function $\varphi\colon \mathbb{R} \to \mathbb{C}$ is a character if $|\varphi(t)| = 1$ and $\varphi(s+t) = \varphi(s)\varphi(t)$, for all $s, t \in \mathbb{R}$. In particular, φ is a homomorphism from the additive group $(R, +)$ to the multiplicative group $(\mathbb{C} \setminus \{0\}, \cdot)$. Hence, for all $\alpha \in \mathbb{R}$, $x \mapsto e^{i\alpha x}$ is a character.

Proposition 9.1.4. Elementary Properties of the Fourier Transform

(a) If $g(x) = e^{i\alpha x}f(x)$, then $\hat{g}(t) = \hat{f}(t-\alpha)$.

(b) If $g(x) = f(x-\alpha)$, then $\hat{g}(t) = e^{-i\alpha t}f(t)$.

(c) If $f, g \in L^1(\mathbb{R})$, then $\frac{1}{\sqrt{2\pi}}\widehat{(f * g)} = \hat{f}\hat{g}$

(d) If $g(x) = \overline{f(-x)}$, then $\hat{g}(t) = \overline{\hat{f}(t)}$.

(e) For $\lambda \neq 0$, if $g(x) = f(\frac{x}{\lambda})$, then $\hat{g}(t) = \lambda\hat{f}(\lambda t)$.

(f) If $g(x) = -ixf(x)$ and $g \in L^1$, then $\hat{f}$ is differentiable and $\hat{f}'(t) = \hat{g}(t)$.

Proof. (c). By **Fubini's Theorem** (8.2.5) and translation invariance,

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}}\widehat{(f*g)}(t) &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\left(\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(x-y)g(y)\,\mathrm{d}y\right)e^{-ixt}\,\mathrm{d}x \\
&= \frac{1}{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} f(x-y)e^{-i(x-y)t}g(y)e^{-iyt}\,\mathrm{d}y\,\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(x-y)e^{-i(x-y)t}\,\mathrm{d}x\cdot\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} g(y)e^{-iyt}\,\mathrm{d}y \\
&= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(x)e^{-ixt}\,\mathrm{d}x\cdot\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} g(y)e^{-iyt}\,\mathrm{d}y \\
&= \hat{f}(t)\hat{g}(t).
\end{aligned}$$

■

(f). Consider $\frac{1}{h}(\hat{f}(t+h)-\hat{f}(t))$, $h\neq 0$. Observe that by Hölder's inequality,

$$\int_{\mathbb{R}}\left|\frac{1}{h}(e^{-ihx}-1)f(x)\right|\mathrm{d}x \leq \frac{2}{h}\|f\|_1 < \infty,$$

and $\frac{1}{h}(e^{-ihx}-1)f(x)\in L^1$. By (a) and linearity,

$$\begin{aligned}
\frac{1}{h}(\hat{f}(t+h)-\hat{f}(t)) &= \frac{1}{h}\mathcal{F}\bigg((e^{-ihx}f(x)-f(x))\bigg)(t) \\
&= \mathcal{F}\bigg(\frac{1}{h}(e^{-ihx}-1)f(x)\bigg)(t). \qquad (9.1.3)
\end{aligned}$$

Note that the RHS of equation (9.1.3) is valid and well-defined. Then, given any sequence $h_n\to 0$,

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{h_n}(e^{-ih_nx}-1)f(x) &= \frac{\mathrm{d}}{\mathrm{d}y}[e^{-iyx}]\bigg|_{y=0} f(x) \\
&= -ixf(x) = g(x)\in L^1.
\end{aligned}$$

By sequential continuity of $\mathcal{F}$,[1]

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{h_n}\bigg(\hat{f}(t+h_n)-\hat{f}(t)\bigg) &= \mathcal{F}\bigg(\lim_{n\to\infty}\frac{1}{h}(e^{-ihx}-1)f(x)\bigg)(t) \\
&= \mathcal{F}(g)(t). \qquad (9.1.4)
\end{aligned}$$

Since equation (9.1.4) holds for all sequence $h_n\to 0$, it holds for $h\to 0$. Therefore, by the definition of derivative,

$$\hat{f}'(t) = \hat{g}(t), \quad t\in\mathbb{R}.$$

■

Other properties can be proved by direct substitution. ■

Remark 9.1.5. Here are some basic observations of the Fourier transform: By (a) and (b), it converts multiplication by a **character into translation**, and vice-versa. By (c), it converts **convolutions to pointwise products**. Property (f) shows that it converts **differentiations to products** with ti.

[1] We can also use Dominated Convergence. Since $|\frac{1}{h_n}(e^{-ih_nx}-1)f(x)|\leq|xf(x)|$ and $g\in L^1$, we can carry the limit inside the integral, without continuity of $\mathcal{F}$ and L^∞.

9.2 The Inversion Theorem

In this section, we are working toward the inverse of the Fourier transform. We will first characterize the range of $\mathcal{F}$. To prepare this, we need a continuity result for translation.

Proposition 9.2.1. *Let $f\colon \mathbb{R} \to \mathbb{C}$, $y \in \mathbb{R}$, and $1 \le p < \infty$. Define the translate of f by $f_y(x) := f(x-y)$. If $f \in L^p$, the map $y \mapsto f_y$ is uniformly continuous from $\mathbb{R}$ to L^p.*

Proof. Given $\varepsilon > 0$, by density of $C_c(\mathbb{R})$ in $L^p(\mathbb{R})$, $p < \infty$, there is $g \in C_c(\mathbb{R})$ such that $\|f - g\|_p < \varepsilon$. Also, there is $[-A, A]$ such that $\operatorname{supp}(g) \subset [-A, A]$, for some $A > 0$. Thus, g is uniformly continuous on $[-A, A]$ and $g = 0$ outside. By uniform continuity, there is $0 < \delta < \min\{1, A\}$, so that for all $y, z \in \mathbb{R}$ with $|y - z| < \delta$, we have $|g(y) - g(z)| < (3A)^{-1/p}\varepsilon$. Then,

$$\begin{aligned}
\int_{\mathbb{R}} |g(x-y) - g(x-z)|^p \,\mathrm{d}x &= \int_{-A-\delta}^{A+\delta} |g(x-y) - g(x-z)|^p \,\mathrm{d}x \\
&< \int_{-A-\delta}^{A+\delta} \Big((3A)^{-1/p}\varepsilon\Big)^p \,\mathrm{d}x \\
&= (3A)^{-1}\varepsilon^p(2A+\delta) < \varepsilon^p.
\end{aligned}$$

Thus, $\|g_y - g_z\|_p < \varepsilon$. Finally by $\triangle$-inequality,

$$\begin{aligned}
\|f_y - f_z\|_p &= \|f_y - g_y + g_y - g_z + g_z - f_z\|_p \\
&\le \|f_y - g_y\|_p + \|g_y - g_z\|_p + \|g_z - f_z\|_p \\
&< 3\varepsilon,
\end{aligned}$$

whenever $|y - z| < \delta$. Therefore, $y \mapsto f_y$ is a uniformly continuous map from $\mathbb{R}$ to L^p. ∎

Proposition 9.2.2. *If $f \in L^1$, then $\hat{f} \in C_0(\mathbb{R})$ and $\|\hat{f}\|_\infty \le \frac{1}{\sqrt{2\pi}}\|f\|_1$.*

Proof. Inequality is shown in Remark (9.1.2). For the continuity of $\hat{f}$, let $t_n \to t$ in $\mathbb{R}$. Note that for all n,

$$|f(x)(e^{-it_nx} - e^{-itx})| \le |2f(x)|.$$

So, $f \in L^1(\mathbb{R})$. By Dominated Convergence,

$$\begin{aligned}
\lim_{n\to\infty} |\hat{f}(t_n) - \hat{f}(t)| &= \lim_{n\to\infty} \frac{1}{\sqrt{2\pi}} \left| \int_{\mathbb{R}} f(x)(e^{-it_nx} - e^{-itx}) \,\mathrm{d}x \right| \\
&= \frac{1}{\sqrt{2\pi}} \left| \int_{\mathbb{R}} \lim_{n\to\infty} f(x)(e^{-it_nx} - e^{-itx}) \,\mathrm{d}x \right| \\
&= 0.
\end{aligned}$$

By sequential continuity, $\hat{f}$ is continuous from $\mathbb{R}$ to $\mathbb{C}$. To see $f \in C_0(\mathbb{R})$, consider $e^{-i\pi} = -1$. By translation invariance and change of variable,

$$\begin{aligned}\hat{f}(t) &= -\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-itx} e^{-i\pi} \,\mathrm{d}x \\ &= -\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-it(x+\pi/t)} \,\mathrm{d}x \\ &= -\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x - \frac{\pi}{t}) e^{-itx} \,\mathrm{d}x\end{aligned}$$

Hence,

$$2\hat{f}(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (f_0 - f_{\pi/t}) e^{-itx} \,\mathrm{d}x,$$

and we conclude $|2\hat{f}(t)| \leq \frac{1}{\sqrt{2\pi}} \|f_0 - f_{\pi/t}\|_1$. By Proposition (9.2.1), the map $y \mapsto f_y$ is uniformly continuous. Therefore, as $t \to \pm\infty$, $|\hat{f}(t)| \to 0$. ■

Definition 9.2.3. Let $H(t) := e^{-|t|}$. For $\lambda > 0$, define

$$h_\lambda(x) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} H(\lambda t) e^{itx} \,\mathrm{d}t = \sqrt{\frac{2}{\pi}} \frac{\lambda}{\lambda^2 + x^2}. \tag{9.2.1}$$

Remark 9.2.4. Note that $0 < H(\lambda t) \leq 1$ and $H(\lambda t) \to 1$ as $\lambda \to 0$. The last equality comes from $\int_0^\infty H(\lambda t) e^{itx} \,\mathrm{d}t = \frac{1}{\lambda - ix}$, and $\int_{-\infty}^0 H(\lambda t) e^{itx} \,\mathrm{d}t = \frac{1}{\lambda + ix}$. Moreover, $\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} h_\lambda(x) \,\mathrm{d}x = 1$.

Lemma 9.2.5. *Let $f \in L^1$, then*

$$(f * h_\lambda)(x) = \int_{\mathbb{R}} H(\lambda t) e^{ixt} \hat{f}(t) \,\mathrm{d}t. \tag{9.2.2}$$

Proof. By definitions,

$$\begin{aligned}(f * h_\lambda)(x) &= \int_{\mathbb{R}} f(x-y) h_\lambda(y) \,\mathrm{d}y \\ &= \int_{\mathbb{R}} f(x-y) \left(\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} H(\lambda t) e^{ity} \,\mathrm{d}t \right) \mathrm{d}y \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} f(x-y) H(\lambda t) e^{ity} \,\mathrm{d}t \,\mathrm{d}y.\end{aligned}$$

Note that for each x, $f(x-y) H(\lambda t) e^{ity} \in L^1(m(t) \times m(y))$. By **Fubini's Theorem** (8.2.5), we can switch the order of integrations. Also, $e^{ity} = e^{-it(x-y)} \cdot e^{itx}$. Hence,

$$\begin{aligned}(f * h_\lambda)(x) &= \int_{\mathbb{R}} H(\lambda t) \left(\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x-y) e^{it(x-y)} \,\mathrm{d}y \right) e^{itx} \,\mathrm{d}t \\ &= \int_{\mathbb{R}} H(\lambda t) e^{itx} \hat{f}(t) \,\mathrm{d}t.\end{aligned}$$

■

Lemma 9.2.6. *If $g \in L^\infty$ and g is continuous at $x \in \mathbb{R}$, then*

$$\lim_{\lambda \to 0} \frac{1}{\sqrt{2\pi}} (g * h_\lambda)(x) = g(x). \tag{9.2.3}$$

Proof. Recall that $\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} h_\lambda(x)\,\mathrm{d}x = 1$. Also, note that for $\lambda > 0$,

$$\begin{aligned}\frac{1}{\lambda} h_1(\frac{x}{\lambda}) &= \frac{1}{\lambda} \cdot \sqrt{\frac{2}{\pi}} \left(\frac{1}{1 + \frac{x^2}{\lambda^2}} \right) \\ &= \sqrt{\frac{2}{\pi}} \frac{\lambda}{\lambda^2 + x^2} = h_\lambda(x).\end{aligned}$$

Now consider the following:

$$\begin{aligned}\frac{1}{\sqrt{2\pi}} (g * h_\lambda)(x) - g(x) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(x-y) h_\lambda(y)\,\mathrm{d}y - \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(x) h_\lambda(y)\,\mathrm{d}y \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (g(x-y) - g(x)) h_\lambda(y)\,\mathrm{d}y \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (g(x-y) - g(x)) \frac{1}{\lambda} h_1(\frac{y}{\lambda})\,\mathrm{d}y \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (g(x-\lambda s) - g(x)) h_1(s)\,\mathrm{d}s,\end{aligned}$$

where the last equality is given by change of variable and translation invariance. Now, suppose $\lambda_n \to 0$. Since $|g(x - \lambda s) - g(x)| \leq 2\|g\|_\infty$, and $h_1 \in L^1$, by Dominated Convergence,

$$\begin{aligned}\lim_{n \to \infty} \frac{1}{\sqrt{2\pi}} (g * h_\lambda)(x) - g(x) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \lim_{n \to \infty} (g(x - \lambda_n s) - g(x)) h_1(s)\,\mathrm{d}s \\ &= 0. \end{aligned} \tag{9.2.4}$$

Note that the continuity of g is used in carrying the limit inside $g(x - \lambda_n s)$. Finally, since equation (9.2.4) holds for all $\lambda_n \to 0$, it also holds for $\lambda \to 0$. ∎

Lemma 9.2.7. *Let $1 \leq p < \infty$ and $f \in L^p$. Then*

$$\lim_{\lambda \to 0} \left\| \frac{1}{\sqrt{2\pi}} (f * h_\lambda) - f \right\|_p = 0. \tag{9.2.5}$$

Proof. **Step 1:** Observe that $h_\lambda \in L^q$, where q is the exponent conjugate to p, $q \in [1, \infty]$. By Hölder's inequality, for all $x \in X$,

$$(f * h_\lambda)(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x-y) h_\lambda(y)\,\mathrm{d}y \leq \|f\|_p \|h_\lambda\|_q < \infty,$$

Hence, $(f(x-y) - f(x)) h_\lambda(y) \in \boldsymbol{L^1(m(y))}$. By **Jensen's inequality** (3.1.3),

$$\left| \frac{1}{\sqrt{2\pi}} (f * h_\lambda)(x) - f(x) \right|^p = \left| \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (f(x-y) - f(x)) h_\lambda(y)\,\mathrm{d}y \right|^p$$

$$\leq \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \Big|(f(x-y)-f(x))h_\lambda(y)\Big|^p \, \mathrm{d}y$$
$$\leq \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |f(x-y)-f(x)|^p h_\lambda(y) \, \mathrm{d}y. \tag{9.2.6}$$

Step 2: Integrating equation (9.2.6) over x, we see that

$$\left|\left|\frac{1}{\sqrt{2\pi}}(f * h_\lambda) - f\right|\right|_p^p \leq \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} |f(x-y)-f(x)|^p h_\lambda(y) \, \mathrm{d}y \, \mathrm{d}x. \tag{9.2.7}$$

Recall that $f \in L^p$. Thus, by Hölder's inequality, $|\frac{1}{\sqrt{2\pi}} f(x-y) - f(x)|^p h_\lambda(y) \in L^1(m(y))$. By **Fubini's Theorem** (8.2.5)(c), we can switch the order of integrations in inequality (9.2.7).

Step 3: Define $g \colon \mathbb{R} \to [0, \infty]$, given by

$$g(y) := \int_{\mathbb{R}} |f(x-y)-f(x)|^p \, \mathrm{d}x = \|f_y - f\|_p^p. \tag{9.2.8}$$

Then, we have

$$\left|\left|\frac{1}{\sqrt{2\pi}}(f * h_\lambda) - f\right|\right|_p^p \leq \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(y) h_\lambda(y) \, \mathrm{d}y. \tag{9.2.9}$$

Observe that $\|g\|_\infty \leq \|f_y\|_p^p + \|f\|_p^p \leq 2\|f\|_p^p < \infty$. Thus, $\boldsymbol{g \in L^\infty}$. Moreover, since $y \mapsto f_y$ is uniformly continuous in L^p, g is **continuous** on $\mathbb{R}$.

Step 4: Finally, note that $g(0) = 0$, by Lemma (9.2.6),

$$\begin{aligned}
\lim_{\lambda \to 0} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(y) h_\lambda(y) \, \mathrm{d}y &= \lim_{\lambda \to 0} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(0-(-y)) h_\lambda(-y) \, \mathrm{d}y \\
&= \lim_{\lambda \to 0} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(0-s) h_\lambda(s) \, \mathrm{d}s \\
&= \lim_{\lambda \to 0} \frac{1}{\sqrt{2\pi}} (g * h_\lambda)(0) \\
&= g(0) = 0.
\end{aligned}$$

Therefore, equation (9.2.5) follows from equation (9.2.9) as $\lambda \to 0$. ∎

Theorem 9.2.8 (The Inversion Theorem). *If $f, \hat{f} \in L^1$, and*

$$g(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(t) e^{itx} \, \mathrm{d}t, \tag{9.2.10}$$

then $g \in C_0(\mathbb{R})$ and $f(x) = g(x)$ for m-almost every $x \in \mathbb{R}$.

Proof. By Lemma (9.2.7),

$$\lim_{\lambda \to 0} \left|\left|\frac{1}{\sqrt{2\pi}}(f * h_\lambda) - f\right|\right|_1 = 0.$$

Pick a sequence $\lambda_n \to 0$. Then, $\{\frac{1}{\sqrt{2\pi}}(f * h_{\lambda_n})\}$ is L^1-convergent to f. By Lemma (3.2.8), there is a subsequence $\{\lambda_{n_k}\}$ such that,[2]

$$\frac{1}{\sqrt{2\pi}} \lim_{k\to\infty} (f * h_{\lambda_{n_k}})(x) = f(x), \tag{9.2.11}$$

for m-almost every $x \in \mathbb{R}$. On the LHS of equation (9.2.11), by Lemma (9.2.5) and Dominated Convergence,

$$\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi}} \lim_{k\to\infty} (f * h_{\lambda_{n_k}})(x) \\
&= \frac{1}{\sqrt{2\pi}} \lim_{n\to\infty} \int_{\mathbb{R}} H(\lambda_n t) e^{ixt} \hat{f}(t)\,\mathrm{d}t \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \lim_{n\to\infty} H(\lambda_n t) e^{ixt} \hat{f}(t)\,\mathrm{d}t \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{ixt} \hat{f}(t)\,\mathrm{d}t \\
&= g(x),
\end{aligned}$$

for m-almost all $x \in \mathbb{R}$. Finally, note that $\hat{f} \in L^1$ and for m-almost all $x \in \mathbb{R}$,

$$g(-x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f} e^{-ixt}\,\mathrm{d}t = \mathcal{F}(\hat{f})(x).$$

By Proposition (9.2.2), $g \in C_0(\mathbb{R})$. ■

Corollary 9.2.9 (The Uniqueness Theorem)**.** *If $f \in L^1$ and $\hat{f} = 0$ a.e., then $f = 0$ a.e.*

Proof. Simply let $\hat{f} = 0$ in equation (9.2.10). Then $f(x) = g(x) = 0$, m-a.e. This is saying the Fourier transform as a linear map is injective. ■

9.3 The Plancherel Theorem

Since $m(\mathbb{R}) = \infty$, $L^2 \not\subset L^1$, and the definition of Fourier transform **cannot be applied** on all $f \in L^2$. However, if $f \in L^1 \cap L^2$, then it turns out that $\hat{f} \in L^2$ and $\|f\|_2 = \|\hat{f}\|_2$. In other words, $\mathcal{F}\colon L^1 \cap L^2 \to L^2$ is a **linear isometry**. In this section, we want to extend the Fourier transform to an isometry from L^2 to L^2. This extension is sometimes called the **Plancherel transform**.

Theorem 9.3.1 **(Plancherel's Theorem).** *If $f \in L^1 \cap L^2$, then $\hat{f} \in L^2$ and $\|f\|_2 = \|\hat{f}\|_2$. Moreover, $\mathcal{F}$ extends uniquely to an Hilbert space isometric isomorphism from L^2 to L^2.*

[2] The main reason we are using subsequence here is to pass L^1-convergence to almost everywhere pointwise convergence, then apply Dominated Convergence on the integral. Hence, it is enough to pick one sequence.

Proof. **Step 1:** Suppose $f \in L^1 \cap L^2$. Define $\tilde{f}(x) := \overline{f(-x)}$ and $g := f * \tilde{f}$. Then,

$$\begin{aligned} g(x) &= \int_{\mathbb{R}} f(x-y)\tilde{f}(y)\,\mathrm{d}y = \int_{\mathbb{R}} f(x-y)\overline{f(-y)}\,\mathrm{d}y \\ &= \int_{\mathbb{R}} f(x+s)\overline{f(s)}\,\mathrm{d}s = \langle f_{-x}, f\rangle_{L^2}. \end{aligned}$$

Step 2: Recall that the maps $x \mapsto f_x$ and $\langle \cdot, y\rangle$ are continuous. Thus, $g\colon \mathbb{R} \to \mathbb{C}$ is a **continuous** map. Also, the Cauchy-Schwarz inequality gives

$$|g(x)| = |\langle f_x, f\rangle| \leq \|f_x\|_2\|f\|_2 = \|f\|_2^2.$$

Hence, g is **bounded**. Moreover, by the Fubini's Theorem (8.2.5),

$$\begin{aligned} \int_{\mathbb{R}} |g|\,\mathrm{d}x &= \int_{\mathbb{R}} \left| \int_{\mathbb{R}} f(x+y)\overline{f(y)}\,\mathrm{d}y \right| \mathrm{d}x \\ &\leq \int_{\mathbb{R}}\int_{\mathbb{R}} |f(x+y)||f(y)|\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_{\mathbb{R}} |f(x+y)|\,\mathrm{d}x \int_{\mathbb{R}} |f(y)|\,\mathrm{d}y \\ &= \|f\|_1^2 < \infty. \end{aligned}$$

We conclude that $\boldsymbol{g \in L^1}$.

Step 3: Since $g \in L^1$, by Lemma (9.2.5),

$$(g * h_\lambda)(0) = \int_{\mathbb{R}} H(\lambda t)\hat{g}(t)\,\mathrm{d}t. \tag{9.3.1}$$

On the LHS of equation (9.3.1), since g is bounded and continuous, by Lemma (9.2.6),

$$\lim_{\lambda \to 0}(g * h_\lambda)(0) = \sqrt{2\pi}g(0) = \sqrt{2\pi}\|f\|_2^2. \tag{9.3.2}$$

On the RHS, observe that by the Fubini's Theorem,

$$\begin{aligned} \hat{g}(t) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x+y)\overline{f(y)}\,\mathrm{d}y \right) e^{-itx}\,\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x+y)e^{-it(x+y)}\,\mathrm{d}x \int_{\mathbb{R}} \overline{f(y)}e^{ity}\,\mathrm{d}y \\ &= \hat{f}(t) \cdot \overline{\int_{\mathbb{R}} f(y)e^{-ity}\,\mathrm{d}y} \\ &= \sqrt{2\pi} \cdot |\hat{f}(t)|^2. \end{aligned}$$

Given any sequence $\lambda_n \searrow 0$, we have $H(\lambda_n t) \nearrow 1$. By Monotone Convergence,

$$\lim_{\lambda_n \to 0} \int_{\mathbb{R}} H(\lambda_n t)\hat{g}(t)\,\mathrm{d}t = \int_{\mathbb{R}} \lim_{\lambda_n \to 0} H(\lambda_n t)\hat{g}(t)\,\mathrm{d}t$$

$$
\begin{aligned}
&= \sqrt{2\pi} \int_{\mathbb{R}} |\hat{f}(t)|^2 \, \mathrm{d}t \\
&= \sqrt{2\pi} \|\hat{f}\|_2^2. \qquad (9.3.3)
\end{aligned}
$$

Hence, equations (9.3.2) and (9.3.3) show that $\|f\|_2 = \|\hat{f}\|_2$, and $\hat{f} \in L^2$.

Step 4: For the extension part, recall that $L^1 \cap L^2$ is dense in L^2 because of simple functions. We will first show that $Y := \mathcal{F}(L^1 \cap L^2)$ is dense in L^2. By continuity, it is equivalent to show $Y^{\perp} = \{0\}$. For all $\alpha \in \mathbb{R}$, $\lambda > 0$, define $g_{\alpha,\lambda}(x) := H(\lambda x)e^{i\alpha x}$. Then, by definition (9.2.3),

$$
\mathcal{F}(g_{\alpha,\lambda})(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} H(\lambda x) e^{ix(\alpha - t)} \, \mathrm{d}x = h_\lambda(\alpha - t).
$$

Since $g_{\alpha,\lambda} \in L^1 \cap L^2$, $h_\lambda(\alpha - t) \in Y$. Now, suppose $w \in Y^{\perp}$. By orthogonality,

$$
\begin{aligned}
(h_\lambda * \overline{w})(x) &= \int_{\mathbb{R}} h_\lambda(x-y)\overline{w}(y) \, \mathrm{d}y \\
&= \langle h_\lambda(x-y), w(y) \rangle_{L^2} = 0,
\end{aligned}
$$

for all $x \in \mathbb{R}$. Recall that h_λ is real-valued, by translation invariance,

$$
\begin{aligned}
\langle w(y), h_\lambda(x-y) \rangle_{L^2} &= \int_{\mathbb{R}} w(y)\overline{h_\lambda(x-y)} \, \mathrm{d}y \\
&= \int_{\mathbb{R}} w(x-s) h_\lambda(s) \, \mathrm{d}s \\
&= (w * h_\lambda)(x) = 0,
\end{aligned}
$$

for all $x \in \mathbb{R}$ and $\lambda > 0$. Finally, by Lemma (9.2.7),

$$
\lim_{\lambda \to 0} \left\| \frac{1}{\sqrt{2\pi}} (w * h_\lambda) - w \right\|_2 = \lim_{\lambda \to 0} \|0 - w\|_2 = 0.
$$

Consequently, $w = 0$ and Y is dense in L^2.

Step 5: To summarize, $\mathcal{F} \colon L^1 \cap L^2 \to Y$ maps a dense subspace in L^2 to a dense subspace in L^2 isometrically. Now, we extend $\mathcal{F} \colon L^2 \to L^2$ naturally by

$$
\mathcal{F}(f) := \lim_{n \to \infty} \mathcal{F}(f_n),
$$

where $f_n \in L^1 \cap L^2$ and $f_n \xrightarrow{L^2} f$. Note that such $\mathcal{F}$ is well-defined, unique, and continuous by sequential limit and isometry. Moreover, given any $g \in L^2$, there is $\hat{f}_n \in Y$ such that $\hat{f}_n \xrightarrow{L^2} g$. By isometry, $\mathcal{F}(f_n) \xrightarrow{L^2} g$, where $\mathcal{F}(f_n) = \hat{f}_n$. Since $\{f_n\}$ is L^2-Cauchy, $f_n \xrightarrow{L^2} f \in L^2$. Then,

$$
\mathcal{F}(f) = \lim_{n \to \infty} \mathcal{F}(f_n) = g.
$$

By Proposition (4.5.2), $\mathcal{F}$ is a Hilbert space isometric isomorphism. ∎

Corollary 9.3.2. *Let $\mathcal{F}\colon L^2 \to L^2$ be the Fourier transform. Then, the following symmetric relations hold: For all $f \in L^2$, with $\hat{f} := \mathcal{F}(f)$,*

$$\|\mathcal{F}(f\chi_{[-A,A]}) - \hat{f}\|_2 \to 0, \tag{9.3.4}$$

and

$$\|\mathcal{F}^{-1}(\hat{f}\chi_{[-A,A]}) - f\|_2 \to 0, \tag{9.3.5}$$

as $A \to \infty$.

Proof. By isometry and Dominated Convergence,

$$\begin{aligned}\|\hat{f} - \mathcal{F}(f\chi_{[-A,A]})\|_2 &= \|\mathcal{F}(f - f\chi_{[-A,A]})\|_2 \\ &= \|f - f\chi_{[-A,A]}\|_2 \to 0,\end{aligned}$$

and

$$\begin{aligned}\|f - \mathcal{F}^{-1}(\hat{f}\chi_{[-A,A]})\|_2 &= \|\mathcal{F}^{-1}(\hat{f} - \hat{f}\chi_{[-A,A]})\|_2 \\ &= \|\hat{f} - \hat{f}\chi_{[-A,A]}\|_2 \to 0,\end{aligned}$$

as $A \to \infty$. ∎

Corollary 9.3.3. *Suppose $f \in L^2$ and $\hat{f} \in L^1$, then*

$$f(x) = \int_{\mathbb{R}} \hat{f}(t)e^{itx}\,\mathrm{d}t, \quad m\text{-}a.e.$$

Proof. From Corollary (9.3.2), we see that $\|\mathcal{F}^{-1}(\hat{f}) - f\|_2 = 0$, and $\mathcal{F}^{-1}(\hat{f}) = f$, m-a.e. Since $\hat{f} \in L^1$, the Inversion Theorem (9.2.8) gives

$$f(x) = \int_{\mathbb{R}} \hat{f}(t)e^{itx}\,\mathrm{d}t, \quad m\text{-a.e.}$$ ∎

Remark 9.3.4 (Important Difference between L^1 and L^2). If $f \in L^1$, then the $\hat{f}(t)$ is defined **for all** $t \in \mathbb{R}$ in definition (9.1.1). However, if $f \in L^2$, the Plancherel Theorem defines $\hat{f}$ uniquely as an element in the Hilbert space L^2. In other words, $\hat{f}(t)$ is only defined **almost everywhere**, not as a pointwise function on $\mathbb{R}$.

Appendix A

Hamlos' Approach in the Construction of Measures

A.1 Preliminaries

Definition A.1.1 (Ring)**.** Let X be a set. A set $\mathcal{R} \subset \mathcal{P}(X)$ is a **ring** if for all $E, F \in \mathcal{R}$, we have $E \cup F \in \mathcal{R}$ and $E \setminus F \in \mathcal{R}$.

Example A.1.2. Let $Q = \{[a, b) \subset \mathbb{R} : -\infty < a \leq b < \infty\}$ and $\mathcal{R}$ denote the collection of finite unions of elements in Q. Then $\mathcal{R}$ is a ring.

Proposition A.1.3. *If $\mathcal{E} \subset \mathcal{P}(X)$, then there exists a **unique ring** $\mathcal{R}(\mathcal{E})$ such that $\mathcal{E} \subset \mathcal{R}(\mathcal{E})$ and if $\mathcal{R}$ is another ring such that $\mathcal{E} \subset \mathcal{R}$, then $\mathcal{R}(\mathcal{E}) \subset \mathcal{R}$.*

Proof. Define $\Omega := \{\mathcal{R} \subset \mathcal{P}(X) : \mathcal{R} \text{ is a ring and } \mathcal{E} \subset \mathcal{R}\}$ and define

$$\mathcal{R}(\mathcal{E}) := \bigcap_{\mathcal{R} \in \Omega} \mathcal{R}.$$

We will show that $\mathcal{R}(\mathcal{E})$ is a ring. Let $E, F \in \mathcal{R}(\mathcal{E})$. Then $E, F \in \mathcal{R}$ for all rings $\mathcal{R}$ in Ω. Since each $\mathcal{R}$ is a ring we have that $\mathcal{E} \cup F \in \mathcal{R}$ and $E \setminus F \in \mathcal{R}$, which implies that $E \cup F \in \mathcal{R}(\mathcal{E})$ and $E \setminus F \in \mathcal{R}(\mathcal{E})$. Hence, $\mathcal{R}(\mathcal{E})$ is a ring.

Since $\mathcal{E} \subset \mathcal{R}$ for each $\mathcal{R} \in \Omega$, it follows that $\mathcal{E} \subset \mathcal{R}(\mathcal{E})$. Now suppose that $\mathcal{R}$ is a ring such that $\mathcal{E} \subset \mathcal{R}$. Then, $\mathcal{R} \in \Omega$ and $\mathcal{R}(\mathcal{E}) \subset \mathcal{R}$.

For the uniqueness, suppose that $\mathcal{R}_1(\mathcal{E})$ and $\mathcal{R}_2(\mathcal{E})$ are two such rings that satisfy the required properties. It follows that $\mathcal{R}_1(\mathcal{E}) \subset \mathcal{R}_2(\mathcal{E})$ and $\mathcal{R}_2(\mathcal{E}) \subset \mathcal{R}_1(\mathcal{E})$. Hence, $\mathcal{R}_1(\mathcal{E}) = \mathcal{R}_2(\mathcal{E})$. ■

Example A.1.4. If $Q = \{[a, b) \subset \mathbb{R} : -\infty < a \leq b < \infty\}$, then $\mathcal{R}(Q)$ is given by all finite unions of elements in Q.

Definition A.1.5 (σ-ring)**.** Let X be a set. A set $\mathcal{S} \subset \mathcal{P}(X)$ is a ***σ*-ring** if

i. $\mathcal{S}$ is a ring.

ii. if $E_1, E_2, \ldots \in \mathcal{S}$, then $\bigcup_{j=1}^{\infty} E_j \in \mathcal{S}$.

Theorem A.1.6. *If $X \in \mathcal{S}$ for a σ-ring $\mathcal{S}$, then $\mathcal{S}$ is a σ-algebra on X.*

Proof. First, by assumption we have that $X \in \mathcal{S}$. Next, let $E \in \mathcal{S}$. Then since $X \in \mathcal{S}$ and $\mathcal{S}$ is a ring, $E^C = X \setminus E \in \mathcal{S}$. Thus, $\mathcal{S}$ is closed under complements. Finally, let $E_1, E_2, \ldots \in \mathcal{S}$. Then by definition of a σ-ring we have that $\bigcup_{j=1}^{\infty} E_j \in \mathcal{S}$, and therefore $\mathcal{S}$ is closed under countable unions. Hence, $\mathcal{S}$ is a σ-algebra. ■

Definition A.1.7. Let $\mathcal{E} \subset \mathcal{P}(X)$. We denote $\mathcal{S}(\mathcal{E})$ the smallest σ-ring containing $\mathcal{E}$ and we call $\mathcal{S}(\mathcal{E})$ the **σ-ring generated** by $\mathcal{E}$.

Definition A.1.8 (Monotone class). A non-empty subset $\mathcal{M} \subset \mathcal{P}(X)$ is called a **monotone class** if

i. $\mathcal{M}$ is closed under unions of increasing sequences.

ii. $\mathcal{M}$ is closed under intersections of decreasing sequences.

Definition A.1.9. Let $\mathcal{E} \subset \mathcal{P}(X)$. Then $\mathcal{M}(\mathcal{E})$ is the smallest monotone class containing $\mathcal{E}$ and we call $\mathcal{M}(\mathcal{E})$ the *monotone class generated* by $\mathcal{E}$.

Theorem A.1.10. *If $\mathcal{R}$ is a ring, then $\mathcal{M}(\mathcal{R})$ is a σ-ring.*

Proof. First, if $\mathcal{M}$ is a monotone class and a ring, then $\mathcal{M}$ is a σ-ring. To see this, let $E_1, E_2, \ldots \in \mathcal{M}$, and for each $n \in \mathbb{N}$, define

$$F_n = \bigcup_{j=1}^{n} E_j.$$

Thus $\{F_n\}_{n=1}^{\infty}$ is an increasing sequence and if $\mathcal{M}$ is a monotone class, then $\bigcup_{n=1}^{\infty} F_n \in \mathcal{M}$. Since $\bigcup_{n=1}^{\infty} F_n = \bigcup_{j=1}^{\infty} E_j$, we have that $\bigcup_{j=1}^{\infty} E_j \in \mathcal{M}$. Hence, $\mathcal{M}$ is closed under countable unions and it is a σ-ring.

Next, we show that $\mathcal{M}(\mathcal{R})$ is a ring. For each $F \in \mathcal{M}(\mathcal{R})$, define $K(F)$ to be

$$K(F) = \{E : E \setminus F, F \setminus E, E \cup F \in \mathcal{M}(\mathcal{R})\}.$$

Note that if $F \in \mathcal{R}$, then for all $E \in \mathcal{R}$, $E \in K(F)$. We claim that $K(F)$ is a monotone class. To see this, suppose that $E_1, E_2, \ldots \in K(F)$ with $E_1 \subset E_2 \subset E_3 \subset \cdots$ and let $E := \bigcup_{j=1}^{\infty} E_j$. We then have that

$$E \setminus F = \bigcup_{j=1}^{\infty} (E_j \setminus F).$$

Since $E_j \setminus F \in \mathcal{M}(\mathcal{R})$, by definition of $K(F)$, it follows that $E \setminus F \in \mathcal{M}(\mathcal{R})$. A similar argument shows that $F \setminus E \in \mathcal{M}(\mathcal{R})$ and $F \cup E \in \mathcal{M}(\mathcal{R})$. Thus, $K(F)$ is closed under unions of increasing sequences. Similarly, one can show that

$K(F)$ is closed under intersections of decreasing sequences, and therefore $K(F)$ is a monotone class.

We have seen that if $F \in \mathcal{R}$, then $\mathcal{R} \subset K(F)$ and since $K(F)$ is a monotone class, it follows that $\mathcal{M}(\mathcal{R}) \subset K(F)$.

If $E \in K(F)$, then by symmetry, $F \in K(E)$. Also, we note that if $E \in \mathcal{M}(\mathcal{R})$ and $F \in \mathcal{R}$, then $F \in K(E)$. If $E \in \mathcal{M}(\mathcal{R})$, then $\mathcal{R} \subset K(E)$. We conclude that

$$F \in \mathcal{M}(\mathcal{R}) \qquad \Rightarrow \qquad \mathcal{M}(\mathcal{R}) \subset K(F).$$

Thus, if $E \in \mathcal{M}(\mathcal{R})$, then $E \in K(F)$ and thus $F \setminus E \in \mathcal{M}(\mathcal{R})$ and $E \cup F \in \mathcal{M}(\mathcal{R})$, by the definition of $K(F)$. This means that $\mathcal{M}(\mathcal{R})$ is a ring, and by the first part of the proof, it follows that $\mathcal{M}(\mathcal{R})$ is a σ-ring. ∎

Theorem A.1.11. *If $\mathcal{R}$ is a ring, then $\mathcal{S}(\mathcal{R}) = \mathcal{M}(\mathcal{R})$.*

Proof. From the previous theorem, we have that $\mathcal{M}(\mathcal{R})$ is a σ-ring such that $\mathcal{R} \subset \mathcal{M}(\mathcal{R})$. Since $\mathcal{S}(\mathcal{R})$ is the smallest σ-ring containing $\mathcal{R}$, it follows that $\mathcal{S}(\mathcal{R}) \subset \mathcal{M}(\mathcal{R})$. But a σ-ring is also a monotone class and since $\mathcal{M}(\mathcal{R})$ is the smallest monotone class containing $\mathcal{R}$, it follows that $\mathcal{M}(\mathcal{R}) \subset \mathcal{S}(\mathcal{R})$. ∎

A.2 Measures

Definition A.2.1 (measure). Let $\mathcal{R}$ be a ring. A **measure** on $\mathcal{R}$ is a function $\mu \colon \mathcal{R} \to [0, \infty]$ satisfying

i. $\mu(\varnothing) = 0$.

ii. if $E_1, E_2, \ldots$ are pairwise disjoint elements in $\mathcal{R}$ and if $\bigcup_{j=1}^{\infty} E_j \in \mathcal{R}$, then $\mu(\bigcup_{j=1}^{\infty} E_j) = \sum_{j=1}^{\infty} \mu(E_j)$.

Example A.2.2. Take $Q = \{[a, b) \subset \mathbb{R} : -\infty < a \leq b < \infty\}$ and let $\mathcal{R}(Q)$ be the ring generated by Q. Then there is a measure μ such that $\mu([a, b)) = b - a$. We will give a construction of μ. For now, define $\varphi \colon Q \to [0, \infty)$, by

$$\varphi([a, b)) = b - a.$$

Lemma A.2.3. *If $[a, b] \subset \bigcup_{j=1}^{n} (a_j, b_j)$, then $b - a \leq \sum_{j=1}^{n} (b_j - a_j)$.*

Lemma A.2.4. *If $I_1, I_2, \ldots \in Q$ are pairwise disjoint and $P := \bigcup_{j=1}^{\infty} I_j \in Q$, then $\varphi(P) = \sum_{j=1}^{\infty} \varphi(I_j)$.*

Proof. First suppose we have $P = [a, b)$ and $I_j = [a_j, b_j)$. Given $\varepsilon > 0$, we have

$$[a, b - \varepsilon] \subset P = \bigcup_{j=1}^{\infty} I_j \subset \bigcup_{j=1}^{\infty} (a_j - \frac{\varepsilon}{2^j}, b_j).$$

By compactness of $[a, b-\varepsilon]$ and the Heine-Borel property, there is an $N \in \mathbb{N}$ such that

$$[a, b-\varepsilon] \subset \bigcup_{j=1}^{N}(a_j - \frac{\varepsilon}{2^j}, b_j),$$

and by the preceding lemma it follows that

$$\begin{aligned} b-a-\varepsilon &\leq \sum_{j=1}^{N}(b_j - a_j + \frac{\varepsilon}{2^j}) \leq \sum_{j=1}^{\infty}(b_j - a_j + \frac{\varepsilon}{2^j}) \\ &= \sum_{j=1}^{\infty}(b_j - a_j) + \varepsilon. \end{aligned}$$

Thus we have that $\varphi(P) = b - a \leq \sum_{j=1}^{\infty}(b_j - a_j) + 2\varepsilon = \sum_{j=1}^{\infty}\varphi(I_j) + 2\varepsilon$. Since it holds for all $\varepsilon > 0$, we obtain

$$\varphi(P) \leq \sum_{j=1}^{\infty}\varphi(I_j).$$

To establish the reverse inequality, fix $N \in \mathbb{N}$. Then we see that

$$\bigcup_{j=1}^{N} I_j \subset P = [a, b).$$

Assume that $a_1 \leq a_2 \leq \cdots \leq a_N$. By disjointness we have $b_j \leq a_{j+1}$, $a \leq a_1$, and $b_N \leq b$. Thus,

$$\begin{aligned} \sum_{j=1}^{N}\varphi(I_j) &= (b_1 - a_1) + (b_2 - a_2) + \cdots + (b_N - a_N) \\ &\leq (a_2 - a_1) + (a_3 - a_2) + \cdots + (a_N - a_{N-1}) + (b_N - a_N) \\ &= b_N - a_1 \leq b - a \\ &= \varphi(P). \end{aligned}$$

As $N \to \infty$ we obtain

$$\sum_{j=1}^{\infty}\varphi(I_j) \leq \varphi(P),$$

and therefore $\varphi(P) = \sum_{j=1}^{\infty}\varphi(I_j)$.

Next we show that φ extends uniquely to finite unions of sets in Q, i.e., to $\mathcal{R}(Q)$. Take

$$P_1, P_2, \ldots, P_n \in Q \quad \text{and} \quad S_1, S_2, \ldots, S_m \in Q$$

as pairwise disjoint sequences and assume $E = \cup_{j=1}^{n}P_j = \cup_{l=1}^{m}S_l$. Then let

$$P_{j,l} = P_j \cap S_l.$$

We note that $P_{j,l} \in Q$ and by the preceding argument we have that

$$\varphi(P_j) = \sum_{l=1}^{m} \varphi(P_{j,l})$$

and so

$$\sum_{j=1}^{n} \varphi(P_j) = \sum_{j=1}^{n} \sum_{l=1}^{m} \varphi(P_{j,l}).$$

But the same number holds for the sequence $S_l = \cup_{j=1}^{n} P_{j,l}$ because from the result that $\varphi(S_l) = \sum_{j=1}^{n} \varphi(P_{j,l})$ we obtain

$$\sum_{l=1}^{m} \varphi(S_l) = \sum_{l=1}^{m} \sum_{j=1}^{n} \varphi(P_{j,l}).$$

Now we can define $\mu(E) = \sum_{j=1}^{n} \varphi(P_j)$ for any disjoint sequence P_j such that $E = \cup_{j=1}^{n} P_j$. Now what remains is to show that μ is countably additive. Let $E_1, E_2, \ldots \in \mathcal{R}(Q)$ be a sequence of pairwise disjoint elements and $E = \cup_{j=1}^{\infty} E_j \in \mathcal{R}(Q)$. Write

$$E_j = \bigcup_{l=1}^{n_j} I_{l,j}.$$

Let's assume first that $E \in Q$. Then by our lemma

$$\begin{aligned} \mu(E) = \varphi(E) &= \sum_{j=1}^{\infty} \sum_{l=1}^{n_j} \varphi(I_{l,j}) \\ &= \sum_{j=1}^{\infty} \varphi(E_j) \\ &= \sum_{j=1}^{\infty} \mu(E_j). \end{aligned}$$

For the general case $E \in \mathcal{R}(Q)$, write

$$E = \bigcup_{s=1}^{m} I_s \quad \text{where} \quad I_s \in Q.$$

We have that

$$I_s = I_s \cap E = I_s \cap (\cup_{j=1}^{\infty} E_j) = \bigcup_{j=1}^{\infty} (I_s \cap E_j).$$

From the above work, we see that

$$\mu(I_s) = \sum_{j=1}^{\infty} \mu(I_s \cap E_j)$$

and summing over all the s we get

$$\begin{aligned}\mu(E) = \sum_{s=1}^{m} \mu(I_s) &= \sum_{s=1}^{m} \sum_{j=1}^{\infty} \mu(I_s \cap E_j) \\ &= \sum_{j=1}^{\infty} \sum_{s=1}^{m} \mu(I_s \cap E_j) \\ &= \sum_{j=1}^{\infty} \mu(\cup_{s=1}^{m} (I_s \cap E_j)) \\ &= \sum_{j=1}^{\infty} \mu(E_j \cap (\cup_{s=1}^{m} I_s)) \\ &= \sum_{j=1}^{\infty} \mu(E_j \cap E) \\ &= \sum_{j=1}^{\infty} \mu(E_j).\end{aligned}$$

We recall that $E = \bigcup_{s=1}^{m} I_s$ and

$$E_j = \bigcup_{s=1}^{m} (I_s \cap E_j) = \bigcup_{s=1}^{m} (I_s \cap (\cup_{l=1}^{n_j} I_{l,j})) = \bigcup_{s=1}^{m} \bigcup_{l=1}^{n_j} (I_s \cap I_{l,j})$$

which represents E_j as a disjoint union, so

$$\mu(E_j) = \sum_{s=1}^{m} \sum_{l=1}^{n_j} \varphi(I_s \cap I_{l,j}).$$

On the other hand we have that $\varphi(E) = \sum_{s=1}^{m} \varphi(I_s)$ and

$$I_s \cap E = I_s = I_s \cap (\cup_{j=1}^{\infty} \cup_{l=1}^{n_j} I_{l,j}).$$

From φ being countably additive on Q, we get that

$$\varphi(I_s) = \sum_{j=1}^{\infty} \sum_{l=1}^{n_j} \varphi(I_s \cap I_{l,j})$$

and summing over s we obtaine

$$\begin{aligned}\mu(E) &= \sum_{s=1}^{m} \sum_{j=1}^{\infty} \sum_{l=1}^{n_j} \varphi(I_s \cap I_{l,j}) \\ &= \sum_{j=1}^{\infty} \sum_{s=1}^{m} \sum_{l=1}^{n_j} \varphi(I_s \cap I_{l,j}) \\ &= \sum_{j=1}^{\infty} \mu(E_j).\end{aligned}$$

Thus, we conclude that μ is countably additive on $\mathcal{R}(Q)$. Next, we want to extend μ to the σ-ring $\mathcal{M}(\mathcal{R})$ where $\mathcal{R} = \mathcal{R}(Q)$. ■

Definition A.2.5 (Hereditary σ-ring)**.** Let $\mathcal{R}$ be a ring. Let $H(\mathcal{R})$ denote the set of all $E \subset X$ such that $E_1, E_2, \ldots \in \mathcal{R}$ and $E \subset \bigcup_{j=1}^{\infty} E_j$. We call $H(\mathcal{R})$ the **hereditary σ-ring** generated by $\mathcal{R}$.

Definition A.2.6 (Outer measure)**.** An **outer measure** ν on $H(\mathcal{R})$ is a set function $\nu\colon H(\mathcal{R}) \to [0, \infty]$ satisfies the followings:

i. If $E, F \in H(\mathcal{R})$ and $E \subset F$, then $\nu(E) \leq \nu(F)$.

ii. If $E_1, E_2, \ldots \in H(\mathcal{R})$, then $\nu(\bigcup_{j=1}^{\infty} E_j) \leq \sum_{j=1}^{\infty} \nu(E_j)$.

iii. $\nu(\varnothing) = 0$.

Theorem A.2.7. *Let $\mathcal{R}$ be a ring and μ a measure on $\mathcal{R}$. Then $\mu^* : H(\mathcal{R}) \to [0, \infty]$ defined by*

$$\mu^*(E) = \inf \left\{ \sum_{j=1}^{\infty} \mu(E_j) : E_j \in \mathcal{R}, E \subset \bigcup_{j=1}^{\infty} E_j \right\}$$

is an outer measure. Moreover, if $E \in \mathcal{R}$, then $\mu^*(E) = \mu(E)$.

Proof. Property iii. is included in the last sentence because $\varnothing \in \mathcal{R}$. For property i., suppose $E, F \in H(\mathcal{R})$, and $E \subset F$. By the hereditary property there are $F_1, F_2, \ldots \in \mathcal{R}$ such that

$$E \subset F \subset \bigcup_{j=1}^{\infty} F_j.$$

By taking infimum of sums with measures of such F_j's we conclude

$$\mu^*(E) \leq \mu^*(F).$$

For property ii., consider $E_1, E_2, \ldots \in H(\mathcal{R})$ and let $E := \bigcup_{j=1}^{\infty} E_j$. If for some $j \in \mathbb{N}$, $\mu^*(E_j) = \infty$, then by monotinicity in property i., we have $\mu^*(E) = \infty$.

Thus suppose $\mu^*(E_j) < \infty$ for all $j \in \mathbb{N}$. By definition of μ^*, given $\varepsilon > 0$, there are sets $E_{j,1}, E_{j,2}, \ldots \in \mathcal{R}$ such that

$$E_j \subset \bigcup_{l=1}^{\infty} E_{j,l}$$

and

$$\mu^*(E_j) + \frac{\varepsilon}{2^j} > \sum_{l=1}^{\infty} \mu(E_{j,l}),$$

for all $j \geq 1$. Note that $E \subset \bigcup_{j=1}^{\infty} E_j \subset \bigcup_{j,l=1}^{\infty} E_{j,l}$. Since $\mu^*(E_{j,l}) = \mu(E_{j,l})$ for $E_{j,l} \in \mathcal{R}$, we see that

$$\mu^*(E) \leq \sum_{j,l=1}^{\infty} \mu^*(E_{j,l}) = \sum_{j,l=1}^{\infty} \mu(E_{j,l})$$

$$\leq \sum_{j=1}^{\infty}(\mu^*(E_j) + \frac{\varepsilon}{2^j}) = \sum_{j=1}^{\infty}\mu^*(E_j) + \varepsilon.$$

Now taking $\varepsilon \searrow 0$, andwe see that

$$\mu^*(E) \leq \sum_{j=1}^{\infty}\mu^*(E_j).$$

We conclude that μ^* is an outer measure. ■

A.3 σ-Finite Measures

Definition A.3.1. A measure μ on a ring $\mathcal{R}$ is called ***σ*-finite** if for each $E \in \mathcal{R}$, there are $E_1, E_2, \ldots \in \mathcal{R}$ with $E \subset \bigcup_{j=1}^{\infty} E_j$ and $\mu(E_j) < \infty$ for all $j \geq 1$.

Definition A.3.2. Let μ be a measure on a ring $\mathcal{R}$. A set $E \in H(\mathcal{R})$ is ***μ**-measurable** if for all $A \in H(\mathcal{R})$,

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^C).$$

Theorem A.3.3. *Let μ be a measure on a ring $\mathcal{R}$. The collection $\mathcal{M}$ of all μ^*-measurable sets is a σ-ring containing $\mathcal{R}$. Moreover, restricting μ^* to $\mathcal{M}$ defines a complete measure.*

Proof. We first prove that $\mathcal{M}$ is a ring. Let $E, F \in \mathcal{M}$. For all $A \in H(\mathcal{R})$, then since E is μ^*-measurable,

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^C).$$

But since $A \cap E \in H(\mathcal{R})$ and $A \cap E^c \in H(\mathcal{R})$, we also have

$$\mu^*(A \cap E) = \mu^*(A \cap E \cap F) + \mu^*(A \cap E \cap F^C),$$

and

$$\mu^*(A \cap E^C) = \mu^*(A \cap E^C \cap F) + \mu^*(A \cap E^C \cap F^C).$$

Adding these last two expressions gives

$$\begin{aligned}\mu^*(A) = \mu^*(A \cap E \cap F) + \mu^*(A \cap E \cap F^C)\\ + \mu^*(A \cap E^C \cap F) + \mu^*(A \cap E^C \cap F^C).\end{aligned}$$

Since $A \cap (E \cup F) \in H(\mathcal{R})$, replacing A by $A \cap (E \cup F)$ gives

$$\mu^*(A \cap (E \cup F)) = \mu^*(A \cap E \cap F) + \mu^*(A \cap E \cap F^C) + \mu^*(A \cap E^C \cap F).$$

So comparing these two expressions we have

$$\mu^*(A) = \mu^*(A \cap (E \cup F)) + \mu^*(A \cap E^c \cap F^C)$$

$$= \mu^*(A \cap (E \cup F)) + \mu^*(A \cap (E \cup F)^C).$$

Thus we have shown that $E \cup F \in \mathcal{M}$. Replacing A by $A \cap (E \cap F^C)$ gives that $E \setminus F = E \cap F^C \in \mathcal{M}$. Therefore, $\mathcal{M}$ is a ring.

To show that $\mathcal{M}$ is a σ-ring, take $E_1, E_2, \ldots \in \mathcal{M}$ pairwise disjoint and let $E = \bigcup_{j=1}^{\infty} E_j$. We then have

$$\begin{aligned}\mu^*(A \cap (E_1 \cup E_2)) =& \mu^*(A \cap (E_1 \cup E_2) \cap E_1) + \mu^*(A \cap (E_1 \cup E_2) \cap E_1^C)\\ =& \mu^*(A \cap E_1) + \mu^*(A \cap E_2).\end{aligned}$$

By induction, it follows that $\mu^*(A \cap (\cup_{j=1}^{n} E_j)) = \sum_{j=1}^{n} \mu^*(A \cap E_j)$. Hence, for each $n \in \mathbb{N}$,

$$\begin{aligned}\mu^*(A) &= \mu^*(A \cap (\bigcup_{j=1}^{n} E_j)) + \mu^*(A \cap (\bigcup_{j=1}^{n} E_j)^C)\\ &\geq \mu^*(A \cap (\bigcup_{j=1}^{n} E_j)) + \mu^*(A \cap (\bigcup_{j=1}^{\infty} E_j)^C)\\ &= \sum_{j=1}^{n} \mu^*(A \cap E_j) + \mu^*(A \cap E^C).\end{aligned}$$

Now letting $n \to \infty$ we obtain

$$\mu^*(A) \geq \sum_{j=1}^{\infty} \mu^*(A \cap E_j) + \mu^*(A \cap E^c). \tag{$\dagger$}$$

Since $(A \cap E) \subset \bigcup_{j=1}^{\infty} (A \cap E_j)$, by subadditivity of μ^* we have

$$\mu^*(A \cap E) \leq \sum_{j=1}^{\infty} \mu^*(A \cap E_j). \tag{$\ddagger$}$$

Combining ($\dagger$) and ($\ddagger$) shows that

$$\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \cap E^C).$$

On the other hand, since $A \subset [(A \cap E) \cup (A \cap E^c)]$, we have

$$\mu^*(A) \leq \mu^*(A \cap E) + \mu^*(A \cap E^C)$$

by subadditivity of μ^*. Hence,

$$\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A \cap E^c) \geq \mu^*(A)$$

and $\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$. We conclude that $E \in \mathcal{M}$ and $\mathcal{M}$ is a σ-ring.

Next, we will show that $\mathcal{R} \subset \mathcal{M}$. For all $E \in \mathcal{R}$, take $A \in H(\mathcal{R})$ with $\mu^*(A) < \infty$. For all $\varepsilon > 0$, there are $E_1, E_2, \ldots \in \mathcal{R}$ such that

$$A \subset \bigcup_{j=1}^{\infty} E_j \qquad \text{and} \qquad \mu^*(A) + \varepsilon > \sum_{j=1}^{\infty} \mu^*(E_j).$$

Also,

$$\mu^*(A \cap E) \leq \sum_{j=1}^{\infty} \mu^*(E_j \cap E).$$

Replacing E by E^C we obtain

$$\mu^*(A \cap E^C) \leq \sum_{j=1}^{\infty} \mu^*(E_j \cap E^C).$$

Consequently,

$$\begin{aligned}
\mu^*(A) &\leq \mu^*(A \cap E) + \mu^*(A \cap E^c) \\
&\leq \sum_{j=1}^{\infty} [\mu^*(E_j \cap E) + \mu^*(E_j \cap E^c)] \\
&= \sum_{j=1}^{\infty} \mu^*(E_j) \\
&\leq \mu^*(A) + \varepsilon.
\end{aligned}$$

Since it holds for any $\varepsilon > 0$,

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^C).$$

Therefore, $E \in \mathcal{M}$ and $\mathcal{M} \supset \mathcal{S}(\mathcal{R})$.

Now, to see that μ^* restricted to $\mathcal{M}$ defines a measure, denote

$$\overline{\mu} = \mu^*|_{\mathcal{M}}.$$

We choose $E = A$ in the above calculation. Given a pairwise disjoint sequence $\{E_j\}_{j=1}^{\infty}$ such that $E = \bigcup_{j=1}^{\infty} E_j$, we have

$$\mu^*(E) \geq \sum_{j=1}^{\infty} \mu^*(E \cap E_j) + \mu^*(E^C \cap E_j) = \sum_{j=1}^{\infty} \mu^*(E_j).$$

But since $E = \bigcup_{j=1}^{\infty} E_j$,

$$\mu^*(E) \leq \sum_{j=1}^{\infty} \mu^*(E_j)$$

also holds by subadditivity of μ^*. Thus we get the desired σ-additivity and $\overline{\mu}$ is a measure.