# Measure Theory

## Contents

**Problem**

Consider a random (infinite) string of 0's and 1's

$$(\ell_1, ..., \ell_n, ...) \quad \ell_n \in \{0, 1\} \forall n \in \mathbb{N}$$

What is the probability that somewhere along the string, I find 100 consecutive zeros?

**Comment**

The statement of the problem is unclear. Lebesgue's machinery will clarify this

Define

$$X = \{x = (\ell_1, ..., \ell_n, ...) \mid \ell_n \in \{0, 1\} \forall n \in \mathbb{N}\}$$

This is, in a natural way, a compact metric space. Use the "Hamming-style" distance:
For $x = (\ell_1, ..., \ell_n, ...)$ and $x' = (\ell_1{}', ..., \ell_n{}', ...)$

$$d(x, x') = \sum_{n \in \mathbb{N}} \frac{|\ell_n - \ell_n{}'|}{2^n}$$

Since this is a metric space, we can consider the topology:

$$\mathcal{T} = \{G \subseteq X \mid G \text{ is open}\} \subseteq 2^X = \mathcal{P}(X)$$

Can consider

$$\mathfrak{B} = \sigma\text{-Alg}(\mathcal{T}) \quad \text{(Borel sigma algebra)}$$

Then a probability measure is

$$P : \mathcal{B} \longrightarrow [0, 1]$$

Problem 1.1 says:
Let

$$S = \left\{x = (\ell_1, ..., \ell_n, ...) \in X \;\middle|\; \exists m \in \mathbb{N} \text{ with } \ell_m = \ell_{m+1} = \cdots = \ell_{m+99} = 0\right\}$$

check that $S \in \mathcal{B}$, and compute $P(S)$

# 1 Algebras of Sets

Fix a non-empty set $X$.

**Definition 1.1** **Algebra of Sets**

$\mathfrak{A} \subseteq 2^X$ is an *algebra of sets* if

1. $\varnothing \in \mathfrak{A}$
2. If $A \in \mathfrak{A}$, then $X \setminus A \in \mathfrak{A}$
3. If $A, B \in \mathfrak{A}$, then $A \cap B \in \mathfrak{A}$

**Proposition 1.2** **Properties of algebras of sets**

Let $\mathfrak{A}$ be an algebra of sets on $X$. Then

1. $X \in \mathfrak{A}$
2. If $A, B \in \mathfrak{A}$, then $A \cup B \in \mathfrak{A}$
3. If $A, B \in \mathfrak{A}$, then $A \setminus B \in \mathfrak{A}$
4. If $\mathfrak{B} \subseteq \mathfrak{A}$ is finite, then $\bigcap \mathfrak{B} \in \mathfrak{A}$
5. If $\mathfrak{B} \subseteq \mathfrak{A}$ is finite, then $\bigcup \mathfrak{B} \in \mathfrak{A}$

**Example 1.1**

Obviously, $2^X$ and $\{\varnothing, X\}$ are algebras of sets on $X$

**Lemma 1.3** **Intersections of algebras**

Let $\mathfrak{F}$ be any collection of algebras of sets on $X$, then $\bigcap \mathfrak{F}$ is an algebra of sets on $X$.

***Proof:*** First note if $\mathfrak{F} = \varnothing$, then $\bigcap \mathfrak{F} = 2^X$ which is an algebra of sets on $X$. We now proceed with checking the axioms of an algebra of sets on $X$.

1. For every $\mathfrak{A} \in \mathfrak{F}$, we have $\varnothing \in \mathfrak{A}$, thus $\varnothing \in \bigcap \mathfrak{F}$.
2. If $A \in \bigcap \mathfrak{F}$, then $A \in \mathfrak{A}$ for every $\mathfrak{A} \in \mathfrak{F}$. Thus $X \setminus A \in \mathfrak{A}$ for every $\mathfrak{A} \in \mathfrak{F}$, and thus $X \setminus A \in \bigcap \mathfrak{F}$.
3. If $A, B \in \bigcap \mathfrak{F}$, then $A, B \in \mathfrak{A}$ for every $\mathfrak{A} \in \mathfrak{F}$. Thus $A \cap B \in \mathfrak{A}$ for every $\mathfrak{A} \in \mathfrak{F}$, and thus $A \cap B \in \bigcap \mathfrak{F}$.

$\square$

**Proposition 1.4** **Algebra generated by a collection**

Let $\mathfrak{B} \subseteq 2^X$ be any collection of subsets of $X$, then there exists a unique algebra of sets $\mathfrak{A}$ on $X$ with the following properties:

1. $\mathfrak{B} \subseteq \mathfrak{A}$
2. Whenever $\mathfrak{C}$ is an algebra of sets on $X$ with $\mathfrak{B} \subseteq \mathfrak{C}$, then we have $\mathfrak{A} \subseteq \mathfrak{C}$

***Proof:*** Define $\mathfrak{F}$ to be the collection of all algebras of sets on $X$ that contain $\mathfrak{B}$. Then $\bigcap \mathfrak{F}$ is an algebra of sets by Intersections of algebras (Lemma 1.3) and satisfies (1) and (2). Uniqueness follows from (2). $\square$

**Definition 1.5** **Algebra of sets generated by a collection of subsets**

Let $\mathfrak{B} \subseteq 2^X$ be any collection of subsets of $X$. We define the *algebra of sets generated by* $\mathfrak{B}$ to be the unique algebra from Algebra generated by a collection (Proposition 1.4). We notate it as $\operatorname{Alg}(\mathfrak{B})$.

## 1.1 Additive set-function on an algebra of sets

**Definition 1.6** **Additive set-function**

Let $\mathfrak{A} \subseteq 2^X$ be an algebra of sets on $X$. A function $\mu : \mathfrak{A} \to [0, \infty]$ is called an *additive set-function on* $\mathfrak{A}$ if

1. If $A, B \in \mathfrak{A}$ with $A \cap B = \varnothing$, then $\mu(A \cup B) = \mu(A) + \mu(B)$
2. $\mu(\varnothing) = 0$

**Remark 1.7**

The arithmetic of $[0, \infty]$ is standard, except for $0 \cdot \infty = \infty \cdot 0 = 0$, which is a useful convention in measure theory.

**Remark 1.8**

We almost don't need to require (2), it's just when $\mu$ is identically $\infty$ do things go wrong without (2)

**Proposition 1.9** **Properties of an additive set function**

Let $\mathfrak{A}$ be an algebra of sets on $X$, and let $\mu$ be an additive set function on $\mathfrak{A}$. Then

1. If $A, C \in \mathfrak{A}$ with $A \subseteq C$, then $\mu(A) \leq \mu(C)$
2. If $A, B \in \mathfrak{A}$, then $\mu(A \cup B) \leq \mu(A) + \mu(B)$ (sub additivity)
3. If $A_1, \dots, A_n \in \mathfrak{A}$ are all pairwise disjoint, then $\mu(A_1 \cup \dots \cup A_n) = \mu(A_1) + \dots + \mu(A_n)$
4. If $A_1, \dots, A_n \in \mathfrak{A}$, then $\mu(A_1 \cup \dots \cup A_n) \leq \mu(A_1) + \dots + \mu(A_n)$

# 2 Sigma Algebras

**Definition 2.1** **Sigma Algebra**

$\mathfrak{M} \subseteq 2^X$ is called a *sigma algebra* on $X$ if

1. $\varnothing \in \mathfrak{M}$
2. If $A \in \mathfrak{M}$, then $X \setminus A \in \mathfrak{M}$
3. If $(A_n)_{n=1}^{\infty} \subseteq \mathfrak{M}$, then $\bigcap_{n=1}^{\infty} A_n \in \mathfrak{M}$

**Remark 2.2**

countable unions work too, and so do finite operations

**Lemma 2.3** **Intersections of sigma algebras**

Let $\mathfrak{F}$ be any collection of sigma algebras on $X$, then $\bigcap \mathfrak{F}$ is a sigma algebra on $X$.

***Proof:*** basically identical to the proof for algebras of sets $\square$

**Proposition 2.4** **Sigma algebra generated by a collection**

Let $\mathfrak{B} \subseteq 2^X$ be any collection of subsets of $X$, then there exists a unique sigma algebra $\mathfrak{A}$ on $X$ with the following properties:

1. $\mathfrak{B} \subseteq \mathfrak{A}$
2. Whenever $\mathfrak{C}$ is a sigma algebra on $X$ with $\mathfrak{B} \subseteq \mathfrak{C}$, then we have $\mathfrak{A} \subseteq \mathfrak{C}$

***Proof:*** basically identical to the proof for algebras of sets □

**Definition 2.5** **Sigma algebra generated by a collection of subsets**

Let $\mathfrak{B} \subseteq 2^X$ be any collection of subsets of $X$. We define the *sigma algebra generated by* $\mathfrak{B}$ to be the unique sigma algebra from Sigma algebra generated by a collection (Proposition 2.4). We notate it as $\sigma\text{-Alg}(\mathfrak{B})$.

**Definition 2.6** **Borel Sigma Algebra**

Given a topological space $(X, \mathcal{T})$, the sigma algebra generated by $\mathcal{T}$ is called the *Borel sigma algebra on* $X$ and written $\mathcal{B}_X := \sigma\text{-Alg}(\mathcal{T})$

## 2.1 Positive measure on a sigma algebra

**Remark 2.7** **Increasing limits and series in $[0, \infty]$**

1. Let $(s_n)_{n=1}^{\infty}$ be an *increasing* sequence in $[0, \infty]$ in the sense that

$$s_1 \leq s_2 \leq \cdots$$

   Such a sequence is sure to approach a limit $L \in [0, \infty]$. With a slight abuse of notation we can also write the formula $L = \sup\{s_n \mid n \in \mathbb{N}\}$ meaning that $L$ is the smallest upper bound for all the $s_n$'s. This is usually written in connection to a bounded increasing sequence of numbers in $[0, \infty)$ – our "slight abuse of notation" is that we are also allowing it to cover the case $L = \infty$, which can occur when some of the $s_n$'s are equal to $\infty$, or when $(s_n)_{n=1}^{\infty}$ is an unbounded sequence in $[0, \infty)$.
2. Let $(t_n)_{n=1}^{\infty}$ be any sequence in $[0, \infty]$. We can consider the finite sums

$$s_n := \sum_{i=1}^{n} t_i$$

   which forms an increasing sequence in $[0, \infty]$. We can consider $L = \lim_{n \to \infty} s_n$ as in (1.), and use the familiar notation $L = \sum_{n=1}^{\infty} t_n$

**Definition 2.8** **Positive measure**

Let $\mathfrak{M} \subseteq 2^X$ be sigma algebra of subsets of $X$. We say $\mu : \mathfrak{M} \to [0, \infty]$ is a *positive measure* on $\mathfrak{M}$ if it satisfies:

1. $\mu(\varnothing) = 0$
2. $$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

   Whenever $(A_n)_{n=1}^{\infty}$ is a sequence of sets in $\mathfrak{M}$ with the property that if $i, j \in \mathbb{N}$ with $i \neq j$ we have $A_i \cap A_j = \varnothing$

**Remark 2.9**

A positive measure is, in particular, finitely additive. Thus we can use all the properties in Properties of an additive set function (Proposition 1.9)

**Definition 2.10** **Nomenclature involving measures**

1. If $\mathfrak{M}$ is a sigma algebra on $X$, then the pair $(X, \mathfrak{M})$ is called a *measurable space*
2. If we also have a positive measure $\mu$ on $\mathfrak{M}$, then the triple $(X, \mathfrak{M}, \mu)$ is called a *measure space*
3. If $\mu(X) < \infty$, then we say $\mu$ is a *finite measure*
4. If $\mu(X) = 1$, then we say $\mu$ is a *probability measure* and that $(X, \mathfrak{M}, \mu)$ is a *probability space*
5. We say that $\mu$ is *sigma finite* if there exists a sequence of sets $(X_n)_{n=1}^{\infty}$ in $\mathfrak{M}$ with
   a. $X_1 \subseteq X_2 \subseteq \cdots$ (The $X_n$'s are said to form an increasing chain)
   b. $\bigcup_{n=1}^{\infty} X_n = X$
   c. $\mu(X_n) < \infty$ for all $n \in \mathbb{N}$

**Proposition 2.11** **Sigma sub-additivity**

Let $\mathfrak{M}$ be a sigma algebra on $X$, and let $\mu$ be a positive measure on $\mathfrak{M}$. If $(A_n)_{n=1}^{\infty}$ is a sequence of sets in $\mathfrak{M}$, then

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n)$$

**Proposition 2.12** **Continuity along chains**

Let $\mathfrak{M}$ be a sigma algebra on $X$, and let $\mu$ be a positive measure on $\mathfrak{M}$.

1. If $(A_n)_{n=1}^{\infty}$ is an increasing chain in $\mathfrak{M}$, then
$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n)$$
2. If $(A_n)_{n=1}^{\infty}$ is a decreasing chain in $\mathfrak{M}$ with $\mu(A_1) < \infty$, then
$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n)$$

**Example 2.1**

Example of "atomic" $\mu$ based on weights of individual points $x \in X$. Suppose you have $w : X \to [0, \infty]$. Let $\mathfrak{M} = 2^X$, for every $A \subseteq X$, put

$$\mu(A) = \sup\left\{\sum_{x\in F} w(x) \;\middle|\; F \subseteq A \text{ finite}\right\}$$

Then $(X, \mathfrak{M}, \mu)$ is a measure space. For the special case that $w(x) = 1, \forall x \in X$ is called the *counting measure* on $X$. For the special case that $w(x_0) = 1$ for some $x_0 \in X$ and $w(x) = 0, \forall x \neq x_0$ is called the *Dirac measure concentrated at* $x_0$

## 2.2 A glimpse of the Lebesgue measure on $\mathbb{R}$

**Note**

1. $\mathcal{B}_{\mathbb{R}}$ can be generated from the half open intervals
2. If $\mu, \nu : \mathcal{B}_{\mathbb{R}} \to [0, \infty]$ are positive measures such that
$$\mu((a, b]) = \nu((a, b]) < \infty, \forall a < b \in \mathbb{R}$$
It follows that $\mu = \nu$

Then the *Lebesgue measure* $\lambda : \mathcal{B}_{\mathbb{R}} \to [0, \infty]$ is determined via with requirement that $\lambda((a, b]) = b - a, \forall a < b \in \mathbb{R}$.
Another approach: $\lambda$ is uniquely determined by the requirement that $\lambda(B + t) = \lambda(B)$ for every $B \in \mathcal{B}_{\mathbb{R}}$ and $t \in \mathbb{R}$ with the normalization that $\lambda((0, 1]) = 1$.

**Note**

For the *existence* of $\lambda$, will invoke an extension theorem of Caratheodory (to be done later)

**Note**

Interesting fact: there is no positive measure $\mu : 2^{\mathbb{R}} \to [0, \infty]$ which is translation invariant and $\mu((0, 1]) = 1$. This shows, in particular that $\mathcal{B}_{\mathbb{R}} \neq 2^{\mathbb{R}}$

# 3 Measurable functions

**Definition 3.1** **Measurable function**

Let $(X, \mathfrak{M})$, $(Y, \mathfrak{N})$ be measurable spaces and $f : X \to Y$. Then $f$ is called $\mathfrak{M}/\mathfrak{N}$-*measurable* if $f^{-1}(N) \in \mathfrak{M}$ for every $N \in \mathfrak{N}$.

**Remark 3.2**

Pre-images behave nicely with respect to the usual set operations.

**Proposition 3.3** **Composition of measurable maps**

Compositions of measurable maps are measurable.

***Proof:*** Obvious □

**Proposition 3.4** **Measurability criterion**

Let $(X, \mathfrak{M})$, $(Y, \mathfrak{N})$ be measurable spaces. $f : X \to Y$, and $\mathfrak{C} \subseteq \mathfrak{N}$ such that $\sigma\text{-Alg}(\mathfrak{C}) = \mathfrak{N}$ and such $f^{-1}(C) \in \mathfrak{M}$ for every $C \in \mathfrak{C}$. Then $f$ is measurable.

***Proof:*** Let $\mathfrak{Q} = \{Q \subseteq Y \mid f^{-1}(Q) \in \mathfrak{M}\} \subseteq 2^Y$

***Claim:*** $\mathfrak{Q}$ is a sigma algebra

Note by assumption $\mathfrak{C} \subseteq \mathfrak{Q}$ thus $\mathfrak{N} = \sigma\text{-Alg}(\mathfrak{C}) \subseteq \mathfrak{Q}$. Thus $f$ is measurable. □

**Corollary 3.5** **Continuous maps are measurable**

Let $X, Y$ be topological spaces, $f : X \to Y$ continuous. Then $f$ is $\mathcal{B}_X/\mathcal{B}_Y$-measurable.

## 3.1 The space of functions $\mathrm{Bor}(X, \mathbb{R})$

**Definition 3.6** **Borel measurable real-valued functions**

Given a measurable space $(X, \mathfrak{M})$. Define

$$\mathrm{Bor}(X, \mathbb{R}) := \{f : X \to \mathbb{R} \mid f \text{ is } \mathfrak{M}/\mathcal{B}_{\mathbb{R}} - \text{measurable}\}$$

**Proposition 3.7** **Algebra and lattice operations**

Let $(X, \mathfrak{M})$ be a measurable space with $f, g \in \operatorname{Bor}(X, \mathbb{R})$. Then

1. $f + g \in \operatorname{Bor}(X, \mathbb{R})$
2. For any $c \in \mathbb{R}$, $cf \in \operatorname{Bor}(X, \mathbb{R})$
3. $f \cdot g \in \operatorname{Bor}(X, \mathbb{R})$
4. $f \vee g, f \wedge g \in \operatorname{Bor}(X, \mathbb{R})$

***Proof of 3:*** Let $f, g \in \operatorname{Bor}(X, \mathbb{R})$. Let $h : X \to \mathbb{R}^2$ given by $h(x) = (f(x), g(x))$. Let $p : \mathbb{R}^2 \to \mathbb{R}$ given by $p(s, t) = s \cdot t$.

***Claim:*** $h$ is measurable

***Claim:*** $p$ is measurable

***Claim:*** $f \cdot g = p \circ h$

Thus $f \cdot g \in \operatorname{Bor}(X, \mathbb{R})$ □

**Exercise 3.1**

Let $\mathfrak{S}$ be the collection of open squares in $\mathbb{R}^2$. Then $\sigma\text{-Alg}(\mathfrak{S}) = \mathcal{B}_{\mathbb{R}^2}$.

**Remark 3.8**

Let $(X, \mathfrak{M})$ be a measurable space. Then $\operatorname{Bor}(X, \mathbb{R})$ is an *algebra of functions* and a *lattice of functions*.

**Remark 3.9**

Operations with only one function. If $f \in \operatorname{Bor}(X, \mathbb{R})$, then $|f| \in \operatorname{Bor}(X, \mathbb{R})$. Use algebraic stability properties to prove this.

**Example 3.1** **Problem 3.1**

***Claim:*** Let $\mathfrak{C} = \{(-\infty, t], \mid t \in \mathbb{R}\} \subseteq \mathcal{B}_{\mathbb{R}}$. Then $\sigma\text{-Alg}(\mathfrak{C}) = \mathcal{B}_{\mathbb{R}}$.

***Proof:*** Denote $\sigma\text{-Alg}(\mathfrak{C}) = \mathfrak{P}$. Clearly, $\mathfrak{P} \subseteq \mathcal{B}_{\mathbb{R}}$. If we can show that $G \in \mathfrak{P}$ for all $G \subseteq \mathbb{R}$ open, then we will be done because we will have that $\mathfrak{P}$ is a sigma-algebra and that $\mathfrak{P} \supseteq \mathcal{T}_{\mathbb{R}}$.
Now, every $\varnothing \neq G \subseteq \mathbb{R}$ open can be written as a countable union of open intervals. So then, it suffices to check that $\mathfrak{P}$ contains all of the open intervals of $\mathbb{R}$. This is left as an exercise. □

Let $(X, \mathfrak{M})$ be a measurable space, and suppose $f : X \to \mathbb{R}$ satisfies $f^{-1}((-\infty, t]) \in \mathfrak{M}$ for all $t \in \mathbb{R}$. That is, $f^{-1}(C) \in \mathfrak{M}$ for all $C \in \mathfrak{C}$. Then by Measurability criterion (Proposition 3.4), $f^{-1}(B) \in \mathfrak{M}$ for all $B \in \mathcal{B}_{\mathbb{R}}$, hence $f \in \operatorname{Bor}(X, \mathbb{R})$.

# 4 Convergence properties of $\mathrm{Bor}(X, \mathbb{R})$

**Proposition 4.1** **Closure under suprema and infima**

Let $(X, \mathfrak{M})$ be a measurable space.

1. Let $f : X \to \mathbb{R}$ be a function. Let $(f_n)_{n=1}^{\infty}$ be a sequence of functions in $\mathrm{Bor}(X, \mathbb{R})$ such that
$$\sup\{f_n(x) \mid n \in \mathbb{N}\} = f(x)$$
for all $x \in X$. Then it follows that $f \in \mathrm{Bor}(X, \mathbb{R})$.
2. Let $g : X \to \mathbb{R}$ be a function. Let $(g_n)_{n=1}^{\infty}$ be a sequence of functions in $\mathrm{Bor}(X, \mathbb{R})$ such that
$$\inf\{g_n(x) \mid n \in \mathbb{N}\} = g(x)$$
for all $x \in X$. Then it follows that $g \in \mathrm{Bor}(X, \mathbb{R})$.

***Proof:***

1. It suffices to show that $f^{-1}((-\infty, t]) \in \mathfrak{M}$ for all $t \in \mathbb{R}$. This is left as an exercise.
2. Exercise.

$\square$

**Remark 4.2**

Let $(t_n)_{n=1}^{\infty}$ be a bounded sequence in $\mathbb{R}$. Then there exists $\ell_+ \in \mathbb{R}$ such that

1. There exists a subsequence of $(t_n)_{n=1}^{\infty}$ which converges to $\ell_+$.
2. There is no subsequence of $(t_n)_{n=1}^{\infty}$ which converges to an $\ell > \ell_+$.

Define
$$\limsup_{n\to\infty} t_n := \ell_+$$
We can obtain $\ell_+$ as
$$\ell_+ = \lim_{k\to\infty}\left(\sup_{n\geq k} t_n\right) = \inf_{k\geq 1}\left(\sup_{n\geq k} t_n\right)$$
We can define $\liminf_{n\to\infty} t_n$ similarly.

Fix a measurable space $(X, \mathfrak{M})$

**Proposition 4.3**

1. Let $f : X \to \mathbb{R}$. Suppose we could find $f_1, f_2, \ldots \in \mathrm{Bor}(X, \mathbb{R})$ such that for every $x \in X$, the sequence $(f_n(x))_{n=1}^{\infty}$ is bounded with
$$\limsup_{n\to\infty} f_n(x) = f(x)$$
Then it follows that $f \in \mathrm{Bor}(X, \mathbb{R})$
2. Let $g : X \to \mathbb{R}$. Suppose we could find $g_1, g_2, \ldots \in \mathrm{Bor}(X, \mathbb{R})$ such that for every $x \in X$, the sequence $(g_n(x))_{n=1}^{\infty}$ is bounded with
$$\liminf_{n\to\infty} g_n(x) = g(x)$$
Then it follows that $g \in \mathrm{Bor}(X, \mathbb{R})$

***Proof of 1:*** For every $k \in \mathbb{N}$, define the helper function $h_k : X \to \mathbb{R}$ by
$$h_k(x) = \sup \underbrace{\{f_n(x) \mid n \geq k\}}_{\text{bounded by hypothesis}}, \ \forall x \in X$$
Observe $f_k, f_{k+1}, \ldots \in \mathrm{Bor}(X, \mathbb{R})$. We can invoke Closure under suprema and infima (Proposition 4.1), to conclude that $h_k \in \mathrm{Bor}(X, \mathbb{R})$. Observe for every $x \in X$, we have
$$h_1(x) \geq h_2(x) \geq \cdots$$
with
$$f(x) = \lim_{k\to\infty} h_k(x) = \inf\{h_k(x) \mid k \in \mathbb{N}\}$$
This is because $f(x) = \limsup_{n\to\infty} f_n(x)$ by hypothesis. But then
$$h_k \in \mathrm{Bor}(X, \mathbb{R}), \ \forall k \in \mathbb{N}$$
and
$$f(x) = \inf\{h_k(x) \mid k \in \mathbb{N}\} \ \forall x \in X$$
thus $f \in \mathrm{Bor}(X, \mathbb{R})$ by Closure under suprema and infima (Proposition 4.1). □

***Proof of 2:*** We can reduce to (1), by introducing $f : X \to \mathbb{R}$, $f(x) = -g(x) \ \forall x \in X$ and introducing $f_n = -g_n \in \mathrm{Bor}(X, \mathbb{R}), \forall n \in \mathbb{N}$, then apply (1) to $f(x) = \limsup_{n\to\infty} f_n(x)$, to get $f \in \mathrm{Bor}(X, \mathbb{R})$, hence $g \in \mathrm{Bor}(X, \mathbb{R})$. □

**Corollary 4.4**

Let $f : X \to \mathbb{R}$. Suppose we could find $(f_n)_{n=1}^{\infty}$ in $\mathrm{Bor}(X, \mathbb{R})$ such that $f(x) = \lim_{n\to\infty} f_n(x), \ \forall x \in X$. Then it follows that $f \in \mathrm{Bor}(X, \mathbb{R})$.

***Proof:*** Use Proposition 4.3. □

**Exercise 4.1**

Suppose that $(X, \mathfrak{M}) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable and consider the function $f' : \mathbb{R} \to \mathbb{R}$ ($f'$ may not be continuous).

***Solution:***
Idea is to produce $(f_n)_{n=1}^{\infty}$ in $\mathrm{Bor}(\mathbb{R}, \mathbb{R})$ such that $\lim_{n\to\infty} f_n(x) = f'(x)$. If we find the $f_n$'s then $f' \in \mathrm{Bor}(\mathbb{R}, \mathbb{R})$ due to Corollary 4.4. How do we find the $f_n$'s? For $x \in \mathbb{R}$

$$f'(x) = \lim_{n\to\infty} \frac{f\left(x + \frac{1}{n}\right) - f(x)}{\frac{1}{n}} = \lim_{n\to\infty} n(u_n(x) - f(x))$$

With $u_n(x) = f\left(x + \frac{1}{n}\right)$. So then, for every $n \in \mathbb{N}$, let $u_n : \mathbb{R} \to \mathbb{R}$ be defined by $u_n(x) = f\left(x + \frac{1}{n}\right),\ \forall x \in \mathbb{R}$ and let $f_n = n(u_n - f)$. Every $f_n$ is continuous, hence in $\mathrm{Bor}(\mathbb{R}, \mathbb{R})$ by Continuous maps are measurable (Corollary 3.5) and $\lim_{n\to\infty} f_n(x) = f'(x)$ as we wanted.

## 4.1 Approximation by simple functions

**Definition 4.5** **Simple functions**

A function $f \in \mathrm{Bor}(X, \mathbb{R})$ is said to be *simple* when it takes finitely many values, that is $f(X)$ is a finite subset of $\mathbb{R}$. We denote

$$\mathrm{Bor}_s(X, \mathbb{R}) := \{f \in \mathrm{Bor}(X, \mathbb{R}) \mid f \text{ is simple}\}$$

We also introduce the spaces

$$\mathrm{Bor}^+(X, \mathbb{R}) = \{f \in \mathrm{Bor}(X, \mathbb{R}) \mid f(x) \geq 0\ \forall x \in X\}$$

and

$$\mathrm{Bor}_s^+(X, \mathbb{R}) := \mathrm{Bor}_s(X, \mathbb{R}) \cap \mathrm{Bor}^+(X, \mathbb{R})$$

**Note**

For $A \in \mathfrak{M}$, we have $\chi_A \in \mathrm{Bor}_s(X, \mathbb{R})$, where $\chi_A : X \to \mathbb{R}$ is defined by

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \in X \setminus A \end{cases}$$

called the *indicator function of* $A$. Also note that $\chi_A$ is $\mathfrak{M}/\mathcal{B}_{\mathbb{R}}$-measurable since for every $S \in \mathcal{B}_{\mathbb{R}}$ we have that

$$\chi_A^{-1}(S) \in \{\varnothing, A, X \setminus A, X\} \subseteq \mathfrak{M}$$

**Note**

It can be shown that

$$\operatorname{Bor}_s(X,\mathbb{R}) = \underbrace{\operatorname{span}\{\chi_A \mid A \in \mathfrak{M}\}}_{\substack{\text{set of all linear combinations} \\ \text{of indicator functions} \\ \chi_A \text{ with } A\in\mathfrak{M}}}$$

**Remark 4.6** **Binning values of a non-negative Borel function**

Let $f \in \operatorname{Bor}^+(X,\mathbb{R})$. For every $n \in \mathbb{N}$, let $f_n : X \to \mathbb{R}$ be the *$n$-th binning of the values of $f$*, defined as follows. For each $x \in X$:

1. If $f(x) \geq n$, define $f_n(x) := n$.
2. If $f(x) < n$, pick the unique $k \in \{1, \ldots, n \cdot 2^n\}$ such that

$$\frac{k-1}{2^n} \leq f(x) < \frac{k}{2^n}$$

   and define $f_n(x) := \frac{k-1}{2^n}$.

**Proposition 4.7** **Binning approximation**

Let $f \in \operatorname{Bor}^+(X,\mathbb{R})$, and for every $n \in \mathbb{N}$, let $f_n$ be the $n$-th binning of the values of $f$. Then

1. $f_n \in \operatorname{Bor}_s^+(X,\mathbb{R})$ for every $n \in \mathbb{N}$
2. $f_n \leq f_{n+1}$ for every $n \in \mathbb{N}$
3. For every $x \in X$,

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

# 5 Integration of functions in $\operatorname{Bor}^+(X,\mathbb{R})$

## 5.1 The map $L_s^+ : \operatorname{Bor}_s^+(X,\mathbb{R}) \longrightarrow [0,\infty]$

Fix a measure space $(X,\mathfrak{M},\mu)$

**Note**

For didactical reasons, it is convenient to first define $\int_X f(x)\,\mathrm{d}\mu(x)$ for $f \in \operatorname{Bor}^+(X,\mathbb{R})$. In this case, the integral is always defined but may be $\infty$. For general $f \in \operatorname{Bor}(X,\mathbb{R})$, the integral may not be defined; to get a finite integral, we need $f \in \mathcal{L}^1(\mu)$.
Moreover, it will be convenient to introduce the integral on $\operatorname{Bor}^+(X,\mathbb{R})$ as a map

$$L^+ : \operatorname{Bor}^+(X,\mathbb{R}) \longrightarrow [0,\infty].$$

We will get $L^+$ in two parts:

1. Define $L_s^+ : \operatorname{Bor}_s^+(X,\mathbb{R}) \to [0,\infty]$
2. Extend $L_s^+$ to a map $L^+ : \operatorname{Bor}^+(X,\mathbb{R}) \to [0,\infty]$

**Remark 5.1** **Easily verified properties of $\mathrm{Bor}_s^+(X, \mathbb{R})$**

1. If $f, g \in \mathrm{Bor}_s^+(X, \mathbb{R})$, then $f + g \in \mathrm{Bor}_s^+(X, \mathbb{R})$
2. If $f \in \mathrm{Bor}_s^+(X, \mathbb{R})$ and $\alpha \in [0, \infty)$, then $\alpha f \in \mathrm{Bor}_s^+(X, \mathbb{R})$
3. If $A \in \mathfrak{M}$, then $\chi_A \in \mathrm{Bor}_s^+(X, \mathbb{R})$

**Note**

As a consequence of (3), taking $A = \varnothing$ and $A = X$, we get $\underline{1}, \underline{0} : X \to \mathbb{R} \in \mathrm{Bor}_s^+(X, \mathbb{R})$.

**Theorem 5.2** **Integral of non-negative simple functions**

There exists a map

$$L_s^+ : \mathrm{Bor}_s^+(X, \mathbb{R}) \longrightarrow [0, \infty],$$

uniquely determined, such that the following conditions are satisfied:

(i) Additivity:

$$L_s^+(f + g) = L_s^+(f) + L_s^+(g)$$

for all $f, g \in \mathrm{Bor}_s^+(X, \mathbb{R})$

(ii) Positive homogeneity:

$$L_s^+(\alpha f) = \alpha L_s^+(f)$$

for all $f \in \mathrm{Bor}_s^+(X, \mathbb{R})$ and $\alpha \in [0, \infty)$

(iii) Relation to the measure:

$$L_s^+(\chi_A) = \mu(A)$$

for all $A \in \mathfrak{M}$

In order to prove the existence of the map $L_s^+$, we will take advantage of the fact that one can describe the functions from $\mathrm{Bor}_s^+(X, \mathbb{R})$ in terms of measurable divisions of $X$.

**Definition 5.3** **Measurable division and intersections of measurable divisions**

1. A *measurable division* of $X$ is a set $\{A_1, ..., A_n\} \subseteq \mathfrak{M}$ of pairwise disjoint, non-empty sets such that $\bigcup_{i=1}^{n} A_i = X$
2. Given two measurable divisions $\Delta = \{A_1, ..., A_n\}$ and $\Gamma = \{B_1, ..., B_m\}$ of $X$, one can "intersect" them, to form a new measurable division of $X$, denoted $\Delta \wedge \Gamma$, defined by

$$\Delta \wedge \Gamma := \left\{A_i \cap B_j \;\middle|\; 1 \leq i \leq n, 1 \leq j \leq m, \text{ and } A_i \cap B_j \neq \varnothing\right\}$$

**Lemma 5.4** **Simple functions and measurable divisions**

Let $f \in \mathrm{Bor}_s^+(X, \mathbb{R})$.

1. There exists a measurable division $\Delta = \{A_1, ..., A_n\}$ of $X$ and numbers $\alpha_1, ..., \alpha_n \in [0, \infty)$ such that
$$f = \alpha_1 \chi_{A_1} + \dots + \alpha_n \chi_{A_n}$$
2. Suppose that $\Delta = \{A_1, ..., A_n\}$ and $\alpha_1, ..., \alpha_n \in [0, \infty)$ are as in (1.), and suppose that $\Gamma = \{B_1, ..., B_m\}$ is a second measurable division of $X$. Then
$$f = \sum_{\substack{1 \leq i \leq n, \\ 1 \leq j \leq m, \\ A_i \cap B_j \neq \varnothing}} \alpha_i \chi_{A_i \cap B_j}$$

***Proof of 2:*** For every $1 \leq i \leq n$, write
$$\chi_{A_i} = \chi_{A_i} \cdot 1 = \chi_{A_i}\left(\chi_{B_1} + \dots + \chi_{B_m}\right) = \sum_{j=1}^{m} \chi_{A_i \cap B_j} = \sum_{\substack{1 \leq j \leq m, \\ A_i \cap B_j \neq \varnothing}} \chi_{A_i \cap B_j}$$

Then
$$f = \sum_{i=1}^{n} \alpha_i \chi_{A_i} = \sum_{i=1}^{n} \alpha_i \left( \sum_{\substack{1 \leq j \leq m, \\ A_i \cap B_j \neq \varnothing}} \chi_{A_i \cap B_j} \right) = \sum_{\substack{1 \leq i \leq n, \\ 1 \leq j \leq m, \\ A_i \cap B_j \neq \varnothing}} \alpha_i \chi_{A_i \cap B_j}$$

as desired. □

**Lemma 5.5** **Well-definedness of the simple integral formula**

Let $f \in \mathrm{Bor}_s^+(X, \mathbb{R})$. Suppose we have two ways of writing $f$ as a linear combination:
$$f = \alpha_1 \chi_{A_1} + \dots + \alpha_n \chi_{A_n}$$
and
$$f = \beta_1 \chi_{B_1} + \dots + \beta_m \chi_{B_m},$$
where $\Delta = \{A_1, ..., A_n\}$ and $\Gamma = \{B_1, ..., B_m\}$ are measurable divisions of $X$, and $\alpha_1, ..., \alpha_n, \beta_1, ..., \beta_m \in [0, \infty)$. Then
$$\sum_{i=1}^{n} \alpha_i \mu(A_i) = \sum_{j=1}^{m} \beta_j \mu(B_j) \in [0, \infty].$$

***Proof:*** For every $1 \leq i \leq n$, let us write

$$A_i = A_i \cap X = A_i \cap (B_1 \cup \cdots \cup B_m) = \bigcup_{j=1}^{m} (A_i \cap B_j) = \bigcup_{\substack{1 \le j \le m, \\ A_i \cap B_j \neq \varnothing}} (A_i \cap B_j)$$

Since the union is disjoint,

$$\mu(A_i) = \sum_{\substack{1 \le j \le m, \\ A_i \cap B_j \neq \varnothing}} \mu(A_i \cap B_j)$$

so

$$\sum_{i=1}^{n} \alpha_i \mu(A_i) = \sum_{i=1}^{n} \alpha_i \sum_{\substack{1 \le j \le m, \\ A_i \cap B_j \neq \varnothing}} \mu(A_i \cap B_j) = \sum_{\substack{1 \le i \le n, \\ 1 \le j \le m, \\ A_i \cap B_j \neq \varnothing}} \alpha_i \mu(A_i \cap B_j)$$

Similarly,

$$\sum_{j=1}^{m} \beta_j \mu(B_j) = \sum_{j=1}^{m} \beta_j \sum_{\substack{1 \le i \le n, \\ B_j \cap A_i \neq \varnothing}} \mu(B_j \cap A_i) = \sum_{\substack{1 \le i \le n, \\ 1 \le j \le m, \\ A_i \cap B_j \neq \varnothing}} \beta_j \mu(A_i \cap B_j)$$

***Claim:*** If $A_i \cap B_j \neq \varnothing$, then $\alpha_i = \beta_j$

Observe that this claim will finish the proof. To verify the claim, suppose $A_i \cap B_j \neq \varnothing$. Pick $x \in A_i \cap B_j$, and evaluate $f(x)$ in two ways:

$$f(x) = \alpha_1 \chi_{A_1}(x) + \cdots + \alpha_n \chi_{A_n}(x) = \alpha_i$$

similarly

$$f(x) = \beta_1 \chi_{B_1}(x) + \cdots + \beta_m \chi_{B_m}(x) = \beta_j,$$

so $\alpha_i = f(x) = \beta_j$ $\square$

***Proof of existence of $L_s^+$ from Integral of non-negative simple functions (Theorem 5.2):*** For every $f \in \mathrm{Bor}_s^+(X, \mathbb{R})$, we define the quantity $L_s^+(f) \in [0, \infty]$ by the following procedure:

Consider a writing

$$f = \alpha_1 \chi_{A_1} + \cdots + \alpha_n \chi_{A_n}$$

with $\Delta = \{A_1, \ldots, A_n\}$ a measurable division of $X$ and $\alpha_1, \ldots, \alpha_n \in [0, \infty)$, $(*)$
and put

$$L_s^+(f) := \sum_{i=1}^{n} \alpha_i \mu(A_i) \in [0, \infty].$$

Well-definedness of the simple integral formula (Lemma 5.5) assures us that the procedure $(*)$ is coherent, in the sense that the proposed quantity $L_s^+(f)$ only depends on $f$ and not on the choice of how we write $f$ as a linear combination of indicators.

With $(*)$, we have thus defined a map

$$L_s^+ : \mathrm{Bor}_s^+(X, \mathbb{R}) \longrightarrow [0, \infty].$$

We must now check that this map has the properties (i), (ii), (iii) indicated in Integral of non-negative simple functions (Theorem 5.2). □

## 5.2 The map $L^+ : \mathrm{Bor}^+(X, \mathbb{R}) \longrightarrow [0, \infty]$

**Definition 5.6** **Integral of non-negative Borel functions**

For every function $u \in \mathrm{Bor}^+(X, \mathbb{R})$, we define

$$L^+(u) := \sup\{L_s^+(f) \mid f \in \mathrm{Bor}_s^+(X, \mathbb{R}), f \leq u\} \in [0, \infty].$$

**Remark 5.7** **Comments on the formula for $L^+$**

- We use letters such as $u, v, \ldots$ for general functions in $\mathrm{Bor}^+(X, \mathbb{R})$, reserving $f, g, \ldots$ for functions in $\mathrm{Bor}_s^+(X, \mathbb{R})$.
- In the formula above, $f \leq u$ means $f(x) \leq u(x)$ for every $x \in X$.
- The set over which we take the supremum is non-empty because it contains $\underline{0}$.
- The supremum is taken in $[0, \infty]$, so it is allowed to be $\infty$.

**Proposition 5.8** **$L^+$ extends $L_s^+$**

$L^+$ is an extension of $L_s^+$. That is, if $u \in \mathrm{Bor}_s^+(X, \mathbb{R})$, then the newly defined quantity $L^+(u)$ is the same as $L_s^+(u)$.

***Proof:*** Must show

$$\sup\{L_s^+(f) \mid f \in \mathrm{Bor}_s^+(X, \mathbb{R}), f \leq u\} = L_s^+(u)$$

"$\geq$" holds because $L_s^+$ is an element of $\{L_s^+(f) \mid f \in \mathrm{Bor}_s^+(X, \mathbb{R}), f \leq u\}$. For "$\leq$", we must show $L_s^+(u)$ is an upper bound for $\{L_s^+(f) \mid f \in \mathrm{Bor}_s^+(X, \mathbb{R}), f \leq u\}$. Indeed, if $f \in \mathrm{Bor}_s^+(X, \mathbb{R})$ with $f \leq u$, then, using Integral of non-negative simple functions (Theorem 5.2),

$$L_s^+(u) = L_s^+(f + (u - f)) = L_s^+(f) + L_s^+(u - f) \geq L_s^+(f)$$

□

**Proposition 5.9** **Monotonicity and homogeneity of $L^+$**

1. $L^+$ is an increasing map: if $u, v \in \mathrm{Bor}^+(X, \mathbb{R})$ with $u \leq v$, then $L^+(u) \leq L^+(v)$
2. $L^+$ is homogeneous: if $u \in \mathrm{Bor}^+(X, \mathbb{R})$, $\alpha \in [0, \infty)$, then $L^+(\alpha u) = \alpha L^+(u)$

***Proof of 1:*** Since $u \leq v$, we have

$$\{f \in \mathrm{Bor}_s^+(X, \mathbb{R}) \mid f \le u\} \subseteq \{f \in \mathrm{Bor}_s^+(X, \mathbb{R}) \mid f \le v\}.$$

Thus

$$\begin{aligned} L^+(u) &= \sup\{L_s^+(f) \mid f \in \mathrm{Bor}_s^+(X, \mathbb{R}), f \le u\} \\ &\le \sup\{L_s^+(f) \mid f \in \mathrm{Bor}_s^+(X, \mathbb{R}), f \le v\} \\ &= L^+(v). \end{aligned}$$

□

***Proof of 2:*** If $\alpha = 0$, then $L^+(\alpha u) = 0 = \alpha L^+(u)$, where we use the convention $0 \cdot \infty = 0$.

Suppose $\alpha > 0$. Then

$$\{g \in \mathrm{Bor}_s^+(X, \mathbb{R}) \mid g \le \alpha u\} = \{\alpha f \mid f \in \mathrm{Bor}_s^+(X, \mathbb{R}), f \le u\}.$$

Thus,

$$\begin{aligned} L^+(\alpha u) &= \sup\{L_s^+(\alpha f) \mid f \in \mathrm{Bor}_s^+(X, \mathbb{R}), f \le u\} \\ &= \sup\{\alpha L_s^+(f) \mid f \in \mathrm{Bor}_s^+(X, \mathbb{R}), f \le u\} \\ &= \alpha \sup\{L_s^+(f) \mid f \in \mathrm{Bor}_s^+(X, \mathbb{R}), f \le u\} \\ &= \alpha L^+(u), \end{aligned}$$

where the second equality uses Integral of non-negative simple functions (Theorem 5.2). □

**Remark 5.10** **Additivity and MCT**

$L^+$ is additive: $L^+(u + v) = L^+(u) + L^+(v)$ for every $u, v \in \mathrm{Bor}^+(X, \mathbb{R})$. We can show $L^+(u + v) \ge L^+(u) + L^+(v)$ by an easy argument using the definition, but the other inequality is not so easy. Instead we move to the Monotone Convergence Theorem (MCT), which will prove it. Note we have to make sure not to use additivity in the proof of MCT.

# 6 The Monotone Convergence Theorem

Fix a measure space $(X, \mathfrak{M}, \mu)$ from chapter 5. We will work with the map $L^+ : \mathrm{Bor}^+(X, \mathbb{R}) \to [0, \infty]$.

**Notation 6.1** **Increasing convergence**

For $u$ and $u_1, u_2, \ldots \in \mathrm{Bor}^+(X, \mathbb{R})$, we will use the shorthand $u_n \nearrow u$ to mean that $u_1 \le u_2 \le \cdots$ and $u_n \to u$ pointwise.

**Theorem 6.2** **Monotone Convergence Theorem**

Let $u, u_1, u_2, \ldots \in \mathrm{Bor}^+(X, \mathbb{R})$ with $u_n \nearrow u$. Then

$$\lim_{n \to \infty} L^+(u_n) = L^+(u)$$

(an increasing limit in $[0, \infty]$)

**Proposition 6.3** **The missing inequality for additivity of $L^+$**

Let $u, v \in \mathrm{Bor}^+(X, \mathbb{R})$. Then

$$L^+(u+v) \leq L^+(u) + L^+(v).$$

***Proof:*** By Binning approximation (Proposition 4.7), choose sequences $(f_n)_{n=1}^\infty$ and $(g_n)_{n=1}^\infty$ in $\mathrm{Bor}_s^+(X, \mathbb{R})$ such that $f_n \nearrow u$ and $g_n \nearrow v$. It follows that $(f_n + g_n)_{n=1}^\infty$ is a sequence in $\mathrm{Bor}_s^+(X, \mathbb{R})$ such that $f_n + g_n \nearrow u + v$.

Thus,

$$\begin{aligned} L^+(u+v) &= \lim_{n\to\infty} L^+(f_n + g_n) \\ &= \lim_{n\to\infty} L_s^+(f_n + g_n) \\ &= \lim_{n\to\infty} (L_s^+(f_n) + L_s^+(g_n)) \\ &= \lim_{n\to\infty} (L^+(f_n) + L^+(g_n)). \end{aligned}$$

The first equality uses Monotone Convergence Theorem (Theorem 6.2), the second and fourth use $L^+$ extends $L_s^+$ (Proposition 5.8), and the third uses Integral of non-negative simple functions (Theorem 5.2).

Now for every $n \in \mathbb{N}$, since $f_n \leq u$ and $g_n \leq v$, Monotonicity and homogeneity of $L^+$ (Proposition 5.9) gives

$$L^+(f_n) + L^+(g_n) \leq L^+(u) + L^+(v).$$

Hence

$$L^+(u+v) \leq L^+(u) + L^+(v),$$

as desired. □

## 6.1 Proof of the Monotone Convergence Theorem

**Remark 6.4**

If $u_n \nearrow u$ in $\mathrm{Bor}^+(X, \mathbb{R})$, then

$$\Lambda := \lim_{n\to\infty} L^+(u_n) = \sup\{L^+(u_n) \mid n \in \mathbb{N}\} \qquad \text{(point-to-prove)}$$

will satisfy the inequality $\Lambda \geq L^+(u)$.

To see that this is enough to prove Monotone Convergence Theorem (Theorem 6.2), suppose $u_n \nearrow u$ in $\mathrm{Bor}^+(X, \mathbb{R})$. Since $u_n \leq u$ for every $n \in \mathbb{N}$, Monotonicity and homogeneity of $L^+$ (Proposition 5.9) gives

$$L^+(u_n) \leq L^+(u)$$

for every $n \in \mathbb{N}$. Hence $L^+(u)$ is an upper bound for $\{L^+(u_n) \mid n \in \mathbb{N}\}$, so

$$\Lambda = \sup\{L^+(u_n) \mid n \in \mathbb{N}\} \leq L^+(u).$$

Thus, if we prove (point-to-prove), then $\Lambda = L^+(u)$, which is exactly the conclusion of Monotone Convergence Theorem (Theorem 6.2).

We will obtain (point-to-prove) in Proposition 6.7 below. Towards that end, we will prove two special cases of (point-to-prove), which are covered by Lemma 6.5 and Lemma 6.6.

**Lemma 6.5**

Let $u_1 \leq u_2 \leq \cdots \leq u_n \leq \cdots$ be in $\mathrm{Bor}^+(X, \mathbb{R})$ and let $A \in \mathfrak{M}$ be such that $u_n \nearrow \chi_A$. Consider the increasing limit

$$\Lambda := \lim_{n\to\infty} L^+(u_n) \in [0, \infty].$$

Then $\Lambda \geq \mu(A)$, where we note that $\mu(A) = L^+(\chi_A)$.

***Proof:*** Assume $A \neq \varnothing$, since if $A = \varnothing$, then $\mu(A) = 0$ and there is nothing to prove. Fix for the moment a $\theta \in (0, 1)$. For every $n \in \mathbb{N}$, define

$$A_n := \{x \in A \mid u_n(x) > \theta\}$$

<u>Observe:</u>

1. $A_n \in \mathfrak{M}$ for all $n \in \mathbb{N}$. This is because $A_n = A \cap u_n^{-1}((\theta, \infty))$ with $A, u_n^{-1}((\theta, \infty)) \in \mathfrak{M}$.
2. $A_n \subseteq A_{n+1}$ for all $n \in \mathbb{N}$. Indeed, if $x \in A_n$, then $u_n(x) > \theta$, and since $u_n \leq u_{n+1}$, we get $u_{n+1}(x) > \theta$, so $x \in A_{n+1}$.
3. $\bigcup_{n=1}^{\infty} A_n = A$. Indeed, $A_n \subseteq A$ for all $n \in \mathbb{N}$, hence $\bigcup_{n=1}^{\infty} A_n \subseteq A$. For "$\supseteq$", pick $x \in A$. Since $u_n \nearrow \chi_A$, we have $u_n(x) \to \chi_A(x) = 1$. Since $\theta < 1$, there exists $n \in \mathbb{N}$ such that $u_n(x) > \theta$, hence $x \in A_n$.

Thus $(A_n)_{n=1}^{\infty}$ is an increasing chain of sets in $\mathfrak{M}$ with $\bigcup_{n=1}^{\infty} A_n = A$. Using Continuity along chains (Proposition 2.12), we have

$$\lim_{n\to\infty} \mu(A_n) = \mu(A) \tag{$*$}$$

Now observe that for every $n \in \mathbb{N}$ we have $u_n \geq \theta\chi_{A_n}$. Indeed, if $x \notin A_n$, then $\theta\chi_{A_n}(x) = 0 \leq u_n(x)$, while if $x \in A_n$, then $\theta\chi_{A_n}(x) = \theta < u_n(x)$.

Hence by Monotonicity and homogeneity of $L^+$ (Proposition 5.9), $L^+$ extends $L_s^+$ (Proposition 5.8), and Integral of non-negative simple functions (Theorem 5.2),

$$L^+(u_n) \geq L^+\left(\theta\chi_{A_n}\right) = \theta L^+\left(\chi_{A_n}\right) = \theta L_s^+\left(\chi_{A_n}\right) = \theta\mu(A_n).$$

So we have

$$\text{For every } n \in \mathbb{N},\ L^+(u_n) \geq \theta\mu(A_n) \tag{$**$}$$

Letting $n \to \infty$ in $(**)$, we get

$$\Lambda \geq \theta\mu(A)$$

where we use $(*)$ for the right-hand side. Finally, unfix $\theta$. We have $\Lambda \geq \theta\mu(A)$ for every $\theta \in (0,1)$, and hence $\Lambda \geq \mu(A)$, as desired. □

**Lemma 6.6**

Let $u_1 \leq u_2 \leq \cdots \leq u_n \leq \cdots$ be in $\mathrm{Bor}^+(X,\mathbb{R})$ and let $f_0 \in \mathrm{Bor}_s^+(X,\mathbb{R})$ be such that $u_n \nearrow f_0$. Consider the increasing limit

$$\Lambda := \lim_{n\to\infty} L^+(u_n) \in [0,\infty].$$

Then $\Lambda \geq L_s^+(f_0)$, where we note that $L_s^+(f_0) = L^+(f_0)$.

**Proposition 6.7**

Let $u_n \nearrow u$ in $\mathrm{Bor}^+(X,\mathbb{R})$, and let us consider the increasing limit

$$\Lambda := \lim_{n\to\infty} L^+(u_n) \in [0,\infty].$$

Then $\Lambda \geq L^+(u)$.

**Theorem 6.8** **Characterization of $L^+$**

The map $L^+ : \mathrm{Bor}^+(X,\mathbb{R}) \to [0,\infty]$ has the following properties, which determine it uniquely:

(i) $L^+(\chi_A) = \mu(A)$ for every $A \in \mathfrak{M}$.
(ii) $L^+(u+v) = L^+(u) + L^+(v)$ for every $u, v \in \mathrm{Bor}^+(X,\mathbb{R})$.
(iii) $L^+(\alpha u) = \alpha \cdot L^+(u)$ for every $u \in \mathrm{Bor}^+(X,\mathbb{R})$ and $\alpha \in [0,\infty)$.
(iv) $\lim_{n\to\infty} L^+(u_n) = L^+(u)$ whenever $u$ and $(u_n)_{n=1}^{\infty}$ are in $\mathrm{Bor}^+(X,\mathbb{R})$ and $u_n \nearrow u$.

**Note**

For the existence part in Characterization of $L^+$ (Theorem 6.8), property (iv) is Monotone Convergence Theorem (Theorem 6.2), while properties (i), (ii), (iii) follow from Integral of non-negative simple functions (Theorem 5.2), combined with the fact that $L^+$ is an extension of the map $L_s^+ : \mathrm{Bor}_s^+(X, \mathbb{R}) \to [0, \infty]$.

For the uniqueness part in Characterization of $L^+$ (Theorem 6.8), it is convenient to record the following lemma, which will also turn out to be useful in subsequent lectures.

**Lemma 6.9** — **Method of the friendly functions**

Let $(X, \mathfrak{M})$ be a measurable space and consider a set of functions $\mathcal{G} \subseteq \mathrm{Bor}^+(X, \mathbb{R})$ which has the properties listed below.

(j) $\chi_A \in \mathcal{G}$ for every $A \in \mathfrak{M}$.
(jj) If $g_1, g_2 \in \mathcal{G}$, then $g_1 + g_2 \in \mathcal{G}$.
(jjj) If $g \in \mathcal{G}$ and $\alpha \in [0, \infty)$, then $\alpha g \in \mathcal{G}$.
(jv) If $g \in \mathrm{Bor}^+(X, \mathbb{R})$, $g_1 \leq g_2 \leq \cdots \leq g_n \leq \cdots$ in $\mathcal{G}$, and $g_n \nearrow g$, then $g \in \mathcal{G}$.

Under these hypotheses, it follows that $\mathcal{G} = \mathrm{Bor}^+(X, \mathbb{R})$.

***Proof:***

- First, every $f \in \mathrm{Bor}_s^+(X, \mathbb{R})$ belongs to $\mathcal{G}$. Indeed, by Simple functions and measurable divisions (Lemma 5.4), we can write

$$f = \alpha_1 \chi_{A_1} + \cdots + \alpha_n \chi_{A_n}$$

  with $A_1, \ldots, A_n \in \mathfrak{M}$ and $\alpha_1, \ldots, \alpha_n \in [0, \infty)$. By properties (j), (jj), and (jjj), it follows that $f \in \mathcal{G}$.
- Now let $u \in \mathrm{Bor}^+(X, \mathbb{R})$. By Binning approximation (Proposition 4.7), there exists a sequence $(f_n)_{n=1}^{\infty}$ in $\mathrm{Bor}_s^+(X, \mathbb{R})$ such that $f_n \nearrow u$. By the previous bullet, $f_n \in \mathcal{G}$ for every $n \in \mathbb{N}$, and hence property (jv) gives $u \in \mathcal{G}$.

$\square$

***Proof of uniqueness in Characterization of $L^+$ (Theorem 6.8):*** Suppose we are given a map

$$M^+ : \mathrm{Bor}^+(X, \mathbb{R}) \to [0, \infty]$$

which also satisfies the conditions (i)-(iv) in Characterization of $L^+$ (Theorem 6.8). Let

$$\mathcal{G} = \{g \in \mathrm{Bor}^+(X, \mathbb{R}) \mid L^+(g) = M^+(g)\} \subseteq \mathrm{Bor}^+(X, \mathbb{R})$$

We will show that $\mathcal{G} = \mathrm{Bor}^+(X, \mathbb{R})$ by checking the hypotheses of Method of the friendly functions (Lemma 6.9).

- If $A \in \mathfrak{M}$, then $\chi_A \in \mathcal{G}$ because

$$L^+(\chi_A) = \mu(A) = M^+(\chi_A)$$

- If $g_1, g_2 \in \mathcal{G}$, then $g_1 + g_2 \in \mathcal{G}$ because

$$L^+(g_1 + g_2) = L^+(g_1) + L^+(g_2) = M^+(g_1) + M^+(g_2) = M^+(g_1 + g_2).$$

- If $g \in \mathcal{G}$ and $\alpha \in [0, \infty)$, then $\alpha g \in \mathcal{G}$ because
$$L^+(\alpha g) = \alpha L^+(g) = \alpha M^+(g) = M^+(\alpha g).$$
- Suppose $g \in \operatorname{Bor}^+(X, \mathbb{R})$, $g_1 \leq g_2 \leq \cdots \leq g_n \leq \cdots$ in $\mathcal{G}$, and $g_n \nearrow g$. Then
$$L^+(g) = \lim_{n\to\infty} L^+(g_n) = \lim_{n\to\infty} M^+(g_n) = M^+(g),$$
so $g \in \mathcal{G}$.

Thus $\mathcal{G} = \operatorname{Bor}^+(X, \mathbb{R})$ by Method of the friendly functions (Lemma 6.9). Hence $L^+(g) = M^+(g)$ for every $g \in \operatorname{Bor}^+(X, \mathbb{R})$, so $L^+$ is uniquely determined. □

# 7 The space of integrable functions $\mathcal{L}^1(\mu)$

## 7.1 Notion of integrable function, the map $L : \mathcal{L}^1(\mu) \to \mathbb{R}$

**Notation and Remark 7.1** **Framework**

In this part, we continue to work with the fixed measure space $(X, \mathfrak{M}, \mu)$ from the previous chapter. There we considered
$$\operatorname{Bor}^+(X, \mathbb{R}) = \{f \in \operatorname{Bor}(X, \mathbb{R}) \mid f(x) \geq 0 \ \forall x \in X\}$$
and defined a map $L^+ : \operatorname{Bor}^+(X, \mathbb{R}) \to [0, \infty]$, which is trying to become the integral with respect to $\mu$.

We now want to work with the full space $\operatorname{Bor}(X, \mathbb{R})$. The idea is to reduce a function $f \in \operatorname{Bor}(X, \mathbb{R})$ to a pair of functions $f_+, f_- \in \operatorname{Bor}^+(X, \mathbb{R})$, called the *positive part* and *negative part* of $f$, defined by
$$f_+ := f \vee \underline{0} \quad \text{and} \quad f_- := (-f) \vee \underline{0}.$$
Here $\vee$ denotes the pointwise maximum of two functions. Equivalently, for every $x \in X$,
$$f_+(x) = \begin{cases} f(x) & \text{if } f(x) \geq 0 \\ 0 & \text{if } f(x) < 0 \end{cases} \quad \text{and} \quad f_-(x) = \begin{cases} 0 & \text{if } f(x) \geq 0 \\ -f(x) & \text{if } f(x) < 0 \end{cases}.$$
From these formulas we have
$$f_+ + f_- = |f| \quad \text{and} \quad f_+ - f_- = f.$$
We also have the useful expressions
$$f_+ = \frac{f + |f|}{2} \quad \text{and} \quad f_- = \frac{|f| - f}{2}.$$

**Lemma 7.2**

For $f \in \operatorname{Bor}(X, \mathbb{R})$, we have the equivalence
$$L^+(|f|) < \infty \iff (L^+(f_+) < \infty \text{ and } L^+(f_-) < \infty).$$

***Proof:*** This is immediate from the fact that

$$L^+(|f|) = L^+(f_+) + L^+(f_-),$$

which follows from $|f| = f_+ + f_-$ and the additivity of $L^+$ from Characterization of $L^+$ (Theorem 6.8). □

**Definition 7.3** **Integrable functions and the map $L$**

A function $f : X \to \mathbb{R}$ is said to be *integrable with respect to* $\mu$ if $f \in \mathrm{Bor}(X, \mathbb{R})$ and the equivalent conditions from Lemma 7.2 hold. We denote

$$\mathcal{L}^1(\mu) := \{f : X \to \mathbb{R} \mid f \ \text{ is integrable with respect to } \ \mu\}.$$

We define a map $L : \mathcal{L}^1(\mu) \to \mathbb{R}$ by putting, for every $f \in \mathcal{L}^1(\mu)$,

$$L(f) := L^+(f_+) - L^+(f_-).$$

**Note**

We will prove in Linearity of $L$ (Theorem 7.5) that the function $L$ is linear. Towards that end, we first prove a lemma which gives some leeway in how $L(f)$ is calculated: instead of having to stick with $f = f_+ - f_-$, the lemma allows us to use a writing $f = h_1 - h_2$.

**Lemma 7.4**

Let $f \in \mathcal{L}^1(\mu)$ and suppose that we can write $f = h_1 - h_2$ with $h_1, h_2 \in \mathrm{Bor}^+(X, \mathbb{R})$ such that $L^+(h_1), L^+(h_2) < \infty$. Then

$$L(f) = L^+(h_1) - L^+(h_2).$$

***Proof:*** Since $f = h_1 - h_2$ and $f = f_+ - f_-$, we have

$$h_1 + f_- = h_2 + f_+.$$

Applying $L^+$ to both sides, and using additivity from Characterization of $L^+$ (Theorem 6.8), gives

$$L^+(h_1) + L^+(f_-) = L^+(h_2) + L^+(f_+).$$

All four quantities appearing here are finite, so we can rearrange this equality to obtain

$$L^+(f_+) - L^+(f_-) = L^+(h_1) - L^+(h_2).$$

The left-hand side is $L(f)$ by definition, and the result follows. □

**Theorem 7.5** **Linearity of $L$**

1. $\mathcal{L}^1(\mu)$ is stable under linear combinations, hence is a linear subspace of $\mathrm{Bor}(X, \mathbb{R})$.
2. The map $L : \mathcal{L}^1(\mu) \to \mathbb{R}$ is linear.

***Proof of 1:*** Pick some $f, g \in \mathcal{L}^1(\mu)$ and $\alpha, \beta \in \mathbb{R}$, and form the linear combination

$$h := \alpha f + \beta g \in \mathrm{Bor}(X, \mathbb{R}).$$

We want to prove that $h \in \mathcal{L}^1(\mu)$. Observe that

$$|h| = |\alpha f + \beta g| \leq |\alpha| \cdot |f| + |\beta| \cdot |g|,$$

which implies that

$$\begin{aligned} L^+(|h|) &\leq L^+(|\alpha| \cdot |f| + |\beta| \cdot |g|) \\ &= |\alpha| \cdot L^+(|f|) + |\beta| \cdot L^+(|g|) \\ &< \infty. \end{aligned}$$

The first inequality follows from the fact that $L^+$ is increasing, and the second equality follows from the additivity and homogeneity properties in Characterization of $L^+$ (Theorem 6.8). Since $f, g \in \mathcal{L}^1(\mu)$, the last expression is finite. It follows that $L^+(|h|) < \infty$, so $h \in \mathcal{L}^1(\mu)$ by Lemma 7.2. □

***Proof of 2:*** We will verify separately the two statements which, together, give linearity:

(a) $L(f+g) = L(f) + L(g)$ for every $f, g \in \mathcal{L}^1(\mu)$.
(b) $L(\alpha f) = \alpha L(f)$ for every $f \in \mathcal{L}^1(\mu)$ and $\alpha \in \mathbb{R}$.

*Verification of (a):*
Pick $f, g \in \mathcal{L}^1(\mu)$. By 1, $f + g \in \mathcal{L}^1(\mu)$. We can write

$$f + g = (f_+ + g_+) - (f_- + g_-),$$

where $f_+ + g_+$ and $f_- + g_-$ are in $\mathrm{Bor}^+(X, \mathbb{R})$. Moreover,

$$L^+(f_+ + g_+) = L^+(f_+) + L^+(g_+) < \infty$$

and

$$L^+(f_- + g_-) = L^+(f_-) + L^+(g_-) < \infty,$$

since $f, g \in \mathcal{L}^1(\mu)$. Thus Lemma 7.4 gives

$$\begin{aligned} L(f+g) &= L^+(f_+ + g_+) - L^+(f_- + g_-) \\ &= L^+(f_+) + L^+(g_+) - L^+(f_-) - L^+(g_-) \\ &= L(f) + L(g). \end{aligned}$$

*Verification of (b):*
Pick $f \in \mathcal{L}^1(\mu)$ and $\alpha \in \mathbb{R}$. By 1, $\alpha f \in \mathcal{L}^1(\mu)$.

If $\alpha \geq 0$, then

$$\alpha f = \alpha f_+ - \alpha f_-,$$

and $L^+(\alpha f_+) = \alpha L^+(f_+) < \infty$, $L^+(\alpha f_-) = \alpha L^+(f_-) < \infty$. Hence Lemma 7.4 gives

$$L(\alpha f) = \alpha L^+(f_+) - \alpha L^+(f_-) = \alpha L(f).$$

If $\alpha < 0$, then

$$\alpha f = (-\alpha) f_- - (-\alpha) f_+,$$

and $L^+((-\alpha)f_-) = (-\alpha)L^+(f_-) < \infty$, $L^+((-\alpha)f_+) = (-\alpha)L^+(f_+) < \infty$. Hence Lemma 7.4 gives

$$L(\alpha f) = (-\alpha)L^+(f_-) - (-\alpha)L^+(f_+) = \alpha L(f).$$

□

**Proposition 7.6**

1. If $f \in \mathcal{L}^1(\mu)$ and also $f \in \mathrm{Bor}^+(X, \mathbb{R})$, then $L(f) = L^+(f)$.
2. If $f, g \in \mathcal{L}^1(\mu)$ are such that $f \leq g$, then $L(f) \leq L(g)$.
3. If $f \in \mathcal{L}^1(\mu)$, then $|f| \in \mathcal{L}^1(\mu)$ as well, and

$$|L(f)| \leq L(|f|).$$

***Proof of 1:*** Since $f \in \mathrm{Bor}^+(X, \mathbb{R})$, we have $f_+ = f$ and $f_- = \underline{0}$. Hence

$$L(f) = L^+(f_+) - L^+(f_-) = L^+(f) - L^+(\underline{0}) = L^+(f).$$

□

***Proof of 2:*** If $f \leq g$, then $g - f \in \mathcal{L}^1(\mu)$ by Linearity of $L$ (Theorem 7.5), and moreover $g - f \in \mathrm{Bor}^+(X, \mathbb{R})$. By part 1,

$$L(g - f) = L^+(g - f) \geq 0.$$

Using the linearity of $L$, this says that $L(g) - L(f) \geq 0$, hence $L(f) \leq L(g)$. □

***Proof of 3:*** Since $f \in \mathcal{L}^1(\mu)$, Lemma 7.2 gives $L^+(|f|) < \infty$, so $|f| \in \mathcal{L}^1(\mu)$.

As functions, we have

$$-|f| \leq f \leq |f|.$$

Applying part 2 gives

$$-L(|f|) \leq L(f) \leq L(|f|),$$

where we also used the linearity of $L$ on the left. Thus

$$|L(f)| \leq L(|f|).$$

□

**Notation and Remark 7.7** **Switching to integral notation**

For our measure space $(X, \mathfrak{M}, \mu)$, we have now constructed two maps

$$L^+ : \mathrm{Bor}^+(X, \mathbb{R}) \to [0, \infty] \quad \text{and} \quad L : \mathcal{L}^1(\mu) \to \mathbb{R},$$

which are trying to become "the integral with respect to $\mu$". If $f \in \mathrm{Bor}(X, \mathbb{R})$ is such that both values $L^+(f)$ and $L(f)$ are defined, then Proposition 7.6 says that $L(f) = L^+(f)$.

It is therefore safe to switch to the customary integral notation:

$$\text{if at least one of } \; L^+(f), L(f) \; \text{ is defined, we denote that value by} \\ \int_X f(x)\,\mathrm{d}\mu(x).$$

This is often abbreviated to just $\int f\,\mathrm{d}\mu$.

While we now have uniform notation for all our integrals, it is useful to remember that the integral notation can have two meanings:
- an $L^+$ meaning, when it has values in $[0, \infty]$ and we use the properties of $L^+$;
- an $L$ meaning, when it has values in $\mathbb{R}$ and we use the properties of $L$.

This usually causes no harm. For instance, the inequality from Proposition 7.6 becomes

$$\left| \int_X f(x)\,\mathrm{d}\mu(x) \right| \le \int_X |f(x)|\,\mathrm{d}\mu(x), \quad \forall f \in \mathcal{L}^1(\mu),$$

where the integral on the right-hand side can be viewed either in the $L^+$ sense or in the $L$ sense.

## 7.2 Integration of subsets

**Notation 7.8**

Let $(X, \mathfrak{M}, \mu)$ be a measure space.

1. For any $f \in \mathrm{Bor}^+(X, \mathbb{R})$ and $A \in \mathfrak{M}$, we denote

$$\int_A f\,\mathrm{d}\mu := \int_X f \cdot \chi_A\,\mathrm{d}\mu \in [0, \infty],$$

   where, as usual, $\chi_A$ denotes the indicator function of $A$.
2. For any $f \in \mathcal{L}^1(\mu)$ and $A \in \mathfrak{M}$, observe that $f \cdot \chi_A \in \mathcal{L}^1(\mu)$, since $|f \cdot \chi_A| \le |f|$. We denote

$$\int_A f\,\mathrm{d}\mu := \int_X f \cdot \chi_A\,\mathrm{d}\mu \in \mathbb{R}.$$

**Proposition 7.9**

Let $(X, \mathfrak{M}, \mu)$ be a measure space and let $h \in \mathrm{Bor}^+(X, \mathbb{R})$. We define a set-function $\nu : \mathfrak{M} \to [0, \infty]$ by putting

$$\nu(A) := \int_A h \, \mathrm{d}\mu, \quad A \in \mathfrak{M}.$$

Then $\nu$ is a positive measure on $\mathfrak{M}$.

***Proof:***

1. $$\nu(\varnothing) = \int_X h \cdot \chi_{\varnothing} \, \mathrm{d}\mu = \int_X \underline{0} \, \mathrm{d}\mu = 0.$$
2. Fix $(A_n)_{n=1}^{\infty}$ in $\mathfrak{M}$ such that $A_i \cap A_j = \varnothing$ for $i \neq j$, and let $U = \bigcup_{n=1}^{\infty} A_n$. We must prove that

$$\nu(U) = \sum_{n=1}^{\infty} \nu(A_n).$$

For every $N \in \mathbb{N}$, let

$$s_N = \sum_{n=1}^{N} h \cdot \chi_{A_n}.$$

Then $(s_N)_{N=1}^{\infty}$ is an increasing sequence in $\mathrm{Bor}^+(X, \mathbb{R})$, and

$$s_N \to h \cdot \chi_U$$

pointwise. Hence, by Monotone Convergence Theorem (Theorem 6.2),

$$\begin{aligned}
\nu(U) &= \int_X h \cdot \chi_U \, \mathrm{d}\mu \\
&= \lim_{N \to \infty} \int_X s_N \, \mathrm{d}\mu \\
&= \lim_{N \to \infty} \sum_{n=1}^{N} \int_X h \chi_{A_n} \, \mathrm{d}\mu \\
&= \sum_{n=1}^{\infty} \nu(A_n).
\end{aligned}$$

□

**Proposition 7.10**

Let $(X, \mathfrak{M}, \mu)$, $h \in \operatorname{Bor}^+(X, \mathbb{R})$, and the positive measure $\nu : \mathfrak{M} \to [0, \infty]$ be as in Proposition 7.9. Then, for every $f \in \operatorname{Bor}^+(X, \mathbb{R})$,

$$\int_X f \, \mathrm{d}\nu = \int_X f \cdot h \, \mathrm{d}\mu \in [0, \infty].$$

This is an equality of integrals considered in the $L^+$ sense.

***Proof:*** We will use the method of the friendly functions (Method of the friendly functions (Lemma 6.9)). Let

$$\mathcal{G} = \left\{ g \in \operatorname{Bor}^+(X, \mathbb{R}) \;\middle|\; \int_X g \, \mathrm{d}\nu = \int_X g \cdot h \, \mathrm{d}\mu \right\}$$

We verify that $\mathcal{G}$ satisfies the conditions of Method of the friendly functions (Lemma 6.9).

- For $A \in \mathfrak{M}$, we have

  $$\int_X \chi_A \, \mathrm{d}\nu = \nu(A) = \int_A h \, \mathrm{d}\mu = \int_X \chi_A \cdot h \, \mathrm{d}\mu,$$

  so $\chi_A \in \mathcal{G}$.
- If $g_1, g_2 \in \mathcal{G}$, then

  $$\begin{aligned} \int_X (g_1 + g_2) \, \mathrm{d}\nu &= \int_X g_1 \, \mathrm{d}\nu + \int_X g_2 \, \mathrm{d}\nu \\ &= \int_X g_1 \cdot h \, \mathrm{d}\mu + \int_X g_2 \cdot h \, \mathrm{d}\mu \\ &= \int_X (g_1 + g_2) \cdot h \, \mathrm{d}\mu, \end{aligned}$$

  so $g_1 + g_2 \in \mathcal{G}$.
- If $g \in \mathcal{G}$ and $\alpha \in [0, \infty)$, then

  $$\begin{aligned} \int_X \alpha g \, \mathrm{d}\nu &= \alpha \int_X g \, \mathrm{d}\nu \\ &= \alpha \int_X g \cdot h \, \mathrm{d}\mu \\ &= \int_X (\alpha g) \cdot h \, \mathrm{d}\mu, \end{aligned}$$

  so $\alpha g \in \mathcal{G}$.
- Suppose that $g_1 \leq g_2 \leq \cdots \leq g_n \leq \cdots$ are in $\mathcal{G}$ and $g_n \to g$ for some $g \in \operatorname{Bor}^+(X, \mathbb{R})$. Since $h \geq 0$, we also have $g_n \cdot h \to g \cdot h$ increasingly. By Monotone Convergence Theorem (Theorem 6.2), applied once with respect to $\nu$ and once with respect to $\mu$,

$$\begin{aligned}\int_X g \,\mathrm{d}\nu &= \lim_{n\to\infty} \int_X g_n \,\mathrm{d}\nu \\ &= \lim_{n\to\infty} \int_X g_n \cdot h \,\mathrm{d}\mu \\ &= \int_X g \cdot h \,\mathrm{d}\mu.\end{aligned}$$

Thus $g \in \mathcal{G}$.

Then Method of the friendly functions (Lemma 6.9) gives $\mathcal{G} = \mathrm{Bor}^+(X, \mathbb{R})$, which means that

$$\int_X g \,\mathrm{d}\nu = \int_X g \cdot h \,\mathrm{d}\mu$$

for all $g \in \mathrm{Bor}^+(X, \mathbb{R})$, as required. □

**Note**

In connection with Proposition 7.10, it is customary to write the connection between the positive measures $\mu$ and $\nu$ in the form

$$\mathrm{d}\nu = h \,\mathrm{d}\mu.$$

This can be thought of as cancelling the common integral expression in

$$\int_X f \,\mathrm{d}\nu = \int_X f \cdot h \,\mathrm{d}\mu.$$

Some further abuse of notation rewrites this relation as

$$h = \frac{\mathrm{d}\nu}{\mathrm{d}\mu},$$

where one thinks of $h$ as a kind of "derivative" of $\nu$ with respect to $\mu$.

# 8 Lebesgue's Dominated Convergence Theorem (LDCT)

Fix a measure space $(X, \mathfrak{M}, \mu)$.

**Theorem 8.1** **LDCT**

Let $f$ and $(f_n)_{n=1}^{\infty}$ be functions in $\mathrm{Bor}(X, \mathbb{R})$ such that

$$\lim_{n\to\infty} f_n(x) = f(x), \quad \forall x \in X.$$

Suppose moreover that there exists a function $h \in \mathcal{L}^1(\mu) \cap \mathrm{Bor}^+(X, \mathbb{R})$ which dominates all the $f_n$, in the sense that

$$|f_n| \leq h, \quad \forall n \in \mathbb{N}.$$

Then $f$ and $f_1, f_2, ..., f_n, ...$ are all in $\mathcal{L}^1(\mu)$, and

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

**Note**

The last equation can be written as

$$\lim_{n\to\infty} \left( \int_X f_n \,\mathrm{d}\mu \right) = \int_X \left( \lim_{n\to\infty} f_n \right) \mathrm{d}\mu,$$

so this is a formula where one interchanges integration with a limit. In principle, this interchange can have issues. The point of LDCT is that the existence of the dominating function $h$ is a useful sufficient condition which makes the issues go away.

It is convenient to think of LDCT as going in parallel with the MCT from Monotone Convergence Theorem (Theorem 6.2). In connection to that, we will observe in Reverse-MCT (Proposition 8.2) that a special case of LDCT gives a "Reverse-MCT" statement.

**Proposition 8.2** **Reverse-MCT**

Let $u$ and $(u_n)_{n=1}^{\infty}$ be functions in $\mathrm{Bor}^+(X, \mathbb{R})$ such that

$$u_1 \geq u_2 \geq \cdots \geq u_n \geq \cdots$$

and such that

$$\lim_{n\to\infty} u_n(x) = u(x), \quad \forall x \in X.$$

(compactly written $u_n \searrow u$). We assume, in addition, that $\int_X u_1 \,\mathrm{d}\mu < \infty$. Then

$$\lim_{n\to\infty} \int_X u_n \,\mathrm{d}\mu = \int_X u \,\mathrm{d}\mu.$$

**Note**

Reverse-MCT is a special case of LDCT because we can use $h = u_1$ as the dominating function. Then LDCT (Theorem 8.1) gives the required conclusion.

For the plan of how we prove LDCT, it is convenient to also consider its special case $\mathrm{LDCT}_0$, which considers a dominated sequence in $\mathrm{Bor}^+(X, \mathbb{R})$ converging pointwise to 0. This is $\mathrm{LDCT}_0$ (Proposition 8.3).

**Proposition 8.3** — $\mathrm{LDCT}_0$

Let $(g_n)_{n=1}^\infty$ be a sequence of functions in $\mathrm{Bor}^+(X, \mathbb{R})$ such that

$$\lim_{n\to\infty} g_n(x) = 0, \quad \forall x \in X.$$

Suppose moreover that there exists a function $h \in \mathcal{L}^1(\mu) \cap \mathrm{Bor}^+(X, \mathbb{R})$ such that

$$g_n \le h, \quad \forall n \in \mathbb{N}.$$

Then

$$\lim_{n\to\infty} \int_X g_n \, \mathrm{d}\mu = 0.$$

**Remark 8.4** — **Scheme of proof of LDCT**

We will go in the following way:

$$\mathrm{MCT} \Longrightarrow \text{Reverse-MCT} \Longrightarrow \mathrm{LDCT}_0 \Longrightarrow \mathrm{LDCT}.$$

The first implication is proved by subtracting everything from the top. The second implication uses the trick of Fatou. The third implication is an immediate reduction, by putting $g_n := |f_n - f|$.

***8.5. Proof of the implication (1).*** Define

$$v_n := u_1 - u_n, \quad n \in \mathbb{N}, \quad \text{and} \quad v := u_1 - u.$$

Then $v_n \in \mathrm{Bor}^+(X, \mathbb{R})$ for every $n$, and $v_n \nearrow v$. By Monotone Convergence Theorem (Theorem 6.2),

$$\lim_{n\to\infty} \int_X v_n \, \mathrm{d}\mu = \int_X v \, \mathrm{d}\mu.$$

Since $u_1 = v_n + u_n$ and $u_1 = v + u$, and since $\int_X u_1 \, \mathrm{d}\mu < \infty$, all the quantities below are finite and we can write

$$\begin{aligned}\int_X u_n \,\mathrm{d}\mu &= \int_X u_1 \,\mathrm{d}\mu - \int_X v_n \,\mathrm{d}\mu \\ &\to \int_X u_1 \,\mathrm{d}\mu - \int_X v \,\mathrm{d}\mu \\ &= \int_X u \,\mathrm{d}\mu.\end{aligned}$$

□

***8.6. Proof of the implication (2).*** The idea is to make a helper sequence $(u_n)_{n=1}^{\infty}$ to which we can apply Reverse-MCT (Proposition 8.2). We break the proof into steps:

1. For every $n \in \mathbb{N}$, define $u_n : X \to [0, \infty)$ by

$$u_n(x) := \sup\{g_k(x) \mid k \geq n\}.$$

   Since each $g_k \in \mathrm{Bor}^+(X, \mathbb{R})$, Closure under suprema and infima (Proposition 4.1) gives $u_n \in \mathrm{Bor}(X, \mathbb{R})$. Also $0 \leq u_n \leq h$, so $u_n \in \mathrm{Bor}^+(X, \mathbb{R})$. In particular,

$$\int_X u_1 \,\mathrm{d}\mu \leq \int_X h \,\mathrm{d}\mu < \infty,$$

   where the first inequality uses the monotonicity of $L^+$ from Characterization of $L^+$ (Theorem 6.8), and the finiteness of the last term follows from Proposition 7.6. Finally, from the definition of the tail suprema we have

$$u_1 \geq u_2 \geq \cdots \geq u_n \geq \cdots.$$

2. We claim that $u_n \searrow \underline{0}$. Fix $x \in X$ and let $\varepsilon > 0$ be given. Since $g_n(x) \to 0$, there is $n_0 \in \mathbb{N}$ such that $g_k(x) < \varepsilon$ for all $k \geq n_0$. For this $n_0$, we have

$$u_{n_0}(x) = \sup\{g_k(x) \mid k \geq n_0\} \leq \varepsilon.$$

   Since the sequence $(u_n(x))_{n=1}^{\infty}$ is decreasing, it follows that

$$0 \leq u_n(x) \leq u_{n_0}(x) \leq \varepsilon, \quad \forall n \geq n_0.$$

   Thus $u_n(x) \to 0$ for every $x \in X$.
3. Now Reverse-MCT (Proposition 8.2) applied to $(u_n)_{n=1}^{\infty}$ gives

$$\lim_{n\to\infty} \int_X u_n \,\mathrm{d}\mu = 0.$$

   Since $g_n \leq u_n$ for every $n \in \mathbb{N}$, the monotonicity of $L^+$ gives

$$0 \leq \int_X g_n \,\mathrm{d}\mu \leq \int_X u_n \,\mathrm{d}\mu, \quad \forall n \in \mathbb{N}.$$

   Hence $\int_X g_n \,\mathrm{d}\mu \to 0$ as required.

□

***8.7. Proof of the implication (3).*** Have $f$ and $(f_n)_{n=1}^{\infty}$ in $\mathrm{Bor}(X, \mathbb{R})$ such that $f_n(x) \xrightarrow[n\to\infty]{} f(x)$ for all $x \in X$. Also, we have $h \in \mathrm{Bor}^+(X, \mathbb{R}) \cap \mathcal{L}^1(\mu)$ such that $|f_n| \leq h$ for all $n \in \mathbb{N}$. We want $f_n \in \mathcal{L}^1(\mu)$ for all $n \in \mathbb{N}$, $f \in \mathcal{L}^1(\mu)$ and

$$\int_X f_n \, \mathrm{d}\mu \xrightarrow[n\to\infty]{} \int_X f \, \mathrm{d}\mu$$

- For every $n \in \mathbb{N}$, we have

$$\begin{aligned} |f_n| \leq h &\implies \int_X |f_n| \, \mathrm{d}\mu \leq \int_X h \, \mathrm{d}\mu < \infty \\ &\implies f_n \in \mathcal{L}^1(\mu) \end{aligned}$$

  How about $f$? Observe that $|f| \leq h$, which holds because

$$|f(x)| = \lim_{n\to\infty} |f_n(x)| \leq h(x) \quad \forall x \in X$$

  Thus

$$\int_X |f| \, \mathrm{d}\mu \leq \int_X h \, \mathrm{d}\mu < \infty$$

  hence $f \in \mathcal{L}^1(\mu)$ as well. It remains to check that

$$\int_X f_n \, \mathrm{d}\mu \xrightarrow[n\to\infty]{} \int_X f \, \mathrm{d}\mu$$

  Towards that end, consider

$$g_n := |f_n - f| \in \mathrm{Bor}^+(X, \mathbb{R}) \quad \forall n \in \mathbb{N}$$

  ***<u>Claim:</u>*** The $g_n$'s satisfy the hypothesis of $\mathrm{LDCT}_0$ (Proposition 8.3). That is, we have

  (i) $$g_n(x) \xrightarrow[n\to\infty]{} 0 \quad \forall x \in X$$

  (ii) There is $\tilde{h} \in \mathrm{Bor}^+(X, \mathbb{R}) \cap \mathcal{L}^1(\mu)$ such that $g_n \leq \tilde{h}$ for all $n \in \mathbb{N}$

  ***<u>Verification of Claim:</u>***

  (i) For every $x \in X$ we have

$$g_n(x) = |f_n(x) - f(x)| \xrightarrow[n\to\infty]{} 0$$

  due to the hypothesis that $f_n(x) \xrightarrow[n\to\infty]{} f(x)$.

  (ii) For every $n \in \mathbb{N}$ and $x \in X$ we have

$$g_n(x) = |f_n(x) - f(x)| \leq |f_n(x)| + |f(x)| \leq 2h(x)$$

  Hence $g_n \leq 2h$ for all $n \in \mathbb{N}$, where

$$\int_X 2h \, \mathrm{d}\mu = 2 \int_X h \, \mathrm{d}\mu < \infty$$

  so we can use $\tilde{h} = 2h$

Due to the claim, we can apply LDCT$_0$ (Proposition 8.3) to $(g_n)_{n=1}^{\infty}$ to get that

$$\int_X g_n \,\mathrm{d}\mu \xrightarrow[n\to\infty]{} 0 \tag{$*$}$$

This implies the fact that

$$\int_X f_n \,\mathrm{d}\mu \xrightarrow[n\to\infty]{} \int_X f \,\mathrm{d}\mu$$

because we can write

$$0 \leq \left|\int_X f_n \,\mathrm{d}\mu - \int_X f \,\mathrm{d}\mu\right| = \left|\int_X f_n - f \,\mathrm{d}\mu\right| \leq \int_X |f_n - f| \,\mathrm{d}\mu = \int g_n \,\mathrm{d}\mu \xrightarrow[n\to\infty]{} 0$$

which completes the proof.

□

## 8.1 The "a.e-$\mu$" annoyance (and the "a.e-$\mu$" version of LDCT)

**Note**

A rule of thumb: changes made on a set of measure 0 will not affect integrability or the value of the integral.

**Definition 8.8**

We say that two functions $f, g : X \to \mathbb{R}$ are *equal almost everywhere with respect to* $\mu$, denoted $f = g$ a.e$-\mu$, to mean there exists a set $Z \in \mathfrak{M}$ such that

(i) $\mu(Z) = 0$.

(ii) $f(x) = g(x)$ for all $x \in X \setminus Z$.

The preceding definition makes sense even if $f$ and $g$ are not Borel functions. In the case of interest when $f, g \in \mathrm{Bor}(X, \mathbb{R})$, we can use the following equivalent description of equality almost everywhere.

**Proposition 8.9**

Let $f, g \in \mathrm{Bor}(X, \mathbb{R})$, and denote

$$M := \{x \in X \mid f(x) \neq g(x)\}.$$

1. We have $M \in \mathfrak{M}$, hence the quantity $\mu(M)$ is well-defined.
2. We have

$$(f = g \text{ a.e}-\mu) \iff (\mu(M) = 0).$$

***Proof of 1:*** We can write

$$M = (f - g)^{-1}(\mathbb{R} \setminus \{0\}).$$

Since $f - g \in \mathrm{Bor}(X, \mathbb{R})$ and $\mathbb{R} \setminus \{0\} \in \mathcal{B}_{\mathbb{R}}$, it follows that $M \in \mathfrak{M}$. $\square$

***Proof of 2:*** "$\Longrightarrow$" Suppose $f = g$ a.e$-\mu$. Then there exists $Z \in \mathfrak{M}$ such that $\mu(Z) = 0$ and $f(x) = g(x)$ for every $x \in X \setminus Z$. It follows that $M \subseteq Z$, so $0 \leq \mu(M) \leq \mu(Z) = 0$, hence $\mu(M) = 0$.

"$\Longleftarrow$" Suppose $\mu(M) = 0$. Taking $Z := M$, we have $Z \in \mathfrak{M}$ by part 1, $\mu(Z) = 0$, and $f(x) = g(x)$ for every $x \in X \setminus Z$. Hence $f = g$ a.e$-\mu$. $\square$

**Lemma 8.10**

Let $f \in \mathrm{Bor}^+(X, \mathbb{R})$ be such that $f = \underline{0}$ a.e$-\mu$. Then

$$\int_X f \,\mathrm{d}\mu = 0.$$

In particular, $f \in \mathcal{L}^1(\mu)$.

***Proof:*** For every $n \in \mathbb{N}$, let $f_n = f \wedge \underline{n}$, where $\underline{n} : X \to \mathbb{R}$ is identically equal to $n$. That is,

$$f_n(x) = \begin{cases} f(x) & \text{if } f(x) \leq n \\ n & \text{if } f(x) > n \end{cases}$$

We have $f_n \in \mathrm{Bor}(X, \mathbb{R})$ because $\mathrm{Bor}(X, \mathbb{R})$ is a lattice of functions. It is also easy to see that $f_n \nearrow f$. By Monotone Convergence Theorem (Theorem 6.2) we have

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

So in order to get $\int_X f \,\mathrm{d}\mu = 0$, it suffices to check that $\int_X f_n \,\mathrm{d}\mu = 0$ for every $n \in \mathbb{N}$. Let $Z = \{x \in X \mid f(x) \neq 0\}$. By Proposition 8.9, we have $\mu(Z) = 0$. Now observe that $f_n \leq n\chi_Z$ for all $n \in \mathbb{N}$: for $x \in Z$ use that $f_n(x) \leq n = n\chi_Z(x)$, while for $x \in X \setminus Z$ we have $f_n(x) = 0 = n\chi_Z(x)$. Thus

$$\int_X f_n \,\mathrm{d}\mu \leq \int_X n\chi_Z \,\mathrm{d}\mu = n\mu(Z) = 0,$$

hence $\int_X f_n \,\mathrm{d}\mu = 0$ as desired. $\square$

**Proposition 8.11**

Let $f, g \in \mathrm{Bor}(X, \mathbb{R})$ be such that $f = g$ a.e$-\mu$. Suppose that one of these functions, say $f$, is in $\mathcal{L}^1(\mu)$. Then $g \in \mathcal{L}^1(\mu)$ as well, and

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

***Proof:*** The hypothesis that $f = g$ a.e$-\mu$ is equivalent to the fact that $|f - g| = \underline{0}$ a.e$-\mu$. Hence Lemma 8.10 applies to $|f - g|$, and tells us that

$$|f - g| \in \mathrm{Bor}^+(X, \mathbb{R}) \cap \mathcal{L}^1(\mu)$$

with

$$\int_X |f-g| \,\mathrm{d}\mu = 0.$$

But then $f-g \in \mathcal{L}^1(\mu)$, since its absolute value is there, and it follows that

$$g = f-(f-g) \in \mathcal{L}^1(\mu)$$

as well, due to the fact that $\mathcal{L}^1(\mu)$ is stable under linear combinations.

Finally, for the statement about the values of the integrals, we write

$$\begin{aligned} \left|\int_X f \,\mathrm{d}\mu - \int_X g \,\mathrm{d}\mu\right| &= \left|\int_X f-g \,\mathrm{d}\mu\right| \\ &\leq \int_X |f-g| \,\mathrm{d}\mu \\ &= 0, \end{aligned}$$

and this shows that

$$\int_X f \,\mathrm{d}\mu = \int_X g \,\mathrm{d}\mu,$$

as required. □

> **Remark 8.12**
>
> The converse of Lemma 8.10 is true as well: if $f \in \mathrm{Bor}^+(X,\mathbb{R})$ and $\int_X f \,\mathrm{d}\mu = 0$, then it follows that $f = 0$ a.e$-\mu$.

***Proof:*** Define

$$Z = \{x \in X \mid f(x) \neq 0\} = \{x \in X \mid f(x) > 0\}$$

and for every $n \in \mathbb{N}$ let

$$Z_n = \left\{x \in X \,\middle|\, f(x) \geq \frac{1}{n}\right\}.$$

Note

$$Z_1 \subseteq Z_2 \subseteq \cdots \subseteq Z_n \subseteq \cdots$$

and

$$Z = \bigcup_{n=1}^{\infty} Z_n.$$

By Continuity along chains (Proposition 2.12), we have $\mu(Z) = \lim_{n\to\infty} \mu(Z_n)$. So in order to prove $\mu(Z) = 0$, it suffices to prove $\mu(Z_n) = 0$ for every $n \in \mathbb{N}$. Indeed, for every $n \in \mathbb{N}$, we have $f \geq \frac{1}{n}\chi_{Z_n}$, hence

$$0 = \int_X f \,\mathrm{d}\mu \geq \int_X \frac{1}{n}\chi_{Z_n} \,\mathrm{d}\mu = \frac{1}{n}\mu(Z_n) \geq 0.$$

so $\mu(Z_n) = 0$ as we wanted. □

**Theorem 8.13** a.e$-\mu$-**version of LDCT (Theorem 8.1)**

Let $f$ and $(f_n)_{n=1}^\infty$ be functions in $\mathrm{Bor}(X, \mathbb{R})$ such that the following hold:

(α) there exists a set $Z_1 \in \mathfrak{M}$ with $\mu(Z_1) = 0$ such that

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

for every $x \in X \setminus Z_1$;

(β) there exists a set $Z_2 \in \mathfrak{M}$ with $\mu(Z_2) = 0$ and a function $h \in \mathrm{Bor}^+(X, \mathbb{R}) \cap \mathcal{L}^1(\mu)$ such that

$$|f_n(x)| \leq h(x)$$

for every $n \in \mathbb{N}$ and every $x \in X \setminus Z_2$.

Then $f$ and $f_1, f_2, \ldots, f_n, \ldots$ are all in $\mathcal{L}^1(\mu)$, and

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

***Proof:*** Let $Z = Z_1 \cup Z_2$. Note $Z \in \mathfrak{M}$ and $\mu(Z) = 0$. Also note the hypotheses $(\alpha), (\beta)$ still hold in connection to $Z$.

Idea: let $\tilde{f} = f\chi_{X\setminus Z}$ and for every $n \in \mathbb{N}$ let $\widetilde{f_n} = f_n\chi_{X\setminus Z}$. Observe $\tilde{f}$ and $\left(\tilde{f}_n\right)_{n=1}^\infty$ satisfy the hypotheses of LDCT (Theorem 8.1): on $X \setminus Z$ this follows from $(\alpha), (\beta)$, while on $Z$ all these functions vanish. Hence LDCT (Theorem 8.1) can be applied to conclude that $\tilde{f}$ and $\tilde{f}_1, \tilde{f}_2, \ldots$ are in $\mathcal{L}^1(\mu)$ and

$$\int_X \widetilde{f_n} \,\mathrm{d}\mu \xrightarrow[n\to\infty]{} \int_X \tilde{f} \,\mathrm{d}\mu.$$

On the other hand, observe that $\tilde{f} = f$ a.e$-\mu$ and $\widetilde{f_n} = f_n$ a.e$-\mu$ for every $n \in \mathbb{N}$. Then Proposition 8.11 allows us to conclude that $f \in \mathcal{L}^1(\mu)$, with $\int_X f \,\mathrm{d}\mu = \int_X \tilde{f} \,\mathrm{d}\mu$. Likewise, for every $n \in \mathbb{N}$, $f_n \in \mathcal{L}^1(\mu)$ and $\int_X f_n \,\mathrm{d}\mu = \int_X \widetilde{f_n} \,\mathrm{d}\mu$. Finally,

$$\int_X \widetilde{f_n} \,\mathrm{d}\mu \xrightarrow[n\to\infty]{} \int_X \tilde{f} \,\mathrm{d}\mu.$$

"becomes"

$$\int_X f_n \,\mathrm{d}\mu \xrightarrow[n\to\infty]{} \int_X f \,\mathrm{d}\mu.$$

□

# 9 Basic stuff about the spaces $\mathcal{L}^p(\mu)$

## 9.1 Definition of $\mathcal{L}^p(\mu)$ and $\|\cdot\|_p$

We fix a measure space $(X, \mathfrak{M}, \mu)$.

**Notation and Remark 9.1** **The $\mathcal{L}^1(\mu)$ seminorm**

For every $f \in \mathcal{L}^1(\mu)$ we denote

$$\|f\|_1 := \int_X |f| \, \mathrm{d}\mu \in [0, \infty).$$

The map $f \mapsto \|f\|_1$ is often referred to as $\|\cdot\|_1$.

***Claim:*** $\|\cdot\|_1$ is a *seminorm* on $\mathcal{L}^1(\mu)$. This means that $\|\cdot\|_1$ has the following properties:

1. (Non-Negative) $\|f\|_1 \geq 0$ for every $f \in \mathcal{L}^1(\mu)$.
2. (Sub-Additive) $\|f+g\|_1 \leq \|f\|_1 + \|g\|_1$ for every $f, g \in \mathcal{L}^1(\mu)$.
3. (Homogeneous) $\|\alpha f\|_1 = |\alpha| \cdot \|f\|_1$ for every $f \in \mathcal{L}^1(\mu)$ and $\alpha \in \mathbb{R}$.

***Proof:*** The properties (Non-Negative) and (Homogeneous) are immediate from how $\|\cdot\|_1$ is defined. The verification of (Sub-Additive) also follows easily from the properties of the integral: for every $f, g \in \mathcal{L}^1(\mu)$, we first write

$$|f+g| \leq |f| + |g|$$

as an inequality of functions in $\mathrm{Bor}^+(X, \mathbb{R})$, and then we integrate this inequality to get

$$\begin{aligned} \|f+g\|_1 &= \int_X |f+g| \, \mathrm{d}\mu \\ &\leq \int_X |f| + |g| \, \mathrm{d}\mu \\ &= \int_X |f| \, \mathrm{d}\mu + \int_X |g| \, \mathrm{d}\mu \\ &= \|f\|_1 + \|g\|_1. \end{aligned}$$

□

**Remark 9.2** **Why only a seminorm**

For $\|\cdot\|_1$ to be a *norm*, we would need the implication

$$[\|f\|_1 = 0] \implies [f = \underline{0}]$$

for every $f \in \mathcal{L}^1(\mu)$. But Lemma 8.10 and Remark 8.12 give the weaker equivalence

$$[\|f\|_1 = 0] \iff [f = \underline{0} \text{ a.e}-\mu].$$

In general, equality almost everywhere does not imply equality at every point of $X$, so $\|\cdot\|_1$ is only a seminorm on $\mathcal{L}^1(\mu)$. In particular, if there exists a non-empty set $A \in \mathfrak{M}$ with $\mu(A) = 0$, then $\|\cdot\|_1$ is not a norm.

**Definition 9.3** **The spaces $\mathcal{L}^p(\mu)$ and the maps $\|\cdot\|_p$**

Let $p \in [1, \infty)$. We denote

$$\mathcal{L}^p(\mu) := \left\{ f \in \mathrm{Bor}(X, \mathbb{R}) \;\middle|\; \int_X |f|^p \,\mathrm{d}\mu < \infty \right\}$$

for $f \in \mathcal{L}^p(\mu)$, we denote

$$\|f\|_p := \left( \int_X |f|^p \,\mathrm{d}\mu \right)^{\frac{1}{p}} \in [0, \infty)$$

Note that for $p = 1$ we get $\|\cdot\|_1$ on $\mathcal{L}^1(\mu)$.

**Proposition 9.4** **$\mathcal{L}^p(\mu)$ is a linear subspace**

For every $p \in [1, \infty)$, we have that $\mathcal{L}^p(\mu)$ is a linear subspace of $\mathrm{Bor}(X, \mathbb{R})$.

***Proof:*** We need to check that $\mathcal{L}^p(\mu)$ is stable under addition and scalar multiplication. Also, $\mathcal{L}^p(\mu)$ is non-empty, since it contains the zero-vector $\underline{0}$ of $\mathrm{Bor}(X, \mathbb{R})$.

*Stability under addition.* We use the elementary inequality

$$|a + b|^p \leq (2 \max\{|a|, |b|\})^p \leq 2^p (|a|^p + |b|^p), \quad \forall a, b \in \mathbb{R}.$$

Given $f, g \in \mathrm{Bor}(X, \mathbb{R})$, applying this pointwise gives the inequality

$$|f + g|^p \leq 2^p \cdot |f|^p + 2^p \cdot |g|^p$$

in $\mathrm{Bor}^+(X, \mathbb{R})$. Integrating both sides in the $L^+$ sense gives

$$\int_X |f + g|^p \,\mathrm{d}\mu \leq 2^p \cdot \int_X |f|^p \,\mathrm{d}\mu + 2^p \cdot \int_X |g|^p \,\mathrm{d}\mu.$$

If $f, g \in \mathcal{L}^p(\mu)$, then the right-hand side is finite, so the left-hand side is finite as well. Hence $f + g \in \mathcal{L}^p(\mu)$.

*Stability under scalar multiplication.* For $\alpha \in \mathbb{R}$ and $f \in \mathcal{L}^p(\mu)$, we have $\alpha f \in \mathrm{Bor}(X, \mathbb{R})$ and

$$\int_X |\alpha f|^p \,\mathrm{d}\mu = \int_X |\alpha|^p \cdot |f|^p \,\mathrm{d}\mu = |\alpha|^p \cdot \int_X |f|^p \,\mathrm{d}\mu < \infty.$$

Thus $\alpha f \in \mathcal{L}^p(\mu)$. Moreover,

$$\|\alpha f\|_p = \left(\int_X |\alpha f|^p \,\mathrm{d}\mu\right)^{\frac{1}{p}} = \left(|\alpha|^p \cdot \int_X |f|^p \,\mathrm{d}\mu\right)^{\frac{1}{p}} = |\alpha| \cdot \|f\|_p.$$

□

**Remark 9.5** **Minkowski inequality**

Pick a number $p \in (1, \infty)$ and consider the question whether $\|\cdot\|_p$ is a seminorm on $\mathcal{L}^p(\mu)$. If this holds, then it would be an analogue of the fact observed at the beginning of the lecture that $\|\cdot\|_1$ is a seminorm on $\mathcal{L}^1(\mu)$.

More precisely, we would like to know whether $\|\cdot\|_p$ has the properties (Non-Negative), (Sub-Additive), and (Homogeneous). The property (Non-Negative) is immediate, and (Homogeneous) was verified during the proof of $\mathcal{L}^p(\mu)$ is a linear subspace (Proposition 9.4). We are left to discuss (Sub-Additive), which is the inequality

$$\|f + g\|_p \le \|f\|_p + \|g\|_p, \quad \forall f, g \in \mathcal{L}^p(\mu).$$

It turns out that this inequality does hold. It is called *Minkowski's inequality*. Its verification poses some difficulties, though: it is not as easy as the corresponding (Sub-Additive) inequality for $\|\cdot\|_1$.

## 9.2 Conjugate Exponents

**Definition 9.6** **Conjugate exponents**

In order to get Minkowski's inequality for an exponent $p \in (1, \infty)$, here is the idea: the exponent $p$ has a conjugate exponent $q \in (1, \infty)$, and we can get Minkowski's inequality in the space $\mathcal{L}^p(\mu)$ by playing it against the conjugate space $\mathcal{L}^q(\mu)$.

Two numbers $p, q \in (1, \infty)$ are said to be *conjugate exponents* when they satisfy

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Thus, if we are given $p \in (1, \infty)$, then there is precisely one conjugate exponent for it, namely $q := \frac{p}{p-1}$. The idea is that even if we want to prove a result which is just about $\mathcal{L}^p(\mu)$, such as Minkowski's inequality, it can still be useful to bring in the sibling space $\mathcal{L}^q(\mu)$ associated to the conjugate exponent.

**Proposition 9.7** **Hölder's inequality**

Let $p, q \in (1, \infty)$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. One has that if $f \in \mathcal{L}^p(\mu)$ and $g \in \mathcal{L}^q(\mu)$ then $f \cdot g \in \mathcal{L}^1(\mu)$ and

$$\|f \cdot g\|_1 \leq \|f\|_p \cdot \|g\|_q. \tag{Hölder 1}$$

As a consequence, we also have

$$\left| \int_X f \cdot g \, \mathrm{d}\mu \right| \leq \|f\|_p \cdot \|g\|_q. \tag{Hölder 2}$$

**Note**

- Hölder's inequality (Proposition 9.7) can be thought of as a generalization of Cauchy-Schwarz.
- Morally, this introduces a duality between $\mathcal{L}^p(\mu)$ and $\mathcal{L}^q(\mu)$: to every $f \in \mathcal{L}^p(\mu)$ we associate the map $\varphi_f : \mathcal{L}^q(\mu) \to \mathbb{R}$ given by

  $$\varphi_f(g) := \int_X f \cdot g \, \mathrm{d}\mu$$

  for all $g \in \mathcal{L}^q(\mu)$.

**Remark 9.8** **Cauchy-Schwarz as a special case of Hölder**

Before jumping at the proof of Hölder's inequality (Proposition 9.7), let us pinpoint what the inequality (Hölder 2) amounts to in the important special case when $p = q = 2$. In that case, Hölder's inequality (Proposition 9.7) claims that

$$[f, g \in \mathcal{L}^2(\mu)] \Longrightarrow [f \cdot g \in \mathcal{L}^1(\mu)].$$

It is customary to introduce an "inner product" notation which takes advantage of the above implication and puts

$$\langle f, g \rangle := \int_X f \cdot g \, \mathrm{d}\mu \in \mathbb{R}, \quad \text{for } f, g \in \mathcal{L}^2(\mu).$$

The incarnation of (Hölder 2) that comes up in this way is known as the *Cauchy-Schwarz inequality* for $\mathcal{L}^2$-functions:

$$|\langle f, g \rangle| \leq \|f\|_2 \cdot \|g\|_2, \quad \forall f, g \in \mathcal{L}^2(\mu). \tag{Cauchy-Schwarz}$$

We now take on the proof of Hölder's inequality (Proposition 9.7). The first step in that direction is to record the elementary inequality stated in the next lemma.

**Lemma 9.9** **Young's inequality**

Let $p, q \in (1, \infty)$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. For every $a, b \in [0, \infty)$ one has the inequality

$$a \cdot b \leq \frac{1}{p} a^p + \frac{1}{q} b^q. \tag{$\bigcirc$}$$

***Proof:*** If one of the numbers $a, b$ is equal to 0, then the inequality is clear. So suppose that $a, b > 0$, and let

$$\alpha = \log a \quad \text{and} \quad \beta = \log b.$$

Then $a = e^\alpha$ and $b = e^\beta$, and what we need to prove in ($\bigcirc$) becomes

$$e^{\alpha+\beta} \leq \frac{1}{p} e^{\alpha p} + \frac{1}{q} e^{\beta q}. \tag{$\bigcirc\bigcirc$}$$

Since $\frac{1}{p} + \frac{1}{q} = 1$, the right-hand side of ($\bigcirc\bigcirc$) is a convex combination of the numbers $e^{\alpha p}$ and $e^{\beta q}$. We use the convexity of the exponential function, which gives

$$e^{\frac{1}{p}s + \frac{1}{q}t} \leq \frac{1}{p} e^s + \frac{1}{q} e^t, \quad \forall s, t \in \mathbb{R}. \tag{$*$}$$

Taking $s = \alpha p$ and $t = \beta q$ gives

$$\frac{1}{p} s + \frac{1}{q} t = \frac{1}{p}(\alpha p) + \frac{1}{q}(\beta q) = \alpha + \beta,$$

hence ($*$) yields ($\bigcirc\bigcirc$), as required. □

Now back to the setting of the spaces $\mathcal{L}^p(\mu)$.

**Lemma 9.10**

Let $p, q \in (1, \infty)$ be such that $\frac{1}{p} + \frac{1}{q} = 1$, and consider two functions $f \in \mathcal{L}^p(\mu)$ and $g \in \mathcal{L}^q(\mu)$. Suppose that $\|f\|_p = \|g\|_q = 1$. Then $f \cdot g \in \mathcal{L}^1(\mu)$ and

$$\left| \int_X f \cdot g \, \mathrm{d}\mu \right| \leq 1.$$

***Proof:*** Observe that

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

Likewise,

$$\int_X |g|^q \, \mathrm{d}\mu = 1.$$

Applying Young's inequality (Lemma 9.9) with $a = |f(x)|$ and $b = |g(x)|$, for $x \in X$, gives

$$|f(x)||g(x)| \leq \frac{|f(x)|^p}{p} + \frac{|g(x)|^q}{q}, \quad \forall x \in X.$$

hence

$$|f \cdot g| \leq \frac{1}{p}|f|^p + \frac{1}{q}|g|^q$$

so

$$\begin{aligned}\int_X |f \cdot g| \, \mathrm{d}\mu &\leq \int_X \frac{1}{p}|f|^p + \frac{1}{q}|g|^q \, \mathrm{d}\mu \\ &= \frac{1}{p}\int_X |f|^p \, \mathrm{d}\mu + \frac{1}{q}\int_X |g|^q \, \mathrm{d}\mu \\ &= \frac{1}{p} \cdot 1 + \frac{1}{q} \cdot 1 = 1.\end{aligned}$$

Thus $f \cdot g \in \mathcal{L}^1(\mu)$, and

$$\left|\int_X f \cdot g \, \mathrm{d}\mu\right| \leq \int_X |f \cdot g| \, \mathrm{d}\mu \leq 1.$$

$\square$

***Proof of Hölder's inequality (Proposition 9.7):*** Let $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$, and let $f \in \mathcal{L}^p(\mu)$, $g \in \mathcal{L}^q(\mu)$. Put

$$\alpha := \|f\|_p \quad \text{and} \quad \beta := \|g\|_q.$$

We split into cases.

1. Suppose that one of $\alpha, \beta$ is 0. Say $\alpha = 0$. Then $\int_X |f|^p \, \mathrm{d}\mu = 0$, so Remark 8.12 gives $|f|^p = \underline{0}$ a.e$-\mu$, hence $f = \underline{0}$ a.e$-\mu$. It follows that $f \cdot g = \underline{0}$ a.e$-\mu$, and hence $|f \cdot g| = \underline{0}$ a.e$-\mu$. Since $f \cdot g \in \mathrm{Bor}(X, \mathbb{R})$, Lemma 8.10 applied to $|f \cdot g|$ gives

$$\int_X |f \cdot g| \, \mathrm{d}\mu = 0.$$

Thus $f \cdot g \in \mathcal{L}^1(\mu)$ and

$$\|f \cdot g\|_1 = 0 = \alpha \cdot \beta.$$

The case $\beta = 0$ is identical.

2. Suppose that $\alpha, \beta \neq 0$. Let $\tilde{f} = \frac{1}{\alpha} f$ and $\tilde{g} = \frac{1}{\beta} g$. Then $\tilde{f} \in \mathcal{L}^p(\mu)$ and $\tilde{g} \in \mathcal{L}^q(\mu)$, with

$$\left\|\tilde{f}\right\|_p = 1 \quad \text{and} \quad \|\tilde{g}\|_q = 1.$$

Applying Lemma 9.10 to $\left|\tilde{f}\right|$ and $|\tilde{g}|$ gives

$$\left\|\tilde{f} \cdot \tilde{g}\right\|_1 \leq 1.$$

Here we used the identity $\left|\tilde{f} \cdot \tilde{g}\right| = \left|\tilde{f}\right| \cdot |\tilde{g}|$. Since

$$f \cdot g = \alpha \cdot \beta \cdot \tilde{f} \cdot \tilde{g},$$

it follows that $f \cdot g \in \mathcal{L}^1(\mu)$ and

$$\|f \cdot g\|_1 \leq \alpha \cdot \beta = \|f\|_p \cdot \|g\|_q.$$

This proves (Hölder 1). The inequality (Hölder 2) follows from

$$\left| \int_X f \cdot g \, \mathrm{d}\mu \right| \leq \int_X |f \cdot g| \, \mathrm{d}\mu = \|f \cdot g\|_1.$$

□

## 9.3 Minkowski's Inequality

We fix $p \in (1, \infty)$, and our goal is to prove Minkowski's inequality for functions in $\mathcal{L}^p(\mu)$. Towards that goal, we also consider the conjugate exponent $q = \frac{p}{p-1}$ for our fixed $p$.

**Notation and Remark 9.11** **Minkowski setup**

Let $f \in \mathcal{L}^p(\mu)$.

1. We define a function $\varphi_f : \mathcal{L}^q(\mu) \to \mathbb{R}$ by the formula

$$\varphi_f(g) := \int_X f \cdot g \, \mathrm{d}\mu, \quad \forall g \in \mathcal{L}^q(\mu).$$

   This definition makes sense due to Hölder's inequality (Proposition 9.7), which assures us that $f \cdot g \in \mathcal{L}^1(\mu)$ for every $g \in \mathcal{L}^q(\mu)$.
2. The map $\varphi_f$ is linear and has the property that

$$\left|\varphi_f(g)\right| \leq \|f\|_p \cdot \|g\|_q, \quad \forall g \in \mathcal{L}^q(\mu).$$

   Indeed, the linearity of $\varphi_f$ follows from the linearity of the integral, while the inequality is precisely (Hölder 2) from Hölder's inequality (Proposition 9.7).
3. The preceding inequality implies that the set

$$\left\{\left|\varphi_f(g)\right| \;\middle|\; g \in \mathcal{L}^q(\mu), \|g\|_q \leq 1\right\} \subseteq [0, \infty)$$

   is bounded from above, with $\|f\|_p$ being an upper bound. We may therefore define

$$\lambda(f) := \sup\left\{\left|\varphi_f(g)\right| \;\middle|\; g \in \mathcal{L}^q(\mu), \|g\|_q \leq 1\right\} \in [0, \infty),$$

   and we know that $\lambda(f) \leq \|f\|_p$.

**Lemma 9.12** **Operator norm formula**

Let $f \in \mathcal{L}^p(\mu)$, and consider the number $\lambda(f) \in [0,\infty)$ defined in Minkowski setup (Notation and Remark 9.11). Then

$$\lambda(f) = \|f\|_p.$$

***Proof:*** The inequality $\lambda(f) \leq \|f\|_p$ was seen in Minkowski setup (Notation and Remark 9.11). In order to prove the reverse inequality, we will display a function $g_0 \in \mathcal{L}^q(\mu)$ such that $\|g_0\|_q \leq 1$ and $\varphi_f(g_0) = \|f\|_p$. For such a function $g_0$, we get

$$\begin{aligned} \lambda(f) &= \sup\{|\varphi_f(g)| \mid g \in \mathcal{L}^q(\mu), \|g\|_q \leq 1\} \\ &\geq |\varphi_f(g_0)| \\ &= \|f\|_p, \end{aligned}$$

as required.

***<u>Claim 1:</u>*** One can find $u \in \operatorname{Bor}(X,\mathbb{R})$ such that $|u| = \underline{1}$ and $f = u|f|$.

***<u>Verification of Claim 1:</u>*** Define $u : X \to \mathbb{R}$ by

$$u(x) = \begin{cases} 1 & \text{if } f(x) \geq 0 \\ -1 & \text{if } f(x) < 0. \end{cases}$$

Since $f \in \operatorname{Bor}(X,\mathbb{R})$, we have $u \in \operatorname{Bor}(X,\mathbb{R})$. It is immediate from the definition that $|u| = \underline{1}$ and $f = u|f|$.

***<u>Claim 2:</u>*** Let $h = u|f|^{\frac{p}{q}}$, with $u$ as in Claim 1. Then $h \in \mathcal{L}^q(\mu)$ and

1. $\|h\|_q = \|f\|_p^{p-1}$.
2. $\int_X fh \,\mathrm{d}\mu = \|f\|_p^p$.

***<u>Verification of Claim 2:</u>*** Since $|u| = \underline{1}$ and $\left(\frac{p}{q}\right)q = p$, we have

$$|h|^q = |f|^p.$$

It follows that $h \in \mathcal{L}^q(\mu)$ and

$$\|h\|_q = \left(\int_X |h|^q \,\mathrm{d}\mu\right)^{\frac{1}{q}} = \left(\int_X |f|^p \,\mathrm{d}\mu\right)^{\frac{1}{q}} = \|f\|_p^{\frac{p}{q}} = \|f\|_p^{p-1}.$$

Also,

$$fh = (u|f|)\left(u|f|^{\frac{p}{q}}\right) = u^2|f|^{1+\frac{p}{q}} = |f|^p,$$

hence

$$\int_X fh \,\mathrm{d}\mu = \int_X |f|^p \,\mathrm{d}\mu = \|f\|_p^p.$$

We now finish the construction of $g_0$. If $\|f\|_p = 0$, we use $g_0 = \underline{0}$. Suppose then that $\|f\|_p \neq 0$. Let $h$ be as in Claim 2, and put

$$g_0 = \frac{1}{\|f\|_p^{p-1}} h.$$

Then $g_0 \in \mathcal{L}^q(\mu)$, and Claim 2 gives

$$\|g_0\|_q = 1 \quad \text{and} \quad \int_X f g_0 \, \mathrm{d}\mu = \|f\|_p.$$

□

**Proposition 9.13** **Minkowski's inequality**

For every $f_1, f_2 \in \mathcal{L}^p(\mu)$, one has

$$\|f_1 + f_2\|_p \leq \|f_1\|_p + \|f_2\|_p.$$

***Proof:*** By $\mathcal{L}^p(\mu)$ is a linear subspace (Proposition 9.4), we know that $f_1 + f_2 \in \mathcal{L}^p(\mu)$. In view of Operator norm formula (Lemma 9.12), it is enough to prove

$$\lambda(f_1 + f_2) \leq \lambda(f_1) + \lambda(f_2). \tag{$\diamondsuit$}$$

By definition,

$$\lambda(f_1 + f_2) = \sup\left\{ \left|\varphi_{f_1+f_2}(g)\right| \;\middle|\; g \in \mathcal{L}^q(\mu), \|g\|_q \leq 1 \right\}.$$

Thus it suffices to check that, for every $g \in \mathcal{L}^q(\mu)$ with $\|g\|_q \leq 1$, one has

$$\left|\varphi_{f_1+f_2}(g)\right| \leq \lambda(f_1) + \lambda(f_2). \tag{$\diamondsuit_-$}$$

So fix $g \in \mathcal{L}^q(\mu)$ with $\|g\|_q \leq 1$. By Hölder's inequality (Proposition 9.7), the functions $f_1 \cdot g$ and $f_2 \cdot g$ are in $\mathcal{L}^1(\mu)$. Hence

$$\begin{aligned} \left|\varphi_{f_1+f_2}(g)\right| &= \left| \int_X f_1 \cdot g + f_2 \cdot g \, \mathrm{d}\mu \right| \\ &= \left| \int_X f_1 \cdot g \, \mathrm{d}\mu + \int_X f_2 \cdot g \, \mathrm{d}\mu \right| \\ &\leq \left| \int_X f_1 \cdot g \, \mathrm{d}\mu \right| + \left| \int_X f_2 \cdot g \, \mathrm{d}\mu \right| \\ &\leq \lambda(f_1) + \lambda(f_2), \end{aligned}$$

where the last line follows from the definition of $\lambda(f_1)$ and $\lambda(f_2)$, since $\|g\|_q \leq 1$. This proves ($\diamondsuit_-$), and therefore ($\diamondsuit$). □

**Corollary 9.14** $\|\cdot\|_p$ **is a seminorm**

The map $\|\cdot\|_p$ is a seminorm on the vector space $\mathcal{L}^p(\mu)$.

***Proof:*** Non-negativity and homogeneity were observed in $\mathcal{L}^p(\mu)$ is a linear subspace (Proposition 9.4), and sub-additivity is precisely Minkowski's inequality (Proposition 9.13). □

## Appendix to Chapter 9

### A. Null-space of a seminorm, and how to quotient by it

**Notation and Remark 90.1** **Null-space of a seminorm**

Let $\mathcal{V}$ be a vector space over $\mathbb{R}$, and let the map

$$\begin{aligned} \mathcal{V} &\longrightarrow [0, \infty) \\ v &\longmapsto \|v\| \end{aligned}$$

be a seminorm. The set

$$\mathcal{N} := \{v \in \mathcal{V} \mid \|v\| = 0\} \subseteq \mathcal{V}$$

is called the *null-space* of the seminorm $\|\cdot\|$.

The term "null-space" is justified by the fact that $\mathcal{N}$ is a linear subspace of $\mathcal{V}$. Indeed, $\|0_{\mathcal{V}}\| = 0$, so $0_{\mathcal{V}} \in \mathcal{N}$. If $v_1, v_2 \in \mathcal{N}$, then

$$0 \leq \|v_1 + v_2\| \leq \|v_1\| + \|v_2\| = 0,$$

hence $v_1 + v_2 \in \mathcal{N}$. Finally, if $v \in \mathcal{N}$ and $\alpha \in \mathbb{R}$, then

$$\|\alpha v\| = |\alpha|\|v\| = 0,$$

so $\alpha v \in \mathcal{N}$.

The seminorm $\|\cdot\|$ is a norm on $\mathcal{V}$ if and only if its null-space is $\mathcal{N} = \{0_{\mathcal{V}}\}$.

**Notation and Remark 90.2** **Quotient by a linear subspace**

Let $\mathcal{V}$ be a vector space over $\mathbb{R}$, and let $\mathcal{W} \subseteq \mathcal{V}$ be a linear subspace. In order to describe the quotient space $\mathcal{V}/\mathcal{W}$, define a relation $\sim$ on $\mathcal{V}$ by

$$(v_1 \sim v_2) \iff (v_1 - v_2 \in \mathcal{W}).$$

This is an equivalence relation. It is reflexive because $v - v = 0_{\mathcal{V}} \in \mathcal{W}$. It is symmetric because $v_1 - v_2 \in \mathcal{W}$ implies $v_2 - v_1 = -(v_1 - v_2) \in \mathcal{W}$, and it is transitive because

$$v_1 - v_3 = (v_1 - v_2) + (v_2 - v_3) \in \mathcal{W}$$

whenever $v_1 \sim v_2$ and $v_2 \sim v_3$.

Let $\mathcal{Q}$ be the collection of equivalence classes with respect to $\sim$. This collection has a natural vector space structure and is called the *quotient space* $\mathcal{V}/\mathcal{W}$. There is a surjective linear map

$$\begin{aligned} \mathcal{V} &\longrightarrow \mathcal{Q} \\ v &\longmapsto \hat{v}, \end{aligned}$$

where $\hat{v}$ denotes the equivalence class of $v$.

Thus we may remember the quotient as

$$\mathcal{Q} = \{\hat{v} \mid v \in \mathcal{V}\},$$

where, for every $v_1, v_2 \in \mathcal{V}$,

$$(\widehat{v_1} = \widehat{v_2}) \iff (v_1 - v_2 \in \mathcal{W}),$$

and “hats respect linear combinations”:

$$\widehat{\alpha_1 v_1 + \alpha_2 v_2} = \alpha_1 \widehat{v_1} + \alpha_2 \widehat{v_2}, \quad \forall v_1, v_2 \in \mathcal{V}, \quad \forall \alpha_1, \alpha_2 \in \mathbb{R}.$$

**Remark 90.3** **How to get a norm out of a seminorm**

Suppose that the vector space $\mathcal{V}$ is endowed with a seminorm $\|\cdot\|$, let $\mathcal{N}$ be its null-space, and consider the quotient space $\mathcal{Q} = \mathcal{V}/\mathcal{N}$. We define a norm on $\mathcal{Q}$ as follows:

$$\begin{aligned}&\text{Take } \xi \in \mathcal{Q}.\\ &\text{Write } \xi = \hat{v} \text{ for some } v \in \mathcal{V}.\\ &\text{Define } \|\xi\|_{\mathcal{Q}} := \|v\| \in [0,\infty).\end{aligned}$$

There are three points to verify.

1. The definition is coherent. Suppose $\xi = \widehat{v_1} = \widehat{v_2}$. Then $v_1 - v_2 \in \mathcal{N}$, so $\|v_1 - v_2\| = 0$. By sub-additivity,
$$\|v_1\| \leq \|v_1 - v_2\| + \|v_2\| = \|v_2\|.$$
Interchanging $v_1$ and $v_2$ gives the reverse inequality, hence $\|v_1\| = \|v_2\|$.
2. The map $\xi \mapsto \|\xi\|_{\mathcal{Q}}$ inherits the seminorm properties from $\|\cdot\|$. Non-negativity is immediate. If $\xi = \hat{v}$ and $\alpha \in \mathbb{R}$, then
$$\|\alpha\xi\|_{\mathcal{Q}} = \|\widehat{\alpha v}\|_{\mathcal{Q}} = \|\alpha v\| = |\alpha|\|v\| = |\alpha|\|\xi\|_{\mathcal{Q}}.$$
Likewise, if $\xi_1 = \widehat{v_1}$ and $\xi_2 = \widehat{v_2}$, then
$$\begin{aligned}\|\xi_1 + \xi_2\|_{\mathcal{Q}} &= \left\|\widehat{v_1 + v_2}\right\|_{\mathcal{Q}}\\ &= \|v_1 + v_2\|\\ &\leq \|v_1\| + \|v_2\|\\ &= \|\xi_1\|_{\mathcal{Q}} + \|\xi_2\|_{\mathcal{Q}}.\end{aligned}$$
3. The resulting seminorm is a norm. Indeed, if $\|\xi\|_{\mathcal{Q}} = 0$ and $\xi = \hat{v}$, then $\|v\| = 0$, so $v \in \mathcal{N}$. Therefore
$$\xi = \hat{v} = \widehat{0_{\mathcal{V}}} = 0_{\mathcal{Q}}.$$

**B. The quotient space $L^p(\mu)$ of $\mathcal{L}^p(\mu)$**

We return to the usual framework of a measure space $(X, \mathfrak{M}, \mu)$. Define

$$\mathcal{N} := \{f \in \mathrm{Bor}(X, \mathbb{R}) \mid f = \underline{0} \text{ a.e}-\mu\}. \tag{90.6}$$

Thus $f \in \mathcal{N}$ means that there exists $Z \in \mathfrak{M}$ with $\mu(Z) = 0$ such that $f(x) = 0$ for every $x \in X \setminus Z$. Equivalently, by Proposition 8.9,

$$\mu(\{x \in X \mid f(x) \neq 0\}) = 0.$$

The definition of $\mathcal{N}$ is independent of the exponent used to define the spaces $\mathcal{L}^p(\mu)$.

**Proposition 90.4**

Let $p \in [1, \infty)$. Then the space $\mathcal{N}$ from (90.6) is contained in $\mathcal{L}^p(\mu)$ and is precisely the null-space of the seminorm $\|\cdot\|_p$ on $\mathcal{L}^p(\mu)$.

***Proof:*** Let $f \in \mathcal{N}$. Then $|f|^p = \underline{0}$ a.e$-\mu$, so Lemma 8.10 gives

$$\int_X |f|^p \, \mathrm{d}\mu = 0.$$

Hence $f \in \mathcal{L}^p(\mu)$ and $\|f\|_p = 0$.

Conversely, suppose that $f \in \mathcal{L}^p(\mu)$ and $\|f\|_p = 0$. Then

$$\int_X |f|^p \, \mathrm{d}\mu = 0.$$

By Remark 8.12, $|f|^p = \underline{0}$ a.e$-\mu$, which is equivalent to $f = \underline{0}$ a.e$-\mu$. Thus $f \in \mathcal{N}$. □

**Definition 90.5**

For every $p \in [1, \infty)$, we denote

$$L^p(\mu) := \mathcal{L}^p(\mu)/\mathcal{N},$$

where $\mathcal{N}$ is the space from (90.6).

By the construction in How to get a norm out of a seminorm (Remark 90.3), the seminorm $\|\cdot\|_p$ on $\mathcal{L}^p(\mu)$ induces a norm on $L^p(\mu)$, which we still denote by $\|\cdot\|_p$. Thus $\left(L^p(\mu), \|\cdot\|_p\right)$ is a normed vector space.

**Remark 90.6** **So what is $L^p(\mu)$, after all?**

The quotient $L^p(\mu)$ is a "space of objects with hats":

$$L^p(\mu) = \left\{\hat{f} \;\middle|\; f \in \mathcal{L}^p(\mu)\right\}.$$

For $f, g \in \mathcal{L}^p(\mu)$, we have

$$\left(\hat{f} = \hat{g}\right) \iff (f = g \text{ a.e}-\mu). \tag{90.9}$$

Indeed,

$$(f - g \in \mathcal{N}) \iff (f - g = \underline{0} \text{ a.e}-\mu) \iff (f = g \text{ a.e}-\mu),$$

so (90.9) is exactly the equivalence relation defining the quotient.

Finally, hats respect linear combinations:

$$\widehat{\alpha f + \beta g} = \alpha \hat{f} + \beta \hat{g}, \quad \forall f, g \in \mathcal{L}^p(\mu), \quad \forall \alpha, \beta \in \mathbb{R},$$

and the seminorm on $\mathcal{L}^p(\mu)$ is transferred to the quotient by putting

$$\left\|\hat{f}\right\|_p := \|f\|_p, \quad \forall f \in \mathcal{L}^p(\mu).$$

By How to get a norm out of a seminorm (Remark 90.3) and Proposition 90.4, this is well-defined and is a norm on $L^p(\mu)$.

# 10 $\mathcal{L}^p(\mu)$-Completeness

## 10.1 Completeness in the framework of a semi-normed vector space

Fix an $\mathbb{R}$-vector space $\mathcal{V}$ and a seminorm $\|\cdot\| : \mathcal{V} \to [0, \infty)$. We call the pair $(\mathcal{V}, \|\cdot\|)$ a *semi-normed vector space.*

**Definition 10.1** **Convergence in a semi-normed space**

Let $(v_n)_{n=1}^{\infty}$ be a sequence of vectors in $\mathcal{V}$. We say $(v_n)_{n=1}^{\infty}$ *converges* to $v \in \mathcal{V}$ when $\lim_{n\to\infty} \|v_n - v\| = 0$, and we write $v_n \to v$. We say $(v_n)_{n=1}^{\infty}$ *converges* if $(v_n)_{n=1}^{\infty}$ converges to some $v \in \mathcal{V}$.

**Warning 10.2** **Non-uniqueness of limits**

The limit of a sequence is not uniquely determined. Indeed, if $v_n \to v \in \mathcal{V}$ and $v_n \to v' \in \mathcal{V}$, then

$$\|v - v'\| = \|v - v_n + v_n - v'\| \leq \|v_n - v\| + \|v_n - v'\| \longrightarrow 0.$$

Thus we can conclude that $\|v - v'\| = 0$, but this does not imply that $v = v'$.

**Definition 10.3** **Cauchy and geometric sequences**

Let $(v_n)_{n=1}^{\infty}$ be a sequence of vectors in $\mathcal{V}$.

1. We say $(v_n)_{n=1}^{\infty}$ is a *Cauchy sequence* when, for every $\varepsilon > 0$, we can find $n_o \in \mathbb{N}$ such that whenever $n, m \geq n_o$, we have $\|v_n - v_m\| < \varepsilon$.
2. We say $(v_n)_{n=1}^{\infty}$ is a $\frac{1}{2}$*-geometric sequence* when

$$\|v_n - v_{n+1}\| \leq \frac{1}{2^n} \quad \forall n \in \mathbb{N}.$$

**Remark 10.4** **Basic sequence facts**

We can prove the following statements using the same arguments used in normed vector spaces:

1. if $(v_n)_{n=1}^{\infty}$ is convergent, then it is Cauchy;
2. if $(v_n)_{n=1}^{\infty}$ is $\frac{1}{2}$-geometric, then it is Cauchy;
3. if $(v_n)_{n=1}^{\infty}$ is Cauchy, then there is a subsequence $\left(v_{n(k)}\right)_{k=1}^{\infty}$ that is $\frac{1}{2}$-geometric.

**Definition 10.5** **Completeness**

A semi-normed space $(\mathcal{V}, \|\cdot\|)$ is said to be *complete* when every Cauchy sequence in $(\mathcal{V}, \|\cdot\|)$ converges.

**Remark 10.6** **Geometric criterion for completeness**

A familiar fact of normed vector spaces remains true in semi-normed spaces: if every $\frac{1}{2}$-geometric sequence in $(\mathcal{V}, \|\cdot\|)$ converges, then $(\mathcal{V}, \|\cdot\|)$ is complete.

We can also weaken the hypothesis slightly: if every $\frac{1}{2}$-geometric sequence $(w_n)_{n=1}^{\infty}$ in $(\mathcal{V}, \|\cdot\|)$ with $w_1 = 0$ converges, then $(\mathcal{V}, \|\cdot\|)$ is complete.

## 10.2 Completeness of $\mathcal{L}^p(\mu)$

**Lemma 10.7** **Convergence mechanism**

Let $f_1, f_2, \dots, f_n, \dots$ be a sequence of functions from $X$ to $\mathbb{R}$. We consider the helper functions $h_n : X \to [0, \infty)$ defined by

$$h_1 = |f_1|, \quad h_2 = |f_1| + |f_2 - f_1|, \quad h_3 = |f_1| + |f_2 - f_1| + |f_3 - f_2|, \cdots$$

and, in general,

$$h_n = |f_1| + \sum_{m=2}^{n} |f_m - f_{m-1}|. \tag{10.1}$$

We observe that

$$0 \leq h_1 \leq h_2 \leq \cdots \leq h_n \leq \cdots$$

and we denote

$$Z := \left\{ x \in X \;\middle|\; \lim_{n \to \infty} h_n(x) = \infty \right\}. \tag{10.2}$$

Then, for every $x \in X \setminus Z$, the sequence $(f_n(x))_{n=1}^{\infty}$ is convergent in $\mathbb{R}$. It therefore makes sense to consider the function $f : X \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 0 & \text{if } x \in Z \\ \lim_{n \to \infty} f_n(x) & \text{if } x \in X \setminus Z. \end{cases} \tag{10.3}$$

***Proof:*** Fix $x \in X \setminus Z$. We will show that the sequence of real numbers $(f_n(x))_{n=1}^{\infty}$ is Cauchy, and hence convergent in $\mathbb{R}$.

Since $x \in X \setminus Z$, the increasing sequence

$$0 \leq h_1(x) \leq h_2(x) \leq \cdots \leq h_n(x) \leq \cdots$$

has a finite limit $\ell \in [0, \infty)$. Let $\varepsilon > 0$. We can find $n_o \in \mathbb{N}$ such that

$$h_{n_o}(x) > \ell - \varepsilon.$$

We claim that $|f_m(x) - f_n(x)| < \varepsilon$ whenever $n, m \geq n_o$. By symmetry, it is enough to consider the case $n > m$. In that case,

$$\begin{aligned}|f_m(x) - f_n(x)| &= \left|\sum_{k=m}^{n-1} (f_k(x) - f_{k+1}(x))\right| \\ &\leq \sum_{k=m}^{n-1} |f_k(x) - f_{k+1}(x)| \\ &= h_n(x) - h_m(x) \\ &< \ell - (\ell - \varepsilon) \\ &= \varepsilon.\end{aligned}$$

Here the strict inequality follows because $h_n(x) \leq \ell$, while

$$h_m(x) \geq h_{n_o}(x) > \ell - \varepsilon.$$

Therefore $(f_n(x))_{n=1}^{\infty}$ is Cauchy and hence converges in $\mathbb{R}$, as required. □

**Proposition 10.8**

Let $(f_n)_{n=1}^{\infty}$ be a sequence of functions in $\mathcal{L}^p(\mu)$ such that $f_1 = \underline{0}$ and

$$\|f_n - f_{n+1}\|_p \leq \frac{1}{2^n}, \quad \forall n \in \mathbb{N}.$$

Upon applying to the functions $(f_n)_{n=1}^{\infty}$ the general convergence mechanism from Convergence mechanism (Lemma 10.7), we get a function $f : X \to \mathbb{R}$, as described in (10.3).

***Claim:*** We have that $f \in \mathcal{L}^p(\mu)$, and $\lim_{n\to\infty} \|f_n - f\|_p = 0$.

***Proof:*** We will use the helper functions $(h_n)_{n=1}^{\infty}$ and the set $Z$ defined in (10.1) and (10.2). In addition, define $h : X \to [0, \infty)$ by

$$h(x) = \begin{cases} 0 & \text{if } x \in Z \\ \lim_{n\to\infty} h_n(x) & \text{if } x \in X \setminus Z. \end{cases} \tag{10.5}$$

We will reach the required conclusion through a number of claims concerning $(f_n)_{n=1}^{\infty}$ and the auxiliary functions $(h_n)_{n=1}^{\infty}$ and $h$.

***Claim 1:*** For every $n \in \mathbb{N}$, we have $h_n \in \mathcal{L}^p(\mu)$ and $\|h_n\|_p \leq 1$.

***Verification of Claim 1:*** We may assume that $n \geq 2$, since the case $n = 1$ is immediate from $h_1 = |f_1| = \underline{0}$.

For every $m$ with $2 \leq m \leq n$, we have $f_m - f_{m-1} \in \mathcal{L}^p(\mu)$ by $\mathcal{L}^p(\mu)$ is a linear subspace (Proposition 9.4). It follows that

$$u_m := |f_m - f_{m-1}|$$

belongs to $\mathcal{L}^p(\mu)$, with

$$\|u_m\|_p = \|f_m - f_{m-1}\|_p \leq \frac{1}{2^{m-1}}.$$

Since $f_1 = \underline{0}$, we have

$$h_n = u_2 + \cdots + u_n.$$

Therefore $\mathcal{L}^p(\mu)$ is a linear subspace (Proposition 9.4) and $\|\cdot\|_p$ is a seminorm (Corollary 9.14) give $h_n \in \mathcal{L}^p(\mu)$, with

$$\begin{aligned}\|h_n\|_p &\leq \|u_2\|_p + \cdots + \|u_n\|_p \\ &\leq \frac{1}{2} + \cdots + \frac{1}{2^{n-1}} \\ &\leq 1.\end{aligned}$$

***Claim 2:*** We have $Z \in \mathfrak{M}$ and $\mu(Z) = 0$. Moreover, the function $h$ from (10.5) belongs to $\mathcal{L}^p(\mu)$ and satisfies $\|h\|_p \leq 1$.

***Verification of Claim 2:*** Since every $h_n$ is Borel, we can write

$$Z = \bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \{x \in X \mid h_n(x) > k\},$$

which shows that $Z \in \mathfrak{M}$.

We use the following consequence of the Monotone Convergence Theorem, whose proof is left as an exercise.

**Exercise 10.1**

Let $(u_n)_{n=1}^{\infty}$ be an increasing sequence in $\mathrm{Bor}^+(X, \mathbb{R})$. If there exists $A \in \mathfrak{M}$ with $\mu(A) > 0$ such that

$$\lim_{n\to\infty} u_n(x) = \infty, \quad \forall x \in A,$$

then

$$\lim_{n\to\infty} \int_X u_n \, \mathrm{d}\mu = \infty.$$

Suppose, towards a contradiction, that $\mu(Z) > 0$. By Claim 1, the functions $h_n^p$ are in $\mathrm{Bor}^+(X, \mathbb{R})$, and

$$0 = h_1^p \leq h_2^p \leq \cdots \leq h_n^p \leq \cdots.$$

Moreover, for every $x \in Z$, we have

$$\lim_{n\to\infty} h_n(x)^p = \infty.$$

Applying the exercise to the sequence $(h_n^p)_{n=1}^{\infty}$ would give

$$\lim_{n\to\infty} \int_X h_n^p \, \mathrm{d}\mu = \infty.$$

On the other hand, Claim 1 gives

$$\int_X h_n^p \, \mathrm{d}\mu = \|h_n\|_p^p \leq 1, \quad \forall n \in \mathbb{N},$$

a contradiction. Hence $\mu(Z) = 0$.

We now consider the sequence $(h_n \chi_{X \setminus Z})_{n=1}^{\infty}$. Every function in this sequence is in $\mathrm{Bor}^+(X, \mathbb{R})$, and

$$h_n \chi_{X \setminus Z} \nearrow h.$$

In particular, Corollary 4.4 gives $h \in \mathrm{Bor}^+(X, \mathbb{R})$. Applying Monotone Convergence Theorem (Theorem 6.2) to the increasing sequence

$$\left((h_n \chi_{X \setminus Z})^p\right)_{n=1}^{\infty}$$

gives

$$\lim_{n \to \infty} \int_X (h_n \chi_{X \setminus Z})^p \, \mathrm{d}\mu = \int_X h^p \, \mathrm{d}\mu.$$

Since $\mu(Z) = 0$, the functions $(h_n \chi_{X \setminus Z})^p$ and $h_n^p$ are equal a.e$-\mu$. Thus Proposition 8.11 and Claim 1 give

$$\int_X h^p \, \mathrm{d}\mu = \lim_{n \to \infty} \int_X h_n^p \, \mathrm{d}\mu = \lim_{n \to \infty} \|h_n\|_p^p \leq 1.$$

Therefore $h \in \mathcal{L}^p(\mu)$ and $\|h\|_p \leq 1$.

***Claim 3:*** The function $f$ belongs to $\mathrm{Bor}(X, \mathbb{R})$.

***Verification of Claim 3:*** The function $f$ is the pointwise limit of the sequence

$$(f_n \chi_{X \setminus Z})_{n=1}^{\infty}.$$

Each function in this sequence belongs to $\mathrm{Bor}(X, \mathbb{R})$, so Corollary 4.4 gives $f \in \mathrm{Bor}(X, \mathbb{R})$.

***Claim 4:*** Fix $n_o \in \mathbb{N}$ and define

$$g = \left(f_{n_o} \chi_{X \setminus Z}\right) - f \in \mathrm{Bor}(X, \mathbb{R}).$$

Then $g \in \mathcal{L}^p(\mu)$, with

$$\|g\|_p \leq \frac{1}{2^{n_o - 1}}.$$

***Verification of Claim 4:*** For every $n \in \mathbb{N}$, define

$$g_n = \left(f_{n_o} - f_n\right) \chi_{X \setminus Z}.$$

By the definitions of $f$ and $g$, we have $g_n(x) \to g(x)$ for every $x \in X$. Moreover,

$$|g_n| \leq h, \quad \forall n \in \mathbb{N}.$$

Indeed, both sides vanish on $Z$, while for $x \in X \setminus Z$, the left-hand side is bounded by the sum of the absolute increments between $f_{n_o}(x)$ and $f_n(x)$, which is at most $h(x)$. Hence

$$|g_n|^p \leq h^p, \quad \forall n \in \mathbb{N}.$$

Since $h^p \in \mathcal{L}^1(\mu)$ by Claim 2, LDCT (Theorem 8.1) applied to the sequence $(|g_n|^p)_{n=1}^{\infty}$ gives

$$\int_X |g|^p \,\mathrm{d}\mu = \lim_{n\to\infty} \int_X |g_n|^p \,\mathrm{d}\mu.$$

On the other hand, for $n > n_o$, the functions $g_n$ and $f_{n_o} - f_n$ are equal a.e$-\mu$. Repeatedly applying $\|\cdot\|_p$ is a seminorm (Corollary 9.14) therefore gives

$$\begin{aligned}\|g_n\|_p &= \left\|f_{n_o} - f_n\right\|_p \\ &\leq \sum_{m=n_o}^{n-1} \left\|f_m - f_{m+1}\right\|_p \\ &\leq \sum_{m=n_o}^{n-1} \frac{1}{2^m} \\ &\leq \frac{1}{2^{n_o-1}}.\end{aligned}$$

It follows that

$$\int_X |g|^p \,\mathrm{d}\mu = \lim_{n\to\infty} \|g_n\|_p^p \leq \left(\frac{1}{2^{n_o-1}}\right)^p < \infty.$$

Thus $g \in \mathcal{L}^p(\mu)$ and $\|g\|_p \leq \frac{1}{2^{n_o-1}}$.

***Claim 5:*** We have $f \in \mathcal{L}^p(\mu)$.

***Verification of Claim 5:*** Fix any $n_o \in \mathbb{N}$. Since $\mu(Z) = 0$, we have

$$f_{n_o}\chi_{X\setminus Z} = f_{n_o} \text{ a.e}-\mu.$$

Hence $f_{n_o}\chi_{X\setminus Z} \in \mathcal{L}^p(\mu)$. Claim 4 also gives

$$\left(f_{n_o}\chi_{X\setminus Z}\right) - f \in \mathcal{L}^p(\mu).$$

Since

$$f = f_{n_o}\chi_{X\setminus Z} - \left(\left(f_{n_o}\chi_{X\setminus Z}\right) - f\right),$$

the linear stability of $\mathcal{L}^p(\mu)$ from $\mathcal{L}^p(\mu)$ is a linear subspace (Proposition 9.4) yields $f \in \mathcal{L}^p(\mu)$.

***Claim 6:*** We have

$$\lim_{n\to\infty} \|f_n - f\|_p = 0.$$

***Verification of Claim 6:*** Apply Claim 4 with $n_o = n$. Since $f$ vanishes on $Z$, we have

$$(f_n\chi_{X\setminus Z}) - f = (f_n - f)\chi_{X\setminus Z} = f_n - f \text{ a.e}-\mu.$$

It follows that

$$\|f_n - f\|_p \leq \frac{1}{2^{n-1}}, \quad \forall n \in \mathbb{N}.$$

Letting $n \to \infty$ proves the claim and completes the proof.

□

**Corollary 10.9**

The semi-normed vector space $(\mathcal{L}^p(\mu), \|\cdot\|_p)$ is complete.

***Proof:*** This follows immediately by combining Geometric criterion for completeness (Remark 10.6) and Proposition 10.8. □

**Corollary 10.10**

The normed vector space $(L^p(\mu), \|\cdot\|_p)$ is a Banach space.

***Proof:*** Let $(\xi_n)_{n=1}^{\infty}$ be a Cauchy sequence in $L^p(\mu)$. For every $n \in \mathbb{N}$, choose $f_n \in \mathcal{L}^p(\mu)$ such that

$$\xi_n = \widehat{f_n}.$$

By the definition of the vector-space operations and the norm on $L^p(\mu)$, for every $m, n \in \mathbb{N}$ we have

$$\begin{aligned} \|f_m - f_n\|_p &= \left\|\widehat{f_m - f_n}\right\|_p \\ &= \left\|\widehat{f_m} - \widehat{f_n}\right\|_p \\ &= \|\xi_m - \xi_n\|_p. \end{aligned}$$

Since $(\xi_n)_{n=1}^{\infty}$ is Cauchy in $L^p(\mu)$, it follows that $(f_n)_{n=1}^{\infty}$ is Cauchy in the semi-normed space $(\mathcal{L}^p(\mu), \|\cdot\|_p)$. By Corollary 10.9, there exists $f \in \mathcal{L}^p(\mu)$ such that

$$\lim_{n\to\infty} \|f_n - f\|_p = 0.$$

Put $\xi = \hat{f} \in L^p(\mu)$. Then, for every $n \in \mathbb{N}$,

$$\begin{aligned} \|\xi_n - \xi\|_p &= \left\|\widehat{f_n} - \hat{f}\right\|_p \\ &= \left\|\widehat{f_n - f}\right\|_p \\ &= \|f_n - f\|_p. \end{aligned}$$

Therefore

$$\lim_{n\to\infty} \|\xi_n - \xi\|_p = 0,$$

so $(\xi_n)_{n=1}^{\infty}$ converges to $\xi$ in $L^p(\mu)$. Thus $L^p(\mu)$ is complete and hence is a Banach space. □

## 10.3 Modes of convergence

A byproduct of the proof of Proposition 10.8 is that convergence with respect to $\|\cdot\|_p$ can be connected to a notion of almost-everywhere convergence.

**Definition 10.11** **Almost-everywhere convergence**

Let $(X, \mathfrak{M}, \mu)$ be a measure space, and let $f$ and $(f_n)_{n=1}^{\infty}$ be functions in $\mathrm{Bor}(X, \mathbb{R})$. We say that $(f_n)_{n=1}^{\infty}$ *converges to $f$ almost everywhere with respect to $\mu$*, or a.e.$-\mu$ for short, if there exists $Z \in \mathfrak{M}$ such that $\mu(Z) = 0$ and

$$\lim_{n\to\infty} f_n(x) = f(x), \quad \forall x \in X \setminus Z.$$

The relevance of this definition can be seen in the proof of Proposition 10.8. The function $f$ constructed there also satisfies $f_n(x) \to f(x)$ for every $x \in X \setminus Z$, and Claim 2 shows that $\mu(Z) = 0$. Thus the sequence converges to $f$ almost everywhere. This gives the following corollary.

**Corollary 10.12**

Let $(X, \mathfrak{M}, \mu)$ be a measure space, let $p \in [1, \infty)$, and let $(f_n)_{n=1}^{\infty}$ be a sequence in $\mathcal{L}^p(\mu)$ such that

$$\|f_n - f_{n+1}\|_p < \frac{1}{2^n}, \quad \forall n \in \mathbb{N}.$$

Then there exists $f \in \mathcal{L}^p(\mu)$ such that $(f_n)_{n=1}^{\infty}$ converges to $f$ almost everywhere with respect to $\mu$.

***Proof:*** Define $g_n = f_n - f_1$ for every $n \in \mathbb{N}$. Then $g_1 = \underline{0}$ and

$$\|g_n - g_{n+1}\|_p = \|f_n - f_{n+1}\|_p < \frac{1}{2^n}, \quad \forall n \in \mathbb{N}.$$

Apply Proposition 10.8 to $(g_n)_{n=1}^{\infty}$. Its proof produces a function $g \in \mathcal{L}^p(\mu)$ and a set $Z \in \mathfrak{M}$ with $\mu(Z) = 0$ such that

$$\lim_{n\to\infty} g_n(x) = g(x), \quad \forall x \in X \setminus Z.$$

Put $f = g + f_1 \in \mathcal{L}^p(\mu)$. For every $x \in X \setminus Z$, we then have

$$\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} (g_n(x) + f_1(x)) = g(x) + f_1(x) = f(x).$$

Hence $(f_n)_{n=1}^{\infty}$ converges to $f$ almost everywhere with respect to $\mu$. $\square$

# 11 Bounded linear functionals on $\mathcal{L}^2(\mu)$

Fix a measure space $(X, \mathfrak{M}, \mu)$, and consider the semi-normed vector space $(\mathcal{L}^2(\mu), \|\cdot\|_2)$.

**Definition 11.1** **Bounded linear functional**

Let $\varphi : \mathcal{L}^2(\mu) \to \mathbb{R}$ be a linear functional. We say that $\varphi$ is *bounded* if there exists a constant $c \geq 0$ such that

$$|\varphi(f)| \leq c \cdot \|f\|_2, \quad \forall f \in \mathcal{L}^2(\mu).$$

**Remark 11.2** **Bounded linear functionals from inner products**

Recall from the special case $p = q = 2$ of Hölder's inequality (Proposition 9.7) that if $f, g \in \mathcal{L}^2(\mu)$, then $f \cdot g \in \mathcal{L}^1(\mu)$. We use the inner-product notation

$$\langle f, g \rangle := \int_X f \cdot g \, \mathrm{d}\mu \in \mathbb{R},$$

and the Cauchy–Schwarz inequality from the beginning of Lecture 9B says

$$|\langle f, g \rangle| \leq \|f\|_2 \cdot \|g\|_2, \quad \forall f, g \in \mathcal{L}^2(\mu).$$

Consequently, every $f \in \mathcal{L}^2(\mu)$ determines a bounded linear functional $\varphi_f : \mathcal{L}^2(\mu) \to \mathbb{R}$ given by

$$\varphi_f(g) := \langle f, g \rangle = \int_X f \cdot g \, \mathrm{d}\mu, \quad \forall g \in \mathcal{L}^2(\mu).$$

Indeed, Cauchy–Schwarz gives

$$|\varphi_f(g)| \leq \|f\|_2 \cdot \|g\|_2, \quad \forall g \in \mathcal{L}^2(\mu).$$

The theorem below, due to Riesz, states that every bounded linear functional on $\mathcal{L}^2(\mu)$ arises in this way.

**Theorem 11.3** **Riesz representation theorem for $\mathcal{L}^2(\mu)$**

Let $\varphi : \mathcal{L}^2(\mu) \to \mathbb{R}$ be a bounded linear functional. Then there exists $f_o \in \mathcal{L}^2(\mu)$ such that

$$\varphi(g) = \langle f_o, g \rangle, \quad \forall g \in \mathcal{L}^2(\mu).$$

We first establish three lemmas.

**Lemma 11.4** **Parallelogram law**

For every $f, g \in \mathcal{L}^2(\mu)$, we have

$$\|f + g\|_2^2 + \|f - g\|_2^2 = 2\|f\|_2^2 + 2\|g\|_2^2.$$

***Proof:*** Using the pointwise identity $(f + g)^2 + (f - g)^2 = 2f^2 + 2g^2$, we obtain

$$\begin{aligned}
\|f + g\|_2^2 + \|f - g\|_2^2 &= \int_X |f + g|^2 \, \mathrm{d}\mu + \int_X |f - g|^2 \, \mathrm{d}\mu \\
&= \int_X \left((f + g)^2 + (f - g)^2\right) \mathrm{d}\mu \\
&= 2\int_X |f|^2 \, \mathrm{d}\mu + 2\int_X |g|^2 \, \mathrm{d}\mu \\
&= 2\|f\|_2^2 + 2\|g\|_2^2.
\end{aligned}$$

$\square$

**Lemma 11.5**

Let $\varphi : \mathcal{L}^2(\mu) \to \mathbb{R}$ be a linear functional which is not identically zero, and let

$$\mathcal{U} := \{u \in \mathcal{L}^2(\mu) \mid \varphi(u) = 1\}.$$

Then:

1. $\mathcal{U}$ is non-empty;
2. the number

$$\alpha_o := \inf\{\|u\|_2 \mid u \in \mathcal{U}\}$$

   is strictly positive;
3. there exists $u_o \in \mathcal{U}$ such that $\|u_o\|_2 = \alpha_o$.

***Proof of 1:*** Since $\varphi$ is not identically zero, choose $f \in \mathcal{L}^2(\mu)$ such that $\varphi(f) \neq 0$. Put

$$u = \frac{1}{\varphi(f)} f.$$

Then $\varphi(u) = 1$, so $u \in \mathcal{U}$. $\square$

***Proof of 2:*** Let $c \geq 0$ satisfy

$$|\varphi(f)| \leq c \cdot \|f\|_2, \quad \forall f \in \mathcal{L}^2(\mu).$$

We must have $c > 0$, since otherwise $\varphi$ would be identically zero. For every $u \in \mathcal{U}$,

$$1 = |\varphi(u)| \leq c \cdot \|u\|_2,$$

so $\|u\|_2 \geq \frac{1}{c}$. Hence $\alpha_o \geq \frac{1}{c} > 0$. $\square$

***Proof of 3:*** By the definition of the infimum, choose a sequence $(u_n)_{n=1}^{\infty}$ in $\mathcal{U}$ such that

$$\lim_{n \to \infty} \|u_n\|_2 = \alpha_o.$$

We claim that $(u_n)_{n=1}^{\infty}$ is Cauchy in $(\mathcal{L}^2(\mu), \|\cdot\|_2)$. Since

$$\varphi\left(\frac{u_n + u_m}{2}\right) = 1,$$

we have $\frac{u_n+u_m}{2} \in \mathcal{U}$, and therefore

$$\left\|\frac{u_n + u_m}{2}\right\|_2 \geq \alpha_o.$$

The parallelogram law now gives

$$\begin{aligned} \|u_n - u_m\|_2^2 &= 2\|u_n\|_2^2 + 2\|u_m\|_2^2 - \|u_n + u_m\|_2^2 \\ &\leq 2\|u_n\|_2^2 + 2\|u_m\|_2^2 - 4\alpha_o^2. \end{aligned}$$

The right-hand side tends to 0 as $n, m \to \infty$, proving the claim.

By Corollary 10.9, there exists $u_o \in \mathcal{L}^2(\mu)$ such that

$$\lim_{n\to\infty} \|u_n - u_o\|_2 = 0.$$

Boundedness of $\varphi$ gives

$$|\varphi(u_o) - 1| = |\varphi(u_o - u_n)| \le c \cdot \|u_o - u_n\|_2 \longrightarrow 0,$$

hence $\varphi(u_o) = 1$ and $u_o \in \mathcal{U}$. Finally, the reverse triangle inequality gives

$$|\|u_n\|_2 - \|u_o\|_2| \le \|u_n - u_o\|_2 \longrightarrow 0.$$

Thus $\|u_o\|_2 = \alpha_o$. □

**Lemma 11.6**

Let $\varphi : \mathcal{L}^2(\mu) \to \mathbb{R}$ be a linear functional which is not identically zero. Let $\mathcal{U}$ and $u_o \in \mathcal{U}$ be as in Lemma 11.5, and let

$$\mathcal{V} := \{v \in \mathcal{L}^2(\mu) \mid \varphi(v) = 0\} = \ker \varphi.$$

Then

$$\langle u_o, v\rangle = 0, \quad \forall v \in \mathcal{V}.$$

***Proof:*** Fix $v \in \mathcal{V}$. For every $t \in \mathbb{R}$, we have

$$\varphi(u_o + tv) = \varphi(u_o) + t\varphi(v) = 1,$$

so $u_o + tv \in \mathcal{U}$. By the minimizing property of $u_o$,

$$\|u_o\|_2^2 \le \|u_o + tv\|_2^2.$$

Expanding the right-hand side yields

$$0 \le 2t\langle u_o, v\rangle + t^2\|v\|_2^2, \quad \forall t \in \mathbb{R}.$$

For $t > 0$, applying this inequality with both $t$ and $-t$ gives

$$|\langle u_o, v\rangle| \le \frac{t}{2}\|v\|_2^2.$$

Letting $t \to 0$ gives $\langle u_o, v\rangle = 0$. □

***Proof of Riesz representation theorem for $\mathcal{L}^2(\mu)$ (Theorem 11.3):*** If $\varphi$ is identically zero, take $f_o = \underline{0}$. Suppose that $\varphi$ is not identically zero. Let $\mathcal{U}$, $\alpha_o$, and $u_o$ be as in Lemma 11.5. Since $\alpha_o > 0$, we may define

$$f_o = \frac{1}{\|u_o\|_2^2} u_o \in \mathcal{L}^2(\mu).$$

Fix $g \in \mathcal{L}^2(\mu)$ and put

$$v = g - \varphi(g)u_o.$$

Since $\varphi(u_o) = 1$, we have $\varphi(v) = 0$, so $v \in \mathcal{V}$. By Lemma 11.6,

$$0 = \langle u_o, v \rangle = \langle u_o, g \rangle - \varphi(g) \langle u_o, u_o \rangle.$$

Since $\langle u_o, u_o \rangle = \|u_o\|_2^2$, it follows that

$$\varphi(g) = \frac{\langle u_o, g \rangle}{\|u_o\|_2^2} = \langle f_o, g \rangle.$$

This holds for every $g \in \mathcal{L}^2(\mu)$, as required. □

# 12 Radon-Nikodym theorem in the special case when $\nu \leq \mu$

Fix a measurable space $(X, \mathfrak{M})$.

**Notation 12.1** **Domination of measures**

Let $\mu, \nu : \mathfrak{M} \to [0, \infty]$ be positive measures. We write $\nu \leq \mu$ if

$$\nu(A) \leq \mu(A), \quad \forall A \in \mathfrak{M}.$$

The inequality defining $\nu \leq \mu$ extends to an inequality between integrals.

**Proposition 12.2** **Monotonicity of integrals with respect to measures**

Let $\mu, \nu : \mathfrak{M} \to [0, \infty]$ be positive measures such that $\nu \leq \mu$. Then

$$\int_X f \,\mathrm{d}\nu \leq \int_X f \,\mathrm{d}\mu, \quad \forall f \in \mathrm{Bor}^+(X, \mathbb{R}).$$

***Proof:*** We use the method of the friendly functions from Method of the friendly functions (Lemma 6.9). Define

$$\mathcal{G} := \left\{ g \in \mathrm{Bor}^+(X, \mathbb{R}) \;\middle|\; \int_X g \,\mathrm{d}\nu \leq \int_X g \,\mathrm{d}\mu \right\}.$$

We verify its four hypotheses.

- If $A \in \mathfrak{M}$, then

  $$\int_X \chi_A \,\mathrm{d}\nu = \nu(A) \leq \mu(A) = \int_X \chi_A \,\mathrm{d}\mu,$$

  so $\chi_A \in \mathcal{G}$.
- If $g_1, g_2 \in \mathcal{G}$, then

$$\begin{aligned}\int_X (g_1 + g_2)\,\mathrm{d}\nu &= \int_X g_1\,\mathrm{d}\nu + \int_X g_2\,\mathrm{d}\nu \\ &\leq \int_X g_1\,\mathrm{d}\mu + \int_X g_2\,\mathrm{d}\mu \\ &= \int_X (g_1 + g_2)\,\mathrm{d}\mu,\end{aligned}$$

so $g_1 + g_2 \in \mathcal{G}$.

- If $g \in \mathcal{G}$ and $\alpha \in [0, \infty)$, then

$$\int_X \alpha g\,\mathrm{d}\nu = \alpha \int_X g\,\mathrm{d}\nu \leq \alpha \int_X g\,\mathrm{d}\mu = \int_X \alpha g\,\mathrm{d}\mu,$$

so $\alpha g \in \mathcal{G}$.

- Suppose that $g \in \mathrm{Bor}^+(X, \mathbb{R})$, $g_1 \leq g_2 \leq \cdots \leq g_n \leq \cdots$ in $\mathcal{G}$, and $g_n \nearrow g$. By Monotone Convergence Theorem (Theorem 6.2), applied with respect to $\nu$ and $\mu$,

$$\int_X g\,\mathrm{d}\nu = \lim_{n\to\infty} \int_X g_n\,\mathrm{d}\nu \leq \lim_{n\to\infty} \int_X g_n\,\mathrm{d}\mu = \int_X g\,\mathrm{d}\mu.$$

Hence $g \in \mathcal{G}$.

Therefore Method of the friendly functions (Lemma 6.9) gives $\mathcal{G} = \mathrm{Bor}^+(X, \mathbb{R})$, proving the result. □

**Remark 12.3** **Domination obtained by integrating a density**

Let $\mu : \mathfrak{M} \to [0, \infty]$ be a positive measure and let $h \in \mathrm{Bor}^+(X, \mathbb{R})$. Recall from Proposition 7.9 that

$$\nu(A) := \int_A h\,\mathrm{d}\mu = \int_X h \cdot \chi_A\,\mathrm{d}\mu, \quad \forall A \in \mathfrak{M},$$

defines a positive measure $\nu$; we write $\mathrm{d}\nu = h\,\mathrm{d}\mu$.

Suppose, moreover, that $h(x) \leq 1$ for every $x \in X$. For every $A \in \mathfrak{M}$, we then have $h \cdot \chi_A \leq \chi_A$, and hence

$$\nu(A) = \int_X h \cdot \chi_A\,\mathrm{d}\mu \leq \int_X \chi_A\,\mathrm{d}\mu = \mu(A).$$

Thus $\nu \leq \mu$.

The following important special case of the Radon–Nikodym theorem gives a converse to Domination obtained by integrating a density (Remark 12.3).

**Theorem 12.4** **Radon–Nikodym theorem in the case when $\nu \leq \mu$**

Let $\mu, \nu : \mathfrak{M} \to [0, \infty)$ be finite positive measures such that $\nu \leq \mu$. Then there exists $h \in \mathrm{Bor}^+(X, \mathbb{R})$ such that $h(x) \leq 1$ for every $x \in X$ and

$$\mathrm{d}\nu = h \, \mathrm{d}\mu;$$

that is,

$$\nu(A) = \int_A h \, \mathrm{d}\mu, \quad \forall A \in \mathfrak{M}.$$

Moreover, $h$ is unique in the a.e$-\mu$ sense: if $\tilde{h} \in \mathrm{Bor}^+(X, \mathbb{R})$ also satisfies $\mathrm{d}\nu = \tilde{h} \, \mathrm{d}\mu$, then $\tilde{h} = h$ a.e$-\mu$.

**Notation and Remark 12.5** **Radon–Nikodym derivative**

1. The function $h$ appearing in Radon–Nikodym theorem in the case when $\nu \leq \mu$ (Theorem 12.4) is called the *Radon–Nikodym derivative of $\nu$ with respect to $\mu$*. It is commonly denoted by

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

2. The uniqueness assertion in Radon–Nikodym theorem in the case when $\nu \leq \mu$ (Theorem 12.4) is comparatively easy, but equality a.e$-\mu$ is the most one can expect. Indeed, if $\mathrm{d}\nu = h \, \mathrm{d}\mu$ and $\tilde{h} = h$ a.e$-\mu$, then Proposition 8.11 gives $\mathrm{d}\nu = \tilde{h} \, \mathrm{d}\mu$ as well.
3. The difficult part is the existence of $h$, which will be proved using an $\mathcal{L}^2(\mu)$ approach. To motivate it, suppose temporarily that the desired derivative has already been found. Then

$$\int_X g \, \mathrm{d}\nu = \int_X g \cdot h \, \mathrm{d}\mu, \quad \forall g \in \mathrm{Bor}^+(X, \mathbb{R}).$$

Since $0 \leq h \leq \underline{1}$ and $\mu$ is finite, we have $h \in \mathcal{L}^2(\mu)$. Restricting the preceding identity to $g \in \mathcal{L}^2(\mu)$ gives

$$\langle h, g \rangle = \int_X g \, \mathrm{d}\nu.$$

This suggests defining a linear functional $\varphi : \mathcal{L}^2(\mu) \to \mathbb{R}$ by

$$\varphi(g) := \int_X g \, \mathrm{d}\nu, \quad \forall g \in \mathcal{L}^2(\mu),$$

and applying Riesz representation theorem for $\mathcal{L}^2(\mu)$ (Theorem 11.3). The next two lemmas ensure that this definition makes sense and that $\varphi$ is bounded.

**Lemma 12.6**

Suppose $\mu, \nu : \mathfrak{M} \to [0, \infty)$ are finite positive measures such that $\nu \leq \mu$. Then $\mathcal{L}^2(\mu) \subseteq \mathcal{L}^2(\nu)$, and

$$\|f\|_{2,\nu} \leq \|f\|_{2,\mu}, \quad \forall f \in \mathcal{L}^2(\mu).$$

***Proof:*** Let $f \in \mathcal{L}^2(\mu)$. Then

$$\int_X |f|^2 \, \mathrm{d}\mu < \infty.$$

Since $|f|^2 \in \mathrm{Bor}^+(X, \mathbb{R})$, Monotonicity of integrals with respect to measures (Proposition 12.2) gives

$$\int_X |f|^2 \, \mathrm{d}\nu \leq \int_X |f|^2 \, \mathrm{d}\mu < \infty.$$

Hence $f \in \mathcal{L}^2(\nu)$. The same inequality says that

$$\|f\|_{2,\nu}^2 \leq \|f\|_{2,\mu}^2,$$

and taking square roots gives the desired norm inequality. □

**Lemma 12.7**

Suppose $\nu : \mathfrak{M} \to [0, \infty)$ is a finite positive measure. Then $\mathcal{L}^2(\nu) \subseteq \mathcal{L}^1(\nu)$. Moreover, for every $f \in \mathcal{L}^2(\nu)$ we have

$$\left| \int_X f \, \mathrm{d}\nu \right| \leq \sqrt{\nu(X)} \cdot \|f\|_{2,\nu}.$$

***Proof:*** The constant function $\underline{1}$ belongs to $\mathcal{L}^2(\nu)$, with

$$\|\underline{1}\|_{2,\nu} = \sqrt{\int_X \underline{1}^2 \, \mathrm{d}\nu} = \sqrt{\nu(X)}.$$

Thus, for every $f \in \mathcal{L}^2(\nu)$, Cauchy–Schwarz gives $f = f \cdot \underline{1} \in \mathcal{L}^1(\nu)$ and

$$\left| \int_X f \, \mathrm{d}\nu \right| \leq \|f\|_{2,\nu} \cdot \|\underline{1}\|_{2,\nu} = \sqrt{\nu(X)} \cdot \|f\|_{2,\nu}.$$

□

***Proof of existence in Radon–Nikodym theorem in the case when $\nu \leq \mu$ (Theorem 12.4):***

***Claim 1:*** It makes sense to define $\varphi : \mathcal{L}^2(\mu) \to \mathbb{R}$ by

$$\varphi(g) := \int_X g \, \mathrm{d}\nu.$$

Moreover, $\varphi$ is a bounded linear functional on $\mathcal{L}^2(\mu)$.

***Verification of Claim 1:*** By Lemma 12.6 and Lemma 12.7,

$$\mathcal{L}^2(\mu) \subseteq \mathcal{L}^2(\nu) \subseteq \mathcal{L}^1(\nu),$$

so the integral defining $\varphi(g)$ is finite for every $g \in \mathcal{L}^2(\mu)$. Linearity of $\varphi$ follows from linearity of the integral with respect to $\nu$. Moreover, the same two lemmas give

$$|\varphi(g)| \leq \sqrt{\nu(X)} \cdot \|g\|_{2,\nu} \leq \sqrt{\nu(X)} \cdot \|g\|_{2,\mu}, \quad \forall g \in \mathcal{L}^2(\mu).$$

Thus $\varphi$ is bounded.

***Claim 2:*** There exists $f_o \in \mathcal{L}^2(\mu)$ such that

$$\int_X g \,\mathrm{d}\nu = \int_X f_o \cdot g \,\mathrm{d}\mu, \quad \forall g \in \mathcal{L}^2(\mu). \tag{$\diamondsuit$}$$

***Verification of Claim 2:*** Apply Riesz representation theorem for $\mathcal{L}^2(\mu)$ (Theorem 11.3) to the bounded linear functional $\varphi$ from Claim 1.

***Claim 3:*** Let $f_o \in \mathcal{L}^2(\mu)$ be as in Claim 2, and define

$$Z_- \coloneqq \{x \in X \mid f_o(x) < 0\} \quad \text{and} \quad Z_+ \coloneqq \{x \in X \mid f_o(x) > 1\}.$$

Then $Z_-, Z_+ \in \mathfrak{M}$ and $\mu(Z_-) = \mu(Z_+) = 0$.

***Verification of Claim 3:*** Since $f_o \in \mathrm{Bor}(X, \mathbb{R})$ and $(-\infty, 0), (1, \infty) \in \mathcal{B}_{\mathbb{R}}$, both $Z_-$ and $Z_+$ belong to $\mathfrak{M}$.

The arguments for $Z_-$ and $Z_+$ are similar, so we verify only that $\mu(Z_+) = 0$. For each $n \in \mathbb{N}$, define

$$Z_n^+ \coloneqq \left\{x \in X \;\middle|\; f_o(x) \geq 1 + \frac{1}{n}\right\}.$$

Then

$$Z_1^+ \subseteq Z_2^+ \subseteq \cdots \subseteq Z_n^+ \subseteq \cdots \quad \text{and} \quad \bigcup_{n=1}^{\infty} Z_n^+ = Z_+.$$

By continuity along increasing chains, it is enough to prove that $\mu(Z_n^+) = 0$ for every $n \in \mathbb{N}$.

Fix $n \in \mathbb{N}$ and let $\chi \coloneqq \chi_{Z_n^+}$. Since $\mu$ is finite, $\chi \in \mathcal{L}^2(\mu)$, so we may substitute it for $g$ in ($\diamondsuit$). From

$$f_o \cdot \chi \geq \left(1 + \frac{1}{n}\right)\chi,$$

we obtain

$$\nu(Z_n^+) = \int_X f_o \cdot \chi \,\mathrm{d}\mu \geq \left(1 + \frac{1}{n}\right)\mu(Z_n^+).$$

On the other hand, $\nu \leq \mu$ gives $\nu(Z_n^+) \leq \mu(Z_n^+)$. Hence

$$\left(1+\frac{1}{n}\right)\mu(Z_n^+) \leq \mu(Z_n^+),$$

which forces $\mu(Z_n^+) = 0$. Therefore $\mu(Z_+) = 0$. The analogous argument gives $\mu(Z_-) = 0$.

Let $f_o$, $Z_-$, and $Z_+$ be as in Claims 2 and 3, and put

$$h := f_o \cdot \left(1 - \chi_{Z_- \cup Z_+}\right).$$

Equivalently,

$$h(x) := \begin{cases} f_o(x) & \text{if } 0 \leq f_o(x) \leq 1 \\ 0 & \text{otherwise} \end{cases}.$$

Then $h \in \mathrm{Bor}^+(X, \mathbb{R})$ and $h(x) \leq 1$ for every $x \in X$. Claim 3 also gives $h = f_o$ a.e$-\mu$. For every $A \in \mathfrak{M}$, substituting $g = \chi_A$ into ($\diamondsuit$) and using Proposition 8.11 yields

$$\nu(A) = \int_X \chi_A \,\mathrm{d}\nu = \int_X f_o \cdot \chi_A \,\mathrm{d}\mu = \int_X h \cdot \chi_A \,\mathrm{d}\mu = \int_A h \,\mathrm{d}\mu.$$

Hence $\mathrm{d}\nu = h\,\mathrm{d}\mu$. □

# 13 Connecting equation, and Radon-Nikodym for $\nu \ll \mu$

Continue to work with the measurable space $(X, \mathfrak{M})$ fixed previously. We will use Radon–Nikodym theorem in the case when $\nu \leq \mu$ (Theorem 12.4) to obtain a version of the Radon–Nikodym theorem under the weaker hypothesis of absolute continuity $\nu \ll \mu$, rather than the stronger hypothesis $\nu \leq \mu$.

## 13.1 The trick of the connecting equation

In this section, let $\mu, \nu : \mathfrak{M} \to [0, \infty)$ be finite positive measures, with no relation assumed between them. The idea is to apply Radon–Nikodym theorem in the case when $\nu \leq \mu$ (Theorem 12.4) using the inequalities

$$\mu \leq \mu + \nu \quad \text{and} \quad \nu \leq \mu + \nu.$$

**Definition and Remark 13.1** **Sum of positive measures**

Let $\mu, \nu : \mathfrak{M} \to [0, \infty]$ be positive measures. Define $\mu + \nu : \mathfrak{M} \to [0, \infty]$ by

$$(\mu + \nu)(A) := \mu(A) + \nu(A), \quad \forall A \in \mathfrak{M}.$$

Then $\mu + \nu$ is itself a positive measure. Moreover,

$$\int_X f \,\mathrm{d}(\mu + \nu) = \int_X f \,\mathrm{d}\mu + \int_X f \,\mathrm{d}\nu, \quad \forall f \in \mathrm{Bor}^+(X, \mathbb{R}).$$

The integral identity follows from the method of the friendly functions.

**Proposition 13.2** **Trick of the connecting equation**

Let $\mu, \nu : \mathfrak{M} \to [0, \infty)$ be finite positive measures. There exist $h_1, h_2 \in \mathrm{Bor}^+(X, \mathbb{R})$ such that

$$h_1(x) + h_2(x) = 1, \quad \forall x \in X,$$

and

$$\int_X g \cdot h_1 \, \mathrm{d}\mu = \int_X g \cdot h_2 \, \mathrm{d}\nu, \quad \forall g \in \mathrm{Bor}^+(X, \mathbb{R}). \tag{$\heartsuit$}$$

***Proof:*** Put $\rho := \mu + \nu$. Then $\rho$ is a finite positive measure and $\nu \leq \rho$. By Radon–Nikodym theorem in the case when $\nu \leq \mu$ (Theorem 12.4), there exists $h_1 \in \mathrm{Bor}^+(X, \mathbb{R})$ such that $h_1 \leq \underline{1}$ and

$$\mathrm{d}\nu = h_1 \, \mathrm{d}\rho.$$

Define $h_2 := \underline{1} - h_1 \in \mathrm{Bor}^+(X, \mathbb{R})$. Then $h_1 + h_2 = \underline{1}$.

Fix $E \in \mathfrak{M}$. Since $\mathrm{d}\nu = h_1 \, \mathrm{d}\rho$ and $\rho = \mu + \nu$, Sum of positive measures (Definition and Remark 13.1) gives

$$\nu(E) = \int_E h_1 \, \mathrm{d}\rho = \int_E h_1 \, \mathrm{d}\mu + \int_E h_1 \, \mathrm{d}\nu.$$

On the other hand, $h_1 + h_2 = \underline{1}$ gives

$$\nu(E) = \int_E (h_1 + h_2) \, \mathrm{d}\nu = \int_E h_1 \, \mathrm{d}\nu + \int_E h_2 \, \mathrm{d}\nu.$$

All the terms are finite, since $0 \leq h_1, h_2 \leq \underline{1}$ and $\mu, \nu$ are finite. Cancelling the common term therefore yields

$$\int_E h_1 \, \mathrm{d}\mu = \int_E h_2 \, \mathrm{d}\nu, \quad \forall E \in \mathfrak{M}.$$

The two sides define positive measures on $\mathfrak{M}$ by Proposition 7.9, and the preceding equality shows that these measures are equal. Applying Proposition 7.10 to each of them gives

$$\int_X g \cdot h_1 \, \mathrm{d}\mu = \int_X g \cdot h_2 \, \mathrm{d}\nu, \quad \forall g \in \mathrm{Bor}^+(X, \mathbb{R}),$$

which is ($\heartsuit$). $\square$

For the connecting equation, it will be useful to understand the points at which one of $h_1(x)$ and $h_2(x)$ is zero and the other is one.

**Lemma 13.3**

In the framework of Trick of the connecting equation (Proposition 13.2), let

$$S := \{x \in X \mid h_2(x) = 0\} = \{x \in X \mid h_1(x) = 1\}.$$

Then $\mu(S) = 0$.

***Proof:*** Substitute $g = \chi_S$ into ($\heartsuit$). Since $h_1 = \underline{1}$ and $h_2 = \underline{0}$ on $S$, we obtain

$$\mu(S) = \int_S h_1 \,\mathrm{d}\mu = \int_S h_2 \,\mathrm{d}\nu = 0.$$

$\square$

## 13.2 Radon-Nikodym, version with absolute continuity $\nu \ll \mu$

**Definition 13.4** **Absolute continuity of measures**

Let $\mu, \nu : \mathfrak{M} \to [0, \infty]$ be positive measures. We say that $\nu$ is *absolutely continuous with respect to* $\mu$, denoted $\nu \ll \mu$, if

$$A \in \mathfrak{M} \quad \text{and} \quad \mu(A) = 0 \Longrightarrow \nu(A) = 0.$$

**Remark 13.5**

Comparing the definitions of $\nu \leq \mu$ and $\nu \ll \mu$, we immediately obtain

$$\nu \leq \mu \Longrightarrow \nu \ll \mu.$$

The converse does not hold in general.

**Remark 13.6** **Absolute continuity obtained from a density**

Let $\mu : \mathfrak{M} \to [0, \infty]$ be a positive measure and let $h \in \mathrm{Bor}^+(X, \mathbb{R})$. Recall that

$$\nu(A) := \int_A h \,\mathrm{d}\mu = \int_X h \cdot \chi_A \,\mathrm{d}\mu, \quad \forall A \in \mathfrak{M},$$

defines a positive measure $\nu$.

Unlike in Domination obtained by integrating a density (Remark 12.3), we do not assume that $h(x) \leq 1$; the function $h$ may even be unbounded. Consequently, $\nu \leq \mu$ need not hold. For example, if

$$A := \{x \in X \mid h(x) \geq 2\}$$

satisfies $\mu(A) > 0$, then

$$\nu(A) \geq 2\mu(A) > \mu(A).$$

Nevertheless, the measure obtained from any such density always satisfies $\nu \ll \mu$.

**Theorem 13.7** **Radon-Nikodym theorem**

Let $\mu, \nu : \mathfrak{M} \to [0, \infty)$ be finite positive measures such that $\nu \ll \mu$. Then there exists $h \in \mathrm{Bor}^+(X, \mathbb{R})$ such that

$$\mathrm{d}\nu = h \, \mathrm{d}\mu;$$

that is,

$$\nu(A) = \int_A h \, \mathrm{d}\mu, \quad \forall A \in \mathfrak{M}.$$

Moreover, $h$ is unique in the a.e$-\mu$ sense: if $\tilde{h} \in \mathrm{Bor}^+(X, \mathbb{R})$ also satisfies $\mathrm{d}\nu = \tilde{h} \, \mathrm{d}\mu$, then $\tilde{h} = h$ a.e$-\mu$.

***Proof:*** Let $h_1, h_2$ be as in Trick of the connecting equation (Proposition 13.2), and put

$$S := \{x \in X \mid h_2(x) = 0\} \quad \text{and} \quad A := X \setminus S.$$

By Lemma 13.3, $\mu(S) = 0$. Since $\nu \ll \mu$, we also have $\nu(S) = 0$. Moreover, $h_2(x) > 0$ for every $x \in A$.

Define $h : X \to \mathbb{R}$ by

$$h(x) = \begin{cases} \frac{h_1(x)}{h_2(x)} & \text{if } x \in A \\ 0 & \text{if } x \in S \end{cases}.$$

The sets $A, S$ are measurable, and the quotient $\frac{h_1}{h_2}$ is non-negative and Borel measurable on $A$, so $h \in \mathrm{Bor}^+(X, \mathbb{R})$.

Fix $f \in \mathrm{Bor}^+(X, \mathbb{R})$ and define $g : X \to \mathbb{R}$ by

$$g(x) = \begin{cases} \frac{f(x)}{h_2(x)} & \text{if } x \in A \\ 0 & \text{if } x \in S \end{cases}.$$

As above, $g \in \mathrm{Bor}^+(X, \mathbb{R})$. Applying ($\heartsuit$) to $g$ gives

$$\int_X f \cdot h \, \mathrm{d}\mu = \int_A g \cdot h_1 \, \mathrm{d}\mu = \int_A g \cdot h_2 \, \mathrm{d}\nu = \int_A f \, \mathrm{d}\nu = \int_X f \, \mathrm{d}\nu.$$

In the first and last equalities, we used $\mu(S) = \nu(S) = 0$. Taking $f = \chi_E$ for each $E \in \mathfrak{M}$ shows that

$$\nu(E) = \int_E h \, \mathrm{d}\mu, \quad \forall E \in \mathfrak{M},$$

and hence $\mathrm{d}\nu = h \, \mathrm{d}\mu$.

For uniqueness, suppose $\tilde{h} \in \mathrm{Bor}^+(X, \mathbb{R})$ also satisfies $\mathrm{d}\nu = \tilde{h} \, \mathrm{d}\mu$. Define

$$P := \left\{x \in X \;\middle|\; h(x) > \tilde{h}(x)\right\} \quad \text{and} \quad N := \left\{x \in X \;\middle|\; h(x) < \tilde{h}(x)\right\}.$$

Both sets are measurable. The functions $h$ and $\tilde{h}$ are integrable because their integrals over $X$ equal the finite number $\nu(X)$. Since their integrals over every measurable set are equal, we have

$$\int_P (h - \tilde{h})\, \mathrm{d}\mu = \nu(P) - \nu(P) = 0.$$

The integrand is non-negative and is strictly positive precisely on $P$. Thus Remark 8.12 and Proposition 8.9 give $\mu(P) = 0$. Applying the same argument to $\tilde{h} - h$ on $N$ gives $\mu(N) = 0$. Since

$$\{x \in X \mid h(x) \neq \tilde{h}(x)\} = P \cup N,$$

we conclude that $h = \tilde{h}$ a.e$-\mu$. □

# 14 Some R-N related stuff

## 14.1 The Lebesgue decomposition theorem

The Lebesgue decomposition theorem can be derived using the same connecting-equation trick as Trick of the connecting equation (Proposition 13.2).

**Definition 14.1** **Concentrated and mutually singular measures**

Let $(X, \mathfrak{M})$ be a measurable space.

1. Let $\mu : \mathfrak{M} \to [0, \infty]$ be a positive measure and let $P \in \mathfrak{M}$. We say that $\mu$ is *concentrated on* $P$ if
$$\mu(X \setminus P) = 0.$$
2. Let $\mu, \nu : \mathfrak{M} \to [0, \infty]$ be positive measures. We say that $\mu$ and $\nu$ are *mutually singular*, denoted $\mu \perp \nu$, if there exist disjoint sets $P, Q \in \mathfrak{M}$ such that $\mu$ is concentrated on $P$ and $\nu$ is concentrated on $Q$.

**Theorem and Definition 14.2** **Lebesgue decomposition**

Let $(X, \mathfrak{M})$ be a measurable space, and let $\mu, \nu : \mathfrak{M} \to [0, \infty)$ be finite positive measures. There exist unique finite positive measures $\nu_1, \nu_2 : \mathfrak{M} \to [0, \infty)$ such that

(i) $\nu_1 + \nu_2 = \nu$;
(ii) $\nu_1 \ll \mu$;
(iii) $\nu_2 \perp \mu$.

The measure $\nu_1$ is called the *absolutely continuous part of* $\nu$ *with respect to* $\mu$, and $\nu_2$ is called the *singular part of* $\nu$ *with respect to* $\mu$. The relation $\nu = \nu_1 + \nu_2$ is called the *Lebesgue decomposition of* $\nu$ *with respect to* $\mu$.

**Remark 14.3**

The uniqueness assertion in Lebesgue decomposition (Theorem and Definition 14.2) is an elementary verification. For existence, we use (♡) for the measures $\mu$ and $\nu$.

***Proof of existence in Lebesgue decomposition (Theorem and Definition 14.2):*** Let $h_1, h_2 \in \mathrm{Bor}^+(X, \mathbb{R})$ be the functions from Trick of the connecting equation (Proposition 13.2), so that $h_1 + h_2 = \underline{1}$ and

$$\int_X g \cdot h_1 \, \mathrm{d}\mu = \int_X g \cdot h_2 \, \mathrm{d}\nu, \quad \forall g \in \mathrm{Bor}^+(X, \mathbb{R}).$$

Recall from Lemma 13.3 that, upon defining

$$S := \{x \in X \mid h_1(x) = 1\} = \{x \in X \mid h_2(x) = 0\} \in \mathfrak{M}, \tag{14.1}$$

we have $\mu(S) = 0$.

Use $S$ to decompose $\nu$: define $\nu_1, \nu_2 : \mathfrak{M} \to [0, \infty)$ by

$$\nu_1(A) := \nu(A \cap (X \setminus S)) \quad \text{and} \quad \nu_2(A) := \nu(A \cap S), \quad \forall A \in \mathfrak{M}. \tag{14.2}$$

These formulas define finite positive measures. Moreover, for every $A \in \mathfrak{M}$,

$$\begin{aligned} \nu_1(A) + \nu_2(A) &= \nu(A \cap (X \setminus S)) + \nu(A \cap S) \\ &= \nu((A \cap (X \setminus S)) \cup (A \cap S)) \\ &= \nu(A), \end{aligned}$$

where additivity applies because the union is disjoint. Hence $\nu_1 + \nu_2 = \nu$.

It remains to verify the other two properties.

***<u>Claim 1:</u>*** $\nu_2 \perp \mu$.

***<u>Verification of Claim 1:</u>*** From its definition, $\nu_2$ is concentrated on $P := S$. Indeed

$$\nu_2(X \setminus S) = \nu((X \setminus S) \cap S) = \nu(\varnothing) = 0.$$

But $\mu$ is concentrated on $Q := X \setminus S$. Indeed

$$\mu(X \setminus Q) = \mu(S) = 0.$$

***<u>Claim 2:</u>*** $\nu_1 \ll \mu$.

***<u>Verification of Claim 2:</u>*** Have $A \in \mathfrak{M}$ with $\mu(A) = 0$, want $\nu_1(A) = 0$, that is $\nu(A \cap (X \setminus S)) = 0$. How? Put $g = \chi_A$ in $\heartsuit$.

$$\int \chi_A h_1 \, \mathrm{d}\mu = \int \chi_A h_2 \, \mathrm{d}\nu = 0$$

since $\chi_A h_1 = \underline{0}$ a.e$-\mu$ (due to $\mu(A) = 0$). Hence have $\chi_A h_2 \in \mathrm{Bor}^+(X, \mathbb{R})$ which has $\int \chi_A h_2 \, \mathrm{d}\nu = 0$. It follows that $\chi_A h_2 = \underline{0}$ a.e$-\nu$ (remark 8.12). Hence

$$Z = \{x \in X \mid \chi_A(x) h_2(x) \neq 0\}$$

has $\nu(Z) = 0$. But

$$\begin{aligned} Z &= \{x \in X \mid \chi_A(x) \neq 0\} \cap \{x \in X \mid h_2(x) \neq 0\} \\ &= A \cap (X \setminus S) \end{aligned}$$

Hence got $\nu(A \cap (X \setminus S)) = 0$ as required.

$\square$

## 14.2 Conditional expectations

Let $(X, \mathfrak{M}, \mu)$ be a probability space. An important application of the Radon–Nikodym theorem is that it allows us to define the conditional expectation of an integrable function onto a sub-sigma-algebra $\mathfrak{N} \subseteq \mathfrak{M}$.

**Notation and Remark 14.4** **Measurability with respect to different sigma-algebras**

Suppose $(X, \mathfrak{M})$ is a measurable space and $\mathfrak{N} \subseteq \mathfrak{M}$ is a sigma-algebra of subsets of $X$. We use the notation

$$\operatorname{Bor}((X, \mathfrak{M}), \mathbb{R}) := \{f : X \to \mathbb{R} \mid f^{-1}(B) \in \mathfrak{M} \text{ for every } B \in \mathcal{B}_{\mathbb{R}}\} \tag{14.3}$$

and, in parallel,

$$\operatorname{Bor}((X, \mathfrak{N}), \mathbb{R}) := \{f : X \to \mathbb{R} \mid f^{-1}(B) \in \mathfrak{N} \text{ for every } B \in \mathcal{B}_{\mathbb{R}}\}. \tag{14.4}$$

This makes explicit which sigma-algebra is used for measurability. Since $\mathfrak{N} \subseteq \mathfrak{M}$, we have

$$\operatorname{Bor}((X, \mathfrak{N}), \mathbb{R}) \subseteq \operatorname{Bor}((X, \mathfrak{M}), \mathbb{R}).$$

The inclusion can be strict, and $\operatorname{Bor}((X, \mathfrak{N}), \mathbb{R})$ can be much smaller than $\operatorname{Bor}((X, \mathfrak{M}), \mathbb{R})$.

**Example 14.5** **Measurability with respect to a finite sigma-algebra**

Let $X = [0, 1]$ and $\mathfrak{M} = \mathcal{B}_{[0,1]}$. Partition

$$[0, 1] = J_1 \cup J_2 \cup J_3,$$

where

$$J_1 = \left[0, \frac{1}{3}\right], \quad J_2 = \left(\frac{1}{3}, \frac{1}{2}\right], \quad J_3 = \left(\frac{1}{2}, 1\right].$$

Define

$$\mathfrak{N} := \{\varnothing, J_1, J_2, J_3, J_1 \cup J_2, J_1 \cup J_3, J_2 \cup J_3, X\}. \tag{14.5}$$

Then $\mathfrak{N}$ is a sigma-algebra and $\mathfrak{N} \subseteq \mathfrak{M}$. The space $\operatorname{Bor}((X, \mathfrak{M}), \mathbb{R})$ is infinite-dimensional; for example, it contains every continuous function $f : [0, 1] \to \mathbb{R}$. On the other hand,

$$g \in \operatorname{Bor}((X, \mathfrak{N}), \mathbb{R}) \iff g \text{ is constant on each of } J_1, J_2, J_3. \tag{14.6}$$

Consequently,

$$\begin{aligned} \operatorname{Bor}((X, \mathfrak{N}), \mathbb{R}) &= \operatorname{span}\left\{\chi_{J_1}, \chi_{J_2}, \chi_{J_3}\right\} \\ &= \left\{\alpha_1 \chi_{J_1} + \alpha_2 \chi_{J_2} + \alpha_3 \chi_{J_3} \;\middle|\; \alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}\right\}. \end{aligned} \tag{14.7}$$

The verification of (14.6) and (14.7) is left as an exercise.

**Lemma 14.6** **Restriction of a measure to a sub-sigma-algebra**

Let $(X, \mathfrak{M}, \mu)$ be a measure space with $\mu(X) < \infty$, and let $\mathfrak{N} \subseteq \mathfrak{M}$ be a sigma-algebra. Define $\mu_o : \mathfrak{N} \to [0, \infty)$ by

$$\mu_o(N) := \mu(N), \quad \forall N \in \mathfrak{N}.$$

Then $\mu_o$ is a finite positive measure on $\mathfrak{N}$, and

$$\int_X g \, \mathrm{d}\mu_o = \int_X g \, \mathrm{d}\mu, \quad \forall g \in \mathrm{Bor}^+((X, \mathfrak{N}), \mathbb{R}). \tag{14.8}$$

The proof of Restriction of a measure to a sub-sigma-algebra (Lemma 14.6) is left as an exercise. One approach is to apply Method of the friendly functions (Lemma 6.9) to

$$\mathcal{G} := \left\{ g \in \mathrm{Bor}^+((X, \mathfrak{N}), \mathbb{R}) \;\middle|\; \int_X g \, \mathrm{d}\mu_o = \int_X g \, \mathrm{d}\mu \right\}.$$

**Definition 14.7** **Conditional expectation**

Let $(X, \mathfrak{M}, \mu)$ be a measure space with $\mu(X) < \infty$, let $\mathfrak{N} \subseteq \mathfrak{M}$ be a sigma-algebra, and let $\mu_o$ be the restriction of $\mu$ to $\mathfrak{N}$. For $f \in \mathcal{L}^1(X, \mathfrak{M}, \mu)$, a function $g : X \to \mathbb{R}$ is called a *(version of the) conditional expectation of $f$ onto $\mathfrak{N}$* if $g \in \mathcal{L}^1(X, \mathfrak{N}, \mu_o)$ and

$$\int_X \chi_N \cdot g \, \mathrm{d}\mu_o = \int_X \chi_N \cdot f \, \mathrm{d}\mu, \quad \forall N \in \mathfrak{N}. \tag{14.9}$$

**Comment**

The function $f$ in (14.9) need not be $\mathfrak{N}/\mathcal{B}_{\mathbb{R}}$-measurable, so the integral of $\chi_N \cdot f$ may not make sense on the measure space $(X, \mathfrak{N}, \mu_o)$. The purpose of $g$ is to replace $f$ by an $\mathfrak{N}/\mathcal{B}_{\mathbb{R}}$-measurable function while preserving its integrals over sets in $\mathfrak{N}$.

**Proposition 14.8** **Existence for non-negative functions**

Let $(X, \mathfrak{M}, \mu)$ be a measure space with $\mu(X) < \infty$, let $\mathfrak{N} \subseteq \mathfrak{M}$ be a sigma-algebra, and let $f \in \mathcal{L}^1(X, \mathfrak{M}, \mu)$ satisfy

$$f(x) \geq 0, \quad \forall x \in X.$$

Then $f$ has a conditional expectation $g$ onto $\mathfrak{N}$, where $g$ may be chosen so that $g(x) \geq 0$ for every $x \in X$.

***Proof:*** [To be shown in class.] □

**Corollary 14.9** **Existence and uniqueness of conditional expectation**

Let $(X, \mathfrak{M}, \mu)$ be a measure space with $\mu(X) < \infty$, let $\mathfrak{N} \subseteq \mathfrak{M}$ be a sigma-algebra, and let $f \in \mathcal{L}^1(X, \mathfrak{M}, \mu)$.

1. The function $f$ has a conditional expectation onto $\mathfrak{N}$.
2. If $g, g'$ are two versions of the conditional expectation of $f$ onto $\mathfrak{N}$, then there exists $Z \in \mathfrak{N}$ with $\mu(Z) = 0$ such that

$$g(x) = g'(x), \quad \forall x \in X \setminus Z.$$

The proof of Existence and uniqueness of conditional expectation (Corollary 14.9) is left as an exercise. For existence, apply Existence for non-negative functions (Proposition 14.8) separately to $f_+$ and $f_-$. For uniqueness, apply the almost-everywhere criterion to $|g - g'| \in \mathrm{Bor}^+((X, \mathfrak{N}), \mathbb{R})$.

**Notation 14.10** **Conditional expectation**

In the framework of Existence and uniqueness of conditional expectation (Corollary 14.9), a conditional expectation of $f$ onto $\mathfrak{N}$ is commonly denoted by

$$g = \mathrm{E}[f \mid \mathfrak{N}].$$

Thus $\mathrm{E}[f \mid \mathfrak{N}] \in \mathcal{L}^1(X, \mathfrak{N}, \mu_o)$. It is not uniquely determined as a function; it is determined only up to equality almost everywhere with respect to $\mu_o$.